Wave Solutions of Evolution Equations

Wolf-Jürgen Beyn[1]
Department of Mathematics
Bielefeld University
33501 Bielefeld
Germany

Denny Otten[2]
Department of Mathematics
Bielefeld University
33501 Bielefeld
Germany

Date: December 20, 2016

Contents

[1]e-mail: beyn@math.uni-bielefeld.de, phone: +49 (0)521 106 4798,
fax: +49 (0)521 106 6498, homepage: `http://www.math.uni-bielefeld.de/~beyn/AG_Numerik/`.
[2]e-mail: dotten@math.uni-bielefeld.de, phone: +49 (0)521 106 4784,
fax: +49 (0)521 106 6498, homepage: `http://www.math.uni-bielefeld.de/~dotten/`.

1. Introduction

Nonlinear waves are a common feature in many applications such as the spread of epidemics, electric signalling in nerve cells and excitable chemical reactions. For example, one of the most common references in Mathematical Biology [22][1] devotes more than 100 pages to explaining and analyzing waves in biological systems.
We emphasize that our main interest is not in classical applications of wave phenomena in evolution equations, such as water waves, shock waves in gas dynamics, and electromagnetic waves, see the reference [31][2]. Rather we concentrate on waves in nonlinear parabolic systems which arise when modelling reaction diffusion systems. One important feature of these systems is that waves have a specific velocity, as opposed to the continuum of waves with different wave lengths and different wave speeds that typical occur in classical wave equations.

[1]James D. Murray, born in Scotland 1931, Professor Emeritus of the University of Oxford, known for his work and his books on Mathematical Biology

[2]Gerald Witham 1927-2014, a British-born American applied mathematician, worked at the California Institute of Technology, known for his work in fluid dynamics and waves

One of the most important applications where this unique speed is essential, is the famous and Nobel-prize winning Hodgkin-Huxley model for conduction and excitation in nerve [15].
Our main topic will be travelling waves in one space dimension and the following problems will provide the guidelines for this lecture:

- Compute explicit travelling wave solutions for specific PDEs from applications.
- Detect travelling waves of PDEs in one space dimension as connecting orbits of ODEs und use phase plane analysis.
- Prove the existence of a travelling wave (and uniqueness of its speed if possible).
- Study the (asymptotic) stability of travelling waves for the underlying evolution equation and investigate its relation to the spectra of linearizations.
- Discuss numerical methods for computing travelling waves and analyze the error arising from the truncation of an unbounded to a bounded domain.
- Study families of travelling waves in systems with conserved quantities, in particular in Hamiltonian PDEs.

In the first chapter we will discuss some formulae of explicitly known wave solutions (see [24]), but also briefly touch upon waves in several space dimensions and show some examples and numerical simulations. However, a rigorous mathematical theory of such dynamic patterns is far from being complete and is well beyond the scope of this lecture.

2. Examples and basic principles

2.1. Travelling wave solutions. Consider a general evolution equation in one space dimension

$$u_t = F(u), \quad u(x,t) \in \mathbb{R}^m, x \in \mathbb{R}, t \geq 0, \tag{2.1}$$

where we think of F being a linear or nonlinear differential operator in $\frac{d}{dx} = \partial_x$.

Definition 2.1. *A solution*

$$u_\star : \mathbb{R} \times [0,\infty) \to \mathbb{R}^m, \quad (x,t) \mapsto u(x,t)$$

of equation (2.1) *of the form*

$$u_\star(x,t) = v_\star(x - c_\star t), x \in \mathbb{R}, t \geq 0 \tag{2.2}$$

is called a travelling wave *with* profile $v_\star : \mathbb{R} \mapsto \mathbb{R}^m$ *and* speed $c_\star \in \mathbb{R}$. *If additionally the limits*

$$\lim_{\xi \to -\infty} v_\star(\xi) = v_- \in \mathbb{R}^m, \quad \lim_{\xi \to \infty} v_\star(\xi) = v_+ \in \mathbb{R}^m \tag{2.3}$$

exist then $u_\star$ is called a travelling front *if $v_- \neq v_+$, and a* travelling pulse *(or a* solitary wave*) if $v_- = v_+$.*

In Definition 2.1 we intentionally left the notion of a solution imprecise and we did also not specify the form of the (nonlinear) differential operator F. This will become clearer with the following examples. Generally, we assume a travelling wave to be continuously differentiable and bounded, i.e. the function $v_\star$ is continuously differentiable and bounded. Figure 2.1 shows some wave profiles $v_\star$ of travelling waves, namely a front in (a), a pulse in (b) and a wavetrain in (c).

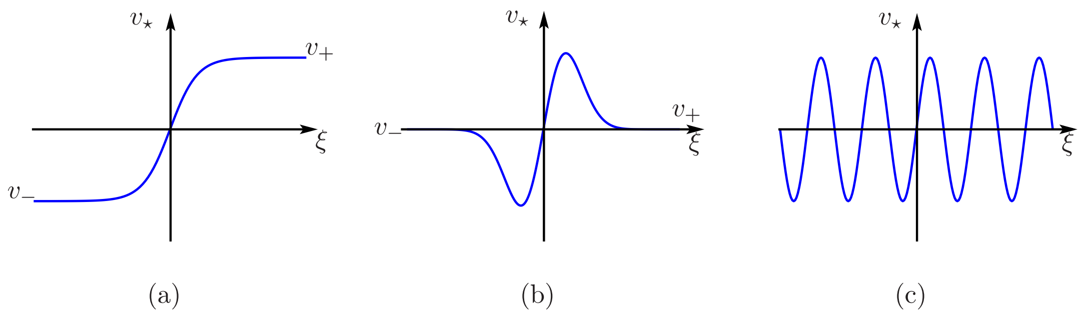

FIGURE 2.1. Profile $v_\star$: front (a), pulse (b), and wavetrain (c).

Finding travelling waves in one space dimension can be reduced to searching for special solutions of ODEs. For that purpose assume (2.1) to have the more specific form

$$u_t = f(u, \partial_x u, \ldots, \partial_x^k u), \quad x \in \mathbb{R}, t \geq 0, \tag{2.4}$$

where $f : \mathbb{R}^{(k+1)m} \to \mathbb{R}^m$ is a given nonlinearity. We insert the *travelling wave ansatz*

$$u(x,t) = v(x - ct), \quad x \in \mathbb{R}, t \geq 0 \tag{2.5}$$

into (2.4) and find

$$-cv'(x-ct) = f(v(x-ct), v'(x-ct), \ldots, v^{(k)}(x-ct)), \quad x \in \mathbb{R}, t \geq 0.$$

Since this should hold for all arguments we can substitute the *wave variable* $\xi = x - ct$ and obtain the equation

$$0 = cv'(\xi) + f(v(\xi), v'(\xi), \ldots, v^{(k)}(\xi)), \quad \xi \in \mathbb{R},$$

which is a k-th order ODE on the real line. We usually suppress arguments and write this *travelling wave ODE* as

$$0 = cv' + f(v, v', \ldots, v^{(k)}), \quad \xi \in \mathbb{R}. \tag{TWODE}$$

Note that the final step of reducing to an autonomous ODEs does not work if the function f in (2.4) depends explicitly on x or t

$$u_t = f(x, t, u, \partial_x u, \ldots, \partial_x^k u), \quad x \in \mathbb{R}, t \geq 0.$$

We will call (TWODE) the *travelling wave ODE* for the given PDE (2.4).

2.2. Examples from linear PDEs.

Example 2.2 (The advection equation)**.** Consider the *advection equation* (or *linear transport equation*)

$$u_t + au_x = 0, \quad x \in \mathbb{R}, t \geq 0 \tag{2.6}$$

for some $0 \neq a \in \mathbb{R}$. The travelling wave ODE of (2.6) is

$$0 = cv' - av', \quad \xi \in \mathbb{R}.$$

For nonconstant v we have $v' \neq 0$ which implies $c = a$. Therefore, any function $u_\star(x,t) = v_\star(x - at)$ with sufficiently smooth profile $v_\star$ (i.e. $v_\star \in C^1(\mathbb{R},\mathbb{R})$) is a travelling wave solution of (2.6) with speed $c_\star = a$. In fact, the associated Cauchy problem of (2.6)

$$\begin{aligned} u_t + au_x &= 0, \qquad x \in \mathbb{R}, t \geq 0, \\ u(x,0) &= u_0(x), \quad x \in \mathbb{R}, \end{aligned} \tag{2.7}$$

admits the solution $u(x,t) = u_0(x - at)$. That is all solutions of this simple equation are travelling waves (which, in a sense, is boring). For constant initial data, the solution is the same, but the speed is arbitrary. We consider this to be a *degenerate* situation since constants are steady states of the evolution equation (2.6) (i.e. $u_x = 0$) while nonconstant profiles do not have this property. We sometimes call the latter ones *true* (or *nontrivial*) *travelling waves*. A travelling pulse solution of (2.6) is shown in Figure 2.2.

Example 2.3 (The heat equation)**.** Consider the *heat equation* (or *diffusion equation*)

$$u_t = au_{xx}, \quad x \in \mathbb{R}, t \geq 0, \tag{2.8}$$

for some $0 < a \in \mathbb{R}$. The travelling wave ODE of (2.8) is

$$0 = cv' + av'', \quad \xi \in \mathbb{R},$$

which has the two linearly independent solutions

$$v_1(\xi) = 1, \quad v_2(\xi) = e^{-\frac{c}{a}\xi}, \quad \xi \in \mathbb{R}. \tag{2.9}$$

Both are either constant or unbounded. We conclude that the heat equation has no true bounded travelling wave solutions. Later on, we will learn that this changes dramatically when we introduce nonlinearities into the heat equation.

Example 2.4 (The Klein Gordon equation)**.** Consider the *Klein Gordon equation*

$$u_{tt} = a^2 u_{xx} - \mu^2 u, \quad x \in \mathbb{R}, t \geq 0 \tag{2.10}$$

for some $0 \neq a, \mu \in \mathbb{R}$. It is an easy exercise to show that the scaling

$$\tilde{u}(y,s) = u(\frac{a}{\mu}y, \frac{1}{\mu}s)$$

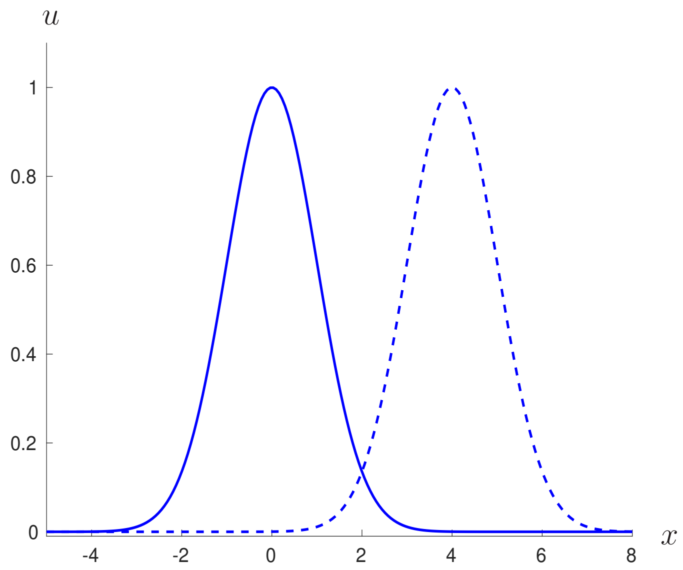

FIGURE 2.2. Travelling pulse solution $u_\star(x,t) = v_\star(x - c_\star t)$ of the advection equation (2.6) with profile $v_\star(\xi) = \exp(-\frac{\xi^2}{2})$ and velocity $c_\star = a = 1$ at time $t = 0$ (solid line) and $t = 4$ (dashed line).

transforms equation (2.10) into

$$u_{tt} = u_{xx} - u, \tag{2.11}$$

where the tilde was dropped and (y, s) was renamed as (x, t) for convenience. Applying the travelling wave ansatz (2.5) to (2.11) leads to the travelling wave ODE of (2.11)

$$0 = (c^2 - 1)v'' + v, \quad \xi \in \mathbb{R}, \tag{2.12}$$

which is now a second order ODE. Note that second order time derivative in (2.11) always imply a second order derivative in the associated traveling wave ODE (2.12). The characteristic polynomial of (2.12)

$$p(\lambda) = (c^2 - 1)\lambda^2 + 1.$$

Solving (2.12) is equivalent to finding zeros of p, i.e. λ with $p(\lambda) = 0$, which are

$$\lambda = \begin{cases} \pm\frac{1}{\sqrt{1-c^2}}, & |c| \neq 1 \\ \text{no zeros}, & |c| = 1 \end{cases} = \begin{cases} \pm\frac{1}{\sqrt{|c^2-1|}}, & |c| < 1 \\ \text{no zeros}, & |c| = 1 \\ \pm i\frac{1}{\sqrt{|c^2-1|}}, & |c| > 1 \end{cases}.$$

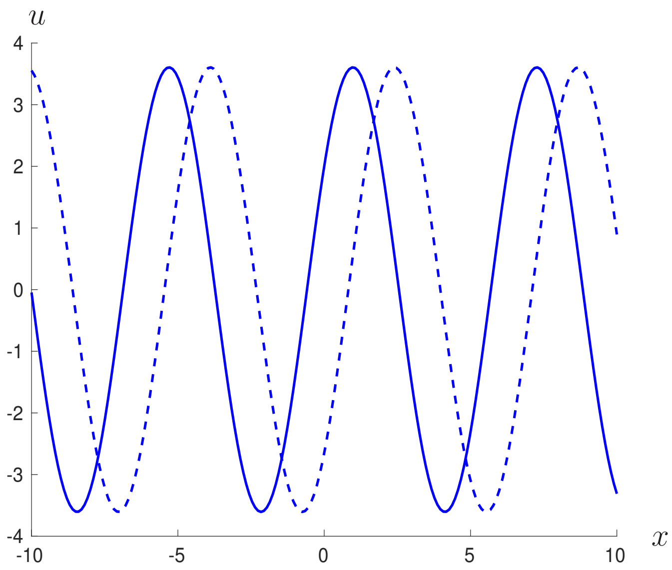

FIGURE 2.3. Travelling wavetrain solution $u_\star(x,t) = v_\star(x - c_\star t)$ of the Klein Gordon equation (2.10) for $k = 1$, $\alpha_1 = 2$, $\alpha_2 = 3$, $c = \frac{\omega(k)}{k} = \sqrt{2}$ at time $t = 0$ (solid line) and $t = 1$ (dashed line).

If $|c| = 1$ the equation $p(\lambda) = 0$, hence also the equation (2.12), has no solutions. If $|c| \neq 1$ the fundamental solutions of the linear equation (2.12) are

$$v_1(\xi) = \begin{cases} \cosh(k\xi), & |c| < 1 \\ \cos(k\xi), & |c| > 1 \end{cases}, \qquad v_2(\xi) = \begin{cases} \sinh(k\xi), & |c| < 1 \\ \sin(k\xi), & |c| > 1 \end{cases},$$

where the quantity k denotes the *angular wave number* defined by

$$k = \frac{1}{\sqrt{|c^2 - 1|}}. \tag{2.13}$$

In case $|c| < 1$ there is no bounded solution whereas in case $|c| > 1$ we have the linear family of bounded solutions

$$v(\xi) = \alpha_1 \cos(k\xi) + \alpha_2 \sin(k\xi), \quad \alpha_1, \alpha_2 \in \mathbb{R}. \tag{2.14}$$

Defining the *angular frequency* $\omega(k) := ck$, we obtain the family of travelling waves

$$u(x,t) = \alpha_1 \cos(kx - \omega(k)t) + \alpha_2 \sin(kx - \omega(k)t), \quad x, t \in \mathbb{R}, \quad \alpha_1, \alpha_2 \in \mathbb{R}. \tag{2.15}$$

The definition of $\omega(k)$ and equation (2.13) imply the relation

$$\omega(k)^2 = k^2 + 1, \tag{2.16}$$

which is called the *dispersion relation*. This relation shows how the speed c of the wave is related to its frequency $\omega(k)$. In fact, multiplying (2.13) by the velocity c

shows that (2.16) is equivalent to

$$\omega(k) = \frac{c}{\sqrt{|c^2 - 1|}}.$$

For completeness, recall the following relations from the theory of wave equations

$$k = \frac{\omega}{c} = \frac{2\pi}{\lambda} = \frac{2\pi\nu}{c},$$

where k is the *angular wave number* (or: *magnitude of the wave vector*), ν the *frequency* of the wave, λ the *wavelength*, $\omega = 2\pi\nu$ the *angular frequency* of the wave and c the *phase velocity*.
Note that the travelling waves (2.15) are neither fronts nor pulses but have an oscillatory character. For this reason they are called *wave trains*. Contrary to Example 2.2, the wave trains (2.15) do not have a fixed speed but can travel at all speeds $c > 1$ and $c < -1$.

Example 2.5 (Symmetric hyperbolic systems)**.** A direct generalization of the advection equation (2.6) is the first order system

$$u_t + Au_x = 0, \quad u : \mathbb{R} \times [0, \infty) \mapsto \mathbb{R}^m, \quad A \in \mathbb{R}^{m,m}. \tag{2.17}$$

The travelling wave ODE reads $(A - cI_m)v' = 0$. Every real eigenvalue with real eigenvector

$$Aw = \lambda w, \quad \lambda \in \mathbb{R}, w \in \mathbb{R}^m \tag{2.18}$$

leads to a solution

$$v(\xi) = \alpha(\xi)w, \tag{2.19}$$

where $\alpha : \mathbb{R} \mapsto \mathbb{R}$ is any bounded sufficiently smooth function. This corresponds to the travelling wave solution

$$u(x, t) = \alpha(x - \lambda t)w. \tag{2.20}$$

If A is real diagonalizable, then it has only real eigenvalues λ_j with linearly independent eigenvectors $w_j, j = 1, \ldots, m$. Correspondingly, equation (2.17) has solutions which are superpositions of waves travelling with speed λ_j in direction w_j

$$u(x, t) = \sum_{j=1}^{m} \alpha_j(x - \lambda_j t)w_j, \tag{2.21}$$

where the α_j are smooth bounded scalar functions. Given an initial condition as in (2.7) with a vector valued function u_0, the solution of the Cauchy problem (2.17),(2.7) is given by (2.21), where the functions α_j are determined from the decomposition of initial data

$$u_0(x) = \sum_{j=1}^{m} \alpha_j(x)w_j, \quad x \in \mathbb{R}.$$

We conclude that all solutions of the first order system (2.17) are linear superpositions of travelling waves with different speeds. The system (2.17) is called symmetric because we can transform (2.17) via

$$u(x,t) = Wv(x,t), \quad W = \begin{pmatrix} w_1 & w_2 & \cdots & w_m \end{pmatrix} \in \mathbb{R}^{m,m}$$

into the diagonal system

$$v_t + \mathrm{diag}(\lambda_j, j = 1, \dots, m)\, v_x = 0,$$

which consists of m copies of the advection equation (2.6) with different propagation speeds.

2.3. Travelling waves in nonlinear parabolic PDEs. We are looking for travelling waves in a nonlinear system of equations

$$u_t = Au_{xx} + f(u), \quad x \in \mathbb{R}, t \geq 0, \tag{2.22}$$

where $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$ and $A \in \mathbb{R}^{m,m}$ is invertible. By (TWODE) the travelling wave ODE is

$$0 = Av'' + cv' + f(v). \tag{2.23}$$

We look for solutions $v \in C^2(\mathbb{R}, \mathbb{R}^m)$ of this system such that the limits

$$v_\pm = \lim_{\xi \to \pm\infty} v(\xi), \quad v'_\pm = \lim_{\xi \to \pm\infty} v'(\xi)$$

exist. As the next Lemma shows, this necessarily implies

$$f(v_\pm) = 0, \quad \lim_{\xi \to \pm\infty} v'(\xi) = 0. \tag{2.24}$$

Lemma 2.6. *With the setting $w = \begin{pmatrix} v \\ v' \end{pmatrix}$ equation (2.23) is equivalent to the first order system*

$$w' = \begin{pmatrix} w_2 \\ -cA^{-1}w_2 - A^{-1}f(w_1) \end{pmatrix} =: G(w). \tag{2.25}$$

Any solution of (2.23) on $[0, \infty)$ for which the limits

$$\lim_{\xi \to \infty} v(\xi) =: v_+, \quad \lim_{\xi \to \infty} v'(\xi) =: v'_+$$

exist, satisfy

$$f(v_+) = 0, \quad v'_+ = 0. \tag{2.26}$$

The same statement holds if the limits $\xi \to -\infty$ exist.

Proof. The transformation to a first order system is standard. From $w(\xi) \to w_+ := \begin{pmatrix} v_+ \\ v'_+ \end{pmatrix}$ as $\xi \to \infty$ we obtain for all $x \geq 0$

$$|G(w_+)| = \left| \int_x^{x+1} G(w_+) - G(w(\xi)) + w'(\xi) d\xi \right|$$

$$\leq \int_x^{x+1} |G(w_+) - G(w(\xi))| d\xi + |w(x+1) - w(x)|$$
$$\leq \sup_{\xi \geq x} |G(w_+) - G(w(\xi))| + |w(x+1) - w_+| + |w_+ - w(x)|.$$

By our assumption and the continuity of G the right hand sides converge to 0 as $x \to \infty$. Hence, w_+ satisfies $G(w_+) = 0$ from which (2.26) follows. □

Remark 2.7. *If all eigenvalues of A have positive real part (i.e. equation (2.22) is parabolic) then one can omit the assumption that the derivative $v'(\xi)$ converges as $\xi \to \pm\infty$. This will be proved in Lemma 2.23 in Section 2.6.*

Another simple observation is the following reflection symmetry.

Lemma 2.8. *If $v_\star, c_\star$ is a travelling wave of the system (2.22) then so is*

$$v^\star(\xi) = v_\star(-\xi), \quad c^\star = -c_\star. \tag{2.27}$$

In the following we restrict to scalar equation with $A = 1$

$$u_t = u_{xx} + f(u), \quad x \in \mathbb{R}, t \geq 0, \tag{2.28}$$

with travelling wave ODE

$$0 = v'' + cv' + f(v). \tag{2.29}$$

According to Lemma 2.6 the potential limits of travelling waves must be zeroes of f. In the following we are particularly interested in the case where f has three zeroes

$$b_1 < b_2 < b_3 \quad \text{and } f(v) \begin{cases} > 0, & v < b_1,\ b_2 < v < b_3 \\ < 0, & b_1 < v < b_2,\ b_3 < v \end{cases}. \tag{2.30}$$

Our model example is the following *Nagumo equation*

Example 2.9 (Nagumo equation)**.** Consider the scalar parabolic equation

$$u_t = u_{xx} + u(1-u)(u-b),\ x \in \mathbb{R},\ t \geq 0, \tag{2.31}$$

where $0 < b < 1$, [21], [22]. It is well known that (2.31) has an explicit travelling front solution $u_\star(x,t) = v_\star(x - c_\star t)$ (called the *Huxley wave*) given by

$$v_\star(\xi) = \frac{1}{1 + \exp\left(-\frac{\xi}{\sqrt{2}}\right)}, \quad c_\star = \sqrt{2}\left(b - \frac{1}{2}\right), \tag{2.32}$$

with asymptotic states $v_- = 0$ and $v_+ = 1$. Note that $c_\star < 0$ if $b < \frac{1}{2}$ and $c_\star > 0$ if $b > \frac{1}{2}$.

Demo: Phase plane analysis of (2.25),(2.29)

In Mathematical Biology this equation is motivated by population models which have three equilibria as in (2.30) in the spatially independent case. An example of this is the spruce budworm model described in [22, Ch.1.2, Ch.11.5]

$$u_t = ru\left(1-\frac{u}{q}\right) - \frac{u^2}{1+u^2}.$$

If one adds spatial spread of populations via diffusion to such a law then a parabolic equation of the type (2.28) arises with the nonlinearity satisfying (2.30).
The following exercise shows that systems with cubic nonlinearities which have 3 consecutive zeros and behave like $-u^3$ (rather than u^3) can always be transformed into the Nagumo equation (2.31).

Exercise 2.10. Show that the general equation

$$u_t = Du_{xx} + B(u-b_1)(b_2-u)(u-b_3), \quad x \in \mathbb{R}, t \geq 0 \tag{2.33}$$

where $D, B > 0$ can be reduced to the Nagumo equation (2.31) via linear transformations $u = \beta_1\tilde{u} + \beta_2$ and scalings of time and space $u(x,t) = \tilde{u}(\alpha_1 x, \alpha_2 t)$. Determine the profile and speed of a travelling wave for (2.33).
Solution: The linear transformation $u = b_1 + \tilde{u}(b_3 - b_1)$ shifts the zeroes b_1, b_2, b_3 to $0, b = \frac{b_2-b_1}{b_3-b_1}, 1$ and gives the equation

$$\begin{aligned} u_t =& (b_3-b_1)\tilde{u}_t \\ =& D(b_3-b_1)\tilde{u}_{xx} + B(b_3-b_1)\tilde{u}(b_2-b_1-(b_3-b_1)\tilde{u})(b_3-b_1)(\tilde{u}-1), \\ \tilde{u}_t =& D\tilde{u}_{xx} + B(b_3-b_1)^2\tilde{u}(b-\tilde{u})(\tilde{u}-1). \end{aligned}$$

Now perform a scaling $\tilde{u}(x,t) = \tilde{\tilde{u}}(\alpha_1 x, \alpha_2 t)$ and obtain

$$\tilde{u}_t = \alpha_2\tilde{\tilde{u}}_t = \alpha_1^2 D\tilde{\tilde{u}}_{xx} + B(b_3-b_1)^2\tilde{\tilde{u}}(b-\tilde{\tilde{u}})(\tilde{\tilde{u}}-1).$$

This is equivalent to (2.31) if we set

$$\alpha_2 = B(b_3-b_1)^2, \quad \alpha_1 = \sqrt{\frac{\alpha_2}{D}} = (b_3-b_1)\sqrt{\frac{B}{D}}.$$

Suppose $\tilde{\tilde{u}}(x,t) = v_\star(x - c_\star t)$ is travelling wave of (2.31) then we have the following solution of (2.33)

$$u(x,t) = b_1 + (b_3-b_1)v_\star(\alpha_1 x - c_\star\alpha_2 t).$$

This is a travelling wave with speed

$$\tilde{c} = c_\star\frac{\alpha_2}{\alpha_1} = c_\star\sqrt{DB}(b_3-b_1) \tag{2.34}$$

and profile

$$\tilde{v}(\xi) = b_1 + (b_3-b_1)v_\star(\alpha_1\xi), \quad \alpha_1 = (b_3-b_1)\sqrt{\frac{B}{D}}. \tag{2.35}$$

With the values of b and $c_\star$ from (2.32) we obtain the final expression for the speed of the wave that belongs to (2.33)

$$\tilde{c} = \sqrt{2}\left(\frac{b_2-b_1}{b_3-b_1} - \frac{1}{2}\right)(b_3-b_1)\sqrt{DB} = \sqrt{\frac{DB}{2}}(-b_1+2b_2-b_3). \tag{2.36}$$

Although it is easy to verify the formula (2.32) for the travelling wave of the Nagumo equation (2.31), we discuss a more general approach that allows to arrive at such an explicit expression. The travelling wave ODE for (2.31) is

$$0 = v'' + cv' + v(v-b)(1-v). \tag{2.37}$$

Note that the standard energy method for ODEs works only in case $c = 0$ (cf. Section 2.4). In order to find a solution which connects $v_- = 0$ to $v_+ = 1$ we try a solution of the first order equation

$$v' = \alpha v(1-v) =: g(v), \tag{2.38}$$

where the parameter $\alpha > 0$ is still to be determined. By separation of variables, the solution of (2.38) with $v(0) = \frac{1}{2}$ is found to be

$$v_\alpha(\xi) = \frac{1}{1+\exp(-\alpha\xi)}, \quad \xi \in \mathbb{R}. \tag{2.39}$$

We insert (2.38) into (2.37):

$$\begin{aligned} v_\alpha'' + cv_\alpha + v_\alpha(b-v_\alpha)(v_\alpha - 1) &= g'(v_\alpha)v_\alpha' + cv_\alpha' - \frac{1}{\alpha}(b-v_\alpha)v_\alpha' \\ &= v_\alpha'\left(-2\alpha v_\alpha + \alpha + c - \frac{b}{\alpha} + \frac{1}{\alpha}v_\alpha\right). \end{aligned}$$

This term vanishes provided we set

$$\alpha = \frac{1}{\sqrt{2}}, \quad c_\star = \frac{b}{\alpha} - \alpha = \sqrt{2}(b - \frac{1}{2}),$$

which together with (2.39) leads to formula (2.32).
The following proposition summarizes the general methodology.

Proposition 2.11. Let $g \in C^1(\mathbb{R}^m, \mathbb{R}^m)$ satisfy for some $c \in \mathbb{R}$

$$(Dg(v) + cI_m)g(v) + f(v) = 0, \quad \forall v \in \mathbb{R}^m. \tag{2.40}$$

Then any solution $v \in C^1(J, \mathbb{R}^m)$ of $v' = g(v)$ on some interval $J \subseteq \mathbb{R}$ satisfies the travelling wave ODE (2.29) on J.

Proof. It is sufficient to note that $v' = g(v)$ on J implies $v'' = Dg(v)v' = Dg(v)g(v)$ on J by the chain rule. □

Remark 2.12. *For the Nagumo equation (2.31) we used this proposition with the settings $f(v) = v(b-v)(v-1)$, $g(v) = \alpha v(1-v)$ where α and c were determined such that (2.40) holds. If, in the scalar case, g is taken to be a polynomial of degree ℓ, then f should be a polynomial of degree $2\ell - 1$. Moreover, relation (2.40) shows that g is a smooth factor of f, i.e. the zeroes of g are also zeroes of f and if f has even more real zeroes, then they must be incorporated into the first factor $g'(v) + c$.*

Exercise 2.13. Consider the quintic Nagumo equation

$$u_t = u_{xx} - \prod_{i=1}^{5}(u - b_i), \tag{2.41}$$

where $b_1 < b_2 < b_3 < b_4 < b_5$. Determine a relation among $b_1, \dots, b_5$ that allows to compute a travelling wave connection of b_1 to b_3 or of b_3 to b_5 from a first order ODE.

2.4. Phase plane analysis of travelling wave ODEs. As shown in Lemma 2.6, the profile of a travelling wave appears as an orbit connecting two steady states of an autonomous ODE.

Definition 2.14. *Let $w \in C^1(\mathbb{R}, \mathbb{R}^m)$ be a solution of the dynamical system*

$$w' = G(w), \quad \textit{where} \quad G \in C^1(\mathbb{R}^m, \mathbb{R}^m), \tag{2.42}$$

such that the limits

$$\lim_{\xi \to \pm\infty} w(\xi) = w_{\pm} \tag{2.43}$$

exist. Then $\mathcal{Q}(w) = \{w(\xi) : \xi \in \mathbb{R}\}$ is called an orbit connecting *the steady state w_- to the steady state w_+. The connecting orbit is called* heteroclinic *if $w_- \neq w_+$, and* homoclinic *if $w_- = w_+$.*

Therefore, a travelling wave of equation (2.22) with speed $c_\star$ and profile $v_\star$ connecting v_- to v_+, corresponds to an orbit $\mathcal{Q}(w_\star), w_\star = \begin{pmatrix} v_\star \\ v_\star' \end{pmatrix}$ of the dynamical system (2.25) with parameter $c = c_\star$ connecting the steady state $w_- = \begin{pmatrix} v_- \\ 0 \end{pmatrix}$ to $w_+ = \begin{pmatrix} v_+ \\ 0 \end{pmatrix}$. We stress the fact that not only the orbit $w_\star$ but also the speed $c_\star$ is unknown. Therefore, we have to drive the system (2.25) by varying c until an orbit connecting two steady states occurs. Proving that such a connection occurs can be quite hard, and we refer to Section 3 for some situations where this is possible. Before treating further examples we add another observation that applies to travelling waves of (2.22) for which the nonlinearity f is a gradient, i.e.

$$f(v) = \nabla F(v), v \in \mathbb{R}^m \quad \text{for some} \quad F \in C^2(\mathbb{R}^m, \mathbb{R}). \tag{2.44}$$

Proposition 2.15. Let $A \in \mathbb{R}^{m,m}$ be symmetric and let $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$ be the gradient of some $F \in C^2(\mathbb{R}^m, \mathbb{R})$. Further, let $v_* \in C^2(\mathbb{R}, \mathbb{R}^m)$ be a travelling wave of (2.22) with speed $c_* \neq 0$, connecting v_- to v_+. Then, $v_\star' \in L^2(\mathbb{R}, \mathbb{R}^m)$ and the following formula holds

$$c_\star \int_{-\infty}^{\infty} |v_\star'(\xi)|^2 d\xi = F(v_-) - F(v_+). \tag{2.45}$$

Proof. Multiply the travelling wave ODE (2.23) by $v_\star'^T$ and integrate over $[-R, R]$

$$\int_{-R}^{R} v_\star'(\xi)^T A v_\star''(\xi) d\xi + c_\star \int_{-R}^{R} |v_\star'(\xi)|^2 d\xi = -\int_{-R}^{R} v_\star'(\xi)^T \nabla F(v_\star(\xi)) d\xi. \tag{2.46}$$

Integration by parts and the symmetry of A show for the first integral

$$\int_{-R}^{R} v_\star'(\xi)^T A v_\star''(\xi) d\xi = \frac{1}{2}\left((v_\star')^2(R) - (v_\star')^2(-R)\right),$$

which converges to zero as $R \to \infty$. The right hand side in (2.46) is a complete integral

$$\int_{-R}^{R} v_\star'^T(\xi)^T \nabla F(v_\star(\xi)) d\xi = \int_{-R}^{R} \frac{d}{d\xi}(F \circ v_\star)(\xi) d\xi = F(v_\star(R)) - F(v_\star(-R)),$$

which converges to $F(v_+) - F(v_-)$. Therefore, we can take the limit $R \to \infty$ in (2.46) and obtain $v_*' \in L^2(\mathbb{R}, \mathbb{R}^m)$ as well as formula (2.45). □

Formula (2.45) shows that the speed $c_\star$ of the wave is positive if $F(v_-) > F(v_+)$ and negative if $F(v_-) < F(v_+)$, i.e. the wave runs from the larger critical value of the potential F to the smaller critical value. This imposes restrictions on the type of transitions that a wave can take. Note that $c_\star = 0$ still follows from our proof in case $F(v_-) = F(v_+)$, but we cannot conclude $v_\star' \in L^2(\mathbb{R}, \mathbb{R}^m)$ anymore. Proposition 2.15 always applies to the scalar case which has the potential

$$F(\xi) = \int_0^\xi f(x) dx. \tag{2.47}$$

In the scalar case (2.28), the system (2.25) becomes two-dimensional

$$w' = \begin{pmatrix} w_1' \\ w_2' \end{pmatrix} = \begin{pmatrix} w_2 \\ -f(w_1) - cw_2 \end{pmatrix} =: G(w), \tag{2.48}$$

and much insight can be gained by so-called *phase plane analysis*.

Lemma 2.16. *Let $v_0 \in \mathbb{R}$ be a zero of f. Then the steady state $w_0 = \begin{pmatrix} v_0 \\ 0 \end{pmatrix}$ of (2.48) is a*

$$\begin{aligned} &\textit{saddle} \quad \textit{if} \quad f'(v_0) < 0, \\ &\textit{sink} \quad \textit{if} \quad f'(v_0) > 0 \textit{ and } c > 0, \\ &\textit{source} \quad \textit{if} \quad f'(v_0) > 0 \textit{ and } c < 0. \end{aligned}$$

More precisely, the eigenvalues of the linearization

$$DG(w_0) = \begin{pmatrix} 0 & 1 \\ -f'(v_0) & -c \end{pmatrix} \tag{2.49}$$

are

$$\lambda_\pm = \frac{1}{2}\left(-c \pm \sqrt{c^2 - 4f'(v_0)}\right). \tag{2.50}$$

In case of a saddle the eigenvalues satisfy $\lambda_- < 0 < \lambda_+$ *and suitable eigenvectors are*

$$y_- = \begin{pmatrix} 1 \\ \lambda_- \end{pmatrix}, \quad y_+ = \begin{pmatrix} 1 \\ \lambda_+ \end{pmatrix}. \tag{2.51}$$

Proof. The proof follows by straightforward computation of the eigenvalues and eigenvectors of the Jacobian (2.49). Note that in the borderline cases $f'(v_0) = 0$ and $f'(v_0) > 0, c = 0$ the linearization is not enough to determine the type of the steady state since some eigenvalues lie on the imaginary axis (in case $f'(v_0) = 0, c < 0$ one can at least conclude that w_0 is unstable). Also note that the eigenvalues become complex if

$$c^2 < 4f'(v_0), \tag{2.52}$$

i.e. the solutions are spiraling in if $c > 0$ and spiraling out if $c < 0$. ☐

Remark 2.17. *The qualitative phase diagram near steady states is preserved from the linear to the nonlinear case provided there are no eigenvalues on the imaginary axis. This is made precise by the famous Hartman-Grobman theorem, see for example* [4, Theorem 19.9]. *Given a dynamical system* $v' = f(v)$ *with* $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$ *and* t*-flow denoted by* φ^t*. Let* $v_0 \in \mathbb{R}^m$ *be a steady state such that* $Df(v_0)$ *has no eigenvalues on the imaginary axis and let* $\Phi^t(v) = \exp(tDf(v_0))v$ *be the linearized flow. Then there exist neighborhoods* $U \subseteq \mathbb{R}^m$ *of* 0*,* $V \subseteq \mathbb{R}^m$ *of* v_0 *and a homeomorphism* $h : U \to V$ *such that*

$$\varphi^t \circ h(v) = h \circ \Phi^t(v) \tag{2.53}$$

holds for all (t, v) *with* $\Phi^s(v) \in U$ *for all* $\min(t, 0) \le s \le \max(t, 0)$*. The relation* (2.53) *is also called* local flow equivalence.

Demo: The following Figure shows the (numerical)

$$\begin{aligned} c_1 = -0.4 <& c_2 = -0.36 < c_3 = -0.35355339 \approx c_\star \\ <& c_4 = -0.3 < c_5 = -0.2 < c_6 = 0. \end{aligned}$$

Since $f'(0) < 0$ and $f'(1) < 0$ both steady states $w_- = (0, 0)$ and $w_+ = (1, 0)$ are saddles. The eigenvectors from (2.51) are locally tangent to the so called *stable and unstable manifolds*

$$\begin{aligned} W^s(w_\pm) =& \{w : \varphi^\xi(w) \to w_\pm \text{ as } \xi \to \infty\} \\ W^u(w_\pm) =& \{w : \varphi^\xi(w) \to w_\pm \text{ as } \xi \to -\infty\}, \end{aligned} \tag{2.54}$$

where φ^ξ denotes the ξ-flow of the system (2.48). We used the variable ξ instead of t since ξ is the wave variable by derivation. Sometimes this approach is called *spatial dynamics*. As we will see in Section 3, the stability of travelling waves for the time-dependent PDE (2.22) has nothing to do with the stability properties of the asymptotic steady states when considered as equilibria of the travelling wave ODE.

In the current example the one-dimensional unstable manifold of w_- and the one-dimensional stable manifold of w_+ intersect at a specific value of c. Then they must coincide because of uniqueness of solutions to initial value problems to form a heteroclinic orbit.
We also observe that there are orbits connecting the source $w_b = (b, 0)$ to the saddle $w_- = (0, 0)$. These are also travelling waves for the original PDE, but as we will see, the corresponding solutions of the PDE (2.22) do not enjoy the same favorable stability properties as the saddle-to-saddle connection. Moreover, they occur over a whole interval of c-values since w_b has a two dimensional unstable manifold and w_- has a one-dimensional stable manifold.
The classical example for such waves is the *Fisher equation* ([22, Ch.11.4],[10, Ch.4.4] for which, unfortunately, there is no general explicit formula for the travelling wave solution (except for the nontypical speed $c = \frac{5}{\sqrt{6}}$, see [24]).

Example 2.18 (The Fisher equation)**.** On assumes logistic growth of a population and diffusion

$$u_t = u_{xx} + u(1-u), \quad x \in \mathbb{R}, t \geq 0, \tag{2.55}$$

with the travelling wave ODE given by

$$0 = v'' + cv' + v(1-v). \tag{2.56}$$

The steady states are $v_0 = 0$ and $v_1 = 1$ with $f'(v_0) = 1, f'(v_1) = 1$. From Lemma 2.16 we immediately infer for the system (2.48) that

$$\begin{aligned} v_+ &:= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{is} \quad \text{a saddle}, \\ v_- &:= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{is} \begin{cases} \text{an unstable node}, & c < -2, \\ \text{a spiral source}, & -2 < c < 0, \\ \text{a spiral sink}, & 0 < c < 2, \\ \text{a stable node}, & 2 < c. \end{cases} \end{aligned} \tag{2.57}$$

We are looking for a connection from v_- to the saddle v_+ and we restrict to $c < 0$ since the case $c > 0$ follows by reflection (cf. Lemma 2.8). Moreover, since the equation models populations, we are only interested in solutions that have $u(\xi) \geq 0$ for all $\xi \in \mathbb{R}$. The phase diagram reveals that the stable manifold of the saddle (approaching v_+ through negative values of v') originates from the unstable node/ spiral source at v_- . The positivity condition on v excludes the case of a spiral source for $-2 < c < 0$. Therefore, we have a whole family of connecting orbits $v(\cdot, c)$ for parameter values $c \leq -2$ (by a limit argument one can show that the orbit from v_- to v_+ is also nonegative for $c = -2$). In view of the phase portraits near an unstable node it is also reasonable to expect that the connecting orbit leaves v_- in the direction of the eigenvector that belongs to the 'slowest eigenvalue'. In view of (2.50) and (2.51) this direction is $y_- = \begin{pmatrix} \lambda_- \\ 1 \end{pmatrix}$ where $\lambda_- = \frac{1}{2}\left(-c - \sqrt{c^2 - 4}\right)$. In the critical case $c_\star = -2$ we have $\lambda_- = \lambda_+ = 1$.

Since we find a whole family of positive travelling waves ($c \le -2$), it is a natural question to ask, which one of these appears in the longtime dynamics of the PDE (2.54) when initial data $u(\cdot,0) = u_0(\cdot)$ far from the true waves are given. This is a delicate matter which depends sensitively on the behavior of the data u_0 at $\pm\infty$. In a sense the critical wave at $c = -2$ is the 'most stable' one concerning these conditions (see the discussion in [22, Ch.11.2]).

2.5. Examples of nonlinear wave equations. As in Section 2.3 for the parabolic case, we now consider some nonlinear versions of the linear wave equations from Section 2.2.

Example 2.19. Consider the nonlinear wave equation

$$u_{tt} = u_{xx} - \sin u, \quad x, t \in \mathbb{R} \tag{2.58}$$

Exercise 2.20. Show that the general equation

$$Au_{tt} = Ku_{xx} - T\sin(u), \quad A, K, T > 0$$

can be cast into the form (2.58) by suitable transformations.

The travelling wave ODE for (2.58) reads

$$(c^2 - 1)v'' = -\sin(v). \tag{2.59}$$

In the following we assume $|c| \neq 1$. The only steady states of the corresponding first order system are $w_n = (v_n, 0) = (n\pi, 0), n \in \mathbb{Z}$. As in Lemma 2.16 saddles occur for even values of n, while w_n are centers for odd values of n (this, however, does not follow from Lemma 2.16). We look for an explicit solution that connects w_2 to w_0.
An explicit solution may be found by the *energy method*: multiply by v' and integrate,

$$\begin{aligned}(c^2-1)v''v' =& (c^2-1)[\frac{1}{2}(v')^2]' = [\cos(v)]' = -\sin(v)v', \\ \text{const} =& \frac{c^2-1}{2}v'(\xi)^2 - \cos(v(\xi)) \quad \forall \xi \in \mathbb{R}.\end{aligned}$$

Since we want a solution with $v(\xi) \to 0$, $v'(\xi) \to 0$ as $\xi \to \infty$ (resp.$v(\xi) \to 2\pi$, $v'(\xi) \to 0$ as $\xi \to -\infty$) we find the constant const $= -1$. Thus we have to solve

$$(v')^2 = \frac{2}{1-c^2}(1 - \cos(v)), \quad \xi \in \mathbb{R}. \tag{2.60}$$

Since the left-hand side is nonnegative and $1 - \cos(v) \ge 0$ the only case left is $|c| < 1$. Then a solution to (2.60) is found by separation of variables as

$$v_\star(\xi) = 4\arctan\left(\exp\left(-\frac{\xi}{\sqrt{1-c^2}}\right)\right), \quad \xi \in \mathbb{R}. \tag{2.61}$$

The corresponding travelling waves $u(x,t) = v_\star(x - ct)$ exist for all $|c| < 1$.
Demo: $c = \frac{1}{2}$

As shown above the system (2.59) has a conserved quantity $E(v,v') = \frac{c^2-1}{2}(v')^2 - \cos(v)$. Hence the phase plane consists of the level curves of E, apart from homoclinic orbits there are also continua of periodic orbits which lead to wave trains.

The next example contains third order spatial derivatives. It is derived by simplifying equations for water waves and to approximately describe the propagation of solitary waves, compare [31, Ch.13.11].

Example 2.21 (The Korteweg-de Vries (KdV) equation)**.** The equation is the following

(2.62) $$u_t = -uu_x - u_{xxx}, \quad x \in \mathbb{R}$$

with travelling wave ODE

(2.63) $$0 = cv' - \frac{1}{2}(v^2)' - v'''.$$

Integrating once leads to

(2.64) $$v'' = cv - \frac{1}{2}v^2 = v(c - \frac{1}{2}v) =: f(v).$$

We took the integration constant to be zero since we look for solutions satisfying $v(\xi) \to 0$ and $v''(\xi) \to 0$ as $\xi \to \pm\infty$.
Equation (2.64) can be handled completely by the energy method since it has the conserved quantity

(2.65) $$E(v) = \frac{1}{2}(v')^2 + P(v), \quad P(v) = -\int_0^v f(s)ds = -\frac{c}{2}v^2 + \frac{1}{6}v^3.$$

Here P is the *potential* satisfying $P' = -f$, and (2.64) may be considered as the Newtonian dynamics of a particle of mass 1 moving in the potential P. The first order system has a homoclinic orbit connecting $(0,0)$ to itself which belongs to the energy level $E(v) = 0$ and, for $c > 0$, is explicitly given by

(2.66) $$v(\xi) = 3c\,\mathrm{sech}^2(\frac{\sqrt{c}}{2}\xi), \quad \xi \in \mathbb{R}.$$

Thus the KdV equation has a whole family of solitary waves given by

$$u(x,t) = 3c\,\mathrm{sech}^2(\frac{\sqrt{c}}{2}(x-ct)), \quad c > 0.$$

Demo: travelling wave and phase plane diagram for KdV

2.6. Travelling waves in parabolic systems.

Example 2.22 (FitzHugh-Nagumo system)**.** Consider the 2-dimensional FitzHugh-Nagumo system

(2.67) $$u_t = \begin{pmatrix} 1 & 0 \\ 0 & \rho \end{pmatrix} u_{xx} + \begin{pmatrix} u_1 - \frac{1}{3}u_1^3 - u_2 \\ \varepsilon(u_1 + a - bu_2) \end{pmatrix}, \; x \in \mathbb{R}, \, t \geq 0,$$

with $u = u(x,t) \in \mathbb{R}^2$ and positive $\rho, a, b, \varepsilon \in \mathbb{R}$, [11]. For the parameters $\rho = 0.1$, $a = 0.7$, $\varepsilon = 0.08$ it is well-known that (2.67) has travelling pulse solutions for $b = 0.8$ and travelling front solutions for $b = 3$. Unfortunately, in both cases there are no explicit representations neither for the wave profile $w_\star$ nor for their velocities $c_\star$. However, we know some approximations of the velocities in both cases. The pulse travels at velocity $c_\star = -0.7892$ and the front at velocity $c_\star = -0.8557$.
In fact, the value of b determines the number of steady states of (2.67) (for b small we have only one steady state but for b large we have three).

The next Lemma improves Lemma 2.6 by showing that the convergence of a travelling wave to its limits implies the convergence of derivatives. In fact, if v satisfies the travelling wave ODE (2.23) and $v_\pm = \lim_{\xi\to\pm\infty} v(\xi)$ exist, then the following Lemma applies to $h(\xi) = -f(v(\xi)), \xi \in \mathbb{R}$ and shows

$$f(v_\pm) = 0, \quad \lim_{\xi\to\pm\infty} v'(\xi) = 0.$$

Lemma 2.23. *Let $A \in \mathbb{R}^{m,m}$ have only eigenvalues with positive real part and suppose $v \in C^2(\mathbb{R}, \mathbb{R}^m)$ and $c \in \mathbb{R}$ solve the second order ODE*

$$Av'' + cv' = h \in C(\mathbb{R}, \mathbb{R}^m), \tag{2.68}$$

such that both limits $\lim_{\xi\to\pm\infty} h(\xi)$ and $\lim_{\xi\to\pm\infty} v(\xi)$ exist. Then the following equalities hold

$$\lim_{\xi\to\pm\infty} h(\xi) = 0 = \lim_{\xi\to\pm\infty} v'(\xi). \tag{2.69}$$

Proof. It is sufficient to prove the result for the positive axis $\mathbb{R}_+ = [0, \infty)$, then the result for the negative axis follows by reflection since we have made no assumption on the sign of c. First multiply (2.68) by A^{-1} to obtain

$$\begin{aligned} v'' + cA^{-1}v' &= A^{-1}h =: r, \\ \lim_{\xi\to\infty} r(\xi) &= A^{-1} \lim_{\xi\to\infty} h(\xi) =: r_+. \end{aligned} \tag{2.70}$$

All solutions of (2.70) can be written as follows

$$v(\xi) = V_1(\xi)\alpha_1 + V_2(\xi)\alpha_2 + v_3(\xi), \quad \xi \in \mathbb{R}, \alpha_1, \alpha_2 \in \mathbb{R}^m, \tag{2.71}$$

where $\begin{pmatrix} V_1 & V_2 \\ V_1' & V_2' \end{pmatrix} \in \mathbb{R}^{2m,2m}$ forms a fundamental system of the first order equation obtained from (2.70) and v_3 is a special solution of the inhomogenous equation. We will construct v_3 in the following form

$$v_3(\xi) = \int_0^\infty G(\xi,\eta) r(\eta) d\eta, \quad \xi \in \mathbb{R} \tag{2.72}$$

with a suitable Green's matrix G. We will use the boundedness of v to show that $v_3(\xi)$ is the dominant part in (2.71) and that $|v_3(\xi)| \to \infty$ as $\xi \to \infty$ if $r_+ \neq 0$.

$c > 0$: Take $V_1(\xi) = I_m$, $V_2(\xi) = A\exp(-cA^{-1}\xi)$ and

$$G(\xi,\eta) = \begin{cases} \frac{1}{c}A\left(I_m - \exp(-cA^{-1}(\xi-\eta))\right), & 0 \le \eta \le \xi, \\ 0, & \xi < \eta. \end{cases} \tag{2.73}$$

One easily verifies that V_1, V_2 generate fundamental solutions and that v_3 satisfies the inhomogeneous equation (2.70) as well as $v_3(0) = v_3'(0) = 0$.

Let us note that all eigenvalues of A^{-1} have positive real part, hence there exists $\rho_1 > 0$ such that $\mathrm{Re}(\mu) < -\rho_1$ for all eigenvalues μ of $-A^{-1}$. By a well-known result (see Numerical Analysis of Dynamical Systems, for example) we then have a constant $C > 0$ such that

$$|\exp(-\tau A^{-1})| \le C\exp(-\rho_1\tau), \quad \tau \ge 0. \tag{2.74}$$

Since V_1, V_2 and v itself are bounded, equation (2.71) implies that v_3 is bounded. We now assume $r_+ \neq 0$ and show that $|v_3(\xi)| \ge C\xi$ for ξ large which is a contradiction. Given $\varepsilon > 0$, take ξ_ε so large that

$$|r(\xi) - r_+| \le \varepsilon \quad \text{for all} \quad \xi \ge \xi_\varepsilon.$$

Then we estimate with (2.74)

$$\begin{aligned}
|v_3(\xi)| \ge & \left|\int_{\xi_\varepsilon}^{\xi} G(\xi,\eta)d\eta r_+\right| - \left|\int_0^{\xi_\varepsilon} G(\xi,\eta)r(\eta)d\eta\right| \\
& - \left|\int_{\xi_\varepsilon}^{\xi} G(\xi,\eta)(r(\eta)-r_+)d\eta\right| \\
\ge & \left|\frac{A}{c}\left[\eta I_m - \frac{A}{c}\exp(-cA^{-1}(\xi-\eta))\right]_{\xi\varepsilon}^{\xi} r_+\right| \\
& -\varepsilon\int_{\xi_\varepsilon}^{\xi} |G(\xi,\eta)|d\eta - \int_0^{\xi_\varepsilon} |G(\xi,\eta)|d\eta\|r\|_\infty \\
\ge & \frac{\xi-\xi_\varepsilon}{c}|Ar_+| - \frac{2C|A|^2|r_+|}{c^2} - \varepsilon(\xi-\xi_\varepsilon)\frac{2C|A|}{c} - \frac{2|A|C\|r\|_\infty\xi_\varepsilon}{c}.
\end{aligned}$$

Since $|Ar_+| > 0$, the third term can be absorbed into the first one by taking ε sufficiently small, which then dominates the second and the fourth one by taking $\xi - \xi_\varepsilon$ sufficiently large. This shows that v_3 is unbounded, a contradiction.

Finally, we obtain from (2.71)

$$v'(\xi) = V_2'(\xi)\alpha_2 + v_3'(\xi), \quad \xi \in \mathbb{R}.$$

The first term decays exponentially due to (2.74). The second term is

$$v_3'(\xi) = \int_0^{\xi} \exp(-cA^{-1}(\xi-\eta))r(\eta)d\eta.$$

With (2.74) we estimate v_3' as follows

$$
\begin{aligned}
|v_3'(\xi)| \leq & \int_0^{\xi/2} |G(\xi,\eta)| d\eta \|r\|_\infty + \int_{\xi/2}^{\xi} |G(\xi,\eta)||r(\eta)| d\eta \\
\leq & C \left\{ \int_0^{\xi/2} \exp(-c\rho_1(\xi-\eta)) d\eta \|r\|_\infty + \int_{\xi/2}^{\xi} \exp(-c\rho_1(\xi-\eta)) d\eta \sup_{\eta \geq \xi/2} |r(\eta)| \right\} \\
\leq & \frac{C}{c\rho_1} \left\{ \left(\exp(-\frac{c\rho_1\xi}{2}) - \exp(-c\rho_1\xi) \right) \|r\|_\infty \right. \\
& \left. + \left(1 - \exp(-\frac{c\rho_1\xi}{2}) \right) \sup_{\eta \geq \xi/2} |r(\eta)| \right\}.
\end{aligned}
$$

Since $r_+ = 0$ and $c, \rho_1 > 0$ all terms on the right-hand side converge to zero as $\xi \to \infty$.

$c = 0$: The fundamental matrices are $V_1(\xi) = I_m$, $V_2(\xi) = \xi I_m$ and Green's matrix is given by

$$
G(\xi,\eta) = \begin{cases} (\xi-\eta) I_m, & \eta \leq \xi, \\ 0, & \eta > \xi. \end{cases} \tag{2.75}
$$

Assume $r_+ \neq 0$ and for a given $\varepsilon > 0$ take ξ_ε such that $|r(\eta) - r_+| \leq \varepsilon$ for all $\eta \geq \xi_\epsilon$. Then we estimate

$$
\begin{aligned}
|v_3(\xi)| \geq & \left| \int_{\xi_\varepsilon}^{\xi} (\xi-\eta) d\eta \, r_+ \right| \\
& - \left| \int_{\xi_\varepsilon}^{\xi} (\xi-\eta)(r_+ - r(\eta)) d\eta \right| - \left| \int_0^{\xi_\varepsilon} (\xi-\eta) r(\eta) d\eta \right| \\
\geq & \frac{1}{2}(|r_+| - \varepsilon)(\xi - \xi_\varepsilon)^2 - \|r\|_\infty \xi_\varepsilon (\xi - \frac{1}{2}\xi_\varepsilon).
\end{aligned}
$$

By taking ε sufficiently small and letting $\xi \to \infty$, we find that v_3 grows quadratically if $r_+ \neq 0$. This is stronger than $\alpha_1 + V_2(\xi)\alpha_2$ which grows at most linearly, and contradicts the boundedness of v_3.

Equation (2.68) has the simple form $v'' = r$ and we have to show that $\lim_{\xi\to\infty} r(\xi) = 0$ and the existence of $\lim_{\xi\to\infty} v(\xi)$ imply $\lim_{\xi\to\infty} v'(\xi) = 0$. Without loss of generality we can assume $m = 1$. Then the mean value theorem implies that

$$
v(n+1) - v(n) = v'(\theta_n) \quad \text{for some} \quad \theta_n \in (n, n+1).
$$

Since $v(n)$ and $v(n+1)$ have the same limit, we infer $v'(\theta_n) \to 0$ as $n \to \infty$. Therefore, for $n \leq \xi \leq n+1$

$$
\begin{aligned}
|v'(\xi)| \leq & |v'(\xi) - v'(\theta_n)| + |v'(\theta_n)| \\
\leq & \sup_{\eta \in [n,n+1]} |v''(\eta)||\xi - \theta_n| + |v'(\theta_n)|
\end{aligned}
$$

$$\leq \sup_{\eta\in[n,n+1]} |v''(\eta)| + |v'(\theta_n)| \to 0 \quad \text{as} \quad n \to \infty.$$

$c < 0$: As in case $c > 0$ we use the fundamental matrices $V_1(\xi) = I_m$ and $V_2(\xi) = A\exp(-cA^{-1}\xi)$, however the Green's matrix is no longer a triangular kernel

$$G(\xi,\eta) = \frac{A}{c} \begin{cases} I_m - \exp(cA^{-1}\eta), & \eta \leq \xi, \\ \exp(cA^{-1}(\eta-\xi))(I_m - \exp(cA^{-1}\xi)), & \eta > \xi. \end{cases} \tag{2.76}$$

Note that (2.74) and $c < 0$ imply for some constant $C_1 > 0$ the estimate

$$|G(\xi,\eta)| \leq C_1 \begin{cases} 1, & \eta \leq \xi, \\ \exp(-|c|\rho_1(\eta-\xi)), & \eta > \xi. \end{cases} \tag{2.77}$$

Therefore, the integral in (2.72) exists and yields the estimate

$$|\int_0^\infty G(\xi,\eta)r(\eta)d\eta| \leq \int_0^\infty |G(\xi,\eta)|d\eta\|r\|_\infty \leq C_1(\xi + \frac{1}{\rho_1|c|})\|r\|_\infty.$$

It is then straightforward to verify that (2.72) yields a solution of (2.68) which grows at most linearly with ξ. Therefore, the dominant term in (2.71) is

$$|V_2(\xi)\alpha_2| = |\exp(-cA^{-1}\xi)A\alpha_2| \geq \exp(|c|\rho_1\xi)|A\alpha_2|.$$

Since v is bounded we conclude that $\alpha_2 = 0$. Let us assume $r_+ \neq 0$ and estimate v_3 from below with ξ_ε chosen as in the previous cases:

$$\begin{aligned} |v_3(\xi)| \geq & |\int_{\xi_\varepsilon}^\infty G(\xi,\eta)d\eta\, r_+| - |\int_0^{\xi_\varepsilon} G(\xi,\eta)r(\eta)d\eta| \\ & -|\int_{\xi_\varepsilon}^\infty G(\xi,\eta)(r_+ - r(\eta))d\eta|. \end{aligned}$$

For the first integral we find

$$\begin{aligned} \int_{\xi_\varepsilon}^\infty G(\xi,\eta)d\eta = & \frac{A}{c}\left[\eta I_m - \frac{A}{c}\exp(cA^{-1}\eta)\right]_{\xi_\varepsilon}^{\xi} \\ & -\frac{A}{c}(I_m - \exp(cA^{-1}\xi))\left[\frac{A}{c}\exp(cA^{-1}(\eta-\xi))\right]_\xi^\infty \\ = & \frac{A}{c}\left\{(\xi-\xi_\varepsilon)I_m + \frac{A}{c}(I_m - \exp(cA^{-1}\xi_\varepsilon))\right\}, \end{aligned}$$

so that for some constants $C_1, C_2 > 0$

$$|\int_{\xi_\varepsilon}^\infty G(\xi,\eta)r_+ d\eta| \geq C_1|\xi - \xi_\varepsilon||Ar_+| - C_2. \tag{2.78}$$

For the third term we obtain from (2.77)

$$
\begin{aligned}
|\int_{\xi_\varepsilon}^{\infty} G(\xi,\eta)(r_+ - r(\eta))d\eta \le & \varepsilon \int_{\xi_\varepsilon} |G(\xi,\eta)|d\eta \\
\le & \varepsilon(C_1(\xi-\xi_\varepsilon)+c_2).
\end{aligned}
$$

This term can be absorbed into (2.78) by taking ε sufficiently small. Finally, equation (2.77) also leads to an estimate of the second term

$$
|\int_0^{\xi_\varepsilon} G(\xi,\eta)r(\eta)d\eta| \le C_1\xi_\varepsilon\|r\|_\infty.
$$

Summing up, if $r_+ \neq 0$ we get a linearly growing lower bound for $v_3(\xi)$ which contradicts the boundedness of v.

Finally we have to show $v'(\xi) \to 0$ as $\xi \to \infty$. For this note that from (2.76) we have

$$
v_3'(\xi) = -\int_{\xi}^{\infty} \exp(cA^{-1}(\eta-\xi))r(\eta)d\eta.
$$

This yields the estimate

$$
\begin{aligned}
|v_3'(\xi)| \le & \sup_{\eta\ge\xi}|r(\eta)| \left| \int_{\xi}^{\infty} \exp(cA^{-1}(\eta-\xi))d\eta \right| \\
\le & \sup_{\eta\ge\xi}|r(\eta)| \int_{\xi}^{\infty} \exp(-\rho_1|c|)d\eta = \frac{1}{\rho_1|c|}\sup_{\eta\ge\xi}|r(\eta)|,
\end{aligned}
$$

which shows $v'(\xi) = v_3'(\xi) \to 0$ as $\xi \to \infty$ and finishes the proof.

□

Remark 2.24. *We notice that the convergence of $v(\xi)$ as $\xi \to \pm\infty$ was only used to derive $v'(\xi) \to 0$ in case $c = 0$. All other conclusions work under the hypothesis that v is bounded.*

2.7. Waves in complex-valued systems in one space dimension. Quite a few models in Physics lead to PDEs for complex-valued functions

$$
u : \mathbb{R} \times [0,\infty) \to \mathbb{C}, \quad (x,t) \to u(x,t). \tag{2.79}
$$

We just mention the well-known Schrödinger equation from quantum mechanics. There u is a wave function with complex amplitude, the absolute value $|u(x,t)|$ of which may be interpreted as the probability of finding the 'particle' at position x at time t.

In this subsection we consider equations of the type

$$
u_t = au_{xx} + g(x,|u|)u, \quad x \in \mathbb{R}, t \in \mathbb{R}, \tag{2.80}
$$

where $a \in \mathbb{C}$ and $g : \mathbb{R}^2 \to \mathbb{C}$ is a smooth function. In case $\mathrm{Re}(a) > 0$ the equation is of parabolic type while in case $a = i$ it is of wave type (the Schrödinger equation). This can be seen from the real-valued version of (2.80), which reads (with $a = a_1 + ia_2$, $u = u_1 + iu_2$, $g = g_1 + ig_2$):

$$(2.81) \qquad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}' = \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_{xx} + \begin{pmatrix} g_1 & -g_2 \\ g_2 & g_1 \end{pmatrix} (x, |u|) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

We consider some characteristic examples:

Example 2.25 (Linear Schrödinger equation (LSE))**.**

$$(2.82) \qquad a = i, \quad g(x, |u|) = -iV(x), x \in \mathbb{R}, \quad V : \mathbb{R} \to \mathbb{R} \text{ potential.}$$

Note that in Physics one usually writes

$$i\hbar\Psi_t = -\frac{\hbar^2}{2\mu}\Delta\Psi + V(x)\Psi$$

where $\hbar$ is *Planck's reduced constant* and Ψ is the *wave function*. Dividing by $i\hbar$ and scaling the space variable allows to reduce this to (2.80),(2.82). We prefer to keep the pure time derivative u_t on the left-hand side in order not to change our general evolution equation (2.1).

Example 2.26 (Nonlinear Schrödinger equation (NLS))**.**

$$(2.83) \qquad a = i, \quad g(x, |u|) = \beta|u|^p, \ \beta = ib, b \in \mathbb{R}, p \geq 2.$$

For $p = 2$ we have the *cubic nonlinear Schrödinger equation (CNLS)*.

Example 2.27 (The Gross-Pitaevskii equation (GPE))**.** This is a mixture of LSE and the cubic NLS and supposed to describe so-called Bose-Einstein condensates:

$$(2.84) \qquad a = \frac{1}{2}i, \quad g(x, |u|) = -iV(x) + \beta|u|^2, \quad \beta = ib, b \in \mathbb{R}.$$

Example 2.28 (Complex Ginzburg-Landau equation)**.** This equation occurs in applications to superconductivity, in nonlinear optics, and in laser physics:

$$(2.85) \qquad rm\,\mathrm{Im}(a) \neq 0, \mathrm{Re}(a) > 0, \quad g(x, |u|) = \mu + \beta|u|^2 + \gamma|u|^4,$$

where typically $\mu \in \mathbb{R}$ but $\beta, \gamma \in \mathbb{C}$. In the case given here, the nonlinearity has quintic terms, therefore it is called the *quintic complex Ginzburg-Landau equation (QCGL)*.

Rather than travelling waves we look for a *standing oscillating pulse* (sometimes called an *oscillon*)

$$(2.86) \qquad u(x, t) = \exp(-i\theta t)v(x), \quad x \in \mathbb{R}, t \geq 0, 0 \leq \theta < \pi.$$

Insert this into (2.80) and obtain

$$-i\theta e^{-i\theta t}v(x) = ae^{-i\theta t}v''(x) + g(x, |v|)e^{-i\theta t}v(x),$$

which leads to the *oscillating pulse ODE*

$$(\text{OPODE}) \qquad 0 = av'' + i\theta v + g(x, |v|)v.$$

Example 2.29 (The cubic NLS)**.** With $p = 2$ and (2.83), equation (OPODE) reads

$$0 = v'' + \theta v + b|v|^2 v. \tag{2.87}$$

We realize that (2.87) is very similar to the travelling wave ODE (2.64) of the KdV equation, and we find real-valued solutions from the conserved quantity

$$\begin{aligned} E(v, v') =& \frac{1}{2}v'^2 + P(v), \\ P(v) =& \int_0^v \theta s + bs^3 ds = \frac{1}{2}\theta v^2 - \frac{1}{4}bv^4. \end{aligned}$$

Note that we expect $v(x) \to 0$ as $x \to \pm\infty$, hence the constant of integration is taken to be zero and we solve $E(v, v') = 0$. In fact the ansatz

$$v(x) = c_1 \mathrm{sech}(c_2 x), \quad x \in \mathbb{R} \tag{2.88}$$

leads to the condition $-\theta = c_2^2, c_1^2 = \frac{2}{b}$. Hence we assume $b > 0$ and find the solution

$$v(x) = \sqrt{\frac{2}{b}}\mathrm{sech}(c_2 x), \quad x \in \mathbb{R},$$

where $c_2 > 0$ is an arbitrary parameter. With θ from above, the travelling waves are then given by the formula

$$u(x, t) = \sqrt{\frac{2}{b}}\exp(ic_2^2 t)\mathrm{sech}(c_2 x), \quad x, t \in \mathbb{R}. \tag{2.89}$$

Let us recall that we found oscillating waves since the nonlinear operator

$$F(u) = au_{xx} + g(x, |u|)u, \quad u \in C^2(\mathbb{R}, \mathbb{C})$$

satisfies the *equivariance condition*

$$F(e^{i\theta}u) = e^{i\theta}F(u), u \in C^2(\mathbb{R}, \mathbb{C}), \theta \in \mathbb{C}.$$

In the travelling wave case we had

$$F(u) = Au_{xx} + f(u), \quad u \in C^2(\mathbb{R}, \mathbb{R}^m),$$

and equivariance with respect to translations

$$F(u(\cdot - \gamma)) = [F(u)](\cdot - \gamma), \quad \gamma \in \mathbb{R}.$$

The notion of *equivariance* will be formulated in abstract terms in Section 3.
Note that this translation equivariance also holds for the complex equations NLS and QCGL, but not for the LSE and GPE which both contain a space dependent potential.
For the first two equations it is therefore reasonable to look for solutions which travel and oscillate simultaneously

$$u(x, t) = e^{-i\theta t}v(x - ct), \quad x, t \in \mathbb{R}. \tag{2.90}$$

Plugging this into the autonomous complex equation

$$u_t = au_{xx} + g(|u|)u, \tag{2.91}$$

leads to the travelling and oscillating wave ODE (TWOSODE)

$$0 = av'' + cv' + (g(|v|) + i\theta)v, \quad x \in \mathbb{R}, v(x) \in \mathbb{C}. \tag{TWOSODE}$$

Exercise 2.30. Determine for the NLS a wave that oscillates and travels by solving (TWOSODE) via the ansatz (2.90) with

$$v(\xi) = c_1 \exp(ic_2\xi)\mathrm{sech}(c_3\xi), \quad \xi \in \mathbb{R}. \tag{2.92}$$

Solution: With the abbreviations

$$\varepsilon = \exp(ic_2\xi), \quad \sigma = \mathrm{sech}(c_3\xi), \quad \tau = \tanh(c_3\xi)$$

we obtain

$$\begin{aligned}
\varepsilon' &= ic_2\varepsilon, \quad \sigma' = -c_3 j\tau\sigma, \quad \tau' = c_3\sigma^2, \quad \sigma^2 = 1-\tau^2 \\
v' &= c_1(ic_2\varepsilon\sigma - \varepsilon c_3\tau\sigma) = v(ic_2 - tc_3), \quad v'' = v(ic_2 - c_3\tau)^2 - vc_3^2\sigma^2, \\
v'' &- icv' + (b|v|^2 + \theta)v = v(ic_2 - c_3\tau)^2 - c_3^2 v\sigma^2 - icv(ic_2 - c_3\tau) + (bc_1^2\sigma^2 + \theta)v \\
&= v[-c_2^2 + cc_2 + \theta] + iv\tau[-2c_2c_3 + cc_3] + v[c_3^2\tau^2 - c_3^2\sigma^2 + bc_1^2\sigma^2] \\
&= v[-c_2^2 + cc_2 + c_3^2 + \theta] + iv\tau c_3[c - 2c_2] + v\sigma^2[bc_1^2 - 2c_3^2].
\end{aligned}$$

Hence we arrive at the relations

$$c = 2c_2, \quad c_1 = c_3\sqrt{\frac{2}{b}}, \quad \theta = -c_2^2 - c_3^2$$

and at the travelling and oscillating wave

$$u(x,t) = \exp(-i\theta t)v(x - 2c_2 t) = c_3\sqrt{\frac{2}{b}}\frac{\exp(ic_2 x + i(c_3^2 - c_2^2)t)}{\cosh(c_3(x - 2c_2 t))}.$$

Since the parameters c_1, c_2, c_3 can be chosen arbitrarily this formula yield a three-parameter family of waves.

2.8. Selected waves in more than one space dimension. In this final subsection we consider reaction diffusion in systems in $d \geq 2$ space dimensions:

$$u_t = A\Delta u + f(u), \quad x \in \mathbb{R}^d, \quad t \geq 0, \quad u(x,t) \in \mathbb{R}^m, \tag{2.93}$$

where $A \in \mathbb{R}^{m,m}$ is assumed to have eigenvalues with positive real part and $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$. For v in a suitable function space (e.g. $H^2(\mathbb{R}^d\mathbb{R}^m)$ or $v \in C^2_{\mathrm{unif}}(\mathbb{R}^d, \mathbb{R}^m)$) the operator F given by

$$[F(v)](x) = A\Delta v(x) + f(v(x)), \quad x \in \mathbb{R}^d$$

satisfies for $\gamma \in \mathbb{R}^d$

$$[F(v)](x - \gamma) = F(v(\cdot - \gamma))(x), \quad x \in \mathbb{R}^d.$$

We look for a travelling wave of the form

$$u(x,t) = v(k^T x - \omega t), \tag{2.94}$$

with *wave vector* $k \in \mathbb{R}^d$, *frequency* $\omega \in \mathbb{R}$ and *one dimensional profile* $v : \mathbb{R} \to \mathbb{R}^d$. Solutions of this form are called *planar waves* since their value is constant in the hyperplane $k^\top$ and they travel in direction k. A short computation reveals that (2.94) defines a solution of (2.93) if v, k and ω satisfy

$$0 = A|k|^2 v'' + \omega v' + f(v), \qquad \text{(PWODE)}$$

compare equation (TWODE). Usually, one normalizes $|k| = 1$, so that (PWODE) agrees with (TWODE) in one space dimension.
But in $\mathbb{R}^d, d \geq 2$ it is also possible to have rotations determined by the orthogonal group

$$\mathrm{O}(\mathbb{R}^d) = \{Q \in \mathbb{R}^{d,d} : Q^T Q = I_d\}. \qquad (2.95)$$

The orthogonal group acts on functions $v : \mathbb{R}^d \to \mathbb{R}^m$ as follows

$$v_Q(x) = v(Q^T x), \quad x \in \mathbb{R}^d, \quad Q \in \mathrm{O}(\mathbb{R}^d). \qquad (2.96)$$

Lemma 2.31. *The operator $F(v) = A\Delta v + f(v)$ satisfies for sufficiently smooth $v : \mathbb{R}^d \to \mathbb{R}^m$ the condition*

$$F(v_Q)(x) = (F(v))(Q^T x), \quad x \in \mathbb{R}^d, \quad Q \in \mathrm{O}(\mathbb{R}^d). \qquad (2.97)$$

Proof. It remains to show for the Laplacian

$$\Delta v_Q(x) = (\Delta v)(Q^T x), x \in \mathbb{R}^d.$$

In fact, the chain rule shows

$$\begin{aligned}
D_i v_Q(x) &= \sum_{j=1}^{d} D_j v(Q^T x)(Q^T)_{ji}, \\
D_\ell D_i v_Q(x) &= \sum_{j,k=1}^{d} D_k D_\ell v(Q^T x)(Q^T)_{k\ell}(Q^T)_{ji} \\
D_i^2 v_Q(x) &= \sum_{j,k=1}^{d} D_k D_j v(Q^T x)(Q^T)_{ji} Q_{ik}, \\
\Delta v_Q(x) &= \sum_{i=1}^{d} D_i^2 v_Q(x) = \sum_{j,k=1}^{d} D_k D_j v(Q^T x) \sum_{i=1}^{d} (Q^T)_{ji} Q_{ik} \\
&= \sum_{j,k=1}^{d} D_k D_j v(Q^T x)\delta_{jk} = \sum_{j=1}^{d} D_j^2 v(Q^T x) = \Delta v(Q^T x).
\end{aligned}$$

□

Example 2.32 (The Barkley model)**.** The following equations were set up by Barkley as a model for excitable media that show the occurence and parametric behavior of spiral waves. For $m = 2, d = 2$ they read

$$u_t = \begin{pmatrix} 1 & 0 \\ 0 & d_2 \end{pmatrix} \Delta u + \begin{pmatrix} \frac{1}{\varepsilon} u_1 (1 - u_1)(u_1 - \frac{u_2 + b}{a}) \\ g(u_1) - u_2 \end{pmatrix}. \qquad (2.98)$$

Here u_1 is the activator which diffuses with constant 1 and u_2 is the inhibitor which diffuses at a much smaller rate $0 < d_2 \ll 1$. The nonlinearity in the first component is of cubic shape as in the Nagumo or the FitzHugh Nagumo equation, see (2.31),(2.67). However, the position $\frac{u_2+b}{a}$ of the intermediate steady state depends on the inhibitor and the parameters $a, b > 0$. The constant $\varepsilon > 0$ is assumed to be small so that the dynamics of the activator become fast. There are 2 common choices for the nonlinearity in the second equation

$$g(w) = w, \quad \text{or} \quad g(w) = w^3, \quad w \in \mathbb{R}.$$

The following figures show a simulation with parameters

$$d_2 = \frac{1}{10}, \quad a = \frac{3}{4}, \quad b = \frac{1}{100}, \quad \varepsilon = \frac{1}{50},$$

and initial data

$$u_1(x,0) = \begin{cases} 0, & x_1 \le 0, \\ 1, & x_1 > 0, \end{cases}, u_2(x,0) = \begin{cases} 0, & x_2 \le 0, \\ \frac{a}{2}, & x_2 > 0 \end{cases}.$$

We observe that a spiral wave develops that seems to rotate about the origin.

Let us make the observation from this example more precise. We consider the solution that develops as $t \to \infty$ as a *rigidly rotating wave*, i.e. a function

$$u(x,t) = v_\star(Q^T(t)x), \quad Q(t) = \begin{pmatrix} \cos(\omega_\star t) & -\sin(\omega_\star t) \\ \sin(\omega_\star t) & \cos(\omega_\star t) \end{pmatrix}, \tag{2.99}$$

where $v_\star : \mathbb{R}^2 \to \mathbb{R}^m$ is the profile of the wave and $\omega_\star$ is its frequency of rotation or its *angular velocity*. In order to match this with the pictures observed, imagine that we fix a specific feature of the profile, e.g. the tip of the spiral, which at time $t = 0$ is located at $x_\star \in \mathbb{R}^2$. At time $t > 0$ we will find the same feature at position $x(t)$ where

$$Q^T(t)x(t) = x_\star, \quad \text{i.e.} \quad x(t) = Q(t)x_\star.$$

Therefore, the profile rotates counterclockwise if $\omega_\star > 0$ and clockwise if $\omega_\star < 0$. Next we set up the equation satisfied by the profile $v_\star$ of a rigidly rotating wave. First note that the rotation $Q(t)$ in (2.99) may be written as the exponential of a simple skew-symmetric matrix

$$Q(t) = \exp(tS), \quad S = \begin{pmatrix} 0 & -\omega_\star \\ \omega_\star & 0 \end{pmatrix}.$$

For a general skew-symmetric matrix $S \in \mathbb{R}^{d,d}, S = -S^T$ the matrix $Q = \exp(S)$ is orthogonal, since

$$Q^TQ = \exp(S)^T\exp(S) = \exp(S^T)\exp(S) = \exp(-S)\exp(S) = \exp(0) = I_d.$$

In Section 3 we will put this relation into the more general framework of Lie groups and Lie algebras.

Proposition 2.33. Let $S \in \mathbb{R}^{d,d}$ be skew-symmetric and let

$$u(x,t) = v(\exp(-tS)x), x \in \mathbb{R}^d, t \geq 0 \tag{2.100}$$

be a solution of the parabolic system (2.93). Then the profile v solves the equation

$$0 = A\Delta v(x) + v_x(x)Sx + f(v(x)), \quad x \in \mathbb{R}^d. \tag{2.101}$$

Conversely, if v satisfies (2.101) then (2.100) defines a solution of the system (2.93).

Proof. Using Lemma 2.31 and the orthogonality of $\exp(tS)$, the proof follows from the straightforward calculation

$$\begin{aligned} u_t(x,t) =& - v_x(\exp(-tS)x)S\exp(-tS)x, \\ A\Delta u(x,t) =& (A\Delta v)(\exp(-tS)x), \\ f(u(x,t)) =& (f \circ v)(x,t), \end{aligned}$$

and the substitution $y = \exp(-tS)x$. □

Remark 2.34. *Note that $v_x \in \mathbb{R}^{m,d}$ is the total derivative so that equation (2.101) contains a first order term with unbounded coefficients Sx. One may also write this term as $\langle Sx, \nabla v(x)\rangle$ where $\nabla v = v_x^T$ is the gradient and the inner product is defined as $\langle Sx, \nabla v\rangle = \sum_{i=1}^d (Sx)_i D_i v$. Using the skew-symmetry we can also rewrite this term as follows*

$$\begin{aligned} (v_x(x)Sx)_i =& \sum_{j=1}^d D_j v_i(x)(Sx)_j = \sum_{j,k=1}^d D_j v_i(x) S_{jk} x_k \\ =& \sum_{j<k} D_j v_i(x) S_{jk} x_k - \sum_{j>k} D_j v_i(x) S_{kj} x_k \\ =& \sum_{j<k} S_{jk}(x_k D_j v_i - x_j D_k v_i)(x), \quad i = 1, \ldots, m. \end{aligned}$$

The expression shows that one takes angular derivatives of the v-components in the (x_j, x_k) plane (as indicated by $x_k D_j - x_j D_k$) with velocity S_{jk}.

Of course, a rotating wave need not rotate about the origin, in general. For example, if we have a rotating wave as in (2.100), then for any $x_0 \in \mathbb{R}^d$ the function

$$u_0(x,t) := u(x - x_0, t) = v(\exp(-tS)(x - x_0)), x \in \mathbb{R}^d, t \geq 0$$

also solves (2.93). Defining the shifted profile by

$$v_0(x) = v(x - x_0), x \in \mathbb{R}^d,$$

we obtain the following representation for the wave rotating about x_0

$$u_0(x,t) = v_0(x_0 + \exp(-tS)(x - x_0)), x \in \mathbb{R}^d, t \geq 0.$$

The wave now satisfies $u_0(x,0) = v_0(x)$, and the shifted profile v_0 solves the equation

$$0 = A\Delta v_0 + v_{0,x}S(x - x_0) + f(v_0), x \in \mathbb{R}^d. \tag{2.102}$$

Hence rotating waves always come in families, a topic that we will take up in a more systematic way in Section 3.
The generalization of the complex valued system (2.80) to d dimensions is

$$u_t = \alpha \Delta u + g(|u|)u, \quad x \in \mathbb{R}^d, \quad u(x,y) \in \mathbb{C}, \tag{2.103}$$

where $\alpha \in \mathbb{C}, \mathrm{Re}(\alpha) > 0$ and $g : \mathbb{R} \to \mathbb{C}$ is smooth. Now the operator

$$F(u) = \alpha \Delta u + g(|u|)u$$

satisfies the equivariance condition

$$F(e^{i\theta}u) = e^{i\theta}F(u), \quad \theta \in \mathbb{R}.$$

Therefore we expect to find oscillating waves

$$u(x,t) = e^{-i\theta t}v(x). \tag{2.104}$$

Inserting this into (2.103) leads to the oscillating wave PDE

$$0 = \alpha \Delta v + i\theta v + g(|v|)v, \quad x \in \mathbb{R}^d. \tag{OWPDE}$$

Example 2.35 (The quintic Ginzburg-Landau equation)**.**

$$u_t = \alpha \Delta u + u(\mu + \beta |u|^2 + \gamma |u|^4). \tag{2.105}$$

The example shows a spinning soliton solution for the parameter settings

$$\alpha = \frac{1}{2}(1+i), \quad \mu = \frac{1}{2}, \quad \beta = \frac{5}{2} + i, \quad \gamma = -1 - \frac{1}{10}i, \tag{2.106}$$

whereas a perturbation to $\beta = \frac{13}{5} + i$ with all other parameters kept fixed, leads to a rotating spiral solution.

Example 2.36 ($\lambda - \omega$ system)**.** These systems are of the form (2.103) with

$$g(r) = \lambda(r) + i\omega(r), \quad \text{e.g.} \quad \lambda(r) = 1 - r^2, \quad \omega(r) = -r^2.$$

Of course this is a special case of a cubic Ginzburg Landau equation, but the name $\lambda - \omega$-system has become customary, in particular in mathematical Biology, see [22]. For the particular choice here, the simulation shows a rigidly rotating spiral wave developing.

We note an observation for the spinning wave solutions of Example 2.35. The simulations suggest that the associated profile $v_\star : \mathbb{R}^2 \to \mathbb{C}$ has the property

$$e^{i\theta}v_\star(x) = v_\star(R_\theta x), \quad x \in \mathbb{R}^2, \quad \theta \in \mathbb{R}, \tag{2.107}$$

where (compare (2.99))

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Therefore, the respresentation of the rotating wave is not unique, but can be given in two different ways

$$u(x,t) = e^{-i\theta t}v_\star(x) = v_\star(R_{-\theta t}x).$$

Finally, we consider the QCGL (2.105) in $d = 3$ space dimensions with the parameter setting from (2.106). We find rotating waves that are of the form

$$u(x,t) = v_{\star}(\exp(-tS)x), \quad x \in \mathbb{R}^3, \quad t \in \mathbb{R}$$

with some skew-symmetric matrix $S \in \mathbb{R}^{3,3}$. In the example shown, the wave rotates about the z-axis, hence we have

$$S = \begin{pmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \exp(-tS) = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) & 0 \\ -\sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3. Equivariant evolution equations, relative equilibria, and Lie groups

3.1. The concept of equivariance and relative equilibria. Before turning to the abstract concept of equivariance let us show that the waves observed in Section 2 may be viewed as steady states in an appropriate comoving frame.
Consider the initial value problem (also called the Cauchy problem) associated with the parabolic system (2.22),

$$\begin{aligned} u_t =& Au_{xx} + f(u), \quad x \in \mathbb{R}, \quad t \geq 0, \\ u(x,0) =& u_0(x), \quad x \in \mathbb{R}. \end{aligned} \tag{3.1}$$

If u solves this equation, then $v : \mathbb{R} \times [0,\infty) \to \mathbb{R}^m$ defined by

$$u(x,t) = v(x-ct,t) \quad \text{or} \quad v(\xi,t) = u(\xi+ct,t) \tag{3.2}$$

solves the equation

$$\begin{aligned} u_t(x,t) =& (-cv_{\xi} + v_t)(x-ct,t) = Au_{xx}(x,t) + f(u(x,t)) \\ =& Av_{\xi,\xi}(x-ct,t) + f(v(x-ct,t)). \end{aligned}$$

Replacing $x - ct = \xi \in \mathbb{R}$, we obtain that v solves the Cauchy problem

$$\begin{aligned} v_t =& v_{\xi,\xi} + cv_{\xi} + f(v), \quad \xi \in \mathbb{R}, \quad t \geq 0, \\ v(\xi,0) =& u_0(\xi), \quad \xi \in \mathbb{R}. \end{aligned} \tag{3.3}$$

We call this the *comoving frame* system. Note that a travelling wave $(v_{\star}, c_{\star})$ now appears as a steady state of (3.3). In particular, the stability of travelling waves (w.r.t. perturbations of initial values) for the original system is reduced to studying the stability of a steady state for the comoving frame equation. This will be the topic of Section 5.
We set up the comoving frame equations for our further examples from (2.80) and (2.93) in Section 2.
With

$$u(x,t) = e^{-i\theta t} v(x,t)$$

the comoving version of (2.80) is

$$v_t = \alpha v_{xx} + i\theta v + g(x, |v|)v. \tag{3.4}$$

For the reaction-diffusion system (2.93) in d dimensions the transformation

$$u(x,t) = v(\exp(-tS)x, t), \quad x \in \mathbb{R}^d, v(x,t) \in \mathbb{R}^m \tag{3.5}$$

leads to the comoving frame equation

$$v_t = A\Delta v + v_x Sx + f(v), \tag{3.6}$$

for which the profile of a rigidly rotating wave (2.100) appears as a steady state, see Proposition 2.33.

Exercise 3.1. Set up the comoving frame equation for the complex autonomous equation (2.91) such that travelling and oscillating waves appear as equilibria, cf. (TWOSODE).

In a nonrigorous sense, the underlying mechanism of these transformations is the following:
Consider a general evolution equation

$$u_t = F(u), \tag{3.7}$$

where $F : Y \subseteq X \to X$ with suitable function spaces $Y \subseteq X$. The operator F has the additional property that it commutes with some subgroup $\Gamma_X \subseteq GL(X)$ of linear homeomorphisms

$$F(\Gamma u) = \Gamma F(u), \quad u \in Y, \quad \Gamma \in \Gamma_X. \tag{3.8}$$

Equation (3.7) is then called *equivariant with respect to* Γ_X. A *relative equilibrium* of (3.7) is defined as a solution of (3.7) of the following form

$$u(t) = \Gamma(t)v_\star, \quad t \in \mathbb{R}, \tag{3.9}$$

where $v_\star \in Y$ and $\Gamma : \mathbb{R} \to \Gamma_X$ is a given function such that $u(t) = \Gamma(t)v_\star$ has some smoothness properties. It will be important in the following to require just this pathwise smoothness, rather than smoothness of the map $\Gamma : \mathbb{R} \to \Gamma_X \subseteq \mathrm{GL}(X)$ where $\mathrm{GL}(X)$ is equipped with the operator topology.

Example 3.2. Consider the system (2.22) and embed it into the setting above by defining

$$\begin{aligned} Y =& H^2(\mathbb{R}, \mathbb{R}^m), \quad X = L^2(\mathbb{R}, \mathbb{R}^m), \\ F(u) =& Au_{xx} + f(u), u \in X, \\ (\Gamma u)(x) =& u(x - \gamma), x \in \mathbb{R}, u \in X \quad \text{for some} \quad \gamma \in \mathbb{R}. \end{aligned} \tag{3.10}$$

Obviously, the translation operator Γ is an isometry of X. For the nonlinearity $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$ one needs the condition $f(0) = 0$ to show that $u(\cdot) \in H^1$ implies $f(u(\cdot)) \in L^2$ (cf. Example 3.37). With $\Gamma(t)v = v(\cdot - \gamma(t))$ a relative equilibrium is of the form

$$u(\cdot, t) = v(\cdot - \gamma(t)).$$

Since it solves the system (2.22) we find

$$0 = Av_{xx} + v_x\gamma_t(t) + f(v), \quad x \in \mathbb{R}.$$

Taking the inner product with v_x we obtain

$$0 = (v_x, Av_{xx})_{L^2} + (v_x, v_x)_{L^2}\gamma_t(t) + (v, f(v))_{L^2},$$

from which we conclude $\gamma_t(t) = c, t \in \mathbb{R}$ for some $c \in \mathbb{R}$ (note that $(v_x, v_x)_{L^2} > 0$, since the only constant function in $L^2(\mathbb{R}, \mathbb{R}^m)$ is identically zero). Hence, $\gamma(t) = ct + b$ for some $b \in \mathbb{R}$, and the relative equilibrium is a travelling wave

$$u(x,t) = v_\star(x - ct), \quad v_\star(\xi) = v(\xi - b), \quad \xi \in \mathbb{R}.$$

As in this example the group Γ_X of transformations is the image of a homomorphism (also called an *action*) $a : G \to GL(X)$ from a group $(G, \circ)$ into $GL(X)$, i.e. $a(g_1 \circ g_2) = a(g_1)a(g_2)$ for all $g_1, g_2 \in G$. In the case above, the group is $(G, \circ) = (\mathbb{R}, +)$ and the action assigns to any $\gamma \in \mathbb{R}$ the translation of functions defined by $a(\gamma)u(\cdot) = u(\cdot - \gamma), u \in L^2(\mathbb{R}, \mathbb{R}^m)$. Such an action is also called a *representation of G on X*.
More interesting examples occur with rotating waves, cf.(2.95), (2.96):

$$\begin{aligned} X &= L^2(\mathbb{R}^d, \mathbb{R}^m), \quad G = \mathrm{O}(\mathbb{R}^d) \quad \text{(the orthogonal group)} \\ [a(g)u](x) &= u(g^T x), \quad x \in \mathbb{R}^d, u \in X, g \in G. \end{aligned}$$

In this case the group is no longer a linear space but a manifold. Moreover, as the example shows, we want to take for fixed $u \in Y$ derivatives of the map

$$g \in G \to a(g)u \in X.$$

The appropriate mathematical framework for this is the theory of Lie groups and Lie algebras which in turn requires some familiarity with the calculus of manifolds. This will be the topic of the following sections.

3.2. Lie groups and manifolds. There are quite a few classical texts on the theory of Lie groups and Lie algebras, see [17],[6],[27],[30]. A particularly elementary and constructive approach to the subject via *linear matrix groups* is provided by [25], which we partly follow. The latter reference also contains a brief introduction into the analysis on manifolds. The general theory of Banach manifolds (where charts map into a Banach space rather than into a finite dimensional space) may be found in the rather abstract book [19]. For application of symmetries to partial differential equations we refer to [7], [13], [9].

Definition 3.3. *A group $(G, \circ)$ is called a Lie group if it is a finite-dimensional C^∞-manifold, for which the composition*

$$\text{comp} : \begin{array}{ccc} G \times G & \to & G, \\ (g, \gamma) & \to & g \circ \gamma \end{array} \tag{3.11}$$

and the inverse

$$\text{inv}: \begin{array}{ccc} G & \to & G, \\ g & \to & g^{-1} \end{array} \tag{3.12}$$

are C^∞-maps.

Remarks. 1. Instead of C^∞-manifolds one may also work in the category of C^ω-manifolds where coordinate transformations between different charts are required to be real analytic (i.e. have a locally convergent power series). This is the approach taken in [25], see Section 3.3.

2. Definition 3.3 implies that multiplication from the left $L_g = \text{comp}(g, \cdot)$ and from the right $R_\gamma = \text{comp}(\cdot, \gamma)$ are C^∞-maps. These operation are given by

$$L_g: \begin{array}{ccc} G & \to & G \\ \gamma & \to & g \circ \gamma = L_g(\gamma) \end{array}, \quad R_\gamma: \begin{array}{ccc} G & \to & G \\ g & \to & g \circ \gamma = R_\gamma(g) \end{array}. \tag{3.13}$$

Below we recall some notions from the analysis on manifolds. Let us first consider an example where the manifold structure is trivial.

Example 3.4. Let X be a real finite-dimensional Banach space and let

$$\text{GL}(X) = \{g \in L[X] : g \text{ is invertible}\} \tag{3.14}$$

be the set of linear automorphisms. Of course, we may identify X with $\mathbb{R}^k, k = \dim(X)$, $L[X]$ with the space of $k \times k$-matrices and $\text{GL}(X)$ with the set of invertible matrices. Then $G = \text{GL}(X)$ is a group with respect to composition (or multiplication of matrices), called the *general linear group on* X. Since G is an open subset of $L[X]$ it inherits its topology and the notions of differentiability from this finite dimensional space of dimension k^2 (in the setting of manifolds below, one has the charts (I_U, U), where $U \subseteq \text{GL}(X)$ is open and $I_U : U \to \text{GL}(X)$ is the identity on U). The composition $\text{comp}(g, \gamma) = g\gamma$ satisfies for $h_1, h_2 \in L[X]$

$$\text{comp}(g + h_1, \gamma + h_2) = \text{comp}(g, \gamma) + h_1\gamma + gh_2 + h_1h_2.$$

Since $h_1h_2 = \mathcal{O}(\|h_1\|^2 + \|h_2\|^2)$, we find the total derivative $d\,\text{comp}(g, \gamma) \in L[L[X] \times L[X], L[X]]$ to be

$$d\,\text{comp}(g, \gamma)(h_1, h_2) = h_1\gamma + gh_2, \tag{3.15}$$

and the partial derivatives $dL_g(\gamma) \in L[L[X], L[X]]$, $dR_\gamma(g) \in L[L[X], L[X]]$ to be as follows

$$dL_g(\gamma) = L_g, \ \forall \gamma \in G, \quad dR_\gamma(g) = R_\gamma, \ \forall g \in G.$$

Therefore the derivatives of L_g and R_γ are constant and their higher derivatives vanish. Furthermore, for the inverse we have from the geometric series for $h \in L[X]$ small,

$$\begin{aligned} \text{inv}(g + h) &= (g + h)^{-1} = (I + g^{-1}h)^{-1}g^{-1} \\ &= (I - g^{-1}h + \mathcal{O}(\|h\|^2))g^{-1} = \text{inv}(g) - g^{-1}hg^{-1} + \mathcal{O}(\|h\|^2), \end{aligned}$$

hence

$$[d\,\mathrm{inv}(g)]h = -g^{-1}hg^{-1}, \quad h \in L[X]. \tag{3.16}$$

We may also write this as $d\,\mathrm{inv}(g) = -L_{g^{-1}} \circ R_{g^{-1}}$.

Definition 3.5. *A C^k-manifold of dimension n is a Hausdorff topological space M together with a family of pairs $(U_\alpha, \varphi_\alpha), \alpha \in \mathfrak{A}$ with the following properties*

(M1) *For all $\alpha \in \mathfrak{A}$, $U_\alpha \subseteq M$, the map*

$$\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha) =: V_\alpha \subseteq \mathbb{R}^n$$

is a homeomorphism, and V_α is open in $\mathbb{R}^n$.

(M2) *For any pair $\alpha, \beta \in \mathfrak{A}$ with $U_\alpha \cap U_\beta \neq \emptyset$, the coordinate transformation*

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to V_\beta$$

is of class C^k.

(M3) $M = \bigcup_{\alpha \in \mathfrak{A}} U_\alpha$.

Some remarks concerning this classical definition are in order.

Remarks. 1. Charts: The pairs $(U_\alpha, \varphi_\alpha)$ are called *charts* (or *cordinate maps*) of the manifold. They play the same role for manifolds as geographical maps do for the earth. Correspondingly, every collection of charts satisfying (M1)-(M3), is also called an *atlas* of M. If an atlas contains all possible charts compatible with the given atlas by (M1),(M2), then it is called *maximal*. It is common to write the images of charts as

$$\varphi_\alpha(p) = x(p) = (x_1(p), x_2(p), \ldots, x_n(p)), \quad p \in U_\alpha,$$

and to call the maps $x_1, \ldots, x_n$ *local coordinates at* $p \in M$.

2. Smoothness: In the definition above we assume $k \geq 1$. From (M2) we find that $\varphi_\beta \circ \varphi_\alpha^{-1}$ has a C^k-inverse given by $\varphi_\alpha \circ \varphi_\beta^{-1}$ on $\varphi_\beta(U_\beta \cap U_\alpha)$, hence it is a diffeomorphism. If these coordinate transformations have smoothness C^∞ or even C^ω, then M is called a C^∞ resp. a C^ω*-manifold*.

3. Topology: Definition 3.5 assumes the topology on M to be given and to be Hausdorff. It is possible to avoid this assumption and assume that just a family of charts $(U_\alpha, \varphi_\alpha)$ is given such that $\varphi_\alpha : U_\alpha \subseteq M \to V_\alpha \subseteq \mathbb{R}^n$ are one-to-one and conditions (M2),(M3) hold. Then one can generate a topology on M by all pre-images $\varphi_\alpha^{-1}(V)$ with $\alpha \in \mathfrak{A}$ and $V \subseteq V_\alpha$ open. It remains to verify that the topology created in this way is Hausdorff. Note, however, that some references don't require M to be Hausdorff at all (cf. [19]).

4. Banach manifolds: Using the calculus of Frechét-derivatives for maps between Banach spaces (see e.g. [5]), one can easily replace the space $\mathbb{R}^n$ in Axioms (M1),(M2) of Definition 3.5 by an arbitrary Banach space X. In this way one obtains a so-called *Banach manifold modeled over* X, see [19]. This generalization

is in fact useful for our applications as the following example shows:
Consider a travelling front $u(x,t) = v_\star(x-ct)$ of (2.22) with

$$\lim_{\xi\to-\infty} v_\star(\xi) = v_- \neq v_+ = \lim_{\xi\to\infty} v_\star(\xi).$$

For this solution an appropriate function space is the *affine space*

$$M = v_\star + L^2(\mathbb{R},\mathbb{R}^m),$$

where $v_\star \in C_b^1(\mathbb{R},\mathbb{R}^m)$ is chosen such that

$$\lim_{\xi\to\pm\infty} v_\star(\xi) = v_\pm, \quad v_{\star,\xi} \in H^1(\mathbb{R},\mathbb{R}^m).$$

The charts consist of open sets $U \subseteq L^2(\mathbb{R},\mathbb{R}^m)$ and the mappings

$$\varphi_U : \begin{array}{ccc} v_\star + U & \to & U \\ v_\star + v & \to & v. \end{array}$$

Obviously, the coordinate transformation $\varphi_{U_1} \circ \varphi_{U_2}^{-1} = I_X$ is arbitrarily smooth on $\varphi_{U_1}((v_\star + U_1) \cap (v_\star + U_2)) = U_1 \cap U_2$. One may argue, that it is easier to reduce the affine case to the linear case by writing the PDE in terms of v rather than $u = v_\star + v$. However, this will make the right-hand side space-dependent and thus introduce complications with the representation of the shift group. Even in this simple case, it seems worthwhile to use the general calculus of Banach manifolds.

Definition 3.6. *Let M be an m-dimensional and N be an n-dimensional C^k-manifold with atlases $(\varphi_\alpha, U_\alpha), \alpha \in \mathfrak{A}$ and $(\psi_\beta, V_\beta), \beta \in \mathfrak{B}$, respectively. Then a function $f : M \to N$ is called of class C^l, $l \le k$, if f is continuous and all coordinate transformations satisfy*

$$\psi_\beta \circ f \circ \varphi_\alpha^{-1} \in C^l(\varphi_\alpha(U_\alpha \cap f^{-1}(V_\beta)), \mathbb{R}^n), \quad \alpha \in \mathfrak{A}, \beta \in \mathfrak{B}. \tag{3.17}$$

Of course, we do not only want to define differentiability but also the derivatives (or tangent maps) themselves. This, however, needs some preparation on tangent spaces (see below).
If M, N are manifolds as in Definition 3.6 then one defines the *product manifold* as the Hausdorff topological space

$$M \times N = \{(p,q) : p \in M, q \in N\},$$

endowed with the product topology and with the charts

$$\varphi_\alpha \otimes \psi_\beta : \begin{array}{ccc} U_\alpha \times V_\beta & \to & \mathbb{R}^m \times \mathbb{R}^n \\ (p,q) & \to & (\varphi_\alpha(p), \psi_\beta(q)), \end{array} \quad \alpha \in \mathfrak{A}, \beta \in \mathfrak{B}.$$

This construction obviously satisfies the properties (M1)-(M3) in Definition 3.5 and $M \times N$ becomes an $m+n$-dimensional manifold with the product atlas defined above.
Summarizing we have recalled all notions necessary for the abstract Definition 3.3 of a Lie group.

3.3. Lie groups of matrices and their Lie algebras. The most important type of Lie groups are subgroups of the general linear group from Example 3.4.

Definition 3.7. *Let X be a finite-dimensional Banach space. Every subgroup G of the general linear group* $\mathrm{GL}(X)$ *is called a linear group.*

Let $\|\cdot\|$ be the norm on X, then we use on $L[X]$ the associated operator norm $\|g\| = \sup\{\|gx\| : x \in X, \|x\| \leq 1\}$. Since X is finite-dimensional, all norms in X and $L[X]$ are equivalent.
Our goal is to recognize every linear group as a Lie group. Let us first consider some standard examples.

Example 3.8 (The orthogonal group)**.** The orthogonal group in $\mathbb{R}^d$ already appeared with rotating waves in (2.95), Lemma 2.31,

$$\mathrm{O}(\mathbb{R}^d) = \{g \in L[\mathbb{R}^d] : g^T g = I_d\}. \tag{3.18}$$

Since $|\det(g)| = 1$ for $g \in \mathrm{O}(\mathbb{R}^d)$, we can decompose $\mathrm{O}(\mathbb{R}^d)$ into two connected components, the *special linear group*

$$\mathrm{SO}(\mathbb{R}^d) = \{g \in L[\mathbb{R}^d] : g^T g = I_d,\ \det(g) = 1\}, \tag{3.19}$$

and its counterpart

$$\mathrm{SO}^-(\mathbb{R}^d) = \{g \in L[\mathbb{R}^d] : g^T g = I_d,\ \det(g) = -1\}.$$

Note that $\mathrm{SO}(\mathbb{R}^d)$ is a linear group itself, but $\mathrm{SO}^-(\mathbb{R}^d)$ is not. For $d = 2$, the group $\mathrm{SO}(\mathbb{R}^d)$ contains rotations

$$\mathrm{SO}(\mathbb{R}^2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in (-\pi, \pi] \right\}, \tag{3.20}$$

while $\mathrm{SO}^-(\mathbb{R}^2)$ contains reflections

$$\mathrm{SO}^-(\mathbb{R}^2) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} : \theta \in (-\pi, \pi] \right\}.$$

In dimensions $d \geq 3$ the group $\mathrm{SO}(\mathbb{R}^d)$ is no longer Abelian.

Example 3.9 (The Euclidean group)**.** Every affine transformation $T : \mathbb{R}^d \to \mathbb{R}^d$ that preserves the Euclidean distance (i.e. $|Tp_1 - Tp_2|_2 = |p_1 - p_2|_2\ \forall p_1, p_2 \in \mathbb{R}^d$) is of the form

$$Tp = Qp + b \quad \text{for some} \quad Q \in \mathrm{O}(\mathbb{R}^d), b \in \mathbb{R}^d.$$

The composition of two such transformations leads to

$$(T_1 \circ T_2)p = Q_1(T_2 p) + b_1 = Q_1(Q_2 p + b_2) + b_1 = Q_1 Q_2 p + Q_1 b_2 + b_1,$$

hence the group operation is

$$(Q_1, b_1) \circ (Q_2, b_2) = (Q_1 Q_2, Q_1 b_2 + b_1). \tag{3.21}$$

These transformations form the *Euclidean group*

$$\mathrm{E}(\mathbb{R}^d) = \{(Q, b) : Q \in \mathrm{O}(\mathbb{R}^d), b \in \mathbb{R}^d\}.$$

Taking $Q \in \mathrm{SO}(\mathbb{R}^d)$ one obtains the *special Euclidean group* $\mathrm{SE}(\mathbb{R}^d)$. Because of the group operation (3.21) (which differs from the standard operation on the product of two groups) the resulting group is called a *semidirect product* and written as

$$\mathrm{E}(\mathbb{R}^d) = \mathrm{O}(\mathbb{R}^d) \ltimes \mathbb{R}^d, \quad \mathrm{SE}(\mathbb{R}^d) = \mathrm{SO}(\mathbb{R}^d) \ltimes \mathbb{R}^d.$$

It is convenient to represent the group $\mathrm{SE}(\mathbb{R}^d)$ as a subgroup of $\mathrm{GL}(\mathbb{R}^{d+1})$,

$$\mathrm{SE}(\mathbb{R}^d) = \left\{ \begin{pmatrix} Q & b \\ 0 & 1 \end{pmatrix} : Q \in \mathrm{SO}(\mathbb{R}^d), b \in \mathbb{R}^d \right\}. \tag{3.22}$$

For the isomorphism between the two representations note that the matrix multiplication

$$\begin{pmatrix} Q_1 & b_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Q_2 & b_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} Q_1 Q_2 & Q_1 b_2 + b_1 \\ 0 & 1 \end{pmatrix}$$

is compatible with the group operation (3.21). We conclude that one may view an affine transformation in $\mathbb{R}^d$ as the restriction of a linear transformation in $\mathbb{R}^{d+1}$ from (3.22) to the affine subspace $\mathbb{R}^d \times \{1\}$ of $\mathbb{R}^{d+1}$.

For the construction of charts on a linear group we use the *exponential map* defined by

$$\exp : \begin{array}{ccc} L[X] & \to & \mathrm{GL}(X), \\ A & \to & \exp(A) = \sum_{j=0}^{\infty} \frac{1}{j!} A^j. \end{array} \tag{3.23}$$

In particular, we need the (local) inverse of this map, that is the logarithm of a matrix.

Proposition 3.10. The following properties hold:

(a) $\det(\exp(A)) = \exp(\mathrm{tr}(A))$ for all $A \in L[X]$,

(b) the exponential function maps a neighborhood of $0 \in L[X]$ homeomorphically onto a neighborhood of $I_X \in \mathrm{GL}(X)$. More precisely, the series

$$\log(B) = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} (B - I_X)^j \tag{3.24}$$

converges for $\|B - I_X\| < 1$. For every $A \in L[X]$ with $\|A\| < \log(2)$ we have $\|\exp(A) - I_X\| < 1$ as well as

$$\log(\exp(A)) = A, \tag{3.25}$$

and for every $B \in L[X]$ with $\|B - I_X\| < \frac{\log(2)}{1+\log(2)}$ we have $\|\log(B)\| < \log(2)$ as well as

$$\exp(\log(B)) = B. \tag{3.26}$$

Proof. Assertion (a) is a special case of Liouville's Theorem for linear ordinary differential systems $y'(t) = A(t)y(t), t \in \mathbb{R}$. If $Y(t)$ denotes a fundamental matrix of this system, then $d(t) = \det(Y(t))$ solves the differential equation $d'(t) = \mathrm{tr}(A(t))d(t)$. For the proof of (b), note that the geometric series $\sum_{j=1}^{\infty} \|B - I_X\|^j$

is a convergent majorant of the log-series (3.24) for $\|B - I_X\| < 1$. Moreover, for $\|A\| < \log 2$ we have

$$\|\exp(A) - I_X\| = \| \sum_{j=1}^{\infty} \frac{1}{j!} A^j \| \leq \sum_{j=1}^{\infty} \frac{\|A\|^j}{j!} = \exp(\|A\|) - 1 < 1.$$

Hence, $\log(\exp(A))$ is defined by (3.24), and the proof of the functional equation (3.25) can be transferred from the scalar to the matrix case. In a similar way, we have for $\|B - I_X\| < \frac{\log(2)}{1+\log(2)}$,

$$\|\log(B)\| \leq \sum_{j=1}^{\infty} \|B - I_X\|^j = \frac{\|B - I_X\|}{1 - \|B - I_X\|} < \log(2),$$

and (3.26) follows again as in the scalar case. Both relations (3.25),(3.26) together show that exp is a homeomorphism from a ball of radius $\log(2)$ about 0 onto its image. The image is a neighborhood of I_X since it contains a ball of radius $\log(2)1 + \log(2)$. $\square$

Remarks. 1. The logarithm may be defined as an analytic function in a much larger domain than the ball of radius 1 around I_X where the power series (3.24) converges. For example, it is well-known from complex analysis that the log-function has a holomorphic extension from the positive real axis $(0, \infty)$ to the slit domain $\mathbb{C} \setminus (-\infty, 0]$. This is called the *principal branch* of the log-function, it has the property

$$\log(z) = \log(r) + i\theta, \text{ where } z = r\exp(i\theta), r > 0, \theta \in (-\pi, \pi). \tag{3.27}$$

For any matrix $B \in L[X]$ with $\sigma(B) \cap (-\infty, 0] = \emptyset$ one can choose a simple closed contour $\mathcal{C}$ in $\mathbb{C} \setminus (-\infty, 0]$ which encloses the spectrum $\sigma(B)$ and define the logarithm of B via the so-called *functional calculus*

$$\log(B) = \frac{1}{2\pi i} \int_{\mathcal{C}} \log(z)(zI_X - B)^{-1} dz. \tag{3.28}$$

Let us show that (3.28) agrees with (3.24) if $\|B - I_X\| < 1$. In order to see this, use Cauchy's theorem and replace $\mathcal{C}$ in (3.28) by a circle $\mathcal{C}_r = \{z \in \mathbb{C} : |z - 1| = r\}$ with $\|B - I_X\| < r < 1$. The geometric series

$$(zI_X - B)^{-1} = (z-1)^{-1}(I_X - (z-1)^{-1}(B - I_X))^{-1} = \sum_{j=0}^{\infty} (z-1)^{-j-1}(B - I_X)^j$$

converges uniformly for $|z - 1| \geq r$. When we insert this into (3.28) (with $\mathcal{C}_r$ instead of $\mathcal{C}$), we can interchange the series with the integral and obtain $\log(B) = \sum_{j=0}^{\infty} b_j (B - I_X)^j$, where by Cauchy's formula,

$$b_j = \frac{1}{2\pi i} \int_{\mathcal{C}_r} \log(z)(z-1)^{-j-1} dz = \frac{1}{j!} \left[\frac{d^j}{dz^j} \log \right] (1) = \frac{(-1)^{j-1}}{j}.$$

The construction via (3.28) works in general for any simply connected domain in $\mathbb{C}$ for which a holomorphic extension exists. With this one can show that for every element $B \in \mathrm{GL}(X)$ there exists a linear operator A such that $\exp(A) = B$. Just take a suitable ray $\varrho_\theta = \{r\exp(i\theta) : r \geq 0\}$ which does not intersect $\sigma(B)$ and then use (3.28) with the holomorphic extension of log to $\mathbb{C} \setminus \varrho_\theta$. However, there is a caveat. The linear operator A need not be real in general (although the local power series (3.24) gives a real operator), and there is no holomorphic extension of this log-function to the whole of $\mathrm{GL}(X)$.
2. An alternative way of deriving (3.25) from the scalar complex case is via diagonalization. For example, assume A to be diagonalizable, i.e. $A = S\mathrm{diag}(\lambda_j)S^{-1}$, and obtain $|\lambda_j| \leq \|A\| < \log(2)$. The power series for exp shows $\exp(A) = S\mathrm{diag}(\exp(\lambda_j))S^{-1}$ and by equation (3.24),

$$\log(\exp(A)) = S\log(\mathrm{diag}(\exp(\lambda_j)))S^{-1} = S\mathrm{diag}(\log(\exp(\lambda_j)))S^{-1} = A.$$

In the general case one selects a sequence of diagonalizable matrices converging to the given matrix and uses a limit argument. A similar reasoning works for (3.26).

Next we show that the special orthogonal group $\mathrm{SO}(\mathbb{R}^d)$ from (3.19) is the image under the exponential function of the linear group of skew symmetric matrices

$$\mathrm{so}(\mathbb{R}^d) = \{S \in L[\mathbb{R}^d] : S^T = -S\}. \tag{3.29}$$

For the two-dimensional case (3.20) this follows from

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \exp(S_\theta) \quad \text{for} \quad S_\theta = \theta\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.30}$$

The relation (3.30) is easily seen by either using the series representations for $\sin, \cos, \exp$ or by diagonalizing S_θ. From Proposition 3.10 and (3.30) we conclude $S_\theta = \log(R_\theta)$ provided $\|S_\theta\|_2 = |\theta| < \log 2$.

Proposition 3.11. For every $g \in \mathrm{SO}(\mathbb{R}^d)$ there exists $S \in \mathrm{so}(\mathbb{R}^d)$ such that

$$g = \exp(S), \quad \sigma(S) \subseteq i[-\pi, \pi]. \tag{3.31}$$

If the eigenvalues $\exp(i\theta_j)$ of g satisfy $\max_j|\theta_j| < \log 2$, then (3.31) holds for $S = \log g$.

Proof. We use the normal form theorem for normal matrices (see [16, Cor.2.5.11]) and note that skew symmetric and orthogonal matrices are normal, i.e. $g^Tg = gg^T$): for every $g \in \mathrm{SO}(n)$ there exists $Q \in \mathrm{O}(\mathbb{R}^d)$ such that

$$Q^TgQ = \mathrm{diag}(R_{\theta_1}, \dots, R_{\theta_p}, \pm 1, \dots \pm 1),$$

where $\theta_j \in (0, \pi)$. Since $\det(g) = 1$ we infer that there is an even number of -1's among the $d - 2p$ values ± 1. Collecting these in matrices $R_\pi = -I_2$ and the remaining 1's in $R_0 = I_2$, we can write the normal form as

$$Q^TgQ = \mathrm{diag}(R_{\theta_1}, \dots, R_{\theta_k}, 1_\star),$$

where $\theta_j \in [0,\pi], j = 1,\dots,k = \lfloor \frac{d}{2} \rfloor$ and

$$1_\star = \begin{cases} 1, & d \text{ odd}, \\ \emptyset, & d \text{ even}. \end{cases}$$

From (3.30) we obtain $Q^T g Q = \exp(\widetilde{S})$ for

$$\widetilde{S} = \operatorname{diag}(S_{\theta_1},\dots,S_{\theta_k},0_\star), \quad 0_\star = \begin{cases} 0, & d \text{ odd}, \\ \emptyset, & d \text{ even}. \end{cases}$$

Finally, $S = Q\widetilde{S}Q^T$ is in $\mathrm{so}(\mathbb{R}^d)$ and satisfies

$$\exp(S) = Q\exp(\widetilde{S})Q^T = g.$$

By Proposition 3.10 the matrix S above agrees with $\log(g)$ defined in (3.24), if

$$\log 2 > \|S\|_2 = \|Q\widetilde{S}Q^T\|_2 = \|\widetilde{S}\|_2 = \max_{j=1,\dots,k}|\theta_j|.$$

□

Let us define the tangent space of a Lie group without recourse to the general theory of manifolds. The connection to this theory will be discussed in Section 3.4.

Definition 3.12. *Let $G \subseteq \mathrm{GL}(X)$ be a linear group with unit element $\mathbb{1}$. We define the tangent space of G at $\mathbb{1}$ to be*

$$T_{\mathbb{1}}G = \{(\mathbb{1},v) : \exists\, \varepsilon > 0, g \in C^1((-\varepsilon,\varepsilon),G) \text{ with } g(0) = \mathbb{1}, g'(0) = v\}. \tag{3.32}$$

The subset

$$\mathfrak{g} = \{v \in L[X] : (\mathbb{1},v) \in T_{\mathbb{1}}G\} \tag{3.33}$$

is called the Lie algebra of G.

Remark 3.13. *The element $\mathbb{1}$ in the pairs of $T_{\mathbb{1}}G$ is included in order to keep track of the point where the tangent space is located. This will be important when we consider the tangent space T_gG at an arbitrary point $g \in G$, see Section 3.4. Of course we can (and actually will) identify the Lie algebra with the tangent space $T_{\mathbb{1}}G$.*

The notion of Lie algebra is motivated by the following proposition.

Proposition 3.14. The Lie algebra $\mathfrak{g}$ of a linear group G has the following properties.

(i) $\mathfrak{g}$ is a real vector space,
(ii) if $\mu_1,\mu_2 \in \mathfrak{g}$, then the commutator satisfies

$$[\mu_1,\mu_2] := \mu_1\mu_2 - \mu_2\mu_1 \in \mathfrak{g}, \tag{3.34}$$

(iii) The bracket $[\cdot,\cdot]$ is skew symmetric $[\mu_1,\mu_2] = -[\mu_2,\mu_1]$ and satisfies the *Jacobi identity*

$$[[\mu_1,\mu_2],\mu_3] + [[\mu_2,\mu_3],\mu_1] + [[\mu_3,\mu_1],\mu_2] = 0 \quad \forall \mu_1,\mu_2,\mu_3 \in \mathfrak{g}. \tag{3.35}$$

Remark 3.15. *An abstract Lie algebra is a real vector space V endowed with a skew-symmetric bilinear bracket operation*

$$[\cdot,\cdot] : V \times V \to V,$$

which satisfies the Jacobi identity.

Proof. **(i):** Let $g_1(\cdot), g_2(\cdot)$ be C^1-curves in G with

$$g_j(0) = \mathbb{1}, \quad g_j'(0) = \mu_j, \quad j = 1, 2.$$

For $\alpha_1, \alpha_2 \in \mathbb{R}$ consider the curve $g_3(t) = g_1(\alpha_1 t)g_2(\alpha_2 t)$ for $|t|$ small and note $g(0) = \mathbb{1}$ as well as

$$g'(t) = \alpha_1 g_1'(\alpha_1 t)g_2(\alpha_2 t) + \alpha_2 g_1(\alpha_1 t)g_2'(\alpha_2 t), \quad g'(0) = \alpha_1\mu_1 + \alpha_2\mu_2,$$

hence $\alpha_1\mu_1 + \alpha_2\mu_2 \in \mathfrak{g}$.
(ii): Let g_1, g_2 be as above and consider for fixed $|\tau|$ small

$$g(t) = g_1(\tau)g_2(t)g_1(\tau)^{-1}, \quad |t| \text{ small}.$$

Then $g(0) = \mathbb{1}$ and $g'(0) = g_1(\tau)g_2'(0)g_1(\tau)^{-1} = g_1(\tau)\mu_2 g_1(\tau)^{-1}$. Setting $\lambda(\tau) = g_1(\tau)\mu_2 g_1(\tau)^{-1}$ we obtain from (3.16)

$$\lambda'(\tau) = g_1'(\tau)\mu_2 g_1(\tau)^{-1} - g_1(\tau)\mu_2 g_1(\tau)^{-1}g_1'(\tau)g_1(\tau)^{-1},$$

and therefore $\lambda'(0) = \mu_1\mu_2 - \mu_2\mu_1 \in \mathfrak{g}$.
(iii) The Jacobi identity is verified by using the definition (3.34) in the expression on the left-hand side of (3.35) and expanding terms. □

Corollary 3.16. *The Lie algebra of the special orthogonal group* $\mathrm{SO}(\mathbb{R}^d)$ *is the additive group* $\mathrm{so}(\mathbb{R}^d)$ *of skew-symmetric matrices.*

Proof. Let $g(\cdot)$ be a C^1-curve in $\mathrm{SO}(\mathbb{R}^d)$ with $g(0) = I_d$. Differentiating $g(t)^T g(t) = I_d$ at $t = 0$ yields

$$0 = g'(t)^T g(t) + g^T(t)g'(t), \quad 0 = g'(0)^T + g'(0),$$

hence $g'(0) \in \mathrm{so}(\mathbb{R}^d)$. Conversely, let $S \in \mathrm{so}(\mathbb{R}^d)$ and define $g(t) = \exp(tS), t \in \mathbb{R}$. Then we have $g(0) = I_d$ and since S and $S^T = -S$ commute,

$$g(t)^T g(t) = \exp(tS)^T \exp(tS) = \exp(tS^T)\exp(tS) = \exp(t(S + S^T)) = I_d.$$

Finally, $\det(\exp(S)) = \exp(\mathrm{tr}(S)) = 1$ by Proposition 3.10 (a). □

The following Theorem shows the general role of the exponential function.

Theorem 3.17. *Let $G \subseteq \mathrm{GL}(X)$ be a linear group and $\mathfrak{g}$ be its Lie algebra. Then* $\exp$ *maps $\mathfrak{g}$ into G.*

Proof. Let $n = \dim(\mathfrak{g})$ and let $\{\mu_1, \dots, \mu_n\}$ be a basis of $\mathfrak{g}$. Further, let $g_j(\cdot)$ be a C^1-curve in G with tangent vector μ_j at $\mathbb{1}$, and define the mapping g by

$$g : \begin{array}{ccc} \mathbb{R}^n & \to & G, \\ x = (x_j)_{j=1}^n & \to & g_1(x_1)\cdots g_n(x_n). \end{array}$$

Then we choose a subspace Y of $L[X]$ with $L[X] = \mathfrak{g} \oplus Y$ and extend g to the mapping

$$f: \begin{array}{ccc} \mathbb{R}^n \otimes Y & \to & L[X], \\ (x,y) & \to & g(x)(I_X + y). \end{array}$$

Clearly, f satisfies $f(0,0) = I_X$ and a straightforward calculation shows

$$Df(0,0)(x,y) = \sum_{j=1}^{n} x_j \mu_j + y, \quad x \in \mathbb{R}^n, y \in Y.$$

By our choice of Y, the linear map $Df(0,0)$ is invertible. The inverse function theorem shows that f has a C^∞-smooth local inverse

$$f^{-1}(B) = (f_1^{-1}(B), f_2^{-1}(B)) \in \mathbb{R}^n \times Y, \quad B \in U(I_X) \subseteq L[X].$$

In particular, by the definition of f,

$$f_2^{-1}(g(x+\tau z)(I_X + y)) = y, \quad x, z \in \mathbb{R}^n, y \in Y, \tau \in \mathbb{R} \quad \text{small.} \tag{3.36}$$

Differentiating (3.36) at $\tau = 0$, we obtain for small $x \in \mathbb{R}^n, y \in Y$,

$$Df_2^{-1}(g(x)(I_X + y))(D_x g(x) z)(I_X + y) = 0. \tag{3.37}$$

This relation is linear in z, hence it does not only hold for small z but for all $z \in \mathbb{R}^n$.

For $z \in \mathbb{R}^n$ the map $\tau \to g(x + \tau z) g(x)^{-1}$ defines a smooth curve in G passing through $\mathbb{1}$ with tangent vector

$$A(x)z := \left[\frac{d}{d\tau} g(x+\tau z) g(x)^{-1}\right]_{|\tau=0} = (D_x g(x) z) g(x)^{-1} \in \mathfrak{g}. \tag{3.38}$$

For $x = 0$ we obtain the linear map $A(0) : \mathbb{R}^n \to \mathfrak{g}$,

$$A(0)z = D_x g(0) z = \sum_{j=1}^{n} z_j \mu_j.$$

This map is invertible, hence $A(x) : \mathbb{R}^n \to \mathfrak{g}$ is also invertible for small x by the Banach perturbation lemma. With $A(x)$ from (3.38) we can write (3.37) as

$$Df_2^{-1}(g(x)(I_X + y))(A(x)z) g(x)(I_X + y) = 0$$

for all $z \in \mathbb{R}^n$ and small $x \in \mathbb{R}^n$, $y \in Y$. Since $A(x)$ is invertible we can replace $A(x)z$ by an arbitrary element $\mu \in \mathfrak{g}$, and since f is a local diffeomorphism, we can replace $g(x)(I_X + y)$ by an arbitrary element B in a neighborhood $U(I_X)$:

$$Df_2^{-1}(B)\mu B = 0 \quad \forall \mu \in \mathfrak{g}, B \in U(I_X). \tag{3.39}$$

For a given $\mu \in \mathfrak{g}$ consider now the curve $B(\tau) = \exp(\tau\mu)$ in $L[X]$. From (3.39) we find for small $|\tau|$,

$$0 = Df_2^{-1}(\exp(\tau\mu))\mu \exp(\tau\mu) = \frac{d}{d\tau}\left[f_2^{-1}(\exp(\tau\mu))\right],$$

hence $f_2^{-1}(\exp(\tau\mu)) = f_2^{-1}(\exp(0\mu)) = f_2^{-1}(I_X) = 0$. By the definition of f, this shows $\exp(\tau\mu) = g(f_1^{-1}(\exp(\tau\mu))) \in G$ for $|\tau|$ in some interval $(-\tau_0, \tau_0)$. For a general $\tau \in \mathbb{R}$ we take $k \in \mathbb{N}$ such that $\frac{|\tau|}{k} < \tau_0$ and find from the functional equation of the exponential function

$$\exp(\tau\mu) = \exp(k\frac{\tau}{k}\mu) = \Big(\exp(\frac{\tau}{k}\mu)\Big)^k \in G.$$

□

Note that this proof uses a local coordinate system near $\mathbb{1}$ (defined by f), where elements of the group G are identified by a vanishing component of the inverse ($f_2^{-1}(B) = 0$). Despite this construction, we do not consider the Lie group G as a submanifold of $GL(X)$. In particular, the topology on G will not be defined as the relative topology of G in $\mathrm{GL}(X)$. The reason is, that a neighborhood of $\mathbb{1}$ in $\mathrm{GL}(X)$ may contain further pieces of the group that do not arise as images of points near zero under f. The following example illustrates this situation.

Example 3.18. With the rotation matrices from (3.30) and two real numbers $a, b > 0$ define the linear group

$$G_{a,b} = \left\{ M(\tau) = \begin{pmatrix} R_{a\tau} & 0 \\ 0 & R_{b\tau} \end{pmatrix} : \tau \in \mathbb{R} \right\} \subseteq \mathrm{GL}(\mathbb{R}^4). \tag{3.40}$$

Geometrically, we may think of this group as a curve on the 2-torus

$$\mathbb{T}^2 = \left\{ \begin{pmatrix} R_{\tau_1} & 0 \\ 0 & R_{\tau_2} \end{pmatrix} : \tau_1, \tau_2 \in \mathbb{R} \right\} \subseteq \mathrm{GL}(\mathbb{R}^4),$$

which rotates with speed a about one axis and with speed b about the second. If $\frac{a}{b} \in \mathbb{Q}$ then the curve is closed (and so is the Lie group) and $G_{a,b}$ may be considered as a submanifold of $\mathbb{T}^2$. However, if $\frac{a}{b} \notin \mathbb{Q}$, then $G_{a,b}$ is dense in $\mathbb{T}^2$, and any $\mathrm{GL}(\mathbb{R}^4)$-neigborhood of a fixed element $M(\tau) \in G_{a,b}$ contains an infinity of branches of $G_{a,b}$ which accumulate on $M(\tau)$.

One can in fact show that the group topology and the relative topology of G in $L[X]$ coincide if the group G is closed (see the Closed subgroup theorem [25, Section 2.7]).
Keeping this warning in mind, we now define a topology and charts on a linear group G which turn G into a C^∞-manifold (in fact, it is of type C^ω). This is achieved by using the relative topology of $\mathfrak{g}$ in $L[X]$ and the exponential map for parametrization.
Let us write balls in $\mathfrak{g}$ as

$$B^{\mathfrak{g}}(0, \varepsilon) = \{\mu \in \mathfrak{g} : \|\mu\| < \varepsilon\}.$$

Definition 3.19. *Let G be a linear group. A set $\mathcal{U} \subseteq G$ is called a neighborhood of an element $g \in G$ if there exists an $\varepsilon > 0$ such that*

$$U_\varepsilon(g) := \{g\exp(\mu) : \mu \in B^{\mathfrak{g}}(0, \varepsilon)\} \subseteq \mathcal{U}.$$

A set $M \subseteq G$ is called open if it is a neighborhood of each of its elements.

It is not difficult to verify that this generates a topology on G, called the *group topology*. Moreover, the topology is Hausdorff: take two different elements $g_1, g_2 \in G$ and then estimate for $\mu_1, \mu_2 \in \mathfrak{g}$,

$$\|g_1 \exp(\mu_1) - g_2 \exp(\mu_2)\| \geq \|g_1 - g_2\| - \|g_1(\mathbb{1} - \exp(\mu_1))\| - \|g_2(\mathbb{1} - \exp(\mu_2))\|$$

Since the exponential function is continuous, the right hand side becomes positive if $\|\mu_1\|$ and $\|\mu_2\|$ are sufficiently small. Hence $U_\varepsilon(g_1)$ and $U_\varepsilon(g_2)$ do not intersect for small ε.
For a chart on $U_\varepsilon(\mathbb{1})$ with $\varepsilon < \log(2)$ we use the local log-function from Proposition 3.10

$$\varphi_{\varepsilon,\mathbb{1}} = \log : \begin{array}{ccc} U_\varepsilon(\mathbb{1}) & \to & B^{\mathfrak{g}}(0,\varepsilon) \\ \gamma = \exp(\mu) & \to & \log(\gamma) = \mu \end{array}.$$

The map is one-to-one by Proposition 3.10. A chart near an arbitrary point $g \in G$ is given by

$$\varphi_{\varepsilon,g} = \varphi_{\varepsilon,\mathbb{1}} \circ L_{g^{-1}} : U_\varepsilon(g) \to B^{\mathfrak{g}}(0,\varepsilon), \quad \varphi_{\varepsilon,g}(g\exp(\mu)) = \mu. \tag{3.41}$$

A coordinate transformation (see condition (M2))

$$\varphi_{\varepsilon_2,g_2} \circ \varphi_{\varepsilon_1,g_1}^{-1} = \varphi_{\varepsilon_2,g_2} \circ L_{g_2^{-1}} \circ L_{g_1} \circ \varphi_{\varepsilon_1,g_1}^{-1} = \log(g_2^{-1} g_1 \exp(\cdot))$$

is of type C^∞ on its domain of definition $\varphi_{\varepsilon_1,g_1}(U_{\varepsilon_1}(g_1) \cap U_{\varepsilon_2}(g_2))$ which is an open subset of $B^{\mathfrak{g}}(0,\varepsilon_1)$. Finally, the covering property $G = \bigcup_{g\in G, \varepsilon < \log 2} U_\varepsilon(g)$ is obvious. Let us summarize the result.

Proposition 3.20. With the topology according to Definition 3.19 and the charts defined in (3.41) every linear group G in $GL[X]$ is a Lie group.

3.4. Tangent maps, tangent bundle, and flows on manifolds. We proceed with the general theory of manifolds and relate it to the previous construction for linear groups.
Let us first define the tangent space. Consider an n-dimensional C^1-manifold M and a point $p \in M$. Suppose that $(U_\alpha, \varphi_\alpha)$ is a chart with $p \in U_\alpha$ and $v(t) \in M, |t| < \varepsilon$ is a C^1-curve in M with $v(0) = p$ (see Definition 3.6 for the meaning of C^1). Such a curve always exists, since one can take $v(t) = \varphi_\alpha^{-1}(z(t))$ where $z(t), |t| < \varepsilon$ is a smooth curve in $V_\alpha = \varphi_\alpha(U_\alpha) \subseteq \mathbb{R}^n$ satisfying $z(0) = \varphi_\alpha(p)$). Then we form the tangent vector of the image curve

$$\frac{d}{d\tau}(\varphi_\alpha(v(\tau)))_{|\tau=0} \in \mathbb{R}^n.$$

If two such curves $v_j \in C^1((-\varepsilon_j, \varepsilon_j)M), j = 1,2$ satisfy $\frac{d}{d\tau}(\varphi_\alpha \circ v_1(\tau))_{|\tau=0} = \frac{d}{d\tau}(\varphi_\alpha \circ v_2(\tau))_{|\tau=0}$, then we call them equivalent and write $v_1 \sim v_2$ (we suppress the dependence of this relation on p). This notion is independent of the chart containing p, since for another index β we have by the chain rule

$$\begin{aligned}\frac{d}{d\tau}(\varphi_\beta\circ v(\tau))_{|\tau=0} =&\frac{d}{d\tau}(\varphi_\beta\circ\varphi_\alpha^{-1}\circ\varphi_\alpha\circ v(\tau))_{|\tau=0}\\ =&D(\varphi_\beta\circ\varphi_\alpha^{-1})(\varphi_\alpha(p))\frac{d}{d\tau}(\varphi_\alpha\circ v(\tau))_{|\tau=0}.\end{aligned}\tag{3.42}$$

Note that the transformation matrix $D(\varphi_\beta\circ\varphi_\alpha^{-1})(\varphi_\alpha(p))$ is invertible and independent of v.

Definition 3.21. *For an n-dimensional manifold M and $p\in M$ the set*

$$T_pM=\{(p,[v]_\sim): v\in C^1((-\varepsilon,\varepsilon),M), v(0)=p\}\tag{3.43}$$

with the equivalence class

$$[v]_\sim=\{w\in C^1((-\varepsilon,\varepsilon),M):\varepsilon=\varepsilon(w)>0, w(0)=p, v\sim w\}$$

is called the tangent space of M at p.

Note that the domain $(-\varepsilon(w),\varepsilon(w))$ of the curve may vary within the equivalence class $[v]_\sim$.
We turn T_pM into a linear space with the help of the map

$$d\varphi_\alpha(p): \begin{array}{ccc} T_pM & \to & \mathbb{R}^n\\ (p,[v]_\sim) & \to & \frac{d}{d\tau}(\varphi_\alpha(v(\tau)))_{|\tau=0}.\end{array}\tag{3.44}$$

By construction, this map is well-defined and one-to-one. It is also onto, since for a given $v_0\in\mathbb{R}^n$ we can take $v(\tau)=\varphi_\alpha^{-1}(\varphi_\alpha(p)+\tau v_0)$ and then find $d\varphi_\alpha(p)(p,[v]_\sim)=\frac{d}{d\tau}(\varphi_\alpha(p)+\tau v_0)_{|\tau=0}=v_0$. Hence $d\varphi_\alpha(p)$ is a bijection. For $a,b\in\mathbb{R}$ and $(p,[v_1]_\sim),(p,[v_2]_\sim)\in T_pM$ one defines

$$a\,(p,[v_1]_\sim)+b\,(p,[v_2]_\sim)=d\varphi_\alpha(p)^{-1}(a\,d\varphi_\alpha(p)(p,[v_1]_\sim)+b\,d\varphi_\alpha(p)(p,[v_2]_\sim)).\tag{3.45}$$

This definition turns out to be independent of the chart at p, and in this way T_pM becomes a vector space isomorphic to $\mathbb{R}^n$. Therefore, it is convenient to suppress the notation of equivalence classes and simply write $(p,v)\in T_pM$ (or even $v\in T_pM$) instead of $(p,[v]_\sim)\in T_pM$.

Definition 3.22. *Let M be an n-dimensional C^1-manifold. Then*

$$TM=\bigcup_{p\in M}T_pM\tag{3.46}$$

is called the tangent bundle of M.

The tangent bundle TM can be given the structure of a manifold by defining the charts $(\mathcal{U}_\alpha,\Phi_\alpha),\alpha\in\Lambda$ as follows

$$\begin{aligned}\mathcal{U}_\alpha =&\{(p,v): p\in U_\alpha,(p,v)\in T_pM\}\\ \Phi_\alpha(p,v)=&(\varphi_\alpha(p),d\varphi_\alpha(p)(p,v))\in\mathbb{R}^n\times\mathbb{R}^n,\ (p,v)\in\mathcal{U}_\alpha.\end{aligned}$$

Note that

$$\Phi_\alpha : \mathcal{U}_\alpha \to V_\alpha \times \mathbb{R}^n$$

is bijective. Moreover the change of coordinates is

$$\Phi_\beta \circ \Phi_\alpha^{-1} : \begin{array}{lcl} (V_\alpha \cap V_\beta) \times \mathbb{R}^n & \to & V_\beta \times \mathbb{R}^n \\ (x, y) & \to & (\varphi_\beta \circ \varphi_\alpha^{-1}(x),\ d\varphi_\beta(p)(p, [d\varphi_\alpha(p)^{-1}(y)]_2)), \end{array}$$

where $p = \varphi_\alpha^{-1}(x)$ and $[\cdot]_2$ indicates the second component of the preimage. From the chain rule (3.42) we finally obtain the formula

$$\Phi_\beta \circ \Phi_\alpha^{-1}(x, y) = (\varphi_\beta \circ \varphi_\alpha^{-1}(x), D(\varphi_\beta \circ \varphi_\alpha^{-1})(x)y). \tag{3.47}$$

This formula is suggestive to keep in mind because it combines the original coordinate change with its derivative applied to a tangent vector. If M is a C^k-manifold then the tangent bundle TM becomes a C^{k-1}-manifold. In this way one can continue to form higher order tangent bundles $T^2M = T(TM)$ and so on.

Definition 3.23. *Let M,N be C^k-manifolds ($k \geq 1$) with tangent bundles TM, TN and let $f : M \to N$ be of class C^1 according to Definition 3.6. Then the tangential of f at $p \in M$ (also called the derivative of f at $p \in M$) is defined as the map*

$$df(p) : \begin{array}{lcl} T_pM & \to & T_{f(p)}N \\ (p, [v]_\sim) & \to & (f(p), [f(v)]_\sim) \end{array} \tag{3.48}$$

Remark 3.24. *By Definition 3.6 a C^1-curve $v(\cdot)$ in M with $v(0) = p$ is mapped into a C^1-curve $w(\cdot) = f(v(\cdot))$ in N satisfying $w(0) = f(p)$. Therefore, $(f(p), [f(v)]_\sim)$ is an element of $T_{f(p)}N$.*
In the literature the notation T_pf is also frequently used for the tangential map of f at p. However, we prefer to write $df(p)$ since the tangential map is a direct generalization of the total derivative to manifolds. For example, let $M \subseteq \mathbb{R}^m$ and $N \subseteq \mathbb{R}^n$ be open subsets with the trivial charts $\varphi_m = I_m$ and $\varphi_n = I_n$, respectively. Then we find

$$df(p)(p, v) = (f(p), Df(p)v), \quad v \in \mathbb{R}^m \tag{3.49}$$

with the Jacobian $Df(p) \in L[\mathbb{R}^m, \mathbb{R}^n]$ of f at p, since

$$\frac{d}{d\tau}\left[\varphi_n \circ f(\varphi_m^{-1}(\varphi_m(p) + \tau v))\right]_{|\tau=0} = \frac{d}{d\tau} f(p + \tau v)_{|\tau=0} = Df(p)v.$$

Definition 3.25. *Let M be a C^k-manifold where $1 \leq k \leq \infty$. A C^r-vector field on M with $0 \leq r \leq k-1$ is a map*

$$f : \begin{array}{lcl} M & \to & TM \\ p & \to & f(p) \end{array}, \tag{3.50}$$

which satisfies $f(p) \in T_pM$ for all $p \in M$ and which is of type C^r.

With this notion we can consider initial value problems on the manifold M

$$p'(\tau) = f(p(\tau)), \tau \in \mathbb{R}, \quad p(0) = p_0, \tag{3.51}$$

where $p_0 \in M$ is given and f is a C^1-vector field on M. Note that we use the suggestive notation $p'(\tau)$ in (3.51) instead of the tangential map $dp(\tau)$ as in Definition 3.23. Moreover, the differential equation in (3.51) is autonomous since the vector field does not depend explicitly on time.
The following theorem is the extension of the classical local Picard-Lindelöf theorem to initial value problems on manifolds.

Theorem 3.26. *Let M be a C^k-manifold ($k \geq 2$) and $f : M \to TM$ be a C^{k-1}-vector field. Then the for any $p_0 \in M$ there exists an $\varepsilon > 0$ sucht that the initial value problem (3.51) has a unique solution $p \in C^{k-1}((-\varepsilon, \varepsilon), M)$.*

A proof in the analytical case may be found in [25, Section 4.3]. The reference [19, IV, § 2] provides not only the local theorem 3.26 for Banach manifolds but also the standard further results such as maximal continuation of local solutions, continuous and differentiable dependence on initial conditions etc. The proof essentially works by applying the classical theorem in the coordinates provided by the charts.
For our purposes it is important to study flows on Lie groups.

3.5. The exponential function for general Lie groups. Our aim is to generalize the exponential function via the differential equation

$$\frac{d}{dt}\exp(tA) = \exp(tA)A, \quad \exp(0) = I. \tag{3.52}$$

Note that this differential equation has a natural meaning in the setting of a linear group and its Lie algebra where the right-hand side of (3.52) is just composition in $L[X]$ (cf. (3.23)). However, in the setting of an abstract group we must define the vector field in (3.52) in a proper way, i.e. the multiplication of elements from G with elements of the Lie algebra $\mathfrak{g}$.

Definition 3.27. *Let $(G, \circ)$ be a Lie group, then the tangent space $T_{\mathbb{1}}G$ of G at $\mathbb{1}$ is called the Lie algebra associated with G and denoted by $\mathfrak{g}$. For every $\mu \in \mathfrak{g}$ we call*

$$g\mu = dL_g(\mathbb{1})\mu, \quad \mu g = dR_g(\mathbb{1})\mu. \tag{3.53}$$

the vector fields induced by right resp. left multiplication with μ.

Let us comment on this definition. The name of a Lie algebra will be justified by Proposition 3.28 below.
According to Definition 3.21, an element $\mu \in \mathfrak{g}$ is of the form $(\mathbb{1}, [v]_\sim)$ with $v(\cdot)$ a path in G passing through $\mathbb{1}$ at $t = 0$. This notation slightly deviates from the notation in Definition 3.12 for linear groups where the element $\mathbb{1}$ was excluded from the symbol. However, as in the remark following Definition 3.12 we can always identify $\mathfrak{g}$ with $\{v : (\mathbb{1}, v) \in \mathfrak{g}\}$

We also note that $dL_g(\mathbb{1})$ is a linear map from $T_{\mathbb{1}}G = \mathfrak{g}$ into T_gG (cf. Definition 3.23), so that (3.53) defines a smooth vector field on G. Using $\mu = (\mathbb{1}, [v]_\sim)$ we may also write the vector field as follows

$$g\mu := \frac{d}{d\tau}(g \circ v(\tau))_{|\tau=0}, \quad \mu g := \frac{d}{d\tau}(v(\tau) \circ g)_{|\tau=0}. \tag{3.54}$$

The following proposition is the analog of Proposition 3.14.

Proposition 3.28. Let $(G, \circ)$ be a Lie group with tangent space $T_{\mathbb{1}}G = \{\mathbb{1}\} \times \mathfrak{g}$. Then the following holds:

(i) For $\mu_j = (\mathbb{1}, [v_j]_\sim) \in \mathfrak{g}, \alpha_j \in \mathbb{R}, j = 1, 2$ the linear combination may be written as

$$\alpha_1\mu_1 + \alpha_2\mu_2 = \frac{d}{d\tau}\left(v_1(\alpha_1\tau) \circ v_2(\alpha_2\tau)\right)_{|\tau=0}.$$

(ii) The vector space $\mathfrak{g}$ is a Lie algebra with respect to the bracket

$$[\mu_1, \mu_2] = \frac{d}{d\tau}\left(v_1(\tau)\mu_2 v_1(\tau)^{-1}\right)_{|\tau=0}.$$

Proof. **(i):** Let φ_α be a chart at $\mathbb{1}$, then the definition (3.45) implies

$$\begin{aligned}
\alpha_1\mu_1 + \alpha_2\mu_2 =& d\varphi_\alpha(\mathbb{1})^{-1}(\alpha_1 d\varphi_\alpha(\mathbb{1})\mu_1 + \alpha_2 d\varphi_\alpha(\mathbb{1})\mu_2)\\
=& d\varphi_\alpha(\mathbb{1})^{-1}(\alpha_1\frac{d}{d\tau}(\varphi_\alpha(v_1(\tau))) + \alpha_2\frac{d}{d\tau}(\varphi_\alpha(v_2(\tau))))_{|\tau=0}\\
=& d\varphi_\alpha(\mathbb{1})^{-1}\frac{d}{d\tau}(\varphi_\alpha(v_1(\alpha_1\tau)) + (\varphi_\alpha(v_2(\alpha_2\tau)))_{|\tau=0}\\
=& d\varphi_\alpha(\mathbb{1})^{-1}\frac{d}{d\tau}(\varphi_\alpha(v_1(\alpha_1\tau) \circ v_2(\alpha_2\tau)))_{|\tau=0}\\
=& d\varphi_\alpha(\mathbb{1})^{-1}d\varphi_\alpha(\mathbb{1})\frac{d}{d\tau}(v_1(\alpha_1\tau) \circ v_2(\alpha_2\tau))_{|\tau=0}\\
=& \frac{d}{d\tau}(v_1(\alpha_1\tau) \circ v_2(\alpha_2\tau))_{|\tau=0}.
\end{aligned}$$

(ii): Similar to the proof of Proposition 3.14 (ii), one considers for fixed τ the curve $v(t) = v_1(\tau)v_2(t)v_1(\tau)^{-1}$ which satisfies $v'(0) = v_1(\tau)\mu_2 v_1(\tau)^{-1} =: \lambda(\tau)$. Then one differentiates λ at $\tau = 0$. The details will be omitted. □

Theorem 3.29. *Let $(G, \circ)$ be a Lie group with Lie algebra $\mathfrak{g}$. Then there exists a unique C^∞-map*

$$\exp : \mathfrak{g} \to G, \tag{3.55}$$

such that for all $\mu \in \mathfrak{g}$ and $\tau \in \mathbb{R}$

$$\frac{d}{d\tau}\exp(\tau\mu) = \exp(\tau\mu)\mu. \tag{3.56}$$

For all $\mu \in \mathfrak{g}, \tau_1, \tau_2 \in \mathbb{R}$ the map satisfies

$$\exp((\tau_1 + \tau_2)\mu) = \exp(\tau\mu_1) \circ \exp(\tau_2\mu). \tag{3.57}$$

Proof. With the vector field $g \to g\mu$ from (3.53), we apply Theorem 3.26 to the initial value problem

$$p'(\tau) = p(\tau)\mu, \quad p(0) = \mathbb{1}$$

and obtain a locally unique solution $p(\tau) =: \exp(\tau\mu), |\tau| < \varepsilon$. The functional equation (3.57) is proved for $|\tau_1|, |\tau_2| < \frac{\varepsilon}{2}$ in the standard way: For example, let $0 < \tau_1, \tau_2$ and define the continuous function

$$p(\tau) = \begin{cases} \exp(\tau\mu), & 0 \le \tau \le \tau_1, \\ \exp(\tau_1\mu) \circ \exp((\tau - \tau_1)\mu), & \tau_1 < \tau \le \tau_1 + \tau_2. \end{cases} \tag{3.58}$$

Differentiating the relation $L_\gamma \circ L_g = L_{\gamma \circ g}$, at $\mathbb{1}$ yields

$$dL_g(\gamma) dL_\gamma(\mathbb{1}) = dL_{g\circ\gamma}(\mathbb{1}) : T_{\mathbb{1}}G \to T_{g\circ\gamma}G$$

Using this when differentiating (3.58) leads to

$$\begin{aligned} p'(\tau) &= \begin{cases} \exp(\tau\mu)\mu, & 0 \le \tau \le \tau_1, \\ dL_{\exp(\tau_1\mu)}(\exp((\tau - \tau_1)\mu)) dL_{\exp(\tau-\tau_1)\mu))}(\mathbb{1})\mu, & \tau_1 < \tau \le \tau_1 + \tau_2, \end{cases} \\ &= \begin{cases} p(\tau)\mu, & 0 \le \tau \le \tau_1, \\ dL_{p(\tau)}(\mathbb{1})\mu & \tau_1 < \tau \le \tau_1 + \tau_2, \end{cases} = p(\tau)\mu. \end{aligned}$$

Since the right-hand side is continuous at $\tau = \tau_1$, so is $p'(\cdot)$. By the uniqueness of solutions to (3.56) we obtain $p(\tau) = \exp(\tau)\mu$ for all $0 \le \tau \le \tau_1 + \tau_2$ and hence

$$\exp((\tau_1 + \tau_2)\mu) = p(\tau_1 + \tau_2) = \exp(\tau_1\mu) \circ \exp(\tau_2\mu).$$

By symmetry we can exchange τ_1 and τ_2 in this relation. So far we have defined $\exp(\tau\mu)$ for τ in some neighborhood $(-\varepsilon, \varepsilon)$ and proved (3.57) for $\tau_1, \tau_2, \tau_1 + \tau_2 \in (-\varepsilon, \varepsilon)$.

For arbitrary $\tau \in \mathbb{R}$ we select $n \in \mathbb{N}$ such that $|\tau| < \varepsilon n$ and define

$$\exp(\tau\mu) := \Big(\exp(\frac{\tau}{n}\mu)\Big)^n.$$

This definition is in fact independent of the choice of n. If $|\tau|q < \varepsilon$ for another $q \in \mathbb{N}$, then we conclude from the local validity of (3.57)

$$\Big(\exp(\frac{\tau}{q}\mu)\Big)^q = \Big(\big(\exp(\frac{\tau}{nq}\mu)\big)^n\Big)^q = \Big(\big(\exp(\frac{\tau}{nq}\mu)\big)^q\Big)^n = \Big(\exp(\frac{\tau}{n}\mu)\Big)^n.$$

In a similar way, for arbitrary $\tau_1, \tau_2 \in \mathbb{R}$ select $n \in \mathbb{N}$ with $2|\tau_1| < \varepsilon n$, $2|\tau_2| < \varepsilon n$ and find

$$\begin{aligned} \exp((\tau_1 + \tau_2)\mu) &= \Big(\exp(\frac{\tau_1 + \tau_2}{n})\Big)^n = \Big(\exp(\frac{\tau_1}{n}\mu)\exp(\frac{\tau_2}{n}\mu)\Big)^n \\ &= \Big(\exp(\frac{\tau_1}{n}\mu)\Big)^n \Big(\exp(\frac{\tau_2}{n}\mu)\Big)^n = \exp(\tau_1\mu)\exp(\tau_2\mu), \end{aligned}$$

since $\exp(\frac{\tau_1}{n}\mu)$ and $\exp(\frac{\tau_2}{n}\mu)$ commute. Finally, we differentiate the relation $\exp((\tau + \sigma)\mu) = L_{\exp(\tau\mu)} \circ \exp(\sigma\mu)$ at $\sigma = 0$ and obtain from the chain rule

$$\frac{d}{d\tau}\exp(\tau\mu) = \frac{d}{d\sigma}\exp((\tau + \sigma)\mu)_{|\sigma=0} = dL_{\exp(\tau\mu)}(\exp(\sigma\mu))\frac{d}{d\sigma}\exp(\sigma\mu)_{|\sigma=0}$$

$$=dL_{\exp(\tau\mu)}(\mathbb{1})\exp(0\mu)\mu = \exp(\tau\mu)\mu.$$

Therefore, the differential equation is satisfied on the whole real line. $\square$

Example 3.30 (The Euclidean group revisited)**.** Consider the representation (3.22) of the Euclidean group $\mathrm{SO}(\mathbb{R}^d)$ as a subgroup of $\mathrm{GL}(\mathbb{R}^{d+1})$. The corresponding Lie algebra is given by

$$\mathrm{se}(\mathbb{R}^d) = \left\{ \begin{pmatrix} S & a \\ 0 & 0 \end{pmatrix} : S \in \mathrm{so}(\mathbb{R}^d), a \in \mathbb{R}^d \right\}. \tag{3.59}$$

This is quite obvious in view of Corollary 3.16. Given a C^1-path $p(t) = \begin{pmatrix} Q(t) & b(t) \\ 0 & 1 \end{pmatrix}$ in $\mathrm{SO}(\mathbb{R}^d)$ with $Q(0) = I_d, b(0) = 0$, we obtain

$$p'(0) = \begin{pmatrix} Q'(0) & b'(0) \\ 0 & 0 \end{pmatrix}, \quad Q'(0)^T = -Q'(0),$$

i.e. $p'(0) \in \mathrm{se}(\mathbb{R}^d)$. Conversely, let $\mu = \begin{pmatrix} S & a \\ 0 & 0 \end{pmatrix} \in \mathrm{se}(\mathbb{R}^d)$ and let us compute

$$\begin{pmatrix} Q(t) & b(t) \\ 0 & 1 \end{pmatrix} := \exp(t\mu) = \exp\left(t\begin{pmatrix} S & a \\ 0 & 0 \end{pmatrix}\right).$$

From the initial value problem (3.56) we have $Q(0) = I_d, b(0) = 0$ and

$$\begin{pmatrix} Q'(t) & b'(t) \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} Q(t) & b(t) \\ 0 & 1 \end{pmatrix}\begin{pmatrix} S & a \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} Q(t)S & Q(t)a \\ 0 & 0 \end{pmatrix}.$$

Therefore, $Q(t) = \exp(tS)$, $b'(t) = Q(t)a$ and by integration,

$$b(t) = \int_0^t Q(s)ds\, a = \int_0^t \exp(sS)ds\, a.$$

If S is invertible, we find $b(t) = S^{-1}(\exp(tS) - I_d)a$. This motivates to define the analytic function

$$\exp_1(x) := \sum_{j=1}^{\infty} \frac{x^{j-1}}{j!} = \begin{cases} \frac{e^x-1}{x}, & x \neq 0, \\ 1, & x = 0. \end{cases} \tag{3.60}$$

With this definition we obtain for all $S \in \mathrm{so}(\mathbb{R}^d)$

$$b(t) = \int_0^t \sum_{j=0}^{\infty} \frac{1}{j!}(sS)^j dsa = \sum_{j=0}^{\infty} \frac{1}{j!} \int_0^t s^j ds S^j a$$

$$= \sum_{j=0}^{\infty} \frac{t^{j+1}}{(j+1)!} S^j a = t\sum_{j=1}^{\infty} \frac{1}{j!}(tS)^{j-1} a = t\exp_1(tS)a.$$

Let us summarize the final formula (which in fact holds for all $S \in \mathbb{R}^{d,d}$),

$$\exp\left(t\begin{pmatrix} S & a \\ 0 & 0 \end{pmatrix}\right) = \begin{pmatrix} \exp(tS) & t\exp_1(tS)a \\ 0 & 1 \end{pmatrix}, \quad t \in \mathbb{R}. \tag{3.61}$$

3.6. Characterization of relative equilibria. Let us return to the general evolution equation (3.7) from Section 3.1, i.e.

$$u_t = F(u), \tag{3.62}$$

under the following assumptions

(A1) $(X, \|\cdot\|_X)$ is a Banach space, $(Y, \|\cdot\|_Y)$ is a dense subspace of X and $F: Y \to X$ is a continuous operator.

(A2) G is a Lie group with Lie algebra $\mathfrak{g}$ and there exists a homomorphism

$$a: \begin{array}{ccc} G & \to & GL(X), \\ \gamma & \to & a(\gamma), \end{array} \tag{3.63}$$

such that for all $\gamma \in G$, $u \in Y$

$$\begin{aligned} a(\gamma)(Y) &= Y, \\ F(a(\gamma)u) &= a(\gamma)F(u). \end{aligned} \tag{3.64}$$

(A3) For fixed $v \in Y$, resp. $v \in X$, the mapping

$$a(\cdot)v : G \to Y \text{ (resp. X)}, \quad \gamma \to a(\gamma)v$$

is continuous.

(A4) For $v \in Y$ the map $a(\cdot)v : G \to X$ is continuously differentiable with differential

$$d_\gamma[a(\gamma)v] : T_\gamma G \to X.$$

We note that this corresponds to equivariance of (3.62) as defined in (3.8) for the representation of G in $\mathrm{GL}(X)$:

$$\Gamma_X = \{a(g) : g \in G\}.$$

As in (3.9) we define:

Definition 3.31. *A relative equilibrium of* (3.62) *is a pair* $(v_\star, \gamma_\star) \in Y \times C^1(\mathbb{R}, G)$ *such that* $\gamma_\star(0) = \mathbb{1}$ *and*

$$u_\star(t) = a(\gamma_\star(t))v_\star(t), \quad t \in \mathbb{R} \tag{3.65}$$

is a solution of (3.62) *on* $\mathbb{R}$*, i.e.* $u_\star \in C(\mathbb{R}, Y) \cap C^1(\mathbb{R}, X)$ *and* (3.62) *holds on* $\mathbb{R}$*.*

In the literature, the whole group orbit

$$\mathcal{O}(v_\star) = \{a(\gamma)v_\star : \gamma \in G\}$$

is sometimes called a relative equilibrium. However, we prefer to keep the orbit $\gamma_\star$ on the group (normalized by $\gamma_\star(0) = \mathbb{1}$) as part of the definition since it satisfies a differential equation and is to be determined in numerical computations.

In the following it will be important to consider those group actions which leave $v_\star$ invariant, i.e.

$$H(v_\star) = \{\gamma \in G : a(\gamma)v_\star = v_\star\}. \tag{3.66}$$

Obviously, $H(v_\star)$ is a subgroup of G.

Definition 3.32. *For any element $v_\star \in X$ the subgroup (3.66) of G is called the isotropy subgroup (or the stabilizer) of $v_\star$ with respect to the group action a.*

Note that with $(\gamma_\star, v_\star)$ every pair $(\gamma_\star, a(\gamma)v_\star), \gamma \in H(v_\star)$ is also a relative equilibrium. There are two extreme cases. If $H(v_\star) = G$ then $u_\star(t) = v_\star$ for all $t \in \mathbb{R}$. Hence $v_\star$ satisfies $0 = u_{\star,t} = F(v_\star)$, i.e. $v_\star$ is an equilibrium of F. This case will usually be excluded in the following, cf. the 'constant travelling waves' in Example 2.2. The other case is $H(v_*) = \{\mathbb{1}\}$, then we expect the profile $v_\star$ to be unique.
Conditions (A3) and (A4) require some smoothness of the group action. From (A3) we conclude that the isotropy subgroup $H(v_\star)$ is closed, hence is a Lie subgroup of G, see the discussion following Example 3.18. It is important that we assume $a(\cdot)v$ to be differentiable only if $v \in Y$. This will be illustrated by the travelling waves example below.

Lemma 3.33. *Let (A1)-(A4) be satisfied and let $v \in Y$. Then the following holds.*

(i) *Let $g \in C^1(J, G)$, $J \subseteq \mathbb{R}$ an open interval, then for all $t \in J$,*

$$d_\gamma[a(g(t))v]dL_{g(t)}(\mathbb{1}) = a(g(t))d_\gamma[a(\mathbb{1})v]. \tag{3.67}$$

(ii) *The isotropy group $H(v)$ is a Lie subgroup of G with Lie algebra*

$$\mathfrak{h} = N(d_\gamma[a(\mathbb{1})v]) = \{\mu \in \mathfrak{g} : d_\gamma[a(\mathbb{1})v]\mu = 0\}. \tag{3.68}$$

Remark 3.34. *Imagine G as a sphere in $\mathbb{R}^3$, $H(v)$ as one of its great circles, passing through a pole, $\mathfrak{g}$ as the tangent plane at the pole and within it $\mathfrak{h}$ as the tangent line to the great circle.*

Proof. **(i):** Differentiate the relation

$$a(g(t)\gamma)v = a(g(t))(a(\gamma)v), \quad t \in J, \gamma \in G$$

with respect to γ at $\mathbb{1}$ and obtain

$$\begin{aligned} d_\gamma[a(g(t)\gamma)v]dL_{g(t)}(\gamma) &= a(g(t))d_\gamma[a(\gamma)v], \\ d_\gamma[a(g(t))v]dL_{g(t)}(\mathbb{1}) &= a(g(t))d_\gamma[a(\mathbb{1})v]. \end{aligned}$$

(ii): Let us abbreviate $A = d_\gamma[a(\mathbb{1})v] \in L[\mathfrak{g}, X]$. As noted above, $H(v)$ is closed and a Lie subgroup of G. By definition its Lie algebra is

$$\mathfrak{h} = T_\mathbb{1}H(v) = \{p'(0) : p \in C^1((-\varepsilon, \varepsilon), H(v)), p(0) = \mathbb{1}\}.$$

Let $\mu = p'(0) \in \mathfrak{g}$ as in this definition. Then differentiate $a(p(\tau))v = v, |\tau| < \varepsilon$ at $\tau = 0$:

$$0 = \frac{d}{d\tau}\left(a(p(\tau))v\right)_{\tau=0} = \left(d_\gamma[a(p(\tau))v]p'(\tau)\right)_{|\tau=0} = A\mu,$$

hence $\mathfrak{h} \subseteq N(A)$. Conversely, let $\mu \in N(A)$ and define $p(\tau) = \exp(\tau\mu)$, $u(\tau) = a(p(\tau))v$. Then $u(0) = v, p'(0) = \mu$ and by Theorem 3.29 and (3.67)

$$u'(\tau) = d_\gamma[a(p(\tau))v]dL_{\exp(\tau\mu)}(\mathbb{1})\mu = a(p(\tau))d_\gamma[a(\mathbb{1})v]\mu = 0.$$

Therefore $u(\tau) = v$ holds for small $|\tau|$. We conclude $p(\tau) \in H(v)$ for small τ and $p'(0) = \mu \in \mathfrak{h}$ by definition. ∎

Theorem 3.35. *Let assumptions (A1)-(A4) hold and in addition assume*

(A5) *For every $u_0 \in Y$ there exists at most one solution $u \in C^1([0,\infty),X) \cap C([0,\infty),Y)$ of the initial value problem*

$$u_t = F(u), t \geq 0, \quad u(0) = u_0. \tag{3.69}$$

Let $(v_\star,\gamma_\star)$ be a relative equilibrium of (3.62). Then there exists $\mu_\star \in \mathfrak{g}$ such that

$$0 = F(v_\star) - d_\gamma[a(\mathbb{1})v_\star]\mu_\star, \tag{3.70}$$

$$a(\gamma_\star(t))v_\star = a(\exp(t\mu_\star))v_\star, \quad t \geq 0. \tag{3.71}$$

If $H(v_\star) = \{\mathbb{1}\}$ then $\mu_\star$ is unique and $\gamma_\star(t) = \exp(t\mu_\star)$. Conversely, let $\mu_\star \in \mathfrak{g}, v_\star \in Y$ satisfy (3.70). Then $v_\star$ and $\gamma_\star(t) = \exp(t\mu_\star)$ are a relative equilibrium of (3.62).

Remark 3.36. *The theorem shows that a relative equilibrium can always be written with a group orbit of the form $\gamma_\star(t) = \exp(t\mu_\star)$ for some $\mu_\star \in \mathfrak{g}$ and that the pair $(v_\star,\mu_\star)$ satisfies a modified stationary equation (3.70). Therefore, the theorem reduces the search for relative equilibria to solving equation (3.70) for $v_\star \in Y$, $\mu_\star \in \mathfrak{g}$. We will show that the abstract form (3.70) comprises all our defining equations for travelling, oscillating or rotating waves in Section 2.*

Proof. Let $(v_\star,\gamma_\star)$ be a relative equilibrium and define $u_\star(t) = a(\gamma_\star(t))v_\star$. Then the chain rule shows

$$\begin{aligned} a(\gamma_\star(t))F(v_\star) &= F(a(\gamma_\star(t))v_\star) = F(u_\star(t)) = u_{\star,t}(t) \\ &= \frac{d}{dt}\left(a(\gamma_\star(t))v_\star\right) = d_\gamma[a(\gamma_\star(t))v_\star]\gamma_\star'(t), \end{aligned} \tag{3.72}$$

where $\gamma_\star'(t) \in T_{\gamma_\star(t)}G$. Since $dL_{\gamma_\star(t)}(\mathbb{1}) : T_{\mathbb{1}}G \to T_{\gamma_\star(t)}G$ is bijective, we can define $\mu_0(t) \in \mathfrak{g}$ via $\gamma_\star'(t) = dL_{\gamma_\star(t)}(\mathbb{1})\mu_0(t)$. From (3.67) and (3.72) we find

$$\begin{aligned} F(v_\star) &= a(\gamma_\star(t))^{-1}d_\gamma[a(\gamma_\star(t))v_\star]dL_{\gamma_\star}(\mathbb{1})\mu_0(t) \\ &= d_\gamma[a(\mathbb{1})v_\star]\mu_0(t). \end{aligned}$$

With the Lie algebra $\mathfrak{h}$ from (3.68) we decompose

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}_\star, \quad \mu_0(t) = \mu_1(t) + \mu_\star(t).$$

Since $\mathfrak{h}$ is the null space of $d_\gamma[a(\mathbb{1})v_\star]$ by Lemma 3.33, we obtain $F(v_\star) = d_\gamma[a(\mathbb{1})v_\star]\mu_\star(t), t \in \mathbb{R}$. But $d_\gamma[a(\mathbb{1})v_\star] : \mathfrak{h}_\star \to X$ is one to one, so we know that $\mu_\star(t)$ is independent of t. Therefore, $\mu_\star \equiv \mu_\star(t)$ solves equation (3.70). In particular, in case $H(v_\star) = \{\mathbb{1}\}$ we have that $\mu_\star$ is unique. We postpone the proof of (3.71).
Conversely, let us assume that $v_\star \in Y$, $\mu_\star \in \mathfrak{g}$ satisfy (3.70). Then we prove that $w_\star(t) = a(g_\star(t))v_\star$ with $g_\star(t) = \exp(t\mu_\star)$ is a solution of (3.69) with $w(0) = v_\star$. Hence by condition (A5) we obtain that $u_\star(t)$ and $w_\star(t)$ agree on their common domain of definition. In this way the postponed equation (3.71) is proved. Moreover, in case $H(v_\star) = \{\mathbb{1}\}$ the relation $a(\gamma_\star(t))v_\star = a(g_\star(t))v_\star$ implies $\gamma_\star(t)^{-1}g_\star(t) = \mathbb{1}$

and hence $\gamma_\star(t) = g_\star(t) = \exp(t\mu_\star)$. Finally, showing that $w_\star$ solves the evolution equation (3.64) is quite similar to the computation in (3.72):

$$a(g_\star(t))F(v_\star) = F(a(g_\star(t))v_\star) = F(w_\star(t)),$$

and by (3.67), (3.70),

$$\begin{aligned} w_{\star,t}(t) =& \frac{d}{dt}\left(a(g_\star(t))v_\star\right) = d_\gamma[a(g_\star(t))v_\star]g_\star'(t) \\ =& d_\gamma[a(g_\star(t))v_\star]dL_{g_\star(t)}(\mathbb{1})\mu_\star = a(g_\star(t))d_\gamma[a(\mathbb{1})v_\star]\mu_\star \\ =& a(g_\star(t))F(v_\star) = F(w_\star(t)). \end{aligned}$$

With $w_\star(0) = a(g_\star(0))v_\star = a(\mathbb{1})v_\star = v_\star$, the proof is complete. □

3.7. Application to parabolic systems. In this section we develop the correct functional setting showing that the travelling and rotating waves from Section 2 fit into the abstract framework of Section 3.6.

Example 3.37. Let $Y = H^1(\mathbb{R}, \mathbb{R}^m)$, $X = L^2(\mathbb{R}, \mathbb{R}^m)$ and consider the shift action

$$[a(\gamma)u](x) = u(x-\gamma), \quad x \in \mathbb{R}, \quad u \in Y, \gamma \in G = \mathbb{R}. \tag{3.73}$$

First, $a(\gamma)$ is an isometry on X which is continuous with respect to γ since (cf. [3, Satz 2.14])

$$\|a(\gamma)v - a(\gamma_0)v\|_{L^2} = \|v(\cdot + \gamma - \gamma_0) - v\|_{L^2} \to 0 \text{ as } \gamma \to \gamma_0.$$

Then we claim the following inequalities for all $h \in \mathbb{R}$ and $v \in H^1(\mathbb{R}, \mathbb{R}^m)$

$$\begin{aligned} \|v(\cdot + h) - v(\cdot)\|_{L^2} &\le |h|\|v_x\|_{L^2}, \\ \|v(\cdot + h) - v(\cdot) + v_x(\cdot)h\|_{L^2} &\le |h| \sup_{|\delta|\le|h|} \|v_x(\cdot + \delta) - v_x(\cdot)\|_{L^2}. \end{aligned} \tag{3.74}$$

It is sufficient to prove (3.74) for $v \in C_0^\infty(\mathbb{R}, \mathbb{R}^m)$ and then use the fact that $C_0^\infty(\mathbb{R}, \mathbb{R}^m)$ is dense in both X and Y. The first inequality in (3.74) follows from the mean value theorem using the Cauchy Schwarz inequality and Fubini's theorem

$$\|v(\cdot + h) - v(\cdot)\|_{L^2}^2 = \int_{\mathbb{R}} h^2 \Big|\int_0^1 v_x(x+\tau h)d\tau\Big|^2 dx \le$$

$$h^2 \int_{\mathbb{R}} \int_0^1 |v_x(x+\tau h)|^2 d\tau dx = h^2 \int_0^1 \int_{\mathbb{R}} |v_x(x+\tau h)|^2 dx d\tau = \quad h^2\|v_x\|_{L^2}^2.$$

In a similar way,

$$\begin{aligned} \|v(\cdot + h) - v(\cdot) - v_x(\cdot)h\|_{L^2}^2 \le& h^2 \int_0^1 \|v_x(\cdot + \tau h) - v_x(\cdot)\|_{L^2}^2 d\tau \\ \le& h^2 \sup_{|\delta|\le|h|} \|v_x(\cdot + \delta) - v_x(\cdot)\|_{L^2}^2. \end{aligned}$$

Now let $h \to 0$ in (3.74) and use invariance and continuity of $\|\cdot\|_{L^2}$ with respect to the shift (3.73) to obtain the derivative of the action

$$d_\gamma[a(\gamma)v]\mu = -v_x(\cdot - \gamma)\mu, \quad \mu, \gamma \in \mathbb{R}, v \in Y. \tag{3.75}$$

The derivative is also continuous since $v_x \in L^2(\mathbb{R}, \mathbb{R}^m)$.
In the language of semigroup theory (cf. Section 8.2), we have shown that H^1 is in the domain of the infinitesimal generator $d_\gamma[a(\mathbb{1})\cdot]$ of the one-parameter group $a(\gamma), \gamma \in \mathbb{R}$ (which here is a group). In fact both spaces agree. For this one has to show that $h^{-1}(v(\cdot + h) - v(\cdot)) \to w$ in L^2 as $h \to 0$ implies $v \in H^1$ and $v_x = w$ (Exercise).

The example shows that the shift action satisfies our assumptions (A3) and (A4). If we want to apply Theorem 3.35 to the nonlinear parabolic system (3.1) we need more conditions on the nonlinearity.

Proposition 3.38. Let $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$ satisfy $f(0) = 0$. Then conditions (A1)-(A4) hold for the action (3.73) and the operator

$$(F(u))(x) = Au_{xx}(x) + f(u(x)), \quad x \in \mathbb{R}^m, \quad u \in Y \tag{3.76}$$

with the setting $Y = H^2(\mathbb{R}, \mathbb{R}^m)$ and $X = L^2(\mathbb{R}, \mathbb{R}^m)$.

Proof. Introduce the Nemitzky operator $\mathcal{F}$ associated with f,

$$(\mathcal{F}(u))(x) = f(u(x)), \quad x \in \mathbb{R}, u \in (\mathbb{R}^m)^{\mathbb{R}}.$$

By the Sobolev imbedding theorem (see Appendix 8.3)

$$\|u\|_{L^\infty} \leq \|u\|_{H^1}, \quad u \in H^1(\mathbb{R}, \mathbb{R}^m). \tag{3.77}$$

Hence for a.e. $x \in \mathbb{R}$ by the mean value theorem

$$|f(u(x))| = |f(u(x)) - f(0)| \leq \sup\{|Df(v)| : |v| \leq \|u\|_{H^1}\}|u(x)|.$$

Taking squares and integrating shows $\mathcal{F}(u) \in L^2(\mathbb{R}, \mathbb{R}^m)$. In a similar way, for any two functions $u_1, u_2 \in H^1(\mathbb{R}, \mathbb{R}^m)$,

$$\|\mathcal{F}(u_1) - \mathcal{F}(u_2)\|_{L^2} \leq \mathcal{K}(u_1, u_2)\|u_1 - u_2\|_{L^2}, \tag{3.78}$$

where $\mathcal{K}(u_1, u_2) = \sup\{|Df(v)| : |v| \leq \max(\|u_1\|_{H^1}, \|u_2\|_{H^1})\}$, i.e. the operator $\mathcal{F}$ is locally Lipschitz bounded. Finally, taking derivatives one finds

$$\begin{aligned} &\|\mathcal{F}(u_1)_x - \mathcal{F}(u_2)_x\|_{L^2} = \|Df(u_1(\cdot))u_{1,x} - Df(u_2(\cdot))u_{2,x}\|_{L^2} \\ \leq&\|(Df(u_1(\cdot)) - Df(u_2(\cdot)))u_{1,x}\|_{L^2} + \|Df(u_2(\cdot))(u_{1,x} - u_{2,x})\|_{L^2} \\ \leq&\|(Df(u_1(\cdot)) - Df(u_2(\cdot))\|_{L^\infty}\|u_{1,x}\|_{L^2} + \mathcal{K}(0, u_2)\|u_{1.x} - u_{2.x}\|_{L^2} \\ \leq&\omega(\max(\|u_1\|_{H^1}, \|u_2\|_{H^1}), \|u_1 - u_2\|_{H^1})\|u_1\|_{H^1} + \mathcal{K}(0, u_2)\|u_1 - u_2\|_{H^1} \end{aligned} \tag{3.79}$$

with the modulus of continuity

$$\omega(R, \delta) = \sup\{|Df(v_1) - Df(v_2)| : |v_1 - v_2| \leq \delta, \quad |v_1|, |v_2| \leq R\}.$$

Since $\omega(R, \delta) \to 0$ as $\delta \to 0$, the operator $\mathcal{F}$ is even continuous when considered as a map from H^1 into H^1. With the choice $Y = H^2$, also the linear differential operator $u \mapsto Au_{xx}$ maps Y into $X = L^2$ and our assertion is proved. □

In the next step we consider the multidimensional reaction diffusion system (2.93) subject to the action of the special Euclidean group $\mathrm{SE}(\mathbb{R}^d)$, see (3.22).

Example 3.39. Consider $X = L^2(\mathbb{R}^d, \mathbb{R}^m)$ and the action

$$[a(\gamma)v](x) = v(Q^T(x-b)), \quad x \in \mathbb{R}^d, \quad \gamma = \begin{pmatrix} Q & b \\ 0 & 1 \end{pmatrix} \in \mathrm{SE}(\mathbb{R}^d). \tag{3.80}$$

We claim that this action is continuous. By the orthogonality of Q and the transformation formula, the operator $a(\gamma)$ is an isometry, i.e. $\|a(\gamma)v\|_{L^2} = \|v\|_{L^2}$. Hence it is sufficient to prove continuity at $\gamma = \mathbb{1} = I_3$. Given $\varepsilon > 0$ we choose $\varphi \in C_0^\infty(\mathbb{R}, \mathbb{R}^m)$ such that $\|v - \varphi\|_{L^2} \le \frac{\varepsilon}{3}$ and obtain

$$\begin{aligned} \|a(\gamma)v - v\|_{L^2} \le& \|a(\gamma)(v-\varphi)\|_{L^2} + \|a(\gamma)\varphi - \varphi\|_{L^2} + \|\varphi - v\|_{L^2} \\ \le& \frac{2}{3}\varepsilon + \|a(\gamma)\varphi - \varphi\|_{L^2}. \end{aligned}$$

Now take $R = 1 + \sup\{|x|_2 : x \in \mathrm{supp}(\varphi)\}$ and $|b|_2 \le 1$ and observe $Q^T(x-b), x \notin \mathrm{supp}(\varphi)$ for $|x|_2 > R$ as well as

$$|Q^T(x-b) - x|_2 \le |Q^T - \mathbb{1}|_2 R + |b|_2 \quad \text{for} \quad |x| \le R.$$

Using the uniform continuity of φ this leads to

$$\begin{aligned} \|a(\gamma)\varphi - \varphi\|_{L^2}^2 =& \int_{\mathbb{R}^d} |\varphi(Q^T(x-b)) - \varphi(x)|_2^2 dx \\ \le& C(R)\omega(R, |Q^T - \mathbb{1}|_2 R + |b|_2)) \to 0, \end{aligned}$$

where the constant $C(R)$ depends on R only and

$$\omega(R, \delta) = \sup\{|\varphi(x) - \varphi(y)|_2^2 : |x|_2, |y|_2 \le R, |x-y|_2 \le \delta\}.$$

Calculating the derivative of the action (3.80) in a formal sense is rather easy (cf. (3.6)). We consider the path (see (3.6), (3.61))

$$\gamma(t) = \exp\left(t\begin{pmatrix} S & a \\ 0 & 0 \end{pmatrix}\right) = \begin{pmatrix} \exp(tS) & b(t) \\ 0 & 1 \end{pmatrix}, \; b(t) = \int_0^t \exp(\tau S) d\tau a, \tag{3.81}$$

passing through $\mathbb{1}$, and evaluate the derivative of $a(\gamma(t))v$ at $t = 0$:

$$\left(d_\gamma(a(\mathbb{1})v)\mu\right)(x) = -v_x(x)(Sx + a), \; x \in \mathbb{R}^d, \; \mu = \begin{pmatrix} S & a \\ 0 & 0 \end{pmatrix} \in \mathrm{se}(\mathbb{R}^d). \tag{3.82}$$

Making this rigorous in suitable function spaces is more involved. Formula (3.82) suggests to introduce for $S \in \mathrm{so}(\mathbb{R}^d)$ the linear differential operator

$$\mathcal{L}_S v(x) = v_x(x) Sx, \quad x \in \mathbb{R}^d, v \in H^1(\mathbb{R}^d, \mathbb{R}^m). \tag{3.83}$$

With this we define the Euclidian function space

$$H^1_{\mathrm{Eucl}}(\mathbb{R}^d, \mathbb{R}^m) = \{v \in H^1(\mathbb{R}^d, \mathbb{R}^m) : \mathcal{L}_S v \in L^2(\mathbb{R}^d, \mathbb{R}^m) \; \forall S \in \mathrm{so}(\mathbb{R}^d)\}, \tag{3.84}$$

which becomes a Banach space with respect to the norm

$$\|v\|^2_{H^1_{\mathrm{Eucl}}} = \|v\|^2_{H^1} + \sup\{\|\mathcal{L}_S v\|^2_{L^2} : S \in \mathrm{so}(\mathbb{R}^d), |S|_2 \le 1\}. \tag{3.85}$$

Of course, instead of taking the supremum over all $S \in \mathrm{so}(\mathbb{R}^d)$ with $|S|_2 \le 1$, one can simply take the maximum over a basis of the $\frac{d(d-1)}{2}$-dimensional space $\mathrm{so}(\mathbb{R}^d)$.

Proposition 3.40. The space $C_0^\infty(\mathbb{R}^d, \mathbb{R}^m)$ is dense in $H^1_{\mathrm{Eucl}}(\mathbb{R}^d, \mathbb{R}^m)$ with respect to the norm $\|\cdot\|_{H^1_{\mathrm{Eucl}}}$. Further, for $v \in H^1_{\mathrm{Eucl}}(\mathbb{R}^d, \mathbb{R}^m)$ the Euclidean action (3.80) is a continuously differentiable map from $\mathrm{SO}(\mathbb{R}^d)$ into $L^2(\mathbb{R}^d, \mathbb{R}^m)$ with the derivative given by (3.82).

Proof. In the following we will frequently omit the underlying $\mathbb{R}^n$-spaces and write $H^1_{\mathrm{Eucl}} = H^1_{\mathrm{Eucl}}(\mathbb{R}^d, \mathbb{R}^m)$, $L^2 = L^2(\mathbb{R}^d, \mathbb{R}^m)$ etc. Balls of radius $R > 0$ are denoted by $B_R = \{x \in \mathbb{R}^d : |x|_2 < R\}$. The proof of the first part adapts standard techniques to prove density of C_0^∞ in Sobolev spaces to our setting (cf. [3, 2.14,2.22,2.23,Ü8.8]). Let

$$\varphi_\delta(x) = \delta^{-d}\varphi_1(\delta^{-1}x), \quad \varphi_1 \in C_0^\infty(B_1, \mathbb{R}) \tag{3.86}$$

be a Dirac sequence satisfying

$$\varphi_1 \ge 0, \quad \int_{\mathbb{R}^d} \varphi_1(x)dx = 1, \quad \varphi_1(x) = \varphi_1(y) \text{ if } |x|_2 = |y|_2. \tag{3.87}$$

As usual we approximate $v \in H^1_{\mathrm{Eucl}}$ by the sequence

$$v_\delta(x) = \int_{\mathbb{R}^d} \varphi_\delta(x-y)v(y)dy, \quad x \in \mathbb{R}^d. \tag{3.88}$$

Then $v_\delta \in C^\infty(\mathbb{R}^d, \mathbb{R}^m)$ holds and we show that

$$\|v - v_\delta\|_{H^1_{\mathrm{Eucl}}} \to 0 \quad \text{as} \quad \delta \to 0. \tag{3.89}$$

For the norm $\|v - v_\delta\|_{H^1}$ this follows by standard arguments, see [3, 2.22]. Hence it is sufficient to prove

$$\sup\{\|\mathcal{L}_S(v - v_\delta)\|_{L^2} : |S|_2 \le 1, S \in \mathrm{so}(\mathbb{R}^d)\} \to 0 \quad \text{as} \quad \delta \to 0. \tag{3.90}$$

From (3.87) we obtain

$$\begin{aligned}
\mathcal{L}_S v_\delta(x) =& \frac{d}{dt} v_\delta(\exp(tS)x)_{|t=0} \\
=& \int_{\mathbb{R}^d} \frac{d}{dt}(\varphi_\delta(\exp(tS)(x - \exp(-tS)y)))_{|t=0} v(y)dy \\
=& \int_{\mathbb{R}^d} \frac{d}{dt}(\varphi_\delta(x - \exp(-tS)y))_{|t=0} v(y)dy \\
=& \int_{\mathbb{R}^d} \sum_{j=1}^d D_j\varphi_\delta(x-y)(Sy)_j v(y)dy \\
=& -\int_{\mathbb{R}^d} \sum_{j=1}^d \sum_{k\ne j} S_{jk}\frac{\partial}{\partial y_j}(y_k\varphi_\delta(x-y))v(y)dy.
\end{aligned}$$

For every fixed $x \in \mathbb{R}^d$ the function $y \to y_k \varphi_\delta(x-y)$ has compact support in $\mathbb{R}^d$, hence by the definition of the weak derivative

$$\begin{aligned}\mathcal{L}_S v_\delta(x) &= \int_{\mathbb{R}^d} \sum_{j=1}^{d} \sum_{k \neq j} S_{jk} y_k \varphi_\delta(x-y) D_j v(y) dy \\ &= \int_{\mathbb{R}^d} \varphi_\delta(x-y) \mathcal{L}_S v(y) dy.\end{aligned}$$

By the standard convolution estimate (cf. [3, Satz 2.14])

$$\|\mathcal{L}_S(v_\delta - v)\|_{L^2} \leq \int_{\mathbb{R}^d} \varphi_\delta(\xi) d\xi \sup_{|h| \leq \delta} \|\mathcal{L}_S v(\cdot + h) - \mathcal{L}_S v\|_{L^2}.$$

Taking squares and the supremum over $|S|_2 \leq 1$ in the finite dimensional space $\mathrm{so}(\mathbb{R}^d)$ then shows (3.90).
Thus we have shown that $C^\infty \cap H^1_{\mathrm{Eucl}}$ is dense in H^1_{Eucl} and it remains to approximate an element $v \in C^\infty \cap H^1_{\mathrm{Eucl}}$ by functions in C_0^∞. For that purpose take a cut-off function $\chi \in C^\infty(\mathbb{R}, \mathbb{R})$ with

$$\chi(r) \begin{cases} = & 1, & |r| \leq 1, \\ \in & [0,1], & 1 \leq |r| \leq 2, \\ = & 0, & |r| \geq 2 \end{cases}$$

and let

$$\chi_R(x) = \chi(R^{-1}|x|_2), \quad x \in \mathbb{R}^d, R > 0.$$

Then the function $\chi_R v$ has compact support and

$$\mathcal{L}_S(v - \chi_R v) = (1 - \chi_R)\mathcal{L}_S v - (\mathcal{L}_S \chi_R) v.$$

The skew-symmetry of S implies

$$(\mathcal{L}_S \chi_R)(x) = (R|x|_2)^{-1} \chi'(R^{-1}|x|_2) \sum_{j=1}^{d} x_j (Sx)_j = 0.$$

Therefore, we obtain for $R \to \infty$

$$\sup_{|S|_2 \leq 1} \|\mathcal{L}_S(v - \chi_R v)\|_{L^2} \leq \sup_{|S|_2 \leq 1} \|\mathcal{L}_S v\|_{L^2(\mathbb{R}^d \setminus B_R, \mathbb{R}^m)} \to 0.$$

The same convergence holds for the norm $\|\cdot\|_{H^1}$ and this shows our first assertion.
Next we prove a Lipschitz estimate for $v \in H^1_{\mathrm{Eucl}}$ and $\gamma = \begin{pmatrix} Q & b \\ 0 & 1 \end{pmatrix}$,

(3.91) $$\|a(\gamma)v - v\|_{L^2} \leq C \|v\|_{H^1_{\mathrm{Eucl}}} \left(|b|_2 + |Q - \mathbb{1}|_2\right).$$

Since both sides are continuous with respect to $\|\cdot\|_{H^1_{\mathrm{Eucl}}}$ it is sufficient to prove (3.91) for $v \in C_0^\infty$. By the triangle inequality,

$$\|a(\gamma)v - v\|_{L^2} \leq \left(\int_{\mathbb{R}^d} |v(Q^T(x-b)) - v(Q^T x)|_2^2 dx \right)^{1/2}$$

$$+\left(\int_{\mathbb{R}^d}|v(Q^Tx)-v(x)|_2^2dx\right)^{1/2}=T_1+T_2.$$

As in Example 3.37 the first term is bounded by

$$T_1\le|Q^Tb|_2\|v_x\|_{L^2}\le|b|_2\|v\|_{H^1}.$$

For the second term take $S\in\mathrm{so}(\mathbb{R}^d)$ with $Q=\exp(S)$ (cf. Proposition 3.11) and estimate

$$\begin{aligned}T_2^2&=\int_{R^d}\left|\int_0^1\frac{d}{d\tau}v(\exp(-\tau S)x)d\tau\right|_2^2dx\\&=\int_{\mathbb{R}^d}\left|\int_0^1v_x(\exp(-\tau S)x)(-S)\exp(-\tau S)xd\tau\right|_2^2dx\\&\le\int_{\mathbb{R}^d}\int_0^1|v_x(\exp(-\tau S)x)S\exp(-\tau S)x|_2^2\,d\tau dx\\&=\int_0^1\int_{\mathbb{R}^d}|v_x(\exp(-\tau S)x)S\exp(-\tau S)x|_2^2\,dxd\tau\\&=\int_{\mathbb{R}^d}|v_x(y)Sy|_2^2\,dy\le|S|_2^2\|v\|^2_{H^1_{\mathrm{Eucl}}}.\end{aligned}$$

By Proposition 3.10 we have

$$|S|_2\le2|Q-\mathbb{1}|_2\quad\text{if}\quad|Q-\mathbb{1}|_2\le\frac{1}{2}.$$

Since (3.91) is trivial for $|Q-\mathbb{1}|_2\ge\frac12$ the assertion follows.
Finally we prove (3.82). Take first $v\in C_0^\infty$ and consider the path (3.81) in $\mathrm{SE}(\mathbb{R}^d)$. In the following we abbreviate

$$(\cdot\ t\tau x)=\exp(-t\tau S)(x-b(t\tau)).$$

$$\begin{aligned}&\|t^{-1}(a(\gamma(t))v-v)+\mathcal{L}_Sv+v_xad\tau\|_{L^2}\\&=\int_{\mathbb{R}^d}\left|\int_0^1t^{-1}\frac{d}{d\tau}v(\cdot\ t\tau x)+v_x(x)(Sx+a)d\tau\right|_2^2dx\\&=\int_{\mathbb{R}^d}\left|\int_0^1-v_x(\cdot\ t\tau x)\left[S(\cdot\ t\tau x)+a\right]+v_x(x)(Sx+a)d\tau\right|^2dx\\&\le2\int_{\mathbb{R}^d}\int_0^1|v_x(x)Sx-v_x(\cdot\ t\tau x)S(\cdot\ t\tau x)|_2^2+|v_x(x)-v_x(\cdot\ t\tau x)|_2^2\,|a|_2^2d\tau dx.\end{aligned}$$

Since $|(\cdot\ t\tau x)-x|_2\to0$ as $t\to0$ uniformly for $\tau\in[0,1]$ and for x in a compact set, the right-hand side converges to zero. For a general $v\in H^1_{\mathrm{Eucl}}$ and $\varepsilon>0$ we choose $v_\varepsilon\in C_0^\infty$ such that $\|v-v_\varepsilon\|_{H^1_{\mathrm{Eucl}}}\le\varepsilon$. Then we estimate with (3.91)

$$\|\frac{1}{t}(a(\gamma(t))v-v)+\mathcal{L}_Sv+v_xa\|_{L^2}$$

$$
\begin{aligned}
&\leq \| \frac{1}{t} \left(a(\gamma(t))(v - v_\varepsilon)\right) - (v - v_\varepsilon) \|_{L^2} + \| \mathcal{L}_S (v - v_\varepsilon) + (v_x - v_{x,\varepsilon}) a \|_{L^2} \\
&+ \| \frac{1}{t} (a(\gamma(t)) v_\varepsilon - v_\varepsilon) + \mathcal{L}_S v_\varepsilon + v_{\varepsilon,x} a \|_{L^2} \\
&\leq \left(C |t|^{-1} (|b(t)|_2 + |\exp(-tS) - \mathbb{1}|_2) + |a|_2 \right) \varepsilon \\
&+ \| \frac{1}{t} (a(\gamma(t)) v_\varepsilon - v_\varepsilon) + \mathcal{L}_S v_\varepsilon + v_{\varepsilon,x} a \|_{L^2}.
\end{aligned}
$$

The first term is bounded by $\mathcal{O}(\varepsilon)$ uniformly in $t \neq 0$ and the second term converges to 0 as $t \to 0$ for fixed $\varepsilon > 0$. This shows that we can bound the whole right-hand side by $\mathcal{O}(\varepsilon)$ for $|t|$ sufficiently small. This finishes the proof. □

Exercise 3.41. Recall that the infinitesimal generator of a one-parameter semigroup $T(t), t \geq 0$ in some Banach space X is given by (cf. Section 8.2)

$$
\mathcal{D}(A) = \{ v \in X : \lim_{t \to 0} t^{-1} (T(t) - I) v =: Av \quad \text{exists} \}.
$$

For every $\mu \in \mathrm{se}(\mathbb{R}^d)$ let $A(\mu)$ denote the infinitesimal generator of the semigroup $T_\mu(t) v = a(\exp(t\mu)) v, v \in L^2(\mathbb{R}^d, \mathbb{R}^m)$, where $a(\gamma)$ is the action (3.80) of the Euclidean group on L^2. Prove that

$$
H^1_{\mathrm{Eucl}}(\mathbb{R}^d, \mathbb{R}^m) = \bigcap \{ \mathcal{D}(A(\mu)) : \mu \in \mathrm{se}(\mathbb{R}^d) \}. \tag{3.92}
$$

Hint: Note that the inclusion '$\subseteq$' follows from Propositions 3.38, 3.40.

Solution: Let $v \in \cap \{ \mathcal{D}(A(\mu)) : \mu \in \mathrm{se}(\mathbb{R}^d) \}$. For every $j = 1, \ldots, d$ let e^j be the j-th Cartesian unit vector and let $v_j \in L^2$ be the limit of $t^{-1}(v(\cdot + te^j) - v(\cdot))$ in L^2 as $t \to 0$. Then we have for every $\varphi \in C_0^\infty$

$$
\begin{aligned}
\int_{\mathbb{R}^d} v_j(x) \varphi(x) dx &= \lim_{t \to 0} \int_{\mathbb{R}^d} t^{-1} (v(x + te^j) - v(x)) \varphi(x) dx \\
&= - \lim_{t \to 0} \int_{\mathbb{R}^d} v(x) t^{-1} (\varphi(x) - \varphi(x - te^j)) dx = - \int_{\mathbb{R}^d} v(x) D_j \varphi(x) dx.
\end{aligned}
$$

Hence v has a weak derivative $D_j v$ and $D_j v = v_j$. Thus we have $v \in H^1$. Next consider the L^2-limit $w_S = \lim_{t \to 0} t^{-1} (v(\exp(tS) \cdot) - v(\cdot))$ for $S \in \mathrm{so}(\mathbb{R}^d)$. Then we obtain for $\varphi \in C_0^\infty$

$$
\begin{aligned}
\int_{\mathbb{R}^d} w_S(x) \varphi(x) dx &= \lim_{t \to 0} \int_{\mathbb{R}^d} t^{-1} (v(\exp(tS)x) - v(x)) \varphi(x) dx \\
&= \lim_{t \to 0} \int_{\mathbb{R}^d} v(x) t^{-1} (\varphi(\exp(-tS)x) - \varphi(x)) dx \\
&= - \int_{\mathbb{R}^d} v(x) \varphi_x(x) Sx dx.
\end{aligned}
$$

Now we use $v \in H^1$, $S = -S^T$ and the fact that $x \mapsto (Sx)_j \varphi(x)$ is a function in C_0^∞ to conclude

$$
\begin{aligned}
\int_{\mathbb{R}^d} v_x(x) Sx \varphi(x) dx &= \int_{\mathbb{R}^d} \sum_{j=1}^d D_j v(x) \left[(Sx)_j \varphi(x) \right] dx \\
&= - \int_{\mathbb{R}^d} v(x) \sum_{j=1}^d D_j \left[(Sx)_j \varphi(x) \right] dx
\end{aligned}
$$

$$= - \int_{\mathbb{R}^d} v(x) \sum_{j=1}^{d} (Sx)_j D_j \varphi(x) dx = - \int_{\mathbb{R}^d} v(x) \varphi_x(x) Sx dx.$$

Comparing both equalities we obtain

$$\int_{\mathbb{R}^d} w_S(x) \varphi(x) dx = \int_{\mathbb{R}^d} v_x(x) Sx \varphi(x) dx$$

for all $\varphi \in C_0^\infty$, hence $\mathcal{L}_S v = w_S \in L^2$.

The verification of the conditions (A1)-(A4) for the multidimensional operator from Lemma 2.31 is almost a copy of Proposition 3.38.

Proposition 3.42. Let $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$ satisfy $f(0) = 0$. Then conditions (A1)-(A4) hold for the action (3.80) and the operator

$$(F(u))(x) = A \Delta u(x) + f(u(x)), \quad x \in \mathbb{R}^m, \quad u \in Y \tag{3.93}$$

with the setting $Y = H^k_{\mathrm{Eucl}}(\mathbb{R}, \mathbb{R}^m)$ and $X = L^2(\mathbb{R}, \mathbb{R}^m)$ provided $k > \frac{d}{2}$.

Proof. Recall the Nemitzky operator $\mathcal{F}$ associated with f,

$$(\mathcal{F}(u))(x) = f(u(x)), \quad x \in \mathbb{R}^d, u \in (\mathbb{R}^m)^{\mathbb{R}}.$$

Since $k > \frac{d}{2}$, the Sobolev imbedding theorem (see Appendix 8.3) implies

$$\|u\|_{L^\infty} \le \|u\|_{H^k}, \quad u \in H^k(\mathbb{R}^d, \mathbb{R}^m). \tag{3.94}$$

Hence for a.e. $x \in \mathbb{R}^d$ by the mean value theorem

$$|f(u(x))| = |f(u(x)) - f(0)| \le \sup\{|Df(v)| : |v| \le \|u\|_{H^k}\} |u(x)|.$$

Taking squares and integrating shows $\mathcal{F}(u) \in L^2(\mathbb{R}^d, \mathbb{R}^m)$. In a similar way, for any two functions $u_1, u_2 \in H^k(\mathbb{R}^d, \mathbb{R}^m)$,

$$\|\mathcal{F}(u_1) - \mathcal{F}(u_2)\|_{L^2} \le \mathcal{K}(u_1, u_2) \|u_1 - u_2\|_{L^2}, \tag{3.95}$$

where $\mathcal{K}(u_1, u_2) = \sup\{|Df(v)| : |v| \le \max(\|u_1\|_{H^k}, \|u_2\|_{H^k})\}$, i.e. the operator $\mathcal{F}$ is locally Lipschitz bounded. Finally, taking derivatives one finds

$$\begin{aligned} \|\mathcal{F}(u_1)_x - \mathcal{F}(u_2)_x\|_{L^2} &= \|Df(u_1(\cdot)) u_{1,x} - Df(u_2(\cdot)) u_{2,x}\|_{L^2} \\ \le& \|(Df(u_1(\cdot)) - Df(u_2(\cdot))) u_{1,x}\|_{L^2} + \|Df(u_2(\cdot))(u_{1,x} - u_{2,x})\|_{L^2} \\ \le& \|(Df(u_1(\cdot)) - Df(u_2(\cdot))\|_{L^\infty} \|u_{1,x}\|_{L^2} + \mathcal{K}(0, u_2) \|u_{1.x} - u_{2.x}\|_{L^2} \\ \le& \omega(\max(\|u_1\|_{H^k}, \|u_2\|_{H^k}), \|u_1 - u_2\|_{H^k}) \|u_1\|_{H^k} + \mathcal{K}(0, u_2) \|u_1 - u_2\|_{H^k} \end{aligned} \tag{3.96}$$

with the modulus of continuity

$$\omega(R, \delta) = \sup\{|Df(v_1) - Df(v_2)| : |v_1 - v_2| \le \delta, \quad |v_1|, |v_2| \le R\}.$$

Since $\omega(R, \delta) \to 0$ as $\delta \to 0$, the operator $\mathcal{F}$ is even continuous when considered as a map from H^k into H^1. With the choice $Y = H^k_{\mathrm{Eucl}}$, $k \ge 2$, also the linear differential operator $u \mapsto A \Delta u$ maps Y into $X = L^2$ and our assertion is proved. □

4. Existence of travelling waves for reaction diffusion systems

In this chapter we return to travelling waves in one space dimension. For the scalar travelling wave equation (2.29) we prove existence of an orbit connecting two steady states for nonlinearities that behave like the cubic in the Nagumo equation (2.3). This requires to make the phase plane analysis from Section 2.4 rigorous in a specific case. Then we will comment on (but not prove) some recent results for the general gradient case, see Proposition 2.15. The proof of existence in the scalar case needs a rather detailed result on stable and unstable manifolds of hyperbolic points in parameterized dynamical systems. This will be the topic of the final Section 4.3.

4.1. Scalar bistable reaction diffusion equations. Consider the equation

$$u_t = u_{xx} + f(u), \quad x \in \mathbb{R}, t \geq 0, \tag{4.1}$$

where $f \in C^1(\mathbb{R}, \mathbb{R})$ has a cubic shape. More precisely, we assume that there exist $a \in (0,1)$ with the following properties (see Figure)

(B1) $f(0) = f(a) = f(1) = 0$,
(B2) $f < 0$ on $(0, a)$, and $f > 0$ on $(a, 1)$,
(B3) $f'(0) < 0$, and $f'(1) < 0$,
(B4) $\int_0^1 f(x)dx > 0$.

In the scalar case f is always a gradient (cf. (2.44)), namely of

$$F(v) = \int_0^v f(x)dx, \quad v \in \mathbb{R}. \tag{4.2}$$

From (B1)-(B4) we find that F has local maxima at $x = 0, 1$ and a local minimum at $x = a$. Moreover, there is a unique value $a_\star \in (a, 1)$ such that $F < 0$ in $(0, a_\star)$ and $F > 0$ in $(a_\star, 1)$ (see Figure). From Proposition 2.15 and (B4) we also obtain that a travelling wave $u_\star(x,t) = v_\star(x - c_\star t)$ of (4.1) connecting $v_0 = 0$ to $v_1 = 1$, satisfies

$$c_\star \int_{-\infty}^{\infty} v_\star'(x)^2 dx = F(0) - F(1) = -F(1) = -\int_0^1 f(x)dx < 0.$$

Hence we look for travelling waves with $c_\star < 0$.
From Section 2.3 we know that $w_\star = (v_\star, v_\star')^T$ is a heteroclinic orbit with $c = c_\star$ of the TWODE

$$\begin{pmatrix} w_1' \\ w_2' \end{pmatrix} = \begin{pmatrix} w_2 \\ -f(w_1) - cw_2 \end{pmatrix} =: G(w, c). \tag{4.3}$$

Accoding to Lemma 2.16 the endpoints of the heteroclinic orbit

$$w^j = \begin{pmatrix} j \\ 0 \end{pmatrix}, \quad j = 0, 1 \tag{4.4}$$

are saddles of (4.3) with eigenvalues and eigenvectors of $DG(w^j, c)$ given by

$$\lambda_\pm^j = \frac{1}{2}(-c \pm \sqrt{c^2 - 4f'(v_j)}), \quad j = 0, 1, \tag{4.5}$$

$$y_\pm^j = \begin{pmatrix} 1 \\ \lambda_\pm^j \end{pmatrix}, \quad j = 0, 1. \tag{4.6}$$

We consider negative values of c. Then the eigenvector y_+^0 belonging to the unstable eigenvalue at w^0 has positive components and points into the positive quadrant, while the negative eigenvector $-y_-^1$ belonging to the stable eigenvalue at w^1 has negative first and positive second component. Then $w^1 - y_-^1$ also lies in the positive quadrant but to the left of w^1. The following schematic phase diagram illustrates the situation: stable and unstable manifolds of the two saddles w^0, w^1 with their tangent spaces and the vertical line $w_1 = a_0$ where $a < a_0 < 1$ and where the unstable manifold of w^0 and the stable manifold of w^1 are expected to intersect. The following is the main result of this section.

Theorem 4.1. *Let conditions (B1)-(B4) be satisfied. Then there exists*

$$c_\star \in (-\widehat{c}, 0), \quad \widehat{c}^2 = 8\max\left(4a^{-2}(F(1) - F(a)), \max_{0<v\leq 1}\frac{f(v)}{v}\right) \tag{4.7}$$

and a solution $w_\star \in C^1(\mathbb{R}, \mathbb{R}^2)$ of system (4.3) at $c = c_\star$ such that

(i) $\lim_{x\to-\infty} w_\star(x) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \lim_{x\to\infty} w_\star(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$

(ii) $w_\star(x) \in (0, 1) \times (0, \infty)$ *for all* $x \in \mathbb{R}$.

Remark 4.2. *The lower bound $-\widehat{c}$ is not the best possible. However, the explicit form shows that such a bound depends only on data determined by the nonlinearity f. In the proof we will frequently derive estimates that are uniform on compact c-intervals. Therefore it is important to have this a-priori bound. One can also show that the heteroclinic orbit is unique in the strip $(0, 1) \times (0, \infty)$, see ??? But we will not go into the details of the proof of such a result.*

The proof of Theorem 4.1 will be divided into several steps.

Step 1: The local unstable manifold.

First, we use Theorem 4.7 from Section 4.4 to study the local unstable manifold of $w^0 \in \mathbb{R}^2$. Let $P_+^0(c) : \mathbb{R}^2 \to \text{span}\{y_+^0\}$ denote the projector onto the unstable subspace of $DG(w^0, c)$ with respect to the decomposition $\mathbb{R}^2 = \text{span}\,\{y_-^0\} \oplus \text{span}\,\{y_+^0\}$. The dependence on $c \in [-\widehat{c}, 0]$ will frequently be suppressed in the sequel.

According to Theorem 4.7 there exist constants $\rho_0, K_1, K_2, K_3, \eta > 0$ such that for all $c \in [-\widehat{c}, 0]$, $0 < \rho \leq \rho_0$ the boundary value problem

$$w' = G(w, c) \quad \text{on} \quad (-\infty, 0], \tag{4.8}$$

$$P_+^0(c)w(0) = \rho y_+^0, \tag{4.9}$$

$$|w(t)| \leq \rho K_1 \quad \text{for all} \quad t \in (-\infty, 0] \tag{4.10}$$

has a unique solution $w = w(\cdot, \rho, c) \in C^1((-\infty, 0], \mathbb{R}^2)$ which is as smooth as G with respect to all variables. Moreover, the following estimates hold

$$|w(t, \rho, c)|_\infty \leq K_2 e^{\eta t}, \;\; t \leq 0, \tag{4.11}$$

$$|(I - P_+^0(c))w(0, \rho, c)|_\infty \leq K_3 \rho^2, \;\; 0 < \rho \leq \rho_0. \tag{4.12}$$

From (4.12), (4.9) we obtain

$$|w(0) - \rho y_+^0|_\infty = |(I - P_+^0(c))w(0)|_\infty \leq K_3 \rho^2.$$

We take $\rho > 0$ sufficiently small such that the following estimates hold:

$$w_1(0) \geq \rho - |\rho y_+^0 - w(0)|_\infty \geq \rho - K_3 \rho^2 \geq \frac{\rho}{2}, \tag{4.13}$$

$$w_1(0) \leq \rho + |\rho y_+^0 - w(0)|_\infty \leq \rho + K_3 \rho^2 < a, \tag{4.14}$$

$$\begin{aligned} w_2(0) - \tfrac{|c|}{2}\, w_1(0) &\geq \lambda_0^+ \rho \; - K_3 \rho^2 - \tfrac{|c|}{2}\, \rho - \tfrac{|c|}{2}\, K_3\, \rho^2 \\ &= \tfrac{\rho}{2} \Big\{ (c^2 - 4 f'(0))^{1/2} - K_3 \rho (2 + |c|) \Big\} \\ &\geq \rho \Big\{ |f'(0)|^{1/2} - K_3 \rho (1 + \tfrac{|c|}{2}) \Big\} > 0 \quad \forall c \in [-\widehat{c}, 0]. \end{aligned} \tag{4.15}$$

As a consequence of these estimates we have $w(0, \rho, c) \in (0, a) \times (0, \infty)$ for all $c \in [-\widehat{c}, 0]$ and for sufficiently small $\rho > 0$. From now on we fix such a $\rho > 0$ and drop the depence of the solutions on ρ.

Step 2: Continuation of the local unstable manifold.

We choose values $a_{_1}, a_0, a_1$ such that

$$a < a_{-1} < a_0 < a_1 < a_* < 1. \tag{4.16}$$

The goal is to match the unstable manifold of w^0 with the stable manifold of w^1 on the line $w_1 = a_0$. In the following let

$$w(t, c), \;\; t \in J(c) \tag{4.17}$$

be the maximally extended solution of the initial value problem

$$w' = G(w, c), w(0) = w(0, c) \tag{4.18}$$

in the domain

$$W^0 := (0, a_1) \times (0, \infty). \tag{4.19}$$

By Step 1 we know $(-\infty, 0] \subset J(c)$ and in fact $(-\infty, \delta] \subset J(c)$ for all $c \in [-\widehat{c}, 0]$ and some $\delta > 0$ by the local existence theorem for initial value problems. We will show that there exists a unique time $T(c) \in J(c) \cap (0, \infty)$ such that

$$w_1(T(c), c) = a_0 \quad \text{for all } c \in [-\widehat{c}, 0]. \tag{4.20}$$

Step 3: The energy function

Consider the energy function

$$E(w) = \frac{1}{2}w_2^2 + F(w_1),\ w = (w_1, w_2) \in \mathbb{R}^2. \tag{4.21}$$

For $t \in J(c)$ we have by (4.3)

$$\frac{d}{dt}\ E(w(t,c)) = F'(w_1)w_1' + w_2 w_2' = -c\ w_2^2(t,c). \tag{4.22}$$

From Step 1 we conclude

$$\lim_{t\to-\infty} E(w(t,c)) = E(w^0) = 0,$$

so that(4.22) implies for all $t \in J(c)$

$$E(w(t,c)) \begin{cases} = 0 & , c = 0, \\ > 0 & , -\widehat{c} \le c < 0. \end{cases} \tag{4.23}$$

In case $c = 0$ the solution lies on the curve

$$\Gamma = \{(w_1, \sqrt{-2F(w_1)}) : 0 < w_1 < a_1\}, \tag{4.24}$$

while for $c < 0$ it lies above it (recall $F < 0$ in $(0, a_*)$).

Step 4: Intersection with the vertical line $w_1 = a_0$.

The function $w_1(\cdot, c)$ is strictly monotone increasing since by (4.3)

$$w_1'(t,c) = w_2(t,c) > 0,\ t \in J(c). \tag{4.25}$$

Let $f_{\min} = \text{Min}\ \{f(x) : x \in [0,1]\}$, then (4.3) implies

$$w_2'(t) = -f(w_1(t)) - cw_2(t) \le |f_{\min}| + |c|w_2(t),\ t \in J(c).$$

Hence by the Gronwall lemma (cf Appendix 8.5)

$$w_2(t) \le e^{|c|t}(w_2(0) + \int_0^t e^{-|c|s}\ \text{ds}\ |f_{\min}|),\ t \in J(c) \cap [0, \infty). \tag{4.26}$$

The theorem on continuation of solutions leaves us with two alternatives:

Case (i): $J(c) = (-\infty, \infty)$.

Then we infer from Step 3 and (4.25) for all $t \ge 0$

$$w_1'(t) = w_2(t) \ge (-2F(w_1(t)))^{1/2} \ge (-2\ \text{Max}(F(w_1(0)),\ F(a_1)))^{1/2} > 0.$$

Hence w_1 grows unboundedly as $t \to \infty$ which contradicts $w_1(t) < a_1$ for all $t \in J(c)$.

Therefore, the following holds:

Case (ii): $J(c) = (-\infty, t_+(c)),\ t_+(c) < \infty,$

and $w(t)$ approaches the boundary of W^0 as $t \to t_+(c)$ or becomes unbounded. The latter case is excluded by (4.26).
Moreover, $w(t),\ t \geq 0$ lies above the curve Γ from (4.24) and thus cannot approach the lower boundary $[0, a_1] \times \{0\}$ of W^0. Further $w_1(t)$ is strictly increasing, so that the left boundary $\{0\} \times [0, \infty)$ is also excluded. We obtain

$$\text{dist } (w(t), \{a_1\} \times [0, \infty)) \to 0 \quad \text{as} \quad t \to t_+(c).$$

Since $w(t)$ lies above Γ and (4.26) holds we find

$$w_1(t) \to a_1 \quad \text{as} \quad t \to t_+(c).$$

Now w_1 is strictly increasing which together with (4.16) and (4.14) implies that there is a unique $T(c) \in (0, t_+(c))$ satisfying (4.20). Moreover, the relation (4.25) shows that the implicit function theorem applies to (4.20) for every fixed $c \in [-\widehat{c}, 0]$. By the uniqueness of $T(c)$ in $(0, t_+(c))$ this implies $T \in C^1([-\widehat{c}, 0],\ (0, \infty))$.

Step 5: Growth on the vertical line.
We discuss the behavior of

$$w_2(T(c), c),\ \ c \in [-\widehat{c}, 0].$$

By construction we have $w(T(0), 0) \in \Gamma$, i. e.

$$w_2(T(0), 0) = (-2F(a_0))^{1/2} > 0. \tag{4.27}$$

For $c = -\widehat{c}$ we show

$$w_2(T(c), c) > 2(2(F(1) - F(a)))^{1/2}. \tag{4.28}$$

We let $\beta = \frac{|\widehat{c}|}{2}$ and prove that

$$w_2(t, -\widehat{c}) > \beta\ w_1(t, -\widehat{c}) \quad \text{for all} \quad t \in [0, T(-\widehat{c})]. \tag{4.29}$$

For $t = 0$ this inequality follows from (4.15). Suppose there exists $t_0 \in (0, T(-\widehat{c})]$ such that (4.29) holds on $[0, t_0)$ but $w_2(t_0, -\widehat{c}) = \beta\ w_1(t_0, -\widehat{c})$. Then from (4.7),

$$\begin{aligned}
0 &\geq (w_2 - \beta\ w_1)'(t_0, -\widehat{c}) = (-\beta\ w_2 + \widehat{c}\ w_2 - f(w_2))\ (t_0, -\widehat{c}) \\
&= w_2(t_0, -\widehat{c})\Big[\frac{\widehat{c}}{2} - \frac{f(w_1)}{\beta\ w_1}\ (t_0, -\widehat{c})\Big] \\
&\geq w_2(t_0, -\widehat{c})\Big[\frac{\widehat{c}}{2} - \frac{2}{\widehat{c}} \operatorname*{Max}_{0<v\leq 1} \frac{f(v)}{v}\Big] \\
&\geq w_2(t_0, -\widehat{c})\Big[\frac{\widehat{c}}{2} - \frac{\widehat{c}}{4}\Big] > 0,
\end{aligned}$$

a contradiction. In this way inequality (4.28) and (4.7) yield

$$\begin{aligned}
w_2(T(-\widehat{c}), \widehat{c}) > \frac{\widehat{c}}{2}\ a_0 &\geq 2(2(F(1) - F(a))^{1/2}\ \ \frac{a_0}{a} \\
&> 2(2(F(1) - F(a)))^{1/2}.
\end{aligned} \tag{4.30}$$

Step 6: The stable manifold of the target saddle.

As in Step 1 we apply Theorem 4.7 and obtain constants $\rho_0, K_1, K_2, K_3, \eta > 0$ such that the boundary value problem

$$z' = G(z,c), \text{ on } [0,\infty), \tag{4.31}$$

$$P^1_-(z(0) - w^1) = -\rho y^1_-, \tag{4.32}$$

$$|z(t) - w^1|_\infty \le K_1\ \rho \text{ for } 0 \le t < \infty \tag{4.33}$$

has a unique solution $z(\cdot,c) \in C^1([0,\infty),\ \mathbb{R}^2)$ depending smoothly on $c \in [-\widehat{c}, 0]$. Moreover, the following estimates hold for all $0 < \rho \le \rho_0$

$$|z(t,c) - w^1| \le K_2 e^{-\eta t},\ t \ge 0, \tag{4.34}$$

$$|(I - P^1_-)(z(0) - w^1)|_\infty \le K_3 \rho^2. \tag{4.35}$$

Taking ρ sufficiently small (compare (4.13), (4.15)) we can guarantee

$$z(0,c) \in (a_{-1}, 1) \times (0,\infty) =: W^1. \tag{4.36}$$

Now let $\widetilde{J}(c) = (t_-(c), \infty) \supseteq [0,\infty)$ be the maximal interval of existence for the solution of (4.31) in W^1. Since $z_1' = z_2 > 0,\ z_1(\cdot,c)$ is strictly increasing on $\widetilde{J}(c)$. Moreover, the solution cannot approach the lower boundary $[a_{-1}, 1] \times \{0\}$ of W^1, because $z_2(t,c) > 0$ for $t_0 < t \le 0$ and $z_2(t_0,c) = 0$ for some $t_0 < 0$ would imply

$$0 = z_2'(t_0,c) = -cz_2(t_0,c) - f(z_1(t_0,c)) = -f(z_1(t_0,c)) < 0.$$

According to (4.22) the energy $E(z(t,c)),\ c \in [-\widehat{c}, 0]$ is monotone increasing with respect to $t \in \widetilde{J}(c)$, hence

$$\frac{1}{2}\ z_2^2(t,c) + F(z_1(t,c)) \le \lim_{\tau\to\infty} E(z(\tau,c)) = F(1). \tag{4.37}$$

with $F(a) \le F(a_{-1}) \le F(z_1(t,c))$ this gives the a–priori bound

$$z_2^2(t,c) \le 2(F(1) - F(z_1(t,c))) \le 2(F(1) - F(a)),\ t \in \widetilde{J}(c). \tag{4.38}$$

In particular $z_2(\cdot,c)$ is bounded from above on $\widetilde{J}(c)$. The theorem on maximal extension of solutions then shows

$$\text{dist } (z(t,c),\ \{a_{-1}\} \times [0,\infty)) \to 0 \text{ as } t \to t_-(c).$$

Since $z_1(\cdot,c)$ is strictly monotone increasing and $a_{-1} < a_0$ there exists a unique $\widetilde{T}(c) \in (t_-(c), 0)$ such that

$$z_1(\widetilde{T}(c), c) = a_0,\ -\widehat{c} \le c \le 0. \tag{4.39}$$

Again, from $z_1' > 0$ on $(t_-(c), 0)$ and the uniqueness of $\widetilde{T}(c)$ we infer $\widetilde{T} \in C^1([-\widehat{c}, 0], (0,\infty))$ from the implicit function theorem. Moreover, (4.38) yields the estimate

$$z_2(\widetilde{T}(c), c) \le (2(F(1) - F(a)))^{1/2}. \tag{4.40}$$

Step 7: Application of the intermediate value theorem .
Consider the difference

$$\sigma(c) = w_2(T(c), c) - z_2(\widetilde{T}(c), c), \;\; c \in [-\widehat{c}, 0].$$

Equations (4.40) and (4.29) show $\sigma(-\widehat{c}) > 0$.
Moreover, at $c = 0$ we have by (B4)

$$\frac{1}{2}\; w_2^2(T(0), 0) + F(a_0) = E(w^0) = F(0) < F(1) = \frac{1}{2}\; z_2^2(\widetilde{T}(0), 0) + F(a_0),$$

hence

$$\sigma(0) = w_2(T(0), 0) - z_2(\widetilde{T}(0), 0) < 0.$$

Since $\sigma(\cdot)$ is continuous the intermediate value theorem gives a value $c_* \in (-\widehat{c}, 0)$ with $\sigma(c_*) = 0$.
Finally, it is easy to prove that

$$w_*(x) = \begin{cases} w(x + T(c_*), c_*), & x \le 0, \\ z(x + \widetilde{T}(c_*), c_*), & x \ge 0 \end{cases}$$

is a solution of (4.3) satisfying

$$\lim_{x \to -\infty} w_*(x) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \; \lim_{x \to \infty} w_*(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

■

4.2. Results for gradient systems. The scalar evolution equation analyzed in Section 4.1 is a special case of the gradient systems which already appeared in Section 2.4

$$u_t = u_{xx} + \nabla F(u), \; x \in \mathbb{R}, \; t \ge 0, \tag{4.41}$$

where $F \in C^2(\mathbb{R}^m, \mathbb{R})$. Suppose $v_-, v_+ \in \mathbb{R}^m$ are two zeroes of ∇F. From Proposition 2.15 we know that every solution (v_*, c_*) of the system

$$0 = v'' + cv' + \nabla F(v) \;\; \text{in} \;\; \mathbb{R}, \tag{4.42}$$

$$\lim_{\xi \to -\infty} v(\xi) = v_-, \; \lim_{\xi \to \infty} v(\xi) = v_+ \tag{4.43}$$

satisfies the relation

$$c_* \int_{-\infty}^{\infty} |v_*'(\xi)|_2^2 \, d\xi = F(v_-) - F(v_+). \tag{4.44}$$

In the following we report about some recent results of Alikakos, Katzourakis [2]. The results have been adapted to our setting and may be considered as generalizations of Theorem 4.1. The main assumption is

$$\begin{aligned} 0 \;\; &= F(v_-) < F(v_+) = \text{Max}\, \{F(v) : v \in \mathbb{R}^m\}, \\ &v_- \;\; \text{is a local maximum of} \;\; F. \end{aligned} \tag{K1}$$

Note that this is almost implied by (B1)–(B4), except for the fact that $F(v_+)$ in (K1) is a global rather than a local maximum. The assumption $F(v_-) = 0$ is just a normalization.

Theorem 4.3. *([2])*
In addition to (K1) assume the following level set conditions. There exists $\alpha_0 > 0$ and for any $\alpha \in [-\alpha_0, 0]$ two convex compact and disjoint subsets $M_\alpha^-, M_\alpha^+ \subset \mathbb{R}^m$ such that

$$\text{(K2)} \qquad \begin{aligned} &F^{-1}[\alpha, \infty) = M_\alpha^- \cup M_\alpha^+, \ v_\pm \in M_\alpha^\pm, \\ &F^{-1}(\{\alpha\}) = \partial M_\alpha^+ \cup \partial M_\alpha^-, \\ &\partial M_\alpha^+ (\textit{ resp. } \ \partial M_\alpha^-) \ \textit{ is of type } \ C^2 \ \textit{ for } \ \alpha \in [-\alpha_0, 0] \\ &(\textit{ resp. } \ \alpha \in (-\alpha_0, 0)), M_0^- = \{v_-\}. \end{aligned}$$

$$\text{(K3)} \qquad \begin{aligned} &\nabla F(v)^T \ n(v) \le -\gamma_0 < 0 \quad \forall \ v \in \partial M_0^+, \ n(v) \ \textit{ outer normal} \\ &D^2 F(v) \le -\gamma_0 I_m \quad \forall \ v \in M_0^+. \end{aligned}$$

$$\text{(K4)} \qquad \begin{aligned} &\frac{d}{dr} \ F(v_+ + r\xi) < 0 \ \textit{ for all } \ r > 0, \ |\xi|_2 = 1, \ \alpha \in [-\alpha_0, 0) \\ &\textit{with } \ v_+ + r\xi \in M_\alpha^+, \ D^2 F(v_-) \le -\gamma_0 I_m. \end{aligned}$$

Then there exists a value $c_* < 0$ and a corresponding solution $v_* \in C^2(\mathbb{R}, \mathbb{R}^m)$ of (4.42), (4.43).

The proof is quite involved and will not be presented here. Note that all assumptions (K2)–(K4) follow from (B1)–(B4) in the one–dimensional case if we additionally assume $F(b_0) < 0, \ f' = F'' < 0$ on $[a_*, b_0]$ for some $b_0 > 1$. Then (K2) holds with suitable intervals $M_\alpha^\pm$ containing $v_- = 0, \ v_+ = 1$ and (K3), (K4) follow from the sign constraints on F. Of course, the proof of Theorem 4.1 did not make use of the behavior of f outside $[0, 1]$.

An important ingredient in the proof of [2] is the energy functional associated with (4.42)

$$E(v, c) = \int_{-\infty}^{\infty} \Big(\frac{1}{2} \ |v_\xi|_2^2 - F(v) \Big) e^{c\xi} d\xi$$

in suitable function spaces (which enforce the integral to exist). A formal calculation shows that a critical point v of this functional satisfies for all $h \in C_0^\infty(\mathbb{R}, \mathbb{R}^m)$

$$
\begin{aligned}
0 &= \frac{d}{d\varepsilon}\, E(v+\varepsilon h, c)|_{\varepsilon=0} \\
&= \int\limits_{-\infty}^{\infty} \frac{d}{d\varepsilon}\left[\frac{1}{2}(v_\xi+\varepsilon h_\xi, v_\xi+\varepsilon h_\xi)_2 \;- F(v+\varepsilon h)(\xi)\right]_{|\varepsilon=0} e^{c\xi} d\xi \\
&= \int\limits_{-\infty}^{\infty} ((v_\xi, h_\xi)_2 - \nabla F(v)^T h) e^{c\xi} d\xi \\
&= -\int\limits_{-\infty}^{\infty} \left(\frac{d}{d\xi}\,(v_\xi e^{c\xi}), h(\xi)\right)_2 + \nabla F(v)^T h(\xi) e^{c\xi}\, d\xi \\
&= -\int\limits_{-\infty}^{\infty} e^{c\xi}(v_{\xi\xi} + c v_\xi + f(v), h)_2\, d\xi.
\end{aligned}
$$

Here we used integration by parts. Since the relation above holds for all $h \in C_0^\infty(\mathbb{R}, \mathbb{R}^m)$, this implies

$$v_{\xi\xi} + c v_\xi + f(v) = 0 \quad \text{on} \quad \mathbb{R}.$$

In general this procedure does not work since the functional E does not have a global minimum in all reasonable Banach spaces. Moreover, the calculation does not reveal how to determine the wave speed c. In fact, the authors of [2] prove that c_* is finally determined from the condition $E(v_*, c_*) = 0$.

4.3. Spectral projections and hyperbolic equilibria. As a preparation for the parameterized stable manifold theorem in Section 4.4 and the stability theory in Section 5 we discuss in this section parameterized families of hyperbolic matrices. First consider a single matrix $A \in \mathbb{C}^{m,m}$ and decompose its spectrum $\sigma(A) \subset \mathbb{C}$ into two disjoint subsets

$$\sigma(a) = \sigma_s \dot\cup \sigma_u. \tag{4.45}$$

Take a closed contour $\Gamma \subset (\mathbb{C} \setminus \sigma(A))$ which has σ_s in its interior but σ_u in its exterior. By definition this means that

$$\frac{1}{2\pi i} \int\limits_{\Gamma} (z-\lambda)^{-1} dz = \begin{cases} 1, & \lambda \in \sigma_s, \\ 0, & \lambda \in \sigma_u \end{cases}. \tag{4.46}$$

It is easy to construct such a contour by taking the sum of sufficiently small circles enclosing exactly one of the finitely many values in σ_s. Then the matrix

$$P = \frac{1}{2\pi i} \int\limits_{\Gamma} (z I_m - A)^{-1} dz \in \mathbb{C}^{m,m} \tag{4.47}$$

is called the **Riesz projector (or the spectral projector) associated with** σ_s.
This notion is motivated by the following proposition.

Proposition 4.4. Under the assumptions above the matrix P is independent of the choice of contour enclosing σ_s and excluding σ_s. Further, P is the unique projector such that the subspaces

$$X_s = P(\mathbb{C}^m), \; X_u = (I - P)(\mathbb{C}^m) \tag{4.48}$$

have the properties

$$\mathbb{C}^m = X_s \oplus X_u, \; A(X_s) \subset X_s, \; A(X_u) \subset X_u, \tag{4.49}$$

$$\sigma(A_{|X_s}) = \sigma_s, \; \sigma(A_{|X_u}) = \sigma_u. \tag{4.50}$$

If $A \in \mathbb{R}^{m,m}$ and $\sigma_s = \overline{\sigma}_s$, then P is a real projector and (4.49), (4.50) hold with $\mathbb{R}^m = X_s \oplus X_u$.

Remark: The Riesz projector (4.47) may be defined for general linear bounded operators in a Banach space with σ_s being a closed subset of $\sigma(A)$ (see e. g. [29, Ch.V]). It even generalizes to closed unbounded operators in a Banach space, see Section 5 and [29].
Proof: Cauchy's Theorem shows that P is independent of the choice of Γ. For the matrix case here we use the Jordan normal form

$$S^{-1}AS = J \in \mathbb{C}^{m,m}, \; S = (S_s \; S_u), \tag{4.51}$$

where

$$J = \begin{pmatrix} J_1 & & & & & \\ & \ddots & & & & \\ & & J_r & & & \\ & & & J_{r+1} & & \\ & & & & \ddots & \\ & & & & & J_k \end{pmatrix}, \; J_j = \begin{bmatrix} \lambda_j & 1 & & \\ & \lambda_j & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{bmatrix} \in \mathbb{C}^{m_j, m_j}.$$

The eigenvalues have been ordered such that

$$\sigma_s = \{\lambda_j : j = 1, \dots, r\}, \; \sigma_u = \{\lambda_j : j = r+1, \dots, k\}, \; \sigma_s \cap \sigma_u = \emptyset. \tag{4.52}$$

With this transformation we obtain

$$P = \frac{1}{2\pi i} \int_\Gamma (zI_m - SJS^{-1})^{-1}\, \mathrm{dz} = S\, \frac{1}{2\pi i} \int_\Gamma (zI_m - J)^{-1}\, \mathrm{dz} S^{-1},$$
$$= S\, \frac{1}{2\pi i} \int_\Gamma \mathrm{diag}((zI_{m_j} - J_j)^{-1} : j = 1, \dots, k)\, \mathrm{dz} S^{-1}.$$

With

$$J_j = \lambda_j I_{m_j} + E_j,\ E_j = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$$

we find

$$(zI_{m_j} - J_j)^{-1} = (z - \lambda_j)^{-1}(I_{m_j} - (z - \lambda_j)^{-1}E_j)^{-1}$$
$$= (z - \lambda_j)^{-1} \sum_{\nu=0}^{m_j-1} (z - \lambda_j)^{-\nu} E_j^\nu = \sum_{\nu=0}^{m_j-1} (z - \lambda_j)^{-\nu-1} E_j^\nu.$$

By the construction of Γ we have the fomula

$$\frac{1}{2\pi i} \int_\Gamma (z - \lambda_j)^{-\nu-1} \mathrm{dz} = \begin{cases} 1, & \text{if } j = 1, \dots, r,\ \nu = 0, \\ 0, & \text{otherwise }, \end{cases}$$

hence

$$\frac{1}{2\pi i} \int_\Gamma (zI_{m_j} - J_j)^{-1} \mathrm{dz} = \begin{cases} I_{m_j}, & j = 1, \dots, r, \\ 0, & j = r+1, \dots, k \end{cases}.$$

This shows that

$$P = S \begin{bmatrix} I_{m_1} & & & & & \\ & \ddots & & & & \\ & & I_{m_r} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} S^{-1} \tag{4.53}$$

is the projector onto $X_s = R(S_s)$ and $I_m - P$ projects onto $X_u = R(S_u)$. The spectral properties (4.49), (4.50) then follow from (4.51) and (4.52).
In the real case $A \in \mathbb{R}^{m,m}$, $\sigma_s = \overline{\sigma}_s$ we can choose a contour which is symmetric with respect to the real axis:

$$\Gamma = \{\varphi(t) : t \in [0, 2]\},\ \overline{\varphi(t)} = \varphi(2 - t),\ t \in [0, 2].$$

Using this and setting $s = 2 - t$ for $t \in [1, 2]$ leads to

$$\begin{aligned} P &= \frac{1}{2\pi i} \left\{ \int_0^1 (\varphi(t) I_m - A)^{-1} \varphi'(t) \mathrm{d}t + \int_1^2 (\varphi(t) I_m - A)^{-1} \varphi'(t) \ \mathrm{d}t \right\} \\ &= \frac{1}{2\pi i} \left\{ \int_0^1 (\varphi(t) I_m - A)^{-1} \varphi'(t) \ \mathrm{d}t - \int_0^1 (\overline{\varphi(s)} I_m - A)^{-1} \overline{\varphi'(s)} \ \mathrm{d}s \right\} \\ &= \frac{1}{\pi} \ \mathrm{Im} \left\{ \int_0^1 (\varphi(t) I_m - A)^{-1} \varphi'(t) \ \mathrm{d}t \right\}. \end{aligned}$$

Therefore, $X_s = P(\mathbb{R}^m)$ and $X_u = (I - P)(\mathbb{R}^m)$ are real subspaces invariant under A:

$$APv = A \frac{1}{2\pi i} \int_\Gamma (z I_m - A)^{-1} \mathrm{d}z \ v = PAv, \ \ v \in \mathbb{R}^m.$$

Moreover, for $v = Pv \in X_s + iX_s$ we have by (4.51), (4.53)

$$\begin{aligned} Av = APv &= SJS^{-1} S \ \mathrm{diag}(I_{m_1}, \ldots, I_{m_r}, 0, \ldots, 0) S^{-1} v \\ &= S \ \mathrm{diag}(J_1, \ldots, J_r, 0, \ldots, 0) S^{-1} v, \end{aligned}$$

which proves the spectral relation (4.50).

■

In the following proposition we continue the spectral projectors for parameterized families of matrices.

Proposition 4.5. Let $Z \subset \mathbb{R}^p$ be a domain (open, connected) and $M \in C^k(Z, \mathbb{R}^{m,m})$, $k \geq 0$ be a matrix family satisfying

$$\sigma(M(\zeta)) \cap i\mathbb{R} = \emptyset \ \text{ for all } \ \zeta \in Z. \tag{4.54}$$

Then there exists a projector–valued function $P_s \in C^k(Z, \mathbb{R}^{m,m})$ with the following properties

$$M(\zeta) X_s(\zeta) \subset X_s(\zeta) \ \text{ for } \ X_s(\zeta) = R(P_s(\zeta)), \ \zeta \in Z, \tag{4.55}$$

$$\mathrm{rank}\ (P_s(\zeta)) = m_s \ \text{ is independent of } \ \zeta \in Z, \tag{4.56}$$

$$\sigma(M(\zeta)_{|X_s(\zeta)}) = \sigma(M(\zeta)) \cap \{z \in \mathbb{C} : \mathrm{Re}\ z < 0\}, \ \zeta \in Z. \tag{4.57}$$

For any compact subset $Z_0 \subseteq Z$ there exist constants $K, \alpha > 0$ such that for all $\zeta \in Z_0, \ t \geq 0$

$$|\exp(tM(\zeta)) P_s(\zeta)| + |\exp(-tM(\zeta)) \ (I_m - P_s(\zeta))| \leq K \ e^{-\alpha t}. \tag{4.58}$$

Remark: $P_s(\zeta)$ is called the **stable projector** associated with $M(\zeta)$. Correspondingly, $P_u(\zeta) = I - P_s(\zeta)$ is the **unstable projector** which satisfies $M(\zeta)X_u(\zeta) \subset X_u(\zeta)$ for $X_u(\zeta) = R(P_u(\eta))$ as well as

$$\sigma(M(\zeta)_{|X_u(\zeta)}) = \sigma(M(\zeta)) \cap \{z \in \mathbb{C} : \text{Re } z > 0\} \text{ for } \zeta \in Z.$$

Instead of taking $i\mathbb{R}$ as the line separating $\sigma(M(\zeta))$ into two parts, one can take any contour Γ which is symmetric with respect to the real axis and runs from $\mathbb{R} - i\infty$ to $\mathbb{R} + i\infty$.

Proof: Let us first consider a compact set $Z_0 \subseteq Z$. Then there exists $R > 0$ such that for all $\zeta \in Z_0$

$$|\lambda| < R, \text{ Re } \lambda < -\frac{1}{R} \text{ for all } \lambda \in \sigma(M(\zeta)), \text{ Re } \lambda < 0. \tag{4.59}$$

If this is not true we find sequences $\zeta_n \in Z_0,\ \lambda_n \in \sigma(M(\zeta_n)),\ v_n \in \mathbb{C}^m$ such that

$$|v_n| = 1,\ M(\zeta_n)v_n = \lambda_n v_n,\ \text{Re } \lambda_n \to 0 \text{ or } |\lambda_n| \to \infty.$$

Since $|\lambda_n| \leq \sup_{\zeta \in Z_0} |M(\zeta)| < \infty$ we conclude Re $\lambda_n \to 0$ as $n \to \infty$. Taking suitable subsequences we find a value $\zeta = \lim_{n\to\infty} \zeta_n \in Z_0$ such that $M(\zeta)$ has an eigenvalue in $i\mathbb{R}$, which contradicts (4.54). With R from (4.59) the semicircle Γ defined by

$$\varphi(t) = -\frac{1}{R} + iR(2t-1),\ 0 \leq t \leq 1,$$
$$\varphi(t) = -\frac{1}{R} + R \exp\Big(i\pi\Big(t - \frac{1}{2}\Big)\Big),\ 1 \leq t \leq 2,$$

encloses all stable eigenvalues of $M(\zeta)$, $\zeta \in Z_0$. Hence, by Proposition 4.4 the projector satisfying (4.55), (4.57) is

$$P_s(\zeta) = \frac{1}{2\pi i} \int_\Gamma (zI_m - M(\zeta))^{-1} \,\text{dz}, \zeta \in Z_0. \tag{4.60}$$

The formula also shows that P_s is of class C^k.
From the spectral property (4.57) we infer

$$|\lambda| < 1 \text{ for all } \lambda \in \sigma(\exp(M(\zeta))P_s(\zeta)) \cup \sigma(\exp(-M(\zeta))P_u(\zeta)).$$

Therefore, for every $\zeta_0 \in Z$ there exists a norm $|\cdot|_0$ in $\mathbb{R}^m$ (depending on ζ_0) such that for the subordinate matrix norm $|\cdot|_0$

$$|\exp(M(\zeta_0))P_s(\zeta_0)|_0,\ |\exp(-M(\zeta_0))P_u(\zeta_0)|_0 < 1. \tag{4.61}$$

Since M and P_s depend continuously on ζ, we find a neighborhood $U(\zeta_0) \subset Z$ such that for some $\alpha_0 > 0$

$$|\exp(M(\zeta))P_s(\zeta)|_0,\ |\exp(-M(\zeta)P_u(\zeta)|_0 \leq e^{-\alpha_0} < 1 \quad \forall\, \zeta \in U(\zeta_0). \tag{4.62}$$

By the submultiplicativity of $|\cdot|_0$ we obtain for all $n \in \mathbb{N},\ \zeta \in U(\zeta_0)$

$$|\exp(nM(\zeta)P_s(\zeta))|_0,\ |\exp(-nM(\zeta)P_u(\zeta))|_0 \le e^{-n\alpha_0}. \tag{4.63}$$

Finally, every $t \ge 0$ is of the form $t = n + \tau,\ 0 \le \tau < 1$ which leads to

$$|\exp(tM(\zeta)P_s(\zeta))|_0,\ |\exp(-tM(\eta)P_u(\zeta))|_0 \le K_0\ e^{-t\alpha_0}, \zeta \in U(\zeta_0) \tag{4.64}$$

with $K_0 = \sup\ \{|\exp(\tau M(\zeta))P_s(\zeta)|_0 + |\exp(-\tau M(\zeta))P_u(\zeta)|_0 : \tau \in [0,1],\ \zeta \in U(\zeta_0)\}$.
Since we can cover Z_0 with finitely many neighborhoods $U(\zeta_j),\ j = 1,\dots,N$ and since all norms in $\mathbb{C}^m$ are equivalent, the estimate (4.64) yields the assertion (4.58).
Next we exhaust Z by a growing sequence of compact subsets

$$Z_n = \Big\{\zeta \in Z : \operatorname{dist}(\zeta, \partial Z) \ge \frac{1}{n},\ |\zeta| \le n\Big\},\ Z_n \subseteq Z_{n+1} \subseteq \ldots \subseteq Z.$$

The corresponding projectors $P_{s,n}(\zeta),\ \zeta \in Z_n$ satisfy

$$P_{s,k}(\zeta) = P_{s,n}(\zeta) \quad \forall\ k \ge n,\ \zeta \in Z_n$$

by the uniqueness of stable projectors. For a fixed $\zeta \in Z$ we take the first $n = n(\zeta) \in \mathbb{N}$ such that $\zeta \in Z_n$ and define $P_s(\zeta) = P_{s,n}(\zeta)$. In this way $P_s : Z \to \mathbb{R}^{m,m}$ is a C^k–smooth projector–valued mapping satisfying (4.55), (4.57) for all $\zeta \in Z$.
Finally, our result applies to any compact curve Z_0 in Z. Since Z is connected and the rank of a projector is a continuous function, condition (4.56) is proved.

■

4.4. Stable and unstable manifolds of equilibria. Let us first recall some notions for dynamical systems of the form

$$\dot{v} = f(v), \quad \text{where} \quad f \in C^k(\Omega, \mathbb{R}^m),\ k \ge 1, \Omega \subseteq \mathbb{R}^m \quad \text{open}\ . \tag{4.65}$$

In the following let $\Phi_t(v_0),\ t \in J(v_0)$ denote the maximally extended solution of (4.65) satisfying $v(0) = v_0 \in \Omega$.

Definition 4.6. *(Notions of invariance)*

(i) *A set $M \subseteq \Omega$ is called*
  - **positive invariant**, *if $v_0 \in M$ implies $[0,\infty) \subseteq J(v_0)$ and $\Phi_t(v_0) \in M$ for all $t \ge 0$,*
  - **negative invariant**, *if $v_0 \in M$ implies $(-\infty, 0] \subseteq J(v_0)$ and $\Phi_t(v_0) \in M$ for all $t \le 0$,*
  - **invariant**, *if it is positive and negative invariant.*

(ii) *Let $M \subset \Omega$ be compact and invariant, and let $V \supseteq M$ be a neighborhood of M. Then the* **stable set of M with respect to V** *is*

$$W_s^V(M) = \{v_0 \in V : [0,\infty) \subset J(v_0),\ \Phi_t(v_0) \in V \quad \forall\ t \ge 0,$$
$$\lim_{t\to\infty} \operatorname{dist}(\Phi_t(v_0), M) = 0\}$$

and the unstable set of M with respect to V is given by

$$W_u^V(M) = \{v_0 \in M : (-\infty, 0] \subset J(v_0),\ \Phi_t(v_0) \in M \quad \forall t \le 0,$$
$$\lim_{t\to-\infty} \operatorname{dist}(\Phi_t(v_0), M) = 0\}.$$

For $V = \Omega$ we obtain the global stable resp. unstable manifold

$$W_s(M) = W_s^\Omega(M),\ W_u(M) = W_u^\Omega(M).$$

Note than in general

$$W_s^V(M) \overset{\neq}{\subset} W_s(M) \cap V.$$

A typical example where these sets differ, is given by a saddle $M = \{v_+\}$ for which a homoclinic orbit $\{v(t)\}_{t\in\mathbb{R}}$ of (4.65) exists (cf. Section 2.4 for examples). Then $W_s^V(M) \cap V$ also contains an initial piece $\{v(t) : t \le T_-\}$.
In the following we consider a parameterized system

$$\dot{v} = f(v, \zeta),\ f \in C^k(\mathbb{R}^m \times \mathbb{R}^p, \mathbb{R}^m),\ k \ge 1 \tag{4.66}$$

and denote its t–flow by $\Phi_t(\cdot, \zeta)$.
Let $Z \subseteq \mathbb{R}^p$ be a domain such that for all $\zeta \in Z$ the following properties hold

$$f(0, \zeta) = 0, \tag{4.67}$$

$$\sigma(D_v f(0, \zeta)) \cap i\mathbb{R} = \emptyset. \tag{4.68}$$

For simplicity we do not only assume the steady state but also the stable and unstable subspaces of $M(\zeta) := D_v f(0, \zeta)$ to be independent of $\zeta \in Z$, i.e. we have $\mathbb{R}^m = X_s \oplus X_u$ such that for all $\zeta \in Z$

$$M(\zeta)X_s \subseteq X_s,\ \operatorname{Re} \lambda < 0 \ \text{ for all } \ \lambda \in \sigma(M(\zeta)_{|X_s}), \tag{4.69}$$

$$M(\zeta)X_u \subseteq X_u,\ \operatorname{Re} \lambda > 0 \ \text{ for all } \ \lambda \in \sigma(M(\zeta)_{|X_u}). \tag{4.70}$$

Let P_s resp. $P_u = I_m - P_s$ be the corresponding projectors onto X_s resp. X_u. Later we show how this situation can be achieved in the proof of Theorem 4.1.

Theorem 4.7. *(Parameterized stable manifold theorem) Let the assumptions above hold and let $Z_0 \subset Z$ be open and bounded with $\bar{Z}_0 \subseteq Z$. Then there exist zero neighborhoods $V_s \subseteq X_s$, $V_u \subseteq X_u$, $V \subseteq \mathbb{R}^m$ with $V_s \oplus V_u \subseteq V$ and a function $h_u \in C^k(V_s \times Z_0,\ V_u)$ such that the following holds.*

(i) *For every $v_s \in V_s$, $\zeta \in Z_0$ the boundary value problem*

$$v' = f(v, \zeta) \quad \text{on} \quad [0, \infty), \tag{4.71}$$

$$P_s v(0) = v_s,\ v(t) \in V \quad \text{for all} \quad t \ge 0, \tag{4.72}$$

has a unique solution $v = v(\cdot, v_s, \zeta) \in C^{k+1}([0, \infty), V)$.

(ii) *For the function* $h_u : V_s,\ \zeta \in Z_0 \to X_u$ *defined by*

$$v(0, v_s, \zeta) = v_s + h_u(v_s, \zeta), \quad v_s \in V_s, \zeta \in Z_0, \tag{4.73}$$

the following assertions hold

$$\begin{aligned} &\{v_s + h_u(v_s, \zeta) : v_s \in V_s\} = W_s^V(0, \zeta) \cap (V_s \oplus V_u) = \\ &\{v_0 \in V_s \oplus V_u : \Phi_t(v_0, \zeta) \in V \quad \forall\, t \geq 0,\ \Phi_t(v_0, \zeta) \to 0 \quad as \quad t \to \infty\}. \end{aligned} \tag{4.74}$$

(iii) *There exist constants* $K, \eta > 0$ *such that for all* $v_s \in V_s$ *and* $\zeta \in Z_0$*,*

$$|v(t, v_s, \zeta)| \leq K\ e^{-\eta t},\ \ t \geq 0, \tag{4.75}$$

$$D_v h_u(0, \zeta) = 0, \tag{4.76}$$

$$|D_v^2 h_u(v_s, \zeta)| \leq K \ (in\ case\ k \geq 2). \tag{4.77}$$

Before proving the theorem let us interpret the result. Equation (4.74) shows that the local stable manifold is the graph of the function $h_u : X_s \to X_u$ which inherits its smoothness from the vector field f. Because of (4.76) the manifold is tangent to X_s at 0, i.e. $T_0\ W_s^v(0, \zeta) = X_s$, see Definition 3.21.
Note also that instead of a single chart we have a single parametrizing map $v_s \to v_s + h_u(v_s, \zeta)$, which is a submersion and hence the inverse of a chart. The local stable manifold is therefore a submanifold of $\mathbb{R}^m$ (see e. g. [5, Ch.VIII,9] for the general theory). However, the global stable manifold $W_s(0, \zeta)$ is not necessarily a submanifold of $\mathbb{R}^m$ and the general manifold concept (see Definition 3.5) is needed. A common procedure is to write

$$W_s(0, \zeta) = \bigcup_{t \leq 0} \Phi_t(W_s^V(0, \zeta)) \tag{4.78}$$

and to define the chart in a neighborhood U of some $\Phi_{-T}(v_0),\ T > 0$, by applying the local chart in $W_s^V(0, \zeta)$ to $\Phi_T(U)$.
We mention that (4.74) characterizes only the intersection of the local stable manifold with the product neighborhood $V_s \oplus V_u$. It is possible to arrange that $V = V_s \oplus V_u$. However, this requires a rather careful geometric construction and is not needed for our application.
Finally, let us formulate the analogous result for the local unstable manifold which holds unter the assumptions of Theorem 4.7. This result can be obtained by reversing time in (4.71) and then applying the stable manifold theorem.
There exists $h_s \in C^k(V_u \times Z_0, V_s)$ with the following properties:

(i') for every $v_u \in V_u,\ \zeta \in Z_0$ the boundary value problem

$$\begin{aligned} v' = f(v, \zeta) \quad &\text{on} \quad (-\infty, 0] \\ P_u v(0) = v_u,\ v(t) \in V \quad &\text{for all} \quad t \leq 0 \end{aligned} \tag{4.79}$$

has a unique solution $v(\cdot, v_u, \zeta) \in C^{k+1}((-\infty, 0], V)$.

(ii') For all $v_u \in V_u,\ \zeta \in Z_0$,

$$v(0, v_u, \zeta) = v_u + h_s(v_u, \zeta), \tag{4.80}$$

$$\begin{aligned} &\{v_u + h_s(v_u, \zeta) : v_u \in V_u\} = W_u^V(0, \zeta) \cap (V_s \oplus V_u) = \\ &\{v_0 \in V_s \oplus V_u : \Phi_t(v_0, \zeta) \in V\ \forall\ t \le 0,\ \lim_{t \to -\infty} \Phi_t(v_0, \zeta) = 0\}. \end{aligned} \tag{4.81}$$

(iii') There exist constants $K, \zeta > 0$ such that for all $v_u \in V_u$ and $\zeta \in Z_0$,

$$|v(t, v_u, \zeta)| \le K\ e^{\eta t},\ \ t \le 0, \tag{4.82}$$

$$D_v h_s(0, \zeta) = 0, \tag{4.83}$$

$$|D_v^2 h_s(v_u, \zeta)| \le K \text{ (in case } k \ge 2). \tag{4.84}$$

Proof of Theorem 4.7. We apply the parameterized Lipschitz inverse mapping Theorem 8.3 with the following settings:

$$\begin{aligned} X =& C_{\mathrm{b}}^1([0, \infty), \mathbb{R}^m), \quad \|v\|_X = \|v\|_\infty + \|v'\|_\infty, \\ Y =& C_{\mathrm{b}}([0, \infty), \mathbb{R}^m) \times X_s, \quad \|(r, v_s)\|_Y = \|r\|_\infty + \|v_s\|, \\ \Lambda =& Z_0 \times V_s, \quad \lambda = (\zeta, v_s^0), \\ L(\lambda)v =& (v' - D_v f(0, \zeta)v, P_s v(0)), \\ F(v, \lambda) =& (D_v f(0, \zeta)v - f(v, \zeta), -v_s^0). \end{aligned}$$

Let us first show that the assumptions (i)-(iii) of Theorem 8.3 are satisfied for suitable constants δ, ℓ, ρ.

Assumption (i): Let $A = A(\zeta) = Df(0, \zeta)$ for $\zeta \in Z$. In the following we will suppress the dependence on $\zeta \in Z$ whenever appropriate. A solution $v \in C_{\mathrm{b}}^1([0, \infty), \mathbb{R}^m)$ of the initial value problem

$$v' + Av = r \in C_{\mathrm{b}}([0, \infty), \mathbb{R}^m), \quad P_s v(0) = v_s \in X_s \tag{4.85}$$

is given by the formula

$$v(t) = \exp(tA)v_s + \int_0^\infty G(t, s)r(s)ds, \quad t \ge 0 \tag{4.86}$$

with the Green's function

$$G(t, s) = \begin{cases} \exp((t - s)A)P_s, & 0 \le s \le t, \\ \exp((t - s)A)(P_s - I), & 0 \le t < s. \end{cases} \tag{4.87}$$

By Proposition 4.5 there exist constants $K, \alpha \ge 0$ such that for all $\tau \ge 0, \zeta \in \overline{Z_0}$

$$\|\exp(\tau A)P_s\| + \|\exp(-\tau A)(P_s - I)\| \le K\exp(-\alpha\tau). \tag{4.88}$$

Because of these estimates the integral in (4.86) exists and the following estimate holds

$$\|v(t)\| \le K \left(\exp(-\alpha t)\|v_s\| + \int_0^\infty \exp(-\alpha|t - s|)\|r\|_\infty ds \right)$$

$$\leq K \left(\exp(-\alpha t)\|v_s\| + \frac{2}{\alpha}\|r\|_\infty \right)$$

It is not difficult to verify that (4.86) actually solves (4.85). Then from the differential equation we find

$$\|v'(t)\| \leq \|A\|_{\infty,Z_0}\|v(t)\| + \|r(t)\|, \quad \|A\|_{\infty,Z_0} = \sup_{\zeta\in\overline{Z_0}} \|A(\zeta)\|.$$

We have shown that $v \in X$ holds and satisfies the estimate

$$\|v\|_X \leq \ell(\|r\|_\infty + \|v_s\|), \quad \ell = 1 + K(1 + \|A\|_{\infty,Z_0})\mathrm{Max}(\frac{2}{\alpha}, 1).$$

It remains to be shown, that v is the unique bounded solution of (4.85). Since the problem is linear it is enough to show that any solution $w \in X$ of the homogeneous problem (4.85) is trivial. For this purpose, first note that

$$w(t) = \exp(tA)w(0) = \exp(tA)(I - P_s)w(0).$$

Since the projectors commute with A we obtain from (4.88)

$$\begin{aligned} \|(I - P_s)w(0)\| =& \|(I - P_s)^2 w(0)\| = \|(I - P_s)\exp(-tA)w(t)\| \\ =& \|\exp(-tA)(I - P_s)w(t)\| \leq K\exp(-\alpha t)\|w\|_\infty. \end{aligned}$$

Since w is bounded this shows $(I - P_s)w(0) = 0$ if we let $t \to \infty$. Therefore, $w(t) = \exp(tA)(I - P_s)w(0) = 0$ holds for all $t \geq 0$. □

5. Stability of travelling waves in parabolic systems

In this chapter we study the stability of relative equilibria with respect to perturbations of initial data. This will not be done in the abstract framework of Section 3.5. Rather we will work with an Abelian Lie group which simplifies the analysis considerably. We will partly follow the approach in [18, Ch.4] where our main application is to travelling waves in parabolic systems as in Section 2.3. The equivariance of the evolution equation implies that the linearized operator of the comoving frame equation (3.3) has eigenvalues on the imaginary axis. The dimension of the corresponding invariant subspace is at least the dimension of the Lie group. In the Abelian case there is only the zero eigenvalue.

A key assumption of the nonlinear stability theorem requires the linearization to have no further eigenvalues on the imaginary axis and all other parts of the spectrum (essential or point spectrum) to be in the left half plane and strictly bounded away from the imaginary axis.

The subsequent sections will therefore be devoted to the spectral theory of second order differential operators on the whole real line.

5.1. Stability with asymptotic phase. As in Section 2.3 we consider a parabolic system

$$u_t = Au_{xx} + f(u), \; x \in \mathbb{R}, \; t \geq 0 \tag{5.1}$$

with $A \in \mathbb{R}^{m,m}$ positive definite and $f : \mathbb{R}^m \to \mathbb{R}^m$ sufficiently smooth. Moreover, assume that there exists a travelling wave solution

$$u(x,t) = \; v_\star(x - c_\star t), \; x \in \mathbb{R}, \; t \in \mathbb{R}, \tag{5.2}$$

where $v_\star \in C_b^2(\mathbb{R}, \mathbb{R}^m)$ and

$$\lim_{\xi \to \pm\infty} v_\star(\xi) = v_\pm, \; f(v_\pm) = 0. \tag{5.3}$$

Then $v_\star$ is an equilibrium of the comoving frame equation (3.3) from Section 3.1,

$$v_t = Av_{\xi\xi} + c_\star v_\xi + f(v), \; \xi \in \mathbb{R}, \; t \geq 0, \tag{5.4}$$

$$v(\cdot, 0) = u_0. \tag{5.5}$$

Definition 5.1. *The travelling wave* $(v_\star, c_\star)$ *is called*

– **orbitially stable** *w.r.t. a norm* $|| \; ||_Y$ *in a Banach space* Y *iff the following condition holds: for all* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that every solution* $v(\cdot, t)$ *of (5.4), (5.5) with* $u_0 - v_\star \in Y$, $||u_0 - v_\star||_Y \leq \delta$ *satisfies*

$$\inf_{\gamma \in \mathbb{R}} ||v(\cdot, t) - v_\star(\cdot - \gamma)||_Y \; \leq \varepsilon \quad \forall t \geq 0, \tag{5.6}$$

– **asymptotically stable with asymptotic phase**, *if it is orbitally stable and for any* u_0 *with* $||u_0 - v_\star||_Y \leq \delta$ *there exists* $\gamma = \gamma(u_0) \in \mathbb{R}$ *such that*

$$||(v(\cdot, t) - v_\star(\cdot - \gamma(u_0))||_Y \to 0 \quad as \quad t \to \infty. \tag{5.7}$$

Note that we did not specify Y yet.
Suitable function spaces will be $Y = H^1(\mathbb{R}, \mathbb{R}^m)$ or $Y = C^1_{\text{unif}}\,(\mathbb{R}, \mathbb{R}^m)$. Moreover, we defined stability in terms of the transformed equation (5.4). For the original equation with

$$v(\xi, t) = u(\xi + c_\star t, t),$$

condition (5.6) translates into

$$\varepsilon \geq \inf_{\gamma \in \mathbb{R}} ||(v(\cdot + c_\star t, t) - v_\star(\cdot - \gamma)|| = \inf_{\gamma \in \mathbb{R}} ||u(\cdot, t) - v_\star(\cdot - \gamma)||_Y,$$

and

$$||u(\cdot, t) - v_\star(\cdot - c_\star t - \gamma(u_0))||_Y \to 0 \; \text{ as } \; t \to \infty.$$

Before starting the general theory, let us make some heuristic arguments that pave the way for the theory.

1. **Asymptotic behaviour of $v_\star$** By Lemma 2.6 the function $w_\star = \binom{v_\star}{v'_\star}$ is an orbit connecting $w_- = \binom{v_-}{0}$ to $w_+ = \binom{v_+}{0}$ for the system

(5.8)
$$w' = G(w) = \begin{pmatrix} w_2 \\ -c_\star A^{-1}\ w_2 - A^{-1}\ f(w_1) \end{pmatrix}.$$

We have the Jacobians

(5.9)
$$DG(w) = \begin{pmatrix} 0 & I_m \\ -A^{-1}\ Df(w_1) & -c_\star A^{-1} \end{pmatrix},\ DG(w_\pm) = \begin{pmatrix} 0 & I_m \\ -A^{-1}\ Df(v_\pm) & -c_\star A^{-1} \end{pmatrix}.$$

Our first lemma guarantees hyperbolicity of $DG(w_\pm)$ uniformly in some parameters.

Lemma 5.2. *Suppose the matrices $A, B \in \mathbb{R}^{m,m}$ satisfy*

(5.10)
$$x^T Ax > 0 > x^T Bx \quad \forall x \in \mathbb{R}^m,\ x \neq 0.$$

Then there exists $\varepsilon > 0$ such that the matrix family

(5.11)
$$M(\tau, c) = \begin{pmatrix} 0 & I_m \\ -A(\tau)^{-1}B & -cA(\tau)^{-1} \end{pmatrix},\ A(\tau) = \tau\ I_m + (1-\tau)A, \tau \in [-\varepsilon, 1+\varepsilon]$$

satisfies the assumptions of Proposition 4.5 with $\zeta = (\tau, c),\ Z = (-\varepsilon, 1+\varepsilon) \times \mathbb{R}$. The corresponding projectors $P_s(\zeta)$ and $P_u(\zeta) = I - P_s(\zeta)$ both have rank m.

Proof. Take $\varepsilon > 0$ such that for all $\tau \in [-2\varepsilon, 1+2\varepsilon]$

(5.12)
$$x^T A(\tau)x = \tau x^T x + (1-\tau)x^T Ax > 0 \ \text{ for all } \ x \neq 0.$$

Let $\lambda \in \mathbb{C}$ be an eigenvalue of $M(\tau, c)$ with eigenvector $\binom{x}{y} \in \mathbb{C}^{2m}$, then we find form (5.11)

(5.13)
$$y = \lambda x,\ (\lambda^2 A(\tau) + c\lambda I_m + B)x = 0.$$

Note that $\binom{x}{y} \neq 0$ implies $x \neq 0$ in this case. Equation (5.13) is the characteristic equation of the differential operator

$$\mathcal{L}v = A(\tau)v'' + cv' + Bv.$$

Suppose that (5.13) holds for some $\lambda = i\omega,\ \omega \in \mathbb{R}$ and $x = x_1 + i\ x_2$, then multiply by x^H and take the real part

$$0 = -\omega^2(x_1^T A(\tau)x_1 + x_2^T\ A(\tau)x_2) + x_1^T Bx_1 + x_2^T Bx_2.$$

However, the right–hand side is negative due to (5.10) and (5.12). This proves condition (4.54) of Proposition 4.5. In order to determine the common rank of $P_s(\zeta), P_u(\zeta)$, it suffices to choose the special parameter set $\tau = 1,\ c = 0$. Then (5.13) reads

(5.14)
$$\lambda^2 x = -Bx.$$

Due to (5.10) the eigenvalues $\mu_j \in \mathbb{C}$ of $-B$, $j = 1, \ldots, m$ (repeated according to multiplicity) satisfy $\mathrm{Re}\ \mu_j > 0$. Hence the quadratic eigenvalue problem (5.14) has eigenvalues $\lambda_{j,\pm} = \pm\sqrt{\mu_j}$, where $\sqrt{\cdot}$ is the principle branch of the complex square root. These eigenvalues satisfy $\mathrm{Re}\ \lambda_{j,-} < 0 < \mathrm{Re}\ \lambda_{j,+}$, $j = 1, \ldots, m$, so that both projectors P_s, P_u have rank m. $\square$

Lemma 5.2 applies to $DG(w_\pm)$ from (5.9) if

(5.15) $$x^T Df(v_\pm)\, x < 0 \quad \forall x \neq 0.$$

Remark 5.3. *In the gradient case (4.41) this condition holds if the Hessians $D^2F(v_\pm) = Df(v_\pm)$ are negative definite, i.e. $v_\pm$ are local minima as in Theorem 4.3.*

From (5.15) we infer that $w_\pm$ are saddles of (5.8). Then the Theorem 4.7 on stable and unstable manifolds shows that for suitable constants $C, \eta > 0$

(5.16) $$\begin{aligned} |w_\star(\xi) - w_+| &\leq C\ e^{-\eta\xi},\ \xi \geq 0, \\ |w_\star(\xi) - w_-| &\leq C\ e^{\eta\xi},\ \xi \leq 0. \end{aligned}$$

For the original wave $v_\star$ this implies

(5.17) $$|v_\star(\xi) - v_\pm| \ + \ |v_\star'(\xi)| \leq C\ e^{-\eta|\xi|},\ \xi \in \mathbb{R}_\pm.$$

Moreover, we obtain for $\xi \in \mathbb{R}_\pm$

$$\begin{aligned} |w_\star'(\xi)| = \left| \begin{pmatrix} v_\star'(\xi) \\ v_\star''(\xi) \end{pmatrix} \right| &= |G(w_\star(\xi), c_\star)| \\ &= |G(w_\star(\xi), c_\star) - G(w_\pm, c_\star| \leq C|w_\star(\xi) - w_\pm| \leq C\ e^{-\eta|\xi|}. \end{aligned}$$

Hence, also the second derivative $v_\star''(\xi)$ decays exponentially. If $v_\star \in C^k(\mathbb{R}, \mathbb{R}^m)$ one can proceed in this way and conclude

(5.18) $$|v_\star^{(j)}(\xi)| \leq C\ e^{-\eta|\xi|} \ \text{ for } \ \xi \in \mathbb{R}_\pm,\ j = 1, \ldots, k.$$

In particular, the derivatives $v_\star^{(j)}$ lie in the Sobolev spaces $W^{k-j,p}(\mathbb{R}, \mathbb{R}^m)$, $p \in \mathbb{N}$ for $j = 1, \ldots, k$.

2. The linearized operator

The linearization of (5.4) at $v = v_\star$ is

(5.19) $$v_t = \mathcal{L}v = Av_{\xi\xi} + c_\star v_\xi + Df(v_\star)v.$$

We expect this equation to govern the dynamics of (5.4), (5.5) close to $v_\star$. For example, let

$$u_0 = v_\star + v_0, \ v(\cdot,t) = v_\star + w(\cdot,t)$$

for $v_0, w(\cdot,t)$ small. Then we find

$$\begin{aligned} v_t = w_t &= Av_{\star,\xi\xi} + c_\star \, v_{\star,\xi} + f(v_\star) + A \, w_{\xi\xi} + c_\star \, w_\xi + f(v_\star + w) - f(v_\star) \\ &= \mathcal{L} \, w + f(v_\star + w) - f(v_\star) - Df(v_\star)w \\ &= \mathcal{L} \, w + R(w, v_\star), \end{aligned}$$

and we expect the remainder $R(w, v_\star)$ to be small if w is small. Differentiating the equation

$$0 = Av_{\star,\xi\xi} + cv_{\star,\xi} + f(v_\star(\xi)), \ \xi \in \mathbb{R} \tag{5.20}$$

with respect to ξ we obtain

$$0 = A(v_{\star,\xi})_{\xi\xi} + c(v_{\star,\xi})_\xi + Df(v_\star(\xi))v_{\star,\xi} = \mathcal{L}v_{\star,\xi}. \tag{5.21}$$

Hence $v_{\star,\xi}$ is in the kernel of $\mathcal{L}$ in any reasonable function space, compare the exponential decay (5.17), (5.18).

Another way to derive (5.21) is to note that $v_\gamma(\xi) = v_\star(\xi - \gamma), \ \xi \in \mathbb{R}$ solves (5.20) for all $\gamma \in \mathbb{R}$ and then differentiate with respect to γ at $\gamma = 0$.

The eigenvalue zero of $\mathcal{L}$ is the reason why we cannot expect asmptotic stability in the classical Lyapunov sense but only asymptotic stability with asymptotic phase.

Our final example shows that the kernel of the linearized operator may have dimension > 1 if the Lie group of equivariance has dimension > 1.

Example 5.4. Consider the comoving frame equation (2.87) for the cubic NLS, (see Example 2.26, (2.80), (2.83))

$$u_t = i \, u_{xx} + i \, |u|^2 u, \tag{5.22}$$

given by the ansatz $u(x,t) = e^{-i\theta_\star t} \, v(x - c_\star t, t)$ as follows

$$v_t = i \, v_{xx} + c_\star v_x + i \, \theta_\star v + i \, v|v|^2. \tag{5.23}$$

Equation (5.23) has a two–parameter family of equilibria

$$v_\star(\xi) = c_3 \, \sqrt{2} \, \exp(ic_2\xi) \, \mathrm{sech}(c_3\xi), \ c_\star = 2c_2, \ \theta_\star = -c_2^2 - c_3^2. \tag{5.24}$$

Since the right–hand side of (5.23) is equivariant with respect to the action

$$[a(\gamma_1, \gamma_2)v](x) = e^{-i\gamma_1} v(x - \gamma_2), \ x \in \mathbb{R} \tag{5.25}$$

the whole family

$$v_\star(\gamma_1,\gamma_2) := a(\gamma_1,\gamma_2)v_\star,\ \ \gamma \in \mathbb{R}^2 \tag{5.26}$$

yields equilibria of (5.23).
Let us write (5.23) as a real system with the settings

$$v = v_1 + iv_2,\ \ |v|^2 = v_1^2 + v_2^2,\ \ J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_t = J\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_{xx} + c_\star \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_x + \theta_\star J \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} -v_2\ |v| \\ v_1\ |v| \end{pmatrix} = \mathcal{L}_0 v + R(v).$$

5.2. Spectral theory for second order operators. The previous considerations suggest to study the spectrum of second order operators

$$\mathcal{L}u = Au_{xx} + B(x)u_x + C(x)u,\ \ x \in \mathbb{R},\ \ u : \mathbb{R} \to \mathbb{R}^m, \tag{5.27}$$

where

$$A \in \mathbb{R}^{m,m},\ \ x^T(A + A^T)x \geq \alpha_0\ |x|^2 \quad \forall\, x \in \mathbb{R}^m \ \text{ for some } \ \alpha_0 > 0, \tag{5.28}$$

$$B, C \in C_b(\mathbb{R}, \mathbb{R}^{m,m}), \tag{5.29}$$

$$B(x) \to B_\pm,\ \ C(x) \to C_\pm \ \text{ as } \ x \to \pm\infty. \tag{5.30}$$

Note $B = c_\star I$, $C(x) = Df(v_\star(x))$, $C_\pm = Df(v_\pm)$ in the traveling wave case. Consider

$$\mathcal{L} : \mathcal{D}(\mathcal{L}) \subset X \to X$$

for the cases

$$\mathcal{D}(\mathcal{L}) = H^2(\mathbb{R}, \mathbb{R}^m),\ \ X = L^2(\mathbb{R}, \mathbb{R}^m),$$
$$\mathcal{D}(\mathcal{L}) = C^2_{\text{unif}}(\mathbb{R}, \mathbb{R}^m),\ \ X = C^0_{\text{unif}}(\mathbb{R}, \mathbb{R}^m).$$

Definition 5.5. *Let X be a Banach space and*

$$\mathcal{L} : \mathcal{D}(\mathcal{L}) \subset X \to X$$

be a linear operator.

(i) $\mathcal{L}$ *is called* **closed**, *if*
$(\mathcal{D}(\mathcal{L}) \ni\ u_n \to u \in X,\ \mathcal{L}u_n \to r\ (n \to \infty))$ *implies*
$u \in \mathcal{D}(\mathcal{L}),\ \mathcal{L}u = r.$

(ii) *Let $\mathcal{L} : \mathcal{D} \subset X \to X$ be closed.*
The **resolvent set** *is defined by*

$$\rho(\mathcal{L}) = \{s \in \mathbb{C} : sI - \mathcal{L} : \mathcal{D}(\mathcal{L}) \to X \ \text{ is bijective and there exists } \ K > 0 \ \text{ with } \ ||(sI - \mathcal{L})^{-1}\ h||_X \leq K\ ||h||_X\}.$$

Further we define:

spectrum: $\sigma(\mathcal{L}) = \mathbb{C} \setminus \ \rho(\mathcal{L})$,

point spectrum: $\sigma_p(\mathcal{L}) = \{s \in \sigma(\mathcal{L}) : sI - \mathcal{L}$ *is Fredholm of index 0 and* s *is an eigenvalue of finite algebraic multiplicity*$\}$,

eigenvalues: $s \in \sigma(\mathcal{L})$ *with* $N(sI - \mathcal{L}) \neq \{0\}$,

algebraic multiplicity: $\text{alg}(s) = \sup\limits_{k\in\mathbb{N}} \ \dim \ N((sI - \mathcal{L})^k)$,

essential spectrum: $\sigma_{\text{ess}}(\mathcal{L}) = \sigma(\mathcal{L}) \setminus \ \sigma_p(\mathcal{L})$.

Example 5.6. $X = C[0,1], \ || \ ||_X = || \ ||_\infty, \ \mathcal{L}u = au, \ a \in C[0,1]$ with a strictly monotone increasing. Then we have

$$\rho(\mathcal{L}) = \mathbb{C} \setminus [a(0), a(1)], \ \sigma_p(\mathcal{L}) = \emptyset, \ \sigma_{\text{ess}}(\mathcal{L}) = [a(0), a(1)].$$

Suppose there exists $u \in X, \ s \in [a(0), a(1)]$ with $au = \mathcal{L}u = su$. This yields $u(x) = 0$ for $x \in [0,1] \setminus \{\bar{x}\}$ where $a(\bar{x}) = s$. Since u is continuous we obtain $u \equiv 0$. Hence there is no point spectrum.

If $\mathcal{L}$ is closed, then $\mathcal{D}(\mathcal{L})$ becomes a Banach space with respect to the graph norm

$$||u||_{\mathcal{D}(\mathcal{L})} = ||u||_X + ||\mathcal{L}u||_X, \ u \in \mathcal{D}(\mathcal{L}), \tag{5.31}$$

and $\mathcal{L} : (\mathcal{D}(\mathcal{L}), || \ ||_{\mathcal{D}(\mathcal{L})}) \to (X, || \ ||_X)$ is bounded.

Lemma 5.7. *Let (5.28), (5.29) be satisfied. Then*

$$\mathcal{L} = A\partial_x^2 + B\partial_x + C : \mathcal{D}(\mathcal{L}) = H^2(\mathbb{R}, \mathbb{R}^m) \to L^2(\mathbb{R}, \mathbb{R}^m)$$

is closed.
Moreover, there exist $\varepsilon_0, C_1, C_0 > 0$ *such that for all*

$$s \in \Omega_0 = \{s \in \mathbb{C} : |s| \geq C_0, \ |\arg(s)| \leq \pi/2 + \varepsilon_0\},$$

the equation $(sI - \mathcal{L})u = h + g_x, \ u \in H^2, \ h \in L^2, \ g \in H^1$ *implies*

$$\begin{aligned} &|s| \ ||u||^2 + ||u_x||^2 \leq C_1(||g||^2 + \frac{1}{|s|} \ ||h||^2), \\ &|s|^2 \ ||u||^2 + |s| \ ||u_x||^2 + ||u_{xx}||^2 \leq C_1(||h||^2 + |s| \ ||g||^2 + ||g_x||^2). \end{aligned} \tag{5.32}$$

Proof. Let us first show that (5.32) implies closedness of $\mathcal{L}$. Suppose $\mathcal{L}u_n = h_n \to h$ in L^2, $u_n \to u$ in L^2. Pick $s_0 \in \Omega_0$ and observe

$$(s_0 I - \mathcal{L})u_n = s_0 u_n - h_n \to s_0 u - h \quad \text{in} \quad L^2. \tag{5.33}$$

Then by (5.32)

$$\begin{aligned} |s_0|^2 \, ||u_n - u_m||^2 + |s_0| \, ||u_{n,x} - u_{m,x}||^2 + ||u_{n,xx} - u_{m,xx}||^2 \\ \leq \; C_1 ||s_0(u_n - u_m) - (h_n - h_m)||^2. \end{aligned}$$

Hence u_n is a Cauchy sequence in H^2 and for some $\widetilde{u} \in H^2$

$$||u_n - \widetilde{u}||_{H^2} \to 0.$$

Since $u_n \to u$ in $\mathcal{L}^2$ we obtain $u = \widetilde{u} \in H^2$, $||u_n - u||_{H^2} \to 0$.
Therefore, from (5.33)

$$s_0 u - h = \lim_{n \to \infty} (s_0 I - \mathcal{L})u_n = s_0 u - \mathcal{L}u \text{ and } \; \mathcal{L}u = h.$$

a-priori estimate

Let $(sI - \mathcal{L})u = h + g_x$ in $\mathbb{R}$, $u \in H^2$, $h \in L^2, g \in H^1$. Multiply by u^H and integrate

$$(u, su) - (u, Au_{xx}) = (u, Bu_x) + (u, Cu) + (u, h) + (u, g_x).$$

Use integration by parts (note that C_0^∞ is dense in $H^k, k \geq 0$) and obtain

$$s||u||^2 + (u_x, Au_x) = (u, Bu_x) + (u, Cu) + (u, h) - (u_x, g). \tag{5.34}$$

In the following we frequently use Young's inequality

$$ab \leq \frac{\alpha}{4} \, a^2 + \frac{1}{\alpha} \, b^2 \quad \forall \; a, b \in \mathbb{R}, \;\; \alpha > 0. \tag{5.35}$$

Take abolute values in (5.34) and find

$$\begin{aligned} |s| \; ||u||^2 \leq ||A||_\infty \; ||u_x||^2 + ||B||_\infty \; ||u|| \; ||u_x|| + ||C||_\infty \; ||u||^2 \\ + ||u|| \; ||h|| + ||u_x|| \; ||g|| \\ \leq K_0 ||u_x||^2 + K_1 ||u||^2 + ||u|| \; ||h|| + \frac{1}{2} \; ||g||^2. \end{aligned} \tag{5.36}$$

Take the real part in (5.34),

$$\begin{aligned} \text{Re } s \; ||u||^2 + \alpha_0 ||u_x||^2 \leq ||B||_\infty \; ||u|| \; ||u_x|| + ||C||_\infty ||u||^2 + ||u|| \; ||h|| + ||u_x|| \; ||g|| \\ \leq \frac{\alpha_0}{4} \; ||u_x||^2 + \frac{1}{\alpha_0} \; ||B||_\infty^2 \; ||u||^2 \\ + ||C||_\infty \; ||u||^2 + ||u|| \; ||h|| + \frac{\alpha_0}{4} \; ||u_x||^2 + \frac{1}{\alpha_0} \; ||g||^2. \end{aligned}$$

We obtain

$$\text{Re } s \; ||u||^2 + \frac{\alpha_0}{2} \; ||u_x||^2 \; \leq \; K_2 \; ||u||^2 + ||u|| \; ||h|| + \frac{1}{\alpha_0} \; ||g||^2. \tag{5.37}$$

Case 1: Re $s \geq |\text{Im } s|, \;\text{Re } s > 0, \;|s| \geq 2\sqrt{2}\; K_2$.
Since $0 < \text{Re } s \leq |s| \leq \sqrt{2}$ Re s, equation (5.37) implies

$$\frac{|s|}{\sqrt{2}}\;||u||^2 + \frac{\alpha_0}{2}\;||u_x||^2 \leq \frac{|s|}{2\sqrt{2}}\;||u||^2 + \frac{1}{\alpha_0}\;||g||^2 + \frac{|s|}{4\sqrt{2}}\;||u||^2 + \frac{\sqrt{2}}{|s|}\;||h||^2,$$

$$\frac{|s|}{4\sqrt{2}}\;||u||^2 + \frac{\alpha_0}{2}\;||u_x||^2 \leq \frac{1}{|\alpha_0|}\;||g||^2 + \frac{\sqrt{2}}{|s|}\;||h||^2 \quad (\text{cf. } (5.32)).$$

Case 2: $|\text{Im } s| \geq \text{ Re } s \geq 0$.
Use (5.37) in (5.36) and note Re $s \geq 0$,

$$\begin{aligned} |s|\;||u||^2 \leq\; & \frac{2||A||_\infty}{\alpha_0}\;(K_2\;||u||^2 + ||u||\;||h|| \;+\; \frac{1}{\alpha_0}\;||g|^2) \\ & + K_1\;||u||^2 + ||u||\;||h|| + \frac{1}{2}\;||g||^2 \\ \leq\; & K_3(||u||^2 + ||u||\;||h|| + ||g||^2). \end{aligned} \tag{5.38}$$

Take $|s| > 2\;K_3$, then $K_3 \leq \frac{|s|}{2}$ and

$$|s|\;||u||^2 \leq \frac{|s|}{2}\;||u||^2 + K_3||g||^2 + \frac{|s|}{4}||u||^2 + \frac{K_3^2}{|s|}\;||h||^2,$$

$$\frac{|s|}{4}\;||u||^2 \leq K_4(||g||^2 + \frac{||h||^2}{|s|}). \tag{5.39}$$

Insert this into (5.37), take $|s| \geq 4K_2$,

$$\begin{aligned} \frac{\alpha_0}{2}\;||u_x||^2 &\leq K_2\;||u||^2 + \frac{|s|}{4}\;||u||^2 + \frac{1}{|s|}\;||h||^2 + \frac{1}{\alpha_0}\;||g||^2 \\ &\leq \frac{|s|}{2}\;||u||^2 + \frac{1}{|s|}\;||h||^2 + \frac{1}{\alpha_0}||g||^2 \\ &\leq\; 2K_4(||g||^2 + \frac{||h||^2}{|s|}) + \frac{1}{|s|}\;||u||^2 + \frac{1}{\alpha_0}\;||g||^2 \\ &\leq\; K_5(\frac{||h||^2}{|s|} + ||g||^2). \end{aligned}$$

Combining this with (5.39) yields (5.32).

Case 3: Re $s \leq 0, \;|\text{Re } s| \leq \varepsilon|\text{Im } s|$.
Take the imaginary part in (5.34) to find by (5.37)

$$\begin{aligned} |\text{Im } s|\;||u||^2 &\leq ||A||_\infty\;||u_x||^2 + K_1||u||^2 + ||u||\;||h|| + \frac{1}{2}\;||g||^2 \\ &\leq\; ||A||_\infty\;\frac{2}{\alpha_0}\;(|\text{Re } s|\;||u||^2 + K_2||u||^2 + ||u||\;||h|| + \frac{1}{\alpha_0}\;||g||^2). \end{aligned}$$

For $\varepsilon \frac{2}{\alpha_0} \|A\|_\infty \leq \frac{1}{2}$ we end up with

$$|\text{Im } s| \; ||u||^2 \leq K_6(||u||^2 + ||u|| \; ||h|| + ||g||^2).$$

Since

$$|s| \leq (1+\varepsilon^2)^{1/2} |\text{Im } s|,$$

we obtain

$$|s| \; ||u||^2 \leq K_6(1+\varepsilon^2)^{1/2}(||u||^2 + ||u|| \; ||h|| + ||g||). \tag{5.40}$$

Now we proceed as in case 2 after (5.38).
Finally, (5.32) and $(sI - \mathcal{L})u = h + g_x$ imply by (5.32)

$$\begin{aligned} ||u_{xx}||^2 &= || - A^{-1}(-su + Bu_x + Cu + h + g_x)||^2 \\ &\leq \; K_7(|s|^2 \; ||u||^2 + ||u_x||^2 + ||h||^2 + ||g_x||^2) \\ &\leq \; K_8(|s| \; ||g||^2 + ||h||^2 + ||h||^2 + ||g_x||^2). \end{aligned}$$

Combining this with the first part of (5.32) shows the second part of (5.32). □

Remark 5.8. *Equation (5.32) implies that there exists $K > 0$ such that*

$$||u||^2_{H^1} \leq K(||g||^2 + ||h||^2) \quad \forall \; s \in \Omega_0.$$

In particular, there are no eigenvalues in Ω_0. We will later show that $sI - \mathcal{L}$, $s \in \Omega_0$ is indeed Fredholm of index 0. Then (5.32) implies $\Omega_0 \subset \rho(\mathcal{L})$.

5.3. Fredholm properties and essential spectrum.

5.4. Linear evolution equations with second order operators.

5.5. The nonlinear stability theorem.

see [8, Ch.4.2], [14, Ch.5.4], [18], [26]

6. Numerical analysis of travelling waves

7. Travelling waves in Hamiltonian PDEs

8. Appendix

In this appendix we summarize various results from functional analysis and its applications which are used throughout the text.

8.1. Linear functional analysis. We assume that the reader is familiar with standard concepts from linear functional analysis as in [3], [29], for example. Here we collect some extra results that are useful when transforming matrices resp. operators into block diagonal form.

Lemma 8.1 (The Sylvester equation)**.** *Let X_1, X_2 be Banach spaces and $\Lambda_1 \in \mathcal{L}[X_1, X_1]$, $\Lambda_2 \in \mathcal{L}[X_2, X_2]$. Then the Sylvester equation*

$$\Lambda_1 P - P\Lambda_2 = R \in L[X_2, X_1] \tag{8.1}$$

has a unique solution $P \in L[X_2, X_1]$ for every $R \in L[X_2, X_1]$ if the spectra of Λ_1 and Λ_2 are disjoint.

Proof. Since the spectra $\sigma(\Lambda_1)$ and $\sigma(\Lambda_2)$ are compact and disjoint we find a simple closed contour $\Gamma \subseteq \mathbb{C}$ which has $\sigma(\Lambda_1)$ in its interior but has $\sigma(\Lambda_2)$ in its exterior. Then we claim that

$$P = \frac{1}{2\pi i}\int_\Gamma (zI - \Lambda_1)^{-1} R (zI - \Lambda_2)^{-1} dz \tag{8.2}$$

solves (8.1). Equation (8.2) is also known as **Rosenblum's formula**, cf [28]. In fact, insert P from (8.2) into (8.1) and use the resolvent identities $\Lambda_j(zI - \Lambda_j)^{-1} = -I + z(zI - \Lambda_j)^{-1}, j = 1, 2$ to find

$$\Lambda_1 P - P\Lambda_2 = \frac{1}{2\pi i}\int_\Gamma -R(zI - \Lambda_2)^{-1} dz + \frac{1}{2\pi i}\int_\Gamma (zI - \Lambda_1)^{-1} R dz$$

By Cauchy's theorem the first integral vanishes since $(zI - \Lambda_2)^{-1}$ is holomorphic in the interior of Γ, and the second integral gives R since the complete spectrum of Λ_1 lies in the interior of Γ. As for uniqueness, suppose that $P \in L[X_2, X_1]$ solves the homogeneous Sylvester equation (8.1). Then we obtain $(zI - \Lambda_1)P - P(zI - \Lambda_2) = 0$ for all $z \in \mathbb{C}$ and hence

$$P(zI - \Lambda_2)^{-1} = (zI - \Lambda_1)^{-1} P \quad \forall z \in \Gamma.$$

By integration we find

$$P = \frac{1}{2\pi i}\int_\Gamma (zI - \Lambda_1)^{-1} P dz = \frac{1}{2\pi i}\int_\Gamma P(zI - \Lambda_2)^{-1} dz = 0.$$

∎

Under the assumptions of Lemma 8.1 the so called **Sylvester operator**

$$\mathcal{S} : \begin{cases} L[X_2, X_1] & \mapsto L[X_2, X_1], \\ P & \to \Lambda_1 P - P\Lambda_2 \end{cases} \tag{8.3}$$

is a linear homeomorphism. The boundedness of $\mathcal{S}^{-1}$ follows from the inverse operator theorem but may also be read off directly from Rosenblum's formula (8.2). The inverse of its norm

$$\|\mathcal{S}^{-1}\|^{-1}_{L[L[X_2,X_1],L[X_2,X_1]]} = \mathrm{sep}(\Lambda_1, \Lambda_2) \tag{8.4}$$

is also called the **separation of** Λ_1 **and** Λ_2, see [12, Chap.7]. Take, for example, a perturbed pair $M_j \in L[X_j, X_j]$ satisfying $\|M_1 - \lambda_1\| + \|M_2 - \Lambda_2\| \le \varepsilon$. Then the perturbed operator $\mathcal{S}_\varepsilon P = M_1 P - P M_2$ satisfies $\|\mathcal{S} - \mathcal{S}_\varepsilon\| \le \varepsilon$. By the Banach Lemma $\mathcal{S}_\varepsilon$ is invertible provided $\varepsilon < \mathrm{sep}(\Lambda_1, \Lambda_2)$ and the inverse satisfies
The quadratic extension of the Sylvester equation is the **Riccati equation**.

Lemma 8.2 (Riccati equation)**.**

8.2. Semigroup theory. [20], [23], [14]

8.3. Sobolev spaces. [1]

8.4. Calculus in Banach spaces. The following result is a parameterized version of a Lipschitz inverse mapping theorem

Theorem 8.3 (Parameterized Lipschitz inverse mapping theorem)**.** *Let X, Y, Z be Banach spaces, let $B_\delta(0) = \{x \in X : \|x\| \le \delta\}, \delta > 0$ be a ball in X and let $\Lambda \subseteq Z$ be open. Further, let operators*

$$L \in C^k(\Lambda, L[X, Y]), \quad F \in C^k(B_\delta(0) \times \Lambda, Y), \quad k \ge 0$$

and constants $\ell > 0, \rho \ge 0$ be given with the following properties:

(i) *The maps $L(\lambda) \in L[X, Y], \lambda \in \Lambda$ are homeomorphisms and*

$$\|L(\lambda)^{-1}\|_{L[X,Y]} \le \ell \quad \forall \lambda \in \Lambda. \tag{8.5}$$

(ii) *For all $x_1, x_2 \in B_\delta(0)$ and $\lambda \in \Lambda$*

$$\|F(x_1, \lambda) - F(x_2, \lambda)\| \le \rho \|x_1 - x_2\|. \tag{8.6}$$

(iii) *The constants satisfy $\rho\ell < 1$ and*

$$\|F(0, \lambda)\| \le (\ell^{-1} - \rho)\delta \quad \forall \lambda \in \Lambda. \tag{8.7}$$

Then the equation

$$L(\lambda)x + F(x, \lambda) = 0 \tag{8.8}$$

has a unique solution $x = x(\lambda) \in B_\delta(0)$ for each $\lambda \in \Lambda$ and the solution function satisfies $x(\cdot) \in C^k(\Lambda, B_\delta(0))$. Moreover, the following inequality holds for all $x_1, x_2 \in B_\delta(0)$ and $\lambda \in \Lambda$

$$\|x_1 - x_2\| \le \frac{\ell}{1 - \ell\rho} \|L(\lambda)x_1 + F(x_1, \lambda) - (L(\lambda)x_2 + F(x_2, \lambda))\|. \tag{8.9}$$

Proof. Let us rewrite (8.8) as a fixed point equation

$$x = -L(\lambda)^{-1}F(x,\lambda) =: T(x,\lambda). \tag{8.10}$$

From (8.5) and (8.6) we have for all $x_1, x_2 \in B_\delta(0), \lambda \in \Lambda$

$$\|T(x_1,\lambda) - T(x_2,\lambda)\| \le \ell \|F(x_1,\lambda) - F(x_2,\lambda)\| \le \ell\rho\|x_1 - x_2\|. \tag{8.11}$$

Using this and (8.7) shows that for all $x \in B_\delta(0)$,

$$\begin{aligned}\|T(x,\lambda)\| \ &\le \|T(0,\lambda)\| + \|T(x,\lambda) - T(0,\lambda)\| \\ &\le \ell\|F(0,\lambda)\| + \ell\rho\|x\| \\ &\le \ell(\ell^{-1} - \rho)\delta + \ell\rho\delta = \delta.\end{aligned}$$

Hence $T(\cdot,\lambda)$ maps $B_\delta(0)$ into itself and is a contraction with constant $q = \ell\rho < 1$. Thus (8.10) has a unique solution $x(\lambda) \in B_\delta(0)$ for all $\lambda \in \Lambda$. Moreover, from (8.5),(8.6), we obtain the following estimate for all $x_1, x_2 \in B_\delta(0), \lambda \in \Lambda$

$$\begin{aligned}\|x_1 - x_2\| \le &\|x_1 - T(x_1,\lambda) - (x_2 - T(x_2,\lambda))\| + \|T(x_1,\lambda) - T(x_2,\lambda)\| \\ \le &\ell\|L(\lambda)x_1 + F(x_1,\lambda)) - (L(\lambda)x_2 + F(x_2,\lambda))\| + \rho\ell\|x_1 - x_2\|,\end{aligned}$$

which implies (8.9). Setting $x_1 = 0, x_2 = x(\lambda)$ in (8.9) yields

$$\|x(\lambda)\| \le \frac{\ell}{1-\ell\rho}\|F(0,\lambda)\|, \quad \lambda \in \Lambda. \tag{8.12}$$

In case $k = 0$ we also obtain continuity of $x(\cdot)$ at $\lambda_0 \in \Lambda$ from (8.9)

$$\begin{aligned}\|x(\lambda) - x(\lambda_0)\| \le &\frac{\ell}{1-\ell\rho}\|L(\lambda)x(\lambda_0) + F(x(\lambda_0),\lambda)\| \\ = &\frac{\ell}{1-\ell\rho}\|(L(\lambda) - L(\lambda_0))x(\lambda_0) + F(x(\lambda_0),\lambda) - F(x(\lambda_0),\lambda_0)\|.\end{aligned}$$

By the continuity of L and F, the right-hand side converges to zero as $\lambda \to \lambda_0$. In case $k \ge 1$ the C^k-smoothness of $x(\cdot)$ follows as in the implicit function theorem for Frechét differentiable functions. For completeness we indicate the first step. Assuming $x(\cdot)$ to be differentiable in Λ, we find from differentiating the equation

$$0 = L(\lambda)x(\lambda) + F(x(\lambda),\lambda), \lambda \in \Lambda$$

formally

$$\begin{aligned}0 = &(DL(\lambda)h)x(\lambda) + L(\lambda)Dx(\lambda)h + D_xF(x(\lambda),\lambda)Dx(\lambda)h \\ &+ D_\lambda F(x(\lambda),\lambda)h \quad \text{for } h \in Z.\end{aligned}$$

Hence we expect the derivative $d_0 := Dx(\lambda_0)$ at $\lambda = \lambda_0$ to satisfy the equation

$$(L(\lambda_0) + D_xF(x(\lambda_0),\lambda_0))d_0h = -(DL(\lambda_0)h)x(\lambda_0) - D_\lambda F(x(\lambda_0),\lambda_0)h. \tag{8.13}$$

With (8.5),(8.6) one shows that the operator $L(\lambda_0) + D_x F(x(\lambda_0), \lambda_0) \in L[X, Y]$ on the left of (8.13) is a linear homeomorphism, and then defines $d_0 \in L[Z, X]$ via (8.13). With this definition one continues to show

$$x(\lambda_0 + h) - x(\lambda_0) - d_0 h = o(\|h\|), \quad h \in Z,$$

which proves that $Dx(\lambda_0)$ exists and coincides with d_0. Higher derivatives are then obtained by further differentiation of (8.13). □

8.5. Analysis on manifolds.

8.6. The Gronwall Lemma.

References

[1] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and applied mathematics series ; 140*. Acad. Press, Amsterdam [u.a.], 2. ed., digital print. edition, 2009.

[2] Nicholas D. Alikakos and Nikolaos I. Katzourakis. Heteroclinic travelling waves of gradient diffusion systems. *Trans. Amer. Math. Soc.*, 363(3):1365–1397, 2011.

[3] Hans Wilhelm Alt. *Lineare Funktionalanalysis*. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2006.

[4] Herbert Amann. *Ordinary differential equations : an introduction to nonlinear analysis*, volume 13 of *de Gruyter studies in mathematics ; 13*. de Gruyter, Berlin [u.a.], 1990.

[5] Herbert Amann and Joachim Escher. *Analysis II*. Grundstudium Mathematik. Birkhäuser Verlag, Basel, zweite, korrigierte auflage edition, 2006.

[6] Nicolas Bourbaki. *Lie groups and Lie algebras/1 = Ch. 1/3*. 1975.

[7] Pascal Chossat and Reiner Lauterbach. *Methods in equivariant bifurcations and dynamical systems*, volume 15 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[8] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate studies in mathematics ; 19*. American Math. Soc., Providence, RI, 2. ed. edition, 2010.

[9] Michael J. Field. *Dynamics and symmetry*, volume 3 of *ICP Advanced Texts in Mathematics*. Imperial College Press, London, 2007.

[10] Paul C. Fife. *Mathematical aspects of reacting and diffusing systems*, volume 28 of *Lecture notes in biomathematics ; 28*. Springer, Berlin [u.a.], 1979.

[11] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.*, 1:445–466, 1961.

[12] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 2012.

[13] Martin Golubitsky and Ian Stewart. *The symmetry perspective*, volume 200 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2002. From equilibrium to chaos in phase space and physical space.

[14] Daniel Henry. *Geometric theory of semilinear parabolic equations*, volume 840 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1981.

[15] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. 117*, 117:500–544, 1952.

[16] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge Univ. Press, Cambridge [u.a.], 2. ed. edition, 2013.

[17] Nathan Jacobson. *Lie algebras*, volume 10 of *Interscience tracts in pure and applied mathematics ; 10*. Wiley, New York [u.a.], 1962.

[18] Todd Kapitula and Keith Promislow. *Spectral and Dynamical Stability of Nonlinear Waves*. Applied Mathematical Sciences ; 185. Springer New York, New York, NY, 2013.

[19] Serge Lang. *Differential and Riemannian manifolds*, volume 160 of *Graduate texts in mathematics ; 160*. Springer, New York [u.a.], 3. ed., 2. corr. printing edition, 1996.

[20] Milan Miklavčič. *Applied functional analysis and partial differential equations*. World Scientific, Singapore [u.a.], 1998.

[21] Robert M. Miura. Accurate computation of the stable solitary wave for the FitzHugh-Nagumo equations. *J. Math. Biol.*, 13(3):247–269, 1981/82.

[22] J. D. Murray. *Mathematical Biology*, volume 19 of *Biomathematics*. Springer-Verlag, Berlin, 1989.

[23] Amnon Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied mathematical sciences ; 44*. Springer, New York [u.a.], 2., corr. print edition, 1983.

[24] A. D. Polyanin, editor. *EQ WORLD The World of Mathematical Equations*. http://eqworld.ipmnet.ru/index.htm, 2004-.

[25] Wulf Rossmann. *Lie groups : an introduction through linear groups*, volume 5 of *Oxford graduate texts in mathematics ; 5*. Oxford Univ. Press, Oxford [u.a.], paperback ed. edition, 2006.

[26] Jens Rottmann-Matthes. Travelling Waves, 2014. Manuscript of Lectures held at the Univ. of Stuttgart.

[27] Jean Pierre Serre. *Lie algebras and Lie groups : 1964 lectures given at Harvard University*, volume 1500 of *Lecture notes in mathematics ; 1500*. Springer, Berlin [u.a.], 1992.

[28] Gilbert W. Stewart and Ji-guang Sun. *Matrix perturbation theory*. Computer science and scientific computing. Acad. Press [u.a.], Boston [u.a.], digital print. edition, 2004.

[29] Angus E. Taylor and David C. Lay. *Introduction to functional analysis*. Wiley, New York [u.a.], 2. ed. edition, 1980.

[30] Frank W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate texts in mathematics ; 94*. Springer, New York, NY [u.a.], repr. edition, 1983.

[31] Gerald B. Whitham. *Linear and nonlinear waves*. Pure and applied mathematics. Wiley, New York [u.a.], 1974.

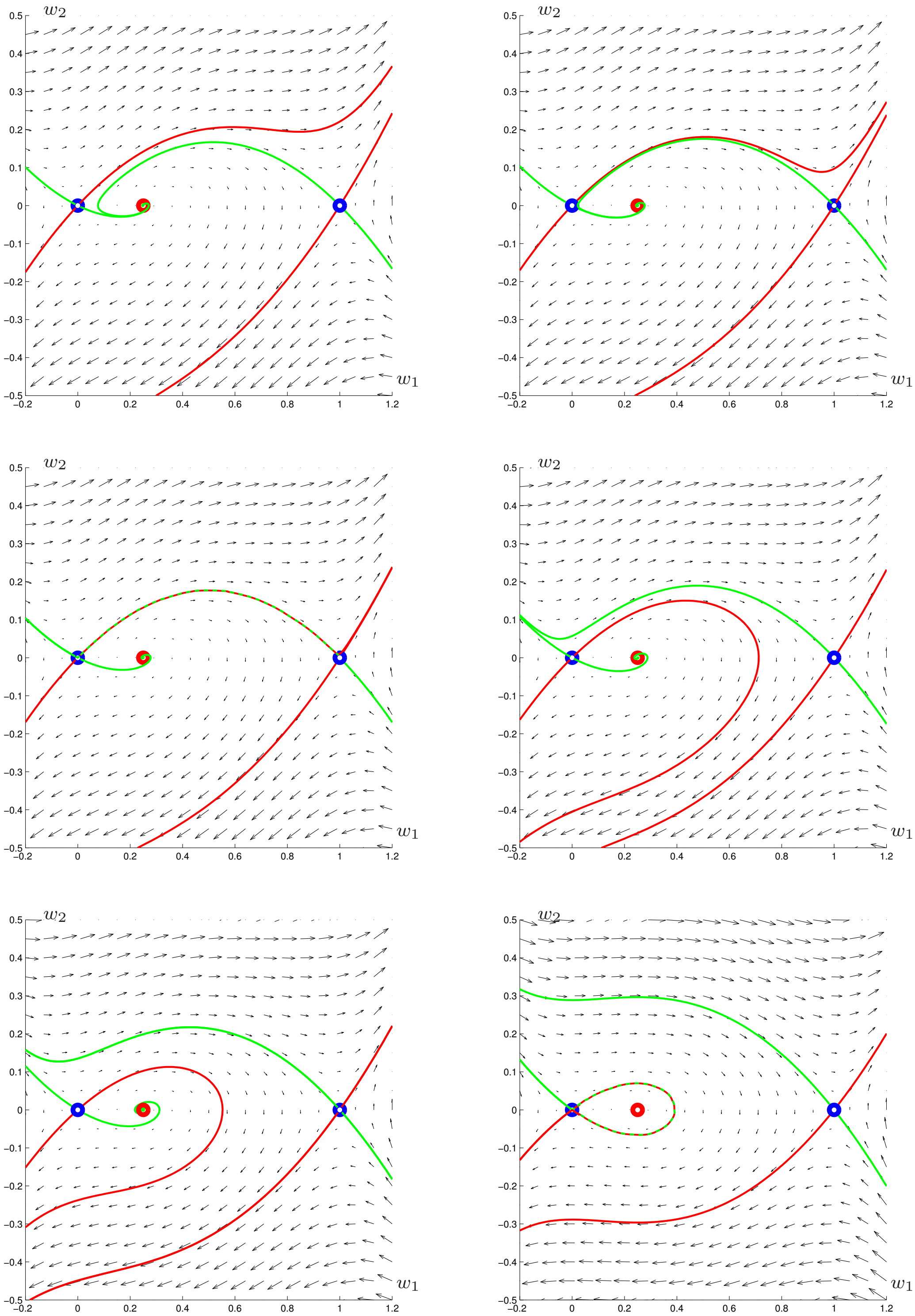

FIGURE 2.4. Phase plane for the Nagumo example with values $b = \frac{1}{4}$ and $c = -0.4,\ -0.35,\ c_{\star},\ -0.3,\ -0.2,\ 0$

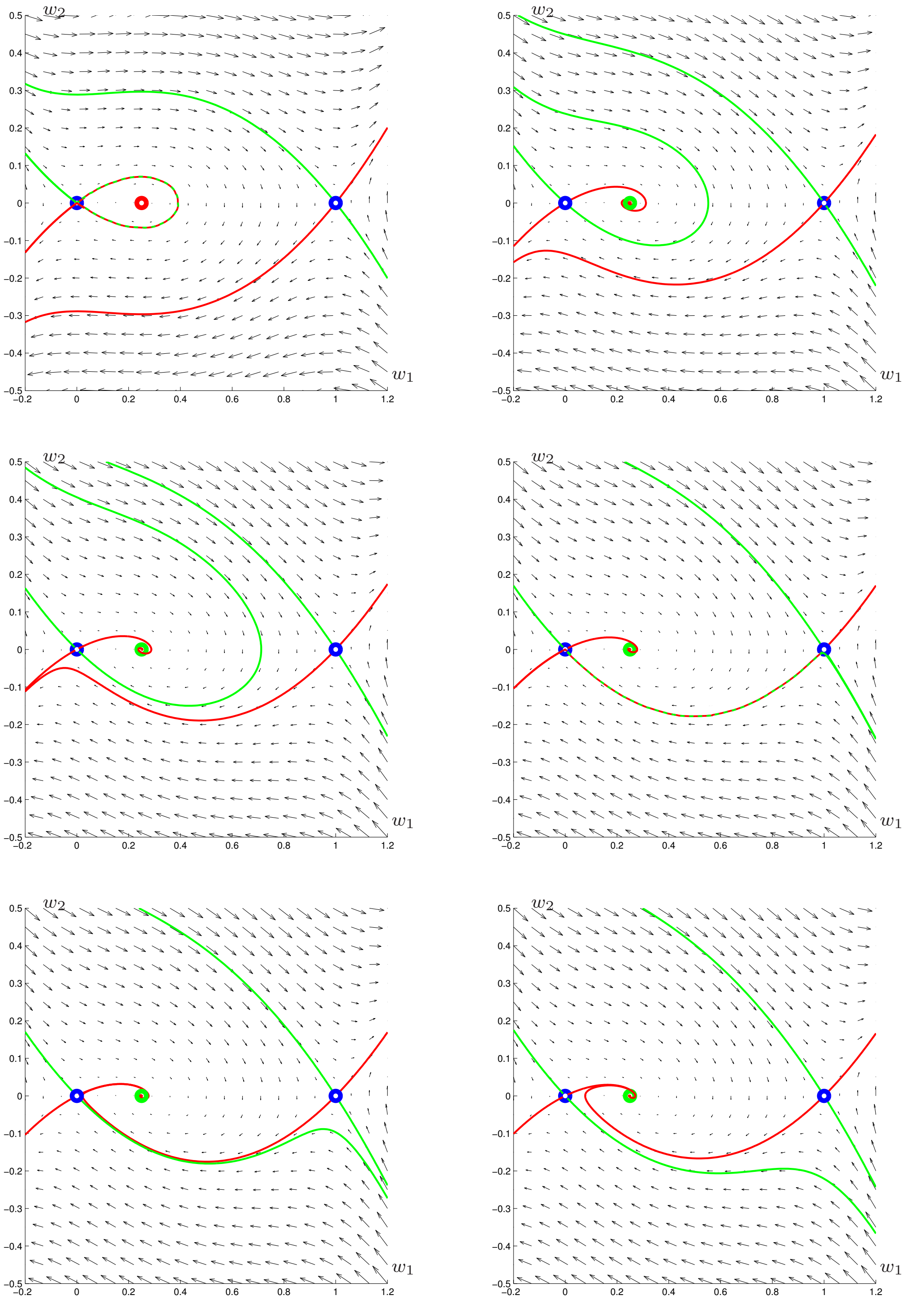

FIGURE 2.5. Phase plane for the Nagumo example with values $b = \frac{1}{4}$ and $c = 0$, 0.2, 0.3, $c^{\star}$, 0.36, 0.4

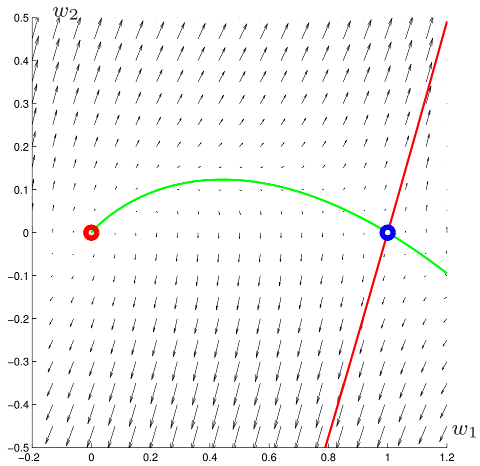

FIGURE 2.6. Phase plane for the Fisher example $c = c_\star = -2$

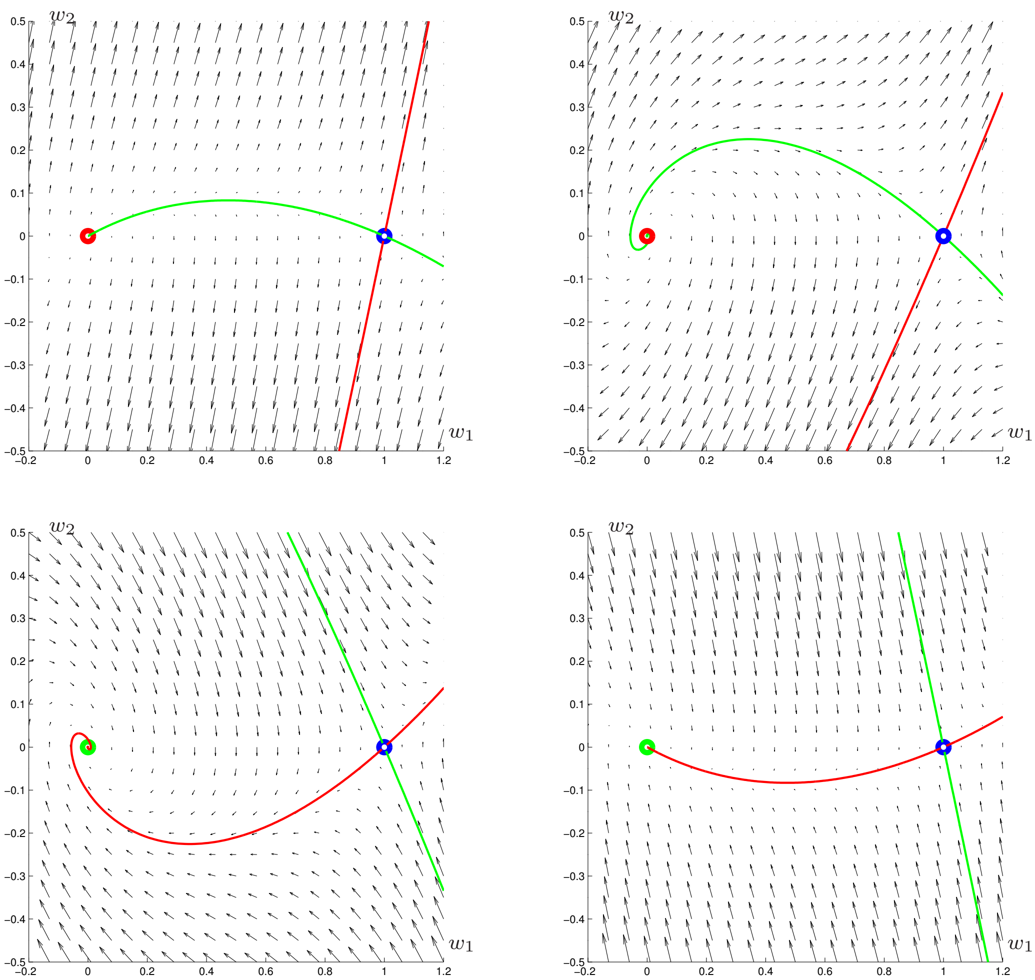

FIGURE 2.7. Phase plane for the Fisher example $c = -3, -1, 1, 3$

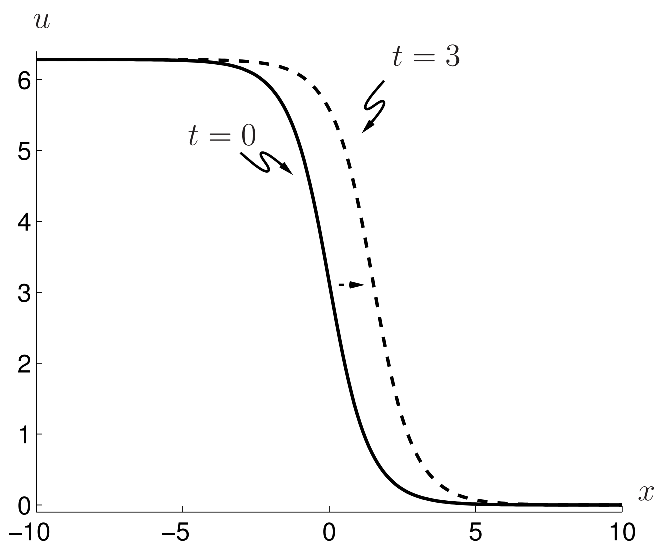

FIGURE 2.8. Front of the sine Gordon equation: $c = \frac{1}{2}$

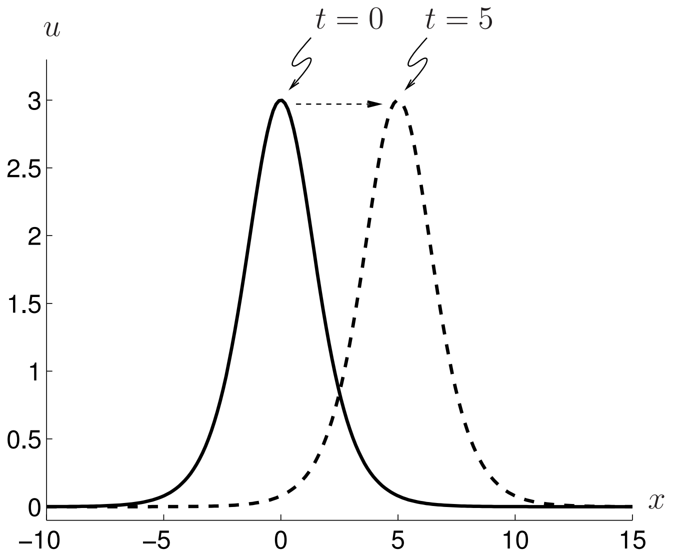

FIGURE 2.9. Pulse in the Korteweg-de Vries (KdV) equation: $c = 1$

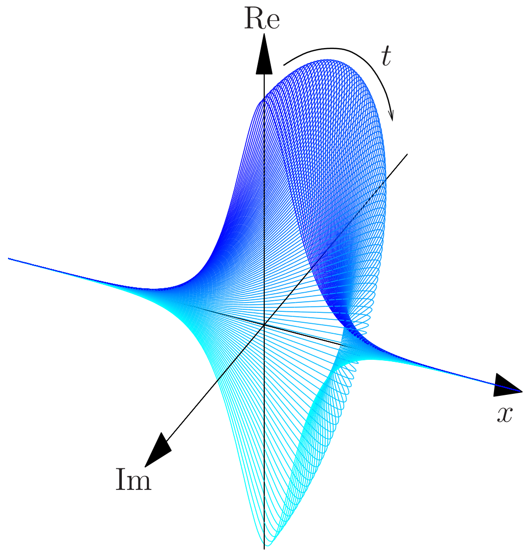

FIGURE 2.10. Oscillating pulse of the cubic NLS with $b = 2$, $c_2 = 1$