Modeling, design and fabrication of diffractive optical elements based on nanostructures operating beyond the scalar paraxial domain

Giang Nam Nguyen

THÈSE / Télécom Bretagne
sous le sceau de l'Université européenne de Bretagne

pour obtenir le grade de Docteur de Télécom Bretagne
En accréditation conjointe avec l'Ecole Doctorale Sicma
mention : Sciences et Technologies de l'Information et de la Communication

présentée par
Giang Nam NGUYEN
préparée dans le département Optique

Modeling, design and fabrication of diffractive optical elements based on nanostructures operating beyond the scalar paraxial domain

Thèse soutenue le 09 décembre 2014
Devant le jury composé de :

Pierre Ambs
Professeur, Université de Haute-Alsace / président & rapporteur

Nicolas Guérineau
Directeur d'études (HDR), ONERA - Palaiseau / rapporteur

Philippe Gérard
Maître de conférences (HDR), Insa - Strasbourg / exeminateur

Pierre-Emmanuel Durand
Maître de conférences, Université de Bretagne-Sud / examinateur

Patrick Meyrueis
Professeur, Université de Strasbourg / Co-directeur de thèse

Kevin Heggarty
Professeur, Télécom Bretagne / directeur de thèse

N° d'ordre : 2014**telb0336**

Sous le sceau de l'Université européenne de Bretagne

Télécom Bretagne

En accréditation conjointe avec l'Ecole Doctorale Sicma

MODELING, DESIGN AND FABRICATION OF DIFFRACTIVE OPTICAL ELEMENTS BASED ON NANOSTRUCTURES OPERATING BEYOND THE SCALAR PARAXIAL DOMAIN

Thèse de Doctorat

Mention : Sciences et Technologies de l'Information et de la Communication

Présentée par **Giang-Nam Nguyen**

Département : Optique

Directeur de thèse : Kevin Heggarty

Soutenue le 9 décembre 2014

Jury :

M. Pierre Ambs – Prof., Université de Haute-Alsace (Rapporteur)
M. Nicolas Guérineau – HDR, ONERA Palaiseau (Rapporteur)
M. Kevin Heggarty – Prof., Télécom Bretagne (Directeur de thèse)
M. Patrick Meyrueis – Prof., Université de Strasbourg (Co-directeur de thèse)
M. Philippe Gérard – HDR, INSA Strasbourg (Encadrant de thèse)
M. Pierre-Emmanuel Durand – MdC., Université de Bretagne-Sud (Examinateur)

Modeling, design and fabrication of diffractive optical elements based on nanostructures operating beyond the scalar paraxial domain

Giang-Nam NGUYEN

February 5, 2015

Acknowledgements

I would like to express my deepest gratitude to my principle advisor, Mr. Kevin Heggarty - Professor at Télécom Bretagne for his excelent guidance, supervision and patience during my doctoral thesis. I would also like to acknowledge Mr. Philippe Gérard - Maître de conférences (HDR) at INSA Strasbourg and Mr. Patrick Meyrueis - Professor at Télécom Physique Strasbourg, for helping me develop my background in diffractive optics and providing me with research directions.

I am grateful to Mr. Andreas Bacher and Mr. Peter-Jürgen Jakobs for their instruction and support during my stay at Karlsruhe Institute of Technology in Germany for the fabrication of submicron Diffractive Optical Elements using Electron Beam Lithography. I would also like to thank Mr. Patrice Baldeck - Professor at Joseph Fourier University in Grenoble for his advice during our collaboration in parallel Two-Photon Polymerization lithography.

I would not be able to finish my dissertation without the kind reviewers: Mr. Pierre Ambs - Professor at University of Upper Alsace in Mulhouse, Mr. Nicolas Guérineau - Maître de conférences (HDR) at ONERA in Palaiseau and Mr. Pierre-Emmanuel Durand - Maître de conférences at University of South Brittany in Lorient. I am very thankful for their patience in correcting my thesis.

Last but not least, I appreciate the help and fruitful discussions from scientists, staffs and fellow students at Télécom Bretagne, Télécom Physique Strasbourg, Joseph Fourier University and Karlsruhe Institute of Technology. I also thank my family and my friends for their support and encouragement during my doctoral thesis.

My sincere thanks to all of you!

Contents

Abbreviations

AFM	Atomic Force Microscope
ASM	Angular Spectrum Method
BS	Beam Splitter
CCD	Charge-Coupled Device
CD	Compact Disc
CGH	Computer Generated Hologram
DBS	Direct Binary Search
DOE	Diffractive Optical Elements
DVD	Digital Video Disc
EBL	Electron Beam Lithography
FDTD	Finite-Difference Time-Domain
FMM	Fourier Modal Method
FT	Fourier Transform
GA	Genetic Algorithm
IASA	Iterative Angular Spectrum Algorithm
Ifremer	Institut Français de Recherche pour l'Exploitation de la Mer
IFT	Inverse Fourier Transform
IFTA	Iterative Fourier Transform Algorithm
JFU	Joseph Fourier University
KIT	Karlsruhe Institute of Technology
LCD	Liquid Crystal Display
MEEP	MIT Electromagnetic Equation Propagation
MPI	Message Passing Interface
MSE	Mean-Square Error
NA	Numerical Aperture
OPD	Optical Path Difference
ORMOCER	ORganically MOdified CERamic
PC	Personal Computer
PI	Physik Instrumente
PMMA	Poly-methyl methacrylate

RAM	Random-Access Memory
RCWA	Rigorous Coupled-Wave Analysis
RIE	Reactive-Ion Etching
RoD	Ratio of Difference
RMSE	Root-Mean-Square Error
RS	Rayleigh-Sommerfeld
RSM	Radiation Spectrum Method
SLM	Spatial Light Modulator
TB	Télécom Bretagne
TI	Texas Instruments
TE	Tranverse Electric
TEA	Thin Element Approximation
TM	Tranverse Magnetic
TPS	Télécom Physique Strasbourg
1D	1-Dimensional
2D	2-Dimensional
2PP	2-Photon Polymerization
3D	3-Dimensional

List of Figures

List of Tables

Abstract

This thesis aims to extend the range of Diffractive Optical Element (DOE) applications by developing models, algorithms and rapid prototyping techniques for DOEs with diffraction angles $> 10°$, which is beyond the limits of scalar paraxial diffraction model. We develop an accurate and efficient scalar non-paraxial far-field propagator to overcome the limits of the conventional scalar diffraction models. An iterative algorithm based on this propagator is then developed for the design of wide-angle Fourier elements. Experimental results confirm that our scalar non-paraxial propagator and design algorithm can be used for the modeling and design of thin Fourier DOEs with diffraction angles up to about 37° and perhaps even higher.

The remaining discrepancies in diffracted power between modeling, design and experiment are then shown to result from both fabrication errors and by the fact that we are approaching the limit of the Thin Element Approximation (TEA). The practical limits of the TEA are investigated by comparison with the rigorous vectorial simulations. We then develop, optimize and parallelize a rigorous diffraction model based on the Finite-Difference Time-Domain method coupled with our scalar non-paraxial propagator to overcome the limits of the TEA and the computational limitations of current vectorial models. A genetic design algorithm based on this model is proposed for the design of thick DOEs and this algorithm is currently being calibrated.

These models and algorithms have now brought us to the resolution limit of our existing photoplotter used for DOE fabrication. Therefore, we investigate the possibility of building a new parallel photoplotter based on Two-Photon Polymerization (2PP) as a way to rapid, cost-effective prototyping of high resolution (submicron) structures. We design and fabricate spot array DOEs at Télécom Bretagne to parallelize the 2PP fabrication process used at Joseph Fourier University in Grenoble by a factor of 625. To further speed up the 2PP fabrication process, another prototype parallel 2PP photoplotter using a Spatial Light Modulator, which can generate up to about 0.5 million parallel beams, is designed and is currently being developed at Télécom Bretagne.

Keywords : diffraction, scalar non-paraxial, vectorial modeling, design and fabrication, parallel 2PP.

Résumé

Cette thèse vise à élargir l'éventail des applications des Elements Optiques Diffractifs (EODs) en développant des models, des algorithmes et des techniques de prototypage rapide pour des EODs avec des angles de diffraction $> 10°$, au delà des limites du model scalaire, paraxiale de diffraction. Nous développons un propagateur non-paraxiale scalaire en champ lointain précis et efficace pour surmonter les limites des modèles classiques de la diffraction scalaire. Un algorithme itératif basé sur ce propagateur est développé pour la conception d'éléments de Fourier à grand angle de diffraction. Les résultats expérimentaux confirment que notre propagateur et notre algorithme scalaire, non-paraxiale peuvent être utilisés pour la modélisation et la conception des EODs minces avec angles de diffraction jusqu'à environ 37° et peut-être encore plus élevé.

Nous montrons que les divergences qui subsistent entre la modélisation et l'expérimentation de la puissance diffractée résultent surtout des erreurs de fabrication et par le fait que nous nous approchons de la limite de l'approximation d'un élément mince (TEA - "Thin Element Approximation"). Les limites pratiques de la TEA sont étudiées en comparaison avec le simulations rigoureuses vectorielles. Nous développons, optimisons et parallélisons un modèle de diffraction rigoureuse basée sur la FDTD ("Finite-Difference Time-Domain") couplée avec notre propagateur non-paraxiale scalaire pour surmonter les limites de la TEA et les limites de calcul des modèles vectoriels actuels. Un algorithme de conception génétique sur la base de ce modèle est proposé pour la conception des EODs "épaisses", il est actuellement en cours d'étalonnage.

Ces modèles et algorithmes nous ont maintenant conduits à la limite de résolution de notre photoplotter existant utilisé pour la fabrication des EODs. Nous étudions la possibilité de construire un nouveau photoplotter parallèle basé sur polymérisation à deux photons (2PP) comme un moyen de prototypage rapide et rentable de structures haute résolution (submicroniques). Nous concevons et fabriquons à Télécom Bretagne, des EODs generant une matrice de points lumineux pour paralléliser le processus de fabrication 2PP utilisé à l'Université Joseph Fourier de Grenoble par un facteur de 625. Pour accélérer encore le processus de fabrication 2PP, nous concevons et assemblons

un autre photoplotter 2PP parallèle à base d'un modulateur spatial de lumière, qui permet de générer jusqu'à environ 0,5 million de faisceaux parallèles. Ce phototracer est actuellement en cours de développement à Télécom Bretagne.

Mots-clés: diffraction, scalaire non-paraxiale, modélisation vectorielle, conception et fabrication, 2PP parallèle.

General introduction

Diffractive Optical Elements (DOEs) are increasingly being used for a broad range of applications [1]. Example DOE applications are beam shapers (for laser welding, cutting, machining), beam splitters (for optical telecommunications couplers), optical disc read-heads (in CD, DVD, Blu-ray), pattern generators (for machine vision) and anti-fraud protection (for security documents), etc. However, traditional theory, which is the scalar paraxial diffraction model, is only valid for the modeling and design of small diffraction angle and thin DOEs. Fabrication technology using high performance facilities now enables manufacturing of wide diffraction angle or thick DOEs, leading to the need for new modeling and design algorithms.

Optical diffraction is a physical phenomenon which occurs when a light beam encounters an obstacle and propagates in many different directions. The smaller the obstacle, the larger the diffraction angles and the stronger the diffraction effects become [2]. Unlike reflection and refraction which can be explained by the corpuscular nature of light (i.e. geometrical optics), diffraction can be best described by the wave nature of light (i.e. electromagnetic theory). The propagation of the diffracted wave can be considered as the interference of Huygens-Fresnel secondary wave sources generated by every point in the obstacle [3]. These waves are superposed together, creating a diffraction pattern with a series of maxima and minima on the observing screen.

DOEs are micro or nanostructures which are designed to modify the spatial distribution of a light beam to generate any desired pattern. DOEs can be categorized into different types based on different criteria:

- **Depending on how the light beam is modified:**
  - **Amplitude element:** the amplitude of light is modulated according to the absorption inside the structure. As part of the light beam is absorbed, this type of DOE has low performance in terms of diffraction efficiency [4].
  - **Phase element:** the light propagates without absorption, only with a modulated phase according to the phase shift introduced by the structure.
- **Depending on where the diffraction pattern is observed:**

  - **Fourier element:** the diffraction pattern is observed at the far-field (typically on the order of decimeters beyond the DOE). This type of DOE is usually a diverging element, e.g. beam splitter, beam shaper, where the diffraction pattern is larger than the DOE size.
  - **Fresnel element:** the diffraction pattern is observed at the near-field (typically on the order of millimeters). This type of DOE is usually a converging element, e.g. microlens, where the diffraction pattern is smaller than the DOE size.

- **Depending on the observation direction of the diffraction pattern**
  - **Transmissive element:** the diffraction pattern is observed at the opposite direction with the light source.
  - **Reflective element:** the diffraction pattern is observed at the same direction with the light source.

- **Depending on the DOE thickness compared to the wavelength of light:**
  - **Thin element:** the DOE thickness is equivalent to a phase shift equal to or smaller than 2π.
  - **Thick element:** the DOE thickness is equivalent to a phase shift significantly bigger than 2π.

Given a DOE structure, there are various diffraction models which allow the diffraction pattern to be calculated. This is the forward problem, or the modeling process, in which diffraction theory often cannot be solved analytically, but can yield numerical solutions with some simulation constraints and validity regions. These models can be categorized depending on how the electromagnetic field is treated:

- **Scalar theory:** only one component (E or H) of the electromagnetic field is calculated. Scalar theory is usually used for thin elements, and can be further classified into:
  - **Scalar paraxial models:** a paraxial approximation is used in calculating the amplitude of the diffracted field. This is the traditional diffraction theory, which is only valid for small diffraction angles.
  - **Scalar non-paraxial models:** the amplitude of the diffracted field is calculated rigorously within the scalar domain.

- **Vectorial theory:** all electromagnetic components of the diffracted field are calculated. This is generally required for sub-wavelength DOEs and thick elements.

More interestingly, from a practical applications viewpoint, the inverse diffraction problem is to create a DOE structure that will produce, by diffraction, a desired target wavefront. This is the DOE design and fabrication process, which generally consists of 3 steps [3]. First, the target image is put into a design algorithm to create a Computer Generated Hologram (CGH). This hologram is then fabricated as a DOE. Finally, the optical function and performance of the DOE are verified experimentally on an optical bench. Traditionally, the CGH has an analog (continuous) profile which is difficult to fabricate accurately, leading to an experimental DOE performance significantly lower than the performance in simulation. For this practical reason, the CGH is often quantized into a digital profile with limited numbers of levels (e.g. 2, 4, 8, ... levels) which are easier to fabricate and result in an experimental DOE performance almost the same as in simulation. However, due to the abrupt analog-to-digital conversion, the quantization process often reduces the CGH performance in simulation [5]. Therefore, an optimization algorithm (e.g. iterative transform [6, 7] or genetic algorithm [8, 9]) is necessary to select the CGH with the best performance from all possible set of discrete levels. In summary, to design a digital DOE, a mathematical model for the propagation of the diffracted wave has to be chosen, and an optimization algorithm based on this model is developed.

Motivation

This thesis aims to design, build and optimise DOEs operating in more complex diffraction regimes than scalar paraxial theory, with a view to obtaining higher performance components (improved diffraction efficiency, larger diffraction angles, new wavelengths, ...) and in this way address applications which are for the moment inaccessible. Example applications are DOEs in integrated optics [10], where the devices are becoming more and more compact, which requires DOEs with large diffraction angles. The model design - fabrication - experimental results - feedback used in this thesis allows us to optimize the new complex diffraction regime DOE algorithms and to determine their practical applicability domains in concrete examples. In this way we go beyond the present limitations and make possible the design and fabrication of new families of DOEs for a wider range of applications: wide angle diffraction Fourier DOEs. The practical applications indicated in the thesis were chosen in close collaboration with the industrial partners of Télécom Bretagne (TB) and of its "start-up" company, Holotetrix, specialised in the commercialisation of DOE prototypes and small series production.

In this work, we focus on digital Fourier phase DOEs, as most of the applications addressed at TB use elements of this type. More specifically, in the preliminary chapters, we will model and design test DOEs producing a diffracted output pattern

that can be easily measured and tested, have practical applications and exploit a two-dimensional output field (much work has already been performed on vectorial theory modeled DOEs producing one-dimensional output fields but there are few publications with more complex output patterns). Two examples are a 5x5 spot array and a grid pattern shown in Fig. 1, in which the spot array will be useful for the verification of diffraction efficiency, whereas the grid pattern will be helpful for the verification of diffraction pattern distortion at high diffraction angles.

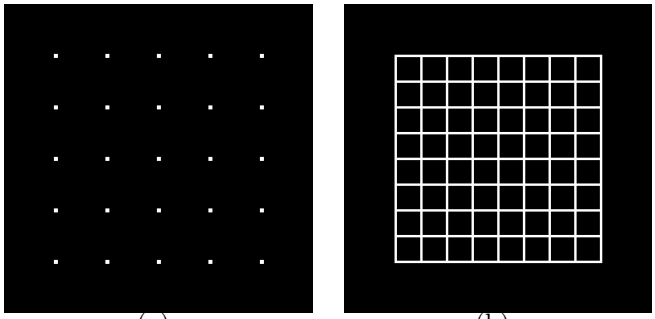


Figure 1 — Output patterns of example DOEs (a) A spot array (for beam splitting applications). (b) A grid pattern (for machine vision applications).

Organisation of the thesis

With respect to the content of this PhD research, the thesis will be organized as follows. Chapter 1 shortly reviews the development history of diffractive optics, including the modeling, design and current fabrication technologies. This chapter also briefly introduces the issues that this thesis attempts to answer, i.e. identifying the limitations of different diffraction models, development of efficient algorithms and new optical lithographic system for the modeling, design and fabrication of high performance DOEs. These issues are further explained in Chapter 2, which analyses in detail the limits of current diffraction theories for the modeling and design of thin and thick elements. For the modeling and design of thin elements, Chapter 3 proposes a scalar non-paraxial propagator and iterative design algorithms based on this model. Chapter 4 presents a vectorial model and a genetic algorithm to overcome the limit of scalar theory in the modeling and design of thick elements. The diffraction models and design algorithms given in these two chapters are verified experimentally in Chapter 5. With these models, we have reached the limit of our fabrication facilities and therefore, some studies on the effects of fabrication limitations on the experimentally observed diffraction pattern are also included in this chapter. To overcome the limits of our current parallel direct-write optical photolithographic system, a new photoplotter based on

parallel Two-Photon Polymerization (2PP) has been built and is described in Chapter 6. Finally, Chapter 7 concludes the results of this work and proposes some directions for the future.

Thesis contributions

- Analysis of the validity regions and computational constraints of different diffraction models in the scalar paraxial, scalar non-paraxial and vectorial regimes.
- Identification of the practical validity of the scalar paraxial regime by the fabrication of test DOEs and the characterisation of their optical performance on an optical bench.
- Development of DOE modeling and design algorithms in the scalar non-paraxial regime. Optimization (design - fabrication - feedback) of each algorithm based on the experimental results.
- Design and fabrication of submicron DOEs (down to 300 nm) where the diffraction angle is up to about 37°. The DOEs were designed using our scalar non-paraxial algorithm and fabricated using Electron Beam Lithography at Karlsruhe Institute of Technology in Germany.
- Investigation of the limits of the Thin Element Approximation by fabricating test binary DOEs and measuring the Hermitian symmetry of their diffraction patterns on the optical bench.
- Development and optimization of a rigorous vectorial diffraction method for the modeling of thick DOEs. Parallelization of the algorithm on a super-computer.
- Study the effects of fabrication errors to the experimental diffraction efficiency and the symmetry of the diffraction pattern.
- Preliminary development of parallel 2PP as a new fabrication technique for rapid fabrication of high resolution (submicron) large diffraction angle DOEs.

List of publications:

- **Journals:**

  1. G. N. Nguyen, K. Heggarty, P. Gérard, B. Serio, and P. Meyrueis, "Computationally efficient scalar non-paraxial modelling of optical wave propagation in the far-field", *Applied Optics*, Vol. 53, Issue 10, pp. 2196-2205, Mar. 2014.

2. G. N. Nguyen, K. Heggarty, A. Bacher, P. J. Jakobs, D. Häringer, P. Gérard, P. Pfeiffer, and P. Meyrueis, "Iterative scalar non-paraxial algorithm for the design of Fourier phase elements", *Optics Letters*, Vol. 39, Issue 19, pp. 5551-5554, Sept. 2014.
3. G. N. Nguyen, K. Heggarty, K. Chikha, P. Gérard, and P. Meyrueis, "Diffraction symmetry of binary Fourier elements with feature sizes on the order of the illumination wavelength and effect of fabrication errors", *in preparation.*
4. A. Liu, G. N. Nguyen, K. Heggarty, and P. Baldeck, "Fabrication of microscale medical devices by parallel two-photon polymerization using Dammann gratings", *in preparation.*

- **Conferences:**

1. G. N. Nguyen, K. Heggarty, P. Gérard, and P. Meyrueis, "Modelling, design and fabrication of diffractive optical elements based on nanostructures operating beyond the paraxial scalar regime " (*Poster presentation*), Journée Futur & Ruptures, 24 Jan., Paris, France, 2013.
2. G. N. Nguyen, K. Heggarty, P. Gérard, and P. Meyrueis, "Iterative scalar algorithm for the rapid design of wide-angle diffraction Fourier elements" (*Oral presentation*), EOSMOC 2013: 3rd EOS Conference on Manufacturing of Optical Components, 13-15 May, Munich, Germany, 2013.

CHAPTER 1 State of the art

In this chapter, the historical development of diffractive optics is quickly reviewed. A brief overview of current fabrication technologies is also included. Finally, in the last section of this chapter, a list of current difficulties and limitations for DOE modeling, design and fabrication is given.

1.1 Development history of diffraction theory

Many books already provide detailed overview of the development of diffractive optics [11–15], only some milestones are mentioned here:

- Diffraction effects were first reported by Grimaldi in 1665, where a small aperture was illuminated by a light source and the light intensity was observed across a plane behind the aperture. Grimaldi discovered a gradual transition from light to dark rather than a sharp geometrical shadow of the aperture.
- In 1673, James Gregory observed the diffraction effects caused by a bird feather. These effects cannot be explained by the corpuscular theory of light, which was the accepted means at that time for explaining rectilinear optical propagation phenomena such as reflection and refraction.
- The first proposal of the wave theory of light that would explain such effects was made by Christian Huygens in 1678. Huygens expressed each point on the wavefront of a diffracted field as a new source of a "secondary" spherical wave. Thus, the wavefront travelling in free space can be found by constructing the "envelope" of all these secondary wavefronts.
- In 1804, Thomas Young performed the double-slit experiment demonstrating interference of light, where light could be added to light and produce darkness. This strengthened the idea that light must propagate as waves.
- In 1818, by making some rather arbitrary assumptions about the amplitudes and phases of Huygens' secondary sources, and by allowing the various wavelets to

mutually interfere, Augustin Jean Fresnel was able to calculate the distribution of light in diffraction patterns with excellent accuracy.

- Over the next two centuries, Rayleigh, Sommerfeld, Fresnel, Fraunhofer and others contributed to the understanding of diffraction, where light is treated as a scalar field. This approach was later shown to be accurate if the structure is large compared to the wavelength and the diffracted field is not observed too close to the structure [16].

- In 1860, Maxwell identified light as an electromagnetic wave. This vectorial approach explains that at the structures, the electric and magnetic fields' components are coupled through Maxwell's equations and cannot be treated independently [17].

- In 1948, Dennis Gabor invented holography, which was originally used in electron microscopy [18, 19]. However, the first optical holograms were only realized in 1962 [20], following the discovery of laser. For his invention and development of holography, Gabor was awarded the Nobel Prize in Physics in 1971.

- The development of diffractive optics design was later influenced by advances in computer technology [21]. In the late 1960s, Adolf Lohmann and Byron Brown calculated the first Computer Generated Holograms (CGHs) and fabricated them using ink as an amplitude absorbing material [22–24]. In 1969, a phase element was reported as more efficient [25].

- In 1970, Goodman showed that the fabrication introduced quantization of the phase and the amplitude in the hologram [5]. To optimize the reconstruction in the presence of quantization, the first design algorithms were implemented [26–28].

- During the 1980s, many techniques were developed for the fabrication of DOEs, e.g. direct laser writing [29] and diamond turning [30]. A significant development was inspired by fabrication technology in electronics, i.e. Electron Beam Lithography (EBL) [31], allowing for manufacturing structures with ever-decreasing feature sizes [32].

- Over the last two decades, Stefan Hell developed stimulated-emission-depletion fluorescence microscopy which overcomes the Abbe diffraction limit [33–35]. For his contribution to the development of super-resolved fluorescence microscopy, Hell was awarded the Nobel Prize in Chemistry in 2014, but his work has also inspired a revolution in optical lithography [36–39].

Design and fabrication continue to work in a push-pull relationship until today. A better design results in the need to improve fabrication and a more accurate fabrication leads to better understanding of design problem. The next section is dedicated to reviewing the existing algorithms for the design of diffractive optical elements.

1.2 Review of design algorithms

Fig. 1.1 illustrates the diffraction geometry of a typical DOE in three dimensions, where U_0 is the illuminating optical field, which is usually approximated as a plane or converging spherical wave [11]. The diffracted field at the DOE plane and at the plane of interest are $U(x_1, y_1; 0)$ and $U(x_2, y_2; z)$, respectively. The forward diffraction problem is to model the diffracted field of a certain structure on the observation plane $U(x_2, y_2; z)$.

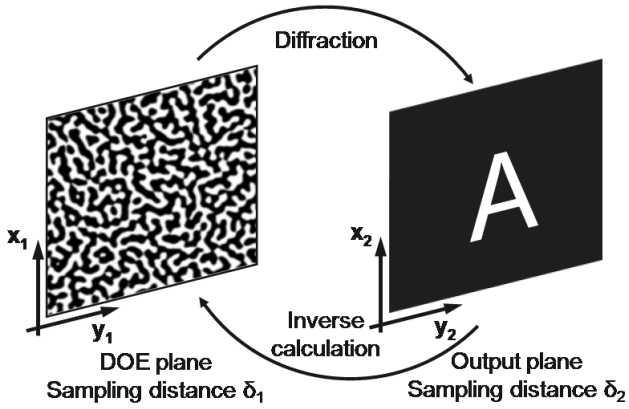


Figure 1.1 — Diffraction geometry.

The inverse diffraction problem is to calculate the DOE that produces a desired output pattern with the best performance. This is the design process, which should take into account as many fabrication constraints as possible. Due to the quantization often occurring in the fabrication [5], the DOE function is quantized into a digital profile with limited number of levels, as shown in Fig. 1.2. For phase elements, the design problem is then equivalent to finding a set of discrete phase values in the DOE plane for which the reconstruction closely matches the target pattern. Because the phase in the target pattern is generally not important for the DOE applications, a random output phase function is usually assigned to the output pattern [3]. This gives an important degree of freedom for optimizing the intensity of the reconstruction pattern.

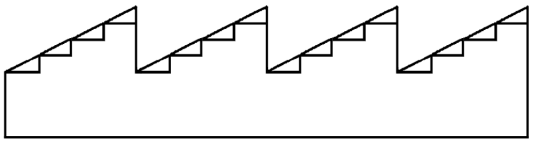


Figure 1.2 — Continuous thickness profile of a blazed grating, and a 4-level quantized profile.

The quality of a CGH is assessed depending on the application, usually by diffraction

efficiency η, uniformity u and/or mean-square-error (MSE) [3, 13, 40]:

$$\eta = \frac{P_{signal}}{P_{total}} \tag{1.1}$$

where P_{signal} and P_{total} are the diffracted power in the the signal window and the total power in the output plane, respectively.

$$u = \frac{P_{max} - P_{min}}{P_{max} + P_{min}} \tag{1.2}$$

where P_{max} and P_{min} represent the maximum and minimum power of the diffraction spots in the reconstruction pattern.

$$MSE = \frac{1}{N^2} \sum_{\mathbf{m}=1}^{N} \sum_{\mathbf{n}=1}^{N} (P_{\mathbf{mn}} - P^{\circ}_{\mathbf{mn}})^2 \tag{1.3}$$

where $P_{\mathbf{mn}}$ and $P^{\circ}_{\mathbf{mn}}$ are the diffraction power of the pixel $(\mathbf{m}, \mathbf{n})$ in the reconstruction and target pattern, respectively.

Depending on whether the optical propagation model from the input field to the output plane can be inverted, a design algorithm can be categorized into:

1.2.1 Unidirectional algorithm

Fig. 1.3 illustrates a general unidirectional algorithm, where the optical propagation can only be calculated in the forward direction. At first, an estimation, usually a random function is generated for the field at the DOE plane. The optical propagation of this field to the output plane is then calculated, and performance constraints, usually intensity requirements, are imposed on the diffracted field. Depending on the specific algorithm, the next estimation for DOE function is made, and the impact of the change on the reconstruction performance is used as the basis for improving the design. Examples of unidirectional design are Direct Binary Search (DBS) [41, 42] and Genetic Algorithm (GA) [8, 9].

Direct Binary Search

The original idea of DBS is to scan all possible set of discrete phase levels for the CGH in a pixel-by-pixel order and find the one with the best reconstruction. This ensures that the algorithm results in the best solution, but at the expense of a huge number of iterations. For example, designing a binary phase DOE with only 8×8 pixels would in theory require $2^{64} \approx 1.8 \times 10^{19}$ cycles (if symmetries are not taken into account). This means an extensive calculation and therefore, the algorithm is limited to DOEs with small number of pixels, even with accelerated versions [41].

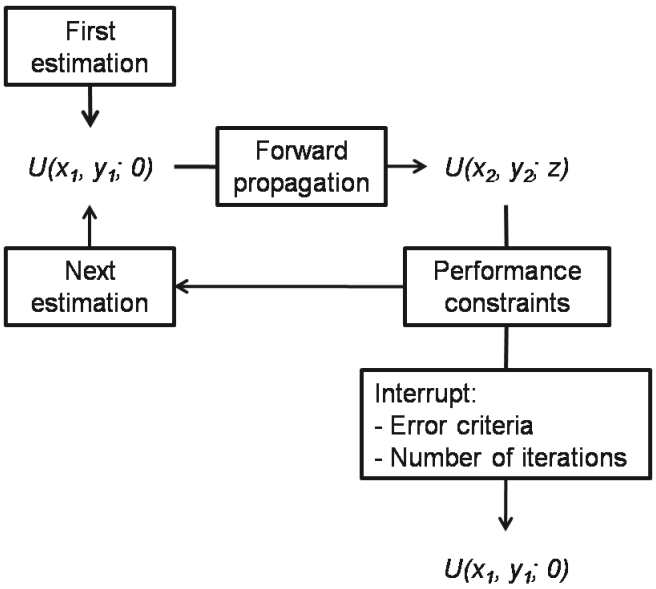


Figure 1.3 — Unidirectional algorithm flowchart.

For this reason, practical DBS algorithms [42–44] generally begin with a random phase function and only make change to one of the CGH's pixel if it has positive effects on the reconstruction. The process is repeated until there are no more single pixel changes, within a limited number of iterations, that can produce a better reconstruction, i.e. the algorithm has converged to an optimum. The drawback of these algorithms is that this optimum is usually a local one, instead of being the global optimum [2], as shown in Fig. 1.4(a).

In order to avoid a local optimum, more complex algorithms have been developed, e.g. Simulated Annealing (SA) [45–47] and Genetic Algorithm (GA) [8, 9].

Simulated Annealing

SA is a stochastic optimization algorithm that was originally modeled after the crystallization of metals when temperature decreases and later adapted for DOE synthesis. Beginning with a first estimation, the forward optical propagation is performed repeatedly as the CGH's pixels are changed randomly. Changes that have positive effects on the reconstruction are accepted. Unlike the DBS, when a change to one pixel has negative effects on the reconstruction, a probability function is used to decide whether the change is also accepted. As the algorithm iterates, the probability of accepting these changes is decreased. In this way, the algorithm can avoid being "trapped" within a local minimum, as would happen with DBS. Finally, the algorithm is stopped when no pixel change is accepted, within a limited number of iterations, i.e. the algorithm has converged to the global optimum. However, due to the random pixel change, SA algorithms converge relatively slowly, often in several thousand iterations [3].

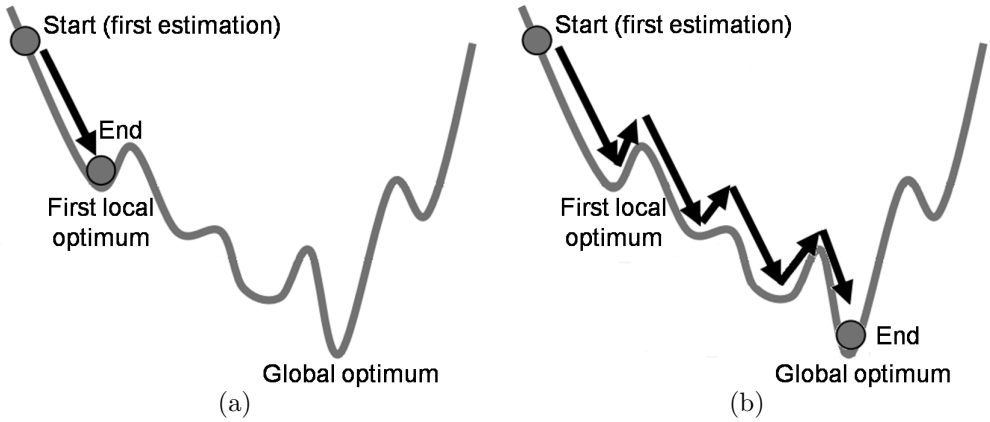


Figure 1.4 — (a) Stagnation effect where the design algorithm stops at a local optimum. (b) Convergence to the global optimum.

Genetic Algorithm

A Genetic Algorithm (GA), as its name implies, is modeled after natural evolution based on breeding, mutation and selection. Typical GAs [8,9] use a large number of CGH estimations, which is called the first generation. The optical propagation of each CGH estimation is calculated and its performance is evaluated. Better-performing estimations are selected for "breeding" to create the next generation. This manner of selective breeding is repeated until an optimum is found. To avoid a local optimum, a few random mutations are inserted to maintain diversity within the successive generations. Therefore, GA generally converges to the global optimum DOE function, usually faster than SA thanks to the pixel change mechanism, often in several hundred iterations [3].

Due to the long calculation time, unidirection algorithms are best suited to the fine optimization of DOEs having a small number of pixels, e.g. spot-array generators. For designing DOEs with a large number of pixels, bidirectional algorithms are often more suitable.

1.2.2 Bidirectional algorithm

If the DOE function can be inversely calculated from the field in the output plane, the designer has the option to use a bidirectional algorithm, which is shown in Fig. 1.5. The first estimation for the output field is usually made by converting the target pattern to amplitude values and generating random phase values. An inverse optical propagation is calculated, and some hologram constraints, such as phase quantization, are applied to get the input field. A forward optical propagation of this field is then simulated, and performance constraints are imposed to update the estimation of the output field. The cycle is repeated until the design converges to the optimum DOE

function, usually in just a few ten to hundred iterations [3]. Examples of bidirectional algorithm are the Iterative Fourier Transform Algorithm (IFTA) [6,48,49] and Iterative Angular Spectrum Algorithm (IASA) [7,50].

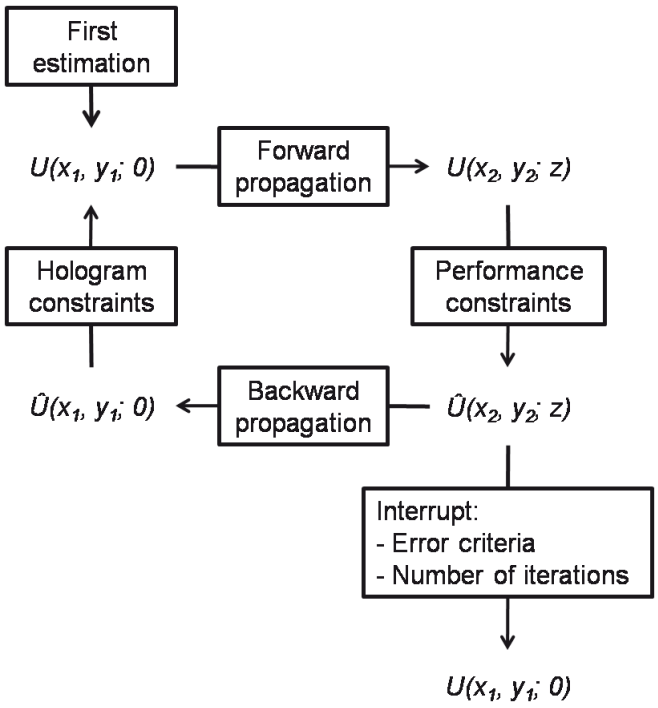


Figure 1.5 — Bidirectional algorithm flowchart.

For designing multilevel DOEs, bidirectional algorithms converge even better and more quickly as the quantization constraints are less severe. Conversely, unidirectional algorithms for multilevel DOEs converge much more slowly as the number of possible structures to test increases greatly. Therefore, bidirectional algorithms are generally accepted as best practical algorithms in most cases [51]. Independent of the algorithm used, the design should account for fabrication parameters and contraints. For this reason, common fabrication technologies and their limitations will be described briefly in the following section.

1.3 Review of fabrication technology

1.3.1 Diamond machining

One of the first techniques for DOE fabrication is diamond machining, where the diffractive microstructures can be generated directly through mechanical removal of optical material. These mechanical methods use a sharp and hard, usually diamond tip to scrape away the optical material in a manner based on computer control. Fig.

1.6(a) illustrates the ruling process of a blazed grating using a diamond tip adapted to grating geometry. Diffractive microstructures can also be generated by turning process, as shown in Fig. 1.6(b).

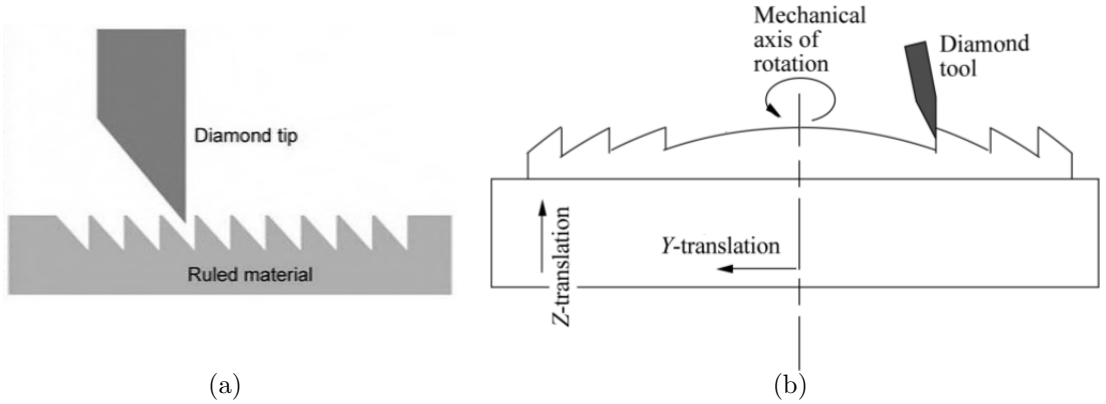


Figure 1.6 — (a) Mechanical ruling of a blazed grating using a diamond tip adapted to grating geometry [2]. (b) Diamond turning of a microlens [3].

Due to the finite size of the tip, the tool cannot create a perfectly accurate surface profile. Advanced diamond machining methods can produce high quality elements, but they are relatively slow and therefore are commonly used for DOE prototyping [3]. Moreover, DOE fabrication using diamond ruling or turning is generally limited to either straight line or circularly symmetrical elements. Most current technologies usually use lithographic processes to fabricate complex micro- or nano-structures by patterning a layer of a photosensitive material, generally photoresist on a substrate, or by etching the substrate itself. The patterning of the photoresist is performed by exposing it to an optical or electron beam, either through an optical mask or directly (without using a mask). A series of chemical treatments then etches the exposed pattern into the photoresist, or the substrate.

1.3.2 Mask based lithography

In general, the most common processes for DOE fabrication involve photolithographic methods that are derived from the electronics industry. They are based on the same processes used to fabricate integrated circuits, which is illustrated in Fig. 1.7. Firstly, a photoresist layer is deposited on a substrate by spin coating. The photoresist is usually a viscous, liquid solution, and the substrate is spun rapidly to produce a relatively uniform layer. The thickness of the photoresist layer depends on the viscosity and the spin coating speed. Secondly, a binary mask of alternating transparent and opaque areas is fabricated using some type of pattern generator. The mask is laid on a substrate coated with a thin layer of photoresist, which is exposed to ultraviolet light

through the mask. The exposure to light causes a chemical change that allows some of the photoresist to be removed by a special solution, called "developer". Positive photoresist becomes soluble in the developer when exposed, while with negative photoresist, unexposed regions are soluble in the developer. After the resist is developed, a pattern is created in the photoresist layer. The substrate is then etched into the substrate until the required depth is reached. The photoresist pattern is then removed, resulting in a binary (two-level) element.

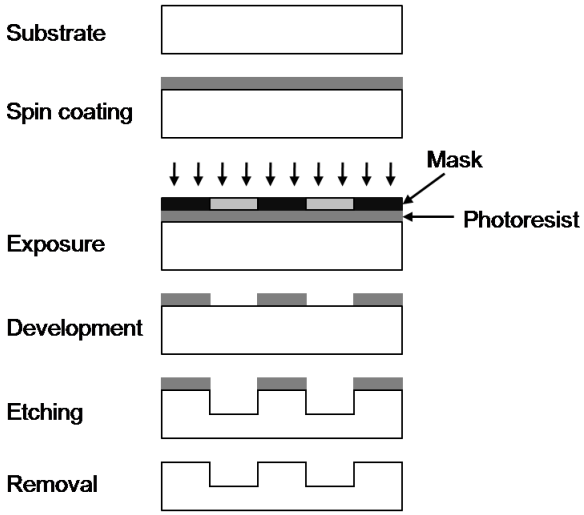


Figure 1.7 — Basic procedure of binary DOE fabrication process using mask based lithography.

If a multi-level DOE is required, the etched substrate from the previous step is re-coated and re-exposed to a second binary mask, as shown in Fig. 1.8. After development, the substrate is again etched until the required depth is reached. The result is a four-level profile. The process is repeated until a 2^L level DOE is created. DOEs fabricated using this approach are commonly referred to as binary optics.

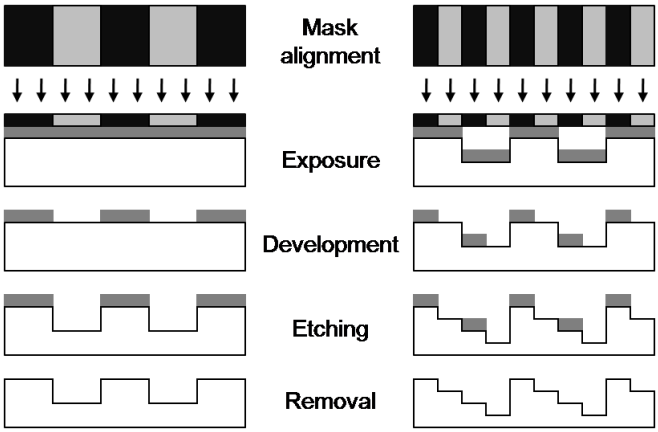


Figure 1.8 — Basic procedure of 2^L level DOE fabrication process using mask based lithography.

One drawback of this approach is the need for multiple processing steps to fabricate multi-level DOEs. This increases production costs over single-step procedures, not to mention that multiple mask alignments can introduce errors which decrease the DOE performance. Another disadvantage is the need for manufacturing fixed photomasks. For these reasons, maskless lithography, where lithography patterns can be changed programmably, is increasingly being used. For DOE fabrication at TB, we generally use a parallel direct-write lithography, which is available in the Optics department's cleanroom.

1.3.3 Parallel direct writing

Fig. 1.9(a) shows the basic principle of our parallel direct-write photoplotter. Firstly, a photoresist layer is spin-coated on a substrate, usually glass. We use the photoresist S1800 series from Micro Resist Technology, which is a positive photoresist. This series has different viscosities and therefore allowing us to put down photoresist layers with different thicknesses ranging from a few hundreds nm to above 10 μm. The error in the uniformity of the photoresist layer in our spin-coating process is about 20 nm. The photoresist is then exposed to a pattern of intense light with a wavelength to which the photoresist is active. In our case, we use a lamp at 436 nm.

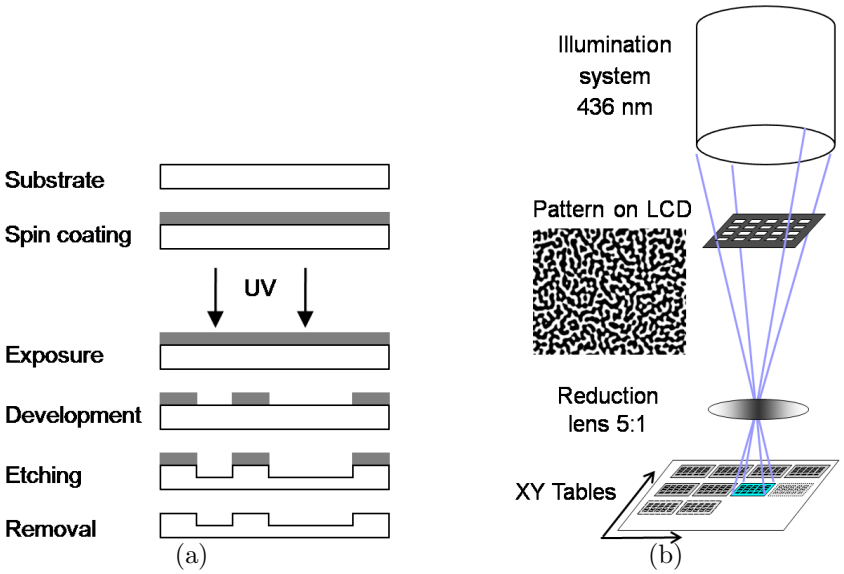


Figure 1.9 — (a) Basic procedure of direct laser writing. (b) Principle of parallel direct-write lithography at TB.

Fig. 1.9(b) illustrates the idea of our parallel direct-write photoplotter, where a programmable Spatial Light Modulator (SLM) is used as a reconfigurable mask. This SLM is in fact a 1050×1400 pixel Liquid Crystal Display (LCD), where the intensity of light passing through each pixel can be controlled. A reduction lens is used to image

the pattern on the LCD into the photoresist layer. By this way, the write beam is parallelized and an area of about 6×4 mm^2 (1.5 Mpixels) can be exposed at the same time. Bigger areas can be exposed in a short time (1 cm^2/min) by moving the nano-precision 2D translational stage.

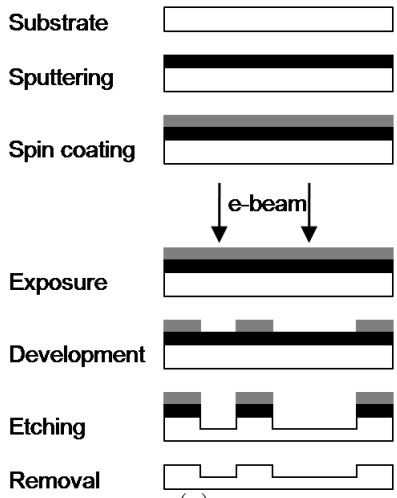
After exposure, the substrate is put into the developer solution to etch the exposed pattern into the photoresist layer. Etching into the substrate might be necessary depending on specific applications, but it is usually not needed in our process, as the controlled thickness of the photoresist layer can itself produce the desired dephasing of the incident wavefront. Hence, to obtain a π phase shift for binary (i.e. 2 levels) phase DOE, the spin-coating speed should be chosen so that the thickness of the photoresist layer is $d = \lambda/2(n-1)$ [3], where n is the refractive index of the photoresist material at the DOE working wavelength λ. The fabricated structure is then measured under an interferometric microscope, where the lateral dimensions and the etching depth can be verified.

Our fabrication process has been shown to be cost-effective and particularly adapted to DOE prototyping [52]. However, its resolution limit is still the same as that of a visible spectrum direct-write lithography, which is about 1 μm. To test our scalar non-paraxial diffraction propagator and design algorithms with visible spectrum DOEs, and to fabricate our submicron DOEs, we also collaborated with Karlsruhe Institute of Technology (KIT) in Germany for the use of Electron Beam Lithography [53].

1.3.4 Electron Beam Lithography

The principle of EBL is similar to that of a serial direct-write lithography, except that an electron beam is used for exposure. To avoid electric charging, a metallic layer (usually Chromium) is deposited on the substrate by sputtering, before spin coating of photoresist, as shown in Fig. 1.10(a). The photoresist used here is Poly-methyl methacrylate (PMMA), which is a negative resist. This photoresist layer is then exposed directly (without a mask) and sequentially by moving the two-dimensional translational stage according to the DOE pattern. After developing, the exposed pattern is etched into the photoresist layer. As the metallic layer is highly reflective, in order to fabricate transmissive elements, a series of chemical treatments, e.g. Reactive Ion Etching (RIE), is necessary to etch the exposed pattern into the substrate and to remove the unexposed photoresist as well as the metallic layer.

Fig. 1.10(b) shows an image of the EBL system at KIT. This system is able to fabricate structures of 200 nm in thick PMMA layer (resist thickness of 3200 nm) or down to 20 nm in thin PMMA layer (resist thickness of 100 nm). Due to the need of RIE in our DOE fabrication, the real critical dimension is the resolution limit of RIE system, which is about 100 nm. The main drawbacks are the high cost and the slow



(a)

(b)

Figure 1.10 — (a) Basic procedure of EBL. (b) An electron beam writer, which is a Vistec VB6 UHR-EWF (Ultra High Resolution-Extra Wide Field, photo courtesy of the KIT).

exposure process of EBL system, resulting in a long writing time (several hours) for relatively small areas (usually a few mm^2).

1.4 Design and experimental requirements

1.4.1 Non-pixelated elements

Although our fabrication facilities have been advancing, it is important to note that the isolated pixels in the designed DOE will be rounded in the fabrication [54], resulting in an experimental performance considerably lower than that predicted by simulation. In order to avoid pixel rounding, the DOEs will have to be fabricated at a pixel size which is several times bigger than the resolution limit. This sort of DOE makes bad use of the available resolution of the fabrication machine, particularly our direct write machine. As a result, the diffraction pattern will contain many higher diffraction orders which reduce the real experimental diffraction efficiency of the DOE in the useful diffraction order [55]. Notice that the simulations generally don't allow for this in their calculation of the diffraction efficiency so the experimental efficiency is usually considerably worse than the simulations suggest.

For these reasons, it is generally much more efficient to avoid isolated pixels in the design, usually by zero padding of the target image in a field of zeros much bigger than required to strongly oversample the DOE. This technique leaves lots of space for amplitude freedom in the output plane (zones outside the signal window) so the design

algorithm has more space to put the noise and it can converge to better solutions [56]. The advantage of this techniques is that it reduces the number of isolated pixels in the DOE design, simplifies the fabrication and greatly improves the practical performance.

Another technique to enhance the DOE practical performance is to replicate the DOE several times in fabrication to simplify illumination by an expanded laser beam and to suppress the speckle noise in the diffraction pattern [26, 57].

1.4.2 Replicated structures

Figure 1.11 shows several configurations of the DOE illumination in practice, where the square represents the DOE area and circle indicates the illuminating area. If the illuminating area is bigger than the DOE size, the light outside the DOE region will contribute to background intensity (for Fresnel elements) or zero order (for Fourier elements) in the diffraction pattern. On the other hand, if parts of the whole DOE region are not illuminated, the diffraction efficiency and uniformity will be reduced compared to the design [58]. One solution is to spatially repeat the original DOEs on a contiguous square grid to obtain a total DOE size several times in fabrication to obtain a bigger total DOE size, which would be easier to illuminate in experiment.

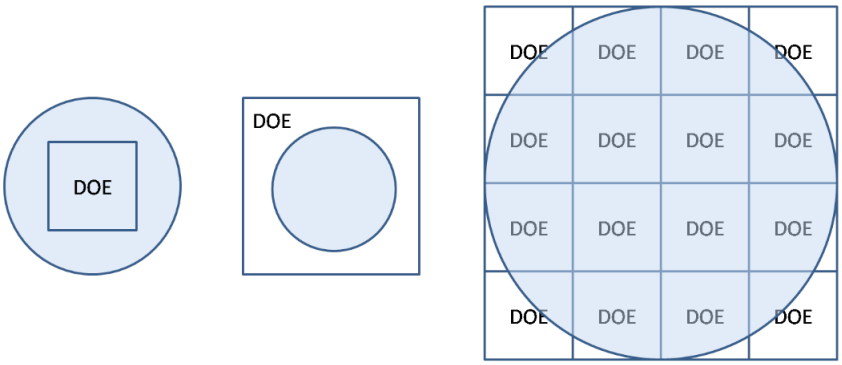
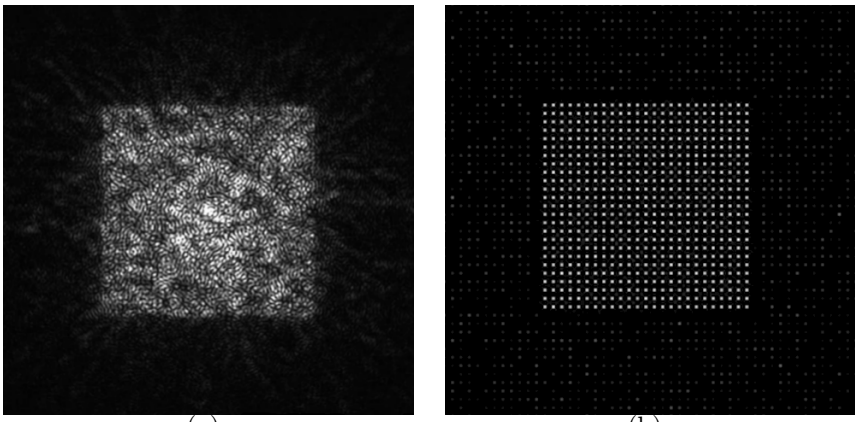


Figure 1.11 — Several configurations of possible DOE illumination.

Another advantage of this technique is that it reduces the speckle effect in the experimental diffraction pattern. This effect is due to the size of the illumination area, which is often considered as infinite in the DOE calculation. As a result, the experimental diffraction pattern is the simulated one convolved with that of the circular aperture. This means that each pixel in the diffraction pattern is broadened by the Airy disk pattern. If there are many closely separated spots in the target image, the fields of these spots will interfere. Since the phase of the pixels is randomly assigned in simulation, the interference leads to a random series of maxima and minima on the observing screen, as shown in Fig. 1.12. As the illumination area is bigger, the diffraction pattern of the aperture is smaller, reducing the speckle noise in the diffraction pattern.



(a) (b)

Figure 1.12 — Diffraction pattern of a spot array DOE: (a) with and (b) without speckle effect.

In summary, for these practical reasons, CGHs are generally calculated by zero padding of the target images in big fields of zeros, leading to large numbers of samples, usually 1024×1024 pixels. Assuming that each pixel is 1 μm, the size is usually about 1×1 mm^2. The designed DOEs are then spatially replicated in fabrication on a contiguous square grid to obtain a total DOE size of at least 4×4 mm^2.

1.5 Thesis problem formulation

As discussed in the previous sections, the push-pull relationship between design and fabrication has driven the development of diffractive optics for many years. Although many high performance design algorithms have been implemented, the quality of the calculated DOEs strongly depends on the diffraction model used in the design. Traditional diffraction theory, which is the scalar paraxial model, is only accurate for thin elements with feature size much bigger than the illumination wavelength [11]. Meanwhile, fabrication technologies are now able to fabricate thin and thick elements with feature size on the order of or smaller than the illumination wavelength. These DOEs operate beyond the validity of scalar paraxial diffraction regime, leading to the need for new modeling and design. Mathematical models of more complex (scalar non-paraxial, vectorial) diffraction theories are available but often have strongly limited validity regions and computational complexity constraints.

This thesis aims to analyse, identify and overcome the limitations of the scalar paraxial and more complex diffraction models. Their practical validity domains are verified experimentally by fabricating large diffraction angle DOEs and characterizing their optical performance on an optical bench. A new propagator is developed for the modeling of far-field diffraction in the scalar non-paraxial domain. Measurement results

of test DOEs fabricated at 1 μm using our parallel direct-write lithography show that our propagator expands the applicable domain of scalar theory beyond the validity of scalar paraxial regime, with very little extra computational expense. We then develope iterative algorithms based on this scalar non-paraxial propagator. Experimental results of test DOEs fabricated at 400 and 300 nm using EBL at KIT verify the accuracy of our modeling and design for submicron structures. As the design has been calibrated, a parallel Two-Photon Polymerization is currently being built as a new fabrication technique for rapid manufacturing of submicron DOEs. On the other hand, as wide angle may also be obtained with microscale structures having deeper etching depth, i.e. those with the phase shift $> 2\pi$, which can now be fabricated using our current photoplotter, we develope a rigorous vectorial method for the modeling and design of thick DOEs. In these ways, we hope to design and fabricate higher performance components for research and industrial applications which are for the moment inaccessible.

The next chapter is dedicated to reviewing and analysing the validity regions and computational constraints of different diffraction models.

CHAPTER 2 Diffraction models

In this chapter, we will introduce the theoretical background of the various different ways of modeling the diffraction process. These models constitute the heart of the algorithms developed to calculate and optimize Diffractive Optical Elements (DOEs). The validity regions and computational constraints will be analysed in details in order to identify the practical limits of different diffraction models.

2.1 Introduction

Fig. 2.1 illustrates the diffraction geometry of a typical DOE in two dimensions, where U_0 is the illuminating optical field, which is usually a plane or converging spherical wave. The diffracted field at the DOE plane and at the plane of interest are $U(x_1, y_1; 0)$ and $U(x_2, y_2; z)$, which are sampled at intervals δ_1 and δ_2, respectively. The rigorous solutions [17] for the diffraction problem are given by the Maxwell's equations (for linear, isotropic and nondispersive medium in the absence of charges $\rho = 0$ and currents $J = 0$):

$$\vec{\nabla} \times \vec{\mathcal{E}} = -\mu \frac{\partial \vec{\mathcal{H}}}{\partial t} \tag{2.1}$$

$$\vec{\nabla} \times \vec{\mathcal{H}} = \epsilon \frac{\partial \vec{\mathcal{E}}}{\partial t} \tag{2.2}$$

$$\vec{\nabla} \cdot \left(\epsilon \vec{\mathcal{E}}\right) = 0 \tag{2.3}$$

$$\vec{\nabla} \cdot \left(\mu \vec{\mathcal{H}}\right) = 0 \tag{2.4}$$

In these expressions, $\vec{\mathcal{E}}$ and $\vec{\mathcal{H}}$ are the electric and magnetic field, which are functions of both position $\vec{r}(x, y; z)$ and time t. The symbols $\times$ and $\cdot$ represent the vector cross product and vector dot product, with $\vec{\nabla} = \frac{\partial}{\partial x}\vec{x} + \frac{\partial}{\partial y}\vec{y} + \frac{\partial}{\partial z}\vec{z}$, where $\vec{x}$, $\vec{y}$, and $\vec{z}$ are unit vectors in the x, y, and z directions, respectively. μ and ϵ are the permeability and permittivity of the medium in which the wave is propagating, respectively, where $\mu = \mu_0$ is a constant. In general, the electromagnetic field is polychromatic, but can be

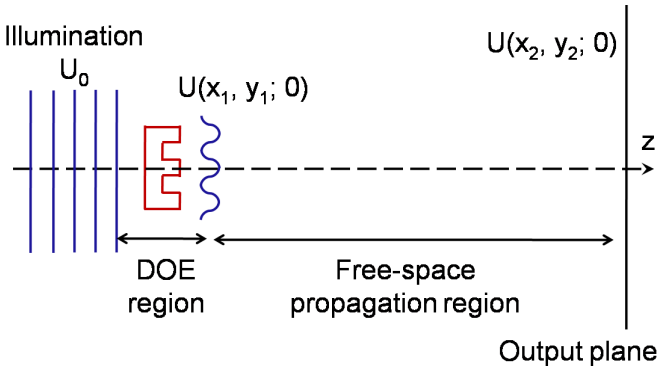


Figure 2.1 — Two-dimensional diffraction geometry, with the problem divided into different regions, where U stands for both the electric and magnetic fields.

decomposed into a sum of monochromatic fields. Such fields are often closely approximated in practice and easier to analyze, where the electric and magnetic field can be written explicitly as:

$$\mathcal{E}(\vec{r}, t) = Re\{E(\vec{r}) \exp(-j2\pi\nu t)\} \tag{2.5}$$

$$\mathcal{H}(\vec{r}, t) = Re\{H(\vec{r}) \exp(-j2\pi\nu t)\} \tag{2.6}$$

where $Re\{\}$ signifies "real part of" and ν is the optical frequency. Respectively, $E(\vec{r})$ and $H(\vec{r})$ are electric- and magnetic-field complex-amplitude vectors: $\vec{E} = (E_x, E_y, E_z)$, $\vec{H} = (H_x, H_y, H_z)$. However, the Maxwell's equations generally have no analytical solutions, but can be solved numerically using several vectorial methods [2], e.g. the Rigorous Coupled Wave Analysis (RCWA) or Finite-Difference Time-Domain (FDTD) methods, as will be reviewed in the next sections.

Although in theory it is possible to calculate directly the diffraction pattern at the output plane $U(x_2, y_2; z)$, these methods in practice are computationally very expensive and often limited to 2D or small/periodic 3D structures [59]. A more common approach to model more complex types of diffractive structures which have a wider range of practical applications is to divide the problem into two regions where the fields $U(x_1, y_1; 0)$ and $U(x_2, y_2; z)$ can be solved separately. Traditional diffraction methods use the TEA for the DOE region and a scalar propagation in the free-space region, such as the Fraunhofer approximation or the Angular Spectrum Method (ASM) [11]. The original Radiation Spectrum Method (RSM) [60] also uses the TEA for the calculation of $U(x_1, y_1; 0)$, but the free-space region is modeled using a vectorial propagation. Some methods based on a vectorial model for the DOE region (FDTD) coupled with a scalar free-space propagation (ASM) have also been described in the literature [7, 61], while the most recent work has focused on the coupling of FDTD in the DOE region with a vectorial method (RSM) for the free-space region [62, 63]. It appears that the FDTD + RSM model is equivalent to the FDTD + vectorial ASM [64], as the RSM and the vectorial ASM obtain the same calculated fields, with the only assumption that the field propagating in free space is considered as the sum of only forward plane waves

(no backward plane waves) [60]. This assumption corresponds to the classical DOE applications where there is only one illumination on one side of the DOE (i.e. no illumination in the inverse direction), as shown in Fig. 2.1. Theoretically the FDTD + vectorial ASM or FDTD + RSM are then fully vectorial models which allows for the accurate calculation of non-paraxial electromagnetic field components in classical DOE applications. However, they are not pratically usable for far-field propagation due to their sampling constraint, as will be shown later in this chapter.

Details on the advantages and disadvantages of the different approaches will be given in the following sections.

2.2 DOE region

2.2.1 Thin Element Approximation

In the Thin Element Approximation (TEA), the diffraction inside the structure is neglected and it is assumed that the illumination wave transmits through the structure without changing the propagation direction. The field after the structure can be calculated approximately as a multiplication of the incident field U_0 with the transmission function of the structure $t(x, y)$ [11]:

$$U(x_1, y_1; 0) = t(x, y) \cdot U_0 \tag{2.7}$$

In general $t(x, y)$ is a complex function:

$$t(x, y) = T(x, y) \cdot \exp\{j\phi(x, y)\} \tag{2.8}$$

where $T(x, y)$ is the amplitude modulation ($0 \leq T(x, y) \leq 1$) and $\phi(x, y)$ is the phase shift due to the transmission through the structure. For example, a square aperture, which is an amplitude DOE, has the transmission function given by:

$$t(x, y) = \begin{cases} 1 & \text{inside the aperture} \\ 0 & \text{outside the aperture} \end{cases} \tag{2.9}$$

For phase elements, the light propagates without absorption, i.e. $T(x, y) = 1$, but with a modulated phase according to the phase shift introduced by the structure. Fig. 2.2 shows a L level phase DOE illuminated by a plane wave at normal incidence in free-space. The optical path difference (OPD) between two rays transmitted through a thickness step δd can be calculated as [65]:

$$OPD = \delta d \cdot n - \delta d \cdot 1 = (n - 1)\delta d \tag{2.10}$$

which results in the phase step:

$$\delta\phi = k \cdot OPD = \frac{2\pi}{\lambda}(n-1)\delta d \tag{2.11}$$

As the phase is periodic with the period of 2π, the phase step should be $\delta\phi = 2\pi/L$. The total thickness of the DOE is given by $d = (L-1)\delta d$, resulting in [2, 65]:

$$d = (L-1)\frac{\lambda}{2\pi(n-1)}\delta\phi = (L-1)\frac{\lambda}{2\pi(n-1)}\frac{2\pi}{L} = \frac{L-1}{L}\frac{\lambda}{n-1} \tag{2.12}$$

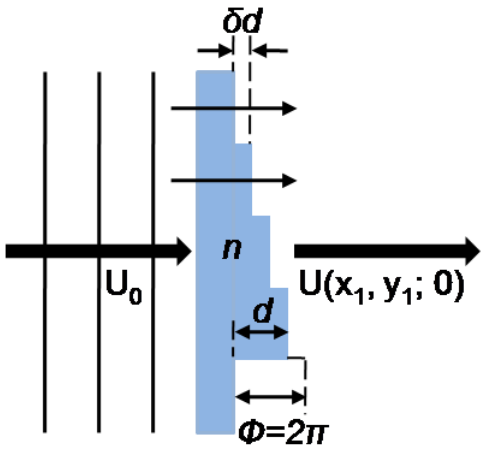


Figure 2.2 — Propagation of two light rays through a L level phase DOE in the TEA.

This approach is sometimes called as the transmission function approximation [66], which has been shown to be accurate and efficient for DOEs having thickness equivalent to a phase shift smaller than 2π and feature size much bigger than the wavelength. However, for the cases where there is interaction between the incident field and the structures (e.g. for DOEs having the sub-wavelength feature sizes, or thickness equivalent to a phase shift significantly bigger than 2π), vectorial theory has to be used in order to solve the diffracted field $U(x_1, y_1; 0)$ exactly [11].

2.2.2 Vectorial theory

The Finite-Difference Time-Domain (FDTD) method is perhaps the most straightforward numerical solution to the Maxwell's equations. The Rigorous Coupled Wave Analysis (RCWA), also called as Fourier Modal Method (FMM) [67], is another vectorial method but can only be applied for infinite periodic structures. In this thesis, as we aimed to develop a vectorial diffraction model for a wider range of applications, the FDTD is more suited. The basic FDTD principle will be summerized here, details for the RCWA/FMM can be found in [68–70].

In the FDTD, finite differences are employed as approximations to both the spatial and temporal derivatives. The second order approximation for the derivative of a function $f(x)$ at a point x_0 is given by:

$$\left.\frac{\partial f(x)}{\partial x}\right|_{x=x_0} \approx \frac{f(x_0+\Delta x)-f(x_0-\Delta x)}{2\Delta x} \tag{2.13}$$

where Δx is a small sampling step. The approximation is more accurate with smaller value of sampling step Δx [71]. Full details and a mathematical derivation of FDTD can be found in [72], the basic principle can be summerized as follows :

1. Sample the calculation region into a high resolution spatial and temporal grid. Discretize the electric and magnetic fields in both space and time.
2. Replace all the derivatives in the Maxwell's equations with finite differences. Solve the difference equations to obtain the relations between the (unknown) future fields in terms of (known) past fields and between the neighboring spatial points.
3. Propagate the electric and magnetic fields one time step into the future so that they are now calculated fields.
4. Repeat the calculation until the fields have been propagated over the desired duration.

Figure 2.3 illustrates one-dimensional FDTD using a more convenient discretization, where the electric and magnetic fields are evaluated as half-step apart (in both space and time) [73]:

$$\left.\frac{\partial f(x)}{\partial x}\right|_{x=x_0} \approx \frac{f(x_0+\frac{\Delta x}{2})-f(x_0-\frac{\Delta x}{2})}{\Delta x} \tag{2.14}$$

The relations between the future fields in terms of past fields and between the neighboring spatial points are as follows:

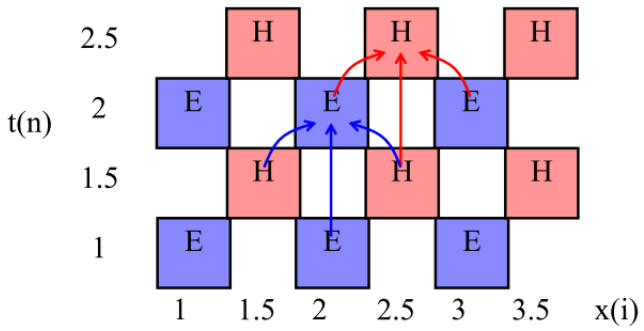


Figure 2.3 — Yee grid and leap-frog time steps.

$$E|_i^{n+1} \approx E|_i^n + \frac{1}{\epsilon|_i}\frac{\Delta t}{\Delta x}\left[H|_{i+0.5}^{n+0.5} - H|_{i-0.5}^{n+0.5}\right] \tag{2.15}$$

$$H_{i+0.5}^{n+0.5} \approx H|_{i+0.5}^{n-0.5} + \frac{1}{\mu_0}\frac{\Delta t}{\Delta x}\left[E|_{i+1}^{n} - E|_{i}^{n}\right] \tag{2.16}$$

This discretization calculates different field components at different grid locations. As a consequence, the output field components must be interpolated to a common point [74].

The spatial grid resolution Δx must be fine enough to display the finest structures of the ϵ distribution and the fields: $\Delta x \leq \lambda/(20n_{max})$, with n_{max} being the highest refractive index in the simulation domain [72]. Furthermore, in order for the method to converge to the correct result, the temporal step must be smaller than the time needed for the electromagnetic wave to propagate to the next spatial grid location $\Delta t \leq \Delta x/c$, where c is the speed of light in the propagation medium. In other words, the temporal resolution must be fine enough to "capture" the propagation of light between two nearest neighbors. This condition is expressed by $\Delta t = S\Delta x/c$, where S is the Courant factor, which must satisfy $S \leq 1/\sqrt{ND}$, with ND is the number of dimensions of the simulation [71].

Due to these strong spatial and temporal converging conditions, in practice the FDTD calculation requires an extremely high memory usage and very long calculation time. For example, a 3D FDTD modeling the traveling of light through a free-space region of 16×16×32 μm using 20 samples/μm requires about 10 GByte of Random-Access Memory (RAM). The simulation time on a Linux computer with an Intel core i5 @ 3.2 GHz processor and 16 GByte of RAM is about 50 minutes. If a structure of refractive index $n = 2$ is added, the resolution has to double. The required amount of RAM increases by 8, which is beyond the amount of available memory of most current desktop PCs. The computational time increases by at least 16 [75], which means more than 10 hours for a single FDTD propagation, not to mention that a design algorithm often requires hundreds of iterative propagations.

For these reasons, FDTD is often limited to a small region, usually about ten times the wavelength around the DOE. For infinite periodic structures, the RCWA/FMM should be used, since the computational effort is much lower than the FDTD, and the far-field diffraction pattern can be obtained directly [68–70], as will be shown in Chapter 4. If an object is very small compared to a wavelength, effective medium theory [63] generally provides more efficient solutions. Alternatively, if the wavelength is very small compared to the physical features of interest, the TEA is a much more efficient way to solve the problem.

Once the field at the DOE plane $U(x_1, y_1; 0)$ has been solved, the diffracted field $U(x_2, y_2; 0)$ on the far-field observation plane (usually about 10 cm beyond the DOE) can be calculated using a free-space propagator. In practice, FDTD is usually not applicable for this region due to the extensive memory and calculation requirements and another method is necessary for the calculation of the free-space propagation to the observation plane.

2.3 Free-space propagation region

2.3.1 From vectorial to scalar theory

In the free-space region, as the propagation medium is homogeneous (ϵ = constant), it is possible to reduce the Maxwell's equations into a single scalar wave equation. Details on the mathematical derivation can be found in [11].

$$\nabla^2 u\left(\vec{r},t\right) - \frac{n^2}{c^2}\frac{\partial^2 u\left(\vec{r},t\right)}{\partial t^2} = 0 \tag{2.17}$$

where the scalar field $u\left(\vec{r},t\right)$ stands for any of the $x-$, $y-$ or $z-$ directed components of the vector fields $\vec{E}$ and $\vec{H}$. This means that all components of the electromagnetic field can be solved, if required, using this equation despite of its scalar being.

For monochromatic wave, the scalar field can be written explicitly as:

$$u\left(\vec{r},t\right) = Re\left\{U\left(\vec{r}\right)\exp\left(-j2\pi\nu t\right)\right\} \tag{2.18}$$

where $U\left(\vec{r}\right)$ is a complex function of position:

$$U\left(\vec{r}\right) = A\left(\vec{r}\right)\exp\left[-j\phi\left(\vec{r}\right)\right] \tag{2.19}$$

where $A\left(\vec{r}\right)$ and $\phi\left(\vec{r}\right)$ are the amplitude and phase of the wave, respectively. It follows from Eq. (2.17) that $U\left(\vec{r}\right)$ must obey the time-independent equation:

$$\left(\nabla^2 + k^2\right) U\left(\vec{r}\right) = 0 \tag{2.20}$$

where $k = 2\pi n\nu/c = 2\pi/\lambda$ is the wave number and $\lambda = c/\left(n\nu\right)$ is the wavelength in the dielectric medium. Eq. (2.20) is known as the Helmholtz equation, which is valid for any monochromatic optical disturbance propagating in homogeneous dielectric medium ($n \geq 1$). This scalar equation gives an exact solution to Maxwell's equations for the propagation of an electromagnetic field in free-space. It can be further reduced to a simpler and more usable expression, which is the Angular Spectrum Method (ASM).

2.3.2 The Angular Spetrum Method

In the ASM, the complex field across the source plane is decomposed into a sum of plane waves propagating with direction cosines α, β, γ, where $\alpha^2 + \beta^2 + \gamma^2 = 1$, as illustrated in Figure 2.4. For this reason, this method is also called as the angular spectrum of plane waves [76–78]. The angular spectrum of the field across a plane at a distance z is given by a Fourier Transform (FT):

$$A\left(f_x, f_y; z\right) = \iint_{-\infty}^{\infty} U\left(x_2, y_2; z\right)\exp\left[-j2\pi\left(f_x x_2 + f_y y_2\right)\right]\,\mathrm{d}f_x\mathrm{d}f_y \tag{2.21}$$

where the FT operation can be seen as a decomposition of a complicated function into a collection of simpler functions, and $f_x = \alpha/\lambda$, $f_y = \beta/\lambda$ are the spatial frequencies [11]. The field can therefore be represented by an Inverse Fourier Transform (IFT) of its spectrum:

$$U\left(x_2, y_2; z\right) \propto \iint_{-\infty}^{\infty} A\left(f_x, f_y; z\right) \exp\left[j2\pi\left(f_x x_2 + f_y y_2\right)\right] \mathrm{d}f_x \mathrm{d}f_y \tag{2.22}$$

Note that the field is expressed as "proportional to" by the symbol $\propto$, since we are not interested in the exact values but in the normalized field distribution. The physical meaning is that in practice, all measured light fields must be normalized. Therefore the constant complex number in front of the integral is not necessary to be written explicitly.

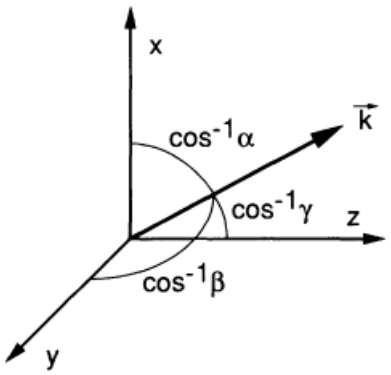


Figure 2.4 — The wave propagation direction.

In addition, $U(x_2, y_2; z)$ must satisfy the Helmholtz equation (2.20). Solving the second order differential equation leads to an elementary solution [11]:

$$A\left(f_x, f_y; z\right) = A\left(f_x, f_y; 0\right) G(f_x, f_y; z) \tag{2.23}$$

where $G(f_x, f_y; z) = \exp\left(jkz\sqrt{1-\lambda^2(f_x^2+f_y^2)}\right)$ is the Transfer Function and $A(f_x, f_y; 0)$ is the FT of the field at the DOE plane $U(x_1, y_1; 0)$. The diffraction pattern at the output plane is therefore given by:

$$U\left(x_2, y_2; z\right) \propto \iint_{-\infty}^{\infty} A(f_x, f_y; 0) G(f_x, f_y; z) \exp\left[j2\pi(f_x x_2 + f_y y_2)\right] \mathrm{d}f_x \mathrm{d}f_y \tag{2.24}$$

In simulation, the Angular Spectrum equation (2.24) can be evaluated directly as:

$$U(x_2, y_2; z) \propto IFT\left\{A(f_x, f_y; 0) G(f_x, f_y; z) \cdot circ\left(\lambda\sqrt{f_x^2+f_y^2}\right)\right\} \tag{2.25}$$

The computational complexity is $O(N^2(1+2\log_2 N))$. Notice that if the source field is sampled uniformly at distance δ_1, the spatial frequencies f_x, f_y will be equally spaced

with $\delta_f = 1/(N\delta_1)$ due to the discrete FT property. The maximum direction cosines α, β are therefore $\alpha_{max} = \beta_{max} = N\lambda\delta_f/2 = \lambda/(2\delta_1)$. In reality, all direction cosines must be real, which means that α, β have to satisfy $\alpha^2 + \beta^2 < 1$. However, this is not always the case in simulation, i.e. if $\delta_1 < \lambda/\sqrt{2}$. Under such a condition, there will be some positions where $\alpha^2 + \beta^2 > 1$, so γ is imaginary since $\alpha^2 + \beta^2 + \gamma^2 = 1$, and α and β are no longer interpretable as direction cosines. These components correspond to evanescent waves, which are rapidly attenuated and cannot propagate farther than a few wavelengths from the DOE. As a consequence, a *circ* function is usually added into the simulation of the ASM [11] so that there is no contribution of the evanescent waves to the field at the observation plane.

The main disadvantage of this method is that the sampling distances at the DOE plane and the output plane are the same $\delta_2 = \delta_1$, which imposes constraints on the simulation. In addition, to adequately sample the transfer function $G(f_x, f_y; z)$, its phase term must satisfy an additional sampling condition [79, 80]. Expressed for one-dimension in the frequency domain, this condition is:

$$\delta_f \left| \frac{\partial \phi_G}{\partial f} \right|_{max} \leq \pi, \quad \text{where } \phi_G(f) = kz\sqrt{1 - \lambda^2 f^2} \tag{2.26}$$

Replacing $|f|_{max} = N\delta_f/2$ and $\delta_f = 1/(N\delta_1)$ results in the validity region

$$z \leq \frac{N\delta_1^2}{\lambda}\sqrt{1 - \frac{\lambda^2}{4\delta_1^2}}, \quad \text{if } \delta_1 \geq \lambda/2 \tag{2.27}$$

which means that this method can only be used without numerical artifact for short propagation distances [81]. For example, with typical values for current DOEs, such as $\lambda = 1.55$ μm, $\delta_1 = 1$ μm and $N = 4096$, the validity distance is $z \leq 1.67$ mm. Notice that for DOEs having feature sizes smaller than half the wavelength ($\delta_1 \leq \lambda/2$), the ASM should not be used because the sampling condition is invalid.

From the above discussions, we can see that the maximum diffraction angle, which can be obtained by inverse cosine of the minimum and real γ, is independent of the number of samples. This means increasing N will only result in finer spacing and a greater validity distance, but not in higher diffraction angle. To model wider angles of diffraction, we can instead decrease the sampling distance in the source plane δ_1, but the validity distance will also be reduced.

Notice that, in case all components of the electromagnetic field are required (e.g. when polarization or Poynting vector have to be taken into account), they can be solved separately using the same scalar ASM equation (see Section 2.3.1). This is the basic idea for the vectorial ASM [82–84], which is equivalent to using the scalar ASM to calculate one component of the electromagnetic field and taking the Maxwell's partial derivatives to get the other components [64]. Another equivalence is the RSM [60],

which is a vectorial modal method originally used for the analysis of integrated optics structures. The drawback of these approaches is the sampling condition $\delta_2 = \delta_1$, which is the same as the scalar ASM. For this reason, these methods have generally been limited to the modeling of near-field focusing [63, 85, 86].

2.3.3 The Rayleigh-Sommerfeld diffraction formula

Unlike the ASM which is a solution of the Helmholtz equation, the Rayleigh-Sommerfeld (RS) diffraction formula is a mathematical model of the Huygens-Fresnel principle, where the observed field $U(x_2, y_2; z)$ is considered as a superposition of diverging spherical waves originating from every secondary point source within the DOE $U(x_1, y_1; 0)$. Fig. 2.5 shows the diffraction geometry in three dimensions. The Rayleigh-Sommerfeld diffraction formula is given in [11] as:

$$U(x_2, y_2; z) \propto \iint_{-\infty}^{\infty} U(x_1, y_1, 0) \left(jk - \frac{1}{r} \right) \frac{\exp(jkr)}{r^2} \, \mathrm{d}x_1 \mathrm{d}y_1 \tag{2.28}$$

where r is the distance between an arbitrary point in the source field plane to an arbitrary point in the observation plane

$$r = \sqrt{z^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{2.29}$$

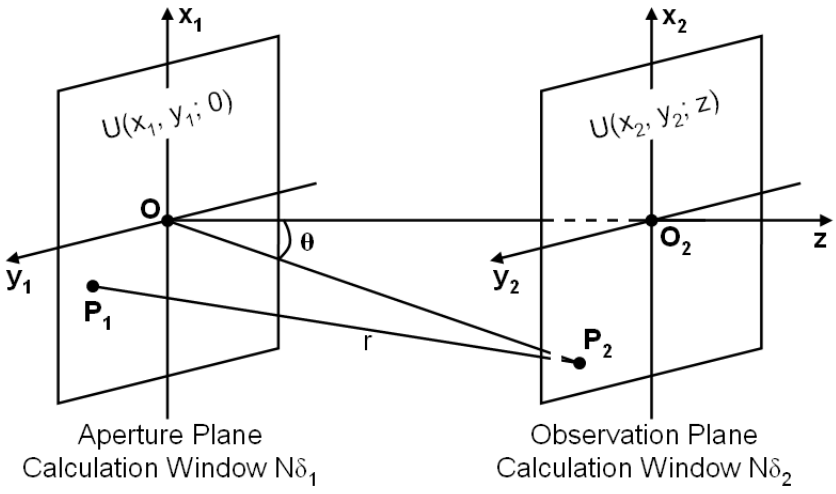


Figure 2.5 — Diffraction geometry.

Despite the apparent differences of their approaches, the ASM and the RS solution yield identical predictions of diffracted fields [87]. Eq. (2.28) can be seen as a convolution of the source-plane field $U(x_1, y_1; 0)$ with an Impulse Response

$$g(x, y; z) = \left(jk - \frac{1}{r} \right) \frac{\exp(jkr)}{r^2} \tag{2.30}$$

where the ASM (2.24) is shown to be the FT representation of the convolution, in which $A(f_x, f_y; 0)$ and the Transfer Function $G(f_x, f_y; z)$ are the Fourier transforms of $U(x_1, y_1; 0)$ and $g(x, y; z)$, respectively.

Similar to the ASM, to adequately sample the impulse response $g(x, y; z)$, its phase term must satisfy a sampling condition [79, 80]. Expressed for one-dimension in the source plane, this condition is:

$$\delta_x \left| \frac{\partial \phi_g}{\partial x} \right|_{max} \leq \pi \tag{2.31}$$

Assuming that $r \gg \lambda$, $g(x; z) \approx jk \exp(jkr)/r^2$ and the phase term can be approximated as $\phi_g = kr + \pi/2$, where $r = \sqrt{z^2 + x^2}$. Replacing $|x|_{max} = N\delta_1/2$ and $\delta_x = \delta_1$ results in the validity region:

$$z \geq \frac{N\delta_1^2}{\lambda} \sqrt{1 - \frac{\lambda^2}{4\delta_1^2}} \quad \text{if } \delta_1 > \lambda/2 \tag{2.32}$$

which means that the RS diffraction formula can be used for the region where the ASM is invalid (see Eq. (2.27)). Notice that if $\delta_1 \leq \lambda/2$, the RS diffraction formula is valid for all values of $z > 0$. Within this validity region, the RS diffraction formula can be evaluated numerically in different ways as follows:

Rayleigh-Sommerfeld direct integral

For a point on the observation plane, the Rayleigh-Sommerfeld integral (2.28) can be evaluated numerically as a Riemann sum:

$$U(x_2|_m, y_2|_n; z) \propto \sum_{i=1}^{N} \sum_{k=1}^{N} U(x_1|_i, y_1|_k; 0) g(x_2|_m - x_1|_i, y_2|_n - y_1|_k; z) \delta_1^2 \tag{2.33}$$

where δ_1 is the sampling interval on the source field plane and m, n are the indices of the observation point $(m, n = 1, \ldots, N)$. The advantage of this simulation is that the sampling distance on the observation plane δ_2 can be adjusted freely. Therefore, we can increase N or δ_2 to model wider angles of diffraction. However, the main drawback of this approach is its high complexity $O(N^4)$, which requires a long computation time, especially for the high resolution/large DOEs that can now be fabricated.

Rayleigh-Sommerfeld convolution

Using the convolution theorem, the Rayleigh-Sommerfeld diffraction formula (2.28) can also be evaluated numerically using the Impulse Response approach:

$$U(x_2, y_2; z) \propto IFT \left\{ FT \left[U(x_1, y_1; 0) \right] FT \left[g(x_1, y_1; z) \right] \right\} \tag{2.34}$$

The computational complexity is $O(N^2(1+3\log_2 N))$, which is much lower than the direct integral approach in Eq. (2.33). However, the disadvantage is that, as for the ASM approach, $\delta_2 = \delta_1$ and we have no control over the grid spacing in the observation plane. This condition limits application to either small diffraction angles or short propagation distances, because the only way to obtain non-paraxial diffraction at the far field is to increase the number of samples, which is not always possible due to the limited computer memory and calculation time. For example, with typical values of current DOEs such as $\delta_1 = \delta_2 = 1$ μm and $N = 4096$, to model up to an angle of 45°, the maximum propagation distance that can be evaluated is specified by:

$$\tan 45^\circ = \frac{O_2 P_{2max}}{z_{max}} \Rightarrow z_{max} = \frac{N\delta_2\sqrt{2}}{2} \approx 2.9 \text{ mm} \tag{2.35}$$

where O_2 and P_{2max} are the origin and the maximum angular extent of the observation plane, respectively (see Fig. 2.5).

Because of these difficulties (i.e. strong sampling constraint and high computational complexity), some approximations are often made to the ASM or RS diffraction formula, to reduce these rigorous free-space propagators to less accurate but more tractable expressions.

2.3.4 The scalar paraxial approximation

The scalar paraxial approximation, as its name implies, is valid for small diffraction angles and has been shown to be relatively accurate and very useful in practice [88]. Depending on whether the approximation is made to the ASM or RS diffraction formula, different versions of the paraxial approximation can be derived, with different sampling requirements and validity regions.

Paraxial Angular Spectrum Method

Using the binomial expansion for the direction cosines $\alpha, \beta \approx 0$ ($\alpha = \lambda f_x, \beta = \lambda f_y$)

$$\gamma = \sqrt{1-\alpha^2-\beta^2} \approx 1 - \frac{\alpha^2+\beta^2}{2} \tag{2.36}$$

the angular spectrum (2.24) can be rewritten as:

$$U(x_2, y_2; z) \propto \iint_{-\infty}^{\infty} A(f_x, f_y; 0) H(f_x, f_y; z) \exp\left[j2\pi\left(f_x x_2 + f_y y_2\right)\right] \mathrm{d}f_x \mathrm{d}f_y \tag{2.37}$$

where $H(f_x, f_y; z) = \exp\left[-j\pi\lambda z\left(f_x^2 + f_y^2\right)\right]$ is the paraxial Transfer Function. This equation is only valid in the paraxial approximation and should not be confused with the non-paraxial one. In simulation, this equation can be evaluated as

$$U(x_2, y_2; z) \propto IFT\left\{A(f_x, f_y; 0) H(f_x, f_y; z)\right\} \tag{2.38}$$

where $A(f_x, f_y; 0)$ is the FT of $U(x_1, y_1; 0)$. The computational complexity is the same as that of the ASM: $O(N^2(1+2\log_2 N))$. The sampling distance is also the same ($\delta_2 = \delta_1$) and we have no control over the grid spacing in the observation plane. Furthermore, Voelz *et al.* [80] showed that the phase term of H must satisfy an additional sampling condition. Expressed for one-dimension in the frequency domain, this condition is:

$$\delta_f \left| \frac{\partial \phi_H}{\partial f} \right|_{max} \leq \pi, \quad \text{where } \phi_H(f) = -\pi\lambda z f^2 \tag{2.39}$$

Replacing $|f|_{max} = N\delta_f/2$ and $\delta_f = 1/(N\delta_1)$ results in the validity region $z \leq N\delta_1^2/\lambda$, which means that this method can only be used for short propagation distances.

Although the paraxial ASM has no advantages over the non-paraxial one, it is useful for evaluating the accuracy of the paraxial approximation. It should be noticed that this validity region can also be derived directly from Eq. (2.27) for the non-paraxial ASM, using the assumption that $\lambda^2/(4\delta_1^2) \ll 1$, or equivalently $\delta_1 \gg \lambda/2$. An alternative, less stringent condition is $\lambda^2/(4\delta_1^2) < 0.1$, leading to $\delta_1 > 1.58\lambda$, where the inequality is now $>$ rather than $\gg$. For example, if λ=0.633 μm and δ_1=1 μm, $\sqrt{1-\lambda^2/(4\delta_1^2)} = 0.949$, resulting in an error in the assumption of about 5%, which will be further demonstrated in the next chapter. This condition suggests that the paraxial approximation is relatively accurate (within a few percent error) for the free-space propagation if the DOE feature size is larger than about 2 times the illumination wavelength.

For paraxial diffraction in longer propagation distances, several approximations for the distance r can be made to the RS equation (2.28), resulting in the Fresnel diffraction formula.

Fresnel approximation

Assuming that the propagation distance is much larger than the wavelength $r \gg \lambda$, the RS equation can be rewritten:

$$U(x_2, y_2; z) \propto \iint_{-\infty}^{\infty} U(x_1, y_1, 0) \frac{\exp(jkr)}{r^2} \mathrm{d}x_1 \mathrm{d}y_1 \tag{2.40}$$

As the accuracy of r in the denominator is generally much less significant than in the exponent, $r \approx z$ can be used for the denominator, while the binomial expansion of (2.29) is used for the exponent term:

$$r \approx z + \frac{(x_2 - x_1)^2 + (y_2 - y_1)^2}{2z} \tag{2.41}$$

Equation (2.40) can be rewritten as

$$U(x_2, y_2; z) \propto \iint_{-\infty}^{\infty} U(x_1, y_1; 0) \exp\left\{ j\frac{k}{2z}\left[(x_2 - x_1)^2 + (y_2 - y_1)^2\right] \right\} \mathrm{d}x_1 \mathrm{d}y_1 \tag{2.42}$$

which is a convolution of the source-plane field $U(x_1, y_1; 0)$ with an Impulse Response $h(x, y; z) = \exp\left[j\frac{k}{2z}(x^2 + y^2)\right]$. Extracting the constant phase term out of the integrals, we obtain the Fresnel diffraction formula for the paraxial approximation:

$$U(x_2, y_2; z) \propto \iint_{-\infty}^{\infty} \{U(x_1, y_1; 0)h(x_1, y_1; z)\} \exp\left[-j\frac{2\pi}{\lambda z}(x_1x_2 + y_1y_2)\right] \mathrm{d}x_1\mathrm{d}y_1 \tag{2.43}$$

In simulation, the Fresnel equation (2.43) can be evaluated directly as:

$$U(x_2, y_2; z) \propto FT\{U(x_1, y_1; 0)h(x_1, y_1; z)\} \tag{2.44}$$

with the computational complexity of $O(N^2(1 + \log_2 N))$, which is much lower than that of the RS direct integral (2.33). The spatial frequencies are specified by $f_x = x_2/(\lambda z)$, $f_y = y_2/(\lambda z)$, or equivalently $\delta f = \delta_2/(\lambda z)$. Due to the discrete FT property: $\delta f = 1/(N\delta_1)$, the grid spacing in the observation plane can be obtained as:

$$\delta_2 = \lambda z \delta_f = \frac{\lambda z}{N\delta_1} \tag{2.45}$$

which is more flexible than that of the ASM or RS convolution (2.34).

According to [80] or equivalently [89], the sampling criterion for Fresnel propagation is that the phase term of the h function satisfies the condition in one direction in the source plane:

$$\delta_x \left|\frac{\partial \phi_h}{\partial x}\right|_{max} \leq \pi, \quad \text{where } \phi_h = \frac{kx^2}{2z} \tag{2.46}$$

Replacing $\delta_x = \delta_1$, $|x|_{max} = N\delta_1/2$ and $k = 2\pi/\lambda$ results in the validity region $z \geq N\delta_1^2/\lambda$, which can also be approximated from Eq. 2.32 for the RS diffraction formula. The accuracy of the Fresnel propagation is therefore similar to that of the paraxial ASM.

Notice that we could also evaluate the Fresnel convolution (2.42) directly as:

$$U(x_2, y_2; z) \propto IFT\{FT[U(x_1, y_1; 0)]FT[h(x_1, y_2; z)]\} \tag{2.47}$$

which is equivalent to the paraxial ASM (2.38) where $A(f_x, f_y; 0)$ and the Transfer Function $H(f_x, f_y; z)$ are the Fourier transforms of $U(x_1, y_1; 0)$ and $h(x, y; z)$, respectively. However, this approach has the same validity distance as the numerical evaluation (2.44), with a fixed grid spacing $\delta_2 = \delta_1$ and a higher computational complexity (three Fourier transforms instead of one). This simulation is therefore rarely used in practice.

If the diffracted field is observed on a plane very far from the DOE plane, the scalar paraxial approximation can be further reduced:

2.3.5 The scalar paraxial approximation in the far field

Fraunhofer approximation

If another assumption is used additionally ($z = \infty$), so that the quadratic phase factor $\exp\left[j\frac{k}{2z}\left(x_1^2 + y_1^2\right)\right]$ in equation (2.43) is approximately unity over the entire aperture, the field at the observation plane can be found directly as a FT of the aperture distribution itself:

$$U\left(x_2, y_2; z\right) \propto \iint_{-\infty}^{\infty} U\left(x_1, y_1; 0\right) \exp\left[-j\frac{k}{z}\left(x_2 x_1 + y_2 y_1\right)\right] \mathrm{d}x_1 \mathrm{d}y_1 \tag{2.48}$$

This is the Fraunhofer diffraction, which is just a special case of the Fresnel diffraction where the propagation distance is at the far-field [11]

$$z > Z_F = \frac{2D^2}{\lambda} \tag{2.49}$$

with D is the diameter of the illumination area. This criterion is known as Fraunhofer distance in antenna theory [90], which can be written in terms of a tolerance on defocus smaller than $\lambda/16$ [91].

For example, at a wavelength of 633 nm and a DOE illumination beam of 4 mm, the observation distance z should strictly be greater than 50 m, which is rare in nearly all practical systems using DOEs. However, the Fraunhofer diffraction pattern can be observed at distances much closer provided the DOE is illuminated by a spherical wave converging toward the observer, or if a positive lens is properly situated right before or after the DOE. These setups aim to bring the image from infinity to the focal point of the converging beam or lens [11, 92].

In simulation, the Fraunhofer diffraction formula can be evaluated as a FT, with the computational complexity of $O(N^2 \log_2 N)$. As it is a special case of the Fresnel approximation, the spacing in the observation plane must obey the same spatial mapping: $\delta_2 = \lambda z/(N\delta_1)$. It is also possible to calculate diffraction patterns in the far-field by using the Fresnel approximation to retain the full accuracy of the paraxial regime [11], at the expense of a slightly higher computational complexity.

Angular Fraunhofer approximation

A different version of the paraxial far-field approximation has been used in [93], where the diffraction pattern is calculated in angular instead of spatial coordinates:

$$U\left(\alpha, \beta; z\right) \propto \iint_{-\infty}^{\infty} U\left(x_1, y_1; 0\right) \exp\left[-j\frac{2\pi}{\lambda}\left(\alpha x_1 + \beta y_1\right)\right] \mathrm{d}x_1 \mathrm{d}y_1 \tag{2.50}$$

This approximation can be simulated by a FT, where the angular coordinates α and β are equally spaced. However in practice, the required output field is usually in spatial

coordinates on a uniform sampling grid with the sampling interval δ_2. A resampling step is therefore necessary to interpolate the field from angular $U(\alpha, \beta; R')$ to spatial coordinates $U(x_2, y_2; z)$ on a uniform sampling grid with the sampling interval δ_2. For example with nearest-neighbor interpolation, which is the simplest method [94], the field in the output plane $U(x_2, y_2; z)$ at the position $(x_2 = m\delta_2, y_2 = n\delta_2)$ is calculated by:

$$U(m\delta_2, n\delta_2; z) \approx U(p\delta_\alpha, q\delta_\beta; R') \tag{2.51}$$

in which m, n, p, q are integers $-N/2 \leq m, n, p, q \leq N/2 - 1$, with $p\delta_\alpha \approx x_2/R', q\delta_\beta \approx y_2/R'$. The sampling interval of the direction cosines is specified by $\delta_\alpha = \delta_\beta = \lambda\delta_f$, where $\delta_f = 1/(N\delta_1)$. The computational complexity of the algorithm is therefore $O(N^2(1 + \log_2 N))$. Although there is no approximation for the spatial distribution of the diffracted field in this method, a paraxial approximation has been used for the amplitude of the field, as will be shown in Chapter 3. Examples of far-field diffraction patterns and numerical artifacts can be found in [11] and [95], respectively.

From the above discussion, we can see that there seems to be no ideal solution for the calculation of far-field non-paraxial diffraction patterns. For this reason, researchers have been searching for different ways of modeling scalar non-paraxial far-field diffraction [96–98]. A potential way to get around this problem is the repeated calculation method, where the above propagators can be used for different calculation windows or consecutive propagation steps.

2.3.6 Repeated propagation method

Multi-window propagation

For diverging elements where the Fourier pattern is much bigger than the DOE size, the output plane can be divided into small zones of the same size as the input plane, and an off-axis propagation method can be repeated over and over at different points to calculate the entire diffraction pattern. For example, Shen *et al.* [99] developed a modified RS convolution which allows calculation of output planes that are not necessarily centered around the optical axis, but the sizes of the calculation windows at the source and destination plane remain the same. This leads to the idea of calculating a repeated RS convolution when the diffraction pattern is bigger than the DOE size, as illustrated in Fig. 2.3.6. However, with this approach, to obtain the same size output plane at the same sampling points as the standard RS integral, the output plane must be calculated at an unnecessarily large number of points, leading to a huge increase in computational effort. If each zone contains one sampling point on the output plane, the convolution has to be repeated $N \times N$ times to obtain the entire diffraction pattern. As the computational complexity of the modified convolution [99] is $O(N_F^2(2+3\log_2 N_F))$,

where $N_F = 2N - 1$, the computational complexity of the repeated convolution is $O(N^2 N_F^2(2 + 3\log_2 N_F))$, which is even higher than that of the RS integral.

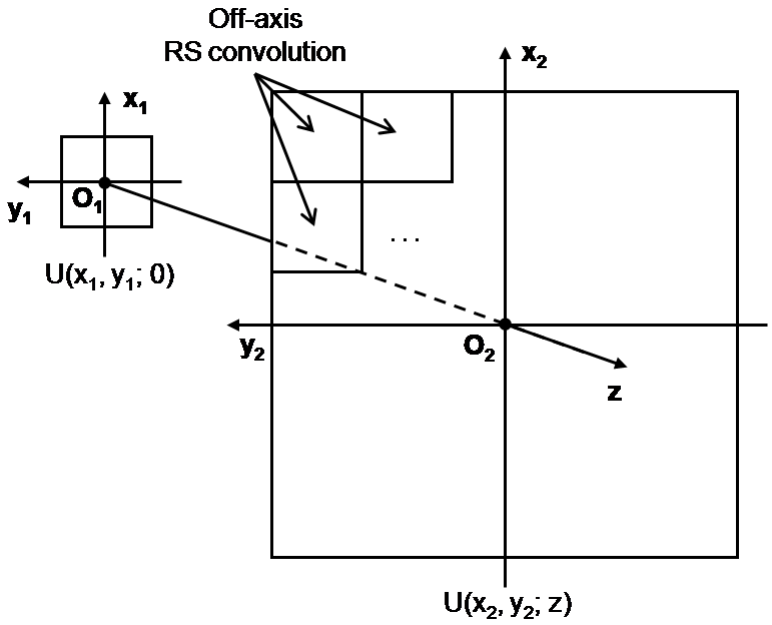


Figure 2.6 — Principle of a repeated calculation based on off-axis RS convolution.

Multi-step propagation

The diffracted field on the source plane can also be propagated to the output plane via multiple intermediate planes using any of the above propagators. For example, a two-step propagation has been used for the Fresnel propagator [100], as illustrated in Fig. 2.7. It is also possible to derive a scaled paraxial ASM where the field is propagated via an intermiate plane before propagating to the observation plane [101]. The main advantage of these scaled scalar paraxial propagators over the original versions is that the sampling distance in the observation plane can be adjusted via the scaling parameter $\delta_2 = m\delta_1$, where $m = (z - z_1)/z_1$ can be changed by choosing the intermediate plane. Attempts have been made to develop scaled versions of the non-paraxial ASM [102, 103] but they are only applicable to high numerical converging elements, which basically means only the near field diffraction pattern. The disadvantage of this approach is the increase in computational complexity and the introduction of other errors due to the cascaded aperture effect [81]. Details on the mathematical derivation of the scaled scalar paraxial propagators and a demonstration for the cascaded aperture effect are given in Appendix A. Once the diffracted field at the observation plane has been calculated, the diffracted power can be estimated in different ways as given in Appendix B.

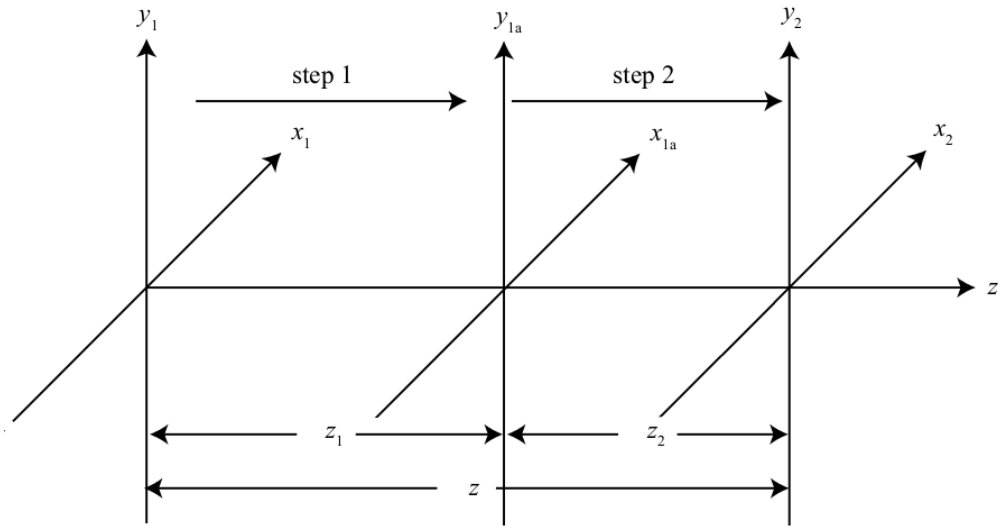


Figure 2.7 — Two-step propagation geometry.

2.3.7 Comparison of different free-space propagation methods

In the previous sections, we have analyzed the numerous free-space propagation methods, whose simulation constraints and validity regions are summerized in Table 2.1. From the above discussion, we can see that there seems to be no ideal solution for the calculation of the very useful practical case of far-field non-paraxial (wide angles) diffraction patterns. While the direct RS integral imposes no constraints, its computational complexity limits its application in many situations. The RS convolution and the ASM greatly reduce the computational complexity, but they impose sampling constraints that limit either the diffraction angle or propagation distance. A multistep RS convolution or ASM where a long distance is split into a sequence of shorter propagations [95,104] offers a potential way to get around this problem, but there are increases in computational complexity and the introduction of other errors due to the cascaded aperture effect [81]. Another approach is a repeated RS convolution or ASM propagation where the output field is divided into small zones of the same size with the DOE plane at the expense of increased computational effort. Due to these reasons, some approximations are often made to the ASM or RS diffraction formula, which are rigorous for the free-space propagation.

On the other hand, scalar paraxial methods are computationally efficient but cannot model correctly wide diffraction angles. The paraxial approximation is relatively accurate (within a few percent error) if the DOE feature size is larger than about 2 times the illumination wavelength, or equivalently diffraction angles smaller than about 10°. It has been reported by Harvey *et al.* [105] that in fact the scalar regime can evaluate, to a high degree of accuracy, non-paraxial diffraction effects of diffractive gratings on an output hemisphere in the far-field by using a simple FT. In a related approach, Cole

Table 2.1 — Simulation constraints and validity regions of different diffraction models for the free-space propagation.

Diffraction model	Sampling constraint	Computational complexity	Validity region	Remakrs
ASM	$\delta_2 = \delta_1$	$N^2(1 + 2\log_2 N)$	$z \leq Z_r$	Rigorous scalar, $Z_r = \frac{N\delta_1^2}{\lambda}\sqrt{1 - \frac{\lambda^2}{4\delta_1^2}}$
RS integral	δ_2 flexible	N^4	$z \geq Z_r$	
RS convol.	$\delta_2 = \delta_1$	$N^2(2 + 3\log_2 N)$		
Repeated RS	$\delta_2 = \delta_1$	$4N^4(2 + 3\log_2 2N)$		
Parax. ASM	$\delta_2 = \delta_1$	$N^2(1 + 2\log_2 N)$	$z \leq Z_p$	Scalar paraxial, $Z_p = N\delta_1^2/\lambda$
Fresnel	$\delta_2 = \lambda z/(N\delta_1)$	$N^2(1 + \log_2 N)$	$z \geq Z_p$	
Scaled Fresnel	$\delta_2 = m\delta_1$	$3N^2(1 + \log_2 N)$	$z \geq sZ_p$	Scalar paraxial, $s = \max(m, \lvert 1 - m\rvert)$
Scaled parax. ASM	$\delta_2 = m\delta_1$	$2N^2(1 + \log_2 N)$	$z \leq mZ_p$, $z \geq \lvert 1 - m\rvert Z_p$	Scalar paraxial, $m \geq 1/2$
Fraunhofer	$\delta_2 = \lambda z/(N\delta_1)$	$N^2 \log_2 N$	$z \geq 2D^2/\lambda$	Scalar paraxial, far-field
Angular Fra.	Resampling	$N^2(1 + \log_2 N)$		

et al. [106] have studied non-paraxial diffraction effects in lens based imaging system, where the applicable range of scalar diffraction theory is extended for numerical apertures up to about 0.6, corresponding to diffraction angles up to about 37° in free space. However, the extension of these approaches to more complex diffractive structures and especially by the projection from the hemisphere to the more usual observation plane has not been studied in detail.

2.4 Conclusion

In conclusion, there is no ideal solution for the calculation of far-field non-paraxial diffraction patterns, as most of the current diffraction models are either fast but inaccurate or rigorous but slow. Vectorial methods are rigorous but can only be used for a very small DOE, usually around the DOE region, due to the extensive calculation required. For the free-space propagation region, the ASM and the RS diffraction formula model the diffraction problem accurately but their applications are often limited due to the imposed sampling constraints and computational complexity. On the other hand, the TEA and the Fraunhofer approximation offer a fast way to calculate the diffracted field at the DOE plane and far-field output plane, respectively. These approximations have been used for the modeling and design of Fourier elements in most cases, but they are only applicable to thin elements and small diffraction angles, respectively. We will now present in Chapter 3 the efficient diffraction propagators we have developed to overcome these limitations and enable the rapid calculation of non-paraxial Fourier

elements. The limit of the TEA and our methods to overcome this limitation in the modeling and design of thick elements will be discussed later in Chapter 4.

CHAPTER 3

Scalar theory for the modeling and design of DOEs operating in the non-paraxial far-field diffraction regime

In this chapter, we aim to extend the applicable range of scalar diffraction theory for general (not limited to lens or one-dimensional grating) far-field elements by presenting a projection step in combination with the Harvey model [105] to estimate the non-paraxial diffraction pattern of a DOE at an observation plane in the far-field. The computational results for a sample DOE are compared with the results obtained by the Rayleigh-Sommerfeld integral [11], which is more rigorous but much more time consuming. It is shown that our proposed method is more accurate than the traditional Fraunhofer diffraction pattern [11], avoids the complex sampling constraints encountered with Angular Spectrum based approaches and gives results that are very close to the experimentally observed non-paraxial patterns, while the computation effort remains very close to that of FT based methods. The results in [107] confirm that, scalar theory can be applied to model DOEs for diffraction angle up to about 33° and perhaps even higher [106] with some decrease in calculation accuracy, provided that the diffractive structures are not so small that vector diffraction theory is required.

3.1 Proposed scalar non-paraxial method

Harvey *et al.* introduced a non-paraxial scalar treatment that is accurate for large diffraction angles [105, 108], where the complex valued diffracted wave field is observed on a hemisphere with radius R from the origin of the DOE plane, as shown in Fig. 3.1. In this model, all terms from the binomial expansion for the distance from the source

point to the observation point

$$r = \sqrt{(x-x_1)^2 + (y-y_1)^2 + z^2} = z\sqrt{1 + \left[\frac{x_1^2+y_1^2}{z^2} + \frac{x_2^2+y_2^2}{z^2} - \frac{2(x_1x_2+y_1y_2)}{z^2}\right]} \tag{3.1}$$

are retained. The Rayleigh-Sommerfeld diffraction formula (2.28) can be rewritten as ($r \gg \lambda$):

$$U(\alpha,\beta;R) \approx \frac{\gamma \exp(jkR)}{j\lambda R} \iint_{-\infty}^{\infty} U(x_1,y_1;0) \frac{\exp(jkW)}{(1+\epsilon)^2} \exp\left[-\frac{j2\pi}{\lambda}(\alpha x_1 + \beta y_1)\right] \mathrm{d}x_1 \mathrm{d}y_1 \tag{3.2}$$

with $\epsilon = (r-R)/R$ and W is a wave-front aberration function which contains all the higher orders of the expansion [105]. The direction cosines of the propagation vector are calculated from the spatial coordinates of the observation point (x, y, z):

$$\alpha = \frac{x}{R}, \quad \beta = \frac{y}{R}, \quad \gamma = \frac{z}{R}, \quad \text{with } R = \sqrt{x^2+y^2+z^2} \tag{3.3}$$

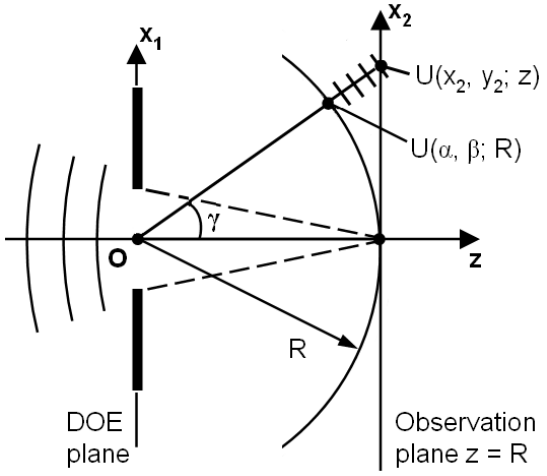


Figure 3.1 — Diffraction geometry of the Harvey model.

This approximation is derived using only the assumption that $r \gg \lambda$. Harvey [109] also showed that the aberration can be negligible ($W \approx 0$) when the illumination is a converging spherical wave, as is often the case with DOEs in practice or very nearly so when the illumination is a plane wave or unfiltered laser beam. Another approximation can be made with the assumption that the radius of the observation hemisphere is much larger than the width of the source field, so that $r \approx R$ and therefore $\epsilon \approx 0$. These assumptions are well satisfied for many practical applications where the DOE size is often less than 1 cm and the propagation distance is generally a few tens cm. The Fourier diffraction pattern can be observed at a finite distance ($z = R$) if a converging spherical wave is used for illumination, which is equivalent to the classical $1f$ setup,

where a focusing lens is placed right after the DOE [11]. These illumination and imaging conditions will be used for all simulation methods discussed in the next section.

In these conditions, the diffracted wave field on the hemisphere can be calculated directly by a single FT:

$$U(\alpha, \beta; R) \approx \frac{\gamma \exp(jkR)}{j\lambda R} FT\left\{U(x_1, y_1; 0)\right\} \tag{3.4}$$

However, the observation space in practice is usually a plane, which means that a projection of the field from the hemisphere to the plane is required.

3.1.1 Hemisphere-to-plane projection

An arbitrary position on the plane $z = R$ has the diffraction angles (α, β, γ) and the spatial coordinates specified by

$$R' = \frac{R}{\gamma}, \quad x_2 = \frac{R\alpha}{\gamma}, \quad y_2 = \frac{R\beta}{\gamma} \tag{3.5}$$

where R' is the distance from the origin of the DOE plane to the observation point in the output plane (x_2, y_2, z). The field at this position can be estimated by one-to-one mapping.

Plane wave projection

As a first approximation for this projection (Fig. 3.1), we assume a local plane wave travelling from every point $(\alpha, \beta; R)$ on the hemisphere to the corresponding position (x_2, y_2, z) on the tangential observation plane with the same magnitude and a phase delay due to an optical path length: $OPL = R' - R$. The field on the plane can be calculated directly from the field on the hemisphere as:

$$U(\alpha, \beta; z) = U(\alpha, \beta; R') = \exp[jk(R' - R)]U(\alpha, \beta; R) \tag{3.6}$$

where $U(\alpha, \beta; R')$ is the field in the output plane expressed in terms of angular coordinates.

In this way, each point on the sphere is mapped onto one point on the plane by considering theoretically straight rays, so for example, a regularly spaced array of diffraction spots on the output hemisphere becomes an irregularly spaced (distorted) array of spots on the observation plane. In reality light does not propagate like this - it will continue diverging/spreading from the hemisphere to the plane, as shown in Fig. 3.2(a). As the light field spreads out, it will cause not only distortion in spot position but also a reduction in the intensity of the off-axis spots when they reach the plane.

Spherical wave projection

A more accurate approximation would be to consider the field at each position on the hemisphere as a local spherical wave $\exp(jkr)/r$ originating from the DOE and travelling towards the observation plane. The field on the plane is therefore:

$$U(\alpha,\beta;R') = \frac{R}{\exp(jkR)}U(\alpha,\beta;R)\frac{\exp(jkR')}{R'} = \frac{R}{R'}\exp[jk(R'-R)]U(\alpha,\beta;R) \quad (3.7)$$

With this more precise approximation we take the diverging/spreading effect into account. This approach remains a one-to-one mapping, and the total power is conserved. Strictly, in reality, there is not only propagation of light from the hemisphere, but also superposition of light from different parts of the hemisphere at the same point on the plane, which means we have another diffraction step from the hemisphere to the tangential plane. In the model we present here, we consider and show that the spherical wave projection is sufficient for our requirements.

Substituting (3.4) in (3.7) results in the diffraction pattern:

$$U(\alpha,\beta;R') \approx \frac{\gamma\exp(jkR')}{j\lambda R'}FT\{U(x_1,y_1;0)\} \quad (3.8)$$

which is in fact equivalent to a multiradii approach, where every point on the observation plane is considered as belonging to different hemispheres with different radii R', as shown in Fig. 3.2(b), and the Harvey equation can be applied directly on the observation plane.

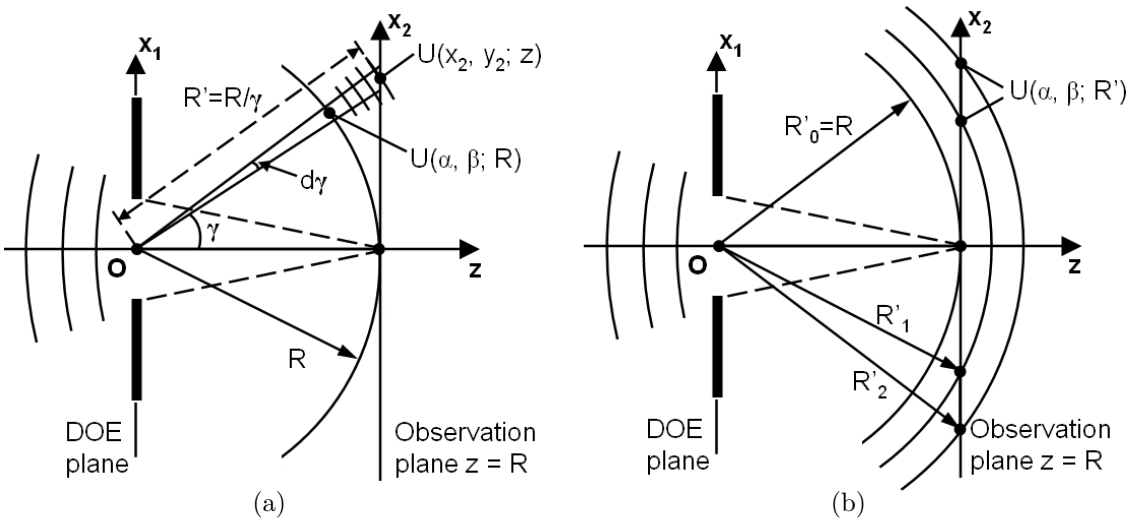


Figure 3.2 — (a) The diverging effect from the hemisphere to the observation plane. (b) Multiradii approach.

3.1.2 Resampling from angular to spatial coordinates

In all cases, in simulation the direction cosines α and β are equally spaced, and $\gamma = (1-\alpha^2-\beta^2)^{1/2}$ is a nonlinear function. The output plane spatial coordinates x_2, y_2 calculated by the spatial mapping (3.5) are therefore unequally spaced. A resampling step is therefore required to interpolate the field in the output plane $U(\alpha, \beta; R')$ to $U(x_2, y_2; z)$ on a uniform sampling grid with the sampling interval δ_2. For example with nearest-neighbor interpolation, which is the simplest method [94], the field in the output plane $U(x_2, y_2; z)$ at the position $(x_2 = m\delta_2, y_2 = n\delta_2)$ is calculated by:

$$U(m\delta_2, n\delta_2; z) \approx U(p\delta_\alpha, q\delta_\beta; R') \tag{3.9}$$

in which m, n, p, q are integers $-N/2 \leq m, n, p, q \leq N/2 - 1$, with $p\delta_\alpha \approx x_2/R', q\delta_\beta \approx y_2/R'$. The sampling interval of the direction cosines is specified by $\delta_\alpha = \delta_\beta = \lambda\delta_f$, where $\delta_f = 1/(N\delta_1)$. The computational complexity of the algorithm is therefore $O(N^2(1 + \log_2 N))$ plus the complexity of the interpolation step. We use a bilinear interpolation because of its low complexity, which is $O(4N^2)$, and show that its performance in preserving fine details of the output field is sufficient.

Note that there is no specific requirement on δ_2, as in the direct RS integral, while the diffraction angle can be widened up to 90° by decreasing the sampling interval in the source plane δ_1, similar to the ASM.

3.1.3 The diffraction position

Due to the diverging of the field from the hemisphere to the plane, the positions of the diffracted spots are distributed as a non-linear function [105]. The diffraction angle of the order $(\mathbf{m}, \mathbf{n})$ in a 2D diffraction pattern can be calculated as:

$$\tan\theta = \frac{\lambda}{\delta_1}\sqrt{\frac{(\mathbf{m}/N)^2 + (\mathbf{n}/N)^2}{1 - (\lambda/\delta_1)^2[(\mathbf{m}/N)^2 + (\mathbf{n}/N)^2]}} \tag{3.10}$$

For low diffraction orders $(\mathbf{m}, \mathbf{m} \ll N)$ or structures with pixel size much bigger than the operating wavelength $(\delta_1 \gg \lambda)$, where the diffraction angle is small, this equation can be simplified into the paraxial approximation:

$$\tan\theta \approx \frac{\lambda}{N\delta_1}\sqrt{\mathbf{m}^2 + \mathbf{n}^2} \tag{3.11}$$

3.1.4 The diffracted power

As a result of the projection from the hemisphere, the propagation direction of non-paraxial diffraction is no longer perpendicular to the observation plane. The total

power incident on a surface area A can be calculated using the scalar equation (B.9):

$$P = \iint_A p\gamma \, \mathrm{d}x_2 \mathrm{d}y_2 \tag{3.12}$$

where γ is the cosine of the angle between the direction of power flow $\vec{k}$ and the normal of the surface $\vec{n}$. When $\vec{k}$ is nearly normal to the surface, $\gamma \approx 1$ and the total power P is simply the integral of the power density p over the detector area, which is a paraxial approximation. Mathematical derivation of the scalar power estimation from the rigorous Poynting vector can be found in Appendix B.

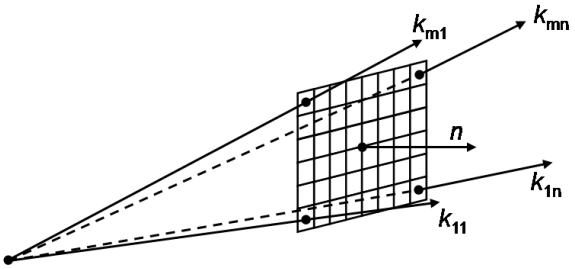


Figure 3.3 — Diffraction power estimation.

In simulation, the spot power is then estimated by the sum of the total light over the spot area:

$$P \propto \sum_{\text{spot area}} |U|^2 \cos(\vec{k}_{ij}, \vec{n}) \delta_2^2 \tag{3.13}$$

where the spot area is defined as a square of arbitrary but constant size for each spot and centered on the spot, as shown in Fig. 3.3. This calculation technique is based on what happens experimentally in practice where the active area of a photodetector is centered on the spot.

Paraxial diffraction in the far field can be seen as a special case of our proposed method, in which $\gamma \approx 1, R' \approx R = z$ and Eq. (3.8) then simplifies into the angular Fraunhofer approximation:

$$U(\alpha, \beta; z) \approx U(\alpha, \beta; R') \approx \frac{\exp(jkz)}{j\lambda z} FT\left\{U(x_1, y_1; 0)\right\} \tag{3.14}$$

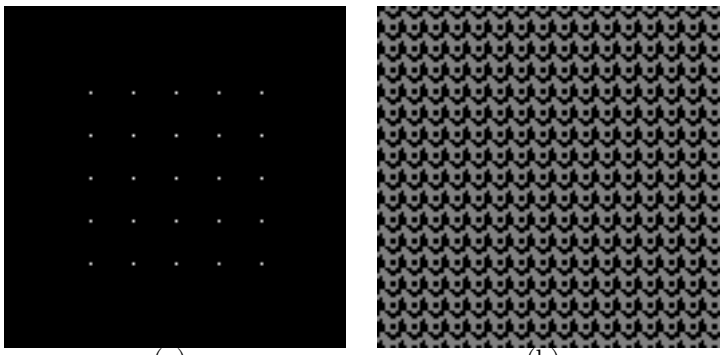
Similar to the multiradii approach, a nonlinear mapping such as Eq. 3.9 can be used to resample the field from angular to spatial coordinates. On the other hand, as $\gamma \approx 1$, the observation plane sampling grid (3.5) is approximately uniform (i.e. linear mapping), and we obtain the Fraunhofer approximation in spatial coordinates [100, 110]:

$$U(x_2, y_2; z) \approx \frac{\exp(jkz)}{j\lambda z} FT\left\{U(x_1, y_1; 0)\right\}, \text{where } x_2 \approx R\alpha \Rightarrow \delta_2 \approx R\delta_\alpha = \frac{z\lambda}{N\delta_1} \tag{3.15}$$

with the power evaluation (3.13) simplified into the calculation of intensity as the squared magnitude of the diffracted field [11].

3.1.5 Computer simulations and discussions

To assess the accuracy of our projections and to compare our proposed model with the scalar methods reviewed in Chapter 2, a binary phase Fourier DOE was designed using a standard multi-stage IFTA [6,51]. The target image was a 5×5 spot array with equal spot intensity and spacing and the designed DOE had a smallest pixel size of 1 μm, as shown in Fig. 3.4. A converging spherical wave at the wavelength of 1.55 μm and a circular aperture were used to generate the far-field diffraction pattern of the replicated DOE at the propagation distance of 18 cm. These parameters were chosen to be compatible with our DOE fabrication facilities for experimental testing. The output field was simulated using different methods, and the RS integral result was used as the reference for our analysis, since no approximation is used and no specific sampling requirement is imposed in this simulation, which allows for a rigorous scalar calculation of the entire output field.



(a) (b)

Figure 3.4 — (a) Target image, which is a 256×256 pixel image at diffraction distance of 18 cm, spot separation of about 3.5 cm. (b) Fourier DOE designed by an IFTA, where a period is 8 μm.

Hemisphere-to-plane projection

In order to see the difference between the diffracted fields calculated from different simulation models, we use the ratio of difference, which is defined similarly to the visibility function [111]:

$$\text{Ratio of difference} = \frac{|U| - |U_{RS}|}{|U| + |U_{RS}|} \tag{3.16}$$

where U is the simulated diffraction using (3.14), (3.6) or (3.7). The advantage of this metric over the absolute difference and the square error is that the difference is normalized and therefore an angular dependent effect can be seen.

Fig. 3.5 shows the ratio of difference between the output field calculated using the Fraunhofer approximation, our proposed projections and the RS integral. In Fig.

3.5(a) for the angular Fraunhofer approximation, we can clearly see different rings corresponding to the angular dependent difference between the calculated field and the RS integral simulation result. This is as expected, due to the paraxial assumptions used in the approximation. The root-mean-square (RMS) of the ratio of difference is 0.345. The Harvey method plus plane wave projection reduces the ratio of difference, where the RMS is 0.199. However, different rings are still clearly visible, as shown in Fig. 3.5(b). The reason is due to the assumption that the field at each position on the hemisphere is a local plane wave travelling towards the observation plane without taking into account the diverging effect. As the light field spreads out, this will cause not only distortion but also a reduction in the intensity of the off-axis spots. The higher the diffraction angle, or the longer the propagation distance from the hemisphere to the plane, the more distorted is the output field.

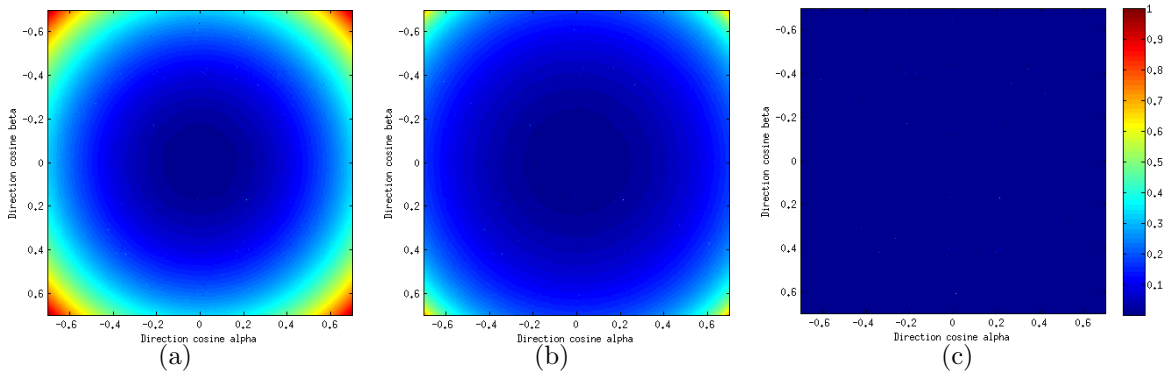


Figure 3.5 — Ratio of difference to the RS calculated diffraction pattern for the observation plane diffraction patterns calculated using: (a) Angular Fraunhofer approximation. (b) Harvey method plus plane wave projection. (c) Harvey method plus spherical wave projection.

On the other hand for our spherical wave projection, where the diverging/spreading effect has been taken into account, Fig. 3.5(c) shows the ratio of difference at almost a constant value over the calculation window, with the RMS of 0.02. Since the projection is still a one-to-one mapping, i.e. the superposition effect on the observation plane is not taken into account, some residual amount of error remains in addition to the aberrations of the Harvey propagation. However, if the propagation distance is much larger than the DOE size, which is the case in the majority of DOE applications, the magnitude of the error between the RS integral and our spherical wave projection is often small, as will be shown in the next section for the test DOE, where the error is about 1% for diffraction angle up to 33°. For higher diffraction angles, corresponding to finer detailed structures, the scalar transmission assumption of the DOE will break down and vector diffraction theory is necessary to take into account the effects of the electromagnetic fields inside the diffractive structure.

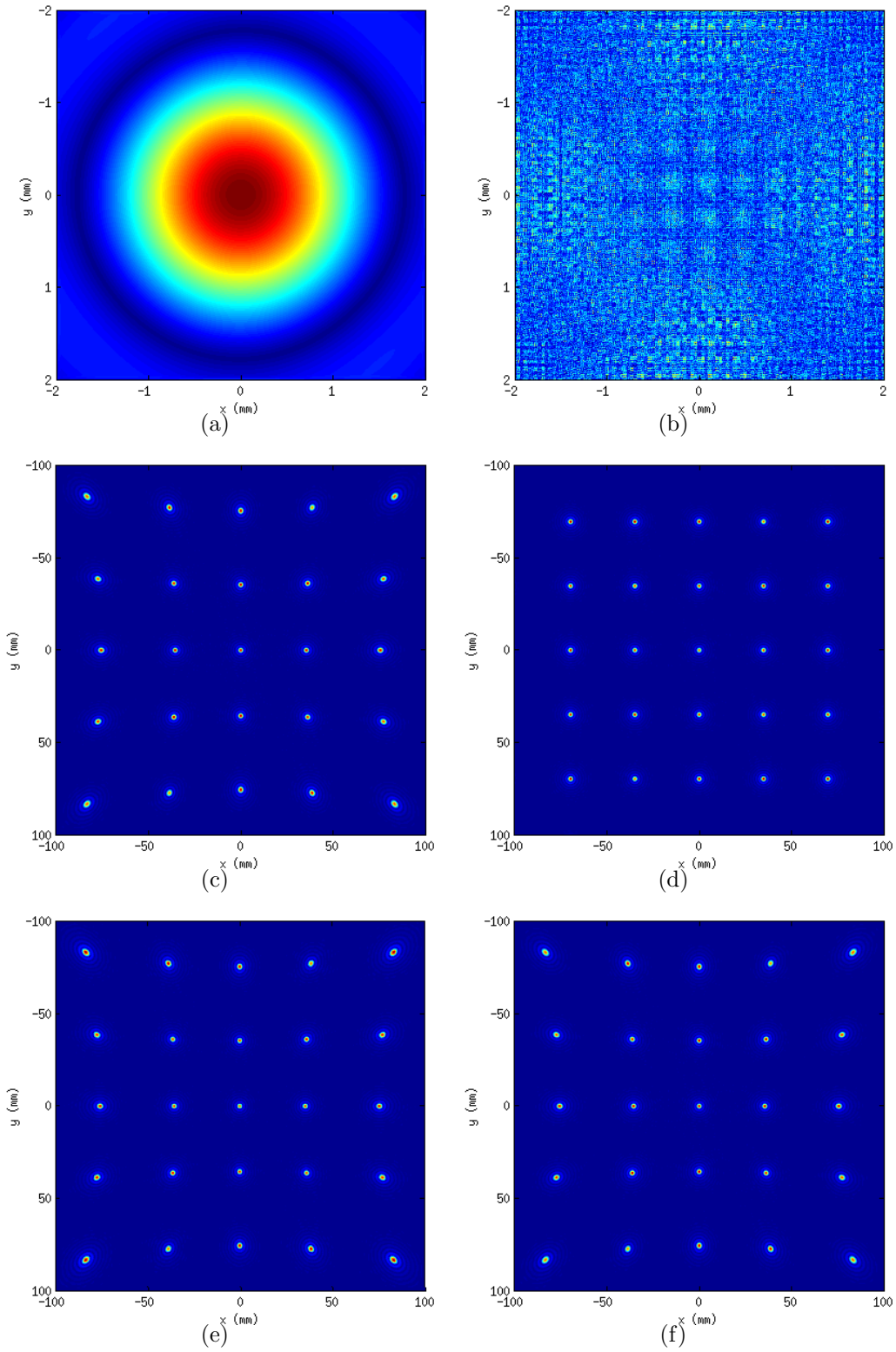


Figure 3.6 — Comparison between the simulated diffraction patterns of the test DOE obtained by using: (a) RS convolution, (b) ASM, (c) RS integral, (d) Fraunhofer approximation, (e) Angular Fraunhofer approximation with nonlinear mapping, (f) Multiradii approach. Due to the sampling constraint, the RS convolution and the ASM only obtain a small area around the optical axis.

Comparison with other methods

To compare our proposed method with the conventional propagation techniques, the far-field diffraction of the test DOE was reconstructed using the ASM (2.25), RS convolution (2.34), RS integral (2.33), Fraunhofer approximation (3.15), angular Fraunhofer approximation (3.14) and multiradii approach (3.8). The output plane sampling interval (3.15) of the Fraunhofer approximation was used for both the RS integral and our model for further analysis. The calculated diffraction patterns on the observation plane are illustrated in Fig. 3.6, with δ_1=1 μm and $N = 4096$ pixels. The zoomed images of the top left corner of the simulated diffraction patterns, except for the RS convolution and ASM, are shown in Fig. 3.7, respectively.

Note that our target image was an equally distributed 5×5 spot array covering an area of a few 10×10 cm^2. Due to the sampling constraint on the output plane ($\delta_2 = \delta_1$), the result of the RS convolution covers an area of only about 4×4 mm^2 around the central spot and is clearly unusable under such conditions, as shown in Fig. 3.6(a). Similarly in Fig. 3.6(b), the ASM models only the central spot area, but incorrectly because the sampling condition (2.27) for this algorithm at this propagation distance is not satisfied. The reference output pattern obtained by the RS integral is shown in Fig. 3.6(c), which is an unequally distributed 5×5 spot array. Fig. 3.6(d) shows the spot array calculated by the Fraunhofer approximation, which is very similar to the target image but incorrect in practice (see Fig. 5.4 for experimetal result), as this simple approach does not correctly model non-paraxial diffraction angles. The result of our proposed method is illustrated in Fig. 3.6(f), which is visually very similar in both spots distribution and spots shape to the pattern obtained by the reference RS integral. These effects are also predicted correctly by the angular Fraunhofer approximation with nonlinear mapping, as shown in Fig. 3.6(e), but this method over estimates the amplitude of the spots at non-paraxial angles, as will be shown later in Table 3.1.

More specifically, two main differences can be seen between the RS integral and the Fraunhofer approximation for non-paraxial diffraction angles. The first difference is in the spot positions, where a linear spatial distribution is obtained in the Fraunhofer pattern, but a nonlinear one is obtained using the RS integral. This behaviour can be seen at the outermost spots which are distributed on parabolic functions instead of the straight lines obtained in the Fraunhofer pattern, as shown in Fig. 3.6(c) and 3.6(d), respectively. The other difference is in spot power, which depends on the spot amplitude and area. The Fraunhofer approximation tends to under estimate the spots size at non-paraxial angles, where all spots are circular in shape of the same size, as shown in Fig. 3.7(b). As in the RS integral calculation, the spots predicted by our method are broader away from the axis due to the diverging effect of the field from the hemisphere to the plane, creating elliptical spots, as shown in 3.7(d). On the other hand, both the Fraunhofer approximations over estimate the amplitude of the spots at

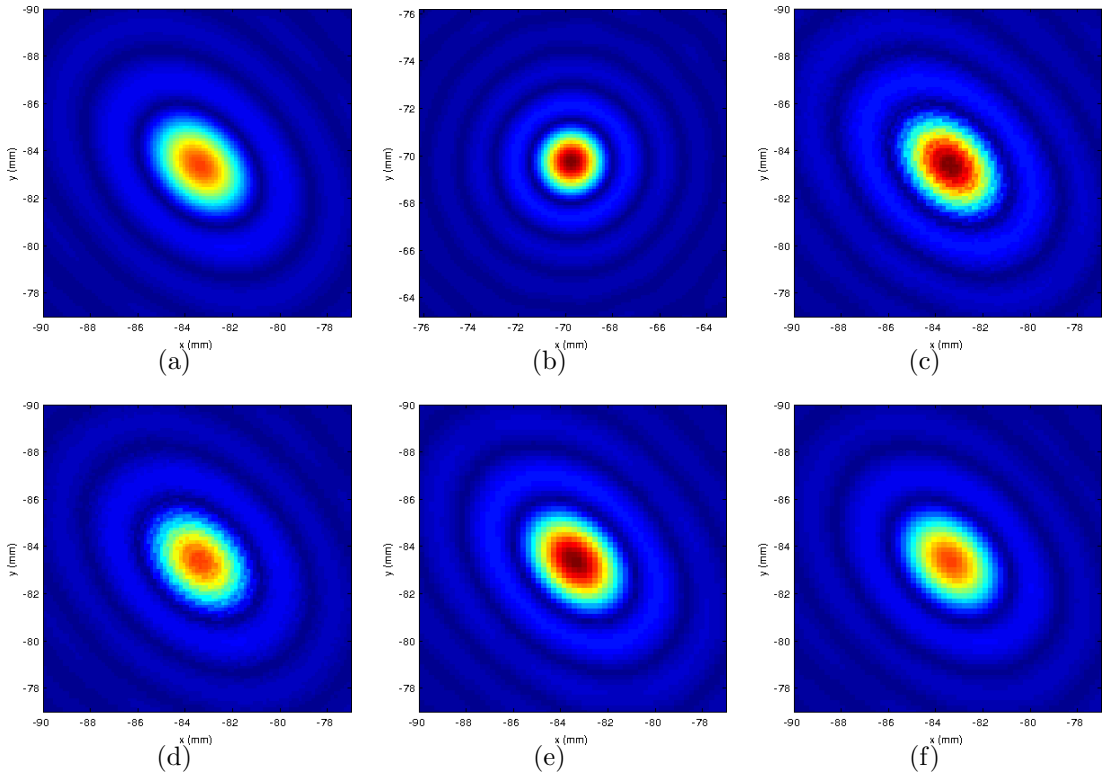


Figure 3.7 — Zoomed images of the top left corner of the simulated diffraction patterns obtained by using: (a) RS integral, (b) Fraunhofer approximation, (c) Angular Fraunhofer approximation with nearest-neighbor interpolation, (d) Multiradii approach with nearest-neighbor interpolation, (e) Angular Fraunhofer approximation with bilinear interpolation, (d) Multiradii approach with bilinear interpolation.

non-paraxial angles, where a ratio of $(\gamma/R')/(1/z) = \gamma^2$ is dropped compared to the multiradii Harvey method (3.8). This amplitude difference can be seen in Fig. 3.7(b) and 3.7(c), where the spots' peaks are higher than those in Fig. 3.7(a) and 3.7(d). These factors can be important for non-paraxial diffractive applications such as beam shaping or beam splitting where high spot power uniformity is often required.

Table 3.1 — Simulated diffraction angles and efficiencies of sample diffraction orders calculated by different methods. The RS convolution and the ASM are not shown as they are inapplicable to calculate the entire output field due to sampling constraint.

Method	**Time**	**Diff. angle (°) of order *mn***					**Diff. eff. (%) of order *mn***				
		01	**11**	**02**	**12**	**22**	**01**	**11**	**02**	**12**	**22**
RS convol.	4.8 s	Inapplicable									
ASM	4.5 s	Inapplicable									
RS integral	2 days	11.1	16.0	22.8	25.7	33.2	2.73	2.32	2.55	2.56	2.05
Fraunhofer	1.8 s	11.0	15.3	21.2	23.4	28.7	2.75	2.35	2.62	2.66	2.23
Angular Fra.	3.5 s	11.1	16.0	22.8	25.7	33.2	2.93	2.68	3.49	3.84	4.16
Multiradii	3.7 s	11.1	16.0	22.8	25.7	33.2	2.72	2.29	2.53	2.53	2.03

Table 3.1 compares the simulation time together with output diffraction angles and spot efficiencies for the different methods, in which $N = 4096$ pixels and Eq. (3.13) is used for power calculation with the spot area of 50×50 pixels. The simulation time was obtained on a PC with an Intel core i5 @ 3.2 GHz processor and 16 GByte of RAM running under Linux. For reasons which will be explained in the next section, the designed DOE was oversampled by a factor of 8, leading to $\delta_1 = 0.125$ μm. Note that there is almost no difference in the spot diffraction angle between the RS integral, angular Fraunhofer approximation and our multiradii Harvey technique, while the maximum difference for the Fraunhofer approach is about 13.6%. For diffraction efficiency, as the angular Fraunhofer approximation predicts correctly the spots' diffraction angle and spots' sizes, but over estimates the amplitude of the spots by a factor of $1/\gamma^2$, the diffraction efficiency is over estimated by $1/\gamma^4$. For example, the ratio of diffraction efficiency of order **22** between the RS integral calculation and the angular Fraunhofer approximation is $4.16/2.05 \approx 1/\cos^4 33.2°$. The Fraunhofer approximation over estimates the spots' magnitude by the same factor, but under estimates the spots sizes, and therefore the error in predicting spots' diffraction efficiency is reduced, with the maximum difference of about 9%. As is to be expected the differences between both the Fraunhofer approximations and the RS integral are greater for the higher diffraction angle spots. The difference between our method and the RS integral is about 1%, which is mainly due to the interpolation step, but can be reduced by increasing the number of samples N or using a more accurate interpolation method, at the expense of slightly increased calculation time. We will later show in Chapter 5 that this simulation error can be negligible compared with the errors often occured in the fabrication process.

3.1.6 Summary of scalar non-paraxial far-field method

In summary, a new technique to calculate the scalar non-paraxial diffraction pattern of a DOE at an observation plane in the far-field by combining a projection step with the Harvey model has been proposed. Numerical results for sample DOEs show that the proposed spherical wave projection, or equivalently the multiradii approach is more accurate than the traditional Fraunhofer approximations in predicting both diffraction order position and diffraction spot power. The error compared with the rigorous RS integral is about 1%, while its computation is much less time-consuming. Moreover, our method does not impose strict sampling requirements on the output plane, unlike the RS convolution and the ASM, so the diffraction pattern in the far-field can be calculated using a simple FT and a projection step. Simulation results show that our method can be used to model DOEs for diffraction angles up to about 33° and perhaps even higher with only some decrease in calculation accuracy. The next section will present an iterative scalar algorithm for designing non-paraxial Fourier DOEs based on our proposed propagation model.

3.2 Iterative scalar non-paraxial algorithms for Fourier elements

As indicated in Chapter 1, recent developments in fabrication technology have enabled the manufacturing of DOEs with nanoscale features [112], which are comparable to or smaller than the illuminating wavelength. These DOEs diffract light at large angles and therefore can be used in a promising field of applications [2]. However, scalar paraxial theory is not valid for the design of such DOEs [11], while rigorous but extensive calculation vector diffraction models are often limited to 1D gratings and relatively simple 2D structures [61,113]. It has been reported [7,50] that scalar non-paraxial theory can be used, within the limit of the Thin Element Approximation (TEA), for the design of some Fresnel DOEs having sub-wavelength features. In this section, we design 2D Fourier DOEs based on our efficient scalar non-paraxial propagator, which has been shown to be valid for far-field diffraction by structures having features on the order of the wavelength.

3.2.1 Limit of the iterative scalar paraxial design

The family of IFTAs has been proven to be one of the most optimized algorithms for designing phase DOEs [6, 51]. The applicable domain of current IFTA is however limited by the validity of the diffraction model, which is the scalar paraxial theory. Fig. 3.8 shows a desired image pattern, which is a grid pattern with the maximal diffraction angle of about 29° at the corners of the grid. A Fourier DOE was designed at δ_1 = 400 nm and λ = 633 nm using a standard multi-stage IFTA [6]. As expected, the reconstruction pattern of this DOE shows two types of distortions. The first distortion is in diffraction position, which is similar to the pincushion distortion in lens systems [111, 114]. The other distortion is in intensity distribution, where the power decreases from the center towards the grid corners, or equivalently with increasing diffraction angle. These distortions are due to the approximations in calculating diffraction position, diffraction amplitude and diffraction power in the scalar paraxial theory.

For a usual IFTA to function correctly, it is necessary to "back propagate" the field from the output plane to the DOE plane. In the case of our diffraction model, to design a non-paraxial IFTA, it is necessary to "back propagate" the field from the output plane to the hemisphere and then from the hemisphere to the DOE plane, as shown in Fig. 3.9. The hemisphere to DOE back propagation is relatively straight-forward: a simple inverse FT. However, the output plane to hemisphere back projection is more complicated. We propose two algorithms based on different projections between the observation plane and the hemisphere. In both algorithms, the spot intensity is corrected according to the diffraction angle ($P \propto \gamma|U|^2$). Other than this, the same

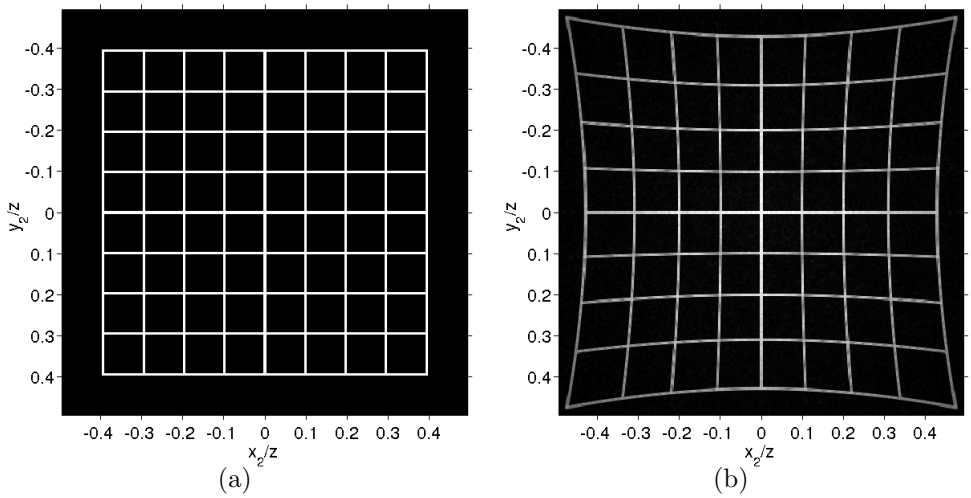


Figure 3.8 — (a) Desired output pattern. (b) Reconstruction pattern of a DOE designed using the standard IFTA.

constraints are used in the DOE plane for the quantization and in the image plane for optimizing the diffraction efficiency and uniformity of the diffraction pattern.

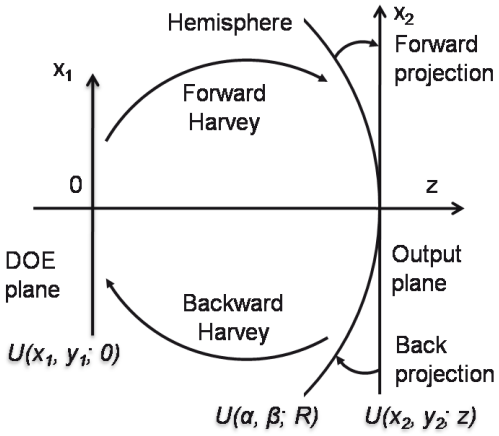


Figure 3.9 — Non-paraxial Fourier DOE design algorithm using single projection from the output plane to the hemisphere and IFTA between the hemisphere and the DOE plane.

3.2.2 Single projection

Fig. 3.10 illustrates the single projection algorithm, where the desired pattern on the output plane is back projected to the hemisphere only once, before an IFTA is operated between the DOE plane and the hemisphere. The projected pattern in angular coordinates on the hemisphere is shown in Fig. 3.11(a), where a nearest neighbour interpolation [94] is used for mapping the field from spatial to angular coordinates. This pattern looks similar to the barrel distortion in lens systems [111, 114], and will

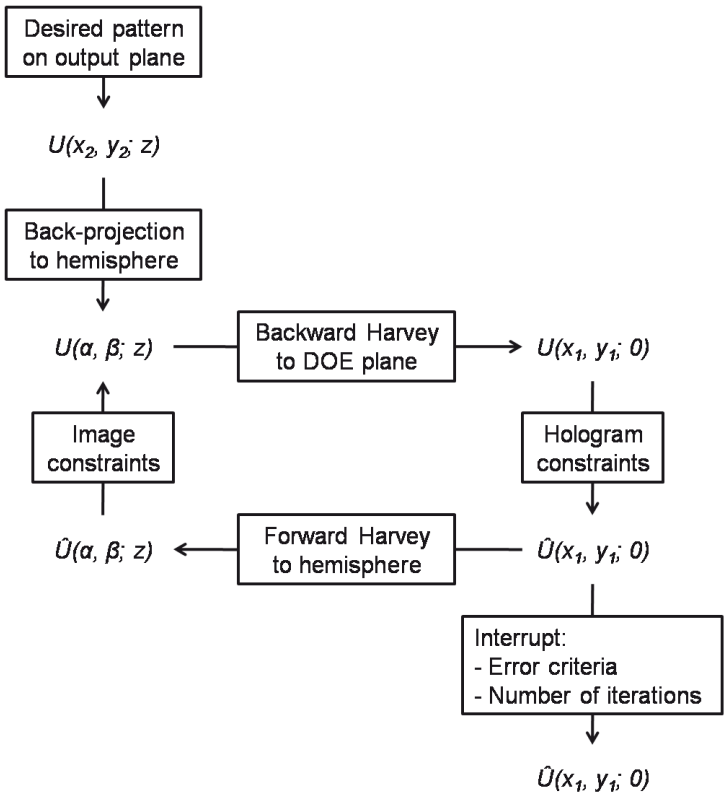


Figure 3.10 — Non-paraxial Fourier DOE design algorithm using single projection from the output plane to the hemisphere and IFTA between the hemisphere and the DOE plane.

compensate the pincushion distortion due to the diverging from the hemisphere to the output plane. In this image, the spot intensity has also been corrected according to the diffraction angle. The DOE is then designed by using the standard multi-stage IFTA using this "distorted" image as the target pattern. Fig. 3.11(b) illustrates the reconstruction on the observation plane of our designed DOE, which shows better reconstruction in both diffraction position and intensity than the one designed by the standard IFTA.

However, as nearest-neighbor interpolation is a one-to-one mapping [94], the output pattern contains an error of up to 1 pixel in diffraction position, as shown in Fig. 3.11(c). In practice, the diffracted beam diverges from the hemisphere to the output plane and therefore, the field at one point on the output plane should be calculated from the field at several points on the hemisphere and vice versa. Attempts to use a many-to-one mapping based on a bicubic interpolation [94] in the single projection algorithm lead to a stagnation (i.e. local optimum) of the iterative algorithm, as shown in Fig. 3.11(d), where the lines are discontinuous. We therefore looked to improve the convergence of the design, so that it can converge to the global optimum at the end of the iteration by using an iterative projection algorithm.

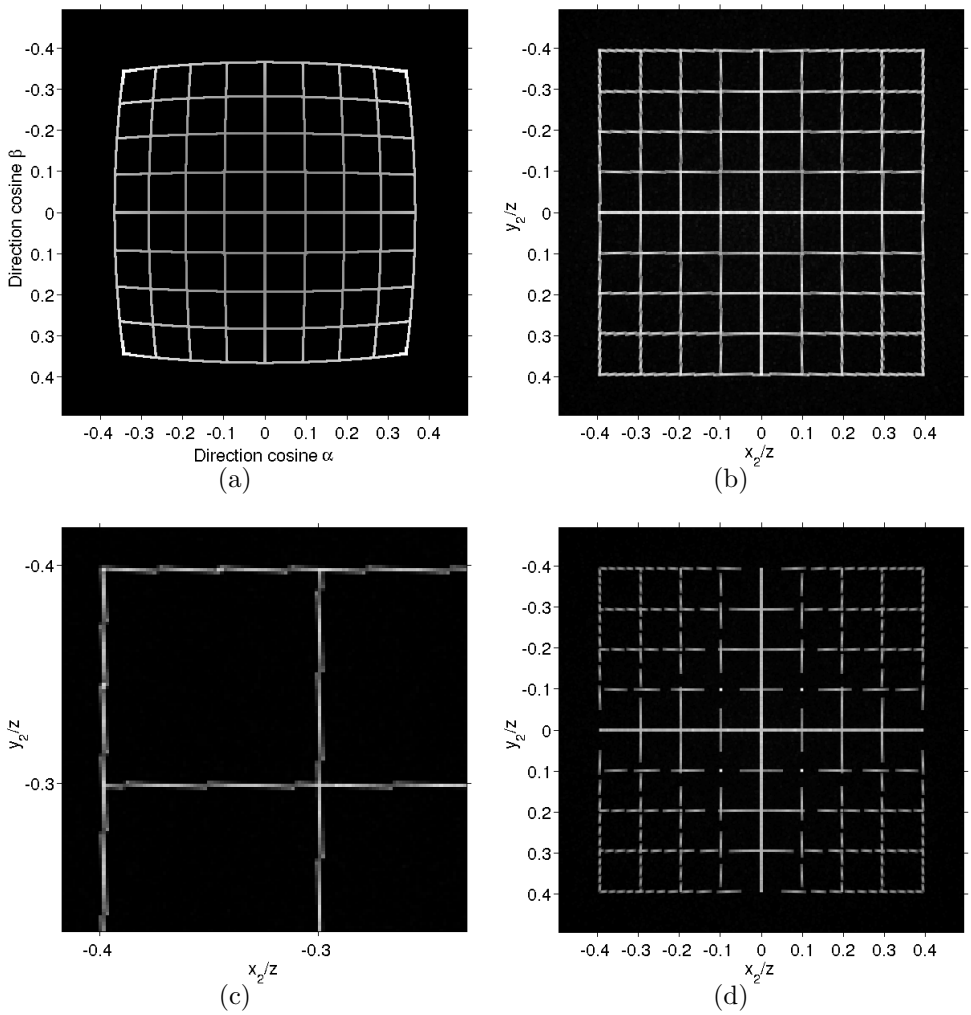


Figure 3.11 — (a) Projection of the grid pattern on the hemisphere using nearest-neighbor interpolation. (b) Reconstruction pattern of the DOE designed using single projection with nearest-neighbor interpolation. (c) Zoomed image of the top left corner of the pattern in (b). (d) Reconstruction pattern of the DOE designed using single projection with bicubic interpolation.

3.2.3 Iterative projection

Fig. 3.12 illustrates the iterative projection algorithm, where the projection is used back and forth between the hemisphere and the output plane, together with the forward and backward FT between the DOE plane and the hemisphere. As the algorithm uses two-step propagation iteratively, the complexity of this algorithm almost doubles that of the single projection one. In return, the algorithm converges to a far more optimized solution, even with the use of a many-to-one mapping, as shown in Fig. 3.13(a) for bicubic interpolation. A closer look on Fig. 3.13(b) confirms that this algorithm produces an output pattern better than the single projection one in diffraction position, although some "ringing", which is due to the bicubic interpolation [94], is still visible.

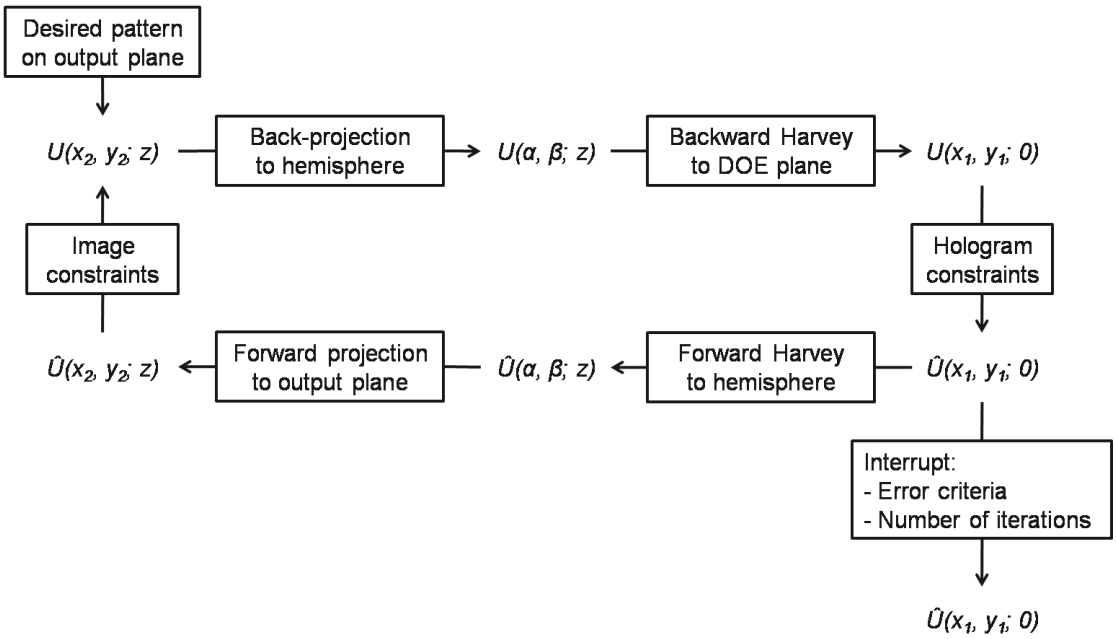


Figure 3.12 — Non-paraxial Fourier DOE design algorithm using iterative projection between the output plane and the hemisphere and IFTA between the hemisphere and the DOE plane.

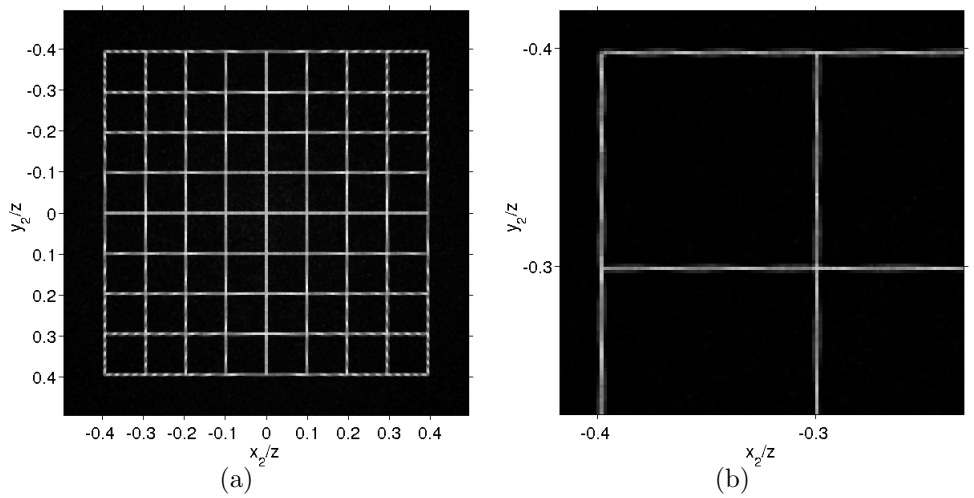


Figure 3.13 — (a) Reconstruction pattern of a DOE designed at $\delta_1 = 400$ nm using iterative projection with bicubic interpolation. (b) Zoomed image of the top left corner of the pattern in (a).

Table 3.2 — Comparison between the performance of different design algorithms (after 110 iterations).

Method	Design time (s)	Output MSE
Standard IFTA	16	2.43
Single projection (nearest)	26	0.39
Single projection (bicubic)	26	0.62
Iterative projection (bicubic)	52	0.24

Table 3.2 compares the performance of different algorithms in designing 1024 x 1024 pixel binary Fourier DOEs at $\lambda = 633$ nm and $\delta_1 = 400$ nm. The design time was obtained on the same computer as before, i.e a PC with an Intel core i5 @ 3.2 GHz processor and 16 GByte of RAM running under Linux. The output patterns of these DOEs were then reconstructed using our scalar non-paraxial diffraction propagator. The mean-squared-error (MSE) of the reconstructions with respect to the target image were calculated, which shows better reconstruction with lower MSE value [3].

$$MSE = \frac{1}{N^2} \sum_{\mathbf{m}=1}^{N} \sum_{\mathbf{n}=1}^{N} (P_{\mathbf{mn}} - P^{\circ}_{\mathbf{mn}})^2 \tag{3.17}$$

where $P_{\mathbf{mn}}$ and $P^{\circ}_{\mathbf{mn}}$ are the diffraction power of pixel $(\mathbf{m}, \mathbf{n})$ in the reconstruction and target pattern, respectively.

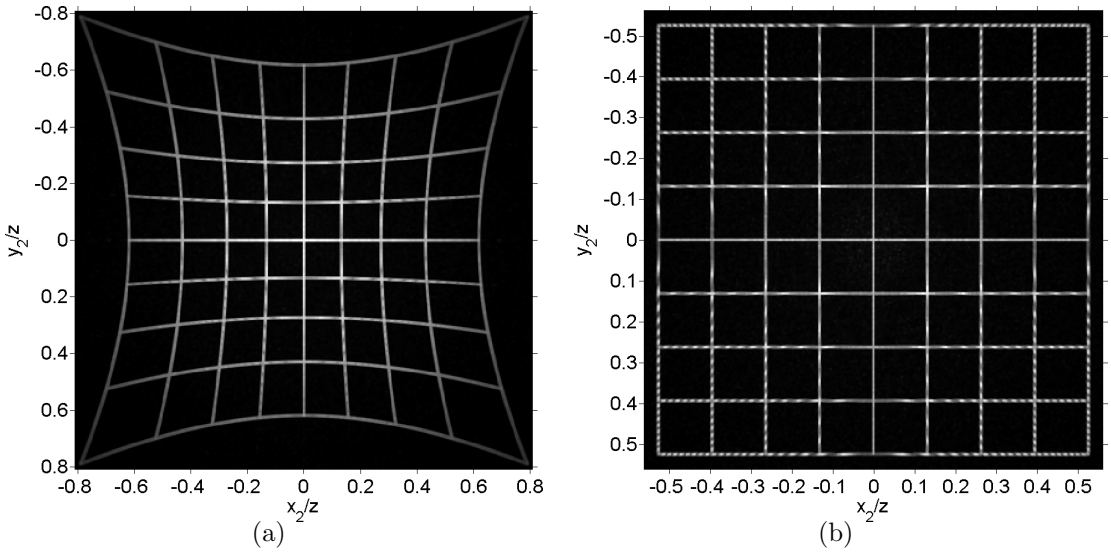


Figure 3.14 — Reconstruction pattern of the DOEs designed at $\delta_1 = 300$ nm using (a) standard IFTA, (b) iterative projection with bicubic interpolation.

Eq. 3.10 suggests that to design DOEs with wider angles of diffraction, we can increase either the ratios $\mathbf{m}/N$, $\mathbf{n}/N$ or λ/δ_1. While the maximum of the former in our

grid pattern is 0.25 (at the corners of the grid), the latter can be increased more freely by using a longer wavelength or a smaller pixel size. In this way, it is possible (within the scalar theory) to design DOEs with diffraction angle up to 90° with the iterative projection algorithm, but in practice, it is limited by the fabrication technology. With this approach, we have designed and fabricated of DOEs at pixel size of 400 and 300 nm, where the maximum diffraction angle at the wavelength of 633 nm is about 29 and 37°, respectively. The simulation results of the DOEs designed at 300 nm using the standard IFTA and our iterative algorithm are shown in Fig. 3.14(a) and 3.14(b), respectively. The experimental verification for all the design algorithms presented in this section will be given in Chapter 5.

3.2.4 Summary of iterative scalar non-paraxial far-field algorithms

In summary, we have developed a fast algorithm but accurate in assigning both diffraction position and diffraction power by using an iterative projection between the output plane and the hemisphere and an iterative FT between the hemisphere and the DOE plane. The algorithm can be used to design non-paraxial Fourier elements which have many applications in practice, e.g. beam splitters, beam shapers, where the propagation distance is often much larger than the DOEs size [3]. The reconstruction pattern of the DOEs designed by our algorithm shows better distribution in both position and intensity than those designed by a standard multi-stage IFTA and very close to the desired pattern, even for large diffraction angle (up to 37°). The validity of our design algorithms no longer requires the paraxial approximation and therefore primarily depends on the TEA, which is used to relate the DOE phase profile and the optical field at the DOE plane [11]. However, the accuracy of the approximation does not yet appear to be fully understood in the literature: some publications indicated that DOE feature sizes should be at least ten times the illuminating wavelength [115, 116], recent work showed that this approximation can be used for some DOEs having features on the order of, or even smaller than the illuminating wavelength [7, 50]. The limit of this approximation will be analyzed using vectorial theory in the next chapter.

3.3 Conclusion

In conclusion, we have developed a scalar non-paraxial diffraction model and an iterative algorithm based on this propagator for the modeling and design of wide-angle Fourier elements. Simulation results showed that our scalar non-paraxial propagator is equivalent to but much faster than the rigorous scalar Rayleigh-Sommerfeld diffraction integral, and that our iterative algorithm converges to much more optimized solutions

than the standard IFTA at the same order of computational complexity. This approach can be used to model and design Fourier DOEs with diffraction angles up to about 37° and perhaps even higher, which is beyond the limit of the scalar paraxial diffraction regime. The validity of our modeling and design algorithms therefore primarily depend on the TEA, which is used to relate the DOE phase profile and the optical field at the DOE plane. In the next chapter, the practical limit of this approximation is examined, and a vectorial approach is proposed to overcome these limitations.

CHAPTER 4

Vectorial theory for the modeling and design of DOEs operating in the non-paraxial far-field diffraction regime

In the previous chapter, we have demonstrated that wide-angle diffraction, which is the aim of this thesis, can be obtained with thin DOEs having feature sizes on the order of, or even smaller than the wavelength, at the cost of a more expensive fabrication process [117]. However, high resolution DOEs are not the only way to obtain large diffraction angles. It is known in optical holography (e.g. volume Bragg gratings [11]) that wide diffraction angles may also be obtained with μm feature size DOEs having deeper etching depth, i.e. those with the thickness much larger than the wavelength. This type of DOE is important to us as it can now be fabricated in our cleanroom, but the synthesis of thick DOEs appears to have been little studied.

Although our scalar non-paraxial propagator is rigorous for free-space propagation [107], the validity of the diffraction model primarily depends on the Thin Element Approximation (TEA), which is used for the DOE region [7, 11]. This approximation cannot be used for phase shift $> 2\pi$ (i.e. thick DOEs), and yet there seems to be no good way to model these structures. In this chapter, we aim to develop a rigorous vectorial method for the modeling and design of thick DOEs. Firstly, the practical limit of the TEA will be examined by comparison with rigorous vectorial simulations, i.e. Finite-Difference Time-Domain (FDTD) and Fourier Modal Method (FMM). We use a well-known and widely used open-source FDTD library implemented by Massachusetts Institute of Technology (MIT), which is called MIT Electromagnetic Equation Propagation (MEEP) [75]. For FMM (also called the Rigorous Coupled Wave Analysis [67,118]), the Grating Toolbox of VirtualLab^{TM}, which is a commercial software developed by LightTrans GmbH, is used [119]. While FMM/RCWA is generally limited to periodic

structures [120], the FDTD method can be used with various boundary conditions. Details on the different boundary conditions are given in Appendix C. Both methods are often limited to 1D or relatively small 2D structures due to the extensive calculation required.

To cope with the differences in the diffraction efficiencies obtained by the various calculation methods and experimentally in practice, we show that diffraction symmetry can be used as a simple yet efficient metric for evaluating the accuracy of the simulations. This metric will be shown as consistent, despite the variations in the simulated and experimental diffraction efficiencies. In what follows, we will present a rigorous diffraction model based on FDTD coupled with our free-space propagator and propose a genetic design algorithm based on this model to overcome the theoretical limits of the TEA and the computational limitations of current vectorial models.

4.1 Limit of the Thin Element Approximation

4.1.1 The diffracted electromagnetic components

Perhaps one of the most noticable examples of the limits of the TEA is the prediction of the direction of the diffracted electromagnetic field after the structure. As a simple illustration of this limitation, Fig. 4.1(a) shows one period of a 1D binary grating illuminated by a linear s-polarized plane wave, i.e. the illumination wave has two components: E_y and H_x. As the electric field oscillates in the direction perpendicular to the two-dimensional plane, it is often called Transverse Electric (TE) field [64, 116].

If we apply the TEA to this structure to calculate its diffraction pattern, the diffracted field after the structure will consist of only the two components, with a phase shift due to the Optical Path Length difference, i.e. $E_y \cdot \exp\{j\phi(x)\}, H_x \cdot \exp\{j\phi(x)\}$. However, a more rigorous FDTD simulation shows that the diffracted field after the grating has three components: E_y, H_x and H_z, as shown in Fig. 4.1(b)-4.1(d). In this simulation, the periodic and perfectly matched boundary conditions (see Appendix C) were used in the $x-$ and $z-$ direction, respectively. The spatial and temporal samplings were chosen fine enough to ensure convergence of the FDTD algorithm [72]. This result suggests that the TEA fails to predict the diffraction of the electromagnetic component perpendicular to the direction of the grooves, i.e. H_x and it should only be used for the component parallel to the grooves, i.e. E_y. The diffracted magnetic components can then be determined by taking partial derivatives of the first Maxwell equation, i.e. $H_x \propto \partial E_y/\partial x$, $H_z \propto \partial E_y/\partial z$, which is rigorous in the vectorial domain [64].

In other words, the TEA fails to predict the generation of the H_z component after the structure. This can be explained due to the oscillation of H_x in the direction

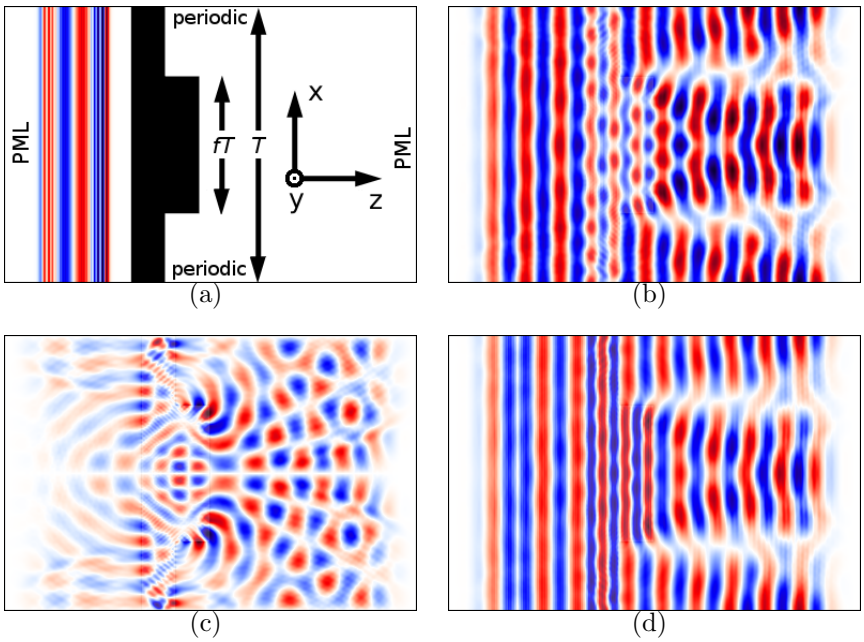


Figure 4.1 — FDTD simulation (MEEP) of a periodic 1D binary grating illuminated by a linear s-polarized plane wave and the diffracted fields. (a) One period of the grating structure and the illumination field. (b) E_y. (c) H_x. (d) H_z.

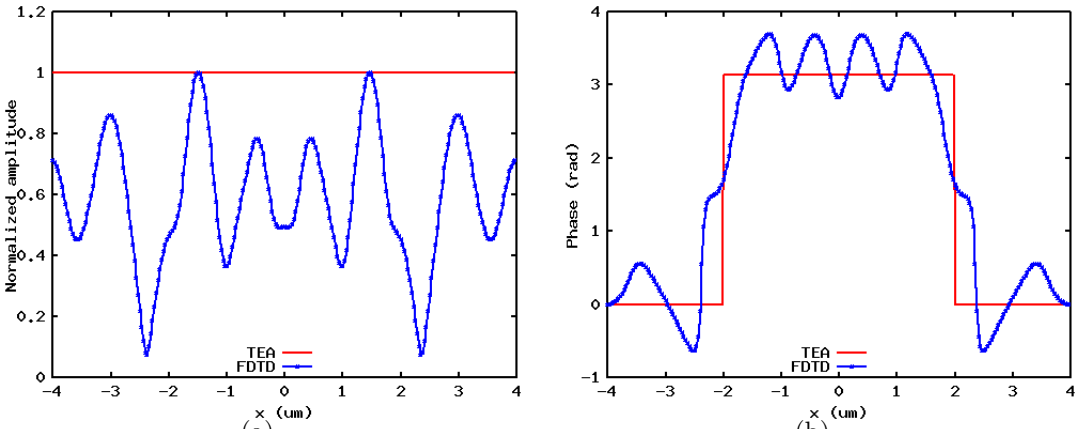
perpendicular to the grooves, leading to diffraction of the magnetic field into the z direction at the interface between different media. Fig. 4.2 shows the state of H_z of the previous example after a few optical cycles, i.e. the electromagnetic field has propagated to the grating and interacted with the structure. Note that the illumination wave on the left hand side of the grating structure has no H_z component, but spherical H_z waves are generated at four corners of the ridge. The effect is similar for the case of linear p-polarized plane waves, i.e. Transverse Magnetic (TM) field, where the illumination has two components H_y, E_x and the diffracted field contains three components H_y, E_x, E_z. For 2D DOEs, the diffracted field after the structures will generally consist of all six electromagnetic components, regardless of the polarization of the illumination wave.

From the above qualitative analysis, it is necessary to quantitatively assess the error of the TEA. Fig. 4.3 shows the normalized amplitude and the phase distribution of E_y immediately after the grating of the above example. The illumination wavelength, the grating period and the fill-factor were $\lambda = 1$ μm, $T = 8$ μm and $f = 0.5$, respectively. The etch depth was chosen as the optimal value predicted by the TEA: $d = \lambda/2(n-1) = 1$ μm, where $n = 1.5$. The FDTD simulation was run for a sufficiently high number of time steps to ensure that the field had stabilized, i.e. the spherical waves generated from neighboring grating periods had contributed to the diffracted field in the simulation region. The phase was unwrapped to avoid 2π discontinuities in the phase response [121]. The amplitude and phase predictions of the TEA are also shown for comparison, which demonstrates that the TEA incorrectly predicts the diffracted field immediately

(a) (b)

Figure 4.2 — The state of H_z obtained by the FDTD simulation (MEEP): (a) after 4.5 optical cycles, and (b) after 6 optical cycles.

after the DOE. It is therefore necessary to quantitatively assess the contribution of the TEA prediction error to the more useful far-field diffraction pattern.



(a) (b)

Figure 4.3 — Normalized amplitude (a) and phase (b) of the diffracted field immediately after the grating. The results were obtained using the TEA ($E_y \cdot \exp\{j\phi(x)\}$) in comparison with the FDTD simulation (MEEP).

Although rigorous vectorial diffraction models have been used for decades [69, 70], there appears to be little or no experimental work verifying the limits of the TEA in predicting the far-field diffraction efficiency, as it is often difficult to separate the simulation errors from the fabrication errors [3]. We have noticed that the symmetry variation of the diffraction pattern of binary DOEs can be used as a simple yet efficient way to determine the limits of the TEA. It can be seen from Fig. 4.3 that although the structure is a binary element, the diffracted field obtained by the FDTD simulation is a complex instead of a real function as predicted by the TEA. As a result, the far-field diffraction pattern will not have Hermitian symmetry as predicted by scalar diffraction theory [2, 11]. However, as the grating has reflection symmetry, the diffracted field after the structure has the same symmetry, as shown in Fig. 4.3. The far-field diffraction

pattern will therefore have the same reflection symmetry which coincides with the Hermitian symmetry in 1D, but this is not the case for general 2D binary structures, as will be shown in the next section.

4.1.2 Symmetry of the far-field diffraction pattern of binary elements

In experiment, we have observed variations in the symmetry of the diffraction pattern by Fourier DOEs fabricated at feature sizes on the order of the illumination wavelength. In this section, we will assess the limit of the TEA in predicting two types of symmetries that are frequently used in practice, which are Hermitian and reflection symmetry [11, 122].

Hermitian symmetry of binary Fourier DOEs

Fig. 4.4(a) shows an asymmetrical binary Fourier DOE, which was designed to generate a 5×5 spot array as illustrated in Fig. 4.4(c). The far-field diffraction pattern of this DOE with the refractive index $n = 1.64$ and the feature size $s = 1$ μm was then reconstructed in simulation at $\lambda = 633$ nm using the TEA for the DOE region and a scalar non-paraxial propagator for the free-space propagation region. For reference, the FMM, which is a rigorous vectorial diffraction method included in LightTrans VirtualLab^{TM}, was used, since the far-field diffraction pattern can be obtained directly by the FMM [69,70], without the need of a near-field to far-field transformation as in the FDTD [123, 124]. As the rigorous Rayleigh-Sommerfeld diffraction integral had been shown to be equivalent to our scalar non-paraxial diffraction model, which is based on a single Fourier Transform (FT), we would expect the symmetry of the far-field diffraction pattern to depend on the symmetry of the FT. At the optimum etching depth predicted by the TEA, the phase shift between the binary levels is π, which means that the diffracted field after a binary DOE can be represented in the TEA by a real function $f(x, y)$ (i.e. $e^{j.0} = 1, e^{j.\pi} = -1$). Therefore the diffraction pattern $F(X, Y)$ should have Hermitian symmetry [11]:

$$f(x, y) = f^*(x, y) \rightarrow F(X, Y) = F^*(-X, -Y) \tag{4.1}$$

where $F(X, Y)$ is the FT of $f(x, y)$ and the symbol * denotes the complex conjugate. This means that the diffracted powers of the spots should be Hermitian symmetric, as $|F(X, Y)|^2 = |F(-X, -Y)|^2$. Table 4.1(a) lists the simulated diffraction efficiency of different diffraction orders $\eta(\mathbf{m}, \mathbf{n})$ obtained by the TEA and our scalar model. Indeed, Hermitian symmetry can clearly be seen, i.e. $\eta(\mathbf{m}, \mathbf{n}) = \eta(-\mathbf{m}, -\mathbf{n})$, with $\mathbf{m}, \mathbf{n} = -2, \ldots, 2$.

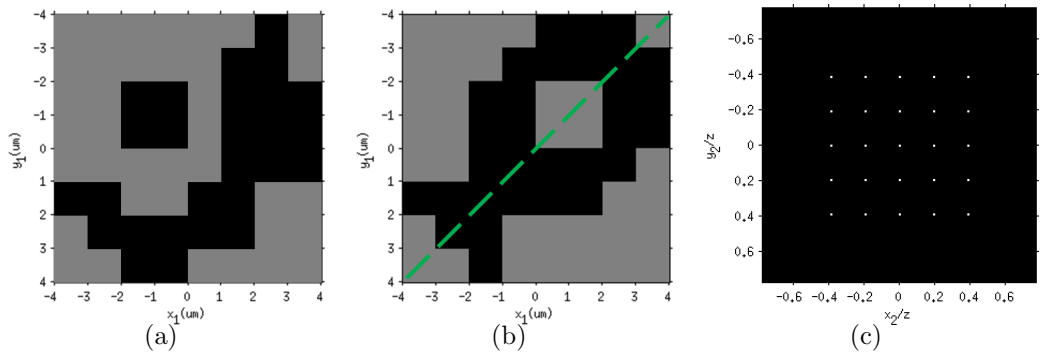


Figure 4.4 — (a) A period of an asymmetrical binary Fourier DOE. (b) A period of a symmetrical binary Fourier DOE, where the dashed line is the reflection symmetry axis. (c) Target image.

However, when the vectorial FMM simulation technique is used, although the predicted diffraction angles are the same as the scalar simulation, the predicted diffraction efficiencies differ considerably. In particular, the symmetry of the diffracted pattern changes, i.e. $\eta(\mathbf{m}, \mathbf{n})$ and $\eta(-\mathbf{m}, -\mathbf{n})$, as shown in Table 4.1(b). This can be explained by the previous example FDTD simulation. Although the asymmetrical DOE is binary at the optimum etching depth predicted by the TEA, the phase of the diffracted field after the structure is a continuous function instead of a π phase shift between binary levels, as shown in Fig. 4.3(b). As the result, the diffracted field after the structure is a complex function, and the far-field diffraction pattern, which can be calculated based on a FT, is not Hermitian symmetric.

Table 4.1 — Simulated diffraction efficiency (%) of diffraction orders (**m**, **n**) of the asymmetrical DOE using: (a) the TEA + our free-space propagator and (b) the FMM (LightTrans VirtualLab^{TM}).

(a)

m \ **n**	**-2**	**-1**	**0**	**1**	**2**
-2	2.33	2.74	2.63	1.87	2.34
-1	2.13	2.27	2.64	2.88	2.33
0	2.61	2.55	2.30	2.55	2.61
1	2.33	2.88	2.64	2.27	2.13
2	2.34	1.87	2.63	2.74	2.33

(b)

m \ **n**	**-2**	**-1**	**0**	**1**	**2**
-2	2.54	2.69	2.71	1.67	2.55
-1	2.04	2.19	2.43	2.60	1.68
0	2.83	2.28	2.51	2.62	2.84
1	2.25	2.67	2.43	2.52	2.22
2	2.41	1.37	2.74	2.45	2.19

In order to show the limit of the TEA in predicting the diffraction symmetry of binary Fourier elements, we define an average Hermitian symmetry factor as follows:

$$S_H = \frac{1}{N_H} \sum s, \text{where } s = \frac{\eta(\mathbf{m}, \mathbf{n})}{\eta(-\mathbf{m}, -\mathbf{n})}, \text{with } s \leq 1 \tag{4.2}$$

In this expression, N_H is the number of Hermitian symmetric pairs of diffraction spots predicted by the TEA and the free-space propagator, e.g. $N_H = 12$ for the case of our 5×5 spot array DOE. The far-field diffraction patterns of the DOE at different feature sizes were reconstructed using the FMM. The average Hermitian symmetry factor was calculated at different feature sizes and shown in Fig. 4.5 in comparison with the TEA prediction, i.e. $S_H = 1$. It can be seen that the FMM average Hermitian symmetry factor drops rapidly after the DOE feature size of 1 μm, which is less than 2 times the wavelength, and the slope is even steeper when the feature size is smaller than the wavelength.

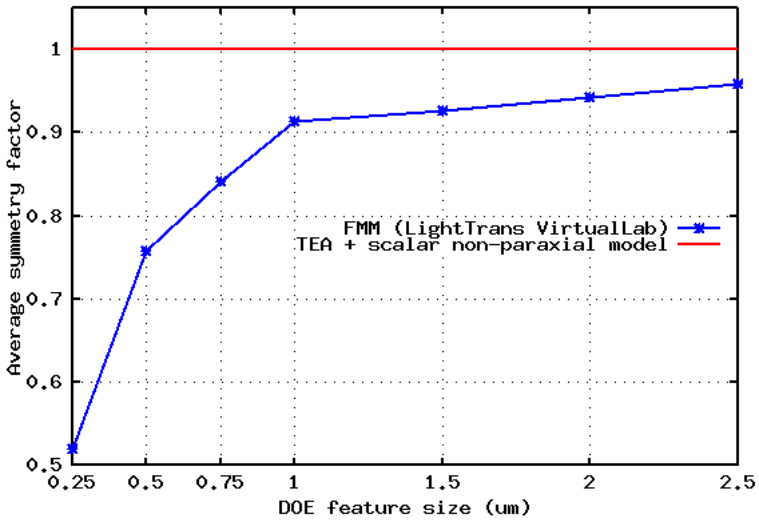


Figure 4.5 — Average Hermitian symmetry factor at different DOE feature sizes. FMM simulations of DOEs having feature sizes larger than 2.5 μm were not possible due to the limited amount of computing memory available (16 GByte of RAM).

From the above analysis, we can see that, in the case of this DOE and based on the Hermitian symmetry of the diffraction pattern, the TEA should not be used for this DOEs with feature sizes smaller than about 2 times the illumination wavelength (within the variation of about 10% in the average Hermitian symmetry factor). In the next section, we will assess the error of the TEA in predicting the reflection symmetry of the diffraction pattern of symmetrical elements.

Reflection symmetry

Assuming that the DOE is reflection symmetric, as shown in Fig. 4.4(b). According to the TEA and our free-space propagator, which is based on a single Fourier Transform (FT), the far-field diffraction pattern should have the same reflection symmetry [125]:

$$f(x, y) = f(-y, -x) \rightarrow F(X, Y) = F(-Y, -X) \tag{4.3}$$

The FMM simulation confirms the reflection symmetry of the diffraction pattern, i.e. $\eta(\mathbf{m}, \mathbf{n})$ and $\eta(-\mathbf{n}, -\mathbf{m})$, as shown in Table 4.2. For this binary phase DOE, based on the TEA + scalar free-space propagator, the diffraction pattern should be Hermitian symmetric so the diffraction efficiency of order (-2, 1) should be the same as that of order (2, -1), but in the reference FMM simulation, they are different from each other (i.e. 2.11% and 2.58%, respectively). Instead, the diffraction efficiencies of order (-2, 1) and order (-1, 2) are almost the same (i.e. 2.11% and 2.10%, respectively). Further analysis shows that instead of Hermitian symmetry, we observe a reflection symmetry of the diffraction pattern in the same axis as the reflection symmetry of the designed DOE, which is the dashed line shown in Fig. 4.4(b).

Table 4.2 — Simulated diffraction efficiency (%) of diffraction orders (**m**, **n**) of the symmetrical DOE with spherical polarization illumination using the FMM (Light-Trans VirtualLab^{TM}).

m	**n**				
	-2	**-1**	**0**	**1**	**2**
-2	2.02	3.53	3.15	2.11	1.28
-1	3.19	3.43	2.87	2.48	2.10
0	3.15	2.88	1.00	2.88	3.13
1	2.58	1.93	2.91	3.39	3.54
2	0.82	2.58	3.23	3.11	2.02

Similar to the average Hermitian symmetry, we define an average reflection symmetry factor of the diffracted spot powers as follows:

$$S_R = \frac{1}{N_R} \sum s, \text{where } s = \frac{\eta(\mathbf{m}, \mathbf{n})}{\eta(-\mathbf{n}, -\mathbf{m})}, \text{with } s \leq 1 \tag{4.4}$$

In this expression, N_R is the number of reflection symmetric pairs predicted by the TEA and the free-space propagator, e.g. $N_R = 10$ for the case of our symmetrical DOE. The average reflection symmetry factor and Hermitian symmetry factor were calculated as $S_R \approx 0.99$ and $S_H \approx 0.90$, respectively. However, based on scalar diffraction theory, the diffraction pattern should have both symmetries, i.e. $S_R = 1$ and $S_H = 1$. The 1% difference in reflection symmetry seems to be the polarization dependent effect of the diffractive structure [116], which is a vectorial electromagnetic phenomenon. Nevertheless, this error is much smaller than the 10% error in the TEA prediction of the Hermitian symmetry. For 1D binary gratings at normal incidence, there is always a reflection symmetry axis, and the diffraction patterns also have reflection symmetry, which coincides with the Hermitian symmetry, as illustrated in Fig. 4.6. The angular diffraction pattern shows the incident wave on the left and the transmitted diffraction orders on the right, where the length of each line represents the diffraction efficiency of indicated orders in the given direction.

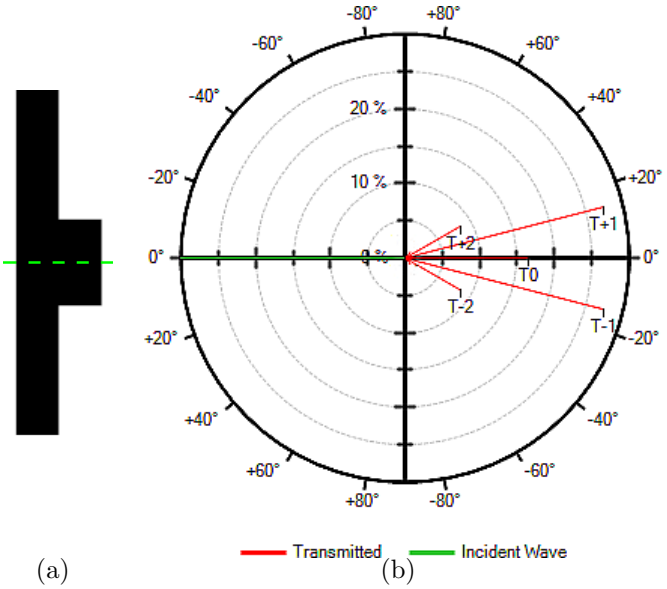


Figure 4.6 — (a) A binary grating with $T = 4$ μm and $f = 0.25$, where the dashed line is the reflection symmetry axis. (b) The angular diffraction pattern of the grating simulated using FMM in VirtualLab^{TM} ($\lambda = 1$ μm, linear s-polarized illumination).

4.1.3 Summary of our TEA investigation

In summary, the limits of the TEA in predicting the electromagnetic field after the structure in comparison with different vectorial methods have been investigated. FDTD simulations show that the TEA fails to predict correctly the diffracted field distribution immediately after the DOE in both amplitude and phase. FMM simulations indicate that the TEA and our free-space propagator model accurately non-paraxial diffraction angles, but incorrectly predict the symmetry of diffraction efficiencies of some DOEs having feature sizes smaller than about 2 times the wavelength. Therefore, to design wide-angle pattern generation DOEs such as those producing the grid patterns in [117], our iterative scalar non-paraxial algorithm may be sufficient. However, for designing DOEs with high uniformity diffraction efficiency, e.g. spot array generators, it is necessary to take the limitations of the TEA into account and include vectorial effects in the DOE region. Although FMM is rigorous and relatively fast (as the far-field diffraction pattern can be obtained directly), it is often limited to infinite periodic structures [69, 70]. Researchers have been looking for a way to model finite and more general structures by coupling the FDTD with some free-space propagators [7, 61, 62]. However, the dependence of the algorithm's accuracy on the coupling parameters has not been studied in detail. We will present in the next section a rigorous vectorial model based on FDTD coupled with our far-field free-space propagator, and compare the results to those obtained by the TEA coupled with the same free-space propagator.

4.2 Rigorous vectorial model

Although FDTD is rigorous, it is generally limited to a small region around the DOEs due to the extensive calculation and memory requirements (the DOE must be spatially sampled at a very fine sampling distance - at least $\lambda/20$ [72], and all 3D sampling points held in memory simultaneously). In order to propagate the electromagnetic wave to a far-field plane, the output field calculated by FDTD is used as the input of a rigorous free-space propagator [7, 61, 62]. To verify the accuracy of the coupling model, the simulation results of the FMM simulation using the commercial LightTrans VirtualLab^{TM} software were used as the reference. However, the FMM implementation of LightTrans VirtualLab^{TM} calculates the far-field diffraction pattern on a sphere, as shown in Fig. 4.6, similarly to the Harvey model, which is also a rigorous free-space propagator [105]. For this reason, only the Harvey model will be used instead of our full scalar propagator (i.e. without the projection step) to propagate the FDTD and TEA field to the reference FMM hemisphere, as illustrated in Fig. 4.7.

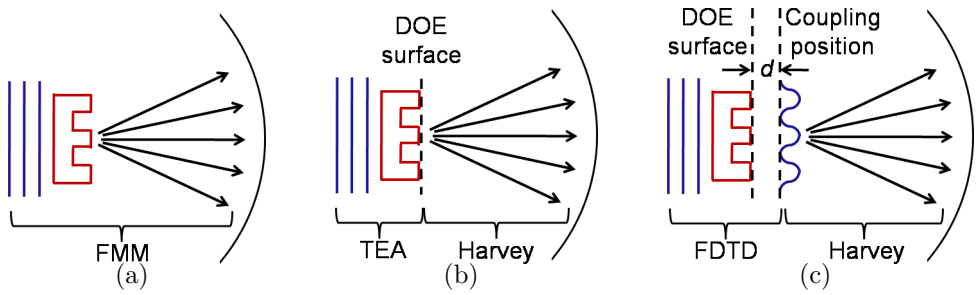


Figure 4.7 — Simulated diffraction efficiency of the coupling model at different FDTD spatial sampling resolutions: (a) The FMM (LightTrans VirtualLab^{TM}). (b) The TEA + Harvey model. (c) The FDTD + Harvey model.

Table 4.3 compares the results of the FDTD + Harvey model with the FMM simulation results. Due to the FDTD and FMM calculation load restrictions, we started with a 1D binary grating having a grating period of 8 μm and the fill-factor of 0.5. The illumination wave was a s-polarized light at $\lambda = 1$ μm. As before, the FDTD simulation was run for a sufficiently high number of time steps (200 optical cycles) to ensure that the field had stabilized, and the FDTD output field consisted of 3 complex-valued electromagnetic components. Each of these components was propagated separately to the far-field using the Harvey propagator, using the same FDTD spatial sampling distance ($\lambda/40$). At the reference FMM hemisphere, the Poynting vector was computed from the 3 far-field electromagnetic components for the calculation of the diffraction power. The effects of the FDTD spatial resolution, FDTD propagation distance d and FDTD simulation time, together with the spatial coupling resolution between the FDTD + Harvey propagator are relatively complex and will be addressed in the following sections.

Table 4.3 — Diffraction efficiency (%) of different diffraction orders of the grating using different simulation methods.

Method	Order						
	0	**±1**	**±2**	**±3**	**± 4**	**±5**	**±6**
FMM (VirtualLab^{TM})	1.30	39.64	1.39	3.48	1.68	0.42	2.15
FDTD + Harvey	1.25	40.12	1.43	3.55	1.66	0.47	2.00
TEA + Harvey	0	42.45	0	4.89	0	1.84	0

From these simulations, it appears that the coupling model based on FDTD and the Harvey free-space propagator achieves more accurate results in diffraction efficiencies (besides obtaining the same diffraction angles) than the purely scalar diffraction model. Notably, the scalar diffraction model tends to under estimate the diffraction power of even orders and over estimate the odd ones. We therefore looked to optimize and parallelize the FDTD + Harvey model to ensure accurate simulation results in a reasonable calculation time, so that the technique can be applied to the modeling and design of general 2D DOEs. Further test cases were examined to verify the convergence and the accuracy of the coupling model.

4.2.1 FDTD spatial resolution

As the amount of computing memory and the FDTD simulation time is exponentially proportional to the spatial sampling resolution [75], it is crucial to reduce the FDTD resolution without reducing the simulation accuracy of the coupling model. Fig. 4.8 shows the calculation results of the previous FDTD + Harvey simulation with different FDTD spatial sampling resolutions. It can be seen that the simulated results slowly converge to asymptotic stable values with higher FDTD resolution. For the next simulations, the spatial sampling of about 40 pixels/wavelength was therefore chosen as a compromise between the simulation accuracy and the amount of computing memory/FDTD simulation time.

4.2.2 FDTD propagation distance

Similar to the spatial sampling, the FDTD propagation distance has a high impact on the amount of computing memory and the simulation time. For this reason, the electromagnetic wave is often propagated for only a few wavelengths beyond the structures before being coupled into a free-space propagator [61–63]. However, the FDTD propagation distance, or in other words, the coupling position, does not yet appear to be fully understood in the literature. Table 4.4 lists the diffraction effiencies we obtained from the FDTD electromagnetic field at different distances from the DOE surface. It

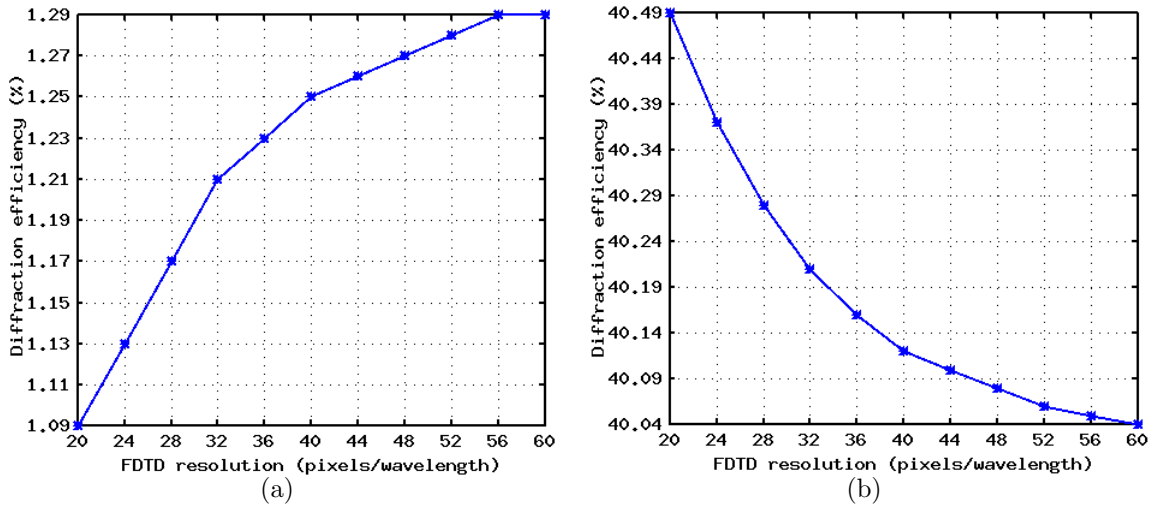


Figure 4.8 — Simulated diffraction efficiency of the coupling model at different FDTD spatial sampling resolutions: (a) 0^{th} order, (b) 1^{st} order.

appears that the coupling postition does not affect the output diffraction efficiencies. Therefore, the field at the DOE surface can be used without the need of propagating to a longer distance, and the amount of computing memory/FDTD simulation time can be reduced.

Table 4.4 — Diffraction efficiency (%) of different diffraction orders of the grating using the FDTD + Harvey model at different distances from the grating surface

Distance from	**Order**						
DOE surface (λ)	**0**	**±1**	**±2**	**±3**	**±4**	**±5**	**±6**
0	1.25	40.12	1.43	3.55	1.66	0.47	2.00
1	1.25	40.13	1.42	3.55	1.66	0.47	2.00
2	1.25	40.12	1.42	3.55	1.66	0.47	2.00
4	1.25	40.12	1.42	3.56	1.66	0.47	2.00
7	1.25	40.12	1.42	3.56	1.66	0.47	2.00
10	1.24	40.14	1.42	3.55	1.66	0.47	2.00

4.2.3 FDTD simulation time

Since the FDTD simulation region is infinitely periodic in transverse (x, y) directions, in theory the FDTD simulation time has to be infinite so that the spherical waves generated from all grating periods have propagated to the field at the coupling distance. However, as the amplitude of a spherical wave at a distance r from the source is inversely proportional to the propagation distance, i.e. $\exp(jkr)/r$, the contribution of the spherical waves generated from a certain point can be neglected. The effect is

illustrated in Fig. 4.9, where the wavefronts of the spherical waves from the neighboring periods, which are indicated by the dashed arrows, are weaker than those of the simulated period, indicated by the solid arrows. This means that, the FDTD simulation time can be limited to a finite value with a minor reduction in the calculation accuracy. Indeed, Fig. 4.10 shows that the simulated results slowly converge to stable levels with higher FDTD calculation times. Therefore, the FDTD simulation should be run for about 100 optical cycles in order for the electromagnetic field in the simulation region to stabilize and the calculated diffraction effiencies to converge.

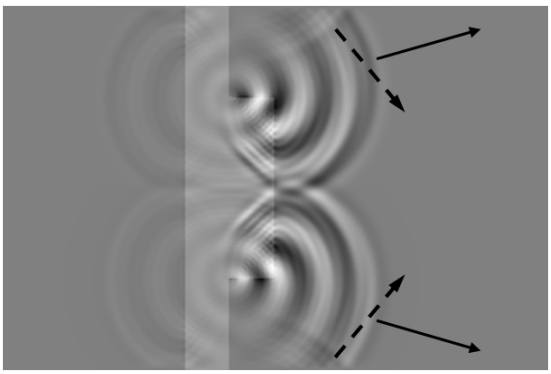


Figure 4.9 — Contribution of the spherical waves from the neighboring periods to the simulation region.

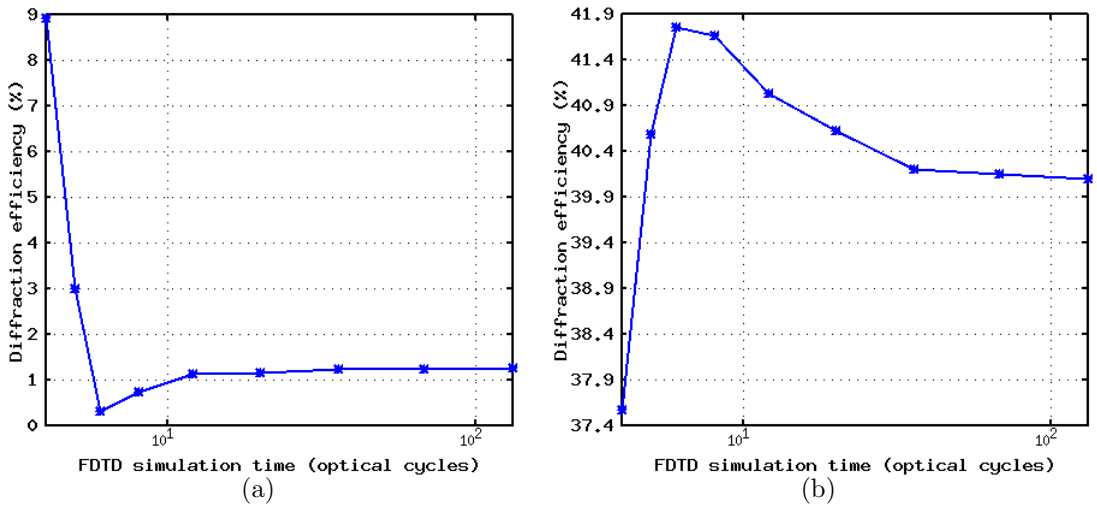


Figure 4.10 — Simulated diffraction efficiency of the coupling model at different FDTD simulation times: (a) 0^{th} order, (b) 1^{st} order.

4.2.4 Coupling resolution

In the previous simulations (Table 4.3, 4.4 and Fig. 4.10), the FDTD spatial sampling grid was used directly for the free-space propagator, i.e. $d_1 = \lambda/40$. Due to the

discrete Fourier transform which is used in the Harvey propagator, the angular spacing in the frequency domain will be relatively high: $d\alpha = \lambda/(Nd_1)$. As the absolute value of α should be smaller than 1, the effective sampling number in the frequency domain will be $N_e \approx 2/d\alpha = N/20$, which is much smaller than the number of samples used for the free-space propagator. In order to increase the effective sampling number and obtain finer sampling on the far-field observation space (i.e. smaller $d\alpha$ and higher N_e), it is necessary to increase the spatial sampling distance d_1. Therefore, the output field of the FDTD simulation should be interpolated on a sparser sampling grid before coupling into the free-space propagator. Table 4.5 lists the calculated diffraction efficiencies of the FDTD + Harvey model with different scaling factors s ($d_1 = s \cdot \lambda/40$) for the coupling interpolation step, where the linear interpolation technique was used. It can be seen that the FDTD calculated field can be interpolated to about a 10 time sparser sampling grid without significant variations in the simulated diffraction efficiencies.

Table 4.5 — Diffraction efficiency (%) of different diffraction orders of the grating using the FDTD + Harvey model at different scaling factors

Scaling	**Order**						
factor s	**0**	**±1**	**±2**	**±3**	**±4**	**±5**	**±6**
1	1.25	40.12	1.43	3.55	1.66	0.47	2.00
2	1.25	40.13	1.43	3.55	1.66	0.47	2.00
4	1.25	40.14	1.42	3.55	1.65	0.47	2.00
8	1.25	40.13	1.42	3.55	1.65	0.47	2.00
10	1.25	40.14	1.42	3.56	1.65	0.47	1.99
20	1.23	39.65	1.40	3.67	1.54	0.27	2.31

4.2.5 2D simulation and FDTD parallelization

2D simulation

In order to test our vectorial model for 2D problems, we simulated the diffraction pattern of the asymmetrical DOE at the feature size of $s = 500$ nm, and the illumination wavelength of $\lambda = 633$ nm. At this wavelength, the refractive index of the DOE material, which was fused silica, is $n \approx 1.457$ and therefore the optimum etching depth predicted by the TEA is $d \approx 692$ nm. The FDTD resolution and simulation time were chosen as 38 samples/wavelength and 100 optical cycles, respectively, and the 6 complex-valued electromagnetic components at the DOE surface were coupled into a far-field free-space propagator. As before, the FMM simulation was used for reference, which obtains the diffraction pattern on a hemisphere. For this reason, the Harvey model was used to propagate the 6 electromagnetic components separately to

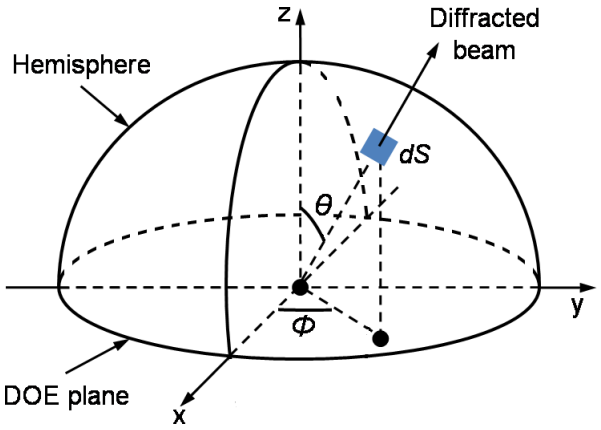


Figure 4.11 — The surface area of the diffraction order on the hemisphere.

the reference hemisphere. The diffraction power was then calculated from these far-field components using the Poynting vector: $P = |\vec{E} \times \vec{H}| \cdot dS$, where dS is the surface area of a certain diffraction order on the reference hemisphere, as shown in Fig. 4.11. The area element on the sphere in spherical coordinates is given by $dS \propto \sin\theta d\theta d\phi$, where $d\theta$ and $d\phi$ are the differentials of $\theta = \arccos\gamma$ and $\phi = \arctan(\beta/\alpha)$, respectively. Using the total differentials of $d\theta, d\phi$ with respect to the direction cosines $d\alpha, d\beta$ [126, 127] results in $dS \propto (d\alpha)^2/\gamma$.

Table 4.6 — Simulated diffraction efficiency (%) of diffraction orders (**m**, **n**) of the asymmetrical DOE using: (a) the FMM (LightTrans VirtualLab^{TM}) and (b) the FDTD + Harvey model.

(a)

m \ **n**	**-2**	**-1**	**0**	**1**	**2**
-2	2.82	2.69	2.56	2.05	3.34
-1	2.39	2.75	3.12	2.77	1.07
0	2.55	2.45	10.57	2.56	3.14
1	3.04	2.99	2.13	3.42	2.77
2	2.43	1.09	3.74	2.77	1.69

(b)

m \ **n**	**-2**	**-1**	**0**	**1**	**2**
-2	2.89	2.58	2.50	2.02	3.29
-1	2.43	2.76	3.09	2.80	1.02
0	2.59	2.50	10.30	2.58	3.02
1	3.16	3.01	2.10	3.52	2.84
2	2.55	1.06	3.85	2.69	1.71

From Table 4.6, we can see that there are differences between the diffraction efficiencies predicted by the FMM and our FDTD + Harvey method. Possible reasons for the differences in the results are:

1. Rounding errors of up to 1 pixel when discretizing the DOE structure in the lateral dimensions (x, y) on the FDTD sampling grid. For example, for the feature size of $s = 500$ nm $\lambda = 633$ nm at the sampling resolution of 38 pixels/wavelength, $s \approx 30.02$ pixels. These errors tend to affect the non-zero diffraction orders more

than the zero order [3], and had already been partly reduced in MEEP by using linear averaging [128, 129]. Particularly for this case we have chosen the sampling resolution of 38 samples/wavelength so that the errors in lateral dimensions is small, i.e. only about 0.02 pixel.

2. Similarly, rounding errors up to 1 pixel when discretizing the etching depth of the structure, i.e. $d \approx 692$ nm ≈ 41.54 pixels, and the undesired reflection of the electromagnetic field at the absorbing layers [130, 131]. These errors usually vary the diffraction efficiency of the zero order more than the non-zero orders [132].

3. By running the FDTD simulation for a finite number of optical cycles, we have neglected the contribution of the spherical waves generated by distant DOE periods (see Section 4.2.3). This is similar to the cutting off of higher order Floquet functions in the expansion of the FMM simulation [68].

4. Perhaps most importantly, the errors due to the direct coupling of the FDTD output field to the free-space propagator. This error can be reduced by using more advanced near-field to far-field transformations [133, 134].

Despite these errors, the calculated Hermitian symmetry factors are very similar, which are about 0.74 for both the FMM and our method. Experimental results of a DOE fabricated using Electron Beam Lithography, which will be given in Chapter 5, show that, although the measurement diffraction efficiencies differ from both the simulated results, the experimental Hermitian symmetry factor is in good agreement with the simulated values.

As expected, the calculation effort of the FMM is much lower than the FDTD + Harvey model: the FMM simulation required 0.2 GByte of RAM and 7 seconds in comparison with 4.1 GByte of RAM and 106 minutes for the case of the FDTD + Harvey simulation. The calculation times were obtained on the same computer model as before (i.e. PCs with an Intel core i5 @ 3.2 GHz processor). Thus, for infinite periodic structures, the FMM should be used, whereas for finite structures, FDTD + Harvey simulation can be used at the cost of higher memory usage and calculation time. To model and design bigger structures in a reasonable calculation time, it is therefore necessary to parallelize the FDTD simulation on a super-computer.

Parallelization on super-computer

As the FDTD simulation requires high memory usage and long calculation time, for modeling bigger structures, it is critical to parallelize the algorithm. MEEP supports distributed-memory parallelism using Message Passing Interface (MPI). For the parallelization of our vectorial model, we had access to a super-computer at the Ifremer

(Institut français de recherche pour l'exploitation de la mer: French research institute for exploitation of the sea) [135]. The super-computer is a "cluster" of 294 computing nodes, where 256 nodes contain two Intel Xeon quad-core @ 2.8 GHz processors and 38 nodes contain two Intel Xeon quad-core @ 3.46 GHz processors. Each computing node has 24 GByte of RAM, and the nodes are inter-connected by a network, resulting in the global memory of 7 TByte.

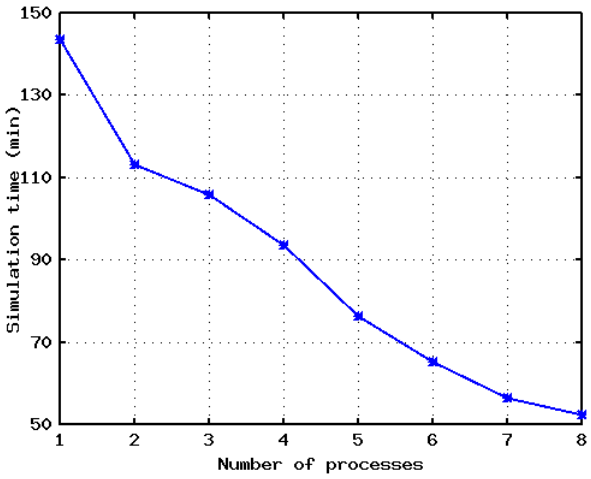


Figure 4.12 — Effect of super-computer parallelization on the simulation time.

In order to parallelize our rigorous model, it is necessary to build the parallel version of MEEP on the super-computer following the instruction [75]. Once MPI and MEEP have been installed, our rigorous model can be readily parallelized. Fig. 4.12 shows the calculation time of our previous FDTD + Harvey simulation using different number of parallel processes on the super-computer. Since the Harvey propagator is very computationally efficient, the simulation time is mostly governed by the FDTD method. It can be seen that the calculation time of the simulation using 8 parallel processes is only about 35% of that using a serial calculation, and may even be further reduced by using higher number of parallel processes. Bigger structures, which require higher memory usage can therefore be simulated and designed using our rigorous model on the super-computer.

4.2.6 Summary of our rigorous vectorial model

In summary, we have developed a rigorous vectorial diffraction model based on FDTD for the DOE region and the Harvey propagator for the free-space region. Simulations for 1D and 2D DOEs show that our model obtains closer results to those calculated by the FMM than the values predicted by the TEA + Harvey model. The diffracted field on the more useful far-field plane can be obtained directly from the field on the Harvey hemisphere by using our spherical wave projection. Parameters for the

FDTD simulation and the coupling have been optimized to reduce the memory usage and calculation time without decreasing the algorithm's accuracy. Parallelization of the algorithm on a super-computer has been performed to further reduce the calculation time to about 35% by using 8 parallel processes. Bigger DOE structures can therefore be simulated and designed using our rigorous model on the super-computer. We propose a genetic algorithm based on this vectorial model for the design of "thick" DOEs in the next section.

4.3 Genetic algorithm for the design of thick DOEs

It has been shown in vectorial simulations that, complementary elements [116] and additional 2π phase shift elements [136], as shown in Fig. 4.13, have different diffraction patterns, unlike the TEA prediction. Another report [137] demonstrated that the optimum etching depth obtained by an extended scalar theory is different from the value predicted by the TEA. Meanwhile, most of the current design algorithms only design DOEs in the lateral (x, y) dimensions without optimizing the etching depth. Fig. 4.14 illustrates an algorithm based on our rigorous vectorial model for the design of binary thick DOEs. As the FDTD propagation is uni-directional, Genetic Algorithms (GA) [138, 139] appear to be a good option for optimizing both lateral and thickness structures, as they are iteratively optimized algorithms applicable to large DOEs and they don't require back-propagation. Direct Binary Search and Simulated Annealing algorithms can also be used but they often converge more slowly, as already addressed in Chapter 1.

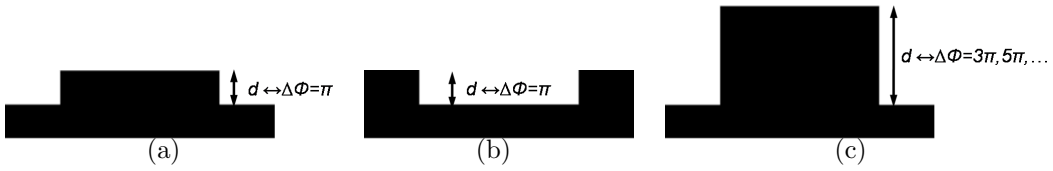


Figure 4.13 — (a) A binary phase element. (b) Complementary element. (c) Additional 2π phase shift elements.

The algorithm can be described as follows. Firstly, the DOE population is initialized. Each DOE is represented by a real-valued $N \times N$ matrix and an etching depth, which are generated by a sequence of random numbers. This population of the algorithm consists of N_p DOEs. Our rigorous vectorial model (where the Harvey model is replaced by our free-space propagator) is then used to obtain the diffraction patterns of the DOE population on the far-field plane. These patterns will be compared with the desired image and the DOE population is ranked based on some performance constraints. A number of the best performed DOEs is kept for the next population. They

will also be randomly selected for "mutation" to create "offsprings". These new DOE structures will replace the DOEs having low performance in the previous population. In order to maintain the population diversity N_p to avoid stagnation, a few randomly generated DOEs will also be created in the next population. The process is iterated until the performance criteria has been satisfied, or the algorithm has been run for a large number of iterations without significantly improving the best DOE performance (i.e. the algorithm has reached an optimum). Due to the lack of time, calibration, optimization of the design and fabrication of thick DOEs were not completed before the end of this thesis. They will form the starting points of continuing investigations by the Optics Department at Télécom Bretagne in this field.

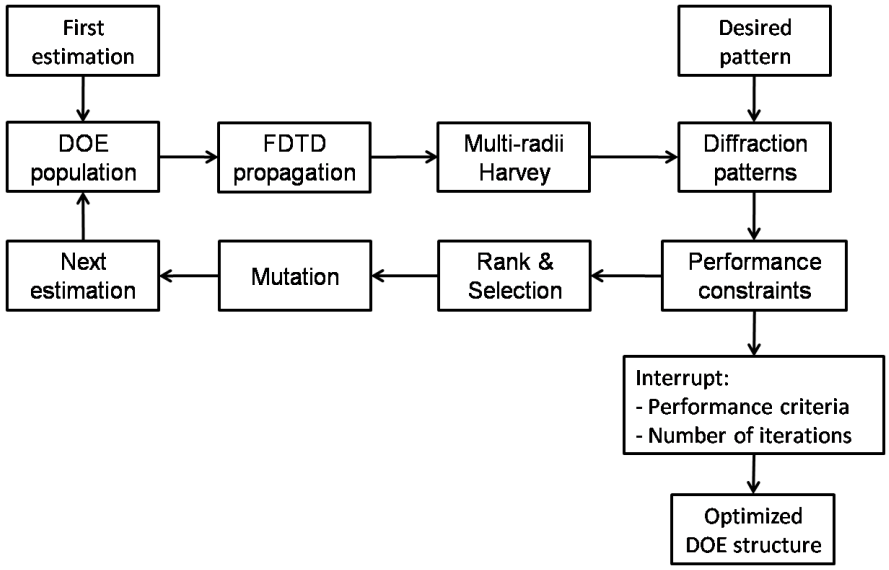


Figure 4.14 — Genetic algorithm for the design of binary thick DOEs.

4.4 Conclusion

In conclusion, we have investigated the practical limits of the TEA in predicting the symmetry of the diffraction pattern by binary phase DOEs by comparison with the rigorous vectorial simulations. The results showed that, strictly speaking, the TEA should not be used for binary DOEs having feature sizes smaller than about 2 times the wavelength, as the error in predicting the Hermitian symmetry factor is more than about 10%. For designing wide-angle pattern generation DOEs, the TEA + our scalar non-paraxial propagator may be sufficient, but for designing high uniformity spot array DOEs, it is necessary to take into account vectorial effects in the DOE itself and its close surroundings. We have developed, optimized and parallelized a rigorous diffraction model based on FDTD coupled with a scalar free-space propagator to overcome the theoretical limits of the TEA and the computational limitations of current vecto-

rial models. A genetic design algorithm based on this model has been proposed for the design of thick DOEs. In addition, our metric for evaluating the accuracy of the simulations, which is Hermitian diffraction symmetry of binary elements, has been shown to be simple yet efficient and consistent, despite the variations in the simulated and experimental diffraction efficiencies. The experimental verification of different modeling and design algorithms and the effects of fabrication errors to the measured diffraction efficiencies will be presented in the next chapter.

CHAPTER 5

Experimental verification and effects of fabrication errors on the experimentally observed diffraction pattern

5.1 Introduction

In the previous chapters, we numerically analyzed the limits of the conventional diffraction models (i.e. the scalar paraxial approximation and the Thin Element Approximation), and developed scalar and vectorial algorithms to overcome the current limitations in the modeling and design of wide-angle diffraction Fourier DOEs. In this chapter, the limits of the TEA and scalar paraxial models, together with the accuracy of our scalar non-paraxial and vectorial simulations are verified experimentally by fabricating test DOEs and measuring their performance on an optical bench.

The fabrication facilities we used to fabricate DOEs for these experimental verifications were our direct-write photoplotter at Télécom Bretagne (TB) [52] for the fabrication of microstructure DOEs, and Electron Beam Lithography (EBL) at Karlsruhe Institute of Technology (KIT) in Germany [53] for the fabrication of submicron elements. The chapter therefore starts with a brief description of these fabrication processes. We then describe successively the experimental confirmation of our scalar non-paraxial model and design algorithms (presented in Chapter 3) through the fabrication and characterization of large diffraction angle Fourier DOEs. The remaining discrepancies between modeling, design and experiment (mainly spot power at large angles) are then investigated and shown to result from both fabrication errors and by the fact that we are approaching the limit of the TEA. We show that the symmetry of the diffraction spot power appears to be a convenient way of evaluating the effects of

fabrication errors and the practical limits of the TEA and is a promising direction of future study. We close the chapter with some example applications of DOEs fabricated for academic and industrial partners of our laboratory.

5.1.1 DOE fabrication at TB

The DOE fabrication process in our cleanroom at TB is described briefly below, full details can be found in [52]. The basic procedure and some images of our cleanroom equipment which were used in this thesis are illustrated in Fig. 5.1.

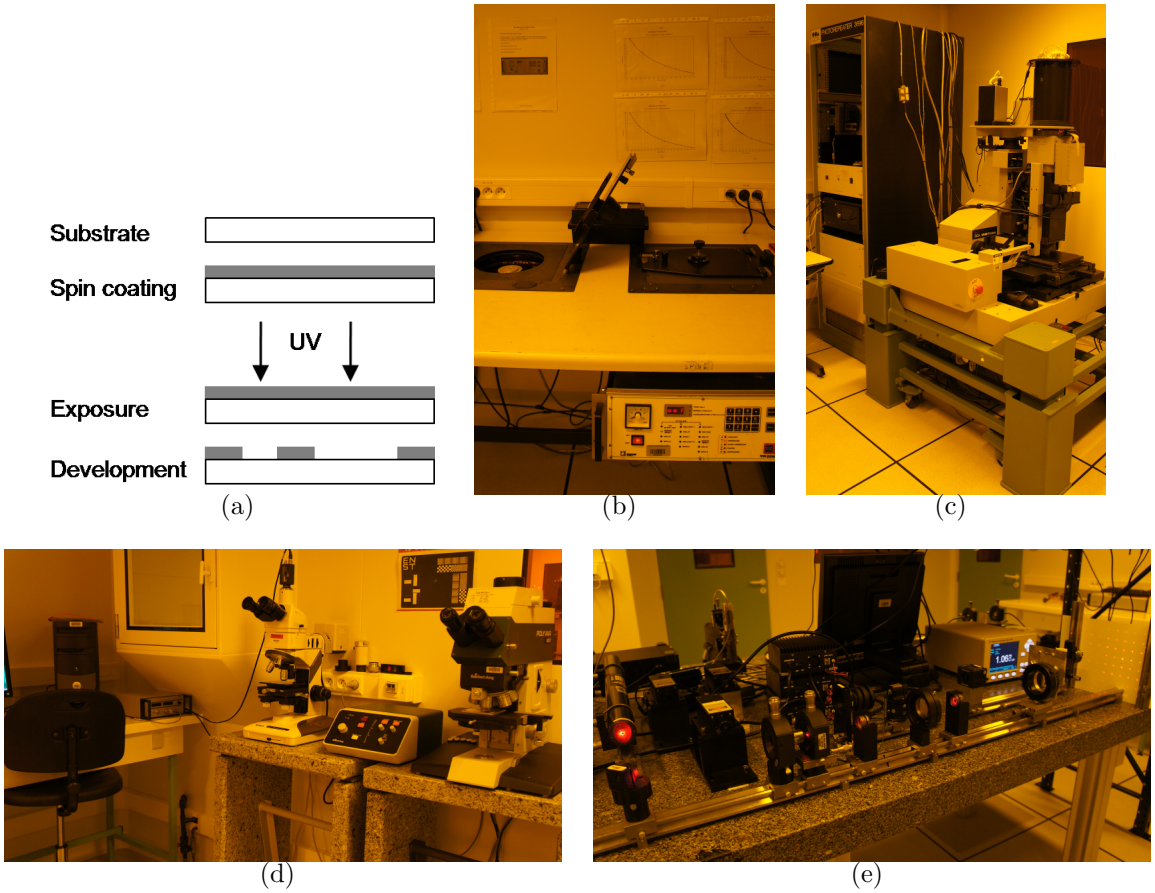


Figure 5.1 — (a) DOE fabrication procedure in the TB cleanroom. (b) Spin-coater. (c) Parallel direct-write photoplotter. (d) Interferometric microscope. (e) Optical setup for DOE characterization.

- Spin coating of a photoresist, which is usually the S1800 series from Micro Resist Technology, onto a clean glass substrate. The spin coating speed is chosen so that the thickness of the photoresist layer will be relatively uniform at the desired etching depth. The thickness variation of the photoresist layer in our spin-coating process is about 20 nm for layer of about 500 nm to a few μm.

- Exposing the photoresist layer using our parallel direct-write photoplotter. The exposed pattern is mapped from the designed DOE image file and controlled via a programmable Spatial Light Modulator. The exposure time can be varied using a mechanical shutter to ensure the optimum etching depth.

- Developing the exposed substrate in a chemical solution (Microposit 303A Developer) to etch the exposed pattern into the photoresist layer.

- Measuring the DOE structure under our 3D interferometric microscope and its performance on an optical bench. The resolution limits of our microscope in measuring the lateral dimensions and the etching depth are about 1 μm and 10 nm, respectively. The schematic diagram of the optical bench is shown in Fig. 5.3(b).

5.1.2 DOE fabrication at KIT

As the resolution limit of our fabrication facilities is about 1 μm, during my PhD we also collaborated with the KIT for the fabrication of nanoscale DOEs. A funding was obtained for me to stay at the KIT in 2 months for the use of high-precision high-resolution Electron Beam Lithography (EBL). The DOE fabrication process and some images of the KIT cleanroom equipment which were used in this thesis are illustrated in Fig. 5.2, the basic procedure is as follows:

- Plasma sputtering of a thin (about 30 nm) Chromium layer on a clean fused silica substrate. This conducting layer is to avoid fabrication errors due to electric charging which occurs during the writing process.

- Spin coating of a photoresist, which is Poly-methyl methacrylate (PMMA), on the substrate. The thickness of this photoresist layer is less important in this case, as the DOE structure will be etched into the substrate later.

- Exposing the substrate using the EBL. This system is able to write structural details down to about 20 nm in PMMA.

- Developing the exposed substrate using a chemical solution. As PMMA is a negative photoresist, the unexposed areas are removed.

- Etching the developed pattern into the Chromium layer and into the substrate using Reactive Ion Etching (RIE). The resolution limit of the RIE system in the lateral dimensions is about 100 nm.

- Removing the undeveloped photoresist and the unetched Chromium layer.

- Measuring the DOE structure under a 3D Atomic Force Microscope (AFM). The lateral resolution of the AFM depends on the radius of the tip which is between 10 - 20 nm, and the vertical resolution is typically better than 1 nm [53].

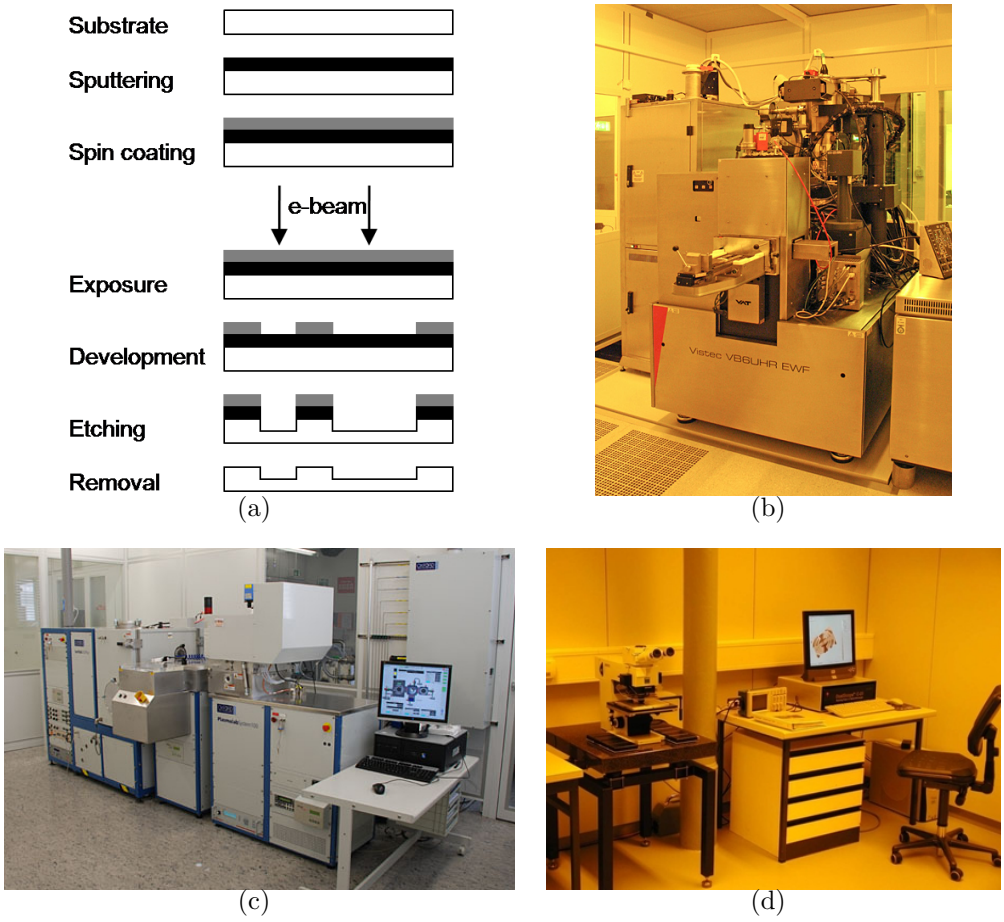


Figure 5.2 — (a) DOE fabrication procedure using EBL. (b) Electron beam writer. (c) RIE cluster, including a Oxford RIE Plasmalab System 100. (d) An AFM. (Photo courtesy of the KIT)

5.2 Experimental verification of the scalar non-paraxial model and design algorithms

5.2.1 Scalar non-paraxial modeling

Test DOE fabrication

To investigate the accuracy of our scalar non-paraxial propagator in predicting the experimental non-paraxial diffraction patterns, we deliberately selected a Fourier DOE

(generating a 5×5 array of spots) with non-paraxial diffraction angles which, as a result, was close to the resolution limit (smallest structures of about 1 μm) of our fabrication facilities and therefore difficult to fabricate. Under these conditions, the elementary period of the binary phase DOE necessarily contained a limited number of pixels (8×8) resulting in an output pattern of relatively low uniformity, as shown in Table 5.1. To accentuate the non-paraxial diffraction, the DOE was designed for operation at a wavelength of 1550 nm. The binary DOE image file was sent to our photoplotter for fabrication [52]. The basic DOE structure was replicated on a regular periodic grid by the photoplotter to give a total fabricated DOE size close to 5×5 mm^2 to avoid speckle effects in the experimental reconstruction [58]. Several DOEs were made at slightly different exposure times to optimize the etch depth and micro-structure critical dimension and compare the influence of fabrication limitation induced variations in diffraction efficiency with the variations resulting from the different diffraction models.

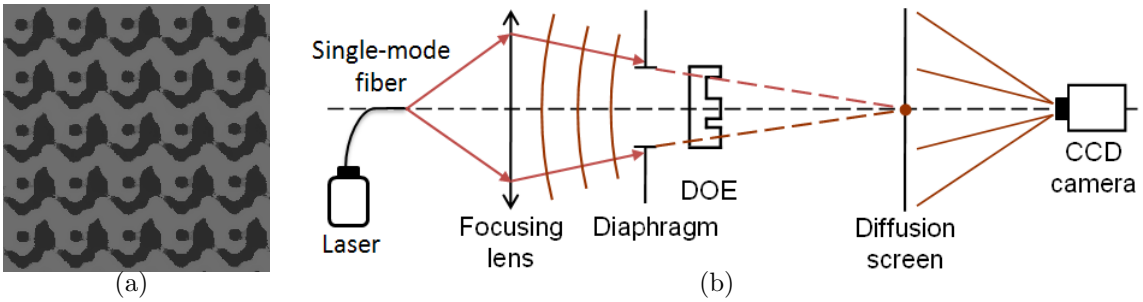


Figure 5.3 — (a) Optical interferometric microscope image of the fabricated DOE, where a period is about 8 μm. (b) Optical DOE playback setup.

Diffraction spot position

For the experimental verification, a single mode fiber was used to guide light from the laser source to the optical setup, as depicted in Fig. 5.3(b). Conforming to our model and simulations (section 3.1), the beam was then converged by a lens into a spherical beam focusing onto the observation plane, which was a diffusion screen parallel to the DOE plane and perpendicular to the optical axis. A diaphragm was used after the lens to limit the illumination to the DOE area. The diffraction pattern on the screen was captured using an InGaAs camera, whereas the spot power was measured using a Newport photodetector.

As shown in Fig. 5.4, the experimentally observed diffraction is very close to the simulated pattern, where both the parabolic distribution and the large diffraction angle broadening effect on the spots shape are clearly visible. A slight deviation in the spot positions can be seen in the superposition of the simulated and experimental diffraction

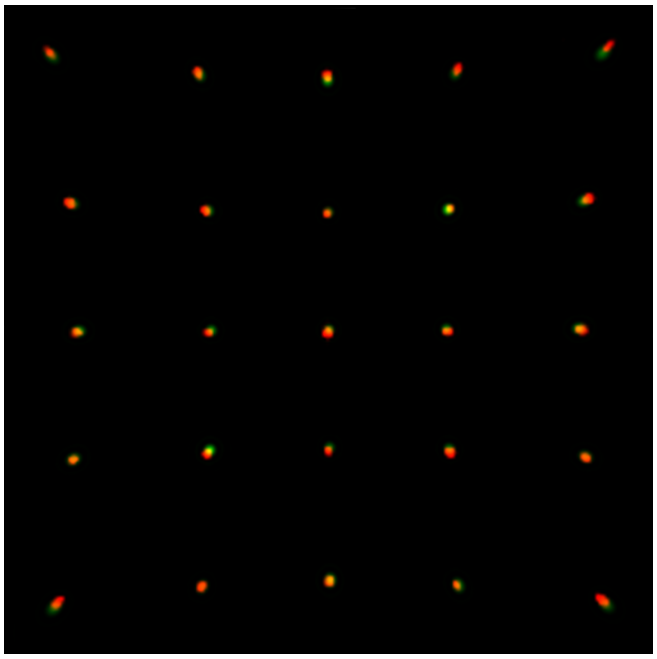


Figure 5.4 — Superposition of the simulated and experimental diffraction patterns. The green spots are those predicted by the multiradii Harvey calculation, whereas the red spots are those of the experimentally observed pattern. The yellow regions are where they overlap.

patterns, probably due to small errors in the relative angles between the DOE plane, the output plane, the camera and the optical axis.

Diffraction spot power

Interpretation of the spots experimental diffraction efficiency is more complex. Since the DOEs contain microstructures at the limit of our fabrication facilities (minimum feature size of 1 μm), as shown in Table 5.1, the spot diffraction efficiency strongly depends on the fabrication limitations (e.g. etching depth, pixel rounding, linewidth), which vary with the exact fabrication parameters - here illustrated by the pattern exposure time. To investigate the effect of such fabrication errors on the spots diffraction efficiency, we simulated the effect of a slight dilation or erosion in the microstructure critical dimension due to under or over exposure, as shown in Fig. 5.5. By oversampling the DOE by a factor 8, a 1 pixel dilation/erosion is equivalent to a variation of $\Delta = 0.125$ μm in the etch width. The simulated results using the RS integral are given in Table 5.2.

Comparison of the reference RS simulation of the DOE with and without fabrication error given in Table 5.2 shows relative variations in diffraction efficiency of at least 4.7%. This variation is significantly higher than the 1% difference (shown in Chapter 3) between the predictions of our model and the RS method. This result is also con-

Table 5.1 — Experimentally observed diffraction efficiencies for sample diffraction orders of the test DOE with fabrication errors.

Method	Diff. eff. (%) of order *mn*				
	01	**11**	**02**	**12**	**22**
Experiment, underexposed DOE	3.59	2.22	2.14	3.47	1.92
Experiment, best DOE	3.64	2.52	2.05	3.57	1.88
Experiment, overexposed DOE	3.44	2.43	1.78	3.45	1.78

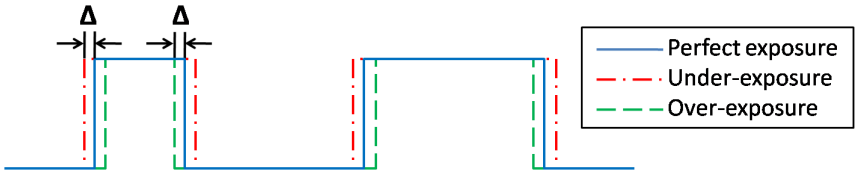


Figure 5.5 — The effect of exposure time to the etch profile and pixel linewidth.

firmed by the variations in the experimentally observed diffraction efficiencies in Table 5.1, which are very frequently significantly greater than 1%. This indicates that the accuracy of the multiradii approach in predicting spot diffraction efficiency is greater than the errors introduced by a DOE linewidth control of 0.125 μm, which is a fabrication accuracy that is already difficult to obtain with current fabrication techniques. In summary, our proposed method can be used to calculate accurately non-paraxial diffraction, both spot position and diffraction efficiency, with a negligible error (lower than that resulting from fabrication limitations) in predicting diffraction efficiency but with a very significantly lower calculation effort than the more rigorous RS integral.

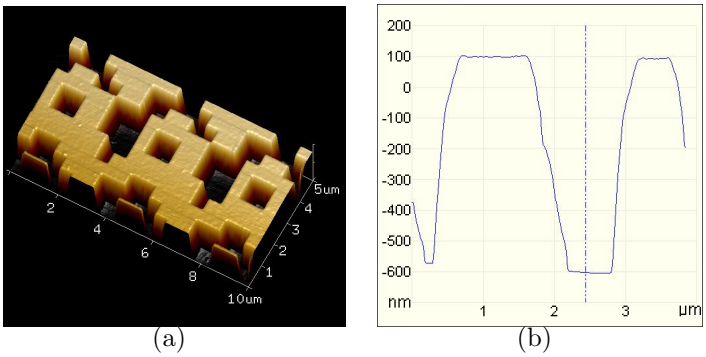
Table 5.2 — Simulated diffraction efficiencies for sample diffraction orders of the test DOE without and with fabrication errors.

Method	Diff. eff. (%) of order *mn*				
	01	**11**	**02**	**12**	**22**
Simulated RS integral, dilated structure	3.13	1.82	2.43	2.74	1.93
Simulated RS integral, perfect structure	2.73	2.32	2.55	2.56	2.05
Simulated RS integral, eroded structure	2.29	2.57	2.40	1.99	1.87

With these results, as we had reached the resolution limit of our own fabrication facilities, we collaborated with the KIT for the fabrication of nanoscale DOEs, which were designed using our iterative scalar non-paraxial algorithms.

5.2.2 Iterative scalar non-paraxial design

To verify the accuracy of our scalar non-paraxial algorithms experimentally, the designed DOEs in Chapter 3.2 at pixel sizes $\delta_1 = 400$ nm and 300 nm were fabricated using EBL at the KIT in Germany [53]. The basic DOE structures were replicated on a regular periodic grid to give a total fabricated DOE size close to 4×4 mm^2 to avoid speckle effects in experiment [49]. The structures were then etched into the fused silica substrate using RIE. At $\lambda = 633$ nm, the refractive index of the substrate is $n \approx 1.46$, which means that the etching depth should be about $d = \lambda/2(n-1) \approx 688$ nm to obtain a π phase shift between binary levels [3]. Fig. 5.6(a) shows an AFM image of a test DOE, with an etching depth of about 700 nm, as can be seen in Fig. 5.6(b).



(a) (b)

Figure 5.6 — (a) AFM image of the test DOE. (b) Etching profile of the test DOE, where the edges look slanted due to the size of the tip used for AFM measurement [53].

The fabricated DOEs were then tested in an optical setup as in the previous section, but with $\lambda = 633$ nm. The experimentally observed patterns of the DOEs at $\delta_1 = 400$ nm are shown in Fig. 5.7, respectively. As is common, a bright central spot is observed due to the etch-depth fabrication errors. In particular, Fig. 5.7(a) confirms the errors in diffraction position and diffraction power of the standard paraxial IFTA, whereas Fig. 5.7(b) - Fig. 5.7(d) confirm the accuracy of our non-paraxial design algorithms, which correctly allow for the geometrical distortions of the output pattern resulting from non-paraxial diffraction, as analyzed in Chapter 3.

A closer look on the experimental diffraction pattern reveals the error of up to 1 pixel in diffraction position of the single projection design using nearest-neighbor interpolation, as shown in Fig. 5.8(a). Fig. 5.8(b) confirms that the iterative projection with bicubic interpolation algorithm produces a better output pattern in diffraction position, although some "ringing", which is due to the bicubic interpolation [94], is still visible, as predicted in the simulation.

Similarly, Fig. 5.9 shows the experimentally observed patterns of the DOEs at $\delta_1 =$ 300 nm. As predicted by our scalar non-paraxial propagator, the diffraction angles is bigger (about 48° at the corners of the pattern) and the distortions in the reconstruc-

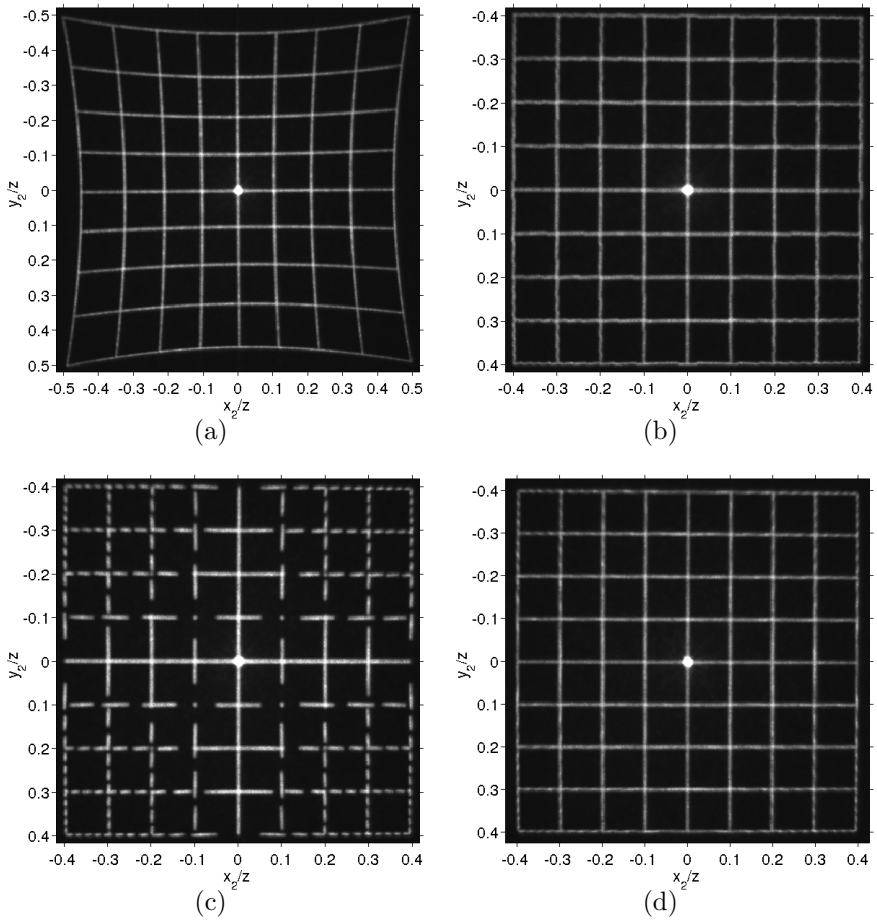


Figure 5.7 — Experimental patterns of the DOEs designed at $\delta_1 = 400$ nm using: (a) Standard paraxial IFTA, (b) Non-paraxial IFTA with single projection + nearest neighbor interpolation, (c) Non-paraxial IFTA with single projection + bicubic interpolation, (d) Non-paraxial IFTA with iterative projection + bicubic interpolation.

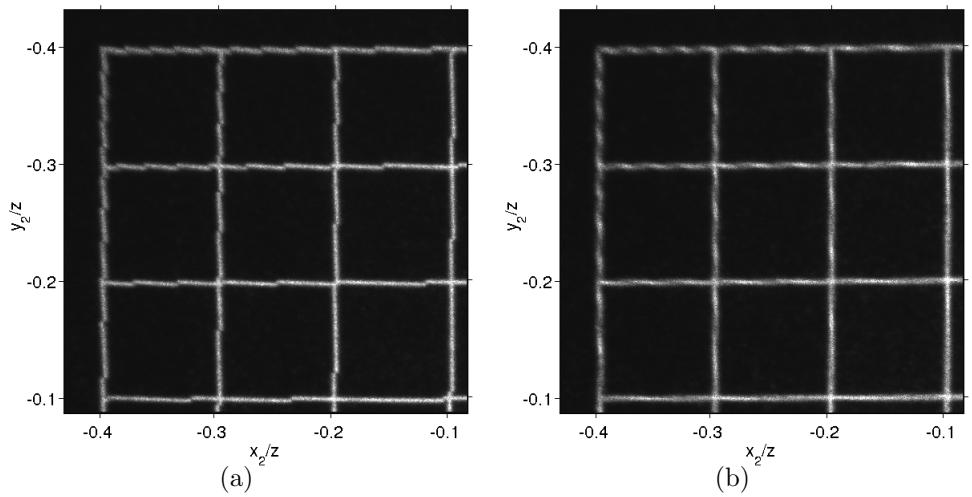


Figure 5.8 — (a) Zoomed image of the top left corner of the pattern in 5.7(b). (b) Zoomed image of the top left corner of the pattern in 5.7(d).

tion of the DOE designed using standard IFTA become stronger. This demonstrates that our scalar non-paraxial propagator can be used to model diffraction angles up to 48°, and perhaps even higher, within the limitations of the TEA. Fig. 5.9(b) shows the reconstruction of the DOE using the iterative projection algorithm with bicubic interpolation, where the maximal diffraction angle is about 37° at the corners of the grid pattern, which is very close to the target image.

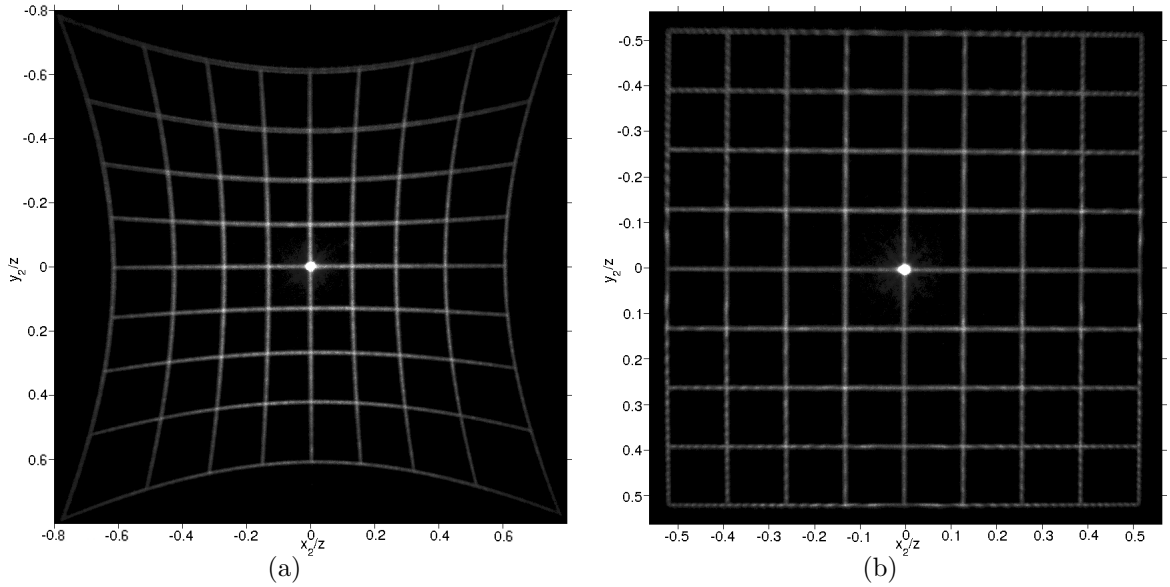


Figure 5.9 — Experimentally observed patterns of the 300 nm feature size DOEs designed using: (a) Standard paraxial IFTA, (b) Non-paraxial IFTA with iterative projection + bicubic interpolation.

These results experimentally confirm the applicability and usefulness of our scalar non-paraxial diffraction model and design algorithms based on this model. However, while our scalar non-paraxial model and design algorithms correctly predict pattern shape, the experimental results frequently show a relatively strong zero order whereas the algorithms show no such strong zero order. These bright zero orders in the experimentally observed patterns seem to be partly due to fabrication errors, which are inevitable even for a high-cost EBL fabrication process used here. The smaller the feature size to be fabricated, the stronger the effects of fabrication errors, which usually result in a brighter un-diffracted central spot.

However, there is another possible explanation of the origin of these strong zero orders: the etch depth of these DOEs is already about 2 times the pixel size, so the TEA used as a basis for our scalar non-paraxial model may not model correctly the interaction between the illumination field and the DOE structures. Although our scalar non-paraxial propagator is rigorous for the free-space region, the diffracted field immediately after the DOEs has been calculated assuming perfect DOE structures and the TEA. For designing spot array DOEs where high uniformity spot power is required,

the effects of fabrication errors and the validity of the TEA to the far-field diffraction power become more important. The effects of fabrication errors and the limit of the TEA are addressed in the next section.

5.3 Effects of fabrication errors and the limit of the TEA on the experimental diffraction symmetry

When building spot array binary Fourier DOEs having feature sizes on the order of the wavelength and close to the fabrication limit of our fabrication facilities, we noticed remarkable variations in the experimental diffraction efficiency compared to the simulation results. Even with the use of a high-cost EBL and the rigorous FMM/FDTD simulations, there appear to be no publications, to our knowledge, showing the close agreement in diffraction efficiency between simulation and experimental results. However, in what follows we show that our preliminary investigation seems to indicate that the diffraction symmetry of binary Fourier DOEs can be used as an efficient and consistant metric for evaluating the effects of fabrication errors and the limit of the TEA.

5.3.1 Effects of fabrication errors

It is widely known that the performance of DOEs strongly depends on the accuracy of manufacturing process, where fabrication errors are inevitable, e.g. etch-depth, feature-size [3,140]. These errors redistribute the diffraction power between the diffraction orders, and therefore the performance of the fabricated DOEs is often degraded compared to the designed elements. It has been determined [141] that an etch-depth error creating a phase error of magnitude ϵ causes the intensities of the non-zero orders to be reduced by $(1+\cos\epsilon)/2$. This reduction is accompanied by a large increase in the zero order, so even a small phase error results in a bright zero order.

For example, assuming that we have designed a perfectly uniform 25×25 binary spot array DOE, where each spot has diffraction efficiency of η. Due to a 5% etch-depth error in fabrication (which is equivalent to about 25 nm for a DOE fabricated using photoresist S1805 at 633 nm), the non-zero orders are reduced by a factor $(1+\cos(0.05\pi))/2 \approx 0.994$. The experimental diffraction efficiency of non-zero orders are therefore 0.994η, which is still close to the design value. However, the diffraction efficiency of the zero order is increased to about $(1+624\cdot 0.006)\eta \approx 4.7\eta$, which is almost 5 times that of the non-zero orders. This etch-depth error is therefore critical for applications where uniformity of the spot array is important, as will be shown in Section 5.4. Jahns *et al.* [141] concluded that Dammann gratings are useful for array

sizes up to about 40×40 spots. Above that size, problems with the computation of the DOEs and with the fabrication resolution become dominant.

Knowledge of these fabrication errors for the zero order has been used for pre-compensation in the design and fabrication process to increase the performance of the fabricated DOEs [132, 142, 143]. However, the impact of fabrication errors on the non-zero orders are less studied and understood. In Section 5.2.1, we showed the effect of error in the lateral critical dimensions of the etched DOE structures on the experimental spots' diffraction efficiency. Although this fabrication error can change the diffraction efficiency of non-zero orders remarkably, within the TEA, it does not vary the diffraction symmetry of binary Fourier DOEs (the same is true with etch-depth error). In this section, we demonstrate that the diffraction symmetry of binary Fourier elements can in practice be modified due to pixel rounding effect. The limit of the TEA in predicting the practical diffraction symmetry is addressed in the next section.

Fig. 5.10(a) shows a reflection symmetric binary phase Fourier DOE, which was designed to generate a 5×5 spot array as illustrated in Fig. 5.10(c). The experimental diffraction efficiency in the different spots by this DOE are shown in Table 5.3(a), where the DOE was fabricated with a pixel size of 1 μm, which is the limit of our direct-write lithography [52], and the DOE illumination wavelength was 1.55 μm. As the designed DOE was binary, we expected Hermitian symmetry [11], but in fact observed reflection symmetry in the experimental diffraction pattern. For instance, based on the TEA, the diffraction efficiency of order (-2, 1) should be the same as that of order (2, -1), but in fact, they are very different from each other (i.e. 0.76% and 2.94%, respectively). Instead, the diffraction efficiencies of order (-2, 1) and order (-1, 2) are very similar (i.e. 0.76% and 0.77%, respectively). Further analysis shows, within experimental errors, a reflection symmetry of the diffraction pattern in the same axis as the reflection symmetry of the designed DOE, which is the dashed line shown in Fig. 5.10(a).

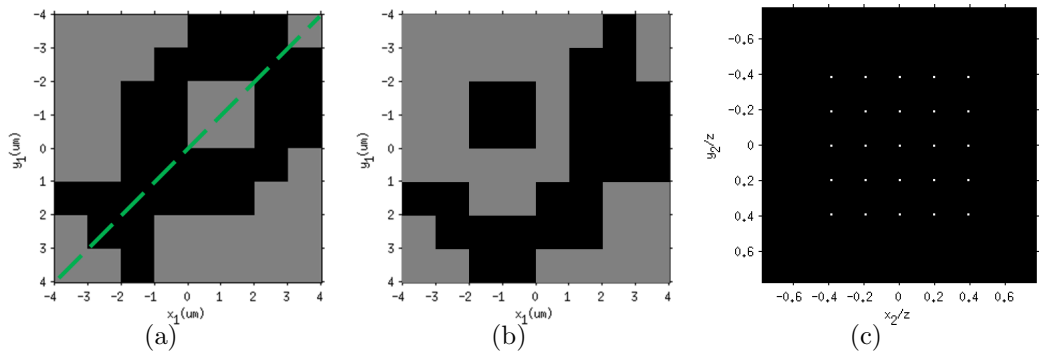


Figure 5.10 — (a) A symmetrical binary Fourier DOE, where the dashed line is the reflection symmetry axis. (b) An asymmetrical binary Fourier DOE. (c) Target image.

Table 5.3 — Diffraction efficiency (%) of diffraction orders (**m**, **n**) for the symmetrical binary DOE: (a) Experimental diffraction efficiencies of the diffraction orders for the DOE fabricated using our photoplotter. (b) TEA + our scalar non-paraxial simulation, assuming flawless binary DOE structure.

(a)

m \ **n**	**-2**	**-1**	**0**	**1**	**2**
-2	1.52	3.93	2.42	0.76	1.76
-1	2.18	5.20	4.19	3.92	0.77
0	2.48	4.18	2.38	4.00	2.62
1	2.89	1.03	4.18	5.10	4.00
2	0.36	2.94	2.38	2.14	1.43

(b)

m \ **n**	**-2**	**-1**	**0**	**1**	**2**
-2	1.67	2.94	2.60	2.08	0.97
-1	2.94	3.69	2.87	2.02	2.08
0	2.58	2.86	0.32	2.86	2.58
1	2.08	2.02	2.87	3.69	2.94
2	0.97	2.08	2.60	2.94	1.67

To investigate this variation in diffraction symmetry, we used the TEA where the field at the DOE plane is calculated as the multiplication of the illumination field and the transmission function of the DOE [11]. The far-field diffraction pattern is then calculated using our scalar non-paraxial propagation [107], which is equivalent to but much faster than the Rayleigh-Sommerfeld diffraction integral, and has been shown to be valid for far-field diffraction of DOEs having features smaller than the illumination wavelength. Since our diffraction model is based on a single Fourier Transform (FT), the symmetry of the calculated diffraction pattern depends on the symmetry of the FT. As the transmission function of a binary DOE is a real function, the diffraction pattern should have Hermitian symmetry ($f(x,y) = f^*(x,y) \rightarrow F(X,Y) = F^*(-X,-Y)$, where $F(X,Y)$ is the FT of $f(x,y)$ and the symbol * denotes the complex conjugate) [11]. Moreover, if the DOE has reflection symmetry, the calculated diffraction pattern should have the same reflection symmetry ($f(x,y) = f(-y,-x) \rightarrow F(X,Y) = F(-Y,-X)$) [125].

However, due to the fabrication errors at the limit of our fabrication facilities, the edge between phase levels of the fabricated DOE is not perfectly sharp but rather smoothed, as illustrated in Fig. 5.11. The fabricated DOE becomes a multilevel element, the transmission function is no longer real and therefore the diffraction pattern is no longer Hermitian symmetric (Note that this is a slightly more sophisticated modeling of the fabrication errors than that used in Section 5.2.1 where the modeled DOE were still binary phase - only the critical dimension changed but no intermediate phase levels were introduced). If the optical setup of the lithography is perfectly aligned, the fabrication errors will be symmetric and the fabricated DOE would still retain the symmetry of the designed DOE, as would the experimental diffraction pattern. This smoothing effect appears to be the reason for the variation in diffraction symmetry of the fabricated DOE given in the previous section. In an attempt to verify this

hypothesis, we tried to take this fabrication error into account in the calculation of spots' diffraction efficiency by using the TEA + our scalar non-paraxial simulation of the smoothed DOE structures.

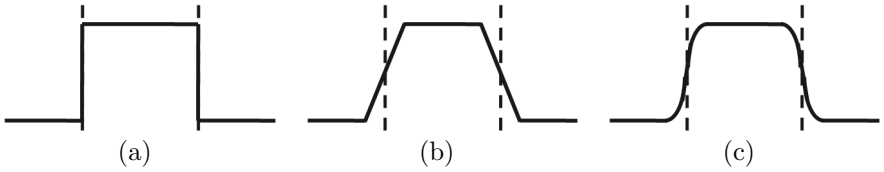


Figure 5.11 — Different models for the DOE structure: (a) Flawless model, (b) Trapezoid model, (c) Rounding model.

To investigate this smoothing effect on diffraction efficiency in simulation, we oversampled the DOE by a factor of 8 and applied different smoothing filters, i.e. averaging filter and Gaussian filter [94], of the same size to the DOE image. These filters correspond to two different models for the proximity effect in fabrication, i.e. trapezoidal [136] and rounding model [54], as shown in Fig. 5.11. The far-field diffraction patterns of these structures were then reconstructed using our scalar non-paraxial model [107]. The calculated diffraction efficiencies of different models were compared with the experimental results in terms of mean-squared-error (MSE).

$$MSE = \frac{1}{25} \sum_{\mathbf{m}=-2}^{2} \sum_{\mathbf{n}=-2}^{2} \left[\eta_{exp}(\mathbf{m}, \mathbf{n}) - \eta_{mod}(\mathbf{m}, \mathbf{n})\right]^2 \tag{5.1}$$

where ($\mathbf{m}$, $\mathbf{n}$) stand for the diffraction orders. η_{exp} is the experimental diffraction efficiency and η_{mod} is the simulation result obtained using different models for the fabricated DOE structure.

As expected, the diffraction spots in the reconstruction of the flawless structure showed both Hermitian and reflection symmetries, as listed in Table 5.3(b), but this is not the case with the experimental results (i.e. with reflection symmetry but without Hermitian symmetry). Although both trapezoidal and rounding models predict correctly the experimental diffraction symmetry, the latter shows better agreement with experimental diffraction efficiencies by a lower MSE value, as given in Table 5.4. This can be explained by the Gaussian shape of the write beam and the nonlinear dissolution rate of the photoresist [144], leading to a rounded transition between phase levels, where trapezoidal shape is only a first-order approximation [136].

From the above discussion, we have shown that pixel rounding effect can strongly affect the diffraction symmetry of binary Fourier DOEs fabricated close to the resolution limit of our fabrication facilities, simply by using the TEA + our scalar non-paraxial simulation of several smoothed DOE structures. On the other hand, it was demonstrated using vectorial simulations in Chapter 4.1.2 that perfectly binary Fourier DOEs

Table 5.4 — MSE between the experimental diffraction efficiencies and the TEA + our scalar non-paraxial simulation results of different DOE structure models.

Model	Filter	Output MSE
Flawless	None	1.09
Trapezoidal	Averaging	0.67
Rounding	Gaussian	0.52

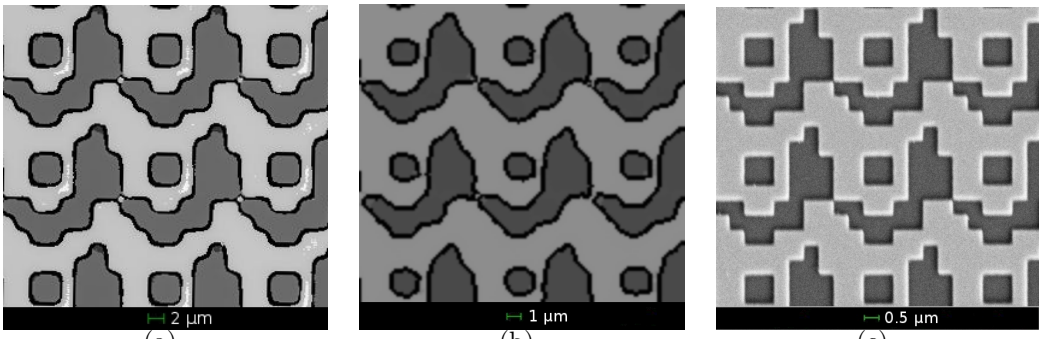
having feature sizes on the order of the wavelength do not necessarily have Hermitian symmetric diffraction pattern as predicted by the TEA. The change in symmetry of the observed diffraction patterns could therefore also be attributed to the fact that we are close to the limit of the TEA. We therefore investigated this possibility through vectorial simulations and experimental verifications. The limit of the TEA and the accuracy of our vectorial simulations are verified experimentally in the next section.

5.3.2 Limit of the TEA

In order to verify the limit of the TEA and the accuracy of our vectorial simulations, we fabricated a binary DOE, as shown in Fig. 5.10(b), using EBL. The DOE was designed at the feature size of 500 nm and the wavelength of 633 nm to generate a 5×5 spot array as illustrated in Fig. 5.10(c). Fig. 5.12 shows an AFM image of the DOE fabricated using EBL in comparison with those fabricated close to the limit of our photoplotter. It can be seen that there is a litte or no pixel rounding in the DOE fabricated using EBL, although other fabrication errors (e.g. etch-depth, feature size) may be present.

The fabricated DOE using EBL was then characterized on an optical bench and the diffraction efficiencies were measured using a Newport photodetector. It can be seen from Table 5.5(a) that the experimental diffraction efficiencies differ from both the FMM and our FDTD + Harvey simulations in Table 4.6 (the simulation results of our rigorous vectorial model are listed again in Table 5.5(b) for the reader's convenience). However, the calculated Hermitian symmetry factor (defined in Section 4.1.2) from the experimental data was about 0.74, which is very different from the TEA + Harvey prediction (i.e. 1) but in good agreement with the vectorial simulations' results (i.e. about 0.74 for both the FMM and our FDTD + Harvey method). This result demonstrates practically the limit of the TEA and verifies the accuracy of our rigorous vectorial model in predicting the Hermitian symmetry of the diffraction pattern of binary elements.

From this analysis, it appears that for the binary Fourier DOE having feature



(a) (b) (c)

Figure 5.12 — (a) and (b) Interferometric microscope images of the DOEs fabricated using our current photoplotter at 2 μm and 1 μm, respectively. (c) AFM image of the DOE fabricated at 500 nm using EBL at KIT.

Table 5.5 — Diffraction efficiency (%) of diffraction orders (**m**, **n**) for the asymmetric binary phase DOE. (a) Experimental results of the DOE fabricated using EBL. (b) FDTD + Harvey simulation results.

(a)

m \ **n**	**-2**	**-1**	**0**	**1**	**2**
-2	2.60	2.48	2.05	1.65	2.53
-1	2.30	3.18	2.98	2.28	1.00
0	2.78	2.10	22.88	3.25	2.70
1	2.43	2.93	1.80	3.33	2.70
2	2.45	0.93	3.30	2.15	1.83

(b)

m \ **n**	**-2**	**-1**	**0**	**1**	**2**
-2	2.89	2.58	2.50	2.02	3.29
-1	2.43	2.76	3.09	2.80	1.02
0	2.59	2.50	10.30	2.58	3.02
1	3.16	3.01	2.10	3.52	2.84
2	2.55	1.06	3.85	2.69	1.71

sizes on the order of the wavelength and fabricated with little or no pixel rounding effect, the TEA fails to predict correctly the Hermitian symmetry of the experimental diffraction pattern. On the other hand, for the DOEs fabricated close to the limit of our fabrication facilities, as shown in Fig. 5.12, it may be questionable whether the TEA simulation error is more dominant than the effect of the pixel rounding. For example, the experimental results of the binary DOE in Fig. 4.4(a) fabricated at pixel sizes of 1 μm and 2 μm using our photoplotter show that the practical Hermitian symmetry factors are 0.55 and 0.70, respectively. The rigorous vectorial simulation results of flawless DOEs showed that the Hermitian symmetry factors of these DOEs should be 0.91 and 0.94, which suggests that the DOEs contain fabrication errors which reduce the experimental Hermitian symmetry by 0.36 and 0.24 from the vectorial simulation results, respectively. These fabrication errors are much bigger than the errors in the TEA + our scalar non-paraxial simulation results compared to the values predicted by vectorial simulations, which are only 0.09 and 0.06, respectively. This means that for our fabrication facilities, the variation in the practical diffraction efficiencies due

to fabrication errors, are usually much bigger than the variation due to the diffraction modeling.

In summary, we have investigated the limits of the TEA and the accuracy of our rigorous diffraction model experimentally by fabricating a test DOE using a high-cost photolithography system and comparing the measured Hermitian symmetry factor with the simulated results. For the fabrication of high-resolution, thin DOEs using a relatively low-cost system such as our photoplotter, which has been specifically adapted to rapid prototyping [52], the variation in the practical diffraction efficiencies due to fabrication errors are very frequently much bigger than the variation due to the diffraction modeling. With this understanding on the effects of fabrication errors, we have improved our fabrication process in the manufacturing of several DOEs for our academic and industrial partners.

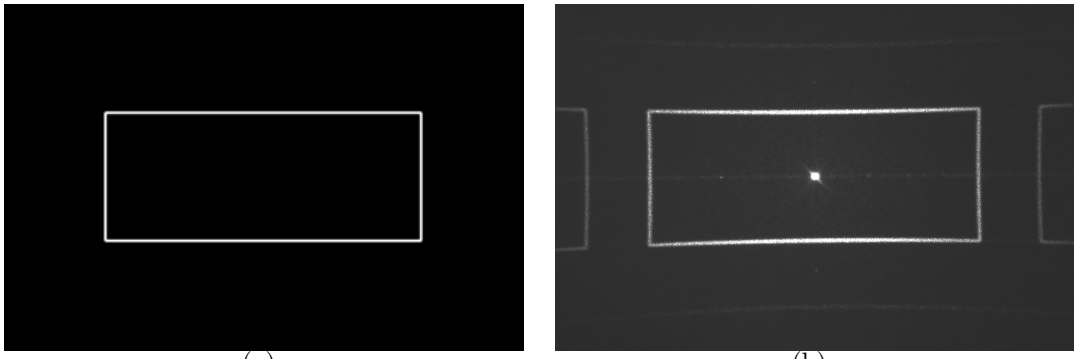
5.4 Example DOE applications

Télécom Bretagne has been designing and manufacturing DOEs for industrial and academic partners for several years. Before this PhD thesis was conducted, the diffraction angle had been limited to about 10°. We are now able to design and fabricate DOEs with much wider diffraction angles, the remainder of this chapter presents a few examples of DOEs that I helped design and fabricate for TB partners during my PhD.

5.4.1 Rectangular pattern

Fig. 5.13(a) shows a target rectangular box diffraction pattern with the maximum diffraction angle at the corners of about 17° that was requested by a microscope manufacturer based in Leuven, Belgium to be used for defining observed area. The experimentally observed pattern of the DOE designed using the standard paraxial IFTA and fabricated at 1 μm feature size is shown in Fig. 5.13(b), where second order rectangulars can be seen on the left and right hand sides of the first order. This effect is due to the discrete FT property which numerically reflects the higher diffraction orders outside of the calculation window back into the simulation region, similar to the Perfectly Conducting Boundary (see Appendix C). In other words, the FT assumes that the diffraction orders outside the calculation window are zero, which overestimates the diffraction orders in the simulation region. In practice, the high diffraction orders reduce the diffraction efficiency of the first order pattern compared to the simulation prediction and produce undesired noise into the optical metrology system.

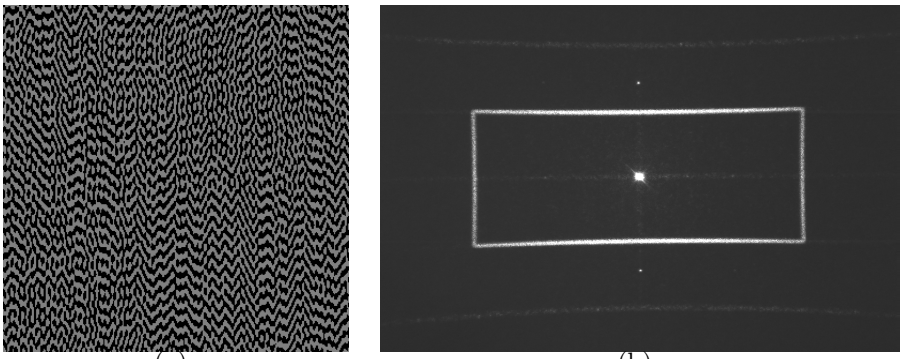
In order to obtain wide diffraction angles with thin DOEs, we have improved the resolution limit of our parallel photoplotter, so that feature size of 500 nm can now be



(a) (b)

Figure 5.13 — (a) Desired diffraction pattern. (b) Experimentally observed pattern of the DOE designed using the iterative scalar paraxial algorithm and fabricated at 1 μm feature size using our photoplotter.

fabricated. To minimize the effects of the pixel rounding effect, the DOE was designed by using the same iterative scalar paraxial algorithm, but with zero padding of the target pattern to reduce the number of isolated pixels, as shown in Fig. 5.14(a), so it is much easier to fabricate at higher resolutions [55]. The experimental result in Fig. 5.14(b) is much closer to the simulation prediction and significantly better than the pattern in Fig. 5.13(b), since less light is diffracted to higher diffraction orders as there are no individual pixels to diffract light at such angles [56].



(a) (b)

Figure 5.14 — (a) The DOE designed using the iterative scalar paraxial algorithm at 500 nm feature size. (b) Experimentally observed pattern of the DOE fabricated using our photoplotter.

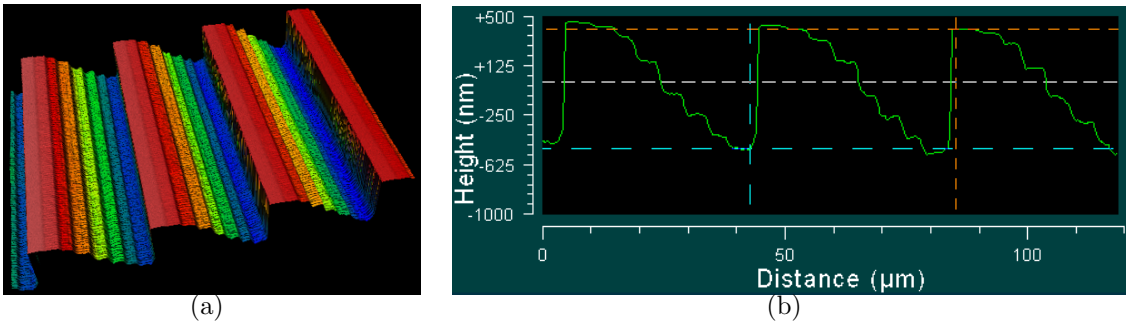
It can be seen that the diffraction patterns of both DOEs contain some distortion due to the use of standard iterative scalar algorithm, which is only valid for paraxial diffraction angle, as discussed in Chapter 3. Our iterative scalar non-paraxial design can be used to correct the distortions, as verified experimentally by the DOEs fabricated using high-resolution EBL at the KIT. However, for the initial tests, the industrial partner prefered cheaper rather than more accurate design and fabrication, as the

errors in the scalar paraxial design in this case is only a few %, which is acceptable for their requirements. If the tests are successful, wider diffraction angles will be proposed and our scalar non-paraxial design shall be used to correct the distortions.

5.4.2 Multilevel DOEs

During this thesis period, we also collaborated with Télécom Physique Strasbourg (TPS) for the fabrication of DOEs for solar cells. The DOEs were designed by TPS using a genetic algorithm based on a FDTD + ASM model, taking into account a broad wavelength range in the solar spectrum [145]. Due to the extensive calculation required, the design was limited to 1D phase elements. In order to increase the practical diffraction efficiency taking into account fabrication constraints, the DOEs were calculated as 8 level elements. We fabricated the DOEs in our department's cleanroom, whereas the microscope measurement and optical characterization were carried out by our academic partner at TPS.

Fig. 5.15 shows interferometric microscope images of a test DOE, which is a 8 level blazed grating, using a Zygo profilometer at TPS [146]. The etching profile of this grating is shown in Fig. 5.15(b), where some rounding can be seen at the top of the structures. Characterization results of the fabricated DOEs show that the experimental diffraction efficiencies are between 70% and 80%, which are about 4% lower than the respective theoretical predictions, presumably due to the fabrication errors. The design algorithm is being calibrated and the complete results are planned to be jointly published by scientists at TPS and TB.



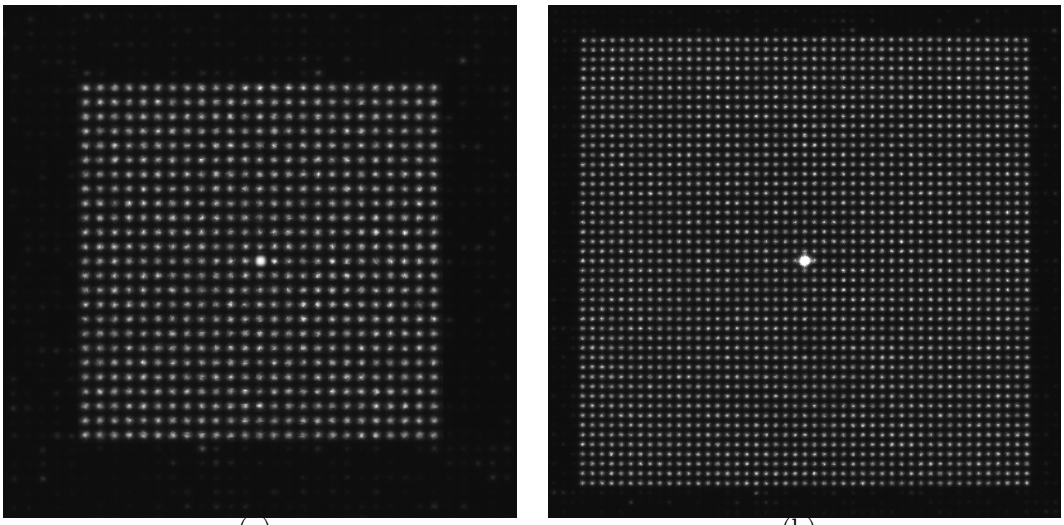
(a) (b)

Figure 5.15 — (a) 3D interferometric microscope image of a 8 level blazed grating. (b) Etching profile of the structure in (a).

5.4.3 Spot array DOEs

During this thesis period, we also collaborated with Joseph Fourier University (JFU) in Grenoble for the parallelization of Two-Photon Polymerization (2PP) lithography

using either DOEs or a Spatial Light Modulator (SLM). The basic principle of 2PP is to use a very high power laser to polymerize a micro- or nanoscale region of a photosensitive material, e.g. ORMOCER, via two-photon absorption process. Recent developments in ultrashort pulsed lasers (see [147] for example) have provided enough laser power for a parallel 2PP process. Details on 2PP together with the use of SLM for 2PP parallelization will are in Chapter 6. In this section, we concentrate on the application of spot array DOEs to split the single 2PP laser beam into multiple parallel beams. The purpose of parallel 2PP laser beams is to parallelize and hence speed up the polymerization process for industrial production of high resolution (< 1 μm) 3D objects of significant size (mm to cm), for example rotary needles that can be used in future dental treatment. With the highly bio-compatible ORMOCER used as the needle material, the occurrence of infections is much lower than with current needles [148].

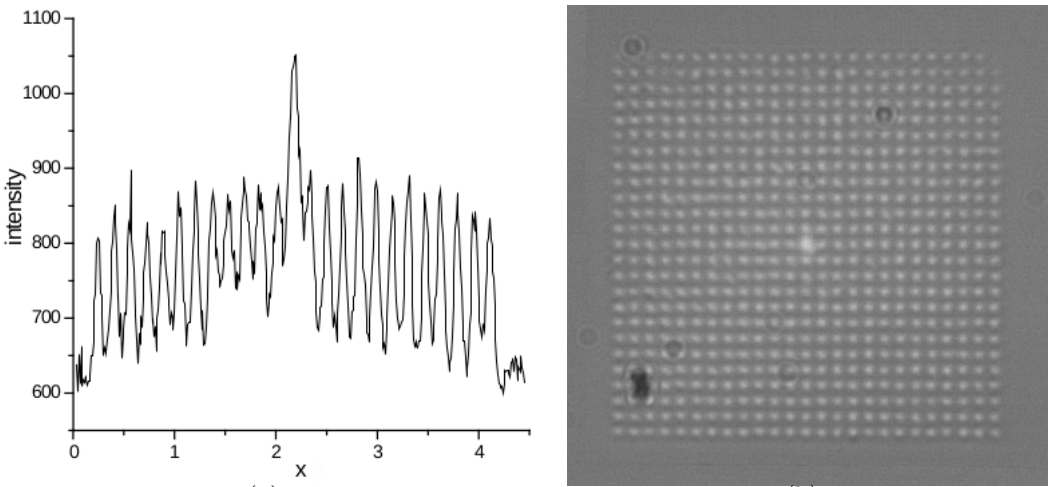


(a) (b)

Figure 5.16 — Experimentally observed diffraction patterns of the fabricated DOEs: (a) 25×25 spot array, (b) 47×47 spot array.

Fig. 5.16 shows the experimentally observed patterns of 2 spot array DOEs that were fabricated using our photoplotter. The DOEs were later used in a parallel 2PP setup at JFU. It has been shown [148] that the uniformity requirement in the 2PP production is relatively strict: the intensity difference between the diffraction spots should be within 10%, which is difficult to obtain, particularly for the zero order, due to the errors in our fabrication process, as analyzed in Section 5.3.1. The intensity distribution over a horizontal line across the zero order of the experimetal 25×25 spot array is shown in Fig. 5.17(a), where each spike represents a diffraction spot, and a relatively higher zero order can be seen. An image of the parallel 2PP fabrication using this DOE is illustrated in Fig. 5.17(b), which shows that the number of parallel 2PP beams is already significantly higher than recently published results [149–151]. The

parallel 2PP production using the 47×47 DOE was not so good, as the zero order is too high. A zero order reduction technique is being studied in our cleanroom to increase the number of uniform parallel beams for 2PP mass production, the complete results are planned to be jointly published by scientists at JFU and TB.



(a) (b)

Figure 5.17 — (a) Intensity distribution over a horizontal line across the zero order of the 25×25 spot array. (b) Microscope image of the parallel 2PP fabrication using our 25×25 spot array DOE. (Images taken from [148])

5.5 Conclusion

In conclusion, we have shown experimentally that our scalar non-paraxial propagator can be used for the modeling and design of thin Fourier DOEs with diffraction angles up to about 37°, where the pattern shapes and diffraction positions are very accurately predicted and controlled. Some discrepancies remain in the prediction of spot power at high diffraction angles but result more from fabrication errors than model errors and the accuracy is thus more than sufficient for many applications.

Fabrication of a thin binary DOE using a high-resolution EBL has allowed us to investigate the limit of the TEA and the accuracy of our vectorial model in predicting Hermitian diffraction symmetry of binary Fourier DOEs having features on the order of the wavelength. Again, at these feature sizes, which are close to the limit of our parallel photoplotter, the variations in the experimental diffraction efficiencies and Hermitian diffraction symmetry due to fabrication errors are very frequently more dominant than the errors resulting from the TEA and our scalar model. Future work will concentrate on the fabrication of thick DOEs and a more complete understanding of the practical limit of the TEA when designing such DOEs. For cases where the TEA is no longer valid

the FMM or FDTD + our scalar non-paraxial model with a genetic design algorithm outlined in Chapter 4.

With these experimental verifications, we have reached the resolution limit of our current photoplotter, while the EBL fabrication process is often time consuming and expensive. Towards the end of this PhD, I also participated in the design and preliminary development of a new parallel-write photoplotter based on Two-Photon Polymerization as a way to fast, cost-effective prototyping of high resolution (submicron) structures, as will be described in the next chapter.

CHAPTER 6

Parallel two-photon polymerization lithography

In the previous chapters, we showed that with our DOE design algorithms, we have reached the resolution limit of our current photoplotter. An EBL based fabrication process allows smaller structures to be produced but is often time consuming and expensive. As part of the process of continually improving DOE modeling and fabrication, we recently started to investigate the possibility of building a new, cost-effective photoplotter that is able to fabricate submicron structures at a reasonable writing time. Two-Photon Polymerization (2PP) appears to be a good candidate in terms of fabrication resolution and cost, but has not been fully parallelized in the literature. Chapter 5.4.3 briefly introduced a parallel 2PP lithography using DOEs at Joseph Fourier University (JFU) in Grenoble; another parallel 2PP system which is currently being built at Télécom Bretagne (TB) in collaboration with JFU using a Spatial Light Modulator (SLM) is addressed in this chapter.

6.1 Review of 2PP lithography

6.1.1 Principle of 2PP

Two-Photon Polymerization is a direct-write technique (i.e. without the use of photomasks) for creating micro- or nano- features in a photosensitive material. This method relies on a two-photon absorption process which is the simultaneous absorption of two photons (usually of identical frequencies) of the write-beam in order to excite a molecule from one energy state to a higher state, as illustrated in Fig. 6.1(b). The energy difference between the involved lower and upper states of the molecule is equal to the sum of the energies of the two photons. The phenomenon results in a chemical change (usually polymerization) at the focal volume of the write-beam, which is not neccessarily

on the surface of the material. This polymerized spot is the volume element in 2PP, which is often referred to as the "voxel" in analogy to the pixel (picture element) [38]. By scanning and properly modulating the write-beam, an arbitrary 3D pattern can be polymerized in the material, which can then be developed in a chemical solution, resulting in a fully 3D structure.

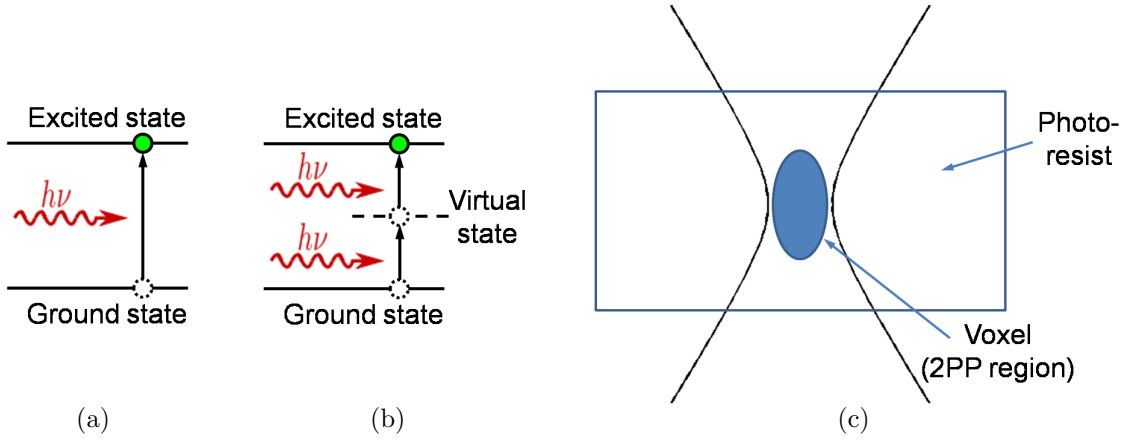


Figure 6.1 — (a) One-photon absorption. (b) Two-photon absorption. (c) 2PP in the photoresist layer and at the focal spot of a focusing beam. (ν is the frequency of the write-beam and h is the Planck constant)

Two-photon absorption differs from one-photon absorption in that it is a nonlinear optical process rather than a linear one [152]. The transition rate (i.e. number of excited molecules per unit volume over time t) by one-photon absorption can be written as follows:

$$\frac{dN_{OP}}{dt} = \sigma N_{GS} F \tag{6.1}$$

where σ is the one-photon absorption cross section and F is the photon flux (number of photons per unit area and time). N_{GS} and N_{OP} are the numbers of molecules per unit volume in the ground state and in the excited state due to one-photon absorption, respectively. Similarly, the transition rate for two-photon absorption is given by:

$$\frac{dN_{TP}}{dt} = \frac{1}{2}\delta N_{GS} F^2 \tag{6.2}$$

where δ is the two-photon absorption cross section and N_{TP} is the number of molecules per unit volume in the excited state due to two-photon absorption. The factor $1/2$ is because two photons are needed to produce one excited molecule. This equation shows that the probability of two-photon absorption depends on the square of the photon flux (or equivalently, to the square of the light intensity).

Since probability for two-photon absorption is several orders of magnitude lower than that of one-photon absorption, very high light intensities are required to increase the number of such rare events. Pulsed laser sources are usually preferred as they

deliver high-intensity pulses while depositing a relatively low average energy, since the energy is reserved for a relatively long period (ms) and released in a very short time (ns or shorter). To enable fully 3D structuring, the write-beam and the photoresist must be adequately chosen so that single-photon absorption is highly suppressed while two-photon absorption is favored. This condition is met if and only if the photoresist is highly transparent for the write-beam's wavelength λ and simultaneously, absorbing at $\lambda/2$ (i.e. no real energy levels corresponding to λ but to $\lambda/2$). A chemical substance, so-called "photoinitiator", is usually added into the photoresist material to further increase the probability for two-photon absorption [153]. At these conditions, the focal regions of the laser beam which exceed a certain intensity threshold of the photosensitive material define the 2PP volume element, which are the axial and lateral resolutions, as illustrated in Fig. 6.2.

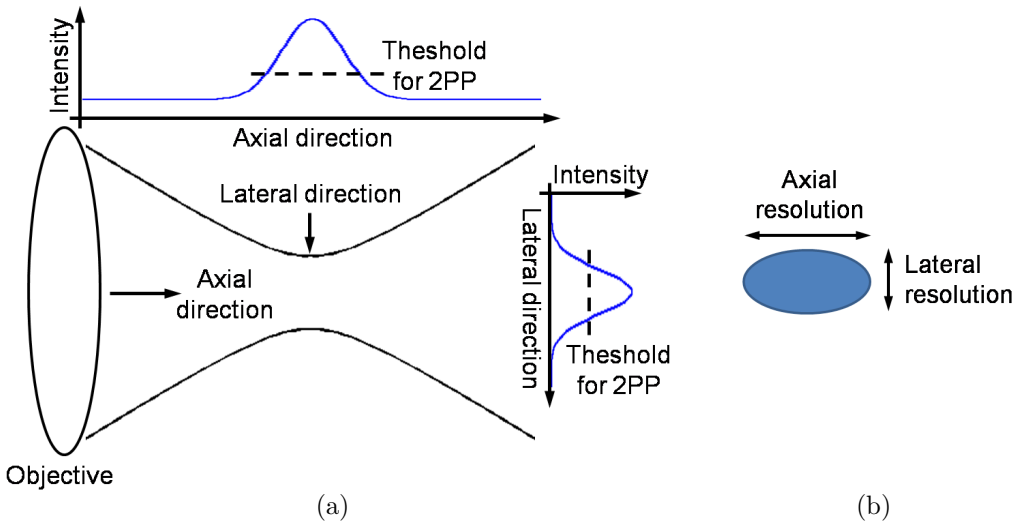


Figure 6.2 — (a) Intensity distribution of a focused beam in the axial and lateral directions. (b) Axial and lateral resolutions of the 2PP voxel.

In general, the resolution of an optical system is limited by the laws of diffraction. In optical microscopy, Ernst Abbe found the lateral resolution to be $d = \lambda/(2NA)$, where NA is the numerical aperture. With typical NA values of about 1, the resolution limit is therefore about half the wavelength [152]. In a first approximation, the use of two-photon absorption shifts this resolution limit by a factor of $\sqrt{2}$ (assuming Gaussian profiles and the same normalized threshold values), since the exposure is proportional to the squared intensity rather than the intensity itself, as illustrated in Fig. 6.3(a). This condition states the lateral resolution limit is $d = \lambda/(2\sqrt{2}NA)$, where $\lambda = 532$ nm for a green laser. Strictly speaking, with minimum feature sizes of about 200 nm, 2PP is not yet a nanotechnology but still a microtechnology, as the borderline between the two is commonly set to be at 100 nm [38].

In addition to the square intensity dependence, there is still another phenomenon

that can yield even better resolution to 2PP process, which is the 2PP threshold, as illustrated in Fig. 6.3(b). This threshold imposes a minimum intensity below which, no polymerization takes place. This phenomenon usually occurs due to the presence of oxygen in the photoresist, which inhibits the action of the photoinitiator, preventing the 2PP from occurring [154]. It can be used to decrease both the lateral and axial resolutions by reducing the write-beam intensity as close to the threshold as possible. This idea has been used for the 2PP fabrication of structures beyond the diffraction limit and down to 9 nm [155–157].

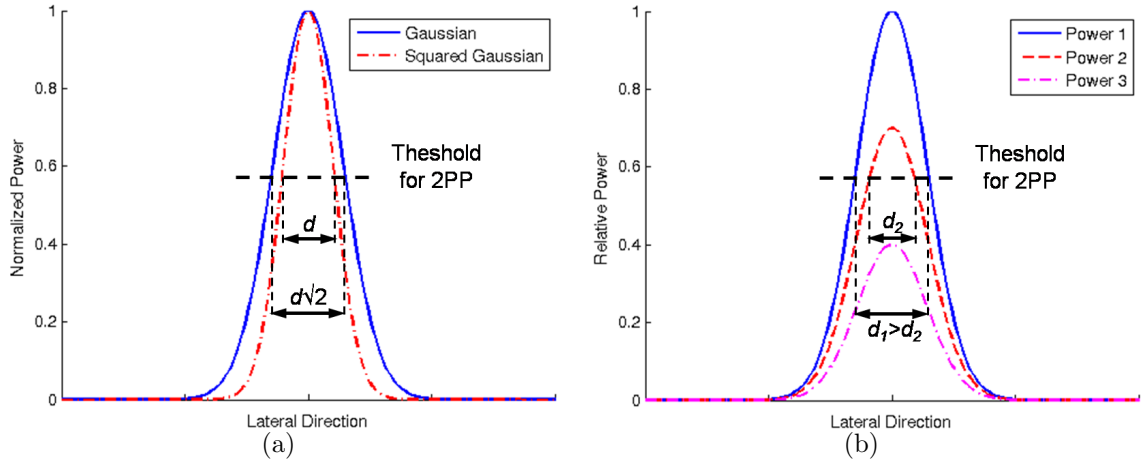


Figure 6.3 — (a) Gaussian and squared gaussian intensity profile. (b) Threshold power relative to various gaussian intensity profiles. The closer the peak power to the threshold power, the smaller the voxel size. 2PP is not observed if peak power is below the threshold power.

Another advantage of 2PP is that it does not necessarily require cleanroom facilities, which greatly reduces the cost in comparison to other conventional micro fabrication techniques [150]. The main disadvantage however is the long writing time, as the write-beam has to be scanned point-by-point in 3D in relative to the photoresist layer in order to polymerize a 3D pattern. For this reason, 2PP has been limited to the fabrication of relatively small structures, such as microneedles, optical waveguides/interconnects, photonic crystals [158–160]. Researchers have been looking for a way to parallelize the 2PP fabrication to speed up the 3D production process. Current parallel 2PP systems are reviewed in the next section.

6.1.2 Current parallel 2PP systems

Recent developments in ultrashort pulsed lasers (see [147] for example) have provided much higher laser power than the threshold required for a single 2PP process. It is therefore of interest to split the laser beam into many parallel beams with relatively

uniform power and higher than the 2PP threshold to speed up the 2PP fabrication process. The parallelization can be done by using a spot array DOE (Section 5.4.3 is the first demonstration, to our knowledge) or a SLM [149–151].

DOEs provide an effective and efficient way for mass production of periodic structures, but they always generate a fixed diffraction pattern, which is not suitable for variable fabrication of complex structures. On the other hand, SLMs often consist of about a million of pixels, where the transparency of each pixel can be controlled by a computer, allowing for variable splitting of the laser beam into multi-focus spot. However, most publications used an SLM as a dynamic spot array Computer Generated Hologram (CGH), where the CGH pattern can be computed and updated on the SLM. In this way, the full functionality of the SLM has not been used, as CGHs have to be computed in real-time, and the parallel beams are limited to multiple foci, which is the far-field diffraction pattern of the spot array CGHs.

Moreover, the number of parallel beams has not been fully optimized. In Section 5.4.3, we showed that with the use of a 25×25 spot array DOE, the 2PP fabrication at JFU can be parallelized by a factor of 625, which is significantly higher than recently published results, where the number of parallel 2PP beams is about 10 [149–151]. We aimed to further increase the number of parallel 2PP beams and use a SLM in a more dynamic way, by setting up a parallel 2PP photoplotter at TB, as addressed in the next section.

6.2 Experimental 2PP photoplotter

The idea of our parallel 2PP photoplotter is similar to the current parallel photoplotter (see Section 1.3.3), where a SLM is used to modulate the write-beam according to a pattern, which is displayed on the SLM and imaged onto the photoresist layer via a microscope objective. Fig. 6.4 shows the schematic diagram of our parallel 2PP photoplotter. The basic principle is as follows:

- A ultrashort pulsed, green laser (400 ps, 532 nm) is used for parallel 2PP. Three mirrors M1, M2, M3 guide the laser beam into the optical setup. Two lenses L1, L2 expand the laser beam size to cover the whole area of the SLM, which is about 1×1 cm^2.
- The SLM is controlled by a computer, in which the polarization of the beam passing through each pixel on the SLM can be modulated. The beam passing through the SLM then consists of multiple parallel beams of either Transverse Electric (s-polarization) or Transverse Magnetic (p-polarization). For preliminary 2PP tests, the SLM can also be replaced by some chrome or film masks, which act as amplitude modulators.

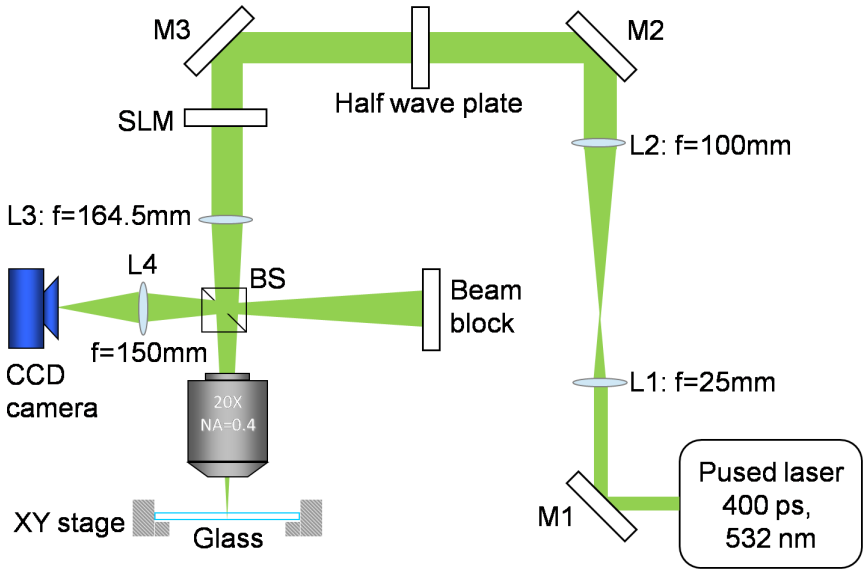


Figure 6.4 — Schematic diagram of our parallel 2PP photoplotter.

- Lens L3 is used to focused the modulated beam into the entrance pupil of the microscope objective.
- A polarization Beam Splitter (BS) is to split the modulated beam into two parts. Only the p-polarized beams go to the microscope objective, whereas the s-polarized beams are absorbed by the beam block.
- A microscope objective images the pattern on the SLM (pixels corresponding to the p-polarization) onto the photoresist layer. We have several objectives, with magnification ranging from 20X to 100X, and NA ranging from 0.4 to 1.4 (oil immersion). Effects of the microscope objectives on the 2PP voxel size are shown in Fig. 6.5.
- A half-wave plate is used to modify the polarization of the laser beam, so that the laser power passing through the BS is not too high (to avoid burning the photoresist instead of 2PP).
- Lens L4 is used to collect the reflection of the parallel beam on the photoresist layer and focus onto the CCD camera, which is connected to a computer. The microscope objective, lens L4 and the CCD act as a microscope to help finding the focus of the laser beam on the photoresist layer and provide a live image of the 2PP fabrication. Since there are different interfaces (air-substrate, substrate-photoresist and photoresist-air), not to mention that the photoresist layer is not neccessarily uniform, this step could be very problematic, as illustrated in Fig. 6.6.
- A nano-precision piezo-controlled XY-stage is used to change the lateral positions of the photoresist layer in relative to the parallel beam. The microscope objective

is attached to another piezo-controller to scan in the axial direction. The XY-stage and the laser are drived by a controller box, which is connected to the computer via RS-232 communication port.

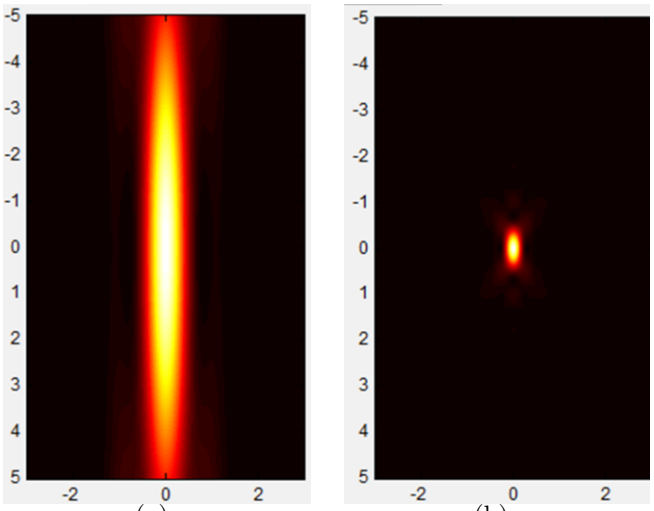


Figure 6.5 — Simulated 2PP voxel size with different microscope objective using a software developed by JFU: (a) 0.7 μm lateral and 8.8 μm axial resolutions for the 20X objective, (b) 0.3 μm lateral and 0.7 μm axial resolutions for 100X objective.

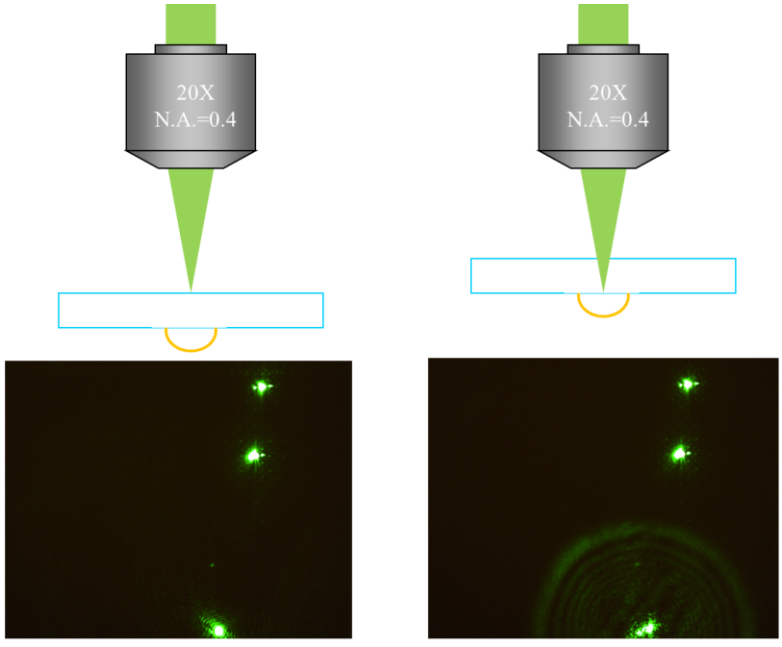


Figure 6.6 — CCD camera images when the focus is (a) at the air-substrate interface and (b) at the substrate-photoresist interface. In the latter case, the image had more rings than that of the air-glass interface focusing.

A photograph of the optical setup in the TB cleanroom can be seen in Fig. 6.7. We use a PowerChip^{TM} passively Q-switched laser from Teem Photonics [147], and a 800×600 pixel SLM from Bildsysteme GmbH [161]. The optical cage system, lenses, mirrors, etc. are from Thorlabs [162]. During the end of this PhD, I helped set up and align the optical cage system, program C code to control the XY-stage, the laser, the SLM and synchronize them. Preliminary 2PP fabrication results are given in the next section.

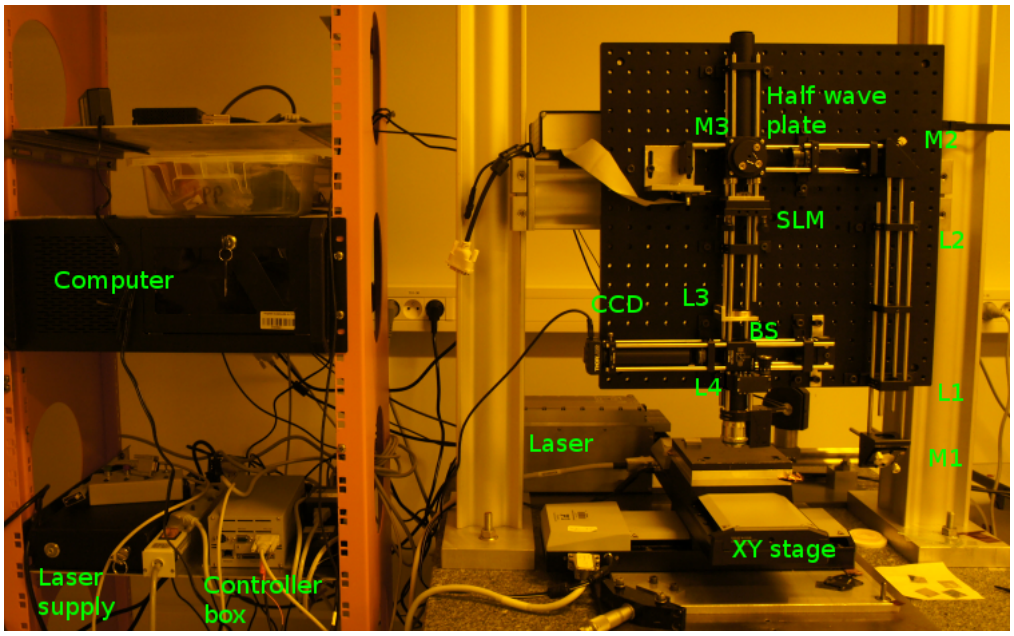


Figure 6.7 — Optical setup of the parallel 2PP photoplotter in TB cleanroom.

6.3 Preliminary fabrication results

For 2PP fabrication, Ormocer was used as the photopolymerizable material. However, since the refractive index of Ormocer is close to that of the glass substrate, the reflected image on the CCD camera is poor, which makes the focusing task very difficult, especially with the high NA oil-immersion objective. To facilitate the preliminary parallel 2PP test, we started with gold ablation, as the reflected image is much clearer, and the threshold power for gold ablation is lower than that for 2PP.

6.3.1 Gold ablation

For gold ablation, a thin gold layer (about 30 nm) is deposited on a glass substrate by sputtering. We started with a sequential (single laser beam) ablation in order to test our codes for controlling the laser, the XY-stage and their synchronization.

Sequential processing

For a sequential processing, the lenses L1, L2 and the SLM are not needed in the optical setup. The laser beam goes via the mirrors directly into the microscope objective. With this high power laser, a single laser pulse and the 20X microscope are enough to ablate the gold layer, provided that the gold layer is at the focal plane of the microscope objective. The XY stage is then moved to a new position to be ablated before the laser generates another pulse. The process continues until the complete ablation pattern is produced, as shown in Fig. 6.8, which demonstrates that our codes for controlling the laser, the XY-stage and synchronizing them worked well.

Figure 6.8 — Microscope image of a pattern fabricated by sequential gold ablation. The white regions are the ablated areas.

Parallel processing

For parallel processing, the lenses L1, L2 are necessary in the optical setup to expand the laser beam to cover the whole area of the SLM. As the SLM has 800×600 pixels, it splits the beam into almost 0.5 million parallel beams, each with a small fraction power of the laser source. For this reason, a higher magnification microscope objective and multiple laser pulses at a fixed position of the XY stage may be necessary to ablate the gold layer. In this test, we used a 40X microscope and tested the number of laser pulses needed for parallel gold ablation, together with our codes for controlling the SLM and synchronizing it with the laser and the XY-stage.

The process is as follows. Firstly, the XY stage is moved to a desired position. An image, which is a chessboard pattern, is sent from the computer to be displayed on the SLM, before a number of laser pulses are generated. The stage may then moved to a new position before the process continues (i.e. a pattern is displayed on the SLM and a number of laser pulses generated), but it is not necessary for this test. Fig. 6.9 shows microscope images of the patterns fabricated by parallel gold ablation using a 40X microscope and different numbers of laser pulses. It can be seen from Fig. 6.9(a) that

laser power with 1 pulse is not enough (the squares in the chessboard pattern are not connected together), whereas the 100000 pulse ablation in Fig. 6.9(d) is over-exposed. A parallel ablation with a number of pulses between 10000 and 50000 should result in the best ablation pattern. It appeared later that using a 63X microscope objective, even a single laser pulse produced a slightly over-exposed ablation pattern (not shown here).

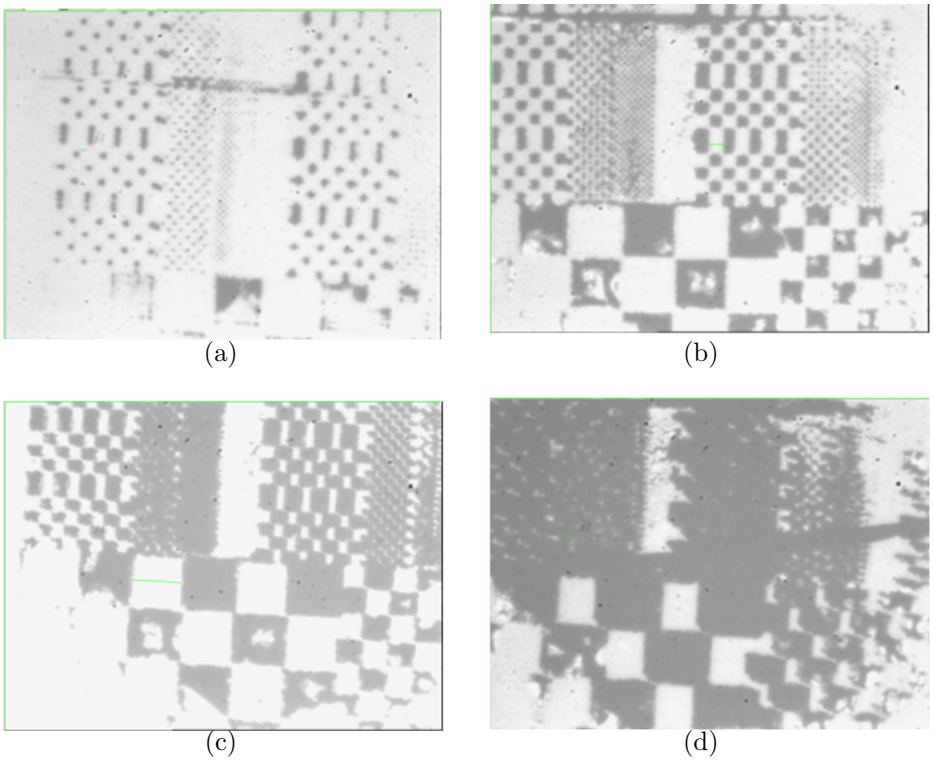


Figure 6.9 — Microscope images of the patterns fabricated by parallel gold ablation using a 40X microscope and: (a) 1 pulse, (b) 10000 pulses, (c) 50000 pulses, (d) 100000 pulses. The black regions are the ablated areas.

These results confirm the working principle of our parallel 2PP photoplotter and the usefulness of a high power laser as well as a high magnification and NA microscope objective for the rapid fabrication. With this understanding on gold ablation, we continued with sequential and parallel 2PP fabrication, some preliminary results are given in the following section.

6.3.2 2PP fabrication

For 2PP fabrication, Ormocer is chosen thanks to its sensitivity to 2PP at 532 nm. Before it is dropped on a glass substrate, this photopolymerizable material is "doped" by adding a suitable photoinitiator to increase the probability for two-photon

absorption [153]. To verify the 2PP threshold power and the voxel size, we also started with a sequential 2PP fabrication.

Similar to the sequential gold ablation, the lenses L1, L2 and the SLM are not present in the optical setup. As it is a single point fabrication, the 20X microscope objective provides enough laser power for 2PP. Fig.6.10 shows microscope images of different lines fabricated by sequential 2PP, where the distance between 2 continuous spots in each line is about 10 μm. The laser power is controlled via the half-wave plate, where each vertical line corresponds to a laser power going into the microscope objective. It can be seen from left to right, corresponding to reducing power that the spot size gets smaller with the lower laser power, as analyzed in Section 6.1.1. The spot diameter on the rightmost line is about 5 μm, which is significantly bigger than the 0.7 μm prediction shown in Fig. 6.5, as the laser power is still too high. The voxel size can of course be reduced by further reducing the laser power and using a higher magnification objective.

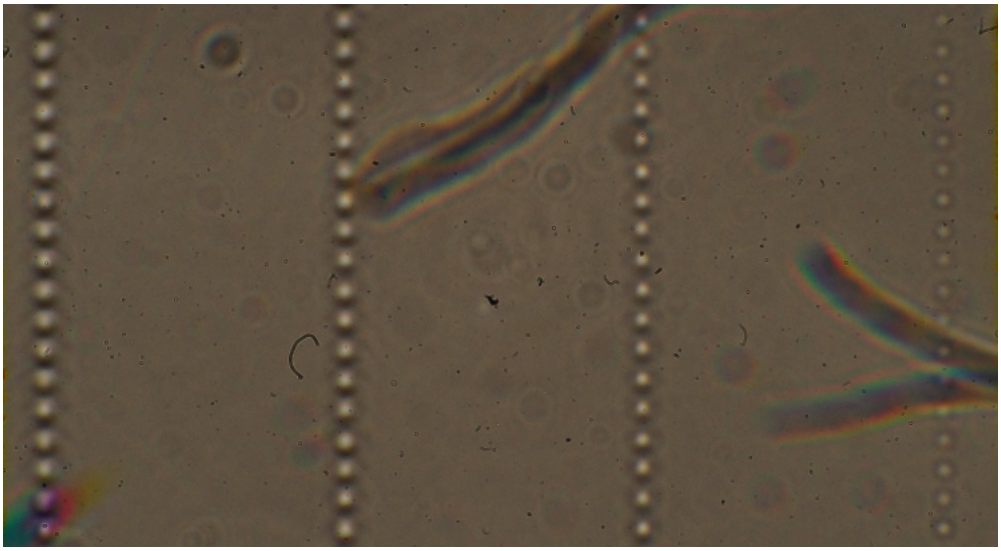


Figure 6.10 — Microscope images of different lines fabricated by sequential 2PP. Each vertical line corresponds to a laser power going into the microscope objective. From left to right is the direction of reducing laser power.

With this result, we continued with a parallel 2PP fabrication test. As before, a higher magnification and NA objective is needed to compensate for the power splitting factor of the beam expander L1, L2 and the SLM. As the 2PP threshold power is higher than that for gold ablation, the 63X objective and 100X oil-immersion objective have to be used. However, problem appears with these objectives as the voxels get smaller, and they can easily be washed out during development. This phenomenon is due to the difficulty in focusing the laser beam on the glass-photoresist interface, especially with these high NA objective. As the voxels get smaller, they may not be well connected to the glass subtrate and will be removed during the development, together with the

unexposed areas (as it is a negative photoresist). Experimental results with S1805, which is a positive photoresist, show that the laser power is not enough, even with the 100X oil-immersion objective, as this photoresist is not sensitive to 2PP at 532 nm.

For these reasons, it is necessary to develop program for the piezo-controller in the axial direction and an algorithm to find the focus of the microscope objective automatically. The work will facilitate the parallel 2PP fabrication and is currently under progress.

CHAPTER 7 Conclusion and perspectives

The main aim of this thesis, as stated in the introduction, was to extend the range of Diffractive Optical Element (DOE) applications by developing models, algorithms and rapid prototyping techniques for large diffraction angle DOEs: diffraction angles $> 10°$ so beyond the limits of scalar paraxial diffraction model. We have developed a scalar non-paraxial far-field propagator, which we have shown to be accurate, computationally efficient and overcomes the limits of the conventional scalar diffraction models. Experimental results using our existing photoplotter show that our scalar non-paraxial propagator predicts correctly diffraction position and pattern shape for diffraction angle up to about 33°, which is beyond the limit of the scalar paraxial diffraction regime. Some discrepancies remain in the prediction of spot power at high diffraction angles but we have shown that they result more from fabrication errors than model errors and the model accuracy is thus more than sufficient for many applications [107]. An iterative algorithm based on this propagator has been developed for the design of wide-angle Fourier elements. It converges to much more optimized solutions than the standard IFTA with the same order of computational complexity. Fabrication of submicron DOEs during my stay at Karlsruhe Institute of Technology in Germany and characterization of these DOEs confirm experimentally that our iterative scalar non-paraxial algorithm can be used to design thin Fourier DOEs with diffraction angles up to about 37° and perhaps even higher [117].

The remaining discrepancies in diffraction power between modeling, design and experiment were then investigated and shown to result from both fabrication errors and by the fact that we are approaching the limit of the Thin Element Approximation (TEA). The practical limits of the TEA in predicting the symmetry of the diffraction pattern by binary phase DOEs were shown by comparison with the rigorous vectorial simulations. In addition, our metric for evaluating the accuracy of the simulations, which is the symmetry of the diffraction pattern by binary elements, has been shown to be simple yet efficient and consistent, despite the variations in the simulated and experimental diffraction efficiencies. It appears that for designing high uniformity spot

array DOEs, it is necessary to take vectorial effects in the DOE itself and its close surroundings into account. We have developed, optimized and parallelized a rigorous diffraction model based on the Finite-Difference Time-Domain (FDTD) method coupled with our scalar non-paraxial propagator to overcome the theoretical limits of the TEA and the computational limitations of current vectorial models. A genetic design algorithm based on this model has been proposed for the design of thick DOEs. Future work will concentrate on the fabrication of thick DOEs and the calibration/optimization of the genetic FDTD + our scalar non-paraxial design algorithm.

As we had reached the resolution limit of our current photoplotter, we investigated the possibility of building a new parallel-write photoplotter based on Two-Photon Polymerization (2PP) as a way to rapid, cost-effective prototyping of high resolution (submicron) structures. With the aid of a 25×25 spot array DOE designed and fabricated by ourselves at Télécom Bretagne (TB), the 2PP fabrication process used at Joseph Fourier University in Grenoble has been successfully parallelized by a factor of 625. A zero order reduction technique is being studied in our cleanroom in order to increase the uniformity of bigger size spot array DOEs (47×47). To further speed up the 2PP fabrication process, another parallel 2PP photoplotter using a Spatial Light Modulator, which can generate up to about 0.5 million parallel beams, has been designed and developed at TB. Preliminary experiments with gold ablation confirm the design of the optical setup and the working principle of our parallel 2PP photoplotter. Future work will also concentrate on further development of the parallel 2PP photoplotter for the fast fabrication of high resolution and 3D structures.

Sommaire

Les éléments optiques diffractifs (EODs) sont des micro ou nano-structures qui sont conçues et réalisées pour générer par diffraction une répartition de champ lumineux désirée. L'utilisation des EODs n'a cessée d'augmenter ces dernières années, et ce, pour une très vaste gamme d'applications. Quelques exemples sont donnés dans l'introduction générale de ce document. Pendant longtemps, le développement de l'optique diffractive a été entretenu par les avancées conjointes et complémentaires des théories de la diffraction et des techniques de fabrication des EODs, qui sont brièvement rappelées dans le chapitre 1. Avant le début de cette thèse, pour concevoir les EODs et simuler les figures de diffraction générées, la théorie paraxiale de la diffraction était classiquement utilisée, ce qui de fait, nous limitait au cas de composants optiquement minces et à de petits angles de diffraction. Cependant, les avancées récentes au niveau des techniques de fabrication hautes performances permettent désormais de fabriquer des composant optiques épais ou diffractant sous de grands angles, ce qui rend nécessaire la mise au point de nouvelles techniques de modélisation et de nouveaux algorithmes de calcul.

Motivation

Le but de cette thèse est de concevoir, réaliser et optimiser des EODs fonctionnant dans des régimes de diffraction plus complexes que ceux permis par la théorie paraxiale, en vue d'obtenir des composants hautes performances (meilleure efficacité et plus grands angles de diffraction...) permettant des applications pour l'instant inenvisageables. L'ensemble du travail que nous avons réalisé (modélisation, fabrication, caractérisation de ces composants, exploitation des résultats expérimentaux), et l'expérience que nous avons acquise, nous permettent d'optimiser les algorithmes de calcul de ces composants et de déterminer leurs domaines d'applications sur des exemples concrets. De cette façon, nous allons au-delà des limitations actuelles et permettons d'accéder au design et à la réalisation d'une nouvelle famille de composants optiques diffractifs autorisant un spectre d'applications plus vaste.

Les différents modèles de la diffraction

Si on considère la structure d'un EOD donné, il existe différents modèles qui permettent de calculer la distribution du champ qu'il diffracte. Ceci provient essentiellement du fait que les équations de Maxwell, point de départ de la modélisation, ne peuvent en général pas être résolues analytiquement, ce qui oblige à faire appel à des approximations et à des méthodes numériques, présentant des contraintes et des limitations, notamment au niveau du domaine de validité. Les différentes approches utilisées classiquement pour modéliser théoriquement la diffraction, scalaires ou vectorielles, sont rappelées au chapitre 2. Dans cette partie, les domaines de validité et les contraintes numériques de chacune de ces méthodes sont analysés en détails, de façon à identifier clairement leurs limites respectives. Nous montrons notamment qu'il n'y a pas de solution idéale pour calculer la figure de diffraction en champ lointain et en régime non paraxial, sachant que les différents modèles résultent d'un compromis entre la vitesse de calcul et la précision du résultat, l'un allant au détriment de l'autre. Les méthodes vectorielles sont rigoureuses, mais ne peuvent être utilisées que pour des EOD très simples ou de très petite taille, sachant qu'elles nécessitent des calculs intensifs. Les théories scalaires non paraxiales couramment utilisées, telles que la méthode du spectre angulaire (MSA) ou la formule de Rayleigh-Sommerfeld (RS), modélisent la diffraction de façon précise, mais leur mise en oeuvre se trouve limitée en pratique par les contraintes imposées, notamment par l'échantillonnage et la complexité des calculs numériques. D'autre part, l'approximation de Fraunhofer permet de calculer rapidement la figure de diffraction dans un plan en champ lointain. C'est pour cette raison que jusqu'à présent, c'est cette méthode qui a été utilisée dans la plupart des cas pour modéliser les EODs fonctionnant en régime de Fourier. Mais ces derniers sont limités à un fonctionnement en régime paraxial, c'est à dire à des angles de diffraction $< 10°$.

Proposition d'un modèle scalaire non-paraxiale

Nous présentons au chapitre 3 le modèle que nous avons développé pour dépasser les limites des approches scalaires et permettre des temps de calcul rapides pour des EODs de Fourier en régime non paraxial. Dans ce chapitre, on étend le domaine de validité de la théorie scalaire de la diffraction, notamment pour calculer la diffraction d'EODs quelconques (pas seulement des lentilles ou des réseaux de diffraction à une dimension) dans un plan en champ lointain et en régime non paraxial, grâce à une projection et en utilisant le modèle de Harvey, comme cela est illustré sur la Figure 7.1. Le propagateur de Harvey est un modèle scalaire non paraxial qui permet de calculer très précisément le champ diffracté sur une calotte sphérique à l'aide d'une simple transformée de Fourier (TF). Cependant, étant donné que la surface d'observation est en pratique généralement

un plan, nous effectuons une projection de la figure de diffraction calculée sur la sphère sur le plan d'observation. De plus, une étape de ré-échantillonnage est nécessaire pour passer du champ exprimé en fonction des coordonnées angulaires sur la sphère (ce qui est dû à la TF) à son interpolation en coordonnées cartésiennes, classiquement utilisée dans le plan d'observation.

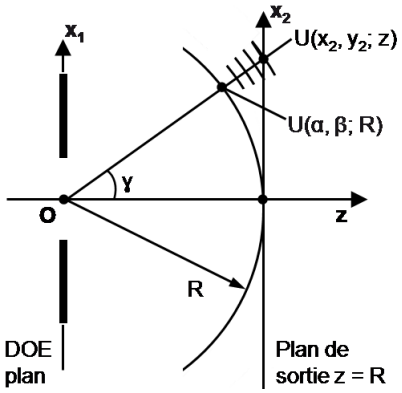


Figure 7.1 — Structure géométrique de la diffraction dans le modèle de Harvey et notre projection d'onde sphérique.

Les résultats de nos simulations numériques de la diffraction d'EODs échantillonnés montrent que le modèle que nous proposons (propagateur de Harvey + projection Sphère-Plan) est plus précis que l'approximation classique de Fraunhofer au niveau de la position et de l'intensité des différents ordres de diffraction. Si on compare les résultats que l'on obtient à l'aide de notre méthode à ceux issus de la théorie rigoureuse de RS, les erreurs sont de l'ordre de 1%, et les temps de calculs nécessaires sont beaucoup moins importants. De plus, notre méthode n'impose pas certaines contraintes strictes au niveau de l'échantillonnage dans le plan d'observation, contrairement à l'approche de RS, basée sur une convolution, et à la méthode du MSA. Si bien que pour obtenir la figure de diffraction en champ lointain une simple transformée de Fourier et une projection suffisent, ce qui signifie qu'en terme d'efforts ou de complexité, le calcul demeure très proche de celui des méthodes basées sur la TF. Nos résultats de simulations montrent que notre méthode peut être utilisée pour modéliser des EODs diffractant à des angles allant jusqu'à 33° et peut être au-delà si l'on accepte une perte au niveau de la précision des calculs, et à condition que les structures diffractantes ne soient pas trop petites, ce qui nécessiterait une théorie vectorielle de la diffraction.

D'un point de vue pratique, le problème inverse consiste à calculer la structure d'un EOD générant par diffraction une distribution de champ désirée dans le plan d'observation. C'est tout l'enjeu du design des EODs. Ceci nécessite un modèle mathématique de la propagation de l'onde diffractée et un algorithme d'optimisation des performances de l'EOD calculé (algorithmes itératives ou génétiques). Bien que de nombreux algorithmes très performants aient été implémentés, qui sont d'ailleurs rapi-

dement présentés dans le chapitre 1, la qualité des EODs calculés dépend fortement du modèle de diffraction utilisé lors de la conception. Par ailleurs, des avancées récentes au niveau des technologies de fabrication des EODs permettent désormais d'atteindre un niveau de résolution en dessous du micron, et par conséquent de l'ordre (voir en dessous) de la longueur d'onde de la lumière diffractée. De tels EODs diffractent la lumière aux grands angles et laissent ainsi entrevoir de nombreux champs d'applications prometteurs. Cependant, la théorie scalaire paraxiale de la diffraction ne peut plus être utilisée dans ce cas pour la conception de tels EODs. Il est alors nécessaire d'avoir recours à des modèles vectoriels de la diffraction, qui sont rigoureux mais nécessitent des calculs numériques intensifs. Ceci restreint leur utilisation à des réseaux de diffraction à une dimension ou des structures 2D extrêmement simples. Nous présentons aussi au chapitre 3 un algorithme itératif dédié au calcul d'EODs de Fourier basé sur notre modèle de propagation scalaire non paraxial, que l'on a montré être valable pour le calcul de la diffraction de la lumière en champs lointain à travers des structures ayant des tailles caractéristiques de l'ordre de la longueur d'onde.

Algorithme itératif reposant sur notre modèle scalaire non paraxial

Il a été montré que les algorithmes itératifs reposant sur la transformée de Fourier (IFTAs) font partie des algorithmes les plus optimisés et performants pour la conception d'EODs de phase. L'utilisation des IFTAs courants se trouve cependant limitée par l'utilisation en leur sein de la théorie scalaire paraxiale de la diffraction. Si malgré tout, on utilise un IFTA pour concevoir un EOD de Fourier fonctionnant en régime non paraxial, la figure de diffraction effectivement générée par cet élément présentera deux types de distorsions par rapport à la figure souhaitée: des déformations spatiales et une chute de l'intensité lumineuse avec l'angle de diffraction. Ces distorsions sont dues aux approximations qui sont faites lorsque l'on utilise la théorie scalaire paraxiale pour calculer la position et l'amplitude de la lumière diffractée. Pour corriger cela, nous avons proposé deux algorithmes basés sur deux techniques différentes de projection entre la sphère de la méthode de Harvey et le plan d'observation, comme cela est schématisé sur la Figure 7.2. Dans les deux cas, l'intensité du point diffracté est corrigée selon l'angle de diffraction ($P \propto \gamma |U|^2$). De plus, les même contraintes pour la quantification sont utilisée dans le plan de l'EOD et dans le plan image, afin d'optimiser l'efficacité de diffraction et l'uniformité de la figure de diffraction.

Dans le cas de l'algorithme (a), la figure de diffraction souhaitée dans le plan d'observation est rétro-projetée sur la sphère du modèle de Harvey, avant que l'on lance l'algorithme IFTA entre le plan de l'EOD et la sphère. La figure projetée est cor-

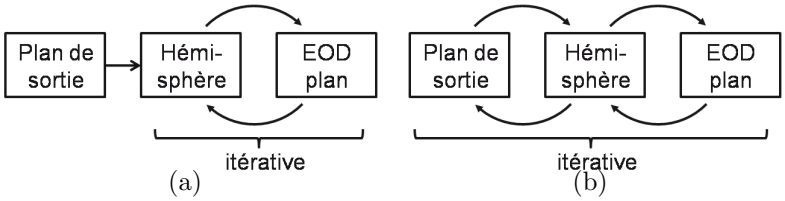


Figure 7.2 — (a) Algorithme itératif utilisant une seule projection. (b) Algorithme Itératif utilisant une projection itérative.

rigée à la fois au niveau de la position et de l'intensité de la lumière diffractée afin de compenser les distorsions dues à la divergence entre la sphère et le plan d'observation. L'EOD est ainsi calculé en utilisant l'algorithme IFTA standard à partir de l'image souhaitée projetée sur la sphère. La reconstruction dans le plan d'observation de notre EOD ainsi calculé s'avère meilleure en terme de position et d'intensité de lumière diffractée que celle que l'on obtient à partir d'un EOD conçu à l'aide de la méthode IFTA classique. Dans le cas de l'algorithme (b), la procédure de projection sphère - plan d'observation est incorporée au sein de l'algorithme itératif. Ainsi à chaque pas du calcul on effectue pour la phase de retour arrière: une projection Plan d'observation - Sphère, puis une transformée de Fourier de la sphère au plan de l'EOD, et enfin ces deux mêmes opérations sont répétées mais en sens inverse pour la phase aller. Comme cet algorithme itératif utilise une propagation en deux temps, sa complexité est environ le double de celle de l'algorithme (a). Par contre, il converge vers une solution mieux optimisée.

Au final, nous avons développé un algorithme rapide et précis au niveau du calcul à la fois de la position et de l'intensité de la lumière diffractée qui utilise une méthode itérative incluant une opération de projection entre le plan d'observation et la sphère de la méthode de Harvey, puis une transformée de Fourier entre la sphère et le plan de l'EOD. Cet algorithme peut être utilisé pour calculer des EODs de Fourier fonctionnant en régime non paraxial qui ont de nombreuses applications pratiques, telles que la séparation et la mise en forme de faisceaux par exemple, où la distance de propagation est souvent beaucoup plus importante que la taille de l'EOD. La figure de reconstruction de l'EOD que nous avons calculé grâce à notre algorithme présente une meilleure distribution en position et en intensité par rapport à ce que l'on peut obtenir via l'algorithme IFTA classique et est très proche de celle désirée, même pour de grands angles allant jusqu'à 37°. L'algorithme que nous avons conçu ne nécessite plus l'approximation paraxiale et ne dépend donc plus que de l'approximation des éléments minces (TEA) qui est utilisée pour relier le profil de phase de l'EOD au champ optique situé dans son plan. La limite de cette approximation est analysée en utilisant la théorie électromagnétique dans le chapitre 4 et nous proposons alors une approche vectorielle pour surmonter ces limitations.

Proposition d'un modèle vectoriel et conception

Au chapitre 3, nous avons montré que la diffraction aux grands angles, qui est l'un des ojectifs de cette thèse, peut être obtenue avec un EOD "mince", ayant des dimensions caractéristiques de l'ordre, voir en dessous, de la longueur d'onde, au prix d'un processus de fabrication plus élevé. Cependant, les EODs à haute résolution ne sont pas la seule façon d'obtenir de grands angles de diffraction. Il est bien connu qu'en holographie optique (par exemple les réseaux de Bragg en volume) de grands angles de diffraction peuvent aussi être obtenus avec des EODs ayant des tailles caractéristiques de l'ordre du micron, et notamment des profondeurs de gravure bien plus grandes que la longueur d'onde. Nous nous intéressons tout particulièrement à ce type d'EOD car nous pouvons désormais les fabriquer dans notre salle blanche, mais il s'est avéré que la réalisation de ce genre de composants n'avait jusqu'alors été que très peu étudié.

Bien que notre propagateur scalaire non paraxial soit parfaitement rigoureux pour la propagation en espace libre, sa validité dépend de l'approximation des éléments minces (TEA) au niveau de l'EOD. Or, cette approximation ne peut pas être utilisée pour des déphasages $> 2\pi$ (cas des EOD épais) et dans ce cas, il ne semble pas y avoir de bonne façon de modéliser ce genre de structure. Dans le chapitre 4, nous avons pour objectif de développer une méthode vectorielle rigoureuse pour la modélisation et la conception des EOD épais. Tout d'abord, nous examinons les limites pratiques de l'emploi de l'approximation des éléments minces (TEA) en comparant les résultats qu'elle donne à ceux obtenus via des simulations vectorielles (électromagnétiques) rigoureuses, utilisant la méthode FDTD (Finite-Difference Time-Domain) et la méthode FMM (Fourier Modal Method). Nous présentons alors un modèle rigoureux pour simuler la diffraction reposant sur la FDTD couplée avec notre propagateur en espace libre. Nous proposons ensuite un algorithme génétique reposant sur ce modèle pour aller au-delà des limites théoriques imposées par la TEA et des limitations de calcul inhérentes aux méthodes vectorielles classiques.

Pour examiner les différences au niveau de l'efficacité de diffraction entre les résultats que l'on obtient grâce aux différentes méthodes de calcul et ceux issus directement de l'expérience, nous montrons que les symétries dans la figure de diffraction peuvent être utilisées comme un critère simple et efficace pour évaluer la précision des simulations. Ce critère se révèle être pertinent malgré les variations de l'efficacité de diffraction entre les simulations et l'expérience. Les simulations FDTD/FMM montrent que la TEA se révèle incapable de prédire correctement la distribution du champ diffracté immédiatement derrière l'EOD à la fois en amplitude et en phase. Les simulations à l'aide de la FMM montrent que la méthode incluant la TEA et notre propagateur modélise précisément la diffraction pour des angles correspondant au cas non paraxial, mais de façon incorrecte les symétries de l'efficacité de diffraction dans le cas d'EODs

ayant des dimensions caractéristiques plus petites que deux fois la longueur d'onde. C'est pourquoi, pour calculer des EODs ayant une grande uniformité au niveau de l'efficacité de diffraction (cas d'un générateur d'une grille de points lumineux par exemple), il est nécessaire de prendre en compte les limitations de l'approximation TEA et d'inclure les effets vectoriels dans la région de l'EOD.

Nous développons un modèle vectoriel rigoureux basé sur la FDTD couplée avec notre propagateur scalaire non paraxial et comparons les résultats qui en découlent à ceux obtenus en utilisant l'approximation des éléments minces (TEA) couplée avec le même propagateur. Les simulations dans le cas d'EODs 1D et 2D montrent que notre modèle permet d'obtenir des résultats plus proches de ceux calculés avec la méthode FMM que ceux issus de la méthode alliant l'approximation TEA alliée au modèle de Harvey. Les paramètres pour la simulation FDTD et le couplage ont été optimisés afin de réduire l'utilisation de la mémoire et le temps de calcul, sans toutefois limiter la précision de l'algorithme. Afin de réduire encore le temps de calcul d'environ 35%, la parallélisation de l'algorithme sur un super calculateur a été réalisée en utilisant 8 processus fonctionnant en parallèle. Des EODs de plus grande taille peuvent ainsi être calculés et conçus en utilisant notre modèle rigoureux sur un super calculateur. Nous proposons aussi un algorithme génétique basé sur cette approche vectorielle pour concevoir des EODs épais.

En résumé, nous avons étudié les limites que l'on rencontre en pratique lorsque l'on utilise l'approximation TEA en estimant les symétries de la figure de diffraction engendrée par un EOD binaire de phase en comparaison de ce que l'on obtient à l'aide des simulations vectorielles rigoureuses. Les résultats montrent que d'un point de vue strict, l'approximation TEA ne devrait pas être utilisée pour des EODs binaires ayant des dimensions caractéristiques plus petites qu'à peu près deux fois la longueur d'onde, sachant que dans ce cas l'erreur au niveau de l'estimation du facteur de symétrie hermitienne est plus de 10%. Nous avons développé, optimisé et parallélisé un modèle rigoureux pour le calcul de la diffraction basé sur la méthode FDTD couplée à un propagateur scalaire en espace libre, afin de dépasser les limites théoriques de l'approximation TEA et les restrictions au niveau du calcul numérique des méthodes vectorielles classiques. De plus, notre critère permettant d'évaluer la précision des simulations, qui s'intéresse à la symétrie hermitienne de la figure de diffraction d'EODs binaires, s'est révélé simple, efficace et fiable, malgré les variations d'efficacité de diffraction entre les simulations et les résultats expérimentaux. Les vérifications expérimentales de nos différents modèles, algorithmes de calcul et des effets des erreurs de fabrication sur les efficacités de diffraction seront présentées au chapitre 5.

Vérifications expérimentales

Dans les précédents chapitres, nous avons étudié numériquement les limites des modèles classiques de diffraction (approximations de la diffraction scalaire paraxiale et des éléments minces), et développé des algorithmes scalaires et vectoriels afin de dépasser les limites habituelles de la modélisation et du calcul des EODs de Fourier diffractant aux grands angles. Au chapitre 5, les limites de l'approximation TEA et des modèles scalaires paraxiaux, de même que la précision de nos simulations scalaires non paraxiales et vectorielles sont vérifiées expérimentalement grace à la fabrication d'EODs tests et la mesure de leurs performances sur un banc optique.

Pour réaliser ces tests expérimentaux nous avons utilisé le photo-traceur à écriture directe de Telecom Bretagne (TB) pour la réalisation d'EODs structurés à l'échelle du micron, et le dispositif de lithographie par faisceau d'électrons (EBL) de l'Institut de Technologies de Karlsruhe (KIT) pour des éléments structurés à plus petite échelle. Nous décrivons successivement les confirmations expérimentales de notre modèle scalaire non paraxial et des algorithmes scalaires itératifs (présentés au chapitre 3) grâce à la fabrication et la caractérisation d'EODs de Fourier diffractant aux grands angles. Nous avons montré expérimentalement que notre propagateur scalaire non paraxial pouvait être utilisé pour modéliser et calculer des EODs de Fourier minces diffractant à des angles allant jusqu'à 37°, engendrant des figures de diffraction extrêmement précises en terme de forme et de position, et ceci de façon parfaitement contrôlée, en total accord avec nos prévisions. Quelques désaccords demeurent au niveau de la prédiction de l'intensité des points lumineux situés aux grands angles, mais ceci provient plus des erreurs au niveau de la fabrication que de notre modèle et de toutes façons, la précision obtenue est plus que suffisante pour de nombreuses applications.

La fabrication d'un EOD binaire à l'aide la technique haute-résolution EBL nous a permis d'étudier les limites de l'approximation TEA et la précision de notre modèle vectoriel en estimant la symétrie hermitienne de la figure de diffraction générée par un EOD binaire de Fourier ayant des dimensions caractéristiques de l'ordre de la longueur d'onde. Nous montrons que la symétrie des taches de diffraction se révèlent être un bon moyen d'évaluer les effets des erreurs de fabrication et des limites en pratique de l'approximation TEA. Il s'agit d'une piste prometteuse pour de futures études. De nouveau, à ces échelles, les écarts au niveau des efficacités de diffraction et des symétries hermitiennes, dûs aux erreurs de fabrication, sont souvent beaucoup plus importantes que les erreurs résultant de l'approximation TEA et de notre modèle scalaire. Nous terminons le chapitre en présentant quelques exemples d'applications des EODs fabriqués pour des partenaires académiques ou industriels de TB. Les travaux à venir se concentreront sur la fabrication d'EODs épais et viseront une compréhension plus large des limites en pratique de l'approximation TEA lors de la conception de tels EODs.

Polymérisation à deux photons parallèle

Les algorithmes de calcul et de conception que nous avons proposés, de même que les vérifications expérimentales que nous avons effectuées, nous ont permis de montrer que nous avons atteint les limites de résolution de notre photo-traceur actuel à TB. Les procédés de fabrication reposant sur la technique de lithographie par faisceau d'électrons permet de produire de plus petites structures, mais cela nécessite beaucoup de temps et c'est très cher. Dans un contexte de recherche d'améliorations constantes au niveau de la conception et de la réalisation des EODs, nous avons étudié la possibilité de construire un nouveau photo-traceur rentable qui serait capable de fabriquer des structures dont les dimensions caractéristiques seraient en dessous du micron et ce en un temps raisonnable. La polymérisation à deux photons (2PP) semble être une bonne technique en terme de résolution au niveau de la fabrication et du prix, mais d'après la littérature, elle n'a pas été totalement parallélisée. Nous présentons au chapitre 6 une description du photo-traceur utilisant la 2PP que nous avons conçu et l'état de l'avancement de sa réalisation. Il s'agira d'un moyen de prototypage de structures submicroniques rapide et rentable. Grâce à un EOD affichant une matrice de 25 x 25 points que nous avons conçu et réalisé à TB, le procédé de fabrication 2PP utilisé à l'Université Joseph Fourier (UJF) de Grenoble a pu être parallélisé avec succès d'un facteur 625. Une technique de réduction de l'ordre 0 est en train d'être étudiée dans notre salle blanche afin d'augmenter l'uniformité d'une matrice de points lumineux plus grande (47 x 47 points) générée par un EOD. Pour pouvoir encore plus augmenter la rapidité du processus de fabrication 2PP, un autre photo-traceur 2PP parallèle utilisant un modulateur spatial de lumière, qui peut engendrer jusqu'à 0.5 millions de faisceaux parallèles, a été conçu et est développé à TB en collaboration avec UJF. Des expériences préliminaires d'ablation d'or valident la conception du dispositif optique et le principe de fonctionnement de notre photo-traceur 2PP parallèle.

Conclusion

Le principal objectif de cette thèse était d'étendre le domaine d'applications des EODs en développant de nouveaux modèles, algorithmes et techniques de prototypage rapide permettant de réaliser des EODs diffractant à des angles bien supérieurs à 10°, c'est à dire au-delà des limites de l'approximation scalaire paraxiale de la diffraction. Nous avons mis au point un propagateur scalaire non paraxial pour le champ lointain, dont nous avons démontré la précision et l'efficacité pour les calculs et qui permet de dépasser les limites des modèles scalaires classiques de la diffraction. Les résultats expérimentaux obtenus à l'aide de notre photo-traceur actuel montrent que notre propagateur non paraxial prédit correctement la position et la forme des figures

de diffraction pour des angles allant jusqu'à 33°. Un algorithme itératif basé sur ce propagateur a été développé pour la conception d'éléments de Fourier diffractant aux grands angles. Il converge vers des solutions bien mieux optimisées que celles que l'on obtient à l'aide de la méthode IFTA classique, tout en ayant une complexité calculatoire du même ordre. La fabrication d'EODs ayant des dimensions caractéristiques en dessous du micron au cours de mon séjour au KIT en Allemagne et leur caractérisation confirment expérimentalement que notre algorithme itératif scalaire non paraxial peut être utilisé pour le calcul d'EODs de Fourier mince permettant d'atteindre des angles de diffraction allant jusqu'à 37° et peut être même au-delà.

Les désaccords qui demeurent au niveau de l'intensité diffractée entre les simulations et les résultats expérimentaux ont été étudiés et il s'est avéré qu'ils provenaient à la fois de défauts de fabrication et des limites de la TEA dont nous sommes proches. Nous avons montré les limites de la TEA au niveau de la prédiction des symétries de la figure de diffraction en comparant les résultats qu'elle fournit à ceux issus de simulations vectorielles rigoureuses. De plus, notre critère de symétrie de la figure diffraction, pour évaluer la précision des simulations, s'est avéré simple, efficace et rigoureux, malgré les variations d'efficacité de diffraction tant au niveau des simulations qu'expérimentalement. Nous avons développé, optimisé et parallélisé un modèle rigoureux pour la diffraction basé sur la méthode FDTD couplée avec notre propagateur scalaire non paraxial afin de dépasser les limites théoriques de l'approximation TEA et des contraintes au niveau du calcul numérique des méthodes vectorielles classiques. Nous avons proposé un algorithme génétique basé sur ce modèle pour la conception d'EODs épais. Les prochains travaux se concentreront sur la fabrication d'EODs épais et la calibration / optimisation de l'algorithme génétique précédemment cité.

Sachant que nous avons atteint les limites de résolution de notre photo-traceur actuel, nous avons étudié la possibilité de construire un nouveau photo-traceur parallèle reposant sur la technique de 2PP pour s'en servir de moyen de prototypage rapide et rentable de structures submicroniques. Une technique lithographique 2PP parallèle utilisant un EOD est en cours de mise au point à UJF de Grenoble. Un autre système 2PP parallèle utilisant un SLM est également en cours de réalisation à TB en collaboration avec UJF. Les travaux à venir consisteront à permettre au photo-traceur parallèle 2PP de servir à la fabrication rapide de structures haute résolution et à trois dimensions. Par ailleurs, sachant que l'on peut également obtenir de grands angles de diffraction avec des structures micrométriques ayant des profondeurs de gravure engendrant des déphasages $> 2\pi$, nous avons proposé un algorithme génétique utilisant un modèle vectoriel rigoureux pour la conception d'EODs épais. Ainsi, nous espérons pouvoir concevoir et réaliser des composants à plus hautes performances pour la recherche et les applications industrielles qui sont à ce jour inaccessibles.

APPENDIX

A Two-step propagation

A.1 Scaled Fresnel

In the Fresnel approximation, as $\delta_2 = \lambda z/(N\delta_1)$, for a given physical problem with fixed λ and z, we can only control the spacing in the observation plane by adjusting N and/or δ_1. However, there is little real freedom since δ_1 must be the smallest detail of the source field, and N is limited by the computer memory. A two-step propagation has been proposed for the Fresnel propagation [100], in which the field is propagated from the source plane to an intermediate plane z_1 before propagating to the observation plane z, as shown in Fig. 2.7. The scaling parameter $m = \delta_2/\delta_1 = (z - z_1)/z_1$ (and therefore δ_2) can be changed by choosing the intermediate plane.

To simplify the notations, let the spatial vectors in the source plane and observation plane be defined as $\vec{r}_1 = (x_1, y_1), \vec{r}_2 = (x_2, y_2)$, respectively. A simpler solution can be obtained for both cases by rewriting the Fresnel convolution (2.42) as:

$$U(\vec{r}_2; z) \propto \int_{-\infty}^{\infty} U(\vec{r}_1; 0) \exp\left[j\frac{k}{2z}(\vec{r}_2 - \vec{r}_1)^2\right] \mathrm{d}\vec{r}_1 \tag{A.1}$$

The scaling parameter m is introduced by manipulating the exponential term:

$$\begin{aligned}
(\vec{r}_2 - \vec{r}_1)^2 &= {\vec{r}_2}^2 - 2\vec{r}_2 \cdot \vec{r}_1 + {\vec{r}_1}^2 \\
&= \left({\vec{r}_2}^2 + \frac{{\vec{r}_2}^2}{m} - \frac{{\vec{r}_2}^2}{m}\right) - 2\vec{r}_2 \cdot \vec{r}_1 + \left({\vec{r}_1}^2 + m{\vec{r}_1}^2 - m{\vec{r}_1}^2\right) \\
&= \left[\frac{{\vec{r}_2}^2}{m} + \left(1 - \frac{1}{m}\right){\vec{r}_2}^2\right] - 2\vec{r}_2 \cdot \vec{r}_1 + \left[m{\vec{r}_1}^2 + (1-m){\vec{r}_1}^2\right] \\
&= m\left[\left(\frac{\vec{r}_2}{m}\right)^2 - 2\left(\frac{\vec{r}_2}{m}\right) \cdot \vec{r}_1 + {\vec{r}_1}^2\right] + \left(1 - \frac{1}{m}\right){\vec{r}_2}^2 + (1-m){\vec{r}_1}^2 \\
&= m\left(\frac{\vec{r}_2}{m} - \vec{r}_1\right)^2 - \frac{1-m}{m}{\vec{r}_2}^2 + (1-m){\vec{r}_1}^2
\end{aligned} \tag{A.2}$$

The diffracted field on the observation plane is therefore given by:

$$U(\vec{r}_2; z) \propto \int_{-\infty}^{\infty} \left\{U(\vec{r}_1; 0) \exp\left[j\frac{k}{2z}(1-m){\vec{r}_1}^2\right]\right\} \exp\left[j\frac{km}{2z}\left(\frac{\vec{r}_2}{m} - \vec{r}_1\right)^2\right] \mathrm{d}\vec{r}_1 \tag{A.3}$$

which can be evaluated as a convolution in the scaled coordinate system: $\vec{r}_2{}' = \vec{r}_2/m$

$$U(\vec{r}_2;z) \propto IFT\left\{FT\left[U(\vec{r}_1;0)f(\vec{r}_1;z)\right]FT\left[h(\vec{r}_1;z)\right]\right\} \tag{A.4}$$

where $f(\vec{r};z)$ and $h(\vec{r};z)$ are the Impulse Responses

$$f(\vec{r};z) = \exp\left[j\frac{k}{2z}(1-m)\vec{r}^2\right]; h(\vec{r};z) = \exp\left(j\frac{km}{2z}\vec{r}^2\right) \tag{A.5}$$

This simulation is equivalent to the Fresnel propagation (2.44), with an advantage that the sampling distance in the observation plane $\delta_2 = m\delta_1$ can be adjusted freely by choosing the scaling parameter m. Furthermore, the requirement for the phase terms in the source field is:

$$\begin{cases} \delta_r \left|\frac{\partial\phi_f}{\partial r}\right|_{max} \leq \pi \\ \delta_r \left|\frac{\partial\phi_h}{\partial r}\right|_{max} \leq \pi \end{cases} \Rightarrow z \geq \frac{N\delta_1^2}{\lambda}\cdot\max(m,|1-m|) \tag{A.6}$$

which means the validity distance can also be changed by m. The disadvantage compared to the Fresnel propagation is a 3-time higher computational complexity, which is $O(3N^2(1+\log_2 N))$. Notice that the Fresnel convolution (2.47) can be considered as a special case of the scaled Fresnel propagation with m=1.

A.2 Scaled paraxial ASM

It is also possible to derive a scaled paraxial ASM where the field is propagated via an intermiate plane before propagating to the observation plane [101]. This approach is equivalent to the Transfer Function simulation of equation (A.3):

$$U(\vec{r}_2;z) \propto IFT\left\{FT\left[U(\vec{r}_1;0)f(\vec{r}_1;z)\right]H(\vec{f};z)\right\} \tag{A.7}$$

with the Transfer Function $H(\vec{f};z) = \exp(-j\pi\lambda z\vec{f}^2/m)$ is the FT of $h(\vec{r};z)$. This method is called as Angular Spectrum Propagation in [100] but it is only valid in scalar paraxial approximation and should not be confused with the non-paraxial one. The computational complexity is $O(2N^2(1+\log_2 N))$, which is slightly higher than that of the paraxial ASM (2.38). The advantage is that the sampling distance in the observation plane can be adjusted via the scaling parameter $\delta_2 = m\delta_1$. The requirement for the phase terms in the source field is:

$$\begin{cases} \delta_r \left|\frac{\partial\phi_f}{\partial r}\right|_{max} \leq \pi \Rightarrow z \geq \frac{|1-m|N\delta_1^2}{\lambda} \\ \delta_f \left|\frac{\partial\phi_H}{\partial f_x}\right|_{max} \leq \pi \Rightarrow z \leq \frac{mN\delta_1^2}{\lambda} \end{cases} \tag{A.8}$$

which means that the propagation is only valid if $|1-m| \leq m$, or equivalently $m \geq 1/2$. Therefore, by choosing a large scaling parameter m, the scaled paraxial Angular

Spectrum can simulate the diffraction pattern at high propagation distances without increasing N. Attempts have been made to develop scaled versions of the non-paraxial ASM [102, 103] but they are only applicable to high numerical converging elements, which basically means only the near field diffraction pattern.

A.3 Cascaded aperture effect

The difference between a multi-step propagation and a scaled propagation is that the diffracted field is calculated numerically instead of derived analytically (see Eq. (A.3) for example). As the field is propagated via multiple steps, it is affected by virtually intermediate apertures, which are numerical generated by the limited computational regions. For diverging diffraction elements, these apertures may block parts of the diffracted field, which usually contribute to the wide-angle diffraction orders. Fig. A.1 illustrates the cascaded aperture effect in a two-step propagation, where an angular spectrum-like approach (e.g. ASM, RS convolution, etc.) is used for the first propagation step, so the calculation window on the intermediate plane is the same as the calculation window of the field at the DOE plane. Therefore, propagation methods with the same sampling distance at the input and output planes ($\delta_2 = \delta_1$) should not be used in a multi-step propagation to propagate a diverging diffracted field.

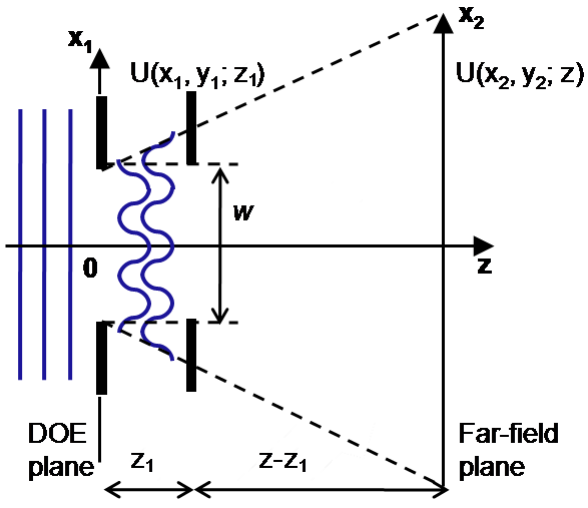


Figure A.1 — Cascaded aperture effect in a numerical multiple-step propagation.

APPENDIX B

Diffracted power estimation

B.1 Vectorial theory

If all electromagnetic components of the diffracted field have been solved, the flow of electromagnetic power can be calculated rigorously by the Poynting vector [152]:

$$\vec{\mathcal{S}} = \vec{\mathcal{E}} \times \vec{\mathcal{H}} \tag{B.1}$$

The direction of power flow is along the direction of the Poynting vector, i.e., orthogonal to both $\vec{\mathcal{E}}$ and $\vec{\mathcal{H}}$. For monochromatic wave, using the expressions (2.5) and (2.6) for the electric and magnetic fields yields:

$$\begin{aligned}\vec{\mathcal{S}} = Re\{\vec{E}e^{-j\omega t}\} \times Re\{\vec{H}e^{-j\omega t}\} &= \frac{1}{2}(\vec{E}e^{-j\omega t} + \vec{E}^* e^{j\omega t}) \times \frac{1}{2}(\vec{H}e^{-j\omega t} + \vec{H}^* e^{j\omega t}) \\ &= \frac{1}{4}(\vec{E} \times \vec{H}^* + \vec{E}^* \times \vec{H} + \vec{E} \times \vec{H}e^{-j2\omega t} + \vec{E} \times \vec{H}e^{j2\omega t})\end{aligned} \tag{B.2}$$

The optical intensity is the magnitude of the time-averaged Poynting vector $I = \left\langle \vec{\mathcal{S}} \right\rangle$. Notice that the terms containing the factors $e^{j2\omega t}$ and $e^{-j2\omega t}$ are fast oscillating compared to the optical cycle and therefore washed out by the averaging process:

$$I = \left\langle \vec{\mathcal{S}} \right\rangle = \frac{1}{4}(\vec{E} \times \vec{H}^* + \vec{E}^* \times \vec{H}) = \frac{1}{2}(\vec{S} + \vec{S}^*) = Re\{\vec{S}\} \tag{B.3}$$

The vector $\vec{S} = \frac{1}{2}\vec{E} \times \vec{H}^*$ is known as the complex Poynting vector. The power density on the observation plane is therefore given as in [7]:

$$p = Re\{\vec{S}\} \cos\left(\vec{k}, \vec{n}\right) \tag{B.4}$$

where $\vec{n}$ is the unit vector pointing into the surface of the detector, and $\vec{k}$ is a unit vector in the direction of power flow. The total power incident on a surface area A is the integral of the power density over the area:

$$P = \iint_A Re\{\vec{S}\} \cos\left(\vec{k}, \vec{n}\right) \, \mathrm{d}x_2 \mathrm{d}y_2 \tag{B.5}$$

For diverging elements where the diffraction pattern is much bigger than the DOE size, the angle between $\vec{k}$ and $\vec{n}$ can be approximated by the diffraction angle, i.e. $\cos\left(\vec{k},\vec{n}\right) = \gamma$. The diffracted power estimation can be further reduced to a simpler expression by using scalar theory.

B.2 Scalar theory

For free-space medium which is isotropic, the monochromatic wave behaves locally as a transverse electromagnetic plane wave (i.e. $\vec{E}$, $\vec{H}$, and $\vec{k}$ form a mutually orthogonal triplet), the electric and magnetic fields can be expressed locally as in [11, 152]:

$$E(\vec{r}) = Re\left\{E_0 \exp\left(-j\vec{k}\cdot\vec{r}\right)\right\} \tag{B.6}$$

$$H(\vec{r}) = Re\left\{H_0 \exp\left(-j\vec{k}\cdot\vec{r}\right)\right\} \tag{B.7}$$

where E_0 and H_0 are locally constant vectors and have complex components. The power flowing in the propagation direction $\vec{k}$ can be expressed as:

$$I = \frac{\vec{E_0}\cdot\vec{E_0^*}}{2\eta} = \frac{|E_0|^2}{2\eta} \tag{B.8}$$

where η is the characteristic impedance of the medium ($\eta = 377\Omega$ in vacuum). Equation (B.5) can therefore be rewritten as:

$$P \propto \iint_A |E_0|^2 \gamma \, \mathrm{d}x_2 \mathrm{d}y_2 \tag{B.9}$$

When $\vec{k}$ is nearly normal to the surface, $\gamma \approx 1$ and the total power P is simply the integral of the power density p over the detector area, which is the paraxial approximation usually used in practice [11, 152]:

$$P \propto \iint_A |E_0|^2 \, \mathrm{d}x_2 \mathrm{d}y_2 \tag{B.10}$$

APPENDIX C

Boundary conditions

The Maxwell's equations require that, in order for the electromagnetic fields to be differentiable, there must be an open neighbourhood around the point they're applied to, i.e. the medium is continuous. The boundary conditions for the electromagnetic field vectors can be derived from the integral forms of Maxwell's equations [17]:

$$\vec{n}_{12} \times (\vec{E}_2 - \vec{E}_1) = 0, \quad \vec{n}_{12} \times (\vec{H}_2 - \vec{H}_1) = \vec{j}_s \tag{C.1}$$

where $\vec{n}_{12}$ is the normal vector from medium 1 to medium 2 and $\vec{j}_s$ is the surface current density between the two media. Therefore the tangential component of $\vec{E}$ is continuous across the interface, and the tangential component of $\vec{H}$ is continuous across the surface if there is no surface current.

On the other hand, we must limit our simulation to a finite region of space, so that it can be simulated on a computer. To determine the field components that are positioned directly at the boundary of the computational domain, the information of the field components outside the simulation region is needed and estimated by choosing a proper boundary condition.

C.1 Perfectly Conducting Boundary

For a totally reflecting material, the perfectly conducting boundary can be used. This is the most simple boundary condition, where the fields are simply forced to be zero on the boundaries, as if the cell were surrounded by a perfectly metallic wall (no absorption and no transmission). The electromagnetic field cannot penetrate the structure, which means the field components are reflected back to the simulation region.

C.2 Floquet-Bloch Periodic Boundary

For finite periodic structures like gratings and photonic crystals, the periodic boundary condition can be applied. With ordinary periodic boundaries in a cell of size L, the

field components satisfy $f(x+L) = f(x)$. The Floquet-Bloch periodicity is a more general boundary condition where $f(x+L) = e^{ik_x L} f(x)$, with k_x is the Bloch wavevector. In general, the structure can be periodic in 3D, with the Bloch wavevector $\vec{k} = (k_x, k_y, k_z)$.

C.3 Perfect Matched Layer

To simulate open boundary conditions, the Perfectly Matched Layer (PML) can be used to absorb all waves incident on them, with no reflections. Strictly speaking, PML is not a boundary condition, but a special absorbing material placed adjacent to the boundaries, which can be set as either of the above boundary conditions. The approach involves surrounding the computational cell with a medium that in theory absorbs without any reflection electromagnetic waves at all frequencies and angles of incidence.

The first implemention of PML absorbing boundary conditions in Finite-Difference Time-Domain method was proposed by Berenger [163]. Berenger showed that it was sufficient to "split" Maxwell's equations into two sets of (unphysical) equations in the absorbing layers, appropriately defined. It was later shown that a similar reflectionless absorbing medium can be constructed as a lossy anisotropic dielectric and magnetic material with "matched" impedance and electrical and magnetic conductivities [164].

Although PML is reflectionless in the theoretical continous system, in the actual discretized system it has some small reflections which make it imperfect [130, 131]. The finite-difference implementation of PML requires the conductivities to be turned on gradually over a distance of a few grid points to avoid numerical reflections from the discontinuity. It is important when using PMLs to make the computational cell sufficiently large otherwise the promixity of the PML to evanescent modes may unnecessarily extract energy from the system and thereby perturb it [75].

Bibliography

[1] J. Turunen and F. Wyrowski, *Diffractive Optics: Industrial and Commercial Applications*. Akademie Verlag, 1997.

[2] B. C. Kress and P. Meyrueis, *Applied Digital Optics: From Micro-optics to Nanophotonics*. Chichester, UK: John Wiley & Sons, Ltd, Oct. 2009.

[3] D. C. O'Shea, T. J. Suleski, A. D. Kathman, and D. W. Prather, *Diffractive Optics: Design, Fabrication, and Test*. SPIE Press, 2004.

[4] L. G. Neto, "Design, fabrication, and characterization of a full complex-amplitude modulation diffractive optical element," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 2, p. 96, Apr. 2003.

[5] J. W. Goodman and A. M. Silvestri, "Some Effects of Fourier-domain Phase Quantization," *IBM Journal of Research and Development*, vol. 14, pp. 478–484, Sept. 1970.

[6] F. Wyrowski, "Diffractive optical elements: iterative calculation of quantized, blazed phase structures," *Journal of the Optical Society of America A*, vol. 7, p. 961, June 1990.

[7] S. Mellin and G. Nordin, "Limits of scalar diffraction theory and an iterative angular spectrum algorithm for finite aperture diffractive optical element design," *Optics Express*, vol. 8, p. 705, June 2001.

[8] U. Mahlab, J. Shamir, and H. J. Caulfield, "Genetic algorithm for optical pattern recognition," *Optics Letters*, vol. 16, p. 648, May 1991.

[9] D. R. Brown and A. D. Kathman, "Multielement diffractive optical designs using evolutionary programming," in *Photonics West '95*, pp. 17–27, International Society for Optics and Photonics, Apr. 1995.

[10] J. Bengtsson, J. Backlund, N. Eriksson, P. Modh, H. Martinsson, J. Vukusic, and A. Larsson, "Diffractive solutions in integrated optics? - OSA Trends in Optics and Photonics," in *Diffractive Optics and Micro-Optics*, vol. 41, (Québec City, Canada), Optical Society of America, 2000.

[11] J. W. Goodman, *Introduction to Fourier Optics*. McGraw-Hill, 1996.

[12] E. Wolf, *Progress in Optics, Volume 28*. Elsevier, 1990.

[13] H. P. Herzig, *Micro-Optics: Elements, Systems And Applications*. CRC Press, 1997.

[14] B. C. Kress and P. Meyrueis, *Digital Diffractive Optics: An Introduction to Planar Diffractive Optics and Related Technology*, vol. 2000. Wiley, 2000.

[15] S. Sinzinger and J. Jahns, *Microoptics*. John Wiley & Sons, 2006.

[16] S. Silver, "Microwave Aperture Antennas and Diffraction Theory," *Journal of the Optical Society of America*, vol. 52, p. 131, Feb. 1962.

[17] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 1999.

[18] D. Gabor, "A New Microscopic Principle," *Nature*, vol. 161, pp. 777–778, May 1948.

[19] D. Gabor, "Microscopy by Reconstructed Wave-Fronts," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 197, pp. 454–487, July 1949.

[20] E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory," *Journal of the Optical Society of America (1917-1983)*, vol. 52, p. 1123, Oct. 1962.

[21] J. N. Mait, "From ink bottles to e-beams: a historical perspective on diffractive optic technology," in *Aerospace/Defense Sensing, Simulation, and Controls*, pp. 75–86, International Society for Optics and Photonics, July 2001.

[22] B. R. Brown and A. W. Lohmann, "Complex spatial filtering with binary masks.," *Applied optics*, vol. 5, pp. 967–9, June 1966.

[23] A. W. Lohmann and D. P. Paris, "Binary fraunhofer holograms, generated by computer.," *Applied optics*, vol. 6, pp. 1739–48, Oct. 1967.

[24] B. R. Brown and A. W. Lohmann, "Computer-generated Binary Holograms," *IBM Journal of Research and Development*, vol. 13, pp. 160–168, Mar. 1969.

[25] L. B. Lesem, P. M. Hirsch, and J. A. Jordan, "The Kinoform: A New Wavefront Reconstruction Device," *IBM Journal of Research and Development*, vol. 13, pp. 150–155, Mar. 1969.

[26] L. B. Lesem, P. M. Hirsch, and J. A. Jordan, "Computer synthesis of Holograms for 3-D display," *Communications of the ACM*, vol. 11, pp. 661–674, Oct. 1968.

[27] R. Gerchberg and O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237 – 246, 1972.

[28] N. C. Gallagher and B. Liu, "Method for Computing Kinoforms that Reduces Image Reconstruction Error.," *Applied optics*, vol. 12, pp. 2328–35, Oct. 1973.

[29] M. T. Gale and K. Knop, "The Fabrication Of Fine Lens Arrays By Laser Beam Writing," in *1983 International Techincal Conference/Europe*, pp. 347–353, International Society for Optics and Photonics, Oct. 1983.

[30] P. P. Clark and C. Londoño, "Production of kinoforms by single point diamond machining," *Optics News*, vol. 15, p. 39_1, Dec. 1989.

[31] S. M. Arnold, "Electron Beam Fabrication Of Computer-Generated Holograms," *Optical Engineering*, vol. 24, p. 245803, Aug. 1985.

[32] G. E. Moore, "Lithography and the future of Moore's law," in *SPIE's 1995 Symposium on Microlithography*, pp. 2–17, International Society for Optics and Photonics, June 1995.

[33] S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Optics Letters*, vol. 19, p. 780, June 1994.

[34] T. A. Klar, S. Jakobs, M. Dyba, A. Egner, and S. W. Hell, "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission," *Proceedings of the National Academy of Sciences*, vol. 97, pp. 8206–8210, July 2000.

[35] S. W. Hell, R. Schmidt, and A. Egner, "Diffraction-unlimited three-dimensional optical nanoscopy with opposing lenses," *Nature Photonics*, vol. 3, pp. 381–387, July 2009.

[36] T. F. Scott, B. A. Kowalski, A. C. Sullivan, C. N. Bowman, and R. R. McLeod, "Two-color single-photon photoinitiation and photoinhibition for subdiffraction photolithography.," *Science (New York, N.Y.)*, vol. 324, pp. 913–7, May 2009.

[37] L. Li, R. R. Gattass, E. Gershgoren, H. Hwang, and J. T. Fourkas, "Achieving lambda/20 resolution by one-color initiation and deactivation of polymerization.," *Science (New York, N.Y.)*, vol. 324, pp. 910–3, May 2009.

[38] J. Fischer and M. Wegener, "Three-dimensional direct laser writing inspired by stimulated-emission-depletion microscopy [Invited]," *Optical Materials Express*, vol. 1, p. 614, Aug. 2011.

[39] J. Fischer and M. Wegener, "Three-dimensional optical laser lithography beyond the diffraction limit," *Laser & Photonics Reviews*, vol. 7, pp. 22–44, Jan. 2013.

[40] H. Kim, B. Yang, and B. Lee, "Iterative Fourier transform algorithm with regularization for the optimal design of diffractive optical elements.," *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 21, pp. 2353–65, Dec. 2004.

[41] B. K. Jennison, J. P. Allebach, and D. W. Sweeney, "Efficient design of direct-binary-search computer-generated holograms," *Journal of the Optical Society of America A*, vol. 8, p. 652, Apr. 1991.

[42] M. Clark, "Enhanced direct-search method for the computer design of holograms using state variables," in *Photonics West '96*, pp. 24–34, International Society for Optics and Photonics, May 1996.

[43] M. A. Seldowitz, J. P. Allebach, and D. W. Sweeney, "Synthesis of digital holograms by direct binary search.," *Applied optics*, vol. 26, pp. 2788–98, July 1987.

[44] S. Teiwes, B. Schillinger, T. Beth, and F. Wyrowski, "Efficient design of paraxial diffractive phase elements with descent search methods," in *Photonics West '95*, pp. 40–49, International Society for Optics and Photonics, Apr. 1995.

[45] M. S. Kim and C. C. Guest, "Simulated annealing algorithm for binary phase only filters in pattern classification.," *Applied optics*, vol. 29, pp. 1203–8, Mar. 1990.

[46] A. G. Kirk and T. J. Hall, "Design of binary computer generated holograms by simulated annealing: coding density and reconstruction error," *Optics Communications*, vol. 94, pp. 491–496, Dec. 1992.

[47] N. Yoshikawa and T. Yatagai, "Phase optimization of a kinoform by simulated annealing.," *Applied optics*, vol. 33, pp. 863–8, Feb. 1994.

[48] F. Wyrowski and O. Bryngdahl, "Iterative Fourier-transform algorithm applied to computer holography," *Journal of the Optical Society of America A*, vol. 5, p. 1058, July 1988.

[49] F. Wyrowski, "Iterative quantization of digital amplitude holograms," *Applied Optics*, vol. 28, p. 3864, Sept. 1989.

[50] S. Yoshida and M. Yamamoto, "Design and evaluation of diffractive optical elements: optimization by using iterative angular spectrum approach and evaluation based on vector diffraction theory," in *SPIE Europe Optics + Optoelectronics*, pp. 73561Z–73561Z–8, International Society for Optics and Photonics, May 2009.

[51] O. Ripoll, V. Kettunen, and H. P. Herzig, "Review of iterative Fourier-transform algorithms for beam shaping applications," *Optical Engineering*, vol. 43, p. 2549, Nov. 2004.

[52] M. V. Kessels, M. El Bouz, R. Pagan, and K. Heggarty, "Versatile stepper based maskless microlithography using a liquid crystal display for direct write of binary and multilevel microstructures," *Journal of Micro/Nanolithography, MEMS and MOEMS*, vol. 6, p. 033002, July 2007.

[53] Karlsruhe Institute of Technology, *Karlsruhe Nano Micro Facility.* Website: www.knmf.kit.edu, 2014.

[54] M. Kessels and K. Heggarty, "Optical proximity correction for a versatile LCD based direct write maskless photoplotter," *Microelectronic Engineering*, vol. 86, pp. 2385–2391, Dec. 2009.

[55] H. Schwarzer, S. Teiwes, and F. Wyrowski, "Non-pixelated design of computer generated diffractive elements for increased diffraction efficiency," in *Diffractive Optics 97*, (Savonlinna), pp. 164–165, EOS, 1997.

[56] H. Aagedal, F. Wyrowski, M. Schmid, J. Müller-Quade, and T. Beth, "Consequence of high resolution lithography for the design in the paraxial domain," in *Diffractive Optics 97*, (Savonlinna), pp. 166–167, EOS, 1997.

[57] F. Wyrowski, R. Hauck, and O. Bryngdahl, "Computer-generated holography: hologram repetition and phase manipulations," *Journal of the Optical Society of America A*, vol. 4, p. 694, Apr. 1987.

[58] F. Wyrowski and O. Bryngdahl, "Speckle-free reconstruction in digital holography," *Journal of the Optical Society of America A*, vol. 6, p. 1171, Aug. 1989.

[59] M. S. Mirotznik, D. W. Prather, J. N. Mait, W. A. Beck, S. Shi, and X. Gao, "Three-Dimensional Analysis of Subwavelength Diffractive Optical Elements with the Finite-Difference Time-Domain Method," *Applied Optics*, vol. 39, p. 2871, June 2000.

[60] P. Gerard, P. Benech, D. Khalil, R. Rimet, and S. Tedjini, "Towards a full vectorial and modal technique for the analysis of integrated optics structures: the Radiation Spectrum Method (RSM)," *Optics Communications*, vol. 140, pp. 128–145, July 1997.

[61] F. Di, Y. Yingbai, J. Guofan, T. Qiaofeng, and H. Liu, "Rigorous electromagnetic design of finite-aperture diffractive optical elements by use of an iterative optimization algorithm," *Journal of the Optical Society of America A*, vol. 20, p. 1739, Sept. 2003.

[62] V. Raulot, P. Gérard, B. Serio, M. Flury, B. Kress, and P. Meyrueis, "Modeling of the angular tolerancing of an effective medium diffractive lens using combined finite difference time domain and radiation spectrum method algorithms," *Optics Express*, vol. 18, p. 17974, Aug. 2010.

[63] V. Raulot, P. Gérard, B. Serio, M. Flury, and P. Meyrueis, "Comparison of the behavior of a subwavelength diffractive lens in TE and TM polarization allowing some nonstandard functions," *Optics Letters*, vol. 36, p. 1194, Mar. 2011.

[64] S. D. Mellin, *Design and analysis of finite-aperture diffractive optical elements*. PhD thesis, University of Alabama in Huntsville, 2001.

[65] G. J. Swanson, "Binary optics technology: Theoretical limits on the diffraction efficiency of multilevel diffractive optical elements," *Technical Report*, vol. 914, Mar. 1991.

[66] H. Zhong, S. Zhang, F. Wyrowski, M. Ahmad, and H. Schweitzer, "Parabasal thin element approximation for the analysis of the diffractive optical elements," in *SPIE MOEMS-MEMS*, p. 897711, International Society for Optics and Photonics, Mar. 2014.

[67] P. Lalanne and E. Silberstein, "Fourier-modal methods applied to waveguide computational problems.," *Optics letters*, vol. 25, pp. 1092–4, Aug. 2000.

[68] M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *Journal of the Optical Society of America*, vol. 71, p. 811, July 1981.

[69] M. G. Moharam, E. B. Grann, D. A. Pommet, and T. K. Gaylord, "Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings," *Journal of the Optical Society of America A*, vol. 12, p. 1068, May 1995.

[70] M. G. Moharam, D. A. Pommet, E. B. Grann, and T. K. Gaylord, "Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach," *Journal of the Optical Society of America A*, vol. 12, p. 1077, May 1995.

[71] J. B. Schneider, *Understanding the Finite-Difference Time-Domain Method.* Website: www.eecs.wsu.edu/˜schneidj/ufdtd, 2010.

[72] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method, Third Edition.* Artech House, 2005.

[73] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas And Propagation*, pp. 302–307, 1966.

[74] A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. Joannopoulos, and S. G. Johnson, "Meep: A flexible free-software package for electromagnetic simulations by the FDTD method," *Computer Physics Communications*, vol. 181, pp. 687–702, Mar. 2010.

[75] Massachusetts Institute of Technology, *MEEP (MIT Electromagnetic Equation Propagation).* Website: ab-initio.mit.edu/wiki/index.php/Meep, 2014.

[76] E. Lalor, "Conditions for the Validity of the Angular Spectrum of Plane Waves," *Journal of the Optical Society of America*, vol. 58, p. 1235, Sept. 1968.

[77] G. C. Sherman and H. J. Bremermiann, "Generalization of the Angular Spectrum of Plane Waves and the Diffraction Transform," *Journal of the Optical Society of America*, vol. 59, p. 146, Feb. 1969.

[78] K. Matsushima, H. Schimmel, and F. Wyrowski, "Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves," *Journal of the Optical Society of America A*, vol. 20, no. 9, p. 1755, 2003.

[79] T. Kozacki, "Numerical errors of diffraction computing using plane wave spectrum decomposition," *Optics Communications*, vol. 281, pp. 4219–4223, Sept. 2008.

[80] D. G. Voelz and M. C. Roggemann, "Digital simulation of scalar optical diffraction: revisiting chirp function sampling criteria and consequences," *Applied Optics*, vol. 48, p. 6132, Nov. 2009.

[81] K. Matsushima and T. Shimobaba, "Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields," *Optics Express*, vol. 17, p. 19662, Oct. 2009.

[82] R. Martínez-Herrero, P. M. Mejías, S. Bosch, and A. Carnicer, "Vectorial structure of nonparaxial electromagnetic beams," *Journal of the Optical Society of America A*, vol. 18, no. 7, p. 1678, 2001.

[83] Y. Zhao, "Vector iterative algorithms for beam shaping," *Optical Engineering*, vol. 42, p. 3080, Nov. 2003.

[84] P. Liu and B. Lü, "The vectorial angular-spectrum representation and Rayleigh-Sommerfeld diffraction formulae," *Optics & Laser Technology*, vol. 39, pp. 741–744, June 2007.

[85] L. CarreteroLópez, M. Pérez Molina, P. Acebal González, S. Blaya Escarré, A. Fimia Gil, R. F. Madrigal, and A. Murciano Cases, "Vectorial Diffraction Analysis of Near-Field Focusing of Perfect Black Fresnel Zone Plates Under Various Polarization States," *J. Lightwave Technol.*, vol. 29, no. 6, pp. 822–829, 2011.

[86] Y. Zhang, C. Zheng, Y. Zhuang, and X. Ruan, "Analysis of near-field subwavelength focusing of hybrid amplitude-phase Fresnel zone plates under radially polarized illumination," *Journal of Optics*, vol. 16, p. 015703, Jan. 2014.

[87] G. C. Sherman, "Application of the Convolution Theorem to Rayleigh's Integral Formulas," *Journal of the Optical Society of America*, vol. 57, p. 546, Apr. 1967.

[88] W. H. Southwell, "Validity of the Fresnel approximation in the near field," *Journal of the Optical Society of America*, vol. 71, p. 7, Jan. 1981.

[89] D. Mendlovic, Z. Zalevsky, and N. Konforti, "Computation considerations and fast algorithms for calculating the diffraction integral," *Journal of Modern Optics*, vol. 44, pp. 407–414, Feb. 1997.

[90] C. A. Balanis, *Antenna Theory: Analysis and Design.* John Wiley & Sons, 2012.

[91] J. E. Harvey, A. Krywonos, and D. Bogunovic, "Tolerance on Defocus Precisely Locates the Far Field (Exactly Where is that Far Field Anyway?)," *Applied Optics*, vol. 41, p. 2586, May 2002.

[92] W. J. Dallas, "Computer-generated holograms," in *The Computer in Optical Research*, vol. 41 of *Topics in Applied Physics*, Berlin/Heidelberg: Springer-Verlag, 1980.

[93] J. P. Perez, *Optique géométrique et ondulatoire: Avec 200 exercices et problèmes résolus.* Masson, 1994.

[94] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition).* Prentice Hall, 2007.

[95] D. G. Voelz, *Computational Fourier Optics: A MATLAB Tutorial (SPIE Tutorial Texts Vol. TT89)*. SPIE Press, 2011.

[96] J. A. C. Veerman, J. J. Rusch, and H. P. Urbach, "Calculation of the Rayleigh-Sommerfeld diffraction integral by exact integration of the fast oscillating factor," *Journal of the Optical Society of America A*, vol. 22, no. 4, p. 636, 2005.

[97] V. Nascov and P. C. Logofatu, "Fast computation algorithm for the Rayleigh-Sommerfeld diffraction formula using a type of scaled convolution," *Applied Optics*, vol. 48, p. 4310, July 2009.

[98] A. Wuttig, M. Kanka, H. J. Kreuzer, and R. Riesenberg, "Packed domain Rayleigh-Sommerfeld wavefield propagation for large targets.," *Optics express*, vol. 18, pp. 27036–47, Dec. 2010.

[99] F. Shen and A. Wang, "Fast-Fourier-transform based numerical integration method for the Rayleigh-Sommerfeld diffraction formula," *Applied Optics*, vol. 45, p. 1102, Feb. 2006.

[100] J. D. Schmidt, *Numerical Simulation of Optical Wave Propagation*. SPIE, 2010.

[101] F. Zhang, I. Yamaguchi, and L. P. Yaroslavsky, "Algorithm for reconstruction of digital holograms with adjustable magnification," *Optics Letters*, vol. 29, p. 1668, July 2004.

[102] S. Odate, C. Koike, H. Toba, T. Koike, A. Sugaya, K. Sugisaki, K. Otaki, and K. Uchikawa, "Angular spectrum calculations for arbitrary focal length with a scaled convolution," *Optics Express*, vol. 19, p. 14268, July 2011.

[103] T. Shimobaba, K. Matsushima, T. Kakue, N. Masuda, and T. Ito, "Scaled angular spectrum method.," *Optics letters*, vol. 37, pp. 4128–30, Oct. 2012.

[104] D. Wang, J. Zhao, F. Zhang, G. Pedrini, and W. Osten, "High-fidelity numerical realization of multiple-step Fresnel propagation for the reconstruction of digital holograms," *Applied Optics*, vol. 47, p. D12, Jan. 2008.

[105] J. E. Harvey, D. Bogunovic, and A. Krywonos, "Aberrations of Diffracted Wave Fields: Distortion," *Applied Optics*, vol. 42, p. 1167, Mar. 2003.

[106] D. C. Cole, E. Barouch, U. Hollerbach, and S. A. Orszag, "Derivation and Simulation of Higher Numerical Aperture Scalar Aerial Images," *Japanese Journal of Applied Physics*, vol. 31, pp. 4110–4119, Dec. 1992.

[107] G.-N. Nguyen, K. Heggarty, P. Gérard, B. Serio, and P. Meyrueis, "Computationally efficient scalar nonparaxial modeling of optical wave propagation in the far-field," *Applied Optics*, vol. 53, p. 2196, Mar. 2014.

[108] J. E. Harvey and R. V. Shack, "Aberrations of diffracted wave fields," *Applied Optics*, vol. 17, p. 3003, Sept. 1978.

[109] J. E. Harvey, "Fourier treatment of near-field scalar diffraction theory," *American Journal of Physics*, vol. 47, p. 974, Nov. 1979.

[110] D. Mas, J. Garcia, C. Ferreira, L. M. Bernardo, and F. Marinho, "Fast algorithms for free-space diffraction patterns calculation," *Optics Communications*, vol. 164, pp. 233–245, June 1999.

[111] E. Hecht, *Optics*. Addison-Wesley, 2002.

[112] S. Kemme, *Microoptics and Nanooptics Fabrication*, vol. 64. CRC Press, 2009.

[113] N. Y. Chang and C. J. Kuo, "Algorithm based on rigorous coupled-wave analysis for diffractive optical element design.," *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 18, pp. 2491–501, Oct. 2001.

[114] L. Alvarez, L. Gómez, and J. R. Sendra, "An Algebraic Approach to Lens Distortion by Line Rectification," *Journal of Mathematical Imaging and Vision*, vol. 35, pp. 36–50, May 2009.

[115] D. A. Gremaux and N. C. Gallagher, "Limits of scalar diffraction theory for conducting gratings.," *Applied optics*, vol. 32, pp. 1948–53, Apr. 1993.

[116] D. A. Pommet, M. G. Moharam, and E. B. Grann, "Limits of scalar diffraction theory for diffractive phase elements," *Journal of the Optical Society of America A*, vol. 11, p. 1827, June 1994.

[117] G.-N. Nguyen, K. Heggarty, A. Bacher, P.-J. Jakobs, D. Häringer, P. Gérard, P. Pfeiffer, and P. Meyrueis, "Iterative scalar nonparaxial algorithm for the design of Fourier phase elements," *Optics Letters*, vol. 39, p. 5551, Sept. 2014.

[118] N. M. Lyndin, O. Parriaux, and A. V. Tishchenko, "Modal analysis and suppression of the Fourier modal method instabilities in highly conductive gratings," *Journal of the Optical Society of America A*, vol. 24, no. 12, p. 3781, 2007.

[119] LightTrans GmbH, *VirtualLab*. Website: www.lighttrans.com, 2014.

[120] V. Liu and S. Fan, "S4 : A free electromagnetic solver for layered periodic structures," *Computer Physics Communications*, vol. 183, pp. 2233–2244, Oct. 2012.

[121] M. A. Herráez, M. A. Gdeisat, D. R. Burton, and M. J. Lalor, "Robust, Fast, and Effective Two-Dimensional Automatic Phase Unwrapping Algorithm Based on Image Decomposition," *Applied Optics*, vol. 41, p. 7445, Dec. 2002.

[122] R. L. Morrison, "Symmetries that simplify the design of spot array phase gratings," *Journal of the Optical Society of America A*, vol. 9, p. 464, Mar. 1992.

[123] J. Francés, C. Neipp, M. Pérez-Molina, and A. Beléndez, "Rigorous interference and diffraction analysis of diffractive optic elements using the finite-difference time-domain method," *Computer Physics Communications*, vol. 181, pp. 1963–1973, Dec. 2010.

[124] J. Muller, G. Parent, G. Jeandel, and D. Lacroix, "Finite-difference time-domain and near-field-to-far-field transformation in the spectral domain: application to scattering objects with complex shapes in the vicinity of a semi-infinite dielectric medium.," *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 28, pp. 868–78, May 2011.

[125] J. F. James, *A Student's Guide to Fourier Transforms: With Applications in Physics and Engineering.* Cambridge University Press, 2011.

[126] E. Goursat, *A Course in Mathematical Analysis Volume 1: Derivatives and Differentials, Definite Integrals, Expansion in Series, and Applications to Geometry.* Dover Publications, 2006.

[127] R. Courant, *Differential and Integral Calculus, Volume 2.* John Wiley & Sons, 2011.

[128] A. Farjadpour, D. Roundy, A. Rodriguez, M. Ibanescu, P. Bermel, J. D. Joannopoulos, S. G. Johnson, and G. W. Burr, "Improving accuracy by subpixel smoothing in the finite-difference time domain," *Optics Letters*, vol. 31, no. 20, p. 2972, 2006.

[129] A. F. Oskooi, C. Kottke, and S. G. Johnson, "Accurate finite-difference time-domain simulation of anisotropic media by subpixel smoothing.," *Optics letters*, vol. 34, pp. 2778–80, Sept. 2009.

[130] A. F. Oskooi, L. Zhang, Y. Avniel, and S. G. Johnson, "The failure of perfectly matched layers, and towards their redemption by adiabatic absorbers," *Optics Express*, vol. 16, p. 11376, July 2008.

[131] P.-R. Loh, A. Oskooi, M. Ibanescu, M. Skorobogatiy, and S. Johnson, "Fundamental relation between phase and group velocity, and application to the failure of perfectly matched layers in backward-wave structures," *Physical Review E*, vol. 79, p. 065601, June 2009.

[132] D. C. O'Shea, "Reduction of the zero-order intensity in binary Dammann gratings.," *Applied optics*, vol. 34, pp. 6533–7, Oct. 1995.

[133] D. J. Robinson and J. B. Schneider, "On the Use of the Geometric Mean in FDTD Near-to-Far-Field Transformations," *IEEE Transactions on Antennas and Propagation*, vol. 55, pp. 3204–3211, Nov. 2007.

[134] A. Taflove, S. G. Johnson, and A. Oskooi, *Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology.* 2013.

[135] Institut Francais de Recherche pour l'Exploitation de la Mer, *Pôle de Calcul Intensif pour la Mer.* Website: www.ifremer.fr/pcim, 2014.

[136] D. A. Pommet, E. B. Grann, and M. G. Moharam, "Effects of process errors on the diffraction characteristics of binary dielectric gratings.," *Applied optics*, vol. 34, pp. 2430–5, May 1995.

[137] G. J. Swanson, "Binary optics technology: The theory and design of multi-level diffractive optical elements," *Technical Report*, Aug. 1989.

[138] T. Shirakawa, K. L. Ishikawa, S. Suzuki, Y. Yamada, and H. Takahashi, "Design of binary diffractive microlenses with subwavelength structures using the genetic algorithm.," *Optics express*, vol. 18, pp. 8383–91, Apr. 2010.

[139] R. Zheng, "Optimization scheme for synthesizing kinoform with genetic algorithm," *Optical Engineering*, vol. 50, p. 097004, Sept. 2011.

[140] J. M. Miller, M. R. Taghizadeh, J. Turunen, and N. Ross, "Multilevel-grating array generators: fabrication error analysis and experiments.," *Applied optics*, vol. 32, pp. 2519–25, May 1993.

[141] J. Jahns, "Dammann Gratings For Laser Beam Shaping," *Optical Engineering*, vol. 28, p. 281267, Dec. 1989.

[142] W. Singer, "Diffractive beamshaping elements at the fabrication limit," *Optical Engineering*, vol. 35, p. 2779, Oct. 1996.

[143] V. Kettunen, J. Simonen, M. Kuittinen, O. Ripoll, and H. P. Herzig, "Diffractive elements designed to suppress unwanted zero order due to surface depth error," in *Diffractive Optics and Micro-Optics*, vol. 75, p. DMD3, Optical Society of America, 2002.

[144] J. Jahns, *Introduction to Micro- and Nanooptics*. John Wiley & Sons, 2012.

[145] A. Albarazanchi, P. Gerard, P. Ambs, and P. Meyrueis, "Design of single layer subwavelength diffractive optical element (G-Fresnel) for spectrum splitting and beam concentration," in *SPIE Photonics Europe*, p. 913126, International Society for Optics and Photonics, May 2014.

[146] Zygo Corporation, *3D Optical Surface Profilers*. Website: www.zygo.com, 2014.

[147] Teem Photonics, *Ultrashort pulsed lasers*. Website: www.teemphotonics.com, 2014.

[148] C.-Y. Liu, *The quantitative analysis in Ni-Ti rotary instrument mechanical properties and the study in large-scale structures manufactured by two-photon polymerization*. PhD thesis, Joseph Fourier University, 2014.

[149] K. Obata, J. Koch, U. Hinze, and B. N. Chichkov, "Multi-focus two-photon polymerization technique based on individually controlled phase modulation.," *Optics express*, vol. 18, pp. 17193–200, Aug. 2010.

[150] S. D. Gittard, A. Nguyen, K. Obata, A. Koroleva, R. J. Narayan, and B. N. Chichkov, "Fabrication of microscale medical devices by two-photon polymerization with multiple foci via a spatial light modulator.," *Biomedical optics express*, vol. 2, pp. 3167–78, Nov. 2011.

[151] G. Vizsnyiczai, L. Kelemen, and P. Ormos, "Holographic multi-focus 3D two-photon polymerization with real-time calculated holograms," *Optics Express*, vol. 22, p. 24217, Sept. 2014.

[152] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. Wiley, 2007.

[153] A. Ovsianikov and B. N. Chichkov, "Two-Photon Polymerization - High Resolution 3D Laser Technology and Its Applications," in *Nanoelectronics and Photonics*, Nanostructure Science and Technology, Springer New York, 2008.

[154] D. S. Correa, L. De Boni, A. J. G. Otuka, V. Tribuzi, and C. R. Mendonça, "Two-Photon Polymerization Fabrication of Doped Microstructures," in *Polymerization*, InTech, Sept. 2012.

[155] W. Haske, V. W. Chen, J. M. Hales, W. Dong, S. Barlow, S. R. Marder, and J. W. Perry, "65 nm feature sizes using visible wavelength 3-D multiphoton lithography," *Optics Express*, vol. 15, no. 6, p. 3426, 2007.

[156] M. Emons, K. Obata, T. Binhammer, A. Ovsianikov, B. N. Chichkov, and U. Morgner, "Two-photon polymerization technique with sub-50 nm resolution by sub-10 fs laser pulses," *Optical Materials Express*, vol. 2, p. 942, June 2012.

[157] Z. Gan, Y. Cao, R. A. Evans, and M. Gu, "Three-dimensional deep sub-diffraction optical beam lithography with 9 nm feature size.," *Nature communications*, vol. 4, p. 2061, Jan. 2013.

[158] H.-B. Sun and S. Kawata, "Two-Photon Photopolymerization and 3D Lithographic Microfabrication," in *Advances in Polymer Science*, vol. 170 of *Advances in Polymer Science*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[159] A. Ostendorf and B. N. Chichkov, "Two-Photon Polymerization: A New Approach to Micromachining," *Photonics Spectra*, 2006.

[160] A. Ovsianikov, M. Farsari, and B. N. Chichkov, "Photonic and Biomedical Applications of the Two-Photon Polymerization Technique," in *Stereolithography*, Boston, MA: Springer, 2011.

[161] Bildsysteme GmbH, *Product Catalog*. Website: www.bbs-bildsysteme.com, 2014.

[162] Thorlabs, *Cage Systems*. Website: www.thorlabs.com, 2014.

[163] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, vol. 114, pp. 185–200, Oct. 1994.

[164] A. Oskooi and S. G. Johnson, "Distinguishing correct from incorrect PML proposals and a corrected unsplit PML for anisotropic, dispersive media," *Journal of Computational Physics*, vol. 230, pp. 2369–2377, Apr. 2011.

Résumé

Cette thèse vise à élargir l'éventail des applications des éléments Optiques Diffractifs (EODs) en développant des modèles, des algorithmes et des techniques de prototypage rapide pour des EODs avec des angles de diffraction > 10°, au delà des limites du modèle scalaire, paraxiale de diffraction. Nous développons un propagateur non-paraxiale scalaire en champ lointain précis et efficace pour surmonter les limites des modèles classiques de la diffraction scalaire. Un algorithme itératif basé sur ce propagateur est développé pour la conception d'éléments de Fourier à grand angle de diffraction. Les résultats expérimentaux confirment que notre propagateur et notre algorithme scalaire, non-paraxiale peuvent être utilisés pour la modélisation et la conception des EODs minces avec angles de diffraction jusqu'à environ 37° et peut-être encore plus élevé.

Nous montrons que les divergences qui subsistent entre la modélisation et l'expérimentation de la puissance diffractée résultent surtout des erreurs de fabrication et par le fait que nous nous approchons de la limite de l'approximation d'un élément mince (TEA - «Thin Element Approximation»). Les limites pratiques de la TEA sont étudiées en comparaison avec les simulations rigoureuses vectorielles. Nous développons, optimisons et parallélisons un modèle de diffraction rigoureuse basée sur la FDTD («Finite-Difference Time-Domain») couplée avec notre propagateur non-paraxiale scalaire pour surmonter les limites de la TEA et les limites de calcul des modèles vectoriels actuels. Un algorithme de conception génétique sur la base de ce modèle est proposé pour la conception des EODs «épaisses», il est actuellement en cours d'étalonnage.

Ces modèles et algorithmes nous ont maintenant conduits à la limite de résolution de notre photo-traceur existant utilisé pour la fabrication des EODs. Nous étudions la possibilité de construire un nouveau phototraceur parallèle basé sur polymérisation à deux photons (2PP) comme un moyen de prototypage rapide et rentable de structures haute résolution (submicroniques). Nous concevons et fabriquons à Télécom Bretagne, des EODs générant une matrice de points lumineux pour paralléliser le processus de fabrication 2PP utilisé à l'Université Joseph Fourier de Grenoble par un facteur de 625. Pour accélérer encore le processus de fabrication 2PP, nous concevons et assemblons un autre phototraceur 2PP parallèle à base d'un modulateur spatial de lumière, qui permet de générer jusqu'à environ 0,5 million de faisceaux parallèles. Ce phototracer est actuellement en cours de développement à Télécom Bretagne.

Mots-clés : Diffraction, Scalaire non-paraxiale, Modélisation vectorielle, Conception et fabrication, 2PP parallèle.

Abstract

This thesis aims to extend the range of Diffractive Optical Element (DOE) applications by developing models, algorithms and rapid prototyping techniques for DOEs with diffraction angles > 10°, which is beyond the limits of scalar paraxial diffraction model. We develop an accurate and efficient scalar non-paraxial far-field propagator to overcome the limits of the conventional scalar diffraction models. An iterative algorithm based on this propagator is then developed for the design of wide-angle Fourier elements. Experimental results confirm that our scalar non-paraxial propagator and design algorithm can be used for the modeling and design of thin Fourier DOEs with diffraction angles up to about 37° and perhaps even higher.

The remaining discrepancies in diffracted power between modeling, design and experiment are then shown to result from both fabrication errors and by the fact that we are approaching the limit of the Thin Element Approximation (TEA). The practical limits of the TEA are investigated by comparison with the rigorous vectorial simulations. We then develop, optimize and parallelize a rigorous diffraction model based on the Finite-Difference Time-Domain method coupled with our scalar non-paraxial propagator to overcome the limits of the TEA and the computational limitations of current vectorial models. A genetic design algorithm based on this model is proposed for the design of thick DOEs and this algorithm is currently being calibrated.

These models and algorithms have now brought us to the resolution limit of our existing photoplotter used for DOE fabrication. Therefore, we investigate the possibility of building a new parallel photoplotter based on Two-Photon Polymerization (2PP) as a way to rapid, cost-efective prototyping of high resolution (submicron) structures. We design and fabricate spot array DOEs at Telecom Bretagne to parallelize the 2PP fabrication process used at Joseph Fourier University in Grenoble by a factor of 625. To further speed up the 2PP fabrication process, another prototype parallel 2PP pho-toplotter using a Spatial Light Modulator, which can generate up to about 0.5 million parallel beams, is designed and is currently being developed at Telecom Bretagne.

Keywords : Diffraction, scalar non-paraxial, Vectorial modeling, Design and fabrication, Parallel 2PP.

TELECOM Bretagne

n° d'ordre : 2014telb0336

Télécom Bretagne

Technopôle Brest-Iroise - CS 83818 - 29238 Brest Cedex 3

Tél : + 33(0) 29 00 11 11 - Fax : + 33(0) 29 00 10 00