



POLITECNICO DI TORINO  
TECHNISCHE UNIVERSITÄT MÜNCHEN

---

DIPARTIMENTO DI ELETTRONICA E TELECOMUNICAZIONI  
MASTER'S DEGREE IN COMMUNICATIONS AND COMPUTER  
NETWORKS ENGINEERING

**DESIGN OF A NUMERICAL SIMULATOR FOR  
OPTICAL SIGNAL PROPAGATION IN A  
MULTIMODE FIBER STRUCTURE**

Supervisor of Politecnico di Torino:  
Prof. ANDREA CARENA

Student:  
PAOLO CARNIELLO  
S277593

Co-Supervisor of Politecnico di Torino:  
Prof. ROBERTO GAUDINO

Supervisor of TUM:  
Prof. Dr.-Ing. NORBERT HANIK

---

ACADEMIC YEAR 2020-2021

## Abstract

The theory behind the implementation of a numerical solver for the generalized multimode nonlinear Schrödinger equation (GMMNLSE), which models the propagation of a modulated electromagnetic field in a fiber, is presented. Both linear and nonlinear coupling effects have been considered, with a focus on the former for which a solver has been implemented. Particular attention is devoted to birefringence and mode coupling in the two polarization mode case, within the framework of polarization mode dispersion (PMD). For the multimode case, spatial mode coupling is generated through both a statistical approach and a physical one modeling some realistic impairments, like core ellipticity, bends and axis rotation.

Concerning the nonlinear effects, the derivation of the GMMNLSE, accounting for linear coupling, Kerr and Raman nonlinear effects, is reviewed. A possible implementation design for a numerical mode solver including the nonlinear phenomena is proposed, based on a Split-Step Fourier method (SSFM) combined with a Runge-Kutta solver for ordinary differential equations of order four (RK4).

Simulations have been carried out throughout the whole study to qualitatively support the theoretical knowledge, drive the intuition and provide feedback about the impact of the various distorting effects on the transmitted pulse shapes.

The present work can be helpful for the development of space division multiplexing (SDM) coherent systems where different independent signals excite different modes of a multimode structure, with the aim of increasing the data rates of the current optical networks.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Thesis Structure . . . . .	2
<b>2</b>	<b>Basics</b>	<b>4</b>
2.1	Examples of MMFs and Modes . . . . .	10
<b>3</b>	<b>Polarization Mode Coupling</b>	<b>15</b>
3.1	Origins of Birefringence . . . . .	16
3.2	Segment of Uniform Birefringence . . . . .	18
3.2.1	Modelling Birefringence and Intrinsic DGD . . . . .	18
3.2.2	Channel Transfer Matrix for Uniform Birefringence . . . . .	21
3.2.3	Propagation in a Segment with Uniform Birefringence . . . . .	23
3.2.4	Frequency Independent Birefringence . . . . .	25
3.2.5	Frequency Dependent Birefringence . . . . .	26
3.3	Concatenation of Segments: Polarization Mode Coupling . . . . .	28
3.3.1	Frequency Independent Birefringence for a Direct-Detection System . . . . .	33
3.3.2	Frequency Independent Birefringence for a Coherent-System . . . . .	34
3.3.3	Principal States of Polarization . . . . .	37
3.3.4	Nonuniform Frequency Dependent Birefringence . . . . .	42
3.3.5	PMD Vector . . . . .	46
3.3.6	Statistics of PMD . . . . .	49
3.4	Harmonization . . . . .	51
<b>4</b>	<b>Spatial Mode Coupling</b>	<b>57</b>
4.1	Importance of Phase-Matching . . . . .	62
4.2	Mode Coupling through Physical Effects . . . . .	64
4.2.1	Axis Rotation . . . . .	65
4.2.2	Bend . . . . .	66
4.2.3	Core ellipticity . . . . .	68
4.2.4	Simulation Model of a Fiber with Physical Coupling Phenomena . . . . .	69

## CONTENTS

4.3	Mode Coupling through a Statistical Approach . . . . .	70
4.4	Primary Modes and Simulations . . . . .	72
<b>5</b>	<b>Generalized Multimode Nonlinear Schrödinger Equation</b>	<b>76</b>
5.1	Derivation . . . . .	77
5.2	Manipulation of the GMMNLSE . . . . .	93
5.3	Brief Description of the GMMNLSE Terms . . . . .	97
5.4	Modelling the Raman Impulse Response . . . . .	98
5.5	Inclusion of Loss . . . . .	100
<b>6</b>	<b>Simulator Design</b>	<b>103</b>
6.1	Split-Step Fourier Method . . . . .	104
6.2	Runge-Kutta Method . . . . .	105
6.3	Hints on Manakov Equations . . . . .	108
<b>7</b>	<b>Conclusions and Outlook</b>	<b>109</b>
	<b>Bibliography</b>	<b>111</b>

# Acknowledgments

It would be easy, at this point, to say that these years of university have all been easy and fun, but it would not be right. Several are the chances I missed and countless the mistakes I made. However, the adventure I began unknowingly five years ago led me to places and personal achievements I would have never thought of beforehand, an adventure which I would enthusiastically start over again if I went back in time. Above all, an adventure which gave me the opportunity to meet people with whom I shared time, ideas and emotions, people who contributed to shape the person I am now.

It would be impossible and restrictive to name all those who made this journey so intense, who inspired me, who supported me, who believed in me. My heartfelt thanks go to them.

Let me express special thanks to Tasnad Kernetzky for helping me throughout the elaboration of this thesis. On top of that, the kicker and bouldering you introduced me to made my working days in Munich more fun.

Last but not least, I would like to thank my family, whose presence besides me was never in question.

München, December 2021

Paolo Carniello

Studia, che te ga 'na campagna in  
testa.

---

Nonno Giovanni



# Notation

Symbol	Meaning
$:=$	defined as
l.h.s.	left-hand side
r.h.s.	right-hand side
i.i.d.	independent and identically distributed
$f$ or $\omega := 2\pi f$	frequency
$x, y$	transversal coordinates of the fiber
$z$	longitudinal coordinate along fiber axis
$\Phi' \equiv \frac{\partial \Phi}{\partial \omega}$	differentiation for scalar functions $\Phi(\omega)$
$\mathbf{T}_\omega := \frac{\partial \mathbf{T}}{\partial \omega}$	differentiation for a matrix $\mathbf{T}$
$\mathbf{T}'$	generic matrix different from $\mathbf{T}$ . Note that $(\mathbf{T}' \neq \mathbf{T}_\omega)$ .
$\mathbf{E}$	generic vector
$\hat{\mathbf{E}}$	generic unit vector
$ E\rangle$	Jones vector
$\mathbf{E}$	Stokes vector
$\tilde{a}(\omega)$ or $\mathfrak{F}\left[a(t)\right]$	Fourier transform of $a(t)$
$\mathfrak{F}^{-1}\left[\tilde{a}(\omega)\right]$	inverse Fourier transform of $\tilde{a}(\omega)$
$\mathbf{I}$	identity matrix of suitable dimensions

The Fourier transform pair is

$$\tilde{E}(f) = \int_{-\infty}^{+\infty} E(t)e^{-j2\pi ft} dt$$

$$E(t) = \int_{-\infty}^{+\infty} \tilde{E}(f)e^{+j2\pi tf} df$$

Several quantities depend on both time  $t$  (or frequency  $\omega$ ) and position  $z$ , but sometimes one or both dependencies have been suppressed for conciseness.

# Chapter 1

## Introduction

Optical networks are the backbone of today's telecommunication infrastructure, consume the least energy per data unit over the other wireless and wired competitors [MS20] and, above all, are the only medium capable of sustaining the current demand for extremely high bit rates. Yet, quite some optical network capacity crunches have been predicted in recent years due to the steadily increasing world network traffic [Chr09; Wal18]. The technique called Space-Division Multiplexing (SDM) is considered to be one of the most promising solutions to keep the pace in the future [ASM16; WN17; Wal18]

SDM consists in transmitting  $M$  parallel independent data streams over the same fiber structure, while today's optical fibers support at most two streams. SDM can rely on different types of fibers, but the two main groups are the multimode fibers (MMFs) and the multicore fibers (MCFs) [ASM16; Wri+17] [Agr19, Ch.14]. The former are not an innovative structure *per se*, but independently modulating a set of  $M$  guided modes of a MMF is a recent use case. The latter are of more recent invention and consist in a dielectric cladding where multiple cores are present, each one usually supporting two modes. The vicinity of the cores is usually such that the modal profiles in the various cores interact with each other, giving rise to new modal profiles called supermodes [Agr19, Ch.14.1.3].

In both structures, the modes are not independent, but are coupled due to linear and nonlinear perturbing effects. The former arise both from geometrical imperfections and stresses, the latter from the nonlinear properties of the medium. The description of these phenomena is still a broad and active area of research.

In order to do research in the field of SDM, the design of a numerical simulator is an essential first step to qualitatively support the theoretical knowledge, drive the intuition and provide feedback about the impact of the various impairments affecting the transmitted signals. Furthermore, the availability of a numerical solver for the exact propagation equations helps the derivation of simplified propagation models and the development of distortion compensation strategies.

Hence, the thesis aims at presenting the theory behind the implementation of a numerical solver for the generalized multimode nonlinear Schrödinger equation (GMMNLSE), which models the propagation of a modulated electromagnetic field in a fiber [KM04; PH08; ASM16; Wri+17]. Both linear and nonlinear coupling effects have been considered, with a focus on the former. A common approach in literature is to either neglect or consider in a purely statistical way the presence of linear coupling in the GMMNLSE [PH08; ASM16; Wri+17]. On the opposite, both a statistical approach and a physical one modeling some realistic impairments are proposed for the linear mode coupling in the GMMNLSE in this thesis. Since linear and nonlinear phenomena are normally treated separately in literature, an effort has been made to keep the notation consistent throughout the present work and to link the different normalizations and conventions, when relevant. A solver for the linear regime has been implemented and a possible design to include nonlinear effects is presented.

A thorough discussion about all linear and nonlinear phenomena is far beyond the scopes of the thesis. However, a desired property of this work is to supply the theory necessary to understand the presented models and the main aspects related to the numerical solver design, from polarization-mode dispersion (PMD) to the GMMNLSE. Toward this target, the derivations of the relevant equations have been reported and graphical examples, produced through the developed simulator, are offered throughout the digression to support the theoretical knowledge. Ultimately, this work can also serve as a foundation for the study and implementation of more advanced models, like the Manakov equations, whose computational burden is much lower than the GMMNLSE presented here [Ryf+12; ASM16].

## 1.1 Thesis Structure

The thesis is organized as follows.

Ch.2 recalls some basic concepts of optical transmission about modes and fiber geometries, and derives the Helmholtz wave equation exploited in the next chapters.

Ch.3 focuses on the phenomenon of polarization-mode dispersion (PMD) for a fiber supporting two polarization modes. The concepts are introduced in a hierarchical way, starting from the origins of birefringence, then considering the four combinations of frequency independent and frequency dependent, uniform and nonuniform birefringence. The model of the Principal States of Polarization (PSPs) is presented and a few concepts of statistical analysis within the framework of PMD are reviewed. The formalism of both Jones and Stokes spaces is considered to strengthen the understanding.

Ch.4 is about mode coupling among modes with different spatial profiles, based on the classical coupled-mode theory. A physical approach to model linear

perturbations, like axis rotations, bends and core ellipticity is compared against a statistical one.

Ch.5 repeats in a detailed manner the derivation of the GMMNLSE starting from Maxwell's equations. The ideas behind the choice of different normalizations and some simplifications are clarified.

Ch.6 sketches the GMNLS numerical solver design, based on a split-step Fourier method combined with a Runge-Kutta method of fourth order. Few basic ideas on the less complex Manakov equations are conveyed.

Ch.7 concludes the thesis and provides an outlook on possible directions for improvement and research.

# Chapter 2

## Basics

In this chapter we recap some basic knowledge about optical fiber transmission necessary for the next chapters. Before starting, we mention the notation followed in the thesis.

In fiber communications a sequence of waveforms which modulate an optical electromagnetic carrier at frequency  $\omega$  are transmitted over the channel. The field propagation in the optical fiber is modeled through Maxwell's equations, with some boundary conditions based on the system geometry. A relevant concept for fiber transmission is a mode, which, restricting the view to the electric field, is defined as

$$\mathbf{E}_k(\mathbf{r}, t) := \mathbf{F}_k(x, y, \omega) e^{-j\beta_k(\omega)z} e^{+j\omega t} \quad (2.1)$$

where  $\mathbf{F}_k(x, y, \omega)$  is the called modal profile or transversal field pattern of the  $k$ -th mode,  $\beta_k(\omega)$  is the propagation constant of the mode and  $\omega$  is the frequency (or wavelength) at which the mode is computed. By definition, the modal profiles are  $z$ -independent, which means that they maintain their shapes  $\mathbf{F}_k$  along the fiber length, only the phases change with  $z$ . The index  $k$  refers to the  $k$ -th mode of the modal set, which in general is composed of  $M = 2N$  elements. The reason for the  $2N$  is going to be clear later on.

Both the modal profiles and the propagation constants change with frequency [Mar74] [Kei11, Ch.2], even though  $\mathbf{F}_k(x, y, \omega)$  is often approximated as frequency independent for  $\omega \in [-\Delta\omega + \omega_0, \omega_0 + \Delta\omega]$ , with  $\Delta\omega \ll \omega_0$  (narrowband approximation). The dependence of  $\beta_k(\omega)$  on frequency is never negligible and gives rise to various dispersion phenomena. The propagation constant can be expanded in a Taylor series as

$$\beta_k(\omega_0 + \Delta\omega) = \beta_k^{(0)} + \beta_k^{(1)}\Delta\omega + \frac{1}{2}\beta_k^{(2)}(\Delta\omega)^2 + \dots \quad (2.2)$$

where  $\beta_k^{(0)} := \beta_k(\omega_0)$  and  $\beta_k^{(a)} := \left( \frac{d^a \beta_k}{d\omega^a} \right)_{\omega_0}$  are the so-called *dispersion coefficients*.

We do not review the mode-independent dispersion phenomena, but just recall

that  $\beta_k^{(1)} = 1/v_{g,m}$  is the inverse group-velocity of the  $m$ -th mode,  $\tau_k = \beta_k^{(1)}z$  is the propagation delay, and  $\beta_k^{(2)}$  is the group-velocity dispersion (GVD).

It might feel confusing at first mixing frequency and time variables in an expression like (2.1). However, the frequency has to be considered only a parameter in (2.1), since the field is a continuous wave (CW) field at a fixed frequency  $\omega$ .

Up to now, the mode is just a definition. However, it can be proven that (2.1) solves Maxwell's equations and also that the modal set is a complete set of orthogonal modes, that is, all the solutions of Maxwell's equations can be expressed as a superposition of orthogonal modes with proper coefficients  $A_k$  [Mar74, p. 83], like

$$\mathbf{E}(\mathbf{r}, t) = \sum_{k=1}^M A_k \mathbf{E}_k(\mathbf{r}, t) = \sum_{k=1}^M A_k \mathbf{F}_k(x, y, \omega) e^{-j\beta_k(\omega)z} e^{+j\omega t} \quad (2.3)$$

Hence, the modal set forms a basis, the modal basis.

We consider  $\mathbf{F}_k$  to be adimensional and  $A_k$  in units of electric field, i.e., V/m. Given that the fiber is ideal and the field is CW, the modal amplitudes  $A_k$  are  $z$ - and  $t$ -independent.

We now perform some steps towards the computation of the ideal modes in a generic fiber structure. We do not derive explicit expressions for the modal profiles (since we would not really need them), but we stop to an equation in  $\mathbf{F}_k$  from which the modal profiles could be computed, and show some important relations which are going to be useful later on.

Let us start by remembering Maxwell's equations in time domain [Agr19, p. 27][Coe10, p. 16][Mid03]:

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (2.4a)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = +\frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} + \mathbf{J}(\mathbf{r}, t) \quad (2.4b)$$

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = \varrho_c; \quad (2.4c)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = \mathbf{0}; \quad (2.4d)$$

where  $\mathbf{E}$  is the electric field vector,  $\mathbf{H}$  is the magnetic field vector,  $\mathbf{D}$  is the electric displacement or electric induction vector,  $\mathbf{B}$  is the magnetic induction or magnetic flux density vector,  $\mathbf{J}$  is the current density,  $\varrho_c$  is the volumetric charge density.

$\nabla$  is the nabla operator, through which are expressed in a symbolic form the curl as  $\nabla \times$ , the divergence as  $\nabla \cdot$ , the gradient as  $\nabla$  and the Laplacian as  $\nabla^2$ .

Given that in optical fibers there are no free charges [Agr19]

$$\mathbf{J} = \mathbf{0} \quad (2.5a)$$

$$\varrho_c = 0 \quad (2.5b)$$

A modes is, by definition (2.1), a complex harmonic at a specific frequency  $\omega$ , concerning the time dependence. Hence, all the time-dependent quantities appearing in Maxwell's equations are complex harmonics at the frequency  $\omega$ . If we indicate with “ $\check{\mathbf{E}}(\mathbf{r}, t)$ ” the  $\mathbf{r}$  dependent amplitude of a complex harmonic, which we call phasor, so that

$$\mathbf{E}(\mathbf{r}, t) = \check{\mathbf{E}}e^{j\omega t} \quad (2.6)$$

then Maxwell's equations become

$$\nabla \times \check{\mathbf{E}}(\mathbf{r}) = -j\omega \check{\mathbf{B}}(\mathbf{r}) \quad (2.7a)$$

$$\nabla \times \check{\mathbf{H}}(\mathbf{r}) = +j\omega \check{\mathbf{D}}(\mathbf{r}) \quad (2.7b)$$

$$\nabla \cdot \check{\mathbf{D}}(\mathbf{r}) = 0; \quad (2.7c)$$

$$\nabla \cdot \check{\mathbf{B}}(\mathbf{r}) = 0; \quad (2.7d)$$

The constitutive relations of the material, written for phasors, are

$$\check{\mathbf{D}}(\mathbf{r}) = \varepsilon_0 \check{\mathbf{E}}(\mathbf{r}) + \check{\mathbf{P}} \quad (2.8a)$$

$$\check{\mathbf{B}}(\mathbf{r}) = \mu_0 \check{\mathbf{H}}(\mathbf{r}) + \mu_0 \check{\mathbf{M}} \quad (2.8b)$$

where  $\check{\mathbf{P}}$  is the induced electric material polarization vector,  $\check{\mathbf{M}}$  is the induced magnetic polarization,  $\varepsilon_0$  is the vacuum dielectric permittivity,  $\mu_0$  is the vacuum magnetic permeability. For a nonmagnetic material, like those of which fibers are made of, [Agr19, p. 27][Coe10, p. 16]

$$\check{\mathbf{M}} = 0 \quad (2.9)$$

The electric material polarization vector  $\mathbf{P}(\mathbf{r}, t)$  can be expanded as [Coe10, p. 2.27]:

$$\begin{aligned} \mathbf{P}(\mathbf{r}, t) := & \varepsilon_0 \int_{-\infty}^{+\infty} \chi^{(1)}(t - t_1) \cdot \mathbf{E}(t_1) dt_1 \\ & + \varepsilon_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi^{(2)}(t_1 - t, t_2 - t) : \mathbf{E}(t_1) \mathbf{E}(t_2) dt_1 dt_2 \\ & + \varepsilon_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi^{(3)}(t_1 - t, t_2 - t, t_3 - t) \vdots \mathbf{E}(t_1) \mathbf{E}(t_2) \mathbf{E}(t_3) dt_1 dt_2 dt_3 \\ & + \dots \end{aligned} \quad (2.10)$$

It can also be written as

$$\mathbf{P}(\mathbf{r}, t) := \mathbf{P}_L(\mathbf{r}, t) + \mathbf{P}_{NL}(\mathbf{r}, t) \quad (2.11)$$

where [Agr19, p. 28]

$$\mathbf{P}_L(\mathbf{r}, t) := \varepsilon_0 \int_{-\infty}^{+\infty} \chi^{(1)}(\mathbf{r}, t - t') \cdot \mathbf{E}(\mathbf{r}, t') dt' \quad (2.12)$$

is the linear part of the polarization vector and  $\chi^{(1)}$  is the first-order nonlinear susceptibility tensor, related to the refractive index of the medium, which characterizes the fiber material and geometry.

For the computation of the modes, we consider the fiber to be ideal, which means that the nonlinear effects are not considered, i.e.,  $\mathbf{P}_{NL}(\mathbf{r}, t) = 0$ . Moreover, it also means that the fiber material and geometry are not even perturbed linearly, which is going to be clarified and exploited soon. In this thesis, with the expression ‘‘perturbation’’ we mean an external or internal stress, or a geometrical or material imperfection of the fiber introducing linear coupling [Mar74; Pal13], as detailed in later chapters. The fact that the modes are computed for an ideal structure does not mean that our whole study is restricted to the ideal case; this procedure is just needed for the computation of the modes.

Since, again, we are interested in the case of complex harmonics, the phasor of the material polarization vector becomes

$$\check{\mathbf{P}}(\mathbf{r}) = \check{\mathbf{P}}_L(\mathbf{r}) = \varepsilon_0 \check{\chi}^{(1)}(\mathbf{r}, \omega) \cdot \check{\mathbf{E}}(\mathbf{r}) \quad (2.13)$$

Let us insert (2.13) in (2.8a),

$$\begin{aligned} \check{\mathbf{D}} &= \varepsilon_0 \check{\mathbf{E}} + \check{\mathbf{P}}_L \\ &= \varepsilon_0 \check{\mathbf{E}} + \varepsilon_0 \check{\chi}^{(1)}(\mathbf{r}, \omega) \cdot \check{\mathbf{E}} \\ &= \varepsilon_0 (\mathbf{I} + \check{\chi}^{(1)}(\mathbf{r}, \omega)) \check{\mathbf{E}} \end{aligned} \quad (2.14)$$

We define [Agr19, p. 2.1.13] [Han21, p. 3.18]

$$\varepsilon_r(\mathbf{r}, \omega) + \delta\varepsilon_r(\mathbf{r}, \omega) := \mathbf{I} + \check{\chi}^{(1)}(\omega) \quad (2.15)$$

where  $\varepsilon_r$  is the material relative dielectric permittivity tensor and  $\delta\varepsilon_r$  is its perturbation induced by the linear perturbations. However, as said, we exclude them for the computation of the ideal fiber, so that

$$\varepsilon_r(\mathbf{r}, \omega) := \mathbf{I} + \check{\chi}^{(1)}(\omega) \quad (2.16)$$

We also define [Agr19, p. 2.1.14] [Coe10]

$$\varepsilon_r(\mathbf{r}, \omega) := \left( n(\mathbf{r}, \omega) \mathbf{I} + j \frac{\alpha c}{2\omega} \right)^2 \quad (2.17)$$

where  $n$  is the material refractive index and  $\alpha$  is the fiber loss. Notice this is also the definition of refractive index. Again, we do not consider fiber loss in an ideal fiber. Hence, (2.17) reduces to

$$\varepsilon_r(\mathbf{r}, \omega) \approx n(\mathbf{r}, \omega)^2 \mathbf{I} \quad (2.18)$$

(2.17) and (2.18) could have been written because an ideal fiber is isotropic, which means that its mechanical and optical properties do not depend on the direction (polarization, formally) of the applied field vector. Hence  $\varepsilon_r$  is diagonal, with diagonal terms approximately equal to each other [Coe10, p. 18].

Inserting firstly (2.16) and then (2.18) in (2.14), we get

$$\check{\mathbf{D}}(\mathbf{r}) = \varepsilon_0 n(\mathbf{r}, \omega)^2 \mathbf{I} \check{\mathbf{E}}(\mathbf{r}) \quad (2.19)$$

which shows that in an isotropic medium  $\check{\mathbf{D}}$  is parallel to  $\check{\mathbf{E}}$ , since, by definition, in an isotropic medium the action and the response of the medium have to be parallel [Mid03, pp. 28, 30].

Up to now we have not started manipulating Maxwell's equations, but just paved down the preliminaries. The next target is to retrieve an equation where only the electric field, and in particular a mode, is present. To do so, we have to eliminate  $\check{\mathbf{D}}$  and  $\check{\mathbf{B}}$ . Thus, let us insert (2.8a) in (2.7b) and cancel  $\check{\mathbf{J}}$  thanks to (2.5), obtaining

$$\nabla \times \check{\mathbf{H}} = \varepsilon_0 j\omega (\check{\mathbf{E}} + \check{\mathbf{P}}_L) \quad (2.20)$$

Similarly to (2.20), inserting (2.8b) in (2.7b), canceling  $\check{\mathbf{M}}$  thanks to (2.9), we retrieve

$$\nabla \times \check{\mathbf{E}} = -\mu_0 j\omega \check{\mathbf{H}} \quad (2.21)$$

Applying the curl to (2.21)

$$\nabla \times \nabla \times \check{\mathbf{E}} = -\mu_0 j\omega \nabla \times \check{\mathbf{H}} \quad (2.22)$$

and inserting (2.20) in (2.22), we obtain

$$\nabla \times \nabla \times \check{\mathbf{E}} = -\mu_0 \varepsilon_0 (j\omega)^2 \check{\mathbf{E}} - \mu_0 (j\omega)^2 \check{\mathbf{P}}_L = -\frac{1}{c^2} (j\omega)^2 \check{\mathbf{E}} - \mu_0 (j\omega)^2 \check{\mathbf{P}}_L \quad (2.23)$$

where the relation of the speed light  $c$  in vacuum has been exploited

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} \quad (2.24)$$

Eq.(2.23) presents a double curl of  $\check{\mathbf{E}}$  which is undesired to us since we intend to retrieve an expression directly relating the space derivatives to the time derivative of the field. Hence, we exploit the identity

$$\nabla \times \nabla \times \equiv \nabla (\nabla \cdot) - \nabla^2 \quad (2.25)$$

to write the l.h.s. of (2.23) as

$$\nabla \times \nabla \times \check{\mathbf{E}} = \nabla (\nabla \cdot \check{\mathbf{E}}) - \nabla^2 \check{\mathbf{E}} \quad (2.26)$$

Neglecting the term  $\nabla (\nabla \cdot \check{\mathbf{E}})$  in the previous expression would simplify our calculations. This is surely satisfied if  $\nabla \cdot \check{\mathbf{E}} = \mathbf{0}$ . To see when this happens, let us apply the divergence to the constitutive relation (2.8a):

$$\nabla \cdot \check{\mathbf{D}} = \nabla \cdot (\varepsilon_0 \check{\mathbf{E}} + \check{\mathbf{P}}_L) = \varepsilon_0 \nabla \cdot \check{\mathbf{E}} + \nabla \cdot \check{\mathbf{P}}_L \quad (2.27)$$

Since  $\nabla \cdot \check{\mathbf{D}} = \mathbf{0}$ , we have

$$\varepsilon_0 \nabla \cdot \check{\mathbf{E}} = -\nabla \cdot \check{\mathbf{P}}_L \quad (2.28)$$

$$\Rightarrow \nabla \cdot \check{\mathbf{E}} = \mathbf{0} \Leftrightarrow \nabla \cdot \check{\mathbf{P}}_L = \mathbf{0} \quad (2.29)$$

Thus, to nullify  $\nabla \cdot \check{\mathbf{E}}$  is necessary to set

$$\nabla \cdot \check{\mathbf{P}}_L = \mathbf{0} \quad (2.30)$$

That is, the linear material polarization has to be nonvarying in space, or at least approximately. The connection of  $\check{\mathbf{P}}_L$  with the refractive index has been shown through (2.14), (2.15) and (2.17). Hence, condition (2.30) corresponds to assuming the refractive index variation over the fiber cross-section to be negligible. This is exact in case of step-index fibers (see Sec.2.1) because the refractive index is constant within the core and within the cladding and (2.23) is solved separately for the two parts. Then, the boundary conditions are exploited to match the field at the core-cladding interface. On the opposite, condition (2.30) is only an approximation for, e.g., graded-index fibers (see Sec.2.1). Depending on the desired level of accuracy, this approximation is not always assumed in literature. Yet, for our scopes it is sufficient and, as shown in Ch.5, it allows to reach the same nonlinear propagation equation as common references in literature [PH08; ASM16; Agr19].

Keeping in mind these considerations, we approximate (2.26) as [Agr19, p. 625] [Coe10, p. 17]

$$\nabla \times \nabla \times \check{\mathbf{E}} \approx -\nabla^2 \check{\mathbf{E}} \quad (2.31)$$

and then insert (2.31) in (2.23), obtaining

$$\nabla^2 \check{\mathbf{E}} = \frac{1}{c^2} (j\omega)^2 \check{\mathbf{E}} + \mu_0 (j\omega)^2 \check{\mathbf{P}}_L = -\frac{\omega^2}{c^2} \check{\mathbf{E}} - \omega^2 \mu_0 \check{\mathbf{P}}_L = -k_0^2 \check{\mathbf{E}} - \omega^2 \mu_0 \check{\mathbf{P}}_L \quad (2.32)$$

where the free-space propagation constant  $k_0$  has been defined as

$$k_0 := \omega/c = 2\pi/\lambda_{\text{vacuum}} \quad (2.33)$$

Inserting (2.13) and (2.16) in (2.32) yields the Helmolzt equation

$$\begin{aligned}\nabla^2 \check{\mathbf{E}} &= -\frac{\omega^2}{c^2} \check{\mathbf{E}} - \omega^2 \mu_0 \check{\mathbf{P}}_L = -k_0^2 \check{\mathbf{E}} - \omega^2 \mu_0 \varepsilon_0 \tilde{\chi}^{(1)} \check{\mathbf{E}}(\mathbf{r}) \\ &= -\frac{\omega^2}{c^2} (\mathbf{I} + \tilde{\chi}^{(1)}) \check{\mathbf{E}} = -\frac{\omega^2}{c^2} \varepsilon_r \check{\mathbf{E}} = -k_0^2 \varepsilon_r \check{\mathbf{E}}\end{aligned}\quad (2.34)$$

We specialize (2.34) for a mode, i.e.,  $\check{\mathbf{E}} = \check{\mathbf{E}}_k = A_k \mathbf{F}_k(x, y, \omega) e^{-j\beta_k(\omega)z}$ , where in the last passage we exploited the mode definition (2.1). Moreover, we recall that the Laplacian operator can be written, in a symbolic form, as

$$\nabla^2 = \nabla_T^2 + \frac{\partial^2}{\partial z^2} \quad (2.35)$$

Hence, exploiting these last two facts, we write (2.34) as

$$\left( \nabla_T^2 + \frac{\partial^2}{\partial z^2} \right) \left( \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k(\omega)z} \right) = -\frac{\omega^2}{c^2} \varepsilon_r \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k(\omega)z} \quad (2.36)$$

$$\Rightarrow \nabla_T^2 \mathbf{F}_k(x, y, \omega_0) - \beta_k^2(\omega) \mathbf{F}_k(x, y, \omega_0) = -\frac{\omega^2}{c^2} \varepsilon_r \mathbf{F}_k(x, y, \omega_0) \quad (2.37)$$

(2.37) is our last equation. Solving it, together with the boundary conditions of the fiber, provides the fiber modal profiles. Moreover, from (2.37) descends the eigenvalue equation which yields the propagation constants of the various modes. Notice that in general the modes are divided in guided modes and radiation modes. The former are confined within the core and propagate without attenuation (neglecting the small fiber losses). The latter transfer power away from the core and, hence, get attenuated due to unavoidable losses of the cladding and the boundary [Mar74, p. 96]. The mentioned completeness of the modal set requires considering also the radiated modes [Mar74, p. 83].

## 2.1 Examples of MMFs and Modes

In this thesis we perform simulations with two common types of multimode fibers: step-index (SI) and graded-index (GRIN) fibers, also abbreviated as GIMMF (graded-index multimode fiber). Both structures have a single core with a refractive index profile  $n(x, y)$  that changes from core to cladding. However, in SI fibers the refractive index is position independent within core and cladding. In GRIN fibers the refractive index varies over the core, with maximum value at the center of it and decreasing usually in a quasi-parabolic fashion toward the cladding [Agr19, p. 625], as depicted in Fig.2.1.

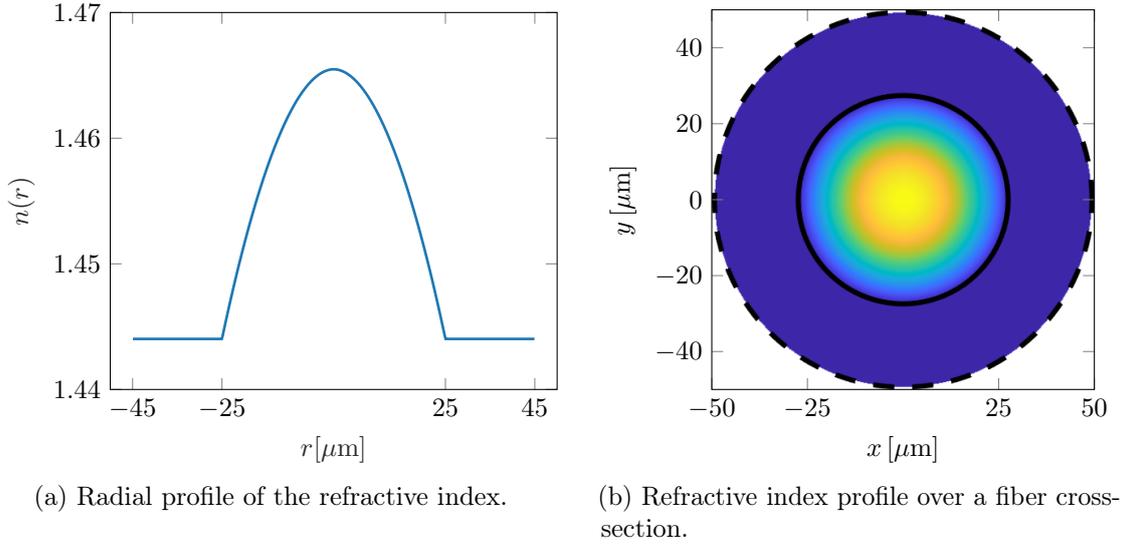


Figure 2.1: Refractive index profile of the GRIN fiber detailed in Sec.2.1.

For the sake of comparison with other papers in the literature (e.g., [SKP21]), the GRIN fiber exploited in this work has been chosen to resemble the one fabricated by Sillard et al. [Sil+16] with diameter of  $50\mu\text{m}$ , cladding of  $90\mu\text{m}$ , a refractive index difference between core and cladding of  $15 \cdot 10^{-3}$  at  $1550\text{ nm}$  and a parameter  $\alpha$  tuning the refractive index profile of  $\alpha = 1.94$ .

This GRIN supports 55 guided modes, out of which we chose at most the first 30 in the various simulations presented in the thesis. The modal profiles of the 15  $x$ -polarized modes are depicted in Fig.2.2, the  $y$ -polarized modes have not been shown since differing only for the polarization. The propagation constant values at  $\lambda_0 = 1550\text{ nm}$  relative to the one of the fundamental mode, i.e.,  $\beta_k(\omega_0) - \beta_1(\omega_0)$ , for the first 30 modes are also shown in Fig.2.3.

For the computation of the modal profiles and the propagation constants it has been exploited a numerical solver developed at the research group.

We consider LP modes, which, as known, are not the exact modes in a fiber, but are an approximation valid in the limit of weakly-guidance [Mar74, Ch.2] [Som06, p. 363]. Every LP mode with a specific spatial pattern supports two orthogonal linear polarizations, say, vertical and horizontal. Notice that, besides the ambiguous language, both polarizations are legitimate modes by themselves, which we call *polarization modes* when we want to emphasize that they differ for the polarization, but have same spatial profile. Hence, we indicate the number of guided modes as  $M = 2N$ , where  $N$  is the number of modes with a different spatial pattern, which we call *spatial modes* when we want to emphasize that they have in general different spatial profiles.

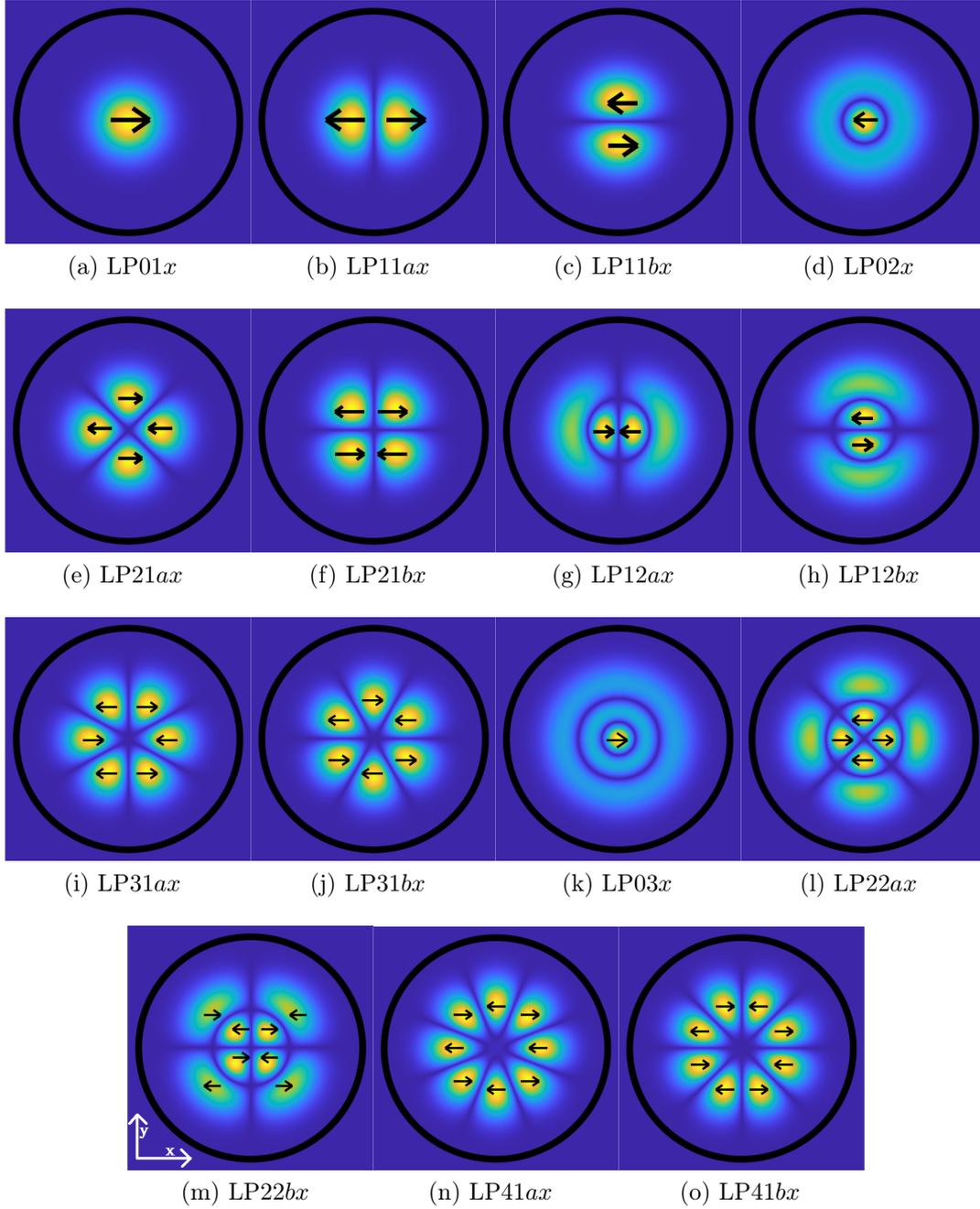


Figure 2.2: Spatial profiles of the first 15  $x$ -polarized guided modes of the GRIN fiber detailed in Sec.2.1. The black circumference represents the core boundary. The color is proportional to the mode field intensity  $\|\mathbf{F}_k(x, y, \omega_0)\|^2$ . The arrows indicate the polarization of the field, i.e., all shown examples are  $x$ -polarized.

With polarization mode coupling we refer to the coupling between two polarizations of the same spatial mode. With spatial mode coupling it is referred to the coupling occurring among modes with different spatial patterns. These concepts are addressed in the next chapters.

We define *(quasi) degenerate group* of modes a set of modes having nearly same value of propagation constants [Som06, pp. 361, 363]. A (quasi) degenerate group of modes is sometimes also referred to as *manifold* [Pal14; SKP21]. For the LP approximation, the modes of a manifold are not just quasi degenerate, but exactly degenerate since having the same propagation constant [Som06, p. 363]. Each LP manifold is formed by either two or four modes [Pal14]. In a manifold with two modes, the two modes have same modal pattern but different polarization ( $\hat{x}$  or  $\hat{y}$ ), and have azimuth order 0. They are the “LP0 $k$ ” modes, where  $k = 1, 2, 3, \dots$ .

In a manifold with four modes, there are two pairs of modes. A pair is characterized for having two modes with same spatial profile, but different polarization ( $\hat{x}$  or  $\hat{y}$ ). One pair is said to have “even” pattern, while the other to have “odd” pattern, which are related one another by a rotation. This last property comes from the modal solution of Maxwell’s equation, which provides either a term  $\cos(n\varphi)$  or  $\sin(n\varphi)$  for the transverse profiles, where  $\varphi$  is the azimuth angle and  $n = 0, 1, 2, \dots$ . The modes containing  $\cos(n\varphi)$  are called even, the others odd [Pal14].

Depending on the particular fiber composition and geometry, there exist LP modes of different manifolds with close values of propagation constants, as evident looking at Fig.2.3. These manifolds are also sometimes considered together [Sil+16; SKP21] and simply referred to as a modal group, being quasi-degenerate. For example, observe the LP02 and LP21 manifolds of Fig.2.3 which have distinct but very close propagation constants.

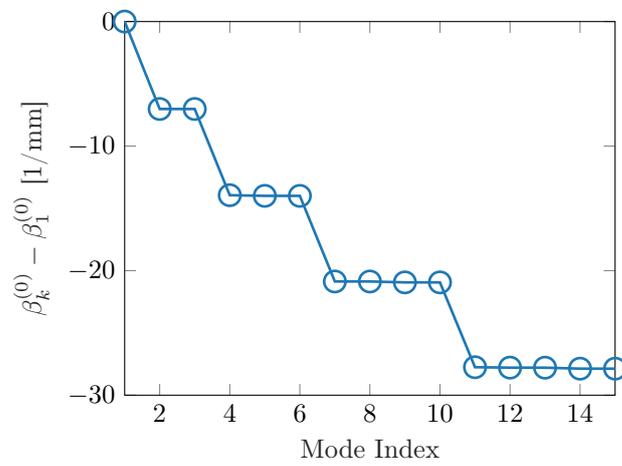


Figure 2.3: Propagation constants  $\beta_k(\omega_0)$  (relative to the fundamental one) for the first 15  $x$ -polarized guided modes of the GRIN fiber detailed in Sec.2.1.

# Chapter 3

## Polarization Mode Coupling

In this chapter we are going to analyze polarization mode coupling, that is, the coupling between the two polarizations of the fundamental modes. In the next chapter we extend the reasoning to a generic multimode structure. We start with a qualitative discussion to give the idea of what mode coupling is.

Prerequisites for this chapter are the concept of Jones and Stokes vectors, whose theory can be found in, e.g., [NJ05; SM05; GK00; Stu18].

We indicate a generic Jones vector as  $|s\rangle$  and its Hermitian conjugate as  $\langle s|$ , a Stokes vector as  $\mathbf{s}$ ; if we want to emphasize they are unit we use the hat, like  $\hat{\mathbf{s}}$  and  $|\hat{\mathbf{s}}\rangle$ . Given the isomorphism between the two spaces, we adopt for every situation the most suitable representation of a vector, either in Jones or Stokes space. The relation between the modal amplitudes  $A_k$  presented in Ch.2 and used in later chapters, and the Jones vector for the 2-polarization case is  $|E\rangle = [A_1, A_2]^T$ .

From an intuitive perspective, mode coupling consists in having modes which do not propagate independently from each other, even though the meaning of the term “mode” is going to be clarified later on. The signal in one mode at a certain position  $z$  depends on both signals launched originally in both modes at the fiber input. Mathematically, this phenomenon can be modeled considering non-zero off-diagonal entries in the channel transfer matrix  $\mathbf{T}(z, \omega)$ . How to choose them is explained in the next sections.

Let us assume that the envelope  $E_x(z = 0, t)$  is being transmitted on the  $x$ -polarization and the envelope  $E_y(z = 0, t)$  on  $y$ -polarization. Then, in presence of mode coupling, since the channel is assumed to be linear in this section, the

input-output relation is expressed as

$$\begin{aligned}
|\tilde{E}(z, \omega)\rangle &= \mathbf{T}(z, \omega) |\tilde{E}(z = 0, \omega)\rangle = \mathbf{T}(z, \omega) \begin{bmatrix} \tilde{E}_x(z = 0, \omega) \\ \tilde{E}_y(z = 0, \omega) \end{bmatrix} \\
\Rightarrow |\tilde{E}(z, \omega)\rangle &= \begin{bmatrix} T_{11}(z, \omega)\tilde{E}_x(z = 0, \omega) + T_{12}(z, \omega)\tilde{E}_y(z = 0, \omega) \\ T_{21}(z, \omega)\tilde{E}_y(z = 0, \omega) + T_{22}(z, \omega)\tilde{E}_x(z = 0, \omega) \end{bmatrix}, \\
\mathbf{T}(z, \omega) &= \begin{bmatrix} T_{11}(z, \omega) & T_{12}(z, \omega) \\ T_{21}(z, \omega) & T_{22}(z, \omega) \end{bmatrix}
\end{aligned} \tag{3.1}$$

where the expression on the second line emphasizes the dependence on both transmitted envelopes of the two received envelopes over  $x$ - and  $y$ -polarizations, conversely to the case of absence of coupling where each transmitted signal evolves independently on the other(s).

Graphically, if we assume to be transmitting Gaussian pulses with different delays along the  $x$ - and  $y$ -polarizations, the envelopes detected over  $x$ - and  $y$ -polarizations at a certain distance  $z$  present two peaks, as visible in Fig.3.1. The two received peaks over each polarization correspond to the two transmitted signals, meaning that the two polarization modes did not propagate independently, but, indeed, coupled into each other.

### 3.1 Origins of Birefringence

An ideal single mode fiber is characterized by a refractive index  $n(x, y, z)$  independent of the position  $z$  along the fiber, i.e.,  $n(x, y, z) = n(x, y)$ , and circular symmetry. Under this assumption, a single mode fiber is known to support two degenerate modes, which differ only by the field polarizations.

As soon as some perturbations are present, the circular symmetry of the fiber is broken and birefringence arises. The perturbations of the ideal structure are due to a non-circular waveguide geometry or due to non-symmetric stresses of the fiber structure [PN97; NJ05; SM05; Kei11]. Both families of phenomena can arise either during the manufacturing process (intrinsic perturbations) or during the cabling phase (extrinsic perturbations) due to externally applied stresses like bending or torsion. The result is that, as experimentally observed, the two modes assume a particular pair of polarizations (called eigenpolarizations) and are no more degenerate. That is, they have different propagation constants and, also, group velocities.

In a real fiber, perturbations and, therefore, birefringence orientation and intensity (whose meaning is going to be detailed below) change with continuity over the fiber position  $z$  and over the frequency  $f$ .

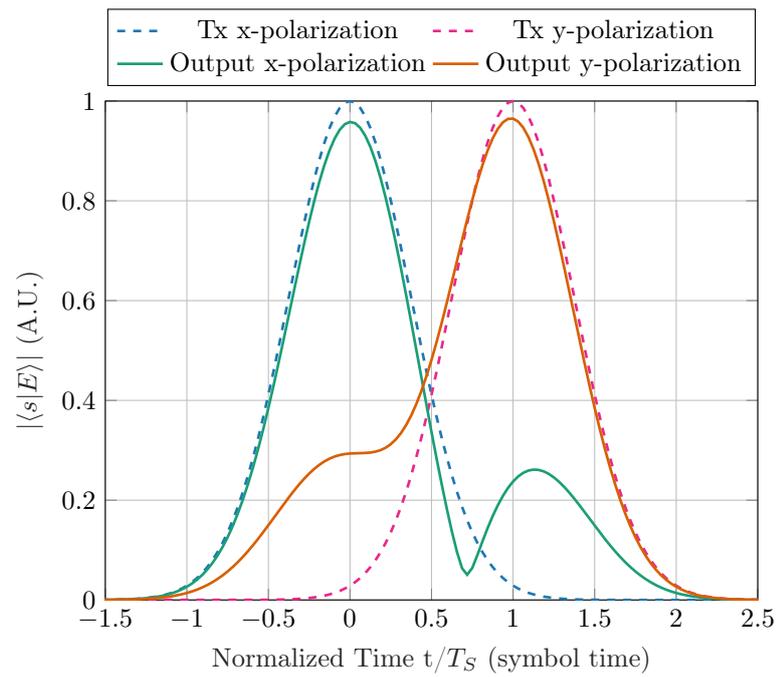


Figure 3.1: Coupling between two different signals (with Gaussian shaping) transmitted at symbol rate  $R_S = 1$  GBaud over the two polarizations of a SMF, due to mode coupling and non-uniform frequency-dependent birefringence over a fiber 100 km long, with birefringence sections 1 km long and  $\Delta n = 10^{-6}$ .

For what concerns the dependence on position, it is common to model the fiber as a series of segments each having uniform birefringence inside, but different orientation of the polarization states and intensity of the birefringence itself, as illustrated in Fig.3.2 ([GK00; NJ05; SM05; PN97; Kei11]). The length of every segment has to be chosen according to the so-called *coupling length* or *correlation length*, which quantifies the length over which the birefringence is approximately constant. A more rigorous definition is provided later on.

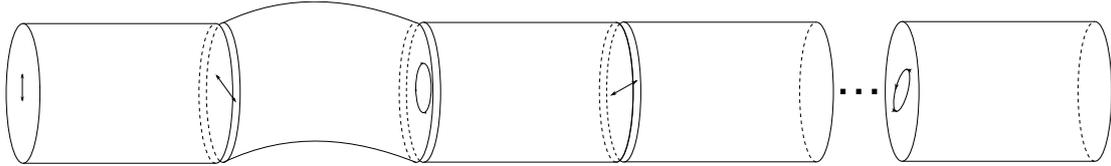


Figure 3.2: A fiber with nonuniform birefringence is modeled as a concatenation of segments with uniform birefringence.

With regards to the dependence of birefringence on frequency, some simplifications are usually assumed as well, in particular the orientation is assumed independent on frequency.

It is the case to observe that it is not the presence of birefringence itself to induce coupling, but the variation of birefringence along the fiber. Furthermore, the mentioned fact that the two eigenpolarizations have different group velocities is responsible for the so-called polarization-mode dispersion (PMD). All these concepts are going to be better explained later on.

## 3.2 Segment of Uniform Birefringence

### 3.2.1 Modelling Birefringence and Intrinsic DGD

It has been found that in presence of birefringence, even if it loses the circular symmetry, the fiber still supports two polarization modes with same spatial profile. They are referred to as *eigenpolarizations* or *local (eigen)modes* and are indicated as  $|\hat{e}_\chi\rangle$  and  $|\hat{e}_\Psi\rangle$  (or simply, without hat,  $|e_\chi\rangle$  and  $|e_\Psi\rangle$ ). When a piece (or segment or section) of fiber has uniform birefringence, like in Fig.3.3, its eigenpolarizations are  $z$ -independent.



Figure 3.3: Fiber segment with uniform birefringence.

We always assume their direction is also frequency independent, even though in reality there is some dependence on frequency even within a uniform birefringence segment. Furthermore, the eigenmodes are always orthogonal and specific for every kind of perturbation ([Han21, p. 24.12]). For example, an external radial pressure on the waveguide determines the presence of two linear orthogonal states (linear birefringence), one parallel and one orthogonal to the direction of the external force; in case of torsion, the two states are circularly polarized (circular birefringence); in the most general case of perturbation they are elliptical. Hence, the modes of a fiber subject to birefringence are no more the linear, but have other polarizations, even though the Jones vector components always refer to the two linearly polarized ideal modes.

To the eigenpolarization states  $|e_\chi\rangle$  and  $|e_\Psi\rangle$  are associated the propagation constants  $\beta_\chi(\omega)$  and  $\beta_\Psi(\omega)$ , respectively. The difference between the propagation constants is commonly referred to as *birefringence*, because it indeed provides a quantitative measure of the birefringence phenomenon. In symbols, birefringence is defined as

$$\Delta\beta := \beta_\chi(\omega) - \beta_\Psi(\omega) = n_\chi k_0 - n_\Psi k_0 = \Delta n k_0 = \Delta n \frac{\omega}{c}, \quad k_0 = \frac{\omega}{c} \quad (3.2)$$

where  $\Delta n = n_\chi - n_\Psi$  is the difference between the refractive indices of the two eigenpolarizations and it commonly lies in  $[10^{-7}; 10^{-5}]$  ([PN97, p. 118], [NJ05, p. 6]), even though other sources (e.g., [Agr19, pp. 14, 190]) indicates its value in  $[10^{-6}; 10^{-4}]$ . Notice that the knowledge of  $\Delta n$  allows to compute  $\Delta\beta$  for simulation purposes. Observe that the birefringence intensity is in general frequency dependent due to the explicit dependence on  $\omega$  and due to material/chromatic dispersion, i.e., the implicit dependence of the refractive index difference  $\Delta n$  on frequency. This second contribution is neglected in our analysis.

The other common way to compute  $\Delta\beta$  is through the *beat length*  $L_B$  defined as the length after which an input polarization returns to its initial state. Indeed, the effect of birefringence on an input polarization is to rotate it with  $z$  due to the accumulation of a phase difference between the two eigenpolarizations over which the input polarization can be decomposed (see (3.16)). Then, we are looking for a phase difference  $\Delta\Phi = \Delta\beta z = 2\pi$  after a length  $z = L_B$ . So,

$$\Delta\beta = 2\pi/L_B \quad (3.3)$$

which is the other common way of computing  $\Delta\beta$  in simulations once  $L_B$  is known (or assumed to be), without the need for estimating  $\Delta n$ . Usually,  $L_B$  is assumed to be in the range of few meters ([Han21; NJ05]), even though other sources (e.g., [Agr19, p. 190]) indicates it to be as small as 1 cm in fibers with high-birefringence ( $\Delta n \approx 10^{-4}$ ).

To complete the digression, combining Eq.3.3 and  $\Delta\beta = \Delta n \frac{\omega}{c} = \Delta n \frac{2\pi}{\lambda}$  (from Eq.3.2), we get the common definition of beat length as [NJ05, p. 4]:

$$L_B = 2\pi/\Delta\beta = \lambda/\Delta n \quad (3.4)$$

Since birefringence is characterized both by a vectorial quantity, the eigenpolarization states  $\{|\hat{e}_\Psi\rangle, |\hat{e}_\chi\rangle\}$ , and by a scalar quantity, the birefringence  $\Delta\beta$ , the birefringence Stokes vector  $\boldsymbol{\beta}$  is defined summarizing both contributions:

$$\boldsymbol{\beta} := \Delta\beta\hat{\boldsymbol{\beta}}, \quad \hat{\boldsymbol{\beta}} := \hat{e}_\chi \quad (3.5)$$

The definition 3.5 associates to the birefringence vector the direction of the *slow principal axis* (or *slow birefringence axis* or *slow eigenpolarization state* or *extraordinary axis*), that is the direction of the eigenpolarization in Stokes space with higher propagation constant and lower group velocity (which is the reason for the term “slow”). We have arbitrarily chosen  $\hat{e}_\chi$ , letter  $\chi$  as subscript, to indicate the slow axis and  $\hat{e}_\Psi$ , letter  $\Psi$  as subscript, to indicate the *fast* (or *ordinary*) axis.

The two polarizations also have different group velocities

$$\begin{aligned} v_{g,\Psi} &:= 1/\left(\frac{d\beta_\Psi}{d\omega}\right) \\ v_{g,\chi} &:= 1/\left(\frac{d\beta_\chi}{d\omega}\right) \end{aligned} \quad (3.6)$$

and, thus, propagation delays

$$\begin{aligned} \tau_\Psi &:= \frac{d\beta_\Psi}{d\omega} z \\ \tau_\chi &:= \frac{d\beta_\chi}{d\omega} z \end{aligned} \quad (3.7)$$

Hence, two replicas of the same transmitted signal over the two eigenmodes accumulate a *Differential Group Delay* (DGD)  $\Delta\tau$  after a distance  $z$  (neglecting higher-order effects in frequency, i.e., the refractive index difference  $\Delta n(\omega) = \Delta n$  is considered frequency independent)

$$\Delta\tau := \tau_\chi - \tau_\Psi = \frac{d\beta_\Psi}{d\omega} z - \frac{d\beta_\chi}{d\omega} z = \left(\frac{d\Delta\beta}{d\omega}\right) z = \frac{\Delta n}{c} z \quad (3.8)$$

The different group delays (GDs) between the two eigenpolarizations are the source of the *Polarization Mode Dispersion* (PMD) phenomenon, which broadens and distorts the transmitted signals, as it will be clear later on.

Since it is going to be useful later on, we also define the average or common propagation constant  $\beta_c(\omega)$  and the average or common group delay  $\tau_c$  between

the two eigenpolarizations as

$$\begin{aligned}\beta_c(\omega) &:= (\beta_\Psi(\omega) + \beta_\chi(\omega))/2 \\ \tau_c &:= \frac{d\beta_c}{d\omega}z = (\tau_\Psi + \tau_\chi)/2\end{aligned}\tag{3.9}$$

The two group delays can then be written as

$$\begin{aligned}\tau_\Psi &= \frac{d\beta_\Psi}{d\omega}z = \tau_c - \frac{\Delta\tau}{2} \\ \tau_\chi &= \frac{d\beta_\chi}{d\omega}z = \tau_c + \frac{\Delta\tau}{2}\end{aligned}\tag{3.10}$$

which, again, justifies the term “slow” associated to the mode with the higher refractive index,  $|\hat{e}_\chi\rangle$  in our conventions. Notice that all the quantities, in particular the DGD in Eq.(3.8) between the two eigenpolarizations, have been defined relative to the case of uniform birefringence, which happens only for small lengths in real fibers. The quantity

$$\frac{d\Delta\beta}{d\omega}\tag{3.11}$$

which is related to the DGD between the two eigenpolarizations in Eq.3.8, is usually referred to as *intrinsic* or *short-length PMD* ([PWN91, p. 372], [PN97, p. 120] [NJ05, p. 5]) to distinguish it from the DGD and the PMD defined for fibers with non-uniform birefringence.

Given that the fiber perturbations are random, every fiber (i.e., every channel transfer matrix) is a realization of an ensemble. Therefore, the DGD is also a random variable fixed a certain fiber length  $z$  and, if we let  $z$  vary, the DGD is a stochastic process over  $z$ . It can be shown that the mean DGD increases linearly with  $z$  for the case of uniform birefringence, or, more precisely, for the case when the fiber length  $z$  is smaller or comparable to the correlation length. In Sec.3.3 the quantities presented above are generalized to the case of non-uniform birefringence and it will be mentioned that the mean DGD increases with the square-root of distance in that case.

### 3.2.2 Channel Transfer Matrix for Uniform Birefringence

The two orthogonal states  $|e_\chi\rangle$  and  $|e_\Psi\rangle$  characterizing a uniform birefringence segment form an orthogonal basis which allows to express the modal amplitude  $|E\rangle$  at any coordinate  $z$ . They are called eigenpolarizations or local (eigen)modes because an input field parallel to one of them keeps the same polarization all over the fiber segment. We indicate with  $\mathbf{T}$  the propagation operator in Jones space of the fiber segment with uniform birefringence. Since the system is considered

linear in this chapter, its propagation operator  $\mathbf{T}$  in the frequency domain is a transfer matrix. Then, the eigenpolarizations  $|e_\chi\rangle$  and  $|e_\Psi\rangle$ , based on the empirical observations as mentioned in Sec.3.1, are the eigenvectors of  $\mathbf{T}$ . However, it is important to stress that the eigenpolarizations are no more the eigenvectors of the propagation operator  $\mathbf{T}$  when the birefringence is not uniform.

Given that the eingepolarizations are the modes for a fiber of uniform birefringence as said in Sec.3.1, their phases change over length as  $e^{-j\beta_{\Psi/\chi}(\omega)z}$ .

Hence, the input-output relation for the fiber segment with uniform birefringence is

$$\begin{aligned} |\tilde{E}(z, \omega)\rangle &= \mathbf{T}(z, \omega) |\tilde{E}(z = 0, \omega)\rangle = \mathbf{C}^{-1} \mathbf{T}' \mathbf{C} |\tilde{E}(z = 0, \omega)\rangle, \\ \mathbf{T} &= \mathbf{C}^{-1} \mathbf{T}' \mathbf{C} \\ \mathbf{C} &= \begin{bmatrix} \langle e_\Psi | \\ \langle e_\chi | \end{bmatrix} \quad \mathbf{T}' = \begin{bmatrix} e^{-j\beta_\Psi z} & 0 \\ 0 & e^{-j\beta_\chi z} \end{bmatrix} \end{aligned} \quad (3.12)$$

where  $\mathbf{T}'$  is not the derivative of  $\mathbf{T}$ , but a different matrix.

Notice that  $\mathbf{T} = \mathbf{C}^{-1} \mathbf{T}' \mathbf{C}$  provides a diagonalization of the transfer matrix  $\mathbf{T}$ , where  $\mathbf{C}$  is a the change-of-coordinate matrix from the Cartesian  $\{x, y\}$  coordinate system to the eigenpolarization coordinate system. As such,  $|e_\chi\rangle$  and  $|e_\Psi\rangle$  have to be the eigenvectors of the propagation operator  $\mathbf{T}$ , consistently with our definitions, and  $e^{-j\beta_\Psi z}$  and  $e^{-j\beta_\chi z}$  are the corresponding eigenvalues.

Besides the  $\mathbf{T}$  matrix, there are two other matrices commonly used to represent the propagation operator. One is the Jones matrix  $\mathbf{U}$  defined as

$$\mathbf{T}(z, \omega) = e^{-j\beta_c(\omega)z} \mathbf{U}(z, \omega) \quad (3.13)$$

where  $\beta_c(\omega)$ , defined by Eq.3.9, is the average value of propagation constant and so  $\beta_c(\omega)z$  is the common phase term which is factored out from  $\mathbf{T}$ . Indeed, the role of the common phase term is to account for a common constant phase rotation and a common propagation delay (beside higher-order effects like CD), while here we are interested in the mode-dependent effect modeled by  $\mathbf{U}$  which are due to the difference between the phases of the eigenpolarizations.

From  $\mathbf{U}$ , it is possible to define the Stokes rotation matrix  $\mathbf{R}$ , that is the propagation operator in Stokes space, as ([GK00, p. 4.2]):

$$\mathbf{R}\sigma = \mathbf{U}^H \sigma \mathbf{U} \quad (3.14)$$

where  $\sigma = (\sigma_1, \sigma_2, \sigma_3)$  is the Pauli spin vector and its components  $\sigma_i$  are the Pauli spin matrix, that can be seen as basis elements of the linear space of  $2 \times 2$  Hermitian matrices [Ant+12].

$$\mathbf{s}(z, \omega) = \mathbf{R}(z, \omega) \mathbf{s}(z = 0, \omega) \quad (3.15)$$

Relation 3.14 is not just important *per se*, but it reveals that the  $\mathbf{R}$  matrix does not account for the absolute phase evolution, only for the relative phase difference between the two eigenpolarization components, since it is directly related to  $\mathbf{U}$  instead of  $\mathbf{T}$ . This is consistent with the fact that a Stokes vector is invariant to a phase offset in the components of the corresponding Jones vector [NJ05, p. 2], which boils down from having only three real parameters/degrees of freedom in Stokes space, while a Jones vector is defined through four real values (two for the amplitudes, two for the phases).

We exploit the propagation matrices  $\mathbf{T}$ ,  $\mathbf{U}$ ,  $\mathbf{R}$  interchangeably, as long as the proper space is selected to represent the input and output polarization vectors. All three matrices are unitary assuming no loss, which we remember it means their inverse equals their Hermitian transpose, i.e.,  $\mathbf{T}^{-1} = \mathbf{T}^H \Leftrightarrow \mathbf{T}\mathbf{T}^H = \mathbf{I}$ . Extending the role of orthogonal matrices as rotation operators for Euclidean spaces\*, unitary matrices can be interpreted as responsible for generalized rotations. Unfortunately, it is harder to visualize such a rotation in Jones space since they work with complex quantities. Yet, the rotation of an input polarization due to the channel matrix  $\mathbf{R}$  is immediately visible in Stokes space through the use of a Poincaré sphere. Then, the relation between  $\mathbf{U}$  and  $\mathbf{R}$  (through, e.g., Eq.3.14) allows to extend the terms “rotation”, “orientation” and “direction” of a polarization vector to the Jones space as well.

As a side remark, to fulfill the unitarity of the propagation matrices  $\mathbf{T}$ ,  $\mathbf{U}$  and  $\mathbf{R}$ ,  $\mathbf{B}$  of Eq.3.1 has to be Hermitian ([Ant+12, p. 9]) because in this way  $-j\mathbf{B}z$  is skew-Hermitian, that is,  $(-j\mathbf{B}z)^H = +j\mathbf{B}z$ , and then the exponential  $\mathbf{T} = e^{-j\mathbf{B}z}$  of a skew-Hermitian matrix is unitary [wik].

### 3.2.3 Propagation in a Segment with Uniform Birefringence

In the framework of PMD, to study the effect of birefringence on the propagating electric field, the evolution with distance and frequency of the input polarization is usually considered over the Poincaré sphere. Indeed, it provides a clear and intuitive picture of the phenomena.

Let us assume to be working with a single modulating signal  $\tilde{a}(\omega)$ , transmitted over the arbitrary input polarization  $|s(z=0)\rangle \notin \{|e_\Psi\rangle, |e_\chi\rangle\}$ , such that the input field is  $|\tilde{E}(z=0, \omega)\rangle = \tilde{a}(\omega)|s(z=0)\rangle$ . The output field can be easily computed through Eq.3.12, firstly by decomposing the input field over the two eigenpolarizations, which are the eigenvectors of  $\mathbf{T}$ , as

$$|\tilde{E}(z=0, \omega)\rangle = \tilde{a}(\omega)|s(z=0)\rangle = c_\Psi \tilde{a}(\omega)|e_\Psi\rangle + c_\chi \tilde{a}(\omega)|e_\chi\rangle \quad (3.16)$$

---

\*Actually; orthogonal matrices account for both rotations and reflections.

where  $c_\chi = \langle e_\chi | s \rangle$  and  $c_\Psi = \langle e_\Psi | s \rangle$  are the projection of  $|s\rangle$  over the eigenpolarizations. Then, the propagation of the electric field is eased by considering the propagation of the eigenpolarizations

$$\begin{aligned}
 |\tilde{E}(z, \omega)\rangle &= \tilde{a}(\omega)c_\Psi e^{-j\beta_\Psi(\omega)z} |e_\Psi\rangle + \tilde{a}(\omega)c_\chi e^{-j\beta_\chi(\omega)z} |e_\chi\rangle \\
 &= \tilde{a}(\omega) \left( c_\Psi e^{-j\beta_\Psi(\omega)z} |e_\Psi\rangle + c_\chi e^{-j\beta_\chi(\omega)z} |e_\chi\rangle \right) \\
 &= \tilde{a}(\omega) e^{-j\beta_c z} \left( c_\Psi e^{-j\frac{\Delta\beta(\omega)}{2}z} |e_\Psi\rangle + c_\chi e^{+j\frac{\Delta\beta(\omega)}{2}z} |e_\chi\rangle \right) \quad (3.17) \\
 &= \tilde{a}(\omega) e^{-j\beta_c z} |s(z, \omega)\rangle, \\
 |s(z, \omega)\rangle &= c_\Psi e^{-j\frac{\Delta\beta(\omega)}{2}z} |e_\Psi\rangle + c_\chi e^{+j\frac{\Delta\beta(\omega)}{2}z} |e_\chi\rangle
 \end{aligned}$$

Hence, due to birefringence, the two eigenpolarization components accumulate a phase difference  $\Delta\beta z$  with varying  $z$  such that the output polarization changes continuously. On the Poincaré sphere, this evolution is a precession of the input polarization around the birefringence axis  $\boldsymbol{\beta} \parallel \hat{e}_\chi$ , as shown in Fig.3.4a. Notice from the figure that the full circle is run over a length equal to the beat length,  $z = L_B$ , as foreseen by its definition.

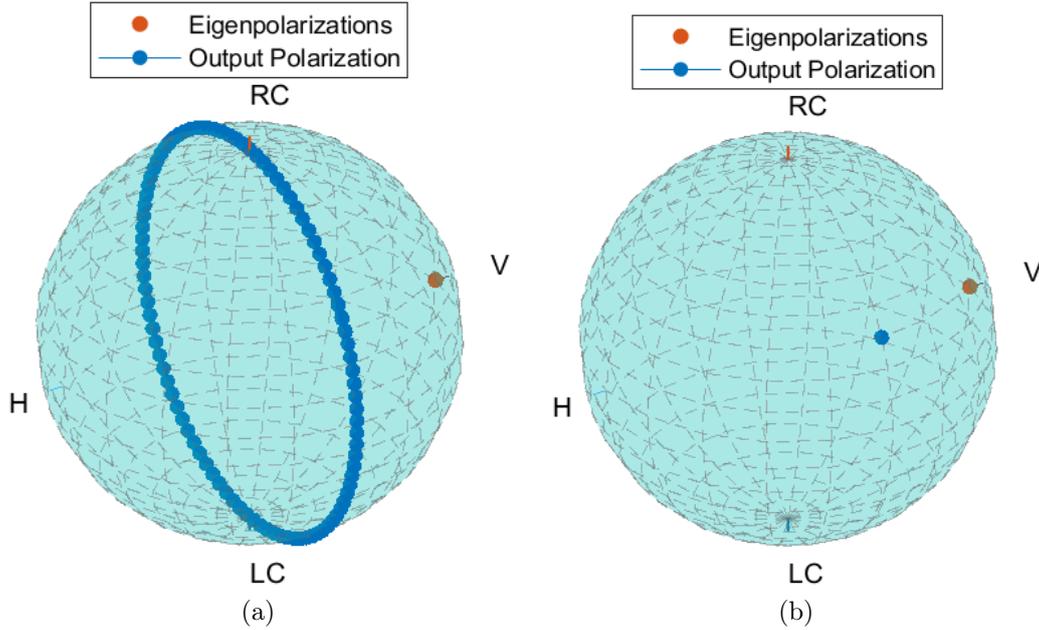


Figure 3.4: Output polarization evolution over the Poincaré sphere for a fiber with frequency-independent uniform birefringence: in 3.4a varying  $z \in [0, L_B]$ , fixed  $f = \omega_0$ , in 3.4b varying  $f = [-B/2, B/2]$  ( $B$  is the signal bandwidth), fixed  $z$ .

The precession is modeled with the following differential equation in Stokes

space:

$$\frac{d\mathbf{s}}{dz} = \boldsymbol{\beta} \times \mathbf{s} \quad (3.18)$$

where  $\boldsymbol{\beta}$  is the local birefringence vector defined in Def.3.5. Eq.3.18 is called the law of infinitesimal rotation and its integration provides an expression for the Stokes rotation matrix  $\mathbf{R}$  ([GK00]).

In the next section we analyze in more detail the distortions due to uniform birefringence distinguishing the case of frequency independence from the one of frequency dependence, already noticing that Eq.3.17 indicates that the received polarization  $|s(z, \omega)\rangle$  is in general frequency dependent due to  $\Delta\beta(\omega) = \Delta n\omega/c$  being frequency dependent.

### 3.2.4 Frequency Independent Birefringence

If the propagation constants of the two eigenmodes are frequency independent, we can write their average value  $\beta_c$  and their difference from Eq.3.9 and Eq.3.2 as

$$\begin{aligned} \beta_c(\omega) &= \beta_c(\omega_0) = (\beta_\chi(\omega_0) + \beta_\Psi(\omega_0))/2 \\ \Delta\beta(\omega) &= \Delta\beta(\omega_0) = \Delta n(\omega_0) \omega_0/c \end{aligned} \quad (3.19)$$

In this way, we are neglecting all higher-order effects, in particular the different group velocity between the two eigenpolarizations. Then, Eq.3.17 allows to easily retrieve the time-domain expression of the received field as

$$\begin{aligned} |E(z, t)\rangle &= a(t)e^{-\beta_c(\omega_0)z} |s(z)\rangle, \\ |s(z)\rangle &= c_\Psi e^{-j\frac{\Delta\beta(\omega_0)}{2}z} |e_\Psi\rangle + c_\chi e^{+j\frac{\Delta\beta(\omega_0)}{2}z} |e_\chi\rangle \end{aligned} \quad (3.20)$$

The simulation of Fig.3.4 considered this assumption and, indeed, Fig.3.4b shows that the output polarization is frequency independent.

Considering the time evolution of the signals, the rotation of the polarization vector determines a reduced detection power when the receiver detects the signal  $\langle s_D | E(z, t) \rangle$  along a specific polarization  $|s_D\rangle$  different from the received polarization  $|s(z)\rangle$ . Fig.3.5a shows a situation where the fiber segment has principal axes aligned with the  $x$ - and  $y$ -axis, while the transmitted and detected polarizations are linear at  $45^\circ$ . This is the case of a coherent detection receiver, where the receiver detects a polarizations which is not, in general, one of the eigenpolarizations. It is visible that the signal detected on the same transmit polarization (yellow curve) has lower amplitude than the actual transmitted signal, due to the phase difference accumulated between the two eigenpolarizations with respect to the input. The phase difference is evident also from the plot in the complex plane, Fig.3.5b, which compares the envelopes of the transmitted signal with the envelopes of the  $x$ -

and  $y$ - components, that are the eigenpolarization in this case. Besides phase rotation leading to a reduced detection power, there is no envelope distortion in this situation, as foreseen by Eq.3.20.

On the opposite, no distortion is observed if the detected polarization is exactly the one which is received, or, equivalently, if the received signal is rotated back, as a simple form of phase distortion compensation. This is the case of direct-detection, as visible in Fig.3.5b where the blue curve of the transmitted envelope is hidden behind the red curve of the output polarization. The reason is that the squaring operation performed by the receiver photodiode coincides with a projection over the actual received polarization. That is,  $\|E(z, t)\|^2 = \langle E|E \rangle = |a(t)|^2 \langle s(z)|s(z) \rangle = |a(t)|^2$ .

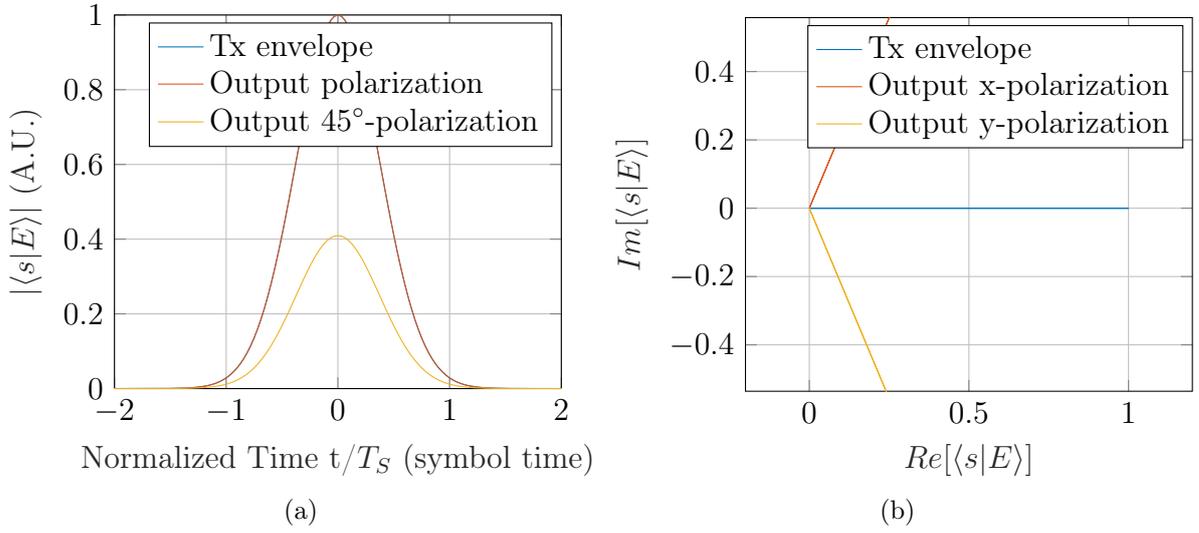


Figure 3.5: Envelope (3.5a) and phase distortion (3.5b) due to frequency-independent uniform birefringence. In Fig.3.5a the blue curve of the transmitted envelope is hidden behind the red curve of the output polarization for the reasons explained in the text.

### 3.2.5 Frequency Dependent Birefringence

We now take into account the first-order frequency dependence of birefringence which implies a differential group delay between the two eigenmodes, as mentioned in Sec.3.2.1. For simulation purposes and for graphical representation, it is possible to neglect the common propagation delay  $\tau_c$  (Def.3.9), as tacitly done in the previous sections.

The electric field at any position  $z$  is computed from Eq.3.17 as

$$\begin{aligned} |\tilde{E}(z, \omega)\rangle &= \tilde{a}(\omega)e^{-j\beta_c(\omega)z} |s(z, \omega)\rangle, \\ |s(z, \omega)\rangle &= c_\Psi e^{-j\frac{\Delta\beta(\omega)}{2}z} |e_\Psi\rangle + c_\chi e^{+j\frac{\Delta\beta(\omega)}{2}z} |e_\chi\rangle = (c_\Psi e^{-j\frac{\Delta\beta(\omega_0)}{2}z} |e_\Psi\rangle + c_\chi e^{+j\frac{\Delta\beta(\omega_0)}{2}z} |e_\chi\rangle) \\ &\quad + (c_\Psi e^{-j\frac{\Delta\beta'(\omega_0)}{2}\Delta\omega z} |e_\Psi\rangle + c_\chi e^{+j\frac{\Delta\beta'(\omega_0)}{2}\Delta\omega z} |e_\chi\rangle) \end{aligned} \quad (3.21)$$

where, within our first-order approximations,

$$\begin{aligned} \beta_c(\omega) &\approx \beta_c(\omega_0) + \left(\frac{d\beta_c}{d\omega}\right)_{\omega_0} \Delta\omega \\ \Delta\beta(\omega) &\approx \Delta\beta(\omega_0) + \left(\frac{d\Delta\beta}{d\omega}\right)_{\omega_0} \Delta\omega \end{aligned} \quad (3.22)$$

and  $\Delta\omega = \omega - \omega_0$  is the discrepancy with respect to the central frequency  $\omega_0$ .

Hence, the time-expression is

$$|E(z, t)\rangle = c_\Psi a(t - \tau_\Psi) e^{-j\beta_\Psi(\omega_0)z} |e_\Psi\rangle + c_\chi a(t - \tau_\chi) e^{-j\beta_\chi(\omega_0)z} |e_\chi\rangle \quad (3.23)$$

which emphasizes that the received signal is a superposition of two differently delayed replicas, with the two eigenpolarizations as basis vectors.

Figure 3.6b shows an example of this situation, where, similarly to the example detailed in the previous section, the signal has been transmitted with equal power on both  $x$ - and  $y$ -components, that is, with a linear  $45^\circ$  polarization. Hence,  $c_\Psi = c_\chi = 1/\sqrt{2}$ . If the field is detected on the same linear  $45^\circ$  polarization (as it happens for a coherent-receiver), the envelope along that polarization appears evidently to be the superposition of the two differently delayed replicas. Yet, notice that the modulus of the sum is lower than the sum of the moduli, hence, one would observe a power reduction. If, instead, direct-detection is performed, the resulting signal is given by the summation of the square moduli of the envelopes. Note that since the channel transfer matrix is unitary, the instantaneous power of the Fourier transform of the signal (along input and output polarizations) is maintained and the average power is conserved both in time and frequency domains thanks to Parseval's theorem. The instantaneous power is not conserved in time domain because the transfer matrix is frequency dependent. Clearly, these properties of power are not specific for the case of uniform birefringence, but are valid whenever the channel transfer matrix is unitary.

The presence of DGD detailed in the above example is just one of two faces of the frequency-dependent birefringence. The other is polarization change with frequency of the electric field. Indeed, if we fix a position  $z$  over fiber length, the output polarization changes with frequency, as modeled by Eq.3.21. The evolution

of the output polarization with frequency can be seen on the Poincaré sphere, as Fig.3.6a illustrates. Notice that for a segment of uniform birefringent fiber, a frequency or a position variation of the same value produces the same polarization rotation due to the same dependence of the phase difference  $\Delta\beta z = \frac{\Delta n}{c}\omega z$  on  $\omega$  and  $z$ . Indeed, Fig.3.4a and Fig.3.6a show qualitatively the same polarization rotation effect on Poincaré sphere.

An analytic link between the DGD and the polarization rotation with frequency is offered later in Ch.3.3.5 through Eq.3.55.

However, if the signal is transmitted over one of the two eigenpolarizations, polarization variation with frequency and pulse broadening due to DGD do not happen, because only one replica of the same signal is received. That is, the received envelope is not distorted, besides a phase retardation  $e^{-j\beta_{\psi/\chi}(\omega_0)z}$ . This is the idea behind the *Principal States of Polarization (PSP)*, that is, a pair of orthogonal polarizations which are frequency-independent to first order in frequency and, thus, a signal transmitted over one of them propagates without coupling to the other PSP and it is received without distortion. This might be unclear at the moment since we have been dealing with a fiber of uniform birefringence so far, but it is going to be clear in the next section. For the time being it suffices to state that PSPs coincide with the eigenpolarizations for a fiber with uniform birefringence where the eigenpolarizations are frequency independent, as we have already explicitly assumed.

### 3.3 Concatenation of Segments: Polarization Mode Coupling

As mentioned in Sec.3.1, birefringence in practice varies with position  $z$ . Hence, modeling a fiber as a single section of uniform birefringence is unrealistic (except for Polarization Maintaining Fibers). On the opposite, a common modeling approach consists in considering the fiber of length  $L$  as a concatenation of  $N$  segments of length  $L_S$  each one uniformly perturbed in a different way, i.e., each one characterized by its own eigenpolarizations and birefringence intensity. Within each segment of uniform birefringence, the evolution of the polarization in Stokes space still consists in a rotation about the corresponding birefringence axis according to the law of infinitesimal rotation 3.18, as shown in Fig.3.7a. Notice that if the beat length is longer than the segment length, as in 3.7b, than the polarization does not complete a full circle on the Poincaré sphere and, thus, the trajectory results qualitatively harder to follow, even though the law of infinitesimal rotation 3.18 holds.

The input-output relation can be again described by a unitary transfer matrix

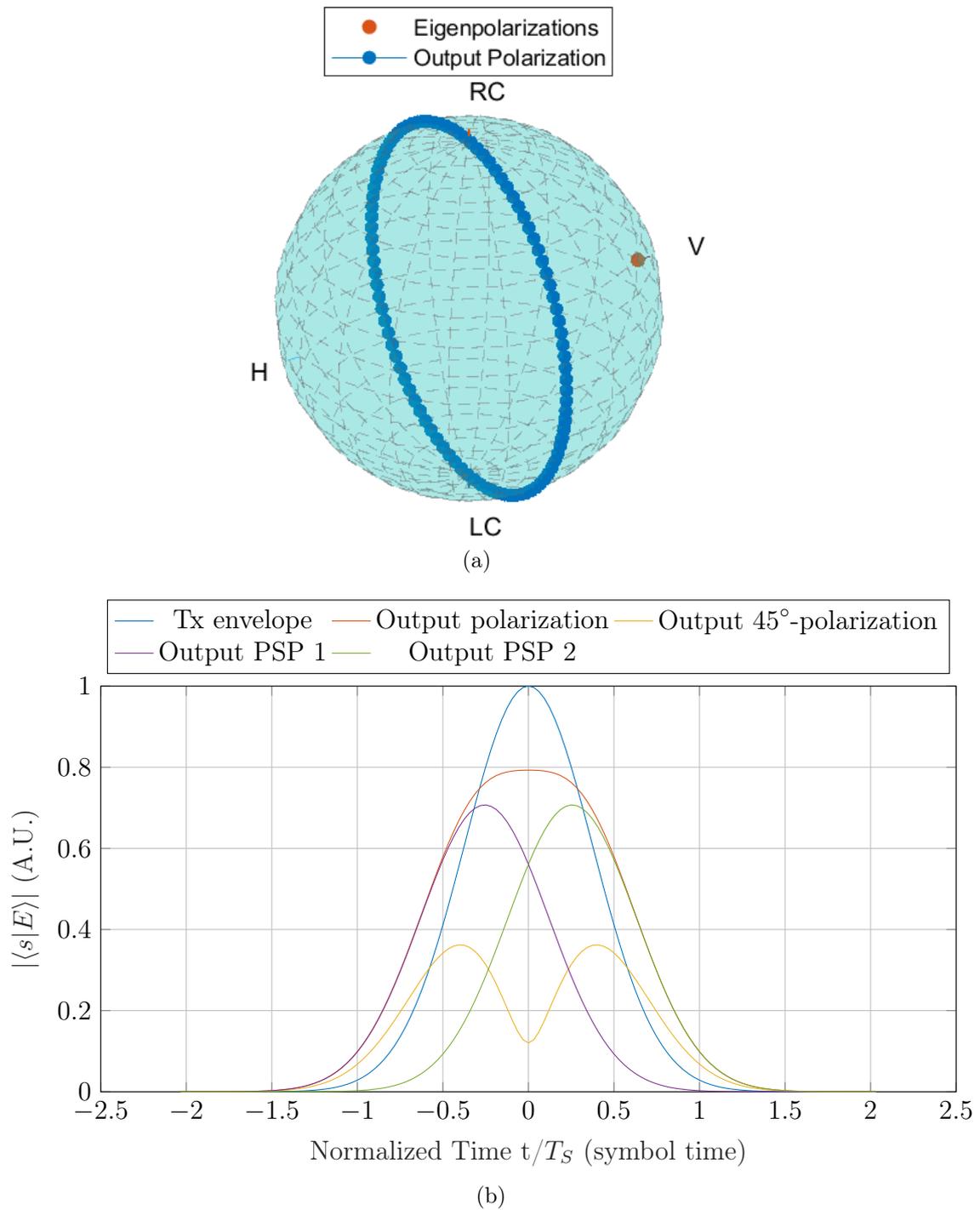


Figure 3.6: Fiber with uniform birefringence, 150 km long,  $\Delta n = 10^{-7}$ ,  $R_B = 10$  Gbps, linear  $45^\circ$  transmit polarization. In 3.6a polarization evolution with  $f$  over the Poincaré sphere, fixed  $z = 150$  km. In 3.6b envelope comparison.

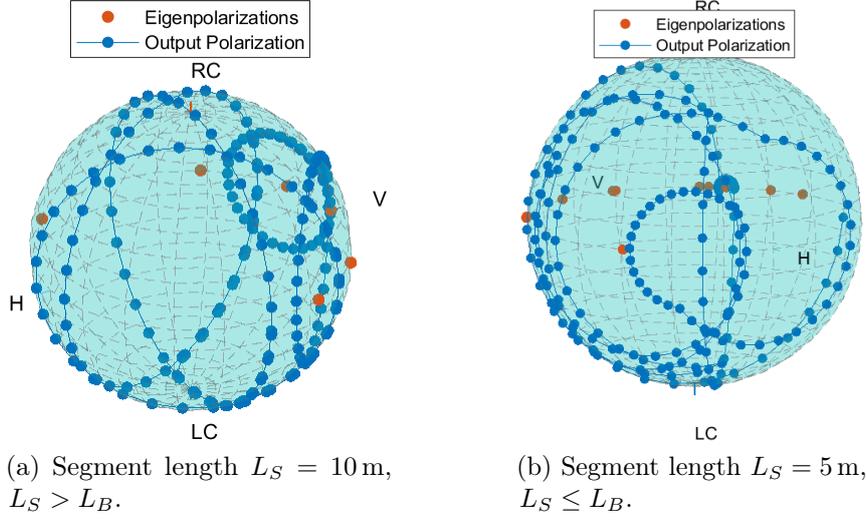


Figure 3.7: Field polarization evolution with  $z$  over the Poincaré sphere (fixed  $\omega = \omega_0$ ), within a fiber made of concatenated segments with different linear birefringence with beat length  $L_B = 7.75$  m.

$\mathbf{T}(z, \omega)$  as

$$\mathbf{T}(z \in [(n-1)L_S, nL_S], \omega) = e^{-j\mathbf{B}(z-(n-1)L_S)} \prod_{i=1}^{n-1} \mathbf{T}_i,$$

$$\mathbf{T}_i = \mathbf{C}^{-1} \mathbf{T}' \mathbf{C}$$

$$\mathbf{C}_i = \begin{bmatrix} \langle e_{\chi,i} | \\ \langle e_{\psi,i} | \end{bmatrix}$$

$$\mathbf{T}'_i = \begin{bmatrix} e^{-j\beta_{\chi,i}L_S} & 0 \\ 0 & e^{-j\beta_{\psi,i}L_S} \end{bmatrix}$$
(3.24)

where  $n = 1, 2, \dots$  is the index of current segment and  $L_S$  is the segment length.

In analogy with Eq.3.9, the concepts of common phase and common group delay are extended to the non-uniform fiber as

$$\Phi_c(z \in [(n-1)L_S, nL_S], \omega) = \beta_{c,n}(\omega) \cdot (z - (n-1)L_S) \sum_{i=1}^{n-1} \Phi_{c,i}(\omega),$$

$$\Phi_{c,i}(\omega) = \beta_{c,i}(\omega) L_S$$

$$\tau_c(z \in [(n-1)L_S, nL_S], \omega) = \frac{d\Phi_c}{d\omega} = \frac{\partial \beta_{c,n}}{\partial \omega} \cdot (z - (n-1)L_S) \sum_{i=1}^{n-1} \tau_{c,i}(\omega),$$

$$\tau_{c,i}(\omega) = \frac{\partial \beta_{c,i}}{\partial \omega} L_S$$
(3.25)

where  $\Phi_{c,i}$  and  $\tau_{c,i}$  are the common phases and delays accumulated from the start of the  $i$ -th segment to its end (see Eq.3.9). If it is assumed  $\Delta n_i$  to be frequency-independent, so is  $\tau_c$ . Then, similarly to Eq.3.13, a common phase term  $\Phi_c(z, \omega)$  can be factored out of the transfer matrix  $\mathbf{T}$  to obtain the  $\mathbf{U}$  matrix:

$$\mathbf{T}(z, \omega) = e^{-j\Phi_c(z, \omega)} \mathbf{U}(z, \omega) \quad (3.26)$$

The previous equation can be directly exploited for simulation purposes in the linear regime, assuming some statistics for the birefringence vector  $\beta_i$  modeled as a stochastic process of the segment index  $i$ . E.g., Poole et al. [PWN91] and Mecozzi et. al. [SM05] modeled  $\beta_i$  as a random Gaussian noise with a certain mean  $\bar{\beta}_M$ . We expect a real fiber to have some elasticity, that is, consecutive segments are expected to have correlated birefringence. Therefore, we assumed a random-walk process in Cartesian coordinates over Poincaré sphere for  $\hat{\beta}$  and a random Gaussian noise process with a certain mean  $\overline{\Delta\beta}$  for the modulus. We recall that a random-walk process means that the increment  $\hat{\beta}_{i+1} - \hat{\beta}_i$  is a random Gaussian noise. The mean value  $\overline{\Delta\beta}$  is generated through the mean refractive index  $\overline{\Delta n}$ . However, to relax the notation in the whole thesis we use  $\Delta n$  instead of  $\overline{\Delta n}$  to refer to the mean value of the refractive index difference.

A realization example of the birefringence vector direction is visible in the Poincaré sphere of Fig.3.17a. Then, from the knowledge of  $\hat{\beta}$  it is easy to retrieve the eigenpolarizations and so the change-of-basis matrices  $\mathbf{C}_i$  to compute the channel transfer matrix  $\mathbf{U}$  of Eq.3.24.

The consequence of birefringence axis variation can be viewed in time-domain as mode coupling, where with “mode” it is referred to the ideal modes exchanging power [KHS12] and, thus, not propagation anymore independently, as explained in the introduction of Ch.3. However, mode coupling can be viewed also as a coupling between local eigenmodes / eigenpolarizations of consecutive segments if one considers the signals along them as the ones which get coupled, and perhaps that is the most intuitive perspective. Indeed, this way, it is clear that the signal over one eigenpolarization at the input of a certain segment is the superposition (with complex coefficients) of the differently delayed pulses over both eigenpolarizations of the previous segment. In other words, mode coupling manifests itself as a “splitting” or “bifurcation” ([KK97]) phenomenon at every interface between consecutive segments which consists in distributing the signal power of every local mode to both local modes of the successive segment. Fig.3.8 illustrates the phenomenon.

Now it is clear that it is not birefringence by itself producing mode coupling, it is its variation over the fiber length, while the signal broadening, also called polarization-mode dispersion in this case, is the result of the combination of frequency-dependent birefringence and mode coupling.

We are now ready to give a more precise definition of correlation length than the one offered in Sec.3.1. If we assume to be transmitting a pulse along a specific

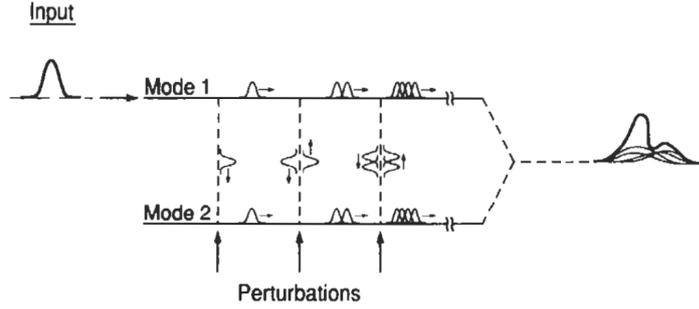


Figure 3.8: Single pulse sent over one input eigenmode of the first segment subject to pulse dispersion and mode coupling (image from [PN97, p. 125]).

polarization, let it be one of the two eigenmodes of the first fiber section, due to birefringence variation the polarization evolves with distance leaking power to the non-excited eigenmode and, after a certain length, it is no more correlated with the input polarization. After a sufficiently long distance, we can imagine the power launched in one mode has spread over both modes. Let us call  $P_{\parallel}(z)$  and  $P_{\perp}(z)$  the average (over time) power carried over the excited input polarization and the orthogonal one which has not been initially excited. The average  $\langle \cdot \rangle$  is considered among an ensemble of fibers with statistically equivalent perturbations, that is, perturbations which act on average the same on every fiber at any time and position [PN97, p. 123] [NJ05, p. 6]. This averaging is needed to be able to provide a definition which does not depend on the specific perturbation profile realization.

If we transmit a power  $P_{\parallel}(z=0) = 1$  over one polarization, the difference between the power on the two polarizations is  $\langle P_{\parallel}(z=0) \rangle - \langle P_{\perp}(z=0) \rangle = 1$ . As the distance increases, so does the coupling and  $\langle P_{\parallel}(z) \rangle - \langle P_{\perp}(z) \rangle$  reduces, because  $\langle P_{\perp}(z) \rangle$  increases. We define *correlation length* or *coupling length*  $L_C$  the distance at which the power difference  $\langle P_{\parallel}(L_C) \rangle - \langle P_{\perp}(L_C) \rangle = \frac{1}{e^2}$ , that is, the power on the orthogonal polarization is within  $\frac{1}{e^2}$  of the the power on the initially excited eigenmode [PN97, p. 123], [NJ05, p. 5]. Another equivalent definition of correlation length is derived directly from the analysis of the width of the correlation functions relative to the birefringence vector [SM05; Ant+12].

The values of correlation length are reported to vary between the order of 100 m for spooled fibers and 1 km for grounded fibers ([PN97, p. 124], [NJ05, p. 6]), while Ref.[SM05, p. 41] indicates around 100 m for grounded fibers and Ref.[Agr19, p. 13] generically claims the value to be in the order of 10 m.

When a fiber segment of length comparable to the correlation length is considered, the birefringence can be treated as approximately constant, which is referred to as *short-distance* [PN97] or *weak-coupling regime* [KHS12], and, thus, the analysis of Sec.3.2.1 applies. Otherwise, we are in the *long-distance* or *strong-coupling regime* discussed in Sec.3.3.

### 3.3.1 Frequency Independent Birefringence for a Direct-Detection System

If we consider to be transmitting a single modulated signal along a certain polarization, as in a direct detection system, then the effect frequency dependent birefringence is similar to what has been observed already for the case of uniform birefringence in Sec.3.2.4. That is, direct detection at the receiver allows to exactly recover the transmitted pulse because the only effect of non-uniform frequency-dependent birefringence is a frequency-independent phase term which disappears after taking the square modulus, as it is proven here.

Let us assume to transmit the field

$$|E(z = 0, t)\rangle = a(t) |s(z = 0)\rangle (c_1 |v_1\rangle + c_2 |v_2\rangle) \quad (3.27)$$

where  $a(t) \in \mathbb{R}$  is the transmit envelope and  $c_i = \langle v_i | s(z = 0) \rangle \in \mathbb{C}$  are the complex projection coefficients. By definition  $\langle s(z = 0) | s(z = 0) \rangle = \langle v_1 | v_1 \rangle = \langle v_2 | v_2 \rangle = 1$ , that is, the polarizations have unit power. Hence,  $\|c_1\|^2 + \|c_2\|^2 = 1$ . We assume  $|v_1\rangle, |v_2\rangle$  are the orthonormal eigenvectors of the propagation operator  $\mathbf{T}$ . In the case of non-uniform frequency-independent birefringence, they are the generalization of the eigenpolarizations used in Sec.3.2.4. Indeed, they represent polarizations which keep parallel to themselves from input to output and which do not varies with frequency (because, clearly, we are assuming frequency-independence). Then, the received signal after  $N$  sections at the fiber length  $L$  can be computed in frequency-domain from Eq.3.24-3.26 as

$$\begin{aligned} |\tilde{E}(z, \omega)\rangle &= \mathbf{T}(z, \omega) a(\omega) |s(z = 0)\rangle = e^{-j\Phi_c(z, \omega)} \mathbf{U}(z) a(\omega) |s(z = 0)\rangle = \\ &= e^{-j\Phi_c(z, \omega)} \tilde{a}(\omega) |s(z)\rangle \end{aligned} \quad (3.28)$$

Neglecting chromatic-dispersion and mode-independent higher-order effects as usual for this chapter, i.e.,  $\frac{\partial^n \Phi_c(z, \omega)}{\partial \omega^n} = 0$  for  $n \geq 2$ , the common phase term can be written as  $\Phi_c(z, \omega) = \Phi_c(z, \omega_0) + \omega \frac{\partial \Phi_c(z, \omega)}{\partial \omega}$ . Then, (3.28) transforms to time-domain as

$$|E(z, t)\rangle = e^{-j\Phi_c(z, \omega_0)} a(t - \tau_g) |s(z)\rangle \quad (3.29)$$

where  $\tau_g = \tau_c = \frac{\partial \Phi_c(z, \omega)}{\partial \omega}$  is the propagation delay which coincides with the common group delay 3.25. The received polarization can be expressed as

$$|s(z)\rangle = \mathbf{U}(z) |s(z = 0)\rangle = c_1 e^{-j\varphi_1(z)} |v_1\rangle + c_2 e^{-j\varphi_2(z)} |v_2\rangle \quad (3.30)$$

where  $e^{-j\varphi_1(z)}, e^{-j\varphi_2(z)}$  are the eigenvalues of the propagation operator  $\mathbf{U}$  corresponding to  $|v_1\rangle, |v_2\rangle$  and are frequency-independent (because  $\mathbf{U}$  is frequency-independent) phase terms (because  $\mathbf{U}$  is unitary).

Finally, performing direct-detection at the receiver, one gets:

$$\begin{aligned} \||E(z, t)\|^2 &= \langle E(z, t) | E(z, t) \rangle = \\ &= |a(t - \tau_g)|^2 e^{+j\Phi_c(z, \omega_0)} e^{-j\Phi_c(z, \omega_0)} \langle s(z) | s(z) \rangle = |a(t - \tau_g)|^2 \end{aligned} \quad (3.31)$$

where the fact that the received polarization has unit norm has been exploited thanks to the unitary transfer matrix  $\mathbf{U}$  (it can also be checked by direct calculation exploiting Eq.3.30 and the orthonormality properties of the eigenvectors). Eq.3.31 shows that after performing direct-detection, the received envelope is undistorted, as it was for the case of uniform birefringence (Sec.3.2.4).

On the opposite, if detection is performed along a specific polarization  $|s_D\rangle$  (which does not happen for a DD system, but which gives the idea of what happens in a CS), as

$$\begin{aligned} \langle s_D | E(z, t) \rangle &= a(t - \tau_g) e^{-j\Phi_c(z, \omega_0)} \langle s_D | s(z = L) \rangle = k a(t - \tau_g), \\ k &= \langle s_D | s(z = L) \rangle \end{aligned} \quad (3.32)$$

then, even taking the absolute value does not avoid to have a scaling factor  $|k| = |\langle s_D | s(z = L) \rangle|$  on the received envelope due to the misalignment between the detection polarization  $s_D$  and the received polarization.

No plots are shown in this case because they would be qualitatively the same as for the case of uniform birefringence, Fig.3.5a and Fig.3.5b.

### 3.3.2 Frequency Independent Birefringence for a Coherent-System

We consider a Polarization Multiplexed system with a coherent receiver and we transmit two quadrature modulated signal with, say, 4-QAM modulation, over two orthogonal input polarizations, say  $x$  and  $y$ , when only birefringence is considered, without PMD. That is,  $\Delta\beta(\omega) = \Delta\beta(\omega_0)$ . This approximation is assumed firstly to understand the role of pure birefringence for a CS, secondly not to have to design a more complex coherent-receiver and thirdly because sometimes PMD can be neglected in some real systems.

The effect of birefringence within a uniform segment is to rotate the two constellations. The effect of birefringence variation (change in the eigenpolarizations) is to couple the two constellations together yielding an unregular constellation made of 16 complex points. The constellation plots below clarify the situation.

Firstly, the constellations at the fiber input are presented in Fig.3.9: in blue in Cartesian coordinates, in red after the projection over the eigenpolarization coordinates of the first segment. As it can be seen, the projection yields four quartets of points being the result of combining every point of a constellation with

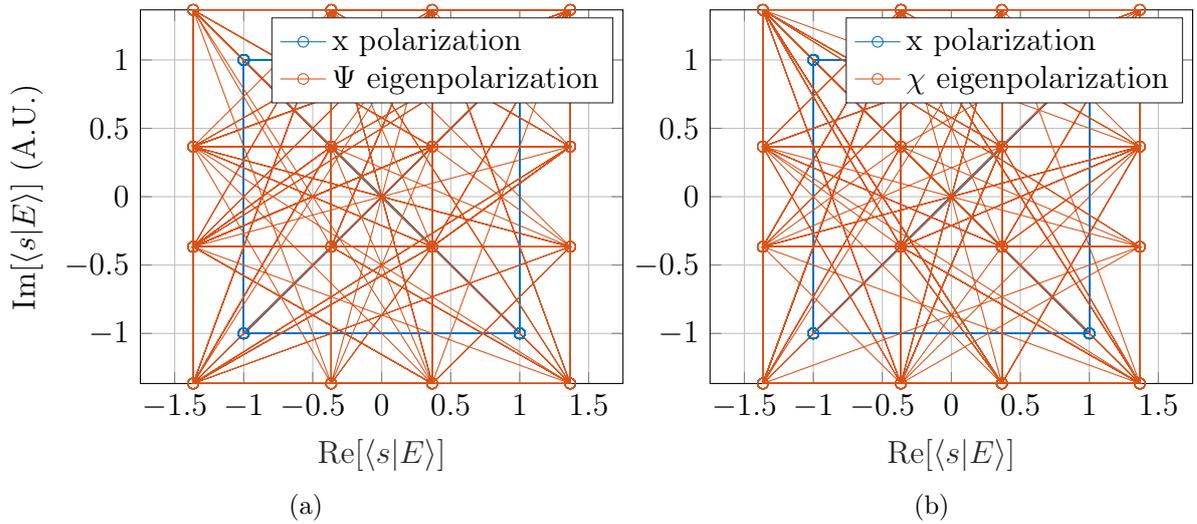


Figure 3.9: Constellations at the beginning of the first segment over the two transmit  $x$ - and  $y$ -polarizations and after the projection over the two eigenpolarizations.

every point of the other. Due to birefringence, the two constellations result to be rotated at the end of the uniform fiber segment, as Fig.3.10 shows.

Then, a change of segment, which means a projection over the new eigenpolarizations, has the effect of producing a different constellation over each eigenpolarization, but still of 16 points. (The number is not exponentially increasing as  $4^n$ , where  $n$  is the segment index, as one could suspect by seeing how the constellation evolves from 4 points to 16 points after the first projection. The reason is that the pairs of points which are combined together at any discontinuity are always the same. So 16 unique pairs in total.) Hence, the final constellations received at the fiber end over  $x$ - and  $y$ -polarizations are made of 16 points distributed over the complex plane in a non-trivial way, as visible in Fig.3.11.

A possible technique for uncoupling the two transmitted modes is to exploit the eigenvectors of the fiber transfer matrix  $\mathbf{T}$  as transmit polarizations. We are not interested in the discussion of distortion compensation techniques, but this fact allows us to describe the signal propagation as a superposition of the transmitted signals over the two eigenvector polarizations. However, this is possible only when the birefringence is frequency-independent. When frequency-dependent perturbations are considered, the eigenvectors of  $\mathbf{T}(z, \omega)$  are frequency-dependent and so no more useful. In this scenario, the role of polarizations able to propagate undistorted and independently on one another is assumed by the principal states of polarization described in the next section.

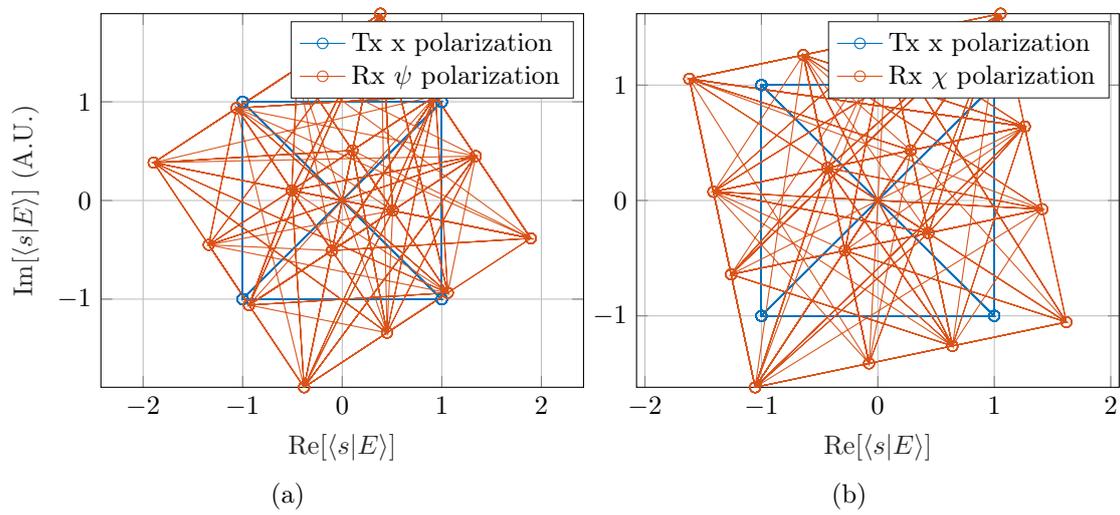


Figure 3.10: Constellations at the end of the first segment over the two eigenpolarizations.

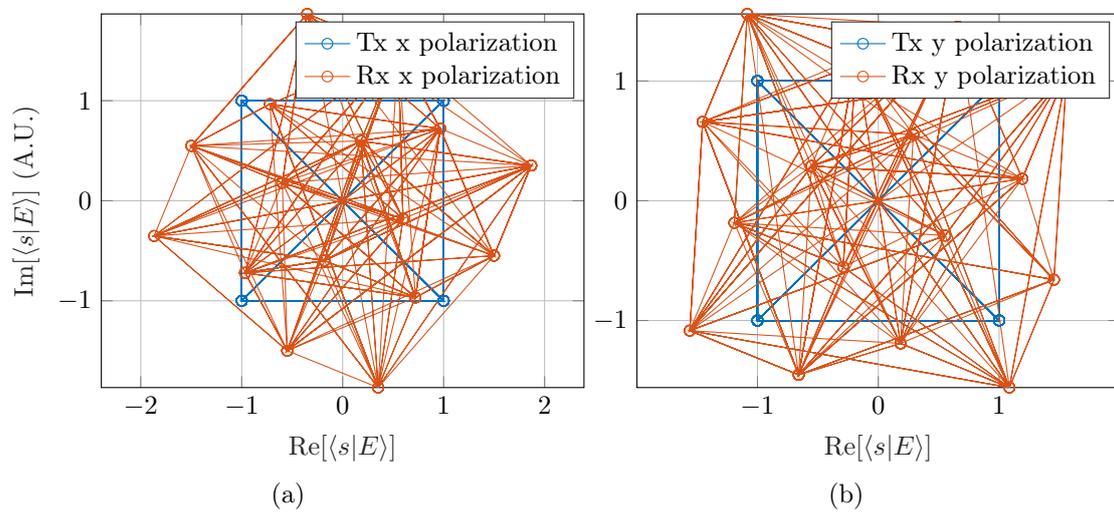


Figure 3.11: Constellations at the fiber end over  $x$ - and  $y$ -polarizations.

### 3.3.3 Principal States of Polarization

For a fiber with non-uniform birefringence modeled as a concatenation of different segments, the eigenpolarizations are no more the eigenvectors of the propagation operator  $\mathbf{U}$ , even though they are the eigenvectors of the individual propagation operators  $\mathbf{U}_i$ . Hence, a signal sent along one input eigenmode does not propagate independently on the signal transmitted on the other mode (if present) or on the delayed replica(s) present on the other eigenmode. Principal States of Polarization, discovered by Poole and Wagner in 1986 ([PW86]), solve the issue (to a first-approximation) of having a polarization along which the signal is received undistorted and are also a helpful theoretical tool in understanding the dispersion phenomenon under discussion. They are defined as the pair of orthogonal input polarizations  $\{|\hat{p}_+(z=0)\rangle, |\hat{p}_-(z=0)\rangle\}$  whose corresponding output polarizations  $\{|\hat{p}_+(z, \omega)\rangle, |\hat{p}_-(z, \omega)\rangle\}$  are frequency-independent to first order in a narrow bandwidth ([PW86]). Let us adopt the symbol  $|\hat{p}(z, \omega)\rangle$  to indicate one of the two PSPs and the shorthand notation  $|\hat{p}(z=0)\rangle \equiv |\hat{p}_0\rangle$ . Then, the definition of PSPs is

$$\frac{d|\hat{p}_D(z, \omega)\rangle}{d\omega} = 0 \quad (3.33)$$

where

$$|\hat{p}(z, \omega)\rangle := e^{-j\Phi_{PSP}(\omega)} |\hat{p}_D(z, \omega)\rangle \quad (3.34)$$

Eq.(3.34) separates the frequency-varying phase  $\Phi_{PSP}$  of a PSP from its direction  $|\hat{p}_D(z, \omega)\rangle$  ( $D$  stands for “direction”), which, thanks to (3.33), is approximately frequency independent for  $\omega \in (-\Delta\omega + \omega_0, \omega_0 + \Delta\omega)$  with  $\Delta\omega \ll \omega_0$ . Imagining to expand  $|\hat{p}_D(z, \omega)\rangle$  in a Taylor series, this last point clarifies the meaning of the statement that the PSP approximation is valid for first-order in frequency. Note that, on the opposite, the direction of a generic input polarization depends on frequency.

To prove the existence of the PSPs, we assume to be transmitting a modulated signal  $a_{IN}(\omega)$  over an input polarization coincident with a PSP:  $|\hat{s}(z, \omega)\rangle \equiv |\hat{p}_0\rangle$ . Since the input polarization and its phase can be freely chosen at the transmitter side, they are always assumed frequency-independent, regardless of the fact that they are PSPs or not. The received field can be expressed as usual as

$$\begin{aligned} |\tilde{E}(z, \omega)\rangle &= \mathbf{T}(z, \omega) |\tilde{E}(z=0, \omega)\rangle, \\ \mathbf{T}(z, \omega) &= e^{-j\Phi_c(z, \omega)} \mathbf{U}(z, \omega) \end{aligned} \quad (3.35)$$

where  $\Phi_c(z, \omega)$ , defined in Eq.3.25, is the common phase factored out from  $\mathbf{T}$ .

Then,

$$\begin{aligned} |\tilde{E}(z, \omega)\rangle &= \mathbf{T}(z, \omega) |E(z=0, \omega)\rangle \Leftrightarrow \\ a_{OUT}(\omega) |\hat{p}(z, \omega)\rangle &= e^{-j\Phi_c(z, \omega)} \mathbf{U}(z, \omega) a_{IN}(\omega) |\hat{p}_0\rangle \end{aligned} \quad (3.36)$$

where  $a_{OUT}(\omega)$  is the complex envelope at the output. Given that, as usual, we are considering unitary transfer matrices  $\mathbf{T}$  and  $\mathbf{U}$ , the transmitted and received envelopes  $a_{OUT}$  and  $a_{IN}$  are equal and so cancel out each other. Hence, from Eq.3.36 we get

$$|\hat{p}(z, \omega)\rangle = e^{-j\Phi_c(z, \omega)} \mathbf{U}(z, \omega) |\hat{p}_0\rangle \quad (3.37)$$

Differentiating Eq.3.37 with respect to  $\omega$  yields:

$$\begin{aligned} \frac{d|\hat{p}(z, \omega)\rangle}{d\omega} &= \frac{d(e^{-j\Phi_c(z, \omega)} \mathbf{U}(z, \omega) |\hat{p}_0\rangle)}{d\omega} \\ \Rightarrow \frac{d(e^{-j\Phi_{PSP}(\omega)} |\hat{p}_D(z, \omega)\rangle)}{d\omega} &= \frac{d(e^{-j\Phi_c(z, \omega)} \mathbf{U}(z, \omega) |\hat{p}_0\rangle)}{d\omega} \end{aligned} \quad (3.38)$$

Exploiting the definition (3.33) of a PSP,

$$\begin{aligned} \frac{de^{-j\Phi_{PSP}(\omega)}}{d\omega} |\hat{p}_D(z, \omega)\rangle + e^{-j\Phi_{PSP}(\omega)} \frac{d|\hat{p}_D(z, \omega)\rangle}{d\omega} &= \frac{de^{-j\Phi_c(z, \omega)}}{d\omega} \mathbf{U}(z, \omega) |\hat{p}_0\rangle + \\ + e^{-j\Phi_c(z, \omega)} \frac{d\mathbf{U}(z, \omega)}{d\omega} |\hat{p}_0\rangle \end{aligned} \quad (3.39)$$

By definition of PSPs,  $\frac{d|\hat{p}_D(z, \omega)\rangle}{d\omega} = 0$ , hence:

$$\begin{aligned} -j\Phi'_{OUT}(z, \omega) e^{-j\Phi_{PSP}(z, \omega)} |\hat{p}_D(z, \omega)\rangle &= -j\Phi'_c(z, \omega) e^{-j\Phi_c(\omega)} \mathbf{U}(z, \omega) |\hat{p}_0\rangle + \\ &+ e^{-j\Phi_c(z, \omega)} \mathbf{U}_\omega(z, \omega) |\hat{p}_0\rangle \end{aligned} \quad (3.40)$$

where  $\Phi'_{OUT}(z, \omega) \equiv \frac{d\Phi_{PSP}(z, \omega)}{d\omega}$ .

Recalling that  $e^{-j\Phi_{PSP}(z, \omega)} |\hat{p}_D(z, \omega)\rangle = |\hat{p}(z, \omega)\rangle = e^{-j\Phi_c(z, \omega)} \mathbf{U}(z, \omega) |\hat{p}_0\rangle$  we have:

$$-j\Phi'_{OUT} e^{-j\Phi_c} \mathbf{U} |\hat{p}_0\rangle = -j\Phi'_c e^{-j\Phi_c} \mathbf{U} |\hat{p}_0\rangle + e^{-j\Phi_c} \mathbf{U}_\omega |\hat{p}_0\rangle \quad (3.41)$$

where the dependencies on  $z$  and  $\omega$  have been hidden for brevity. Now, simplifying  $e^{-j\Phi_c}$  on both sides and moving the first term on the right-hand side to the left-hand side:

$$(-j\Phi'_{OUT} + j\Phi'_c) \mathbf{U} |\hat{p}_0\rangle = \mathbf{U}_\omega |\hat{p}_0\rangle \quad (3.42)$$

Setting  $k = -\Phi'_{OUT} + \Phi'_c$  and multiplying by  $\mathbf{U}^{-1} = \mathbf{U}^H$  (unitarity) on both sides, yields the final eigenvalue equation:

$$k |\hat{p}_0\rangle = -j\mathbf{U}^H \mathbf{U}_\omega |\hat{p}_0\rangle \quad (3.43)$$

Solving the eigenvalue equation 3.43 allows to determine the PSPs of the fiber, which are the eigenvectors of

$$\mathbf{G} = -j\mathbf{U}^H\mathbf{U}_\omega \quad (3.44)$$

referred to as the group delay operator. The corresponding output PSPs can be computed simply through the transfer matrix  $\mathbf{T}$ , since the following relation has been used in the proof above:

$$|\hat{p}(z, \omega)\rangle = \mathbf{T}(z, \omega) |\hat{p}_0\rangle \quad (3.45)$$

In this way we have proven that there exist in any fiber subject to non-uniform birefringence a pair of input polarizations, the PSPs, whose corresponding output states are frequency-independent to first order. That is, the PSPs are the analogous of the eigenpolarizations for the case of fiber with uniform birefringence. However, it is important to remark that PSPs and eigenpolarizations coincide only for a fiber with uniform birefringence and for the first segment of a fiber characterized by a cascade of segments. In general, PSPs are  $z$ -dependent since being the eigenvectors of  $-j\mathbf{U}^H\mathbf{U}_\omega$ , which is  $z$ -dependent itself. This is consistent with the fact that the PSPs are not the eigenvectors of the propagation operator  $\mathbf{U}$  and, thus, do not have the same polarization at both input and output. PSPs are also in general frequency dependent, but in the so-called *first-order PMD approximation*, whose validity is discussed in Sec.3.3.6, they are considered frequency-independent.

Furthermore, the PSPs at any  $z$  can be shown to form an orthonormal basis which allows to represent any polarization as a combination of them [PW86]. As it going to be clear later on, this makes the PSP a powerful tool to describe signal evolution in a linearly perturbed fiber.

The matrix  $-j\mathbf{U}^H\mathbf{U}_\omega$  is referred to as the group delay operator because its eigenvalues are related to the GDs of the PSPs. Indeed, if we identify the first-order derivative of a phase term with respect to frequency as a propagation delay (cfr. Eq.3.9 and Eq.3.25), then,  $\tau_\pm := \Phi'_{OUT\pm}$  ( $\Phi_{OUT\pm}$  defined in Eq.3.34) is interpreted as the group delay of the two output principal states. Moreover,  $k = -\Phi'_{OUT} + \Phi'_c = \tau_\pm - \tau_c$  ( $k$  defined in (3.43)) is the difference between the group delay of a pulse sent over one PSP ( $\tau_+$  for the slow PSP,  $\tau_-$  for the fast PSP) and the common delay  $\Phi'_c$  defined in Eq.3.25.

In addition, it can be observed that  $k = \pm\Delta\tau/2$  by directly computing  $k$ , as done by Poole and Wagner [PW86], which results to be:

$$k_\pm = \pm\sqrt{\|u'_1\|^2 + \|u'_2\|^2} \quad (3.46)$$

where  $u'_1, u'_2$  are the derivatives (w.r.t.  $\omega$ ) of the element of position (1, 1) and (1, 2) of  $\mathbf{U}(z, \omega)$ .

Or, following the derivation of Gordon and Kogelnik [GK00], by first proving that  $-j\mathbf{U}^H\mathbf{U}_\omega = +j\mathbf{U}_\omega\mathbf{U}^H$  (a way is rearranging some passages above, another is offered in [GK00]), which then implies that  $-j\mathbf{U}^H\mathbf{U}_\omega$  is Hermitian. Furthermore, the GD operator has zero-trace and so the last two properties (hermitianity and zero-trace) imply its eigenvalues are real and sum to zero [GK00]. Thus, they can be indicated as  $k = \pm\Delta\tau/2$  and, hence, the common group delay is also the average of the PSP GDs  $\Phi'_c = \tau_c = (\Phi'_{OUT,+} + \Phi'_{OUT,-})/2$ , as expected from Eq.3.25.

The difference between the delays  $\tau_\pm = \Phi'_{OUT,\pm}$  of the two output PSPs is referred to as *differential group delay* (DGD), i.e.,

$$\Delta\tau = \tau_+ - \tau_- \quad (3.47)$$

Hence, we just proven that the DGD can be directly computed from the eigenvectors of the GD operator as

$$\Delta\tau = k_+ - k_- = 2\sqrt{\|u'_1\|^2 + \|u'_2\|^2} \quad (3.48)$$

which provides the desired link between the GD operator and the (differential) group delay of the PSPs and justifies the name of “Group Delay” operator. Notice that, fixed the fiber length, the DGD is in general a time-varying random quantity ([Poo+88; NJ05]), since the fiber perturbations are assumed random along the fiber and time-varying. But if the perturbations were exactly known to the point of being able to compute the transfer matrix of the fiber, then the DGD could be computed as 3.48.

The group delay operator can also be introduced in a form different to the eigenvalue equation (see, e.g., [GK00; Pal14]). Such a form follows from the derivative of a generic received polarization  $|\hat{s}(z, \omega)\rangle$ , first line of Eq.3.38, which we repeat here for convenience:

$$\frac{d|\hat{s}(z, \omega)\rangle}{d\omega} = \frac{d(e^{-j\Phi_c(z, \omega)}\mathbf{U}(z, \omega)|\hat{s}(z=0)\rangle)}{d\omega}$$

Working on the right-hand-side, we can write:

$$\begin{aligned} \frac{d|\hat{s}(z, \omega)\rangle}{d\omega} &= \frac{d(e^{-j\Phi_c(z, \omega)}\mathbf{U}(z, \omega)|\hat{s}(z=0)\rangle)}{d\omega} \\ &= (-j\Phi'_c e^{-j\Phi_c}\mathbf{U} + e^{-j\Phi_c}\mathbf{U}_\omega)|\hat{s}_0\rangle \\ &= (-j\Phi'_c e^{-j\Phi_c} + e^{-j\Phi_c}\mathbf{U}_\omega\mathbf{U}^{-1})\mathbf{U}|\hat{s}_0\rangle \\ &= (-j\Phi'_c + \mathbf{U}_\omega\mathbf{U}^{-1})e^{-j\Phi_c}\mathbf{U}|\hat{s}_0\rangle \\ &= -j(\Phi'_c + j\mathbf{U}_\omega\mathbf{U}^{-1})\mathbf{T}|\hat{s}_0\rangle \\ &= -j(\Phi'_c + j\mathbf{U}_\omega\mathbf{U}^{-1})|\hat{s}(z, \omega)\rangle \end{aligned} \quad (3.49)$$

with  $|\hat{s}_0\rangle \equiv |\hat{s}(z=0)\rangle$ . Thus, we can conclude that the derivative of a generic Jones polarization vector  $|\hat{s}(z, \omega)\rangle$  is

$$\frac{d|\hat{s}(z, \omega)\rangle}{d\omega} = -j\mathbf{G}'|\hat{s}(z, \omega)\rangle \quad (3.50)$$

where

$$\mathbf{G}' = (\Phi'_c + j\mathbf{U}_\omega\mathbf{U}^{-1}) \quad (3.51)$$

is the group-delay operator. At first it might seem that  $\mathbf{G}'$  is different from the group delay operator defined in Eq.3.44 due to the presence of  $\Phi'_c$  and due to  $j\mathbf{U}_\omega\mathbf{U}^{-1}$  instead of  $-j\mathbf{U}^H\mathbf{U}_\omega$ . The former is a scalar, so it does not modify the direction of the output vector of  $j\mathbf{U}_\omega\mathbf{U}^{-1}$ , whatever the input, in particular, in the relevant case of the eigenvectors. However, it does modify the amplitude.  $j\mathbf{U}_\omega\mathbf{U}^{-1}$  is actually identical to  $-j\mathbf{U}^H\mathbf{U}_\omega$  thanks to the hermitianity of the matrix, as explained above. So, the eigenvectors of  $\mathbf{G}'$  are parallel to the eigenvectors of  $\mathbf{G}$  while the eigenvalues of  $\mathbf{G}'$  are bigger than the eigenvalues of  $\mathbf{G}$  by the average delay  $\tau_c = \Phi'_c$ , which means that the eigenvalues of  $\mathbf{G}'$  represent the absolute delays of the two PSPs (while the eigenvalues of  $\mathbf{G}$  represent the delay deviations from the mean value).

From another perspective, if the output polarization is one of the two PSPs,  $|\hat{s}(z, \omega)\rangle \equiv |\hat{p}_\pm(z, \omega)\rangle$ , we expect  $\frac{d|\hat{p}(z, \omega)\rangle}{d\omega}$  to be parallel to  $|\hat{p}(z, \omega)\rangle$  in order to satisfy the definition of PSP as frequency-independent to first order. This means that the PSP  $\frac{d|\hat{p}(z, \omega)\rangle}{d\omega}$  has to be an eigenvector of  $\mathbf{G}'$ , which is identical to be an eigenvector of  $j\mathbf{U}_\omega\mathbf{U}^{-1}$ .

The orthonormality of the Principal States, together with the fact that a pulse sent along one of them does not suffer from pulse broadening to first order, allows to write the field transmitted along a generic input polarization:

$$|\hat{s}(z=0)\rangle = c_+|\hat{p}_+(z=0)\rangle + c_-|\hat{p}_-(z=0)\rangle$$

as a superposition of the replicas over the two PSPs (neglecting the second- and higher-order terms on the PSP output phase, so that  $\Phi_{PSP}(\omega) = \Phi_{PSP}(\omega_0) + \omega\Phi'_{PSP}(\omega_0)$ ), i.e.,

$$|E(z, t)\rangle = c_+e^{-j\Phi_{PSP+}(z, \omega_0)}|E(z=0, t-\tau_+)\rangle|\hat{p}_{D+}(z)\rangle + c_-e^{-j\Phi_{PSP-}(z, \omega_0)}|E(z=0, t-\tau_-)\rangle|\hat{p}_{D-}(z)\rangle \quad (3.52)$$

where  $c_\pm$  are the complex weighting/projection coefficients and Eq.3.34 has been exploited. Eq. 3.52 also implies that the PSPs exhibit maximum and minimum propagation delays among all possible transmit polarizations.

### 3.3.4 Nonuniform Frequency Dependent Birefringence

An example of signal transmission through a fiber with nonuniform frequency-varying birefringence is shown in Fig.3.12a, where a Gaussian pulse has been transmitted over a 100 km fiber with segment length  $L_S = 1$  km. The envelope detected along the same polarization employed at the transmitter side, that is, linear  $45^\circ$ , is qualitatively a superposition of the envelopes over the two PSPs, yet it is not just a summation between the envelopes due to the phase mismatch related to the terms  $c_\pm$  of Eq.3.52. The envelope obtained by direct-detection (which however can no more be considered to be detected along the “received polarization” because, due to frequency-dependent birefringence, polarization is frequency dependent and so there is no more a “received polarization”, conversely to the case frequency-independent birefringence discussed in Sec.3.2.4) is shown as well and it results to be less distorted.

Finally, Fig.3.12b plots the same case as Fig.3.12a, except that the birefringence is set uniform for the whole fiber length. The comparison of the two figures confirms that the presence of non-uniform birefringence, and thus of mode coupling, is beneficial in reducing the DGD between the two PSPs, as was anticipated in Sec.3.2.1. It can be shown that the DGD reduces by reducing the segment length (fixed the fiber length) or increasing the birefringence vector variance ([Shemirani:2009; Jua+14]) or changing the statistics of the birefringence stochastic process in order to, loosely speaking, increase the birefringence “randomization” (e.g., choosing the eigenpolarizations to be uniformly distribution on the Poincaré sphere rather than following a random walk process as we did).

From another perspective, the mixing of pulses with different amplitudes and delays at every interface between consecutive sections intuitively suggests that the broadening of the signal is lower than it would be for a fiber with the same length but uniform birefringence.

What has been stated about Fig.3.12b and Fig.3.12a is consistent with the theoretical knowledge that the mean DGD increases linearly with distance when the birefringence is uniform. Under this circumstance, the DGD can be computed with (3.8). On the opposite, the mean DGD increases with the square-root of distance when the fiber is nonuniform.

The effectiveness of the PSPs as polarizations along which the signal propagation is undistorted to first-order can be qualitatively appreciated in the examples below. We have considered a fiber of length  $L = 100$  km, with sections of length  $L_S = 1$  km and mean refractive index difference between the two eigenpolarizations  $\Delta n = 10^{-6}$ ; the symbol rate is  $R_S = 1$  GBaud. These parameters are enough to consider the polarization dispersion negligible, as can be observed by graphically inspecting the received envelope corresponding to a single transmitted Gaussian pulse over a linear  $45^\circ$  polarization, Fig.3.13.

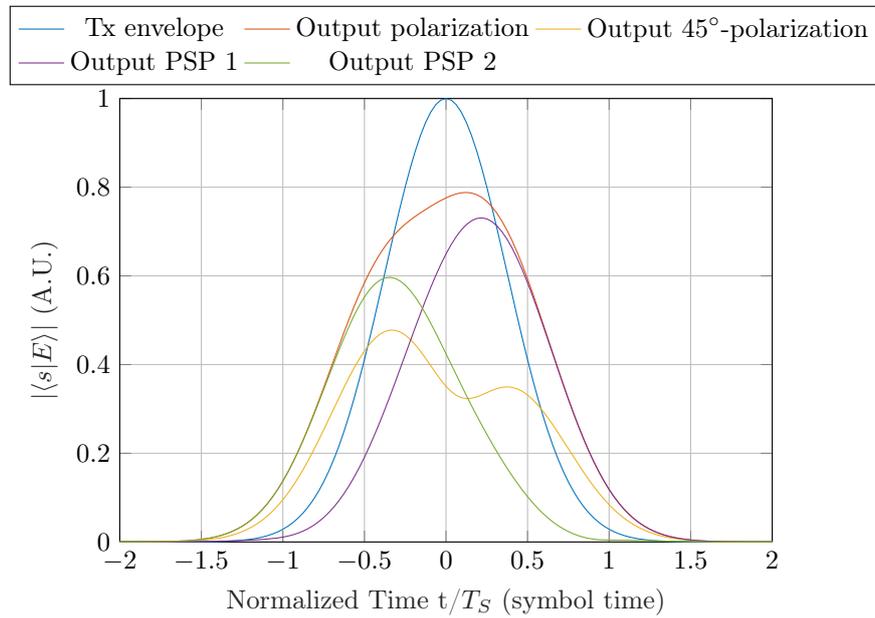
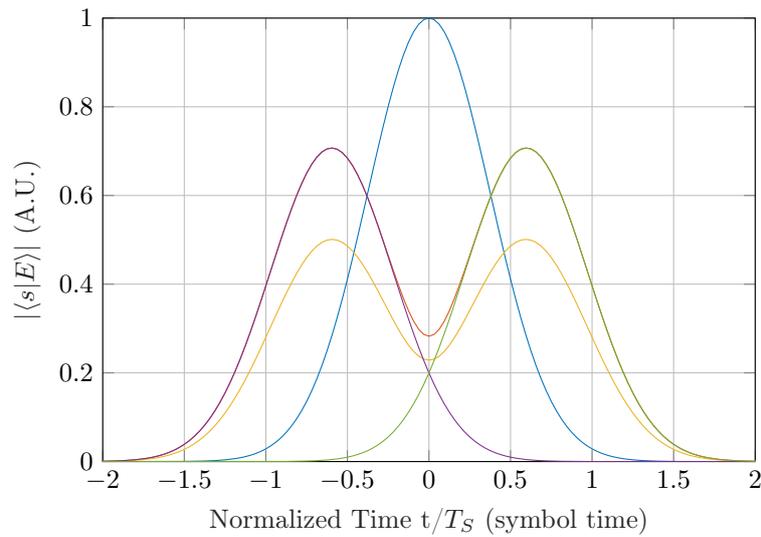
(a)  $L_S = 1\text{km} \ll L$ (b)  $L_S > L$ 

Figure 3.12: Envelope distortion comparison between the case of relatively high mode coupling (3.12a) and no mode coupling (3.12b) for a fiber with average refractive index difference  $\Delta n = 3.5 \cdot 10^{-6}$  and length  $L = 100\text{ km}$ ; symbol rate  $R_S = 1\text{ GBaud}$ .

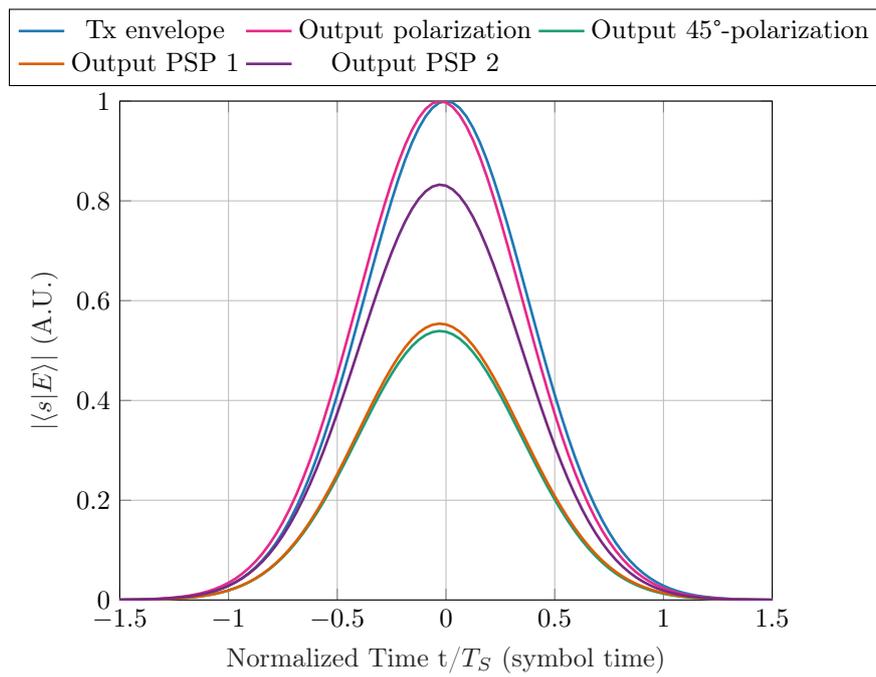


Figure 3.13: Single transmitted Gaussian pulse over a fiber of length  $L = 100$  km, with sections of length  $L_S = 1$  km and mean refractive index difference between the two eigenpolarizations  $\Delta n = 10^{-6}$ . Observe the low DGD.

However, if one considers a PM multiplexed system with unipolar PAM and transmits two different signals along two orthogonal polarizations  $x$  and  $y$ , as depicted in Fig.3.14, then the received envelopes over  $x$ - and  $y$ -polarizations show significant distortion due to mode coupling. In particular, because of the presence of a pulse on the other mode, pulses appear in correspondence to symbol times where a '0' has been transmitted. Moreover, when two pulses are contemporary present on the two modes, some energy exchange is present producing overshoots and undershoots, see Fig.3.14. Detection over the 2 PSPs does not help alone because the transmission happened along  $x$  and  $y$  polarizations, not along the input PSP. Hence, over the two input PSPs a combination of both signals is present and so at the output PSPs show a combination of them as well.

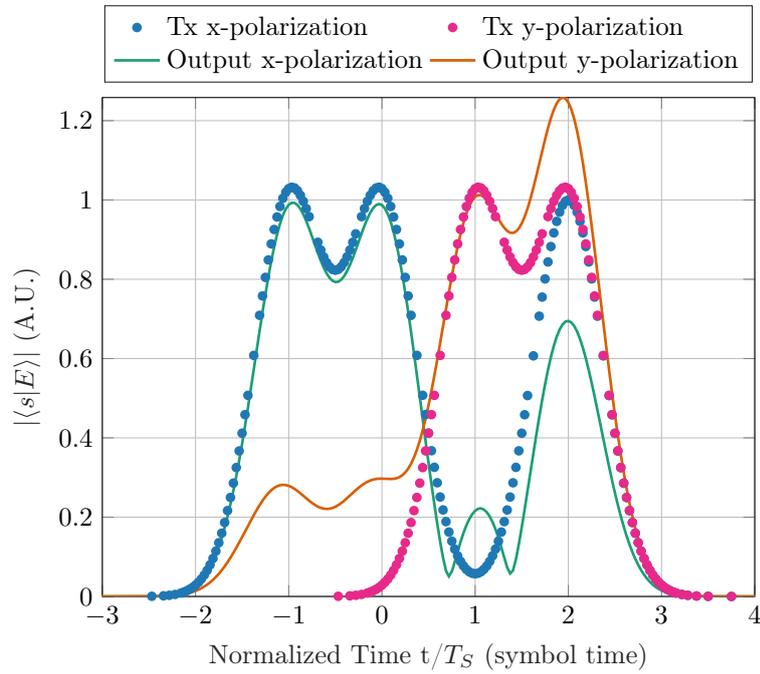


Figure 3.14: PM system transmitting and receiving two independently modulated signals at symbol rate  $R_S = 1$  GBaud over the two orthogonal  $x$ - and  $y$ - polarizations. Fiber of length  $L = 100$  km, with sections of length  $L_S = 1$  km and mean refractive index difference between the two eigenpolarizations  $\Delta n = 10^{-6}$ .

However, if the channel input PSPs are estimated in advance, chosen as pair of transmit polarizations and detection performed along the output PSPs, then the received signals show significantly less distortion, as visible in Fig.3.15. (Notice that in case frequency-independent birefringence, as seen in Sec.3.3.2, the same analysis could be done exploiting the eigenvectors of  $\mathbf{T}$  instead of the PSPs, obtaining perfect compensation.)

This qualitative graphical analysis shows the effectiveness of PSPs as to limit signal distortion due to frequency-varying birefringence and mode coupling, being the PSPs independent on one another and not suffering from polarization dispersion to first order. Clearly, two issues related to this method are the availability of a channel estimate at the transmitter side and the rate at which the PSPs have to be estimated again because of variations in the channel response. However, assessing the feasibility of PSPs for distortion compensation is not the focus of the present work. It is worth remarking that when the first-order approximation is not fulfilled, that is, when the signal band is too large (strictly speaking, larger than the PSP band, see Sec.3.3.6), the PSPs are not much effective since higher-order effects become predominant.

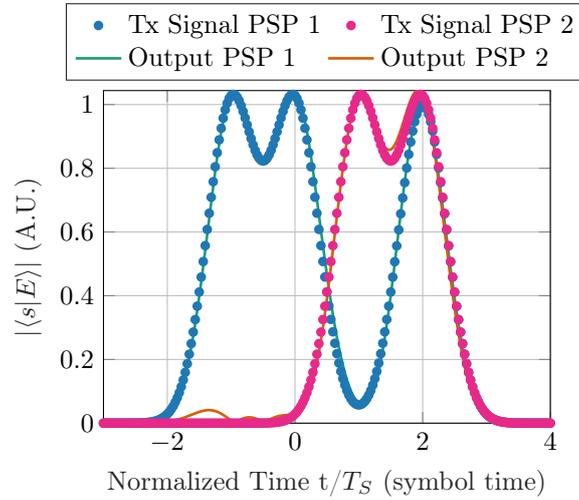


Figure 3.15: PM system transmitting and receiving two independently modulated signals at symbol rate  $R_S = 1$  GBaud over the two orthogonal PSPs. Fiber of length  $L = 100$  km, with sections of length  $L_S = 1$  km and mean refractive index difference between the two eigenpolarizations  $\Delta n = 10^{-6}$ .

### 3.3.5 PMD Vector

The Polarization-Mode Dispersion (PMD) vector is a non-unit Stokes vector defined as

$$\boldsymbol{\tau}(z, \omega) := \Delta\tau(z, \omega)\hat{\boldsymbol{p}}_+(z, \omega) \quad (3.53)$$

where  $|\hat{\boldsymbol{p}}_-(z, \omega)\rangle$  is the slow PSP for the position  $z$  of interest. Notice that the PMD vector is, in general, frequency-dependent, because both magnitude (DGD) and direction (slow PSP) can vary with frequency. It is considered frequency-independent when the first-order PMD is assumed. The PMD vector can be

expanded as [SM05; NJ05]:

$$\boldsymbol{\tau}(\omega_0 + \omega) = \boldsymbol{\tau}_0 + \omega \boldsymbol{\tau}_1 + \frac{1}{2} \omega^2 \boldsymbol{\tau}_2 + \dots \quad (3.54)$$

where  $\boldsymbol{\tau}_0 \equiv \boldsymbol{\tau}(\omega = \omega_0)$  is the term referred to as first-order PMD vector, while  $n$ -th order derivative  $\boldsymbol{\tau}_n \equiv \frac{d^n \boldsymbol{\tau}}{d\omega^n}$  is referred to as  $(n + 1)$ -th order PMD vector [SM05; NJ05; PN97].

The PMD vector is useful because it allows to draw a link between the time-domain description of the polarization dispersion phenomenon portrayed in the previous section and the frequency domain one, which we present now. In Sec.3.2.5 it has been observed that, for the case of uniform birefringence, the field polarization rotates with frequency about the principal axes (i.e., the eigenpolarizations) on the Poincaré sphere, in case the transmit polarization is different from the eigenpolarizations. The theoretical and experimentally verified existence of PSPs (e.g., [Poo+88]) allows to extend that description to the case of fibers with non-uniform birefringence. Indeed, starting from the expression for the derivative of a polarization vector, Eq.3.50, it is possible to prove (see [GK00] or, in a different way, [FP91]) that the field output polarization evolves with frequency (fixed a position  $z$ ) as [Poo+88; PN97; NJ05; SM05; GK00]

$$\frac{d\hat{\mathbf{s}}}{d\omega} = \boldsymbol{\tau} \times \hat{\mathbf{s}} \quad (3.55)$$

Notice that in the first-order PMD approximation, the PMD vector has fixed direction  $\boldsymbol{\tau}_0$  and so the previous equation describes a precession of the polarization state about the first-order PMD vector, that is, about the PSP. In other words, Eq.3.55 is the law of infinitesimal rotation for the polarization in the frequency domain. Observe the duality with the law of infinitesimal rotation for the position-domain, Eq.3.18.

If one considers the higher-order terms of the PMD vector, then the evolution is more complex and it is no more exactly a precession about a fixed axis. For example, if the PMD vector is composed of the first two terms of the expansion 3.54, the exact expression of the polarization state derivative becomes (applying Eq.3.55):

$$\frac{d\hat{\mathbf{s}}}{d\omega} = \boldsymbol{\tau}_0 \times \hat{\mathbf{s}} + \omega \boldsymbol{\tau}_1 \times \hat{\mathbf{s}} \quad (3.56)$$

However, one can always narrow the frequency range in order for the first-order approximation to be accurate enough and describe the polarization evolution as a precession about the PSP at a certain fixed frequency  $\omega = \omega_0$ .

Eq.3.55 tells that the higher the DGD  $|\boldsymbol{\tau}| = \Delta\tau$  (rate of rotation of the polarization  $\hat{\mathbf{s}}$  on the Poincaré sphere) and the misalignment between the polarization

$\hat{\mathbf{s}}$  and the PSP direction  $\hat{\boldsymbol{\tau}}$ , the higher the frequency-variation of the polarization vector. If, e.g., the transmitted polarization coincides with one PSP, then the cross-product  $\boldsymbol{\tau} \times \hat{\mathbf{s}}$  goes to zero and no polarization variation with frequency happens. On the opposite, if  $\hat{\mathbf{s}}$  is equally split between the two PSPs, there is maximum angular excursion for  $\hat{\mathbf{s}}$  on the Poincaré sphere. This is consistent with the time-domain description of Eq.3.52 where the higher the DGD and the decomposition among the two PSPs, the higher the signal broadening.

For completeness, we report another insightful relation which can be assumed as an alternative definition of PSPs ([NJ05]) or derived (see [GK00]):

$$\tau_g = \tau_c + \frac{1}{2} \boldsymbol{\tau} \cdot \mathbf{s}(z=0) \quad (3.57)$$

where  $\tau_g$  is the group delay of the received signal  $|\tilde{E}(z, \omega)\rangle$  of Eq.3.52 and  $\tau_c$  is the common group delay defined in Eq.3.25. Eq.3.57 basically expresses the signal delay  $\tau_g$  as an average between the two PSPs delays, weighted by the coefficients of the decomposition of  $\mathbf{s}$  along the two PSPs.

Notice that  $\tau_g$  is not a propagation delay associated to a replica of the transmitted signal, that is, it is not possible to write something like  $|E(z=0, t - \tau_g)\rangle$ , because the signal is received distorted and so it cannot be just a replica of the transmitted one. Then, it is the propagation delay of the distorted envelope. However, if the signal polarization  $\mathbf{s}(z=0)$  is parallel to one of the input PSPs, then Eq.3.57 provides  $\tau_g \in \{\tau_+, \tau_-\}$  (because two orthogonal Jones vectors, like the eigenpolarizations or the PSPs, are antiparallel in Stokes space), consistently with the fact that pulses transmitted along the PSPs have maximum and minimum delay, as already observed from Eq.3.52.

### Higher-Order PMD

We will not discuss about the PMD terms in Eq.3.54 of order greater than the first in detail. It is enough to mention that their presence describes a frequency-variation of the PMD vector  $\boldsymbol{\tau}$ . Concerning the second-order PMD, the dependence of  $\Delta\tau$  on the frequency gives rise to *polarization-dependent chromatic dispersion* (PCD) which broadens or compresses the received pulses [PN97; NJ05]. The dependence of the PMD vector direction  $\hat{\boldsymbol{\tau}} = \hat{\mathbf{p}}_-(z, \omega)$  on frequency accounts for the *PSP depolarization* (i.e., rotation) with frequency, which yields envelope distortions like overshoots [NJ05]. An example of higher-order effects is given in Fig.3.16a for the time-domain, in Fig.3.16b for the PSP evolution over the Poincaré sphere in frequency-domain. In this case, the combination of birefringence, signal band and distance produces a significant distortion of the received envelope such that we are no more in the first-order PMD assumption. In particular, notice that the received signals appear to be a train of pulses with different delays, which

is consistent with the bifurcation idea of polarization mode dispersion and mode coupling. Nevertheless, the received envelope can be recognized to be a sort of superposition of the two replicas over the PSPs.

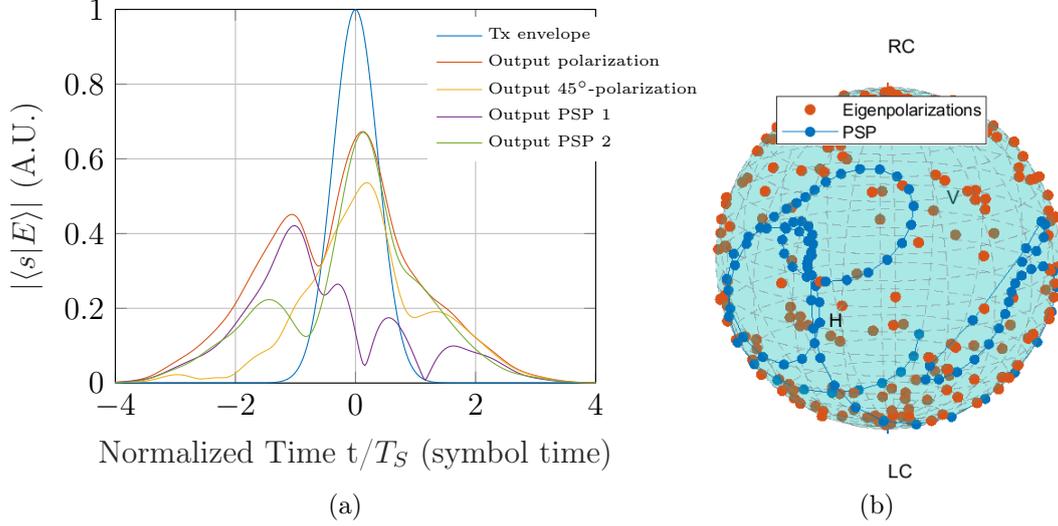


Figure 3.16: Higher-order PMD effects on the received envelope (Fig.3.16a) and PSP direction in frequency (Fig.3.16b) at a symbol rate  $R_S = 1$  GBaud for a fiber with length  $L = 100$  km, section length  $L_S = 1$  km and mean refractive index difference  $\Delta n = 2 \cdot 10^{-5}$

### 3.3.6 Statistics of PMD

It has been mentioned in Sec.3.3 that the birefringence vector  $\beta$  is modeled as a stochastic process owing to the fact that any fiber suffers a different profile of perturbations and it changes rapidly in a single fiber owing to the change of environmental conditions (e.g., vibrations, temperature). This implies that the PMD vector  $\tau$  is a stochastic processes as well, whose statistics have been extensively studied in the literature [PWN91; FP91; SM05; NJ05; PN97]. A core equation for the statistical analysis is the universal or dynamical PMD equation [PWN91; FP91; PN97; GK00; NJ05; SM05] which directly connects the PMD and birefringence Stokes vectors as

$$\frac{\partial \tau(z, \omega)}{\partial z} = \frac{\partial \beta(z, \omega)}{\partial \omega} + \beta(z, \omega) \times \tau(z, \omega) \quad (3.58)$$

The derivation of 3.58 starts by differentiating the equation for the birefringence vector Eq.3.18 w.r.t.  $\omega$  and the equation for the dispersion vector Eq.3.55 w.r.t.  $z$  and then requires little algebraic manipulation [PWN91].

Usually, the way the dynamical equation is used consists in fixing the statistics of the birefringence stochastic process and retrieving the statistics of the PMD stochastic process, either by numerical solution [PWN91] or through analytic methods [FP91].

An interesting result is that the DGD of the overall fiber is Maxwellian distributed [PWN91; FP91; PN97; NJ05; SM05]. Indeed, if the fiber birefringence is modeled as a random-walk process, the PMD vector  $\boldsymbol{\tau}$  turns out to converge to a Gaussian vector (i.e., a vector with i.i.d. Gaussian components [PWN91]), given that the PMD vector can be seen to be in some sense the sum of Gaussian vectors, the PMD vectors of the single segments [FP91; GK00]. Hence, the DGD, being the modulus of a Gaussian vector, is Maxwellian distributed. Being the DGD Maxwellian distributed, it can assume particularly high values, even though with low probability. This last aspect has two consequences. The first is that it is not possible to adopt a worst-case approach in designing systems which are PMD-tolerant, but a statistical approach considering the outage probability for a certain threshold of PMD, i.e., the probability that the DGD is beyond a certain value considered the boundary for the fault-rejection capabilities of the system, has to be employed. The second implication is that extreme events are rare to happen in simulation if every time a random realization of fiber channel is considered. Then, a possible solution is to simulate beforehand a large number of realizations for a specific fiber and store the worst case (or the worst-acceptable case), together with other realizations of interest (e.g., corresponding to an average value of DGD).

Fig.3.17b presents the DGD histogram obtained through simulating  $10^5$  fiber realizations of 1000 fiber segments each of length 100 m and  $\Delta n = 10^{-6}$  (the eigenpolarizations of a fiber realization are depicted in the Poincaré sphere in Fig.3.17a) compared against the theoretical best-fitting Maxwellian curve. The two curves are close to each other, confirming the theoretical expectations.

Other two parameters of great importance in the statistical analysis of PMD are the mean value of the DGD  $\langle \Delta\tau \rangle$  and the mean square value of the DGD  $\langle \Delta\tau^2 \rangle$ , where, according to the definition, the averaging is performed over an ensemble of statistically equivalent fibers and for a fixed frequency [NJ05, p. 12]. However, the ensemble averages are hard to compute in practice due to the lack of an ensemble of fibers. Hence, it has been shown [PWN91; NJ05] (through comparison of simulation and experimental results) that the average can be computed in the frequency domain  $\Delta\tau(\omega)$  from a single fiber realization [NJ05, p. 17].

Indeed, as already mentioned, the DGD is frequency-dependent. This computation is most common in experimental analysis where usually only one or few fibers are available. It is even possible to compute the average from fiber realizations at different time instants of the same physical fiber, because of the time-varying property of the stochastic fiber channel [NJ05, p. 25]. Yet, the time scale of such

channel variations can be particularly long depending on the situation, hence the method might be practically unfeasible [NJ05, p. 17]. Finally, an analytic expression for  $\langle \Delta\tau^2 \rangle$  is [NJ05]:

$$\langle \Delta\tau^2 \rangle = 2(\Delta\tau_b L_C / L_B)^2 (L / L_C + e^{-L/L_C} - 1) \quad (3.59)$$

where  $L_C$  is the correlation length and  $L_B$  is the beat length.

We tried to estimate  $\overline{\Delta\tau}$  through  $\sqrt{\langle \Delta\tau^2 \rangle}$ , which is not based on any rigorous derivation, but it is just a reasonable estimation, and through spectral averaging as a way of applying the definition and we obtained close values. Poole et. al. observed in his simulations [PWN91] that the discrepancy between the root-mean square value  $\sqrt{\langle \Delta\tau^2 \rangle}$  and  $\overline{\Delta\tau}$  is at most of the 8%, confirming the sufficient quality of the approximation for our interests.

Another relevant concept which emerges from the study of the PMD statistics is the *PSP bandwidth* [NJ05] which is defined as the signal bandwidth  $\Delta f_{PSP}$  over with the first-order PMD approximation holds. Hence, values of DGD separated by a frequency spacing  $\Delta f > 6\Delta f_{PSP}$  are uncorrelated. A practical formula to compute  $\Delta f_{PSP}$  is the following [NJ05, p. 12]:

$$\Delta f_{PSP} = 125\text{GHz} / \overline{\Delta\tau} \quad (3.60)$$

The PSP band can be equivalently estimated as the bandwidth of the PMD correlation function [SM05; NJ05]:

$$\langle \tau(f)\tau(f + \Delta f) \rangle$$

This is reasonable since the correlation function informs on the strength of the dependence of two values of a certain function,  $\tau(f)$  here, at a certain distance,  $\Delta f$  in this case.

### 3.4 Harmonization

In the previous sections we have shown that it is possible to model and simulate the fiber channel through the use of the transmission matrix  $\mathbf{T}$  generated as concatenation of segment matrices  $\mathbf{T}_i = \mathbf{C}^H \mathbf{T}' \mathbf{C}$ , where  $\mathbf{T}'$  is related to the ordinary and extraordinary propagation constants by  $\mathbf{T}' = \text{diag}(e^{-j\beta_y z}, e^{-j\beta_x z})$  and  $\mathbf{C}$  contains the eigenpolarization vectors (see Eq.3.12 and Eq.3.24). We have introduced  $\mathbf{T}$  through reasonable and experimental-based considerations about birefringence in a fiber, but we did not show its consistency with the Helmholtz wave equation (2.34) presented in Ch.2, which is fundamental since deriving from the Maxwell's equations. Furthermore, in the previous sections we have presented the local birefringence vector  $\boldsymbol{\beta}(z, \omega)$  which describes the local properties of the

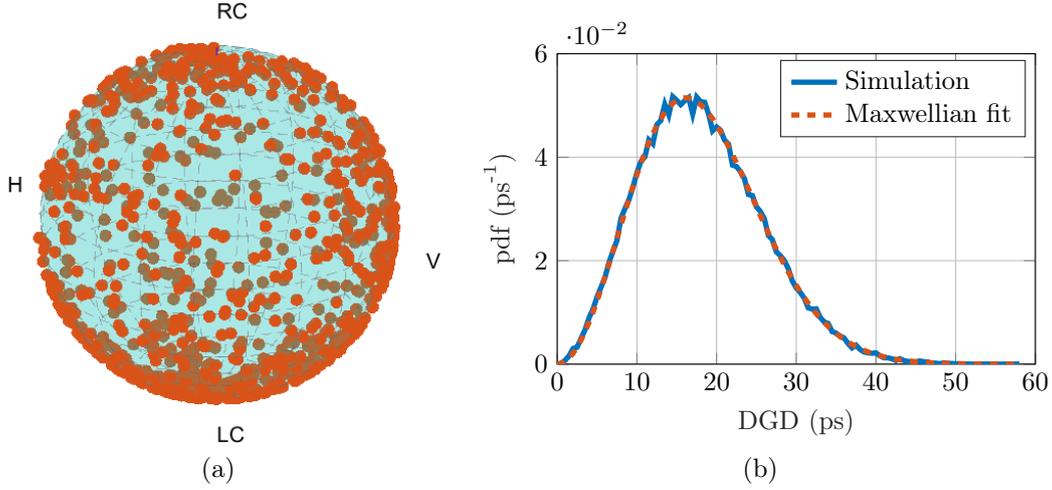


Figure 3.17: Simulation of  $10^5$  fiber realizations of  $N = 1000$  sections each of length  $L_S = 100$  m, total fiber length  $L = 100$  km, with non-uniform birefringence and mean refractive index difference  $\Delta n = 10^{-6}$ . The statistical properties of the generated birefringence vector are described in Sec.3.3. In 3.17a a realization of the eigenpolarizations is shown; in 3.17b the Probability Density Function retrieved through the simulation is compared against the Maxwellian fitting curve.

birefringence which vary along fiber length. Such a vector describes also the evolution with  $z$  of a generic input polarization state  $\mathbf{s}$  through Eq.3.18.

Aim of the present section is to show the connections among  $\mathbf{T}$  proposed in Eq.3.12 and in Eq.3.24, the birefringence vector  $\boldsymbol{\beta}$  (to understand in which way the components of  $\boldsymbol{\beta}$  are related to the propagation constants used in  $\mathbf{T}'$  of Eq.3.12) and the LSE of Ch.2.

Let us start with rewriting the Helmholtz equation (5.35) with the Jones vector notation, as

$$\frac{\partial^2 |\tilde{E}(z, \omega)\rangle}{\partial z^2} + \varepsilon_r(x, y, z) k_0^2 |\tilde{E}(z, \omega)\rangle = 0 \quad (3.61)$$

where we recall that  $\varepsilon_r(x, y, z)$  is the relative dielectric permittivity tensor.

In truth, the Helmholtz equation (2.34) of Ch.2 was written for the ideal fiber, in which case  $\varepsilon_r$  is the ideal dielectric tensor and there is no birefringence, and for a single frequency, while here we are considering a modulated field. The Helmholtz equation for the case of linear perturbations and a generic transmitted signal is (5.35) derived in Ch.5. Nonetheless, concerning the linear part, it is similar to (2.34), with the differences that in (5.35)  $\varepsilon_r$  is the actual perturbed dielectric and not the ideal one, and instead of phasor there is the transmitted signal.

We factor out the contribution of the common phase  $e^{-j\beta_c(\omega)z}$  ( $\beta_c(\omega)$  defined in (3.9)), which is a fast-varying phase, as already done several times (e.g., in (3.13)),

as

$$|\tilde{E}(z, \omega)\rangle = e^{-j\beta_c z} |s(z, \omega)\rangle \quad (3.62)$$

where  $|s\rangle$  still contains information on the complex envelope like  $|\tilde{E}\rangle$ , and a slow-varying phase. We indeed assume that  $|s\rangle$  and  $\varepsilon_r$  are slow-varying with  $z$  (adiabatic approximation [GK00]), which in practice means that the second-order derivatives of these quantities with respect to  $z$  are negligible.

In addition, we express the dielectric tensor  $\varepsilon_r$  as a function of the local birefringence Stokes vector  $\boldsymbol{\beta}$ , like [GK00, Sec.6]

$$\boldsymbol{\varepsilon} k_0^2 = \beta_c^2 \mathbf{I} + \beta_c \boldsymbol{\beta} \cdot \boldsymbol{\sigma} \quad (3.63)$$

where  $\boldsymbol{\sigma}$  is the vector collecting the three Pauli spin matrices and the product has to be intended as [GK00]

$$\boldsymbol{\beta} \cdot \boldsymbol{\sigma} = \beta_1 \boldsymbol{\sigma}_1 + \beta_2 \boldsymbol{\sigma}_2 + \beta_3 \boldsymbol{\sigma}_3 \quad (3.64)$$

where  $\boldsymbol{\sigma}_i$  are the Pauli spin matrices and  $\beta_i$  are the three components of  $\boldsymbol{\beta}$ .

Equation 3.63 expresses the dielectric tensor as a sum between a common component represented by the common propagation constant  $\beta_c$  and the deviations of the components due to birefringence.

Inserting 3.62 and 3.63 in 3.61 and exploiting adiabaticity we retrieve the following propagation equation, that is the linear Schrödinger equation (LSE) written with a mixture of Jones and Stokes vector formalism [GK00, Sec.6]

$$\frac{\partial |s(z, \omega)\rangle}{\partial z} + \frac{1}{2} j \boldsymbol{\beta} \cdot \boldsymbol{\sigma} |s(z, \omega)\rangle = 0 \quad (3.65)$$

Let us now consider a piece of fiber with uniform birefringence so that  $\varepsilon_r(x, y, z) = \varepsilon_r(x, y)$  and  $\boldsymbol{\beta}(z) = \boldsymbol{\beta}$ . The extension of the reasoning to a series of differently perturbed sections is treated afterwards. Thanks to the uniformity assumption, Eq.3.65 becomes a first-order linear differential equation with constant coefficients and, thus, admits the exponential solution

$$|s(z, \omega)\rangle = e^{-j\frac{1}{2}\boldsymbol{\beta}(\omega) \cdot \boldsymbol{\sigma} z} |s(z=0, \omega)\rangle = e^{-j\frac{1}{2}\Delta\boldsymbol{\beta}(\omega) \hat{\boldsymbol{\beta}} \cdot \boldsymbol{\sigma} z} |s(z=0, \omega)\rangle \quad (3.66)$$

where the reason for the factorization of  $\boldsymbol{\beta}(\omega) = \Delta\boldsymbol{\beta}(\omega) \hat{\boldsymbol{\beta}}$  will be clear soon. Notice that  $-\hat{\boldsymbol{\beta}} \cdot \boldsymbol{\sigma}$  and  $e^{-\frac{1}{2}j\Delta\boldsymbol{\beta} \hat{\boldsymbol{\beta}} \cdot \boldsymbol{\sigma} z}$  are  $2 \times 2$  Jones matrices.

The matrix exponential is defined through a matrix power series, but in the case of a diagonalizable matrix, the expression is simpler. Let us first diagonalize  $\hat{\boldsymbol{\beta}} \cdot \boldsymbol{\sigma}$  as

$$\boldsymbol{\beta} \cdot \boldsymbol{\sigma} = \mathbf{C}^{-1} \boldsymbol{\Lambda} \mathbf{C} \quad (3.67)$$

where  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2)$  is a diagonal matrix whose diagonal elements  $\lambda_1, \lambda_2$  are the eigenvalues of  $\hat{\boldsymbol{\beta}} \cdot \boldsymbol{\sigma}$  and  $\mathbf{C} = \begin{bmatrix} \langle v_1 | \\ \langle v_2 | \end{bmatrix}$  has as rows the Hermitian conjugate of the eigenvectors  $\{|v_1\rangle, |v_2\rangle\}$  of  $\mathbf{C}$ . Then, the exponential matrix of Eq.3.66 becomes ([wik]):

$$e^{-j\frac{1}{2}\Delta\beta\hat{\boldsymbol{\beta}}\cdot\boldsymbol{\sigma}z} = \mathbf{C}^{-1}e^{-j\frac{1}{2}\Delta\beta\mathbf{\Lambda}z}\mathbf{C} \quad (3.68)$$

Notice that since  $-j\frac{1}{2}\Delta\beta\mathbf{\Lambda}z$  is a diagonal matrix, its exponential  $e^{-j\frac{1}{2}\Delta\beta\mathbf{\Lambda}z}$  is a diagonal matrix being the element-wise exponential of  $-j\frac{1}{2}\Delta\beta\mathbf{\Lambda}z$  ([wik]). Hence, Eq.3.68 already provides a diagonalization of  $-j\frac{1}{2}\Delta\beta\hat{\boldsymbol{\beta}} \cdot \boldsymbol{\sigma}z$ . Then, Eq. 3.68 also implies that a diagonalizable matrix and its exponential have the same eigenvectors. Moreover, the eigenvalues of the exponential matrix are the exponential of the eigenvalues of the argument matrix.

The last step consists in finding the expression of the eigenvectors  $|v_1\rangle, |v_2\rangle$  to compute  $\mathbf{C}$  and the eigenvalues  $\lambda_1, \lambda_2$  to compute  $\mathbf{\Lambda}$  to calculate the diagonalization (3.67) of  $\hat{\mathbf{s}} \cdot \boldsymbol{\sigma}$ . To do so, it is useful to exploit the property [GK00]

$$|\hat{s}\rangle = \hat{\mathbf{s}} \cdot \boldsymbol{\sigma} |\hat{s}\rangle \quad (3.69)$$

where  $|\hat{s}\rangle$  is the unit Jones vector corresponding to the generic unit Stokes vector  $\hat{\mathbf{s}}$ . Eq.3.69 states that an eigenvector of the matrix  $\hat{\mathbf{s}} \cdot \boldsymbol{\sigma}$  is simply  $|\hat{s}\rangle$  with eigenvalue 1. It can be easily shown (from Eq.3.10 of [GK00]) that the Stokes vector  $\hat{\mathbf{s}}_{\perp}$  perpendicular to  $\hat{\mathbf{s}}$  is the other eigenvector of  $\hat{\mathbf{s}} \cdot \boldsymbol{\sigma}$ , with eigenvalue  $-1$ .

Hence, the eigenvectors  $|v_1\rangle, |v_2\rangle$  of  $\hat{\boldsymbol{\beta}} \cdot \boldsymbol{\sigma}$  are the Jones vectors  $|\hat{e}_{\chi}\rangle$  and  $|\hat{e}_{\Psi}\rangle$  corresponding to the Stokes vectors  $\hat{\boldsymbol{\beta}} \equiv \hat{\mathbf{e}}_{\chi}$  and  $\mathbf{e}_{\Psi}$  (which is perpendicular to  $\mathbf{e}_{\chi}$ ), with eigenvalues  $+1$  and  $-1$ . Then, the  $\mathbf{C}$  and  $\mathbf{\Lambda}$  of Eq.3.67, containing the eigenvectors and eigenvalues of  $\hat{\boldsymbol{\beta}} \cdot \boldsymbol{\sigma}$ , respectively, are

$$\mathbf{C} = \begin{bmatrix} \langle e_{\Psi} | \\ \langle e_{\chi} | \end{bmatrix} \quad \mathbf{\Lambda} = \begin{bmatrix} -1 & 0 \\ 0 & +1 \end{bmatrix} \quad (3.70)$$

Thus, Eq.3.68 assumes the form:

$$\mathbf{C}^{-1}e^{-j\frac{1}{2}\Delta\beta\mathbf{\Lambda}z}\mathbf{C} = \mathbf{C}^{-1}e^{-j\frac{1}{2}\Delta\beta\begin{bmatrix} -1 & 0 \\ 0 & +1 \end{bmatrix}z}\mathbf{C} = \mathbf{C}^{-1}\begin{bmatrix} e^{+j\frac{1}{2}\Delta\beta z} & 0 \\ 0 & e^{-j\frac{1}{2}\Delta\beta z} \end{bmatrix}\mathbf{C} \quad (3.71)$$

Hence, the solution (3.66) of the LSE can be written as

$$|s(z, \omega)\rangle = e^{-j\frac{1}{2}\beta(\omega)\cdot\boldsymbol{\sigma}z} |s(z=0, \omega)\rangle = \mathbf{C}^{-1}\begin{bmatrix} e^{+j\frac{1}{2}\Delta\beta(\omega)z} & 0 \\ 0 & e^{-j\frac{1}{2}\Delta\beta(\omega)z} \end{bmatrix}\mathbf{C} |s(z=0, \omega)\rangle \quad (3.72)$$

which shows that the transfer matrix for a fiber segment with uniform birefringence is

$$\mathbf{T} = \mathbf{C}^{-1} \begin{bmatrix} e^{+j\frac{1}{2}\Delta\beta z} & 0 \\ 0 & e^{-j\frac{1}{2}\Delta\beta z} \end{bmatrix} \mathbf{C} \quad (3.73)$$

Eq.3.73, derived from the solution of the LSE 3.65, is exactly the transfer matrix which has been presented in Eq.3.12 through reasonable and experimental-based considerations on the physics of system and has been used in the simulations. Moreover, its formulation is consistent with the definition of the Stokes vector of local birefringence  $\boldsymbol{\beta}$  since it has been used in the above proof.

Moreover, defining

$$\mathbf{B} := \frac{1}{2}j\boldsymbol{\beta} \cdot \boldsymbol{\sigma} \quad (3.74)$$

we obtain from (3.65), (3.67) and (3.70) that

$$\mathbf{B} = \boldsymbol{\beta} \cdot \boldsymbol{\sigma} = \mathbf{C}^{-1} \begin{bmatrix} -\Delta\beta & 0 \\ 0 & +\Delta\beta \end{bmatrix} \mathbf{C} \quad (3.75)$$

and so the LSE (3.65) can be written equivalently as

$$\frac{\partial |s(z, \omega)\rangle}{\partial z} + \frac{1}{2}j\mathbf{B} |s(z, \omega)\rangle = 0 \quad (3.76)$$

Besides the Jones vector formalism, Eq.(3.76) is the most common way of expressing the LSE for the 2-polarization mode case in literature and, from a symbolic perspective, identical to the LSE for the multimode case (see (4.1) in Ch.4).

Eq.(3.75) provides again the same explanation of the out-diagonal / coupling coefficients of  $\mathbf{B}$  matrix given in Sec.3.2.2: they are the result of both the birefringence intensity  $\Delta\beta$  and the misalignment between the pair of input orthogonal polarizations and the pair of eigenpolarizations of the present fiber segments. The higher the birefringence and the misalignment, the higher the coupling coefficient.

The exact knowledge of  $\Delta\beta$  and  $\mathbf{C}$  for every fiber segment in a real fiber is practically impossible and, actually, irrelevant for long-haul fibers where the statistical properties of PMD depend only on the mean square value of  $\boldsymbol{\beta}$  (or, equivalently, of  $\boldsymbol{\tau}$ ). Yet, Eq.3.75 offers an interesting insight on the origin and scaling of coupling coefficients.

Furthermore, it is possible to prove (see [GK00, Sec.6]) that the infinitesimal law of rotation 3.18 can be derived from Eq.3.65, which confirms once more the agreement between the LSE, the birefringence vector and its law of infinitesimal rotation.

To conclude the analysis, when the birefringence is not uniform, the fiber can be considered as a series of sections short enough to consider the birefringence constant within each section. Then, the differential equation Eq.3.65 can be solved

by dividing the integration in intervals corresponding to the sections of uniform birefringence and solving each differential equation through the methods described above. Finally, the continuity of the field has to be imposed at the boundary of each interval which, in this case, means that the output of one segment is the input of the next one. In this way, the connection between the LSE 3.65 and the channel transfer matrix computed as the product of the matrices of the single sections, Eq.3.24, is justified.

# Chapter 4

## Spatial Mode Coupling

The previous chapter has been dedicated to explain birefringence and mode coupling phenomena focusing on the case of 2-polarization modes, i.e., polarization-mode coupling. We now extend these concepts to the generic multimodal case of  $M = 2N$  supported modes, where there are  $N$  spatial modes and 2 polarizations for each of them. In this multimode scenario, both polarization mode coupling and spatial mode coupling are present.

With recall from Ch.2 that spatial mode coupling refers to the coupling occurring among modes with different spatial patterns. The coupling among spatial modes of the same manifold is referred to as intramode coupling, while the coupling among difference manifolds is called intermode coupling.

The origins of spatial mode coupling are analogous to the polarization mode ones, i.e., geometrical imperfections and external or internal fiber stresses, like bends, core ellipticity, torsion or external magnetic fields [Mar74; Pal13; Pal14]. However, a perturbation introducing birefringence does not have to induce also strong spatial mode coupling, and viceversa. That is, every perturbation has a different impact on coupling polarizations and spatial modes, with different perturbations coupling different modal groups and with a different intensity.

The difference in propagation constants  $\Delta\beta_{mk} = \beta_k - \beta_m$  between two modes is called phase mismatch and, for reasons which will be clear later on, it plays a relevant role in mode coupling. In particular, it is generally true that the higher the phase mismatch between two modes, the lower the coupling efficiency between them [Mar74; KHS12; Pal13; Pal14; Agr19]. Hence, polarization mode coupling is usually stronger than intramode coupling, which is in turn stronger than intermode coupling. Exceptions are perturbations with a specific spatial frequency which enhances the coupling between selected modal groups because of the particular phase mismatch [Mar74; Pal14].

A better understanding of the coupling mechanisms is offered by coupled mode theory [Mar74; Mar75; PN97], the broad theory under which falls the PSP theory.

Coupled mode theory models the interaction of modes through complex coefficients  $\kappa_{mn}$  appearing in the multimode linear Schrödinger equation (LSE) written in frequency domain as [Mar75, p. 44]

$$\frac{\partial \tilde{A}_m(z, \omega)}{\partial z} = -j\beta_m(\omega)\tilde{A}_m(z, \omega) - j \sum_{k=1}^M \kappa_{mk}(z, \omega)\tilde{A}_k(z, \omega) \quad (4.1)$$

where  $\tilde{A}_m(z, t)$  is the excitation coefficients, or modal amplitude, and  $\beta_m(\omega)$  the propagation constant of the  $m$ -th mode.

In general the coupling effects vary along fiber length. Hence, the coupling coefficients  $\kappa_{mk}$  are  $z$ -dependent. However, as for the PMD case, we can define the correlation length as the length over which such phenomena are approximately constant and choose to model the fiber again as a concatenation of uniform segments (or sections) of length  $L_S$  equal to the correlation length. Within each segment the coefficients  $\kappa_{mk}$  are constant, while  $\kappa_{mk}$  change from a segments to another.

The segment length can be assumed to be in the same range as the section length for birefringence (see the end of Sec.3.3), i.e., from 10 m to 1 km, depending on the specific scenario.

The concepts of weak- and strong-coupling regime, which are related to the correlation length and which have been presented at the end of Sec.3.3, can be extended to the multimode case, as well.

The LSE 4.1 has the exponential solution for a uniform segment and, exploiting the same reasoning of Sec.3.3 and Sec.3.4, for a generic series of segments the channel transfer matrix is the product of the matrices of the single segments as

$$\mathbf{T}(z \in [(n-1)L_S, nL_S], \omega) = e^{-j(\mathbf{B}_{ideal} + \mathbf{K}_n)(z - (n-1)L_S)} \prod_{i=1}^{n-1} e^{-j(\mathbf{B}_{ideal} + \mathbf{K}_i)L_S} \quad (4.2)$$

where  $n = 1, 2, \dots$  is the index of current segment,  $L_S$  is the segment length,  $\mathbf{B}_{ideal} = \text{diag}(\beta_1(\omega), \dots, \beta_M(\omega))$  is the matrix of ideal propagation,  $\mathbf{K}_n$  is the matrix whose elements are the coupling coefficients  $\kappa_{mk}$  for the current segment. Eq.(4.2) has a slightly cumbersome expression in order to be formally correct, yet the idea behind it is simple, as explained above.

The coupling among different modes is due to the off-diagonal elements of  $\mathbf{T}$  which are present due to the nonzero off-diagonal elements of the coupling matrices  $\mathbf{K}_n$ . Such elements introduce the so-called distributed coupling [SKP21], right because their effect accumulates along fiber length. However, when rotations of the segment axes or offsets between consecutive segments are considered, we talk about discrete coupling [SKP21]. Indeed, the matrices correspondent to these effects are directly multiplied to  $\mathbf{K}_n$  at the interface between two segments. To

draw a parallel, coupling in the 2-polarization case appeared due to a change of eigenpolarizations between consecutive segments and so it was a discrete coupling.

One might suspect that this description of discrete coupling is not consistent with (4.2) and it is partially true. While the presence of axis rotation is still included in  $\mathbf{K}_n$  and so in (4.2), the offset is not because an offset changes the boundary conditions between two consecutive segments requiring a projection of the modal basis of one segment onto the basis of the subsequent. However, we do not consider offsets in our model, so (4.2) is the valid solution of the LSE for us.

In order to mention a relevant phenomenon in multimode propagation, let us remember that the propagation constant can be expanded as

$$\beta_m(\omega_0 + \Delta\omega) = \beta_m^{(0)} + \beta_m^{(1)}\Delta\omega + \frac{1}{2}\beta_m^{(2)}(\Delta\omega)^2 + \dots \quad (4.3)$$

where  $\beta_m^{(0)} := \beta_m(\omega_0)$  and  $\beta_m^{(a)} := \left(\frac{d^a \beta_m}{d\omega^a}\right)_{\omega_0}$  are the dispersion coefficients. Observe that in the absence of coupling, i.e., when  $\kappa_{mk} \approx 0 \forall (m, k)$ , the LSE (4.1) indicates that the different modes of a multimode fiber travel at different group velocity  $v_{g,m} = \frac{1}{\beta_m^{(1)}}$  (see (3.6)) because LP modes belonging to different manifolds have different propagation constants. Note that for the exact fiber modes, the propagation constants are different also within a manifold. The different velocities among modes imply that pulses sent in parallel over them walk-off with distance, similarly to what has been shown for the case of birefringence in Fig.3.6b in Sec.3.2.5. This distortion phenomenon is called modal dispersion and it is graphically visible in Fig.4.1, where a Gaussian pulse at the transmitter excites the first  $M = 16$  modes of a GRIN fiber, which propagate for 20 km.

However, the presence of coupling tends to reduce modal dispersion because, intuitively, the various pulses over the modes start mixing due to the presence of distributed and discrete coupling effects, similarly to what has been discussed for the PMD case throughout Ch.3. The effectiveness of mode coupling in reducing the modal dispersion is graphically analyzed later in Sec.4.4.

Eq.(4.1) and its coefficients  $\kappa_{mk}$  result from the manipulation of Maxwell's equations (as proven in Ch.5 where nonlinear effects are considered as well), after having assumed the electric field to be expressible through a modal expansion as [Mar74; Mar75]

$$\mathbf{E}(\mathbf{r}, t) = \sum_k A_k(z, t) \mathbf{F}_k(\mathbf{r}, \omega_0) \quad (4.4)$$

where the summation runs, in principle, over all guided and radiation modes and  $\mathbf{F}_k(\mathbf{r}, \omega_0)$  is the modal profile of the  $k$ -th mode.  $\mathbf{F}_k$  is assumed to be frequency-independent within the narrowband spectrum of the propagating signal  $\mathbf{E}(\mathbf{r}, t)$ .

A fundamental concept of coupled mode theory is that, since the actual fiber is perturbed by  $z$ -dependent disturbances, while the modes are solution of an

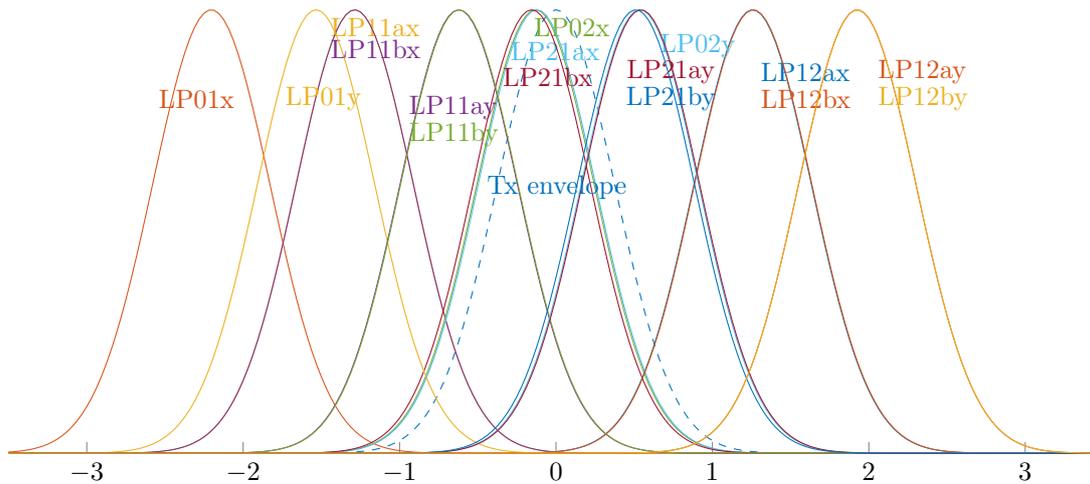


Figure 4.1: Modal dispersion for a Gaussian transmitted pulse with a symbol rate of 10 GHz in a GRIN fiber supporting  $M = 16$  modes. No mode coupling, only birefringence is present. The fiber length is  $L = 20$  km and the birefringence is  $\Delta n = 10^{-6}$ .

unperturbed structure, it is necessary to choose a fictitious ideal structure for which the modes are computed. The way the fictitious ideal structure can be chosen depends on the particular modal expansion. Several modal expansions are possible, all of them fulfilling some completeness theorem to ensure that any solution of the Maxwell's equations in the fiber is expressible as a superposition of modes [Mar74]. However, backward propagating modes and radiated modes are necessary for the completeness to hold, while in most of the cases one assumes the coupling to them to be negligible [Mar74; She+09; Jua+14]. It is not always clear when this approximation holds [KM04], but the simplification of the calculations is enough to justify its use. Since only forward guided modes are detected in the field of optical communications, backward waves and radiated modes, if excited, represent a power loss called coupling loss. Hence, when one wants to neglect the coupling losses, it is natural to neglect the presence of backward waves and radiated modes and consider an increased fiber attenuation  $\alpha$ . Another option is to consider the coupling with the first nonguided mode and periodically discard its energy to account for coupling losses, as done by [Jua+14]. We will just neglect the nonguided modes and assume the linear perturbations to be small to consider the coupling loss negligible.

Marcuse [Mar74] presents and discusses in detail two modal expansions for the isotropic waveguide: the ideal modal expansion (IM) and the local normal mode expansion (LNM). In [Mar75] he extends the IM technique to the case of anisotropic waveguides.

A material is said to be anisotropic if, considering linear effects only, the electric induction  $\tilde{\mathbf{D}}(\mathbf{r}, \omega) = \varepsilon_0 \varepsilon_r(\mathbf{r}, \omega) \tilde{\mathbf{E}}(\mathbf{r}, \omega)$  (see (5.11)) is not parallel to the electric

field  $\tilde{\mathbf{E}}(\mathbf{r}, \omega)$ . This implies that the relative dielectric permittivity  $\varepsilon_r(\mathbf{r}, \omega)$  is not a scalar, but a tensor.

The IM method assumes the fictitious ideal waveguide is characterized by a  $z$ -independent refractive index  $n_0(x, y)$ . Thus, the modal profiles  $\mathbf{F}_k(x, y)$  are also  $z$ -independent. The choice of  $n_0(x, y)$  is arbitrary in the sense that the coupling coefficients can be computed for any  $n_0(x, y)$ . The only discriminant is the convenience of a certain refractive index in simplifying the computation of the coefficients based on the specific perturbations acting on the fiber. For example in some circumstances a particular choice of geometry and refractive index profile for the ideal waveguide could lead to coupling coefficients that are zero for most of the fiber length, while another choice could have coefficients which never vanish.

Conversely to the IM method, in the LNM approach the fictitious ideal fiber is assumed to be at every point  $z$  coincident with the actual waveguide in geometry and material, that is,  $n_0(x, y, z) = n(x, y, z)$ , where  $n(x, y, z)$  is the actual perturbed refractive index. Hence, the modal profiles are also  $z$ -dependent and this is the reason for which in (4.4) we indicated a  $\mathbf{r}$ -dependent modal profile  $\mathbf{F}_k(\mathbf{r}, \omega_0)$  for the most general case.

The two approaches are both valid, besides some criticality for the boundary conditions, and the choice of one or the other is driven by reasons of convenience of calculations of the coupling coefficients on the specific situation, similarly to the hints given for the choice of the ideal waveguide in the IM case.

As mentioned, the expression for the coupling coefficient depends on the specific expansion chosen. In the general case of an anisotropic waveguide, the coupling coefficients for the IM approach are computed as [Mar75, Eq.48]

$$\kappa_{mk}(\omega) = \frac{\omega}{2c\sqrt{D_m D_k n_m n_k}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{F}_m^*(x, y, \omega_0) \cdot (\delta\varepsilon_r \mathbf{F}_k(x, y, \omega_0)) dx dy \quad (4.5)$$

where  $n_k$  is the effective refractive index of the  $k$ -th mode,  $D_k$  is the mode area of the  $k$ -th mode defined in (5.47), reported here for convenience

$$D_k = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \|\mathbf{F}_k(x, y, \omega_0)\|^2 dx dy$$

and  $\delta\varepsilon_r = \varepsilon_p - \varepsilon_r$  is relative the dielectric permittivity perturbation tensor, i.e., the difference between the actual (perturbed) dielectric tensor  $\varepsilon_p$  and the tensor  $\varepsilon_r = n_0^2(x, y)$  of the chosen ideal fictitious fiber.

Notice that  $\kappa_{mk} = \kappa_{mk}^*$  because  $\varepsilon_r$  is symmetric, given that the material is assumed lossless [Pal14]. This confirms that the coupling matrix  $\mathbf{K}$  is Hermitian and the transfer matrix  $\mathbf{T}$ , whose expression is given by (4.2), is unitary.

Marcuse provides the formula in [Mar75, Eq.48] with a different notation and with in a slightly different prefactor due to both a different normalization of the

modal profiles and to account also for backward waves and nonguided modes. On the opposite, we have expressed (4.5) to conform with the normalization of the modal amplitudes shown in (5.90) and with the notation we adopt in Ch.5, where we also provide its derivation from Maxwell's equations. Our proof is different from Marcuse's one, yet the results are equivalent, once the same notation and normalization for the modal profiles are used.

Eq.(4.5) is an approximation valid in the limit of small linear perturbations, that is,  $(\delta\varepsilon_r)_{ij} \ll \varepsilon_r$ . The more general and complicated formula, from which (4.5) is derived, is offered in [Mar75, Eq.46].

Notice that  $\kappa_{mk}(\omega)$  is in general frequency dependent due to the presence of the factor  $\omega$ . Strictly speaking, also the refractive indices  $n_m, n_k$  and the modal profiles  $\mathbf{F}_m, \mathbf{F}_k$  are frequency dependent, but the approximation with the values at the central frequency is less harmful in this case. Nonetheless, the frequency dependence of  $\kappa_{mk}$  is often neglected since being a small variation with respect to the contribution at the central frequency  $\kappa_{mk}(\omega_0)$  [MAS12b].

(4.5) should not confuse the reader with the orthogonality condition among modes, which, within the weakly-guiding approximation, is

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{F}_m^*(x, y, \omega_0) \cdot \mathbf{F}_k(x, y, \omega_0) dx dy = \delta_{mk} \quad (4.6)$$

where  $\delta_{mk}$  is Kronecker's delta.

Such condition is not in contrast with (4.5) because in (4.5) the polarization of  $\mathbf{F}_k$  is modified by the tensor  $\delta\varepsilon_r$ . Moreover, even in the isotropic case (i.e.,  $\delta\varepsilon_r$  is scalar),  $\delta\varepsilon_r$  varies along the section and, thus, the integral result is in general different from 0.

## 4.1 Importance of Phase-Matching

We have mentioned before that phase matching is a crucial concept for mode coupling. Assuming to be working with the IM expansion and to consider a segment with uniform perturbations, let us rewrite the modal expansion (4.4) in frequency domain factoring out the propagation constant, i.e.,

$$\mathbf{E}(\mathbf{r}, \omega) = \sum_k \tilde{A}'_k(z, \omega) \mathbf{F}_k(\mathbf{r}, \omega_0) e^{-j\beta_k(\omega)z} \quad (4.7)$$

where  $\tilde{A}'_k(z, \omega)$  is a slowly-varying envelope with  $z$ , since the fast oscillations induced by  $\beta(\omega)$  have been factored out.

The relation between the fast- and slowly-varying envelopes  $\tilde{A}_k(z, \omega)$  and  $\tilde{A}'_k(z, \omega)$  is immediate comparing the two expansions (4.4) (in frequency domain)

and (4.7), i.e.,

$$\tilde{A}_k(z, \omega) = \tilde{A}'_k(z, \omega)e^{-j\beta_k(\omega)z} \quad (4.8)$$

Inserting (4.8) in (4.1), we obtain the LSE for the slowly-varying amplitude [Mar74, p. 105]

$$\begin{aligned} \frac{\partial(e^{-j\beta_m(\omega)z}A'_m(z, \omega))}{\partial z} &= -j\beta_m(\omega)e^{-j\beta_m(\omega)z}A'_m(z, \omega) - j\sum_k \kappa_{mk}e^{-j\beta_k(\omega)z}A'_m(z, \omega) \\ &\Downarrow \\ -j\beta_m(\omega)e^{-j\beta_m(\omega)z}A'_m(z, \omega) + e^{-j\beta_m(\omega)z}\frac{\partial A'_m(z, \omega)}{\partial z} \\ &= -j\beta_m(\omega)e^{-j\beta_m(\omega)z}A'_m(z, \omega) - j\sum_k \kappa_{mk}e^{-j\beta_k(\omega)z}A'_k(z, \omega) \\ &\Downarrow \\ \frac{\partial A'_k(z, \omega)}{\partial z} &= -j\sum_k \kappa_{mk}e^{-j(\beta_k(\omega)-\beta_m(\omega))z}A'_k(z, \omega) \end{aligned} \quad (4.9)$$

Integrating the previous equation, we get [Mar74, p. 106]

$$A'_k(z, \omega) - A'_k(z_0, \omega) = -j\int_{z_0}^z \sum_k \kappa_{mk}e^{-j(\beta_k(\omega)-\beta_m(\omega))z}A'_k(z', \omega) dz' \quad (4.10)$$

Eq.(4.10) shows that, whenever there is a phase mismatch between two modes, a fast-oscillating term  $e^{-j(\beta_k(\omega)-\beta_m(\omega))z}$  appears in the the propagation equation and in its solution, so that the various small contributions to the integral have different phase and tend to average out (given that  $A'_k(z, \omega)$  is slowly-varying and  $\kappa_{mk}$  is constant in a uniform segment). Hence, the higher the phase mismatch, the lower the coupling efficiency. Provided that modes belonging to different manifolds have different propagation constants, intermode coupling is expected to be smaller than intramode coupling.

Observe that in a real fiber the situation is a bit more complicated than the one just described since the perturbations inducing the presence of the coupling coefficients  $\kappa_{mk}$  are  $z$ -dependent (yet, slowly-varying with  $z$  for unintentional perturbations in a typical communication system), as said before. However, the described principle is valid.

As a last remark, the previous analysis was considering IMs, while for LNMs the propagation constants are  $z$ -varying since the LNMs are  $z$ -varying. If a step-wise approximation like the previous one is undesired, an integration in the exponential is necessary to account for the changes in  $\beta_k$  along  $z$ , i.e., [Mar74, pp. 107, 108]

$$A'_k(z, \omega) - A'_k(z_0, \omega) = -j\int_{z_0}^z \sum_k \kappa_{mk}e^{-j\int_{z_0}^{z'}(\beta_k(z'', \omega)-\beta_m(z'', \omega)) dz''} A'_k(z, \omega) dz' \quad (4.11)$$

## 4.2 Mode Coupling through Physical Effects

The coupling coefficient (4.5) allows to model the coupling effect introduced by a perturbation, once a model for the dielectric tensor difference  $\varepsilon_r$  is known. Modeling linear coupling is still an active area of research because, despite the availability of physical models for the dielectric tensor perturbations, it is still necessary to assess how these models relate to reality through experimental analysis [Pal13; SKP21]. Every model requires the measurement of some physical parameters and it is not always clear how to choose them in a realistic way. For few disturbances Spenner [SKP21] carried out some experiments to find some realistic values for the coupling parameters and so, for the case of core ellipticity, we adopt his results as a reference for comparison with ours.

In this sections we present some common models for the relative dielectric tensor perturbation  $\varepsilon_r$ , derived mainly from Palmieri's work [Pal13; Pal14], which have been adapted to the different normalization we use (see (5.90) in Ch.5). Moreover, the factor  $\varepsilon_0$ , the dielectric permittivity of vacuum, has been removed from his equations since we work with the relative dielectric tensor, not the absolute one as him.

We consider axis rotations, bends and core ellipticity. Twist [Pal13] is not included because it requires considering the longitudinal components of the modes, which we do not do since working with LP modes. Offset [SKP21] is not considered since introducing losses.

Note that coupling can occur between the transverse components of two modes, between the longitudinal components, or between the transfer component of a mode and the longitudinal of the other. Given that the longitudinal component of a mode is significantly smaller than its transverse component(s), reason for which we also neglect it in the LP modal approximation, the intensity of the coupling coefficients is generally highest for transverse-transverse coupling, lower for transverse-longitudinal coupling, even lower for the longitudinal-longitudinal case [Pal13].

The transverse-longitudinal coupling, and in general the coupling between two orthogonal polarizations, happen only for anisotropic distortion effects, i.e., when  $\varepsilon_r$  is a tensor. In this case, even if two modes  $\mathbf{F}_m$  and  $\mathbf{F}_k$  have orthogonal polarizations so that  $\mathbf{F}_m \cdot \mathbf{F}_k = 0$ , the presence of  $\varepsilon_r$  acts as if the polarization of a mode was modified. Hence, the integrand  $\mathbf{F}_m \cdot (\varepsilon_r \cdot \mathbf{F}_k)$  of the coupling coefficient does not vanish and the coupling coefficient  $\kappa_{km}$  might not vanish.

On the opposite, if the model for the perturbation is scalar, i.e.,  $\varepsilon_r$  is a scalar, then given two orthogonal polarizations  $\mathbf{F}_m$  and  $\mathbf{F}_k$ ,  $\mathbf{F}_m \cdot (\varepsilon_r \cdot \mathbf{F}_k) = \varepsilon_r \mathbf{F}_m \cdot \mathbf{F}_k$  and so  $\kappa_{km} = 0$ .

If the longitudinal field component is neglected, which happens when working with LP modes as we do, only the elements of  $\varepsilon_r$  which are not relative to a  $z$

component are of interest.

Finally, for reasons of symmetry of the modes, the coupling coefficients  $\kappa_{mn}$  are expected to vanish for some indexes  $(m, n)$  for certain modal groups, as observed by Palmieri [Pal13]. This happens because both the modal profiles and the perturbations are periodic functions of the azimuth angle  $\Phi$  through a dependence of the type  $\sin(n\Phi)$  or  $\cos(n\Phi)$ . Such theoretical expectations, proposed by [Pal13], match the experimental results visible in the figures below depicting the power coupling matrices for each analyzed phenomenon. With *power matrix* we mean a matrix whose elements are the square norm of the elements of another matrix.

### 4.2.1 Axis Rotation

The coupling coefficients of a perturbing effect depend on the reference system adopted to express them. Let us assume that the modes are ordered for increasing modal order with all  $x$ -polarized waves first, and then all  $y$ -polarizations in the same order as the  $x$ -polarizations. Within a certain manifold, the even modes (i.e.,  $\cos(n\Phi)$  dependency) are put before the odd ones (i.e.,  $\sin(n\Phi)$  dependency). If we assume that the coupling matrix  $\mathbf{K}'$  of a certain perturbation is computed in a reference system parallel to the axes  $\{x', y'\}$  of the perturbation, then the coupling matrix in the reference frame  $\{x, y\}$  along which the modes are expressed can be shown to be computed as [Pal13; Pal14]

$$\mathbf{K} = \mathbf{R}^{-1} \mathbf{K}' \mathbf{R} = \mathbf{R}^T \mathbf{K}' \mathbf{R} \quad (4.12)$$

where in the last passage the orthogonality of  $\mathbf{R}$  has been exploited.  $\mathbf{R}$  is defined as a block diagonal matrix as

$$\mathbf{R} = \text{diag}(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M) \quad (4.13)$$

where  $\mathbf{R}_i$  with  $i = 1, \dots, M$  is the projection matrix for the  $i$ -th manifold. If the manifold has two modes (polarization degeneracy only), which means the azimuth order is 0 (as explained in Sec.2.1), then

$$\mathbf{R}_i = \begin{bmatrix} \cos(\alpha) & 0 & \sin(\alpha) & 0 \\ 0 & \cos(\alpha) & 0 & \sin(\alpha) \\ -\sin(\alpha) & 0 & \cos(\alpha) & 0 \\ 0 & -\sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} \quad (4.14)$$

If the manifold has four modes (polarization degeneracy and geometrical degeneracy) of azimuth order  $n$ , then

$$\mathbf{R}_i = \mathbf{R}_{pol} \mathbf{R}_{spat,i}$$

$$\mathbf{R}_{pol} = \begin{bmatrix} \cos(\alpha) & 0 & \sin(\alpha) & 0 \\ 0 & \cos(\alpha) & 0 & \sin(\alpha) \\ -\sin(\alpha) & 0 & \cos(\alpha) & 0 \\ 0 & -\sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} \quad (4.15)$$

$$\mathbf{R}_{spat,i} = \begin{bmatrix} \cos(n\alpha) & \sin(n\alpha) & 0 & 0 \\ -\sin(n\alpha) & \cos(n\alpha) & 0 & 0 \\ 0 & 0 & \cos(n\alpha) & \sin(n\alpha) \\ 0 & 0 & -\sin(n\alpha) & \cos(n\alpha) \end{bmatrix} \quad (4.16)$$

$\mathbf{R}_i$  accounts for the rotation (in a four dimensional space) of both the polarizations and the spatial modes, represented by  $\mathbf{R}_{pol}$  and  $\mathbf{R}_{spat,i}$ , respectively. An important remark is that the expression of the rotation matrices have been provided for the specific mode order adopted by us. If a different order is chosen, like alternating  $x$ - and  $y$ -waves, the matrices  $\mathbf{R}_{pol}$  and  $\mathbf{R}_{spat,i}$  have to be sorted accordingly. E.g., Palmieri [Pal13] provides the matrices for alternated  $x$ - and  $y$ -polarizations with even modes first and odd modes second. Moreover, our definition of  $\mathbf{R}_i$  coincides with the Hermitian of his rotation matrix because, conversely to Palmieri, we chose  $\mathbf{R}_i$  as the projection from  $\{x', y'\}$  to  $\{x, y\}$ , in accordance with our convention for the change-of-basis for the PMD case (see Ch.3).

The disposition of the elements of  $\mathbf{R}_{spat}$  makes it clear that such matrix introduces only intramode coupling. The intuition behind its expression is that a mode whose reference system is rotated w.r.t. another one where we want to express it, can be decomposed over only the modes of the same manifold of the second reference system. For instance, if the relative orientation between the two cartesian pairs of axes  $\{x, y\}$  and  $\{x', y'\}$  is  $90^\circ$ ,  $LP11ax$  becomes  $LP11by$ , with  $\hat{y}$  polarization. Expressing a modes in a different reference system cannot change its properties related to its propagation constant, so it is not possible that there is coupling to a different manifold. Hence,  $\mathbf{R}_i$  and  $\mathbf{R}$  do not introduce intermode coupling, given that  $\mathbf{R}_{pol}$  and  $\mathbf{R}_{spat,n}$  are only responsible for intramode coupling.

Finally, observe that the rotation matrices can be applied to any coupling matrix of a certain perturbation effect referred to its own perturbation axes, with the result of modeling rotated perturbations.

### 4.2.2 Bend

Several models for bends are available, like the scalar model used by Shemirani et al. [She+09] or the one used by Juarez et al. [Jua+14]. The model we report here

is from Palmieri [Pal13]. It considers bends in the  $x$ - $z$  plane and the correspondent dielectric perturbation tensor is

$$\varepsilon_r(x, y) = -an(x, y)^4 r \cos \Phi \begin{bmatrix} q_1 & 0 & 0 \\ 0 & q_1 & 0 \\ 0 & 0 & q_2 \end{bmatrix} \quad (4.17)$$

where  $a = 1/b$  is the curvature of the bend, i.e., the inverse of the bending radius  $b$ ,  $n$  is the fiber refractive index,  $r$  and  $\Phi$  are the radial and azimuth coordinates of a reference system centered in the center of the core,  $q_1$  and  $q_2$  are two parameters related to the fiber material. In fused silica  $q_1 \approx 0.206$  and  $q_2 \approx 0.031$ .

Observe that neglecting the longitudinal component of the modes,  $\varepsilon_r = -an(x, y)^4 r \cos \Phi q_1$  becomes a scalar and so polarization mode coupling is not possible.

Palmieri provides also the tensor of a second-order contribution arising from bends, which we neglect for simplicity.

In Fig.4.2a is depicted a realization of a normalized power coupling matrix (in dB) generated with the model (4.17) (with a certain axis rotation), for the first 15 spatial modes of the GRIN fiber presented in Ch.2.1. Containing values normalized to the maximum matrix element, it is independent on the chosen bending radius  $b$ . As mentioned before, some groups have vanishing coupling coefficients for reasons of symmetry of the modes and of the bend tensor 4.17.

However, since as argued in the previous section, the actual coupling strength is given by the elements of the channel transfer matrix  $\mathbf{T}$  which depend on the phase mismatch between the modes. Hence, to test the coupling introduced by bend we have simulated the channel matrix as

$$\mathbf{T} = e^{-j(\mathbf{B}_{ideal} + \mathbf{K})L} \quad (4.18)$$

where  $\mathbf{B}_{ideal}$  is the ideal propagation matrix and  $\mathbf{K}$  is the coupling matrix resulting from bend, that is, its elements  $\kappa_{mk}$  have been computed inserting the tensor model for bends (4.17) into the expression for the coupling coefficient (4.5). The fiber length has been set to  $L = 4.4$  km and the perturbation has been considered uniform all over the fiber length, i.e.,  $\kappa_{mk}$  have been considered  $z$ -independent. After some simulations, it has been noticed that the order of magnitude for the bending radius to have some non-negligible coupling is 1 cm. Such a value loses its physical sense because, if we model the fiber as a cascade of segments of length  $L_S$  in the order of tens to hundreds of meters, we expect the bend radius for each segment to be at least 100 m. To remedy, a possibility would be to consider the curvature  $a$  in (4.17) not for its physical sense of bending radius, but only as a model parameter to be tuned to introduce the desired level of coupling during simulations. In conclusion, this aspect needs further investigation. For the moment

we limit ourselves to show in Fig.4.2b the power transfer matrix for the case of bending radius  $b = 1$  cm. Observe that the coupling is limited to intramode or, at most, among adjacent modal groups for which the phase mismatch does not produce cancellation.

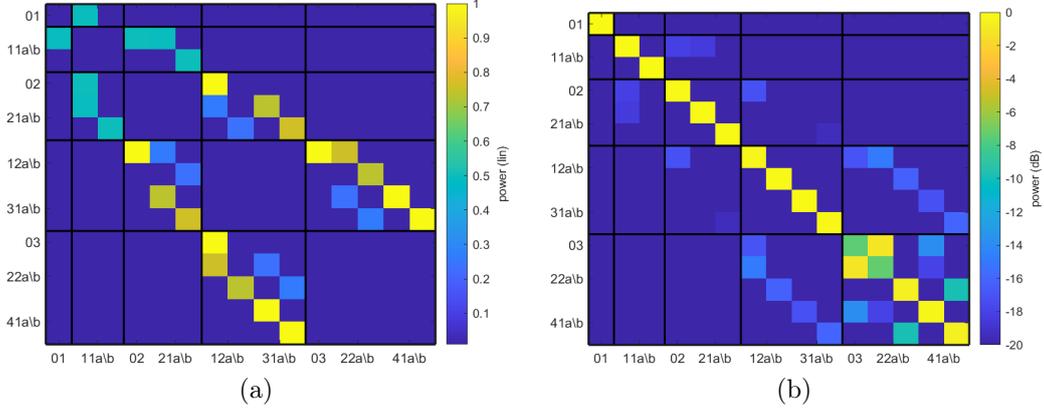


Figure 4.2: Power matrices for the bend model (4.17), for a bending radius  $b = 1$  cm. Fig.4.2a shows the power coupling matrix, Fig.4.2b shows the power transfer matrix for a fiber of length of 4.4 km, with uniform perturbation.

### 4.2.3 Core ellipticity

The model for the tensor perturbation associated to an elliptically deformed core along the reference  $x$ - $y$  axis is offered by Ulrich et al. [US79], as

$$\varepsilon_r(r, \Phi) = -\eta r \cos(2\Phi) \quad (4.19)$$

where, besides the symbols already defined for (4.17),  $\eta$  is a physical parameter to be tuned. In Fig.4.3a we show a realization of the coupling coefficients for the model (4.19) (with a certain axis rotation), where it is possible to observe the coefficients that are nonvanishing for the case of core ellipticity. It is not clear to us which is a realistic range of values for  $\eta$ , but from simulations it has been observed that  $\eta = 1.5$  yields a power transfer matrix qualitatively similar to the one presented in [SKP21]. For a fair comparison with [SKP21], the fiber length has been set to 4.4 km and the perturbation has been assumed uniform along the fiber length, following the same approach already described in the previous section. The resulting power transfer matrix is shown in Fig.4.3b. Again, notice that the actual coupling exists only within some manifold or among adjacent manifolds due to the phase-mismatch.

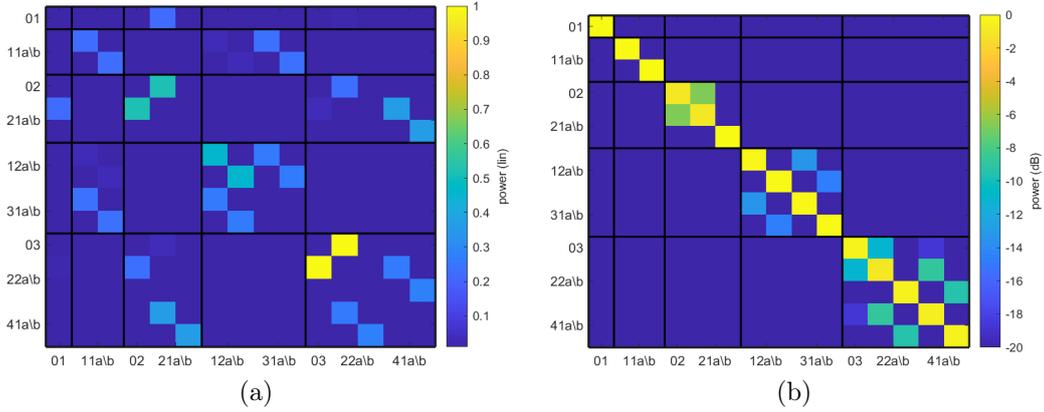


Figure 4.3: Power matrices for the model (4.19) of elliptical core, with  $\eta = 1.5$ . Fig.4.3a shows the power coupling matrix, Fig.4.3b shows the power transfer matrix for a fiber of length of 4.4 km, with uniform perturbation.

#### 4.2.4 Simulation Model of a Fiber with Physical Coupling Phenomena

Based on the analysis of the previous sections, we simulated a fiber channel as a concatenation of segments each one suffering from:

- birefringence, modeled by the matrix  $\mathbf{B}_{biref,i}(\omega)$
- bend, modeled by  $\mathbf{K}_{bend,i}$
- core ellipticity, modeled by  $\mathbf{K}_{ellip,i}$
- axis rotation, modeled by  $\mathbf{R}_i$ , as discussed in Sec.4.2.1, with random orientations
- eingepolarization variation, modeled by and  $\mathbf{C}_i$

This last effect is generated through a random walk on the Poincare sphere in Cartesian coordinates as described in Sec.3.3. A statistical generation approach has been preferred here to a physical one because birefringence and polarization mode coupling are experimentally known to be always present in fibers, but the physical models considered above do not introduce them if the longitudinal component of the modes is neglected.

The bend and the core ellipticity effects are also differently oriented with respect to the common axes of the segment through rotation matrices like  $\mathbf{R}_{bend,i}$  and  $\mathbf{R}_{ellip,i}$ , with the orientations randomly generated. That is,

$$\mathbf{K}_{bend,i} = \mathbf{R}_{bend,i}^H \mathbf{K}_{bend} \mathbf{R}_{bend,i}$$

and

$$\mathbf{K}_{ellip,i} = \mathbf{R}_{ellip,i}^H \mathbf{K}_{ellip} \mathbf{R}_{ellip,i}$$

where  $\mathbf{K}_{bend}$  and  $\mathbf{K}_{ellip}$  are the  $z$ -independent coupling matrices discussed in Sec.4.2.3 and Sec.4.2.2, respectively. The role of the rotation matrices  $\mathbf{R}_{ellip,i}$  and  $\mathbf{R}_{bend,i}$  is to simulate the possibility that the perturbation axes of the two phenomena are not always aligned.

Finally, the transfer matrix of this physical model for a segment with uniform perturbations is

$$\mathbf{T}_i(z, \omega) = e^{-j \left( \mathbf{C}_i^H \mathbf{R}_i^H \left( \mathbf{B}_{ideal}(\omega) + \mathbf{B}_{biref,i}(\omega) + \mathbf{K}_{bend,i} + \mathbf{K}_{ellip,i} \right) \mathbf{R}_i \mathbf{C}_i \right) z} \quad (4.20)$$

This model is exploited later in Sec.4.4 to perform some simulations.

A realization of the power transfer matrix for a fiber of length  $L = 4.4$  km, with uniform perturbations with bend parameter  $b = 1$  cm and ellipticity parameter  $\eta = 1.5$ , birefringence  $\Delta n = 10^{-6}$  is depicted in Fig.4.4a. Observe the present of significant intramode coupling and little coupling among adjacent modal groups, while, due to the phase-mismatch, almost no intermode coupling among groups far apart. Notice also the polarization dependence of the coupling due to the axis rotation and the eingepolarizations presence.

A realization of a power transfer matrix for a fiber of length  $L = 4.4$  km, with the same parameters as before, but with segments of length  $L_S = 100$  m, is provided in Fig.4.4b. Notice the higher level of coupling thanks to the nonuniformity of the modeled perturbations.

### 4.3 Mode Coupling through a Statistical Approach

The physical models for mode coupling explained so far are not the one commonly employed to introduce mode coupling in signal propagation. Mode coupling is often modeled through a statistical approach for long-haul communications [Pal13], in particular in combination with the study of nonlinear effects, usually considering strong intramode coupling [MMW97; MAS12b; MAS12a] and, if desired, also strong intermode coupling [MEA13; ASM16]. The basic idea behind the statistical approach is that over distances much longer than the correlation length of the linear coupling disturbances the details of the physical coupling models are irrelevant.

The statistical approach consists in considering as channel matrix of the  $i$ -th segment

$$\mathbf{T}_i(z, \omega) = \mathbf{R}_i e^{-j \mathbf{B}_{ideal}(\omega) z} \quad (4.21)$$

where  $\mathbf{B}_{ideal} = \text{diag}(\beta_1(\omega), \dots, \beta_M(\omega))$  accounts for the ideal propagation and  $\mathbf{R}_i$  is a random unitary matrix. A numerically stable algorithm to generate random

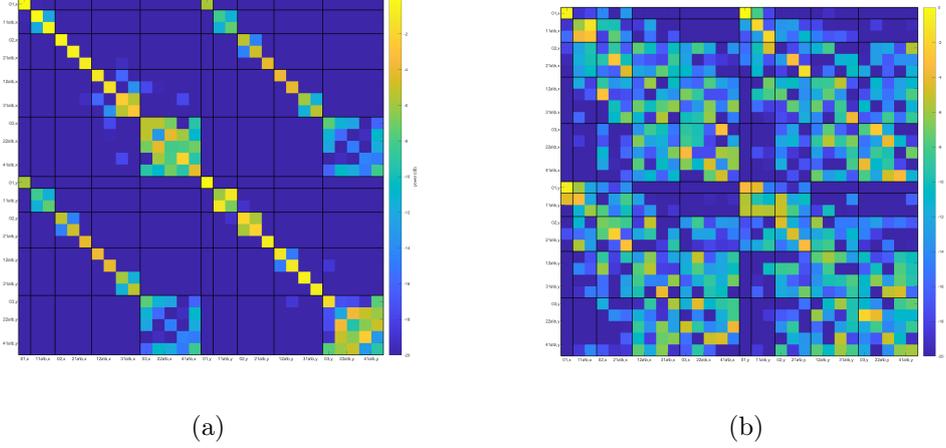


Figure 4.4: Realizations of power matrices for the multimode fiber model of Sec.4.2.4. Fiber length  $L = 4.4$  km, bend parameter  $b = 1$  cm, ellipticity parameter  $\eta = 1.5$ , birefringence  $\Delta n = 10^{-6}$ . Fig.4.4a is for uniform perturbations along  $L$ . Fig.4.4b is for nonuniform perturbations, with segment length  $L_S = 100$  m.

coupling matrices uniformly distributed on the the set  $U(M)$  of unitary matrices of dimensions  $M \times M$  follows the next steps (see [Mez06] for the formal background):

1. generate a  $M \times M$  matrix  $\mathbf{A}$  with elements drawn by a complex standard normal distribution.
2. decompose  $\mathbf{A}$  as  $\mathbf{A} = \mathbf{Q}\mathbf{R}$  through a QR decomposition algorithm.
3. build the diagonal matrix

$$\mathbf{\Lambda} = \begin{bmatrix} \frac{r_{11}}{|r_{11}|^2} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \frac{r_{MM}}{|r_{MM}|^2} \end{bmatrix}$$

where  $r_{ii}$  is the  $i$ -th diagonal element of  $\mathbf{R}$ .

4. the matrix  $\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}$  is the desired random uniform unitary matrix.

An example of realization of a random unitary matrix  $\mathbf{R}$  generated through the mentioned algorithm is depicted in Fig.4.5 (power matrix), where it is evident that all coupling coefficients have similar strength, conversely to the matrices for the physical models which tend to couple only few groups. Notice that the a discrete coupling like the one introduced by  $\mathbf{R}$  does not suffer of the canceling effect of the phase-mismatch because it is independent on it. Also for this reason the modes

are expected to be more strongly coupled when a statistical approach is employed compared to the case in which physical models are used, if an adequate number of segments are employed.

A power transfer matrix is not shown here since it would be qualitatively the same as Fig.4.5.

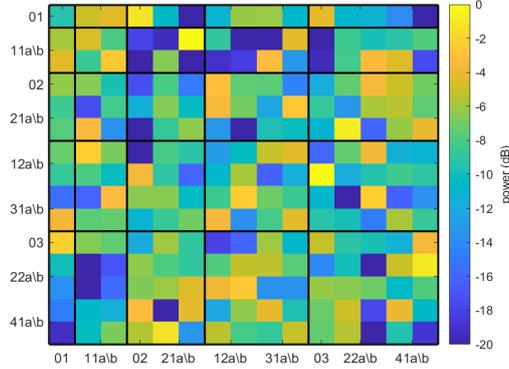


Figure 4.5: Power matrix of the realization of a random unitary matrix for the statistical approach of Sec.4.3 following the Algorithm 4.3.

## 4.4 Primary Modes and Simulations

The concept of Principal States of Polarization can be extended to the multimode case, where the modes fulfilling the definition of being undistorted to first order in frequency are called Primary Modes (PMs) [FK05; KHS12; Ant+12; CES16]. The definitions of Jones and Stokes spaces can also be extended to the multimode case [Ant+12], with only few differences with the single mode case.

Therefore, it is not surprising that given the GD operator (formally identical to (3.44) the GD operator for the PSPs) [FK05; KHS12; Ant+12]

$$\mathbf{G} = -j\mathbf{U}^H\mathbf{U}_\omega \quad (4.22)$$

the modes can be proved to be the eigenvectors of (4.22) and the GDs are their eigenvalues, as for the PSPs (see Sec.3.3.3). The proof proceeds formally the same way it has been shown in Sec.(3.3.3) for the PSPs, with the only difference that the Jones vectors this time contain  $M \geq 2$  modes.  $\mathbf{U}$  is again the the transfer matrix without the common phase, see (3.26).

Fig.4.6 illustrates the property of the PMs not to suffer from GD dispersion within a perturbed fiber channel, by comparing the envelopes detected over the PMs with the envelopes detected over the ideal modes. In this particular case, the first  $M = 6$  modes of a GRIN fiber have been excited with a Gaussian envelope at a

symbol rate of 10 GHz, exploiting the physical fiber model of Sec.4.2.4 for a fiber of length  $L = 10$  km, with segments of 1 km. Group-velocity dispersion has also been considered. Observe that the LP modes at the receiver are differently delayed and distorted due to modal dispersion and mode coupling. On the opposite, the PMs retain better the Gaussian shape. Their amplitudes are differently since depending on the power coupled to each input PM.

Notice also that the ideal modes are divided between LP01 and LP11 groups for what concerns the propagation delays. This boils down from the fact that there is lower intermode coupling than intramode one.

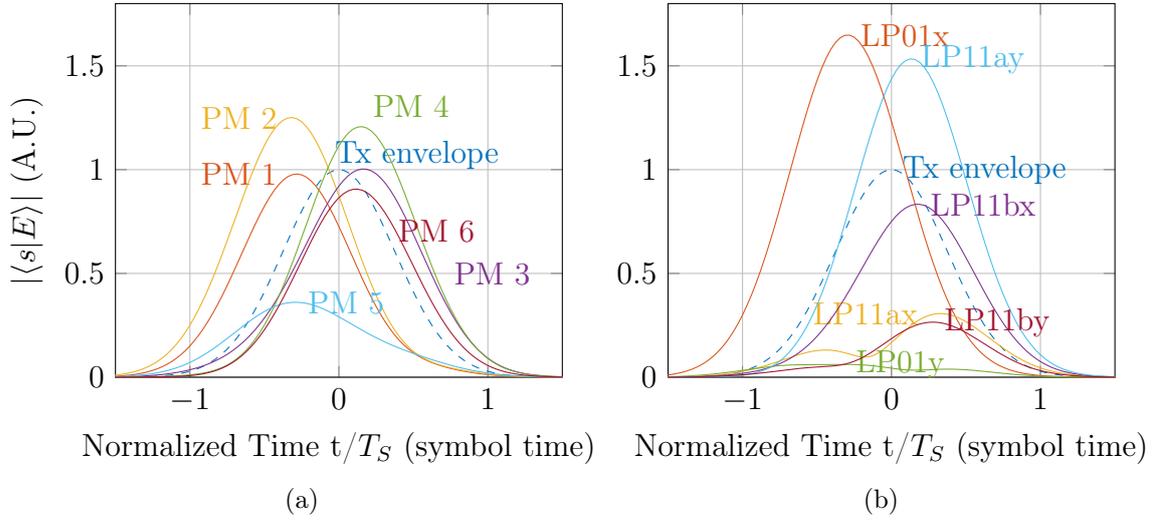


Figure 4.6: Comparison between the envelopes detected over the ideal modes (Fig.4.6a) and over the PMs (Fig.4.6b). The first  $M = 6$  modes of a GRIN fiber have been excited with a Gaussian envelope at a symbol rate of 10 GHz, exploiting the physical fiber model of Sec.4.2.4 for a fiber of length  $L = 10$  km, with segments of 1 km. Group-velocity dispersion has also been considered.

It can be shown that, analogous to the PMD case, the mean of the maximum delay difference (MDGD) between two PMs scales linearly with the fiber length in the weak-coupling regime, while it scales with the square root of the fiber length for the strong-coupling regime [KHS12]. That is, strong-coupling reduces the modal dispersion because the MDGD is lower.

Khan [KHS12] showed that the p.d.f. of the MDGD is no more Maxwellian as for the 2-polarizations case, but semicircular. Given that the complexity of a MIMO receiver depends on the duration of the channel impulse response and the duration is proportional to the MDGD, knowing the statistics of the MDGD is useful for designing the MIMO compensation receiver. Moreover, since the MDGD can span thousands of symbols [KHS12] for a typical long-haul transmission in a

weak-coupling scenario, the presence of strong coupling is usually desired to reduce the distortions and the receiver complexity [KHS12; Ant+12].

The general principles mentioned for the PMD that increasing the variance of the perturbations [KHS12; She+09] and reducing the segment length tend to increase coupling are still valid. Indeed, in such circumstances the correlation length reduces and the coupling is “stronger”.

The effect of mode coupling in reducing the distortions on the received envelopes is visible comparing the previous case of Fig.4.6b with Fig.4.7a. In both cases a Gaussian pulse at 10 GHz has been sent over the first  $M = 6$  modes of a GRIN fiber of length  $L = 10$  km, modeled according to the physical approach of Sec.4.2.4 (considering GVD as well). The difference between the two cases is that the segment length has been reduced from 1 km for Fig.4.6b, to 100 m for Fig.4.7a. In this second case, the DGDs among the PMs are lower than in the first one and, thus, also the distortions suffered by a pulse over a generic mode.

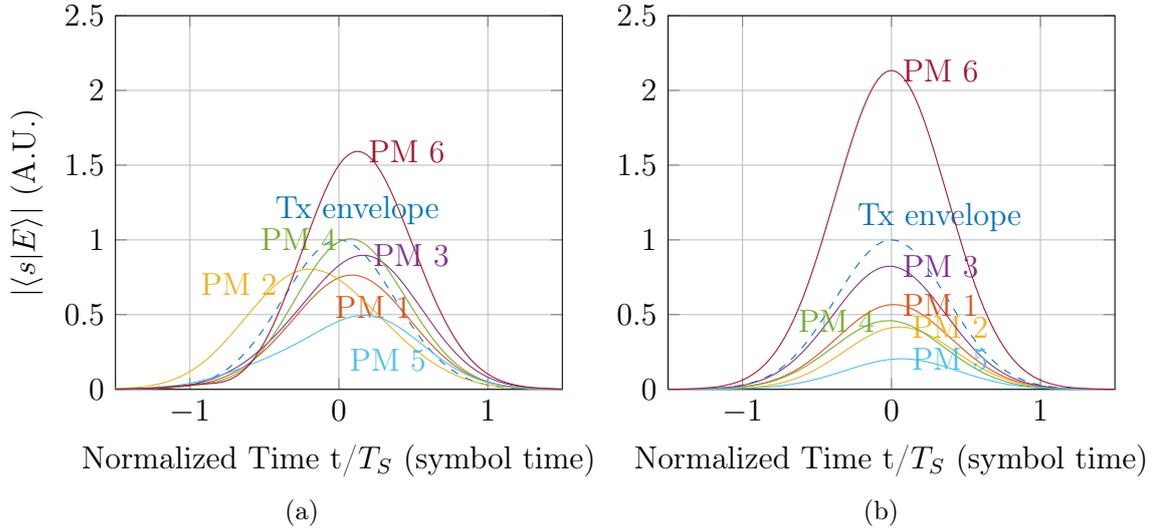


Figure 4.7: Comparison between the physical coupling approach, Fig.4.7a, and the random coupling one, Fig.4.7b. The first  $M = 6$  modes of a GRIN fiber have been excited with a Gaussian envelope at a symbol rate of 10 GHz, for a fiber of length  $L = 10$  km, with segments of 100 m. Group-velocity dispersion has also been considered.

The physical model of Sec.4.2.4 has also been compared, through simulations, against the statistical model of Sec.4.3. As expected from the considerations of Sec.4.3, the statistical approach introduces higher coupling and so lower distortion. This is visible comparing Fig.4.7a (physical approach of Sec.4.2.4) with Fig.4.7b (statistical approach of Sec.4.3). Both simulations have been run transmitting a Gaussian pulse at 10 GHz over the first  $M = 6$  modes of a GRIN fiber of length

$L = 10$  km with segment length  $L_S = 100$  m.

The higher coupling effects of the random approach are even more pronounced considering a transmitted Gaussian pulse at 100 GHz, as visible comparing the Fig.4.8a, for the physical approach, with Fig.4.8b, for the random matrix approach. An intuitive explanation in time domain is that the pulses walk-off quicker due to the lower duration, for which the coupling introduced by the physical perturbations cannot “compensate”. In frequency one can draw an analogy with the PSP case, where a higher band meant exceeding the PSPs coherence band and, thus, higher-order distortion effects kicked in.

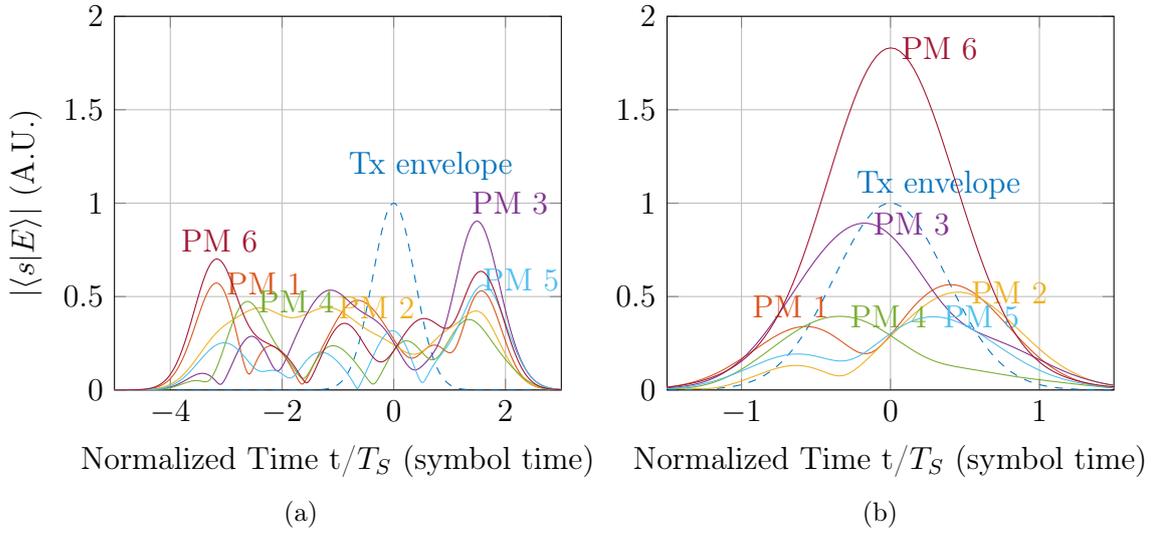


Figure 4.8: Comparison between the physical coupling approach, Fig.4.8a, and the random coupling one, Fig.4.8b, for a symbol rate of 100 GHz over  $M = 6$  modes of a GRIN fiber of length  $L = 10$  km, with segments of 100 m. Group-velocity dispersion has also been considered.

The significant discrepancy in DGDs and distortion of the envelopes between the physical and the random models emerging from the presented plots raises the question of which scenarios and to which extend the assumption of strong intermode coupling is reasonable. More investigation is needed in this sense.

# Chapter 5

## Generalized Multimode Nonlinear Schrödinger Equation

The present chapter aims at deriving the propagation equation for the case when the nonlinearities are of interest, as it is for the case of long-reach optical communications. Such propagation equation is called in the literature in various way, like coupled nonlinear Schrödinger equation (CNLSE) ([ASM16]), vectorial nonlinear Schrödinger equation (VNLSE) ([Han21]), multimode nonlinear Schrödinger equation (MMNLSE) ([Agr19]) or simply nonlinear Schrödinger equation (NLS / NLSE) ([Agr19]). We adopt the acronym GMMNLSE, which stands for generalized multimode nonlinear Schrödinger equation (also MM-GNLSE [PH08; HP12]) used by Wright et al. [Wri+17].

One of the first and most popular works to derive the GMMNLSE as used in optical communications has been done by Poletti and Horak [PH08], based on a previous paper of Kolesik [KM04]. However, the derivation proposed here mainly follows the procedure of Agrawal [Agr19] and Antonelli et al. [ASM16].

Before digging into the math, the main “tools” we use are:

1. Maxwell’s equations
2. Constitutive relations of the material
3. Expression for the material polarization vector
4. Fiber modal expansion

The basic idea of the derivation is to exploit the constitutive relations and the material polarization vector to obtain a wave equation containing only an unknown, the electric field. Then, the electric field is decomposed through a modal expansion, which allows to retrieve the nonlinear propagation equation for every modal amplitude.

A number of approximations are going to be made to simplify a problem which would be otherwise too general and complex. In the following, the mathematical formality is kept to a minimum, which means that, e.g., the hypotheses which allow to exchange order between space and time derivatives, as well as the conditions for the existence of some integrals and of some Fourier transforms are always tacitly assumed to hold for the cases of interest.

## 5.1 Derivation

The derivation starts from Maxwell's equations [Agr19, p. 27][Coe10, p. 16][Mid03]

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (5.1a)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = +\frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} + \mathbf{J}(\mathbf{r}, t) \quad (5.1b)$$

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = \varrho_c; \quad (5.1c)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = \mathbf{0}; \quad (5.1d)$$

$$(5.1e)$$

where  $\mathbf{E}$  is the electric field vector,  $\mathbf{H}$  is the magnetic field vector,  $\mathbf{D}$  is the electric displacement or electric induction vector,  $\mathbf{B}$  is the magnetic induction or magnetic flux density vector,  $\mathbf{J}$  is the current density,  $\varrho_c$  is the volumetric charge density.

$\nabla$  is the nabla operator, through are expressed in a symbolic form the curl as  $\nabla \times$ , the divergence as  $\nabla \cdot$ , the gradient as  $\nabla$  and the Laplacian as  $\nabla^2$ .

In optical fibers there are no free charges, hence [Agr19]

$$\mathbf{J} = \mathbf{0} \quad (5.2a)$$

$$\varrho_c = 0 \quad (5.2b)$$

The constitutive relations of the material are

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \quad (5.3a)$$

$$\mathbf{B} = \mu_0 \mathbf{H} + \mu_0 \mathbf{M} \quad (5.3b)$$

where  $\mathbf{P}$  is the induced electric material polarization vector,  $\mathbf{M}$  is the induced magnetic polarization,  $\varepsilon_0$  is the vacuum dielectric permittivity,  $\mu_0$  is the vacuum magnetic permeability. For a nonmagnetic material, like those of which fibers are made of, [Agr19, p. 27][Coe10, p. 16]

$$\mathbf{M} = \mathbf{0} \quad (5.4)$$

The electric material polarization vector can be expanded as [Coe10, p. 2.27]:

$$\begin{aligned}
 \mathbf{P}(\mathbf{r}, t) := & \varepsilon_0 \int_{-\infty}^{+\infty} \chi^{(1)}(t - t_1) \cdot \mathbf{E}(t_1) dt_1 \\
 & \varepsilon_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi^{(2)}(t_1 - t, t_2 - t) : \mathbf{E}(t_1) \mathbf{E}(t_2) dt_1 dt_2 \\
 & \varepsilon_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi^{(3)}(t_1 - t, t_2 - t, t_3 - t) : \mathbf{E}(t_1) \mathbf{E}(t_2) \mathbf{E}(t_3) dt_1 dt_2 dt_3 \\
 & + \dots
 \end{aligned} \tag{5.5}$$

It can also be written as

$$\mathbf{P}(\mathbf{r}, t) := \mathbf{P}_L(\mathbf{r}, t) + \mathbf{P}_{NL}(\mathbf{r}, t) \tag{5.6}$$

where [Agr19, p. 28]

$$\mathbf{P}_L(\mathbf{r}, t) := \varepsilon_0 \int_{-\infty}^{+\infty} \chi^{(1)}(\mathbf{r}, t - t') \cdot \mathbf{E}(\mathbf{r}, t') dt' \tag{5.7}$$

is the linear part of the polarization vector and  $\chi^{(1)}$  is the first-order nonlinear susceptibility tensor, related to the refractive index of the medium. Its expression in the frequency domain is

$$\tilde{\mathbf{P}}_L(\mathbf{r}, \omega) = \varepsilon_0 \tilde{\chi}^{(1)}(\mathbf{r}, \omega) \cdot \tilde{\mathbf{E}}(\mathbf{r}, \omega) \tag{5.8}$$

Moreover,

$$\mathbf{P}_{NL}(\mathbf{r}, t) := \varepsilon_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi^{(3)}(t_1 - t, t_2 - t, t_3 - t) : \mathbf{E}(t_1) \mathbf{E}(t_2) \mathbf{E}(t_3) dt_1 dt_2 dt_3 \tag{5.9}$$

is the nonlinear part of the polarization vector, where  $\chi^{(3)}$  is the third-order nonlinear susceptibility tensor, from which arise the Kerr nonlinear effect, also called Four Wave Mixing (FWM), and the Raman nonlinear effect [HP12; Agr19; Wri+17; Han21]. Notice  $\mathbf{P}_{NL}$  accounts for the third-order nonlinear effects only because the second order-term of (5.5) is negligible in the optical fibers for symmetry reasons [Coe10, p. 18], while the terms of order higher than the third are neglected. Let us insert (5.5) in (5.3a)

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}_L + \mathbf{P}_{NL} \tag{5.10}$$

and transform it to frequency domain

$$\tilde{\mathbf{D}} = \varepsilon_0 \tilde{\mathbf{E}} + \tilde{\mathbf{P}}_L + \tilde{\mathbf{P}}_{NL} = \varepsilon_0 \tilde{\mathbf{E}} + \varepsilon_0 \tilde{\chi} \cdot \tilde{\mathbf{E}}(\mathbf{r}, \omega) + \tilde{\mathbf{P}}_{NL} = \varepsilon_0 (\mathbf{I} + \tilde{\chi}^{(1)}) \tilde{\mathbf{E}} + \tilde{\mathbf{P}}_{NL} \tag{5.11}$$

We define [Agr19, p. 2.1.13] [Han21, p. 3.18]

$$\varepsilon_r(\mathbf{r}, \omega) + \delta\varepsilon_r(\mathbf{r}, \omega) := \mathbf{I} + \tilde{\chi}^{(1)}(\omega) \quad (5.12)$$

where  $\varepsilon_r$  is the material relative dielectric tensor and  $\delta\varepsilon_r$  is its perturbation induced by the linear perturbations. We also define [Agr19, p. 2.1.14] [Coe10]

$$\varepsilon_r(\mathbf{r}, \omega) := \left( n(\mathbf{r}, \omega) \mathbf{I} + j \frac{\alpha c}{2\omega} \right)^2 \quad (5.13)$$

where  $n$  is the material refractive index and  $\alpha$  is the fiber loss. Notice this is also the definition of refractive index. Since  $\alpha$  is small for a fiber of usual material and frequency range, we neglect it and insert it in the final equation, when needed, through a perturbative approach [Agr19, p. 29]. Hence, (5.13) reduces to

$$\varepsilon_r(\mathbf{r}, \omega) \approx n(\mathbf{r}, \omega)^2 \mathbf{I} \quad (5.14)$$

An explicit expression for  $\delta\varepsilon_r$  is available only when a specific distortion effect is considered. Moreover,  $\delta\varepsilon_r$  is in general a tensor to account for anisotropic disturbances. However, when the stress is isotropic, it reduces to be proportional to a scalar, i.e., the refractive index perturbation  $\delta n^2$ . Nonetheless, for the sake of clarity, we can write the tensor  $\delta\varepsilon$  as

$$\delta\varepsilon_r = \begin{bmatrix} \delta\varepsilon_{xx} & \delta\varepsilon_{xy} & \delta\varepsilon_{xz} \\ \delta\varepsilon_{yx} & \delta\varepsilon_{yy} & \delta\varepsilon_{yz} \\ \delta\varepsilon_{zx} & \delta\varepsilon_{zy} & \delta\varepsilon_{zz} \end{bmatrix} \quad (5.15)$$

In an anisotropic material  $\delta\varepsilon_r$  contains off-diagonal elements, while  $\delta\varepsilon_r$  is proportional to the identity matrix if no anisotropy is present. The anisotropy corresponds to have a response of the medium which is not parallel to the action on it [Mid03; Coe10], that is,  $\mathbf{D}$  is not parallel to  $\mathbf{E}$ , as can be seen from (5.11). As a side note, when LP modes are considered (weakly-guiding approximation), the  $z$  component of the electromagnetic field is negligible and so the elements the third column and the third row of (5.15) are not of interest.

Inserting firstly (5.12) and then (5.14) in (5.11), we get

$$\tilde{\mathbf{D}} = \varepsilon_0(n^2 \mathbf{I} + \delta\varepsilon_r) \tilde{\mathbf{E}} + \tilde{\mathbf{P}}_{NL} \quad (5.16)$$

Up to now we have not started manipulating Maxwell's equations, but just paved down the preliminaries. The next target is to retrieve a propagation equation where only the electric field is figuring. To do so, we have to get rid of  $\mathbf{D}$  and  $\mathbf{B}$ . Thus, let us insert (5.3a) in (5.1b) and cancel  $\mathbf{J}$  thanks to (5.2), obtaining

$$\nabla \times \mathbf{H} = \frac{\partial(\varepsilon_0 \mathbf{E} + \mathbf{P})}{\partial t} = \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t} \quad (5.17)$$

Similarly to (5.17), inserting (5.3b) in (5.1b), canceling  $\mathbf{M}$  thanks to (5.4), we retrieve

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \quad (5.18)$$

Applying the curl to (5.18)

$$\nabla \times \nabla \times \mathbf{E} = -\mu_0 \frac{\partial(\nabla \times \mathbf{H})}{\partial t} \quad (5.19)$$

and inserting (5.17) in 5.19, it is obtained

$$\nabla \times \nabla \times \mathbf{E} = -\mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu_0 \frac{\partial \mathbf{P}}{\partial t} = -\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu_0 \frac{\partial \mathbf{P}}{\partial t} \quad (5.20)$$

where it has been exploited the relation for the speed light in vacuum  $c$

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} \quad (5.21)$$

Eq.(5.20) presents a double curl of  $\mathbf{E}$  which is undesired to us since we intend to retrieve an expression directly relating the space derivatives to the time derivative of the field. Hence, we exploit the identity

$$\nabla \times \nabla \times \equiv \nabla (\nabla \cdot) - \nabla^2 \quad (5.22)$$

to write the l.h.s. of (5.20) as

$$\nabla \times \nabla \times \mathbf{E} = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} \quad (5.23)$$

Neglecting the term  $\nabla (\nabla \cdot \mathbf{E})$  in the previous expression would simplify our calculations. This is surely satisfied if  $\nabla \cdot \mathbf{E} = \mathbf{0}$ . To see when this happens, apply the divergence to the constitutive relation (5.3a)

$$\nabla \cdot \mathbf{D} = \nabla \cdot (\varepsilon_0 \mathbf{E} + \mathbf{P}) = \varepsilon_0 \nabla \cdot \mathbf{E} + \nabla \cdot \mathbf{P} \quad (5.24)$$

Since  $\mathbf{D} = \mathbf{0}$ , we have

$$\varepsilon_0 \nabla \cdot \mathbf{E} = -\nabla \cdot \mathbf{P} \quad (5.25)$$

$$\Rightarrow \nabla \cdot \mathbf{E} = \mathbf{0} \Leftrightarrow \nabla \cdot \mathbf{P} = \mathbf{0} \quad (5.26)$$

Thus, to nullify  $\nabla \cdot \mathbf{E}$  is necessary to set

$$\nabla \cdot \mathbf{P} = \nabla \cdot \mathbf{P}_L + \nabla \cdot \mathbf{P}_{NL} = \mathbf{0} \quad (5.27)$$

That is, the material polarization vector contribution to (5.23) is neglected. If we observe that in general  $\mathbf{P}_L \neq \mathbf{P}_{NL}$ , condition (5.27) means

$$\nabla \cdot \mathbf{P}_{NL} \approx \mathbf{0} \quad (5.28)$$

$$\nabla \cdot \mathbf{P}_L \approx \mathbf{0} \quad (5.29)$$

That is, basically, both the linear and nonlinear polarizations are required to be nonvarying in space, or at least approximately. Since  $\mathbf{P}_{NL}$ , if present, is  $\mathbf{r}$  dependent (see, e.g., 5.86), condition (5.28) means assuming  $\mathbf{P}_{NL}$  has a negligible contribution to the divergence. Concerning  $\mathbf{P}_L$ , its connection with the refractive index and the linear distortions has been shown through (5.11), (5.12) and (5.13). Hence, condition (5.28) corresponds to not only assuming the linear distortions (which are  $\mathbf{r}$  dependent) vary little over space, but also the refractive index variation over the fiber cross-section is negligible. This last assumption is exact in case of SI fibers (see Sec.2.1) because the refractive index is constant within the core and within the cladding and (5.23) is solved separately for the two parts. Then, the boundary conditions are exploited to match the field at the core-cladding interface. On the opposite, condition (5.28) is only an approximation for, e.g., GRIN fibers. Depending on the desired level of accuracy, this approximation is not always assumed in literature. Yet, for our scopes it is sufficient and it allows to reach the same nonlinear propagation equation as common references in literature [PH08; ASM16; Agr19].

With these considerations in mind, we approximate (5.23) as

$$\nabla \times \nabla \times \mathbf{E} \approx \nabla^2 \mathbf{E} \quad (5.30)$$

and then insert (5.30) in (5.20), obtaining

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu_0 \frac{\partial \mathbf{P}}{\partial t} \quad (5.31)$$

We transform (5.31) to frequency domain

$$\nabla^2 \tilde{\mathbf{E}} = \frac{1}{c^2} (j\omega)^2 \tilde{\mathbf{E}} + (j\omega)^2 \mu_0 \tilde{\mathbf{P}} = -\frac{\omega^2}{c^2} \tilde{\mathbf{E}} - \omega^2 \mu_0 \tilde{\mathbf{P}} = -k_0^2 \tilde{\mathbf{E}} - \omega^2 \mu_0 \tilde{\mathbf{P}} \quad (5.32)$$

where we have exploited the definition of the free-space propagation constant  $k_0$

$$k_0 := \omega/c = 2\pi/\lambda_{\text{vacuum}} \quad (5.33)$$

and the derivation rule

$$\mathfrak{F} \left[ \frac{d^n a(t)}{dt^n} \right] = (j\omega)^n \mathfrak{F} [a(t)] \quad (5.34)$$

Inserting (5.5), (5.8) and (5.12) in (5.32) yields

$$\begin{aligned}\nabla^2 \tilde{\mathbf{E}} &= -k_0^2 \tilde{\mathbf{E}} - \omega^2 \mu_0 (\tilde{\mathbf{P}}_L + \tilde{\mathbf{P}}_{NL}) = -k_0^2 \tilde{\mathbf{E}} - \omega^2 \mu_0 \varepsilon_0 \tilde{\chi}^{(1)} \tilde{\mathbf{E}} - \omega^2 \mu_0 \tilde{\mathbf{P}}_{NL} \\ &= -\frac{\omega^2}{c^2} (\mathbf{I} + \tilde{\chi}^{(1)}) \tilde{\mathbf{E}} - \omega^2 \mu_0 \tilde{\mathbf{P}}_{NL} = -\frac{\omega^2}{c^2} (\varepsilon_r + \delta\varepsilon_r) \tilde{\mathbf{E}} - \omega^2 \mu_0 \tilde{\mathbf{P}}_{NL}\end{aligned}\quad (5.35)$$

At this point, we have to observe that in all the previous equations  $\mathbf{E}(\mathbf{r}, t)$  is a real-passband signals. Indeed, Maxwell's equations and the polarization vectors have been (tacitly) written for the real-passband signals. Let us recall that a real-passband signal is a signal assuming real values in time, like the actual electric field in a fiber. Thus, its Fourier transform is Hermitian and it has both positive and negative frequency contributions. The analytic signal of a certain real-passband signal is defined as twice the positive-frequency part only of the real-passband signal, i.e.,

$$\tilde{\mathbf{E}}_A(\mathbf{r}, \omega) = 2u(\omega) \tilde{\mathbf{E}}_{RP}(\mathbf{r}, \omega) \quad (5.36)$$

where the subscript  $RP$  has been assigned to the real-passband signal and the subscript  $A$  for the analytic signal and

$$u(\omega) := \begin{cases} 1 & \text{if } \omega \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (5.37)$$

The relation between the two signals can be shown to be

$$\mathbf{E}_{RP}(\mathbf{r}, t) = \Re[\mathbf{E}_A(\mathbf{r}, t)] = \frac{1}{2} (\mathbf{E}_A(\mathbf{r}, t) + \mathbf{E}_A^*(\mathbf{r}, t)) \quad (5.38)$$

Given the analytic signal, the complex envelope or baseband equivalent signal  $\mathbf{E}_C$  can be defined as

$$\tilde{\mathbf{E}}_C(\mathbf{r}, \omega) = \tilde{\mathbf{E}}_A(\mathbf{r}, \omega + \omega_0) \quad \Leftrightarrow \quad \tilde{\mathbf{E}}_A(\mathbf{r}, \omega) = \tilde{\mathbf{E}}_C(\mathbf{r}, \omega - \omega_0) \quad (5.39)$$

where the subscript  $C$  has been used for the complex envelope, and  $\omega_0$  is a conventional frequency, normally the central frequency of the spectrum of  $\mathbf{E}_A(\mathbf{r}, \omega)$ . The complex envelope is often a convenient signal since it is slowly-varying with time being centered at zero frequency, while the analytic signal is fast varying being centered at  $\omega_0$ .

The passages we are going to perform are simpler with the analytic signal for the electric field rather than the real-passband signal. Later we will make us also of the modal amplitudes that are complex envelopes, as already tacitly done in the previous chapters.

For simplicity of notation, in all previous equations the real-passband signal has been indicated as  $\mathbf{E}$  instead of  $\mathbf{E}_{RP}$ , and similarly for  $\mathbf{P}_{NL}$ . For simplicity again,

we are not going to use  $\mathbf{E}_A$  to refer to the analytic signal in the next equations, but just with  $\mathbf{E}$ , and similarly for  $\mathbf{P}_{NL}$ . Confusion should not arise given that in the following we use  $\mathbf{E}$  and  $\mathbf{P}_{NL}$  to refer to the analytic signals, unless otherwise specified.

Hence, (5.35) has to be considered now for analytic signals, i.e., we keep only (twice) the positive frequency components of  $\tilde{\mathbf{E}}$  and  $\tilde{\mathbf{P}}_{NL}$  appearing inside, as for definition of analytic signal. However, the equation is formally the same.

We also resort to the following modal expansion for  $M$  guided modes

$$\tilde{\mathbf{E}}(\mathbf{r}, \omega) = \sum_{k=0}^M \tilde{B}_k(z, \omega) \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k(\omega)z} \quad (5.40)$$

which will allow us to obtain the propagation equations for the slowly-varying excitation coefficients  $\tilde{B}_k$ . Note that the modal amplitudes  $\tilde{B}_k$  are  $z$ -dependent right because of the presence of linear and nonlinear distortion effects.  $\tilde{B}_k$  is said to be slowly-varying in space because the fast space-oscillating term  $e^{-j\beta_k z}$  has been factored out in (5.40). However, it is not slowly-varying in frequency since the term  $e^{-j\omega_0 t}$  has not been factored out yet. That is,  $\tilde{B}_k$  is an analytic signal, not a complex basenband signal. This aspect is going to be addressed later on, when a change of variable is introduced.

As already mentioned in other chapters, the modal set of guided and radiation modes can be shown to constitute a complete set of solutions for the linear problem. That is, if the nonlinear perturbations were not present, the electric field in the fiber resulting from the presence of the linear distortion introduced by  $\delta\varepsilon$  could be exactly described as a superposition of the ideal fiber modes, including both the discrete set of guided modes and the continuum of radiation modes [Mar74, p. 98] [Mar75]. Since in this study we neglect the contribution of radiation modes, the modal expansion is only an approximation for the exact field which can be regarded as accurate if

$$(\delta\varepsilon_r)_{i,j} \ll \varepsilon_r \quad (5.41)$$

Hence, we are assuming both the linear and nonlinear perturbations do not affect the transverse mode profiles.

Similar considerations hold for the nonlinear polarization vector whose perturbations have to be small. These two conditions are consistent with (5.28).

We remember that the Laplacian operator can be written, in a symbolic form, as

$$\nabla^2 = \nabla_T^2 + \frac{\partial^2}{\partial z^2} \quad (5.42)$$

The modal expansion (5.40) and the property (5.42) are inserted in (5.35), leaving

$$\begin{aligned}
 & \left( \nabla_T^2 + \frac{\partial^2}{\partial z^2} \right) \left( \sum_{k=0}^M \tilde{B}_k(z, \omega) \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k(\omega)z} \right) \\
 &= -\frac{\omega^2}{c^2} (\varepsilon_r + \delta\varepsilon_r) \left( \sum_{k=0}^M \tilde{B}_k(z, \omega) \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k(\omega)z} \right) - \omega^2 \mu_0 \tilde{\mathbf{P}}_{NL} \\
 \Rightarrow & \sum_{k=0}^M \tilde{B}_k(z, \omega) \nabla_T^2 \left( \mathbf{F}_k(x, y, \omega_0) \right) e^{-j\beta_k(\omega)z} + \sum_{k=0}^M \mathbf{F}_k(x, y, \omega_0) \frac{\partial^2 \left( \tilde{B}_k(z, \omega) e^{-j\beta_k(\omega)z} \right)}{\partial z^2} \\
 = & -\frac{\omega^2}{c^2} \varepsilon_r \sum_{k=0}^M \tilde{B}_k(z, \omega) \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k(\omega)z} + \frac{\omega^2}{c^2} \delta\varepsilon_r \sum_{k=0}^M \tilde{B}_k(z, \omega) \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k(\omega)z} \\
 & - \omega^2 \mu_0 \tilde{\mathbf{P}}_{NL} \quad (5.43)
 \end{aligned}$$

Let us write the second-order derivative term on the left-hand side of (5.43) as

$$\begin{aligned}
 & \frac{\partial^2 (\tilde{B}_k e^{-j\beta_k z})}{\partial z^2} = \frac{\partial \left( \frac{\partial \tilde{B}_k}{\partial z} e^{-j\beta_k z} - j\beta_k e^{-j\beta_k z} \tilde{B}_k \right)}{\partial z} \\
 = & \frac{\partial^2 \tilde{B}_k}{\partial z^2} e^{-j\beta_k z} + \frac{\partial \tilde{B}_k}{\partial z} (-j\beta_k) e^{-j\beta_k z} - j\beta_k e^{-j\beta_k z} \frac{\partial \tilde{B}_k}{\partial z} + (-j\beta_k)^2 e^{-j\beta_k z} \tilde{B}_k \quad (5.44) \\
 \approx & -2j\beta_k e^{-j\beta_k z} \frac{\partial \tilde{B}_k}{\partial z} - \beta_k^2 e^{-j\beta_k z} \tilde{B}_k
 \end{aligned}$$

where in the last passage we neglected  $\frac{\partial^2 \tilde{B}_k}{\partial z^2}$  because we assume that then linear and nonlinear distortions are little so that the modal envelopes vary slowly with  $z$ . This is referred to as adiabatic approximation [GK00] or slow varying wave approximation [Han21, p. 221] [Agr19, p. 37].

Substituting (5.44) in (5.43) yields

$$\begin{aligned}
 & \sum_{k=0}^M \tilde{B}_k \nabla_T^2 \mathbf{F}_k e^{-j\beta_k z} - \sum_{k=0}^M 2j\beta_k e^{-j\beta_k z} \frac{\partial \tilde{B}_k}{\partial z} \mathbf{F}_k - \sum_{k=0}^M \beta_k^2 e^{-j\beta_k z} \tilde{B}_k \mathbf{F}_k \\
 = & -\frac{\omega^2}{c^2} \varepsilon_r \sum_{k=0}^M \tilde{B}_k \mathbf{F}_k e^{-j\beta_k z} - \frac{\omega^2}{c^2} \delta\varepsilon_r \sum_{k=0}^M \tilde{B}_k \mathbf{F}_k e^{-j\beta_k z} - \omega^2 \mu_0 \tilde{\mathbf{P}}_{NL} \quad (5.45)
 \end{aligned}$$

Observe that the colored terms in (5.45) cancel out each others because they fulfill the relation 2.37 of Ch.2 for ideal modal profiles

Hence, from (5.45) we are left with

$$\begin{aligned}
 - \sum_{k=0}^M 2j\beta_k e^{-j\beta_k z} \frac{\partial \tilde{B}_k}{\partial z} \mathbf{F}_k = \\
 - \frac{\omega^2}{c^2} \delta\varepsilon_r \sum_{k=0}^M \tilde{B}_k \mathbf{F}_k e^{-j\beta_k z} - \omega^2 \mu_0 \tilde{\mathbf{P}}_{NL} \quad (5.46)
 \end{aligned}$$

Notice that in case the linear perturbations were not considered, the term proportional to  $\delta\varepsilon_r$  would not be present anymore.

The previous equation contains on both sides a summation over all the  $M$  modes and the modal profiles  $F_k(x, y, \omega_0)$ , while we would like to retrieve an equation where only one modal amplitude  $B_m$  is figuring. Let us first define the mode area of the  $k$ -th mode as [Agr19, p. 39]

$$D_k = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \|\mathbf{F}_k(x, y, \omega_0)\|^2 dx dy \quad (5.47)$$

To do so, we can introduce a scalar multiplication by  $\mathbf{F}_m^*(x, y, \omega_0)$  on both sides of (5.46), normalize by the modal area  $D_m$  and integrate over a cross-section  $(x, y)$  of infinite area, like this

$$\begin{aligned}
 - \sum_{k=0}^M 2j\beta_k(\omega) e^{-j\beta_k(\omega)z} \frac{\partial \tilde{B}_k}{\partial z}(z, \omega) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\mathbf{F}_m^*(x, y, \omega_0) \cdot \mathbf{F}_k(x, y, \omega_0)}{D_m} dx dy = \\
 - \frac{\omega^2}{c^2} \sum_{k=0}^M \tilde{B}_k(z, \omega) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\mathbf{F}_m^*(x, y, \omega_0) \cdot (\delta\varepsilon_r \mathbf{F}_k(x, y, \omega_0))}{D_m} e^{-j\beta_k(\omega)z} dx dy \\
 - \omega^2 \mu_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\mathbf{F}_m^*(x, y, \omega_0) \cdot \tilde{\mathbf{P}}_{NL}}{D_m} dx dy \quad (5.48)
 \end{aligned}$$

The orthonormality relation among the modes is

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\mathbf{F}_m^*(x, y, \omega_0) \cdot \mathbf{F}_k(x, y, \omega_0)}{D_m} dx dy = \delta_{mk} \quad (5.49)$$

where  $\delta_{mn}$  is Kronecker's delta defined as

$$\delta_{mn} = \begin{cases} 1 & \text{if } m \neq n \\ 0 & \text{if } m = n \end{cases} \quad (5.50)$$

This orthonormality condition (5.49) is less formal than the actual orthonormality condition we should be requiring for generic modes. Eq.(5.49) holds only in the

limit of the weakly-guiding approximation [ASM16; Mar75; PH08; HP12], assuming modes fulfill

$$\mathbf{H}_k = \frac{\mathbf{E}_k \times \hat{z}}{Z_k} \quad (5.51)$$

where with  $Z_k$  it is referred to the impedance of the  $k$ -th mode [ASM16], i.e.:

$$Z_k = \frac{Z_0}{n_k} \quad (5.52)$$

where  $Z_0$  is the impedance of the vacuum and  $n_k$  is the effective refractive index of the  $k$ -th mode.

Exploiting the orthonormality relation (5.49) in (5.48), we get

$$\begin{aligned} -2j\beta_m(\omega)e^{-j\beta_m(\omega)z}\frac{\partial\tilde{B}_m}{\partial z}(z,\omega) + j\sum_{k=0}^M\kappa_{mk}\tilde{B}_k(z,\omega)e^{-j\beta_k(\omega)z} = \\ -\omega^2\mu_0\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\mathbf{F}_m^*(x,y,\omega_0)\cdot\tilde{\mathbf{P}}_{NL}}{D_m}dx dy \end{aligned} \quad (5.53)$$

where the linear mode coupling coefficient has been defined as

$$\kappa_{mk} = \frac{\omega^2}{c^2 D_m} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{F}_m^*(x,y,\omega_0) \cdot (\delta\varepsilon_r \mathbf{F}_k(x,y,\omega_0)) dx dy \quad (5.54)$$

Notice that the presence of the dielectric tensor  $\delta\varepsilon_r$  does not allow to exploit the orthonormality relation for the second term on the r.h.s. of (5.53). Indeed, that term accounts for the anisotropic linear distortion effects.

Eq.(5.53) is manipulated to express it in a “nicer” form

$$\begin{aligned} \frac{\partial\tilde{B}_m}{\partial z}(z,\omega) + j\sum_{k=0}^M\kappa_{mk}\tilde{B}_k(z,\omega)e^{-j(\beta_k(\omega)-\beta_m(\omega))z} = \\ -j\frac{e^{+j\beta_m(\omega)z}}{2\beta_m(\omega)}\omega^2\mu_0\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\mathbf{F}_m^*(x,y,\omega_0)\cdot\tilde{\mathbf{P}}_{NL}}{D_m}dx dy \end{aligned} \quad (5.55)$$

where the linear mode coupling coefficient (5.54) has been redefined as

$$\kappa_{mk} = \frac{\omega^2}{2\beta_m(\omega)c^2 D_m} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{F}_m^*(x,y,\omega_0) \cdot (\delta\varepsilon_r \mathbf{F}_k(x,y,\omega_0)) dx dy \quad (5.56)$$

The prefactor of (5.56) can be manipulated as

$$\frac{\omega^2}{2\beta_m(\omega)c^2 D_m} = \frac{\omega}{2cD_m} \frac{k_0}{\beta_m(\omega)} = \frac{\omega}{2cD_m} \frac{1}{n_m(\omega)} \approx \frac{\omega}{2cn_m(\omega)D_m} \quad (5.57)$$

where it has been exploited  $\omega/c = k_0$ ,  $\beta_0/k_0 = n_m$  is the effective refractive index of the  $m$ -th mode and  $n_m(\omega) \approx n_m(\omega_0)$ .

The prefactor of the nonlinear term in (5.55) can be manipulated as

$$\begin{aligned} -j \frac{e^{+j\beta_m(\omega)z}}{2\beta_m(\omega)} \omega^2 \mu_0 &= -j \frac{e^{+j\beta_m(\omega)z}}{2\beta_m(\omega)} \omega^2 \frac{\mu_0 \varepsilon_0}{\varepsilon_0} = -j \frac{e^{+j\beta_m(\omega)z}}{2\beta_m(\omega)} \omega^2 \frac{1}{c^2 \varepsilon_0} = \\ -j \frac{e^{+j\beta_m(\omega)z}}{2\beta_m(\omega)} \frac{\omega}{c} k_0 \frac{1}{\varepsilon_0} &= -j \frac{e^{+j\beta_m(\omega)z} \omega}{2n_m(\omega) \varepsilon_0 c} \approx -j \frac{e^{+j\beta_m(\omega)z} \omega}{2n_m(\omega_0) \varepsilon_0 c} \end{aligned} \quad (5.58)$$

where it has been used  $1/(\mu_0 \varepsilon_0) = c^2$ ,  $\omega/c = k_0$ ,  $\beta_0/k_0 = n_m$ ,  $n_m(\omega) \approx n_m(\omega_0)$ .

Expressing the prefactor of the nonlinear term of (5.58) as (5.55), yields

$$\begin{aligned} \frac{\partial \tilde{B}_m}{\partial z}(z, \omega) + j \sum_{k=0}^M \kappa_{mk} \tilde{B}_k(z, \omega) e^{-j(\beta_k(\omega) - \beta_m(\omega))z} = \\ -j \frac{e^{+j\beta_m(\omega)z} \omega}{2n_m(\omega_0) \varepsilon_0 c} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\mathbf{F}_m^*(x, y, \omega_0) \cdot \tilde{\mathbf{P}}_{NL}}{D_m} dx dy \end{aligned} \quad (5.59)$$

Expressing the prefactor of the linear coupling coefficient (5.54) as (5.57), yields

$$\kappa_{mk} = \frac{\omega_0}{2cn_m(\omega_0)D_m} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{F}_m^*(x, y, \omega_0) \cdot (\delta\varepsilon_r \mathbf{F}_k(x, y, \omega_0)) dx dy \quad (5.60)$$

Notice that (5.59) and (5.60) are equivalent to (5.55) and (5.56), respectively. The only difference is that the prefactors have been written in another (equivalent) form which is going to be handier for later. Moreover, we have assumed  $\kappa_{mk}$  to be frequency independent since its frequency dependence is only a weak correction to the frequency independent part [MAS12b].

At this point, it is necessary to provide an expression for  $\tilde{\mathbf{P}}_{NL}$  where the dependence on the modal amplitudes is explicit. However, since it is more convenient for us to find such an expression in time domain, we have to transform (5.55) to time domain. To ease this task, we move to a slightly different modal expansion, because the expansion used up to now, (5.40), does not allow for an easy expression in time domain due to the presence of the dispersion term  $e^{-j\beta_k(\omega)z}$ . The time domain expression of the new modal expansion is

$$\mathbf{E}(z, t) = \sum_{k=1}^M A_k(z, t) \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k^{(0)}z} e^{+j\omega_0 t} \quad (5.61)$$

where, besides all the already-introduced symbols,  $A_k$  are the new modal amplitudes and we recall the notation

$$\beta_k^{(0)} \equiv \beta_k(\omega_0) \beta_k^{(n)} \equiv \frac{d^n \beta_k(\omega_0)}{d\omega^n}(\omega = \omega_0), \quad \text{for } n = 0, 1, 2, \dots, N_D \quad (5.62)$$

where  $N_D$  is called dispersion order.

Note that the modal amplitudes  $A_k$  are  $z$ -dependent because of the presence of linear and nonlinear distortion effects.  $A_k$  is said to be slowly-varying in time because the fast time-oscillating term  $e^{+j\omega_0 t}$  has been factored out since the beginning since the signal is the equivalent baseband, and in space since the fast space-oscillating term  $e^{-j\beta_k z}$  has been factored out in (5.61).

In order to express (5.55) with respect to  $A_k$ , we need to find the relation between  $A_k$  and  $B_k$ , which is easier to be obtained in frequency. Hence, let us transform (5.61) to frequency domain, i.e.,

$$\tilde{\mathbf{E}}(z, \omega) = \sum_{k=1}^M \tilde{A}_k(z, \omega - \omega_0) \mathbf{F}_k(x, y, \omega_0) e^{-j\beta_k^{(0)} z} \quad (5.63)$$

where it has been used the rule

$$\mathfrak{F} \left[ a(t) e^{-j2\pi f_0 t} \right] (f) = \mathfrak{F} \left[ a(t) \right] (f - f_0) \quad (5.64)$$

Comparing (5.40) and (5.63), we get

$$\tilde{B}_k(z, \omega) = \tilde{A}_k(z, \omega - \omega_0) e^{+j(\beta_k(\omega) - \beta_k^{(0)})z} \quad (5.65)$$

Taking the derivative with respect to  $z$  yields

$$\begin{aligned} \frac{\partial \tilde{B}_k}{\partial z}(z, \omega) &= \frac{\partial (\tilde{A}_k(z, \omega - \omega_0) e^{+j(\beta_k(\omega) - \beta_k^{(0)})z})}{\partial z} \\ &= \frac{\partial \tilde{A}_k(z, \omega - \omega_0)}{\partial z} e^{+j(\beta_k(\omega) - \beta_k^{(0)})z} + j(\beta_k(\omega) - \beta_k^{(0)}) e^{+j(\beta_k(\omega) - \beta_k^{(0)})z} \tilde{A}_k(z, \omega - \omega_0) \end{aligned} \quad (5.66)$$

Hence, substituting (5.65) and (5.66) in (5.59),

$$\begin{aligned} \frac{\partial \tilde{A}_m(z, \Delta\omega)}{\partial z} e^{+j(\beta_m(\omega) - \beta_m^{(0)})z} + j(\beta_m(\omega) - \beta_m^{(0)}) e^{+j(\beta_m(\omega) - \beta_m^{(0)})z} \tilde{A}_m(z, \Delta\omega) \\ + j \sum_{k=0}^M \kappa_{mk} \tilde{A}_k(z, \Delta\omega) e^{+j(\beta_k(\omega) - \beta_k^{(0)})z} e^{-j(\beta_k(\omega) - \beta_m(\omega))z} \\ = -j \frac{e^{+j\beta_m(\omega)z} (\omega_0 + \Delta\omega)}{2n_m(\omega_0)\varepsilon_0 c} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\mathbf{F}_m^*(x, y, \omega_0) \cdot \tilde{\mathbf{P}}_{NL}}{D_m} dx dy \end{aligned} \quad (5.67)$$

where we defined

$$\Delta\omega := \omega - \omega_0 \quad (5.68)$$

Remind that the propagation constant  $\beta_k(\omega)$  of the  $k$ -th mode can be expanded in a Taylor series around  $\omega_0$  as

$$\beta_m(\omega) = \beta_m^{(0)} + \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \Delta\omega \quad (5.69)$$

Diving by  $e^{+j(\beta_m(\omega) - \beta_m^{(0)})z}$  and exploiting the Taylor series (5.69), we obtain from (5.67)

$$\begin{aligned} \frac{\partial \tilde{A}_m(z, \Delta\omega)}{\partial z} + j \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \Delta\omega \tilde{A}_m(z, \Delta\omega) + j \sum_{k=0}^M \kappa_{mk} \tilde{A}_k(z, \Delta\omega) e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z} \\ = -j \frac{e^{+j\beta_m^{(0)}z} (\omega_0 + \Delta\omega)}{2n_m(\omega_0)\varepsilon_0 c} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\mathbf{F}_m^*(x, y, \omega_0) \cdot \tilde{\mathbf{P}}_{NL}}{D_m} dx dy \end{aligned} \quad (5.70)$$

The previous equation is ready to be transformed to time domain. However, the frequency variable in this case is considered to be  $\Delta\omega$  which eases the calculations since the equation is mostly written with respect to  $\Delta\omega$ . That is,

$$\mathfrak{F}^{-1} \left[ A(\Delta\omega) \right] (t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A(\Delta\omega) e^{+j\Delta\omega t} d\Delta\omega \quad (5.71)$$

We just recall that  $\tilde{\mathbf{P}}_{NL}(\mathbf{r}, \omega)$  is an analytic passband signal and it holds the property

$$\begin{aligned} \tilde{\mathbf{P}}_{NL}(\mathbf{r}, \omega) &= \tilde{\mathbf{P}}_{NL}(\mathbf{r}, \omega_0 + \Delta\omega) \\ &\downarrow \\ \mathfrak{F}^{-1} \left[ \tilde{\mathbf{P}}_{NL}(\omega_0 + \Delta\omega, z) \right] &= \mathbf{P}_{NL}(\mathbf{r}, t) e^{-j\omega_0 t} \end{aligned} \quad (5.72)$$

With the help of (5.71), (5.72) and the derivation rule (5.34) applied to the second term on the l.h.s. of (5.70), we transform (5.70) to time domain

$$\begin{aligned} \frac{\partial A_m(z, t)}{\partial z} + j \left( \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \frac{\partial^k}{\partial t^k} \right) A_m(z, t) + j \sum_{k=0}^M \kappa_{mk} A_k(z, t) e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z} \\ = -j \frac{e^{+j\beta_m^{(0)}z} \omega_0 \left( 1 + \frac{1}{j\omega_0} \frac{\partial}{\partial t} \right)}{2n_m(\omega_0)\varepsilon_0 c} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\mathbf{F}_m^*(x, y, \omega_0) \cdot \mathbf{P}_{NL}}{D_m} dx dy \end{aligned} \quad (5.73)$$

We now have to provide an explicit expression for  $\mathbf{P}_{NL}(\mathbf{r}, t)$ , which we repeat is an analytic signal. The definition of nonlinear polarization vector, (5.9), is however for the real-passband version of the vector

$$\mathbf{P}_{NL,RP}(\mathbf{r}, t) := \varepsilon_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \chi^{(3)}(t_1 - t, t_2 - t, t_3 - t) \mathbf{E}_{RP}(t_1) \mathbf{E}_{RP}(t_2) \mathbf{E}_{RP}(t_3) dt_1 dt_2 dt_3 \quad (5.74)$$

where, conversely to (5.9), we have emphasized that the vectors are real-passband signals through the subscript  $RP$ .

The third-order nonlinear susceptibility tensor can be shown to assume a simplified form in silica fibers [Agr19, p. 687]

$$\chi^{(3)}(t_1 - t, t_2 - t, t_3 - t) = \chi^{(K)} \delta(t - t_1) \delta(t - t_2) \delta(t - t_3) + \chi^{(R)}(t_1 - t_2) \delta(t - t_1) \delta(t_2 - t_3) \quad (5.75)$$

where  $\delta(t - t_i)$  is the Dirac's delta.

The term proportional to  $\chi^{(K)}$  in (5.75), Kerr constant, accounts for the (approximately) instantaneous response of the electrons of the medium, which is referred to as Kerr effect. The term proportional to  $\chi^{(R)}$  accounts for the response of nuclei of the material, which being heavier, have a noninstantaneous response [Agr19, p. 687] [RP05; ASM16; HP12; Wri+17].

Hence, it makes sense to separate  $\mathbf{P}_{NL}$  in his two Kerr and Raman contributions, as

$$\mathbf{P}_{NL,RP} = \mathbf{P}_{NL}^K + \mathbf{P}_{NL,RP}^R \quad (5.76)$$

where, inserting (5.75) in (5.9), [Agr19, p. 687]

$$\mathbf{P}_{NL,RP}^K(\mathbf{r}, t) := \varepsilon_0 \chi^{(K)} \mathbf{E}_{RP}(\mathbf{r}, t) \mathbf{E}_{RP}(\mathbf{r}, t) \mathbf{E}_{RP}(\mathbf{r}, t) \quad (5.77)$$

and [Agr19, p. 687]

$$\mathbf{P}_{NL,RP}^R := \varepsilon_0 \mathbf{E}_{RP}(t) \int_{-\infty}^{+\infty} \chi^{(R)}(t - t_1) \mathbf{E}_{RP}(t_1) \mathbf{E}_{RP}(t_1) dt_1 \quad (5.78)$$

Owing to the isotropicity of the silica, the elements of the tensor  $\chi^{(K)}$  are (approximately) dependent on one single value, as

$$\chi_{ijkl}^{(K)} = \frac{\sigma_K}{3} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \quad (5.79)$$

while the elements of the tensor  $\chi^{(R)}$  reduce to be combinations of only two independent contributions, like [RP05] [Agr19, p. 687]

$$\chi_{ijkl}^{(R)} = \sigma_R [f_a h_a(t) \delta_{ij} \delta_{kl} + \frac{1}{2} f_b h_b(t) (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})] \quad (5.80)$$

Exploiting the previous two relations, after few algebraic passages, (5.77) becomes [RP05] [Agr19, p. 687]

$$\mathbf{P}_{NL,RP}^K(\mathbf{r}, t) := \varepsilon_0 \sigma_K (\mathbf{E}_{RP}(\mathbf{r}, t) \cdot \mathbf{E}_{RP}(\mathbf{r}, t)) \mathbf{E}_{RP}(\mathbf{r}, t) \quad (5.81)$$

and (5.78), after few algebraic passages, becomes [RP05] [Agr19, p. 687]

$$\begin{aligned} \mathbf{P}_{NL,RP}^R(\mathbf{r}, t) := & \varepsilon_0 \sigma_R \left( \mathbf{E}_{RP}(\mathbf{r}, t) \int_{-\infty}^{+\infty} f_a h_a(t - t_1) (\mathbf{E}_{RP}(\mathbf{r}, t_1) \cdot \mathbf{E}_{RP}(\mathbf{r}, t_1)) dt_1 \right. \\ & \left. + \int_{-\infty}^{-\infty} f_b h_b(t - t_1) (\mathbf{E}_{RP}(\mathbf{r}, t_1) \cdot \mathbf{E}_{RP}(\mathbf{r}, t_1)) \mathbf{E}_{RP}(\mathbf{r}, t_1) dt_1 \right) \quad (5.82) \end{aligned}$$

where

$$\begin{aligned} f_a + f_b &= 1 \\ \int_{-\infty}^{+\infty} h_a(t) dt &= \int_{-\infty}^{+\infty} h_b(t) dt = 1 \end{aligned}$$

We recall that the relation between the real passband signal  $\mathbf{E}_{RP}$  and the analytic signal  $\mathbf{E}$  is

$$\mathbf{E}_{RP} = \frac{1}{2} (\mathbf{E} + \text{c.c.}) \quad (5.83)$$

Substituting the previous equation in (5.81) produces terms at  $\pm\omega_0$  and  $\pm 3\omega_0$ . We keep only the positive frequencies of the result since, as said, we are looking for the analytic version,  $\mathbf{P}_{NL}$ , of the polarization vector  $\mathbf{P}_{NL,RP}$ . Moreover, the terms at  $\pm 3\omega_0$  can be shown to be strongly non phase matched and so negligible. Hence, the analytic part of (5.81) at frequency  $+\omega_0$  is [Agr19, p. 688]

$$\mathbf{P}_{NL}^K(z, t) = \frac{2}{8} \sigma_K \varepsilon_0 (2 \|\mathbf{E}\|^2 \mathbf{E} + (\mathbf{E} \cdot \mathbf{E}) \mathbf{E}) \quad (5.84)$$

Similarly for (5.82), if we keep only the terms at  $+\omega_0^*$ , we arrive to [ASM16]

$$\mathbf{P}_{NL}^R(z, t) = \left( \frac{\sigma_R \varepsilon_0}{4} \right) 2 \mathbf{E}(\mathbf{r}, t) \int_{-\infty}^{+\infty} h(t - t_1) (\mathbf{E}(\mathbf{r}, t_1) \cdot \mathbf{E}^*(\mathbf{r}, t_1)) dt_1 \quad (5.85)$$

Now, it is possible to substitute (5.61) in  $\mathbf{P}_{NL}^K(z, t)$  and  $\mathbf{P}_{NL}^R(z, t)$ . For the former we obtain

$$\mathbf{P}_{NL}^K(z, t) = \frac{\varepsilon_0 \sigma_K}{4} e^{+j\omega_0 t} \sum_j \sum_k \sum_l A_j A_k A_l^* e^{+j(-\beta_j^{(0)} - \beta_k^{(0)} + \beta_l^{(0)})z} [2(\mathbf{F}_j \cdot \mathbf{F}_l^*) \cdot \mathbf{F}_k + (\mathbf{F}_j \cdot \mathbf{F}_k) \mathbf{F}_l^*] \quad (5.86)$$

---

\*This corresponds to keeping the first term of Eq.B.9 of [Agr19, p. 688] setting  $f_a = 1$ ,  $f_b = 0$

Inserting (5.61) in  $\mathbf{P}_{NL}^R(z, t)$  we get

$$\mathbf{P}_{NL}^R(z, t) = \left(\frac{\varepsilon_0 \sigma_R}{2}\right) e^{+j\omega_0 t} \sum_k \sum_j \sum_l \mathbf{F}_k(\mathbf{F}_j \cdot \mathbf{F}_l^*) e^{-j(\beta_j^{(0)} + \beta_k^{(0)} - \beta_l^{(0)})z} A_k(z, t) \cdot \int_{-\infty}^{+\infty} h(t - t_1) A_j(t_1) A_l^*(t_1) dt_1 \quad (5.87)$$

Notice the presence of the term  $e^{+j\omega_0 t}$  in both the previous equations because only the contributes at frequency  $\omega_0$  have been saved. Observe also that both polarization vectors contain a vector term  $\mathbf{F}_k(\mathbf{F}_j \cdot \mathbf{F}_l^*)$ .

Plugging (5.86) and (5.87) into (5.73), via (5.76), results immediately in the following propagation equation

$$\begin{aligned} \frac{\partial A_m(z, t)}{\partial z} + j \left( \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \frac{\partial^k}{\partial t^k} \right) A_m(z, t) + j \sum_{k=0}^M \kappa_{mk} A_k(z, t) e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z} = \\ - j\omega_0 \left( 1 + \frac{1}{j\omega_0} \frac{\partial}{\partial t} \right) \frac{1}{2n_m(\omega_0) \varepsilon_0 c D_m} \sum_j \sum_k \sum_l \left( \frac{\varepsilon_0 \sigma_K}{4} C'_{jklm} A_j A_k A_l^* \right. \\ \left. + \frac{\varepsilon_0 \sigma_R}{2} Q'_{jklm} A_k \int_{-\infty}^{+\infty} h(t - t_1) A_j A_l^* \right) e^{-j\Delta\beta_{jklm}z} \end{aligned} \quad (5.88)$$

where the linear coupling coefficients are given by (5.54) and

$$C'_{jklm} = 2Q'_{jklm} + R'_{jklm} \quad (5.89a)$$

$$Q'_{jklm} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\mathbf{F}_j \cdot \mathbf{F}_l^*)(\mathbf{F}_k \cdot \mathbf{F}_m^*) dx dy \quad (5.89b)$$

$$R'_{jklm} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [(\mathbf{F}_j \cdot \mathbf{F}_k)(\mathbf{F}_l^* \cdot \mathbf{F}_m^*)] dx dy \quad (5.89c)$$

$$\Delta\beta_{jklm} = -\beta_m^{(0)} + \beta_j^{(0)} + \beta_k^{(0)} - \beta_l^{(0)} \quad (5.89d)$$

Notice the exchange of the double integrals with the summations moving from (5.73) to (5.88).

The derivation of the GMMNLSE could stop at (5.88) since all the information about the signal propagation is already present in (5.88). However, we perform few more passages to rearrange the equation in a form closer to the common literature notation and to clarify the idea behind the choice of a certain normalization. The target is to reach the same form as Antonelli et al. [ASM16].

## 5.2 Manipulation of the GMMNLSE

Eq.(5.88) has been expressed referring to the modal expansion (5.61), which assumes the modal profiles  $\mathbf{F}_k$  to be adimensional and so, given that the electric field is in units of V/m in the S.I. (International System of Units), the modal amplitudes are also in units of V/m. However, it is common practice to handle modal amplitudes whose modulus square is the power of the considered mode in units of  $\sqrt{W}$ . In order to achieve it, it is necessary to consider a slightly different modal expansion from (5.61). In particular, we consider

$$\mathbf{E}(z, t) = \sum_{k=1}^M \frac{A_k(z, t) \mathbf{F}_k(x, y, \omega_0)}{N_k} e^{-j\beta_k^{(0)} z} e^{+j\omega_0 t} \quad (5.90)$$

where  $N_k$  is a normalization constant in units of  $\frac{m}{\sqrt{\Omega}}$  and its square value is defined as

$$N_k^2 := \frac{D_k}{2Z_k} \quad (5.91)$$

where  $D_k$  and  $Z_k$  are the mode area (5.47) in  $m^2$  and the impedance (5.52) in  $\Omega$ , respectively, of the  $k$ -th mode.

Let us define the modal electric field of the  $k$ -th mode as

$$\mathbf{E}_k = \frac{A_k(z, t) \mathbf{F}_k(x, y, \omega_0)}{N_k} e^{-j\beta_k^{(0)} z} e^{+j\omega_0 t} \quad (5.92)$$

and the modal power  $P_k$  as the power of the  $k$ -th mode. Within the weak-guidance and narrowband signal approximation, i.e., the signal frequency is much lower than the carrier, the power of the  $k$ -th mode is

$$P_k(z, t) = \int \int \frac{\|\mathbf{E}_k(z, t)\|^2}{2Z_k} dx dy \quad (5.93)$$

The previous formula can be derived following the ideas in Appendix B of [ASM16].

Substituting (5.92) in (5.93)

$$\begin{aligned} P_k(z, t) &= \int \int \frac{\|A_k(z, t) \mathbf{F}_k(x, y, \omega_0)\|^2}{2Z_k N_k^2} dx dy = \|A_k(z, t)\|^2 \int \int \frac{\|\mathbf{F}_k(x, y, \omega_0)\|^2}{2Z_k N_k^2} dx dy \\ &\quad \Downarrow \\ P_k(z, t) &= \|A_k(z, t)\|^2 \int \int \frac{\|\mathbf{F}_k(x, y, \omega_0)\|^2}{2Z_k \frac{D_k}{2Z_k}} dx dy = \|A_k(z, t)\|^2 \end{aligned} \quad (5.94)$$

where in the last passage it has been exploited also the Definition (5.47) of  $D_k$ . As we wanted to show, the modal expansion (5.92) allows to have as  $\|A_k\|^2(z, t)$  the power carried over a single mode in units of Watts [PH08; ASM16] [Agr19, p. 39].

In truth, (5.91) makes sense only in the limit of the weakly-guiding approximation [PH08; HP12; ASM16] because the modes can no more be approximated as plane waves and the orthogonality relation becomes more complex than (5.49). The more general normalization coefficient  $N_k$  can be found in [ASM16]. However, for our purposes (5.91) is enough.

If the new modal expansion (5.92) is employed, the GMMNLSE (5.88) assumes a slightly different expression. In order to reduce possible misunderstandings, let us rename the modal amplitudes of the previous expansion (5.61) and of the GMMNLSE (5.88) as  $A'_m$  and keep the symbol  $A_m$  for the amplitudes of the new expansion (5.92). Comparing the two expansions (5.61) and (5.92), the relation between the amplitudes is immediate

$$A'_m = \frac{A_m}{N_m} \quad (5.95)$$

Performing the change of variable (5.95) in (5.88) yields (in red the new terms, in green some old terms used later)

$$\begin{aligned} & \frac{\partial A_m}{\partial z} + j \left( \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \frac{\partial^k}{\partial t^k} \right) A_m + j \sum_{k=0}^M \frac{N_m}{N_k} \kappa_{mk} A_k(z, t) e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z} \\ & = -j\omega_0 \left( 1 + \frac{1}{j\omega_0} \frac{\partial}{\partial t} \right) \frac{1}{8n_m(\omega_0)cD_m} \sum_j \sum_k \sum_l \frac{N_m}{N_j N_k N_l} \left( \sigma_K C'_{jklm} A_j A_k A_l^* \right. \\ & \left. + 2\sigma_R Q'_{jklm} A_j \int_{-\infty}^{+\infty} h(t - t_1) A_k A_l^* \right) e^{-j\Delta\beta_{jklm}z} \end{aligned} \quad (5.96)$$

Including the factor  $N_m/N_k$  in the previous definition (5.60) of linear coupling coefficients, we get that the prefactor of the new coupling coefficient is

$$\frac{N_m}{N_k} \frac{\omega_0}{2cn_m(\omega_0)D_m} = \frac{\sqrt{D_m n_m}}{\sqrt{D_k n_k}} \frac{\omega_0}{2cn_m(\omega_0)D_m} = \frac{\omega_0}{2c\sqrt{D_m D_k n_m n_k}} \quad (5.97)$$

where the definition of mode area (5.47) has been exploited.

The colored (green and red) factor appearing in the nonlinear part of (5.96) can be manipulated as

$$\begin{aligned} & \frac{\omega_0}{8n_m c D_m} \frac{1}{N_j N_k N_l} \frac{N_m}{N_j N_k N_l} = \frac{\omega_0}{8n_m c} \frac{1}{2Z_m N_m^2} \frac{1}{N_j N_k N_l} \frac{N_m}{N_j N_k N_l} = \frac{\omega_0}{16} \frac{1}{n_m c} \frac{1}{Z_m} \frac{1}{N_m N_j N_k N_l} \\ & = \frac{\omega_0}{16} \frac{\sqrt{\varepsilon_0 \mu_0}}{n_m} \frac{1}{Z_m} \frac{1}{N_m N_j N_k N_l} = \frac{\omega_0}{16} \varepsilon_0 Z_m \frac{1}{Z_m} \frac{1}{N_m N_j N_k N_l} = \frac{\omega_0 \varepsilon_0}{16} \frac{1}{N_m N_j N_k N_l} \end{aligned} \quad (5.98)$$

Hence, the colored term of (5.88) becomes

$$\frac{\omega_0 \varepsilon_0}{16} \frac{1}{N_m N_j N_k N_l} \quad (5.99)$$

Let us define the nonlinear parameter [ASM16] [Agr19, p. 39]

$$\gamma := \frac{\omega_0 n_2}{c A_{\text{eff}}} = \frac{\omega_0}{c A_{\text{eff}}} \frac{3 \chi^{(3)}}{4 n_{\text{eff}}} Z_1 \quad (5.100)$$

where [ASM16]

$$n_2 := \frac{3 \chi^{(3)}}{4 n_{\text{eff}}} Z_1 \quad (5.101)$$

and

$$\chi^{(3)} = \sigma_K + \frac{2}{3} \sigma_R \quad (5.102)$$

$$n_{\text{eff}} \equiv n_1 \quad (5.103)$$

$$A_{\text{eff}} = \frac{\int \int \left( \|\mathbf{F}_1^2(x, y, \omega_0)\| \right)^2 dx dy}{\int \int \mathbf{F}_1^4(x, y, \omega_0) dx dy} \quad (5.104)$$

The nonlinear coefficients  $\gamma$  and  $n_2$  are the same as the single mode nonlinear propagation equation where the nonlinear term is given by [ASM16]

$$- \gamma \frac{|E|^2}{2Z_1} E \quad (5.105)$$

or by

$$- \gamma |E|^2 E \quad (5.106)$$

depending on whether the term  $2Z_1$  is included or not in  $n_2$  [ASM16]. Observe that a more common definition of  $n_2$  considers the core refractive index  $n$  at the denominator [Agr19, p. 39], instead of the effective refractive index  $n_{\text{eff}}$  of the fundamental mode as for (5.101). Such definition is reasonable since one desires the nonlinear parameter  $n_2$  to be linked to the fiber material only, not the specific fiber geometry. Yet, given our derivation it makes more sense to follow the definition (5.101) we adopted, which is the same of Antonelli et al. [ASM16].

Given the previous definitions, the nonlinear term (5.99) can be expressed as

$$\begin{aligned} \frac{\omega_0 \varepsilon_0}{16} \frac{1}{N_m N_j N_k N_l} &= \gamma \left( \frac{A_{\text{eff}} n_{\text{eff}}}{Z_1} \frac{1}{\chi^{(3)}} \frac{4}{3} c \right) \frac{\varepsilon_0}{16} \frac{1}{N_m N_j N_k N_l} = \gamma \frac{A_{\text{eff}} n_{\text{eff}}^2}{12 Z_0} \frac{1}{\chi^{(3)}} \frac{c}{\varepsilon_0} \frac{1}{N_m N_j N_k N_l} \\ &= \gamma \frac{A_{\text{eff}} n_{\text{eff}}^2}{12 Z_0} \frac{1}{\chi^{(3)}} \frac{1}{Z_0} \frac{1}{N_m N_j N_k N_l} = \gamma \frac{A_{\text{eff}} n_{\text{eff}}^2}{12 Z_0^2} \frac{1}{\chi^{(3)}} \frac{1}{N_m N_j N_k N_l} \end{aligned} \quad (5.107)$$

Thus, substituting the nonlinear colored terms in (5.96) with (5.107) provides the final form of the GMMNLSE, when loss is neglected, i.e.,

$$\begin{aligned} \frac{\partial A_m}{\partial z} + j \left( \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \frac{\partial^k}{\partial t^k} \right) A_m + j \sum_{k=0}^M \kappa_{mk} A_k e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z} = \\ - j \left( 1 + \frac{1}{j\omega_0} \frac{\partial}{\partial t} \right) \gamma \sum_j \sum_k \sum_l \left( C_{jklm} A_j A_k A_l^* + H_{jklm} A_k \int_{-\infty}^{+\infty} h(t - t_1) A_j A_l^* dt_1 \right) e^{-j\Delta\beta_{jklm}z} \end{aligned} \quad (5.108)$$

where the linear coupling coefficient has been redefined as explained in relation to (5.97), i.e.,

$$\kappa_{mk} = \frac{\omega_0}{2c\sqrt{D_m D_k n_m n_k}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{F}_m^*(x, y, \omega_0) \cdot (\delta\varepsilon_r \mathbf{F}_k(x, y, \omega_0)) dx dy \quad (5.109)$$

and the other parameters are

$$C_{jklm} = \frac{A_{\text{eff}} n_{\text{eff}}^2}{12Z_0^2} (1 - f_R) (2Q_{jklm} + R_{jklm}) \quad (5.110)$$

$$H_{jklm} = \frac{A_{\text{eff}} n_{\text{eff}}^2}{12Z_0^2} (f_R) (3Q_{jklm}) \quad (5.111)$$

$$Q_{jklm} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{[(\mathbf{F}_j \cdot \mathbf{F}_l^*)(\mathbf{F}_k \cdot \mathbf{F}_m^*)]}{N_j N_k N_l N_m} dx dy \quad (5.112)$$

$$R_{jklm} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{[(\mathbf{F}_j \cdot \mathbf{F}_k)(\mathbf{F}_j^* \cdot \mathbf{F}_m^*)]}{N_j N_k N_l N_m} dx dy \quad (5.113)$$

$$f_R = \frac{\frac{2}{3}\sigma_R}{\chi^{(3)}} \quad (5.114)$$

$$(1 - f_R) = \frac{\sigma_K}{\chi^{(3)}} \quad (5.115)$$

$$\Delta\beta_{jklm} = -\beta_m^{(0)} + \beta_j^{(0)} + \beta_k^{(0)} - \beta_l^{(0)} \quad (5.116)$$

Observe that  $f_R$  is the relative intensity of the Raman parameter relative to the sum of the Raman and Kerr parameter. In other words, it is the Raman contribution to the Kerr effect.

Notice that the GMMNLSE in the form (5.108) with the parameters (5.110) is very close in content and notation to the one shown in [ASM16], besides an approximation of the Raman term which is going to be discussed soon and besides the fact that the authors does not consider a physical model for linear coupling,

conversely from our approach through (5.109). The GMMNLSE of [PH08], one of the first works to derive it, is also similar, with the differences that no linear mode coupling is modeled, while some more nonlinear contributions of the Raman term (5.82), are considered in that paper. Agrawal [Agr19], on which the first part of the derivation is based, does not provide the equation with all the terms (5.108), but several versions considering the separate effects and, anyhow, it paves down the way to derive this equation. [HP12] and [Wri+17] show a similar version of the equation, but without considering linear coupling and with the latter restricting to the single polarization case. All the above mentioned works adopt the convention for the phase retardation opposite to us, i.e., they treat a forward-propagating mode with  $e^{+j\beta_k z}$ , resulting in different signs in some terms of the equation. However, this does not affect the interpretation of the phenomena since, when the real passband electric field is considered, the description is the same.

### 5.3 Brief Description of the GMMNLSE Terms

The terms on the l.h.s. of the GMMNLSE account all for the linear effects. The term proportional to the time derivatives accounts for the mode-independent dispersion terms up to the desired order, often two. E.g., the first-order derivative is responsible for the propagation delay of a mode, the second for chromatic dispersion.

The third term on term on the l.h.s. of the equation models linear mode coupling, whose discussion happened in the previous chapters. The presence of the phase-mismatch between a pair of modes determines the presence of a fast-oscillating (with  $z$ ) term  $e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z}$  which tends to average out the coupling term and, thus, has the effect of reducing the actual coupling strength between the modes, as discussed in Ch.4.

On the r.h.s. of the equation there are the nonlinear terms.  $\frac{1}{j\omega_0} \frac{\partial}{\partial t}$  is called self-steepening and it generally does not have a great impact on the signal propagation [Wri+17].

The nonlinear term proportional to  $C_{jklm}$  is responsible for Kerr effect. Particular cases of it are the self-phase modulation (SPM), proportional to  $|A_m|^2 A_m$ , where the phase of  $A_m$  is modulated by the power of  $A_m$  itself; the cross-phase modulation (XPM), proportional to  $|A_j|^2 A_m$ , where the phase of  $A_m$  is modulated by the power of another modal amplitude  $A_j$ ; the more complex four-wave mixing (FWM) where not only the phase of  $A_m$  is modulated, but there is an exchange of power between the modes given the generality of the terms  $A_j A_k A_l^*$ .

The nonlinear term proportional to  $H_{jklm}$  is responsible for Raman effect, which determines a transfer of power among the modes and with the material, which implies a loss of energy [Agr19, p. 16].

Similar to the linear coupling, observe the presence of the phase-mismatch  $\Delta\beta$  for the nonlinear coupling which determines the actual intensity of the coupling terms.

In the following we consider two final aspects: the approximation of the Raman term for signals with band below 10THz, and the inclusion of fiber loss.

## 5.4 Modelling the Raman Impulse Response

The Raman impulse response  $h(t)$  defining the Raman part of the nonlinear polarization vector (5.85) and entering the GMMNLSE (5.108), has to be experimentally assessed. However, some simplified models requiring few measured parameters are possible [Agr19, Sec.2.3.3] [RP05]. For example, a basic one is the damped harmonic oscillator, whose impulse response is [Agr19, p. 2.3.39] [RP05]

$$h(t) = \frac{\tau_1^2 + \tau_2^2}{\tau_1 \tau_2} e^{\frac{-t}{\tau_2}} \sin\left(\frac{t}{\tau_1}\right) \quad (5.117)$$

where  $\tau_1$  and  $\tau_2$  are two experimental parameters. For silica fibers,  $\tau_1 = 12$  fs and  $\tau_2 = 32$  fs; for germanosilicate fibers,  $\tau_1 = 12$  fs and  $\tau_2 = 83$  fs [RP05; Agr19].

The impulse and frequency responses are depicted in Fig.5.1.

From the latter is evident that the model is not too accurate in representing the typical Raman gain curves (see, e.g., [Agr19, Fig.2.2]). At the same time, the Raman response has a bandwidth of approximately 15 THz, which means that the model is overcomplicated for bandwidth lower than, at least, few THz. Moreover, if such model is employed for simulations, a simulation bandwidth of at least 30 THz is needed in simulations, otherwise strong aliasing occurs. For signals of bandwidth much lower than few THz, a huge oversampling factor would be required slowing down the computations. Hence, in this circumstances another simpler model is more suited. It is based on approximating the Raman frequency response with a Taylor expansion of the first order [ASM16; Was02] [Agr19, p. 44]

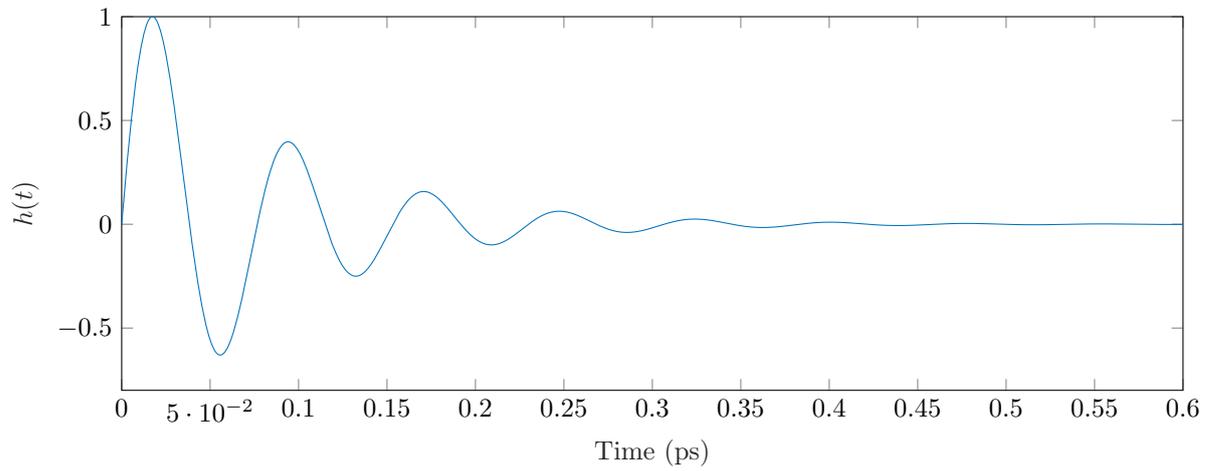
$$\tilde{h}(\Delta\omega) \approx 1 - jT_R\Delta\omega \quad (5.118)$$

where

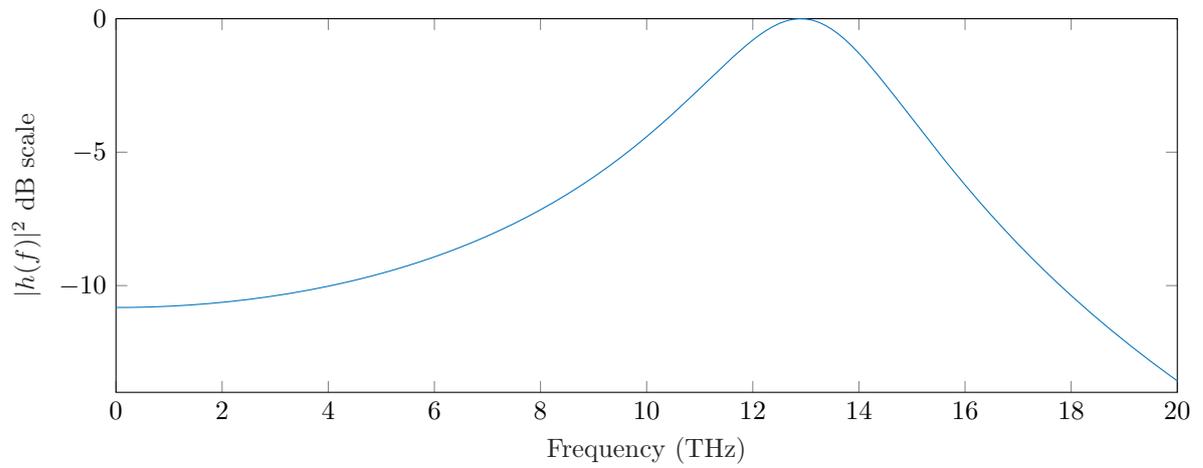
$$T_R := -\frac{1}{j} \left( \frac{dh}{d\omega} \right)_0 = - \int_{-\infty}^{+\infty} th(t) dt \quad (5.119)$$

is the Raman time constant. From a physical perspective, it represents the Raman gain around the central frequency [Was02]. Notice that (5.119) has first summand equal to one since the integral of  $h(t)$  has been normalized to be unity, as required by (5.1). Inverse transforming (5.118) we obtain the Raman impulse response, symbolically, as

$$h(t) = \delta(t) - T_R \frac{\partial}{\partial t} \quad (5.120)$$



(a) Normalized impulse response.



(b) Normalized frequency response.

Figure 5.1: Oscillator model (5.117) with  $\tau_1 = 12$  fs and  $\tau_2 = 83$  fs.

Exploiting (5.118), the argument of the triple summation on the r.h.s. of the GMMNLSE (5.108) becomes

$$\begin{aligned}
 & C_{jklm}A_jA_kA_l^* + H_{jklm}A_k \int_{-\infty}^{+\infty} h(t-t_1)A_jA_l^* dt_1 \\
 &= (C_{jklm} + H_{jklm})A_jA_kA_l^* - T_R H_{jklm}A_k \frac{\partial(A_jA_l^*)}{\partial t} \\
 &= \tilde{C}_{jklm}A_jA_kA_l^* - T_R H_{jklm}A_k \frac{\partial(A_jA_l^*)}{\partial t}
 \end{aligned} \tag{5.121}$$

where

$$\tilde{C}_{jklm} := C_{jklm} + H_{jklm} = \frac{A_{\text{eff}}n_{\text{eff}}^2}{12Z_0^2} \left[ (1-f_R)(2Q_{jklm} + R_{jklm}) + 3f_R Q_{jklm} \right] \tag{5.122}$$

Observe that (5.122) indicates that the instantaneous part of the Raman effect has a similar contribution to the Kerr effect, through  $f_R$ .

## 5.5 Inclusion of Loss

The last missing piece of our interest of the GMMNLSE is the fiber loss, which is not negligible for the fiber lengths typical of long-haul communication systems. For simplicity, we assume it to be mode and frequency independent.

Let  $\alpha$  be the loss coefficient in units of 1/m and let  $\alpha_{dB}$  be the loss in dB/m, whose typical value in current communication systems is around 0.20 dB/km. The loss coefficient  $\alpha$  is defined relative to the signal power as [Agr19, Eq.1.2.3]

$$P(z) = e^{-\alpha z} P(z=0) \tag{5.123}$$

where, thanks to our expansion (5.92) for which the power of a mode is the amplitude square of its modal coefficient (5.2),

$$P(z, t) = \sum_{k=1}^M P_k(z, t) = \sum_{k=1}^M |A_k(z, t)|^2 \tag{5.124}$$

Notice that, being  $e^{-\alpha z}$  the loss factor relative to the power, the attenuation factor on the field  $\mathbf{E}$  is

$$\sqrt{e^{-\alpha z}} = e^{-\frac{\alpha}{2}z} \tag{5.125}$$

In view of the presence of fiber loss, a term  $\frac{\alpha}{2}A_m$  has to be inserted on the r.h.s.

of the GMMNLSE (5.108), obtaining (in red the the addition)

$$\begin{aligned} \frac{\partial A_m}{\partial z} + j \left( \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \frac{\partial^k}{\partial t^k} \right) A_m + j \sum_{k=0}^M \kappa_{mk} A_k e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z} + \frac{\alpha}{2} A_m = \\ - j \left( 1 + \frac{1}{j\omega_0} \frac{\partial}{\partial t} \right) \gamma \sum_j \sum_k \sum_l \left( C_{jklm} A_j A_k A_l^* + H_{jklm} A_k \int_{-\infty}^{+\infty} h(t - t_1) A_j A_l^* dt_1 \right) e^{-j\Delta\beta_{jklm}z} \end{aligned} \quad (5.126)$$

where all the other terms are as in (5.110)-(5.109).

Then, we consider a new modal expansion where a term

$$e^{-\frac{\alpha}{2}z} \quad (5.127)$$

is factored out from the modal amplitude. That is, the new modal expansion is

$$\mathbf{E}_k = \frac{A_k(z, t) \mathbf{F}_k(x, y, \omega_0)}{N_k} e^{-j\beta_k^{(0)}z} e^{+j\omega_0 t} e^{-\frac{\alpha}{2}z} \quad (5.128)$$

Renaming  $A_m''$  the modal amplitude of the previous modal expansion (5.92) and keeping  $A_m$  for the new modal expansion (5.128), the relation between the two is

$$A_m'' = A_m e^{-\frac{\alpha}{2}z} \quad (5.129)$$

Hence, the GMMNLSE for the modal expansion (5.128) can be obtained by inserting (5.129) in (5.126), i.e. (in red the additions),

$$\begin{aligned} e^{-\frac{\alpha}{2}z} \frac{\partial A_m}{\partial z} - \frac{\alpha}{2} e^{-\frac{\alpha}{2}z} A_m + j \left( \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \frac{\partial^k}{\partial t^k} \right) e^{-\frac{\alpha}{2}z} A_m \\ + j \sum_{k=0}^M \kappa_{mk} e^{-\frac{\alpha}{2}z} A_k e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z} + e^{-\frac{\alpha}{2}z} A_m \\ = -j e^{-\frac{\alpha}{2}z} e^{-\frac{\alpha}{2}z} e^{-\frac{\alpha}{2}z} \left( 1 + \frac{1}{j\omega_0} \frac{\partial}{\partial t} \right) \gamma \sum_j \sum_k \sum_l \left( C_{jklm} A_j A_k A_l^* \right. \\ \left. + H_{jklm} A_k \int_{-\infty}^{+\infty} h(t - t_1) A_j A_l^* dt_1 \right) e^{-j\Delta\beta_{jklm}z} \end{aligned} \quad (5.130)$$

After basic simplifications, we retrieve the following GMMNLSE

$$\begin{aligned} \frac{\partial A_m}{\partial z} + j \left( \sum_{k=1}^{N_D} \frac{1}{k!} \beta_m^{(k)} \frac{\partial^k}{\partial t^k} \right) A_m + j \sum_{k=0}^M \kappa_{mk} A_k e^{-j(\beta_k^{(0)} - \beta_m^{(0)})z} \\ = -j e^{-\alpha z} \left( 1 + \frac{1}{j\omega_0} \frac{\partial}{\partial t} \right) \gamma \sum_j \sum_k \sum_l \left( C_{jklm} A_j A_k A_l^* + H_{jklm} A_k \int_{-\infty}^{+\infty} h(t - t_1) A_j A_l^* dt_1 \right) e^{-j\Delta\beta_{jklm}z} \end{aligned} \quad (5.131)$$

where the coefficients are defined as in (5.110)-(5.109).

Notice that, compared to (5.126), the term  $+\frac{\alpha}{2}A_m$  has (clearly) disappeared and an equivalent contribution popped up on the nonlinear term. Indeed, the nonlinearities depend on the field power. If the modal expansion (5.90) is considered for the GMMNLSE (5.126), the modal amplitudes decrease with  $z$  and so the effect of the nonlinearities, depending on the power of the modal amplitudes, is intrinsically weaker. However, if the expansion (5.128) is adopted, the modal amplitudes do not reduce with  $z$  due to loss, given that the term  $e^{-\frac{\alpha}{2}z}$  has been factored out. However, one still intuitively expects that the strength of the nonlinearities decreases with distance. This is indeed the case, because now the nonlinear terms in the GMMNLSE (5.126) (correspondent to the expansion (5.128)) explicitly reduce over distance.

The mentioned fact that the modal amplitudes do not reduce with  $z$  due to loss is sometimes an advantage in simulations and the reason for which the expansion (5.128) is sometimes preferable over (5.90).

# Chapter 6

## Simulator Design

In this chapter we provide indications for the implementation of a numerical solver for the GMMNLSE comprising both linear and nonlinear coupling effects, giving to the reader the chance to choose the preferred expression of the GMMNLSE among the various presented in the previous chapter. A comparison of the available techniques and a rigorous theoretical description of the chosen one is beyond the scope of this work. Hence, we limit ourselves to select and present in a light way an algorithm already successfully exploited in literature for optical communications, that is, the Split-Step Fourier method (SSFM) combined with a Runge-Kutta (RK) algorithm used by Poletti et al. [PH08; HP12] and Wright et al. [Wri+17]. Wright et al. started from the GMMNLSE of [HP12] and, among the other things, provided an implementation of the numerical solver for the GMMNLSE, without linear coupling and restricting to the single-polarization case, while in this thesis we are concerned with both polarizations and the inclusion of linear coupling. They implemented both the SSFM with the RK method, and the so-called Massive Parallel Algorithm (MPA), which outperforms the former if a GPU is exploited. Nonetheless, both algorithms show a significantly reduced computational time (up to two orders of magnitude) when the efficient GPU implementation is used instead of the one relying on the CPU only. The implementation and an excellent documentation is freely available at [Wri+18] and it has been a reference for this chapter.

The method we briefly present for the solution of the GMMNLSE is the SSFM with a RK method of fourth order (RK4) employed by Wright et al., adding the linear coupling effects discussed in Ch.3-4.

## 6.1 Split-Step Fourier Method

The SSFM is a widely known and used method (e.g., [Was02] [Agr19, p. 46]) to efficiently and rather easily solve the nonlinear Schrödinger equation for the single mode case, like

$$\frac{\partial A(z, t)}{\partial z} = -j\beta^{(2)}A(z, t) - j\gamma|A(z, t)|^2A(z, t) \quad (6.1)$$

where  $A$  is the modal amplitude,  $\beta^{(2)}$  is the only surviving dispersion term after having factored out the frequency-independent part of the propagation constant and having chosen a reference system moving with the group delay,  $\gamma$  is the nonlinearity constant. The previous equation can be rewritten in a symbolic way as

$$\frac{\partial A(z, t)}{\partial z} = (\hat{D} + \hat{N})A(z, t) \quad (6.2)$$

where  $\hat{D} = -j\beta^{(2)}$  is the linear dispersion operator and  $\hat{N} = -j\gamma|A(z, t)|^2$  is the nonlinear operator. The exact solution of (6.2) is, in a symbolic form, [Agr19, p. 46]

$$A(z + \Delta z, t) = e^{(\hat{D} + \hat{N})\Delta z} A(z, t) \quad (6.3)$$

where  $\Delta z$  is the longitudinal step size.

For applying the SSFM we assume the linear and nonlinear effects to be independent on one another along the longitudinal step, which clearly is just an approximation. Hence, the solution of (6.2) is given by the product of the solutions of the two phenomena computed independently on one another, i.e.,

$$A(z + \Delta z, t) = e^{\hat{D}\Delta z} e^{\hat{N}\Delta z} A(z, t) \quad (6.4)$$

With the help of the Baker-Hausdorff formula (see, e.g., [Was02] [Agr19, p. 46]), the error of the approximation is found to be of second-order in the step size  $\Delta z$  (symbolically  $O((\Delta z)^2)$ ), i.e., proportional to  $(\Delta z)^2$  and hence the method is said to be of first order. Indeed, a solution method having an error which is  $O((\Delta z)^{n+1})$  is called an  $n$ -th order method. A refined version of the SSFM (6.4) is the so-called symmetrized split-step Fourier method which consists in splitting the linear step in two half-steps, one before and one after the nonlinear step, whose symmetric structure gives the name to the method. That is,

$$A(z + \Delta z, t) = e^{\hat{D}\frac{\Delta z}{2}} e^{\hat{N}\Delta z} e^{\hat{D}\frac{\Delta z}{2}} A(z, t) \quad (6.5)$$

With the help of, again, the Baker-Hausdorff formula, the approximation is found to have an error which is  $O((\Delta z)^3)$ , and hence the symmetrized SSFM is a second-order method, compared to the original formulation which is first-order.

For the solution of (6.5) a common approach is to compute the dispersive part in the Fourier domain with the efficient FFT and the nonlinear part in time domain. Indeed, as known from the previous chapters, the solution to the LSE is expressed in the frequency domain as

$$\tilde{A}(z + \Delta z, \omega) = e^{\hat{D}_\omega \Delta z} \tilde{A}(z, \omega) \quad (6.6)$$

where  $\hat{D}_\omega$  is the dispersion operator in frequency domain which in this case is

$$\hat{D}_\omega = e^{-j\frac{\beta^{(2)}}{2}\omega^2} \quad (6.7)$$

Hence, the overall solution for one step is

$$A(z + \Delta z, t) = \mathfrak{F}^{-1} \left[ e^{\hat{D}_\omega \Delta z / 2} \mathfrak{F} \left[ e^{\hat{N} \Delta z} \mathfrak{F}^{-1} \left[ e^{\hat{D}_\omega \Delta z / 2} \tilde{A}(z, \omega) \right] \right] \right] \quad (6.8)$$

Moving from the single mode to the multimode case, the linear solution is replaced by the transfer matrix  $\mathbf{T}$  described in Sec.4.2.4 and Sec.4.3, depending on whether the physical or statistical approach is employed.

Moreover, in the multimode case instead of computing the nonlinear step as  $e^{\hat{N}z}$ , an alternative is, given the result of the first linear-half step, to numerically integrate the nonlinear equation through a method like the RK4 to obtain the nonlinear increment. The common idea between the two approaches is to have an algorithm which advances for a nonlinear step, that is, an algorithm which solves only the nonlinear part of the equation for a step. We do not provide a formal proof to back it up, but we limit to observe this is a common approach in literature [PH08; HP12; Wri+17].

## 6.2 Runge-Kutta Method

We now briefly describe the RK4 method following the notions provided by [Pre+07]. The RK4 method is a standard numerical solution algorithm for ordinary differential equations. Its basic configuration is rather simple, but some improvements, like the stepsize adaption to the estimated error, are possible. To introduce it, let us consider an ordinary differential equation like

$$\frac{dA(z)}{dz} = f(z, A) \quad (6.9)$$

where  $A$  is the unknown function and  $z$  is the independent variable. Notice that the GMMNLSE without the linear dispersive terms can be written like (6.9).

Discretizing the derivative on the l.h.s. of (6.9), that is, substituting it with  $\Delta A/\Delta z$  and multiplying  $\Delta z$  on both sides, (6.9) becomes an iterative numerical solution method

$$A_{n+1} \approx A_n + f(z_n, A_n)\Delta z, \quad \text{for } n = 0, 1, \dots \quad (6.10)$$

where  $(z_0, A_0)$  represent the initial conditions of the system and  $\Delta z$  is the integration stepsize.

In other words, at every step  $n$  we are approximating  $A$  with a Taylor expansion truncated to the first order term, where the derivative is immediately provided by the r.h.s.  $f(z, A)$  of the differential equation, evaluated in the previous point. From the Taylor expansion of the exact solution

$$A_{n+1} = A_n + f(z_n, A_n)\Delta z + O(\Delta z^2) \quad (6.11)$$

we can observe that the error in using the Euler method (6.10) is proportional to the second power of the stepsize  $\Delta z$ . A solution method having an error which is  $O((\Delta z)^{n+1})$  is called an  $n$ -th order method. Hence, the Euler method is first-order accurate in  $\Delta z$ . Its weak point is to exploit only the derivative of  $A$  computed at the initial point of the current step to compute the value of  $A$  at the end of the step, while the derivative of  $A$  changes along the step. Here is where Runge-Kutta methods come into play. RK evaluates the r.h.s. of (6.9) at multiple points for a single step and exploits such evaluations to improve the accuracy. A Runge-Kutta method of order 2, RK2, performs two evaluations of  $f$  for a single step, one at the beginning of the step (like the Euler method), as

$$k_1 = \Delta z f(z_n, A_n) \quad (6.12)$$

and one at the middle of the step exploiting the previous computation of the derivative

$$k_2 = \Delta z f(z_n + \Delta z/2, A_n + k_1/2) \quad (6.13)$$

Then, this second derivative allows for a more accurate estimation of  $A_{n+1}$  as

$$A_{n+1} = A_n + k_2 \quad (6.14)$$

Computing the error for the RK2, one finds that it is  $O((\Delta z)^3)$  and so the RK2 is second-order in  $\Delta z$ . However, notice that the RK2 requires two evaluations of  $f$  for a single step, while the Euler method just one. Hence, RK2 is superior to the Euler method only if, for the same accuracy, it allows to increase the stepsize more than twice compared to the Euler method. This indicates that a higher-order method does not always mean a better method.

Combining multiple evaluations of  $f$  at intermediate points of a step allows for higher order RK methods. The construction of the RK formulas is not trivial

because the evaluations have to be chosen to cancel the error terms of the Taylor expansion, leaving an error  $O((\Delta z)^{n+1})$  for a RK method to be called of order  $n$ . The RK4 method is one of the most popular since being fourth-order, yet still easy to implement and with still a limited number of evaluations, four, needed. That is, [Pre+07]

$$k_1 = \Delta z f(z_n, A_n) \quad (6.15)$$

$$k_2 = \Delta z f(z_n + \Delta z/2, A_n + k_1/2) \quad (6.16)$$

$$k_3 = \Delta z f(z_n + \Delta z/2, A_n + k_2/2) \quad (6.17)$$

$$k_4 = \Delta z f(z_n + \Delta z, A_n + k_3) \quad (6.18)$$

$$A_{n+1} = A_n + 1/6k_1 + 1/3k_2 + 1/3k_3 + 1/6k_4 \quad (6.19)$$

The same considerations about the comparison between RK2 and Euler method hold for the comparison between RK2 and RK4. However, it is generally the case that RK4 outperforms RK2 in the sense that it allows for a stepsize more than the double of RK2 for the same accuracy.

The stepsize  $\Delta z$  can also be adaptively adjusted from step to step, based on some error metric which can be computed from  $A_{n+1}$ , through the so-called double stepping or the embedded Runge-Kutta formulas. The basic idea is to exploit two estimates of  $A_{n+1}$  computed independently with two different methods.

In the double stepping, one estimate is computed through the chosen RK for the chosen stepsize, while the other estimate is computed by halving the stepsize. The subtraction between the two estimates is an estimate of the error, based on which one can exploit algorithms for stepsize adaption. The most basic consists in recomputing  $A_{n+1}$  (and the error) for a reduced stepsize, if the error is higher than a threshold.

In the embedded RK formulas, the two estimates of  $A_{n+1}$  come from two RKs of different order, but based on the same  $k_i$  evaluations (combined in a different way). Hence, the difference between the two provides an estimate of the truncation error for the lower order RK. Even if formally the error applies to the lower order RK, the higher order estimate of  $A_{n+1}$  can be used, in which case it is said that  $A_{n+1}$  has been obtained by local extrapolation. The first to be discovered has been the Runge-Kutta-Fehlberg method which exploits a RK4 and a RK5 both computed through six evaluations  $k_i$  arranged in different order.

Finally, from our first attempts of implementation and from running the implementation of [Wri+17] it emerged that the computational time is in the order of days per kilometer of fiber, if a GPU is not used. The reason for it is the tiny step sizes needed to account for the mode beating arising from the phase-mismatch between the modes [Wri+17] and a sufficient number of time/frequency points to prevent the signal broadening (both in time and frequency) to introduce artifacts

[Was02]. When an efficient GPU implementation is considered, the time reduces between one and two orders of magnitude.

### 6.3 Hints on Manakov Equations

Another option to significantly speed up the integration consists in exploiting a simplified version of the GMMNLSE, the Manakov equations. These are a family of equations first invented and used within the framework of SMF transmission [MMW97] for simplifying the numerical solution of the nonlinear Schrödinger equation. They have been later extended to the multimode case [MAS12b; MAS12a; Ryf+12; ASM16] considering firstly only the Kerr terms and later the Raman as well, for different scenarios of linear mode coupling. Nowadays they have been derived for weak, intermediate and strong coupling [Agr19, p. 638].

These propagation equations are derived by averaging the coefficients of the nonlinear evolution with respect to the fast evolution of the frequency-independent birefringence and mode coupling. The idea behind this approach is that the nonlinear perturbations affect the signal propagation on length scales much longer than the fast (hyper)polarization variations due to birefringence and mode coupling. Hence, instead of considering the actual nonlinear terms of the GMMNLSE, it is possible to average their parameters over the (hyper)polarization variations. In this way, we are missing the exact local description of the signal propagation, which for long-haul communications is irrelevant, but we are enormously reducing the solution time. Indeed, the Manakov equations have less terms and allow to use much longer stepsizes compared to the GMMNLSE, not having to deal anymore with the short beat length.

More details on the the topic can be found, e.g., in [Ryf+12; ASM16].

# Chapter 7

## Conclusions and Outlook

This thesis has been composed of three main parts. In the first we have reviewed polarization-mode coupling, in the second spatial mode coupling.

Particular attention has been devoted to the 2-polarization case and the theory of the principal states of polarization, because they are basic elements for the more complex multimode scenario and to understand advanced concepts, like the Manakov equations. In addition, they are naturally generalized to the multimode scenario.

We have provided and implemented two possible models for the multimode fiber channel in linear regime, one modeling some distortion effects acting on the fibers and one based on a statistical approach. The latter seems to introduce a too strong level of coupling compared to the former when the same segment length is assumed. Yet, it is a fast simulation tool and it might be tweaked to bring about a level of coupling closer to reality.

There is room for research in modeling physical perturbing phenomena. In particular, alternatives should be investigated for bend since the model employed by us loses significance, if realistic level of coupling are desired.

In the last part of the thesis, the derivation of the GMMNLSE has been reviewed starting from Maxwell's equations, highlighting the role of the different terms and of the various possible normalizations. The reader shall adapt it to the considered scenario. For instance, the Raman term can be simplified when the signal band is below 10 THz, as illustrated. A possible design for the numerical solver, comprising both linear and nonlinear effects, has been proposed. The next step consists in its implementation, giving attention to exploit a GPU to speed up the computations.

As a future direction of research, it has been mentioned that the Manakov equations significantly reduce the computational burden for long-haul applications, as space-division multiplexing. However, the intuition and derivation of this technique requires a broad view on both the linear and nonlinear phenomena described by the GMMNLSE. This thesis covers all of them to various degree of

detail. Hence, it provides all the necessary tools to approach the Manakov model and to implement it. The availability of both the GMMNLSE, where the linear effects are modeled through a physical approach, and the Manakov model, which is usually based on statistical considerations on linear coupling, would allow to assess the accuracy of the Manakov model in various scenarios of coupling.

# Bibliography

- [Agr19] Govind Agrawal. *Nonlinear Fiber Optics*. Academic Press Inc, 2019.
- [Ant+12] Cristian Antonelli et al. “Stokes-space analysis of modal dispersion in fibers with multiple mode transmission”. In: *Optics Express* (2012).
- [ASM16] Cristian Antonelli, Mark Shtaif, and Antonio Mecozzi. “Modeling of Nonlinear Propagation in Space-Division Multiplexed Fiber-Optic Transmission”. In: *Journal of Lightwave Technology* (2016).
- [CES16] Joel Carpenter, Benjamin Eggleton, and Jochen Schroder. “Comparison of principal modes and spatial eigenmodes in multimode optical fibre”. In: *Laser And Photonics Reviews* (2016).
- [Chr09] Andrew Chrallyvy. “Plenary paper: The coming capacity crunch”. In: *2009 35th European Conference on Optical Communication*. 2009, pp. 1–1.
- [Coe10] Leonardo Didier Coelho. “Modeling, Simulation and Optimization of Optical Communication Systems using Advanced Modulation Formats”. In: PhD thesis. Munich, 2010. Chap. 2.9, pp. 16–33.
- [FK05] Shanhui Fan and Joseph Kahn. “Principal modes in multimode waveguides”. In: *Optics Letters* (2005).
- [FP91] G.J. Foschini and C.D. Poole. “Statistical theory of polarization dispersion in single mode fibers”. In: *Journal of Lightwave Technology* 9.11 (1991), pp. 1439–1456. DOI: 10.1109/50.97630.
- [GK00] J. P. Gordon and H. Kolgenik. “PMD fundamentals: Polarization mode dispersion in optical fibers”. In: *PNAS* (2000).
- [Han21] Norbert Hanik. *Notes of Optical Communications and Seminars on PMD and PMD-Manakov equations*. Paolo Carniello’s Notes based on Hanik’s slides. Seminar Notes. Insitute for Communications Engineering, Technische Universität München, May 2021.

- [HP12] Peter Horak and Francesco Poletti. “Multimode Nonlinear Fibre Optics: Theory and Applications”. In: *Recent Progress in Optical Fiber Research*. Ed. by Dr Moh. Yasin. 2012.
- [Jua+14] Adrian A. Juarez et al. “Modeling of Mode Coupling in Multimode Fibers With Respect to Bandwidth and Loss”. In: *Journal of Lightwave Technology* 32.8 (2014), pp. 1549–1558. DOI: 10.1109/JLT.2014.2308059.
- [Kei11] Gerd Keiser. *Optical Fiber Communications, Fourth Edition*. McGraw-Hill, 2011.
- [KHS12] Joseph Kahn, Keang-Po Ho, and Mahdieh Bagher Shemirani. “Mode Coupling Effects in Multi-Mode Fibers”. In: *Optical Fiber Communication Conference*. Optical Society of America, 2012, OW3D.3. DOI: 10.1364/OFC.2012.OW3D.3. URL: <http://www.osapublishing.org/abstract.cfm?URI=OFC-2012-OW3D.3>.
- [KK97] Ivan Kaminov and Thomas Koch. *Optical Fiber Telecommunications IIIA*. Academic Press, 1997.
- [KM04] M. Kolesik and J. V. Moloney. “Nonlinear optical pulse propagation simulation: From Maxwell’s to unidirectional equations”. In: *Phys. Rev. E* 70 (3 Sept. 2004), p. 036604. DOI: 10.1103/PhysRevE.70.036604. URL: <https://link.aps.org/doi/10.1103/PhysRevE.70.036604>.
- [Mar74] Dietrich Marcuse. *Theory of Dielectric Optical Waveguides*. Quantum electronics—principles and applications. Academic Press Inc, 1974. ISBN: 0124709508,9780124709508. URL: <http://gen.lib.rus.ec/book/index.php?md5=AAFA53BB0D76D824792989BFC9A0FB76>.
- [Mar75] D. Marcuse. “Coupled-mode theory for anisotropic optical waveguides”. In: *The Bell System Technical Journal* 54.6 (1975), pp. 985–995. DOI: 10.1002/j.1538-7305.1975.tb02878.x.
- [MAS12a] Antonio Mecozzi, Cristian Antonelli, and Mark Shtaif. “Coupled Manakov equations in multimode fibers with strongly coupled groups of modes”. In: *Opt. Express* 20.21 (Oct. 2012), pp. 23436–23441. DOI: 10.1364/OE.20.023436. URL: <http://www.opticsexpress.org/abstract.cfm?URI=oe-20-21-23436>.
- [MAS12b] Antonio Mecozzi, Cristian Antonelli, and Mark Shtaif. “Nonlinear propagation in multi-mode fibers in the strong coupling regime”. In: *Opt. Express* 20.11 (May 2012), pp. 11673–11678. DOI: 10.1364/OE.20.011673. URL: <http://www.opticsexpress.org/abstract.cfm?URI=oe-20-11-11673>.

- [MEA13] Sami Mumtaz, René-Jean Essiambre, and Govind P. Agrawal. “Non-linear Propagation in Multimode and Multicore Fibers: Generalization of the Manakov Equations”. In: *Journal of Lightwave Technology* 31.3 (2013), pp. 398–406. DOI: 10.1109/JLT.2012.2231401.
- [Mez06] Francesco Mezzadri. “How to generate random matrices from the classical compact groups”. In: *Notices of the American Mathematical Society* 54 (Oct. 2006).
- [Mid03] Michele Midrio. *Campi Elettromagnetici*. SGE Editoriali, 2003. ISBN: 88-86281-82-X.
- [MMW97] D. Marcuse, C.R. Menyuk, and P.K.A. Wai. “Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence”. In: *Journal of Lightwave Technology* 15.9 (1997), pp. 1735–1746. DOI: 10.1109/50.622902.
- [MS20] Messe Muenchen and SPECTARIS. *Light as the key to global environment sustainability. High-tech photonics solutions to protect the environment and preserve resources*. Study. 2020. URL: [https://www.spectaris.de/fileadmin/Content/Verband/Themenspecial/Study\\_GreenPhotonics\\_2020\\_eng.pdf](https://www.spectaris.de/fileadmin/Content/Verband/Themenspecial/Study_GreenPhotonics_2020_eng.pdf).
- [NJ05] L. E. Nelson and R. M. Jopson. “Introduction to polarization mode dispersion in optical systems”. In: *Polarization Mode Dispersion*. Ed. by Curtis Menyuk and Andrea Galtarossa. Optical and Fiber Communications Reports. Springer, 2005.
- [Pal13] Luca Palmieri. “Coupling mechanism in multimode fibers”. In: vol. 9009. Dec. 2013, 90090G. DOI: 10.1117/12.2042763.
- [Pal14] Luca Palmieri. “Coupling mechanism in multimode fibers”. In: *Next-Generation Optical Communication: Components, Sub-Systems, and Systems III*. Vol. 9009. International Society for Optics and Photonics. 2014, 90090G.
- [PH08] Francesco Poletti and Peter Horak. “Description of ultrashort pulse propagation in multimode optical fibers”. In: *J. Opt. Soc. Am. B* 25.10 (Oct. 2008), pp. 1645–1654. DOI: 10.1364/JOSAB.25.001645. URL: <http://www.osapublishing.org/josab/abstract.cfm?URI=josab-25-10-1645>.
- [PN97] Craig Poole and Jonathan Nagel. “Polarization Effects in Lightwave Systems”. In: *Optical Fiber Telecommunications IIIA*. Ed. by Ivan Kaminov and Thomas Koch. Academic Press, 1997. Chap. 6.

- [Poo+88] C. D. Poole et al. “Polarization Dispersion and Principal States in a 147-km Undersea Lightwave Cable”. In: *Journal of Lightwave Technology* (1988).
- [Pre+07] William H. Press et al. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3rd ed. USA: Cambridge University Press, 2007. Chap. 17. ISBN: 0521880688.
- [PW86] C.D. Poole and R.E. Wagner. “Phenomenological Approach to Polarisation Dispersion in Long Single-Mode Fibres”. In: *Electronics Letters* (1986).
- [PWN91] C. D. Poole, J. H. Winters, and J. A. Nagel. “Dynamical equation for polarization dispersion”. In: *Opt. Lett.* 16.6 (Mar. 1991), pp. 372–374. DOI: 10.1364/OL.16.000372. URL: <http://ol.osa.org/abstract.cfm?URI=ol-16-6-372>.
- [RP05] K. Rottwitt and J.H. Povlsen. “Analyzing the fundamental properties of Raman amplification in optical fibers”. In: *Journal of Lightwave Technology* 23.11 (2005), pp. 3597–3605. DOI: 10.1109/JLT.2005.857776.
- [Ryf+12] Roland Ryf et al. “Mode-Division Multiplexing Over 96 km of Few-Mode Fiber Using Coherent  $6 \times 6$  MIMO Processing”. In: *Journal of Lightwave Technology* 30.4 (2012), pp. 521–531. DOI: 10.1109/JLT.2011.2174336.
- [She+09] Mahdiah B. Shemirani et al. “Principal Modes in Graded-Index Multimode Fiber in Presence of Spatial- and Polarization-Mode Coupling”. In: *Journal of Lightwave Technology* 27.10 (2009), pp. 1248–1261. DOI: 10.1109/JLT.2008.2005066.
- [Sil+16] Pierre Sillard et al. “50 um Multimode Fibers for Mode Division Multiplexing”. In: *Journal of Lightwave Technology* 34.8 (2016), pp. 1672–1677. DOI: 10.1109/JLT.2015.2507442.
- [SKP21] Christian M. Spenner, Peter M. Krummrich, and Klaus Petermann. “Measurement Based Modeling of Mode Coupling Effects in a 50 um Graded Index Multi Mode Fiber”. In: *Photonic Networks; 22th ITG Symposium*. 2021, pp. 1–5.
- [SM05] Mark Shtaif and Antonio Mecozzi. “Modelling of polarization mode dispersion in optical communications systems”. In: *Polarization Mode Dispersion*. Ed. by Curtis Menyuk and Andrea Galtarossa. Optical and Fiber Communications Reports. Springer, 2005.
- [Som06] Carlo G. Someda. *Electromagnetic Waves*. TaylorAndFrancis, 2006.

- [Stu18] Warren Stutzman. *Polarization in Electromagnetic Systems Second Edition*. Artech House, 2018.
- [US79] R. Ulrich and A. Simon. “Polarization optics of twisted single-mode fibers”. In: *Appl. Opt.* 18.13 (July 1979), pp. 2241–2251. DOI: 10.1364/AO.18.002241.
- [Wal18] Helio Waldman. “The Impending Optical Network Capacity Crunch”. In: *2018 SBFoton International Optics and Photonics Conference (SBFoton IOPC)*. 2018, pp. 1–4. DOI: 10.1109/SBFoton-IOPC.2018.8610949.
- [Was02] Brian R. Washburn. “Dispersion and Nonlinearities associated with super continuum generation in microstructure fibers,” in: PhD thesis. Georgia Tech, 2002. Chap. 4.
- [wik] wikipedia.org. *Matrix exponential*. URL: [https://en.wikipedia.org/wiki/Matrix\\_exponential](https://en.wikipedia.org/wiki/Matrix_exponential).
- [WN17] Peter J. Winzer and David T. Neilson. “From Scaling Disparities to Integrated Parallelism: A Decathlon for a Decade”. In: *Journal of Lightwave Technology* 35.5 (2017), pp. 1099–1115. DOI: 10.1109/JLT.2017.2662082.
- [Wri+17] Logan Wright et al. “Multimode Nonlinear Fiber Optics: Massively Parallel Numerical Solver, Tutorial and Outlook”. In: *IEEE Journal of Selected Topics in Quantum Electronics* PP (Aug. 2017). DOI: 10.1109/JSTQE.2017.2779749.
- [Wri+18] Logan Wright et al. *GMMNLSE Implementation*. 2018. URL: <https://github.com/WiseLabAEP/GMMNLSE-Solver-FINAL>.