# Dynamical Systems and Nonequilibrium Pattern Formation

L.A. Peletier
W. van Saarloos
G.B. van den Berg
J. Müller
C. Storm
G. Tripathy

Instituut-Lorentz for Theoretical Physics/Mathematics Institute,

Universiteit Leiden

1999

# Preface

While there has been quite a bit of work on nonequilibrium pattern formation at the Instituut–Lorentz for the last decade, our theory group has never offered a formal course on nonlinear dynamics or on nonequilibrium pattern formation. Moreover, my mathematics colleague Peletier and I had often realized that it would be nice at some point to offer such a course together, since in his group there has, over the years, been lots of work which is closely related to topics of interest within the physics literature. When the idea arose in 1998 to organize a "studentenseminarium" or "werkgroep" for undergraduate students interested in doing their undergraduate research ("afstudeeronderzoek") at the Instituut–Lorentz, the idea quickly emerged to do the "werkgroep" on this subject, and to make it into a joint project.

The idea of such a "werkgroep" is that the participants do not learn a field by consuming passively what a teacher explains on a blackboard, but that they learn by activily solving problems and discussing them with faculty members in a setting that is much more reminiscent of an actual research environment than a standard course. E.g., we intend to discuss all or most of the solutions on the blackboard with all participants present, and to have discussions about possible other routes to solving the problems, shortcuts, etc. At each meeting, we intend to discuss one of the chapters of this syllabus.

The project of preparing this syllabus has grown into a joint effort of a diverse group of people, two professors, two postdocs, and two graduate students. Their names are listed on the next page. But the diversity does not stop there: for some of us the field was actually rather new, while others had worked in it for a long time. Moreover, during the writing process, we have become more and more aware of the differences in style, between mathematicians and physicists, of approaching a problem. You will quickly discover these differences in style, training and outlook for yourself. For, although we have done the necessary editing of each other's chapters, we have decided not to try to artificially suppress these differences. In fact, I personally view these differences as a strength, *not* a weakness of this course: they do reflect real differences in ways to approach a problem and the need a physicist feels to advance in bigger steps towards understanding a complicated physical problem. Since we have profited from discussing these and from being exposed to each other's way of thinking, I believe that you will do so as well. If you actively participate in this course, you will have an excellent starting point to write an "afstudeerscriptie" either in mathematics or in theoretical physics or, who knows, at the interface between the two.

We hope that you will have as much fun with this course as we have already had

preparing it!

Wim van Saarloos
Leiden, July 5 1999

**Authors**

1. An Introduction to Maps: Jan Bouwe van den Berg.
2. Chaotic Dynamics: Jan Bouwe van den Berg.
3. The Buckling Instability: Bert Peletier.
4. The Hopf Bifurcation: Judith Müller.
5. Reaction Diffusion Equations: Bert Peletier.
6. The Turing Instability and Pattern Formation in Chemical Systems: Judith Müller.
7. The Navier-Stokes equations of Fluid Dynamics: Wim van Saarloos.
8. The Rayleigh-Bénard Instability: Wim van Saarloos.
9. Derivation of Amplitude Equations: Weakly Nonlinear Theory: Kees Storm.
10. Implications of the Amplitude Equation Description: Analysis of Pattern Dynamics near Threshold: Kees Storm.
11. Front Formation In The Nonlinear Diffusion Equation: Goutam Tripathy.
12. Front Propagation into Unstable States: Goutam Tripathy.

For further information about the course, contact:
L. A. Peletier, phone 071-5277136, email: `peletier@wi.leidenuniv.nl`
W. van Saarloos, phone 071-5275501, email: `saarloos@lorentz.leidenuniv.nl`

# Contents

# Chapter 1

# An introduction to maps

The purpose of chapters 1 and 2 is to convey the idea that even the simplest of maps can behave extremely surprising. Without further ado, we introduce the concept of a *map* given by a difference equation:

$$x_{n+1} = f(x_n), \tag{1.0.1}$$

where $x_n \in \mathbb{R}^k$ (for example). One also says that $x$ *maps to* $f(x)$, notation: $x \mapsto f(x)$. What we mean is that when we start with some initial value $x_0$, then we iterate the map $f$ to obtain a sequence of points

$$x_0\ ,\ x_1 = f(x_0)\ ,\ x_2 = f(x_1) = f(f(x_0))\ ,\ \ldots. \tag{1.0.2}$$

The *forward orbit* of $x$ is $\gamma^+(x) = \{f^n(x) \,|\, n \geq 0\}$, where

$$f^n = \underbrace{f \circ f \circ \cdots \circ f \circ f}_{n} \tag{1.0.3}$$

is the $n^{\text{th}}$ composite of $f$, and $f^0$ is the identity map. If the inverse $f^{-1}$ is well-defined, then the *backward orbit* of $x$ is defined by $\gamma^-(x) = \{f^{-n}(x) \,|\, n \geq 0\}$. Finally, the *orbit* of $x$ is the sequence of all positions visited by $x$, $\gamma(x) = \gamma^-(x) \cup \gamma^+(x)$.

We will come back to maps in much more detail in Section 1.3. For now we just notice that maps are very similar to ordinary differential equations (ODEs), but with a discrete time. We now put this in the context of *dynamical systems*.

## 1.1 Flows

A first-order system of differential equations is written as

$$\frac{du}{dt} = f(u, t; \mu). \tag{1.1.1}$$

where $u \in \mathbb{R}^n$ is the dependent variable, t is the independent time variable, and $\mu$ is the set of parameters for the system. Sometimes the parameter dependence is denoted by a subscript as in $f_\mu(u, t)$.

A vector field is defined by a map $F : A \subset \mathbb{R}^n \to \mathbb{R}^n$ that assigns a vector $F(x)$ to each point $x$ in its domain $A$. A system governed by a time-independent vector field is called autonomous; otherwise it is called non-autonomous. Of course, any non-autonomous vector field can be converted to an autonomous vector field of a higher dimension (by adding the equation $\frac{dt}{dt} = 1$).

The flow $\phi$ of the vector field $F$ is defined by:

$$\frac{\partial}{\partial t}\phi(x,t) = F(\phi(x,t)) \tag{1.1.2}$$

$$\phi(x,0) = x. \tag{1.1.3}$$

The position $x$ is the *initial condition*. The collection of all states $u$ of the dynamical system is called the phase space. A *solution curve* $u(t)$ (also called an *orbit*) of the flow is a single solution of the above differential equation based at some point $u_0$ in phase space. In other words, $u(t)$ is the solution of

$$\tfrac{du(t)}{dt} = F(u(t),t) \quad \text{with} \quad u(t_0) = u_0, \tag{1.1.4}$$

hence $u(t) = \phi(u_0, t - t_0)$.

The term 'flow' describing the evolution of the system in phase space comes from the motion of a fluid. The flow $\phi(x,t)$ is regarded as a function of the initial condition $x$ and the single parameter time $t$. The flow $\phi$ tells us the position of the initial condition $x$ after a time $t$.

The flow is often written as $\phi_t(x)$ to highlight its dependence on the single parameter $t$. The flow $\phi_t$ is also known as the evolution operator. Composition of the evolution operator is defined in a natural way: starting at the state $x$ at time $s = 0$, it first flows to the point $\phi_s = \phi(x,s)$ and then on to the point $\phi_{t+s} = \phi(\phi(x,s),t)$. The evolution operator satisfies the group properties

$$\phi_0 = \text{identity}, \qquad \phi_{t+s} = \phi_t \circ \phi_s, \tag{1.1.5}$$

which are taken as the defining relations of a flow for a *dynamical system*. We note that the flow $\phi_t$ is not always defined for all $t \in \mathbb{R}$. For example, a common situation is that the flow flows forward in time, but not backward.

**Exercise 1** The harmonic oscillator is described by $\frac{d^2u}{dt^2} = -u$. What is the phase space of this dynamical system? Calculate the flow $\phi$. Is the flow defined for all $t \in \mathbb{R}$?

**Exercise 2** Maps lead to dynamical systems with a discrete time, which will in general exist for all $t \in \mathbb{N}$. Give the discrete flow $\phi(x,n)$ of the following maps $f, g : \mathbb{R} \to \mathbb{R}$:

$$x \mapsto f(x) = \tfrac{x}{2} \quad \text{and} \quad x \mapsto g(x) = x^2. \tag{1.1.6}$$

Is the flow defined for all $n \in \mathbb{Z}$? And when we restrict the domain of $f$ and $g$ to the interval $[0,1]$?

**Exercise 3** The notion of a dynamical system can be generalised to some types of partial differential equations. What replaces the phase space in this case? Does the flow exist for positive and negative time in the case of the Wave equation, and the Heat equation?

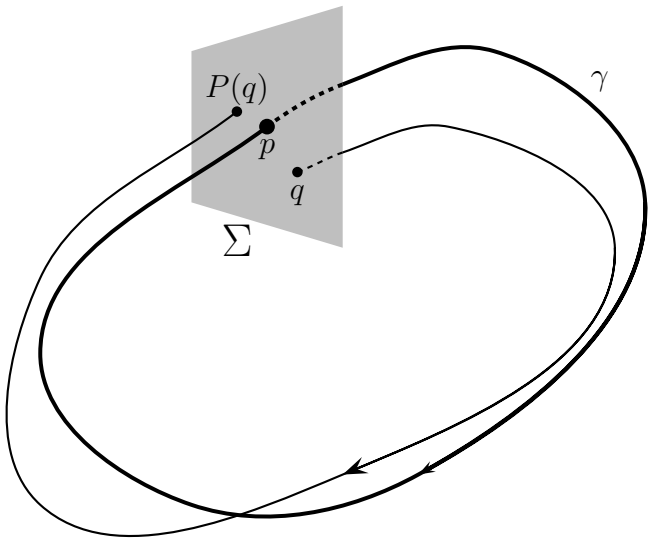


Figure 1.1. Construction of a Poincaré map from a local cross section.

## 1.2 Poincaré Maps

Flows and maps are closely related as the following construction of a so-called Poincaré map illustrates. Let $\gamma$ be an orbit of a flow $\phi_t$ in $\mathbb{R}^n$. As illustrated in Figure 1.1 it is often possible to find a local cross section $\sigma \in \mathbb{R}^n$ about $\gamma$, which is of dimension $n-1$.

The cross section need not be planar; however, it must be transverse to the flow. All the orbits in the neighbourhood of $\gamma$ must pass through $\Sigma$ (the requirement is that $F(x) \cdot N(x) \neq 0$ for all $x \in \Sigma$, where $N(x)$ is the unit normal vector to $\Sigma$ at $x$). Let $p$ be a point where $\gamma$ intersects $\Sigma$, and let $q \in \Sigma$ be a point in the neighbourhood of $p$. Then the *Poincaré map* (or *first return map*) is defined by

$$P : \Sigma \to \Sigma, \qquad P(q) = \phi_\tau(q), \tag{1.2.1}$$

where $\tau = \tau(q)$ is the time taken for an orbit starting at $q$ to return to $\Sigma$. It is especially useful to define a Poincaré map in the neighborhood of a periodic orbit. If the orbit $\gamma$ is periodic of period $T$, then $\tau(p) = T$. A periodic orbit that returns directly to itself is a fixed point of the Poincaré map. Moreover, an orbit starting at $q$ close to $p$ will have a return time close to $T$. The map $P$ now contains all the information about the dynamics in the neighbourhood of the periodic orbit $\gamma$.

**Exercise 4** Consider the flow in Exercise 1. Choose a periodic orbit and a cross section $\Sigma$, and determine the Poincaré map.

**Exercise 5** Consider the following vector field in $\mathbb{R}^2$:

$$\frac{dx}{dt} = \mu x - y - x(x^2 + y^2), \tag{1.2.2}$$

$$\frac{dy}{dt} = x + \mu y - y(x^2 + y^2), \tag{1.2.3}$$

where $\mu \in \mathbb{R}$ is a parameter. Draw the phase plane. Switch to polar coordinates $(r, \theta)$ and show that there is a unique periodic solution $\gamma$. Take a cross section (for example $\theta = 0$), and compute the Poincaré map of $\gamma$.

## 1.3 Maps

A simple model of a nonlinear system is given by the difference equation known as a the *logistic* map (or *quadratic* map)

$$x_{n+1} = \lambda x_n(1 - x_n). \tag{1.3.1}$$

It was studied as early as 1845 by Verhulst as a model for population growth. He was led to this difference equation by the following reasoning. Suppose that in any given year, indexed by the subscript $n$, the (normalized) population is $x_n$. Then to find the population $x_{n+1}$ in the next year it seems reasonable to assume that it will be proportional both to the current population $x_n$, and to the remaining inhabitable space $1-x_n$. The product of these two factors and $\lambda$ gives the quadratic map, where $\lambda$ is some parameter that depends on the fertility rate, the initial living area, the average disease rate, and so on.

Given the logistic map as our model for population dynamics, it now looks like an easy problem to predict the future population. Will it grow, decline, or vary in a cyclic pattern? As we will see, this question is easy to answer for some values of $\lambda$, but not for others. The dynamics are difficult to predict because, in addition to exhibiting cyclic behavior, it is also possible for the population to vary in a chaotic manner. An examination of the dynamics of the logistic map provides an excellent introduction to the rich behavior that can exist in nonlinear systems.

We have introduced the map

$$x \mapsto f_\lambda(x) = \lambda x(1 - x), \tag{1.3.2}$$

and we recall that we think of a map $f : x_n \to x_{n+1}$ as generating a sequence of points. With the seed $x_0$, define $x_n = f^n(x_0)$ and consider the sequence $x_0, x_1, x_2, \ldots$, as an orbit of the map. To find the itinerary of an individual orbit all we need is a computer program (see Section 1.6).

In addition to doing a calculation, there is a graphical procedure for finding the itinerary of an orbit. This graphical method is illustrated in Figure 1.2 (a so-called *web* diagram) and can be described as follows:

1. Start at $x_0$ on the horizontal axis.
2. Move vertically, up or down, until you hit the graph $f(x)$.
3. Move horizontally until you hit the diagonal line $y = x$.
4. Repeat steps 2 and 3.

**Exercise 6** Show by graphical analysis that if $x_0 \notin [0, 1]$ and $\lambda > 1$, then $f^n(x_0) \to -\infty$ as $n \to \infty$. What happens when $\lambda \leq 1$?

**Exercise 7** Investigate the logistic map numerically, i.e., generate web diagrams for different values of $\lambda$. see Section 1.6

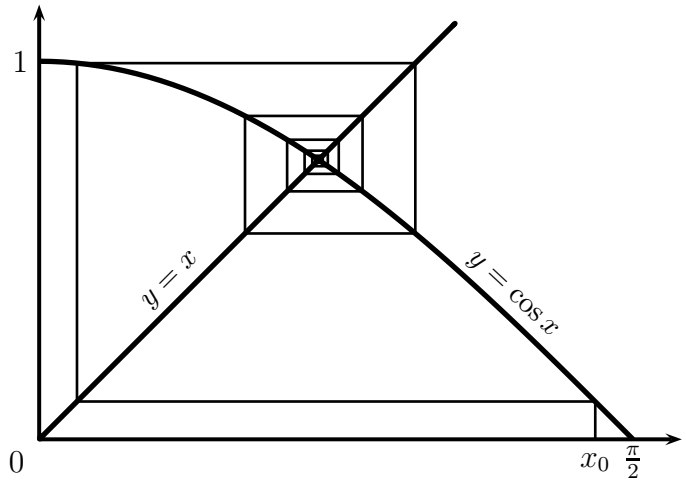


Figure 1.2. The graphical method for iterating a map: a web diagram.

## 1.4 Fixed points and periodic orbits

A linear map of $\mathbb{R}$ to itself is given by $f(x) = \alpha x$. Clearly, for $x_0 \neq 0$ we have

$$\begin{array}{ll} \lim_{n\to\infty} x_n = \pm\infty & \text{if } |\alpha| > 1. \\ \lim_{n\to\infty} x_n = 0 & \text{if } |\alpha| < 1. \end{array} \tag{1.4.1}$$

A point that maps to itself is called a *fixed point*. If $|\alpha| < 1$ then the origin is an *attracting* fixed point or a *sink*, since nearby points tend to 0. If $|\alpha| > 1$, then the origin is a *repelling* fixed point or a *source*, since points near the origin always tend away from it.

Graphical analysis allows us to see why certain fixed points are locally attracting and others repelling. Here we encounter an example of the phenomenon that (in)stability can often be determined by looking at the linear approximation.

**Exercise 8** Examine the fixed points of $f(x) = x^2$ with the help of a web diagram.

**Exercise 9** Argue by drawing web diagrams that the local stability of a fixed point is often determined by the slope of the curve passing through the fixed point. What is the criterion for stability?

**Exercise 10** What happens at the boundary of stability and instability? Examine the maps $f(x) = x + x^3$ and $f(x) = x - x^3$.

**Exercise 11** The *tent-map* is defined by

$$T_\lambda(x) = \lambda \min\{x, 1-x\} \qquad \text{for } x \in [0,1] \text{ and } \lambda > 0. \tag{1.4.2}$$

Examine the stability of the fixed point.

An orbit of a map is periodic if it repeats itself after a finite number of iterations. A period $n$ point repeats itself after $n$ iterations and thus is a solution of the equation

$$f^n(x) = x. \tag{1.4.3}$$

In other words, a period $n$ point is a fixed point of the $n^{\text{th}}$ composite function of $f$. Accordingly, the stability of this fixed point and of the corresponding period $n$ orbit is determined by the derivative of $f^n(x_0)$.

**Exercise 12** If a periodic orbit is given by $x_0, x_1, \ldots x_{n-1}$. Show that $(f^n)'(x_0) = (f^n)'(x_i)$ for $i = 1 \ldots n-1$.

Our discussion about the fixed points of a map is summarized in the following two definitions concerning fixed points, periodic points, and their stability.

**Definition 1** Let $f : \mathbb{R} \to \mathbb{R}$. The point $x_0$ is a *fixed point* if $f(x_0) = x_0$. The point $x_0$ is a *periodic point* of period $n$ if $f^n(x_0) = x_0$ but $f^i(x_0) \neq x_0$ for $0 < i < n$. The point $x_0$ is *eventually periodic* if $f^m(x_0) = f^{m+n}(x_0)$ for some $m, n > 0$, but $x_0$ is not itself periodic.

A point $x_0$ is eventually periodic if its orbit becomes periodic after a finite number of iterates. For example, when $f(x) = x^2$ then $x_0 = -1$ is eventually periodic with period 1 (also called *eventually fixed*).

**Definition 2** A periodic point $x_0$ of period $n$ is *attracting* if $|(f^n)'(x_0)| < 1$. The periodic point $x_0$ is *repelling* if $|(f^n)'(x_0)| > 1$. The point $x_0$ is neutral if $|(f^n)'(x_0)| = 1$.

Notice that an attracting periodic orbit is stable and a repelling periodic orbit is unstable, but a neutral periodic orbit may be either stable or unstable.

**Exercise 13** Looking at the Poincaré map, determine the stability of the periodic orbit in Exercise 5.

## 1.5 The logistic map I

We will now consider the logistic map (1.3.2) on the interval $[0, 1]$, which implies that we restrict ourselves to $\lambda \in (0, 4]$ (this is suggested by the biological interpretation, namely that $x$ is a population which cannot be negative).

**Exercise 14** Calculate the fixed point(s) of the logistic map and its/their stability (for all $0 < \lambda \leq 4$). Also solve this problem graphically by drawing $y = f(x)$ and $y = x$ in the same graph for several values of $\lambda$. Draw a bifurcation diagram.

To investigate the period 2 orbits, we calculate

$$f^2(x) = f(f(x)) = \lambda(\lambda x(1-x))(1-\lambda x(1-x)). \tag{1.5.1}$$

**Exercise 15** Calculate the period 2 orbits and their stability. Draw a bifurcation diagram for the fixed points of the second iterate $f^2$.

Clearly the period 2 orbits can be found graphically as the intersection of $f^2(x)$ with the line $y = x$.

**Exercise 16** Draw graphs of $f(x)$ and $f^2$ (together with the line $y = x$) for $\lambda$ just below, at, and just above the bifurcation value $\lambda = 3$. What is typical for the graphs of $f$ and $f^2$ as the parameter $\lambda$ passes the bifurcation value(s).

The transition from a fixed point (period 1 orbit) to a period 2 orbit is called *period doubling*.

**Exercise 17** Looking at the analogy between $f(x)$ and $f^2(x)$, guess what happens as the 2-period orbit becomes unstable. Can you substantiate this using graphical methods?

To explore the dynamics of the logistic map further, we can choose an initial condition $x_0$ and a parameter value $\lambda$, and then iterate the map using a computer program to see where the orbit goes. By playing this game long enough we notice the following. First, if $1 < \lambda < 3$ then the graphical analysis shows that the dynamics of the logistic map are simple: there is only one attracting fixed point and one repelling fixed point. Second, the initial condition we pick is usually not important in determining the attractor (although the value of $\lambda$ is very important). Therefore, when studying the logistic map, it will usually suffice to pick a single initial condition from the unit interval. By the *attractor* we mean the 'limit' of an orbit for large $n$. If $0 < \lambda \leq 1$ then the only attractor is a stable fixed point at zero, whereas for $1 < \lambda \leq 3$ the only attractor is the other fixed point.

A bifurcation diagram is a particularly powerful method for studying the attractors in the logistic map. A bifurcation diagram is a plot of an asymptotic solution on the vertical axis and a control parameter on the horizontal axis. To construct a bifurcation diagram for the logistic map one only needs a simple computer) program (see Section 1.6) for iterating the logistic map. The algorithm consists of the following steps:

1. Set $\lambda = 0$, and $x_0 = 0.1$ (almost any $x_0$ will do).

2. Iterate the logistic map 200 times to remove the transient solution, and then print the next 200 points, which are presumably part of the attractor;

3. Increment $\lambda$ by a small amount, and reset $x_0 = 0.1$.

4. Repeat steps 2 and 3 until $\lambda = 4$.

When plotted in Figure 1.3, the output of our simple program produces a bifurcation diagram of stunning complexity .

The diagram shows a so-called period doubling route to chaos. An infinite number of period doublings occur for $3 < \lambda < 3.57$. It also shows periodic windows (white bands) within the chaotic regions. A period three window begins at $\lambda \approx 3.83$, and a

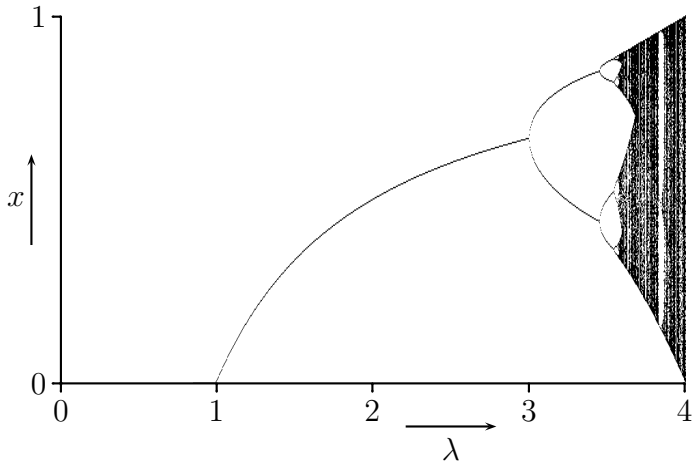


Figure 1.3. Bifurcation diagram for the logistic map (the Feigenbaum diagram).

period five window begins at $\lambda \approx 3.74$. Looking closely at the periodic windows we see that each branch of these periodic windows also undergoes a sequence of period doublings.

**Exercise 18** Calculate the bifurcation diagram numerically for the map $f_\lambda(x) = \lambda \sin x$ (see Section 1.6). Observe the similarities with the bifurcation diagram for the logistic map in Figure 1.3.

**Exercise 19** Calculate the bifurcation diagram numerically for the tent-map $f_\lambda(x) = \lambda \min\{x, 1-x\}$ (see Section 1.6). Observe the difference with the bifurcation diagram for the logistic map in Figure 1.3.

We will see more chaotic behaviour in the next Chapter.

## 1.6 Computer programs

For the logistic map some links to nice interactive applets on the internet can be found at
http://www.wi.leidenuniv.nl/~gvdberg/maps.html
Three short maple programs (1. iterating maps and web diagrams; 2. bifurcation diagrams; 3. Lyapunov exponents) are available on the same webpage. These are not very fast, but they have the advantage that you have a lot of control.

You can also use the (slightly inflexible) program on the computers on the third floor. You could also write your own program.

# Chapter 2

# Chaotic dynamics

In this chapter we investigate how maps can lead to chaotic dynamics.

## 2.1 The logistic map II

In the diagram in Figure 1.3 we encounter several types of bifurcations.

- At $\lambda = 1$ a second fixed point is generated, and the origin becomes unstable. Of course the second fix point already existed for $\lambda < 1$, but it was negative. At $\lambda = 1$ the stability is *exchanged* between the origin and this second fixed point. It is called a *transcritical* bifurcation.

- At $\lambda = 3$ a stable 2-periodic orbit bifurcates from the fixed point, and the fixed point becomes unstable. Considering the second iterate $f^2(x)$, we see that at $\lambda = 3$ a stable fixed point of $f^2$ bifurcates to three fixed points, two stable and one unstable. This is a *pitchfork* bifurcation. The bifurcations is also called a *period doubling* bifurcation. The 2-periodic orbit undergoes a period doubling bifurcation at $\lambda = 1 + \sqrt{6}$.

- At $\lambda \approx 3.83$ two period three orbits appear out of the blue, one stable and one unstable. This is called a *saddle-node* bifurcation.

We warn that a bifurcation diagrams showing only attracting solutions can be somewhat misleading. Much of the structure in the bifurcation diagram can only be understood by keeping track of both the stable attracting solutions and the unstable repelling solutions. For example, two 2-periodic orbits exist for all $3 < \lambda \leq 4$ (see Exercise 15).

**Exercise 1** Use a computer algebra package to find the exact parameter value for which period 3 orbits are born. How many period 3 orbits are there?

For all $\lambda$ for which a period 3 orbit exists we can apply the following general result

**Theorem 1** *Let $f : [0, 1] \to [0, 1]$ be a continuous function such that a periodic orbit of period 3 exists. Then there exist periodic orbits of any period.*

**Exercise 2** Prove Theorem 1 along the following lines:

- If $J$ is a closed subinterval of $[0,1]$ and $J \subset f^m(J)$ for some $m > 0$, then there exists a point $x \in J$ such that $f^m(x) = x$ (in words: if $J$ is contained the image of $J$ under $f^m$, then there is a fixed point of $f^m$ in $J$).

- If $J_0, J_1, \ldots J_m$ are closed subintervals of $[0,1]$ such that $J_{i+1} \subset f(J_i)$ for $i = 0, \ldots, m-1$, then there is a closed interval $J \subset J_0$ such that $f^i(J) \subset J_i$ for $i = 0, \ldots, m$ and such that $f^m(J) = J_m$.

- Let $p_1 = f(p_3)$, $p_2 = f(p_1)$, $p_3 = f(p_2)$ be a period 3 orbit. Suppose that $p_1 < p_2 < p_3$. Then $[p_1, p_3] \subset f([p_2, p_3])$ and $[p_2, p_3] \subset f([p_1, p_2])$.

- For each $m > 0$ choose appropriate $J_i$ for $i = 0, \ldots, m-1$, to derive that there exists an orbit of period $m$ (which is not periodic with period smaller than $m$)

- Deal with the case that $p_1 < p_3 < p_2$ and finish the proof.

**Exercise 3** For $\lambda = 4$, draw $f^n$, for $n = 1, 2, 3, 4$. How many fixed points does $f^n$ have? How many different periodic orbits of period $n$ are there for $n = 1, 2, \ldots 10$?

The diagram in Figure 1.3 reveals not one, but rather an infinite number of period doubling bifurcations. As $\lambda$ is increased a stable period 2 orbit becomes a stable period 4 orbit, and this in turn becomes a stable period 8 orbit, and so on. This sequence of period doubling bifurcations is known as a period doubling cascade. This process appears to converge at a finite value of about 3.57, beyond which a non-periodic motion appears to exist. The period 3 orbit also undergoes a period doubling cascade, and it occurs in many nonlinear systems. The period doubling route is one common way (but certainly not the only way) by which a (low-dimensional) nonlinear system can progress from a simple behavior (one or a few periodic orbits) to a complex behavior (chaotic motion and the existence of an infinity of unstable periodic orbits).

Feigenbaum discovered this period doubling cascade in 1976. He recorded the values of $\lambda$ at which the first few period doubling bifurcations occur and he recognized that the convergence appears to follow that of a geometric series. Let $\lambda_n$ be the value of the $n^{\text{th}}$ period doubling bifurcation, and define $\lambda_\infty$ as $\lim_{n\to\infty} \lambda_n$. Feigenbaum guessed that this sequence obeys a geometric convergence, i.e., $\lambda_\infty - \lambda_n = c\delta^{-n}$ as $n \to \infty$ for some $c > 0$ and $\delta > 1$. This implies that

$$\delta = \lim_{n\to\infty} \frac{\lambda_n - \lambda_{n-1}}{\lambda_{n+1} - \lambda_n}. \tag{2.1.1}$$

The constant $\delta \approx 4.6692$ is nowadays called the Feigenbaum constant, because Feigenbaum went on to show that this number is universal in that it arises in a wide class of systems that are close to the single-humped map, not only in mathematical equations, but also in a variety of physical systems. An example is shown in Figure 2.1 where the physical phenomenon is Rayleigh-Bénard convection in a very small cell (which will be the subject of Chapter 8.

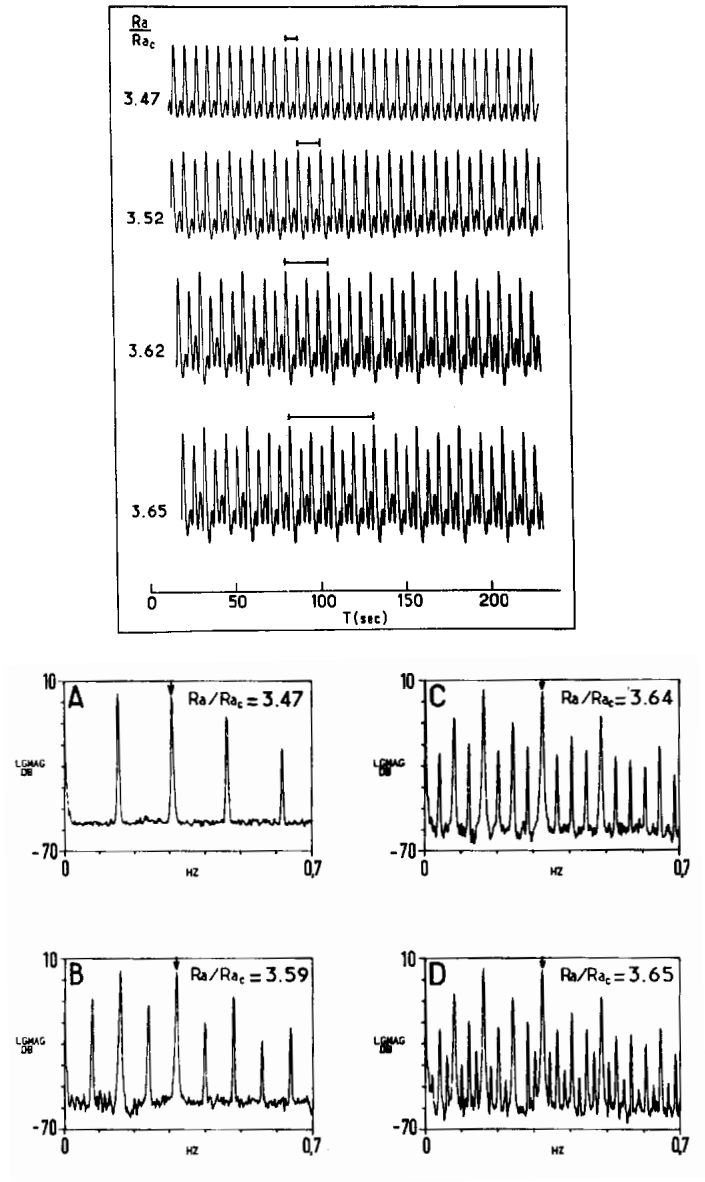


Figure 2.1. Period doubling cascade in a Rayleigh-Bénard experiment. The changing shape of the signals in the top half of the picture clearly shows the period-doubling process that takes place as the control parameter $\frac{\mathrm{Ra}}{\mathrm{Ra}_c}$ is increased. The line segments indicate the length of one period. The Fourier analysis in the bottom half of the picture unambiguously shows that the system undergoes period doublings . *From: Libchaber et al., Physica D, vol. 7, p. 73 (1983).*

## 2.2 Chaos

Another simple example of a nonlinear one-dimensional map is

$$x \mapsto E_2(x) \stackrel{\text{def}}{=} 2x \bmod 1 \qquad \text{for } x \in [0,1). \tag{2.2.1}$$

Although this map is of course discontinuous, one may also think of it as smooth map from the unit circle to itself. Namely, set $z = e^{2\pi i x}$, then

$$z \mapsto z^2 \qquad \text{for all } z \in \mathbb{C} \text{ with } |z| = 1. \tag{2.2.2}$$

**Exercise 4** Find the fixed points of $E_2$ and the periodic orbits of period 2,3 and 4. Determine the stability of the periodic points.

**Exercise 5** Set $y_n = \sin^2(\pi x_n)$ and show that $y_n$ obeys the logistic map with $\lambda = 4$.

When we write $x \in [0,1)$ in binary expansion, then $E_2$ acts on an initial condition as a shift map. At each iteration we multiply the previous iterate by two, which is a left shift, and then apply the (mod 1)-function, which erases the integer part. For example

$$E_2(0.1110101\ldots) = 1.110101\ldots \bmod 1 = 0.110101\ldots. \tag{2.2.3}$$

**Exercise 6** Identify all points that are eventually fixed points, i.e., $E_2^{n+1}(x) = E_2^n(x)$ for some $n > 0$.

**Exercise 7** Identify all eventually periodic points of $E_2$.

**Exercise 8** Identify a dense orbit of $E_2$, i.e., find an $x \in [0,1)$ such that the orbit starting at $x$ comes arbitrarily close to any point in $[0,1)$.

The map has a very sensitive dependence on initial conditions. If our starting value has a precision of $n$ digits in binary notation, then after $n+1$ iterations all information has evaporated. Assuming that our initial condition has some small error $\epsilon$, the growth rate of the error is

$$|E_2^n(x_0) - E_2^n(x_0+\epsilon)| = |2^n x_0 - 2^n(x_0+\epsilon)| = 2^n\epsilon = \epsilon e^{n\ln 2}. \tag{2.2.4}$$

If we think of $n$ as time, then the example the error grows at a constant exponential rate of $\ln 2$.

We can quantify the amount of sensitive dependence the system exhibits, that is, the rate at which an initial error grows. For this purpose we define the Lyapunov exponent of an orbit:

$$\lambda(x) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \ln|f'(f^{k-1}(x))| \tag{2.2.5}$$

The Lyapunov exponent thus measures the mean value of the logarithm of the derivative of $f$ along the orbit. As mentioned above, $\lambda(x)$ is the rate at which orbits grow apart.

**Exercise 9** What is the relation between (in)stability and the Lyapunov exponent of a periodic orbit or fixed point? Show that $\lambda(x) = \lim_{n\to\infty} \frac{\ln|(f^n)'(x)|}{n}$.

**Exercise 10** Calculate the Lyapunov exponent of an orbit of $E_2$.

**Exercise 11** Investigate numerically the Lyapunov exponent of orbits of the logistic map for various values of $\lambda$ (see Section 1.6). Try to explain what you see.

The map $E_2$ is a very simple map, completely deterministic, but it has the characteristics of *chaotic* dynamics:

1. There is a (countable) infinite number of periodic orbits having arbitrarily large period.
2. There is an uncountable number of non-periodic orbits.
3. There is a dense orbit
4. There is sensitive dependence on initial values; (almost) all orbits have positive Lyapunov exponent.

We will now investigate what happens to the logistic map when $\lambda > 4$. In Exercise 6 we saw that if $x_0 \notin [0,1]$, then the orbit goes to $-\infty$.

**Exercise 12** Fix $\lambda > 4$. Which points inside the interval $[0,1]$ are mapped outside the interval after one iteration? And after two iterations?

To make things a little easier to compute, we will look at the following tent-map:

$$x \mapsto T_3(x) \stackrel{\text{def}}{=} 3\min\{x, 1-x\} \qquad \text{for } x \in \mathbb{R}. \tag{2.2.6}$$

The dynamics of the map $T_3$ are completely analogous to those of the logistic map for $\lambda > 4$, but for $T_3$ the analysis is slightly easier. We define the sets

$$A_i \stackrel{\text{def}}{=} \{x \in [0,1] \,|\, T_3^i(x) \in [0,1]\} \qquad \text{for all } i \geq 1. \tag{2.2.7}$$

**Exercise 13** Show that $A_{i+1} \subset A_i$. Determine the measure of $A_i$.

We now define

$$\Lambda \stackrel{\text{def}}{=} \bigcap_{i=1}^{\infty} A_i. \tag{2.2.8}$$

**Exercise 14** Expanding $x \in [0,1]$ in base three, characterize $\Lambda$. Show that $\Lambda$ is closed, totally disconnected, and that every point $x$ in $\Lambda$ is a limit point (i.e., in every neighbourhood of $x$ there are infinitely many points of $\Lambda$). Such a set is called a *Cantor set* (in this case it is the middle-third Cantor set).

Now, $\Lambda$ is a so-called forward invariant set.

**Definition 3** A set $\Gamma$ is said to be *invariant* under the map $f$ if for any $x_0 \in \Gamma$ we have $f^n(x_0) \in \Gamma$ for all $n \in \mathbb{Z}$. It is *forward invariant* if this only holds for all $n \geq 0$. It is *backward invariant* if this only holds for all $n \leq 0$.

It follows from this definition that it makes sense to consider a map $f$ restricted to an invariant set.

**Exercise 15** Show that $T_3$ restricted to $\Lambda$ has all the properties (listed above) that characterize chaotic dynamics (use Exercise 14).

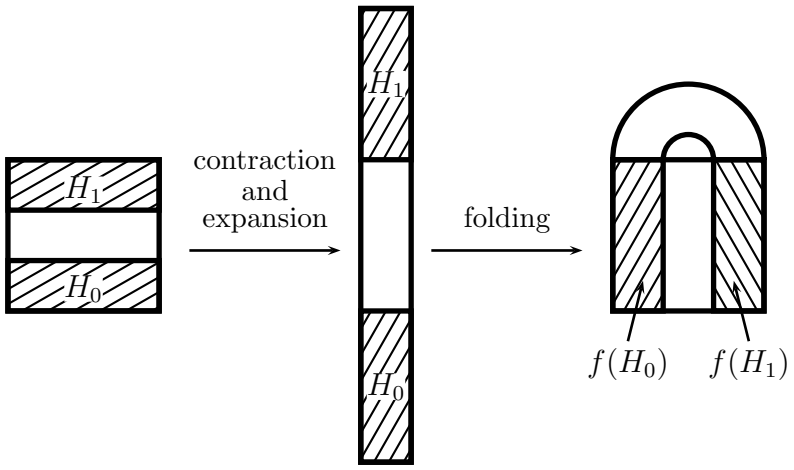


Figure 2.2. Forward iteration of the horseshoe map.

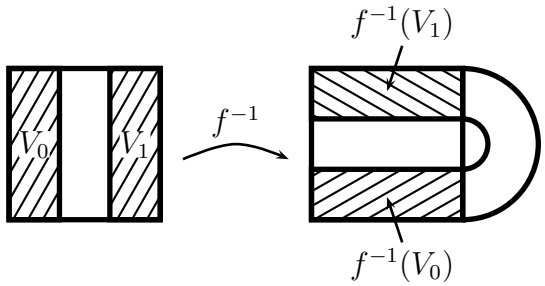


Figure 2.3. Backward iteration of the horseshoe map.

## 2.3 The Smale horseshoe

To finish our discussion of maps we investigate a map in $\mathbb{R}^2$ that was first studied by Smale in 1963. The Smale *horseshoe* is a mapping of the unit square $D = [0,1]\times[0,1]$, which contracts the horizontal directions, expands in the vertical direction, and then folds. The forward iteration of the horseshoe map is shown in Figure 2.2. Note that the image of $D$ under $f$ lies only partly in $D$. Points that leave the square are ignored. We define two horizontal strips

$$H_0 = \{(x,y) \in D \,|\, 0 \le y \le \tfrac{1}{3}\}, \tag{2.3.1}$$
$$H_1 = \{(x,y) \in D \,|\, \tfrac{2}{3} \le y \le 1\}. \tag{2.3.2}$$

The horseshoe map takes these horizontal strips to two vertical strips

$$f(H_0) = V_0 = \{(x,y) \in D \,|\, 0 \le x \le \tfrac{1}{3}\}, \tag{2.3.3}$$
$$f(H_1) = V_1 = \{(x,y) \in D \,|\, \tfrac{2}{3} \le x \le 1\}. \tag{2.3.4}$$

The inverse $f^{-1}$ of the horseshoe map is shown in Figure 2.3. Clearly $f^{-1}(V_0) = H_0$ and $f^{-1}(V_1) = H_1$. The invariant set $\Lambda$ of the horseshoe map is the collection of all

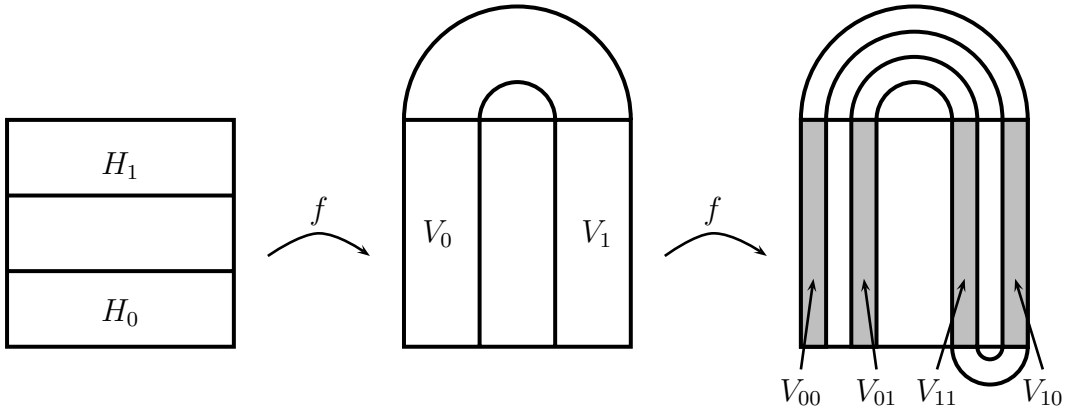


Figure 2.4. Second forward iteration of the horseshoe map.

points that remain in $D$ under all iterations of $f$:

$$\Lambda = \ldots \cap f^{-2}(D) \cap f^{-1}(D) \cap D \cap f^1(D) \cap f^2(D) \cap \ldots = \bigcap_{n=-\infty}^{\infty} f^n(D). \qquad (2.3.5)$$

We are now going to encode the points in $\Lambda$. Let us look at the second iterate of $f$ in Figure 2.4. We see that the horizontal rectangles $H_0$ and $H_1$ are mapped under $f^2$ to four vertical rectangles. We have labeled them $V_{00}$, $V_{01}$, $V_{10}$, and $V_{11}$, so that for $a_0, a_{-1} \in \{0, 1\}$

$$V_{a_0,a_{-1}} = \{p \in D \,|\, f^0(p) \in V_{a_0} \text{ and } f^{-1}(p) \in V_{a_{-1}}\}. \qquad (2.3.6)$$

We see that $f(V_0) \cap D = V_{00} \cup V_{10}$.

**Exercise 16** Draw the third iterate of $D$ under the horseshoe mapping. Extend the labeling to this case.

This process can be extended to any iterate of $f$ and we obtain

$$V_{a_0,a_{-1},\ldots a_{-n}} = \{p \in D \,|\, f^{-i}(p) \in V_{a_{-i}} \text{ for } i = 0, 1, \ldots, n\}. \qquad (2.3.7)$$

**Exercise 17** What does $V_{a_0,a_{-1},\ldots a_{-n}}$ look like if we take $n = \infty$?

**Exercise 18** Draw $f^{-1}(D)$ and $f^{-2}(D)$ and label the four/eight horizontal rectangles that constitute the image of $H_0 \cup H_1$ suitably by $H_{b_0,b_1}$ and $H_{b_0,b_1,b_2}$ with $b_0, b_1, b_2 \in \{0, 1\}$. How can you extend this to all iterates of $f^{-1}$?

**Exercise 19** Show that every point $p \in \Lambda$, can be uniquely labeled by a bi-infinite sequence $s(p) = \{s_n(p)\}_{n=-\infty}^{\infty}$ with $s_n \in \{0, 1\}$ for all $n \in \mathbb{Z}$, such that

$$f^n(p) \in H_{s_n} \qquad \text{for all } n \in \mathbb{Z}. \qquad (2.3.8)$$

How are the $s_n$ related to $a_n$ and $b_n$?

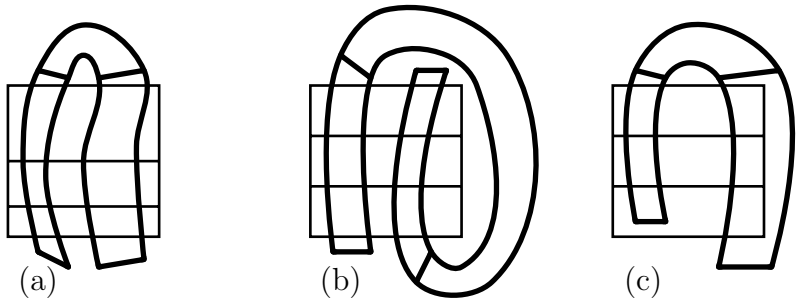


Figure 2.5. (a) A deformed horseshoe map; (b) A different scenario leading to similar dynamics; (c) Not a suitable horseshoe.

**Exercise 20** What does the invariant set $\Lambda$ look like? Can you detect periodic orbits in $\Lambda$? Can you find a dense orbit?

In general, the exact shape of the horseshoe map is not important. What is essential is that there is contraction in one direction, expansion in the other direction, and folding. For example the mapping in Figure 2.5(a) will lead to the same type of invariant set with similar (chaotic) dynamics. The horseshoe map is of importance because it turns up in a wide variety of dynamical systems, and if you encounter a situation in which a (Poincaré) map has (on some domain) the structure of a horseshoe map, then it can be concluded that the underlying dynamical system had chaotic behaviour.

## 2.4 References

The Bergé–Pomeau–Vidal book mostly gives physical intuition. The book by Guckenheimer and Holmes (*Nonlinear oscillations, dynamical systems, and bifurcations of vector fields* (1983)) is a nice starting point for the more mathematically inclined. If you want something in between you might find it at `http://www.nbi.dk/ChaosBook/` . A nice popular historical account is *Chaos: making a new science* (1987) by J. Gleick.

# Chapter 3

# The buckling instability

Buckling of a thin rod under axial thrust $P$ is a classical example of *Bifurcation*. The trivial state in which the rod is straight loses stability when the thrust exceeds a critical value $P_c$ and when the thrust is further increased, the stable state is one in which the rod is bent. The transition at $P_c$ is called bifurcation.

This is an old problem that goes back to Daniel Bernouilli and to Euler, who first investigated it around 1744. An elegant account of the formulation of the problem, as well as some history, can be found in the classical book by Love [L]. We also refer to the book by Koiter [K], to whom much important work on buckling is due, the proceedings of a seminar held in 1966-67 at New York University [R] and a survey paper by Stakgold [S].

## 3.1 The model

We consider a homogeneous thin rod of length $L$ constrained to move in a plane. Its end points are constrained to lie on a fixed straight line, the $x$ - axis: the left end point is asssumed fixed, and this point we choose as the origin $x = 0$, and the other end point is free to slide along the $x$ - axis. At this point a force will be applied to the rod in the direction of negative $x$ - axis (see Fig. 1). We shall study the effect of this force on the shape of the rod.

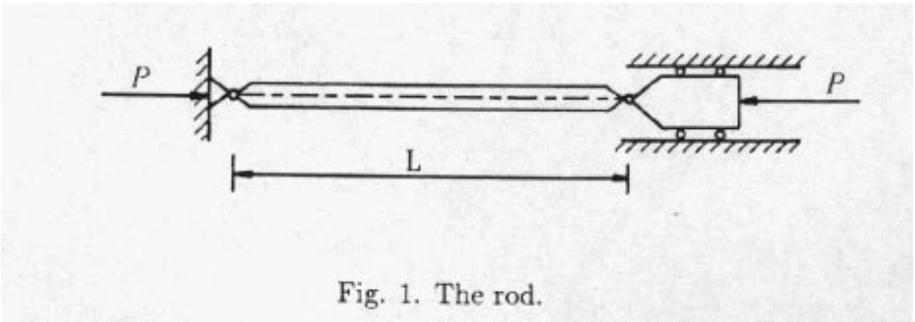


Fig. 1. The rod.

When no force is applied, the rod will be positioned along the $x$ - axis, taking up the interval $[0, L]$. If we then impose an axial load $P$, as described above, then for sufficiently large values of $P$ the rod will bend or *buckle*.

To describe this phenomenon, we need to write down the equations describing

the equilibrium positions of the rod. We use a Langrangian coordinate frame in that we shall follow a particle of the rod, whilst it changes position. We shall assume throughout that the movement of the rod takes place in the $xy$ - plane, the $y$ axis being chosen perpendicular to the $x$ - axis. If initially a particle is placed at the point $(X, 0)$, then its new position will be denoted by $(X + u, v)$. We assume that the rod is incompressible, so that we always have

$$s = X, \tag{3.1.1}$$

where $s$ denotes the arc length along the rod. This will serve as the independent variable.

The equation describing the position of the rod is obtained from the Bernouilli-Euler law relating the moment $M$ on a cross section of the rod and the curvature $d\theta/ds$ of the rod, where $\theta$ denotes the slope of the the rod (see Fig. 2).

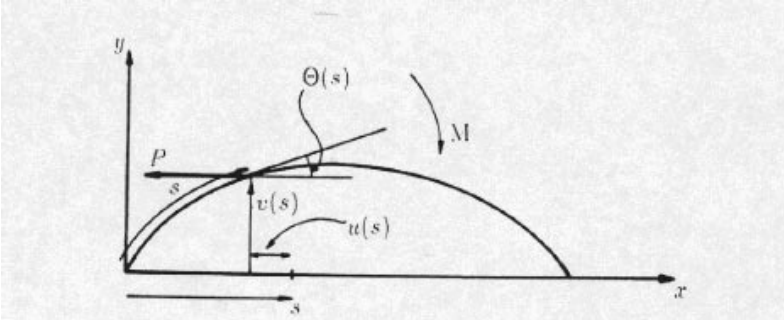

Fig. 2. Forces and moments

This relation is given by

$$M = -EI\frac{d\theta}{ds}, \tag{3.1.2}$$

where $E$ denotes modulus of elasticity and $I$ the moment of inertia of the cross section. To achieve equilibrium, this moment needs to balance the imposed moment at the point $s$:

$$M = Pv. \tag{3.1.3}$$

The system of equations is completed by the geometric condition

$$\frac{dv}{ds} = \sin\theta. \tag{3.1.4}$$

Thus, eliminating $M$ and $v$ from equations (3.2)-(3.4) we obtain the second order differential equation

$$\frac{d^2\theta}{ds^2} + \lambda \sin\theta = 0, \quad 0 < s < L, \tag{3.1.5}$$

where

$$\lambda = \frac{P}{EI}. \tag{3.1.6}$$

At the end-points of the rod, different conditions may prevail. They may be clamped; the boundary conditions then become

$$\theta(0) = \theta_0 \quad \text{and} \quad \theta(L) = \theta_1. \tag{3.1.7}$$

It may also happen that the moments are prescribed. In that case we have

$$\frac{d\theta}{ds}(0) = \alpha_0 \quad \text{and} \quad \frac{d\theta}{ds}(L) = \alpha_1. \tag{3.1.8}$$

A particularly important case is when the end points are free to rotate. We then have (3.8) with $\alpha_0 = 0$ and $\alpha_1 = 0$.

Of course one may also have a mixture of these two, say clamped at the origin and prescribed moment at $s = L$:

$$\theta(0) = \theta_0 \quad \text{and} \quad \frac{d\theta}{ds}(L) = \alpha_1. \tag{3.1.9}$$

**Remark.** Note that this is the same equation as the one describing the pendulum. However, for a pendulum we normally have conditions at one end of the interval, whereas here we have conditions at both ends of the interval.

**Exercise 1** Show that having found the function $\theta(s)$, we may find the displacements $u(s)$ and $v(s)$ through the relations

$$\frac{du}{ds} = \cos\theta(s) - 1 \quad \text{and} \quad \frac{dv}{ds} = \sin\theta(s), \tag{3.1.10}$$

combined with the conditions $u(0) = 0$ and $v(0) = 0$ at the origin.

## 3.2 The linear problem

The equation (3.5) for $\theta$ is nonlinear. This means that many standard methods for finding explicit expressions for solutions cannot be applied, and often that such expressions cannot be found at all.

One way to get around this difficulty is to *linearize* the equation. One observes that

$$\sin\theta \approx \theta \quad \text{for} \quad |\theta| << \pi \tag{3.2.1}$$

and so for small values of $\theta$ the equation

$$\theta'' + \lambda\theta = 0, \quad 0 < s < L \tag{3.2.2}$$

may be a good approximation of (3.5).

Suppose that we allow the end-points to rotate freely, that is we set

$$\theta'(0) = 0 \quad \text{and} \quad \theta'(L) = 0. \tag{3.2.3}$$

Equation (3.12) possesses a nonzero solution which satisfies (3.13) if and only if $\lambda$ equals one of the eigenvalues $\mu_n$ of the operator $-d^2/dx^2$ on $(0, L)$ with Neumann boundary conditions. They are given by

$$\mu_n = \left(\frac{n\pi}{L}\right)^2, \qquad n = 1, 2, 3, \ldots. \tag{3.2.4}$$

In that case we have

$$\theta(s) = A\varphi_n(s), \qquad 0 \le s \le L, \tag{3.2.5}$$

where $A$ is an arbitrary constant and $\varphi_n$ the eigenfunction corresponding to $\mu_n$, suitably normalized:

$$\varphi_n(s) = \cos\Big(\frac{n\pi s}{L}\Big), \qquad 0 \le s \le L, \tag{3.2.6}$$

Thus, we can draw the following conclusion.

**Theorem 2** *Problem (3.12),(3.13) has a solution if and only if*

$$\lambda \in \{\mu_1, \mu_2, \ldots\}. \tag{3.2.7}$$

*If $\lambda = \mu_n$ for some $n \ge 1$, then Problem (3.12),(3.13) has infinitely many solutions.*

From a physical point of view this answer is not very satisfactory.

In this context it is interesting to consider a related problem, in which one end-point, say the origin ($s = 0$) is free to rotate, but a small torque is applied at the other end point:

$$\theta'(0) = 0 \quad \text{and} \quad \theta'(L) = \varepsilon, \qquad \varepsilon > 0. \tag{3.2.8}$$

**Exercise 2** Investigate Problem (3.12) (3.18): when does this problem have a solution, and if it does, how many solutions does it have? Find explicit expression(s).

Since the eigenvalues $\mu_n$ will play an important role in what follows we recall here their variational characterization. Let

$$X_1 = \{w \in H^1(0, L) : \int w = 0\}, \tag{3.2.9}$$

where $H^1$ denotes the so-called *Sobolev space* of function in $L^2(0, L)$ which have generalised derivatives that are also functions in $L^2(0, L)$. The integral in (3.19), like all later integrals, is taken over $(0, L)$. We set

$$I(w) = \frac{\int w'^2}{\int w^2}. \tag{3.2.10}$$

Then we can characterise $\mu_1$ by

$$\mu_1 = \inf\{I(w) : w \in X_1,\ w \ne 0\}. \tag{3.2.11}$$

The second eigenvalue $\mu_2$ is characterised by taking the infimum of $I(w)$ over a smaller set of functions:

$$X_2 = \{w \in X_1 : \int w\varphi_1 = 0\} \tag{3.2.12}$$

Thus we have

$$\mu_2 = \inf\{I(w) : w \in X_2,\ w \ne 0\}. \tag{3.2.13}$$

Continuing in this manner all the eigenvalues $\mu_n$ can be characterised:

$$\mu_n = \inf\{I(w) : w \in X_n,\ w \neq 0\}, \tag{3.2.14}$$

where

$$X_n = \{w \in X_{n-1} : \int w\varphi_{n-1} = 0\}, \quad n \geq 2. \tag{3.2.15}$$

In this way we obtain an infinite orthogonal sequence of functions $\{\varphi_n\}$ corresponding to the eigenvalues $\{\mu_n\}$.

## 3.3 The full problem

We now return to the problem as it was originally formulated. We begin with the situation in which the rod is free to rotate at both end points. This yields the *Two-Point Boundary Value Problem*:

$$\text{(P)} \quad \begin{cases} \theta'' + \lambda \sin\theta = 0 & 0 < s < L, \\ -\pi \leq \theta \leq \pi & 0 < s < L, \\ \theta'(0) = 0 \quad \text{and} \quad \theta'(L) = 0. \end{cases} \tag{3.3.1}$$

To study this problem, we replace the condition at $x = L$ by a new one at $x = 0$, turning it into an *Initial Value Problem*:

$$\text{(PI)} \quad \begin{cases} \theta'' + \lambda \sin\theta = 0 & 0 < s < L, \\ \theta(0) = \gamma \in [0, \pi] \quad \text{and} \quad \theta'(0) = 0. \end{cases} \tag{3.3.2}$$

By standard ODE theory there exists for every $\gamma \in (0, \pi)$ a unique local solution, which we denote by $\theta(x, \gamma)$. We begin with some preliminary observations. Plainly, $\theta = 0$ and $\theta = \pm\pi$ are solutions. We shall call them the *trivial solutions*. Suppose that $\theta(s, \gamma)$ is a nontrivial solution. Then $\theta > 0$ in a neighbourhood of $s = 0$ and in fact

$$\theta' < 0 \ \text{ and } \ \theta'' < 0 \quad \text{as long as} \quad \theta > 0. \tag{3.3.3}$$

**Exercise 3** Prove (3.28).

Thus, the graph of $\theta$ is concave down as long as $\theta$ is positive and therefore $\theta$ must have a first zero $\sigma_1(\gamma, \lambda)$:

$$\sigma_1(\gamma, \lambda) = \sup\{x > 0 : \theta(\cdot, \gamma) > 0 \text{ on } [0, x)\}.$$

Beyond $\sigma_1$ we may continue $\theta$ by

$$\theta(s) = -\theta(2\sigma_1 - s), \qquad \sigma_1 < s < 2\sigma_1. \tag{3.3.4}$$

Note that

$$\theta'(s) = \theta'(2\sigma_1 - s), \tag{3.3.5}$$

and in particular

$$\theta'(2\sigma_1) = 0. \tag{3.3.6}$$

Thus, if we can find a value $\gamma_1 \in (0, \pi)$ such that:

$$2\sigma_1(\gamma_1, \lambda) = L, \tag{3.3.7}$$

then we have found a solution of Problem (P).

**Exercise 4** Use the continuity of solutions of the solution $\theta(x,\gamma)$ with respect to $\gamma$ on bounded set of $x$ to prove that

$$\sigma_1(\gamma,\lambda) \to \infty \quad \text{as} \quad \gamma \to \pi. \tag{3.3.8}$$

**Exercise 5 Exercise 3.5** Show that

$$\sigma_1(\gamma,\lambda) \to \frac{\pi}{2\sqrt{\lambda}} \quad \text{as} \quad \gamma \to 0 \tag{3.3.9}$$

**Exercise 6** Prove by means of an appropriate scaling of the variable $x$ that

$$\sigma_1(\gamma,\lambda) = \frac{\sigma_1(\gamma,1)}{\sqrt{\lambda}}.$$

To keep the notation simple we shall usually write

$$\sigma_1(\gamma) = \sigma_1(\gamma,1)$$

**Exercise 7** Prove that $\sigma_1(\gamma)$ is a continuous function on $(0,\pi)$.

Because $\sigma_1(\gamma)$ is continuous on $(0,\pi)$ and has a range which includes the interval $(\pi/(2\sqrt{\lambda}),\infty)$, we can conclude that Problem (P) possesses a solution *with one zero*, when

$$\frac{\pi}{\sqrt{\lambda}} < L.$$

Thus we have proved:

**Lemma 1** *Problem (P) has a nontrivial solution if*

$$\lambda > \mu_1 = \left(\frac{\pi}{L}\right)^2. \tag{3.3.10}$$

Of course we can now continue $\theta$ over a further distance $2\sigma_1$ to obtain a solution of the diffeential equation in (3.26), with vanishing slope at $s = 0$, on the interval $[0, 4\sigma_1]$. If we can find a $\gamma_2 \in (0,\pi)$ such that

$$4\sigma_1(\gamma_2) = L,$$

we have found a solution of Problem (P) *with two zeros*. Clearly this is possible if

$$\lambda > \mu_2 = \left(\frac{2\pi}{L}\right)^2.$$

Continuing in this fashion we find that there exists a solution *with n zeros* if

$$\lambda > \mu_n = \left(\frac{n\pi}{L}\right)^2, \quad n = 1, 2 \ldots.$$

**Exercise 8** Prove that if $\theta$ is a solution of Problem (P), then

$$\int_0^L \theta = 0. \tag{3.3.11}$$

Having shown in Lemma 1 that Problem (P) has a nontrivial solution if $\lambda > \mu_1$, we show in the next lemma that it has none if $0 < \lambda \leq \mu_1$.

**Lemma 2** *Problem (P) has no nontrivial solution if*

$$\lambda \leq \mu_1 = \Big(\frac{\pi}{L}\Big)^2.$$

**Exercise 9** Prove Lemma 2.

Hints. (a) Multiply the differential equation by $\theta$ and integrate over $(0, L)$ to obtain

$$\int (\theta')^2 = \lambda \int \theta \, \sin\theta < \lambda \int \theta^2, \tag{3.3.12}$$

(b) Use the variational characterisation of $\mu_1$ in (3.21) to show that

$$\int (\theta')^2 \geq \mu_1 \int \theta^2. \tag{3.3.13}$$

With Lemma 2 the question of existence of solutions is clear. We can also say something about uniqueness by means of the following monotonicity property.

**Lemma 3** *We have*

$$\gamma_1 < \gamma_2 \quad \Rightarrow \quad \sigma_1(\gamma_1) < \sigma_1(\gamma_2). \tag{3.3.14}$$

Lemma 3 implies that Problem (3.26) has precisely *one* solution $\theta_1(x)$ with $\theta_1(0) = \gamma_1 \in (0, \pi)$, provided that $L$ is big enough.

*Proof.* We assume to the contrary that the two graphs intersect when $\theta_1 > 0$, and thus that

$$S = \sup\{s > 0 : \theta_1 < \theta_2 \text{ on } [0, s)\} < \sigma_1(\gamma_1). \tag{3.3.15}$$

We now multiply the equation for $\theta_1$ by $\theta_2$ and the equation for $\theta_2$ by $\theta_1$, subtract, and integrate over $(0, S)$. This yields

$$(\theta_2\theta_1' - \theta_1\theta_2')\Big|_0^S + \lambda \int_0^S \Big(\frac{\sin\theta_1}{\theta_1} - \frac{\sin\theta_2}{\theta_2}\Big)\theta_1\theta_2 ds = 0. \tag{3.3.16}$$

At $s = S$, we have $\theta_1 = \theta_2$ and $\theta_1' > \theta_2'$. Hence the first term in (3.41) is positive. On $(0, S)$ , $\theta_1 < \theta_2$. Because the function $t^{-1}\sin t$ is strictly decreasing on $(0, \pi)$, this implies that the second term in (3.41) is also positive. We therefore have a contradiction, and (3.40) cannot be true.

The set of values of $\lambda$ and $\gamma$ for which there exists a solution is given by the monotone curves

$$\Gamma_n : \quad \lambda = \lambda_n = \Big(\frac{2n\sigma_1(\gamma)}{L}\Big)^2, \qquad 0 < \gamma < \pi, \quad n = 1, 2, \ldots. \tag{3.3.17}$$

These curves all *bifurcate* from the $\gamma = 0$ axis at the points $(0, \mu_n)$; the *bifurcation points*. Note that the solutions corresponding to $\Gamma_n$ have $k$ zeros.

It so happens that we can find an explicit expression for $\sigma_1(\gamma)$ in terms of the *complete elliptic integral of the first kind*:

$$K(k) = \int_0^{\pi/2} \frac{d\psi}{\sqrt{1 - k^2 \sin^2 \psi}}, \qquad 0 < k < 1. \tag{3.3.18}$$

Details about elliptic integrals can be found in [AS, p. 589].

**Lemma 4** *We have*

$$\sigma_1(\gamma) = K(\sin \frac{\gamma}{2}). \tag{3.3.19}$$

**Exercise 10** Draw the bifurcation diagram using the sine and the elliptic integral $K(k)$, and MAPLE or Mathematica.

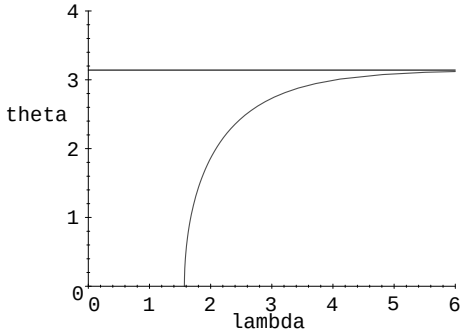


Fig. 3. The bifurcation diagram

*Proof.* We proceed along a series of exrecises.

**Exercise 11** Let $\theta$ be the solution of differential equation in (3.26) which satisfies

$$\theta(0) = \gamma \in (0, \pi) \quad \text{and} \quad \theta'(0) = 0$$

and let $\lambda = 1$. Prove then that

$$\theta'^2 = 2(\cos\theta - \cos\gamma).$$

Next we set

$$\sin \frac{\theta}{2} = k \sin\psi, \quad \text{where} \quad k = \sin \frac{\gamma}{2}.$$

**Exercise 12** Prove that $\psi$ satisfies the equation

$$\psi'^2 = 1 - k^2 \sin^2 \psi. \tag{3.3.20}$$

**Exercise 13** Deduce (3.44) from (3.45).

## 3.4 The shape of the rod

To conclude this investigation, it is necessary to return to the variables $u$ and $v$ which describe the shape of the rod. One can proceed as follows: (a) Write the differentail equation in (3.26) and (3.10) as the first order system for the variables $\theta(s)$, $\zeta(s) = \theta'(s)$, $x(s) = s + u(s)$ and $y(s) = v(s)$:

$$\begin{cases} \theta' &= \zeta \\ \zeta' &= -\lambda \sin\theta \\ x' &= \cos\theta \\ y' &= \sin\theta \end{cases} \tag{3.4.1}$$

and shoot with the initial values

$$\theta(0) = \gamma, \quad \zeta(0) = 0, \quad x(0) = 0, \quad y(0) = 0.$$

(b) Fix $\lambda$ and adjust $\gamma \in (0, \pi)$ so that $\theta'(1) = \zeta(1) = 0$.
(c) Draw the graph

$$\mathcal{C} = \{(x(s), y(s)) : 0 \le s \le 1\}.$$

**Exercise 14** Carry out the above programme for

$$\lambda = 10, 11, 12, 15, 20, 30, 50, 80, 120.$$

## 3.5 Energy

The total energy $\mathcal{E}$ stored in the strut consists of two parts: the bending energy $\mathcal{E}_B$ and the energy $\mathcal{E}_P$ used to oppose the thrust $P$ to bend the strut. Thus

$$\mathcal{E}_B(\theta) = \int_0^L \frac{dM}{ds}\frac{d\theta}{ds} ds = \frac{1}{2} EI \int_0^L \{\theta'(s)\}^2 \, ds. \tag{3.5.1}$$

and for $\mathcal{E}_P$ we find

$$\mathcal{E}_B(\theta) = P \int_0^L (1 - \cos\theta) \, ds. \tag{3.5.2}$$

Thus

$$\mathcal{E}(\theta) = EI \int_0^L \Big\{ \frac{1}{2}\{\theta'(s)\}^2 - \lambda(1 - \cos\theta) \Big\} \, ds. \tag{3.5.3}$$

**Exercise 15** Show that if $\lambda \le (\frac{\pi}{L})^2$, then $\mathcal{E}(\theta) \ge \mathcal{E}(0) = 0$ for any $\theta \in X_1$. Use *Poincare's inequality*

$$\int_0^L u^2(s) \, ds \le \Big(\frac{L}{\pi}\Big)^2 \int_0^L \{u'(s)\}^2 \, ds \quad \text{for} \quad u \in X_1. \tag{3.5.4}$$

**Exercise 16** Show that if $\lambda > (\frac{\pi}{L})^2$, then

$$\mathcal{E}^* = \inf\{\mathcal{E}(\theta) : \theta \in X_1\} < 0.$$

Use a test function $\varphi(x) = \varepsilon \sin(\frac{\pi x}{L})$, in which $\varepsilon$ is sufficiently small.

**Remark.** The buckling instability has an interesting analogue in liquid crystals, the so called *Fréedericks Transition*. In this case, the parameter $\lambda$ in the buckling problem is changed by an electric (or magnetic) field. This instability underlies the operation of liquid crystal displays.

## References

[AS] Abramowitz, M. and I.A. Stegun, *Handbook of mathematical functions*, Dover Publications, New York, 1965.

[K] Koiter, W.T., *Inleiding tot de leer van stijfheid en sterkte* Epsilon, Utrecht. 1985.

[L] Love, A., *The mathematical theory of elasticity*, pp 401-412, Cambridge at the University Press, 1959.

[R] Reiss. E.L., *Column buckling – an elementary example of bifurcation*, in *Bifurcation theory and nonlinear eigenvalue problems*, (Eds. J.B. Keller and S. Antman), pp 1-16, W.A. Benjamin, New York, 1969.

[S] Stakgold, I., *Branching of solutions of nonlinear equations*, SIAM Review **13** (1971) 289-332.

# Chapter 4

# Hopf Bifurcation

The Hopf bifurcation is a bifurcation of a branch of periodic solutions from a trivial state. It is of particular interest in nonlinear chemical systems, nonlinear optics and biology. In the first section we will study the bifurcation more qualitatively to get an intuitive understanding of the phenomena. For that we analyze the behavior of a forced oscillator, the Van der Pol oscillator. In the second section we will treat the bifurcation in more quantitative terms where we consider a system of chemical reactions, known as the Brusselator.

## 4.1 Van der Pol oscillator

### 4.1.1 Free oscillator

The simple pendulum is the oldest known example of the free oscillator. Its equation of motion is known to be:

$$\frac{d^2\theta}{dt^2} + \omega^2 \sin\theta = 0\,, \tag{4.1.1}$$

where $\theta$, defined on the whole real axis, measures the angle between the rod of length $l$ and the vertical, $\omega^2 = g/l$ measures the frequency of small oscillations which is given by the length $l$ and the gravitation $g$. To analyze the motion of the pendulum it is convenient to study the motion in its phase space $(\theta, \dot\theta)$. The energy $E$:

$$E(\theta, \dot\theta) = \frac{1}{2}\dot\theta^2 + \omega^2(1 - \cos\theta) \tag{4.1.2}$$

is conserved and hence a constant of motion:

$$\frac{dE}{dt} = \left(\frac{d^2\theta}{dt^2} + \omega^2 \sin\theta\right)\dot\theta = 0\,. \tag{4.1.3}$$

It follows that the trajectories in phase space are curves of constant energies.

**Exercise 1** Draw the phase portrait for a simple pendulum, i.e. lines of constant energy in the $(\theta, \dot\theta)$ plane. Draw a typical curve for the linear regime where oscillations are small so that $\sin\theta$ can be approximated by $\theta$. Draw a curve where this approximation does not hold anymore but where the pendulum still oscillates. Draw

a curve where the energy is so high that the pendulum does not oscillate but rotate. Draw the curve which separates the oscillatory behavior from the rotational. This line is called separatrix.

### 4.1.2 Damped oscillator

In practice however, the motion of a real pendulum stops eventually due to friction. Friction breaks the time reversal symmetry. By including a friction term in the equation of motion of a simple pendulum, we obtain:

$$\frac{d^2\theta}{dt^2} + \gamma\frac{d\theta}{dt} + \omega^2 \sin\theta = 0\,, \tag{4.1.4}$$

where $\gamma > 0$ is the friction coefficient. The energy is no longer conserved.

**Exercise 2** Show that

$$\frac{dE}{dt} = -\gamma\dot{\theta}^2\,. \tag{4.1.5}$$

Let $\gamma$ now be a parameter which can be positive as well as negative. From (4.1.5) it follows that the energy decreases if $\gamma > 0$ and increases if $\gamma < 0$. For $\gamma = 0$ we have seen that the energy was conserved and that the trajectories in phase space form a continuous family of closed curves about the equilibrium point $\theta = \dot{\theta} = 0$. If $\gamma$ is not zero, this family of curves is transformed into a set of trajectories that either converge to or diverge from the origin, depending whether $\gamma > 0$ or $\gamma < 0$. To show that these trajectories are spirals for weak damping, i.e. $|\gamma| \ll \omega$, we consider the following ansatz:

$$\theta(t) \simeq \rho(t)\cos(\omega t + \theta)\,, \tag{4.1.6}$$

where $\rho(t)$ is a function of time which varies slowly as compared to the period $T = 2\pi/\omega$. Let an overbar be an averaging over one period of the motion. The variation of the energy over one period can be evaluated in two different ways:

$$\overline{\frac{dE}{dt}} = \frac{1}{T}\int_{t_0}^{t_0+T} dt\frac{dE}{dt} = -\gamma\overline{\dot{\theta}^2}\,. \tag{4.1.7}$$

**Exercise 3** Show that using (4.1.6) the change of energy over one period is:

$$\overline{\frac{dE}{dt}} = -\frac{1}{2}\gamma\rho^2\omega^2\,. \tag{4.1.8}$$

**Exercise 4** Show that using (4.1.8) in (4.1.2) the energy is given by:

$$E = \frac{1}{2}\rho^2\omega^2\,, \tag{4.1.9}$$

and, that its mean variation over a period is:

$$\overline{\frac{dE}{dt}} = \frac{1}{2}\omega^2\overline{\frac{d\rho^2}{dt}}\,. \tag{4.1.10}$$

Combining (4.1.8) and (4.1.10) we see that over a period the dynamics of the average is given by:

$$\overline{\frac{d\rho^2}{dt}} = -\gamma\rho^2 \,. \tag{4.1.11}$$

**Exercise 5** For $\gamma$ very small compared to $\omega$, integrate this relation over time. Then derive the parametric equation of the trajectories in the $(\theta, \dot{\theta})$ plane. What kind of trajectory do you obtain? Draw their phase portrait near the origin for $\gamma < 0$, $\gamma = 0$ and, $\gamma > 0$.

### 4.1.3 Forced oscillator

If the friction coefficient is $\theta$ dependent:

$$\gamma = \epsilon[1 - \frac{\theta^2}{\theta_0^2}] \,, \tag{4.1.12}$$

with $\epsilon$ playing the role of a *control parameter*, we obtain the *Van der Pol equation*:

$$\frac{d^2\theta}{dt^2} + \epsilon(1-\theta^2)\frac{d\theta}{dt} + \theta = 0 \,. \tag{4.1.13}$$

$\gamma$ appears as an effective friction coefficient. In the linear regime, $\theta \ll 1$ and $\dot{\theta} \ll 1$, the Van der Pol oscillator behaves as the damped oscillator. Hence, the system's behavior changes from damped oscillations for $\epsilon < 0$ to an amplified oscillation if $\epsilon > 0$, as was discussed in exercise 5. In the nonlinear regime, when $\theta^2$ can no longer be neglected, the system's behavior is more complicated. Although we cannot solve the Van der Pol oscillator analytically, we can get a qualitative understanding of its behavior.

**Exercise 6** For $\epsilon > 0$ show that the phase space of initial conditions can be devided up in two regimes:

- $\dot{\theta}(0) \ll 1$ and $\theta^2(0) \ll 1$: The effective friction coefficient is negative, hence $\theta$ gets amplified and the trajectory spirals out.
- $\dot{\theta}(0) \gg 1$ and $\theta^2(0) \gg 1$: The effective friction coefficient is positive, hence $\theta$ is damped and the trajectory in phase space spirals in.

Intuitively, we can hypothesize that there exists a closed trajectory between these two extremes encircling the origin. To study the different possible trajectories for different $\epsilon$ and different initial conditions you can use the simulation package on the 3rd floor.

**Exercise 7** Show that the energy dissipation of the Van der Pol oscillator is given by:

$$\frac{dE}{dt} = \epsilon\dot{\theta}^2(1-\theta^2) \,. \tag{4.1.14}$$

**Exercise 8** Show that in the *limit cycle*, i.e. closed trajectories, the average of the energy dissipation over one period is given by:

$$\overline{\frac{dE}{dt}} = 0\,. \tag{4.1.15}$$

This just expresses that averaged over one cycle, the dissipation is cancelled by the energy pumped in.

**Exercise 9** Consider now small $\epsilon$. Argue why to first order in $\epsilon$ we can write:

$$0 = \epsilon \int_{t_0}^{t_0+T} dt\, \dot{\theta_0}^2 (1-\theta_0^2) + O(\epsilon^2)\,, \tag{4.1.16}$$

where $\theta_0$ and $\dot{\theta_0}$ are the expressions to zeroth order in $\epsilon$, i.e. no friction.

**Exercise 10** Show that "energy conservation" for zeroth order (no friction) orbits imply:

$$\theta_0^2 + \dot{\theta_0}^2 = \theta_{max}^2\,, \tag{4.1.17}$$

where $\theta_{max}$ is the maximal angle.

**Exercise 11** Show by using (4.1.16) and (4.1.17) that for $\epsilon \to 0$ the orbit is a circle with radius $r = 2$. Hint: Introduce $x = \theta/\theta_{max}$ as integration variable, and then substitute $x = \sin z$.

## 4.2 Brusselator (chemical reactions)

Chemical reactions can generate interesting patterns. Theoretically they can be studied by reaction-diffusion models. The Brusselator, introduced by Lefever, Nicolis, and Prigogine is a toy model which played an important role in the understanding of self-organization phenomena in nonlinear chemical systems. It is probably the simplest one to present the basic ingredients which lead to destabilize uniform steady states. Various theoretical methods and hypotheses have been tested on this model. It is based on the following chemical reactions:

$$\begin{aligned} A &\to X\,, \\ B + X &\to Y\,, \\ 2X + Y &\to 3X\,, \\ X &\to E\,, \end{aligned}$$

where the concentrations of the species $A$, $B$ and $E$ are maintained constant by connecting them to a reservoir and are thus the real control parameters of the system. For a stirred system which mixes so well that the system is homogenuous, the evolution of the system can be described by a set of kinetic equations for the concentrations of the active species $X$ and $Y$ which comprise the different reactions.

The kinetic equations for the reactions are given by[1]:

$$\dot{X} = f(X,Y) = A - (B+1)X + X^2Y\,, \tag{4.2.1}$$
$$\dot{Y} = g(X,Y) = BX - X^2Y\,, \tag{4.2.2}$$

where $A$ and $B$ are concentrations which are being fixed. The dot means time derivative. Traditionally $B$ is used as bifurcation parameter.

**Exercise 12** Find the uniform steady states of the system, i.e. $\dot{X} = \dot{Y} = 0$.

**Exercise 13** Draw the null clines $f(X,Y) = 0$ and $g(X,Y) = 0$ in phase space $(X,Y)$ and discuss possible trajectories.

### 4.2.1 Linear stability analysis

In order that the steady state is observable and hence a physically relevant state, it has to be *stable*. The stability of a state can be studied via a *linear stability analysis*. In a linear stability analysis the evolution of a small perturbation $x, y$ of the steady state $(X_0, Y_0)$ is analyzed to linear order. The evolution equation of the small perturbation is obtained by replacing $X = X_0 + x$ in (4.2.1) and $Y = Y_0 + y$ in (4.2.2).

**Exercise 14** Show that the evolution of a small perturbation is given by:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \bar{\bar{L}}_c \begin{pmatrix} x \\ y \end{pmatrix}\,, \tag{4.2.3}$$

with

$$\bar{\bar{L}}_c = \begin{pmatrix} (B-1) & A^2 \\ -B & -A^2 \end{pmatrix}\,. \tag{4.2.4}$$

**Exercise 15** Since equation (4.2.3) is linear, the ansatz:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} e^{st} \tag{4.2.5}$$

transforms equation (4.2.3) into an eigenvalue problem. Find the eigenvalues $s = \sigma + i\omega$ of equation (4.2.3).

Whether the reference state is stable or unstable is determined by the real part $\sigma$ of $s$ which is called *growth rate*:

- $\sigma < 0$: The perturbation decays and the steady state is *linearly stable*.
- $\sigma > 0$: The perturbation grows and the steady state is *unstable*.
- $\sigma = 0$: The steady state is called *marginal* or *neutral*.

[1] How to write down the kinetic equations for a system of reactions will be explained in the next session.

The point in parameter space where $\sigma$ changes from negative to positive is the *instability* or *bifurcation point*. A bifurcation point corresponds to situations in which small changes in the coefficients of the defining equations can lead to qualitatively different behavior of solutions. In this case the steady state changes from being stable to being unstable.

**Exercise 16** Find the bifurcation point $B_H$ as a function of $B$.

**Exercise 17** Show that the imaginary eigenvalues at the bifurcation point are not zero.

This is the characteristic for a *Hopf Bifurcation*. Since at the bifurcation point the unstable mode appears as an oscillatory mode we expect a transition from a uniform steady state to limit cycle oscillations.

## 4.3 Amplitude equation

What happens to the system beyond the instability?
Since the linear regime predicts an exponential growth, nonlinear contributions have to be considered. In the vicinity of the bifurcation point the new solution that appears beyond the bifurcation may be determined through a perturbation expansion of the kinetic equations in series of a small parameter $\epsilon$ which characterizes the distance to the threshold $B = B_H$. The method is referred to as *multiple scale expansion*.
Close to the instability $B \simeq B_H$ the concentrations $\mathbf{X}(t) = (X(t), Y(t))$ are expressed as expansions in the deviations from the uniform steady state $\mathbf{X_0}$:

$$\mathbf{X}(t) = \mathbf{X_0} + \epsilon \mathbf{X_1} + \epsilon^2 \mathbf{X_2} + \dots , \tag{4.3.1}$$

where the small parameter $\epsilon$ is related to the distance to the instability via the expansion:

$$B = B_H + \epsilon B_1 + \epsilon^2 B_2 + \dots , \tag{4.3.2}$$

The evolution of the variables takes place on large time scales. For the moment we do not anticipate a particular scaling behavior for the time scale and assume that:

$$\mathbf{X_1}(t) = A(\tau_1, \tau_2, \dots) e^{i\omega_0 \tau_0} + c.c. , \tag{4.3.3}$$

where the $\tau_i$ scale as $\tau_i = \epsilon^i t$. We may thus write the time derivative $d_t$ as

$$d_t = d_{\tau_0} + \epsilon d_{\tau_1} + \epsilon^2 d_{\tau_2} + \dots . \tag{4.3.4}$$

Substituting (4.3.1) to (4.3.4) in (4.2.1) and (4.2.2) and identifying the contributions at each order in $\epsilon$ leads to a set of linear equations that have to be solved iteratively.

**Exercise 18** Show that at the order $\epsilon^0$, one recovers the reference state $(X_0, Y_0)$.

**Exercise 19** Show that at the order $\epsilon^1$, the linear equation (4.2.3) of the linear stability analysis is recovered.

**Exercise 20** Determine the critical left $\vec{v}$ and right $\vec{u}$ eigenvector of $\bar{\bar{L}}_c$.

Thus, the solution $\mathbf{X_1}$ of (4.2.3) is a linear combination of the right eigenvectors of $\bar{\bar{L}}_c$:

$$\mathbf{X_1} = \begin{pmatrix} 1 \\ \frac{i-A}{A} \end{pmatrix} Z(t) \exp iAt + \begin{pmatrix} 1 \\ -\frac{i+A}{A} \end{pmatrix} Z^*(t) \exp -iAt\,. \tag{4.3.5}$$

Hence, we have to determine the amplitudes $Z(t)$ and $Z^*(t)$. To achieve this we have to consider higher order contributions in $\epsilon$.

**Exercise 21** Show that at the order $\epsilon^2$, we obtain the following linear equation:

$$(\partial_{\tau_0} - \bar{\bar{L}}_c) \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} I_2^x \\ I_2^y \end{pmatrix}\,, \tag{4.3.6}$$

where $\vec{I_2}$ is:

$$\begin{pmatrix} I_2^x \\ I_2^y \end{pmatrix} = \left[-\partial_{\tau_1} + \begin{pmatrix} B_1 & 0 \\ -B_1 & 0 \end{pmatrix}\right] \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \left(\frac{B_H}{A} x_1^2 + 2Ax_1y_1\right) \begin{pmatrix} 1 \\ -1 \end{pmatrix}\,, \tag{4.3.7}$$

or

$$\bar{\bar{L}}_0 \vec{u}_2 = \vec{I}_2\,. \tag{4.3.8}$$

$\bar{\bar{L}}_0$ is not invertible since it has a zero eigenvalue, and so the only possibility to have nontrivial solutions is for $I_2$ to be orthogonal to the kernel of $\bar{\bar{L}}_0^+$ or $< \vec{v}|\vec{I_2} >= 0$. This *solvability condition* is known as the *Fredholm alternative*[2]. In the present case, this condition is:

$$< \vec{v}|\vec{I_2} >= \frac{1}{2\pi} \int dt \left[(v_x^*, v_y^*) \exp -iAt \begin{pmatrix} I_2^x \\ I_2^y \end{pmatrix}\right]\,. \tag{4.3.9}$$

**Exercise 22** Show that

$$\int dt \exp i\omega t = 0 \qquad \text{for} \quad \omega \neq 0\,. \tag{4.3.10}$$

The only contribution to this scalar product comes from the so called resonant terms.

**Exercise 23** Show that from the solvability condition we obtain the following conditions:

$$B_1 = 0\,, \tag{4.3.11}$$

and

$$d_{\tau_1} Z = d_{\tau_1} Z^* = 0\,. \tag{4.3.12}$$

[2]This will be explained in more detail in chapter 9.

Note that in (4.3.3) we did not anticipate a particular scaling behavior for the time scale. Here, we see that the evolution of $Z$ occurs on a larger time scale than $\tau_1^{-1}$. In order to obtain the dynamics of $Z$ we have to go to the order $\epsilon^3$. Hence, without anticipating a particular time scale, the correct time scale drops out automatically. However, with a bit of experience the correct time scale can be "guessed" which simplifies the algebra a bit. To obtain this equation we need first to determine $\mathbf{X_2}$ which is given by (4.3.6) and has to have the form:

$$\mathbf{X_2} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} c_0 \\ d_0 \end{pmatrix} + \begin{pmatrix} c_1 \\ d_1 \end{pmatrix} \exp iAt + \begin{pmatrix} c_1^* \\ d_1^* \end{pmatrix} \exp -iAt \tag{4.3.13}$$

$$+ \begin{pmatrix} c_2 \\ d_2 \end{pmatrix} \exp i2At + \begin{pmatrix} c_2^* \\ d_2^* \end{pmatrix} \exp -i2At\,. \tag{4.3.14}$$

**Exercise 24** Show that:

$$c_0 = 0\,, \tag{4.3.15}$$

$$d_0 = \frac{2}{A^3}(A^2-1)ZZ^*\,, \tag{4.3.16}$$

$$c_2 = -\frac{2i}{3A^2}(1+iA)^2Z^2\,, \tag{4.3.17}$$

$$d_2 = \frac{1}{3A^3}(1+iA)^2(1+2iA)Z^2\,. \tag{4.3.18}$$

As will come clear in a few steps, $d_1$ and $c_1$ are not needed.

**Exercise 25** Show that at the order $\epsilon^3$, we obtain the following linear equation:

$$(\partial_{\tau_0} - \bar{\bar{L}}_c)\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} I_3^x \\ I_3^y \end{pmatrix}\,, \tag{4.3.19}$$

where $\vec{I}_3$ is:

$$\begin{pmatrix} I_3^x \\ I_3^y \end{pmatrix} = \left[-\partial_{\tau_2} + \begin{pmatrix} B_2 & 0 \\ -B_2 & 0 \end{pmatrix}\right]\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \tag{4.3.20}$$

$$+ \left(2\frac{B_H}{A}x_1x_2 + 2A(x_1y_2+x_2y_1) + x_1^2y_1\right)\begin{pmatrix} 1 \\ -1 \end{pmatrix}\,, \tag{4.3.21}$$

or

$$\bar{\bar{L}}_0\vec{u}_3 = \vec{I}_3\,. \tag{4.3.22}$$

**Exercise 26** Show that imposing the corresponding solvability condition:

$$< \vec{v}|\vec{I}_3 >= 0 \tag{4.3.23}$$

translates to:

$$I_3^{x+} + \frac{A(A-i)}{B_H}I_3^{y+} = 0\,, \tag{4.3.24}$$

$$I_3^{x-} + \frac{A(A-i)}{B_H}I_3^{y-} = 0\,, \tag{4.3.25}$$

which results in:

$$\partial_{\tau_2} Z = \frac{B_2}{2} Z - \left[ \frac{2 + A^2}{2A^2} + i \frac{4 - 7A^2 + 4A^4}{6A^3} \right] |Z|^2 Z \,. \tag{4.3.26}$$

$B_2$ is unknown, but if we do not consider higher orders, we will get:

$$\epsilon^2 B_2 = B - B_H \,, \tag{4.3.27}$$

$$\epsilon^2 \partial_{\tau_2} = \partial_t \,, \tag{4.3.28}$$

$$Z_H = \epsilon \sqrt{\frac{2 + A^2}{2A^2}} Z \,, \tag{4.3.29}$$

and the *amplitude equation* for $A_H$ takes the final form:

$$\partial_t Z_H = \mu Z_H - (1 + i\beta) |Z_H|^2 Z_H \,, \tag{4.3.30}$$

with

$$\mu = \frac{B - B_H}{2} \,, \tag{4.3.31}$$

$$\beta = \frac{4 - 7A^2 + 4A^4}{A(2 + A^2)} \,. \tag{4.3.32}$$

This equation is a generic amplitude equation of any system close to a Hopf instability.

### 4.3.1 Limit cycle

The amplitude equation describes the dynamics of the amplitude of the critical perturbation. There are two steady state solutions:

- $Z_H = 0$.
- $|Z_H|^2 = \mu$.

$\mu$ is the control parameter:

- $\mu < 0$: Only $Z_H = 0$ is a steady state solution.
- $\mu > 0$: Both solutions are steady state solutions.

As argued before, only a stability analysis determines whether a steady state is physically observable/relevant or not.

**Exercise 27** Perform a linear stability analysis of the two steady states. Show that for:

- $\mu < 0$, $Z_H = 0$ is a stable solution, and
- $\mu > 0$, $Z_H = 0$ is unstable but $|Z_H|_\pm = \pm\sqrt{\mu}$ is stable.

Draw a bifurcation diagram $Z_H$ vs $\mu$.

The bifurcation is called *normal* or *supercritical* since the nontrivial solution exists for $\mu > 0$, i.e., above the linear threshold.

**Exercise 28** $Z_H$ is the complex amplitude, hence it can be written as $Z_H = |Z_H| \exp i\theta$. Show that the evolution of modulus $|Z_H|$ and phase $\theta$ are given by:

$$\frac{d|Z_H|}{dt} = (\mu - |Z_H|^2)|Z_H| \,, \tag{4.3.33}$$

$$\frac{d\theta}{dt} = -\beta |Z_H|^2 \,. \tag{4.3.34}$$

The modulus is completely decoupled from the phase and can be solved independently. Once the dynamics of $Z_H$ is known, the evolution of the phase can be determined. In the asymptotic regime where $Z_H$ has reached its fixed point value the dynamics of $\theta$ reduces to:

$$\frac{d\theta}{dt} = \beta |Z_H|^2_{fix} \,, \tag{4.3.35}$$

which describes a periodic motion with a frequency $\beta$, which represents a *limit cycle*.

**Exercise 29** Draw the motion of $Z_H$ in the complex plane.

To conclude, in the first section we have seen in a qualitative discussion that an oscillator with an amplitude dependent friction coefficient branches from a trivial state to a limit cycle behavior. In the second section we have derived by means of a systematic expansion an amplitude equation which allowed a quantitative calculation of the frequency as well as the amplitude of the limit cycle in the vicinity of the bifurcation point.

## 4.4 Suggested further reading

We have based our presentation mostly on [Wal] — see in particular appendix 14.3.

*[Wal]* D. Walgraef, *Spatio-Temporal Pattern Formation*, Springer, New York, 1997

# Chapter 5

# Reaction diffusion equations

In many chemical processes, reactions take place in combination with diffusion. A typical situation is a substance that diffuses into a medium and reacts there with either the medium itself, or with another substance that is present in the medium. Medicines diffuse into tissue and combine there with biological substances, pollutants enter the soil and combine with substances present in the subsoil. One can think of many examples.

**Exercise 1** Think up a few examples.

The basic equation describing the evolution of the concentration $c$ of a process involving reaction and diffusion are [A, BE]:

- *Mass balance*

$$\frac{\partial c}{\partial t} + \text{div}\mathbf{J} = -R(c), \tag{5.0.1}$$

where $\mathbf{J}$ denotes the mass flux and $R(c)$ the rate at which the substance is spent or created by the chemical reaction.

- *Flow equation (or Constitutive equation)*

$$\mathbf{J} = -D(c)\ \text{grad}\ c, \tag{5.0.2}$$

where $D(c)$ denotes the diffusion coefficient of the medium. Combining equations (5.1) and (5.2) we obtain the nonlinear diffusion equation

$$\frac{\partial c}{\partial t} = \text{div}\{D(c)\text{grad}c\} - R(c). \tag{5.0.3}$$

In the following section we shall discuss different functions $R(c)$. Throughout we assume that $D$ is a positive constant.

## 5.1 Reaction kinetics

The particular form of the rate function $R(c)$ in equation (5.1) depends on the type of reaction involved. We give some examples.

(1) Suppose that the reaction is of the form

$$A + B \xrightarrow{k} C. \tag{5.1.1}$$

Then, as time progresses, the concentrations $a(t)$ and $b(t)$ of the substances $A$ and $B$ will decrease, whilst the concentration $c(t)$ of $C$ will increase. If mass action kinetics applies, the rate is proportional to the product of the concentrations of the reactants involved. The dynamics of the reaction is then described by

$$\frac{da}{dt} = -kab, \qquad \frac{db}{dt} = -kab \qquad \text{and} \qquad \frac{dc}{dt} = +kab, \tag{5.1.2}$$

where $k$ is a positive constant, called the *rate constant*.

If there is an excess amount of the chemical $B$, i.e. if $b \gg a$, then during the reaction the relative change of $b$ will be small. In that case we may replace $b$ by the initial concentration $b_0$, so that

$$R(a) = Ka, \quad \text{where} \quad K = k\, b_0. \tag{5.1.3}$$

This is called a first order reaction. To understand the reduction of the three rate laws for $a(t)$, $b(t)$ and $c(t)$ to a single rate law for $a(t)$, we consider the system (5.5) and denote the initial concentrations of $A$, $B$ and $C$ by respectively $a_0$, $b_0$ and $c_0$. We assume that $b_0 \gg a_0$. Plainly, to obtain $a$ and $b$ we only need the first two equations of (5.5). Once these are known, $c$ follows from the third equation of (5.5). We introduce dimensionless variables:

$$t^* = kb_0 t, \qquad x(t^*) = \frac{a(t)}{a_0}, \qquad y(t^*) = \frac{b(t)}{b_0}, \qquad \varepsilon = \frac{a_0}{b_0} \ll 1.$$

The equations for $a$ and $b$ then become

$$x' = -xy \quad \text{and} \quad y' = -\varepsilon xy \tag{5.1.4}$$

with initial values $x(0) = 1$ and $y(0) = 1$. For convenience we drop the asterisk again.

**Exercise 2** (a) Show that the solution orbit

$$\gamma_\varepsilon^+ = \{(x_\varepsilon(t), y_\varepsilon(t)) : 0 \le t < \infty\}$$

fills the line segment $\ell$ in the phase plane (the $xy$-plane):

$$\ell = \{(x, y) : 0 < x \le 1,\ y = 1 - \varepsilon(1 - x)\}.$$

(b) Show that

$$x_\varepsilon(t) = (1 - \varepsilon)\{\varepsilon + (1 - 2\varepsilon)e^{(1-\varepsilon)t}\}^{-1}.$$

(c) Show that for each $t > 0$,

$$\lim_{\varepsilon \to 0} x_\varepsilon(t) = e^{-t},$$

which corresponds to

$$a(t) = a_0 e^{-kb_0 t},$$

and hence to the rate law corresponding to (5.6).

If instead, we have the reaction

$$pA + B \xrightarrow{k} C, \tag{5.1.5}$$

then we obtain under the same assumptions

$$R(a) = Ka^p \tag{5.1.6}$$

and we speak of a $p^{th}$ order reaction. Although in this example, $p$ should be a positive integer, there exist reactions with rate functions like (5.9), in which $p$ is a positive constant, but not an integer.

**Exercise 3** Show that for the reversible reaction,

$$2A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C, \qquad k_1 > 0, \quad k_{-1} > 0, \tag{5.1.7}$$

the reaction equations becomes

$$\begin{cases} a' &= -2k_1a^2b + 2k_{-1}c, \\ b' &= -k_1a^2b + k_{-1}c, \\ c' &= +k_1a^2b - k_{-1}c. \end{cases} \tag{5.1.8}$$

**Remark.** In gase phase reactions, the correspondence between reaction laws, such as (5.4) and (5.10), and reaction rates as in (5.5) and (5.11) is quite direct and accurate. On the other hand, in liquids and solids the correspondence is often much less clear and it is better to think of the terms on the right hand side of (5.5) and (5.11) as effective empically determined reaction rates.

**Exercise 4** Suppose that the initial concentrations of $A$, $B$ and $C$ are given by respectively $a_0$, $b_0$ and $c_0$. Show that the concentrations $a(t)$, $b(t)$ and $c(t)$ converge to an equilibrium state $(a^*, b^*, c^*)$. Hint: Derive a differential equation for $a(t)$ alone.

(2) A very well known chemical reaction in enzyme kinetics is the *Michaelis-Menten reaction*

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SE \overset{k_2}{\longrightarrow} E + P, \tag{5.1.9}$$

in which $S$ denotes the substrate, $E$ the enzyme, $SE$ the enzyme-substrate complex, $P$ the product and $k_1$, $k_{-1}$ and $k_2$ are the reaction constants [R].

**Exercise 5** (a) Assuming mass action kinetics, show that the kinetic equations for the the Michaelis-Menten reaction are

$$\begin{cases} s' &= -k_1se - k_{-1}c, \\ e' &= -k_1se + (k_2 + k_{-1})c, \\ c' &= k_1se - (k_2 + k_{-1})c, \\ p' &= k_2c, \end{cases} \tag{5.1.10}$$

where $s(t)$, $e(t)$, $c(t)$ and $p(t)$ denote the concentrations of respectively $S$, $E$, $SE$ and $P$.

(b) Show that the total amount of enzyme is constant:

$$e(t) + c(t) = \text{constant} = e_1.$$

(c) Reduce the system (5.13) to the second order system:

$$\begin{array}{rl} s' & = -k_1 s e_1 + (k_1 s + k_{-1})c, \\ c' & = +k_1 s e_1 - (k_1 s + k_2 + k_{-1})c. \end{array} \tag{5.1.11}$$

When the total amount of enzyme $e_1$, (free enzyme, and enzyme that is bound in the complex $SE$), is small compared to the amount of substrate $s$, one can show that, except during a short time at the beginning of the reaction,

$$k_1 s e_1 - (k_1 s + k_2 + k_{-1})c \approx 0. \tag{5.1.12}$$

Using this in the expression for $s'$ we find the classical Michaelis-Menten rate function

$$R_1(s) = \frac{Vs}{K_1 + s}, \tag{5.1.13}$$

where $V = k_2 e_1$ and $K_1 = (k_{-1} + k_2)/k_1$ (see [P1]).

To understand the approximation (5.15), we introduce dimensionless variables:

$$t^* = k_1 e_1, \qquad x(t^*) = \frac{s(t)}{s_0}, \qquad y(t^*) = \frac{c(t)}{e_1},$$

and the constants

$$\nu = \frac{k_2}{k_1 s_0}, \qquad \kappa = \frac{k_{-1} + k_2}{k_1 s_0}, \qquad \varepsilon = \frac{e_1}{s_0}.$$

**Exercise 6** Show that in terms of the new variables, the equations (5.14) for $s$ and $c$ become

$$\left\{ \begin{array}{rcl} x' & = & -x + \{x + (\kappa - \nu)\}y, \\ \varepsilon y' & = & +x - (x + k_2 + \kappa)y, \end{array} \right. \tag{5.1.14}$$

where we have dropped the asterisks again. Follow the orbits in the phase plane numerically when $\varepsilon \ll 1$. For an analytical justification of (5.15) we refer to Part 2 of [P1].

(d) If there is *Substrate Inhibition*, that is the complex $SE$ is consumed by the substrate. Then we have we have to add to (5.12) a second reaction equation

$$S + SE \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} S^2 E, \tag{5.1.15}$$

in which $S^2E$ denotes a different complex. One can show that when the total amount of enzyme is small compared to the amount of substrate, the rate function $R$ becomes

$$R_2(s) = \frac{Vs}{K_1 + s + (s^2/K_3)}, \tag{5.1.16}$$

where $K_3 = k_{-3}/k_3$ and $V$ and $K_1$ are as before [R].

Note that

$$R_1(s) \to V \quad \text{and} \quad R_2(s) \to 0 \quad \text{as} \quad s \to \infty.$$

Suppose that the reaction takes place in a pellet. The questions one would then like to ask are for instance:

1. If the pellet is initially free from chemical and suddenly it is immersed in a bath in which the chemical is present. How fast does the chemical penetrate the pellet?
2. How does the concentration profile in the pellet evolve with time. Does it converge to a steady state?
3. What are the characteristics of the final concentration profile (if it exists)?
4. How deeply does the chemical penetrate into the pellet?

The answers to these questions will depend on the diffusion coefficient $D$, the the rate function $R(c)$ and the size of the pellet.

## 5.2 Evolution of concentration profiles in a pellet

For convenience we choose a one-dimensional geometry and assume that the reaction takes place in an infinite slab. We put the $x$-axis perpendicular to the parallel faces, and set $x = -L$ at one face and $x = L$ at the other. We assume that the concentration $c$ of the reactant only depends on the distance from the faces and on time, i.e. $c = c(x,t)$.

Thus, over a period of time $[0,T]$, the dynamics of the concentration profiles in the slab is governed by the equation

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} - R(c), \qquad Q_T = \{(x,t) : |x| < L,\ 0 < t < T\}, \tag{5.2.1}$$

in which $R$ is the rate function. We assume it to be smooth. We assume that at $t = 0$ the slab is immersed in a bath in which the concentration of the reactant is kept at a constant value $c_1$:

$$c(-L,t) = c_1 \quad \text{and} \quad c(+L,t) = c_1. \tag{5.2.2}$$

The initial concentration profile in the slab is given by

$$c(x,0) = c_0(x) \quad \text{for} \quad -L < x < L. \tag{5.2.3}$$

in which we assume that $c_0 \in C(-L,L)$.

**Definition.** We say that $c(x,t)$ is a solution of Problem (5.20)-(5.22) on $Q_T$ if

*(i)* $c_t$ and $c_{xx}$ are continuous in $Q_T$ and $c$ is continuous in $\overline{Q}_T$, possibly with the exception of the corner points $(x,t) = (\pm L, 0)$. Here subscripts denote partial differentiation.

*(ii)* $c$ satisfies (5.20), (5.21) and (5.22).

Throughout we shall assume that Problem (5.20)-(5.22) has a unique solution on $\overline{Q}_T$ for arbitrary large $T$.

In what follows we shall also need the notions of a super– and subsolution of Problem (5.20)-(5.22).

**Definition.** We say that $c(x,t)$ is a *supersolution* of Problem (5.20)-(5.22) if it has the regularity properties of a solution and of Problem (5.20)-(5.22) on $Q_T$ if

$$\begin{aligned} c_t - Dc_{xx} + R(c) &\geq 0 \qquad \text{in} \quad Q_T, \\ c(\pm L, t) &\geq c_1 \quad \text{for} \qquad 0 < t < T, \\ c(x,0) &\geq c_0(x) \quad \text{for} \quad -L < x < L. \end{aligned} \tag{5.2.4}$$

We say that $c(x,t)$ is a *subsolution* of Problem (5.20)-(5.22) if it has the regularity properties of a solution and satisfies (5.23) with the inequalities reversed.

**Comparison Principle.** *Let $\overline{c}$ be a supersolution and $\underline{c}$ a subsolution of Problem (5.20)-(5.22). Then*

$$\underline{c} \leq \overline{c} \quad \textit{in} \quad Q_T. \tag{5.2.5}$$

*If any of the inequalities in (5.23) are strict at some point in $Q_T$ or on the lateral boundary $\Gamma_T$ of $Q_T$, then the inequality in (5.24) is also strict.*

Recall that

$$\Gamma_T = \{(x,t) : x \in [-L, L],\ t = 0\} \cup \{(x,t) : x \in \{\pm L\},\ t \in (0,T]\}.$$

For the proof of the Comparison Principle we refer to Protter and Weinberger [PW], or to [P2].

As a first example we suppose that the reaction in the slab is a first order reaction, so that

$$R(c) = -k_1 c, \tag{5.2.6}$$

and that initially the slab contains no reactant, i.e.

$$c_0(x) = 0 \quad \text{for} \quad -L < x < L. \tag{5.2.7}$$

**Exercise 7** (a) Show that there exists a unique stationary concentration profile $c^*(x)$, which is given by

$$c^*(x) = c_1 \frac{\cosh(x\sqrt{(k_1/D)})}{\cosh(L\sqrt{(k_1/D)})}. \tag{5.2.8}$$

(b) Discuss the influence of the two constants $D$ and $k_1$ on the concentration profile $c^*$. Specifically, set

$$I_\alpha = \{x \in (-L, L) : c^*(x) \leq \alpha c_1\}, \qquad \alpha \in (0,1).$$

– Show that $I_\alpha$ is either the empty set, or a connected interval of the form $I_\alpha = [-a, a]$ where $a \in [0, L)$.

– Fix $\alpha \in (0,1)$, and write $D/k_1 = \varepsilon^2$. Show that when $\varepsilon > 0$ is small enough, then there exists a point $a(\varepsilon) \in [0, L)$ such that $I_\alpha$ is given by

$$I_\alpha = [-a(\varepsilon), a(\varepsilon)].$$

– Show that

$$a(\varepsilon) \sim L - \varepsilon \log\Big(\frac{1}{\alpha}\Big) \quad \text{as} \quad \varepsilon \to 0.$$

Thus, as $\varepsilon \to 0$, *Boundary Layers* develop near $x = -L$ and $x = L$.

Next we turn to the dynamics of the problem. Let $c(x,t)$ be the solution of Problem (5.20)-(5.22) with $R$ given by (5.25), and the initial concentration profile by (5.26).

**Exercise 8** (a)Show that

$$0 < c(x,t) < c^*(x) \quad \text{for} \quad x \in (-L,L), \quad t > 0. \tag{5.2.9}$$

Hint. Use the Comparison Principle.

(b) Show that at each $x_0 \in (-L,L)$ the concentration $c(x_0,t)$ increases monotonically to $c^*(x_0)$:

$$c(x_0,t) \ \nearrow \ c^*(x_0) \quad \text{as} \quad t \to \infty. \tag{5.2.10}$$

Hint. Use the Comparison Principle to prove that

$$c(x,t+h) > c(x,t) \quad \text{for} \quad x \in (-L,L), \quad t > 0, \ h > 0.$$

We conclude from (5.29) that the profile $c^*(x)$ is asymptotically stable from below.

(c) Show that $c^*(x)$ is asymptotically stable from above.

**Exercise 9** Suppose that the reaction in the slab follows the Michaelis-Menten kinetics

$$R(c) = \frac{Vc}{c+K}$$

and that, as in the previous exercise, the initial concentration profile $c_0$ is given by

$$c_0(x) = 0 \quad \text{for} \quad -L < x < L.$$

(a) Use the Comparison Principle to show that

$$0 < c(x,t) < c_1 \quad \text{for} \quad x \in (-L,L), \quad t > 0. \tag{5.2.11}$$

(b) Show that at each $x_0 \in (-L,L)$ the concentration $c(x_0,t)$ increases monotonically to a limit $c_-(x_0)$:

$$c(x_0,t) \ \nearrow \ c_-(x_0) \quad \text{as} \quad t \to \infty. \tag{5.2.12}$$

One can show that $c_-$ is a solution of the stationary problem:

$$c'' - D^{-1}R(c) = 0 \ \text{ on } \ (-L,L) \quad \text{and} \quad c(\pm L) = c_1. \tag{5.2.13}$$

Thus, this proves that Problem (5.32) has a solution.

(c) Now replace the initial condition by

$$c_0(x) = c_1 \quad \text{for} \quad x \in (-L,L).$$

Show then that the solution $\tilde{c}(x,t)$ of Problem (5.20)-(5.22) has the following properties:
– Bounds:

$$0 < c(x,t) < \tilde{c}(x,t) < c_1 \quad \text{for} \quad x \in (-L,L), \quad t > 0. \tag{5.2.14}$$

– At each $x_0 \in (-L,L)$ the concentration $\tilde{c}(x_0,t)$ decreases monotonically to a limit $c_+(x_0)$:

$$\tilde{c}(x_0,t) \ \searrow \ c_+(x_0) \quad \text{as} \quad t \to \infty. \tag{5.2.15}$$

As with the function $c_-$, one can show that $c_+$ is a stationary solution of Problem (5.20)-(5.22). It follows from (5.33) that

$$0 < c_-(x) \le c_+(x) < c_1 \quad \text{for} \quad x \in (-L,L).$$

(d) Prove that

$$c_-(x) = c_+(x) \quad \text{for} \quad x \in [-L,L].$$

Hint. Write $z_\pm = c_1 - c_\pm$. Then $z_+$ and $z_-$ are solutions of the problem

$$z'' + D^{-1}R(c_1 - z) = 0 \ \text{ on } \ (-L,L) \quad \text{and} \quad z(\pm L) = 0. \tag{5.2.16}$$

Multiply equation (5.35) for $z_-$ by $z_+$ and multiply equation (5.35) for $z_+$ by $z_-$, subtract, and integrate over $(-L,L)$ (See also the proof of Lemma 3 in Chapter 3.

**Exercise 10** Next we investigate a reaction with Michaelis-Menten kinetics in which the constant $K_1$ is very small, i.e. $k_{-1} \ll k_1$ and $k_2 \ll k_1$. This leads to the function

$$R_\varepsilon(c) = \frac{c}{\varepsilon + c}, \qquad 0 < \varepsilon \ll 1. \tag{5.2.17}$$

and we consider the problem

$$c'' - \lambda \frac{c}{\varepsilon + c} = 0 \quad \text{on} \quad (-1,1), \qquad c(\pm 1) = 1, \quad \lambda = D^{-1}. \tag{5.2.18}$$

We denote the unique solution of problem (5.37) by $c_\varepsilon$.
(a) Use the Comparison Principle to prove that

$$\varepsilon_1 < \varepsilon_2 \quad \Rightarrow \quad c_{\varepsilon_1}(x) < v_{\varepsilon_2}(x) \quad \text{for} \quad x \in (-1,1). \tag{5.2.19}$$

(b) Since $v_\varepsilon(x) > 0$ for all $x \in [-1,1]$ and all $\varepsilon > 0$, it follows from (5.38) that for every $x \in [-1,1]$.

$$\lim_{\varepsilon \to 0} c_\varepsilon(x) \text{ exists } = c^*(x). \tag{5.2.20}$$

(c) Find the solution $c^*$ of the problem

$$c'' - \lambda = 0 \quad \text{on} \quad (-1,1), \qquad c(\pm 1) = 1,$$

and determine for which values of $\lambda$ for which

$$c^*(x) \ge 0 \quad \text{for} \quad x \in [-1,1].$$

(d) Prove that.

$$c_\varepsilon(x) \ge c^*(x) \quad \text{for} \quad x \in [-1,1].$$

(e) Determine the limit profile $c^*(x)$ of $c_\varepsilon(x)$ as $\varepsilon \to 0$.

$$(i) \ \text{ when } \ \lambda < 6 \qquad (ii) \ \text{ when } \ \lambda > 6.$$

## 5.3 A bi-stable system

We consider an infinitely extended material which exchanges reactant with its environment at a rate $\rho$. We assume that $\rho$ is positive when the concentration $c$ in the slab lies below a reference value $q$, and negative when it lies above $q$:

$$\rho = \mu(q - c).$$

Here $\mu$ and $q$ are positive constants. If there is no reaction in the medium, the evolution of the concentration is described by the equation

$$c_t = c_{xx} + \mu(q - c), \qquad x \in \mathbf{R},\ t > 0, \tag{5.3.1}$$

where the diffusion coefficient has been absorbed in the spatial coordinate ($x' = x/\sqrt{D}$).

Suppose that the initial concentration profile is given by

$$c(c, 0) = c_0(x), \qquad x \in \mathbf{R}, \tag{5.3.2}$$

where $c_0 \in C(\mathbf{R})$,and that

$$0 \le c_0(x) \le M \quad \text{for} \quad x \in \mathbf{R},$$

for some constant $M$. Problem (5.40),(5.41) has a unique solution $c(x, t)$ on $S = \mathbf{R} \times [0, \infty)$, which is continuous in $\overline{S}$.

**Exercise 11** Show then that $c(x, t)$ satisfies

$$\|c(\cdot, t) - c_0\| = \max\{|c(x, t) - c_0(x)| : x \in \mathbf{R}\} \le Ke^{-\mu t},$$

where

$$K = \max\{q, M - q\}.$$

Next, we suppose that a reaction takes place in the material which follows that rate law

$$R(c) = \frac{c}{(1 + c)^2} \qquad \text{(cf. (5.19))}$$

Then

$$c_t = c_{xx} + f(c, q), \qquad x \in \mathbf{R},\ t > 0, \tag{5.3.3}$$

where

$$f(c, q) = \mu(q - c) - R(c).$$

We choose the feeding rate $\mu$ so small that $\partial f/\partial c$ has two sign changes. Let

$$\begin{array}{rcl} q_- & = & \inf\{q : f(\cdot, q) \ \text{ has three zeros}\}, \\ q_+ & = & \sup\{q : f(\cdot, q) \ \text{ has three zeros}\}. \end{array} \tag{5.3.4}$$

Then for $q_- < q < q_+$ the function $f(c, q)$ has three zeros: $\alpha < \beta < \gamma$ and

$$\begin{array}{rcccl} q & = & q_- & \Longrightarrow & \beta = \gamma \quad \text{and} \quad f(\cdot, q) < 0 \ \text{ on } \ (\alpha, \gamma), \\ q & = & q_+ & \Longrightarrow & \alpha = \beta \quad \text{and} \quad f(\cdot, q) > 0 \ \text{ on } \ (\alpha, \gamma). \end{array} \tag{5.3.5}$$

**Exercise 12** Show that the uniform stationary states $c = \alpha$ and $c = \gamma$ are asymptotically stable and that $c = \beta$ is unstable.

**Definition.** We say that a stationary solution $c^*$ is asymptotically stable if there exists an $\varepsilon_0 > 0$ such that if

$$\|c(\cdot, 0) - c^*\| < \varepsilon_0,$$

then

$$\|c(\cdot, t) - c^*\| \to 0 \quad \text{as} \quad t \to \infty.$$

Thus, if $q \in (q_-, q_+)$ then the reaction diffusion equation (5.42) describes a *Bi-Stable* system, and it is natural to look for Traveling Wave or *Fronts* with one stable state in front of it, and the other stable state behind it. They are of the form

$$c(x, t) = \varphi(x - vt),$$

in which $v$ denotes the speed whith which the wave moves through the material. Plainly the function $\varphi$ and the speed $v$ satisfy

$$\begin{cases} \varphi'' + v\varphi' + f(\varphi, q) = 0 \quad \text{on} \quad \mathbf{R} \\ \varphi(-\infty) = \alpha, \qquad \varphi(+\infty) = \gamma, \\ \varphi' > 0 \quad \text{on} \quad \mathbf{R}. \end{cases} \tag{5.3.6}$$

The existence and uniqueness (modulo translations) for the source function $f$ is well known [FM].

**Theorem.** *For every $q \in (q_-, q_+)$ there exists a unique travelling wave solution $(v, \varphi)$ of Problem (5.45).*

**Exercise 13** It is to be expected that when $q \in (q_-, q_+)$ is close to $q_-$, then fronts will move forward, so that at a given point, the concentration *drops* from the higher stable state ($c = \gamma$) to the lower stable state ($c = \alpha$), as the wave passes. On the other hand, if $q \in (q_-, q_+)$ is close to $q_+$, then, by the same argument, one expects the fronts to move backward.

Verify this *Conjecture.* Hint. Show that

$$\operatorname{sign} v = -\operatorname{sign} \int_\alpha^\gamma f(r, q)\, dr.$$


- [A] Aris, R., *The mathematical theory of diffusion and reaction in permeable catalysis*, Vol. 1 and 2, Clarendon Press, Oxford, 1975.

- [BE] Bebernes, J. and D. Eberly, *Mathematical problems from combustion theory*, Appl. Math. Sci. 83, Springer, New York, 1989.

- [F] Fife, P.C., *Mathematical aspects of reacting and diffusing systems*, Lecture Notes in Biomathematics **28**, Springer Verlag, New York, 1979.

- [FM] Fife, P.C. and J.B. McLeod, *The approach of solutions of nonlinear diffusion equations to travelling front solutions*, Arch. Rational Mech. Anal. **65** (1977) 335-361.

- [P1] Peletier, L.A., *Ordinary differential equations* Vol. 1 and 2, Lecture Notes Leiden University, 1990/91.

- [P2] Peletier, L.A., *Partial differential equations*, Lecture Notes Leiden University, 1996.

- [PW] Protter M.H, and H. Weinberger, *Maximum Principles in Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1967.

- [R] Roberts, D.V., *Enzyme kinetics*, Cambridge University Press, Cambridge, UK, 1977.

# Chapter 6

# The Turing Instability and Patterns in Reaction-Diffusion Systems

In order to understand the origin of *morphogenesis*[1], Alan Turing (1952) proposed a simple model of coupled chemical reactions in which starting from a homogeneous chemical mixture, when certain conditions are met, spontaneous spatially inhomogeneous patterns are formed. The mechanism which generates this stationary spatial structure is based on the interplay between reaction and diffusion in nonlinear chemical systems. This is rather surprising since one intuitively thinks that diffusion suppresses spatial inhomogeneities and does not favor them.

Chemical reactions with the following general properties may generate a Turing instability:

- An activator $X$ that participates in an autocatalytic reaction, where it enhances its own production or exerts a positive feedback.
- An inhibitor $Y$ that induces a slowing down of the previous activation process or exerts a negative feedback.
- The inhibitor has to diffuse more rapidly than the activator ($D_Y > D_X$).

The characteristic features of the Turing structures are:

- They occur as the result of chemical reaction and diffusion only, and do not involve convection.
- They are spatially inhomogeneous, steady in time and have an intrinsic wavelength that is independent of the dimension and the geometry of the system, provided the system is larger than at least a few wavelengths. This is quite different from convective instabilities.

[1] the formation of spatial structures in multicellular living organisms, starting from a homogeneous colony of identical cells.

Although Turing proposed the instability already in 1952, it took almost 40 years till Turing patterns for the first time could be observed experimentally in 1990 by a group in Bordeaux [Bor]. The Turing mechanism demands two conditions which are difficult to achieve experimentally. For one, convection has to be suppressed, and in addition the diffusion constant of the inhibitor has to be much larger than the diffusion constant of the activator. However, the diffusion constant is for most chemical systems in the same environment about the same. The Bordeaux group could solve this problem by designing a reactor in which the reaction takes place in a gel which suppresses mass transport via convection. In addition they used starch as an indicator which makes a large macromolecule complex with the activator and reduces its mobility considerably.

Though real biological systems are far too complex to be represented by simple models, Turing's ideas have been tested in a number of controlled chemical reactions involving a small number of key species. Nevertheless, there are many biological systems which show pattern formation, some are simple such as stripes, others are more complex. For example, genes essential for the segmentation of insects are activated in narrow stripes that surround the embryo in a beltlike manner. In monkeys, the nerves of the right and left eye project onto adjacent stripes in the cortex. We will concentrate here on simple models which elucidate the basic mechanism. They are also often the basic building blocks for more complex patterns which are often generated by hierarchical superposition of several pattern-forming systems. For more information see [Lan]. We will consider here two systems which are able to generate Turing patterns, one is the Brusselator which has been studied already in the context of the Hopf bifurcation, the other is an activator-inhibitor system which is a key building block for biological systems.

## 6.1 Example: The Brusselator - again!

We have already seen in chapter 4 that by tuning parameters (e.g. by changing concentrations of certain control chemicals in a chemical reaction) one can obtain temporal oscillations in the concentrations of different species. All the reactions were carried out in a CSTR (continuous flow stirred tank reactor) where constant stirring ensured spatial homogeneity at all times. Now, suppose we perform the experiment in a CFUR (continuous flow unstirred reactor) where the solution is not stirred continuously, then molecular diffusion becomes important. Accordingly, we must add *diffusion* terms to the set of equations 4.2.1. Hence we obtain the following reaction-diffusion system:

$$\dot{X} = D_X \nabla^2 X + A - (B+1)X + X^2 Y \,, \tag{6.1.1}$$

$$\dot{Y} = D_Y \nabla^2 Y + BX - X^2 Y \,, \tag{6.1.2}$$

with $D_X$ and $D_Y$ being the diffusion constants for $X$ and $Y$, respectively, and $A$ and $B$ are concentrations which are kept constant by coupling them to a reservoir. The dot refers to the time derivative.

**Exercise 1** Why is $X$ called an activator? How does $Y$ inhibit the selfactivation process of $X$?

## 6.2 Linear Stability Analysis

In chapter 4 we saw that the Brusselator, without the diffusion terms, had a uniform steady state ($X_0 = A, Y_0 = B/A$), which became unstable with respect towards small but *spatially uniform* fluctuations in the concentrations of $X$ or $Y$ when the control parameter $B$ was increased beyond the critical value $B_H = 1 + A^2$. The newly emerging steady state showed temporal oscillations in $X$ and $Y$. The diffusion terms do not change the uniform, steady state $(X_0, Y_0)$ which was calculated in chapter 4. However, with diffusion terms added we have the additional possibility that *spatially inhomogeneous* fluctuations destabilize the uniform steady state $(X_0, Y_0)$. In order to check under what conditions the spatially homogeneous state becomes unstable with respect to small spatially inhomogeneous fluctuations, we follow the same logic as was done in chapter 4 and perturb the steady state $(X_0, Y_0)$ with small perturbation $(x(\vec{r}, t), y(\vec{r}, t))$ so that

$$X(\vec{r}, t) = X_0 + x(\vec{r}, t)\,, \tag{6.2.1}$$
$$Y(\vec{r}, t) = Y_0 + y(\vec{r}, t)\,. \tag{6.2.2}$$

**Exercise 2** To study the stability of the uniform steady state to small perturbations, proceed along the same lines as in chapter 4 and write down the linear evolution equation for the perturbations $x, y$.
Show that the linear evolution matrix for a Fourier mode:

$$x(\vec{q}, s) = u_1 \exp\left(st + i\vec{q}\cdot\vec{r}\right), \tag{6.2.3}$$
$$y(\vec{q}, s) = u_2 \exp\left(st + i\vec{q}\cdot\vec{r}\right) \tag{6.2.4}$$

is now given by:

$$\bar{\bar{L}} = \begin{pmatrix} (B-1) - D_X q^2 & A^2 \\ -B & -A^2 - D_Y q^2 \end{pmatrix}. \tag{6.2.5}$$

**Exercise 3** Calculate the eigenvalues of the linear evolution equation. Note that now the eigenvalues are a function of the wavenumber $q$ of the mode considered. Show that they are given by the following *dispersion relation*

$$s_\pm = \frac{1}{2}[\Sigma \pm \sqrt{\Sigma^2 - 4\Delta}]\,, \tag{6.2.6}$$

where

$$\Sigma = B - 1 - A^2 - q^2(D_X + D_Y)\,,$$
$$\Delta = A^2 + q^2[A^2 D_X + (1-B)D_Y] + q^4(D_X D_Y)\,. \tag{6.2.7}$$

Thus, $s_\pm(q) = \sigma \pm i\omega$ are the growth rates of the two eigenvectors $u_\pm$ of $\bar{\bar{L}}$ for a given Fourier mode $q$. As discussed in chapter 4, the stability of the reference state is determined by the real part $\sigma$ of the growth rate:

- $\sigma < 0$: The perturbation decays and the steady state is linearly *stable*.
- $\sigma > 0$: The perturbation grows and the steady state is *unstable*.
- $\sigma = 0$: The steady state is *marginal* or *neutral* to linear perturbations.

There are two different mechanism through which the uniform steady state can loose its stability:

1. $\sigma > 0$ and $\omega = 0$, hence $s$ is purely real: Some modes are unstable and give rise to spatially inhomogeneous, stationary (Turing) patterns.
2. $\sigma > 0$ and $\omega \neq 0$, hence $s$ is complex: Some modes are unstable and due to their imaginary part give rise to temporal oscillations. The Hopf instability refers to the $q = 0$ mode being one of the unstable modes.

**Exercise 4** Draw the dispersion relation for the different cases listed above.

## 6.3 Turing instability versus Hopf bifurcation

In a generic reaction-diffusion system both Turing and Hopf instabilities are present. However, depending on the system parameters the two may be far from each other or close together so that they give rise to interesting oscillating spatial patterns.

To get a better understanding for the above mentioned possibilities we return to the example of the Brusselator. To make things simpler, we keep the parameters $A, D_X$ and $D_Y$ fixed and study how $s_\pm(q)$ vs $q$ changes if we increase the parameter $B$ from "low" values. In fact, we need to focus on $s_+(q)$ only since its real part is always larger than that of $s_-(q)$ and hence is the more relevant one for the linear analysis.

**Exercise 5** Show that the critical value of $B$ at which the Turing instability sets in for the Brusselator is given by:

$$B_T = (1 + A\eta)^2 \,, \tag{6.3.1}$$

with $\eta$ being the ratio of the diffusion constants:

$$\eta = \sqrt{\frac{D_X}{D_Y}} \,. \tag{6.3.2}$$

**Exercise 6** Show that the critical wave vector $q_c$ is given by:

$$q_c = \left( \frac{A^2}{D_x D_Y} \right)^{\frac{1}{4}} \,. \tag{6.3.3}$$

For the Brusselator we have seen that the critical value of $B$ at which the Hopf-bifurcation occurs is $B_H = 1 + A^2$.

**Exercise 7** Whether the Turing instability or the Hopf bifurcation appears first when we increase $B$ depends on the ratio $\eta = D_X/D_Y$. Show that the Turing instability occurs first if

$$\eta < \frac{\sqrt{A^2+1}-1}{A}. \tag{6.3.4}$$

Otherwise the Hopf bifurcation sets in first.

**Exercise 8** Show that if $D_X = D_Y$, i.e. $\eta = 1$ the Hopf bifurcation sets in before the Turing instability can occur.

Note that for fixed $A$, the two instabilities are well separated in parameter space, when the diffusion constants are much different $D_X \ll D_Y$, which is one of the requirements to observe Turing patterns.
On the other hand, if $\sqrt{D_X/D_Y} \simeq \frac{\sqrt{A^2+1}-1}{A}$ then the two bifurcations are close together, leading to a combination of spatial patterns superposed with temporal oscillations.

## 6.4 Amplitude equation

The linear stability analysis around the uniform steady state (reference sate) determined the critical value $B_T$ of the bifurcation parameter beyond which Fourier modes of wavenumber $q_c$ grow exponentially and hence destabilize the reference state.

In the vicinity of an instability point, a *separation of time and length* scales occurs. This is due to the fact that the real parts of the eigenvalues of the unstable modes tend to zero at the bifurcation point, whereas they remain finite in case of stable modes. Hence the evolution of the unstable modes is much slower than the evolution of the stable ones. As a result, the faster evolution of the stable modes usually allows them to relax on the evolution of the unstable ones and may be adiabatically eliminated from the dynamics. Hence, close to instability points, the dynamics may be reduced to the unstable or slow modes dynamics. This reduction leads to a dramatic simplification of the dynamics but nevertheless captures the asymptotic properties of the system's evolution.

The reduction of the dynamics can be achieved by *multiple scale analysis* which leads to evolution equations for the amplitudes of the unstable modes which is of the Ginzburg-Landau type where the unstable modes play the role of the order parameter. This procedure was already employed to derive an amplitude equation for the Hopf bifurcation. These equations allow to analyze the selection and stability of the spatial patterns that develop beyond the instability.

Each possible planform may be characterized by $m$ pairs of wave vectors $(\vec{q}_i, -\vec{q}_i)$. We follow here the same procedure as was introduced to study the Hopf bifurcation. However, this time one has also to include spatial dependencies due to diffusion terms.

Close to the instability $B \simeq B_c$ the concentrations $\mathbf{X}(\vec{r},t) = (X(\vec{r},t), Y(\vec{r},t))$ are expressed as expansions in the derivations from the uniform steady state $\mathbf{X_0}$:

$$\mathbf{X}(\vec{r},t) = \mathbf{X_0} + \epsilon \mathbf{X_1} + \epsilon^2 \mathbf{X_2} + \dots, \tag{6.4.1}$$

where the small parameter $\epsilon$ is related to the distance to the instability via the expansion:

$$B = B_H + \epsilon B_1 + \epsilon^2 B_2 + \dots . \tag{6.4.2}$$

We make again use of the fact that close to the bifurcation point a separation of length and time scales occurs. Hence we express the $\mathbf{X_i}$ as:

$$\mathbf{X_1}(\vec{r}, t) = \mathbf{A}(\vec{r}_1, \vec{r}_2, .., \tau_1, \tau_2) e^{i(\omega_c \tau_0 + \vec{q}_c \cdot \vec{x}_0)} + c.c. . \tag{6.4.3}$$

Due to the product rule, one may thus write the time derivative $\partial_t$ as

$$\partial_t = \partial_{\tau_0} + \epsilon \partial_{\tau_1} + \epsilon^2 \partial_{\tau_2} + \dots . \tag{6.4.4}$$

Similarly, the spatial differential operator may be written as:

$$\partial_x = \partial_{x_0} + \epsilon \partial_{x_1} + \epsilon^2 \partial_{x_2} + \dots , \tag{6.4.5}$$

$$\partial_y = \partial_{y_0} + \epsilon \partial_{y_1} + \epsilon^2 \partial_{y_2} + \dots . \tag{6.4.6}$$

By introducing these expressions in the nonlinear equations for the $X$ and $Y$, and identifying the contribution in each order of $\epsilon$ leads to a set of linear equations that have to be solved iteratively. The zeroth order term recovers the reference state $(X_0, Y_0)$ as was already worked out for the Hopf bifurcation. The first order term recovers the linear equation of the linear stability analysis (6.2.5).

**Exercise 9** Show that the right eigenvector associated with the + eigenvalue is given by:

$$\vec{u}(\vec{q}_c) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \exp\left(i\vec{q}_c \cdot \vec{r}\right) = \begin{pmatrix} 1 \\ -\frac{\eta}{A}(1 + A\eta) \end{pmatrix} \exp\left(i\vec{q}_c \cdot \vec{r}\right), \tag{6.4.7}$$

and the left eigenvector by:

$$\vec{v}(\vec{q}_c) = (v_1, v_2) \exp\left(i\vec{q}_c \cdot \vec{r}\right) = \left(1, \frac{A\eta}{1 + A\eta}\right) \exp\left(i\vec{q}_c \cdot \vec{r}\right). \tag{6.4.8}$$

$\mathbf{X_1}$ is a linear combination of the unstable modes. In order to determine the active modes in spatially extended systems, one has to take into account two types of degeneracy:

1. rotational degeneracy, and
2. for $B > B_T$ a whole band of unstable modes.

At $B = B_T$ one has to consider only the rotational degeneracy. Since the concentrations of the chemical species are real, active modes have to appear by pairs $(\vec{q}_i, -\vec{q}_i)$ with $|\vec{q}_i| = q_c$. Different patterns are possible:

- stripes and walls:

$$\mathbf{X_1} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \left[Z \exp\left(iq_c x\right) + Z^* \exp\left(-iq_c x\right)\right], \tag{6.4.9}$$

  where $\vec{q}_c$ lies parallel to the $x$-direction.

- squares (bimodal structures):

$$\mathbf{X_1} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} [Z_1 \exp(i\vec{q}_c \cdot \vec{x}) + Z_2 \exp(i\vec{q}_c \cdot \vec{y}) + c.c.] , \tag{6.4.10}$$

- hexagonal structures:

$$\mathbf{X_1} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} [Z_1 \exp(i\vec{q}_1 \cdot \vec{r}) + Z_2 \exp(i\vec{q}_2 \cdot \vec{r}) + Z_3 \exp(i\vec{q}_3 \cdot \vec{r}) + c.c.] , \tag{6.4.11}$$

with $|\vec{q}_1| = |\vec{q}_2| = |\vec{q}_3| = q_c$ and $\vec{q}_1 + \vec{q}_2 + \vec{q}_3 = 2n\pi$ .

The evolution of stripe patterns in a bidimensional system reduces to a *one-dimensional problem* with $\mathbf{X_1}$ given by (6.4.9). To determine $Z(t)$ one has to go to higher orders in $\epsilon$. The procedure is analogous to the derivation of the amplitude equation of the Hopf bifurcation. However, now we have to include also spatial dependencies. We are going to analyze the stripe pattern which is the simplest case.

**Exercise 10** Show that at the order $\epsilon^2$, one obtains the following linear equation:

$$\bar{\bar{L}}_0 \vec{u}_2 = \vec{I}_2 , \tag{6.4.12}$$

where $\vec{I}_2$ is:

$$\begin{aligned} \begin{pmatrix} I_2^x \\ I_2^y \end{pmatrix} &= \left[ -\partial_{\tau_1} + \begin{pmatrix} B_1 + 2D_X \nabla_0 \nabla_1 & 0 \\ -B_1 & 2D_Y \nabla_0 \nabla_1 \end{pmatrix} \right] \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \\ &+ \left( \frac{B_T}{A} x_1^2 + 2A x_1 y_1 \right) \begin{pmatrix} 1 \\ -1 \end{pmatrix} . \end{aligned} \tag{6.4.13}$$

**Exercise 11** Show that from the solvability condition:

$$< \vec{v} | \vec{I}_2 >= 0 , \tag{6.4.14}$$

the following constraints follow:

$$B_1 = 0 , \tag{6.4.15}$$

$$\partial_{\tau_1} Z = \partial_{\tau_1} Z^* = 0 . \tag{6.4.16}$$

**Exercise 12** $\mathbf{X_2}$ is determined by (6.4.7). Show that

$$\mathbf{X_2} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \exp(iq_c x) + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \exp(i2q_c x) + c.c. , \tag{6.4.17}$$

with

$$a_0 = 0 , \tag{6.4.18}$$

$$b_0 = -\frac{2}{A^3}(1 - A^2\eta^2) Z Z^* , \tag{6.4.19}$$

$$a_1 + \frac{A}{\eta(1 + A\eta)} b_1 = -\frac{2iq_c \sqrt{D_X D_Y}}{A(1 + A\eta)} \nabla_1 Z , \tag{6.4.20}$$

$$a_2 = \frac{4}{9} \frac{(1 - A^2\eta^2)}{A^2\eta} Z^2 , \tag{6.4.21}$$

$$b_2 = -\frac{1}{9} \frac{(1 - A^2\eta^2)(1 + 4A\eta)}{A^3} Z^2 . \tag{6.4.22}$$

Here, we see that the evolution of $Z$ occurs on a larger time scale than $\tau_1^{-1}$. As has been discussed in chapter 4, we obtain the correct time scale, although we did not anticipate a particular scaling behavior for the time and the length scale. This time we could have guessed already the correct time scale. To obtain the evolution of $Z$ we have to go to order $\epsilon^3$.

**Exercise 13** Show that at order $\epsilon^3$, one obtains the following evolution equation:

$$\bar{\bar{L}}_0 \vec{u}_3 = \vec{I}_3 \,, \tag{6.4.23}$$

where $\vec{I}_3$ is:

$$\begin{aligned} \begin{pmatrix} I_3^x \\ I_3^y \end{pmatrix} &= \left[ -\partial_{\tau_2} + \begin{pmatrix} B_2 + 2D_X\nabla_1^2 & 0 \\ -B_2 & 2D_Y\nabla_1^2 \end{pmatrix} \right] \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \\ &+ \begin{pmatrix} 2D_X\nabla_0\nabla_1 & 0 \\ 0 & 2D_Y\nabla_0\nabla_1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \\ &+ \left( 2\frac{B_T}{A} x_1 x_2 + 2A(x_1 y_2 + x_2 y_1) + x_1^2 y_1 \right) \begin{pmatrix} 1 \\ -1 \end{pmatrix} . \end{aligned} \tag{6.4.24}$$

**Exercise 14** Show that from the solvability condition:

$$< \vec{v} | \vec{I}_3 >= 0 \,, \tag{6.4.25}$$

the following amplitude equation follows:

$$\begin{aligned} (1-\eta^2)\frac{\partial Z}{\partial \tau_2} &= \frac{B_2}{1+A\eta} Z + \frac{4D_X}{1+A\eta}\nabla_1^2 Z \\ &- \frac{1}{9A^3\eta}(-8A^3\eta^3 + 5A^2\eta^2 + 38A\eta - 8)|Z|^2 Z \,. \end{aligned} \tag{6.4.26}$$

**Exercise 15** Discarding higher order, show that one has:

$$\epsilon^2 B_2 = B - B_T \,, \tag{6.4.27}$$

$$\epsilon^2 \partial_{\tau_2} = \partial_t \,, \tag{6.4.28}$$

$$\epsilon \nabla_1 = \nabla \,, \tag{6.4.29}$$

$$A = \epsilon Z \,, \tag{6.4.30}$$

and that the amplitude equation for $A$ has the final form:

$$\partial_t A = \mu A - g|A|^2 A + D\nabla^2 A \,, \tag{6.4.31}$$

with

$$\mu = \frac{1+A\eta}{1-\eta^2}\left(\frac{B-B_T}{B_T}\right) , \tag{6.4.32}$$

$$g = \frac{-8A^3\eta^3 + 5A^2\eta^2 + 38A\eta - 8}{9A^3\eta(1-\eta^2)} \,, \tag{6.4.33}$$

$$D = \frac{4D_X(1+A\eta)}{B_T(1-\eta^2)} \,. \tag{6.4.34}$$

**Exercise 16** What are the fixed point solutions of the amplitude equation (6.4.31)?

**Exercise 17** Explain why $\mu$ plays the role of a control parameter.
How does the solution change from $\mu < 0$ to $\mu > 0$ ?

The nontrivial solution exists for $\mu > 0$ and is stable if $g > 0$. Hence the leading order nonlinearity saturates the instability and the bifurcation is said to be *supercritical*. If however $g < 0$, the nontrivial solution exists already for $\mu < 0$ but is unstable. The bifurcation is said to be *subcritical*. In this case higher order terms neglected in (6.4.31) have to be taken into account in order to explain the saturation which is observed at large amplitudes.

**Exercise 18** Draw the bifurcation diagram $A$ vs $\mu$ for the supercritical and subcritical bifurcation.

Similar amplitude equations can be derived for other structures, such as the squares and hexagonal structures mentioned above. For squares starting with (6.4.10) one obtains after some algebra in the same spirit as for the stripe structure derived above the following amplitude equation:

$$\frac{\partial A_i}{\partial t} = \mu A_i - g|A_i|^2 A_i - h(\theta)|A_j|^2 A_i + D\nabla_i^2 A_i \,, \tag{6.4.35}$$

with $(i, j = 1, 2)$, $\theta$ being the angle between the two wavevectors $\vec{k_1}, \vec{k_2}$, which for squares is $\theta = \frac{\pi}{2}$, and where $\mu$, $g$ and $D$ are as given above and

$$\begin{aligned} h(\theta) &= \frac{2(2+A\eta)}{A^2(1-\eta^2)} - \frac{8(1-A\eta)}{A^2(1-\eta^2)}\left[\frac{2(1+A\eta-A^2\eta^2)}{A\eta(1-4\cos^2\theta)^2} - \frac{(1+4\cos^2\theta)}{(1-4\cos^2\theta)^2}\right], \\ \nabla_i &= \left(\frac{\partial}{\partial_{x_i}} + \frac{1}{2iq_c}\frac{\partial^2}{\partial^2_{x_j}}\right). \end{aligned}$$

The linear and nonlinear terms play different roles in the selection mechanisms of spatial patterns. The linear terms determine the possible wave vectors that may give rise to pattern formation and, sometimes, select their orientation, while the nonlinear terms saturate the growth of the linearly unstable modes and act as a selection mechanism by reflecting the symmetries of the patterns that may stabilize.

To see that, we consider the amplitude equation of the squared structure in more detail. We leave out the diffusion term which is for the time being not important and obtain a coupled amplitude equation:

$$\dot{A}_1 = \mu A_1 - g|A_1|^2 A_1 - h_{\pi/2} g|A_2|^2 A_1 \,, \tag{6.4.36}$$

$$\dot{A}_2 = \mu A_2 - g|A_2|^2 A_2 - h_{\pi/2} g|A_1|^2 A_2 \,. \tag{6.4.37}$$

The nonlinear terms represent self ($g$) and cross ($gh_{\pi/2}$) interactions between the excited modes.

**Exercise 19** Show that when $g > 0$ and $h_{\pi/2} > 0$, four types of steady state solutions are possible:

1. The trivial reference state: $A_1^0 = A_2^0 = 0$.
2. Rolls or walls perpendicular to $\vec{k_1}$: $A_1^0 = \sqrt{\frac{\mu}{g}}, A_2^0 = 0$.
3. Rolls or walls perpendicular to $\vec{k_2}$: $A_1^0 = 0, A_2^0 = \sqrt{\frac{\mu}{g}}$.
4. Squares: $A_1^0 = A_2^0 = \sqrt{\frac{\mu}{g(1+h_{\pi/2})}}$.

**Exercise 20** Show that for $h_{\pi/2} > 1$ the roll structure is stable and the squares are unstable, whereas for $h_{\pi/2} < 1$ the squares are stable but the rolls are unstable. This can be shown easily by performing a linear stability analysis of the roll structure, for example $A_1\sqrt{\mu/g}, A_2 = 0$. Show that the linear evolution equations of the perturbation $a_1$ and $a_2$ ($A_i = A_i^0 + a_i$) are in this case given by:

$$\dot{a_1} = -2\mu a_1 \,, \tag{6.4.38}$$

$$\dot{a_2} = (1 - h_{\pi/2})\mu a_2 \,. \tag{6.4.39}$$

Hence, the stability of the roll structure is determined by the sign of the term $(1 - h_{\pi/2})$.

The amplitude equation for hexagonal structures are given in [Wal]. There it is also shown that the hexagons are the first to appear subcritically when the bifurcation parameter $\mu$ is increased. If $h(\theta) > 1$ and $\mu$ is increased further, the hexagons become unstable and a first-orderlike transition occurs to roll structures. If $\mu$ is now decreased, a roll-to-hexagon transition occurs, but with hysteresis. Similar transitions occur between hexagons and squares when $h(\theta) < 1$.

## 6.5 Suggested further reading

*[Tur]* A. Turing, Phil. Trans. of the Royal Soc. London, **237B**, 37 (1952)

*[Bor]* V. Castets, E. Dulos, J.Boissonade and P. DeKepper, Phys. Rev. Lett. **64**, 2953 (1990)

*[Lan]* Koch and Meinhardt, Rev. Mod. Phys. **66**, 1481 (1994)

*[Ber]* P. Bergé, Y. Pomeau, Ch. Vidal, *Order within Chaos*, Hermann, Paris, France, 1984

*[Wal]* D. Walgraef, *Spatio-Temporal Pattern Formation*, Springer, New York, 1997

# Chapter 7

# The Navier Stokes equations of Fluid Dynamics

In this and the next chapter, we will investigate the hydrodynamic equations for a fluid. In this chapter, we introduce the equations and explore some of its applications to simple flow geometries, in the next chapter we will use the equations to analyze the so-called Rayleigh-Bénard instability, as an example of a pattern forming instability in convection.

## 7.1 Derivation of the Navier-Stokes equation

Of course, when we look on a very small scale (of order Å), a fluid consists of atoms and molecules, that move in a seemingly random fashion. Nevertheless, when there is an overall fluid flow, the average value of the velocity of each molecule is nonzero. To describe hydrodynamic behavior on large length and time scales (of order centimeters and seconds, say), we are not interested in this motion of the individual molecules and atoms on a rapid time scale, but in their gross "average" behavior. In a hydrodynamic description, we focus right away on these "hydrodynamic scales" by introducing continuous fields that describe this putative average and slow behavior on lengthscales which are much longer than the microscopic scales but much smaller than most macroscopic scales of interest to us.

*Remark* The strict meaning of the word "hydrodynamic equations" is really the equations for a fluid flow, in particular the so-called Navier Stokes equations. However, in the last quarter of this century, physicists have come to realize that quite often the behavior on a mesoscopic scale (hundreds of Å to micrometers, say) is often well described by an effective field theory, and the equations describing the behavior of a system on this scale is sometimes referred to more generally as a "hydrodynamic description". The insight that the behavior on a mesoscopic scale is often quite independent on the precise microscopic details of the physical system, and that it can be described by an effective field theory (even if fluctuations are important), essentially dates back to Landau. The importance of fluctuations was only realized later, however, and the conceptual underpinning of this idea came only in the seventies

as a result of the Renormalization Group theory of critical phenomena[1].

We now turn to the Navier-Stokes equations for a "simple" one-component fluid, i.e., a fluid that consists of one type of molecules that somehow don't have special features. E.g., a fluid of polymer molecules won't do, as these may have strange visco-elastic properties, but just water or a simple oil is okay. The hydrodynamic fields are associated with the quantities which microscopically are *conserved*, i.e., conserved in isolated collisions of the constituent molecules. This implies that if we imagine the fluid to be divided up into small boxes (but which each are large enough to contain many molecules, so that the fluctuations in their number are small), the change in the hydrodynamic variables can only be due to what flows in or out at the walls of these little hypothetical boxes. Associated with each conserved variable, we then expect one hydrodynamic field.

**Exercise 1** What are the three conserved quantities for a simple one-component fluid.

**Exercise 2** If you count each component of a vector as one variable, how many hydrodynamic fields do you expect for a one-component fluid?

**Exercise 3** Argue, along the lines of the previous two exercises, that you will need to describe a simple two-component fluid (e.g. a mixture of water and alcohol) with six hydrodynamic equations.

According to the above considerations, to arrive at the hydrodynamical equations, we introduce the field $\rho$ for the mass density of the fluid, the velocity field $\mathbf{v}$ (throughout, we use boldface to indicate vectors) and a variable for the energy density (we shall actually use the temperature $T$ below — different variables can be related through thermodynamic relations).

**Exercise 4** Consider a small cube of size $V = \Delta x \Delta y \Delta z$ in a Cartesian coordinate system. Argue that the change in the total mass inside this cube equals $V \partial\rho/\partial t$, and that the mass flux through the surface of the cube equals

$$\begin{aligned}
&\Delta y \Delta z \left(\rho(x,y,z,t)v_x(x,y,z,t) - \rho(x+\Delta x,y,z,t)v_x(x+\Delta x,y,z,t)\right)\\
&\Delta x \Delta z \left(\rho(x,y,z,t)v_y(x,y,z,t) - \rho(x,y+\Delta y,z,t)v_y(x,y+\Delta y,z,t)\right)\\
&\Delta x \Delta y \left(\rho(x,y,z,t)v_z(x,y,z,t) - \rho(x,y,z+\Delta z,t)v_z(x,y,z+\Delta z,t)\right)
\end{aligned} \tag{7.1.1}$$

Throughout this chapter, $v_x$, $v_y$ and $v_z$ are the $x$, $y$, and $z$ components of the velocity field. Sometimes, we shall also indicate the cartesian coordinates by $x_i$ and the velocity components by $v_i$.

**Exercise 5** Combine the results of the previous exercise to show that in the limit $V \to 0$, mass conservation is expressed by

$$\partial_t \rho(\mathbf{r},t) + \nabla \cdot \rho(\mathbf{r},t)\mathbf{v}(\mathbf{r},t) = 0 \ , \tag{7.1.2}$$

where, as usual, $\nabla \cdot \mathbf{a} = \sum_i \partial_{x_i} a_i$.

[1]For a more extensive discussion along these lines, as well as an introduction to some of the main ideas of the Renormalization Group, see e.g. the article by M. E. Fisher, Rev. Mod. Phys. **70**, 653 (1998).

Next, we consider the equation for the momentum density $\rho\mathbf{v}$ of the fluid.

**Exercise 6** Argue along similar lines as above in exercises 4 and 5 that in the continuum limit momentum conservation leads to the result

$$\partial_t \rho\mathbf{v} + \nabla\cdot(\rho\mathbf{v}\mathbf{v}) = -\nabla\cdot\sigma + \mathbf{f}_{ext} \ . \tag{7.1.3}$$

Here $\sigma$ is what is usually referred to as the *stress tensor* whose components $\sigma_{ij}$ give the force per unit area in the direction $j$ on the surface with normal in the $i^{th}$ direction. Thus, $\nabla\cdot\sigma$ is a vector whose $j^{th}$ component is $\sum_i \partial_{x_i}\sigma_{ij}$, and likewise the $j^{th}$ component of $\nabla\cdot(\rho\mathbf{v}\mathbf{v})$ is $\sum_i \partial_{x_i}(\rho v_i v_j)$. Finally, $\mathbf{f}_{ext}$ is the external force per unit volume.

Usually, the only external force in the system is gravity, and in this case, we have simply

$$\mathbf{f}_{ext} = -\rho g\hat{\mathbf{z}} \ , \tag{7.1.4}$$

where $g$ is the gravitational acceleration and $\hat{\mathbf{z}}$ a unit vector in the vertical (upward) $z$ direction.

As it stands, equation (7.1.3) is general, but it is incomplete as long as we do not specify how the stress tensor $\sigma$ depends on the fluid properties. Now, in an inviscid liquid (without friction), there is just the hydrostatic pressure, which is a scalar quantity. In this case, we have $\sigma_{ij} = p\delta_{ij}$ wheren $\delta_{ij}$ is the Kronecker $\delta$ ($= 1$ when $i = j$ and 0 for $i \neq j$). In general, this hydrostatic pressure $p$ will itself be a function of the density $\rho$ of the fluid and of some other thermodynamic variable specifying the fluid (e.g., the temperature). Below, we will only retain the density dependence of $p$.

In the presence of viscous effects, we will write

$$\sigma = p\delta_{ij} + \sigma'_{ij} \ , \tag{7.1.5}$$

where $\sigma'$ accounts for the viscous effects. The general form of this tensor $\sigma'$ can be established as follows. Processes of internal friction in a fluid can only occur when different "fluid particles" move with different velocities, so that there is relative motion between different parts of the fluid. Hence $\sigma'$ should vanish for a spatially homogeneous velocity field: it can only depend on the spatial derivatives of the velocity field. Now, when the gradients are small, we may suppose that the momentum transfer due to the viscosity depends only on the first derivatives of the velocity. In lowest order, $\sigma'_{ij}$ will then be a linear function of the *gradients* $\partial v_i/\partial x_j$. Furthermore, $\sigma'$ should also vanish when the fluid is in a uniform rotation, since then there is no friction either. When a fluid is rotating uniformly with rotation rate $\mathbf{\Omega}$, the velocity field is equal to the vector product $\mathbf{\Omega}\times\mathbf{r}$, and hence $\sigma'$ should vanish for such a velocity field.

**Exercise 7** Check that the symmetric expression

$$\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \tag{7.1.6}$$

vanishes for a velocity field corresponding to a uniform rotation.

Exercise 7 shows that $\sigma'$ should be a linear function of the symmetric combination of gradients (7.1.6), so that we can write

$$\sigma'_{ij} = \eta \left( \frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\delta_{ij} \sum_k \frac{\partial v_k}{\partial x_k} \right) + \zeta \delta_{ij} \sum_k \frac{\partial v_k}{\partial x_k} \ . \tag{7.1.7}$$

The coefficient $\eta$ is usually called *the* viscosity or, more precisely, the dynamic viscosity, while $\zeta$ is sometimes referred to as the *second viscosity*. Both of these coefficients are positive. As we shall see below, for incompressible fluids only $\eta$ arises. The last term in the expression between parentheses is conventional: it makes the term between parentheses into a traceless symmetric tensor.

In passing, we note that there is, implicitly, a lot of physics underlying an expression like (7.1.7). For, the assumption is that the microscopic processes like molecular collisions, are so fast that on a coarse-grained scale we can write the viscous stress tensor $\sigma'$ in terms of the *local* gradients of *instantaneous* values of the velocity field. *This is nontrivial!* E.g., for a polymer solution or a polymer melt, the molecular relaxation is so slow that many expressions for the stress tensor (there is no consensus on what is the right one!) have memory terms, i.e., lead to integro-differential equations rather than *o.d.e.*'s.

The viscosities are phenomenological coefficients that depend on the properties of the fluid. Since a single component thermodynamic system is characterized by two thermodynamic quantities, the temperature and the pressure (or density), say, $\eta$ and $\zeta$ are in general also functions of the pressure and the temperature. Quite often, however, the viscosities do not vary appreciably throughout the fluid, and their variation with temperature and density may be neglected. In this case, Eq. (7.1.3) reduces with (7.1.2), (7.1.5) and (7.1.7) to

$$\rho \left[ \partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right] = -\nabla p + \eta \nabla^2 \mathbf{v} + (\zeta + \frac{1}{3}\eta)\nabla(\nabla \cdot \mathbf{v}) + \mathbf{f}_{ext} \ . \tag{7.1.8}$$

This is the form in which the equation is often refered to as the Navier-Stokes equation.

**Exercise 8** Check the derivation of (7.1.8).

**Exercise 9** In classical physics, the equations describing a system should be Galilean invariant, i.e., should be the same independent of whether they are formulated in a fixed frame or in a frame moving with an arbitrary constant velocity $\mathbf{u}$. Check that Eqs. (7.1.2) and (7.1.8) are Galilean invariant. (Hint: note that when we transform to a frame moving with velocity $\mathbf{u}$, the fluid velocity $\mathbf{v}$ transforms into $\mathbf{v} + \mathbf{u}$). Also, the derivative

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \tag{7.1.9}$$

is sometimes called the "material derivative" or "convective derivative". Argue that this is indeed the derivative of a quantity when we follow a "fluid particle" as it convects with the flow. Finally, give the interpretation of the terms if we write the mass conservation equation (7.1.2) in the form $d\rho/dt + \rho \nabla \cdot \mathbf{v} = 0$.

Eqs. (7.1.2) and (7.1.8) are not yet a closed set of equations as long as we do not specify how the pressure $p$ depends on the fluid properties. If temperature variations play an important role, we certainly have to consider the fifth equation associated with the last conserved quantity, energy [see the discussion at the beginning of this section and exercise (3)]. Also, if the fluid is incompressible, we have to specify how the pressure depends on the density $\rho$ of the fluid. We will postpone discussion of these cases to later in Section 7.4, and first consider the important class of problems where the equations are closed, incompressible flows.

## 7.2 The equations for incompressible flows

In many situations, it is a good approximation to regard the fluid as incompressible. This means that we can consider the mass density $\rho$ of the fluid to be constant. Eqs. (7.1.2) and (7.1.8) then reduce to[2]

$$\nabla \cdot \mathbf{v} = 0 \,, \tag{7.2.1}$$

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\rho^{-1}\nabla p + \nu \nabla^2 \mathbf{v} + \mathbf{f}_{ext} \,. \tag{7.2.2}$$

Here we have introduced the so-called kinematic viscosity

$$\nu = \eta/\rho \,. \tag{7.2.3}$$

These equations for an incompressible fluid flow might look innocent, but the nonlinearity on the left hand side of the second equation make them extremely hard to solve except in a few trivial cases where this term is identically zero (see the exercises below) or relatively trivial. That these nonlinear equations contain a lot of interesting physics is illustrated by the fact that it is believed (but not proven) that these equations should be sufficient to describe fully developed turbulence — but turbulence is still considered one of the open fundamental questions of fluid dynamics!

Although one does not normally follow this route in analyzing the equations for a particular problem, it is conceptually important to realize that for incompressible flows, one can closed equations for the velocity field $\mathbf{v}$ only, and then obtain the pressure field directly from the velocity field, once that is known.

**Exercise 10** Demonstrate the above assertion by showing that if one takes the **curl** of Eq. (7.2.2) [in other words, take $\nabla\times$ of this equation], one obtains an equation involving only $\mathbf{v}$, while if one takes the divergence of (7.2.2) [in other words, take $\nabla\cdot$ of this equation] one obtains a Poisson-type equation for the pressure field.

In the presence of boundaries or walls, we have to specify suitable boundary conditions as well. In practice, the so-called stick boundary conditions are most important: the fluid velocity at the wall equals the velocity of the wall itself. For a free surface, which can not support stress, we need a no-stress boundary condition, however (see exercise 15).

[2]Note that if the external forces is just gravity, Eq. (7.1.4), its effect can be accounted for by a redefinition of the pressure, $p + \rho g z \longrightarrow p$.

We now consider some examples. Flow between fixed walls is often called Poisseuille flow — we will consider planar Poisseuille flow and Poisseuille flow in a pipe below. Flow between walls which are moving relative to each other is often called Couette flow — exercises 13 and 14 consider the Couette flow between rotating cylinders. In all the examples below we just consider steady, so-called laminar, flow. In exercises 11–14 we ignore gravity for simplicity, although it can easily be included in the analysis (see footnote 2).

**Exercise 11** *Planar Poisseuille flow.* Consider an incompressible steady flow between two fixed planar plates, a distance $d$ in the $y$-direction apart (at $y = 0$ and $y = d$), with the $x$-axis in the direction of the flow. Starting from the observation that the velocity $\mathbf{v}$ can be a function of $y$ only, $\mathbf{v} = v(y)\hat{\mathbf{x}}$, show that the velocity field is

$$v = -\frac{1}{2\eta}\frac{dp}{dx}y(d-y) \ , \tag{7.2.4}$$

and that the mean fluid velocity (i.e., averaged over the vertical distance) is

$$\bar{v} = -\frac{d^2}{12\eta}\frac{dp}{dx} \ . \tag{7.2.5}$$

The frictional force at one of the plates equals $\sigma_{xy}$ evaluated at the plate (Cf. exercise 6). Show that the friction per unit area is $-d(dp/dx)$.

**Exercise 12** *Poisseuille flow in a pipe.* Repeat the calculation of the previous exercise for Poisseuille flow in a cylindrical pipe whose radius is $R$. (Hint: use the expression for the Laplacian in cylindrical coordinates). Show that the mass $Q$ of fluid flowing per unit time through the cross-section of a pipe of length $L$ is given by

$$Q = \frac{\pi\Delta p}{8\nu L}R^4 \ . \tag{7.2.6}$$

Here $\Delta p$ is the pressure difference across the pipe. Thus, the "discharge" $Q$ increases with the fourth power of $R$, an important result to know for a plumber!

**Exercise 13** *Couette flow between rotating cylinders.* Calculate the steady velocity field of incompressible flow between two rotating concentric cylinders. The inner cylinder has radius $R_1$ and angular rotation rate $\Omega_1$, the outer one radius $R_2$ ($> R_1$) and rotation rate $\Omega_2$. Hint: 1) use again cylindrical coordinates, and use the symmetry of the problem to make the Ansatz $v_z = v_r = 0$, $v_\phi = v(r)$, $p = p(r)$. 2) Note that the homogeneous linear ordinary differential equation for $v(r)$ admits solutions of the type $r^n$. Determine two independent value of $n$ by substitution.

**Exercise 14** The friction force per unit area on the inner cylinder in the previous example is given by the component $\sigma'_{r\phi}$ of the stress tensor, evaluated at the cylinder. Using

$$\sigma'_{r\phi}|_{r=R_1} = \eta\left[\frac{\partial v}{\partial r} - \frac{v}{r}\right]_{r=R_1} \ , \tag{7.2.7}$$

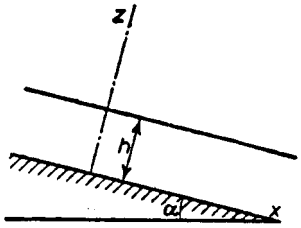


Figure 7.1. Coordinate system for our analysis of flow of a liquid layer down an inclined plane.

show that the total moment $M_1$ on the inner cylinder, due to the viscous forces, is, per unit length

$$M_1 = -\frac{4\pi\eta(\Omega_1 - \Omega_2)R_1^2 R_2^2}{R_2^2 - R_1^2} \ . \tag{7.2.8}$$

In practice, this result is often used to determine the viscosity of a fluid by measuring the torque on a cylinder in Couette flow.

**Exercise 15** A fluid layer of thickness $d$ is flowing down an plane inclined with angle $\alpha$ relative to the horizontal direction (see Fig. 7.1). The upper surface is a free surface. Determine the flow due to gravity, and show that the total amount of mass flow per unit length along the plane in the direction perpendicular to the flow is

$$Q = \frac{\rho g d^3 \sin\alpha}{3\nu} \ . \tag{7.2.9}$$

Hint: introduce a coordinate system is indicated in the figure, and use that the boundary conditions at the upper free surface are $\sigma_{xz} = 0$ (no shear stress on a free surface), $\sigma_{zz} = -p_0$, where $p_0$ is the ambient atmospheric pressure.

In the above examples, we have only considered the stable laminar flow states. For high enough pressures or rotation rates, all these flow states become unstable, and beyond the threshold for these unstabilities various nontrivial flowpatterns result. In almost all of these cases, the flow is highly nonlinear. The Couette system is a particularly rich experimental system with many nontrivial flow states. Some of these are understood quite well, others not that well.

## 7.3 An example of compressible flow: sound waves

In the previous section, we focussed on incompressible flow. A simple example of an inherently compressible flow is a sound wave in a fluid — a sound wave is a periodic variation of the density. We briefly consider here the derivation of the wave equation for these sound waves; the analysis illustrates that in such general cases we have to specify how the pressure depends on the local thermodynamic variables.

Since heat exchange in a fluid is relatively slow, when the density varies in a sound wave, we may, to a good approximation, consider the variation to be adiabatic (i.e. without local exchange of energy). This means that we can write to linear approximation for small variations $\Delta p$ in the pressure

$$\Delta p \approx (\partial p/\partial\rho)_s \Delta\rho \ , \tag{7.3.1}$$

where the derivative is evaluated in the steady reference state. We now consider small amplitude sound waves, sound waves whose amplitude is so small that the Navier-Stokes equations can be linearized about the steady state without flow, $\rho = \rho_0, \mathbf{v} = 0$.

**Exercise 16** Linearize Eqs. (7.1.2) and (7.1.8) in the deviations $\Delta\rho$ and $\Delta\mathbf{v}$ about the state $\rho = \rho_0, \mathbf{v} = 0$, using (7.3.1). Show that the equations can be combined into one single equation for $\Delta\rho$,

$$\frac{\partial^2 \Delta\rho}{\partial t^2} - \left(\frac{\partial p}{\partial \rho}\right)_s \nabla^2(\Delta\rho) + \rho_0^{-1}\left(\zeta + \frac{4}{3}\eta\right)\nabla^2 \frac{\partial \Delta\rho}{\partial t} = 0 \ . \tag{7.3.2}$$

**Exercise 17** Note that when we ignore the viscous damping terms in the above equation (7.3.2), it reduces to the second order wave equation. What is the speed of sound $c$ in a fluid according to this equation? Do you have an idea of the order of magnitude of $c$ in air or water?

**Exercise 18** Give the dispersion relation for a simple fourier mode $\Delta\rho \sim e^{-i\omega t + ikx}$ in the presence of viscous effects.

Sound modes in a fluid can actually also be probed by laser light scattering; in this case, the light waves get Doppler shifted by the sound modes; the damping of the modes, and hence the sum $\zeta + 4/3\eta$ of the viscosities can then be obtained from the linewidth of the scattered light.

We can now also answer the question when a fluid can, to a good approximation, be regarded incompressible. Basically whenever the typical flow speeds $v$ are much less than the speed of sound determined in exercise 17, $v \ll c$, the pressure variations occurring as a result of the motion are so small (and are essentially following the changes in the flow "instantaneously") that the variations in the density associated with this (or in other thermodynamic quantities) can be neglected — in other words, the fluid can be considered incompressible.

## 7.4 The equation for the temperature

We have mentioned several times that for a fluid, there is a fifth hydrodynamic equation associated with energy also being a conserved quantity. For an ideal fluid (without viscosity), the equation is simply that the entropy is a conserved quantity. For a non-ideal fluid, the general formulation can be given in generality (See [LL], sec. 49), but the formulation is too cumbersome for our purposes, especially since it involves several thermodynamic manipulations. However, as we already discussed above, when the fluid velocity is much smaller than the speed of sound in the fluid, the effect of the pressure variations on the density can normally be neglected. Likewise, in this limit $v \ll c$, we can also neglect the variation of the entropy with pressure, and instead focus on the variation of the entropy with temperature. In this case, the energy or entropy equation simplifies considerably. Since the specific heat at constant pressure, $C_p$ is nothing but the derivative of the entropy at constant pressure,

$$C_p = T\left(\frac{\partial s}{\partial T}\right)_p \ . \tag{7.4.1}$$

It is therefore not surprising that in this approximation, our fifth equation becomes

$$\left(\frac{\partial T}{\partial t} + \mathbf{v}\cdot\nabla T\right) = \nabla\cdot(\kappa\nabla T) + \frac{1}{\rho C_p}\sum_{ij}\sigma'_{ij}\frac{\partial v_i}{\partial x_j}\ . \tag{7.4.2}$$

Here $\kappa$ is the thermal diffusion coefficient[3]. The interpretation of this equation is actually quite transparant. On the left hand side, we again see the "material derivative" of exercise 9, i.e. the derivative of the temperature as we follow a "fluid particle" as it flows. The right hand side expresses that the temperate can change due to two effects. The first term is associated with heat diffusion, the fact that in the presence of a temperature gradient $\nabla T$ there is a "heat current" proportional to this gradient (this is sometimes called Fick's law), very much like when we consider a collection of particles, there is generally a diffusion current proportional to the gradient of the number of particles). The second term accounts for the heat generated by the internal viscous "friction" in the fluid. In view of equation (7.1.7) for $\sigma'$, this term is quadratic in the velocities, very much like the energy dissipation in Newtonian mechanics of a particle in the presence of a friction force which itself is proportional to the velocity, is quadratic in the velocity.

The temperature equation (7.4.2) together with the Navier-Stokes equation (7.1.8) and the mass conservation equation (7.1.2) will be the starting point for our analysis of Rayleigh-Bénard convection in the next chapter.

## 7.5 Suggested further reading

There are numerous introductory books on hydrodynamics that cover the subject of this chapter. Our presentation has been inspired mostly by [LL]; the article *An introduction to hydrodynamics* by Castaing in [GM] is also very nice — became available to us after completing the chapter. Its spirit is very much like the one of this chapter, and I recommend it strongly if you want to read more. It is less overwhelming than the book by [LL].

[3]Beware! We have followed the notation of [Man] in which $\kappa$ is the thermal diffusion coefficient. It is more common, however, to use $\kappa$ for the heat transport coefficient, which in the present notation equals $\rho C\kappa$

# Chapter 8

# The Rayleigh-Bénard instability

In this chapter, we analyze an important fluid dynamical system which shows a pattern forming instability, the Rayleigh-Bénard problem. The analysis that we will perform is exemplary for many other pattern forming instabilities in fluids and other systems. Before embarking on the analysis, we will go through some preliminaries.

## 8.1 Dimensional Analysis and similarity

So far, we have not paid attention to the fact that all physical quantities in the equations that we have discussed, actually have a *physical dimension*. E.g., if we use $M$ to denote a mass, $L$ to denote a length, and $T$ to denote a time, and indicate the physical dimension of the velocity $v$ by $[v]$, then the fact that a velocity is a length per unit time, is expressed by

$$[v] = LT^{-1} \tag{8.1.1}$$

**Exercise 1** Give the physical dimensions of the other quantities that were considered in chapter 7, $\rho$, $p$, $\sigma$, $\eta$, $\nu$, and $\kappa$.

In physics, but in fluid dynamics in particular, dimensional analysis can be a powerful tool to understand the general form that one expects for the solutions of the equations describing a given problem, and of the scaling of the various quantities with, e.g., the system size.

The basic observation underlying dimensional analysis is the fact that when we have formulated a physical problem in mathematical terms correctly, it should not matter whether we measure time in seconds or hours, lengths in centimeters or meters, etc. Of course, if we solve our equations and write an expression for a velocity field, then this quantity will scale with our basic length scale and inversely with our timescale: if we change our basic unit of length from meters to centimeters, say, the the velocity field in the new units will become 100 times as large, since the physical dimension of $v$ is $[v] = L/T$. But if we change our mass scale from grams to kilograms, the velocity can not change. These arguments can be formalized by applying a bit of group theory to the scale transformations — see [Ba1], [ Ba2]) — but what it simply comes down to is that apart from factors that set the overall scale of the quantity

we consider, the essential dependence on parameters like the viscosity $\nu$ can *only be through dimensionless quantities*.

A simple example is probably best to illustrate this. Consider incompressible flow in a pipe, the Poisseuille flow problem that we analyzed in the laminar regime in chapter 7, exercise 12. We assume the pipe is very long, so long that boundary effects near the inlet and outlet can be neglected, and ignore gravity. Then, the incompressible Navier-Stokes equation (7.2.2) just contains one dimensional parameter, the kinematic viscosity $\nu$, whose dimension is $[\nu] = L^2T^{-1}$.

**Exercise 2** The incompressible Navier-Stokes equation (7.2.2) also contains a factor $\rho^{-1}$ in front of the pressure term. Why can we nevertheless say that the equation contains only one dimensional parameter?

The only dimensional parameter characterizing the pipe is its radius $R$ with obvious dimension $[R] = L$. If we use a velocity $U$ to characterize the scale of the flow velocity (e.g., $U$ can be the average downstream velocity), then the *only* dimensionless number we can form from the available quantities is the so-called Reynolds number

$$Re = \frac{UR}{\nu} \ , \tag{8.1.2}$$

and if we use $r$ as a radial coordinate in the pipe, it follows that the solution of the flow equations can be written in the form

$$\mathbf{v} = U \ \mathbf{f}(r/R, Re) \ , \tag{8.1.3}$$

where $\mathbf{f}$ is some function, which can only be determined by solving the equations explicitly.

Eq. (8.1.3) illustrates an important consequence of dimensional analysis: since the flow field essentially depends only on the dimensionless Reynolds number, the flow of fluids with different viscosities or in pipes with different diameter is *similar* if their Reynolds number is the same. In other words, we can reach a given Reynolds number with a very viscous fluid and a pipe of large diameter, or with a fluid with small viscosity in a thin pipe. The only thing that matters is whether the Reynolds number is the same! This observation is the reason we can use a scale model of a plane in a windtunnel to study the flow around a real airplane.

The use of dimensional analysis has another important consequence, which is nonrigorous, but often helps in finding suitable approximations. Usually, whenever a dimensionless number is very small or very large, that indicates that a particular physical effect may probably be neglected. E.g., the Reynolds number is essentially a measure for the flow velocity. When the Reynolds number is small, the nonlinearities in the Navier-Stokes equation can often be neglected, and one may work with the linearized equation — this approximation is so important in colloids, where one has small particles in a fluid (remember that a small size helps to make the Reynolds number small!), that the field has gotten its own name: low Reynolds number hydrodynamics [HB]. Likewise, in Poisseuille flow in a pipe the flow is laminar at small Reynolds numbers, but becomes turbulent at large ones. As another example, note

that if gravity has an important effect on the flow, we may at define another dimensionless number, the so-called Froude number $F = U^2/Rg$. Depending on the typical value of this number in a given flow problem, certain approximations may be appropriate.

**Exercise 3** Consider the temperature equation (7.4.2). If we ignore the last term on the right hand side that accounts for heating through the viscosity, what other dimensionless number can you define for a fluid in which temperature effects play a role?

## 8.2 The Rayleigh-Bénard instability

Consider the experimental setup sketched in Fig. 8.1: two planar parallel plates, between which we have a fluid. The spacing between the plates is $h$, and the lateral extent of the plates is so big, that we can think of the system as being laterally infinite. The two plates are held at different temperatures, $T_b$ at the bottom and $T_t$ at the top. When the bottom plate is hotter, $T_b > T_t$, the fluid density becomes smaller at the bottom than at the top, since a fluid expands when the temperature increases (except water below 4 degrees Celsius). What is found experimentally is that if the temperature difference $\Delta T = T_b - T_t$ is small enough, there is no convection in the fluid. Heat is transported from the bottom to the top plate, of course, but this only happens through heat conduction. At some point, however, it is found that the fluid starts to convect, i.e., there is an appreciable fluid flow, whose amplitude increases with increasing temperature difference.

The physical origin of this effect, which is called the Rayleigh-Bénard instability, is very simple. We already saw above that when the bottom plate is heated relative to the upper plate, the liquid on the top becomes more dense than on the bottom. This clearly is an unstable situation! The denser fluid tends to "fall down", making the lighter fluid rise upwards. In particular, if we consider a little hot blob of lighter fluid rising upwards from the bottom, the difference with the density of the surrounding fluid becomes bigger the higher the blob rises. So the effect gets amplified. In practice, there is of course another effect as well: as the blob rises, the temperature difference with the surrounding fluid also gets smoothed out due to heat diffusion. If heat diffusion is fast (large $\kappa$), this effect is quite important.

Since we can only have upward flow at one point if the flow is downwards at some other point, we expect the emergence of periodic flow patterns. As we shall see, the wavelength of this pattern is of the order of the plate spacing $h$, not surprisingly.

We shall see later in this chapter that the emergence of flow patterns is indeed associated with the existence of a linear instability: above a precise value of the temperature difference, the nonconvective state is linearly unstable to convection modes.

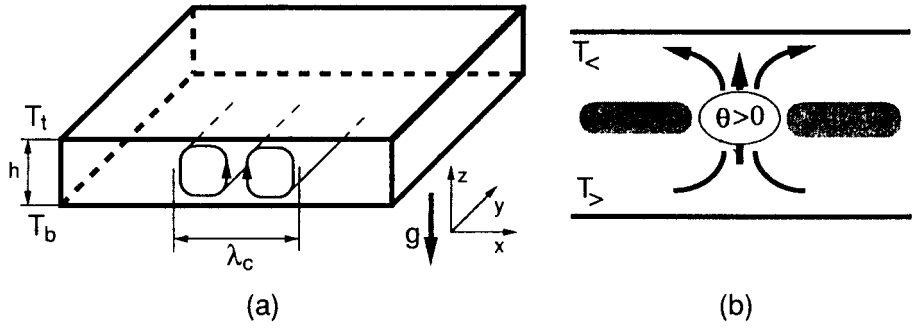


Figure 8.1. Geometry and mechanism of the Rayleigh-Bénard convection.

## 8.3 The Boussinesq approximation

To analyze the Rayleigh-Bénard instability quantitatively, it is clearly crucial to include the variation of the density with temperature. We will do so in the so-called Boussinesq approximation, which takes the temperature variation of the fluid into account in the boyancy term associated with the gravity, but otherwise treats the fluid flow as incompressible. It arises as follows. Consider first the conductive state without flow ($\mathbf{v} = 0$). Putting the plates at $z = h$ and $z = 0$, we write the vertical temperature field $T_0(z)$ in this case as

$$T_0(z) = T_b - \Delta T z/h \ . \tag{8.3.1}$$

When the temperature variations are not too big, we can write to first approximation

$$\rho_0(z) = \rho(T_0(z)) = \rho_b(1 + \alpha \Delta T z/h) \ , \tag{8.3.2}$$

where

$$\alpha = -\frac{1}{\rho_b}\frac{d\rho}{dT} \tag{8.3.3}$$

is the thermal expansion coefficient of the liquid, and where $\rho_b$ is the reference density at the bottom.

If we write the pressure in the conductive state as $p_0(z)$, we simply have for the Navier-Stokes equation (7.2.2) with (7.1.4) in the conductive state $\mathbf{v} = 0$

$$0 = -\frac{1}{\rho_0(z)}\nabla p_0(z) - g\hat{\mathbf{z}} \ . \tag{8.3.4}$$

We now arrive at the so-called Boussinesq approximation for the conductive state $\mathbf{v} \neq 0$ by taking the fluid flow field itself incompressible, while keeping the density variation due to the temperature variations in the terms on the left hand side of (7.2.2) to lowest order. If we write the total pressure as $p_0(z) + p$ (so that from here on $p$ denotes the change in pressure relative to the reference conductive state), the temperature as $T = T_0(z) + \theta$ and the density variation with temperature as $\rho = \rho_0(z)(1 - \alpha\theta)$, then we arrive at the so-called Boussinesq equations

$$\nabla \cdot \mathbf{v} = 0 \ , \tag{8.3.5}$$

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\rho_b^{-1}\nabla p + \nu \nabla^2 \mathbf{v} + \alpha g \theta \hat{\mathbf{z}} \ , \tag{8.3.6}$$

$$\partial_t \theta + \mathbf{v} \cdot \nabla \theta = \kappa \nabla^2 \theta + \Delta T v_z/h \ . \tag{8.3.7}$$

**Exercise 4** Check equations (8.3.5)-(8.3.7).

**Exercise 5** What is the dimension of $\alpha$. Which two dimensionless numbers can you form with the fluid parameters in the Boussineq approximation and the layer spacing $h$?

## 8.4 The boundary conditions

One of course also needs to specify the boundary conditions. For the temperature, we will assume that the upper and lower boundary temperatures are held exactly constant. For the change in temperature $\theta$, this implies

$$\theta(z=0) = \theta(z=h) = 0 \ . \tag{8.4.1}$$

For the velocity field, the experimentally relevant boundary conditions is that the fluid velocity is zero at the plates

$$\text{stick boundary conditions:} \quad \mathbf{v}(z=0) = \mathbf{v}(z=h) = 0 \ . \tag{8.4.2}$$

These boundary conditions are usually refered to as the stick boundary conditions. This is to distinguish them from the experimentally unrealistic but computationally convenient slip boundary conditions,

$$\text{slip boundary conditions:} \quad v_z(z=0) = v_z(z=h) = 0 \ , \quad \partial \mathbf{v}_h/\partial z|_{z=0,h} = 0 \ . \tag{8.4.3}$$

From here on, we will use $v_h$ for the horizontal component of the velocity field

$$\mathbf{v}_h = v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}} \tag{8.4.4}$$

**Exercise 6** Show that (8.4.3) is the appropriate boundary condition for a free fluid surface that can not support any viscous stress.

Eq. (8.4.2) is always the realistic boundary condition; by clever tricks (e.g. using a material for the plates that is a very good thermal conductor) experimentalists can get close to realizing the temperature boundary conditions (8.4.2). It is also experimentally possible to fix the heat flow through the upper and lower plates. For a discussion of non-ideal boundary conditions, we refer to [Man], section 4.1.2.

**Exercise 7** Translate the boundary condition of fixed heat flow into mathematical terms.

**Exercise 8** Show that the Boussinesq equations (8.3.5)-(8.3.7) and are invariant (both for stick and for slip boundary conditions) under the transformation

$$\begin{aligned} v_{x,y}(x,y,z,t) &\longrightarrow v_{x,y}(x,y,-z+h,t) \ , \\ v_z(x,y,z,t) &\longrightarrow -v_z(x,y,-z+h,t) \ , \\ \theta(x,y,z,t) &\longrightarrow -\theta(x,y,-z+h,t) \ . \end{aligned} \tag{8.4.5}$$

The fact in the Boussinesq approximation the the fields have an up-down ($z \longrightarrow -z$) symmetry will play a role in later chapters. In practice, of course, the symmetry is weakly broken, e.g. due to terms quadratic in $\theta$ in the expansion of the density with temperature, or due to the dependence of the viscosity on temperature. As we shall see, this weak symmetry breaking implies that very close to threshold (and even slightly below it), one finds hexagonal patterns in the Rayleigh-Bénard instability, *not* "roll" or "stripe" patterns. Actually, hexagonal patterns generally arise near threshold in models where the basic equations are not invariant under a reversal of the sign of the fields, such as the models for Turing patters that we discussed in chapter 6.

## 8.5 The Rayleigh-Bénard instability: simple arguments

To get a feel for the physics, we first do some back-of-the-envelope estimates of the various effects involved in the Rayleigh-Bénard instability.

**Exercise 9** Consider a little blob of fluid of temperature $T_b$ at near the upper plate, where the surrounding fluid is of temperature $T_t$. Give an order of estimate for the *buoyancy force* (the force due to gravity and the density difference between the blob and the surrounding fluid). What is the resulting accelaration due to this force? Now estimate the timescale $\tau_{boy}$ for this blob to get accelerated to the bottom plate, a distance $h$ below.

**Exercise 10** Now estimate the relaxation of the velocity field $\mathbf{v}$ as follows. Consider the setup of the Rayleigh-Bénard instability with the two planar plates. To study small velocities in the absence of a temperature field, linearize the hydrodynamic equation (8.3.6). Consider a velocity field $v_x(z,t)$ in the $x$ direction which is homogeneous in the horizontal direction, in the absence of an overall pressure gradient, $\nabla p =$ and of gravity. Study the relaxation of this field, by performing a Fourier transform $v_x = \sum_{n=1} a_n(t) \sin(n\pi z/h)$, appropriate for stick boundary conditions. How do the coefficients $a_n$ relax to zero? Give an expression for the slowest viscous relaxation time $\tau_{visc}$ for this type of flow.

**Exercise 11** Do an analysis similar to the one in the previous analysis, but now for the temperature field. Give an expression for the temperature relaxation time $\tau_{temp}$.

The discussion of section 8.2 suggests that the Rayleigh-Benard instability will occur if the temperure difference is large enough that the time $\tau_{boy}$ that it takes a cold blob to sink to the bottom is shorter than both the viscous relaxation time $\tau_{visc}$ and the temperature relaxation time $\tau_{temp}$:

$$\frac{\tau_{visc}\tau_{temp}}{\tau_{boy}^2} \gtrsim 1 \quad \Longrightarrow \quad \text{instability} \; . \tag{8.5.1}$$

This result suggests that it is useful to define the dimensionless *Rayleigh* number

$$R = \frac{\alpha g \Delta T h^3}{\nu \kappa} \; . \tag{8.5.2}$$

**Exercise 12** Show that 8.5.1 suggests that the instability will occur for $R$ of order $2\pi^4 \approx 200$.

**Exercise 13** As we shall see later, in reality the Rayleigh-Bénard instability arises for $R = R_c = 1707.76$. Can you think of an argument why you in (8.5.1) we have underestimated the critical Rayleigh number $R_c$ (i.e., the Rayleigh number where the instability sets in)? Hint: can you think of a reason why you've overestimated $\tau_{visc}$ and $\tau_{temp}$?

**Exercise 14** $R$ is not the only dimensionless number for the Rayleigh-Bénard problem. Can you give a second one?

Note that the Rayleigh number $R$ is the experimental "control parameter" that an experimentalist can easily change in a given experimental setup: you turn up the Rayleigh number by increasing $\Delta T$ — all other quantities are fixed.

The above discussion nicely illustrates the emergence and usefulness of dimensionless numbers. The Rayleigh number increases with the third power of $h$, and so the instability occurs very easily in a thick cell, and experimentalists can play with their cell dimensions to have the instability arise for an appropriate temperature difference. If they want to compare their results to calculations based on the Boussinesq approximation, they should not make their experimental cells too thick, since then the density variation with pressure becomes too big. On the other hand, if they make their cell too thin, then the necessary temperature differences $\Delta T$ becomes so large that the temperature dependence of the transport coefficients might become important.

## 8.6 Calculation of the linear instability for slip boundary conditions

We now turn to an explicit calculation of the instability threshold in the Rayleigh-Bénard problem.

Returning to Eqs. (8.3.5)-(8.3.7), we first introduce dimensionless variables by choosing $h$ and $\tau_\theta = h^2/\kappa$ as the unit of length and time, $h/\tau_\theta = \kappa/h$ as the unit of velocity, $\kappa\nu/\alpha g h^3$ as the unit of temperature, and $\kappa^2/h^2$ as the unit for $p/\rho_b$. From here on, these dimensionless units will be used, but we will not distinguish them notationally from the dimensional ones used up to now. In addition, we split the velocity in the horizontal component $\mathbf{v}_h$ and the vertical component $v_z$ and write $\nabla_h = \hat{\mathbf{x}}\partial_x + \hat{\mathbf{y}}\partial_y$. The equations then become

$$\nabla_h \cdot \mathbf{v}_h + \partial_z v_z = 0 \ , \tag{8.6.1}$$

$$\frac{1}{P}[\partial_t \mathbf{v} + (\mathbf{v}\cdot\nabla)\mathbf{v}_h + \nabla_h p] = \nabla^2 \mathbf{v}_h \tag{8.6.2}$$

$$\frac{1}{P}[\partial_t v_z + (\mathbf{v}\cdot\nabla)v_z + \partial_z p] = \nabla^2 v_z + \theta \ , \tag{8.6.3}$$

$$(\partial_t \theta + \mathbf{v}\cdot\nabla)\theta = \nabla^2\theta + R v_z \ , \tag{8.6.4}$$

where we have introduced the dimensionless Prandtl number

$$P = \nu/\kappa \tag{8.6.5}$$

which you may already have come up with before in exercise 14.

**Exercise 15** For water, we have at 40 degrees Celsius $\alpha \approx 410^{-4}$ per degree Celsius, $\nu = 6.510^{-3} cm^2/s$, $\kappa = 1.510^{-3} cm^2/s$; of course, $g = 9.8 m/s^2$. How thick should a Reighleigh-Bénard cell be in order that the instability occurs for a temperature difference of order 10 degrees Celsius? What is the Prandtl number for water at this temperature?

**Exercise 16** Check equations (8.6.1)-(8.6.4).

Following up on the discussion after exercise 14, we note that when one considers a problem characterized by two (or more) dimensionless numbers, their choice can not be fixed by mathematical conideractions: the most useful choice depends on physical considerations and sometimes a bit of estatics, and often the best choice only becomes clear on hindsight. E.g., we are free to introduce the Clinton number $RP^{5/2} = \alpha g \Delta T h^3 \nu^{3/2}/\kappa^{7/2}$ — it is a fine dimensionless number, but it is not very useful for analyzing the Rayleigh-Bénard instability. As the above equations show, instead $R$ and $P$ naturally emerge in this problem. As we discussed, in an experiment, $R$ is changed simply by increasing the temperature, while $P$ remains fixed for a given choice of fluid.

**Exercise 17** Can you think of a molecular reason why for almost all fluids, $P$ is of order of one or larger? An exception to this rule are metallic fluids, like liquid mercury or gallium, for which $P$ is very small. Can you think of the physical reason for this? Actually, for gases, $P$ is often of order one.

When we perform the linear stability analysis by linearizing the equations in $\mathbf{v}$ and $\theta$, we see that the linear part of the temperature equation (8.6.4) only involves $\theta$ and $v_z$. The linear part of the equation for the velocity $v_z$ involves the pressure $p$ as well, and through this term it is coupled to the equations for the horizontal flow velocity $\mathbf{v}_h$. As we already saw before in exercise 11, however, it is possible to eliminate the pressure by taking the curl of the Navier-Stokes equation. It is useful to apply this trick here, since as we shall see, we then get a set of two closed linearized equations in $v_z$ and $\theta$ that is easy to work with.

**Exercise 18** Show that by considering the horizonal components of the vorticity $\nabla \times \mathbf{v}$, we obtain from Eqs. (8.6.2), (8.6.3) for the linear parts

$$(\partial_t - P\nabla^2)(\partial_y v_z - \partial_z v_y) - P\partial_y \theta = \text{nonlinear terms} , \tag{8.6.6}$$

$$(\partial_t - P\nabla^2)(\partial_z v_x - \partial_x v_z) + P\partial_x \theta = \text{nonlinear terms} . \tag{8.6.7}$$

**Exercise 19** Take the $y$ derivative of Eq. (8.6.6) and the $x$ derivative of Eq. (8.6.7), subtract the two results and eliminate the horizontal components with the aid of the incompressibility condition (8.6.1). Show that this yields equations (8.6.8) and (8.6.9) below.

Upon eliminating the horizontal components of the velocity field, the linearized equations for $\theta$ and $v_z$ become

$$\partial_t \nabla^2 v_z = P(\nabla^4 v_z + \nabla_h^2 \theta) \ , \tag{8.6.8}$$

$$\partial_t \theta = \nabla^2 \theta + R v_z \ . \tag{8.6.9}$$

Since we ignore sidewalls and take the setup to be laterally unbounded, we can now do the stability analysis by using simple fourier modes,

$$\begin{pmatrix} v_z \\ \theta \end{pmatrix} = \begin{pmatrix} V(z) \\ \Theta(z) \end{pmatrix} e^{i\mathbf{k}_h \cdot \mathbf{r}_h + \sigma t} \ . \tag{8.6.10}$$

As usual, our goal is to get the dispersion relation for the growth rate $\sigma$ as a function of the wavenumber $k$. Since we are now dealing with a three-dimensional problem, this dispersion relation depends on the functions $V(z)$ and $\Theta(z)$, which in turn depend on the boundary conditions.

## 8.7 Slip boundary conditions

The easiest boundary conditions are the slip boundary conditions, since here the eigenfunctions $V(z)$ and $\Theta(z)$ can be obtained analytically in closed form.

**Exercise 20** Show that in addition to (8.4.3), the slip boundary conditions also imply the boundary condition $\partial_z^2 v_z = 0$ at the plates, i.e. at $z = 0$ and $z = 1$ in dimensionless units.

In this case, the eigenfunctions which obey the boundary conditions are simply $V_n(z) = \Theta_n(z) = \sin(n\pi z)$, and the corresponding dispersion relation $\sigma_n$ for each mode $n$ becomes

$$(\sigma_n + (k^2 + n^2\pi^2)) \left(\sigma_n (k^2 + n^2\pi^2) + P(k^2 + n^2\pi^2)^2\right) - RPk^2 = 0 \ , \tag{8.7.1}$$

where $k = |\mathbf{k}_h|$.

**Exercise 21** Check the dispersion relation (8.7.1) for the slip boundary conditions.

**Exercise 22** Show from (8.7.1) that the $n^{th}$ mode has an finite wavelength instability at a critical Rayleigh number

$$R_n(k) = \frac{(k^2 + n^2\pi^2)^3}{k^2} \ , \tag{8.7.2}$$

in other words, $\sigma_n(k) > 0$ for $R > R_n(k)$. Moreover, show that the instability is stationary, i.e., that $\mathrm{Im}\sigma_n = 0$ for the unstable modes.

The above result shows that if we increase the temperature difference, and hence $R$, the first mode to become unstable is the $n = 1$ mode. The instability threshold is thus given by the minimum of the function $R_1(k)$. This gives the critical Rayleigh number $R_c$ and critical wavenumber $k_c$

$$R_c = \frac{27\pi^4}{4} \approx 657 \ , \qquad k_c = \frac{\pi}{\sqrt{2}} \approx 2.22 \tag{8.7.3}$$

**Exercise 23** Check Eq. (8.7.3).

Within the context of the present slip boundary condition approximation, the above analysis confirms that there is indeed a true instability, so that for $R < R_c$ the conductive state is linearly stable, and that for $R > R_c$ the conductive state is unstable. Since we are using dimensionless units, the result $k_c \approx 2.22$ shows that the lateral size of one convective cells is of the order of the plate spacing $h$.

**Exercise 24** As we shall see, the critical Reynolds number with stick boundary conditions is about a factor three larger than the one for slip boundary conditions. Can understand why slip boundary conditions yield a lower critical Rayleigh number?

**Exercise 25** $R_1(k)$ diverges as $k^4$ for large $k$ according to (8.7.2. Can you understand this scaling with an argument in the spirit of the argument of section 8.5?

## 8.8 Stick boundary conditions

We now turn to a brief discussion of the realistic case of stick boundary conditions.

**Exercise 26** Show that in the case of stick boundary conditions, the condition that $\mathbf{v}_h = 0$ at the plates implies for $v_z$ the boundary condition $\partial_z v_z|_{z=0,1} = 0$ in addition to the boundary condition $v_z|_{z=0.1} = 0$.

Clearly, in this case taking $V(z) = \sin(n\pi z)$ does not satisfy the boundary conditions discussed in the exercise above, and we can not simply get the dispersion relation in closed form. Conceptually, however, the problem is not really much more difficult than before.

If we substitute the general mode Ansatz (8.6.10) into the linearized equation (8.6.8), (8.6.9), we get a system of two coupled *o.d.e.*'s for the functions $V$ and $\Theta$,

$$(\partial_z^2 - k^2)\sigma V = P\left((\partial_z^2 - k^2)^2 V - k^2\Theta\right) , \tag{8.8.1}$$

$$\sigma\Theta = (\partial_z^2 - k^2)\Theta + RV . \tag{8.8.2}$$

These equations can be written as a single linear sixth order *o.d.e.* for one of the fields; e.g., if we elimate $\Theta$ in favor of $V$, we get

$$\left[\left(\partial_z^2 - k^2 - \sigma\right)\left(P(\partial_z^2 - k^2)^2 - \sigma(\partial_z^2 - k^2)\right) + RPk^2\right] V(z) = 0 . \tag{8.8.3}$$

Clearly, the general solution of such a linear equation can be written as the sum of six exponential of the type

$$V(z) = \sum_{m=1}^{6} V_m e^{iq_m z} . \tag{8.8.4}$$

**Exercise 27** The exponents $q_m$ in this expression come in pairs: if $q_1$ is a solution, then $-q_1$ is also a solution. Why?

If we solve for these roots $q_m$, and then apply the boundary conditions, we can solve for $\sigma(k; R, P)$. This can easily be done numerically, but we will not do so here. By determining the minimum of $\sigma$ as a function of $k$ and $R$ for fixed $P$, we then obtain the critical Rayleigh number and $k_c$. It is found that

$$R_c = 1707.76 \ , \qquad k_c = 3.11632 \ . \tag{8.8.5}$$

Note that since $k_c$ is very close to $\pi$, the diameter of the convection cells near threshold is almost quare, i.e. they are only slightly wider than the cell thickness.

We finally briefly consider a simple approximation, that is often very useful in more complicated problems, and that even lies at the basis of many dynamical numerical studies of Rayleigh-Bénard convection.

The difficulty with the case of stick boundary conditions are that the boundary conditions discussed in exercise 26 are not obeyed by simple basisfunctions, like the trigonometric functions sine and cosine. However, they do motivate a simple expansion of the form

$$\Theta(z) = (1/4 - z^2) \sum_{n=0}^{\infty} \Theta_n z^n \ , \tag{8.8.6}$$

$$V(z) = (1/4 - z^2)^2 \sum_{n=0}^{\infty} V_n z^n \ . \tag{8.8.7}$$

If we truncate such an expansion at a particular value of $n$, the coefficients obey sets of linear equations. The coefficients in these equations have to be obtained by projecting the differential equations in the $z$-direction onto certain basisfunctions or weight functions. For the present convection problem, it turns out that a very good approximation can be obtained already by only retaining the lowest order $n = 0$ terms in the above expansion, and projecting the equations onto the same factors $(1/4 - z^2)$ (for the temperature equation) and $(1/4 - z^2)^2$ (for the velocity equation) [projection means that we multiply these equations from the left by these factors, and then integrate over $z$]. That this leads to a very accurate result can be seen in Fig. 8.2, where the marginal stability lines $\sigma(k; R, P) = 0$ are shown in a $R$-$k$ diagram. Curve $A$ corresponds to the slip boundary conditions, curve $B$ to the numerical solution for the case of stick boundary conditions, and curve $B'$ to the above approximation.

We note that while the threshold for the instability does *not* depend on the Prandtl number, the nonlinear behavior above threshold does depend strongly on Prandtl number. In particular, a recent set of experiments have shown that in the regime where $P$ is of order unity, so that the temperature diffusion timescale and the viscous damping timescale are of the same order, quite complicated dynamics can occur, including a chaotic regime dominated by spiral defect chaos. We shall review some to the nonlinear behavior in our meetings.

## 8.9 Further information from the dispersion relation

In the chapter on the derivation of the amplitude equations (chapter 9), we shall see that there is more important information hidden in the behavior of the dispersion

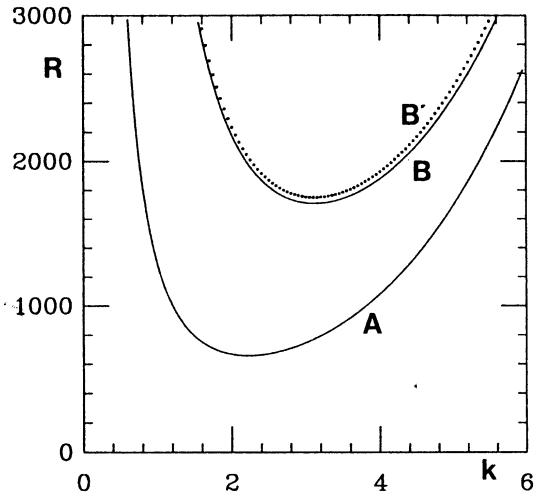


Figure 8.2. The marginal stability curves in the $k$-$R$ plane for three case. Curve $A$ corresponds to the case of slip boundary conditions, curve $B$ to the case of stick boundary conditions, and curve $B'$ to the approximation discussed in the text. Above the marginal stability curve, the conductive state is unstable.

relation near threshold. Intuitively, we can already understand this: the faster $\sigma$ increases upon increasing $R$, the stronger the instability becomes, and the smaller the timescale $\tau_0$ for the growth of the linear mode. Moreover, the larger the curvature of the marginal stability line is at the critical value (at the minimum, see Fig. 8.2), the more the instability is suppressed when we consider modes with wavenumber $k$ slightly different from $k_c$, and the more the closer the wavenumber of nonlinear patterns will lie to $k_c$ in the nonlinear regime. This will imply that the length scale $\xi_0$ over which variations in the amplitude and wavenumber become noticeable, is then large.

When we make these ideas more precise in the later chapter on amplitude equations, we shall see that the derivatives of $\sigma$ at the critical value $R_c$ indeed define a time scale $\tau$ and length scale $\xi$ by

$$\tau_0^{-1} = R_c \left.\frac{\partial\sigma(k;R,P)}{\partial R}\right|_{R_c,k_c} , \qquad \xi_0^2 = -\frac{\tau_0}{2} \left.\frac{\partial^2\sigma(k;R,P)}{\partial k^2}\right|_{R_c,k_c} . \tag{8.9.1}$$

For the case of slip boundary conditions, one finds

$$\tau_0 = \frac{2}{3\pi^2}\frac{1+P}{P} , \qquad \xi_0^2 = \frac{8}{3\pi^2} , \tag{8.9.2}$$

while from the numerical analysis of the case of stick boundary conditions, one finds

$$\tau_0 \approx \frac{1+1.9544P}{38.4429P} , \qquad \xi_0^2 = 0.1479 . \tag{8.9.3}$$

**Exercise 28** Check (8.9.2) from the dispersion relation (8.7.1).

Comparison of the two results shows that the slip boundary conditions are numerically somewhat different, but that again the trends as a function of $P$ for $\tau_0$ and the independence of $\xi_0$ of $P$ are basically the same. For this reason, the slip boundary conditions are often used as a first approximation, when one considers a complicated extension of these calculations. As an illustration, we note that a few years ago, our group extended some of the analysis for the case of an annular Rayleigh-Bénard cell in a rotating frame. In this case, the Coriolis force has to be taken into account, and it is found that there then is a Hopf transition to traveling waves localized near each side wall for Prandtl numbers of order 0.4. In this case, the eigenmodes in one of the horizontal directions are nontrivial, due to presence of the sidewalls. This problem was so complicated that only the case of slip boundary conditions at the upper and lower plate could be treated. See M. van Hecke and W. van Saarloos, Phys. Rev. E *55* R1259 (1997).

The fact that $\xi$ is only about 0.4, has important consequences. We shall see in later chapters that $\xi$ plays the role of the basic length scale for variations of the amplide of the pattern above threshold. Since $\xi_0$ is of the order of the wavelength of the pattern, this means that in Rayleigh-Bénard convection, an experimentalist can relatively easily get a laterally large system, i.e., large in units of $\xi_0$. In some pattern forming systems $\xi_0$ is much larger than the basic lengthscale of the pattern, and then it is much harder to get effectively large systems experimentally.

The Rayleigh-Bénard problem is a very rich system, in which many types of different instabilities occur above threshold, both as a function of Rayleigh number and as a function of Prandtl number. We will illustrate some of the possible behavior in the course meetings. Because of this, and because experimentalists have been able to refine their experimental cells to such a degree so that they could not only test theoretical predictions *quantitatively*, but also identify new dynamical phenomena, the Rayleigh-Bénard problem has in many ways been the working horse of nonequilibrium pattern formation in the last two decades. The most precise confrontation of the weakly nonlinear theory (amplitude description) that we will discuss in the next chapter, with experiments, have been done with this system. See e.g. [CH].

## 8.10 Suggested further reading

We have based our presentation mostly on [Man] — see in particular sections 3.1, 4.1 and 4.2. A more extensive introduction towards some of the pattern forming aspects can be found in [HSH] or in the articles by Hakim and Fauve in [GM]. A detailed review of the comparison between the predictions from weakly nonlinear theory and experiments can be found in [CH].

*[CH]* M. C. Cross and P. C. Hohenberg, *Pattern formation outside of equilibrium*, Rev. Mod. Phys. **65**, 851 (1993).

*[GM]* *Hydrodynamics and Nonlinear Instabilities*, C. Godreche and P. Manneville, eds. (Cambridge University Press, Cambridge, 1998).

*[HB]* J. Happel and H. Brenner, *Low Reynolds number hydrodynamics* (Martinus Nijhoff, Dordrecht, 1973).

*[HSH]* M. van Hecke, W. van Saarloos and P. C. Hohenberg, *Amplitude equations for pattern forming systems*, in: Fundamental Problems in Statistical Mechanics VIII, H. van Beijeren and M. H. Ernst, eds. (North Holland, Amsterdam, 1994).

*[Man]* P. Manneville, *Dissipative Structures and Weak Turbulence* (Academic Press, Boston, 1990).

*[LL]* L. D. Landau and E. M. Lifshitz, *Fluid Dynamics* (Course of Theoretical Physics, vol 6, 2nd edition (Pergamon, New York, 1987).

*[Ba1]* G. I. Barenblatt, *Dimensional Analysis* (Gordon and Breach, New York, 1987).

*[Ba2]* G. I. Barenblatt, *Scaling, Self-Similarity and Intermediate Asymptotics* (Cambridge University Press, Cambridge, 1996).

# Chapter 9

# Derivation of Amplitude Equations: Weakly Nonlinear Theory

Over the course of the last two chapters, we have become familiar with the some of the basic concepts of fluid dynamics. We have seen that the equations that describe hydrodynamic systems are the Navier-Stokes equation, supplemented with a mass-conservation law and the heat equation. Although these equations might look rather innocent, a nonlinear term $\sim (\mathbf{v} \cdot \nabla)\mathbf{v}$ spoils the fun (although it could just as well be described as being responsible for all the fun), and makes hydrodynamics into the complex problem that it is. In chapter 8, we have acquainted ourselves with one of the (conceptually) simplest effects of this nonlinearity, the Rayleigh-Bénard instability. A closed 'box' of fluid is submitted to a vertical temperature gradient, by heating the lower plate, and we have seen that at some critical temperature difference $\Delta T_c$, the homogenous state, where heat is transferred to the top plate purely by conduction, loses stability and the fluid starts to move or *convect*. At onset, a basic state consisting of stationary rolls (in the absence of boundary effects) emerges, and the amplitude of this state varies smoothly when the Rayleigh number $R$ is increased through $R_c$. The nature of the *bifurcation* underlying the appearance of these rolls is therefore *supercritical* (also referred to as a *forward* or *pitchfork* bifurcation), and it is stationary. In this chapter, we will try to go one step beyond the linear stability theory, and find out what happens to these rolls when we increase $R$ even further. We will see that it is possible to construct a theory that adequately describes the behavior above, but near onset. In this regime, the nonlinearities do play a key role, but are not yet as violent as they are in for instance turbulent (= large $R$) regimes. The theory is therefore called a *weakly nonlinear theory*.

## 9.1 A Toy Model : The Swift-Hohenberg equation

Although it is possible (and has been done) to perform the weakly nonlinear analysis on the full set of hydrodynamic equations, the sheer amount of algebra involved tends to cloud the underlying structure of the argument. We will therefore do what scien-

tists often do when they prefer qualitative understanding over quantitative accuracy, and introduce a toy model. A toy model has to be a model that posesses some of the necessary features of the full model, but is mathematically easier to handle, the key question being of course what these necessary features are. One of these is obviously the bifurcation structure. Our model equation should also have a stationary, forward bifurcation when a control parameter is increased through some critical value. Furthermore, we would like the dispersion relation to behave the same, at least near threshold. Finally, our model equation is required to posess the same symmetries as the hydrodynamic equations. When the system is rotationally invariant, the direction of the wavevector can be chosen at will, and we take advantage of this freedom by choosing it to lie along the $x$-axis, and furthermore neglect all variation in the $y$-direction (our model is therefore 1-dimensional). We will take our system to be infinitely long so as to get rid of the effect of the boundaries, and moreover we will assume left-right reflection symmetry. In doing so, we have effectively reduced our three coupled hydrodynamic equations to a single, one-dimensional equation for just one scalar quantity.

**Remark** Although the toy model thus constructed does reflect some of the most important properties of the hydrodynamic equations, the description is far from complete. The dimensions we have 'thrown away' also contain valuable information (cf. the nontrivial modes in the $z$-direction discussed in the last chapter). As we shall see, the linear problem in our toy model is quite trivial, while this already becomes quite rich in the actual RB-system.

**Exercise 1** Demanding that the dispersion relation of our toy model near threshold behaves qualitatively like it does for the full hydrodynamic equations effectively fixes the linear part of the equation. For $R > R_c$, the growth-rate becomes positive for $q = \pm q_c$, and is locally quadratic around these $q$'s. Argue that the linear part of the equation should therefore be of the form

$$\partial_t u(x,t) = \varepsilon u - D(\partial_x^2 + q_c^2)^2 u \quad . \tag{9.1.1}$$

With $u$ a real function. Furthermore, show that the bifurcation in an equation with this linear part is indeed stationary (*i.e.* $\mathtt{Re}s_c = 0$). Sketch the dispersion relation both in an $(\mathtt{Im}s, q)$ and a $(q, \varepsilon)$ plot, to get a feel for what it looks like.

**Exercise 2** As we have seen in the previous chapter, the Navier-Stokes equations in the Boussinesq approximation posess an up-down symmetry[1]. In our one-dimensional model, this translates into a $u \to -u$ symmetry. The simplest nonlinearity we can include is therefore a cubic one:

$$\partial_t u(x,t) = \varepsilon u - D(\partial_x^2 + q_c^2)^2 u - K u^3 \quad , \tag{9.1.2}$$

[1] As was also remarked, the up-down symmetry is a consequence of the Boussinesq-approximation. As the name suggests, this is only an approximation, and the symmetry is therefore always weakly broken in real life. This can be accounted for by the inclusion of small terms that do not respect the up-down symmetry in our model equation (like $u^2$). Such terms give rise to so-called non-Boussinesq effects, and their consequences for the Amplitude equation will be investigated in an exercise later on in this chapter

where by a proper rescaling of $x$ and $u$ (which one?) $D$ and $K$ can be put to 1 and $-1$ respectively. In principle, the $u \to -u$ symmetry allows us to also include a nonlinear term like $u^2\partial_x u$, but why is this not allowed in this case?

We will however forget about other allowed nonlinear terms, and stick with the cubic one. The equation (9.1.2) thus constructed is known as the Swift-Hohenberg equation, and has many interesting properties, some of which we will investigate now.

We will now continue our discussion of the SH-equation by showing that for $\varepsilon > 0$, there exists a band of stationary solutions. In order to do so, let us first examine the existence of periodic solutions to the linearized equation

**Exercise 3** Substitute an Ansatz for stationary periodic solutions $u(x,t) = a_0 e^{iqx} + c.c.$ into the linearized SH-equation. Show that such solutions exists for

$$q = \pm\sqrt{q_c^2 \pm \sqrt{\varepsilon}} \equiv \pm q_m^{\pm} \quad . \tag{9.1.3}$$

Argue that in the neigbourhood of $q_m^{\pm}$, periodic solutions with a small amplitude will exist. Hint : set up a perturbation expansion about the $q_m$-solution in a relevant small parameter.

Let us now turn to the full problem for small, and study the existence of periodic solutions. In order to do so, it will prove useful to write the stationary SH-equation as a 4-component dynamical system, so that we can work in phase-space. The phase-space equations are

$$\left\{ \begin{array}{rcl} y_0' &=& y_1 \\ y_1' &=& y_2 \\ y_2' &=& y_3 \\ y_3' &=& -(\varepsilon + q_c^4)y_0 - 2q_c^2 y_2 + y_0^3 \quad , \end{array} \right. \tag{9.1.4}$$

where $y_0 = u(x), y_1 = u'(x), y_2 = u''(x)$ and so on. What we are dealing with is therefore a 4-dimensional phase space. We can now apply what is often referred to as a counting argument, to find out some useful things about the periodic solutions.

**Exercise 4** Let us start in phase space at $x = 0$, and demand that $y_0(0) = y_2(0) = 0$. This specifies a 2-D plane in the full space, and in this plane we are still free to choose $y_1$ and $y_3$. Suppose now we want to shoot to some point $x$, and demand that at this point again $y_0(x) = y_2(x) = 0$. Count the number of degrees of freedom at the starting point, and the number of constraints at the endpoint. Argue that one expects there to be a unique trajectory fulfilling all the demands. Next, use the symmetries of the SH-equation ($u \to -u, x \to -x$) to demonstrate the existence of a second trajectory connecting $x = 0$ and $-x$, fulfilling the same conditions at the $-x$ endpoint. By 'gluing' these two together, show that one obtains a closed curve in phase space. This corresponds to a periodic solution. Identify the one free parameter left in this argument to argue that one expects these periodic solutions to come in a one-parameter family, parametrised for instance by $q$.

The line of argument of the previous exercise is known as a counting argument, and has proven its worth on several occasions. One counts the number of degrees of freedom and constraints to obtain an upper bound on the multiplicity of trajectories, and therefore solutions. In no way does argument prove the existence of solutions, it just tells one, under the assumption that the flow is continuous, whether one expects 'nearby' solutions *when* one is found. It is however useful before embarking on a full-blown analysis (which we won't do here) to have an idea of what the outcome will be. One has to be careful when the system posesses certain symmetries, because they, in general, add to the number of degrees of freedom. In one instance in the literature where a counting argument yielded that a certain solution should come in a discrete set, an exact solution was found that described a family of solutions. It is thought that this can be ascribed to a certain "hidden symmetry".

Now let us see what happens when one investigates periodic states in the SH-equation. The linear terms of course preserve the period, but the nonlinear term does something funny. Substituting in a function of the form

$$u(x) \sim \cos[qx] \tag{9.1.5}$$

We see that the nonlinear term produces $\cos^3[qx]$. This can be reworked to give $\frac{1}{4}\cos[3qx] + \frac{3}{4}\cos[qx]$ from which the action becomes clear : the nonlinearity produces higher harmonics of any periodic function one puts in. These higher harmonics themselves will again produce higher harmonics, and we are left 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. Rather, inspired by the aforementioned higher harmonics generation, we will look for a solution of the form

$$u(x,t) = \sum_{n=1}^{\infty} a_{2n-1} \cos[(2n-1)(q_c x + \phi)] \quad , \tag{9.1.6}$$

in the next exercise, you may anticipate that $a_{2n-1}$ is of $\mathcal{O}(\varepsilon^{\frac{2n-1}{2}})$.

**Exercise 5** Comment on the appearance of the arbitrary phase-factor in the above expansion. Set $\phi$ to zero for the moment and take only the first two terms in the above expansion, and substitute this in the SH-equation. Collect order by order in $\varepsilon$, and show that the ordering of the $a_n$ in $\varepsilon$ is consistent. Furthermore, show that (at lowest order) we recover the following restriction on the wavenumber $q$

$$\sqrt{q_c^2 - \sqrt{\varepsilon}} \leq q \leq \sqrt{q_c^2 + \sqrt{\varepsilon}} \tag{9.1.7}$$

which defines the band.

Note that $a_1$ does indeed go to zero at the edge of the band (that is, near $q_m$), as we saw in one of the previous exercises. The result of our counting argument is therefore seen to hold, we find perturbatively a band (or 1-parameter family) of periodic solutions.

**Exercise 6** Draw the band in a $(q, \varepsilon)$ plot. To lowest order in $\varepsilon$, the band is parabolic (just expand expression (9.1.7) to see this). Can you argue why, for larger $\varepsilon$, the band bends off to the left?

Far as we may have gotten, we still don't know what will actually happen to the pattern above threshold, and it is precisely here that the amplitude formalism alluded to in the title of this chapter, will prove its worth.

## 9.2 The Amplitude Expansion

Central to the Amplitude approach is the observation that the pattern above, but near threshold (small $\varepsilon$), can be decomposed into a 'fast' and a 'slow' part, the meaning of these terms will become clear in a moment. To see this, consider a solution to the full SH-equation containing a mode with period $q$ :

$$u(x,t) \sim ce^{iqx} + c^* e^{-iqx} \quad . \tag{9.2.1}$$

This expression can be trivially reworked to give

$$u(x,t) = (ce^{i(q-q_c)x})e^{iq_c x} + (c^* e^{-i(q-q_c)x})e^{-iq_c x} \quad . \tag{9.2.2}$$

This in itself does not tell us much, but when we add the information you've probably obtained in exercise 9.1.3, that $|q - q_c| \leq \sqrt{\varepsilon}$, we see that the first exponential $\sim (q - q_c)x$ is very slowly varying compared to the one $\sim q_c x$. In other words, as long as $\varepsilon$ is small, the solutions to the full SH-equation can to good approximation be thought of as the critical mode, slowly modulated (at least as far as their *spatial* dependence is concerned). These modulations typically happen on a lengthscale $x_{\text{slowspace}} = 1/(q - q_c) \sim \varepsilon^{-1/2}$, which is the spatial scale on which the slow exponent becomes of $\mathcal{O}(1)$.

At the heart of the amplitude expansion now lies that we take advantage of this phenomenon, known as *separation of scales*, by explicitly separating the scales through the introduction of a separate slow coordinate $X$, defined by

$$X = \sqrt{\varepsilon}x \quad . \tag{9.2.3}$$

This slow coordinate is treated like an *independent* variable throughout the rest of the argument. It is important to remember that although the pattern *at threshold* is time-independent, we cannot rule out that above threshold it will become so. Since we are interested in expressing the patterns above threshold in terms of modulations of the critical pattern, we would like to obtain a similar scale-separation for the fast and slow timescales.

**Exercise 7** Does the the pattern above threshold depend on the *fast* timescale? Compare this to the case of the Hopf bifurcation. What is the typical timescale $t_{\text{slowtime}}$ on which you expect the slow temporal modulations to take place? (Hint : what is the growth-rate of the fastest growing mode?). Use your previous answers to argue that the correct slow time variable is

$$T = \varepsilon t \quad . \tag{9.2.4}$$

Let us now pause for a moment and shortly recap what we've found up to now.

- 1 Just above onset, the amplitude of the periodic solutions scales like $\sqrt{\varepsilon}$
- 2 For small $\varepsilon$, it makes sense to introduce slow scales $X = \sqrt{\varepsilon}x$ and $T = \varepsilon t$, on which the modulations of the basic pattern take place.

The above considerations now inspire the actual amplitude expansion, which reads for a solution to the full problem

$$u(x,t) = \sum_{i=1}^{\infty} \varepsilon^{n/2} U_n \quad . \tag{9.2.5}$$

An expansion of this form is what is known as a weakly nonlinear expansion, and is effectively a power series in the small parameter $\varepsilon$. The $U_n$ can be expressed as a product of a slow and a fast part[2] , and in particular,

$$U_1 = A_1(X,T)e^{iq_c x} + A_1^*(X,T)e^{-iq_c x} \quad . \tag{9.2.6}$$

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 especially that the slow amplitudes $A_n$ are complex quantities, and that they depend on the slow variables only. There is, as you might have already concluded in exercise 9.2.1, no dependence on the fast time, and the functions are $2\pi/q_c$-periodic in fast space, a fact we will need later on. We will also demand that the functions are bounded for large and small $x$. Although the construction of the amplitude equation is similar to the procedure in chapters 4 and 6, we will take it one step further here and see what happens beyond the lowest non-trivial order.

Now without further ado let us see what all of this implies for the SH-equation

$$\partial_t u(x,t) = \varepsilon u - (\partial_x^2 + q_c^2)^2 u - u^3 \quad \equiv \varepsilon u - \mathcal{L}u - u^3 \quad . \tag{9.2.7}$$

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 (9.2.6):

$$\begin{aligned} \partial_t &\rightarrow \varepsilon\partial_T \quad , \\ \partial_x &\rightarrow \partial_x + \sqrt{\varepsilon}\partial_X \quad , \end{aligned} \tag{9.2.8}$$

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 ($2\pi/q_c$-periodic) part of a function. The linear differential operator transforms under this change-of-variables as

$$\begin{aligned} \mathcal{L} &\rightarrow \left[ \underbrace{(\partial_x^2 + q_c^2)}_{\equiv \mathcal{L}_{\mathrm{f}}} + 2\varepsilon^{\frac{1}{2}}\partial_x\partial_X + \varepsilon\partial_X^2 \right]^2 \\ &= \left[ \mathcal{L}_{\mathrm{f}}^2 + 4\varepsilon^{\frac{1}{2}}\mathcal{L}_{\mathrm{f}}\partial_x\partial_X + \varepsilon(2\mathcal{L}_{\mathrm{f}} + 4\partial_x^2)\partial_X^2 + 4\varepsilon^{\frac{3}{2}}\partial_x\partial_X^3 + \varepsilon^2\partial_X^4 \right] \quad , \end{aligned} \tag{9.2.9}$$

[2]This observation is what makes this expansion completely different from expansion (9.1.6)!

**Exercise 8** Verify Eq. (9.2.9).

where we have given the purely fast part of the linear operator it's own name $\mathcal{L}_{\rm f}$ for future notational convenience. Substituting the amplitude expansion (9.2.5) in this equation brings the SH-equation to the following, rather cumbersome but as we shall see powerful, form

$$\sum_{n=1}^{\infty}\varepsilon^{\frac{n+2}{2}}\partial_T U_n = \sum_{n=1}^{\infty}\varepsilon^{\frac{n+2}{2}}U_n - \sum_{n=1}^{\infty}\varepsilon^{\frac{n}{2}}\Big[\mathcal{L}_{\rm f}^2 + 4\varepsilon^{\frac{1}{2}}\mathcal{L}_{\rm f}\partial_x\partial_X + \varepsilon(2\mathcal{L}_{\rm f}+4\partial_x^2)\partial_X^2 + 4\varepsilon^{\frac{3}{2}}\partial_x\partial_X^3 + \varepsilon^2\partial_X^4\Big]U_n - \sum_{l,m,n=1}^{\infty}\varepsilon^{\frac{l+m+n}{2}}U_lU_mU_n \quad . \tag{9.2.10}$$

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

$$\mathcal{O}(\varepsilon^{\frac{1}{2}}) \;:\; -\mathcal{L}_{\rm f}^2U_1 = 0 \quad , \tag{9.2.11}$$

$$\mathcal{O}(\varepsilon^{1}) \;:\; -\mathcal{L}_{\rm f}^2U_2 - 4\mathcal{L}_{\rm f}\partial_x\partial_X U_1 = 0 \quad , \tag{9.2.12}$$

$$\mathcal{O}(\varepsilon^{\frac{3}{2}}) \;:\; \partial_T U_1 = U_1 - \big[\mathcal{L}_{\rm f}^2U_3 + 4\mathcal{L}_{\rm f}\partial_x\partial_X U_2 + (2\mathcal{L}_{\rm f}+4\partial_x^2)\partial_X^2U_1\big] - U_1^3 \tag{9.2.13}$$

**Exercise 9** Verify Eqs. (9.2.11), (9.2.12) and (9.2.13). Can you write down the equation obtained at $\mathcal{O}(\varepsilon^2)$?

Our task now is to solve these equations one by one, using the results of the previous order to solve the next order. Eq. (9.2.11) determines the linearized solution (cf.(9.2.6))[3]:

$$\mathcal{L}_{\rm f}U_1 = 0 \quad \rightarrow \quad U_1 = A_1(X,T)e^{iq_cx} + A_1^*(X,T)e^{-iq_cx} \quad , \tag{9.2.14}$$

where the (complex) function $A_1(X,T)$ is still completely arbitrary, since $\mathcal{L}_{\rm f}$ works only on the fast scales. The next order also yields no information about the slow-scale dynamics

$$\mathcal{L}_{\rm f}U_2 = 0 \quad \rightarrow \quad U_2 = A_2(X,T)e^{iq_cx} + A_2^*(X,T)e^{-iq_cx} \quad , \tag{9.2.15}$$

but the third equation produces the first interesting result.

**Exercise 10** Show that the third equation can be cast into the following form

$$\mathcal{L}_{\rm f}^2U_3 = \left[e^{iq_cx}\left\{-\partial_T + 1 + 4q_c^2\partial_X^2 - 3|A_1|^2\right\}A_1 + c.c.\right] - \left[e^{3iq_cx}A_1^3 + c.c.\right] \quad . \tag{9.2.16}$$

In order to get this, you will need to use the results obtained from solving the first and second equation.

[3]Note that actually, the first order demands that $\mathcal{L}_{\rm f}^2U_1 = 0$, and not $\mathcal{L}_{\rm f}U_1 = 0$. Although the former implies the latter, the 'squared' equation also posesses solutions of the type $U_1 \sim xAe^{iq_cx} + c.c.$. These are however bounded nor periodic, and are therefore discarded here

The operator $\mathcal{L}_{\rm f}$ is linear, and acts on the fast scales only, and the RHS of Eq. (9.2.16) is of the form $F(X,T)e^{iq_cx} + G(X,T)e^{3iq_cx} + c.c.$. This implies that the fast space dependence of $U_3$ is also a sum of these exponentials, and this is where, as promised, the higher harmonics enter our expansion naturally

$$U_3 = A_3(X,T)e^{iq_cx} + B_3(X,T)e^{3iq_cx} + c.c. \quad . \tag{9.2.17}$$

Note that the amplitude expansion for this particular nonlinearity generates *in principle* not only odd, but also even higher 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 non-Boussinesq effect inspired $u^2$ term in our SH model.

Sofar 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

**Theorem 3 [Fredholm]** Let $\mathcal{L} = \mathcal{L}^\dagger$ be a self-adjoint operator. Then the equation

$$\mathcal{L}u = v \quad \mathcal{L} : S \to S \tag{9.2.18}$$

is solvable if and only if the vector $v \in S$ is orthogonal to the zero-space $\mathrm{Ker}[\mathcal{L}] = \{w \in S | \mathcal{L}w = 0\}$ of the operator $\mathcal{L}$:

$$\langle w|v\rangle = 0 \quad \forall w \in \mathrm{Ker}[\mathcal{L}] \quad \bigstar \tag{9.2.19}$$

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

$$\langle a|b\rangle = \int_0^{\frac{2\pi}{q_c}} \mathrm{d}x a^* b \quad . \tag{9.2.20}$$

**Exercise 11** Prove the 'if' part of the Fredholm theorem, that is prove that the equation has a solution if the theorem is satisfied. The 'only if' part is a lot more work...

**Exercise 12** What is the kernel of our fast operator $\mathcal{L}_{\rm f}$? Apply the Fredholm theorem to equation (9.2.16), and perform the integrations to obtain the *solvability condition*

$$\partial_T A_1(X,T) = A_1 + 4q_c^2\partial_X^2 A_1 - 3|A_1|^2 A_1 \quad , \tag{9.2.21}$$

which is exactly what we set out to obtain : an equation for the dynamics of the slow modulations, valid for small $\varepsilon$. This is called an amplitude equation, for reasons you will probably understand by now.(Of course, one also obtains an equation for $A_1^*$, which is just the *c.c.* of the above equation)

The dynamics of the amplitude $A_1$ is therefore governed by a second order, nonlinear PDE. When one has obtained a solution to this equation (which is the tricky part), the amplitude of the third harmonic $B_3(X,T)$ is also known. Combining Eq. (9.2.16) with (9.2.17), we find

$$B_3(X,T) = \frac{A_1^3}{8q_c^2} \tag{9.2.22}$$

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 $\varepsilon$ does not appear in our equation shows that we've chosen the correct scales of space, time and amplitude. We prefer however to keep $\varepsilon$ explicit, in order to get a good idea of what happens as it is increased through zero. After some transformations of space, time and amplitude, we arrive finally at the amplitude equation in its standard form.

$$\partial_t A = \varepsilon A + \partial_x^2 A - |A|^2 A \quad , \tag{9.2.23}$$

where we have dropped the subscripts, and transformed back to the fast variables.

**Exercise 13** What transformations need to be done to get (9.2.21) to the form (9.2.23)?

One might wonder what happens to the higher order terms in the expansion. As you may have already obtained in exercise 9.2.2, the equation at $\mathcal{O}(\varepsilon^2)$ reads

$$\begin{aligned} \mathcal{L}_{\mathrm{f}}^2 U_4 &= \left[e^{iq_c x}\left\{-\partial_T A_2 + A_2 + 4q_c^2\partial_X^2 A_2 - 4iq_c\partial_X^3 A_1 - 3A_1^2 A_2^* - 6|A_1|^2 A_2\right\}\right. \\ &\quad \left. + c.c.\right] + \left[e^{3iq_c x}\left\{96iq_c^3\partial_X B_3 - 3A_1^2 A_2\right\} + c.c.\right] \quad . \end{aligned} \tag{9.2.24}$$

We conclude from this that $U_4$ can again be written as

$$U_4 = A_4(X,T)e^{iq_c x} + B_4(X,T)e^{3iq_c x} + c.c. \quad , \tag{9.2.25}$$

while applying Fredholm's theorem produces the dynamical equation for $A_2$ :

$$\partial_T A_2 = \left[1 - 6|A_1|^2\right] A_2 + 4q_c^2\partial_X^2 A_2 - 4iq_c\partial_X^3 A_1 - 3A_1^2 A_2^* \quad . \tag{9.2.26}$$

**Exercise 14** Show that the linear coefficient of $A_2$ $(1 - 6|A_1|^2)$ is negative everywhere in the band of allowed wavenumbers. This is nice because otherwise the third harmonic would itself be unstable.

The nice thing about this equation is that it is a linear one, and it depends only on $A_1$. Once one has solved the equation for $A_1$ (which is, unfortunately, in general not possible), the solution can be put into this equation, which is then easily solved. That is to say, the dynamics of $A_2$ is completely determined by $A_1$, a phenomenon known as *slaving*. In general, higher order equations will always be of the form

$$\mathcal{L}_{\mathrm{f}}^2 U_n = F[U_{n-1}, \ldots, U_1] \quad , \tag{9.2.27}$$

and every subsequent amplitude is slaved by the previous ones. Not only does this hold for the $A_i$, the $B_i, C_i, D_i \ldots$ (*i.e.* the amplitudes of the higher harmonics) are also slaved. It is therefore fair to say that $A_1$ contains all the physical information, the higher order terms of the expansion are all driven by it.

**Exercise 15** When we allow for non-Bousinesq effects to play a role, we have seen that we need to include terms that break the up-down symmetry. In this exercise, we will briefly investigate the effect of such a term on the level of the amplitude equations. Write down the analogue to Eq. (9.2.10) for the following modified Swift-Hohenberg equation

$$\partial_t u(x,t) = \varepsilon u - (\partial_x^2 + q_c)^2 u + \alpha u^2 - \beta u^3 \quad . \tag{9.2.28}$$

Collect the various orders in $\varepsilon$, and solve order by order. You will find that, as opposed to the example in the text, already the $\mathcal{O}(\varepsilon)$ equation is non-trivial. When done correctly, your $\mathcal{O}(\varepsilon)$ equation should read

$$\mathcal{L}_{\mathrm{f}}^2 U_2 = 2\alpha|A_1|^2 + \left\{\alpha A_1^2 e^{2iq_c x} + c.c.\right\} \quad , \tag{9.2.29}$$

which implies that $U_2$ is of the form

$$U_2(X,T) = O_2(X,T) + A_2(X,T)e^{iq_c x} + B_2(X,T)e^{2iq_c x} + c.c. \quad . \tag{9.2.30}$$

Solve the $U_2$ equation for general $A_1$. What all of this implies is that the inclusion brings the even harmonics into play as well. As you can see however, these are also slaved. This time though, the even harmonics are driven by symmetry breaking expressions like $A_1^2$!. When the amplitude expansion is carried to higher orders, this manifests itself as symmetry-breaking terms like $|A|^2\partial_X A$. At lowest order, they only renormalize the coefficient of $|A|^2 A$. As the industrious student may want to verify, the dynamical equation for $A_1$ becomes

$$\partial_T A_1 = A_1 + 4q_c^2\partial_X^2 A_1 - 3\left(\beta - \frac{38\alpha^2}{27q_c^4}\right)|A_1|^2 A_1 \tag{9.2.31}$$

Here we can see that when

$$\alpha^2 > \frac{27\beta q_c^4}{38} \quad , \tag{9.2.32}$$

the coefficient of the cubic term becomes positive, and the nonlinearity is no longer saturating. In order to account for this we will need to include higher order terms in the amplitude equation (like $|A_1|^4 A_1$). The inclusion of these terms will make the bifurcation *subcritical*. Equation (9.2.23) 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. The scientists who first derived it in the context of nonlinear hydrodynamics (Newell and Whitehead, 1969), did therefore not get the equation named after them, but instead it is 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. (9.2.5) 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

$$\tau_0 \partial_t A = \varepsilon A + \xi_0^2 \partial_x^2 A - g_0 |A|^2 A \quad , \tag{9.2.33}$$

in which the coefficients $\tau_0, \xi_0$ and $g_0$ reflect the physical properties of the actual system under study. They are *the same* correlation length and time discussed in section 8.9, and here's how to compute them.

**Exercise 16** Consider Eq. (9.2.33). Suppose that the underlying system has a linear dispersion relation $s(k)$.Of course, when the amplitude approach is applied to a given system or equation, the linear terms in the amplitude will reproduce the expansion of the dispersion relation about the critical value of that equation. We can use this as a shortcut to obtain these coefficients directly from the dispersion relation of the full problem. . Derive the linear dispersion relation of Eq. (9.2.33) by setting $A \sim e^{s(k)t+ikx}$, with $s = \sigma + i\omega$. Differentiate once with respect to $\varepsilon$ to obtain

$$\frac{1}{\tau_0} = \left. \frac{\partial \sigma(k)}{\partial \varepsilon} \right|_{k_c} \quad . \tag{9.2.34}$$

Use a similar procedure to derive the relation for the lengthscale

$$\xi_0^2 = - \left. \frac{\tau_0}{2} \frac{\partial^2 \sigma(k)}{\partial k^2} \right|_{k_c} \quad . \tag{9.2.35}$$

What this shows is that the coefficients that set the basic length- and timescale for the system are essentially determined by the linear dispersion relation of the underlying equations.

**Exercise 17** Verify that eqs. (9.2.34,9.2.35) give the same $\xi_0$,$\tau_0$ that come out of the amplitude expansion.

The fact that the general form of the amplitude equation for a supercritical stationary bifurcation is always the same has earned Eq. (9.2.23) the title Normal Form. Other types of bifurcations also have their own normal forms, but we will not comment on that here.

In the next section, we will get to work with the RGLE, and see what happens in the weakly nonlinear regime. New instabilities will show up (called *secondary instabilities* because they are instabilities of the basic pattern, which itself was born through an instability of the $u = 0$ state), and we will see some experimental evidence of exactly how good the amplitude approach performs.

## 9.3 Suggested Further Reading

The derivation of the Real Ginzburg-Landau equation follows the lines of [HSH]. Interested readers may want to consult [Man] for a slightly different formulation, and [CH] for a wealth of information on all kinds of pattern-formation related topics. A short and sweet discussion of normal forms (very global) can be found in [GM]. The effect of the inclusion of other terms in the SH-equation on the ensuing Amplitude equations can be found in the appendix of [vS].

*[HSH]* M. van Hecke, W. van Saarloos and P. C. Hohenberg, *Amplitude equations for pattern forming systems*, in: Fundamental Problems in Statistical Mechanics VIII, H. van Beijeren and M. H. Ernst, eds. (North Holland, Amsterdam, 1994).

*[Man]* P. Manneville, *Dissipative Structures and Weak Turbulence* (Academic Press, Boston, 1990).

*[CH]* M. C. Cross and P. C. Hohenberg, *Pattern formation outside of equilibrium*, Rev. Mod. Phys. **65**, 851 (1993).

*[GM]* P. Manneville, *Overview*, in: Hydrodynamics and Nonlinear Instabilities, C. Godrèche and P. Manneville, eds. (Cambridge University Press, Cambridge 1998).

*[vS]* W. van Saarloos, Phys. Rev. A **39**, 6367 (1989).

# Chapter 10

# Implications of the Amplitude Equation Description : Analysis of Pattern Dynamics near Threshold

In the previous chapter, we have derived the (lowest order) Amplitude equation for the Swift-Hohenberg equation. The Amplitude expansion is essentially a double expansion in $\varepsilon$ and the higher harmonics, when we write

$$u(x,t) = \sum_{n=1,m=0}^{\infty} \varepsilon^{\frac{n}{2}} A_{nm} e^{imq_c x} + c.c. \quad . \tag{10.0.1}$$

The $A_{mn}$ are assumed to be functions of the slow variables $(X,T)$ only. We have seen that we obtain a PDE for the $A_{11}$ through a solvability condition, as well as PDE's for the $A_{n1}$ that are slaved by $A_{11}$. The amplitudes of the harmonics *i.e.* the $A_{nm}, m \neq 1$ turned out to be slaved as well. (Note the slight difference in notation compared to the previous chapter). To lowest order, we found the Real Ginzburg-Landau (RGL) equation

$$\partial_t A = \varepsilon A + \partial_x^2 A - |A|^2 A \quad , A \in \mathbb{C} \quad , \tag{10.0.2}$$

which we'll look at in some more detail in the present chapter. Let us start our investigations with the simplest non-trivial solutions to the RGL equation, the so-called *phase-winding* solution.

## 10.1 Phase Winding Solutions and their Stability

The RGLE (10.0.2) admits spatial plane-wave solutions of the form

$$A = a_0 e^{iqx} \quad . \tag{10.1.1}$$

These are also referred to as phase-winding solutions (when one plots a function of the form (10.1.1) in 3D like $(\mathtt{Re}(A), \mathtt{Im}(A), x)$, the result will be a 'corkscrew' or helical curve, hence the name). In terms of solutions to the Swift-Hohenberg equation, they describe stationary periodic patters 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 $q_c$, a property we expect to find in the RGLE as well.

**Exercise 1** Study the existence of phase winding solutions by substituting Ansatz (10.1.1) in the RGLE. What is the relation between the amplitude $a_0$ and the wavenumber $q$? Show that these solutions are again restricted to lie in a band whose width is proportional to $\sqrt{\varepsilon}$ (as we already found to lowest order in the previous chapter!). The band is centered about $q = 0$, why is this?

The appearance of periodic solutions with wavenumbers within a certain band is thus recovered, also on the level of the amplitude equations. The existence of certain types of solutions however says nothing about their stability, and as we have seen before, states that are likely to be reached from physical initial conditions are the stable ones. Once again we therefore whip out our trusty linear stability analysis, and apply it to the phase winding solutions. If we first split the (complex) amplitude $A$ into modulus and phase (both real)

$$A(x,t) = a(x,t)e^{i\varphi(x,t)} \quad , \tag{10.1.2}$$

and substitute this in eq. (10.0.2), we arrive at the equivalent set of equations

$$\begin{aligned} \partial_t a &= \varepsilon a + \partial_x^2 a - a^3 - a(\partial_x \varphi)^2 \quad , \\ a\partial_t \varphi &= 2(\partial_x a)(\partial_x \varphi) + a\partial_x^2 \varphi \quad . \end{aligned} \tag{10.1.3}$$

**Exercise 2** Check Eqs. (10.1.3).

The phase-winding solution is characterised by $a = a_0, \varphi = qx$, and as a starting point for our stability analysis we study, as usual, the response to periodic perturbations :

$$\begin{aligned} a &= a_0 + \delta a e^{iQx+\lambda t} \quad , \\ \varphi &= qx + \delta\varphi e^{iQx+\lambda t} \quad . \end{aligned} \tag{10.1.4}$$

**Exercise 3** Substitute (10.1.4) in Eqs. (10.1.3), and linearize in $\delta a, \delta\varphi$. Write the resulting equation in matrix-form as

$$\lambda \begin{pmatrix} \delta a \\ \delta\varphi \end{pmatrix} = \mathbf{D} \cdot \begin{pmatrix} \delta a \\ \delta\varphi \end{pmatrix} \quad , \tag{10.1.5}$$

write out the eigenvalue problem

$$\det[\mathbf{D} - \lambda \mathbf{I}] = \mathbf{0} \quad , \tag{10.1.6}$$

to yield

$$\lambda^2 + 2\lambda(Q^2 + a_0^2) + Q^2(2(\varepsilon - 3q^2) + Q^2) = 0 \quad . \tag{10.1.7}$$

**Exercise 4** We are interested in the stability of phase-winding solutions, and therefore we need only the signs of the two roots of this equation. When one of them (or both) is positive, we conclude that the state concerned is unstable, since the associated perturbation will grow in time. Show that we find one unstable mode if $q^2 > \varepsilon/3$, and no unstable modes if $q^2 < \varepsilon/3$, and that the instability is a long wavelength instability, *i.e.* that it first occurs for small $Q$.

As you have shown in the previous exercises, the phase-winding solutions go linearly unstable for $|q| > \sqrt{\varepsilon/3}$. This kind of instability is called a *secondary* instability, since the phase-winding solutions themselves emerged via a linear instability of the $A = 0$ state. It is called the Eckhaus instability, and has some interesting consequences for the dynamics. In the previous chapter, we found a band of allowed wave numbers for the periodic states, and now it turns out that a smaller band within the allowed band is actually stable. There is another way of looking at the same problem, which is also based on the linear stability analysis.

**Exercise 5** Consider the eigenvalue problem (10.1.5). Show that in the limit of $Q \to 0$ (long wavelength), the two eigenvalues are given by

$$\begin{aligned} \lambda_+ &= \mathcal{O}(Q^2) \quad , \\ \lambda_- &= -2a_0^2 + \mathcal{O}(Q^2) \quad . \end{aligned} \tag{10.1.8}$$

Calculate the eigenvectors corresponding to these two eigenvalues, and argue that in the long wavelength limit, phase perturbations $\delta\varphi$ will relax arbitrarily slow, while amplitude perturbations still relax exponentially fast. What is the reason for the fact that a phase perturbation with $Q = 0$ does not relax at all (*i.e.* has eigenvalue 0)?

The difference in relaxation times suggests that there is again a separation of timescales present in this system, the fast one being the timescale on which the magnitude of the amplitude relaxes, and the slow one associated with the relaxation of the phase. As we have seen, it is actually translation invariance that is responsible for this phenomenon. On the fast timescale, $a$ relaxes to the value dictated by the phase

$$a^2 = \varepsilon - (\partial_x \varphi)^2 \quad . \tag{10.1.9}$$

**Exercise 6** Verify Eq. (10.1.9). While for the phase winding solutions we had $a_0^2 = \varepsilon - q^2$, with $q$ the wavenumber of the phase winding solution, we see that here $q$ is replaced by $\partial_x \varphi$. The derivative of $\varphi$ is usually referred to as the *local* wavenumber. Comment on the difference between the local wavenumber and the 'global' wavenumber, and show that for phase winding solutions, the two are equivalent.

Eq. (10.1.9) is again an example of slaving. Due to the different timescales present, $a$ will reach its asymptotic value, which is dictated by the (gradient of the) phase, exponentially fast. One says that the amplitude follows the phase *adiabatically*. The only important quantity is therefore the phase, and substituting Eq. (10.1.9) in Eq. (10.1.3) yield the dynamic equation for the phase, which can be put in the following form

$$\partial_t \varphi(x,t) = D(q_{\rm loc}) \partial_x^2 \varphi \tag{10.1.10}$$

**Exercise 7** Derive Eq. (10.1.10), and give an expression for the (local wavenumber dependent) diffusion coefficient $D(q_{\rm loc})$. Show that the diffusion coefficient changes sign exactly at the Eckhaus boundary.

When a diffusion equation has a positive diffusion coefficient, its fundamental solutions are simply the well-known Gaussian ones. Diffusion with positive coefficient will

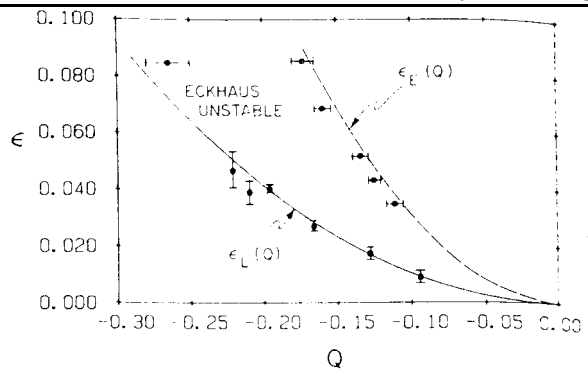


Figure 10.1. Measured stability boundaries as a function of the distance $\varepsilon$ above onset and the dimensionless wavenumber $Q$. The solid line is a parabolic fit to the data for the marginal stability curve $\varepsilon_L(Q)$, while the dashed line is the prediction for the Eckhaus boundary $\varepsilon_E(Q) = 3\varepsilon_L(Q)$. Figure taken from [LG]

try to even out any gradients present in the initial condition. A sharply peaked initial condition will eventually 'flatten out' to a homogenous distribution. Diffusion with a *negative* diffusion coefficient does just the opposite. As it favors sharp gradients, an initially flat state (with a small perturbation) will start to heap up at some point, until the final state, an infinitely narrow peak ('delta function') is achieved. Therefore, we expect the phase diffusion equation to produce qualitatively very different behaviors depending on the sign of the diffusion coefficient. The fact that in our case, the diffusion coefficient passes through zero at $q^2 = \varepsilon/3$ is no coincidence, and shows that this is just another way of interpreting the Eckhaus instability.

What now will happen to a pattern that has an Eckhaus-unstable wavenumber? Somehow it has to get back into the stable band, and in order to do so it has to either lose or gain rolls. Rolls can only be created or annihilated in pairs, which corresponds to a change in the phase of magnitude $2\pi$. However, since the phase is well-defined and continuous whenever $|A| \neq 0$, the only way for it to change is when at the locus of the creation/annihilation of the roll pair $|A| = 0$. Then the phase is undefined, and is therefore allowed to 'slip' by $2\pi$. Points where this happens are called phase slip centers. In simulations one sees that phase slips are indeed the prime mechanism for getting $q$ back in the stable band.

The Eckhaus instability is not just a theoretical concept, it has also been widely studied experimentally. Some results of this are plotted in Fig. 10.1, from an article by Lowe and Gollub (1985), where they measured the stability boundaries we have also derived here, the linear stability curve and the Eckhaus curve. These experiments were conducted on a convecting layer of liquid crystals. The experimentally observed values are seen to agree quite well with predictions from amplitude equations. Phase slips are also observed, and described in the same paper.

## 10.2 Lyapunov Functionals and Dynamics

The Swift-Hohenberg equation has one very interesting property, which is that it can be derived from a so-called Lyapunov functional $\mathcal{F}_{\rm SH}[u]$, in that it can be written as

$$\partial_t u = -\frac{\delta \mathcal{F}_{\rm SH}[u]}{\delta u} \quad , \tag{10.2.1}$$

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

$$\mathcal{F}_{\rm SH}[u] = \frac{1}{2}\int \mathrm{d}x \left\{ \left[(\partial_x^2 + q_c^2)u\right]^2 - \varepsilon u^2 + \frac{1}{2}u^4 \right\} \quad . \tag{10.2.2}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}_{\rm SH} \leq 0 \quad . \tag{10.2.3}$$

**Exercise 8** Verify Eq. (10.2.1), with $\mathcal{F}_{\rm SH}[u]$ as in (10.2.2). Prove relation (10.2.3).

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 (the physicists have probably calculated it at some point for the harmonic oscillator), 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 $\mathcal{F}_{\rm SH}[u]$) implies that these states correspond to minima of $\mathcal{F}_{\rm SH}[u]$. 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 there may be some dynamics that leave $\mathcal{F}_{\rm SH}[u]$ invariant.

**Exercise 9** Since the SH-equation itself derives from a Lyapunov functional, it will come as no surprise that the RGLE can also be derived from a Lyapunov functional. Note that this works in one direction, but nonvariational equations may actually produce amplitude equations that do posess a Lyapunov functional. Since $A$ is a complex quantity however, it looks slightly different :

$$\begin{aligned} \frac{\partial A}{\partial t} &= -\frac{\delta \mathcal{F}_{\rm GL}}{\delta A^*} \quad , \\ \frac{\partial A^*}{\partial t} &= -\frac{\delta \mathcal{F}_{\rm GL}}{\delta A} \quad , \end{aligned} \tag{10.2.4}$$

with the Ginzburg-Landau Lyapunov functional now a function of both $A$ and $A^*$. Can you find the expression for $\mathcal{F}_{\rm GL}$?

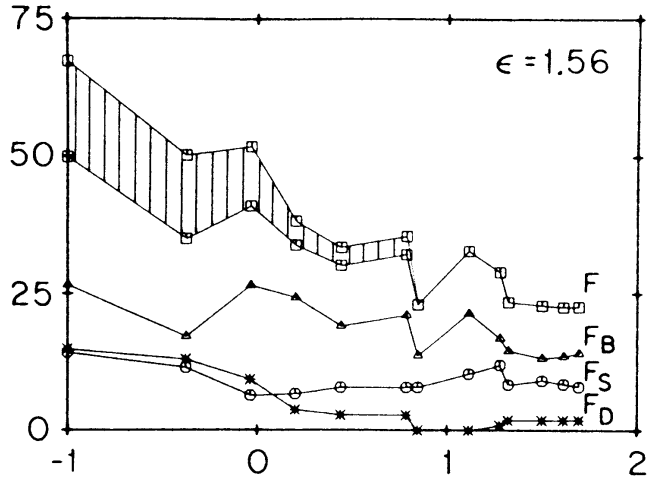


Figure 10.2. Swift-Hohenberg Lyapunov functional as a function of time (logarithmically) at $\varepsilon = 1.56$, as extracted from an experiment on RB-convection in a cylindrical geometry. The separate contributions due to the bulk ($F_B$), surface ($F_S$) and defects ($F_D$) are also shown. Slow wavenumber adjustments at the end of this run lead to fluctuations is $F$. Figure taken from [HG].

Although the Lyapunov functional does appear to be a purely theoretical construct, it has actually been measured experimentally (Heutmaker and Gollub, 1987) for convective patterns in a cylindrical geometry. The results of this experiment are plotted in Fig. 10.2, and as you can see there is indeed an overall trend towards decreasing $\mathcal{F}_{\rm SH}[u]$, with occasional regions of increase. Note that time is on a log scale, and the plot suggests that towards the end of the measurement, $\mathcal{F}_{\rm SH}[u]$ changes very little.

**Exercise 10** Why do you think $\mathcal{F}_{\rm SH}[u]$ does not decrease uniformly in an experiment like it does in the SH-equation?

## 10.3 An Exact Solution to the RGL-Equation

Sofar, we have seen one exact solution, the phase-winding solution. Now we will derive one more exact solution. Experimentally it is relevant to study the effects of boundaries on convection. As we have seen, the usual boundary conditions for a problem concerning flows with walls are either the stick or slip boundary conditions. Let us mimic these in our amplitude equation by demanding $A(0,t) = 0$. Furthermore, so as not to complicate things, we will look for a stationary, real solution.

**Exercise 11** rescale the stationary RGL-equation, so as to get rid of $\varepsilon$. In order to do this, you will need to rescale both the $x$ and $A$. Make an Ansatz of the form

$$A(x) = \tanh(\beta x) \quad , \tag{10.3.1}$$

and solve the equation. The tanh ensures that the boundary condition at $x = 0$ is automatically fulfilled. Plot the resulting profiles for various values of $\varepsilon$

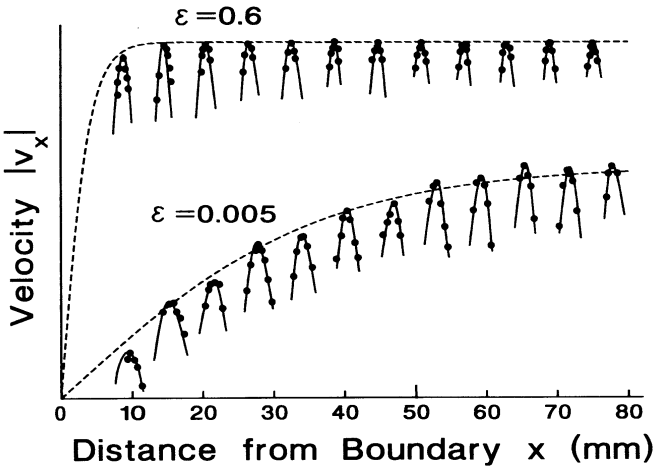


Figure 10.3. Suppression of convection near a sidewall: comparison of predictions of amplitude equation with experiments, for two values of the control parameter $\varepsilon$. Solid points represent measured values of the magnitude of the horizontal component of the fluid velocity at the mid-plane. The lines drawn through the points are guides to the eye to display the roll structure. The dashed lines are the *predictions* of the amplitude equation; $A = A_0 \tanh(x/\sqrt{2}\xi)$, with $\xi = \varepsilon^{-1/2}\xi_0$ and $\xi_0 = 0.385$, with the asymptotic amplitude $A_0$ fit to the data. Figure taken from [CH].

This solution has also been compared to experiments (Pocheau *et al*, 1987), and the results of these are plotted in Fig. 10.3. As you can see, the agreement is excellent. The scaling factors of the plot are composed from the prefactors of the different terms in the amplitude equation. As was already remarked in the previous chapter, these prefactors indicate the basic length- and timescales in the pattern, and are determined by the linear dispersion relation of the full equations. The basic time- ($\tau_0$) and lengthscale ($\xi_o$) can be computed from the dispersion relation as follows

$$\frac{1}{\tau_0} = \left.\frac{\partial\sigma}{\partial\varepsilon}\right|_{\varepsilon_c,k_c}; \quad \xi_0^2 = -\left.\frac{\tau_0}{2}\frac{\partial^2\sigma}{\partial q^2}\right|_{\varepsilon_c,k_c}, \tag{10.3.2}$$

as you have shown in the previous chapter. Note also the excellent agreement between the $\xi_0$ computed in chapter 8, where we found $\xi_0 = \sqrt{0.1479} = 0.3846$.

As we shall see in the next section, knowledge of these basic scales, supplied with symmetry considerations, can get you a long way in getting out the correct equations without doing the explicit calculation.

## 10.4 Symmetries and Amplitude Equations

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.

Therefore, for *any* pattern forming system with a stationary bifurcation and a finite wavelength instability will produce an equation of the same form. The only difference will be system-dependent prefactors that give the basic length- and timescales, and we have seen that these can be extracted from the linear dispersion relation.

**Exercise 12** Our model (in the 1D case) has two basic symmetries, reflection symmetry and translation invariance. Argue that translation invariance implies invariance under $A \to A^{i\varphi}$. Use the linear dispersion relation (of the SH-equation) to determine how space and time scale with $\varepsilon$ (you've done this before in the previous chapter). Write down to lowest order in $\varepsilon$ all the terms that are allowed. Then proceed by discarding all those that are forbidden by either one of the symmetries. Give all terms arbitrary prefactors, and scale away $\varepsilon$ and the prefactors to end up with the RGL equation.

This last exercise demonstrates that (at least in the case of the SH-equation) the 'if it's not forbidden, it is allowed' tactics produce the correct equation. Somewhat less trivial is the 2-D case with hexagons (you've seen this before in chapter 6). For these hexagons, the fields are expressed in terms of amplitudes as

$$\text{fields} \sim A_1 e^{i\vec{k}_1^c \cdot \vec{r}} + A_2 e^{i\vec{k}_2^c \cdot \vec{r}} + A_3 e^{i\vec{k}_3^c \cdot \vec{r}} \quad , \tag{10.4.1}$$

in which $\vec{k}_1^c + \vec{k}_2^c + \vec{k}_3^c = 0$, and the $\vec{k}_i^c$ are at $120^o$ angles to each other. A pattern in which all three $A_i$ are of equal magnitude corresponds therefore to a perfect hexagonal pattern.

**Exercise 13** Show that translation invariance in the $\vec{k}_1^c$-direction implies invariance under

$$A_1 \to A_1 e^{i\varphi}, \quad A_2 \to A_2 e^{-i\frac{\varphi}{2}}, \quad A_3 \to A_3 e^{-i\frac{\varphi}{2}} \quad . \tag{10.4.2}$$

Assume the same scaling of $x, t$ and $A$ as in the previous exercise, and list all symmetry-allowed terms. Show that one is left with

$$\partial_t A_1 = (\hat{e}_1 \cdot \vec{\nabla})^2 A_1 + \varepsilon A_1 \pm C_2 A_2^* A_3^* \pm C_3[\text{cubic terms}] \quad (\text{cyclic}) \quad . \tag{10.4.3}$$

**Exercise 14** Argue that in the Boussinesq approximation, the quadratic terms are absent.

As the name would suggest, the Boussinesq approximation is only an approximation. In the full equations, our guessing argument suggests therefore that quadratic terms, although possibly small, will always be present. One striking consequence of this stems from the fact that an amplitude equation with quadratic terms does not undergo a supercritical bifurcation, but a *subcritical* one. One therefore expects any 2-D system with the above symmetries (not the Boussinesq one of course) to display hexagons already below threshold. This prediction has also been studied experimentally, and has been shown to hold. Hexagons *always* show up subcritically and we now know that this is a consequence of basic symmetries of the system, without the somewhat artificial Boussinesq one. Note that in chapter 6, the equation did *not* have the $A \to -A$ symmetry either.

## 10.5 The Complex Ginzburg-Landau Equation

As we have seen in this and the last chapter, the Amplitude formalism provides a useful tool in studying nonlinear equations near onset. For more complicated problems (like actual RB-convection), the procedure is analogous and produces qualitatively the same amplitude equations. We have only looked at a stationary bifurcation (with a critical frequency equal to zero), but the formalism can be carried out for other types of bifurcations as well. One obvious generalization is to the case of a Hopf bifurcation. You have actually already done this in one of the previous chapters, and you will remember that this yielded the so-called Complex Ginzburg-Landau equation (CGL)

$$\partial_t A = \varepsilon A + (1 + ic_1)\partial_x^2 A - (1 - ic_3)|A|^2 A \quad . \tag{10.5.1}$$

The inclusion of complex parts in the coefficients brings a new phenomenon into play.

**Exercise 15** Study phase-winding solutions in the CGL equation. Derive the dispersion relation, and interpret the coefficients $c_1$ and $c_3$. As you see, the frequency itself now varies with $k$ and $a$. Show that $c_1$ measures the dependence of $\omega$ on $k$, while $c_3$ is responsible for the variation of $\omega$ with $a$. This is called *dispersion*, the new phenomenon referred to above. Also, show that the RGL does not have dispersion.

The difference between RGL and CGL might appear to be small, but is in fact huge. One of the most dramatic consequences is that the inclusion of complex coefficients breaks the variational structure, that is, it is no longer possible to find a Lyapunov functional. The dynamics are, for $c_1 \neq c_3$, no longer relaxational, and chaotic states occur. It posesses also an Eckhaus-like instability (usually referred to as the Benjamin-Feir instability in CGL-context), with the added complication that the width of the stable band depends on the coefficients, and this width can even shrink to zero! This means that there is no stable wavenumber possible. This combined with the nonvariational structure leads to persistent chaotic behavior. In fact, the CGL equation is widely used to model spatiotemporal chaos in a number of experimental systems. Another very interesting property of the CGL is that, in the limit of $c_i \to \infty$, it reduces to the so-called non-linear Schrödinger equation, another widely studied equation that posesses soliton solutions.

## 10.6 Suggested Further Reading

The discussion of the RGLE was mostly inspired by [MvH]. For those interested in more information about both the experimental and theoretical methods in pattern formation I can highly recommend [CH]. The two other papers form which the plots in this chapter were taken are [LG] and [HG].

*[MvH]* M. van Hecke, Ph.D. Thesis *The Amplitude description of nonequilibrium patterns*, Universiteit Leiden, 1996.

*[CH]* M. C. Cross and P. C. Hohenberg, *Pattern formation outside of equilibrium*, Rev. Mod. Phys. **65**, 851 (1993).

*[LG]* M. Lowe and J.P. Gollub, Phys. Rev. Lett. **23**, 2575 (1985).

*[HB]* M.S. Heutmaker and J.P. Gollub, Phys. Rev. A **35**, 242 (1987).

# Chapter 11

# Front formation in the nonlinear diffusion equation.

A front is the localized region separating two distinct phases. We will see that usually one phase invades the other leading to a traveling front. In this chapter and the next we will be concerned with *planar* fronts, i.e. when the junction between the two phases is straight. We do this in order to avoid more complicated effects such as that of curvature on the properties of fronts.

## 11.1 The Nonlinear Diffusion Equation (NLDE)

The mathematical treatment of planar fronts is simplified due to the fact that only one direction (the direction normal to the front) is relevant. The front is described by a localized variation of the "phase-field" $\phi(x,t)$ which takes constant but distinct values in the two phases separated by the front. Spatio-temporal evolution of $\phi(x,t)$ is governed by partial differential equations in one spatial dimension. As an example we consider the so-called *nonlinear diffusion equation* (NLDE),

$$\frac{\partial \phi}{\partial t} = D\frac{\partial^2 \phi}{\partial x^2} + f(\phi). \tag{11.1.1}$$

The above equation (with slightly different notations) was already introduced in section 5.3 with the concentration of a reacting species denoted by $\phi(x,t)$. The diffusion term models the molecular diffusion at a macroscopic level and leads to smoother profiles. The function $f$ represents the reaction terms. As we will see, the feature which is important in order to allow for traveling front solutions is the multi-stability of the nonlinear function $f$ – it has two or more zeroes each of which is interpreted as a distinct phase/state of the system. For most illustrations in this chapter and the next we will chose the following one-parameter family of nonlinear functions:

$$f(\phi) = \phi(1-\phi)(\phi+\alpha). \tag{11.1.2}$$

Though one could have chosen a more complicated form of $f$ with many more zeroes, we will see that (11.1.2) is adequate to capture most of the features of propagating front solutions in the NLDE of interest to us.

**Exercise 1** Can you think of a hypothetical scheme of chemical reactions which leads to the choice of $f$ above?

Since $\phi$ represents a physical field, e.g., the concentration of a chemical, we will always be concerned with bounded solutions of (11.1.1). The uniform states $\phi_0(x,t) = 0$, $\phi_1(x,t) = 1$ and $\phi_\alpha(x,t) = -\alpha$ are the three time-independent, homogeneous solutions of (11.1.1) and are interpreted as the three phases of the system.

**Exercise 2** *Linear stability analysis of the uniform states.* Along the same lines as in Section 6.2, do a linear stability analysis of (11.1.1) around the three uniform solutions $\phi(x,t) = \phi_0, \phi_1, \phi_\alpha$. Show that the stability of a uniform state $\phi_i$ against small, long wavelength fluctuations depends upon the sign of $f'(\phi_i)$. Chose, for example, $-1/2 < \alpha < 1/2$ and show that while the state $\phi_1(x,t)$ is linearly stable to small perturbations of any wavelength, the state $\phi_0$ is linearly stable (unstable) to sufficiently small wavelength perturbations if $\alpha$ is negative (positive).

**Exercise 3** *Global stability analysis of the uniform states.*

(a) Show that (11.1.1) can be written in the form

$$\frac{\partial \phi}{\partial t} = -\frac{\delta}{\delta \phi}\mathcal{F}_{ND}, \tag{11.1.3}$$

with $\frac{\delta}{\delta\phi}$ being a functional derivative and [1]

$$\mathcal{F}_{ND}[\phi(x,t)] = \int dx \left\{ \frac{D}{2}\left(\frac{\partial \phi}{\partial x}\right)^2 + F(\phi) \right\}, \tag{11.1.4}$$

where $F(u) = -\int_0^u f(u)du$.

(b) In the physics literature, function $\mathcal{F}$ is often referred to as the Ginzburg-Landau "free energy" of the system. Show that under the dynamics of (11.1.3)

$$\frac{d}{dt}\mathcal{F}_{ND} \leq 0.$$

Thus, the "free energy" monotonically decreases with time under dynamics (11.1.3). [$\mathcal{F}_{ND}$ is more generally referred to as the Lyapunov functional and we have already seen such a functional in connection with the Swift-Hohenberg equation in Section 10.2.]

(c) Draw $F(\phi)$ vs $\phi$ corresponding to $f(\phi)$ in (11.1.2) with, say, $\alpha \in (0, 1/2)$. Argue, in view of (b) above, that any initial state $\phi(x,0) \in (0,\infty)$ evolves to the state $\phi(x, t \to \infty) = 1$. Similarly, initial states with $-\infty < \phi_i < 0$ evolve to $\phi(x, t \to \infty) = -\alpha$. Note that this extends the linear stability analysis of exercise 2 to large perturbations.

Further, note that the state $\phi_0 = 0$ remains an unstable attractor of the dynamics (11.1.3), as was concluded in the linear stability analysis. Thus, initial conditions, however close to this state, will always move away from this state in time. The asymptotic state reached depends on the precise initial condition.

[1] Note that $\mathcal{F}_{ND}$ defined this way is not finite, even for the homogeneous bounded solutions of (11.1.1), if the range of integration is not finite. Hence (11.1.4) is to be considered as for a large but finite range.

Thus, states corresponding to the minima of the free energy are stable to perturbations which need not be small and are termed as *stable* states. In the physics literature a stable state normally refers to the state corresponding to the absolute minimum of the free-energy. States corresponding to all other local minima are referred to as *meta-stable* [2] (see exercise 4 below).

**Exercise 4** Consider an initial state $\phi(x,0)$ which links the two stable states $\phi_1 = 1$ and $\phi_\alpha = -\alpha$. Argue that the state corresponding to the absolute minima of the free energy (i.e. $\phi_1 = 1$ for $0 < \alpha < 1/2$) will grow at the expense of the other stable state (i.e. $\phi_\alpha$). Thus, $\phi_\alpha$ is a meta-stable state for $0 < \alpha < 1/2$.

By contrast, the state corresponding to a maxima of the free-energy is called *unstable* (e.g. the state $\phi_0 = 0$ for $0 < \alpha < 1/2$) and is unstable to small perturbations.

**Exercise 5** Based on the above definitions of stable, meta-stable and unstable states, determine which category each of the three states $\phi_1, \phi_0$ and $\phi_\alpha$ falls into for $-\infty < \alpha < \infty$. Sketch the three stationary states as functions of $\alpha$. (e.g. denote stable, meta-stable and unstable states with different colors.)

## 11.2 Traveling front solutions of the NLDE

Having discussed the various homogeneous, time-independent solutions of (11.1.1) with (11.1.2), let us now turn to the traveling-front type solutions of the nonlinear diffusion equation [3]. Mathematically, a traveling front is a solution of the form $\phi(x,t) = \Phi(x - vt)$ with $\Phi(-\infty) = \phi_i; \Phi(\infty) = \phi_j$ where $\phi_i$'s are the uniform stationary solutions of (11.1.1). In this chapter we will be concerned with the front solutions of (11.1.1) which connect the states $\phi_i = \phi_1$ at $\xi = -\infty$ to the state $\phi_j = \phi_0$ at $\xi = \infty$ for $-1/2 < \alpha < 0$. Thus, the front is between a stable ($\phi_1$) and a meta-stable ($\phi_0$) state. In the next chapter we will study the fronts between the same two states but for $\alpha > 0$, so that we have a stable state $\phi_1 = 1$ invading an unstable state $\phi_0 = 0$. We will discover that there are striking qualitative differences depending upon whether the front is between a stable and a meta-stable state or a stable and an unstable state.

Some of the general questions one is faced with are the following.

1. Do traveling-front type solutions exist for the nonlinear diffusion equation (11.1.1)?
2. Are these solutions stable against "small" perturbations? This has implications whether a given front solution will be approached dynamically starting from a generic initial condition.
3. Starting from an arbitrary initial state $\phi(x,0)$ how is the final front solution approached? This is related to the question of dynamical selection of one particular front solution when more than one solution exist.

[2] In the mathematics literature, the term meta-stable is used in a different way.

[3] The front solutions of the NLDE are all of "kink" type, i.e. they connect uniform states. Fronts can also occur between spatially modulated states, for example, see chapter 9 of [Wal].

By an analogy we will argue below that, at least for the choice of $f$ in (11.1.2), the answer to (1) is always affirmative. This analogy applies equally well to more complicated forms of the nonlinear function $f$.
The answer to (2) is slightly involved, even if we consider small perturbations within a linearized theory, due to the fact that the reference state (i.e. the front solution) is spatially nonuniform. We will follow a line of argument which borrows ideas from quantum mechanics to address the question of stability of the front solutions.
The answer to (3) is the most difficult and considerable amount of work has been done to show rigorously, starting with quite general initial conditions, how the front solution is approached as $t \to \infty$ [FM,AW,EvS].

In this chapter we will concentrate more on methods and arguments that are found in the physics literature.

## 11.3 Fronts between a stable and a meta-stable state

### The particle-on-a-hill analogy

**Exercise 6** Since we are dealing with traveling front solutions $\phi(x,t) = \Phi(x - vt)$ of (11.1.1), show that $\Phi$ must satisfy

$$\Phi'' = -v\Phi' + \frac{dF}{d\Phi}, \tag{11.3.1}$$

by a straightforward insertion of the ansatz into (11.1.1).

Equation (11.3.1) is conveniently interpreted as the equation of motion of a particle in the inverted potential $V(\Phi) = -F(\Phi)$, where $\Phi$ is the position of the particle, $\xi$ plays the role of time and $v$ the role of friction. A front solution corresponds to the trajectory which starts at the top of the hill at $\Phi = 1$ at $\xi \to -\infty$ and reaches the top of the hill at $\Phi = 0$ at $\xi \to \infty$. It is intuitively clear that such a trajectory exists only for a particular value of the "friction" $v$ (see exercise 7 below). If $v$ is too small, then enough energy is not dissipated during the motion from $\Phi = 1$ to $\Phi = 0$ and the particle continues past $\Phi = 0$. On the other hand, if $v$ is too large, the particle never reaches $\Phi = 0$ and ends in the valley between $\Phi = 1$ and $\Phi = 0$.

**Exercise 7** It is possible that, for a different choice of the nonlinearity $f$ and thus the potential $V(\Phi)$, there is more than one value of friction for which particle trajectories connecting $\Phi(-\infty) = 1$ and $\Phi(\infty) = 0$ exist. For an example, consider the nonlinearity $f(\Phi) = \mu\Phi + \Phi^3 - \Phi^5$ for $\mu < 0$.

(a) Sketch the corresponding potential $V(\Phi)$. Check that it has three "hills": one at $\Phi = 0$ and two (degenerate) ones at $\Phi_\pm = \pm\sqrt{1/2 + \sqrt{1/4 + \mu}}$ if $\mu > -1/4$ and that $V(\Phi_\pm) > V(0)$ for $\mu > -3/16$. Thus, for $-3/16 < \mu < 0$, in the free-energy terminology, we have two (absolutely) stable (degenerate) states at $\Phi_\pm$ and a metastable state at $\Phi = 0$. (Note that there are two unstable states at $\Phi = \pm\sqrt{1/2 - \sqrt{1/4 + \mu}}$ as well but we will only be concerned with solutions connecting stable states.)

(b) Let us look for particle trajectories between, say, the stable state $\Phi = \Phi_+$ and the meta-stable state $\Phi = 0$. As argued above, for a very large value of $v$, the particle ends up in the valley between the two hills and the friction has to be lowered to a critical value $v_c$ for the particle to be able to climb up to the hill $\Phi = 0$. Argue that, now, however, there could be still smaller (isolated) values of $v$ for which allowed particle trajectories exist.

[Hint: For $v = 0$ the particle climbs over the $\Phi = 0$ hill and reaches the hill-top $\Phi_-$ at $\xi \to \infty$. Hence for a very small $v > 0$ the particle will reach very close to the peak of the hill $\Phi_-$ and roll back. It will end up right at the top of the hill at $\Phi = 0$ if the value $v$ chosen correctly.]

(c) Thus, for a general choice of $f$ in (11.1.1), there could me more than one (but countable) number of front solutions for $v < v_c$. However, make a rough sketch of a particle trajectory for one such value $v < v_c$ and observe that while the trajectory corresponding to $v = v_c$ is monotonic, those corresponding to $v < v_c$ are not. We will see later that this non-monotonicity of the later front solutions makes them unstable to perturbations and are thus will not usually be approached starting from a generic initial condition.

Now, let us come back to our choice of the nonlinearity $f$ in (11.1.2).

**Exercise 8** Show that for $f$ in (11.1.2) with $\alpha = -1/2$, the potential $V(\Phi)$ is such that its values at the two hill tops are equal and that in this case the critical friction is $v_c = 0$.

(a) Write down the front solution in this case as an explicit quadrature and perform the integration to obtain the solution in closed form. [Hint: In terms of the analogy use "energy conservation".]

(b) Show that the solution (let us call it $\Phi_0(\xi)$) obtained in (a) above is also a solution of (11.3.1) for $v = \sqrt{2}(\alpha + 1/2)$. Check that $\Phi_0(\xi)$ is monotonic in $\xi$. Now, from the particle-on-a-hill picture we already know that there is only one monotonic front solution corresponding to a critical speed $v_c$. Thus, we have in fact found the exact front solution and $v_c = \sqrt{2}(\alpha + 1/2)$ for arbitrary value of $\alpha \in [-1/2, 0)$.

(c) Check that the results of (a) and (b) above can be extended to $\alpha \in (-1, -1/2]$ as well. The negative sign of $v_c$ in this case means that the relative stability of the uniform states $\Phi = 0$ and $\Phi = 1$ have been interchanged. Now, $\Phi = 0$ being the absolutely stable state invades the metastable state $\Phi = 1$.

## 11.4 Beyond the Particle-on-a-Hill analogy

In the particle-on-a-hill analogy we appealed to intuition to arrive at some results for the nonlinear diffusion equation (11.1.1). In the particular choice (11.1.2) of the nonlinearity $f$ it turned out that an explicit solution is possible. It is easily seen that this would not be the case for other choices of $f$. Furthermore, it is not difficult to realize that this mechanical analogy breaks down if one considers, for example, higher order Pde's. Thus, in the following we will present somewhat more general techniques which have the possibility of being extended to more complicated situations.

### Perturbative approach

As we saw in exercise 8, the front solution can be found by a quadrature for a potential $V_0(\Phi)$ for which $V(\Phi = 1) = V(\Phi = 0)$ (1 and 0 being the stable and the metastable states respectively). The front speed can be computed perturbatively if the potential difference between the stable and the meta-stable state $\epsilon \equiv V(1) - V(0)$ is small, i.e. for $|\epsilon| << 1$.

Let the front solution be $\Phi(\xi) = \Phi_0(\xi) + \Phi_1(\xi) + ...$ for a potential $V(\Phi) = V_0(\Phi) + \delta V(\Phi)$ which is slightly different from $V_0(\Phi)$. Here the perturbation series is in the small parameter $\epsilon$, i.e. $\Phi_n(\xi)$ is of the order of $\epsilon^n$. For the example of $f$ in (11.1.2), $\epsilon = (1/2 + \alpha)/6$.

**Exercise 9** Substitute the above perturbative expression for $\Phi(\xi)$ up to first order in (11.3.1) and keep only linear terms in $\Phi_1(\xi)$. Show that we then arrive at the following linear equation for $\Phi_1(\xi)$

$$\hat{H_0}\Phi_1(\xi) = h_0(\xi), \tag{11.4.1}$$

where

$$\hat{H_0} = d^2/d\xi^2 - V''[\Phi_0(\xi)]; \quad h_0(\xi) = \delta V'[\Phi_0(\xi)] - v\Phi_0'(\xi). \tag{11.4.2}$$

**Exercise 10** Show that $\Phi_0' = d\Phi_0(\xi)/d\xi$ is an eigenfunction of the above linear operator $\hat{H_0}$ with eigenvalue 0, i.e. $\int_{-\infty}^{\infty} \hat{H_0}\Phi_0'(\xi')d\xi' = 0$. [Hint: Use translational invariance of (11.3.1).]

**Exercise 11** Show that the operator $\hat{H_0}$ is Hermitian, i.e. for arbitrary square integrable (possibly complex) functions $u(\xi), v(\xi)$ the following holds: $\int_{-\infty}^{\infty} u^*(\xi)\hat{H_0}v(\xi)dx = \int_{-\infty}^{\infty} v^*(\xi)\hat{H_0}u(\xi)d\xi$.

(a) Now, using the result of the previous exercise show that the function $h_0(\xi)$ in (11.4.2) must satisfy

$$\int_{-\infty}^{\infty} dx\Phi_0'(\xi)h_0(\xi) = 0. \tag{11.4.3}$$

This condition arises in other contexts as well and is referred to as the *Fredholm alternative* (see, for example, section 9.2).

(b) Using $h_0(\xi)$ as given in (11.4.2), obtain the speed $v_c$ of the front up to first order in $\alpha + 1/2$:

$$v_c{}^{(1)} = (V[1] - V[0]) \,/ \int [\Phi_0'(\xi)]^2 d\xi. \tag{11.4.4}$$

Calculate the speed of the front for the function $f$ in (11.1.2) for $\alpha$ close to $-1/2$.

(c) Argue that (11.4.4) above, in the particle-on-a-hill analogy, is an approximate statement of the fact that the energy dissipated by friction is equal to the potential energy difference between the two hill-tops.

[Hint: (11.3.1) can be integrated by multiplying both sides by $\Phi'(\xi)$ to obtain (11.4.4) with $\Phi_0'(\xi)$ replaced with $\Phi'(\xi)$.]

### Properties of the front solution

**Exercise 12** *The front solution as strongly heteroclinic orbit.* (a) Write (11.3.1) as two first order equations in the "phase space" $(\Phi, \Phi' \equiv d\Phi/d\xi)$

$$\begin{aligned} \frac{d\Phi}{d\xi} &= \Phi', \\ \frac{d\Phi'}{d\xi} &= -v\Phi' - f(\Phi). \end{aligned} \tag{11.4.5}$$

The above may be looked upon as a (coupled) map and, in the language of maps in Chapter 1 and 2, the points $(0,0)$ and $(1,0)$ are the two fixed points of the map. Argue that a front solution corresponds to a trajectory in the phase space connecting these two fixed points. Such a trajectory is called a *heteroclinic* orbit.

(b) Write (11.4.5) as $\mathbf{v} = M\mathbf{v}$, where $\mathbf{v}$ is the column vector $(\Phi, \Phi')$ and $M$ is the coefficient matrix on the right hand side of (11.4.5). Find the flows close to the fixed point $(0,0)$ and argue that any trajectory ending at this fixed point has to approach it exactly along the eigendirection of the above operator $M$ corresponding to the eigenvalue $\lambda_+$.

Though, in this case (i.e. for $f'(0) < 0$) it turns out that $M$ has only one contracting eigendirection, we will see (section 12.3.2) that even when both the eigendirections of $M$ are contracting (e.g. for $f'(0) > 0$) there may exist a trajectory which approaches the $(0,0)$ fixed point along the most strongly contracting eigendirection. Such a trajectory is referred to as the *strongly heteroclinic* orbit. In the next chapter we will see that strongly heteroclinic orbits, when they exist, play an important role in front propagation into unstable states.

**Exercise 13** *Monotonicity of the front solution*

To show that a traveling front solution of (11.1.1), with appropriate boundary conditions, is monotonic first show that

(a) Any smooth initial state $\phi(x,0)$ confined to the strip $[1,0]$ remains in this strip under the evolution (11.1.1). [Hint: In order that a profile $\phi(x,t) \in [0,1]$ crosses the line $\phi = 1$ it has to "touch" this line first from below. This is not possible as the growth rate precisely at that point is non-positive.]

(b) Now, with the help of the phase-flow equations (11.4.5) show that along any allowed trajectory connecting the fixed points $(\Phi = 0, \Phi' = 0)$ and $(\Phi = 1, \Phi' = 0)$ $\Phi$, which can be reached from an initial condition confined to the strip $[0,1]$, is monotonically increasing (or decreasing if traversed in the reverse direction).

[Hint: Flows in the $\Phi' > 0$ plane is always to the right (i.e. in the increasing $\Phi$ direction) and is to the left for $\Phi' < 0$.]

(c) Now, show that, if the front solution is not confined to the strip $[1,0]$, then it is necessarily nonmonotonic. We will see in the next section that such a front solution is usually unstable to small perturbations and hence would not be approached starting from a generic initial condition.

## 11.5 Stability of the front solution

We have stressed earlier that in general explicit front solutions, even for simple nonlinearities, are nontrivial. As may be guessed already, the study the stability of such front solutions presents even a higher order of difficulty. However, we will illustrate a general technique, which can be applied to a wide variety of situations, to say something about the stability of the front solutions against "small" perturbations even without knowing the front solution explicitly. By "small" perturbations it is meant that we will focus on the stability issues within linearized equations. The question of stability of the front solutions is intimately connected to the question how different initial conditions approach the traveling front solution. An arbitrary initial condition $\phi(x,0)$ may be thought of as a superposition of the final front solution $\Phi(x)$ and a "perturbation" $\eta(x,0) = \phi(x,0) - \Phi(x)$. If this perturbation decays in time then the given initial condition will approach the front solution.

Let us consider a traveling-front solution $\Phi_v(x - vt)$ of the nonlinear diffusion equation (11.1.1) corresponding to speed $v$. We would like to check the stability of such a solution against perturbations $\eta(x,t)$ which obey the boundary conditions

$$|\eta(x \to \pm\infty, t)| = 0, \quad \forall t. \tag{11.5.1}$$

We will see that for a front between a meta-stable and a stable state, the unique monotonic solution $\Phi^\dagger(\xi)$ (the strongly heteroclinic orbit in exercise 12) is stable against perturbations satisfying (11.5.1 above. Thus, in this case, generic initial conditions satisfying the same boundary conditions (12.1.2) as the front solution will approach $\Phi^\dagger(\xi)$ (or its translate) in time.

**Exercise 14** (a) Show that the evolution of a small perturbation $\eta(\xi,t)$ in the frame of the moving front is governed by (keeping only terms linear in $\eta$)

$$\partial_t \eta(\xi,t) = \hat{L}_v \eta(\xi,t), \tag{11.5.2}$$

where

$$\hat{L}_v = \frac{d^2}{d\xi^2} + v\frac{d}{d\xi} + f'[\Phi_v(\xi)] \tag{11.5.3}$$

is called the (linear) stability operator and has already been encountered in (11.4.1): $\hat{H}_0 = \hat{L}_{v=0}$.

(b) Check that for $v \neq 0$, the linear operator $\hat{L}_v$ is not self-adjoint. Write down the eigenvalue equation for $\hat{L}_v$ from (11.5.2),

$$-\sigma\eta_\sigma(\xi) = \hat{L}_v \eta_\sigma(\xi) \tag{11.5.4}$$

Note that the eigenfunctions $\eta_\sigma$ of interest to us have to satisfy the same boundary conditions as $\eta(\xi,t)$ and thus, (11.5.4), along with boundary conditions $\eta_\sigma(\xi \to \pm\infty) = 0$, is in fact the *Sturm-Liouville* problem for the linear operator $\hat{L}_v$ (see, e.g., chapter 9 of [Arf]).

(c) Prove that the (right) eigenfunctions $\eta_\sigma$ form a complete orthogonal set and hence we can decompose any general perturbation $\eta(\xi,t)$ as $\eta(\xi,t) = \sum_\sigma e^{-\sigma t}\eta_\sigma(\xi)$.

Thus, we know the evolution of a generic perturbation if we know the full eigenvalue spectrum of $\hat{L}_v$. Now, this task is difficult since $\hat{L}_v$ is not a self-adjoint operator (why?).

**Exercise 15** (a) By making the transformation $\psi_\sigma(\xi) = e^{v\xi/2}\eta_\sigma$ show that (11.5.4) is transformed to

$$\sigma\psi_\sigma(\xi) = \hat{H}_v\psi_\sigma(\xi), \tag{11.5.5}$$

where

$$\hat{H}_v = -\frac{d^2}{d\xi^2} + U(\xi); \quad U(\xi) = v^2/4 - f'[\Phi_v(\xi)].$$

(b) Show that $\hat{H}_v$ is self-adjoint.

(c) Show that the Hilbert space of $\hat{H}_v$ does not contain all perturbations satisfying (11.5.1), namely those which do not satisfy in addition

$$[e^{v\xi/2}\eta]_{\xi\to\infty} = 0.$$

(d) Sketch the "potential" $U(\xi)$ for a "typical" monotonic front. For example, use the explicit front solution from exercise 8 in the last section. Note that $U(\xi)$ is a bounded function: $|U(\pm\infty)| < \infty$.

Though, by making the transformation in exercise 15(a) above we have lost track of certain set of relevant perturbations, the Hermiticity of $\hat{H}_v$ operator makes it convenient for us to use known results for the eigenvalue spectrum of Hermitian operators. In particular, we will use the following results for the spectrum of a Hermitian operator with a bounded "potential" $U(\xi)$ (see, e.g., [Mes] for a proof of these results):
(1): The spectrum is real and has two parts: a continuum one for $\sigma > min[U(-\infty), U(\infty)]$ and a discrete one for $\sigma < min[U(-\infty), U(\infty)]$.
(2): The eigenfunctions in the discrete part of the spectrum are in one-to-one correspondence with the number of nodes of the eigenfunction. The eigenfunction corresponding to the lowest eigenvalue is node-less, the next eigenfunction has one node and so on.

**Exercise 16** In the following we would argue that the front solution between a metastable and a stable state, is linearly stable against perturbations from the Hilbert space of $\hat{H}_v$.

(a) Show that the translation mode $\Phi_v'(\xi)$ is in the Hilbert space of $\hat{H}_v$, i.e. $[e^{v\xi/2}\Phi_v'(\xi)]_{\xi\to\infty} = 0$ and that $\Phi_v'(\xi)$ is an eigenfunction of $\hat{H}_v$ with eigenvalue 0. [Hint: Argue along similar lines as in exercise 10.]

(b) Argue, on the basis of the results (1) and (2) quoted above, that if $f'(0) < 0$ (which is the case when the invaded state is metastable) then there exist eigenfunctions of $\hat{H}_v$ with negative eigenvalue (and are thus destabilizing) if and only if the front is nonmonotonic. [Hint: The translation mode $\Phi_v'(\xi)$ has no nodes if and only if the front is monotonic.]

**Exercise 17** *Perturbations outside the Hilbert space of $\hat{H}_v$.* In the previous exercise we argued that monotonic fronts are stable and nonmonotonic fronts are unstable to perturbations which are *inside* the Hilbert space of $\hat{H}_v$. However, to ensure absolute stability of monotonic fronts we have to consider those perturbations which lie outside this Hilbert space. For this we go back to the original equation (11.5.4).

(a) Write the eigenvalue equation (11.5.4) as

$$\left[\frac{d^2}{d\xi^2} + v\frac{d}{d\xi} + f'[\Phi_v(\xi)] + \sigma\right]\eta_\sigma = 0. \tag{11.5.6}$$

Check that, for $\xi \to \infty$, the eigenfunction $\eta_\sigma$ has the asymptotic form

$$\eta_\sigma(\xi \to \infty) \sim a_v e^{-\Lambda_- \xi} + b_v e^{-\Lambda_+ \xi},$$

where

$$\Lambda_\pm(v) = \frac{v}{2} \pm \sqrt{\frac{v^2}{4} - [\sigma + f'(0)]}.$$

(b) Consider the two cases (i) $\sigma + f'(0) > 0$ and (ii) $\sigma + f'(0) < 0$ separately.

(i) $\sigma > -f'(0)$: Show that corresponding perturbations lie outside the Hilbert space and are not destabilizing.

(ii) $\sigma < -f'(0)$: Argue that the corresponding eigenfunctions lie inside the Hilbert space and use the analysis of the previous exercise to conclude that while monotonic fronts are stable, non-monotonic ones are unstable to perturbations.

**Exercise 18** *Approach to the front solution.* A generic initial condition may be written as $\phi(x,0) = \Phi(x) + \eta(x,0)$ where $\Phi(x - vt)$ is the asymptotic solution. The subsequent evolution is governed by (11.1.1) and at a later time we have $\phi(x,t) = \Phi(x - vt) + \eta(x,t)$. As argued in exercise 14, decompose the perturbation into eigenfunctions of the stability operator $\hat{L}_v$ and using the stability analysis in the previous two exercises show that a generic initial condition approaches the final front solution (or a translation of it) exponentially fast in time. What is the time scale of the slowest mode? Note that you have to disallow zero mode as a perturbation (why?).

### Stability of bound states of two fronts.

Let us look at the possibility of existence of a localized bound state of two fronts, which is essentially a bubble of one state in the other one. Using the particle-on-a-hill analogy described earlier, it is simple to see that one can include the stable state into the meta-stable one. Indeed, considering our case of $f$ in (11.1.2) for $\alpha \in (-1/2, 0)$, we release the particle at the top of the hill $\Phi = 0$ at $\xi = -\infty$. For $v = 0$, the motion is conservative. The particle will move down the $\Phi = 0$ hill and then climb up the $\Phi = 1$ hill but not up to the top. When the initial releasing height is reached the velocity changes sign and the particle returns to the initial position. For a small height difference between the two hills, the particle spends a long time near $\Phi = 1$ before returning to $\Phi = 0$ and the motion clearly describes the inclusion of $\Phi = 1$ state into the $\Phi = 0$ state.

**Exercise 19** We would like to see if such a state, which can be thought of as the bound state of two fronts, is linearly stable with respect to small perturbations.

(a) Using the approach of the previous section, show that there exist eigenmodes of the Hamiltonian (describing the evolution of small perturbations) with negative energy and thus the bound state of two fronts is inherently unstable to the corresponding eigenmodes. [Hint: Such a bound state is necessarily nonmonotonic.]

(b) The conclusion above can also be reached in the following way. Argue, on the basis of the principle of free-energy minimization of exercise (3), that while a small "bubble" of the stable state in the metastable one will shrink in time and vanish, a large enough stretch will actually always grow. Thus, there is a critical bubble size (i.e. a critical distance between the two fronts) which neither grows nor decays but it is unstable to perturbations.

## 11.6 Suggested further reading

A complete and rigorous treatment of existence, uniqueness and dynamical selection for a front between a metastable and a stable state is done in [FM]. In the article by Hakim in [GM], on which some of the sections in this chapter is based on, more illustrative examples are considered. The reader is encouraged to work through some of these examples.

# Chapter 12

# Front propagation into unstable states

In this section we consider front solutions of the nonlinear diffusion equation when one of the states is unstable. This happens, as discussed in exercise 3, when $\alpha > 0$ in (11.1.2). It is usually difficult to observe single propagating fronts since the unstable state is very sensitive to small perturbations. Nevertheless, there have been a number of experiments where this type of fronts are observed. In [Expts] we list a number of experimental studies where traveling fronts have been observed between a stable and an unstable state. One of the most striking observations of the theory of front propagation in these cases is that there exist a continuum of front speeds and shapes, in contrast to the unique front solution observed between a stable and a metastable state. Hence, one of the main questions in front propagation into unstable states is the selection criterion by which a particular front solution is approached starting from a given initial condition.

Historically, the invasion of an unstable state by a stable state seems to have been considered by Fisher [Fisher] in the thirties as a model for population dynamics. It was subsequently analyzed by Kolmogorov, Petrovsky and Piscunoff [KPP] for the nonlinear diffusion equation (11.1.1) for $f'(0) = 1$. The example treated by Fisher dealt with $f(\phi) = \phi(1-\phi)$. This can be considered as the $\alpha \to \infty$ of our choice of $f$ in (11.1.2) (with a subsequent rescaling of time and the diffusion constant: $t \to t/\alpha$; $D \to D/\alpha$ in (11.1.1).

## 12.1 Continuum of front speeds

As discussed in the beginning of the last chapter, for $\alpha > 0$ the homogeneous state $\phi_0 = 0$ becomes unstable. As in the last section we are looking for travelling front solutions $\Phi(x - vt)$ of (11.1.1) and hence the particle-on-a-hill picture applies here as well:

$$\frac{d^2\Phi}{d\xi^2} = -v\frac{d\Phi}{d\xi} - f(\Phi). \tag{12.1.1}$$

The boundary conditions being

$$\Phi(-\infty) = 1; \quad \Phi(\infty) = 0. \tag{12.1.2}$$

As before, this can be interpreted as the motion of the particle in the potential $V(\Phi) = \int^{\Phi} f(\Phi)d\Phi$, the role of time being played by $\xi$.

**Exercise 20** (a) Note that now we are looking for a particle trajectory which starts from $\Phi = 1$ at $\xi = -\infty$ and ends at the bottom of the "valley" $\Phi = 0$ at time $\xi = \infty$. Argue, in the particle-on-a-hill picture, that such a trajectory in fact exists for a *range* of values of friction $v$. Distinguish between the cases $\alpha < 1/2$ and $\alpha > 1/2$. [Hint: The lower limit of the range of possible $v$'s depends upon the relative heights of the two hills $\Phi = 1$ and $\Phi = -\alpha$.]

(b) Another way to see the existence of multiple solutions of (12.1.1) for $f'(0) > 0$ is the following. Refer to the phase-flow equations in exercise (12). Compute the eigenvalues and sketch the flows close to the fixed points (0,0) and (1,0). Compare with the case $f'(0) < 0$ of exercise (12) and argue for the existence of a continuum of front solutions.

**Exercise 21** *Linearization around the leading edge.*

(a) Linearize (12.1.1) close to the unstable state $\Phi = 0$ to look at the asymptotic behaviour of the front as $\xi \to \infty$.

$$\frac{d^2\Phi}{d\xi^2} = -v\frac{d\Phi}{d\xi} - \alpha\Phi. \tag{12.1.3}$$

(b) Write down the general solutions of (12.1.3) as

$$\begin{aligned} \Phi_v(\xi) &= A_v e^{-\lambda_- \xi} + B_v e^{-\lambda_+ \xi}, \quad v > v^* \\ &= K_v e^{-\lambda_v \xi} \cos q_v(\xi - \xi_0), \quad v < v^* \end{aligned} \tag{12.1.4}$$

where

$$\lambda_\pm(v) = v/2 \pm \sqrt{(v/2)^2 - \alpha}; \quad \lambda_v = v/2; \quad q_v = \sqrt{\alpha - (v/2)^2}.$$

What is $v^*$? What is the asymptotic behaviour of $\Phi_v$ for $v = v^*$?

(c) Interpret the qualitatively different solutions found above depending upon whether $v <, =$ or $> v^*$ in terms of the particle motion in the potential.

## 12.2 Stability of the front solutions.

**Exercise 22** Show, along the lines of section 11.5, that the oscillating front solutions for $v < v^*$ are linearly unstable to a continuum band of perturbations inside the Hilbert space. Check that the arguments in exercise 13 applies here as well and hence, oscillating front solutions anyway cannot be dynamically reached starting from an initial condition $\phi(x, 0) \in [0, 1]$.

**Exercise 23** *Perturbations outside Hilbert space.* Reformulate the arguments of exercise 17 to the present case with $f'(0) > 0$. Show that for monotonic fronts $\Phi_v$ with $v \geq v^*$, the only destabilizing ($\sigma < 0$) eigenfunctions of the stability operator $\hat{L}_v$ are the ones which are asymptotically "flatter" than the front itself and will eventually converge to a solution faster than $\Phi_v$, i.e. $v_{sel} > v$. In this sense the perturbations are

"too large". Nonmonotonic fronts, with $v \geq v^*$, remain unstable to eigenfunctions in the discrete part of the spectrum of $\hat{L}_v$. This discrete solutions are outside the Hilbert space of $\hat{H}_v$ and need special considerations (For a proof see [EvS]).

**Exercise 24** *Approach to the front solution.* Argue, along the lines of exercise 18, that generic initial conditions approaching the front solutions corresponding to $v > v^*$ do so exponentially fast in time. Check that this is not the case with approach to the front solution $v = v^*$. In fact, the velocity relaxes *algebraically* in this case and the convergence to the $\Phi_{v^*}$ is not uniform. For a detailed discussion of the approach to the front solution the reader is referred to [EvS].

## 12.3 Velocity selection.

We would now like to address the question of velocity selection starting from a general initial condition satisfying $[\phi(x,0)]_{x\to\infty} = 0$ and $\phi(x,0) > 0$ for some $x$. We will see that the selected front speed is determined by the "steepness" of the leading edge of the initial condition. The steepness $\lambda$ of an initial condition $\phi(x,0)$ is defined by

$$\lambda = -\lim_{x\to\infty} \frac{\partial\phi(x,0)}{\partial x}, \qquad \text{i.e.,} \quad \phi(x\to\infty,0) \sim e^{-\lambda x}. \tag{12.3.1}$$

**Exercise 25** (a) Argue that a nonzero steepness of the leading edge is conserved under the dynamics of (11.1.1), linearized around $\phi = 0$, i.e., if $\phi(x\to\infty,0) \sim e^{-\lambda x}$ then $\phi(x\to\infty,t) \sim e^{-\lambda x}$, $\forall t < \infty$.

(b) Show that the speed of the invading front is related to the steepness of the leading edge, i.e., for a front asymptotically behaving as $\Phi(\xi\to\infty) \sim e^{-\lambda\xi}$, the speed is given by $v_\lambda = \lambda + 1/\lambda$.

(c) Argue that the behaviour of $v \sim 1/\lambda$ for small $\lambda$ reflects the independent growth of $\Phi$ at different $\xi$ set up by the initial condition. It has nothing to do with real propagation. [Hint: By setting $D = 0$ in (11.1.1), $\phi$'s at different $x$ become decoupled.]

The previous exercise makes it clear that the velocity of the invading front depends on the initial condition $\phi(x, t = 0)$ and it can be as fast as one wants if the initial condition decreases sufficiently slowly as $\xi \to \infty$. Since in most physical systems the initial condition is localized [1](see exercise 27 below), the interesting question is therefore to determine the selected front velocity for an initial condition that decreases sufficiently rapidly as $\xi \to \infty$. In the case of Fisher's equation with $f'(0) = 1$ the problem has been rigorously solved [KPP,AW]. The result is that the selected front, starting from a sufficiently localized initial condition, is the one with the smallest speed and which remains positive everywhere $\Phi(\xi) > 0, \forall x$. In rigorous mathematical texts a "comparison theorem" is used to prove this [AW]. Below we sketch a non-rigorous version of the theorem.

[1]In most experiments, the initial condition is of the form $\phi(x,0) = 0(\neq 0),\ x > x_0(x < x_0)$, thus, $\lambda = 0$ in such cases.

**Exercise 26** *Comparison theorem.*

(a) Consider the nonlinear diffusion equation (11.1.1) with $f(\phi)$ replaced by $c(x,t)\phi(x,t)$

$$\partial_t \phi = \partial_{xx}\phi + c(x,t)\phi, \tag{12.3.2}$$

where $c(x,t)$ is a bounded function $\forall x,t$. Prove that if an initial condition $\phi(x,0)$ is everywhere non-negative then it remains so for all later times. i.e. if $\phi(x,0) \geq 0, \forall x$ then $\phi(x,t) \geq 0,\ \forall x; t > 0$. (Check that a similar result holds if $\phi(x,0) \leq 0$).

(b) Now, let us consider two (time-dependent) solutions $\phi_1(x,t)$ and $\phi_2(x,t)$ of (11.1.1) such that $\phi_1(x,0) \geq \phi_2(x,0)$. Using the positivity property in (a) prove that $\phi_1(x,t) \geq \phi_2(x,t),\quad \forall x; t > 0$. [Hint: Consider the time evolution of the difference $\phi_1(x,t) - \phi_2(x,t)$ and use the result of (a) above.]

(c) Now argue that, since a sufficiently localized initial condition can be bounded above, at $t = 0$, by the slowest front solution (which has an exponential leading edge) it cannot propagate faster than this solution. To rigorously prove that such an initial condition actually approaches the slowest front solution (or its translate) for $t \to \infty$, one needs to use a "maximum principle" which is beyond the scope of this text. Interested readers are referred to [PW].

**Exercise 27** Based on the results of exercise 25 and the fact that $v < v^*$ solutions are unstable, argue that for initial conditions which are steeper than $e^{-\lambda^* x}$ in the leading edge the selected front is $\Phi_{v^*}$ (provided it is monotonic and positive) corresponding to the speed $v_{sel} = v^*$. What is $\lambda^*$ for our case?

For the nonlinear diffusion equation, the comparison theorem provides a powerful method to determine the selected front solution for localized initial conditions. Unfortunately, this theorem cannot easily be extended to more general situations (convince yourself of this by adding, for example, the term $\partial^4\phi/\partial x^4$ to the right hand side of (11.1.1).) Below we will see that, in more general situations, the *linear spreading speed* determined by the linearized version of the original (nonlinear) equation is the selected front speed if the nonlinearities are not "large".

### 12.3.1 Linear velocity selection: pulled fronts.

**Exercise 28** Consider a front solution traveling with speed $v$. A small perturbation of the front can be written as a Fourier integral:

$$\eta(x,t) = \int dk a(k) e^{ikx - \tilde{s}_k t}, \tag{12.3.3}$$

where $a(k)$ is a smoothly varying function for the perturbation $\eta$ to be localized in space. [The dispersion relation $\tilde{s}_k$ is to be obtained from the linearized evolution equation for the perturbation.]

Check that, in the frame of the moving front, the perturbation transforms to

$$\eta(\xi,t) = \int dk a(k) e^{ik(\xi+vt) - \tilde{s}_k t}. \tag{12.3.4}$$

For large $t$, do a (formal) saddle-point/stationary-phase/steepest-descent evaluation of the above integral to approximate $\eta$ as

$$\eta(\xi, t \to \infty) \simeq \Gamma_{k^*} a(k^*) e^{ik^*\xi - s_{k^*}t}, \tag{12.3.5}$$

where $s_k = \tilde{s}_k - ikv$ and $k^*$ is given by the condition $(ds_k/dk)_{k=k^*} = iv$ and $\Gamma_{k^*}$ is a finite number.

[Hint: Extend the integral into complex $k$-plane. Write $k = k_1 + ik_2$ and $s_k = \sigma_k + i\omega_k$ with $k_1, k_2, \sigma_k, \omega_k$ real. The saddle point evaluation of the integral (12.3.5) above is based on the observation that, for large $t$, the principal contribution comes from the point in $k$ plane where $\sigma_k$ is the smallest and the integration contour is chosen such that $\omega_k$ has the slowest variation along it. Refer to any standard text on saddle-point analysis, e.g., chapter 7 in [Arf].]

Now, looking at (12.3.5) we make the following observations:

(a) If $\sigma_{k^*} < 0$ the perturbation grows at a fixed position in the moving frame. This is known as *absolute instability*.

(b) If $\sigma_{k^*} > 0$ the perturbation decays at a fixed position in the moving frame. But we know that any perturbation will grow (in the fixed frame) since the reference state is the unstable state. Thus, what happens is that the perturbation actually grows in this case but is convected away (to $\xi \to -\infty$) sufficiently rapidly in the moving frame so that at any fixed position of the moving frame we see a decay. This is known as *convective instability*.

The linear spreading speed is the speed for which $\sigma_{k^*} = 0$ [2] [WvS]. In a frame moving at this speed, the leading edge of the front, which may be viewed as a perturbation of the unstable state, remains stationary.

**Exercise 29** Consider (12.1.2) with $f$ in (11.1.2) and $\alpha = 1$ and show that the linear spreading speed is given by $v^* = 2\sqrt{\alpha}$.

**Exercise 30** Can you find out the linear spreading speed for the Swift-Hohenberg equation (9.1.2) discussed in Chapter 9?

The linear spreading speed is the speed determined by the leading edge of the front. If the leading edge dominates the dynamics of the front we have a *pulled* front. Depending on the nonlinearities it may happen the nonlinear part of the front has a speed larger than the linear spreading speed and thus the selected front is the strongly heteroclinic orbit $\Phi_{v^\dagger}$ with speed $v^\dagger > v^*$. The front in this case is referred to as a *pushed* front since it is the nonlinear part of the front which dominates the dynamics and "pushes" the leading edge.

### 12.3.2 Nonlinear velocity selection: pushed fronts.

As discussed earlier, if the nonlinearities are strong enough that there exists a strongly heteroclinic orbit then the later becomes the selected front solution for sufficiently

[2] This is also known as the *linear marginal stability criterion*, see [B-J].

localized initial conditions. We will see that this happens when $\alpha \in (0, 1/2)$. We will see that in this case the front solutions corresponding to $v < v^\dagger$ don't remain non-negative and hence are unstable.

**Exercise 31** *Selected front for $\alpha < 1/2$.*

(a) Plot $\lambda$ vs. $v$ from exercise (21).

(b) Show that $\Phi_{v\dagger}$ ($\equiv \Phi_0$ of exercise 8) is a solution for $v = v^\dagger = \sqrt{2}(1/2 + \alpha)$ even for $\alpha > 0$. Draw the $\lambda = \lambda^\dagger = 1/\sqrt{2}$ in the graph of (a) above and observe that for $\alpha > 1/2$ ($\alpha < 1/2$) it intersects the left branch $\lambda_-$ (right branch $\lambda_+$).

(c) Comparing with the general solution (12.1.4) check that $A_{v^\dagger} = 0$ ($B_v^\dagger = 0$) for $\alpha < 1/2$ ($\alpha > 1/2$). Now, argue that for very large $v$ the coefficient $A_v > 0$ and that, by continuity of the coefficients in (12.1.4) as function of $v$, one must have $A_v < 0$ ($B_v < 0$) for $v < v^\dagger$ if $\alpha < 1/2$ ($\alpha > 1/2$). Note that we have assumed in the above that there is no more than one strongly heteroclinic solution and thus the change of sign of the coefficient $A_v$ can occur only once.

(d) Argue, using the result from (c) above, that for $\alpha < 1/2$, $\Phi_{v^\dagger}$ is the selected front solution. For $\alpha > 1/2$, there are no strongly heteroclinic orbits and thus $\Phi_{v^*}$ is the selected front solution with speed $v_{sel} = v^*$.

(e) Plot the selected front speed $v_{sel}$, for localized initial conditions, as a function of $\alpha$.

In the example of the nonlinearity $f$ in (11.1.2) we have been dealing with, we happen to know the strongly heteroclinic front solution explicitly for any value of $\alpha$. However, this may not be the generic situation. In the following we sketch a technique which works for a certain class of $f$.

**Exercise 32** *Finding the front solution when $v_{sel} > v^*$.* Our starting point is the requirement that for a strongly heteroclinic orbit $A_{v^\dagger} = 0$.

(a) Argue that, since the corresponding front solution decays as a pure exponential $\Phi_{v^\dagger}(\xi \to \infty) \simeq e^{-\lambda_+ \xi}$ it must be the solution of a first order differential equation

$$\frac{d\Phi_{v^\dagger}}{d\xi} = g(\Phi_{v^\dagger}),$$

where the function $g$ is yet to be determined.

(b) The function $g$ can be found as follows. Substitute the above equation into (12.1.3) to obtain

$$[g' + v]g = -f[\Phi].$$

Try a polynomial solution for $g(\Phi)$. Since $f$ in (11.1.2) is a cubic polynomial in $\Phi$ argue that $g$ has to be of second order and obtain $g(\Phi)$ by matching powers on both sides of the above equation.

Integrate the above equation to obtain the front solution and check that it is identical to the solution obtained in exercise 8.

Fronts propagating into unstable states are encountered in many situations. Most of the interesting experiments [Expts] are done on systems where the invading state

and the invaded state, either or both, are patterned. In such situations the front is referred to as an *envelop* front or a *domain wall* respectively. The Swift-Hohenberg equation (Chapters 9, 10) and the associated amplitude equation are used to study the front propagation in these cases [B-J].

## 12.4 Suggested further reading

We have prepared the text of the last two chapters along the lines of the article by V. Hakim in [GM] and the article by U. Ebert and W. van Saarloos [EvS]. For more rigorous derivation of some of the results the original articles [KPP,AW,FM] should be referred to. In [EvS] many of the techniques used here are described in detail.

In [EvS] a comprehensive study of the question of relaxation to the asymptotic front solution is done. The reader is encouraged to have a look at this article. Algebraic convergence for pulled fronts in higher order differential equations, in difference equations and in couple equations are also discussed in detail here.

Also, in the present study we have been confined to the one dimensional aspect of propagating fronts. For a brief account of issues involved in higher dimensions have a look at the book [Wal].

*[Wal]* D. Walgraef, *Spatio-Temporal Pattern Formation*, Springer-Verlag, New York, 1997.

*[FM]* P. C. Fife and J. B. McLeod, *Arch. Rational Mech. and Anal.* **65**, 335 (1997).

*[AW]* D. G. Aronson and H. F. Weinberger, *Adv. Math*, **30**, 33 (1978).

*[EvS]* U. Ebert and W. van Saarloos, *Front Propagation into Unstable States*, preprint. Leiden, 1999.

*[Arf]* G. Arfken, *Mathematical Methods for Physicists* (3rd. edition), Academic Press, San Diego, 1985.

*[Mes]* A. Messiah, *Quantum Mechanics*, North-Holland, Amsterdam, 1974.

*[GM]* V. Hakim in *Hydrodynamics and Nonlinear Instabilities*, C. Godreche and P. Manneville, eds. (Cambridge University Press, Cambridge, 1998).

*[Expts]* G. Ahlers and D. S. Cannell, Phys. Rev. Lett. **50**, 1583(1983), J. Fineberg and V. Steinberg, Phys. Rev. Lett. **58**, 1332 (1987), L. Limat *et. al.*, Physica D **61**, 166 (1992).

*[Fisher]* R. A. Fisher, *Ann. Eugenics*, **7**, 355 (1937).

*[KPP]* A. Kolmogorov, I. Perovsky, and N. Piskunoff, Bulletin de l'université d'état à Moscou, Ser. int., Section A, Vol. 1 (1937).

*[PW]* M. H. Protter and H. Weinberger, *Maximum Principles in Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1967.

*[WvS]* W. van Saarloos, Phys. Rev. A **37**, 211 (1988).

*[B-J]* E. Ben-Jacob et. al., *Physics D*, **14**, 348 (1985).