

Dynamics and Interactions of
Coherent Structures
in Nonlinear Systems

Cornelis Storm

The cover depicts a close-up view of the structure of the matrix that determines the propagation in Fourier space of the laser modes considered in Chapter 6. Parameter values used were $M = 2$ and $N = 1800$.

Dynamics and Interactions of
Coherent Structures
in Nonlinear Systems

P R O E F S C H R I F T

TER VERKRIJGING VAN
DE GRAAD VAN DOCTOR AAN DE UNIVERSITEIT LEIDEN,
OP GEZAG VAN DE RECTOR MAGNIFICUS DR. D. D. BREIMER,
HOGLERAAR IN DE FACULTEIT DER WISKUNDE EN
NATUURWETENSCHAPPEN EN DIE DER GENEESKUNDE,
VOLGENS BESLUIT VAN HET COLLEGE VOOR PROMOTIES
TE VERDEDIGEN OP DINSDAG 19 JUNI 2001
TE KLOKKE 15.15 UUR

DOOR

Cornelis Storm

GEBOREN TE GROENLO IN 1973

Promotiecommissie:

Promotor: Prof. dr. ir. W. van Saarloos
Referent: Prof. dr. ir. W. van de Water (TU Eindhoven)
Overige leden: Prof. dr. P. van Baal
dr. U. Ebert (CWI Amsterdam)
dr. M.L. van Hecke
Prof. dr. P.H. Kes
Prof. dr. J.M.J. van Leeuwen

*Don't think twice,
it's all right -*

Bob Dylan

Contents

1	Pattern Formation	1
1.1	Introduction	1
1.2	Convection in a fluid layer heated from below	1
1.3	Some phenomenology	6
1.4	Basic questions	7
1.5	Organization of this thesis	10
2	Hydrodynamic Instabilities and Amplitude Equations	13
2.1	Introduction	13
2.2	Fluid dynamics and the Navier-Stokes equation	13
2.3	The Rayleigh-Bénard instability	19
2.4	The linear instability	23
2.5	The Swift-Hohenberg equation	28
2.6	The amplitude expansion	31
2.7	Implications of the amplitude description	40
2.8	The complex Ginzburg-Landau equation	41
2.9	Coherent structures	46
2.10	Amplitude equations and symmetries	47
3	Sources and sinks in traveling wave systems	51
3.1	Introduction	51
3.2	The heated wire experiment	51
3.3	Amplitude equations for the heated wire	55
3.4	Definition of sources and sinks	56
3.5	Coherent structures and counting arguments	59
3.5.1	General formulation and main results	59
3.5.2	Comparison between shooting and direct simulations	64
3.5.3	Multiple discrete sources	65

3.6	Scaling properties of sources and sinks for small ε	67
3.6.1	Coherent sources: analytical arguments	69
3.6.2	Sources: numerical simulations	71
3.6.3	Sinks	76
3.6.4	The limit $s_0 \rightarrow 0$	77
3.7	Conclusion	79
3.A	Coherent structures in the single CGLE	81
3.A.1	The flow equations	81
3.A.2	Fixed points and linear flow equations in their neighborhood	83
3.A.3	The linear fixed points	84
3.A.4	The nonlinear fixed points	85
3.B	Detailed counting for the coupled CGL equations	87
3.B.1	General considerations	87
3.B.2	Multiplicities of sources and sinks	89
3.B.3	The role of ε	91
3.B.4	The role of the coherent structure velocity v	91
3.B.5	Normal sources always come in discrete sets	92
3.B.6	Counting for anomalous $v = 0$ sources	93
3.B.7	Counting for anomalous structures with $\varepsilon_{\text{eff}} > 0$ for the suppressed mode	94
3.C	Asymptotic behavior of sinks for $\varepsilon \rightarrow 0$	95
4	Dynamical Properties of Source/Sink Patterns	97
4.1	Introduction	97
4.2	Convective and absolute sideband-instabilities	98
4.3	Instability to bimodal states: source-induced bimodal chaos	106
4.4	Mixed mechanisms	109
4.4.1	Core instabilities and unstable waves	110
4.4.2	Phase slips and bimodal instabilities	110
4.4.3	Intermittency and bimodal instabilities	111
4.4.4	Periodic and other states	115
4.5	Interactions between sources and sinks	116
4.5.1	Setup of the problem	122
4.5.2	Zero-modes of the linear operator	124
4.5.3	Solvability conditions	126
4.6	Outlook	126
4.7	Experimental implications	127
4.8	Comparison of results with experimental data	128

4.8.1	Heated wire experiments	128
4.8.2	Binary mixtures	133
4.9	Conclusion	135
4.A	Details of the interactions calculation	136
5	Universal algebraic relaxation in pulled front propagation	137
5.1	Introduction	137
5.2	Fronts in the nonlinear diffusion equation	137
5.3	Velocity selection.	140
5.4	Uniformly translating pulled fronts	143
5.5	Coherent pattern generating fronts	145
5.6	Incoherent or chaotic fronts	146
5.7	Choosing the proper frame and transformation	147
5.8	Understanding the intermediate asymptotics	148
5.9	Systematic expansion	150
5.10	Conclusion	152
6	Fractal Lasers	153
6.1	Introduction	153
6.2	Setup of the problem	156
6.3	Analytical results	157
6.3.1	Fourier transform	157
6.3.2	Approximate evaluation	160
6.4	Numerical results	164
6.5	Magnitude of the largest eigenvalue	166
6.6	k -Space matrix for an even state	167
6.7	Conclusion	169
6.A	Exact expressions for the k -space matrix elements	172
	Bibliography	173
	Samenvatting	181
	Publications	187
	Curriculum Vitae	189

Pattern Formation

1.1 Introduction

As this thesis is going to be about the theoretical description of the phenomenon known as pattern formation, we would first like to define what exactly is meant by pattern formation. In this first chapter we hope to give the reader some feel for the field, its phenomenology and for some of the basic issues. What better place for such a discussion to start, than the experiment which has been called the “granddaddy of canonical examples used to study pattern formation and behavior in spatially extended systems” by Alan Newell [1], one of the pioneers of the amplitude approach that will be a (if not *the*) central topic of this work. The example Newell alluded to is that of the motion of a fluid in a large, shallow container which is uniformly heated from below. This setup has become known in the literature as the Rayleigh-Bénard experiment, and we will begin here with a brief discussion of its history and main characteristics.

1.2 Convection in a fluid layer heated from below

The actual discovery of thermal convection in fluids is generally attributed to Sir Benjamin Thompson, Count Rumford [2], who devoted most of his scientific career to investigations into the nature of heat (and was, in fact, the first to show that heat is not a liquid form of matter, but rather a form of *energy*). It wasn’t until some 200 years after the initial discovery of

the phenomenon though, that the first quantitative experimental research on convection was performed by Bénard [3]. He did so in a setup in which a thin layer of fluid is heated from below. The actual apparatus he used is displayed in Fig. 1.1. As was already known in those days, upon a steady increase of the temperature of the bottom plate supporting the fluid the fluid would at some point start to move or *convect*. What initially escaped the notice of many however was that convection only set in when the temperature exceeded some finite value, and *not* for arbitrarily small heating. The physical mechanism responsible for the convection in this experiment is readily identified: since the fluid near the bottom of the container is heated, and fluids generally expand upon heating, the local mass density is lower near the bottom than it is at the top. This situation is inherently unstable, as the heavier fluid at the top will try to fall down into the less dense fluid under the influence of gravity, resulting in a large-scale motion in the fluid. Bénard however was the first to realize that the situation is stabilized to some extent by viscous effects, and started doing experiments on highly viscous fluids such as melted spermaceti, a fatty byproduct of the whaling industry, and paraffin. It was precisely the use of these high-viscosity fluids that enabled him to see clearly that the convective state did, in fact, only appear for sufficiently high values of the bottom plate temperature. This was however not his main interest in the experiment.

Previous experiments, notably those by James Thomson [5], had already established some of the curious properties of the convection as it takes place in the setup Bénard was also using, the most striking of which being that the convection generally takes place in the form of a very regular *cellular pattern*, as one of Bénard's original figures, reproduced in Fig. 1.2 illustrates. Moreover, he was able to determine the nature of the fluid motion within these cells: it turned out that the fluid was rising in the center, while it was coming down again at the boundaries of the individual cells. Bénard observed that for temperatures sufficiently high to sustain convection, the cells would initially quickly form and take on convex polygonal shapes with four to seven sides, but that after this initial phase came a phase in which all of the cells would slowly become approximately equal in size, and form a remarkably well-aligned hexagonal structure, not unlike the structure some regions in Fig. 1.2 exhibit. Furthermore, this second phase was seen to change only over timescales much larger than those characteristic of the initial phase, in which the

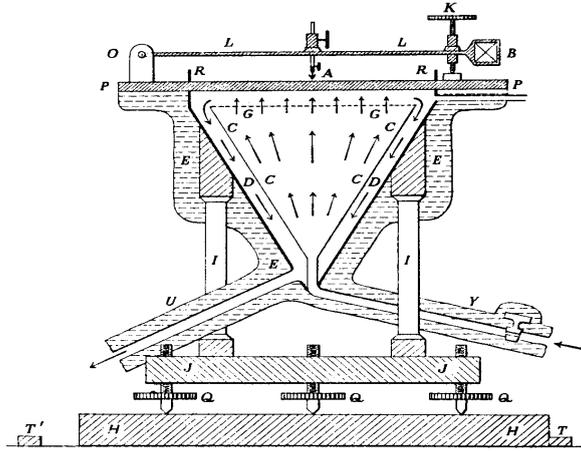


Figure 1.1: The original apparatus used by Bénard. Uniform heating from below is provided by a steady flow of hot liquid pumped through the device along the trajectory YCGDU. The fluid, which in contrast to the 'modern' Rayleigh-Bénard setup has a free surface, is on the levelled metallic plate P, within a circular container whose boundary is marked R. Figure taken from [4].

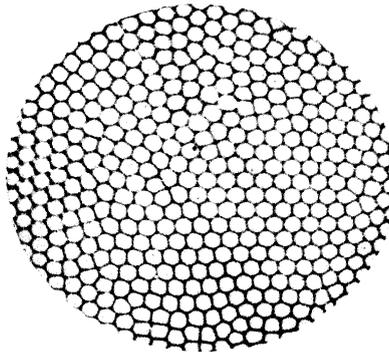


Figure 1.2: Reproduction of one of Bénard's original photographs of convection in spermaceti. Dark regions correspond to fluid moving downward, while light regions indicate fluid is being convected to the surface. Figure taken from [4].

cells were formed. As he kept increasing the temperature, Bénard went on to discover an astonishingly rich variety of convection patterns, some of which were highly ordered, but also others which correspond to states which we would nowadays describe as *turbulent* (or space-time chaotic). [6].

The “interesting results obtained by Bénard’s careful and skillful experiments” (as Lord Rayleigh would later refer to them) were begging for a solid mathematical analysis, but it was not before another 20 years had passed that Lord Rayleigh actually published his seminal paper on the subject [7]. The paper contains an in-depth analysis of the hydrodynamical instability that underlies the appearance of the convective pattern, starting from the Navier-Stokes equations of hydrodynamics. Although Bénard’s experiments were done in an *open* container, Rayleigh chose to simplify matters slightly by studying the system in a closed geometry, which was uniformly heated at the bottom plate and where the ‘lid’, or top plate, was also kept at a constant (but of course lower) temperature. The main result of Lord Rayleigh’s effort is a solid prediction for the temperature difference at which convection would set in, accompanied by a prediction for the *wavenumber* of the pattern at onset. This wavenumber is to be interpreted as the typical size of the convection cells in the asymptotic ‘permanent’ cellular pattern. But, not only did Rayleigh derive these values, he was also able to give them in terms of *dimensionless quantities*, such as the number that nowadays bears his name, the Rayleigh number $R = \frac{\alpha g \Delta T h^3}{\xi \nu}$ which is a dimensionless measure of the temperature difference ΔT , involving the thermal expansion coefficient α , the acceleration of gravity g , the container height h , the thermal diffusion coefficient ξ and the dynamical viscosity ν . The advantage of the use of dimensionless quantities lies in the fact that because all specifics of the materials and setup used are absorbed into them, they allow us to compare a multitude of experiments to each other. In particular, his analysis led Rayleigh to predict that *no matter what the experimental realization*, the instability should set in at $R = 1708 \equiv R_c$. As stated before, the actual onset of convection was not Bénard’s main interest, but as he would later claim after analyzing his old data, it did seem to occur roughly where Lord Rayleigh had predicted it. This is surprising to say the least, and is an indication that the surface tension effects induced by the free surface (effects *completely absent* in Rayleigh’s analysis) play only a very minor role in this particular setup. Subsequent experimental studies *did* however produce

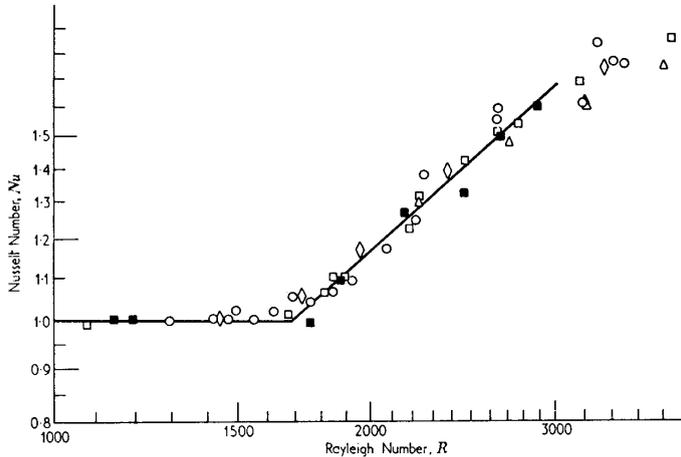


Figure 1.3: Silveston’s experimental results for the Nusselt number Nu as a function of the Rayleigh number R for various liquids (solid squares: silicone oil AK350, diamonds: silicone oil AK3, open squares: glycol, triangles: heptane, circles: water). The data indicate that for all of these liquids (which vary appreciably in viscosity), the instability sets in at a critical Rayleigh number of $R_c = 1700 \pm 51$. Figure from [9]

accurate measurements of the critical Rayleigh number in this experiment that by this time was generally known by the name it still carries today: the Rayleigh-Bénard experiment. Early experiments, such as those conducted by Schmidt and Milverton [8] already confirmed that Rayleigh’s results were essentially correct, yielding the result that $R_c = 1770 \pm 140$. Perhaps one of the most striking verifications of Rayleigh’s predictions can be found in an experimental paper by Silveston [9], the main result of which is plotted in Fig. 1.3. Plotted in this figure is the Nusselt number Nu , a dimensionless quantity measuring the ratio of the *total* heat transported from bottom to top over the heat transported purely by conduction; $Nu = \frac{Q_{conv} + Q_{cond}}{Q_{cond}}$ versus the Rayleigh number R . The full advantage of using dimensionless quantities becomes apparent from this figure, in which data are collected from several *different* experimental realizations, all of which are seen to collapse nicely onto one curve. For values of R below the critical value, the Nusselt number will be equal to one (as there is no convection), but when convection *is* present, the Nusselt number will be *larger* than one. As Fig. 1.3 clearly indicates, the instability does indeed

set in at, or at least very close to the predicted value.

Thus, the first milestone in pattern formation was reached. Both intuitively and analytically the linear instability could be understood from hydrodynamic first principles, and experimental findings were in excellent accord with the theory as it stood. Understanding the linear instability however, as much of an achievement in its own right as this may be, is only part of understanding pattern formation as a whole, considering it also encompasses the rich pattern dynamics (possibly far) beyond threshold. Over the subsequent years more and more examples of pattern forming systems were discovered, and although the underlying physical mechanism was in many cases completely unlike that of the Rayleigh-Bénard system, the actual *patterns* observed turned out to be remarkably alike. Let us have a quick look what we are “up against”...

1.3 Some phenomenology

The Rayleigh-Bénard system is by no means the only system to exhibit patterns like the ones shown in Fig. 1.2. Although convection experiments in a variety of particular setups are still the most widely used, this is to large extent because of their relative simplicity¹. As we will demonstrate in Section 2.10, the hexagons that Bénard observed are actually the very first pattern one expects to see in such a system. They are however not the simplest modes that the system supports, which instead are patterns consisting of straight rolls such as the one displayed in Fig. 1.4. As is obvious from Fig. 1.4, a straight roll pattern can essentially be described as a one-dimensional (the direction perpendicular to the rolls) system with a well-defined spatial periodicity. These straight-roll patterns change completely upon further heating, and different types of convection patterns take over. One of the most common ones consisting of spirals. Fig. 1.5(a) shows a straight-roll state being invaded by a pattern known as spiral defect chaos, which has almost taken over the entire convection cell in Fig. 1.5(b). The spiral defect chaotic state is very dynamic: all individual spirals rotate, and new spiral cores are being created while others are annihilated. Although the spiral defect state is very much a two-dimensional specialty, a single spiral *does* possess a one dimensional analogue, which consists of a point-like core sending out *traveling waves* to either side.

¹Even olive oil in a frying pan will display the characteristic cellular pattern upon heating on a stove.

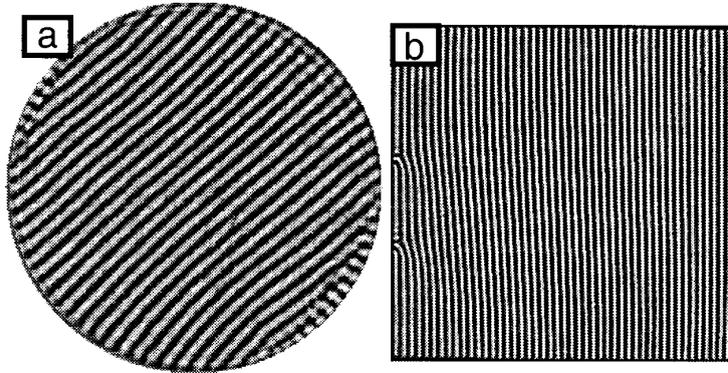


Figure 1.4: Nearly ideal straight roll patterns in Rayleigh-Bénard convection for circular and a square container. Note the slight imperfections in the patterns at the edges, due to the influence of the boundaries ((a) taken from [10], (b) from [11]).

Structures like these *sources* will be discussed in detail in Chapters 3 and 4. What also occurs in one dimensional systems is chaos invading otherwise quiescent systems. An example of such a *chaotic front* is shown in Fig. 1.7. Spiral states turn out to be very common, and although Fig. 1.6 looks about the same as the plot for the convection experiment, the system it displays could hardly be more different: we are looking at the slime mold *Dictyostelium Discoideum*, a living one-celled organism that forms these kind of structures when it is starved of nutrition. The individual cells are *signaling* to each other by the excretion of the chemical cyclic AMP. The cells are capable of sensing gradients in the concentration of this chemical in their immediate surroundings, and move in the direction of these gradients, a process known as *chemotaxis*.

1.4 Basic questions

As we have seen, a variety of physical systems produces similar patterns when driven sufficiently far from equilibrium. It is however not the optical similarity that we will be interested in here, but their behavior close to threshold. In all of the systems mentioned in the previous section, it is possible to identify a single *control parameter* (such as the Rayleigh

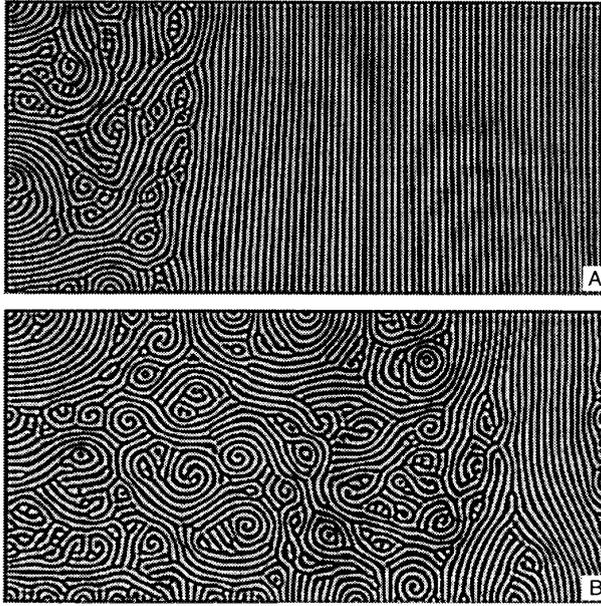


Figure 1.5: (a) Spiral defect chaos propagating into a straight-roll state in a rectangular Rayleigh-Bénard cell. Displayed in (b) is the same system at a later time, and the chaotic state has almost taken over the entire cell. Figure taken from [12].

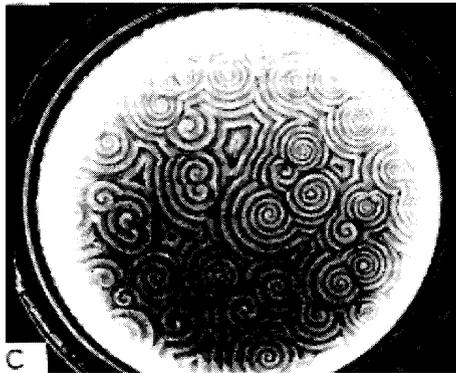


Figure 1.6: Spiral patterns in the aggregation phase of *Dictyostelium Discoideum*. Figure reproduced from [13].

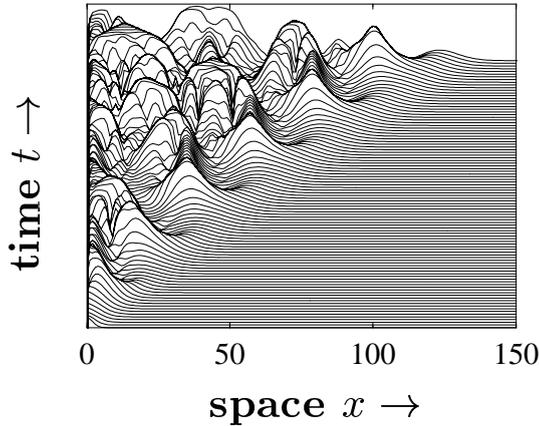


Figure 1.7: Space-timeplot of a chaotic state invading an ordered state in a one-dimensional model equation. For details, see Chapter 5

number R for the Rayleigh-Bénard system) which measures the distance from the onset of the instability, provided of course one knows the critical value. In the study of any pattern forming system, it is essential that one knows the nature of the primary instability leading to the formation of patterns. Once this has been identified, as we shall see, it is possible to construct a workable *weakly nonlinear* theory for the behavior of pattern forming systems close to this threshold, by the use of what is essentially an expansion in the *deviation* of the control parameter from its critical value. This description is known as the amplitude description, and will be treated in some detail in Chapter 2. The main reason why this approach works is because it so happens that the behavior close to onset can be described by considering the slow and long-length scale *modulations* of the pattern precisely *at* onset. The goal of any physical theory of pattern formation should be to adequately describe the nature of the states that the system is likely to reach from physical initial conditions, and that is exactly what we demand from this theory as well. On the level of the linearized equations however, all of the pattern forming systems mentioned up to now possess not just one, but rather a continuous family or *band* of unstable states beyond threshold. Linear theory cannot differentiate between the competing allowed modes. Rather, the mode that is actually *observed* is determined by a combination of *external biases*, such as boundary conditions, imperfections or impurities, and the *nonlinear coupling* between

the various competing configurations, each of which is equally likely to occur in a linearized theory. Without the identification of the *nonlinear selection mechanism* at play therefore, no theory can be complete.

The patterns that arise in nature very rarely consist of simple small variations of straight roll patterns or perfect hexagonal cells. Instead, the interplay of rotational symmetry and boundary conditions tends to yield complicated patterns much more like those in Fig. 1.5(b), riddled with structures known as *defects*, localized regions of discontinuity in the patterns such as boundaries between rolls of differing orientation or centers of spiral waves that can impossibly be considered small variations to a pattern consisting of straight rolls. It is therefore a great challenge to develop a weakly nonlinear theory that is able to take maximal advantage of the underlying periodic structure as obtained from the linear instability, but which is also able to handle the singularities that invariably come with a macroscopic description of such patterns.

1.5 Organization of this thesis

In Chapter 2, we will present a detailed analysis of the linear instability in the Rayleigh-Bénard system, after which we will introduce the Swift-Hohenberg equation, a toy model that is designed to reproduce the essentials of the hydrodynamic equations used for the analysis of the Rayleigh-Bénard model, but which is mathematically much easier to handle. It is with these equations that we demonstrate the so-called Amplitude formalism, along the lines of the technique originally employed by Newell and Whitehead to arrive at an effective theory of the pattern dynamics above, but near the instability threshold. Chapter 2 concludes with a brief account some of the implications of this amplitude description. In Chapter 2, we derive the amplitude equations that should adequately describe a particular convection experiment known as the heated wire experiment, based on the basic symmetries and some empirical input from experiments. As it turns out, particular types of solutions belonging to the class of *coherent structures* are extremely relevant in organizing the overall behavior and stability of this system, and in Chapter 3 we investigate a number of properties for these *sources and sinks*, as they are known. Because the equations we use to study this system are determined by basic symmetry considerations, they should describe a much *wider* class of experiments however. Chapter 4 is an investigation into the implications

that these sources and sinks have for the spatial and temporal dynamical behavior of such systems, up to and including chaotic regimes. Also, we study the interactions between sources and sinks. In Chapter 5, we focus on a different type of coherent structure, known as a *front*. Fronts are regions that separate one state or phase in a system from another, and in general these fronts move or *propagate*. The gradual invasion of one state by the other is always accompanied by a front, and we study the velocity of propagation for a variety of nonlinear model equations, each of which should cover a number of actual experimental situations. Even though these equations all have quite different characteristics, and the invading states range from completely trivial to chaotic, we show that there is nonetheless a remarkable degree of universality both in the asymptotic velocity reached and in the way these fronts approach their asymptotic states. Chapter 6, the final chapter is somewhat detached in subject from the rest of this thesis, as it is dealing with a remarkable property of unstable laser cavities that was recently discovered by the quantum optics group in Leiden. What they found was that the eigenmodes of such laser cavities were *fractal* in nature. We show from the underlying equations that this is indeed the case, and give an intuitively simple explanation of the effect. In addition, we obtain the fractal *dimension* of these modes, which is verified numerically.

Hydrodynamic Instabilities and Amplitude Equations

2.1 Introduction

In this chapter, we will attempt to clarify some of the underlying principles of pattern formation using one of the oldest and most widely studied examples of a pattern forming hydrodynamic AI instability, the Rayleigh-Bénard instability. Starting from the Navier-Stokes equation, we will use this instability to demonstrate various physical ideas and mathematical techniques commonly used to analyze a variety of nonlinear phenomena associated with instabilities of nonlinear systems. What all this will eventually lead to is a rather general description of pattern forming dynamics close to (but just above) the onset of the instability, where the nonlinearities are still weak. This will be achieved by means of a procedure known as the amplitude equation formalism, which will play a key role throughout this thesis.

2.2 Fluid dynamics and the Navier-Stokes equation

Although the equations that adequately describe the motion of a fluid (a *fluid* meaning a liquid or a gas) were known already quite a long time ago, studying and possibly solving them in a staggering number of particular

settings has been one of the main problems of classical physics, receiving huge attention before the advent of quantum mechanics. One of the other main fields in those days was of course the study of the motion of individual bodies or *classical mechanics*. The main difference between classical mechanics and fluid dynamics is that fluids are supposed to be *continuous* media, by which we mean to convey the idea that we will only be interested in fluid properties manifesting themselves at scales at which the properties of the individual constituents (molecules or atoms) of the fluid do not matter any more. A continuous medium is therefore characterized not by discrete equations of motion for each of the constituent particles separately, but rather by *fields*, the hydrodynamic variables, taking on values in the whole of space and time. Another way of stating this same idea is that we average over the individual motion of the fluid constituents to end up with a coarse-grained description of the fluid on length scales much larger than the typical size of the fluid particles. In fact, although in its strictest sense the term “hydrodynamic equations” is reserved for the Navier-Stokes equations describing fluid flow, over the past years it has also become more widely used as the name for an effective theory describing a system which is in principle discrete at lengthscales and timescales larger than those characteristic of its individual entities, whether these be particles, spins or of different nature. In order to arrive at such an effective theory, one has to take what is called the “thermodynamic limit”, *i.e.* one lets the number of constituents tend to infinity.

Back now to the problem at hand: setting up the equations of motion for a fluid. Let us begin this derivation by restricting ourselves to a simple, one-component fluid. To describe such a fluid in full, one needs as many equations as there are conserved quantities¹. For the following discussion, it will prove most instructive to choose as these independent conserved quantities the pressure, the three components of the fluid momentum and the fluid energy.

Associated with each of these five quantities will be three hydrodynamic *fields*, one of a vector nature and two scalar fields. Their static and

¹To see this, consider a small volume of fluid. Any change in a globally conserved quantity in this volume can only be caused by a flux into or out of the volume at its boundaries. Associating exactly one hydrodynamic field with each of these conserved quantities will therefore yield a complete description of the system, since the other nonconserved quantities will decay fast.

dynamical properties will be our main concern this chapter. They are

$$p(\vec{r}, t) \quad : \quad \text{pressure}, \quad (2.2.1a)$$

$$\vec{v}(\vec{r}, t) \quad : \quad \text{velocity}, \quad (2.2.1b)$$

$$T(\vec{r}, t) \quad : \quad \text{temperature}. \quad (2.2.1c)$$

The hydrodynamic or continuum nature of these quantities is illustrated by their dependence on the (continuous) spatial variable \vec{r} . When referring to these fields, we will in the future omit the explicit dependence on \vec{r} and t . Although the natural variable associated with energy conservation would be the entropy, different quantities are of course related by thermodynamic relations. Mass conservation is expressed by the *equation of continuity*

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot [\rho \vec{v}] = 0, \quad (2.2.2)$$

while momentum conservation leads to the vector equation

$$\frac{\partial \rho \vec{v}}{\partial t} + \vec{\nabla} \cdot [\rho \vec{v} \vec{v}] = - \vec{\nabla} \cdot \vec{\sigma} + \vec{f}_{\text{ext}}. \quad (2.2.3)$$

We use here the (dyadic) tensorial product of two 3-vectors, which when expressed in components reads

$$(\vec{v} \vec{v})_{ij} = v_i v_j. \quad (2.2.4)$$

The quantity $\vec{\sigma}$ is generally known as the *stress tensor*, whose (ij) th component is the amount of force per unit area in direction j on the surface with normal in the i -direction. \vec{f}_{ext} is the external force per unit volume. While in an inviscid (frictionless) fluid, hydrostatic pressure is the only contribution to the stress tensor, when we allow for viscous effects to be present it is customary to distinguish between the hydrostatic and viscous contributions as follows

$$\sigma_{ij} = p \delta_{ij} + \sigma'_{ij}, \quad (2.2.5)$$

where δ_{ij} is the Kronecker delta symbol, p is the hydrostatic pressure (which appears, as it should, on the diagonal of $\vec{\sigma}$), and all viscous contributions are accounted for in $\vec{\sigma}'$, sometimes referred to as the *viscosity stress tensor*. We can construct this tensor as follows. Firstly, we note that in order to get processes of internal friction, different parts of the fluid need to be moving at different velocities. Therefore, $\vec{\sigma}'$ has to vanish

for a spatially homogeneous velocity field, and can only depend on the derivatives of the velocity field. Assuming now that there exists an expansion in these derivatives, which is local in time (implying a dependence on the *spatial* derivatives only), we find to lowest order that $\vec{\sigma}'$ is a linear function of the *gradients* $\frac{\partial v_i}{\partial x_j}$

$$\sigma'_{ij} = A_{ijkl} \frac{\partial v_k}{\partial x_l}, \quad (2.2.6)$$

where the summation over indices appearing twice is implied (the Einstein summation convention), and we still need to determine the precise form of the tensor A_{ijkl} , which we assume to be independent of position. We can split σ'_{ij} in a symmetric and an antisymmetric part

$$\sigma'_{ij} = \frac{A_{ijkl}}{2} \left(\frac{\partial v_k}{\partial x_l} + \frac{\partial v_l}{\partial x_k} \right) + \frac{A_{ijkl}}{2} \left(\frac{\partial v_k}{\partial x_l} - \frac{\partial v_l}{\partial x_k} \right) \quad (2.2.7)$$

$$\equiv A_{ijkl} e_{kl} - \frac{1}{2} A_{ijkl} \varepsilon_{klm} \omega_m, \quad (2.2.8)$$

where we have introduced the *vorticity* $\vec{\omega} = \vec{\nabla} \times \vec{v}$, and ε_{ijk} is the completely antisymmetric, or Levi-Civita tensor in three dimensions. Since a fluid in uniform rotation also has no internal friction, σ'_{ij} should also vanish when $\vec{v} = \vec{\Omega} \times \vec{r}$. The vorticity of this velocity field is simply $\vec{\omega} = 2\vec{\Omega}$ while the symmetric combination e_{kl} vanishes identically, leaving us with

$$A_{ijkl} \varepsilon_{klm} \Omega_m = 0 \quad \forall \Omega, \quad (2.2.9)$$

which is satisfied when A_{ijkl} is symmetric in its last two indices. In an isotropic fluid, *i.e.* one that does not distinguish between the different spatial directions, A_{ijkl} should be what is known as a completely isotropic tensor. Such a tensor in arbitrary dimensions can always be expressed in terms of the Kronecker δ -tensor, and in four dimensions its most general form can be shown to be [14]

$$A_{ijkl} = \mu \delta_{ik} \delta_{jl} + \mu' \delta_{il} \delta_{jk} + \mu'' \delta_{ij} \delta_{kl}. \quad (2.2.10)$$

Since σ'_{ij} is a symmetric tensor μ and μ' are necessarily equal, and we find for σ'_{ij}

$$\sigma'_{ij} = 2\mu \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) + \mu'' \frac{\partial v_k}{\partial x_k}, \quad (2.2.11)$$

but it is customary to split of a traceless part by defining

$$\eta \equiv 2\mu, \quad (2.2.12)$$

$$\zeta \equiv \mu'' - \frac{4}{3}\mu \quad (2.2.13)$$

which yields the standard definition of the viscosity stress tensor

$$\sigma'_{ij} = \eta \left[\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\delta_{ij} \sum_k \frac{\partial v_k}{\partial x_k} \right] + \zeta \delta_{ij} \sum_k \frac{\partial v_k}{\partial x_k}. \quad (2.2.14)$$

The coefficients η and ζ are known as the *dynamic viscosity* (or simply *the viscosity*) and *second viscosity* respectively.

A general one-component thermodynamic system can be fully characterized by two thermodynamic quantities, temperature and density for instance. Therefore, generally speaking quantities like η and ζ should also be considered functions of temperature and density. In real life however, it turns out that the viscosities are usually to very good approximation constant throughout the fluid, and their dependence on the thermodynamic quantities can be safely ignored. Assuming this holds, we can combine Eqs. (2.2.3), (2.2.2), (2.2.5) and (2.2.14) to obtain the famous Navier-Stokes equation

$$\rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} \right] = -\vec{\nabla} p + \eta \nabla^2 \vec{v} + \left(\zeta + \frac{1}{3}\eta \right) \vec{\nabla} (\vec{\nabla} \cdot \vec{v}) + \vec{f}_{\text{ext}}. \quad (2.2.15)$$

All we need now is the equation associated with energy conservation. The full derivation is quite lengthy, and it will suffice here to skip most of the details, which are well documented among others in [15]. The so-called *general equation of heat transfer* in hydrodynamics expresses the balance of entropy in the presence of viscous effects²

$$\rho T \left[\frac{\partial s}{\partial t} + (\vec{v} \cdot \vec{\nabla}) s \right] = \vec{\sigma}' : (\vec{\nabla} \vec{v}) + \vec{\nabla} \cdot (\kappa \vec{\nabla} T), \quad (2.2.16)$$

where κ is the thermal conductivity. We can understand the significance of the different terms in Eq. (2.2.16) when we realize that the quantity on the left is nothing but the total time derivative $\frac{ds}{dt}$ of the entropy of the fluid multiplied by ρT , which is the amount of *heat* gained per unit

²Systems that “leak away” energy are said to be *dissipative*. In this case, energy is dissipated into heat by viscosity.

volume. This heat can be gained (or lost) either by viscous dissipation, measured by the first term on the right hand side, or it can diffuse away from the volume under consideration. The last effect is accounted for in the second term on the right hand side, which is simply Fick's law for thermal diffusion. We will use the colon to denote contraction of two 3-tensors of rank 2 to yield a scalar

$$\vec{A}:\vec{B} \equiv \sum_{i=1}^3 \sum_{j=1}^3 A_{ij}B_{ij}. \quad (2.2.17)$$

In view of the systems we will be considering further on however, it would be more convenient to have an equation for the *temperature* field. We can obtain such an equation by using the following relations from thermodynamics

$$\frac{\partial s}{\partial t} = \left(\frac{\partial s}{\partial T} \right)_p \frac{\partial T}{\partial t} = \frac{c_p}{T} \frac{\partial T}{\partial t}, \quad (2.2.18a)$$

$$\vec{\nabla} s = \left(\frac{\partial s}{\partial T} \right)_p \vec{\nabla} T = \frac{c_p}{T} \vec{\nabla} T, \quad (2.2.18b)$$

which hold under the assumption that the variations in the density caused by changes in the pressure field are small enough to be neglected. A practical condition for this is that the fluid velocity should be small compared to the velocity of sound. c_p is the specific heat at constant pressure. Substituting Eqs. (2.2.18) into Eq. (2.2.16), we find the equation for the temperature

$$\frac{\partial T}{\partial t} + (\vec{v} \cdot \vec{\nabla})T = \frac{\kappa}{\rho c_p} \nabla^2 T + \frac{1}{\rho c_p} \vec{\sigma}':(\vec{\nabla}\vec{v}), \quad (2.2.19)$$

which completes the set of equations we will be using. In summary, the 5 equations describing the motion of a fluid in the presence of viscous effects read

$$\left\{ \begin{array}{l} \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot [\rho \vec{v}] = 0, \\ \frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla)\vec{v} = -\frac{1}{\rho} \vec{\nabla} p + \nu \nabla^2 \vec{v} + \left(\frac{\zeta}{\rho} + \frac{1}{3}\nu \right) \vec{\nabla}(\vec{\nabla} \cdot \vec{v}) + \vec{f}_{\text{ext}}, \\ \frac{\partial T}{\partial t} + (\vec{v} \cdot \vec{\nabla})T = \chi \nabla^2 T + \frac{1}{\rho c_p} \vec{\sigma}':(\vec{\nabla}\vec{v}). \end{array} \right. \quad (2.2.20)$$

Here, $\nu \equiv \frac{\eta}{\rho}$ is the *kinematic viscosity*, $\chi = \frac{\kappa}{\rho c_p}$ is the *thermometric conductivity* or *thermal diffusion coefficient*. In all applications treated in this thesis, it will be appropriate to consider the fluid *incompressible*, that is, the mass density ρ is assumed to be a constant, although it *is* allowed to vary with temperature. Mass conservation (2.2.2) in an incompressible fluid simply amounts to the requirement

$$\vec{\nabla} \cdot \vec{v} = 0. \quad (2.2.21)$$

Although this condition does reduce the full set of hydrodynamic equations (2.2.20) slightly, it does *not* remove the nonlinearities. Even for incompressible flows, very few settings allow for exact solutions, although there exist some particular examples where symmetries cause the nonlinear terms to vanish, such as Poiseuille flow between plates or in a pipe. Eqs. (2.2.20) for an incompressible fluid reduce to

$$\left\{ \begin{array}{l} \vec{\nabla} \cdot \vec{v} = 0, \\ \frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \vec{v} = -\frac{1}{\rho} \vec{\nabla} p + \nu \nabla^2 \vec{v} + \vec{f}_{\text{ext}}, \\ \frac{\partial T}{\partial t} + (\vec{v} \cdot \vec{\nabla}) T = \chi \nabla^2 T + \frac{2\nu}{c_p} (\vec{\nabla} \vec{v}) : (\vec{\nabla} \vec{v}). \end{array} \right. \quad (2.2.22)$$

It is in this form that we will use the hydrodynamic equations throughout the remainder of this chapter.

2.3 The Rayleigh-Bénard instability

We will now analyze in more detail the convection problem introduced in Chapter 1. A layer of fluid, sandwiched between two parallel plates is subjected to a vertical temperature gradient ΔT . The excess heat at the bottom plate has somehow to be transported to the top plate, which can happen by two distinct mechanisms. These mechanisms are *conduction* and *convection*. What sets these two apart is that in convective transport, the fluid itself is in motion while for conductive transport it is at rest. In experimental studies of the Rayleigh-Bénard system, which was introduced in Chapter 1, it was found that for small temperature differences heat is only transported by means of conduction, but for values of ΔT larger than some critical value ΔT_c convection suddenly sets in—the fluid starts to move in a very distinct manner, in that *rolls* of moving fluid

appear. These rolls have a well-defined wavelength, and near the onset of convection this wavelength is very close to the container height h . We can gain some qualitative understanding of this behavior by considering the physical mechanisms that inhibit and promote fluid flow in this particular setup.

Since the fluid near the bottom plate is hotter than that at the top plate, and fluids in general expand upon heating, the mass density at the bottom plate will be lower than at the top. It is this mass density difference that, in the presence of gravity, destabilizes the stationary state of the fluid—the heavier fluid will tend to fall down into the lighter. On the other hand, there are viscous effects present that suppress convection. These viscous effects succeed in suppressing convection up to the critical value of ΔT , after which the fluid starts convecting. The Rayleigh-Bénard (we will sometimes abbreviate this to RB) instability can thus be considered the result of a competition between causes promoting opposite effects, a feature frequently encountered in nonlinear physics.

We will try to analyze the RB-instability in more detail using the equations derived in the previous section. An important ingredient in this description will obviously be the variation of the density with temperature, as this is the ultimate cause of the instability. Although a fluid that displays that behavior is clearly compressible, we will get around the extra complications this incurs by using the so-called Boussinesq approximation. In this approximation, we retain the temperature dependence of the density only in a buoyancy term associated with gravity in the external force part of the Navier-Stokes equation, but otherwise assume an incompressible fluid.

We will start with the conductive state, which has $\vec{v} = 0$ everywhere and is stationary, *i.e.* all time derivatives vanish identically. Under these assumptions, we have to solve a simple second order equation for the temperature, supplied of course with the appropriate boundary conditions

$$\begin{cases} \frac{\partial^2 T(z)}{\partial z^2} = 0, \\ T(0) = T_b, \quad T(h) = T_t, \end{cases} \quad (2.3.1)$$

which of course is solved by the linear temperature profile

$$T(z) \equiv T_0(z) = T_b - \frac{\Delta T z}{h}. \quad (2.3.2)$$

We will use the subscript 0 to refer to quantities in the conductive state. All of these are homogenous in x and y , so depend only on z . Using Eq.

(2.3.2), the density profile is readily obtained to lowest order by expanding

$$\rho_0(z) = \rho_0(0) + \frac{\partial \rho}{\partial z} z + \dots, \quad (2.3.3a)$$

$$= \rho_b + \frac{\partial \rho}{\partial T} \frac{\partial T}{\partial z} z + \dots, \quad (2.3.3b)$$

$$\approx \rho_b \left[1 + \alpha \frac{\Delta T z}{h} \right], \quad (2.3.3c)$$

with $\alpha = -\frac{1}{\rho} \frac{d\rho}{dT}$ the thermal expansion coefficient, and again using the density at the bottom ρ_b as a reference point. Using these expressions, we can obtain the pressure field from the Navier-Stokes equation, which now reads

$$\frac{1}{\rho_0(z)} \vec{\nabla} p + g \hat{z} = 0, \quad (2.3.4)$$

From which we find for the equilibrium pressure field

$$p_0(z) = p_b - g \rho_b \left[z + \frac{\alpha \Delta T}{2h} z^2 \right]. \quad (2.3.5)$$

As we have seen, all hydrodynamic fields are known for the conductive state, which we will take as a reference point from here on. Let us now focus on the *deviations* that occur when the fluid is allowed to start moving. In order to do this, we define

$$p(\vec{r}, t) = p_0(z) + \tilde{p}(\vec{r}, t), \quad (2.3.6a)$$

$$\vec{v}(\vec{r}, t) = v_0(z) + \tilde{\vec{v}}(\vec{r}, t), \quad (2.3.6b)$$

$$T(\vec{r}, t) = T_0(z) + \tilde{T}(\vec{r}, t). \quad (2.3.6c)$$

Note that $v_0(z) = 0$. We take the fluid to be incompressible, but account for the variation of density with temperature by writing

$$\begin{aligned} \rho(\vec{r}, t) &= \rho_0(z) + \frac{d\rho}{dT} \tilde{T} + \dots, \\ &\approx \rho_0(z) \left[1 - \alpha \tilde{T}(\vec{r}, t) \right], \end{aligned} \quad (2.3.7)$$

valid to first order in $\alpha \tilde{T}$. For temperature differences above but near the threshold value ΔT_c we do not expect the quantities with tildes to be very large. We can therefore safely substitute Eqs. (2.3.6) and (2.3.7) into the incompressible fluid equations (2.2.22), to arrive finally at the Boussinesq-

equations (we will drop the tildes on \vec{v} and p from now on, and to avoid confusion adopt the notation $\theta = \tilde{T}$)

$$\begin{cases} \vec{\nabla} \cdot \vec{v} & = 0, \\ \frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \vec{v} & = -\frac{1}{\rho_b} \vec{\nabla} p + \nu \nabla^2 \vec{v} + \alpha g \theta \hat{z}, \\ \frac{\partial \theta}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \theta & = \chi \nabla^2 \theta + \frac{\Delta T v_z}{h}. \end{cases} \quad (2.3.8)$$

As always, in order to solve the Boussinesq equations we will need to supply the appropriate boundary conditions. The most realistic boundary conditions are those known as “stick”, where we take the fluid to be completely stationary at the boundaries of the container. Here, we will take so-called “slip” boundary conditions, for which some analytical results can be obtained. We will allow the fluid to slip at the boundaries, but the velocity component perpendicular to the top and bottom plates is taken to be zero. Stick boundary conditions are always closer to the experimental reality, but as they are analytically much harder to incorporate, and in general one has to resort to numerical methods to solve the equations. In addition to the slip boundary conditions, we assume that the temperature at the top and bottom plates is constant, which is experimentally quite feasible.

$$v_z(0) = v_z(h) = 0, \quad (2.3.9)$$

$$\left. \frac{\partial \vec{v}_h}{\partial z} \right|_{0,h} = 0, \quad (2.3.10)$$

$$\theta(0) = \theta(h) = 0, \quad (2.3.11)$$

where we have adopted the notation \vec{v}_h for the horizontal component of the fluid velocity

$$\vec{v}_h = v_x \hat{x} + v_y \hat{y}. \quad (2.3.12)$$

Briefly making a small sidestep from our derivations, we note that the Boussinesq equations (2.3.8) are invariant under the following transformation of the fields

$$\vec{v}_h(x, y, z, t) \rightarrow \vec{v}_h(x, y, h - z, t), \quad (2.3.13a)$$

$$v_z(x, y, z, t) \rightarrow -v_z(x, y, h - z, t), \quad (2.3.13b)$$

$$\theta(x, y, z, t) \rightarrow -\theta(x, y, h - z, t), \quad (2.3.13c)$$

which one can think of as an up-down symmetry (since it relates fields at z to those at $h - z$). In Fig. 2.1 the situation is sketched for the

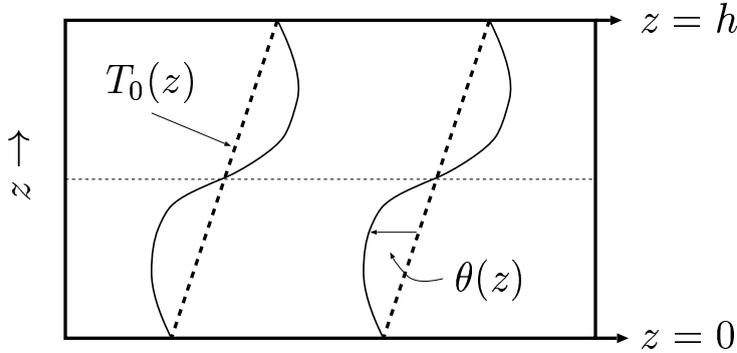


Figure 2.1: Linear temperature profile $T_0(z)$ and the deviation from it $\theta(z)$.

temperature field $\theta(z)$. It should be noted that this symmetry is an artifact of the Boussinesq approximation, and in real life is always (weakly) broken. That this is not without consequences is something we will see later on in this chapter. We now have everything we need to analyze the onset of convection in some detail.

2.4 The linear instability

Before we get deeper into our analysis of the linear instability, it will be convenient to introduce dimensionless variables as follows

$$x \rightarrow xh \quad \theta \rightarrow \frac{\theta\chi\nu}{\alpha gh^3} \quad t \rightarrow \frac{th^2}{\chi} \quad (2.4.1a)$$

$$\frac{p}{\rho_b} \rightarrow \frac{p\chi^2}{h^2} \quad \vec{v} \rightarrow \frac{\vec{v}\chi}{h}. \quad (2.4.1b)$$

Which brings the Boussinesq equations to the following form

$$\vec{\nabla} \cdot \vec{v} = 0, \quad (2.4.2a)$$

$$\frac{1}{\text{P}} \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \vec{v} + \vec{\nabla} p \right] = \nabla^2 \vec{v} + \theta \hat{z}, \quad (2.4.2b)$$

$$\left[\frac{\partial \theta}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \theta \right] = \nabla^2 \theta + \text{R}v_z, \quad (2.4.2c)$$

where we have introduced two important dimensionless numbers, the first being the *Rayleigh number* $\text{R} = \frac{\alpha g \Delta T h^3}{\chi \nu}$, which can be thought of as measuring the ratio of the strength of the destabilizing mechanism, which

is buoyancy which increases with increasing temperature, versus that of the stabilizing mechanisms, *i.e.* viscous relaxation (proportional to ν) and thermal relaxation (χ). The other dimensionless number is $\mathbf{P} = \frac{\nu}{\chi}$, which is the ratio of the thermal and viscous diffusivities. \mathbf{P} is therefore a material constant³, while \mathbf{R} is varied by varying the temperature difference between the plates. We can therefore think of \mathbf{R} as the control parameter in the above set of equations. Our main goal in this section will be to perform a normal mode analysis, which amounts to determining the stability of simple fourier modes in this system. In order to do this, we will first have to linearize the system of equations (2.4.2a), and subsequently derive the *dispersion relation* $\sigma(k)$, which will give us the growth rate σ of a fourier mode with wave number k . If the system does indeed possess an instability, this should be reflected by positive values of the growthrate for some modes.

We can now look at the equation that describes the dynamics of the *vorticity* field, which as we have seen earlier on in this chapter is defined as the curl of the velocity field:

$$\vec{\omega} = \vec{\nabla} \times \vec{v}. \quad (2.4.3)$$

Taking the curl of Eq. (2.4.2b) then yields

$$\frac{\partial \vec{\omega}}{\partial t} - \mathbf{P} \nabla^2 \vec{\omega} + \mathbf{P} \left[\frac{\partial \theta}{\partial x} \hat{y} - \frac{\partial \theta}{\partial y} \hat{x} \right] = (\vec{\omega} \cdot \vec{\nabla}) \vec{v} - (\vec{v} \cdot \vec{\nabla}) \vec{\omega}. \quad (2.4.4)$$

From now on, we will assume the velocity field is small enough to justify linearization, and get rid of the right hand side of the previous equation. When we project what is left onto the z -axis we find for the vertical vorticity the following

$$\frac{\partial \omega_z}{\partial t} = \mathbf{P} \nabla^2 \omega_z, \quad (2.4.5)$$

revealing the essentially diffusive behavior of the vertical vorticity in the linear regime. The fact that this field is completely decoupled from all other fields implies that it is not necessary to take vertical vorticity modes (or equivalently, the *horizontal* components of the velocity) into account in the linear stability analysis, and in keeping with that we will not do so here, and consider only the vertical velocity field v_z . The equation

³For all fluids, \mathbf{P} is typically of order one or larger. Its value for water at 20°C for instance is 6.75.

governing the dynamics of that field is obtained by taking the curl of the linearized Eq. (2.4.2b) *twice* and projecting again on the z -axis, to yield

$$\frac{\partial \nabla^2 v_z}{\partial t} = \mathbf{P} \left[\nabla^4 v_z + \nabla_h^2 \theta \right]. \quad (2.4.6)$$

This equation, supplied with the linearized version of Eq. (2.4.2c)

$$\frac{\partial \theta}{\partial t} = \nabla^2 \theta + \mathbf{R}v_z, \quad (2.4.7)$$

will be the starting point of the actual linear stability calculation. Let us look now at the stability of a fourier-mode with horizontal wave vector \vec{k}_h , by writing

$$v_z(\vec{r}, t) = V(z) e^{i\vec{k}_h \cdot \vec{r}_h + \sigma t}, \quad (2.4.8)$$

$$\theta(\vec{r}, t) = \Theta(z) e^{i\vec{k}_h \cdot \vec{r}_h + \sigma t}, \quad (2.4.9)$$

which should enable us to extract the desired dispersion relation $\sigma(\vec{k}_h)$. When $\sigma(k)$ is positive for some value of \vec{k}_h , one can see from the Ansatz (2.4.8) that that mode will grow exponentially in time. We call such a mode *linearly unstable*. Substituting (2.4.8) into (2.4.6) and (2.4.7), we find

$$\sigma (\partial_z^2 - k^2) V(z) = \mathbf{P} \left[(\partial_z^2 - k^2)^2 V(z) - k^2 \Theta(z) \right], \quad (2.4.10a)$$

$$\sigma \Theta(z) = (\partial_z^2 - k^2) \Theta(z) + \mathbf{R}V(z), \quad (2.4.10b)$$

where we write $\partial_z f = \frac{\partial f}{\partial z}$ (throughout this thesis, we will be using both notations) and $k = |\vec{k}_h|$. We can combine Eqs. (2.4.10) into one by eliminating $\Theta(z)$, to obtain

$$\left[(\partial_z^2 - k^2 - \sigma) \left(\mathbf{P} (\partial_z^2 - k^2)^2 - \sigma (\partial_z^2 - k^2) \right) + \mathbf{R}P k^2 \right] V(z) = 0. \quad (2.4.11)$$

Now let us briefly consider the boundary conditions on $V(z)$. Conditions (2.3.9) translate into

$$V(0) = V(1) = \frac{\partial V(z)}{\partial z} \Big|_{z=0,1} = \frac{\partial^2 V(z)}{\partial z^2} \Big|_{z=0,1} = 0. \quad (2.4.12)$$

The eigenfunctions that obey these boundary conditions are simply

$$V_n(z) = \sin(n\pi z), \quad (2.4.13)$$

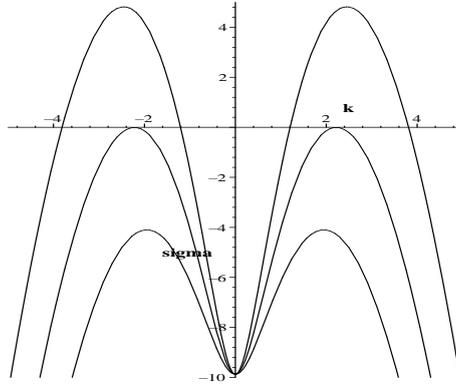


Figure 2.2: Dispersion relation $\sigma(k; R, P)$ as a function of k for $R = 400, 657, 1000$ (lower, middle and upper graph respectively). For $R < R_c$, the instability is absent. The curves are computed from Eq. (2.4.14).

which allows us to determine the dispersion relation for mode n by solving

$$\sigma_n^2 [k^2 + n^2\pi^2] + \sigma_n [(P + 1)(k^2 + n^2\pi^2)^2] + [P(k^2 + n^2\pi^2)^3 - RPk^2] = 0. \quad (2.4.14)$$

which defines for each mode n a dispersion curve $\sigma_n(k; R, P)$. We can locate the instability by looking at the so-called *marginal* modes, which are those that neither grow nor decay, and are therefore characterized by $\sigma_n(k; R, P) = 0$, which, employing Eq. (2.4.14) gives us R_n , the Rayleigh number at which a fourier mode with wavevector k becomes marginal, as

$$R_n(k) = \frac{(k^2 + n^2\pi^2)^3}{k^2}, \quad (2.4.15)$$

from which we read off that the $n = 1$ mode is the first to acquire a nonzero growth rate. In the remainder of this discussion, we will focus on the $n = 1$ mode, and drop the index n . Fig. 2.2 plots $\sigma(k; R, P)$ as a function of k for three values of R , one below, one precisely at and one above the critical Rayleigh number, which is defined as the minimum value of R .

$$R_c \equiv \min_k(R(k)) = \frac{27\pi^4}{4} \approx 657. \quad (2.4.16)$$

The value of the wavevector k at which R attains this minimum is called the *critical* wavevector, and its magnitude shall be denoted by k_c . Note

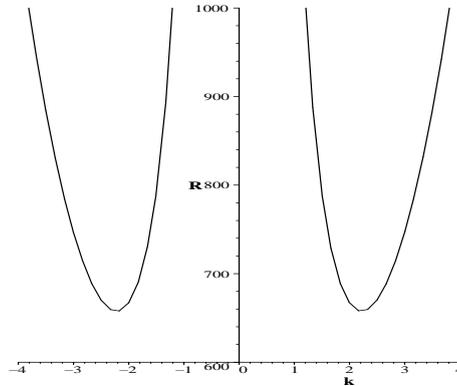


Figure 2.3: Rayleigh number R versus wavevector k . R_c is the minimum of the curve, which occurs at $k = k_c$.

that for unstable modes, $\text{Im } \sigma(k; R, P) = 0$, so that the unstable modes do not oscillate in time.

$$R(k_c) = R_c \quad \Rightarrow \quad k_c = \pm \frac{\pi}{\sqrt{2}} \approx 2.22. \quad (2.4.17)$$

Of course, since the velocity field is a *real* quantity, the system will possess a $k \rightarrow -k$ symmetry. Fig. 2.2 shows the graph of R vs. k as determined by Eq. (2.4.15), with R_c and k_c drawn in. The physical picture emerging from the analysis up to this point is the following. For values of the Rayleigh number below the critical value R_c , all perturbations of the conducting state decay, and we call the stationary state *stable*. For $R > R_c$ however, there exist Fourier modes that do not die down but instead grow exponentially, taking the system further and further away from its initial stationary state and never returning it to this state. For such values of R , we call this system *linearly unstable*. The modes that acquire a positive growth rate all correspond to fluid velocity fields that are periodic in space, and the first mode to go unstable has a finite wavelength.

Although these claims are based on an analysis using the relatively simple slip boundary conditions, detailed numerical work on the more realistic case of stick boundary conditions has revealed that the essential properties listed above do not change. The precise numerical values for quantities like R_c and k_c do however. Actual values, computed numerically

for stick boundary conditions are

$$R_c = 1707.76, \quad (2.4.18a)$$

$$k_c = 3.11632. \quad (2.4.18b)$$

In mathematics, the sudden loss of stability of certain solutions and the appearance of different solutions at that same point is known as a *bifurcation*. Our focus has so far been on the analysis of the linearized equations of motion, but in light of the interesting phenomena *beyond* threshold as such as those encountered in Chapter 1 we would like to get some feel for what happens in the nonlinear regime as well. One way to do this, is to look at the control parameter regime close to threshold, where the nonlinearities are still relatively small (but not negligible!). As we shall see, it is possible to construct an effective, so-called *weakly nonlinear* theory for this regime. We will outline the construction of such a theory in the next section. We choose to demonstrate the procedure on a toy model called the Swift-Hohenberg equation rather than use the Boussinesq equations, in order to ensure the mathematical procedure is clear.

2.5 The Swift-Hohenberg equation

We have seen that a linear analysis of the hydrodynamic equations in an appropriate approximation can give us already a lot of information about the nature of the bifurcation. The most important result of this chapter up to now has been the dispersion relation Eq. (2.4.14). An important demand on the toy model is that its dispersion relation should possess the same characteristics. On the other hand, in order to keep the mathematics straightforward, a simple form for the nonlinearity would be desired. The main features of the dispersion relation that we wish to reproduce are its finite- k maxima, and its $k \rightarrow -k$ symmetry. A fourth order polynomial would fulfill these demands, and a sensible choice for this dispersion relation would be

$$\sigma(k; \varepsilon) = \varepsilon - (k_c^2 - k^2)^2, \quad (2.5.1)$$

Our new dimensionless control parameter ε (which can be thought of as $\frac{R-R_c}{R_c}$ in the context of the RB system, so that the instability is now at $\varepsilon = 0$) shifts the dispersion curve vertically, while we also capture the “double hump” structure with locally quadratic behavior around $k = \pm k_c$.

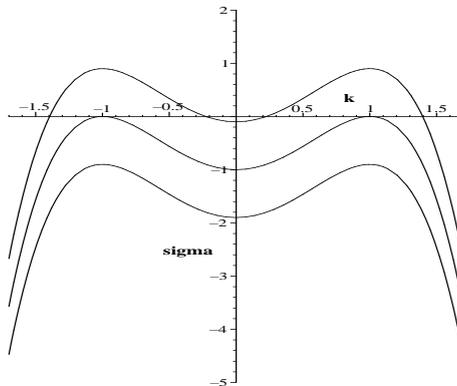


Figure 2.4: Dispersion relation $\sigma(k; \varepsilon)$ for the Swift-Hohenberg equations a function of k for $\varepsilon = -0.7, 0, 0.7$ (lower, middle and upper graph respectively).

Plotted in Fig. 2.4 is the dispersion curve for various values of ε , to be contrasted against Fig. 2.2. We may similarly look at the equivalent of Fig. 2.3, and plot ε_c vs k , which is done in Fig. 2.5. The dispersion relation for our toy model fixes the linear part of the equation we are constructing to be

$$\partial_t u(x, t) = \varepsilon u - (\partial_x^2 + k_c^2)^2 u. \quad (2.5.2)$$

Now, the $z \rightarrow -z$ symmetry (2.3.13a) induced by the Boussinesq approximation is mimicked in our toy model by a $u \rightarrow -u$ symmetry, and therefore the simplest nonlinearity we are allowed to take is a cubic one⁴. The equation thus constructed is known as the Swift-Hohenberg (SH) equation, which we will use as a springboard for further analysis.

$$\partial_t u(x, t) = \varepsilon u - (\partial_x^2 + k_c^2)^2 u - u^3. \quad (2.5.3)$$

The analogues of the rolls found at the onset of convection in the Rayleigh-Bénard system are periodic solutions to the SH equation. We can put in a purely periodic function to show what a nonlinearity might do to solutions

⁴As the name suggests, the Boussinesq approximation is only an approximation, and the induced symmetry is therefore always weakly broken in real life. This can also be simulated by the inclusion of small terms that do not respect the up-down symmetry in our model equation (like u^2). The consequences of the inclusion of such terms on the ensuing amplitude equation will be investigated later on in this chapter.

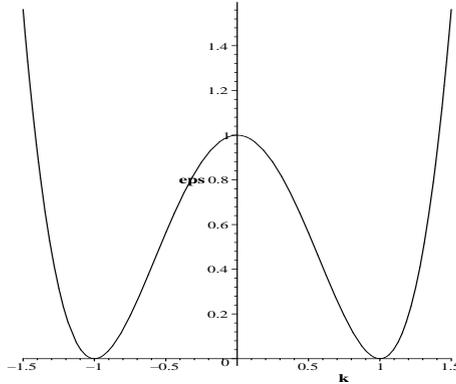


Figure 2.5: Dimensionless control parameter ε versus wavevector k for the Swift-Hohenberg equation. Above the curve, defined by $\sigma(k; \varepsilon) = 0$, the $u = 0$ state is linearly unstable. The merging of the two branches is a feature not present in the Boussinesq equations.

of this type. To this end, set

$$u(x) \sim \cos(kx). \quad (2.5.4)$$

Observe that the nonlinear term produces $\cos^3[kx] = \frac{1}{4} \cos(3kx) + \frac{3}{4} \cos(kx)$. The nonlinearity produces higher harmonics of any periodic function one puts in. These higher harmonics themselves will again produce higher harmonics, leaving us with a veritable zoo of harmonics in the end. This is the reason for the fact that solutions of the SH-equations cannot be written down in what is called a 'closed form' *i.e.* a simple expression in terms of elementary functions. Notwithstanding this fact, fourth order equations such as the SH-equation have received considerable attention in the literature [16], and a lot is known about their solutions, periodic or otherwise. It is important to note though that the higher harmonics are *ordered* in magnitude. To illustrate: if we label the amplitude of the primary wave (2.5.4) by α then the term proportional to $\cos 3kx$ has typical amplitude α^3 , and so on. In the weakly nonlinear regime that we are interested in here, α is a small quantity, and the harmonics are indeed seen to be ordered. Let us now investigate the shape of the dispersion curve in a bit more detail. From now on, we will just consider *positive* values of k . This can be done without loss of generality, as the dispersion curve is symmetric under reflections. As we can see from the linearized dispersion relation, the critical wavenumber is k_c (whose value may be

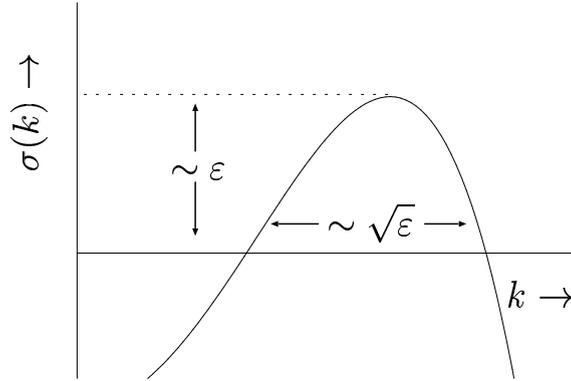


Figure 2.6: Qualitative scaling behavior from the dispersion relation $\sigma(k)$.

chosen at will), and it undergoes a linear instability for $\varepsilon > \varepsilon_c$. For values of ε beyond the threshold, a band of unstable wavevectors develops. The width of this band is proportional to $\sqrt{\varepsilon}$. The maximum is always located at k_c , at which $\sigma(k_c; \varepsilon) = \varepsilon$. For small ε , we therefore have the following inequalities which will be useful later on

$$|k - k_c| \leq \sqrt{\varepsilon}, \quad (2.5.5a)$$

$$\sigma(k_c; \varepsilon) \leq \varepsilon. \quad (2.5.5b)$$

In Fig 2.6, we sketch qualitatively how the scaling behavior can be understood from the dispersion curve. We will need this in the next and final step of the analysis, the amplitude expansion.

2.6 The amplitude expansion

Central to the amplitude approach is the observation that the patterns as they occur above, but still near threshold (*i.e.* for small ε), can be decomposed into a 'fast' and a 'slow' part. Consider a solution to the full SH-equation containing a mode with wavenumber k :

$$u(x, t) \sim ce^{ikx} + c^*e^{-ikx}. \quad (2.6.1)$$

This expression may be rewritten as

$$u(x, t) = (ce^{i(k-k_c)x})e^{ik_cx} + (c^*e^{-i(k-k_c)x})e^{-ik_cx}. \quad (2.6.2)$$

Although this in itself does not tell us much, adding the information obtained at the end of the previous section, namely that $|k - k_c| \leq \sqrt{\varepsilon}$, we see that the first enveloping function $e^{i(k-k_c)x}$ is very slowly varying compared to the rapid $k_c x$ oscillations. In other words, as long as ε is small, the spatial dependence of solutions to the full SH-equation can to good approximation be thought of as that of the critical mode, only slowly modulated. These modulations typically happen on a length-scale $x_{\text{slow space}} = \frac{1}{(k-k_c)} \sim \varepsilon^{-1/2}$, which is the spatial scale on which the slow exponent becomes of $\mathcal{O}(1)$. A similar argument can be made for the temporal dependence of solution near threshold, with the difference that the critical mode has *no* time dependence, as we are dealing with a stationary bifurcation here. This is however strictly valid only *at* threshold, and does not exclude the possibility of dynamics beyond threshold. Since $\sigma(k; \varepsilon) \leq \varepsilon$, the typical *timescale* of this dynamical behavior is expected to be $t_{\text{slow time}} = \frac{1}{\max(\sigma)} \sim \varepsilon^{-1}$. At the heart of the amplitude expansion now is the *separation of scales*, which we can use to our advantage by explicitly separating the temporal and spatial scales through the introduction of slow coordinates X and T , defined by

$$X \equiv \sqrt{\varepsilon}x, \quad (2.6.3a)$$

$$T \equiv \varepsilon t, \quad (2.6.3b)$$

We can now make a weakly nonlinear expansion for the field $u(x, t)$, in writing

$$u(x, t) = \sum_{i=1}^{\infty} \varepsilon^{i/2} U_n. \quad (2.6.4)$$

We assume the parameter ε to be a small quantity, so that the terms in the series are properly ordered. The fact that we expand in powers of $\sqrt{\varepsilon}$, can be understood by considering for instance a Fourier mode with wavenumber q_c near onset. Since we know the mode *at* onset to be stationary, the time dependence is expected to still be very small near threshold. This requires the linear term εu and the nonlinear one u^3 to balance each other approximately; which in turn implies that the amplitude of waves near onset should scale roughly as $\sqrt{\varepsilon}$. It follows that the *bifurcation diagram* should look as roughly sketched in Fig. 2.7(a), for $\varepsilon < 0$ the homogenous solution $u = 0$ is the only stable one, while for positive ε , periodic solutions appear whose amplitude increases continuously, and does so like $\sqrt{\varepsilon}$. This bifurcation diagram is characteristic of a supercritical

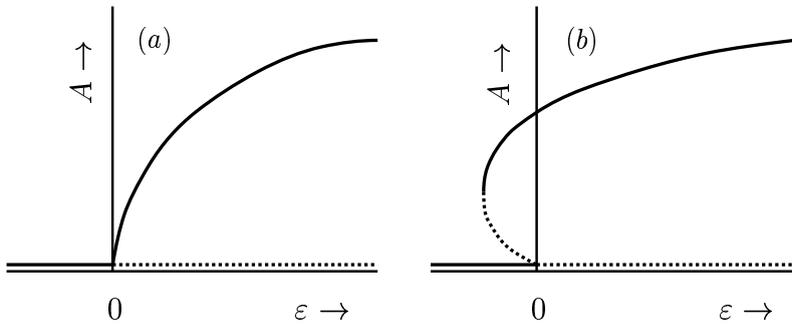


Figure 2.7: Bifurcation diagrams for (a): subcritical and (b): supercritical bifurcations. Stable branches are indicated by solid lines, unstable ones by dotted lines.

bifurcation (also known as a *forward* or *pitchfork* bifurcation). Because the order parameter varies continuously as ε passes through its critical value, the supercritical bifurcation is similar, in a sense, to a second order phase transition. Another possibility for the bifurcation diagram is sketched in Fig. 2.7(b), and is characteristic of the *subcritical* bifurcation. Here, the order parameter immediately jumps to a finite value when the control parameter is increased beyond its critical value. When the control parameter is subsequently *decreased* below the critical value again however, the system will remain on the upper branch up to some ε_1 which is *smaller* than the critical value before jumping back onto the lower branch. A system with a subcritical bifurcation structure is therefore characterized by the presence of *hysteresis*, a phenomenon usually associated with first order phase transitions. The analogy between these bifurcations and phase transitions is however not complete, as we shall see a little later on.

The U_n can be expressed as a product of a slow and a fast part, and in particular,

$$U_1 = A_1(X, T)e^{ik_c x} + A_1^*(X, T)e^{-ik_c x}. \quad (2.6.5)$$

This will indeed come out of the analysis. It might also seem necessary to include the higher harmonics in the amplitude expansion, but as we will see these arise naturally in the expansion. Note that the slow amplitudes A_n are complex quantities, and that they depend on the slow variables only. The functions are $\frac{2\pi}{k_c}$ -periodic in fast space, a fact we will need later on. We shall also demand that the functions are bounded for large and

small x . Let us now look at what Ansatz (2.6.4) implies for the Swift-Hohenberg equation, which we will slightly rewrite for ease of notation as

$$\partial_t u(x, t) = \varepsilon u - (\partial_x^2 + k_c^2)^2 u - u^3 \quad \equiv \varepsilon u - \mathcal{L}u - u^3. \quad (2.6.6)$$

Using the chain-rule, we find that in terms of the fast and slow variables the following replacements need to be made, when working on product functions of the type (2.6.5)

$$\partial_t \rightarrow \varepsilon \partial_T, \quad (2.6.7a)$$

$$\partial_x \rightarrow \partial_x + \sqrt{\varepsilon} \partial_X. \quad (2.6.7b)$$

Note that the small x, t on the LHS are *not* the same as the small x, t on the RHS, the small variables on the left work on all of space while the ones on the right only act on the fast ($\frac{2\pi}{kc}$ -periodic) part of a function. From now on, we will therefore consider the fast and slow variables to be *independent* quantities. The linear differential operator transforms under this change-of-variables as

$$\begin{aligned} \mathcal{L} &\rightarrow \left[\underbrace{(\partial_x^2 + k_c^2)}_{\equiv \mathcal{L}_f} + 2\varepsilon^{\frac{1}{2}} \partial_x \partial_X + \varepsilon \partial_X^2 \right]^2 \\ &= \left[\mathcal{L}_f^2 + 4\varepsilon^{\frac{1}{2}} \mathcal{L}_f \partial_x \partial_X + \varepsilon (2\mathcal{L}_f + 4\partial_x^2) \partial_X^2 + 4\varepsilon^{\frac{3}{2}} \partial_x \partial_X^3 + \varepsilon^2 \partial_X^4 \right], \quad (2.6.8) \end{aligned}$$

where we have given the purely fast part of the linear operator its own name \mathcal{L}_f for future notational convenience. Substituting the amplitude expansion (2.6.4) into this equation brings the SH-equation to the following rather cumbersome, but as we shall see quite potent, form

$$\begin{aligned} \sum_{n=1}^{\infty} \varepsilon^{\frac{n+2}{2}} \partial_T U_n &= \sum_{n=1}^{\infty} \varepsilon^{\frac{n+2}{2}} U_n - \sum_{n=1}^{\infty} \varepsilon^{\frac{n}{2}} \left[\mathcal{L}_f^2 + 4\varepsilon^{\frac{1}{2}} \mathcal{L}_f \partial_x \partial_X + \varepsilon (2\mathcal{L}_f + 4\partial_x^2) \partial_X^2 \right. \\ &\quad \left. + 4\varepsilon^{\frac{3}{2}} \partial_x \partial_X^3 + \varepsilon^2 \partial_X^4 \right] U_n - \sum_{l,m,n=1}^{\infty} \varepsilon^{\frac{l+m+n}{2}} U_l U_m U_n. \quad (2.6.9) \end{aligned}$$

As ε is a small parameter, we will now try to solve this equation order by order. Collecting the various orders in ε up to $\mathcal{O}(\varepsilon^{\frac{3}{2}})$, we find the following

hierarchy of equations

$$\mathcal{O}(\varepsilon^{\frac{1}{2}}) : -\mathcal{L}_f^2 U_1 = 0, \quad (2.6.10a)$$

$$\mathcal{O}(\varepsilon^1) : -\mathcal{L}_f^2 U_2 - 4\mathcal{L}_f \partial_x \partial_X U_1 = 0, \quad (2.6.10b)$$

$$\begin{aligned} \mathcal{O}(\varepsilon^{\frac{3}{2}}) : \partial_T U_1 = U_1 - [\mathcal{L}_f^2 U_3 + 4\mathcal{L}_f \partial_x \partial_X U_2 + \\ + (2\mathcal{L}_f + 4\partial_x^2) \partial_X^2 U_1] - U_1^3. \end{aligned} \quad (2.6.10c)$$

Our task now is to solve these equations recursively, using the results of the previous order to solve the next order. Eq. (2.6.10a) determines the linearized solution [cf. Eq.(2.6.5)]⁵:

$$\mathcal{L}_f U_1 = 0 \quad \rightarrow \quad U_1 = A_1(X, T) e^{ik_c x} + A_1^*(X, T) e^{-ik_c x}, \quad (2.6.11)$$

where the (complex) function $A_1(X, T)$ is still completely arbitrary, since \mathcal{L}_f works only on the fast scales. We should point out here however that the translational invariance in our original SH-equation translates into an invariance of A under phase rotations, a symmetry that should obviously be conserved throughout the argument

$$A \rightarrow A e^{i\varphi}, \quad (2.6.12)$$

which corresponds to a translation of the entire pattern by a distance $\frac{\varphi}{k_c}$. While the $\mathcal{O}(\varepsilon)$ equation (2.6.10b) yields no information about the slow-scale dynamics

$$\mathcal{L}_f U_2 = 0 \quad \rightarrow \quad U_2 = A_2(X, T) e^{ik_c x} + A_2^*(X, T) e^{-ik_c x}, \quad (2.6.13)$$

the third equation does produce an interesting result. It can be cast into the following form

$$\mathcal{L}_f^2 U_3 = \left[e^{ik_c x} \{ -\partial_T + 1 + 4k_c^2 \partial_X^2 - 3|A_1|^2 \} A_1 + c.c. \right] - \left[e^{3ik_c x} A_1^3 + c.c. \right]. \quad (2.6.14)$$

The operator \mathcal{L}_f is linear, and acts on the fast scales only, and the RHS of Eq. (2.6.14) is of the form $F(X, T) e^{ik_c x} + G(X, T) e^{3ik_c x} + c.c.$. This implies that the fast space dependence of U_3 is also a sum of these exponentials,

⁵Note that actually, the first order demands that $\mathcal{L}_f^2 U_1 = 0$, and not $\mathcal{L}_f U_1 = 0$. Although the former implies the latter, the 'squared' equation also possesses solutions of the type $U_1 \sim x A e^{ik_c x} + c.c.$. These are however neither bounded nor periodic, and are therefore discarded here.

and this is where, as promised, the higher harmonics enter the expansion naturally

$$U_3 = A_3(X, T)e^{ik_c x} + B_3(X, T)e^{3ik_c x} + c.c.. \quad (2.6.15)$$

Note that the amplitude expansion for this particular nonlinearity generates *in principle* not only odd, but also higher order even harmonics. The amplitudes of these will however turn out to be zero, so we need not include them. This is different if we include a term that breaks the $u \rightarrow -u$ symmetry u^2 in our SH model. So far however, we still do not know anything about the dynamics on the slow scale, and for this we need to use a theorem due to Fredholm. It states that for a general operator \mathcal{L} , the equation

$$\mathcal{L}u = v \quad \mathcal{L} : S \rightarrow S \quad (2.6.16)$$

is solvable if and only if the vector $v \in S$ is orthogonal to the kernel $\text{Ker}[\mathcal{L}^\dagger] = \{w \in S | \mathcal{L}^\dagger w = 0\}$ of the adjoint operator \mathcal{L}^\dagger :

$$\langle w | v \rangle = 0 \quad \forall w \in \text{Ker}[\mathcal{L}^\dagger]. \quad (2.6.17)$$

The space S that we are working in here is the space of functions that are $\frac{2\pi}{k_c}$ -periodic in fast space (nothing is said about the slow dependence), and an appropriate inner-product is therefore

$$\langle a | b \rangle = \int_0^{\frac{2\pi}{k_c}} dx a^* b. \quad (2.6.18)$$

The kernel of the fast operator (which, fortunately, is self-adjoint) \mathcal{L}_f is determined by

$$\mathcal{L}_f w = 0 \quad \rightarrow \quad w = C e^{ik_c x} + c.c., \quad (2.6.19)$$

and an application of the Fredholm theorem to Eqs. (2.6.14) and (2.6.15) produces the dynamical equation for the slow amplitude $A_1(X, T)$

$$\partial_T A_1(X, T) = A_1 + 4k_c^2 \partial_X^2 A_1 - 3|A_1|^2 A_1. \quad (2.6.20)$$

This is known as a *solvability condition*, and the technique is a very useful one indeed. It will be used again in a very different context in Section 4.5. The manner by which the Fredholm theorem produces this solvability condition is perhaps illustrated best by considering its action on Eq. (2.6.14). The first term, proportional to $e^{iq_c x}$ is what is called *resonant*: upon multiplication with the (conjugated) zero modes of \mathcal{L}_f , it produces

a term that is independent of the fast scales. While all other nonresonant terms necessarily yield zero when integrated over one fast period, it is by demanding that the resonant ones also vanish that we actually obtain the solvability condition. Eq. (2.6.20) is exactly what we set out to obtain: an equation for the dynamics of the slow modulations, valid for small ε . Such an equation is what is called an *amplitude equation*. Of course, one also obtains an equation for A_1^* , which is just the complex conjugate of the above equation. We have derived now that the dynamics of the amplitude A_1 is governed by a second order, nonlinear PDE. When one has obtained a solution to this equation which is often the hard part, the amplitude of the third harmonic $B_3(X, T)$ is also known. Combining Eq. (2.6.14) with (2.6.15), we find

$$B_3(X, T) = \frac{A_1^3}{8k_c^2}, \quad (2.6.21)$$

which implies that the amplitude of the higher harmonics is what is called *slaved* by the A_1 . This holds for all higher harmonics. We will comment more on this slaving further on. The fact that ε does not appear in our equation shows that we have chosen the correct scales of space, time and amplitude. We prefer however to keep ε explicit, in order to get a good idea of what happens as it is increased through zero and to avoid control parameter-dependent rescaling. After the appropriate transformations of space, time and amplitude, we arrive finally at the amplitude equation in its standard form

$$\partial_t A = \varepsilon A + \partial_x^2 A - |A|^2 A, \quad (2.6.22)$$

where we have dropped the subscripts, and transformed back to the fast variables. We can now also consider the higher order terms in the expansion, and along similar lines one can show that the dynamical equation for $A_2(X, T)$ reads

$$\partial_T A_2 = [1 - 6|A_1|^2] A_2 + 4k_c^2 \partial_X^2 A_2 - 4ik_c \partial_X^3 A_1 - 3A_1^2 A_2^*. \quad (2.6.23)$$

What sets this equation apart from the one for $A_1(X, T)$ is that is completely *linear*, and depends solely on $A_1(X, T)$. Formally, once one has solved Eq. (2.6.20), that solution completely determines A_2 , yet another example of slaving. Since *both* the higher order terms and the higher harmonics are slaved to the dynamics of $A_1(X, T)$, it makes sense to say that all physical information is encoded in A_1 , since all other amplitudes are driven by it.

Returning for a moment to the $u \rightarrow -u$ symmetry, we will briefly explore the consequences of the *breaking* of this symmetry for the ensuing amplitude equations. In order to break the symmetry, we could for instance modify the Swift-Hohenberg equation by introducing a quadratic nonlinearity as follows

$$\partial_t u(x, t) = \varepsilon u - (\partial_x^2 + k_c)^2 u + \alpha u^2 - \beta u^3. \quad (2.6.24)$$

Carrying out the entire program for this equation produces in the end the lowest-order amplitude equation

$$\partial_T A_1 = A_1 + 4k_c^2 \partial_X^2 A_1 - 3 \left(\beta - \frac{38\alpha^2}{27k_c^4} \right) |A_1|^2 A_1, \quad (2.6.25)$$

from which we read off that for sufficiently small values of α , the symmetry breaking term only renormalizes the coefficient of the nonlinearity. Although it might seem surprising at first that the effect of a quadratic nonlinearity affects the *cubic* term in the amplitude equation, this is nonetheless correct. Considering again the *resonant* terms only, one sees that although the square nonlinearity produces terms proportional to $e^{2ik_c x}$, which are of order ε , these are never resonant by themselves, as a single multiplication with $e^{ik_c x}$ can never produce an expression independent of the fast scales. The lowest order combination involving the square nonlinearity to do so is $e^{2ik_c x} \times e^{-ik_c x}$, which is of order $\varepsilon \times \sqrt{\varepsilon}$, the same order of magnitude as the cubic term in the original SH-equation. Upon closer inspection, we find that when

$$\alpha^2 > \frac{27\beta k_c^4}{38}, \quad (2.6.26)$$

the coefficient of the cubic term becomes positive, and the nonlinearity is no longer saturating, causing any solutions to grow exponentially without bounds. This is of course an unphysical result, and in order to account for this we will need to include higher order terms in the amplitude equation (e.g. $|A_1|^4 A_1$). It is known that the inclusion of such terms will render the bifurcation *subcritical*. This is problematic, since the fact that the order parameter is finite immediately at the onset of the instability is in apparent contradiction with one of the main assumptions in deriving the amplitude equation that there exists a weakly nonlinear expansion of the order parameter close to threshold.

Equation (2.6.22) has the same form as the Ginzburg-Landau equation for superconductivity in the absence of a magnetic field, which was discovered long before this one. Newell and Whitehead [17], who first derived it in the context of nonlinear hydrodynamics in 1969, did therefore not get the equation named after them, and instead it is usually known as the Real Ginzburg-Landau equation. Note that the term 'real' here does not refer to the amplitude, which is a complex quantity, but to the fact that the coefficients are real.

This concludes the derivation of the amplitude equation for the Swift-Hohenberg equation. Although the equation itself was introduced as a toy model, the procedure is essentially the same for the equations governing the Rayleigh-Bénard system (one obvious difference being the non-trivial modes in the z -direction in Rayleigh-Bénard). One separates the space- and time dependence into a slow and a fast part, and makes an Ansatz like Eq. (2.6.4) for solutions $u(x, t)$ to the full nonlinear problem. The resulting series of equations is solved order by order, the first one determining the linear solution, the second one the second order term and the third one yielding as a solvability condition the amplitude equation for the $A_1(X, T)$. The nice thing about this amplitude expansion is that for a stationary, forward bifurcation the ensuing equation is *always* of the form

$$\tau_0 \partial_t A = \varepsilon A + \xi_0^2 \partial_x^2 A - g_0 |A|^2 A, \quad (2.6.27)$$

in which the coefficients τ_0 , ξ_0 and g_0 reflect the physical properties of the actual system under study. The *correlation length* ξ_0 and *correlation time* τ_0 can be derived from the linear dispersion relation as follows

$$\frac{1}{\tau_0} = \left. \frac{\partial \sigma(k)}{\partial \varepsilon} \right|_{k_c}, \quad \xi_0^2 = - \left. \frac{\tau_0}{2} \frac{\partial^2 \sigma(k)}{\partial k^2} \right|_{k_c}. \quad (2.6.28)$$

Using this, one can substitute results obtained numerically (or even experimental data) for these coefficients into the Ginzburg-Landau equation, while forgetting about all other hydrodynamic complications, and still get a good description of the pattern dynamics near threshold. As an example, if we calculate the correlation length and time for the Boussinesq equations, supplied with slip boundary conditions, we find

$$\frac{1}{\tau_0} = \frac{2}{3\pi^2} \frac{1+P}{P}, \quad \xi_0^2 = \frac{8}{3\pi^2}. \quad (2.6.29)$$

2.7 Implications of the amplitude description

Now that we have obtained the equation that should adequately describe the dynamics of the roll pattern in a Rayleigh-Bénard, we will take a closer look at what exactly it is this equations predicts will happen beyond threshold. The RGLE (2.6.22) admits spatial plane-wave solutions of the form

$$A = a_0 e^{iqx}. \quad (2.7.1)$$

It should be noted that we will use q to denote the wavenumber from now on. Such solutions are also referred to as phase-winding solutions, because when one plots a function of the form (2.7.1) in 3D like $(\text{Re}(A), \text{Im}(A), x)$, the result will be a 'corkscrew' or helical curve. In terms of solutions to the Swift-Hohenberg equation, they describe stationary periodic patterns with wavenumbers slightly above ($q > 0$) or below ($q < 0$) q_c . As we have seen in the previous chapter, periodic solutions exists within a band around q_c , and upon substitution of (2.7.1) in (2.6.22) we find that

$$a_0 = \varepsilon - q^2 \quad \Rightarrow \quad |q| \leq \sqrt{\varepsilon} \quad (2.7.2)$$

The appearance of periodic solutions with wavenumbers within a certain band is thus recovered, also on the level of the amplitude equations. This is what makes an instability like the Rayleigh-Bénard instability very much different from phase transitions. Whereas phase transitions usually occur between two well-defined, but most importantly *unique* states, the $A = 0$ solution here loses stability in favor of not one, but instead a *continuous family* of phase-winding solutions. In actual realizations of the system, only one member of this family will usually be present, and therefore in addition to an instability mechanism we will need to uncover a *selection mechanism* to be able to say something about the asymptotic state we expect the system to end up in. The existence of certain types of solutions however says nothing about their stability, and it is precisely the states that are stable that are likely to be reached from physical initial conditions. Therefore, let us investigate the linear stability of the phase winding solutions. If we first split the (complex) amplitude A into modulus and phase (both real)

$$A(x, t) = a(x, t) e^{i\varphi(x, t)}, \quad (2.7.3)$$

and substitute this in Eq. (2.6.22), we arrive at the equivalent set of equations

$$\partial_t a = \varepsilon a + \partial_x^2 a - a^3 - a(\partial_x \varphi)^2, \quad (2.7.4a)$$

$$a \partial_t \varphi = 2(\partial_x a)(\partial_x \varphi) + a \partial_x^2 \varphi. \quad (2.7.4b)$$

Note the absence of any dependence on the phase φ itself. This was to be expected, since translational invariance requires the system of equations to be invariant under $\varphi \rightarrow \varphi + c$. The phase-winding solution is characterized by $a = a_0$, $\varphi = qx$, and as a starting point for our stability analysis we study, as usual, the response to periodic perturbations :

$$a = a_0 + \delta a e^{iQx + \lambda t}, \quad (2.7.5a)$$

$$\varphi = qx + \delta \varphi e^{iQx + \lambda t}. \quad (2.7.5b)$$

When we now substitute (2.7.5a) in Eqs. (2.7.4a), and linearize in δa , $\delta \varphi$, we find the following matrix equation for the perturbations

$$\lambda \begin{pmatrix} \delta a \\ \delta \varphi \end{pmatrix} = \mathbf{D} \cdot \begin{pmatrix} \delta a \\ \delta \varphi \end{pmatrix}. \quad (2.7.6)$$

Solving this eigenvalue problem yields the secular equation for the growth-rate λ

$$\lambda^2 + 2\lambda(Q^2 + a_0^2) + Q^2(2\varepsilon - 3q^2) + Q^2 = 0. \quad (2.7.7)$$

In order to be able to determine whether a particular mode is stable or not, we only need the *signs* of the two roots of this equation. These signs are determined by the sign of the $(\varepsilon - 3q^2)$ term, and it is easy to show that the system is stable for $q^2 < \frac{\varepsilon}{3}$, and possesses one unstable mode when $q^2 > \frac{\varepsilon}{3}$. The situation is sketched in Fig. 2.8. We can therefore conclude that *within* the band of allowed wavenumbers, there is a smaller band of phase winding solutions that are actually stable. This instability is a long-wavelength one, in that it occurs first for small Q . It is generally known as the Eckhaus instability [18]. It is important to note at this point that this is, in fact, a *secondary* instability, since the mode that loses stability (the phase winding solution) *itself* originated from an instability of the Swift-Hohenberg equation, the primary one.

2.8 The complex Ginzburg-Landau equation

We have seen that the real GL equation is the one generically describing a stationary, supercritical bifurcation, *i.e.* one that is characterized by the

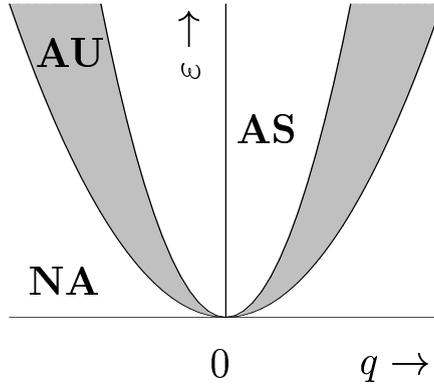


Figure 2.8: The bands of allowed stable (AS), allowed but unstable (AU) and not-allowed (NA) wavenumbers.

fact that the critical frequency is zero at onset. There is however also a large class of systems that does not show this behavior, and instead has nonzero critical frequency. This means that the pattern formed at onset is periodic not only in space, but also in time – such systems possess *traveling wave* solutions. Provided the bifurcation is supercritical, a weakly nonlinear analysis completely analogous to the one developed in the previous sections reveals that these systems also have a generic amplitude equation to describe them, an equation that is known as the Complex Ginzburg-Landau equation or CGLE⁶

$$\partial_t A = \varepsilon A + (1 + ic_1)\partial_x^2 A - (1 - ic_3)|A|^2 A. \quad (2.8.1)$$

Although its form is very reminiscent of the RGLE, the dynamical behavior of this equation is completely different. The physical mechanism that is introduced by the addition of complex coefficients is *dispersion*, which is most clearly illustrated by substituting a traveling wave solution $a_0 e^{i(qx - \omega t)}$ into Eq. (2.8.1), to yield

$$\omega = c_1 q^2 - c_3 a_0^2, \quad (2.8.2)$$

$$q^2 = \varepsilon - a_0^2. \quad (2.8.3)$$

⁶In principle, one should also include a complex coefficient for the linear term, $\varepsilon(1 + ic_0)A$. This can however be removed by redefining $\tilde{A} = e^{-i\varepsilon c_0 t}$ in a co-rotating frame. Furthermore, one might expect an advective (corresponding to an overall group velocity) term proportional to $\partial_x A$. This can also be removed here by a Galilean transformation to a co-moving frame. In later Chapters, we will see examples where this is no longer possible and one needs to include such terms.

The coefficient c_1 measures the linear dependence of the frequency on the wavenumber, while c_3 couples the wavenumber and the amplitude. A very instructive way of illustrating the essential difference between the RGLE and the CGLE is by means of a variational formulation of the problem. The Swift-Hohenberg equation has the interesting property that it can be derived from what is known as a Lyapunov functional, in that

$$\partial_t u = -\frac{\delta \mathcal{F}_{\text{SH}}[u]}{\delta u}, \quad (2.8.4)$$

where $\frac{\delta}{\delta u}$ denotes the functional derivative. The explicit form of \mathcal{F}_{SH} is given by

$$\mathcal{F}_{\text{SH}}[u] = \frac{1}{2} \int dx \left\{ [(\partial_x^2 + q_c^2)u]^2 - \varepsilon u^2 + \frac{1}{2} u^4 \right\}. \quad (2.8.5)$$

From this, a very useful property of the SH-equation follows, namely that its dynamics can never increase \mathcal{F}_{SH} :

$$\frac{d}{dt} \mathcal{F}_{\text{SH}} \leq 0. \quad (2.8.6)$$

The behavior of the Lyapunov functional is very reminiscent of the free energy in thermodynamics, in that its minima correspond to equilibrium states of a given system. The thermodynamic free energy is however a function because one looks at homogenous phases, as opposed to the Lyapunov functional. Because of the similarity, the Lyapunov functional is sometimes also referred to as the Free Energy functional. When one is interested in the states that the SH-system is likely to reach from relevant initial conditions, the 'gradient dynamics' (dynamics in the direction of decreasing $\mathcal{F}_{\text{SH}}[u]$) implies that these states correspond to minima of $\mathcal{F}_{\text{SH}}[u]$. Because the dynamics is very thermodynamics-like, it is also called relaxational. In particular, because there is something like a free energy to minimize, we will not find chaos in the SH-equation. Note that we cannot derive the precise temporal evolution of the SH-system from the Lyapunov functional description, since requiring a quantity to decrease does not give any information on how exactly it does so, and there may indeed be some dynamics that leave $\mathcal{F}_{\text{SH}}[u]$ invariant. It will come as no surprise that the amplitude equation formalism does not break the variational structure of the underlying equation, and that the RGLE can be derived from its own Lyapunov functional, which reads

$$\mathcal{F}_{\text{GL}}[A] = \int dx \left[|\partial_x A|^2 - \varepsilon |A|^2 + \frac{1}{2} |A|^4 \right]. \quad (2.8.7)$$

The RGLE can therefore never display chaotic behavior. This is where the main difference with the CGLE lies, since it turns out that the CGLE cannot be derived from a variational principle unless $c_1 = -c_3$ [19]. The lack of such a principle opens up the possibility of having dynamics that are non-relaxational, *i.e.* go on indefinitely without tending to some well defined asymptotic state. We call such dynamics *persistent*, and in its most dramatic form it is known as chaos. As we have seen, in the limit $c_i \rightarrow 0$, the CGLE tends to the RGLE, which does possess a variational structure. In the opposite limit, that of very large imaginary coefficients, we can rewrite it as follows

$$i\partial_t A = [-c_1 \partial_x^2 - c_3 |A|^2] A, \quad (2.8.8)$$

which is known as the *nonlinear Schrödinger equation*, which is not only Hamiltonian but also integrable and has been widely studied, particularly for the solitonic solutions it is known to possess. One way of bringing about chaos in the CGLE can be understood by looking at the stability of phase winding solutions in the CGLE. A similar derivation as the one for the RGLE⁷ then reveals a similar phenomenon: within the band of allowed wavevectors $|q| < \varepsilon$ there is smaller band of actually stable solutions. In the complex case however, the width of this band depends on the value of the coefficients c_1 and c_3 . In the long-wavelength limit, one finds

$$q^2 \leq \frac{\varepsilon(1 - c_1 c_3)}{3 - c_1 c_3 + 2c_3^2}, \quad (2.8.9)$$

from which we read off that the size of the stable band shrinks down to zero when

$$c_1 c_3 > 1. \quad (2.8.10)$$

This is known as the Newell criterion, and it effectively divides the plane (c_1, c_3) into two parts; in the part *below* the Newell line the CGLE possesses stable phase winding solutions, while all plane waves are unstable *above* it. Consider now the situation where we prepare a system governed by the RGLE in a plane wave state with a wavenumber q outside of the stable band. We know the RGLE dynamics to be relaxational, and consequently the system somehow has to end up in the stable band again. For it to do so in a finite system requires the phase to change in a discontinuous manner, since in order to change its wavenumber the quantity

⁷Only for *traveling* phase winding modes $A = ae^{i(qx - \omega t)}$.

$\int dx \varphi(x)$, the equivalent of a topological winding number in this system, has to change. The winding number can only change by discrete multiples, and therefore getting back into the stable band amounts to losing (or gaining) 2π in total phase. This however can only happen when the *amplitude* is zero, as at such a point the phase is not defined. As it turns out, this is precisely the way in which the RGLE achieves the return to the stable band, and such an event (strongly located in both space and time) is called a phase slip, and is one of the most commonly encountered examples of a *topological defect*. Topological defects in general are characterized by the presence of a core region, where order is destroyed, and a far-field region where variables change only slowly in space. One can for instance consider a two-dimensional pattern for which the integer valued integral

$$\frac{1}{2\pi} \oint_C \vec{\nabla} \varphi \cdot \vec{dl}, \quad (2.8.11)$$

is equal to 1. This is clearly not consistent with a smoothly varying phase everywhere inside the closed contour C . To see this, shrink the contour down to infinitesimal size in a smooth manner. The value of the integral should then also vary in a smooth manner, and since it is an integer-valued quantity, that implies it should remain constant, even for arbitrarily small contours. This indicates that there should be at least one discontinuous point within the contour. The existence of such topological defects is closely related to the presence of broken continuous symmetries, but we will not go into that in detail here [20]. It is well-known that the RGLE is not the only system that can phase-slip, the CGLE does so as well, and in fact it is what unstable modes generically do. If we now consider the CGLE above the Newell line, we can see how this would generate chaos: all modes are unstable and will therefore phase slip, altering the wavenumber by some amount, but the new wavenumber is unstable as well and itself slips elsewhere in the band of allowed wavevectors. In this manner, the system will explore the whole band over and over without ever finding stability. Note that although the wavenumber has up to now been defined only as a *global* quantity, it is also possible to define a *local* wavenumber as the spatial derivative of the phase at some point. In essence, this wavenumber behaves in the same way as the global wavenumber does, and for traveling wave solutions the two definitions coincide. In the following chapter, we will see many examples of chaotic behavior in CGL-like equations.

2.9 Coherent structures

An important role throughout this thesis will be played by solutions to nonlinear equations such as the CGLE known as *coherent structures*. We will use this term to describe structures that are either themselves localized, or consist of domains of regular patterns connected by localized defects, that do not change shape over time and move at constant velocities. Because we will at times allow them to oscillate uniformly though. The most general form of such a solution will therefore be

$$A(x, t) = e^{-i\omega t} \hat{A}(x - vt). \quad (2.9.1)$$

Although this might not seem like too much of a simplification as compared to solving the full nonlinear PDE, the fact that the structure moves at a constant velocity means that by considering the equation under study in the frame co-moving with the structure's velocity, our coherent structure will manifest itself as a stationary solution, which turns the PDE into an ODE, and those are in general much easier to handle than PDE's. An additional bonus is that one can now employ the well-equipped toolbox of ODE research, of which especially phase space methods will prove very useful. Coherent structures are however interesting for many more reasons apart from the fact that they are relatively easy to study. There exist many regions of CGLE parameter space in which it is known that coherent structures organize much of the dynamical behavior of the system, and can in a sense be thought of as the building blocks of such dynamical states. Knowledge of the particular type of coherent structure one expects to find, as well as their interactions for a given set of parameter values can be a valuable tool in analyzing the actual dynamical states one encounters. Especially in the context of the CGLE, coherent structures and their interactions have been the subject of intense study. In [21], an attempt was made to classify the structures according to the topology of their orbits in phase space. Basically, four different types of structures were found in this way for a single CGLE. Those four are sources, sinks, fronts and pulses, and they are sketched in Fig. 2.9. In subsequent studies, a fifth relevant structure called the *homoclinic hole* was discovered [22]. All of these coherent structures play an important part in various regions of parameter space.

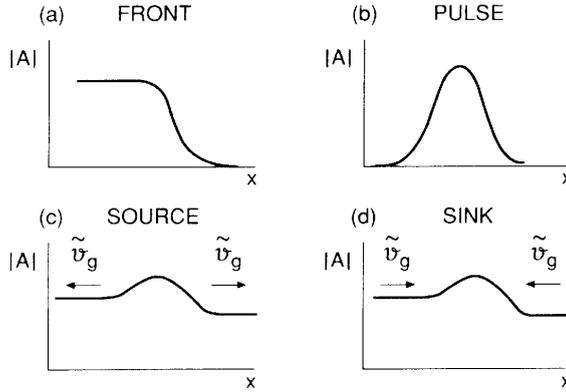


Figure 2.9: Classification of coherent structures in [21].

2.10 Amplitude equations and symmetries

Up to now in our discussion, we have frequently encountered certain symmetries, and seen how these are reflected in the resulting amplitude equation. One can also skip the formal derivation, and basically 'guess' the terms of the amplitude equation from basic symmetries and scaling considerations. The assumption underlying such a construction is essentially that *any term that is not forbidden by symmetries will be present*. Collecting such terms to lowest order (which is where the scaling comes in), then produces an educated, and in most cases correct, amplitude equation. As an example, let us look in some more detail at the hexagonal pattern in 2 dimensions, such as the one depicted in Fig. 1.2 in Chapter 1. As a closer inspection of the pattern reveals, it can be thought of as a superposition of three separate roll patterns, that are at 120° angles to each other. In an amplitude-type description of such a pattern, we would accordingly split of three separate slowly varying amplitudes by writing

$$\text{fields} \sim A_1 e^{i\vec{k}_1^c \cdot \vec{r}} + A_2 e^{i\vec{k}_2^c \cdot \vec{r}} + A_3 e^{i\vec{k}_3^c \cdot \vec{r}}, \quad (2.10.1)$$

and the different \vec{k}_i^c 's obey

$$\sum_{i=1}^3 \vec{k}_i^c = 0, \quad (2.10.2a)$$

$$|\vec{k}_1^c| = |\vec{k}_2^c| = |\vec{k}_3^c|. \quad (2.10.2b)$$

A pattern in which all three amplitudes A_i , $i = 1 \dots 3$ are equal would correspond to a perfect hexagonal pattern. Translation invariance in the \vec{k}_1^c -direction now would require the fields be invariant under $\vec{r} \rightarrow \vec{r} + \alpha \vec{k}_1^c$, or

$$\begin{aligned} A_1 e^{i\vec{k}_1^c \cdot \vec{r}} + A_2 e^{i\vec{k}_2^c \cdot \vec{r}} + A_3 e^{i\vec{k}_3^c \cdot \vec{r}} = \\ A'_1 e^{i\alpha |\vec{k}_1^c|^2} e^{i\vec{k}_1^c \cdot \vec{r}} + A'_2 e^{i\alpha \vec{k}_1^c \cdot \vec{k}_2^c} e^{i\vec{k}_2^c \cdot \vec{r}} + A'_3 e^{i\alpha \vec{k}_1^c \cdot \vec{k}_3^c} e^{i\vec{k}_3^c \cdot \vec{r}}. \end{aligned} \quad (2.10.3)$$

If we now define $\varphi = \alpha |\vec{k}_1^c|^2$, we see that translation invariance in the \vec{k}_1^c -direction requires invariance of the three amplitudes under the transformation

$$A_1 \rightarrow A_1 e^{i\varphi}, \quad A_2 \rightarrow A_2 e^{-i\frac{\varphi}{2}}, \quad A_3 \rightarrow A_3 e^{-i\frac{\varphi}{2}}. \quad (2.10.4)$$

We know that the amplitudes A_i are of $\mathcal{O}(\varepsilon)$, that $x \sim \sqrt{\varepsilon}^{-1}$ and that $t \sim \varepsilon^{-1}$. Furthermore, we require reflection symmetry in the \vec{k}_1^c -direction. If we now write down which terms are allowed to lowest nontrivial order (which is, as we have seen, $\mathcal{O}(\varepsilon^{\frac{3}{2}})$), we end up with

$$\partial_t A_1 = (\vec{e}_1 \cdot \vec{\nabla})^2 A_1 + \varepsilon A_1 \pm C_2 A_2^* A_3^* + [\text{cubic terms}]. \quad (2.10.5)$$

The equations for A_2 and A_3 can be obtained by cyclic permutation of the indices. To lowest order, Eq. (2.10.5) is the only amplitude equation for hexagonal patterns that respects all the symmetries and has the correct scaling behavior. The appearance of quadratic terms in this equation is important, as amplitude equations with quadratic terms undergo a *subcritical* bifurcation instead of a supercritical one. What this symmetry argument is telling us therefore is that *any* system with the above symmetries should display the hexagonal pattern even before the actual convection threshold, which is indeed what Bénard already discovered. One notable symmetry that was not included in our argument is the up-down symmetry, roughly associated with the Boussinesq approximation. When we also incorporate it, the quadratic term drops out and we are left with the supercritical bifurcation we derived earlier. The subcritical appearance of hexagons 'in real life' can therefore be thought of as a direct consequence of the weakly broken Boussinesq symmetry.

In one dimension we can play the same game, and indeed one can show that the RGLÉ is the generic equation for a reflection and up-down

symmetric system that possesses translational invariance. This goes to illustrate once again that there can be a big payoff if one carefully considers the essentials of the system one plans to study, before starting any detailed calculations. In Chapter 3, we will use such symmetry considerations to derive the appropriate amplitude equation in a slightly more complicated system, which allows us to skip hydrodynamics altogether and get straight to the pattern formation.

Sources and sinks in traveling wave systems

3.1 Introduction

In this chapter, we will apply some of the concepts developed in Chapter 2 to analyze the pattern formation in a convection experiment known as the *heated wire*, using the amplitude approach. The focus of this analysis will be the properties of two types of coherent structures, *sources* and *sinks*. The main objective of this research was to provide the experimentalists working on this experiment and comparable setups with qualitative predictions for experimentally accessible quantities. Wherever this is possible, we will try to make contact with those experiments to see whether our predictions are in fact accurate.

3.2 The heated wire experiment

In the heated wire experiment, a thin wire is suspended below the free surface of a liquid. A variable power Q is sent through this wire, resulting in an increase of its temperature. In Fig. 3.1, the setup is sketched. Apart from the obvious similarities to the Rayleigh-Bénard system discussed in Chapter 2, there are also some different features, which are not without consequences. The main *difference* is that this problem involves a free

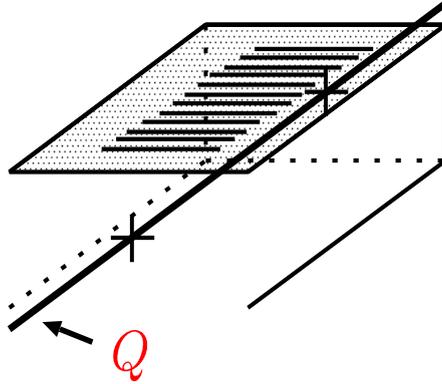


Figure 3.1: Schematic setup of the heated wire experiment. A thin wire is suspended below (typically $\sim 2\text{mm}$) the free surface of a liquid. The wire is heated by a power Q . When this Q is increased beyond some critical value, traveling waves appear at the surface. These waves travel in the direction along the wire.

liquid surface, whereas in Rayleigh-Bénard¹ we were dealing with a closed volume of fluid. Free surfaces are notoriously difficult to deal with in hydrodynamics, especially when temperature gradients are involved. One has to take into account deviations from a flat surface, as well as local variations in the surface tension due to the temperature gradients (the so-called Marangoni effect). We will address these issues in a moment, but let us first concentrate on the phenomenology of the heated wire experiment.

As opposed to the case of RB-convection, in the heated wire convection sets in immediately upon turning on the heating. The principal mode of convection consists of two relatively small rolls that cause the surface to bulge slightly on either side of the wire. Upon increasing the temperature further however, at a certain onset value of the heating power *traveling waves* appear. These waves are traveling in the direction along the wire, and they can travel either to the left or to the right. Both directions occur

¹The *modern* Rayleigh-Bénard experiment, that is. As we have seen in Chapter 1, Bénard's original setup also had a free surface.

and, more importantly, both occur at the same time. It is the bifurcation to these traveling waves that is the one we will be most interested in here. At any one time the system will be completely filled with traveling waves separated into patches of left- or right traveling ones. A typical experimental picture is plotted in Fig. 3.2 (although it should be noted that this is actually a rather small system, and the boundaries therefore have a big influence on the dynamics). The patches are clearly visible, as well as the relatively narrow regions that separate the different patches.

These traveling waves are precisely what sets the heated wire experiment apart from the pattern forming systems we have encountered so far, in that the primary pattern is not a stationary spatially periodic one, but instead consists of *traveling waves*. The most detailed comparisons between the predictions of an amplitude description and experiments have been made [24] for the type of systems we encountered in the previous chapter, *i.e.* hydrodynamic systems that bifurcate to a stationary periodic pattern (critical wavenumber $q_c \neq 0$ and critical frequency $\omega_c = 0$). The corresponding amplitude equation as we have seen has real coefficients and takes the form of a Ginzburg-Landau equation; it is therefore often referred to as the real Ginzburg-Landau equation. As we have also seen, the coefficients occurring in this equation set length and time scales only, and for a theoretical analysis of an infinite system, they can be scaled away. We have shown that, for the RB-system, it is possible to derive the amplitude equation from the hydrodynamic equations in the Boussinesq approximation, but that simple symmetry considerations already yielded the correct lowest-order equation. This implies, that not only should the dynamics near the onset of convection in the RB-system obey this equation, but also *all other* systems with same symmetry properties. Hence one equation describes a variety of experimental situations and the theoretical predictions have been compared in detail with the experimental observations in a number of cases [24, 1, 25, 26].

For traveling wave systems (critical wavenumber $q_c \neq 0$ and critical frequency $\omega_c \neq 0$) such as the heated wire, there are, however, few examples of a direct confrontation between theory and experiment, since the qualitative dynamical behavior depends *strongly* on the various coefficients that enter the resulting amplitude equations². The calculations of these coefficients from the underlying equations of motion are rather

²In practice complications may also arise due to the presence of additional important slow variables [27, 28, 29, 30].

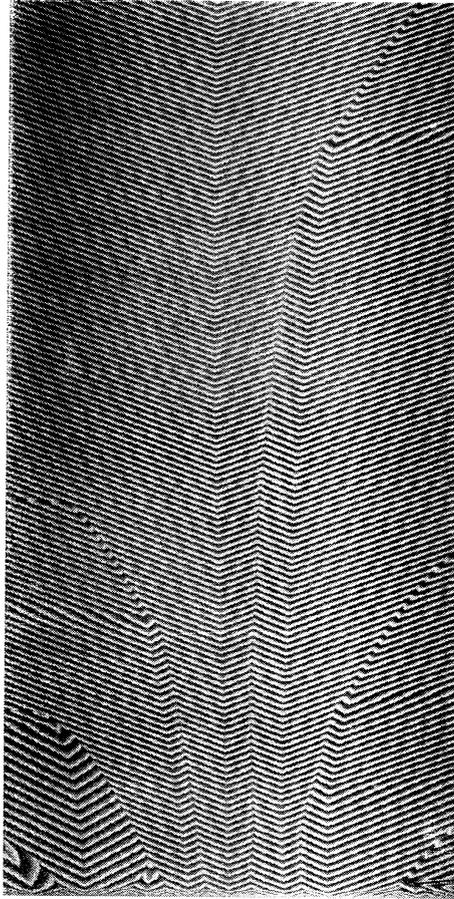


Figure 3.2: Temporal evolution of the heated wire system. Plotted here is a sequence of cross-sections of the traveling waves in a plane parallel to the wire, and time is running in the upward direction.. Darker regions correspond to wave crests, while the lighter regions correspond to the troughs. The vertical direction is time, so that the slope of the lines we see corresponds to the *phase velocity*. Several sources (regions emanating waves) and sinks (regions with incoming waves on either side) are visible. After [23].

involved and have only been carried out for a limited number of systems [31, 32, 33, 34, 35], and in many experimental cases the values of these coefficients are not known. We will therefore adopt a more heuristic approach here, and instead of deriving the amplitude equations from the hydrodynamic equations we will base our choice of equations on symmetry considerations, very much like the example given in Sec.2.10. A different problem generally arises when dealing with systems of counter-propagating waves, where in many cases the standard coupled amplitude equations (3.3.3a,3.3.3b) are not uniformly valid in ε . Therefore one has to be cautious about the interpretation of results based on these equations [36, 37, 38, 39].

3.3 Amplitude equations for the heated wire

All we need to do in order to write down the amplitude equation for this system is to list the basic symmetries. First of, we stress that we are looking for an equation to describe the *traveling waves* in the direction along the wire. While the waves *do* of course also have a variation in the perpendicular direction, this is the same all along the wire and will be neglected here. The system as we describe is therefore essentially one-dimensional, and in the following we will use x to denote the direction along the wire.

The bifurcation to traveling waves in the heated wire system is supercritical (an oscillatory supercritical bifurcation such as this one is known as a Hopf bifurcation). This implies that the amplitude of the traveling waves has to increase smoothly when the control parameter is increased beyond its critical value. The generic amplitude for a single amplitude will therefore be the a single CGL equation

$$\partial_t A + s_0 \partial_x A = \varepsilon A + (1 + ic_1) \partial_x^2 A - (1 - ic_3) |A|^2 A. \quad (3.3.1)$$

Now, because in the heated wire system we have both left- and right-traveling waves, and these modes are basically independent, we will require two amplitude equations, one for each mode separately. We will label the amplitudes of the left- and right traveling modes respectively by an index L or R . The observable fields one actually *sees* in an experiment are then related to these amplitudes like

$$\text{physical fields} \propto A_R e^{-i(\omega_c t - q_c x)} + A_L e^{-i(\omega_c t + q_c x)} + c.c. , \quad (3.3.2)$$

One of the key features of the experiments is that one actually sees traveling waves. If the situation were such that both modes were indeed completely decoupled, one would expect them to simultaneously grow beyond threshold, which would result in *standing* waves (the wavenumbers of left- and right traveling modes have to be equal because of the left-right symmetry). The fact that one doesn't actually see these standing waves implies that the two modes are coupled. Moreover, it tells us that they are coupled in such a way that the modes suppress each other. All of this leaves us with the set of equations that we will be using in this chapter and the next to describe the patterns close to onset in the heated wire system, a set of 2 coupled CGL equations

$$\begin{aligned} \partial_t A_R + s_0 \partial_x A_R &= \varepsilon A_R + (1 + ic_1) \partial_x^2 A_R \\ &\quad - (1 - ic_3) |A_R|^2 A_R - g_2 (1 - ic_2) |A_L|^2 A_R, \end{aligned} \quad (3.3.3a)$$

$$\begin{aligned} \partial_t A_L - s_0 \partial_x A_L &= \varepsilon A_L + (1 + ic_1) \partial_x^2 A_L \\ &\quad - (1 - ic_3) |A_L|^2 A_L - g_2 (1 - ic_2) |A_R|^2 A_L. \end{aligned} \quad (3.3.3b)$$

Note that for these coupled equations, we can no longer get rid of the linear group velocity terms by means of a Galilean transformation to a comoving frame, as there is obviously no one single frame that rids us of the s_0 -term in both equations. We prefer therefore to choose a frame in which both linear group velocities are equal, but of opposite sign. The left-right symmetry in the full (coupled) system is therefore broken *only* by the opposing signs of the respective advective terms.

3.4 Definition of sources and sinks

Sources and sinks arise when the coupling coefficient g_2 is sufficiently large that one mode suppresses the other. Then the system tends to form domains of either left-moving or right-moving waves, separated by domain walls or shocks. The distinction between *sources* or *sinks* according to whether the nonlinear group velocity points s of the asymptotic plane waves points *outwards* or *inwards* — see Fig. 3.3 — is crucial here. From a physical point of view, the group velocity determines the propagation of small perturbations. In our definition, a source is an “active” coherent structure which sends out waves to both sides, while a sink is sandwiched between traveling wave states with the group velocity pointing inwards; perturbations travel away from sources and into sinks. Mathematically,

it will turn out that the distinction between sources and sinks in terms of the group velocity s is also precisely the one that is natural in the context of the counting arguments.

In an actual experiment concerning traveling waves, when one measures an order parameter and produces space-time plots of its time evolution (compare for instance Fig. 3.2), lines of constant intensity indicate lines of constant phase of the traveling waves (see for example [40, 41, 65, 89]). The direction of the *phase velocity* v_{ph} of the waves in each single-mode domain is then immediately clear. But, since s and v_{ph} do not necessarily have to have the same sign, one can not distinguish sources and sinks based on this data alone. As was shown by Alvarez *et al.* [40] however, in the heated wire experiment v_{ph} and s are parallel, so that structures which to the eye look like sources, are *indeed* sources according to our definition.

In the coupled CGL equations (3.3.3a,3.3.3b), s_0 is the *linear* group velocity, i.e. the group velocity $\frac{\partial\omega}{\partial q}$ evaluated at q_c , of the fast modes³.

It is important to realize [21] that for positive ε , the group velocity s is *different* from s_0 . To see this, note that the coupled CGL equations admit single mode traveling waves of the form

$$A_R = ae^{-i(\omega_R t - qx)}, \quad A_L = 0, \quad (3.4.1)$$

or

$$A_L = ae^{-i(\omega_L t - qx)}, \quad A_R = 0. \quad (3.4.2)$$

Substitution of these wave solutions in the amplitude equations (3.3.3a) and (3.3.3b) yields

$$\omega_{R,L} = \pm s_0 q + c_1 q^2 - c_3 a_{R,L}^2, \quad (3.4.3a)$$

$$a_{R,L}^2 = \varepsilon - q^2, \quad (3.4.3b)$$

From this we obtain the nonlinear dispersion relation

$$\omega_{R,L} = \pm s_0 q + (c_1 + c_3)q^2, \quad (3.4.4)$$

³We stress that the indices R and L of the amplitudes A_R and A_L are associated with the sign of the *linear group velocity* s_0 . In writing Eq. (3.3.2) with q_c and ω_c positive, we have also associated a wave whose phase velocity v_{ph} is to the right with A_R , and one whose v_{ph} is to the left with A_L , but this choice is completely arbitrary: At the level of the amplitude equations, the sign of the phase velocity of the critical mode plays no role.

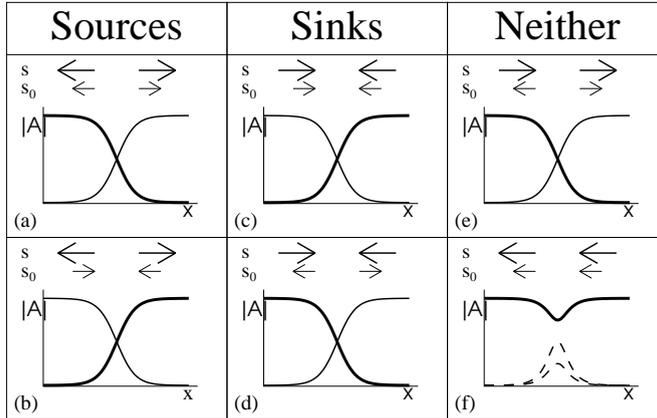


Figure 3.3: Schematic representations of the various coherent structures that we will encounter in this and the next chapter. The amplitude of the left (right) traveling waves is indicated by a thick (thin) curve, while the linear group velocity and total group velocity are denoted by s_0 and s respectively, and their direction is indicated by arrows. (a) and (b) are, in our definition, both sources, since the nonlinear group velocity s points outward; the majority of cases that we will encounter will be of type (a). Similarly, (c) and (d) both represent sinks. Finally, one may in principal encounter structures that are neither sources nor sinks. We never have observed a structure of the form shown in (e) in our simulations, but structures like shown in (f) occur quite generally in the chaotic regimes. The dotted curve for the A_R mode indicates that we can have many different possibilities here, including the case where $A_R=0$; in that case a description in terms of a single CGL equation suffices. Note that Figure (f) does not exhaust all possibilities which are essentially single-mode structures. E.g., in our simulations presented in Fig. 3.5, we encounter a case where in between a source of type (a) and one of type (b) there is a single-mode sink, for which s points inwards.

so that the nonlinear group velocity $s = \frac{\partial\omega}{\partial q}$ of these traveling waves becomes

$$s_R = s_{0,R} + 2(c_1 + c_3)q, \quad \text{with } s_{0,R} = s_0, \quad (3.4.5)$$

$$s_L = s_{0,L} + 2(c_1 + c_3)q, \quad \text{with } s_{0,L} = -s_0. \quad (3.4.6)$$

When $\varepsilon \downarrow 0$, we have seen that the band of allowed q -values shrinks to zero, and s approaches the linear group velocity $\pm s_0$, as it should. The term $2(c_1 + c_3)q$ accounts for the change in the group velocity away from threshold where the total wave number may differ from the critical value q_c . This term involves both the linear and the nonlinear dispersion coefficient, and its importance increases with increasing ε . We will therefore sometimes refer to s as the *nonlinear* or *total* group velocity, to emphasize the difference between s_0 and s .

Clearly it is possible for s_0 and s to have opposite signs. Since the labels R and L of A_R and A_L refer to the signs of *linear* group velocity s_0 , if this occurs, the mode A_R corresponds to a wave whose total group velocity s is to the left! The various possibilities concerning sources and sinks are illustrated in Fig. 3.3. It is important to stress that our analysis will focus on sources and sinks near the primary supercritical Hopf bifurcation from a homogeneous state to traveling waves. Experimentally, sources and sinks have been studied in detail by Kolodner [41] in his experiments on traveling waves in binary mixtures. Unfortunately, for this system a direct comparison between theory and experiments is hindered by the fact that the transition to traveling waves is *subcritical*, not supercritical.

3.5 Coherent structures and counting arguments

3.5.1 General formulation and main results

In many patterns that occur in experiments on traveling wave systems or numerical simulations of the single and coupled CGL equations (3.3.3a) and (3.3.3b), coherent structures (see Chapter 2, local structures that have an essentially time-independent shape and propagate with a constant velocity v) play an important role. For these coherent structures, the spatial and temporal degrees of freedom are not independent: apart from an overall phase factor, they are stationary in the frame co-moving with the coherent structure's velocity $\xi = x - vt$. Since the appropriate functions that describe the profiles of these coherent structures depend

only on the single variable ξ , these functions can be determined by ordinary differential equations (ODE's). These are obtained by substitution of the appropriate Ansatz in the original CGL equations, which of course are partial differential equations. Since the ODE's can themselves be written as a set of first order flow equations in a simple phase space, the coherent structures of the amplitude equations correspond to certain orbits of these ODE's. Note that plane waves, since they have constant profiles, are trivial examples of coherent structures; in the flow equations they correspond to fixed points. Sources and sinks connect, asymptotically, plane waves, and so the corresponding orbits in the ODE's connect fixed points. Many different coherent structures have been identified within this framework [22, 21, 42, 43].

The counting arguments that give the multiplicity of such solutions are essentially based on determining the dimensions of the stable and unstable manifolds near the fixed points. These dimensions, together with the parameters of the Ansatz such as v and the assumption that the phase-space flow is continuous, determine for a certain orbit the number of constraints and the number of free parameters that can be varied to fulfill these constraints. The theoretical importance of counting arguments can be illustrated by recalling that for the single CGL equation a continuous family of hole solutions has been known to exist for some time [42]. Later, however, counting arguments showed that these source type solutions were on general grounds expected to come as discrete sets, not as a continuous one-parameter family [21]. This suggested that there is some accidental degeneracy or hidden symmetry in the single CGL equation, so that by adding a seemingly innocuous perturbation to the CGL equation, the family of hole solutions should collapse to a discrete set. This was indeed found to be the case [44, 45]. For further details of the results and implications of these counting arguments for coherent structures in the single CGL equation, the reader is referred to [21].

It should be stressed that counting arguments can not prove the existence of certain coherent structures, nor can they establish the dynamical relevance of the solutions. They can only establish the multiplicity of the solutions, assuming that the equations have no hidden symmetries. Imagine that we know — either by an explicit construction or from numerical experiments — that a certain type of coherent structure solution does exist. The counting arguments then establish whether this should be an isolated or discrete solution (at most a member of a discrete set of them),

or a member of a one-parameter family of solutions, etc. In the case of an isolated solution, there are no nearby solutions if we change one of the parameters (like the velocity v) somewhat. For a one-parameter family, the counting argument implies that when we start from a known solution and change the velocity, we have enough other free parameters available to make sure that there is a perturbed trajectory that flows into the proper fixed point as $\xi \rightarrow \infty$.

For the two coupled CGL equations (3.3.3a,3.3.3b) the counting can be performed by a straightforward extension of the counting for the single CGL equation [21]. The Ansatz for coherent structures of the coupled CGL equations (3.3.3a,3.3.3b) is the following generalization of the Ansatz for the single CGL equation:

$$A_L(x, t) = e^{-i\omega_L t} \hat{A}_L(x - vt) , \quad (3.5.1a)$$

$$A_R(x, t) = e^{-i\omega_R t} \hat{A}_R(x - vt) . \quad (3.5.1b)$$

Note that we take the velocities of the structures in the left and right mode equal, while the frequencies ω are allowed to be different. This is due to the form of the coupling of the left- and right-traveling modes, which is through the moduli of the amplitudes. It obviously does not make sense to choose the velocities of the A_L and A_R differently: for large times the cores of the structures in A_L and A_R would then get arbitrarily far apart, and at the technical level, this would be reflected by the fact that with different velocities we would not obtain simple ODE's for \hat{A}_L and \hat{A}_R . Since the phases of A_L and A_R are not directly coupled, there is no a priori reason to take the frequencies ω_L and ω_R equal; in fact we will see that in numerical experiments they are not always equal (see for instance the simulations presented in Fig. 3.5). Allowing $\omega_L \neq \omega_R$, the Ansatz (3.5.1) clearly has three free parameters, ω_L, ω_R and v .

Substitution of the Ansatz (3.5.1) into the coupled CGL equations

(3.3.3a,3.3.3b) yields the following set of ODE's:

$$\partial_\xi a_L = \kappa_L a_L, \quad (3.5.2a)$$

$$\begin{aligned} \partial_\xi z_L = & -z_L^2 + \frac{1}{1+ic_1} [-\varepsilon - i\omega_L + (1-ic_3)a_L^2 \\ & + g_2(1-ic_2)a_R^2 - (v+s_0)z_L], \end{aligned} \quad (3.5.2b)$$

$$\partial_\xi a_R = \kappa_R a_R, \quad (3.5.2c)$$

$$\begin{aligned} \partial_\xi z_R = & -z_R^2 + \frac{1}{1+ic_1} [-\varepsilon - i\omega_R + (1-ic_3)a_R^2 \\ & + g_2(1-ic_2)a_L^2 - (v-s_0)z_R], \end{aligned} \quad (3.5.2d)$$

where we have written

$$\hat{A}_L = a_L e^{i\phi_L}, \quad (3.5.3a)$$

$$\hat{A}_R = a_R e^{i\phi_R}. \quad (3.5.3b)$$

and where the phase-space variables q , κ and z are defined as

$$q \equiv \partial_\xi \phi \quad (3.5.4a)$$

$$\kappa \equiv \frac{1}{a} \partial_\xi a, \quad (3.5.4b)$$

$$z \equiv \partial_\xi \ln(\hat{A}) = \kappa + iq. \quad (3.5.4c)$$

We prefer to use κ , the logarithmic derivative, rather than an ordinary derivative, in order to be able to fully resolve regions of exponential decay of the amplitudes. For such regions, the ordinary derivative tends to zero while κ remains finite. Compared to the flow equations for the single CGL equation (see appendix 3.A), there are two important differences that should be noted: (i) Instead of the velocity v we now have velocities $v \pm s_0$; this is simply due to the fact that the linear group velocity terms can not be transformed away. (ii) The nonlinear coupling term in the CGL equations shows up only in the flow equations for the z 's.

The fixed points of these flow equations, the points in phase space at which the right hand sides of Eqs. (3.5.2a)-(3.5.2d) vanish, describe the asymptotic states for $\xi \rightarrow \pm\infty$ of the coherent structures. What are these fixed points? From Eq. (3.5.2a) we find that either a_L or κ_L is equal to zero at a fixed point, and similarly, from Eq. (3.5.2c) it follows that either a_R or κ_R vanishes. For the sources and sinks of (3.3.3a,3.3.3b) that we wish to study, the asymptotic states are left- and right-traveling waves.

Therefore the fixed points of interest to us have either both a_L and κ_R to zero, and we search for heteroclinic orbits connecting these two fixed points.

As explained before, in a counting argument one determines the multiplicity of a certain class of solutions (coherent structures in our case) by comparing

- (i) the dimension $\mathcal{D}_{\text{out}}^-$ of the outgoing (“unstable”) manifold of the fixed point describing the state on the left ($\xi = -\infty$),
- (ii) the dimension $\mathcal{D}_{\text{out}}^+$ of the outgoing manifold at the fixed point characterizing the state on the right ($\xi = \infty$) and
- (iii) the number $\mathcal{N}_{\text{free}}$ of free parameters in the flow equations.

Note that every flowline of the ODE’s corresponds to a particular coherent solution, with a fully determined spatial profile but with an *arbitrary* position; if we would also specify the location of the point associated with $\xi = 0$ on the flowline, the position of the coherent structure would be fixed. When we refer to the multiplicity of the coherent solutions, however, we only care about the profile and not the position. We therefore need to count the multiplicity of the *orbits*. In terms of the quantities given above, one thus expects a $(\mathcal{D}_{\text{out}}^- - 1 - \mathcal{D}_{\text{out}}^+ + \mathcal{N}_{\text{free}})$ -parameter family of solutions; the factor -1 is associated with the invariance of the ODE’s with respect to a shift in the pseudo-time ξ which leaves the flowlines invariant. In other words, the coherent structures are translation invariant, as they should be since the amplitude equations are as well.

When the number $(\mathcal{D}_{\text{out}}^- - 1 - \mathcal{D}_{\text{out}}^+ + \mathcal{N}_{\text{free}})$ is zero, one expects a discrete set of solutions, while when this number is negative one expects there to be no solutions at all, generically. *Proving* the existence of solutions, within the context of an analysis of this type, amounts to proving that the outgoing manifold at the $\xi = -\infty$ fixed point and the incoming manifold at the $\xi = \infty$ fixed point intersect. Such proofs are in practice far from trivial — if at all possible — and will not be attempted here.

Counting arguments are conceptually simple, since the dimensions $\mathcal{D}_{\text{out}}^-$ and $\mathcal{D}_{\text{out}}^+$ are just determined by studying the linear flow in the neighborhood of the fixed points. Technically, the analysis of the coupled equations is a straightforward but somewhat involved extension of the earlier findings for the single CGL. We therefore prefer to only quote the main result of the analysis, and to relegate all technicalities to appendix 3.B.

For sources and sinks, one of the two modes always vanishes at the relevant fixed points. We are especially interested in the case in which the effective value of ε , defined as

$$\varepsilon_{\text{eff}}^L := \varepsilon - g_2 |a_R|^2, \quad \varepsilon_{\text{eff}}^R := \varepsilon - g_2 |a_L|^2. \quad (3.5.5)$$

is *negative* for the mode which is suppressed. In this case small perturbations of the suppressed mode decay to zero in each of the single-amplitude domains, rendering the configuration a *stable* one. E.g., for a stable source configuration as sketched in Fig. 3.3, $\varepsilon_{\text{eff}}^R$ should be negative on the left, and $\varepsilon_{\text{eff}}^L$ should be negative on the right of the source. We will focus below on the results for this regime of fully effective suppression of one mode by the other.

The basic result of our counting analysis for the multiplicity of source and sink solutions is that when $\varepsilon_{\text{eff}} < 0$ the counting arguments for “*normal*” sources and sinks (the linear group velocity s_0 and the nonlinear group velocity s of the same sign), is simply that

- *Sources occur in discrete sets. Within these sets, as a result of the left-right symmetry for $v = 0$, we expect a stationary, symmetric source to occur.*
- *Sinks occur in a two parameter family.*

Notice that apart from the conditions formulated above, these findings are completely independent of the precise values of the coefficients of the equations. This gives these results their predictive power. Essentially all of the results of the remainder of this chapter and the next are based on the first finding that sources come in discrete sets, so that they fix the properties of the states in the domains they separate.

As discussed in Appendix 3.B the multiplicity of *anomalous* sources (see for instance Fig. 3.5) is the same as for normal sources and sinks in large parts of parameter space, but larger multiplicities *can* occur. Likewise, sources with $\varepsilon_{\text{eff}} > 0$ may occur as a two-parameter family, although most of these are expected to be unstable (Appendix 3.B.7). We shall see in Section 4.1 that in this case, which happens in particular when g_2 is only slightly larger than 1, new nontrivial dynamics can occur.

3.5.2 Comparison between shooting and direct simulations

Clearly, the coherent structure solutions are by construction *special* solutions of the original partial differential equations. The question then

arises whether these solutions are also dynamically relevant, in other words, whether they emerge naturally in the long time dynamics of the CGL equation or as “nearby” transient solutions in nontrivial dynamical regimes. For the single CGL equation, this has indeed been found to be the case [22, 21, 46, 47, 48, 49]. To check that this is also the case here, we have performed simulations of the coupled CGL equations and compared the sinks and sources that are found there to the ones obtained from the ODE’s (3.5.2a-3.5.2d). Direct integration of the coupled CGL equations was done using a pseudo-spectral code. The profiles of uniformly translating coherent structures were obtained by direct integration of the ODE’s (3.5.2a-3.5.2d), shooting from both the $\xi = +\infty$ and $\xi = -\infty$ fixed points and matching in the middle.

In Fig. 3.4(a), we show a space-time plot of the evolution towards sources and sinks, starting from random initial conditions. The grey shading is such that patches of A_R mode are light and A_L mode are dark. Clearly, after a quite short transient regime, a stationary sink/source pattern emerges. In Fig. 3.4(b) we show the amplitude profiles of $|A_R|$ (thin curve) and $|A_L|$ (thick curve) in the final state of the simulations that are shown in Fig. 3.4(a). In Fig. 3.4(c) and (d), we compare the amplitude and wavenumber profile of the source obtained from the CGL equations around $x = 440$ (boxes) to the source that is obtained from the ODE’s (3.5.2a-3.5.2d) (full lines). The fit is excellent, which illustrates our finding that sources are stable and stationary in large regions of parameter space and that their profile is completely determined by the ODE’s associated with the Ansatz (3.5.1).

However, the CGL equations possess a large number of coefficients that can be varied, and it will turn out that there are several mechanisms that can render sources and source/sink patterns unstable. We will encounter these scenarios in sections 3.6 and 4.1.

3.5.3 Multiple discrete sources

As we already pointed out before, the fact that sources come in a discrete set does not imply that there exists only one unique source solution. There could in principle be more solutions, since the counting only tells us that infinitesimally close to any given solution, there will not be another one.

Fig. 3.5 shows an example of the occurrence of two different isolated source solutions. The figure is a space-time plot of a simulation where we obtained two different sources, one of which is an anomalous one (s and

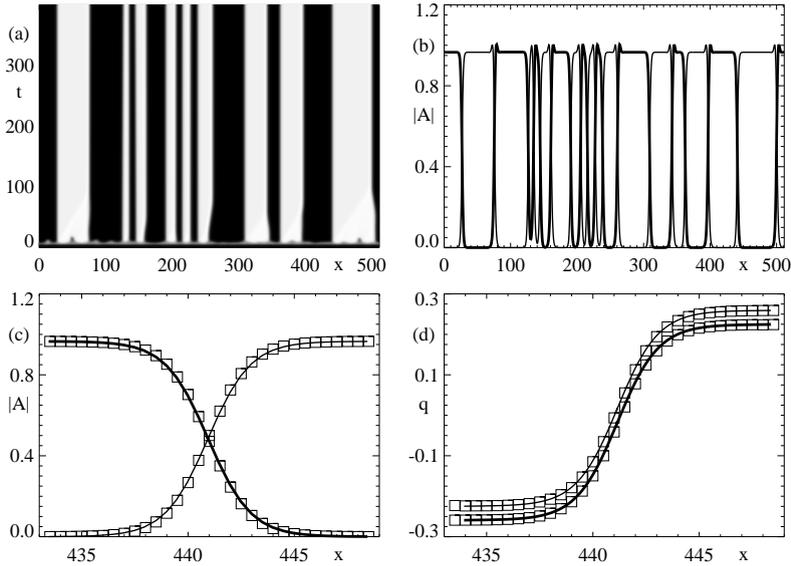


Figure 3.4: (a) Space-time plot showing the evolution of the amplitudes $|A_L|$ and $|A_R|$ in the CGL equations starting from random initial conditions. The coefficients were chosen as $c_1 = 0.6, c_2 = 0.0, c_3 = 0.4, s_0 = 0.4, g_2 = 2$ and $\varepsilon = 1$. The grey shading is such that patches of A_R mode are light and the A_L mode are dark. (b) Amplitude profiles of the final state of (a), showing a typical sink/source pattern. (c) Comparison between the source obtained from direct simulations of the CGL equations as shown in (b) (squares) and profiles obtained by shooting in the ODE's (3.5.2a-3.5.2d) (full curves). (d) Similar comparison, now for the wavenumber profiles. In (c) and (d), the thick (thin) curves correspond to the left (right) traveling mode.

s_0 of opposite sign). One clearly sees the different wavenumbers emitted by the two structures, and sandwiched in between these two sources is a single amplitude sink, whose velocity is determined by the difference in incoming wavenumbers. We have checked that the wavenumber selected by the anomalous source is such that the counting still yields a discrete set. If we follow the spatio-temporal evolution of this particular configuration, we find highly nontrivial behavior which we do not fully understand yet (not shown in Fig. 3.5).

These findings illustrate our belief that the "normal" sources and sinks are the most relevant structures one expects to encounter. It therefore appears to be safe to ignore the possible dynamical consequences of the more esoteric structures, which one *a priori* cannot rule out. The main complication of the possible occurrence of multiple discrete sources, as in Fig. 3.5, is that single amplitude sinks can arise in the patches separating them. The motion of these sinks can dominate the dynamics for an appreciable time.

3.6 Scaling properties of sources and sinks for small ε

In this section we study the scaling properties and dynamical behavior of sources and sinks in the limit where ε is small. This is a nontrivial issue, since due to the presence of the linear group velocity s_0 , the coupled CGL equations do not scale uniformly with ε . We focus in particular on the width of the sources and sinks. The results we obtain are open for experimental testing, since the control parameter ε can usually be varied quite easily. The behavior of the sources is the most interesting, and we will discuss this in sections 3.6.1 and 3.6.2. Using arguments from the theory of front propagation, we recover the result from Couillet *et al.* [50] that there is a finite threshold value for ε , below which no *coherent* sources exist (section 3.6.1). For ε below this critical value, there are, depending on the initial conditions, roughly two different possibilities. For well-separated sink/source patterns, we find *non-stationary* sources whose average width scales as $1/\varepsilon$ (possibly in agreement with the experiments of Vince and Dubois [51], but recently also seen convincingly Westra and van de Water [52]; see section 4.8.1). These sources can exist for arbitrarily small values of ε . For patterns with less-well separated sources and sinks, we typically find that the sources and sinks annihilate each other and

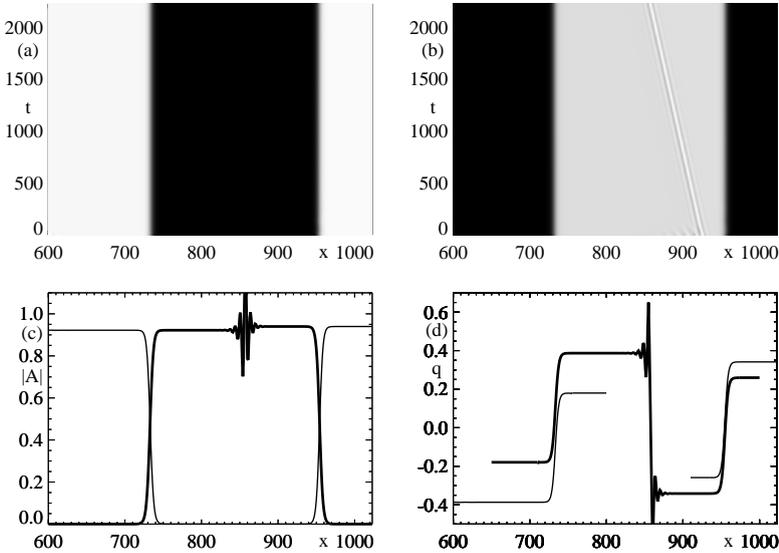


Figure 3.5: (a),(b) Space-time plots showing $|A_R|$ (a) and $|A_L|$ (b) in a situation in which there are two different sources present. Coefficients in this simulation are $c_1 = 3.0$, $c_2 = 0$, $c_3 = 0.75$, $g_2 = 2.0$, $s_0 = 0.2$ and $\varepsilon = 1.0$. Initial conditions were chosen such that a well-separated source-source pair emerges, and a short transient has been removed. The source at $x \approx 730$ is anomalous, i.e., its linear and nonlinear group velocity s_0 and s have opposite signs. Sandwiched between the sources is a single-mode sink, traveling in the direction of the anomalous source; this sink is visible in (b). (c) Snapshot of the amplitude profiles of the two sources and the single mode sink at the end of the simulation shown in (a-b). (d) The wavenumber profiles of the two sources in their final state. Note that when the modulus goes to zero, the wavenumber is no longer well-defined; we can only obtain q up to a finite distance from the sources. The selected wavenumber emitted by the anomalous source is $q_{\text{sel}} = 0.387$, while the wavenumber emitted by the ordinary source is $q_{\text{sel}} = 0.341$. The velocity of the sink in between agrees with the velocity that follows from a phase-matching rule, i.e., the requirement that the phase difference across the sink remains constant. In (c) and (d), thick (thin) curves correspond to left (right) traveling modes.

disappear altogether. The system evolves then to a single mode state. These scenarios are discussed in section 3.6.2 below. By some simple analytical arguments we obtain that the width of coherent sinks diverges as $1/\varepsilon$; typically these structures remain stationary (see section 3.6.3).

3.6.1 Coherent sources: analytical arguments

By balancing the linear group velocity term with the second order spatial derivative terms, we see that the coupled amplitude equations (3.3.3a-3.3.3b) may contain solutions whose widths approach a finite value of order $1/s_0$ as $\varepsilon \rightarrow 0$. As pointed out in particular by Cross [36, 37], this behavior might be expected near end walls in finite systems; in principle, it could also occur for coherent structures such as sources and sinks which connect two oppositely traveling waves. Solutions of this type are *not* consistent with the usual assumption of separation of scales (length scale $\sim \varepsilon^{-1/2}$) which underlies the derivation of amplitude equations. One should interpret the results for such solutions with caution.

As we shall discuss below, the existence of stationary, coherent sources is governed by a finite critical value ε_c^{so} , first identified by Coulet *et al.* [50]. Since the coupled amplitude equations (3.3.3a-3.3.3b) are only valid to lowest order in ε , the question then arises whether the existence of this finite critical value ε_c^{so} is a peculiarity of the lowest order amplitude equations. Since this threshold is determined by the interplay of the linear group velocity and a front velocity, which are both defined for arbitrary ε , we will argue that the existence of a threshold is a robust property indeed.

We now proceed by deriving this critical value ε_c^{so} from a slightly different perspective than the one that underlies the analysis of Coulet *et al.* [50], by viewing wide sources as weakly bound states of two widely separated fronts. Indeed, consider a sufficiently wide source like the one sketched in Fig. 3.6a in which there is quite a large interval where both amplitudes are close to zero⁴. Intuitively, we can view such a source as a weakly bound state of two fronts, since in the region where one of the amplitudes crosses over from nearly zero to some value of order unity, the other mode is nearly zero. Hence as a first approximation in describing the fronts that build up the wide source of the type sketched in Fig. 3.6a, we can neglect the coupling term proportional to g_2 in the core-region. The

⁴It is not completely obvious that wide sources necessarily have such a large zero patch, but this is what we have found from numerical simulations. Wide sinks actually will turn out not to have this property.

resulting fronts will now be analyzed in the context of the single CGL equation.

Let us look at the motion of the A_R front on the right (by symmetry the A_L front travels in the opposite direction). As argued above, its motion is governed by the single CGL equation in a frame moving with velocity s_0

$$(\partial_t + s_0 \partial_x) A_R = \varepsilon A_R + (1 + ic_1) \partial_x^2 A_R - (1 - ic_3) |A_R|^2 A_R . \quad (3.6.1)$$

The front that we are interested in here corresponds to a front propagating "upstream", i.e., to the left, into the *unstable* $A_R = 0$ state. Such fronts have been studied in detail [21], both in general and for the single CGL equation specifically.

Fronts propagating into unstable states come in two classes, depending on the nonlinearities involved. Typically, when the nonlinearities are saturating, as in the cubic CGL equation (3.6.1), the asymptotic front velocity v_{front} equals the *linear spreading velocity* v^* . This v^* is the velocity at which a small perturbation around the unstable state grows and spreads according to the *linearized* equations. For Eq. (3.6.1), the velocity v^* of the front, propagating into the unstable $A = 0$ state, is given by [21]

$$v^* = s_0 - 2\sqrt{\varepsilon(1 + c_1^2)} . \quad (3.6.2)$$

The parameter regime in which the selected front velocity is v^* is often referred to as the "linear marginal stability" [53, 54] or "pulled fronts" [55, 56, 57, 58] regime, as in this regime the front is "pulled along" by the growing and spreading of linear perturbations in the tip of the front.

For small ε , the velocity $v^* = v_{\text{front}}$ is positive, implying that the front moves to the right, while for large ε , v^* is negative so that the front moves to the left. Intuitively, it is quite clear that the value of ε where $v^* = 0$ will be an important critical value for the dynamics, since for larger ε the two fronts sketched in Fig. 3.6a will move towards each other, and some kind of source structure is bound to emerge. For $\varepsilon < \varepsilon_c^{\text{so}}$, however, there is a possibility that a source splits up into two retracting fronts. Hence the critical value of ε is defined through $v^*(\varepsilon_c^{\text{so}}) = 0$, which, according to Eq. (3.6.2) yields

$$\varepsilon_c^{\text{so}} = \frac{s_0^2}{(4 + 4c_1^2)} . \quad (3.6.3)$$

We will indeed find that the width of *coherent* sources diverges for this value of ε ; however, the sources will not disappear altogether, but are

replaced by *non-stationary* sources which can not be described by the coherent structures Ansatz (3.5.1).

3.6.2 Sources: numerical simulations

By using the shooting method, i.e., numerical integration of the ODE's (3.5.2a-3.5.2d), to obtain coherent sources, we have studied the width of the coherent sources as a function of ε . The width is defined here as the distance between the two points where the left- and right traveling amplitudes reach 50 % of their respective asymptotic values. In Fig. 3.6b, we show how the width of coherent sources varies with ε . For the particular choice of coefficients here ($c_1 = c_3 = 0.5, c_2 = 0, g_2 = 2$ and $s_0 = 1$), $\varepsilon_c^{s_0} = 0.2$, and it is clear from this figure, that the width of stationary source solutions of Eqs. (3.6.1) diverges at this critical value⁵.

In dynamical simulations of the full coupled CGL equations however, this divergence is cut off by a crossover to the dynamical regime characteristic of the $\varepsilon < \varepsilon_c^{s_0}$ behavior. Fig. 3.6c is a space-time plot of $|A_L| + |A_R|$ that illustrates the incoherent dynamics we observe for $\varepsilon < \varepsilon_c^{s_0}$. The initial condition here is source-like, albeit with a very small width. In the simulation shown, we see the initial source flank diverge as we would expect since $s_0 > v^*$. As time progresses, right ahead of the front a small 'bump' appears: as we mentioned before, both amplitudes are to a very good approximation zero in that region, so the state there is unstable (remember that though small, ε is still nonzero). This bump will therefore start to grow, and will be advected in the direction of the flank. The flank and bump merge then and the flank jumps forward. The average front velocity is thus enhanced. The front then slowly retracts again, and the process is repeated, resulting in a "breathing" type of motion. For longer times these oscillations become very, very small. For this particular choice of parameters, they become almost invisible after times of the order 3000; however, a close inspection of the data yields that the sources never become stationary but keep performing irregular oscillations. Since these fluctuations are so small, it is very likely that to an experimentalist such sources appear to be completely stationary.

From the point of view of the stability of sources, we can think of the change of behavior of the sources as a core-instability. This instability is basically triggered by the fact that wide sources have a large core where

⁵Note that by a rescaling of the CGL equations, one can set $s_0 = 1$ without loss of generality.

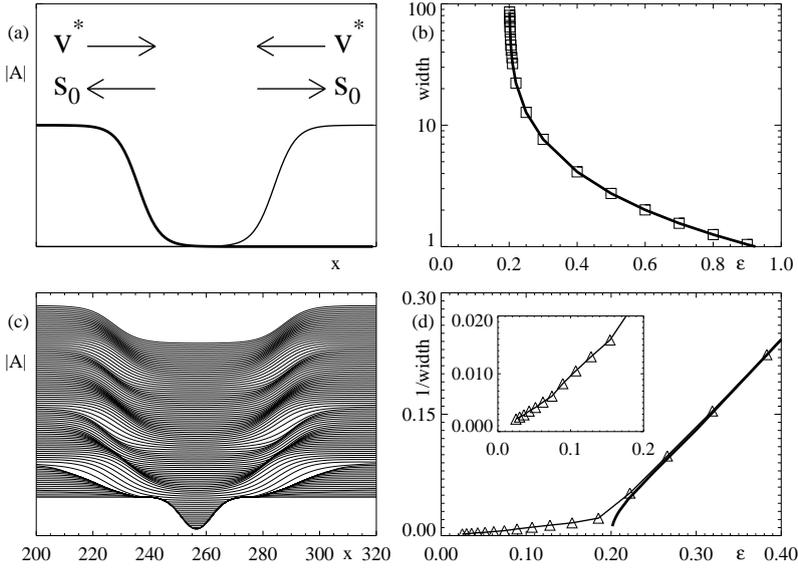


Figure 3.6: (a) Sketch of a wide source, indicating the competition between the linear group velocity s_0 and the front velocity v^* . (b) Width of coherent sources as obtained by shooting, for $c_1 = c_3 = 0.5$, $c_2 = 0$, $g_2 = 2$ and $s_0 = 1$. (c) Example of dynamical source for same values of the coefficients and $\epsilon = 0.15$. The order parameter shown here is the sum of the amplitudes $|A_L|$ and $|A_R|$, and the total time shown here is 1000. (d) Average inverse width of sources for the same coefficients as (b) as a function of ϵ . The thick curve corresponds to the coherent sources as shown in (b). For ϵ close to and below $\epsilon_c^{so} = 0.2$, there is a crossover to dynamical behavior. The inset shows the region around $\epsilon = 0$, where the average width roughly scales as ϵ^{-1} .

both A_L and A_R are small, and since the neutral state is unstable, this renders the sources unstable. The difference between the critical value of ε where the instability sets in and $\varepsilon_c^{\text{so}}$ is minute, and we will not dwell on the distinction between the two.⁶ Although all our numerical results are in accord with this scenario, one should be aware, however, that it is not excluded that other types of core-instabilities exist in some regions of parameter space⁷. Furthermore, it should be pointed out that when ε is below $\varepsilon_c^{\text{so}}$, there is absolutely *no* stationary, but unstable source! The dynamical sources can then *not* be viewed as oscillating around an unstable stationary source.

The weak fluctuations of the source flanks are very similar to the fluctuations of domain walls between single and bimodal states in inhomogeneously coupled CGL equations as studied in [59]. Completely analogous to what is found here, there is a threshold given in terms of ε and s_0 for the existence of stationary domain walls, which we understand now to result from a similar competition between fronts and linear group velocities. Beyond the threshold, dynamical behavior was shown to set in, which, depending on the coefficients, can take qualitatively different forms; similar scenarios can be obtained for the sources here.

The main ingredient that generates the dynamics seems to be the following. For a very wide source, we can think of the flank of the source as an isolated front. However, the *tip* of this front will always feel the other mode, and it is precisely this tip which plays an essential role in the propagation of “pulled” fronts [53, 54, 55, 56, 57, 58]! Close inspection of the numerics shows that near the crossover between the front regime and the interaction regime, oscillations, phase slips or kinks are generated, which are subsequently advected in the direction of the flank. These perturbations are a *deterministic* source of perturbations, and it is these perturbations that make the flank jump forward, effectively narrowing down the source.

⁶For a similar scenario in the context of non-homogeneously coupled CGL equations, see [59].

⁷An example of a similar scenario is provided by pulses in the single quintic CGL equation. Pulses are structures consisting of localized regions where $|A| \neq 0$. The existence and stability of pulse solutions can, to a large extent, be understood by thinking of a pulse as a bound state of two fronts [21]. However, recent perturbative calculations near the non-dissipative (Schrödinger-like) limit [60] have shown that in some parameter regimes a pulse can become unstable against a localized mode. This particular instability can not simply be understood by viewing a pulse as a bound state of two fronts.

The jumping forward of the flank of the source for ε just below ε_c^{so} is reminiscent to the mechanism through which traveling pulses were found to acquire incoherent dynamical behavior, if their velocity was different from the linear group velocity [61]. In extensions of the CGL equation, it was found that if a pulse would travel slower than the linear spreading speed v^* , fluctuations in the region just ahead of the pulse could grow out and make the pulse at one point "jump ahead". In much the same way the fronts can be viewed to "jump ahead" in the wide source-type structures below ε_c^{so} when the fluctuations ahead of it grow sufficiently large.

In passing, we point out that we believe these various types of "breathing dynamics" to be a general feature of the interaction between local structures and fronts. Apart from the examples mentioned above, a well known example of incoherent local structures are the oscillating pulses observed by Brand and Deissler in the quintic CGL [62]. Also in this case we have found that these oscillations are due to the interaction with a front, but instead of a pulled front it is a *pushed* front that drives the oscillations here [63].

Returning to the discussion of the behavior of the wide non-stationary sources, we show in Fig. 3.6d the (inverse) average width of the dynamical sources for small ε . These simulations were done in a large system (size 2048), with just one source and, due to the periodic boundary conditions, one sink. If one slowly decreases ε , one finds that the average width of the sources diverges roughly as ε^{-1} (see the inset of Fig. 3.6d). However, if one does not take such a large system, i.e., sources and sinks are not so well separated, we often observed that, after a few oscillations of the sources, they interact with the sinks and annihilate. In many cases, especially for small enough ε , *all* sources and sinks disappear from the system, and one ends up with a state of only right or left traveling wave.

In conclusion, we arrive at the following scenario.

- For $\varepsilon > \varepsilon_c^{so}$, sources are *stationary* and stable, provided that the waves they send out are stable. The structure of these stationary source solutions is given by the ODE's (3.5.2a-3.5.2d), and their multiplicity is determined by the counting arguments.
- When $\varepsilon \downarrow \varepsilon_c^{so}$, the source width rapidly increases, and for $\varepsilon = \varepsilon_c^{so}$, the size of the coherent sources (i.e., solutions of the ODE's (3.5.2a-3.5.2d)) diverges, in agreement with the picture of a source consisting of two weakly bound fronts. For a value of ε just above ε_c^{so} , the

sources have a wide core where both A_R and A_I are close to zero, and these sources turn unstable. Our scenario is that in this regime a source consists essentially of two of the “nonlinear global modes” of Couairon and Chomaz [64]. Possibly, their analysis can be extended to study the divergence of the source width as $\varepsilon \downarrow \varepsilon_c^{\text{SO}}$.

- For $\varepsilon < \varepsilon_c^{\text{SO}}$, *wide, non-stationary* sources can exist. Their dynamical behavior is governed by the continuous emergence and growth of fluctuations in the region where both amplitudes are small, resulting in an incoherent “breathing” appearance of the source. For long times, these oscillations may become very mild, especially when ε is not very far below $\varepsilon_c^{\text{SO}}$.
- In the limit for $\varepsilon \downarrow 0$, there are, depending on the initial conditions, two possibilities. For random initial conditions, pairs of sources and sinks annihilate and the system often ends up in a single mode state. This happens in particular in sufficiently small systems. Alternatively, in large systems, one may generate well-separated sources and sinks. In this case the average width of the incoherent sources diverges as $1/\varepsilon$, in apparent agreement with the experiments of Vince and Dubois [51] (see section 4.8.1 for further discussion of this point).

We finally note that our discussion above was based on the fact that near a supercritical bifurcation, fronts propagating into an unstable state are “pulled” [55, 56, 57, 58] or “linear marginal stability” [53, 54] fronts: $v_{\text{front}} = v^*$ (for more details, please consult Section 5.3). It is well-known that when some of the nonlinear terms tend to enhance the growth of the amplitude, the front velocity can be higher: $v_{\text{front}} > v^*$ [53, 54, 55, 56, 57, 58]. These fronts, which occur in particular near a subcritical bifurcation, are sometimes called “pushed” [55, 56, 57, 58] or “nonlinearly marginal stability” [21, 54] fronts. In this case it can happen that the front velocity remains large enough for stable stationary sources to exist all the way down to $\varepsilon = 0$. We believe that this is probably the reason that Kolodner [65] does not appear to have seen any evidence for the existence of a critical $\varepsilon_c^{\text{SO}}$ in his experiments on traveling waves in binary mixtures, as in this system the transition is weakly subcritical [31, 66].

3.6.3 Sinks

As is demonstrated in detail in section 3.B.2, counting arguments show that there generically exists a two-parameter family of uniformly translating sink solutions. The scaling of their width as a function of ε is not completely obvious, since the figures of Cross [36]⁸ indicate that their width approaches a finite value as $\varepsilon \downarrow 0$, while Coulet *et al.* found a class of sink solutions whose width diverges as ε^{-1} for $\varepsilon \downarrow 0$.

In appendix 3.C we demonstrate, by examining the ODE's (3.5.2a-3.5.2d) in the $\varepsilon \downarrow 0$ limit, that the asymptotic scaling of the width of sinks as ε^{-1} follows naturally.

If we now focus again on uniformly translating sink structures of the form

$$A_{R,L} = e^{-i\omega_{R,L}t} \hat{A}_{R,L}(\xi), \quad (3.6.4)$$

and explicitly carry out this scaling by introducing the scaled variables

$$\bar{\xi} = \varepsilon\xi, \quad (3.6.5a)$$

$$\bar{\omega}_{R,L} = \frac{\omega_{R,L}}{\varepsilon}, \quad (3.6.5b)$$

$$\bar{A}_{R,L} = \frac{\hat{A}_{R,L}}{\sqrt{\varepsilon}}, \quad (3.6.5c)$$

we find that, *if* the limit $\varepsilon \rightarrow 0$ is regular we can (to lowest order in ε), approximate the ODE's (3.5.2a-3.5.2d) by the following reduced set of equations

$$\begin{aligned} (-i\bar{\omega} + s_0\partial_{\bar{\xi}})\bar{A}_R &= \bar{A}_R - (1 - ic_3)|\bar{A}_R|^2\bar{A}_R - \\ &\quad - g_2(1 - ic_2)|\bar{A}_L|^2\bar{A}_R, \end{aligned} \quad (3.6.6)$$

$$\begin{aligned} (-i\bar{\omega} - s_0\partial_{\bar{\xi}})\bar{A}_L &= \bar{A}_L - (1 - ic_3)|\bar{A}_L|^2\bar{A}_L - \\ &\quad - g_2(1 - ic_2)|\bar{A}_R|^2\bar{A}_L, \end{aligned} \quad (3.6.7)$$

where we have set $\bar{\omega}_R = \bar{\omega}_L = \bar{\omega}$ and $v = 0$, to study symmetric, stationary sinks. As one can see by comparing Eqs. (3.6.6-3.6.7) with the original equations (3.5.2a-3.5.2d), the taking of the $\varepsilon \rightarrow 0$ limit effectively amounts to the removal of the diffusive term $\sim \partial_{\xi}^2$. One could *a priori* wonder

⁸The work of Cross was motivated by experiments on traveling waves in binary mixtures. In such systems, the bifurcation is weakly subcritical; experimentally, the sinks width is then expected to be finite for small ε .

whether this procedure is justified, since we are removing the highest order derivative from the equations, which could very well constitute a singular perturbation. This matter will be resolved below with the aid of our counting argument.

Equations (3.6.6-3.6.7) admit an exact solution for the sink profile, first obtained by Coulet *et al.* When we substitute

$$\bar{A}_{R,L} = \bar{a}_L e^{i\bar{\phi}_{R,L}}, \quad \bar{q}_{R,L} = \partial_{\bar{\xi}} \bar{\phi}_{R,L}, \quad (3.6.8)$$

the explicit solution is given by

$$a_R(x) = \sqrt{\frac{\varepsilon}{1 + e^{(2(g_2-1)\varepsilon x)/s_0}}} = \sqrt{\varepsilon - a_L^2}. \quad (3.6.9)$$

The width of these solutions is easily seen to indeed diverge as ε^{-1} . Since we can still vary $\bar{\omega}$ continuously to give various values for the asymptotic wavenumber, which is for solutions of the type (3.6.9) given by

$$\bar{q}_R = \frac{1}{s_0}(\bar{\omega} + c_3) \text{ for } \bar{\xi} = -\infty \text{ and } \bar{q}_L = \frac{-1}{s_0}(\bar{\omega} + c_3) \text{ for } \bar{\xi} = \infty, \quad (3.6.10)$$

we see that we still have a 1-parameter family of $v=0$ sinks. Since this is in accord with the full counting argument, the limit $\varepsilon \downarrow 0$ is indeed regular.

In passing we note that source solutions of finite width are completely absent in the scaled Eqs. (3.6.6-3.6.7). This is because the only orbit that starts from the $A_R = 0$ single mode fixed point and flows to the $A_L = 0$ single mode fixed point passes through the $A_L = A_R = 0$ fixed point, and therefore takes an infinite pseudo-time ξ ; such a source has an infinitely wide core regime where A_L and A_R are both zero. This also agrees with our earlier observations, since the coherent sources already diverge at finite ε_c^{so} .

In Fig. 3.7 we plot the sink width versus ε for the full set of ODE's, as obtained from our shooting. It is clear that the sink indeed diverges at $\varepsilon=0$, and that it asymptotically approaches the theoretical prediction from the above analysis.

3.6.4 The limit $s_0 \rightarrow 0$

In this paper, we focus mainly on the experimentally most relevant limit s_0 finite, ε small. For completeness, we also mention that Malomed [67]

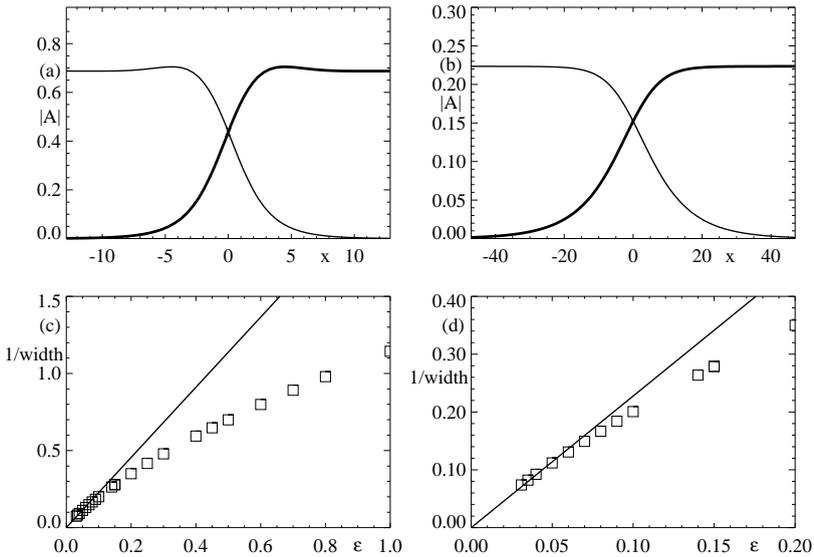


Figure 3.7: The width of stationary sinks obtained from the ODE's (3.5.2a,3.5.2d) as a function of ε , for $c_1 = 0.6$, $c_3 = 0.4$, $c_2 = 0$, $s_0 = 0.4$, $g_0 = 1$ and $g_2 = 2$. (a) Example of the stationary sink which has an incoming wavenumber corresponding to the wavenumber that is selected by the sources, for $\varepsilon = 0.5$. (b) *Idem*, now for $\varepsilon = 0.05$. Notice the differences in scale between (a) and (b). These two sinks are not related by simple scale transformations; this illustrates again the absence of uniform ε scaling of the coupled CGL equations. (c) As ε is decreased, the sink width initially roughly increases as $\varepsilon^{-1/2}$. When ε becomes sufficiently small, the group-velocity terms dominate over the diffusive/dispersive terms, and the sink-width is seen to obey an asymptotic ε^{-1} scaling (see (d) for a blowup around $\varepsilon = 0$). The straight line indicates the analytic result for the 50% width as obtained from Eq. (3.6.9), i.e. $\text{width}^{-1} = 5 \varepsilon / (2 \ln 3)$.

has also investigated the limit where ε is nonzero and $s_0 \rightarrow 0$, $c_i \rightarrow 0$, perturbatively. In this limit, which is relevant for some laser systems [68], sinks are found to be *wider* than sources. This finding can easily be recovered from the results of our appendix: From (3.A.11) it follows that to first order in s_0 the change in the exponential growth rate κ of the suppressed mode away from zero is

$$\delta\kappa_L^\pm = -s_0/2, \quad \delta\kappa_R^\pm = s_0/2. \quad (3.6.11)$$

where according to our convention of the appendices, κ^- corresponds to the negative root of (3.A.11), and κ^+ to the positive one. For a sink, the left traveling mode is suppressed on the left of the structure, and so this mode grows as $\exp(\kappa_L^+ \xi)$, while on the right of the sink the right-traveling mode decays to zero as $\exp(\kappa_L^- \xi)$. For the sources, the right and left traveling modes are interchanged. According to (3.6.11), upon increasing s_0 the relevant rate of spatial growth and decay decreases for sinks and increases for sources. Hence in this limit, somewhat counter-intuitively, sinks are wider than sources. For a further discussion of the limit $s_0 \rightarrow 0$, we refer to the paper by Malomed [67].

3.7 Conclusion

By considering only the basic symmetries of the heated wire system, we have been able to derive what should be, to lowest order, the amplitude equations that describe the behavior of this traveling wave system close to threshold. For the class of systems to which the heated wire belongs, these equations are a set of two coupled CGL equations.

Using a variety of techniques ranging from the use of exact solutions to numerical simulation, we have been able to obtain information about the coherent structures called sources and sinks in these coupled equations. We have tried to come up with predictions that would be relatively easy to verify experimentally, such as the dependence of their width on the control parameter and the uniqueness of the source solution.

This will turn out to be a particularly important finding, as we will see when we explore its consequences for the dynamics of the system as a whole in the next Chapter. The power of the counting arguments we have applied to derive some of the results in this Chapter lies in the fact that they are largely independent of the values of the various coefficients as they appear in the coupled equations.

In terms of some of the basic questions in pattern formation as posed in Section 1.4, we haven't done too bad so far. Although we have not identified the linear instability mechanism in the underlying hydrodynamic system, we were still able to construct the weakly nonlinear near-threshold theory. The subsequent discovery that stationary sources are unique effectively solves the question of selection in this system. Up to now, defects have played a relatively minor part but in the next Chapter we will see more of them.

3.A Coherent structures in the single CGLE

3.A.1 The flow equations

In this appendix, we lay the groundwork for our analysis of the coupled equations by summarizing and simplifying the main ingredients of the analysis of [21] of the single CGL equation

$$\partial_t A = \varepsilon A + (1 + ic_1)\partial_x^2 A - (1 - ic_3)|A|^2 A . \quad (3.A.1)$$

Note that if a single mode is present, the coupled equations reduce to a single CGL written in the frame moving with the linear group velocity of this mode, *not* in the stationary frame.

As in Eqs.(3.5.1), a coherent structure is defined as a solution whose time dependence amounts, apart from an overall time-dependent phase factor, to a uniform translation with velocity v :

$$A(x, t) \equiv e^{-i\omega t} \hat{A}(x - vt) = e^{-i\omega t} \hat{A}(\xi) . \quad (3.A.2)$$

Note that if the coherent structure approaches asymptotically a plane wave state for $\xi \rightarrow \infty$ or for $\xi \rightarrow -\infty$, the phase velocity of these waves would equal the propagation velocity of the coherent structures if ω would be 0. When $\omega \neq 0$, these two velocities differ.

For solutions of the form (3.A.2) we have that $\partial_t = -i\omega - v\partial_\xi$, so when we substitute the Ansatz (3.A.2) into the single CGL equation (3.A.1), we obtain the following ODE:

$$(-i\omega - v\partial_\xi)\hat{A} = \varepsilon\hat{A} + (1 + ic_1)\partial_\xi^2 \hat{A} - (1 - ic_3)|\hat{A}|^2 \hat{A} . \quad (3.A.3)$$

Solutions of this ODE correspond to coherent structures of the CGL equation (3.A.1) and vice-versa [21].

To analyze the orbits of the ODE (3.A.3), it is useful to rewrite it as a set of coupled first order ODE's. To do so, it is convenient to write A in terms of its amplitude and phase

$$\hat{A}(\xi) \equiv a(\xi)e^{i\varphi(\xi)} , \quad (3.A.4)$$

where a and φ are real-valued. Substituting the representation (3.A.4) into the ODE (3.A.3) yields, after some trivial algebra

$$\partial_\xi a = \kappa a , \quad (3.A.5a)$$

$$\partial_\xi \kappa = \mathcal{K}(a, q, \kappa) , \quad (3.A.5b)$$

$$\partial_\xi q = \mathcal{Q}(a, q, \kappa) , \quad (3.A.5c)$$

where q and κ are defined as

$$q \equiv \partial_\xi \varphi, \quad \kappa \equiv (1/a) \partial_\xi a. \quad (3.A.6)$$

The fact that there is no fourth equation is due to the fact that the CGL equation is invariant under a uniform change of the phase of A , so that φ itself does not enter in the equations. The functions \mathcal{K} and \mathcal{Q} are given by [21]

$$\mathcal{K} \equiv \frac{1}{1+c_1^2} [c_1(-\omega - vq) - \varepsilon - v\kappa + (1 - c_1c_3)a^2] + q^2 - \kappa^2, \quad (3.A.7a)$$

$$\mathcal{Q} \equiv \frac{1}{1+c_1^2} [(-\omega - vq) + c_1(v\kappa + \varepsilon) - (c_1 + c_3)a^2] - 2\kappa q. \quad (3.A.7b)$$

At first sight it may appear somewhat puzzling that we write the equations in a form containing $\kappa = \partial_\xi \ln a$ instead of simply $\partial_\xi a$. One advantage is that it allows us to distinguish more clearly between various structures whose amplitudes vanish exponentially as $\xi \rightarrow \pm\infty$ — these are then still distinguished by different values of κ . Secondly, the choice of κ in favor of $\partial_\xi a$ allow us to combine κ and q as the real and imaginary part of the logarithmic derivative of \hat{A} : we can rewrite (3.A.5b) and (3.A.5c) more compactly as

$$\partial_\xi z = -z^2 + \frac{1}{1+ic_1} [-\varepsilon - i\omega + (1 - ic_3)a^2 - vz]. \quad (3.A.8)$$

where $z \equiv \partial_\xi \ln(\hat{A}) = \kappa + iq$.

The fixed points of the ODE's have, according to (3.A.5a), either $a=0$ or $\kappa=0$. The values of q and κ for the $a=0$ fixed points are related through the dispersion relation of the linearized equation, or, what amounts to the same, by the equation obtained by setting the right hand side of (3.A.8) equal to zero and taking $a=0$. Following [21] we will refer to these fixed points as *linear fixed points*. We will denote them by L_\pm , where the index indicates the sign of κ . This means that the behavior near an L_+ fixed point corresponds to a situation in which the amplitude is growing away from zero to the right, while the behavior near an L_- fixed point describes the situation in which the amplitude a decays to zero.

Since a fixed point with $a \neq 0, \kappa=0$ corresponds to nonlinear traveling waves, the corresponding fixed points are referred to as *nonlinear fixed points* [21]. We denote these by N_\pm , where the index now indicates the sign of the *nonlinear group velocity* s of the corresponding traveling wave [21]. Thus, since the index of N denotes the sign of the group velocity,

the amplitude near an N_+ fixed point can either grow ($\kappa > 0$) or decay ($\kappa < 0$) with increasing ξ .

The coherent structures correspond to orbits which go from one of the fixed points to another one or back to the original one, and the counting analysis amounts to establishing the dimensions of the in- and outgoing manifolds of these fixed points. In combination with the number of free parameters (in this case v and ω), this yields the multiplicity of orbits connecting these fixed points, and therefore of the multiplicity of the corresponding coherent structures.

3.A.2 Fixed points and linear flow equations in their neighborhood

Since there are three flow equations (3.A.5a), there are three eigenvalues of the linear flow near each fixed point. When we perform the counting analysis for these fixed points we will only need the signs of the real parts of the three eigenvalues, since these determine whether the flow along the corresponding eigendirection is inwards ($-$) or outwards ($+$). We will denote the signs by pluses and minuses, so that $L_-(+, +, -)$ denotes an L_- fixed point with two eigenvalues which have a positive real part, and one which has a negative real part.

From Eqs.(3.A.5a) and (3.A.8), we obtain as fixed point equations

$$\begin{aligned} a\kappa &= 0, \\ (1 + ic_1)z^2 + vz + \varepsilon + i\omega - (1 + ic_3)a^2 &= 0, \end{aligned} \quad (3.A.9)$$

where $z = \kappa + iq$. From (3.A.9) we immediately obtain that fixed points either have $a=0$ (linear fixed points denoted as L) or $a \neq 0, \kappa=0$ (nonlinear fixed points denoted as N). Defining $\tilde{v} \equiv v/(1 + c_1^2)$ and $\tilde{a} \equiv a/(1 + c_1^2)$, the derivative of the flow (3.A.5a) is given by the matrix:

$$\mathcal{D} = \begin{pmatrix} \kappa & a & 0 \\ 2\tilde{a}(1 - c_1c_3) & -2\kappa - \tilde{v} & 2q - c_1\tilde{v} \\ -2\tilde{a}(c_1 + c_3) & -2q + c_1\tilde{v} & -2\kappa - \tilde{v} \end{pmatrix}. \quad (3.A.10)$$

Solving the fixed point equations (3.A.9,3.A.9) and calculating the eigenvalues of the matrix \mathcal{D} (3.A.10) yields the dimensions of the incoming and outgoing manifolds of these fixed points. Note that according to our convention, a fixed point with a two-dimensional outgoing and one-dimensional ingoing manifold is denoted as $(+, +, -)$.

We can restrict the calculations to the case of positive v , since the case of negative v can be found by the left-right symmetry operation: $\xi \rightarrow -\xi$, $v \rightarrow -v$, $z \rightarrow -z$.

3.A.3 The linear fixed points

For the linear fixed points $a=0$, and from (3.A.9) we obtain as fixed-point equation:

$$(1 + ic_1)z^2 + vz + \varepsilon + i\omega = 0, \quad (3.A.11)$$

which has as solutions

$$z = \frac{-v \pm \sqrt{v^2 - 4(1 + ic_1)(\varepsilon + i\omega)}}{2(1 + ic_1)}. \quad (3.A.12)$$

The linear fixed points come as a pair, and the left-right symmetry implies that for $v=0$, the eigenvalues of these fixed points are opposite.

At these fixed points, the eigenvalues are given by

$$\kappa \text{ or } -\tilde{v} - 2\kappa \pm i(c_1\tilde{v} - 2q). \quad (3.A.13)$$

To establish the signs of the real parts of the eigenvalues, we need to determine the signs of κ and $-\tilde{v} - 2\kappa$.

Let us first establish the signs of κ ; this is important in establishing whether the evanescent wave decays to the left (L_+) or to the right (L_-). For $v=0$, the equation (3.A.11) is purely quadratic, and so its solutions come in pairs $\pm(\kappa + iq)$. By expanding the square-root (3.A.13) for large v one obtains that in this case $\kappa = -v$ or $\kappa = -\varepsilon/v$; for large v , both κ 's are negative. Solving equation (3.A.11) we find that κ changes sign when

$$q = \pm\sqrt{\varepsilon}, \quad v = \frac{c_1\varepsilon - \omega}{\sqrt{\varepsilon}}. \quad (3.A.14)$$

For $\varepsilon < 0$, these equations have no solutions, and in that case there always is a L_- and a L_+ fixed point. For $\varepsilon > 0$ and $v < (c_1\varepsilon - \omega)/\sqrt{\varepsilon}$ there also is a L_- and a L_+ fixed point; for large v , there are two L_- fixed points.

To determine the sign of $-\tilde{v} - 2\kappa$ note that from the solution (3.A.12), we obtain that $\kappa = -\tilde{v}/2 \pm \text{Re}(\sqrt{\dots}/\dots)$. After some trivial rearranging this yields that $-\tilde{v} - 2\kappa$ has opposite sign for the pair of L fixed points; when one of them has two + 's, the other has two - 's.

In the case that we have a L_+ and a L_- fixed point the counting is as follows. For the L_+ fixed point, $-\tilde{v} - 2\kappa$ is negative since both v and κ

are positive, and the eigenvalue structure is then $(+, -, -)$. The L_- fixed point then has one negative eigenvalue κ , and two positive eigenvalues coming from the $-\tilde{v} - 2\kappa$. For large v , both κ 's are negative, and we obtain a $L_-(+, +, -)$ and a $L_-(+, -, -)$ fixed point.

In summary, then, the counting for the linear fixed points is as follows:

$$\begin{array}{l} \varepsilon < 0 \\ \varepsilon > 0 \end{array} \quad \left\{ \begin{array}{ll} \text{all } v : & L_-(+, +, -) \quad L_+(+, -, -) , \\ v < -v_{cL} : & L_+(+, -, -) \quad L_+(+, +, +) , \\ |v| < v_{cL} : & L_-(+, +, -) \quad L_+(+, -, -) , \\ v > v_{cL} : & L_-(+, +, -) \quad L_-(-, -, -) , \end{array} \right. \quad (3.A.15)$$

where $v_{cL} = |c_1\varepsilon - \omega|/\sqrt{\varepsilon}$.

3.A.4 The nonlinear fixed points

The analysis of the nonlinear fixed points goes along the same lines. Since the nonlinear fixed point has $\kappa = 0$, $z = iq$, the fixed point equations become:

$$a^2 = \varepsilon - q^2, \quad q^2(c_1 + c_3) - vq - \omega - c_3\varepsilon = 0. \quad (3.A.16)$$

which yields

$$q = \frac{v \pm \sqrt{v^2 + 4(\omega + c_3\varepsilon)(c_1 + c_3)}}{2(c_1 + c_3)}. \quad (3.A.17)$$

So the nonlinear fixed points come as a pair.

To obtain the eigenvalues, we substitute $\kappa = 0$ in the (3.A.10) and obtain as a secular equation:

$$(1 + c_1^2)\lambda^3 + 2v\lambda^2 + [2a^2(c_1c_3 - 1) + 4q^2(1 + c_1^2) - 4c_1qv + v^2]\lambda + [4a^2(c_1 + c_3)q - 2a^2v] = 0. \quad (3.A.18)$$

We only need to know the number of solution of the secular equation that have positive real part, and instead of solving the equation explicitly, we can proceed as follows. For a we cubic equation of the form

$$p_3\lambda^3 + p_2\lambda^2 + p_1\lambda^1 + p_0, \quad (3.A.19)$$

where $p_3 > 0$, we may read off the signs of the real parts of the solution to this equation from the following table [21]:

$$\begin{array}{l}
 p_0 > 0 \left[\begin{array}{ll} p_2 > 0, & p_1 p_2 > p_0 p_3 : & (-, -, -) & (\text{case } i) , \\ \text{else:} & & (+, +, -) & (\text{case } ii) , \end{array} \right. \\
 p_0 < 0 \left[\begin{array}{ll} p_2 < 0, & p_1 p_2 < p_0 p_3 : & (+, +, +) & (\text{case } iii) , \\ \text{else:} & & (+, -, -) & (\text{case } iv) . \end{array} \right. \quad (3.A.20)
 \end{array}$$

According to these rules, there are three combinations of the coefficients that we need to now the sign of, being

$$p_0 = 4a^2 q(c_1 + c_3) - 2a^2 v , \quad (3.A.21a)$$

$$p_2 = 2v , \quad (3.A.21b)$$

$$\begin{aligned}
 p_1 p_2 - p_0 p_3 = & - (1 + c_1^2) [4a^2(c_1 + c_3)q - 2a^2 v] + \\
 & + 2v [2a^2(c_1 c_3 - 1) + 4q^2(1 + c_1^2) - \\
 & - 4c_1 q v + v^2] . \quad (3.A.21c)
 \end{aligned}$$

As before, we will take $v > 0$, which makes $p_2 > 0$.

The sign of p_0 is equal to the sign of $2q(c_1 + c_3) - v$, which according to Eq. (3.A.17) is either $\pm\sqrt{\dots}$. The group velocity $\partial_q \omega$ of the the plane waves corresponding to the N fixed points is found from (3.A.16) to be $2q(c_1 + c_3) - v$, which can be rewritten as $p_0/(2a^2)$. So, we always have one N_- fixed point with $p_0 < 0$ and one N_+ fixed point with $p_0 > 0$.

When $p_0 < 0$, since p_2 is positive, the fixed point is $N_-(+, -, -)$ (case (iv)). When $p_0 > 0$, the eigenvalues depend on the sign of $p_1 p_2 - p_0 p_3$; when it is positive the eigenvalues are $(-, -, -)$, when it is negative, the eigenvalues are $(+, +, -)$. Defining v_{cN} as the value of $|v|$ where $p_1 p_2 - p_0 p_3$ changes sign, we obtain for the nonlinear fixed points:

$$\begin{array}{ll}
 v < -v_{cN} : & N_-(+, +, +) \text{ and } N_+(+, +, -) , \\
 |v| < v_{cN} : & N_-(+, -, -) \text{ and } N_+(+, +, -) , \\
 v > v_{cN} : & N_-(+, -, -) \text{ and } N_+(-, -, -) .
 \end{array} \quad (3.A.22)$$

Eqs. (3.A.15) and (3.A.22) express the dimensions of the stable and unstable manifolds of the fixed points of the single CGL equation, and these are the basis for the counting arguments for coherent structures in this equation [21]. We now turn to the extension of these results to the coupled CGL equations.

3.B Detailed counting for the coupled CGL equations

3.B.1 General considerations

While the counting for the coupled CGL equations follows unambiguously from that for the single CGL, there are various nontrivial subtleties in the extension of those results to the coupled CGL equations that require careful discussion.

Suppose we want to perform the counting for the $a_L=0, \kappa_R=0$ fixed point, which corresponds to the case in which only a right-traveling wave is present. The fixed point equations that follow from (3.5.2d) are, up to a change of $v \rightarrow v - s_0$, equal to the fixed point equation for the nonlinear fixed points of the single CGL equation, and can be solved accordingly. To solve the fixed point equations that follow from (3.5.2b), note that a_R is a constant at the fixed point and so the term $-g_2(1 - ic_2)a_R^2$ can be absorbed in the $-\varepsilon - i\omega_L$ term. Since we may choose ω_L freely, for the counting analysis we can forget about the $ig_2c_2a_R^2$ as we may think of it as having been absorbed into the frequency. The sign of $\varepsilon_{\text{eff}}^L$, defined in (3.5.5) to be $\varepsilon_{\text{eff}}^L = \varepsilon - g_2a_R^2$ will, however, be important. Likewise, at the other fixed point where $a_R = \kappa_L = 0$ the effective ε is $\varepsilon_{\text{eff}}^R = \varepsilon - g_2a_L^2$.

Since the fixed points we are interested in for sources and sinks always have either $a_L=0$ or $a_R=0$, the linearization around them largely parallels the analysis of the single CGL equation. For, when we linearize about the $a_L=0$ fixed point, we do not have to take into account the variation of a_R in the coupling term and this allows us, for the counting argument, to absorb these terms into an effective ε and redefined ω as discussed above. Once this is done, the linear equations for the mode whose amplitude a vanishes at the fixed point *do not involve the other mode variables at all*. As a result, the matrix of coefficients of the linearized equations has a block structure, and most of the results follow directly from those of the single CGL equation. We will below demonstrate this explicitly, using a symbolic notation for various terms whose precise expression we do not need explicitly.

If we consider the 6 variables $a_L, \kappa_L, q_L, a_R, \kappa_R$ and q_R as the elements of a vector w , and linearize the flow equations (3.A.5a) about a fixed point where one of the modes is nonzero, we can write the linearized equations

in the form $\dot{w}_i = \sum_j \mathcal{M}_{ij} w_j$, where the 6×6 matrix \mathcal{M} has the structure

$$\mathcal{M} = \begin{pmatrix} \kappa_L & a_L & 0 & 0 & 0 & 0 \\ "a_L" & X & X & "a_R" & 0 & 0 \\ "a_L" & X & X & "a_R" & 0 & 0 \\ 0 & 0 & 0 & \kappa_R & a_R & 0 \\ "a_L" & 0 & 0 & "a_R" & X & X \\ "a_L" & 0 & 0 & "a_R" & X & X \end{pmatrix}. \quad (3.B.1)$$

In this expression, all quantities assume their fixed point values. Furthermore, " a_R " and " a_L " represent terms that are linear in a_R or a_L , and the X stand for longer expressions that we do not need at the moment. At the fixed points, either a_R or a_L is zero, so either the upper-right block is identical to zero, or the lower-left block is zero. *In either case, the eigenvalues are simply given by the eigenvalues of the upper-left and lower-right block-matrices.* This implies that for each of the 3×3 blocks, we can use the results of the counting for a single CGL equation, provided we take into account that v and ε should be replaced by $v \pm s_0$ and $\varepsilon_{\text{eff}}^L$ or $\varepsilon_{\text{eff}}^R$ at the appropriate places!

As discussed in Appendix 3.A, the fixed point structure of the single CGL depends on two "critical" velocities, v_{cL} and v_{cN} . In general, these are different for the two fixed points which the orbit connects, so there is in principle a large number of possible regimes, each with their own combination of eigenvalue structures at the fixed points. An exhaustive list of all possibilities can be given, but it does not appear to be worthwhile to do so here. For, many of the exceptional cases occur for large values of the propagation velocity v and the relevance of the results for these solutions of the coupled CGL equations is questionable — after all, as we explained before, the counting can at most only demonstrate that certain solutions might be possible in some of these presumably somewhat extreme ranges of parameter values, but they by no means prove the existence of such solutions or their stability or dynamical relevance. Indeed, as discussed in section 3.6.2, for small ε the sources are intrinsically dynamical and are not given by the *coherent* sources as obtained from the ODE's (3.5.2a-3.5.2d).

For these reasons, our discussion will be guided by the following observations. The sinks and sources observed in the heated wire experiments have velocities that are smaller than the group velocity [40]⁹; this also seems to hold for other typical experiments with finite linear group

⁹In the experiments of [40], it was estimated from the data that $s_0 \approx v_{ph}/3$, where

velocity s_0 . This motivates us to start the discussion by investigating the regime that the velocity v is smaller than the linear group velocity, $|v| < s_0$. The sources are now as sketched in Fig. 3.3a and the sinks are as in Fig. 3.3c; this restriction already leads to a tremendous simplification. Furthermore, we are especially interested in the case that the two modes suppress each other sufficiently that the effective ε of the mode which is suppressed is negative, i.e., $\varepsilon_{\text{eff}}^{L/R} < 0$. This requirement is certainly fulfilled for sufficiently strong cross-coupling. The technical simplification of taking $\varepsilon_{\text{eff}}^{L/R} < 0$ is that in this case the structure of the linear fixed points is completely independent of the parameters v and ω — see Eq. (3.A.15). It should be noted, however, that in section 4.3 we will encounter source/sink patterns where ε_{eff} is positive; these patterns are chaotic. Also, the *anomalous* sources and sinks, mentioned at the end of section 3.5, can in some parameter ranges defy the general rules obtained here (see section 3.B.7 of this appendix). Furthermore, in section 3.B.6 we will discuss the cases $s_0 < 2q(c_1 + c_3)$ (i.e., sources and sinks corresponding to those of Fig. 3.3b and d), and the $s_0 = 0$ limit.

3.B.2 Multiplicities of sources and sinks

We will first perform the analysis starting with the restrictions given above. From Fig. 3.3 we can read off what the building blocks of sources and sinks are. We refer to the fixed point corresponding to $x \rightarrow -\infty(\infty)$ as fixed point 1 (2). In the coupled CGL equation case, we refer to the total group velocity of the nonlinear waves, which is given by $2q(c_1 + c_3) + v \pm s_0$ [see Eqs.(3.4.5), (3.4.6)]; since by the substitution $v \rightarrow v \pm s_0$ we absorb the s_0 in the v , the indexes of the N_- and N_+ fixed points correspond to the nonlinear group velocities in the co-moving frame of the coherent structures. For sinks of the type sketched in Fig. 3.3c, $A_L = 0$ for large negative x and $A_R = 0$ for large positive x . Consequently, the flow is

$$\text{from } \begin{cases} N_+ & (v - s_0) \\ L_+ & (v + s_0) \end{cases} \text{ to } \begin{cases} L_- & (v - s_0) \\ N_- & (v + s_0) \end{cases} . \quad (3.B.2)$$

For sources of the type sketched in Fig. 3.3a, $A_R = 0$ for large negative x and $A_L = 0$ for large positive x . Consequently, the flow is

$$\text{from } \begin{cases} N_- & (v + s_0) \\ L_+ & (v - s_0) \end{cases} \text{ to } \begin{cases} L_- & (v + s_0) \\ N_+ & (v - s_0) \end{cases} . \quad (3.B.3)$$

v_{ph} is the phase velocity, while typical sinks had a velocity v which could be as small as $v_{ph}/50$.

As in appendix 3.A, we will denote the real parts of the three eigenvalues of the fixed points by a string of plus or minus signs; e.g. $(+, -, -)$.

For $\varepsilon_{eff} < 0$ and arbitrary velocities, we obtain for the L fixed points (see Eqs. (3.A.15)):

$$L_-(+, +, -), \quad L_+(+, -, -). \quad (3.B.4)$$

For now we assume that $|v| < s_0$, $v - s_0 < 0$ and $v + s_0 > 0$. This yields, according to (3.A.22) for the N fixed points:

$$N_-(+, -, -), \quad N_+(+, +, -). \quad (3.B.5)$$

For sources we find that the combined (N_-, L_+) fixed point 1 has a two-dimensional outgoing manifold, which yields one free parameter. We can think of this parameter as a coordinate parameterizing the “directions” on the unstable manifold¹⁰. Now, the only other freedom we have for the trajectories out of fixed point 1 is associated with the freedom to view v , ω_L and ω_R as parameters in the flow equations that we can freely vary. This yields a total of four free parameters. Fixed point 2 (a (N_+, L_-) combination) has, according to Eqs. (3.B.3-3.B.5), a four-dimensional outgoing manifold. An orbit starting from fixed point 1 has to be “perpendicular” to this manifold in order to flow to fixed point 2; this yields four conditions. Assuming that these conditions can be obeyed for some values of the free parameters, it is clear that as long as there are no accidental degeneracies, we expect that there is at most only a discrete set of solutions possible — in other words, solutions will be found for sets of isolated values of the angle, v , ω_L and ω_R . One refers to this as a discrete set of sources.

When we fix $v = 0$, there is the following symmetry that we have to take into account: $\xi \rightarrow -\xi, z_L \leftrightarrow -z_R, a_L \leftrightarrow a_R$. Furthermore, this left-right symmetry yields that we should take $\omega_L = \omega_R$, so, in comparison to the general case, we have two free parameters less. When the outgoing manifold of fixed point 1 intersects the hyper-plane $z_L = -z_R, a_L = a_R$, this yields, by symmetry, a heteroclinic orbit to fixed point 2. Therefore we only need to intersect the hyper-plane to obtain a heteroclinic orbit, which yields two conditions (instead of four in the general case). For the sources we have now two conditions and two free parameters; and this

¹⁰Note that a one-dimensional manifold yields no free parameters other than the one associated with the trivial translation symmetry of the solution, and, in general, a p -dimensional outgoing manifold yields $p - 1$ nontrivial free parameters

yields a discrete set of $v=0$ sources. In other words, within the discrete set of sources we generically expect there to be a $v=0$ source solution.

For a sink we obtain, combining (3.B.2, 3.B.4) and (3.B.5), that fixed point 1 (a (N_+, L_+) combination) has a three-dimensional outgoing manifold, which yields two free parameters, while fixed point 2 (a (N_-, L_-) combination) has a three-dimensional outgoing manifold, which yields three conditions to be satisfied. Together with the three free parameters v, ω_L and ω_R , this yields a two-parameter family of sinks.

3.B.3 The role of ε

When discussing the counting for the single CGL equation, the value of ε is uniquely determined. In the coupled equations however, one needs to work with the *effective* value of ε when studying the linear fixed points, since the growth of the linear modes are determined by renormalized values of ε which are given by $\varepsilon_{\text{eff,L}} = \varepsilon - g_2 a_R^2$, $\varepsilon_{\text{eff,R}} = \varepsilon - g_2 a_L^2$ for the left- and right-traveling modes respectively [see Eq. (3.5.5)]. While the inclusion of the sign structure of the linear fixed points for positive values of ε may have seemed somewhat superfluous for the *single* CGL equation, in the case of the coupled equations this is relevant. In the analysis in sections 3.B.4–3.B.6 we assume that both effective values of ε are negative. Some comments on the counting for positive values of ε_{eff} are given in section 3.B.7.

3.B.4 The role of the coherent structure velocity v

In the counting for the single CGL equation, we were able to remove the group velocity term $\sim s_0$ by means of a Galilean transformation to the comoving frame. In the coupled equations this is not possible, however, and we need to incorporate the s_0 -terms when studying the fixed point structure.

In particular, when translating the result for the single CGL into coupled CGL variables, we need to make the following replacements where v is concerned

$$\text{For the } a_R \text{ mode} : v \rightarrow v - s_0 \equiv v_R, \quad (3.B.6)$$

$$\text{For the } a_L \text{ mode} : v \rightarrow v + s_0 \equiv v_L. \quad (3.B.7)$$

Just like the possible occurrence of positive values of ε could possibly affect the linear fixed points, this may well affect the nonlinear fixed points. In the single CGL equation we were allowed to take $v \geq 0$, but we can no longer do this in the coupled case. Let us focus on the case $v = 0$, i.e., consider stationary coherent structures. Since s_0 is by definition positive, the a_L mode has $v_L = s_0 > 0$, while the a_R has $v_R = -s_0 < 0$. The statement that we can always take $v > 0$ therefore no longer holds here, and we need to exercise caution when evaluating the nonlinear fixed points as well. In particular, *moving sources* ($v > 0$) with $|v_R| > v_{cN}$ or $v_L > v_{cN}$ can have different multiplicities than the stationary one with $v = 0$.

In the formulas for the counting, one should keep in mind that the linear group velocities have opposite signs for the left- and right moving modes: this is also apparent from Eqs. (3.4.5,3.4.6), where we defined $s_{0,R} = s_0 = -s_{0,L}$, so that we may write the nonlinear group velocities as

$$s_R = s_{0,R} + 2q_R(c_1 + c_3) , \quad s_L = s_{0,L} + 2q_L(c_1 + c_3) . \quad (3.B.8)$$

3.B.5 Normal sources always come in discrete sets

In this section, we show that it is not possible for normal stationary sources, i.e., sources whose s and s_0 have the same sign, and for whom $\varepsilon_{\text{eff}} < 0$ for the linear modes, to come in families. The flow for a normal source is

$$\text{from } \begin{cases} A_L : N_- & (v + s_0) \\ A_R : L_+ & (v - s_0) \end{cases} \quad \text{to } \begin{cases} A_L : L_- & (v + s_0) \\ A_R : N_+ & (v - s_0) \end{cases} . \quad (3.B.9)$$

According to the counting, we have for the $N_-(v + s_0)$ fixed point on the left that (we take $v = 0$)

$$\begin{aligned} p_0 &= 4a_L^2 q_L (c_1 + c_3) - 2a_L^2 v_L = 2a_L^2 [-s_0 + 2q_L (c_1 + c_3)] , \\ &= 2a_L^2 s_L < 0 , \end{aligned} \quad (3.B.10)$$

because for a normal source s_L has the same sign as $s_{0,L}$. Furthermore we have

$$p_2 = 2v_L = 2s_0 > 0 . \quad (3.B.11)$$

This implies, according to Eq. (3.A.20), that the sign structure of the left fixed point is a $(N_-(+, -, -), L_+(+, -, -))$ combination, independent

of the selected wavenumber of the nonlinear mode and the sign of the combination $p_1 p_2 - p_0 p_3$. The dimension of the outgoing manifold is therefore always equal to 2, yielding one free parameter. For the right fixed point, a completely similar argument yields an $(N_+(+, +, -), L_-(+, +, -))$ fixed point, again independent of the selected wavenumber or $\text{sgn}[p_1 p_2 - p_0 p_3]$. We therefore have to satisfy 4 conditions at this fixed point.

Combining this with the free parameters we already had and the additional symmetry at $v = 0$ we find that the sources *always* come in discrete sets, independent of the selected wavenumbers and the parameters.

3.B.6 Counting for anomalous $v = 0$ sources

When the signs of the linear group velocity s_0 and the nonlinear group velocity s are opposite, we are dealing with anomalous sources. This section investigates the consequences this has for the counting of such sources.

For an anomalous source, cf. Fig. 3.3b, the flow is (again we only consider $\varepsilon_{\text{eff}} < 0$ for the linear modes)

$$\text{from } \begin{cases} A_L : L_+ & (v + s_0) \\ A_R : N_- & (v - s_0) \end{cases} \quad \text{to} \quad \begin{cases} A_L : N_+ & (v + s_0) \\ A_R : L_- & (v - s_0) \end{cases}, \quad (3.B.12)$$

which yields for the nonlinear fixed point on the left

$$\begin{aligned} p_0 &= 4a_R^2 q_R (c_1 + c_3) - 2a_R^2 v_R = 2a_R^2 [s_0 + 2q_R (c_1 + c_3)], \\ &= 2a_R^2 s_R < 0. \end{aligned} \quad (3.B.13)$$

where $\text{sgn}[s_R] = -\text{sgn}[s_{0,R}]$. Furthermore

$$p_2 = 2v_R = -2s_0 < 0, \quad (3.B.14)$$

so that both p_0 and p_2 are negative, which implies that, according to Eq. (3.A.20), the sign structure of the N_- fixed point depends on $\text{sgn}[p_1 p_2 - p_0 p_3]$. In particular, when $p_1 p_2 - p_0 p_3$ is negative it is $N_-(+, +, +)$, and if it is positive it is $N_-(+, -, -)$. If $p_1 p_2 - p_0 p_3 < 0$, we can perform a similar calculation for the right fixed point, and we find that the counting then yields a 2-parameter family of anomalous sinks. If the expression is positive, however, we find that the anomalous sources also come in a discrete set.

The sign of this expression depends, for any given set of coefficients, on the selected wavenumber q_{sel} of the nonlinear mode, and therefore the

wavenumber selection mechanism will determine whether we can actually get to a regime where sources come as a family. In practice, we have not found any examples where this happens. This suggests to us that the possible regions of parameters space where this might happen, are small.

3.B.7 Counting for anomalous structures with $\varepsilon_{\text{eff}} > 0$ for the suppressed mode

As mentioned before, another situation that can change the counting is realized when the suppression of the effective ε by the nonlinear mode is not sufficiently large at the linear fixed points, so that $\varepsilon_{\text{eff}} > 0$. If we restrict ourselves to the $v = 0$ case, Eq. (3.A.15) tell us that the counting may indeed change when in addition $|s_0| > v_{cL}$. This implies that the multiplicity of sources and sinks changes dramatically under these circumstances. An insufficient suppression may happen in particular when g_2 is only slightly bigger than 1, while the selected wavenumber is large enough to lower the asymptotic value of the nonlinear amplitude significantly below its maximal value $\sqrt{\varepsilon}$. The zero mode then no longer stays suppressed; instead, it starts to grow, and we then typically get chaotic dynamics, see, e.g., section 4.3. For this reason, we confine ourselves to a few brief observations concerning the $v = 0$ case.

For $v = 0$ and $\varepsilon_{\text{eff}} > 0$, we can, according to Eq. (3.A.15), have a $L_-(- - -)$ fixed point of the A_L mode when $s_0 > v_{cL}$. The A_R mode then has a $L_+(+, +, +)$ fixed point. Since the index of L denotes the sign of the asymptotic value of κ , with these fixed points we could in principle build a 2-parameter family of stationary sources, provided s and s_0 have the same sign in the nonlinear region; otherwise the structures would be anomalous sinks.

Although we have not pursued the possible properties of such sources, we expect almost all members of this double family to be unstable. The reason for this is that when ε_{eff} is positive, the dynamics of the leading edge of the suppressed mode is essentially like that of a front propagating into an unstable state. As is well known [21], in that case there is also a 2-parameter family of fronts in the CGL equation, but almost all of them are dynamically irrelevant.

3.C Asymptotic behavior of sinks for $\varepsilon \rightarrow 0$

In this appendix, we will discuss the scaling of the width of sinks in the small- ε limit.

We will assume that in the domain to the left of the sink, the A_R -mode is suppressed, i.e., $\varepsilon_{\text{eff}}^L < 0$ (likewise to the right of the sink). As will be discussed in section 4.3 below, we may get anomalous behavior when $\varepsilon_{\text{eff}} > 0$, which can occur when $g_2 a_R^2 < \varepsilon$; in that case the A_L mode is (weakly) unstable and various types of disordered behavior occur.

Assuming $\varepsilon_{\text{eff}}^L$ to be negative to the left of a sink, the amplitude of the left-traveling mode grows exponentially for increasing ξ as $|A_L|(\xi) \sim e^{\kappa_L^+ \xi}$. The spatial growth rate κ_L is given, by definition, by the value of κ at the linear fixed point. According to Eq. (3.A.12), one finds for $z_L = \kappa_L + iq_L$:

$$z_L = \frac{-(v + s_0) \pm \sqrt{(v + s_0)^2 - 4(1 + ic_1)(\varepsilon_{\text{eff},L} + i\omega)}}{2(1 + ic_1)}, \quad (3.C.1)$$

where we have used the fact that for the left-traveling mode, v as used in the appendix is replaced by $v + s_0$, and $\varepsilon_{\text{eff},L} = \varepsilon - g_2 a_R^2$. If we expand the square-root in the small ε regime, where ω also tends to zero, we obtain

$$z_L \approx \frac{-(v + s_0)}{2(1 + ic_1)} \pm \frac{(v + s_0)}{2(1 + ic_1)} \left[1 - \frac{2(1 + ic_1)(\varepsilon_{\text{eff},L} + i\omega)}{(v + s_0)^2} \right]. \quad (3.C.2)$$

Since $\varepsilon_{\text{eff},L}$ is negative, and of order ε , the root z_L^+ with the positive real part is therefore

$$z_L^+ \approx \frac{-\varepsilon_{\text{eff},L} - i\omega}{(v + s_0)}, \quad (3.C.3)$$

so that κ_L^+ scales with ε as

$$\kappa_L^+ = \text{Re}[z_L^+] \sim \varepsilon. \quad (3.C.4)$$

In order for the exponent in $|A_L(\xi)| \sim e^{\kappa_L^+ \xi}$ to be of order unity, $\xi \sim \kappa_L^{+^{-1}} \sim \varepsilon^{-1}$, which shows that the width of the sinks will asymptotically scale as ε^{-1} for small ε .

Dynamical Properties of Source/Sink Patterns

4.1 Introduction

In the previous section, we have seen that in the parameter regimes where phase winding solutions to the coupled CGL equations are stable, sources and sinks are the relevant coherent structures to look at. The sources in particular are important, as they provide the mechanism for wavenumber selection in this system. Furthermore, we have seen that the source core undergoes an instability for values of the control parameter ε below an analytically known $\varepsilon_c^{\text{so}}$. Beyond this value, the stationary source solution disappears altogether, and is superseded by a dynamical structure, which executes a breathing motion. As we shall see however, there are at least two other mechanisms that lead to nontrivial dynamics of source/sink patterns, and this Chapter is devoted to a description of such states. Due to the high dimensionality of the parameter space (one has to consider, in principle, the coefficients c_1, c_2, c_3, g_2 and ε or s_0), we aim at presenting some typical examples and uncovering general mechanisms, rather than attempting a complete overview.

The starting point of our analysis here is the discrete nature of the sources (see Section 3.B.2), which implies that the wavenumber of the laminar patches is often uniquely determined [69, 67, 70].

As we have already briefly seen in Section 2.7, the stability of such waves depends on their wavenumber in a single CGL equation. Things

are no different here, and the first instability mechanism is therefore the Benjamin-Feir instability. When the waves emitted by the sources are unstable to long wavelength modes, it is the nature of this instability, i.e., whether it is *convective* or *absolute*, that determines the global dynamical behavior. The dynamical states that occur in this case are discussed in section 4.2.

What we have also seen for the coupled CGL equations, however, is that the selected wavenumber can alter the *effective* value of ε that the other mode experiences. Generally, the selected wavenumber is such that the suppressed mode in a single mode patch (*i.e.*, the mode that is zero), is *indeed* suppressed and feels a negative effective ε . In some cases however, this is not the case and the system can yield to a *bimodal* instability. The essential observation is that for a selected wavenumber q_{sel} there exists a range $1 < g_2 < \varepsilon/(\varepsilon - q_{\text{sel}}^2)$ for which *both* single and bimodal states are unstable. Provided that there are sources in the system, we find then a regime of *source-induced bimodal* chaos (see section 4.3). This type of chaos is indeed source-induced, as the selected wavenumber tunes the instability.

Furthermore, both of these instabilities can occur simultaneously, as seems to be the case in experiments of the Saclay group [71], and both can be combined with the small- ε instability of the sources, discussed in section 3.6. This leads to quite a rich palette of dynamical and chaotic states (section 4.4). We have summarized the various disordered states that are typical for the coupled amplitude equations in Table 4.1 above. The first three types of dynamics are source-driven. Sources are not essential for the last three types of dynamics, which are driven by the coupling between the A_L and A_R modes. We will discuss all of them briefly in this chapter, and give examples of what these dynamical states typically look like.

4.2 Convective and absolute sideband-instabilities

Let us begin by recalling a result first quoted in Chapter 2.7, which states that plane waves in the single CGL equation with wavenumber q exhibit

Table 4.1: Overview of disordered and chaotic states.

Type	Section	Fig.	Parameters
Core-instabilities	3.6.1,3.6.2	3.6	$\varepsilon < \varepsilon_c^{so} = s_0^2/(4 + 4c_1^2)$
Absolute instabilities	4.2	4.3,4.4	$v_{BF}^* > 0$
Bimodal chaos	4.3	4.5	$1 < g_2 < \varepsilon/(\varepsilon - q_{sel})$
Defects + Bimodal	4.4.2	4.6	g_2 just above 1
Intermittent + Bimodal	4.4.3	4.7	g_2 just above 1
Periodic patterns	4.4.4	4.3,4.4,4.8	c_2, c_3 : opposite signs and not small

sideband instabilities when [24]¹

$$q^2 > \frac{\varepsilon(1 - c_1 c_3)}{3 - c_1 c_3 + 2c_3^2}, \quad (4.2.1)$$

and when the curve $c_1 c_3 = 1$ (the Newell line) is crossed, all plane waves become unstable, and one encounters various types of spatio-temporal chaos [24, 72, 73, 74]. For the coupled CGL equations under consideration here, the condition for linear stability of a single mode is still given by Eq. (4.2.1), since the mode which is suppressed is coupled quadratically to the one which is nonzero. Since the sources in general select a wavenumber unequal to zero, the relevant stability boundary for the plane waves in source/sink patterns typically lies below the $c_1 c_3 = 1$ curve.

Consider now a linearly unstable plane wave. Perturbations of this wave grow, spread and are advected by the group velocity. The instability of the wave is called convective when the perturbations are advected away faster than they grow and spread; when monitored at a fixed position, all perturbations eventually decay. In the case of absolute instability, the perturbations spread faster than they are advected; such an instability often results in persistent dynamics. To distinguish between these two cases one has to compare, therefore, the group velocity and the spreading velocity of perturbations. In Fig. 4.1, we sketch qualitatively the difference between absolute and convective instability. For a general introduction

¹When both numerator and denominator are negative, as may occur for large c_1 , this equation seems to suggest that one might have a stable band of wavenumbers. However, when $1 - c_1 c_3$ is negative, no waves are stable; the flipping of the sign of the denominator for large c_1 bears no physical relevance, but is due to a long-wavelength expansion performed to obtain Eq. (4.2.1). Note that the denominator is always positive as long as $1 - c_1 c_3$ is positive.

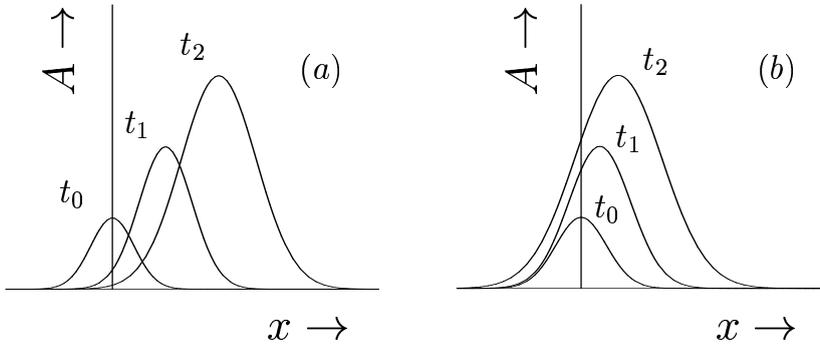


Figure 4.1: Convective (a) and absolute (b) instability. Plotted is the growth and advection of a perturbation for subsequent times $t_0 < t_1 < t_2$. In the first case, perturbations are advected faster than they grow, while in the latter the converse is true.

to the concepts of convective and absolute instabilities, the reader might want to confer [75]. Numerical simulations of the coupled CGL equations presented below show that the distinction between the two types of instabilities is important for the dynamical behavior of the source/sink patterns. When the waves that are selected by the sources are convectively unstable, we find that, after transients have died out, the pattern typically “freezes” in an irregular juxtaposition of stationary sources and sinks. When the waves are absolutely unstable², however, persistent chaos occurs.

The wavenumber selection and instability scenario sketched above for the coupled CGL equations is essentially the one-dimensional analogue to the “vortex-glass” and defect chaos states in the 2D CGL equation [76, 77]; in that case the wavenumber is selected by so-called spiral or vortex solutions. As we shall discuss, there are, however, also some differences between these cases.

Below we will briefly indicate how the threshold between absolute and convective instabilities is calculated (see also [77]). The advection of a small perturbation is given by the nonlinear group velocity $s = \partial\omega/\partial q$

²It should be noted that the criterion for absolute instability concerns the propagation of perturbations in an ideal, homogeneous background. For typical source/sink patterns, one has finite patches; the criterion can also not determine when perturbations are strong enough to really affect the core of the sources. Analogous to the 2D case, we have found that persistent dynamics sets in slightly *above* the threshold between convective and absolute instabilities.

which is the sum of the linear group velocity s_0 and the nonlinear term $s_q := 2q(c_1 + c_3)$:

$$s_L = -s_0 + 2q_L(c_1 + c_3), \quad s_R = s_0 + 2q_R(c_1 + c_3). \quad (4.2.2)$$

The spreading velocity of perturbations is conveniently calculated in the linear marginal stability/pulled front framework (We will just use the results here, but the framework will be discussed in more detail in Chapter 5), once one has obtained a dispersion relation for these perturbations. Since we consider single mode patches, we are allowed to restrict ourselves to a single CGL equation, in which the linear group velocity term $\pm s_0 \partial_x A$ is easily incorporated, as it just gives a constant boost. Considering a perturbed plane wave of the form $A = (a + u) \exp i(qx - \omega t)$, where u is a small complex-valued perturbation $\sim \exp i(kx - \sigma t)$ and $a^2 = \varepsilon - q^2$. Upon substituting this Ansatz into a single CGL equation, linearizing and going to a Fourier representation, one obtains a dispersion relation $\sigma(k)$ [78]. From this relation one then finally calculates the spreading velocity v_{BF}^* of the Benjamin-Feir perturbations in the linear marginal stability or saddle-point framework [53].

Since in general we can only calculate the selected wavenumber q by a shooting procedure of the ODE's (3.5.2a-3.5.2d) for a source, obtaining a full overview of the stability of the plane waves as a function of the coefficients necessarily involves extensive numerical calculations. Therefore, we will focus now on a single sweep of c_2 . For reasons to be made clear below, we choose $\varepsilon = 1, c_1 = c_3 = 0.9, s_0 = 0.1$ and $g_2 = 2$. Since we fix all coefficients but c_2 , the stability boundary (4.2.1) is fixed. By sweeping c_2 , the selected wavenumber varies over a range of order 1, and one encounters both convective and absolute instabilities.

We have found that after a transient, patterns in the stable or convectively unstable case are indistinguishable³. When there is no inherent source of noise or perturbations, there is nothing that can be amplified, and the convective instability is rendered powerless (see however, section 4.4).

Although the transition between stable and convectively unstable waves is not very relevant for the source/sinks patterns here, the transition between convectively and absolutely unstable waves is interesting. To obtain an absolute instability one needs to carefully choose the parameters; when

³Except, of course, when we prepare a very large system with widely separated sources and sinks.

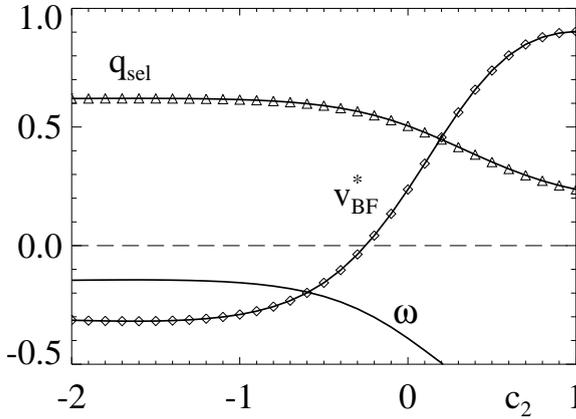


Figure 4.2: Frequency ω , corresponding selected wavenumber q_{sel} and perturbation velocity v_{BF}^* as a function of c_2 , for $\varepsilon=1$, $c_1=c_3=0.9$, $s_0=0.1$ and $g_2=2$. For $c_2 < -0.25$, $v_{BF}^* < 0$, and perturbations in the right flank of the source propagate to the left, so that the waves are absolutely unstable.

q increases, the contribution to the group velocity of the nonlinear term s_q increases, and we have to take c_1 and c_3 quite close to the $c_1 c_3 = 1$ curve to find absolute instabilities. This is the reason for our choice of coefficients. In Fig. 4.2 we have plotted the selected frequency (obtained by shooting), corresponding wavenumber and propagation velocity v_{BF}^* of the mode to the right of the source, as a function of c_2 . For this choice of coefficients the single mode waves turn Benjamin-Feir convectively unstable when, accordingly to Eq. (4.2.1) $|q| > 0.223$, which is the case for all values of c_2 shown in Fig. 4.2. The waves turn absolutely unstable when $|q| > 0.553$, and this yields that the waves become absolutely unstable for $c_2 < -0.25$.

When the selected waves becomes absolutely unstable, the sources may be destroyed since perturbations can no longer be advected away from them; the system typically ends up in a chaotic state. In Fig. 4.3 we show what happens when we choose the coefficients as in Fig. 4.2, and decrease c_2 deeper and deeper into the absolutely unstable regime. All runs start from random initial conditions, and a transient of $t = 10^4$ was deleted. Although the left- and right traveling waves do not totally suppress each other, it was found that pictures of $|A_L|$ and $|A_R|$ are, to

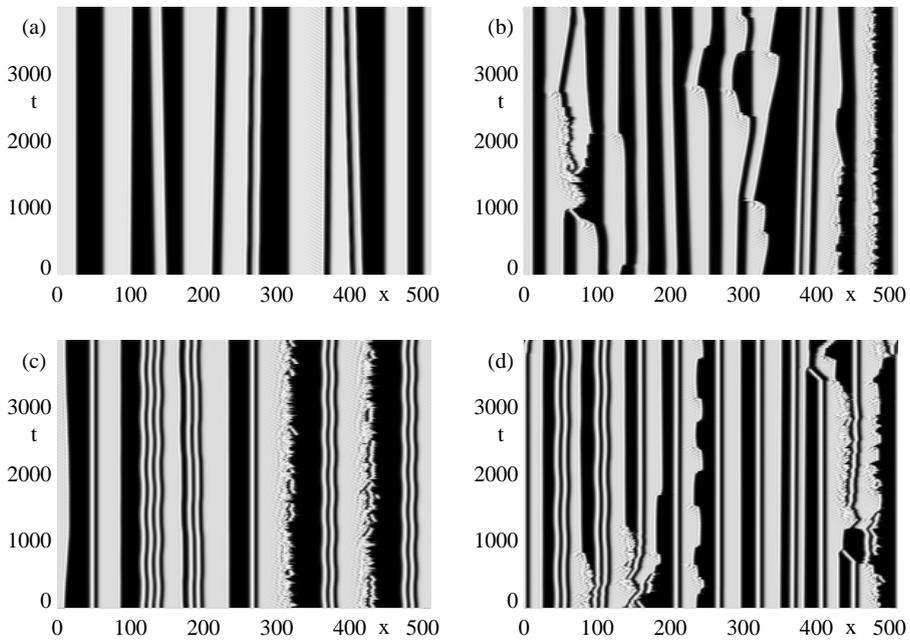


Figure 4.3: Source/sink patterns with absolutely unstable selected wavenumbers for the same coefficients as in Fig. 4.2 and various values of c_2 . (a) $c_2 = -0.3$, (b) $c_2 = -0.4$, (c) $c_2 = -0.6$, (d) $c_2 = -0.8$. For more information see text.

within good approximation, each others negative (see also the final states in Fig. 4.4). In accordance with this, we choose our greyscale coding to correspond to $|A_R|$, such that light areas corresponds to right-traveling waves and dark ones to left-traveling waves.

In Fig. 4.3(a), $c_2 = -0.3$ and the waves have just turned absolutely unstable, but the only nontrivial dynamics is a very slow drift of some of the sources and sinks. Note that this does not invalidate our counting results that isolated sources are typically stationary, because the drifting occurs only for structures that are close together. When c_2 is lowered to -0.4 (Fig. 4.3(b)), one can see now the Benjamin-Feir perturbations spreading out in the opposite direction of the group velocity, eventually affecting the sources (for example around $x = 230, t = 2700$). Some of the sinks become very irregular. When c_2 is decreased even further to -0.6 (Fig. 4.3(c)), the sources and sinks show a tendency to form periodic states [79] (see also Fig. 4.4). These states seem at most weakly unstable since only some very mild oscillations are observed. The two sinks with the largest patches around them show most dynamics, and one sees the irregular creation and annihilation of small source/sink pairs here (around $x = 320$ and 440). Finally, when c_2 is decreased to -0.8 (Fig. 4.3(d)) the state becomes more and more disordered; the irregular “jumping” sink at $x \approx 230$ is worth noting here.

It is interesting to note that, in particular for large negative c_2 closely bound, uniformly drifting sink-source pairs are formed (see for instance around $x = 430, t = 700$ in Fig. 4.3(d)). Another frequently occurring type of solution are periodic states, corresponding to an array of alternating patches of A_L and A_R mode (see also Fig. 4.4). The source/sink pairs and in particular the periodic states occur over a quite wide range of coefficients; their existence has been reported before by Sakaguchi [79]. In a coherent structures framework, periodic states correspond to limit cycles of the ODE’s (3.5.2a-3.5.2d). In many cases they can be seen as strongly nonlinear standing waves, and they show an interesting destabilization route to chaos (see section 4.4.4).

Apart from the similarities between the mechanisms here and the spiral chaos of the 2D CGL equation, it is also enlightening to notice the differences. The first difference is that our sources, in contrast to the spirals in 2D, are not topologically stable. In the states we have shown so far this does not play a role; in the following section we will see examples where instabilities of the sources themselves play a role. While

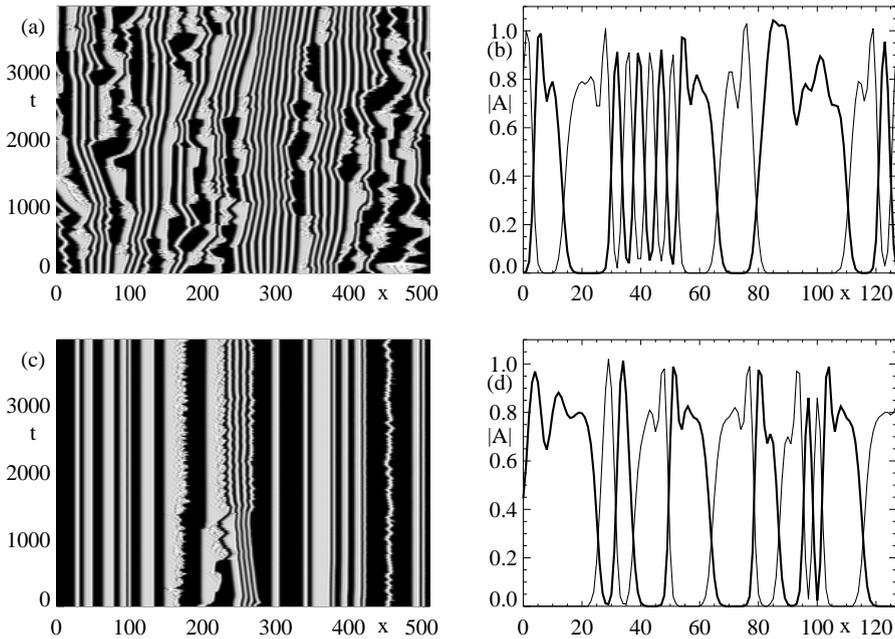


Figure 4.4: Two more examples of nontrivial dynamics in the absolutely unstable case. Both cases: $c_1 = c_3 = 0.9$, $c_2 = -2.6$, $g_2 = 2$, and a transient of 10^4 is deleted. (a-b): $s_0 = 0.1$. Here the periodic states are quite dominant. It appears that these states themselves are prone to drifting and slow dynamics. (b) Snapshots of $|A_L|$ (thick curve) and $|A_R|$ (thin curve) in the final state. Obviously, the two modes, although disordered, suppress each other completely. (c-d) Here we have increased s_0 to 0.2. The plane waves are still absolutely unstable, and the dynamics is disordered, but much less than in case (a-b).

in the 2D case the spiral cores that play the role of a source are created and annihilated in pairs, it is here only the sources and sinks that are created or annihilated in pairs. Furthermore, in the spiral case, the linear analysis that determines whether the waves are absolutely or convectively unstable is performed for plane waves. This means one neglects curvature corrections of the order $1/r$, where r is the distance to the core of the source. Here, the only correction comes from the asymptotic, exponential approach of the wave to a plane wave; this exponential decay rate is given by the decay rate κ (see the appendix). Finally, in the spiral case, for fixed c_1 and c_3 , both the group velocity and the selected wavenumber are fixed, while here the selected wavenumber can be tuned by c_2 , without influencing the stability boundaries of the single mode state. The group velocity can be tuned by s_0 . Although the selected wavenumber influences the group velocity, cf. Eqs. (4.2.2), and s_0 influences the selected wavenumber, this large number of coefficients gives us more freedom to tune the instabilities.

4.3 Instability to bimodal states: source-induced bimodal chaos

The dynamics we study in this section are intrinsically due to a competition between the single source-selected waves and bimodal states. Therefore, this state is in an essential way different from what can be found in a single CGL equation framework.

The wavenumber selection by the sources is of importance to understand the competition between single mode and bimodal states. In the usual stability analysis of the single mode and bimodal states, it is assumed that both the A_L and A_R modes have equal wavenumber [80]. Therefore, this analysis does not apply to the case of a single mode, say the right-traveling mode, with nonzero wavenumber. The left-traveling mode is then in the zero amplitude state and has no well-defined wavenumber; one should consider therefore its fastest growing mode, i.e., a wavenumber of zero. The following, limited analysis, already shows that for g_2 just above 1, instabilities are expected to occur. Restricting ourselves to long wavelength instabilities, the analysis is simply as follows. Write the left- and right-traveling waves as the product of a time dependent amplitude

and a plane wave solution:

$$A_L = a_L(t)e^{i(q_L x - \omega_L t)}, \quad A_R = a_R(t)e^{i(q_R x - \omega_R t)}, \quad (4.3.1)$$

and substitute this Ansatz in the coupled CGL equations. One obtains then the following set of ODE's

$$\partial_t a_L = (\varepsilon - q_L^2 - a_L^2 - g_2 a_R^2) a_L, \quad (4.3.2a)$$

$$\partial_t a_R = (\varepsilon - q_R^2 - a_R^2 - g_2 a_L^2) a_R. \quad (4.3.2b)$$

Consider the single mode state with $a_R \neq 0, a_L = 0$ and take $q_L = 0$. The maximum linear growth rate of a_L now follows from Eq. (4.3.2a) to be the one with $q_L = 0$; this mode has a growth rate given by $\varepsilon - g_2 a_R^2 = \varepsilon - g_2(\varepsilon - q_R^2)$. From this it follows that a single mode state with wavenumber q_R is unstable when $g_2 < \varepsilon/(\varepsilon - q_R^2)$. In source/sink patterns, the selected wavenumber can get as large as $\sqrt{\varepsilon/3}$ at the edge of the stability band for $c_1 = c_3 = 0$; it is as large as $0.6\sqrt{\varepsilon}$ in Fig. 4.2. In extreme cases, the value of g_2 necessary to stabilize plane waves can be at least 50% larger than the value 1 that one would expect naively.

On the other hand, the stability analysis of the bimodal states shows that they are certainly unstable for $g_2 > 1$. A naive analysis for general q_L and q_R , based on Eqs. (4.3.2a) can be performed as follows. Solving the fixed point equations of Eqs. (4.3.2a) for the bimodal state (i.e., a_L and a_R both unequal to zero), and linearizing around this fixed point yields a 2×2 matrix. From an inspection of the eigenvalues we find that the bimodal states turn unstable when $g_2 < \varepsilon - q_1^2/(\varepsilon - q_2^2)$, where q_1 is the largest and q_2 is the smallest of the wavenumbers q_L, q_R . When both wavenumbers are equal this critical value of g_2 is one; it is smaller in general.

It should be noted that this analysis does not capture sideband instabilities that may occur, and therefore waves in a much wider range might be unstable. For sideband-instabilities of bimodal states, the reader may consult [80] and [81]. However, our analysis shows already that there is certainly a regime around $g_2 = 1$ where *both* the single and bimodal states are unstable. This regime at least includes the range

$$1 < g_2 < \varepsilon/(\varepsilon - q_{\text{sel}}^2) \quad (4.3.3)$$

The distinction between convective and absolute instabilities becomes slightly blurred here. Suppose for instance we inspect a single-mode state that turns unstable against bimodal perturbations. Initially, these perturbation will be advected by the group velocity of the nonlinear mode,

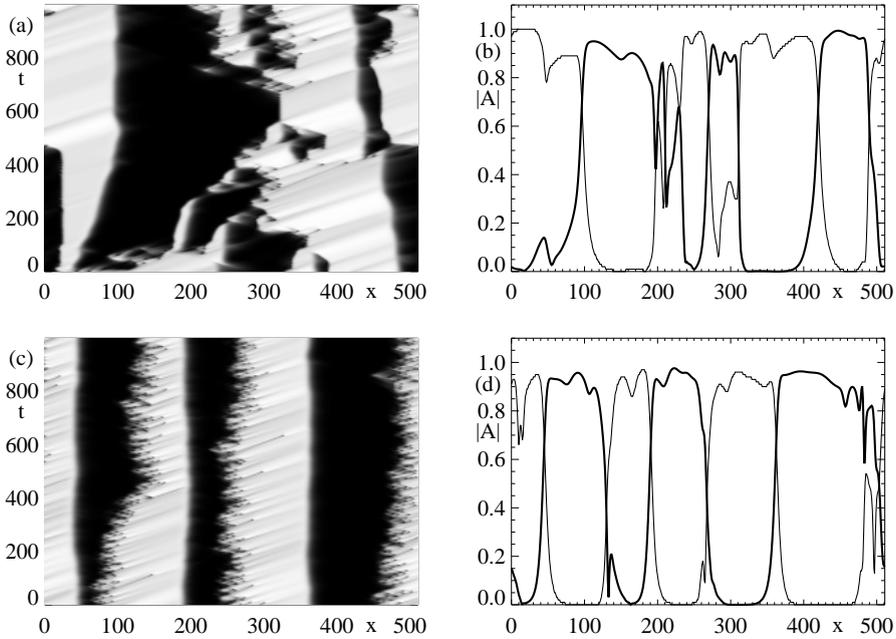


Figure 4.5: Two examples of bimodal chaos. (a) and (c) show space time plots, and the grey shading is the same as before. Both simulations started from random initial conditions, and a transient of $t=10^4$ has been deleted from these pictures. For a detailed description, see text. Note that the final states of runs (a) and (c), depicted in (b) and (d), clearly show that the two modes no longer suppress each other completely.

but as the perturbations grow, both modes will start to play a role, and since they feel a group velocity of opposite sign, the perturbations are effectively slowed down. Roughly speaking, the instability might be linearly convectively unstable but nonlinearly absolutely unstable [75].

Without going into further detail we will now show two examples of the bimodal chaos that occurs when g_2 is just above 1. For examples of similar dynamics, also for $g_2 < 1$, see [81]. In the first example (Fig. 4.5(a-b)) we have taken $\varepsilon = 1, c_1 = c_3 = 0.5, c_2 = -0.7, s_0 = 1$ and $g_2 = 1.1$. The selected wavenumber is almost independent of the value of g_2 and approximately equal to 0.35, which yields a critical value of g_2 of 1.14. For g_2 just below this value, the instability appears convective, and after a transient the system ends up in a mildly fluctuating source/sink pattern. When g_2 is

decreased, the instability becomes stronger and, presumably, absolute in nature. The *sources* behave then very irregularly, while the sinks drift according to their incoming, disordered waves. Note that sources and sinks are created and annihilated in this state. In Fig. 4.5(c-d) we show the disordered dynamics for $\varepsilon = 1, c_1 = 1, c_3 = -1, c_2 = 1, s_0 = 0.5$ and $g_2 = 1.1$. Note that in the laminar patches, since $c_1 = -c_3$, the dynamics is relaxational [24, 25]. In this state, no creation or annihilation of sources and sinks is found; the sinks drift slowly, while the sources behave very irregularly.

The dynamical states as shown in Fig. 4.5 are different from the chaotic states that we are familiar with from the single CGL equation [72, 73, 74, 22], and so they are of some interest in their own right. Note that it is possible to get persistent dynamics for values of c_1 and c_3 that in a single CGL equation-framework would lead to completely orderly dynamics. As the two examples in Fig. 4.5 show, qualitatively different states seem to be possible in this regime; the question of classification of the various dynamical states is completely open as far as we are aware.

Finally, it should be pointed out that when, as is the case here, the left- and right-traveling mode no longer suppress each other, ε_{eff} becomes positive. In principle this might change the multiplicity of the sources, since the eigenvalues coming from the linear fixed point can have a different structure for positive ε_{eff} (see appendix 3.B.7). However, this is only true when the effective velocity $v \pm s_0$ is larger than the critical velocity v_{cL} ; for the cases considered above, this does not happen. Hence, the sources are here still unique and select a unique wavenumber.

4.4 Mixed mechanisms

In the previous sections we have described three mechanisms by which sink/source patterns can be destabilized. First of all, in section 3.6 we found that due to a competition between the linear group velocity s_0 and the propagation of linear fronts, the cores of the sources become unstable when $\varepsilon < \varepsilon_c^{\text{so}}$. In section 4.2 we have shown that the waves that are sent out by the sources can be convectively or even absolutely unstable, and in section 4.3 we found that these waves may also be unstable to bimodal perturbations when g_2 is not very far above 1. Since the mechanisms that lead to these instabilities are independent, these instabilities might occur together. This is the subject of this section. In particular, one can in

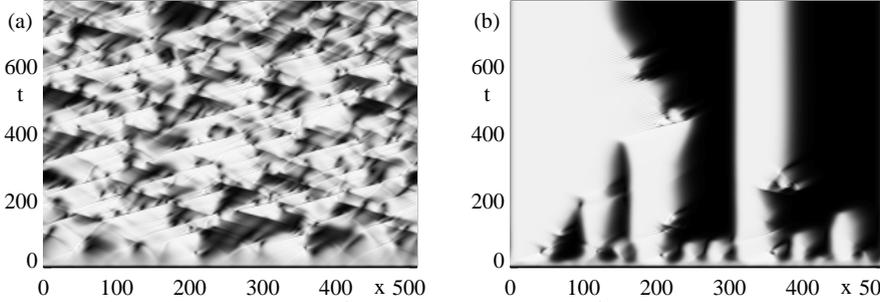


Figure 4.6: Two examples of the combination of phase slips and a value of g_2 just above 1. The coefficients are $c_1 = 1, c_3 = 1.4, c_2 = 1, \varepsilon = 1, s_0 = 0.5$. Grey shading as before (right (left) traveling waves are light (dark)). In (a), $g_2 = 1.05$, while in (b) $g_2 = 1.2$.

principle always lower the control parameter ε in an experiment to make the sources become core-unstable (section 4.4.1). A second combination of instabilities occurs when g_2 is close to 1 and the plane waves are unstable and generate phase slips (section 4.4.2); a particular interesting case occurs when the single mode waves are in the so-called intermittent regime (section 4.4.3).

4.4.1 Core instabilities and unstable waves

As discussed in section 3.6.2, the cores of the source may start to fluctuate when $\varepsilon < \varepsilon_c^{so}$. As is visible in Fig. 3.6(c), the perturbations that are generated in the core are then advected away into the asymptotic plane waves. In the discussions in section 3.6 above, we have focused on the case where these waves are stable, but obviously, when they are unstable, this will amplify the perturbations emitted by the source core. In particular, when the waves are convectively unstable, a stable core for $\varepsilon > \varepsilon_c^{so}$ leads to stationary patterns, but a fluctuating core can fuel the convective instabilities. This yields a simple experimental protocol to check for convective instabilities; simply lower ε and follow the perturbations sent by the sources for $\varepsilon > \varepsilon_c^{so}$.

4.4.2 Phase slips and bimodal instabilities

Let us for definiteness suppose we have that $A_L = 0$, and the right-traveling mode is active. When this A_R mode is chaotic and displays phase slips, the

effective growth rate of the A_L mode, $\varepsilon_{\text{eff}}^L$, may become positive for some period. A_L only grows during this period; it depends then on the duration and spatial extension of the positive $\varepsilon_{\text{eff}}^L$ “pocket” whether A_L can grow on average. Clearly, one should look at a properly averaged value of $\varepsilon_{\text{eff}}^L$, and therefore at the averages of $\varepsilon - g_2 a_R^2$ [79]. When g_2 is sufficiently large, the averaged effective growth rate always becomes negative, so that even a heavily phase slipping wave can still suppress its counter-propagating partner.

We show two examples of the dynamics when phase slips occur and g_2 is not large enough to strictly suppress the near-zero mode. As coefficients we choose $c_1 = 1, c_3 = 1.4, c_2 = 1, \varepsilon = 1, s_0 = 0.5$, and the dynamics is illustrated in Figs. 4.6. It should be noted that in Fig. 4.6(b) the sources are stationary, while some of the sinks drift. This seems to be due to the fact that near the sink, i.e., far away from the sources, the wave emitted by the sources has undergone phase slips, and the incoming wavenumbers of the sink can therefore be different from the source-selected wavenumbers. For slightly different coefficients we have observed patterns of stationary sources, with sinks in between that by this mechanism move in zig-zag fashion, i.e., alternating to the left and to the right.

4.4.3 Intermittency and bimodal instabilities

Recently, Amengual *et al.* studied the case of spatio-temporal intermittency in the coupled CGL equations for a linear group velocity $s_0 = 0$ and $c_2 = c_3$ [82]. This particular sub-case of the coupled CGL equations is of importance in the description of some laser systems [82, 68]. When g_2 is increased from zero, the authors of [82] found that for $g_2 < 1$ one finds intermittency, with the A_L and A_R obviously becoming more and more correlated as the cross-coupling increases. Furthermore, the authors observed that for $g_2 > 1$, the two modes become “synchronized”, i.e., the intermittency disappears and the systems ends up in a state that we recognize now as a stationary source/source pattern (not source/sink, see below). Since the intermittency “disappears” the authors question the applicability of a single CGL equation for patches of single modes in the coupled CGL equations (3.3.3a-3.3.3b).

The purpose of this section is to clarify, correct and extend their results, using our results for the wavenumber selection, the bimodal instabilities and the discussion in section 4.4.2. In particular we will show that, (i) for sufficiently large g_2 , the intermittency can persist, (ii) when the

intermittency disappears it can do so by at least two distinct mechanisms, *(iii)* more complicated states can occur. We conclude then that for single mode patches the single CGL is a correct description, provided one is sufficiently far away from bimodal instabilities and one takes the source-selected wavenumber and correct boundary conditions into account.

For the case considered in [82] the group-velocity s_0 is equal to zero, so the two modes A_L and A_R are completely equivalent. The distinction between sources and sinks depends therefore on the nonlinear group velocity, which follows from the selected wavenumber. The counting arguments yield in this case again a discrete $v=0$ source and a two parameter family of sinks (see section 3.5). In simulations we typically find stationary sources that separate the patches of A_L and A_R mode, and *single amplitude sinks* sandwiched in between these sources.

We will show now a variety of scenarios for intermittency in the coupled CGL equations (3.3.3a-3.3.3b). The coefficients used in [82] are $c_1 = 0.2, c_2 = c_3 = 2, \varepsilon = 1$ and $s_0 = 0$. The coefficients c_1 and c_3 are chosen such that a single mode is in the so-called intermittent regime. In this regime, depending on initial conditions, one may either obtain a plane wave attractor or a chaotic, “intermittent” state; the latter one is typically built up from propagating homoclinic holes and phase slips [72, 73, 74, 22].

In Fig. 4.7(a) we take $g_2 = 2$ and start from an ordered pair of sources. By a rapidly changing c_1 to a value of 1.2 and then back to the value 0.2, we generate phase slips that nucleate a typical intermittent state. This intermittent state persists for long times; there is no “synchronization” whatsoever. We found that we can also first let the source develop completely, and then introduce some phase slips; also in this case the intermittency clearly persists. To understand this, note that in this case g_2 is sufficiently large, and so ε_{eff} is negative (see Section 4.4.2); although there are phase slips, the two modes suppress each other completely.

In contrast, when g_2 is lowered, ε_{eff} can become positive, and this corresponds to a scenario described in [82]. In Fig. 4.7(b) we start from state obtained for $g_2=2$, and then quench g_2 to a value of 1.5. In this case, ε_{eff} becomes positive every now and then, and after a while, in the patch originally the exclusive domain of A_L , small blobs of A_R mode grow. After a sufficient period has elapsed, these blobs nucleate new sources, and the system ends up in a stationary source/source pattern. The laminar patches in between the sources are quite small and the intermittency disappears.

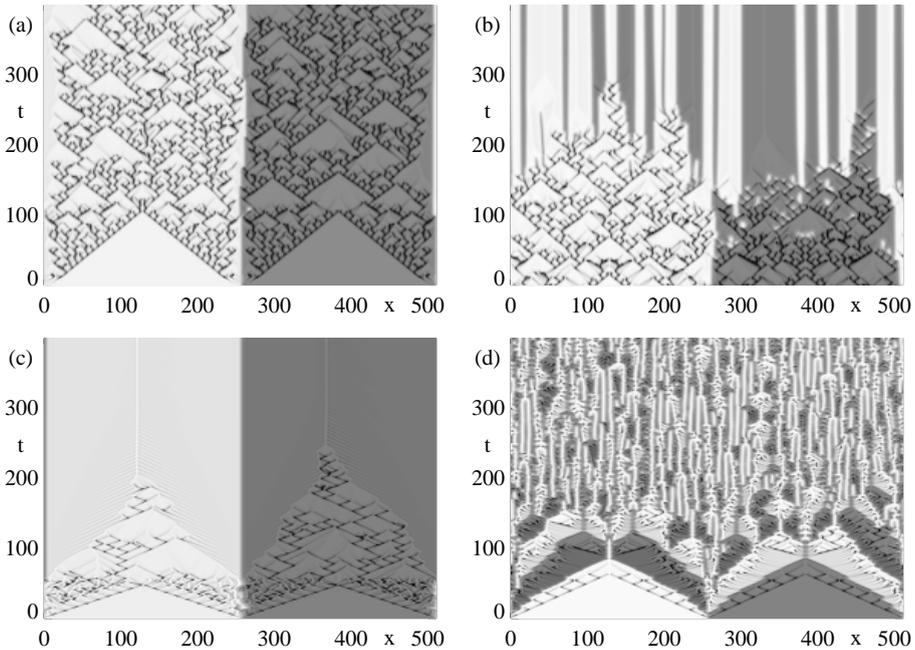


Figure 4.7: Space-time plots in the coupled-intermittent regime. To be able to show both the dynamics in the A_L and A_R mode, the grey shading corresponds to $2|A_R| + |A_L|$. This yields that right traveling patches are brighter in shade than left-traveling patches. (a) $c_1 = 0.2, c_2 = c_3 = 2, \varepsilon = 1, s_0 = 0$ and $g_2 = 1.2$. (b) Same coefficients as (a), except for $g_2 = 1.5$. (c) $c_1 = 0.6, c_3 = 1.4, c_2 = 1, \varepsilon = 1, s_0 = 0.1$ and $g_2 = 2$. (d) Same coefficients as (c), except for $c_2 = 0$. For as more detailed description see text.

The system switches from the intermittent to the plane wave attractor when the new sources are formed; this does not mean that the CGL equation is incorrect here, since both plane waves and intermittent states are attractors for these coefficients. The disappearance of the intermittency can be easily understood as follows: the main mechanism by which intermittency spreads through the single CGL equation is by the propagation of homoclinic holes that are connected by phase slip events [22]. If the phase slips now generate sources, there is no generation of new homoclinic holes and the intermittency dies out.

It should be noted that for this particular choice of the coefficients c_1 and c_3 , the homoclinic holes have a quite deep minimum in $|A|$, which increases the value of the average of ε_{eff} ; therefore one needs quite a large g_2 to guarantee the mutual suppression of the A_L and A_R modes.

Finally, we found that the selected wavenumber for the coefficients of this particular example is ≈ 0.1 . As a consequence, the transition to stationary domains as observed in [82] can *not* occur at g_2 precisely equal to 1, but occurs for $g_2 \approx 1.01$ (see section 4.3).

This generation of sources due to phase slips of the nonlinear mode is not the only way in which the intermittency can disappear. Considerer the example shown in Fig. 4.7(c). We have chosen the coefficients as $c_1 = 0.6$, $c_3 = 1.4$, $c_2 = 1$, $\varepsilon = 1$, $s_0 = 0.1$ and $g_2 = 2$. The sources select now a wavenumber of 0.3783, and the plane wave emitted by the source simply “eats up” the intermittent state; note the single amplitude sinks visible for late times. It should be realized that many dynamical states are sensitive to a background wavenumber, and that the spatio-temporal intermittent state is particularly sensitive to this [22]; when describing a patch in the coupled CGL equations by a single CGL equation, one should take into account that one has wave-selection at the boundaries due to the sources.

Finally, when c_2 is lowered to a value of 0, the sources themselves become unstable and the system displays the tendency to form periodic patterns; these are however not stable, and an example of the peculiar dynamical states one finds is shown in Fig. 4.7(d).

In conclusion, when one is far away from any bimodal instabilities, i.e., when g_2 is sufficiently large, a description in terms of a single CGL equation is sufficient for the patches separating the sources, provided one takes into account the group velocity, boundary effects and, most importantly, the selected wavenumber. It is amusing to note that the question under which conditions a single amplitude equation is a correct description of

these waves depends on the coefficients g_2 and c_2 of the *cross-coupling* term.

To clarify this situation, let us consider a large patch where the A_L mode is zero while the A_R mode is active; this is a typical situation when g_2 is sufficiently large. It is often stated loosely that a single CGL equation is a correct description for this A_R mode that occur when g_2 is larger than one. The question under what conditions such a single CGL equation applies is of considerable importance, since most theoretical studies have focused on this case while many experimental systems show both left- and right-traveling waves. Obviously, we should restrict ourselves to rather large patches; otherwise the effects from the edges of the domains are obviously too severe. Furthermore the group velocity terms cannot be transformed away due to these boundaries, so one should take that into account too.

4.4.4 Periodic and other states

We would like to conclude this section by showing an example of the wide range of different states that occur in the coupled amplitude equations when we sweep c_2 . We choose the other coefficients as follows: $g_2 = 1.1, c_1 = 0.9, c_3 = 2, s_0 = -0.1, \varepsilon = 1$. Our main finding is that for large positive or negative c_2 , there is no sustained dynamics, while for small c_2 we find a strongly chaotic state. In between there are at least two transitions between laminar and disordered state (see Figs. 4.8 and 4.9).

For sufficiently negative c_2 , all initial conditions evolve to a spatially periodic state, with rapidly alternating A_L and A_R patches. We can view these states as an example of highly nonlinear standing wave patterns. Depending on initial conditions, these states may either be stationary or have a small drift. For our particular choice of coefficients it is empirically found that these states are linearly stable for $c_2 \leq -0.72$. In Fig. 4.8(a) we see the evolution from a slightly perturbed initial condition for this value of c_2 . Qualitatively, we observe that when the “local wavenumber” of the standing wave is lowered, this leads to oscillations, that may or may not lead to “defects”. After some reasonably long transient (note the perturbation at $x \approx 320, t \approx 2600$), the dynamics settles down in a slowly drifting standing wave. This shows that these generalized standing waves are stable here.

In Fig. 4.8(b) we start from such a coherent standing wave state and have lowered c_2 to a value of -0.71 . In this case perturbations of the

waves are spontaneously formed, indicating a linear instability. Since the state is unstable, these perturbations then spread to the system in a way that is reminiscent of the intermittent patterns obtained, for instance, in experiments on intermittency in Rayleigh-Bénard convection [83]. It should be noted that, due to the instability of the laminar state, one does not have an absorbing state, so strictly speaking this state should not be referred to as intermittent. Interestingly enough, the transition between laminar and chaotic behavior seems to be second order, i.e., we could not find any hysteresis. The transition is simply triggered by the linear stability of the periodic/standing waves, and when these waves are stable, they are the only type of attractor.

If c_2 is further increased to a value of -0.5 (Fig. 4.8(c)), we find a state that we might call defect-chaos of a standing wave pattern. For $c_2=0$ (Fig. 4.8(d)), the dynamics evolves on much faster time-scales, and no clear structures are visible by eye.

On the other hand, when we keep increasing c_2 , we again find regular states, but these ones correspond to stationary source/sink patterns. This is illustrated in Fig. 4.9, where we show four space-time plots for increasing, positive values of c_2 . In comparison with the dynamics as shown in Fig. 4.8(d), the time scales become slower and slower when c_2 is increased. This slowing down becomes quite clear for $c_2=0.8$ (Fig. 4.9(a) and $c_2=0.9$ (Fig. 4.9(b). For $c_2 = 0.95$ (Fig. 4.9(c), the dynamics becomes even more slow and regular. We clearly see now stationary sources, with irregularly moving sinks in between. Due to the smallness of g_2 , phase slips in one of the single modes leads in some case to the formation of new sources and sinks. When c_2 is increased to a value of 1 (Fig. 4.9(d), some slow dynamics sets in, that may or may not be a long living transient. For values of c_2 above 1.1, all initial conditions seem to evolve to a stationary, regular source/sink pattern.

4.5 Interactions between sources and sinks

In all of our discussions of sources and sinks up to now, we have considered sources and sinks to be independent and noninteracting structures. It is however a well known fact [84, 85] that generically, coherent structures *will* interact and, in particular, that they do so over long ranges. In single CGL equations, special attention has been given in the literature to the interactions between two *fronts*, and as it turns out, the interac-

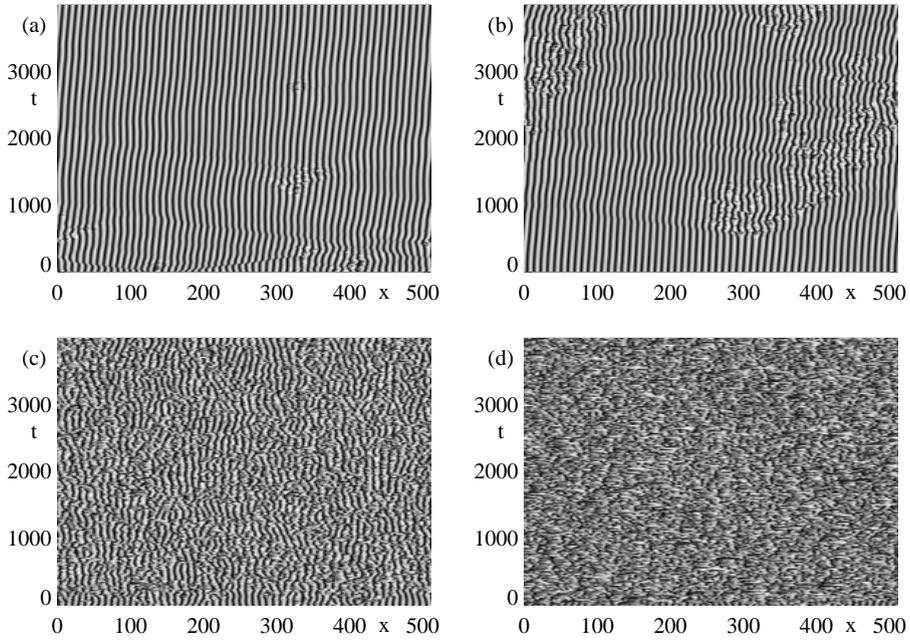


Figure 4.8: Four space-time plots, showing the transition from standing waves to disordered patterns, for $g_2 = 1.1$, $c_1 = 0.9$, $c_3 = 2$, $s_0 = -0.1$, $\varepsilon = 1$, and (a) $c_2 = -0.72$, (b) $c_2 = -0.71$, (c) $c_2 = -0.5$, (d) $c_2 = 0$. See text.

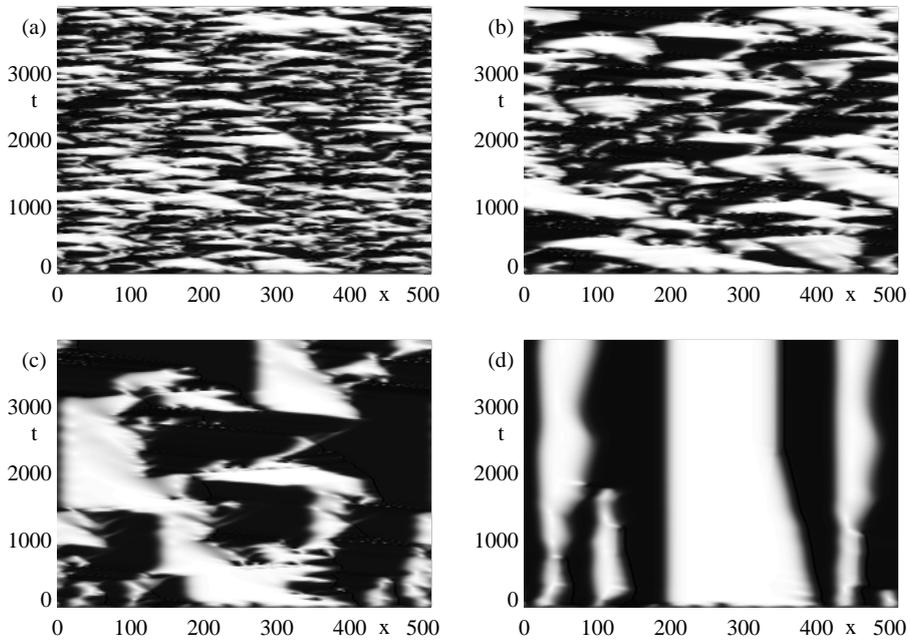


Figure 4.9: Four space-time plots for the same coefficients as in Fig. 4.8, but now for positive values of c_2 . (a) $c_2 = 0.8$, (b) $c_2 = 0.9$, (c) $c_2 = 0.95$, (d) $c_2 = 1$.

tions between fronts can be both attractive and repulsive, depending on the distance between them two, and in certain cases there exists a stable equilibrium distance [84, 85]. In fact, one can understand the formation of finite sized *domains* from the existence of such stable two-front configurations. The interest in interactions in our systems stems from the fact that in experiments on traveling wave systems, it was already noted by some experimentalists that the source/sink state is actually a *transient* state: after a considerable time all sources and sinks will disappear from the system by means of successive pair annihilations. The system would then end up completely filled with a single traveling wave⁴.

We will now briefly analyze the interactions using the same method as the one successfully applied to the case of interacting fronts in the CGL equation. This method is based on deriving solvability conditions, in much of the same way as we did in Section 2.6. To get some idea for the sort of situation we are dealing with here, we have plotted in Fig. 4.10 a typical pair of a source and a sink and the corresponding wavenumber profile. As one can see, the *amplitude* profiles match nicely, but there is a *wavenumber* mismatch in the single mode patches. This wavenumber mismatch however occurs in the mode that is almost zero. If one were to track the positions of an initially well-separated source/sink pair, a picture typically like the one shown in Figs. 4.11 appears: initially, the source is gliding slowly towards the sink, while the sink moves very little at all. It is only in the final stages of the annihilation process that the sink suddenly starts to rapidly approach the source.

In this section, we will try to perform a perturbative study of the long-range interactions between sources and sinks, hoping to capture the initial stages of such an annihilation process. The start of this perturbative study will be the non-interacting case, *i.e.* we need the profiles of an isolated source and sink (or, equivalently, those of a source/sink pair with infinite separation). Those are of course the *same* ones as we obtained from the shooting procedure we employed to extract, for instance, the dependence of the width on ε .

If we start from the coupled CGL Eqs. (3.3.3a,3.3.3b), and split the complex amplitudes $A_{R,L}$ in a modulus and phase part as follows,

$$A_{R,L} = e^{-i\omega t} a_{R,L} e^{i\phi_{R,L}}, \quad (4.5.1)$$

⁴In fact, some of these experiments actually focused in particular on the properties of these single traveling waves, and the source/sink pattern was considered an unwanted transient

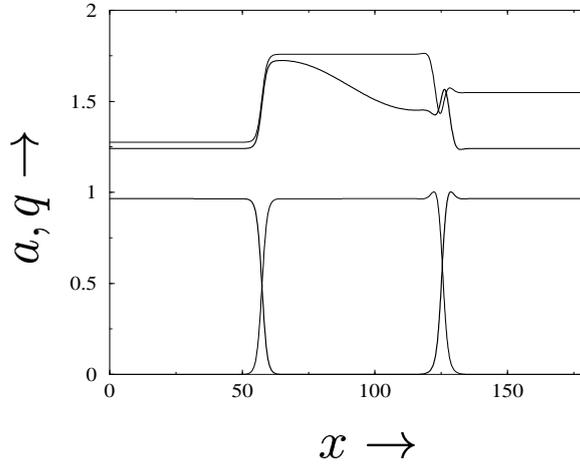


Figure 4.10: A source/sink pair at $\varepsilon = 1, c_1 = 0.6, c_2 = 0, c_3 = 0.4$ and $s_0 = 0.4$. Note the mismatch in selected wavenumber in the linear (nearly zero) mode in the center. The left (right) amplitude is plotted with a thick (thin) line. The wavenumber profile has been shifted upward by an amount of 1.5.

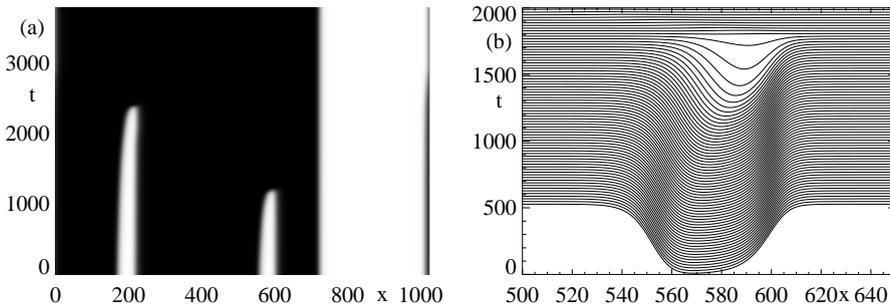


Figure 4.11: (a) Space-time plot of $|A_R|$ (large values of $|A_R|$ show up white) illustrating the interaction between sources and sinks. The runs started from random initial conditions, and the coefficients were chosen as $c_1 = 0.6, c_3 = 0.4, c_2 = 0, g_2 = 2.0, s_0 = 0.4$ and at $\varepsilon = 0.07$. Note that ε is well above the critical value $\varepsilon_c^{\text{so}} = 0.029$, and the sources are stable. Hence, any movement of the coherent structures is solely due to their interactions. Note that in the final stage of an annihilation event, the source moves most, while the sink stays almost put. Note also the similarity to Fig. 24 of [65]. (b) Hidden line plot of $|A_L|$ showing the annihilation process in detail.

we can write the two coupled complex equations equivalently as the following set of four *real* equations

$$\begin{aligned} \partial_t a_R = & \epsilon a_R - s_0 \partial_x a_r + \partial_x^2 a_R - 2c_1 \partial_x a_R \partial_x \phi_R \\ & - c_1 a_R \partial_x^2 \phi_R - a_R (\partial_x \phi_R)^2 - a_R^3 - g_2 a_L^2 a_R, \end{aligned} \quad (4.5.2a)$$

$$\begin{aligned} \partial_t \phi_R = & \omega - s_0 \partial_x \phi_R + \frac{c_1}{a_R} \partial_x^2 a_R + \frac{2}{a_R} \partial_x a_R \partial_x \phi_R \\ & + \partial_x^2 \phi_R - c_1 (\partial_x \phi_R)^2 + c_3 a_R^2 + g_2 c_2 a_L^2, \end{aligned} \quad (4.5.2b)$$

$$\begin{aligned} \partial_t a_L = & \epsilon a_L + s_0 \partial_x a_L + \partial_x^2 a_L - 2c_1 \partial_x a_L \partial_x \phi_L \\ & - c_1 a_L \partial_x^2 \phi_L - a_L (\partial_x \phi_L)^2 - a_L^3 - g_2 a_R^2 a_L, \end{aligned} \quad (4.5.2c)$$

$$\begin{aligned} \partial_t \phi_L = & \omega + s_0 \partial_x \phi_L + \frac{c_1}{a_L} \partial_x^2 a_L + \frac{2}{a_L} \partial_x a_L \partial_x \phi_L \\ & + \partial_x^2 \phi_L - c_1 (\partial_x \phi_L)^2 + c_3 a_L^2 + g_2 c_2 a_R^2. \end{aligned} \quad (4.5.2d)$$

When ϵ is close to onset, the wavenumber selected by the source will be close to zero as well. Since the wavenumber is the spatial derivative of the phase, we can safely assume quantities like $\partial_x \phi_{R,L}$ to be small in the following. Note again that the phases $\phi_{R,L}$ themselves do not and should not appear in these equations. As we have seen this is a consequence of translation invariance. We will work in lowest order here, and neglect variations in time of the phase profiles, or rather the *wavenumber* profiles $q_{R,L}(x)$, which allows us to eliminate the equations for $\varphi_{R,L}$ and write

$$\begin{aligned} \partial_t a_R = & [\epsilon - q_R^2(x) - c_1 \partial_x q_R(x)] a_R - [s_0 + 2c_1 q_R(x)] \partial_x a_R + \partial_x^2 a_R \\ & - a_R^3 - g_2 a_L^2 a_R, \end{aligned} \quad (4.5.3a)$$

$$\begin{aligned} \partial_t a_L = & [\epsilon - q_L^2(x) - c_1 \partial_x q_L(x)] a_L + [s_0 - 2c_1 q_L(x)] \partial_x a_L + \partial_x^2 a_L \\ & - a_L^3 - g_2 a_R^2 a_L. \end{aligned} \quad (4.5.3b)$$

It will be useful to introduce an *effective* ϵ and s_0 (not to be confused with the ϵ_{eff} introduced elsewhere in this Chapter) as follows

$$\begin{aligned} \tilde{\epsilon}_{R,L}(x) = & \epsilon - q_{R,L}^2(x) - c_1 \partial_x q_{R,L}(x), \\ \tilde{s}_{0R,0L}(x) = & s_0 \pm 2c_1 q_{R,L}(x), \end{aligned} \quad (4.5.4a)$$

where in this last definition, it is understood that the minus sign is to be taken for the left mode, and the plus sign for the right mode. Furthermore, from now on we will assume the wavenumber (which is itself

a small quantity), to vary very slowly spatially, so that we can drop the term proportional to $\partial_x q_{R,L}$. This reduces our equations to a modified set of coupled *real* GL equations

$$\partial_t a_R = \tilde{\varepsilon}_R(x) a_R - \tilde{s}_{0R}(x) \partial_x a_R + \partial_x^2 a_R - a_R^3 - g_2 a_L^2 a_R, \quad (4.5.5a)$$

$$\partial_t a_L = \tilde{\varepsilon}_L(x) a_L + \tilde{s}_{0L}(x) \partial_x a_L + \partial_x^2 a_L - a_L^3 - g_2 a_R^2 a_L. \quad (4.5.5b)$$

When we obtain source and sink profiles via the shooting method, we also get the wavenumber profile that belongs to those particular coefficient values. Once such profiles are known, it is easy to calculate the spatial dependence of the control parameter and group velocity terms. In the following, we will drop the tildes and the explicit x -dependence and just write ε and s_0 again. Eqs.4.5.5 can be put in the compact form

$$\partial_t \vec{a} = \{\hat{\mathcal{L}} + \hat{\mathcal{N}}[\vec{a}]\} \vec{a}, \quad (4.5.6)$$

where \vec{a} is the vector

$$\vec{a} = \begin{pmatrix} a_R \\ a_L \end{pmatrix}. \quad (4.5.7)$$

The linear ($\hat{\mathcal{L}}$) and non-linear ($\hat{\mathcal{N}}$) parts of the RHS of Eq. (4.5.6 are respectively the matrices

$$\hat{\mathcal{L}} = \begin{pmatrix} \varepsilon_R - s_{0,R} \partial_x + (1 + ic_1) \partial_x^2 & 0 \\ 0 & \varepsilon_L + s_{0,L} \partial_x + (1 + ic_1) \partial_x^2 \end{pmatrix}, \quad (4.5.8)$$

and

$$\hat{\mathcal{N}}[\vec{a}] = \begin{pmatrix} -(1 - ic_3) |a_R|^2 - g_2 (1 - ic_2) |a_L|^2 & 0 \\ 0 & -(1 - ic_3) |a_L|^2 - g_2 (1 - ic_2) |a_R|^2 \end{pmatrix}. \quad (4.5.9)$$

Shooting, as stated before, provides us with exact, stationary solutions for a source and a sink. By definition, those profiles satisfy

$$0 = \partial_t \vec{a}^{\text{so}}(\xi) = \{\hat{\mathcal{L}} + \hat{\mathcal{N}}[\vec{a}^{\text{so}}(\xi)]\} \vec{a}^{\text{so}}(\xi), \quad (4.5.10a)$$

$$0 = \partial_t \vec{a}^{\text{si}}(\xi) = \{\hat{\mathcal{L}} + \hat{\mathcal{N}}[\vec{a}^{\text{si}}(\xi)]\} \vec{a}^{\text{si}}(\xi). \quad (4.5.10b)$$

4.5.1 Setup of the problem

To start looking at the interactions between the structures, we can simply construct a well-separated pair from the isolated profiles obtained from

the shooting and a small perturbation amplitude, which we *do not know* a priori. We therefore write

$$\vec{a}(x, T) = \vec{a}^{\text{so}}(x - x_{\text{so}}(T)) + \vec{a}^{\text{si}}(x - x_{\text{si}}(T)) - \vec{a}_0 + \mu \vec{a}^{\text{p}}(x, T), \quad (4.5.11)$$

where x_{so} and x_{si} are the positions of the source and sink respectively (well separated therefore means $|x_{\text{so}} - x_{\text{si}}| \gg 1$), and μ is a small quantity (the perturbation amplitude \vec{a}^{p} itself is now an $\mathcal{O}(1)$ quantity). Much in the spirit of the amplitude approach, we assume that the dynamics of such a well-separated source/sink pair will be slow and take the typical timescale to be $T = \mu t$. Furthermore, for definiteness we shall always take a source on the left hand side of the pair, which means that we should take

$$\vec{a}_0 = \begin{pmatrix} a_0 \\ 0 \end{pmatrix}. \quad (4.5.12)$$

Writing out the coupled modified real GL equations 4.5.5 for profiles of the type 4.5.11 leaves us (after quite a bit of rather straightforward algebra, which is presented in some more detail in Appendix 4.A) with an equation for the perturbation amplitude of the form

$$\mu \hat{\mathcal{L}}_1 \vec{a}^{\text{p}} = - \{ v_{\text{so}} \partial_\xi \vec{a}^{\text{so}} + v_{\text{si}} \partial_\xi \vec{a}^{\text{si}} \} + \vec{\mathcal{P}}. \quad (4.5.13)$$

The source and sink *velocities* are given by $v_{\text{so,si}} = \partial_\xi x_{\text{so,si}}$. The operator $\hat{\mathcal{L}}_1$ is of the form

$$\hat{\mathcal{L}}_1 = \begin{pmatrix} \varepsilon_R - s_{0,R} \partial_x + \partial_x^2 - \mathcal{E}_R - g_2 \mathcal{G}_R & -g_2 \mathcal{H}_R \\ -g_2 \mathcal{H}_L & \varepsilon_L + s_{0,L} \partial_x + \partial_x^2 - \mathcal{E}_L - g_2 \mathcal{G}_L \end{pmatrix}, \quad (4.5.14)$$

and

$$\vec{\mathcal{P}} = \begin{pmatrix} \varepsilon a_0 + \mathcal{D}_R + g_2 \mathcal{F}_R \\ \mathcal{D}_L + g_2 \mathcal{F}_L \end{pmatrix} = \varepsilon \vec{a}_0 + \vec{\mathcal{D}} + g_2 \vec{\mathcal{F}}, \quad (4.5.15)$$

$\mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}$ and \mathcal{H} are all polynomials in the source and sink profiles, explicit expressions can be found in Appendix 4.A. We can solve for these velocities by applying again the Fredholm alternative (see Section 2.6 for details). There is however one slight complication here, and that is that we need the zero-modes of the *adjoint* linear operator. This linear operator is not self-adjoint, and in general there is no simple relation between zero-modes and adjoint zero-modes. The term causing the troubles is the

first derivative (multiplicative terms are always self-adjoint), but fortunately there is a way around this. The essential part of the equation for a zero-mode of $\hat{\mathcal{L}}_1$ looks like

$$\begin{pmatrix} \partial_x^2 - s_0 \partial_x & 0 \\ 0 & \partial_x^2 + s_0 \partial_x \end{pmatrix} \begin{pmatrix} a_R \\ a_L \end{pmatrix} = 0 \quad (4.5.16)$$

And is not self-adjoint. If we apply now a transformation

$$\hat{\mathcal{H}}_1 = -\mathbf{M} \hat{\mathcal{L}}_1 \mathbf{M}^{-1}, \quad (4.5.17a)$$

$$\vec{\psi} = \mathbf{M} \vec{a}, \quad (4.5.17b)$$

with \mathbf{M} the $SL(2)$ matrix (with unit determinant, therefore no zero eigenvalue, and invertible)

$$\mathbf{M} = \begin{pmatrix} e^{-s_0 x/2} & 0 \\ 0 & e^{s_0 x/2} \end{pmatrix}, \quad \mathbf{M}^{-1} = \begin{pmatrix} e^{s_0 x/2} & 0 \\ 0 & e^{-s_0 x/2} \end{pmatrix}. \quad (4.5.18)$$

We can ask how the new $\hat{\mathcal{H}}_1$ acts on functions $\vec{\varphi}$

$$\hat{\mathcal{H}}_1 \vec{\varphi} = -\mathbf{M} \hat{\mathcal{L}}_1 \mathbf{M}^{-1} \vec{\varphi}, \quad (4.5.19)$$

to find after some calculation that

$$\hat{\mathcal{H}}_1 \vec{\varphi} = \begin{pmatrix} -\partial_x^2 + s_0^2/4 & 0 \\ 0 & -\partial_x^2 + s_0^2/4 \end{pmatrix} \begin{pmatrix} \varphi_R \\ \varphi_L \end{pmatrix}, \quad (4.5.20)$$

and we see that the $\hat{\mathcal{H}}_1$ thus defined is indeed self-adjoint. All we need to do therefore is look for zero-modes of $\hat{\mathcal{L}}_1$ itself, and transform these. As we shall see in the next section, these zero modes are rather easy to obtain.

4.5.2 Zero-modes of the linear operator

Since the entire system is translation invariant, displacing the source/sink pair over any distance should not have any influence. Such a displacement corresponds to a shift in coordinates

$$x_{s_0} \rightarrow x_{s_0} + \Delta x, \quad x_{s_i} \rightarrow x_{s_i} + \Delta x, \quad (4.5.21)$$

which implies for the pair profile

$$\begin{aligned}\vec{A} &\rightarrow \vec{a}^{\text{so}}(x - x_{\text{so}} - \Delta x) + \vec{a}^{\text{si}}(x - x_{\text{si}} - \Delta x) - \vec{A}_0 \\ &= \vec{a}^{\text{so}}(x - x_{\text{so}}) - \partial_x \vec{a}^{\text{so}} \Delta x + \vec{a}^{\text{si}}(x - x_{\text{si}}) - \partial_x \vec{a}^{\text{si}} \Delta x - \vec{A}_0 \\ &= \vec{A} - \Delta x \partial_x \{ \vec{a}^{\text{so}} + \vec{a}^{\text{si}} \}.\end{aligned}\quad (4.5.22)$$

We can therefore think of

$$\vec{N}_{\text{t}}(x) = \partial_x \begin{pmatrix} a_R^{\text{so}} + a_R^{\text{si}} \\ a_L^{\text{so}} + a_L^{\text{si}} \end{pmatrix}, \quad (4.5.23)$$

the translation zero mode, as the generator of translations of the entire pair in this system. The other relevant zero mode is a breathing mode of the pair, where the source and sink are displaced in opposite directions.

$$x_{\text{so}} \rightarrow x_{\text{so}} + \Delta x, \quad x_{\text{si}} \rightarrow x_{\text{si}} - \Delta x. \quad (4.5.24)$$

A similar calculation yields the corresponding generator

$$\vec{N}_{\text{b}}(x) = \partial_x \begin{pmatrix} a_R^{\text{so}} - a_R^{\text{si}} \\ a_L^{\text{so}} - a_L^{\text{si}} \end{pmatrix}, \quad (4.5.25)$$

the breathing zero mode. From these $\hat{\mathcal{L}}_1$ -zero modes, we can construct the $\hat{\mathcal{L}}_1^\dagger$ zero modes by noting the following. Since, as we have seen, we have

$$\hat{\mathcal{L}}_1 = -\mathbf{M}^{-1} \hat{\mathcal{H}}_1, \quad (4.5.26)$$

we can write the adjoint linear operator as

$$\hat{\mathcal{L}}_1^\dagger = -\mathbf{M} \hat{\mathcal{H}}_1 \mathbf{M}^{-1}. \quad (4.5.27)$$

The equation solved by the zero modes we just constructed is

$$\hat{\mathcal{L}}_1 \vec{N} = 0 \quad \text{or, equivalently,} \quad -\mathbf{M}^{-1} \hat{\mathcal{H}}_1 \mathbf{M} \vec{N} = 0. \quad (4.5.28)$$

Now let $\vec{\psi}_{\vec{N}} \equiv \mathbf{M} \vec{N}$, so that $\vec{N} = \mathbf{M}^{-1} \vec{\psi}_{\vec{N}}$. In terms of $\hat{\mathcal{H}}_1$, we have therefore solved

$$\hat{\mathcal{H}}_1 \vec{\psi}_{\vec{N}} = 0. \quad (4.5.29)$$

Now to the adjoint equation $\hat{\mathcal{L}}_1^\dagger \vec{N}^\dagger = 0$. This implies that $-\mathbf{M} \hat{\mathcal{H}}_1 \mathbf{M}^{-1} \vec{N}^\dagger = 0$. If we now define $\vec{\psi}_{\vec{N}}^\dagger \equiv \mathbf{M}^{-1} \vec{N}^\dagger$, or equivalently $\vec{N}^\dagger = \mathbf{M} \vec{\psi}_{\vec{N}}^\dagger$, we get

$$\hat{\mathcal{H}}_1 \vec{\psi}_{\vec{N}}^\dagger = 0 \quad (4.5.30)$$

This equation we have already solved above, which tells us that

$$\vec{\psi}_{\vec{N}} = \vec{\psi}_{\vec{N}}^\dagger \Rightarrow M\vec{N} = M^{-1}\vec{N}^\dagger \quad (4.5.31)$$

So that we can construct the left eigenvectors from the right eigenvectors by applying M twice :

$$\vec{N}^\dagger = M^2\vec{N} \quad (4.5.32)$$

4.5.3 Solvability conditions

We now have all we need to write down the solvability conditions that determine the velocities of both structures. Applying the Fredholm theorem yields the following equation for the velocities

$$v_{\text{so}} \langle M^2 \vec{N}_{\text{t}}, \partial_\xi \vec{a}^{\text{so}} \rangle + v_{\text{si}} \langle M^2 \vec{N}_{\text{t}}, \partial_\xi \vec{a}^{\text{si}} \rangle = \langle M^2 \vec{N}_{\text{t}}, \vec{\mathcal{P}} \rangle, \quad (4.5.33a)$$

$$v_{\text{so}} \langle M^2 \vec{N}_{\text{b}}, \partial_\xi \vec{a}^{\text{so}} \rangle + v_{\text{si}} \langle M^2 \vec{N}_{\text{b}}, \partial_\xi \vec{a}^{\text{si}} \rangle = \langle M^2 \vec{N}_{\text{b}}, \vec{\mathcal{P}} \rangle. \quad (4.5.33b)$$

The scalar product is defined as

$$\langle a, b \rangle = \int_{-\infty}^{+\infty} dx a^*(x)b(x) \quad (4.5.34)$$

The solvability conditions now completely determine v_{so} and v_{si} . Although we are still in the process of studying these equations numerically, preliminary results show that the interaction is purely attractive, as simulations of the full system of PDE's and indeed experiments also strongly suggest.

4.6 Outlook

In this and the previous Chapter, we have extended the coherent structures framework and the counting arguments to the coupled CGL equations, and obtained important information on the dynamical states that are independent of the precise values of the coefficients and bear experimental relevance. In general, these considerations lead to the conclusion that sources are often unique, sometimes come in pairs but in any case are at most members of a discrete set of solutions. As a result, they are instrumental for the wavenumber selection of both regular and chaotic patterns. Many of the instability mechanisms and dynamical regimes of the coupled CGL equations can be understood qualitatively from this point of view, and we have shown several examples of hitherto unexplored

regimes of persistent spatio-temporal chaotic dynamics (see Table 4.1). In this closing section, we wish to discuss some of these findings in the light of experimental observations, and summarize the most important open theoretical problems.

4.7 Experimental implications

In short, the experimental predictions that we make, based on our study of the coupled CGL equations are the following :

- *Multiplicity.* Our analysis shows that sources are expected to come in a discrete set, which would experimentally amount to a *unique*, stationary source. Furthermore, this source is expected to be *symmetric*, in that it sends out waves of the same wavenumber to both sides.

Sinks are non-unique. This means that one could have sinks with different velocities present at the same time. In light of the previous remark on the uniqueness of sources, this might prove hard to observe experimentally.

- *Wavenumber selection.* One important consequence of the uniqueness of sources is that they select an asymptotic wavenumber, just as spirals do in the 2D-case. Since the traveling-wave system is quasi-one-dimensional however, we expect the wavenumber selection to be much easier to study.

- *Scaling Behavior.* We have made definite predictions for the small- ε scaling of the width of sources and sinks. Moreover, we predict the stationary sources to disappear at some finite value of ε , which is the point where the non-stationary sources take over. These sources scale as ε^{-1} , as do the sinks.

- *Instabilities and Dynamical Behavior.* Apart from the non-stationary sources that occur when ε is decreased sufficiently, we have found that there are at least two other mechanisms leading to dynamical states. The central observation is that the waves that are selected and sent out by the sources may become unstable. Similar to what happens in the single CGL equation, these waves can become convectively or absolutely unstable; the latter case in particular yields chaotic states (section 4.2). When the cross-coupling coefficient is not too far above one, and the selected wavenumber is unequal to zero, there is a regime where both single and bimodal states are unstable.

4.8 Comparison of results with experimental data

Most research in the field of traveling wave systems has focussed on the properties of the single-mode states, i.e., the states where the entire experimental cell is filled up by either the left- or the right-traveling wave. From such a perspective, it is natural to disregard the source/sink patterns that generally occur initially above onset as unwanted transient states. Consequently they have not been studied as extensively as we think they deserve to be. It is the aim of this section to confront a number of the theoretical findings of this article with some of the experimental observations in the heated wire experiments [40, 86, 51, 87, 52, 88] and in the experiments on traveling waves in binary liquids [41, 65, 89, 90, 91, 92]. In no way do we claim this comparison to be exhaustive — the main aim of our discussion is to show that our results put various earlier observations in a new light, and that it should be feasible to settle various of the issues we raise with further systematic experiments.

4.8.1 Heated wire experiments

When a wire which is put a distance of the order of a millimeter under the free surface of a liquid layer is heated, traveling waves occur beyond some critical value of the heating power [40, 86, 51, 87, 52, 88]. In deriving our amplitude equations, we have assumed that this bifurcation to traveling waves is supercritical. This is also what several experimental groups [87, 52] have found. Fig 4.12 shows nicely that in recent experiments by Westra *et al.* the spectral power, a quantity related to the *square* of the wave amplitude, grows linearly in ε beyond its threshold value. This linear law, corresponding to the characteristic square root behavior of the order parameter near the onset of a supercritical bifurcation, is seen to hold for values of ε roughly up to 0.5, which should be a good indication of the regime of applicability of the amplitude equations we have used. Moreover, it was established experimentally that the group velocity and phase velocity have the same sign [40]⁵. The paper by Vince and Dubois [87] is one of the few older papers we know of that discusses the ε -dependence of the width of sources. The authors show that the inverse width scales linearly with the heating power Q , and associate this with a scaling of the source width as ε^{-1} . This is correct if the value of Q at

⁵Fig. 11 of [87] also illustrates quite nicely that the group velocity and phase velocity are parallel.

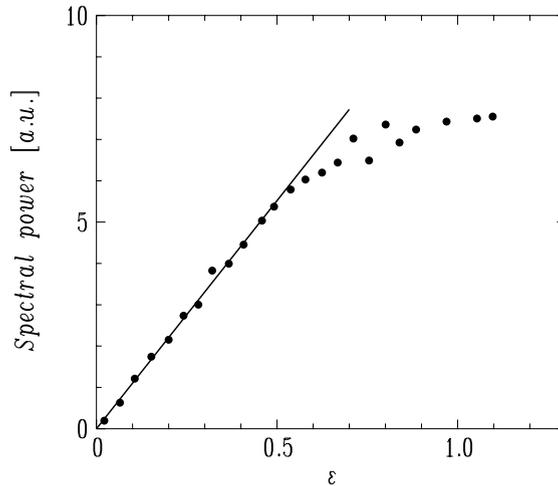


Figure 4.12: Scaling of the spectral power (proportional to the wave amplitude squared) versus control parameter ε , revealing the supercritical nature of the bifurcation to traveling waves in the heated wire system. Figure reproduced from [52] with permission.

which the source width diverges coincides with the threshold value for the linear instability, but whether this is actually the case is unfortunately not quite clear from the published data⁶. Formulated differently, in terms of our numerical data shown in Fig. 3.6d, the question arises whether in the experiments the approximate linear scaling of the inverse width with the heating power was associated with that of the thick line above ε_c^{so} , or with the linear scaling $\sim \varepsilon$ below ε_c^{so} . If indeed the experiments are consistent with an ε^{-1} scaling of the width, then according to our analysis the sources should be (weakly) non-stationary and prone to pinning to inhomogeneities in the cell. Recently however, new experimental data have become available from the Eindhoven group, and these *do* show clear evidence for the existence of the critical control parameter ε_c^{so} .

Plotted in Fig. 4.13 is the width of sources versus ε . Clearly visible in this plot is the existence of two distinct regimes. In both regimes, the source width scales with the inverse control parameter, but with different

⁶In the experiments shown in Fig. 10 of [87], the source width diverged at $Q \approx 4.2$ Watts. Unfortunately, the distance h between the wire and the fluid surface is not given for the data shown. All other measurements in the paper are made at $h = 1.34$ mm and $h = 1.97$ mm, and these values correspond to $Q_c \approx 2.5$ Watts and $Q_c \approx 2$ Watts.

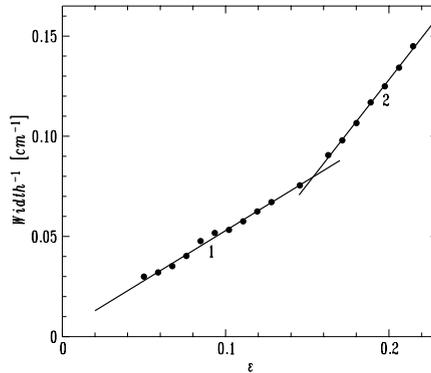


Figure 4.13: Scaling of the source width with ϵ . Clearly visible in this figure are the two regimes also present in Fig. 3.6. From this figure, we estimate the critical value ϵ_c^{so} to be approximately 0.1. Figure reproduced from [52] with permission.

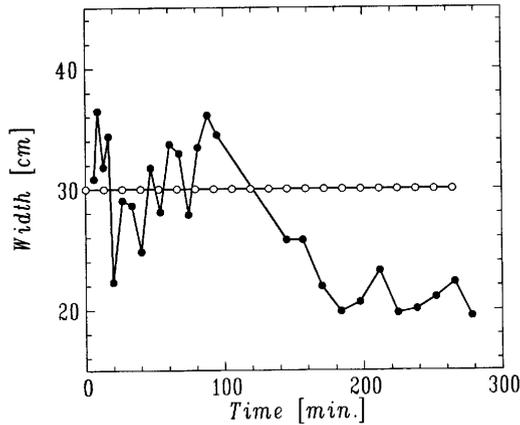


Figure 4.14: Source behavior above ($\epsilon = 0.28$, open circles) and below ($\epsilon = 0.04$, filled circles) the critical $\epsilon_c^{so} \approx 0.1$. The structure above ϵ_c^{so} is perfectly stationary, while below the threshold value dynamical behavior is seen to set in. Figure reproduced from [52] with permission.

prefactors. This is also what we observed in our coupled CGL equations. The theoretical analysis also predicted that the transition from one regime to the other when ε is decreased below its critical value $\varepsilon_c^{\text{so}}$ is one from stationary sources to fluctuating ones. This is again confirmed by the experiments, as Fig. 4.14 illustrates. We see that for a value of ε well above $\varepsilon_c^{\text{so}}$ (empirically determined to lie roughly at $\varepsilon = 0.1$, as can be inferred from Fig. 4.13), the source is completely stationary, and stays perfectly put. For $\varepsilon < \varepsilon_c^{\text{so}}$ however, we see that the source starts to move about quite appreciably. We take this as evidence for the stationary-nonstationary transition also observed in the coupled equations.

In [86], Dubois *et al.* also note that “. . .sources may be large when the sinks are always very narrow . . .” in their heated wire experiments. This agrees with our finding that sinks are always less wide than the sources but the published data do not allow us to extract the scaling of the sink width with ε .

In the experiments by Alvarez *et al.* [40], sources were found to be stationary and symmetric but non-unique, i.e., each source sends out the same waves to both sides, but different sources send out different waves. As a result, patches with different wavenumbers were found to be present in the system (at any one time), and the sources were seen to move in response to the fact that they were sandwiched between waves of different frequency. We have already seen in section 3.5.3 that there are certain regions of parameter space where there were two different sources present at the same time (for one of them, the linear group velocity s_0 and nonlinear group velocity s had opposite signs). However, the fact that we can have various discrete source solutions can not explain the experimental observations. First of all, in our simulations two of such sources were separated by a sink-type structure in one single mode patch, *not* by a sink separating two oppositely traveling waves, as in the experiments. Secondly, in the experiments there were always slight differences between any two pair of sources, which appears inconsistent with the existence of a finite number of discrete source solutions.

It appears likely to us that the occurrence of slight differences between different sources results from the fact that there are always some impurities or inhomogeneities present in any experimental setup. Very much like the spirals and target patterns one encounters in the 2D CGL equation [93], coherent structures might well be pinned to such imperfections⁷. This

⁷An example of how sources can be pinned near cell boundaries below $\varepsilon_c^{\text{so}}$ is discussed

would of course not invalidate the results of the counting arguments for the homogeneous case, as it is precisely on the basis of this counting argument that one would expect the properties of the discrete source solution(s) to depend sensitively on the local parameter values.

The sinks which in the experiments of [40] were sandwiched between two patches with different wavenumbers, were found to move according to what was referred to as a “phase matching rule”: during the motion, a constant phase difference is maintained across the sink profile, so that no phase slip events occur. This commonly occurs for sinks in the *single* CGL equation, and Fig. 3.5 provides an example of this, but there is one important difference here: sinks in the experiments separate two oppositely traveling waves, so phase matching in the actual experiments involves the *fast* scales represented by the critical wavelength q_c of the pattern at onset. In the amplitude approach all information about this q_c is lost since we eliminated the fast scales and only consider the difference between the actual wavenumber q of the pattern and this q_c . At least in the experiments of [40] the coupling between the fast and the slow scales is important. Experimentally, it is still not completely clear whether the “phase matching rule” was a peculiarity of [40], or whether it holds quite generally.

As we have seen in this Chapter, the wavenumber selection by sources entails various scenarios for instabilities and chaotic dynamics in the single-mode patches that are separated by sources and sinks. In the experiments, there are regimes in parameter space where the dynamics is reminiscent of what one expects when the mode selected by the sources becomes convectively or absolutely unstable. Whether the data are consistent with this scenario has remained unexplored, however.

We finally note that it has recently become apparent that traveling waves in convection cells with a free surface which are heated from the side [95, 96, 97], are intimately related to those occurring in the heated wire experiments [71]. Clearly, both the heated wire experiments and this system appear to be very suitable setups to study the dynamics of sources and sinks; in addition, both do show rich dynamical behavior.

in [94].

4.8.2 Binary mixtures

One of the best studied experimental traveling wave systems is binary fluid convection [41, 65, 89, 91, 92]. Since the bifurcation in this case has been shown to be weakly subcritical [31], the description of the behavior in this system is strictly speaking beyond the scope of the coupled CGL equations we consider. A brief discussion is nevertheless warranted, not only because some of the behavior of sources and sinks is quite generic, in that it does not strongly depend on the underlying bifurcation structure (e.g., sources still form a discrete set according to the counting arguments), but also because the additional complications of the binary mixture convection experiments are an interesting subject for future study.

Kaplan and Steinberg have shown that the transition from localized traveling wave patterns (pulses) to extended traveling waves is essentially governed by the convective instability of the edges of the pulses [98]⁸. The fact that the relevant front velocity is given by linear marginal stability arguments, suggests that the subcritical character of the bifurcation is not very strong here. On the other hand, the nonadiabatic effects, such as locking, observed in [99], point in the other direction, namely that the subcritical nature of the transition is rather strong. Hence, the importance of the subcritical effects in these experiments can not be trivially established.

Kolodner [65] has observed a wide variety of source/sink behavior. In some cases, there appears to be a stable source/sink pair where the sink is clearly wider than the source. This of course contradicts what we typically find (except close to the relaxational limit — see section 3.6.4). This may have to do with the subcritical nature of the bifurcation, but one should also keep in mind that in other cases there is evidence that such behavior could still be a transient, because there are still phase slip events occurring. E.g., Fig. 5 of [65] shows a notable example of a case in which the sink is initially wider than the source, but in which a process clearly involving the fast scales narrows it down, so that in the end it is smaller than the source.

Another interesting state that is encountered in the experiments are drifting source/sink patterns (see, e.g., Fig.7 of [65]). The sources here move slowly but with a constant velocity, and are non-symmetric in that the wavenumbers on either side are different. However, there is again a

⁸This is similar to the behavior of sources near $\varepsilon_c^{\text{so}}$.

one-to-one correspondence between the drift velocity and the difference in wavenumbers. In [65], this is referred to this process as “Doppler shifting”, to indicate that in the frame co-moving with it, the drifting source sends out waves with the same frequency to the left and the right. This is completely equivalent to the “phase matching rule” of [40]. When such a moving source is present, the sinks are also found to obey the phase matching rule and so they move with exactly the same drift velocity as the sources. Clearly, it is still the source that selects the wave number and hence plays the active role here — as usual, the sink motion is essentially determined by the properties of the waves that come in. A priori, one could imagine that the sources and sinks in the binary fluid experiments are more prone towards obeying the phase matching rule due to the subcritical nature of the bifurcations to traveling waves, but one can find various examples in the experiments where they do not obey this rule. Obviously, this question deserves further study.

The fact that Kolodner [65] observes in his Fig. 7 a steadily moving source is not necessarily in contradiction with our counting arguments, as these do allow for the existence of a discrete set of $v \neq 0$ sources. In practice, however, for a proper analysis of such source solutions in the binary fluid experiments it is probably necessary to include the coupling to the slow concentration field, as in the work of Riecke and coworkers on traveling pulse solutions [27, 100, 101].

Although several of the experiments of Kolodner have been done at very small values of ε , there is no visible evidence of the divergence of the width of any of the sources and sinks. Presumably, this is due to the subcritical nature of the bifurcation — in section 3.6.2 we already argued that in this case the width of neither the sources nor the sinks need to diverge as $\varepsilon \rightarrow 0$.

In passing, we note that, quite impressively, Kolodner has also been able to extract the spatial amplitude profiles of his sources and sinks (Figs. 8, 18 and 21 of [65]). These agree remarkably well with the profiles we obtained numerically using the shooting method described earlier. Even the characteristic overshoots of the amplitudes near the edges of sinks are clearly observable in all cases.

In conclusion, although a detailed comparison between the sources and sinks in binary fluid experiments and those analyzed theoretically here, is not justified, many qualitative features (multiplicity, wavenumber selection, etc.) are quite similar. We expect that the ε dependence of

the width of these structures is very different in the two cases, due to the subcritical nature of the bifurcation in binary mixtures and due to the coupling to the slow concentration field. The latter effect probably also plays an important role in the drift of the sources.

4.9 Conclusion

In this chapter and the previous one, we have tried to analyze in some detail the properties of the coherent structures called sources and sinks, as these appear in a variety of traveling wave systems. We have attempted to come up with predictions that should not be too hard to verify (or indeed falsify) experimentally, and in fact in very recent experiments by the Eindhoven group, some of these predictions have already been corroborated. Especially the recent experimental observation of the transition to non-stationary sources at finite ε , which was also found in the equations is remarkable to say the least. To our knowledge, never before has use of the CGL equations been able to predict in such detail the behavior of similar systems. Even though there are still a number of inconsistencies between theory and experiment, this gives us great confidence that we are on the right path. We are currently collaborating with the Eindhoven group to study in detail the implications of their recent results, in particular to see whether we can extract values of the various coefficients from the experimental data. Knowledge of these quantities should facilitate quantitative confrontations between theory and experiment.

4.A Details of the interactions calculation

The polynomials appearing in Eqs. (4.5.14) and (4.5.15) are explicitly given in terms of the *isolated* source and sink profiles by

$$\begin{aligned} \mathcal{D}_R = & -a_0^3 + 3a_0^2 a_R^{\text{si}} - 3a_0 a_R^{\text{si}^2} + 3a_0^2 a_R^{\text{so}} - 6a_R^{\text{si}} a_R^{\text{so}} a_0 \\ & + 3a_R^{\text{si}^2} a_R^{\text{so}} - 3a_0 a_R^{\text{so}^2} + 3a_R^{\text{si}} a_R^{\text{so}^2}, \end{aligned} \quad (4.A.1a)$$

$$\mathcal{E}_R = 3a_0^2 - 6a_0 a_R^{\text{si}} + 3a_R^{\text{si}^2} - 6a_0 a_R^{\text{so}} + 6a_R^{\text{si}} a_R^{\text{so}} + 3a_R^{\text{so}^2}, \quad (4.A.1b)$$

$$\begin{aligned} \mathcal{F}_R = & -a_0 a_L^{\text{si}^2} - 2a_0 a_L^{\text{si}} a_L^{\text{so}} - a_0 a_L^{\text{so}^2} + 2a_L^{\text{si}} a_L^{\text{so}} a_R^{\text{si}} + a_L^{\text{so}^2} a_R^{\text{si}} \\ & + a_L^{\text{si}^2} a_R^{\text{so}} + 2a_L^{\text{si}} a_L^{\text{so}} a_R^{\text{so}}, \end{aligned} \quad (4.A.1c)$$

$$\mathcal{G}_R = a_L^{\text{si}^2} + 2a_L^{\text{si}} a_L^{\text{so}} + a_L^{\text{so}^2}, \quad (4.A.1d)$$

$$\begin{aligned} \mathcal{H}_R = & -2a_0 a_L^{\text{si}} - 2a_0 a_L^{\text{so}} + 2a_L^{\text{si}} a_R^{\text{si}} + 2a_L^{\text{so}} a_R^{\text{si}} + 2a_L^{\text{si}} a_R^{\text{so}} \\ & + 2a_L^{\text{so}} a_R^{\text{so}}. \end{aligned} \quad (4.A.1e)$$

And

$$\mathcal{D}_L = 3a_L^{\text{si}^2} a_L^{\text{so}} + 3a_L^{\text{si}} a_L^{\text{so}^2}, \quad (4.A.1f)$$

$$\mathcal{E}_L = 3a_L^{\text{si}^2} + 6a_L^{\text{si}} a_L^{\text{so}} + 3a_L^{\text{so}^2}, \quad (4.A.1g)$$

$$\begin{aligned} \mathcal{F}_L = & a_0^2 a_L^{\text{si}} + a_0^2 a_L^{\text{so}} - 2a_0 a_L^{\text{si}} a_R^{\text{si}} - 2a_0 a_L^{\text{so}} a_R^{\text{si}} \\ & a_L^{\text{so}} a_R^{\text{si}^2} - 2a_0 a_L^{\text{si}} a_R^{\text{so}} - 2a_0 a_L^{\text{so}} a_R^{\text{so}} + 2a_L^{\text{si}} a_R^{\text{si}} a_R^{\text{so}} \\ & + 2a_L^{\text{so}} a_R^{\text{si}} a_R^{\text{so}} + a_L^{\text{si}} a_R^{\text{so}^2}, \end{aligned} \quad (4.A.1h)$$

$$\mathcal{G}_L = a_0^2 - 2a_0 a_R^{\text{si}} + a_R^{\text{si}^2} - 2a_0 a_R^{\text{so}} + 2a_R^{\text{si}} a_R^{\text{so}} + a_R^{\text{so}^2}, \quad (4.A.1i)$$

$$\begin{aligned} \mathcal{H}_L = & -2a_0 a_L^{\text{si}} - 2a_0 a_L^{\text{so}} + 2a_L^{\text{si}} a_R^{\text{si}} + 2a_L^{\text{so}} a_R^{\text{si}} + 2a_L^{\text{si}} a_R^{\text{so}} \\ & + 2a_L^{\text{so}} a_R^{\text{so}}. \end{aligned} \quad (4.A.1j)$$

*Universal algebraic relaxation in
pulled front propagation*

5.1 Introduction

In this Chapter, we will investigate some of the properties of *fronts*. As we have very briefly seen in Section 2.9, fronts are coherent structures that are localized in space, and connect two distinct phases on either side. In the systems we will be looking at here, one of these phases will always be unstable with respect to the other, which causes the stable preferred state to *invade* the other one. In this Chapter, we will be interested in determining the velocity at which such invasions occur in a variety of model systems, and furthermore we will look at the relaxation to this velocity. We will show that this relaxation behavior is remarkably similar in all of the systems studied in this Chapter, even though their spatio-temporal behavior is completely different, ranging from completely ordered to chaotic. We will begin here by briefly introducing some of the basic concepts and questions, before turning our attention to the analysis of the universal relaxation behavior.

5.2 Fronts in the nonlinear diffusion equation

One of the simplest models to display fronts propagation is the one known as the *nonlinear diffusion equation* (NLDE)

$$\partial_t \varphi(x, t) = D \partial_x^2 \varphi + f(\varphi). \quad (5.2.1)$$

We will use here one particular instance of this equation, which has for its nonlinear function $f(\varphi) = \varepsilon\varphi + \varphi^3$. In this form (and with the diffusion coefficient D rescaled to 1), the NLDE is better known as the Fisher-Kolmogorov¹ equation.

It is quite easy to check that the FK-equation possesses three stationary, spatially homogenous solutions, which are

$$\varphi(x, t) = 0, \text{ and } \varphi = \pm\sqrt{\varepsilon}, \quad (5.2.2)$$

the latter two of which obviously exist only for positive values of the control parameter ε . Without loss of generality, we can and will only consider strictly positive solutions here. A simple linear stability calculation reveals that for $\varepsilon > 0$, the $\varphi = 0$ solutions become unstable and the $\varphi = \sqrt{\varepsilon}$ state is the stable one. The front solutions that we are looking for are coherent structures, and therefore should move at a constant velocity v . They connect a region in the stable state to a region that is still in the unstable state, and such asymptotic front solutions $\Phi(x, t)$ can therefore be characterized by

$$\Phi(x, t) = \Phi(x - vt) \equiv \Phi(\xi) \quad , \text{ with } \Phi(\xi) > 0 \quad \forall \xi, \quad (5.2.3a)$$

$$\lim_{\xi \rightarrow -\infty} \Phi(\xi) = \sqrt{\varepsilon} \quad , \quad \lim_{\xi \rightarrow \infty} \Phi(\xi) = 0. \quad (5.2.3b)$$

We can substitute this form for the solutions in the FK equations, to find that the profile should obey the following ODE

$$-v \frac{d\Phi}{d\xi} = \frac{d^2\Phi}{d\xi^2} + \varepsilon\Phi\Phi^3, \quad (5.2.4)$$

When we look at this equation deep in the tip region, where the order parameter is small, we can effectively ignore the nonlinear term and work instead with a linearized version of Eq. (5.2.4). We can look for solutions of this linearized equation by making the usual Ansatz $\Phi \sim e^{-\mu\xi}$, to find that the values μ can take on are

$$\mu = \lambda_{\pm} = \frac{v}{2} \pm \sqrt{\left(\frac{v}{2}\right)^2 - \varepsilon}. \quad (5.2.5)$$

The above relation distinguishes two different velocity regimes, each of which has its own characteristic type of solutions. The first regime is the

¹Although it looks identical to the RGL equation of earlier Chapters, this equation is for *real* functions $\varphi(x, t)$.

one where the argument of the square root in Eq. (5.2.5) is zero

$$v < v^c = 2\sqrt{\varepsilon} \quad : \quad \Phi_v(\xi) = K_v e^{-\frac{v}{2}\xi} \cos q_v(\xi - \xi_0), \quad (5.2.6)$$

where the critical velocity v^c is determined by equating the square root to zero in Eq. (5.2.5). When the velocity is *larger* than the critical velocity we get the second type of solutions, a sum of two distinct real exponentials

$$v > v^c = 2\sqrt{\varepsilon} \quad : \quad \Phi_v(\xi) = A_v e^{-\lambda_+\xi} + B_v e^{-\lambda_-\xi}. \quad (5.2.7)$$

This means, that at $v = v^c$ there is transition from non-monotonic (or oscillating) solutions to monotonically decaying ones. Right *at* the transition point, the two eigenvalues λ_+ and λ_- exactly coincide, and we get solutions of the form

$$v = v^c \quad : \quad \Phi^c(\xi) = (\alpha\xi + \beta)e^{-\lambda^c\xi} \quad , \quad \text{with} \quad \lambda^c = \frac{v^c}{2} \quad (5.2.8)$$

One can now once again apply the machinery of linear stability theory to investigate the actual *stability* of these different types of front solutions. The outcome of this analysis (which we will not perform here) is that the non-monotonic front solutions, *i.e.* the ones with velocities below the critical value v^c are linearly unstable. The emerging picture for front solutions is therefore a very familiar one: For velocity values below a known critical value, no stable front solutions exist. For values above this critical value, we see that there exists a continuum of stable front solutions, parametrized by their velocity. The question is once again one of selection, much as it was in the case of the wavenumber in our traveling wave system. What velocity does the front choose from the continuum of allowed ones?

In order to answer this question, we should note that the answer to that question does in fact depend on initial conditions. The *steepness* λ of an initial condition, defined by

$$\lambda = - \lim_{x \rightarrow \infty} \frac{\partial \varphi(x, 0)}{\partial x} \quad \text{so that} \quad \lim_{x \rightarrow \infty} \varphi(x, 0) \sim e^{-\lambda x}, \quad (5.2.9)$$

is a conserved quantity, and that the propagation velocity depends on this steepness as follows

$$v_\lambda = \lambda + \frac{1}{\lambda}. \quad (5.2.10)$$

It is therefore possible to obtain propagating fronts that are arbitrarily fast, by taking very weakly decaying initial conditions. In most physical

systems however, the initial conditions are localized, *i.e.* are nonzero only in a finite region of space. The question of selection therefore remains, but should be posed as follows: what is the asymptotic propagation velocity for a front starting from sufficiently localized initial conditions?

5.3 Velocity selection.

To derive this asymptotic velocity, consider the following general, linearized equation

$$(\partial_t - \hat{\mathcal{L}}_0)\varphi(x, t) = 0, \quad (5.3.1)$$

where for definiteness, one could consider the linearized FK equation, which would have for its linear operator $\mathcal{L}_0 = \partial_x^2 - \varepsilon$. We will assume the linear differential operator to be a sum of different order spatial derivatives, and possibly multiplication with a constant, *i.e.* $\hat{\mathcal{L}}_0 = \hat{\mathcal{L}}_0(1, \partial_x^n)$. In a Green's function formulation of this problem, initial conditions $\varphi(y, 0)$ evolve as

$$\varphi(x, t) = \int_{-\infty}^{+\infty} dy \mathcal{G}(x - y, t) \varphi(y, 0), \quad (5.3.2)$$

with

$$\mathcal{G}(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{ikx - i\omega(k)t}, \quad (5.3.3)$$

the Green's function. $\omega(k)$ is the linear dispersion relation, as determined from the linearized equation 5.3.1.

$$\omega(k) = ie^{-ikx} \hat{\mathcal{L}}_0 e^{ikx}. \quad (5.3.4)$$

As we will be interested in *asymptotic* states of this system, we consider now the large-time asymptotics of the Green's function. Furthermore, we assume that asymptotically the front velocity approaches some constant velocity which we shall call v^* . In the asymptotic frame, *i.e.* the one comoving with velocity v^* with comoving coordinate $\xi = x - v^*t$, Eq. (5.3.3) reads

$$\mathcal{G}(\xi, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{ik\xi - i[\omega(k) - v^*k]t}. \quad (5.3.5)$$

In the large-time limit, the Green's function will be dominated by the saddle point contribution. Let us call the locus of the saddlepoint k^* , and expand the exponent appearing in the Green's function around that point

$$\begin{aligned}
ik\xi - i[\omega(k) - v^*k]t &\approx ik^*\xi - i[\omega(k^*) - v^*k^*]t \\
&+ \left(i\xi - i \left[\frac{\partial\omega}{\partial k} \Big|_* - v^* \right] t \right) (k - k^*) \\
&+ \frac{1}{2} \left(-i \frac{\partial^2\omega}{\partial k^2} \Big|_* t \right) (k - k^*)^2 \\
&+ \frac{1}{6} \left(-i \frac{\partial^3\omega}{\partial k^3} \Big|_* t \right) (k - k^*)^3 + \dots, \quad (5.3.6)
\end{aligned}$$

the notation $|_*$ indicates the derivative concerned is to be evaluated at k^* . Introducing now the new variable

$$k - k^* = \frac{\kappa}{\sqrt{t}} \quad (5.3.7)$$

and the abbreviation

$$D_n = \frac{i}{n!} \frac{\partial^n \omega}{\partial k^n} \Big|_* \quad , \quad n \geq 2, \quad (5.3.8)$$

we can expand the entire Green's function around k^* as follows

$$\begin{aligned}
\mathcal{G}(\xi, t) &\approx \exp(ik^*\xi - i[\omega(k^*) - v^*k^*]t) \int_{-\infty}^{+\infty} \frac{d\kappa}{2\pi\sqrt{t}} \otimes \\
&\otimes \exp \left(i \frac{\kappa}{\sqrt{t}} \xi - i\kappa\sqrt{t} \left[\frac{\partial\omega}{\partial k} \Big|_* - v^* \right] - D_2\kappa^2 + \mathcal{O} \left(\frac{D_3\kappa^3}{\sqrt{t}} \right) \right). \quad (5.3.9)
\end{aligned}$$

The point k^* that we are expanding around is indeed a saddle point when the *phase* of the integrand at that point is stationary for $t \rightarrow \infty$, which amounts here to demanding that the coefficient of \sqrt{t} vanishes identically

$$\frac{\partial\omega}{\partial k} \Big|_* = v^*, \quad (5.3.10)$$

or equivalently

$$\frac{\partial \operatorname{Im} \omega}{\partial \operatorname{Im} k} \Big|_* = v^*, \quad (5.3.11a)$$

$$\frac{\partial \operatorname{Im} \omega}{\partial \operatorname{Re} k} \Big|_* = 0. \quad (5.3.11b)$$

If this saddle point condition is indeed fulfilled, the asymptotic Green's function assumes the following form

$$\mathcal{G}(\xi, t) \approx e^{ik^*\xi - i[\omega(k^*) - v^*k^*]t} \int_{-\infty}^{+\infty} \frac{d\kappa}{2\pi\sqrt{t}} e^{-\frac{\xi^2}{4D_2t}} e^{-D_2\left(\kappa - \frac{i\xi}{2D\sqrt{t}}\right)^2} + \text{h.o.t} \quad (5.3.12)$$

$$\approx \frac{1}{\sqrt{4\pi D_2t}} e^{-\frac{\xi^2}{4D_2t}} \left(1 + \mathcal{O}\left(\frac{D_3\xi}{D_2^2t}\right)\right) e^{ik^*\xi - i[\omega(k^*) - v^*k^*]t} \quad (5.3.13)$$

When we apply the Green's function in this form to Eq. (5.3.2), we see that we can obtain the asymptotic profiles from the linear evolution of initial conditions using the following equation

$$\lim_{t \rightarrow \infty} \varphi(\xi, t) = \bar{\varphi}(k^*, t=0) e^{ik^*\xi} e^{-i[\omega(k^*) - v^*k^*]t} \frac{e^{-\frac{\xi^2}{4D_2t}}}{\sqrt{4\pi D_2t}} \left(1 + \mathcal{O}\left(\frac{D_3\xi}{D_2^2t}\right)\right). \quad (5.3.14)$$

For what we have derived up to now to be consistent, we have to realize that our coordinate ξ should be chosen such that we are in the frame comoving with the asymptotic velocity v^* . In that frame, the (asymptotic) profile as we have just derived it should by definition neither grow nor decay. Besides the saddlepoint condition, there is therefore one more condition to fulfill, the *self-consistency condition*

$$\text{Im}(\omega(k^*) - v^*k^*) = 0. \quad (5.3.15)$$

This condition and Eq. (5.3.10) constitute what is known in the literature as the Linear Marginal Stability (LMS) criterion, which enables one to derive the asymptotic spreading speed of fronts. For sufficiently steep initial conditions, the propagation velocity is therefore surprisingly determined by the *linear equations* only. It is as if these such fronts are pulled along by the region where the linearized equations suffice, which is of course the tip region or the *leading edge* as we shall call it. The class of fronts for which the recipe as derived in this section does indeed produce the correct v^* is therefore known as the class of *pulled fronts*². The actual velocity that comes out of this argument for the FK equation is $v^* = 2\sqrt{\varepsilon}$, a value we encountered in the previous section where it was called v^c , the lower

²As opposed to the class of *pushed fronts*, which consists of those fronts for which the nonlinearities *do* play an important role in determining the propagation velocity. Such pushed fronts are characterized by a propagation velocity $v^\dagger > v^*$. We will not go into those in any detail here, and instead refer to [58]

bound on the velocity of a stable front. As we have seen, the front propagating at this velocity asymptotically behaves like $(\alpha\xi + \beta)e^{-\frac{v^*c}{2}\xi}$. The spatial decay rate in the comoving frame is the same quantity as $\text{Im}k^*$. From now on, to avoid confusion, we will write $k^* = q^* + i\lambda^*$. It is important to realize that Eqs. (5.3.10) and (5.3.15) determine *all* of these quantities. In calculations further on in this chapter, we will sometimes also use the dispersion relation in the comoving frame $\Omega(k) = \omega(k) - v^*k$, for which the LMS equations read

$$\left. \frac{\partial \Omega(k)}{\partial k} \right|_* = 0, \quad \text{Im}(\Omega(k^*)) = 0. \quad (5.3.16)$$

The fact that the propagation velocity is determined by the linear equations does not imply that the nonlinearities do not play any part in the propagation of pulled fronts. In fact, we shall see they enter in a very subtle way into the relaxation process.

5.4 Uniformly translating pulled fronts

The simplest types of fronts are those which have already encountered in this chapter, those for which the dynamical field $\varphi(x, t)$ asymptotically approaches a uniformly translating profile $\varphi \equiv \Phi_{v^*}(\xi)$, $\xi = x - v^*t$. If we define *level curves* as the lines in an x, t diagram where $\varphi(x, t)$ has a particular value, we can define the velocity $v(t)$ as the slope of a level curve. As we already seen, these fronts asymptotically have a propagation velocity v^* . An important thing to know besides that is how the front velocity *relaxes* to this value. Does this happen exponentially fast, or is there some other mechanism? For the case of uniformly translating pulled fronts, this question was studied and answered in detail in [58]. The outcome for the FK-equation (with ε set to 1) is

$$v(t) = v^* + \dot{X}(t) \quad (5.4.1a)$$

$$\dot{X}(t) = -\frac{3}{2t} + \frac{3\sqrt{\pi}}{2t\sqrt{t}} + \mathcal{O}\left(\frac{1}{t^2}\right), \quad (5.4.1b)$$

which shows that relaxation in this system is in fact *algebraic* and *not* exponential. The front *profile* also relaxes in time, and to lowest order was shown to do so like

$$\varphi(x, t) = \Phi_{v(t)}(\xi_X) + \mathcal{O}(t^{-2}) \quad \xi_X \ll \sqrt{t}, \quad (5.4.2a)$$

$$\xi_X = x - v^*t - X(t), \quad (5.4.2b)$$

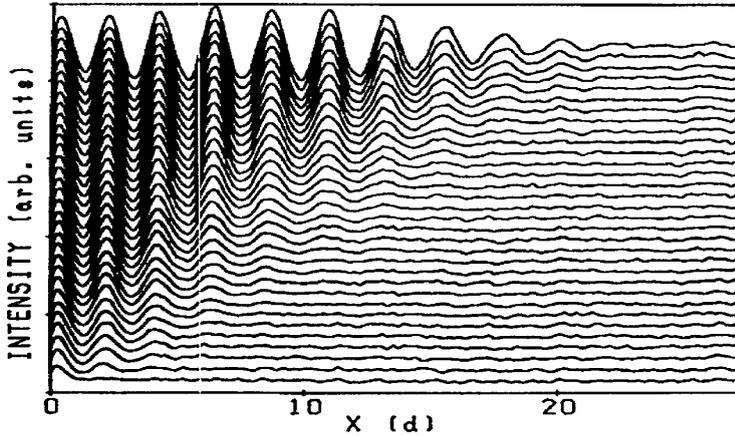


Figure 5.1: Contour plot of the light intensity profiles as a function of position along the length of a Rayleigh-Bénard cell. The light intensity is a measure of the local fluid velocity, *i.e.* rolls are penetrating the unstable conducting state. This is an example of a pattern forming pulled front. Figure taken from [103]

The algebraic relaxation has important implications, as it indicates that relaxation processes for such propagating fronts do not have an intrinsic timescale for transients to die out, which can therefore take arbitrarily long. The power-law relaxation was checked in detail, and found to be accurate even up to the subdominant correction term (not surprisingly so, since the results in Eqs. (5.4.1) are *exact*). Moreover, this relaxation was found to be *universal* up to that same order, and a large class of equations (more precisely the class of equations that possess pulled uniformly translating fronts) was shown to display the same relaxation behavior.

However, the most relevant *experimental* realizations of pulled fronts propagating into unstable states do not form uniformly translating fronts, but instead tend to generate *patterns*, leaving behind a (nearly) periodic pattern. Examples of such systems are Taylor vortex fronts [102], fronts in Rayleigh-Bénard convection [103] and the pearling instability [104, 105]. Fig. 5.1 for instance shows experimental data of Rayleigh-Bénard rolls penetrating the unstable conducting state. Since the derivation of the relaxation behavior for uniformly translating fronts did make use of the fact that the state behind the front was relatively simple, the calculation

cannot simply be translated into a prediction for such pattern forming fronts. In this Chapter, we will present a formalism that treats *all* pulled fronts on equal footing, irrespective of the type of state they leave behind. As we will show, we will not need any knowledge of the state behind the front *at all*. Instead, it allows us to focus directly on the leading edge region. This argument will produce the same results as those that were derived for uniformly translating fronts in [58], but in addition the new formalism can be directly implemented for the cases of pattern forming fronts and even chaotic fronts, we shall see. Let us first introduce the model equations we will be using for the pattern forming and chaotic fronts, as well as some other useful definitions.

5.5 Coherent pattern generating fronts

As an example of coherent pattern generating fronts, we consider again the Swift-Hohenberg (SH) equation

$$\partial_t u = \varepsilon u - (1 + \partial_x^2)^2 u - u^3, \quad \varepsilon > 0. \quad (5.5.1)$$

The space-time plot of Fig. 5.2(a) illustrates how SH-fronts with steep initial conditions (falling off faster than $e^{-\lambda^* x}$ as $x \rightarrow \infty$ into the unstable state $u=0$) generate a periodic pattern. It is known that they are pulled [106, 107, 108]. In this case, new level curves in an x, t plot are constantly being generated. If we define in this case the velocity as the slope of the uppermost level curve, one gets an oscillatory function. Its average is what we shall call $v(t)$ in this case. It is however difficult to extract $v(t)$ this way, and numerically it is better to determine the velocity from an empirical envelope obtained by interpolating the positions of the maxima. Since these pattern forming front solutions for long times have a temporal periodicity $u(\xi, t) = u(\xi, t + T)$ in the frame $\xi = x - vt$ moving with the velocity v of the front, the asymptotic profiles can be written in the form $\sum_{n=1} e^{-2\pi i n t/T} U_v^n(\xi) + c.c.$. In terms of these complex modes U , our result for the shape relaxation of the pulled front profile becomes in analogy to (5.4.2)

$$u(x, t) \simeq \sum_{n=1} e^{-ni\Omega^* t - ni\Gamma(t)} U_{v(t)}^n(\xi_X) + c.c. + \dots \quad (5.5.2)$$

with the frequency Ω^* given below. Underlying (5.5.2) is the assumption that the (5.5.1) admits a two-parameter family of front solutions. This was shown for small ε in [109], and is demonstrated by counting arguments

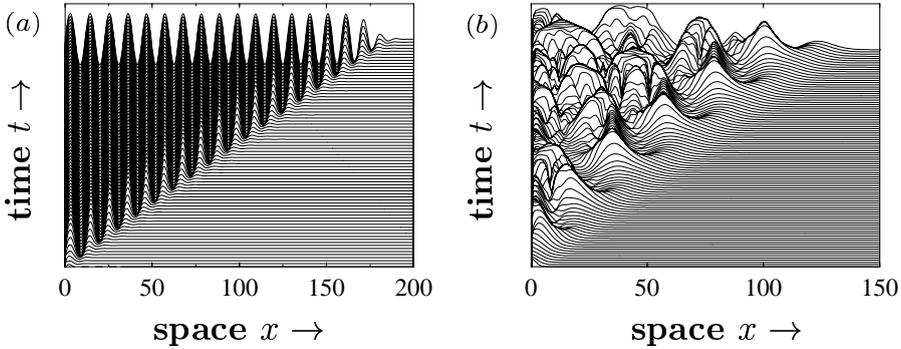


Figure 5.2: (a) Space-time plot of a pulled front in the SH eq. (5.5.1) with $\varepsilon = 5$ and Gaussian initial conditions. Time steps between successive lines are 0.1. (b) A pulled front in the QCGL eq. (5.6.1) with $\varepsilon = 0.25$, $C_1 = 1$, $C_3 = C_5 = -3$, and Gaussian initial conditions. Plotted is $|A(x, t)|$. Time steps between lines are 1.

for arbitrary ε in [108]. Eq. (5.5.2) shows that $\Gamma(t)$ is the *global phase* of the relaxing profile, as the functions U_v^n only have a ξ_X -dependence. We stress that while for $\varepsilon \rightarrow 0$, an Ansatz like (5.5.2) leads to an amplitude equation for the $n = \pm 1$ terms, our analysis applies for *any* $\varepsilon > 0$.

5.6 Incoherent or chaotic fronts

The third class we consider consists of fronts which leave behind chaotic states. They occur in some regions of parameter space in the cubic Complex Ginzburg-Landau equation [110] or in the quintic extension (QCGL) [21] that we consider here,

$$\begin{aligned} \partial_t A = \varepsilon A + (1 + iC_1)\partial_x^2 A &+ (1 + iC_3)|A|^2 A \\ &- (1 - iC_5)|A|^4 A. \end{aligned} \quad (5.6.1)$$

Fig. 5.2(b) shows an example of a pulled front in this equation. Level curves in a space-time diagram can now also both start and end. If we calculate the velocity from the slope of the uppermost level line, then its average value is what we shall call $v(t)$ in this case, but the oscillations can be quite large. This is true for chaotic fronts provided that the temporal correlation function for the chaotic variable falls off at least as fast as t^{-2} , so that the temporal change of the average velocity $v(t)$ can be considered

adiabatically. However, our analysis confirms what is already visible in Fig. 5.2(b), namely that even a chaotic pulled front becomes more coherent the further one looks into the leading edge of the profile. Indeed we will see that in the leading edge where $|A| \ll 1$ the profile is given by an expression reminiscent of (5.5.2),

$$A(x, t) \approx e^{-i\Omega^* t - i\Gamma(t)} e^{ik^* \xi_X} \psi(\xi_X), \quad 1 \ll \xi_X \ll \sqrt{t}. \quad (5.6.2)$$

The fluctuations about this expression become smaller the larger ξ_X .

5.7 Choosing the proper frame and transformation

Eq. (5.3.13), especially when we slightly rewrite it as

$$G(\xi, t) \approx e^{ik^* \xi - i\Omega^* t} \frac{e^{-\frac{\xi^2}{4Dt}}}{\sqrt{4\pi Dt}} \quad (5.7.1)$$

not only confirms that a localized initial condition will grow out and spread asymptotically with the velocity v^* given by (5.3.16), but the Gaussian factor also determines how the asymptotic velocity is approached in the fully linear case. Our aim now is to understand the convergence of a pulled front due to the interplay of the linear spreading and the nonlinearities. The Green's function expression (5.3.13) gives three important hints in this regard: First of all, $G(\xi, t)$ is asymptotically of the form $e^{ik^* \xi - i\Omega^* t}$ times a crossover function whose diffusive behavior is betrayed by the Gaussian form in (5.3.13). Hence if we write our dynamical fields as $A = e^{ik^* \xi - i\Omega^* t} \psi(\xi, t)$ for the QCGL (5.6.1) or $u = e^{ik^* \xi - i\Omega^* t} \psi(\xi, t) + c.c.$ for the real field u in (5.5.1), we expect that the dynamical equation for $\psi(\xi, t)$ obeys a diffusion-type equation. Second, as we have argued in [111], for the relaxation analysis one wants to work in a frame where the crossover function ψ becomes asymptotically time independent. This is obviously not true in the ξ frame, due to the factor $1/\sqrt{t}$ in (5.3.13). However, this term can be absorbed in the exponential prefactor, by writing $t^{-\nu} e^{ik^* \xi - i\Omega^* t} = e^{ik^* \xi - i\Omega^* t - \nu \ln t}$. Hence, we introduce the logarithmically shifted frame $\xi_X = \xi - X(t)$ [111] as already used in (5.4.2b). Third, we find a new feature specific for pattern forming fronts: the complex parameters, and D in particular, lead us to introduce the global phase $\Gamma(t)$. We

expand $\dot{\Gamma}(t)$ like $\dot{X}(t)$ [111]

$$\dot{X}(t) = \frac{c_1}{t} + \frac{c_{3/2}}{t^{3/2}} + \dots, \quad \dot{\Gamma}(t) = \frac{d_1}{t} + \frac{d_{3/2}}{t^{3/2}} + \dots \quad (5.7.2)$$

and analyze the long time dynamics by performing a so-called “leading edge transformation” [58] to the field ψ ,

$$\begin{aligned} \text{QCGL: } A &= e^{ik^*\xi_X - i\Omega^*t - i\Gamma(t)} \psi(\xi_X, t), \\ \text{SH: } u &= e^{ik^*\xi_X - i\Omega^*t - i\Gamma(t)} \psi(\xi_X, t) + c.c. \end{aligned} \quad (5.7.3)$$

Step initial conditions imply that $\psi(\xi_X, t) \rightarrow 0$ as $\xi_X \rightarrow \infty$. The determination of the coefficients in the expansions (5.7.2) of \dot{X} and $\dot{\Gamma}$ will be the main goal of the subsequent analysis.

5.8 Understanding the intermediate asymptotics

Substituting the leading edge transformation (5.7.3) into the nonlinear dynamical equations, we get

$$\begin{aligned} \partial_t \psi &= D \partial_{\xi_X}^2 \psi + \sum_{n=3} D_n \partial_{\xi_X}^n \psi \\ &+ [\dot{X}(t)(\partial_{\xi_X} + ik^*) + i\dot{\Gamma}(t)]\psi - N(\psi), \end{aligned} \quad (5.8.1)$$

with

$$D_n = \left. \frac{-i}{n!} \frac{\partial^n \Omega}{\partial ik^n} \right|_*. \quad (5.8.2)$$

Note the slight difference in definition with 5.3.8 (of course, for the QCGL, $\Omega(k)$ is quadratic in k , so $D_n = 0$). In this equation, N accounts for the nonlinear terms; e.g., for the QCGL, we simply have

$$N = e^{-2\lambda^*\xi_X} |\psi|^2 \psi [1 - iC_3 + (1 - iC_5)e^{-2\lambda^*\xi_X} |\psi|^2]. \quad (5.8.3)$$

The expression for the SH equation is similar.

The structure of Eq. (5.8.1) is that of a diffusion-type equation with $1/t$ and higher order corrections from the \dot{X} and $\dot{\Gamma}$ terms, and with a nonlinearity N . The crucial point to recognize now is that for fronts, N is nonzero *only in a region of finite width*: For $\xi_X \rightarrow \infty$, N decays exponentially due to the explicit exponential factors in (5.8.3). For $\xi_X \rightarrow -\infty$, N also decays exponentially, since u and A remain finite, so that ψ

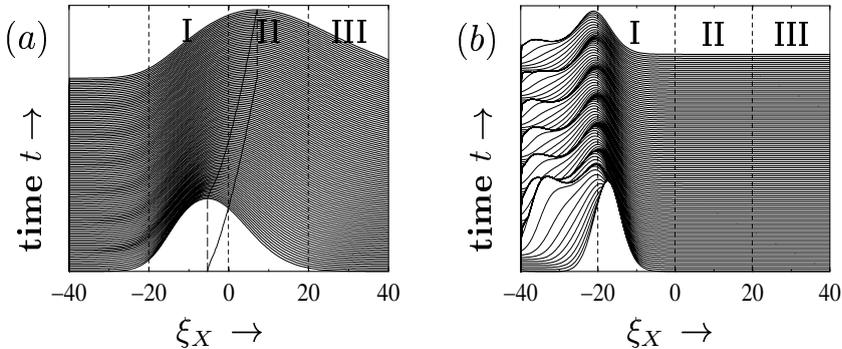


Figure 5.3: (a) and (b): Simulation of the QCGL Equation as in Fig. 5.2(b) for times $t = 35$ to 144. (a) shows $|N|$ (5.8.3) as a function of ξ_X . (b) shows $|\psi|$, which in region I builds up a linear slope $\psi \propto \alpha \xi_X$, and in region III decays like a Gaussian widening in time. The lines in region II show the maxima of $\psi(\xi_X, t)$ for fixed t and their projection $\xi_X \sim \sqrt{t}$ into the (ξ_X, t) plane.

decays as $e^{-\lambda^*|\xi_X|}$ according to (5.7.3). Intuitively, therefore, we can think of (5.8.1) as a diffusion equation in the presence of a sink N localized at some finite value of ξ_X . The ensuing dynamics of ψ to the right of the sink can be understood with the aid of Figs. 5.3, which are obtained directly from the time-dependent numerical simulations of the QCGL (5.6.1). To extract the intermediate asymptotic behavior illustrated by these plots, we integrate (5.8.1) once to get

$$\begin{aligned} \partial_t \int_{-\infty}^{\xi_X} d\xi'_X \psi &= D \partial_{\xi_X} \psi + \sum_{n=3} \frac{D_n}{n-1} \partial_{\xi_X}^{n-1} \psi + \\ &+ i[k^* \dot{X}(t) + \dot{\Gamma}(t)] \int_{-\infty}^{\xi_X} d\xi'_X \psi + \dot{X}(t) \psi - \int_{-\infty}^{\xi_X} d\xi'_X N(\psi) \end{aligned} \quad (5.8.4)$$

Now, in the region labeled I in Figs. 5.3, we have for fixed ξ_X and $t \rightarrow \infty$ that the terms proportional to \dot{X} and $\dot{\Gamma}$ can be neglected upon averaging over the fast fluctuations; the same holds for the term on the left. Since the integral converges quickly to the right due to the exponential factors in N , we then get immediately, irrespective of the presence of higher order spatial derivatives

$$\lim_{t \rightarrow \infty} D \frac{\overline{\partial \psi}}{\partial \xi_X} = \int_{-\infty}^{\infty} d\xi_X \overline{N(\psi)} \equiv \alpha D. \quad (5.8.5)$$

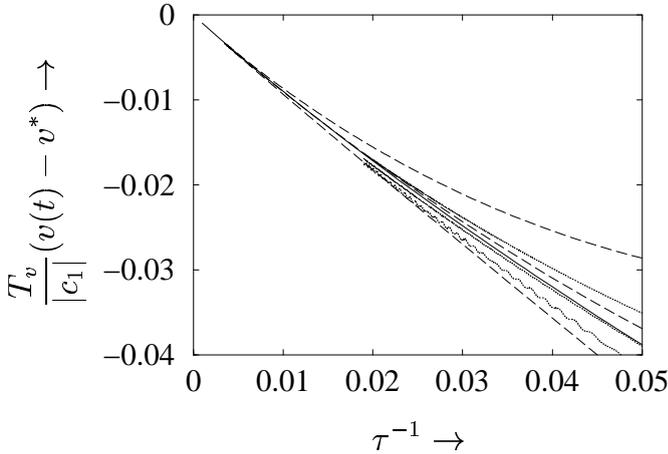


Figure 5.4: Scaling plot of the velocity relaxation $(v(t) - v^*) \cdot T_v / |c_1|$ vs. $1/\tau$ with $\tau = t/T_v$ and characteristic time $T_v = (c_{3/2}/c_1)^2$. Plotted are from top to bottom the data for the SH eq. for heights $u = 0.0001\sqrt{\epsilon}$, $0.01\sqrt{\epsilon}$, and $\sqrt{\epsilon}$ ($\epsilon = 5$) as dashed lines, and for the QCGL eq. (5.6.1) for heights $|A| = 0.00002$, 0.0002 , and 0.002 as dotted lines. The solid line is the universal asymptote $-1/\tau + 1/\tau^{3/2}$.

Here, the overbar denotes a time average (necessary for the case of a chaotic front). Thus, for large times in region I, $\bar{\psi} \approx \alpha \xi_X + \beta$ in dominant order. Moreover, from the diffusive nature of the equation, our assertion that the fluctuations of ψ rapidly decrease to the right of the region where N is nonzero comes out naturally. In other words, provided that the time-averaged sink strength α is nonzero, $\alpha \neq 0$, one will find a buildup of a linear gradient in $|\bar{\psi}|$ in region I, *independent of the precise form of the nonlinearities or of whether or not the front dynamics is coherent*. This behavior is clearly visible in Fig. 5.3(b). We can understand the dynamics in regions II and III along similar lines. In region III the dominant terms in (5.8.1) are the one on the left and the first one on the second line, and the cross-over region II which separates regions I and III moves to the right according to the diffusive law $\xi_X \sim D\sqrt{t}$.

5.9 Systematic expansion

These considerations are fully corroborated by our extension of the analysis of [58]. Anticipating that ψ falls off for $\xi_X \gg 1$, we split off a Gaussian

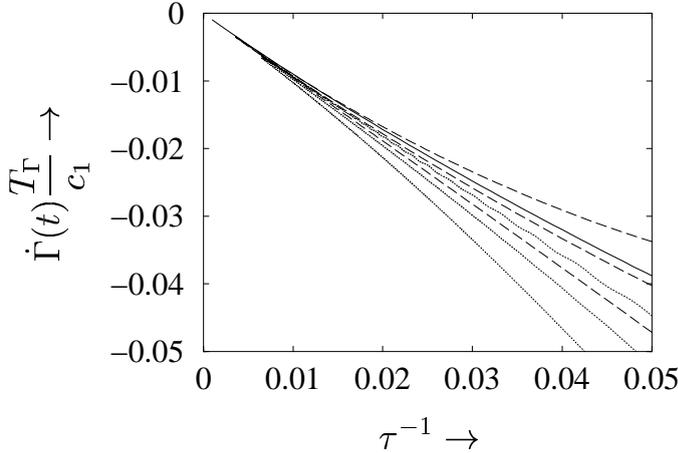


Figure 5.5: Scaling plot for the phase relaxation. From top to bottom: SH (dashed) for $u = 0.0001\sqrt{\epsilon}$, $0.01\sqrt{\epsilon}$, and $\sqrt{\epsilon}$ ($\epsilon = 5$), and QCGLE (dotted) for $|A| = 0.002$, 0.0002 , and 0.00002 . Plotted is $\dot{\Gamma}(t) \cdot T_{\Gamma}/c_1$ vs. $1/\tau$. Here $\tau = t/T_{\Gamma}$, and $T_{\Gamma} = T_v \cdot [1 + \lambda^* \text{Im}D^{-1/2}/(q^* \text{Re}D^{-1/2})]$. The solid line again is the universal asymptote $-1/\tau + 1/\tau^{3/2}$.

factor by writing $\psi(\xi_X, t) = G(z, t) e^{-z}$ in terms of the similarity variable $z = \xi_X^2/(4Dt)$, and expand

$$G(z, t) = t^{1/2} g_{-\frac{1}{2}}(z) + g_0(z) + t^{-1/2} g_{\frac{1}{2}}(z) + \dots \quad (5.9.1)$$

This, together with the expansion (5.7.2) for $X(t)$ and $\Gamma(t)$, the left “boundary condition” that $\psi(\xi_X, t \rightarrow \infty) = \alpha\xi_X + \beta$ and the condition that the functions $g(z)$ do not diverge exponentially, then results in the following expressions for $\dot{X}(t)$ and for $\dot{\Gamma}$ (we will not report the expansion in detail here, but instead refer the reader to [58] for details)

$$v(t) \equiv v^* + \dot{X}(t) \quad (5.9.2)$$

$$\dot{X}(t) = -\frac{3}{2\lambda^* t} + \frac{3\sqrt{\pi}}{2\lambda^* t^{3/2}} \text{Re} \left(\frac{1}{\sqrt{D_2}} \right) + \mathcal{O} \left(\frac{1}{t^2} \right), \quad (5.9.3)$$

$$\dot{\Gamma}(t) = -q^* \dot{X}(t) - \frac{3\sqrt{\pi}}{2\lambda^* t^{3/2}} \text{Im} \left(\frac{1}{\sqrt{D_2}} \right) + \mathcal{O} \left(\frac{1}{t^2} \right). \quad (5.9.4)$$

These predictions were checked numerically for the case of the QCGLE and the SH equation. The results are gathered in Figs. 5.4 and 5.5,

which indeed show the correctness of our results for these two types of equations. For the QCGL, the analysis immediately implies the result (5.6.2) for the front profile in the leading edge. In addition for the SH equation, one arrives at (5.5.2) for the shape relaxation in the front interior along the lines of [111]: Starting from the *o.d.e.*'s for the U_v^n , one finds upon transforming to the frame ξ_X that to $\mathcal{O}(t^{-2})$, the time dependence only enters parametrically through $v(t)$. This then yields (5.5.2). Finally, we note that the velocity and phase relaxation also imply that another experimentally accessible quantity should display power-law relaxation. This is the wavenumber $\Lambda(t)$ of pattern forming fronts directly behind the front, and for this quantity Eqs. (5.9.2) and (5.9.4) imply that

$$\Lambda(t) = 2\pi \left| \frac{v^* + \dot{X}(t)}{\Omega^* + \dot{\Gamma}(t)} \right| + \mathcal{O}\left(\frac{1}{t^2}\right). \quad (5.9.5)$$

5.10 Conclusion

In conclusion, we have shown that the long time relaxation of pulled fronts is remarkably universal: independent of whether fronts are uniformly translating, pattern generating or chaotic, the velocity and phase relaxation is governed by one simple formula, with universal dominant and subdominant power law expressions. Our results apply to the entire class of pulled fronts, which also includes front-forming systems that are neither uniformly translating, pattern forming *or* chaotic. One notable example for which we checked the results was the Cahn-Hilliard equation, a fourth order equation with conserved order parameter widely used to describe spinodal decomposition and coarsening.

Fractal Lasers

6.1 Introduction

Lasers have been with us for over forty years now. They have become an indispensable tool for research and technology, and their physics is well understood. Nevertheless, every so often new aspects are still discovered, and this is precisely what was done recently by Karman *et al.* [112, 113].

Usually, lasers are constructed in such a way as to operate at one single transverse mode. This regime may be the most obvious one when lasers are used as a preferably stable tool in experiments and applications, but even in some single mode regimes lasers sometimes do exhibit dynamical behavior which is of intrinsic interest. For example, lasers have been studied as an experimental realization of low-dimensional nonlinear dynamics, associated with the fact that the laser equations map onto the Lorenz equations. Furthermore, for strong detuning multiple transverse modes can become active. One can then enter a regime where patterns in the transverse direction can emerge spontaneously. Such transverse patterns have been studied as an example of non-equilibrium pattern formation, and the laser systems are linked to other pattern forming system by the complex Ginzburg-Landau equation, which emerges as the universal amplitude equation near threshold for the instability.

The interesting twist to the ongoing laser story added recently by Karman *et al.* is the possibility of obtaining transverse intensity profiles with fractal scaling properties. Their focus in this case was not on the dynamics, but rather they showed that in an unstable cavity laser, depending on

the mirror shape (triangular, octagonal, etc.), the two-dimensional transverse intensity profiles show self-similar structure reminiscent of fractals like the Sierpinsky gasket.

Indeed, Karman and Woerdman found numerically that the eigenmodes of unstable cavity lasers possess fractal scaling, and using the box-counting method they estimate the fractal dimension D_f to be about 1.6 – 1.7 in one dimension. For a two-dimensional systems with a circular aperture, a similar study produced a value of $D_f \approx 1.3$.

The fractal scaling properties of unstable cavity lasers can be approached from various angles. In the confocal unstable cavity, sketched in Fig. 6.1, the intensity profile is magnified in the transverse direction (one might say “stretched”) upon each round-trip through the cavity. At the same time, that part of the initial intensity profile which is projected outside of the mirror is lost. Courtial and Padgett have termed this the “monitor-outside-a-monitor effect”. It leads one to consider sequences of functions of the form

$$f_n(x) = \sum_{i=0}^{n-1} p\left(\frac{x}{M^i}\right), \quad (6.1.1)$$

starting from some initial function $p(x)$. In Eq.(6.1.1), M is the amplification (or stretching) factor, so that Eq.(6.1.1) expresses the n^{th} iteration $f_n(x)$ as the sum of the initial function $p(x)$, the original function stretched by a factor M , $p(x)$ stretched by M^2 , and so forth. As Courtial and Padgett discuss, Eq.(6.1.1) captures an important ingredient of the unstable cavity laser mode profiles, although it appears that from an expression like Eq.(6.1.1) by itself one cannot establish the fractal dimension. The mathematical formulation (6.1.1) is a very convenient way to focus on the basic features of the unstable cavity laser, but as it stands, it does not account for the fact that besides magnification *diffraction* also plays a major role in the development of fractal structures in this setup. Diffractive effects are in essence nonlocal, as they involve the interference of waves originating from different locations. Courtial and Padgett get around this by endowing the initial function $p(x)$ with a complex phase $e^{i\varphi(x)}$, where they obtain $\varphi(x)$ from the actual eigenfunction of the cavity, which has therefore to be calculated first using different methods. Presumably, the diffraction is crucial in both the generation of the fractal modes and the determination of their dimension.

New *et al.* have studied this aspect of the fractal laser modes in more detail. Their formulation in terms of the virtual source method expresses

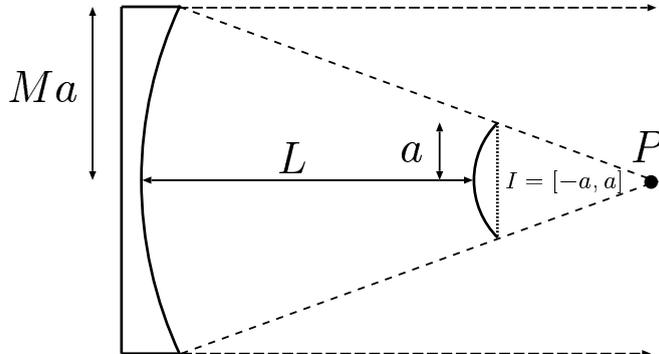


Figure 6.1: A confocal unstable cavity : The two mirrors share a common focal point P . We calculate mode profiles along the interval $I = [-a, a]$.

the cavity eigenmodes as a superposition of Fresnel diffraction patterns, determined by the virtual images of the defining aperture of the system. This formulation is especially useful to understand several prominent features of the detailed features of the power spectra of the mode profiles. Moreover, New *et al.* have analyzed the fractal dimension of the modes by considering the power spectrum, and what they find is a dimension D_f very close to 1.5. These authors also noted that the Fresnel spectrum of a *single* slit falls off with the wavenumber k like k^{-2} , and that this suggests that the spectrum of the unstable cavity modes would also fall off like k^{-2} . Provided the *phases* of the different k -modes are essentially random, a power spectrum decaying like k^{-2} is indeed consistent with a fractal dimension $D_f = 1.5$.

In this Chapter, we will present both analytical arguments, as well as numerical evidence for a fractal dimension of exactly $3/2$ for these unstable resonator eigenmodes. Although our analysis falls short of a proof, our results together with those of [114] yield quite compelling evidence that the fractal dimension is indeed $3/2$ for *any* magnification factor $M > 1$. Moreover our analysis strongly suggests that also in higher dimensions, the intensity variations along a typical cross-section should show fractal scaling with dimension $3/2$.

In section 6.2, we introduce the Huygens-Fresnel equation for the problem in one dimension, and discuss the connection between the spectrum of the eigenmodes and the fractal dimension. Then, in section 6.3 we dis-

cuss an analytical approximation to the Huygens-Fresnel integral for the (lowest-loss) eigenmode, which yields $D_f = 3/2$. Finally, in section 6.4 we present numerical results that are also consistent with this value of D_f .

6.2 Setup of the problem

Throughout this Chapter, we will focus on one particular realization of the unstable-cavity laser, a one-dimensional hard-edged confocal unstable cavity. This setup is depicted schematically in Fig. 6.1.

The effect of one round trip in such a cavity is described in the paraxial approximation by the Huygens-Fresnel integral [115]

$$\begin{aligned} u_{i+1}(x) &= \sqrt{\frac{i}{B\lambda}} \int_{-a}^{+a} dx' u_i(x') e^{-\frac{i\pi}{B\lambda}(Ax'^2 - 2xx' + Dx^2)} \\ &\equiv \int_{-a}^{+a} dx' \mathbb{T}(x, x') u(x'). \end{aligned} \quad (6.2.1)$$

where $u_i(x)$ and $u_{i+1}(x)$ are the transverse field before and after the roundtrip, respectively. The size of the small mirror is measured by $2a$, and λ is the wavelength of the light. For the case of the confocal unstable cavity, the $ABCD$ -matrix, widely used in optics to characterize the effect a given setup of mirrors and lenses has on a beam passing through it, is given by

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} = \begin{pmatrix} M & \frac{(M+1)L}{M} \\ 0 & \frac{1}{M} \end{pmatrix}, \quad (6.2.2)$$

Here, M is the round-trip linear magnification and L is the cavity length. The eigenmodes are fully characterized by only two numbers, the magnification M and the equivalent Fresnel number N , given by

$$N = \frac{1}{2}(M-1)\frac{a^2}{\lambda L}. \quad (6.2.3)$$

The eigenmodes $u_n(x)$ and eigenvalues (γ_n) of this cavity are defined by

$$\gamma_n u_n(x) = \int_{-a}^{+a} dx' \mathbb{T}(x, x') u_n(x'). \quad (6.2.4)$$

with the kernel \mathbb{T} given explicitly by

$$\mathbb{T}(x, x') = \sqrt{\frac{2iMN}{(M^2-1)a^2}} e^{\frac{-2i\pi M^2 N}{(M^2-1)a^2} \left(x' - \frac{x}{M}\right)^2}. \quad (6.2.5)$$

Experimentally, the most relevant mode is the one corresponding to the largest (absolute) eigenvalue γ_0 , which is called the *lowest loss mode*. As mentioned in the introduction, the numerical evidence for fractal scaling that Karman and Woerdman found, was obtained using the so-called box-counting algorithm. This method is known to work well for self-similar curves. For self-affine curves however, which scale anisotropically (like, for instance, the trace of a random walker as a function of time), this method in general does *not* produce the correct dimension. For the latter class of curves, the proper way of defining the fractal dimension is through the power spectrum. This is therefore the method we will employ here. Suppose the Fourier series of some function $u(x)$ reads

$$u(x) = \sum_k \bar{u}_k e^{ikx}, \quad (6.2.6)$$

where the \bar{u}_k 's have random or pseudo-random phases. If the power spectrum $\mathcal{P}(k) \equiv |\bar{u}_k|^2$ behaves asymptotically like

$$\lim_{|k| \rightarrow \infty} \mathcal{P}(k) \sim |k|^{-\beta}, \text{ where } 1 < \beta \leq 3, \quad (6.2.7)$$

the graphs of $\text{Re}(u)$ and $\text{Im}(u)$ will be continuous but non-differentiable, with fractal dimension

$$D_f = \frac{1}{2}(5 - \beta). \quad (6.2.8)$$

Almost always (in absence of close connections between $\text{Re}(u)$ and $\text{Im}(u)$), the graph of $|u(x)|^2$ also has the same fractal dimension D_f [116]. In practice, and particularly for the case considered here, the power law scaling (6.2.7) is only observed over a limited range of k -values, in which case there is also a limited range of length scales over which the curve possesses fractal scaling.

6.3 Analytical results

6.3.1 Fourier transform

In this section we will investigate some properties of the Huygens-Fresnel integral, which we consider to define the eigenvalue equation for the lowest-loss mode as follows

$$\gamma_0 u(x) = \sqrt{\frac{iA}{2\pi M}} \int_{-a}^{+a} dx' e^{-\frac{iA}{2}(x' - \frac{x}{M})^2} u(x') \quad (6.3.1)$$

where we have introduced the quantity $A = \frac{4\pi M^2 N}{(M^2 - 1)a^2}$ for ease of notation. We want to study the fractal behavior of this lowest loss mode by considering its power spectrum. Note that the eigenfunction $u(x)$ in Eq.(6.3.1) is defined for $|x| \leq a$ only. Therefore, one has some freedom in extending the definition of $u(x)$ *outside* of the principal interval $[-a, a]$, as long as Eq.(6.3.1) is obeyed for $|x| \leq a$. In this section, it will be advantageous to consider the wavenumber k a continuous variable. In order to do so, we define $u(x)$ on the entire real axis by extending Eq.(6.3.1) to arbitrary x (a different extension will be considered in section 6.6).

Using the following conventions for the Fourier transformation

$$\bar{u}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx u(x) e^{ikx}, \quad (6.3.2)$$

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dk \bar{u}(k) e^{-ikx}, \quad (6.3.3)$$

which fixes the Dirac delta function to be

$$\delta(x-x') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{ik(x-x')}, \quad (6.3.4)$$

we obtain upon Fourier transforming Eq.(6.3.1)

$$\begin{aligned} \bar{u}(k) &= \frac{1}{\gamma_0(2\pi)^{3/2}} \sqrt{\frac{iA}{M}} \int_{-a}^{+a} dx' \int_{-\infty}^{+\infty} dk_1 \otimes \\ &\otimes \int_{-\infty}^{+\infty} dx \bar{u}(k_1) e^{-\frac{iA}{2}(x' - \frac{x}{M})^2 - ikx + ik_1 x'}. \end{aligned} \quad (6.3.5)$$

We can evaluate the Gaussian integrals to arrive at the eigenvalue equation in k -space

$$\begin{aligned} \bar{u}(k) &= e^{\frac{ik^2 M^2}{2A}} \frac{\sqrt{M}}{\gamma_0} \int_{-\infty}^{+\infty} dk_1 \otimes \\ &\otimes \left[\frac{1 \sin(a(k_1 - kM))}{k_1 - kM} \right] \bar{u}(k_1) \end{aligned} \quad (6.3.6)$$

Since the eigenvalue equation is invariant under spatial reflection $x \rightarrow -x$, the eigenmodes have to be either symmetric or antisymmetric. Numerically the lowest-loss mode $u(x)$ is found to be a symmetric one, *i.e.* it

obeys $u(x) = u(-x)$. If we now use the identity

$$\begin{aligned} \frac{\sin(ax)}{\pi x} &= \frac{1}{2\pi} \int_{-a}^{+a} dq e^{iqx} \\ &= \delta(x) - \frac{1}{\pi} \int_{+a}^{+\infty} dq \cos(qx), \end{aligned} \quad (6.3.7)$$

we obtain for the eigenvalue equation the following, *exact* expression

$$\begin{aligned} \bar{u}(k) &= e^{\frac{ik^2 M^2}{2A}} \frac{\sqrt{M}}{\gamma_0} \left[\bar{u}(kM) - \right. \\ &\quad \left. - \sqrt{\frac{2}{\pi}} \int_{+a}^{+\infty} dx' u(x') \cos(kMx') \right]. \end{aligned} \quad (6.3.8)$$

Eq.(6.3.8) has several interesting features. As it stands, it suggests that the entire action of the unstable cavity consists in fact of two processes, the first being simple magnification, represented by the term proportional to $\bar{u}(kM)$. The other main feature is that the information contained in the part of the field which is projected outside of the principal interval $[-a, a]$ is lost, represented by the second term (note also the correct minus sign), which, when we rework it slightly is easily seen to be the part of the Fourier transform outside the main interval but in terms of the magnified k 's

$$\begin{aligned} \sqrt{\frac{2}{\pi}} \int_{+a}^{+\infty} dx' u(x') \cos(kMx') &= \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R} \setminus [-a, a]} dx' u(x') e^{ikMx'} \end{aligned} \quad (6.3.9)$$

Since $\bar{u}(kM)$ itself is the entire transform, the first term minus the correction term effectively leaves the 'inside' or restricted Fourier transform. Another way of writing Eq.(6.3.8) is therefore

$$\bar{u}(k) = \frac{\sqrt{M}}{\gamma_0 \sqrt{2\pi}} e^{\frac{ik^2 M^2}{2A}} \int_{-a}^{+a} dx' u(x') e^{ikMx'} \quad (6.3.10)$$

In this formulation, the eigenfunction outside of the principal interval still contains some aspects of the fractal structure, and it is therefore instructive to look at it a little more closely. Fig. 6.2 shows the eigenfunction for $M = 2, N = 100$ and $a = 1$ on $[-4, 4]$. Seen clearly is that the function

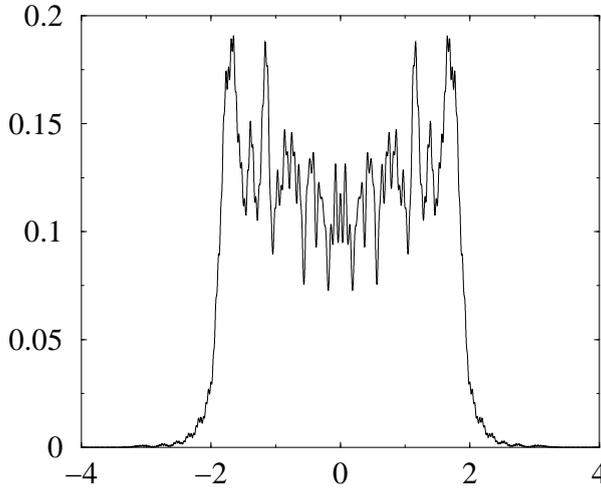


Figure 6.2: Shape of the lowest-loss mode for $M = 2, N = 10, a = 1$. Note that $u(x)$ is nonzero even outside of $[-2, 2]$.

spreads out wider than it would under magnification only (we would expect the function to be confined to $[-2, 2]$, which is caused by diffraction. Also plotted in Fig. 6.3 are several magnified images of the eigenfunction, in a manner inspired by [117], to demonstrate the presence of smaller copies of the eigenfunction in its own interior. When we hit the smallest detail scale however, this procedure breaks down as expected.

The interpretation of Eq. (6.3.8) in terms of splitting the field in terms of parts retained and discarded is reminiscent of the Baker's Map or variants known in the literature as the Arnol'd Cat Map or the Smale Horseshoe. Such one-dimensional maps consist of stretching and folding of an interval, and their asymptotic states are generically fractals. The stretching is analogous to the magnifying action of the cavity, while the folding could be associated with diffraction at the hard edges of the cavity, which in a sense folds the principal interval back onto itself.

6.3.2 Approximate evaluation

If we now crudely approximate $u(x)$ outside of the interval by its value on the boundary, $u(a)$ (since there cannot be any physical information stored in $u(x)$ for $|x| > a$, it can be argued that the actual value one takes to

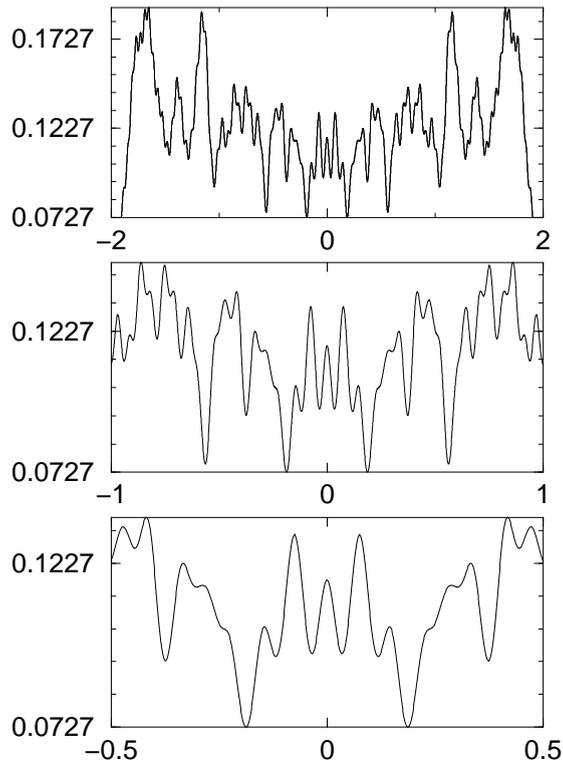


Figure 6.3: Magnified versions of Fig. 6.2. Upper panel : $\times 2$. Middle panel : $\times 4$. Lower panel : $\times 8$. Figures are scaled such that maximal value within displayed interval is at top of graph. Note the approximate constancy of the location of the maximum.

approximate it is immaterial, as long as it is away from zero) in Eq.(6.3.8), we obtain the approximate identity

$$\bar{u}(k) \approx e^{\frac{ik^2M^2}{2A}} \frac{\sqrt{M}}{\gamma_0} \left[\bar{u}(kM) + \sqrt{\frac{2}{\pi}} \frac{\sin(akM)}{kM} u(a) \right], \quad (6.3.11)$$

where we have used the asymptotic result that

$$\int_{-a}^{+\infty} dx' \cos(kx') \xrightarrow{\text{large } k} \frac{\sin ka}{k}. \quad (6.3.12)$$

We now write $\bar{u}(k)$ as

$$\bar{u}(k) = \frac{f(k)}{|k|}, \quad (6.3.13)$$

and will show that the $f(k)$'s remain finite for some range of k -values, and that they have (pseudo-)random phases. This then implies that the power spectrum $\mathcal{P}(k)$ asymptotically scales like $1/k^2$, corresponding to a fractal with fractal dimension $D_f = 1.5$ according to Eqs.(6.2.7) and (6.2.8). Upon substituting Eq.(6.3.13) into Eq.(6.3.11), we get

$$f(k) = C_1(k^2M^2) \left[f(kM) + C_2 \sin(akM) \right], \quad (6.3.14)$$

where the abbreviations

$$C_1(k^2M^2) = \frac{1}{\gamma_0 \sqrt{M}} e^{\frac{ik^2M^2}{2A}}, \quad C_2 = \sqrt{\frac{2}{\pi}} u(a), \quad (6.3.15)$$

have been introduced for brevity. Since Eq.(6.3.14) relates $f(k)$ to $f(kM)$, it constitutes a recursive definition of $f(k)$, which we can iterate to obtain

$$f(k) = C_2 \sum_{n=1}^{+\infty} \left(\prod_{m=1}^n C_1(k^2M^{2m}) \right) \sin(akM^n). \quad (6.3.16)$$

The product over the Gaussian factors $C_1(k^2M^2)$ can be worked out to give

$$\prod_{m=1}^n C_1(k^2M^{2m}) = \frac{1}{\gamma_0^n M^{n/2}} e^{\frac{ik^2M^2(1-M^{2n})}{2A(1-M^2)}}, \quad (6.3.17)$$

all of which leaves us with the following expression for $f(k)$

$$f(k) = \sqrt{\frac{2}{\pi}} u(a) \sum_{n=1}^{+\infty} \frac{1}{\gamma_0^n M^{n/2}} e^{\frac{ik^2M^2(1-M^{2n})}{2A(1-M^2)}} \sin(akM^n). \quad (6.3.18)$$

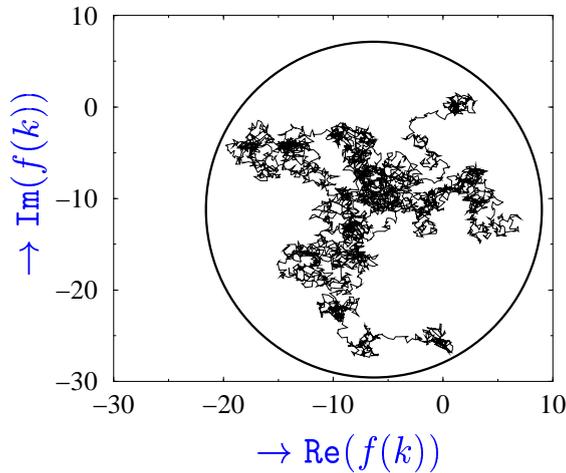


Figure 6.4: $\text{Re}(f)$ vs. $\text{Im}(f)$ for k -values up to 10^6 . At first sight this curve may look like a random walk, but for a random walk the root mean square radius would grow with the square root of the number of steps, while here $|f(k)|$ remains bounded. The fact that $|f(k)|$ remains finite proves the validity of Ansatz (6.3.13).

The sum clearly converges due to the factors $M^{-n/2}$, fulfilling one of the conditions mentioned below Eq.(6.3.13). Moreover, because of the rapid variations of the phases of the exponential factors and of the signs and values of the sine term, the phase of $f(k)$ will be a rapidly varying and pseudo-random quantity. This is verified numerically in Fig. 6.4, where successive values of $f(k)$ are plotted in the complex plane. The fact that $|f(k)|$ remains bounded, does not shrink to zero and has a pseudo-random phase shows with Eq.(6.3.11) that the power spectrum will fall off as k^{-2} , and hence that the fractal dimension is $3/2$.

Upon closer inspection of Fig. 6.4, we see that the successive points of $f(k)$ are not totally random, but instead appear to be clustered together. This reflects the fact that there is actually additional structure in the spectrum. As New *et al.* [114] have argued, this additional structure can be associated with the various virtual images. We will discuss these features in more detail below.

This is a good point to compare with the monitor-outside-a-monitor effect embodied by Eq.(6.1.1). For the cavity, the stretching of the interval

through the magnifying action of the cavity in essence does nothing other than blowing up the central part of the eigenfunction and projecting it back onto the interval. Also seen here is that besides that, there is another effect at play here which is the folding. This of course, is nothing other than diffraction at the hard edges of the cavity, the most important feature of this system *not* captured by classical ray-optics.

The approximate expression Eq.(6.3.11), or the equivalent iterated form (6.3.18), is actually the Fourier transform analogue of Courtial and Padgett's expression Eq.(6.1.1) which they introduce to study the monitor-outside-a-monitor effect: Eq.(6.3.18) gives the Fourier transform $\bar{u}(k)$ as a sum over terms with wavenumbers kM , kM^2 , and so on. These terms can be associated with length scales decreasing as a power of M .

6.4 Numerical results

To verify the predictions for the dimension of the eigenmodes, we have also studied them numerically. The actual (lowest loss) eigenmode was obtained by the Fox and Li method (see e.g. [115]), which consists of iterating Eq.(6.2.1) in real space, and boosting the norm of the function back to 1 after each iteration. This way, one ends up with the eigenmode corresponding to the largest absolute eigenvalue, and the asymptotic boost factor is the inverse of the corresponding eigenvalue. This method is in general accurate and quick, provided one takes enough gridpoints to resolve the smallest scales. The convergence to the eigenvalue is shown in Fig. 6.5 for a typical run. Some of the actual eigenmodes are shown in Fig. 6.6, which clearly shows an increase in the amount of structure with N , as expected: in optics, the size of the smallest structure is set basically by $1/N$. The eigenmodes were also obtained for moderate values of A by diagonalizing the k -space matrix V defined below. Since larger values of A greatly increase the number of Fourier modes necessary to get good resolution, the diagonalization method slows down much faster for larger A than the real space iteration method. Wherever a comparison was possible, both methods agreed excellently.

To obtain the fractal dimension, we plot the power spectrum $|u(k)|^2$ vs k in Fig. 6.7. This clearly shows that the power spectrum does indeed fall off as k^{-2} —corresponding to a fractal dimension $D_f = 3/2$ —but that there is also a lot of additional fine structure. As discussed by New *et al.*, the structure in the spectrum can be associated with successive images of

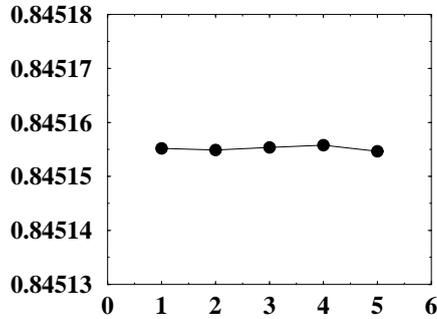


Figure 6.5: Convergence of the largest eigenvalue γ_0 . Shown are the last 5 iterations for the system specified in Fig. 6.7. Note that in this case $M = 1.4$, and γ_0 converges to the value $1/\sqrt{M} = 0.845154$.

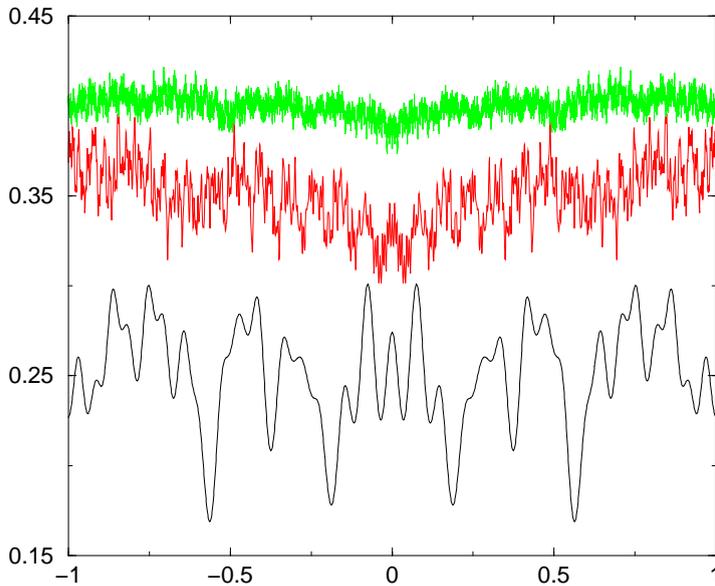


Figure 6.6: Shape of the eigenmodes for $N = 10$ (lower), $N = 100$ (middle) and $N = 1000$ (upper). The magnification M is equal to 2 for all graphs. Note the increasing detail with increasing N , and the fact that the modes get increasingly localized for larger values of N

the virtual sources.

The data shown in Fig. 6.7 were taken after ten iterations, with magnification $M = 1.4$ and the largest value of N we have studied, which is $N = 5000$. Data for other values of M but smaller Fresnel number N show the same structure. In general, after the first few iterations the structure visible in Fig. 6.7 for values $\log k \gtrsim 3.2$ is present, and upon each successive iteration one additional peak for smaller values of k appears.

6.5 Magnitude of the largest eigenvalue

In this section, we will derive the magnitude of the largest eigenvalue in the large- A limit. An equivalent definition of this largest eigenvalue which will be useful here is to take it to be the *ratio* between subsequent iterates

$$|\gamma_0|^2 = \frac{|u_{n+1}(x)|^2}{|u_n(x)|^2}. \quad (6.5.1)$$

We can work this out, to yield

$$\begin{aligned} |\gamma_0|^2 = & \frac{2MN}{(M^2 - 1)a^2} \int_{-a}^{+a} dx' \int_{-a}^{+a} dx'' e^{-\frac{iA}{2}(x'^2 - x''^2) - \frac{iAx}{M}(x' - x'')} \otimes \\ & \otimes \frac{u_n(x') u_n^*(x'')}{u_n(x) u_n^*(x)}. \end{aligned} \quad (6.5.2)$$

The crucial observation now is the following. As is apparent from Fig. 6.6, as we increase the equivalent Fresnel number N (and therefore our parameter A), the modes become more and more localized in the intensity direction. That is, the intensity profiles have an extremely well-defined average and fluctuate only mildly about that value. We can therefore, in the limit of large A , safely replace the intensity *ratios* appearing in Eq. (6.5.2) by 1. Doing so leaves us with the following expression for the eigenvalue

$$|\gamma_0|^2 = \frac{A}{2\pi M} \mathcal{I}(x) \mathcal{I}^*(x), \quad (6.5.3)$$

with

$$\mathcal{I}(x) = e^{\frac{2iAx_0^2}{M^2}} \int_{-a}^{+a} dy \exp \left\{ - \left(\sqrt{\frac{iA}{2}} \left(y - \frac{2x}{M} \right) \right)^2 \right\}. \quad (6.5.4)$$

Upon a change-of-variables to $z = \sqrt{\frac{iA}{2}}(y - \frac{2x}{M})$ the integral becomes

$$\mathcal{I}(x) = e^{\frac{2iAx^2}{M^2}} \int_{-\sqrt{\frac{iA}{2}}(a+\frac{2x}{M})}^{\sqrt{\frac{iA}{2}}(a-\frac{2x}{M})} dz e^{-z^2}. \quad (6.5.5)$$

We can now take the large- A limit of Eq. (6.5.3), using the fact that the integral appearing in Eq. (6.5.5) tends to $\sqrt{\pi}$ in that limit. Combined, this argument then yields that

$$\lim_{A \rightarrow \infty} |\gamma_0| = \frac{1}{\sqrt{M}}, \quad (6.5.6)$$

as is indeed seen from the numerics, for instance in Fig. 6.5.

6.6 k -Space matrix for an even state

Since the equation for the intensity profile is linear, the map that connects the mode profile after one round trip in the cavity to the previous one can also be written down explicitly in k -space. In this section, we briefly illustrate the structure of the matrix that effectuates the map in k -space.

In order to be able to work with discrete k -values, we use the freedom we have to define $u(x)$ outside the interval $[-a, a]$ discussed in section 6.3. If we set $a = 1$ for definiteness, and periodically extend $u(x)$ outside $[-1, 1]$ (for the symmetric modes), we can describe the profile using its discrete Fourier transform with wavenumbers $k = n\pi$, $n = 0, 1, 2, \dots$. Labeling the Fourier amplitudes with n by writing them as \bar{u}_n we thus have

$$u(x) = \bar{u}_0 + 2 \sum_{n=1}^{+\infty} \bar{u}_n \cos n\pi x, \quad (6.6.1)$$

and the Huygens-Fresnel map assumes the form

$$\bar{u}_n^{i+1} = \sum_{m=-\infty}^{+\infty} \langle n | \mathbb{T} | m \rangle \bar{u}_m^i. \quad (6.6.2)$$

If we now define a variable \bar{w} as

$$\bar{w}_{n>0} = \bar{u}_{n>0} \quad \text{and} \quad \bar{w}_0 = \frac{\bar{u}_0}{2}, \quad (6.6.3)$$

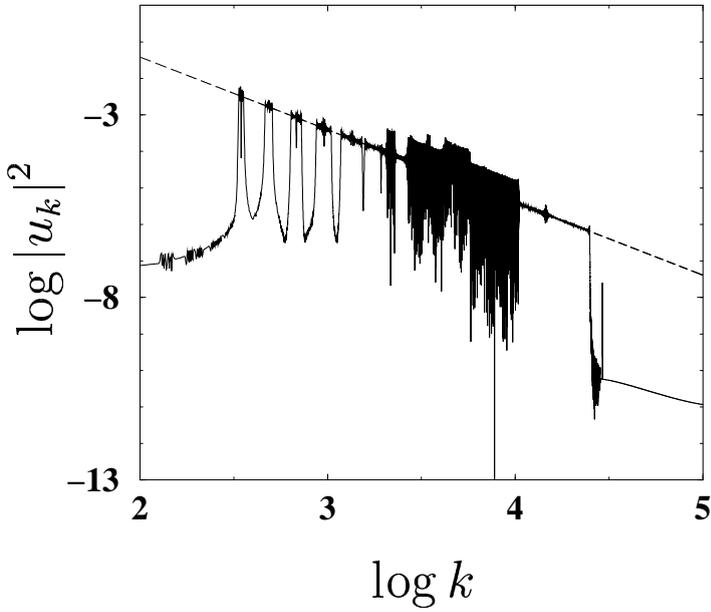


Figure 6.7: Power spectrum of $u(x)$ for tenth iterate for $N = 5000$ and $M = 1.4$, calculated on a spatial grid consisting of $5 \cdot 10^5$ points. The integral is evaluated using the method of Fox and Li, and the dashed line drawn in has slope -2 .

We can write the map as

$$\begin{cases} \bar{w}_{n>0}^{i+1} = \sum_{m=0}^{+\infty} \langle n|\mathbf{V}|m\rangle \bar{w}_m^i \\ \bar{w}_0^{i+1} = \frac{1}{2} \sum_{m=0}^{+\infty} \langle 0|\mathbf{V}|m\rangle \bar{w}_0^i \end{cases} \quad (6.6.4)$$

where $\langle n|\mathbf{V}|m\rangle = \langle n|\mathbf{T}|m\rangle + \langle n|\mathbf{T}|-m\rangle$. The matrix elements $\langle n|\mathbf{V}|m\rangle$ can be calculated exactly, this is done in Appendix 6.A. From the expressions obtained there, one sees that for the structure of the matrix, relevant values for m and n are those where the coefficients z_{++} etcetera change sign, which is at

$$\begin{aligned} m_{\pm} &= \frac{A}{\pi} \left(1 \pm \frac{1}{M} \right) \\ n_{\pm} &= \frac{A}{M\pi} \left(1 \pm \frac{1}{M} \right) \end{aligned} \quad (6.6.5)$$

These four special values of m and n divide the matrix \mathbf{V} into nine different parts, each with their own characteristic asymptotics. Fig. 6.8 shows the different regions, while Fig. 6.9 shows the general structure of this matrix. In Fig. 6.9, the matrix elements which are large in absolute value show up in white. The bright line in the center of the matrix indicates that for some range of wavenumbers, the map is dominated by the stretching effect, reflected by the fact that matrix elements with $n \approx Mm$ are large. Furthermore, the plot confirms that the values m_{\pm} and n_{\pm} are important cross-over values, in agreement with the structure of the matrix shown in Fig. 6.8 that one determines from an asymptotic analysis. Moreover, from a detailed investigation of the matrix structure one can self-consistently show that the dominant eigenvalue will have a $\bar{u}_n \sim 1/n$ scaling, in agreement with the results discussed in section 6.3.

6.7 Conclusion

We have introduced an analytical approximation that shows that the lowest loss mode of an unstable cavity laser is a fractal with dimension $3/2$. Numerical results for the spectrum are fully consistent with this conclusion. Our results also suggest very strongly that, when the intensity profile

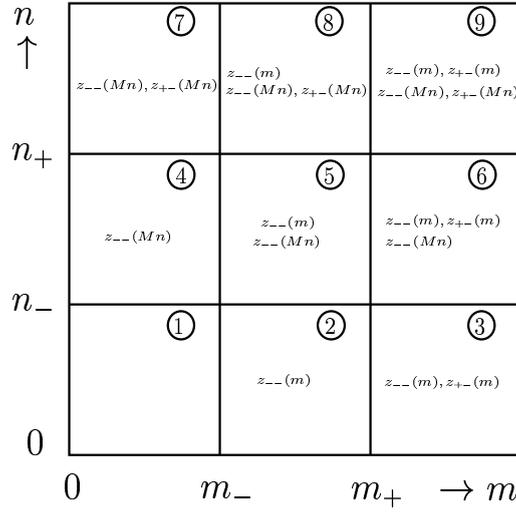


Figure 6.8: General structure of the first quadrant of the matrix V . In each of the 9 regions, the negative z 's are marked.

is traced along an arbitrary line in more than one dimension, the mode amplitude as a function of position will show fractal scaling, with *the same* fractal dimension $D_f = 3/2$. Apart from the overall fractal structure, much more detailed structure is present in these eigenmodes, as was also found by New *et al.* Recently, we have gained some new insights into the behavior of this system that seems to suggest that the fractal behavior in its present form disappears when the magnification M approaches 1 from above. We are still working towards a full understanding of the behavior of this intriguing system, in particular the origin and role of the fine structure in the spectrum.

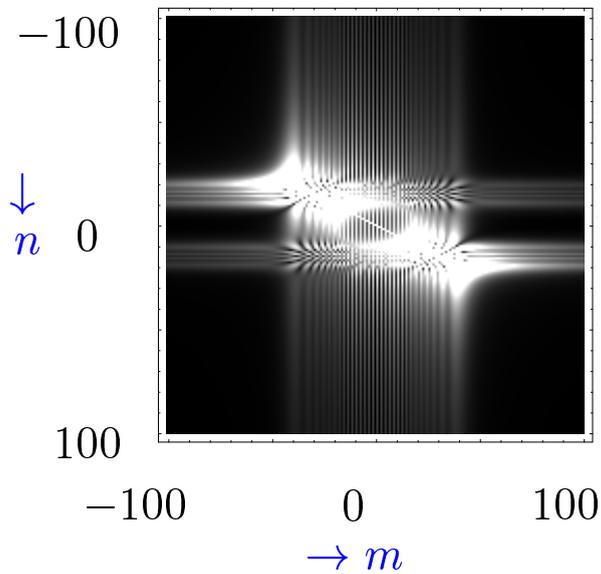


Figure 6.9: General structure of the k -space matrix for $M = 2$ and $N = 100$. Grey scaling is such that large absolute values of the matrix elements show up white.

6.A Exact expressions for the k -space matrix elements

When we set

$$D(z) = \int_0^z dt e^{i\pi \frac{t^2}{2}}, \quad (6.A.1)$$

and define

$$\begin{aligned} z_{++}(m) &= \sqrt{\frac{A}{\pi}} \left(1 + \frac{1}{M} + \frac{\pi m}{A}\right), \\ z_{-+}(m) &= \sqrt{\frac{A}{\pi}} \left(1 - \frac{1}{M} + \frac{\pi m}{A}\right), \\ z_{+-}(m) &= \sqrt{\frac{A}{\pi}} \left(1 + \frac{1}{M} - \frac{\pi m}{A}\right), \\ z_{--}(m) &= \sqrt{\frac{A}{\pi}} \left(1 - \frac{1}{M} - \frac{\pi m}{A}\right) \end{aligned} \quad (6.A.2)$$

the matrix elements are explicitly given by

$$\begin{aligned} \langle n|\mathbf{V}|m\rangle &= \sqrt{\frac{-iM}{2}} \frac{1}{\pi(m^2 - M^2 n^2)} \left\{ (-1)^n m e^{\frac{i\pi^2 m^2}{2A}} \otimes \right. \\ &\otimes \left(\cos\left(\frac{\pi m}{M}\right) [D_{-+}(m) + D_{+-}(m) - D_{--}(m) - D_{++}(m)] + \right. \\ &+ \left. i \sin\left(\frac{\pi m}{M}\right) [D_{++}(m) + D_{-+}(m) + D_{+-}(m) + D_{--}(m)] \right) + \\ &+ (-1)^m F e^{\frac{-i\pi^2 F^2}{2A}} \otimes \\ &\otimes \left(\cos(F\pi) [D_{--}(F) - D_{+-}(F) + D_{++}(F) - D_{-+}(F)] + \right. \\ &+ \left. i \sin(F\pi) [D_{--}(F) - D_{+-}(F) - D_{++}(F) + D_{-+}(F)] \right) \left. \right\} \quad (6.A.3) \end{aligned}$$

Here $F = Mn$. This is the expression used to calculate the matrix in for instance Fig. 6.9.

Bibliography

- [1] A. C. Newell, T. Passot, J. Lega, *Ann. Rev. Fluid Mech.* **25**, 399 (1993).
- [2] Count Rumford, '*Of the propagation of heat in fluids*', Complete works **1**, 239, American Academy of Arts and Sciences, Boston (1870).
- [3] H. Bénard, '*Les tourbillons cellulaires dans un nappe liquide*', *Revue générale des Sciences pures et appliquées* **11**, 1261-1271 and 1309-1328 (1900).
- [4] S. Chandrasekhar, *Hydrodynamic and Hydromagnetic instability* (Dover publications, inc., New York 1981).
- [5] J. Thomson, '*On a changing tessellated structure in certain liquids*', *Proc.Phil.Soc.Glasgow* **13**, 464-468 (1882).
- [6] H. Bénard, '*Les tourbillons cellulaires dans un nappe liquide transportant de la chaleur par convection en régime permanent*', *Annales de Chimie et de Physique* **23**, 62-144 (1901).
- [7] Lord Rayleigh, '*On convective currents in a horizontal layer of fluid when the higher temperature is on the under side*', *Phil.Mag.* **32**, 529-546 (1916); also *Scientific Papers* **6**, 432-446, Cambridge (England, 1920).
- [8] R. J. Schmidt and S. W. Milverton, '*on the instability of a fluid when heated from below*', *Proc. Roy. Soc. (London) A* **152**, 586-594 (1935).
- [9] P. L. Silveston, '*Wärmedurchgang in waagerechten Flüssigkeitsschichten*', Part 1, *Forsch.Ing.Wes.* **24**, 29-32 and 59-69 (1958).

-
- [10] J. Liu and G. Ahlers, Phys. Rev. Lett **77** 3126 (1996).
- [11] R. V. Cakmur, D. A. Egolf, B. B. Plapp and E. Bodenschatz, Phys. Rev. Lett. **79** 1853 (1997).
- [12] S. Jeanjean, Tech.rep., Cornell University (1997).
- [13] P. Newell, *Biology of the Chemotactic Response*, J. M. Lackie ed. (Cambridge: Cambridge University Press 1986).
- [14] H. Jeffreys, *Cartesian Tensors*, Cambridge University Press, Cambridge (England, 1931).
- [15] L. D. Landau and E.M. Lifschitz, *Fluid Mechanics* (Pergamon, New York, 1978).
- [16] J.B.van den Berg, Ph.D. Thesis Leiden (2000) and references therein.
- [17] A. C. Newell and J. A. Whitehead, J. Fluid Mech. **38** 279 (1969).
- [18] W. Eckhaus, *Nonlinear stability theory* (Springer, Berlin 1965).
- [19] R. Montagne, E. Hernández-García and M. San Miguel, Physica D **96**, 47 (1996).
- [20] P. M. Chaikin and T. C. Lubensky, *Principles of condensed matter physics* (Chap 9), Cambridge University Press (1995).
- [21] W. van Saarloos and P. C. Hohenberg, Physica D **56**, 303 (1992); **69**, 209 (1993) [Errata].
- [22] M. van Hecke, Phys. Rev. Lett. **80**, 1896 (1998).
- [23] J. M. Vince, Ph.D. Thesis, Saclay.
- [24] M. C. Cross and P. C. Hohenberg, Rev. Mod. Phys. **65** 851 (1993).
- [25] For an elementary introduction to the amplitude equation approach, see, e.g., M. van Hecke, P. C. Hohenberg and W. van Saarloos, in *Fundamental Problems in Statistical Mechanics VIII*, H. van Beijeren and M. H. Ernst, eds. (North-Holland, Amsterdam, 1994).
- [26] D. Walgraef, *Spatio-temporal pattern formation* (Springer, Berlin, 1996).

- [27] For an example in the context of binary mixtures, see H. Riecke, Phys. Rev. Lett. **68**, 301 (1992).
- [28] For an example in the context of chemical waves, see M. Ipsen, 'Amplitude equations and normal forms', PhD Thesis, University of Copenhagen (unpublished); M. Ipsen, F. Hynne and P. G. Sørensen, "Amplitude equation for reaction-diffusion systems with a Hopf bifurcation and a slow real mode", preprint (1998).
- [29] In the context of convection in liquid crystals, see M. Dennin, M. Treiber, L. Kramer, G. Ahlers and D. S. Cannell, Phys. Rev. Lett. **76**, 319 (1995).
- [30] In the context of Marangoni convection, see A. A. Golovin, A. A. Nepomnyashchy, L. M. Pismen and H. Riecke, Physica D **106**, 131 (1997).
- [31] D. Bensimon, A. Pumir and B. I. Shraiman, J. Physique France **50**, 3089 (1989).
- [32] W. Schöpf and W. Zimmerman, Phys. Rev. E **47**, 1739 (1993).
- [33] E. Y. Kuo and M. C. Cross, Phys. Rev. E **47**, R2245 (1993).
- [34] M. van Hecke and W. van Saarloos, Phys. Rev. E **55**, R1259 (1997).
- [35] J. Herrmann and F. H. Busse, J. Fluid Mech. **255**, 183 (1993).
- [36] M. C. Cross, Phys. Rev. A **38**, 3593 (1988).
- [37] M. C. Cr
- [38] E. Knobloch and J. de Luca, Nonlinearity **3**, 975 (1990); E. Knobloch, 'Nonlocal amplitude equations', in 'Pattern Formation in Complex Dissipative Systems', S. Kai (editor), World Scientific, (1992).
- [39] C. Martel and J. M. Vega, Nonlinearity **9**, 1129 (1996); J. M. Vega, SIAM J. Math. Anal. **24**, 603 (1993); C. Martel and J. M. Vega, Nonlinearity **11**, 105 (1998).
- [40] R. Alvarez, M. van Hecke and W. van Saarloos, Phys. Rev. E **56**, R1306 (1997).

-
- [41] P. Kolodner, D. Bensimon and C. M. Surko, *Phys. Rev. Lett* **60**, 1723 (1988).
- [42] N. Bekki and K. Nozaki, *Phys. Lett. A* **110**, 133 (1985).
- [43] R. Conte and M. Musette, *Physica D* **69** 1 (1993); P. Marcq, H. Chaté and R. Conte, *Physica D* **73**, 305 (1994).
- [44] S. Popp, O. Stiller, I. Aranson, A. Weber and L. Kramer, *Phys. Rev. Lett.* **70**, 3880 (1993).
- [45] A. Doelman, *Physica* **97**, 398 (1996).
- [46] O. Thual and S. Fauve, *J. Phys. France* **49**, 1829 (1988).
- [47] S. Fauve and O. Thual, *Phys. Rev. Lett.* **64**, 282 (1990).
- [48] V. Hakim, P. Jakobsen and O. Pomeau, *Europhys. Lett.* **11** 19 (1990); *Eur. J. Mech. B* **10**, 137 (1991).
- [49] C. Elphick and E. Meron, *Phys. Rev. A* **40**, 3226 (1989); *Phys. Rev. Lett.* **65**, 2476 (1990).
- [50] P. Couillet, T. Frisch and T. Plaza, *Physica D* **59**, 75 (1993).
- [51] J. M. Vince and M. Dubois, *Europhys. Lett.* **20**, 505 (1992).
- [52] M.-T. Westra, Ph.D. Thesis, Eindhoven (2001).
- [53] L. Kramer, E. Ben-Jacob and H. Brand, *Phys. Rev. Lett.* **49**, 1891 (1982); G. Dee and J. S. Langer, *Phys. Rev. Lett.* **50**, 383 (1983); E. Ben-Jacob, H. Brand, G. Dee, L. Kramer and J. S. Langer, *Physica D* **14**, 348 (1985), W. van Saarloos, *Phys. Rev. A* **37**, 211 (1988).
- [54] W. van Saarloos, *Phys. Rev. A* **39**, 6367 (1989).
- [55] A. N. Stokes, *Math. Biosci.* **31**, 307 (1976).
- [56] G. C. Paquette, L.-Y. Chen, N. Goldenfeld and Y. Oono, *Phys. Rev. Lett.* **72**, 76 (1994).
- [57] U. Ebert and W. van Saarloos, *Phys. Rev. Lett.* **80**, 1650 (1998).
- [58] U. Ebert and W. van Saarloos, *Physica D* **146**, 1 (2000).

- [59] M. van Hecke and B. A. Malomed, *Physica D* **101**, 131 (1997).
- [60] T. Kapitula and B. Sandstede, *Physica D* **124**, 58-103 (1998); T. Kapitula and B. Sandstede, *J. Opt. Soc. of America B* **15**, 2757-2762 (1998); J. Yang and D. J. Kaup, *Solitary waves in perturbed generalized nonlinear Schroedinger equations*, preprint, Dec. 1997.
- [61] M. van Hecke, E. de Wit and W. van Saarloos, *Phys. Rev. Lett.* **75**, 3830 (1995).
- [62] R. J. Deissler and H. R. Brand, *Phys. Rev. Lett.* **72**, 478 (1994); **74**, 4847 (1995).
- [63] E. de Wit, Masters Thesis, Leiden (1995).
- [64] A. Couairon and J.-M. Chomaz, *Primary and secondary nonlinear global instability* (preprint 1999).
- [65] P. Kolodner, *Phys. Rev. A*, **46**, 6452 (1992).
- [66] D. Bensimon, B. I. Shraiman and V. Croquette, *Phys. Rev. A* **38**, 5461 (1988).
- [67] B. A. Malomed, *Phys. Rev. E* **50**, R3310 (1994).
- [68] L. Gil, *Phys. Rev. Lett.* **70**, 162 (1993); M. San Miguel, *Phys. Rev. Lett.* **75**, 425 (1995).
- [69] P. Coullet, C. Elphick, L. Gil and J. Lega, *Phys. Rev. Lett.* **59**, 884 (1987).
- [70] I. Aranson and L. Tsimring, *Phys. Rev. Lett.* **75**, 3273 (1995).
- [71] N. Mukolobwicz, A. Chiffaudel and F. Daviaud, private communications.
- [72] B. I. Shraiman, A. Pumir, W. van Saarloos, P. C. Hohenberg, H. Chaté and M. Holen, *Physica D* **57**, 241 (1992);
- [73] H. Chaté, *Nonlinearity* **7**, 185 (1994).
- [74] D. A. Egolf and H. S. Greenside, *Phys. Rev. Lett.* **74**, 1751 (1995).
- [75] P. Huerre and P. A. Monkewitz, *Annu. Rev. Fluid Mech.* **22**, 473 (1990); A. Couairon and J.-M. Chomaz, *Physica D* **108**, 236 (1997).

- [76] P. Hagan, SIAM J. Appl. Math. **42**, 762 (1982); A. Weber, L. Kramer, I. S. Aranson and L. Aranson, Physica D **61**, 279 (1992); T. Bohr, G. Huber and E. Ott, Physica D **106**, 95 (1997).
- [77] I. S. Aranson, L. Aranson, L. Kramer and A. Weber, Phys. Rev. A **46**, 2992 (1992).
- [78] T. Bohr, M. H. Jensen, G. Paladin and A. Vulpiani, “*Dynamical Systems Approach to Turbulence*”, Cambridge, in press (1998).
- [79] H. Sakaguchi, Prog. Theor. Phys. **95**, 823 (1996); Physica Scripta **T67**, 148 (1996).
- [80] P. Couillet, S. Fauve and E. Tirapegui, J. Physique Lett. **46**, L-787 (1985).
- [81] H. Riecke and L. Kramer, “The stability of standing waves”, ptt-sol/9805005 (1998).
- [82] A. Amengual, E. Hernández-García, R. Montagne, and M. San Miguel, Phys. Rev. Lett. **78**, 4379 (1997).
- [83] S. Ciliberto and P. Bigazzi, Phys. Rev. Lett. **60**, 286 (1988); F. Daviaud, J. Lega, P. Bergé, P. Couillet and M. Dubois, Physica D **55**, 287 (1992).
- [84] B.A. Malomed and A.A. Nepomnyashy, Phys. Rev. A **42**, 6009 (1990).
- [85] V. Hakim and Y. Pomeau, Eur. J. Mech. B Suppl **10** 137 (1991).
- [86] M. Dubois, F. Daviaud, O. Ronsin and P. Bergé, Physica D **61**, 140 (1992).
- [87] J. M. Vince and M. Dubois, Physica D **102**, 93 (1997).
- [88] J. M. van Boxmeer, Masters Thesis, Eindhoven (1999).
- [89] P. Kolodner, Phys. Rev. E **50**, 2731 (1994).
- [90] E. Moses, J. Fineberg and V. Steinberg, Phys. Rev. A. **35**, 2757, 1987.
- [91] G. W. Baxter, K. D. Eaton and C. M. Surko, Phys. Rev. A **46**, 1735 (1992).

- [92] E. Kaplan, V. Steinberg, Phys. Rev. E, **48**, R641 (1993).
- [93] See sections V.B.c.(iii) and X.A.6.a of [24], and references therein.
- [94] A. B. Rovinsky, A. M. Zhabotinsky, and I. R. Epstein, Phys. Rev. E **56**, 2412 (1997).
- [95] A. Garcimartín, N. Mukolobwicz and F. Daviaud, Phys. Rev. E **56**, 1699 (1997).
- [96] N. Mukolobwicz, thesis, Université Paris XI (1998) [unpublished].
- [97] N. Mukolobwicz, A. Chiffaudel and F. Daviaud, Phys. Rev. Lett. **80**, 4661 (1998).
- [98] E. Kaplan and V. Steinberg, Phys. Rev. A **46**, R2996 (1992)
- [99] E. Kaplan and V. Steinberg, Phys. Rev. Lett. **71**, 3291 (1993).
- [100] H. Herrero and H. Riecke, Physica D **85**, 79 (1995); see also chaosdyn 9701017.
- [101] H. Riecke and W.-J. Rappel, Phys. Rev. Lett. **76**, 4035 (1996).
- [102] G. Ahlers and D.S. Canell, Phys. Rev. Lett. **50**, 1583 (1983).
- [103] J. Fineberg and V. Steinberg, Phys. Rev. Lett. **50**, 1583 (1987).
- [104] R. Bar-Ziv and E. Moses, Phys. Rev. Lett. **73**, 1392 (1994).
- [105] R. Powers and R. E. Goldstein, Phys. Rev. Lett. **78**, 2555 (1997).
- [106] W. van Saarloos, Phys. Rev. A **39**, 6367 (1989).
- [107] G. Dee and J. S. Langer, Phys. Rev. Lett. **50**, 383 (1983).
- [108] U. Ebert, W. Spruijt and W. van Saarloos (unpublished).
- [109] P. Collet and J.-P. Eckmann, Commun. Math. Phys. **107**, 39 (1986).
- [110] K. Nozaki and N. Bekki, Phys. Rev. Lett. **51**, 271 (1983).
- [111] U. Ebert and W. van Saarloos, Phys. Rev. Lett. **80**, 1650 (1998).
- [112] G.P. Karman and J.P. Woerdman, Optics Letters **23** 1909-1911 (1998).

- [113] G.P. Karman, G.S. McDonald, G.H.C. New, and J.P. Woerdman, *Nature (London)* **402**, 138 (1999).
- [114] G.H.C. New, M.A. Yates, J.P. Woerdman, and G.S. McDonald, unpublished.
- [115] A.E. Siegman, in *Lasers* (University Science Books, Mill Valley, CA, 1986).
- [116] M.V. Berry, *J. Phys. A* **29** 6617-6629 (1996).
- [117] J. Courtial and M.J. Padgett, *Phys. Rev. Lett.* **85**, 5320.

Samenvatting

Natuurkunde is zo eenvoudig nog niet. Verschijnselen, die op het eerste gezicht kinderlijk eenvoudig aandoen blijken, bij nadere bestudering, niet zelden een verbluffende veelheid aan mechanismen in zich te verenigen. Het gedrag van dergelijke systemen kan dan ook alleen goed begrepen worden wanneer men om te beginnen die mechanismen kent, maar wellicht nog belangrijker ook de wijze waarop al die verschillende aspecten elkaar *beïnvloeden*. Een systeem, waarvan men het gedrag als geheel opgebouwd kan denken uit vele gekoppelde processen is dus allerm minst éénvoudig. Deze wisselwerkende samengesteldheid wordt in de natuurkunde wel aangegeven met het woord *complexiteit*.

De natuurkunde probeert de processen die ervoor zorgen dat de wereld om ons heen zich gedraagt zoals zij doet te identificeren en te beschrijven. In zekere zin beschouwt zij daartoe die wereld, of beter gezegd de werkelijkheid zoals wij die kunnen waarnemen (al dan niet met hulpmiddelen), als een machine die net als een computer “input” kan omzetten in “output”. De interne processen in die machine liggen vast, en kunnen vergeleken worden met het programma dat op een computer draait. Eenvoudig gesteld is de reconstructie van dat programma door het meten en interpreteren van de respons op een veelheid aan stimuli de kerntaak van de natuurkunde. De bekende natuurkundige wetten, waar we er allemaal wel een paar van kennen, zijn onderdelen van dit programma. Nu heeft de natuur echter een eigenschap die alle computers missen, en dat is de rol die in veel van haar processen gespeeld wordt door het toeval¹. Dat toeval maakt, dat een gegeven systeem niet altijd op dezelfde manier reageert,

¹Het is onmogelijk een computer zó te programmeren, dat deze een oneindige rij willekeurig gekozen getallen produceert. Omdat men de computer eerst in begrijpelijke taal moet vertellen *hoe* hij deze moet construeren, kan hij ten hoogste een reeks produceren die erg veel *lijkt* op een willekeurige.

ook al geeft men het dezelfde stimulus. Het zoeken naar de achterliggende wetmatigheden wordt er hierdoor niet eenvoudiger op. Zie daar dus het probleem waar de natuurkunde als geheel zich mee geconfronteerd ziet: het voorspellen van de output van een afgesloten machine waarin de raderen op al dan niet willekeurige wijze draaien en in elkaar grijpen.

Dat is natuurlijk een onmogelijke opgave. Nog afgezien van de fundamentele problemen die volgen uit het feit dat ook wij mensen niet *buiten* deze machine staan, maar er juist middenin zitten, is het nog maar zeer de vraag of het toeval zich in ons soort wetten laat vangen.

Toch is er hoop. Ingenieuze experimenten kunnen ons soms een kijkje verschaffen in het binnenste van de machine, waarbij men bijvoorbeeld kan denken aan de Scanning Tunneling microscopen die individuele atomen kunnen “zien”. Om in de machine-analogie te blijven kunnen we dus soms raderen waarvan het bestaan voorheen slechts geponeerd was direct waarnemen. Maar er is meer. Een van de mooie aspecten van de natuur is dat het grote zich vaak spiegelt in het kleine, en men door het bestuderen van eenvoudige modelsystemen vaak veel kan leren over meer complexe fenomenen. Het is wellicht goed om op dit punt op te merken dat het werk dat voor U ligt een theoretische studie behelst, en we onder modelsystemen dus handig gekozen wiskundige vergelijkingen zullen verstaan. We kunnen in die modellen alle effecten die niet van direct belang zijn uitschakelen, en toch door studie van een dergelijk “gestript” model iets nuttigs leren. Het feit dat vele, op het oog zeer verschillende systemen kwalitatief hetzelfde gedrag vertonen staat bekend als *universaliteit*. Het is deze universaliteit die het ons mogelijk maakt een connectie maken tussen de modellen zoals die in dit werk onderzocht worden, en experimentele resultaten. Resultaten van experimenten die overigens vaak, door slim ontwerp, óók in staat zijn systemen tot in zeer goede benadering te isoleren van storende uitwendige invloeden, en zich dus bij uitstek lenen voor vergelijking met onze modelsystemen.

Wat in dit proefschrift beschreven staat, is een onderzoek naar een aantal verschillende *niet-lineaire* systemen. In zulke systemen is de wisselwerking tussen de afzonderlijke ingrediënten, die men in dit geval wel met de term *modes* aanduidt, van dien aard dat de *sterkte* van die wisselwerking niet recht-evenredig is met de *grootte* van de wisselwerkende modes. Dit opent de deur voor allerlei niet-lineaire fenomenen, zoals bijvoorbeeld resonantie, en in sommige extreme gevallen zelfs chaos, een toestand van (al dan niet schijnbare) volledige willekeur. Wij zijn echter niet

in de eerste plaats geïnteresseerd in het gedrag van dit soort systemen onder extreme omstandigheden, maar meer in het regime waarin we voor het eerst wat beginnen te merken van het niet-lineaire karakter. In dat regime manifesteren de niet-lineariteiten zich vaak op een verrassende manier: de systemen ontwikkelen spontaan regelmatige structuren, een proces dat we wel met *spontane patroonformatie* aanduiden. In Hoofdstuk 1 geven een aantal figuren een idee van hoe die patronen eruit zien. Hoewel deze keurig geordende toestanden op het eerste gezicht weinig hebben uit te staan met de associatie die velen met het begrip chaos hebben, kan een studie naar hun gedrag toch veel verhelderen over het gedrag van niet-lineaire systemen, waaronder ook “echte” hard-chaotische.

In dit werk zijn wij vooral geïnteresseerd in de rol die een bepaald soort structuren speelt in de dynamica (oftewel hoe het systeem in de tijd evolueert) van onze modelsystemen. We noemen de structuren *coherent*, omdat ze een intrinsieke samenhang vertonen: na verloop van tijd veranderen ze niet meer van vorm, en bewegen met een constante snelheid. Het bestaan van zulke structuren in een gegeven model is niet triviaal, aangezien ze er niet “met de hand” ingestopt zijn. We moeten ze zien als een product van het subtiele samenspel tussen de verschillende modes in het systeem. Dat neemt niet weg dat ze, nadat ze eenmaal ontstaan zijn, vaak een lange levensduur hebben, en we ze dus kunnen beschouwen als een soort deeltjes in ons systeem. Deze deeltjes beïnvloeden niet alleen elkaar, maar ook de ruimte om zich heen, zoals ook een magneet de ruimte om zich heen verandert. Dit proefschrift is een poging in de eerste plaats de relevante coherente structuren voor een aantal modelsystemen te identificeren, en vervolgens hun effect op de dynamica van het systeem als geheel en op andere coherente structuren in kaart te brengen. Dat is dan ook de titel van dit werk: *“De Dynamica en Interacties van Coherente Structuren in Niet-Lineaire Systemen”*.

In hoofdstuk 1 geef ik een kort overzicht van de historie van de spontane patroonformatie aan de hand van een van de standaard-experimenten, te weten het Rayleigh-Bénard experiment. Hierin wordt een vloeistof van onder verhit, en wanneer de temperatuur van de onderkant een bepaalde drempelwaarde overschrijdt zet de vloeistof zich in beweging. Dit is eenvoudig te begrijpen als men bedenkt dat vloeistoffen (behalve water beneden 4°C) uitzetten indien zij verhit worden. Per volume-eenheid is er dus minder water en dus is heet water lichter dan koud water. Het hete water wil dus opstijgen, en aldus krijg je de vloeistof in beweging. Ver-

der laat ik een paar voorbeelden van spontane patroonformatie uit andere vakgebieden zien,

Hoofdstuk 2 gaat nader in op de wiskundige technieken die nodig zijn om de klasse van systemen waarin het Rayleigh-Bénard experiment valt zo goed mogelijk te beschrijven. In het bijzonder laat ik zien hoe we die hele klasse kunnen beschrijven met één en dezelfde vergelijking. We buiten hier het universaliteitsprincipe uit, door uit een grote verscheidenheid aan kwalitatief equivalente systemen het voor ons makkelijkste te kiezen.

In Hoofdstuk 3 behandelen we voor een bepaald systeem de coherente structuren. We leiden ruimtelijke profielen af en gebruiken een elegant tel-argument om de multipliciteit van oplossingen af te leiden. Het feit, dat de meest relevante coherente structuur *uniek* blijkt te zijn, blijkt van grote waarde als we in Hoofdstuk 4 de effecten van de eerder afgeleide coherente structuren op de dynamica bekijken, en hun onderlinge interacties nader bestuderen. Ook vergelijken we onze resultaten met recente (en minder recente) experimenten. Die vergelijking valt in een aantal opzichten verbazend gunstig uit, hoewel er zeker ook nog een paar vraagtekens blijven bestaan.

In hoofdstuk 5 behandelen we in weer andere modelsystemen zogenaamde invasieprocessen. Vaak is het in de natuur zo dat er een competitie is tussen twee toestanden, waarbij de ene uitstekend gedijt in de andere. De jaarlijkse griep epidemie is er een goed voorbeeld van. In een gezonde populatie kan het griepvirus razendsnel om zich heen grijpen, en we kunnen op de kaart de verspreiding van zo'n virus volgen. Wat we dan zullen zien is (afgezien van de nucleatie van nieuwe verspreidingsgebieden door bijvoorbeeld vliegverkeer), dat er een duidelijke scheidslijn is tussen de zieke en de gezonde populatie. Dat is een eigenschap die vele van dit soort invasieprocessen hebben, en we noemen zo'n scheidingsgebied een *front*. De invasie van de ene toestand in de andere kan goed beschreven worden door naar de dynamica van het front dat de twee toestanden scheidt te kijken. Wij doen dat in Hoofdstuk 5, en kijken in het bijzonder naar de snelheid waarmee dit gebeurt en de wijze waarop deze snelheid bereikt wordt. Dit is interessant, omdat dit voor de klasse van systemen die wij bekijken heel langzaam blijkt te gaan. Omdat deze langzame *relaxatie* aantoonbaar een gedeelde eigenschap van vele niet-lineaire systemen is, spreken we van *universele algebraïsche relaxatie*.

Hoofdstuk 6 tenslotte staat qua onderwerp los van de rest van dit proefschrift. Het bevat een theoretische verklaring voor een opvallende recente

ontdekking uit de quantum-optica groep te Leiden. Zij toonden aan de het laserlicht dat door een bepaald soort optisch apparaat uitgezonden wordt, eigenschappen vertoont die wijzen op *fractaliteit*. Fractalen zijn wiskundige figuren die men opgebouwd kan denken uit kleinere kopieën van zichzelf. Zulke figuren worden gekarakteriseerd door een gebroken dimensie. Ik geef het mechanisme aan dat de fractalen genereert, en leid bovendien een waarde van 1.5 af voor die dimensie. Dit is inderdaad wat er ook uit numerieke simulaties komt.

Publications

- M. van Hecke, C. Storm and W. van Saarloos, *Sources, sinks and wavenumber selection in coupled CGL equations and experimental implications for traveling wave systems*, Physica D **134** 1-47.
- C. Storm, W. Spruijt, U. Ebert and W. van Saarloos, *Universal algebraic relaxation of velocity and phase in pulled fronts generating periodic or chaotic states*, Phys. Rev. E **61**, R6063-R6066.
- C.Storm, M.V. Berry and W. van Saarloos, *Fractal eigenmodes in an unstable cavity laser*, in preparation.
- J. Kockelkoren, C. Storm and W. van Saarloos, *A re-examination of front propagation in Rayleigh-Bénard convection*, in preparation.

Curriculum Vitae

Ik ben op 23 mei 1973 geboren te Groenlo. Mijn middelbare schooltijd heb ik doorgebracht op het Erasmus College te Zoetermeer, waar ik in 1991 het eindexamen VWO behaalde.

Vervolgens ben ik in 1991 natuurkunde aan de Universiteit Leiden gaan studeren. In het kader van de introductieperiode in het laboratorium, een verplicht onderdeel in het curriculum, deed ik enkele maanden in de groep van prof. dr. G. Frossati onderzoek in het kader van het GRAIL project. Dit project heeft het ontwerp en de bouw van een antenne voor gravitatiestraling uit het heelal tot doel. Mijn afstudeeronderzoek in de theoretische natuurkunde verrichtte ik onder begeleiding van prof. dr. ir. W. van Saarloos en betrof analytische en numerieke onderzoeken aan de complexe Ginzburg-Landau vergelijking. Met een scriptie over dit onderwerp rondde ik in mei van het jaar 1997 mijn studie af.

In juni 1997 trad ik in dienst van de Universiteit Leiden, eerst als beurspromovendus en een jaar later in 1998 als assistent in opleiding. Onder begeleiding van prof. dr. ir. W. van Saarloos deed ik theoretisch onderzoek naar spontane patroonformatie in niet-evenwichts systemen. De resultaten van dit onderzoek zijn in dit proefschrift verzameld.

Tijdens mijn aanstelling bezocht ik zomerscholen en conferenties te Altenberg, Budapest, Leiden, Magdeburg, Atlanta en Sitges. Daarnaast heb ik mijn werk gepresenteerd in voordrachten gegeven te Bayreuth, Utrecht, Amsterdam, San Diego, Tucson en Philadelphia. Ik heb bovendien het Niels Bohr Instituut in Kopenhagen en het Commissariat à l'Énergie Atomique te Saclay mogen bezoeken. In 1999 bracht ik van april tot juni een werkbezoek aan de Universiteit van Bayreuth, waar ik met prof. dr. W. Pesch heb samengewerkt.

Tijdens de vier jaren van mijn promotie heb ik werkcolleges behorende bij de vakken Electromagnetisme II (vier jaar) en Quantumtheorie I (één

jaar) gegeven. Daarnaast was ik één van de docenten van het interdisciplinaire college “Dynamical Systems and Pattern Formation”, waaraan ik ook een tweetal hoofdstukken voor de syllabus bijgedragen heb.

Ik ben prof. van Saarloos en de stichting FOM dankbaar voor het feit dat zij mij onlangs in de mogelijkheid gesteld hebben een reis door de Verenigde Staten te maken, om aldaar diverse universiteiten te bezoeken in het kader van een vervolgaanstelling. Deze reis heb ik in december van 2000 gemaakt, en heeft er mede toe geleid dat ik vanaf september 2001 werkzaam zal zijn als postdoc aan de University of Pennsylvania te Philadelphia.