Lecture Notes: Dynamical Systems and PDEs

Christian Kuehn
Institute for Analysis and Scientific Computing,
Vienna University of Technology, 1040 Vienna, Austria

April 29, 2015

Preface

This document is eventually (i.e. hopefully) going to develop into an introductory book on how to use various dynamical systems ideas to study properties of solutions of partial differential equations (PDEs). Part of these lecture notes have been used in teaching at TU Vienna in 2014/2015.

As background knowledge a first course in PDEs is essential, e.g. based upon parts of [Eva02]. Some basic knowledge of functional analysis [Rud91], ODEs and/or dynamical systems [HSD03] would be desirable but not strictly necessary as one may look up the required results that we use as tools relatively easily along the way. Furthermore, the cited literature references are definitely not exhaustive and just provide some pointers to the literature; these notes provide basic ideas and the reader is strongly encouraged to explore particular topics in more detail. Currently, the notes are only available without figures, which have been drawn in the lectures.

Note carefully: The lecture notes are a work in progress and should be used with caution! Please report any errors or inaccuracies you find to

ck274@cornell.edu

So far, I have the pleasure to thank Annalisa Iuorio, Tobias Jawecki, Stefan Portisch, Pedro Aceves Sanchez, Elisabeth Schiessler, and Lara Trussardi for alerting me of several mistakes and misprints in previous versions of these notes.

Contents

1 A Whirlwind Introduction

The goal is to study prototypical models of PDEs for some unknown function u with

$$x \in \Omega \subseteq \mathbb{R}^N, \qquad t \in [0,T), \qquad u = u(x,t), \qquad u : \Omega \times [0,T) \to \mathbb{R}$$

where Ω is a bounded domain with sufficiently smooth boundary and $T > 0$ is a final time up to which we study the dynamics; frequently, the one-dimensional spatial case $x \in \mathbb{R}$ and long-time dynamics with $T = +\infty$ will be considered. The following notations for derivatives will be used

$$\Delta u := \sum_{n=1}^{N} \frac{\partial^2 u}{\partial x_n^2}, \quad u_t := \frac{\partial u}{\partial t}, \quad \partial_{x_n} u := \frac{\partial u}{\partial x_n}, \quad \nabla u := (u_{x_1}, \ldots, u_{x_N})^\top,$$

where $(\cdot)^\top$ denotes transpose (so the **gradient** ∇u is a column vector) and Δu is the **Laplacian** of u. We shall use $|\cdot|$ to denote absolute value and $\|\cdot\|$ for norms; where the Euclidean norm is always understood as the default one on any finite-dimensional normed space.

Example 1.1. The **Nagumo** or **Real-Ginzburg-Landau (RGL) equation**

$$u_t = \Delta u + u(1-u)(u-p), \tag{1.1}$$

where $p \in \mathbb{R}$ is a parameter, is a classical model in nonlinear PDE theory. The Nagumo equation initially arose as a simplification of the Hodgkin-Huxley model in neuroscience for electric impulse propagation in axons. The RGL name is part of the important class of amplitude equations to be discussed in Lecture 9. Sometimes (1.1) is also referred to as the **Allen-Cahn equation**. We augment (1.1) by the **initial condition** $u(x,0) = u_0(x)$ for a given sufficiently smooth function $u_0 : \Omega \to \mathbb{R}$. Furthermore, if $\Omega \subsetneq \mathbb{R}^N$ has a boundary then we consider suitable boundary conditions such as **Dirichlet boundary conditions**

$$u(x,t) = g(x), \quad \text{for } x \in \partial\Omega,$$

where $\partial\Omega$ denotes the boundary of Ω, or **Neumann boundary conditions**

$$(\vec{n} \cdot \nabla u)(x,t) = g(x), \quad \text{for } x \in \partial\Omega,$$

where $\vec{n}$ is the outer unit normal vector to $\partial\Omega$, and $g : \mathbb{R}^N \to \mathbb{R}$ is assumed to be sufficiently smooth. Usually, the precise boundary conditions for the Nagumo equation, as well as for all the other equations to be discussed, will not be our main focus in these notes and we shall use **homogeneous** conditions $g(x) \equiv 0$ or even **periodic boundary conditions**

$$x \in \mathbb{T}^N := \mathbb{R}^N/\mathbb{Z}^N \text{ (= the } N\text{-dimensional torus)}$$

to simplify the problem to study its main dynamical features. ♦

In particular, we shall always implicitly assume from now on that "sufficiently nice" initial and boundary conditions are chosen for the problem at hand. Once the boundary and/or initial condition are needed for a particular calculation, we specify them explicitly.

Example 1.2. The **stationary** version of the Nagumo/RGL equation (1.1) is obtained by setting $\partial_t u = 0$ and given by

$$0 = \Delta u + u(1-u)(u-p), \qquad u = u(x). \tag{1.2}$$

Nonlinear elliptic PDE, such as (1.2), arise frequently in applications, e.g. nonlinear elasticity, mathematical biology or theoretical physics. ♦

Although the cubic nonlinearity of the Nagumo equation and the Laplace operator Δ are natural choices to model **reaction** and **diffusion** respectively, there are many other choices that are natural to study.

Example 1.3. One option is to consider a quadratic nonlinearity instead, which yields the so-called Fisher- Kolmogorov-Petrovskii-Piskounov **(FKPP)** equation

$$u_t = \Delta u + u(1-u). \tag{1.3}$$

Instead of changing the reaction-term, one frequently encounters other linear operators, not just Δ. A typical example is the **Swift-Hohenberg equation**

$$\partial_t u = -(1+\Delta)^2 u + f(u), \tag{1.4}$$

where $f(u)$ is usually a cubic nonlinearity in u, and depends usually on a parameter $p \in \mathbb{R}$ as well. In applications, it is frequently natural to consider additional components. A classical example is the **FitzHugh-Nagumo equation**

$$\begin{aligned} \partial_t u &= \Delta u + u(1-u)(u-p_1) - v + p_2, \\ \partial_t v &= p_3(u - p_4 v), \end{aligned} \tag{1.5}$$

where $p_j \in \mathbb{R}$, $j \in \{1,2,3,4\}$, are parameters and $v = v(x,t)$. ♦

The field of nonlinear spatio-temporal evolution equations and their dynamical analysis is vast (to say the least). Here we shall focus on some examples but one should always keep in mind that even the list of important examples is extremely long. Here we shall list a few examples, where dynamical systems techniques have turned out tremendously helpful (just look at the differences as well as similarities in structure of the equations for now):

$$
\begin{aligned}
\textbf{Kuramoto-Sivashinsky} \quad u_t &= -u_{xxxx} - u_{xx} - uu_x, && (1.6)\\
\textbf{Burgers'} \quad u_t &= -uu_x, && (1.7)\\
\textbf{Nonlinear Wave} \quad u_{tt} &= \Delta u + f(u), && (1.8)\\
\textbf{Porous Medium} \quad u_t &= \Delta(u^p), \quad p \in (0,+\infty) && (1.9)\\
\textbf{Korteweg-de Vries} \quad u_t &= -uu_x + -u_{xxx}, && (1.10)\\
\textbf{Neural Field} \quad u_t &= -u + \int_\Omega \omega(x,y) f(u(y))\ \mathrm{d}y, && (1.11)\\
\textbf{Cahn-Hilliard} \quad u_t &= \Delta(f(u) - p\Delta u), && (1.12)\\
\textbf{Allen-Cahn equation} \quad u_t &= \Delta u - f(u), && (1.13)\\
\textbf{Nonlinear Schrödinger} \quad \mathrm{i}\psi_t &= -\Delta\psi + p\psi|\psi|^2, \quad \psi \in \mathbb{C}, && (1.14)\\
\textbf{Gross-Pitaevskii} \quad \mathrm{i}\psi_t &= -\Delta\psi + V(x)\psi + p\psi|\psi|^2, && (1.15)\\
\textbf{Keller-Segel} \quad \begin{pmatrix} u_t \\ v_t \end{pmatrix} &= \begin{pmatrix} \nabla\cdot(\nabla u - u\nabla v) + f(u,v) \\ \Delta v + g(u,v) \end{pmatrix}, && (1.16)
\end{aligned}
$$

where the list could be continued with Sine-Gordon, Boltzmann, Gray-Scott, Gierer-Meinhardt, Landau-Lifshitz-Gilbert, Euler, Navier-Stokes, and many more! The main point is: A dynamical systems viewpoint can be useful to understand all of these equations better, and we shall try to motivate here with a few of these examples, why this is the case. But first, we have to recall some important results for dynamics of ODEs

$$u' := u_t = f(u), \qquad u(0) = u_0 \in \mathbb{R}^d. \tag{1.17}$$

Theorem 1.4. ***(Local existence and uniqueness**; see [HSD03]) Consider* (1.17) *with $f \in C^1 = C^1(\mathbb{R}^d, \mathbb{R}^d)$. Then there exists $t_0 > 0$ and $u : (-t_0, t_0) \to \mathbb{R}^d$ such that*

$$u'(t) = f(u(t)),\ \text{for } t \in (-t_0, t_0) \quad \text{and} \qquad u(0) = u_0, \tag{1.18}$$

i.e., u solves (1.17) *for some open time interval containing $t = 0$.*

Theorem 1.4 is completely local in time, and says *nothing* about what the solutions to (1.17) actually do. Of course, it is critical to have such knowledge for any application, for which the model was written down in the first place. The next classical theorem does only slightly better.

Theorem 1.5. ***(Continuous dependence**; see [HSD03]) Consider* (1.17) *with $f \in C^1$ and f with Lipschitz constant κ on an open bounded set $\mathcal{U} \subset$*

$\mathbb{R}^d$. *Suppose* u, v *both solve* (1.17) *and remain in* $\mathcal{U}$ *for all* $t \in [0, T]$ *then*

$$\|u(t) - v(t)\| \leq \|u(0) - v(0)\| \mathrm{e}^{\kappa T}, \tag{1.19}$$

i.e., solutions may diverge at most exponentially from each other; see also Figure TODO.

Theorems 1.4-1.5 do not provide enough information on solutions. Nonlinear ODEs are already very difficult and the dynamics can be tremendously complicated, even one-dimensional examples are interesting.

Example 1.6. Consider the one-dimensional ODE

$$u' = f(u), \qquad u : \mathbb{R} \to \mathbb{R}. \tag{1.20}$$

If the equation is linear $f(u) = pu$ for a parameter $p \in \mathbb{R}$, then $u(t) = u(0)\mathrm{e}^{pt}$ and we can easily draw the phase portraits as in Figure TODO. For general nonlinear equations, closed-form solutions may not exist. Even if they do, it is frequently more insightful to argue abstractly and/or geometrically. For example, consider the ODE

$$u' = u(p - u) = f(u). \tag{1.21}$$

The steady states (or equilibrium points) u^* are obtained by setting $u' = 0$, so that $u_1^* = 0$ and $u_2^* = p$ as shown in Figure TODO. It is an extremely important idea to first study the dynamics *locally* near steady states. Consider

$$u = u_2^* + \varepsilon w = p + \varepsilon w$$

for some small $\varepsilon > 0$. Then we have

$$u' = (p + \varepsilon w)' = \varepsilon w'$$

as well as

$$u(p - u) = (p + \varepsilon w)(p - p - \varepsilon w) = -\varepsilon p w + \mathcal{O}(\varepsilon^2) \approx -\varepsilon p w.$$

So we can hope to study the linear system $w' = -pw$ locally near $w = 0$ to obtain stability results for $u_2^* = p$. Note that we can also derive this system via direct linearization as

$$w' = (\mathrm{D}_u f)(u_2^*) w = f'(p) w = (p - 2p) w = -pw. \tag{1.22}$$

From (1.22), the local flow near p is given as in Figure TODO. The stability of $u_2^* = p$ changes as p passes through zero, i.e., u_2^* is **locally asymptotically unstable** for $p < 0$ and **locally asymptotically stable** for $p > 0$. Similarly, we can obtain results for $u_1^* = 0$. Then another very helpful view is to consider the (u, p)-plane and draw a **bifurcation diagram** as shown in Figure TODO. ♦

To define what we really mean by a bifurcation, we need a definition.

Definition 1.7. A dynamical system (A) is **topologically equivalent** (or **topologically conjugate**) to another one (B) if there is a homeomorphism mapping trajectories of (A) to (B) preserving the direction of time.

For example, an ODE (1.17) generates a **flow**

$$\varphi : \mathbb{R}^d \times (-t_0, t_0) \to \mathbb{R}^d, \qquad \varphi(u_0, t) = u(t) \tag{1.23}$$

and we can ask, when two flows are topologically conjugate. Suppose $h : \mathbb{R}^d \to \mathbb{R}^d$ is a diffeomorphism, $v := h(u)$ then

$$v' = (\mathrm{D}h)u' = (\mathrm{D}h)f(u) = (\mathrm{D}h)f(h^{-1}(v)) \tag{1.24}$$

and it is easy to show that $v' = g(v) := (\mathrm{D}h)f(h^{-1}(v))$ and $u' = f(u)$ are topologically conjugate. Unfortunately, the converse is not true, and existence of a topological conjugacy does not imply the existence of h; furthermore, there are subtle differences between continuous-time and discrete-time dynamical systems.

Example 1.8. Consider as another example the **fold** (or **saddle-node**) **bifurcation** , which is exemplified by the one-dimensional system

$$u' = p + u^2 = f(u). \tag{1.25}$$

Steady states are $u_1^* = \sqrt{-p}, u_2^* = -\sqrt{-p}$, which exist for $p < 0$, collide at $p = 0$ and disappear for $p > 0$. Local stability is easily checked from the linearization

$$w' = (\mathrm{D}_u f)(u^*)w = 2u^* w = \pm 2\sqrt{-p}w \tag{1.26}$$

so u_1^* is unstable and u_2^* is stable as shown in the bifurcation diagram in Figure TODO. Clearly, the phase portraits for $p < 0$, $p = 0$ and $p > 0$ are not topologically equivalent. ♦

It should be noted that (1.25) is a **normal form** for a fold bifurcation, i.e., fold bifurcations in other coordinates can be qualitatively 'reduced'/'transformed' to (1.25).

Definition 1.9. The appearance of a topologically nonequivalent phase portrait under parameter variation is called a **bifurcation**.

Studying bifurcations for PDEs will be one of the main themes in the lectures to follow. Another main theme is to take a geometric viewpoint of phase space.

Example 1.10. Consider the two-dimensional linear ODE

$$\begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}}_{=:A} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{1.27}$$

which has solutions $u_1(t) = u_1(0)\mathrm{e}^{-t}$, $u_2(t) = u_2(0)\mathrm{e}^{t}$ and a **saddle** steady state at $u^* = (0,0)$; see Figure TODO. The eigenspaces of A

$$E^{\mathrm{s}}(u^*) := \{u = (u_1, u_2)^\top \in \mathbb{R}^2 : u_2 = 0\}, \tag{1.28}$$

$$E^{\mathrm{u}}(u^*) := \{u = (u_1, u_2)^\top \in \mathbb{R}^2 : u_1 = 0\}, \tag{1.29}$$

are **invariant** under the flow, i.e., trajectories cannot enter or leave. Furthermore, the dynamics is directed away from u^* in $E^{\mathrm{s}}(u^*)$ and towards u^* in $E^{\mathrm{s}}(u^*)$; see Figure TODO. ♦

More generally, one should not only look at linear spaces but smooth manifolds.

Definition 1.11. Let $\varphi(u_0, t)$ be a flow associated to an ODE (1.17) with a steady state u^*. Define the **stable** and **unstable manifolds** by

$$\begin{aligned} W^{\mathrm{s}}(u^*) &:= \{v \in \mathbb{R}^d : \varphi(v,t) \to u^*, \quad \text{as } t \to +\infty\}, \\ W^{\mathrm{u}}(u^*) &:= \{v \in \mathbb{R}^d : \varphi(v,t) \to u^*, \quad \text{as } t \to -\infty\}. \end{aligned}$$

Definition 1.12. A steady state u^* of the ODE (1.17) is called **hyperbolic** if $(\mathrm{D}_u f)(u^*) \in \mathbb{R}^{d\times d}$ has eigenvalues λ_i with $\mathrm{Re}(\lambda_i) \neq 0$ for all $i \in \{1, 2, \ldots, d\}$.

The next classical theorem shows that hyperbolicity of steady states implies that the local geometry perturbs very nicely, when the linearized and the fully nonlinear system are compared.

Theorem 1.13. *(**Stable-Unstable Manifold Theorem**; see [KH95]) Suppose the ODE* (1.17) *has a hyperbolic steady state u^* and* $(\mathrm{D}_u f)(u^*)$ *has k real-part negative and $d-k$ real-part positive eigenvalues with corresponding eigenspaces $E^{\mathrm{s}}(u^*)$ and $E^{\mathrm{u}}(u^*)$ for the linearized system. Then there exists a neighbourhood $\mathcal{U}$ of u^* with local stable and unstable manifolds $W^{\mathrm{s}}_{\mathrm{loc}}(u^*)$ and $W^{\mathrm{u}}_{\mathrm{loc}}(u^*)$*

$$\begin{aligned} W^{\mathrm{s}}_{\mathrm{loc}}(u^*) &= \{v \in \mathcal{U} : \varphi(v,t) \to u^* \ \text{as } t \to \infty \text{ and } \varphi(v,t) \in \mathcal{U}\ \forall t \geq 0\}, \\ W^{\mathrm{u}}_{\mathrm{loc}}(u^*) &= \{v \in \mathcal{U} : \varphi(v,t) \to u^* \ \text{as } t \to -\infty \text{ and } \varphi(v,t) \in \mathcal{U}\ \forall t \leq 0\}. \end{aligned}$$

Furthermore, $W^{\mathrm{s}}_{\mathrm{loc}}(u^*)$ and $W^{\mathrm{u}}_{\mathrm{loc}}(u^*)$ are tangent to $E^{\mathrm{s}}(u^*)$ and $E^{\mathrm{u}}(u^*)$ at u^* and are as smooth as f.

Note that Theorem 1.13 does relate algebra, geometry, analysis and dynamics near hyperbolic steady states; see also Figure TODO.

Exercise 1.14. Use separation of variables to show that the ODE $u' = u^2$ for $u \in \mathbb{R}$ has solutions becoming unbounded in finite time, i.e. $t_0 \neq \infty$ in Theorem 1.4. $\Diamond$

Exercise 1.15. Consider $u' = Au$ for some matrix $A \in \mathbb{R}^{2\times 2}$ and classify the stability of $u^* = 0$ based upon the trace and determinant of A. $\Diamond$

Exercise 1.16. Derive the bifurcation diagram for the **pitchfork bifurcation** normal form $u' = u(p - u^2)$ with $u \in \mathbb{R}$ and $p \in \mathbb{R}$. $\Diamond$

Background and Further Reading: There are many books excellent books on dynamical systems and ODEs. Very readable introductions suited for self-study are [Str00, HSD03]. Bifurcation theory for ODEs is well-documented in the monograph [Kuz04]. Deriving all the results and techniques presented in this section is definitely a course by itself but we shall simply assume the results here as motivating starting points.

2 Implicit Functions and Lyapunov-Schmidt

A first step to try to use dynamical systems ideas to PDEs is to focus on local bifurcations and this requires some technical background from functional analysis.

Let X, Y, Z be real Banach spaces. We want to study the problem

$$F(u, v) = 0, \qquad F : X \times Y \to Z, \quad (u, v) \in X \times Y. \tag{2.1}$$

It helps to think of the simple case $Y = \mathbb{R}$ and v as a parameter, then one may write, e.g., the stationary Nagumo equation (1.2) in the form (2.1)

$$F(u, p) = \Delta u + u(1 - u)(u - p) = 0 \tag{2.2}$$

with $p = v \in \mathbb{R}$ and a suitable Banach space X.

Definition 2.1. $F : X \times Y \to Z$ is **Fréchet differentiable** in X at (u_0, v_0), if there exists a bounded linear operator $\mathrm{D}_u F(u_0, v_0) \in \mathcal{L}(X, Z)$ such that

$$\lim_{h \to 0} \frac{\|F(u_0 + h, v_0) - F(u_0, v_0) - \mathrm{D}_u F(u_0, v_0)h\|_Z}{\|h\|_X} = 0.$$

Sometimes we shall write $(\mathrm{D}_u F)(u_0, v_0)$ to indicate with correct brackets, where the derivative is evaluated. Mostly, the shorter notation $\mathrm{D}_u F(u_0, v_0)$, or even just $\mathrm{D}_u F$, will be used.

Theorem 2.2. ***(Implicit Function Theorem**, [Dei10, Kie04]) Suppose (u_0, v_0) satisfies (2.1), $(\mathrm{D}_u F)(u_0, v_0)$ is bijective, $F \in C(X \times Y, Z)$ and $\mathrm{D}_u F \in C(X \times Y, \mathcal{L}(X, Z))$. Then there exists a neighbourhood $\mathcal{U} \times \mathcal{V}$ of (u_0, v_0) and a continuous map $f : \mathcal{V} \to \mathcal{U}$ such that $f(v_0) = u_0$ and*

$$F(f(v), v) = 0 \quad \text{for all } v \in \mathcal{V}. \tag{2.3}$$

Moreover, all solutions in $\mathcal{U} \times \mathcal{V}$ are of the form (2.3).

Theorem 2.2 is a natural generalization of the implicit function theorem on $\mathbb{R}^d$ [Rud76]. All the assumptions on F and its derivatives only have to hold locally and f is smoother if F is.

Example 2.3. Consider the **transcritical bifurcation** (1.21) again

$$F(u, p) = u(p - u), \qquad u \in \mathbb{R} = X,\ p \in \mathbb{R} = Y,\ Z = \mathbb{R}. \tag{2.4}$$

Consider $(u, p) = (0, p)$, which always solves (2.4) and calculate

$$(\mathrm{D}_u F)(0, p) = p - 2 \cdot 0 = p \qquad \text{and} \qquad (\mathrm{D}_p F)(0, p) = 0.$$

So the implicit function theorem applies using $D_u F$ as long as $p \neq 0$ and fails at the bifurcation value $p = 0$. In particular, for $p \neq 0$ the last part of the implicit function theorem guarantees local *uniqueness* of a *branch* of solutions while two solution branches cross at $p = 0$; see Figure TODO.

Definition 2.4. Consider $F : \mathcal{U} \subset X \to Z$ with F Fréchet differentiable and let $u_0 \in \mathcal{U} \subset X$. F is a nonlinear **Fredholm operator** if the following conditions hold:

- $\dim(\mathcal{N}[(D_u F)(u_0)]) < \infty$, where $\mathcal{N}[\cdot]$ denotes the **nullspace**,
- $\operatorname{codim}(\mathcal{R}[(D_u F)(u_0)]) < \infty$, where $\mathcal{R}[\cdot]$ denotes the **range**,

and where $\operatorname{codim}(S) := \dim(Z - S)$. Then define the **Fredholm index** as

$$\text{Fredholm index } := \dim(\mathcal{N}[(D_u F)(u_0)]) - \operatorname{codim}(\mathcal{R}[(D_u F)(u_0)]).$$

It can be shown that the Fredholm index is independent of $u_0 \in \mathcal{U}$; essentially Fredholm operators have 'relatively small nullspace' and miss a 'relatively small part' of the range; see Figure TODO.

The next goal is to reduce the general infinite-dimensional problem

$$F(u, v) = 0, \qquad F : \mathcal{U} \times \mathcal{V} \subset X \times Y \to Z \tag{2.5}$$

to a more tractable finite-dimensional problem. We usually assume $\mathcal{V} \subset \mathbb{R}$ (or $\mathbb{R}^d$ for some d) and it remains to reduce the u-component. This will be achieved using the **Lyapunov-Schmidt method** . Assume that

$$F(u_0, v_0) = 0, \quad F, D_u F \in C, \quad F(\cdot, v_0) : X \to Z \text{ is a Fredholm operator.}$$

Then it is relatively easy to show that

$$\begin{aligned} X &= \mathcal{N}[(D_u f)(u_0, v_0)] \oplus X_0 =: \mathcal{N} \oplus X_0, \\ Z &= \mathcal{R}[(D_u f)(u_0, v_0)] \oplus Z_0 =: \mathcal{R} \oplus Z_0, \end{aligned}$$

where $\mathcal{N}$ and Z_0 are finite-dimensional. Then define projections

$$\begin{aligned} P &: X \to \mathcal{N}, && \text{along } X_0, \\ Q &: Z \to Z_0, && \text{along } \mathcal{R}. \end{aligned}$$

Lemma 2.5. *P, Q are continuous.*

Proof. Apply the closed graph theorem (recall: $T : X \to Z$ is continuous iff $\{(x, z) : Tx = z\}$ is closed). $\square$

Theorem 2.6. *(**Lyapunov-Schmidt Reduction**) Under the assumptions in this section, there exists a neighbourhood $\mathcal{U} \times \mathcal{V}$ of (u_0, v_0) such that $F(u, v) = 0$ is equivalent in $\mathcal{U} \times \mathcal{V}$ to the finite-dimensional problem*

$$\Phi(\tilde{u}, v) = 0, \qquad (\tilde{u}, v) \in \tilde{\mathcal{U}_1} \times \mathcal{V}_1 \subset \mathcal{N} \times \mathcal{V} \tag{2.6}$$

and Φ is continuous with $\Phi(\tilde{u}_1, v_1) = 0$ for $(\tilde{u}_1, v_1) \in \tilde{\mathcal{U}_1} \times \mathcal{V}_1$.

Remark: Note that $\tilde{u} \in \mathcal{N}$, and $v_1 \in \mathbb{R}^d$ by assumption, really means that (2.6) is finite-dimensional. A more precise expression for the **bifurcation function** Φ will be provided below.

Proof. (of Theorem 2.6) The equation $F(u, v) = 0$ is equivalent to the system

$$\begin{aligned} QF(Pu + (\mathrm{Id} - P)u, v) &= 0, \\ (\mathrm{Id} - Q)F(Pu + (\mathrm{Id} - P)u, v) &= 0. \end{aligned} \tag{2.7}$$

Define

$$\tilde{u} := Pu, \qquad w := (\mathrm{Id} - P)u, \qquad G(\tilde{u}, w, v) := (\mathrm{Id} - Q)F(\tilde{u} + w, v). \tag{2.8}$$

We work locally near (u_0, v_0) so it is natural to also define

$$\tilde{u}_0 := Pu_0, \qquad w_0 := (\mathrm{Id} - P)u_0 \tag{2.9}$$

and observe that $G(\tilde{u}_0, w_0, v_0) = 0$ by (2.7). The key observation of the proof is that

$$(\mathrm{D}_w G)(\tilde{u}_0, w_0, v_0) = (\mathrm{Id} - Q)(D_u F)(u_0, v_0) : X_0 \to \mathcal{R} \quad \text{is bijective.}$$

In particular, $\mathrm{D}_w G$ acts 'nicely' on the infinite-dimensional parts of X and Z; here the finiteness assumptions of the Fredholm property are absolutely crucial. Now we may just apply the Implicit Function Theorem 2.2 and obtain

$$\psi : \tilde{\mathcal{U}_1} \times \mathcal{V}_1 \to X_0, \quad \psi(\tilde{u}, v) = w, \quad \psi(\tilde{u}_0, v_0) = w_0. \tag{2.10}$$

Inserting $\psi(\tilde{u}, v) = w$ into (2.7) yields

$$0 = QF(\tilde{u} + w, v) = QF(\tilde{u} + \psi(\tilde{u}, v), v), \tag{2.11}$$

$$0 = (\mathrm{Id} - Q)F(\tilde{u} + w, v) = (\mathrm{Id} - Q)F(\tilde{u} + \psi(\tilde{u}, v), v), \tag{2.12}$$

where (2.12) holds by construction and (2.11) is a finite-dimensional problem

$$\Phi(\tilde{u}, v) := QF(\tilde{u} + \psi(\tilde{u}, v), v) = 0. \tag{2.13}$$

Continuity of Φ also follows from the Implicit Function Theorem. $\square$

Corollary 2.7. *Consider the same setup as in Theorem 2.6 and* $F \in C^1$*. Then* $\psi \in C^1$*,* $\Phi \in C^1$ *and*

$$\mathrm{D}_{\tilde{u}}\psi(\tilde{u}_0, v_0) = 0 \in \mathcal{L}(\mathcal{N}, X_0), \tag{2.14}$$

$$\mathrm{D}_{\tilde{u}}\Phi(\tilde{u}_0, v_0) = 0 \in \mathcal{L}(\mathcal{N}, Z_0). \tag{2.15}$$

Proof. Regularity is clear and just follows from the Implicit Function Theorem. The idea is to just differentiate (2.12) with respect to $\tilde{u}$ so

$$(\mathrm{Id} - Q)(\mathrm{D}_u F)(\tilde{u} + \psi(\tilde{u}, v), v)[\mathrm{D}_{\tilde{u}}\psi + \mathrm{Id}_{\mathcal{N}}] = 0.$$

Evaluating at $(\tilde{u}_0, v_0)$, and using $(\mathrm{D}_u F)\mathrm{Id}_{\mathcal{N}} = 0$, yields

$$(\mathrm{Id} - Q)(\mathrm{D}_u F)(\tilde{u}_0 + w_0, v_0)[\mathrm{D}_{\tilde{u}}\psi(\tilde{u}_0, v_0)] = 0.$$

Since $\mathrm{D}_{\tilde{u}}\psi(\tilde{u}_0, v_0)$ maps into X_0, which is complementary to $\mathcal{N}$, the last expression will only vanish if $\mathrm{D}_{\tilde{u}}\psi(\tilde{u}_0, v_0) = 0$ so (2.14) holds. Differentiating Φ is similar and yields (2.15). □

Exercise 2.8. Carry out the last step in the proof of Corollary 2.7. ◊

Exercise 2.9. Let $F(u, p) := -\Delta u + f(u, p)$, where $f(u, p) = u(p - u)$ for all $p \in \mathbb{R}$. Consider Δ as an operator with Dirichlet boundary conditions for $x \in [0, \pi]$, i.e.,

$$-\Delta u = -\frac{\partial^2 u}{\partial x^2}, \quad u(0) = 0 = u(\pi). \tag{2.16}$$

Prove that the Fréchet derivative $L(u_0, p_0) := (\mathrm{D}_u F)(u_0, p_0)$ is $L(u_0, p_0) = -\Delta + p_0 - 2u_0$. ◊

Exercise 2.10. Calculate the **eigenvalues** λ_j and **eigenfunctions** $e_j = e_j(x)$ of $L(0, p_0)$, i.e., when do we have $L(0, p_0)e_j = \lambda_j e_j$. ◊

Background and Further Reading: The material in this section is based upon [Kie04], which is a very detailed, but not necessarily easy to digest, account of bifurcation theory for wide classes PDEs. Other classic accounts of this topic are [CH82, IJ97].

3 Crandall-Rabinowitz and Local Bifurcations

The next step is to prove a result, which we can actually apply to find local bifurcations of certain classes of PDEs. Consider

$$F(u,p)=0, \qquad F: X\times\mathbb{R}\to Z, \qquad F(0,p)\equiv 0, \tag{3.1}$$

where the last condition means that we always have a **trivial** (or **homogeneous**) **solution branch**. Furthermore, we assume throughout this section that

(A1) $\dim(\mathcal{N}[\mathrm{D}_uF(0,p_0)]) = 1 = \mathrm{codim}(\mathcal{R}[\mathrm{D}_uF(0,p_0)])$ so $F(\cdot,p_0)$ is a nonlinear Fredholm operator of index zero at p_0,

(A2) $F\in C^3$ in an open neighbourhood of the trivial branch.

We view D^2_{up} as an element of $\mathcal{L}(X,Z)$. We may shift the parameter, if necessary, and assume without loss of generality that the interesting point, where the Implicit Function Theorem fails, is at $p_0=0$. The next theorem is a fundamental result in the field:

Theorem 3.1. *(**Crandall-Rabinowitz Theorem**; [CR71]) Consider* (3.1) *and assume that (A1)-(A2) hold as well as*

$$\mathcal{N}[\mathrm{D}_uF(0,0)] = span[e_0], \quad (\mathrm{D}^2_{up}F)(0,0)e_0 \notin \mathcal{R}[\mathrm{D}_uF(0,0)]. \tag{3.2}$$

*for $e_0\in X$ and $\|e_0\|_X=1$. Then there is a **nontrivial branch** of solutions described by a C^1-curve through* $(u,p)=(0,0)$

$$\{(u(s),p(s)) \;:\; s\in(-s_0,s_0),\; (u(0),p(0))=(0,0)\}, \tag{3.3}$$

which satisfies $F(u(s),p(s))=0$ locally, and all solutions in a neighbourhood of $(0,0)$ *are either the trivial solution or on the nontrivial curve* (3.3)*.*

Remark: An illustration of the bifurcation point at $(u,p)=(0,0)$ is shown in Figure TODO, together with a few different situations we expect to appear from the finite-dimensional cases discussed in Section 1.

Proof. (of Theorem 3.1) The idea is, as we have done before in similar situations, to apply the Implicit Function Theorem to get the non-trivial curve. Using Lyapunov-Schmidt Reduction, the bifurcation function Φ satisfies

$$\Phi(\tilde{u},p)=0, \qquad \Phi:\tilde{\mathcal{U}}_1\times\mathcal{V}_1\to Z_0, \quad \dim(Z_0)=1.$$

Using Theorem 2.6 and Corollary 2.7, as well as the associated notations, we also have

$$\psi(0,p)=0, \qquad \mathrm{D}_p\psi(0,p)=0, \qquad \Phi(0,p)=0, \tag{3.4}$$

in a suitable neighbourhood of $p = 0$. Using the last observation about Φ gives us

$$\Phi(\tilde{u}, p) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\Phi(t\tilde{u}, p)\ \mathrm{d}t \quad \Rightarrow \quad \Phi(\tilde{u}, p) = \int_0^1 \mathrm{D}_{\tilde{u}}\Phi(t\tilde{u}, p)\tilde{u}\ \mathrm{d}t. \tag{3.5}$$

Next, let $\tilde{u} = se_0$, $s \in (-s_0, s_0)$ and observe that we get nontrivial solutions to $\Phi(\tilde{u}, p) = 0$ if we can solve

$$\tilde{\Phi}(s, p) := \int_0^1 \mathrm{D}_{\tilde{u}}\Phi(tse_0, p)e_0\ \mathrm{d}t = 0 \tag{3.6}$$

for $s \neq 0$. To solve the last equation, we want to use the Implicit Function Theorem and this requires computing a derivative, so we carry out the following main computation of the proof. We need $\mathrm{D}_p((\mathrm{D}_{\tilde{u}}\Phi)(\tilde{u}, p)e_0) =$

$$\begin{aligned}
&\overset{(2.13)}{=} \mathrm{D}_p(Q(\mathrm{D}_u F)(\tilde{u} + \psi(\tilde{u}, p), p)(e_0 + (\mathrm{D}_{\tilde{u}}\psi)(\tilde{u}, p)e_0)) \\
&= \underbrace{Q(\mathrm{D}^2_{uu}F)(\tilde{u} + \psi(\tilde{u}, p), p)[e_0 + (\mathrm{D}_{\tilde{u}}\psi)(\tilde{u}, p)e_0, \mathrm{D}_p\psi(\tilde{u}, p)]}_{:=(\mathrm{T1})} \\
&\quad + \underbrace{Q(\mathrm{D}^2_{up}F)(\tilde{u} + \psi(\tilde{u}, p), p)(e_0 + (\mathrm{D}_{\tilde{u}}\psi)(\tilde{u}, p)e_0)}_{:=(\mathrm{T2})} \\
&\quad + \underbrace{Q(\mathrm{D}_u F)(\tilde{u} + \psi(\tilde{u}, p), p)((\mathrm{D}^2_{p\tilde{u}}\psi)(\tilde{u}, p)e_0))}_{:=(\mathrm{T3})}
\end{aligned}$$

and evaluate the last expression at $(\tilde{u}, p) = (0, 0)$. Then $(\mathrm{T1}) = 0$ since $(\mathrm{D}_p\psi)(0, 0) = 0$ by (3.4). Furthermore, $(\mathrm{T3}) = 0$ since Q projects along $\mathcal{R}$ to the complement of the range of the linearized problem. So we are left with (T2). Now we can compute

$$\begin{aligned}
(\mathrm{D}_p\tilde{\Phi})(0, 0) &= \int_0^1 (\mathrm{D}_p\mathrm{D}_{\tilde{u}}\Phi)(0, 0)e_0\ \mathrm{d}t \\
&= Q(\mathrm{D}^2_{\tilde{u}p}F)(0, 0)e_0 \neq 0 \in Z_0,
\end{aligned} \tag{3.7}$$

where the last conclusion about the non-equality with zero follows from the assumption (3.2). Finally, applying the Implicit Function Theorem 2.2 and get a curve $\varphi : (-s_0, s_0) \to \mathcal{V}_1$, such that $\varphi(0) = 0$ and $\tilde{\Phi}(s, \varphi(s)) = 0$ near $s = 0$. This implies

$$\Phi(se_0, \varphi(s)) = s\tilde{\Phi}(s, \varphi(s)) = 0$$

and the curve given by

$$(u(s), p(s)) = (se_0 + \psi(se_0, \varphi(s)), \varphi(s)), \qquad s \in (-s_0, s_0),$$

has all the required properties we stated in the theorem. $\square$

Frequently, the situation in the Crandall-Rabinowitz Theorem is also described as **bifurcation from a simple eigenvalue**, which makes sense as $\dim \mathcal{N} = 1$ by assumption so the interesting eigenspace indeed has a simple eigenvalue.

Corollary 3.2. *The tangent vector to the nontrivial solution curve at* $(u,p) = (0,0)$ *is given by* $(e_0, \dot{p}(0))$.

The proof of the corollary will be left as an exercise; see also Figure TODO. The next step is to determine the shape of the nontrivial solution curve more precisely, which requires us to evaluate the derivative $\dot{p}(0) = \frac{\mathrm{d}}{\mathrm{d}s}p(s)\big|_{s=0}$. It is helpful to get another representation of the projection $Q : Z \to Z_0$ to do this calculation.

Lemma 3.3. *Suppose* $Z_0 = span[g_0]$, $g_0 \in Z$, $\|g_0\|_Z = 1$, *then*

$$Qz = \langle z, g_0' \rangle g_0, \qquad \textit{for some } g_0' \in Z' \textit{ and for all } z \in Z, \tag{3.8}$$

where $\langle \cdot, \cdot \rangle$ *denotes the* ***duality pairing*** *between* Z *and the* ***dual space*** Z'*; furthermore, we have* $\langle g_0, g_0' \cdot \rangle = 1$.

Proof. We apply the **Hahn-Banach Theorem** [Rud91] (or more precisely an immediate consequence of the Hahn-Banach Theorem) to obtain a vector $g_0' \in Z'$ such that

$$\langle g_0, g_0' \rangle = 1, \qquad \langle z, g_0' \rangle = 0 \quad \forall z \in \mathcal{R}(\mathrm{D}_u F(0,0)).$$

The result now follows immediately. $\square$

The following result is an elegant formula that can, however, be difficult to evalute in some situations.

Theorem 3.4. *The derivative* $\dot{p}(0)$ *of the nontrivial solution curve from Theorem 3.1 is given by*

$$\dot{p}(0) = -\frac{1}{2}\frac{\langle (\mathrm{D}^2_{uu}F)(0,0)[e_0,e_0], g_0' \rangle}{\langle (\mathrm{D}^2_{up}F)(0,0)e_0, g_0' \rangle}. \tag{3.9}$$

Proof. Recall that $\tilde{\Phi}(s, p(s)) = 0$, locally near $s = 0$, and differentiate

$$\frac{\mathrm{d}}{\mathrm{d}s}\tilde{\Phi}(s,p(s))\bigg|_{s=0} = (\mathrm{D}_s\tilde{\Phi})(0,0) + (\mathrm{D}_p\tilde{\Phi})(0,0)\dot{p}(0) = 0. \tag{3.10}$$

We proved in (3.7) that $(\mathrm{D}_p\tilde{\Phi})(0,0) \neq 0$ and computed an expression for it in terms of F. So it remains to calculate

$$(\mathrm{D}_s\tilde{\Phi})(0,0) = \int_0^1 (\mathrm{D}^2_{\tilde{u}\tilde{u}}\Phi)(0,0)[e_0, te_0]\ \mathrm{d}t = \frac{1}{2}Q(\mathrm{D}^2_{uu}F)(0,0)[e_0,e_0] \tag{3.11}$$

where we have used that $(\mathrm{D}_{\tilde{u}}\Phi) = Q(\mathrm{D}_u F)\mathrm{Id}_{\mathcal{N}}$ and bilinearity of the second derivative operator. Using (3.10), (3.11) and Lemma 3.3 yields the result. $\square$

If $\dot{p}(0) \neq 0$ then we expect the structure of a **transcritical bifurcation** similar to Example 1.6. So let us make a sanity check.

Example 3.5. Consider the ODE (1.21) again

$$u' = u(p-u), \qquad F(u,p) := u(p-u). \tag{3.12}$$

So we have $X = \mathbb{R}$, $Y = \mathbb{R}$ and $Z = \mathbb{R}$. The Fréchet derivatives just become usual partial derivatives and the dual pairing just becomes the inner product (i.e. just multiplication in $\mathbb{R}$ here). Furthermore, the point of interest clearly is $(u,p) = (0,0)$, where the implicit function theorem fails as

$$(\mathrm{D}_u F)(0,0) = (\partial_u F)(0,0) = 0 = (\partial_p F)(0,0) = (\mathrm{D}_p F)(0,0).$$

So $\mathrm{span}[e_0] = \mathcal{N}[\mathrm{D}_u F(0,0)]$ for $e_0 = 1$ and

$$(\mathrm{D}^2_{up}F)(0,0)e_0 = (\partial^2_{up}F)(0,0)e_0 = 1 \cdot 1 = 1 \notin \mathcal{R}[\mathrm{D}_u F(0,0)] = \{0\}.$$

So the Crandall-Rabinowitz Theorem 3.1 applies. In addition, we can easily use the formula (3.9)

$$\dot{p}(0) = -\frac{1}{2}\frac{\langle (\mathrm{D}^2_{uu}F)(0,0)[e_0,e_0], g_0'\rangle}{\langle (\mathrm{D}^2_{up}F)(0,0)e_0, g_0'\rangle} = -\frac{1}{2}\frac{(-2)\cdot 1}{1 \cdot 1} = 1.$$

Hence, we recover the result in Figure TODO.

Example 3.5 demonstrates that we may expect formula (3.9) to be quite widely applicable, and easy to calculate, if X is a Hilbert space. Sometimes the explicit calculations can even be done in quite general Banach spaces but these examples are beyond, what we can present here.

Theorem 3.6. *(see [Kie04]) The second derivative $\ddot{p}(0)$ of the nontrivial solution curve from Theorem* (3.1) *is given by*

$$\ddot{p}(0) = -\frac{1}{3}\frac{\langle (\mathrm{D}^3_{\tilde{u}\tilde{u}\tilde{u}}\Phi)(0,0)[e_0,e_0,e_0], g_0'\rangle}{\langle (\mathrm{D}^2_{up}F)(0,0)e_0, g_0'\rangle}. \tag{3.13}$$

In the Crandall-Rabinowitz Theorem case, when $\dot{p}(0) = 0$ and $\ddot{p}(0) \neq 0$ we have a **pitchfork bifurcation**. The pitchfork is **subcritical** if $\ddot{p}(0) < 0$ and **supercritical** if $\ddot{p}(0) > 0$.

One last step, we need to verify is that many of the operators we are dealing with are indeed Fredholm. For example, consider again the problem

$$\partial_t u = \Delta u + f(u,p) = F(u,p), \tag{3.14}$$

where we could, e.g., take $f(u,p) = u(1-u)(u-p)$ or $f(u,p) = u(p-u)$. Then the ansatz $u = u^* + \epsilon w$ for a steady state u^*, say for $p = 0$, easily leads to the linearized problem

$$\partial_t w = \Delta w + (D_u f)(u^*,0)w = (D_u F)(u^*,0)w. \tag{3.15}$$

So a typical class of operators we should deal with are of the form "Laplacian + lower order terms".

Definition 3.7. Consider the differential operator

$$Lu := -\sum_{i,j=1}^{N} a^{ij}(x)\partial_{x_i}\partial_{x_j}u + \sum_{j=1}^{N} b^j(x)\partial_{x_j}u + c(x)u.$$

for $x \in \Omega \subset \mathbb{R}^N$ with sufficiently regular (say: smooth) coefficient functions. Then L is called **uniformly elliptic** if there exists a constant $K > 0$ such that

$$\sum_{i,j=1}^{N} a^{ij}(x)\xi_i\xi_j \geq K\|\xi\|^2 \qquad \forall x \in \Omega, \xi \in \mathbb{R}^N.$$

The following is quite a remarkable, and extremely useful, fact.

Theorem 3.8. *Let the $\mathcal{D}(L) := H^2(\Omega) \cap H_0^1$ be the domain of L and consider it as an operator $L : \mathcal{D}(L) \to H^0(\Omega)$, then L is a Fredholm operator of index zero.*

Remark: Of course, by elliptic regularity [Eva02], we already know that the eigenfunctions we are going to get are not only in some Sobolev space but are actually classical smooth solutions.

Proof. (Sketch; for more details see [Kie04]) By standard elliptic PDE theory, one may see that the operator $L - c\ \mathrm{Id} : \mathcal{D}(L) \to H^0(\Omega)$ is bounded and bijective for a suitable constant $c \geq 0$ (indeed, just look at solving $Lu - cu = g$). Then one may see that the operator $(L - c\ \mathrm{Id})^{-1} =: K_c : H^0(\Omega) \to H^0(\Omega)$ is compact (images of convergent sequences have a convergent subsequence). For $g \in H^0(\Omega)$ we can compute

$$Lu = g \quad \Leftrightarrow (\mathrm{Id} + cK_c)u = K_c g. \tag{3.16}$$

One alternative characterization of Fredholm operators is that they are precisely those, which are invertible up to a compact operator (which accounts for the kernel and range properties we used above to define Fredholm operators). Hence, one sees that $\mathrm{Id} + cK_c$ is Fredholm. Furthermore, (3.16) then implies

$$\dim(\mathcal{N}[L]) = \dim(\mathcal{N}[\mathrm{Id} + cK_c]) = n < \infty$$

as well as

$$g \in \mathcal{R}[L] \Leftrightarrow K_c g \in \mathcal{R}[\mathrm{Id} + cK_c].$$

One now just has to show that the range also has dimension n to get that the index is zero. From the decomposition

$$g = (\mathrm{Id} + cK_c)g - cK_c g$$

one finds

$$H^0(\Omega) = \mathcal{R}[\mathrm{Id} + cK_c] + \mathcal{R}[K_c] \quad \Rightarrow \; H^0(\Omega) = \mathcal{R}[\mathrm{Id} + cK_c] + K_c(Z_0)$$

for some n-dimensional space $Z_0 \subset H^0(\Omega)$ with $\mathcal{R}[L] \cap Z_0 = \{0\}$. Then it is relatively easy to conclude with a few further steps (check it!) that $\mathrm{codim}(\mathcal{R}[L]) = n$. $\square$

Exercise 3.9. Prove Corollary 3.2. $\Diamond$

Exercise 3.10. Write down a one-dimensional ODE $u' = f(u, p_1, p_2)$, which has a pitchfork bifurcation upon varying p_1 through 0 and the pitchfork changes from sub- to super-critical if p_2 is varied through 0. $\Diamond$

Exercise 3.11. Let $x \in (0, \pi) =: \Omega$ and consider the problem

$$\partial_t u = \partial_{xx}^2 u + f(u, p), \tag{3.17}$$

where Δ is understood as an operator with the domain $L^2(0, \pi)$ and homogeneous Dirichlet boundary conditions. Now try to apply the results presented so far to the two cases

$$f(u, p) = u(p - u) \qquad \text{and} \qquad f(u, p) = u(p - u^2). \tag{3.18}$$

As a more advanced question: What do you expect to happen if we perturb f by $f(u, p) + \delta \tilde{f}(u)$ with $\tilde{f}(0) \neq 0$? This last case is sometimes called **imperfection**. $\Diamond$

Background and Further Reading: The material in this section is based upon [Kie04] and the last exercise/question is motivated by [IJ97]. It can also be useful to explore to study associated numerical algorithms as calculating local bifurcation points can beome tedious and/or impossible quite quickly; for 1D-boundary-value problems see [DCD$^+$07] and for elliptic PDEs in $\mathbb{R}^2$ see [UWR14, Ban07].

4 Stability and Spectral Theory

The last question we have to address is stability for the evolution problem

$$\partial_t u = F(u,p), \qquad u : [0,+\infty) \to X, \quad u = u(t) \in X, \quad p \in \mathbb{R}. \tag{4.1}$$

Recall that the **spectrum** $\sigma(A)$ of a linear operator A consists of all elements $\lambda \in \mathbb{C}$ such that $(A - \lambda \mathrm{Id})$ is not invertible or the inverse $(A-\lambda\mathrm{Id})^{-1}$ is not a bounded operator; careful: this may crucially depend on the choice of spaces for A, the associated domain for A and implicitly we understand that if A is not a closed operator, we always look at its closure.

Definition 4.1. A solution branch $(u^*(p),p)$ for $F(u,p) = 0$ is called (**linearly**) **stable** at p^* if $\sigma(\mathrm{D}_u F(u^*,p^*))$ is properly contained in the left half of the complex plane.

We continue with the notation from Section 3 and assume in addition that the Banach space X is continuously embedded into the Banach space Z. Consider the case of a simple eigenvalue $\lambda(s)$ with $\lambda(0) = 0$, which occurs in the Crandall-Rabinowitz Theorem 3.1. In particular, simple eigenvalue means that $\mathcal{N}[\mathrm{D}_u F(0,0)] = \mathrm{span}[e_0]$ implies $e_0 \notin \mathcal{R}[\mathrm{D}_u F(0,0)]$. This implies we have a decomposition

$$Z = \mathcal{N}[\mathrm{D}_u F(0,0)] \oplus \mathcal{R}[\mathrm{D}_u F(0,0)]$$

and an associated induced decomposition

$$X = \mathcal{N}[\mathrm{D}_u F(0,0)] \oplus (X \cap \mathcal{R}[\mathrm{D}_u F(0,0)]).$$

Note that this identifies the projection Q used in the Lyapunov-Schmidt method to prove Crandall-Rabinowitz as $Q : Z \to \mathcal{N}[\mathrm{D}_u F(0,0)]$.

We can determine local stability of the trivial branch from this eigenvalue if we assume that $\sigma(\mathrm{D}_u F(0,p(s))) - \{\lambda(s)\}$ is properly contained in the left-half complex plane for $s \approx 0$; see Figure TODO. Parametrize λ by p such that $\lambda = \lambda(p)$, $\lambda(0) = 0$, and consider the **eigenvalue perturbation**

$$(\mathrm{D}_u F)(0,p)(e_0 + w(p)) = \lambda(p)(e_0 + w(p)), \qquad w(0) = 0. \tag{4.2}$$

Differentiating (4.2) with respect to p and evaluating at $p = 0$ yields

$$(\mathrm{D}^2_{up} F)(0,0)e_0 + (\mathrm{D}_u F)(0,0)\frac{\mathrm{d}w}{\mathrm{d}p}(0) = \frac{\mathrm{d}\lambda}{\mathrm{d}p}(0)\ e_0. \tag{4.3}$$

Recall the dual pairing with $g_0' \in Z'$ satisfies the property that $\langle z, g_0' \rangle = 0$ for all $z \in \mathcal{R} = \mathcal{R}[(\mathrm{D}_u F)(0,0)]$. Due to the identification of Z_0 with

$\mathcal{N}[\mathrm{D}_u F(0,0)]$ as discussed above, we can just apply the pairing $\langle \cdot, g_0' \rangle$ to (4.3) obtain

$$\frac{\mathrm{d}\lambda}{\mathrm{d}p}(0) = \langle (\mathrm{D}^2_{up} F)(0,0) e_0, g_0' \rangle.$$

In particular, the second condition in (3.2) in the Crandall-Rabinowitz Theorem has a reformulation

$$(\mathrm{D}^2_{up} F)(0,0) e_0 \notin \mathcal{R}[\mathrm{D}_u F(0,0)] \quad \Leftrightarrow \quad \frac{\mathrm{d}\lambda}{\mathrm{d}p}(0) \neq 0.$$

In particular, we may interpret the non-vanishing condition of the derivative of $\lambda(p)$ at $p = 0$ as an eigenvalue **crossing with nonzero speed** upon variation of p or as a **transversality condition**.

So far, we have implicitly assumed that we know the spectrum of the linear operator

$$(\mathrm{D}_u F)(0,0) \in \mathcal{L}(X, Z). \tag{4.4}$$

However, for practical problems this may be far from trivial so we consider a few standard cases, where explicit calculations turn out to be possible.

Example 4.2. Let $A \in \mathbb{R}^{d \times d}$ be a symmetric matrix and consider the linear ODE

$$u' = Au, \quad u = u(t) \in \mathbb{R}^d, \qquad \Rightarrow \; u(t) = \mathrm{e}^{tA} u(0) \tag{4.5}$$

using the matrix exponential. Without loss of generality (upon applying a coordinate change) we may assume that A is in diagonal form with eigenvalues λ_j, $j \in \{1, 2, \ldots, d\}$. If $\lambda_j < 0$ for all j then $u = 0 := (0, 0, \ldots, 0) \in \mathbb{R}^d$ is a (globally asymptotically) stable steady state. Ordering the eigenvalues

$$\cdots \leq \lambda_2 \leq \lambda_1 < 0$$

we see that $u(t) \sim K \mathrm{e}^{\lambda_1 t} e_1$ as $t \to +\infty$, where $A e_1 = \lambda_1 e_1$. In particular, the behaviour is dominated by the weakest attracting direction in the long-time limit; see Figure TODO. The rate of collapse onto this direction is given by the **spectral gap** $\lambda_2 - \lambda_1$. ♦

It is very helpful to just calculate a few examples.

Example 4.3. Consider the eigenvalue problem for the Laplacian on an interval

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} u = \lambda u \tag{4.6}$$

for $x \in [0, \rho]$. Denote the eigenfunctions by e_j and the associated eigenvalues by λ_j. We find

$$\begin{array}{lll} u(0) = 0 = u(\rho) & \Rightarrow & \lambda_j = -\left(j\pi/\rho\right)^2, \ e_j(x) = \sin(j\pi x/\rho), \ j \geq 1, \\ \frac{\mathrm{d}u}{\mathrm{d}x}(0) = 0 = \frac{\mathrm{d}u}{\mathrm{d}x}(\rho) & \Rightarrow & \lambda_j = -\left(j\pi/\rho\right)^2, \ e_j(x) = \cos(j\pi x/\rho), \ j \geq 0. \end{array}$$

Hence, the eigenvalues and eigenfunctions clearly depend crucially on the chosen boundary conditions. ♦

Example 4.4. Consider the Laplacian on a rectangle $\Omega = [0, \rho_1] \times [0, \rho_2]$ and the eigenvalue problem

$$\Delta u = \lambda u, \qquad u(x) = 0 \text{ for } x \in \partial\Omega. \tag{4.7}$$

Then the eigenvalues and eigenfunctions are easily checked to be

$$e_{jk}(x) = \sin(j\pi x_1/\rho_1)\sin(j\pi x_2/\rho_2), \quad \lambda_{jk} = -\left(j\pi/\rho_1\right)^2 - \left(j\pi/\rho_2\right)^2,$$

for $j, k \geq 1$. ♦

However, beyond simple rectangular (or more generally hypercube) domains, giving explicit formulas for the eigenfunctions and eigenvalues of the Laplacian is usually not possible. Nevertheless, there are very remarkable abstract results.

Theorem 4.5. *(**Spectrum of Elliptic Operators**; see [Eva02]) Consider the eigenvalue problem*

$$-Lu = \lambda u, \ \text{in } \Omega \qquad u = 0, \ \text{on } \partial\Omega, \tag{4.8}$$

where L is uniformly elliptic, then L has an at most countable set of eigenvalues λ_j with $\lambda_j \to -\infty$.

Remark: One should be careful with sign-conventions as one also frequently finds in the literature results for $Lu = \lambda u$.

If one consider L on the domain $\mathcal{D}(L) = H^1_0(\Omega) \cap H^2(\Omega)$ with $L : \mathcal{D}(L) \to H^0(\Omega)$, it is not difficult to see that there are also associated bases of orthonormal eigenfunctions for the relevant Hilbert spaces associated to the eigenvalues.

Example 4.6. Let us return for $x \in \Omega$ to the problem

$$\partial_t u = \Delta u + f(u, p) = F(u, p), \qquad u = 0 \text{ on } \partial\Omega \tag{4.9}$$

from equation (3.14), where we assume that the trivial branch $(u, p) = (0, p)$ exists. The eigenvalue problem of the linearized PDE (see also (3.15)) for the trivial branch is

$$-Lu = (\Delta + (D_u f)(0, p))u = \lambda u \quad \text{or} \quad \Delta u = (\lambda - (D_u f)(0, p))u \tag{4.10}$$

where L is an elliptic operator. So we only have to deal with discrete eigenvalues in the spectrum by Theorem 4.5. On some domains, we can even calculate stability explicitly! For example, if $\Omega = [0, \rho]$ then we see from (4.10) that the eigenvalues of the Laplacian $\Delta u = \tilde{\lambda} u$ are just shifted

$$\begin{aligned} \tilde{\lambda}_j &= \lambda_j - (D_u f)(0, p) \\ \Rightarrow \ \lambda_j &= \tilde{\lambda}_j + (D_u f)(0, p) \\ \Rightarrow \ \lambda_j &= -(j\pi/\rho)^2 + (D_u f)(0, p) \end{aligned}$$

for $j \geq 1$. In particular, we see that the **critical eigenvalue**, which is going to pass through the imaginary axis first is given by $j = 1$. The condition for instability is $\lambda_1 > 0$ or

$$(\pi/\rho)^2 < (D_u f)(0, p)$$

so increasing the domain size ρ will eventually destabilize the steady state, which is then referred to as a **long-wave instability**. Of course, we may also consider the case when ρ is fixed then varying the parameter p could eventually destabilize the system. ⧫

The next example shows that we do not only have to focus on elliptic PDE and can treat other classes with very similar bifurcation-theoretic methods.

Example 4.7. Consider the **thin-film equation** (with constant surface tension) on a periodic domain

$$\partial_t u = -\partial^4_{xxxx} u - \partial_x (f(u) \partial_x u) =: F(u), \quad x \in [0, 2\pi]/(0 \sim 2\pi) = \mathbb{T}^1, \tag{4.11}$$

where $u = u(x, t)$ models the height of a thin fluid film on a substrate and f takes into account the substrate fluid interactions. Clearly any constant $u \equiv u^* \in \mathbb{R}^+$ is a steady state for (4.11). The linearization at u^* is

$$\begin{aligned} \partial_t w &= (\mathrm{D}_u F)(u^*) w = [[\mathrm{D}(-\partial^4_{xxxx} - f(\cdot)\partial^2_{xx} - f'(\cdot)(\partial_x)^2)](u^*)] w, \\ &= -\partial^4_{xxxx} w - f(u^*) \partial^2_{xx} w + \underbrace{f'(u^*) w \cdot \partial^2_{xx} u^* + \cdots}_{=0}, \end{aligned}$$

where we have used that evaluating operators such as ∂_x and ∂^2_{xx} on constants u^* yields zero (it is a good exercise to check the last calculation by setting $u := u^* + \epsilon w$). In the context, we work here, it makes sense to only study perturbations w, which have mean zero. Indeed, from the thin film equation we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \int_0^{2\pi} u(x, t) \ \mathrm{d}x &= -\int_0^{2\pi} \partial^4_{xxxx} u + \partial_x (f(u) \partial_x u) \ \mathrm{d}x \\ &= \left(-\partial^3_{xxx} u - f(u) \partial_x u\right)\Big|_0^{2\pi} = 0 \end{aligned}$$

by periodicity, so we have **mass conservation**, so meaningful perturbations should also have mass conservation. Substituting a **Fourier mode**

$$w_k = \exp(\lambda t)\exp(\mathrm{i}kx), \quad k \neq 0,$$

into the linearized problem yields $\lambda = -k^2(k^2 - f(u^*))$ so if f depends upon parameters, e.g. due to a parametrically changing substrate-fluid interaction, then we can determine the stability of steady states. Note that the sign of f is crucial: (a) $f > 0$ then the second-order term acts destabilizing, which is no surprise as it represents a 'backward-heat-equation'-term and (b) $f < 0$ then $\lambda < 0$ so the 'forward-heat-equation'-term stabilizes. ♦

Exercise 4.8. Show that the general solution of the second-order ODE

$$a\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + b\frac{\mathrm{d}u}{\mathrm{d}x} + cu = 0, \qquad u = u(x),\ x \in \Omega \subset \mathbb{R}$$

is given by $u(x) = \mathrm{e}^{\alpha x}(K_1\cos(\beta x) + K_2\sin(\beta x))$, where $r_\pm = \alpha \pm \mathrm{i}\beta$ with $\alpha, \beta \in \mathbb{R}$ are solutions to $ar^2 + br + c = 0$. Use this result to derive the solutions of (4.6) for homogeneous Dirichlet, homogeneous Neumann as well as homogeneous Robin,

$$-\frac{\mathrm{d}u}{\mathrm{d}x}(0) + u(0) = 0 = -\frac{\mathrm{d}u}{\mathrm{d}x}(L) + u(L),$$

boundary conditions. ♢

Exercise 4.9. Consider (4.9) on $\Omega = (0, \pi)$ and look at the two cases

$$f(u,p) = u(1-u)(u-p), \qquad f(u,p) = u(p-u).$$

Investigate the stability of steady states for both cases under parameter variation of p. ♢

Exercise 4.10. Recall from your PDE course, how to solve the **heat equation** $\partial_t u = \partial^2_{xx}u$ and the **wave equation** $\partial^2_{tt}u = \partial^2_{xx}u$ using **separation of variables**, which is the ansatz $u(x,t) = u_1(x)u_2(t)$. In particular, recall how the eigenvalues of the Laplacian discussed above enter in a formal series solution obtained by separation of variables. ♢

Background and Further Reading: The material for supplementing Crandall-Rabinowitz is from [Kie04], while the spectral theory and thin-film example follows [Lau12]. A good source for classical spectral theory results in applied mathematics is [DL00].

5 Existence of Travelling Waves

Suppose we consider the one-dimensional reaction-diffusion PDE

$$\partial_t u = \partial^2_{xx} u + f(u), \qquad u = u(x,t) \in \mathbb{R},\ x \in \mathbb{R}, \tag{5.1}$$

which we have studied in Sections 3 and 4 as an example several times. We know already a bit about how to deal with the steady state solutions, their stability, and some of their bifurcations. Another natural question is to look for important non-stationary solutions. A key ingredient is the **travelling wave ansatz**

$$u(x,t) = u(x - st) =: u(\xi), \qquad i.e. \quad \xi := x - st, \tag{5.2}$$

where $s \in \mathbb{R}$ is the **wave speed** to be determined. Essentially (5.2) postulates that we only want to look for solutions u, which depend upon a **moving frame** variable $\xi = x - st$. A wave profile moves 'to the left' if $s < 0$, it moves 'to the right' if $s > 0$ and it is a **standing wave** if $s = 0$. Plugging (5.2) into (5.1) yields

$$-s\frac{\mathrm{d}u}{\mathrm{d}\xi} = \frac{\mathrm{d}^2 u}{\mathrm{d}\xi^2} + f(u), \tag{5.3}$$

where we just used the chain rule. It is clear that steady states $u(x,t) \equiv u^* \in \mathbb{R}$ of (5.1) are also steady states (or equilibria) for (5.3). It is those solutions where $u(\xi)$ is not constant, which are interested in here; see Figure TODO. One may naively hope that (5.3) has explicit solution formulas. In generic situations, this is not the case. However, there are examples other than (5.1), where one may be lucky to find nice closed-form solutions.

Example 5.1. Probably the most famous example, where explicit formulas for waves exist is the **Korteweg-deVries** (**KdV**) equation

$$\partial_t u = -u\partial_x u - \partial^3_{xxx} u. \tag{5.4}$$

Substituting (5.2) into (5.4) and re-arranging terms yields

$$-su' + uu' + u''' = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}\xi} =' \tag{5.5}$$

Since $uu' = \frac{1}{2}(u^2)'$, the last equation can be integrated once

$$-su + \frac{1}{2}u^2 + u'' = c_1, \tag{5.6}$$

where c_1 is a constant of integration. If we are only interested in waves which are doubly asymptotic as $\xi \to \pm\infty$ to zero (see also Figure TODO), with vanishing derivatives, we have the conditions

$$\lim_{\xi\to\pm\infty} u(\xi) = 0, \qquad \lim_{\xi\to\pm\infty} u'(\xi) = 0, \qquad \lim_{\xi\to\pm\infty} u''(\xi) = 0, \qquad \cdots \tag{5.7}$$

so the integration constant must be $c_1 = 0$ in this case. Multiplying (5.6) by u' gives

$$-suu' + \frac{1}{2}u^2u' + u''u' = -s\frac{1}{2}(u^2)' + \frac{1}{6}(u^3)' + \frac{1}{2}((u')^2)' = 0, \tag{5.8}$$

which can be integrated. The integration constant is again zero by (5.7) so

$$-s\frac{1}{2}u^2 + \frac{1}{6}u^3 + \frac{1}{2}(u')^2 = 0 \quad \Leftrightarrow \quad 3(u')^2 = (3s-u)u^2. \tag{5.9}$$

Under the further assumptions $u(\xi) \in (0, 3s)$ and taking the positive square-root of the last expression, we get

$$\frac{\sqrt{3}}{u\sqrt{3s-u}}u' = 1. \tag{5.10}$$

This equation is not quite easy enough to be integrated by-hand but upon the substitutions $v^2 = 3s - u$, $u' = -2vv'$ one finds

$$\frac{2\sqrt{3}}{3s-v^2}v' = -1 \quad \Rightarrow \quad \ln\left(\frac{\sqrt{3s}+v}{\sqrt{3s}-v}\right) = -\sqrt{s}\xi + c_2 \tag{5.11}$$

where integration and the method of partial fractions have been used in the last step (see Exercise (5.6)(a)). After a few further calculations, we find up to a shift of the profile of the wave that

$$u(\xi) = 3s\ \mathrm{sech}^2\left[\frac{\sqrt{s}}{2}\xi\right] \quad \Rightarrow \ u(x,t) = 3s\ \mathrm{sech}^2\left[\frac{\sqrt{s}}{2}(x-st)\right], \tag{5.12}$$

which is also called a **solitary wave** or **soliton**. The calculation showed that are several special features of the KdV equation:

(K1) it is "**integrable**" (formally: it can be viewed as an infinite-dimensional Hamiltonian dynamical system with a lot of conserved quantities) which is echoed by the fact that we were able to determine some of the integrals of the moving frame ODE exactly,

(K2) there is a wave for every wave speed s, (K3) higher solitary waves move faster due to the relation between the speed s and the amplitude $3s$.

In general, we cannot expect such a special calculation to hold but it is a great starting point for **perturbation arguments**, i.e., one is interested in small perturbations of the KdV equation, which may not be integrable, but some integrable features persist under perturbation. ♦

The travelling wave ansatz is interesting not only for equations, which are first-order in the time derivative as demonstrated by the next example.

Example 5.2. Consider the **Sine-Gordon** equation

$$\partial^2_{tt} u = \partial^2_{xx} u - \sin u. \tag{5.13}$$

A similar procedure as for KdV works. Using the travelling wave ansatz and the assumption (5.7), a calculation (see Exercise 5.6) shows that

$$u(x,t) = 4 \arctan \left[\exp\left(-\frac{x - st}{\sqrt{1-s^2}} \right) \right] \tag{5.14}$$

is a family of travelling wave solutions. ♦

We continue to make very important general observations about the ODEs obtained in the travelling wave frame and the full PDE system. To illustrate this in a more concrete case, consider again (5.3), which we can re-write as a first-order system

$$\begin{aligned} u_1' &= u_2, \\ u_2' &= -su_2 - f(u_1). \end{aligned} \tag{5.15}$$

So which solutions of (5.15) correspond to nice bounded travelling wave profiles $u(x - st)$? Basically, bounded solutions connecting between the steady states $(u_1, u_2) = (a, 0)$ and $(u_1, u_2) = (b, 0)$ of (5.15) correspond to travelling wave solutions such that

$$\lim_{\xi \to -\infty} u(\xi) = a, \qquad \lim_{\xi \to +\infty} u(\xi) = b$$

as shown in Figure TODO with velocities tending to zero at infinity

$$\lim_{\xi \to -\infty} u_1'(\xi) = \lim_{\xi \to -\infty} u_2(\xi) = 0, \qquad \lim_{\xi \to +\infty} u_1'(\xi) = \lim_{\xi \to +\infty} u_2(\xi) = 0.$$

We also call a and b the **endstates** of the wave; see Figure TODO. The next definition is valid for general ODEs and will be helpful for classifying travelling waves.

Definition 5.3. Consider the ODE $\frac{\mathrm{d}u}{\mathrm{d}\xi} = f(u)$, $u \in \mathbb{R}^d$, with steady states u^* and $\tilde{u}^*$.

- A solution $u(\xi)$ is called a **periodic orbit** (or periodic trajectory) of minimal **period** $\xi_T > 0$ if

$$u(\xi) = u(\xi + \xi_T) \tag{5.16}$$

  and there is no smaller ξ_T such that (5.16) holds.

- A solution $u(\xi)$ is called a **heteroclinic orbit** (or heteroclinic trajectory, or just heteroclinic) between u^* and $\tilde{u}^*$ if

$$\lim_{\xi \to -\infty} u(\xi) = u^*, \qquad \lim_{\xi \to +\infty} u(\xi) = \tilde{u}^*. \tag{5.17}$$

- A solution $u(\xi)$ is called a **homoclinic orbit** (or homoclinic trajectory, or just homoclinic) to u^* if

$$\lim_{\xi \to -\infty} u(\xi) = u^* = \lim_{\xi \to +\infty} u(\xi). \tag{5.18}$$

Remark: Alternatively, one could also have expressed the previous definition using the definitions of α**- and** ω**-limit sets**, for the flow $\varphi(u_0, t)$ associated to $u' = f(u)$,

$$\begin{aligned} \alpha(\mathcal{U}) &:= \{u \in \mathbb{R}^d : \exists t_j,\, t_j \to -\infty \text{ s.t. } \varphi(u_0, t_j) \to u \text{ for some } u_0 \in \mathcal{U}\}, \\ \omega(\mathcal{U}) &:= \{u \in \mathbb{R}^d : \exists t_j,\, t_j \to +\infty \text{ s.t. } \varphi(u_0, t_j) \to u \text{ for some } u_0 \in \mathcal{U}\}. \end{aligned} \tag{5.19}$$

For example, any point on a heteroclinic orbit from u^* to $\tilde{u}^*$ has as the α-limit set u^* and as the ω-limit set $\tilde{u}^*$; see also Figure TODO.

Using Definition 5.3, the following observations/definitions are clear:

(H1) A periodic orbit of the travelling wave ODEs corresponds to a **travelling wave train** solution of the associated PDE; see Figure TODO(a).

(H2) A homoclinic orbit of the travelling wave ODEs corresponds to a **travelling pulse** solution of the associated PDE; see Figure TODO(b).

(H3) A heteroclinic orbit of the travelling wave ODEs corresponds to a **travelling front** solution of the associated PDE; see Figure TODO(c).

Example 5.4. Suppose we study the Nagumo equation, i.e., (5.1) with $f(u) = u(1-u)(u-p)$, say with $p \in (0,1)$. Then (5.15) implies that the travelling wave ODE is

$$\begin{aligned} u_1' &= u_2, \\ u_2' &= -su_2 - u_1(1-u_1)(u_1 - p), \end{aligned} \tag{5.20}$$

where the only steady states occur for $u^l = (0,0)$, $u^m = (p,0)$ and $u^r = (1,0)$; see Figure TODO. The linearization at a steady state u^* is given by

$$w' = \begin{pmatrix} 0 & 1 \\ p - 2(1+p)u_1^* + 3(u_1^*)^2 & -s \end{pmatrix} w =: A(u^*)w. \tag{5.21}$$

For example, if $u^* = u^l$ then the eigenvalues λ^l of $A(u^l)$ satisfy the equation

$$(\lambda^l)^2 + s\lambda^l - p = 0, \qquad \Rightarrow \quad \lambda^l_\pm = -\frac{s}{2} \pm \sqrt{\frac{s^2}{4} + p}$$

and it now follows that u^l is a saddle point as $\lambda^l_- < 0 < \lambda^l_+$. Similarly, it can be checked that u^r is a saddle point as well and u^m is completely unstable for $s < 0$ and stable for $s > 0$. It is instructive to sketch the phase portrait as shown in Figure TODO. Suppose we are interested in left-moving front solutions connecting the states $u = 0$ and $u = 1$. This means looking for a heteroclinic orbit between u^l and u^r, which could potentially exist for some $s < 0$; see Figure TODO. There are analytical proofs that there exists precisely one s for which there is a trajectory

$$\gamma = \gamma(\xi), \text{ such that } \gamma(-\infty) = (0,0) \text{ and } \gamma(+\infty) = (1,0).$$

In particular, it can be shown that the unstable manifold $W^{\mathrm{u}}(u^l)$ and the stable manifold $W^{\mathrm{s}}(u^r)$ are the same curve in the region $\{u_1 > 0, u_2 > 0\}$; see Figure TODO. A similar result holds regarding the existence of a unique wave speed $s > 0$ for a right-moving front.

A viewpoint useful for analytical and numerical arguments is the following procedure:

(L1) Pick a codimension one submanifold $\mathcal{M}$ in phase space 'between' the two steady states we want to connect; for (5.20) a vertical half-line

$$\mathcal{M} = \{u \in \mathbb{R}^2 : u_1 = \kappa, u_2 > 0\}$$

for some suitable fixed $\kappa \in (0,1)$ works well. Observe that the vector field for (5.20) is tangent to $\mathcal{M}$ if and only if $u_2 = 0$, so if we stay in the positive quadrant this degenerate case does not concern us here.

(L2) Compute the intersection submanifolds of the manifolds $W^{\mathrm{u}}(u^l)$, $W^{\mathrm{s}}(u^r)$ with $\mathcal{M}$

$$\mathcal{P}^l := \mathcal{M} \cap W^{\mathrm{u}}(u^l), \qquad \mathcal{P}^r := \mathcal{M} \cap W^{\mathrm{s}}(u^r).$$

For (5.20) these are generically just two points. Note that all objects obviously depend upon the choice of the wave speed parameter s, although this is not shown in the notation.

(L3) Define a bifurcation function by

$$\Phi(s) := \mathrm{d}(\mathcal{P}^l, \mathcal{P}^r) \tag{5.22}$$

where $\mathrm{d}(\cdot,\cdot)$ is a suitable metric measuring the distance on $\mathcal{M}$; for (5.20) we can just take the Euclidean distance between two points; see Figure TODO.

If we can find a zero $s = s_0$ of the bifurcation function ($\Phi(s_0) = 0$) then we have found a connecting orbit between the two steady states as desired. This idea naturally generalizes into higher dimensions and is also called **Lin's method** and $\Phi(s)$ is the **Lin gap**. ♦

Another strategy worth being aware of can be useful if there is some 'integrability' available in the system.

Example 5.5. Consider (5.20) in the standing wave case $s = 0$ and pick for simplicity $p = 1/4$. Then we see that the resulting ODEs in the travelling wave frame form a **Hamiltonian system**

$$\begin{aligned} u_1' &= u_2, \\ u_2' &= -u_1(1-u_1)\left(u_1 - \tfrac{1}{4}\right), \end{aligned} \quad \leftrightarrow \quad \begin{aligned} u_1' &= -\tfrac{\partial H}{\partial u_2}(u_1,u_2), \\ u_2' &= \tfrac{\partial H}{\partial u_1}(u_1,u_2), \end{aligned} \tag{5.23}$$

where $H(u_1,u_2) := -\frac{1}{2}(u_2)^2 + \frac{1}{8}(u_1)^2 - \frac{5}{12}(u_1)^3 + \frac{1}{4}(u_1)^4$ is the **Hamiltonian function** or just 'the Hamiltonian'. The Hamiltonian is always a **first integral** of the flow since

$$\frac{\mathrm{d}}{\mathrm{d}t}H(u_1,u_2) = \frac{\partial H}{\partial u_1}\cdot u_1' + \frac{\partial H}{\partial u_2}\cdot u_2' \overset{(5.23)}{=} u_2'\cdot u_1' - u_1'\cdot u_2' = 0$$

so H is constant along trajectories. Hence, the level curves $\{H(u_1,u_2) = \text{constant}\}$ are trajectories of (5.23). If we want to find a stationary/standing pulse solution of the original PDE (5.1) with the cubic Nagumo nonlinearity, then it suffices to find a homoclinic orbit of the Hamiltonian ODEs (5.23), which is just a bounded level curve of the Hamiltonian function containing the origin. ♦

In summary, we have introduced several important dynamical systems methods to find travelling waves: explicit integration, phase plane analysis, Lin's method, and Hamiltonian structure. However, there are many more, such as the Conley index relating the area to algebraic topology or analytical sub- and super-solution techniques; see references more details.

Exercise 5.6. (a) Justify the steps to go from (5.10) to the soliton solution (5.12). (b) Derive the existence of the family of travelling waves (5.14) for the Sine-Gordon equation. ♢

Exercise 5.7. Consider (5.1) with the cubic nonlinearity $f(u) = u(1 - u)(u - p)$ with $p \in (0, 1)$. Prove that $u \equiv 0, 1$ are stable stationary states for this PDE, while $u \equiv p$ is unstable for the PDE. Therefore, the situation is also referred to as **bistability**. Hence, the stability of the travelling wave ODEs do *NOT* provide information about the actual stability of solutions for the PDE! $\lozenge$

Exercise 5.8. Consider Example 5.4 and study with any numerical method of your choice, how the wave speed of a front depends upon the value of $p \in (0, 1)$. $\lozenge$

Background and Further Reading: The material in this section is based upon [Eva02, GK09]. An excellent description of Lin's method can be found in [KR08]. For the Conley index and bistable equations see [Smo94, MM02] and for sub- and super-solutions in the bistable case [Che97].

6 Pushed and Pulled Fronts

We again consider the one-dimensional reaction-diffusion PDE

$$\partial_t u = \partial_{xx}^2 u + f(u), \qquad u = u(x,t) \in \mathbb{R},\ x \in \mathbb{R}, \tag{6.1}$$

and we can again make the ansatz (5.2), i.e., $u(x,t) = u(x - st) = u(\xi)$, to study the travelling wave. In this section we will focus on the actual value of the wave speed s.

Example 6.1. The main example in this section will be the FKPP nonlinearity

$$f(u) = u(1 - u). \tag{6.2}$$

This quadratic nonlinearity will turn out to have fundamentally different properties from the bistable cubic nonlinearity discussed in detail in Section 5. Indeed, note that we have only two constant steady states $u \equiv 0$ and $u \equiv 1$. The travelling wave frame ODEs are given by

$$\begin{aligned} u_1' &= u_2, \\ u_2' &= -su_2 - u_1(1 - u_1), \end{aligned} \tag{6.3}$$

where $' = \frac{\mathrm{d}}{\mathrm{d}\xi}$ and it is easily checked that $u_l = (0,0)$ and $u_r = (1,0)$ have associated eigenvalues of the linearized system

$$\lambda_\pm^l = \frac{1}{2}\left(\pm\sqrt{s^2 - 4} - s\right), \quad \lambda_\pm^r = \frac{1}{2}\left(\pm\sqrt{s^2 + 4} - s\right). \tag{6.4}$$

Therefore, u^l is a **stable node** if $s \geq 2$ and an **unstable source** if $s \leq -2$, while u^r is always a hyperbolic saddle; see Figure TODO. From Figure TODO, it is apparent that there are infinitely many wave speeds s for which front solutions could exist. ♦

One technique to understand the wave speed better is to try to linearize directly on the level of the PDE (6.1). Suppose u^* is a steady state of (6.1) and also one of the endstates of the wave (e.g. $u(\xi) \to u^*$ as $\xi \to +\infty$) we want to analyze. Consider

$$\partial_t w = \partial_{xx}^2 w + \underbrace{(D_u f)(u^*)}_{=:a^*} w, \qquad w = w(x,t) \in \mathbb{R},\ x \in \mathbb{R}, \tag{6.5}$$

where linear equations describe the local stability and dynamics near u^*. Note that we currently work on an *unbounded* domain. Hence, it is natural to consider the spatial **Fourier transform**

$$\hat{w}(k,t) := \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}kx} w(x,t)\ \mathrm{d}x, \qquad k \in \mathbb{R} \tag{6.6}$$

as the problem is linear. Applying the Fourier transform to (6.5) yields

$$\partial_t \hat{w} = (-\mathrm{i}k)^2 \hat{w} + a^* \hat{w} \quad \Rightarrow \quad \hat{w}(t) = \mathrm{e}^{((-\mathrm{i}k)^2 + a^*)t} \hat{w}(0). \tag{6.7}$$

The spatial Fourier modes decay if $\mathrm{Re}[(-\mathrm{i}k)^2 + a^*] = -k^2 + a^* < 0$ for all $k \in \mathbb{R}$, while a mode $k \in \mathbb{R}$ grows if $-k^2 + a^* > 0$. Hence, we can also check if u^* is stable if all Fourier modes decay to zero.

Example 6.2. Returning to the FKPP equation from Example 6.1, we find

$$\partial_t w = \partial_{xx}^2 w + (D_u f)(u^*) w = \begin{cases} \partial_{xx}^2 w + w, & \text{if } u^* = 0, \\ \partial_{xx}^2 w - w, & \text{if } u^* = 1, \end{cases} \tag{6.8}$$

so $a^* = \pm 1$ in the notation above and we see that $u^* = 0$ is unstable while $u^* = 1$ is stable. The travelling front $u(\xi)$ we want to analyze is a heteroclinic with the following properties (cf. Example 6.1)

$$\begin{aligned} &u(-\infty) = 0,\ u(+\infty) = 1,\ \text{if } s < 0, \text{ front moves 'left'}, \\ &u(-\infty) = 1,\ u(+\infty) = 0,\ \text{if } s > 0, \text{ front moves 'right'}, \end{aligned}$$

see also Figure TODO. In particular, the front always propagates into the unstable state. ♦

In the general case, another way to think about the linearized problem is to substitute an ansatz similar to variation-of-constants idea

$$\hat{w}(k,t) = \hat{w}(k,0)\mathrm{e}^{-\mathrm{i}\omega(k)t},$$

where $\omega = \omega(k)$ will be called the (angular) **frequency** and k the **wave number**, which yields

$$-\mathrm{i}\omega = (-k^2 + a^*) \quad \Leftrightarrow \quad \omega = -\mathrm{i}(k^2 - a^*). \tag{6.9}$$

Remark: $\mathrm{Im}[\omega]$ is the actual 'physical' angular frequency defined as 2π/period. Also, observe that $\hat{w}(k,0)$ is just the Fourier transform of the initial condition $w(x, t = 0)$. Furthermore, there is an implicit sign convention as one may equally well use $-\omega$ instead of ω.

Definition 6.3. A relation between wave numbers (or Fourier modes) k and a frequency ω is called a **dispersion relation**.

In general, there is quite some confusion in the literature, what one should call 'the' dispersion relation. For example, another way to get 'a' dispersion relation is to substitute a single Fourier-mode wave ansatz directly into the linearized problem

$$w_k(x,t) = \mathrm{e}^{\mathrm{i}kx + \sigma t} \quad \Rightarrow \quad \sigma = -k^2 + a^* \tag{6.10}$$

is also usually called a dispersion relation, where $\mathrm{Im}[\omega] = \sigma$. Hence, there is a choice whether we want to look at σ or just its imaginary part.

Remark: The frequency $\omega = \mathrm{Im}[\sigma]$ of the wave is related to the speed. Indeed, if $\omega(k) = ks(k)$ is real then $\mathrm{e}^{\mathrm{i}(kx+\omega t)} = \mathrm{e}^{\mathrm{i}k(x+st)}$ is travelling with 'phase velocity' $s = s(k)$. Since different wave numbers may have different speeds, it is possible that an initial wave form disperses; this explains the name **dispersion relation**. Also, note that

$$\mathrm{e}^{\mathrm{i}k(x+st)} + \text{c.c.} = \mathrm{e}^{\mathrm{i}k(x+st)} + \mathrm{e}^{-\mathrm{i}k(x+st)} = 2\cos(k(x-st))$$

is then the real representation of the wave where c.c. denotes complex conjugate.

The next step is to finally tackle the wave speed problem. Consider the following setup, which is assumed to hold throughout this section:

(P1) $u^* = 0$ is an unstable state of (6.1),

(P2) the dispersion relation $\omega(k)$ is an analytic function when $k \in \mathbb{C}$,

(P3a) $w(\cdot, 0) \in C_c^\infty(\mathbb{R}, \mathbb{R})$ (smooth with compact support),

(P3b) $w(x, 0) > 0$ for some $x \in \mathbb{R}$, $w(x, 0) \geq 0$ for all $x \in \mathbb{R}$,

(P4) $x_\kappa(t)$ is a (level) curve in $\mathbb{R} \times [0, \infty)$ such that $w(x_\kappa(t), t) = \kappa$.

Observe that the time derivative of the level curve x_κ can be used to track the speed at which waves of an initial compact support spread to the left or right into the unstable state as shown in Figure TODO.

Definition 6.4. A **linear spreading speed** $s^* \in (-\infty, +\infty)$ is any value, which can be obtained as a well-defined limit of the form

$$s^* = \lim_{t \to +\infty} \frac{\mathrm{d}x_\kappa}{\mathrm{d}t}(t)$$

and can be calculated by just using the linearized equation at the unstable state.

Usually, we expect to have one linear spreading speed to the right and one to the left if a front propagates into an unstable state; see Figure TODO.

Proposition 6.5. *Assume (P1)-(P4). Then any generic linear spreading speed* $s^* \neq 0$ *satisfies*

$$s^* = \frac{\mathrm{d}\omega}{\mathrm{d}k}(k^*), \tag{6.11}$$

$$s^* = \frac{\mathrm{Im}[\omega(k^*)]}{\mathrm{Im}[k^*]}, \tag{6.12}$$

where $\omega = \omega(k)$ *is the dispersion relation and* $k^* \in \mathbb{C}$ *is a constant.*

It will be made clear in the following proof, what the role of $k^* \in \mathbb{C}$ is.

Proof. Using variation-of-constants ansatz and the inverse Fourier transform, we write the solution of the linearized problem in the original variables as

$$w(x,t) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{w}(k,0)\mathrm{e}^{\mathrm{i}kx-\mathrm{i}\omega(k)t}\ \mathrm{d}k. \tag{6.13}$$

Consider a travelling wave frame $\xi := x - s^*t$ with some linear spreading speed s^*. Then we obtain from (6.13) that

$$\begin{aligned} w(\xi,t) &= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{w}(k,0)\mathrm{e}^{\mathrm{i}k(\xi+s^*t)-\mathrm{i}\omega(k)t}\ \mathrm{d}k \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{w}(k,0)\mathrm{e}^{\mathrm{i}k\xi}\mathrm{e}^{(\mathrm{i}ks^*-\mathrm{i}\omega(k))t}\ \mathrm{d}k \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \underbrace{\hat{w}(k,0)\mathrm{e}^{\mathrm{i}k\xi}}_{=:h(k)}\mathrm{e}^{\rho(k)t}\ \mathrm{d}k, \end{aligned}$$

where $\rho(k) := -\mathrm{i}\omega(k) + \mathrm{i}s^*k$. By (P3a), we know that $\hat{w}(k,0)$ is analytic as a function of $k \in \mathbb{C}$ and since $\mathrm{e}^{ik\xi}$ is analytic it follows that $h(k)$ is analytic. We are interested in a fixed ξ and $t \to +\infty$ limit as we look for linear asymptotic spreading speeds. Then, the main contribution of the integral

$$\frac{1}{2\pi}\int_{\mathbb{R}} h(k)\mathrm{e}^{\rho(k)t}\ \mathrm{d}k, \qquad \text{as } t \to +\infty, \tag{6.14}$$

can be found by deforming the integration contour into the upper- or lower-half in $\mathbb{C}$ as $h(k)$ and $\rho(k)$ are analytic by Cauchy's Theorem [Gam01]. As a fact, we accept that the **method of steepest descents** [BO99] shows that the main contribution can be calculated if the integration contour contains a point k^* where $\rho(k)$ varies least (so-called 'saddle points'; here we call k^* the **linear spreading point**) but this just means

$$\frac{\mathrm{d}\rho}{\mathrm{d}k}(k^*) = 0.$$

By genericity, we know that the minimum is non-degenerate (any slight perturbation does make it non-degenerate. Furthermore, we see that

$$\frac{\mathrm{d}\rho}{\mathrm{d}k}(k^*) = -\mathrm{i}\frac{\mathrm{d}\omega}{\mathrm{d}k}(k^*) + \mathrm{i}s^* \overset{!}{=} 0 \tag{6.15}$$

from which (6.11) follows. For (6.12), we now evaluate the integral and the dominant contribution is the integrand involving t evaluated at the saddle point

$$\frac{1}{2\pi}\int_{\mathbb{R}} h(k)\mathrm{e}^{\rho(k)t}\ \mathrm{d}k = \mathrm{e}^{\rho(k^*)t} + \text{h.o.t}, \qquad \text{as } t \to +\infty.$$

To have a linear spreading speed a consistency requirement is that the leading-order term does neither *grow* or *decay* as we are in the moving frame of the asymptotic linear spreading speed. This implies

$$\mathrm{Re}[\rho(k^*)] = 0 \quad \Rightarrow \mathrm{Re}[-\mathrm{i}\omega(k^*) + \mathrm{i}s^*k^*] = \mathrm{Im}[\omega(k^*) - s^*k^*] = 0$$

from which the second part in (6.12) follows easily. □

Usually, there is more than just one linear spreading speed, e.g., two linear spreading speeds for the left and right front. If we lift the genericity requirement on ρ, i.e., we require less of the structure on the dispersion relation and do not have that the minimum of $\rho(k)$ is non-degenerate, there are a lot more possibilities.

Example 6.6. It is helpful to work out, what Proposition 6.5 says about our FKPP example (6.1)-(6.2). The dispersion relation follows from (6.9) with $a^* = 1$ as (6.8) has to be applied for the unstable state $u^* = 0$

$$\omega(k) = \mathrm{i}(-k^2 + 1) \quad \Rightarrow \quad \frac{\mathrm{d}\omega}{\mathrm{d}k} = -2\mathrm{i}k$$

so $s^* = -2\mathrm{i}k^*$ by (6.11). Since s^* has to be real it follows that k^* is purely imaginary, say $k^* = \mathrm{i}\beta$. Then (6.11) implies

$$2\beta = s^* = \frac{\mathrm{Im}[\omega(k^*_\pm)]}{\mathrm{Im}[k^*_\pm]} = \frac{\beta^2 + 1}{\beta},$$

so $\beta = \pm 1$ and we have

$$k^*_\pm = \pm\mathrm{i}, \qquad s^*_\pm = -2\mathrm{i}k^*_\pm = \pm 2.$$

So if we believe that the propagation into the unstable state of the FKPP equation is governed by the behaviour of the front near the unstable state, then we would expect that the selected wave speed satisfies $|s| = 2$. ♦

Definition 6.7. A travelling front propagating into an unstable state is called a **pulled front** if the wave speed of the full nonlinear system equals the linear spreading speed near the unstable state. Otherwise, the wave is called a **pushed front**.

Giving a rigorous proofs for precise wave speed selection is a non-trivial task and there are many results in the literature. It is a folklore result in applied mathematics that for the FKPP equation, there can only be waves for $|s| \geq 2$, and that all 'sufficiently rapidly decaying' initial conditions converge to waves with minimal speed $|s| = 2$, i.e., the practically stable nonlinear fronts are pushed fronts for the classical FKPP equation.

Exercise 6.8. Prove the following asymptotics of an integral

$$\int_0^1 \frac{\mathrm{e}^{\mathrm{i}kt}}{1+k}\mathrm{d}k = -\frac{\mathrm{i}}{2t}\mathrm{e}^{\mathrm{i}t} + \frac{\mathrm{i}}{t} + \text{h.o.t.} \qquad \text{as } t \to +\infty \tag{6.16}$$

using integration by parts repeatedly. It is frequently very useful in mathematics to be able to estimate **Fourier-type integrals** such as (6.16); for more background see [BO99, Mil06]. $\lozenge$

Exercise 6.9. Consider the FKPP equation with **convection/advection** given by

$$\partial_t u = \partial_{xx}^2 u - \kappa \partial_x u + u(1-u) \tag{6.17}$$

and calculate the dispersion relation and the linear spreading speed. Fact: (6.17) does generate pushed or pulled fronts depending on the parameter $\kappa \in \mathbb{R}$. $\lozenge$

Exercise 6.10. Consider the **complex Ginzburg-Landau equation (cGL)**

$$\partial_t u = (1+\mathrm{i}c_1)\partial_{xx}^2 u + c_2 u - (1-\mathrm{i}c_3)|u|^2 u, \qquad u = u(x,t) \in \mathbb{C} \tag{6.18}$$

where $c_j \in \mathbb{R}$, $j \in \{1,2,3\}$. Calculate the dispersion relation for the steady state $u^* = 0$, show that u^* is unstable, and calculate the linear spreading speed and the spreading point depending upon the c_j's. $\lozenge$

Background and Further Reading: The material in this section is based upon [vS03]. Detailed proofs for the full nonlinear case of many equations using sub- and super-solutions can be found in [VVV94]. The FKPP equation is a cornerstone example in mathematical biology [Mur02].

7 Sturm-Liouville and Stability of Travelling Waves

Having established some basic techniques and properties regarding existence and speed of travelling waves, it is natural to also look at **stability** of waves.

Example 7.1. For $u = u(x,t)$, $(x,t) \in \mathbb{R} \times [0, +\infty)$, consider the reaction-diffusion PDE

$$\partial_t u = \partial_{xx}^2 u - u + 2u^3. \tag{7.1}$$

Then it can actually be checked that there is a **standing** (speed $s = 0$) pulse solution

$$u(x,t) = u(x - st) = u(x - 0 \cdot t) = U(\xi) = \operatorname{sech}(\xi), \tag{7.2}$$

i.e., $U(\xi)$ solves the stationary problem for (7.1) with x replaced by ξ. Following the usual local linearization paradigm, we consider $u(x,t) = U(\xi) + \epsilon w(\xi, t)$ and obtain

$$\partial_t w = \partial_{\xi\xi}^2 w + (6U(\xi)^2 - 1)w =: Lw. \tag{7.3}$$

Hence, we might expect that the eigenvalue problem

$$Lw = \lambda w, \qquad \lambda \in \mathbb{C},\ w \in X,\ w = w(\xi), \tag{7.4}$$

where X is a suitable Banach space, can help us to determine the stability of the pulse. In fact, the example shows that we should study linear operators L, which are defined on an unbounded spatial domain and which depend in general upon the travelling waves coordinate ξ as we have linearized around the wave profile. ♦

On bounded domains, it was actually sufficient to just look at the eigenvalues. On unbounded domains, we have to consider a more general setup. Let $L : X \to Y$ be a linear operator between two Banach spaces X, Y.

Definition 7.2. If $\mathcal{N}[L - \lambda \mathrm{Id}] \neq 0$ then $\lambda \in \mathbb{C}$ is called an **eigenvalue**.

The following definition is only relevant for the infinite-dimensional operator case; see also Figure TODO.

Definition 7.3. The **essential spectrum** $\sigma_{\text{ess}}(L)$ consists of those $\lambda \in \mathbb{C}$ such that $L - \lambda \mathrm{Id}$ is not a Fredholm operator of index zero. The point spectrum is the complement, i.e., $\sigma_{\text{pt}}(L) := \sigma(L) - \sigma_{\text{ess}}(L)$ where $\sigma(L)$ denotes the spectrum of L as defined before.

As a key classical example consider **Sturm-Liouville operators** (7.5)

$$Lw = \partial^2_{\xi\xi}w + a_1(\xi)\partial_\xi w + a_0(\xi)w, \qquad w = w(\xi),\ \xi \in \mathbb{R}, \tag{7.5}$$

where $a_{0,1}(\xi)$ are smooth coefficients decaying exponentially at infinity to asymptotic values

$$\lim_{\xi\to\pm\infty} \mathrm{e}^{\nu|\xi|}|a_1(\xi) - a_1^\pm| = 0, \quad \lim_{\xi\to\pm\infty} \mathrm{e}^{\nu|\xi|}|a_0(\xi) - a_0^\pm| = 0, \tag{7.6}$$

for some $\nu > 0$ and $a^\pm_{0,1} \in \mathbb{R}$. If we would have $\xi \in \Omega$ for some bounded domain/interval Ω, then the results in Section 4 imply that L has only point spectrum. However, on an unbounded domain, Sturm-Liouville operators may have essential spectrum. For now, we shall not worry about this problem and look at the point spectrum on unbounded domains.

Consider the Sturm-Liouville operator (7.5) as a mapping $L : H^2_\mu(\mathbb{R}) \to L^2_\mu(\mathbb{R})$, the subscript μ indicates that we endow the usual H^2- and L^2-spaces with the **weighted inner product**

$$\langle v, w\rangle_\mu := \int_{\mathbb{R}} v(x)\overline{w(x)}\mu(x)\ \mathrm{d}x, \qquad \mu(x) := \mathrm{e}^{\int_0^x a_1(y)\ \mathrm{d}y}. \tag{7.7}$$

Then classical **Sturm-Liouville theory** provides the following result.

Theorem 7.4. *(see [KP13]) Consider the eigenvalue problem $Lw = \lambda w$ for the Sturm-Liouville operator L defined in* (7.5) *with the condition* (7.6)*. The point spectrum $\sigma_{pt}(L)$ is given by a finite number of eigenvalues such that*

$$\lambda_0 > \lambda_1 > \cdots > \lambda_N, \qquad \lambda_j \in \mathbb{R} \textit{ for all } j \in \{0, 1, 2, \ldots, N\}.$$

Furthermore, the eigenfunction e_j for λ_j has j simple zeros, the eigenfunctions are orthonormal in the inner product (7.7)*, and we have the formula*

$$\lambda_0 = \sup_{\|w\|_\mu=1} \langle Lw, w\rangle_\mu \tag{7.8}$$

where the supremum is achieved at $w = e_0$.

Frequently, one also refers to e_0 associated to (7.8) as the **ground state**.

Example 7.5. We continue with Example 7.1. Observe that (7.4) is an eigenvalue problem for a Sturm-Liouville operator (7.5) with

$$a_1(\xi) = 0, \qquad a_0(\xi) = 6(U(\xi)^2 - 1) = 6(\mathrm{sech}^2(\xi) - 1).$$

The condition (7.6) holds with $a_0^{\pm} = -6$ as $\mathrm{sech}(\xi) = 2/(\mathrm{e}^{-\xi} + \mathrm{e}^{\xi})$ decays exponentially at $\xi = \pm\infty$ and easily with $a_1^{\pm} = 0$. Next, differentating the steady state equation

$$0 = \partial_{\xi\xi}^2 U - U + 2U^3$$

it follows that

$$0 = \partial_{\xi\xi}^2(\partial_\xi U) - \partial_\xi U + 6U^2\partial_\xi U = L(\partial_\xi U).$$

So we have $L(\partial_\xi U) = 0 \cdot (\partial_\xi U)$. Since

$$(\partial_\xi U)(\xi) = -\mathrm{sech}(\xi)\,\tanh(\xi) \qquad \text{recall: } \tanh(\xi) = \frac{\mathrm{e}^{\xi} - \mathrm{e}^{-\xi}}{\mathrm{e}^{\xi} + \mathrm{e}^{-\xi}}$$

it is easy to see that $(\partial_\xi U)$ decays exponentially at infinity. We conclude that $\lambda = 0$ is an eigenvalue of L in $H_\mu^2(\mathbb{R})$. Furthermore, $(\partial_\xi U)(\xi)$ has precisely one zero at $\xi = 0$ ('the top of the pulse', see Figure TODO). So we must have $\lambda_1 = 0$ in the notation of Theorem 7.4. Since Theorem 7.4 also yields the existence of a ground state eigenvalue $\lambda_0 > \lambda_1 = 0$ we see that the point spectrum contains a positive eigenvalue, which implies that the standing pulse solution of (7.1) is unstable. ♦

The goal is to prove a more general theorem about stability for equations of the form

$$\partial_t u = \partial_{xx}^2 u - \mathcal{F}'(u), \qquad u = u(x,t),\ x \in \mathbb{R},\ \mathcal{F} : \mathbb{R} \to \mathbb{R}, \tag{7.9}$$

and $\mathcal{F}'$ denotes the derivative of the smooth **potential** $\mathcal{F}$. Assume that $-\mathcal{F}$ has three critical points at $u = 0, p, 1$ for $p \in (0, \frac{1}{2})$ and $u = 0, 1$ are minima, i.e.,

$$-\mathcal{F}''(0) < 0, \qquad -\mathcal{F}''(1) < 0$$

just means that $u = 0, 1$ are stable stationary states and we should think of the classical **bistable** Nagumo equation with $\mathcal{F}'(u) = -u(1-u)(u-p)$ as an example. We can apply the phase-plane and Lin's-type methods from Section 5 to investigate the existence of travelling waves and summarize the results here:

Theorem 7.6. *Consider the one-dimensional bistable reaction-diffusion equation* (7.9) *in one space dimension. Then there exists*

- *a standing pulse solution $U(\xi) \geq 0$ with $U(\pm\infty) = 0$,*
- *a travelling front solution $U(\xi) \geq 0$ with $U(-\infty) = 0$, $U(+\infty) = 1$, $U'(\xi) > 0$, for a unique wave speed $s = s^* > 0$.*

Both waves $U(\xi)$ and their derivatives $(\partial_\xi U)(\xi)$ decay exponentially to their asymptotic limits as $\xi \to \pm\infty$.

Proof. The result follows using the methods in Section 5. Indeed, we use Lin's method to find the unique heteroclinic orbit that yields the positive monotone travelling front. We can adapt the Hamiltonian argument from Example 5.5 to get the homoclinic orbit representing the standing pulse. The exponential decay at the endstates just follows from the local linearization at the hyperbolic saddle points of the travelling wave ODEs in coordinates $(u_1, u_2) = (u, \partial_\xi u) \in \mathbb{R}^2$ associated to (7.9); see also Figure TODO. $\square$

Theorem 7.7. *Consider the same setup as in Theorem 7.6. Denote the point spectra for the standing pulse and the travelling front by σ_{pt}^{pulse} and σ_{pt}^{front} respectively, and let $z = a + \mathrm{i}b \in \mathbb{C}$, then*

$$\sigma_{pt}^{pulse} \cap \{a > 0, b = 0\} \neq \{\} \tag{7.10}$$

$$\sigma_{pt}^{front} \subset \{a^* < a < 0, b = 0\} \cup \{0\}, \tag{7.11}$$

for some constant $a^* < 0$ when the eigenvalue problem is considered on $H^2_\mu(\mathbb{R})$.

Proof. We start with the standing pulse denoted by $U(\xi)$, which satisfies

$$\partial^2_{\xi\xi} U = \mathcal{F}'(U). \tag{7.12}$$

The relevant eigenvalue problem is

$$Lw = \lambda w, \quad L = \partial^2_{\xi\xi} - \mathcal{F}''(U),$$

so L is a Sturm-Liouville operator with $a_0(\xi) = -\mathcal{F}''(U(\xi))$ and $a_1(\xi) = 0$. As in Example 7.5, we differentiate (7.12) and obtain

$$0 = \partial^2_{\xi\xi}(\partial_\xi U) - \mathcal{F}''(U)(\partial_\xi U) = L(\partial_\xi U) \tag{7.13}$$

so $(\partial_\xi U)(\xi)$ is an eigenfunction with eigenvalue $\lambda = 0$. Since $(\partial_\xi U)(\xi)$ vanishes precisely once for $\xi \in \mathbb{R}$ by its phase-plane construction, we conclude from Theorem 7.4 that $\lambda_1 = 0$ and the ground state has a positive eigenvalue $\lambda_0 > 0$, which proves (7.10).

For the second part, we use again the notation $U(\xi)$ for the travelling front, which solves

$$0 = \partial^2_{\xi\xi} U + s^* \partial_\xi U - \mathcal{F}'(U). \tag{7.14}$$

The relevant eigenvalue problem is

$$Lw = \lambda w, \quad L = \partial^2_{\xi\xi} + s^* \partial_\xi - \mathcal{F}''(U),$$

so L is a Sturm-Liouville operator with $a_0(\xi) = -\mathcal{F}''(U(\xi))$ and $a_1(\xi) = s^*$. As before, direct differentiation of (7.14) shows that $\partial_\xi U$ is an eigenfunction with eigenvalue $\lambda = 0$. However, since $U(\xi)$ is monotone (see Theorem (7.6)) it follows that $\partial_\xi U$ is the ground state so $\lambda_0 = 0$ and Theorem 7.4 implies that all eigenvalues must be real and negative. The lower bound in (7.11) follows also from Theorem 7.4 as the number of eigenvalues is finite. $\square$

Remark: The eigenvalue $\lambda = 0$ always occurs due to **translation invariance**, i.e., if $U(\xi)$ is a travelling wave so is $U(\xi + \xi_0)$ for any $\xi_0 \in \mathbb{R}$; see also Section 8. More generally, zero eigenvalues may arise due to other symmetries of the problem as well.

Basically, Theorem 7.7 is a first step towards the folklore result that in one-dimensional reaction diffusion equations in one space dimension, monotone waves are stable while non-monotone waves are unstable. However, this result fails miserably when we increase the dimension.

Example 7.8. Consider a version of the FitzHugh-Nagumo equation

$$\begin{aligned} \partial_t u &= \partial_{xx}^2 u + u(1-u)(u-0.1) - v, \\ \partial_t v &= \epsilon(u - v), \end{aligned} \tag{7.15}$$

for $x \in \mathbb{R}$. It is a well-known fact that for $0 < \epsilon \ll 1$ sufficiently small there exists a travelling pulse solution for (7.15), which is stable. However, the proof of this result is difficult; see references.

The last example shows that we cannot, despite the elegance, rely on studying just one-dimensional cases based upon Sturm-Liouville theory.

Exercise 7.9. Prove that the Sturm-Liouville operator L with smooth coefficient functions (7.5) is **self-adjoint** in H^2_μ, i.e., $\langle Lu, v\rangle_\mu = \langle u, L^* v\rangle_\mu$ with **adjoint operator** $L^* = L$. $\diamondsuit$

Exercise 7.10. Consider (7.9) and define the **energy** as $E(u) := \frac{1}{2}(\partial_x u)^2 - \mathcal{F}(u)$. What is the relation between the energy and the Hamiltonian function of Example 5.5? $\diamondsuit$

Exercise 7.11. Show that studying the point spectrum of travelling fronts in FKPP equation (6.1)-(6.2) for compactly supported initial perturbations leads to an eigenvalue problem with Dirichlet boundary conditions. $\diamondsuit$

Background and Further Reading: This section was mainly based upon the excellent book [KP13]. Example 7.1 can be found in [Kap05]. An important survey of travelling wave stability results is [San01]. For classical Sturm-Liouville theory a good source is [Wei87]. For the existence proof of the FitzHugh-Nagumo pulse see [JKL01] and for stability [Jon84].

8 Exponential Dichotomies and the Evans Function

For the results in Section 7 we did rely crucially on classical Sturm-Liouville theory to understand the spectrum. In this section, we develop the basics of a more general theory for equations of the form

$$\partial_t u = \mathcal{A}(\partial_x)u + f(u), \qquad u = u(x,t),\ x \in \mathbb{R} \tag{8.1}$$

where $\mathcal{A}(\cdot)$ is a polynomial of its argument and $\mathcal{A}(\partial_x) : X \to X$ is a linear operator on a suitable Banach space X and f is a mapping representing the nonlinear terms.

Example 8.1. An excellent example to keep in mind for (8.1) are reaction-diffusion systems

$$\partial_t u = D\Delta u + f(u), \qquad u = u(x,t) \in \mathbb{R}^N, \tag{8.2}$$

where D is a diagonal matrix with positive entries. ♦

Let $U(\xi) = u(x - st)$ be a travelling wave solution for (8.1). Then the first step to investigate (linear) stability of the wave is to consider the operator

$$L := \mathcal{A}(\partial_\xi) + s\partial_\xi + [(\mathrm{D}_u f)(U(\xi))] \tag{8.3}$$

and study the eigenvalue problem $Lw = \lambda w$.

Example 8.2. For (8.2) we find that

$$Lw = D\partial^2_{\xi\xi}w + s\partial_\xi w + (\mathrm{D}_u f)(U(\xi))w.$$

The eigenvalue problem $Lw = \lambda w$ can be written as a first-order system using $\frac{\mathrm{d}w}{\mathrm{d}\xi} = \tilde{w}$

$$\begin{pmatrix} w' \\ \tilde{w}' \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{Id} \\ D^{-1}[\lambda - (\mathrm{D}_u f)(U(\xi))] & -sD^{-1} \end{pmatrix} \begin{pmatrix} w \\ \tilde{w} \end{pmatrix}. \tag{8.4}$$

where $' = \frac{\mathrm{d}}{\mathrm{d}\xi}$. Note that the eigenvalue problem for $v := (w, \tilde{w}) \in \mathbb{R}^{2N}$ has the structure

$$v' = A(\xi;\lambda)v = (\tilde{A}(\xi) + B(\lambda))v \tag{8.5}$$

for matrix-valued functions $\tilde{A}$ and B, which follow directly from (8.4). ♦

Since we work in one spatial dimension, we can always reduce to an operator L which is a first-order differential operator acting on a space of more function components as illustrated in Example 8.2. We still use L

to denote this operator, i.e., the re-writing of (8.3) to a first-order ODE system is understood. This spectrum of L can be studied by considering

$$Lw = \lambda w \qquad \leftrightarrow \qquad (L - \lambda \ \mathrm{Id})w = 0,$$

which is equivelent to a first-order ODE system of the form

$$\frac{\mathrm{d}v}{\mathrm{d}\xi} = A(\xi;\lambda)v \tag{8.6}$$

for some matrix-valued function $A(\xi;\lambda)$ and for some $v \in \mathbb{R}^n$; it helps again to think of Example 8.2 to illustrate the general theory. Hence, we have to study the family of linear operators

$$\mathcal{L}(\lambda)v := \frac{\mathrm{d}v}{\mathrm{d}\xi} - A(\cdot;\lambda)v.$$

We shall usually assume that we work in the spaces

$$X = C^1_{\mathrm{unif}}(\mathbb{R},\mathbb{C}^n), \quad Y = C^0_{\mathrm{unif}}(\mathbb{R},\mathbb{C}^n), \quad \mathcal{L} = \mathcal{L}(\lambda) : X \to Y$$

so that $\mathcal{L}$ is a closed and densely-defined operator. We assume that the matrix-valued function $A \in \mathbb{C}^{n\times n}$ decomposes as

$$A(\xi;\lambda) = \tilde{A}(\xi) + B(\lambda) \tag{8.7}$$

where $\tilde{A}, B \in \mathbb{R}^{n\times n}$ are smooth; frequently one just has $B(\lambda) = \lambda B$ for a constant matrix B. The spectral properties of $\mathcal{L}$ can obviously be studied by looking at the ODEs

$$v' = A(\xi;\lambda)v, \tag{8.8}$$

which is a **non-autonomous** linear system as A depends upon the 'time' variable ξ. In particular, we have to look for those λ, where $\mathcal{L}$ is not invertible.

Definition 8.3. Denote by $\phi = \phi(\xi,\zeta)$ the **fundamental solution** (or **propagator**) for the system (8.8), i.e.,

$$v(\xi) = \phi(\xi,0)v(0)$$

solves (8.8), $\phi(\xi,\xi) = \mathrm{Id}$ for all $\xi \in \mathbb{R}^n$, and $\phi(\xi,\chi)\phi(\chi,\zeta) = \phi(\xi,\zeta)$ holds for all $\xi,\chi,\zeta \in \mathbb{R}$.

Example 8.4. If the linear system would be autonomous

$$v' = A(\lambda)v, \qquad v \in \mathbb{C}^n, \tag{8.9}$$

then we could easily solve (8.9) using the matrix exponential and solutions can be classified depending on their behaviour near the steady-state $v \equiv 0$. For example, if the steady state is hyperbolic for some λ then there is a splitting

$$\mathbb{C}^n = E^{\mathrm{s}}(0;\lambda) \oplus E^{\mathrm{u}}(0;\lambda) = \mathcal{R}[P_0^s(\lambda)] \oplus \mathcal{N}[P_0^s(\lambda)]$$

where P_0^s is the **spectral projection** onto the stable eigenspace $E^{\mathrm{s}}(0;\lambda)$ for $v \equiv 0$. Note also that the stable and unstable subspaces are invariant under the propagator $\phi(\xi,\zeta) = \mathrm{e}^{(\xi-\zeta)A(\lambda)}$ and solutions decay in forward time in the stable space and in backward time in the unstable space. ♦

The right concept to generalize the splitting from the last example is given in the next definition.

Definition 8.5. Let $I = \mathbb{R}^+, \mathbb{R}^-$ or $\mathbb{R}$ and fix $\lambda \in \mathbb{C}$. The ODE (8.8) has an **exponential dichotomy** on I if there exist constants $K > 0$, $\kappa_{\mathrm{s}} < 0 < \kappa_{\mathrm{u}}$ and a continuous family of projectors $P(\xi)$ for $\xi \in I$ such that for $\xi, \zeta \in I$ the following holds:

- $\|\phi(\xi,\zeta)P(\zeta)\| \leq K\ \mathrm{e}^{\kappa_s(\xi-\zeta)}$ for $\xi \geq \zeta$,
- $\|\phi(\xi,\zeta)[\mathrm{Id} - P(\zeta)]\| \leq K\ \mathrm{e}^{\kappa_u(\xi-\zeta)}$ for $\xi \leq \zeta$,
- projections commute with evolution, i.e., $\phi(\xi,\zeta)P(\zeta) = P(\xi)\phi(\xi,\zeta)$.

The ξ-independent dimension $\mathcal{N}[P(\xi)]$ is also called the **Morse index**. If the exponential dichotomies hold on $\mathbb{R}^+$ and $\mathbb{R}^-$, the associated Morse indices are denoted by $m_+(\lambda)$ and $m_-(\lambda)$; see also Figure TODO.

Remark: It can be shown that $\mathcal{L}$ is a Fredholm operator if and only if it has exponential dichotomies on $\mathbb{R}^+$ and $\mathbb{R}^-$. Furthermore, then $m_-(\lambda) - m_+(\lambda)$ is equal to the Fredholm index.

Theorem 8.6. ***(Palmer's Theorem*** *[Pal84, Pal88]) The following characterization of spectrum and resolvent set of L holds:*

(P1) $\lambda \notin \sigma(L)$ *if and only if* (8.6) *has an exponential dichotomy on* $\mathbb{R}$*,*

(P2) $\lambda \in \sigma_{pt}(L)$ *if and only if* (8.6) *has an exponential dichotomy on* $\mathbb{R}^+$ *and* $\mathbb{R}^-$ *with the same Morse index and* $\dim(\mathcal{N}[\mathcal{L}(\lambda)]) > 0$*,*

(P3) If (P2) holds then $\mathcal{N}[P_-(0;\lambda)] \cap \mathcal{R}[P_+(0;\lambda)] \cong \mathcal{N}[\mathcal{L}(\lambda)]$*, where* $P_\pm(\xi;\lambda)$ *denote the projections for the exponential dichotomies on* $\mathbb{R}^\pm$*,*

(P4) $\lambda \in \sigma_{ess}(L)$ *if we are not in the situation (P1) or (P2).*

Suppose we are interested in applying Palmer's Theorem to the case of a travelling front $U(\xi)$ with

$$\lim_{\xi \to \pm\infty} U(\xi) = U^*_\pm,$$

with homogeneous endstates $U^*_\pm \in \mathbb{R}^N$. Consider (8.6) with $A(\xi;\lambda)$ and define

$$A_\pm(\lambda) := \lim_{\xi \to \pm\infty} A(\xi;\lambda). \tag{8.10}$$

It turns out that the asymptotic matrices $A_\pm(\lambda)$ can be used to characterize the exponential dichotomies and hence, by Palmer's Theorem, also the spectrum.

Theorem 8.7. *([Cop78, San01]) Fix $\lambda \in \mathbb{C}$, consider a travelling front then the following results hold:*

(F1) The ODE (8.6) *has an exponential dichotomy on $\mathbb{R}^+$ if and only if the matrix A_+ is hyperbolic. In this case $m_+(\lambda) = \dim E^{\mathrm{u}}_+(\lambda)$, where $E^{\mathrm{u}}_+(\lambda)$ is the unstable eigenspace of $A_+(\lambda)$.*

(F2) (F1) holds with $\mathbb{R}^+$ replaced by $\mathbb{R}^-$ and $A_+(\lambda)$ replaced by $A_-(\lambda)$.

(F3) The ODE (8.6) *has an exponential dichotomy on $\mathbb{R}$ if and only if it has exponential dichotomies on $\mathbb{R}^+$ and $\mathbb{R}^-$ with projections $P_\pm(\xi;\lambda)$ such that $\mathcal{N}[P_-(0;\lambda)] \oplus \mathcal{R}[P_+(0;\lambda)] = \mathbb{C}^n$.*

Example 8.8. Here we use the previous results to gain more insight on stability for reaction-diffusion systems. Consider the bistable Nagumo equation

$$\partial_t u = \partial^2_{xx} u + u(1-u)(u-p), \qquad p \in \left(0, \frac{1}{2}\right), \; x \in \mathbb{R}, \; u = u(x,t), \tag{8.11}$$

with a right-moving travelling front solution $U(\xi)$, $\xi = x - st$ with $s > 0$, as discussed in Section 5 with endstates $U^*_- \equiv 0$ and $U^*_+ \equiv 1$. A direct calculation using the formulas (8.4)-(8.5) yields

$$A(\xi;\lambda) = \begin{pmatrix} 0 & 1 \\ \lambda - f'(U(\xi)) & -s \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \lambda + 3U(\xi)^2 - U(\xi)(2+2p) + p & -s \end{pmatrix}$$

so that the asymptotic matrices are

$$A_-(\lambda) = \begin{pmatrix} 0 & 1 \\ \lambda + p & -s \end{pmatrix}, \quad A_+(\lambda) = \begin{pmatrix} 0 & 1 \\ \lambda + 1 - p & -s \end{pmatrix}.$$

Suppose we are interested in determining the essential spectrum and which points lie outside of the spectrum, i.e., we look at point spectrum later. By Palmer's Theorem 8.6 and Theorem 8.7, we should check when the matrices $A_\pm(\lambda)$ are hyperbolic. It is helpful to compute trace and determinant

$$\det(A_\pm(\lambda)) = -\lambda \pm p - U^*_\pm, \qquad \operatorname{tr}(A_\pm(\lambda)) = -s < 0.$$

The trace is real and negative and since it is the sum of the eigenvalues $\mu_1^\pm, \mu_2^\pm$ of $A_\pm(\lambda)$ it follows that if complex conjugate eigenvalues $\mu_1 = \overline{\mu_2}$ occured then those have negative real part so those eigenvalues must be contained in the left-half plane. Hence, we focus on real eigenvalues $\mu_1^\pm, \mu_2^\pm \in \mathbb{R}$, which implies $\lambda \in \mathbb{R}$ as $\det(A_\pm(\lambda)) = \mu_1^\pm \mu_2^\pm$. By using the trace-determinant analysis from Exercise 1.15 it follows that hyperbolicity with real eigenvalues fails only when the determinant vanishes, which happens when

$$-\lambda \pm p - U^*_\pm = 0, \quad \leftrightarrow \quad -\lambda - p = 0 \text{ or } -\lambda + p - 1 = 0$$

and since $p \in (0, 1/2)$ by assumption, this can only happen when $\lambda < 0$. This implies that for fixed p we have

$$\sigma_{\text{ess}}(L) \subset \{\lambda \in \mathbb{C} : \operatorname{Re}(\lambda) \leq \lambda_b < 0\}.$$

for some fixed negative λ_b. Therefore, linear instability can only arise via the point spectrum. ♦

The last example shows that we would like to have a good tool to locate the point spectrum, as it is usually quite easy to show that the essential spectrum is contained in the left-half plane. Observe that Palmer's Theorem (P3) implies that one possibility is to look at the intersection

$$\mathcal{N}[P_-(0;\lambda)] \cap \mathcal{R}[P_+(0;\lambda)] \tag{8.12}$$

and check when it is non-empty to characterize the point spectrum. This means looking at bounded solutions as those lying in the intersection (8.12). Indeed, $\mathcal{R}[P_+(0;\lambda)]$ consists of all initial conditions $v(0)$ with solutions $v(\xi)$ for (8.8), which decay exponentially as $\xi \to +\infty$, while $\mathcal{N}[P_-(0;\lambda)]$ consists of all initial conditions $v(0)$ with solutions $v(\xi)$, which decay exponentially as $\xi \to -\infty$; see Figure TODO and also compare this to the weighted space in the Sturm-Liouville problems in Section 7.

Suppose the essential spectrum is in the left-half plane and let Ω denote the connected component of $\mathbb{C} - \sigma_{\text{ess}}$ intersecting the right-half plane. One very nice device to study the intersection (8.12) for $\lambda \in \Omega$ is to consider bases

$$\begin{array}{ll} v_1(\lambda), \ldots, v_k(\lambda) & \text{of } \mathcal{N}[P_-(0;\lambda)], \\ v_{k+1}(\lambda), \ldots, v_n(\lambda) & \text{of } \mathcal{R}[P_+(0;\lambda)], \end{array}$$

where $k = \dim(\mathcal{N}[P_-(0;\lambda)])$ is the Morse index; it can be shown that the Morse index does not change in Ω and the bases can be chosen to depend analytically [Kat80] upon λ.

Definition 8.9. The **Evans function** is defined as

$$E(\lambda) := \det[v_1(\lambda), \ldots, v_n(\lambda)]. \tag{8.13}$$

Theorem 8.10. *([AGJ90, Eva72]) The Evans function $E(\lambda)$ is analytic for $\lambda \in \Omega$ and $E(\lambda) = 0$ if and only if λ is an eigenvalue of L.*

Proof. If λ is an eigenvalue, then (8.12) is non-empty. Therefore, the bases are linearly dependent and the determinant (8.13) vanishes. The converse is equally simple. Analyticity of the Evans function follows from the analyticity of the bases. □

In general, it is difficult to compute the Evans function explicitly, except for certain special cases and integrable/conservative systems. However, it provides a useful abstract theoretical as we as practical numerical tool.

Example 8.11. To see the difficulty in the computation of the Evans function for reaction-diffusion systems consider the case from Example 7.1 given by

$$\partial_t u = \partial_{xx}^2 u - u + 2u^3. \tag{8.14}$$

with standing pulse solution $U(\xi) = \mathrm{sech}(\xi)$. Then we must eventually understand the exponential dichotomy properties of the linear system

$$v' = A(\xi;\lambda)v = \begin{pmatrix} 0 & 1 \\ 1 + \lambda - 6U(\xi)^2 & 0 \end{pmatrix}$$

which can be turned into an autonomous system via setting $\frac{\mathrm{d}\xi}{\mathrm{d}\tau} = 1$ using the new (dummy) time variable; note $\xi(\tau) = \tau$ if we assume $\xi(0) = 0$. However, in the autonomous case we still have the nonlinear term $U(\xi) = \mathrm{sech}(\xi)$ in the problem. Hence, we are still left with understanding a nonlinear ODE which is a highly non-trivial task. ♦

Example 8.12. The previous example can be generalized. For an N-component reaction-diffusion system (8.2) with a nonlinearity $f : \mathbb{R}^N \to$

$\mathbb{R}^N$, we are facing the problem

$$\begin{pmatrix} \dot{u} \\ \dot{\tilde{u}} \end{pmatrix} = \begin{pmatrix} \tilde{u} \\ D^{-1}(-s\tilde{u} - f(u)) \end{pmatrix} \tag{8.15}$$

$$\begin{pmatrix} \dot{w} \\ \dot{\tilde{w}} \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{Id} \\ D^{-1}[\lambda - (\mathrm{D}_u f)(u)] & -sD^{-1} \end{pmatrix} \begin{pmatrix} w \\ \tilde{w} \end{pmatrix}, \tag{8.16}$$

$$\dot{\xi} = 1, \tag{8.17}$$

which is an autonomous, fully-coupled, $(4N+1)$-dimensional, nonlinear system of ODEs. The $2N$ equations (8.15) come from the existence problem in the travelling wave frame, the $2N$ equations (8.16) from the eigenvalue problem to determine stability and $\dot{\xi} = \frac{\mathrm{d}\xi}{\mathrm{d}\tau}$ makes the system autonomous but represents another difficulty. Indeed, the ξ equation shows directly that there are no steady states so we are always dealing, in some sense, with trying to understand transients. ♦

However, for some systems with a special structure, one may actually find the Evans function explicitly. Here we just quote a result to illustrate this fact.

Example 8.13. Consider the following coupled hierarchy of **Nonlinear Schrödinger Equations (NLS)** given by

$$\begin{aligned} \partial_t u &= \mathrm{i}\left(\partial_{xx}^2 u - 2u - 2u^2 v\right), \\ \partial_t v &= \mathrm{i}\left(\partial_{xx}^2 v - 2v - 2v^2 u\right), \end{aligned} \tag{8.18}$$

where $u = u(x,t), v = v(x,t) \in \mathbb{C}$. Then it turns out that a highly non-trivial explicit calculation yields for a so-called 'stationary soliton solution' the Evans function

$$E(\lambda) = 8a(\lambda)^2 b(\lambda)^2 \sqrt{\lambda - \mathrm{i}}\sqrt{\lambda + \mathrm{i}}$$

where there are explicit formulas for a, b given by

$$a(\lambda) = \frac{\mathrm{e}^{\mathrm{i}\pi/4}\sqrt{\lambda - \mathrm{i}} - 1}{\mathrm{e}^{\mathrm{i}\pi/4}\sqrt{\lambda - \mathrm{i}} + 1}, \quad b(\lambda) = \frac{\mathrm{e}^{-\mathrm{i}\pi/4}\sqrt{\lambda + \mathrm{i}} - 1}{\mathrm{e}^{-\mathrm{i}\pi/4}\sqrt{\lambda + \mathrm{i}} + 1}.$$

Now one has an explicit formula but then a new problem arises as $E(\pm\mathrm{i}) = 0$ so one is always forced to deal with spectrum on the imaginary axis, which - a priori - neither indicates stability or instability. ♦

However, we stress that a fundamental step has been discussed here, as we have converted a full PDE problem for waves in one-dimensional spatial domains into systems of ODEs, which elegantly links dynamics of ODEs to existence and stability of travelling patterns of spatially-extended systems posed on the spatial domain $\mathbb{R}$.

Exercise 8.14. Consider the reaction-diffusion system (8.2) and suppose it has a travelling pulse or a travelling front solution $U(\xi)$. Show that $\lambda = 0$ is always in the spectrum with associated eigenfunction $U'(\xi)$. $\diamondsuit$

Exercise 8.15. Prove that an autonomous linear system having a hyperbolic equilibrium point at the origin always satisfies an exponential dichotomy. What are the values of the Morse index for an autonomous linear system? $\diamondsuit$

Exercise 8.16. Use Palmer's Theorem to characterize the spectrum for homogeneous rest states $U(\xi) \equiv U^* \in \mathbb{R}^N$ of (8.1). In particular, prove that the point spectrum of U^* is empty. This illustrates a fundamental difference between spectra for problems on bounded domains in contrast to unbounded domains. $\diamondsuit$

Background and Further Reading: The exposition here summarizes several key aspects from the review [San01] with the last example from [Kap05]; consider also [SS04]. A key paper developing, and naming, the Evans function is [AGJ90]. The original work of Evans [Eva72] was actually focused on trying to understand the stability of waves in FitzHugh-Nagumo type reaction-diffusion systems and associated nerve impulse equations.

9 Onset of Patterns and Multiple Scales

In this section, we are going to start with a concrete model problem to motivate the development of amplitude equations on a formal level. Consider the **Swift-Hohenberg equation**

$$\partial_t u = [p - (\Delta + 1)^2]u - u^3, \quad u = u(x,y,t) \in \mathbb{R}, \quad (x,y) \in \mathbb{R}^2, \tag{9.1}$$

where $p \in \mathbb{R}$ is a parameter and Δ is the Laplacian; note that we deviate from our convention to use x as the spatial variable as using (x,y) as coordinates is going to simplify the notation later on. The Swift-Hohenberg equation is an idealized model of convective instabilities in fluid dynamics. Note carefully that the domain is $\Omega = \mathbb{R}^2$. Therefore, we cannot expect to capture the dynamical effects for bifurcations of steady states using the point spectrum as we did in Sections 3-4 and the essential spectrum plays a crucial role (cf. Exercise 8.16). Suppose we are interested in the stability of $u \equiv 0$ and linearize (9.1) around $u \equiv 0$, then we obtain

$$\partial_t w = pw - (\Delta + 1)^2 w, \qquad w = w(x,y,t). \tag{9.2}$$

Substituting the dynamics of an individual Fourier mode

$$w_k(x,y,t) = \mathrm{e}^{\sigma t + \mathrm{i}k\cdot(x,y)^\top} + \text{c.c.}, \qquad k = (k_x, k_y)^\top \in \mathbb{R}^2,$$

where k is a **wavevector**, into (9.2) yields the following dispersion relation (check it!)

$$\sigma = p - (k^2 - 1)^2, \qquad k^2 := \|k\|^2. \tag{9.3}$$

If $p < 0$ then $\sigma < 0$, so all modes decay. When $p = 0$, then the critical modes, which are no longer linearly stable occur for wave numbers with $\|k\| = 1$. When $p > 0$ a whole band of wavenumbers is linearly unstable and could occur in a **pattern** bifurcating from the homogeneous solution; see Figure TODO. Let us assume we only look for a simple 'stripe' pattern at the bifurcation $p = 0$ of the form

$$u(x,y) = A\mathrm{e}^{\mathrm{i}x} + \overline{A}\mathrm{e}^{-\mathrm{i}x} = 2\,\mathrm{Re}(A)\cos(x) - 2\,\mathrm{Im}(A)\sin(x) \tag{9.4}$$

with **amplitude** A. A formal perturbation ansatz based at the critical wave number for the pattern (9.4) is given by

$$k = (1 + \hat{k}_x)\begin{pmatrix}1\\0\end{pmatrix} + \hat{k}_y\begin{pmatrix}0\\1\end{pmatrix}.$$

If we also let $p = \epsilon^2\hat{p}$, and substitute the wavevector perturbation into the dispersion relation, we obtain

$$\sigma = \epsilon^2\hat{p} - (2\hat{k}_x + (\hat{k}_x)^2 + (\hat{k}_y)^2)^2. \tag{9.5}$$

Observe carefully that both directions x and y contribute to the same order in (9.5) if we set $\hat{k}_x \sim \epsilon$, $\hat{k}_y \sim \sqrt{\epsilon}$ and $\sigma \sim \epsilon^2$; see also Figure TODO. This result provides a formal way to look at the size of the unstable band of wavenumbers.

The main idea is now that very close to the bifurcation point, near the onset of the pattern, the amplitude A in (9.4) is not stationary but will be slowly modulated.

Remark: The assumption of slow modulation of the amplitude, also leads to the names **amplitude equation** and **modulation equation**, which are the PDEs we are going to derive below. In this section, the derivation will be formal, while a rigorous justification will be provided in Section 10.

Therefore, we have several scales involved in the problem: the quickly varying **carrier wave** e^{ix} and the slowly varying amplitude (or **envelope**) A; see Figure TODO. Motivated by the dispersion relation, we consider the slow/small variables

$$T := \epsilon^2 t, \quad X := \epsilon x, \quad Y := \sqrt{\epsilon} y,$$

and the fast/large variables $\tilde{x} := x$, $\tilde{t} := t$, for the ansatz

$$u(x, y, t) = u(\tilde{x}, \tilde{y}, \tilde{t}; X, Y, T) = A(X, Y, T)e^{i\tilde{x}} + \text{c.c.}, \tag{9.6}$$

which is also called a **multiple scales** or **two-scale** ansatz. Furthermore, we consider a regime near onset with $p = \epsilon^2 \hat{p}$ and note that the chain rule formally prescribes the relations

$$\partial_x = \partial_{\tilde{x}} + \epsilon \partial_X, \quad \partial_y = \partial_{\tilde{y}} + \sqrt{\epsilon} \partial_Y, \quad \partial_t = \partial_{\tilde{t}} + \epsilon^2 \partial_T. \tag{9.7}$$

Substituting everything into the Swift-Hohenberg equation (9.1) and collecting terms leads to

$$\begin{aligned} \epsilon^2 \partial_T u = \epsilon^2 \hat{p} u - \big[&(1 + \partial^2_{\tilde{x}\tilde{x}})^2 + 2\epsilon(\partial^2_{\tilde{x}\tilde{x}} + 1)(2\partial^2_{\tilde{x}X} + \partial^2_{YY}) \\ &+ 2\epsilon^2 \partial^2_{XX}(\partial^2_{\tilde{x}\tilde{x}} + 1) + \epsilon^2 (2\partial^2_{\tilde{x}X} + \partial^2_{YY})^2 + \cdots \big] u - u^3, \end{aligned} \tag{9.8}$$

where it has been used that u does not depend upon $\tilde{y}$ and $\tilde{t}$. It is now convenient to drop the tildes and revert to the notation x for $\tilde{x}$. A classical ansatz to study equations such as (9.8), is to use an **asymptotic expansion**

$$u = u_0 + \epsilon u_1 + \epsilon^2 u_2 + \epsilon^3 u_3 + \cdots \tag{9.9}$$

where each $u_j = u_j(x, X, Y, T)$ does not depend upon ϵ and $u_0 \equiv 0$ as we **perturb** near the trivial solution.

Remark: One also refers to $\{\epsilon^j\}_{j=0}^{\infty}$ as an asymptotic sequence, i.e., $\lim_{\epsilon\to 0}\epsilon^{j+1}/\epsilon^j = 0$ for all $j\in\mathbb{N}_0$. In principle, other asymptotic sequences depending upon ϵ could work as well such as $\{\epsilon^{j/2}\}_{j=0}^{\infty}$. Usually this involves some trial-and-error for each problem at hand.

Inserting (9.9) into (9.8) we can collect terms at different orders of ϵ and obtain for the first two orders

$$\begin{aligned} \text{at } \mathcal{O}(\epsilon): \qquad & 0 = (\partial_{xx}^2+1)^2u_1 =: L(u_1),\\ \text{at } \mathcal{O}(\epsilon^2): \qquad & L(u_2) = -2(\partial_{xx}^2+1)(2\partial_{xX}^2+\partial_{YY}^2)u_1, \end{aligned}$$

The procedure is to try to find $u_1 = u_1(x,X,Y,T)$ by solving the equations 'order-by-order' so we start with

$$0 = (\partial_{xx}^2+1)^2u_1. \tag{9.10}$$

Using the Fourier transform in the x-component with Fourier variable $\xi\in\mathbb{R}$, we find

$$0 = [(-\mathrm{i}\xi)^4+2(-\mathrm{i}\xi)^2+1)]\hat{u}_1(\xi,X,Y,T) = (\xi^2-1)^2\hat{u}_1(\xi,X,Y,T). \tag{9.11}$$

One possibility is that $\xi=\pm 1$ so in the x-component u_1 is just a linear combination of the Fourier modes $\mathrm{e}^{\pm\mathrm{i}x}$ and if we write

$$u_1(x,X,Y,T) = \tilde{A}(X,Y,T)\mathrm{e}^{\mathrm{i}x}+\text{c.c.}, \tag{9.12}$$

where the amplitude $\tilde{A}(X,Y,T)$ naturally incorporates the trivial solution $u_1\equiv 0$ if $\tilde{A}\equiv 0$. Note that $\tilde{A}=A/\epsilon$ in comparison to the amplitude A in the ansatz (9.4); again, we shall drop the tilde from now on so if we derive an equation for A it is understood that variations of A on the scale $\mathcal{O}(1)$ become small modulations on the original scale.

For the order $\mathcal{O}(\epsilon^2)$, we observe that it does not provide additional information on u_1 as we find the equation $L(u_2)=0$. We now assume that the amplitude from solving $L(u_2)=0$ is zero, i.e., $u_2\equiv 0$. The interesting part occurs at order $\mathcal{O}(\epsilon^3)$ which is given by

$$L(u_3) = -\partial_T u_1 + [\hat{p} - 2\partial_{XX}^2(\partial_{xx}^2+1) - (2\partial_{xX}^2+\partial_{YY}^2)^2]u_1 - u_1^3. \tag{9.13}$$

If we assume that u_3 is bounded as $x\to\pm\infty$ and the expansion is uniformly valid in space-time, then the coefficients in front of $\mathrm{e}^{\pm ix}$ of the right-hand side of (9.13) must vanish. Instead of directly substituting (9.12) into (9.13) a few preliminary calculations are helpful

$$\begin{aligned} \partial_T u_1 &= (\partial_T A)\mathrm{e}^{\mathrm{i}x} + (\partial_T\overline{A})\mathrm{e}^{-\mathrm{i}x},\\ (\partial_{xx}^2+1)u_1 &= (\mathrm{i}^2+1)A\mathrm{e}^{\mathrm{i}x} + ((-\mathrm{i})^2+1)\overline{A}\mathrm{e}^{-\mathrm{i}x} = 0,\\ (2\partial_{xX}^2+\partial_{YY}^2)u_1 &= (2\mathrm{i}\partial_X A+\partial_{YY}^2A)\mathrm{e}^{\mathrm{i}x} + (-2\mathrm{i}\partial_X\overline{A}+\partial_{YY}^2\overline{A})\mathrm{e}^{-\mathrm{i}x}. \end{aligned}$$

So we already see that one term is going to vanish. Furthermore, we have to differentiate the last expression once more, and also calculate the cubic term, which yields

$$\begin{aligned}(2\partial_{xX}^2+\partial_{YY}^2)^2u_1 &= (4\mathrm{i}^2\partial_{XX}A+4\mathrm{i}\partial_{YYX}^3A+\partial_{YYYY}^4A)\mathrm{e}^{\mathrm{i}x}\\ &\quad+(4\mathrm{i}^2\partial_{XX}^2\overline{A}-4\mathrm{i}\partial_{YYX}^3\overline{A}+\partial_{YYYY}^4\overline{A})\mathrm{e}^{-\mathrm{i}x}\\ &= -4\mathrm{e}^{\mathrm{i}x}\left(\partial_X-\frac{\mathrm{i}}{2}\partial_{YY}^2\right)^2A-4\mathrm{e}^{-\mathrm{i}x}\left(\partial_X+\frac{\mathrm{i}}{2}\partial_{YY}^2\right)^2\overline{A},\\ u_1^3 &= (A\mathrm{e}^{\mathrm{i}x}+\overline{A}\mathrm{e}^{-ix})^3=3|A|^2A\mathrm{e}^{\mathrm{i}x}+3|A|^2\overline{A}\mathrm{e}^{-\mathrm{i}x}+\cdots,\end{aligned}$$

which shows that the coefficient of $\mathrm{e}^{\mathrm{i}x}$ is just the complex conjugate of the coefficient for $\mathrm{e}^{-\mathrm{i}x}$ so there is only one amplitude equation to satisfy. Indeed, with the preparations inserting (9.12) into (9.13) easily yields

$$\partial_T A=\hat{p}A-3A|A|^2+4\left(\partial_X-\frac{\mathrm{i}}{2}\partial_{YY}^2\right)^2A. \tag{9.14}$$

There is a slight simplification of the previous equation if we re-scale

$$\hat{X}:=X/2,\qquad \hat{Y}:=Y/\sqrt{2},\qquad \hat{A}:=\sqrt{3}A.$$

Using this scaling, and dropping all the hats from the variables in equation (9.14), finally results in the classical version of the **Newell-Whitehead-Segel equation**

$$\partial_T A=\hat{p}A-A|A|^2+\left(\partial_X-\frac{\mathrm{i}}{2}\partial_{YY}^2\right)^2A. \tag{9.15}$$

Using this equation, we can now study parameter variations of $\hat{p}$; note that $\epsilon^2\hat{p}=p$ so $p=\mathcal{O}(\epsilon^2)$ corresponds to $\hat{p}=\mathcal{O}(1)$. Another amplitude equation is obtained if we negelect the Y-dependence of the amplitude in (9.14) so that

$$\partial_T A=\hat{p}A-3A|A|^2+4\partial_{XX}^2A, \tag{9.16}$$

which in this context is called the **(real) Ginzburg-Landau equation**. The 'real' prefix is chosen, although $A\in\mathbb{C}$, the coefficients for each term are real; see also Section 10. Note carefully that (9.15) and (9.16) again have the cubic nonlinearity we have become accustomed to, when dealing with problems involving bifurcations from trivial branches.

It is emphasized that the calculation so far has been formal and we still need to address the question, in what sense the Newell-Whitehead-Segel and/or the Ginzburg-Landau equations really approximate true solutions of the original system we started with. Indeed, it can be shown that the particular form of the Swift-Hohenberg equation we used as a starting point does not matter as much as one would think: other pattern-forming problems lead to similar (classes of) amplitude equations. Hence, in some sense, an amplitude equation can also be viewed as a 'normal form' in the context for pattern-formation on unbounded domains; see Exercise 9.2

Exercise 9.1. Consider the Swift-Hohenberg equation (9.1) and replace $\Delta + 1$ by $\Delta + k_c$ for some $k_c > 0$. What changes in the derivation of the amplitude equation? $\Diamond$

Exercise 9.2. Suppose that an abstract dispersion relation of the form $\sigma = p - (k^2 - 1)^2 + \mathcal{O}([k^2 - 1]^3)$ as $\|k\| \to 1$. Use the formal methods in this section to show that the linear part of the amplitude equation (9.15) is already prescribed by this dispersion relation. $\Diamond$

Exercise 9.3. Substitute $A = R_0 \mathrm{e}^{\mathrm{i}qX}$ into (9.15) to determine a relationship between p,q and R_0. Explain, why this is related to the occurence of solutions of the form $u = R_0 \mathrm{e}^{\mathrm{i}(1+\epsilon q)x}$; this is called a **roll solution**. $\Diamond$

Background and Further Reading: This section closely follows part of the book [Hoy06], where a lot more details on pattern formation and amplitude equations can be found. The classical survey of the area is [CH93]. Other sources with similar flavour are [CG09, Pis06].

10 Validity of Amplitude Equations

In Section 9 we have seen how to formally derive amplitude equations in the context of the Swift-Hohenberg equation. Here we provide a proof of the amplitude approximation. As before, we consider as a starting point the **Swift-Hohenberg equation**

$$\partial_t u = \mathcal{L}_p u - u^3, \qquad \mathcal{L}_p := p \text{ Id} - (\partial_{xx}^2 + 1)^2, \tag{10.1}$$

where $p \in \mathbb{R}$ is a parameter and $u = u(x,t)$ for $x \in \mathbb{R}$. For the scalings

$$p = \epsilon^2, \qquad T = \epsilon^2 t, \qquad X = \epsilon x,$$

and the formal ansatz

$$u_A(x,t) := \epsilon(A(X,T)\mathrm{e}^{\mathrm{i}x} + \text{c.c.}) \tag{10.2}$$

we already know from Section 9 that the amplitude $A = A(X,T)$ is expected to satisfy the **(real) Ginzburg-Landau equation**

$$\partial_T A = A - 3A|A|^2 + 4\partial_{XX}^2 A. \tag{10.3}$$

The question now is in what sense does u_A approximate u?

Theorem 10.1. *Let u denote the solution of* (10.1). *For each $T_0 > 0$ and $\kappa > 0$ there exist $\epsilon_0, K > 0$ such that for all $\epsilon \in (0, \epsilon_0)$ the following holds: If*

$$|u(x,0) - u_A(x,0)| \leq \kappa\epsilon^2$$

then we have the estimate

$$|u(x,t) - u_A(x,t)| < K\epsilon^2, \qquad \textit{for all } (x,t) \in \mathbb{R} \times [0, T_0/\epsilon^2], \tag{10.4}$$

where u_A is given by (10.2) *and A solves* (10.3).

Remark: Note that u_A and u are $\mathcal{O}(\epsilon)$ near criticality so the error is one order higher in Theorem 10.1. Furthermore, it is natural that we only get a finite-time estimate as the Ginzburg-Landau equation only approximates the Swift-Hohenberg equation. However, the closer we get to criticality as $\epsilon \to 0$, the better the approximation becomes.

Before proving the result, we state a helpful lemma and motivate the strategy.

Lemma 10.2. *For the linearized operator $\mathcal{L}_0$ at criticality we have*

$$\begin{aligned} -\mathcal{L}_0(B(\epsilon x)\mathrm{e}^{ki x}) = [&(1-k^2)^2 B + \epsilon 4\mathrm{i}k(1-k^2)B' \\ &+\epsilon^2(2-6k^2)B'' + \epsilon^3 4\mathrm{i}kB''' + \epsilon^4 B'''']\mathrm{e}^{\mathrm{i}kx} \end{aligned}$$

The proof is left as an exercise calculating derivatives. Using Lemma 10.2 it is easy to see that substituting u_A into (10.1) yields

$$\epsilon^3 \mathrm{e}^{\mathrm{i}x} \partial_T A = [\epsilon^3 \mathrm{e}^{\mathrm{i}x} A + 4\epsilon^3 \mathrm{e}^{\mathrm{i}x} \partial^2_{XX} A - 3\epsilon^3 \mathrm{e}^{\mathrm{i}x} A|A|^2] + \epsilon^3 \mathrm{e}^{3\mathrm{i}x} A^3 + \mathcal{O}(\epsilon^4) \quad (10.5)$$

with the complex conjugate terms understood on both sides. Therefore, the residual error is of order $\mathcal{O}(\epsilon^3)$ and given by the term $\epsilon^3 \mathrm{e}^{3\mathrm{i}x} A^3$. Upon integrating the equation up to time T_0/ϵ^2, a total error of $\mathcal{O}(\epsilon)$ seems to remain and this is not good enough. The idea of the following proof is to use a modified approximation.

Proof. (of Theorem 10.1) The improved approximation is defined as

$$v_A(x,t) = \epsilon A(X,T)\mathrm{e}^{\mathrm{i}x} - \frac{\epsilon^3}{64} A(X,T)^3 \mathrm{e}^{3\mathrm{i}x} + \text{c.c.}$$

The idea is to study the error by looking at

$$R(x,t) := \frac{u(x,t) - v_A(x,t)}{\epsilon^2}.$$

If we can show that

$$\|R(x,t)\|_\infty = \sup_{x\in\mathbb{R}} |R(x,t)|$$

is bounded for $t \in [0, T_0/\epsilon^2]$ by a constant independent of ϵ, then we have

$$|u(x,t) - u_A(x,t)| = |\epsilon^2 R(x,t) - \epsilon^3 (A^3 \mathrm{e}^{3\mathrm{i}x} + \text{c.c.})/64| = \mathcal{O}(\epsilon^2)$$

for $(x,t) \in \mathbb{R} \times [0, T_0/\epsilon^2]$ and the result follows. Hence, it remains to analyze $R(x,t)$. We calculate

$$\begin{aligned}
\epsilon^2 \partial_t R &= \partial_t u - \partial_t v_A \\
&= \underbrace{\partial_t u - \mathcal{L}_0 u - \epsilon^2 u + u^3}_{=0} + \mathcal{L}_0 u + \epsilon^2 u - u^3 - \partial_t v_A \\
&= \mathcal{L}_0(\epsilon^2 R + v_A) + \epsilon^2(\epsilon^2 R + v_A) - (\epsilon^2 R + v_A)^3 - \partial_t v_A \\
&= \epsilon^2 \mathcal{L}_0 R + \epsilon^2 R(\epsilon^2 - 3v_A^2) + \epsilon^3(-\epsilon^3 R - 3\epsilon R^2 v_A) \\
&\quad -\partial_t v_A - v_A^3 + \epsilon^2 v_A + \mathcal{L}_0 v_A.
\end{aligned}$$

Upon grouping terms and dividing through by ϵ^2 we find

$$\partial_t R = \mathcal{L}_0 R + \epsilon^2 a(x,t;\epsilon) R + \epsilon^3 b(x,t,R;\epsilon) + \epsilon^2 r(x,t;\epsilon) \quad (10.6)$$

where $a(x,t;\epsilon) = 1 - 3(v_A/\epsilon)^2$, $b(x,t,R;\epsilon) = -\epsilon R^3 - 3(v_A/\epsilon)R^2$ and $r(x,t;\epsilon) = [-\partial_t v_A - v_A^3 + \epsilon^2 v_A + \mathcal{L}_0 v_A]/\epsilon^4$. The last term is the interesting one since

$$
\begin{aligned}
-\epsilon^4 r(x,t;\epsilon) &= \partial_t v_A + v_A^3 - \epsilon^2 v_A - \mathcal{L}_0 v_A \\
&= \epsilon^3 \left[(\partial_T A - 4\partial_{XX}^2 A - A)\mathrm{e}^{\mathrm{i}x} - (1-3^2)^2 \frac{1}{64} A^3 \mathrm{e}^{3\mathrm{i}x} + \text{c.c.} \right] \\
&\quad + \epsilon^3 (A\mathrm{e}^{\mathrm{i}x} + \text{c.c.})^3 + \mathcal{O}(\epsilon^4)
\end{aligned}
$$

by applying Lemma 10.2. Indeed, the $\mathcal{O}(\epsilon^3)$-term in the last expression complete vanishes (which explains the choice of prefactor 1/64) and this shows that $r(x,t;\epsilon)$ is bounded over $(0,\epsilon_0) \times [0,\infty) \times \mathbb{R}$. Furtermore, we have for our given $\kappa > 0$ that

$$
|R(x,0)| \leq 2\kappa
$$

if ϵ_0 is chosen sufficiently small. To study (10.6) we are going to use some basic semigroup theory; see Section TODO. In particular, we drop the x-variable in the notation and write (10.6) as

$$
\partial_t R = \mathcal{L}_0 R + h(t), \qquad R(0) = R_0. \tag{10.7}
$$

The linear part $\partial_t R = \mathcal{L}_0 R$ is solved by a uniformly bounded strongly continuous semigroup $\mathrm{e}^{t\mathcal{L}_0}$ when the equation is considered on $C_b(\mathbb{R})$ with the usual supremum norm $\|\cdot\|_\infty$; see Lemma 10.3. Hence we may write the solution of (10.7), respectively (10.6), via the **variation-of-constants formula** or **Duhamel's formula**

$$
\begin{aligned}
R(t) &= \mathrm{e}^{t\mathcal{L}_0} R(0) + \int_0^t \mathrm{e}^{(t-s)\mathcal{L}_0} h(s)\ \mathrm{d}s \qquad (10.8) \\
&= \mathrm{e}^{t\mathcal{L}_0} R(0) + \epsilon^2 \int_0^t \mathrm{e}^{(t-s)\mathcal{L}_0} [a(s;\epsilon)R(s) + \epsilon b(s,R(s);\epsilon) + r(s;\epsilon)]\ \mathrm{d}s
\end{aligned}
$$

For a given $\delta > 0$ we have $\|b(s,R(s);\epsilon)\|_\infty \leq K_b$ for all R with $\|R(s)\|_\infty \leq \delta$ and $\epsilon \in (0,\epsilon_0)$. Since the semigroup is uniformly bounded we have $\|\mathrm{e}^{s\mathcal{L}_0}\|_\infty \leq K$ for some $K > 0$. We can also ensure that $\|r(s;\epsilon)\|_\infty \leq K$ by the argument above about the boundedness of r and clearly $|a(s;\epsilon)| \leq K$. Then one estimates (10.8) and obtains

$$
\|R(t)\|_\infty \leq 2K\kappa + \int_0^t \epsilon^2 K^2 \|R(s)\|_\infty\ \mathrm{d}s + \epsilon^2 t K(\epsilon K_b + K) \tag{10.9}
$$

as long as $\|R(s)\|_\infty \leq \delta$ holds. **Gronwall's inequality** (see Exercise 10.6) applied to (10.9) yields

$$
\|R(t)\|_\infty \leq [2K\kappa + T_0 K(\epsilon K_b + K)]\mathrm{e}^{\epsilon^2 K^2 t} \tag{10.10}
$$

and for $t \leq T_0/\epsilon^2$, we can easily see that upon making ϵ small so that $\epsilon K_b \leq K$, we find that $\|R(t)\|_\infty$ is bounded independent of ϵ. □

Lemma 10.3. *([KSM92, Lem.2.3]) The semigroup* $\mathrm{e}^{t\mathcal{L}_0} : C_b(\mathbb{R}) \to C_b(\mathbb{R})$ *is strongly continuous and uniformly bounded.*

We shall not prove this result here but instead remark that the main idea of the proof to insert a clever approximation to remove higher-order correction terms, is not limited to the Ginzburg-Landau equation. Consider the **sine-Gordon equation**

$$\partial_{tt}^2 u = \partial_{xx}^2 u - \sin u, \qquad u = u(x,t), \; x \in \mathbb{R}. \tag{10.11}$$

Then it is shown in Exercise (10.7) that the dispersion relation for the linearized sine-Gordon equation is given by $\omega^2 = k^2 + 1$.

Definition 10.4. For a given dispersion replation $\omega(k)$, the **group velocity** ν is defined by $\nu := \frac{\mathrm{d}\omega}{\mathrm{d}k}$.

The solutions we are interested in for the sine-Gordon equation are of the form

$$u_A(x,t) = \epsilon A(\epsilon t, \epsilon(x - \nu t))\mathrm{e}^{\mathrm{i}(kx - \omega t)} + \text{c.c.}. \tag{10.12}$$

It can be shown using a formal calculation via the method of multiple scales as in Section 9 that the amplitude A satisfies the **nonlinear Schrödinger equation (NLS)**

$$2i\omega\partial_T A = (\nu^2 - 1)\partial_T A + \frac{1}{2}A|A|^2. \tag{10.13}$$

Theorem 10.5. *([KSM92]) Let u denote the solution of* (10.11)*. For each $T_0 > 0$ and $\kappa > 0$ there exist $\epsilon_0, K > 0$ such that for all $\epsilon \in (0, \epsilon_0)$ the following holds: If*

$$\|u(\cdot,0) - u_A(\cdot,0)\|_{L^2(\mathbb{R})} \leq \kappa\epsilon^{3/2}$$

then we have the estimate

$$\|u(\cdot,t) - u_A(\cdot,t)\|_{L^2(\mathbb{R})} < K\epsilon^{3/2}, \qquad \textit{for all } (x,t) \in \mathbb{R} \times [0, T_0/\epsilon^2], \tag{10.14}$$

where u_A is given by (10.12) *and A solves* (10.13)*.*

Theorem 10.5 can be proven in a similar way as Theorem 10.1. One has to consider the linearized semigroup on a different Banach space. Furthermore, a new improved approximation has to be used as discussed in Exercise 10.8.

Exercise 10.6. Prove **Gronwall's inequality**, i.e., show that for a given time interval $(a, b) \subset \mathbb{R}$ and continuous functions α, β, g the inequality

$$g(t) \leq \alpha(t) + \int_a^t \beta(s) g(s) \; \mathrm{d}s, \qquad \forall t \in (a, b),$$

implies that

$$g(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s) \exp\left(\int_s^t \beta(r) \; \mathrm{d}r\right) \; \mathrm{d}s,$$

for $t \in (a, b)$. Show that if in addition $\alpha(t)$ non-decreasing then

$$g(t) \leq \alpha(t) \exp\left(\int_a^t \beta(r) \; \mathrm{d}r\right),$$

which is the version we used above. $\Diamond$

Exercise 10.7. Show that the dispersion relation for the linearized sine-Gordon equation is given by $\omega^2 = k^2 + 1$ using the ansatz $\mathrm{e}^{\mathrm{i}(kx - \omega t)}$. $\Diamond$

Exercise 10.8. Show that the improved approximation for the sine-Gordon equation leading to the NLS equation is given by

$$v_A(x, t) = \epsilon A(X, T) \mathrm{e}^{\mathrm{i}(kx - \omega t)} - \frac{1}{54k^2 - 54\omega^2 + 6} \epsilon^3 A(X, T) \mathrm{e}^{3\mathrm{i}(kx - \omega t)} + \text{c.c.},$$

where $T = \epsilon^2 t$ and $X = \epsilon(x - \nu t)$. $\Diamond$

Background and Further Reading: This section is based upon the paper [KSM92] which developed the idea of using improved approximations to get rid of certain error terms. Other approaches to prove validity of amplitude equations can be found in [CE90, vH91]. For more on the sine-Gordon equation see [CMKW14]. The GLE and NLS equations are amplitude equations for many different problems, e.g., the NLS equation is an amplitude equation for certain water wave problems [Lan13].

11 Semigroups and Sectorial Operators

For some time-dependent dynamical problems, it will be convenient to have the language and some basic results from semigroup theory available.

Definition 11.1. Let X be a Banach space. A **(strongly continuous) semigroup** on X is a family of continuous linear operators $\{S(t)\}_{t\geq 0}$ on X such that

(S1) $S(0) = \mathrm{Id}$,

(S2) $S(t)S(s) = S(t+s)$ for $t, s \geq 0$,

(S3) $\|S(t)u - u\|_X \to 0$ as $t \searrow 0$ for every $u \in X$.

Remark: Flows generated by sufficiently smooth ODEs are automatically semigroups. In fact, they are groups as (S2) also holds for negative times.

All semigroups we consider will be strongly continuous so we omit this prefix. However, Definition 11.1 can be modified by requiring additional regularity, e.g., $S(t)$ is an **analytic semigroup** if $t \mapsto S(t)u$ is real analytic for $t \in (0, +\infty)$ and every $u \in X$.

Definition 11.2. The **infinitesimal generator** of a semigroup $S(t) : X \to X$ is defined by

$$Au := \lim_{t \searrow 0} \frac{1}{t}(S(t)u - u)$$

with domain $D(A)$ consisting of all $u \in X$, where the limit exists. If the generator can be identified, then we write $S(t) = \mathrm{e}^{tA}$.

An important example is the linear ODE $\frac{\mathrm{d}u}{\mathrm{d}t} = Au$, where the matrix $A \in \mathbb{R}^{d\times d}$ is easily seen to be the infinitesimal generator of the flow $\phi(u_0, t) = u(t)$, which also defines a semigroup $S(t)$.

Example 11.3. The natural example from the context of PDEs is the heat equation

$$\partial_t u = \Delta u, \qquad (x,t) \in \Omega \times [0, +\infty),\ u = u(x,t) \in \mathbb{R}, \tag{11.1}$$

say for $\Omega = [0,1]$ and with Dirichlet boundary conditions for simplicity. We shall see below that $A = \Delta$ is the infinitesimal generator for an analytic semigroup $\mathrm{e}^{t\Delta} : L^2(\Omega) \to L^2(\Omega)$ with $D(\Delta) = H_0^1(\Omega) \cap H^2(\Omega)$. In fact, Δ is also a closed and densely defined operator on $L^2(\Omega)$. ♦

The next definition is useful to set up an abstract framework to check, which operators generate (analytic) semigroups.

Definition 11.4. A linear operator $A : X \to X$ is called **sectorial** if it is closed, densely defined, and for some $\theta \in (\frac{\pi}{2}, \pi)$, $M \geq 1$, $a \in \mathbb{R}$, the sector

$$\mathcal{S}_{a,\theta} := \{\lambda \in \mathbb{C} : |\arg(\lambda - a)| \leq \theta, \lambda \neq a\}$$

is contained in the **resolvent set** $\rho(A) := \mathbb{C} - \sigma(A)$, and furthermore

$$\|(\lambda \,\mathrm{Id} - A)^{-1}\| \leq \frac{M}{|\lambda - a|} \tag{11.2}$$

for all $\lambda \in \mathcal{S}_{a,\theta}$.

We refer to Figure TODO for an illustration of the sector $\mathcal{S}_{a,\theta}$.

Remark: Slightly different versions of the definition of sectorial operators exist, particularly regarding sign conventions.

Example 11.5. Continuing Example 11.3 with $\Delta = \partial_{xx}^2$ on $\Omega = [0, 1]$ we know from Section 4 that the eigenvalues of ∂_{xx}^2 are $\lambda_n = -(\pi n)^2$ with $n \in \{1, 2, 3, \ldots\}$. So ∂_{xx}^2 is a sectorial operator on $L^2(\Omega)$, with

$$D(\partial_{xx}^2) = \{u \in L^2(\Omega) : \partial_{xx}^2 u \in L^2(\Omega)\} = H_0^1(\Omega) \cap H^2(\Omega), \tag{11.3}$$

where we could e.g. take $a = 0$ and any $\theta \in (\frac{\pi}{2}, \pi)$ in Definition 11.4. In fact, the Laplacian considered with suitable boundary conditions on a sufficiently regular domain $\Omega \subset \mathbb{R}^d$ and many other elliptic operators turn out to be sectorial. ♦

Theorem 11.6. *([Hen81]) If A is sectorial then A is the infinitesimal generator of an analytic semigroup e^{tA} with the representation formula*

$$\mathrm{e}^{tA} = \frac{1}{2\pi \mathrm{i}} \int_\Lambda (\lambda \,\mathrm{Id} - A)^{-1} \mathrm{e}^{\lambda t} \, \mathrm{d}\lambda, \tag{11.4}$$

where Λ is a contour in $\rho(A)$ with $\arg(\lambda) \to \pm\theta^*$ as $|\lambda| \to +\infty$ for $\theta^* \in (\frac{\pi}{2}, \pi)$. Furthermore, if $\mathrm{Re}(\lambda) < a$ for $\lambda \in \sigma(A)$ then

$$\|\mathrm{e}^{tA}\| \leq K \mathrm{e}^{at}, \quad \textit{for } t > 0 \tag{11.5}$$

and some $K > 0$.

See also Figure TODO for an illustration. The **Dunford integral** (11.4) involving the **resolvent** $(\lambda \mathrm{Id} - A)^{-1}$ of A is frequently key to work with semigroups as it often convenient to work with resolvents.

Example 11.7. We continue with Example 11.5. Let $u_0 = u(x,0)$ be the initial value. Since the operator $\Delta = \partial^2_{xx}$ is sectorial, Theorem 11.6 implies that it generates the analytic semigroup $\mathrm{e}^{t\Delta}$. Therefore, the heat equation (11.1) is solved by

$$u(t) = \mathrm{e}^{t\Delta} u_0. \tag{11.6}$$

However, by separation of variables we can solve the heat equation by

$$u(x,t) = \sum_{n=1}^{\infty} \mathrm{e}^{\lambda_n t} \langle e_n, u_0 \rangle_{L^2(\Omega)} e_n(x),$$

where $e_n(x) = \sqrt{2}\sin(n\pi x)$ is the eigenfunction for λ_n. This gives concrete formulas to work with A and e^{tA} by their action on basis functions. For example, we have

$$(\lambda \,\mathrm{Id} - \Delta)^{-1} v = \sum_{n=1}^{\infty} (\lambda - \lambda_n)^{-1} \ \langle e_n, v \rangle_{L^2(\Omega)} \ e_n(x)$$

for $v \in L^2(\Omega)$. Another example defining $(-\Delta)^\alpha$ and computing $D((-\Delta)^\alpha)$ for $\alpha \geq 0$ is considered in Exercise 11.13. ♦

One may generalize the previous observations about $(-\Delta)^\alpha$.

Definition 11.8. Suppose A is a sectorial operator and $\mathrm{Re}(\sigma(A)) < 0$. Then for any $\alpha > 0$ define

$$(-A)^{-\alpha} := \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} \mathrm{e}^{tA} \,\mathrm{d}t,$$

where $\Gamma(\alpha)$ denotes the **gamma function** (recall $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$ and $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}\mathrm{e}^{-x} \,\mathrm{d}x$ for $\alpha \in \mathbb{C} - \{0, -1, -2, \ldots\}$).

Theorem 11.9. *([Hen81, Thm.1.4.2]) If A is sectorial with $\mathrm{Re}(\sigma(A)) < 0$, then for any $\alpha > 0$, the operator $(-A)^{-\alpha}$ is bounded and injective.*

The last theorem shows, why it is sometimes nicer to analyze A via its negative powers. Furthermore, one can also use a similar idea with positive exponents to define spaces, which turns out to be very suitable for many dynamical problems.

Definition 11.10. Let A be sectorial on X and fix $a \in \mathbb{R}$ such that $A_1 := A + a\,\mathrm{Id}$ satisfies $\mathrm{Re}(\sigma(A_1)) < 0$. For $\alpha \geq 0$, define the **fractional power space** $X^\alpha := D(A_1^\alpha)$ with the norm $\|x\|_\alpha := \|(-A_1)^\alpha x\|_X$ for $x \in X^\alpha$.

The norms $\|\cdot\|_\alpha$ turn out to be equivalent for different feasible choices of $a \in \mathbb{R}$. It is easy to see that X^α are Banach spaces. The main reason to use the spaces X^α is that they provide well-defined domains for operators. Furthermore, there are **embedding theorems** for $X = L^p(\Omega)$ which show, under suitable conditions, that X^α is a subset of smooth functions $C^k(\Omega)$ or inside some Sobolev space $W^{k,q}(\Omega)$ for some values k, q. As before, we shall not worry about regularity assumptions and just select the domain, the boundary and initial data, and those equations, which are sufficiently regular.

The next step is to move from the linear problem $\frac{\mathrm{d}u}{\mathrm{d}t} = Au$ to the inhomogeneous problem

$$\frac{\mathrm{d}u}{\mathrm{d}t} = Au + f(t), \qquad u \in X,\ f : (0,T) \to X,\ u(0) = u_0. \tag{11.7}$$

If A generates a strongly continuous semigroup e^{tA}, then we may always consider a **mild solution** to (11.7) given by

$$u(t) = \mathrm{e}^{tA} u_0 + \int_0^t \mathrm{e}^{(t-s)A} f(s)\ \mathrm{d}s. \tag{11.8}$$

Theorem 11.11. *([Hen81, Thm.3.2.2]) Consider* (11.7) *and assume that A is a sectorial operator. Furthermore, suppose $f : (0,T) \to X$ is continuous and $\int_0^\rho \|f(t)\|\ \mathrm{d}t < +\infty$ for some $\rho > 0$. Then there is a unique strong (classical) solution of* (11.7) *which coincides with the mild solution* (11.8).

Although the Laplacian, classical elliptic operators on L^2 and the related nonlinear problems are key examples in semigroup theory for PDEs, the framework is obviously not limited to this setup as the next example illustrates.

Example 11.12. Consider the linear evolution equation

$$\partial_t u = \mathcal{L}_0 u := -(1 + \partial_{xx}^2)^2 u, \qquad u = u(x,t),\ x \in \mathbb{R}, \tag{11.9}$$

arising in the context of the **Swift-Hohenberg equation**; see also Section (10). In fact, it is relatively easy to check that

$$\mathrm{e}^{t\mathcal{L}_0} : C_\mathrm{b}(\mathbb{R}) \to C_\mathrm{b}(\mathbb{R})$$

is a strongly continuous semigroup. Lemma 10.3 states that the semigroup is uniformly bounded. Although we do not give a full proof, let us motivate this fact formally. The dispersion relation (9.3) yields that at criticality the spectrum touches the imaginary axis precisely at zero and the rest of the (essential) spectrum is contained in the left-half complex plane. Hence, it is natural to consider Definition 11.4 with a sector $\mathcal{S}_{0,\theta}$. It turns out that $\theta = 2\pi/3$ is a good choice for the angle; see Figure TODO. Motivated by (11.2), we aim to estimate the resolvent. Consider $\omega := (-\lambda)^{1/2}$, $\mathrm{Im}(\omega) > 0$ and the operator splitting

$$\mathcal{L}_0 v - \lambda v = -(1 + \partial_{xx}^2 + \omega)(1 + \partial_{xx}^2 - \omega)v = h$$

for some h. This yields the solution

$$v = (\mathcal{L}_0 - \lambda)^{-1} h = -[G_{\beta+} \circ G_{\beta-}]h, \quad \beta_\pm = (-1 \pm \omega)^{1/2},\ \mathrm{Re}(\beta_\pm) > 0,$$

where the operator

$$[G_\beta h](x) = -\int_{\mathbb{R}} \frac{1}{2\beta} \mathrm{e}^{-\beta|x-\xi|} h(\xi)\ \mathrm{d}\xi \tag{11.10}$$

can be derived using Fourier transforms; see Exercise 11.15. Therefore, we obtain

$$\begin{aligned} \|(\mathcal{L}_0 - \lambda\ \mathrm{Id})^{-1}\| &\leq \frac{1}{|\beta_+|\mathrm{Re}(\beta_+)} \frac{1}{|\beta_-|\mathrm{Re}(\beta_-)} \\ &\leq \frac{1}{\mathrm{Re}(\beta_+)\mathrm{Re}(\beta_-)} \frac{1}{\sqrt{|1+\lambda|}}. \end{aligned} \tag{11.11}$$

From this inequality it is not too difficult to obtain that

$$\|(\mathcal{L}_0 - \lambda\ \mathrm{Id})^{-1}\| \leq \frac{M}{|\lambda|} \tag{11.12}$$

for all $\lambda \in \mathcal{S}_{0,\theta}$. Then we may conclude sectoriality and employ the Dunford integral (11.4) to see that the semigroup is uniformly bounded. ♦

Exercise 11.13. Consider the 1-dimensional Laplacian $\Delta = \partial_{xx}^2$ with Dirichlet boundary conditions on $\Omega = [0,1]$ from Example 11.7 and define

$$(-\Delta)^\alpha v = \sum_{n=1}^{\infty} (-\lambda_n)^\alpha\ \langle e_n, v\rangle_{L^2(\Omega)}\ e_n(x)$$

for $v \in L^2(\Omega)$ where λ_n denotes the eigenvalue of Δ with eigenfunction e_n. Prove that

$$D((-\Delta)^\alpha) = \left\{ v \in L^2(\Omega) : \sum_{n=1}^{\infty} \lambda_n^{2\alpha} \langle e_n, v\rangle_{L^2(\Omega)}^2 < +\infty \right\}. \tag{11.13}$$

Furthermore, prove that $D((-\Delta)^{1/2}) = H_0^1(\Omega)$. Note that this exercise generalizes to a domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$. ♢

Exercise 11.14. Show that if $-A$ is positive definite and self-adjoint then so is $(-A)^\alpha$ for all $\alpha > 0$. $\Diamond$

Exercise 11.15. Consider Example 11.12 and formally solve the equation

$$(1 + \omega + \partial_{xx}^2)v = h, \qquad v = v(x),\ x \in \mathbb{R},$$

using the Fourier transform; this calculation essentially yields (11.10). Furthermore, prove that (11.11) implies (11.12) for $\lambda \in \mathcal{S}_{0,\frac{2}{3}\pi}$. $\Diamond$

Background and Further Reading: This section is mainly based upon the introduction to analytic semigroups in [Hen81]. Other important sources for semigroup theory are [EN00, Paz83]. The uniform boundedness of the linearized Swift-Hohenberg equation at criticality is taken from [KSM92].

12 Dissipation and Absorbing Sets

Roughly speaking, dissipation refers to the frequent occurence in partial differential equations, that the dynamics of the system contracts or reduces to a subset of phase space.

Example 12.1. Prototypical examples for dissipative systems are reaction-diffusion equations such as

$$\partial_t u = \Delta u + f(u), \quad (x,t) \in \Omega \times [0,+\infty),\ u = u(x,t), \tag{12.1}$$

for a suitable sufficiently smooth nonlinearity $f : \mathbb{R} \to \mathbb{R}$ on a bounded smooth domain $\Omega \subset \mathbb{R}^d$ with Dirichlet boundary conditions $u(x) = 0$ for $x \in \partial\Omega$. ♦

Let X be a Banach space and suppose we may view the solution of the PDE $u(x,t)$ with initial condition $u_0 \in X$ as a **semiflow** (or **nonlinear semigroup**)

$$S(t)u_0 = u(x,t), \qquad S(t) : X \to X.$$

In particular, $S(t)$ satisfies $S(0) = \mathrm{Id}$, S is continuous in t and in the argument u_0, and $S(t+s) = S(t)S(s)$ for $t, s \geq 0$; see also Definition 11.1. There are several different notions of dissipativity. For us, the following will suffice:

Definition 12.2. $S(t)$ is called **(bounded) dissipative** if there exists a bounded set $\mathcal{B} \subset X$, such that for each bounded set $\mathcal{Y}$ there exists a time $t_{\mathcal{Y}}$ such that $S(t)\mathcal{Y} \subset \mathcal{B}$ for all $t \geq t_{\mathcal{Y}}$.

The previous statement can be re-phrased loosely by saying that there exists a bounded **absorbing set** for the dynamics.

Remark: Frequently one also finds explicit requirements on the dynamics such as a **dissipative estimate**

$$\|S(t)u_0\|_X \leq Q(\|u_0\|_X)\mathrm{e}^{-\alpha t} + K$$

for some constants $K, \alpha > 0$ and a monotone increasing function Q.

The natural question to ask is, which PDEs are actually dissipative?

Example 12.3. (Example 12.1) We continue with the reaction-diffusion equation (12.1) and make the additional assumptions

$$-K - \alpha_1|v|^p \leq f(v)v \leq K - \alpha_2|v|^p, \qquad f'(v) \leq K_d, \tag{12.2}$$

for all $v \in \mathbb{R}$ and some constants $K, K_d, \alpha_{1,2} > 0$. It also helps to simplify the calculations to require $f(0) = 0$. A typical example is the case

$$f(v) = v - v^3.$$

Indeed, in this case we have

$$-|v|^4 \leq f(v)v = v^2 - v^4 \leq K - \frac{1}{2}|v|^4$$

and the upper bound on the derivative is trivial to show. Under the assumptions (12.2) it is not difficult to see that $S(t)u_0 = u(x,t)$ yields a semiflow with

$$u(t) \in C(0,T;L^2(\Omega)) \cap L^2(0,T;D((-\Delta)^{1/2})) \tag{12.3}$$

and we know from Exercise 11.13 that $D((-\Delta)^{1/2}) = H_0^1(\Omega)$. Furthermore, we expect from Theorem 11.11 that solutions should be classical ones for many cases of the nonlinearity f. ♦

A main strategy to prove the existence of a suitable bounded absorbing set is to consider a PDE on a suitable Banach space X on which the differential operators appearing in the equation are at least densely-defined and which is "tractable" analytically. For the reaction-diffusion equation (12.1) one excellent guess is to take $X = L^2(\Omega)$.

Theorem 12.4. *Suppose* (12.1) *satisfies* (12.2). *Then there exists constants* $K_1, K_2 > 0$ *and a time* $t_{u_0} > 0$ *such that*

$$\|u(t)\|_{L^2(\Omega)} \leq K_1 \tag{12.4}$$

for all $t \geq t_{u_0}$. *Furthermore, we have the bound*

$$\int_t^{t+1} \|u(s)\|^2_{H_0^1(\Omega)} \,\mathrm{d}s \leq K_2 \tag{12.5}$$

for all $t \geq t_{u_0}$.

Proof. The domain Ω will be dropped as a subscript in the spatial norm notation throughout the proof. We write (12.1) in the form

$$\frac{\mathrm{d}u}{\mathrm{d}t} - \Delta u = f(u), \qquad u = u(t) = u(x,t), \tag{12.6}$$

and take the L^2-inner product with u, which yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|^2_{L^2} + \|u\|^2_{H_0^1} = \int_\Omega f(u(x))u(x)\,\mathrm{d}x \leq \int_\Omega K - \alpha_2|u|^p\,\mathrm{d}x, \tag{12.7}$$

where we used (12.2) in the last inequality. In the next step, we use a version of the **Poincaré inequality** given by

$$\|u\|_{L^2} \leq \lambda^{-1/2}\|\nabla u\|_{L^2} = \lambda^{-1/2}\|u\|_{H_0^1},$$

where $\lambda > 0$ is the smallest eigenvalue of $-\Delta$ on Ω. Using this inequality and dropping the negative term on the right-hand side in (12.7) leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{L^2}^2 + 2\lambda\|u(t)\|_{L^2}^2 \leq K, \tag{12.8}$$

for some constant $K > 0$. By Gronwall's inequality we may conclude that

$$\begin{aligned}
\|u\|_{L^2}^2 &\leq \mathrm{e}^{-2\lambda t}\|u_0\|_{L^2}^2 + \int_0^t \mathrm{e}^{-2\lambda t+2\lambda s} K \ \mathrm{d}s \\
&= \mathrm{e}^{-2\lambda t}\|u_0\|_{L^2}^2 + K\mathrm{e}^{-2\lambda t}\int_0^t \mathrm{e}^{2\lambda s} \ \mathrm{d}s \leq \mathrm{e}^{-2\lambda t}\|u_0\|_{L^2}^2 + \frac{K}{2\lambda}.
\end{aligned}$$

The first result (12.4) now follows. The second result (12.5) is left as Exercise 12.9. □

Theorem 12.4 is a typical first step to check whether there is a possibility that the long-term dynamics of a PDE is finite-dimensional. The second estimate (12.5) helps to construct a better absorbing set in H_0^1. The idea is to multiply (12.6) by Δu and integrate, which yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{H_0^1}^2 + \|\Delta u\|_{L^2}^2 \leq K\|u\|_{H_0^1}^2, \tag{12.9}$$

for some constant $K > 0$, by using the upper bound of the derivative f' from (12.2). Integrating the last inequality between s and t we get

$$\|u(t)\|_{H_0^1}^2 \leq 2K\int_s^t \|u(r)\|_{H_0^1}^2 \ \mathrm{d}r + \|u(s)\|_{H_0^1}^2.$$

Integrating again, now with respect to s and between $t-1$ and t (with $t \geq 1$) yields

$$\|u(t)\|_{H_0^1}^2 \leq (K+1)\int_{t-1}^t \|u(s)\|_{H_0^1}^2 \ \mathrm{d}s \leq K_2(1+K), \tag{12.10}$$

where the last inequality follows from (12.5). Note that the last calculation was formal since we required regularity of $u(t)$ that we may not have but this can be made rigorous by either proving directly that solutions are regular enough or by using a Galerkin approximation, which we do not detail here.

Theorem 12.5. *Suppose* (12.1) *satisfies* (12.2) *and* $u(x,t)$ *is sufficiently smooth. Then there exists an absorbing set in* H_0^1*.*

Corollary 12.6. *Suppose* (12.1) *satisfies* (12.2) *and* $u(x,t)$ *is sufficiently smooth. Then there exists a compact absorbing set in* L^2*.*

Proof. Using Theorem 12.5 and the compact embedding of H^1 into L^2 the result follows. □

Reaction-diffusion equations are a good benchmark scenario for other classes of dissipative PDEs.

Example 12.7. Another classical, albeit technically more involved, example are the **(incompressible) Navier-Stokes equations**

$$\partial_t u - \nu \Delta u + (u \cdot \nabla) u + \nabla p = f(x,t), \qquad \nabla \cdot u = 0,$$

where $u = u(x,t)$ represents the velocity of a fluid in a domain $\Omega \subset \mathbb{R}^d$ ($d = 2,3$), $\nu > 0$ controls the strength of the viscosity, f is a forcing term, and $\nabla \cdot u = 0$ is the incompressibility condition. One may actually eliminate the pressure term and, with quite a few calculations, end up with

$$\frac{\mathrm{d}u}{\mathrm{d}t} + \nu A u + B(u,u) = f \tag{12.11}$$

where A is a linear operator and B represents the (quadratic!) nonlinearity of the Navier-Stokes equations. The simplest domain to consider is $\Omega = [0,L] \times [0,L] \subset \mathbb{R}^2$ with further simplifications

$$\int_\Omega u_0(x)\ \mathrm{d}x = 0, \qquad \int_\Omega f(x,t)\ \mathrm{d}t = 0$$

for the initial condition and the forcing term. For this case, the existence, uniqueness and regularity theory for the Navier-Stokes equations is complicated but still possible to carry out via classical techniques. Then one may carry out similar arguments to show that (12.11) in the two-dimensional periodic domain Ω generates a semiflow on

$$L^2_{\mathrm{NS}} := \left\{ u \in L^2(\Omega) : \int_\Omega u(x)\ \mathrm{d}x = 0, \nabla \cdot u = 0, u \text{ periodic} \right\}$$

as well as on the suitable Sobolev space analog

$$H^1_{\mathrm{NS}} := \left\{ u \in H^1(\Omega) : \int_\Omega u(x)\ \mathrm{d}x = 0, \nabla \cdot u = 0, u \text{ periodic} \right\}.$$

In this setup, similar arguments as demonstrated above for the reaction-diffusion case can be used to establish the existence of absorbing sets in L^2_{NS} as well as in H^1_{NS}. ♦

Exercise 12.8. A semiflow $S(t) : X \to X$ is called **point dissipative** if there exists a bounded set $\mathcal{B}$ such that for every initial condition u_0 there exists a time t_{u_0} such that $S(t)u_0 \subset \mathcal{B}$ for all $t \geq t_{u_0}$. Prove that if $X = \mathbb{R}^N$ then point dissipativity implies bounded dissipativity. Hint: Use the Heine-Borel Theorem. ♢

Exercise 12.9. Prove the inequality (12.5) from Theorem (12.4). $\Diamond$

Exercise 12.10. Prove the inequality (12.9) under the assumptions that u is sufficiently smooth and f' is bounded above. $\Diamond$

Background and Further Reading: The presentation here is based upon [Rob01]. Other important source are [BV92, Tem97]. For more details on the Navier-Stokes we refer to [DG95].

13 Nonlinear Saddles and Invariant Manifolds

In this chapter we study the class of PDEs

$$\partial_t u + Au = f(u), \qquad u_0 \in H, \tag{13.1}$$

where $-A : H \to H$ is a linear, negative, self-adjoint, and sectorial operator on a Hilbert space H and f is a sufficiently smooth globally Lipschitz nonlinearity. From Section 11 it follows that there exists a well-defined solution

$$u \in C^0(0, T; H) \cap L^2(0, T; D(A^{1/2})), \quad u(t) = S(t)u_0 \tag{13.2}$$

for a semiflow $S(t) : H \to H$. Furthermore, we require that f is globally bounded in H. This is not really a restriction here as we are going to be interested here in local dynamics as well as globally attracting manifolds. In Section 12 we showed that certain classes of PDEs have global bounded absorbing sets and if this fact has been shown, then we can just cut off the nonlinearity outside some large ball.

The final goal of our analysis here and in Lecture 14 is to establish that certain PDEs of the form (13.1) essentially reduce to finite-dimensional ODE problems.

Definition 13.1. Consider a semiflow $S(t) : H \to H$ and define an **inertial manifold** $\mathcal{M}$ as a finite-dimensional, exponentially attracting, and invariant (sufficiently) smooth manifold for $S(t)$.

Note that **invariance** means here that $S(t)\mathcal{M} \subset \mathcal{M}$ for $t \geq 0$, i.e., formally one should say **positively invariant**. Instead of tackling the inertial manifold problem directly, we are going to focus on local stable and unstable manifolds for a saddle point for the problem

$$\partial_t u + Bu = F(t, u) \tag{13.3}$$

for a nonlinearity F and a self-adjoint sectorial operator $B : H \to H$. One approach is to study the related inhomogeneous linear problem

$$\partial_t u + Bu = h(t) \tag{13.4}$$

for a given smooth forcing $h(t)$. Suppose H is split into an orthogonal sum

$$H = H_+ \oplus H_- \tag{13.5}$$

and denote the associated projections $P_+ : H \to H_+$ and $P_- : H \to H_-$. Furthermore, assume $B = \operatorname{diag}(B_+, B_-)$ is split accordingly where

$$\begin{aligned} \langle B_+ u, u\rangle_H &\leq -\theta \|u\|_H^2 \qquad \text{for } u \in H_+ \cap D(B), \\ \langle B_- u, u\rangle_H &\geq \theta \|u\|_H^2 \qquad \text{for } u \in H_- \cap D(B). \end{aligned} \tag{13.6}$$

Note that (13.4) is equivalent to

$$\begin{aligned} \partial_t u_+ + B_+ u_+ &= h_+(t), \\ \partial_t u_- + B_- u_- &= h_-(t), \end{aligned} \tag{13.7}$$

where $u_\pm = P_\pm u$ and $h_\pm = P_\pm h$; note that we have encountered a similar projection splitting already in Section 2. For $h \equiv 0$, it follows that (13.7) is solved by

$$u_+(t) = \mathrm{e}^{-tB_+} u_+(0), \qquad u_-(t) = \mathrm{e}^{-tB_-} u_-(0).$$

Due to (13.6) we may then conclude that

$$\begin{aligned} \|\mathrm{e}^{-tB_+}\|_{\mathcal{L}(H,H)} &\leq \mathrm{e}^{\theta t}, \qquad \text{for } t \leq 0, \\ \|\mathrm{e}^{-tB_-}\|_{\mathcal{L}(H,H)} &\leq \mathrm{e}^{-\theta t}, \quad \text{for } t \geq 0, \end{aligned} \tag{13.8}$$

which just means that $u \equiv 0$ is a saddle point for the linear system with unstable eigenspace H_+ and stable eigenspace H_-; see also Section 1 for the finite-dimensional saddle case and Section 8 for exponential dichotomies of ODEs. It is instructive to see that we may bound $u_\pm$ in the case nonzero forcing $h \neq 0$. Consider u_+ and observe that Theorem 11.11 implies

$$u_+(t) = \mathrm{e}^{(t_0 - t)B_+} u_+(t_0) + \int_{t_0}^{t} \mathrm{e}^{(s-t)B_+} h_+(s) \, \mathrm{d}s.$$

However, we know that e^{-tB_+} contracts on H_+ as $t \to -\infty$ so if we let $t_0 \to \infty$ we have

$$u_+(t) = -\int_{t}^{\infty} \mathrm{e}^{(s-t)B_+} h_+(s) \, \mathrm{d}s. \tag{13.9}$$

Suppose that $h \in C_\mathrm{b}(\mathbb{R}, H)$ so h is bounded and continuous with values in H. Then it is easy to see that

$$\|u_+\|_{C_\mathrm{b}(\mathbb{R},H)} \leq \frac{K}{\theta} \|h_+\|_{C_\mathrm{b}(\mathbb{R},H)}$$

for some constant K and for $\theta > 0$ as introduced above. A similar estimate also holds for u_-. Therefore, it follows that

$$\|u\|_{C_\mathrm{b}(\mathbb{R},H)} \leq \frac{K}{\theta} \|h\|_{C_\mathrm{b}(\mathbb{R},H)}$$

for some generic constant $K > 0$ independent of θ. It turns out that working in a different norm, and also working a bit more, we can explicitly determine a sharp value of K.

Lemma 13.2. *([Zel13, Lem.2.2]) Consider the inhomogeneous problem* (13.4)*, suppose $h \in L^2(\mathbb{R}, H)$ and $h \neq 0$. Then*

$$\|u\|_{L^2(\mathbb{R},H)} \leq \frac{1}{\theta}\|h\|_{L^2(\mathbb{R},H)},$$

where $\theta > 0$ is the constant from (13.6)*. In particular, the solution operator $R : L^2(\mathbb{R}, H) \to L^2(\mathbb{R}, H)$ of* (13.4) *has norm bounded by $1/\theta$.*

Remark: Note that here $\|u\|^2_{L^2(\mathbb{R},H)} := \int_{\mathbb{R}} \|u(t)\|^2_H \,\mathrm{d}t$. Furthermore, Lemma 13.2 can be used to obtain the estimate $\|u\|_{C_\mathrm{b}(\mathbb{R},H)} \leq K\|u\|_{L^2(\mathbb{R},H)}$ for some constant $K > 0$.

We return to the full problem (13.3) and assume that F is globally Lipschitz

$$\|F(t, u) - F(t, v)\|_H \leq \kappa\|u - v\|_H. \tag{13.10}$$

uniformly for $t \in \mathbb{R}$. Furthermore, we assume that the nonlinear problem has a hyperbolic saddle point at $u \equiv 0$, i.e., the linear operator B still satisfies (13.8) and

$$F(t, 0) \equiv 0.$$

The next step is to prove an analog of Theorem 1.13 showing persistence of the linear spaces $H_\pm$ as local stable and unstable manifolds. However, the construction turns out to be more technical than for the finite-dimensional case.

Definition 13.3. Fix $\tau \in \mathbb{R}$. The **unstable set** $\mathcal{M}_+(\tau) \subset H$ consists of all $u_\tau \in H$ such that there exists a backward trajectory $u(t)$ with $t \leq \tau$ with

$$u(\tau) = u_\tau, \qquad \|u\|_{L^2((-\infty,\tau],H)} < \infty.$$

Similarly, the **stable set** $\mathcal{M}_-(\tau) \subset H$ consists of all $u_\tau \in H$ such that there exists a forward trajectory $u(t)$ with $t \geq \tau$ with

$$u(\tau) = u_\tau, \qquad \|u\|_{L^2([\tau,+\infty),H)} < \infty.$$

Essentially, the sets $\mathcal{M}_+(\tau)$ and $\mathcal{M}_-(\tau)$ turn out to be the unstable and stable manifolds of the saddle point at $u \equiv 0$. The time τ is somewhat arbitrary and we shall mostly work with $\tau = 0$ to simplify the notation. However, we keep in mind that the following arguments work in more generality, in fact, even uniformly in τ. The key point is that the sets $\mathcal{M}_\pm = \mathcal{M}_\pm(0)$ turn out to be (Lipschitz) manifolds.

Theorem 13.4. *Suppose the* ***spectral gap condition***

$$\theta > \kappa \tag{13.11}$$

holds where θ controls the linear contraction/expansion rates in (13.8) *and κ is the Lipschitz constant of the nonlinearity F in* (13.10)*. Then $\mathcal{M}_\pm$ are Lipschitz manifolds, i.e.,*

$$\mathcal{M}_\pm = \{u_\pm + M_\pm(u_\pm),\ u_\pm \in H_\pm\},$$

where the maps $M_\pm : H_\pm \to H_\mp$ satisfy

$$\|M_\pm(v_1) - M_\pm(v_2)\|_{H_\mp} \leq K\|v_1 - v_2\|_{H_\pm}$$

for some constant $K > 0$.

Proof. (Sketch) We only consider $\mathcal{M}_+$ since the same arguments can be adapted to $\mathcal{M}_-$. Suppose we could verify that for every $u_0 \in H_+$ there exists a unique solution $u \in L^2((-\infty, 0], H)$ to

$$\partial_t u + Bu = F(t, u), \quad (P_+u)(0) = u_0, \tag{13.12}$$

and that u depends in a Lipschitz continuous way on u_0. In this scenario, one may simply define

$$M_+(u_0) := (P_-u)(0).$$

This mapping simply sends the element u_0 to the suitable stable set via solving (13.12); see Figure TODO. Hence it remains to solve (13.12). A natural idea is to measure the deviation from the linear problem and introduce for $t \leq 0$

$$w(t) := u(t) - v(t), \qquad v(t) := \mathrm{e}^{-tB_+}u_0.$$

In particular, w then solves

$$\partial_t w + Bw = F(t, w + \mathrm{e}^{-tB_+}u_0), \quad (P_+w)(0) = 0. \tag{13.13}$$

One may extend (13.13) to an equivalent equation for $t \in \mathbb{R}$ by defining

$$\tilde{F}(t, u_0, w) := \begin{cases} F(t, w + v(t)) & \text{for } t < 0, \\ 0 & \text{for } t \geq 0, \end{cases}$$

and consider the problem to find

$$\partial_t w + Bw = \tilde{F}(t, u_0, w), \qquad w \in L^2(\mathbb{R}, H), \tag{13.14}$$

see also Exercise 13.5. To solve (13.14) one considers the equivalent fixed-point problem

$$w = R \circ \tilde{F}(\cdot, u_0, w), \tag{13.15}$$

where R is the solution operator defined in Lemma 13.2. It is natural to try to solve (13.15) using the Banach fixed point theorem on the Banach space $L^2(\mathbb{R}, H)$. A key step is to derive the estimate

$$\begin{aligned}\|\tilde{F}(t, u_{0,1}, w_1) - \tilde{F}(t, u_{0,2}, w_2)\|_{L^2(\mathbb{R},H)} &\leq \kappa\|w_1 + v_1 - w_2 - v_2\|_{L^2((-\infty,0],H)} \\ &\leq \kappa\left(\|w_1 - w_2\|_{L^2(\mathbb{R},H)}\right. \qquad (13.16)\\ &\qquad\left. + \frac{1}{\theta}\|u_{0,1} - u_{0,2}\|_H\right),\end{aligned}$$

where Lipschitz continuity of F was used in the first inequality and the second inequality follows by using the fact

$$\|v\|_{L^2((-\infty,0],H)} \leq \frac{1}{\theta}\|u_0\|_H,$$

which follows from the definition of v and expansion properties of B_+ considered in (13.8). Since $F(t, 0) \equiv 0$, one observes by taking $w_2 = 0 = u_{0,2}$ in (13.16) that $\tilde{F}(\cdot, u_0, w) \in L^2(\mathbb{R}, H)$ if $u_0 \in H_+$ and $w \in L^2(\mathbb{R}, H)$. Therefore, (13.15) is a well-defined mapping on $L^2(\mathbb{R}, H)$. In addition, Lemma 13.2 yields that the operator norm of R is bounded by $1/\theta$ which implies in combination with (13.16) that the mapping (13.15) is a contraction if $\kappa/\theta < 1$. Since this is precisely the spectral gap condition (13.11), the existence of M_+ follows. The Lipschitz continuity is discussed in Exercise 13.6. □

Lastly, we note that a little bit of extra work shows that solutions in M_+ indeed decay in backward time

$$\|u(t)\|_H \leq K\mathrm{e}^{\delta t}\|u_0\|_H, \qquad t \leq 0,$$

for $|\delta| < \theta - \kappa$, $\delta > 0$. Hence, we could also formally write $\mathcal{M}_+ = W^{\mathrm{u}}(0)$ as the unstable manifold, and similarly for the stable manifold $\mathcal{M}_- = W^{\mathrm{s}}(0)$.

Exercise 13.5. Prove that solving (13.13) is equivalent to (13.14). ♢

Exercise 13.6. Show that the mapping M_+ is Lipschitz. ♢

Exercise 13.7. Construct a non-trivial (and nonlinear!) example for (13.1), which satisfies all the assumptions used in this chapter. ♢

Background and Further Reading: The main line of argument follows the lecture notes [Zel13]. There are many important resources developing invariant manifold theory for PDEs, see for example the material and references in [BJ89, BLZ98].

14 Spectral Gap and Inertial Manifolds

We continue the topic from Lecture 13 studying PDEs

$$\partial_t u + Au = f(u), \qquad u_0 \in H, \tag{14.1}$$

where $-A : H \to H$ is a linear, negative, self-adjoint, and sectorial operator on a Hilbert space H and f is a sufficiently smooth globally Lipschitz nonlinearity. Recall also that (14.1) generates a semiflow $S(t) : H \to H$ as defined in (13.2) and that we required that f is globally bounded in H. In this lecture, we want to establish the existence of an **inertial manifold** $\mathcal{M}$ for (14.1) to demonstrate that the dynamics is low-dimensional; see Definition 13.1.

By the Hilbert-Schmidt Theorem, it follows that A has a complete orthonormal system in H

$$Ae_n = \lambda_n e_n, \qquad 0 < \lambda_1 \le \lambda_2 \le \cdots$$

with eigenfunctions e_n and eigenvalues λ_n. Therefore, we have

$$v = \sum_{n=1}^{\infty} v_n e_n, \qquad v_n := \langle v, e_n \rangle_H$$

for every $v \in H$. A natural idea to construct a low-dimensional inertial manifold is simply to project functions onto the first few (Fourier) modes

$$P_N v := \sum_{n=1}^{N} v_n e_n,$$

and let $Q_N := \mathrm{Id} - P_N$ which leads to the linear spaces

$$H_+ := P_N H, \qquad H_- := Q_N H, \qquad H = H_- \oplus H_+.$$

As discussed in a similar setting in Lecture 13, it helps to present the PDE 14.1 as a system

$$\begin{aligned} \partial_t u_+ + Au_+ &= f_+(u_+ + u_-), & (14.2)\\ \partial_t u_- + Au_- &= f_-(u_+ + u_-), & (14.3) \end{aligned}$$

where $f_+ := P_N F$, $f_- := Q_N F$, $u_+ := P_N u$ and $u_- := Q_N u$. By construction we have

$$\begin{array}{ll} \langle Av, v \rangle_H \le \lambda_N \|v\|_H^2 & \text{for } v \in H_+ \cap D(A), \\ \langle Av, v \rangle_H \ge \lambda_{N+1} \|v\|_H^2 & \text{for } v \in H_- \cap D(A). \end{array} \tag{14.4}$$

The main idea is to construct a mapping $\Phi : H_+ \to H_-$ such that the inertial manifold is given by

$$\mathcal{M} := \{u_+ + \Phi(u_+),\ u_+ \in H_+\}, \qquad u_- = \Phi(u_+),$$

so $\mathcal{M}$ is parametrized by a finite-dimensional set of variables and the dynamics on $\mathcal{M}$ reduces to the ODEs

$$\partial_t u_+ + Au_+ = f_+(u_+ + \Phi(u_+)) \tag{14.5}$$

which is also called the **inertial form**.

Example 14.1. We may view the situation in analogy to multiple time-scale systems on $\mathbb{R}^N$. A **fast-slow system** is given by

$$\begin{aligned} \varepsilon\tfrac{\mathrm{d}x}{\mathrm{d}\tau} &= f(x,y,\varepsilon), \\ \tfrac{\mathrm{d}y}{\mathrm{d}\tau} &= g(x,y,\varepsilon), \end{aligned} \tag{14.6}$$

where $\tau \in \mathbb{R}$, $(x,y) \in \mathbb{R}^{m+n}$, the maps f, g are sufficiently smooth and $\varepsilon > 0$ is assumed to be small. The variables $x \in \mathbb{R}^m$ are **fast** while the variables $y \in \mathbb{R}^n$ are **slow**. The set

$$\mathcal{C}_0 := \{(x,y) \in \mathbb{R}^{m+n} : f(x,y,0) = 0\} \tag{14.7}$$

is called the **critical manifold**. $\mathcal{C}_0$ is called **normally hyperbolic** if the matrix

$$\mathrm{D}_x f(x,y,0)|_{(x,y)\in\mathcal{C}_0} \in \mathbb{R}^{m\times m}$$

has no eigenvalues with zero real parts. Then **Fenichel's Theorem** (see background below) states that for sufficiently small ε there exists a perturbed invariant manifold $\mathcal{C}_\varepsilon$, called a **slow manifold**. The manifold $\mathcal{C}_\varepsilon$ as well as the dynamics on $\mathcal{C}_\varepsilon$ converges to $\mathcal{C}_0$ as $\varepsilon \to 0$; see also Figure TODO. In particular, if $\mathcal{C}_\varepsilon = \{x = h_\varepsilon(y)\}$ for some map $h_\varepsilon : \mathbb{R}^m \to \mathbb{R}^n$ then the effective slow dynamics on $\mathcal{C}_\varepsilon$ is just given by

$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = g(h_\varepsilon(y), y).$$

Note the direct analogy to the PDE case discussed above, where we also required exponential attraction for the inertial manifold $\mathcal{M}$, which would correspond to requiring attraction of $\mathcal{C}_\varepsilon$ for $0 \le \varepsilon \ll 1$. Indeed, the map Φ above is the direct analog of the parametrization h_ε here. ⧫

The last example shows that it is helpful to think of $\mathcal{M}$ as an attracting slow manifold where the fast variables u_- decay very quickly and the effective long-term dynamics is given by the slow variables u_+.

Theorem 14.2. *Consider* (14.1) *under the assumptions stated above and suppose furthermore that for some N there exists a spectral gap*

$$\lambda_{N+1} - \lambda_N > 2\kappa \tag{14.8}$$

where $\kappa > 0$ is the Lipschitz constant of f. Then there exists an N-dimensional inertial manifold $\mathcal{M}$ defined via $\Phi : H_+ \to H_-$. In particular, for each u_0 there exists $v_0 \in \mathcal{M}$ such that

$$\|S(t)u_0 - S(t)v_0\|_H \leq K\mathrm{e}^{-\lambda_N t}\|u_0 - v_0\|_H$$

for some constant $K > 0$.

Proof. (Sketch) The idea is to use the previous work from Lecture 13 on invariant manifolds for nonlinear saddles and to obtain $\mathcal{M}$ as a certain unstable manifold. We define

$$B := A - \frac{\lambda_N + \lambda_{N+1}}{2}\mathrm{Id}$$

and observe that

$$\begin{aligned} \langle B_+ u, u\rangle_H &\leq -\theta\|u\|_H^2 \qquad \text{for } u \in H_+ \cap D(A), \\ \langle B_- u, u\rangle_H &\geq \theta\|u\|_H^2 \qquad \text{for } u \in H_- \cap D(A). \end{aligned} \tag{14.9}$$

holds due to (14.4) with $\theta = \frac{\lambda_{N+1} - \lambda_N}{2}$. For simplicity we assume $f(0) \equiv 0$. Define

$$\tilde{u}(t) := \mathrm{e}^{\alpha t}u(t), \qquad \alpha := \frac{\lambda_N + \lambda_{N+1}}{2}.$$

This transforms (14.1) to

$$\partial_t \tilde{u} + B\tilde{u} = F(t, \tilde{u}), \qquad F(t, \tilde{u}) := \mathrm{e}^{\alpha t} f(\mathrm{e}^{-\alpha t}\tilde{u}). \tag{14.10}$$

It can be checked that F is also Lipschitz continuous with the same Lipschitz constant $\kappa > 0$. We want to apply Theorem 13.4, which holds for equations of the form (14.10). The assumptions about A and the definition of B easily lead to the correct contraction and expansion rates (13.6) for B. The spectral gap condition in Theorem 13.4

$$\theta > \kappa$$

is immediately guaranteed by (14.8). Futhermore, $F(t, 0) \equiv 0$ since $f(0) \equiv 0$. Hence, the existence of an unstable manifold $\mathcal{M}_+$ at $\tilde{u} = 0$ for (14.10) follows. One may then check that $\mathcal{M}_+$ is indeed invariant under the semi-flow $S(t)$ of the original problem (14.1). It requires quite a bit of extra work to then also show that trajectories must track it exponentially in forward time; the idea is to use that $\mathrm{e}^{-\alpha t}\tilde{u}(t) = u(t)$ by construction so $\mathcal{M}_+ = \mathcal{M}$ is indeed the inertial manifold we wanted to construct. $\square$

Remark: Theorem 14.2 can be generalized in various directions. For example, the smoothness of $\mathcal{M}$ can be improved in many cases but usually invariant manifolds only have finite smoothness as this phenomenon already occurs for ODEs. There are also other common alternative proof techniques, e.g., a method based upon **invariant cones** and the **squeezing property**. See background references below for more details.

Example 14.3. A classical example for the existence of inertial manifolds are dissipative reaction-diffusion equations, such as

$$\partial_t u - \partial^2_{xx} u = f(u), \quad (x,t) \in \mathbb{R}^d \times [0,+\infty),\ u = u(x,t), \tag{14.11}$$

for a suitable sufficiently smooth nonlinearity $f : \mathbb{R} \to \mathbb{R}$ on a bounded interval $\Omega \subset \mathbb{R}^1$ with Dirichlet boundary conditions $u(x) = 0$ for $x \in \partial\Omega$. We have seen in Lecture 12 that under certain assumptions on the nonlinearity an absorbing set exists. In such a case, we may cut off the nonlinearity and assume that f is indeed globally Lipschitz. Since $A = -\Delta$, we know from Example 4.3 that the eigenvalues satisfy

$$\lambda_N = KN^2, \qquad \text{as } N \to +\infty, \tag{14.12}$$

for some constant $K > 0$ depending on the length of the interval; in fact, (14.12) is a version of **Weyl's law**

$$\lambda_N \sim KN^2, \qquad \text{as } N \to +\infty,$$

which holds for many other situations involving the Laplacian. We find from (14.12) that

$$\lambda_{N+1} - \lambda_N = K(N+1)^2 - KN^2 = K(2N+1) > 2\kappa$$

for some sufficiently large N. Therefore, the spectral gap condition (14.8) holds and Theorem 14.2 implies the existence of an inertial manifold. ♦

It should be noted that although it is theoretically very important to know that a certain PDE is effectively finite-dimensional, it is not always immediately useful in practical applications. For example, the dimension of the inertial form ODE system (14.5) could be extremely large or the mapping Φ is difficult to compute.

Example 14.4. Consider the **Kuramoto-Sivashinsky equation**

$$\partial_t u + \partial^4_{xxxx} u + 2p\partial^2_{xx} u = \partial_x(u^2) \tag{14.13}$$

with a parameter $p \in \mathbb{R}$ posed on an interval $\Omega = [0,\pi]$, $(x,t) \in \Omega \times [0,\infty)$, $u = u(x,t)$ and boundary conditions

$$u(x) = 0, \quad (\partial^2_{xx} u)(x) = 0, \quad \text{for } x \in \partial\Omega = \{0,\pi\}.$$

One may show that (14.13) is well-posed and dissipative on $H := L^2(\Omega)$ with an absorbing ball in a suitable space. However, Theorem 14.2 is not directly applicable as the nonlinearity $f(u) = \partial_x(u^2)$ is not a map from H to H. ♦

The last example shows the need to slightly generalize Theorem 14.2. Consider (14.1) on the Hilbert space H and define

$$H^s := D(A^s),\ s \in \mathbb{R}, \qquad \|v\|_{H^s}^2 = \sum_{n=1}^{\infty} \lambda_n^{2s} \langle v, e_n \rangle_H^2.$$

Note that for $s > 0$ the definition of H^s is just Definition 11.10 of fractional operator norms while for $s < 0$ one just takes H^s as the completion of H with respect to the norm $\|\cdot\|_{H^s}$; furthermore $H^0 = H$.

Remark: One also frequently finds the definition $H^s := D(A^{s/2})$ with $\|v\|_{H^s}^2 = \sum_{n=1}^{\infty} \lambda_n^s \langle v, e_n \rangle_H^2$ in the literature and one just has to keep track of the factor of 2 to match the different definitions.

Suppose now f is a Lipschitz map from H^{α_1} to H^{α_2}

$$\|f(u) - f(v)\|_{H^{\alpha_1}} \leq \kappa \|u - v\|_{H^{\alpha_2}} \tag{14.14}$$

for $\alpha_1 < \alpha_2$ and $u, v \in H^{\alpha_2}$.

Theorem 14.5. *Consider* (14.1) *with the assumptions stated above and suppose f satisfies* (14.14) *for $\alpha_2 = 0$ and some $\alpha_1 \in (-2, 0]$. Furthermore, suppose the spectral gap condition*

$$\frac{\lambda_{N+1} - \lambda_N}{\lambda_{N+1}^{-\alpha_1/2} + \lambda_N^{-\alpha_1/2}} > \kappa \tag{14.15}$$

holds, then there exists a Lipschitz inertial manifold.

The exercises give a guide how to apply the last result to the Kuramoto-Sivashinsky equation to prove the existence of an inertial manifold.

Exercise 14.6. Consider Example 14.4 and show that $A := \partial_{xxxx}^4 + 2p\partial_{xx}^2 + p^2 + 1$ is positive and self-adjoint on $D(A)$. Re-write (14.13) in the form (14.1) using A, i.e., compute f. ♢

Exercise 14.7. Continue with Exercise 14.6 and prove that f is Lipschitz with $\alpha_1 = -1/2$ and $\alpha_2 = 0$ once the nonlinearity f has been cut off properly. ♢

Exercise 14.8. Continue with Exercises 14.6-14.7 and check that A has eigenvalues $\lambda_N = (N^2 + p)^2 + 1$ and use this result to verify that the spectral gap condition (14.15) must hold for sufficiently large N. ♢

Background and Further Reading: We mainly followed the lecture notes [Zel13] in this section. Other important sources on inertial manifold theory are [Rob01, Tem97]. An overview of finite-dimensional multiple time scale dynamics can be found in [Kue15].

References

[AGJ90] J.C. Alexander, R.A. Gardner, and C.K.R.T. Jones. A topological invariant arising in the stability analysis of travelling waves. *J. Reine Angew. Math.*, 410:167–212, 1990.

[Ban07] R.E. Bank. *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations. Users' Guide 10.0.* Department of Mathematics, University of California at San Diego, 2007.

[BJ89] P.W. Bates and C.K.R.T. Jones. Invariant manifolds for semilinear partial differential equations. In U. Kirchgraber and H.O. Walther, editors, *Dynamics Reported*, volume 2, pages 1–37. Wiley, 1989.

[BLZ98] P.W. Bates, K. Lu, and C. Zeng. Existence and persistence of invariant manifolds for semiflows in Banach spaces. *Mem. Amer. Math. Soc.*, 135, 1998.

[BO99] C.M. Bender and S.A. Orszag. *Asymptotic Methods and Perturbation Theory.* Springer, 1999.

[BV92] A.V. Babin and M.I. Vishik. *Attractors of Evolution Equations.* North-Holland, 1992.

[CE90] P. Collet and J.P. Eckmann. The time dependent amplitude equation for the Swift-Hohenberg problem. *Comm. Math. Phys.*, 132(1):139–153, 1990.

[CG09] M. Cross and H. Greenside. *Pattern Formation and Dynamics in Nonequilibrium Systems.* CUP, 2009.

[CH82] S.-N. Chow and J.K. Hale. *Methods of Bifurcation Theory.* Springer, 1982.

[CH93] M.C. Cross and P.C. Hohenberg. Pattern formation outside of equilibrium. *Rev. Mod. Phys.*, 65(3):851–1112, 1993.

[Che97] X. Chen. Existence, uniqueness, and asymptotic stability of travelling waves in nonlocal evolution equations. *Adv. Differential Equations*, 2:125–160, 1997.

[CMKW14] J. Cuevas-Maraver, P.G. Kevrekidis, and F. Williams, editors. *The Sine-Gordon Model and Its Applications: From Pendula and Josephson Junctions to Gravity and High-energy Physics.* Springer, 2014.

[Cop78] W.A. Coppel. *Dichotomies in Stability Theory.* Springer, 1978.

[CR71] M.G. Crandall and P.H. Rabinowitz. Bifurcation from simple eigenvalues. *J. Functional Analysis*, 8(2):321–340, 1971.

[DCD+07] E.J. Doedel, A. Champneys, F. Dercole, T. Fairgrieve, Y. Kuznetsov, B. Oldeman, R. Paffenroth, B. Sandstede,

X. Wang, and C. Zhang. Auto 2007p: Continuation and bifurcation software for ordinary differential equations (with homcont). *http://cmvl.cs.concordia.ca/auto*, 2007.

[Dei10] K. Deimling. *Nonlinear Functional Analysis.* Dover, Mineola, NY, 2010.

[DG95] C.R. Doering and J.D. Gibbon. *Applied Analysis of the Navier-Stokes Equations.* CUP, 1995.

[DL00] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 3 Spectral Theory and Applications.* Springer, 2000.

[EN00] K.-J. Engel and R. Nagel. *Semigroups for Linear Evolution Equations.* Springer, 2000.

[Eva72] J. Evans. Nerve axon equations III: stability of nerve impulses. *Indiana U. Math. J.*, 22:577–594, 1972.

[Eva02] L.C. Evans. *Partial Differential Equations.* AMS, 2002.

[Gam01] T.W. Gamelin. *Complex Analysis.* Springer, 2001.

[GK09] J. Guckenheimer and C. Kuehn. Homoclinic orbits of the FitzHugh-Nagumo equation: The singular limit. *DCDS-S*, 2(4):851–872, 2009.

[Hen81] D. Henry. *Geometric Theory of Semilinear Parabolic Equations.* Springer, Berlin Heidelberg, Germany, 1981.

[Hoy06] R. Hoyle. *Pattern Formation: An Introduction to Methods.* Cambridge University Press, 2006.

[HSD03] M.W. Hirsch, S. Smale, and R. Devaney. *Differential Equations, Dynamical Systems, and an Introduction to Chaos.* Academic Press, 2nd edition, 2003.

[IJ97] G. Iooss and D.D. Joseph. *Elementary stability and bifurcation theory.* Springer, 1997.

[JKL01] C.K.R.T. Jones, N. Kopell, and R. Langer. Construction of the FitzHugh-Nagumo pulse using differential forms. In *Multiple-Time-Scale Dynamical Systems*, pages 101–113. Springer, 2001.

[Jon84] C.K.R.T. Jones. Stability of the travelling wave solution of the FitzHugh-Nagumo system. *Trans. Amer. Math. Soc.*, 286(2):431–469, 1984.

[Kap05] T. Kapitula. Stability analysis of pulses via the Evans function: dissipative systems. In *Dissipative Solitons*, volume 661 of *Lecture Notes in Physics*, pages 407–427. Springer, 2005.

[Kat80] T. Kato. *Perturbation Theory for Linear Operators.* Springer, 1980.

[KH95] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems.* CUP, 1995.

[Kie04] H. Kielhoefer. *Bifurcation Theory: An Introduction with Applications to PDEs.* Springer, 2004.

[KP13] T. Kapitula and K. Promislow. *Spectral and Dynamical Stability of Nonlinear Waves.* Springer, 2013.

[KR08] B. Krauskopf and T. Riess. A Lin's method approach to finding and continuing heteroclinic connections involving periodic orbits. *Nonlinearity*, 21(8):1655–1690, 2008.

[KSM92] P. Kirrmann, G. Schneider, and A. Mielke. The validity of modulation equations for extended systems with cubic nonlinearities. *Proc. R. Soc. Edinburgh A*, 122(1):85–91, 1992.

[Kue15] C. Kuehn. *Multiple Time Scale Dynamics.* Springer, 2015.

[Kuz04] Yu.A. Kuznetsov. *Elements of Applied Bifurcation Theory.* Springer, New York, NY, 3rd edition, 2004.

[Lan13] D. Lannes. *The Water Waves Problem: Mathematical Analysis and Asymptotics.* AMS, 2013.

[Lau12] R.S. Laugesen. *Spectral Theory of Partial Differential Equations.* arXiv:1203.2344v1, 2012.

[Mil06] P.D. Miller. *Applied Asymptotic Analysis.* AMS, 2006.

[MM02] K. Mischaikow and M. Mrozek. Conley Index Theory. In B. Fiedler, editor, *Handbook of Dynamical Systems*, volume 2, pages 393–460. North-Holland, 2002.

[Mur02] J.D. Murray. *Mathematical Biology I: An Introduction.* Springer, 3rd edition, 2002.

[Pal84] K.J. Palmer. Exponential dichotomies and transversal homoclinic points. *J. Differential Equat.*, 55:225–256, 1984.

[Pal88] K.J. Palmer. Exponential dichotomies and Fredholm operators. *Proc. Amer. Math. Soc.*, 104:149–156, 1988.

[Paz83] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, New York, 1983.

[Pis06] L.M. Pismen. *Patterns and Interfaces in Dissipative Dynamics.* Springer, 2006.

[Rob01] J.C. Robinson. *Infinite-Dimensional Dynamical Systems.* CUP, 2001.

[Rud76] W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, 1976.

[Rud91] W. Rudin. *Functional Analysis.* McGraw-Hill, 1991.

[San01] B. Sandstede. Stability of travelling waves. In B. Fiedler, editor, *Handbook of Dynamical Systems*, volume 2, pages 983–1055. Elsevier, 2001.

[Smo94] J. Smoller. *Shock Waves and Reaction-Diffusion Equations.* Springer, 1994.

[SS04] B. Sandstede and A. Scheel. Evans function and blow-up methods in critical eigenvalue problems. *Discr. Cont. Dyn. Syst.*, 10:941–964, 2004.

[Str00] S.H. Strogatz. *Nonlinear Dynamics and Chaos.* Westview Press, 2000.

[Tem97] R. Temam. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics.* Springer, 1997.

[UWR14] H. Uecker, D. Wetzel, and J.D.M. Rademacher. pde2path - A Matlab package for continuation and bifurcation in 2D elliptic systems. *Num. Math.: Th. Meth. Appl.*, 7:58–106, 2014.

[vH91] A. van Harten. On the validity of the Ginzburg-Landau equation. *J. Nonlinear Sci.*, 1(4):397–422, 1991.

[vS03] W. van Saarloos. Front propagation into unstable states. *Physics Reports*, 386:29–222, 2003.

[VVV94] A.I. Volpert, V. Volpert, and V.A. Volpert. *Traveling Wave Solutions of Parabolic Systems.* Amer. Math. Soc., 1994.

[Wei87] J. Weidmann. *Spectral Theory of Ordinary Differential Equations.* Springer, 1987.

[Zel13] S. Zelik. Inertial manifolds and finite-dimensional reduction for dissipative PDEs. *arXiv:1303.4457v1*, pages 1–60, 2013.

Index