TEL AVIV UNIVERSITY

Raymond and Beverly Sackler Faculty of Exact Sciences

School of Mathematical Sciences

Vectorial Effects and Multiple Filamentation in Self-Focusing of Laser Beams

Thesis submitted for the degree of “Doctor of Philosophy”

by

BOAZ ILAN

Submitted to the Senate of Tel-Aviv University

July 2002

The work on this thesis was carried out under the supervision of

DR. GADI FIBICH

Acknowledgments

Some five years ago, when I began my graduate studies, Gadi Fibich arrived at Tel Aviv University and gave a course on NLS theory, which I attended. At the end of the course Gadi offered the students a list of assignments to choose from. When I chose one, little did I know that Gadi would become my PhD advisor and that my assignment would evolve into this thesis.

Gadi has fueled our journey with his insight, support, and encouragement. Always leading to new challenges and guiding me to the best path, Gadi has taught me that research can be both prolific and enjoyable.

During these years Gadi and I have enjoyed fruitful collaborations with George Papanicolaou at Stanford University, who was Gadi's PhD advisor and my scientific "grandfather", with Semyon Tsynkov at North Carolina State University, and with Alexandro Aceves at the University of New Mexico. I benefitted from my discussions with Alex Gaeta, who does intense-laser experiments at Cornell University. I am deeply grateful to George, Semyon, and Alex, who assisted me on several occasions.

I will always remember the numerous friends I made while occupying the OpenSpace. I will miss their distractions, humor, and moments of serious discussions.

I am indebted to the Wise, Marejn, Schreiber, and Sackler families for the generous support I received from their foundations during my graduate studies.

My parents and brother have taught me much about science, more than they will take credit for. It is through their love, and the patient support from my parents, that I could evolve my assignment into this thesis.

Boaz Ilan

Tel Aviv, Israel, 2002

Abstract

Intense lasers have applications in medicine (*e.g.*, laser surgery), communications, remote sensing of the atmosphere, and more. When an intense laser beam propagates in a Kerr medium, such as air, water, or glass, the beam can become a long and very narrow filament as a result of its nonlinear interaction with the medium, a phenomenon called self-focusing. Laser experiments since the 1960's show that a sufficiently intense laser beam propagating in a Kerr medium can break up into several long and narrow filaments, a phenomenon known as multiple filamentation or beam breakup. Multiple filamentation can be either a "curse" (e.g., in eye surgery) or a blessing, depending on the application. In both cases we need to know what causes this phenomenon in order to control it. The nonlinear Schrödinger equation (NLS), which is derived from the nonlinear Maxwell equations, is the model equation for self-focusing. The NLS model, however, has several weaknesses. One weakness is that the NLS model predicts a "catastrophic" self-focusing process (*i.e.*, singularity formation), which corresponds to infinite energy density. Physical quantities, however, do not become infinite. Indeed, in experiments, the self-focusing process is arrested by physical mechanisms that are neglected in the derivation of the NLS model. In addition, when the input beam is cylindrically symmetric (*e.g.*, a Gaussian beam) then according to the NLS model it remains cylindrically symmetric during the propagation. However, the beam cannot remain cylindrically symmetric during multiple filamentation. Therefore, the self-focusing model should include a physical mechanism that can break-up cylindrical symmetry and lead to multiple filamentation. Our study focuses on vectorial effects, which result from the interaction between the three components of the beam's electric field during propagation. We show that these effects, which are neglected in the NLS model, relate to all the weaknesses mentioned above.

Much of our attention has been directed to possible explanations for multiple filamentation. For over thirty years the standard (and only) explanation for multiple filamentation has been that it is initiated by *noise* in the input beam ("random" multiple filamentation). In this study we propose a novel *deterministic* explanation for multiple filamentation based on vectorial effects. Consequently, even if

the input beam has a perfectly cylindrically-symmetric profile, it can still undergo multiple filamentation. To show this, we derive a new scalar equation from the nonlinear Maxwell equations for linearly polarized beams, which shows how vectorial effects can deterministically break up the cylindrical symmetry of the beam. Numerical simulations of this (2+1)-dimensional equation show that when the input beam is sufficiently powerful, vectorial effects lead to multiple filamentation. Furthermore, we compare multiple filamentation due to vectorial effects with multiple filamentation due to noise, from which we conclude that multiple filamentation can result from either mechanism. Which of the two mechanisms leads to multiple filamentation in experiments remains an open question. We suggest a simple experiment to answer this question.

Having studied self-focusing of *linearly polarized* beams, we were motivated to study the self-focusing of *circularly polarized* beams. The primary motivation for this study was that, in contrast to linear polarization, circular polarization does not induce a preferred direction in the transverse plane. Therefore, we suspected that multiple filamentation is less likely to occur with circular polarization. Because the NLS model is derived for linear polarization, a systematic study of multiple filamentation of circularly polarized beams requires a different mathematical model. After inspecting the literature we discovered that previous studies used the wrong model to study self-focusing of circularly polarized beams. In this study we use a systematic perturbation analysis to derive a new model from Maxwell's equations for self-focusing of circularly polarized beams in a Kerr medium. Using this new model we show that circular polarization is stable. Finally, we compare the possibilities of multiple filamentation with the linear and circular polarization states. Based on this comparison we predict that multiple filamentation is suppressed with circular polarization.

Contents

List of Figures

List of Tables

0.1 Notations

The principal constants in this work are listed in Table 1. Additional constants and variables for the circular polarization state are listed in Table 2. The gradient, divergence, curl, and Laplace operators act on (x, y, z), *e.g.*, $\Delta = \partial_{xx} + \partial_{yy} + \partial_{zz}$. When these operators are suffixed by "$_\perp$" they act only on (x, y), *e.g.*, $\Delta_\perp = \partial_{xx} + \partial_{yy}$. In the derivations in Chapter 3 and Chapter 4 we nondimensionalize the variables according to Eq. (3.4). We use calligraphic letters ($\vec{\mathcal{E}}$, $\vec{\mathcal{P}}_{\rm NL}$, etc) to denote the dimensional variables and non-calligraphic letters ($\vec{E}$, $\vec{P}$, etc) to denote dimensionless variables [see Eqs. (3.4) and (3.5c)].

Constant	**meaning**	**equation**
μ_0	electric permeability	(2.1)
ϵ_0	electric permittivity	(2.2)
c	light speed in vacuum	(2.3)
ω_0	light frequency	(2.5)
n_0	linear index of refraction	(2.6)
k_0	wavenumber	(2.7)
$\bar{n}_2$	nonlinear (Kerr) index of refraction	(2.12)
γ	Kerr-mechanism constant	(2.12)
r_0	input beam width	(3.1)
f	ratio of wavelength to input beam width	(3.1)
N_c	threshold power for collapse	(2.21)

Table 1: Principal constants, their meaning, and the first equation in which they appear.

Variable/constant	**dimensional**	**nondimensional**	**equation**
deviation from circular polarization	—	ε	(4.6)
electric field components	$\mathcal{E}_\pm$	$E_\pm$	(4.1)
electric amplitudes	$\mathcal{A}_\pm$	$A_\pm$	(4.8)
input power	$\mathcal{P}_\pm(0)$	$N_\pm(0)$	(4.5)
threshold power for collapse	$\mathcal{P}_c^{\rm circ}$	$N_c^{\rm circ}$	(4.23)

Table 2: Variables and constants defined for the circular-polarization state and the first equation in which they appear.

Chapter 1

Outline

The nonlinear Schrödinger equation (**NLS**) is the model equation for the propagation of intense *linearly polarized* laser beams in a Kerr medium. Based on the NLS equation, Kelley predicted in 1965 that intense laser beams whose input power is above a certain threshold would undergo "catastrophic" self-focusing in a finite propagation distance [39], *i.e.*, become narrow much more than the linear diffraction limit (see Figure 1.1). Although this prediction was confirmed experimentally, the NLS, as a physical model for self-focusing, has several weaknesses:

1. According to the NLS model, the beam intensity becomes infinite at a finite distance (this singularity formation is termed beam collapse or blowup). As a result, the NLS model fails to describe beam propagation beyond that point.

2. The NLS is the leading order approximation to Maxwell's equations. However, because the NLS is extremely sensitive to the addition of small terms [33], one cannot ignore the remaining terms simply because they are small.

3. According to the NLS model, beams with cylindrically-symmetric input profile should remain cylindrically symmetric during propagation. However, self-focusing experiments in solids, liquids, and gases have shown that catastrophic self-focusing is often preceded by *multiple filamentation*[*], in which a single input beam breaks-up into several long and narrow filaments (see Figure 1.2) [1, 3, 12, 13, 14, 15, 16, 18, 19, 41, 43, 52, 53, 55]. Therefore, a self-focusing model should include a mechanism that breaks-up the cylindrical symmetry and leads to multiple filamentation.

[*] Multiple filamentation is also called multiple-foci, small-scale filaments, or beam-breakup.

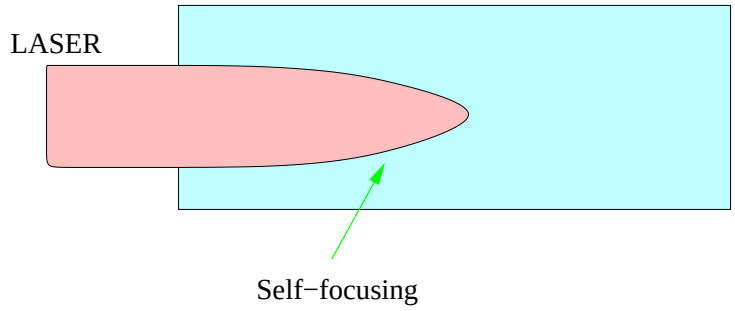

Figure 1.1: Self-focusing.

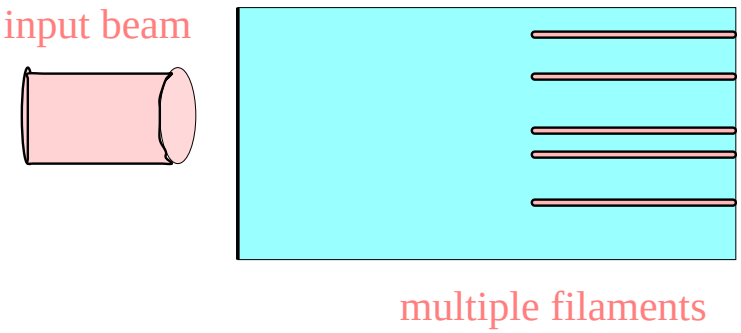

Figure 1.2: Multiple filamentation.

In Chapter 2 we review some facts from Electromagnetic Theory, NLS theory, and discuss the connection between multiple filamentation and cylindrical symmetry. In Chapter 3 we focus on the role of vectorial effects in self-focusing of *linearly polarized* input beams. We show that these effects, which are neglected in the derivation of the NLS model from Maxwell's equations, relate to all of the above weaknesses of the NLS model[†]:

1. Vectorial effects arrest the blowup. As a result, the NLS model with vectorial effects can be used to describe beam propagation beyond the NLS blowup point.

2. Vectorial effects, although small in magnitude, have a large effect on beam propagation. In fact, we show that vectorial effects have a larger effect than nonparaxiality.

3. In the case of linear polarization the *deterministic* vectorial effects can lead to multiple filamentation.

For over thirty years, the standard explanation for multiple filamentation, due to Bespalov and Talanov [5], has been that it is initiated by *random* noise in the input-beam profile ("random multiple

[†]We quote [26, 28, 29, 30] in this work without further citation.

filamentation"). Our results thus show that multiple filamentation can result from either vectorial effects or from noise in the input beam. Which of the two mechanisms leads to multiple filamentation in experiments remains an open question. To answer this question we suggest a simple experiment for deciding whether multiple filamentation is due to vectorial effects or not.

We recall that the NLS model is derived under the assumption that the beam is *linearly polarized*. In 1966 Close, Giuliano, Hellwarth, Hess, McClung, and Wagner [21] conducted experiments with intense *circularly polarized* input beams propagating in Kerr media, which suggested that circular polarization is unstable. Close et al. also proposed a mathematical model for self-focusing of circularly polarized beams, which they used to explain the observed instability of circular polarization. Their model is a system of two coupled NLS equations for the two circular polarization components. Subsequent theoretical studies have used the same system of equations as Close *et al.*, but obtained contradictory results with regard to circular-polarization (in)stability. As a result, to these days there is some confusion in the literature with regard to circular polarization stability. Remarkably, the only thing that was always agreed upon was the Close et al. model itself. As we show in this study, however, this model is based on problematic assumptions, and it can lead to wrong results. Hence, previous studies used the wrong model for studying circular polarization stability.

In Chapter 4 we present a systematic study of propagation of circularly polarized beams in a Kerr medium. In contrast to previous studies of circular polarization, our derivation does not neglect nonparaxiality, the coupling to the axial component of the electric field, and the contribution from the grad-div term. This leads to a new system of equations that takes into account nonparaxiality, coupling to the axial component, the contributions from the grad-div term, and coupling to the opposite circular component (*i.e.*, the one rotating in the opposite direction). Using this system we show that circular polarization is stable during self-focusing. We also show that nonparaxiality, the coupling to the axial component, and the contribution from the grad-div term arrest collapse, leading instead to focusing-defocusing oscillations. In addition, we show that circularly polarized beams are much less likely to undergo multiple filamentation than linearly polarized beams.

In Chapter 5 we compare self-focusing of circularly- and linearly-polarized beams. Based on our results we predict that circularly polarized beams are much less likely to undergo multiple filamentation than linearly polarized beams.

Numerical simulations are used frequently in this study, either to complete the picture obtained from the asymptotic results, or when analytic and asymptotic results are too difficult to obtain. In Chapter 6 we discuss specific issues related to the numerical simulations of self-focusing beams, such as numerical realization of nonparaxiality and the reliability of numerical simulations of multiple filamentation. For example, we show that when the computational domain is not sufficiently large,

reflections from the boundary of the numerical mesh can lead to what may appear as breakup of cylindrical symmetry and even as multiple filamentation.

We conclude in Chapter 7 with several open problems and suggest possible directions for further study.

Chapter 2

The vector Helmholtz and the NLS

2.1 Review of linear propagation

We begin by reviewing several facts from linear Electromagnetic Theory. Let $\vec{E}(x,y,z,t) = (E_1, E_2, E_3)$, $\vec{D}(x,y,z,t) = (D_1, D_2, D_3)$, and $\vec{B}(x,y,z,t) = (B_1, B_2, B_3)$ be the electric field intensity, electric field induction, and the magnetic field induction, respectively. When the medium is isotropic, homogeneous, non-magnetic, and has no free electromagnetic sources (*i.e.*, charges or currents) these fields satisfy the Maxwell system* [38]

$$\frac{1}{\mu_0}\nabla \times \vec{B} - \vec{D}_t = 0 \,, \tag{2.1a}$$

$$\nabla \times \vec{E} + \vec{B}_t = 0 \,, \tag{2.1b}$$

$$\nabla \cdot D = 0 \,, \tag{2.1c}$$

$$\nabla \cdot B = 0 \,, \tag{2.1d}$$

where $\nabla = (\partial_x, \partial_y, \partial_z)$, μ_0 is magnetic permeability, and the electric induction field satisfies

$$\vec{D} = \epsilon_0(\vec{E} + \vec{P}_{\mathrm{L}}) \,, \tag{2.2}$$

where ϵ_0 is electric permittivity and $\vec{P}_{\mathrm{L}} = (P_1, P_2, P_3)$ is the linear (electric) polarization field, which depends on the interaction between the field and the medium. Taking the curl of (2.1b), substituting $\vec{B}$ from (2.1b), and using (2.2) yields the Maxwell equation

$$\frac{1}{c^2}\vec{E}_{tt} + \nabla \times \nabla \times \vec{E} = -\frac{1}{c^2}\vec{P}_{\mathrm{L,tt}} \,, \tag{2.3}$$

*Unless otherwise mentioned, in this work we use the SI or Rationalized MKSA units. For conversion between different electromagnetic units see, *e.g.*, [38, Appendix].

where $c = 1/\sqrt{\epsilon_0 \mu_0}$ is light speed in vacuum. It is customary to use the vectorial identity

$$\nabla \times \nabla \times \vec{E} \equiv -\Delta \vec{E} + \nabla(\nabla \cdot \vec{E}) \, ,$$

where $\Delta = \partial_{xx} + \partial_{yy} + \partial_{zz}$, in order to rewrite Eq. (2.3) as the vectorial wave equation

$$\frac{1}{c^2}\vec{E}_{tt} - \Delta \vec{E} + \nabla(\nabla \cdot \vec{E}) = -\frac{1}{c^2}\vec{P}_{\mathrm{L,tt}} \, . \tag{2.4}$$

Electromagnetic waves, such as sunlight and laser beams, are solutions of Maxwell's equations. One of the features of laser beams is that they are monochromatic, *i.e.*, they emit radiation at a certain wavelength. Let us consider the case of a "cw" laser beam (short for continuous-wave), also called "time-harmonic" in the mathematical literature. The electric and polarization fields of such beams can be represented by

$$\vec{E}(x,y,z,t) = \vec{\mathcal{E}}(x,y,z)\, e^{i\omega_0 t} + c.c. \, , \qquad \vec{P}_{\mathrm{L}}(x,y,z,t) = \vec{\mathcal{P}}_{\mathrm{L}}(x,y,z)\, e^{i\omega_0 t} + c.c. \, , \tag{2.5}$$

where ω_0 is frequency, $\vec{\mathcal{E}}(x,y,z) = (\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3)$ and $\vec{\mathcal{P}}_{\mathrm{L}}(x,y,z) = (\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3)$ are time-independent (complex-valued) vector fields, and "$c.c.$" denotes the complex-conjugate of the term to its left. In order to provide a complete description of the field dynamics, it is necessary to describe the relation between the polarization field and the electric field. In a linear medium", *i.e.*, when the intensity of the laser beam is not high enough to induce a nonlinear response, the polarization field can be described by the *linear* constitutive relation

$$\vec{\mathcal{P}}_{\mathrm{L}} = (n_0^2 - 1)\vec{\mathcal{E}} \, , \tag{2.6}$$

where $n_0 = n_0(\omega_0)$ is the linear index of refraction. In particular, in vacuum $n_0 = 1$ and $\vec{P}_{\mathrm{L}} \equiv 0$. Substituting (2.5) into the Maxwell equation (2.4) and using (2.6) gives the *linear vector Helmholtz equation*

$$\Delta \vec{\mathcal{E}} - \nabla(\nabla \cdot \vec{\mathcal{E}}) + k_0^2 \vec{\mathcal{E}} = 0 \, , \tag{2.7}$$

where $k_0 = \omega_0 n_0 / c$ is wavenumber. Using (2.1c) leads to

$$\nabla \cdot \vec{\mathcal{E}} = 0 \, . \tag{2.8}$$

Maxwell's equations (2.4) [and the vector Helmholtz (2.7)] have exact time-harmonic solutions that are *linearly polarized, i.e.*, their electric field vector points to a fixed direction that is perpendicular to the direction of propagation. For example, the following are exact time-harmonic plane waves solutions of (2.4)

$$\vec{E}^x = (\cos(\omega_0 t - k_0 z), 0, 0) = e^{i(\omega_0 t - k_0 z)}(1,0,0) + c.c. \, ,$$

$$\vec{E}^y = (0, \sin(\omega_0 t - k_0 z), 0) = e^{i(\omega_0 t - k_0 z)}(0,i,0) + c.c. \, ,$$

which are linearly polarized in the x and y directions, respectively. We note that there is a $\pi/2$ phase difference between $\vec{E}^x$ and $\vec{E}^y$. As (2.4) is a linear equation, the sum (superposition) of any two solutions is also a solution. Thus, the following fields

$$\frac{1}{\sqrt{2}}(\vec{E}^x \pm \vec{E}^y) = \frac{1}{\sqrt{2}}(\cos(\omega_0 t - k_0 z), \pm\sin(\omega_0 t - k_0 z), 0) = \frac{1}{\sqrt{2}}\, e^{i(\omega_0 t - k_0 z)}(1, \pm i, 0) + c.c. \quad (2.9)$$

are also plane wave solutions of (2.4), which are called *circularly polarized*, because their electric-field vector forms an imaginary circle in the (x, y) plane during the propagation. In view of (2.9) it is customary to define for time-harmonic fields the left $(+)$ and right $(-)$ circular components

$$\mathcal{E}_\pm(x, y, z) = \frac{1}{\sqrt{2}}\,(\mathcal{E}_1 \pm i\mathcal{E}_2)\,.$$

Accordingly, a left-circular beam is defined by $\mathcal{E}_- \equiv 0$ and a right-circular beam is defined by $\mathcal{E}_+ \equiv 0$.

2.2 The nonlinear vector Helmholtz

When an *intense* laser beam propagates in a medium its interaction with the medium is nonlinear, in which case the electric induction field is given by

$$\vec{D} = \epsilon_0(\vec{E} + \vec{P}_{\rm L} + \vec{P}_{\rm NL})\,, \quad (2.10)$$

where $\vec{P}_{\rm NL}$ is the nonlinear polarization field. In that case the propagation of cw beams is governed by the *nonlinear vector Helmholtz equations*

$$\Delta\vec{\mathcal{E}}(x, y, z) - \nabla(\nabla\cdot\vec{\mathcal{E}}) + k_0^2\vec{\mathcal{E}} \;=\; -\frac{k_0^2}{\epsilon_0 n_0^2}\vec{\mathcal{P}}_{\rm NL}\,, \quad (2.11a)$$

$$\nabla\cdot\vec{\mathcal{E}} \;=\; -\frac{1}{\epsilon_0 n_0^2}\nabla\cdot\vec{\mathcal{P}}_{\rm NL}\,, \quad (2.11b)$$

where $\vec{\mathcal{P}}_{\rm NL}$ is the time-harmonic nonlinear polarization field. For convenience, we refer to the second term on the left-hand side of (2.11a) as the "grad-div term".

In vacuum or when beam intensity is not sufficiently high, $\vec{\mathcal{P}}_{\rm NL}$ is negligible, in which case Eqs. (2.10), (2.11a), and (2.11b) reduce to (2.2), (2.7), and (2.8), respectively. When, however, beam intensity is sufficiently high, such as with intense laser beams, $\vec{\mathcal{P}}_{\rm NL}$ cannot be neglected. In that case, it is customary in Nonlinear Optics to expand $\vec{\mathcal{P}}_{\rm NL}$ in a power series, where each term corresponds to a different power or product of the components of the electric field [11]. The first term in this series, which is called the second-order nonlinear polarization, is a weighted sum of the products of any two components of $\vec{\mathcal{E}}$ (*e.g.*, $\mathcal{E}_1\mathcal{E}_3^*$, $|\mathcal{E}_2|^2$, etc). Because of inversion symmetry, the second-order nonlinear

polarization is zero in isotropic media [11]. Therefore, in the weakly nonlinear regime, *i.e.*, when $\vec{\mathcal{P}}_{\rm NL} \ll \vec{P}_{\rm L}$, the third-order nonlinear polarization is the predominant nonlinear interaction in isotropic media. When the medium is isotropic and homogeneous this nonlinear polarization can be described by the vectorial relation [11, 44, 45]

$$\vec{\mathcal{P}}_{\rm NL}(\vec{\mathcal{E}}) \;=\; \frac{4\epsilon_0 n_0 \bar{n}_2}{1+\gamma}\left[\,(\vec{\mathcal{E}}\cdot\vec{\mathcal{E}}^*)\vec{\mathcal{E}} + \gamma(\vec{\mathcal{E}}\cdot\vec{\mathcal{E}})\vec{\mathcal{E}}^*\,\right] ,$$

where $\bar{n}_2$ is the Kerr coefficient[†] and γ is a constant, whose value depends on the physical origin of the Kerr effect[‡] (see Table 2.1).

It is instructive to rewrite relation (2.12) as

$$\vec{\mathcal{P}}_{\rm NL}(\vec{\mathcal{E}}) \;=\; \frac{4\epsilon_0 n_0 \bar{n}_2}{1+\gamma}\left[|\vec{\mathcal{E}}|^2\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + \gamma\begin{pmatrix}|\mathcal{E}_1|^2 & \mathcal{E}_2\mathcal{E}_1^* & \mathcal{E}_3\mathcal{E}_1^*\\ \mathcal{E}_1\mathcal{E}_2^* & |\mathcal{E}_2|^2 & \mathcal{E}_3\mathcal{E}_2^*\\ \mathcal{E}_1\mathcal{E}_3^* & \mathcal{E}_2\mathcal{E}_3^* & |\mathcal{E}_3|^2\end{pmatrix}\right]\begin{pmatrix}\mathcal{E}_1\\ \mathcal{E}_2\\ \mathcal{E}_3\end{pmatrix} .$$

When $\gamma = 0$ the Kerr effect is only "semi-vectorial", in the sense that $\vec{\mathcal{P}}_{\rm NL} = 4\epsilon_0 n_0 \bar{n}_2 |\vec{\mathcal{E}}|^2 \vec{\mathcal{E}}$. In other words, the semi-vectorial Kerr effect is given by the scalar index of refraction

$$n^2 = n_0^2 + 4 n_0 \bar{n}_2 |\vec{\mathcal{E}}|^2 . \tag{2.12}$$

However, when $\gamma \neq 0$, the Kerr effect is genuinely vectorial, *i.e.*, it cannot be written in the form (2.12).

Kerr mechanism	γ
electrostriction	0
nonresonant electrons	0.5
molecular orientation	3

Table 2.1: Values of γ for common Kerr mechanisms (see, *e.g.*, [11]).

2.3 From vector Helmholtz to NLS

Because of the couplings between the field components, direct analysis or simulations of the vector Helmholtz model (2.11a)–(2.12) is hard. Therefore, it is customary to approximate the vector Helmholtz model with the NLS model, which is much more amenable to analysis and simulations.

[†] See Boyd [11] for the various definitions of the Kerr coefficient.

[‡] Note that γ is positive for most liquids [56, Chp. 17]

We now outline the derivation of the NLS from the vector Helmholtz model. A more detailed derivation is given in Subsection 3.1.1.

We set the coordinate system such that the Kerr medium is in the half-space $z \geq 0$, the beam enters the Kerr medium at $z = 0$ and propagates in the positive z direction (see Figure 2.1). We denote by $\vec{\mathcal{E}}^0(x, y) = \vec{\mathcal{E}}(x, y, z = 0)$ the input beam at the medium's interface.

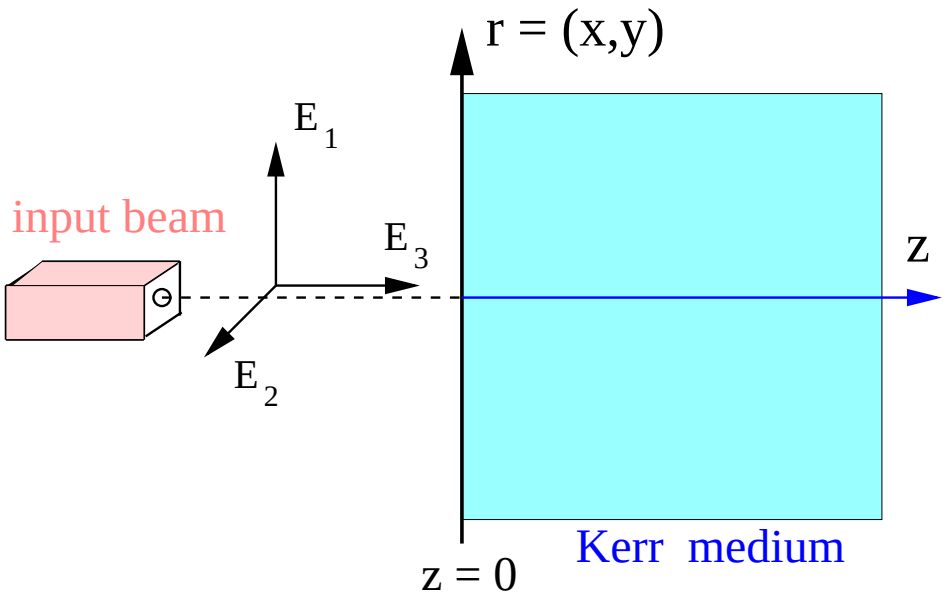

Figure 2.1: The physical setup.

Strictly speaking, when an intense linearly polarized input beam propagates in a Kerr medium the beam does not remain linearly polarized for all $z > 0$, because of the nonlinear polarization couplings in the vector Kerr effect (2.12) and the linear coupling through the grad-div term in Eq. (2.11a). Nevertheless, it is usually assumed that the beam remains (almost) linearly polarized as it propagates inside the Kerr medium, *i.e.*, that $\mathcal{E}_2 \equiv \mathcal{E}_3 \equiv 0$ for $z \geq 0$. In that case, the Kerr effect (2.12) is given by the scalar relation $n^2 = n_0^2 + 4n_0\bar{n}_2|\mathcal{E}_1|^2$. If one further uses the approximation $\nabla(\nabla \cdot \vec{\mathcal{E}}) = 0$, then, to leading order, the vector Helmholtz equation reduces to the scalar nonlinear Helmholtz equation for $\mathcal{E}_1$:

$$\Delta\mathcal{E}_1(x, y, z) + k^2\mathcal{E}_1 = 0\,, \qquad k^2 = k_0^2\left(1 + \frac{4\bar{n}_2}{n_0}|\mathcal{E}_1|^2\right)\,. \tag{2.13}$$

Separating the fast oscillations from the slowly-varying amplitude, *i.e.*, $\mathcal{E}_1 = \mathcal{A}_1(x, y, z)e^{ik_0 z}$, equation (2.13) can be rewritten as

$$\mathcal{A}_{1,zz} + 2ik_0\mathcal{A}_{1,z} + \Delta_\perp\mathcal{A}_1 + \frac{4k_0^2\bar{n}_2}{n_0}|\mathcal{A}_1|^2\mathcal{A}_1 = 0\,, \tag{2.14}$$

where $\Delta_\perp = \partial_{xx} + \partial_{yy}$. The scalar nonlinear Helmholtz equation can be further simplified by using the paraxial approximation $|\mathcal{A}_{1,zz}| \ll |k_0 \mathcal{A}_{1,z}|$. The resulting equation for the beam amplitude $\mathcal{A}_1$ is the (dimensional) NLS

$$2ik_0 \mathcal{A}_{1,z} + \Delta_\perp \mathcal{A}_1 + \frac{4k_0^2 \bar{n}_2}{n_0} |\mathcal{A}_1|^2 \mathcal{A}_1 = 0 \,. \tag{2.15}$$

2.4 Review of NLS theory

We now briefly review several results of NLS theory. For more comprehensive presentations, see [33, 59]. Let ψ be the solution of the dimensionless NLS (see rescaling in Subsection 3.1.1)

$$i\psi_z + \Delta_\perp \psi + |\psi|^2 \psi = 0 \,, \qquad \psi(x, y, z = 0) = \psi_0(x, y) \,. \tag{2.16}$$

Two important invariants of the NLS (2.16) are the *power*

$$N(z) = N(\psi(\cdot, z)) := \frac{1}{2\pi} \int |\psi|^2 \, dxdy \equiv N(0) \tag{2.17}$$

and the *Hamiltonian*

$$H(z) = H(\psi(\cdot, z)) := \frac{1}{2\pi} \left(\int |\nabla_\perp \psi|^2 \, dxdy - \frac{1}{2} \int |\psi|^4 \, dxdy \right) \equiv H(0) \,. \tag{2.18}$$

2.4.1 Variance argument for collapse

Let the variance of the solution be defined as

$$V(z) = \int |\psi|^2 (x^2 + y^2) \, dxdy.$$

Then solutions of the NLS (2.16) satisfy the *variance identity* [62]:

$$V_{zz} = H(z). \tag{2.19}$$

From Hamiltonian conservation (2.18) and (2.19) it follows that $V_{zz} = 8H(0)$. Hence,

$$V(z) = 4H(0)z^2 + V'(0)z + V(0).$$

Therefore, when $H(0) < 0$ then $V(z)$ should vanish at a finite propagation distance $z = Z^*$ and V should become negative afterward. Since, however, V is positive by definition, it follows that the solution ψ must become singular (*i.e.*, collapse) at or before Z^*. This proves that $H(0) < 0$ is a sufficient condition for collapse. We note, however, that the collapse occurs at a distance $Z_{\rm cr} < Z^*$, where the variance is still positive (*i.e.*, partial beam collapse [27]). We also remark that the condition $H(0) < 0$ typically overestimates the actual threshold power for collapse [25].

2.4.2 The Townes Soliton

The NLS has the cylindrically-symmetric waveguide solution

$$\psi = R(r)e^{iz}\,, \qquad r = \sqrt{x^2+y^2}\,,$$

where $R(r)$, the so-called *Townes soliton*, is the ground-state positive solution of

$$\Delta_\perp R(r) - R + R^3 = 0, \qquad R'(0) = 0, \quad \lim_{r\to\infty} R(r) = 0, \tag{2.20}$$

A necessary condition for blowup in the NLS is that the input power be above the critical power N_c [64], *i.e.*, that $N(0) \geq N_c$, where N_c is equal to the power of the Townes soliton

$$N_c = \int_0^\infty R^2\, r dr \;\approx\; 1.86\,. \tag{2.21}$$

In dimensional variables [*i.e.*, for Eq. (2.15)] this threshold power is given by[§]:

$$\mathcal{P}_c = \frac{\lambda^2}{4\pi n_0 n_2} N_c, \tag{2.22}$$

where λ is wavelength [¶].

In addition, the Townes profile has exactly zero Hamiltonian, *i.e.*, $H = 0$. Because the Townes profile satisfies both the necessary and sufficient conditions for collapse, *i.e.*, $N = N_c$ and $H = 0$, it is the borderline case for singularity formation. Hence, if the NLS is solved with the Townes input beam, *i.e.*, $\psi^0(x,y) = R(r)$, then any small numerical perturbations can lead either to collapse or to global existence, which explains why the Townes soliton is an unstable solution.

2.4.3 Critical dimension/exponent

The NLS (2.16) is a special case of the focusing NLS in d dimensions and with a general power-law nonlinearity

$$i\psi_t(\mathbf{x},t) + \Delta_\perp \psi + |\psi|^{2\sigma}\psi = 0, \qquad \psi(\mathbf{x},0) = \psi_0(\mathbf{x}), \tag{2.23}$$

where $\mathbf{x} = (x_1, \ldots, x_d)$, t is time[‖], and $\Delta_\perp = \partial_{x_1 x_1} + \cdots + \partial_{x_d x_d}$.

[§]Expression (2.22) is the same in both mks and cgs units, because n_2 has units of inverse intensity [*i.e.*, $n(I) = n_0 + n_2 I$, where I is intensity]. In mks and cgs units the relations between $\bar{n}_2$ in (2.12) and n_2 in (2.22) are $n_2[\text{mks}] = 2Z_0\bar{n}_2/n_0$, where $Z_0 = 377\Omega$ and $n_2[\text{cgs}] = 4\pi\bar{n}_2/n_0 c$, respectively [11, Appendix].

[¶]The condition $N(0) \geq N_c$ is necessary for blowup but not sufficient, *i.e.*, the actual threshold power for NLS collapse of input beams is generically a few percent above N_c [25].

[‖]In the NLS (2.16) z plays the role of time.

A key question in NLS theory is whether solutions can become singular (blowup) in finite time. It is well known that the answer to this question depends on the product σd as follows. When $\sigma d < 2$ (subcritical NLS), all solutions of (2.23) exist globally. When $\sigma d > 2$ (supercritical NLS) solutions of (2.23) can become singular and this singularity formation is insensitive to small perturbations. The *critical dimension/exponent* $\sigma d = 2$ is a borderline case between global existence and singularity formation. In this case, singularity formation is a very sensitive and even small perturbations can arrest the blowup process. The reason for this sensitivity is that in the critical NLS the Laplacian, which defocuses the solution, and nonlinear term, which focuses the solution, almost balance each other. Consequently, the blowup process is very sensitive and even a small perturbation can shift the balance.

At present, rigorous theory for singularity formation in critical *perturbed* NLS equations is difficult. As a result, in this study we mostly rely on asymptotic and numerical methods. When beam power is not much higher than N_c, the self-focusing dynamics can be analyzed using *modulation theory*, which is an asymptotic theory for analyzing the effects of small perturbations on the critical NLS (2.16) [32, 33]. Modulation theory is based on the observation that, after some propagation has taken place, a self-focusing filament rearranges itself as a modulated Townesian[**], *i.e.*,

$$|\psi(x,y,z)| \sim \frac{1}{L(z)} R(\frac{r}{L(z)}) \, ,$$

where $R(r)$ is defined in Eq. (2.20). Using modulation theory it has been shown, for example, that nonparaxiality, *i.e.*, the $A_{1,zz}$ term in Eq. (2.14), can arrest beam collapse, leading instead to focusing-defocusing oscillations [24]. In Section 3.3 and Section 4.4 we use modulation theory to study the self-focusing dynamics according to perturbed NLS models, in which the small perturbation terms correspond to nonparaxiality and vectorial effects.

2.5 Multiple filamentation and cylindrical symmetry

It is customary to assume that the input beam has a cylindrically-symmetric profile, *i.e.*, $\psi_0(x,y) = \psi_0(r)$. According to the NLS model, beams with cylindrically-symmetric input profile should remain cylindrically symmetric during propagation, because the NLS equation (2.16) is invariant under rotation of the (x,y) plane. In self-focusing experiments, however, a sufficiently intense laser beam can break-up into several long and narrow filaments, a phenomenon known as multiple filamentation [1, 3, 12, 13, 14, 15, 16, 18, 19, 41, 43, 52, 53, 55] (see Figure 1.2). We note that in 1D NLS models

[**] It is essential to use the Townes profile in the asymptotics as opposed, for example,to a Gaussian profile, because only the Townes profile captures the delicate balance between the diffraction and the nonlinearity [27].

(*e.g.*, propagation in fibers) "multiple filamentation" can occur *without* breakup of spatial-symmetry[††]. In contrast, in 2D NLS models for beam propagation in a bulk Kerr medium, multiple filamentation *cannot* occur without breakup of cylindrical symmetry (see Figure 2.2). Therefore, in order to explain the phenomenon of multiple filamentation, where cylindrical symmetry is clearly lost, one has to add a symmetry-breaking mechanism to the NLS model. It is therefore natural to ask what is the symmetry-breaking mechanism that is responsible for multiple filamentation.

The standard explanation for multiple filamentation (see Subsection 3.4.2), is that breakup of cylindrical symmetry is initiated by *random* noise in the input beam. In Chapter 3 we show that in the case of *linearly polarized* input beams the *deterministic* vectorial effects can also lead to multiple filamentation. Therefore, in general, multiple filamentation can be caused either by breakup of cylindrical-symmetry in the input profile (*e.g.*, by noise) or by the preferred direction that is induced by the input polarization state (*e.g.*, the direction of the input linear polarization). In contrast, in Chapter 4 we show that a perfect *circularly polarized* cylindrically-symmetric input beam will not undergo (deterministic) multiple filamentation.

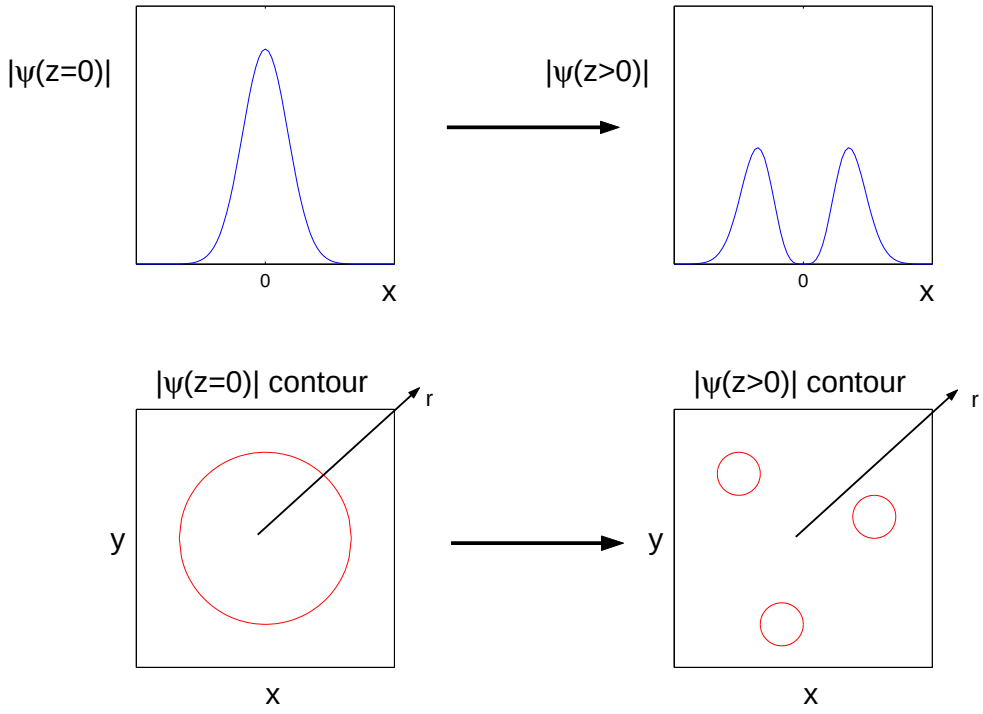

Figure 2.2: Spatial symmetry can be maintained in 1D multiple filamentation (top) but not in 2D multiple filamentation (bottom).

[††]For instance, when two identical 1D solitons move toward each other and overlap at $z = 0$ then at $z = 0$ the input intensity is spatially symmetric, yet, according to the 1D NLS, there would be two distinct solitons at a sufficiently large propagation distance.

Chapter 3

Linearly polarized input beams

In this chapter we analyze the role of vectorial effects in self-focusing of *linearly polarized* input beams. In Section 3.1 we approximate the vector Helmholtz equation with a scalar PDE, which describes self-focusing in the presence of vectorial and nonparaxial effects. Because the perturbation analysis is involved and because other studies of vectorial effects on self-focusing have obtained different equations from ours, we present a careful and systematic derivation. In particular, we derive the scalar equation using two different methods, namely, perturbation analysis at the equation level and at the Lagrangian level, and verify that the two methods give consistent results.

The new model equation leads to a novel explanation for multiple filamentation based on *deterministic* vectorial effects. In Section 3.2 we present numerical simulations that show that the preferred direction induced by linearly polarized beams can lead to multiple filamentation, even of cylindrically symmetric input beams ("deterministic multiple filamentation"). Note that since 1966 the standard (and only) explanation for multiple filamentation was that it is caused by *random* noise in the input beam.

In Section 3.3 we discuss the case where the input power $N(0)$ is only moderately above the critical power N_c. We use modulation theory to reduce the scalar PDE to an ODE for self-focusing dynamics of a single filament in the presence of vectorial and nonparaxial effects. We show that vectorial effects can arrest beam collapse and, as a result, the NLS model with vectorial effects can be used to describe beam propagation beyond the blowup point of the NLS. We also show that vectorial effects have a larger effect than nonparaxiality.

In Section 3.4 we discuss the possibility that noise in the input beam leads to multiple filamentation ("random multiple filamentation"). In Section 3.5 we compare the noise and the vectorial effects explanations for multiple filamentation. In Section 3.6 we suggest a simple experiment for deciding whether multiple filamentation is due to vectorial effects or not.

3.1 Scalar equations for linearly polarized beams

In this section we identify and use the small parameter f of the model to reduce the vector Helmholtz equation (2.11a) for linearly polarized input beams to a scalar equation(s) that takes vectorial and non-paraxial effects into account. The results in this section are as follows (see also Figure 3.1). In Subsection 3.1.1 we apply perturbation analysis to the vector Helmholtz equation (2.11) and derive the scalar equation (3.7). In Subsection 3.1.2 we apply a similar perturbation analysis to the Lagrangian (3.11) of the vector Helmholtz equation, resulting in the Lagrangian (3.12), whose variational derivative leads to the scalar equation (3.13). Although Eqs. (3.7) and (3.13) are not identical, the two equations agree with each other to the order of accuracy of their derivation (Subsubsection 3.1.2).

Chi and Guo were the first to obtain a scalar equation that takes vectorial effects on self-focusing into account [17]. Subsequent studies by Ciattoni, DiPorto, Crosignani, and Yariv [20, 22], Blair and Wagner [7], de la Fuente, Varela, and Michinel [23], and Malomed, Boardman, Marinov, Pushkarov, and Shivarova [9, 46] used similar perturbation analysis to obtain scalar a equation with vectorial effects for 1D propagation dynamics. However, all the scalar equations obtained in [7, 9, 17, 20, 22, 23, 46] differ from Eqs. (3.7), (3.13), [and from (6.2) in Section 6.2], as well as from each other, in the $O(f^2)$ terms. Because of that, we present in this study a careful systematic derivation of the equations. The agreement of Eqs. (3.7) and (3.13), which are derived independently, provides further support that these equations are indeed correct.

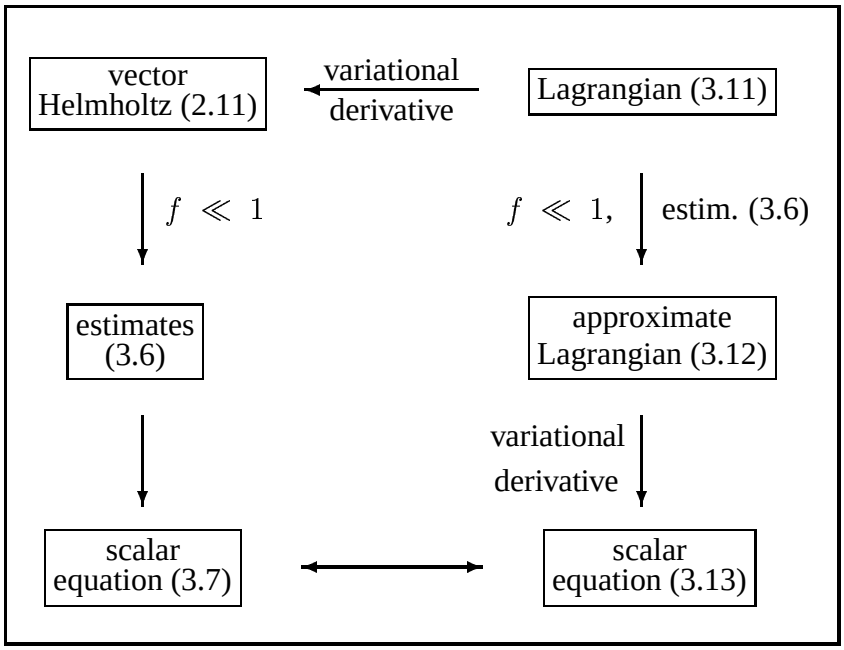

Figure 3.1: Schematic outline of the derivation in Section 3.1.

3.1.1 From vector Helmholtz to a scalar equation

The key dimensionless parameter of the model is

$$f = \frac{1}{k_0 r_0} = \frac{\lambda}{2\pi r_0} \,, \tag{3.1}$$

where λ is wavelength and r_0 is input-beam width. Since the wavelength is much smaller than the input-beam width, the parameter f is small, *i.e.*,

$$f \ll 1 \,. \tag{3.2}$$

The existence of a small parameter enables us to apply perturbation analysis on Eq. (2.11) and reduce it to a scalar equation. To do that, we first use Eq. (2.11b) to rewrite Eq. (2.11a) as

$$\Delta\vec{\mathcal{E}} + k_0^2\vec{\mathcal{E}} + \frac{k_0^2}{\epsilon_0 n_0^2}\vec{\mathcal{P}}_{\rm NL} = -\frac{1}{\epsilon_0 n_0^2}\nabla(\nabla\cdot\vec{\mathcal{P}}_{\rm NL}) \,. \tag{3.3}$$

We rescale the variables according to

$$\tilde{x} \;=\; \frac{x}{r_0} \,, \quad \tilde{y} \;=\; \frac{y}{r_0} \,, \quad \tilde{z} \;=\; \frac{z}{2L_{\rm DF}} \,, \quad \vec{\mathcal{E}} \;=\; \frac{1}{2r_0 k_0}\sqrt{\frac{n_0}{\bar{n}_2}}\vec{A}(x,y,z)e^{ik_0 z} \,, \tag{3.4}$$

where $\vec{A} = (A_1, A_2, A_3)$ is the dimensionless electric-field amplitude and $L_{\rm DF} = k_0 r_0^2$ is the diffraction length. For convenience we drop the tilde signs from now on.

Substitution of Eqs. (3.4) in the vector Helmholtz system leads to the dimensionless system

$$\begin{aligned} i\vec{A}_{,z} + \Delta_\perp\vec{A} + \frac{1}{4}f^2\vec{A}_{,zz} + \vec{N} = & \qquad (3.5a)\\ & -\left[f\nabla_\perp + \hat{e}_3\left(i + \frac{1}{2}f^2\partial_z \right)\right]\left[f\nabla_\perp\cdot\vec{N} + iN_3 + \frac{1}{2}f^2 N_{3,z} \right] \,, \\ f\nabla_\perp\cdot\vec{A} + iA_3 + \frac{1}{2}f^2 A_{3,z} \;=\; & -f^2\left(f\nabla_\perp\cdot\vec{N} + iN_3 + \frac{1}{2}f^2 N_{3,z} \right) \,, \qquad (3.5b) \end{aligned}$$

where

$$\vec{P}(\vec{A}) \;:=\; \frac{1}{1+\gamma}\left[(\vec{A}\cdot\vec{A}^*)\vec{A} + \gamma(\vec{A}\cdot\vec{A})\vec{A}^* \right] \,, \tag{3.5c}$$

$\hat{e}_3 = (0,0,1)$, and $\nabla_\perp = (\partial_x, \partial_y, 0)$. Eqs. (3.5a)–(3.5c) correspond to Eqs. (3.3)–(2.12), respectively.

Using a careful perturbation analysis, we show in Appendix A.1 that over propagation distances of several diffraction lengths the dimensionless amplitudes satisfy

$$\begin{aligned} A_1 \;&=\; O(1) \,, \\ A_2 \;&=\; O(f^2) \,, \end{aligned} \tag{3.6a}$$

and

$$A_3 \;=\; if A_{1,x} + O(f^3)\,. \tag{3.6b}$$

These relations provide a quantitative measure of the degree to which the beam remains linearly polarized. In particular, relations (3.6a) and (3.6b) indicate that the second transverse component $\mathcal{E}_2$ is significantly smaller than the axial component $\mathcal{E}_3$. The fact that $\mathcal{E}_2/\mathcal{E}_1 = O(f^2)$ plays an important role in the perturbation analysis, because it means that $\mathcal{E}_2$ does not contribute to the perturbation terms in the scalar equations for A_1 (see Appendix A.2).

Remark. Lax, Louisell, and McKnight [42] studied the case were the Kerr effect is "semi-vectorial", *i.e.*, that the refractive index n is given by a relation of the form (2.12). Using perturbation analysis they derived the relation

$$A_3 \approx if(A_{1,x} + A_{2,y})\,,$$

which, upon assuming that $A_2 \equiv 0$, they approximated with (3.6b). Thus, the derivation in Appendix A.1 improves on [42] in that relation (3.6a) is proved for a "genuine" vectorial Kerr effect and then used to show that relation (3.6b) remains valid even when $A_2 \not\equiv 0$.

Substituting relations (3.6a) and (3.6b) in the vector Helmholtz equation (3.5a) leads to the following scalar equation for A_1 (Appendix A.2):

Proposition 3.1. *When $f \ll 1$, Eqs. (2.11)–(2.12) can be approximated with the scalar equation*

$$\underbrace{iA_{1,z} + \Delta_\perp A_1 + |A_1|^2 A_1}_{\text{NLS}} = -f^2 \Bigg[\underbrace{\frac{1}{4} A_{1,zz}}_{\text{nonparax}} + \underbrace{\frac{4+6\gamma}{1+\gamma}|A_{1,x}|^2 A_1 + (A_{1,x})^2 A_1^* + \frac{1+2\gamma}{1+\gamma}\left(|A_1|^2 A_{1,xx} + A_1^2 A_{1,xx}^* \right)}_{\text{vectorial}} \Bigg] + O(f^4)\,. \tag{3.7}$$

When $f = 0$, Eq. (3.7) reduces to the NLS (2.16). The $A_{1,zz}$ term is the nonparaxial term, which comes from the *scalar* Helmholtz equation (2.13). The remaining terms on the right-hand-side correspond to "vectorial effects", *i.e.*, they result from the combined effects of the linear and nonlinear coupling between $\mathcal{E}_1$ and $\mathcal{E}_3$ in Eqs. (2.11)–(2.12). Note that $\mathcal{E}_2$ does not contribute to the vectorial terms in Eq. (3.7), because, in light of Eq. (3.6a), its effect is only $O(f^4)$.

We note that the "vectorial" terms on the right-hand-side of Eq. (3.7) do not vanish even if one assumes that $\mathcal{E}_2 \equiv \mathcal{E}_3 \equiv 0$ in the derivation. Indeed, a close inspection* of the derivation of Eq. (3.7)

*See Eq. (A.12) in Appendix A.2.

shows that if we assume that $\mathcal{E}_2 \equiv \mathcal{E}_3 \equiv 0$, the resulting equation for A_1 is given by

$$iA_{1,z} + \Delta_\perp A_1 + |A_1|^2 A_1 = -f^2 \frac{1}{4} A_{1,zz} - \underbrace{f^2 (|A_1|^2 A_1)_{xx}}_{\text{grad-div } E_1} , \tag{3.8}$$

where the second term on the right-hand side corresponds to the contribution of $\mathcal{E}_1$ in the grad-div term in the vector Helmholtz equation (2.11a). Note that this term can also be written as

$$-\underbrace{f^2 (|A_1|^2 A_1)_{xx}}_{\text{grad-div } E_1} \equiv -f^2 \left[4|A_{1,x}|^2 A_1 + 2(A_{1,x})^2 A_1^* + 2|A_1|^2 A_{1,xx} + A_1^2 A_{1,xx}^* \right] .$$

Thus, strictly speaking, the vectorial-effects mechanism also corresponds, in part, to a scalar mechanism. The remaining vectorial terms on the right-hand side of Eq. (3.7) correspond to the coupling to $\mathcal{E}_3$. If only those terms are taken into account (*i.e.*, neglecting nonparaxial and the grad-div-E_1 terms) the resulting equation is

$$iA_{1,z}+\Delta_\perp A_1+|A_1|^2 A_1 = \underbrace{-\frac{f^2}{1+\gamma} \left[2\gamma |A_{1,x}|^2 A_1 - (1+\gamma)(A_{1,x})^2 A_1^* - |A_1|^2 A_{1,xx} + \gamma A_1^2 A_{1,xx}^* \right]}_{\text{coupling to } \mathrm{E_3}} . \tag{3.9}$$

To recap, the vectorial terms on the right-hand side of (3.7) correspond to the contribution of the grad-div-E_1 term and the coupling to $\mathcal{E}_3$:

$$\underbrace{\frac{4+6\gamma}{1+\gamma}|A_{1,x}|^2 A_1 + (A_{1,x})^2 A_1^* + \frac{1+2\gamma}{1+\gamma} \left(|A_1|^2 A_{1,xx} + A_1^2 A_{1,xx}^* \right)}_{\text{vectorial}} \tag{3.10}$$

$$\equiv \underbrace{(|A_1|^2 A_1)_{xx}}_{\text{grad-div } E_1} + \underbrace{\frac{1}{1+\gamma} \left[2\gamma |A_{1,x}|^2 A_1 - (1+\gamma)(A_{1,x})^2 A_1^* - |A_1|^2 A_{1,xx} + \gamma A_1^2 A_{1,xx}^* \right]}_{\text{coupling to } \mathrm{E_3}} .$$

3.1.2 A variational approach

We now give an alternative derivation of a scalar equation for A_1 , based on the Lagrangian of the vector Helmholtz equation (2.11a). It is straightforward to verify that the vector Helmholtz equation (2.11a) has the Lagrangian density

$$\mathcal{L}_{\mathrm{VH}}(\vec{\mathcal{E}}, \vec{\mathcal{E}}^*) = \mathcal{E}_{k,k}\mathcal{E}_{j,j}^* - \mathcal{E}_{j,k}\mathcal{E}_{j,k}^* + k_0^2 \mathcal{E}_k \mathcal{E}_k + \frac{2k_0^2 \bar{n}_2}{n_0(1+\gamma)} \left(\delta_i^l \delta_j^k + \gamma \delta_i^k \delta_j^l \right) \mathcal{E}_i \mathcal{E}_j^* \mathcal{E}_k \mathcal{E}_l^* , \tag{3.11}$$

where

$$\mathcal{E}_{i,k} := \frac{\partial \mathcal{E}_i}{\partial x_k} , \qquad \delta_i^j := \begin{cases} 1 , & i = j \\ 0 , & i \neq j . \end{cases}$$

The Lagrangian in (3.11) is written using the standard summation convention. Alternatively, $\mathcal{L}_{\rm VH}$ can also be rewritten as

$$\mathcal{L}_{\rm VH}(\vec{\mathcal{E}},\vec{\mathcal{E}}^*) = -\sum_{\substack{i,j=1\\ i\neq j}}^{3} |\mathcal{E}_{i,j}|^2 + k_0^2|\vec{\mathcal{E}}|^2 + \frac{2k_0^2\bar{n}_2}{n_0(1+\gamma)}\left[\,(\vec{\mathcal{E}}\cdot\vec{\mathcal{E}}^*)^2 + \gamma(\vec{\mathcal{E}}\cdot\vec{\mathcal{E}})^2\,\right] .$$

Clearly, the quadratic and the quartic terms of $\mathcal{L}_{\rm VH}$ correspond to the linear and nonlinear terms in the vector Helmholtz equation, respectively.

Using the estimates (3.6a) and (3.6b), we show in Appendix A.3 that $\mathcal{L}_{\rm VH}$ can be approximated with terms that depend only on A_1:

$$\begin{aligned}\mathcal{L}_{A_1}(A_1,A_1^*) = \frac{i}{2}\left(\,A_1^*A_{1,z} - A_1A_{1,z}^*\,\right) - |A_{1,x}|^2 - |A_{1,y}|^2 + \frac{1}{2}|A_1|^4 \\ + f^2\left\{-\frac{1}{4}|A_{1,z}|^2 - \,|A_{1,xx}|^2 - |A_{1,xy}|^2 + \frac{i}{2}\left(\,A_{1,xz}A_{1,x}^* - A_{1,xz}^*A_{1,x}\,\right)\right. \\ \left. + \frac{1}{2(1+\gamma)}\left[\,2|A_1|^2|A_{1,x}|^2 - \gamma\left(\,(A_{1,x})^2A_1^{*2} + (A_{1,x}^*)^2A_1^2\,\right)\right]\right\} \;+\; O(f^4)\,.\end{aligned} \tag{3.12}$$

Taking the variational derivative of $\mathcal{L}_{A_1}$ leads to the following equation:

$$\begin{aligned}iA_{1,z} + \Delta_\perp A_1 + |A_1|^2A_1 = f^2\left\{-\frac{1}{4}A_{1,zz} + \partial_{xx}\left(\,iA_{1,z} + \Delta_\perp A_1\,\right)\right. \\ \left. - \frac{1}{1+\gamma}\left[\,2\gamma|A_{1,x}|^2A_1 - (1+\gamma)(A_{1,x})^2A_1^* - |A_1|^2A_{1,xx} + \gamma A_1^2A_{1,xx}^*\,\right]\right\} + O(f^4)\,.\end{aligned} \tag{3.13}$$

Thus, the scalar equation (3.13) for self-focusing in the presence of vectorial and nonparaxial effects preserves the Lagrangian structure of the vector Helmholtz equation (2.11a).

Invariants

We can use Noether's Theorem to find conservation laws for Eq. (3.13). Invariance of the action $\int \mathcal{L}_{A_1}\,dxdydz$ under phase change $A_1 \to A_1e^{i\varepsilon}$ leads to

$$\int\left(\,|A_1|^2 + f^2|A_{1,x}|^2\,\right)\,dxdy - \frac{i}{4}f^2\int\left(\,A_1A_{1,z}^* - A_1^*A_{1,z}\,\right)\,dxdy \equiv const\,.$$

This relation can also be obtained by multiplying (3.13) by A_1^*, subtracting the complex-conjugate and integrating over the (x,y) plane. Clearly, when $f=0$ this relation reduces to Eq. (2.17). The term $f^2|A_{1,x}|^2$ corresponds to vectorial effects. Indeed, from (3.6a) and (3.6b) we have that

$$\int|\vec{A}|^2\,dxdy = \int\left(\,|A_1|^2 + |A_2|^2 + |A_3|^2\,\right)\,dxdy = \int\left(\,|A_1|^2 + f^2|A_{1,x}|^2\,\right)\,dxdy\;+O(f^4)\,,$$

which has the meaning of total beam power. The right integral corresponds to nonparaxiality, *i.e.*, it comes from balance of power in the scalar Helmholtz equation (2.13).

Invariance of the action under "time dilation", *i.e.*, $z \to z + \varepsilon$, leads to

$$\int \left(|A_{1,x}|^2 + |A_{1,y}|^2 - \frac{1}{2}|A_1|^4 \right) dxdy - f^2 \int \left\{ -\frac{1}{4}|A_{1,z}|^2 + |A_{1,xx}|^2 + |A_{1,xy}|^2 \right.$$
$$\left. -\frac{1}{2(1+\gamma)} \left[2|A_1|^2|A_{1,x}|^2 - \gamma \left((A_{1,x})^2 A_1^{*2} + (A_{1,x}^*)^2 A_1^2 \right) \right] \right\} dxdy \equiv const .$$

This relation can also be obtained by multiplying (3.13) by $A_{1,z}^*$, adding the complex-conjugate and integrating over the (x, y) plane.

Consistency of equations (3.7) and (3.13)

The two scalar equations for A_1 , Eqs. (3.7) and (3.13), are derived from the same 'mother' equation and by using the same assumptions and approximations. In both Eqs. (3.7) and (3.13), the $O(1)$ and nonparaxial terms are the same, but the $O(f^2)$ terms that come from vectorial effects are not. This apparent inconsistency can be resolved by showing that the two equations differ only by $O(f^4)$ terms, which is the order of accuracy of these equations:

Proposition 3.2. *Equations (3.7) and (3.13) differ only by $O(f^4)$ terms.*

Proof: see Appendix A.4.

Indeed, we checked that applying the same perturbation analysis to the vector Helmholtz equation in the form (2.11a), rather than on the form (3.3), leads to Eq. (3.13), rather than to Eq. (3.7). The consistency of the derivations is summarized in Figure 3.1.

3.2 Deterministic multiple filamentation

The standard explanation for multiple filamentation (see Section 2.5 and Subsection 3.4.2), is that breakup of cylindrical symmetry is initiated by *random* noise in the input beam. We now show that *deterministic* vectorial effects can also lead to multiple filamentation.

In order to understand why vectorial effects might lead to multiple filamentation, we note that the asymmetry in the x and y derivatives of the vectorial perturbation terms in either Eq. (3.7), (3.13), or (6.2) implies that vectorial effects are a symmetry-breaking mechanism. Clearly, this, by itself, does not imply that vectorial effects lead to multiple filamentation. However, the following simulations show that when the input beam is sufficiently powerful, vectorial effects do lead to multiple filamentation (see Chapter 6 for details about the numerical methods).

We note that at present there is no theory for the NLS in the high-power regime $N(0) \gg N_c$. Therefore, our results in this high-power regime on vectorial effects in general, and on multiple filamentation in particular, are only numerical. The arrest of collapse and the focusing-defocusing oscillations that are observed in the following simulations have some theoretical basis, which is discussed in Section 3.3.

We begin by presenting the results of simulations of Eq. (6.2) with $f = 0.05$ and $\gamma = 1/2$, where we gradually increase the input power $N(0)$. The input beam is a cylindrically-symmetric Gaussian,

$$A_1(x, y, z = 0) = 2\sqrt{N(0)}\, e^{-(x^2+y^2)} , \tag{3.14}$$

where the constant $N(0)$ is equal to the input power of A_1 .

When $N(0) = 2N_c$, the beam propagates as a single filament which undergoes focusing-defocusing oscillations (Figs. 3.2 and 3.3). Although the beam appears to be symmetric during its propagation, a more careful inspection shows a small deviation from cylindrical symmetry.

When $N(0) = 3N_c$, the beam initially goes through the following stages: (*i*) self-focusing, (*ii*) defocusing into a symmetric ring (crater), (*iii*) second self-focusing, (*iv*) defocusing and formation of two small sub-peaks, and (*v*) focusing with a single peak (Figure 3.4). During further propagation, the beam undergoes focusing-defocusing oscillations (Figure 3.5).

When the input power is raised to $N(0) = 3.75N_c$, the beam initially self-focuses and defocuses into an asymmetric ring with two peaks on its rim (Figure 3.6). After the second focusing-defocusing cycle, a complete breakup of cylindrical symmetry occurs as the beam splits into two filaments. Shortly after, however, the two filaments reunite and continue to propagate as a single filament, as can be seen in Figure 3.7. We call this phenomenon *pseudo multiple-filamentation*, in order to distinguish it from (genuine) multiple filamentation, in which the filaments do not reunite.

As the input power is further increased to $N(0) = 4N_c$, a similar dynamics leads to the emergence of two filaments (Figure 3.8). This time, however, the two filaments do not reunite. Rather, they propagate forward in the z direction, while moving away from each other along the x axis (Figure 3.9). When the power is increased to $N(0) = 10N_c$, the beam goes through the same stages, but in this case the two filaments move away from each other along the y axis (Figs. 3.10 and 3.11).

We note that vectorial effects induce a preferred direction in the transverse (x, y) plane: The direction on initial polarization (the x-axis direction in our model). Therefore, in the case of cylindrically-symmetric input beams, when vectorial effects lead to breakup of the beam into two filaments, the two filaments can move away from each other (in the transverse plane) either along the direction of initial polarization (as in Figure 3.9) or perpendicular to it (as in Figure 3.11)

Finally, with $N(0) = 20N_c$ we observe multiple filamentation into five filaments: One that continues to propagate along the z-axis and four other filaments that propagate slightly 'sideways' (Figs. 3.12 and 3.13).

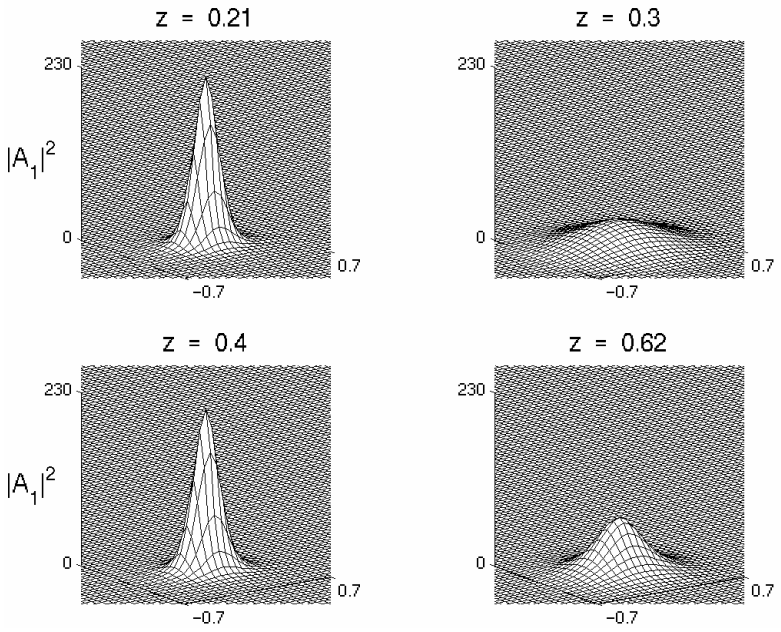

Figure 3.2: Solution of Eq. (6.2) with the input beam (3.14), $f = 0.05$, $\gamma = 0.5$, and $N(0) = 2N_c$.

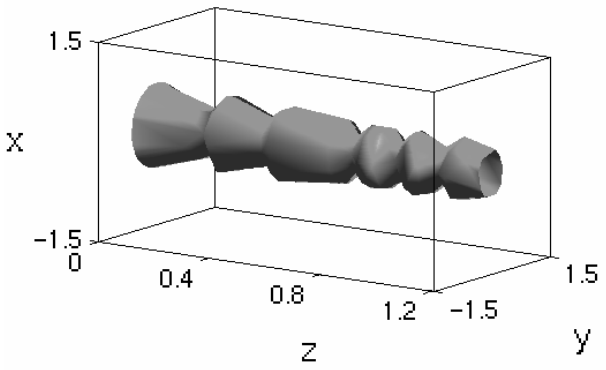

Figure 3.3: Iso-surface $|A_1|^2 \equiv 7$ of the data in Figure 3.2.

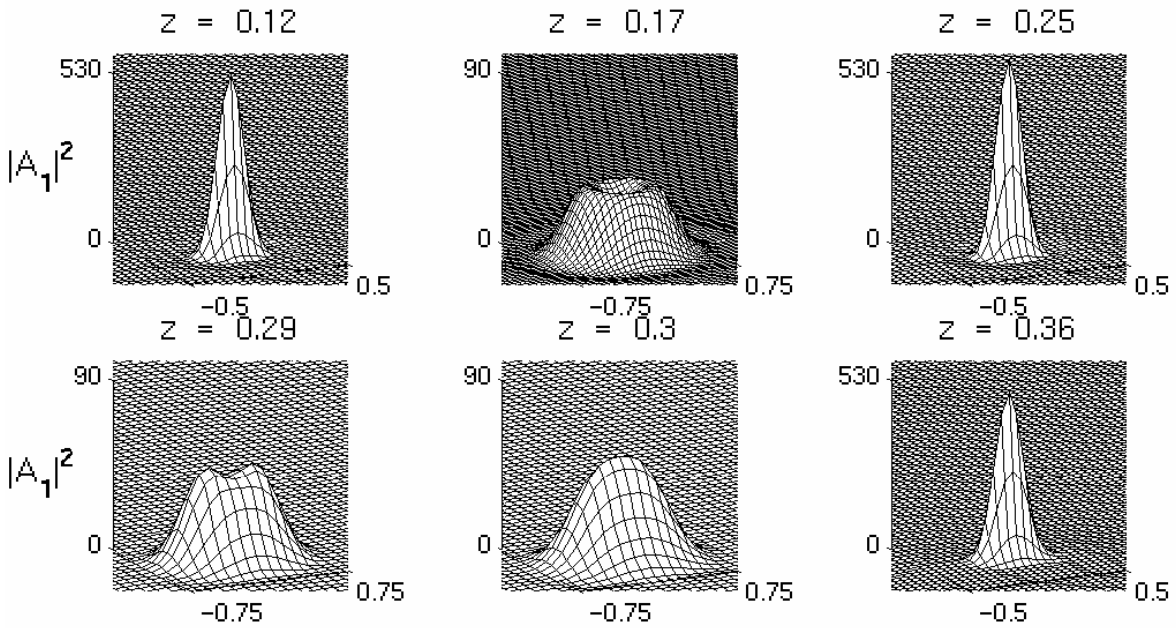

Figure 3.4: Same as Figure 3.2 with $N(0) = 3N_c$.

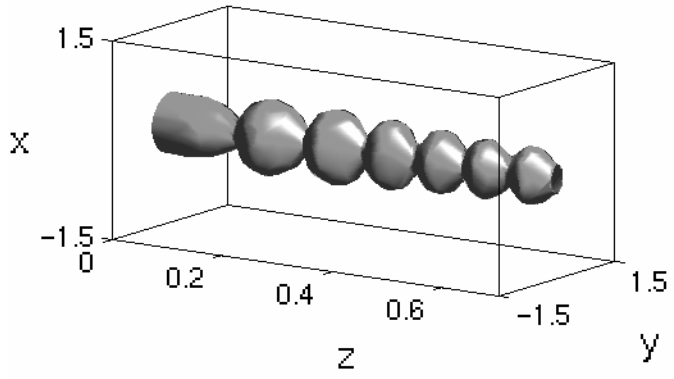

Figure 3.5: Iso-surface $|A_1|^2 \equiv 7$ of the data in Figure 3.2.

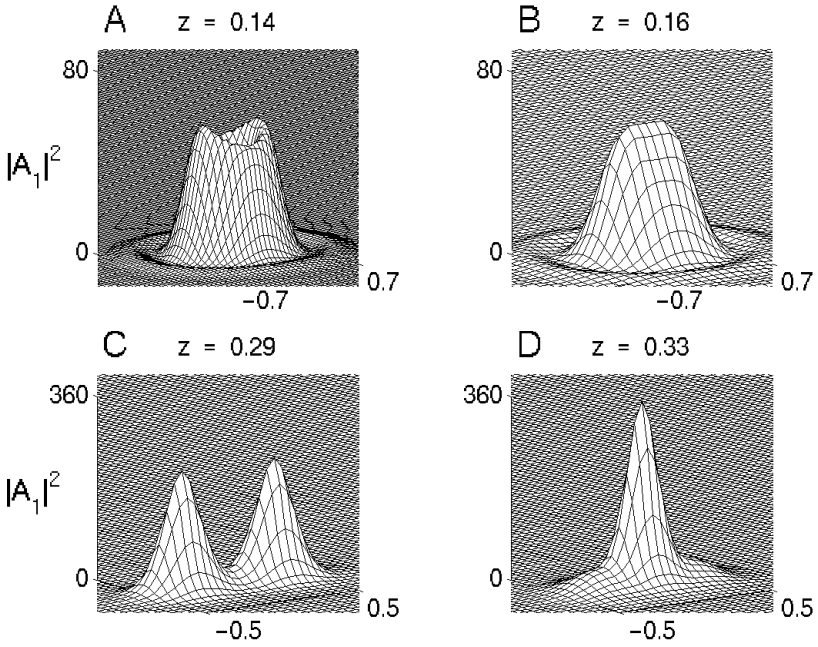

Figure 3.6: Same as Figure 3.2 with $N(0) = 3.75N_c$.

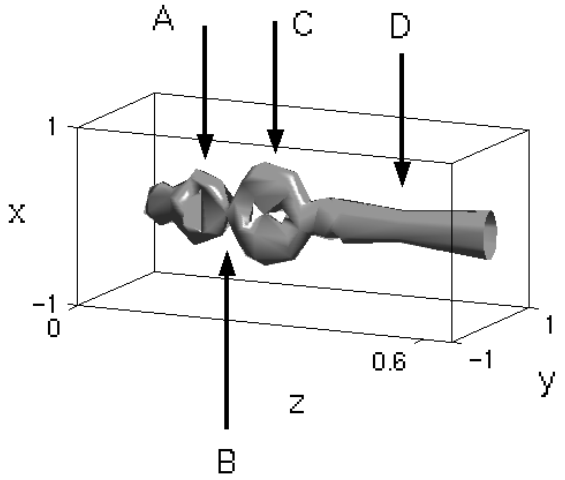

Figure 3.7: Iso-surface $|A_1|^2 \equiv 48$ of the data in Figure 3.6. Capital letters to mark corresponding z-slices on the 3D plots in Figure 3.6.

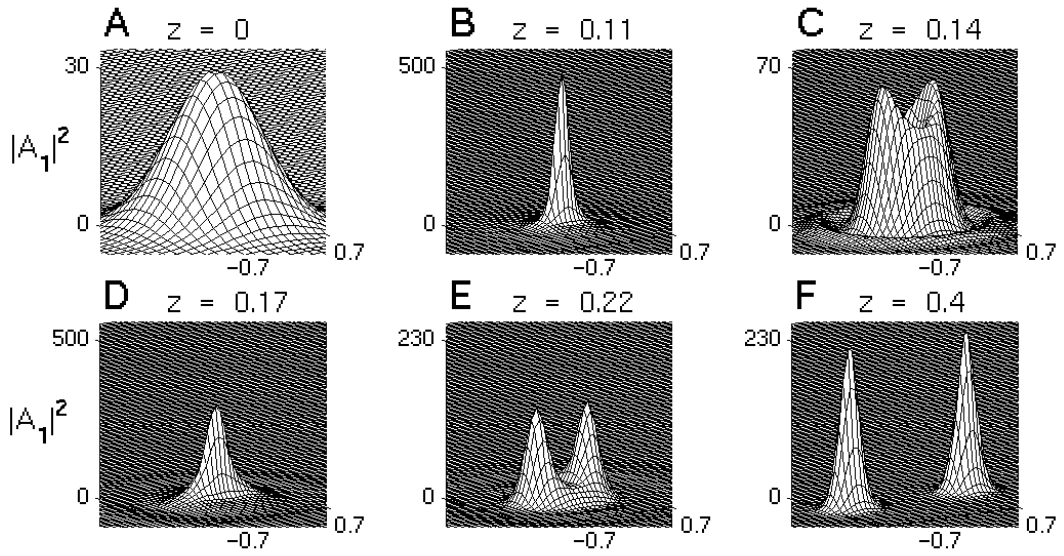

Figure 3.8: Same as Figure 3.2 with $N(0) = 4N_c$. Viewing angle in the (x, y) plane is $-35°$.

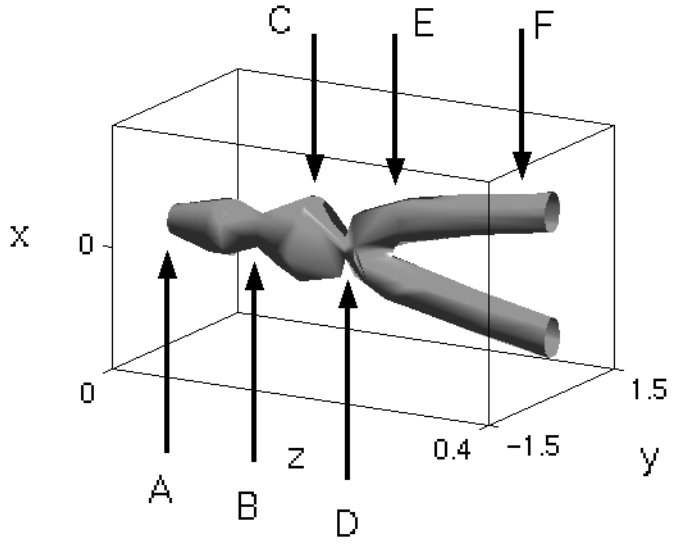

Figure 3.9: Iso-surface $|A_1|^2 \equiv 30$ of the data in Figure 3.6. Capital letters to mark corresponding z-slices on the 3D plots in Figure 3.8.

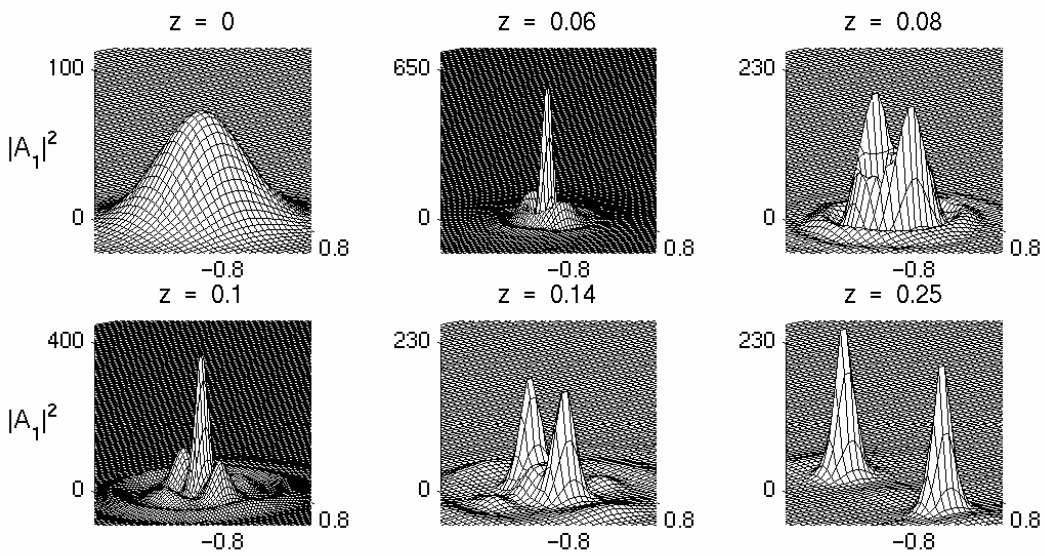

Figure 3.10: Same as Figure 3.2 with $N(0) = 10N_c$.

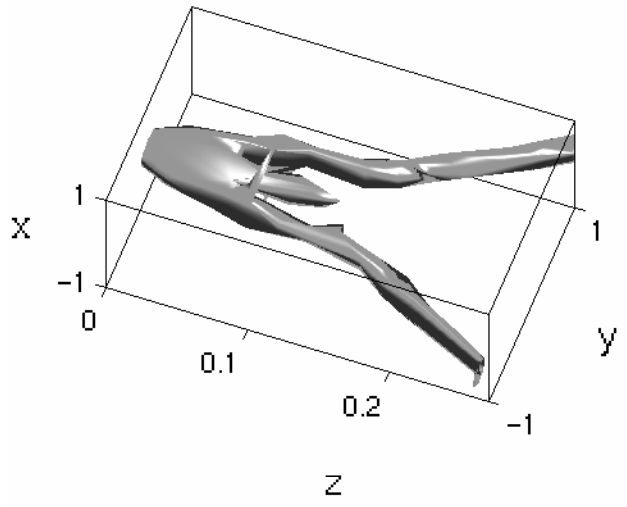

Figure 3.11: Iso-surface $|A_1|^2 \equiv 70$ of the data in Figure 3.10.

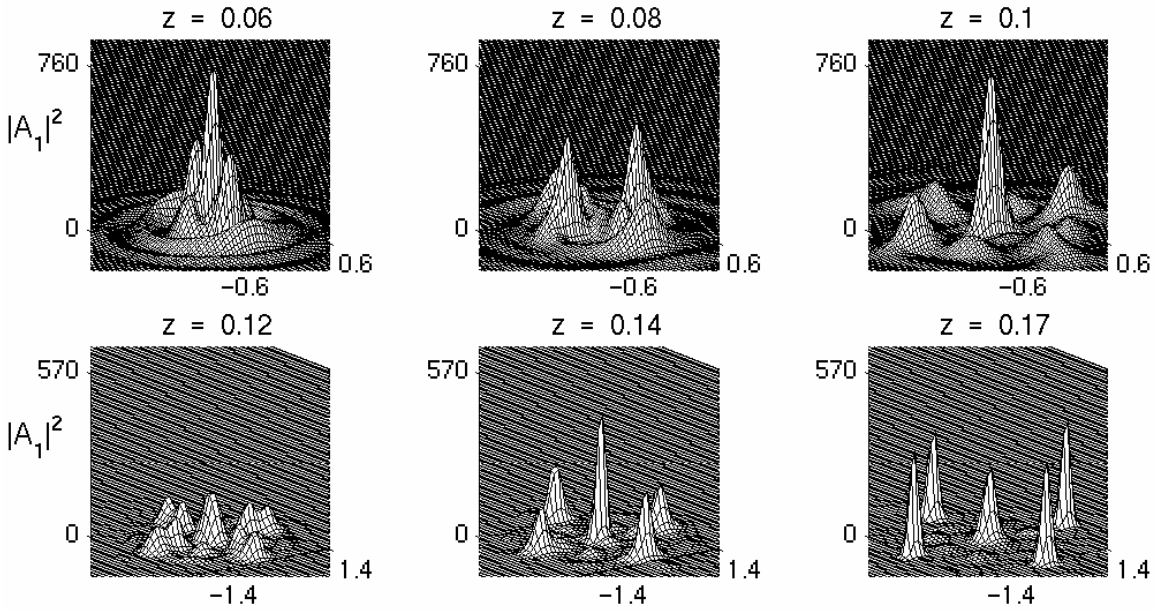

Figure 3.12: Same as Figure 3.2 with $N(0) = 20N_c$. Viewing angle in the (x, y) plane is $-35°$.

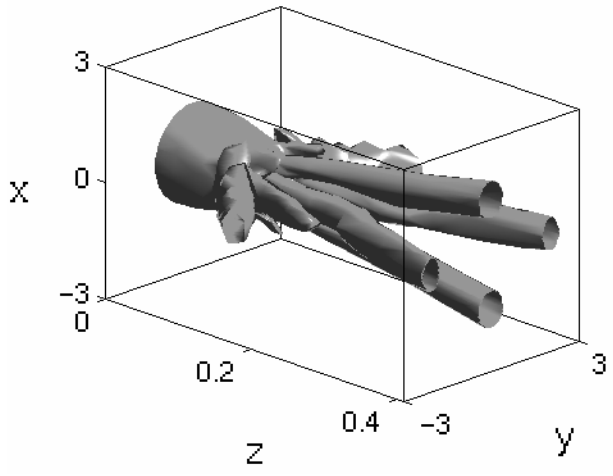

Figure 3.13: Iso-surface $|A_1|^2 \equiv 50$ of the data in Figure 3.10.

We now vary some other parameters in the simulations. In Figure 3.14 we add a focusing lens at the medium's interface $z = 0$ to the input beam of Figure 3.6, *i.e.*, the initial condition is

$$A_1(x, y, z = 0) = 2\sqrt{2N(0)}\, e^{-(x^2+y^2)} e^{-i(x^2+y^2)/4} , \tag{3.15}$$

and $N(0) = 3.75N_c$. In this simulation we observe the same qualitative dynamics as with the unfocused beam, except that the pseudo multiple-filamentation stage is much longer.

In Figs. 3.15 and 3.16 we repeat the simulation of Figure 3.8 with $f = 0.1$. In this case we observe pseudo multiple-filamentation, rather than a genuine one. We do, however, observe genuine multiple filamentation for $f = 0.1$ when $N(0) = 5N_c$. Therefore, the threshold power for multiple filamentation is higher for $f = 0.1$ than for $f = 0.05$. This result is surprising, since a larger f corresponds to stronger vectorial effects. In Figs. 3.17 and 3.18 we repeat the simulations of Figs. 3.12 and 3.13 with $f = 0.08$. In this case, the beam splits into four filaments, all of which propagate slightly off the z-axis.

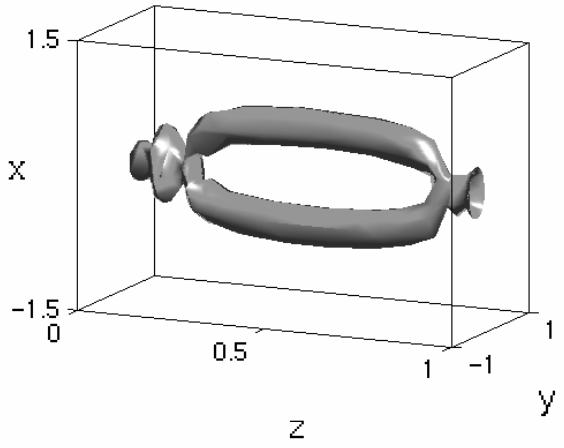

Figure 3.14: Iso-surface $|A_1|^2 \equiv 45$ of the same data as in Figure 3.6 with a focusing lens [*i.e.*, the initial condition (3.15)].

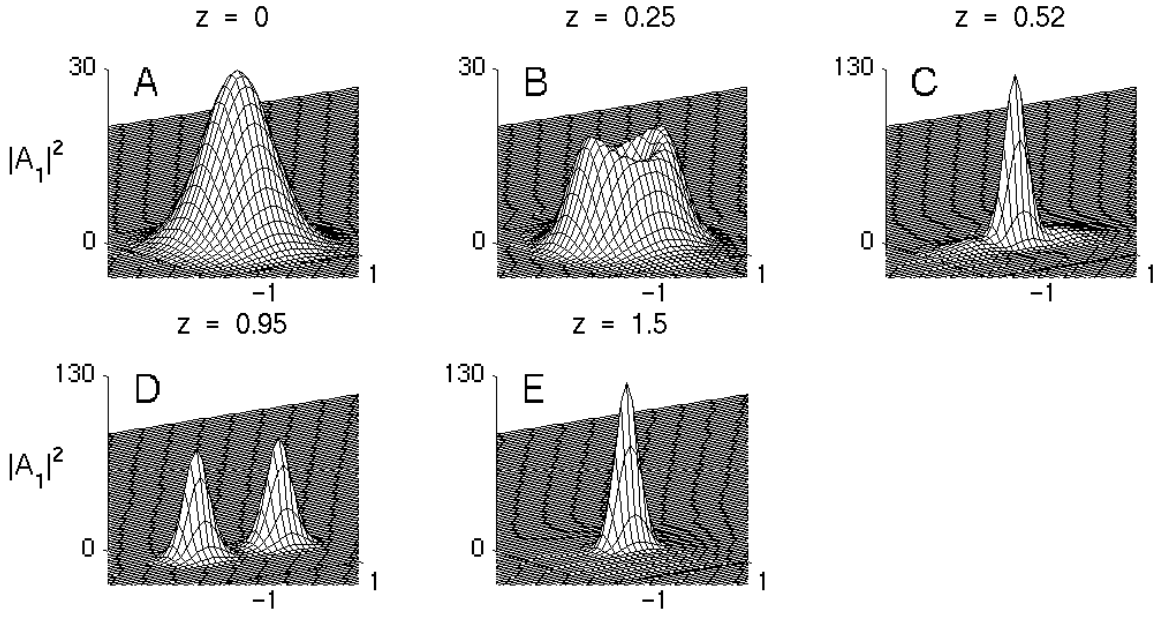

Figure 3.15: Same as Figure 3.8 with $f = 0.1$.

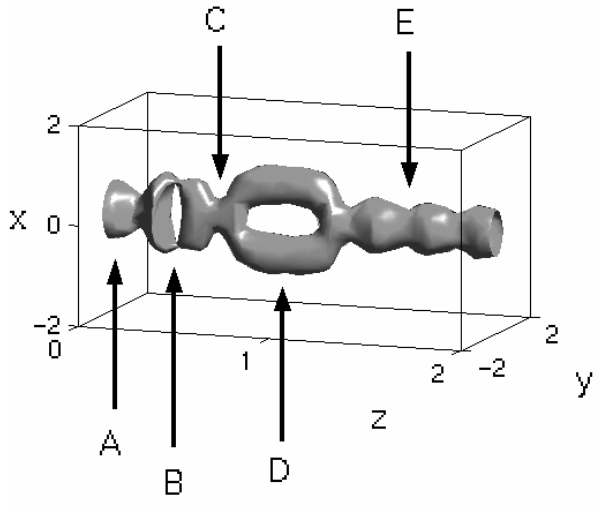

Figure 3.16: Iso-surface $|A_1|^2 \equiv 15$ of the data in Figure 3.15.

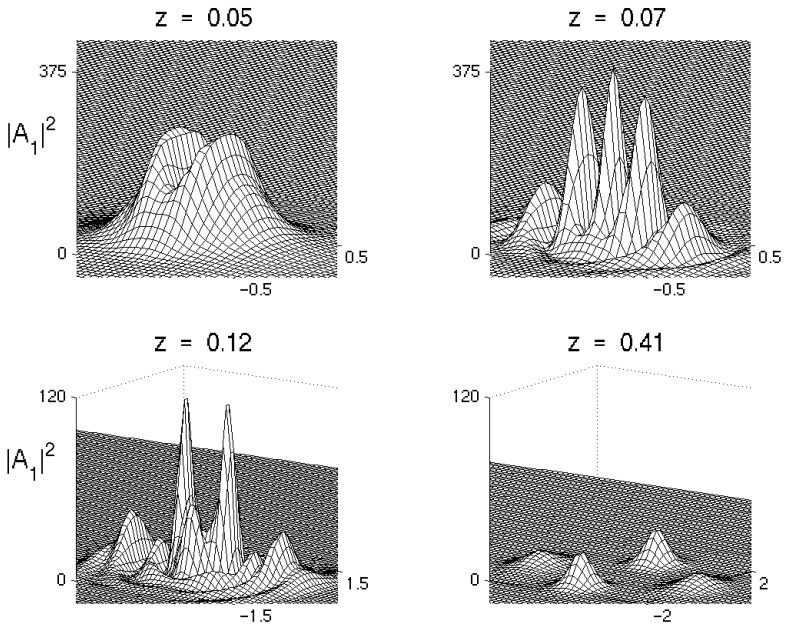

Figure 3.17: Same as Figure 3.12 with $f = 0.08$. Viewing angle in the (x, y) plane of the 3D graphs is $-55°$.

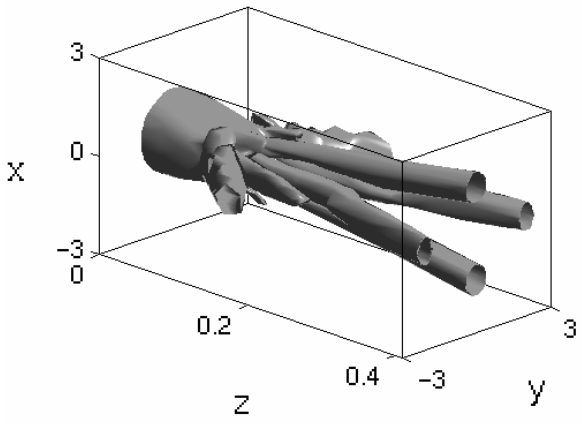

Figure 3.18: Iso-surface $|A_1|^2 \equiv 5$ of the data in Figure 3.17.

3.2.1 Summary of simulations

Although the picture of vectorial effects on self-focusing is only partial, the simulations in Section 3.2, as well as additional ones which we do not show, suggest the following. For given model parameters, there is a threshold power for genuine multiple-filamentation (that appears to be in the range $3N_c$–$5N_c$), such that,

1. When the input-power is sufficiently below this threshold, the beam propagates as a single filament, undergoing focusing-defocusing oscillations.

2. When the input-power is slightly below the threshold, an asymmetric ring is formed during the defocusing stage, followed by pseudo multiple-filamentation (*i.e.*, the filaments reunite into one lobe).

3. When the input-power is moderately above the threshold, an asymmetric ring is formed during the defocusing stage, followed by beam splitting into two disjoint filaments (genuine multiple-filamentation).

4. When the input-power is highly above the threshold, the beam can split into more than two filaments. In this case, all filaments split from the central beam (rather than a fractal process where the beam splits into two filaments and then each filament splits again).

In both (3) and (4) all filaments are of comparable power, which is roughly between $1.5N_c$ and $3.6N_c$. Thus, the power of each filament is below the threshold for additional splitting.

As we have already said, vectorial effects play an essential role in the multiple filamentation in Figs. 3.8–3.13, as they are the only mechanism that breaks-up the axial-symmetry. Because nonparaxiality preserves axial-symmetry, a natural question is whether it is 'needed' for multiple filamentation. To answer this question, we repeat the simulations of Figs. 3.8–3.13 but without the nonparaxial terms. In these simulations we observe some breakup of axial-symmetry but no multiple filamentation. Therefore, it is possible that nonparaxiality is also needed for multiple filamentation.

Remark. In the iso-intensity plots of multiple filamentation (Figs. 3.9, 3.11, and 3.13), the filaments are not parallel to each other. Rather, there is an angle of several degrees between the filaments. Such an angle has not been reported in the multiple filamentation experiments literature, nor in previous numerical studies, where noise in the input beam leads to multiple filamentation. This may suggest that multiple filamentation due to noise is more similar to experimental observations than multiple filamentation due to vectorial effects. This is not the case, however, because of the following reasons:

1. Most experiments of multiple filamentation measure the beam intensity at a fixed z. Therefore, it is possible that such an angle went unnoticed.

2. Because of the different rescaling in the axial and transverse coordinates [see Eq. (3.4)], the physical angle is approximately f times the angle in our iso-intensity plots. Thus, the physical angle is much smaller than what appears in the plots.

3. In our simulations of noise-driven multiple filaments in Section 3.4, these filaments are also not parallel (see Figure 3.26).

3.3 Single filament dynamics

The numerical simulations in Section 3.2 show that when the input beam is sufficiently powerful, vectorial effects can lead to multiple filamentation. In addition, these simulations show that, regardless of whether multiple filamentation occurs or not, beam collapse is arrested and a single filament or several filaments are formed, which propagate over long distances while maintaining roughly a constant width.

At present there is no theory for analyzing self-focusing in the high-power regime $N(0) \gg N_c$, which is why we rely on numerical simulations in the exploration of multiple filamentation in Section 3.2. However, when the power of a single filament is not much higher than N_c, its propagation can be analyzed using *modulation theory*, which is an asymptotic theory for analyzing the effects of small perturbations on critical self-focusing [32, 33]. Modulation theory is based on the observation that, after some propagation has taken place, a self-focusing filament rearranges itself as a modulated Townesian, *i.e.*,

$$|A_1(x,y,z)| \sim \frac{1}{L(z)} R(\frac{r}{L(z)}) \ ,$$

where $R(r)$ is defined in Eq. (2.20). Therefore, self-focusing dynamics is described by the modulation variable $L(z)$, which is proportional to beam-width and also to 1/(on-axis amplitude). In particular, $L \to 0$ and $L \to \infty$ correspond to blowup and to complete defocusing, respectively.

By applying modulation theory to Eq. (3.7) we prove in Appendix B that

Proposition 3.3. *When $f \ll 1$ the self-focusing dynamics of a single filament propagating in the presence of vectorial and nonparaxial effects [*i.e.*, Eq. (3.7)] is given, to leading order, by the reduced*

system of ODEs

$$\begin{cases} L_{zz}(z) &= -\dfrac{\beta}{L^3}\,, \\ \beta_z(z) &= -\dfrac{f^2(C_{\rm nonparax}+C_{\rm vec})N_c}{2M}\left(\dfrac{1}{L^2}\right)_z\,, \end{cases} \tag{3.16}$$

where

$$M=\frac{1}{4}\int_0^\infty \rho^2 R^2 \rho\, d\rho \;\approx\; 0.55\,, \tag{3.17}$$

$$C_{\rm nonparax}=1 \qquad \textit{and} \qquad C_{\rm vec}(\gamma)\approx\frac{16}{3}\left(1+\frac{\gamma}{1+\gamma}\right)\,. \tag{3.18}$$

Inspection of the derivation of eq. (3.16) shows that the terms with $C_{\rm nonparax}$ and with $C_{\rm vec}$ correspond to nonparaxial and vectorial effects, respectively. Therefore, the reduced system (3.16) shows that nonparaxiality and vectorial effects have the same qualitative effect on the self-focusing dynamics of a single filament. This observation is surprising, because at the PDE level [*i.e.*, Eq. (3.7)] the expressions for nonparaxiality and for vectorial effects are completely different [†].

As can be seen from Table 3.1, γ is zero or positive for most common physical mechanisms leading to the Kerr effect. Therefore, from (3.18) we have that $C_{\rm vec}(\gamma)\geq 16/3$. Thus, the reduced system (3.16) shows that vectorial effects dominate over nonparaxiality. This observation implies that models of physical self-focusing that include nonparaxiality should also include vectorial effects. We note that this has not been done in most previous studies.

Kerr mechanism	γ	$C_{\rm vec}(\gamma)$	$C_{E_3}(\gamma)$	$C_{{\rm grad-div}\ E_1}$
electrostriction	0	5.3	-2.7	8
nonresonant electrons	0.5	7.1	-0.9	8
molecular orientation	3	9.8	1.8	8

Table 3.1: Approximate values of the vectorial constants $C_{\rm vec}(\gamma)$ [in Eq. (3.18)], $C_{\mathcal{E}_3}(\gamma)$, and $C_{{\rm grad-div}\ \mathcal{E}_1}$ [in Eq. (3.21)] for common Kerr mechanisms.

We can follow [24, 32] and integrate Eqs. (3.16) to get

$$(y_z)^2=-\frac{4H_0}{My}(y_M-y)(y-y_m), \qquad y(z):=L^2(z)\,, \tag{3.19}$$

[†]It is interesting to note that the reduced system (3.16) also appears in the study of self-focusing dynamics with a saturated-nonlinearity effect such as Eq. (3.25) [24, 32, 33].

where

$$\begin{aligned} y_m &\approx \frac{M\beta(0)}{-2H_0}\left(1-\sqrt{1-4\delta}\right)\sim \frac{f^2N_c(C_{\rm nonparax}+C_{\rm vec})}{4M\beta(0)}\left[1+O(\delta)\right], \\ y_M &\approx \frac{M\beta(0)}{-2H_0}\left(1+\sqrt{1-4\delta}\right)\sim \frac{M\beta(0)}{-H_0}\left[1+O(\delta)\right], \end{aligned} \tag{3.20}$$

$\delta = -f^2N_c(C_{\rm nonparax}+C_{\rm vec})H_0/4M^2\beta^2(0)$ and $H_0 \approx H(0)$, where $H(0)$ is the input Hamiltonian (2.18).

We recall that a necessary condition for blowup in the unperturbed NLS, *i.e.*, Eq. (2.16), is that the input power is above critical, *i.e.*, $N(0) \geq N_c$ (see Section 2.4). In modulation theory variables [32], this condition amounts to $\beta(0) \approx [N(0) - N_c]/M \geq 0$. However, when $\beta(0) > 0$ we see from Eqs. (3.19) and (3.20) that $y(z) \geq y_m > 0$. Therefore blowup is arrested by vectorial and nonparaxial effects and the minimal beam width is $L_m \sim L(0)f\sqrt{N_c(C_{\rm nonparax}+C_{\rm vec})/4M\beta(0)}$, which, in physical variables, corresponds several wavelengths. Even at this stage the magnitude of the nonparaxial and vectorial terms in Eq. (3.7) is $O(\beta)$ smaller than that of the NLS terms $\Delta_\perp A_1$ and $|A_1|^2A_1$, providing an *a-posteriori* justification for treating them as small perturbations.

In addition, a sufficient condition for blowup in the unperturbed NLS (2.16) is $H(0) < 0$. However, from Eqs. (3.19) and (3.20) we see that if $\beta(0) > 0$ and $H(0) < 0$ then $y_m \leq y(z) \leq y_M$, *i.e.*, arrest of blowup is followed by focusing-defocusing oscillations. When nonadiabatic radiation is added to (3.16) the oscillations decay with propagation [24].

The qualitative picture predicted by (3.16), *i.e.*, arrest of blowup followed by focusing-defocusing oscillations, can be observed in the simulations of Figure 3.19, where the input power is only moderately above N_c, as well as in previous numerical studies of vectorial effects [37, 50]. We verify the prediction of the reduced system (3.16) that the effect of nonparaxiality is defocusing and small compared with that of vectorial effects, by comparing the dynamics with and without nonparaxial effects. Indeed, as can be in Figure 3.19, the effect of neglecting nonparaxiality is slightly focusing. We also verify that when nonparaxial effects are kept but vectorial effects are neglected in the simulations of Figure 3.19, then the qualitative dynamics remains the same, but the peak intensities are significantly higher (data not shown).

We recall that the vectorial terms in Eq. (3.7) correspond, in part, to the contribution of $\mathcal{E}_1$ from the grad-div term, which does not vanish even when $\mathcal{E}_2 \equiv \mathcal{E}_3 \equiv 0$ [see Eq. (3.8)]. A closer inspection of the derivation of (3.16) reveals that $C_{\rm vec}(\gamma) = C_{{\rm grad-div}\ E_1} + C_{E_3}(\gamma)$, where the constants

$$C_{{\rm grad-div}\ E_1} \approx 8 \quad \text{and} \quad C_{E_3}(\gamma) \approx -\frac{8(1-\gamma)}{3(1+\gamma)} \tag{3.21}$$

correspond to the contribution of $\mathcal{E}_1$ from the grad-div term and the coupling to $\mathcal{E}_3$, respectively [see Eq. (3.10)]. This result is surprising, because it shows that the defocusing effect of the grad-div-E_1 term is *eight times stronger* than nonparaxiality. Also surprising is that the effect of the coupling to $\mathcal{E}_3$ is weakly *focusing* for $|\gamma| < 1$, since $C_{E_3} < 0$ (see Table 3.1).

We now verify these predictions using numerical simulations. In Figure 3.20A we show that NLS blowup is slightly accelerated by the coupling effects to $\mathcal{E}_3$ when $\gamma = 1/2$. In addition, as $C_{E_3}(0.5) \approx -0.9$, we get that $C_{\text{nonparax}} + C_{E_3}(0.5) = 0.1$, which shows that the combined effect of nonparaxiality and coupling to $\mathcal{E}_3$ for $\gamma = 1/2$ is *slightly* defocusing. In order to verify this prediction we compare the solution of Eq. (3.7) with $\gamma = 1/2$ (which takes nonparaxiality, the coupling to $\mathcal{E}_3$, and the contribution of the grad-div-E_1 term into account) with the solution of Eq. (3.8) without the nonparaxial term, *i.e.*,

$$iA_{1,z} + \Delta_\perp A_1 + |A_1|^2 A_1 = -f^2(|A_1|^2 A_1)_{xx} \,, \tag{3.22}$$

which takes only the contribution of the grad-div-E_1 term into account. Figure 3.20B shows that, as predicted, the peak focusing is slightly lower with the nonparaxial and the coupling to $\mathcal{E}_3$ terms than without them. As an additional test, we verify that the solution of the NLS with the nonparaxial and coupling to $\mathcal{E}_3$ terms (without the grad-div-E_1 term) with the same parameters and input beam as in Figure 3.20B does not blowup (data not shown).

Finally, when the input power is above the threshold for multiple filamentation, a more careful inspection of the data in Figs. 3.8, 3.10, and 3.12 reveals that after the splitting has taken place each filament undergoes almost-periodic focusing-defocusing oscillations. For example, in Figure 3.21 multiple filamentation occurs after the first two oscillations, and the subsequent oscillations are the focusing-defocusing cycles of each filament.

Remark. As in the case of the angle between the filaments (see the remark at the end of Section 3.2), in physical variables the oscillations are much slower than what may appear from Figs. 3.3, 3.5, and 3.16–3.21. These focusing-defocusing oscillations can be interpreted as self-trapping, *i.e.*, the formation of a long and narrow filament. Indeed, such oscillations were observed in the cw self-trapping experiments of Bjorkholm and Ashkin [6].

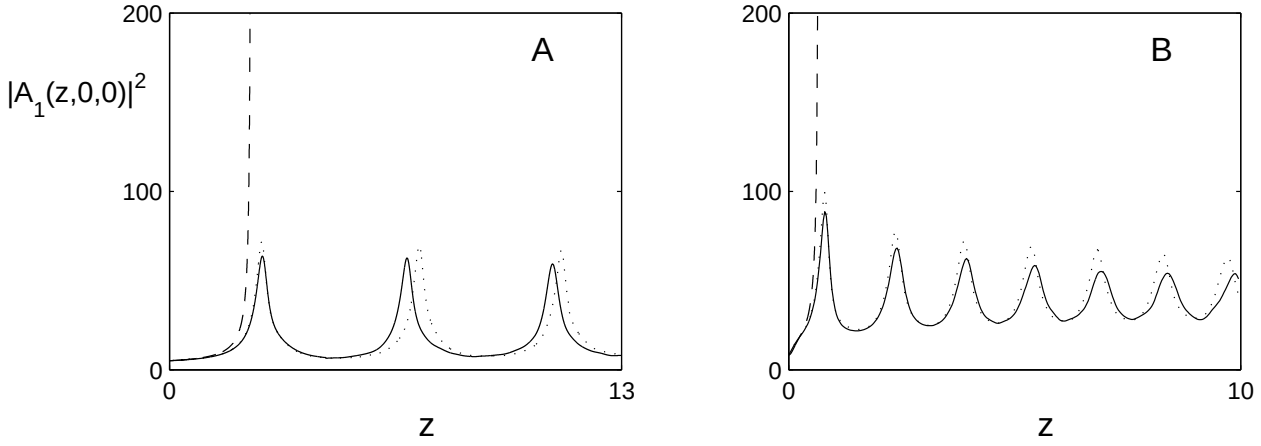

Figure 3.19: Blow-up in NLS [*i.e.*, Eq. (2.16), dashed line] is arrested by vectorial effects, with nonparaxiality [*i.e.*, Eq. (6.2), solid line] and without nonparaxiality [*i.e.*, setting $A_{1,zz} \equiv 0$ in Eq. (3.7), dots], resulting instead in focusing-defocusing oscillations. Here $f = 0.04$, $\gamma = 0.5$, and (A) $A_1(z=0,r) = \sqrt{1.05}R(r)$ [*i.e.*, $N(0) = 1.05N_c$]; (B) $A_1(z=0,r) = 2\sqrt{1.1N_c}e^{-r^2}$ [*i.e.*, $N(0) = 1.1N_c$].

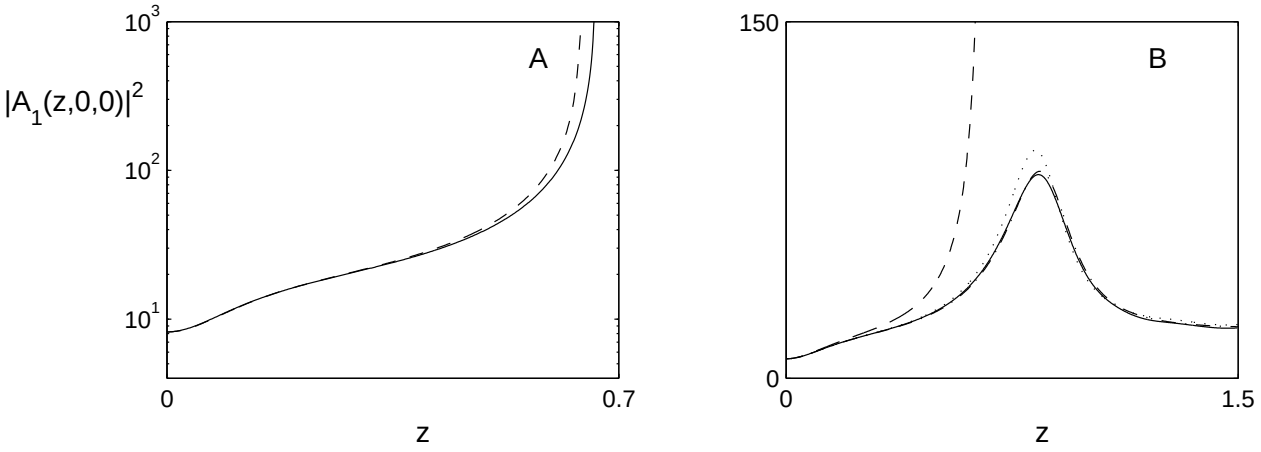

Figure 3.20: (A) NLS blowup (solid line) is slightly accelerated by the coupling effects to $\mathcal{E}_3$ [*i.e.*, Eq. (3.9), dashes], same data as in Figure 3.19B. (B) Same data (and lines) as in Figure 3.19B with, additionally, the solution of Eq. (3.22) (dash-dots '–.'), which reaches a slightly higher peak, and is almost indistinguishable from the solution of Eq. (3.7) (solid line).

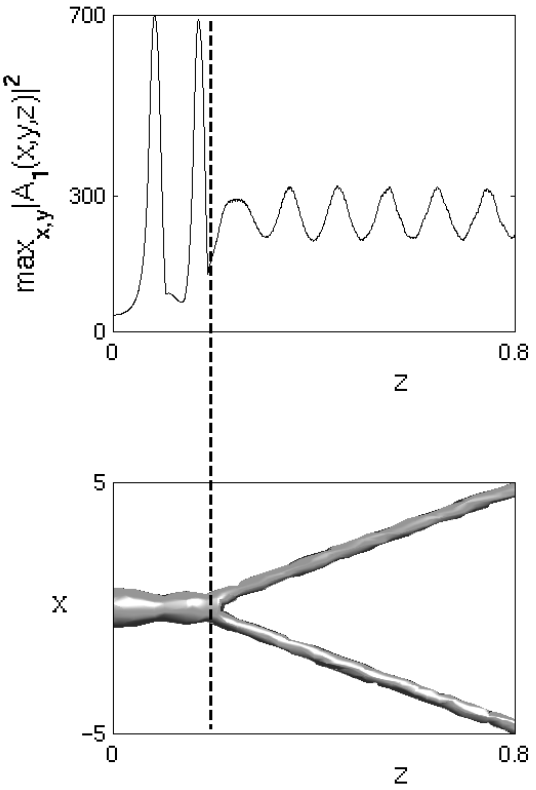

Figure 3.21: Same as Figure 3.2 with $N(0) = 5N_c$. Top: peak intensity. Bottom: Iso-surface $|A_1|^2 = 10$. The dashed line is the point where the beam breaks-up into two filaments.

3.4 Random multiple filamentation

3.4.1 The Bespalov-Talanov model

In 1966 Bespalov and Talanov suggested that noise in the input beam is the symmetry-breaking mechanism that leads to multiple filamentation [5]. Their analysis was based on the assumption that, to leading order, the electric field is a plane-wave, *i.e.*,

$$\mathcal{E}_1(x,y,z) \;\sim\; \alpha\, e^{-i\alpha^2 z}\;, \qquad\qquad \alpha \;\equiv\; const\;. \tag{3.23}$$

Using this assumption, they showed that certain frequencies are linearly unstable. From this they concluded that:

> *"... in a nonlinear dielectric [Kerr medium] amplitude-phase perturbations of a plane electromagnetic wave bring about its decay into individual beams".*

In other words, the noise-driven instabilities would breakup the cylindrical symmetry of the beam, ultimately leading to multiple filamentation.

At present, the Bespalov-Talanov explanation is the standard, and only, explanation for multiple filamentation of laser beams. Surprisingly, until now this explanation has not been corroborated using numerical simulations. Testing their explanation numerically requires a (2+1)D simulation because of the input noise. Although even (1+1)D simulations (*i.e.*, using a cylindrically-symmetric code) were difficult in the 1960's, (2+1)D simulations became possible in the 1980's. One can only speculate as to why this explanation has not been tested until now. One possible reason for this is that by the 1980's their explanation had already been accepted as the true explanation for multiple filamentation without doubt.

We now numerically test the Bespalov-Talanov model for multiple filamentation. To do that, we solve the unperturbed NLS (2.16) with a high-power [*i.e.*, $N(0) \gg N_c$] cylindrically-symmetric Gaussian input-beam, to which we add random noise both in amplitude and in phase, *i.e.*,

$$A_1(x,y,z=0) = 2\sqrt{N(0)}\, e^{-(x^2+y^2)} \left[\; 1 + c * noise(x,y) \;\right]\;, \tag{3.24}$$

where $N(0)$ is the noiseless input power, $noise(x,y)$ is a random complex-valued function (see Section 6.1 for more details), and the constant c determines the noise level ($c \ll 1$). Although we have made many simulations with high-power input beams and random noise, we see neither evidence for multiple filamentation nor even mild instabilities. Rather, the beam converges to a cylindrically-symmetric profile as it blows-up (see Figure 3.22 and Figure 3.23).

Apparently, the major weakness of the Bespalov-Talanov argument is that it assumes that, to leading order, the electric field is a plane-wave (3.23). Under this assumption (which implies infinite input power) the self-focusing process of the field does not depend on the transverse dynamics, *i.e.*, $\mathcal{E}_{1,z} + |\mathcal{E}_1|^2\mathcal{E}_1 = 0$. As a result, instabilities can grow while the leading order solution remains unchanged. This is not the case, however, for a propagating beam, where the transverse dynamics of the beam dominates the evolution of the noise.

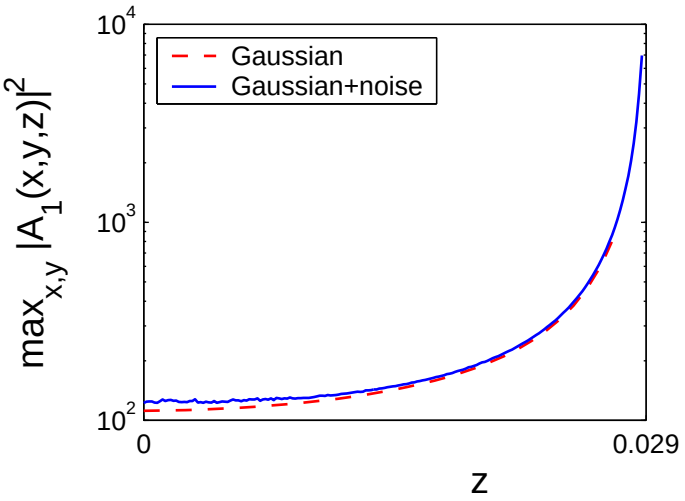

Figure 3.22: Blow-up in the NLS (2.16) of high-power input beams (3.24) with noise (solid line, $c = 0.1$) and in the absence of noise (dot-dashed line, $c = 0$) is very similar. Here $N(0) = 15N_c$.

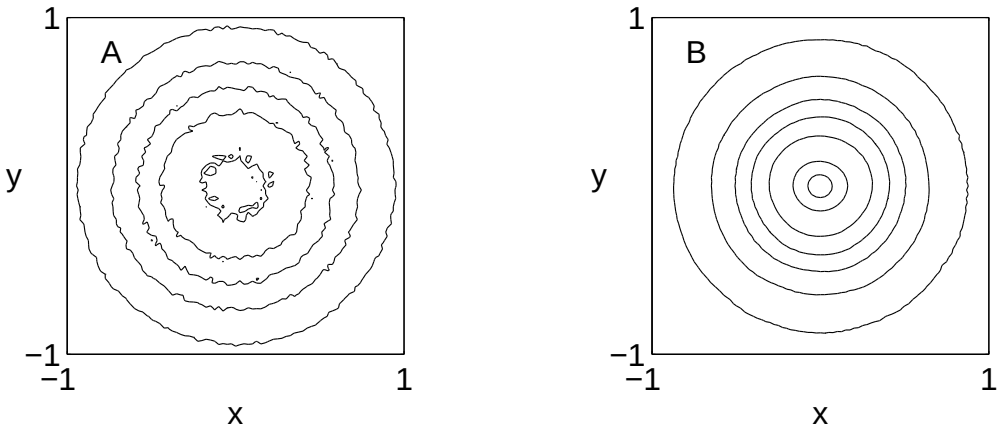

Figure 3.23: Contour plots of $|A_1(x, y, z)|^2$ of the solution of Figure 3.22 at (A) $z = 0$ and (B) $z = 0.026$.

3.4.2 Noise and a saturating nonlinearity

In 1968, Marburger and Dawes [48] showed numerically that intense Gaussian beams propagating in a Kerr medium with saturable nonlinearity do not collapse, but instead go through focusing-defocusing cycles. In addition, they showed that the transverse profile of the beam can develop a concentric ring structure. Although they pointed out the relation between a ring structure and the formation of small-scale filaments, they could not demonstrate this numerically, because their code was cylindrically symmetric. This relation was established in 1979, when Konno and Suzuki solved the saturated NLS using a Cartesian grid and showed that the ring structure is indeed unstable [40]. Later, using both numerical simulations and an approximate stability analysis, Soto-Crespo, Wright and Akhmediev [58] showed that the transition from cylindrical symmetry to multiple filamentation is associated with the appearance of a spatial ring. Multiple filamentation due to noise in the input beam and nonlinear saturation was also observed in [2, 35, 47, 51, 60].

In order to compare multiple filamentation due to vectorial effects with the one to noise, we solve the saturated NLS

$$iA_{1,z} + \Delta_\perp A_1 + \frac{|A_1|^2 A_1}{1+\varepsilon|A_1|^2} = 0\ , \quad 0 < \varepsilon \ll 1 \tag{3.25}$$

with high-power noisy input beams (3.24). For example, when $\varepsilon = 0.01/\ln 2$ and $c = 0.02$, we observe pseudo multiple-filamentation when $N_0 \approx 10N_c$ (Figure 3.24) and genuine multiple filamentation when $N_0 \approx 14N_c$ (Figs. 3.25 and 3.26). Therefore, for these values of ε and c, the threshold for multiple filamentation lies between $10N_c$ and $14N_c$.

We carry additional simulations of Eq. (3.25) with the same values of ε and c in the range $14N_c \leq N(0) \leq 120N_c$. These simulations suggest that noise-induced multiple filamentation is characterized by a powerful central on-axis filament, from which less-powerful off-axis filaments split. After multiple filamentation occurs, above half the input power remains in the on-axis filament. The off-axis filaments are less powerful, as the power of each is below one-tenth of the input power. In addition, a significant amount of power is radiated to the background, as was also observed in [60]. We note that even when the power of the on-axis filament is much larger than $14N_c$ (*i.e.*, above the initial threshold for multiple filamentation), no additional splitting occurs. This observation shows that, as can be expected, the effect of input noise diminishes with propagation.

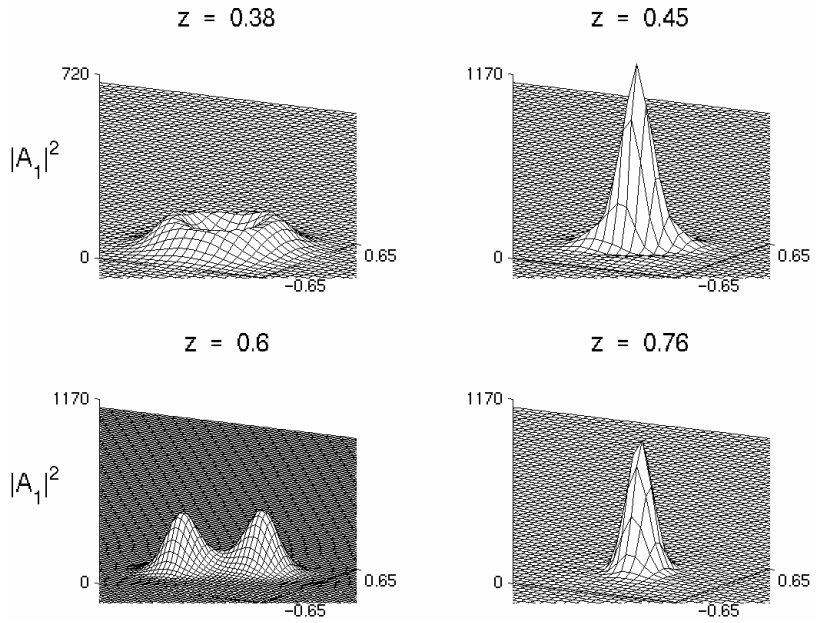

Figure 3.24: Solution of Eqs. 3.25 with $\varepsilon = 0.01/\ln 2$ with the input beam (3.24), where $N(0) \approx 9.75N_c$ and $c = 0.02$.

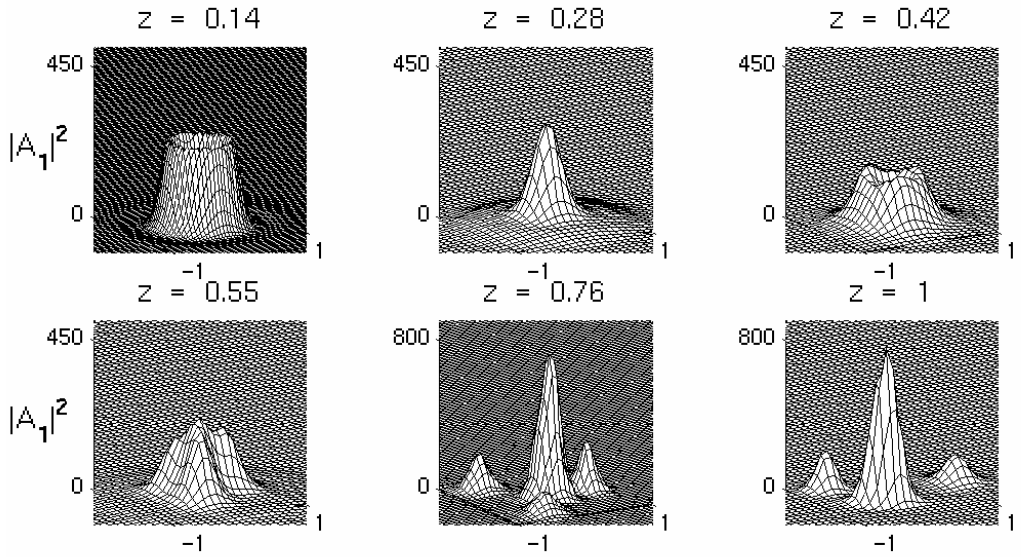

Figure 3.25: Same as Figure 3.24 with $N(0) \approx 14N_c$.

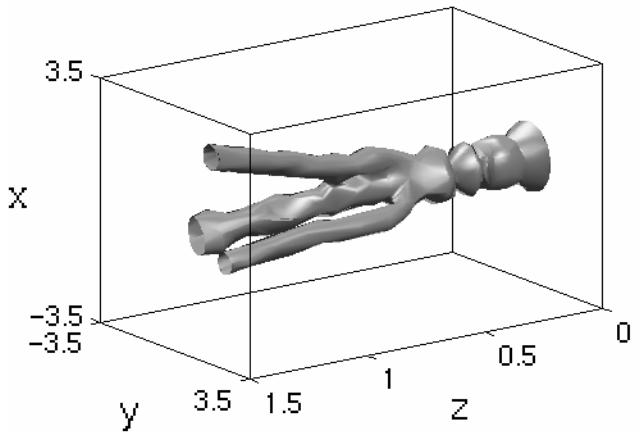

Figure 3.26: Iso-intensity surface $|A_1|^2 = 30$ of the data in Figure 3.25. Note that the beam propagates from right to left.

3.4.3 Astigmatism and saturation

In the previous section we showed that input noise can lead to multiple filamentation in a saturated-Kerr medium. In this section we show that input astigmatism can also lead to (pseudo and genuine) multiple filamentation in a saturated-Kerr medium.

To study whether astigmatism can also lead to multiple filamentation we consider the input beam

$$A_1^0(x,y) = 2\sqrt{CN_c}\,\exp\left[\,-(ex)^2 - y^2\,\right], \tag{3.26}$$

where C is constant and $0 \leq e \leq 1$ is input astigmatism parameter ($e = 1$ corresponds to a cylindrically-symmetric input beam). Our simulations of Eq. (3.25) show that input astigmatism can lead to multiple filamentation when the input power is several times N_c[‡]. For example, Figure 3.27 shows an astigmatic input beam that break-up into two filaments. As with noise-initiated multiple filamentation, during further propagation each of the filaments undergoes focusing-defocusing oscillations and is roughly cylindrically symmetric.

[‡]We note that the effective threshold power for collapse of astigmatic beams is higher than for cylindrically-symmetric beams by a factor of approximately $[0.2(e + 1/e) + 0.6]$ (see [27]).

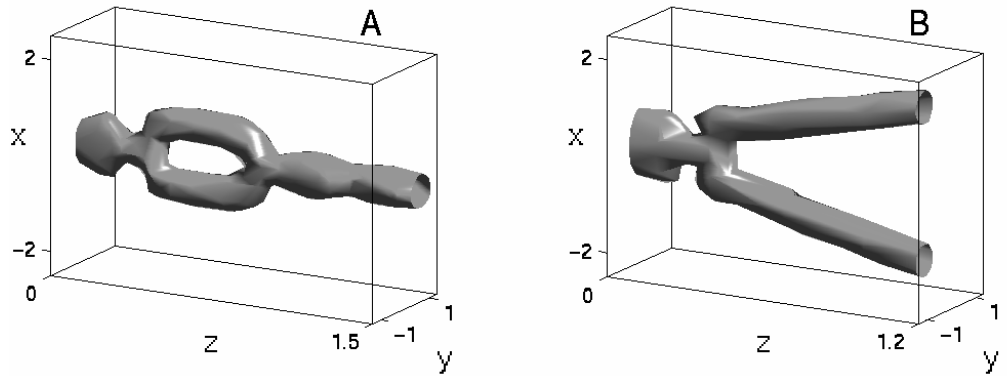

Figure 3.27: Isosurface of $|A_1|^2$ of the solution of Eq. (3.25) with the astigmatic input beam (3.26), $\varepsilon = 0.01$, $e = 0.6$, and (A) $C = 2.5$ [*i.e.*, $N(0) \approx 4.17N_c$]; (B) $C = 3$ [*i.e.*, $N(0) = 5N_c$].

3.4.4 Noise and nonparaxiality?

The saturation of the Kerr nonlinearity has been observed experimentally for only some materials. For example, Yau, Lee and Wang reported recently that the nonlinearity for Sapphire crystal is of the form (3.25). However, there are other materials for which the assumption of nonlinear saturation does not have a solid physical justification. Therefore, a natural question is whether noise can lead to multiple filamentation in the absence of saturation. To answer this question we study the possibility of multiple filamentation for a beam propagating in the presence of nonparaxiality (but without vectorial effects), *i.e.*,

$$iA_{1,z} + \Delta_\perp A_1 + |A_1|^2 A_1 = -\frac{1}{4} f^2 A_{1,zz} \, . \tag{3.27}$$

We remark that vectorial effects are always present in physical self-focusing. Our purpose here is to analyze the ability of the noise mechanism to cause multiple filamentation. In (3.27) we neglect vectorial effects in order to assure that if multiple filamentation occurs, it occurs because of the input noise and not because of vectorial effects.

When we solve Eq. (3.27) we do not see multiple filamentation. For example, we solve (3.27) with an input beam with $N(0) = 20N_c$ and 10% noise. As can be seen in Figure 3.28, the beam forms a ring at an early stage of the propagation, which is followed by convergence into a single filament that undergoes focusing-defocusing oscillations during further propagation. However, in spite our simulations, we believe that multiple filamentation may be possible with nonparaxiality and noise, for the following reason. All of the multiple filamentation simulations show that multiple filamentation is preceded by ring formation, which strongly suggest that the ring structure is unstable to small perturbations. As Figure 3.28 does show ring formation, it is possible that noise-driven multiple filamentation in the presence of nonparaxiality does occur with higher input power.

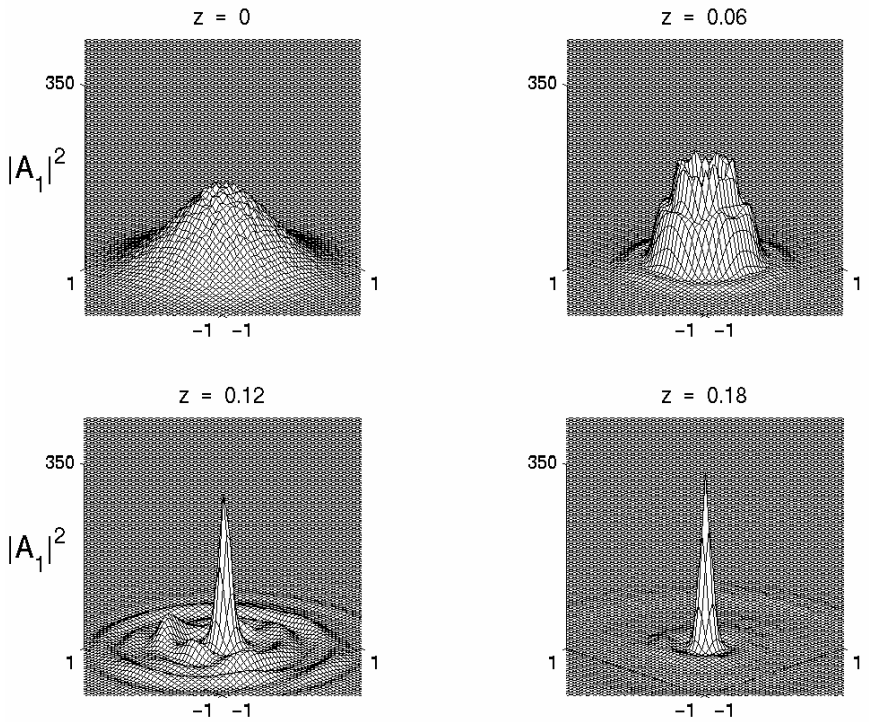

Figure 3.28: Intensity of solution of Eq. (3.27) with $f = 0.1$ and a noisy high-power input beam as in Figure 3.12.

Remark. As we have seen, in the case of linear polarization vectorial effects result in anisotropic terms in the model equation (3.7). In contrast, in Section 4.6 we show that for the case of *circular polarization* vectorial effects result in *isotropic* terms in the model equation (4.27). Thus, those vectorial effects cannot lead to multiple filamentation by themselves. However, in Section 4.6 we show that noise with nonparaxial and vectorial (circular-polarization) effects can lead to multiple filamentation, even in the absence of a saturation mechanism.

3.5 Vectorial effects and/or noise?

As we have seen, regardless of whether the symmetry-breaking mechanism is vectorial effects or noise in the input beam, multiple filamentation always occurs after the formation of a ring during the defocusing stage. Because a ring structure is unstable, it can be broken into multiple filaments by symmetry-breaking mechanisms.

There are, however, significant differences between multiple filamentation induced by vectorial effects and by noise. The most important difference is that the former is a deterministic process whereas the later is a random one. Therefore, when the input beam is cylindrically symmetric ('clean' input

beam) the filamentation pattern is reproducible in the case of vectorial-induced filaments but varies from experiment to experiment in the case of noise-induced filaments. Another difference, which can be utilized to determine experimentally the mechanism of multiple filamentation (see Section 3.6) is that vectorial effects induce a preferred direction in the transverse (x, y) plane, which is the direction of input beam polarization.

Our simulations suggest that the threshold power for multiple filamentation caused by vectorial effects is significantly lower than the one for multiple filamentation caused by noise. For example, simulations of Eq. (3.25) with $\varepsilon = 0.01$ and noise level of $c = 0.02$ do not lead to multiple filamentation when $N(0) < 10N_c$, but do lead to multiple filamentation when $N(0) > 14N_c$. In contrast, in the vectorial case with $f = 0.05$, we already observe multiple filamentation when $N(0) = 4N_c$. Thus, multiple filamentation at these lower powers is more likely to be the result of vectorial effects.

Another difference is that in vectorial-induced multiple filamentation, regardless of whether there is a central filament after the splitting (Figure 3.12) or not (Figure 3.8), all filaments are of comparable power, which is below $3.6N_c$. In contrast, noise-induced multiple filamentation is characterized by a powerful central filament which has about half the input power, and significantly less-powerful off-axis filaments.

In addition, because both vectorial effects and noise take part in multiple filamentation experiments and each of these two mechanisms can, in principle, lead to multiple filamentation on its own, it is natural to ask whether they can cooperate to cause multiple filamentation. To answer this question we solve Eq. (3.7) with a noisy input beam whose power is slightly below the threshold for multiple filamentation. We recall that in the case where the input beam is symmetric (*i.e.*, without noise), $f = 0.05$, and $N(0) = 3.75N_c$ then ring formation is followed by *pseudo* multiple filamentation (see Figure 3.7). In contrast, when noise is added to the input beam the beam undergoes genuine multiple filamentation (see Figure 3.29). Therefore, in this simulation, noise cooperates with the vectorial effects to lower the threshold power for (genuine) multiple filamentation.

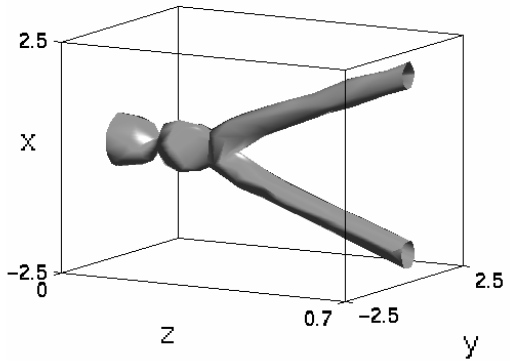

Figure 3.29: Iso-surface of $|A_1|^2 = 15$ of the solution of Eq. (3.7) with $f = 0.05$ and the noisy input beam (3.24) with $c = 0.1$ and $N(0) \approx 3.78 N_c$.

3.6 An experimental test

The results presented so far show that multiple filamentation can result from either vectorial effects or from noise in the input beam. Therefore, in theory, when all parameters of a multiple-filamentation experiment are known, one can use numerical simulations to determine whether the mechanism behind multiple filamentation is vectorial effects or noise. However, at the high input powers associated with multiple filamentation, other physical mechanisms (*e.g.*, plasma generation, time-dispersion, etc.), which are neglected in both models, can also become important. Thus, our model might not capture all the relevant physics, in which case, its reliability in determining the mechanism behind multiple filamentation is less clear.

In order to overcome this difficulty and be able to determine whether vectorial or random effects are the physical mechanism behind multiple filamentation, we propose the following experimental test. This test is based on the observation that *vectorial effects are the only mechanism neglected in the derivation of the NLS model that breaks-up the cylindrical symmetry while inducing a preferred direction* in an isotropic homogeneous medium (the direction of input beam polarization). Therefore, if multiple filamentation is caused by vectorial effects, then

1. The filamentation pattern in the transverse plane should persist between experiments.

2. When the direction of linear polarization of the input beam is rotated in the transverse plane between experiments, the filamentation pattern should follow the same rotation.

3. When a beam splits into two filaments, the splitting should occur either along the direction of initial polarization or perpendicular to it (see Figs. 3.9 and 3.11).

In contrast, when multiple filamentation results from random instabilities, the filamentation pattern

should vary between experiments and be independent of the direction of initial polarization.

We note that in the multiple-filamentation experiments in [52], Nowak and Ham observed that

> *"the ... [filament] patterns..., although random in appearance, were perfectly reproducible shot to shot"*.

A similar observation was reported by Brodeur, Ilkov and Chin [14]. Because of this observation, Nowak and Ham conjectured that multiple filamentation was due to small in-homogeneities in the medium. However, our study shows that this behavior is also consistent with the vectorial effects explanation for multiple filamentation.

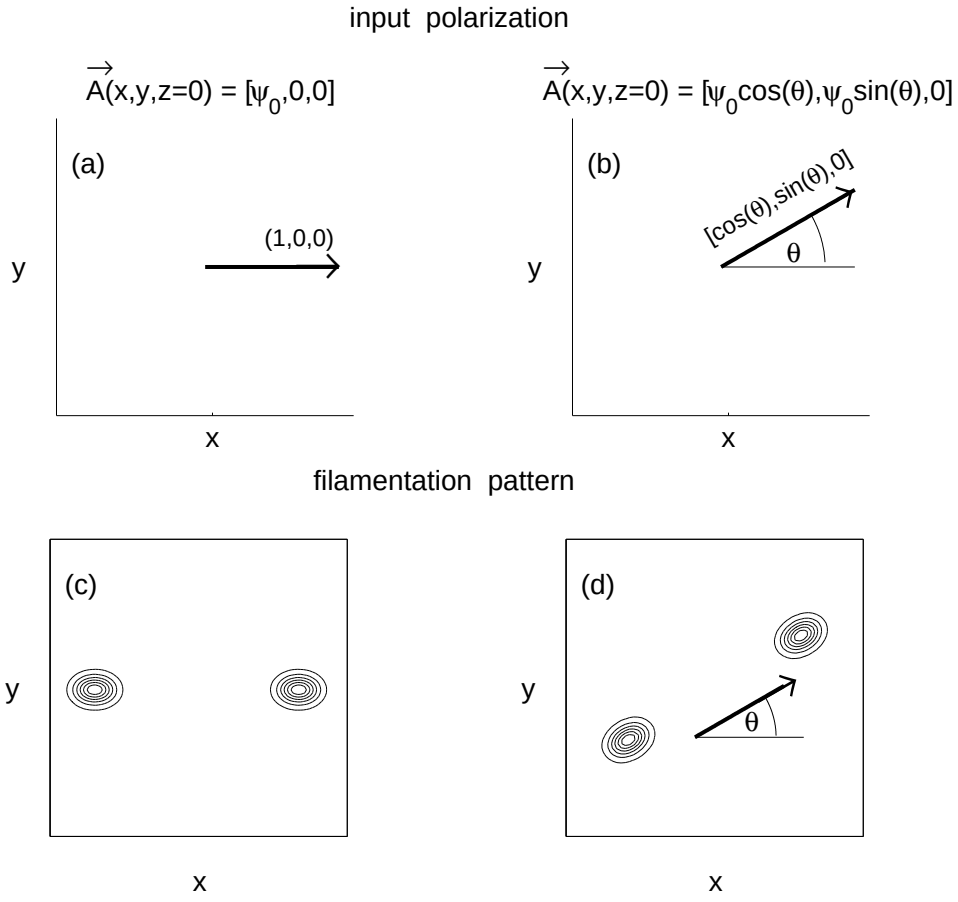

Figure 3.30: An experimental test for identifying deterministic multiple filamentation caused by vectorial effects. Let a linearly polarized input beam (a) whose direction of linear polarization is $(1, 0, 0)$ propagate in a Kerr medium and let (c) be filamentation pattern at some propagation distance. Rotate the input polarization to $(\cos\theta, \sin\theta, 0)$ (b) and repeat the experiment. If multiple filamentation is caused by vectorial effects, the filamentation pattern should follow the same rotation (d).

Chapter 4

Circularly polarized input beams

The NLS model is derived under the assumption that the beam is *linearly polarized*. In 1966 Close, Giuliano, Hellwarth, Hess, McClung, and Wagner [21] conducted experiments with intense *circularly polarized* input beams propagating in Kerr media, which suggested that circular polarization is unstable. Close *et al.*also proposed a mathematical model for self-focusing of circularly polarized beams, which they used to explain the observed instability of circular polarization. Their model was a system of two coupled NLS equations for the two circular polarization components. Subsequent theoretical studies have used the same system of equations as Close *et al.*, but obtained contradictory results with regard to circular-polarization (in)stability. As a result, to these days there is some confusion in the literature with regard to circular polarization stability. Remarkably, the only thing that was always agreed upon was the Close et al. model itself. As we show in this study, however, this model is based on problematic assumptions, and it can lead to wrong results. Hence, previous studies used the wrong model for studying circular polarization stability.

This chapter is organized as follows (see also Figure 4.1 and see Table 2 in Section 0.1 for notations). In Section 4.1 we describe the contradictory results of previous studies on circular polarization stability, all of which were based on the Close *et al.* system (4.3). In Section 4.2 we systematically reduce the vector Helmholtz equation to the new system (4.10) that models self-focusing of circularly polarized beams. Similarly to the Close *et al.* system (4.3), the new system (4.10) takes into account the coupling to the opposite-circular component (*i.e.*, the one rotating in the opposite direction). Unlike (4.3), however, the new System (4.10) also takes into account beam nonparaxiality, the coupling to the axial component, and the contribution from the grad-div term. Using (4.10) we show that circular polarization is stable. We also show that the assumptions on which the derivation of the Close *et al.* System (4.3) is based can be physically incorrect. In Section 4.3 we prove that when nonparaxiality, the coupling to the axial component, and the grad-div terms are negligible, Systems (4.3), (4.10), and

the simpler system (4.20) are asymptotically equivalent, and that solutions of these systems can undergo collapse. In Section 3.3 we use *modulation theory* to describe the dynamics of a single filament with the reduced system of ODEs (3.16), which shows that nonparaxiality, the coupling to the axial component, and the contribution from the grad-div term arrest beam collapse and lead to focusing-defocusing oscillations. In Section 4.5 we use numerical simulations of System (4.10) to confirm the stability of circular polarization and the predictions of modulation theory. These simulations also demonstrate that System (4.3) can lead to wrong predictions.

As we have shown in Section 3.2, the preferred direction in the transverse plane that is induced by *linearly polarized* input beams can lead to *deterministic* multiple filamentation, even when the input beams are perfectly cylindrically symmetric. In this chapter we show that, in contrast, cylindrically symmetric, *circularly polarized* input beams cannot undergo multiple filamentation, because in that case the input polarization state does not induce a preferred direction in the transverse plane. In Section 4.6 we show that when the input profile is cylindrically symmetric then a small deviation from circular polarization is unlikely to lead to multiple filamentation. However, small imperfections in the input profile, such as input noise or astigmatism, can lead to multiple filamentation.

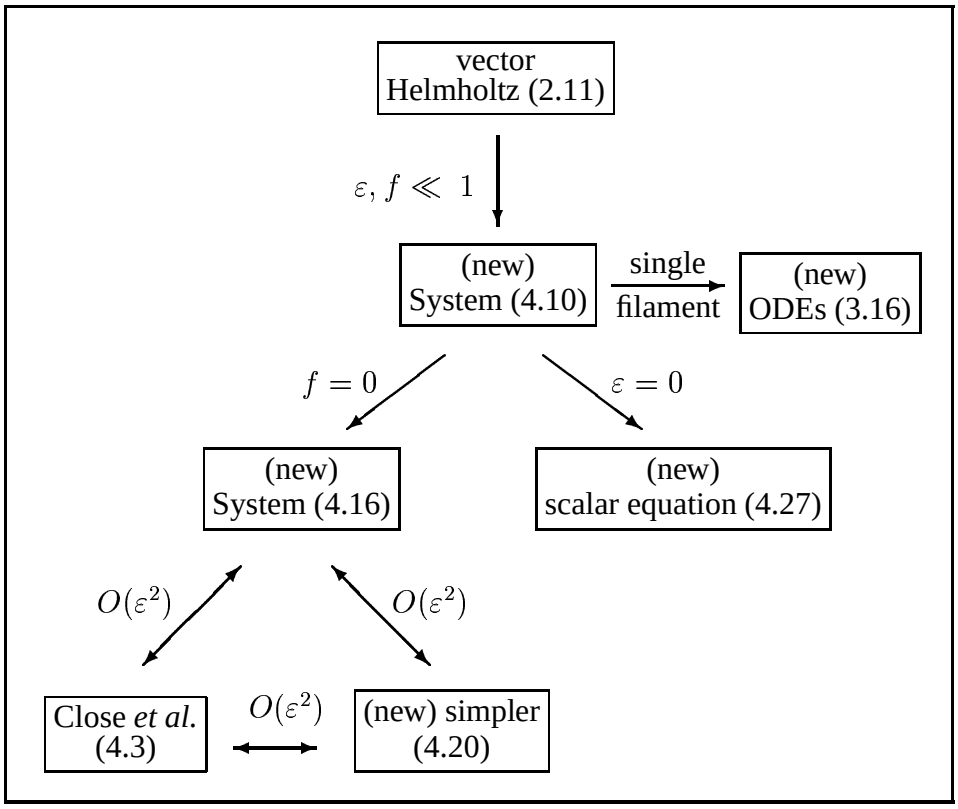

Figure 4.1: Schematic outline of the derivation. $f = 1/r_0 k_0$ and ε measures the deviation from perfect circular polarization [see Eq. (4.6)].

4.1 Polarization (in)stability - historical background

Strictly speaking, an intense circularly polarized input beam that propagates in a Kerr medium does not remain circularly polarized. Indeed, the vector Kerr effect (2.12) can be rewritten in terms of the circular components as

$$\vec{\mathcal{P}}_{\rm NL} = \frac{4\epsilon_0 n_0 \bar{n}_2}{1+\gamma} \left[(|\mathcal{E}_+|^2 + |\mathcal{E}_-|^2 + |\mathcal{E}_3|^2)\vec{\mathcal{E}} + \gamma(2\mathcal{E}_+\mathcal{E}_- + \mathcal{E}_3^2)\vec{\mathcal{E}}^* \right],$$

where the left-circular $(+)$ and right-circular $(-)$ components are denoted by

$$\mathcal{E}_\pm(x,y,z) = \frac{1}{\sqrt{2}} \left(\mathcal{E}_1 + i\mathcal{E}_2 \right), \tag{4.1}$$

with the corresponding input components $\mathcal{E}_\pm^0(x,y) = \mathcal{E}_\pm(x,y,z=0)$. Thus, $\mathcal{E}_+$, $\mathcal{E}_-$, and $\mathcal{E}_3$ are nonlinearly coupled through the vector Kerr effect (2.12) and linearly coupled through the grad-div term in Eq. (2.11a). Therefore, even if initially $\mathcal{E}_-^0 \equiv 0$ then $\mathcal{E}_-$ does not remain zero for $z > 0$. In addition, in practice the input beam is never perfectly circularly polarized. Therefore, a more realistic representation of circularly polarized input beams is $\mathcal{E}_+^0 \gg \mathcal{E}_-^0$.

All the previous studies of circular-polarization stability included the coupling of $\mathcal{E}_+$ to $\mathcal{E}_-$, but neglected the coupling of $\mathcal{E}_+$ to the axial component $\mathcal{E}_3$ and the grad-div term in (2.11a), *i.e.*, they assumed that

$$\mathcal{E}_3 \ll \mathcal{E}_+, \qquad \mathcal{E}_3 \ll \mathcal{E}_-, \qquad \nabla(\nabla\cdot\vec{\mathcal{E}}) \ll k_0^2\mathcal{E}_+. \tag{4.2}$$

Under the assumptions (4.2), Eqs. (2.11)–(2.12) reduce to the two coupled equations

$$\Delta\mathcal{E}_\pm + k_0^2\mathcal{E}_\pm + \frac{4k_0^2 n_2}{n_0(1+\gamma)} \left[|\mathcal{E}_\pm|^2 + (1+2\gamma)|\mathcal{E}_\mp|^2 \right] \mathcal{E}_\pm = 0.$$

These studies also used the slowly-varying envelopes $\mathcal{E}_\pm = \mathcal{A}_\pm(x,y,z)\, e^{ik_0 z}$ and the paraxial approximation $\mathcal{E}_{\pm,zz} \ll k_0\mathcal{E}_{\pm,z}$, which lead to the two coupled equations for $\mathcal{A}_+$ and for $\mathcal{A}_-$:

$$2ik_0\mathcal{A}_{\pm,z} + \Delta_\perp\mathcal{A}_\pm + \frac{4k_0^2 n_2}{n_0(1+\gamma)} \left[|\mathcal{A}_\pm|^2 + (1+2\gamma)|\mathcal{A}_\mp|^2 \right] \mathcal{A}_\pm = 0. \tag{4.3}$$

System (4.3) has been used to determine whether circular polarization is stable. As we now show, the results have been controversial.

4.1.1 The analysis of Close *et al.*

In 1966 Close, Giuliano, Hellwarth, Hess, McClung, and Wagner [21] conducted experiments with intense circularly polarized input beams propagating in Kerr media. They observed that *"in every*

case studied, the trapped light from a beam, circularly polarized to better than 1 part in 200, was markedly, if not completely, depolarized as soon as self-trapping could be detected". Moreover, *"the filament pattern [...] suggested that each filament might consist mainly of light linearly polarized in some random direction".* In other words, they observed that circular polarization is unstable and that during self-focusing circularly polarized beams formed filament(s) that are linearly polarized in randomly oriented directions (see Figure 4.2).

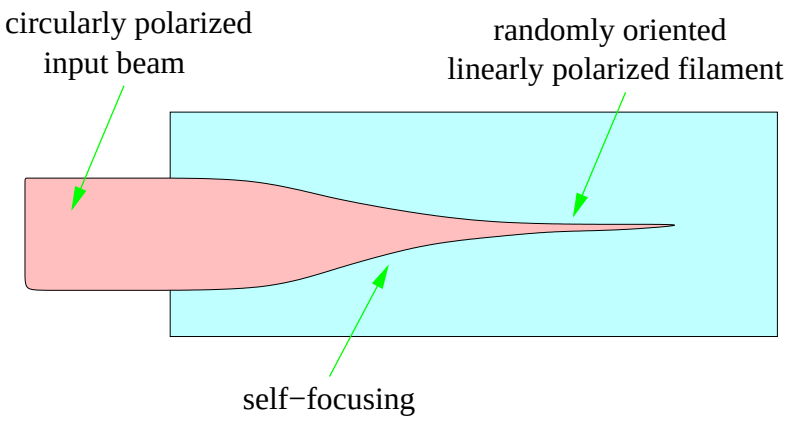

Figure 4.2: The experimental observation of Close *et al.* [21].

Close *et al.* suggested the following theoretical explanation for instability of circular polarization observed in their experiments. From (4.3) it follows that the effective change of the nonlinear refractive index of $\mathcal{A}_\pm$ is given by

$$\delta n_\pm = \frac{4k_0^2 n_2}{n_0(1+\gamma)} \left[|\mathcal{A}_\pm|^2 + (1+2\gamma)|\mathcal{A}_\mp|^2 \right] . \tag{4.4}$$

When $\gamma > 0$ the coefficient of the second term in the brackets of Eq. (4.4) is larger than the coefficient of the first term*. Therefore, Close *et al.* concluded that when $\gamma > 0$ then:

1. Self-focusing of the weaker-circular component is faster than that of the stronger circular component. As a result, eventually a balance is reached where $|\mathcal{A}_+| \approx |\mathcal{A}_-|$, which corresponds to linear polarization.

2. Linear polarization is stable whereas circular polarization is unstable.

*Note that γ is positive for most liquids [56, Chp. 17].

4.1.2 Subsequent studies of (4.3)

Wagner, Haus, and Marburger used the aberrationless approximation to approximate (4.3) with a system of ODEs [63]. Based on these ODEs they concluded that the self-focusing distance (*i.e.*, the distance for beam collapse) is "sensitive" to small departures from circular polarization. This result suggests that circular polarization is unstable. However, it was noted in [63] that the aberrationless approximation is a only rough approximation of self-focusing dynamics[†]. Prakash and Chandra [54] as well as Vlasov, Korobkin, and Serov [61] studied (4.3) and, using arguments similar to those in [21], reached the conclusion that both linear and circular polarizations are stable. These conclusions are consistent with the experimental observations of Meyer [49], as well as of Skinner and Kleiber [57], and of Golub, Shuker, and Erez [36].

In 1970 Berkhoer and Zakharov [4] showed that the power of each circular amplitude in (4.3) is conserved during the propagation, *i.e.*, that

$$N_{\pm}(z) = \int |\mathcal{A}_{\pm}|^2 \, dxdy = N_{\pm}(0). \tag{4.5}$$

This *proves* that according to (4.3) circular polarization is stable[‡] for both $\gamma > 0$ and $\gamma = 0$, because $N_+(z)/N_-(z) = N_+(0)/N_-(0) \gg 1$. In spite of this, the explanation of Close et al. for instability of circular polarization has persisted in the Nonlinear Optics literature long after 1970 (e.g., in the classic book of Shen [56, Chp. 17]) and even up to these days. Remarkably, the only thing that was always agreed upon is the System (4.3) itself. The derivation of (4.3) is based, however, on the assumptions (4.2), which can be physically incorrect (see Section 4.2). Indeed, our simulations in Section 4.5 show that System (4.3) can lead to wrong predictions.

4.2 New model for self-focusing of circular beams

In this section we present a systematic derivation of the new System (4.10) for propagation of circularly polarized input beams, which we derive from the vector Helmholtz model (2.11). System (4.10) consists of two coupled equations for the two circular polarization amplitudes, which, unlike the standard Close *et al.* model (4.3), takes into account nonparaxiality, the coupling to the axial component, and the grad-div term.

[†]It is now known that applying the aberrationless approximation to the 2D NLS can lead to completely incorrect predictions [27]. Indeed, our numerical simulations in Section 4.5 show that the self-focusing distance is relatively insensitive to the deviation from circular polarization (see footnote in Section 4.5).

[‡]Strictly speaking, power conservation does not imply that the *intensity* of A_- does not become comparable to the intensity of A_+. However, the numerical simulations in Section 4.5 show that A_- does remain much smaller than A_+ during the propagation, in contrast to the qualitative argument of Close et al. .

As in the case of linear polarization (see Subsection 3.1.1), a key dimensionless parameter of the model is $f = 1/k_0 r_0$. We are interested in the case where input beam is almost left-circularly polarized, *i.e.*,

$$\mathcal{E}_-^0/\mathcal{E}_+^0 = O(\varepsilon), \qquad \mathcal{E}_3^0 \ll \mathcal{E}_+^0, \tag{4.6}$$

where ε measures the deviation from perfect input circular-polarization. Therefore, there are two small parameters in the problem, *i.e.*,

$$f \ll 1 \quad \text{and} \quad \varepsilon \ll 1. \tag{4.7}$$

Because f and ε are small we can use perturbation analysis to simplify the vector Helmholtz equation (2.11). To do that, we rescale the variables (as in the case of linear polarization) according to (3.4) and obtain the nondimensional system (3.5). In analogy with (4.1) we denote the nondimensional circular amplitudes by

$$A_\pm(x,y,z) = \frac{1}{\sqrt{2}}\left(A_1 \pm iA_2 \right) \tag{4.8}$$

and the corresponding input amplitudes by $A_\pm^0(x,y) = A_\pm(x,y,z=0)$.

Using a careful perturbation analysis of Eqs. (3.5) we prove in Appendix C.1 the following result:

Lemma 4.1. *Let an almost circularly polarized input beam [*i.e., *that satisfies (4.6)] propagate in a Kerr medium. Assume that the rescaling (3.4) is valid and that f and ε are small. Then the dimensionless amplitudes satisfy*

$$A_-/A_+ = O(f^2, \varepsilon), \tag{4.9a}$$

and

$$A_3/A_+ = \frac{f}{\sqrt{2}}\left(iA_{+,x} + A_{+,y} \right) + O(f^3, \varepsilon f). \tag{4.9b}$$

Since from Lemma 4.1 it follows that $A_- \ll A_+$ and $A_3 \ll A_+$ for $z > 0$, we have the following result:

Proposition 4.2 (Stability of circular polarization - part I). *Under the assumption that the rescaling (3.4) is valid, an almost circularly polarized input beam remains almost circularly polarized during the propagation.*

Proposition 4.2 shows that circular polarization is stable, in agreement with Berkhoer and Zakharov [4] and in contrast to Close *et al.* [21]. Moreover, whereas the standard explanation of Close *et al.*'s for instability of circular polarization assumes that $\gamma > 0$, Proposition 4.2 is independent of the value of γ. An obvious weakness of Proposition 4.2 is that it is based on the assumption that the scaling (3.4) remains valid during the propagation. *A-priori*, the validity of this assumption is questionable, because of the high intensities that can be reached during the self-focusing. However, in Section 3.3 we substantiate Proposition 4.2 by proving that (3.4) remains valid during self-focusing. Numerical simulations in Section 4.5 also corroborate this result.

We recall that all the previous studies used System (4.3), whose derivation is based on the assumption that $\mathcal{E}_3/\mathcal{E}_- \ll 1$. However, from estimates (4.9) it follows that $\mathcal{E}_-/\mathcal{E}_3 = A_-/A_3 = O(f, \varepsilon/f)$. Thus,

Corollary 4.3. *When $\varepsilon \ll f$ the assumption that $\mathcal{E}_3 \ll \mathcal{E}_-$ is wrong.*

Indeed, in Section 4.4 and Section 4.5 we show that System (4.3) leads to completely wrong predictions when $\varepsilon \ll f$. Moreover, we show that even when $f \ll \varepsilon$ this system can lead to wrong predictions (see, *e.g.*, Figure 4.5).

Using the estimates (4.9) we prove in Appendix C the following result:

Proposition 4.4. *Let an almost circularly-polarized input beam [*i.e.*, that satisfies (4.6)] propagate in a Kerr medium. Then to leading order $A_\pm(x, y, z)$ satisfy the coupled system*

$$\underbrace{iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma}|A_+|^2 A_+}_{\text{NLS for } A_+} = -\underbrace{\frac{1}{4} f^2 A_{+,zz}}_{\text{nonparax}} - \underbrace{\frac{1+2\gamma}{1+\gamma}|A_-|^2 A_+}_{\text{coupling to } E_-} \tag{4.10a}$$

$$-\frac{f^2}{2(1+\gamma)}\left[\underbrace{4|\nabla_\perp A_+|^2 A_+ + (\nabla_\perp A_+)^2 A_+^* + |A_+|^2 \Delta_\perp A_+ + A_+^2 \Delta_\perp A_+^*}_{\text{grad}-\text{div}+\text{coupling to } E_3}\right],$$

$$\underbrace{iA_{-,z} + \Delta_\perp A_-}_{\text{LS for } A_-} + \underbrace{\frac{1+2\gamma}{1+\gamma}|A_+|^2 A_-}_{\text{coupling to } E_+} = 0, \tag{4.10b}$$

where

$$|\nabla_\perp A_+|^2 = |A_{+,x}|^2 + |A_{+,y}|^2, \quad (\nabla_\perp A_+)^2 = A_{+,x}^2 + A_{+,y}^2.$$

The terms that are neglected in Eqs. (4.10a) and (4.10b) are $O(f^4, \varepsilon f^2)$ and $O(f^2, \varepsilon^3)$, respectively.

Let us explain the origin of the terms in System (4.10). Eq. (4.10b) is a linear Schrödinger equation for A_- with a nonlinear coupling to A_+ that results from the nonlinear coupling of $\mathcal{E}_+$ to $\mathcal{E}_-$ in (2.11)–(2.12). When $\varepsilon = f = 0$ System (4.10) reduces to the following NLS for A_+:

$$iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma}|A_+|^2 A_+ = 0. \tag{4.11}$$

The $A_{+,zz}$ term on the right-hand-side of Eq. (4.10a) is the nonparaxial term that comes from the scalar Helmholtz equation. The second term results from linear and nonlinear couplings of $\mathcal{E}_+$ to $\mathcal{E}_-$ in (2.11)–(2.12)[§]. The remaining terms in Eq. (4.10a) result from the coupling of $\mathcal{E}_+$ to $\mathcal{E}_3$ and the grad-div term. In Section 3.3 we show that the effect of the $O(f^2)$ terms, which are neglected in the Close *et al.* system (4.3), dominates the effect of the coupling to $\mathcal{E}_-$.

As in the case of linear polarization (see Subsection 3.1.1), here too the $O(f^2)$ on the right-hand side of (4.10a) do not vanish even if one assumes that $\mathcal{E}_- \equiv \mathcal{E}_3 \equiv 0$. Indeed, a close inspection of the derivation of this equation shows that if we assume that $\mathcal{E}_- \equiv \mathcal{E}_3 \equiv 0$, then the resulting equation for A_+ [instead of Eq. (4.10a)] is

$$iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma}|A_+|^2 A_+ = -\frac{f^2}{4}A_{+,zz} - \underbrace{\frac{f^2}{2(1+\gamma)}\Delta_\perp(|A_+|^2 A_+)}_{\text{grad}-\text{div E}_+}\ , \tag{4.12}$$

where the second term on the right-hand side corresponds to the contribution of $\mathcal{E}_+$ from the grad-div term in the vector Helmholtz equation (2.11a), which can also be written as

$$-\underbrace{\frac{f^2}{2(1+\gamma)}\Delta_\perp(|A_+|^2 A_+)}_{\text{grad}-\text{div E}_+} \equiv -\frac{f^2}{2(1+\gamma)}\left[\,4|\nabla_\perp A_+|^2 A_+ + 2(\nabla_\perp A_+)^2 A_+^* + 2|A_+|^2\Delta_\perp A_+ + A_+^2\Delta_\perp A_+^*\,\right].$$

The remaining $O(f^2)$ terms in the square brackets on the right-hand side of Eq. (4.10a) correspond to the coupling to $\mathcal{E}_3$. If only those terms are taken into account (*i.e.*, neglecting nonparaxiality and the grad-div-E_+ terms) the resulting equation is

$$iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma}|A_+|^2 A_+ = \underbrace{+\frac{f^2}{2(1+\gamma)}\left[\,(\nabla_\perp A_+)^2 A_+^* + |A_+|^2\Delta_\perp A_+\,\right]}_{\text{coupling to E}_3}. \tag{4.13}$$

To recap, the $O(f^2)$ terms in the square brackets on the right-hand side right-hand side of (4.10a)

[§]A close inspection of the derivation of Eq. (4.10a) shows that the nonlinear interaction between $\mathcal{E}_+$ and itself in the divergence term in Eqs. (2.11)–(2.12) also contributes to these $O(f^2)$ terms, *i.e.*, these terms do not vanish even if one assumes that $\mathcal{E}_- \equiv \mathcal{E}_3 \equiv 0$ [see Eq. (C.19) in Appendix C.2].

correspond to the contribution of the grad-div-E_+ term and and the coupling to $\mathcal{E}_3$:

$$\frac{f^2}{2(1+\gamma)}\left[\underbrace{4|\nabla_\perp A_+|^2A_+ + (\nabla_\perp A_+)^2A_+^* + |A_+|^2\Delta_\perp A_+ + A_+^2\Delta_\perp A_+^*}_{\text{grad}-\text{div}+\text{coupling to E}_3}\right] \tag{4.14}$$

$$\equiv \underbrace{\frac{f^2}{2(1+\gamma)}\Delta_\perp(|A_+|^2A_+)}_{\text{grad}-\text{div E}_+} - \underbrace{\frac{f^2}{2(1+\gamma)}\left[(\nabla_\perp A_+)^2A_+^* + |A_+|^2\Delta_\perp A_+\right]}_{\text{coupling to E}_3}.$$

4.2.1 Power conservation

As mentioned in Subsection 4.1.2, Berkhoer and Zakharov [4] showed that the power of each circular amplitude is conserved in the System (4.3). The same holds, to leading order, in (4.10). To see that, we multiply (4.10) by $A_\pm^*$, integrate over the (x, y) plane, and subtract the complex conjugate to obtain that

$$\int\left[i(A_\pm^*A_{\pm,z} \pm A_\pm A_{\pm,z}^*) + (A_\pm^*\Delta_\perp A_\pm - A_\pm\Delta_\perp A_\pm^*)\right]dxdy = O(f^2),$$

where the $O(f^2)$ error corresponds to the $O(f^2)$ terms on the right-hand side of (4.10a). Using the identity $(|A_\pm|^2)_z \equiv A_\pm^*A_{\pm,z} + A_\pm A_{\pm,z}^*$ and Green's Theorem leads to the conservation law

$$N_\pm(z) = \int |A_\pm|^2\,dxdy = N_\pm(0) + O(f^2). \tag{4.15}$$

In other words, the power of each circular component is conserved with $O(f^2)$ accuracy. Since (4.6) implies that $N_-(0)/N_+(0) = O(\varepsilon^2)$, we conclude from Eq. (4.15) that $N_-(z)/N_+(z) = O(f^2, \varepsilon^2)$, i.e., that almost all of the beam's power remains in the left-circular component during the propagation. As we have noted, strictly speaking, power conservation does not imply that the *intensity* of A_- does not become comparable to the intensity of A_+. However, the numerical simulations of System (4.10) show that A_- does remain $O(\varepsilon)$ smaller than A_+ during the propagation (see Section 4.5).

4.3 Early stage of propagation

During the early stage of the propagation the $O(f^2)$ terms in System (4.10) are small, and the model can be simplified by setting $f = 0$. The resulting system of equations is

$$iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma}\left[|A_+|^2 + (1+2\gamma)|A_-|^2\right]A_+ = 0, \tag{4.16a}$$

$$iA_{-,z} + \Delta_\perp A_- + \frac{1+2\gamma}{1+\gamma}|A_+|^2A_- = 0. \tag{4.16b}$$

Below we prove that solutions of (4.16) can undergo catastrophic collapse. Therefore, the simplified model (4.16) can be used for the early stage of the propagation, but fails to describe the propagation near and beyond the blowup point.

We note the Systems (4.3) and (4.16) are almost identical[¶], the only difference being that the equation for A_- in (4.3) includes the $|A_-|^2 A_-$ term. This terms is negligible, however, because it is $O(\varepsilon^2)$ smaller than the $|A_+|^2 A_-$ term in (4.10b). Indeed, numerical simulations in Section 4.5 show that solutions of Systems (4.3) and (4.16) are almost indistinguishable.

4.3.1 Collapse of circularly polarized beams

We now prove that solutions of (4.16) can collapse at a finite propagation distance. To do that, we first observe that (4.16) conserves the two powers $N_\pm(z)$ [see (4.5)] as well as the Hamiltonian

$$H(z) = \int (|\nabla A_+|^2 + |\nabla A_-|^2)\, dxdy - \frac{1}{2(1+\gamma)} \int |A_+|^4\, dxdy$$
$$-\frac{1+2\gamma}{2(1+\gamma)} \int |A_+|^2 |A_-|^2\, dxdy. \tag{4.17}$$

The latter is proved by multiplying (4.16) by $A^*_{z,\pm}$, adding the two equations and their complex conjugates, integrating over the (x, y) plane, and taking the real part.

In addition, we have the following result:

Lemma 4.5 (Variance Identity for (4.16)). *Let $A_\pm(x, y, z)$ be the solution of (4.16) and let*

$$V(z) = \int (|A_+|^2 + |A_-|^2)(x^2 + y^2)\, dxdy \tag{4.18}$$

be its variance. Then

$$V_{zz} = 8H(z), \tag{4.19}$$

where $H(z)$ is defined by (4.17).

Eq. (4.19) can be proved by differentiating $V(z)$ twice with respect to z, using (4.16) to replace z derivatives with transverse derivatives, and integrating by parts.

As in the case of the NLS, from Hamiltonian conservation (4.17) and the variance identity (4.19) it follows that when $H(0) < 0$ the variance would become negative at a finite propagation distance. Since by definition the variance has to be positive, this implies that the solution blows up at a finite propagation distance:

[¶]Indeed, roughly speaking, $f = 0$ corresponds to assumptions (4.2) and the paraxial approximation.

Proposition 4.6. *Let $H(z)$ be given by (4.17). Then $H(0) < 0$ is a sufficient condition for collapse of solutions of (4.16).*

A similar result, of course, holds for System (4.3):

Proposition 4.7. *Let $A_\pm(x, y, z)$ be the solution of (4.3) and let $H(z)$ be given by*

$$H(z) = \int (|\nabla A_+|^2 + |\nabla A_-|^2)\, dxdy - \frac{1}{2(1+\gamma)} \int (|A_+|^4 + |A_-|^4)\, dxdy$$
$$-\frac{1+2\gamma}{2(1+\gamma)} \int |A_+|^2 |A_-|^2\, dxdy.$$

Then $H(z) = H(0)$ and $V_{zz} = 8H(z)$, where $V(z)$ is given by (4.18). Therefore, $H(0) < 0$ is a sufficient condition for collapse of solutions of (4.3)

Remark. Propositions 4.6 and 4.7 show that the coupling to $\mathcal{E}_-$ does not arrest the collapse. In fact, because the coupling term to A_- in the square brackets in (4.16a) appears with a positive coefficient, it *accelerates* the collapse. In Section 3.3 we show that, in contrast, the combined effect of the coupling to $\mathcal{E}_3$ and the grad-div term can arrest the collapse.

4.3.2 Threshold power

Because the $|A_-|^2 A_+$ term in (4.16a) is $O(\varepsilon^2)$ small compared with the $|A_+|^2 A_+$ term, to leading order this term is negligible and System (4.16) can be further approximated with the semi-decoupled system

$$iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma}|A_+|^2 A_+ = 0, \tag{4.20a}$$

$$iA_{-,z} + \Delta_\perp A_- + \frac{1+2\gamma}{1+\gamma}|A_+|^2 A_- = 0. \tag{4.20b}$$

Here the equation for A_+ is decoupled from A_-. If we rescale A_+ in (4.20) as

$$\psi(x, y, z) = (1+\gamma)^{-1/2} A_+(x, y, z)\,, \tag{4.21}$$

then (4.20a) becomes the NLS (2.16), showing that the threshold power for collapse of circularly polarized beams is

$$N_c^{\text{circ}} = (1+\gamma) N_c. \tag{4.22}$$

In dimensional variables this threshold power is given by

$$\mathcal{P}_c^{\text{circ}} = \frac{\lambda^2}{4\pi n_0 n_2}(1+\gamma) N_c, \tag{4.23}$$

where $n = n_0 + n_2 I$, and I is the intensity.

4.4 Single filament dynamics

Similar to Section 3.3, we can use modulation theory to describe the dynamics of a single circularly-polarized filament, whose power is not much higher than N_c^{circ}. By applying modulation theory to System (4.10) we prove in Appendix D that

Proposition 4.8. *When $f, \varepsilon \ll 1$ self-focusing dynamics of a single circularly-polarized filament is given, to leading order, by the reduced system of ODEs*

$$\begin{cases} L_{zz}(z) &= -\dfrac{\beta}{L^3}, \\ \beta_z(z) &= -\dfrac{f^2(C_{\mathrm{nonparax}} + C_{\mathrm{vec}})N_c}{2M}\left(\dfrac{1}{L^2}\right)_z + C_{E_-}(z), \end{cases} \tag{4.24}$$

where $\beta \ll 1$, $M = \frac{1}{4}\int\limits_0^\infty \rho^2 R^2 \rho\, d\rho \approx 0.55$, $C_{\mathrm{nonparax}} = 1$, $C_{\mathrm{vec}} \approx 16/3$, and $C_{E_-}(z) = O(f^2, \varepsilon)$.

Inspection of the derivation of (4.24) shows that the terms with C_{nonparax}, C_{vec}, and $C_{E_-}(z)$ correspond to the nonparaxiality, grad-div term and coupling to $\mathcal{E}_3$, and coupling to $\mathcal{E}_-$ in Eq. (4.10a), respectively.

Based on (4.24) we make the following observations:

1. To leading order, nonparaxiality, the grad-div term, and coupling to $\mathcal{E}_3$ have the same qualitative effect on self-focusing of a single filament. This observation is surprising, because at the PDE level [*i.e.*, System (4.10)] the expressions corresponding to nonparaxiality and to coupling to $\mathcal{E}_3$ are completely different.

2. $C_{\mathrm{vec}} \approx 5.3 C_{\mathrm{nonparax}}$. Thus, System (4.24) shows that the combined effect of the grad-div term and the coupling to $\mathcal{E}_3$ is more than five times stronger than nonparaxiality.

3. When $\varepsilon \ll 1$ the term corresponding to C_{E_-} is much smaller than the other terms in (4.24). Therefore, System (4.24) shows that the coupling to $\mathcal{E}_-$ is negligible compared with nonparaxiality, the coupling to $\mathcal{E}_3$, and the grad-div term. We thus see that *the Close* et al. *model (4.3) is wrong even when $\varepsilon \gg f$.*

4. Except for the small term that corresponds to the coupling to $\mathcal{E}_-$, System (4.24) is independent of γ. Indeed, if one neglects the coupling to $\mathcal{E}_-$ in Eq. (4.10a) and rescales A_+ as in (4.21) then γ is "eliminated" from the rescaled equation. Thus, we see again that *the value of γ (*i.e.*, whether it is zero or positive) has no effect on circular polarization stability*, in agreement with Berkhoer and Zakharov [4] and in contrast to Close *et al.* [21] (see Section 4.1).

If we neglect the coupling to $\mathcal{E}_-$ (*i.e.*, set $C_{E_-} = 0$), then (4.24) is the same as (3.16), the only difference is in the constants corresponding to the vectorial effects. Therefore, we can conclude that collapse is arrested by nonparaxiality and the contribution of $\mathcal{E}_+$ from the grad-div term, followed by focusing-defocusing oscillations when $H(0) < 0$. In addition, the minimal filament width is $L_m \sim L(0) f \sqrt{N_c^{\rm circ}(C_{\rm nonparax} + C_{\rm vec})/4M\beta(0)}$. Since $\beta(0) \ll 1$, even at this stage the magnitude of the $O(f^2)$ terms in (4.10) is $O(\beta(0))$ smaller than that of the NLS terms $\Delta_\perp A_+$ and $|A_+|^2 A_+$, providing an *a-posteriori* justification for treating the right-hand side terms in (4.10) as small perturbations and showing that

Corollary 4.9. *The scaling of the variables (3.4) remains valid during the propagation.*

This enables us to remove the assumption in Proposition 4.2:

Proposition 4.10 (Stability of circular polarization - part II). *When an almost circularly polarized filament [*i.e.*, that satisfies (4.6)] with power moderately above $N_c^{\rm circ}$ propagate in a Kerr medium, the filament remains almost circularly polarized (*i.e.*, $A_- \ll A_+$) for all $z > 0$.*

Finally, we recall that the $O(f^2)$ terms in Eq. (4.10a) correspond, in part, to the contribution of $\mathcal{E}_+$ from the grad-div term in the vector Helmholtz equation (2.11a), which does not vanish even when $\mathcal{E}_- \equiv \mathcal{E}_3 \equiv 0$. A closer inspection of the derivation reveals that $C_{\rm vec} = C_{{\rm grad-div}\ E_+} + C_{E_3}$, where the constants

$$C_{{\rm grad-div}\ E_+} \approx 8 \quad \text{and} \quad C_{E_3} \approx -\frac{8}{3}$$

correspond to contribution of $\mathcal{E}_+$ from the grad-div term and the coupling effects to $\mathcal{E}_3$ [see Eq. (4.14)]. This result is surprising, because it shows that the defocusing effect of the contribution of $\mathcal{E}_+$ from the grad-div is *eight times stronger* than nonparaxiality. Also surprising is that the effect of the coupling to $\mathcal{E}_3$ is weakly *focusing*, since $C_{E_3} < 0$.

4.5 Stability of circular-polarization - simulations

In this section we confirm Proposition 4.10 and the predictions of modulation theory, by solving (4.10) for the almost circularly polarized, Gaussian input beams

$$A_\pm^0(x,y) = \sqrt{N(0)}\, e^{-x^2-y^2}(1 \pm e^{i\varepsilon}), \tag{4.25}$$

where the input power $N(0)$ is moderately above $N_c^{\rm circ}$ (4.22) and ε is the input "ellipticity angle" (*i.e.*, $\varepsilon = 0$ corresponds to a perfectly left-circularly polarized beam). Below we study different asymptotic regimes of the parameters f and ε.

Figure 4.3A confirms that when nonparaxiality and the coupling to the axial component are neglected (*i.e.*, $f = 0$) solutions of Systems (4.3), (4.20), and (4.16) can undergo catastrophic collapse, and that the solutions of these three systems are almost indistinguishable (see Section 4.3). Figs. 4.3B and 4.4 show that $A_-/A_+ = O(\varepsilon)$, *i.e.*, that the solutions remain circularly polarized to leading order. As the intensity of A_+ grows during the self-focusing process, the intensity of A_- also grows because of the coupling to A_+. However, A_- always remains smaller than A_+, because it has insufficient power for an independent collapse (see Subsection 4.2.1). Figure 4.4 also confirms that A_+ approaches a modulated Townes profile during the collapse, *i.e.*,

$$|A_+(x,y,z)| \sim \sqrt{1+\gamma}\,\frac{1}{L(z)}R(\frac{r}{L(z)}), \tag{4.26}$$

where $R(r)$ is defined in Eq. (2.20), which justifies the application of modulation theory (see Section 3.3).

When $0 < f \ll \varepsilon$ one might expect that the coupling to A_- in (4.10) would dominate the coupling to A_3 and nonparaxiality. In fact, quite the opposite is true. When $f > 0$ the picture of self-focusing completely changes, because the $O(f^2)$ terms in (4.10), which correspond to nonparaxiality and the coupling to A_3, arrest beam collapse and lead to focusing-defocusing oscillations, as predicted by modulation theory in Section 3.3 (see Figure 4.5A). In addition, the beam remains almost circularly polarized during propagation (see Figure 4.5B). Thus, even in this regime System (4.3) leads to wrong predictions.

When $f = O(\varepsilon)$ and when $f \gg \varepsilon$ the picture is qualitatively similar to the case $0 < f \ll \varepsilon$ (see Figs. 4.6 and 4.7). Moreover, the value of ε seems to have negligible quantitative effect on the dynamics of A_+, as can be seen by comparing Figs. 4.5A and 4.6A[‖]. Finally, focusing-defocusing oscillations and circular-polarization stability are also observed when the input power is much higher than $N_c^{\rm circ}$, a regime that is formally beyond the validity of modulation theory (see Figure 4.8).

[‖] This comparison also shows that the deviation from circular polarization has a small effect on the location of the (first) focal point (see footnote in Subsection 4.1.2).

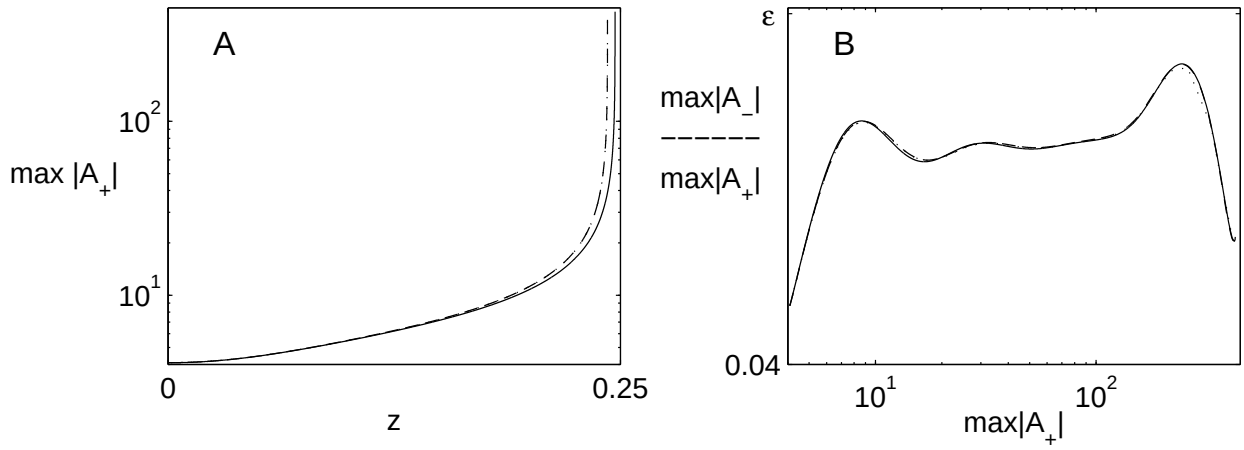

Figure 4.3: Solutions of System (4.3) (dashed), System (4.16) (dotted), and System (4.20) (solid): (A) on-axis amplitude; (B) deviation from circular polarization. Here $\gamma = 0.5$ and the input beam is (4.25) with $\varepsilon = 0.1$ and $N(0) = 1.5N_c^{\text{circ}}$.

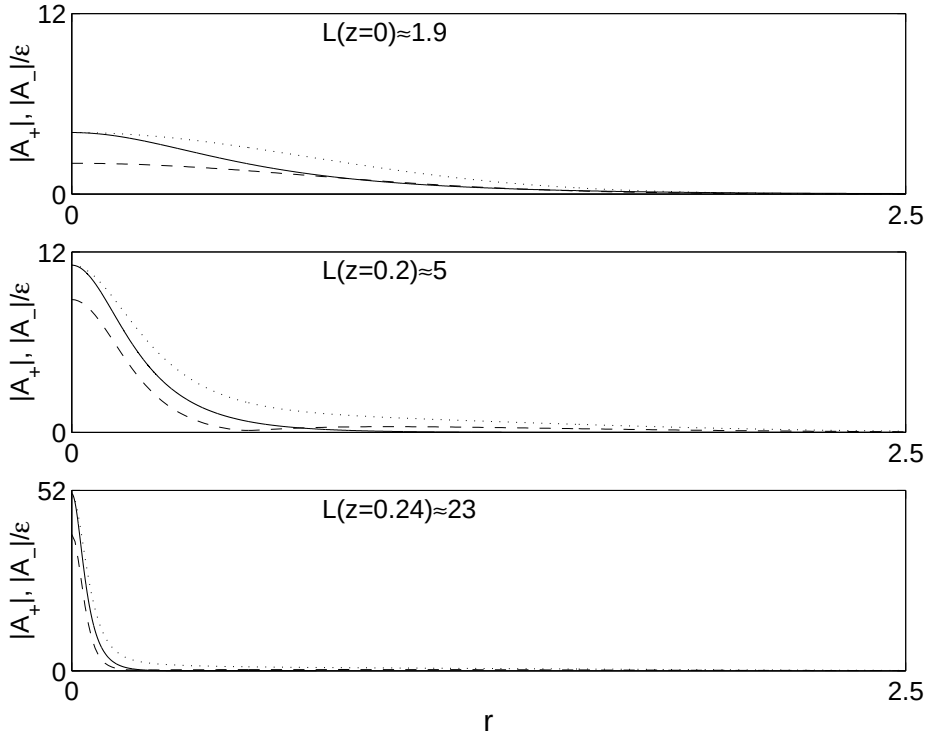

Figure 4.4: $|A_+|$ (solid) converges to the modulated Townes profile (4.26) (dots) during beam collapse. Dashed line is $|A_-|/\varepsilon$. Shown is the solution of (4.16) from Figure 4.3.

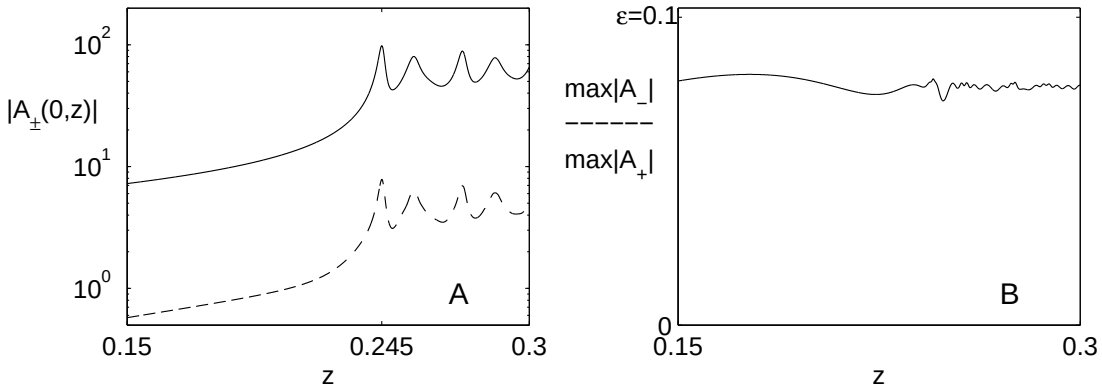

Figure 4.5: Solutions of System (4.10): (A) peak amplitudes of A_+ (solid) and A_- (dashed); (B) deviation from circular polarization. Same γ and input beam as in Figure 4.3 with $f = 0.01$ and $\varepsilon = 0.1$.

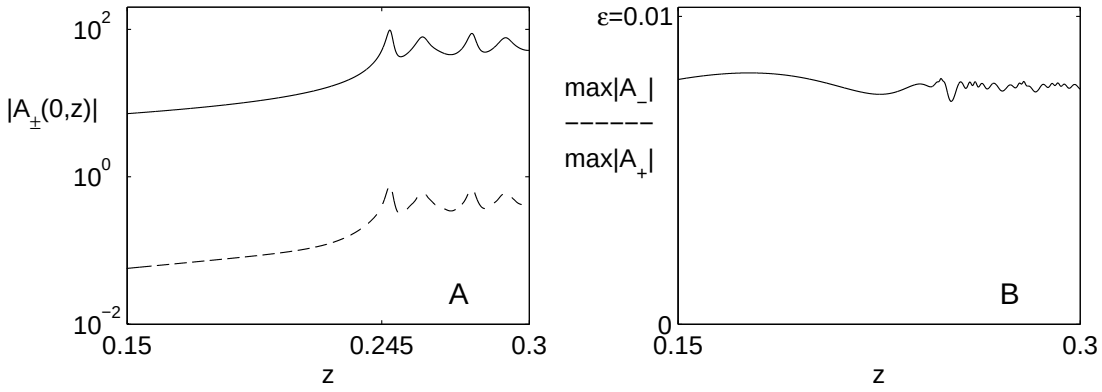

Figure 4.6: Same as Figure 4.5 with $f = \varepsilon = 0.01$.

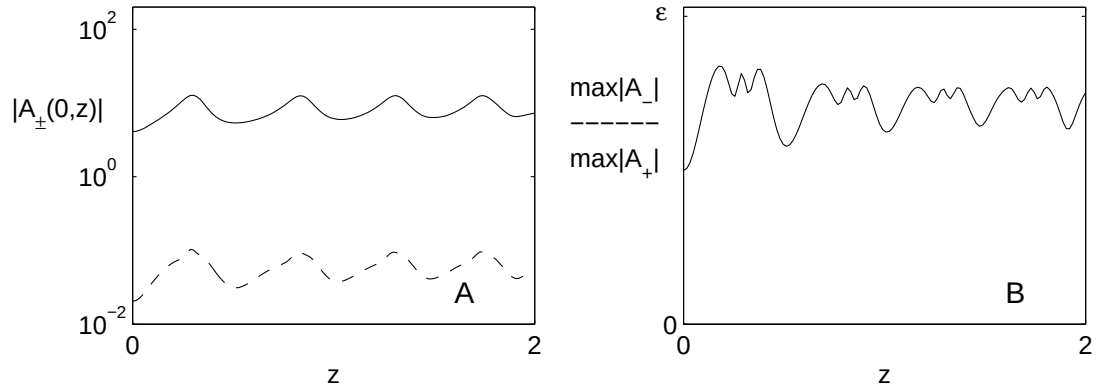

Figure 4.7: Same as Figure 4.5 with $f = 0.1$ and $\varepsilon = 0.01$.

4.6 Multiple filamentation

In Section 3.2 we showed that the preferred direction induced by *linearly polarized* beams can lead to multiple filamentation, even of cylindrically symmetric input beams. Let us now consider the "ideal" case of a circularly polarized input beam with a cylindrically-symmetric profile. Since in that case neither the medium nor the input beam induce a preferred direction in the (x, y) plane, we can make the following observation:

Corollary 4.11. *Let a circularly polarized, cylindrically-symmetric input beam (*i.e., $A_+^0 = A_+^0(r)$ *and* $A_-^0 = 0$*) propagate in a Kerr medium. Then the beam remains cylindrically symmetric during the propagation. In particular, the beam does not undergo multiple filamentation.*

In the case of a cylindrically-symmetric input profile (*i.e.*, $A_\pm = A_\pm^0(r)$) with a small deviation from circular polarization state (*i.e.*, $A_-^0 \ll A_+^0$), the beam will not remain cylindrically symmetric during its propagation, because the initial condition $\vec{\mathcal{E}}^0$ of the vector Helmholtz model (2.11) is not rotation invariant as a vector entity. However, because Eqs. (4.10) are isotropic, when the input profile is cylindrically symmetric, then according to (4.10) the beam would remain cylindrically symmetric, *i.e.*, $A_\pm = A_\pm(r, z)$ for all $z > 0$. This difference is caused by the anisotropic $O(\varepsilon f^2)$ terms that are neglected in (4.10a), which account for the symmetry-breaking in the Helmholtz model (2.11).** Thus, we conclude that

Corollary 4.12. *Let an almost circularly polarized, cylindrically-symmetric input beam [*i.e., $A_-^0 \ll A_+^0$ *and* $A_\pm^0 = A_\pm^0(r)$*] propagate in a Kerr medium. Then to leading order the beam remains cylindrically symmetric during the propagation,* i.e., $A_+ = A_+(r, z) + O(\varepsilon f^2)$ *for all* $z > 0$.

Corollary 4.12 suggests that cylindrically symmetric, almost circularly polarized beams would not undergo multiple filamentation. This result is not conclusive, of course, as (4.10) neglects the $O(\varepsilon f^2)$ symmetry-breaking terms in the vector Helmholtz model (2.11). Indeed, in the "extreme" case of linear polarization (*i.e.*, $\varepsilon = 1$) these $O(f^2)$ symmetry-breaking terms can lead to multiple filamentation (see Chapter 5).

4.6.1 Scalar equation for circularly polarized beams

In Corollaries 4.11 and 4.12 we assumed that the input beam is cylindrically symmetric. Such idealization, however, is unrealistic, as there is always some degree of imperfection when generating an input

**A close inspection of the derivation of (4.10) shows that the $O(f^4)$ terms that are neglected in (4.10a) are isotropic, because they correspond to higher-order effects of nonparaxiality and the coupling of $\mathcal{E}_+$ to $\mathcal{E}_3$.

beam. Because the coupling to A_- in Eq. (4.10a) is $O(\varepsilon^2)$ small (see Subsection 4.2.1), System (4.10) can be approximated with the scalar equation

$$iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma}|A_+|^2 A_+ = -\frac{1}{4}f^2 A_{+,zz} \tag{4.27}$$
$$-\frac{f^2}{2(1+\gamma)}\left[\, 4|\nabla_\perp A_+|^2 A_+ + (\nabla_\perp A_+)^2 A_+^* + |A_+|^2 \Delta_\perp A_+ + A_+^2 \Delta_\perp A_+^* \,\right].$$

In Sections 4.6.2 and 4.6.3 we use (4.27) to study whether small imperfections in the input profile can lead to multiple filamentation of circularly polarized beams.

4.6.2 Noise induced multiple filamentation

As discussed in Subsection 3.4.1, the standard theoretical explanation of multiple filamentation (of linearly polarized beams) was suggested by Bespalov and Talanov [5]. In order to test whether noise can lead to multiple filamentation of circularly polarized beams we first solve the Close *et al.* system (4.3) with very high-power input beams, to which we add noise both in amplitude and in phase, *i.e.*,

$$A_\pm^0(x,y) = \sqrt{N(0)}\, e^{-(x^2+y^2)}(1 \pm e^{i\varepsilon})\,\left[\, 1 + c * noise(x,y)\,\right], \tag{4.28}$$

where $N(0)$ is the noiseless input power, $noise(x,y)$ is a random complex-valued function, and the constant c determines the noise level ($c \ll 1$). In our simulations we see neither evidence for multiple filamentation nor even for mild instabilities. Rather, the beam collapses while converging to a cylindrically-symmetric profile (see Figs. 4.9 and 4.10). We do see, however, multiple filamentation when we solve Eq. (4.27) with noisy very powerful input beams, *i.e.*,

$$A_+^0(x,y) = \sqrt{N(0)}\, e^{-(x^2+y^2)}\,\left[\, 1 + c * noise(x,y)\,\right], \tag{4.29}$$

where $N(0)$ is several times $N_c^{\rm circ}$. For example, in Figure 4.11 we show a noisy beam with ten times the threshold power that breaks-up into three filaments.[††] Note that the difference between the scalar equation (4.27) and the Close *et al.* system (4.3) is the $O(f^2)$ nonparaxial and coupling to $\mathcal{E}_3$ terms, both of which are isotropic. Thus, these terms cannot lead to multiple filamentation by themselves. Nevertheless, as Figure 4.9 shows, they are necessary for noise-induced multiple filamentation, (see Section 5.2 for further discussion).

[††] A closer inspection of the results reveals that after the breakup has occurred each of the filaments undergoes focusing-defocusing oscillations, as predicted by modulation theory (see Section 3.3) and is roughly cylindrically symmetric (see Figure 4.13).

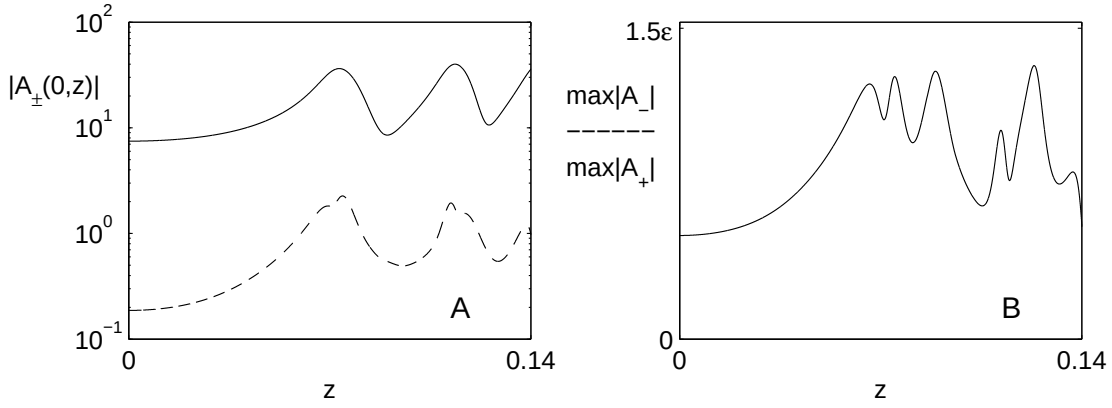

Figure 4.8: Same as Figure 4.5 with $N(0) = 5N_c^{\rm circ}$ and $f = \varepsilon = 0.05$.

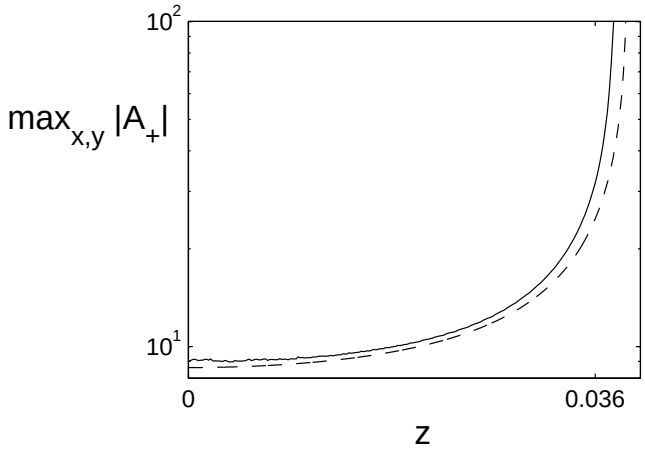

Figure 4.9: Peak amplitude of the solution of System (4.3) with the noisy input beam (4.28) with $\varepsilon = 0.1$, $c = 0.1$ (*i.e.*, 10% noise), and $N(0) = 10N_c^{\rm circ}$ (solid), and the same input beam without the noise (*i.e.*, $c = 0$, dashed).

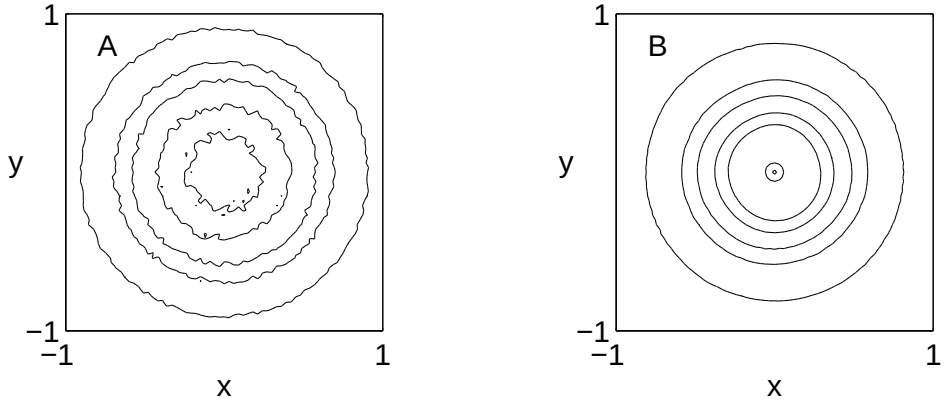

Figure 4.10: Contour plots of (A) $|A_+(x, y, z = 0)|$ and (B) $|A_+(x, y, z = 0.036)|$ of the solution of Figure 4.9 (solid).

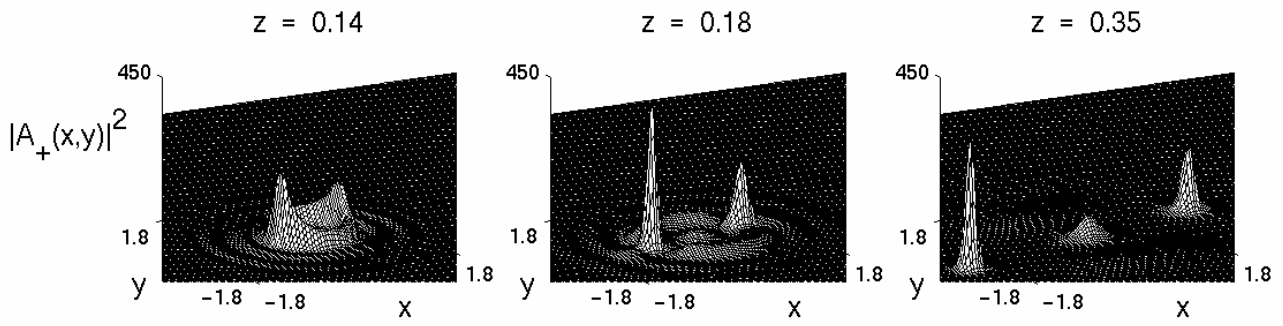

Figure 4.11: Intensity of the solution of Eq. (4.27) with $f = 0.05$ and the noisy input beam (4.28) with $c = 0.1$ and $N(0) = 10N_c^{\text{circ}}$.

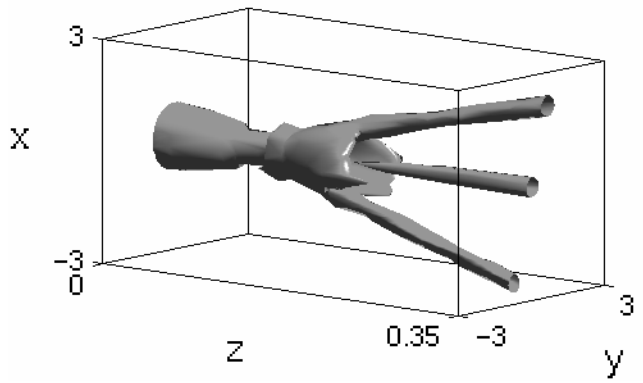

Figure 4.12: Isosurface of $|A_+|^2$ of the solution in Figure 4.11.

4.6.3 Astigmatism induced multiple filamentation

Optical devices, such as those used for producing circularly-polarized beams, are known to produce astigmatic beams (see, *e.g.*, [8]). To study whether astigmatism can also lead to multiple filamentation we consider the input beam

$$A_+^0(x,y) = 2\sqrt{CN_c^{\rm circ}}\,\exp\left[-(ex)^2 - y^2\right], \tag{4.30}$$

where C is constant and $0 \leq e \leq 1$ is input astigmatism parameter ($e = 1$ corresponds to a cylindrically-symmetric input beam). Our simulations of Eq. (4.27) show that input astigmatism can lead to multiple filamentation when the input power is several times $N_c^{\rm circ}$[‡‡]. For example, Figure 4.14 shows astigmatic input beams that break-up into two and three filaments. As with noise-initiated multiple filamentation, during further propagation each of the filaments undergoes focusing-defocusing oscillations and is roughly cylindrically symmetric.

[‡‡]We note that the effective threshold power for collapse of astigmatic beams is higher than for cylindrically-symmetric beams by a factor of approximately $[0.2(e + 1/e) + 0.6]$ (see [27]).

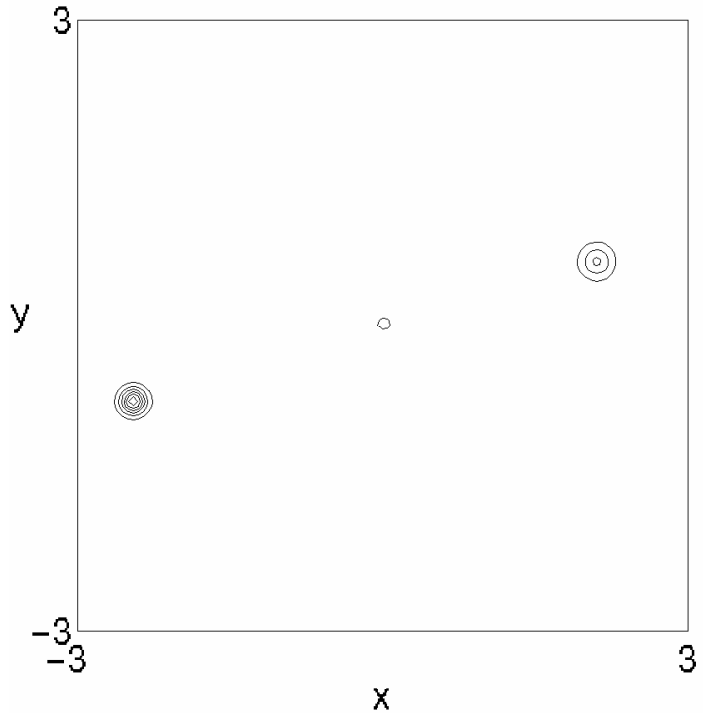

Figure 4.13: Contour plot of $|A_+(x, y, z = 0.35)|$ of the solution of Figure 4.11.

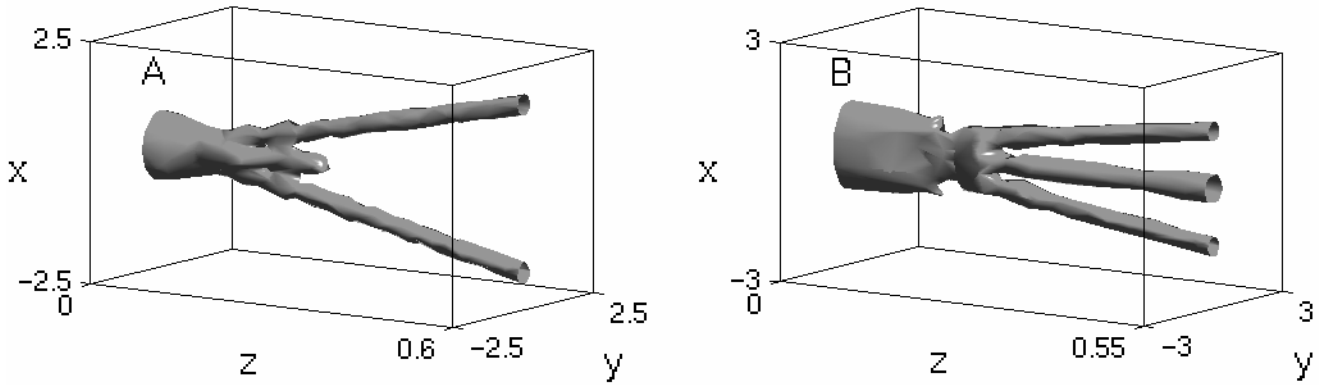

Figure 4.14: Isosurface of $|A_+|^2$ of the solutions of Eq. (4.27) with $f = 0.05$ and astigmatic input beams (4.30) with (A) $C = 7.5$ and $e = 0.9$ [*i.e.*, $N(0) = 8.3N_c^{\text{circ}}$]; (B) $C = 3.75$ and $e = 0.6$ [*i.e.*, $N(0) = 6.25N_c^{\text{circ}}$].

Chapter 5

Comparison of linear and circular polarizations

It is instructive to compare the results of linearly polarized input beams with those of circularly polarized ones. In Subsection 3.1.1 we showed that when the input beam is linearly polarized in the x direction and $f \ll 1$, Eqs. (2.11)–(2.12) can be approximated with the scalar equation (3.7):

$$\underbrace{iA_{1,z} + \Delta_\perp A_1 + |A_1|^2 A_1}_{\text{NLS}} = -f^2 \left[\underbrace{\frac{1}{4} A_{1,zz}}_{\text{nonparax}} + \underbrace{\frac{4+6\gamma}{1+\gamma} |A_{1,x}|^2 A_1 + (A_{1,x})^2 A_1^* + \frac{1+2\gamma}{1+\gamma} \left(|A_1|^2 A_{1,xx} + A_1^2 A_{1,xx}^* \right)}_{\text{grad}-\text{div}+\text{coupling to } E_3} \right] ,$$

where the variables are rescaled according to Eq. (3.4) and A_1 is the nondimensional amplitude in the x direction. Using the same rescaling, we showed in Subsection 4.6.1 that when the input beam is (perfectly) circularly polarized and $f \ll 1$, Eqs. (2.11)–(2.12) can be approximated with the scalar equation (4.27):

$$\underbrace{iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma} |A_+|^2 A_+}_{\text{NLS}} = -f^2 \left[\underbrace{\frac{1}{4} A_{+,zz}}_{\text{nonparax}} + \underbrace{\frac{1}{2(1+\gamma)} \left[4|\nabla_\perp A_+|^2 A_+ + (\nabla_\perp A_+)^2 A_+^* + |A_+|^2 \Delta_\perp A_+ + A_+^2 \Delta_\perp A_+^* \right]}_{\text{grad}-\text{div}+\text{coupling to } E_3} \right] .$$

In the limiting case of $f = 0$ these equations show that the threshold power for self-focusing of circularly polarized beams is higher by $(1+\gamma)$ than for the linearly polarized beams [see Eq. (4.22)]. In addition, we recall that in the case of circular polarization γ can be "factored out" of the equation

[see Eq. (4.16a)]. Therefore, when $\mathcal{E}_-^0$ is negligible then, to leading order, the constant γ does not affect the beam dynamics, other than to increase the threshold power (4.23). In contrast, γ cannot be factored out of the corresponding scalar equation for *linearly*-polarized input beams (3.7).

5.1 Single filament dynamics

Applying modulation theory for the case of linear polarization [*i.e.*, to Eq. (3.7)] and for the case of circular polarization [*i.e.*, to System (4.10)] leads to the reduced systems (3.16) and (4.24), respectively. These reduced systems are qualitatively the same and very similar quantitatively: in both systems the modulation constant corresponding to nonparaxiality is $C_{\rm nonparax} = 1$ and the only (small) difference between these reduced systems is in the constant corresponding to vectorial effects, which for linear polarization is

$$C_{\rm vec}(\gamma) \approx \frac{16}{3}\left(1 + \frac{\gamma}{1+\gamma}\right)$$

and for circular polarization is $C_{\rm vec} \approx 16/3$. Therefore, the reduced systems show that both nonparaxiality and vectorial effects are defocusing, arrest collapse, and lead to focusing-defocusing oscillations.

We recall that in the case of linear polarization the $O(f^2)$ terms in the reduced PDE (3.7)] corresponds, in part, to the contribution of $\mathcal{E}_1$ from the grad-div term, which does not vanish even when $\mathcal{E}_2 \equiv \mathcal{E}_3 \equiv 0$. Similarly, in the case of circular polarization the $O(f^2)$ terms in (4.10a) correspond, in part, to the contribution of $\mathcal{E}_+$ from the grad-div term, which does not vanish even when $\mathcal{E}_- \equiv \mathcal{E}_3 \equiv 0$. A closer inspection of the derivation of the corresponding reduced systems reveals that the constants corresponding to the contribution $\mathcal{E}_1$ and $\mathcal{E}_+$ from the grad-div term are $C_{\rm grad-div\ E_1} = C_{\rm grad-div\ E_+} \approx 8$. These results are surprising, because they show that for both linear and circular polarizations the contribution of $\mathcal{E}_1$ or $\mathcal{E}_+$ from the grad-div term is a defocusing defocusing mechanism that is *eight times stronger* than nonparaxiality.

Also surprising is that the coupling to $\mathcal{E}_3$ is considerably weaker than the contribution of $\mathcal{E}_1$ or $\mathcal{E}_+$ from the grad-div term, and that this coupling is *focusing*: for linear polarization it is focusing for $|\gamma| < 1$, because

$$C_{E_3}(\gamma) \approx -\frac{8(1-\gamma)}{3(1+\gamma)}\,,$$

and for circular polarization it is focusing independently of γ, because $C_{E_3} \approx -8/3$. Based on previous studies of vectorial effects on self-focusing, the general notion has been that vectorial effects are defocusing and are (mainly) due to the coupling to $\mathcal{E}_3$. Thus, the interpretation of the results in previous studies suggested that the coupling to $\mathcal{E}_3$ is a defocusing mechanism. In contrast, our study

shows that the dominant (defocusing) effects results from the contribution of $\mathcal{E}_1$ or $\mathcal{E}_+$ from the grad-div term (which are, strictly speaking, scalar effects), whereas the coupling to $\mathcal{E}_3$ is, in fact, weakly defocusing. In retrospect, this result may not seem so surprising, as it shows that the coupling to $\mathcal{E}_3$ is slightly focusing, which is similar to the coupling to $\mathcal{E}_-$ in the case of circular polarization.

5.2 Multiple filamentation

In Chapter 3 we point out that the preferred direction induced by linear polarization of an input-beam breaks up the cylindrical symmetry in the vector Helmholtz model (2.11). Numerical simulations in Section 3.2 show that this symmetry breaking, which is manifested by the anisotropic $O(f^2)$ terms in (3.7), can lead to multiple filamentation, even when the input profile is perfectly cylindrically symmetric, *i.e.*, $\mathcal{E}_1^0 = \mathcal{E}_1^0(r)$ and $\mathcal{E}_2^0 = 0$. In contrast, circular polarization does not induce a preferred direction. Therefore, in Chapter 4 we show that a cylindrically-symmetric circularly polarized beam would not undergo multiple filamentation (see Corollary 4.11). Moreover, Corollary 4.12 suggests that even cylindrically symmetric, *almost* circularly polarized beams are unlikely to undergo multiple filamentation, because the anisotropic terms are much weaker ($O(\varepsilon f^2)$).

In Subsection 3.4.1 we tested numerically the original Bespalov-Talanov model for multiple filamentation, by solving the *unperturbed* NLS (2.16) with high-power cylindrically-symmetric Gaussian input-beams, to which we added random noise (4.29). We saw neither evidence for multiple filamentation nor even for mild instabilities. Rather, the beams converged to a cylindrically-symmetric profile and collapsed. However, when additional physical mechanisms, such as saturation of the Kerr nonlinearity, are added to the NLS model, then input noise can lead to multiple filamentation of very high-power input beams (see Subsection 3.4.2). For circularly polarized beams we reach similar conclusions: When we solve the Close *et al.* model (4.3) with high-power noisy input beams the beams collapse while converging to a cylindrically-symmetric profile (see Subsection 4.6.2). However, when nonparaxiality and the coupling to the axial component are included, input noise can lead to multiple filamentation of circularly polarized input beams. Thus, noise can lead to multiple filamentation only in the presence of a regularizing mechanism (such as nonlinear saturation, nonparaxiality or coupling to the axial component) that lead to an unstable ring structure.

Table 5.1 summarizes the possibility of multiple filamentation under various input beam characteristics. To recap, ideal cylindrically-symmetric circularly polarized input beams will not undergo multiple filamentation. Small ellipticity of the input polarization is unlikely to lead to multiple filamentation of circularly polarized beams, whereas input beam noise/astigmatism can lead to multiple filamentation. Therefore, suppression of multiple filamentation of circularly polarized beams should

focus on producing a cylindrically-symmetric input beam, rather than on producing perfect circular polarization. In contrast, one cannot suppress multiple filamentation of linearly polarized beams by producing a clean cylindrically-symmetric input beam. Finally, circularly polarized beams are less likely to undergo multiple filamentation than linearly polarized beams.

Input beam characteristics	Linear polarization	Circular polarization
Perfect polarization state and cylindrically-symmetric profile	yes (deterministic)	no
Small deviation from preferential polarization state	yes	unlikely
Small imperfections in the input profile (noise/astigmatism)	yes	yes

Table 5.1: Possibility of multiple filamentation under various input beam characteristics.

Chapter 6

Numerical issues

In Section 6.1 we present an overview of the methods used for the numerical simulations. In Section 6.2 we explain how we approximate the nonparaxial term in the model equations to obtain initial-value problems, which are more amenable for numerical simulations. In Section 6.3 we provide a numerical test to make sure that the splitting the deterministic multiple filamentation simulations is due to vectorial effects, rather than to grid effects. In Section 6.4 we show that when the computational domain is not sufficiently large, reflections from the boundary of the numerical mesh can lead to what may appear as breakup of cylindrical symmetry and even as multiple filamentation.

6.1 Numerical methods - overview

In our simulations of the (2+1)D NLS-type equations (2.16), (3.25), (4.3), and (4.10), we use a finite-difference scheme on a rectangular Cartesian grid with fourth-order accuracy in space. Time-stepping (*i.e.*, z-stepping) is achieved using a fourth-order Runge-Kutta algorithm. This procedure, although not the most efficient from a numerical standpoint, is chosen mainly because of its robustness. We impose zero-Dirichlet boundary conditions at the outer boundaries. Because Dirichlet boundary conditions are reflecting rather than absorbing, special care is taken to assure that reflections from the numerical boundaries have no effect (see also Section 6.4). Furthermore, the results are verified by enlarging the computational domain, as well as by refining the grid and reducing the time-step. In addition, we confirmed the validity of the Runge-Kutta algorithm by using the Crank-Nicholson algorithm.

In most of the simulations the symmetries in the x and y directions enable us to solve the equation on one quadrant of the (x, y) plane. However, in the rotation simulation of Figure 6.1 these symmetries cannot be exploited. Therefore, those simulations are carried on all four quadrants of the (x, y). The same applies to the "noisy" simulations in Section 3.4 and Subsection 4.6.2, because otherwise the noise would be symmetric in the x and y directions, in which case it would be less likely to lead

to a complete break-up of cylindrical symmetry. The noise in the initial conditions [*e.g.*, in (3.14)] is realized using MATLAB's *rand* function, which generates random numbers that are uniformly distributed in the interval $[0, 1]$.

Remark.

Because MATLAB is an interpreted language, *"for"* loops take considerably more run-time compared with complied languages (such as C). Therefore, special care is taken to "vectorize" the MATLAB code. This means that the only *for* loop in the main part of the code is the time-stepping loop. All other algebraic operations in the scheme are accomplished using matrix addition and multiplication, which are internally vectorized (*i.e.*, parallelized) in MATLAB. Furthermore, we use MATLAB's *spdiags* function to generate the sparse differentiation matrices, because MATLAB's operations on such matrices are considerably faster than on full matrices. In particular, the differentiation matrices take the boundary conditions into account.

6.2 "Eliminating" the nonparaxial term

Eqs. (3.7), (3.13), (4.10), and (4.27) contain the nonparaxial terms $A_{1,zz}$ and $A_{+,zz}$. These terms are not related to vectorial effects, as they come from the substitution $\mathcal{E}_1 = \mathcal{A}_1(x, y, z)e^{ik_0 z}$ (for linear polarization) or $\mathcal{E}_+ = \mathcal{A}_+(x, y, z)e^{ik_0 z}$ (for circular polarization) in the *scalar* nonlinear Helmholtz equation (2.13). Because the Helmholtz equation is a boundary value problem, solving it numerically on the half-plane $z \geq 0$ requires setting appropriate radiation boundary conditions at $z \to \infty$. Since, in addition, this equation is nonlinear, solving it as a true boundary-value problem is difficult*. Therefore, the standard approach in numerical simulations is to approximate the nonparaxial term $A_{1,zz}$ with terms that do not have z-derivatives.

In Appendix A.5 we show that the nonparaxial term in Eqs. (3.7) and (3.13) can be approximated with

$$A_{1,zz} = -\Big[\, \Delta_\perp^2 A_1 + 4|A_1|^2 \Delta_\perp A_1 + 4(\nabla_\perp A_1)\cdot(\nabla_\perp A_1^*)A_1 + 2(\nabla_\perp A_1)\cdot(\nabla_\perp A_1)A_1^* + |A_1|^4 A_1 \,\Big] + O(f^2)\,, \tag{6.1}$$

where $\Delta_\perp^2 = (\partial_{xx} + \partial_{yy})^2$ is the biharmonic operator. Substituting the approximation (6.1) in Eq. (3.7)

*For a recent numerical study of the scalar nonlinear Helmholtz equation as a true boundary value problem, see [31, 34].

leads to the following *initial value problem*:

$$\underbrace{iA_{1,z} + \Delta_\perp A_1 + |A_1|^2 A_1}_{\text{NLS}} = \tag{6.2}$$

$$\underbrace{\frac{1}{4} f^2 \left[\Delta_\perp^2 A_1 + 4|A_1|^2 \Delta_\perp A_1 + 4(\nabla_\perp A_1)\cdot(\nabla_\perp A_1^*)A_1 + 2(\nabla_\perp A_1)\cdot(\nabla_\perp A_1)A_1^* + |A_1|^4 A_1 \right]}_{\text{nonparaxial}}$$

$$\underbrace{-f^2 \left[\frac{4+6\gamma}{1+\gamma}|A_{1,x}|^2 A_1 + (A_{1,x})^2 A_1^* + \frac{1+2\gamma}{1+\gamma} \left(|A_1|^2 A_{1,xx} + A_1^2 A_{1,xx}^* \right) \right]}_{\text{vectorial}} + O(f^4) \; .$$

For convenience, we note the origin of the terms in Eq. (6.2).

Clearly, the two scalar equations for A_1 obtained in Section 3.1 and Eq. (6.2) agree with each other to the order of their accuracy (see also Subsubsection 3.1.2):

Corollary 6.1. *Equations (3.7) and (6.2) differ only by $O(f^4)$ terms.*

We can apply the same "treatment" for the circular polarization models. First, it is convenient to rescale A_+ in Eq. (4.27) as in (4.21) in order to eliminate γ from the equation. We then approximate the nonparaxial term with (6.1) [the proof in the same as in Appendix A.5]. Substituting approximation (6.1) after rescaling A_+ in Eq. (4.27), and collecting terms leads to the following *initial value problem*,

$$i\psi_z + \Delta_\perp \psi + |\psi|^2 \psi = \tag{6.3}$$

$$\frac{1}{4} f^2 \left[\Delta_\perp^2 \psi + 2|\psi|^2 \Delta_\perp \psi - 4|\nabla_\perp \psi|^2 \psi - 2\psi^2 \Delta_\perp \psi^* + |\psi|^4 \psi \right],$$

that is used in our numerical simulations. Because the $A_{+,zz}$ term in (4.27) is multiplied by f^2, it follows from (6.1) that Eqs. (6.3) and (4.27) are consistent to $O(f^4)$, *i.e.*, to the order of their derivation. We also use the substitution (6.1) in our simulations of System (4.10) in Section 4.5.

6.3 Physical or grid-induced splitting?

As noted in Section 3.2, vectorial effects induce a preferred direction in the transverse (x, y) plane, which, in our model, is the x-axis direction. Therefore, in the case of cylindrically-symmetric input beams, when vectorial effects lead to breakup of the beam into two filaments, the two filaments can

move away from each other either along the direction of initial polarization or perpendicular to it. Indeed, in our simulations the filaments move away in either of these directions. However, the Cartesian grid that we use in our simulations also has the preferred x and y directions. Therefore, we would like to make sure that the splitting in these simulations is due to vectorial effects, rather than to grid effects.

To do that, we solve Eq. (6.2) using the same parameters and input beam as in Figure 3.8, but in the rotated coordinate system:

$$(\xi, \eta) = (x\cos\theta_0 - y\sin\theta_0, x\sin\theta_0 + y\cos\theta_0) \ ,$$

where θ_0 is the angle of rotation. Since in the rotated (ξ, η) system the preferred direction of vectorial effects forms an angle of θ_0 with $\hat{e}_\xi$, there is now a clear distinction between the preferred direction of vectorial effects and that of grid effects. We note that NLS and nonparaxial terms in Eq. (6.2) remain the same under rotations. The vectorial perturbation terms in Eq. (6.2) do change under rotations, according to:

$$\begin{aligned} A_{1,x} &= A_{1,\xi}\cos\theta_0 + A_{1,\eta}\sin\theta_0 \ , \\ A_{1,xx} &= A_{1,\xi\xi}\cos^2\theta_0 + A_{1,\xi\eta}\sin(2\theta_0) + A_{1,\eta\eta}\sin^2\theta_0 \ . \end{aligned}$$

In the simulation in Figure 6.1 we take $\theta_0 = 30°$. We observe the same dynamics as when $\theta_0 = 0°$, except that the direction of beam-splitting and filament propagation in the (ξ, η) plane follows the preferred direction of the vectorial effects. Thus, this simulation shows that the multiple filamentation observed in our simulations is a feature of the PDE (6.2), rather than a numerical artifact.

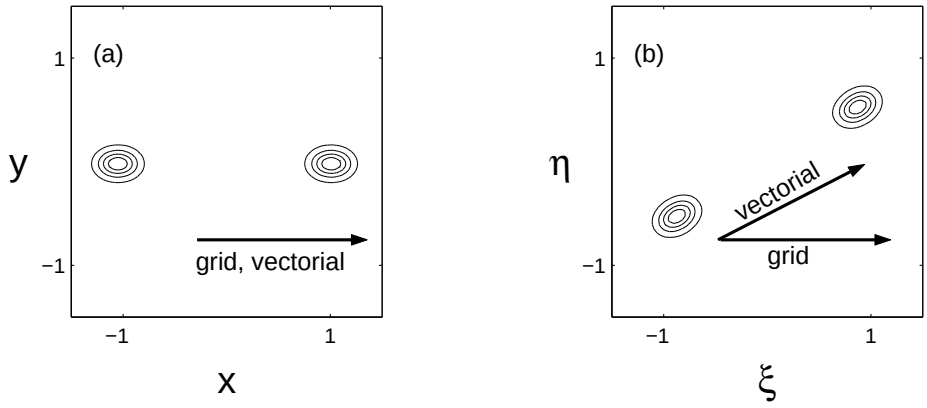

Figure 6.1: Contour plots at $z = 0.32$ of the solution of (6.2) with the same parameters as in Figure 3.21 and $N(0) = 5N_c$. (a) Computed in the (x, y) plane. (b) Computed in the $(\tilde{x}, \tilde{y})$ plane ($\theta_0 = 30°$). Arrows show preferred directions of the numerical grid and of the vectorial perturbation terms.

6.4 Boundary-induced filaments

We now show that when the computational domain is not sufficiently large, reflections from the boundary can lead to what may appear as breakup of cylindrical symmetry and even as multiple filamentation. To see that, we carry three simulations of Eq. (3.25) with the same high-power input beam. The first simulation, which serves as a benchmark, is over a larger domain with Dirichlet boundary conditions. The other two simulations are over a smaller computational domain, on which we impose either Dirichlet or periodic boundary conditions. For consistency, we use *the same noise function realization* in the two simulations over the smaller domain and extend this noise function in the (x, y) plane with zero values for the simulation over the larger domain.

In the simulation over the larger domain we do not observe multiple filamentation for $0 \leq z \leq 0.9$ (Figure 6.2A, top). In addition, at $z = 0.9$ the beam has an almost-symmetric ring structure (Figure 6.2A, bottom). In contrast, we observe filament-type patterns in the iso-surfaces of the two simulations over the smaller domain (top of Figure 6.2B and 6.2C). Moreover, in the simulation over the smaller domain with Dirichlet boundary conditions the ring at $z = 0.9$ is asymmetric (Figure 6.2B, bottom). Thus, the filaments observed in the simulations over the smaller domain, as well as the breakup of symmetry of the ring, are a numerical artifact.

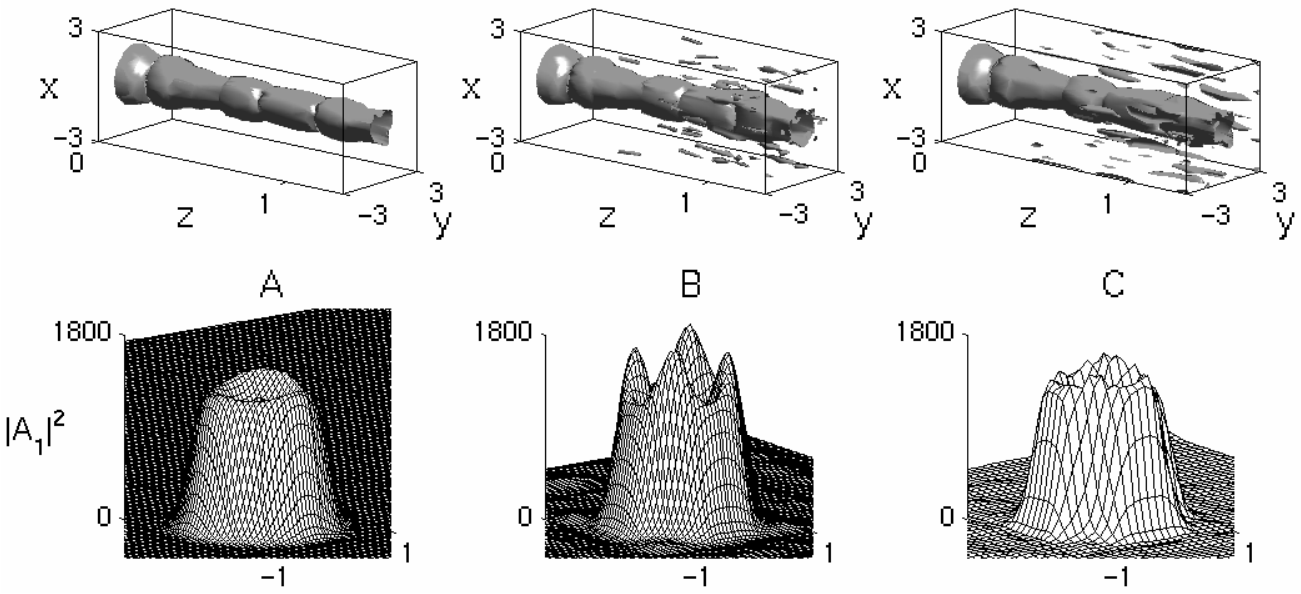

Figure 6.2: Solution of Eq. (3.25) with $\epsilon = 0.01$ and with the input beam $A_1(x, y, z = 0) = 36\, e^{-(x^2+y^2)\ln(2)}[1 + 0.02 * noise(x, y)]$ (*i.e.*, $N(0) \approx 253 N_c$). Computational domain and boundary conditions are (A) $(x, y) \in [-8, 8]^2$, Dirichlet boundary conditions. (B) $(x, y) \in [-3, 3]^2$, Dirichlet boundary conditions. (C) $(x, y) \in [-3, 3]^2$, periodic boundary conditions. Top: Iso-surface $|A_1|^2 = 30$. Bottom: 3D plot at $z = 0.9$.

Chapter 7

Suggestions for further study

During the course of our study several problems remained unsolved and several new questions were raised. Below we mention some of these problems and suggest possible directions for further study.

When the input power $N(0)$ is only moderately above the critical power, the propagation dynamics can be analyzed using modulation theory, and is thus fairly well understood. Unfortunately, there is no such theory for the high-power regime [*i.e.*, $N(0) \gg N_c$] at which multiple filamentation takes place. The reason for this is that modulation theory assumes that the beam is close to a modulated Townes profile, which is cylindrically symmetric. In contrast, multiple filamentation means a complete breakup of cylindrical symmetry. Furthermore, all the multiple filamentation simulations show that multiple filamentation is preceded by ring formation, in contrast to the Townes profile, which is monotonically decreasing. Therefore, a future goal is to offer *a new asymptotic theory for self-focusing that does not assume cylindrical-symmetry, can predict ring formation, and can explain multiple filamentation at the high-power regime*. Because, at present, there is no such theory, we can only rely on numerical simulations to explore vectorial (and other) effects in the high-power regime. Several key questions still await an answer, such as: *How does the threshold power for multiple filamentation depend on the model parameters?* and *How do the number and pattern of the filaments depend on the input power and on the input profile?*

Many laser experiments are conducted with ultrashort laser pulses. The propagation of such pulses is characterized by additional phenomena, such as time-dispersion, that can be neglected for cw beams. A key question is: *What is the effect, if any, of time dispersion and other ultrashort phenomena on multiple filamentation of ultrashort pulses?* Similarly, at the extremely high power regime (*e.g.*, TeraWatt laser beams [65]) plasma formation and other multi-photon processes take place, which raises the question *What is the effect, if any, of plasma formation on multiple filamentation of extremely powerful beams?*

Perhaps the most important open question related to our study is: *What is the mechanism that leads*

to multiple filamentation? We believe that the results of this study support the explanation based on the input linear polarization state, at least when the input power is only a few times N_c. A definite answer, however, would probably emerge from the experimental test suggested in Section 3.6.

Another key question in this study is: *Are circularly polarized beams less likely to undergo multiple filamentation than linearly polarized beams?* Based on our results we predict a positive answer to this question. This prediction can be tested experimentally. For example, let us suppose that the threshold input power for multiple filamentation of a *linearly polarized* beam in a Kerr medium is $N_{\rm th}^{\rm linear}$. One can then repeat the experiment using the same beam, whose polarization state is changed to *circular* before it enters the Kerr medium. If multiple filamentation is suppressed for circular polarization, the corresponding threshold power should be lower. Because the threshold power for self-focusing with circular polarization is $(1+\gamma)$ higher than with linear polarization (4.22), we expect that, at the very least, $N_{\rm th}^{\rm circular} \leq (1+\gamma) N_{\rm th}^{\rm linear}$. As we have predicted, however, aside from the threshold power for self-focusing, multiple filamentation is less likely with circular polarization, because linear polarization induces a preferred direction in the transverse plane, whereas circular polarization does not. Thus, we expect that $N_{\rm th}^{\rm circular}$ is strictly smaller than $(1+\gamma) N_{\rm th}^{\rm linear}$, which would support our prediction.

Finally, the possibility to control beam propagation by changing the polarization state can be useful in many applications. We studied linear and circular polarization states; however, other polarization states are possible, such as azimuthal and radial polarizations (see, *e.g.*, [10]). This raises the question: *How do other polarization states affect self-focusing and multiple filamentation of laser beams?*

Appendix A

Scalar equations for linear polarization

A.1 Derivation of (3.6a) and (3.6b)

The input beam is linearly polarized in the x direction, *i.e.*,

$$A_2 = A_3 = 0 \quad \text{at} \quad z = 0 \,. \tag{A.1}$$

Therefore, to leading order, the beam remains linearly polarized over propagation distances of several diffraction lengths, *i.e.*,

$$A_1 = O(1) \,, \quad A_2 = o(1) \,, \quad A_3 = o(1) \,. \tag{A.2}$$

From Eq. (3.5c) and (A.2), we have that

$$P_1 = O(1) \,, \tag{A.3}$$

$$P_2 = G_1(A_1) \cdot A_2 + G_2(A_1) \cdot A_2^* + o(1) \,, \tag{A.4}$$

$$P_3 = G_3(A_1) \cdot A_3 + G_4(A_1) \cdot A_3^* + o(1) \,, \tag{A.5}$$

where the functions G_i ($i = 1, 2, 3, 4$) are $O(1)$. Using Eqs. (A.2)–(A.5), we can rewrite Eq. (3.5b) as

$$A_3 = if \nabla_\perp \cdot \vec{A} + f^2 \left[-\frac{1}{2} A_{3,z} + iG_3(A_1) \cdot A_3 + iG_4(A_1) \cdot A_3^* \right] + O(f^3) \,.$$

From this equation and Eq. (A.2), we can conclude that

$$A_3 = if(A_{1,x} + A_{2,y}) + O(f^3) = O(f) \,. \tag{A.6}$$

Using Eqs. (3.5c) and (A.6), we can rewrite Eq. (A.4) as

$$P_2 = G_1(A_1) \cdot A_2 + G_2(A_1) \cdot A_2^* + O(f^2) \,. \tag{A.7}$$

The second component of the Helmholtz equation (3.5a) reads

$$iA_{2,z} + \Delta_\perp A_2 + P_2 + \frac{1}{4}f^2 A_{2,zz} = -f\partial_y \left(f\nabla_\perp \cdot \vec{P} + iP_3 + \frac{1}{2}f^2 P_{3,z} \right) . \tag{A.8}$$

In light of Eqs. (A.2)–(A.6), the right-hand-side of Eq. (A.8) is $O(f^2)$. Therefore, using Eq. (A.7) we can rewrite Eq. (A.8) as

$$iA_{2,z} = -\Delta_\perp A_2 - G_1(A_1) \cdot A_2 + G_2(A_1) \cdot A_2^* + O(f^2) . \tag{A.9}$$

From (A.1) and (A.9) we obtain (3.6a). The estimate (3.6b) follows from (A.6) and (3.6a).

A.2 Proof of Proposition 3.1

The first component of the Helmholtz equation (3.5a) is

$$iA_{1,z} + \Delta_\perp A_1 + P_1 + \frac{1}{4}f^2 A_{1,zz} + f\partial_x \left(f\nabla_\perp \cdot \vec{P} + iP_3 + \frac{1}{2}f^2 P_{3,z} \right) = 0 . \tag{A.10}$$

In light of Eqs. (A.5) and (A.6), we have that $P_3 = O(f)$. Therefore, Eq. (A.10) can be rewritten as

$$iA_{1,z} + \Delta_\perp A_1 + P_1 + \frac{1}{4}f^2 A_{1,zz} + f\partial_x \left(f\nabla_\perp \cdot \vec{P} + iP_3 \right) = O(f^4) . \tag{A.11}$$

In order to simplify the terms that depend on $\vec{P}$, we use the estimates (3.6a) and (3.6b) to obtain from Eq. (3.5c) that

$$\begin{aligned} P_1 &= |A_1|^2 A_1 + \frac{f^2}{1+\gamma}\left[|A_{1,x}|^2 A_1 - \gamma (A_{1,x})^2 A_1^* \right] + O(f^4) , \\ P_3 &= \frac{if}{1+\gamma}\left(|A_1|^2 A_{1,x} - \gamma A_1^2 A_{1,x}^* \right) + O(f^3) , \\ \nabla_\perp \cdot \vec{P} &= P_{1,x} + P_{2,y} = 2|A_1|^2 A_{1,x} + A_1^2 A_{1,x}^* + O(f^2) . \end{aligned} \tag{A.12}$$

Substituting the estimates (A.12) in Eq. (A.11), we obtain the scalar equation

$$\begin{aligned} &iA_{1,z} + \Delta_\perp A_1 + |A_1|^2 A_1 + \frac{f^2}{1+\gamma}\left[|A_{1,x}|^2 A_1 - \gamma (A_{1,x})^2 A_1^* \right] + \frac{1}{4}f^2 A_{1,zz} \\ &\quad + f^2 \partial_x \left[\left(2|A_1|^2 A_{1,x} + A_1^2 A_{1,x}^* \right) - \frac{1}{1+\gamma}\left(|A_1|^2 A_{1,x} - \gamma A_1^2 A_{1,x}^* \right) \right] = O(f^4) , \end{aligned}$$

which can be simplified to give Eq. (3.7).

A.3 Derivation of Eq. (3.12)

Lemma A.1. *The components of the Lagrangian $\mathcal{L}_{\rm VH}$ (3.11) can be approximated with*

$$\mathcal{E}_{k,k}\mathcal{E}^*_{j,j} = O(f^6)\,, \tag{A.13}$$

$$\begin{aligned}\mathcal{E}_{j,k}\mathcal{E}^*_{j,k} = \frac{f^2 n_0}{4r_0^2\bar{n}_2}\Bigg\{\Bigg[\frac{1}{f^2}|A_1|^2+\frac{1}{f^2}|A_2|^2+|A_{1,x}|^2\Bigg]+\Bigg[-\frac{i}{2}\Big(A_1^*A_{1,z}-A_1A_{1,z}^*\Big)\\ +|A_{1,x}|^2+|A_{1,y}|^2\Bigg]+f^2\Bigg[\frac{1}{4}|A_{1,z}|^2+|A_{1,xx}|^2+|A_{1,xy}|^2\\ -\frac{i}{2}\Big(A_{1,xz}A_{1,x}^*-A_{1,xz}^*A_{1,x}\Big)\Bigg]\Bigg\}+O(f^6)\,,\end{aligned} \tag{A.14}$$

$$k_0^2\mathcal{E}_k\mathcal{E}_k^* = \frac{f^2 n_0}{4r_0^2\bar{n}_2}\Bigg[\frac{1}{f^2}|A_1|^2+\frac{1}{f^2}|A_2|^2+|A_{1,x}|^2\Bigg]+O(f^6) \tag{A.15}$$

and

$$\begin{aligned}\frac{2k_0^2\bar{n}_2}{n_0(1+\gamma)}\Big(\delta_i^l\delta_j^k+\gamma\delta_i^k\delta_j^l\Big)\mathcal{E}_i\mathcal{E}_j^*\mathcal{E}_k\mathcal{E}_l^* = \frac{f^2n_0}{4r_0^2\bar{n}_2}\frac{1}{2}|A_1|^4\\ +\frac{f^2n_0}{4r_0^2\bar{n}_2}\frac{f^2}{2(1+\gamma)}\Big\{2|A_1|^2|A_{1,x}|^2-\gamma\Big[(A_{1,x})^2A_1^{*2}+(A_{1,x}^*)^2A_1^2\Big]\Big\}+O(f^6)\,.\end{aligned} \tag{A.16}$$

Proof. Using the rescaling (3.4) and substituting the estimates (3.6a) and (3.6b) we have that

$$\mathcal{E}_{k,k}\mathcal{E}^*_{j,j} \overset{(3.4)}{=} \frac{f^2n_0}{4r_0^2\bar{n}_2}\left|A_{1,x}+\frac{i}{f}A_3+O(f^2)\right|^2 \overset{(3.6a),(3.6b)}{=} \frac{f^2n_0}{4r_0^2\bar{n}_2}\cdot O(f^4)=O(f^6)\,, \tag{A.17}$$

which proves (A.13). Similarly,

$$\begin{aligned}\mathcal{E}_{j,k}\mathcal{E}^*_{j,k} &\overset{(3.4)}{=} \frac{f^2n_0}{4\bar{n}_2}\Bigg(\left|\frac{A_{1,x}}{r_0}\right|^2+\left|\frac{A_{1,y}}{r_0}\right|^2+\left|\frac{A_{1,z}}{2k_0r_0^2}+ik_0A_1\right|^2+\left|\frac{A_{2,x}}{r_0}\right|^2+\left|\frac{A_{2,y}}{r_0}\right|^2\\ &\quad+\left|\frac{A_{2,z}}{2k_0r_0^2}+ik_0A_2\right|^2+\left|\frac{A_{3,x}}{r_0}\right|^2+\left|\frac{A_{3,y}}{r_0}\right|^2+\left|\frac{A_{3,z}}{2k_0r_0^2}+ik_0A_3\right|^2\Bigg)\\ &\overset{(3.6a),(3.6b)}{=} \frac{f^2r_0^4n_0}{4\bar{n}_2}\Bigg(|A_{1,x}|^2+|A_{1,y}|^2+\left|\frac{1}{2}fA_{1,z}+\frac{i}{f}A_1\right|^2+\left|\frac{i}{f}A_2\right|^2\\ &\quad+|ifA_{1,xx}|^2+|ifA_{1,xy}|^2+\left|\frac{1}{2}if^2A_{1,xz}+iA_{1,x}\right|^2\Bigg)+O(f^6)\,,\end{aligned}$$

which gives (A.14). Relation (A.15) is straightforward. Finally,

$$\frac{2k_0^2\bar{n}_2}{n_0(1+\gamma)}\left(\delta_i^l\delta_j^k+\gamma\delta_i^k\delta_j^l\right)\mathcal{E}_i\mathcal{E}_j^*\mathcal{E}_k\mathcal{E}_l^* \overset{(3.4)}{=} \frac{2k_0^2\bar{n}_2}{n_0(1+\gamma)}\frac{f^4n_0^2}{16\bar{n}_2^2}\left(A_iA_i^*A_kA_k^*+\gamma A_iA_iA_k^*A_k^*\right)$$

$$=\frac{f^2n_0}{4r_0^2\bar{n}_2}\frac{f^2}{2(1+\gamma)}\left[\left(|A_1|^2+|A_2|^2+|A_3|^2\right)^2+\gamma\left|A_1^2+A_2^2+A_3^2\right|^2\right]+O(f^4)$$

$$\overset{(3.6a),(3.6b)}{=}\frac{f^2n_0}{4r_0^2\bar{n}_2}\frac{f^2}{2(1+\gamma)}\left[\left(|A_1|^2+f^2|A_{1,x}|^2\right)^2+\gamma\left|A_1^2-if^2A_{1,x}^2\right|^2\right]+O(f^4)\,,$$

which gives (A.16). □

Substituting Eqs. (A.13)–(A.16) in (3.11) and dividing by $f^2n_0/4r_0^2\bar{n}_2$ gives (3.12).

A.4 Proof of Proposition 3.2

Eqs. (3.7) and (3.13) have the same $O(1)$ and nonparaxial terms. Therefore, the difference between these equations comes only from the vectorial terms, and is given by

$$\text{Eq. (3.7)} - \text{Eq. (3.13)} =$$

$$f^2\left\{-\frac{4+6\gamma}{1+\gamma}|A_{1,x}|^2A_1-(A_{1,x})^2A_1^*-\frac{1+2\gamma}{1+\gamma}\left(|A_1|^2A_{1,xx}+A_1^2A_{1,xx}^*\right)\right.$$

$$+\frac{1}{1+\gamma}\left[2\gamma|A_{1,x}|^2A_1-(1+\gamma)(A_{1,x})^2A_1^*-|A_1|^2A_{1,xx}+\gamma A_1^2A_{1,xx}^*\right]$$

$$\left.-\partial_{xx}\left(iA_{1,z}+\Delta_\perp A_1\right)\right\}+O(f^4)\,. \tag{A.18}$$

We now prove that this difference is $O(f^4)$, rather than $O(f^2)$. To do that, we differentiate Eq. (3.13) twice with respect to x, to obtain

$$\partial_{xx}\left(iA_{1,z}+\Delta_\perp A_1+|A_1|^2A_1\right)=O(f^2)\,.$$

Therefore,

$$\partial_{xx}\left(iA_{1,z}+\Delta_\perp A_1\right)=-\left[4|A_{1,x}|^2A_1+2(A_{1,x})^2A_1^*+2|A_1|^2A_{1,xx}+A_1^2A_{1,xx}^*\right]+O(f^2)\,.$$

Substituting this equation in Eq. (A.18) gives that

$$\begin{aligned}
&\text{Eq. (3.7)} - \text{Eq. (3.13)} = \\
&f^2 \left\{ -\frac{4+6\gamma}{1+\gamma}|A_{1,x}|^2 A_1 - (A_{1,x})^2 A_1^* - \frac{1+2\gamma}{1+\gamma} \left(|A_1|^2 A_{1,xx} + A_1^2 A_{1,xx}^* \right) \right. \\
&+\frac{1}{1+\gamma} \left[2\gamma|A_{1,x}|^2 A_1 - (1+\gamma)(A_{1,x})^2 A_1^* - |A_1|^2 A_{1,xx} + \gamma A_1^2 A_{1,xx}^* \right] \\
&\left. + \left[4|A_{1,x}|^2 A_1 + 2(A_{1,x})^2 A_1^* + 2|A_1|^2 A_{1,xx} + A_1^2 A_{1,xx}^* \right] \right\} + O(f^4) \; .
\end{aligned}$$

Technical calculations show that the $O(f^2)$ terms on the right-hand-side cancel each other. Therefore, the difference between Eqs. (3.7) and (3.13) is only $O(f^4)$.

A.5 Derivation of Eq. (6.1)

Let us denote

$$W := \Delta_\perp A_1 + |A_1|^2 A_1 \; . \tag{A.19}$$

Using either of the scalar equations (3.13) or (3.7), we have that

$$A_{1,z} = iW + O(f^2) \; . \tag{A.20}$$

Therefore, differentiating (A.20) with respect to z and using (A.19) and (A.20), we get that

$$\begin{aligned}
A_{1,zz} &\overset{(A.20)}{=} iW_z + O(f^2) \overset{(A.19)}{=} i\Delta_\perp A_{1,z} + 2i|A_1|^2 A_{1,z} + iA_1^2 A_{1,z}^* + O(f^2) \\
&\overset{(A.20)}{=} -\Delta_\perp W - 2|A_1|^2 W + A_1^2 W^* + O(f^2) \; .
\end{aligned} \tag{A.21}$$

Substituting (A.19) in the right-hand-side of (A.21), we obtain

$$A_{1,zz} = -\Delta_\perp^2 A_1 - \Delta_\perp(|A_1|^2 A_1) - 2|A_1|^2 \Delta_\perp A_1 + A_1^2 \Delta_\perp A_1^* - |A_1|^4 A_1 + O(f^2) \; . \tag{A.22}$$

Eq. (6.1) follows from Eq. (A.22) and the vectorial identity

$$\Delta_\perp(|A_1|^2 A_1) \equiv 2|A_1|^2 \Delta_\perp A_1 + 4(\nabla_\perp A_1) \cdot (\nabla_\perp A_1^*) A_1 + 2(\nabla_\perp A_1) \cdot (\nabla_\perp A_1) A_1^* + A_1^2 \Delta_\perp A_1^* \; .$$

Appendix B

Modulation theory for vectorial and nonparaxial effects

B.1 Modulation theory

In order to conform to the notations of [33], we denote $\psi = A_1$, $\varepsilon = f^2$, and rewrite Eq. (3.7) as the perturbed NLS

$$i\psi_z + \Delta_\perp \psi + |\psi|^2\psi + \varepsilon F[\psi] = 0 \,, \tag{B.1}$$

where $F[\psi] = P[\psi] + G[\psi] + H[\psi] + K[\psi]$ and

$$\begin{aligned} P[\psi] &= \frac{1}{4}\psi_{zz} \,, & G[\psi] &= \frac{4+6\gamma}{1+\gamma}|\psi_x|^2\psi \,, \\ H[\psi] &= (\psi_x)^2\psi^* \,, & K[\psi] &= \frac{1+2\gamma}{1+\gamma}\left(|\psi|^2\psi_{xx} + \psi^2\psi^*_{xx} \right) \,. \end{aligned} \tag{B.2}$$

Here $P[\psi]$ corresponds to nonparaxiality, and $G[\psi]$, $H[\psi]$, and $K[\psi]$ correspond to vectorial effects.

Modulation theory is based on the following assumptions:

- The focusing part of a filament is close to the asymptotic profile $\psi_R(r,z)$, which is given by

$$\psi_R(r,z) := \frac{1}{L(z)} R(\rho) e^{iS} \,, \tag{B.3}$$

  where $R(\rho)$ is defined in Eq. (2.20),

$$\rho(r,z) := \frac{r}{L(z)} \,, \quad S(r,z) := \zeta(z) + \frac{r^2 L_z(z)}{4L(z)} \,, \quad \zeta_z(z) := \frac{1}{L^2(z)} \,. \tag{B.4}$$

- The filament's power is close to the critical power N_c.
- The perturbation terms are small, *i.e.*, $|\varepsilon F| \ll |\Delta_\perp \psi|$ and $|\varepsilon F| \ll |\psi|^3$.

Under these assumptions, self-focusing dynamics of the perturbed NLS (B.1) is described, to leading order, by

$$L_{zz}(z) = -\frac{\beta}{L^3}\,, \qquad \beta_z(z) = \frac{\varepsilon}{2M}\left(f_{1,z} - 4f_2 \right)\,, \tag{B.5}$$

where N_c and M are defined in Eqs. (2.21) and (3.17), respectively, and the auxiliary functions for $F[\psi]$ are given by

$$\begin{aligned} f_1(z) &= \frac{L}{\pi}\,\mathrm{Re}\int F[\psi_R][R(\rho)+\rho R'(\rho)]e^{-iS}\,dxdy\,, \\ f_2(z) &= \frac{1}{2\pi}\,\mathrm{Im}\int F[\psi_R]\psi_R^*\,dxdy\,. \end{aligned} \tag{B.6}$$

Because Eq. (B.6) is linear in F, the reduced system (B.5) is additive in the perturbation terms. Therefore, in our case we get that

$$f_1 = p_1 + g_1 + h_1 + k_1\,, \qquad f_2 = p_2 + g_2 + h_2 + k_2\,, \tag{B.7}$$

where p_i ($i = 1, 2$) are the auxiliary functions with $P[\psi_R]$ instead of $F[\psi_R]$ in Eqs. (B.6), and similarly for g_i, h_i, and k_i (see Appendix B.4).

Lemma B.1. *The auxiliary functions corresponding to P, G, H, and K in Eq. (B.2) satisfy:*

$$\begin{aligned} p_2 &\sim \frac{N_c}{4}\left(\frac{1}{L^2}\right)_z\,, & p_{1,z} &\ll p_2\,, \\ g_2 &\equiv 0\,, & g_{1,z} &\sim \frac{4+6\gamma}{3(1+\gamma)}N_c\left(\frac{1}{L^2}\right)_z\,, \\ h_2 &= \frac{N_c}{4}\left(\frac{1}{L^2}\right)_z\,, & h_{1,z} &\sim \frac{N_c}{3}\left(\frac{1}{L^2}\right)_z\,, \\ k_2 &\equiv 0\,, & k_{1,z} &\sim \frac{2+4\gamma}{3(1+\gamma)}\left(3N_c - 2I_6\right)\left(\frac{1}{L^2}\right)_z\,, \end{aligned}$$

where $I_6 \approx 6N_c$.

In order to prove Lemma B.1, we first make some preliminary calculations.

B.2 Preliminary calculations

In the following calculations we switch back and forth between Cartesian and cylindrical coordinates, given by

$$(x, y) = (r\cos\theta, r\sin\theta) = \rho L(\cos\theta, \sin\theta)\,.$$

We also denote $(\cdot)' := \frac{d}{d\rho}$.

The small parameter in modulation theory is $\beta(z)$. We can utilize this to simplify the expressions in Lemma B.3, by using the following relations:

Lemma B.2. *Let $|\beta| \ll 1$. Then, the following relations hold:*

$$[L^2 L_z^2] = [x^2 L_z^2] = O(\beta) \ll 1 \,, \tag{B.8}$$

$$R_x^2 - (RS_x)^2 \sim R_x^2 \,, \tag{B.9}$$

where $[\cdot]$ stands for the characteristic size.

Proof. These relations follow from $[x] = [L]$, $[S_x] = [LL_z]$, and $|\beta| = |L^3 L_{zz}| \ll 1$. □

Lemma B.3. *Let ψ_R be given by (B.3) and (B.4). Then:*

$$|\psi_{R,x}|^2 \;\sim\; L^{-4} R'^2 \cos^2\theta \,, \tag{B.10}$$

$$\mathrm{Im}\left[\,(\psi_{R,x})^2 e^{-2iS}\,\right] \;=\; L^{-2}(R^2)_x S_x \,, \tag{B.11}$$

$$\mathrm{Re}\left[\,(\psi_{R,x})^2 e^{-2iS}\,\right] \;\sim\; L^{-4} R'^2 \cos^2\theta \,, \tag{B.12}$$

$$\mathrm{Re}(\psi_R^* \psi_{R,xx}) \;\sim\; L^{-4}\left(\, RR''\cos^2\theta + \frac{1}{\rho} RR' \sin^2\theta \,\right) \,. \tag{B.13}$$

Proof. From (B.3) and (B.4) we obtain that

$$S_x = \frac{xL_z}{2L} \,, \quad R_x = \frac{R'\cos\theta}{L} \,, \quad R_{xx} = \frac{1}{L^2}\left(\, R''\cos^2\theta + \frac{1}{\rho} R' \sin^2\theta \,\right) \,. \tag{B.14}$$

Using Eqs. (B.3) and (B.4) we get that

$$\psi_{R,x} = L^{-1}(R_x + iRS_x) e^{iS} \,, \tag{B.15}$$

from which we obtain

$$|\psi_{R,x}|^2 \overset{(B.15)}{=} L^{-2}(R_x^2 + R^2 S_x^2) \overset{(B.14)}{=} L^{-4}\cos^2\theta \left(\, R'^2 + \frac{1}{4} R^2 \rho^2 L^2 L_z^2 \,\right) \,. \tag{B.16}$$

This expression can be further simplified using (B.8), because

$$R'^2 + \frac{1}{4}\rho^2 R^2 L^2 L_z^2 = R'^2 + O(L^2 L_z^2) \sim R'^2 \,.$$

Substituting this approximation in (B.16) leads to (B.10). In addition, from (B.15) we have that

$$(\psi_{R,x})^2 = L^{-2} e^{2iS}\left[\, R_x^2 - (RS_x)^2 + i(R^2)_x S_x \,\right] \,, \tag{B.17}$$

from which (B.11) follows. From (B.17) we also have that

$$\mathrm{Re}\left[\ (\psi_{R,x})^2 e^{-2iS}\ \right] = L^{-2}\left[\ R_x^2 - (RS_x)^2\ \right] \overset{(B.9)}{\sim} L^{-2}R_x^2 \overset{(B.14)}{=} L^{-4}R'^2\cos^2\theta\ ,$$

which proves (B.12). Finally, we have that

$$\psi_{R,xx} \overset{(B.15)}{=} L^{-1}e^{iS}\left[\ (R_{xx} - RS_x^2) + i(2R_xS_x + RS_{xx})\ \right]\ .$$

From this relation we get that

$$\begin{aligned}\mathrm{Re}(\psi_R^*\psi_{R,xx}) &\overset{(B.3),(B.4)}{=} \frac{R}{L^2}\left(\ R_{xx} - RS_x^2\ \right) \overset{(B.14)}{=} \frac{R}{L^4}\left[\left(\ R''\cos^2\theta + \frac{1}{\rho}R'\sin^2\theta\ \right) - \frac{1}{4}x^2RL_z^2\ \right] \\ &\overset{(B.8)}{\sim} \frac{R}{L^4}\left(\ R''\cos^2\theta + \frac{1}{\rho}R'\sin^2\theta\ \right)\ ,\end{aligned}$$

which proves (B.13).

B.3 Integral relations

Here we obtain relations (B.18), (B.20), (B.21), and (B.26), in order to reduce the number of constants that appear in the reduced system (3.16).

The following relation is well-known:

$$\int R^4\rho\,d\rho = 2N_c\ . \tag{B.18}$$

We now prove some additional relations.

Lemma B.4. *Let $R(\rho)$ be defined by Eq. (2.20) and let*

$$\begin{aligned} I_1 &:= 3\int R^2R'^2\rho\,d\rho\ , & I_2 &:= 3\int \rho RR'^3\rho\,d\rho\ , \\ I_3 &:= 3\int \rho R^2R'R''\rho\,d\rho\ , & I_4 &:= 3\int R^3R''\rho\,d\rho\ , & I_6 &:= \int R^6\rho\,d\rho\ . \end{aligned} \tag{B.19}$$

Then the following relations hold:

$$I_1 + I_2 = N_c\ , \tag{B.20}$$

$$I_1 + I_3 + I_4 - \frac{3}{4}R^4(0) = 3N_c - 2I_6\ . \tag{B.21}$$

Proof. The proof will follow from the four linear equations (B.22)–(B.25) with I_1, I_2, I_3, I_4, I_6, $R(0)$, and N_c, which we now derive. Multiplying Eq. (2.20) by $(3\rho^2 R^2 R'\, d\rho)$ and integrating, we get that

$$\int \left(3\rho^2 R^2 R' R'' + 3\rho R^2 R'^2 + 3\rho^2 R^5 R' - 3\rho^2 R^3 R' \right) \, d\rho = 0 \, .$$

Using (B.19) we arrive at

$$I_3 + I_1 + \int \left[\frac{1}{2}\rho^2 (R^6)' - \frac{3}{4}\rho^2 (R^4)' \right] \, d\rho \stackrel{\text{IBP}}{=} I_3 + I_1 - \int \left(\rho R^6 - \frac{3}{2}\rho R^4 \right) \, d\rho = 0 \, ,$$

where "IBP" stands for integration-by-parts. Using (B.18), we obtain that

$$I_1 + I_3 = I_6 - 3N_c \, . \tag{B.22}$$

Multiplying (2.20) by $(R^3 \rho \, d\rho)$ and integrating we obtain

$$\int \left(\rho R^3 R'' + R^3 R' + \rho R^6 - \rho R^4 \right) \, d\rho = 0 \, .$$

Using (B.18) we obtain the second relation

$$\frac{1}{3} I_4 + \int_0^\infty \frac{1}{4} (R^4)' \, d\rho + I_6 - 2N_c = \frac{1}{3} I_4 - \frac{1}{4} R^4(0) + I_6 - 2N_c = 0 \, . \tag{B.23}$$

From the definition of I_1 we have that

$$\begin{aligned} I_1 \;&:=\; 3\int R^2 R' R' \rho \, d\rho \stackrel{\text{IBP}}{=} -3\int (\rho R^2 R')' R \, d\rho \stackrel{(B.19)}{=} -\frac{3}{4}\int_0^\infty (R^4)' \, d\rho - 2I_1 - I_4 \\ &=\; \frac{3}{4} R^4(0) - 2I_1 - I_4 \, , \end{aligned}$$

which leads to the third relation

$$3I_1 + I_4 = \frac{3}{4} R^4(0) \, . \tag{B.24}$$

Integration of the differential identity

$$\frac{3}{2}\frac{d}{d\rho} \left(\rho R R' \right)^2 \;\equiv\; 3\rho R^2 R'^2 + 3\rho^2 R R'^3 + 3\rho^2 R^2 R' R'' \, ,$$

leads to the fourth relation

$$I_1 + I_2 + I_3 = 0 \, . \tag{B.25}$$

Solving the linear system (B.22)–(B.25) we obtain

$$\begin{aligned} I_1 &= I_6 - 2N_c\,, \qquad & I_2 &= 3N_c - I_6\,, \\ I_3 &= -N_c\,, & I_4 &= 6N_c - 3I_6 + \frac{3}{4}R^4(0)\,, \end{aligned}$$

from which (B.20) and (B.21) follow. □

Finally, our numerical calculations (see also [33]) show that $I_6 \approx 6.07N_c$. Therefore, we can use the approximation

$$I_6 \approx 6N_c\,, \tag{B.26}$$

which has about 1% accuracy.

B.4 Proof of Lemma B.1

The first perturbation term in Eq. (B.2) is the nonparaxial term $P[\psi]$. In [24] it was shown that

$$p_{1,z} \ll p_2 := \frac{1}{2\pi}\,\mathrm{Im}\int \frac{1}{4}\psi_{R,zz}\psi_R^*\,dxdy \sim \frac{N_c}{4}\left(\frac{1}{L^2}\right)_z\,. \tag{B.27}$$

Therefore, to leading order, p_1 can be neglected in the reduced system.

According to Eq. (B.6), the second auxiliary function corresponding to $G[\psi]$ is

$$g_2 := \frac{1}{2\pi}\frac{4+6\gamma}{1+\gamma}\int \mathrm{Im}(|\psi_{R,x}|^2|\psi_R^*|^2)\,dxdy = 0\,.$$

Calculating g_1 gives that

$$\begin{aligned} g_1 \;:=\;& \frac{4+6\gamma}{1+\gamma}\frac{L}{\pi}\,\mathrm{Re}\int |\psi_{R,x}|^2\psi_R(R+\rho R')e^{-iS}\,dxdy \\ \overset{(B.3)}{=}\;& \frac{4+6\gamma}{1+\gamma}\pi\int |\psi_{R,x}|^2 R(R+\rho R')\,dxdy \\ \overset{(B.10)}{\sim}\;& \frac{4+6\gamma}{1+\gamma}\frac{1}{\pi L^4}\int R'^2\cos^2\theta(R^2+\rho RR')\,dxdy \\ =\;& \frac{4+6\gamma}{1+\gamma}\frac{1}{\pi L^2}\int_0^{2\pi}\cos^2\theta\,d\theta\int (R^2R'^2+\rho RR'^3)\rho\,d\rho \\ \overset{(B.19)}{=}\;& \frac{4+6\gamma}{1+\gamma}\frac{1}{3L^2}\left(I_1+I_2\right)\overset{(B.20)}{=}\frac{4+6\gamma}{3(1+\gamma)}\frac{N_c}{L^2}\,. \end{aligned} \tag{B.28}$$

For $H[\psi]$ we have that

$$\begin{aligned} h_1 := \; & \frac{L}{\pi}\,\mathrm{Re}\int (\psi_{R,x})^2\psi_R^*(R+\rho R')e^{-iS}\,dxdy \overset{(B.3)}{=} \frac{1}{\pi}\int \mathrm{Re}[(\psi_{R,x})^2]R(R+\rho R')\,dxdy \\ \overset{(B.12)}{=} \; & \frac{1}{\pi L^2}\int RR'^2(R+\rho R')\cos^2\theta\,\rho\,d\rho d\theta \overset{(B.19)}{\sim} \frac{1}{3L^2}\left(\, I_1+I_2\,\right) \overset{(B.20)}{=} \frac{N_c}{3L^2}\,, \end{aligned} \tag{B.29}$$

and

$$\begin{aligned} h_2 \; := \; & \frac{1}{2\pi}\,\mathrm{Im}\int (\psi_{R,x})^2(\psi_R^*)^2\,dxdy \overset{(B.3)}{=} \frac{1}{2\pi L^2}\int R^2\,\mathrm{Im}\left[\,(\psi_{R,x})^2e^{-2iS}\,\right]\,dxdy \\ \overset{(B.11)}{=} \; & \frac{1}{2\pi L^4}\int R^2(R^2)_xS_x\,dxdy = \frac{1}{4\pi L^4}\int (R^4)_xS_x\,dxdy \\ \overset{\text{IBP}}{=} \; & -\frac{1}{4\pi L^4}\int R^4S_{xx}\,dxdy \overset{(B.14)}{=} -\frac{L_z}{8\pi L^5}\int R^4\,dxdy \\ = \; & -\frac{L_z}{8L^3}\int R^4\rho\,d\rho \overset{(B.18)}{=} \frac{N_c}{4}\left(\,\frac{1}{L^2}\,\right)_z\,. \end{aligned} \tag{B.30}$$

For $K[\psi]$ we have that

$$K[\psi_R]\psi_R^* = \frac{1+2\gamma}{1+\gamma}\left(\,|\psi_R|^2\psi_{R,xx}+\psi_R^2\psi_{R,xx}^*\,\right)\psi_R^* \overset{(B.3)}{=} \frac{2+4\gamma}{1+\gamma}\frac{R^2}{L^2}\,\mathrm{Re}\left(\,\psi_R^*\psi_{R,xx}\,\right)\,. \tag{B.31}$$

In light of (B.31), the integrand in k_2 (B.6) is real valued. Therefore, $k_2=0$. Calculating k_1 gives that

$$\begin{aligned} k_1 \; := \; & \frac{L}{\pi}\,\mathrm{Re}\int K[\psi_R]\left(\,R+\rho R'\,\right)e^{-iS}\,dxdy \\ \overset{(B.3)}{=} \; & \frac{L}{\pi}\int \mathrm{Re}\left(\,K[\psi_R]\psi_R^*\frac{L}{R}\,\right)\left(\,R+\rho R'\,\right)\,dxdy \\ \overset{(B.31)}{=} \; & \frac{1+2\gamma}{1+\gamma}\frac{2L^2}{\pi}\int \mathrm{Re}\left(\,\psi_R^*\psi_{R,xx}\,\right)\left(\,R^2+\rho RR'\,\right)\rho\,d\rho d\theta \\ \overset{(B.13)}{\sim} \; & \frac{1+2\gamma}{1+\gamma}\frac{2}{\pi L^2}\int \left(\,RR''\cos^2\theta+\frac{1}{\rho}RR'\sin^2\theta\,\right)\left(\,R^2+\rho RR'\,\right)\rho\,d\rho d\theta \\ \overset{(B.19)}{=} \; & \frac{1+2\gamma}{1+\gamma}\frac{2}{3L^2}\left[\,I_4+I_3-\frac{3}{4}R^4(0)+I_1\,\right] \overset{(B.21)}{=} \frac{2+4\gamma}{3(1+\gamma)}\left(\,3N_c-2I_6\,\right)\frac{1}{L^2}\,. \end{aligned} \tag{B.32}$$

Combining (B.27)–(B.30) and (B.32) proves Lemma B.1.

B.5 Proof of Proposition 3.3

To obtain Eqs. (3.16), we substitute the auxiliary functions of Lemma B.1 in Eq. (B.5), use Eq. (B.7), and use $\varepsilon = f^2$. Doing that, to leading order we obtain

$$\beta_z = \frac{f^2}{2M}\left(-4p_2 + g_{1,z} + h_{1,z} - 4h_2 + k_{1,z}\right) = \underbrace{-\frac{f^2 N_c}{2M}\left(\frac{1}{L^2}\right)_z}_{\text{nonparaxial}}$$

$$\underbrace{+\frac{f^2}{2M}\left[\frac{4+6\gamma}{3(1+\gamma)}N_c + \frac{1}{3}N_c - N_c + \frac{2+4\gamma}{3(1+\gamma)}\left(3N_c - 2I_6\right)\right]\left(\frac{1}{L^2}\right)_z}_{\text{vectorial}},$$

from which we obtain the second equation in (3.16), with $C_{\text{nonparax}} = 1$ and

$$C_{\text{vec}}(\gamma) = \frac{4}{3}\cdot\frac{(I_6/N_c - 2) + (2I_6/N_c - 4)\gamma}{1+\gamma} \overset{(B.26)}{\approx} \frac{16}{3}\left(1 + \frac{\gamma}{1+\gamma}\right) .$$

Appendix C

Proof of Proposition 4.4

The starting point for the derivation of (4.10) is Eqs. (3.5), which were also used in the derivation of Eq. (3.7). Below we omit the technical details that were already obtained in Appendix A.1 and Appendix A.2.

C.1 Derivation of estimates (4.9)

We assume that the input beam is almost left circularly-polarized (4.6), *i.e.*, that $A_-^0 \ll A_+^0$ and $A_3^0 \ll A_+^0$. Therefore, it follows from Eqs. (3.5) that over propagation distances of several diffraction lengths

$$A_+(x,y,z) = O(1)\ , \quad A_-(x,y,z) = o(1)\ , \quad A_3(x,y,z) = o(1)\ . \tag{C.1}$$

In analogy with (4.3) we define the nondimensional circular nonlinear-polarizations

$$P_\pm := \frac{1}{\sqrt{2}} \left(P_1 \pm iP_2 \right)\ , \tag{C.2}$$

where $\vec{P} = (P_1, P_2, P_3)$ is defined in (3.5c).

Similarly to the derivation in [29] it follows from Eqs. (3.5b), (3.5c), and (C.1) that

$$A_3 = O(f)\ . \tag{C.3}$$

Using (C.3) we get from Eq. (3.5c) that

$$P_3 = O(f)\ . \tag{C.4}$$

Therefore, substituting (C.4) in Eq. (3.5a) gives that

$$iA_{1,z} + \Delta_\perp A_1 + \frac{1}{4} f^2 A_{1,zz} + P_1 = -f\partial_x \left(f\nabla_\perp \cdot \vec{P} + iP_3 \right) + O(f^4)\ , \tag{C.5a}$$

and

$$iA_{2,z} + \Delta_\perp A_2 + \frac{1}{4} f^2 A_{2,zz} + P_2 = -f\partial_y \left(f\nabla_\perp \cdot \vec{P} + iP_3 \right) + O(f^4) \,. \tag{C.5b}$$

Subtracting Eq. (C.5b) from i times Eq. (C.5a), dividing by $\sqrt{2}$, and using (C.2) gives that

$$iA_{-,z} + \Delta_\perp A_- + \frac{1}{4} f^2 A_{-,zz} + P_- = O(f^2) \,. \tag{C.6}$$

Throughout the derivation in Appendix C we use the following identities, whose proof is straightforward:

$$\begin{aligned} \vec{A} \cdot \vec{A}^* &= |A_1|^2 + |A_2|^2 + |A_3|^2 \overset{(4.8)}{=} |A_+|^2 + |A_-|^2 + |A_3|^2 \,, \\ \vec{A} \cdot \vec{A} &= A_1^2 + A_2^2 + A_3^2 \overset{(4.8)}{=} 2A_+A_- + A_3^2 \,, \end{aligned} \tag{C.7}$$

and

$$A_1 \overset{(4.8)}{=} \frac{1}{\sqrt{2}} \left(A_+ + A_- \right) \,, \qquad A_2 \overset{(4.8)}{=} -\frac{i}{\sqrt{2}} \left(A_+ - A_- \right) \,. \tag{C.8}$$

It follows from Eqs. (4.8), (3.5c), (C.2), and (C.7) that

$$\begin{aligned} P_- &= \frac{1}{1+\gamma} \left[(\vec{A} \cdot \vec{A}^*) A_- + \gamma (\vec{A} \cdot \vec{A}) A_+^* \right] \\ &\overset{(C.7)}{=} \frac{1}{1+\gamma} \left[(|A_+|^2 + |A_-|^2 + |A_3|^2) A_- + \gamma (2A_+A_- + A_3^2) A_+^* \right] \,. \end{aligned} \tag{C.9}$$

Therefore, we get from (C.3) and (C.9) that

$$P_- = \frac{1}{1+\gamma} \left[(1+2\gamma)|A_+|^2 + |A_-|^2 \right] A_- + O(f^2) \,. \tag{C.10}$$

We can rewrite Eq. (C.6) using estimate (C.10) as

$$\left(i\partial_z + \Delta_\perp + \frac{1+2\gamma}{1+\gamma} |A_+|^2 + \frac{1}{1+\gamma} |A_-|^2 \right) A_- = O(f^2) \,. \tag{C.11}$$

We note that Eq. (C.11) is a homogeneous equation in A_- with an $O(1)$ operator on the left hand side. Because $A_-^0 / A_+^0 = O(\varepsilon)$ [Eqs. (C.1)] and the driving terms on the right hand-side of Eq. (C.11) are $O(f^2)$, estimate (4.9a) follows.

Using Eq. (3.5b) we get from (C.3) and (C.4) that

$$A_3 = if(A_{1,x} + A_{2,y}) + O(f^3) \,. \tag{C.12}$$

Using (4.9a) we obtain from identities (C.8) that

$$A_1 = \frac{1}{\sqrt{2}} A_+ + O(f^2, \varepsilon), \qquad A_2 = -\frac{i}{\sqrt{2}} A_+ + O(f^2, \varepsilon) \,. \tag{C.13}$$

Substituting (C.13) into (C.12) yields estimate (4.9b).

C.2 Derivation of System (4.10)

In order to obtain Eq. (4.10b) we first use (4.9a) to get that

$$|A_-|^2 A_- = O(f^6, \varepsilon^3)\ . \tag{C.14}$$

Eq. (4.10b) follows from substituting estimate (C.14) into Eq. (C.11).

Below we derive Eq. (4.10a). Summing Eq. (C.5b) with i times Eq. (C.5a) and dividing by $\sqrt{2}$ leads to

$$iA_{+,z} + \Delta_\perp A_+ + \frac{1}{4} f^2 A_{+,zz} + P_+ = -\frac{f}{\sqrt{2}}(\partial_x + i\partial_y)(f\nabla_\perp \cdot \vec{P} + iP_3) + O(f^4)\ . \tag{C.15}$$

Using identities (C.7) and estimates (4.9) it follows that

$$\begin{aligned} P_+ &\overset{(3.5c)}{=} \frac{1}{1+\gamma}\left[\ (\vec{A}\cdot\vec{A}^*)A_+ + \gamma(\vec{A}\cdot\vec{A})A_-^*\ \right] \\ &\overset{(C.7)}{=} \frac{1}{1+\gamma}\left[\ (|A_+|^2 + |A_-|^2 + |A_3|^2)A_+ + \gamma(2A_+A_- + A_3^2)A_-^*\ \right] \\ &\overset{(4.9)}{=} \frac{1}{1+\gamma}|A_+|^2A_+ + \frac{1+2\gamma}{1+\gamma}|A_-|^2A_+ + \frac{f^2}{2(1+\gamma)}|(i\partial_x + \partial_y)A_+|^2A_+ + O(f^4, \varepsilon f^2)\ . \end{aligned} \tag{C.16}$$

and that

$$\begin{aligned} \vec{A}\cdot\vec{A}^* &\overset{(C.7)}{=} |A_+|^2 + |A_-|^2 + |A_3|^2 \overset{(4.9)}{=} |A_+|^2 + O(f^2, \varepsilon^2)\ , \\ \vec{A}\cdot\vec{A} &\overset{(C.7)}{=} 2A_+A_- + A_3^2 \overset{(4.9)}{=} O(f^2, \varepsilon)\ . \end{aligned} \tag{C.17}$$

Similarly, we obtain that

$$\begin{aligned} P_3 &\overset{(3.5c)}{=} \frac{1}{1+\gamma}\left[\ (\vec{A}\cdot\vec{A}^*)A_3 + \gamma(\vec{A}\cdot\vec{A})A_3^*\ \right] \\ &\overset{(C.17)}{=} \frac{1}{1+\gamma}|A_+|^2A_3 + O(f^3, \varepsilon f) \overset{(3.6b)}{=} \frac{if}{\sqrt{2}(1+\gamma)}|A_+|^2(\partial_x - i\partial_y)A_+\ , + O(f^3, \varepsilon f) \end{aligned} \tag{C.18}$$

and

$$\begin{aligned}
\nabla_\perp \cdot \vec{P} &\equiv P_{1,x} + P_{2,y} \\
&\overset{(3.5c)}{=} \frac{1}{1+\gamma}\partial_x \left[(\vec{A}\cdot\vec{A}^*)A_1 + \gamma(\vec{A}\cdot\vec{A})A_1^* \right] + \frac{1}{1+\gamma}\partial_y \left[(\vec{A}\cdot\vec{A}^*)A_2 + \gamma(\vec{A}\cdot\vec{A})A_2^* \right] \\
&\overset{(C.17)}{=} \frac{1}{1+\gamma}\partial_x(|A_+|^2A_1) + \frac{1}{1+\gamma}\partial_y(|A_+|^2A_2) + O(f^2,\varepsilon) \\
&\overset{(C.13)}{=} \frac{1}{\sqrt{2}(1+\gamma)}\partial_x(|A_+|^2A_+) - \frac{i}{\sqrt{2}(1+\gamma)}\partial_y(|A_+|^2A_+) + O(f^2,\varepsilon) \\
&= \frac{1}{\sqrt{2}(1+\gamma)}(\partial_x - i\partial_y)(|A_+|^2A_+) + O(f^2,\varepsilon)\,. \qquad \text{(C.19)}
\end{aligned}$$

Substituting Eqs. (C.16), (C.18) and (C.19) into Eq. (C.15) leads to

$$\begin{aligned}
&iA_{+,z} + \Delta_\perp A_+ + \frac{1}{1+\gamma}|A_+|^2A_+ + \frac{1+2\gamma}{1+\gamma}|A_-|^2A_+ = \\
&\quad -\frac{1}{4}f^2A_{+,zz} - \frac{f^2}{2(1+\gamma)}|(i\partial_x+\partial_y)A_+|^2A_+ \\
&\quad -\frac{f^2}{2(1+\gamma)}\left(\partial_x + i\partial_y\right)\left[(\partial_x - i\partial_y)(|A_+|^2A_+) - |A_+|^2(\partial_x - i\partial_y)A_+ \right] + O(f^4, \varepsilon f^2)\,.
\end{aligned}$$

Rearranging the right hand side of this equation gives Eq. (4.10a).

Appendix D

Proof of Proposition 4.8

We briefly outline the derivation of the reduced system (4.24). In order to apply modulation theory to (4.10) it is convenient to use the rescaling (4.21), *i.e.*, $\psi(x,y,z) = (1+\gamma)^{-1/2}A_+(x,y,z)$, and rewrite Eq. (4.10a) as the following perturbed NLS:

$$i\psi_z + \Delta_\perp\psi + |\psi|^2\psi + C_e|a_-|^2\psi + f^2F[\psi] = 0, \tag{D.1}$$

where $C_e = O(f^2,\varepsilon)$ is a small positive constant, $a_-(x,y,z) = A_-/C_e = O(1)$, $F[\psi] = P[\psi] + G[\psi] + H[\psi] + K[\psi] + K^{'}[\psi]$, and

$$\begin{aligned}
P[\psi] &= \frac{1}{4}\psi_{zz}\,, & G[\psi] &= 2\left(\,|\psi_x|^2 + |\psi_y|^2\,\right)\psi\,,\\
H[\psi] &= \frac{1}{2}\left(\,\psi_x^2 + \psi_y^2\,\right)\psi^*\,, & K[\psi] &= \frac{1}{2}\left(\,|\psi|^2\psi_{xx} + \psi^2\psi^*_{xx}\,\right)\,,\\
K^{'}[\psi] &= \frac{1}{2}\left(\,|\psi|^2\psi_{yy} + \psi^2\psi^*_{yy}\,\right)\,.
\end{aligned} \tag{D.2}$$

Here $P[\psi]$ corresponds to nonparaxiality and the other functionals correspond to circular polarization effects.

With minor adjustments, the calculations of the auxiliary functions corresponding to $F[\psi]$ in Eq. (D.1) are carried in Appendix B. The results are as follows.

Lemma D.1. *The auxiliary functions corresponding to P, G, H, K, and K' in (D.2) satisfy: $p_{1,z} \ll p_2$, $g_2 \equiv k_2 \equiv k_2' \equiv 0$, and*

$$\begin{aligned}
p_2 &\sim \frac{N_c}{4}\left(\frac{1}{L^2}\right)_z\,, & g_{1,z} &\sim \frac{4}{3}N_c\left(\frac{1}{L^2}\right)_z\,,\\
h_{1,z} &\sim \frac{1}{3}N_c\left(\frac{1}{L^2}\right)_z\,, & h_2 &= \frac{N_c}{4}\left(\frac{1}{L^2}\right)_z\,,\\
k'_{1,z} = k_{1,z} &\sim \left(N_c - \frac{2}{3}I_6\right)\left(\frac{1}{L^2}\right)_z\,,
\end{aligned}$$

where $I_6 = \int R^6 \rho \, d\rho$, p_i *(*$i = 1, 2$*) are the auxiliary functions with* $P[\psi_R]$ *instead of* $F[\psi_R]$ *in Eqs. (B.6), and similarly for* g_i, h_i, k_i, *and* $k_i^{'}$.

For the fourth term in Eq. (D.1), which corresponds to the mismatch from perfect circular input polarization, it straightforward to show that $f_2 = 0$. In addition, for this term we get that

$$f_1 = \frac{C_e L}{\pi} \mathrm{Re} \int |a_-(x, y, z)|^2 L^{-1} R(R + \rho R') \, dxdy = C(z) \, , \tag{D.3}$$

where $C(z) = O(\varepsilon, f^2)$.

We note that the first equation in (4.24) is the first one in (B.5). To obtain the second equation, we substitute the auxiliary functions of Lemma D.1 and (D.3) in the second equation in (B.5) and use the additivity of the reduced system (B.5) to perturbation terms. To leading order we obtain that

$$\begin{aligned} \beta_z &= \frac{f^2}{2M} \left(-4p_2 + g_{1,z} + h_{1,z} - 4h_2 + k_{1,z} + k_{1,z}^{'} \right) + \frac{C'(z)}{2M} \\ &= \underbrace{-\frac{f^2 N_c}{2M} \left(\frac{1}{L^2} \right)_z}_{\text{nonparax}} \underbrace{- \frac{f^2}{2M} \frac{4I_6 - 8N_c}{3} \left(\frac{1}{L^2} \right)_z}_{\text{vec}} + \underbrace{\frac{C'(z)}{2M}}_{E_-} \, . \end{aligned} \tag{D.4}$$

Finally, the second equation in (B.5) is obtained from (D.4) by defining

$$C_{\text{nonparax}} = 1 \, , \qquad C_{E_-}(z) = \frac{C'(z)}{2M} = O(f^2, \varepsilon) \, ,$$

and

$$C_{\text{vec}} = \frac{4I_6}{3N_c} - \frac{8}{3} \approx \frac{16}{3} \, ,$$

where we have used the approximation $I_6 \approx 6N_c$ that has about 1% accuracy (see [33]).

Bibliography

[1] S.C. Abbi and H. Mahr. Correlation of filaments in Nitrobenzene with laser spikes. *Phys. Rev. Lett.*, 26(11):604–606, 1971.

[2] J. Atai, Y. Chen, and J.M. Soto-Crespo. Stability of three-dimensional self-trapped beams with a dark spot surrounded by bright rings of varying intensity. *Phys. Rev. A*, 49:3170–3173, 1994.

[3] T. Bergqvist and B. Kleman sand P. Wahren. Breakdown phenomena in CS_2 caused by Nd-laser radiation. *Ark. Fys.*, 34:81–95, 1967.

[4] A.L. Berkhoer and V.E. Zakharov. Excitation of waves with different polarizations in nonlinear media. *Sov. Phys. JETP*, 31:486–490, 1970. Zh. Eksp. Teor. Fiz. - Pis'ma Red. (USSR) 58:903–911 (1970).

[5] V.I. Bespalov and V.I. Talanov. Filamentary structure of light beams in nonlinear media. *JETP Lett.*, 3:307–310, 1966. Zh. Eksp. Teor. Fiz. - Pis'ma Red. (USSR) 3:471–476 (1966).

[6] J.E. Bjorkholm and A. Ashkin. cw self-focusing and self-trapping of light in Sodium vapor. *Phys. Rev. Lett.*, 32:129–132, 1974.

[7] S. Blair and K. Wagner. (2+1)-D propagation of spatio-temporal solitary waves including higher-order corrections. *Opt. Quant. Elec.*, 30:697–737, 1998.

[8] J.L. Blows, J.M. Dawes, and P.A. Piper. A simple, thermally-stabilized, diode end-pumped, planar Nd : YAG laser. *Opt. Commun.*, 162:247–250, 1999.

[9] A.D. Boardman, K. Marinov, D.I. Pushkarov, and A. Shivarova. Influence of nonlinearly induced diffraction on spatial solitary waves. *Opt. Quant. Elec.*, 32:49–62, 2000.

[10] Ze'ev Bomzon, Vladimir Kleiner, and Erez Hasman. Formation of radially and azimuthally polarized light using space-variant subwavelength metal stripe gratings. *Appl. Phys. Lett.*, 79:1587–1589, 2000.

[11] R.W. Boyd. *Nonlinear Optics*. Academic Press, Boston, 1992.

[12] A. Braun, G. Korn, X. Liu, D. Du, J. Squier, and G. Mourou. Self-channeling of high-peak-power femtosecond laser pulses in air. *Opt. Lett.*, 20:73–75, 1995.

[13] R.G. Brewer and J.R. Lifsitz. Narrow optical waveguides and instabilities induced in liquids. *Phys. Lett.*, 23:79–81, 1966.

[14] A. Brodeur, F.A. Ilkov, and S.L. Chin. Beam filamentation and the white light continuum divergence. *Opt. Commun.*, 129:193–198, 1996.

[15] A.V. Butenin, V.V. Korobkin, A.A. Malyutin, and M. Ya. Shchelv. Investigation of the kinetics of self-focusing in liquids. *JETP*, 6:173–176, 1967. Zh. Eksp. Teor. Fiz. - Pis'ma Red. (USSR) 6: 687–690, 1967.

[16] A.J. Campillo, S.L. Shapiro, and B.R. Suydam. Periodic breakup of optical beams due to self-focusing. *Appl. Phys. Lett.*, 23:628–630, 1973.

[17] S. Chi and Q. Guo. Vector theory of self-focusing of an optical beam in Kerr media. *Opt. Lett.*, 20(15):1598–1560, 1995.

[18] R.Y. Chiao, M.A. Johnson, S. Krinsky, C.H. Townes, and E. Garmire. 6A1 - a new class of trapped light filaments. *IEEE J. Quant. Elec.*, QE-2(9):467–469, 1966.

[19] Yu.S. Chilingarian. Self-focusing of inhomogeneous laser beams and its effect on stimulated scattering. *Sov. Phys. JETP*, 28:832–835, 1968. Zh. Eksp. Teor. Fiz. - Pis'ma Red. (USSR) 55(5):832–835.

[20] A. Ciattoni, P. DiPorto, B. Crosignani, and A. Yariv. Vectorial nonparaxial propagation equation in the presence of tensorial refractive-index perturbation. *J. Opt. Soc. Am. B*, 17:809–819, 2000.

[21] D.H. Close, C.R. Giuliano, R.W. Hellwarth, L.D. Hess, F.J. McClung, and W.G. Wagner. 8A2 - the self-focusing of light of different polarizations. *IEEE J. Quant. Elec.*, QE-2(2):553–556, 1966.

[22] B. Crosignani, P. DiPorto, and A. Yariv. Nonparaxial equation for linear and nonlinear optical propagation. *Opt. Lett.*, 22:778–780, 1997. Errata: 22:1820.

[23] R. de la Fuente, O. Varela, and H. Michinel. Fourier analysis of non-paraxial self-focusing. *Opt. Commun.*, 173:403–411, 2000.

[24] G. Fibich. Small beam nonparaxiality arrests self-focusing of optical beams. *Phys. Rev. Lett.*, 76:4356–4359, 1996.

[25] G. Fibich and A. Gaeta. Critical power for self-focusing in bulk media and in hollow waveguides. *Opt. Lett.*, 25:335–337, 2000.

[26] G. Fibich and B. Ilan. Self-focusing of circularly polarized beams. *Phys. Rev. A.* – submitted.

[27] G. Fibich and B. Ilan. Self-focusing of elliptic beams: an example for the failure of the aberrationless approximation. *J. Opt. Soc. Am. B*, 17:1749–1758, 2000.

[28] G. Fibich and B. Ilan. Deterministic vectorial effects lead to multiple filamentation. *Opt. Lett.*, 26:840–842, 2001.

[29] G. Fibich and B. Ilan. Vectorial and random effects in self-focusing and in multiple filamentation. *Physica D*, 157:113–147, 2001.

[30] G. Fibich and B. Ilan. Multiple filamentation of circularly polarized beams. *Phys. Rev. Lett.*, 89:#013901, 2002.

[31] G. Fibich, B. Ilan, and S. Tsynkov. Computation of nonlinear backscattering using a high-order numerical method. *J. Sci. Comput.*, 17:351–364, 2002.

[32] G. Fibich and G.C. Papanicolaou. A modulation method for self-focusing in the perturbed critical nonlinear Schrödinger equation. *Phys. Lett. A*, 239:167–173, 1998.

[33] G. Fibich and G.C. Papanicolaou. Self-focusing in the perturbed and unperturbed nonlinear Schrödinger equation in critical dimension. *SIAM J. Appl. Math.*, 60:183–240, 1999.

[34] G. Fibich and S. Tsynkov. High-order two-way artificial boundary conditions for nonlinear wave propagation with backscattering. *J. Comp. Phys.*, 171:632–677, 2001.

[35] S. Gatz and J. Hermann. Propagation of optical beams and the properties of two-dimensional spatial solitons in media with a local saturable nonlinear refractive index. *J. Opt. Soc. Am. B*, 14:1795–1805, 1997.

[36] I. Golub, R. Shuker, and G. Erez. On the optical characteristics of the conical emission. *Opt. Commun.*, 57:143–145, 1986.

[37] Q. Guo. Private communication.

[38] J.D. Jackson. *Classical Electrodynamics*. Wiley, New-York, 1975.

[39] P.L. Kelley. Self-focusing of optical beams. *Phys. Rev. Lett.*, 15(26):1005–1008, 1965.

[40] K. Konno and H. Suzuki. Self-focusing of a laser bean in nonlinear media. *Physica Scripta*, 20:382–386, 1979.

[41] V.V. Korobkin and R.V. Serov. Investigation of self-focusing of Neodymium-laser radiation. *JETP Lett.*, 6:135–137, 1967. Zh. Eksp. Teor. Fiz. - Pis'ma Red. (USSR) 6:642–644, 1967.

[42] M. Lax, W.H. Louisell, and W.B. McKnight. From Maxwell to paraxial wave optics. *Phys. Rev. A*, 11(4):1365–1370, 1975.

[43] G.J. Lord. Small-scale filaments in liquids and tracks of moving foci. *Phys. Rev. Lett.*, 22:994–997, 1969.

[44] P.D. Maker and R. W. Terhune. Study of optical effects due to an induced polarization third order in the electric field strength. *Phys. Rev.*, 137:A801–A819, 1965.

[45] P.D. Maker, R.W. Terhune, and C.M. Savage. Intensity-dependent changes in the refractive index of liquids. *Phys. Rev. Lett.*, 12:507–509, 1964.

[46] B.A. Malomed, K. Marinov, D.I. Pushkarov, and A. Shivarova. Stability of narrow beams in bulk Kerr-type nonlinear media. *Phys. Rev. A*, 64(023814):1–7, 2001.

[47] A.V. Mamaev, M. Saffman, D.Z. Anderson, and A.A. Zozulya. Propagation of light beams in anisotropic nonlinear media: From symmetry breaking to spatial turbulence. *Phys. Rev. A*, 54:870–879, 1996.

[48] J.H. Marburger and E.L. Dawes. Dynamical formation of small-scale filaments. *Phys. Rev. Lett.*, 21:556–558, 1968.

[49] Y.H. Meyer. Multiple conical emissions from near resonant laser propagation in dense Sodium vapor. *Opt. Commun.*, 34:439–444, 1980.

[50] C.S. Milsted and C.D. Cantrell. Vector effects in self-focusing. *Phys. Rev. A*, 53:3536–3542, 1996.

[51] M. Mlejnek, M. Kolesik, J.V. Moloney, and E.M. Wright. Optically turbulent femtosecond light guide in air. *Phys. Rev. Lett.*, 83:2938–2941, 1999.

[52] A.V. Nowak and D.O. Ham. Self-focusing of $10\mu m$ laser pulses in SF_6. *Opt. Lett.*, 6:185–187, 1981.

[53] N.F. Pilipetskii and A.R. Rustamov. Observation of self-focusing of light in liquids. *JETP Lett.*, 2:55–56, 1965. Zh. Eksp. Teor. Fiz. - Pis'ma Red. (USSR) 2:88–89 (1965).

[54] H. Prakash and N. Chandra. Change in polarization of electromagnetic waves in nonlinear isotropic dielectric. *Nuovo Cimento*, 57 (LVII B):215–217, 1968.

[55] J. Schwartz, P. Rambo, J.C. Diels, M. Kolesik, E.M. Wright, and J.V. Moloney. Ultraviolet filamentation in air. *Opt. Commun.*, 180:383–390, 2000.

[56] Y.R. Shen. *The Principles of Nonlinear Optics*. John Wiley & Sons, Inc., New-York, 1984.

[57] C.H. Skinner and P.D. Kleiber. Observation of anomalous conical emission from laser-excited Barium vapor. *Phys. Rev. A*, 21:151–157, 1980.

[58] J.M. Soto-Crespo, E.M. Wright, and N.N. Akhmediev. Recurrence and azimuthal-symmetry breaking of a cylindrical Gaussian beam in a saturable self-focusing medium. *Phys. Rev. A*, 45:3168–3174, 1992.

[59] C. Sulem and P.L. Sulem. *The Nonlinear Schrödinger Equation*. Springer, New-York, 1999.

[60] F. Vidal and T.W. Johnston. Electromagnetic beam breakup - multiple filaments, single beam equilibria and radiation. *Phys. Rev. Lett.*, 77:1282–1285, 1996.

[61] D.V. Vlasov, V.V. Korobkin, and R.V. Serov. Nonlinear precession of elliptically polarized Gaussian beams. *Sov. J. Quant. Electron.*, 9:904–907, 1979. Kvantovaya Elektron. (Moscow) 6, 1542–1546, 1979.

[62] S.N. Vlasov, V.A. Petrishchev, and V.I. Talanov. Averaged description of wave beams in linear and nonlinear media. *Izv. Vyssh. Radiofiz.*, 14:1353–1363, 1971. RadioPhy. Quant. Electon. (RPQEAC) 14 1062-1070 (1971).

[63] W.G. Wagner, H.H. Haus, and J.H. Marburger. Large-scale self-trapping of optical beams in the paraxial-ray approximation. *Phys. Rev.*, 175:256–266, 1968.

[64] M.I. Weinstein. Nonlinear Schrödinger equations and sharp interpolation estimates. *Commun. Math. Phys.*, 87:567–576, 1983.

[65] L. Wöste, C. Wedekind, H. Wille, P. Rairoux, B. Stein, S. Nikolov, C. Werner, S. Niedermeier, F. Ronnenberger, H. Schillinger, and R. Sauerbrey. Femtosecond atmospheric lamp. *Laser Optoelektron.*, 29:51–53, 1997.