FIELD TRACING IN GRADED-INDEX MEDIA

DISSERTATION
FOR THE ACQUISITION OF THE ACADEMIC TITLE
doctor rerum naturalium (Dr. rer. Nat.)

submitted to the Council of the Faculty of Physics and Astronomy
of the *Friedrich-Schiller-Universität Jena*

by M. Sc. Huiying Zhong
born in Shandong, China, on 09 June 1989

Supervisor: Prof. Dr. Frank Wyrowski, Friedrich-Schiller-Universität Jena
Reviewer: Prof. Dr. Karl-Heinz Brenner, Universität Heidelberg
Reviewer: Prof. Dr. Pasi Vahimaa, University of Eastern Finland
Date of the Disputation: 18 February 2021
Dissertation, Friedrich-Schiller-Universität Jena, 2021

ABSTRACT

The graded-index (GRIN) media have been widely used in modern optical systems, such as imaging lenses, optical beam delivery components, and beam shaping elements. Accurate and efficient modeling of electromagnetic fields propagating in GRIN media is essential for developing innovative and high-quality optical products. It is the goal of this thesis.

We first give an overview of several existing modeling techniques, including the rigorous ones to model symmetric GRIN structures, e.g., Mie theory (spherically symmetric) and several beam propagation methods (BPMs) with approxiamations to model general GRIN structures, which are numerically more efficient than rigorous methods. However, many of BPMs are limited by considering only a small variance of the refractive index.

To overcome the limitations, we develop a unified field solver, known as the Runge-Kutta (RK) k-domain field propagation method. It can be used to calculate general input fields propagating through an arbitrary GRIN medium, accurately, without physical approximations, such as the scalar field approximation or the paraxial approximation. We convert Maxwell's equations in spatial (x-) domain into ordinary differential equations (ODEs) in the spatial-frequency (k-) domain, which can be solved iteratively using the RK method. In our numerical calculation, taking advantage of the convolution theorem, the total numerical effort is linear in the number of sampling points.

As many GRIN components are modeled or designed using ray tracing techniques, the field passing through the GRIN components must show negligible diffraction, which satisfies the geometric field assumption. By substituting geometric field ansatz into the ODEs of RK k-domain field propagation method and Maxwell's equations, we obtain two ODEs in x-domain. After solving them by using again RK method, the RK x-domain field propagation method is established. The calculation speed is further improved, even as fast as ray tracing techniques.

The straightforward criteria for selecting the RK k- or x-domain field propagation method is whether the diffraction is essential or not. Together with other field solvers, e.g., free-space propagation operators, we can calculate the electromagnetic field propagating through an optical system composed of GRIN components. As an example, we present several GRIN media related applications, including designing a generalized Luneburg lens with multi-focus and a GRIN bend waveguide. Furthermore, we have developed a workflow to model the thermal-mechanical-optical effects by using software Ansys and VirtualLab Fusion. In Ansys, we model the inhomogeneous temperature distribution and the resulted surface deformation. Then, in VirtualLab Fusion, we interpolate and transfer the data into the refractive index distribution, reconstruct the surface morphology, and simulate the field propagation using the developed solvers.

ZUSAMMENFASSUNG

Gradientenindex-Medien (GRIN-Medien) sind in modernen optischen Systemen wie Abbildungslinsen, Komponenten für die optische Strahlabgabe und Strahlformungselementen weit verbreitet. Eine genaue und effiziente Modellierung der Ausbreitung elektromagnetischer Felder in GRIN-Medien ist für die Entwicklung innovativer und qualitativ hochwertiger optischer Produkte von entscheidender Bedeutung. Es ist das Ziel dieser These.

Wir führen zunächst einige existierende Modellierungstechniken ein, einschließlich der strengen zur Modellierung symmetrischer GRIN-Strukturen, z. B. der Mie-Theorie (sphärisch symmetrisch) und mehrerer Beam-Propagation-Verfahren (BPMs) mit Annäherungen zur Modellierung allgemeiner GRIN-Strukturen, die numerisch effizienter sind als strenge Methoden. Viele BPMs sind jedoch begrenzt, indem nur eine kleine Abweichung des Brechungsindex berücksichtigt wird.

Um die Einschränkung, entwickeln wir einen neuen Feldlöser, der als Runge-Kutta (RK) k-Bereichs-Feld-Propagation-Verfahren bekannt ist. Er kann verwendet werden, um allgemeine Eingabefelder zu berechnen, die sich durch ein beliebiges GRIN-Medium ohne physikalische Näherungen wie die Skalarfeldnäherung oder die paraxiale Näherung ausbreiten. Wir konvertieren Maxwells Gleichungen im räumlichen (x-) Bereich in gewöhnliche Differentialgleichungen (ODEs) im räumlichen Frequenzbereich (k-), die iterativ (durch schrittweise wiederholende Rechengänge) mit der RK-Methode gelöst werden können. Bei unserer numerischen Berechnung unter Verwendung des Faltungssatzes ist der gesamte numerische Aufwand in Bezug auf die Anzahl der Abtastpunkte linear.

Da viele GRIN-Komponenten unter Verwendung von Raytracing-Techniken modelliert oder entworfen werden, muss das durch die GRIN-Komponenten hindurchtretende Feld eine vernachlässigbare Beugung aufweisen, was die Annahme eines geometrischen Feldes erfüllt. Durch Einsetzen des geometrischen Feldansatzes in die ODEs der RK-k-Bereichsmethode und der Maxwellschen Gleichungen erhalten wir zwei ODEs im x-Bereich. Nach dem Lösen durch Einsatz der RK-Methode wird die RK x-Bereichs-Feld-Propagation-Methode eingerichtet. Die Berechnungsgeschwindigkeit wird weiter verbessert, sogar so schnell wie bei Raytracing-Techniken.

Durch Kombination der entwickelten RK k- und x-Bereichs-Methoden mit anderen Feldlösern, z. B. Freiraum-Ausbreitungsoperatoren, können wir die Ausbreitung des elektromagnetischen Feldes durch ein optisches System berechnen, das aus GRIN-Komponenten besteht. Als Beispiel präsentieren wir verschiedene GRIN-Medienanwen-dungen, einschließlich des Entwurfs eines verallgemeinerten Lüneburger Objektivs mit Mehrfachfokus und eines GRIN-Biegewellenleiters. Darüber hinaus haben wir einen Workflow entwickelt, um die thermo-mechanisch-optischen Effekte durch den Einsatz der Software Ansys und VirtualLab Fusion vorzuführen. In Ansys führen wir die inhomogene Temperaturverteilung und die daraus resultierende Oberflächenverformung vor. Anschließend sorgen wir für die Einfügung und Übertragung der Daten in die Brechungsindexverteilung in VirtualLab Fusion, rekonstruieren die Oberflächenmorphologie und simulieren die Feldausbreitung unter Einsatz der entwickelten Löser.

CONTENTS

1 INTRODUCTION

The graded-index, or gradient-index (GRIN) medium denotes a medium having a refractive index that varies smoothly from point to point. GRIN media have been widely used in modern optical systems, such as imaging lenses and optical beam delivery and beam shaping components. **1) Imaging lenses.** In the natural world, GRIN lenses frequently occur in animal eyes, e.g., in the lenses of octopus and human eyes [1], which has inspired us to use GRIN media in constructing imaging systems. The first study of the GRIN lens dates back to 1854 [2], when Maxwell proposed a mathematical model for the GRIN ball lens, known as Maxwell's fisheye, to image a point on the spherical surface of the lens onto its opposite/conjugate point. In 1944, Luneburg described another type of ball lens, known as the Luneburg lens, to image a point object at infinity to a point on the spherical surface. He also presented a generalized ball lens model, known as the generalized Luneburg lens, to image an arbitrary point object to another point (or two points) at arbitrary positions [3–7]. In addition to the ball lens, a flat lens consisting of two parallel planes and a cylindrically symmetric GRIN medium between them was constructed by Wood. The focal length of this lens could be adjusted by varying its thickness [8]. Many other types of GRIN lenses have been explored recently [9–12]. GRIN media offer more degrees of freedom in the design of imaging lenses, which increases image quality and makes an optical system lighter and more compact. **2) Beam delivery components.** GRIN fibers are widely used in telecommunications because of their low-loss and low-modal dispersion [13]. In the fiber optics community, GRIN components are used as fiber couplers or collimators. In the endoscope, GRIN rod lenses not only extend the length of the system, but also work as relay lenses to invert the image by producing an intermediate focal plane [14–16]. At present, many works have been undertaken in the field of conformal transformation optics [17–19] based on the concept of GRIN media, which was probably first mentioned by Luneburg in 1944 [3]. GRIN media are also designed to produce cloak devices, collimators, or bending waveguides [20, 21]. **3) Beam shaping components.** For optical metrology or fabrication, GRIN components are designed to shape the input beam into a uniform line and are commercially available [15]. In recent research, the use of GRIN components to gradually shape an arbitrary coherent beam to another beam profile, e.g., the top-hat, has been explored [22, 23].

Over the years, the above GRIN media applications have been limited by the difficulty in the fabrication process. Recently, however, fabrication techniques for GRIN media have been developed so as to achieve arbitrary index profiles and enlarge the range of variation in the refractive index (more details in Appendix A). These fabrication techniques enable the practical application of GRIN media to meet the demands of modern optical systems for precisely manipulating light.

In optical engineering, the most important question is how to design a GRIN component, or even an optical system that contains GRIN components to achieve different functionalities. However, before the optical design step, the GRIN medium should be modeled, and the light propagating through it should be simulated by reliable methods. We will come back to this point, as this is the task that will be attempted in this thesis.

In addition to the man-made GRIN components, GRIN media may also occur unintentionally. For a hundred years, astronomers have been exploring the effect of atmospheric fluctuations in telescopic images. The most familiar example of such atmospheric fluctuation in our daily lives is the mirage, which is caused by a decrease in the air refractive index near the ground [11]. In high-powered laser experiments, the increased temperature created by optical heating in the optical components gives rise to a GRIN distribution in the media, e.g., a thermal lensing effect. We may ask, how do we simulate light propagation and evaluate the effects of a turbulent atmosphere, a mirage, or thermal lenses?

1.1 OBJECTIVE

How do we simulate light propagating through an optical system containing GRIN components? Light can be represented as an electromagnetic field, and the parameters used to describe the optical properties of GRIN components are the relative permittivity ϵ, which is equal to the square of the refractive index, and the permeability μ *. After solving Maxwell's equations, the electromagnetic fields in the optical system are calculated. This process is referred to as field tracing [24, 25]. However, one might ask, what about representing the light as optical rays, and simulating the ray propagation through the optical system? We will return to this point later.

From a practical point of view, the objective of field tracing is to enable fast physical optics [26], which means calculating the electromagnetic fields accurately and efficiently. To achieve this goal, several tricks are explored. One of the most important suggestions is that it is not advisable to use the universal field solvers to model the entire system, especially when the system scale extends well beyond the working wavelengths. As an example, the finite element method (FEM) requires a spatial sampling distance smaller than half a wavelength [27], which induces very heavy sampling in a three-dimensional optical system, even at a millimeter scale, for a visible light source. Alternatively, fast physical optics breaks the whole system down into pieces, finds specific field solvers for each piece, and connects all the field solvers according to the light path.

The specific field solvers should guarantee the accuracy of the field calculation and work as efficiently as possible. The universal field solvers, e.g., FEM, can be selected as specific solvers when the optical structure is on the scale of several wavelengths. Taking advantage of the symmetry/periodicity of the optical structure, many rigorous field solvers, e.g., the Fourier modal method (FMM) [28] provide accurate results for peri-

* The general description of ϵ and μ is in tensor form as a function of position and frequency in linear optics. We will come to a detailed discussion in Chapter 2.

odic structures, and the optical structure only needs sampling inside one period instead of the whole structure. When the structure or the field shows some specific properties, field solvers with proper mathematical approximations can give results much more efficiently with acceptable accuracy. As an example, as the grating period becomes larger, calculation by FMM becomes slower. However, if the field is a geometric field, the phase variation dominates over the amplitude variation. In this case, the field can be treated locally as a plane wave, known as the local plane wave approximation (LPWA). Geometric field propagations through gratings of large periods can be considered as local interactions between plane waves and plane interfaces, known as local plane interface approximations (LPIAs) [29, 30]. If the local plane interfaces are almost perpendicular to the normal of the element, thin element approximation (TEA) [31] can be used. Specific field solvers for different optical elements have been developed, such as field solvers for free space (homogeneous media) [32–34], plane surfaces with coatings, spherical/aspherical/freeform surfaces [29, 30], diffractive optical elements [31], and crystals [35, 36].

Fast physical optics has well-developed techniques for finding light paths to connect field solvers, either sequentially or non-sequentially, and here we do not delve into the details. Instead, let us come back to the point that we are representing light as optical rays, and calculating ray propagation through an optical system. Ray tracing is an element of fast physical optics. We have mentioned that geometric fields can be locally represented as plane waves: local plane waves can be considered rays with additional field properties, such as amplitude and polarization. Mathematically, a homeomorphic Fourier transform produces a spatially smooth connection between electromagnetic fields and optical rays [33].

Which specific field solvers can be selected to simulate GRIN components? A few field solvers for GRIN media have been discussed in the literature [37–43], however, there has been no systematic study of field solvers within the framework of fast physical optics. Thus, the major goal of this thesis is to explore different field solvers for GRIN components, and to develop new field solvers where necessary. Finally, a systematic workflow for field solver selection is proposed for the efficient simulation of optical systems with GRIN components.

1.2 SCOPE OF THIS THESIS

In Chapter 2, we begin with a discussion of fundamental theory, i.e., Maxwell's equations, and further derivation, e.g., second-order wave equations and boundary conditions. Then, we discuss several relevant field solvers, such as the beam propagation method (BPM) and the Mie scattering theory. The concepts and capabilities of those relevant field solvers, including their brief mathematical derivations, are reviewed based on the following three aspects.

- Input fields — what kind of input fields the solver can be used for, e.g., for an ideal plane wave, geometric field, or general field

- GRIN media — what kind of GRIN distribution the solver can be used with, i.e., spherically/cylindrically symmetric, periodic, or general GRIN distributions
- Modeling scope — to what extent the solver can include, such as the optical surface effects between the surrounding medium and the GRIN medium, or applicability to non-paraxial fields, polarization crosstalk, diffraction, reflection, or evanescent waves generated inside GRIN media.

After reviewing the existing solvers, we develop a new field solver to include multiple functionalities and to merge the capabilities of several other solvers. In Chapter 3, we derive one field solver directly from Maxwell's equations that is computationally efficient for general input fields, and for a general GRIN distribution in which reflection and evanescent waves can be ignored. We sequentially present the mathematical derivation, the numerical discussion, and the validity tests. The mathematical derivation produces an ordinary differential equation (ODE) in the angular spectrum domain (k-domain), which can be solved using the Runge-Kutta method, so we name this solver the Runge-Kutta k-domain field propagation method (RKK-FPM). In the k-domain, the convolution between the angular spectrum of ϵ and that of the field components needs to be calculated. To reduce the numerical effort, in the RKK-FPM we calculate this convolution using the Fourier transforms and computing the multiplication in the spatial domain (x-domain), based on the convolution theorem.

What happens if we restrict the field to be geometric in the RKK-FPM? As most GRIN lenses and GRIN beam shaping elements are still modeled or designed using ray tracing techniques, a field passing through the GRIN components must have negligible diffraction, or, thus, be known as a geometric field. In Chapter 4, we study specific field solvers for geometric fields, beginning with the ODE of the RKK-FPM. We replace the fast Fourier transform with the specific transform for geometric fields, known as the homeomorphic Fourier transform [33]. Then, from the fundamental theory of ray optics, the eikonal equation is derived. Substituting the geometric field ansatz and the eikonal equation into Maxwell's equations, we obtain another set of ODEs represented in the x-domain. We solve them by using our developed RK x-domain field propagation method (RKX-FPM).

In Chapter 5 we present some numerical examples using the RKK- and RKX-FPMs for GRIN components. The first example designs and simulates generalized Luneburg lenses. The mathematical description was given by Luneburg [3] and was analytically solved by Sochacki [5–7]. We implement the formula and extend it to design a generalized multi-focus Luneburg lens. We then simulate field tracing through the whole system, from the spherical input field to the focal planes. The second example models a bending waveguide, designed using conformal transformation optics. The third example simulates the thermal-mechanical-optic effect. We first use an analytical model to simulate a thermal lens inside a laser system. Then we use the software Ansys [44] to model the thermal-mechanical effect.

2

REVIEW OF MODELING TECHNIQUES FOR GRADED-INDEX MEDIA

In this chapter, we begin with a discussion of the fundamental theory, i.e., Maxwell's equations, and further derivation, e.g., second-order wave equations and boundary conditions. We then discuss several existing field solvers for GRIN media, including brief mathematical derivations of the solvers, as well as their capabilities.

Many graded-index (GRIN) components show high symmetry. GRIN lenses, such as the generalized Luneburg lens and Maxwell's fisheye, are spherically symmetric, having a relative permittivity ϵ that only varies radially. The widely used GRIN fiber is cylindrically symmetric, and its ϵ changes radially across its cross section, and it is invariant along its cylindrical axis. Taking advantage of the symmetry of the structure, specific field solvers can be developed by solving Maxwell's equations in the proper coordinate systems. Periodic GRIN components, e.g., the holographic grating, can be modeled using the Fourier modal method (FMM). To model general GRIN distribution, a series of beam propagation methods are developed. We select several solvers here, and show their brief mathematical derivations and their capabilities based on the following three aspects.

- Input fields — what kind of input fields the solver can be used for, e.g., for an ideal plane wave, geometric field, or general field
- GRIN media — which kind of GRIN distribution the solver can be used with, i.e., spherically/cylindrically symmetric, periodic, or general GRIN distributions.
- Modeling scope - to what extend the solver can include, such as the optical surface effects between the surrounding medium and the GRIN medium, or applicability to non-paraxial fields, polarization crosstalk, diffraction, reflection, or evanescent waves generated inside GRIN media.

2.1 FUNDAMENTAL THEORY

2.1.1 *Maxwell's Equations*

In physical optics, light is represented as electromagnetic fields having behavior that is governed by Maxwell's equations. All field tracing algorithm derivations begin with these equations. Maxwell's equations can be represented in both integral and differential form. In this thesis, as the GRIN media is differentiable, we work with the differential form.

2.1.1.1 *Mathematical representation of electromagnetic fields*

Electromagnetic field quantities are real-valued quantities in the spatial and time domain. Following the book Fundamentals of Photonics [45], Maxwell's equation for the real fields are written as

$$\begin{aligned} \boldsymbol{\nabla} \times \bar{\boldsymbol{E}}^{(\mathrm{r})}(\boldsymbol{r},t) &= -\partial_t \bar{\boldsymbol{B}}^{(\mathrm{r})}(\boldsymbol{r},t), \\ \boldsymbol{\nabla} \times \bar{\boldsymbol{H}}^{(\mathrm{r})}(\boldsymbol{r},t) &= \partial_t \bar{\boldsymbol{D}}^{(\mathrm{r})}(\boldsymbol{r},t) + \bar{\boldsymbol{j}}^{(\mathrm{r})}(\boldsymbol{r},t), \\ \boldsymbol{\nabla} \cdot \bar{\boldsymbol{D}}^{(\mathrm{r})}(\boldsymbol{r},t) &= \bar{\rho}^{(\mathrm{r})}(\boldsymbol{r},t), \\ \boldsymbol{\nabla} \cdot \bar{\boldsymbol{B}}^{(\mathrm{r})}(\boldsymbol{r},t) &= 0. \end{aligned} \tag{2.1}$$

In Eq. (2.1) $\bar{\boldsymbol{E}}^{(\mathrm{r})}(\boldsymbol{r},t)$ and $\bar{\boldsymbol{H}}^{(\mathrm{r})}(\boldsymbol{r},t)$ are three-dimensional (3D) field vectors that denote the electric field and magnetic field, respectively, at position $\boldsymbol{r}$ (3D vector) and at time t. The bar over the field symbols indicates a quantity in the spatial and time domain, and $^{(\mathrm{r})}$ denotes that the quantity is real-valued in the spatial and time domain. $\bar{\boldsymbol{D}}^{(\mathrm{r})}(\boldsymbol{r},t)$ and $\bar{\boldsymbol{B}}^{(\mathrm{r})}(\boldsymbol{r},t)$ are auxiliary fields, known as dielectric displacement and magnetic induction,

$$\begin{aligned} \bar{\boldsymbol{D}}^{(\mathrm{r})}(\boldsymbol{r},t) &= \epsilon_0 \bar{\boldsymbol{E}}^{(\mathrm{r})}(\boldsymbol{r},t) + \bar{\boldsymbol{P}}^{(\mathrm{r})}(\boldsymbol{r},t) \\ \bar{\boldsymbol{B}}^{(\mathrm{r})}(\boldsymbol{r},t) &= \mu_0 \bar{\boldsymbol{H}}^{(\mathrm{r})}(\boldsymbol{r},t) + \bar{\boldsymbol{M}}^{(\mathrm{r})}(\boldsymbol{r},t), \end{aligned} \tag{2.2}$$

with $\bar{\boldsymbol{P}}^{(\mathrm{r})}(\boldsymbol{r},t)$ and $\bar{\boldsymbol{M}}^{(\mathrm{r})}(\boldsymbol{r},t)$ denoting the electric and magnetic polarizations, respectively.

In this thesis, we assume $\bar{\rho}^{(\mathrm{r})}(\boldsymbol{r},t) = 0$ and $\bar{\boldsymbol{j}}^{(\mathrm{r})}(\boldsymbol{r},t) = 0$; we are only interested in non-magnetizable material, which means $\bar{\boldsymbol{M}}^{(\mathrm{r})}(\boldsymbol{r},t) = 0$.

Fields can be represented and solved not only in the spatial and time domain, but also in the conjugates of this domain. The mathematical tool for connecting domains is the Fourier transform between the time and the frequency domains.
The one-dimensional (1D) Fourier transform from the time to the frequency domains is defined as

$$\boldsymbol{f}(\boldsymbol{r},\omega) = \mathcal{F}_\omega \bar{\boldsymbol{f}}(\boldsymbol{r},t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}t \; \bar{\boldsymbol{f}}(\boldsymbol{r},t) \exp(\mathrm{i}\omega t)\,, \tag{2.3}$$

with $\boldsymbol{f}$ representing one of the 3D field vectors.
The inverse transform is defined as

$$\bar{\boldsymbol{f}}_\ell(\boldsymbol{r},t) = \mathcal{F}_\omega^{-1} \boldsymbol{f}(\boldsymbol{r},\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}\omega \; \boldsymbol{f}(\boldsymbol{r},\omega) \exp(-\mathrm{i}\omega t)\,. \tag{2.4}$$

In substituting Eq. (2.4) into Eqs. (2.1,2.2), the field vectors $\bar{\boldsymbol{E}}^{(\mathrm{r})}$, $\bar{\boldsymbol{H}}^{(\mathrm{r})}$, $\bar{\boldsymbol{D}}^{(\mathrm{r})}$, $\bar{\boldsymbol{B}}^{(\mathrm{r})}$ and $\bar{\boldsymbol{P}}^{(\mathrm{r})}$ can be represented by the field in the spatial and frequency domains. As an example,

$$\begin{aligned} 0 &= \boldsymbol{\nabla} \times \tfrac{1}{\sqrt{2\pi}} \textstyle\int_{-\infty}^{+\infty} \mathrm{d}\omega \; \boldsymbol{E}^{(\mathrm{r})}(\boldsymbol{r},\omega) \exp(-\mathrm{i}\omega t) + \partial_t \tfrac{1}{\sqrt{2\pi}} \textstyle\int_{-\infty}^{+\infty} \mathrm{d}\omega \; \boldsymbol{B}^{(\mathrm{r})}(\boldsymbol{r},\omega) \exp(-\mathrm{i}\omega t) \\ &= \tfrac{1}{\sqrt{2\pi}} \textstyle\int_{-\infty}^{+\infty} \mathrm{d}\omega [\boldsymbol{\nabla} \times \boldsymbol{E}^{(\mathrm{r})}(\boldsymbol{r},\omega) + \boldsymbol{B}^{(\mathrm{r})}(\boldsymbol{r},\omega) \partial_t \exp(-\mathrm{i}\omega t)] \exp(-\mathrm{i}\omega t) \\ &= \tfrac{1}{\sqrt{2\pi}} \textstyle\int_{-\infty}^{+\infty} \mathrm{d}\omega [\boldsymbol{\nabla} \times \boldsymbol{E}^{(\mathrm{r})}(\boldsymbol{r},\omega) - \mathrm{i}\omega \boldsymbol{B}^{(\mathrm{r})}(\boldsymbol{r},\omega)] \exp(-\mathrm{i}\omega t). \end{aligned} \tag{2.5}$$

Because of the orthogonality of the exponential basis function $\exp(-\mathrm{i}\omega t)$, the integral equals zero only when

$$\boldsymbol{\nabla} \times \boldsymbol{E}^{(\mathrm{r})}(\boldsymbol{r},\omega) = \mathrm{i}\omega \boldsymbol{B}^{(\mathrm{r})}(\boldsymbol{r},\omega).$$

Similary, Eqs. (2.1) and (2.2) can be rewritten as

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{E}^{(\mathrm{r})}(\boldsymbol{r},\omega) &= \mathrm{i}\omega \boldsymbol{B}^{(\mathrm{r})}(\boldsymbol{r},\omega), \\ \boldsymbol{\nabla} \times \boldsymbol{H}^{(\mathrm{r})}(\boldsymbol{r},\omega) &= -\mathrm{i}\omega \boldsymbol{D}^{(\mathrm{r})}(\boldsymbol{r},\omega), \\ \boldsymbol{\nabla} \cdot \boldsymbol{D}^{(\mathrm{r})}(\boldsymbol{r},\omega) &= 0, \\ \boldsymbol{\nabla} \cdot \boldsymbol{B}^{(\mathrm{r})}(\boldsymbol{r},\omega) &= 0, \end{aligned} \tag{2.6}$$

and

$$\begin{aligned} \boldsymbol{D}^{(\mathrm{r})}(\boldsymbol{r},\omega) &= \epsilon_0 \boldsymbol{E}^{(\mathrm{r})}(\boldsymbol{r},\omega) + \boldsymbol{P}^{(\mathrm{r})}(\boldsymbol{r},\omega) \\ \boldsymbol{B}^{(\mathrm{r})}(\boldsymbol{r},\omega) &= \mu_0 \boldsymbol{H}^{(\mathrm{r})}(\boldsymbol{r},\omega). \end{aligned} \tag{2.7}$$

Note that the bar over each field vector symbol, which signifies that the field is in the vector's frequency domain, is removed, while $^{(\mathrm{r})}$ remains to indicate that the field is the Fourier counterpart of the corresponding real-valued fields in the spatial and time domain. From the mathematical property of the Fourier transform, $\boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)$, the Fourier counterpart of a real-valued function, $\bar{\boldsymbol{F}}^{(\mathrm{r})}(\boldsymbol{r},t)$, is Hermitian symmetric, i.e.,

$$\boldsymbol{F}^{(\mathrm{r})\star}(\boldsymbol{r},\omega) = \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},-\omega), \tag{2.8}$$

with $\boldsymbol{F}^{(\mathrm{r})\star}$ as the complex conjugate of $\boldsymbol{F}^{(\mathrm{r})}$. Therefore, the field vectors at negative frequency contain no more information than the ones at positive frequency. In order to simplify the mathematical procedures, the complex representation of field vectors is defined as

$$\boldsymbol{F}(\boldsymbol{r},\omega) := \begin{cases} 2\boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega) & \text{if } \omega \geq 0\,, \\ 0 & \text{else}\,. \end{cases} \tag{2.9}$$

Replacing $\boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)$ with the new field definition $\boldsymbol{F}(\boldsymbol{r},\omega)$ in Eqs. (2.6) and (2.7), we obtain

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r},\omega) &= \mathrm{i}\omega \boldsymbol{B}(\boldsymbol{r},\omega), \\ \boldsymbol{\nabla} \times \boldsymbol{H}(\boldsymbol{r},\omega) &= -\mathrm{i}\omega \boldsymbol{D}(\boldsymbol{r},\omega), \\ \boldsymbol{\nabla} \cdot \boldsymbol{D}(\boldsymbol{r},\omega) &= 0, \\ \boldsymbol{\nabla} \cdot \boldsymbol{B}(\boldsymbol{r},\omega) &= 0, \end{aligned} \tag{2.10}$$

and

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{r},\omega) &= \epsilon_0 \boldsymbol{E}(\boldsymbol{r},\omega) + \boldsymbol{P}(\boldsymbol{r},\omega) \\ \boldsymbol{B}(\boldsymbol{r},\omega) &= \mu_0 \boldsymbol{H}(\boldsymbol{r},\omega). \end{aligned} \tag{2.11}$$

Let's explore the relation between $\bar{\boldsymbol{F}}^{(\mathrm{r})}(\boldsymbol{r},t)$ and $\bar{\boldsymbol{F}}(\boldsymbol{r}.t)$ by substituting Eqs. (2.8) and (2.9) into the Fourier transform (2.4),

$$\begin{aligned}
\bar{\boldsymbol{F}}^{(\mathrm{r})}(\boldsymbol{r},t) &= \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)\exp(-\mathrm{i}\omega t)\\
&= \tfrac{1}{\sqrt{2\pi}}\left[\int_{-\infty}^{0}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)\exp(-\mathrm{i}\omega t)+\int_{0}^{+\infty}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)\exp(-\mathrm{i}\omega t)\right]\\
&= \tfrac{1}{\sqrt{2\pi}}\left[\int_{0}^{+\infty}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},-\omega)\exp(\mathrm{i}\omega t)+\int_{0}^{+\infty}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)\exp(-\mathrm{i}\omega t)\right]\\
&= \tfrac{1}{\sqrt{2\pi}}\left[\int_{0}^{+\infty}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})\star}(\boldsymbol{r},\omega)\exp(\mathrm{i}\omega t)+\int_{0}^{+\infty}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)\exp(-\mathrm{i}\omega t)\right]\\
&= \tfrac{1}{\sqrt{2\pi}}\left\{\left[\int_{0}^{+\infty}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)\exp(-\mathrm{i}\omega t)\right]^{\star}+\int_{0}^{+\infty}\mathrm{d}\omega\ \boldsymbol{F}^{(\mathrm{r})}(\boldsymbol{r},\omega)\exp(-\mathrm{i}\omega t)\right\}\\
&= \tfrac{1}{2}\left\{\left[\mathcal{F}_{\omega}^{-1}\boldsymbol{F}(\boldsymbol{r},\omega)\right]^{\star}+\mathcal{F}_{\omega}^{-1}\boldsymbol{F}(\boldsymbol{r},\omega)\right\}\\
&= \mathbb{R}\mathrm{e}\left\{\bar{\boldsymbol{F}}(\boldsymbol{r},t)\right\}.
\end{aligned} \tag{2.12}$$

Eq. (2.12) shows that the real-value field quantities $\bar{\boldsymbol{F}}^{(\mathrm{r})}(\boldsymbol{r},t)$ can be obtained from the real part of the complex-valued ones. On this basis, we can now carry out the mathematical treatment of physical optics by representing light objects with complex-valued field quantities. Maxwell's equations and the constitutive equations should be similar to Eqs. (2.1) and (2.2), but with all the field symbols replaced by $\bar{\boldsymbol{F}}(\boldsymbol{r},t)$,

$$\begin{aligned}
\boldsymbol{\nabla}\times\bar{\boldsymbol{E}}(\boldsymbol{r},t) &= -\partial_t\bar{\boldsymbol{B}}(\boldsymbol{r},t),\\
\boldsymbol{\nabla}\times\bar{\boldsymbol{H}}(\boldsymbol{r},t) &= \partial_t\bar{\boldsymbol{D}}(\boldsymbol{r},t),\\
\boldsymbol{\nabla}\cdot\bar{\boldsymbol{D}}(\boldsymbol{r},t) &= 0,\\
\boldsymbol{\nabla}\cdot\bar{\boldsymbol{B}}(\boldsymbol{r},t) &= 0,
\end{aligned} \tag{2.13}$$

and

$$\begin{aligned}
\bar{\boldsymbol{D}}(\boldsymbol{r},t) &= \epsilon_0\bar{\boldsymbol{E}}(\boldsymbol{r},t)+\bar{\boldsymbol{P}}(\boldsymbol{r},t)\\
\bar{\boldsymbol{B}}(\boldsymbol{r},t) &= \mu_0\bar{\boldsymbol{H}}(\boldsymbol{r},t).
\end{aligned} \tag{2.14}$$

As a short summary, we start with the real-valued representation of electromagnetic field quantities in the spatial and time domain (2.1) and end with the complex-valued representation (2.13). With a Fourier transform between the time and frequency domains, we can determine the relation between both representations of a field, as in Eqs. (2.9) and (2.12). Maxwell's equations, as well as the constitutive equations for the complex-valued field quantities in the spatial and time domain are shown in Eqs. (2.13) and (2.14), while those in the spatial and frequency domains are Eqs. (2.10) and (2.11).

2.1.1.2 *Constitutive equations*

The relation between dielectric displacement $\bar{\boldsymbol{D}}$ and electric field $\bar{\boldsymbol{E}}$, as well as that between magnetic induction $\bar{\boldsymbol{B}}$ and magnetic field $\bar{\boldsymbol{H}}$ are used to describe the iteration between field and medium. In general, the mathematical relation can be complicated, but here we use the linear response theory for an isotropic medium, i.e.,

$$\bar{\boldsymbol{P}}(\boldsymbol{r},t) = \frac{\epsilon_0}{\sqrt{2\pi}} \int_{-\infty}^{t} \mathrm{d}t' \bar{R}(\boldsymbol{r}, t-t') \bar{\boldsymbol{E}}(\boldsymbol{r},t'). \tag{2.15}$$

with $\bar{R}(\boldsymbol{r},t)$ as the linear response function. Moreover, the effect is not instantaneous. Using the inverse Fourier transform, Eq. (2.4), to represent $\bar{R}(\boldsymbol{r},t)$ in Eq. (2.15), i.e.,

$$\bar{R}(\boldsymbol{r},t) = \mathcal{F}_{\omega}^{-1}\chi(\boldsymbol{r},\omega), \tag{2.16}$$

it is easy to derive the relation in the frequency domain, i.e.,

$$\boldsymbol{P}(\boldsymbol{r},\omega) = \epsilon_0 \chi(\boldsymbol{r},\omega)\boldsymbol{E}(\boldsymbol{r},\omega). \tag{2.17}$$

In Eqs. (2.16) and (2.17), $\chi(\boldsymbol{r},\omega)$ is the electric susceptibility. Defining the relative permittivity as

$$\epsilon(\boldsymbol{r},\omega) = 1 + \chi(\boldsymbol{r},\omega). \tag{2.18}$$

$\epsilon(\boldsymbol{r},\omega)^*$ is the parameter for describing the optical medium. For GRIN media, $\epsilon(\boldsymbol{r},\omega)$ is smoothly changed versus $\boldsymbol{r}$. In applications, another parameter for describing the optical medium is the refractive index $n(\boldsymbol{r},\omega) := \sqrt{\epsilon(\boldsymbol{r},\omega)}$. In this thesis, we mainly use $\epsilon(\boldsymbol{r},\omega)$.

The constitutive equation in the spatial and the frequency domains can be rewritten as

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{r},\omega) &= \epsilon_0 \epsilon(\boldsymbol{r},\omega)\boldsymbol{E}(\boldsymbol{r},\omega) \\ \boldsymbol{B}(\boldsymbol{r},\omega) &= \mu_0 \boldsymbol{H}(\boldsymbol{r},\omega). \end{aligned} \tag{2.19}$$

Substituting Eq. (2.19) into Eq. (2.10), we obtain Maxwell's equation for the space and frequency domains, which is the most frequently used equation in this thesis

$$\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r},\omega) = \mathrm{i}\omega\mu_0 \boldsymbol{H}(\boldsymbol{r},\omega), \tag{2.20}$$

$$\boldsymbol{\nabla} \times \boldsymbol{H}(\boldsymbol{r},\omega) = -\mathrm{i}\omega\epsilon_0 \epsilon(\boldsymbol{r},\omega)\boldsymbol{E}(\boldsymbol{r},\omega), \tag{2.21}$$

$$\boldsymbol{\nabla} \cdot [\epsilon(\boldsymbol{r},\omega)\boldsymbol{E}(\boldsymbol{r},\omega)] = 0, \tag{2.22}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{H}(\boldsymbol{r},\omega) = 0, \tag{2.23}$$

Note that by solving Eqs. (2.20-2.23), we obtain the distribution of electromagnetic fields with a single wavelength/frequency (ω) in an inhomogeneous isotropic medium. However, this does not mean we cannot deal with chromatic light, because chromatic light is composed of several ωs with no crosstalk in the field of linear optics. We could solve each ω separately, and then reconstruct the output fields.

Maxwell's equations clearly present the relation between the electric field, magnetic field, and optical medium. Modeling techniques can be physically and mathematically derived from these equations.

---

* In the notation of some studies, they use ϵ_r to represent the relative permittivity. In this work, we omit $_r$ to simplify the notation.

2.1.2 *Second-Order Field Equations*

The further derivation of Maxwell's equations (2.20) to (2.23) to decouple $\boldsymbol{E}$ and $\boldsymbol{H}$ is done by replacing $\boldsymbol{H}(\boldsymbol{r},\omega)$ in Eq. (2.21) by Eq. (2.20), or replacing $\boldsymbol{E}(\boldsymbol{r},\omega)$ in Eq. (2.20) by Eq. (2.21). Then we obtain

$$\Delta\boldsymbol{E}(\boldsymbol{r},\omega)+\frac{\omega^2}{c^2}\epsilon(\boldsymbol{r},\omega)\boldsymbol{E}(\boldsymbol{r},\omega)=-\boldsymbol{\nabla}[\boldsymbol{\nabla}\ln\epsilon(\boldsymbol{r},\omega)\cdot\boldsymbol{E}(\boldsymbol{r},\omega)], \tag{2.24}$$

$$\Delta\boldsymbol{H}(\boldsymbol{r},\omega)+\frac{\omega^2}{c^2}\epsilon(\boldsymbol{r},\omega)\boldsymbol{H}(\boldsymbol{r},\omega)=-\boldsymbol{\nabla}\ln\epsilon(\boldsymbol{r},\omega)\times[\boldsymbol{\nabla}\times\boldsymbol{H}(\boldsymbol{r},\omega)], \tag{2.25}$$

with $c:=(\epsilon_0\mu_0)^{-1/2}$ as the speed of light in a vacuum.

In a homogeneous medium with relative permittivity $\epsilon(\boldsymbol{r},\omega)$, the wave equations can be written as

$$\Delta\boldsymbol{E}(\boldsymbol{r},\omega)+k_0^2\epsilon(\omega)\boldsymbol{E}(\boldsymbol{r},\omega)=0, \tag{2.26}$$

$$\Delta\boldsymbol{H}(\boldsymbol{r},\omega)+k_0^2\epsilon(\omega)\boldsymbol{H}(\boldsymbol{r},\omega)=0, \tag{2.27}$$

with vacuum wave number $k_0:=\dfrac{\omega}{c}$.

In further discussion, we concentrate on the calculation of an electromagnetic field with a single ω, and ω in further equations will be neglected. However, please note that we still consider the dispersion property of a medium, and that the chromatic input field can be decomposed to fields with different ω and calculated separately.

2.1.3 *Boundary Condition at a Surface of Discontinuity*

Although this work models a GRIN medium having a permittivity $\epsilon(\boldsymbol{r},\omega)$ that is continuous, we still need boundary conditions in two cases:

1. Modeling electromagnetic fields through the interface between the surrounding medium and the GRIN medium;
2. Discretizing the GRIN into a sufficient number of layers separated by surfaces of discontinuities and modeling the electromagnetic field through the layered structure.

For a charge-free medium, the boundary condition at the interface between media with permittivity ϵ_1 and ϵ_2 is

1. The tangential component of the electric field is continuous

$$\hat{\boldsymbol{e}}(\boldsymbol{r})\times[\boldsymbol{E}_1(\boldsymbol{r})-\boldsymbol{E}_2(\boldsymbol{r})]=0, \tag{2.28}$$

with $\hat{\boldsymbol{e}}(\boldsymbol{r})$ denoting the normal vector of the surface at position $\boldsymbol{r}$, and $\boldsymbol{E}_i(\boldsymbol{r})$, $i=1,2$ denoting the electric fields in medium i.

2. The tangential component of the magnetic field is continuous

$$\hat{\boldsymbol{e}}(\boldsymbol{r}) \times [\boldsymbol{H}_1(\boldsymbol{r}) - \boldsymbol{H}_2(\boldsymbol{r})] = 0, \tag{2.29}$$

with $\boldsymbol{H}_i(\boldsymbol{r}),\ i = 1, 2$ denoting the magnetic fields in medium i.

3. The normal component of the displacement is continuous

$$\hat{\boldsymbol{e}}(\boldsymbol{r}) \cdot [\epsilon_1(\boldsymbol{r})\boldsymbol{E}_1(\boldsymbol{r}) - \epsilon_2(\boldsymbol{r})\boldsymbol{E}_2(\boldsymbol{r})] = 0. \tag{2.30}$$

4. The normal component of the magnetic field is continuous

$$\hat{\boldsymbol{e}}(\boldsymbol{r}) \cdot [\boldsymbol{H}_1(\boldsymbol{r}) - \boldsymbol{H}_2(\boldsymbol{r})] = 0. \tag{2.31}$$

2.1.4 *Fourier Transform*

Equations. (2.3) and (2.4) are Fourier transforms between the time and frequency domains. A Fourier transform between the spatial and spatial-frequency domain is also frequently used in field solvers. This Fourier transform acts in plane z.

The Fourier transform from the spatial domain to the spatial-frequency domain is defined as

$$\tilde{f}(\boldsymbol{\kappa}) = \mathcal{F}_k f(\boldsymbol{\kappa}) = \frac{1}{2\pi} \iint_{-\infty}^{+\infty} \mathrm{d}x\, \mathrm{d}y\, f(\boldsymbol{\rho}) \exp(-\mathrm{i}\boldsymbol{\kappa} \cdot \boldsymbol{\rho}), \tag{2.32}$$

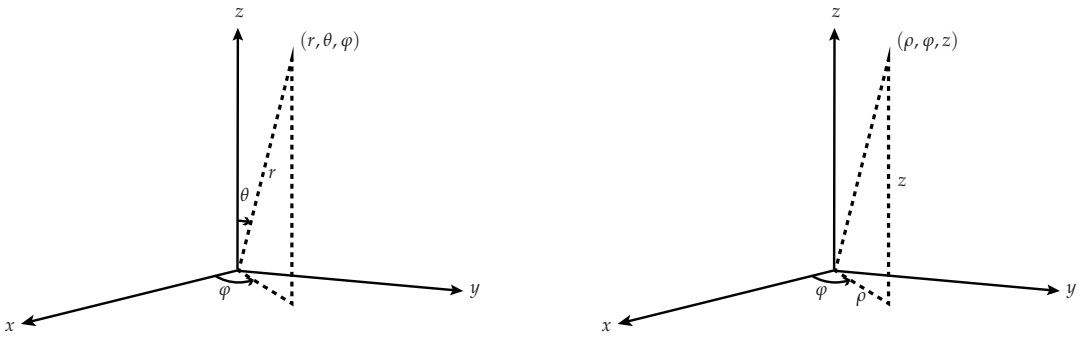
and the inverse transform is defined as

$$f(\boldsymbol{\rho}) = \mathcal{F}_k^{-1} \tilde{f}(\boldsymbol{\kappa}) = \frac{1}{2\pi} \iint_{-\infty}^{+\infty} \mathrm{d}k_x\, \mathrm{d}k_y\, \tilde{f}(\boldsymbol{\kappa}) \exp(\mathrm{i}\boldsymbol{\kappa} \cdot \boldsymbol{\rho}). \tag{2.33}$$

Here $\boldsymbol{\rho} = (x, y)$ and $\boldsymbol{\kappa} = (k_x, k_y)$. f is an arbitrary scalar function in the spatial domain, while $\tilde{f}$ denotes the function in the spatial-frequency domain.

2.2 HIGHLY SYMMETRIC GRADED-INDEX COMPONENT

We are able to find analytical solvers for highly symmetric GRIN components by solving Maxwell's Eqs. (2.20-2.23) or field Eqs. (2.24-2.25) in a proper coordinate system. In the Cartesian coordinate system, the position vector $\boldsymbol{r} = (x, y, z)$, with x, y, z denoting coordinates on three orthogonal axes. In a spherical coordinate system, $\boldsymbol{r} = (r, \theta, \varphi)$, with r denoting radius, θ denoting the azimuth angle, and φ denoting the inclination angle. In a cylindrical coordinate system $\boldsymbol{r} = (\rho, \varphi, z)$, with ρ denoting the radial distance from the z axis, and φ denoting the azimuth angle. The relations between the spherical and Cartesian coordinate systems, as well as that of the cylindrical and Cartesian coordinate system are shown in Fig. 1.



(a) (b)

Figure 1: Coordinate systems. (a) Relation between spherical coordinate system and Cartesian coordinate system; (b) relation between cylinder coordinate system and Cartesian coordinate system.

2.2.1 *Spherically Symmetric Graded-Index Component*

A spherically symmetric GRIN medium has the highest symmetry, having a permittivity that varies only along the radial direction $\epsilon(r)$ in a spherical coordinate system, as shown in Fig. 1 (a). The field solver for Maxwell's equation in a spherical coordinate system is known as Mie theory [46], which is rigorous and analytical. In modeling a GRIN sphere, we discretize $\epsilon(r)$ into a sufficient number of concentric spherical shells radially, and then solve the field propagation in this multilayered sphere.

Task description

Assume that the multilayered sphere, shown in Fig. 2, is centered at the origin of the coordinate system in Fig. 1 (a). In the multilayered sphere in Fig. 2, we see several concentric spherical shells/boundaries between two different permittivities. r_l, $l = 1, 2, \cdots, N$ represents the radius of each spherical shell from the smallest to largest, whereas ϵ_l, $l = 1, 2, \cdots, N$ denotes the relative permittivity between shells.

We calculate an x-polarized ideal plane wave propagating through the multilayered sphere. The mathematical representation of an ideal plane wave is

$$\boldsymbol{E}^{\text{in}}(\boldsymbol{r} = x, y, z) = E_0 \exp(\mathrm{i}k_0\sqrt{\epsilon_{N+1}}z)\hat{\boldsymbol{e}}_x, \tag{2.34}$$

with E_0 as the amplitude, ϵ_{N+1} as the relative permittivity of the medium outside of the sphere, and $\hat{\boldsymbol{e}}_x$ as the unit vector along the x-axis. The related magnetic field $\boldsymbol{H}^{\text{in}}(\boldsymbol{r})$ can be calculated via Eq. (2.20).

The electromagnetic field in the whole three-dimensional (3D) space, i.e., $\boldsymbol{E}(\boldsymbol{r})$ and $\boldsymbol{H}(\boldsymbol{r})$, should be calculated rigorously.

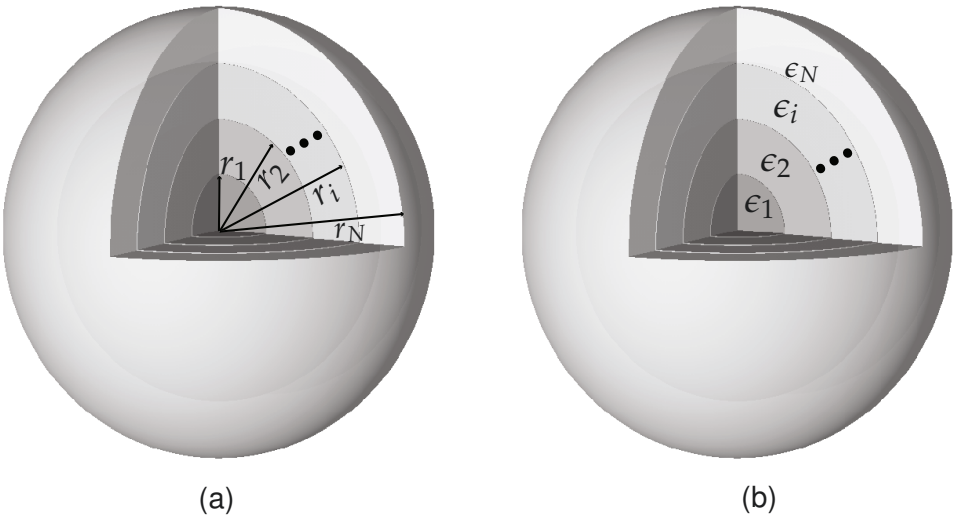


Figure 2: Specifications of multilayered sphere: (a) the radius of each spherical boundary, and (b) the permittivity of each layer.

Please note that this technique can be extended to a general input field by decomposing the input field into plane waves and using a coordinate system transform. Here we show only the derivation process when the input field is an ideal plane wave.

Mie theory derivation

The specific solver, which is derived in a spherical coordinate system, is the well-known Mie theory. The complete derivation is shown in Appendix B. Here we talk about the derivation logic. Two main steps are performed: (1) Representing the electromagnetic field in homogeneous media with relative permittivity ϵ_l, $l = 1, 2, \cdots, N, N+1$ in a spherical coordinate system. These representations contain unknown coefficients to scale the electromagnetic fields. (2) Matching the boundary conditions between neighboring layers to calculate the unknown coefficients.

Representing fields: vectorial spherical harmonics. The basis functions in spherical coordinate systems are the vectorial spherical harmonics $\boldsymbol{M}_l(\boldsymbol{r})$ and $\boldsymbol{N}_l(\boldsymbol{r})$, defined as

$$\begin{aligned} \boldsymbol{M}_l(\boldsymbol{r}) &= \boldsymbol{\nabla} \times [\boldsymbol{r}\psi_l(\boldsymbol{r})], \\ \boldsymbol{N}_l(\boldsymbol{r}) &= \frac{1}{k_0\sqrt{\epsilon_l}} \times \boldsymbol{M}_l(\boldsymbol{r}), \end{aligned} \tag{2.35}$$

where $l = 1, 2, \cdots, N+1$ denotes the index of different layers, and $\psi_l(\boldsymbol{r})$ fulfils the field equations (2.26) and (2.27) in a homogeneous medium with relative permittivity ϵ_l

$$\Delta\psi_l(\boldsymbol{r}) + k_0^2\epsilon_l\psi_l(\boldsymbol{r}) = 0 \tag{2.36}$$

Eq. (2.36) is solved by separating the variables

$$\psi_l(\boldsymbol{r}) = R_l(r)\Theta(\theta)\Phi(\varphi)$$

with $\Theta(\theta)$ as the function of θ, $\Phi(\varphi)$ as the function of φ, and $R_l(r)$ as the function of r. Only $R_l(r)$ is related to the relative permittivity ϵ_l. Solving $\Theta(\theta)$, $\Phi(\varphi)$, and $R_l(r)$, as shown in Appendix B, discrete functions $\psi_{l,omn}(\boldsymbol{r})$ and $\psi_{l,emn}$ are consequently obtained as

$$\begin{aligned}\psi_{l,emn}(\boldsymbol{r}) &= \cos(m\varphi)P_n^m(\cos\theta)z_n(k_0\sqrt{\epsilon_l}r),\\ \psi_{l,omn}(\boldsymbol{r}) &= \sin(m\varphi)P_n^m(\cos\theta)z_n(k_0\sqrt{\epsilon_l}r).\end{aligned} \tag{2.37}$$

where m and n are integer indices, $m \geqslant 0$, $n \geqslant m$, whereas e denotes the even terms $\cos(m\varphi)$, and o denotes the odd terms $\sin(m\varphi)$. P_n^m is the associated Legendre function of degree n and order m, and z_n is the spherical Bessel function [47]. Then, the discrete vectorial spherical harmonics $\boldsymbol{M}_{l,omn}(\boldsymbol{r})$, $\boldsymbol{M}_{l,emn}(\boldsymbol{r})$, $\boldsymbol{N}_{l,omn}(\boldsymbol{r})$, and $\boldsymbol{N}_{l,emn}(\boldsymbol{r})$ are calculated using Eq. (2.35), as shown in Tab. 7 Appendix B.

The vectorial spherical harmonics are **complete** and mutually **orthogonal** for all (m, m', n, n'). We represent the electromagnetic field as a combination of $\boldsymbol{M}_{l,omn}(\boldsymbol{r})$, $\boldsymbol{M}_{l,emn}(\boldsymbol{r})$, $\boldsymbol{N}_{l,omn}(\boldsymbol{r})$, and $\boldsymbol{N}_{l,emn}(\boldsymbol{r})$.

$$\boldsymbol{E}_l(\boldsymbol{r}) = \sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\left[a_{l,mn}\boldsymbol{M}_{l,emn}(\boldsymbol{r}) + b_{l,mn}\boldsymbol{M}_{l,omn}(\boldsymbol{r}) + c_{l,mn}\boldsymbol{N}_{l,emn}(\boldsymbol{r}) + d_{l,mn}\boldsymbol{N}_{l,omn}(\boldsymbol{r})\right], \tag{2.38}$$

and

$$\boldsymbol{H}_l(\boldsymbol{r}) = -\mathrm{i}\sqrt{\frac{\epsilon_0\epsilon_l}{\mu_0}}\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\left[a_{l,mn}\boldsymbol{N}_{l,emn}(\boldsymbol{r}) + b_{l,mn}\boldsymbol{N}_{l,omn}(\boldsymbol{r}) + c_{l,mn}\boldsymbol{M}_{l,emn}(\boldsymbol{r}) + d_{l,mn}\boldsymbol{M}_{l,omn}(\boldsymbol{r})\right], \tag{2.39}$$

with free coefficients $a_{l,mn}$, $b_{l,mn}$, $c_{l,mn}$, and $d_{l,mn}$.

Matching the boundary conditions. In our task, the input field is an $x-$polarized ideal plane wave, which can be represented as

$$\boldsymbol{E}^{\mathrm{in}}(r,\theta,\varphi) = E_0\sum_{n=1}^{\infty}\mathrm{i}^n\frac{2n+1}{n(n+1)}(\boldsymbol{M}_{N+1,o1n} - \mathrm{i}\boldsymbol{N}_{N+1,e1n}). \tag{2.40}$$

From Eq. (2.34), only vectorial spherical harmonics with $m = 1$ can be exited. This considerably reduce the numerical effort required. Then, substituting Eqs. (2.38) and (2.39) into the boundary conditions, Eqs. (2.28-2.31), all coefficients $a_{l,1n}$, $b_{l,1n}$, and $c_{l,1n}$ can be calculated. The final implemented formulas are Eqs. (B.44-B.51) [48, 49].

Properties of Mie theory

- In the current derivation, the input field is an $x-$polarized ideal plane wave having a wave vector along the z-axis.
- It only works for a spherically symmetric GRIN medium.
- It is a rigorous solver for modeling a multilayered sphere. When the layer number is large enough to produce a converged result, we say the solver is rigorous for modeling a GRIN sphere.

- The optical surface effect, which is introduced by an abrupt change in relative permittivity $\epsilon(\boldsymbol{r})$ between the surrounding medium and the GRIN sphere boundary, is rigorously included.
- All physical effects (linear optics domain) generated inside the GRIN medium, i.e., non-paraxial fields, polarization crosstalk, diffraction, reflection, evanescent wave, are rigorously considered.

2.2.2 *Cylindrically Symmetric Graded-Index Fiber*

Cylindrically symmetric GRIN components frequently appear in applications, e.g., Wood GRIN lens, relay lens in an endoscope, GRIN rod lens, and so on. However, the most widely used cylindrical GRIN component is the GRIN fiber.

There are different specific field solvers that calculate the propagation modes for GRIN fibers. If the fiber is long enough, only parts of the input field, which can be projected to the modes, remain in the fiber, while the remaining part disappears during the propagation. Therefore, the process of the field solver for a long GRIN fiber is

- Calculating the propagating fiber modes

$$\boldsymbol{V}_{lm}(\boldsymbol{\rho}, z=0) = \{E_{x,lm}, E_{y,lm}, E_{z,lm}, H_{x,lm}, H_{y,lm}, H_{z,lm}\}(\boldsymbol{\rho}, z=0),$$

  with mode index (l,m), $l = 0, 1, \cdots, L$, $m = 0, 1, \cdots, M$, and $\boldsymbol{\rho} = (x, y)$. L and M can be fixed during the calculation. The modes are only related to fiber structure $\epsilon(\boldsymbol{r})$ and the working wavelength.

- Representing the input field $\boldsymbol{E}^{\text{in}}(\boldsymbol{\rho}, z=0)$ at fiber end as the summation of modes $\boldsymbol{E}_{lm}(\boldsymbol{\rho}, z=0)$ with weight coefficients a_{lm}, as

$$\boldsymbol{E}^{\text{in}}(\boldsymbol{\rho}, z=0) = \sum_{l=0}^{L}\sum_{m=0}^{M} a_{lm}\boldsymbol{E}_{lm}(\boldsymbol{\rho}, z=0) + \boldsymbol{E}^{\text{rest}}(\boldsymbol{\rho}, z=0). \tag{2.41}$$

  Taking advantage of the orthogonality of fiber modes, coefficients a_{lm} can be calculated using the overlap integral.

$$a_{lm}^2 = \frac{\iint \boldsymbol{E}^{\text{in}}(\boldsymbol{\rho}, z=0) \times \boldsymbol{H}_{lm}^{\star}(\boldsymbol{\rho}, z=0)\,\mathrm{d}x\,\mathrm{d}y}{\iint \boldsymbol{E}_{lm}(\boldsymbol{\rho}, z=0) \times \boldsymbol{H}_{lm}^{\star}(\boldsymbol{\rho}, z=0)\,\mathrm{d}x\,\mathrm{d}y}, \tag{2.42}$$

  and $a_{lm} = \sqrt{a_{lm}^2}$, a complex value.

- The output field is the coherent summation of all propagated modes

$$\boldsymbol{E}^{\text{out}}(\boldsymbol{\rho}, z) = \sum_{l=0}^{L}\sum_{m=0}^{M} a_{lm}\boldsymbol{E}_{lm}(\boldsymbol{\rho}, z=0)\exp(\mathrm{i}\beta_{lm}z), \tag{2.43}$$

  with β_{lm} calculated during the calculation of modes.

To calculate these propagating fiber modes $\boldsymbol{V}_{lm}(\boldsymbol{\rho}, z = 0)$, field Eqs. (2.24) and (2.25) are solved in the cylindrical coordinate system. In practice, most fibers are weakly guiding fibers, which means $\epsilon(\rho)$ is slightly variant, so $\boldsymbol{\nabla} \ln \epsilon(\boldsymbol{r})$ in Eqs. (2.24) and (2.25) is approximately zero.

The analytical mode solver for the GRIN fiber can be developed for several types of $\epsilon(\rho)$, e.g., having an infinite parabolic profile [50]

$$\epsilon(\rho) = \epsilon_c \left[1 - 2\delta \left(\frac{\rho}{\rho_0}\right)^2\right]. \tag{2.44}$$

with δ as a constant. The modes have a distribution similar to the Gaussian Laguerre at its waist. In Appendix C, we give a detailed derivation of it, to demonstrate a fiber mode calculation, as well as the specific field solver based on modes. In [50], several mode solvers for specific values of $\epsilon(\rho)$ are presented.

Properties of fiber solvers

- The input field is general with no limitations, such as paraxial or geometric.
- The optical surface effects induced by an abrupt change in relative permittivity $\epsilon(\boldsymbol{r})$ between the surroundings and the GRIN fiber are not included.
- The limitation of analytical fiber solvers mainly comes from their mode calculation.
  - It works for a cylindrically symmetric structure with a specific $\epsilon(\rho)$, e.g., having an infinite parabolic profile.
  - Modes calculation induces approximations, e.g., weak guidance with small $\epsilon(\rho)$ variance, and no polarization crosstalk between E_x, E_y, or E_z is considered ($\boldsymbol{\nabla} \ln \epsilon(\boldsymbol{r}) \approx 0$).
  - Reflections from the GRIN medium are not included.
  - Evanescent waves are not considered.

2.3 BEAM PROPAGATION METHODS FOR GENERAL GRADED-INDEX COMPONENTS

There are a few numerical field solvers for modeling general GRIN components that split the medium into slices along z-axis and solve Eqs. (2.24) and (2.25) iteratively along the z-axis. They are called beam propagation methods.

2.3.1 *Fourier Transform Beam Propagation Method*

In 1978, Feit and Fleck [39] described a method that solves Eq. (2.24), with the assumption of small variance in $\epsilon(\boldsymbol{r})(\boldsymbol{\nabla} \ln \epsilon(\boldsymbol{r}) \approx 0)$. Here, we rewrite the equation as

$$\Delta \boldsymbol{E}(\boldsymbol{r}) + k_0^2 \epsilon(\boldsymbol{r}) \boldsymbol{E}(\boldsymbol{r}) = 0. \tag{2.45}$$

In the Cartesian coordinate system, Eq. (2.45) can be decoupled into three identical equations to describe $E_x(\boldsymbol{r})$, $E_y(\boldsymbol{r})$, and $E_z(\boldsymbol{r})$, and **no crosstalk between** $E_x(\boldsymbol{r})$, $E_y(\boldsymbol{r})$, **or** $E_z(\boldsymbol{r})$ **can be predicted**.

Mathematical description. Separating $\Delta = \Delta_\perp + \dfrac{\partial^2}{\partial z^2}$, with $\Delta_\perp = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$, one can rewrite Eq. (2.45) as

$$\frac{\partial^2}{\partial z^2}\boldsymbol{E}(\boldsymbol{r}) = -\left[\Delta_\perp + k_0^2\epsilon(\boldsymbol{r})\right]\boldsymbol{E}(\boldsymbol{r}), \tag{2.46}$$

and if we **consider only the forward propagation,**

$$\boldsymbol{E}(\boldsymbol{\rho}, z+\Delta z) = \exp \mathrm{i}\Delta z\sqrt{\Delta_\perp + k_0^2\epsilon(\boldsymbol{\rho}, z)}\boldsymbol{E}(\boldsymbol{\rho}, z). \tag{2.47}$$

As $\Delta_\perp$ is a linear differential operator, the rigorous linear mathematical transform $\sqrt{a+b^2} = \dfrac{a}{\sqrt{a+b^2}+b} + b$ can be used, considering $\Delta_\perp$ as the free parameter a

$$\sqrt{\Delta_\perp + k_0^2\epsilon(\boldsymbol{\rho}, z)} = \frac{\Delta_\perp}{\sqrt{\Delta_\perp + k_0^2\epsilon(\boldsymbol{\rho}, z)} + k_0\sqrt{\epsilon(\boldsymbol{\rho}, z)}} + k_0\sqrt{\epsilon(\boldsymbol{\rho}, z)}. \tag{2.48}$$

Now, we use an approximation in the denominator of Eq. (2.48) to replace $\epsilon(\boldsymbol{\rho}, z)$ with a constant ϵ_c because of the small variance in $\epsilon(\boldsymbol{\rho}, z)$. We rewrite the operator as

$$\begin{aligned}\sqrt{\Delta_\perp + k_0^2\epsilon(\boldsymbol{\rho}, z)} &\approx \frac{\Delta_\perp}{\sqrt{\Delta_\perp + k_0^2\epsilon_c} + k_0\sqrt{\epsilon_c}} + k_0\sqrt{\epsilon_c} - k_0\sqrt{\epsilon_c} + k_0\sqrt{\epsilon(\boldsymbol{\rho}, z)} \\ &= \sqrt{\Delta_\perp + k_0^2\epsilon_c} + k_0\left[\sqrt{\epsilon(\boldsymbol{\rho}, z)} - \sqrt{\epsilon_c}\right].\end{aligned} \tag{2.49}$$

Substituting Eq. (2.49) into Eq. (2.47), and splitting the operator into a symmetric form, we have

$$\boldsymbol{E}(\boldsymbol{\rho}, z+\Delta z) = \exp\left(\mathrm{i}\frac{\Delta z}{2}\sqrt{\Delta_\perp + k_0^2\epsilon_c}\right)\exp\left\{\mathrm{i}\Delta z k_0\left[\epsilon(\boldsymbol{\rho}, z) - \epsilon_c\right]\right\}\exp\left(\mathrm{i}\frac{\Delta z}{2}\sqrt{\Delta_\perp + k_0^2\epsilon_c}\right)\boldsymbol{E}(\boldsymbol{\rho}, z). \tag{2.50}$$

Eq. (2.50) is the final mathematical formula for this field solver, which describes the field behavior from z to $z+\Delta z$.

Physical interpretation. Again, we consider only the calculation from z to $z+\Delta z$, which has the mathematical formula Eq. (2.50). The three operators work on field $\boldsymbol{E}(\boldsymbol{\rho}, z)$ one by one. Comparing operator $\exp\left(\mathrm{i}\dfrac{\Delta z}{2}\sqrt{\Delta_\perp + k_0^2\epsilon_c}\right)$ with Eq. (2.47), we know that this operator means a free space propagation of distance $\Delta z/2$ in homogeneous media with relative permittivity ϵ_c. Thus, the physical interpretation of the operations are,

- Propagation of $\boldsymbol{E}(\boldsymbol{\rho}, z)$ in the homogeneous media with ϵ_c for $\Delta z/2$
  - Fourier transform of $\boldsymbol{E}(\boldsymbol{\rho}, z)$ to the k domain by Eq. (2.32) gives $\tilde{\boldsymbol{E}}(\boldsymbol{\kappa}, z)$.
  - Multiplying by the free space propagation operator

$$\tilde{\boldsymbol{E}}(\boldsymbol{\kappa}, z+\frac{\Delta z}{2}) = \tilde{\boldsymbol{E}}(\boldsymbol{\kappa}, z)\exp\left[\mathrm{i}k_z(\boldsymbol{\kappa})\frac{\Delta z}{2}\right],$$

    with $k_z(\boldsymbol{\kappa}) = \sqrt{k_0^2\epsilon_c - ||\boldsymbol{\kappa}||^2}$.

  - Inverse Fourier transform of $\tilde{\boldsymbol{E}}(\boldsymbol{\kappa}, z+\frac{\Delta z}{2})$ using Eq. (2.33) gives $\boldsymbol{E}(\boldsymbol{\rho}, z+\frac{\Delta z}{2})$

- Multiplying the phase factor, which has a form similar to the thin element approximation (TEA) [31, 51] for a paraxial beam

$$\boldsymbol{E}^{\mathrm{TEA}}(\boldsymbol{\rho}, z+\frac{\Delta z}{2}) = \boldsymbol{E}(\boldsymbol{\rho}, z+\frac{\Delta z}{2}) \exp\left\{\mathrm{i}\Delta z k_0 \left[\sqrt{\epsilon(\boldsymbol{\rho}, z)} - \sqrt{\epsilon_c}\right]\right\} \tag{2.51}$$

- Propagation of $\boldsymbol{E}^{\mathrm{TEA}}(\boldsymbol{\rho}, z+\frac{\Delta z}{2})$ in the homogeneous media with ϵ_c for $\Delta z/2$ gives $\boldsymbol{E}(\boldsymbol{\rho}, z+\Delta z)$

During the free space propagation, a Fourier transform is used, so in some studies this method is called the Fourier transform beam propagation method (FT-BPM).

Properties of FT-BPM

- The input field can be a general field, either geometric or diffractive.
- The GRIN media has a general $\epsilon(\boldsymbol{r})$, with no restriction in symmetry or periodicity, but the variance of $\epsilon(\boldsymbol{r})$ should be sufficiently small.
- The optical surface effect, where there is an abrupt change in $\epsilon(\boldsymbol{r})$ between the surrounding medium and the GRIN medium, is not included.
- Effects generated by the GRIN medium:
  - Reflection is not modeled, because only forward propagation is considered in Eq. (2.47).
  - Polarization crosstalk between E_x, E_y, and E_z is not considered, as we neglect the right-hand side of Eq. (2.24).
  - Diffraction is considered, so the focus inside GRIN media can be calculated.
  - It is not easy to evaluate the capability of this method to calculate non-paraxial. In the mathematical derivation, no explicit approximation shows that this method has constraints for non-paraxial calculation. However, in the physical interpretation, TEA, which offers a paraxial approximation, is used to compensate the phase variance in the GRIN medium. That gives us an indication that this method is better used for modeling paraxial fields. In [38], the authors showed an experiment that worked for angle range of -10° to 10° when $\Delta\sqrt{\epsilon(\boldsymbol{r})} \approx 0.5$. Here we give another example for propagating a plane wave through a Luneburg lens, shown in Fig. 3 (a). The light should focus on a rear point with a large convergent angle, but the FT-BPM does not predict the focus correctly. Thus, we suggest using the method to model paraxial beam propagation.

– Evanescent waves are not predicted with this method. In Fig. 3 (b), a plane wave is shown propagating through a volume grating. Compared with the result calculated using the Fourier modal method (FMM, a rigorous method), the evanescent wave orders (red circle) are not well predicted.

2.3.2 *Wave Propagation Method*

In 1993, Brenner and Singer [38] proposed the wave propagation method (WPM), which also solves Eq. (2.45), but overcomes the limitation of the small variance in $\epsilon(\boldsymbol{r})$, and can be used to model non-paraxial beam propagation.

The mathematical interpretation is based on pseudo-differential operator, shown in the appendix of [52], and is not shown again here. However, it is interesting to derive this method's physical interpretation. From the discussion of FT-BPM, the guess of paraxial limitation is based on the TEA. However, Turunen extended it to treat inclined incident propagation through a microstructure[31], and in Appendix D, we also extend the TEA to treat inclined beam propagation through a GRIN layer,

$$\boldsymbol{E}(\boldsymbol{\rho}, z+\Delta z) = \boldsymbol{E}(\boldsymbol{\rho}, z)\exp(\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{\rho})\exp(\mathrm{i}k_z(\boldsymbol{\rho}, z; \boldsymbol{\kappa})\Delta z), \tag{2.52}$$

with $\boldsymbol{E}(\boldsymbol{\rho}, z)$ as a field having a narrow angular spectrum centered on $\boldsymbol{\kappa}$, and $k_z(\boldsymbol{\rho}, z; \boldsymbol{\kappa}) = \sqrt{k_0^2\epsilon(\boldsymbol{\rho}, z) + ||\boldsymbol{\kappa}||^2}$.

Let us replace TEA by Eq. (2.52) and discuss the physical interpretation from z to $z+\Delta z$

- A Fourier transform of $\boldsymbol{E}(\boldsymbol{\rho}, z)$ to the k domain by Eq. (2.32) gives $\tilde{\boldsymbol{E}}(\boldsymbol{\kappa}, z)$. This step can be understood as the plane wave decomposition.
- For each $\boldsymbol{\kappa}$, the related complex amplitude in Eq. (2.52) is a constant along $\boldsymbol{\rho}$

  $$\boldsymbol{E}(\boldsymbol{\rho}, z; \boldsymbol{\kappa}) = \tilde{\boldsymbol{E}}(\boldsymbol{\kappa}, z), \tag{2.53}$$

  and the field in plane $z+\Delta z$ is

  $$\boldsymbol{E}(\boldsymbol{\rho}, z+\Delta z; \boldsymbol{\kappa}) = \boldsymbol{E}(\boldsymbol{\rho}, z; \boldsymbol{\kappa})\exp(\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{\rho})\exp(\mathrm{i}k_z(\boldsymbol{\rho}, z, \boldsymbol{\kappa})\Delta z), \tag{2.54}$$

  with $k_z(\boldsymbol{\rho}, z; \boldsymbol{\kappa}) = \sqrt{k_0^2\epsilon(\boldsymbol{\rho}, z) + ||\boldsymbol{\kappa}||^2}$. This is the first scanning, which scans all positions $\boldsymbol{\rho}$.
- Reconstruct the field in plane $z+\Delta z$

  $$\boldsymbol{E}(\boldsymbol{\rho}, z+\Delta z) = \frac{1}{2\pi}\iint \boldsymbol{E}(\boldsymbol{\rho}, z+\Delta z; \boldsymbol{\kappa})\,\mathrm{d}k_x\,\mathrm{d}k_y \tag{2.55}$$

  This is the second scanning, which scans all $\boldsymbol{\kappa}$.

The optical surface effects were covered by Fertig and Brenner in 2010 [40]. Before performing the extended TEA, the authors considered the Fresnel coefficients [53] of the local change from $\epsilon(\boldsymbol{\rho}, z^-)$ to $\epsilon(\boldsymbol{\rho}, z^+)$. This process includes not only the optical surface, but also the polarization crosstalk produced by the abrupt change from $\epsilon(\boldsymbol{\rho}, z^-)$ to $\epsilon(\boldsymbol{\rho}, z^+)$.

Properties of WPM

Please note that here we mainly discuss the WPM introduced in [38, 40]. The WPM has potential for including more effects [54].

- the input field can be a general field.
- The GRIN media has a general $\epsilon(\boldsymbol{r})$, with no restriction in symmetry or periodicity, but the variance in $\epsilon(\boldsymbol{r})$ should be sufficiently small.
- The optical surface effect, where an abrupt change in the $\epsilon(\boldsymbol{r})$ between surrounding medium and the GRIN medium occurs, is accounted for.
- Effects generated inside a GRIN medium:
  - Reflection is not modeled, because only forward propagation is considered.
  - Polarization crosstalk between E_x, E_y, and E_z is modeled by including the Fresnel matrices of local change from $\epsilon(\boldsymbol{\rho}, z^-)$ to $\epsilon(\boldsymbol{\rho}, z^+)$
  - Diffraction is considered, thus the focus inside GRIN media can be calculated.
  - It works for non-paraxial beams. Here, we also propagate a plane wave through a Luneburg lens, as in Fig. 3 (a). This method predicts the focus with a large convergent angle accurately.
  - The evanescent wave is not included. In Fig. 3 (b), a plane wave is propagating through a volume grating. Compared with the result calculated by the Fourier modal method (rigorous method), the evanescent wave is not well-predicted.

2.3.3 *Finite Difference Beam Propagation Methods*

The other kinds of beam propagation methods are also straightforward. They replace the partial differential operator ∂x by Δx, and ∂y by Δy, and they solve the ordinary differential equation (ODE) along the z-axis. They are called finite difference beam propagation methods (FD-BPM). We first introduce the most fundamental BPM, which is paraxial and does not include the crosstalk between different field components in the Cartesian coordinate system, i.e., $E_x(\boldsymbol{r})$, $E_y(\boldsymbol{r})$, and $E_z(\boldsymbol{r})$. Then, depending on the physical effect to be considered, we add patches to the mathematical derivation to develop, e.g., non-paraxial BPM or vectorial BPM.

2.3.3.1 *Fundamental FD-BPM*

The most fundamental FD-BPM solves Eq. (2.46) with the assumption of small variance in $\epsilon(r)(\boldsymbol{\nabla} \ln \epsilon(\boldsymbol{r}) \approx 0)$. We emphasize here that because of the approximation, **crosstalk between the field components in the Cartesian coordinate system, i.e., $E_x(\boldsymbol{r})$, $E_y(\boldsymbol{r})$, and $E_z(\boldsymbol{r})$, cannot be predicted**.

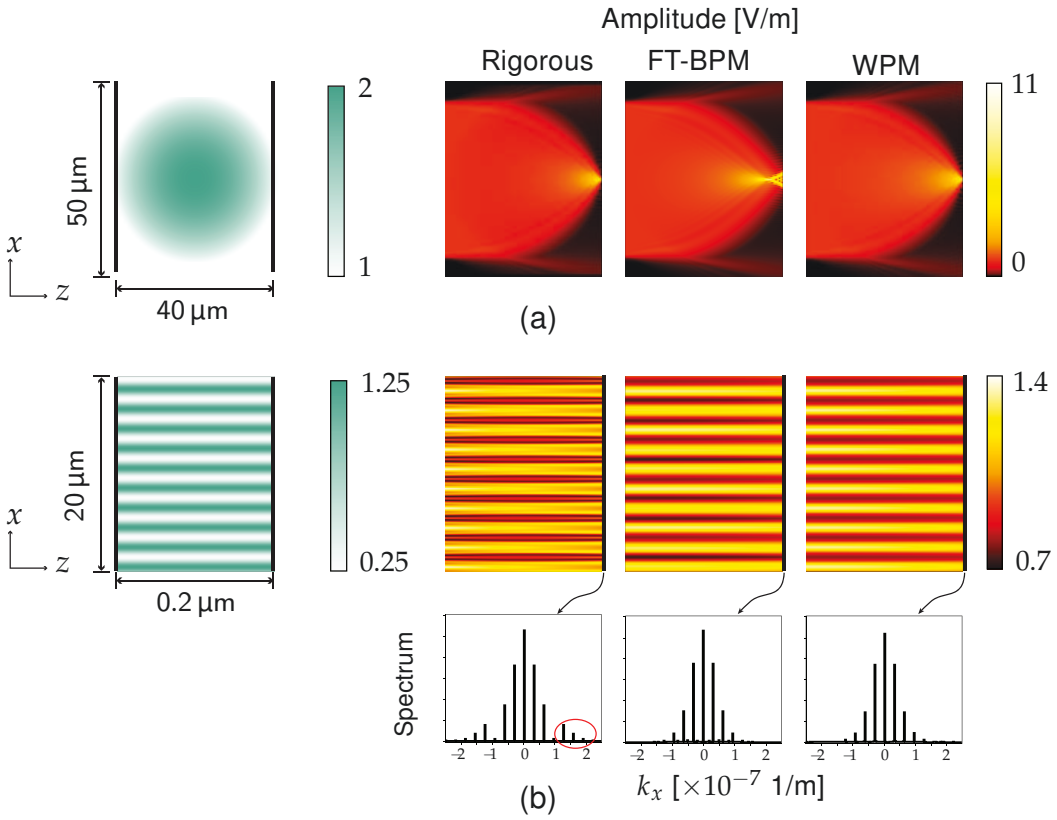


Figure 3: Two test examples are modeled using a rigorous technique, FT-BPM, and WPM. (a) A plane wave is focused by a Luneburg lens, and the amplitude (E_y) of the field in the xz-plane is calculated. (b) A plane wave propagates through a volume grating. The amplitude (E_y) in the xz-plane, and the angular spectrum of the field in the plane z =0.2 µm are calculated. The color at the left in the figures represents the relative permittivity $\epsilon(\boldsymbol{r})$ of the optical systems. The orders in the red circle (bottom) are evanescent wave ones.

Now let us replace $\Delta_\perp f(\boldsymbol{r})$ with a finite difference representation, as

$$\begin{cases} \dfrac{\partial^2 f(\boldsymbol{r})}{\partial x^2} = \dfrac{f(x+\Delta x,y,z)-2f(\boldsymbol{r})+f(x-\Delta x,y,z)}{\Delta x^2} \\ \dfrac{\partial^2 f(\boldsymbol{r})}{\partial y^2} = \dfrac{f(x,y+\Delta y,z)-2f(\boldsymbol{r})+f(x,y-\Delta y,z)}{\Delta y^2}. \end{cases} \tag{2.56}$$

From now on, we continue to use $\Delta_\perp$ in the equations, but please note that it is not a differential operator but an operator related to Δx and Δy. Thus, Eq. (2.46) can be represented as an ODE,

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\boldsymbol{E}(\boldsymbol{r}) = \left[\Delta_\perp + k_0^2\epsilon(\boldsymbol{r})\right]\boldsymbol{E}(\boldsymbol{r}) \tag{2.57}$$

We know that the phase change along the z axis is quite fast, so the field is rewritten as

$$\boldsymbol{E}(\boldsymbol{\rho}, z) = \boldsymbol{e}(\boldsymbol{\rho}, z) \exp(\mathrm{i}k_0\sqrt{\epsilon_c}z), \tag{2.58}$$

with $k_0\sqrt{\epsilon_c}z$ denoting the quickly changing phase, ϵ_c a constant, and $\boldsymbol{e}(\boldsymbol{\rho}, z) := \boldsymbol{E}(\boldsymbol{\rho}, z)\exp(-\mathrm{i}k_0\sqrt{\epsilon_c}z)$. Then, substituting the field ansatz (2.58) into Eq. (2.57), we get

$$-\frac{\mathrm{i}}{2k_0\sqrt{\epsilon_c}}\frac{\mathrm{d}^2}{\mathrm{d}z^2}\boldsymbol{e}(\boldsymbol{\rho}, z) + \frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{e}(\boldsymbol{\rho}, z) = \frac{\mathrm{i}}{2k_0\sqrt{\epsilon_c}}\left[\Delta_\perp + k_0^2\epsilon(\boldsymbol{r}) - k_0^2\epsilon_c\right]\boldsymbol{e}(\boldsymbol{\rho}, z). \tag{2.59}$$

Here we emphasis that there is no approximation introduced by the field ansatz or the rewritten ODE (2.59). Now we introduce another approximation, which assumes that $\boldsymbol{e}(\boldsymbol{\rho}, z)$ varies slowly along the z-axis, i.e., a slowly varying envelop approximation, so $\frac{\mathrm{d}^2}{\mathrm{d}z^2}\boldsymbol{e}(\boldsymbol{\rho}, z) \approx 0$. Eq.(2.59) becomes a first-order ODE,

$$\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{e}(\boldsymbol{\rho}, z) = \frac{\mathrm{i}}{2k_0\sqrt{\epsilon_c}}\left[\Delta_\perp + k_0^2\epsilon(\boldsymbol{r}) - k_0^2\epsilon_c\right]\boldsymbol{e}(\boldsymbol{\rho}, z). \tag{2.60}$$

This approximation introduces the **paraxial limitation** of the method.

When the initial field $\boldsymbol{e}(\boldsymbol{\rho}, z = 0)$ is known, Eq. (2.60) can be solved iteratively using the Euler method or the Runge–Kutta methods[55].

To formulate the numerical algorithm, topics related to the boundary conditions are frequently discussed, i.e., how to define $\boldsymbol{e}(x = x_{\min}, y, z)$, $\boldsymbol{e}(x = x_{\max}, y, z)$, $\boldsymbol{e}(x, y = y_{\min}, z)$, and $\boldsymbol{e}(x, y = y_{\max}, z)$ in the calculation of Eq. (2.56). The boundary conditions can be, e.g., (1) electric wall boundary, i.e., all the mentioned values are zero; (2) transparent boundary conditions [56]; (3) perfectly matched layers [57].

2.3.3.2 *Non-paraxial BPM*

To include the non-paraxial situation, the slowly varying envelop approximation should be avoided. The most popular non-paraxial BPM was proposed by Hadley in 1992 [41].

Defining an operator P

$$P = \Delta_\perp + k_0^2\epsilon(\boldsymbol{r}) - k_0^2\epsilon_c, \tag{2.61}$$

substituting P into Eq. (2.59), and solving $\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{e}(\boldsymbol{\rho}, z)$ using a quadratic equation

$$AX^2 + BX + C = 0, \tag{2.62}$$

with

$$A = -\frac{\mathrm{i}}{2k_0\sqrt{\epsilon_c}},\ B = 1,\ C = -\frac{\mathrm{i}}{2k_0\sqrt{\epsilon_c}}P, \tag{2.63}$$

we get the forward propagation part as

$$\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{e}(\boldsymbol{\rho}, z) = \mathrm{i}\left[\sqrt{P + k_0^2\epsilon_c} - k_0\sqrt{\epsilon_c}\right]\boldsymbol{e}(\boldsymbol{\rho}, z). \tag{2.64}$$

The complete proof is not as simple as Eqs. (2.62) and (2.63), and Hadley did a more complete derivation. **The reflection is not included**.

Then, a Padé approximation, which is better converged than the Taylor expansion, is applied to the operation

$$f(P) = \left[\sqrt{P + k_0^2 \epsilon_c} - k_0 \sqrt{\epsilon_c}\right], \tag{2.65}$$

and $f(P)$ is approximated by $R(P)$

$$f(P) \approx R(P) = \frac{\sum_{m=0}^{M} a_m P^m}{1 + \sum_{n=1}^{N} b_n P^n}, \tag{2.66}$$

with all free parameters a_m and b_n calculated by the definition of the Padé approximation [58], i.e.,

$$\frac{\mathrm{d}^{(j)}}{\mathrm{d}P^{(j)}} f(P) = \frac{\mathrm{d}^{(j)}}{\mathrm{d}P^{(j)}} R(P), \; j = 0, \cdots, M+N \tag{2.67}$$

As the term $\boldsymbol{\nabla} \ln \epsilon(\boldsymbol{r}) \approx 0$, the non-paraxial BPM does **not include crosstalk between the field components in the Cartesian coordinate system, i.e.,** $E_x(\boldsymbol{r})$, $E_y(\boldsymbol{r})$, **and** $E_z(\boldsymbol{r})$. However, compared with the paraxial BPM, it does not use a slowly varying envelop approximation, so the non-paraxial case can be treated. If only the lowest order of Padé approximation is considered, the method is reduced to the fundamental BPM. **Evanescent waves are still not treated**, as proved in [42].

2.3.3.3 *Vectorial BPM [43, 59]*

Now, we need to return to Eq. (2.24), and separate ∂z and $\partial x, y$ into the left and right hands of the equation in the Cartesian coordinate system

$$\frac{\partial^2}{\partial z^2}\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}(\boldsymbol{r}) + \begin{pmatrix} 0 \\ 0 \\ \frac{\partial}{\partial z} \end{pmatrix}[\nabla \ln \epsilon(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r})] = -[\Delta_\perp + k_0^2 \epsilon(\boldsymbol{r})]\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}(\boldsymbol{r}) - \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ 0 \end{pmatrix}[\nabla \ln \epsilon(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r})], \tag{2.68}$$

with $\Delta_\perp = (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})$ and

$$\nabla \ln \epsilon(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r}) = \frac{\partial}{\partial x} \ln \epsilon(\boldsymbol{r}) E_x(\boldsymbol{r}) + \frac{\partial}{\partial y} \ln \epsilon(\boldsymbol{r}) E_y(\boldsymbol{r}) + \frac{\partial}{\partial z} \ln \epsilon(\boldsymbol{r}) E_z(\boldsymbol{r})$$

At this point, it is rigorous. Next, we introduce the first approximation of the small change in ϵ along the z-axis, so $\frac{\partial}{\partial z} \ln \epsilon(\boldsymbol{r}) \approx 0$. After substituting this approximation into Eq. (2.68), we see that the behavior of transverse components E_x and E_y is not related to E_z anymore. We then analyze E_x and E_y, and Eq. (2.68) can be simplified as

$$\frac{\partial^2}{\partial z^2}\begin{pmatrix} E_x \\ E_y \end{pmatrix}(\boldsymbol{r}) = -[\Delta_\perp + k_0^2 \epsilon(\boldsymbol{r})]\begin{pmatrix} E_x \\ E_y \end{pmatrix}(\boldsymbol{r}) - \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}\left[\frac{\partial}{\partial x} \ln \epsilon(\boldsymbol{r}) E_x(\boldsymbol{r}) + \frac{\partial}{\partial y} \ln \epsilon(\boldsymbol{r}) E_y(\boldsymbol{r})\right]. \tag{2.69}$$

Again, we replace ∂x and ∂y by Δx and Δy using Eq. (2.56), and

$$\begin{cases} \dfrac{\partial}{\partial x} g(\boldsymbol{r}) &= \dfrac{g(x+\Delta x, y, z) - g(x, y, z)}{\Delta x} \\ \dfrac{\partial}{\partial y} g(\boldsymbol{r}) &= \dfrac{g(x, y+\Delta y, z) - g(x, y, z)}{\Delta y}, \end{cases} \tag{2.70}$$

Eq. (2.68) becomes an ODE with respect to z. Now, using the same field ansatz in Eq. (2.58), and slowly varying the envelop approximation such that $\dfrac{\partial^2}{\partial z^2}\boldsymbol{e}(\boldsymbol{r}) \approx 0$, we get a first-order ODE,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix} e_x \\ e_y \end{pmatrix}(\boldsymbol{r}) = \; & \frac{\mathrm{i}}{2k_0\sqrt{\epsilon_c}}[\Delta_\perp + k_0^2\epsilon(\boldsymbol{r}) - k_0^2\epsilon_c]\begin{pmatrix} e_x \\ e_y \end{pmatrix}(\boldsymbol{r}) + \\ & \frac{\mathrm{i}}{2k_0\sqrt{\epsilon_c}}\begin{pmatrix} \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} \end{pmatrix}\left[\frac{\partial}{\partial x}\ln\epsilon(\boldsymbol{r})e_x(\boldsymbol{r}) + \frac{\partial}{\partial y}\ln\epsilon(\boldsymbol{r})e_y(\boldsymbol{r})\right]. \end{aligned} \tag{2.71}$$

The ODE can be solved using the numerical approach, e.g., the Euler or the Runge–Kutta method [55]. Because of the slowly varying envelop approximation, this method works for the paraxial situation.

2.3.3.4 *Other BPMs*

In the literature, there is also an extended FD-BPM that treat reflection and evanescent waves [60, 61].

The simple way to include reflection is to introduce an optical surface effect between neighboring z steps and treat the backward propagation using Eq. (2.64), but replace $\sqrt{P + k_0^2\epsilon_c}$ with $-\sqrt{P + k_0^2\epsilon_c}$.

The way to include evanescent waves is to let ϵ_c be a complex value [60].

Properties of FD-BPMs

One advantage of FD-BPMs is that they can be extended to include the desired physical effects. Until now, none of them have modeled the polarization and the non-paraxial case at the same time.

The properties are

- The input field can be a general field, and some FD-BPMs can only deal with a paraxial input beam.
- GRIN media have a general $\epsilon(\boldsymbol{r})$, with no restriction in symmetry or periodicity, but most FD-BPMs require small variance in $\epsilon(\boldsymbol{r})$, and the vectorial FD-BPM requires small variance in $\epsilon(\boldsymbol{r})$ along the z axis.
- Effects generated by the GRIN medium:

- Most BPMs only deal with forward propagation, and reflection is not included. To include reflection, a special bidirectional BPM is developed [61].
- Polarization crosstalk between E_x, E_y, and E_z is generally not included because of $\boldsymbol{\nabla} \ln \epsilon(\boldsymbol{r}) \approx 0$. Vectorial BPM considers the crosstalk between E_x and E_y.
- Diffraction is considered, and the focus inside GRIN media can be calculated.
- The fundamental BPM and vectorial BPM deal only with the paraxial case, and the non-paraxial case is covered by using the Padé approximation.
- Evanescent waves are not dealt with by most BPMs, but one can include them by using a complex ϵ_c in Eq. (2.58).

2.4 SUMMARY

In this chapter, we described some field solvers for GRIN media, including Mie theory, the fiber mode solver, and several BPMs. The properties of the solvers described are shown in Tab. 1.

Many other solvers can be used to model GRIN components.

- Rigorous field solvers, e.g., finite element methods (FEM) [44] or Fourier modal methods with perfectly matched layers (FMM+PMLs) [62], can calculate general input field propagation through general GRIN components, with all the mentioned effects included.
- The Fourier modal method (FMM) can be used to model periodic GRIN components.

However, from Tab. 1, we also find that although there are many field solvers that model different types of GRIN distributions, or that include different physical effects, it is not always convenient to switch between solvers.

Table 1: Overview of different techniques

Techniques\Features		Input field	$\epsilon(\boldsymbol{r})$		Effects generated in GRIN media					
			sym.	$\nabla \ln \epsilon \approx 0$	surface	polari.	diffra.	non-para.	reflec.	eva. wave
Mie theory		plane wave†	sphe.		✓	✓	✓	✓	✓	✓
Fiber mode solver		general	cylind.	!		✓	✓	✓		
FT BPM		paraxial		!			✓			
WPM		general		!‡	✓	✓	✓	✓		
FD BPM	basic	paraxial		!			✓			
	wide angle	general		!			✓	✓		
	vectorial	paraxial		$\partial_z \epsilon \approx 0$		✓	✓			
	bidirectional	general		!			✓	✓	✓	✓
FMM		general	periodic		✓	✓	✓	✓	✓	✓
FMM+PMLs		general			✓	✓	✓	✓	✓	✓

† It can be extended to model the general field using plane wave decomposition and coordinate system transformation.
‡ It is compensated by including the Fresnel matrices of local change from $\epsilon(\boldsymbol{\rho}, z^-)$ to $\epsilon(\boldsymbol{\rho}, z^+)$.
! means that this solver has constraints.
✓ denotes that this solver correctly predicts the relevant effect.

3 RUNGE-KUTTA k-DOMAIN FIELD PROPAGATION METHOD

In Chapter 2, we reviewed several field solvers for modeling an electromagnetic field propagating through a graded-index (GRIN) medium. We mentioned three rigorous solvers that could predict all linear physical effects, e.g., polarization crosstalk, non-paraxial, or reflection, rigorously. However, Mie theory is only valid for spherically symmetric GRIN media, and Fourier modal method (FMM) only works for periodic GRIN media. FMM with perfectly matched layers (FMM+PMLs) is capable of modeling general GRIN media, but the numerical effort it demands, which is linear at N^3 (N is the number of sampling points), is quite high, and the sampling distance should be smaller than half a wavelength. Then, the concept and capabilities of beam propagation methods (BPMs), including the Fourier transform (FT-) BPM, the wave propagation method (WPM), and finite difference (FD-) BPMs, are discussed. BPMs are easy to implement and highly extensible. However, most BPMs have the constraint $\boldsymbol{\nabla} \ln \epsilon(\boldsymbol{r}) \approx 0$, which introduces inaccuracy in predicting polarization crosstalk, or even ignores it. Furthermore, too many BPMs require frequent selection and switching between proper solvers. To overcome the constraint, and to simplify the technique selection process, we would like to develop a conceptionally straightforward field solver to include multiple functionalities and to unify the benefits of several other solvers.

In this chapter, we derive a field solver directly from Maxwell's equations. From a theoretical perspective, we convert Maxwell's curl equations in the spatial domain (x-domain) into ordinary differential equations (ODEs) in the spatial-frequency domain (k-domain) [63]. The ODEs can be solved iteratively using the Runge-Kutta method. In our numerical calculation, we take advantage of the fast Fourier transform (FFT) to convert the convolution-type calculation involving permittivity and field components, which requires a linear numerical effort of N^2, to a simple multiplication that requires N operations, where N is the number of sampling points [64].

We first introduce the fundamental technique, including the derivation of the ODEs in the k-domain, using the Runge-Kutta method to solve them, and introduce the validity proof by comparing the result with that calculated by rigorous field solvers. Then, we propose ways to further reduce the numerical effort.

3.1 FUNDAMENTAL TECHNIQUE

3.1.1 *Ordinary Differential Equations (ODEs) in the k-Domain*

To derive a field solver, we start with the two Maxwell curl-equations, Eqs. (2.20) and (2.21),

$$\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}) = \mathrm{i}\omega\mu_0 \boldsymbol{H}(\boldsymbol{r}), \tag{3.1}$$

and

$$\boldsymbol{\nabla} \times \boldsymbol{H}(\boldsymbol{r}) = -\mathrm{i}\omega\epsilon_0\epsilon(\boldsymbol{r})\boldsymbol{E}(\boldsymbol{r}). \tag{3.2}$$

with $\boldsymbol{E}(\boldsymbol{r})$ and $\boldsymbol{H}(\boldsymbol{r})$ denoting the electric and magnetic field at position $\boldsymbol{r} = (x, y, z)$. In Eqs. (3.1) and (3.2), i is the imaginary unit, ϵ_0 and μ_0 are the vacuum permittivity and permeability constants, and $\epsilon(\boldsymbol{r})$ is the relative permittivity of the GRIN media, which is position-dependent. Please note that we neglect ω in the notation, and focus on the formula for single-frequency ω, but it is not an approximation, because all general fields can be represented as a linear combination of different spectrum modes with different ω, and in linear optics, there is no energy crosstalk between different spectrum modes.

If we represent all six electromagnetic field components as

$$\boldsymbol{V}(\boldsymbol{r}) = \{E_x, E_y, E_z, \sqrt{\frac{\mu_0}{\epsilon_0}}H_x, \sqrt{\frac{\mu_0}{\epsilon_0}}H_y, \sqrt{\frac{\mu_0}{\epsilon_0}}H_z\}^T(\boldsymbol{r}), \tag{3.3}$$

Eqs. (3.1) and (3.2) can be rewritten as

$$\begin{pmatrix} \partial_y V_3(\boldsymbol{r}) - \partial_z V_2(\boldsymbol{r}) \\ \partial_z V_1(\boldsymbol{r}) - \partial_x V_3(\boldsymbol{r}) \\ \partial_x V_2(\boldsymbol{r}) - \partial_y V_1(\boldsymbol{r}) \end{pmatrix} = \mathrm{i}k_0 \begin{pmatrix} V_4(\boldsymbol{r}) \\ V_5(\boldsymbol{r}) \\ V_6(\boldsymbol{r}) \end{pmatrix}, \tag{3.4}$$

and

$$\begin{pmatrix} \partial_y V_6(\boldsymbol{r}) - \partial_z V_5(\boldsymbol{r}) \\ \partial_z V_4(\boldsymbol{r}) - \partial_x V_6(\boldsymbol{r}) \\ \partial_x V_5(\boldsymbol{r}) - \partial_y V_4(\boldsymbol{r}) \end{pmatrix} = -\mathrm{i}k_0\epsilon(\boldsymbol{r}) \begin{pmatrix} V_1(\boldsymbol{r}) \\ V_2(\boldsymbol{r}) \\ V_3(\boldsymbol{r}) \end{pmatrix}, \tag{3.5}$$

with $\partial_i := \frac{\partial}{\partial i}$, $i = x, y, z$.

In an arbitrary plane z, we represent $V_\ell(\boldsymbol{r})$ with $\ell = 1, 2, 3, 4, 5, 6$ using the inverse Fourier transform between the x-domain and the k-domain, as in Eq. (2.33),

$$V_\ell(\boldsymbol{\rho}, z) = \mathcal{F}_k^{-1}\tilde{V}_\ell(\boldsymbol{\kappa}, z) = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} \mathrm{d}k_x\,\mathrm{d}k_y\; \tilde{V}_\ell(\boldsymbol{\kappa}, z)\exp(\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{\rho}), \tag{3.6}$$

with $\tilde{V}_\ell(\boldsymbol{\kappa}, z)$ as the field in the k-domain, $\boldsymbol{\rho} = (x, y)$ and $\boldsymbol{\kappa} = (k_x, k_y)$.

Similarly, $\epsilon(\boldsymbol{r})$ is represented by its inverse Fourier transform

$$\epsilon(\boldsymbol{\rho}, z) = \mathcal{F}_k^{-1}\tilde{\epsilon}(\boldsymbol{\kappa}, z) = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} \mathrm{d}k_x\,\mathrm{d}k_y\; \tilde{\epsilon}(\boldsymbol{\kappa}, z)\exp(\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{\rho}). \tag{3.7}$$

Substituting Eqs. (3.6) and (3.7) into Eqs. (3.4) and (3.5), we get

$$\begin{pmatrix} \mathrm{i}\kappa_y\tilde{V}_3(\boldsymbol{\kappa}, z) - \partial_z\tilde{V}_2(\boldsymbol{\kappa}, z) \\ \partial_z\tilde{V}_1(\boldsymbol{\kappa}, z) - \mathrm{i}\kappa_x\tilde{V}_3(\boldsymbol{\kappa}, z) \\ \mathrm{i}\kappa_x\tilde{V}_2(\boldsymbol{\kappa}, z) - \mathrm{i}\kappa_y\tilde{V}_1(\boldsymbol{\kappa}, z) \end{pmatrix} = \mathrm{i}k_0 \begin{pmatrix} \tilde{V}_4(\boldsymbol{\kappa}, z) \\ \tilde{V}_5(\boldsymbol{\kappa}, z) \\ \tilde{V}_6(\boldsymbol{\kappa}, z) \end{pmatrix}, \tag{3.8}$$

and

$$\begin{pmatrix} \mathrm{i}\kappa_y\tilde{V}_6(\boldsymbol{\kappa}, z) - \partial_z\tilde{V}_5(\boldsymbol{\kappa}, z) \\ \partial_z\tilde{V}_4(\boldsymbol{\kappa}, z) - \mathrm{i}\kappa_x\tilde{V}_6(\boldsymbol{\kappa}, z) \\ \mathrm{i}\kappa_x\tilde{V}_5(\boldsymbol{\kappa}, z) - \mathrm{i}\kappa_y\tilde{V}_4(\boldsymbol{\kappa}, z) \end{pmatrix} = -\mathrm{i}k_0\tilde{\epsilon}(\boldsymbol{\kappa}, z) * \begin{pmatrix} \tilde{V}_1(\boldsymbol{\kappa}, z) \\ \tilde{V}_2(\boldsymbol{\kappa}, z) \\ \tilde{V}_3(\boldsymbol{\kappa}, z) \end{pmatrix}, \tag{3.9}$$

with $*$ indicating convolution, so that $\tilde{\epsilon}(\boldsymbol{\kappa},z) * \tilde{V}_\ell(\boldsymbol{\kappa},z) = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} \mathrm{d}k_x'\,\mathrm{d}k_y'\,\tilde{\epsilon}(\boldsymbol{\kappa}',z)\tilde{V}_\ell(\boldsymbol{\kappa}-\boldsymbol{\kappa}',z)$.

Compared with Eqs. (3.4) and (3.5), Eqs. (3.8) and (3.9) contain a single differential variable, z, after ∂_x and ∂_y are replaced by $\mathrm{i}k_x$ and $\mathrm{i}k_y$, respectively. Thus, ∂_z can be replaced by $\mathrm{d}/\mathrm{d}z$, and the equations become ODEs. Furthermore, of the six equations in Eqs. (3.8) and (3.9), there are four ODEs, while the other two serve to conclude that field components $\tilde{V}_3$ and $\tilde{V}_6$ can be represented in terms of $\tilde{V}_1$, $\tilde{V}_2$, $\tilde{V}_4$ and $\tilde{V}_5$. Therefore, there are four independent field components, while the other two can be calculated from those four.

After some rearranging of the equations, the ODE in matrix form can be written as

$$\frac{d}{dz}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\\ \tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z) = \mathrm{i}k_0\begin{bmatrix} 0 & 0 & \frac{k_x}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_y}{k_0} & 1-\frac{k_x}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_x}{k_0}\\ 0 & 0 & \frac{k_y}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_y}{k_0}-1 & -\frac{k_y}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_x}{k_0}\\ -\frac{k_xk_y}{k_0^2} & \frac{k_x^2}{k_0^2}-\underline{\tilde{\boldsymbol{\epsilon}}} & 0 & 0\\ \underline{\tilde{\boldsymbol{\epsilon}}}-\frac{k_y^2}{k_0^2} & \frac{k_yk_x}{k_0^2} & 0 & 0\end{bmatrix}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\\ \tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z). \tag{3.10}$$

In Eq. (3.10), operators $\underline{\tilde{\boldsymbol{\epsilon}}}$ and $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}$ are convolution operators, i.e.,

$$\underline{\tilde{\boldsymbol{\epsilon}}}\tilde{V}_\ell(\boldsymbol{\kappa},z) = \tilde{\epsilon}(\boldsymbol{\kappa},z) * \tilde{V}_\ell(\boldsymbol{\kappa},z),\ \text{with}\ \ell = 1,2, \tag{3.11}$$

$$\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_i\tilde{V}_\ell(\boldsymbol{\kappa},z) = \widetilde{\epsilon^{-1}}(\boldsymbol{\kappa},z) * [k_i\tilde{V}_\ell(\boldsymbol{\kappa},z)],\ \text{with}\ i = x,y,\ \ell = 4,5, \tag{3.12}$$

where $\widetilde{\epsilon^{-1}}(\boldsymbol{\kappa},z) = \mathcal{F}_k\,\dfrac{1}{\epsilon(\boldsymbol{\rho},z)} = \dfrac{1}{2\pi}\iint_{-\infty}^{+\infty}\mathrm{d}x\,\mathrm{d}y\,\dfrac{\exp(-\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{\rho})}{\epsilon(\boldsymbol{\rho},z)}$.

Eq. (3.10) rigorously describes field behavior in GRIN media, without any physical approximation whatsoever as long as the problem is considered using the framework of linear optics.

In some practical applications, where both the structure and the fields are y-invariant, a two-dimensional (2D) treatment suffices. Substituting $\partial_y = 0$ into Eqs. (3.4) and (3.5), and following a derivation similar to that of Eq. (3.6) to (3.10), we get a simplified version of the ODE,

$$\frac{d}{dz}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\\ \tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(k_x,z) = \mathrm{i}k_0\begin{bmatrix} 0 & 0 & 0 & 1-\frac{k_x}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_x}{k_0}\\ 0 & 0 & -1 & 0\\ 0 & \frac{k_x^2}{k_0^2}-\underline{\tilde{\boldsymbol{\epsilon}}} & 0 & 0\\ \underline{\tilde{\boldsymbol{\epsilon}}} & 0 & 0 & 0\end{bmatrix}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\\ \tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(k_x,z). \tag{3.13}$$

with operators $\underline{\tilde{\boldsymbol{\epsilon}}}$ and $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}$ as the analogues to those of Eqs. (3.11) and (3.12), but replacing $\boldsymbol{\kappa}$ by k_x.

One point in the derivation of the 2D case is worthy of mention: after substituting $\partial_y = 0$ into Eqs. (3.4) and (3.5), we note that the six field components are separated into two independent groups,

$$\begin{pmatrix} \partial_x V_6(\boldsymbol{r}) - \partial_z V_4(\boldsymbol{r}) \\ -\partial_z V_2(\boldsymbol{r}) \\ \partial_x V_2(\boldsymbol{r}) \end{pmatrix} = \mathrm{i}k_0 \begin{pmatrix} V_2(\boldsymbol{r}) \\ \epsilon(\boldsymbol{r}) V_4(\boldsymbol{r}) \\ V_6(\boldsymbol{r}) \end{pmatrix}, \tag{3.14}$$

and

$$\begin{pmatrix} \partial_z V_5(\boldsymbol{r}) \\ -\partial_x V_5(\boldsymbol{r}) \\ \partial_z V_1(\boldsymbol{r}) - \partial_x V_3(\boldsymbol{r}) \end{pmatrix} = \mathrm{i}k_0 \epsilon(\boldsymbol{r}) \begin{pmatrix} V_1(\boldsymbol{r}) \\ V_3(\boldsymbol{r})/\epsilon(\boldsymbol{r}) \\ V_5(\boldsymbol{r}) \end{pmatrix}. \tag{3.15}$$

Eq. (3.14) contains only one electric field component, $V_2(\boldsymbol{r})$, which is transverse to the main propagation direction, and that is why this group of fields constitutes the so-called *transverse electric mode (TE)*. Similarly, Eq. (3.15) represents the behavior of *transverse magnetic mode (TM)*. The TE and TM modes are decoupled without any energy crosstalk between the two modes.

3.1.2 *Stability Analysis of the ODEs*

From this point onward, we illustrate how to solve Eq. (3.10). Eq. (3.13) is merely a simplified version of Eq. (3.10), and as such, the solution will be the same. We would like to emphasize that both 2D and 3D cases can be solved using the following technique.

Equation (3.10) is in the typical vectorial form of ODEs.

$$\frac{\mathrm{d}}{\mathrm{d}z} \tilde{\boldsymbol{V}}_\perp(z) = \boldsymbol{f}(z, \tilde{\boldsymbol{V}}_\perp), \tag{3.16}$$

with $\tilde{\boldsymbol{V}}_\perp(z) = (\tilde{V}_1, \tilde{V}_2, \tilde{V}_4, \tilde{V}_5)^{\mathrm{T}}(\boldsymbol{\kappa}, z)$, and $\boldsymbol{f}(z, \tilde{\boldsymbol{V}}_\perp)$ denoting the right-hand side of Eq. (3.10),

$$\boldsymbol{f}(z, \tilde{\boldsymbol{V}}_\perp) = \underline{\mathbf{M}}(\boldsymbol{\kappa}, z) \tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa}, z), \tag{3.17}$$

with

$$\underline{\mathbf{M}}(\boldsymbol{\kappa}, z) = \mathrm{i}k_0 \begin{bmatrix} 0 & 0 & \frac{k_x}{k_0} \underline{\tilde{\boldsymbol{\epsilon}}}^{-1} \frac{k_y}{k_0} & 1 - \frac{k_x}{k_0} \underline{\tilde{\boldsymbol{\epsilon}}}^{-1} \frac{k_x}{k_0} \\ 0 & 0 & \frac{k_y}{k_0} \underline{\tilde{\boldsymbol{\epsilon}}}^{-1} \frac{k_y}{k_0} - 1 & -\frac{k_y}{k_0} \underline{\tilde{\boldsymbol{\epsilon}}}^{-1} \frac{k_x}{k_0} \\ -\frac{k_x k_y}{k_0^2} & \frac{k_x^2}{k_0^2} - \underline{\tilde{\boldsymbol{\epsilon}}} & 0 & 0 \\ \underline{\tilde{\boldsymbol{\epsilon}}} - \frac{k_y^2}{k_0^2} & \frac{k_y k_x}{k_0^2} & 0 & 0 \end{bmatrix} \tag{3.18}$$

$\tilde{V}_\ell$ is a function of $\boldsymbol{\kappa}$, so the dimensionality of $\tilde{\boldsymbol{V}}_\perp(z)$ depends on the number of sampling points of $\tilde{V}_\ell$ in the k-domain. Assuming the number of sampling points of each $\tilde{V}_\ell$, $\ell = 1, 2, 4, 5$ is $N_x \times N_y$; then $\tilde{V}_\perp(z)$ is an M-dimensional vector, where $M = N_x \times N_y \times 4$. As z is a real value, the domain and range of $\tilde{V}_\perp(z)$ and the function $\boldsymbol{f}$ are

$$\begin{aligned} \tilde{\boldsymbol{V}}_\perp : &\quad \mathbb{R} \to \mathbb{C}^M \\ \boldsymbol{f} : &\quad \mathbb{R} \times \mathbb{C}^M \to \mathbb{C}^M \end{aligned}. \tag{3.19}$$

Now, let us solve Eq. (3.16). Mathematically, the unique solution of an ODE can be obtained by adding either "boundary conditions," if $\tilde{\boldsymbol{V}}_\perp$ on the boundaries of the domain is assigned, or "initial conditions," if $\tilde{\boldsymbol{V}}_\perp$ is specified at a single value of z [55]. The former solution with "boundary conditions" is formulated as an "eigenvalue problem," which is known to be a mode solver, and the latter with "initial conditions" is formulated as an "initial value problem." The FMM, which is also called rigorous coupled wave analysis, solves Eq. (3.16) as an "eigenvalue problem" [63, 65, 66]. Our main interest in this work is with the latter, more specifically, solving Eq. (3.16) to obtain $\tilde{\boldsymbol{V}}_\perp(z)$ for all $z > z_0$, under the assumption that the initial condition $\tilde{\boldsymbol{V}}_\perp(z_0)$ is known. The fact that there is no practical mechanism to include the reflected field in this initial condition results in the first constraint of the technique: **the reflection field inside the GRIN media is not calculated**.

To obtain the unique solution of the "initial value problem," the mathematical stability of the ODE needs to be checked [55], i.e., whether an arbitrary perturbation in the fields can be suppressed, or at least not enlarged. This stability check is not required when we solve an ODE as an "eigenvalue problem."

Now let us discuss whether Eq. (3.10) is stable, or discuss when it is stable, so that we obtain the validity criteria of the solution.

Mathematically, let us assume the perturbation is $\tilde{\boldsymbol{P}}(\boldsymbol{\kappa}, z)$, which has the same dimensionality as $\tilde{\boldsymbol{V}}(\boldsymbol{\kappa}, z)$, i.e., $\tilde{\boldsymbol{P}}_\perp : \mathbb{R} \to \mathbb{C}^M$. After replacing $\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa}, z)$ by $\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa}, z) + \tilde{\boldsymbol{P}}(\boldsymbol{\kappa}, z)$ in Eq. (3.10), we obtain the formula for $\tilde{\boldsymbol{P}}(\boldsymbol{\kappa}, z)$ in the form of a linear first-order ODE,

$$\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\boldsymbol{P}}(z) = \underline{\mathbf{M}}(\boldsymbol{\kappa}, z)\tilde{\boldsymbol{P}}(\boldsymbol{\kappa}, z), \tag{3.20}$$

with $\underline{\mathbf{M}}(\boldsymbol{\kappa}, z)$ in Eq. (3.18). The right-hand side of Eq. (3.20) can be described using linear algebra,

$$\underline{\mathbf{M}}(\boldsymbol{\kappa}, z)\tilde{\boldsymbol{P}}(\boldsymbol{\kappa}, z) = \eta\tilde{\boldsymbol{P}}(\boldsymbol{\kappa}, z), \tag{3.21}$$

with η as the eigenvalue of $\underline{\mathbf{M}}$*.

The mathematical solution of Eqs. (3.20-3.21) can be represented as,

$$\tilde{\boldsymbol{P}}(z) = \sum_i c_i \tilde{\boldsymbol{p}}_i(\boldsymbol{\kappa}) \exp(\eta_i z). \tag{3.22}$$

with η_i the eigenvalue and $\tilde{\boldsymbol{p}}_i(\boldsymbol{\kappa})$ the corresponding eigenvector. If a perturbation is introduced in position z_0, we need first to decompose $\tilde{\boldsymbol{P}}(z_0)$ into eigenvectors $\tilde{\boldsymbol{p}}_i(\boldsymbol{\kappa})$ to calculate c_i,

$$c_i = \iint \tilde{\boldsymbol{P}}(z_0)\tilde{\boldsymbol{p}}_i^\star(\boldsymbol{\kappa})\,\mathrm{d}k_x\,\mathrm{d}k_y, \tag{3.23}$$

with $\tilde{\boldsymbol{p}}^\star$ the conjugate vector of $\tilde{\boldsymbol{p}}$.

To ensure that the perturbation $\tilde{\boldsymbol{P}}(z_0)$ is not enlarged by Eq. (3.10), it is necessary that all eigenvalues η_i satisfy

$$\exp(\eta_i z) \leqslant 1. \tag{3.24}$$

* The symbol for an eigenvalue according to mathematical convention is λ. However, in this work, we use λ to denote the wavelength, so here we use η to denote eigenvalue.

Or

$$\mathbb{R}\mathrm{e}\{\eta_i\} \leqslant 0. \tag{3.25}$$

Then, we need to **calculate the maximum eigenvalue of $\underline{\mathbf{M}}$ for GRIN media, and check if $\eta_{\max}$ satisfies Eq. (3.25).**

The detailed calculation of the maximum eigenvalue of $\underline{\mathbf{M}}$ is in Appendix E. Here, we show the conclusions:

- The eigenvalue of the ODE for a homogeneous medium, with relative permittivity ϵ_c, is

$$\eta = \pm \mathrm{i} k_z = \pm \mathrm{i}\sqrt{k_0^2 \epsilon_c - ||\boldsymbol{\kappa}||^2}, \tag{3.26}$$

  and the real part is

$$\mathbb{R}\mathrm{e}\{\eta\} = \begin{cases} 0 & ||\boldsymbol{\kappa}||^2 \leqslant k_0^2 \epsilon_c \quad \text{propagating wave} \\ \pm |k_z| & ||\boldsymbol{\kappa}||^2 > k_0^2 \epsilon_c \quad \text{evanescent wave.} \end{cases} \tag{3.27}$$

  Thus, the ODE is stable when $||\boldsymbol{\kappa}||^2 \leqslant k_0^2 \epsilon_c$.

- For the ODE of general GRIN media, the range of the eigenvalue is

$$-k_0^2 \epsilon_{\max} \leqslant \eta^2 \leqslant ||\boldsymbol{\kappa}||_{\max}^2 - k_0^2 \epsilon_{\min}. \tag{3.28}$$

  when $||\boldsymbol{\kappa}||_{\max}^2 \leqslant k_0^2 \epsilon_{\min}$, $\mathbb{R}\mathrm{e}\{\eta\} = 0$.

At this point, we can conclude that in order to obtain a stable solution for the ODE, (3.16), the width of the angular spectrum should be limited. Thus, to ensure that the algorithm is stable going forward, we rewrite the ODE as

$$\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa}, z) = \begin{cases} \underline{\mathbf{M}}(\boldsymbol{\kappa}, z)\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa}, z) & \text{for } ||\boldsymbol{\kappa}||^2 \leqslant k_0^2 \epsilon_{\min} \\ 0 & \text{for } ||\boldsymbol{\kappa}||^2 > k_0^2 \epsilon_{\min} \end{cases}, \tag{3.29}$$

with $\underline{\mathbf{M}}$ as the matrix in Eq. (3.18).

From the above discussion about stability, we realize the second constraint of the technique, i.e., **it cannot correctly predict the evanescent wave, which has $||\boldsymbol{\kappa}|| > k_0\sqrt{\epsilon_{\min}}$.** Thus, if the evanescent wave plays an important role in the application, this technique is not the right choice.

3.1.3 *Numerical Solver Development*

The main idea in the approach to solving the initial value problem is that if a fine enough grid along z is defined, i.e., $\{z_0, z_1, \cdots z_{i-1}, z_i, z_{i+1}, \cdots z_{N_z-1}\}$, $\tilde{\boldsymbol{V}}_\perp(z_{i+1})$ can be calculated from $\tilde{\boldsymbol{V}}_\perp(z_i)$ using different methods, and using different approximations which are reasonable for small stretches of z. In this work, we use the 4^{th}-order Runge-Kutta method, which is formulated as follows.

$$\tilde{\boldsymbol{V}}_\perp(z_{i+1}) = \tilde{\boldsymbol{V}}_\perp(z_i) + \frac{1}{6}(\boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2 + 2\boldsymbol{\alpha}_3 + \boldsymbol{\alpha}_4), \tag{3.30}$$

with

$$\begin{cases} \boldsymbol{\alpha}_1 &= \Delta z_i \boldsymbol{f}(z_i, \tilde{\boldsymbol{V}}_\perp(z_i)) \\ \boldsymbol{\alpha}_2 &= \Delta z_i \boldsymbol{f}(z_i + \frac{1}{2}\Delta z_i, \tilde{\boldsymbol{V}}_\perp(z_i) + \frac{1}{2}\boldsymbol{\alpha}_1) \\ \boldsymbol{\alpha}_3 &= \Delta z_i \boldsymbol{f}(z_i + \frac{1}{2}\Delta z_i, \tilde{\boldsymbol{V}}_\perp(z_i) + \frac{1}{2}\boldsymbol{\alpha}_2) \\ \boldsymbol{\alpha}_4 &= \Delta z_i \boldsymbol{f}(z_{i+1}, \tilde{\boldsymbol{V}}_\perp(z_i) + \frac{1}{2}\boldsymbol{\alpha}_3) \end{cases}, \tag{3.31}$$

and

$$\Delta z_i = z_{i+1} - z_i. \tag{3.32}$$

To implement the iterative equations (3.30-3.32), we first need to know how to calculate $\boldsymbol{f}(z, \tilde{\boldsymbol{V}}_\perp)$. At a given plane z, $\boldsymbol{\kappa}$ is discretized as $\boldsymbol{\kappa}_{ij}$,

$$\boldsymbol{\kappa}_{ij} = (i\delta k_x, j\delta k_y) + (k_{x,\min}, k_{y,\min}),\ i, j \in \mathbb{Z},\ i \in [0, N_x - 1],\ j \in [0, N_y - 1], \tag{3.33}$$

with $(\delta k_x, \delta k_y)$ as the sampling distances, and $k_{x,\min}$ and $k_{y,\min}$ as the minimum value of k_x and k_y, respectively. $\tilde{\boldsymbol{V}}_\perp$ is a function of $\boldsymbol{\kappa}_{ij}$. Most of the operations in $\boldsymbol{f}(z, \tilde{\boldsymbol{V}}_\perp)$ are pointwise with respect to $\boldsymbol{\kappa}_{ij}$, which has a linear numerical effort (runtime), $(N_x \times N_y)$, other than operators $\underline{\tilde{\boldsymbol{\epsilon}}}$ and $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}$. From Eqs. (3.11) and (3.12), we know that $\underline{\tilde{\boldsymbol{\epsilon}}}$ and $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}$ are convolution operators, which have a numerical effort proportional to $(N_x \times N_y)^2$. However, based on the convolution theorem, the convolution in the k-domain can be realized in the x-domain with a Fourier transform pair (2.32-2.33) and an algebraic multiplication [47]. As a result, we rewrite Eqs. (3.11) and (3.12) as

$$\underline{\tilde{\boldsymbol{\epsilon}}} \tilde{V}_\ell(\boldsymbol{\kappa}, z) = \mathcal{F}_k \left\{ \epsilon(\boldsymbol{\rho}, z) \times \mathcal{F}_k^{-1} \left[\tilde{V}_\ell(\boldsymbol{\kappa}, z) \right] \right\} \text{ with } \ell = 1, 2, \tag{3.34}$$

$$\underline{\tilde{\boldsymbol{\epsilon}}}^{-1} k_i \tilde{V}_\ell(\boldsymbol{\kappa}, z) = \mathcal{F}_k \left\{ \epsilon^{-1}(\boldsymbol{\rho}, z) \times \mathcal{F}_k^{-1} \left[k_i \tilde{V}_\ell(\boldsymbol{\kappa}, z) \right] \right\},\ i = x, y, \text{with } \ell = 4, 5, \tag{3.35}$$

From Eqs. (3.34) and (3.35), if a Fourier transform (inverse) is performed using an FFT, the numerical effort of the operators $\underline{\tilde{\boldsymbol{\epsilon}}}$ and $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}$ becomes $[2 \times (N_x \times N_y) \log(N_x \times N_y) + (N_x \times N_y)]$, which is almost linear in $(N_x \times N_y)$. This offers the significant advantage of reducing numerical effort in the general field calculation.

Consequently, the total numerical effort of the solver becomes almost linear as $(N_x \times N_y \times N_z)$

Because we use the Runge-Kutta (RK) method to solve Eq. (3.29) in the k-domain, to calculate the field propagation through GRIN media, we name the algorithm RK k-domain field propagation method (RKK-FPM). Then we implement the algorithm using a customized programmable component in the physical-optics design software VirtualLab Fusion [26].

3.1.4 *Validation: y-Invariant Luneburg Lens*

In this example, the input plane wave is y-invariant and x-polarized with a wavelength of 532 nm, and the aperture size is 38 µm, as shown in Fig. 4. This plane wave propagates through a Luneburg lens [67] having a relative permittivity given by

$$\epsilon(x, z) = 2.0 - \frac{x^2 + z^2}{R^2}. \tag{3.36}$$

Here, (x, z) corresponds to the spatial coordinates and $R = 20\,\mu\text{m}$ to the radius of the Luneburg lens. We use both the RKK-FPM and a rigorous algorithm, i.e., the FMM+PMLs [62], available in the optical software VirtualLab Fusion [68], to calculate the field distribution in the xz-plane and at the output plane.

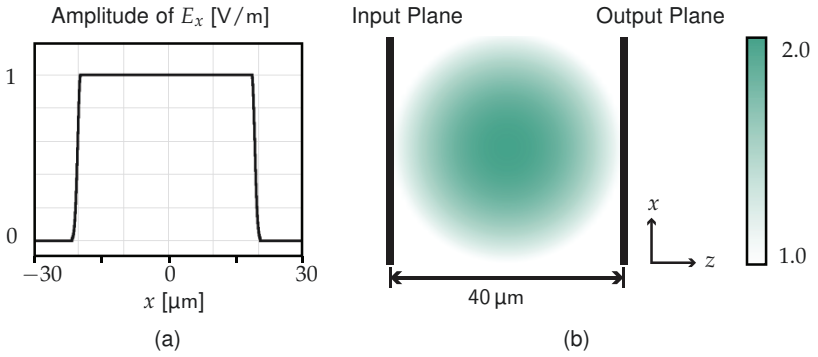


Figure 4: Illustration of the optical system. (a) The input field is an x-polarized truncated plane wave, with aperture size 38 µm. (b) The distribution of $\epsilon(x, z)$ of the Luneburg lens.

In Fig. 5, we show the plots of the amplitude of the field components E_x in the xz-plane, as computed by the RKK-FPM (top left) and the FMM+PMLs (bottom left). The plane wave is focused onto a rear point of the Luneburg lens with a convergence angle larger than 60°. Both methods give identical results, demonstrating the accuracy of the RKK-FPM to calculate the field in a non-paraxial situation. We compute the deviation to the referent field by using

$$\sigma = \frac{\sum_{x,y} |E_x(x,y) - E_x^{\text{ref}}|^2 + |E_y(x,y) - E_y^{\text{ref}}|^2}{|E_x^{\text{ref}}|^2 + |E_y^{\text{ref}}|^2}, \tag{3.37}$$

and the deviation is 0.1 %. The error comes from the two constraints of the RKK-FPM, i.e., no reflection, or evanescent waves are included.

The number of sampling points in this case is $N_x = 245$ (sampling distance 250 nm), with $N_z = 2000$ for the RKK-FPM. In this case, the FMM+PMLs also needs a few layers in the z-direction to resolve the variation in $\epsilon(x, z)$. Both calculations were performed on a personal computer with an Intel Core i7-7700HQ CPU @2.80 GHz and a total 32 GB RAM. The computational time for the FMM is about 150 minutes, while the RKK-FPM takes several seconds.

Fast and accurate calculation of the y-invariant case provides several benefits: (1) for the analysis of field properties in waveguides [37], y-invariant simulations are, in most cases, good enough; (2) y-invariant simulations are a means for offering a convergent prediction or the estimation of numerical parameters, i.e., sampling parameters in the xy plane or the step number along the propagation (z) axis for most three-dimensional (3D) cases.

Here we show how we select the sampling parameters:
(1) The Luneburg lens generates a high-numerical-aperture (high-NA) focus in the detec-

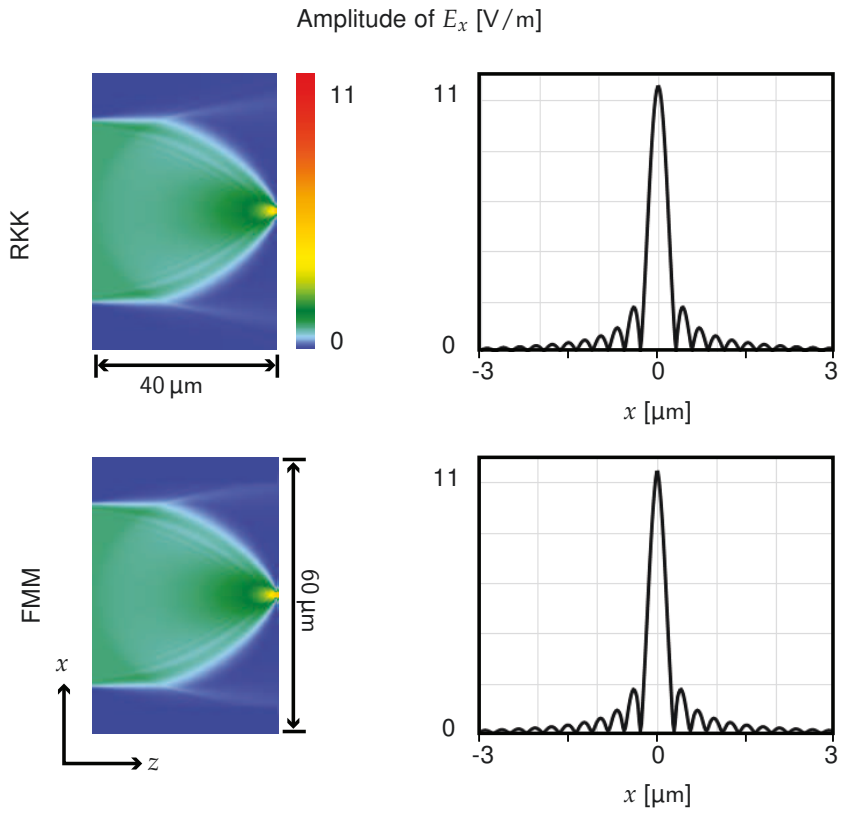


Figure 5: Amplitude distributions of the E_x field component in the xz plane (left column), as well as at the output plane (right column), calculated by the RKK-FPM (upper row) and FMM+PMLs (bottom row).

tor plane, so the width of k_x is the largest value we can work with, i.e., $\Delta k_x = 2k_0\sqrt{\epsilon_{\min}}$. From the Nyquist theory [69], the sampling distance in the x-domain is

$$\delta x = \frac{2\pi}{\Delta k_x} = \frac{\lambda}{2\sqrt{\epsilon_{\min}}} \approx 250 \text{ nm}. \tag{3.38}$$

(2) The windows size Δx should be large enough to avoid the aliasing effect. The input beam size is almost equal to the lens diameter, and some light escapes from the lens. Thus, we select $\Delta x = 60$ μm. Then the sampling point N_x is fixed as

$$N_x = \Delta x/\delta x \approx 245. \tag{3.39}$$

(3) We enlarge N_z and run the code for the RKK-FPM and save the result fields. Next, we select the N_z when the result field is convergent. Eq. (3.37) is used to evaluate the convergence. In this example, we select $N_z = 2000$.

3.1.5 *Validation: 3D Luneburg Lens*

In this example, we extend the second example to a 3D case, i.e., the input plane wave is polarized along x-axis, has a wavelength of 532 nm, and is truncated by a round aperture

of diameter 45 µm (larger than the sphere); the relative permittivity of the rotationally symmetric spherical Luneburg lens is

$$\epsilon(r) = 2.0 - \frac{r^2}{R^2} \tag{3.40}$$

with $r = \sqrt{x^2 + y^2 + z^2}$ and $R = 20\,\mu\text{m}$ as the radius of the Luneburg lens. The optical system in the xz-plane is still as illustrated by Fig. 4 (b). Instead of the FMM+PMLs, we use Mie theory, another rigorous solver introduced in Chapter 2, to calculate the reference result. To simulate the 3D field propagation in the GRIN sphere, the sampling points $N_x = N_y = 245$ and $N_z = 2000$ for the RKK-FPM. For Mie theory, we use 280 orders of spherical harmonics, and we use 500 layers along the radial direction to resolve $\epsilon(r)$, and in the detector, we calculate 200×200 sampling points within the window size $4\,\mu\text{m} \times 4\,\mu\text{m}$. In Fig. 6, we show the plots of the amplitude of the field compo-

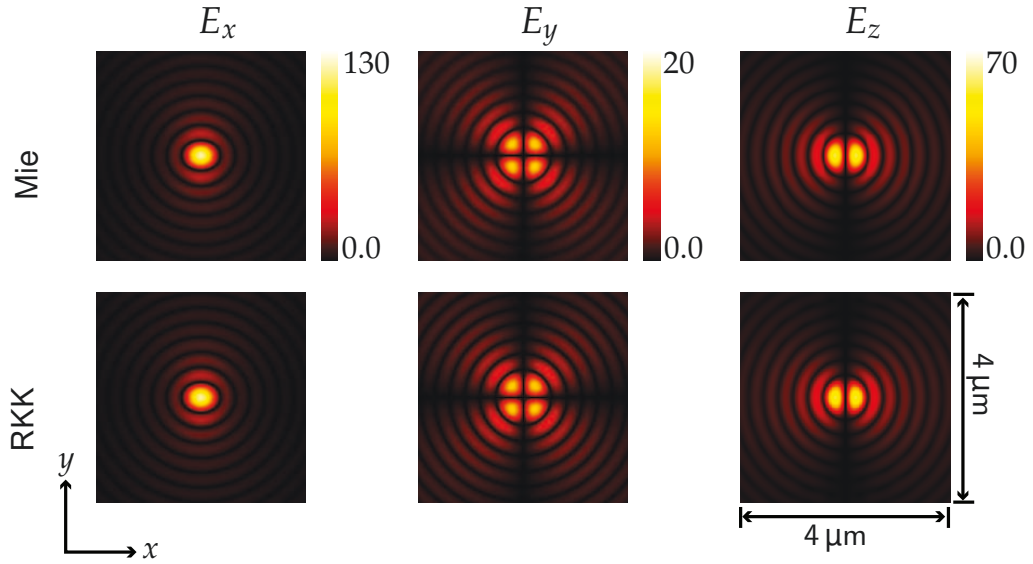


Figure 6: Amplitude distribution of the field components E_x, E_y and E_z at the output plane (Fig. 4 (b)).

nents E_x, E_y, and E_z at the output plane (Fig. 4 (b)), computed by the Mie theory (top) and RKK-FPM (bottom). E_x is not rotationally symmetric in the xy-plane, but instead is stretched along the x-axis. This stretching is expected for any imaging system with a high numerical aperture (NA) and is also related to the incident polarization state. Note that the relative permittivity at the interface is equal to that of the exterior embedding medium, $\epsilon(R) = 1.0$, which means that there is no interface between the surrounding medium and the Luneburg lens, so in this case, E_y is generated exclusively by the polarization crosstalk in the GRIN medium. Meanwhile, at the focal plane, E_z has a magnitude comparable to that of E_x. For a high-NA imaging system, polarization analysis becomes vitally important. Both methods give similar results (the deviation calculated by Eq. (3.37) is 0.4 %), demonstrating the accuracy of the RKK-FPM in calculating fields

in a non-paraxial situation. Meanwhile, it calculates accurately the polarization crosstalk between E_x, E_y, and E_z.

We also computed the focused field by using the FT-BPM and WPM, introduced in Chapter 2. FT-BPM ignores $\nabla \ln \epsilon$, so the calculated E_x is rotationally symmetric but has no stretching along the x-axis. WPM predicts the stretching by including the local Fresnel matrices between neighboring layers. However, WPM requires the numerical effort $(N_x \times N_y)^2$. Compared with the WPM, our RKK-FPM still has the advantage of lower numerical effort, which is linear at $N_x \times N_y$.

3.2 ODES WITH ANALYTICAL FAST-CHANGING PHASE TERM

We have discussed the fundamental RKK-FPM, and that its numerical effort is linear at $(N_x \times N_y \times N_z)$, with N_x, N_y, N_z as the number of sampling points in the x-, y-, and z-directions, respectively. We can further enhance the calculation efficiency in two ways: (1) reduce the number of sampling points N_z by extracting/analyzing the quickly changing phase term along the z-axis, and (2) reduce the number of sampling points $N_x \times N_y$ by extracting/analyzing the linear phase, or introducing the semi-analytical Fourier transform to replace the FFT where applicable. In this work, we realized the former.

The fast-changing phase is represented as $\exp(\mathrm{i}k_0\bar{n}z)$, with $\bar{n} = \sqrt{\epsilon_{\min}}$. Thus, the fields in Eq. (3.3) can be represented as

$$\boldsymbol{V}(\boldsymbol{r}) = \boldsymbol{U}(\boldsymbol{r}) \exp(\mathrm{i}k_0\bar{n}z), \tag{3.41}$$

and the field in the k-domain is

$$\tilde{\boldsymbol{V}}(\boldsymbol{\kappa}, z) = \tilde{\boldsymbol{U}}(\boldsymbol{\kappa}, z) \exp(\mathrm{i}k_0\bar{n}z). \tag{3.42}$$

Substituting the field representation (3.42) into the ODE (3.10), we obtain

$$\frac{d}{dz}\begin{pmatrix} \tilde{U}_1 \\ \tilde{U}_2 \\ \tilde{U}_4 \\ \tilde{U}_5 \end{pmatrix}(\boldsymbol{\kappa}, z) = \mathrm{i}k_0 \begin{bmatrix} -\bar{n} & 0 & \frac{k_x}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_y}{k_0} & 1 - \frac{k_x}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_x}{k_0} \\ 0 & -\bar{n} & \frac{k_y}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_y}{k_0} - 1 & -\frac{k_y}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_x}{k_0} \\ -\frac{k_x k_y}{k_0^2} & \frac{k_x^2}{k_0^2} - \underline{\tilde{\boldsymbol{\epsilon}}} & -\bar{n} & 0 \\ \underline{\tilde{\boldsymbol{\epsilon}}} - \frac{k_y^2}{k_0^2} & \frac{k_y k_x}{k_0^2} & 0 & -\bar{n} \end{bmatrix} \begin{pmatrix} \tilde{U}_1 \\ \tilde{U}_2 \\ \tilde{U}_4 \\ \tilde{U}_5 \end{pmatrix}(\boldsymbol{\kappa}, z). \tag{3.43}$$

In Eq. (3.46), operators $\underline{\tilde{\boldsymbol{\epsilon}}}$ and $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}$ are convolutional operators, analogs to those of Eqs. (3.11) and (3.12).

$$\underline{\tilde{\boldsymbol{\epsilon}}}\tilde{U}_\ell(\boldsymbol{\kappa}, z) = \tilde{\epsilon}(\boldsymbol{\kappa}, z) * \tilde{U}_\ell(\boldsymbol{\kappa}, z), \text{ with } \ell = 1, 2, \tag{3.44}$$

$$\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_i\tilde{U}_\ell(\boldsymbol{\kappa}, z) = \widetilde{\epsilon^{-1}}(\boldsymbol{\kappa}, z) * [k_i\tilde{U}_\ell(\boldsymbol{\kappa}, z)], \text{ with } i = x, y, \ \ell = 4, 5, \tag{3.45}$$

where $\widetilde{\epsilon^{-1}}(\boldsymbol{\kappa}, z) = \mathcal{F}_k \dfrac{1}{\epsilon(\boldsymbol{\rho}, z)} = \dfrac{1}{2\pi}\iint_{-\infty}^{+\infty} \mathrm{d}x\,\mathrm{d}y\, \dfrac{\exp(-\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{\rho})}{\epsilon(\boldsymbol{\rho}, z)}$.

Eq. (3.43) rigorously describes field behavior in GRIN media, without any physical approximation whatsoever, as long as the problem is considered in the framework of linear optics.

In this section, our purpose is to solve Eq. (3.46) to achieve $\tilde{\boldsymbol{U}}_\perp(z) = (\tilde{U}_1, \tilde{U}_2, \tilde{U}_4, \tilde{U}_5)^{\mathrm{T}}(\boldsymbol{\kappa}, z)$, and then $\tilde{\boldsymbol{V}}_\perp(z) = (\tilde{V}_1, \tilde{V}_2, \tilde{V}_4, \tilde{V}_5)^{\mathrm{T}}(\boldsymbol{\kappa}, z)$ can be calculated analytically using Eq. (3.42). Here, we would like to emphasis that the fast phase term $\exp(\mathrm{i}k_0\bar{n}z)$ is treated analytically without any approximation.

3.2.1 *Stability Analysis of the ODEs*

Eq. (3.43) is also in a typical vectorial form of ODE,

$$\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\boldsymbol{U}}_\perp(z) = \boldsymbol{f}_u(z, \tilde{\boldsymbol{U}}_\perp), \tag{3.46}$$

with $\tilde{\boldsymbol{U}}_\perp(z) = (\tilde{U}_1, \tilde{U}_2, \tilde{U}_4, \tilde{U}_5)^{\mathrm{T}}(\boldsymbol{\kappa}, z)$ and $\boldsymbol{f}_u(z, \tilde{\boldsymbol{U}}_\perp)$ denoting the right-hand side of Eq. (3.43),

$$\boldsymbol{f}_u(z, \tilde{\boldsymbol{U}}_\perp) = \underline{\mathbf{M}}_u(\boldsymbol{\kappa}, z)\tilde{\boldsymbol{U}}_\perp(\boldsymbol{\kappa}, z), \tag{3.47}$$

with

$$\underline{\mathbf{M}}_u(\boldsymbol{\kappa}, z) = \mathrm{i}k_0 \begin{bmatrix} -\bar{n} & 0 & \frac{k_x}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_y}{k_0} & 1 - \frac{k_x}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_x}{k_0} \\ 0 & -\bar{n} & \frac{k_y}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_y}{k_0} - 1 & -\frac{k_y}{k_0}\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\frac{k_x}{k_0} \\ -\frac{k_x k_y}{k_0^2} & \frac{k_x^2}{k_0^2} - \underline{\tilde{\boldsymbol{\epsilon}}} & -\bar{n} & 0 \\ \underline{\tilde{\boldsymbol{\epsilon}}} - \frac{k_y^2}{k_0^2} & \frac{k_y k_x}{k_0^2} & 0 & -\bar{n} \end{bmatrix} \tag{3.48}$$

$\tilde{U}_\ell$ is a function of $\boldsymbol{\kappa}$, so the dimensionality of $\tilde{\boldsymbol{U}}_\perp(z)$ depends on the number of sampling points of $\tilde{U}_\ell$ in the k-domain. Assuming the number of sampling points of all $\tilde{U}_\ell$, $\ell = 1, 2, 4, 5$ is $N_x \times N_y$, $\tilde{U}_\perp(z)$ is an $M-$dimensional vector, and $M = N_x \times N_y \times 4$. As z is a real value, the domain and range of $\tilde{U}_\perp(z)$, and the function $\boldsymbol{f}_u$ are

$$\begin{aligned} \tilde{\boldsymbol{U}}_\perp: &\quad \mathbb{R} \to \mathbb{C}^M \\ \boldsymbol{f}_u: &\quad \mathbb{R} \times \mathbb{C}^M \to \mathbb{C}^M \end{aligned}. \tag{3.49}$$

Again, we solve Eq. (3.46) as an "initial value problem." To achieve a converged, reliable result, we need not only a sufficient number of sampling points, but also the mathematical stability of the ODEs, analogs to the discussion of the fundamental RKK-FPM. More specifically, the eigenvalue η of $\underline{\mathbf{M}}_u$ should satisfy the criteria in Eq. (3.25). The detailed mathematical derivation is in Appendix E. The range of η is

$$-k_0^2\epsilon_{\min} \leqslant (\eta + 2\mathrm{i}k_0\bar{n})^2 \leqslant ||\boldsymbol{\kappa}||_{\max}^2 - k_0^2\epsilon_{\min}. \tag{3.50}$$

Thus, $||\boldsymbol{\kappa}||_{\max}^2 \leqslant k_0^2\epsilon_{\min}$ grants that $\mathbb{R}\mathrm{e}\{\eta\} \leqslant 0$.

At this point, we can conclude that in order to obtain a stable solution for the ODE (3.46), the width of angular spectrum should be limited. Thus, to ensure that the algorithm is stable going forward, we rewrite the ODE as

$$\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\boldsymbol{U}}_{\perp}(\boldsymbol{\kappa},z) = \begin{cases} \underline{\mathbf{M}}_u(\boldsymbol{\kappa},z)\tilde{\boldsymbol{U}}_{\perp}(\boldsymbol{\kappa},z) & \text{for } ||\boldsymbol{\kappa}||^2 \leqslant k_0^2\epsilon_{\min} \\ 0 & \text{for } ||\boldsymbol{\kappa}||^2 > k_0^2\epsilon_{\min} \end{cases}, \tag{3.51}$$

with $\underline{\mathbf{M}}_u$ as the matrix in Eq. (3.48).

From the above discussion about stability, the constraint of this technique is identical to that of the fundamental technique of the previous section, i.e., it cannot correctly predict the evanescent wave, which has $||\boldsymbol{\kappa}|| > k_0\sqrt{\epsilon_{\min}}$. Thus, if the evanescent wave plays an important role, this technique is not the correct choice.

3.2.2 *Validation: Multimode Fiber*

Here we use a fiber example to show the workflow when using the RKK-FPM.

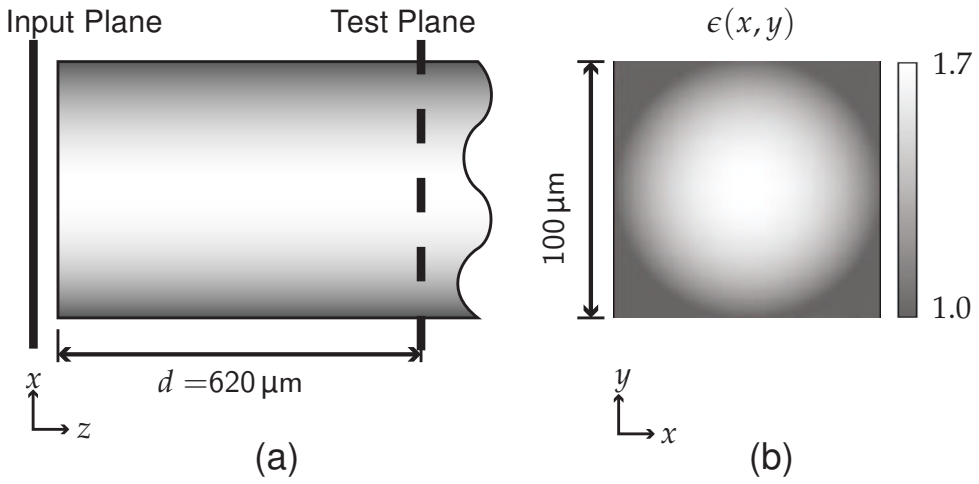


Figure 7: Illustration of the optical system. (a) The dashed line represents the test output plane located at a distance of $d = 620\,\mu\text{m}$ from the input plane (long solid vertical line). (b) The relative permittivity inside the fiber, which has a core diameter of 100 µm.

In this example, we model a y-polarized fundamental Gaussian mode (i.e., $E_x(x,y) = 0$) with a beam waist radius of 5 µm and with a wavelength of 532 nm propagating in a GRIN optical fiber, as shown in Fig. 7 (a). The spatial distribution of the relative permittivity $\epsilon(\boldsymbol{r})$ is presented in Fig. 7 (b). We use the RKK-FPM to calculate the field distribution in the test plane after the field has propagated a distance of $d = 620\,\mu\text{m}$ from the input plane.

Let us first fix the sampling parameters in this case by assuming a y-invariant situation, i.e., propagating a y-polarized and y-invariant Gaussian with a beam waist radius

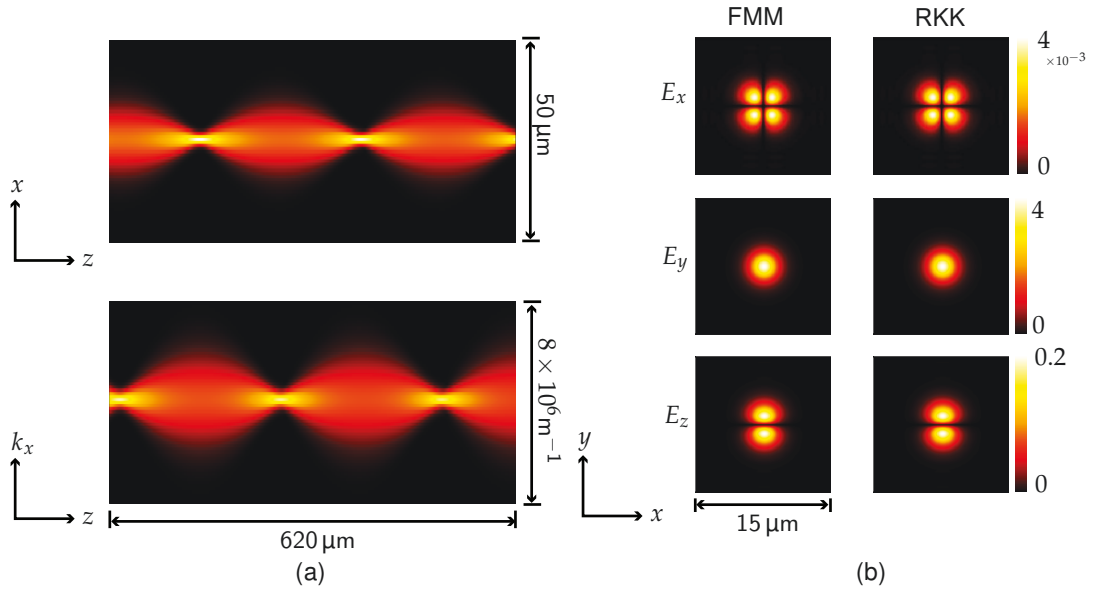


Figure 8: Fields in the multimode fiber. (a) The amplitude of E_y in the xz-plane (top) and $\tilde{E}_y$ in the $k_x z$ plane (bottom) in the y-invariant case. The right part shows the amplitude distributions of the field components E_x, E_y, and E_z in the test plane, calculated using the RKK-FPM; the unit is [$\mathrm{V\,m^{-1}}$]. The field components are interpolated using Sinc interpolation [26, 70].

of 5 µm through a fiber with an $\epsilon(x, z)$ identical to $\epsilon(x, y = 0, z)$ in Fig. 7. The amplitude of E_y and angular spectrum $\tilde{E}_y$ are shown in Fig. 8 (a).
(1) The spectrum width of $\tilde{E}_y$ is $\Delta k_x = 8 \times 10^6$ $\mathrm{m^{-1}}$, and from the Nyquist sampling theory, we calculate the sampling distance δx as

$$\delta x = \frac{2\pi}{\Delta k_x} \approx 800 \text{ nm}.$$

(2) The window size in the x-domain is about $\Delta x = 50\,\mu\text{m}$, so the number of sampling points is

$$N_x = \frac{\Delta x}{\delta x} \approx 63.$$

(3) Then, we perform the simulation with different N_z until the result is converged, and we select $N_z = 6200$.

Using the sampling parameters, i.e., $N_x = N_y = 63$, $\delta x = \delta y = 800\,\text{nm}$, and $N_z = 6200$, we perform a simulation of the 3D example. In Fig. 8 (b), we show the plots of the amplitudes of the field components E_x, E_y and E_z in the test plane, computed using the RKK-FPM (right column).

It is worth mentioning that the RKK-FPM itself does not include the modeling of the optical surface between the surrounding media and the GRIN component. In this example, the plane is modeled using a local plane interface approximation [29, 30].

We also use the FMM+PMLs to simulate this example, and the result field is shown in the left column of Fig. 8 (b). The relative deviation between the two result fields

calculated by Eq. (3.37) is 0.02 %, demonstrating the accuracy of the RKK-FPM. Compared with the FMM+PMLs, the RKK-FPM has a significant advantage in terms of numerical effort (computational time). The numerical effort of the RKK-FPM is linear at $(N_x \times N_y \times N_z)$, while that of the FMM is generally proportional to $(N_x \times N_y)^3$. To calculate the results presented in Fig. 8 (b) using the FMM on a workstation with an Intel Xeon CPU E5-2643 v2@3.5 GHz and a total 128 GB RAM, the simulation ran for over 40 hours, while with the RKK-FPM, only 5 minutes are required on a personal computer with an Intel Core i7-7700HQ CPU @2.80 GHz and a total 32 GB RAM.

3.2.3 *Comparison with the Fundamental Technique*

We use the RKK-FPM with a reduction of N_z to simulate the 3D Luneburg lens example presented in Section 3.1.5. The result fields are identical to those in Fig. 6. The numerical sampling parameters, and the simulation time when performed on a personal computer with an Intel Core i7-7700HQ CPU @2.80 GHz and a total 32 GB RAM, are shown in Tab. 2. From the table, the simulation time is shortened by extracting the fast-changing phase.

Table 2: 3D Luneburg lens simulation

RKK-FPM	$N_x \times N_y$	N_z	Time
Fundamental	245×245	2000	$\sim$ 20 min
Reduce N_z	245×245	400	$\sim$ 4 min

It is worth mentioning that the field ansatz and further derivation do not introduce any other physical approximations but do introduce the advantage of reducing the numerical effort required. Thus, in further discussion, we do not distinguish the fundamental method, or the method with reduced N_z. The RKK-FPM always analyzes the analytical phase term $\exp(\mathrm{i}k_0\bar{n}z)$ with $\bar{n} = \sqrt{\epsilon_{\min}}$. In this method, we always use the field ansatz (3.41) and solve the ODE (3.47) using the Runge-Kutta method.

3.3 SUMMARY

We have proposed a conceptually straightforward method for the fast calculation of electromagnetic fields propagating through GRIN media. This method converts Maxwell's curl equations into ordinary differential equations in the k-domain, via the FFT. We solve the equations using the Runge-Kutta method, hence the chosen denomination, RK k-domain field propagation method (RKK-FPM).

The properties of the RKK-FPM are:

- The input field can be a general field, either geometric or diffractive, and paraxial or non-paraxial.

- The GRIN media have general $\epsilon(\boldsymbol{r})$, with no restriction in symmetry or periodicity.
- Effects generated by a GRIN medium:
  - It works for a non-paraxial field, and always takes the polarization into account so that the vectorial field components are well-calculated.
  - Reflection is not included.
  - Evanescent waves, where $||\boldsymbol{\kappa}|| > k_0\sqrt{\epsilon_{\min}}$, are not included.

The RKK-FPM has a significant advantage in terms of numerical effort (calculation time), i.e., it is linear at $(N_x \times N_y \times N_z)$, with N_x, N_y, and N_z as the numbers of sampling points in the x-, y-, and z-directions, respectively.

Compared with other modeling techniques presented in Chapter 2, the RKK-FPM has no restriction with respect to the GRIN distribution ϵ, e.g., symmetry or small variance. We rearrange Tab. 1: The RKK-FPM can completely replace the FT-BPM, FD-BPM basic, wide angle method, and vectorial method. Although the bidirectional FD-BPM can model reflection and evanescent waves, it neglects $\nabla \ln \epsilon$, and introduces inaccuracy in the polarization crosstalk calculation, which is not negligible, especially when the beam is non-paraxial. Thus, in Tab. 3, we remove the bidirectional FD-BPM. Currently, when reflection and evanescent waves play an important role in the simulation, we suggest using the rigorous field solvers Mie, FMM, or FMM+PMLs.

Table 3: Field solvers to model GRIN media (Chapter 3)

Solvers/Feat.	Input field	$\epsilon(\boldsymbol{r})$		Effects generated in GRIN media			
		sym.	$\boldsymbol{\nabla} \ln \epsilon \approx 0$	surface	polari.	diffr.	ref/eva.
Mie theory	plane wave[†]	sphe.		✓	✓	✓	✓
FMM	general	periodic		✓	✓	✓	✓
FMM+PMLs	general			✓	✓	✓	✓
Fiber solver	general	cylind.	!		✓	✓	
WPM	general		!‡	✓	✓	✓	
RKK-FPM	general				✓	✓	

[†] It can be extended to model a general field by using plane wave decomposition and coordinate system transformation.

‡ It is compensated by including the Fresnel matrices of local change from $\epsilon(\boldsymbol{\rho}, z^-)$ to $\epsilon(\boldsymbol{\rho}, z^+)$.

! means that this solver has constraints.

✓ denotes that this solver correctly predicts the relevant effect.

In addition to the physical effects, the numerical effort needs to be considered when a field solver is selected to model general GRIN media. At this time, the FMM+PMLs, WPM, and RK k-domain field propagation method can be selected: the FMM+PMLs is mainly proportional to $(N_x \times N_y)^3 \times N_z^{\mathrm{FMM}}$, the WPM is proportional to $(N_x \times N_y)^2 \times N_z$, whereas the RK k-domain field propagation method is linear at $(N_x \times N_y \times N_z)$. Here, we distinguish N_z and N_z^{FMM}, because the selection of N_z^{FMM} depends on how many layers are needed to resolve $\epsilon(\boldsymbol{r})$, and generally the number of layers is not as large as N_z, as required the other two solvers, e.g., for z-invariant GRIN media, $N_z^{\mathrm{FMM}} = 1$.

However, the RK k-domain field propagation method does not include the surface effects efficiently. Thus, when the surface effects are important, we suggest using the WPM as an alternative.

4

RUNGE-KUTTA x-DOMAIN FIELD PROPAGATION METHOD

As most graded-index (GRIN) lenses and GRIN beam-shaping elements are still modeled or designed using ray tracing techniques, the field passing through the GRIN components must be a field with a negligible diffraction, which is known as a geometric field. In this chapter, we study specific field solvers for geometric fields. Starting with the ordinary differential equation (ODE) of the Runge-Kutta k-domain field propagation method (RKK-FPM), we replace the fast Fourier transform (FFT) with the specific transform for geometric fields, known as the homeomorphic Fourier transform (HFT) [33]. Then, the fundamental theory of ray optics, the eikonal equation, is derived. Substituting the geometric field ansatz and the eikonal equation into Maxwell's equations, we obtain the geometric field equations. After further derivation with GRIN media constraints, we obtain another set of ODEs represented in the x-domain. We solve them using our developed Runge-Kutta x-domain field propagation method (RKX-FPM).

4.1 FUNDAMENTAL TECHNIQUE TO MODEL GEOMETRIC FIELDS

4.1.1 *Geometric Fields: Eikonal and Field Equations*

The electromagnetic fields can be represented as

$$\boldsymbol{V}(\boldsymbol{r}) = \boldsymbol{U}(\boldsymbol{r}) \exp[i\psi(\boldsymbol{r})], \tag{4.1}$$

with $\boldsymbol{V}(\boldsymbol{r}) = \{E_x, E_y, E_z, \sqrt{\frac{\epsilon_0}{\mu_0}} H_x, \sqrt{\frac{\epsilon_0}{\mu_0}} H_y, \sqrt{\frac{\epsilon_0}{\mu_0}} H_z\}(\boldsymbol{r})$. $\psi(\boldsymbol{r})$ is a common phase function, which is extracted from the electromagnetic fields $\boldsymbol{V}(\boldsymbol{r})$ at position $\boldsymbol{r}$. Extracting the common phase function $\psi(\boldsymbol{r})$ leaves out the residual fields $\boldsymbol{U}(\boldsymbol{r}) := \boldsymbol{V}(\boldsymbol{r}) \exp[-i\psi(\boldsymbol{r})]$, which are still generally complex-valued. The choice of $\psi(\boldsymbol{r})$ is free in principle, and Eq. (4.1) is the general representation of electromagnetic fields without any approximation.

A well-chosen $\psi(\boldsymbol{r})$ may greatly simplify the simulation, and therefore it is of practical relevance. In optics, it is often possible to select $\psi(\boldsymbol{r})$ so that the behavior of the electromagnetic field is dominated by the phase part. More specifically, $\boldsymbol{U}(\boldsymbol{r})$ is slowly varying spatially, and $\psi(\boldsymbol{r})$ is more rapidly varying. Note that it is not always possible to find a common phase function that satisfies the phase-dominate constraint, and only when the constraint holds can we say that the field is in its **geometric field zone** [32], or that it is a **geometric field**. Otherwise, the field is said to be in the diffractive zone. In this algorithm, we focus on electromagnetic field propagation through GRIN media when the field is in the geometric zone.

After defining the geometric field, we explore its propagation behavior in inhomogeneous and isotropic media. In this derivation, we need to use two tools, i.e., the homeomorphic Fourier transform, which is the Fourier transform of a geometric field in an arbitrary plane z [33] and the ODEs in the k-domain (3.10), which rigorously describes field propagation in isotropic inhomogeneous media.

Homeomorphic Fourier Transform (HFT)

Based on the field ansatz (4.1), and the definition of geometric fields, in [33] the authors derive the Fourier transform for a geometric field using the stationary phase method [71, 72]. The Fourier transform is called the **homeomorphic Fourier transform** (HFT),

$$\tilde{V}_\ell(\boldsymbol{\kappa}, z) = \boldsymbol{\mathcal{F}}_k^h[V_\ell(\boldsymbol{\rho}, z)] = a[\boldsymbol{\rho}(\boldsymbol{\kappa}), z] U_\ell[\boldsymbol{\rho}(\boldsymbol{\kappa}), z] \exp\{\mathrm{i}\psi[\boldsymbol{\rho}(\boldsymbol{\kappa}), z] - \mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{\rho}\}, \tag{4.2}$$

with $\boldsymbol{\mathcal{F}}_k^h$ denoting the HFT, $\boldsymbol{\rho} = (x, y)$, and $\boldsymbol{\kappa} = (k_x, k_y)$. V_ℓ and U_ℓ, with $\ell = 1, 2, 3, 4, 5, 6$ represent the field components in Eq. (4.1). $\boldsymbol{\rho}(\boldsymbol{\kappa})$ is the mapping relation between $\boldsymbol{\rho}$ and $\boldsymbol{\kappa}$, which can be calculated by

$$\boldsymbol{\nabla}_\perp \psi(\boldsymbol{\rho}, z) = (\frac{\partial}{\partial x}\psi, \frac{\partial}{\partial y}\psi)(\boldsymbol{\rho}, z) = \boldsymbol{\kappa}. \tag{4.3}$$

a is a scaling factor, which is calculated by the second derivative of ψ,

$$a(\boldsymbol{\rho}, z) = \begin{cases} \sqrt{\dfrac{\mathrm{i}}{\psi_{xx}(\boldsymbol{\rho}, z)}} \sqrt{-\dfrac{\mathrm{i}\psi_{xx}(\boldsymbol{\rho}, z)}{\psi_{xy}^2(\boldsymbol{\rho}, z) - \psi_{xx}(\boldsymbol{\rho}, z)\psi_{yy}(\boldsymbol{\rho}, z)}} & , \quad \psi_{xx}(\boldsymbol{\rho}, z) \neq 0 \\ \dfrac{1}{\psi_{xy}(\boldsymbol{\rho}, z)} & , \quad \psi_{xx}(\boldsymbol{\rho}, z) = 0 \end{cases} \tag{4.4}$$

where $\psi_{ij} := \dfrac{\partial^2 \psi}{\partial i \partial j}$.

It is interesting that the field value in the x-domain is directly mapped/projected to the k-domain via the mapping relation (4.3). This phenomenon is quite different from our conventional understanding of the Fourier transform integral, which is that the field value at one $\boldsymbol{\kappa}$ is related to the field values at all $\boldsymbol{\rho}$s. However, for the geometric field, the field value at one $\boldsymbol{\kappa}$ is only related to that at one $\boldsymbol{\rho}$.

From Eq. (4.2), we see that $\tilde{\boldsymbol{V}}(\boldsymbol{\kappa}, z)$ can also be written as a multiplication of a slowly varying complex amplitude and a phase term,

$$\tilde{\boldsymbol{V}}(\boldsymbol{\kappa}, z) = \tilde{\boldsymbol{A}}(\boldsymbol{\kappa}, z) \exp\{\mathrm{i}\tilde{\psi}(\boldsymbol{\kappa}, z)\}, \tag{4.5}$$

with

$$\tilde{\boldsymbol{A}}(\boldsymbol{\kappa}, z) = a[\boldsymbol{\rho}(\boldsymbol{\kappa}), z] U_\ell[\boldsymbol{\rho}(\boldsymbol{\kappa}), z], \tag{4.6}$$

and

$$\tilde{\psi}(\boldsymbol{\kappa}, z) = \psi[\boldsymbol{\rho}(\boldsymbol{\kappa}), z] - \boldsymbol{\kappa}\cdot\boldsymbol{\rho}. \tag{4.7}$$

The inverse HFT from the k-domain to the x-domain is

$$V_\ell(\boldsymbol{\rho}, z) = \boldsymbol{\mathcal{F}}_k^{\mathrm{h},-1}[\tilde{V}_\ell(\boldsymbol{\kappa}, z)] = \tilde{a}[\boldsymbol{\kappa}(\boldsymbol{\rho})] \tilde{A}_\ell[\boldsymbol{\kappa}(\boldsymbol{\rho})] \exp\{\mathrm{i}\tilde{\psi}[\boldsymbol{\kappa}(\boldsymbol{\rho})] + \mathrm{i}\boldsymbol{\rho}\cdot\boldsymbol{\kappa}\}, \tag{4.8}$$

with the mapping relation

$$\tilde{\boldsymbol{\nabla}}_{\perp}\tilde{\psi}(\boldsymbol{\kappa},z)=(\frac{\partial}{\partial k_x}\tilde{\psi},\frac{\partial}{\partial k_y}\tilde{\psi})(\boldsymbol{\kappa},z)=-\boldsymbol{\rho}, \tag{4.9}$$

and with

$$\tilde{a}(\boldsymbol{\kappa},z)=\begin{cases}\sqrt{\dfrac{\mathrm{i}}{\tilde{\psi}_{k_xk_x}(\boldsymbol{\kappa},z)}}\sqrt{-\dfrac{\mathrm{i}\tilde{\psi}_{k_xk_x}(\boldsymbol{\kappa},z)}{\tilde{\psi}^2_{k_xk_y}(\boldsymbol{\kappa},z)-\tilde{\psi}_{k_xk_x}(\boldsymbol{\kappa},z)\tilde{\psi}_{k_yk_y}(\boldsymbol{\kappa},z)}} & , \quad \tilde{\psi}_{k_xk_x}(\boldsymbol{\kappa},z)\neq 0\\ \dfrac{1}{\tilde{\psi}_{k_xk_y}(\boldsymbol{\kappa},z)} & , \quad \tilde{\psi}_{k_xk_x}(\boldsymbol{\kappa},z)=0.\end{cases} \tag{4.10}$$

The properties of a geometric field in plane z are described by the HFT, i.e., (1) the field at one position $\boldsymbol{\rho}$ contributes to the field at one $\boldsymbol{\kappa}$, and vice versa; (2) the mapping relation $\boldsymbol{\rho}(\boldsymbol{\kappa})$ and $\boldsymbol{\kappa}(\boldsymbol{\rho})$ are only determined by the local gradient of the phases $\psi(\boldsymbol{\rho})$ and $\tilde{\psi}(\boldsymbol{\kappa})$, respectively.

Second-order ODEs in the k-domain

The properties of a geometric field in plane z are described by the HFT. Next, we explore its propagation in GRIN media. This property can be derived using the ODE (3.10) in the k-domain. First, the GRIN medium is described as having a sufficient number of thin layers along the z axis, and for each layer, the $z-$variance can be neglected, i.e., $\frac{\mathrm{d}\epsilon(\boldsymbol{r})}{\mathrm{d}z}\approx 0$ (for z-invariant media, the number of layers is 1, e.g., fiber). Then, for each layer, we derive the second-order ODE from Eq. (3.10)

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\\ \tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z)=\boldsymbol{M}(\boldsymbol{\kappa},z)\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\\ \tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z), \tag{4.11}$$

with

$$\boldsymbol{M}(\boldsymbol{\kappa},z)=\begin{bmatrix}-k_0^2\tilde{\underline{\boldsymbol{\epsilon}}}+k_y^2+k_x\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_x\tilde{\underline{\boldsymbol{\epsilon}}} & k_x\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_y\tilde{\underline{\boldsymbol{\epsilon}}}-k_xk_y & 0 & 0\\ k_y\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_x\tilde{\underline{\boldsymbol{\epsilon}}}-k_xk_y & -k_0^2\tilde{\underline{\boldsymbol{\epsilon}}}+k_x^2+k_y\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_y\tilde{\underline{\boldsymbol{\epsilon}}} & 0 & 0\\ 0 & 0 & -k_0^2\tilde{\underline{\boldsymbol{\epsilon}}}+k_x^2+\tilde{\underline{\boldsymbol{\epsilon}}}k_y\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_y & -\tilde{\underline{\boldsymbol{\epsilon}}}k_y\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_x+k_xk_y\\ 0 & 0 & -\tilde{\underline{\boldsymbol{\epsilon}}}k_x\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_y+k_xk_y & -k_0^2\tilde{\underline{\boldsymbol{\epsilon}}}+k_y^2+\tilde{\underline{\boldsymbol{\epsilon}}}k_x\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_x\end{bmatrix}. \tag{4.12}$$

Substituting Eqs. (4.2-4.10) into the ODE (4.11), Eq. (4.11) can be rewritten as

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{\kappa},z)=-k_z^2[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{\kappa},z), \tag{4.13}$$

with $k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]=k_0^2\epsilon[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]-k_y^2-k_x^2$ and $\tilde{\boldsymbol{V}}_{\perp}=(\tilde{V}_1,\tilde{V}_2,\tilde{V}_4,\tilde{V}_5)^T$. The details of the mathematical derivation can be found in Appendix G.

If we consider the forward propagation, then the angular spectrum in plane $z+\Delta z$, with Δz approaching 0 ($\Delta z\to 0$), is

$$\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{\kappa},z+\Delta z)=\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{\kappa},z)\exp\{\mathrm{i}k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]\Delta z\}. \tag{4.14}$$

The related field in the x-domain is

$$\begin{aligned} \boldsymbol{V}_\perp(\boldsymbol{\rho}, z+\Delta z) &= \mathcal{F}_k^{\mathrm{h},-1}\left\{\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa}, z+\Delta z)\right\} \\ &= \mathcal{F}_k^{\mathrm{h},-1}\left\{\tilde{\boldsymbol{A}}_\perp(\boldsymbol{\kappa}, z)\exp\{\mathrm{i}\tilde{\psi}(\boldsymbol{\kappa}, z)+\mathrm{i}k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z]\Delta z\}\right\} \end{aligned}. \tag{4.15}$$

For further derivation, we need to formulate the mapping relation in plane $z+\Delta z$ as

$$\begin{aligned} \tilde{\boldsymbol{\nabla}}_\perp\{\tilde{\psi}(\boldsymbol{\kappa}, z)+k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z]\Delta z\} &= \tilde{\boldsymbol{\nabla}}_\perp\tilde{\psi}(\boldsymbol{\kappa}, z)+\Delta z\tilde{\boldsymbol{\nabla}}_\perp k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z] \\ &= -\boldsymbol{\rho}+\Delta z\left\{\frac{k_0^2}{2k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z]}\tilde{\boldsymbol{\nabla}}_\perp\epsilon[\boldsymbol{\rho}(\boldsymbol{\kappa})]-\frac{\boldsymbol{\kappa}(\boldsymbol{\rho})}{k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z]}\right\}. \\ &\overset{\Delta z\to 0}{\approx} -\boldsymbol{\rho} \quad \{...\} \text{ has limited magnitude} \end{aligned} \tag{4.16}$$

With the unchanged mapping relation $\boldsymbol{\rho}(\boldsymbol{\kappa})$, Eq. (4.15) is continued as

$$\begin{aligned} \boldsymbol{V}_\perp(\boldsymbol{\rho}, z+\Delta z) &= \boldsymbol{V}_\perp(\boldsymbol{\rho}, z)\exp\{\mathrm{i}k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z]\Delta z\} \\ &= \boldsymbol{U}_\perp(\boldsymbol{\rho}, z)\exp\{\mathrm{i}\psi(\boldsymbol{\rho}, z)+\mathrm{i}k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z]\Delta z\} \end{aligned}. \tag{4.17}$$

Comparing $\boldsymbol{V}_\perp(\boldsymbol{\rho}, z+\Delta z)$ and $\boldsymbol{V}_\perp(\boldsymbol{\rho}, z)$, the z derivative of the phase is

$$\begin{aligned} \frac{\partial}{\partial z}\psi(\boldsymbol{\rho}, z) &= \lim_{\Delta z\to 0}\frac{\psi(\boldsymbol{\rho}, z+\Delta z)-\psi(\boldsymbol{\rho}, z)}{\Delta z} \\ &= k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z] \end{aligned}. \tag{4.18}$$

Now we can conclude

$$\boldsymbol{\nabla}\psi(\boldsymbol{r}) = \boldsymbol{k}(\boldsymbol{r}), \tag{4.19}$$

with $\boldsymbol{k}(\boldsymbol{r}) = (k_x, k_y, k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}), z])$. Eq. (4.19) is the vectorial version of the eikonal equation, which is the most fundamental equation in geometric optics. From this equation, we can also determine that one important property of the geometric field is that the coordinate mapping relation $\boldsymbol{r}\to\boldsymbol{k}$ is one-to-one.

The eikonal equation, Eq. (4.19), describes the phase change. Furthermore, the behavior of $\boldsymbol{U}(\boldsymbol{\rho}, z)$, or $\tilde{\boldsymbol{A}}(\boldsymbol{\kappa}, z)$ can be derived by substituting the geometric field ansatz, and the eikonal equation, Eq. (4.19) into Maxwell's equations (2.23). To simplify the notation, from now on, we replace $\boldsymbol{V}$ by $\boldsymbol{E}$ and $\boldsymbol{H}$,

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}_0(\boldsymbol{r})\exp[i\psi(\boldsymbol{r})], \tag{4.20}$$

and

$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}_0(\boldsymbol{r})\exp[i\psi(\boldsymbol{r})]. \tag{4.21}$$

Maxwell's equations (2.23) can be rewritten as

$$\nabla\times\boldsymbol{E}_0(\boldsymbol{r})+i\boldsymbol{k}(\boldsymbol{r})\times\boldsymbol{E}_0(\boldsymbol{r})-i\omega\mu_0\boldsymbol{H}_0(\boldsymbol{r}) = 0, \tag{4.22}$$

$$\nabla\times\boldsymbol{H}_0(\boldsymbol{r})+i\boldsymbol{k}(\boldsymbol{r})\times\boldsymbol{H}_0(\boldsymbol{r})+i\omega\epsilon_0\epsilon_r(\boldsymbol{r})\boldsymbol{E}_0(\boldsymbol{r}) = 0, \tag{4.23}$$

$$\nabla\cdot\boldsymbol{E}_0(\boldsymbol{r})+i\boldsymbol{k}(\boldsymbol{r})\cdot\boldsymbol{E}_0(\boldsymbol{r})+\boldsymbol{E}_0(\boldsymbol{r})\cdot\nabla\ln\epsilon(\boldsymbol{r}) = 0, \tag{4.24}$$

$$\nabla \cdot \boldsymbol{H}_0(\boldsymbol{r}) + i\boldsymbol{k}(\boldsymbol{r}) \cdot \boldsymbol{H}_0(\boldsymbol{r}) = 0. \tag{4.25}$$

For spatially slowly varying $\boldsymbol{E}_0(\boldsymbol{r})$, $\boldsymbol{H}_0(\boldsymbol{r})$ and relative permittivity $\epsilon(\boldsymbol{r})$, the terms $\nabla \cdot \boldsymbol{E}_0(\boldsymbol{r})$, $\nabla \cdot \boldsymbol{H}_0(\boldsymbol{r})$, $\nabla \times \boldsymbol{E}_0(\boldsymbol{r})$, $\nabla \times \boldsymbol{H}_0(\boldsymbol{r})$, and $\nabla\epsilon(\boldsymbol{r})$ can be neglected. Then, we obtain Maxwell's equations for geometric fields,

$$\boldsymbol{k}(\boldsymbol{r}) \times \boldsymbol{E}_0(\boldsymbol{r}) = \omega\mu_0\boldsymbol{H}_0(\boldsymbol{r}), \tag{4.26}$$

$$\boldsymbol{k}(\boldsymbol{r}) \times \boldsymbol{H}_0(\boldsymbol{r}) = -\omega\epsilon_0\epsilon(\boldsymbol{r})\boldsymbol{E}_0(\boldsymbol{r}), \tag{4.27}$$

$$\boldsymbol{k}(\boldsymbol{r}) \cdot \boldsymbol{E}_0(\boldsymbol{r}) = 0, \tag{4.28}$$

$$\boldsymbol{k}(\boldsymbol{r}) \cdot \boldsymbol{H}_0(\boldsymbol{r}) = 0. \tag{4.29}$$

Eqs. (4.26-4.29) are called **geometric field equations**. These equations can also be found in Chapter 3 in the book [53], and we show it in Appendix F.

Here we summarize that the geometric fields are a construction of slowly varying complex amplitude $\boldsymbol{E}_0(\boldsymbol{r})$ and $\boldsymbol{H}_0(\boldsymbol{r})$, and a rapidly varying phase $\psi(\boldsymbol{r})$ (phase-dominate), as shown in Eqs. (4.20) and (4.21). The eikonal equation, Eq. (4.19), describes the behavior of $\psi(\boldsymbol{r})$, while the geometric field equations, Eqs. (4.26-4.29), describe that of $\boldsymbol{E}_0(\boldsymbol{r})$ and $\boldsymbol{H}_0(\boldsymbol{r})$. Although $\epsilon(r)$ variance need to be small, but it does not effect the including of polarization crosstalk, and we will show it later.

4.1.2 *ODEs in the x-Domain and Energy Conservation*

In order to solve Eqs. (4.19) and (4.26-4.29) for GRIN media, we first compare these equations with those for plane waves, which are derived using a plane wave ansatz in Maxwell's equations (2.20-2.23). The plane wave ansatzes are

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}_0 \exp(i\boldsymbol{k} \cdot \boldsymbol{r}), \tag{4.30}$$

and

$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}_0 \exp(i\boldsymbol{k} \cdot \boldsymbol{r}), \tag{4.31}$$

with a constant wave vector $\boldsymbol{k}$, and complex amplitude in the electric field $\boldsymbol{E}_0$ and in the magnetic field $\boldsymbol{H}_0$.

Plugging the plane wave ansatz, (4.30) and (4.31) into Maxwell's equations (2.23), we can derive the plane wave equations for homogeneous media with ϵ,

$$\boldsymbol{k} \times \boldsymbol{E}_0 = \omega\mu_0\boldsymbol{H}_0, \tag{4.32}$$

$$\boldsymbol{k} \times \boldsymbol{H}_0 = -\omega\epsilon_0\epsilon\boldsymbol{E}_0, \tag{4.33}$$

$$\boldsymbol{k} \cdot \boldsymbol{E}_0 = 0, \tag{4.34}$$

$$\boldsymbol{k} \cdot \boldsymbol{H}_0 = 0. \tag{4.35}$$

Comparing Eqs. (4.26-4.29) to Eqs. (4.32-4.35), we can see that the geometric field shows plane wave behavior locally in the x-domain, with wave vector $\boldsymbol{k}(\boldsymbol{r})$. Thus, we

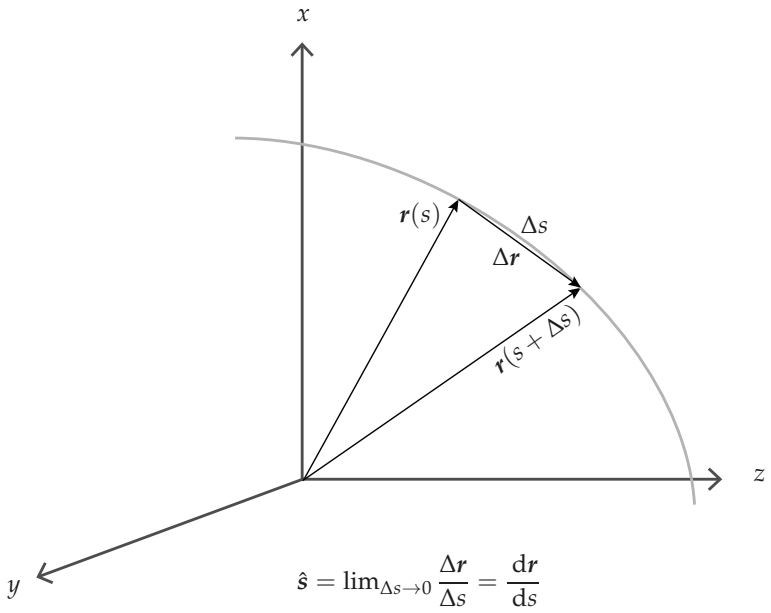


Figure 9: Definition of the tangential vector $\hat{\boldsymbol{s}}$ of a ray curve.

name the local field that shows the plane wave behavior as the **local plane wave**. Therefore, the geometric field is a composition of local plane waves.

Replacing $\boldsymbol{k}$ by $\boldsymbol{k}(\boldsymbol{r})$ in Eqs. (4.30) and (4.31), we obtain the mathematical representation of the geometric fields,

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}_0(\boldsymbol{r}) \exp\left\{ i \left[\psi(\boldsymbol{r}_0) + \int_{\boldsymbol{r}_0}^{\boldsymbol{r}} \boldsymbol{k}(\boldsymbol{r}) \cdot d\boldsymbol{r} \right] \right\}, \tag{4.36}$$

and

$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}_0(\boldsymbol{r}) \exp\left\{ i \left[\psi(\boldsymbol{r}_0) + \int_{\boldsymbol{r}_0}^{\boldsymbol{r}} \boldsymbol{k}(\boldsymbol{r}) \cdot d\boldsymbol{r} \right] \right\}. \tag{4.37}$$

Comparing Eqs. (4.36) and (4.37) with those for plane waves, (4.30) and (4.31), we see that in the phase of the geometric fields, the plane wave phase $\boldsymbol{k} \cdot \boldsymbol{r}$ is replaced by $\int \boldsymbol{k}(\boldsymbol{r}) \cdot d\boldsymbol{r}$, a path integral that integrates over the position. In a GRIN medium, $\boldsymbol{k}(\boldsymbol{r})$ of each local plane wave changes smoothly during propagation, and we can define **the path of one local plane wave as a ray**. Now we give several mathematical terms/properties related to a parametric curve (ray) [73]: (1) **arc length** s is the distance between two points along the curve; (2) the smooth curve can be parametrized with respect to the arc length, i.e., $\boldsymbol{r}(s)$; (3) the tangential vector of the curve $\hat{\boldsymbol{s}}(\boldsymbol{r}) := \dfrac{d\boldsymbol{r}}{ds}$, as shown in Fig. 9.

Furthermore, we define

$$\boldsymbol{k}(\boldsymbol{r}) = k_0 \sqrt{\epsilon(\boldsymbol{r})} \hat{\boldsymbol{s}}(\boldsymbol{r}), \tag{4.38}$$

where $\hat{\boldsymbol{s}}(\boldsymbol{r})$ denotes the propagation direction (normalized 3D vector) of the local plane wave at position $\boldsymbol{r}$.

Let us now consider the geometric field representation in the context of a modeling task: which parameters are known, and which are to be calculated (Fig. 10). Typically, we assume to be known: (1) the electric field distribution in the input plane, i.e.,

$$\boldsymbol{E}(\boldsymbol{r}^{\text{in}}) = \boldsymbol{E}_0(\boldsymbol{r}^{\text{in}}) \exp[i\psi(\boldsymbol{r}^{\text{in}})], \tag{4.39}$$

where $\boldsymbol{r}^{\text{in}}$ denotes the position in the input plane. Both $\boldsymbol{E}_0(\boldsymbol{r}^{\text{in}})$ and $\psi(\boldsymbol{r}^{\text{in}})$ are known in the input plane; (2) the wave vector in the input plane through Eq. (4.3),

$$\boldsymbol{\kappa}(\boldsymbol{r}^{\text{in}}) = \nabla_{\perp}\psi(\boldsymbol{r}^{\text{in}}), \tag{4.40}$$

with $\boldsymbol{\kappa} = (k_x, k_y)$ and $\nabla_{\perp} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$, and

$$k_z(\boldsymbol{r}^{\text{in}}) = \sqrt{k_0^2\epsilon(\boldsymbol{r}^{\text{in}}) - ||\boldsymbol{\kappa}(\boldsymbol{r}^{\text{in}})||^2}. \tag{4.41}$$

Then, $\hat{\boldsymbol{s}}(\boldsymbol{r}^{\text{in}}) = \boldsymbol{k}(\boldsymbol{r}^{\text{in}})/[k_0\sqrt{\epsilon(\boldsymbol{r}^{\text{in}})}]$. In the following, we no longer mention the magnetic field with the electric field because it can always be calculated using Eqs. (4.20-4.21) and Eq. (4.26), i.e.,

$$\boldsymbol{H}(\boldsymbol{r}) = \frac{1}{\omega\mu_0}\boldsymbol{k}(\boldsymbol{r}) \times \boldsymbol{E}(\boldsymbol{r}). \tag{4.42}$$

Here we emphasis that if there is an abrupt change in $\epsilon(\boldsymbol{r})$ before and after the input plane, the local plane interface approximation (LPIA) [29, 30], which calculates the geometric field propagating through an optical surface, can be used to calculate $\boldsymbol{E}(\boldsymbol{r}^{\text{in}})$ inside the GRIN media on the input plane.

In a modeling task, we often seek to calculate the field distribution in the output plane, i.e.,

$$\boldsymbol{E}(\boldsymbol{r}^{\text{out}}) = \boldsymbol{E}_0(\boldsymbol{r}^{\text{out}}) \exp[i\psi(\boldsymbol{r}^{\text{out}})]. \tag{4.43}$$

We start with calculating $\psi(\boldsymbol{r}^{\text{out}})$. From Eq. (4.36), we know

$$\psi(\boldsymbol{r}^{\text{out}}) = \psi(\boldsymbol{r}^{\text{in}}) + \int_{\boldsymbol{r}^{\text{in}}}^{\boldsymbol{r}^{\text{out}}} \boldsymbol{k}(\boldsymbol{r}) \cdot d\boldsymbol{r}. \tag{4.44}$$

To compute $\int_{\boldsymbol{r}^{\text{in}}}^{\boldsymbol{r}^{\text{out}}} \boldsymbol{k}(\boldsymbol{r}) \cdot d\boldsymbol{r}$, we need to determine the integration path from $\boldsymbol{r}^{\text{in}}$ to $\boldsymbol{r}^{\text{out}}$, which is the ray path.

To calculate the ray path, we differentiate Eq. (4.19) with respect to arc length s [53] to obtain the first ODE in the x-domain (Appendix G),

$$\frac{\mathrm{d}}{\mathrm{d}s}[\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})] = \boldsymbol{\nabla}\sqrt{\epsilon(\boldsymbol{r})}. \tag{4.45}$$

It describes how $\hat{\boldsymbol{s}}(\boldsymbol{r})$ varies along the ray path. From this equation, we also know that when the initial position $\boldsymbol{r}^{\text{in}}$ and direction $\hat{\boldsymbol{s}}(\boldsymbol{r}^{\text{in}})$ are known, and the relative permittivity $\epsilon(\boldsymbol{r})$ is fixed in space, the path is unique. This is a well-known ray equation for a GRIN medium [53]. Meanwhile, the unique ray path introduces a one-to-one mapping between the positions in the input plane and those in the output plane $\boldsymbol{r}^{\text{out}}(\boldsymbol{r}^{\text{in}})$. By solving Eq. (4.45), we obtain the ray path and the mapping relation $\boldsymbol{r}^{\text{out}}(\boldsymbol{r}^{\text{in}})$, as shown

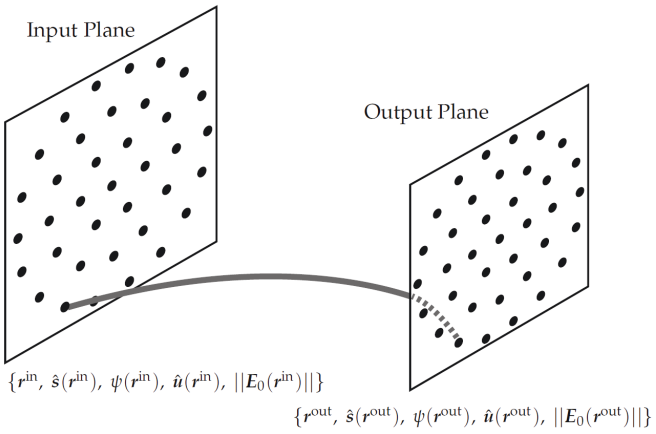


Figure 10: Illustration of a ray path from the input plane to the output plane of a GRIN medium. The dots in the input plane denote the starting positions of the ray path, while those in the output plane denote the target positions. The dots in the input and output plane are mapped one-to-one. The amplitude-related parameters $||\boldsymbol{E}_0(\boldsymbol{r})||$ and $\hat{\boldsymbol{u}}(\boldsymbol{r})$ are calculated along the ray path.

in Fig. 10. The detailed algorithm for solving Eq. (4.45) will be described in the next section.

After the ray path has been calculated, $\psi(\boldsymbol{r}^{\text{out}})$ is computed by plugging Eq. (4.38) into Eq. (4.44),

$$\begin{aligned}\psi(\boldsymbol{r}^{\text{out}}) &= \psi(\boldsymbol{r}^{\text{in}}) + \int_{\boldsymbol{r}^{\text{in}}}^{\boldsymbol{r}^{\text{out}}} k_0\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot \mathrm{d}\boldsymbol{r} \\ &= \psi(\boldsymbol{r}^{\text{in}}) + k_0\int_{\boldsymbol{r}^{\text{in}}}^{\boldsymbol{r}^{\text{out}}}\sqrt{\epsilon(\boldsymbol{r})}\,\mathrm{d}s.\end{aligned} \tag{4.46}$$

With this, we know $\psi(\boldsymbol{r}^{\text{out}})$, and all the information about the ray path. Next, we need to calculate the amplitude $\boldsymbol{E}_0(\boldsymbol{r}^{\text{out}})$ in Eq. (4.43), which can be expressed as a product,

$$\boldsymbol{E}_0(\boldsymbol{r}^{\text{out}}) = ||\boldsymbol{E}_0(\boldsymbol{r}^{\text{out}})||\hat{\boldsymbol{u}}(\boldsymbol{r}^{\text{out}}), \tag{4.47}$$

with $\hat{\boldsymbol{u}}(\boldsymbol{r}^{\text{out}})$ describing the normalized field vector, which is also the polarization vector, and $||\boldsymbol{E}_0(\boldsymbol{r}^{\text{out}})||$ describing the magnitude of the electric field. Next, we explore $\hat{\boldsymbol{u}}(\boldsymbol{r}^{\text{out}})$ and $||\boldsymbol{E}_0(\boldsymbol{r}^{\text{out}})||$ step by step.

Differentiate Eq. (4.34) with respect to arc length s to obtain the second ODE in the x-domain (Appendix G),

$$\sqrt{\epsilon(\boldsymbol{r})}\frac{\mathrm{d}\hat{\boldsymbol{u}}(\boldsymbol{r})}{\mathrm{d}s} = -\left[\hat{\boldsymbol{u}}(\boldsymbol{r})\cdot\boldsymbol{\nabla}\sqrt{\epsilon(\boldsymbol{r})}\right]\hat{\boldsymbol{s}}(\boldsymbol{r})\ . \tag{4.48}$$

This equation describes the variation in $\hat{\boldsymbol{u}}(\boldsymbol{r})$ along the ray path. As shown in Fig. 10, the initial information for field $\hat{\boldsymbol{u}}(\boldsymbol{r}^{\text{in}})$ is known; thus, one can solve $\hat{\boldsymbol{u}}(\boldsymbol{r})$ along the ray path until $\hat{\boldsymbol{u}}(\boldsymbol{r}^{\text{out}})$ in the output plane.

To calculate $||\boldsymbol{E}_0(\boldsymbol{r}^{\text{out}})||$, the energy conservation law is used in the case where the medium is free of gain and absorption, which is mathematically represented as,

$$\nabla\cdot <\boldsymbol{S}(\boldsymbol{r})>= 0. \tag{4.49}$$

where $< \boldsymbol{S}(\boldsymbol{r}) >$ denotes the time-averaged Poynting vector.

Using Gauss's theorem, Eq. (4.49) can be rewritten as

$$\oint_F < \boldsymbol{S}(\boldsymbol{r}) > \cdot \hat{\boldsymbol{a}}(\boldsymbol{r}) \, \mathrm{d}f = 0, \tag{4.50}$$

with $\hat{\boldsymbol{a}}(\boldsymbol{r})$ denoting the normal vector to the surface at position $\boldsymbol{r}$. This equation denotes a surface integral over a closed interface F.

Because the energy of the geometric field flows only through the input and output planes, while no energy is lost through the (figurative) lateral boundaries because there is no diffraction, Eq. (4.50) may be rewritten as

$$\int_{F^{\mathrm{in}}} < \boldsymbol{S}(\boldsymbol{r}^{\mathrm{in}}) > \cdot \hat{\boldsymbol{a}}(\boldsymbol{r}^{\mathrm{in}}) \, \mathrm{d}f = \int_{F^{\mathrm{out}}} < \boldsymbol{S}(\boldsymbol{r}^{\mathrm{out}}) > \cdot \hat{\boldsymbol{a}}(\boldsymbol{r}^{\mathrm{out}}) \, \mathrm{d}f. \tag{4.51}$$

For the propagation of geometric fields, the mapping between $\boldsymbol{r}^{\mathrm{in}}$ and $\boldsymbol{r}^{\mathrm{out}}$ is one-to-one. Therefore, Eq. (4.51) should be true for each path. The integral can then be removed, and Eq. (4.51) becomes

$$< \boldsymbol{S}(\boldsymbol{r}^{\mathrm{in}}) > \cdot \hat{\boldsymbol{a}}(\boldsymbol{r}^{\mathrm{in}}) \, \mathrm{d}f(\boldsymbol{r}^{\mathrm{in}}) =< \boldsymbol{S}(\boldsymbol{r}^{\mathrm{out}}) > \cdot \hat{\boldsymbol{a}}(\boldsymbol{r}^{\mathrm{out}}) \, \mathrm{d}f(\boldsymbol{r}^{\mathrm{out}}), \tag{4.52}$$

where $\mathrm{d}f(\boldsymbol{r})$ denotes the differential element of the area at position $\boldsymbol{r}$. In [53], this equation is called "intensity law of geometric optics."

The time-averaged Poynting vector is defined as

$$< \boldsymbol{S}(\boldsymbol{r}) >= \frac{1}{2} \Re[\boldsymbol{E}(\boldsymbol{r}) \times \boldsymbol{H}^*(\boldsymbol{r})]. \tag{4.53}$$

Now we calculate $< \boldsymbol{S}(\boldsymbol{r}) >$ by plugging the field representations (4.20) and (4.21), as well as the geometric field equation (4.32) into Eq. (4.53),

$$< \boldsymbol{S}(\boldsymbol{r}) >= \sqrt{\epsilon(\boldsymbol{r})} \sqrt{\frac{\epsilon_0}{\mu_0}} ||\boldsymbol{E}_0(\boldsymbol{r})||^2 \hat{\mathbf{s}}(\boldsymbol{r}). \tag{4.54}$$

By plugging Eq. (4.54) into Eq. (4.52), $||\boldsymbol{E}(\boldsymbol{r}^{\mathrm{out}})||$ can be calculated [24],

$$||\boldsymbol{E}(\boldsymbol{r}^{\mathrm{out}})|| = ||\boldsymbol{E}(\boldsymbol{r}^{\mathrm{in}})|| \sqrt{\frac{\hat{\mathbf{s}}(\boldsymbol{r}^{\mathrm{in}}) \cdot \hat{\boldsymbol{a}}(\boldsymbol{r}^{\mathrm{in}}) \, \mathrm{d}f(\boldsymbol{r}^{\mathrm{in}})}{\hat{\mathbf{s}}(\boldsymbol{r}^{\mathrm{out}}) \cdot \hat{\boldsymbol{a}}(\boldsymbol{r}^{\mathrm{out}}) \, \mathrm{d}f(\boldsymbol{r}^{\mathrm{out}})}} \left[\frac{\epsilon(\boldsymbol{r}^{\mathrm{in}})}{\epsilon(\boldsymbol{r}^{\mathrm{out}})} \right]^{\frac{1}{4}}. \tag{4.55}$$

With this, all the information for the geometric field in the output plane has been calculated,

$$\boldsymbol{E}(\boldsymbol{r}^{\mathrm{out}}) = ||\boldsymbol{E}_0(\boldsymbol{r}^{\mathrm{out}})|| \hat{\boldsymbol{u}}(\boldsymbol{r}^{\mathrm{out}}) \exp[i\psi(\boldsymbol{r}^{\mathrm{out}})]. \tag{4.56}$$

Next, the magnetic field in the output plane is computed using Eqs. (4.32) and (4.56),

$$\boldsymbol{H}(\boldsymbol{r}^{\mathrm{out}}) = \sqrt{\frac{\epsilon_0 \epsilon(\boldsymbol{r}^{\mathrm{out}})}{\mu_0}} \hat{\mathbf{s}}(\boldsymbol{r}^{\mathrm{out}}) \times \boldsymbol{E}(\boldsymbol{r}^{\mathrm{out}}). \tag{4.57}$$

4.1.3 *Numerical Solver Development*

From the theory section, we can summarize the calculation procedure for geometric field propagation from the input plane to the output plane in a GRIN medium: the field in the output plane is represented as Eq. (4.56); to calculate $\psi(\boldsymbol{r}^{\text{out}})$, we need first to calculate the ray path by solving Eq. (4.45), and to calculate the phase function $\psi(\boldsymbol{r}^{\text{out}})$ by solving Eq. (4.46); then the amplitude calculation is separated into the calculation of $\hat{\boldsymbol{u}}(\boldsymbol{r})$ by solving Eq. (4.48), and the calculation of $||\boldsymbol{E}_0(\boldsymbol{r})||$ by Eq. (4.55).

To solve Eqs. (4.45) and (4.48), two additional parameters are defined to simplify these two equations [74],

$$\boldsymbol{D}(\boldsymbol{r}) = \boldsymbol{\nabla}\sqrt{\epsilon(\boldsymbol{r})}, \tag{4.58}$$

$$\boldsymbol{T}(\boldsymbol{r}) = \sqrt{\epsilon(\boldsymbol{r})}\frac{d\boldsymbol{r}}{ds} = \sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r}). \tag{4.59}$$

Both $\boldsymbol{D}(\boldsymbol{r})$ and $\boldsymbol{T}(\boldsymbol{r})$ are three-dimensional (3D) vectors. With the definition given in (4.58) and (4.59), Eqs. (4.45) and (4.48) can be rewritten as

$$\frac{d}{ds}\boldsymbol{T}(\boldsymbol{r}) = \boldsymbol{D}(\boldsymbol{r}), \tag{4.60}$$

and

$$\frac{d\hat{\boldsymbol{u}}(\boldsymbol{r})}{ds} = -[\hat{\boldsymbol{u}}(\boldsymbol{r})\cdot\boldsymbol{D}(\boldsymbol{r})]\frac{\boldsymbol{T}(\boldsymbol{r})}{\epsilon(\boldsymbol{r})}. \tag{4.61}$$

Both Eqs. (4.60) and (4.61) are first-order differential equations, so it is possible to use the Runge-Kutta (RK) method to solve them iteratively. As shown in Fig. 11, we first process a ray path calculation step from $\boldsymbol{r}_i$ to $\boldsymbol{r}'_i$ with step size $\frac{\Delta s}{2}$. The same process is performed from $\boldsymbol{r}'_i$ to $\boldsymbol{r}_{i+1}$. Then, we process a field calculation step from $\boldsymbol{r}_i$ to $\boldsymbol{r}_{i+1}$ with step size Δs.

Phase calculation

In this process, we calculate the ray path by solving the auxiliary quantities $\boldsymbol{D}(\boldsymbol{r})$ and $\boldsymbol{T}(\boldsymbol{r})$, which may be useful for the amplitude calculation process. Then, the phase $\psi(\boldsymbol{r})$ can be calculated by taking the integral in Eq. (4.46) along the ray path.

Starting from one point $(\boldsymbol{r}_i, \boldsymbol{D}(\boldsymbol{r}_i), \boldsymbol{T}(\boldsymbol{r}_i))$, one can calculate the next point $(\boldsymbol{r}'_i, \boldsymbol{D}(\boldsymbol{r}'_i), \boldsymbol{T}(\boldsymbol{r}'_i))$ using

$$\boldsymbol{T}(\boldsymbol{r}'_i) = \boldsymbol{T}(\boldsymbol{r}_i) + \frac{\Delta s}{12}\left[\boldsymbol{D}(\boldsymbol{r}_i) + 4\boldsymbol{D}(\boldsymbol{r}''_i) + \boldsymbol{D}(\boldsymbol{r}'_i)\right], \tag{4.62}$$

where $\boldsymbol{r}''_i$ and $\boldsymbol{r}'_i$ are calculated by Taylor expansion,

$$\boldsymbol{r}''_i = \boldsymbol{r}_i + \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}s}\frac{\Delta s}{4} + \frac{1}{2}\frac{\mathrm{d}^2\boldsymbol{r}_i}{\mathrm{d}s^2}\left(\frac{\Delta s}{4}\right)^2,$$

$$\boldsymbol{r}'_i = \boldsymbol{r}_i + \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}s}\frac{\Delta s}{2} + \frac{1}{2}\frac{\mathrm{d}^2\boldsymbol{r}_i}{\mathrm{d}s^2}\left(\frac{\Delta s}{2}\right)^2,$$

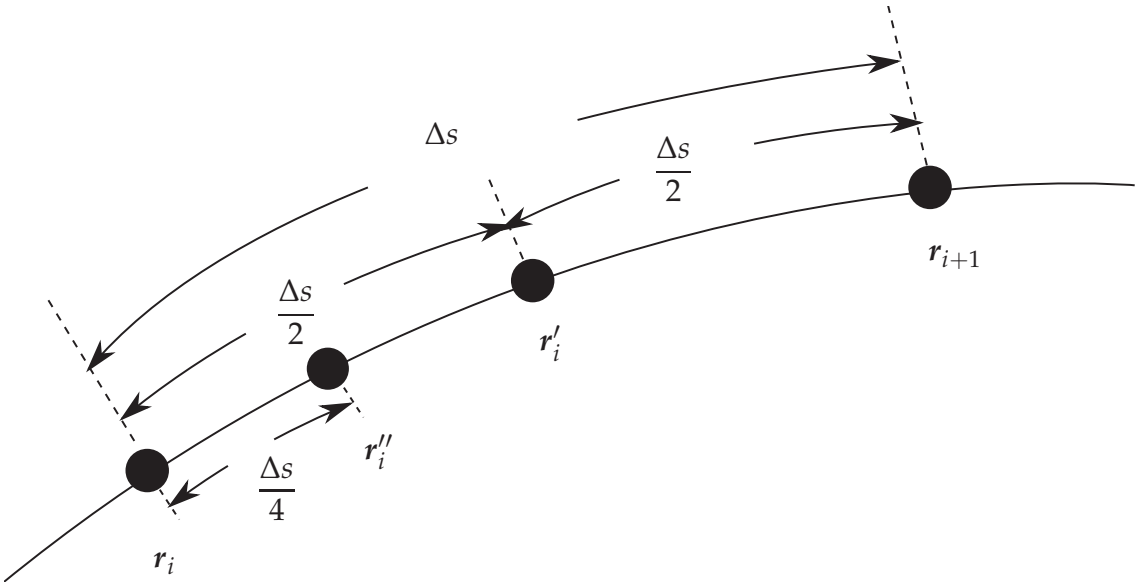


Figure 11: Notation of the parameters in the algorithm. $\boldsymbol{r}$ denotes the spatial position along the ray path. The step size of the phase calculation process is $\frac{\Delta s}{2}$, while that of the amplitude is Δs.

with $\dfrac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}s} = \dfrac{\boldsymbol{T}(\boldsymbol{r}_i)}{\sqrt{\epsilon(\boldsymbol{r}_i)}}$ and $\dfrac{\mathrm{d}^2\boldsymbol{r}_i}{\mathrm{d}s^2} = \epsilon^{2/3}(\boldsymbol{r}_i)\left\{\epsilon(\boldsymbol{r}_i)\boldsymbol{D}(\boldsymbol{r}_i) - [\boldsymbol{T}(\boldsymbol{r}_i)\cdot\boldsymbol{D}(\boldsymbol{r}_i)]\boldsymbol{T}(\boldsymbol{r}_i)\right\}$.

To calculate the point $(\boldsymbol{r}_{i+1}, \boldsymbol{D}(\boldsymbol{r}_{i+1}), \boldsymbol{T}(\boldsymbol{r}_{i+1}))$ from the point $(\boldsymbol{r}'_i, \boldsymbol{D}(\boldsymbol{r}'_i), \boldsymbol{T}(\boldsymbol{r}'_i))$, one can use Eq. (4.62) by simply replacing $\boldsymbol{r}'_i$ $(\boldsymbol{r}_i)$ with $\boldsymbol{r}_{i+1}$ $(\boldsymbol{r}'_i)$. Following Eq. (4.46), the phase $\psi(\boldsymbol{r})$ from $\boldsymbol{r}_i$ to $\boldsymbol{r}_{i+1}$ can be computed as

$$\psi(\boldsymbol{r}_{i+1}) = \psi(\boldsymbol{r}_i) + \frac{k_0\Delta s}{4}\left(\sqrt{\epsilon(\boldsymbol{r}_i)} + 2\sqrt{\epsilon(\boldsymbol{r}'_i)} + \sqrt{\epsilon(\boldsymbol{r}_{i+1})}\right). \tag{4.63}$$

Using this iterative calculation procedure starting with $\boldsymbol{r}^{\mathrm{in}}$, one can ultimately calculate $\psi(\boldsymbol{r}^{\mathrm{out}})$.

Amplitude calculation

In this process, we first solve Eq. (4.61) to obtain the normalized field vector $\hat{\boldsymbol{u}}(\boldsymbol{r})$ using the RK method. Starting with $\hat{\boldsymbol{u}}(\boldsymbol{r}_i)$, one can calculate $\hat{\boldsymbol{u}}(\boldsymbol{r}_{i+1})$ using

$$
\begin{aligned}
\boldsymbol{\alpha}_1 &= -\{\hat{\boldsymbol{u}}(\boldsymbol{r}_i)\cdot\frac{\boldsymbol{D}(\boldsymbol{r}_i)}{\epsilon(\boldsymbol{r}_i)}\}\boldsymbol{T}(\boldsymbol{r}_i),\\
\boldsymbol{\alpha}_2 &= -\{[\hat{\boldsymbol{u}}(\boldsymbol{r}_i)+\tfrac{\Delta s}{2}\boldsymbol{\alpha}_1]\cdot\frac{\boldsymbol{D}(\boldsymbol{r}'_i)}{\epsilon(\boldsymbol{r}'_i)}\}\boldsymbol{T}(\boldsymbol{r}'_i),\\
\boldsymbol{\alpha}_3 &= -\{[\hat{\boldsymbol{u}}(\boldsymbol{r}_i)+\tfrac{\Delta s}{2}\boldsymbol{\alpha}_2]\cdot\frac{\boldsymbol{D}(\boldsymbol{r}'_i)}{\epsilon(\boldsymbol{r}'_i)}\}\boldsymbol{T}(\boldsymbol{r}'_i),\\
\boldsymbol{\alpha}_4 &= -\{[\hat{\boldsymbol{u}}(\boldsymbol{r}_i)+\Delta s\boldsymbol{\alpha}_3]\cdot\frac{\boldsymbol{D}(\boldsymbol{r}_{i+1})}{\epsilon(\boldsymbol{r}_{i+1})}\}\boldsymbol{T}(\boldsymbol{r}_{i+1}),\\
\hat{\boldsymbol{u}}(\boldsymbol{r}_{i+1}) &= \hat{\boldsymbol{u}}(\boldsymbol{r}_i)+\tfrac{\Delta s}{6}(\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4).
\end{aligned}
\tag{4.64}
$$

Note that $\boldsymbol{D}(\boldsymbol{r}_i)$, $\boldsymbol{D}(\boldsymbol{r}'_i)$, $\boldsymbol{T}(\boldsymbol{r}_i)$, and $\boldsymbol{T}(\boldsymbol{r}'_i)$ are still unknown parameters if we solve Eq. (4.61) individually. However, they are all phase-related parameters, which were calculated in the phase calculation process. Therefore, Eqs. (4.60) and (4.61) are coupled to each other and must be solved simultaneously.

Then, we calculate $||\boldsymbol{E}_0(\boldsymbol{r})||$ from the initial condition $||\boldsymbol{E}_0(\boldsymbol{r}^{\text{in}})||$. In Eq. (4.55), all parameters are known or calculated except $\mathrm{d}f(\boldsymbol{r}^{\text{in}})/\mathrm{d}f(\boldsymbol{r}^{\text{out}})$. However, because we already know the one-to-one mapping relation $\boldsymbol{r}^{\text{out}}(\boldsymbol{r}^{\text{in}})$, $\mathrm{d}f(\boldsymbol{r}^{\text{in}})/\mathrm{d}f(\boldsymbol{r}^{\text{out}})$ can be calculated using a triangulation algorithm, which is already implemented in field tracing in the physical optics software VirtualLab Fusion [26].

4.1.4 *Validation: Multimode Fiber*

Here we are modeling a y-polarized fundamental Gaussian mode (i.e., $E_x(x,y)=0$), with a beam waist radius of 5 µm, and at a wavelength of 532 nm propagating in a GRIN optical fiber, as shown in Fig. 12 (a). The spatial distribution of the relative permittivity $\epsilon(x,y)$ is presented in Fig. 12 (b). The input plane is located at a distance of 160 µm after the position of the beam waist to ensure the field is in its geometric zone. We use both the RKX-FPM and the Fourier modal method with perfectly matched layers (FMM+PMLs), to calculate the field distribution in the test plane after the field has propagated a distance of $d=100$ µm from the input plane.

In Fig. 13, an illustration of rays propagating in the fiber, as well as the dot diagram in the test plane is shown. In Fig. 14, we show the plots for the amplitudes of the field components E_x, E_y, and E_z in the test plane, which are computed using the RKX-FPM (upper row) and FMM+PMLs (bottom row). Note that the methods give similar results. The relative deviation between the electric field calculated by the two approaches is 1% (calculated using Eq. (3.37)), demonstrating the practical strength of the RKX-FPM. Incidentally, the input Gaussian field is y-polarized, but the field in the test plane contains a relatively small E_x component. This is due to the crosstalk between the different field components, which is introduced by the field propagating through the interface as well as through the GRIN medium. The computation of field propagation through an

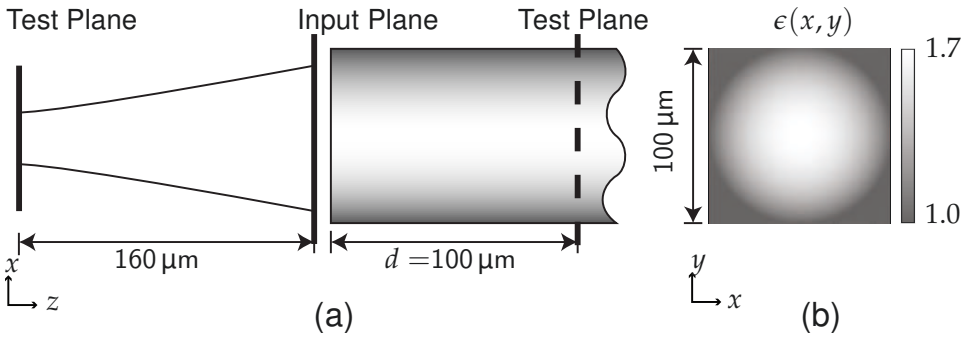


Figure 12: Illustration of the optical system. (a) The input plane is located at a distance of 160 µm after the beam waist (short, solid vertical line). The dashed line represents the test output plane located at a distance of $d = 100\,\mu m$ after the input plane (long solid vertical line). (b) The relative permittivity $\epsilon(x, y)$ inside the fiber, which has a core radius of 50 µm.

interface is carried out using the LPIA [29]. As the divergent angle of the input beam is small, the crosstalk effect after the first interface, the input plane, is extremely weak, with an amplitude of E_x lower than 42 μV/m. We have found that the modification of the polarization state of the input beam is mainly caused by the field propagating in the GRIN medium, with a propagation distance $d = 100\,\mu m$, which is adequately solved by using RKX-FPM.

Compared with the FMM+PMLs, our approach offers a significant advantage in terms of computational time. To calculate the results presented in Fig. 14, which were attained using FMM+PMLs on a workstation with an Intel Xeon CPU E5-2643 v2 @3.5 GHz and a total 128 GB RAM, the simulation ran for over 40 hours, whereas by exploiting the RKX-FPM put forth in this section, several seconds are required on a personal computer with an Intel Core i7-7700HQ CPU @2.80 GHz and a total 32 GB RAM.

4.1.5 *Validation: Ball Lens*

Here we are modeling an x-polarized fundamental Gaussian mode (i.e., $E_y(x, y) = 0$) with a Rayleigh length of 5 µm and at a wavelength of 532 nm, propagating in a GRIN ball lens, as shown in Fig. 15 (a). The spatial distribution of the relative permittivity $\epsilon(\boldsymbol{r})$ is represented by the color map, which shows no abrupt change between the surrounding medium and the GRIN sphere. We use both the RKK-FPM and RKX-FPM to calculate the field distribution in the test plane after the field propagates through the ball lens. The accuracy of the RKK-FPM has been proved in Chapter 3, so we use it here to offer a reference result.

In Fig. 15 (b), we show the ray tracing result from the input plane to the test plane, which is computed using the phase calculation process of the RKX-FPM. Inside the ball

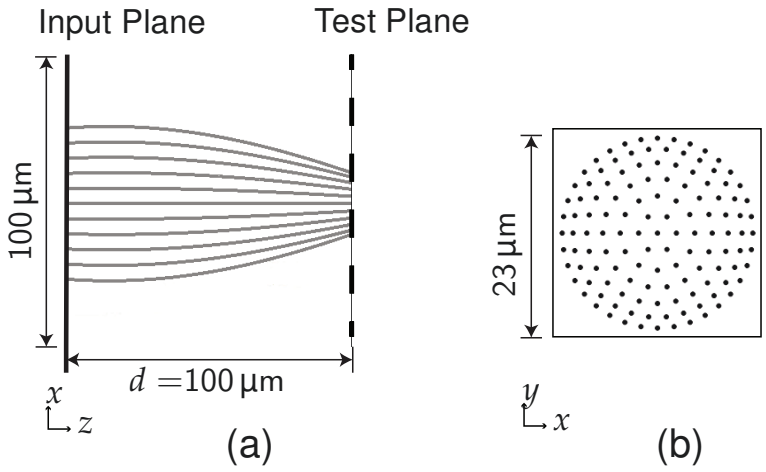


Figure 13: Results calculated by the phase calculation process of our approach. (a) Illustration of the rays propagating in the fiber. (b) Dot diagram in the test plane.

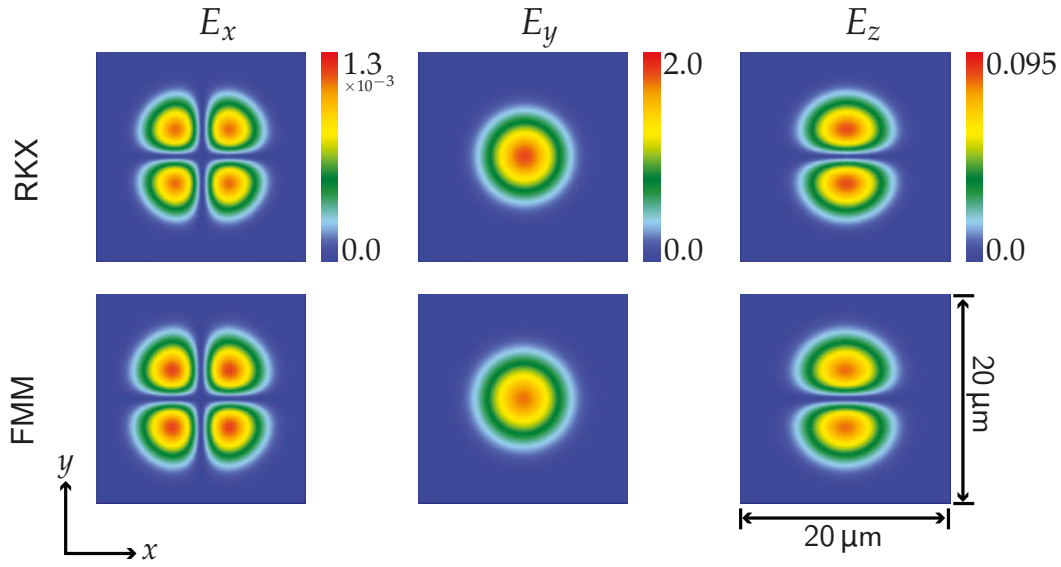


Figure 14: Distributions of the field components E_x, E_y, and E_z in the test plane calculated by the RKX-FPM (top row) and FMM (bottom row).

lens, the Gaussian beam propagates first divergently and then convergently. In Fig. 15 (c), we show the plots for the amplitudes of the field components E_x, E_y, and E_z in the test plane, which are computed by both the RKK-FPM (left column) and the RKX-FPM (right column). Both methods give similar results. The relative deviation between the electric fields, calculated using Eq. (3.37), is less than 1 %, demonstrating the practical strength of the RKX-FPM.

Compared with the RKK-FPM, the RKX-FPM has a further advantage in terms of computational time. To calculate the result fields in the test plane on a personal computer

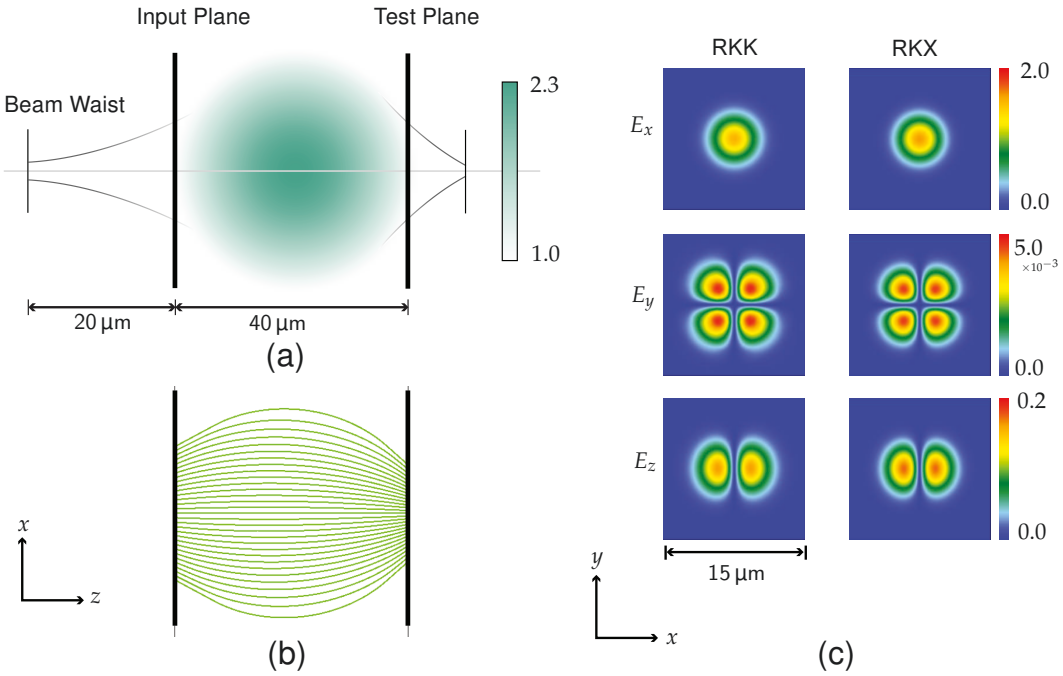


Figure 15: Modeling of a GRIN ball lens. (a)The Gaussian beam waist is located 20 µm before the input plane. The color represents the $\epsilon(\boldsymbol{r})$ of the ball lens. (b) The ray tracing between the input and test plane. (c) Amplitude distribution of field components E_x, E_y, and E_z in the test plane, calculated by the RKK-FPM (left) and the RKX-FPM (right).

with an Intel Core i7-7700HQ CPU @2.80 GHz and a total 32GB RAM, the RKK-FPM runs for about 4 min, while the RKX-FPM requires only a few seconds.

4.2 EXTENSION TO INCLUDE DIFFRACTION

The RKX-FPM is valid for simulation of a geometric field in GRIN media, and it is sufficient for many applications that are relevant to the GRIN imaging lens. The question is when the diffractive zones/focal regions are inside the GRIN media, and whether we can extend the RKX-FPM to calculate this diffractive field.

Combining Both RK Methods

The most straightforward way is to combine the RKK-FPM and RKX-FPM. More specifically, the geometric field propagation is simulated using the RKX-FPM, while the other parts are simulated by the RKK-FPM. We give an example in Fig. 16. In this example, the input plane wave is x-polarized, and has a wavelength of 532 nm and an aperture size of 100 µm with a smooth edge of 10 µm down to zero (in total 120 µm). The radius of

the Luneburg lens is 50 µm. We use both the Mie theory and the combined RK methods to calculate the focus field in the plane through the rear pole of the ball lens.

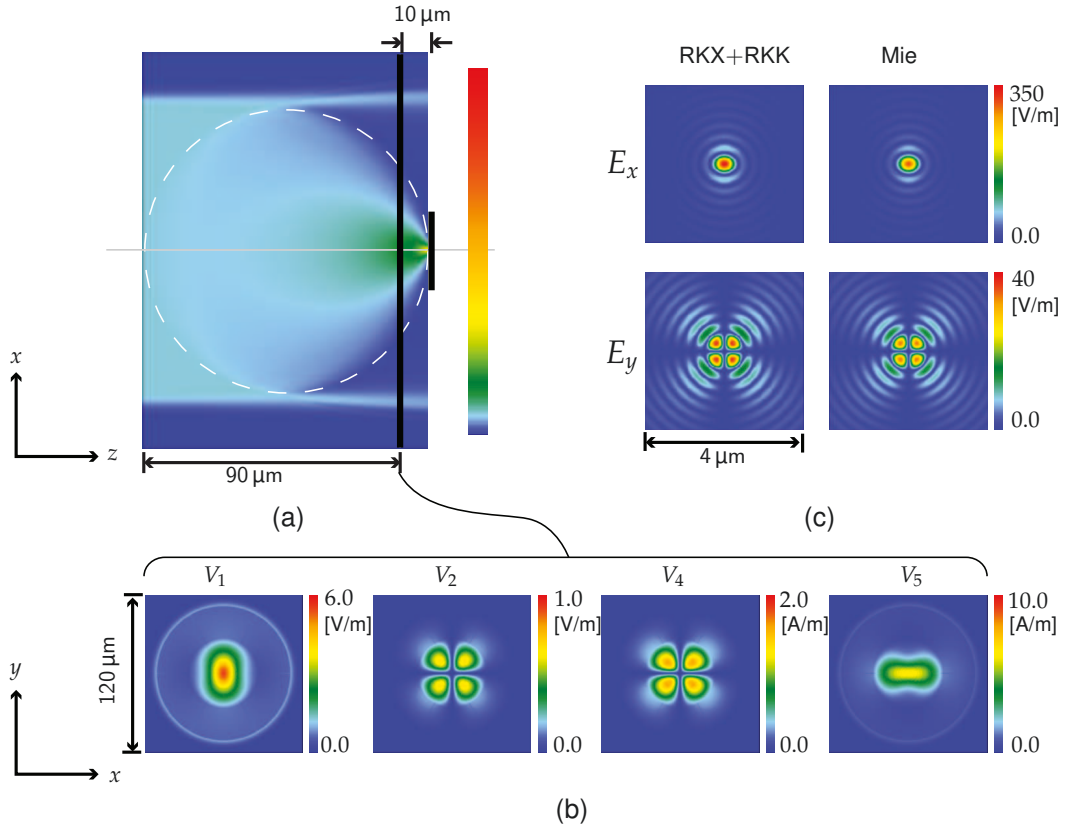


Figure 16: Simulation of a plane wave propagating through a Luneburg lens by combining both RK methods. (a) Field distribution given by simulating a y-invariant plane wave focused by a y-invariant Luneburg lens (radius is 50 µm); the color map is a logarithmic reverse rainbow. The switching plane is denoted by the long vertical solid line, while the focal plane is denoted by the short one. (b) The amplitude of the field components V_1, V_2, V_4, and V_5, is calculated by the RKX-FPM in the switching plane. (c) The amplitude of field components E_x and E_y, calculated by combining the RK methods(left column), and Mie theory(right column), in the focal plane.

In the calculation, we first use the RKX-FPM to simulate the input plane wave propagating a distance of 90 µm, and calculate E_x, E_y, $\sqrt{\frac{\epsilon_0}{\mu_0}}H_x$, and $\sqrt{\frac{\epsilon_0}{\mu_0}}H_y$ in the switching plane (long vertical black bar in Fig. 16 (a)). In Fig. 16 (b), the amplitude of four field components, which can also be denoted as $\boldsymbol{V}_\perp = \{V_1, V_2, V_4, V_5\}$, is shown. Then, we use the RKK-FPM to simulate the $\boldsymbol{V}_\perp$ propagating the remaining 10 µm to calculate the focus field, shown in the left column of Fig. 16 (c).

The amplitude of the focus field components E_x and E_y are shown in Fig. 16 (c), calculated by combining the RK methods(left column) and Mie theory (right column).

The strength of the resulting field calculated by combining the RK methods is somewhat higher than that of the Mie result. The reason is that a small part of the field escapes the ball lens because of the lens aperture diffraction, which is not predicted by the RKX-FPM. To compute the deviation between the two methods, we modify Eq. (3.37) to include a complex scaling factor s,

$$\sigma = \frac{\sum_{x,y} |sE_x(x,y) - E_x^{\text{ref}}|^2 + |sE_y(x,y) - E_y^{\text{ref}}|^2}{|E_x^{\text{ref}}|^2 + |E_y^{\text{ref}}|^2}, \tag{4.65}$$

with s as a constant to give the minimum σ. The deviation computed by Eq. (4.65) is 0.6 % with $s = 0.895$, demonstrating that the method is valid. The calculations were performed on a personal computer with an Intel Core i7-7700HQ CPU @2.80 GHz and a total 32 GB RAM. The combined RK methods runs for about 4 min in total, and most of the time is used by the RKK-FPM. We can imagine that the shorter the distance simulated by the RKK-FPM, the shorter the computing time required. To select the switching plane (long vertical black bar), we need to ensure that the field in the switching plane is geometric. The method we use to evaluate whether the field is geometric is by determining whether the bandwidth of $\tilde{E}_x$ in Eq. (4.20) is more than 50 times larger than that of $\tilde{E}_{x0}$.

Combining RKX-FPM with free space propagation

Another way to extend the RKX-FPM is found in the example in Section 4.1.4. Here we model the same y-polarized fundamental Gaussian propagating in a GRIN optical fiber for a longer distance of 1 mm, using both the RKX-FPM and the FMM+PMLs. The fields on planes $z = 0\,\mu\text{m}, 2\,\mu\text{m}, 4\,\mu\text{m}, \cdots, 1\,\text{mm}$ are recorded for two purposes: (1) to calculate the relative deviations $\sigma(z)$ using Eq. (4.65); (2) to plot the amplitude of field components E_x and E_y.

As $E_x = 0$ on the x- and y-axes in Fig. 17, $E_x = 0$, we extract the fields on the diagonal from $(x, y) = (-30\,\mu\text{m}, -30\,\mu\text{m})$ to $(30\,\mu\text{m}, 30\,\mu\text{m})$ and plot the amplitudes of field components E_x (left column) and E_y (right column) in the plane determined by the diagonal and z-axis. Comparing the amplitudes calculated by the FMM and RKX-FPM reveals an obvious difference in the focal regions, which is because the RKX-FPM cannot predict the diffraction there. However, the interesting thing is that after the focal region, the RKX-FPM gives the correct fields again, which are also realized by the curve $\sigma(z)$ in Fig. 17 (c), calculated by Eq. (4.65). Please note that the curve is the relative deviation, including E_x and E_y in the xy-plane at different z-positions. We plot it under E_y to show that the peaks of $\sigma(z)$ coincide with the positions of the focal planes. We don't have solid proof for this phenomena yet: In principle, the RKX-FPM is derived under the geometric field assumption, but this example shows that the resulting geometric field can also be correctly predicted, even when there is a diffractive zone during propagation.

Taking advantage of this phenomena, we can simulate the same example in Fig. 16. In Fig. 18 (a), we show an illustration of field propagation through the Luneburg lens and a free space ($\epsilon_c = 1.0$), which is the amplitude of E_x calculated by the RKK-FPM

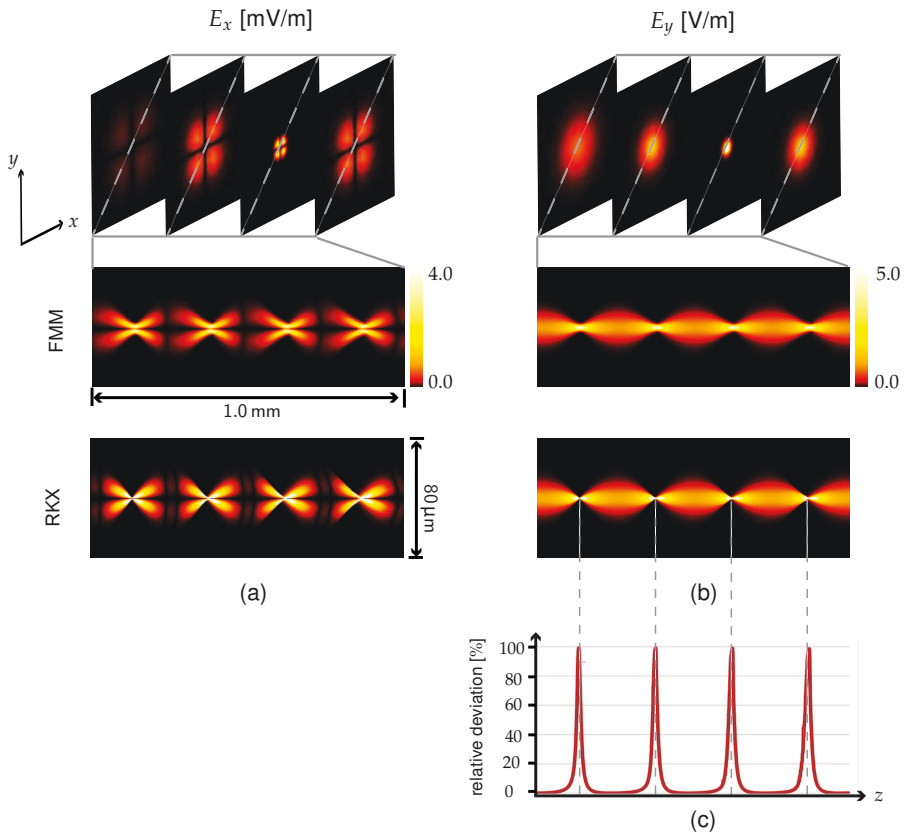


Figure 17: Simulation of field propagation in GRIN fiber (Fig. 12) by FMM+PMLs(second row) and RKX-FPM (third row). (a) Amplitude of field component E_x, and (b) that of E_y in the plane determined by the z-axis and the diagonal in the xy-plane are plotted. (c) Relative deviation $\sigma(z)$ between fields, calculated using two methods, is calculated using Eq. (4.65).

(y-invariant case). In the domain diagram in Fig. 18 (b), we can see that the calculation process contains two main steps. We first use the RKX-FPM to simulate the input plane wave propagating a distance of 110 μm into the switching plane, which is 10 μm after the focal plane. Here, the field is geometric. (2) Then we perform a free space propagation for $\Delta z = -10\,\mu\text{m}$, which performs an HFT of the result field in step (1), and we obtain $\tilde{\boldsymbol{E}}(\boldsymbol{\kappa})$; then, the multiplication of the phase term is performed, i.e., $\mathcal{P}\{\tilde{\boldsymbol{E}}(\boldsymbol{\kappa})\} = \tilde{\boldsymbol{E}}(\boldsymbol{\kappa}) \cdot \exp(ik_z(\boldsymbol{\kappa})\Delta z)$, with $\kappa_z(\boldsymbol{\kappa}) = \sqrt{k_0^2\epsilon_c - ||\boldsymbol{\kappa}||^2}$, and inverse FFT for a diffractive field. In Fig. 18 (c), we show the amplitude of field components E_x and E_y in the switching plane (left column), and those in the focus plane (right column). The relative deviation σ between this focus field and the one calculated by Mie theory is about 2 %. The error mainly comes from the interpolation of fields in the switching plane. The mesh of rays through the focus and that outside the ball lens are degenerated. To avoid the mesh degeneration-induced interpolation issue, we use an aperture and only allow the focused rays through. Thus, the field in the switching plane shows a sharp edge,

which induces the error. However, when lens or GRIN components become larger, the aperture-induced error is relatively reduced.

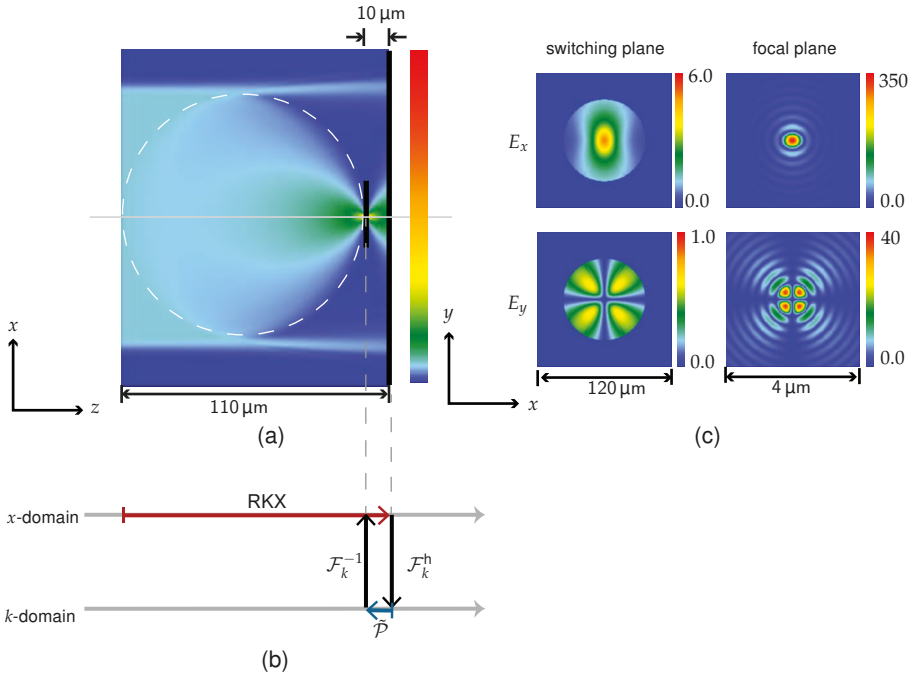


Figure 18: Simulation of a plane wave being focused using the RKX-FPM and inverse propagation. (a) The amplitude of a y-invariant plane wave propagating for 110 µm, including a y-invariant Luneburg lens and a free space with $\epsilon = 1.0$. The switching plane is denoted by the long vertical solid line, whereas the focal plane is denoted by the shorter line. (b) The domain diagram, and the techniques we use to calculate the fields in the focal plane. (c) The amplitude of field components E_x and E_y in the switching plane (left column), and those on the focal plane (right column).

The calculation takes no more than 1 min on a personal computer with an Intel Core i7-7700HQ CPU @2.80 GHz and a total 32 GB RAM.

4.3 SUMMARY

Starting from the ODEs of the RKK-FPM, after using the geometric field constraints (geometric field ansatzes, and replacing FFT by HFT), we derived the eikonal equation (4.19) and geometric field equations (4.26-4.29). Then, two ODEs in the x-domain were derived and solved using the Runge-Kutta method. This field solver was named the RK x-domain field propagation method (RKX-FPM), which is the calculation of geometric field propagation through GRIN media.

The properties of the RKX-FPM are:

- The input field and output field must be geometric

- The GRIN media have a general $\epsilon(\boldsymbol{r})$, with no restriction in symmetry or periodicity.
- The effects of the optical surface between the surrounding medium and GRIN medium is treated using LPIA. [29]
- The effects generated by a GRIN medium
  - It works for a non-paraxial field and always take the polarization into account, so the vectorial field components are well-predicted.
  - Reflection is not included
  - Evanescent waves are not included.

The RKX-FPM shows a remarkable advantage in that the calculation efficiency is remarkably high, especially for large-scale components having a size well beyond their wavelength. That can be understood as the RKX-FPM avoiding the redundant sampling induced by the wrapped phase.

Coming back to Tab. 3, it can be now updated to

Table 4: Field solvers to model GRIN media (Chapter 4)

Solvers/Feat.	Input field	$\epsilon(\boldsymbol{r})$		Effects in GRIN media				effort
		sym.	$\nabla \ln \epsilon$	surf.	pola.	diff.	ref/eva.	
Mie theory	plane wave[†]	sphe.		✓	✓	✓	✓	$N_x \times N_y \times M$
FMM	general	peri.		✓	✓	✓	✓	$(N_x \times N_y)^3 \times N_z$
FMM+PMLs	general			✓	✓	✓	✓	$(N_x \times N_y)^3 \times N_z$
Fiber solver	general	cyli.	!		✓	✓		$(N_x \times N_y)^3 \times M$
WPM	general		!‡	✓	✓	✓		$(N_x \times N_y)^2 \times N_z$
RKK-FPM	general				✓	✓		$N_x \times N_y \times N_z$
RKX-FPM	geometric		!§	✓	✓			$N_x \times N_y \times N_z$

[†] It can be extended to model the general field using plane wave decomposition and coordinate system transformation.

[‡] It is compensated by including the Fresnel matrices of local change from $\epsilon(\boldsymbol{\rho}, z^-)$ to $\epsilon(\boldsymbol{\rho}, z^+)$.

[§] The polarization crosstalk is included by solving the ODE of the normalized field vector (4.48).

! means that this solver has constraints.

✓ denotes that this solver correctly predicts the relevant effect.

M is the number of modes.

Compared with the RKK-FPM, the RKX-FPM only calculates the geometric field, and has three advantages: (1) The optical surface effect between the surrounding medium and the GRIN component can be treated using LPIA; (2) The calculation does not require two parallel planes, because local plane waves are traced and reconstructed in planes of any orientation (3) This method has the remarkable advantage of further numerical reduction, which makes the calculation very efficient, making this method the only choice when the scale of a GRIN component is exceptionally large, e.g., air turbulence.

As we mentioned at the beginning of the chapter, most GRIN components are designed and modeled using a ray tracing technique, and the field passing through the GRIN components must be a geometric field. The RKX-FPM is widely chosen in practical applications.

When there are focal regions inside a GRIN component, the RKX-FPM needs to be extended. The straightforward way to do this is by using the RKX-FPM to simulate the propagation of geometric fields, and then switch to the RKK-FPM to simulate the other parts. However, we also find that the RKX-FPM correctly predicts the output geometric field, even when there are focal regions between the input and output planes. This phenomenon can be taken to simulate different cases with much higher efficiency. As an example, when we simulate the Luneburg lens with $R = 50\,\mu m$, the RKK-FPM method takes 25 min, whereas the combined RK methods takes 4 min, and the RKX-FPM with inverse free space propagation takes several seconds. Unfortunately, however, we still have not found a solid mathematical proof for this phenomena, and we will continue to explore it.

5 APPLICATIONS OF GRADED-INDEX MEDIA

In this chapter, we introduce three different applications for graded-index (GRIN) media, including (1) analytical design of a perfect imaging GRIN ball lens, (2) design of a GRIN device using conformal transformation optics, and (3) analysis of the thermal lens effect. In this chapter, we mainly use the Runge-Kutta k-domain (Chapter 3, RKK-FPM), and the x-domain field propagation methods (Chapter 4, RKX-FPM) for modeling the GRIN components and connecting other field solvers, e.g., the free space propagation operator in homogeneous media, to model the whole system.

5.1 DESIGN AND SIMULATION OF GENERAL LUNEBURG LENS

A perfect imaging GRIN lens is one of the most popular applications of GRIN media. In this section, we introduce the design of general Luneburg lens, and extend it to design a multi-focus system, which is used not only for imaging, but also for other purposes, e.g., illumination.

5.1.1 *Single-Focus Luneburg Lens*

In 1944, Luneburg presented a theoretical description of a graded, radially symmetric lens that creates a perfect image of a point object [3]. The derivation is based on the eikonal equation in a spherical coordinate system. The illustration of the imaging system with the general Luneburg lens is in Fig. 19. The relative permittivity, $\epsilon(r)$, of the Luneburg lens can be written as

$$\epsilon(\bar{r}) = \begin{cases} \exp 2\left[\omega(\bar{r}, \bar{f}_0) + \omega(\bar{r}, \bar{f}_1)\right] & 0 \leqslant \bar{r} \leqslant 1 \\ 1 & \bar{r} > 1 \end{cases}, \tag{5.1}$$

with

$$\omega(\bar{r}, \bar{f}_i) = \frac{1}{\pi} \int_{\bar{r}}^{1} \frac{\arcsin x/\bar{f}_i}{(x^2 - \bar{r}^2)^{\frac{1}{2}}} \, \mathrm{d}x, \; i = 0, 1 \tag{5.2}$$

and

$$\begin{cases} \bar{r} = r/R\sqrt{\epsilon(r)} \\ \bar{f}_i = f_i/R, \; f_i \geqslant 1, \; i = 0, 1 \end{cases}. \tag{5.3}$$

r is the radial coordinate. The ball lens with radius R is centered at the origin, as shown in Fig. 19. f_0 is the distance between the object point and the sphere center, while f_1 is the distance between the image and the sphere center. One may have questions about the expression of $\bar{r}$, which is a function of r and $\epsilon(r)$. $\epsilon(r)$ is always continuous, even

when $r = R$ at the boundary of lens, $\epsilon(R) = 1$, and $\bar{r}(R) = 1$. Later when we discuss the solution Eq. (5.2), we will give a description for calculating r, as well as $\epsilon(r)$ from $\epsilon(\bar{r})$.

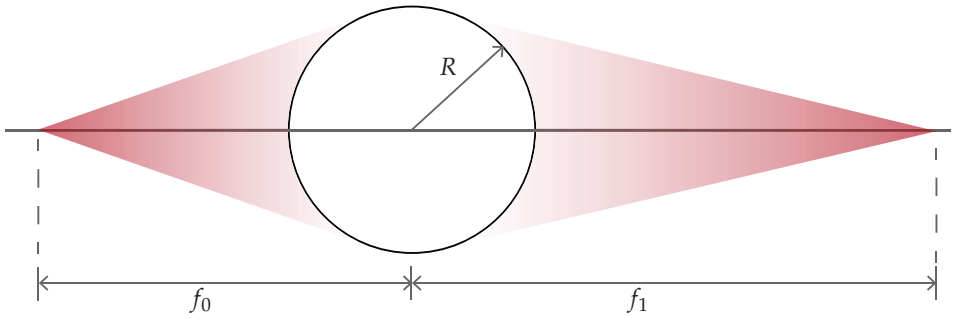


Figure 19: Illustration of a general Luneburg lens, which gives a perfect image of a point object. R is the radius of the Luneburg lens, f_0 is the distance between the object point and the sphere center, whereas f_1 is the distance between the image and the sphere center.

This is the fundamental design of a general Luneburg lens, which includes the following special cases:

1. Maxwell fisheye, $f_0 = f_1 = R$. In this case, $\omega(\bar{r}, \bar{f}_0) = \omega(\bar{r}, \bar{f}_1) = \ln \sqrt{1 + (1 - \bar{r}^2)^{1/2}}$. Similarly, $\epsilon(r)$ for $r \leqslant R$ can be analytically derived,

$$\epsilon(r) = \frac{4}{\left(1 + \dfrac{r^2}{R^2}\right)^2}. \tag{5.4}$$

Here, we show an example in Fig. 20. **System configuration.** We use an x-polarized spherical field as the source, with a 532 nm wavelength, propagating through a Maxwell's fisheye lens ($R = 20\,\mu$m), and calculate the focal field. The spherical field's point source is in the front pole of the lens, while the focal plane is through the rear pole. The spherical field is sampled 5 µm before the front pole. **Algorithm selection.** Diffraction is strong over the whole propagation. Thus, we use the RKK-FPM. **Numerical parameters.** The sampling distance in the xy-plane is $\lambda/2$ to include enough range of the angular spectrum, which is at least from $-k_0$ to k_0. The window size in the xy-plane should be large enough to avoid the aliasing effect, i.e., 70 µm. The number of sampling points along the x, y, and z axes are $N_x \times N_y \times N_z = 309 \times 309 \times 550$, and the calculation takes 10 mins. **Result**

discussion. As we mentioned in Chapter 3, y-invariant simulation can be used to determine numerical parameters and roughly check the field behavior. Thus, in this example, we calculate it and show the amplitude of E_x in Fig. 20 (b) and that of $\tilde{E}_x$ in (c). In Fig. 20 (b), the field propagates from one focus (front pole) to another focus(rear pole). In Fig. 20 (c), the angular spectrum is quite wide in the focal planes, while it is very narrow in the middle. In Fig. 20 (d), ray tracing through the fisheye lens is shown, which shows a phenomenon consistent with (b). In Fig. 20 (e), the focal field is calculated, with the diameter of the first dark ring as 0.53 µm.

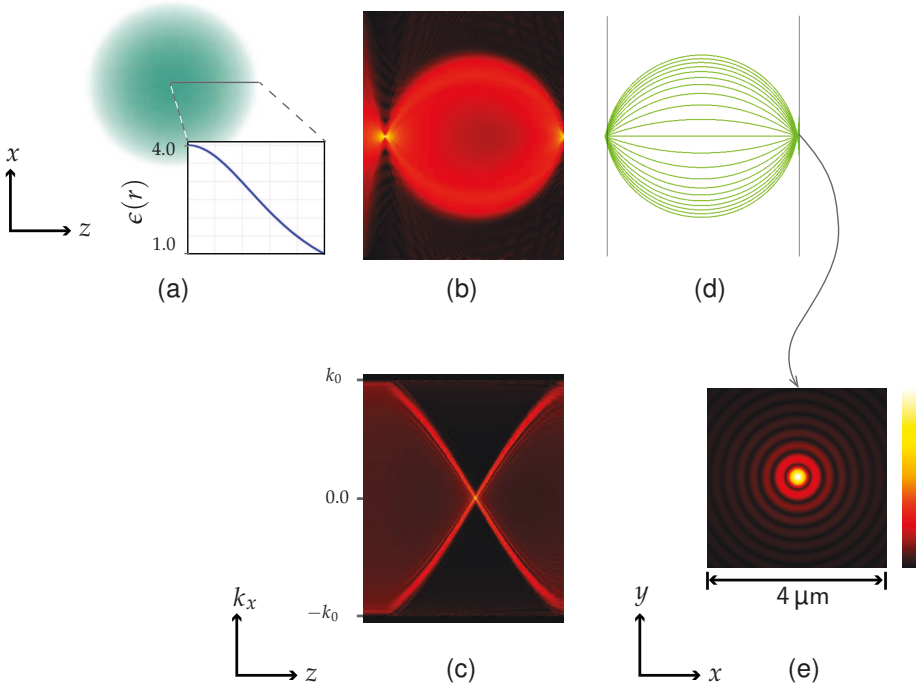


Figure 20: Maxwell fisheye. (a) $\epsilon(r)$ of the lens. A y-invariant case is simulated: (b) amplitude of E_x, and (c) amplitude of $\tilde{E}_x$ in the xz-plane. In the 3D simulation, we plot (d) the ray tracing result, and (e) the amplitude of E_x in the focal plane.

2. Luneburg lens, $f_0 = \infty$ and $f_1 = R$. In this case, $\omega(\bar{r}, \bar{f}_0) = 0$, and $\omega(\bar{r}, \bar{f}_1) = \ln \sqrt{1 + (1 - \bar{r}^2)^{1/2}}$. Using Eq. (5.1), $\epsilon(r)$ for $r \leqslant R$ can be analytically derived,

$$\epsilon(r) = 2 - (r/R)^2, \; r \leqslant R. \tag{5.5}$$

This is the example we used to prove the validity of algorithms in Section 3.1.5. Here, we show it in Fig. 21. **System configuration.** The source is an x-polarized plane wave with aperture size of 38 µm and a 532 nm wavelength. The radius of the Luneburg lens is 20 µm. **Algorithm selection.** This example can be simulated by the RKK-FPM or the combined RKX- and RKK-FPMs, as shown in Section

4.2. **Numerical parameters.** The number of sampling points along the x, y, and z axes are $N_x \times N_y \times N_z = 245 \times 245 \times 450$, and the calculation takes 4 mins. **Result discussion.** In Fig. 21 (b), the amplitude of E_x, that plane wave propagating through the lens and focused at the rear pole, is shown. The diffraction in the boundary caused by the aperture truncation (38 µm) starts from the source plane. In Fig. 21 (c), the amplitude in the k domain is plotted: the vertical axis is the k_x axis, from $-k_0$ to k_0. The angular spectrum of the initial plane wave is an array pattern, which is very narrow along the k_x axis and becomes broader during propagation. In Fig. 21 (d), ray tracing through the lens is shown, and the phenomenon is consistent with that shown in (b). In Fig. 21 (e), the focal field is calculated. The spot diameter is 0.6 µm along the y-axis, while it is 0.8 µm along the x axis.

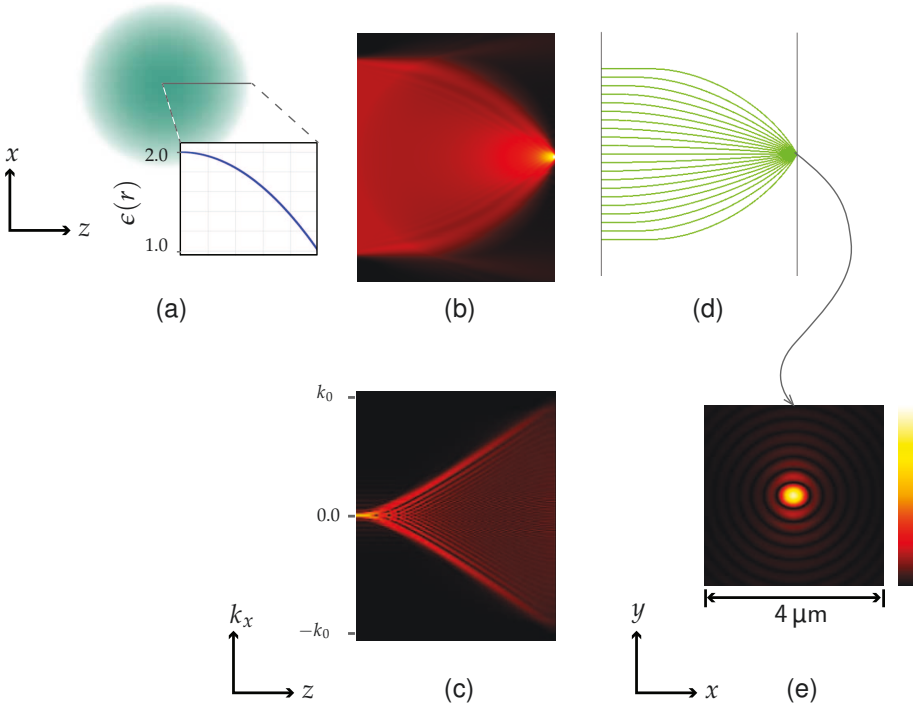


Figure 21: Luneburg lens. (a) $\epsilon(r)$ of the lens. Next, a y-invariant case is simulated: (b) amplitude of E_x, and (c) amplitude of $\tilde{E}_x$ in the xz-plane. In the 3D simulation, we plot (d) the ray tracing result, and (e) the amplitude of E_x in the focal plane.

In 1983, Sochacki [5, 6] proposed an analytical solution, shown as Eq. (5.2), for any $f_0 \geqslant R$ and $f_1 \geqslant R$, based on the Taylor expansion, which is

$$\omega(\bar{r}, \bar{f}_i) = \frac{1}{\pi}(1-\bar{r})^{1/2} \sum_{k=0}^{\infty} s_k(\bar{f}_i)\bar{r}^{2k}, \tag{5.6}$$

with

$$s_k(\bar{f}_i) = \sum_{l=0}^{\infty} \frac{(2l)!}{4^l (l!)^2 [2(k+l)+1]^2 \bar{f}_i^{2(k+l)+1}}. \tag{5.7}$$

After substituting Eqs. (5.6) and (5.7) into Eq. (5.1), the permittivity of the Luneburg lens $\epsilon(\bar{r})$ can be calculated. We conclude the procedure of the calculation of $\epsilon(r)$:

1. Define an equidistant sampling grid for $\bar{r}$, i.e., $\bar{r} = 0, \Delta\bar{r}, 2\Delta\bar{r}, \cdots, (N-2)\Delta\bar{r}, 1$ with $\Delta\bar{r} = \dfrac{1}{N-1}$ and N as the number of sampling points.

2. Implement Eqs. (5.6) and (5.7) to calculate $\omega(\bar{r}, \bar{f}_i)$, with $i = 0, 1$. In this process, a threshold needs to be defined, e.g., 10^{-8}, to truncate l and k. As an example, when $|s_k(\bar{f}_i)|_{l_{\max}=L} - s_k(\bar{f}_i)|_{l_{\max}=L-1}| < 10^{-8}$, the calculation of Eq. (5.7) stops, and returns the value $s_k(\bar{f}_i)|_{l_{\max}=L}$.

3. Substitute $\omega(\bar{r}, \bar{f}_i)$ into Eq. (5.1) to calculate $\epsilon(\bar{r})$.

4. Calculate $r(\bar{r})$ by using
$$r(\bar{r}) = \bar{r} / \sqrt{\epsilon(\bar{r})} R. \tag{5.8}$$
After mapping $\epsilon(\bar{r})$ to the relevant r, we finally obtain $\epsilon(r)$, with non-equidistant sampled r.

5. Linear interpolation is implemented to calculate $\epsilon(r)$ at any r. As an example, when $r_i \leqslant r < r_{i+1}$,
$$\epsilon(r) = \epsilon(r_i) + \frac{r - r_i}{r_{i+1} - r_i}[\epsilon(r_{i+1}) - \epsilon(r_i)] \tag{5.9}$$

Here, we show an example in Fig. 22. **System configuration**. The source field is an x-polarized spherical wave, which is sampled in the plane 80 µm after the point source, with a 532 nm wavelength. The generalized Luneburg lens has $R = 20$ µm, $f_0 = 100$ µm, and $f_1 = 30$ µm. After substituting the parameters into the design procedure just presented, $\epsilon(r)$ can be calculated, and it is shown in Fig. 22 (a). **Algorithm selection.** In Fig. 22 (b), the necessary solvers to model the whole system are given. To model the field propagation through the Luneburg lens, the RKX-FPM is selected, because both input and output fields of the GRIN lens are geometric. After the lens, light further propagates in free space (homogeneous media with $\epsilon_c = 1.0$). The free space propagation operator [32–34], which is composed of an (inverse) Fourier transform and a linear operation in the k-domain $\times \exp(ik_z(\boldsymbol{\kappa})\Delta z)$ is used. This combination of different field solvers to model the whole optical system is the general idea of field tracing, in which different solvers are connected [26]. The calculation takes several ms. **Result analysis.** In Fig. 22 (a), rays propagate from an object point to a focal spot. The RKX-FPM offers ray tracing of the ball lens [74]. The diffraction pattern, i.e., amplitudes of E_x, E_y, and E_z, are shown in (c). An extremely small E_y, which is the polarization crosstalk, is created during field propagation within the GRIN lens. The amplitude of E_x is also compressed along the y axis, but not as obviously as in Fig. 21. This is because the numerical aperture (NA) decreases in this case.

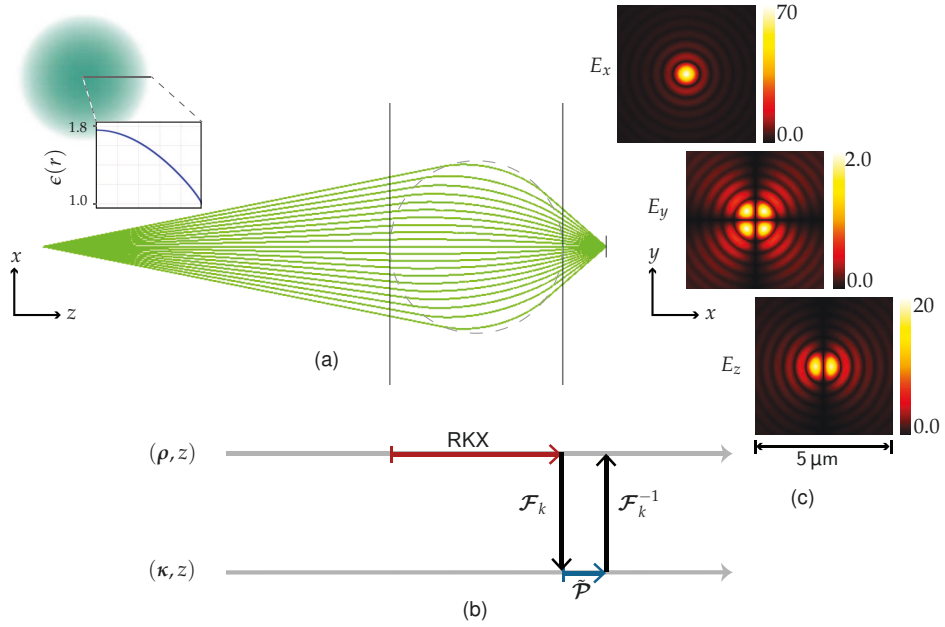


Figure 22: Single-focus generalized Luneburg lens. (a)The designed $\epsilon(r)$, and ray tracing from point object to focal plane. (b) Domain diagram with selected solvers that are used to model the system. (c) Amplitude of E_x, E_y, and E_z in the focal plane.

5.1.2 *Double-Focus Luneburg Lens*

In 1984, Sochacki [7] extended the single-focus Luneburg lens to a double-focus lens and gave a straightforward formula that showed a clear physical understanding. In Fig. 23, the Luneburg lens contains two concentric spheres, with radii R_1 and R. The light field from a point source propagates through the outer shell-central sphere-the outer shell is focused at the point at a distance f_1, while light that propagates through the outer shell only is focused at the position f_2.

The permittivity $\epsilon(r)$ for $r \leqslant R$ is written as

$$\epsilon(\bar{r}) = \begin{cases} \exp 2\left[\omega(\bar{r}, \bar{f}_0) + \omega(\bar{r}, \bar{f}_2)\right] & \bar{r}_1 \leqslant \bar{r} \leqslant 1 \\ \exp 2\left[\omega(\bar{r}, \bar{f}_0) + \omega(\bar{r}, \bar{f}_2) + \omega(\frac{\bar{r}}{\bar{r}_1}, \frac{\bar{f}_1}{\bar{r}_1}) - \omega(\frac{\bar{r}}{\bar{r}_1}, \frac{\bar{f}_2}{\bar{r}_1})\right] & \bar{r} < \bar{r}_1 \end{cases} \tag{5.10}$$

with $\bar{r}_1 = \sqrt{\epsilon(R_1)} R_1 / R$, and arbitrary value $\bar{r}_1 \subset [0, 1]$. ω is the function in Eq. (5.2), and the solutions to ω are Eqs. (5.6) and (5.7). When we implement Eq. (5.10), we need to set the value of $\bar{r}_1$ as an input value, and follow the procedure for calculating $\epsilon(r)$, discussed in Section 5.1.1.

Here, we give an example in Fig. 24. The source field is an x-polarized spherical wave, which is sampled in the plane 80 µm after the point source, with a 532 nm wavelength. The Luneburg lens has $R = 20\,\mu\text{m}$, $f_0 = 100\,\mu\text{m}$, $f_1 = 30\,\mu\text{m}$, $f_2 = 60\,\mu\text{m}$, and $\bar{r}_1 =$

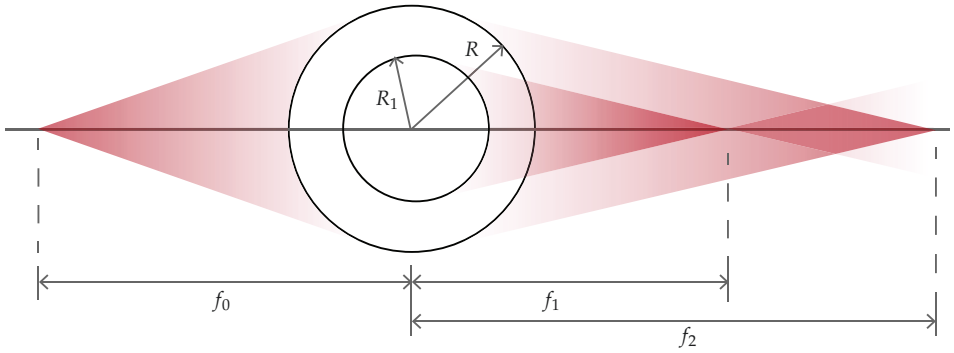


Figure 23: Illustration of a general Luneburg lens, which generates two image points for one object. R_1 is the radius of the central Luneburg lens, while R is the radius of the whole Luneburg lens. f_0 is the distance between the object point and the lens center, while f_1 and f_2 are the distances between two image points and the lens center, respectively.

0.55. Here $\bar{r}_1$ can be used to adjust the energy distribution of the two focuses. After substituting the parameters into the design procedure just presented, $\epsilon(r)$ can be calculated and is shown in Fig. 24 (a). We use the RKX-FPM to model the GRIN lens, and free space propagation is done in the k-domain with the proper Fourier transforms. The simulation takes ms. In Fig. 24 (b), the ray tracing result is shown. We get two focal spots at positions 10 µm and 40 µm after the lens rear pole. The two beams related to two focal spots, respectively, coexist everywhere. In Fig. 24 (c), we plot the energy density ($|E_x|^2 + |E_y|^2 + |E_z|^2$) at positions 10 µm, 25 µm, 40 µm, and 50 µm (gray bars in (b)) by using the logarithmic color map. In Fig. 25 (b), the illuminance of both focal spots is shown. Both are small, and the diameters of the spots are approximately 1 µm. The second focus has a stronger side lobe, which can be understood as a smaller NA, and this focal spot is created by a spherical field truncated by a ring-shaped aperture. The position of the first focus is identical to that in Fig. 22, but here, the focus spot is larger because of the smaller NA. Next, we calculate the energy density on 50 planes at different values of z, and extract the values along the x-axis to obtain the distribution in the xz-plane, as shown in Fig. 24 (d). Compared with the ray tracing figure, the energy density shows analog behavior, but the interference pattern between the two beams is clearly presented. In Fig. 25, the illuminance of the xz-plane is plotted, where the two focal regions are clearly illustrated. The energy density along the z-axis is extracted and shown in (a). By adjusting $\bar{r}_1 = 0.55$, the largest energy density of both focuses are made

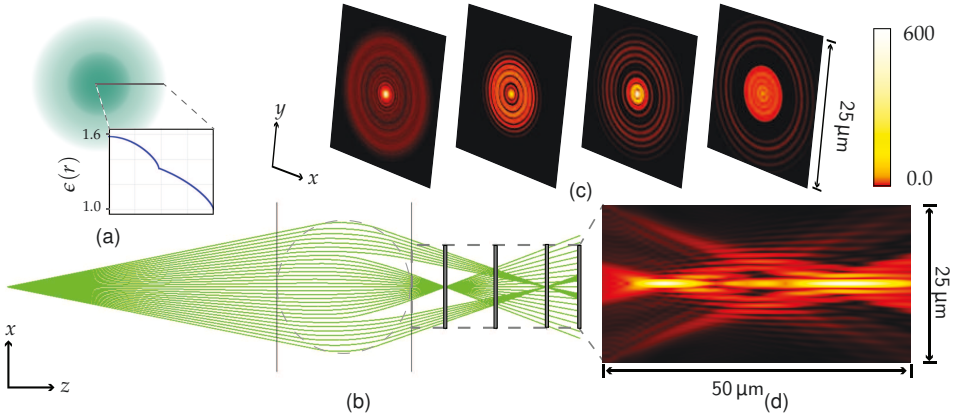


Figure 24: Double-focus lens. The Luneburg lens has $R = 20\,\mu m$, $f_0 = 100\,\mu m$, $f_1 = 30\,\mu m$, and $f_2 = 60\,\mu m$. (a) $\epsilon(r)$. (b) Ray tracing result from the point source to the detector plane, which is 50 µm after the lens rear pole. Here, there are four detector planes (gray vertical bars), for which the energy density is plotted in (c) with logarithmic color maps. (d) The energy density in the xz-plane, within the dashed box of (b).

similar, and the depth of the first focus is about 10 µm, whereas that of the second is 20 µm.

5.1.3 *Multi-Focus Luneburg Lens*

Based on the derivation of Sochacki in 1984 [7], we can extend the design of the generalized Luneburg lens to generate a multi-focus lens. In Fig. 26, each spherical shell is a Luneburg lens that generates an image point. R_i is the radius of the i^{th} Luneburg lens shell, with $i = 1, 2, .., N-1$. R as the radius of the whole Luneburg lens. f_0 is the distance between the object point and the lens center, while f_i is the distance between the i^{th} image point and the lens center.

The relative permittivity $\epsilon(r)$ for $r \leqslant R$ is written as

$$\epsilon(\bar{r}) = \begin{cases} \exp 2\left[\omega(\bar{r}, \bar{f}_0) + \omega(\bar{r}, \bar{f}_N)\right] & \bar{r}_{N-1} \leqslant \bar{r} \leqslant 1 \\ \exp 2\left\{\omega(\bar{r}, \bar{f}_0) + \omega(\bar{r}, \bar{f}_N) + \sum_{j=i}^{N-1}\left[\omega(\frac{\bar{r}}{\bar{r}_i}, \frac{\bar{f}_i}{\bar{r}_i}) - \omega(\frac{\bar{r}}{\bar{r}_i}, \frac{\bar{f}_{i+1}}{\bar{r}_i})\right]\right\} & \bar{r}_{i-1} \leqslant \bar{r} < \bar{r}_i \end{cases} \tag{5.11}$$

with $i = 1, 2, ..., N-1$, and $\bar{r}_0 = 0$. $0 \leqslant \bar{r}_1 \leqslant \bar{r}_2 \leqslant \cdots \leqslant \bar{r}_{N-1} \leqslant 1$. ω is the function in Eq. (5.2), and the solutions for ω are Eqs. (5.6) and (5.7). When we implement Eq. (5.11), we need to set the value of $\bar{r}_i$ as an input value. All the others follow the implementation procedure for calculating $\epsilon(r)$, discussed in Section 5.1.1.

Here, we show an example of a triple-focus Luneburg lens in Fig. 27. The source field is an x-polarized spherical wave, which is sampled in the plane 80 µm after the point

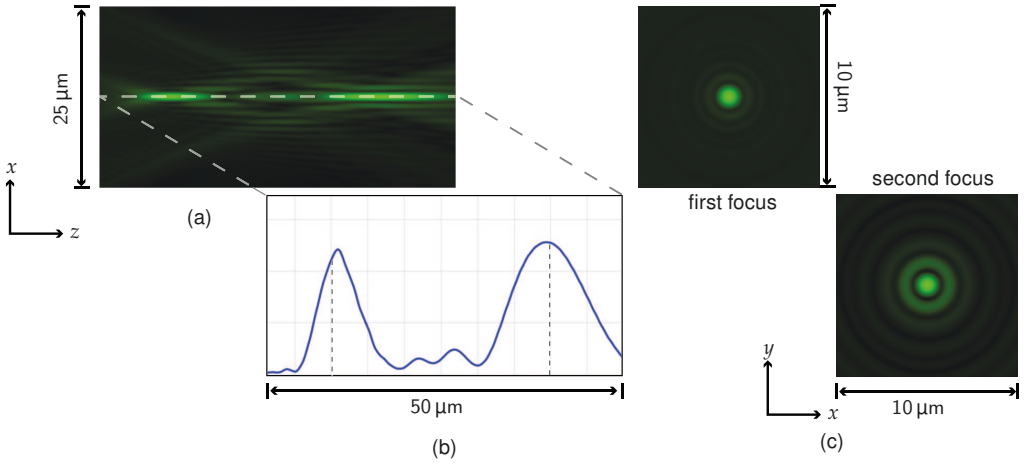


Figure 25: Light propagation after the double-focus lens, plotted from the perception of the human eye. (a) Illuminance in the xz-plane. (b) Energy density along the z-axis. Two vertical dot lines denote the positions of the focuses, i.e., 10 µm and 40 µm, (c) Illuminance of two focal spots.

source, with a 532 nm wavelength. The Luneburg lens has $R = 20\,\mu m$, $f_0 = 100\,\mu m$, $f_1 = 25\,\mu m$, $f_2 = 40\,\mu m$, and $f_3 = 60\,\mu m$. We set $\bar{r}_1 = 0.4$, and $\bar{r}_2 = 0.7$ to obtain a uniform energy allocation for the three beams respectively related to the three focuses. In the following discussion, we name the three beams as beams 1, 2, and 3. After substituting the parameters into the design procedure just presented, $\epsilon(r)$ can be calculated, as shown in Fig. 27 (a). In Fig. 27 (b), the ray tracing result is shown. There are three focal spots at positions 5 µm, 20 µm, and 40 µm after the rear pole of the lens. The three beams, 1, 2, and 3, are coherent with each other because they come from the same monochromatic point source. The three beams coexist everywhere, and the interactions among them are explored with a logarithmic color map: (c) shows the energy density in three focal planes, while (d) shows the energy density in the xz-plane. In the ray tracing result shown in (b), before the first focal plane, three beams are spatially separated. The same behavior can be observed in (d). To explore the influence of these interference patterns, which can be considered as stray light, we plot the illuminance (perception of the human eye) in the xz-plane (Fig. 28 (a)) and three focal planes (Fig. 28 (c)). The diameter of all three focal spots are extremely small, on a scale of 1 µm, which is satisfactory. The energy density on the z-axis is extracted in Fig. 28 (b), and the three focal regions are separated, with focal depths of 10 µm, 15 µm, and 20 µm, respectively.

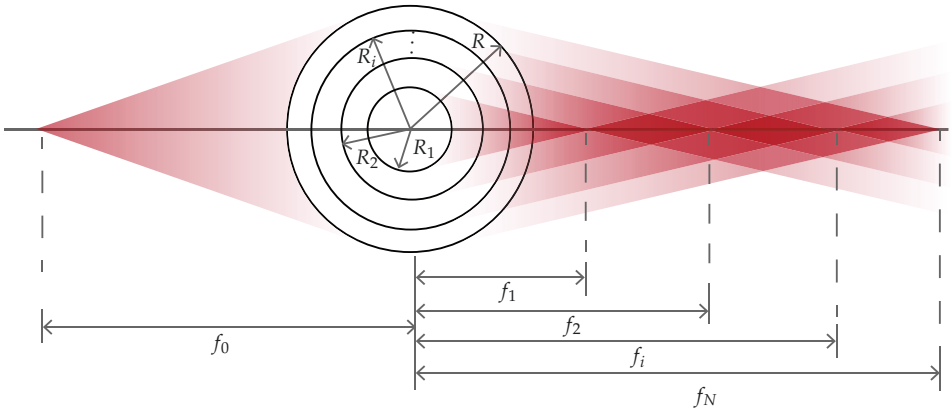


Figure 26: Illustration of a general Luneburg lens, which gives N image points to one object. R_i is the radius of the i^{th} Luneburg lens shell, with $i = 1, 2, .., N-1$. R as the radius of the whole Luneburg lens. f_0 is the distance between the object point and the lens center, whereas f_i is the distance between the i^{th} image point and the lens center.

5.1.4 *Discussion of the General Luneburg Lens*

In the examples, $\epsilon(r)$ when $r > R$ is 1.0, which can be set to any constant value ϵ_c. Afterwards, the design of ϵ for the generalized Luneburg lens does not change, but needs an additional multiplication with ϵ_c. Currently, $\epsilon(r = R)$ is identical to that of the surrounding media. In further research, we will include the discontinuity of $\epsilon(r)$, which is discussed by Morgan [4].

The most straightforward application of the designed lens is use in an imaging system to increase image depth, or even create images for near, far, and intermediate distances.

The other application can be imagined from the xz-plane in Figs. (24-28). Here, the three focal regions are still separated, but we can bring them closer to each other until they are not resolved, which can be used in optical fabrication.

Mao [75] explores the generalized Luneburg lens(up to the double-focus) in a photonic nanojet, which creates a narrow, high-energy electromagnetic field after the sphere. Photonic nanojets can be used in high-resolution optical detection(as an imaging system) or nanolithography (fabrication).

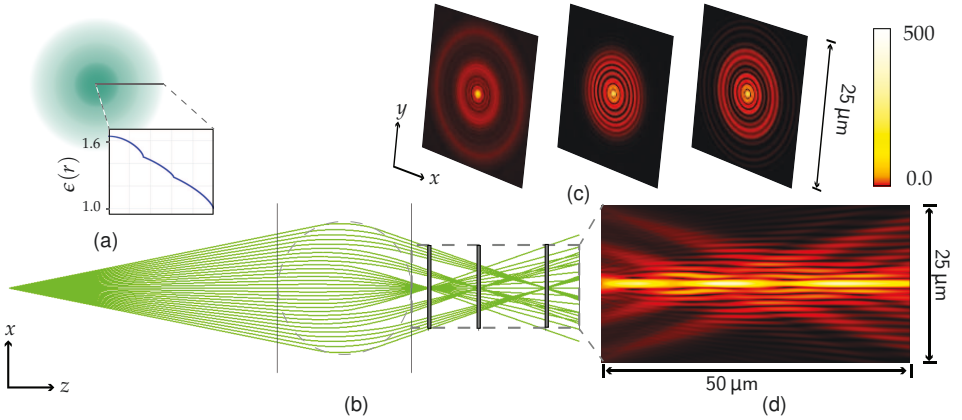


Figure 27: Triple-focus lens. The Luneburg lens has $R = 20\,\mu m$, $f_0 = 100\,\mu m$, $f_1 = 25\,\mu m$, $f_2 = 40\,\mu m$, and $f_3 = 60\,\mu m$. (a) $\epsilon(r)$, and there are three lens shells. (b) Ray tracing result from the point source to the detector plane, which is $50\,\mu m$ after the lens rear pole. Here, there are three detector planes (focal planes), in which the energy density is plotted (c) with a logarithmic color map. The energy density in the xz-plane, within the dashed box, is plotted in(d).

5.2 DESIGN AND SIMULATION OF COMPONENTS USING CONFORMAL TRANSFORMATION OPTICS

Conformal transformations [47] are important for solving problems in physics and engineering. By using a proper mapping, complex geometry is transformed into a simpler one, which greatly simplifies further calculation. Inversely, in 1944, Luneburg proposed the idea to design a GRIN medium analytically by using the concept of conformal transformation (Eq. 28.56) [3]. In 2006, Leonhardt[17] gave the same formula and proposed the design of a cylindrically symmetric invisibility device based on the formula. As it is a remarkable design method for GRIN components, we repeat the design task here, and its workflow, which follows Luneburg and Leonhardt faithfully.

5.2.1 *Design Workflow*

The conformal transformation optical design is based on the ray representation of light, and the design formula is derived based on the eikonal equation (4.19). Therefore, we can imagine that rays in different domains are mapped. Please note that this method can only be used to design a y-invariant GRIN component and can be further extended to cylindrically symmetric components.

In a real physical space in the xz-plane, ray distribution is determined by purpose. As an example, in telecommunications, we need to design bending waveguides to change

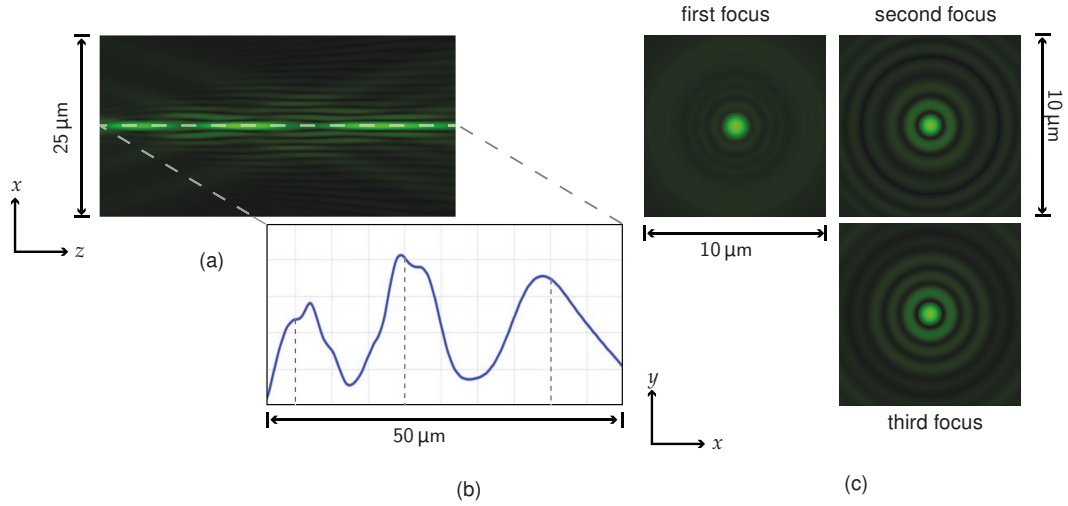


Figure 28: Light propagation after the triple-focus lens, plotted from the perception of the human eye. (a) Illuminance in the xz-plane. (b) Energy density along the z-axis. Three vertical dot lines denote the positions of the focuses, i.e., 5 µm, 20 µm and 40 µm, (c) Illuminance of the three focal spots.

the propagation direction of light to avoid bend loss. So, in the xz-plane, the desired rays are curved as circles, as shown in Fig. 29 (a). The virtual space, with coordinates (u, v), gives a simpler geometry of ray distribution. As in (a), all rays move along the azimuthal direction, we decide that in the uv-plane, straight rays are parallel to the v-axis, as shown in Fig. 29 (b). The simplest relative permittivity to achieve this ray distribution is $\epsilon_w(u, v) = \epsilon_c$, a constant. Thus, we need to find the mapping relation of rays from the uv-plane to those in the xz-plane, and then $\epsilon(x, z)$ is calculated.

Now let us interpret the task mathematically. Two complex planes are defined: one t-plane, with $t = x + \mathrm{i}z$, and one w-plane, with $w = u + \mathrm{i}v$. We need to search the transformation $w(t)$ and $t(w)$ from the ray mappings, and calculate the relative permittivity using

$$\epsilon(x, z) = \epsilon_w(u, v) \left| \frac{\mathrm{d}w(t)}{\mathrm{d}t} \right|^2, \tag{5.12}$$

which is derived based on the eikonal equation [3, 17].

Find $w(t)$ and $t(w)$

This requires knowledge of the transformation properties of several fundamental functions, e.g., power function, or exponential function. In 2015, Xu *et al.*[21] reviewed several different conformal transformations that could be used as a database. In this example, the mapping is quite straightforward. In the t-plane, the position coordinates on a single ray (red ray in Fig. 29 (a)) satisfy $|t| = \sqrt{x^2 + z^2} = t_c$, while in the w-plane, the

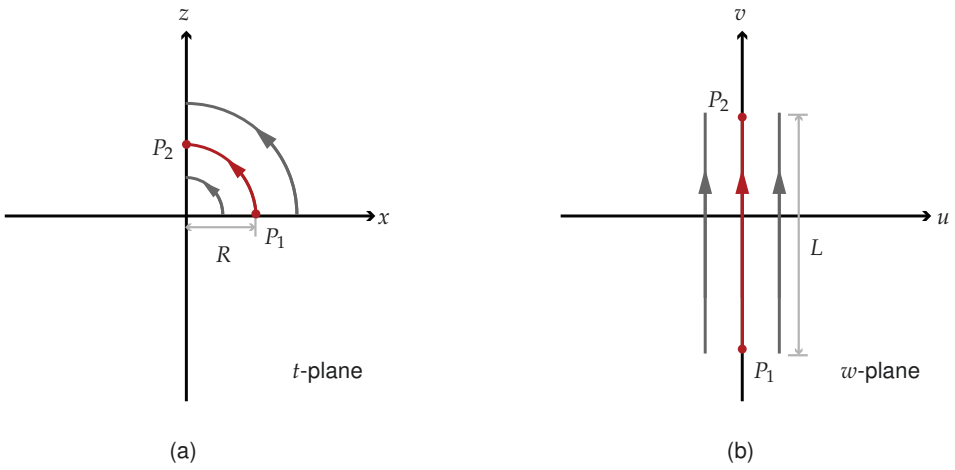


Figure 29: Conformal transformation between t-plane and w-plane. (a) Rays in t-plane propagate along circles centered at the origin, and the red ray starts from P_1 and ends in P_2. (b) Rays in the w-plane propagate along the v-axis, and the red ray from P_1 to P_2 is mapped to the one in (a).

real part of the position coordinates of the related ray (red ray in Fig. 29 (b)) is also a constant, i.e., $u = u_c$, so the mapping function can be

$$t(w) = A\exp(Bw + C), \tag{5.13}$$

with A, B, and C as the real-valued free parameters*.

To fix A, B, and C, we map P_1 and P_2 in the t- and w- plane, with

$$\begin{cases} t|_{P_1} &= R \\ t|_{P_2} &= \mathrm{i}R \\ w|_{P_1} &= -\mathrm{i}\dfrac{L}{2} \\ w|_{P_2} &= \mathrm{i}\dfrac{L}{2}. \end{cases} \tag{5.14}$$

Substituting the mapping into Eq. (5.13), we get

$$t(w) = R\exp(\frac{\pi}{2L}w + \mathrm{i}\frac{\pi}{4}), \tag{5.15}$$

and

$$w(t) = \frac{2L}{\pi}\ln\left[\frac{t}{R}\exp(-\mathrm{i}\frac{\pi}{4})\right]. \tag{5.16}$$

* One can assume these three parameters to be complex-valued, and more points are needed to find a more accurate mapping relation.

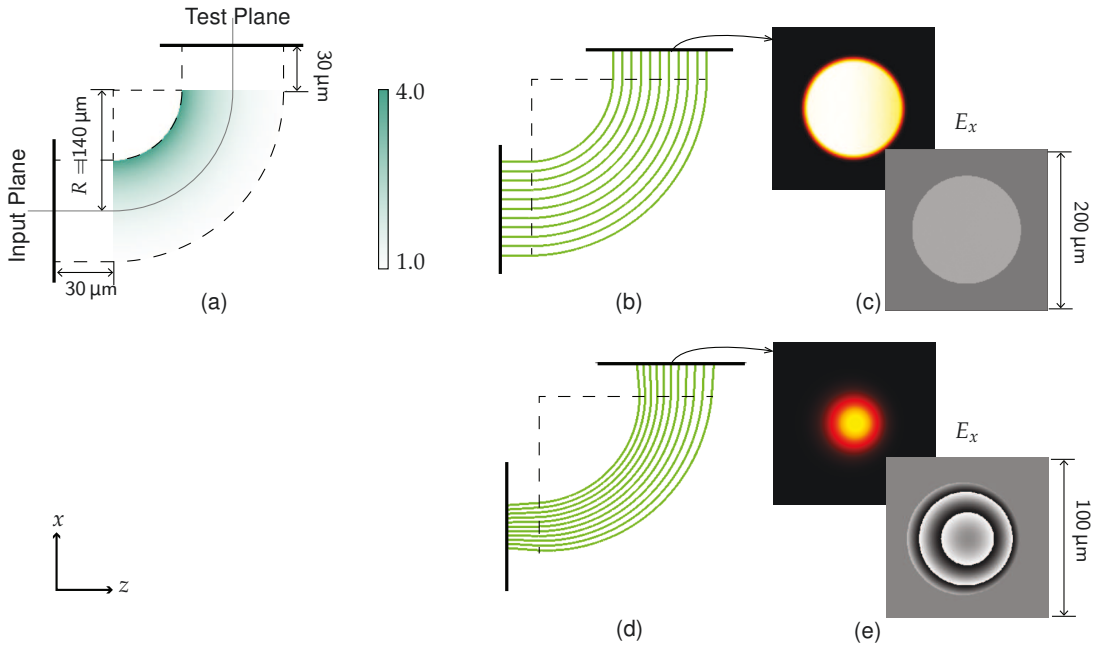


Figure 30: Modeling of the y-invariant bending waveguide. (a) Illustration of the optical setup, with the color map representing $\epsilon(x,z)$. (b) Ray tracing result for a plane wave source, with an aperture diameter of 100 µm. (c) Amplitude and phase of E_x in the test plane. (d) Ray tracing result for a Gaussian source, with a waist radius of 5 µm. (e) Amplitude and phase of E_x in the test plane.

$\epsilon(x,z)$

Now, the relative permittivity can be calculated using Eq. (5.12),

$$\epsilon(x,z) = \frac{4L^2\epsilon_c}{\pi^2(x^2+z^2)} \tag{5.17}$$

5.2.2 *Modeling of the Bending Waveguide*

Eq. (5.17) gives the relative permittivity $\epsilon(x,z)$ of the y-invariant bending waveguide. Now we set

$$L = \frac{\pi R}{2}, \tag{5.18}$$

with $R = 140\,\mu\text{m}$, and $\epsilon_c = 1.69$. The optical setup is illustrated in Fig. 30 (a).

We model a E_x-polarized plane wave propagating through the optical system, using an aperture diameter of 100 µm and the RKX-FPM. Fig. 30 (b) shows the ray tracing result, in which all rays are parallel, and the curved parts are concentric quarter circles, which meets the design purpose. In the test plane, the amplitude and phase of E_x are shown in Fig. 30 (c). The gradient distribution of the amplitude along the x-axis is

caused by the optical surface effect between the surrounding medium and the GRIN bending waveguide, while the phase is a constant within the aperture, which meets the design purpose.

Next, we model a E_x-polarized Gaussian beam with a waist radius of 5 µm propagating through the optical system. To ensure the field in the input plane is geometric, the beam waist is located 160 µm before the input plane. In Fig. 30 (d), the rays are divergent, while (e) shows the amplitude and phase of E_x in the test plane. The phase is spherical, as it is a divergent beam, and the M^2 of the beam in the test plane is 1.019 in the x-direction, whereas it is 1.0079 in the y-direction, demonstrating that the bending waveguide keeps the property of the beam after turning the direction of propagation 90°.

Please note that conformal transformation is one branch of transformation optics, which can be used to design y-invariant, or further extend cylindrically symmetric GRIN components. In 2006, Pendry et al. [18] showed that Maxwell's equations are invariant against coordinate transformations. This generalizes the transformation optics between any domain. However, this design strategy results in an inhomogeneous anisotropic component, for which the modeling is beyond the scope of the thesis. We have the research plan to do further exploration of the transformation optics and extend our modeling techniques to include anisotropic media.

5.3 MODELING OF THERMAL LENSING EFFECT

Thermal-optical simulation is one of the applications in which GRIN modulation is undesired, but it must be taken into consideration. When an optical beam has high power, the induced temperature distribution changes the $\epsilon(\boldsymbol{r})$ of the medium. In this section we present two examples and use two different ways to calculate $\epsilon(\boldsymbol{r})$. The first one simulates a Gaussian beam focused by the thermal lens effect induced by a high-power laser, where $\epsilon(\boldsymbol{r})$ is given by an analytical modal [76]. The second one analyses the thermally induced tolerance of a fiber coupling lens, where $\epsilon(\boldsymbol{r})$, as well as the mechanical distortion of the optical surfaces are calculated using the engineering software Ansys [44].

5.3.1 *Mathematical Model of Thermal Lens for Laser Rod*

This experiment is simulated with an Nd:YAG laser rod, as shown in Fig. 31. The laser power is from 8 kW to 20 kW. The focal lens and focal spot changing caused by the thermal lensing effect will be explored. In the thermal-optical simulation, we first do a thermal analysis to calculate the thermal effects induced in $\epsilon(\boldsymbol{r})$ versus the incident power P^{in}. Then, we do an optical simulation and propagate the electromagnetic field through the optical system with the laser rod.

Thermal analysis. In the 1970s, Koechner [76] described a theoretical model for doing a thermal analysis of an optically pumped Nd:YAG laser rod, and $\epsilon(\boldsymbol{r})$ was written as a first-order Taylor expansion,

$$\epsilon(\boldsymbol{r}) = \epsilon(T_0) + \frac{\mathrm{d}\epsilon}{\mathrm{d}T}[T(\boldsymbol{r}) - T_0], \tag{5.19}$$

with $\epsilon(T_0)$ as the relative permittivity at temperature T_0, which is the temperature of the rod center. $T(\boldsymbol{r})$ is the temperature at position $\boldsymbol{r} = (\rho, \theta, z)$ in the cylindrical coordinate system. The center of the first optical surface of the laser rod is the origin of the cylindrical coordinate system. Koechner gave another tensor, $\Delta\underline{\epsilon}(\boldsymbol{r})$, which is the stress-dependent change required to make the media anisotropic. In this work, this term is not considered. In Eq. (5.19), the temperature distribution $T(\boldsymbol{r})$ is only $\rho-$dependent.

$$T(\rho) = T_0 - \frac{\eta P^{\mathrm{in}} \rho^2}{4\pi K \rho_0^2 L}, \tag{5.20}$$

with P^{in} as the variable, and the other related parameters are shown in Tab. 5.

Table 5: Parameters of laser rod [76]

Nd:YAG laser rod			
Physical properties			
Rod length	L	7.5	cm
Rod radius	ρ_0	0.31	cm
Permittivity in rod center	$\epsilon(T_0)$	3.32	
	$\frac{\mathrm{d}\epsilon}{\mathrm{d}T}$	14.6	$^{\circ}\mathrm{C}^{-1}$
Thermal coefficients			
Thermal conductivity	K	0.111	$\mathrm{W\,cm^{-1}\,^{\circ}C^{-1}}$
fraction of heat dissipated	η	0.05	

Optical simulation. Here we simulate a Gaussian beam focused by the thermal lens effect induced by a high-power laser. A laser crystal rod under a thermal lens effect with the parameters in Tab. 5 is shown in Fig. 31. A field propagating through the thermal lens can be modeled by the RKX-FPM, while the field propagating after the thermal lens to the focal region should be calculated by other appropriate techniques to include the optical diffraction, i.e., the free space propagation operator. This combination of different field solvers is a typical example of the idea of field tracing, in which different solvers are connected [26].

We use a fundamental Gaussian beam, with a waist radius of 750 µm, and at a wavelength of 632.8 nm as the input beam. The input plane of the thermal lens is located at 300 cm after the beam waist to ensure that the field is in its geometric field zone (Fig. 31). As the power of the input beam increases from 8 kW to 20 kW, the focal plane, which

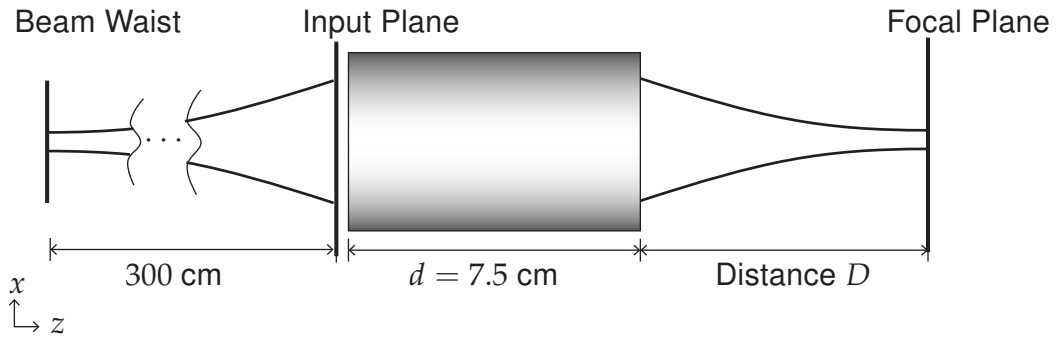


Figure 31: Illustration of the optical system to focus a Gaussian beam onto a focal plane by the thermal lens effect.

gives the smallest beam diameter, is found, and the beam waist diameter in the focal plane is computed. In Fig. 32, the dependence of the distance D, which is measured between the output interface of the thermal lens and the focal plane, and the beam waist diameter of the focused field in the focal plane are presented. As the input power increases, the spatial variation in $\epsilon(\rho)$ becomes more pronounced, which introduces a stronger thermal lens effect. Therefore, D becomes shorter. This coincides with the result presented in [76]. Meanwhile, the divergence angle of the output beam turns out to be large, leading to a smaller waist diameter in the focal plane.

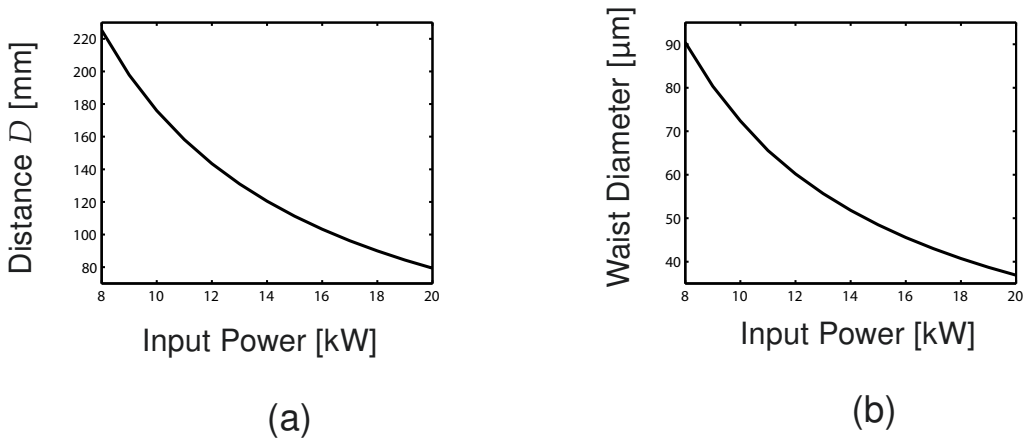


Figure 32: Optical simulation results (a) distance D, which is measured between the output interface of the thermal lens and the focal plane, as a function of the input power; (b) beam waist radius, which is detected in the focal plane, as a function of the input power.

5.3.2 *Analysis of Thermal Effect of a Fiber Coupling Lens by using Ansys*

When a light field has a high-power, the thermal-mechanic effects are not negligible. The incident electromagnetic field causes a heat flux distribution $\boldsymbol{\phi}_q(\boldsymbol{r})$ in the optical component, which then leads to its inhomogeneous temperature distribution $T(\boldsymbol{r})$. Furthermore, $T(\boldsymbol{r})$ causes the inhomogeneity of $\epsilon(\boldsymbol{r})$ and optical surface deformation.

In this section, we demonstrate a workflow to do thermal-mechanical-optic simulation via an application use case.

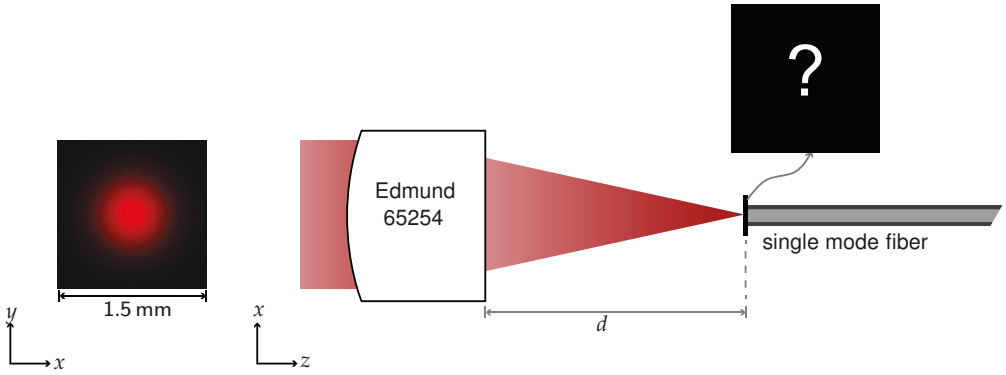


Figure 33: The Edmund 65254 lens is used to couple a fundamental Gaussian mode with the wavelength of 780 nm(left) into a single-mode fiber. d is the working distance between the lens and the fiber end. The energy density in the fiber end (position of the question mark) and the coupling efficiency will be calculated.

5.3.2.1 *Task description*

We use a commercial lens Edmund 65254 [77] to couple a fundamental Gaussian mode with a wavelength of 780 nm and a waist radius of 330 µm to a single-mode fiber with a mode field diameter of 3 µm, as shown in Fig. 33. The energy density ($|E_x|^2 + |E_y|^2 + |E_z|^2$) in the fiber end, and the coupling efficiency, which is calculated by using the overlap integral, will be calculated as shown in Fig. 33 question mark. When the Gaussian beam-induced thermal effect is negligible, we set the working distance $d = 1.585\,\text{mm}$, and get the energy density shown in Fig. 35 (a), and the coupling efficiency 88.4 %. The task now is to calculate the energy density and coupling efficiency when the Gaussian input field causes a heat flux distribution, $\boldsymbol{\phi}_q(x, y)$, in the spherical surface of the lens

$$\boldsymbol{\phi_q} = A \exp\left(2\frac{x^2 + y^2}{\omega_0^2}\right)\hat{z}, \tag{5.21}$$

with $\omega_0 = 330\,\mu\text{m}$ as the input Gaussian beam and $A = 2000\,\text{mW}\,\text{mm}^{-2}$(This value is selected to be sufficiently large, such that the thermal effect is obvious. In practice,

the value A could be much smaller, and thermal analysis is used to check the thermal tolerance).

5.3.2.2 *Thermal-Mechanical-Optic Analysis in Ansys and VirtualLab Fusion [26]*

Thermal analysis provides two optical parameters: the temperature distribution of the lens $T(\boldsymbol{r})$, and surface deformation. To do the analysis by Ansys, the following information needs to be provided: (1)The lens geometry, which is offered via a STP (STandard for the exchange of Product) file, offers a 3D graphic field used with CAD software. The file can be downloaded from the Edmund website [77]. (2) The thermal-mechanical properties of the lens material, which is a SCHOTT material N-LASF 9[78] and the parameters are listed in Tab. 6. (3) The boundary condition of the lens, i.e., the surrounding air media with a convection heat coefficient of $0.5\,\mathrm{W\,m^{-2}\,^{\circ}C^{-1}}$ and a copper holder, which is used to fix the cylindrical surface of the lens, with the convection heat coefficient of $386\,\mathrm{W\,m^{-2}\,^{\circ}C^{-1}}$. Here we assume that the holder was wrapped around the whole lens cylindrical surface, and with the holder, the lens expansion only happens in the front spherical surface and the plane interface. In a more accurate analysis, the holder model with real materials could be included in the thermal analysis [79, 80].

Table 6: Data sheet of SCHOTT material N-LASF9 [78]

N-LASF9			
Physical properties			
Density		4.41	$\mathrm{g\,cm^{-3}}$
Initial temperature	T_0	20	°C
Initial permittivity @T_0@ 780 nm	$\epsilon(T_0)$	3.362	
Thermal-mechanical coefficients			
Thermal conductivity	k	0.79	$\mathrm{W\,m^{-1}\,K^{-1}}$
Specific heat capacity	c_p	0.530	$\mathrm{J\,g^{-1}\,K^{-1}}$
Coefficient of expansion	$\alpha_{+20/+300^{\circ}C}$	8.4×10^{-6}	$\mathrm{K^{-1}}$
Young's modulus	E	1.09×10^{5}	$\mathrm{N\,mm^{-2}}$
Poisson's ratio	μ	0.288	
Constants of formula for $\mathrm{d}\epsilon/\mathrm{d}T$			
	D_0	1.05×10^{-6}	
	D_1	1.02×10^{-8}	
	D_2	-2.38×10^{-11}	
	E_0	9.19×10^{-7}	
	E_1	1.18×10^{-9}	
	λ_{TK}	0.257	µm

In Ansys, finite element analysis [44] is used to do the thermal-mechanical analysis. The related theory is not discussed here, and we take advantage of Ansys to vary the mesh density until we get a convergent result (after interpolation) for both $T(\boldsymbol{r})$ and the deformed surface.

Then, we export the data $T(\boldsymbol{r}_i)$ with i, the mesh node index, for further processing.

Temperature analysis in Ansys. Fig. 34 (a) shows the temperature (steady state) distribution $T(\boldsymbol{r})$. To visualize $T(\boldsymbol{r})$ inside the lens, the section view is displayed. The highest temperature appears at the center of the spherical surface, on which the heat flux in Eq. (5.21) is loaded. The $T(\boldsymbol{r})$ distribution is rotationally symmetric about the z-axis, so we simply export the $T(\boldsymbol{r}_i)$ data (i is the mesh node index) in one planar section, for further processing. $T(\boldsymbol{r})$ dissipates quickly in the z direction, and the temperature at the rear surface center approaches the surrounding temperature, which is 20 °C. Thus, in further analysis we are only interested in the deformation of the spherical surface.

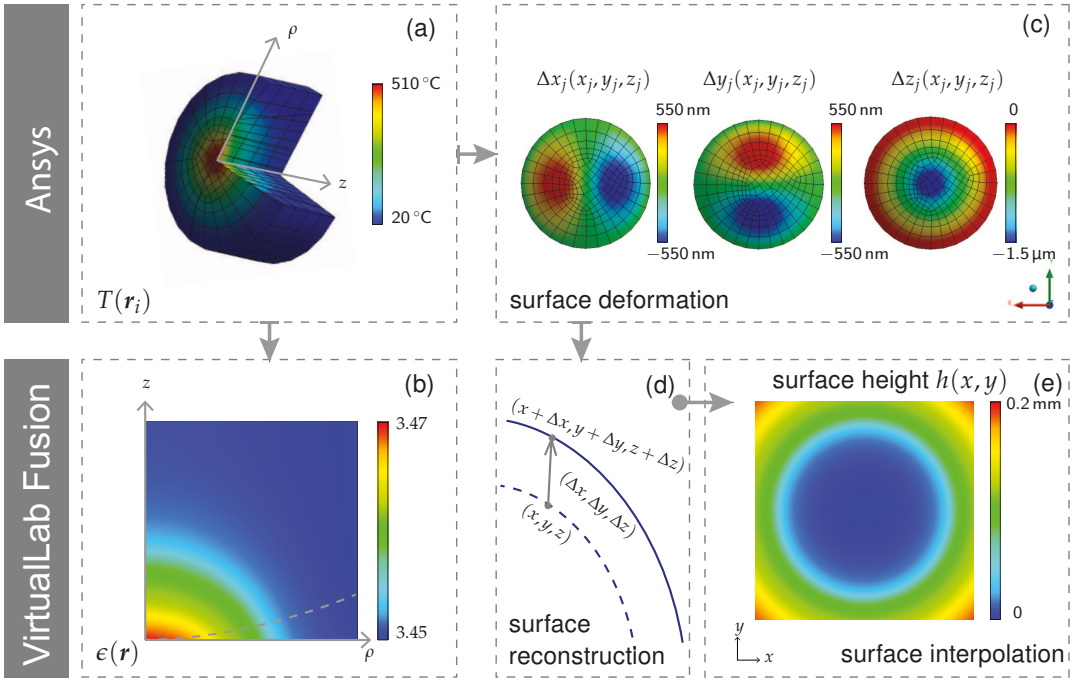


Figure 34: Thermal-mechanical-optic model. In the software Ansys, thermal-mechanical analysis is performed. (a) Temperature $T(\boldsymbol{r})$; (b) $\epsilon(\rho, z)$; (c) surface deformation $(\Delta x, \Delta y, \Delta z)$ on mesh nodes; (d) reconstructed surface; (e) interpolated surface.

Calculation of $\epsilon(\boldsymbol{r})$ in VirtualLab Fusion. After exporting $T(\boldsymbol{r}_i)$, a linear interpolation method is used to calculate $T(\boldsymbol{r})$ in the ρz plane (cylindrical coordinate system) on an equidistant grid. The relative permittivity $\epsilon(\boldsymbol{r})$ needs to be calculated using the method in [81]

$$\epsilon(\boldsymbol{r}) = \epsilon(\lambda, T_0) + \frac{\mathrm{d}\epsilon}{\mathrm{d}T}(\lambda, T)[T(\boldsymbol{r}) - T_0], \tag{5.22}$$

with

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}T}(\lambda, T) = [\epsilon(\lambda, T_0) - 1]\left(D_0 + 2D_1(T - T_0) + 3D_2(T - T_0)^2 + \frac{E_0 + 2E_1(T - T_0)}{\lambda^2 - \lambda_{\mathrm{TK}}^2}\right), \tag{5.23}$$

with parameters $T_0, D_0, D_1, D_2, E_0, E_1,$ and λ_{TK} shown in Tab. 6. $\epsilon(\boldsymbol{r})$ in the ρz plane, having a value from 3.45 to 3.47, is shown in Fig. 34 (b). The dashed curve denotes the optical surface between the lens medium and the air. The value under the dashed line is purely an interpolation result, which will be replaced by the homogeneous air permittivity later in the optical simulation. Again, here we simply show $\epsilon(\boldsymbol{r})$ in the ρz-plane, and in the optical simulation, we assume $\epsilon(\boldsymbol{r})$ is rotationally symmetric about the z-axis.

Expanded optical surface. After we have $T(\boldsymbol{r}_i)$, the surface expansion can then be calculated in Ansys. Mesh node j, with j as the node index on the spherical surface, has its original position coordinate (x_j, y_j, z_j) in the global coordinate system, the origin of which is at the center of the spherical surface. After importing $T(\boldsymbol{r}_i)$ and the structure analysis, each node j has a shift, $(\Delta x_j, \Delta y_j, \Delta z_j)$. In Fig. 34 (c), $\Delta x_j(x_j, y_j, z_j)$, $\Delta y_j(x_j, y_j, z_j)$, and $\Delta z_j(x_j, y_j, z_j)$ are shown. To reconstruct the deformed surface, the processing in Fig. 34 (d) is done in VirtualLab Fusion: The dashed curve represents the original spherical surface with coordinates (x, y, z), and the solid curve is the reconstructed surface with coordinates $(x + \Delta x, y + \Delta y, z + \Delta z)$. This is a one-to-one mapping from the dashed curve to the solid curve. Then, a spline-interpolation is performed to calculate the height profile $h(x, y)$ of the reconstructed surface, as shown in (e).

Optical simulation of the fiber coupling. By using $\epsilon(\boldsymbol{r})$ and the surface height profile $h(x, y)$ in Fig. 34 (e), the deformed lens is reconstructed. Replacing the Edmund 65254 with the reconstructed lens in Fig. 33, and at the same working distance $d = 1.585\,\mathrm{mm}$, the energy density in the fiber end and the coupling efficiency can be calculated, as shown in Fig. 32 (b), and the coupling efficiency is only 31 %. However, the energy density/coupling efficiency is not as large as in the original situation, because this is not the best working distance, and we need to find the best working distance $d = 1.549\,\mathrm{mm}$, which gives the highest energy density (Fig. 32 (c)). The coupling efficiency is 96.6 %, which is even larger than the original setup. This can be understood as the deformation giving the surface the proper aspherical shape, which focuses light into a smaller spot. The light spot is somewhat not rotationally symmetric along the z axis in (b). This is because the mesh distribution is not rotationally symmetric (Fig. 34 surface deformation), and the surface interpolation introduces asymmetry. The author also performs an optical simulation, which is not shown here, for two cases: (1) reconstructed $\epsilon(\boldsymbol{r})$+ spherical surface, and (2) homogeneous N-LASF 9 + reconstructed surface. The inhomogeneity in this simulation causes a slight effect because of the small change in $\Delta\epsilon(\boldsymbol{r}) = 0.02$. The main effect is caused by the deformation of the spherical surface.

5.4 SUMMARY

In this chapter, we presented three practical examples relevant to GRIN media.

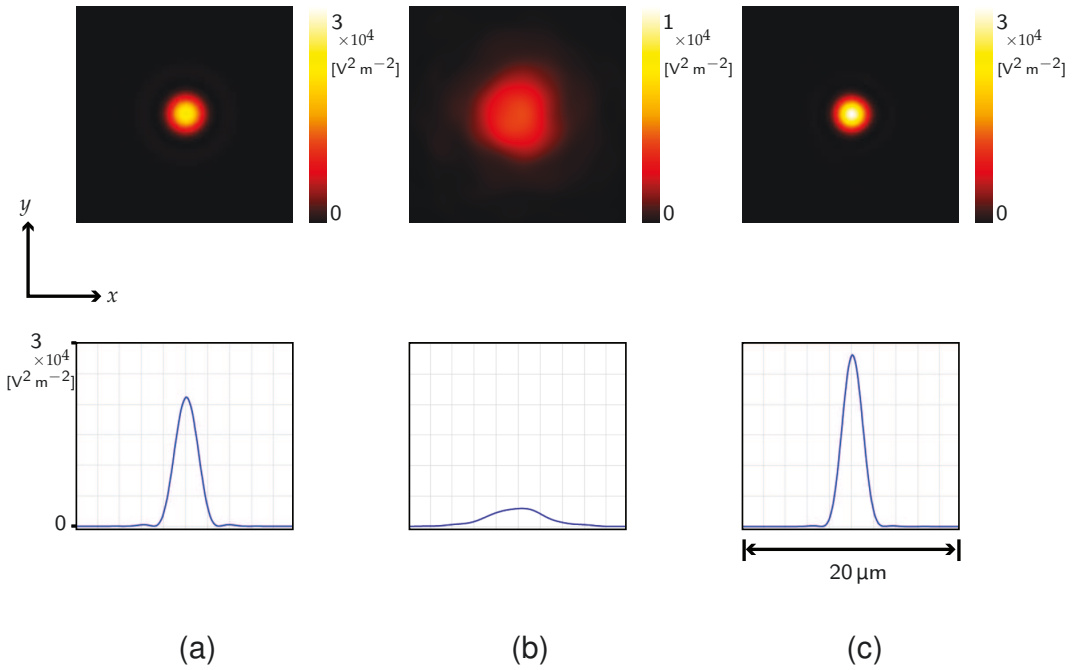


Figure 35: Optical simulation results. The upper row shows the energy density in the fiber end (question mark in Fig. 33, while the bottom row shows the energy density along the x axis. (a) Result for the original setup, without considering thermal effects. The working distance $d = 1.585\,\text{mm}$. (b) Result after considering the thermal effects at the same working distance $d = 1.585\,\text{mm}$. (c) Result after considering the thermal effects at the proper working distance $d = 1.585\,\text{mm}$.

The first example was designing and simulating generalized Luneburg lenses. The mathematical description was given by Luneburg [3], and it was analytically solved by Sochacki [5–7]. We implemented the formula and extended it to design a generalized Luneburg lens with multi-focus. Then, we simulated field tracing through the whole system, from a spherical input field to the focal plane. One shortcoming of this example is that the optical surface, which is an abrupt change in the relative permittivity between the surrounding medium and the ball lens, is not currently included. This is not a constraint of the formula derived by Luneburg; instead, the analytical solution did not include the interface term yet. However, Morgen [4] has included the interface term, and in further research we will explore whether it can be combined with the analytical solution by Sochacki.

The second example was designing a bending waveguide using conformal transformation and modeling the light propagation through it. Although conformal transformation can only design y-invariant or cylindrical symmetric components, it is sufficient for most imaging systems. In further research, we may try to combine the generalized

Luneburg lens and the conformal transformation design strategy to design a multi-focal lens with a simpler GRIN distribution, or more easily fabricated surfaces.

The third example was the simulation of the thermal-mechanic-optic effect. We either used an analytical model to simulate the thermal lens inside a laser system, or we used the software Ansys [44] to model the thermal-mechanical effect. It is important for modern optical engineering to include thermo-mechanical-optic simulation.

6 CONCLUSION AND OUTLOOK

In summary, in this thesis, we explored field solvers to simulate an electromagnetic field propagating through isotropic graded-index (GRIN) media in the framework of linear optics.

We introduced several existing solvers based on three aspects, i.e., type of input fields, spatial GRIN distribution, and capability of predicting different optical phenomena, e.g., polarization crosstalk, non-paraxial effect, and so on. Table 1 lists the properties of the introduced solvers. We mentioned three rigorous solvers capable of predicting all linear optical effects. However, Mie theory is only valid for spherically symmetric GRIN media, and the Fourier modal method (FMM) only works for periodic structures. The FMM with perfectly matched layers (FMM+PMLs) can model the general GRIN distribution. However, the FMM+PMLs can only model small optical components that are several wavelengths in size, because its numerical effort is linear to cubic of the number of sampling points, N ($\sim N^3$). We also discussed the concept and the capabilities of beam propagation methods (BPMs), including the Fourier transform (FT-) BPM, the wave propagation method (WPM), and the finite-difference (FD-) BPMs. The BPMs are easy to implement and highly extensible. However, many of them have the constraint that the variance of ϵ should be negligible, i.e., $\nabla \ln \epsilon \approx 0$, which introduces inaccuracy in calculating the polarization crosstalk.

We developed a conceptually straightforward field solver, i.e., the Runge-Kutta (RK) k-domain field propagation method (RKK-FPM), which can be used to efficiently model the general input field through arbitrary GRIN media without the constraint of small variance in $\epsilon(\boldsymbol{r})$. This method unified multiple benefits of other field solvers. Starting with Maxwell's equations, we derived an ordinary differential equation (ODE) in the spatial-frequency (k-) domain, and solved it for an initial field iteratively by using the RK method. Taking advantage of the convolution theorem, the total numerical effort was linear at N ($\sim N$). Non-paraxial fields, as well as the polarization crosstalk generated by the GRIN media, can be solved efficiently. However, the RKK-FPM does not take the optical surface effect (between the surrounding and the GRIN media) into account. Furthermore, the reflection or evanescent wave is currently not included in this method. As a result, if the optical surface effect is not negligible, one can use the WPM ($\sim N^2$) as the field solver, and if the reflection/evanescent wave is important, we suggest using one of the rigorous field solvers, e.g., Mie, FMM, or FMM+PMLs, to tackle it.

We developed another field solver, known as the RK x-domain field propagation method (RKX-FPM), to model a geometric field having a diffraction that is negligible when propagating through GRIN media. As most GRIN lenses and GRIN beam-shaping elements are still modeled or designed using ray tracing techniques, the geometric field assumption is valid in these cases. By using the geometric field ansatz, and replacing the

Fourier transforms with the transforms for the geometric field, i.e., the homeomorphic Fourier transform (HFT) [33], in the ODE of the RKK-FPM, we arrived at the fundamental theory of ray optics, the eikonal equation. After substituting the geometric field ansatz and eikonal equation into Maxwell's equations, we obtained the geometric field equations. The eikonal equation and geometric field equations were further derived, specifically for GRIN media, to two ODEs in the x-domain. Although the derivation process was different, Born and Wolf gave the same ODEs in [53]. Thus, the ODEs were solved by the RK method, together with the energy conservation law, and the RKX-FPM was developed. Furthermore, the RKX-FPM always included optical surface effects by using local plane interface approximation (LPIA) [29]. One remarkable advantage of the RKX-FPM is that it can be as fast as ray tracing techniques, which makes it the perfect choice when the scale of a GRIN component is well beyond its wavelength, such as air turbulence.

After we obtained both RK methods, a Luneburg lens was modeled by combining the RKK- and RKX-FPMs. The RKX-FPM was used to calculate the field propagation in the geometric zone and the RKK-FPM to simulate the other parts.

By using the field solvers, we calculated the electromagnetic field propagation through an optical system with GRIN components. We also discussed two analytical design methods for GRIN components: (1) The generalized Luneburg lens's analytical design (the formula was derived by Luneburg [3], and solved analytically by Sochacki [5–7]) was discussed and extended to design a multi-focus ball lens; (2) The analytical design of a bending waveguide using conformal transformation was introduced. Afterwards, we calculated electromagnetic field propagation through an optical system with the designed GRIN components.

Finally, we developed a workflow to model the thermal-mechanical-optic effects using the software Ansys and VirtualLab Fusion. In Ansys, we modeled the inhomogeneous temperature distribution and surface deformation, and in VirtualLab Fusion we interpolated and transferred the data to relative permittivity, and reconstructed the surface to simulate electromagnetic field propagation through the thermal lens system.

In future research, we will explore how to include optical surface effects, reflection, and evanescent waves in the RKK-FPM. Meanwhile, the numerical effort can be further reduced by using analytical phase processing of, e.g., linear/quadratic phase. Afterward, we will extend both RK-FPMs to model anisotropic inhomogeneous media, which is explored in some pioneer's work [63, 82–84]. We will extend the design to include the optical surface between the surrounding and the generalized Luneburg lens, and transform it to a cylindrically symmetric lens using the conformal transformation.

A

FABRICATION OF GRADED-INDEX MEDIUM

Many literatures and patents from 1970s to nowadays present several methods of fabricating graded-index(GRIN) medium [85, 86]. The most important properties of the fabrication method is: (1) Which kind of materials are used, glass, polymer or ceramic; (2) How flexible is the graded-index direction, one-dimensional graded, rotational symmetric, or flexible? (3) What is the scale of the component, in µm, mm or even larger? (4) The possible range of refractive index Δn, with n the refractive index ($n = \sqrt{(\epsilon)}$). The typical value of (3) and (4) is found via authors summary or in the manufactured examples, and they might have slight difference in different literatures. Recently there are many discussion about GRIN medium in infrared applicaiton, which we don't dive into details [87]. Here we list the frequently used methods as follows.

- **(Modified) Chemical vapor deposition (MCVD)** is widely used to fabricate GRIN fiber.[88–90]. It is possible to obtain $\Delta n \approx 0.1$ at a radial distance in the scale of 0.2 mm. It can also be used to fabricate one-dimensional GRIN glass, because CVD can be used to fabricate thin films.

- **(Double) ion exchange** is used to fabricate the GRIN rod (radially symmetric index profile) or spherical GRIN lens [91–94]. It is almost the most widely used for commercial production of GRIN glass [15]. The typical radial distance is in the scale of 100 mm, while $\Delta n \approx 0.05$.

- **Sol-gel method** [95–97] introduces GRIN distribution into specific sol-gel and then sintered into a GRIN glass rod (radially symmetric). The typical radial distance is in the scale of 3 mm, while $\Delta n \approx 0.05$.

- **Three-dimensional(3D) printing process**, either using silica powder[98], or from sol-gel ink[99]. The former one fabricates rotationally symmetric GRIN glass or ceramic with diameter 4 mm with $\Delta n \approx 0.002$. The latter one shows flexibility of GRIN direction, at least two dimensional gradient can be produced (no examples of axial gradient). The diameter of the fabricated GRIN glass is 10 mm, with $\Delta n \approx 0.01$.

- **Lithography direct-write technology(two-photon polymerization)** fabricates GRIN medium with arbitrary index distribution, and meanwhile the micro-topological structure. [100] The size of the element is approximately 50 µm, and $\Delta n \approx 0.01$. However, in this method, we didn't found the resolution of the fabrication process, which should be an important parameter.

- **Nanolayer polymer extrusion technology** generates flexible GRIN profile (polymer) with scale of several mm and large range of refractive index $\Delta n \approx 0.17$ [101, 102].
- **Metamaterials** are widely explored to achieve equivalent GRIN [103]. Especially when the conformal transformation optics [21] driven GRIN medium has special refractive index near zero [104].

There are also some other techniques to manufacture GRIN glass, e.g., neutron irradiation[105], ion stuffing [106] and drop-on-demand technology [107].

B

FIELD SOLVER IN SPHERICAL COORDINATE SYSTEM

Let us solve field equation in spherical coordinate system.

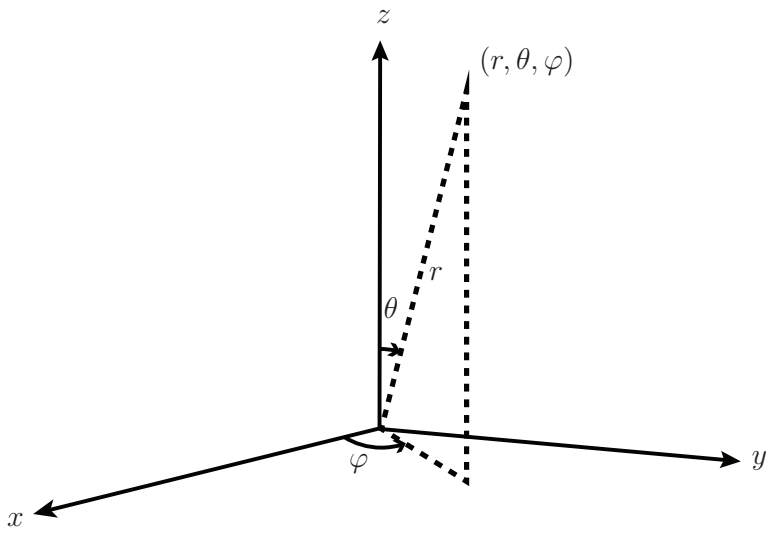


Figure 36: Specifying spherical coordinate system in Cartesian coordinate system.

B.1 FIELD REPRESENTATION IN SPHERICAL COORDINATE SYSTEM

In the spherical coordinate system, electric field is represented as

$$\boldsymbol{E}(\boldsymbol{r}) = \begin{bmatrix} E_r \\ E_\theta \\ E_\varphi \end{bmatrix} (\boldsymbol{r}), \tag{B.1}$$

with $\boldsymbol{r} = (r, \theta, \varphi)$, and E_r, E_θ, E_φ denoting field components along radial, inclination angle, and azimuthal angle direction.

Substituting the representation into Helmholtz equation (2.24), we obtain a very complex equations, because all E_r, E_θ and E_φ mixed up, because the Laplace operation to a vector B.1 is written as

$$\Delta \boldsymbol{E}(\boldsymbol{r}) = \begin{bmatrix} \Delta E_r + \dfrac{2E_r}{r^2} - \dfrac{2}{r^2\sin\theta}\partial_\theta\left(E_\theta \sin\theta\right) - \dfrac{2}{r^2\sin\theta}\partial_\varphi E_\varphi \\ \Delta E_\theta - \dfrac{E_\theta}{r^2\sin^2\theta} + \dfrac{2}{r^2}\partial_\theta E_r - \dfrac{2\cos\theta}{r^2\sin^2\theta}\partial_\varphi E_\varphi \\ \Delta E_\varphi - \dfrac{E_\varphi}{r^2\sin\theta} + \dfrac{2}{r^2\sin\theta}\partial_\varphi E_r + \dfrac{2\cos\theta}{r^2\sin^2\theta}\partial_\varphi E_\theta \end{bmatrix} \tag{B.2}$$

but if we represent the field in Cartesian coordinate system

$$\boldsymbol{E}(\boldsymbol{r}) = \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}(\boldsymbol{r}), \tag{B.3}$$

Eq. (2.24) becomes separable for E_x, E_y and E_z and we could concentrate on solving the following equation

$$\Delta\psi(\boldsymbol{r}) + k_0^2\epsilon\psi(\boldsymbol{r}) = 0, \tag{B.4}$$

with $\psi(\boldsymbol{r})$ denoting E_x, E_y or E_z.

Representing the Laplace operator in spherical coordinate system to be

$$\Delta = \frac{1}{r^2}\partial_r(r^2\partial_r) + \frac{1}{r^2\sin^2\theta}\partial_\varphi^2 + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta), \tag{B.5}$$

and separate variables as

$$\psi(\boldsymbol{r}) = R(r)\Theta(\theta)\Phi(\varphi), \tag{B.6}$$

we obtain

$$\begin{aligned} &\frac{1}{r^2}\partial_r[r^2\partial_r R(r)]\Theta(\theta)\Phi(\varphi) + \frac{1}{r^2\sin^2\theta}\partial_\varphi^2\Phi(\varphi)R(r)\Theta(\theta) + \\ &\frac{1}{r^2\sin\theta}\partial_\theta[\sin\theta\partial_\theta\Theta(\theta)]R(r)\Phi(\varphi) + k_0^2\epsilon R(r)\Theta(\theta)\Phi(\varphi) = 0. \end{aligned} \tag{B.7}$$

Divided Eq. (B.7) by $\frac{1}{r^2\sin^2\theta}R(r)\Theta(\theta)\Phi(\varphi)$, and shifting the term related to φ to the right hand side, we get

$$\frac{1}{R(r)}\sin^2\theta\partial_r[r^2\partial_r R(r)] + \frac{1}{\Theta(\theta)}\sin\theta\partial_\theta[\sin\theta\partial_\theta\Theta(\theta)] + k_0^2\epsilon r^2\sin^2\theta = -\frac{1}{\Phi(\varphi)}\partial_\varphi^2\Phi(\varphi). \tag{B.8}$$

If Eq. (B.8) is valid, the left hand side and right hand side must equal to a constant, here we assume m^2. The term with φ can be isolated and written as

$$\partial_\varphi^2\Phi(\varphi) + m^2\Phi(\varphi) = 0, \tag{B.9}$$

and the solution is

$$\Phi(\varphi) = A\cos(m\varphi) + B\sin(m\varphi). \tag{B.10}$$

To get a single value $\Phi(\varphi)$ ($\Phi(\varphi + 2\pi) = \Phi(\varphi)$), m is an integer. Conventionally, the solution is denoted as

$$\Phi_e(\varphi) = \cos(m\varphi),\ \Phi_o(\varphi) = \sin(m\varphi), \tag{B.11}$$

with e and o represent even and odd. Similarly, dividing the left hand side of Eq. (B.8) by $\sin^2\theta$ and shift the terms related to θ to the right hand side, we get

$$\frac{1}{R(r)}\partial_r[r^2\partial_r R(r)] + k_0^2\epsilon r^2 = -\frac{1}{\Theta(\theta)\sin\theta}\partial_\theta[\sin\theta\partial_\theta\Theta(\theta)] + \frac{m^2}{\sin^2\theta} \tag{B.12}$$

Similar argument, if both side equal, then it should be a constant, here we assume $n(n+1)$, so we obtain two separated equations

$$\partial_r[r^2\partial_r R(r)] + [k_0^2\epsilon r^2 - n(n+1)]R(r) = 0, \tag{B.13}$$

and

$$\frac{1}{\sin\theta}\partial_\theta[\sin\theta\partial_\theta\Theta(\theta)] + \left[n(n+1) - \frac{m^2}{\sin^2\theta}\right]\Theta(\theta) = 0. \tag{B.14}$$

Eq. (B.14) shows exactly identical form to the associated Legendre equation[47]

$$(1-x^2)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}y}{\mathrm{d}x} + \left[n(n+1) - \frac{m^2}{1-x^2}\right]y = 0, \tag{B.15}$$

when $x = \cos\theta$. The mathematical solution to Eq. (B.15) is a linear combination of the first and second kind of Legendre polynomial. However, we are only interested in the converge solution for $x \subset [-1, 1]$, so the solution to Eq. (B.14) is the first kind Legendre polynomial $P_n^m(\cos\theta)$ (the second kind Legendre polynomial is only converge when $|x| < 1$).

Similarly, Eq. (B.13) shows exactly identical form to the equation of the Bessel function

$$x^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + x\frac{\mathrm{d}y}{\mathrm{d}x} + [x^2 - (n+0.5)^2]y = 0 \tag{B.16}$$

when $x = k_0\sqrt{\epsilon}r$ and $y = \sqrt{r}R(r)$. The mathematical solution to Eq. (B.16) is the linear combination of Bessel functions of first kind $J_{n+1/2}(x)$ and second kind $Y_{n+1/2}(x)$. The function $r^{1/2}J_{n+1/2}(r)$ and $r^{1/2}Y_{n+1/2}(r)$, can be normalized and written as spherical Bessel functions of the first and second kind, respectively, as follows

$$\begin{aligned} j_n(x) &= \sqrt{\frac{\pi}{2x}}J_{n+1/2}(x), \\ y_n(x) &= \sqrt{\frac{\pi}{2x}}Y_{n+1/2}(x). \end{aligned} \tag{B.17}$$

Later we define $z_n(x)$ to represent either $j_n(x)$ or $y_n(x)$ or the combination of $j_n(x)$ and $y_n(x)$, which is caud spherical Hankel functions

$$\begin{aligned} h_n^{(1)}(x) &= j_n(x) + iy_n(x) \\ h_n^{(2)}(x) &= j_n(x) - iy_n(x) \end{aligned} \tag{B.18}$$

To combine all solutions together we obtain

$$\begin{aligned} \psi_{emn}(\boldsymbol{r}) &= \cos(m\varphi)P_n^m(\cos\theta)z_n(k_0\sqrt{\epsilon}r), \\ \psi_{omn}(\boldsymbol{r}) &= \sin(m\varphi)P_n^m(\cos\theta)z_n(k_0\sqrt{\epsilon}r). \end{aligned} \tag{B.19}$$

B.2 FIELD SOLVER FOR A SINGLE SPHERE

Till now, we might think about E_x, E_y and E_z is linear combination of $\psi_{emn}(\boldsymbol{r})$ and $\psi_{omn}(\boldsymbol{r})$. And try to match the boundary condition to calculate the coefficients of each order. However, conventionally, vector spherical harmonics are defined to simplify the progress

$$\begin{aligned} \boldsymbol{M}(\boldsymbol{r}) &= \boldsymbol{\nabla} \times [\boldsymbol{c}\psi(\boldsymbol{r})], \\ \boldsymbol{N}(\boldsymbol{r}) &= \frac{1}{k_0\sqrt{\epsilon}}\boldsymbol{\nabla} \times \boldsymbol{M}(\boldsymbol{r}). \end{aligned} \tag{B.20}$$

with $\boldsymbol{c}$ denoting a constant vector, $\psi(\boldsymbol{r})$ denoting a scalar function. Now we ask the question, if $\boldsymbol{M}(\boldsymbol{r})$ and $\boldsymbol{N}(\boldsymbol{r})$ can be used to represent electromagnetic field. First, $\boldsymbol{M}(\boldsymbol{r})$ and $\boldsymbol{N}(\boldsymbol{r})$ are divergent-free, which follows Eqs. (2.22) and (2.23) in homogeneous medium $\epsilon = \text{const}$. Second, $\boldsymbol{M}(\boldsymbol{r})$ and $\boldsymbol{N}(\boldsymbol{r})$ fulfil the wave equations (2.24) and (2.25),

$$\begin{aligned} \Delta\boldsymbol{M}(\boldsymbol{r}) + k_0^2\epsilon\boldsymbol{M}(\boldsymbol{r}) &= \boldsymbol{\nabla} \times \left\{\boldsymbol{c}\left[\Delta\psi(\boldsymbol{r}) + k_0^2\epsilon\psi(\boldsymbol{r})\right]\right\} = 0, \\ \Delta\boldsymbol{N}(\boldsymbol{r}) + k_0^2\epsilon\boldsymbol{N}(\boldsymbol{r}) &= \frac{1}{k_0\sqrt{\epsilon}}\boldsymbol{\nabla} \times \left[\Delta\boldsymbol{M}(\boldsymbol{r}) + k_0^2\epsilon\boldsymbol{M}(\boldsymbol{r})\right] = 0. \end{aligned} \tag{B.21}$$

with $\boldsymbol{c} = \boldsymbol{r}$ and the scalar function $\psi(\boldsymbol{r})$ follows Eq. (B.4) which can be represented as Eq. (B.19). Follows the convention, let's define the quantities as follow

$$\begin{aligned} \boldsymbol{M}_{emn}(\boldsymbol{r}) &= \boldsymbol{\nabla} \times [\boldsymbol{r}\psi_{emn}(\boldsymbol{r})], \\ \boldsymbol{M}_{omn}(\boldsymbol{r}) &= \boldsymbol{\nabla} \times [\boldsymbol{r}\psi_{omn}(\boldsymbol{r})], \\ \boldsymbol{N}_{emn}(\boldsymbol{r}) &= \frac{1}{k_0\sqrt{\epsilon}}\boldsymbol{\nabla} \times \boldsymbol{M}_{emn}(\boldsymbol{r}), \\ \boldsymbol{N}_{omn}(\boldsymbol{r}) &= \frac{1}{k_0\sqrt{\epsilon}}\boldsymbol{\nabla} \times \boldsymbol{M}_{omn}(\boldsymbol{r}) \end{aligned} \tag{B.22}$$

One should know that $\boldsymbol{M}_{emn}(\boldsymbol{r})$, $\boldsymbol{M}_{omn}(\boldsymbol{r})$, $\boldsymbol{N}_{emn}(\boldsymbol{r})$ and $\boldsymbol{N}_{omn}(\boldsymbol{r})$ are mutually orthogonal, which means

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \,\mathrm{d}\theta A_{m'n'}B_{mn} = 0 \tag{B.23}$$

when $m' \neq m$ or $n' \neq n$, and here A and B denotes arbitrary two quantities of $\boldsymbol{M}_e(\boldsymbol{r})$, $\boldsymbol{M}_o(\boldsymbol{r})$, $\boldsymbol{N}_e(\boldsymbol{r})$ and $\boldsymbol{N}_o(\boldsymbol{r})$.

Till now, we found the basis functions $\boldsymbol{M}_{emn}(\boldsymbol{r})$, $\boldsymbol{M}_{omn}(\boldsymbol{r})$, $\boldsymbol{N}_{emn}(\boldsymbol{r})$ and $\boldsymbol{N}_{omn}(\boldsymbol{r})$ which are complete and orthogonal. Electric field can be represented as

$$\boldsymbol{E}(\boldsymbol{r}) = \sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\left[a_{emn}\boldsymbol{M}_{emn}(\boldsymbol{r}) + a_{omn}\boldsymbol{M}_{omn}(\boldsymbol{r}) + b_{emn}\boldsymbol{N}_{emn}(\boldsymbol{r}) + b_{omn}\boldsymbol{N}_{omn}(\boldsymbol{r})\right] \tag{B.24}$$

The coefficients can be calculated by

$$a_{emn} = \frac{\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi}\sin\theta\,\mathrm{d}\theta\boldsymbol{E}(\boldsymbol{r})\cdot\boldsymbol{M}_{emn}(\boldsymbol{r})}{\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi}\sin\theta\,\mathrm{d}\theta\left|\boldsymbol{M}_{emn}(\boldsymbol{r})\right|^2} \tag{B.25}$$

which is similar for a_{omn}, b_{emn} and b_{omn}

Now let's consider a task as follows: an arbitrary incident field propagates to a sphere, assume

$$\epsilon = \begin{cases} \epsilon_1, & r > r_0 \\ \epsilon_2, & r \leqslant r_0 \end{cases} \tag{B.26}$$

The incident electric field can be represented as

$$\boldsymbol{E}^{\text{inc}} = \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \left[a_{emn}^{\text{inc}} \boldsymbol{M}_{emn}^{\epsilon_1}(\boldsymbol{r}) + a_{omn}^{\text{inc}} \boldsymbol{M}_{omn}^{\epsilon_1}(\boldsymbol{r}) + b_{emn}^{\text{inc}} \boldsymbol{N}_{emn}^{\epsilon_1}(\boldsymbol{r}) + b_{omn}^{\text{inc}} \boldsymbol{N}_{omn}^{\epsilon_1}(\boldsymbol{r}) \right] \tag{B.27}$$

The reflected field, which is outside of the sphere, can be named as scattered electric field

$$\boldsymbol{E}^{\text{sca}} = \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \left[a_{emn}^{\text{sca}} \boldsymbol{M}_{emn}^{\epsilon_1}(\boldsymbol{r}) + a_{omn}^{\text{sca}} \boldsymbol{M}_{omn}^{\epsilon_1}(\boldsymbol{r}) + b_{emn}^{\text{sca}} \boldsymbol{N}_{emn}^{\epsilon_1}(\boldsymbol{r}) + b_{omn}^{\text{sca}} \boldsymbol{N}_{omn}^{\epsilon_1}(\boldsymbol{r}) \right] \tag{B.28}$$

and the electric field inside the sphere, which can be named as the internal electric field

$$\boldsymbol{E}^{\text{int}} = \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \left[a_{emn}^{\text{int}} \boldsymbol{M}_{emn}^{\epsilon_2}(\boldsymbol{r}) + a_{omn}^{\text{int}} \boldsymbol{M}_{omn}^{\epsilon_2}(\boldsymbol{r}) + b_{emn}^{\text{int}} \boldsymbol{N}_{emn}^{\epsilon_2}(\boldsymbol{r}) + b_{omn}^{\text{int}} \boldsymbol{N}_{omn}^{\epsilon_2}(\boldsymbol{r}) \right] \tag{B.29}$$

Then we could calculate the related magnetic fields based on Eqs. (2.20, B.21, B.22)

$$\boldsymbol{H}^{\text{inc}} = c_1 \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \left[a_{emn}^{\text{inc}} \boldsymbol{N}_{emn}^{\epsilon_1}(\boldsymbol{r}) + a_{omn}^{\text{inc}} \boldsymbol{N}_{omn}^{\epsilon_1}(\boldsymbol{r}) + b_{emn}^{\text{inc}} \boldsymbol{M}_{emn}^{\epsilon_1}(\boldsymbol{r}) + b_{omn}^{\text{inc}} \boldsymbol{M}_{omn}^{\epsilon_1}(\boldsymbol{r}) \right], \tag{B.30}$$

$$\boldsymbol{H}^{\text{sca}} = c_1 \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \left[a_{emn}^{\text{sca}} \boldsymbol{N}_{emn}^{\epsilon_1}(\boldsymbol{r}) + a_{omn}^{\text{sca}} \boldsymbol{N}_{omn}^{\epsilon_1}(\boldsymbol{r}) + b_{emn}^{\text{sca}} \boldsymbol{M}_{emn}^{\epsilon_1}(\boldsymbol{r}) + b_{omn}^{\text{sca}} \boldsymbol{M}_{omn}^{\epsilon_1}(\boldsymbol{r}) \right], \tag{B.31}$$

$$\boldsymbol{H}^{\text{int}} = c_2 \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \left[a_{emn}^{\text{int}} \boldsymbol{N}_{emn}^{\epsilon_2}(\boldsymbol{r}) + a_{omn}^{\text{int}} \boldsymbol{N}_{omn}^{\epsilon_2}(\boldsymbol{r}) + b_{emn}^{\text{int}} \boldsymbol{M}_{emn}^{\epsilon_2}(\boldsymbol{r}) + b_{omn}^{\text{int}} \boldsymbol{M}_{omn}^{\epsilon_2}(\boldsymbol{r}) \right]. \tag{B.32}$$

with $c_i = -\mathrm{i}\sqrt{\frac{\epsilon_0 \epsilon_i}{\mu_0}}$, $i = 1, 2$.

Now we could introduce the boundary condition at the sphere boundary $r = r_0$

$$\begin{aligned} \left[\boldsymbol{E}^{\text{inc}}(r_0) + \boldsymbol{E}^{\text{sca}}(r_0) - \boldsymbol{E}^{\text{int}}(r_0) \right] \times \hat{\boldsymbol{e}}_r = 0 \\ \left[\boldsymbol{H}^{\text{inc}}(r_0) + \boldsymbol{H}^{\text{sca}}(r_0) - \boldsymbol{H}^{\text{int}}(r_0) \right] \times \hat{\boldsymbol{e}}_r = 0 \end{aligned} \tag{B.33}$$

In principle, after solving the boundary equations (B.33), the coefficients of scatter field and internal field can be calculated, especially, only the spherical Bessel functions will be used in the boundary equations. However, conventionally, the derivation is simplified by assuming an x-polarized ideal plane wave as incident field

$$\begin{aligned} \boldsymbol{E}^{\text{inc}}(\boldsymbol{r} = x, y, z) &= E_0 \exp(\mathrm{i} k_0 \sqrt{\epsilon_1} z) \hat{\boldsymbol{e}}_x, \\ \boldsymbol{E}^{\text{inc}}(\boldsymbol{r} = r, \theta, \varphi) &= E_0 \exp(\mathrm{i} k_0 \sqrt{\epsilon_1} r \cos\theta)(\sin\theta \cos\varphi \hat{\boldsymbol{e}}_r + \cos\theta \cos\varphi \hat{\boldsymbol{e}}_\theta - \sin\varphi \hat{\boldsymbol{e}}_\varphi) \end{aligned} \tag{B.34}$$

a_{emn}^{inc}, a_{omn}^{inc}, b_{emn}^{inc} and b_{omn}^{inc} for the plane wave (B.34) is quite simple. Because of the orthogonality of $\cos m\varphi$ and $\sin m\varphi$, if we decompose the plane wave into the basis functions,

the coefficients for terms $m \neq 1$ are zero. Meanwhile, because of y_n in Eq. (B.17) is infinite large when $x = 0$ which is not properly describe the plane wave, we just use the first kind of Bessel function j_n here and add a notation $^{(1)}$ to denote the first kind Bessel function. Then by using Eq. (B.25), we can get

$$\boldsymbol{E}^{\text{inc}}(r,\theta,\varphi) = E_0 \sum_{n=1}^{\infty} \mathrm{i}^n \frac{2n+1}{n(n+1)} (\boldsymbol{M}_{o1n}^{\epsilon_1,(1)} - \mathrm{i}\boldsymbol{N}_{e1n}^{\epsilon_1,(1)}). \tag{B.35}$$

Similarly, the internal field also just use first kind spherical Bessel function, so the internal electric field can be simplified as

$$\boldsymbol{E}^{\text{int}}(r,\theta,\varphi) = E_0 \sum_{n=1}^{\infty} \mathrm{i}^n \frac{2n+1}{n(n+1)} (c_n \boldsymbol{M}_{o1n}^{\epsilon_2,(1)} - \mathrm{i}d_n \boldsymbol{N}_{e1n}^{\epsilon_2,(1)}) \tag{B.36}$$

To avoid the infinity value in the position $r = \infty$, the spherical Bessel function of scatter field is chosen as $h_n^{(1)}$, which is called third kind of Bessel function.

$$\boldsymbol{E}^{\text{sca}}(r,\theta,\varphi) = E_0 \sum_{n=1}^{\infty} \mathrm{i}^n \frac{2n+1}{n(n+1)} (a_n \boldsymbol{M}_{o1n}^{\epsilon_1,(3)} - \mathrm{i}b_n \boldsymbol{N}_{e1n}^{\epsilon_1,(3)}) \tag{B.37}$$

The related magnetic fields can be calculated based on Eqs. (B.30-B.32). After using the boundary condition (B.33) coefficients a_n, b_n, c_n and d_n can be calculated, which is a function of $\sin\varphi$, $\cos\varphi$, $P_n^1(\cos\theta)$, $j_n(k_0\sqrt{\epsilon_i}r)$ and $h_n^{(1)}(k_0\sqrt{\epsilon_i}r)$. We don't give further mathematical discussion, but directly the conclusion formula for code implementation. The coefficients a_n, b_n, c_n and d_n can be represented in form of

$$\begin{aligned}
a_n &= \frac{m\psi_n(\rho)\,\mathrm{d}_{\rho'}\psi_n(\rho') - \psi_n(\rho')\,\mathrm{d}_{\rho}\psi_n(\rho)}{m\zeta_n(\rho)\,\mathrm{d}_{\rho'}\psi_n(\rho') - \psi_n(\rho')\,\mathrm{d}_{\rho}\zeta_n(\rho)} \\
b_n &= \frac{\psi_n(\rho)\,\mathrm{d}_{\rho'}\psi_n(\rho') - m\psi_n(\rho')\,\mathrm{d}_{\rho}\psi_n(\rho)}{\zeta_n(\rho)\,\mathrm{d}_{\rho'}\psi_n(\rho') - m\psi_n(\rho')\,\mathrm{d}_{\rho}\zeta(\rho)} \\
c_n &= \frac{\psi_n(\rho)\,\mathrm{d}_{\rho}\zeta_n(\rho) - \zeta_n(\rho)\,\mathrm{d}_{\rho}\psi_n(\rho)}{\psi_n(\rho')\,\mathrm{d}_{\rho}\zeta(\rho) - \zeta_n(\rho)\,\mathrm{d}_{\rho'}\psi(\rho')} \\
d_n &= \frac{m\psi_n(\rho)\,\mathrm{d}_{\rho}\zeta_n(\rho) - m\zeta_n(\rho)\,\mathrm{d}_{\rho}\psi_n(\rho)}{m\psi_n(\rho')\,\mathrm{d}_{\rho}\zeta(\rho) - \zeta_n(\rho)\,\mathrm{d}_{\rho'}\psi_n(\rho')}
\end{aligned} \tag{B.38}$$

with $\rho = k_0\sqrt{\epsilon_1}r$, $\rho' = k_0\sqrt{\epsilon_2}r$ and $m = \sqrt{\dfrac{\epsilon_2}{\epsilon_1}}$ and the Riccati-Bessel functions are

$$\begin{aligned}
\psi_n(\rho) &= \rho j_n(\rho) \\
\zeta_n(\rho) &= \rho h^{(1)}(\rho)
\end{aligned} \tag{B.39}$$

Table 7: Vector spherical harmonics

	$\hat{\boldsymbol{e}}_r$	$\hat{\boldsymbol{e}}_\theta$	$\hat{\boldsymbol{e}}_\varphi$
$\boldsymbol{M}_{emn}$	0	$-\frac{m}{\sin\theta}\sin m\varphi P_n^m(\cos\theta)z_n(\rho)$	$-\cos m\varphi\frac{\mathrm{d}P_n^m(\cos\theta)}{\mathrm{d}\theta}z_n(\rho)$
$\boldsymbol{M}_{omn}$	0	$\frac{m}{\sin\theta}\cos m\varphi P_n^m(\cos\theta)z_n(\rho)$	$-\sin m\varphi\frac{\mathrm{d}P_n^m(\cos\theta)}{\mathrm{d}\theta}z_n(\rho)$
$\boldsymbol{N}_{emn}$	$\frac{z_n(\rho)}{\rho}n(n+1)\cos m\varphi P_n^m(\cos\theta)$	$\cos m\varphi\frac{\mathrm{d}P_n^m(\cos\theta)}{\mathrm{d}\theta}\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}[\rho z_n(\rho)]$	$-m\sin m\varphi\frac{P_n^m(\cos\theta)}{\sin\theta}\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}[\rho z_n(\rho)]$
$\boldsymbol{N}_{omn}$	$\frac{z_n(\rho)}{\rho}n(n+1)\cos m\varphi P_n^m(\cos\theta)$	$\cos m\varphi\frac{\mathrm{d}P_n^m(\cos\theta)}{\mathrm{d}\theta}\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}[\rho z_n(\rho)]$	$-m\sin m\varphi\frac{P_n^m(\cos\theta)}{\sin\theta}\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}[\rho z_n(\rho)]$

with $\rho = k_0\sqrt{\epsilon}r$, and $z_n(\rho)$ can be the first kind Bessel function $j_n(\rho)$, the second kind Bessel function $y_n(\rho)$ or the third kind Bessel function $h_n^{(1)}(\rho)$ or $h_n^{(2)}(\rho)$.

B.3 FIELD SOLVER FOR A MULTILAYERED SPHERE

This section is more interesting for us, because a racially graded index media (rotational symmetric) can be discreted as a multilayered sphere, as shown in Fig. 37. And we could use similar method as section B.2.

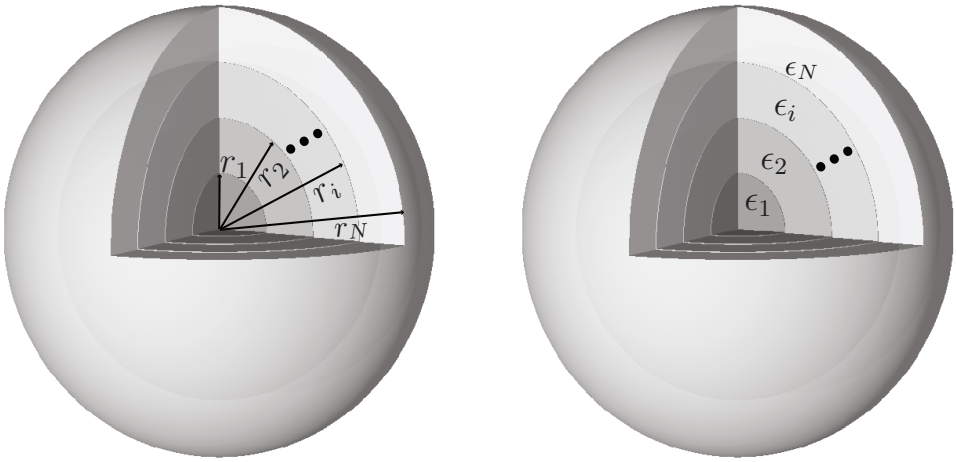


Figure 37: Configuration of a multilayered sphere.

We still assume a x-polarized ideal plane wave as incident field, as in Eq. B.35. Therefore, we still just consider spherical harmonics $\boldsymbol{M}$ and $\boldsymbol{N}$ for $m = 1$. The general field representation in each shell is

$$\boldsymbol{E}_l(r,\theta,\varphi) = E_0 \sum_{n=1}^{\infty} \mathrm{i}^n \frac{2n+1}{n(n+1)} (c_{l,n}\boldsymbol{M}_{o1n}^{\epsilon_l,(1)} - \mathrm{i}d_{l,n}\boldsymbol{N}_{e1n}^{\epsilon_l,(1)} + a_{l,n}\boldsymbol{M}_{o1n}^{\epsilon_l,(3)} - \mathrm{i}b_{l,n}\boldsymbol{N}_{e1n}^{\epsilon_l,(3)}) \quad \text{(B.40)}$$

with $l = 1, 2, \cdots, N+1$. The special cases are within $r <= r_1$ and $r > r_N$:

- When $r \leqslant r_1$, $a_{l,n} = b_{l,n} = 0$ to avoid the infinity value caused by $y_n(r = 0)$;
- When $r > r_N$, $c_{N+1,n} = d_{N+1,n} = 1$ which represent the input ideal plane wave in Eq. B.35. Please note, $\boldsymbol{E}_{N+1}(r,\theta,\varphi) = \boldsymbol{E}^{\mathrm{inc}}(r,\theta,\varphi) + \boldsymbol{E}^{\mathrm{sca}}(r,\theta,\varphi)$

The related magnetic field in each layer is

$$\boldsymbol{H}_l(r,\theta,\varphi) = E_0 \sum_{n=1}^{\infty} \mathrm{i}^n \frac{2n+1}{n(n+1)} (c_{l,n}\boldsymbol{N}_{o1n}^{\epsilon_l,(1)} - \mathrm{i}d_{l,n}\boldsymbol{M}_{e1n}^{\epsilon_l,(1)} + a_{l,n}\boldsymbol{N}_{o1n}^{\epsilon_l,(3)} - \mathrm{i}b_{l,n}\boldsymbol{M}_{e1n}^{\epsilon_l,(3)}) \quad \text{(B.41)}$$

Then we introduce the boundary conditions for each shell r_l

$$\begin{aligned} [\boldsymbol{E}_l(r_l) - \boldsymbol{E}_{l+1}(r_l)] \times \hat{\boldsymbol{e}}_r = 0 \\ [\boldsymbol{H}_l(r_l) - \boldsymbol{H}_{l+1}(r_l)] \times \hat{\boldsymbol{e}}_r = 0 \end{aligned} \quad \text{(B.42)}$$

with $l = 1, 2, \cdots, N$. Compare Eqs.(B.42) and Tab. 7, and using the definition of Riccati-Bessel functions Eq. (B.39), Eqs. (B.40) and (B.41) can be written as

$$
\begin{aligned}
a_{l,n}\bar{\rho}_l\zeta(\rho_l) + c_{l,n}\bar{\rho}_l\psi(\rho_l) - a_{l+1,n}\rho_l\zeta(\bar{\rho}_l) - c_{l+1,n}\rho_l\psi(\bar{\rho}_l) = 0 \\
a_{l,n}\zeta'(\rho_l) + c_{l,n}\psi'(\rho_l) - a_{l+1,n}\zeta'(\bar{\rho}_l) - c_{l+1,n}\psi'(\bar{\rho}_l) = 0 \\
b_{l,n}\bar{\rho}_l\zeta'(\rho_l) + d_{l,n}\bar{\rho}_l\psi'(\rho_l) - b_{l+1,n}\rho_l\zeta'(\bar{\rho}_l) - d_{l+1,n}\rho_l\psi'(\bar{\rho}_l) = 0 \\
b_{l,n}\zeta(\rho_l) + d_{l,n}\psi(\rho_l) - b_{l+1,n}\zeta(\bar{\rho}_l) - d_{l+1,n}\psi(\bar{\rho}_l) = 0
\end{aligned}
\tag{B.43}
$$

with $l = 1, 2, \cdots, N$, $\rho_l = k_0\sqrt{\epsilon_l}r_l$ and $\bar{\rho}_l = k_0\sqrt{\epsilon_{l+1}}r_l$. As $a_{l,n} = b_{l,n} = 0$ and $r > r_N$, $c_{N+1,n} = d_{N+1,n} = 1$, we have in total $4N$ unknowns (coefficients) and $4N$ equations, so all coefficients can be solved. I am just interested in a_{N+1} and b_{N+1} to represent the scattered field outside of the sphere, so let's further define variables

$$
\begin{aligned}
A_{l,n} &= a_{l,n}/c_{l,n} \\
B_{l,n} &= b_{l,n}/d_{l,n}
\end{aligned}
\tag{B.44}
$$

After further derivation, we could get the iterative equations

$$
\begin{aligned}
A_{l+1,n} &= -R_n(\bar{\rho}_l)\frac{H_n^a(A_{l,n},\rho_l)\rho_l - D_n^{(1)}(\bar{\rho}_l)\bar{\rho}_l}{H_n^a(A_{l,n},\rho_l)\rho_l - D_n^{(3)}(\bar{\rho}_l)\bar{\rho}_l}, \\
B_{l+1,n} &= -R_n(\bar{\rho}_l)\frac{H_n^b(B_{l,n},\rho_l)\bar{\rho}_l - D_n^{(1)}(\bar{\rho}_l)\rho_l}{H_n^b(B_{l,n},\rho_l)\bar{\rho}_l - D_n^{(3)}(\bar{\rho}_l)\rho_l}, \\
H_n^a(A_{l,n},\rho_l) &= \frac{A_{l,n}D_n^{(3)}(\rho_l) + D_n^{(1)}(\rho_l)R_n(\rho_l)}{A_{l,n} + R_n(\rho_l)}, \\
H_n^b(B_{l,n},\rho_l) &= \frac{B_{l,n}D_n^{(3)}(\rho_l) + D_n^{(1)}(\rho_l)R_n(\rho_l)}{B_{l,n} + R_n(\rho_l)},
\end{aligned}
\tag{B.45}
$$

with the mathematical terms which are related to the Riccati-Bessel function

$$
\begin{aligned}
D_n^{(1)}(\rho) &= \psi_n'(\rho)/\psi_n(\rho), \\
D_n^{(3)}(\rho) &= \zeta_n'(\rho)/\zeta_n(\rho), \\
R_n(\rho) &= \psi_n(\rho)/\zeta_n(\rho),
\end{aligned}
\tag{B.46}
$$

and the initial $A_{l,n} = B_{l,n} = 0$. Until now, if we calculate field out of the multilayer sphere, we need to calculate $a_{N+1,n} = A_{N+1,n}$ and $b_{N+1,n} = B_{N+1,n}$ iteratively and substitute $a_{N+1,n}$ and $b_{N+1,n}$ into Eq. (B.40). However, to implement the code, we may also suffer from several numerical issues because of the round-off errors. We take advantage of the computational algorithm in [48] and [49] to compute $D_n^{(1)}(\rho)$, $D_n^{(3)}(\rho)$ and $R_n(\rho)$:

n in Eq. (B.40) cannot be infinite large, the truncation $N_{\max}$ is

$$N_{\max} = \begin{cases} \bar{\rho}_N + 4\bar{\rho}_N^{1/3} + 1 & 0.02 \leqslant \bar{\rho}_N < 8, \\ \bar{\rho}_N + 4.05\bar{\rho}_N^{1/3} + 1, & 8 \leqslant \bar{\rho}_N < 4200, \\ \bar{\rho}_N + 4\bar{\rho}_N^{1/3} + 2, & 4200 \leqslant \bar{\rho}_N < 20000. \end{cases} \tag{B.47}$$

with $\bar{\rho}_N = k_0\sqrt{\epsilon_{N+1}}r_N$.

To compute $D_n^{(1)}(\rho)$, we start from $D_{N_{\max}}^{(1)} = 0$, and

$$D_{n-1}^{(1)}(\rho) = \frac{n}{\rho} - \frac{1}{D_n^{(1)}(\rho) + n/\rho} \tag{B.48}$$

with $n = N_{\max}, \cdots, 1$.

To compute $D_n^{(3)}(\rho)$, we start from $D_0^{(3)}(\rho) = \mathrm{i}$, and

$$D_n^{(3)}(\rho) = D_n^{(1)}(\rho) + \frac{\mathrm{i}}{\psi_n(\rho)\zeta_n(\rho)} \tag{B.49}$$

with

$$\begin{aligned} \psi_n(\rho)\zeta_n(\rho) &= \psi_{n-1}(\rho)\zeta_{n-1}(\rho)\left[n/\rho - D_{n-1}^{(1)}(\rho)\right]\left[n/\rho - D_{n-1}^{(3)}(\rho)\right] \\ \psi_0(\rho)\zeta_0(\rho) &= \frac{1}{2} - \frac{1}{2}\left\{\cos\left[2\Re(\rho)\right] + \mathrm{i}\sin\left[2\Re(\rho)\right]\right\}\exp\left[-2\Im(\rho)\right] \end{aligned} \tag{B.50}$$

To compute $R_n(\rho)$

$$\begin{aligned} R_n(\rho) &= R_{n-1}(\rho)\frac{D_n^{(3)}(\rho) + n/\rho}{D_n^{(1)}(\rho) + n/\rho} \\ R_0(\rho) &= \frac{1}{2} - \frac{1}{2}\exp[-2\mathrm{i}\Re(\rho)]\exp[2\Im(\rho)] \end{aligned} \tag{B.51}$$

Till now, the coefficients $a_{N+1,n} = A_{N+1,n}$ and $b_{N+1,n} = B_{N+1,n}$ are calculated, and after substituting into Eq. (B.40), the electric field outside of the multilayered sphere can be calculated.

C

FIELD SOLVER IN CYLINDRICAL COORDINATE SYSTEM

Similar as spherical harmonics, in cylindrical coordinate system, one can solve cylindrical harmonics, but it is not as straightforward as that of spherical harmonics. As we know, these mode solvers are efficient for scale smaller than hundreds of wavelength. In this scale, the most important application is graded-index (GRIN) fiber. So here we want to mention an approximated but specific solver for GRIN fiber. Let's solve field equation in cylindrical coordinate system. This solver only works for fiber with permittivity

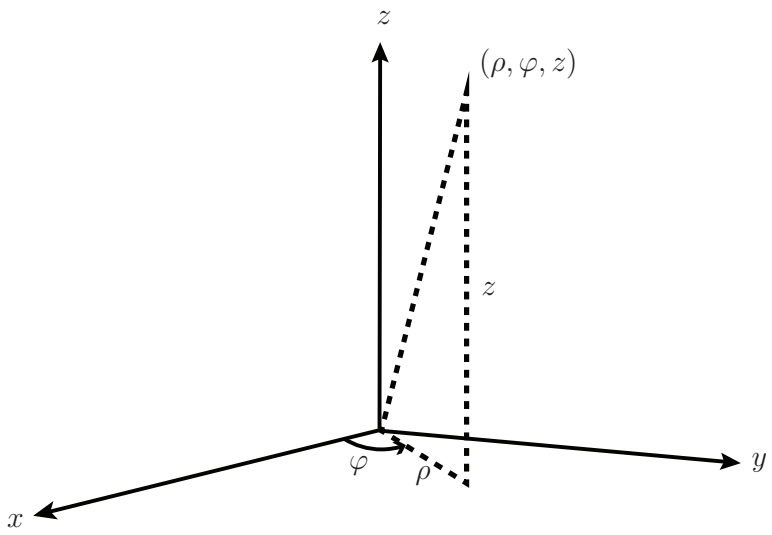


Figure 38: Specifying cylindrical coordinate system in Cartesian coordinate system.

$$\epsilon(\rho) = \epsilon_c \left[1 - 2\delta \left(\frac{r}{r_0}\right)^2\right], \tag{C.1}$$

with r_0 the radius of fiber core, and ϵ_c the largest relative permittivity of core δ is a constant.

The fiber solver contains three steps:

- Representing E_z and H_z with only one free coefficient
- Representing E_x, E_y, H_x and H_y by using E_z and H_z
- Using overlap integral to calculate the free coefficient

C.1 SOLVING E_z AND H_z IN CYLINDRICAL COORDINATE SYSTEM

In the cylindrical coordinate system, electric field is represented as

$$\boldsymbol{E}(\boldsymbol{r}) = \begin{bmatrix} E_\rho \\ E_\varphi \\ E_z \end{bmatrix} (\boldsymbol{r}), \tag{C.2}$$

with $\boldsymbol{r} = (\rho, \varphi, z)$, and E_ρ, E_φ, E_z denoting field components along radial, inclination angle, and z direction.

Substituting the representation into Helmholtz equation (2.24), and approximately consider $\boldsymbol{\nabla} \ln \epsilon(\boldsymbol{r}) \approx 0$, because of small variation of ϵ. which is in the magnitude of 10^{-3}, we obtain three separated equations

$$\Delta \boldsymbol{E}(\boldsymbol{r}) = \begin{bmatrix} \Delta E_\rho - \dfrac{E_\rho}{\rho^2} - \dfrac{2}{\rho^2}\partial_\rho \varphi E_\varphi \\ \Delta E_\varphi - \dfrac{E_\varphi}{\rho^2} + \dfrac{2}{\rho^2}\partial_\rho \varphi E_\rho \\ \Delta E_z \end{bmatrix} \tag{C.3}$$

Similarly, replace $\boldsymbol{E}$ by $\boldsymbol{H}$, identical form is achieved. From Eq. (C.3), only z component of field is decouple. Next step, we focus on solver E_z and H_z by solving the following equation

$$\Delta \psi(\boldsymbol{r}) + k_0^2 \epsilon(\boldsymbol{r}) \psi(\boldsymbol{r}) = 0, \tag{C.4}$$

with $\psi(\boldsymbol{r})$ denoting $E_z(\boldsymbol{r})$ and $H_z(\boldsymbol{r})$.

Representing the Laplace operator in cylinder coordinate system to be

$$\Delta = \frac{1}{\rho}\partial_\rho(\rho\partial_\rho) + \frac{1}{\rho^2}\partial_\varphi^2 + \partial_z^2, \tag{C.5}$$

and separate variables as

$$\psi(\boldsymbol{r}) = P(\boldsymbol{\rho})\Phi(\varphi)Z(z), \tag{C.6}$$

we obtain

$$\begin{aligned} \frac{1}{\rho}\partial_\rho \left[\rho \partial_\rho P(\boldsymbol{\rho})\right] \Phi(\varphi) Z(z) + \frac{1}{\rho^2}\partial_\varphi^2 \Phi(\varphi) P(\boldsymbol{\rho}) Z(z) + \partial_z^2 Z(z) P(\boldsymbol{\rho}) \Phi(\varphi) + \\ k_0^2 \epsilon(\boldsymbol{r}) P(\boldsymbol{\rho}) \Phi(\varphi) Z(z) = 0 \end{aligned} \tag{C.7}$$

Divided Eq. (C.7) by $P(\boldsymbol{\rho})\Phi(\varphi)Z(z)$, and shifting the term related to z to the right hand side, we get

$$\frac{1}{\rho P(\rho)}\partial_\rho \left[\rho \partial_\rho P(\boldsymbol{\rho})\right] + \frac{1}{\rho^2 \Phi(\varphi)}\partial_\varphi^2 \Phi(\varphi) + k_0^2 \epsilon(\boldsymbol{r}) = -\frac{1}{Z(z)}\partial_z^2 Z(z) = \beta^2 \tag{C.8}$$

If Eq. (C.8) is valid, the left hand side and right hand side must equal to a constant, here we assume β^2. The term with z can be isolated and written as

$$\partial_z^2 Z(z) + \beta^2 Z(z) = 0, \tag{C.9}$$

and the solution is

$$Z(z) = \exp(\mathrm{i}\beta z). \tag{C.10}$$

In Eq. (C.10) we didn't give any free parameter, i.e., Z_0 which is multiplied to $Z(z)$ or z_0 inside the bracket. Both of them can be realized by a single coefficient multiplied to ψ after all derivation.

Then we divide ρ^2 to isolate φ, and shift the terms related to φ to the right hand side, we get

$$\frac{\rho}{P(\rho)}\partial_\rho\left[\rho\partial_\rho P(\boldsymbol{\rho})\right] + (k_0^2\epsilon(\boldsymbol{r}) - \beta^2)\rho^2 = -\frac{1}{\Phi(\varphi)}\partial_\varphi^2\Phi(\varphi) = l^2 \tag{C.11}$$

To get a single value $\Phi(\varphi)$ ($\Phi(\varphi + 2\pi) = \Phi(\varphi)$), l is an integer. Conventionally, the solution is denoted as

$$\Phi(\varphi) = \cos(l\varphi)\text{or } \sin(l\varphi) \tag{C.12}$$

The left hand side of Eq. (C.11) is finally written as

$$\mathrm{d}_\rho^2 P(\rho) + \frac{1}{\rho}\mathrm{d}_\rho P(\rho) + (k_0^2\epsilon(\boldsymbol{r}) - \beta^2 - \frac{l^2}{\rho^2})P(\rho) = 0. \tag{C.13}$$

If the fiber epsilon is represented as C.1, we follow a smart mathematical transforms to denote $P(\rho)$ as another function $g(s)$

$$P(\rho) = \exp\left(-\frac{s}{2}\right)s^{\frac{l}{2}}g(s), \tag{C.14}$$

with

$$s = \tau\rho^2 = \frac{k_0\sqrt{2\delta\epsilon_c}}{r_0}\rho^2. \tag{C.15}$$

From Eq. (C.15) we can derive

$$\begin{aligned}\mathrm{d}_\rho &= 2\sqrt{\tau s}\,\mathrm{d}_s,\\ \mathrm{d}_\rho^2 &= 4\tau s\,\mathrm{d}_s^2 + 2\tau\,\mathrm{d}_s.\end{aligned} \tag{C.16}$$

Now substituting Eqs. (C.14)-(C.16) and the relative permittivity (C.1) into Eq. (C.13), we get

$$s\frac{\mathrm{d}^2 g(s)}{\mathrm{d}s^2} + (l+1-s)\frac{\mathrm{d}g(s)}{\mathrm{d}s} + mg(s) = 0, \tag{C.17}$$

with

$$m = \frac{1}{4}\left[\frac{r_0}{k_0\sqrt{2\epsilon_c\delta}}\left(k_0^2\epsilon_c - \beta^2\right) - 2l - 2\right]. \tag{C.18}$$

When m is a non-negative integer, Eq. (C.17) has exactly identical form with the associated Laguerre equation [47] and $g(s)$ is the associated Laguerre polynomials $L_m^l(s)$. To simplify the notation, we introduce

$$\omega_0 = \sqrt{2/\tau} = \frac{2r_0}{k_0\sqrt{2\delta\epsilon_c}}. \tag{C.19}$$

And $P(\rho)$ is

$$P_{lm}(\rho) = L_m^l\left(\frac{2\rho^2}{\omega_0^2}\right)\left(\frac{\sqrt{2}\rho}{\omega_0}\right)^l \exp\left(-\frac{\rho^2}{\omega_0^2}\right). \tag{C.20}$$

From Eq. (C.18), β is discrete and can be calculated from l and m

$$\beta_{lm} = k_0\sqrt{\epsilon_c}\left[1 - 4\frac{2m+l+1}{k_0^2\epsilon_c\omega_0^2}\right]^{\frac{1}{2}} \tag{C.21}$$

From

Till now we can represent E_z and H_z as follows:

$$\begin{bmatrix} E_z \\ H_z \end{bmatrix}_{lm} = \begin{bmatrix} A \\ B \end{bmatrix}_{lm} P_{lm}(\rho)\exp(\mathrm{i}l\varphi)\exp(\mathrm{i}\beta_{lm}z) \tag{C.22}$$

with $P_{lm}(\rho)$ in Eq. (C.20) and β_{lm} in Eq. (C.21).

However, there are still two free coefficients A and B. Here we use another relation between E_z and H_z of mode (l,m) [108] to reduce the number of free coefficients to 1.

$$\eta = -\mathrm{i}\frac{k_0}{\beta_{lm}}\sqrt{\frac{\mu_0}{\epsilon_0}}\frac{H_z}{E_z}. \tag{C.23}$$

η can be 0, ∞, 1, -1 and which denotes transverse magnetic (TM) mode($H_z = 0$), transverse electric (TE) mode($E_z = 0$), EH mode, and HE mode. Here we again induce approximation by using η. Rigorously, $|\eta|$ approaching to 1, but generally not exactly 1. The approximation works for weakly guided fiber, which means δ in Eq. (C.1) should be very small.

Substituting η into Eq. (C.22), we get the representation of E_z and H_z with only one free coefficient A

$$\begin{bmatrix} E_z \\ H_z \end{bmatrix}_{lm} = A_{lm}\begin{bmatrix} 1 \\ \eta\frac{\mathrm{i}\beta_{lm}}{k_0}\sqrt{\frac{\epsilon_0}{\mu_0}} \end{bmatrix} P_{lm}(\rho)\exp(\mathrm{i}l\varphi)\exp(\mathrm{i}\beta_{lm}z). \tag{C.24}$$

C.2 FIELD REPRESENTATION IN CARTESIAN COORDINATE SYSTEM

Till now, E_z and H_z of mode (l,m) is represented. Only one free coefficient A is left, which will be fixed by overlap integral in the next section. However, to do overlap integral, all field components should be represented. In most of the techniques, a general input field is sampled in Cartesian coordinate system. So from now on, we come back the Cartesian coordinate system and just discuss mode (l,m). There are two physical objects needs to be transfer to Cartesian coordinate system, i.e., field components $\boldsymbol{E}$ and $\boldsymbol{H}$, and coordinate $\boldsymbol{r}$.

Firstly, we calculate E_x, E_y, H_x, H_y by using the two curl equations in (2.10). Although we didn't come to detail derivation in last section, but from Eq. (C.3), all $Z(z)$ of all field

components can be separated and represented as Eq. (C.10), so all ∂_z can be replaced by $\mathrm{i}\beta$.

$$\begin{cases} E_x &= \dfrac{\mathrm{i}}{k_0^2\epsilon(\boldsymbol{r})-\beta^2}(\beta\partial_x E_z + \kappa_0\sqrt{\dfrac{\mu_0}{\epsilon_0}}\partial_y H_z) \\ E_y &= \dfrac{\mathrm{i}}{k_0^2\epsilon(\boldsymbol{r})-\beta^2}(\beta\partial_y E_z - \kappa_0\sqrt{\dfrac{\mu_0}{\epsilon_0}}\partial_x H_z) \\ H_x &= \dfrac{\mathrm{i}}{k_0^2\epsilon(\boldsymbol{r})-\beta^2}(-\kappa_0\sqrt{\dfrac{\epsilon_0}{\mu_0}}\epsilon(\boldsymbol{r})\partial_y E_z + \beta\partial_x H_z) \\ H_y &= \dfrac{\mathrm{i}}{k_0^2\epsilon(\boldsymbol{r})-\beta^2}(\kappa_0\sqrt{\dfrac{\epsilon_0}{\mu_0}}\epsilon(\boldsymbol{r})\partial_x E_z + \beta\partial_y H_z) \end{cases} \tag{C.25}$$

Next step, because E_z and H_z are represented in the cylinder coordinate system, we replace ∂_x and ∂_y by ∂_ρ and ∂_φ

$$\begin{cases} \partial_x &= \cos\varphi\partial_\rho - \rho^{-1}\sin\varphi\partial_\varphi \\ \partial_y &= \sin\varphi\partial_\rho + \rho^{-1}\cos\varphi\partial_\varphi \end{cases}. \tag{C.26}$$

E_x, E_y, H_x, H_y can be calculated by substituting Eq. (C.26) and (C.24) into Eq. (C.25), with

$$\mathrm{d}_\rho L_m^l(s) = \begin{cases} -L_{m-1}^{l+1}(s), & m \geqslant 1 \\ 0 & \text{otherwise} \end{cases}. \tag{C.27}$$

Here we don't write them done in details.

Next step is to transform coordinates $\boldsymbol{r}$ back to Cartesian coordinate system by

$$\begin{cases} \rho &= \sqrt{(x^2+y^2)} \\ \varphi &= \arctan(y,x) \end{cases}. \tag{C.28}$$

The last step is calculation of A_{lm} to normalize the power of mode (l,m). More specifically,

$$\frac{1}{2}\iint_{-\infty}^{\infty}(\boldsymbol{E}_{lm}(x,y)\times\boldsymbol{H}_{lm}^{\star}(x,y))\cdot\hat{\boldsymbol{e}}_z\,\mathrm{d}x\,\mathrm{d}y = 1. \tag{C.29}$$

Therefore,

$$A_{lm} = \left[\frac{1}{2}\iint_{-\infty}^{\infty}(\boldsymbol{E}_{lm}(x,y)\times\boldsymbol{H}_{lm}^{\star}(x,y))\cdot\hat{\boldsymbol{e}}_z\,\mathrm{d}x\,\mathrm{d}y\right]^{-\frac{1}{2}}. \tag{C.30}$$

C.3 OVERLAP INTEGRAL TO FIX THE FREE COEFFICIENT

Now we have represented all $(E_x,E_y,E_z)_{lm}^{\mathrm{T}}$ and $(H_x,H_y,H_z)_{lm}^{\mathrm{T}}$ in Eq. (C.22) and Eq. (C.25). Assume the general input field $\boldsymbol{E}^{\mathrm{in}}(x,y,z=0)$ propagating through a GRIN fiber with modes (l,m), what is the output field $\boldsymbol{E}^{\mathrm{out}}(x,y,z)$. The modes (l,m) are propagate modes, and all other field which cannot project to these modes, are attenuated during the propagation. The projection from $E^{\mathrm{in}}(x,y)$ to modes (l,m) uses the overlap integral.

First let's represent the input field as

$$E^{\mathrm{in}}(x,y) = \sum_{m=0}^{M}\sum_{l=-L}^{L} a_{lm}\boldsymbol{E}_{lm}(x,y), \tag{C.31}$$

with

$$a_{lm} = \frac{1}{2} \iint_{-\infty}^{\infty} (\boldsymbol{E}^{\text{in}}(x,y) \times \boldsymbol{H}^{\star}_{lm}(x,y)) \cdot \hat{\boldsymbol{e}}_z \, \mathrm{d}x \, \mathrm{d}y. \tag{C.32}$$

Finally $\boldsymbol{E}^{\text{out}}(x,y,z)$ is

$$E^{\text{out}}(x,y,z) = \sum_{m=0}^{M} \sum_{l=-L}^{L} a_{lm} \boldsymbol{E}_{lm}(x,y) \exp(\mathrm{i}\beta z). \tag{C.33}$$

D

THIN ELEMENT APPROXIMATION FOR GRADED-INDEX LAYER

In 1991, Turunen [31] extended the thin element approximation, which is originally just work for paraxial incident beam, to model oblique/inclined input field. Here we derive the extended thin element approximation for a thin graded-index(GRIN) layer.

More specifically, the task is that we propagate a source field through a thin GRIN layer from plane Σ^{in} to Σ^{out}, as shown in Fig. (39). The source field is located INSIDE the GRIN layer in $\Sigma^{\text{in},+}$, which is

$$\boldsymbol{E}^{\text{in}}(\boldsymbol{\rho}')|_{\Sigma^{\text{in},+}} = \boldsymbol{E}_c \exp(\mathrm{i}\boldsymbol{\kappa} \cdot \boldsymbol{\rho}'), \tag{D.1}$$

with $\boldsymbol{\rho}' = (x', y')$ and $\boldsymbol{E}_c$ the constant complex amplitude of the field. This field only has one $\boldsymbol{\kappa}$. The output field INSIDE the GRIN layer in plane $\Sigma^{\text{out},-}$ is to be calculated.

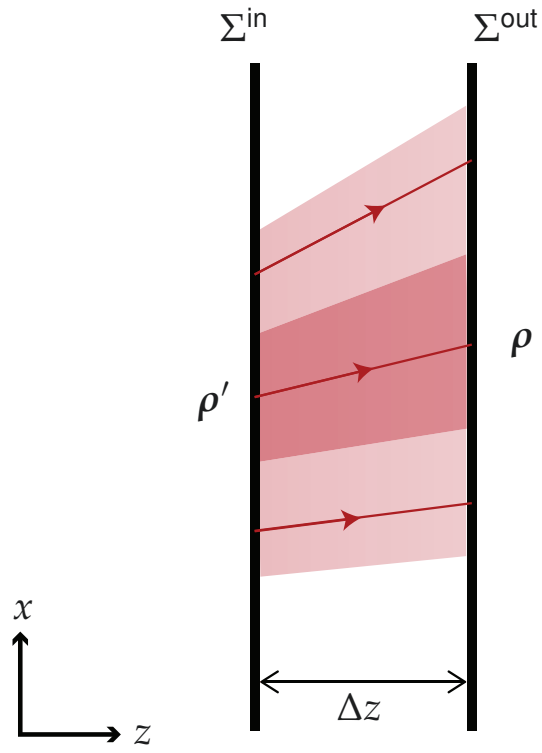


Figure 39: The illustration of the extended thin element approximation in GRIN layer.

The input field can be treated locally as plane wave, similar as geometric field. Each local plane wave just see/interact with the GRIN media locally. We can first calculate $k_z(\boldsymbol{\rho}')$

$$k_z(\boldsymbol{\rho}') = \sqrt{k_0^2 \epsilon(\boldsymbol{\rho}') - ||\boldsymbol{\kappa}||^2}. \tag{D.2}$$

In GRIN media, the skeleton of local plane wave should be a curve, but because the layer is very thin (thin element), we approximate the skeleton to be straight, from coordinate $\boldsymbol{\rho}'$ to $\boldsymbol{\rho}$, and the one-to-one mapping relation is recorded as $\boldsymbol{\rho}(\boldsymbol{\rho}')$. The thin element approximation just include the phase change during propagation, and keeps amplitude invariant. The phase change from $\boldsymbol{\rho}'$ to $\boldsymbol{\rho}$ is

$$\Delta\psi[\boldsymbol{\rho}(\boldsymbol{\rho}')] = \boldsymbol{\kappa} \cdot (\boldsymbol{\rho} - \boldsymbol{\rho}') + k_z(\boldsymbol{\rho}')\Delta z. \tag{D.3}$$

And the final output field is

$$\begin{aligned} \boldsymbol{E}^{\text{out}}(\boldsymbol{\rho})|_{\Sigma^{\text{out,-}}} &= \boldsymbol{E}^{\text{in}}(\boldsymbol{\rho}')|_{\Sigma^{\text{in,+}}} \exp \mathrm{i}\Delta\psi[\boldsymbol{\rho}(\boldsymbol{\rho}')] \\ &= \boldsymbol{E}_c \exp(\mathrm{i}\boldsymbol{\kappa} \cdot \boldsymbol{\rho}) \exp[\mathrm{i}\kappa_z(\boldsymbol{\rho}')\Delta z] \\ &\approx \boldsymbol{E}_c \exp(\mathrm{i}\boldsymbol{\kappa} \cdot \boldsymbol{\rho}) \exp[\mathrm{i}\kappa_z(\boldsymbol{\rho})\Delta z] \text{ assume } \quad \boldsymbol{\rho} \approx \boldsymbol{\rho}' \end{aligned} \tag{D.4}$$

In summary, if we know $\boldsymbol{E}_c$ and $\boldsymbol{\kappa}$ of input field, output field can be calculated in the same grid with an additional phase term $\kappa_z(\boldsymbol{\rho})\Delta z$.

E

STABILITY ANALYSIS OF THE ODES

E.1 ANALYSIS OF THE ODES FOR HOMOGENEOUS MEDIA

Let's start from a simple case to calculate the eigenvalue η for homogeneous media with permittivity ϵ. In this case, in $\underline{\mathbf{M}}$, $\underline{\tilde{\epsilon}}$ and $\underline{\tilde{\epsilon}}^{-1}$ are constant, which can be replaced by ϵ and $1/\epsilon$.

$$
\begin{aligned}
\det\left[\underline{\mathbf{M}}-\eta\boldsymbol{\mathcal{I}}\right] &= (\mathrm{i}k_0)^4 \begin{vmatrix} -\frac{\eta}{\mathrm{i}k_0} & 0 & \frac{k_x}{k_0}\frac{1}{\epsilon}\frac{k_y}{k_0} & 1-\frac{k_x}{k_0}\frac{1}{\epsilon}\frac{k_x}{k_0} \\ 0 & -\frac{\eta}{\mathrm{i}k_0} & \frac{k_y}{k_0}\frac{1}{\epsilon}\frac{k_y}{k_0}-1 & -\frac{k_y}{k_0}\frac{1}{\epsilon}\frac{k_x}{k_0} \\ -\frac{k_xk_y}{k_0^2} & \frac{k_x^2}{k_0^2}-\epsilon & -\frac{\eta}{\mathrm{i}k_0} & 0 \\ \epsilon-\frac{k_y^2}{k_0^2} & \frac{k_yk_x}{k_0^2} & 0 & -\frac{\eta}{\mathrm{i}k_0} \end{vmatrix} \\
&= (\mathrm{i}k_0)^4 \left[(-\tfrac{\eta}{\mathrm{i}k_0}) \begin{vmatrix} -\frac{\eta}{\mathrm{i}k_0} & \frac{k_y}{k_0}\frac{1}{\epsilon}\frac{k_y}{k_0}-1 & -\frac{k_y}{k_0}\frac{1}{\epsilon}\frac{k_x}{k_0} \\ \frac{k_x^2}{k_0^2}-\epsilon & -\frac{\eta}{\mathrm{i}k_0} & 0 \\ \frac{k_yk_x}{k_0^2} & 0 & -\frac{\eta}{\mathrm{i}k_0} \end{vmatrix} + \frac{k_x}{k_0}\frac{1}{\epsilon}\frac{k_y}{k_0} \begin{vmatrix} 0 & -\frac{\eta}{\mathrm{i}k_0} & -\frac{k_y}{k_0}\frac{1}{\epsilon}\frac{k_x}{k_0} \\ -\frac{k_xk_y}{k_0^2} & -\frac{\eta}{\mathrm{i}k_0} & 0 \\ \epsilon-\frac{k_y^2}{k_0^2} & \frac{k_yk_x}{k_0^2} & -\frac{\eta}{\mathrm{i}k_0} \end{vmatrix} \right. \\
&\left. -(1-\frac{k_x}{k_0}\frac{1}{\epsilon}\frac{k_x}{k_0}) \begin{vmatrix} 0 & -\frac{\eta}{\mathrm{i}k_0} & \frac{k_y}{k_0}\frac{1}{\epsilon}\frac{k_y}{k_0}-1 \\ -\frac{k_xk_y}{k_0^2} & \frac{k_x^2}{k_0^2}-\epsilon & -\frac{\eta}{\mathrm{i}k_0} \\ \epsilon-\frac{k_y^2}{k_0^2} & \frac{k_yk_x}{k_0^2} & 0 \end{vmatrix} \right] \\
&= -k_0^2(\eta^2+k_z^2) = 0,
\end{aligned}
\tag{E.1}
$$

with

$$
k_z^2 = k_0^2\epsilon - (k_x^2+k_y^2).
$$

The eigenvalue is

$$
\eta = \pm\mathrm{i}k_z.
$$

So the ODE (3.10) for homogeneous media is not always stable

$$
\mathbb{R}\mathrm{e}\left\{\eta\right\} = \begin{cases} 0 & k_x^2+k_y^2 \leqslant k_0^2\epsilon \quad \text{propagating wave} \\ \pm|k_z| & k_x^2+k_y^2 > k_0^2\epsilon \quad \text{evanescent wave} \end{cases}
\tag{E.2}
$$

From Eq. (E.2), the eigenvalue for propagation wave always satisfies Eq. (3.25), but that of evanescent wave does not. In the further test of the algorithm, if we include

evanescent wave $k_x^2 + k_y^2 > k_0^2\epsilon$, even when the initial value $\tilde{\boldsymbol{V}}(\boldsymbol{\kappa}, z_0) = 0$, we get an infinity large result field, which is the enlarged numerical noise. So to keep the stable calculation, we need to add a criteria

$$\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa}, z) = 0, \text{ for } ||\boldsymbol{\kappa}||^2 > k_0^2\epsilon \tag{E.3}$$

Therefore, although we don't use physical approximation during the derivation of Eq. (3.10), except for the initial condition, it is not stable for evanescent wave calculation, when we solve it as the "initial value problem", even for homogeneous media.

E.2 ANALYSIS OF THE ODES FOR $y-$INVARIANT CASE

Now we consider the inhomogeneity of the GRIN media, and assume both electromagnetic field and the inhomogeneity is y-invariant. Please note, we don't need to calculate the exact eigenvalues, which is already done by Fourier modal method. **What we are interested in is the largest eigenvalues.** The ODE for y-invariant (3.13) is an anti-diagonal block matrix, whose entries are zero except those on the diagonal going from the lower left corner to the upper right corner. If we just analyse the matrix for one thin layer, where the z-variance of permittivity can be also neglected. Then Eq. (3.20) can be rewritten to be a second-order ODE as

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}z^2}\tilde{P}_1 &= \underline{\mathbf{M}}_1\tilde{P}_1, \\
\frac{\mathrm{d}^2}{\mathrm{d}z^2}\tilde{P}_2 &= \underline{\mathbf{M}}_2\tilde{P}_2, \\
\frac{\mathrm{d}^2}{\mathrm{d}z^2}\tilde{P}_4 &= \underline{\mathbf{M}}_4\tilde{P}_4, \\
\frac{\mathrm{d}^2}{\mathrm{d}z^2}\tilde{P}_5 &= \underline{\mathbf{M}}_5\tilde{P}_5,
\end{aligned} \tag{E.4}$$

with

$$\underline{\mathbf{M}}^{\mathrm{TE}}(\boldsymbol{\kappa}, z) = \underline{\mathbf{M}}_2(\boldsymbol{\kappa}, z) = \underline{\mathbf{M}}_4(\boldsymbol{\kappa}, z) = k_x^2 - k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}, \tag{E.5}$$

and

$$\underline{\mathbf{M}}_1^{\mathrm{TM}}(\boldsymbol{\kappa}, z) = \underline{\mathbf{M}}_1(\boldsymbol{\kappa}, z) = -k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}, \tag{E.6}$$

and

$$\underline{\mathbf{M}}_2^{\mathrm{TM}}(\boldsymbol{\kappa}, z) = \underline{\mathbf{M}}_5(\boldsymbol{\kappa}, z) = -k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + \underline{\tilde{\boldsymbol{\epsilon}}}k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x, \tag{E.7}$$

What is advantages to convert the first-order ODE into second-order one, and analyse the stability of the second-order one (E.4)?

- Eqs. (E.4) are equivalent to the ODE in Eq. (3.20) for a thin layer(z-variance of permittivity can be also neglected), and different polarization field components are decoupled.

- The matrix $\underline{\mathbf{M}}^{\mathrm{TE}}$ and $\underline{\mathbf{M}}^{\mathrm{TM}}$ have higher and symmetry, and mathematically properties can be used to analyse the extreme value of eigenvalues.

Now the mathematical solution for matrices $\underline{\mathbf{M}}^{\mathrm{TE}}$ and $\underline{\mathbf{M}}^{\mathrm{TM}}$ are

$$\tilde{\boldsymbol{P}}(z) = \sum_i c_i^+ \tilde{\boldsymbol{p}}_i^+(\boldsymbol{\kappa}) \exp(\eta_i z) + c_i^- \tilde{\boldsymbol{p}}_i^-(\boldsymbol{\kappa}) \exp(-\eta_i z). \tag{E.8}$$

with η_i^2 the eigenvalue of $\underline{\mathbf{M}}^{\mathrm{TE}}$ or $\underline{\mathbf{M}}^{\mathrm{TM}}$. To make sure that the perturbation is not enlarged, criteria (3.24) can be rewritten as

$$\eta_i^2 \subset \mathbb{R} \text{ and } \eta_i^2 \leqslant 0. \tag{E.9}$$

Eigenvalue η_i^2 of $\underline{\mathbf{M}}^{\mathrm{TE}}$

In this work, we limit $\epsilon(x)$ to be real-valued, so the spatial frequency $\tilde{\epsilon}(k_x)$ has the property

$$\tilde{\epsilon}(-k_x) = \tilde{\epsilon}^\star(k_x), \tag{E.10}$$

with * denoting the conjugate. If we sample k_x from $k_{x\min}$ to $k_{x\max}$ with sampling points N and sampling distance δk_x, i.e., $\delta k_x = (k_{x\max} - k_{x\min})/(N-1)$. Then $\underline{\mathbf{M}}^{\mathrm{TE}}$ can be explicitly written in the matrix form

$$\underline{\mathbf{M}}^{\mathrm{TE}} = k_x^2 - k_0^2 \underline{\tilde{\boldsymbol{\epsilon}}} =$$

$$\begin{bmatrix} k_{x\min}^2 & 0 & \cdots & 0 & 0 \\ 0 & (k_{x\min}+\delta k_x)^2 & 0 & \cdots & 0 \\ 0 & 0 & (k_{x\min}+2\delta k_x)^2 & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & \cdots & 0 & k_{x\max}^2 \end{bmatrix} - k_0^2 \begin{bmatrix} \tilde{\epsilon}(0) & \tilde{\epsilon}^\star(\delta k_x) & \tilde{\epsilon}^\star(2\delta k_x) & \cdots & \tilde{\epsilon}^\star((N-1)\delta k_x) \\ \tilde{\epsilon}(\delta k_x) & \tilde{\epsilon}(0) & \tilde{\epsilon}^\star(\delta k_x) & \cdots & \tilde{\epsilon}^\star((N-2)\delta k_x) \\ \tilde{\epsilon}(2\delta k_x) & \tilde{\epsilon}(\delta k_x) & \tilde{\epsilon}(0) & \cdots & \tilde{\epsilon}^\star((N-3)\delta k_x) \\ & & \cdots & & \\ \tilde{\epsilon}((N-1)\delta k_x) & & \cdots & \tilde{\epsilon}(\delta k_x) & \tilde{\epsilon}(0) \end{bmatrix} \tag{E.11}$$

To analyse the eigenvalues of matrices, we need to know some definition and properties of specified matrices, as follows

Definition 1. *In linear algebra, a **diagonal matrix** is a matrix in which the entries outside the main diagonal are all zero.*

Theorem 1. *The eigenvalues of a diagonal matrix are the entries in the main diagonal.*

Definition 2. *In linear algebra, a **Hermitian matrix** is a complex square matrix that is equal to its own conjugate transpose, i.e., entry $m_{ij} = m_{ji}^\star$.*

Theorem 2. *The eigenvalues of a Hermitian matrix are real-valued.*

From definition 1 and theorem 1, we know the matrix k_x^2 is a diagonal matrix, whose eigenvalues are the entries of the diagonal, i.e., $\eta_i^2(k_x^2) = k_{x\min}^2$, $(k_{x\min}+\delta k_x)^2$, ..., $k_{x\max}^2$. This matrix also satisfies definition 2, so k_x^2 is a Hermitian matrix. So

$$\eta_i^2(k_x^2) \subset \mathbb{R} \text{ and } 0 \leqslant \eta_i^2(k_x^2) \leqslant |k_x|_{\max}^2. \tag{E.12}$$

To study the matrix $\underline{\tilde{\boldsymbol{\epsilon}}}$, the chapter 5.2 of [109] studied this kind of matrix and gives definition and theorem as follows:

Definition 3. *Let $f(x)$ be a real-valued function, and the entries of the special* ***Toeplitz matrix*** *$T_N(f)$ are*

$$t_n = \frac{1}{p}\int_{-p/2}^{p/2} f(x)\exp^{-\mathrm{i}n\delta kx}\,\mathrm{d}x$$

with n the diagonal index of the matrix $T_N(f)$ where $n = 0, \pm 1, \pm 2, \cdots, N-1$. p is the period of $f(x)$ and $\delta k = \dfrac{2\pi}{p}$.

$$T_n(f) = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \cdots & t_{-(N-1)} \\ t_1 & t_0 & t_{-1} & \cdots & t_{-(N-2)} \\ t_2 & t_1 & t_0 & \cdots & t_{-(N-3)} \\ & & \cdots & & \\ t_{N-1} & t_{N-2} & t_{N-3} & \cdots & t_0 \end{bmatrix}$$

Theorem 3. *We denote by L and U the essential lower and upper bound of $f(x)$, respectively. If $l_1, l_2, \cdots, l_N$ are the eigenvalues of the matrix $T_N(f)$, and $l_1 \geqslant l_2 \geqslant \cdots \geqslant l_N$, then*

$$l_1 \leqslant U,\ l_N \geqslant L.$$

And the equal sign holds when $N \to \infty$.

$\underline{\tilde{\boldsymbol{\epsilon}}}$ satisfies the definition 3 of $T_N(f)$, as $f(x) = \epsilon(x)$, and $t_0 = \tilde{\epsilon}(0), t_1 = \tilde{\epsilon}(\delta k_x), t_{-1} = \tilde{\epsilon}^{\star}(\delta k_x), \cdots, a_{N-1} = \tilde{\epsilon}((N-1)\delta k_x)$. Meanwhile, $\underline{\tilde{\boldsymbol{\epsilon}}}$ also satisfies definition 2, so we know that **all eigenvalues of $\underline{\tilde{\boldsymbol{\epsilon}}}$ are real, i.e., $\eta_i^2(\underline{\tilde{\boldsymbol{\epsilon}}}) \subset \mathbb{R}$.** From the theorem 3, the mathematical concept "essential lower" and "essential upper" bound of $f(x)$ is $\epsilon_{\min}$ and $\epsilon_{\max}$ in the TE case, so **the eigenvalue $\eta^2_{\mathbf{min}}(\underline{\tilde{\boldsymbol{\epsilon}}}) \geqslant \epsilon_{\mathbf{min}}$ and $\eta^2_{\mathbf{max}}(\underline{\tilde{\boldsymbol{\epsilon}}}) \leqslant \epsilon_{\mathbf{max}}$.** So

$$\eta_i^2(k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}) \subset \mathbb{R} \text{ and } k_0^2\epsilon_{\min} \leqslant \eta_i^2(k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}) \leqslant k_0^2\epsilon_{\max} \tag{E.13}$$

To analyse the eigenvalues η_i^2 of $\underline{\mathbf{M}}^{\mathrm{TE}}$ we need Weyl's Inequalities theorem to study the eigenvalues of the summation of two Hermitian matrices[110].

Theorem 4. *A and B be $N \times N$ Hermitian matrices.*

- *$l_1(A), l_2(A), \cdots, l_N(A)$ are the eigenvalues of the matrix A, and $l_1(A) \geqslant l_2(A) \geqslant \cdots \geqslant l_N(A)$.*
- *$l_1(B), l_2(B), \cdots, l_N(B)$ are the eigenvalues of the matrix B, and $l_1(B) \geqslant l_2(B) \geqslant \cdots \geqslant l_N(B)$.*
- *$l_1(A+B), l_2(A+B), \cdots, l_N(A+B)$ are the eigenvalues of the matrix $A+B$, and $l_1(A+B) \geqslant l_2(A+B) \geqslant \cdots \geqslant l_N(A+B)$*

Then,

$$\begin{aligned} l_j(A+B) &\leqslant l_i(A) + l_{j-i+1}(B) \text{ for } i \leqslant j \\ l_j(A+B) &\geqslant l_i(A) + l_{j-i+N}(B) \text{ for } i \geqslant j \end{aligned} \tag{E.14}$$

If we put $i = j$ in the inequalities of theorem 4 (E.14), we could easily get

$$l_j(A) + l_N(B) \leqslant l_j(A+B) \leqslant l_j(A) + l_1(B) \text{ for } j = 1, 2, \cdots, N.$$

The matrix $\underline{\mathbf{M}}^{\mathrm{TE}} = k_x^2 - k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}$. Matrices k_x^2 and $-k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}$ are both Hermitian matrices, so the properties of eigenvalues $\eta_i^2(\underline{\mathbf{M}}^{\mathrm{TE}})$ are

- $\eta_i^2(\underline{\mathbf{M}}^{\mathrm{TE}}) \subset \mathbb{R}$ because the summation of Hermitian matrices is still Hermitian.
- $-k_0^2\epsilon_{\max} \leqslant \eta_i^2(\underline{\mathbf{M}}^{\mathrm{TE}}) \leqslant |k_x|_{\max}^2 - k_0^2\epsilon_{\min}$.

Compared to the stable criteria (E.9), the criteria of the stable algorithm for TE case is

$$|k_x|_{\max}^2 \leqslant k_0^2\epsilon_{\min} \tag{E.15}$$

To make sure that the further algorithm is stable, we need to add one requirement, similar as that of homogeneous media Eq. (E.3)

$$\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{\kappa}, z) = 0, \text{ for } ||\boldsymbol{\kappa}||^2 > k_0^2\epsilon_{\min}. \tag{E.16}$$

Eigenvalue η_i^2 of $\underline{\mathbf{M}}^{\mathrm{TM}}$

Let's first analyse $\underline{\mathbf{M}}_1^{\mathrm{TM}}$ (E.6). Eq. (E.13) gives the analysis of $\eta_i^2(k_0^2\underline{\tilde{\boldsymbol{\epsilon}}})$.

Then how to calculate the extreme eigenvalues of $k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}$?

First, let's see if $k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}$ has real eigenvalues. From paper [111], we get theorem 5 as follows

Theorem 5. *For a diagonalizable matrix H to have all real eigenvalues, a necessary and sufficient condition is that there exists a Hermitian matrix $W = W^H$ such that*

$$WH = H^H W, \tag{E.17}$$

and all the eigenvalues of W are positive real-valued.

In this case, $H = k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}$ and $H^H = \underline{\tilde{\boldsymbol{\epsilon}}}k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x$. Compared with Eq. (E.17), it is obvious that $W = \underline{\tilde{\boldsymbol{\epsilon}}}$ can be one solution, Because from theorem 3, we also know that the eigenvalue are positive real-valued. Therefore $k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}$ **satisfies theorem 5, and all eigenvalues are real**.

Definition 3. *In linear algebra, two n by n matrices A and B are called similar if there exsits an invertible n by n matrix P such that*

$$B = P^{-1}AP$$

The matrix $k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}$ can be considered as the product of matrices $A = k_x$ and $B = \underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}$, and following the definition 3, these two matrices A and B are similar matrices, when $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}\underline{\tilde{\boldsymbol{\epsilon}}} = I$, with I identical matrix.

Theorem 6. *Similar matrices have identical eigenvalues.*

From theorem 6, two matrices $A = k_x$ and $B = \underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}$ have identical eigenvalues, which are $-k_{\max}, -k_{\max} + \delta k_x, -k_{\max} + 2\delta k_x, ..., -k_{\max} + (N-1)\delta k_x, k_{\max}$. But it is still not clear what is the eigenvalue of the product of two similar matrices. The eigenvalue

of the production of two similar matrices, especially one of them is diagonal, must be known. **We have our guess, that the eigenvalue $\eta_i^2(k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}) \approx k_x^2$.** We have several numerical tests about different $\epsilon(x)$. In Fig. (40) one of the test results are shown. The left figure is $\epsilon(x)$, while the right figure shows eigenvalue $\eta_i^2(k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}})$ and $\eta_i^2(k_x^2)$. We use the matrix analysis software MATLAB to calculate all eigenvalues. In this calculation,

- k_x is equidistant sampled by 1500 points. $-1.5k_0\sqrt{\epsilon_{\min}} \leqslant k_x \leqslant 1.5k_0\sqrt{\epsilon_{\min}}$
- There should be 1500 eigenvalues. The eigenvalues are sorted from minimum to maximum, and plot.

We can see that our guess is correct. Honestly, we didn't find a perfect mathematical proof of our guess. We tried two methods to find the inequality of $\eta_i^2(k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}})$.

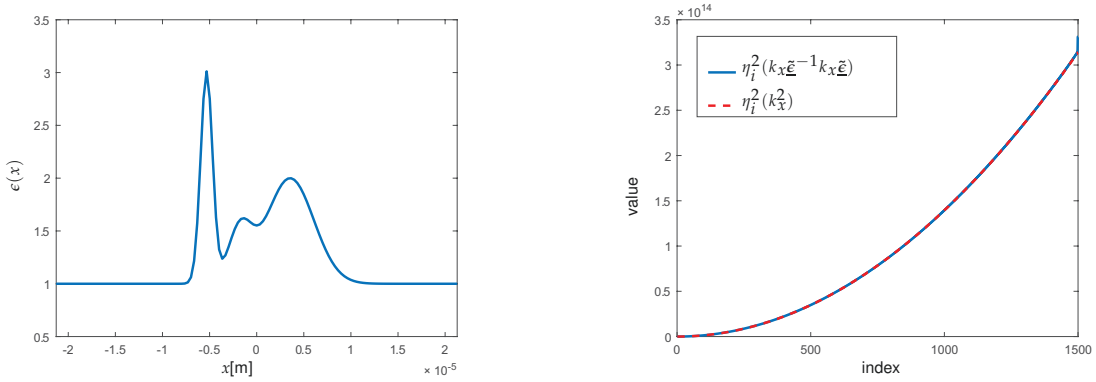


Figure 40: Matrix analysis. The left figure is the distribution of $\epsilon(x)$, while the right figure shows the comparison of sorted eigenvalue of two matrices.

Method 1

The useful theorems can be found in [112, 113].

Theorem 7. *Let matrices M_1 and M_2 be $n \times n$ Hermitian matrices and non-negative(all eigenvalues are non-negative). Then for every integer k, $1 \leqslant k \leqslant n$ the largest eigenvalue of matrix HP satisfies*

$$u_{\max}(M_1)u_{\min}(M_2) \leqslant u_{\max}(M_1M_2) \leqslant u_{\max}(M_1)u_{\max}(M_2), \tag{E.18}$$

with u denoting eigenvalue of matrix.

- $k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}$ can be considered as the production of matrices $A = k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x$ and $B = \underline{\tilde{\boldsymbol{\epsilon}}}$. Both A and B are hermitian matrices, following definition 2.
- If we well adjust the sampling and $k_x \neq 0$, it can be invertible. Another matrix $C = k_x^{-1}Ak_x = \underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x^2$ is similar to A, following the definition 3. A and C have identical eigenvalue, following theorem 6.

- Matrix C is the production of two non-negative hermitian matrices $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}$ and k_x^2, so the maximum eigenvalue of C satisfies $u_{\max}(C) \leqslant u_{\max}(\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}) u_{\max}(k_x^2) = k_{x\max}^2/\epsilon_{\min}$, and $u_{\max}(C) \geqslant 0$ following definition 1, 3, and theorem 1, 3, 7.
- Following theorem 6, $0 \leqslant u_{\max}(A) \leqslant u_{\max}(\underline{\tilde{\boldsymbol{\epsilon}}})$.
- Then the inequality of eigenvalue $u_{\max}(k_x \underline{\tilde{\boldsymbol{\epsilon}}}^{-1} k_x \underline{\tilde{\boldsymbol{\epsilon}}}) = u_{\max} AB$ can be derived, by using theorem 7

$$k_{x\max}^2 \leqslant u_{\max}(k_x \underline{\tilde{\boldsymbol{\epsilon}}}^{-1} k_x \underline{\tilde{\boldsymbol{\epsilon}}}) \leqslant k_{x\max}^2 \frac{\epsilon_{\max}}{\epsilon_{min}} \tag{E.19}$$

From method 1, we can only limit the eigenvalue in the range as shown in the inequality (E.19), still not prove our guess that eigenvalue $u_{\max}(k_x \underline{\tilde{\boldsymbol{\epsilon}}}^{-1} k_x \underline{\tilde{\boldsymbol{\epsilon}}}) \equiv k_{x\max}^2$

Method 2

As all the matrices comes from integral operator, so we go back to the integration and re-formulate the matrix

$$\begin{aligned}
&k_x \underline{\tilde{\boldsymbol{\epsilon}}}^{-1} k_x \underline{\tilde{\boldsymbol{\epsilon}}} \tilde{V}_1(k_x) \\
&= \frac{1}{(2\pi)^4} k_x \int \mathrm{d}x_2 \frac{1}{\epsilon(x_2)} \exp(-\mathrm{i}k_x x_2) \int \mathrm{d}k_{x2} k_{x2} \exp(\mathrm{i}k_{x2} x_2) \int \mathrm{d}x_1 \epsilon(x_1) \exp(-\mathrm{i}k_{x2} x_1) \int \mathrm{d}k_{x1} \tilde{V}_1(k_{x1}) \exp(\mathrm{i}k_{x1} x_1) \\
&= \frac{1}{(2\pi)^4} k_x \int \mathrm{d}k_{x1} \int \mathrm{d}x_1 \int \mathrm{d}x_2 \cdots \int \mathrm{d}k_{x2} k_{x2} \exp[\mathrm{i}k_{x2}(x_2 - x_1)] \qquad \text{exchange the order of integral} \\
&= \frac{1}{(2\pi)^3} k_x \int \mathrm{d}k_{x1} \int \mathrm{d}x_1 \int \mathrm{d}x_2 \cdots - \mathrm{i} \frac{\mathrm{d}}{\mathrm{d}(x_2 - x_1)} \delta(x_2 - x_1) \\
&= \frac{1}{(2\pi)^3} k_x \int \mathrm{d}k_{x1} \int \mathrm{d}x_2 \cdots \int \mathrm{d}x_1 \epsilon(x_1) \exp(\mathrm{i}k_{x1} x_1) \cdot (-\mathrm{i}) \delta'(x_2 - x_1) \\
&= \frac{1}{(2\pi)^2} k_x \int \mathrm{d}k_{x1} \int \mathrm{d}x_2 \cdots (-\mathrm{i}) \left[\frac{\mathrm{d}}{\mathrm{d}x_2} \epsilon(x_2) + \epsilon(x_2) \mathrm{i}k_{x1} \right] \exp(\mathrm{i}k_{x1} x_2) \qquad \frac{\mathrm{d}}{\mathrm{d}x} \delta(x) * f(x) = \frac{\mathrm{d}}{\mathrm{d}x} f(x) \\
&= \frac{1}{(2\pi)^2} k_x \int \mathrm{d}k_{x1} \cdots \int \mathrm{d}x_2 (-\mathrm{i}) \frac{1}{\epsilon(x_2)} \left[\frac{\mathrm{d}}{\mathrm{d}x_2} \epsilon(x_2) + \epsilon(x_2) \mathrm{i}k_{x1} \right] \exp[\mathrm{i}(k_{x1} - k_x) x_2] \\
&= \frac{1}{(2\pi)^2} k_x \int \mathrm{d}k_{x1} \cdots \int \mathrm{d}x_2 \left[(-\mathrm{i}) \frac{\mathrm{d}}{\mathrm{d}x_2} \ln \epsilon(x_2) + k_{x1} \right] \exp[\mathrm{i}(k_{x1} - k_x) x_2] \\
&= \frac{1}{2\pi} k_x \int \mathrm{d}k_{x1} \tilde{V}_1(k_{x1}) \left[-\mathrm{i}^2 (k_{x1} - k_x) \underline{\tilde{\boldsymbol{g}}}(k_{x1} - k_x) + k_{x1} \delta(k_{x1} - k_x) \right] \\
&= [k_x^2 \underline{\tilde{\boldsymbol{g}}} - k_x \underline{\tilde{\boldsymbol{g}}} k_x + k_x^2] \tilde{V}_1(k_x),
\end{aligned} \tag{E.20}$$

with

$$\underline{\tilde{\boldsymbol{g}}}(k_x, z) = \boldsymbol{\mathcal{F}}_k \ln \epsilon(x, z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}k_x \ln \epsilon(x, z) \exp(-\mathrm{i}k_x x). \tag{E.21}$$

Now the matrix $\underline{\mathbf{M}}_1^{\mathrm{TM}}$ can be rewritten as

$$\underline{\mathbf{M}}_1^{\mathrm{TM}} = -k_0^2 \underline{\tilde{\boldsymbol{\epsilon}}} + k_x^2 + k_x^2 \underline{\tilde{\boldsymbol{g}}} - k_x \underline{\tilde{\boldsymbol{g}}} k_x = \underline{\mathbf{M}}^{\mathrm{TE}} + k_x^2 \underline{\tilde{\boldsymbol{g}}} - k_x \underline{\tilde{\boldsymbol{g}}} k_x \tag{E.22}$$

If we analyse matrix $k_x^2 \underline{\tilde{\boldsymbol{g}}} - k_x \underline{\tilde{\boldsymbol{g}}} k_x = k_x (k_x \underline{\tilde{\boldsymbol{g}}} - \underline{\tilde{\boldsymbol{g}}} k_x)$ mathematically, it is very difficult. We tried several different mathematical definitions and theorems, but none of them gives us accurate desired proof. So let's try it again by using physical thinking.

$$\begin{aligned}
(k_x \underline{\tilde{\boldsymbol{g}}} - \underline{\tilde{\boldsymbol{g}}} k_x) \tilde{V}_1(k_x) &= k_x \underline{\tilde{\boldsymbol{g}}} \tilde{V}_1(k_x) - \underline{\tilde{\boldsymbol{g}}} k_x \tilde{V}_1(k_x) \\
&= \frac{1}{2\pi} \int \mathrm{d}x \exp^{-\mathrm{i}k_x x} \left\{ \frac{\mathrm{d}}{\mathrm{d}x} [\ln \epsilon(x) V_1(x)] - \ln \epsilon(x) \frac{\mathrm{d}}{\mathrm{d}x} V_1(x) \right\} \\
&= \frac{1}{2\pi} \int \mathrm{d}x \exp^{-\mathrm{i}k_x x} \left\{ \frac{\mathrm{d}}{\mathrm{d}x} \ln \epsilon(x) V_1(x) \right\}
\end{aligned} \tag{E.23}$$

As in current available material, ϵ is generally not larger than 10, the term $\frac{\mathrm{d}}{\mathrm{d}x}\ln\epsilon(x)$ gives very slightly effect to final result. So we neglect this term, which means $\eta_i^2(k_x^2\underline{\tilde{\boldsymbol{g}}} - k_x\underline{\tilde{\boldsymbol{g}}}k_x) \approx 0$. Fig. (41) shows $\eta_i^2(k_x^2\underline{\tilde{\boldsymbol{g}}} - k_x\underline{\tilde{\boldsymbol{g}}}k_x)$. Here

- k_x is equidistant sampled by 1000 points. $-k_0\sqrt{\epsilon_{\min}} \leqslant k_x \leqslant k_0\sqrt{\epsilon_{\min}}$
- There should be 1000 eigenvalues. The eigenvalues are sorted from minimum to maximum, and plot.

The real part is as what we guess, almost all eigenvalues are equal to 0. There are some non-zero value, but later in the final analysis of matrix $\underline{\mathbf{M}}_1^{\mathrm{TM}}$, the effects of them are negligible.

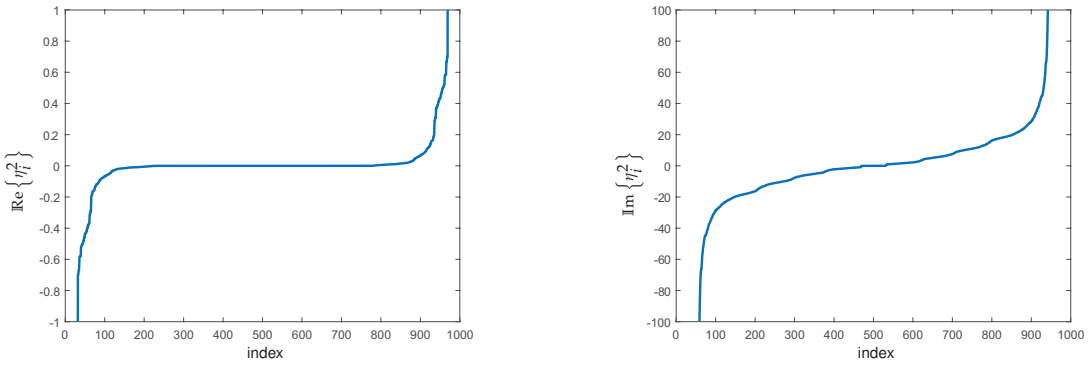


Figure 41: Eigenvalues of the additional matrix in TM case.

Now we calculate the eigenvalues of both $\underline{\mathbf{M}}^{\mathrm{TE}}$ and $\underline{\mathbf{M}}_1^{\mathrm{TM}}$. Fig. (42) is the calculation when $-k_0\sqrt{\epsilon_{\min}} \leqslant k_x \leqslant k_0\sqrt{\epsilon_{\min}}$ by 1000 sampling points, while Fig. (43) is $-1.5k_0\sqrt{\epsilon_{\min}} \leqslant k_x \leqslant 1.5k_0\sqrt{\epsilon_{\min}}$ with 1500 sampling points. Figs. (42) and (43) show that the eigenvalues of $\underline{\mathbf{M}}^{\mathrm{TE}}$ and $\underline{\mathbf{M}}_1^{\mathrm{TM}}$ are almost same. Although $\eta_i^2(\underline{\mathbf{M}}_1^{\mathrm{TM}})$ contains a imaginary part, but it is negligible, compared to the magnitude of the real part. Here we finally see that the effect of the matrix $k_x^2\underline{\tilde{\boldsymbol{g}}} - k_x\underline{\tilde{\boldsymbol{g}}}k_x$ can be neglected.

After neglecting the matrix $k_x^2\underline{\tilde{\boldsymbol{g}}} - k_x\underline{\tilde{\boldsymbol{g}}}k_x$, the eigenvalue analysis is identical with TE case. And the criteria when all eigenvalues are real and negative-valued, as Eq. (E.9), is also identical with TE case, as Eq. (E.15). Fig. (42) gives the eigenvalues when $|k_x| \leqslant k_0\sqrt{\epsilon_{\min}}$, which is given in Eq. (E.15). In this case, the largest eigenvalue approaching to 0, but still negative. Fig. (43) shows eigenvalues when k_x contains value which is larger than $k_0\sqrt{\epsilon_{\min}}$, and the largest eigenvalue is larger than 0, which will introduce instability of further algorithm.

There are two matrices in TM case (E.6) and (E.7). What is the relation between the two matrices?

Theorem 8. *Let matrix A is the transpose conjugate of matrix B, i.e., $A_{ij} = B_{ji}^{\star}$, with i, j index of matrix entries. The relation is denoted as $A^H = B$, Then the eigenvalue of A is the complex conjugate of B.*

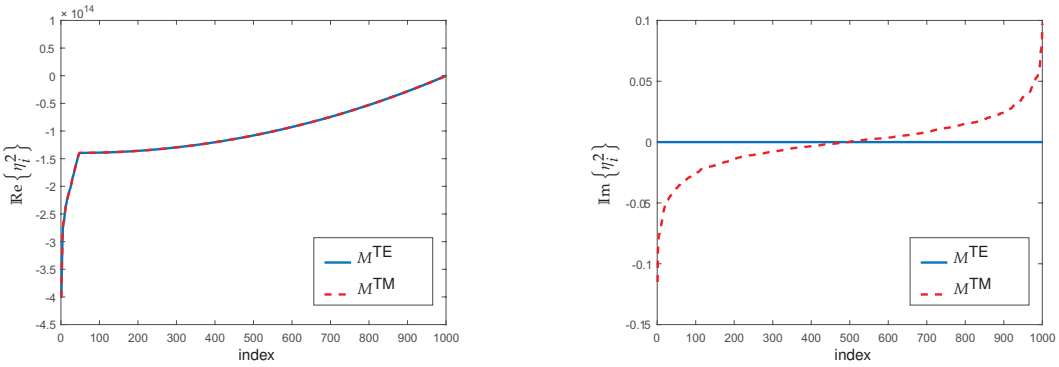


Figure 42: Eigenvalues of TE and TM matrices for $\epsilon(x)$ in Fig. (40). $|k_x| \leqslant k_0\sqrt{\epsilon_{\min}}$. The left figure is real part, and the right one is imaginary part.

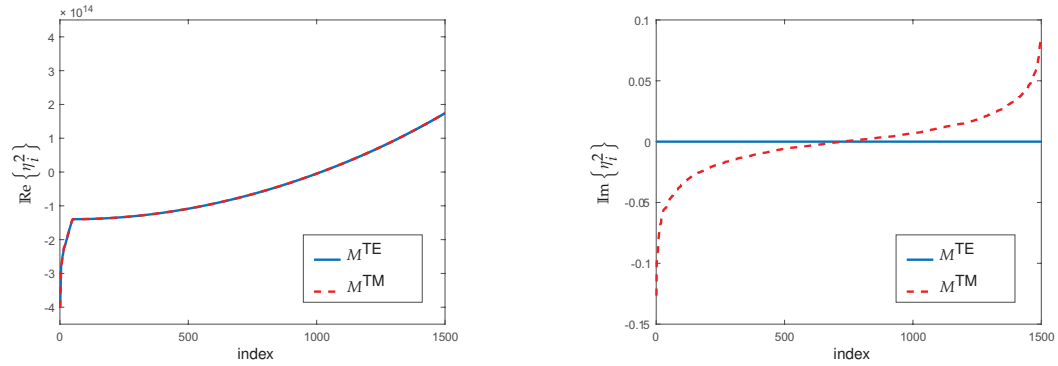


Figure 43: Eigenvalues of TE and TM matrices for $\epsilon(x)$ in Fig. (40). $|k_x| \leqslant 1.5k_0\sqrt{\epsilon_{\min}}$. The left figure is real part, and the right one is imaginary part.

The matrices $\underline{\tilde{\boldsymbol{\epsilon}}}$,$k_x$,and $\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}$ are Hermitian matrices, which satisfy definition 2. And from basic linear algebra, we know the conjugate transpose of $\{\underline{\mathbf{M}}_1^{\mathrm{TM}}\}^H$ is

$$\begin{aligned}\{\underline{\mathbf{M}}_1^{\mathrm{TM}}\}^H &= \{-k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}\}^H \\ &= \{-k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}\}^H + \{k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}}\}^H \\ &= -k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + \underline{\tilde{\boldsymbol{\epsilon}}}k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x = \underline{\mathbf{M}}_2^{\mathrm{TM}}.\end{aligned} \tag{E.24}$$

We have discussed that the imaginary part of $\eta_i^2(\underline{\mathbf{M}}_1^{\mathrm{TM}})$ is negligible. Together with theorem 8, we know that $\eta_i^2(\underline{\mathbf{M}}_1^{\mathrm{TM}}) = \eta_i^2(\underline{\mathbf{M}}_2^{\mathrm{TM}})$.

Conclusion in $y-$ invariant case

The criteria of the stable algorithm for is

$$|k_x|^2_{\max} \leqslant k_0^2\epsilon_{\min}. \tag{E.25}$$

To make sure that the further algorithm is stable, the ODE becomes

$$\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa},z) = \begin{cases} \underline{\mathbf{M}}(\boldsymbol{\kappa},z)\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa},z) & \text{for } ||\boldsymbol{\kappa}||^2 \leqslant k_0^2\epsilon_{\min} \\ 0 & \text{for } ||\boldsymbol{\kappa}||^2 > k_0^2\epsilon_{\min} \end{cases}, \tag{E.26}$$

with $\underline{\mathbf{M}}$ the matrix in Eq. (3.18).

E.3 ANALYSIS OF THE ODES FOR GENERAL 3D CASE

Now we consider the inhomogeneity of the GRIN media in the general case. As the analysis in the $y-$invariant case, we derive the second order differential equation

$$\frac{d^2}{dz^2}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\end{pmatrix}(\boldsymbol{\kappa},z) = \begin{bmatrix} -k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + k_y^2 + k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}} & k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_y\underline{\tilde{\boldsymbol{\epsilon}}} - k_xk_y \\ k_y\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x\underline{\tilde{\boldsymbol{\epsilon}}} - k_xk_y & -k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + k_x^2 + k_y\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_y\underline{\tilde{\boldsymbol{\epsilon}}} \end{bmatrix}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\end{pmatrix}(\boldsymbol{\kappa},z). \tag{E.27}$$

$$\frac{d^2}{dz^2}\begin{pmatrix}\tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z) = \begin{bmatrix} -k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + k_x^2 + \underline{\tilde{\boldsymbol{\epsilon}}}k_y\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_y & -\underline{\tilde{\boldsymbol{\epsilon}}}k_y\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x + k_xk_y \\ -\underline{\tilde{\boldsymbol{\epsilon}}}k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_y + k_xk_y & -k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + k_y^2 + \underline{\tilde{\boldsymbol{\epsilon}}}k_x\underline{\tilde{\boldsymbol{\epsilon}}}^{-1}k_x \end{bmatrix}\begin{pmatrix}\tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z). \tag{E.28}$$

After exchanging the integral order, Eqs. (E.27) and (E.28) can be rewritten as

$$\frac{d^2}{dz^2}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\end{pmatrix}(\boldsymbol{\kappa},z) = (-k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} + k_x^2 + k_y^2)\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\end{pmatrix}(\boldsymbol{\kappa},z) + \underline{\mathbf{M}}_{12}\begin{pmatrix}\tilde{V}_1\\ \tilde{V}_2\end{pmatrix}(\boldsymbol{\kappa},z). \tag{E.29}$$

with

$$\underline{\mathbf{M}}_{12} = \begin{bmatrix} k_x^2\underline{\tilde{\boldsymbol{g}}} - k_x\underline{\tilde{\boldsymbol{g}}}k_x & k_xk_y\underline{\tilde{\boldsymbol{g}}} - k_y\underline{\tilde{\boldsymbol{g}}}k_x \\ k_xk_y\underline{\tilde{\boldsymbol{g}}} - k_x\underline{\tilde{\boldsymbol{g}}}k_y & k_y^2\underline{\tilde{\boldsymbol{g}}} - k_y\underline{\tilde{\boldsymbol{g}}}k_y \end{bmatrix} \tag{E.30}$$

$$\frac{d^2}{dz^2}\begin{pmatrix}\tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z)=(-k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}+k_x^2+k_y^2)\begin{pmatrix}\tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z)+\underline{\mathbf{M}}_{45}\begin{pmatrix}\tilde{V}_4\\ \tilde{V}_5\end{pmatrix}(\boldsymbol{\kappa},z). \tag{E.31}$$

with

$$\underline{\mathbf{M}}_{45}=\begin{bmatrix}\underline{\tilde{\boldsymbol{g}}}k_y^2-k_y\underline{\tilde{\boldsymbol{g}}}k_y & \underline{\tilde{\boldsymbol{g}}}k_xk_y-k_y\underline{\tilde{\boldsymbol{g}}}k_x\\ \underline{\tilde{\boldsymbol{g}}}k_xk_y-k_x\underline{\tilde{\boldsymbol{g}}}k_y & \underline{\tilde{\boldsymbol{g}}}k_x^2-k_x\underline{\tilde{\boldsymbol{g}}}k_x\end{bmatrix} \tag{E.32}$$

The entries of matrices $\underline{\mathbf{M}}_{12}$ and $\underline{\mathbf{M}}_{45}$ are all in the form of subtraction between two similar matrices, which is discussed in the last section. As we discussed in last section, the effects of $\underline{\mathbf{M}}_{12}$ and $\underline{\mathbf{M}}_{45}$ are neglected.

The criteria of the stable algorithm for is

$$||\boldsymbol{\kappa}||_{\max}^2\leqslant k_0^2\epsilon_{\min}. \tag{E.33}$$

To make sure that the further algorithm is stable, the ODE becomes

$$\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa},z)=\begin{cases}\underline{\mathbf{M}}(\boldsymbol{\kappa},z)\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa},z) & \text{for } ||\boldsymbol{\kappa}||^2\leqslant k_0^2\epsilon_{\min}\\ 0 & \text{for } ||\boldsymbol{\kappa}||^2>k_0^2\epsilon_{\min}\end{cases}, \tag{E.34}$$

with $\underline{\mathbf{M}}$ the matrix in Eq. (3.18).

E.4 ANALYSIS OF THE ODES AFTER EXTRACTING FAST-CHANGING PHASE

First of all, we still just check the stability in a tiny z step, which means z-variance of $\epsilon(\boldsymbol{r})$ is negligible. Second, we rewrite the ODE (3.43) into a second-order ODE as

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\begin{pmatrix}\tilde{U}_1\\ \tilde{U}_2\end{pmatrix}+2\mathrm{i}k_0\bar{n}\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix}\tilde{U}_1\\ \tilde{U}_2\end{pmatrix}=(-k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}+k_0^2\bar{n}^2+k_x^2+k_y^2)\begin{pmatrix}\tilde{U}_1\\ \tilde{U}_2\end{pmatrix}+\underline{\mathbf{M}}_{12}\begin{pmatrix}\tilde{U}_1\\ \tilde{U}_2\end{pmatrix}. \tag{E.35}$$

with

$$\underline{\mathbf{M}}_{12}=\begin{bmatrix}k_x^2\underline{\tilde{\boldsymbol{g}}}-k_x\underline{\tilde{\boldsymbol{g}}}k_x & k_xk_y\underline{\tilde{\boldsymbol{g}}}-k_y\underline{\tilde{\boldsymbol{g}}}k_x\\ k_xk_y\underline{\tilde{\boldsymbol{g}}}-k_x\underline{\tilde{\boldsymbol{g}}}k_y & k_y^2\underline{\tilde{\boldsymbol{g}}}-k_y\underline{\tilde{\boldsymbol{g}}}k_y\end{bmatrix} \tag{E.36}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\begin{pmatrix}\tilde{U}_4\\ \tilde{U}_5\end{pmatrix}+2\mathrm{i}k_0\bar{n}\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix}\tilde{U}_4\\ \tilde{U}_5\end{pmatrix}=(-k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}+k_0^2\bar{n}^2+k_x^2+k_y^2)\begin{pmatrix}\tilde{U}_4\\ \tilde{U}_5\end{pmatrix}+\underline{\mathbf{M}}_{45}\begin{pmatrix}\tilde{U}_4\\ \tilde{U}_5\end{pmatrix}. \tag{E.37}$$

with

$$\underline{\mathbf{M}}_{45}=\begin{bmatrix}\underline{\tilde{\boldsymbol{g}}}k_y^2-k_y\underline{\tilde{\boldsymbol{g}}}k_y & \underline{\tilde{\boldsymbol{g}}}k_xk_y-k_y\underline{\tilde{\boldsymbol{g}}}k_x\\ \underline{\tilde{\boldsymbol{g}}}k_xk_y-k_x\underline{\tilde{\boldsymbol{g}}}k_y & \underline{\tilde{\boldsymbol{g}}}k_x^2-k_x\underline{\tilde{\boldsymbol{g}}}k_x\end{bmatrix} \tag{E.38}$$

The entries of matrices $\underline{\mathbf{M}}_{12}$ and $\underline{\mathbf{M}}_{45}$ are all in the form of subtraction between two similar matrices, which is discussed in the last section. As we discussed in last section, the effects of $\underline{\mathbf{M}}_{12}$ and $\underline{\mathbf{M}}_{45}$ are neglected. And Eqs. (E.35) and (E.37) can be rewritten as

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\tilde{U}_i+2\mathrm{i}k_0\bar{n}\frac{\mathrm{d}}{\mathrm{d}z}\tilde{U}_i+(k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}-k_0^2\bar{n}^2-k_x^2-k_y^2)\tilde{U}_i=0, \tag{E.39}$$

with $i = 1, 2, 4, 5$.

Eigenvalue η_i of Eq. (E.39) is calculated by

$$\begin{aligned} \eta &= \frac{-2\mathrm{i}k_0\bar{n} \pm \sqrt{(2\mathrm{i}k_0\bar{n})^2 - 4(k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} - k_0^2\bar{n}^2 - k_x^2 - k_y^2)}}{2} \\ &= -2\mathrm{i}k_0\bar{n} \pm \mathrm{i}\sqrt{k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} - k_x^2 - k_y^2} \end{aligned} \tag{E.40}$$

To analyze Eq. (E.40), we need the matrix definition and theorem from last section.

- The matrix $k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}$ is Toeplitz matrix (definition 3), and its eigenvalue $k_0^2\epsilon_{\min} \leqslant \eta(k_0^2\underline{\tilde{\boldsymbol{\epsilon}}}) \leqslant k_0^2\epsilon_{\max}$ (theorem 3).
- The matrices k_x^2 and k_y^2 are diagonal matrices (definition 1) and the eigenvalue $0 \leqslant \eta(k_x^2) \leqslant |k_x|^2_{\max}$ and $0 \leqslant \eta(k_y^2) \leqslant |k_y|^2_{\max}$ (theorem 1)
- All matrices mentioned in the above terms are Hermitian matrices (definition 2) and the eigenvalue of sum of Hermitian matrices is $\eta(k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} - k_x^2 - k_y^2) \geqslant k_0^2\epsilon_{\min} - |k_x|^2_{\max} - |k_y|^2_{\max}$ (theorem 4)
- When $\eta(k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} - k_x^2 - k_y^2) \geqslant 0$, $\eta(\sqrt{k_0^2\underline{\tilde{\boldsymbol{\epsilon}}} - k_x^2 - k_y^2}) \subset \mathbb{R}$.

Therefore, η in Eq. (E.40) consists of a pure imaginary values when

$$||\boldsymbol{\kappa}||^2 \leqslant k_0^2\epsilon_{\min}. \tag{E.41}$$

F

DERIVATION OF EIKONAL EQUATION FROM MAXWELL'S EQUATIONS

After plugging the field representations (4.20) and (4.21) into Maxwell equations (2.23), we get

$$\nabla \times \boldsymbol{E}_0(\boldsymbol{r}) + i\nabla\psi(\boldsymbol{r}) \times \boldsymbol{E}_0(\boldsymbol{r}) - i\omega\mu_0\boldsymbol{H}_0(\boldsymbol{r}) = 0, \tag{F.1}$$

$$\nabla \times \boldsymbol{H}_0(\boldsymbol{r}) + i\nabla\psi(\boldsymbol{r}) \times \boldsymbol{H}_0(\boldsymbol{r}) + i\omega\epsilon_0\epsilon_r(\boldsymbol{r})\boldsymbol{E}_0(\boldsymbol{r}) = 0, \tag{F.2}$$

$$\nabla \cdot \boldsymbol{E}_0(\boldsymbol{r}) + i\boldsymbol{E}_0(\boldsymbol{r}) \cdot \nabla\psi(\boldsymbol{r}) + \boldsymbol{E}_0(\boldsymbol{r}) \cdot \nabla \ln \epsilon_r(\boldsymbol{r}) = 0, \tag{F.3}$$

$$\nabla \cdot \boldsymbol{H}_0(\boldsymbol{r}) + i\boldsymbol{H}_0(\boldsymbol{r}) \cdot \nabla\psi(\boldsymbol{r}) = 0. \tag{F.4}$$

By applying the geometric field approximations for slowly varying refractive index, mathematically represented as $\boldsymbol{\nabla}\epsilon_r(\boldsymbol{r}) \approx 0$, the term which contains $\nabla \cdot \boldsymbol{E}_0(\boldsymbol{r})$, $\nabla \cdot \boldsymbol{H}_0(\boldsymbol{r})$ and $\nabla\epsilon_r(\boldsymbol{r})$ can be neglected. Then we get the geometric field equations [53]

$$\nabla\psi(\boldsymbol{r}) \times \boldsymbol{E}_0(\boldsymbol{r}) = \omega\mu_0\boldsymbol{H}_0(\boldsymbol{r}), \tag{F.5}$$

$$\nabla\psi(\boldsymbol{r}) \times \boldsymbol{H}_0(\boldsymbol{r}) = -\omega\epsilon_0\epsilon_r(\boldsymbol{r})\boldsymbol{E}_0(\boldsymbol{r}), \tag{F.6}$$

$$\nabla\psi(\boldsymbol{r}) \cdot \boldsymbol{E}_0(\boldsymbol{r}) = 0, \tag{F.7}$$

$$\nabla\psi(\boldsymbol{r}) \cdot \boldsymbol{H}_0(\boldsymbol{r}) = 0. \tag{F.8}$$

Eqs. (F.5-F.8) describe the interaction of a geometric field with a GRIN medium, and their solution yields the theory for field propagation in GRIN medium.

In order to solve Eqs. (F.5-F.8), we first try to understand the geometric field by comparing these equations with those for plane waves, which are derived by using a plane-wave ansatz in Maxwell's equations (2.23) for a homogeneous medium. The plane wave ansatz is

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}_0 \exp(i\boldsymbol{k} \cdot \boldsymbol{r}), \tag{F.9}$$

and

$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}_0 \exp(i\boldsymbol{k} \cdot \boldsymbol{r}), \tag{F.10}$$

with a constant wave vector $\boldsymbol{k}$ and the complex amplitude of the electric field $\boldsymbol{E}_0$ and that of the magnetic field $\boldsymbol{H}_0$.

Plugging the plane wave ansatz (F.9) and (F.10) into Maxwell's equations (2.23), we can derive the plane wave equations for homogeneous media

$$\boldsymbol{k} \times \boldsymbol{E}_0 = \omega\mu_0\boldsymbol{H}_0, \tag{F.11}$$

$$\boldsymbol{k} \times \boldsymbol{H}_0 = -\omega\epsilon_0 n^2\boldsymbol{E}_0, \tag{F.12}$$

$$\boldsymbol{k} \cdot \boldsymbol{E}_0 = 0, \tag{F.13}$$

$$\boldsymbol{k} \cdot \boldsymbol{H}_0 = 0, \tag{F.14}$$

where n denotes the refractive index.

Comparing Eqs. (F.5-F.8) to Eqs. (F.11-F.14), we see that $\nabla\psi(\boldsymbol{r})$ in the care of the geometric field exhibits a role analogous to that of $\boldsymbol{k}$ for the plane wave

$$\nabla\psi(\boldsymbol{r}) \sim \boldsymbol{k}. \tag{F.15}$$

The plane wave ansatz (F.9) and (F.10) contains the wave vector $\boldsymbol{k}$, while Eqs. (4.20) and (4.21) are rewritten using the gradient theorem to include the term $\nabla\psi(\boldsymbol{r})$,

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}_0(\boldsymbol{r}) \exp\left\{ i \left[\psi(\boldsymbol{r}_0) + \int_{\boldsymbol{r}_0}^{\boldsymbol{r}} \nabla\psi(\boldsymbol{r}) \cdot d\boldsymbol{r} \right] \right\}, \tag{F.16}$$

and

$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}_0(\boldsymbol{r}) \exp\left\{ i \left[\psi(\boldsymbol{r}_0) + \int_{\boldsymbol{r}_0}^{\boldsymbol{r}} \nabla\psi(\boldsymbol{r}) \cdot d\boldsymbol{r} \right] \right\}. \tag{F.17}$$

Here $\psi(\boldsymbol{r}_0)$ denotes the phase at position $\boldsymbol{r}_0$.

Then we use Eqs. (F.16) and (F.17) to solve Eqs. (F.5-F.8), following the same procedure that is employed to solve Eqs. (F.11-F.14) via the plane-wave ansatz (F.9) and (F.10): Calculating $\boldsymbol{k} \times \boldsymbol{k} \times \boldsymbol{E}_0$ using Eqs. (F.11-F.12) yields the well known dispersion relation [53, 114]

$$\boldsymbol{k}^2 = k_0^2 \epsilon_r. \tag{F.18}$$

with $k_0 = 2\pi/\lambda$, where λ denotes the vacuum wavelength. Therefore, a plane wave (F.9) and (F.10) with a wave vector $\boldsymbol{k}$ which fulfills (F.18) solves Eqs. (F.11-F.14). Furthermore, we define

$$\boldsymbol{k} = k_0 \sqrt{\epsilon_r} \hat{\boldsymbol{s}}, \tag{F.19}$$

where $\hat{\boldsymbol{s}}$ denotes the propagation direction of the plane wave. From Eqs. (F.11-F.14), we also know that $\hat{\boldsymbol{s}}$ is perpendicular to the vectorial electromagnetic field $\boldsymbol{E}(\boldsymbol{r})$ and $\boldsymbol{H}(\boldsymbol{r})$, whose directions are identical to those of the corresponding amplitudes $\boldsymbol{E}_0$ and $\boldsymbol{H}_0$.

Analogously for a geometric field, we calculate $\nabla\psi(\boldsymbol{r}) \times \nabla\psi(\boldsymbol{r}) \times \boldsymbol{E}_0(\boldsymbol{r})$ using Eqs. (F.5-F.6), so that the **eikonal equation** is derived,

$$(\nabla\psi(\boldsymbol{r}))^2 = k_0^2 \epsilon_r(\boldsymbol{r}). \tag{F.20}$$

Eqs. (F.16) and (F.17) with $\nabla\psi(\boldsymbol{r})$ according to Eq. (F.20) are solutions of Eqs. (F.5-F.8). Similarly to what we did with plane waves, we define

$$\nabla\psi(\boldsymbol{r}) = \boldsymbol{k}(\boldsymbol{r}) = k_0 \sqrt{\epsilon_r}(\boldsymbol{r}) \hat{\boldsymbol{s}}(\boldsymbol{r}), \tag{F.21}$$

where $\hat{\boldsymbol{s}}(\boldsymbol{r})$ denotes the propagation direction of geometric field locally at position $\boldsymbol{r}$. From Eqs. (F.5-F.8), we also know that $\hat{\boldsymbol{s}}(\boldsymbol{r})$ is perpendicular to the vectorial electromagnetic field $\boldsymbol{E}(\boldsymbol{r})$ and $\boldsymbol{H}(\boldsymbol{r})$, whose directions are identical with those of the corresponding amplitudes $\boldsymbol{E}_0(\boldsymbol{r})$ and $\boldsymbol{H}_0(\boldsymbol{r})$. Eq. (F.21) is the vectorial version of **eikonal equation**.

G

SOME MATHEMATICAL DERIVATION RELATED TO RK x-DOMAIN METHOD

G.1 DERIVATION OF ODE IN k-DOMAIN FOR GEOMETRIC FIELD

Substituting Eq. (4.2) into the ODE (4.11), and calculate $\boldsymbol{M}_{11}\tilde{V}_1$ by the following process. We firstly calculating the convolution $\tilde{\underline{\boldsymbol{\epsilon}}}\tilde{V}_1(\boldsymbol{\kappa},z)$

$$\begin{aligned}\tilde{\underline{\boldsymbol{\epsilon}}}\tilde{V}_1(\boldsymbol{\kappa},z) &= \boldsymbol{\mathcal{F}}_k[\epsilon(\boldsymbol{\rho},z)V_1(\boldsymbol{\rho},z)] \quad \epsilon(\boldsymbol{\rho},z)U_1(\boldsymbol{\rho},z) \text{ is slowly varying}\\ &= \boldsymbol{\mathcal{F}}_k^{\mathrm{h}}[\epsilon(\boldsymbol{\rho},z)U_1(\boldsymbol{\rho},z)\exp[\mathrm{i}\psi(\boldsymbol{\rho},z)]] \quad \epsilon \text{ doesn't change } \psi\\ &\overset{(4.2)}{=} a[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]\epsilon[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]U_1[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]\exp\{\mathrm{i}\psi[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]-\mathrm{i}\boldsymbol{\kappa}\cdot\boldsymbol{\rho}\}\\ &= \epsilon[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]\tilde{V}_1(\boldsymbol{\kappa},z).\end{aligned} \tag{G.1}$$

For geometric field, the convolution between $\tilde{\epsilon}(\boldsymbol{\kappa})$ and $\tilde{V}_1(\boldsymbol{\kappa})$ becomes a multiplication. Then we calculate $\boldsymbol{\mathcal{F}}_k^{-1}[k_x\tilde{\underline{\boldsymbol{\epsilon}}}\tilde{V}_1](\boldsymbol{\rho},z)$

$$\begin{aligned}\boldsymbol{\mathcal{F}}_k^{-1}[k_x\tilde{\underline{\boldsymbol{\epsilon}}}\tilde{V}_1](\boldsymbol{\rho},z) &= \boldsymbol{\mathcal{F}}_k^{-1}\{k_x\epsilon[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]\tilde{V}_1(\boldsymbol{\kappa},z)\}(\boldsymbol{\rho},z) \quad k_x\underline{\epsilon}\tilde{A}_1 \text{ is slowly varying}\\ &= \boldsymbol{\mathcal{F}}_k^{\mathrm{h},-1}\{k_x\epsilon[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]\tilde{A}_1(\boldsymbol{\kappa},z)\exp[\mathrm{i}\tilde{\psi}(\boldsymbol{\kappa},z)]\}(\boldsymbol{\rho},z) \quad k_x\epsilon \text{ doesn't change } \tilde{\psi}\\ &= k_x(\boldsymbol{\rho},z)\epsilon(\boldsymbol{\rho},z)V_1(\boldsymbol{\rho},z),\end{aligned} \tag{G.2}$$

and finally calculating $\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_x\tilde{\underline{\boldsymbol{\epsilon}}}\tilde{V}_1(\boldsymbol{\kappa},z)$

$$\begin{aligned}\tilde{\underline{\boldsymbol{\epsilon}}}^{-1}k_x\tilde{\underline{\boldsymbol{\epsilon}}}\tilde{V}_1(\boldsymbol{\kappa},z) &= \boldsymbol{\mathcal{F}}_k[\epsilon^{-1}(\boldsymbol{\rho},z)k_x(\boldsymbol{\rho},z)\epsilon(\boldsymbol{\rho},z)V_1(\boldsymbol{\rho},z)]\\ &= \boldsymbol{\mathcal{F}}_k[k_x(\boldsymbol{\rho},z)V_1(\boldsymbol{\rho},z)] \quad k_x \text{ doesn't change } \psi\\ &= k_x\tilde{V}_1(\boldsymbol{\kappa},z).\end{aligned} \tag{G.3}$$

Substituting of Eqs. (G.1) and (G.3) into $\boldsymbol{M}_{11}\tilde{V}_1(\boldsymbol{\kappa},z)$, we get

$$\boldsymbol{M}_{11}\tilde{V}_1(\boldsymbol{\kappa},z) = \{-k_0^2\epsilon[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]+k_y^2+k_x^2\}\tilde{V}_1(\boldsymbol{\kappa},z). \tag{G.4}$$

Using the same calculation process as Eqs. (G.1-G.4), the 2nd-order ODE (4.11) can be rewritten as

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa},z) = -k_z^2[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]\tilde{\boldsymbol{V}}_\perp(\boldsymbol{\kappa},z), \tag{G.5}$$

with $k_z[\boldsymbol{\rho}(\boldsymbol{\kappa}),z] = k_0^2\epsilon[\boldsymbol{\rho}(\boldsymbol{\kappa}),z]-k_y^2-k_x^2$ and $\tilde{\boldsymbol{V}}_\perp = (\tilde{V}_1,\tilde{V}_2,\tilde{V}_4,\tilde{V}_5)^T$.

G.2 ODE OF RAY PATH

To derive the first ODE in x-domain, which is relevant to the ray path, we differentiate Eq. (4.19) with respect to arc length s [53] to obtain the equation

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}[\boldsymbol{k}(\boldsymbol{r})] &= \frac{\mathrm{d}}{\mathrm{d}s}[\nabla\psi(\boldsymbol{r})]\\ &= \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s}\cdot\nabla[\nabla\psi(\boldsymbol{r})]\\ &= \frac{1}{k_0\sqrt{\epsilon(\boldsymbol{r})}}\nabla\psi(\boldsymbol{r})\cdot[\nabla\psi(\boldsymbol{r})]\\ &= \frac{1}{2k_0\sqrt{\epsilon(\boldsymbol{r})}}\nabla\left\{[\nabla\psi(\boldsymbol{r})]^2\right\}\\ &= k_0\nabla\sqrt{\epsilon(\boldsymbol{r})}\end{aligned} \tag{G.6}$$

After eliminating k_0 on both sides, the ray equation is derived

$$\frac{\mathrm{d}}{\mathrm{d}s}[\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})] = \boldsymbol{\nabla}\sqrt{\epsilon(\boldsymbol{r})}. \tag{G.7}$$

G.3 ODE OF THE NORMALIZED FIELD VECTOR

To achieve an ODE of the normalized field vector $\hat{\boldsymbol{u}}(\boldsymbol{r})$, we differentiate Eq. (4.34) with respect to arc length s to obtain

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}s}[\boldsymbol{k}(\boldsymbol{r})\cdot\boldsymbol{E}_0(\boldsymbol{r})]\\ =\ &\frac{\mathrm{d}}{\mathrm{d}s}[k_0\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\boldsymbol{E}_0(\boldsymbol{r})]\\ =\ &k_0\{\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{E}_0(\boldsymbol{r})+\boldsymbol{E}_0(\boldsymbol{r})\cdot\underbrace{\frac{\mathrm{d}}{\mathrm{d}s}[\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})]}_{\text{using Eq. (4.45)}}\}\\ =\ &k_0\{\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{E}_0(\boldsymbol{r})+\boldsymbol{E}_0(\boldsymbol{r})\cdot\boldsymbol{\nabla}\sqrt{\epsilon(\boldsymbol{r})}\}\\ =\ &k_0\{\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\frac{\mathrm{d}}{\mathrm{d}s}[||\boldsymbol{E}_0(\boldsymbol{r})||\hat{\boldsymbol{u}}(\boldsymbol{r})]+||\boldsymbol{E}_0(\boldsymbol{r})||\hat{\boldsymbol{u}}(\boldsymbol{r})\cdot\boldsymbol{\nabla}\sqrt{\epsilon(\boldsymbol{r})}\}\\ =\ &k_0\{||\boldsymbol{E}_0(\boldsymbol{r})||\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{u}}(\boldsymbol{r})+\sqrt{\epsilon(\boldsymbol{r})}\underbrace{\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\hat{\boldsymbol{u}}(\boldsymbol{r})}_{\boldsymbol{k}(\boldsymbol{r})\cdot\boldsymbol{E}_0(\boldsymbol{r})=0\Rightarrow\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\hat{\boldsymbol{u}}(\boldsymbol{r})=0}\frac{\mathrm{d}}{\mathrm{d}s}||\boldsymbol{E}_0(\boldsymbol{r})||\\ &+||\boldsymbol{E}_0(\boldsymbol{r})||\hat{\boldsymbol{u}}(\boldsymbol{r})\cdot\nabla\sqrt{\epsilon(\boldsymbol{r})}\}\\ =\ &k_0||\boldsymbol{E}_0(\boldsymbol{r})||\{\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{u}}(\boldsymbol{r})+\hat{\boldsymbol{u}}(\boldsymbol{r})\cdot\boldsymbol{\nabla}\sqrt{\epsilon(\boldsymbol{r})}\}\\ =\ &0,\end{aligned} \tag{G.8}$$

Therefore,

$$\sqrt{\epsilon(\boldsymbol{r})}\hat{\boldsymbol{s}}(\boldsymbol{r})\cdot\frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{u}}(\boldsymbol{r})+\hat{\boldsymbol{u}}(\boldsymbol{r})\cdot\nabla\sqrt{\epsilon(\boldsymbol{r})}=0 \tag{G.9}$$

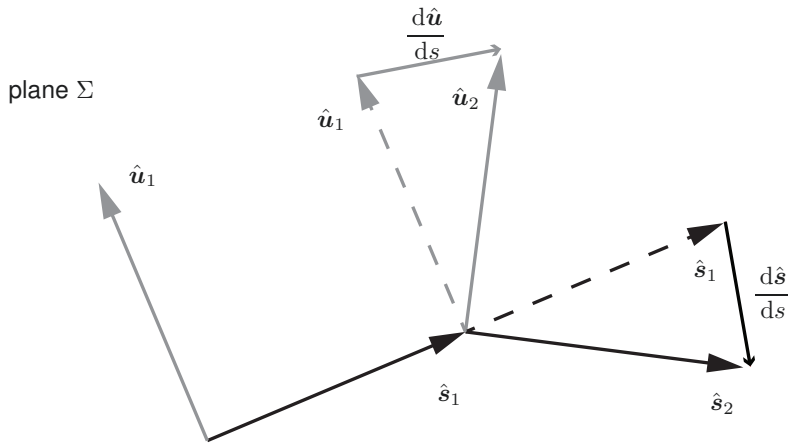


Figure 44: Geometric relation of the adjacent ray segments and the related normalized field vectors. $\hat{\boldsymbol{s}}_1$ and $\hat{\boldsymbol{s}}_2$ are two adjacent ray segments, while $\hat{\boldsymbol{u}}_1$ and $\hat{\boldsymbol{u}}_2$ are the related normalized field vectors.

The information of $\frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{u}}(\boldsymbol{r})$ is still not enough, we need to add more physical discussion. More specifically, we need to know the relationship between $\hat{\boldsymbol{s}}(\boldsymbol{r})$ and $\frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{u}}(\boldsymbol{r})$. In Fig. 44, we use the geometric relation, and boundary condition to prove $\frac{\mathrm{d}\hat{\boldsymbol{s}}(\boldsymbol{r})}{\mathrm{d}s} \parallel \hat{\boldsymbol{s}}(\boldsymbol{r})$.

- $\boldsymbol{k}(\boldsymbol{r}) \cdot \boldsymbol{E}_0(\boldsymbol{r}) = 0 \Rightarrow \hat{\boldsymbol{s}}(\boldsymbol{r}) \cdot \hat{\boldsymbol{u}}(\boldsymbol{r}) = 0 \Rightarrow \hat{\boldsymbol{s}}(\boldsymbol{r}) \perp \hat{\boldsymbol{u}}(\boldsymbol{r})$ is always true.
- Any two adjacent steps ray segments form one plane Σ, as shown in Fig. 44.
- As $\hat{\boldsymbol{s}}(\boldsymbol{r})$ is a normalized vector, $\frac{\mathrm{d}}{\mathrm{d}s}(\hat{\boldsymbol{s}}(\boldsymbol{r}) \cdot \hat{\boldsymbol{s}}(\boldsymbol{r})) = 2\hat{\boldsymbol{s}}(\boldsymbol{r}) \cdot \frac{\mathrm{d}\hat{\boldsymbol{s}}(\boldsymbol{r})}{\mathrm{d}s} = 0 \Rightarrow \frac{\mathrm{d}\hat{\boldsymbol{s}}(\boldsymbol{r})}{\mathrm{d}s} \perp \hat{\boldsymbol{s}}(\boldsymbol{r})$. Here $\frac{\mathrm{d}\hat{\boldsymbol{s}}}{\mathrm{d}s}$ is also in plane Σ.
- We can decompose the electric field vector related to $\hat{\boldsymbol{s}}_1$ into two modes: one is in plane Σ, and the other one is perpendicular to Σ. Then we could discuss the two modes separately.
  1. If $\hat{\boldsymbol{u}}_1$ is in plane Σ

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{u}} \times \hat{\boldsymbol{s}} = \frac{\mathrm{d}}{\mathrm{d}s}(\hat{\boldsymbol{u}} \times \hat{\boldsymbol{s}}) - \hat{\boldsymbol{u}} \times \frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{s}} \tag{G.10}$$

     In Fig. (44), the adjacent $\hat{\boldsymbol{u}}_1 \times \hat{\boldsymbol{s}}_1$ and $\hat{\boldsymbol{u}}_2 \times \hat{\boldsymbol{s}}_2$ are the normal vector of plane Σ, and not change, so $\frac{\mathrm{d}}{\mathrm{d}s}(\hat{\boldsymbol{u}} \times \hat{\boldsymbol{s}}) = 0$. Meanwhile, from the previous items, we

know $\hat{\boldsymbol{u}} \perp \hat{\mathbf{s}}$, and $\frac{\mathrm{d}}{\mathrm{d}s}\hat{\mathbf{s}} \perp \hat{\mathbf{s}}$, and all three vectors are in the same plane Σ, so $\hat{\boldsymbol{u}} \times \frac{\mathrm{d}}{\mathrm{d}s}\hat{\mathbf{s}} = 0$. Then we know that in Eq. (G.10), $\frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{u}} \times \hat{\mathbf{s}} = 0$

2. If $\hat{\boldsymbol{u}}_1 \perp \Sigma$. $\hat{\boldsymbol{u}}_2$ has no reason to change. So $\frac{\mathrm{d}}{\mathrm{d}s}\hat{\boldsymbol{u}} = 0$.

For general case, any $\hat{\boldsymbol{u}}(\boldsymbol{r})$ can be decomposed into Case 1 and Case 2, $\frac{\mathrm{d}\hat{\boldsymbol{u}}(\boldsymbol{r})}{\mathrm{d}s} \parallel \hat{\mathbf{s}}(\boldsymbol{r})$ is always true.

Then Eq. (G.9) can be rewritten as

$$\sqrt{\epsilon(\boldsymbol{r})}\frac{\mathrm{d}\hat{\boldsymbol{u}}(\boldsymbol{r})}{\mathrm{d}s} = -\left[\hat{\boldsymbol{u}}(\boldsymbol{r}) \cdot \boldsymbol{\nabla}\sqrt{\epsilon(\boldsymbol{r})}\right]\hat{\mathbf{s}}(\boldsymbol{r}) \,. \tag{G.11}$$

BIBLIOGRAPHY

[1] Eugene Hecht. *Optics*. Addison-Wesley; 4 edition, 2001.

[2] JC Maxwell. "Solution of a problem." In: *Cambridge & Dublin Mathematical Journal* 9 (1854), pp. 9–11.

[3] R.K. Luneburg. *Mathematical Theory of Optics*. University of California Press, 1964.

[4] Samuel P Morgan. "General solution of the Luneberg lens problem." In: *J. Appl. Phys.* 29.9 (1958), pp. 1358–1368.

[5] Jacek Sochacki. "Exact analytical solution of the generalized Luneburg lens problem." In: *J. Opt. Soc. Am.* 73.6 (1983), pp. 789–795.

[6] Jacek Sochacki. "Generalized Luneburg lens problem solution: a comment." In: *J. Opt. Soc. Am.* 73.12 (1983), pp. 1839–1839.

[7] Jacek Sochacki. "Multiple-foci Luneburg lenses." In: *Appl. Opt.* 23.23 (1984), pp. 4444–4449.

[8] R. W. Wood. *Physical Optics*. Macmillan, New York, 1905. Chap. 4, pp. 88–89.

[9] Peter McCarthy and Duncan T Moore. "Optical design with gradient-index elements constrained to real material properties." In: *Optical Fabrication and Testing*. Optical Society of America. 2012, OTu4D–2.

[10] Peter McCarthy. "Gradient-index materials, design, and metrology for broadband imaging systems." In: (2015).

[11] Duncan T. Moore. *Handbook of optics*. Ed. by Michael Bass. Vol. I. Optical society of America, 1996. Chap. 24, pp. 24.1–24.10.

[12] RHH Kröger, MCW Campbell, RD Fernald, and H-J Wagner. "Multifocal lenses compensate for chromatic defocus in vertebrate eyes." In: *J. Comp. Physiol. A* 184.4 (1999), pp. 361–369.

[13] Graded-index Fibers, RP Photonics Encyclopedia, https://www.rp-photonics.com.

[14] Gradient Lens Corporation, https://www.gradientlens.com.

[15] GRINTECH GmbH,http://www.grintech.de.

[16] Wei Yan, Xiao Peng, Danying Lin, Qi Wang, Jian Gao, Teng Luo, Jie Zhou, Tong Ye, Junle Qu, and Hanben Niu. "Fluorescence microendoscopy imaging based on GRIN lenses with one-and two-photon excitation modes." In: *Frontiers of Optoelectronics* 8.2 (2015), pp. 177–182.

[17] Ulf Leonhardt. "Optical conformal mapping." In: *Science* 312.5781 (2006), pp. 1777–1780.

[18] John B Pendry, David Schurig, and David R Smith. "Controlling electromagnetic fields." In: *Science* 312.5781 (2006), pp. 1780–1782.

[19] Jensen Li and John B Pendry. "Hiding under the carpet: a new strategy for cloaking." In: *Phys. Rev. Lett.* 101.20 (2008), p. 203901.

[20] Martin Schmiele, Vineeth S. Varma, Carsten Rockstuhl, and Falk Lederer. "Designing optical elements from isotropic materials by using transformation optics." In: *Phys. Rev. A* 81 (3 2010), p. 033837.

[21] Lin Xu and Huanyang Chen. "Conformal transformation optics." In: *Nat. Photonics* 9 (2015), pp. 15–23.

[22] Di Lin and James R Leger. "Numerical gradient-index design for coherent mode conversion." In: *Advanced Optical Technologies* 1.3 (2012), pp. 195–202.

[23] W Minster Kunkel and James R Leger. "Gradient-index design for mode conversion of diffracting beams." In: *Opt. Express* 24.12 (2016), pp. 13480–13488.

[24] F. Wyrowski and M. Kuhn. "Introduction to field tracing." In: *J. Mod. Opt.* 58.5-6 (2011), pp. 449–466.

[25] Z. Wang, S. Zhang, C. Hellmann, and F. Wyrowski. "Fast physical optics by semi-analytical and geometric Fourier transform." In: *EOS Topical Meeting on Diffractive Optics 2017* (2017).

[26] Physical optics simulation and design software "Wyrowski VirtualLab Fusion", developed by Wyrowski Photonics UG, distributed by LightTrans GmbH, Jena, Germany.

[27] Finite element method, JCM Wave, https://www.rp-photonics.com.

[28] M. G. Moharam and T. K. Gaylord. "Rigorous coupled-wave analysis of planar-grating diffraction." In: *J. Opt. Soc. Am. A* 71.7 (1981), pp. 811–818.

[29] Rui Shi, Christian Hellmann, and Frank Wyrowski. "Physical-optics propagation through curved surfaces." In: *J. Opt. Soc. Am. A* 36.7 (2019), pp. 1252–1260.

[30] Rui Shi and Frank Wyrowski. "Comparison of aplanatic and real lens focused spots in the framework of the local plane interface approximation." In: *J. Opt. Soc. Am. A* 36.10 (2019), pp. 1801–1809.

[31] Jari Turunen, Antto Vasara, Hiroyuki lchikawa, Eero Noponen, and Jan Westerholm. "Storage of multiple images in a thin synthetic Fourier hologram." In: *Opt. Commun.* 84.5-6 (1991), pp. 383–392.

[32] Zongzhao Wang, Site Zhang, Olga Baladron-Zorita, Christian Hellmann, and Frank Wyrowski. "Application of the semi-analytical Fourier transform to electromagnetic modeling." In: *Opt. Express* 27.11 (2019), pp. 15335–15350.

[33] Zongzhao Wang, Olga Baladron-Zorita, Christian Hellmann, and Frank Wyrowski. "Theory and algorithm of the homeomorphic Fourier transform for optical simulations." In: *Opt. Express* 28.7 (2020), pp. 10552–10571.

[34] Zongzhao Wang, Olga Baladron-Zorita, Christian Hellmann, and Frank Wyrowski. "Generalized Debye integral." In: *Opt. Express* 28.17 (2020), pp. 24459–24470.

[35] S. Zhang, C. Hellmann, and F. Wyrowski. "Algorithm for the propagation of electromagnetic fields through etalons and crystals." In: *Appl. Opt.* 56.15 (2017), pp. 4566–4576.

[36] Site Zhang. "Field tracing in optically anisotropic media." PhD thesis. Friedrich Schiller University Jena, 2018.

[37] R. Scarmozzino, A. Gopinath, R. Pregla, and S. Helfert. "Numerical techniques for modeling guided-wave photonic devices." In: *IEEE Journal of Selected Topics in Quantum Electronics* 6.1 (2000), pp. 150–162. ISSN: 1077-260X.

[38] K.-H. Brenner and W. Singer. "Light propagation through microlenses: a new simulation method." In: *Appl. Opt.* 32.26 (1993), pp. 4984–4988.

[39] M. D. Feit and J. A. Fleck. "Light propagation in graded-index optical fibers." In: *Appl. Opt.* 17.24 (1978), pp. 3990–3998.

[40] Matthias Fertig and K-H Brenner. "Vector wave propagation method." In: *J. Opt. Soc. Am. A* 27.4 (2010), pp. 709–717.

[41] G. Ronald Hadley. "Wide-angle beam propagation using Padé approximant operators." In: *Opt. Lett.* 17.20 (1992), pp. 1426–1428.

[42] Hatem El-Refaei, Ian Betty, and David Yevick. "The application of complex Padé approximants to reflection at optical waveguide facets." In: *IEEE Photon. Technol. Lett.* 12.2 (2000), pp. 158–160.

[43] W. P. Huang and C. L. Xu. "Simulation of three-dimensional optical waveguides by a full-vector beam propagation method." In: *IEEE J. Quantum Electron.* 29.10 (1993), pp. 2639–2649. ISSN: 0018-9197.

[44] Ansys Inc ANSYS. 2020.

[45] Bahaa EA Saleh, Malvin Carl Teich, and Bahaa E Saleh. *Fundamentals of photonics*. Vol. 22. Wiley New York, 1991.

[46] Graig F. Bohren and Donald R. Huffman. *Absorption and Scattering by an Arbitrary Particle*. John Wiley & Sons, Ltd, 2007. Chap. 3, pp. 57–81.

[47] K.F. Riley, M.P. Hobson, and S.J. Bence. *Mathematical Methods for Physics and Engineering: A Comprehensive Guide*. Cambridge University Press, 2006. ISBN: 9781139450997.

[48] Wen Yang. "Improved recursive algorithm for light scattering by a multilayered sphere." In: *Appl. Opt.* 42.9 (2003), pp. 1710–1720.

[49] O. Peña and U. Pal. "Scattering of electromagnetic radiation by a multilayered sphere." In: *Computer Physics Communications* 180.11 (2009), pp. 2348 –2354. ISSN: 0010-4655.

[50] Allan W Snyder and John Love. *Optical waveguide theory*. Springer Science & Business Media, 2012.

[51] Joseph W Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[52] S Schmidt, T Tiess, S Schröter, R Hambach, M Jäger, H Bartelt, A Tünnermann, and H Gross. "Wave-optical modeling beyond the thin-element-approximation." In: *Opt. Express* 24.26 (2016), pp. 30188–30200.

[53] Max Born and Emil Wolf. *Principles of Optics*. 7th. Cambridge University Press, 1999.

[54] Matthias Wilhelm Fertig. "Vector Wave Propagation Method." PhD thesis. Universität Mannheim, 2011.

[55] J.C. Butcher. *Numerical Methods for Ordinary Differential Equations*. Wiley, 2008.

[56] G Ronald Hadley. "Transparent boundary condition for the beam propagation method." In: *IEEE J. Quantum Electron.* 28.1 (1992), pp. 363–370.

[57] Weng Cho Chew and William H Weedon. "A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates." In: *Microwave and optical technology letters* 7.13 (1994), pp. 599–604.

[58] George A Baker and PR Graves-Morris. *Padé Approximants: Extensions and Applications*. Vol. 2. Addison-Wesley, 1981.

[59] W. Huang, C. Xu, S. . Chu, and S. K. Chaudhuri. "The finite-difference vector beam propagation method: analysis and assessment." In: *J. Light. Technol* 10.3 (1992), pp. 295–305. ISSN: 0733-8724.

[60] Hongling Rao, Rob Scarmozzino, and Richard M Osgood. "A bidirectional beam propagation method for multiple dielectric interfaces." In: *IEEE Photon. Technol. Lett.* 11.7 (1999), pp. 830–832.

[61] Hongling Rao, Michael J Steel, Rob Scarmozzino, and Richard M Osgood. "Complex propagators for evanescent waves in bidirectional beam propagation method." In: *J. Light. Technol* 18.8 (2000), p. 1155.

[62] Jean Paul Hugonin and Philippe Lalanne. "Perfectly matched layers as nonlinear coordinate transforms: a generalized formalization." In: *J. Opt. Soc. Am. A* 22.9 (2005), pp. 1844–1849.

[63] Evgeny Popov and Michel Nevière. "Maxwell equations in Fourier space: fast-converging formulation for diffraction by arbitrary shaped, periodic, anisotropic media." In: *J. Opt. Soc. Am. A* 18.11 (2001), pp. 2886–2894.

[64] Shinan-Chur Sheng and AE Siegman. "Nonlinear-optical calculations using fast-transform methods: Second-harmonic generation with depletion and diffraction." In: *Phys. Rev. A* 21.2 (1980), p. 599.

[65] MG Moharam, Eric B Grann, Drew A Pommet, and TK Gaylord. "Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings." In: *J. Opt. Soc. Am. A* 12.5 (1995), pp. 1068–1076.

[66] Evgeni Popov and Michel Nevière. "Grating theory: new equations in Fourier space leading to fast converging results for TM polarization." In: *J. Opt. Soc. Am. A* 17.10 (2000). FMM, pp. 1773–1784.

[67] R.K. Luneburg. *Mathematical Theory of Optics, by R.K. Luneburg. Supplementary Notes by M. Herzberger and Others*. 1944.

[68] Electromagnetic Field Interaction with Nanocylinders, www.LightTrans.com.

[69] C.E. Shannon. "Communication in the Presence of Noise." In: *Proceedings of the IRE* 37.1 (1949), pp. 10–21.

[70] Thomas Schanze. "Sinc interpolation of discrete periodic signals." In: *IEEE Transactions on Signal Processing* 43.6 (1995), pp. 1502–1503.

[71] Leonard Mandel and Emil Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995.

[72] Jochen Brüning and Victor W Guillemin. *Mathematics Past and Present Fourier Integral Operators*. Springer-Verlag Berlin Heidelberg, 1994.

[73] Gerald Farin. *Curves and Surfaces for CAGD: A Practical Guide*. Morgan Kaufmann, 2001.

[74] Anurag Sharma, D. Vizia Kumar, and A. K. Ghatak. "Tracing rays through graded-index media: a new method." In: *Appl. Opt.* 21.6 (1982), pp. 984–987.

[75] Haitao Dai Dan Luo Baoli Yao Xiurun Mao Yang Yang and Shaohui Yan. "Tunable photonic nanojet formed by generalized Luneburg lens." In: *Opt. Express* 23.20 (2015), pp. 26426–26433.

[76] W. Koechner. "Thermal Lensing in a Nd:YAG Laser Rod." In: *Appl. Opt.* 9.11 (1970), pp. 2548–2553.

[77] Edmund Plano-Convex Lenses Edmund lens. 2020.

[78] Data Sheets Optical Glasses from SCHOTT N-LASF9. 2014.

[79] Tobias Bonhoff, Lasse Büsing, Jochen Stollenwerk, and Peter Loosen. "Modeling of optical aberrations due to thermal deformation using finite element analysis and ray-tracing." In: *Optical Systems Design 2015: Optical Design and Engineering VI*. Ed. by Laurent Mazuray, Rolf Wartmann, and Andrew P. Wood. Vol. 9626. International Society for Optics and Photonics. SPIE, 2015, pp. 462 –471.

[80] Tobias Bonhoff, Jochen Stollenwerk, and Peter Loosen. "Multi-physical modeling of thermal and dispersive effects in optical systems for high power ultra-short laser pulses." In: *High-Power Laser Materials Processing: Applications, Diagnostics, and Systems VII*. Vol. 10525. International Society for Optics and Photonics. SPIE, 2018.

[81] SCHOTT Technical Information TIE-19 "Temperature coefficient of refractive index" schott link.

[82] Maarten Sluijter, Dick K. G. de Boer, and Joseph J. M. Braat. "General polarized ray-tracing method for inhomogeneous uniaxially anisotropic media." In: *J. Opt. Soc. Am. A* 25.6 (2008), pp. 1260–1273.

[83] Maarten Sluijter, Dick K. de Boer, and H. Paul Urbach. "Ray-optics analysis of inhomogeneous biaxially anisotropic media." In: *J. Opt. Soc. Am. A* 26.2 (2009), pp. 317–329.

[84] Maarten Sluijter, Angie Herzog, Dick K. G. de Boer, Marcel P. C. M. Krijn, and H. Paul Urbach. "Ray-tracing simulations of liquid-crystal gradient-index lenses for three-dimensional displays." In: *J. Opt. Soc. Am. B* 26.11 (2009), pp. 2035–2043.

[85] Duncan T Moore. "Gradient-index optics: a review." In: *Appl. Opt.* 19.7 (1980), pp. 1035–1038.

[86] Jeremy Teichman, Jenny Holzer, Bohdan Balko, Brent Fisher, and Leonard Buckley. *Gradient index optics at DARPA*. Tech. rep. INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA VA, 2013.

[87] Kathleen A Richardson, Myungkoo Kang, Laura Sisken, Anupama Yadav, Spencer Novak, Antoine Lepicard, Isabelle Martin, Hugues Francois-Saint-Cyr, Casey M Schwarz, Theresa S Mayer, et al. "Advances in infrared gradient refractive index (GRIN) materials: a review." In: *Opt. Eng.* 59.11 (2020), p. 112602.

[88] JB MacChesney, PB O'connor, and HM Presby. "A new technique for the preparation of low-loss and graded-index optical fibers." In: *Proceedings of the IEEE* 62.9 (1974), pp. 1280–1281.

[89] Suzanne R Nagel, John B MacChesney, and Kenneth L Walker. "An overview of the modified chemical vapor deposition (MCVD) process and performance." In: *IEEE Transactions on Microwave Theory and Techniques* 30.4 (1982), pp. 305–322.

[90] Michael A Pickering, Raymond L Taylor, and Duncan T Moore. "Gradient infrared optical material prepared by a chemical vapor deposition process." In: *Appl. Opt.* 25.19 (1986), pp. 3364–3372.

[91] David P Hamblen. *Gradient refractive index optical lenses*. US Patent 3,486,808. 1969.

[92] Sigeaki Ohmi, Hiroyuki Sakai, Yoshiyuki Asahara, Shin Nakayama, Yoshitaka Yoneda, and Tetsuro Izumitani. "Gradient-index rod lens made by a double ion-exchange process." In: *Appl. Opt.* 27.3 (1988), pp. 496–499.

[93] Andrey A Lipovskii, Peter C Schultz, Dmitry K Tagantsev, and Boris V Tatarintsev. "Phosphate glasses for GRIN structures by ion exchange." In: *2005 International Symposium on Glass in Conjunction with the International Commission on Glass, Shanghai, China*. 2005, pp. 10–14.

[94] Gregor Matz, Bernhard Messerschmidt, Ulf Possner, Jan Kobelke, Marcel Kunze, and Herbert Gross. "Reliable profile reconstruction of GRIN lenses produced by ion-exchange processes." In: *Opt. Lett.* 43.3 (2018), pp. 407–410.

[95] Lisa C Klein. *Sol-gel optics: processing and applications*. Vol. 259. Springer Science & Business Media, 2013.

[96] S Konishi, K Shingyouchi, and A Makishima. "r-GRIN glass rods prepared by a sol-gel method." In: *J. Non-Cryst. Solids* 100.1-3 (1988), pp. 511–513.

[97] Maria Laczka, J Najman, D Wegrzynek, and Marek Wychowaniec. "GRIN glasses prepared by sol-gel method." In: *Gradient-Index Optics in Science and Engineering*. Vol. 2943. International Society for Optics and Photonics. 1996, pp. 95–104.

[98] Hong-Ren Wang, Michael J Cima, Brian D Kernan, and Emanuel M Sachs. "Alumina-doped silica gradient-index (GRIN) lenses by slurry-based three-dimensional printing (S-3DP™)." In: *J. Non-Cryst. Solids* 349 (2004), pp. 360–367.

[99] K Sasan, T Yee, Nikola Dudukovic, Michael Johnson, R Dylla-Spears, J Yoo, O Herrera, D Nguyen, A Lange, T Fears, et al. *Developing 3D-printed optical glasses from sol-gel feedstocks*. Tech. rep. Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), 2019.

[100] Albertas Žukauskas, Ieva Matulaitienė, Domas Paipulas, Gediminas Niaura, Mangirdas Malinauskas, and Roaldas Gadonas. "Tuning the refractive index in 3D direct laser writing lithography: towards GRIN microoptics." In: *Laser & Photonics Reviews* 9.6 (2015), pp. 706–712.

[101] JS Shirk, M Sandrock, D Scribner, E Fleet, R Stroman, E Baer, and A Hiltner. *Biomimetic gradient index (GRIN) lenses*. Tech. rep. Naval Research Lab Washington DC, 2006.

[102] Shanzuo Ji, Kezhen Yin, Matthew Mackey, Aaron Brister, Michael Ponting, and Eric Baer. "Polymeric nanolayered gradient refractive index lenses: technology review and introduction of spherical gradient refractive index ball lenses." In: *Opt. Eng.* 52.11 (2013), p. 112105.

[103] Nathan Kundtz and David R Smith. "Extreme-angle broadband metamaterial lens." In: *Nat. Mater.* 9.2 (2010), pp. 129–132.

[104] Pekka Alitalo and Sergei Tretyakov. "Electromagnetic cloaking with metamaterials." In: *Materials today* 12.3 (2009), pp. 22–29.

[105] P Sinai. "Correction of optical aberrations by neutron irradiation." In: *Appl. Opt.* 10.1 (1971), pp. 99–104.

[106] RK Mohr, JA Wilder, PB Macedo, and PK Gupta. "Digest of Topical Meeting on Gradient-Index Optical Imaging Systems." In: *Optical Society of America, Washington DC, paper WA1* (1979).

[107] HJ Trost, S Ayers, T Chen, WR Cox, ME Grove, and R Hoenigman. "Using drop-on-demand technology for manufacturing grin lenses." In: *Proc. 2001 Ann. Mtg. ASPE* (2001), pp. 533–536.

[108] Elias Snitzer. "Cylindrical dielectric waveguide modes." In: *J. Opt. Soc. Am.* 51.5 (1961), pp. 491–498.

[109] U. Grenander and G. Szegö. *Toeplitz forms and their applications*. Chelsea Pub. Co., 1984. ISBN: 9780828403214.

[110] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics. Springer New York, 2013.

[111] Jia wen Deng, Uwe Guenther, and Qing hai Wang. "General PT-Symmetric Matrices." In: (2012).

[112] F. Zhang and Q. Zhang. "Eigenvalue Inequalities for Matrix Product." In: *IEEE Transactions on Automatic Control* 51.9 (2006), pp. 1506–1509. ISSN: 2334-3303.

[113] Kumar Ravinder Merikoski Jorma K. "Inequalities for spreads of matrix sums and products." eng. In: *Applied Mathematics E-Notes* 4 (2004), pp. 150–159.

[114] C. Bao C. Gomez-Reino M.V. Perez. *Gradient-Index Optics Fundamentals and Applications*. Springer-Verlag Berlin Heidelberg, 2002.

LIST OF FIGURES

LIST OF TABLES

LISTINGS

ACKNOWLEDGEMENTS

I realize the six years as a Ph.D. student at the Faculty of Physics and Astronomy of Friedrich Schiller University Jena passed very quickly. It was not always smooth, but I had a fruitful, busy, and joyful journey with all of your's help.

First of all, I am deeply indebted to my supervisor, Prof. Dr. Frank Wyrowski. Frank passed me his wisdom and passion for research, creative thinkings for physics. Meanwhile, he found and affirmed my abilities, and trustfully gave me lots of opportunities and freedom during my work and study. Without his dedicated guidance and encouragement, it would be impossible for the existence of this thesis. Frank also offered a surprisingly good opportunity to communicate with worldwide scientists and engineers, which widen my horizons and make my six-year life colorful.

I would also like to extend my deepest gratitude to my colleague, mentor, friend, Dr. Site Zhang. In the passing years, Site offered me invaluable help with details learning of scientific writing, programming, and communicating with others. He opened the door of the RK k-domain method, one of my thesis's essential parts. I also appreciate his contribution to team-building, including different templates and workflows, where I get many benefits.

It was always a pleasure to come to work every day with such lovely, passionate, and talented colleagues. I would like to say thank you to all of you. Olga Baladron-Zorita is very professional and careful in physical thinking and expression. She helped me dramatically to improve the quality of the publications and scientific reports. Her friendly and patience are unparalleled. My office mate Zongzhao Wang uses his profound scientific skills and humor to make the serious discussion relax and productive. Special thanks to Shih-Te Hung to offer me the code of Mie theory for a multilayered sphere, Rui Shi to offer me the rigorous field solver FMM+PMLs, and Huaiyu Chen for teaching me the software Ansys. Additionally, I also like to acknowledge the young students, Xiaoyan Yu, Wenxiu Wang, Zhouping Lyu, Shiyao Wang, and Yichen Liu, from whom I obtained valuable knowledge.

Furthermore, I am also grateful to Mr. Christian Hellmann and all the other colleagues in LightTrans International GmbH and Wyrowski Photonics UG, who have been continuously supporting my research by providing useful codes and insights into the software VirtualLab Fusion.

Last but not least, I would like to thank my parents and my sister for their endless support. They make me feel being loved and needed. My sincere thanks to Renwen, who has accompanied me throughout these years at each moment, no matter happy or rough.

EHRENWÖRTLICHE ERKLÄRUNG

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Die aus anderen Quellen direkt order indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen:

1. Herr Prof. Dr. Frank Wyrowski als Betreuer dieser Arbeit.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen.

Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten ergalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die geltende Promotionsordnung der Physikalisch-Astronomishcen Fakultät ist mir bekannt.

Ich versichere ehrenwörtlich, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Ort, Datum

Unterschrift