16th Internet Seminar
on Evolution Equations

Operator Semigroups and Dispersive Equations

Lecture Notes

Dirk Hundertmark
Lars Machinek
Martin Meyries
Roland Schnaubelt

Karlsruhe, Halle, February 21, 2013

Contents

LECTURE 1

Introduction: evolution equations, operator semigroups and dispersive equations

Evolution equations govern the dynamic behavior of deterministic systems in the sciences. This is done in a surprisingly simple way, only based on a few fundamental assumptions.

One first requires that at time $t \geq 0$ the *state* of the given system can uniquely be described by an element of the *state space* X, where we assume that X is a Banach space. Given an *initial state* $u_0 \in X$ at time $t = 0$, we denote the state of the system at time $t \geq 0$ by $u(t) = u(t; u_0) \in X$ if $u(0) = u_0$.

We second suppose that the change of the state at time t uniquely depends on the present state $u(t)$, and that the law which links the present state to the present change is given by a (linear or nonlinear) map $F : \mathrm{D}(F) \to X$, where $\mathrm{D}(F) \subseteq X$.

Under these assumptions, the evolution of the system is determined by the *evolution equation* (more precisely, the initial value problem)

$$u'(t) = F(u(t)), \quad t \geq 0, \qquad u(0) = u_0. \tag{1.1}$$

We abbreviate $u' = \frac{\mathrm{d}}{\mathrm{dt}} u$ for the time derivative of the unknown function $u : \mathbb{R}_+ \to X$. Observe that here time evolves only into one direction, i.e., that $t \geq 0$. However, in many cases it makes sense to allow for $t \in \mathbb{R}$. Moreover, we have restricted to *autonomous* (of *time-invariant*) systems, when assuming that the map F only depends on the state and not on the time $t \geq 0$.

For a given natural phenomenon, it is a priori not clear which state space X and map F describe the phenomenon appropriately. One cannot determine such X and F within mathematics alone; this task of modeling clearly requires the expertise of the relevant science. In this course, we will not be concerned with this issue though we will discuss the physical background of certain equations.

Probably, you have already encountered systems of type (1.1) given by ordinary differential equations. These arise if the state of a system can be described by finitely many numbers so that one can choose $X = \mathbb{C}^n$ for some $n \in \mathbb{N}$. This type of equations covers for instance Newton's theory of classical mechanics. Below, we allow for Banach spaces X of infinite dimension in which one can formulate and work with partial differential equations.

In our course we first investigate the case of linear maps F. By means of so–called *strongly continuous semigroups*, one can establish a comprehensive and elegant solution theory for (1.1) in the *linear case*. Even more importantly, this theory provides the framework for the systematic investigation of the qualitative properties of the solutions to (1.1). In the second half of the course we then treat nonlinear equations. Here one does not have such a unifying and powerful general theory as in the linear case. Instead, we focus on certain "semilinear" equations which we can study starting from the linear theory, and which still give a flavor of the new challenges arising in nonlinear problems.

Within the vast realm of evolution equations, we want to focus on a specific class called *dispersive equations*. The most prominent examples in this class are the *Schrödinger equation* and the *wave equation* which will be the main applications treated in our lectures. Before continuing with an introductory discussion of the general problem (1.1), we describe these equations and explain how they fit into the general framework above.

The linear Schrödinger equation is given by

$$\begin{aligned} \mathrm{i}\frac{\partial}{\partial t}v(t,x) &= (-\Delta + V(x))v(t,x), \quad && x \in \mathbb{R}^d, \quad t \in \mathbb{R}, \\ v(0,x) &= v_0(x), && x \in \mathbb{R}^d, \end{aligned} \tag{1.2}$$

where $v : \mathbb{R} \times \mathbb{R}^d \to \mathbb{C}$ is the unknown. For the dimension $d = 3$ it describes the evolution of the probability wave function of a single particle in a quantum mechanical system. Here, $\Delta = \partial_{11} + \ldots + \partial_{dd}$ is the Laplace operator, $V : \mathbb{R}^d \to \mathbb{R}$ is a given *potential* which is determined by the physics of the system, and $\mathrm{i} \in \mathbb{C}$ is the imaginary unit. One has $V(x) = -\frac{c}{|x|}$ with a constant $c \in \mathbb{R}$ for the electron in the hydrogen atom, for instance. To cast (1.2) into the form (1.1), one chooses $X := L^2(\mathbb{R}^3)$ for the state space, requires that $u(t) := v(t, \cdot) \in X$ for each $t \in \mathbb{R}$ and sets $F(u) := \mathrm{i}(\Delta - V)u$. Depending on V, the domain $\mathrm{D}(F)$ of F could be a suitable Sobolev space. (These function spaces are explained later.)

The (damped) linear wave equation with Dirichlet boundary conditions on a bounded domain $U \subseteq \mathbb{R}^d$ is given by

$$\begin{aligned} \frac{\partial^2}{\partial t^2}w(t,x) &= \Delta w(t,x) - b(x)\frac{\partial}{\partial t}w(t,x), \quad && x \in U, \quad t \in \mathbb{R}, \\ w(t,x) &= 0, && x \in \partial U, \quad t \in \mathbb{R}, \\ w(0,x) &= w_0(x), \quad \frac{\partial}{\partial t}w(0,x) = w_1(x), && x \in U, \end{aligned} \tag{1.3}$$

where $b : U \to \mathbb{R}_+$ is the damping and $w : \mathbb{R} \times U \to \mathbb{C}$ is the unknown. As for ordinary differential equations, one can rewrite this equation for w (which is second order in time) into a system of two equations of first order in time for the unknowns $(w, \frac{\partial}{\partial t}w)$. One chooses $X := \mathring{H}^1(U) \times L^2(U)$ as the state space, where $\mathring{H}^1(U)$ is the first order Sobolev space with Dirichlet boundary conditions, and one writes

$$u(t) := (u_1(t), u_2(t)) := (w(t, \cdot), \tfrac{\partial}{\partial t}w(t, \cdot)) \in X$$

as above. One then obtains (1.1) by setting

$$F\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} := \begin{pmatrix} 0 & I \\ \Delta_D & -b \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

for the Dirichlet Laplacian Δ_D on $L^2(U)$. We remark that the Dirichlet boundary condition $w(t,\cdot)=0$ on ∂U here is included in the state space X. In many other situations boundary conditions are encoded in the domain of the map F.

The problems (1.2) and (1.3) are linear. In our lectures about nonlinear equations we will concentrate on the semilinear wave and Schrödinger equations. More precisely, we investigate the wave equation

$$\begin{aligned} &\frac{\partial^2}{\partial t^2}w(t,x) = \Delta w(t,x) - aw(t,x)|w(t,x)|^2, && x\in U, \quad t\in\mathbb{R},\\ &w(t,x)=0, && x\in\partial U, \quad t\in\mathbb{R},\\ &w(0,x)=w_0(x), \quad \frac{\partial}{\partial t}w(0,x)=w_1(x), && x\in U, \end{aligned} \tag{1.4}$$

with a cubic forcing term $-aw|w|^2$, where $a\in\mathbb{R}$ and $d=3$. (One can view the nonlinearity as a truncated power series of a general nonlinear force term.) The last part of the course is devoted to the nonlinear Schrödinger equation

$$\begin{aligned} &\mathrm{i}\frac{\partial}{\partial t}v(t,x) = -\Delta v(t,x) + \mu|v(t,x)|^{\alpha-1}v(t,x), && x\in\mathbb{R}^d, \quad t\in\mathbb{R},\\ &v(0,x)=v_0(x), && x\in\mathbb{R}^d, \end{aligned} \tag{1.5}$$

where $\mu\in\{-1,1\}$ and $\alpha>1$.[1] This equation (and its variants) appear in quantum field theory, e.g., in the study of so-called Bose–Einstein condensates. A discretized version models phenomena in DNA. It is also used to describe (approximately) the amplitudes of wave packages in nonlinear materials. The nonlinear problems (1.4) and (1.5) can be formulated as the evolution equation (1.1) in similar way as their linear counterparts (1.2) and (1.3).

We continue with the discussion of the general problem (1.1). To some extent our basic strategy is guided by the theory of ordinary differential equations, but there will be many fundamental differences when implementing this strategy. *Wellposedness* is the first fundamental question one has to answer:

> Given an initial state $u_0\in X$, is there a unique solution $u = u(\cdot\,;u_0)$ of (1.1) that continuously depends on u_0?

In the present context, by a solution u of (1.1) we understand a function $u\in C^1([0,\infty),X)$ such that $u(t)\in \mathrm{D}(F)$ for each $t\geq 0$ and (1.1) is fulfilled. Of course, such a solution can only exist for $u_0\in\mathrm{D}(F)$.[2] To allow for "many" initial data, we assume that $\mathrm{D}(F)$ is dense in X. The continuous dependence on the initial state means that the map $u_0\mapsto u(t;u_0)$ is continuous in the norm of X for all $t\geq 0$. Actually, we will further require that this map is uniformly continuous for $t\in[0,t_0]$ and every $t_0>0$.

As for ordinary differential equations, in the linear case we will indeed obtain solutions that exist for all times $t>0$. However, for nonlinear F in general one can only expect a solution on a possibly bounded time interval $[0,t^+(u_0))$, whose length $t^+(u_0)\in(0,\infty]$ (the *maximal existence time*) may depend on the initial state u_0. Solutions with a finite existence time (having *blow-up*) already

---

[1] For the experts: We will stay in the "subcritical range" and assume that $\alpha<\frac{d+2}{(d-2)_+}$.

[2] Often one can weaken the solution concept to obtain solutions in a generalized sense for every $u_0\in X$. We will not study this point systematically.

occur for the simple scalar equation $u' = u^2$. The reader is invited to compute explicitly the solution of this equation for any initial value $u_0 > 0$ and to derive $t^+(u_0) = \frac{1}{u_0} < \infty$. As we shall see, blow-up also occurs for (1.4) if $a < 0$ and for (1.5) if $\mu = -1$ (the so-called "focusing" case). One thus has to distinguish between *local* and *global* wellposedness. For convenience, we still write "$t \geq 0$" or "$t \in \mathbb{R}$" as the range for the time when posing the problem (1.1) although some solutions may not exist for all times.

Once we have solutions, we can (and should) study their qualitative behavior with respect to a large variety of possible properties. Still in the context of local wellposedness one can investigate the regularity of solutions. We will instead focus on the long-term behavior of solutions especially in the case of nonlinear F, where we first ask:

> Do we have a *global solution* of (1.1) for each (or some) $u_0 \in \mathrm{D}(F)$; i.e., do we have $t^+(u_0) = \infty$?

The answer will depend heavily on the concrete structure of F. For instance, the solutions of $u' = -u^2$ with $u(0) > 0$ exist for all $t \geq 0$. If a solution u exists for all times $t > 0$, we can then inquire:

> What is the behavior of $u(t)$ as $t \to \infty$?

Here a basic scenario is the *convergence to an equilibrium*, i.e., $u(t) \to u^*$ as $t \to \infty$ and $u^* \in \mathrm{D}(F)$ satisfies $F(u^*) = 0$.

So far, we did not discuss the concept of an *operator semigroup* from the title of the Internet Seminar. To explain its central role for evolution equations, assume that (1.1) is a linear problem, i.e., it is given by

$$u'(t) = Au(t), \quad t \geq 0, \qquad u(0) = u_0, \tag{1.6}$$

where $A : \mathrm{D}(A) \to X$ is a linear operator and its domain $\mathrm{D}(A)$ is a dense linear subspace of X. Here the reader may think of an appropriate realization of the Laplacian Δ.

We first assume that for each $u_0 \in \mathrm{D}(A)$ we have a unique solution $u = u(\,\cdot\,; u_0)$ of (1.6), i.e., u belongs to $C^1(\mathbb{R}_+, X)$, $u(t) \in \mathrm{D}(A)$ for all $t \geq 0$ and (1.6) holds. Thanks to uniqueness, we can then define the map

$$T(t) : \mathrm{D}(A) \to \mathrm{D}(A), \quad T(t)u_0 = u(t; u_0),$$

for $t \geq 0$. Of course, $T(0)$ is just the identity on $\mathrm{D}(A)$. Note that $t \mapsto T(t)u_0$ is continuous from $\mathbb{R}_+$ to X by our assumption. If $u_0, v_0 \in \mathrm{D}(A)$ and $\alpha, \beta \in \mathbb{C}$, then the function $\alpha u(\,\cdot\,; u_0) + \beta u(\,\cdot\,; v_0)$ solves (1.6) for the initial value $\alpha u_0 + \beta v_0$ since A is linear. Again by uniqueness, it follows $T(t)(\alpha u_0 + \beta v_0) = \alpha T(t)u_0 + \beta T(t)v_0$ so that $T(t)$ is a *linear* map on $\mathrm{D}(A)$ for each fixed $t \geq 0$.

Let us now further assume that $u(t; u_0)$ depends continuously (with respect to the norm of X) on u_0, locally uniformly for $t \geq 0$. By a standard density argument one can then extend $T(t)$ to a *bounded* linear operator on X such that the *orbit map* $t \mapsto T(t)x$ is continuous from $\mathbb{R}_+$ to X for each $x \in X$. We denote the resulting one parameter family of bounded linear operators on X by $(T(t))_{t \geq 0}$ or simply $T(\cdot)$.

Finally, take $t, s \geq 0$ and restart the problem (1.6) with initial value $u(s)$. We then have the two solutions $t \mapsto u(t; u(s))$ and $t \mapsto u(t + s; u_0)$ of (1.6) for

$t \geq 0$, which have to coincide by uniqueness. Thus $T(\cdot)$ satisfies the *semigroup law*

$$T(t+s) = T(t)T(s) \quad \text{and} \quad T(0) = I.$$

We call such operator families $T(\cdot)$ *strongly continuous operator semigroups* or shortly C_0*–semigroups*. Moreover, $T(\cdot)$ is sometimes called the *solution semigroup* for (1.6). If we can reverse time and the above properties hold for all $t, s \in \mathbb{R}$, then we have a C_0*–group*. As it turns out, this will indeed be the case for the linear Schrödinger equation (1.2) and for the linear wave equation (1.3) (at least if b is bounded).

As we have seen above, the concept of a C_0–semigroup naturally arises from the basic assumptions

(1) existence of unique solutions,
(2) continuous dependence on the initial data,
(3) linearity

for the solutions of (1.6). Therefore, if one requires these properties, semigroups come automatically into play. We stress that the points (1) and (2) above are fundamental for any scientific investigation of systems evolving in time. It must be possible to test a statement in science by an experiment, at least in principle. For an evolutionary problem this requirement forces one to describe the states in a unique way and to give a definite prediction about the future. We point out that one can be very flexible in choosing the states of a system. For the Schrödinger equation (1.2) they are given by a probability wave function in $L^2(\mathbb{R}^3)$ as the state, for instance. This example indicates how a probabilistic viewpoint can also be included in the above deterministic approach by choosing an appropriate state space.

Moreover the initial states will only be known approximately in practice so that predictions must be made with a certain error tolerance, which requires the continuous dependence on the initial values.

The linearity of each map $T(t)$ is a consequence of the linearity of A. Also, in the general case (1.1) when F is nonlinear, one can define a solution semigroup (here usually called *semiflow*) if the solutions satisfy (1) and (2). However, in this case $T(t)$ will not be linear, and one always has to keep in mind that solutions may exists only on finite time intervals depending on the respective initial value.

There is an astonishingly rich and elegant theory on C_0–semigroups for linear evolution equations. Its foundation is the theorem of Hille-Yosida. This result characterizes those operators A which "generate" a C_0–semigroup that solves (1.6) for A. We will also show that the linear problem (1.6) has the two fundamental properties (1) and (2) for a densely defined linear operator A *if and only if* A generates a C_0–semigroup, which then solves (1.6). In this sense, semigroup theory provides the *natural framework* for linear evolution equations and the Hille-Yosida theorem closes the circle between the given linear operator A, the evolution equation (1.6) and the corresponding solution semigroup $T(\cdot)$. Even more importantly, this framework serves as a suitable setting for our further investigations of the long-term behavior of (1.1) as $t \to \infty$.

In our applications the norm on X corresponds to the physical energy of the system (modulo constants). The energy does not increase in time in (1.2) and (1.3), as well as in many other systems in sciences; i.e., it holds $\|T(t)x\| \le \|x\|$ for each $t \ge 0$ and each $x \in X$. As a result, the corresponding operators $T(t)$ are *contractions* for each $t \ge 0$. For contractions the Hille-Yosida generation result has a particular simple form (the Lumer-Phillips theorem) whose conditions are relatively easy to check in examples. On a Hilbert space X, the Lumer-Phillips theorem further implies Stone's theorem about the generation of unitary C_0–groups (which is needed for the Schrödinger equation (1.2)). In view of our applications, we will focus on the contraction case and on Hilbert spaces at least when more general results require significantly more effort.

Actually, one could establish the generation results rather quickly. But we think that we should take a slower start and present the basic arguments in detail and provide several illustrating simple examples. We hope that this helps the beginner to become familiar with (at first linear) analysis in Banach spaces and the functional analytic approach to evolution equations.

In the framework for linear evolution equations provided by the Hille-Yosida theorem we treat three more topics:

(a) perturbation theory,
(b) exponential stability,
(c) inhomogeneous problems.

The equations (1.2) and (1.3) are typically solved by perturbation arguments. The Laplacian in an L^2–setting can be treated by functional analytic methods (Fourier transform, Lemma of Lax-Milgram) in a very elegant way. The potential and the damping terms in (1.2) and (1.3) can then often be handled as "lower order" perturbations which do not affect the existence of a solution semigroup for these evolution equations.

A C_0–semigroup $T(\cdot)$ is called *exponentially stable* if $\|T(t)\| \to 0$ exponentially as $t \to \infty$. As we will see, on a Hilbert space X a C_0–semigroup is exponentially stable if (and only if) the resolvent $(\lambda I - A)^{-1}$ of its generator exists and is bounded for $\operatorname{Re}\lambda > 0$. This powerful theorem (basically due to Gearhart) links the spectral properties of A to the dynamic properties of the solutions of (1.6). It is a generalization of a corresponding result for linear ordinary differential equations due to Lyapunov, which the reader might know from undergraduate courses.

If a given forcing or control function $g : \mathbb{R}_+ \to X$ is present in a linear system, instead of equation (1.6), one has to investigate the *inhomogeneous problem*

$$u'(t) = Au(t) + g(t), \quad t \ge 0, \qquad u(0) = u_0. \tag{1.7}$$

It can be solved by *Duhamel's formula* (or the *variation of constants formula*)

$$u(t) = T(t)u_0 + \int_0^t T(t-s)g(s)\,\mathrm{d}s, \qquad t \ge 0, \tag{1.8}$$

where $T(\cdot)$ is the solution semigroup to the *homogeneous problem* (1.6). This formula builds the bridge to the *semilinear problem*

$$u'(t) = Au(t) + f(u(t)), \quad t \ge 0, \qquad u(0) = u_0, \tag{1.9}$$

for a nonlinearity $f : X \to X$ which is Lipschitz continuous on bounded sets. Indeed, by means of (1.8) the equation (1.9) can be reformulated as the integrated fixed point problem

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(u(s))\,\mathrm{d}s. \tag{1.10}$$

Here the reader may recall the proof of the Picard-Lindelöf theorem for ordinary differential equations. This setting already covers the nonlinear wave equation (1.4) for the space dimension $d = 3$, as we will see later. But even more importantly, for problem (1.9) one can establish a prototypical theory for local wellposedness and for the asymptotic behavior, which then serves as a guideline for much more complicated nonlinear systems when f is not defined on the whole space X. As for the linear problems, we want to spell out the basic methods, ideas and results for the nonlinear analysis first on the level of a more accessible problem, namely (1.9).

The nonlinear Schrödinger equation (1.5) is a much more complicated system. Here the nonlinearity $f(u) = -\mathrm{i}\mu|u|^{\alpha-1}u$ does not map $X = L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^d)$ (since $\alpha > 1$), and thus one has to exploit the specific properties of the operator $\mathrm{i}\Delta$ and the nonlinearity to obtain wellposedness.

Here, finally, the dispersive character of Schrödinger and wave type problems really comes into play. Physically, dispersion means that waves of different frequencies propagate with different velocities. In our context therefore wave packages smear out as time evolves, thereby improving their integrability in some sense. We will discuss this concept in much greater detail later in the corresponding lectures. Here we just note that mathematically this property is most easily seen for the (free) linear Schrödinger equation

$$\frac{\partial}{\partial t}v(t,x) = \mathrm{i}\Delta v(t,x), \quad x \in \mathbb{R}^d, \ \ t \in \mathbb{R}, \qquad v(0,x) = v_0(x), \quad x \in \mathbb{R}^d.$$

This problem is solved by a C_0–group $T(\cdot)$ on $L^2(\mathbb{R}^d)$ which has the explicit representation formula

$$(T(t)v_0)(x) = \frac{1}{(4\pi\mathrm{i}t)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\frac{|x-y|^2}{4t}} v_0(y)\,\mathrm{d}y,$$

for $v_0 \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ and $t \neq 0$. The formula directly implies that

$$\|T(t)v_0\|_{L^\infty} \le \frac{1}{(4\pi|t|)^{d/2}}\|v_0\|_{L^1}. \tag{1.11}$$

Hence the states are bounded away from the initial time and tend uniformly to zero as $|t| \to \infty$. On the other hand, one can show that $\|T(t)v_0\|_{L^2} = \|v_0\|_{L^2}$ holds for $v_0 \in L^2(\mathbb{R}^d)$ and $t \in \mathbb{R}$. These two properties of $T(\cdot)$ imply that the solution of any localized initial state smears out as time evolves and its "mass" will be pressed out towards infinity.

Using methods from functional analysis, the simple estimate (1.11) can be upgraded to the famous Strichartz' estimates, saying that

$$\|T(\cdot)v_0\|_{L^q(\mathbb{R},L^p(\mathbb{R}^d))} \le c\|v_0\|_{L^2(\mathbb{R}^d)},$$
$$\|T(\cdot) *_+ f\|_{L^q(\mathbb{R},L^p(\mathbb{R}^d))} \le c\|f\|_{L^{q'}(\mathbb{R},L^{p'}(\mathbb{R}^d))}$$

for a constant $c > 0$ independent of $v_0 \in L^2(\mathbb{R}^d)$ and $f \in L^{q'}(\mathbb{R}, L^{p'}(\mathbb{R}^d))$. Here $(T(\cdot) *_+ f)(t) := \int_0^t T(t-s)f(s)\,\mathrm{d}s$ is the half-line convolution, the exponents $q, p \geq 2$ satisfy a certain relation dictated by scaling which involves the underlying dimension d and q', p' are the conjugate (dual) exponents.

These estimates will be crucial for the study of the nonlinear Schrödinger equation (1.5) which can be written as the fixed point problem

$$u(t) = T(t)u_0 - \mathrm{i}\mu \int_0^t T(t-s)u(s)|u(s)|^{\alpha-1}\,\mathrm{d}s, \tag{1.12}$$

cf. (1.10). Observe that Strichartz's estimates deal with precisely the two terms on the right-hand side. Based on the function spaces arising in Strichartz' estimates, we will find a fixed point of the equation (1.12) and in this way establish the local wellposedness theory for (1.5). In certain cases we will also show global wellposedness of (1.5) based on the fact that the "energy" of solutions of (1.5) is constant in time. We also study the long-term behavior if $\mu = 1$.

This course is meant as a starting point for a deeper investigation of dispersive or wave type equations. Based on our course, we intend to treat various more advanced topics (or other subjects for which we will not have enough time in the lectures) in the projects of Phase 2 of the Internet Seminar.

Prerequisites. Concluding this introduction we want to explain the prerequisites needed for this course and the role of the appendices and the exercises.

We assume some familiarity with functional analysis. Besides basic facts and the standard function spaces, e.g. L^p spaces, we will freely use the principle of uniform boundedness, the open mapping theorem, the Hahn-Banach theorem, the theorem of Banach-Alaoglu, Riesz' representation theorem for Hilbert space duals and their standard consequences. For this material we refer to, e.g., [**Bre11**], [**Con90**], [**RS72**], [**Rud91**], [**TL80**], [**Wer07**], [**Yos80**]. Several other topics from functional analysis, namely closed operators, basic spectral theory, selfadjoint operators, Sobolev spaces, the Fourier transform and the Bochner integral, will be treated in the appendices. In the lectures we introduce briefly the necessary concepts, state the relevant results and discuss them a bit. It should be possible to follow the course without reading the appendices which contain the proofs and include more background material. Of course many of you already know (part of) this extra material. In the text we will cite from time to time additional results from the literature. These citations are made just for your information and are not needed later on. The exercises are mostly meant to provide additional examples and give you an opportunity to test your understanding. Some exercises provide variants of the results treated in the lectures. Very rarely, we use simple facts from the exercises later on.

We gratefully thank the participants of the Internet Seminar whose comments helped to correct mistakes and to improve the presentation.

LECTURE 2

Strongly continuous semigroups and their generators

In the first part of the Internet Seminar we investigate the linear evolution equation (or more precisely, the initial value or Cauchy problem)

$$u'(t) = Au(t), \quad t \geq 0, \qquad u(0) = u_0, \tag{2.1}$$

on a Banach space X. Here A is a given linear operator with domain $\mathrm{D}(A)$ and $u_0 \in X$ (often $u_0 \in \mathrm{D}(A)$) is the initial value. We want to develop a systematic theory for (2.1) which will also be the basis for our study of semilinear problems. As explained in Lecture 1, the solution u of (2.1) will be given by $u(t) = T(t)u_0$ for an operator semigroup $(T(t))_{t\geq 0} = T(\cdot)$ on X. Our analysis starts with the study of such semigroups. For each semigroup we introduce its generator A as the derivative of $t \mapsto T(t)$ at $t = 0$, roughly speaking. We next investigate some of their basic properties, solve problem (2.1) for the generator A, and study a simple but instructive class of examples, the translation semigroups.

In the following lectures, we will tackle the more difficult problem to characterize those (given) operators A which are generators of a strongly continuous semigroup.

A few words about our **notation**: Throughout we assume that $(X, \|\cdot\|_X)$ is a complex Banach space with $X \neq \{0\}$, where we mostly write $\|\cdot\|$ instead of $\|\cdot\|_X$, if no confusion is to be expected. By $\mathcal{B}(X, Y)$ we denote the space of all bounded linear operators from X into another Banach space Y, setting $\mathcal{B}(X) = \mathcal{B}(X, X)$. Mostly, the operator norm is also designated by $\|\cdot\|$. Further, X^* is the dual space of X and I is the identity map on X. We write $\langle x, x^*\rangle := x^*(x)$ for $x^* \in X^*$ and $x \in X$. The scalar product on a Hilbert space H is denoted by $(x|y)_H$, or simply $(x|y)$, for $x, y \in H$. By $\mathbb{1}_M$ we designate the characteristic function of a set M. We put $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}_- = (-\infty, 0]$.

For $M \subseteq \mathbb{R}^d$ we write $C(M, X)$ for the vector space of continuous functions $f : M \to X$. We use the subspaces

$$C_b(M, X) = \{f \in C(M, X) \,|\, f \text{ is bounded}\},$$
$$C_0(M, X) = \{f \in C(M, X) \,|\, f(s) \to 0 \text{ as } |s| \to \infty \text{ or as } s \to \partial M \setminus M\},$$
$$C_c(M, X) = \{f \in C(M, X) \,|\, \operatorname{supp} f \subseteq M \text{ is compact}\},$$

where $C_b(M) := C_b(M, \mathbb{C})$ etc. Here $\operatorname{supp} f$ is the closure in $\mathbb{R}^d$ of the set $\{s \in M \,|\, f(s) \neq 0\}$. Below we will repeat these definitions in specific cases. The spaces $C_b(M, X)$ and $C_0(M, X)$ are always equipped the sup–norm $\|f\|_\infty := \sup_{s\in M} \|f(s)\|$ and then become Banach spaces.

We employ analogous notations for spaces of differentiable functions. For instance, $C_0^1([0, 1)) = \{f \in C^1([0, 1)) \,|\, f, f' \in C_0([0, 1))\}$.

If M is a Borel subset of $\mathbb{R}^d$ and $p \in [1,\infty]$, we write $L^p(M)$ for the usual Lebesgue space of complex valued functions with respect to the Lebesgue measure on M and endow it with the p–norm, given by $\|f\|_p^p := \int_M |f|^p \,\mathrm{d}x$ if $p < \infty$ and by the essential supremum for $p = \infty$. If $M = (a,b) \subseteq \mathbb{R}$, we write $C_0(a,b)$, $L^p(a,b)$ and so on.

Definition 2.1. *A map* $T(\cdot) : \mathbb{R}_+ \to \mathcal{B}(X)$ *is called* strongly continuous operator semigroup *or just* C_0–semigroup *if the following conditions are fulfilled:*

(a) $T(0) = I$ *and we have* $T(t+s) = T(t)T(s)$ *for all* $t, s \geq 0$.

(b) For each $x \in X$ *the* orbit, *defined as the map*

$$T(\cdot)x : \mathbb{R}_+ \to X, \quad t \mapsto T(t)x,$$

is continuous.

The generator A of $T(\cdot)$ *is given by setting*

$$\mathrm{D}(A) := \{x \in X \mid \text{ the limit } \lim_{t\to 0^+} \tfrac{1}{t}(T(t)x - x) \text{ exists in } X\}$$

and defining

$$Ax := \lim_{t\to 0^+} \tfrac{1}{t}(T(t)x - x)$$

for $x \in \mathrm{D}(A)$. *We also say that* A generates $T(\cdot)$.

If one replaces throughout in this definition $\mathbb{R}_+$ *by* $\mathbb{R}$ *and "*$t \to 0^+$*" by "*$t \to 0$*", one obtains the concept of a* C_0–group *with* generator A.

Property (a) in Definition 2.1 is called the *semigroup law* and (b) is the *strong continuity*. We point out that the semigoup does not need to be continuous with respect to the operator norm, cf. Example 2.6. Observe that the above definition directly implies that $\mathrm{D}(A)$ is a linear subspace of X and A is a linear map in X which is uniquely determined by the semigroup.

Let $T(\cdot)$ be a C_0–semigroup. The semigroup operators then commute since

$$T(t)T(s) = T(t+s) = T(s+t) = T(s)T(t) \tag{2.2}$$

holds for all $t, s \geq 0$. By induction, we obtain

$$T(nt) = T\Big(\sum_{j=1}^{n} t\Big) = \prod_{j=1}^{n} T(t) = T(t)^n \tag{2.3}$$

for all $n \in \mathbb{N}$ and $t \geq 0$. If $T(\cdot)$ is even a C_0–group, it satisfies

$$T(t)T(-t) = T(0) = I = T(-t)T(t) \tag{2.4}$$

for all $t \in \mathbb{R}$. Hence, $T(t)$ is invertible with inverse $T(-t)$ for every $t \in \mathbb{R}$.

We first look at the finite dimensional situation, which is rather simple since here one can construct the semigroup from the given operator A by a power series. This does not work for unbounded A.

Example 2.2. Let $X = \mathbb{C}^d$, $A \in \mathcal{B}(X) \cong \mathbb{C}^{d\times d}$ and set

$$T(t) := \mathrm{e}^{tA} := \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n$$

for $t \in \mathbb{R}$. It is known from analysis courses that the series converges in $\mathcal{B}(X)$, $T(\cdot)$ satisfies (a) in Definition 2.1 and $T(\cdot)$ is continuously differentiable even in

$\mathcal{B}(X)$ with $\frac{\mathrm{d}}{\mathrm{dt}} \mathrm{e}^{tA} = A\mathrm{e}^{tA}$ for all $t \in \mathbb{R}$. In particular, $T(\cdot)$ is a C_0–group with generator A. Moreover, for any given $u_0 \in X$ the function $u : \mathbb{R}_+ \to X$ defined by $u(t) = \mathrm{e}^{tA}u_0$ solves the linear ordinary differential equation

$$u'(t) = Au(t), \quad t \in \mathbb{R}, \qquad u(0) = u_0.$$

The same results hold for any bounded linear operator A on a Banach space X, see Exercise 2.1. ♢

The simple Definition 2.1 has many astonishing consequences. We first observe that every C_0–semigroup is exponentially bounded. This fact then leads to the subsequent basic definition.

LEMMA 2.3. *Let $T(\cdot) : \mathbb{R}_+ \to \mathcal{B}(X)$ satisfy condition (a) in Definition 2.1 as well as $\limsup_{t\to 0^+} \|T(t)x\| < \infty$ for each $x \in X$. Then there are constants $M \geq 1$ and $\omega \geq 0$ such that $\|T(t)\| \leq M\mathrm{e}^{\omega t}$ for all $t \geq 0$. This fact holds in particular if $T(\cdot)$ is a C_0–semigroup.*

PROOF. Suppose that $\|T(t_n)\| \to \infty$ as $n \to \infty$ for a null sequence $(t_n)_n$ in $\mathbb{R}_+$. The principle of uniform boundedness then gives a vector $x \in X$ satisfying $\lim_{n\to\infty} \|T(t_n)x\| = \infty$, contradicting the boundedness assumption. There thus exist constants $M \geq 1$ and $t_0 > 0$ such that $\|T(\tau)\| \leq M$ for all $\tau \in [0, t_0]$. Let $t \geq 0$. Take $n \in \mathbb{N}_0$ and $\tau \in [0, t_0)$ with $t = \tau + nt_0$. Setting $\omega = \frac{\log M}{t_0} \geq 0$, we deduce from Definition 2.1 (a) and (2.3) that

$$\|T(t)\| = \|T(\tau)T(nt_0)\| = \|T(\tau)T(t_0)^n\| \leq M \cdot M^n = M\mathrm{e}^{nt_0\omega} \leq M\mathrm{e}^{\omega t}.$$

If $T(\cdot)$ is a C_0–semigroup, each orbit is bounded on $[0, 1]$ by the strong continuity. So the last assertion holds. □

DEFINITION 2.4. *Let $T(\cdot)$ be a C_0–semigroup with generator A. Then*

$$\omega_0(T) := \omega_0(A) := \inf \Big\{\omega \in \mathbb{R} \,\big|\, \exists M_\omega \geq 1 : \|T(t)\| \leq M_\omega \mathrm{e}^{\omega t} \text{ for all } t \geq 0\Big\}$$

is called the growth bound *of $T(\cdot)$.*

Lemma 2.3 says that $\omega_0(T) < \infty$. It may happen that $\omega_0(T) = -\infty$, see Example 2.6. In general the infimum in Definition 2.4 is not a minimum even in the matrix case. For instance, for $X = \mathbb{C}^2$ (endowed with the 1–norm) and $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, we have $T(t) = \mathrm{e}^{tA} = \left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right)$, so that $\|T(t)\| = t + 1$ for $t \geq 0$, while $\omega_0(T) = 0$. The notation $\omega_0(A)$ will be justified in the next lecture where we show that an operator A can generate at most one C_0–semigroup.

The next lemma often helps to verify the strong continuity of an operator semigroup.

LEMMA 2.5. *Let $T(\cdot) : \mathbb{R}_+ \to \mathcal{B}(X)$ be a map satisfying condition (a) in Definition 2.1. Then the following assertions are equivalent.*

(a) $T(\cdot)$ is strongly continuous (and thus a C_0–semigroup).
(b) It holds $\lim_{t\to 0^+} T(t)x = x$ for all $x \in X$.
(c) There are a number $t_0 > 0$ and a dense subspace $D \subseteq X$ such that $\sup_{0\leq t\leq t_0} \|T(t)\| < \infty$ and $\lim_{t\to 0^+} T(t)x = x$ for all $x \in D$.

The analogous equivalences hold in the group case.

Proof. The implication "(a)⇒(c)" is a consequence of Lemma 2.3. Assertion (b) follows from (c) by a standard approximation argument.

To conclude (a) from (b), we fix $x \in X$ and $t > 0$. For $h > 0$ the semigroup property implies

$$\|T(t+h)x - T(t)x\| = \|T(t)(T(h)x - x)\| \le \|T(t)\| \cdot \|T(h)x - x\|,$$

where the right-hand side of this inequality converges to 0 as h tends to 0. In the case that $h \in (-t, 0]$, we note that Lemma 2.3 yields

$$\|T(t+h)\| \le M\mathrm{e}^{\omega(t+h)} \le M\mathrm{e}^{\omega t}$$

for some constants $M \ge 1$ and $\omega \ge 0$. As a result,

$$\|T(t+h)x - T(t)x\| = \|T(t+h)(x - T(-h)x)\| \le M\mathrm{e}^{\omega t} \cdot \|x - T(-h)x\| \to 0$$

as $h \to 0^-$, and (a) holds. The assertions about groups are shown similarly. □

In the above lemma the implication "(c)⇒(a)" can fail if one omits the boundedness assumption (cf. Exercise I.5.9(4) in [**EN99**]). We now examine a basic class of examples for C_0–semigroups, the translation semigroups. They are given by an explicit formula (which is a rare exception) and are thus very convenient to illustrate various aspects of the theory. Their generators will be determined in the next lecture.

Example 2.6 (Left translation semigroups on $\mathbb{R}$ and $[0,1)$).
(a) Let $X = C_0(\mathbb{R}) = \{f \in C(\mathbb{R}) \mid f(s) \to 0 \text{ as } |s| \to \infty\}$ and $T(\cdot)$ be given by

$$(T(t)f)(s) := f(s+t) \qquad \text{for } \ t \in \mathbb{R},\ f \in X,\ s \in \mathbb{R}.$$

We claim that $T(\cdot)$ is a C_0–group on X. Clearly, $T(0) = I$ and $T(t)$ is a linear isometry on X so that $\|T(t)\| = 1$. We further obtain

$$T(t)T(r)f = (T(r)f)(\,\cdot\, + t) = f(\,\cdot\, + t + r) = T(t+r)f$$

for all $f \in X$ and $r, t \in \mathbb{R}$. Hence, $T(t)T(r) = T(t+r)$. We employ Lemma 2.5 to verify the strong continuity. For $f \in C_c(\mathbb{R})$ the function $T(t)f$ converges uniformly to f as t tends to 0 since f is uniformly continuous. It thus remains to check $\overline{C_c(\mathbb{R})} = C_0(\mathbb{R})$. For each $n \in \mathbb{N}$ take a "cut–off" function $\varphi_n \in C(\mathbb{R})$ satisfying $\varphi_n = 1$ on $[-n, n]$, $0 \le \varphi_n \le 1$ and $\operatorname{supp} \varphi_n \subseteq (-n-1, n+1)$. For $f \in C_0(\mathbb{R})$, we then have $\varphi_n f \in C_c(\mathbb{R})$ and

$$\|f - \varphi_n f\|_\infty = \sup_{|s| \ge n} |(1 - \varphi_n(s))f(s)| \le \sup_{|s| \ge n} |f(s)| \ \longrightarrow\ 0$$

as $n \to \infty$. We now conclude that $T(\cdot)$ is a C_0–group by means of Lemma 2.5.

The same assertions hold for $X = L^p(\mathbb{R})$ with $1 \le p < \infty$ by similar arguments, see Exercise 2.2.

In contrast to these results, $T(\cdot)$ is not strongly continuous on $X = L^\infty(\mathbb{R})$. Indeed, consider $f = \mathbb{1}_{[0,1]}$ and observe that

$$T(t)f(s) = \mathbb{1}_{[0,1]}(s+t) = \left.\begin{cases} 1, & s+t \in [0,1] \\ 0, & s+t \notin [0,1] \end{cases}\right\} = \mathbb{1}_{[-t,1-t]}(s)$$

for $s, t \in \mathbb{R}$. Thus, $\|T(t)f - f\|_\infty = 1$ for every $t \neq 0$.

In addition, $T(\cdot)$ is not continuous as a $\mathcal{B}(X)$-valued function for X being $L^p(\mathbb{R})$ (see Exercise 2.2) or $C_0(\mathbb{R})$. In fact, for $X = C_0(\mathbb{R})$ consider for each $n \in \mathbb{N}$ functions $f_n \in C_c(\mathbb{R})$ with $0 \le f_n \le 1$, $f_n(n) = 1$ and $\operatorname{supp} f_n \subseteq (n-\frac{1}{n}, n+\frac{1}{n})$. We then have $\operatorname{supp} T(\frac{2}{n})f_n \subseteq (n-\frac{3}{n}, n-\frac{1}{n})$ for $n \in \mathbb{N}$, which implies

$$\|T(\tfrac{2}{n}) - I\| \ge \|T(\tfrac{2}{n})f_n - f_n\|_\infty = 1 \qquad \text{for all } n \in \mathbb{N}. \qquad \diamondsuit$$

(b) Let $X = C_0([0,1)) = \{f \in C([0,1)) \mid \lim_{s\to 1} f(s) = 0\}$. For $t \ge 0$ and $s \in [0,1)$, we define

$$(T(t)f)(s) := \begin{cases} f(s+t), & s+t < 1, \\ 0, & s+t \ge 1. \end{cases}$$

We show that $T(\cdot)$ is a C_0–semigroup on X. Since $f(s+t) \to 0$ as $s+t \to 1$, we have $T(t)f \in X$. Clearly, $T(t)$ is linear and $\|T(t)\| \le 1$. Note that $T(t) = 0$ whenever $t \ge 1$. In this case, one says that $T(\cdot)$ is *nilpotent*. As a consequence, $\omega_0(T) = -\infty$. Let $t, r \ge 0$ and $s \in [0,1)$. We then obtain

$$\begin{aligned} (T(t)T(r)f)(s) &= \begin{cases} (T(r)f)(s+t), & \text{if } s+t < 1, \\ 0, & \text{else}, \end{cases} \\ &= \begin{cases} f(s+t+r), & \text{if } s+t<1, \quad s+t+r<1, \\ 0, & \text{else}, \end{cases} \\ &= (T(t+r)f)(s). \end{aligned}$$

Hence, $T(\cdot)$ is a semigroup (which cannot be extended to a group since e.g. $T(1) = 0$ is not bijective). As in (a) one sees that $C_c([0,1)) = \{f \in C([0,1)) \mid \exists b_f \in (0,1) : \operatorname{supp} f \subseteq [0, b_f]\}$ is a dense subspace of X. For $f \in C_c([0,1))$ and $t \in (0, 1-b_f)$, we compute

$$T(t)f(s) - f(s) = \begin{cases} f(s+t) - f(s), & s \in [0, 1-t), \\ 0, & s \in [1-t, 1) \subseteq [b_f, 1], \end{cases}$$

and deduce $\lim_{t\to 0} \|T(t)f - f\|_\infty = 0$ using the uniform continuity of f. According to Lemma 2.5, $T(\cdot)$ is a C_0–semigroup on X. $\diamondsuit$

We next state the solution concept for equation (2.1).

Definition 2.7. *Let A be a linear operator on X with domain $\mathrm{D}(A)$ and let $x \in \mathrm{D}(A)$. We say a function $u : \mathbb{R}_+ \to X$* solves *the Cauchy problem*

$$u'(t) = Au(t), \quad t \ge 0, \qquad u(0) = x, \tag{2.5}$$

if $u \in C^1(\mathbb{R}_+, X)$ satisfies $u(t) \in \mathrm{D}(A)$ for all $t \ge 0$ and fulfills (2.5).

We next show that if A generates a C_0–semigroup, then the semigroup gives the unique solution of (2.5). Moreover, $T(t)$ and A commute on $\mathrm{D}(A)$.

Proposition 2.8. *Let A generate the C_0–semigroup $T(\cdot)$ and $x \in \mathrm{D}(A)$. Then $T(t)x \in \mathrm{D}(A)$, $AT(t)x = T(t)Ax$ for all $t \ge 0$ and the function*

$$u : \mathbb{R}_+ \to X, \quad t \mapsto T(t)x,$$

is the unique solution of (2.5).

Proof. 1) Let $t \geq 0$, $h > 0$ and $x \in \mathrm{D}(A)$. We obtain

$$\tfrac{1}{h}(T(t+h)x - T(t)x) = \tfrac{1}{h}(T(h) - I)T(t)x = T(t)\tfrac{1}{h}(T(h)x - x) \ \longrightarrow\ T(t)Ax$$

as $h \to 0$. The very definition of A yields $T(t)x \in \mathrm{D}(A)$ and $AT(t)x = T(t)Ax$. Moreover, $T(\cdot)x$ is differentiable from the right. Let $0 < h < t$. It further holds

$$\tfrac{1}{-h}(T(t-h)x - T(t)x) = T(t-h)\tfrac{1}{h}(T(h)x - x) \ \longrightarrow\ T(t)Ax$$

as $h \to 0$, where we have used Lemma 2.9 below (with $S(t,h) = T(t-h)$). Since $T(\cdot)Ax$ is continuous, we have shown that $T(\cdot)x \in C^1(\mathbb{R}_+, X)$ with derivative $\frac{\mathrm{d}}{\mathrm{d}t} T(\cdot)x = AT(\cdot)x$; i.e., u solves (2.5).

2) Let v be another solution of (2.5) and $t > 0$. Set $w(s) := T(t-s)v(s)$ for $s \in [0,t]$. Lemma 2.9 (with $S(t,s) = T(t-s)$ and $Y = \mathrm{D}(A)$) and the first step now imply that

$$\tfrac{\mathrm{d}}{\mathrm{d}s} w(s) = T(t-s)v'(s) - T(t-s)Av(s) = 0,$$

where the last equality follows from the assumption that v solves (2.5). So for every $x^* \in X^*$ the scalar function $\langle w(\cdot), x^*\rangle$ is differentiable with vanishing derivative and thus constant, which leads to

$$\langle T(t)x, x^*\rangle = \langle w(0), x^*\rangle = \langle w(t), x^*\rangle = \langle v(t), x^*\rangle$$

for all $t \geq 0$ and $x^* \in X^*$. The Hahn-Banach theorem now yields $T(\cdot)x = v$. $\square$

In general one really needs the extra condition that $x \in \mathrm{D}(A)$ to obtain a solution of (2.5). For instance, if $f \in C_0(\mathbb{R}) \setminus C^1(\mathbb{R})$, then the orbit $T(\cdot)f$ of the translation semigroup on $C_0(\mathbb{R})$ is not differentiable, cf. Example 2.6. We continue with the lemma used in the previous proof.

Lemma 2.9. *Let $b > a$ be real numbers, $M = \{(t,s) \in [a,b]^2 \,|\, t \geq s\}$, $S : M \to \mathcal{B}(X)$ be strongly continuous and $f \in C([a,b], X)$. Then the function*

$$g : M \to X, \quad (t,s) \mapsto S(t,s)f(s),$$

is also continuous.

Further, let $Y \subseteq X$ be a subspace and let the map $[a,t] \to X$, $s \mapsto S(t,s)y$, have the derivative $\partial_s S(t,s)y$ for each $t \in (a,b]$ and $y \in Y$. Let $f \in C^1([a,b], X)$ take values in Y. Then the map $[a,t] \ni s \mapsto g(t,s)$ is differentiable in X with

$$\partial_s g(t,s) = S(t,s)f'(s) + \partial_s S(t,s)f(s).$$

Proof. Observe that $\sup_{(t,s)\in M} \|S(t,s)x\| < \infty$ for every $x \in X$ by continuity. The number $c := \sup_{(t,s)\in M} \|S(t,s)\|$ is finite by the uniform boundedness principle. For $(t,s), (t',s') \in M$ we thus obtain

$$\|S(t',s')f(s') - S(t,s)f(s)\| \leq c\,\|f(s') - f(s)\| + \|(S(t',s') - S(t,s))f(s)\|,$$

where the right-hand side of this inequality tends to 0 as $(t',s') \to (t,s)$.

To show the second assertion, fix $t \in (a,b]$ and take $s, s+h \in [a,t]$ for $h \in \mathbb{R} \setminus \{0\}$. We compute

$$\begin{aligned}&\tfrac{1}{h}(S(t,s+h)f(s+h) - S(t,s)f(s))\\ &\qquad = S(t,s+h)\tfrac{1}{h}(f(s+h) - f(s)) + \tfrac{1}{h}(S(t,s+h) - S(t,s))f(s).\end{aligned}$$

As $h \to 0$, the second claim follows from the first part and our assumptions. $\square$

In the next lecture we want to study further properties of generators. To this aim, we will need several concepts which we now explain. Here we only recall the basic definitions, results and examples; most proofs and more details can be found in Appendices A and B.

Intermezzo 1: Closed operators and their spectra

Let $\mathrm{D}(A) \subseteq X$ be a linear subspace and $A : \mathrm{D}(A) \to X$ be linear. The operator A is called *closed* if it holds:

If $(x_n)_n$ is any sequence in $\mathrm{D}(A)$ such that the limits $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} Ax_n = y$ exist in X, then $x \in \mathrm{D}(A)$ and $Ax = y$.

Note that any operator $A \in \mathcal{B}(X)$ is closed, where $\mathrm{D}(A) = X$. In the next example we introduce the prototype of an unbounded closed operator.

EXAMPLE 2.10. Let $X = C([0,1])$ and $Af := f'$ with $\mathrm{D}(A) = C^1([0,1])$. Take a sequence $(f_n)_n$ in $\mathrm{D}(A)$ such that $(f_n)_n$, respectively $(f_n')_n$, converge to f, respectively g, in X. It is a well known fact that then f belongs to $C^1([0,1])$ and $f' = g$ (see also Remark 2.11 (f) below), which means that A is closed.

Second, consider $A_0 f := f'$ with $\mathrm{D}(A_0) := \{f \in C^1([0,1]) \,|\, f'(0) = 0\}$. If $(f_n)_n$ is a sequence in $\mathrm{D}(A)$ such that $f_n \to f$ and $f_n' \to g$ in X as $n \to \infty$, then we obtain $f \in C^1([0,1])$ with $f' = g$ as above. Furthermore, $g(0) = f'(0) = \lim_{n\to\infty} f_n'(0) = 0$. Consequently, $f \in \mathrm{D}(A_0)$ and $A_0 f = g$, i.e., A_0 is closed. $\diamondsuit$

Next, we define the Riemann integral for vector valued functions. Let $a < b$ be real numbers. A *(tagged) partition* Z of the interval $[a,b]$ is given by finite sequences $(t_k)_{k=0}^m$ and $(\tau_k)_{k=1}^m$ in $[a,b]$ satisfying $t_{k-1} < t_k$ and $\tau_k \in [t_{k-1}, t_k]$ for all $k \in \{1, \dots, m\}$, where $t_0 = a$ and $t_m = b$. We set $\delta(Z) := \max_{k=1,\dots,m}(t_k - t_{k-1})$. For a function $g \in C([a,b], X)$ we define the *Riemann sum* $S(g,Z)$ (of g with respect to Z) by

$$S(g,Z) := \sum_{k=1}^{m} g(\tau_k)(t_k - t_{k-1}) \in X.$$

As for continuous real valued functions, it can be shown that for every sequence $(Z_n)_n$ of (tagged) partitions with $\lim_{n\to\infty} \delta(Z_n) = 0$ the sequence $(S(g, Z_n))_n$ converges in X and that the limit J does not depend on the choice of such $(Z_n)_n$. In this sense we say that $S(g,Z)$ converges in X to J as $\delta(Z) \to 0$. The *Riemann integral* $\int_a^b g(t)\,\mathrm{d}t$ is now defined as this limit, i.e.,

$$\int_a^b g(t)\,\mathrm{d}t := \lim_{\delta(Z)\to 0} S(g,Z).$$

The integral has the usual properties known from the real valued case (with similar proofs) like linearity, additivity and validity of the standard estimate

$$\left\| \int_a^b g(t)\,\mathrm{d}t \right\| \le (b-a)\|g\|_\infty.$$

The same definition and results work for piecewise continuous functions.

In the following remark we collect the properties of closed operators and the Riemann integral we need later. We especially emphasize that the fundamental theorem of calculus is valid also in the vector valued case (see part (e)) so that the substitution rule extends to this setting. The simple property (g) is used very often in these lectures.

REMARK 2.11. Let A be a linear operator on X. Then the following assertions hold.

(a) The operator A is closed if and only if the *graph* of A, i.e., the set

$$\mathrm{gr}(A) := \{(x, Ax) \,|\, x \in \mathrm{D}(A)\},$$

is closed in $X \times X$ (endowed with the norm given by $\|(x, y)\| = \|x\| + \|y\|$) if and only if $\mathrm{D}(A)$ is a Banach space with respect to the *graph norm* $\|x\|_A := \|x\| + \|Ax\|$. We write $[\mathrm{D}(A)]$ for $(\mathrm{D}(A), \|\cdot\|_A)$.

(b) If A is closed with $\mathrm{D}(A) = X$, then A is even continuous ("closed graph theorem").

(c) Let A be injective and set $\mathrm{D}(A^{-1}) = \mathrm{R}(A) := \{Ax \,|\, x \in \mathrm{D}(A)\}$. Then A is closed if and only if A^{-1} is closed.

(d) Let A be closed and $f \in C([a, b], X)$ with $f(t) \in \mathrm{D}(A)$ for each $t \in [a, b]$ such that $Af \in C([a, b], X)$, where $(Af)(t) := Af(t)$. We then have

$$\int_a^b f(t)\,\mathrm{d}t \in \mathrm{D}(A) \quad \text{and} \quad A\int_a^b f(t)\,\mathrm{d}t = \int_a^b Af(t)\,\mathrm{d}t.$$

An analogous result holds for piecewise continuous functions.

(e) For $f \in C([a, b], X)$ the function

$$[a, b] \to X, \quad t \mapsto \int_a^t f(\tau)\,\mathrm{d}\tau,$$

is differentiable with

$$\frac{\mathrm{d}}{\mathrm{dt}}\int_a^t f(\tau)\,\mathrm{d}\tau = f(t) \quad \text{for all } t \in [a, b]. \tag{2.6}$$

For $g \in C^1([a, b], X)$ and $t \in [a, b]$ we have

$$\int_a^t g'(\tau)\,\mathrm{d}\tau = g(t) - g(a). \tag{2.7}$$

(f) Let $(f_n)_n$ be a sequence in $C^1(J, X)$ and $f, g \in C(J, X)$ for an interval J such that $f_n \to f$ and $f_n' \to g$ uniformly on J as $n \to \infty$. Then $f \in C^1(J, X)$ and $f' = g$.

(g) Let $f \in C([a, b], X)$ and $t \in [a, b)$. Then $\frac{1}{h}\int_t^{t+h} f(s)\,\mathrm{d}s \to f(t)$ as $h \to 0^+$.

PROOF. Parts (a) and (c) are proved in Lemma A.6 of the appendix. Part (b) can be found in Theorem A.7.

(d) Let f be as in the statement. Clearly, $S(f, Z) \in \mathrm{D}(A)$ for any partition Z of $[a, b]$ and

$$AS(f, Z) = \sum_{k=1}^m (Af)(\tau_k)(t_k - t_{k-1}) = S(Af, Z) \;\longrightarrow\; \int_a^b Af(t)\,\mathrm{d}t$$

as $\delta(Z) \to 0$ because Af is continuous. The assertion now follows from the closedness of A.

(e) Let $t \in [a,b]$ and $h \neq 0$ such that $t+h \in [a,b]$. We estimate

$$\begin{aligned}\Big\| \frac{1}{h}\Big(\int_a^{t+h} f(\tau)\,\mathrm{d}\tau - \int_a^t f(\tau)\,\mathrm{d}\tau\Big) - f(t)\Big\| &= \Big\| \frac{1}{h}\int_t^{t+h} (f(\tau)-f(t))\,\mathrm{d}\tau\Big\| \\ &\le \tfrac{|h|}{|h|} \sup_{|\tau-t|\le h} \|f(\tau)-f(t)\| \;\longrightarrow\; 0\end{aligned}$$

as $h \to 0$. So we have shown (2.6). Putting $a = t$, we also derive (g).

To show (2.7), set $\varphi(t) = \int_a^t g'(\tau)\,\mathrm{d}\tau$ for $t \in [a,b]$. Equation (2.6) implies that $\varphi \in C^1([a,b],X)$ with $\varphi' = g'$. Therefore, $\varphi - g$ belongs to $C^1([a,b],X)$ with vanishing derivative. In the proof of Proposition 2.8 we have seen that thus $\varphi - g$ is constant, and hence (2.7) is true.

(f) Formula (2.7) gives

$$f_n(t) = f_n(a) + \int_a^t f_n'(\tau)\,\mathrm{d}\tau$$

for all $t \in J$. Letting $n \to 0$, we deduce that

$$f(t) = f(a) + \int_a^t g(\tau)\,\mathrm{d}\tau$$

for all $t \in J$. Hence, $f \in C^1(J,X)$ and $f' = g$ due to (2.6). □

It is a delicate mater to add or multiply closed operators. The situation is simpler if one operator is bounded, see Proposition A.9 in Appendix A.

Remark 2.12. Let A be closed and $T \in \mathcal{B}(X)$. Then the operators $B = A+T$ with $\mathrm{D}(B) = \mathrm{D}(A)$ and $C = AT$ with $\mathrm{D}(C) = \{x \in X \mid Tx \in \mathrm{D}(A)\}$ are closed. This applies in particular to the operator $\lambda I - A$ for $\lambda \in \mathbb{C}$. ♢

For a closed operator A, we define the *resolvent set*

$$\rho(A) := \{\lambda \in \mathbb{C} \mid \lambda I - A : \mathrm{D}(A) \to X \text{ is bijective}\}.$$

We write $R(\lambda, A)$ for $(\lambda I - A)^{-1}$ if $\lambda \in \rho(A)$. This operator is called *resolvent*. The *spectrum of* A is given by $\sigma(A) := \mathbb{C} \setminus \rho(A)$. Since $\lambda I - A$ is closed, $R(\lambda, A)$ is closed with domain X and thus bounded thanks to the closed graph theorem (see Remark 2.11(b)). It is known that $\rho(A)$ is open in $\mathbb{C}$ (and so $\sigma(A)$ is closed). More precisely, for $\lambda \in \rho(A)$ we have

$$B(\lambda, \|R(\lambda,A)\|^{-1}) \subseteq \rho(A), \tag{2.8}$$

as one can see by a Neumann series. Moreover, if $T \in \mathcal{B}(X)$, then $\sigma(T)$ is even compact and always non-empty, and the *spectral radius* of T is given by

$$r(T) := \max\{|\lambda| \mid \lambda \in \sigma(T)\} = \inf_{n\in\mathbb{N}} \|T^n\|^{\frac{1}{n}} = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}.$$

There are closed operators A with $\sigma(A) = \mathbb{C}$ or $\sigma(A) = \emptyset$ (see Example B.3 (b) and (c)). We have the *resolvent equation*

$$R(\mu, A) - R(\lambda, A) = (\lambda - \mu) R(\lambda, A) R(\mu, A) = (\lambda - \mu) R(\mu, A) R(\lambda, A)$$

for all $\lambda, \mu \in \rho(A)$. Furthermore, the map $\rho(A) \to \mathcal{B}(X)$, $\lambda \mapsto R(\lambda, A)$, is infinitely often differentiable (even analytic) with

$$\Big(\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big)^n R(\lambda, A) = (-1)^n\, n!\, R(\lambda, A)^{n+1} \tag{2.9}$$

for all $\lambda \in \rho(A)$ and $n \in \mathbb{N}_0$. These results are shown in Theorems B.4 and B.6 of the appendix.

Exercises

EXERCISE 2.1. Let $A \in \mathcal{B}(X)$ and $t \in \mathbb{R}$. Show that the series

$$\mathrm{e}^{tA} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

converges absolutely in $\mathcal{B}(X)$ uniformly for $t \in [-r, r]$, for any $r > 0$. Further show that $(\frac{\mathrm{d}}{\mathrm{dt}})^n \mathrm{e}^{tA} = A^n \mathrm{e}^{tA} = \mathrm{e}^{tA} A^n$ for all $t \in \mathbb{R}$ and $n \in \mathbb{N}$.

Let $A, B \in \mathcal{B}(X)$ with $AB = BA$. Show that $\mathrm{e}^{A+B} = \mathrm{e}^A \mathrm{e}^B = \mathrm{e}^B \mathrm{e}^A$. In particular $\mathrm{e}^{(t+s)A} = \mathrm{e}^{tA} \mathrm{e}^{sA}$ for $t, s \in \mathbb{R}$ and $\mathrm{e}^{\lambda I + A} = \mathrm{e}^{\lambda} \mathrm{e}^A$ for $\lambda \in \mathbb{C}$.

EXERCISE 2.2. Let $p \in [1, \infty)$ and $X = L^p(\mathbb{R})$. Set $T(t)f = f(\cdot + t)$ for $t \in \mathbb{R}$ and $f \in L^p(\mathbb{R})$. Show that $(T(t))_{t \in \mathbb{R}}$ is a C_0–group of isometries on X and that the map $\mathbb{R} \to \mathcal{B}(X)$, $t \mapsto T(t)$, is not continuous.

EXERCISE 2.3. Let $p \in [1, \infty)$ and $X = L^p(0, 1)$. For $t \ge 0$, $s \in (0, 1)$ and $f \in L^p(0, 1)$ set

$$(T(t)f)(s) := \begin{cases} f(s+t), & s + t < 1, \\ 0, & s + t \ge 1. \end{cases}$$

Show that $(T(t))_{t \ge 0}$ is a C_0–semigroup on X.

EXERCISE 2.4. Let $\emptyset \ne \Omega \subseteq \mathbb{R}^d$ be open, $X = C_0(\Omega)$ and $m \in C(\Omega)$ such that $\sup_{s \in \Omega} \operatorname{Re}(m(s)) < \infty$. Define $T(t)f = \mathrm{e}^{tm} f$ for $t \ge 0$ and $f \in X$. Show that $T(\cdot)$ is a C_0–semigroup on X generated by the operator

$$Af = mf \quad \text{with} \quad \mathrm{D}(A) = \{f \in X \,|\, mf \in X\}.$$

EXERCISE 2.5. Let A be a closed operator, $a \in \mathbb{C} \setminus \{0\}$ and $b \in \mathbb{C}$. Define $B = aA + b$ with $\mathrm{D}(B) = \mathrm{D}(A)$. Show that $\sigma(B) = a\sigma(A) + b$ and $R(\mu, B) = \frac{1}{a} R(\frac{\mu - b}{a}, A)$ for $\mu \in \rho(B)$.

LECTURE 3

Characterization of generators

In this lecture we establish the fundamental Hille-Yosida generation theorem. It characterizes those linear operators A that generate a C_0–semigroup $T(\cdot)$ of contractions. We recall that A should be considered as the given object and $T(\cdot)$ as the unknown. Much of the linear theory and the applications of our course will depend on this theorem and its companion, the Lumer-Phillips theorem shown in Lecture 4. We further state the version of the Hille-Yosida theorem for general semigroups at the end of the present lecture, without proving it.

We first derive necessary conditions for generation. These results help us to discuss the generators of certain translation semigroups. In addition, as in every mathematical field, there are several simple but useful facts and methods, which are used everywhere in this field and which we present here in detail. The first of these lemmas is concerned with an important rescaling procedure.

LEMMA 3.1. *Let A generate the C_0–semigroup $T(\cdot)$, $\lambda \in \mathbb{C}$, and $a > 0$. Then $(S(t))_{t\geq 0} := (\mathrm{e}^{\lambda t}T(at))_{t\geq 0}$ is also a C_0–semigroup with generator $B = \lambda I + aA$ and $\mathrm{D}(B) = \mathrm{D}(A)$.*

PROOF. For $t, s \geq 0$ we have $S(t+s) = \mathrm{e}^{\lambda t}T(at)\mathrm{e}^{\lambda s}T(as) = S(t)S(s)$, and also $S(0) = I$. The strong continuity of $S(\cdot)$ is clear so that $S(\cdot)$ is a C_0–semigroup. Next, let B be the generator of $S(\cdot)$. Because of

$$\tfrac{1}{t}(S(t)x - x) = a\mathrm{e}^{\lambda t}\tfrac{1}{at}(T(at)x - x) + \tfrac{1}{t}(\mathrm{e}^{\lambda t} - 1)x,$$

we have $x \in \mathrm{D}(B)$ if and only if $x \in \mathrm{D}(A)$, and in this case it holds $Bx = aAx + \lambda x$. □

The next result is a variant of the fundamental theorem of calculus for C_0–semigroups. Observe that the first part of Lemma 3.2 allows to produce elements of $\mathrm{D}(A)$.

LEMMA 3.2. *Let A generate the C_0–semigroup $T(\cdot)$, $t > 0$, and $x \in X$. Then $\int_0^t T(s)x\,\mathrm{d}s$ belongs to $\mathrm{D}(A)$ and*

$$T(t)x - x = A\int_0^t T(s)x\,\mathrm{d}s. \tag{3.1}$$

Furthermore, for $x \in \mathrm{D}(A)$ we have

$$T(t)x - x = \int_0^t T(s)Ax\,\mathrm{d}s. \tag{3.2}$$

PROOF. For $h > 0$ we compute

$$\begin{aligned}\tfrac{1}{h}(T(h)-I)\int_0^t T(s)x\,\mathrm{d}s &= \frac{1}{h}\Big(\int_0^t T(s+h)x\,\mathrm{d}s - \int_0^t T(s)x\,\mathrm{d}s\Big)\\ &= \frac{1}{h}\Big(\int_h^{t+h} T(r)x\,\mathrm{d}r - \int_0^t T(s)x\,\mathrm{d}s\Big)\\ &= \frac{1}{h}\int_t^{t+h} T(s)x\,\mathrm{d}s - \frac{1}{h}\int_0^h T(s)x\,\mathrm{d}s, \qquad (3.3)\end{aligned}$$

substituting $r = s+h$ in the second line. The final difference tends to $T(t)x - x$ as $h \to 0$ due to the continuity of the orbits and Remark 2.11 (g). But this fact precisely means that $\int_0^t T(s)x\,\mathrm{d}s$ is an element of $\mathrm{D}(A)$ and (3.1) holds.

If in addition $x \in \mathrm{D}(A)$, we have $T(\cdot)x \in C^1(\mathbb{R}_+, X)$ with $\frac{\mathrm{d}}{\mathrm{d}t}T(\cdot)x = T(\cdot)Ax$ by Proposition 2.8. Hence, (2.7) implies (3.2). □

We now derive the first necessary conditions for generation and show that the semigroup is uniquely determined by its generator.

PROPOSITION 3.3. *Let A generate a C_0–semigroup $T(\cdot)$. Then A is closed and densely defined. Moreover, $T(\cdot)$ is the only C_0–semigroup generated by A. Finally, we have $T(\cdot)x \in C(\mathbb{R}_+, [\mathrm{D}(A)])$ for each $x \in \mathrm{D}(A)$.*

PROOF. 1) To show the closedness of A, we take any sequence $(x_n)_n$ from $\mathrm{D}(A)$ with limit x in X such that $\lim_{n\to\infty} Ax_n = y$ for some $y \in X$. For $t > 0$ the equation (3.2) yields

$$\tfrac{1}{t}(T(t)x_n - x_n) = \frac{1}{t}\int_0^t T(s)Ax_n\,\mathrm{d}s$$

for all $n \in \mathbb{N}$, which leads to

$$\tfrac{1}{t}(T(t)x - x) = \frac{1}{t}\int_0^t T(s)y\,\mathrm{d}s$$

in the limit $n \to \infty$. Remark 2.11 (g) now implies that

$$\lim_{t\to 0}\tfrac{1}{t}(T(t)x - x) = y.$$

Consequently, $x \in \mathrm{D}(A)$ and $Ax = y$; i.e., A is closed.

2) For $x \in X$ and $n \in \mathbb{N}$ the vectors

$$x_n = n\int_0^{\frac{1}{n}} T(s)x\,\mathrm{d}s$$

belong to $\mathrm{D}(A)$ by Lemma 3.2. Remark 2.11 (g) says that $x_n \to x$ as $n \to \infty$. So we arrive at $\overline{\mathrm{D}(A)} = X$.

3) Let A generate another C_0–semigroup $S(\cdot)$. Then $S(\cdot)x$ solves (2.5) for every $x \in \mathrm{D}(A)$. But the uniqueness of such a solution (see Proposition 2.8) forces $T(t)x = S(t)x$ for all $t \geq 0$ and $x \in \mathrm{D}(A)$. Since $\mathrm{D}(A)$ is dense in X, the bounded operators $T(t)$ and $S(t)$ are equal for each $t \geq 0$.

4) For the final assertion, take $x \in \mathrm{D}(A)$ and $t, s \geq 0$. Proposition 2.8 yields

$$\begin{aligned}\|T(t)x - T(s)x\|_A &= \|T(t)x - T(s)x\| + \|A(T(t)x - T(s)x)\| \\ &= \|T(t)x - T(s)x\| + \|T(t)Ax - T(s)Ax\| \longrightarrow 0\end{aligned}$$

as $t \to s$, as asserted. □

PROPOSITION 3.4. *Let A generate the C_0–semigroup $T(\cdot)$ and $\lambda \in \mathbb{C}$. Then the following assertions hold.*

(a) If the improper integral

$$R(\lambda)x := \int_0^\infty \mathrm{e}^{-\lambda s} T(s)x \, \mathrm{d}s := \lim_{t\to\infty} \int_0^t \mathrm{e}^{-\lambda s} T(s)x \, \mathrm{d}s$$

exists in X for all $x \in X$, then $\lambda \in \rho(A)$ and $R(\lambda) = R(\lambda, A)$.

(b) The integral in (a) exists even absolutely for all $x \in X$ if $\operatorname{Re}\lambda > \omega_0(A)$. Hence, the spectral bound *of A, defined by*

$$s(A) := \sup \{\operatorname{Re}\lambda \mid \lambda \in \sigma(A)\},$$

is less than or equal to $\omega_0(A) < \infty$. (Here we put $\sup \emptyset = -\infty$.)

(c) Let $M \geq 1$ and $\omega \in \mathbb{R}$ such that $\|T(t)\| \leq M\mathrm{e}^{\omega t}$ for all $t \geq 0$. (One can take $\omega > \omega_0(A)$, see Lemma 2.3.) For all $n \in \mathbb{N}$ and $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$ we then have

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\operatorname{Re}\lambda - \omega)^n}. \tag{3.4}$$

By Proposition 3.4 (a) the resolvent of the generator is given as the “Laplace transform” of the semigroup. Moreover, the above propositions show that a generator A

$$\text{is closed and satisfies } \overline{\mathrm{D}(A)} = X, \;\; s(A) < \infty \;\text{ and }\; (3.4).$$

Actually, these conditions already imply that a linear operator A generates a C_0–semigroup, as we will see below in Theorem 3.11.

PROOF OF PROPOSITION 3.4. (a) Let $h > 0$ and $x \in X$. Recall from Lemma 3.1 that $T_\lambda(s) := \mathrm{e}^{-\lambda s}T(s)$ is generated by $A - \lambda I$ with domain $\mathrm{D}(A)$. Equation (3.3) yields

$$\begin{aligned}\tfrac{1}{h}(T_\lambda(h) - I)R(\lambda)x &= \lim_{t\to\infty} \tfrac{1}{h}(T_\lambda(h) - I)\int_0^t T_\lambda(s)x \, \mathrm{d}s \\ &= \lim_{t\to\infty} \frac{1}{h}\int_t^{t+h} T_\lambda(s)x \, \mathrm{d}s - \frac{1}{h}\int_0^h T_\lambda(s)x \, \mathrm{d}s = -\frac{1}{h}\int_0^h T_\lambda(s)x \, \mathrm{d}s,\end{aligned}$$

employing the convergence of $\int_0^\infty T_\lambda(s)x \, \mathrm{d}s$. By Remark 2.11 (g), we can now let $h \to 0$ to obtain that $R(\lambda)x \in \mathrm{D}(A - \lambda I) = \mathrm{D}(A)$ and

$$(A - \lambda I)R(\lambda)x = -x. \tag{3.5}$$

If $x \in \mathrm{D}(A)$, we deduce from $T(s)Ax = AT(s)x$ for $s \geq 0$ (see Proposition 2.8), Remark 2.11 (d) and (3.5) that

$$\begin{aligned} R(\lambda)(\lambda I - A)x &= \lim_{t\to\infty} \int_0^t \mathrm{e}^{-\lambda s} T(s)(\lambda I - A)x \,\mathrm{d}s \\ &= \lim_{t\to\infty} (\lambda I - A) \int_0^t \mathrm{e}^{-\lambda s} T(s)x \,\mathrm{d}s \\ &= (\lambda I - A) \lim_{t\to\infty} \int_0^t \mathrm{e}^{-\lambda s} T(s)x \,\mathrm{d}s = (\lambda I - A)R(\lambda)x = x, \end{aligned}$$

using that A is closed by Proposition 3.3. Hence, $\lambda \in \rho(A)$ and $R(\lambda) = R(\lambda, A)$.

(b) Observe that $\|\mathrm{e}^{-\lambda s} T(s)\| \leq M\mathrm{e}^{(\omega - \operatorname{Re}\lambda)s}$ for all $s \geq 0$, $\omega \in (\omega_0(A), \operatorname{Re}\lambda)$ and some $M \geq 1$. For $x \in X$ and $0 < a < b$ we can thus estimate

$$\Big\| \int_0^b T_\lambda(s)x \,\mathrm{d}s - \int_0^a T_\lambda(s)x \,\mathrm{d}s \Big\| \leq \int_a^b \|T_\lambda(s)x\| \,\mathrm{d}s \leq M \int_a^b \mathrm{e}^{(\omega - \operatorname{Re}\lambda)s} \,\mathrm{d}s \|x\| \to 0$$

as $a, b \to \infty$. Consequently, $\int_0^t T_\lambda(s)x \,\mathrm{d}s$ converges (absolutely) in X as $t \to \infty$ for each $x \in X$. Assertion (a) then implies that $\lambda \in \rho(A)$ if $\operatorname{Re}\lambda > \omega_0(A)$, and thus the second part of (b) also follows.

(c) For $n = 1$, as in part (b), one sees that

$$\|R(\lambda, A)x\| \leq \int_0^\infty M\mathrm{e}^{(\omega - \operatorname{Re}\lambda)t} \|x\| \,\mathrm{d}t = \frac{M}{\operatorname{Re}\lambda - \omega} \|x\|$$

for all $x \in X$, $\operatorname{Re}\lambda > \omega$ and the given $M \geq 1$, $\omega \in \mathbb{R}$. The general case can be shown employing also (2.9), see e.g. Corollary II.1.11 in [**EN99**]. □

For $\lambda \in \mathbb{C}$ and any interval $J \subseteq \mathbb{R}$ we set $\mathrm{e}_\lambda(t) = \mathrm{e}^{\lambda t}$ for $t \in J$. We next show that the first derivative $\frac{\mathrm{d}}{\mathrm{d}s}$ on suitable domains generates the left translation semigroups from Example 2.6 and that there are domains on which $\frac{\mathrm{d}}{\mathrm{d}s}$ is not a generator (since it violates the necessary spectral conditions). Roughly speaking, the domain has to include a boundary condition on the right end point of the underlying interval J where the left translation 'enters' J.

Example 3.5 (The first derivative as generator).

(a) Let $T(t)f = f(\,\cdot\, + t)$ be the left translation group on $X = C_0(\mathbb{R})$ and A be its generator. We show that $Au = u'$ with

$$\mathrm{D}(A) = C_0^1(\mathbb{R}) = \Big\{ f \in C^1(\mathbb{R}) \,\big|\, f, f' \in X \Big\}$$

and $\sigma(A) = \mathrm{i}\mathbb{R}$. In fact, for all $f \in \mathrm{D}(A)$ and $s \in \mathbb{R}$, the limits

$$Af(s) = \lim_{t\to 0} \tfrac{1}{t}(T(t)f(s) - f(s)) = \lim_{t\to 0} \tfrac{1}{t}(f(s+t) - f(s)) \tag{3.6}$$

exist in $\mathbb{C}$, so that f is differentiable with $f' = Af \in C_0(\mathbb{R})$, i.e., $\mathrm{D}(A) \subseteq C_0^1(\mathbb{R})$. Conversely, let $f \in C_0^1(\mathbb{R})$. For $s \in \mathbb{R}$, we obtain

$$\begin{aligned} \Big| \tfrac{1}{t}(T(t)f(s) - f(s)) - f'(s) \Big| &= \Big| \tfrac{1}{t}(f(s+t) - f(s)) - f'(s) \Big| \\ &= \Big| \frac{1}{t} \int_0^t (f'(s+\tau) - f'(s)) \,\mathrm{d}\tau \Big| \\ &\leq \tfrac{|t|}{|t|} \sup_{0\leq|\tau|\leq|t|} |f'(s+\tau) - f'(s)| \;\longrightarrow\; 0, \end{aligned}$$

as $t \to 0$ uniformly in $s \in \mathbb{R}$, since $f' \in C_0(\mathbb{R})$ is uniformly continuous. As a result, $f \in \mathrm{D}(A)$ and therefore $A = \frac{\mathrm{d}}{\mathrm{d}s}$ with $\mathrm{D}(A) = C_0^1(\mathbb{R})$.

Proposition 3.4 yields $s(A) \le \omega_0(A) = 0$, because of $\|T(t)\| = 1$. In the same way as above one sees that $-A$ generates the contraction C_0–semigroup $(S(t))_{t\ge 0} = (T(-t))_{t\ge 0}$. Hence, $s(-A) \le 0$. Due to $\lambda I - (-A) = -(-\lambda I - A)$, we have $\rho(-A) = -\rho(A)$ as well as $R(\lambda, -A) = -R(-\lambda, A)$. So we have shown that $\sigma(A) \subseteq \mathrm{i}\mathbb{R}$. To verify the converse inclusion, let $\tau \in \mathbb{R}$. Take functions $\varphi_n \in C_c^1(\mathbb{R})$ with $0 \le \varphi_n \le 1$, $\varphi_n(0) = 1$ and $\|\varphi_n'\|_\infty \le \frac{1}{n}$ for $n \in \mathbb{N}$. The function $f_n = \varphi_n \mathrm{e}_{\mathrm{i}\tau} \in \mathrm{D}(A)$ then satisfies $\|f_n\|_\infty = 1$ and $\mathrm{i}\tau f_n - Af_n = -\varphi_n' \mathrm{e}_{\mathrm{i}\tau}$. Consequently, $\|\mathrm{i}\tau f_n - Af_n\|_\infty \to 0$ as $n \to \infty$ and so $\mathrm{i}\tau I - A$ cannot have a bounded inverse. We thus obtain $\sigma(A) = \mathrm{i}\mathbb{R}$. ◊

(b) The nilpotent left translation semigroup on $X = C_0([0,1))$ is given by

$$(T(t)f)(s) = \begin{cases} f(s+t), & s+t < 1, \\ 0, & s+t \ge 1, \end{cases}$$

for $f \in X$, $t \ge 0$ and $s \in [0,1)$. We claim that $Au = u'$ with

$$\mathrm{D}(A) = C_0^1([0,1)) = \left\{ f \in C^1([0,1)) \,\middle|\, f, f' \in X \right\}$$

generates $T(\cdot)$ and that $\sigma(A) = \emptyset$. Indeed, as in (3.6), one shows that $Af = f'$ for $f \in \mathrm{D}(A) \subseteq C_0([0,1))$. Note that now in (3.6) one can only consider $t \to 0^+$. Therefore one has to use the following elementary fact, which follows from Corollary 2.1.2 of [**Paz83**]: If $g \in C([a,b))$ is differentiable from the right such that the right-hand side derivative $\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^+ g$ is continuous, then one already has $g \in C^1([a,b))$.

Conversely, let $f \in C_0^1([0,1))$. We set $\tilde f(s) = f(s)$ for $s \in [0,1)$ and $\tilde f(s) = 0$ for $s \ge 1$. Observe that $\tilde f \in C_0^1(\mathbb{R}_+)$ and $\tilde f'|_{[0,1)} = f'$. As in part (a), it follows

$$\begin{aligned} \tfrac{1}{t}(T(t)f(s) - f(s)) &= \begin{cases} \frac{1}{t}(f(s+t) - f(s)), & 0 \le s < 1-t, \\ -\frac{1}{t} f(s), & 1-t \le s < 1, \end{cases} \\ &= \tfrac{1}{t}(\tilde f(s+t) - \tilde f(s)) \;\longrightarrow\; \tilde f'(s) = f'(s) \end{aligned}$$

as $t \to 0^+$ uniformly in $s \in [0,1)$ since $\tilde f'$ is uniformly continuous. Hence, $\mathrm{D}(A) = C_0^1([0,1))$ and $Af = f'$. Here we have $\omega_0(A) = -\infty$ so that $\sigma(A) = \emptyset$ and $\rho(A) = \mathbb{C}$ due to $s(A) \le \omega_0(A)$ by Proposition 3.4. ◊

(c) The operator $Af = f'$ with $\mathrm{D}(A) = C^1([0,1])$ on $X = C([0,1])$ has the spectrum $\sigma(A) = \mathbb{C}$. In fact, the function e_λ belongs to $\mathrm{D}(A)$ with $A\mathrm{e}_\lambda = \lambda \mathrm{e}_\lambda$ for each $\lambda \in \mathbb{C}$. In view of Proposition 3.4, A thus cannot be a generator. ◊

(d) Let $X = C_0(\mathbb{R}_-) = \{ f \in C((-\infty, 0]) \mid \lim_{s\to-\infty} f(s) = 0 \}$ and $A = \frac{\mathrm{d}}{\mathrm{d}s}$ with $\mathrm{D}(A) = C_0^1(\mathbb{R}_-)$. Then A is not a generator. Indeed, for all λ with $\operatorname{Re}\lambda > 0$ we have $\mathrm{e}_\lambda \in \mathrm{D}(A)$ and $A\mathrm{e}_\lambda = \lambda \mathrm{e}_\lambda$ so that $\lambda \in \sigma(A)$, violating $s(A) < \infty$ in Proposition 3.4. ◊

(e) On the space $X = C([0,1])$ the operator $Af = f'$ with $\mathrm{D}(A) = \{ f \in C^1([0,1]) \mid f(1) = 0 \}$ is not a generator since $\overline{\mathrm{D}(A)} \subseteq \{ f \in X \mid f(1) = 0 \} \ne X$ (see Proposition 3.3). Actually, it holds $\overline{\mathrm{D}(A)} = \{ f \in X \mid f(1) = 0 \}$. ◊

Before we come to the Hille-Yosida theorem we show a few auxiliary facts. For linear operators A, B on X we write $A \subseteq B$ if $\mathrm{gr}(A) \subseteq \mathrm{gr}(B)$, i.e., if $\mathrm{D}(A) \subseteq \mathrm{D}(B)$ and $Ax = Bx$ for all $x \in \mathrm{D}(A)$. In this case we call B an *extension of* A. The next lemma is often used to compute a generator.

LEMMA 3.6. *Let A and B be linear operators with $A \subseteq B$ such that A is surjective and B is injective. Then $A = B$. In particular, we have $A = B$ whenever $A \subseteq B$ and $\rho(A) \cap \rho(B) \neq \emptyset$ are satisfied.*

PROOF. We have to show that $\mathrm{D}(B) \subseteq \mathrm{D}(A)$. Let $x \in \mathrm{D}(B)$. By the assumptions, there is a vector $y \in \mathrm{D}(A)$ with $Bx = Ay = By$. The injectivity of B then implies $x = y \in \mathrm{D}(A)$.

The addendum follows by considering $\lambda I - A$ and $\lambda I - B$ for some $\lambda \in \rho(A) \cap \rho(B)$. Clearly, $\lambda I - A \subseteq \lambda I - B$ if $A \subseteq B$. The statement just shown thus gives $\lambda I - A = \lambda I - B$, so that $A = B$. $\square$

LEMMA 3.7. *Let A be closed and $M, \omega \geq 0$ such that $(\omega, \infty) \subseteq \rho(A)$ and $\|R(\lambda, A)\| \leq \frac{M}{\lambda - \omega}$ for all $\lambda > \omega$. Then $\lambda R(\lambda, A)x \to x$ as $\lambda \to \infty$ for all $x \in \overline{\mathrm{D}(A)}$ and $\lambda AR(\lambda, A)y \to Ay$ as $\lambda \to \infty$ for all $y \in \mathrm{D}(A)$ with $Ay \in \overline{\mathrm{D}(A)}$.*

PROOF. Let $x \in \mathrm{D}(A)$ and $\lambda > \omega$. Observe that

$$\lambda R(\lambda, A)x - x = AR(\lambda, A)x = R(\lambda, A)Ax, \tag{3.7}$$

cf. Theorem B.4. The assumed estimate thus yields

$$\|\lambda R(\lambda, A)x - x\| \leq \tfrac{M}{\lambda - \omega} \|Ax\| \longrightarrow 0$$

as $\lambda \to \infty$. Since $\|\lambda R(\lambda, A)\| \leq M$ for $\lambda \geq \omega + 1$ by our assumption, the first assertion follows by approximation. The second claim is an immediate consequence of the first assertion, taking $x = Ay$ and using (3.7). $\square$

We now briefly describe the idea of the proof the Hille-Yosida theorem given below, which is due to Yosida. As already explained, for a bounded operator A we can construct the corresponding semigroup $T(t) = \mathrm{e}^{tA}$ by an exponential series. For unbounded A, such a construction fails. A different approach is suggested by the above lemma: If the resolvent of A satisfies the decay property (3.4) for $n = 1$, one can approximate A pointwise on $\mathrm{D}(A)$ by the bounded operators A_n on X, defined by

$$A_n := nAR(n, A) = n^2 R(n, A) - nI.$$

The semigroup e^{tA_n} is now defined, and we will see that it converges pointwise to a semigroup $T(\cdot)$ whose generator is A.

Here is the main result of this lecture. By a *contraction semigroup* we mean a C_0–semigroup $T(\cdot)$ such that each $T(t)$ is a contraction, i.e., $\|T(t)\| \leq 1$.

THEOREM 3.8 (Hille, Yosida 1948). *A linear operator A generates a contraction semigroup $T(\cdot)$ if and only if A is closed, densely defined, $(0, \infty) \subseteq \rho(A)$ and the* Hille-Yosida estimate

$$\|R(\lambda, A)\| \leq \frac{1}{\lambda} \tag{3.8}$$

holds for all $\lambda > 0$. In this case $\mathbb{C}_+ := \{z \in \mathbb{C} \mid \operatorname{Re} z > 0\}$ belongs to $\rho(A)$ and we have $\|R(\lambda, A)^n\| \leq (\operatorname{Re} \lambda)^{-n}$ for all $n \in \mathbb{N}$ and $\lambda \in \mathbb{C}_+$.

Proof. The necessity of the conditions and also the addendum follow from Propositions 3.3 and 3.4 (with $M = 1$ and $\omega = 0$). In order to prove sufficiency, we take $A_n = nAR(n, A) = n^2R(n, A) - nI$ for $n \in \mathbb{N}$. Lemma 3.7 and (3.8) imply that $\lim_{n\to\infty} A_n y = Ay$ for all $y \in \mathrm{D}(A)$. Let $t \geq 0$ and $n \in \mathbb{N}$. By means of Exercise 2.1 and (3.8), we obtain the crucial estimate

$$\begin{aligned} \|\mathrm{e}^{tA_n}\| = \|\mathrm{e}^{-tn}\mathrm{e}^{tn^2R(n,A)}\| &\leq \mathrm{e}^{-tn}\sum_{j=0}^{\infty}\frac{(tn^2\|R(n,A)\|)^j}{j!} \\ &= \mathrm{e}^{-tn}\mathrm{e}^{tn^2\|R(n,A)\|} \leq \mathrm{e}^{-tn}\mathrm{e}^{tn} = 1. \end{aligned} \tag{3.9}$$

Next, take $n, m \in \mathbb{N}$, $t_0 > 0$, $y \in \mathrm{D}(A)$ and $t \in [0, t_0]$. We have $A_nA_m = A_mA_n$ and hence

$$A_n\mathrm{e}^{tA_m} = A_n\sum_{j=0}^{\infty}\frac{t^j}{j!}A_m^j = \sum_{j=0}^{\infty}\frac{t^j}{j!}A_m^jA_n = \mathrm{e}^{tA_m}A_n.$$

Recall from Exercise 2.1 that $\frac{\mathrm{d}}{\mathrm{dt}}\mathrm{e}^{tA_n} = A_n\mathrm{e}^{tA_n}$. Using (2.7), we then compute

$$\begin{aligned} \mathrm{e}^{tA_n}y - \mathrm{e}^{tA_m}y &= \int_0^t \frac{\mathrm{d}}{\mathrm{ds}}(\mathrm{e}^{(t-s)A_m}\mathrm{e}^{sA_n}y)\,\mathrm{d}s = \int_0^t \mathrm{e}^{(t-s)A_m}(A_n - A_m)\mathrm{e}^{sA_n}y\,\mathrm{d}s \\ &= \int_0^t \mathrm{e}^{(t-s)A_m}\mathrm{e}^{sA_n}(A_n - A_m)y\,\mathrm{d}s. \end{aligned}$$

Estimate (3.9) now implies that

$$\|\mathrm{e}^{tA_n}y - \mathrm{e}^{tA_m}y\| \leq t_0\,\|A_ny - A_my\| \;\longrightarrow\; 0 \tag{3.10}$$

as $n, m \to \infty$. Hence, $(\mathrm{e}^{tA_n}y)_n$ is Cauchy and so the limit $T(t)y = \lim_{n\to\infty}\mathrm{e}^{tA_n}y$ exists in X. Since $\mathrm{D}(A)$ is dense in X and (3.9) holds, we obtain contractions $T(t) \in \mathcal{B}(X)$ given by $T(t)x = \lim_{n\to\infty}\mathrm{e}^{tA_n}x$ for all $t \geq 0$ and $x \in X$. Clearly, $T(0) = I$ and[1]

$$T(t+s)x = \lim_{n\to\infty}\mathrm{e}^{(t+s)A_n}x = \lim_{n\to\infty}\mathrm{e}^{tA_n}\mathrm{e}^{sA_n}x = T(t)T(s)x$$

for all $t, s \geq 0$. Letting $m \to \infty$ in (3.10), we further deduce that

$$\|\mathrm{e}^{tA_n}y - T(t)y\| \leq t_0\,\|A_ny - Ay\|$$

for all $t \in [0, t_0]$. As a result, $\mathrm{e}^{tA_n}y$ converges to $T(t)y$ uniformly for $t \in [0, t_0]$ so that $T(\cdot)y$ is continuous for all $y \in \mathrm{D}(A)$. Because of the density of $\mathrm{D}(A)$, $T(\cdot)$ is strongly continuous and thus a contraction semigroup.

Let B be the generator of $T(\cdot)$. We still have to establish $B = A$. Observe that $(0, \infty) \subseteq \rho(A) \cap \rho(B)$ due to Proposition 3.4 and the assumptions. In view of Lemma 3.6, it thus remains to show $A \subseteq B$. For $t > 0$ and $y \in \mathrm{D}(A)$, we conclude from (2.7) that

$$\tfrac{1}{t}(T(t)y - y) = \lim_{n\to\infty}\tfrac{1}{t}(\mathrm{e}^{tA_n}y - y) = \lim_{n\to\infty}\frac{1}{t}\int_0^t \mathrm{e}^{sA_n}A_ny\,\mathrm{d}s = \frac{1}{t}\int_0^t T(s)Ay\,\mathrm{d}s.$$

As $t \to 0$, Remark 2.11 (g) yields $y \in \mathrm{D}(B)$ and $By = Ay$, i.e., $A \subseteq B$. ☐

[1] Here we use that $T_nx_n \to Tx$ as $n \to \infty$ if $T_n, T \in \mathcal{B}(X)$, $x_n \to x$ in X and $T_ny \to Ty$ for each $y \in X$ as $n \to \infty$. This fact can be shown as Lemma 2.9 and is employed below without notice.

We point out that the Hille-Yosida theorem reduces the task of solving the evolution equation (2.5) to the study of the *stationary problem*

$$\lambda u - Au = f, \tag{3.11}$$

where $\lambda > 0$ and $f \in X$ are given. But besides the unique solvability of (3.11), we need the estimate $\|u\| = \|R(\lambda, A)f\| \le \frac{\|f\|}{\lambda}$. In the next lecture we will see that in the context of contraction semigroups one gets this estimate for free if one can solve (3.11) for some $\lambda > 0$ and all $f \in X$.

We continue with two illustrating examples, where the first one complements Examples 2.6 and 3.5.

EXAMPLE 3.9. Let $X = C_0(\mathbb{R}_-)$ and $A = -\frac{\mathrm{d}}{\mathrm{ds}}$ with $\mathrm{D}(A) = C_0^1(\mathbb{R}_-)$. We claim that A generates the right translation semigroup on X given by $T(t)f = f(\cdot - t)$.

We first check the assumptions of the Hille-Yosida theorem. Take $f_n \in \mathrm{D}(A)$ such that $f_n \to f$ and $f_n' \to g$ in X. Then $f \in C^1(\mathbb{R}_-)$ and $f' = g \in X$ so that $f \in \mathrm{D}(A)$. As a result, A is closed. Clearly, $C_c^1(\mathbb{R}_-) \subseteq \mathrm{D}(A)$. We want to show that $C_c^1(\mathbb{R}_-)$ is dense in X which yields the density of $\mathrm{D}(A)$. As in Example 2.6 one can see that $C_c(\mathbb{R}_-)$ is dense in $C_0(\mathbb{R}_-)$. Pick $f \in C_c(\mathbb{R}_-)$ with $\operatorname{supp} f \subseteq [a, 0]$. Weierstraß' approximation theorem gives a sequence of polynomials g_n converging to f uniformly on $[a-1, 0]$. Take a function $\varphi \in C_c^1(\mathbb{R}_-)$ with $\varphi = 1$ on $[a, 0]$ and $\operatorname{supp} \varphi \subseteq (a-1, 0]$. Then, $h_n = \varphi g_n \in C_c^1(\mathbb{R}_-)$ and

$$\|h_n - f\|_\infty \le \sup_{a-1 \le t < a} |\varphi(t) g_n(t) - 0| + \sup_{a \le t \le 0} |g_n(t) - f(t)| \longrightarrow 0$$

as $n \to \infty$, so that $C_c^1(\mathbb{R}_-)$ is dense in X. Let $f \in X$ and $\lambda > 0$. We then have

$$\begin{aligned} u \in \mathrm{D}(A) \text{ and } \lambda u - Au = f \iff{}& u' = -\lambda u + f,\ u \in C^1(\mathbb{R}_-) \text{ and } u \in X \\ \iff{}& u(s) = \int_{-\infty}^{s} \mathrm{e}^{-\lambda(s-\tau)} f(\tau)\,\mathrm{d}\tau =: R(\lambda)f(s) \\ & \text{for } s \le 0 \ \text{ and } \lim_{s \to -\infty} u(s) = 0. \end{aligned}$$

For any $\varepsilon > 0$, there is an $s_\varepsilon \le 0$ such that $|f(\tau)| \le \varepsilon$ for all $\tau \le s_\varepsilon$. For $s \le s_\varepsilon$ we thus estimate

$$|R(\lambda)f(s)| \le \int_{-\infty}^{s} \mathrm{e}^{-\lambda(s-\tau)} |f(\tau)|\,\mathrm{d}\tau \le \varepsilon \int_0^\infty \mathrm{e}^{-\lambda r}\,\mathrm{d}r = \frac{\varepsilon}{\lambda}.$$

As a result, $R(\lambda)f(s) \to 0$ as $s \to -\infty$ so that $\lambda \in \rho(A)$ and $R(\lambda) = R(\lambda, A)$. Moreover,

$$\|R(\lambda, A)f\|_\infty \le \sup_{s \le 0} \int_{-\infty}^{s} \mathrm{e}^{-\lambda(s-\tau)} \|f\|_\infty\,\mathrm{d}\tau = \|f\|_\infty \int_0^\infty \mathrm{e}^{-\lambda r}\,\mathrm{d}r = \frac{\|f\|_\infty}{\lambda}$$

for all $f \in X$ and $\lambda > 0$, i.e., $\|R(\lambda, A)\| \le \frac{1}{\lambda}$. Theorem 3.8 thus implies that A generates a contraction semigroup $T(\cdot)$.

To compute $T(t)$, we take $f \in \mathrm{D}(A)$ and set $u(t) = T(t)f$ for $t \ge 0$. Proposition 2.8 says that u is the unique function satisfying $u \in C^1(\mathbb{R}_+, X)$, $u(t) \in \mathrm{D}(A)$ for all $t \ge 0$, $u(0) = f$ and

$$u'(t) = Au(t) = -\frac{\mathrm{d}}{\mathrm{ds}} u(t), \quad t \ge 0. \tag{3.12}$$

We make a guess for u and put $v(t) = f(\,\cdot - t)$ for $t \geq 0$. Clearly, $v(t) \in X$ and $v(0) = f$. We will check that v solves (3.12) and so $u = v$. We compute

$$\begin{aligned} \Big\| \frac{v(t') - v(t)}{t' - t} + f'(\,\cdot - t) \Big\|_\infty &= \sup_{s \in \mathbb{R}_-} \Big| \frac{-1}{t' - t} \int_t^{t'} f'(s - \tau)\,\mathrm{d}\tau + f'(s - t) \Big| \\ &\leq \sup_{s \in \mathbb{R}_-} \sup_{|\tau - t| \leq |t' - t|} |f'(s - t) - f'(s - \tau)| \end{aligned}$$

for $t', t \geq 0$ with $t' \neq t$. Since f' is uniformly continuous, the right-hand side tends to 0 as $t' \to t$. Hence, $v : \mathbb{R}_+ \to X$ is differentiable with $\frac{\mathrm{d}}{\mathrm{d}t} v(t) = -f'(\,\cdot - t)$ for $t \geq 0$. As in Example 2.6, one verifies that $t \mapsto f'(\,\cdot - t) \in X$ is continuous and thus $v \in C^1(\mathbb{R}_+, X)$. One similarly sees that $v(t) \in C^1(\mathbb{R}_-)$ and $\frac{\mathrm{d}}{\mathrm{d}s} v(t) = f'(\,\cdot - t) \in X$ so that $v(t) \in \mathrm{D}(A)$ for all $t \geq 0$ and v solves (3.12). Therefore $T(t)f = v(t) = f(\,\cdot - t)$ for all $f \in \mathrm{D}(A)$ and $t \geq 0$. By density, this equation holds for all $f \in X$, as required.

Finally, we prove that $\sigma(A) = \{z \in \mathbb{C} \,|\, \operatorname{Re} z \leq 0\}$. If $\operatorname{Re}\lambda < 0$, then $\mathrm{e}_{-\lambda} \in \mathrm{D}(A)$ satisfies $A\mathrm{e}_{-\lambda} = -(\mathrm{e}_{-\lambda})' = \lambda \mathrm{e}_{-\lambda}$ so that $\lambda \in \sigma(A)$. Since $s(A) \leq \omega_0(A) = 0$, the claim follows from the closedness of $\sigma(A)$. ♢

The next example indicates that the assumptions in the Hille-Yosida theorem are essentially optimal.

Example 3.10. Let $X = C_0(\mathbb{R}) \times C_0(\mathbb{R})$ with $\|(f,g)\| := \max\{\|f\|_\infty, \|g\|_\infty\}$, let $m(s) := \mathrm{i}s$ and

$$A \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} mu + mv \\ mv \end{pmatrix} = \begin{pmatrix} m & m \\ 0 & m \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

with $\mathrm{D}(A) = \{(u,v) \in X \,|\, (mu, mv) \in X\}$. Since $C_c(\mathbb{R}) \times C_c(\mathbb{R}) \subseteq \mathrm{D}(A)$, the domain $\mathrm{D}(A)$ is dense in X. One can check that A is closed and that for $\operatorname{Re}\lambda > 0$ one has $\lambda \in \rho(A)$ with

$$R(\lambda, A) \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} \frac{1}{\lambda - m} & \frac{m}{(\lambda - m)^2} \\ 0 & \frac{1}{\lambda - m} \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} \frac{f}{\lambda - m} + \frac{mg}{(\lambda - m)^2} \\ \frac{g}{\lambda - m} \end{pmatrix},$$

cf. Proposition B.5. For $\lambda > 0$ and $\|(f,g)\| \leq 1$ we then estimate

$$\begin{aligned} \|R(\lambda, A)\left(\begin{smallmatrix} f \\ g \end{smallmatrix}\right)\| &\leq \max\Big\{ \Big\| \frac{f}{\lambda - m} \Big\|_\infty + \Big\| \frac{mg}{(\lambda - m)^2} \Big\|_\infty, \Big\| \frac{g}{\lambda - m} \Big\|_\infty \Big\} \\ &\leq \sup_{s \in \mathbb{R}} \frac{1}{|\lambda - \mathrm{i}s|} + \sup_{s \in \mathbb{R}} \frac{|s|}{|\lambda - \mathrm{i}s|^2} = \frac{1}{\lambda} + \sup_{s \in \mathbb{R}} \frac{|s|}{\lambda^2 + s^2} = \frac{3/2}{\lambda}. \end{aligned}$$

On the other hand, for $a > 0$ and $n \in \mathbb{N}$ we choose $g_n \in C_0(\mathbb{R})$ such that $g_n(n) = 1$ and $\|g_n\|_\infty = 1$. It then follows

$$\|R(a + \mathrm{i}n, A)\| \geq \Big\| R(a + \mathrm{i}n, A) \begin{pmatrix} 0 \\ g_n \end{pmatrix} \Big\| \geq \Big| \frac{\mathrm{i}n g_n(n)}{(a + \mathrm{i}n - \mathrm{i}n)^2} \Big| = \frac{n}{a^2}.$$

As a result, $R(\lambda, A)$ is unbounded on every imaginary line $\operatorname{Re}\lambda = a$, violating the estimate in Proposition 3.4. Thus A does not generate a C_0–semigroup though it almost satisfies the assertions of Theorem 3.8. ♢

There are even operators A that are not generators but satisfy $\|R(\lambda, A)\| \leq \frac{c}{\operatorname{Re}\lambda}$ for all $\lambda \in \mathbb{C}_+$ and some constant $c > 1$ (see Example 2 in §12.4 of [**HP57**]).

Of course, not all C_0–semigroups are contractive, cf. Exercise 3.4. We state below the characterization theorem for generators of general C_0–semigroups which is a cornerstone of semigroup theory. Since we focus on contraction semigroups in our course, we omit its proof. It can be found e.g. in Theorem II.3.8 of [**EN99**], where the theorem is reduced to the Hille-Yosida theorem by a renorming and rescaling argument.

Theorem 3.11 (Feller, Miyadera, Phillips 1952). *A linear operator A generates a C_0–semigroup $T(\cdot)$ satisfying $\|T(t)\| \leq M\mathrm{e}^{\omega t}$ for all $t \geq 0$ and some $M \geq 1$ and $\omega \in \mathbb{R}$ if and only if A is closed, $\overline{\mathrm{D}(A)} = X$, $(\omega, \infty) \subseteq \rho(A)$ and*

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\lambda - \omega)^n}$$

holds for all $\lambda \in (\omega, \infty)$ and all $n \in \mathbb{N}$. If this is the case, we also have $\{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda > \omega\} \subseteq \rho(A)$ as well as $\|R(\lambda, A)^n\| \leq \frac{M}{(\operatorname{Re}\lambda - \omega)^n}$ for all $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega$ and all $n \in \mathbb{N}$.

Exercises

EXERCISE 3.1. Let A generate the C_0–semigroup $T(\cdot)$ on a Banach space X. Let $J : X \to E$ be an isomorphism to another Banach space E, and let Y be a Banach space which is densely embedded into X (where we identify Y with a subspace of X). Assume that $T(t)Y \subseteq Y$ for all $t \geq 0$ and $T(\cdot)y \in C(\mathbb{R}_+, Y)$ for all $y \in Y$. For $t \geq 0$ we define the operators

$$S(t) = JT(t)J^{-1} \text{ on } E, \quad T_Y(t)y = T(t)y \text{ on } Y.$$

Show that $S(\cdot)$ and $T_Y(\cdot)$ are C_0–semigroups on E and Y, respectively, and compute their generators. How does the result simplify if $\|\cdot\|_Y$ is equivalent to $\|\cdot\|_X$ (and thus Y is closed in X)?

For Exercise 3.2 and 3.3 we need Sobolev spaces on open intervals $J \subseteq \mathbb{R}$. Sobolev spaces on open subsets of $\mathbb{R}^d$ are discussed in Lecture 5. Just for these exercises we define them on intervals in an equivalent way by requiring that the fundamental theorem of calculus holds. Let $p \in [1, \infty]$. We set

$$W_p^1(J) = \left\{ f \in C(\overline{J}) \cap L^p(J) \,\middle|\, \exists\, g \in L^p(J)\ \forall\, t, s \in J : f(t) - f(s) = \int_s^t g(\tau)\, \mathrm{d}\tau \right\}.$$

(To be precise: By $f \in C(\overline{J}) \cap L^p(J)$ we mean that $f \in L^p(J)$ possesses a continuous representative which has a continuous extension to $\overline{J}$.) For $f \in W_p^1(J)$ one sets $g := f'$, where g is given by the definition of $f \in W_p^1(J)$. By Lebesgue's differentiation theorem (see e.g. Theorem 7.7 in [**Rud87**]) the "generalized derivative" f' is uniquely determined. The definition implies the following useful fact (show it!).

If $u_n \in W_p^1(J)$ and u_n' converge to u and f in $L^p(J)$, respectively, then $u \in W_p^1(J)$ and $u' = f$.

EXERCISE 3.2. Let $X = L^p(\mathbb{R})$, $p \in [1, \infty)$ and $T(t)f = f(\cdot + t)$ for $t \in \mathbb{R}$. From Exercise 2.2 we know that $T(\cdot)$ is a C_0–group. Show that its generator is given by $Au = u'$ on $\mathrm{D}(A) = W_p^1(\mathbb{R})$. Further show that $\sigma(A) = \mathrm{i}\mathbb{R}$.

EXERCISE 3.3. Let $X = L^p(0, 1)$ and $p \in [1, \infty)$. For $t \geq 0$ and $s \in (0, 1)$ set

$$T(t)f(s) = \begin{cases} f(s+t), & s + t < 1, \\ 0, & s + t \geq 1. \end{cases}$$

From Exercise 2.3 we know that $T(\cdot)$ is a C_0–semigroup. Show that its generator is given by $Au = u'$ on $\mathrm{D}(A) = \left\{ u \in W_p^1(0, 1) \,\middle|\, u(1) = 0 \right\}$.

EXERCISE 3.4. A C_0–semigroup $T(\cdot)$ is called *quasicontractive* if there is a number $\omega \in \mathbb{R}$ such that $\|T(t)\| \leq \mathrm{e}^{\omega t}$ holds for all $t \geq 0$. Characterize the generators of such C_0–semigroups as in Theorem 3.8 (without using Theorem 3.11).

Find an equivalent norm on $X = C_0(\mathbb{R})$ such that the left translation group is not quasicontractive for this norm.

LECTURE 4

Dissipative operators and the Lumer-Phillips theorem

In this lecture we first characterize the generators of C_0–*groups* in the spirit of Theorem 3.11. It turns out that an operator A generates a C_0–group if and only if A and $-A$ generate C_0–semigroups corresponding to forward and backward time. Afterwards we focus again on contraction semigroups and describe their generators by means of the Lumer-Phillips theorem (which will be extended to the group case in the next lecture). To explain the relevance of the Lumer-Phillips theorem, we recall the Hille-Yosida theorem from the previous lecture.

THEOREM 3.8. *A linear operator A generates a contraction semigroup $T(\cdot)$ if and only if A is closed, densely defined, $(0,\infty) \subseteq \rho(A)$ and*

$$\|R(\lambda, A)\| \leq \tfrac{1}{\lambda} \tag{3.8}$$

holds for all $\lambda > 0$. In this case $\mathbb{C}_+ = \{z \in \mathbb{C} \mid \operatorname{Re} z > 0\}$ belongs to $\rho(A)$ and we have $\|R(\lambda, A)^n\| \leq (\operatorname{Re}\lambda)^{-n}$ for all $n \in \mathbb{N}$ and $\lambda \in \mathbb{C}_+$.

Often the estimate (3.8) is hard to verify in examples, in particular since it involves the usually unknown resolvent. In the Lumer-Phillips theorem the assumption (3.8) will be replaced by conditions on A itself, namely its "dissipativity" and a range condition. As we will see, this result is very powerful in many important applications. For this reason we discuss the notion of dissipativity in more detail below.

Our first lemma provides the essential step for the subsequent characterization theorem for generators of C_0–groups.

LEMMA 4.1. *Let $T(\cdot)$ be a C_0–semigroup and $t_0 > 0$ such that $T(t_0)$ is invertible. Then $T(\cdot)$ can be extended to a C_0–group $(T(t))_{t\in\mathbb{R}}$.*

PROOF. Take $M \geq 1$ and $\omega \in \mathbb{R}$ such that $\|T(t)\| \leq M\mathrm{e}^{\omega t}$ holds for all $t \geq 0$ and set $c := \|T(t_0)^{-1}\|$. Let $0 \leq t \leq t_0$. We then have

$$\begin{aligned} T(t_0) &= T(t_0 - t)T(t) = T(t)T(t_0 - t), \\ I &= T(t_0)^{-1}T(t_0 - t)T(t) = T(t)T(t_0 - t)T(t_0)^{-1}, \end{aligned}$$

so that $T(t)$ has the inverse $T(t_0)^{-1}T(t_0 - t)$ with norm less than or equal to $M_1 := cM\mathrm{e}^{|\omega|t_0}$. Furthermore, let $t = nt_0 + \tau$ for some $n \in \mathbb{N}$ and $\tau \in [0, t_0)$. In this case $T(t) = T(\tau)T(t_0)^n$ has the inverse $T(t_0)^{-n}T(\tau)^{-1}$. Consequently, $T(t)$ is invertible for all $t \geq 0$ and we can define $T(t) := T(-t)^{-1}$ for $t < 0$.

This definition gives a group since for $t, s \geq 0$ we can calculate

$$\begin{aligned} T(-t)T(-s) &= T(t)^{-1}T(s)^{-1} = (T(s)T(t))^{-1} = T(s+t)^{-1} = T(-s-t), \\ T(-t)T(s) &= (T(s)T(t-s))^{-1}T(s) = T(t-s)^{-1}T(s)^{-1}T(s) \\ &= T(t-s)^{-1} = T(s-t) \qquad \text{for} \quad t \geq s, \\ T(-t)T(s) &= T(t)^{-1}T(t)T(s-t) = T(s-t) \qquad \text{for} \quad s \geq t, \end{aligned}$$

and similarly for $T(s)T(-t)$. Let $t \in [0, t_0]$ and $x \in X$. We then obtain

$$\|T(-t)x - x\| = \|T(-t)(x - T(t)x)\| \leq M_1\|x - T(t)x\| \longrightarrow 0$$

as $t \to 0$. Hence, $(T(t))_{t\in\mathbb{R}}$ is a C_0–group by Lemma 2.5. □

THEOREM 4.2. *Let A be a linear operator, $M \geq 1$ and $\omega \geq 0$. The following assertions are equivalent.*

(a) *The operator A generates a C_0–group $(T(t))_{t\in\mathbb{R}}$ with $\|T(t)\| \leq M\mathrm{e}^{\omega|t|}$ for all $t \in \mathbb{R}$.*
(b) *The operator A generates a C_0–semigroup $(T_+(t))_{t\geq 0}$ and $-A$ with $\mathrm{D}(-A) := \mathrm{D}(A)$ generates a C_0–semigroup $(T_-(t))_{t\geq 0}$ such that $\|T_\pm(t)\| \leq M\mathrm{e}^{\omega t}$ for all $t \geq 0$.*
(c) *The operator A is closed, densely defined and for all $\lambda \in \mathbb{R}$ with $|\lambda| > \omega$ we have $\lambda \in \rho(A)$ and $\|R(\lambda, A)^n\| \leq M(|\lambda| - \omega)^{-n}$ for all $n \in \mathbb{N}$.*

If any (and thus all) of these conditions is (are) fulfilled, then $T_+(t) = T(t)$ and $T_-(t) = T(-t)$ for all $t \geq 0$. Moreover, for all $\lambda \in \mathbb{C}$ with $|\operatorname{Re}\lambda| > \omega$ we obtain $\lambda \in \rho(A)$ and $\|R(\lambda, A)^n\| \leq M(|\operatorname{Re}\lambda| - \omega)^{-n}$ for all $n \in \mathbb{N}$.

PROOF. "(a)⇒(b)": Setting $T_+(t) := T(t)$ and $T_-(t) := T(-t)$ for $t \geq 0$, one obtains C_0–semigroups with generators $A_\pm$. Since there exists $\frac{\mathrm{d}}{\mathrm{dt}}T(0)x = Ax$ for all $x \in \mathrm{D}(A)$, we have $A \subseteq \pm A_\pm$. To show the converse, first let $x \in \mathrm{D}(A_+)$ and $t > 0$. We compute

$$\begin{aligned} \tfrac{1}{t}(T(t)x - x) &= \tfrac{1}{t}(T_+(t)x - x) \longrightarrow A_+x, \\ -\tfrac{1}{t}(T(-t)x - x) &= T(-t)\tfrac{1}{t}(T_+(t)x - x) \longrightarrow A_+x \end{aligned}$$

as $t \to 0$, where we use Lemma 2.9. Hence, $x \in \mathrm{D}(A)$. Similarly, for $x \in \mathrm{D}(A_-)$ and $t > 0$ we derive

$$\begin{aligned} -\tfrac{1}{t}(T(-t)x - x) &= -\tfrac{1}{t}(T_-(t)x - x) \longrightarrow -A_-x, \\ \tfrac{1}{t}(T(t)x - x) &= -T(t)\tfrac{1}{t}(T_-(t)x - x) \longrightarrow -A_-x \end{aligned}$$

as $t \to 0$, so that $x \in \mathrm{D}(A)$. Therefore $A = \pm A_\pm$.

"(b)⇒(c)": For $\lambda > \omega$, the assertion follows from Propositions 3.3 and 3.4. For $\lambda < -\omega$, we recall that $\rho(A) = -\rho(-A)$ with $R(\lambda, A) = -R(-\lambda, -A)$. So we also obtain the asserted estimate for $\lambda < -\omega$. In the same way one shows the last assertion involving $\operatorname{Re}\lambda$.

"(c)⇒(a)": By Theorem 3.11, A and $-A$ generate the C_0–semigroups $(T_+(t))_{t\geq 0}$ and $(T_-(t))_{t\geq 0}$, respectively (arguing for $-A$ as in the previous step). Let $x \in \mathrm{D}(A) = \mathrm{D}(-A)$ and $t \geq s \geq 0$. Proposition 2.8 and Lemma 2.9 imply

$$\frac{\mathrm{d}}{\mathrm{ds}}(T_+(s)T_-(s)x) = T_+(s)(-A)T_-(s)x + T_+(s)AT_-(s)x = 0,$$

and hence $T_+(t)T_-(t)x = x$. Analogously, one obtains $T_-(t)T_+(t)x = x$. By approximation, it follows that $I = T_+(t)T_-(t) = T_-(t)T_+(t)$. Lemma 4.1 thus allows to extend the C_0–semigroup $T_+(\cdot)$ to a C_0–group with generator B. We have $B \subseteq A$ by definition and so $A = B$ due to Lemma 3.6 and $s(B) < \infty$. ☐

We note that the above proof relies on Theorem 3.11 which was only shown for $M = 1$. However, we will use Theorem 4.2 just for this case.

We now change the topic and discuss the concept of dissipativity which is crucial for the Lumer-Phillips theorem.

Definition 4.3. *The* duality set $J(x)$ *of* $x \in X$ *is defined by*

$$J(x) := \left\{ x^* \in X^* \,|\, \langle x, x^* \rangle = \|x\|^2 = \|x^*\|^2 \right\}.$$

A linear operator A *is called* dissipative *if for each* $x \in \mathrm{D}(A)$ *there is an* $x^* \in J(x)$ *such that* $\mathrm{Re}\langle Ax, x^* \rangle \le 0$.

The Hahn-Banach theorem ensures that $J(x) \neq \emptyset$ for every $x \in X$ (see Exercise 4.2). We often use that $\|x\| = \|x^*\|$ if $x^* \in J(x)$. In the standard function spaces, elements in the duality set can be computed explicitly.

Example 4.4. (a) Let X be a Hilbert space with scalar product $(\cdot\,|\,\cdot)$ and let $x \in X$. We show that $J(x) = \{(\cdot\,|x)\}$. Clearly, $y^* = (\cdot\,|x) \in J(x)$. Let $y^* \in J(x)$. Riesz' representation theorem gives a unique $y \in X$ such that $\langle z, y^* \rangle = (z|y)$ holds for all $z \in X$, and $\|y\| = \|y^*\|$. From $y^* \in J(x)$ we then deduce

$$\|x\| = \|y\| \qquad \text{and} \qquad (x|y) = \|x\|\,\|y\| \tag{4.1}$$

The last identity shows that x and y are linearly dependent (due to the characterization of equality in the Cauchy-Schwarz inequality). In view of the first equality in (4.1) there thus exists $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ and $x = \alpha y$, so that $x = y$ by the second part of (4.1). ♢

(b) Let $E = L^p(B)$ for $p \in [1, \infty)$ and a Borel set $\emptyset \neq B \subseteq \mathbb{R}^d$. Further, let $f \in E \setminus \{0\}$. We isometrically identify E^* with $L^{p'}(B)$ (via the usual duality pairing), where $p' := \frac{p}{p-1}$ for $p > 1$ and $1' := \infty$. We show that the function

$$g := \|f\|_p^{2-p} \overline{f} |f|^{p-2}$$

belongs to $J(f)$, where $\frac{0}{0} := 0$. (More precisely the functional $h \mapsto \int_B gh \,\mathrm{d}x$ on E is contained in $J(f)$.) We set $f^* := \overline{f}|f|^{p-2}$. For $p = 1$, we have $\|f^*\|_\infty = 1$. For $p > 1$, it holds

$$\|f^*\|_{p'} = \Big(\int_B |f|^{(p-1)\cdot \frac{p}{p-1}} \,\mathrm{d}x \Big)^{\frac{p-1}{p}} = \|f\|_p^{p-1}.$$

In both cases it follows that $\|g\|_{p'} = \|f\|_p$. Finally,

$$\langle f, g \rangle_{L^p} := \int_B fg \,\mathrm{d}x = \|f\|_p^{2-p} \int_B |f|^{p-2} \cdot \overline{f} \cdot f \,\mathrm{d}x = \|f\|_p^{2-p} \cdot \|f\|_p^p = \|f\|_p^2$$

so that $g \in J(f)$. We even have $J(f) = \{g\}$ if $p \in (1, \infty)$, see Exercise 4.2. ♢

(c) Let $\emptyset \neq U \subseteq \mathbb{R}^d$ be open and $E = C_0(U)$. For $f \in E$ there is an $x_0 \in U$ with $|f(x_0)| = \|f\|_\infty$. Set $\varphi(g) := \overline{f}(x_0) g(x_0)$ for $g \in E$. Then $\varphi \in J(f)$

since $\varphi \in E^*$, $\|\varphi\| = |f(x_0)| = \|f\|_\infty$ and $\varphi(f) = |f(x_0)|^2 = \|f\|_\infty^2$. The same construction works on $E = C(K)$ for compact $K \subseteq \mathbb{R}^d$. $\diamondsuit$

We next characterize dissipativity linking it to the Hille-Yosida estimate (3.8).

Proposition 4.5. *A linear operator A is dissipative if and only if*

$$\|\lambda x - Ax\| \geq \lambda \|x\|$$

holds for all $\lambda > 0$ and $x \in \mathrm{D}(A)$. If A generates a contraction semigroup $T(\cdot)$, then A is dissipative and we have $\mathrm{Re}\langle Ax, x^*\rangle \leq 0$ for all $x \in \mathrm{D}(A)$ and even for all $x^* \in J(x)$.

Proof. 1) Let A generate the contraction semigroup $T(\cdot)$, $x \in \mathrm{D}(A)$ and $x^* \in J(x)$. The contractivity implies

$$\begin{aligned}\mathrm{Re}\langle Ax, x^*\rangle &= \lim_{t\to 0} \mathrm{Re}\langle \tfrac{1}{t}(T(t)x - x), x^*\rangle = \lim_{t\to 0} \tfrac{1}{t}(\mathrm{Re}\langle T(t)x, x^*\rangle - \|x\|^2) \\ &\leq \limsup_{t\to 0} \tfrac{1}{t}(\|x\| \cdot \|x^*\| - \|x\|^2) = 0.\end{aligned}$$

2) Let A be dissipative and $x \in \mathrm{D}(A)$. There is an $x^* \in J(x)$ such that $\mathrm{Re}\langle Ax, x^*\rangle \leq 0$. So we obtain for each $\lambda > 0$ that

$$\lambda\|x\|^2 \leq \mathrm{Re}(\lambda\langle x, x^*\rangle) - \mathrm{Re}\langle Ax, x^*\rangle \leq |\langle \lambda x - Ax, x^*\rangle| \leq \|\lambda x - Ax\| \cdot \|x^*\|.$$

Hence, $\lambda\|x\| \leq \|\lambda x - Ax\|$ since $\|x\| = \|x^*\|$.

3) Assume that $\|\lambda x - Ax\| \geq \lambda\|x\|$ holds for all $\lambda > 0$ and $x \in \mathrm{D}(A)$. We first show the dissipativity of A if X is a Hilbert space since this case is very easy to treat. Here the assumption leads to

$$\lambda^2\|x\|^2 \leq \|\lambda x - Ax\|^2 = \lambda^2\|x\|^2 - 2\lambda\,\mathrm{Re}(Ax|x) + \|Ax\|^2$$

which is equivalent to

$$\mathrm{Re}(Ax|x) \leq \tfrac{1}{2\lambda}\|Ax\|^2.$$

Letting $\lambda \to \infty$, the dissipativity follows in view of Example 4.4.

We turn back to the case of a general Banach space X. Assume without loss of generality that $\|x\| = 1$. Take $y_\lambda^* \in J(\lambda x - Ax)$. Because of $\|y_\lambda^*\| = \|\lambda x - Ax\| \geq \lambda\|x\| = \lambda > 0$, we have $y_\lambda^* \neq 0$. Set $x_\lambda^* = \frac{1}{\|y_\lambda^*\|} y_\lambda^*$ for $\lambda > 0$. We then obtain $\|x_\lambda^*\| = 1$ as well as

$$\begin{aligned}\lambda \leq \|\lambda x - Ax\| &= \frac{1}{\|y_\lambda^*\|}\langle \lambda x - Ax, y_\lambda^*\rangle = \mathrm{Re}\langle \lambda x - Ax, x_\lambda^*\rangle \\ &= \lambda\,\mathrm{Re}\langle x, x_\lambda^*\rangle - \mathrm{Re}\langle Ax, x_\lambda^*\rangle \leq \min\{\lambda - \mathrm{Re}\langle Ax, x_\lambda^*\rangle,\ \lambda\,\mathrm{Re}\langle x, x_\lambda^*\rangle + \|Ax\|\}.\end{aligned}$$

As a result, $\mathrm{Re}\langle Ax, x_\lambda^*\rangle \leq 0$ and $1 - \frac{1}{\lambda}\|Ax\| \leq \mathrm{Re}\langle x, x_\lambda^*\rangle$ for all $\lambda > 0$. We restrict x_λ^* to the two dimensional subspace $E = \mathrm{lin}\{x, Ax\}$ of X. Since $\|x_\lambda^*\| \leq 1$, there are a functional $y^* \in E^*$ and a sequence $\lambda_j \to \infty$ such that $x_{\lambda_j}^* \to y^*$ in E^* as $j \to \infty$. Then $\mathrm{Re}\langle Ax, y^*\rangle \leq 0$ and $\mathrm{Re}\langle x, y^*\rangle \geq 1$. The Hahn-Banach theorem now yields $x^* \in X^*$ satisfying $\|x^*\| \leq 1$, $\mathrm{Re}\langle Ax, x^*\rangle \leq 0$ and $1 \leq \mathrm{Re}\langle x, x^*\rangle$. It remains to check $x^* \in J(x)$. We note that

$$1 \leq \mathrm{Re}\langle x, x^*\rangle \leq |\langle x, x^*\rangle| \leq \|x\| \cdot \|x^*\| = \|x^*\| \leq 1.$$

Hence, $1 = \|x^*\| = \|x\| = \langle x, x^*\rangle$ and so $x^* \in J(x)$. $\square$

We next discuss the dissipativity of certain differential operators of first order.

Example 4.6. (a) Let $X = C_0(\mathbb{R})$. Let $b, c \in C_b(\mathbb{R})$ be real-valued. Define $Au = bu' + cu$ with $\mathrm{D}(A) = C_0^1(\mathbb{R})$. We show that $A - \|c\|_\infty I$ is dissipative. In fact, take $u \in \mathrm{D}(A)$ and let $\varphi \in J(u)$ be given by $\varphi(v) = \overline{u}(s_0)v(s_0)$, where $|u(s_0)| = \|u\|_\infty$ (see Example 4.4). We then have

$$\begin{aligned}\operatorname{Re}\langle Au - \|c\|_\infty u, \varphi\rangle &= b(s_0)\operatorname{Re}(\overline{u}(s_0)u'(s_0)) + (c(s_0) - \|c\|_\infty)\operatorname{Re}(\overline{u}(s_0)u(s_0)) \\ &\le b(s_0)\operatorname{Re}(\overline{u}(s_0)u'(s_0)).\end{aligned}$$

We set $h(s) := \operatorname{Re}(\overline{u}(s_0)u(s))$ for $s \in \mathbb{R}$. Clearly, $h \in C_0^1(\mathbb{R})$ is real-valued and

$$|u(s_0)|^2 = h(s_0) \le \|h\|_\infty \le |u(s_0)| \cdot \|u\|_\infty = |u(s_0)|^2$$

so that $h(s_0) = \max h$. Hence, $h'(s_0) = 0$ and $\operatorname{Re}\langle Au - \|c\|_\infty u, \varphi\rangle \le 0$. ♢

(b) Let $X = L^2(\mathbb{R})$ and $A = \frac{\mathrm{d}}{\mathrm{d}s}$ with $\mathrm{D}(A) = C_c^1(\mathbb{R})$. For $u \in \mathrm{D}(A)$ we have $\overline{u} \in J(u)$ by Example 4.4 (b). Integration by parts yields

$$2\operatorname{Re}\langle Au, \overline{u}\rangle = \langle Au, \overline{u}\rangle + \overline{\langle Au, \overline{u}\rangle} = \int_{\mathbb{R}} u'\overline{u}\,\mathrm{d}s + \int_{\mathbb{R}} \overline{u}'u\,\mathrm{d}s = 0.$$

Hence, A is dissipative (but not closed as one easily checks). ♢

(c) Let $X = L^2(0,1)$, $A_j = \frac{\mathrm{d}}{\mathrm{d}s}$ and $\mathrm{D}(A_j) = \{u \in C^1([0,1]) \mid u(j) = 0\}$ for $j \in \{0,1\}$. ($\mathrm{D}(A_j)$ is considered as a subspace of X.) For $u \in \mathrm{D}(A_j)$, we take again $\overline{u} \in J(u)$ and obtain

$$2\operatorname{Re}\langle A_j u, \overline{u}\rangle = \int_0^1 u'\overline{u}\,\mathrm{d}s + \int_0^1 \overline{u}'u\,\mathrm{d}s = u\overline{u}\big|_0^1 = |u(1)|^2 - |u(0)|^2.$$

Hence, A_1 is dissipative. Moreover, $A_0 + \omega I$ is not dissipative for any $\omega \in \mathbb{R}$, since we can find a function $u \in \mathrm{D}(A_0)$ (depending on ω) such that

$$\operatorname{Re}\langle A_0 u + \omega u, \overline{u}\rangle = \tfrac{1}{2}|u(1)|^2 + \omega\|u\|_2 > 0.$$

(Note that here $\{\overline{u}\}$ can be identified with $J(u)$ by Example 4.4.) ♢

The examples (b) and (c) can be extended to L^p with $p \in [1,\infty)$. In these examples, we have encountered dissipative, but non-closed operators. To find closed extensions of them, we will need Sobolev spaces introduced in the next lecture, and the concepts of closability and of the closure of a linear operator A. These concepts also play an important role in the Lumer-Phillips theorem. We discuss them in the next intermezzo.

Intermezzo 2: Closable operators

Often one initially defines an operator on a subspace of “good vectors” to simplify calculations. It may happen that the operator is not closed, and one then looks for closed extensions. So we are led to the next definition.

Definition 4.7. *A linear operator A is called* closable *if it possesses a closed extension. Using part (c) of the next lemma, for a closable operator A we define its* closure $\overline{A}$ *by*

$$\begin{aligned}\mathrm{D}(\overline{A}) &:= \Big\{x \in X \,\big|\, \exists (x_n)_n \subseteq \mathrm{D}(A), y \in X : \lim_{n\to\infty} x_n = x, \lim_{n\to\infty} Ax_n = y\Big\}, \\ \overline{A}x &:= y, \quad \text{where } y \text{ is the vector from the definition of } \mathrm{D}(\overline{A}).\end{aligned} \tag{4.2}$$

Lemma 4.8. *For a linear operator A, the following assertions are equivalent.*

(a) *The operator A is closable.*
(b) *If $(x_n)_n \subseteq \mathrm{D}(A)$ tends to 0 in X and $(Ax_n)_n$ converges in X to some y in X, then $y = 0$ must hold.*
(c) *The definition (4.2) gives a closed extension of A.*

If this is the case, then $\overline{\mathrm{gr}(A)} = \mathrm{gr}(\overline{A})$, $\mathrm{D}(A)$ is dense in $[\mathrm{D}(\overline{A})]$ and $\overline{A}$ is the smallest closed extension of A. Moreover, the linear operator A is closed if and only if it is closable and $A = \overline{A}$.

Proof. "(a)⇒(b)": Let B be a closed extension of A. If the vectors $x_n \in \mathrm{D}(A) \subseteq \mathrm{D}(B)$ tend to 0 in X and their images $Ax_n = Bx_n$ converge to some y in X, the closedness of B implies that $y = B(0) = 0$.

"(b)⇒(c)": Let $(x_n)_n$ and $(z_n)_n$ be sequences in $\mathrm{D}(A)$ tending to x in X such that $(Ax_n)_n$ converges to y and $(Az_n)_n$ converges to w in X. Then $x_n - z_n$ belongs to $\mathrm{D}(A)$ as well as $x_n - z_n \to 0$ and $A(x_n - z_n) \to y - w$ in X as $n \to \infty$. We conclude $y = w$ using the assumption (b). Therefore (4.2) defines a map $\overline{A}$. One easily verifies that $\overline{A}$ is linear with $\overline{\mathrm{gr}(A)} = \mathrm{gr}(\overline{A})$, which shows the first part of the addendum. Hence, $\overline{A}$ is closed due to Remark 2.11 (a) and $\overline{A}$ extends A.

"(c)⇒(a)": This implication is obvious.

If B is another closed extension of A, we have $\mathrm{gr}(A) \subseteq \mathrm{gr}(B)$ and thus $\mathrm{gr}(\overline{A}) = \overline{\mathrm{gr}(A)} \subseteq \mathrm{gr}(B)$ because of the closedness of B, so that $\overline{A} \subseteq B$. The other assertions are an immediate consequence of $\overline{\mathrm{gr}(A)} = \mathrm{gr}(\overline{A})$, the definition of the graph norm and the closedness of $\overline{A}$. □

In general, it is a tricky business to compute the closure of a linear operator. We give some basic examples for illustration. They also highlight the crucial impact of the choice of boundary conditions for a fixed differential operator.

Example 4.9. (a) Let $X = L^1(0,1)$ and $Af = f(0)\mathbb{1}$ with $\mathrm{D}(A) = C([0,1])$. This operator is not closable. In fact, the functions $f_n(s) = \max\{1 - ns, 0\}$ satisfy $\|f_n\|_1 = \frac{1}{2n} \to 0$ as $n \to \infty$, but $Af_n = \mathbb{1}$ for all $n \in \mathbb{N}$, contradicting Lemma 4.8 (b). ♢

(b) Let $X = C([0,1])$ and $A_0 u = u'$ with $\mathrm{D}(A_0) = C_c^1(0,1)$. We claim that A_0 has the closure $A := \frac{\mathrm{d}}{\mathrm{ds}}$ with $\mathrm{D}(A) := C_0^1(0,1)$. It is clear that A is closed and that $A_0 \subseteq A$. Hence, A_0 is closable and $\overline{A_0} \subseteq A$. To show $A \subseteq \overline{A_0}$, we pick $f \in C_0^1(0,1)$ and check that $f \in \mathrm{D}(\overline{A_0})$. Take $\varphi_n \in C_c^1(0,1)$ with $\varphi = 1$ on $[\frac{1}{n}, 1 - \frac{1}{n}]$, $0 \le \varphi_n \le 1$ and $\|\varphi_n'\|_\infty \le cn$ for some $c > 0$ and all $n \in \mathbb{N}$ with $n \ge 2$. Set $J_n = [0, \frac{1}{n}] \cup [1 - \frac{1}{n}, 1]$. Then $f_n := \varphi_n f$ belongs to $\mathrm{D}(A_0)$, and we deduce

$$\|f_n - f\|_\infty = \sup_{s \in J_n} |(\varphi_n(s) - 1) f(s)| \le \sup_{s \in J_n} |f(s)| \longrightarrow 0,$$
$$\|\varphi_n f' - f'\|_\infty \le \sup_{s \in J_n} |f'(s)| \longrightarrow 0$$

as $n \to \infty$ since $f, f' \in C_0(0,1)$. We further obtain

$$\begin{aligned}\|\varphi_n' f\|_\infty &\le \sup_{0\le s\le \frac{1}{n}} |\varphi_n'(s)f(s)| + \sup_{1-\frac{1}{n}\le s\le 1} |\varphi_n'(s)f(s)| \\ &\le \sup_{0\le s\le \frac{1}{n}} cn\Big|\int_0^s f'(\tau)\,\mathrm{d}\tau\Big| + \sup_{1-\frac{1}{n}\le s\le 1} cn\Big|\int_s^1 f'(\tau)\,\mathrm{d}\tau\Big| \\ &\le cn\int_0^{\frac{1}{n}} |f'(\tau)|\,\mathrm{d}\tau + cn\int_{1-\frac{1}{n}}^1 |f'(\tau)|\,\mathrm{d}\tau \;\longrightarrow\; 0\end{aligned}$$

as $n \to \infty$, again because of $f, f' \in C_0(0,1)$. Hence, $A_0(\varphi_n f) = \varphi_n' f + \varphi_n f'$ converges to $Af = f'$. Consequently, $f \in \mathrm{D}(\overline{A_0})$ and thus $\overline{A_0} = A$.

There are further closed extensions of A_0. By Exercise 4.3, $A_1 = \frac{\mathrm{d}}{\mathrm{ds}}$ with $\mathrm{D}(A_1) = \{u \in C^1([0,1]) \,|\, u'(1) = 0\}$ generates a C_0–semigroup on X and $\sigma(A_1) = \{0\}$. Moreover, $A_3 = \frac{\mathrm{d}}{\mathrm{ds}}$ with $\mathrm{D}(A_3) = C^1([0,1])$ has the spectrum $\sigma(A_3) = \mathbb{C}$ by Example 3.5. Clearly, $A_0 \subsetneqq \overline{A_0} = A \subsetneqq A_1 \subsetneqq A_3$. We claim that A is not a generator. In fact, Lemma 3.6 and $A \subsetneqq A_1$ yield $\rho(A) \cap \rho(A_1) = \emptyset$. Since $Au \neq \mathbb{1}$ for all $u \in C_0^1(0,1)$, A is not surjective and so $0 \in \sigma(A)$. It follows that $\rho(A) = \emptyset$, implying the claim.

On the other hand, also $A_2 = \frac{\mathrm{d}}{\mathrm{ds}}$ with $\mathrm{D}(A_2) = \{u \in C^1([0,1]) \,|\, u(0) = 0\}$ is not a generator on X (because $\overline{\mathrm{D}(A_2)} \neq X$), and we have $A \subsetneqq A_2 \subsetneqq A_3$. Moreover, A_1 and A_2 are not comparable.

Summing up, here the "minimal" operator A and the "maximal" operator A_3 do not generate C_0–semigroups. Between them there are various, partly noncomparable operators with differing boundary conditions (so–called "realizations" of $\frac{\mathrm{d}}{\mathrm{ds}}$) which may or may not be generators. ◊

We now use the concept of a closure in the context of generators A showing a criterion for "cores" $\mathcal{D}$ of A needed later. A *core* is a linear subspace of $\mathrm{D}(A)$ which is dense for the graph norm of A. It is easy to see that for a closable operator A and $\mathcal{D} \subseteq \mathrm{D}(A)$ one has $\overline{A|_{\mathcal{D}}} = \overline{A}$ if and only if $\mathcal{D}$ is a core of A (and then also for $\overline{A}$). In this sense, A is determined by its restriction to a core.

Proposition 4.10. *Let A generate the C_0–semigroup $T(\cdot)$ on X. Let $\mathcal{D} \subseteq \mathrm{D}(A)$ be a linear subspace which is dense in X and satisfies $T(t)\mathcal{D} \subseteq \mathcal{D}$ for all $t \ge 0$. Then $\mathcal{D}$ is a core for A.*

Proof. Let $x \in \mathrm{D}(A)$. By Proposition 3.3, the function $T(\cdot)x$ belongs to $C(\mathbb{R}_+, [\mathrm{D}(A)])$. For each $n \in \mathbb{N}$ we can thus fix $\tau = \tau(n,x) \in (0,1]$ such that

$$\Big\|\frac{1}{\tau}\int_0^\tau T(t)x\,\mathrm{d}t - x\Big\|_A \le \frac{1}{2n}.$$

Observe that the above Riemann integral coincides with the Riemann integral of the function $T(\cdot)x : \mathbb{R}_+ \to (X, \|\cdot\|)$ because $\|y\| \le \|y\|_A$ holds for all $y \in \mathrm{D}(A)$ and thus the respective Riemann sums have the same limit.

Since $\mathcal{D}$ is dense in X, there is a sequence $(y_m)_m$ in $\mathcal{D}$ converging to x in X. We define $x_m := \frac{1}{\tau}\int_0^\tau T(t)y_m\,\mathrm{d}t$ for $m \in \mathbb{N}$. Due to $T(t)y_m \in \mathcal{D} \subseteq \mathrm{D}(A)$ for all $t \ge 0$, the vector x_m belongs to the closure $\overline{\mathcal{D}}^A$ of $\mathcal{D}$ in $[\mathrm{D}(A)]$. We have

$\sup_{t\in[0,1]} \|T(t)\| =: C < \infty$. Employing Lemma 3.2, we can now estimate

$$\begin{aligned}
\|x - x_m\|_A &\le \Big\|x - \frac{1}{\tau}\int_0^\tau T(t)x \,\mathrm{d}t\Big\|_A \\
&\qquad + \frac{1}{\tau}\Big\|\int_0^\tau T(t)(x-y_m)\,\mathrm{d}t\Big\| + \frac{1}{\tau}\Big\|A\int_0^\tau T(t)(x-y_m)\,\mathrm{d}t\Big\| \\
&\le \tfrac{1}{2n} + C\|x-y_m\| + \tfrac{1}{\tau}\|(T(\tau)-I)(x-y_m)\| \\
&\le \tfrac{1}{2n} + (C + \tfrac{1}{\tau}(1+C))\|x-y_m\|.
\end{aligned}$$

It follows that $\limsup_{m\to\infty} \|x - x_m\|_A \le \frac{1}{2n}$. Since $x_m \in \overline{\mathcal{D}}^A$, for each $n \in \mathbb{N}$ we can now find a vector $z_n \in \mathcal{D}$ such that $\|x - z_n\|_A \le \frac{1}{n}$. □

We now come back to the main theme of this lecture and prove several properties of dissipative operators including their closability.

Proposition 4.11. *Let A be dissipative. Then the following assertions hold.*

(a) For all $\lambda > 0$ the operator $\lambda I - A$ is injective and for $y \in \mathrm{R}(\lambda I - A) = (\lambda I - A)(\mathrm{D}(A))$ we have $\|(\lambda I - A)^{-1}y\| \le \frac{1}{\lambda}\|y\|$.

(b) Let $\lambda_0 I - A$ be surjective for some $\lambda_0 > 0$. Then A is closed, $(0,\infty) \subseteq \rho(A)$ and $\|R(\lambda, A)\| \le \frac{1}{\lambda}$ for all $\lambda > 0$.

(c) If $\mathrm{D}(A)$ is dense in X, then A is closable and $\overline{A}$ is also dissipative.

Proof. Assertion (a) immediately follows from Proposition 4.5.

(b) If the assumptions in (b) hold, then part (a) implies that $\lambda_0 I - A$ has an inverse with a norm less than or equal to $\frac{1}{\lambda_0}$; in particular, A is closed. For any $\lambda \in (0, 2\lambda_0)$, it holds $|\lambda - \lambda_0| < \lambda_0 \le \|R(\lambda_0, A)\|^{-1}$ and hence $\lambda \in \rho(A)$ due to (2.8). This argument shows that we have $(0, 2\lambda) \subseteq \rho(A)$ and thus $(0, \frac{3}{2}\lambda] \subseteq \rho(A)$, whenever $\lambda \in \rho(A)$ is fulfilled. As a consequence, one inductively obtains $(0, (\frac{3}{2})^n \lambda_0] \subseteq \rho(A)$ for all $n \in \mathbb{N}$ arriving at $(0,\infty) \subseteq \rho(A)$. The asserted estimate for $R(\lambda, A)$ is an immediate consequence of part (a).

(c) Assume that $\overline{\mathrm{D}(A)} = X$. We want to prove the closability of A by means of Lemma 4.8. Let $(x_n)_n$ be a sequence in $\mathrm{D}(A)$ such that $x_n \to 0$ and $Ax_n \to y$ in X as $n \to \infty$ for some $y \in X$. Due to the density assumption, there is a sequence $(y_n)_n$ in $\mathrm{D}(A)$ converging to y in X. Take any $\lambda > 0$ and $n, m \in \mathbb{N}$. Proposition 4.5 then implies

$$\|\lambda(\lambda I - A)x_n + (\lambda I - A)y_m\| = \|(\lambda I - A)(\lambda x_n + y_m)\| \ge \lambda\|\lambda x_n + y_m\|$$

Letting $n \to \infty$, we deduce $\|-\lambda y + (\lambda I - A)y_m\| \ge \lambda\|y_m\|$, or equivalently,

$$\|(I - \tfrac{1}{\lambda}A)y_m - y\| \ge \|y_m\|.$$

As $\lambda \to \infty$, it follows that $\|y_m - y\| \ge \|y_m\|$. Taking the limit $m \to \infty$, we conclude $y = 0$. Lemma 4.8 then shows that the operator A is closable and that for $x \in \mathrm{D}(\overline{A})$ there are $z_n \in \mathrm{D}(A)$ such that $\lim_{n\to\infty} z_n = x$ and $\lim_{n\to\infty} Az_n = \overline{A}x$ in X. From Proposition 4.5 we can now infer the estimate

$$\|\lambda x - \overline{A}x\| = \lim_{n\to\infty} \|\lambda z_n - Az_n\| \ge \lambda \lim_{n\to\infty}\|z_n\| = \lambda\|x\|,$$

and thus the dissipativity of $\overline{A}$. □

Propositions 4.5 and 4.11 and the Hille-Yosida theorem now imply the Lumer-Phillips theorem. It characterizes generators of contraction semigroups by the density of their domain, their dissipativity and a range condition. In applications one often uses part (a) for closed A, where $A = \overline{A}$.

THEOREM 4.12 (Lumer-Phillips, 1961). *Let A be linear and densely defined. Then the following assertions hold.*

(a) *If A is dissipative and $\lambda_0 I - A$ has dense range for some $\lambda_0 > 0$, then A is closable and $\overline{A}$ generates a contraction semigroup.*
(b) *If A is dissipative and $\lambda_0 I - A$ is surjective for some $\lambda_0 > 0$, then A generates a contraction semigroup.*
(c) *If A generates a contraction semigroup, then A is dissipative, $\mathbb{C}_+ \subseteq \rho(A)$ and $\|R(\lambda, A)\| \le \frac{1}{\operatorname{Re}\lambda}$ for all $\lambda \in \mathbb{C}_+$.*

PROOF. (a) Proposition 4.11 tells us that A possesses a dissipative closure. Moreover, the range of $\lambda_0 I - \overline{A}$ contains that of $\lambda_0 I - A$ and it is thus dense. For a given $y \in X$, we can now find elements x_n of $\mathrm{D}(\overline{A})$ such that $y_n := \lambda_0 x_n - \overline{A} x_n$ converge to y in X as $n \to \infty$. Since $\overline{A}$ is dissipative, we have

$$\|x_n - x_m\| \le \frac{1}{\lambda_0} \|(\lambda_0 I - \overline{A})(x_n - x_m)\| = \frac{1}{\lambda_0} \|y_n - y_m\|$$

for all $n, m \in \mathbb{N}$ thanks to Proposition 4.5. Therefore $(x_n)_n$ has a limit x in X. Hence, the vectors $\overline{A} x_n = \lambda_0 x_n - y_n$ tend to $\lambda_0 x - y$ as $n \to \infty$. Due to the closedness of $\overline{A}$, the vector x belongs to $\mathrm{D}(\overline{A})$ and $\overline{A} x = \lambda_0 x - y$ so that $\lambda_0 I - \overline{A}$ is surjective. Proposition 4.11 and Theorem 3.8 now imply assertion (a).

(b) Proposition 4.11 says that A is closed if $\lambda_0 I - A$ is surjective. Part (a) now shows that $A = \overline{A}$ generates a contraction semigroup.

(c) This assertion can be deduced from Propositions 4.5 and 3.4. □

We conclude with a simple example for Theorem 4.12, namely the second derivative with Neumann boundary conditions on $C([0,1])$. More examples and further discussions are given in the next lecture.

EXAMPLE 4.13. Let $X = C([0,1])$. We want to show that $Au = u''$ with $\mathrm{D}(A) = \{u \in C^2([0,1]) \,|\, u'(0) = u'(1) = 0\}$ satisfies the conditions of the Lumer-Phillips theorem and thus generates a C_0–semigroup on X.

To see $\overline{\mathrm{D}(A)} = X$, let $f \in C([0,1])$. Fix a function $\varphi \in C^2([0,1])$ with $\varphi(0) = 1$ and $\varphi'(0) = 0$ and $\operatorname{supp}\varphi \subseteq [0, \frac{1}{2})$. Set $\psi(t) = \varphi(1-t)$ for $t \in [0,1]$ and $g = f - \varphi f(0) - \psi f(1) \in C_0(0,1)$. As in Example 3.9 one obtains functions $v_n \in C_c^\infty(0,1)$ that converge to g uniformly. Hence, the maps $u_n := v_n + \varphi f(0) + \psi f(1) \in \mathrm{D}(A)$ tend to f in X as $n \to \infty$.

To show the dissipativity, we proceed as in Example 4.6. Let $u \in \mathrm{D}(A)$ and fix a point $s_0 \in [0,1]$ such that $|u(s_0)| = \|u\|_\infty$. Then the functional $\varphi(f) = \overline{u}(s_0) f(s_0)$ belongs to $J(u)$. Moreover, the function $h = \operatorname{Re}(\overline{u}(s_0)u) \in C^2([0,1])$ attains its maximum at s_0 and $\operatorname{Re}\langle Au, \varphi\rangle = \operatorname{Re} u''(s_0)\overline{u}(s_0) = h''(s_0)$. Hence, $\operatorname{Re}\langle Au, \varphi\rangle \le 0$ if $s_0 \in (0,1)$. If $s_0 = 0$, then for each $s > 0$ Taylor's formula gives a number $\sigma \in [0,s]$ such that $h(s) = h(0) + \frac{s^2}{2} h''(\sigma)$ since $u \in \mathrm{D}(A)$ yields $h'(0) = 0$. Thus, $h''(\sigma) = \frac{2}{s^2}(h(s) - h(0)) \le 0$. Letting $s \to 0^+$, we deduce $\operatorname{Re}\langle Au, \varphi\rangle = h''(0) \le 0$. The case $s_0 = 1$ is treated in the same way.

For the surjectivity of $I - A$, take $f \in C([0,1])$ and set

$$u(s) = a\mathrm{e}^{s} + b\mathrm{e}^{-s} + \frac{1}{2}\int_0^1 \mathrm{e}^{-|s-\tau|} f(\tau)\,\mathrm{d}\tau \tag{4.3}$$

for $s \in [0,1]$ and constants $a, b \in \mathbb{C}$ to be determined, depending on f. A straightforward computation shows that $u \in C^2([0,1])$, $u - u'' = f$ and

$$u'(0) = a - b + \frac{1}{2}\int_0^1 \mathrm{e}^{-\tau} f(\tau)\,\mathrm{d}\tau, \qquad u'(1) = \mathrm{e}a - \tfrac{1}{\mathrm{e}}b - \frac{1}{2\mathrm{e}}\int_0^1 \mathrm{e}^{\tau} f(\tau)\,\mathrm{d}\tau.$$

There are $a, b \in \mathbb{C}$ with $u'(0) = u'(1) = 0$, i.e., $u \in \mathrm{D}(A)$ and $u - Au = f$. $\diamondsuit$

A more detailed analysis of such differential operators on intervals (including a motivation of the ansatz (4.3)) can be found in Section VI.4 of [**EN99**].

Exercises

Exercise 4.1. Let $T(t) \in \mathcal{B}(X)$, $t \geq 0$, have locally bounded norms such that $T(0) = I$ and $T(t+s) = T(t)T(s)$ for $t, s \geq 0$. We set

$$X_0 = \{x \in X \,|\, T(t)x \to x \text{ as } t \to 0\}.$$

Show that X_0 is a closed subspace of X, that $T(t)X_0 \subseteq X_0$ for $t \geq 0$ and that the restrictions $T_0(t) := T(t)|_{X_0}$, $t \geq 0$, form a C_0–semigroup on X_0.

Exercise 4.2. Let X be a Banach space and $x \in X$. Show that

$$J(x) = \Big\{x^* \in X^* \,|\, \langle x, x^*\rangle = \|x\|^2 = \|x^*\|^2\Big\}$$

is not empty.

A Banach space Y is called *strictly convex* if for all $x, y \in X$ with $\|x\| = \|y\| = 1$ and $\|x + y\| = 2$ we have $x = y$. Show that $J(x)$ contains exactly one element for each $x \in X$ if X^* is strictly convex. (Remark: L^p spaces are strictly convex for $p \in (1, \infty)$, see (15.5) and (15.8) in [**HS65**].)

Exercise 4.3. Let $X = C([0,1])$. Define $Au = u'$ with $\mathrm{D}(A) = \{u \in C^1([0,1]) \,|\, u'(1) = 0\}$. Show that A is dissipative, $\sigma(A) = \{0\}$ and that A generates a contraction semigroup. Determine this semigroup.

Exercise 4.4. Let $F \in C^1(\mathbb{R}^d, \mathbb{R}^d)$ have a bounded derivative. It is known that there is a function $\varphi \in C^1(\mathbb{R} \times \mathbb{R}^d, \mathbb{R}^d)$ such that

$$\partial_t \varphi(t,x) = F(\varphi(t,x)) \quad \text{for } (t,x) \in \mathbb{R} \times \mathbb{R}^d,$$

$\varphi(t, \varphi(s,x)) = \varphi(t+s, x)$ and $\varphi(0,x) = x$ for all $s, t \in \mathbb{R}$ and $x \in \mathbb{R}^d$. Set $E = C_0(\mathbb{R}^d)$ and

$$(T(t)f)(x) = f(\varphi(t,x)) \quad \text{for } (t,x) \in \mathbb{R} \times \mathbb{R}^d \text{ and } f \in E.$$

Show that $T(\cdot)$ is a contractive C_0–group on E with generator A, where A is the closure of A_0 given by $(A_0u)(x) = F(x) \cdot \nabla u(x)$ and $\mathrm{D}(A_0) = C_c^1(\mathbb{R}^d)$.

LECTURE 5

Stone's theorem and the Laplacian

In the previous lecture we have shown the Lumer-Phillips Theorem 4.12, which says among other things that

if A is dissipative, $\overline{\mathrm{D}(A)} = X$ and $\mathrm{R}(\lambda_0 I - A) = X$ for some $\lambda_0 > 0$, then A generates a contraction semigroup.

In this lecture we first establish several variants of this result and deduce a characterization of generators of isometric groups. This characterization then implies Stone's famous theorem stating that precisely the skewadjoint operators generate unitary C_0–groups on Hilbert spaces. In the remainder of the lecture we then discuss the Laplacian in two simple settings: in $L^2(\mathbb{R}^d)$ and in $L^2(U)$ with Dirichlet boundary conditions, where $U \subseteq \mathbb{R}^d$ is open and bounded. We show that these "realizations" of the Laplacian are dissipative and selfadjoint (and thus generate C_0–semigroups).

These facts play a crucial role in our main applications to wave and Schrödinger equations. Here the Hilbert space setting arises naturally in view of the physical background. It also allows to establish the desired properties of the Laplacian by elegant functional analytic tools. This approach relies on Sobolev spaces and weak derivatives. Actually, we do not need many deep results about these spaces. The relevant information is recalled below in an intermezzo. We also provide a rather long appendix containing the proofs and an introduction to this large subject. We further use basic facts about the Fourier transform and selfadjoint operators which are collected and proved in two more appendices. Again we list the relevant results in the intermezzo.

We first come back to the Lumer-Phillips Theorem 4.12 and discuss the range condition via duality theory. To this aim, for a linear operator A in X with dense domain, we define its *adjoint* A^* by

$$\begin{aligned} A^*x^* &:= y^* \quad \text{for all } x^* \in \mathrm{D}(A^*), \text{ where } y^* \text{ is taken from} \\ \mathrm{D}(A^*) &:= \big\{x^* \in X^* \,\big|\, \exists\, y^* \in X^* \;\forall\, x \in \mathrm{D}(A) : \langle Ax, x^*\rangle = \langle x, y^*\rangle\big\}\,. \end{aligned} \tag{5.1}$$

This means that $\langle Ax, x^*\rangle = \langle x, A^*x^*\rangle$ for all $x \in \mathrm{D}(A)$ and all $x^* \in \mathrm{D}(A^*)$, and $\mathrm{D}(A^*)$ consists of all $x^* \in X^*$ for which this operation works. It turns out that this is the correct definition to develop the theory. To get some experience with adjoints, we discuss several basic properties of adjoints in the next remark.

Remark 5.1. Let A be a densely defined linear operator in X.

(a) Since $\mathrm{D}(A)$ is dense, there is at most one vector $y^* = A^*x^*$ as in (5.1), so that $A^* : \mathrm{D}(A^*) \to X^*$ is a map. It is clear that A^* is linear. If $A \in \mathcal{B}(X)$, then $\mathrm{D}(A^*) = X^*$ and (5.1) coincides with the definition of A^* usually given in courses on functional analysis. ◊

(b) The operator A^* is closed in X^* (though A does not need to be closed). In fact, let $x_n^* \in \mathrm{D}(A^*)$, $x^* \in X^*$, and $y^* \in X^*$ such that $x_n^* \to x^*$ and $A^*x_n^* \to y^*$ in X^* as $n \to \infty$. For every $x \in \mathrm{D}(A)$ we then compute

$$\langle x, y^* \rangle = \lim_{n\to\infty} \langle x, A^*x_n^* \rangle = \lim_{n\to\infty} \langle Ax, x_n^* \rangle = \langle Ax, x^* \rangle.$$

As a result, $x^* \in \mathrm{D}(A^*)$ and $A^*x^* = y^*$. ◊

(c) If $T \in \mathcal{B}(X)$, then the sum $A+T$ with $\mathrm{D}(A+T) = \mathrm{D}(A)$ has the adjoint $(A+T)^* = A^* + T^*$ with $\mathrm{D}((A+T)^*) = \mathrm{D}(A^*)$. To verify this fact, let $x \in \mathrm{D}(A)$ and $x^* \in X^*$. We obtain

$$\langle (A+T)x, x^* \rangle = \langle Ax, x^* \rangle + \langle x, T^*x^* \rangle.$$

Hence, $x^* \in \mathrm{D}((A+T)^*)$ if and only if $x^* \in \mathrm{D}(A^*)$, and then $(A+T)^*x^* = A^*x^* + T^*x^*$. ◊

(d) If $T \in \mathcal{B}(X)$, then the product TA with domain $\mathrm{D}(TA) = \mathrm{D}(A)$ has the adjoint A^*T^* with domain $\mathrm{D}(A^*T^*) = \{x^* \in X^* \,|\, T^*x^* \in \mathrm{D}(A^*)\}$. This fact can be shown analogously as part (c). From (c) and (d) it follows that $(\lambda I - A)^* = \lambda I - A^*$ for $\lambda \in \mathbb{C}$. ◊

The next result allows to replace the range condition in the Lumer-Phillips theorem by the injectivity of $\lambda I - A^*$ for some $\lambda > 0$. *If* one knows the adjoint of A, the injectivity should be much easier to check than the range condition. However, in applications it is often very hard to compute A^* (namely $\mathrm{D}(A^*)$) directly. Nevertheless, Corollary 5.2 is used in the proof of Stone's theorem.

Corollary 5.2. *Let A be dissipative and densely defined. If $\lambda I - A^*$ is injective for some $\lambda > 0$, then $\overline{A}$ generates a contraction semigroup. (By Proposition 4.11, $\lambda I - A^*$ is injective for all $\lambda > 0$ if A^* is dissipative.)*

Proof. Due to Theorem 4.12, it suffices to show that $\lambda I - A$ has dense range. Let $x^* \in X^*$ annihilate $\mathrm{R}(\lambda I - A)$; i.e., $\langle (\lambda I - A)x, x^* \rangle = 0 = \langle x, 0 \rangle$ for all $x \in \mathrm{D}(A)$. This fact leads to $x^* \in \mathrm{D}(A^*)$ and $\lambda x^* - A^*x^* = 0$. The injectivity assumption now yields $x^* = 0$. A corollary to the Hahn-Banach theorem then implies the density of $\mathrm{R}(\lambda I - A)$. □

Usually the density of the domain of an operator is relatively easy to check. Still it is a nice fact that one gets it for free in the context of the Hille-Yosida theorem if X is reflexive.

Proposition 5.3. *Let X be a reflexive Banach space and A be a closed operator in X such that $(\omega, \infty) \subseteq \rho(A)$ and $\|R(\lambda, A)\| \le \frac{M}{\lambda - \omega}$ for all $\lambda > \omega$ and some constants $M, \omega \ge 0$. Then $\mathrm{D}(A)$ is dense in X. In particular, if A is a dissipative operator on a reflexive Banach space such that $\lambda_0 I - A$ is surjective for some $\lambda_0 > 0$, then A generates a contraction semigroup.*

Proof. Let $x \in X$ and $n \in \mathbb{N}$ with $n > \omega$. Then the vectors $x_n = nR(n, A)x$ belong to $\mathrm{D}(A)$ and are uniformly bounded in X. Since X is reflexive, the Banach-Alaoglu theorem gives a subsequence $(x_{n_j})_j$ and a vector $z \in X$ such that x_{n_j} converges weakly to z as $j \to \infty$. Since $\overline{\mathrm{D}(A)}$ is closed and convex, we obtain $z \in \overline{\mathrm{D}(A)}$, see e.g. Theorem 3.7 in [**Bre11**]. We now show $x = z$ to deduce $\overline{\mathrm{D}(A)} = X$. We first note that the vectors $y_n = R(\omega + 1, A)x_n =$

$nR(n,A)R(\omega+1,A)x$ also belong to $\mathrm{D}(A)$ and converge to $R(\omega+1,A)x =: y$ in X as $n\to\infty$ due to Lemma 3.7 (because of $y\in \mathrm{D}(A)$). We thus obtain the weak limits

$$y_{n_j} \rightharpoonup y \quad \text{and} \quad Ay_{n_j} = (1+\omega)y_{n_j} - x_{n_j} \rightharpoonup (1+\omega)y - z =: v$$

as $j\to\infty$. As above, we infer that (y,v) is contained in $\overline{\mathrm{gr}(A)} = \mathrm{gr}(A)$ so that $Ay = v$. Consequently, $x = (1+\omega)y - Ay = z$ belongs to $\overline{\mathrm{D}(A)}$.

If A is dissipative and $\mathrm{R}(\lambda_0 I - A) = X$ for some $\lambda_0 > 0$, then A is closed, $(0,\infty)\subseteq \rho(A)$ and $\|\lambda R(\lambda,A)\|\le 1$ for all $\lambda>0$ due to Proposition 4.11. The first part thus yields that $\overline{\mathrm{D}(A)} = X$ and the second assertion now follows from the Hille-Yosida Theorem 3.8. □

We next use the Lumer-Phillips theorem to characterize the generators of isometric C_0–groups.

Corollary 5.4. *Let A be linear. The following assertions are equivalent.*

- *(a) The operator A generates an isometric C_0–group $T(\cdot)$, i.e., $\|T(t)x\| = \|x\|$ for all $x\in X$ and $t\in\mathbb{R}$.*
- *(b) The operator A is closed and densely defined, A and $-A$ are dissipative and $\lambda I - A$ and $\lambda I + A$ are surjective for some $\lambda > 0$.*
- *(c) The operator A is closed and densely defined, $\mathbb{R}\setminus\{0\}\subseteq\rho(A)$ and $\|R(\lambda,A)\|\le \frac{1}{|\lambda|}$ for all $\lambda\in\mathbb{R}\setminus\{0\}$.*

Proof. The Lumer-Phillips Theorem 4.12 says that (b) holds if and only if A and $-A$ generate contraction semigroups. Theorem 4.2 then implies the equivalence of assertion (b) and (c), and it shows that (b) holds if and only if A generates a C_0–group of contractions $T(t)$, $t\in\mathbb{R}$. It remains to prove that a contractive C_0–group $T(\cdot)$ is already isometric. Indeed, we have

$$\|T(t)x\| \le \|x\| = \|T(-t)T(t)x\| \le \|T(-t)\|\,\|T(t)x\| \le \|T(t)x\|$$

for all $x\in X$ and all $t\in\mathbb{R}$, so that $T(t)$ is isometric. □

In the Hilbert space setting, the above result leads to Stone's theorem on unitary groups. To this aim, we have to recall a few more concepts.

Definition 5.5. *Let X be a Hilbert space with scalar product $(\cdot|\cdot)$. A linear operator A on X is called* symmetric *if*

$$\forall\, x,y\in \mathrm{D}(A):\ (Ax|y) = (x|Ay).$$

Let A be a densely defined linear operator on X. Then its Hilbert space adjoint *A' is given by*

$$A'y := z \quad \text{for all } y\in \mathrm{D}(A'), \text{ where } z \text{ is taken from}$$
$$\mathrm{D}(A') := \{y\in X \,|\, \exists\, z\in X\ \forall\, x\in \mathrm{D}(A) : (Ax|y) = (x|z)\},$$

cf. (5.1). A densely defined linear operator A is called selfadjoint *if $A = A'$ and* skewadjoint *if $A' = -A$. Finally, an operator $T\in\mathcal{B}(X)$ is called* unitary *if it is invertible with $T^{-1} = T'\in\mathcal{B}(X)$.*

Besides A' there also exists the adjoint A^* in X^*. These two operators are related via Riesz' representation theorem, cf. (C.1) in Appendix C.

Observe that a densely defined linear operator A is symmetric if and only if $A \subseteq A'$. Moreover, A is skewadjoint if and only if $\mathrm{i}A$ is selfadjoint. We point out that selfadjointness and skewadjointness mean in particular that $\mathrm{D}(A) = \mathrm{D}(A')$. We further remark that symmetry is relatively easy to check in many examples, whereas selfadjointness or skewadjointness (mainly the equality $\mathrm{D}(A) = \mathrm{D}(A')$) often is very hard to verify. On the other hand, these properties have far reaching consequences as in Stone's theorem below.

We collect several properties of the Hilbert space adjoints and of selfadjoint operators in the next remark, see Appendix C for more details. Property (f) is used below to prove the selfadjointness of the Dirichlet Laplacian.

Remark 5.6. Let A be a densely defined operator in a Hilbert space X. Then the following assertions hold.

(a) As in Remark 5.1 one sees that A' is closed in X. In particular, a symmetric, densely defined, linear operator is closable with $\overline{A} \subseteq A'$ (since $A \subseteq A'$).

(b) Parts (c) and (d) of Remark 5.1 hold for A' in a similar way. But note that $(aI)' = \overline{a}I$ for $a \in \mathbb{C}$.

(c) Let A be closed. Then $\sigma(A') = \{\overline{\lambda} \mid \lambda \in \sigma(A)\}$ by (C.2) in Appendix C.

(d) If A is symmetric, then also $\overline{A}$ is symmetric. In fact, for $x, y \in \mathrm{D}(\overline{A})$ there are $x_n, y_n \in \mathrm{D}(A)$ with $x_n \to x$, $y_n \to y$, $Ax_n \to \overline{A}x$ and $Ay_n \to \overline{A}y$ in X as $n \to \infty$. It follows

$$(\overline{A}x|y) = \lim_{n\to\infty}(Ax_n|y_n) = \lim_{n\to\infty}(x_n|Ay_n) = (x|\overline{A}y).$$

(e) There are symmetric closed operators which are not selfadjoint, see Example C.10 in Appendix C.

(f) Let A be closed and symmetric. Then A is selfadjoint if and only if $\sigma(A) \subseteq \mathbb{R}$. Moreover, if $\rho(A) \cap \mathbb{R} \neq \emptyset$, then $\sigma(A) \subseteq \mathbb{R}$. See Theorem C.9.

We can now derive Stone's theorem from the Lumer-Phillips theorem. Alternatively Stone's theorem also follows from the spectral theorem, see Section VII.4 in [**RS72**]. We will use Stone's theorem to solve the linear Schrödinger equation in a later lecture. In fact, this theorem is a cornerstone of the mathematical foundations of quantum mechanics.

Theorem 5.7 (Stone, 1930). *Let X be a Hilbert space and A be a linear operator on X with dense domain. Then A generates a C_0–group of unitary operators if and only if A is skewadjoint.*

Proof. 1) Let $A' = -A$. For $x \in \mathrm{D}(A) = \mathrm{D}(A')$, we have $J(x) = \{\varphi_x\}$ where $\varphi_x := (\,\cdot\,|x)$ (see Example 4.6). We compute

$$\langle Ax, \varphi_x \rangle = (Ax|x) = -(x|Ax) = -\overline{(Ax|x)} = -\overline{\langle Ax, \varphi_x \rangle},$$

so that $\mathrm{Re}\langle Ax, \varphi_x \rangle = 0$. Hence, A and $A' = -A$ are dissipative as well as $A'' = (-A)' = A$. We note that for a Hilbert space X, Corollary 5.2 holds with the same proof if one replaces A^* by A'. Hence, A and A' generate contraction semigroups, so that A generates a C_0–group $(T(t))_{t\in\mathbb{R}}$ of invertible isometric operators due to Corollary 5.4. A result from functional analysis then implies that each $T(t)$ is unitary (see Proposition C.7 in Appendix C).

2) Let A generate a unitary C_0–group $(T(t))_{t\in\mathbb{R}}$. Since $T(t)' = T(t)^{-1} = T(-t)$ for $t \geq 0$, the family $(T(t)')_{t\geq 0}$ is a contraction semigroup with the generator $-A$ by Theorem 4.2. For $x, y \in \mathrm{D}(A)$ we thus obtain

$$(Ax|y) = \lim_{t\to 0}(\tfrac{1}{t}(T(t)x - x)|y) = \lim_{t\to 0}(x|\tfrac{1}{t}(T(t)'y - y)) = (x| - Ay),$$

i.e., $-A \subseteq A'$. We further know from Theorem 4.2 that $\sigma(A) \subseteq \mathrm{i}\mathbb{R}$. Since $\sigma(A') = \{\overline{\lambda} \mid \lambda \in \sigma(A)\} \subseteq \mathrm{i}\mathbb{R}$ by Remark 5.6 (c), Lemma 3.6 yields $-A = A'$. $\square$

In our main applications the Laplacian in an L^2–setting plays a crucial role. To investigate this operator, we need weak derivatives, Sobolev spaces and the Fourier transform. These topics are discussed in the next intermezzo. The proofs and much more background material can be found in the corresponding appendices D and E.

Intermezzo 3: Weak derivatives, Sobolev spaces and the Fourier transform

The classical derivative does not fit well to L^p spaces since it is based on a pointwise limit. Instead, one uses on L^p spaces the so called weak derivative. In its definition one requires that one can integrate by parts against functions $\varphi \in C_c^\infty$, which is well adapted to integrable functions.

Let $\emptyset \neq U \subseteq \mathbb{R}^d$ be open, $k \in \mathbb{N}$ and $p \in [1, \infty]$. A function $u \in L^p(U)$ has a *weak derivative* v in $L^p(U)$ with respect to the jth coordinate for some $j \in \{1, \dots, d\}$ if there is a function $v \in L^p(U)$ such that

$$\int_U u\partial_j\varphi \,\mathrm{d}x = -\int_U v\varphi \,\mathrm{d}x \tag{5.2}$$

for all $\varphi \in C_c^\infty(U)$. The function v is uniquely determined (up to a null function) due to Lemma D.5 in Appendix D, and we put $\partial_j u := v$. The set $W_p^1(U)$ of all functions u in $L^p(U)$ possessing weak derivatives in $L^p(U)$ with respect to all coordinates is called a *Sobolev space*. The linear space $W_p^1(U)$ becomes a Banach space when endowed with the norm

$$\|u\|_{1,p} := \begin{cases} \Big(\|u\|_p^p + \sum_{j=1}^d \|\partial_j u\|_p^p\Big)^{1/p}, & \text{if } p < \infty, \\ \max_{j=1,\dots,d}\{\|u\|_\infty, \|\partial_j u\|_\infty\}, & \text{if } p = \infty, \end{cases}$$

see Proposition D.3. Here, as usual in the L^p context, we identify functions which are equal almost everywhere. For each $p \in [1, \infty]$, this norm is equivalent to the norm given by

$$\|u\|_p + \sum_{j=1}^d \|\partial_j u\|_p.$$

Weak derivatives of higher order and the Sobolev spaces $W_p^k(U)$ are defined analogously. We set $W_p^0(U) = L^p(U)$ and $H^k(U) = W_2^k(U)$. Observe that (5.2) leads to

$$\langle u, \partial_j\varphi\rangle_{L^p} = \int_U u\partial_j\varphi \,\mathrm{d}x = -\int_U (\partial_j u)\varphi \,\mathrm{d}x = -\langle \partial_j u, \varphi\rangle_{L^p}$$

for all $u \in W_p^1(U)$ and $\varphi \in C_c^\infty(U)$. This definition via duality already allows to deduce various properties of weak derivatives (e.g. their linearity). Other properties follow by approximation. (The corresponding density results are proved via “cut-off” functions and “mollifiers”, see Appendix D.)

In the next remark we summarize the properties of weak derivatives and Sobolev spaces needed later on. Roughly speaking, many of the simple facts about (classical) derivatives can be extended to weak ones if properly translated into the L^p setting. But, of course, there are several new phenomena also discussed below.

REMARK 5.8. Let $U \subseteq \mathbb{R}^d$ be open, $k \in \mathbb{N}$ and $p \in [1, \infty]$. Then the following assertions hold.

(a) If $u \in C^k(U)$ and u and all its derivatives up to order k belong to $L^p(U)$, then $u \in W_p^k(U)$ and the classical and the weak derivatives coincide (see Remark D.2).

(b) Let $p < \infty$. A function $u \in L^p(U)$ belongs to $W_p^k(U)$ if and only if there are $u_n \in C^k(U) \cap W_p^k(U)$ such that $u_n \to u$ in $L^p(U)$ and all derivatives of u_n up to order k converge in $L^p(U)$. We then have $\partial_j u = \lim_{n\to\infty} \partial_j u_n$ in $L^p(U)$ and analogously for higher derivatives up to order k (see Lemma D.6 and Theorem D.13).

(c) The space $C_c^\infty(\mathbb{R}^d)$ is dense in $W_p^k(\mathbb{R}^d)$ if $p < \infty$ (see Theorem D.13).

(d) A function $u \in L^p(a, b)$ (where $-\infty \le a < b \le \infty$) belongs to $W_p^1(a, b) := W_p^1((a, b))$ if and only if u has a continuous representative and there is a function $v \in L^p(a, b)$ such that

$$u(t) = u(s) + \int_s^t v(\tau)\,\mathrm{d}\tau \qquad \text{for all } t, s \in (a, b). \tag{5.3}$$

It then holds $u' := \partial_1 u = v$ and u has a continuous extension to a (or b) if $a > -\infty$ (or $b < \infty$). See Theorem D.10. For instance, let $u(t) = |t|$ and $v(t) = 1$ for $t > 0$ and $v(t) = -1$ for $t < 0$. Then (5.3) holds, and thus $u \in W_p^1(-1, 1)$ with $u' = v$. See also Example D.4 for further explicit examples.

(e) (Product rule) If $u \in W_p^1(U)$ and $v \in W_{p'}^1(U)$ with $\frac{1}{p} + \frac{1}{p'} = 1$, then $uv \in W_1^1(U)$ and $\partial_j(uv) = u\partial_j v + v\partial_j u$ for $j = 1, \dots, d$ (see Proposition D.7).

Let $p \in [1, \infty)$ and $k \in \mathbb{N}$. In view of Remark 5.8 (c), we define

$$\mathring{W}_p^k(U) = \text{closure of } C_c^\infty(U) \text{ in } W_p^k(U),$$

where we set $\mathring{H}^k(U) = \mathring{W}_2^k(U)$. Remark 5.8 (c) shows that $\mathring{W}_p^k(\mathbb{R}^d) = W_p^k(\mathbb{R}^d)$. We say that functions $u \in \mathring{W}_p^1(U)$ have the *trace* 0 on ∂U. This definition is justified by the following result: If ∂U is sufficiently smooth (e.g. ∂U is compact and of class C^1), then the restriction map $u \mapsto u|_{\partial U}$ from $W_p^1(U) \cap C(\overline{U})$ to $L^p(\partial U, \sigma)$ can continuously be extended to the trace operator $\operatorname{tr} : W_p^1(U) \to L^p(\partial U, \sigma)$, where σ is the surface measure on ∂U. Moreover, $\mathring{W}_p^1(U)$ is the kernel of tr. (See Theorem D.27 in Appendix D.)

We continue with an upgraded version of (5.2) which will allow us to prove the symmetry of the Laplacian on $L^2(\mathbb{R}^d)$. Let $p \in (1, \infty)$. Let $F \in \mathring{W}_p^1(U)^d$

and $\varphi \in W^1_{p'}(U)$ or let $F \in W^1_p(U)^d$ and $\varphi \in \mathring{W}^1_{p'}(U)$. We then have Gauß' formula

$$\int_U \varphi \operatorname{div} F \,\mathrm{d}x = -\int_U F \cdot \nabla\varphi \,\mathrm{d}x. \tag{5.4}$$

In fact, this identity holds for, say, $F \in C_c^\infty(U)^d$ and $\varphi \in W^1_{p'}(U)$ and thus (5.4) follows by approximation. (See Theorem D.28 for a much more precise result.)

One of the most important features of Sobolev spaces is that their elements enjoy better regularity properties than their definition directly implies. This behavior is encoded in the embedding theorems of Sobolev and Morrey (see Theorem D.15 and Corollary D.21). Here we only state certain basic versions:

$$W^k_p(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d) \qquad \text{if } k - \tfrac{d}{p} \ge -\tfrac{d}{q}, \quad q \in [p,\infty), \tag{5.5}$$

$$W^k_p(\mathbb{R}^d) \hookrightarrow C_0(\mathbb{R}^d) \qquad \text{if } k - \tfrac{d}{p} > 0. \tag{5.6}$$

For open sets $U \subseteq \mathbb{R}^d$ it further holds

$$\mathring{W}^k_p(U) \hookrightarrow L^q(U) \qquad \text{if } k - \tfrac{d}{p} \ge -\tfrac{d}{q}, \quad q \in [p,\infty), \tag{5.7}$$

$$\mathring{W}^k_p(U) \hookrightarrow C_0(U) \qquad \text{if } k - \tfrac{d}{p} > 0. \tag{5.8}$$

As a byproduct, one further obtains Poincaré's inequality

$$\int_U |\nabla u|^p \,\mathrm{d}x \ge \delta \|u\|_p^p \tag{5.9}$$

for each bounded open subset $U \subseteq \mathbb{R}^d$, $p \in [1,\infty)$, some $\delta > 0$ and all $u \in \mathring{W}^1_p(U)$. (See Theorem D.15 and Corollary D.19 in Appendix D.)

The Sobolev space $W^k_2(\mathbb{R}^d) = H^k(\mathbb{R}^d)$ can be treated by means of the Fourier transform in a very efficient way. To that purpose, we recall that for a function $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ the *Fourier transform* is given by

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) := (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}^d, \tag{5.10}$$

where $\xi \cdot x = \xi_1 x_1 + \ldots + \xi_d x_d$. Clearly, $\|\widehat{f}\|_\infty \le (2\pi)^{-\frac{d}{2}} \|f\|_1$. In fact, it holds $\widehat{f} \in C_0(\mathbb{R}^d)$ (see Corollary E.8 in Appendix E). Plancherel's theorem says that one can extend $\mathcal{F}$ to a unitary operator on $L^2(\mathbb{R}^d)$ which is again denoted by $\mathcal{F}$. Note that formula (5.10) does not hold with a Lebesgue integral if f is not integrable on $\mathbb{R}^d$.

Theorem 5.9. *The Fourier transform extends to a unitary operator $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ satisfying $(\mathcal{F}^{-1}g)(y) = (\mathcal{F}g)(-y)$ for $g \in L^2(\mathbb{R}^d)$ and $y \in \mathbb{R}^d$.*

Let $f, g \in L^2(\mathbb{R}^d)$, $h \in L^1(\mathbb{R}^d)$, $k \in \mathbb{N}$ and $j \in \{1, \ldots, d\}$. We then obtain the following assertions.

- *(a) (Plancherel)* $\quad (\mathcal{F}f|\mathcal{F}g)_{L^2} = (f|g)_{L^2}, \qquad \int_{\mathbb{R}^d} f\widehat{g}\,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{f} g\,\mathrm{d}x.$
- *(b)* $\mathcal{F}(h * f) = (2\pi)^{\frac{d}{2}} \widehat{h}\widehat{f}, \quad \mathcal{F}^{-1}(\widehat{h}\widehat{f}) = (2\pi)^{-\frac{d}{2}} h * f.$
- *(c)* $H^k(\mathbb{R}^d) = \Big\{ u \in L^2(\mathbb{R}^d) \,\Big|\, |\xi|_2^k \widehat{u} \in L^2(\mathbb{R}^d) \Big\}.$
- *(d)* $\partial_j u = \mathrm{i}\mathcal{F}^{-1}(\xi_j \widehat{u})$ *for* $u \in H^1(\mathbb{R}^d)$.

This result is proved in Theorems E.11 and E.14 in Appendix E. Here the symbols $|\xi|_2^k$ and ξ_j denote the functions $\xi \mapsto |\xi|_2^k$ and $\xi \mapsto \xi_j$, respectively. The convolution $h * f \in L^2(\mathbb{R}^d)$ is given by

$$(h * f)(x) = \int_{\mathbb{R}^d} h(x-y) f(y)\,\mathrm{d}y, \quad x \in \mathbb{R}^d,$$

see (E.3) in Appendix E.

Based on the Fourier transform we now treat the Laplacian Δ on $L^2(\mathbb{R}^d)$ and show in particular its selfadjointness. Therefore, $\mathrm{i}\Delta$ is skewadjoint which is crucial for the investigation of Schrödinger equations.

EXAMPLE 5.10. We set $X = L^2(\mathbb{R}^d)$, $m(\xi) = |\xi|_2^2$ for $\xi \in \mathbb{R}^d$,

$$\mathrm{D}(A) := \{u \in X \mid m\widehat{u} \in X\} = H^2(\mathbb{R}^d),$$

and $Au := -\mathcal{F}^{-1} m\widehat{u} = \Delta u$. The latter identities follow from Theorem 5.9. Observe that Δu is a sum of weak derivatives of second order and that $\Delta u = \operatorname{div}(\nabla u)$. We want to show that A is dissipative, selfadjoint and $\sigma(A) = \mathbb{R}_-$.

Recall that $H^2(\mathbb{R}^d) = \mathring{H}^2(\mathbb{R}^d)$ by Remark 5.8 (c), so that (5.4) implies that

$$\int_{\mathbb{R}^d} \Delta u\, \overline{v}\,\mathrm{d}x = -\int_{\mathbb{R}^d} \nabla u \cdot \nabla \overline{v}\,\mathrm{d}x = \int_{\mathbb{R}^d} u \Delta \overline{v}\,\mathrm{d}x,$$

$$\int_{\mathbb{R}^d} \Delta u\, \overline{u}\,\mathrm{d}x = -\int_{\mathbb{R}^d} |\nabla u|^2\,\mathrm{d}x \le 0$$

for all $u, v \in \mathrm{D}(A)$. This means that A is symmetric and dissipative.

Let $\lambda \in \mathbb{C} \setminus \mathbb{R}_-$. To show the surjectivity of $\lambda I - A$, we set $r_\lambda(\xi) = \frac{1}{\lambda + |\xi|^2}$ for $\xi \in \mathbb{R}^d$. Clearly, r_λ and mr_λ are bounded. We define $u = \mathcal{F}^{-1} r_\lambda \widehat{f}$ for $f \in X$. Theorem 5.9 yields $u \in X$, $m\widehat{u} = mr_\lambda \widehat{f} \in X$ and

$$\lambda u - \Delta u = \mathcal{F}^{-1}(\lambda + m)\widehat{u} = \mathcal{F}^{-1}(\lambda + m) r_\lambda \widehat{f} = f.$$

Hence, $u \in H^2(\mathbb{R}^d)$ and $\lambda I - A$ is surjective. Proposition 5.3 now shows that A generates a contraction semigroup, and thus A is closed, densely defined and satisfies $(0, \infty) \subseteq \rho(A)$. Since A is symmetric, Remark 5.6 (f) then implies that A is selfadjoint and that $\sigma(A) \subseteq \mathbb{R}$. So we arrive at $\sigma(A) \subseteq \mathbb{R}_-$. The equality $\sigma(A) = \mathbb{R}_-$ is shown in Exercise 5.1. $\Diamond$

We state two important consequences of the above result. First, the graph norm of the operator A of Example 5.10 is complete on $H^2(\mathbb{R}^d)$ since A is closed on $\mathrm{D}(A) = H^2(\mathbb{R}^d)$. The open mapping theorem then implies that $\|\cdot\|_A$ is equivalent to the H^2 norm. There thus exists a constant $c > 0$ with

$$\|u\|_2^2 + \sum_{k=1}^d \|\partial_k u\|_2^2 + \sum_{k,l=1}^d \|\partial_{kl} u\|_2^2 \le c(\|u\|_2^2 + \|\Delta u\|_2^2) \tag{5.11}$$

holds for all $u \in H^2(\mathbb{R}^d)$. In other words, the graph norm of the Laplacian dominates in the L^2 sense all derivatives of second and first order. This is a truly astonishing result in view of the possible cancellations in the sum Δu.

Thanks to the Lumer-Phillips theorem, the operator A generates a C_0–semigroup $T(\cdot)$. Hence, the function u given by $u(t) = T(t)u_0$ for $t \ge 0$ and

$u_0 \in H^2(\mathbb{R}^d)$ belongs to $C^1(\mathbb{R}_+, L^2(\mathbb{R}^d)) \cap C(\mathbb{R}_+, H^2(\mathbb{R}^d))$ and it is the unique solution of the diffusion equation

$$u'(t) = \Delta u(t), \quad t \geq 0, \qquad u(0) = u_0,$$

in this function class. We will not study this or similar parabolic equations in detail. In fact, their solutions have much better regularity properties than provided by the theory presented here. For the relevant theory on "analytic semigroups" we refer in particular to the monograph [**Lun95**].

To handle the Dirichlet Laplacian on a domain, we need another famous result (due to Lax and Milgram) which turns Riesz' representation theorem for Hilbert space duals into one of the most useful tools for applied analysis. A map T on X is called *antilinear* if $T(\alpha x + y) = \overline{\alpha}\, Tx + Ty$ for $\alpha \in \mathbb{C}$ and $x, y \in X$.

THEOREM 5.11 (Lax-Milgram lemma). *Let Y be a Hilbert space and $a : Y \times Y \to \mathbb{C}$ be a sesquilinear form (i.e., $x \mapsto a(x,y)$ is linear and $y \mapsto a(x,y)$ is antilinear for $x, y \in Y$) such that*

$$|a(x,y)| \leq c\|x\| \cdot \|y\| \qquad \text{(boundedness)},$$
$$\operatorname{Re} a(x,x) \geq \delta \|x\|^2 \qquad \text{(strict accretivity)}$$

hold for all $x, y \in Y$ and some constants $c, \delta > 0$. Then for each $\psi \in Y^*$ there is a unique $z \in Y$ such that $a(y,z) = \psi(y)$ for all $y \in Y$. The map $\psi \mapsto z$ is antilinear and bounded.

PROOF. We first establish a connection between the form a and the scalar product on Y. The map $\varphi_y := a(\,\cdot\,, y)$ belongs to Y^* with $\|\varphi_y\| \leq c\|y\|$ for each $y \in Y$ since a is bounded. Riesz' representation theorem gives a unique $Sy \in Y$ such that

$$(x|Sy) = \varphi_y(x) = a(x,y) \quad \text{for all } x \in Y$$

and $\|Sy\| = \|\varphi_y\| \leq c\|y\|$. As a result, $S \in \mathcal{B}(Y)$. Moreover, the strict accretivity yields

$$\delta\|y\|^2 \leq \operatorname{Re} a(y,y) = \operatorname{Re}(y|Sy) \leq |(y|Sy)| \leq c\|y\| \cdot \|Sy\|$$

for every $y \in Y$ which implies $\frac{\delta}{c}\|y\| \leq \|Sy\|$. It is then easy to see that S is injective and has a closed range $\mathrm{R}(S)$. If $x \in Y$ is orthogonal to $\mathrm{R}(S)$, we conclude

$$0 = (x|Sx) = \operatorname{Re}(x|Sx) = \operatorname{Re} a(x,x) \geq \delta\|x\|^2,$$

so that $x = 0$. Thus, $\mathrm{R}(S) = \overline{\mathrm{R}(S)} = Y$ and S is invertible with $\|S^{-1}\| \leq \frac{c}{\delta}$.

Let $\psi \in Y^*$. There is a unique $v \in Y$ such that $\psi = (\,\cdot\,|v)$ thanks to Riesz' theorem. Hence,

$$a(y, S^{-1}v) = (y|SS^{-1}v) = (y|v) = \psi(y)$$

for all $y \in Y$ and the vector $z := S^{-1}v = S^{-1}T\psi \in Y$ is the desired solution, where $T : Y^* \to Y$ denotes the antilinear isomorphism from Riesz' theorem. If also $\widetilde{z} \in Y$ satisfies $a(y, \widetilde{z}) = \psi(y)$ for all $y \in Y$, then $0 = a(z - \widetilde{z}, z - \widetilde{z}) \geq \delta\|z - \widetilde{z}\|^2$ as above, and thus $z = \widetilde{z}$. $\square$

We now use the Lax-Milgram lemma to define the Dirichlet Laplacian on a domain by means of a corresponding form.

Example 5.12 (Dirichlet Laplacian).

Let $\emptyset \neq U \subseteq \mathbb{R}^d$ be open and bounded and $X = L^2(U)$. We define the sesquilinear form

$$a(u,v) = \int_U \nabla u \cdot \nabla \overline{v}\,\mathrm{d}x$$

for $u, v \in \mathring{H}^1(U) =: Y$. We construct a selfadjoint, dissipative and invertible operator A corresponding to a. Due to Hölder's inequality and Poincaré's estimate (5.9), the form a satisfies the conditions of the Lax-Milgram lemma. We next introduce

$$\begin{aligned} \mathrm{D}(A) &:= \{u \in Y \,|\, \exists\, f \in X\ \forall v \in Y : a(u,v) = (f|v)_{L^2}\}, \\ Au &:= -f, \quad \text{where } f \text{ is given by } \mathrm{D}(A). \end{aligned}$$

Observe that here the function f is unique since Y is dense in X. Moreover, $\mathrm{D}(A)$ is dense in X since $C_c^\infty(U) \subseteq \mathrm{D}(A)$. Clearly, A is linear. To show its bijectivity, we take $f \in L^2(U)$. The map $\varphi_f : v \mapsto (v|f)_{L^2}$ belongs to Y^* with

$$\|\varphi_f\|_{Y^*} = \sup_{\|v\|_{1,2} \leq 1} |(v|f)_{L^2}| \leq \sup_{\|v\|_{1,2} \leq 1} \|v\|_2 \|f\|_2 \leq \|f\|_2.$$

Theorem 5.11 gives a unique $u \in Y$ satisfying $\|u\|_2 \leq \|u\|_{1,2} \leq c\|\varphi_f\|_{Y^*} \leq c\|f\|_2$ and

$$\overline{a(u,v)} = a(v,u) = \varphi_f(v) = \overline{(f|v)}_{L^2}, \quad \text{i.e., } a(u,v) = (f|v)_{L^2},$$

for all $v \in Y$. Hence, $u \in \mathrm{D}(A)$ and $Au = -f$ so that A is bijective with a bounded inverse. It follows that A is closed and that there is a point $\lambda_0 > 0$ in $\rho(A)$ since $\rho(A)$ is open. For $u, v \in \mathrm{D}(A)$ we further compute

$$\begin{aligned} (Au|v)_{L^2} &= -a(u,v) = -\overline{a(v,u)} = \overline{(Av|u)}_{L^2} = (u|Av)_{L^2}, \\ (Au|u)_{L^2} &= -a(u,u) \leq 0. \end{aligned}$$

Consequently, A is densely defined, symmetric, dissipative and $\lambda_0 I - A$ is surjective. Theorem 4.12 and Remark 5.6 now imply that A generates a contraction semigroup and that it is selfadjoint. Below, we write Δ_D instead of A. $\Diamond$

In the above example we have constructed a selfadjoint operator that we call "Dirichlet Laplacian". Let us justify this name. First, we recall that a function $u \in C(\overline{U})$ satisfies the "homogeneous Dirichlet boundary condition" if it vanishes on ∂U. As explained after Remark 5.8 this condition is replaced by "$u \in \mathring{H}^1(U)$" in the L^2 setting.

To explain the operator A itself, we set $A_0 u = Au$ for $u \in \mathrm{D}(A_0) = H^2(U) \cap \mathring{H}^1(U)$. One may consider A_0 as the "natural" Dirichlet Laplacian in $L^2(U)$. Gauß' formula (5.4) then yields

$$(-A_0 u|v) = -\int_U (\Delta u)\overline{v}\,\mathrm{d}x = \int_U \nabla u \cdot \nabla \overline{v}\,\mathrm{d}x = a(u,v)$$

for all $u \in \mathrm{D}(A_0)$ and $v \in \mathring{H}^1(U) = Y$. So the operator A extends A_0. If we assume more, namely $\partial U \in C^2$, known results about elliptic partial differential equations imply that $I - A_0$ is surjective in $L^2(U)$, see Theorems 8.3 and 8.12 in [**GT01**]. Therefore, $A = A_0$ if $\partial U \in C^2$. In this case the graph norm of A again controls the 2–norms of all derivatives of first and second order, cf. (5.11).

The above generation results can be extended to more general elliptic differential operators with suitable boundary, acting in L^p spaces or in spaces of continuous functions. Among the vast literature we refer to [**Lun95**] and [**Tan97**], and to [**Ouh05**] for form methods.

Example 5.12 allows to solve the heat equation on the domain U with Dirichlet boundary conditions. As in Example 5.10, we omit this result. We rather consider an operator matrix involving the Dirichlet Laplacian which will later be used to investigate the wave equation (cf. Lecture 1).

EXAMPLE 5.13. Let $U \subseteq \mathbb{R}^d$ be open and bounded. We use the Dirichlet Laplacian Δ_D in $L^2(U)$ introduced in the previous example. Recall that $\mathrm{D}(\Delta_D)$ consists of those $u \in \mathring{H}^1(U)$ such that there is a function $f \in L^2(U)$ with

$$\forall\, v \in \mathring{H}^1(U): \ \int_U \nabla u \cdot \nabla \overline{v}\, \mathrm{d}x = \int_U f \overline{v}\, \mathrm{d}x,$$

and then $\Delta_D u = -f$. To treat the wave equation (1.3) (at first with $b = 0$), we introduce the Hilbert space $X = \mathring{H}^1(U) \times L^2(U)$ endowed with the scalar product

$$\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \middle| \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right) = \int_U (\nabla u_1 \cdot \nabla \overline{v_1} + u_2 \overline{v_2})\, \mathrm{d}x.$$

By Poincaré's estimate (5.9) the corresponding norm is equivalent to the usual norm on X given by $(\|u_1\|_{1,2}^2 + \|u_2\|_2^2)^{\frac{1}{2}}$. On X we define the operator

$$A = \begin{pmatrix} 0 & I \\ \Delta_D & 0 \end{pmatrix} \ \text{ with } \ \mathrm{D}(A) = \mathrm{D}(\Delta_D) \times \mathring{H}^1(U).$$

As mentioned in the introduction, the Cauchy problem for A in X should correspond to the wave equation (1.3) with $b = 0$. We explain this in detail in the next lecture. In this example we show that A is skewadjoint.

The domain $\mathrm{D}(A)$ is dense in X since $C_c^\infty(U) \times C_c^\infty(U) \subseteq \mathrm{D}(A)$. For $(u_1, u_2)^\top, (v_1, v_2)^\top \in \mathrm{D}(A)$ we first compute

$$\begin{aligned} \left(A \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \middle| \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right) &= \left(\begin{pmatrix} u_2 \\ \Delta_D u_1 \end{pmatrix} \middle| \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right) = \int_U (\nabla u_2 \cdot \nabla \overline{v_1} + (\Delta_D u_1) \overline{v_2})\, \mathrm{d}x \\ &= \int_U (\nabla u_2 \cdot \nabla \overline{v_1} - \nabla u_1 \cdot \nabla \overline{v_2})\, \mathrm{d}x \\ &= -\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \middle| \begin{pmatrix} v_2 \\ \Delta_D v_1 \end{pmatrix} \right) = -\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \middle| A \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right). \end{aligned}$$

Hence, A is skewsymmetric (i.e., $\mathrm{i}A$ is symmetric). Moreover, $\mathrm{Re}(Aw|w) = 0$ for all $w \in \mathrm{D}(A)$, so that A is dissipative. We define the bounded operator

$$R = \begin{pmatrix} 0 & \Delta_D^{-1} \\ I & 0 \end{pmatrix}$$

on X, where the inverse Δ_D^{-1} exists by Example 5.12. It is easy to see that $RX \subseteq \mathrm{D}(A)$ and $AR = I$, as well as $RAw = w$ for all $w \in \mathrm{D}(A)$. As a result, $\mathrm{i}A$ is invertible. Remark 5.6 (f) now yields the selfadjointness of $\mathrm{i}A$ and so A is skewadjoint.

Exercises

Exercise 5.1. On $X = L^2(\mathbb{R}^d)$ let $Au = \Delta u$ with $\mathrm{D}(A) = H^2(\mathbb{R}^d)$. Show that $\mathbb{R}_- \subseteq \sigma(A)$ (and hence $\sigma(A) = \mathbb{R}_-$ by Example 5.10).

Exercise 5.2. Let X be a Hilbert space with an orthonormal basis $\{u_n \,|\, n \in \mathbb{N}\}$. Let $a_n \in \mathbb{C}$ be given. Define the operator A on X by $Ax = \sum_{n=1}^{\infty} a_n(x|u_n)u_n$ for $x \in \mathrm{D}(A) = \{x \in X \,\big|\, (a_n(x|u_n))_n \in \ell^2\}$. Compute the spectrum and the resolvent of A. Show that

(i) A generates a C_0–semigroup $T(\cdot)$ if and only if $\sup_{n\in\mathbb{N}} \operatorname{Re} a_n < \infty$,

(ii) A generates a unitary group $T(\cdot)$ if and only if $a_n \in \mathrm{i}\mathbb{R}$ for all $n \in \mathbb{N}$.

How does $T(\cdot)$ look like?

Exercise 5.3. Let A generate a C_0–semigroup $T(\cdot)$ on a reflexive Banach space X. Show that A^* generates a C_0–semigroup $S(\cdot)$ on X^* and that $S(t) = T(t)^*$ for all $t \geq 0$. What can be said if X is not reflexive?

Exercise 5.4. Let $X = L^2(0,\infty)$ and $Au = -u'$ for $u \in \mathrm{D}(A) = \mathring{H}^1(0,\infty)$. Show that A is dissipative, closed and $\sigma(A) = \{\lambda \in \mathbb{C} \,\big|\, \operatorname{Re}\lambda \leq 0\}$. Compute A'. (We note that A generates a contraction semigroup $T(\cdot)$ by the Lumer-Phillips theorem. As in Exercise 3.3 one can show that $T(t)f(s) = f(s-t)$ if $s-t > 0$ and $T(t)f(s) = 0$ if $s - t \leq 0$, where $f \in X$, $t \geq 0$ and $s > 0$.)

LECTURE 6

Wellposedness and inhomogeneous equations

In this lecture we complete the linear existence theory. In the introduction we have explained the concept of wellposedness and stressed that only well-posed evolution equations are truly relevant for the description of systems in the sciences. So far we have characterized those operators that generate C_0–semigroups and solved the Cauchy problems associated with them. We now prove that a Cauchy problem for a closed linear operator A is wellposed if and only if A generates a C_0–semigroup. In this sense, semigroup theory provides the natural framework for linear evolution equations.

As a second main topic we treat inhomogeneous problems where one adds a given forcing (or control) function f to the differential equation. Such equations are solved by means of Duhamel's (or the variation of parameters) formula that involves the C_0–semigroup solving the homogeneous problem.

We then apply these results to the linear wave equation. Here we will reformulate the given equation (having second order in time) as an evolution equation of first order, which belongs to the class we have studied so far.

We first repeat our basic linear evolution equation. Let A be a closed operator on X. For each given $u_0 \in \mathrm{D}(A)$, we consider the Cauchy problem

$$u'(t) = Au(t), \quad t \geq 0, \qquad u(0) = u_0. \tag{6.1}$$

Recall that a *solution* of (6.1) is a function $u \in C^1(\mathbb{R}_+, X)$ such that $u(t) \in \mathrm{D}(A)$ for all $t \in \mathbb{R}$ and u satisfies (6.1). Observe that then $Au \in C(\mathbb{R}_+, X)$ and thus $u \in C(\mathbb{R}_+, [\mathrm{D}(A)])$. We next introduce the concept of wellposedness.

DEFINITION 6.1. *Let A be closed. The Cauchy problem* (6.1) *is called* wellposed *if*

(a) $\mathrm{D}(A)$ *is dense in X,*
(b) for each $u_0 \in \mathrm{D}(A)$ there is a unique solution $u = u(\cdot\,; u_0)$ of (6.1)*,*
(c) if $u_{0,n}, u_0 \in \mathrm{D}(A)$ and the initial values $u_{0,n}$ tend to u_0 in X as $n \to \infty$, then the solutions $u(\cdot\,; u_{0,n})$ converge to $u(\cdot\,; u_0)$ uniformly on compact subsets of $\mathbb{R}_+$ (continuous dependence on initial data).

We can now establish the announced characterization of wellposedness in terms of the given operator A.

THEOREM 6.2. *Let A be a closed linear operator. Then* (6.1) *is wellposed if and only if A generates a C_0–semigroup $T(\cdot)$. In this case, the function $u = T(\cdot)u_0$ solves* (6.1) *for each given initial value $u_0 \in \mathrm{D}(A)$.*

PROOF. If A is a generator, then $T(\cdot)u_0$ is the unique solution of (6.1) according to Proposition 2.8, $\mathrm{D}(A)$ is dense in X by Proposition 3.3 and the solution depends continuously on the initial data since $T(\cdot)$ is locally bounded.

Conversely, let (6.1) be wellposed. We define the operator $T(t) : \mathrm{D}(A) \to X$ by $T(t)x = u(t;x)$ for $x \in \mathrm{D}(A)$ and $t \geq 0$ using uniqueness. For $x, y \in \mathrm{D}(A)$ and $\alpha, \beta \in \mathbb{C}$, the function v given by $v(t) = \alpha u(t;x) + \beta u(t;y)$ for $t \geq 0$ solves (6.1) with the initial value $\alpha x + \beta y$ since A is linear. Uniqueness now yields $v(t) = u(t; \alpha x + \beta y) = T(t)(\alpha x + \beta y)$, so that $T(t)$ is linear for every $t \geq 0$.

We claim that for each $t_0 > 0$ there is a $c > 0$ such that $\|T(t)x\| \leq c\|x\|$ for all $x \in \mathrm{D}(A)$ and all $t \in [0, t_0]$. In fact, if this assertion were wrong, there would exist $t_0 > 0$, a sequence $(x_n)_n$ in $\mathrm{D}(A)$ and a sequence $(t_n)_n$ in $[0, t_0]$ such that $\|x_n\| = 1$ and $\|T(t_n)x_n\| =: c_n \to \infty$ as $n \to \infty$. Set $y_n := \frac{1}{c_n} x_n \in \mathrm{D}(A)$ for every $n \in \mathbb{N}$. The initial values y_n tend to 0 as $n \to \infty$, but the norms $\|u(t_n; y_n)\| = \frac{1}{c_n}\|T(t_n)x_n\| = 1$ do not converge to 0. This contradicts the wellposedness of (6.1) and thus $T(\cdot)$ is locally bounded. So we can extend each single operator $T(t)$ to a continuous linear operator on $\overline{\mathrm{D}(A)} = X$ (also denoted by $T(t)$) having the same operator norm.

Clearly, $T(0) = I$. Since $t \mapsto T(t)x \in X$ is continuous on $\mathbb{R}_+$ for every $x \in \mathrm{D}(A)$, $\overline{\mathrm{D}(A)} = X$ and $T(\cdot)$ is locally bounded, the strong continuity of $T(\cdot)$ on X follows by approximation. Furthermore, let $t, s \geq 0$ and $x \in \mathrm{D}(A)$. Then $u(s;x)$ belongs to $\mathrm{D}(A)$ so that $v(t) := T(t)u(s;x) = u(t; u(s;x))$ for $t \geq 0$ is the unique solution of (6.1) with initial value $u(s;x)$. On the other hand, $u(t+s;x) = T(t+s)x$ for $t \geq 0$ solves this problem, too. Because solutions are unique, we obtain $T(t)T(s)x = T(t+s)x$ which gives the semigroup property by approximation.

Let B be the generator of $T(\cdot)$. By definition, we have $A \subseteq B$. Since $\mathrm{D}(A)$ is dense in X and $T(t)\,\mathrm{D}(A) \subseteq \mathrm{D}(A)$ for all $t \geq 0$, Proposition 4.10 shows that $\mathrm{D}(A)$ is a core of B. So for any $x \in \mathrm{D}(B)$, there are $x_n \in \mathrm{D}(A)$ such that $x_n \to x$ and $Ax_n = Bx_n \to Bx$ in X as $n \to \infty$. The closedness of A now implies $x \in \mathrm{D}(A)$ and $A = B$. □

One cannot drop the continuous dependence on initial data in Theorem 6.2, as seen by the next simple example.

Example 6.3. Let B be a closed, densely defined, unbounded operator on a Banach space Y. Set $X = Y \times Y$ and $A = \left(\begin{smallmatrix} 0 & B \\ 0 & 0 \end{smallmatrix}\right)$ with dense domain $\mathrm{D}(A) = Y \times \mathrm{D}(B)$. For $(x,y) \in \mathrm{D}(A)$ one has the unique solution $u(t) = \left(\begin{smallmatrix} x+tBy \\ y \end{smallmatrix}\right)$ of (6.1) with $u(0) = (x,y)$. But for $t > 0$ the map $T(t) : (\mathrm{D}(A), \|\cdot\|_X) \to X, (x,y) \mapsto u(t)$ is not continuous, since $T(t)\left(\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} tBy \\ y \end{smallmatrix}\right)$. We further note that $\lambda I - A$ is not surjective for every $\lambda \in \mathbb{C}$, see Example II.6.5 of [**EN99**]. ◇

Actually, it can be shown that (6.1) has a unique solution for a closed operator A and each $x \in \mathrm{D}(A)$ if and only if the operator A_1 on $X_1 = [\mathrm{D}(A)]$ given by $A_1 x = Ax$ with $\mathrm{D}(A_1) = \{x \in \mathrm{D}(A) \,|\, Ax \in \mathrm{D}(A)\}$ generates a C_0–semigroup on X_1, see Proposition II.6.6 in [**EN99**]. Moreover, if $\rho(A) \neq \emptyset$ and (6.1) has a unique solution for each $x \in \mathrm{D}(A)$, then A is a generator (and in particular densely defined), see Theorem II.6.7 in [**EN99**].

For $u_0 \in X$ one calls the orbit $T(\cdot)u_0$ the *mild solution* of (6.1), see Definition 6.6. In Exercise 6.3 it is shown in which sense the function $T(\cdot)u_0$ solves (6.1).

We now come to the second main topic of this lecture. In the remainder of this lecture, let $J \subseteq \mathbb{R}$ be a closed interval containing 0 and having non-empty interior. Further, let $u_0 \in X$, $f \in C(J,X)$ and A be a closed linear operator. We study the *inhomogeneous Cauchy problem* or *inhomogeneous evolution equation*

$$u'(t) = Au(t) + f(t), \quad t \in J, \qquad u(0) = u_0. \tag{6.2}$$

It is convenient to allow for finite time intervals J here since possibly f is not given for all times. (As noted in the introduction this situation occurs when treating nonlinear problems.) Moreover, for later use we include backward time. Our solution concept for (6.1) directly extends to (6.2).

Definition 6.4. *A function $u : J \to X$ is a* solution *of (6.2) if u belongs to $C^1(J,X)$, $u(t) \in \mathrm{D}(A)$ for all $t \in J$ and (6.2) holds.*

This definition implies that the initial value u_0 of a solution must belong to $\mathrm{D}(A)$. Observe that a solution of (6.2) is contained in $C(J,[\mathrm{D}(A)])$. Our solutions are often called "classical" or "strict" solutions in the literature.

We first derive "Duhamel's formula" for the solutions of (6.2). Since the case $f = 0$ is included, it is natural to assume from the beginning that A is a generator. The next proof is just a variant of the uniqueness part of the proof of Proposition 2.8.

Proposition 6.5. *Let A generate the C_0–semigroup $T(\cdot)$, $u_0 \in \mathrm{D}(A)$, and $f \in C(J,X)$. If $J \not\subseteq \mathbb{R}_+$, we require that $T(\cdot)$ can be extended to a C_0–group. If u solves (6.2), then u is given by*

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\,\mathrm{d}s, \quad t \in J. \tag{6.3}$$

In particular, solutions of (6.2) are unique.

Proof. For simplicity, we concentrate on the case that $J \subseteq \mathbb{R}_+$. Let $t \in J$ with $t > 0$ and set $v(s) = T(t-s)u(s)$ for $0 \le s \le t$, where u solves (6.2). Using Lemma 2.9, one shows that v is continuously differentiable with derivative

$$v'(s) = T(t-s)u'(s) - T(t-s)Au(s) = T(t-s)f(s)$$

for all $0 \le s \le t$. By integration we deduce

$$\int_0^t T(t-s)f(s)\,\mathrm{d}s = v(t) - v(0) = u(t) - T(t)u_0. \qquad \square$$

We point out that by Duhamel's formula (6.3) one can *define* a function $u \in C(J,X)$ for any given $u_0 \in X$ and $f \in C(J,X)$. This leads to a weaker solution concept, which turns out to be very useful.

Definition 6.6. *Let A generate the C_0–semigroup $T(\cdot)$, $u_0 \in X$ and $f \in C(J,X)$. If $J \not\subseteq \mathbb{R}_+$, we require that $T(\cdot)$ can be extended to a C_0–group. Then the function $u \in C(J,X)$ given by*

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\,\mathrm{d}s, \quad t \in J,$$

is called the mild solution *of (6.2).*

Proposition 6.5 says that every solution of (6.2) is a mild one, whereas the next example shows that the converse implication may fail.

EXAMPLE 6.7. Let $X = C_0(\mathbb{R})$, $A = \frac{\mathrm{d}}{\mathrm{ds}}$ with $\mathrm{D}(A) = C_0^1(\mathbb{R})$ and let $\varphi \in X$ be non-differentiable. The operator A generates the C_0–group $T(\cdot)$ given by $T(t)g = g(\,\cdot\, + t)$, see Example 3.5. Clearly, $T(t)\varphi \notin \mathrm{D}(A)$ for all $t \in \mathbb{R}$. Set $f(s) = T(s)\varphi$ for $s \in \mathbb{R}$. The function f belongs to $C(\mathbb{R}, X)$ and the mild solution of (6.2) with $u_0 = 0$ is given by

$$u(t) = \int_0^t T(t-s)T(s)\varphi \,\mathrm{d}s = tT(t)\varphi$$

for $t \in \mathbb{R}$. Hence, $u(t) \notin \mathrm{D}(A)$ for $t \neq 0$, i.e., u does not solve (6.2). $\diamondsuit$

We will derive conditions on f and u_0 such that the mild solution of (6.2) in fact solves (6.2). At first, we treat the case $u_0 = 0$.

LEMMA 6.8. *Let A generate the C_0–semigroup $T(\cdot)$ and $f \in C(J, X)$. If $J \nsubseteq \mathbb{R}_+$, we require that $T(\cdot)$ can be extended to a C_0–group. Define*

$$v(t) = \int_0^t T(t-s)f(s)\,\mathrm{d}s, \quad t \in J.$$

Then the following assertions are equivalent.

(a) $v \in C^1(J, X)$.

(b) $v(t) \in \mathrm{D}(A)$ for all $t \in J$ and $Av \in C(J, X)$.

If (a) or (b) are valid, v solves (6.2) with $u_0 = 0$.

PROOF. We concentrate on the case that $J \subseteq \mathbb{R}_+$. The general case can be treated similarly. Since f is locally bounded, the function v belongs to $C(J, X)$ and $v(0) = 0$. To show the asserted equivalence, we need a few preparations. We fix $t \in J$ and take $h > 0$ such that $t \pm h \in J$. We then define

$$D_1(h) := \tfrac{1}{h}(T(h) - I)v(t), \qquad D_2^{\pm}(h) := \tfrac{1}{\pm h}(v(t \pm h) - v(t)),$$

$$I^+(h) := \frac{1}{h}\int_t^{t+h} T(t+h-s)f(s)\,\mathrm{d}s, \qquad I^-(h) := \frac{1}{h}\int_{t-h}^{t} T(t-s)f(s)\,\mathrm{d}s,$$

assuming that $t > 0$ for $D_2^-(h)$ and $I^-(h)$. (These formulas are asymmetric since $T(\cdot)$ is not assumed to be a group.) Observe that

$$D_1(h) = D_2^+(h) - I^+(h), \tag{6.4}$$

$$\begin{aligned} D_2^-(h) &= \tfrac{1}{h}(T(h) - I)v(t-h) + I^-(h) \\ &= \tfrac{1}{h}A\int_0^h T(\tau)v(t-h)\,\mathrm{d}\tau + I^-(h), \end{aligned} \tag{6.5}$$

where we use Lemma 3.2 in the last equality. We start by investigating $I^{\pm}(h)$. Employing the continuity of f and Lemma 2.9, we obtain

$$\begin{aligned} \|I^+(h) - f(t)\| &= \Big\|\frac{1}{h}\int_t^{t+h} (T(t+h-s)f(s) - f(t))\,\mathrm{d}s\Big\| \\ &\le \max_{t \le s \le t+h} \|T(t+h-s)f(s) - f(t)\| \;\longrightarrow\; 0 \end{aligned}$$

as $h \to 0^+$. Similarly, one sees that $I^-(h) \to f(t)$ as $h \to 0^+$.

First, assume that v satisfies (a). Hence, $D_2^{\pm}(h) \to v'(t)$ as $h \to 0^+$. Equality (6.4) then implies that $D_1(h)$ converges to $v'(t) - f(t)$ as $h \to 0^+$. As a result, $v(t)$ belongs to $\mathrm{D}(A)$ for all $t \in J$ and $Av = v' - f$ is continuous.

Second, let (b) hold so that $D_1(h) \to Av(t)$ as $h \to 0^+$. From (6.4) we now infer that v is differentiable from the right and $(\frac{\mathrm{d}}{\mathrm{dt}})^+ v = Av + f$. Moreover, (6.5) and (b) yield

$$
\begin{aligned}
D_2^-(h) &= \frac{1}{h}\int_0^h T(\tau)Av(t-h)\,\mathrm{d}\tau + I^-(h) \\
&= \frac{1}{h}\int_0^h T(\tau)Av(t)\,\mathrm{d}\tau + \frac{1}{h}\int_0^h T(\tau)(Av(t-h) - Av(t))\,\mathrm{d}\tau + I^-(h) \\
&\to Av(t) + f(t)
\end{aligned}
$$

as $h \to 0^+$, thanks to Remark 2.10 (d) and (g) and the continuity of Av. Summing up, v is differentiable with $v' = Av + f \in C(J, X)$ so that (a) is true. In both cases we have also shown that v solves (6.2) with $u_0 = 0$. □

The above lemma now implies that we obtain solutions of (6.2) if the initial value u_0 belongs to $\mathrm{D}(A)$ and if the inhomogeneity has either more time regularity (i.e., $f \in C^1(J, X)$) or more "space" regularity (i.e., $f \in C(J, [\mathrm{D}(A)])$).

Theorem 6.9 (Existence result for inhomogeneous evolution equations). *Let A generate the C_0–semigroup $T(\cdot)$, $u_0 \in \mathrm{D}(A)$ and $J \subseteq \mathbb{R}$ be a closed interval containing 0. If $J \nsubseteq \mathbb{R}_+$, we require that $T(\cdot)$ can be extended to a C_0–group. Assume either that $f \in C^1(J, X)$ or that $f \in C(J, [\mathrm{D}(A)])$. Then the mild solution u given by (6.3) is the unique solution of (6.2) on J.*

Proof. Uniqueness was already shown in Proposition 6.5. By Proposition 2.8, the function $T(\cdot)u_0$ is contained in $C^1(J, X) \cap C(J, [\mathrm{D}(A)])$ and solves (6.2) with $f = 0$. It remains to show that the map $t \mapsto v(t) = \int_0^t T(t-s)f(s)\,\mathrm{d}s$ defined in Lemma 6.8 also belongs to $C^1(J, X) \cap C(J, [\mathrm{D}(A)])$ and solves (6.2) with $u_0 = 0$, since then $u = T(\cdot)u_0 + v$ solves (6.2). In view of this lemma, we have to verify (a) or (b) in Lemma 6.8.

Let $f \in C^1(J, X)$. Since $v(t) = \int_0^t T(s)f(t-s)\,\mathrm{d}s$ for all $t \in J$, it follows that $v \in C^1(J, X)$ and hence (a) in Lemma 6.8 is satisfied.

Let $f \in C(J, [\mathrm{D}(A)])$. Since A is closed and commutes with $T(t-s)$ on $\mathrm{D}(A)$, we obtain $v(t) \in \mathrm{D}(A)$ and $Av(t) = \int_0^t T(t-s)Af(s)\,\mathrm{d}s$ so that $Av \in C(J, X)$. In this case, (b) in Lemma 6.8 is fulfilled. □

We want to apply the above results to the wave equation on a bounded open set $\emptyset \neq U \subseteq \mathbb{R}^d$ with Dirichlet boundary conditions. This equation describes the displacement $w(t, x)$ of a vibrating body at a time $t \in \mathbb{R}$ and at a point $x \in U$. Here we consider the system

$$
\begin{aligned}
&\partial_{tt}w(t,x) = \Delta w(t,x) - b(x)\partial_t w(t,x) + g(t,x), && x \in U, \quad t \in \mathbb{R}, \\
&w(t,x) = 0, && x \in \partial U, \quad t \in \mathbb{R}, \\
&w(0,x) = w_0(x), \quad \partial_t w(0,x) = w_1(x), && x \in U,
\end{aligned}
\tag{6.6}
$$

for the given initial displacement w_0 and initial velocity distribution w_1. The functions g and $b \geq 0$ are also given. Let us sketch the physical background of this equation. (This is not meant to be an honest derivation.)

In the differential equation we assume for simplicity that the mass density of the body is equal to 1 everywhere. The differential equation in (6.6) then describes the balance of forces at the space point x and at the time t. The acceleration $\partial_{tt}w(t,x)$ is equal to the sum of the "forces" on the right-hand side, where $\Delta w(t,x)$ describes the mechanical force due to tension, $-b(x)\partial_t w(t,x)$ is a damping proportional to the velocity $\partial_t w(t,x)$ and $g(t,x)$ corresponds to an external force. The term $\Delta w(t,x)$ can be justified if one assumes that the material is perfectly elastic, homogeneous (with material constant 1) and if only small deflections occur.

So far we have implicitly assumed that $w \in C^2(J\times\overline{U})$ so that (6.6) would hold in a pointwise sense. For the analysis of the problem it is much more convenient to reformulate (6.6) as an evolution equation (of second order in time) in the space $L^2(U)$. To this aim, we use the Dirichlet Laplacian Δ_D introduced in Example 5.12. Recall that Δ_D acts in $L^2(U)$, that $\mathrm{D}(\Delta_D) \subseteq \mathring{H}^1(U)$ and that

$$(-\Delta_D u|v)_{L^2} = \int_U \nabla u \cdot \nabla \overline{v}\,\mathrm{d}x \qquad \text{for } u \in \mathrm{D}(\Delta_D) \text{ and } v \in \mathring{H}^1(U). \tag{6.7}$$

Let $A_0 u = \Delta u$ with $\mathrm{D}(A_0) = H^2(U) \cap \mathring{H}^1(U)$. As indicated after Example 5.10, we have $A_0 \subseteq \Delta_D$ and $A_0 = \Delta_D$ if $\partial U \in C^2$. However, we will stick to the case of a general bounded open $U \subseteq \mathbb{R}^d$ and to the operator Δ_D given by (6.7). Using this setup, we rewrite (6.6) as

$$\begin{aligned} &w''(t) = \Delta_D w(t) - bw'(t) + g(t), \quad t \in J, \\ &w(0) = w_0, \qquad w'(0) = w_1, \end{aligned} \tag{6.8}$$

where w' and w'' denote the first and second derivatives of w with respect to t. The unknown w now is a function from J to $L^2(U)$ and similarly for g. The Dirichlet boundary conditions and the differential operator Δ are incorporated in the operator Δ_D in a somewhat generalized form. We are looking for *solutions*

$$w \in C^2(J, L^2(U)) \cap C^1(J, \mathring{H}^1(U)) \qquad \text{with } w(t) \in \mathrm{D}(\Delta_D) \text{ for all } t \in J$$

of (6.8), with the given initial values

$$w_0 \in \mathrm{D}(\Delta_D) \quad \text{and} \quad w_1 \in \mathring{H}^1(U).$$

For simplicity we assume that $b \in L^\infty(U)$. Moreover, we take $g \in C(J, L^2(U))$. We note that the time interval J is dictated by g. If $g = 0$, we will take $J = \mathbb{R}$. Observe that a solution w of (6.8) satisfies the integrated equation

$$\int_U w''(t)\,\overline{\varphi}\,\mathrm{d}x + \int_U \nabla w(t) \cdot \nabla\overline{\varphi}\,\mathrm{d}x + \int_U bw'(t)\,\overline{\varphi}\,\mathrm{d}x = \int_U g(t)\,\overline{\varphi}\,\mathrm{d}t$$

for all $t \in J$ and each $\varphi \in \mathring{H}^1(U)$, see (6.7). We will come back to such "weak formulations" in later lectures.

We call the Cauchy problem (6.8) with $g = 0$ *wellposed* if the following conditions hold.

(a) For all $w_0 \in \mathrm{D}(\Delta_D)$ and $w_1 \in \mathring{H}^1(U)$ there is a unique solution $w = w(\,\cdot\,; w_0, w_1)$ of (6.8) with $g = 0$.
(b) Let $w_{0,n}, w_0 \in \mathrm{D}(\Delta_D)$ and $w_{1,n}, w_1 \in \mathring{H}^1(U)$, $n \in \mathbb{N}$. Assume that $w_{0,n} \to w_0$ in $\mathring{H}^1(U)$ and that $w_{1,n} \to w_1$ in $L^2(U)$ as $n \to \infty$. Then $w(t; w_{0,n}, w_{1,n})$ converge to $w(t; w_0, w_1)$ in $\mathring{H}^1(U)$ and $w'(t; w_{0,n}, w_{1,n})$ tend to $w'(t; w_0, w_1)$ in $L^2(U)$ as $n \to \infty$, both locally uniformly in $t \in J$.

Comparing the above concept with Definition 6.1 you may miss the density condition. In fact, we already know from Example 5.13 that $\mathrm{D}(\Delta_D) \times \mathring{H}^1(U)$ is dense in $\mathring{H}^1(U) \times L^2(U)$ which is the appropriate density assumption in view of (b) above.

To solve (6.8) and to show its wellposedness, we want to use the theory established in this lecture. As in the case of ordinary differential equations, we thus introduce the new state

$$u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \begin{pmatrix} w(t) \\ w'(t) \end{pmatrix}.$$

The state space will be

$$X = \mathring{H}^1(U) \times L^2(U)$$

endowed with the scalar product

$$(u|v) = ((u_1, u_2)|(v_1, v_2)) = \int_U (\nabla u_1 \cdot \nabla \overline{v_1} + u_2 \overline{v_2})\, \mathrm{d}x.$$

This choice fits well to condition (b) above, where solutions are required to converge in the corresponding norm on X given by

$$\|u\|^2 = \int_U |\nabla u_1|_2^2 \,\mathrm{d}x + \int_U |u_2|^2 \,\mathrm{d}x.$$

(As noted in Example 5.13 this norm is equivalent to the usual norm on X.) Physically one can interpret $\|u(t)\|^2$ as the total energy of the solution $u(t) = (w(t), w'(t))$ modulo constants, where $\int_U |\nabla w(t)|_2^2 \,\mathrm{d}x$ corresponds to the potential energy and $\int_U |w'(t)|^2 \,\mathrm{d}x$ to the kinetic energy. We further define the operator

$$A = \begin{pmatrix} 0 & I \\ \Delta_D & -b \end{pmatrix} \quad \text{with } \mathrm{D}(A) = \mathrm{D}(\Delta_D) \times \mathring{H}^1(U) \tag{6.9}$$

in X, where $-b$ denotes the bounded multiplication operator $\varphi \mapsto -b\varphi$ on $L^2(U)$. Finally, we put $u_0 = (w_0, w_1)$ and $f = (0, g) \in C(J, X)$. We recall from Example 5.13 that for $b = 0$ the operator A is skewadjoint in X.

We next describe in which sense (6.8) is equivalent to the evolution equation (6.2) for the operator matrix A defined in (6.9).

Lemma 6.10. *Let $U \subseteq \mathbb{R}^d$ be open and bounded and $b \in L^\infty(U)$. Let $u_0 = (w_0, w_1) \in \mathrm{D}(A)$ and $g \in C(J, L^2(U))$, and set $f = (0, g) \in C(J, X)$. Then the following assertions hold.*

(a) The problem (6.8) has a solution w if and only if the problem (6.2) with A from (6.9) has a solution u. If this is the case, we have $u = (w, w')$. Also the uniqueness of solutions to these two problems is equivalent.

(b) *The problem* (6.8) *with* $g = 0$ *is wellposed if and only if the problem* (6.1) *with* A *from* (6.9) *is wellposed.*

PROOF. (a) Let $w \in C^2(J, L^2(U)) \cap C^1(J, \mathring{H}^1(U))$ solve (6.8). Then $u := (w, w')$ belongs to $C^1(J, X)$, $u(t) \in \mathrm{D}(A)$ for all $t \in J$ and

$$u'(t) = \begin{pmatrix} w'(t) \\ w''(t) \end{pmatrix} = \begin{pmatrix} w'(t) \\ \Delta_D w(t) - bw'(t) + g(t) \end{pmatrix} = Au(t) + f(t)$$

holds for all $t \in J$. Moreover, $u(0) = (w(0), w'(0)) = u_0$. Thus, u solves (6.2).

Conversely, let $u = (u_1, u_2)$ solve (6.2) for A. We set $w = u_1$ obtaining $w \in C^1(J, \mathring{H}^1(U))$ and $w(t) \in \mathrm{D}(\Delta_D)$ for all $t \in J$. It further follows

$$\begin{pmatrix} w'(t) \\ u_2'(t) \end{pmatrix} = A \begin{pmatrix} w(t) \\ u_2(t) \end{pmatrix} + f(t) = \begin{pmatrix} u_2(t) \\ \Delta_D w(t) - bu_2(t) + g(t) \end{pmatrix}$$

for all $t \in J$. As a consequence, $w' = u_2 \in C^1(J, L^2(U))$ and $u = (w, w')$, so that $w \in C^2(J, L^2(U))$, $(w(0), w'(0)) = (w_0, w_1)$ and w solves (6.8). This equivalence also yields that the solutions to (6.2) for our A are unique if and only if the solutions to (6.8) are unique.

(b) It follows from Example 5.13 that $\mathrm{D}(A)$ is dense in X. Part (a) then easily implies the equivalence of the two wellposedness assertions. □

Combining the above lemma with Example 5.13 and the previous theorems, we can now solve the undamped wave equation with $b = 0$. The damping $b \neq 0$ will be treated in the next lecture by a perturbation argument.

PROPOSITION 6.11. *Let* $U \subseteq \mathbb{R}^d$ *be open and bounded,* $J \subseteq \mathbb{R}$ *be a closed interval containing 0,* $(w_0, w_1) \in \mathrm{D}(\Delta_D) \times \mathring{H}^1(U)$ *and either* $g \in C^1(J, L^2(U))$ *or* $g \in C(J, \mathring{H}^1(U))$*. Then the wave equation* (6.8) *with* $b = 0$ *has a unique solution. Moreover, problem* (6.8) *with* $b = 0$ *and* $g = 0$ *is wellposed.*

PROOF. Thanks to Example 5.13, the operator A defined in (6.9) with $b = 0$ is skewadjoint on $X = \mathring{H}^1(U) \times L^2(U)$ and thus generates a (unitary) C_0–group on X by Stone's Theorem 5.7. The assertions now follow from Theorems 6.2 and 6.9 and Lemma 6.10. □

We finally consider the "free" Schrödinger equation given by

$$\begin{aligned} \partial_t u(t,x) &= \mathrm{i}\Delta u(t,x), \quad x \in \mathbb{R}^d, \ t \in \mathbb{R}, \\ u(0,x) &= u_0(x), \qquad\quad x \in \mathbb{R}^d. \end{aligned} \tag{6.10}$$

Here we look for solutions $u \in C^1(\mathbb{R}, L^2(\mathbb{R}^d)) \cap C(\mathbb{R}, H^2(\mathbb{R}^d))$. To obtain such solutions we introduce in $L^2(\mathbb{R}^d)$ the operator A given by $Au = \mathrm{i}\Delta u$ with $\mathrm{D}(A) = H^2(\mathbb{R}^d)$. We say that (6.10) is *wellposed* if the Cauchy problem (6.1) for this operator A is wellposed in $L^2(\mathbb{R}^d)$. Since A is skewadjoint by Example 5.10, Stone's Theorem 5.7 and Theorem 6.2 immediately yield the next result.

PROPOSITION 6.12. *The free Schrödinger equation* (6.10) *is wellposed.*

Exercises

Exercise 6.1. Let $X = C_0(0,1)$, $Au(s) = s(1-s)u''(s)$ for $s \in (0,1)$ and $u \in \mathrm{D}(A) = \{u \in C^2(0,1) \cap X \,|\, Au \in X\}$. Show that A is densely defined, dissipative, invertible and generates a contraction semigroup on X.

Exercise 6.2. Let X be a Banach space and A generate a C_0–semigroup $T(\cdot)$ on X such that $T(t) \to I$ in $\mathcal{B}(X)$ as $t \to 0$. Show that $\lambda R(\lambda, A) \to I$ in $\mathcal{B}(X)$ as $\lambda \to \infty$ and deduce that A is bounded.

Exercise 6.3. Let X be a Banach space, A generate a C_0–semigroup $T(\cdot)$ on X, $f \in C(\mathbb{R}_+, X)$ and $u_0 \in X$. An *integrated solution* u of

$$u'(t) = Au(t) + f(t), \quad t \geq 0, \qquad u(0) = u_0, \tag{6.11}$$

is a function $u \in C(\mathbb{R}_+, X)$ such that $\int_0^t u(s)\,\mathrm{d}s \in \mathrm{D}(A)$ and

$$u(t) = A\int_0^t u(s)\,\mathrm{d}s + u_0 + \int_0^t f(s)\,\mathrm{d}s \quad \text{for all } t \geq 0.$$

Show that the mild solution

$$v(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\,\mathrm{d}s$$

is the unique integrated solution of (6.11).

LECTURE 7

Perturbation and exponential stability

So far we have introduced the basic concepts of linear semigroup theory, characterized wellposedness and solved inhomogeneous evolution equations. We can already handle the wave equation without damping and the free Schrödinger equation. One can now develop the theory of linear evolution equations in various directions. In this lecture we establish fundamental results in two major subjects: *perturbation theory* and *asymptotic behavior*.

We first prove the main perturbation theorem for contraction semigroups. It allows us to treat the wave equation (1.3) with damping and the Schrödinger equation (1.2) with a potential. The bounded perturbation theorem for general C_0–semigroups will also be deduced from the contraction case.

The long-term behavior of linear evolution equations is a vast field which could easily fill another Internet Seminar. Here, we only consider the arguably simplest case. We want to find conditions on the generator A such that $\|T(t)\| \to 0$ as $t \to \infty$. If a C_0–semigroup $T(\cdot)$ is "uniformly stable" in this sense, all solutions of the Cauchy problem

$$u'(t) = Au(t), \quad t \geq 0, \qquad u(0) = x, \tag{7.1}$$

tend to 0 as $t \to \infty$, uniformly for x in bounded subsets of X. In view of Lyapunov's theorem for ordinary differential equations, one may hope that the above behavior can be characterized by the spectral condition $s(A) < 0$. Unfortunately, this guess is wrong as various examples show. However, if X is a Hilbert space, "Gearhart's stability theorem" will tell us that uniform stability is equivalent to $s(A) < 0$ plus a resolvent estimate.

We start with the perturbation problem. Let A generate a C_0–semigroup and B be linear. We want to show that also $A + B$ is a generator if B is sufficiently "small". When writing down $A + B$, one first has to think about the domain of this sum. In our course we only consider the simplest case, where $\mathrm{D}(A) \subseteq \mathrm{D}(B)$. Therefore the sum $A + B$ will always be defined on $\mathrm{D}(A + B) := \mathrm{D}(A)$.

But also in this case $A + B$ can fail to generate a C_0–semigroup. Consider for instance $Au = \Delta u$ with $\mathrm{D}(A) = H^2(\mathbb{R}^d)$ in $L^2(\mathbb{R}^d) = X$ and $B = -bA$ for any $b > 1$. We then have $A + B = (1 - b)A$ and thus $\sigma(A + B) = \mathbb{R}_+$, cf. Example 5.10. In particular, $A + B$ is *not* a generator. Clearly, for $b \in (0, 1)$ the sum $A + B$ generates a C_0–semigroup. In the borderline case $b = -1$ we obtain $A + B = 0$ on $\mathrm{D}(A + B) = H^2(\mathbb{R}^2)$. This operator is not closed, but has the closure $0 \in \mathcal{B}(X)$ which generates the trivial group $T(t) = I$.

To measure the "smallness" of B with respect to A, we introduce the following concept.

Definition 7.1. *Let A and B be linear operators in X with $\mathrm{D}(A) \subseteq \mathrm{D}(B)$. Then B is called* A-bounded *(or* relatively bounded*) if there are constants $a, b \geq 0$ such that*

$$\|Bx\| \leq a\|Ax\| + b\|x\| \qquad \text{for all } x \in \mathrm{D}(A). \tag{7.2}$$

The infimum of the numbers $a \geq 0$ for which (7.2) holds for some $b = b(a) \geq 0$ is called the A-bound *of B.*

We remark that B is A-bounded if and only if $B \in \mathcal{B}([\mathrm{D}(A)], X)$. In particular, if B is A-bounded and $\lambda \in \rho(A)$, then $BR(\lambda, A)$ is well defined and a bounded operator on X. Observe that a bounded operator B is A-bounded with constants $a = 0$ and $b = \|B\|$.

We first show that $A + B$ inherits the closedness of A if B has an A-bound strictly less than 1.

Lemma 7.2. *Let A be closed and B be A-bounded with A-bound strictly less than 1. Then the graph norms of A and $A + B$ on $\mathrm{D}(A)$ are equivalent and hence $A + B$ with $\mathrm{D}(A + B) = \mathrm{D}(A)$ is closed.*

Proof. By our assumption the estimate (7.2) holds for some $a \in [0, 1)$ and $b \geq 0$. Let $x \in \mathrm{D}(A)$. We then derive

$$\begin{aligned} \|(A + B)x\| &\leq \|Ax\| + \|Bx\| \leq (1 + a)\|Ax\| + b\|x\|; \\ \|Ax\| &\leq \|(A + B)x\| + \|Bx\| \leq \|(A + B)x\| + a\|Ax\| + b\|x\|, \\ \|Ax\| &\leq \frac{1}{1 - a}\|(A + B)x\| + \frac{b}{1 - a}\|x\|. \end{aligned}$$

In particular, $\mathrm{D}(A)$ is a Banach space for the graph norm of $A + B$, so that $A + B$ is closed by Remark 2.11 (a). □

We continue with a useful condition implying that B has the A-bound 0.

Lemma 7.3. *Let A and B be linear operators with $\mathrm{D}(A) \subseteq \mathrm{D}(B)$ and $\|Bx\| \leq c\|Ax\|^{\alpha}\|x\|^{1-\alpha}$ for all $x \in \mathrm{D}(A)$ and for some constants $c \geq 0$ and $\alpha \in (0, 1)$. Then B is A-bounded with the A-bound 0.*

Proof. Recall Young's elementary inequality $ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'}$ for all $a, b \geq 0$, $p \in (1, \infty)$ and $p' = \frac{p}{p-1}$. Using it with $p = \frac{1}{\alpha}$ and $p' = \frac{1}{1-\alpha}$, we estimate

$$\|Bx\| \leq \varepsilon\|Ax\|^{\alpha}\tfrac{c}{\varepsilon}\|x\|^{1-\alpha} \leq \alpha\varepsilon^{\frac{1}{\alpha}}\|Ax\| + (1 - \alpha)c^{\frac{1}{1-\alpha}}\varepsilon^{-\frac{1}{1-\alpha}}\|x\|$$

for all $x \in \mathrm{D}(A)$ and $\varepsilon > 0$. □

Our approach to the perturbation theorem relies on the following result on the resolvent which is proved by means of the Neumann series.

Lemma 7.4. *Let A be closed, $\lambda \in \rho(A)$ and B be A-bounded with $\|BR(\lambda, A)\| < 1$. Then $1 \in \rho(BR(\lambda, A))$, $A + B$ is closed and $\lambda \in \rho(A + B)$ with*

$$R(\lambda, A + B) = R(\lambda, A)\sum_{n=0}^{\infty}(BR(\lambda, A))^n = R(\lambda, A)(I - BR(\lambda, A))^{-1}, \tag{7.3}$$

$$\|R(\lambda, A + B)\| \leq \frac{\|R(\lambda, A)\|}{1 - \|BR(\lambda, A)\|}.$$

Proof. Let $\lambda \in \rho(A)$. Since $\|BR(\lambda, A)\| =: q < 1$, the operator $I - BR(\lambda, A)$ is invertible and its inverse is given by the (Neumann) series in (7.3). Moreover, the norm of the inverse is bounded by $(1-q)^{-1}$. Combining these results with the identity

$$\lambda I - A - B = (I - BR(\lambda, A))(\lambda I - A),$$

we derive the remaining assertions. □

We can now prove the dissipative perturbation theorem.

Theorem 7.5. *Let A generate the contraction semigroup $T(\cdot)$ on X. Let B be an A-bounded, dissipative operator with A-bound strictly less than 1. Then the following assertions hold.*

(a) *The operator $A + B$ with $\mathrm{D}(A + B) = \mathrm{D}(A)$ generates a contraction semigroup $S(\cdot)$ on X.*

(b) *For $x \in \mathrm{D}(A)$ and $t \geq 0$ we have the integral equations*

$$S(t)x = T(t)x + \int_0^t T(t-s)BS(s)x \,\mathrm{d}s, \tag{7.4}$$

$$S(t)x = T(t)x + \int_0^t S(t-s)BT(s)x \,\mathrm{d}s. \tag{7.5}$$

(c) *In addition, let X be a Hilbert space, A be skewadjoint and $-B$ be dissipative. Then $A + B$ generates a unitary group.*

Proof. 1) Observe that $\mathrm{D}(A+B) = \mathrm{D}(A)$ is dense in X. Let $x \in \mathrm{D}(A) \subseteq \mathrm{D}(B)$. Since B is dissipative, there is an $x^* \in J(x)$ such that $\mathrm{Re}\langle Bx, x^*\rangle \leq 0$. Proposition 4.5 shows that $\mathrm{Re}\langle Ax, x^*\rangle \leq 0$, and hence $A + B$ is dissipative.

2) We want to check the range condition in the Lumer-Phillips theorem. Here we use (7.2) which holds for some constants $a \in [0, 1)$ and $b \geq 0$. We first assume that $a < \frac{1}{2}$. Take some $\lambda_0 > \frac{b}{1-2a} \geq 0$. Inequality (7.2) and the Hille-Yosida estimate yield

$$\begin{aligned}\|BR(\lambda_0, A)\| &\leq a\|(A - \lambda_0 I + \lambda_0 I)R(\lambda_0, A)\| + b\|R(\lambda_0, A)\| \\ &\leq a + a\lambda_0\|R(\lambda_0, A)\| + b\|R(\lambda_0, A)\| \leq a + a + \tfrac{b}{\lambda_0} < 1.\end{aligned}$$

Lemma 7.4 now implies that $A + B$ is closed and $\lambda_0 \in \rho(A + B)$. Therefore $A + B$ generates a contraction semigroup by the Lumer-Phillips Theorem 4.12.

3) If $a \geq \frac{1}{2}$, we fix $k \in \mathbb{N}$ with $k > \frac{2a}{1-a}$. Then the operator $\frac{1}{k}B$ with domain $\mathrm{D}(B)$ is dissipative and A-bounded with the constant $a' = \frac{a}{k} < \frac{1-a}{2} \leq \frac{1}{2}$. By step 2), $A + \frac{1}{k}B$ thus generates a contraction semigroup. Inductively, we assume that $C_j := A + \frac{j}{k}B$ generates a contraction semigroup for some $j \in \{1, \dots, k-1\}$. Let $x \in \mathrm{D}(A)$. Inequality (7.2) yields

$$\begin{aligned}\|Bx\| &\leq a\|Ax\| + b\|x\| \leq a\|C_j x\| + a\tfrac{j}{k}\|Bx\| + b\|x\|, \\ (1-a)\|Bx\| &\leq (1 - a\tfrac{j}{k})\|Bx\| \leq a\|C_j x\| + b\|x\|.\end{aligned}$$

It follows that

$$\|\tfrac{1}{k}Bx\| \leq \frac{a}{k(1-a)}\|C_j x\| + \frac{b}{k(1-a)}\|x\|$$

for all $x \in \mathrm{D}(A)$. Since $\frac{a}{k(1-a)} < \frac{1}{2}$, step 2) implies that $C_j + \frac{1}{k}B = C_{j+1}$ generates a contraction semigroup. By induction, we derive that $A + C_k = A + B$ generates a contraction semigroup $S(\cdot)$. So we have shown assertion (a).

4) Let $x \in \mathrm{D}(A) = \mathrm{D}(A+B)$. Observe that $B \in \mathcal{B}([\mathrm{D}(A+B)], X)$ due to Lemma 7.2, so that the function $f = BS(\cdot)x$ is well defined and belongs to $C(\mathbb{R}_+, X)$. The map $u = S(\cdot)x$ solves the problem

$$u'(t) = (A+B)u(t) = Au(t) + f(t), \quad t \geq 0, \qquad u(0) = x.$$

Proposition 6.5 then implies (7.4). Similarly, we set $g = BT(\cdot)x \in C(\mathbb{R}_+, X)$ and note that $v = T(\cdot)x$ solves

$$v'(t) = Av(t) = (A+B)v(t) - g(t), \quad t \geq 0, \qquad u(0) = x.$$

Hence, also (7.5) follows from Proposition 6.5.

5) Under the assumptions of (c), we can apply assertion (a) to $-A$ and $-B$, so that $-(A+B)$ generates a contraction semigroup. Since also $A+B$ generates a contraction semigroup, as in Corollary 5.4 we can now derive that $A+B$ generates an isometric C_0–group, which is unitary due to Proposition C.7. □

There is a result in the borderline case $a = 1$ in (7.2). It can be shown that $A+B$ is closable and that $\overline{A+B}$ generates a contraction semigroup on X if X is reflexive, A generates a contraction semigroup, B is dissipative and (7.2) holds with $a = 1$. See e.g. Corollary III.2.9 in [**EN99**].

We next deduce the bounded perturbation theorem by means of a simple rescaling and renorming procedure.

Theorem 7.6. *Let $B \in \mathcal{B}(X)$ and A generate the C_0–semigroup $T(\cdot)$ on X. Fix constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that $\|T(t)\| \leq M\mathrm{e}^{\omega t}$ for all $t \geq 0$. Then the sum $A+B$ with $\mathrm{D}(A+B) = \mathrm{D}(A)$ generates a C_0–semigroup $S(\cdot)$ on X satisfying*

$$\|S(t)\| \leq M\mathrm{e}^{(\omega + M\|B\|)t} \quad \text{for all } t \geq 0.$$

Furthermore, the equations (7.4) and (7.5) hold for all $x \in X$ and $t \geq 0$. Finally, if A generates a C_0–group, then $S(\cdot)$ can be extended to a C_0–group.

Proof. Let $x \in X$. We define

$$|||x||| = \sup_{s \geq 0} \|\mathrm{e}^{-\omega s}T(s)x\|.$$

It is straightforward to check that $|||\cdot|||$ is a norm on X, that $\|x\| \leq |||x||| \leq M\|x\|$, and that $|||\mathrm{e}^{-\omega t}T(t)x||| \leq |||x|||$ for all $t \geq 0$ and $x \in X$. Using Lemma 3.1, we derive that $A - \omega I$ generates the contraction semigroup $(e^{-\omega t}T(t))_{t \geq 0}$ on $(X, |||\cdot|||)$. To relate B to $|||\cdot|||$, we estimate

$$|||Bx||| \leq M\|Bx\| \leq M\|B\| \, \|x\| \leq M\|B\| \, |||x|||.$$

Set $\beta = M\|B\|$. Observe that

$$\mathrm{Re}\langle (B - \beta I)x, x^*\rangle = \mathrm{Re}\langle Bx, x^*\rangle - \beta |||x|||^2 \leq |||Bx||| \, |||x||| - \beta |||x|||^2 \leq 0$$

for all $x \in X$ and $x^* \in J(x)$, where X^* is equipped with the norm induced by $|||\cdot|||$. Theorem 7.5 now yields that $A - \omega I + B - \beta I$ with domain $\mathrm{D}(A)$ generates a contraction semigroup $\widetilde{S}(\cdot)$ on $(X, |||\cdot|||)$, so that $A+B$ generates the C_0–semigroup given by $S(t) = \mathrm{e}^{(\omega+\beta)t}\widetilde{S}(t)$ on X. The asserted estimate now follows from

$$\|S(t)x\| \le \|\mathrm{e}^{(\omega+\beta)t}\widetilde{S}(t)x\| \le \mathrm{e}^{(\omega+\beta)t}\|x\| \le M\mathrm{e}^{(\omega+\beta)t}\|x\|.$$

If A generates a group, then the above results yield that also $-(A+B)$ generates a C_0–semigroup. Hence, $S(\cdot)$ can be extended to a C_0–group due to Theorem 4.2. Finally, the equations (7.4) and (7.5) can be shown for $x \in \mathrm{D}(A)$ as in the proof of Theorem 7.5. We can then extend them to all $x \in X$ since B is continuous. □

The above results imply the desired generation properties for the damped wave equation and the Schrödinger equation.

Example 7.7 (damped wave equation).
As in Example 5.12, let $U \subseteq \mathbb{R}^d$ be open and bounded and let Δ_D be the Dirichlet Laplacian in $L^2(U)$. Consider the damping $0 \le b \in L^\infty(U)$ and the forcing function $g \in C(J, L^2(U))$ with a closed interval $J \subseteq \mathbb{R}$ containing 0. We assume that either $g \in C^1(J, L^2(U))$ or $g \in C(J, \mathring{H}^1(U))$. Finally, let $w_0 \in \mathrm{D}(\Delta_D)$ and $w_1 \in \mathring{H}^1(U)$ be given. In Lecture 6 we have explained how one reformulates the damped wave equation (6.6) as the second order problem

$$\begin{aligned} &w''(t) = \Delta_D w(t) - bw'(t) + g(t), \qquad t \in J,\\ &w(0) = w_0, \qquad w'(0) = w_1. \end{aligned} \tag{7.6}$$

We claim that this problem has a unique solution

$$w \in C^2(J, L^2(U)) \cap C^1(J, \mathring{H}^1(U)) \cap C(J, [\mathrm{D}(\Delta_D)]).$$

Moreover, if $g = 0$, then the problem (7.6) is wellposed and the "energy" of the solution

$$\int_U |\nabla w(t)|_2^2 \,\mathrm{d}x + \int_U |w'(t)|^2 \,\mathrm{d}x = \|(w(t), w'(t))\|_X^2$$

decreases in $t \in J$. To show these claims, we consider the Hilbert space $X = \mathring{H}^1(U) \times L^2(U)$ with the above norm $\|\cdot\|_X$ and introduce the operators

$$A_0 = \begin{pmatrix} 0 & I \\ \Delta_D & 0 \end{pmatrix} \qquad \text{with } \mathrm{D}(A_0) = \mathrm{D}(\Delta_D) \times \mathring{H}^1(U),$$

$$B = \begin{pmatrix} 0 & 0 \\ 0 & -b \end{pmatrix} \qquad \text{on } X,$$

where $-b$ denotes the operator $v \mapsto -bv$. By Example 5.13, A_0 is skewadjoint. The operator B is bounded since $b \in L^\infty(U)$ and dissipative since

$$\left(B\binom{u}{v} \Big| \binom{u}{v}\right)_X = -\int_U bv\overline{v}\,\mathrm{d}x \le 0$$

for all $(u, v) \in X$. Theorems 7.6 and 7.5 now show that $A_0 + B$ with domain $\mathrm{D}(A_0)$ generates a C_0–group $T(\cdot)$ on X and that $\|T(t)\| \le 1$ for $t \ge 0$. Lemma 6.10 thus yields the unique solvability of (7.6) and the wellposedness. Moreover, the solution is given by $(w(t), w'(t)) = T(t-s)(w(s), w'(s))$ for $t \ge s$ in J so that the "energy" decays due to contractivity. ◇

The above result can be extended to certain unbounded b, see Exercise 7.2. We next treat the Schrödinger equation for the hydrogen atom, cf. (1.2).

Example 7.8 (Schrödinger equation).
Let $V(x) = \frac{-b}{|x|_2}$ for $b \in \mathbb{R}$ and $x \in \mathbb{R}^3 \setminus \{0\}$ and $V(0) = 0$. Take $u_0 \in H^2(\mathbb{R}^3)$. We consider the Schrödinger equation

$$\begin{aligned} \partial_t u(t,x) &= \mathrm{i}\Delta u(t,x) - \mathrm{i}V(x)u(t,x), \quad && x \in \mathbb{R}^3, t \in \mathbb{R}, \\ u(0,x) &= u_0(x), && x \in \mathbb{R}^3. \end{aligned} \tag{7.7}$$

Example 5.10 implies the skewadjointness of the operator A_0 given by $A_0 u = \mathrm{i}\Delta u$ with $\mathrm{D}(A_0) = H^2(\mathbb{R}^3)$. We first show that the multiplication operator $u \mapsto -\mathrm{i}Vu$ has the A_0-bound 0. Recall that Sobolev's embedding (5.6) yields $H^2(\mathbb{R}^3) \hookrightarrow C_0(\mathbb{R}^3)$ since $2 - \frac{3}{2} > 0$. Let $\varepsilon \in (0,1]$. Using also polar coordinates and (5.11), we estimate

$$\begin{aligned} \int_{\mathbb{R}^3} |Vu|^2 \,\mathrm{d}x &= b^2 \int_{B(0,\varepsilon)} \frac{|u(x)|^2}{|x|_2^2} \,\mathrm{d}x + b^2 \int_{\mathbb{R}^3 \setminus B(0,\varepsilon)} \frac{|u(x)|^2}{|x|_2^2} \,\mathrm{d}x \\ &\le c \int_0^\varepsilon \frac{r^2}{r^2} \,\mathrm{d}r \|u\|_\infty^2 + \frac{b^2}{\varepsilon^2} \int_{\mathbb{R}^3 \setminus B(0,\varepsilon)} |u(x)|^2 \,\mathrm{d}x \\ &\le c\varepsilon \|u\|_{2,2}^2 + \frac{b^2}{\varepsilon^2} \|u\|_2^2 \le c\varepsilon \|A_0 u\|_2^2 + c\varepsilon \|u\|_2^2 + \frac{b^2}{\varepsilon^2} \|u\|_2^2, \end{aligned}$$

for constants $c > 0$ independent of u and ε. Moreover,

$$\operatorname{Re}(-\mathrm{i}Vu|u) = -\operatorname{Re} \mathrm{i} \int_{\mathbb{R}^3} V|u|^2 \,\mathrm{d}x = 0$$

for all $u \in \mathrm{D}(A_0) = H^2(\mathbb{R}^3)$. Theorem 7.5 now shows that the operator A given by $Au = \mathrm{i}\Delta u - \mathrm{i}Vu$ with $\mathrm{D}(A) = H^2(\mathbb{R}^3)$ generates a unitary group in $L^2(\mathbb{R}^3)$. Hence, the Schrödinger equation (7.7) is wellposed (in the same sense as for (6.10)), where the solution satisfies $\|u(t)\|_2 = \|u_0\|_2$ for all $t \in \mathbb{R}$.

In quantum mechanics, if $\|u_0\|_2 = 1$ the function $\int_G |u(t,x)|^2 \,\mathrm{d}x$ describes the probability that the electron in a hydrogen atom is contained in a Borel set $G \subseteq \mathbb{R}^3$ at time $t \in \mathbb{R}$ (if $b > 0$ is chosen correctly). ♢

We now come to the asymptotic theory and start with a definition and a few preliminary observations.

Definition 7.9. *A C_0–semigroup $T(\cdot)$ is called* exponentially stable *if there exist constants $M, \beta > 0$ such that*

$$\|T(t)\| \le M\mathrm{e}^{-\beta t} \quad \textit{for all } t \ge 0.$$

Let A be the generator of $T(\cdot)$. We note that $\|T(t)\| \le \mathrm{e}^{-\beta t}$ for all $t \ge 0$ if and only if $A + \beta I$ is dissipative, by the Lumer-Phillips theorem and rescaling. Proposition 3.4 implies that $s(A) < 0$ if A generates an exponentially stable C_0–semigroup.

We first characterize exponential stability on the level of the semigroup. To this aim, we recall the formula

$$r(T) := \max \{|\lambda| \mid \lambda \in \sigma(T)\} = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \inf_{n\in\mathbb{N}} \|T^n\|^{\frac{1}{n}} \le \|T\|, \tag{7.8}$$

for the spectral radius of $T \in \mathcal{B}(X)$, see Theorem B.6 in Appendix B.

PROPOSITION 7.10. *Let $T(\cdot)$ be a C_0–semigroup with generator A. Then the following assertions are equivalent.*

(a) *$T(\cdot)$ is exponentially stable.*
(b) *$\|T(t_0)\| < 1$ for some $t_0 > 0$.*
(c) *$r(T(t_1)) < 1$ for some $t_1 > 0$.*
(d) *$\omega_0(A) < 0$.*

If this is the case, then (b) is valid for all sufficiently large $t_0 > 0$, assertion (c) is true for all $t_1 > 0$ and we have $s(A) < 0$.

Moreover, $\mathrm{e}^{ts(A)} \le \mathrm{e}^{t\omega_0(A)} = r(T(t))$ for all $t > 0$ and (with $\log(0) := -\infty$)

$$\omega_0(A) = \lim_{t\to\infty} \tfrac{1}{t}\log\|T(t)\| = \inf_{t>0} \tfrac{1}{t}\log\|T(t)\|. \tag{7.9}$$

PROOF. Since $\log\|T(t+s)\| \le \log\|T(t)\| + \log\|T(s)\|$, the elementary Lemma IV.2.3 in [**EN99**] shows that the limit $\lim_{t\to\infty} \frac{1}{t}\log\|T(t)\|$ exists and is equal to $\omega := \inf_{t>0} \frac{1}{t}\log\|T(t)\|$. Hence, $\mathrm{e}^{t\omega} \le \|T(t)\|$ for all $t \ge 0$, and we obtain $\omega \le \omega_0(A)$. Take any $\omega_1 > \omega$. Then there exists $t_0 \ge 0$ such that $\|T(t)\| \le \mathrm{e}^{\omega_1 t}$ for all $t \ge t_0$ so that $\|T(t)\| \le M\mathrm{e}^{\omega_1 t}$ for all $t \ge 0$ and $M := \sup\{\mathrm{e}^{-\omega_1 t}\|T(t)\| \,|\, 0 \le t \le t_0\} \ge 1$. This estimate leads to $\omega \ge \omega_0(A)$, and so (7.9) holds. Using (7.8) and (7.9), we infer

$$r(T(t)) = \lim_{n\to\infty} \exp(t\tfrac{1}{nt}\log\|T(nt)\|) = \exp(t \lim_{n\to\infty}(\tfrac{1}{nt}\log\|T(nt)\|) = \mathrm{e}^{t\omega_0(A)}$$

for each $t > 0$. All other assertions about $T(\cdot)$ now follow. Since $s(A) \le \omega_0(A)$ by Proposition 3.4, we also have $\mathrm{e}^{ts(A)} \le \mathrm{e}^{t\omega_0(A)}$. □

If $X = \mathbb{C}^d$, a theorem by Lyapunov says that $s(A) = \omega_0(A)$, and thus $T(\cdot)$ is exponentially stable if and only if $s(A) < 0$. Here the spectrum of the *given* operator A determines (this aspect of) the long-term behavior of the semigroup. Unfortunately, in infinite dimensions it may happen that $s(A) < \omega_0(A)$ as the following example shows.

EXAMPLE 7.11 (Greiner-Voigt-Wolff, 1981). Let $X = C_0(\mathbb{R}_+) \cap L^1(\mathbb{R}_+, \mathrm{e}^s\,\mathrm{d}s)$ be endowed with the norm

$$\|f\| = \|f\|_\infty + \int_0^\infty |f(s)|\mathrm{e}^s\,\mathrm{d}s =: \|f\|_\infty + \|f\|_{1,\mathrm{w}}.$$

Define $T(t)f = f(\,\cdot\, + t)$ for $t \ge 0$ and $f \in X$. Observe that

$$\begin{aligned}\|T(t)f\| &= \sup_{s\ge 0}|f(s+t)| + \int_0^\infty |f(s+t)|\mathrm{e}^s\,\mathrm{d}s \le \|f\|_\infty + \mathrm{e}^{-t}\int_t^\infty |f(\tau)|\mathrm{e}^\tau\,\mathrm{d}\tau \\ &\le \|f\|_\infty + \mathrm{e}^{-t}\|f\|_{1,\mathrm{w}} \le \|f\|. \end{aligned} \tag{7.10}$$

Hence, $T(t)$ is a contraction on X. It is easy to see that $T(\cdot)$ is a semigroup on X and strongly continuous on $C_c(\mathbb{R}_+)$ for $\|\cdot\|$, and that $C_c(\mathbb{R}_+)$ is dense in X. As a result, $T(\cdot)$ is a contraction semigroup on X. We denote its generator by A. Given $t > 0$ and $q \in (0,1)$, take $f \in X$ with $\|f\| = 1$ and $\|f\|_\infty = |f(t)| = q$. Then $\|T(t)\| \ge \|T(t)f\| \ge |T(t)f(0)| = |f(t)| = q$. So we obtain that $\|T(t)\| = 1$ and therefore $\omega_0(A) = 0$.

We want to show that $s(A) = -1$. Note that $s(A) \le \omega_0(A) = 0$. First let $\operatorname{Re}\lambda < -1$. (Recall that $\mathrm{e}_\lambda(t) = \mathrm{e}^{\lambda t}$.) We then have $\mathrm{e}_\lambda \in X$ and $T(t)\mathrm{e}_\lambda = \mathrm{e}^{\lambda t}\mathrm{e}_\lambda$, for all $t \ge 0$ so that $\frac{1}{t}(T(t)\mathrm{e}_\lambda - \mathrm{e}_\lambda)$ tends to $\lambda \mathrm{e}_\lambda$ as $t \to 0^+$. It follows that $\mathrm{e}_\lambda \in \mathrm{D}(A)$ with $A\mathrm{e}_\lambda = \lambda \mathrm{e}_\lambda$, Hence, $\{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda \le -1\} \subseteq \sigma(A)$ since $\sigma(A)$ is closed. Next, let $\operatorname{Re}\lambda \in (-1, 0]$ and $f \in X$. The estimate (7.10) shows that $\|T(t)f\|_{1,\mathrm{w}} \le \mathrm{e}^{-t}\|f\|_{1,\mathrm{w}}$ for all $t \ge 0$, so that the integrals $\int_0^b \mathrm{e}^{-\lambda t} T(t)f \,\mathrm{d}t =: J(b)$ converge in $L^1(\mathbb{R}_+, \mathrm{e}^s \,\mathrm{d}s)$ as $b \to \infty$. Moreover, for $b' > b \ge 0$ we have

$$\Big\| \int_b^{b'} \mathrm{e}^{-\lambda t} T(t) f \,\mathrm{d}t \Big\|_\infty \le \sup_{s\ge 0} \int_b^{b'} \mathrm{e}^{-\operatorname{Re}\lambda t} |f(s+t)| \,\mathrm{d}t$$
$$= \sup_{s\ge 0} \int_{b+s}^{b'+s} \mathrm{e}^{\operatorname{Re}\lambda s} \mathrm{e}^{-\operatorname{Re}\lambda \tau} |f(\tau)| \,\mathrm{d}\tau \le \int_b^\infty \mathrm{e}^{\tau} |f(\tau)| \,\mathrm{d}\tau.$$

As a consequence, $J(b)$ also converges in $C_0(\mathbb{R}_+)$ as $b \to \infty$. Proposition 3.4 now shows that $\lambda \in \rho(A)$, and hence $s(A) = -1 < \omega_0(A) = 0$. $\diamondsuit$

We note that the above example can be modified such that $s(A) = -\infty$ (see Exercise IV.2.13(5) in [**EN99**]). There are analogous examples on Hilbert spaces using growing Jordan blocks on $X = \ell^2$ (see [**Zab75**]) or a perturbed wave equation (see [**Ren94**]). On the other hand, various positive results are known in this context. One could impose additional assumptions on the semigroup, see e.g. Corollary IV.3.12 in [**EN99**]. Since such results cannot be used for, say, the damped wave equation, we rather combine spectral conditions with resolvent estimates. Below we show the exponential stability of a C_0–semigroup on a *Hilbert* space if the resolvent $R(\lambda, A)$ of its generator A exists and is uniformly bounded for $\operatorname{Re}\lambda > 0$.

In the proof we need the Fourier transform of Hilbert space-valued functions and thus the Bochner integral, which is introduced in the next intermezzo. More details can be found in Appendix F.

Intermezzo 4: Bochner's integral and the Fourier transform

Let X be a Banach space and $J \subseteq \mathbb{R}$ be an interval. Simple functions $f : J \to X$ and their integral are defined as in the case $X = \mathbb{C}$. A function $f : J \to X$ is called *strongly measurable* if there are simple functions $f_n : J \to X$ converging to f pointwise. Observe that then the function $t \mapsto \|f(t)\|$ is measurable. Hence, we can define

$$L^p(J, X) := \{ f : J \to X \mid f \text{ is strongly measurable and } \|f(\cdot)\|_X \in L^p(J) \},$$
$$\|f\|_p := \| \|f(\cdot)\|_X \|_{L^p(J)},$$

for $p \in [1, \infty]$, where we identify functions that coincide almost everywhere. It can be seen that $f \in L^1(J, X)$ if and only if there are simple functions f_n converging to f pointwise such that $\int_J \|f_n - f\| \,\mathrm{d}t$ tends to 0 as $n \to \infty$, see Lemma F.4 in Appendix F. This fact implies that the integrals $\int_J f_n(t) \,\mathrm{d}t$ converge in X and that their limit is independent of the choice of such a sequence $(f_n)_n$. This limit is denoted by $\int_J f(t) \,\mathrm{d}t$ and called the *(Bochner) integral* of

f. The integral is linear and we have

$$\Big\|\int_J f\,\mathrm{d}t\Big\| \le \int_J \|f\|\,\mathrm{d}t \quad \text{and} \quad T\int_J f(t)\,\mathrm{d}t = \int_J Tf(t)\,\mathrm{d}t$$

for $f \in L^1(J,X)$ and $T \in \mathcal{B}(X,Y)$, where Y is another Banach space. It can be shown that $L^p(J,X)$ is a Banach space and the analogues of Hölder's inequality and the theorems of Riesz-Fischer, Lebesgue and Fubini hold for the Bochner integral, see Appendix F. Moreover, if X is a Hilbert space, then $L^2(J,X)$ is a Hilbert space for the scalar product given by $(f|g) = \int_J (f(t)|g(t))_X\,\mathrm{d}t$.

For $f \in L^1(\mathbb{R},X) \cap L^2(\mathbb{R},X)$ we define the Fourier transform

$$\widehat{f}(\tau) = \mathcal{F}f(\tau) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\tau t} f(t)\,\mathrm{d}t, \quad \tau \in \mathbb{R}.$$

If X is a Hilbert space, then $\mathcal{F}$ extends to a unitary operator

$$\mathcal{F} : L^2(\mathbb{R},X) \to L^2(\mathbb{R},X)$$

whose inverse is given by $\mathcal{F}^{-1}g(t) = \mathcal{F}g(-t)$ for $g \in L^2(\mathbb{R},X)$ and $t \in \mathbb{R}$. See Theorem F.16 in Appendix F.

The next result is a special case of a theorem proved by Gearhart in 1978 for contraction semigroups and independently by Herbst (1983), Howland (1984) and Prüss (1984) for general semigroups.

Concerning the boundedness condition in Theorem 7.12 we recall that $\|R(\lambda,A)\| \le c_\delta$ for all $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > \omega_0(A) + \delta$ and all $\delta > 0$, due to the Hille-Yosida estimate. Thus in Theorem 7.12 the behavior of $\|R(\lambda,A)\|$ as $\operatorname{Re}\lambda \to 0^+$ determines the long-term behavior of $T(\cdot)$.

Theorem 7.12 (Gearhart's stability theorem). *Let X be a Hilbert space. A C_0–semigroup $T(\cdot)$ with generator A is exponentially stable if and only if $s(A) < 0$ and $C = \sup_{\lambda\in\mathbb{C}_+} \|R(\lambda,A)\| < \infty$, where $\mathbb{C}_+ = \{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda > 0\}$.*

Proof. The necessity of the condition was shown in Proposition 3.4 for a general Banach space X. Assume that the resolvent exists and is bounded on $\mathbb{C}_+$. In view of Datko's Lemma 7.13 below, we want to show that $T(\cdot)x \in L^2(\mathbb{R}_+,X)$ for all $x \in X$. As we will see, this fact can be deduced from resolvent estimates. We first relate $T(\cdot)$ and $R(\,\cdot\,,A)$ by means of the Fourier transform.

Set $\omega_1 = \max\{\omega_0(A), 0\}$ and take $\omega > \omega_1$. Let $x \in X$. We set $T_\omega(t) = \mathrm{e}^{-\omega t}T(t)$ for $t \ge 0$ and $T_\omega(t) = 0$ for $t < 0$. Moreover, we put $r_\lambda(\tau) = R(\lambda + \mathrm{i}\tau, A)x$ for $\lambda \in \mathbb{C}_+$ and $\tau \in \mathbb{R}$. Then $u = T_\omega(\cdot)x$ belongs to $L^2(\mathbb{R},X) \cap L^1(\mathbb{R},X)$ for all $x \in X$. Fix $\overline{\omega} > w_1$. It follows that $\|T_{\overline{\omega}}(\cdot)x\|_2 \le c_0\|x\|$ for some $c_0 > 0$ only depending on $\overline{\omega}$ and the exponential estimate for $T(\cdot)$. Using Proposition 3.4, we compute

$$\mathcal{F}u(\tau) = \frac{1}{\sqrt{2\pi}}\int_0^\infty \mathrm{e}^{-\mathrm{i}\tau t}\mathrm{e}^{-\omega t}T(t)x\,\mathrm{d}t = \frac{1}{\sqrt{2\pi}}R(\omega + \mathrm{i}\tau, A)x, \tag{7.11}$$

for $\tau \in \mathbb{R}$ and $\omega > \omega_1$. We next estimate the resolvent. Since $\mathcal{F}$ is unitary, equation (7.11) implies that r_ω belongs to $L^2(\mathbb{R},X)$ if $\omega > \omega_1$. Moreover, $\|r_{\overline{\omega}}\|_2 \le \sqrt{2\pi}c_0\|x\|$. From the resolvent equation we further deduce that

$$r_\omega(\tau) = R(\omega + \mathrm{i}\tau, A)x = R(\overline{\omega} + \mathrm{i}\tau, A)x + (\overline{\omega} - \omega)R(\omega + \mathrm{i}\tau, A)R(\overline{\omega} + \mathrm{i}\tau, A)x$$

for all $\omega > 0$ and $\tau \in \mathbb{R}$. The assumption thus leads to the basic estimate

$$\|r_\omega\|_2 \le \sqrt{2\pi} c_0 (1 + C|\omega - \overline{\omega}|)\|x\| \tag{7.12}$$

for all $\omega > 0$ and $x \in X$. From (7.11), (7.12) and the unitarity of $\mathcal{F}$, we now deduce that

$$\|T_\omega(\cdot)x\|_2 = \frac{1}{\sqrt{2\pi}}\|r_\omega\|_2 \le c_0(1 + C|\omega - \overline{\omega}|)\|x\| \tag{7.13}$$

for all $\omega > \omega_1$. Due to Fatou's lemma, the estimate (7.13) also holds for $\omega = \omega_1$. Datko's Lemma 7.13 then implies that $(\mathrm{e}^{-\omega_1 t} T(t))_{t \ge 0}$ is exponentially stable which cannot be true for $\omega_1 = \omega_0(A)$. Hence $\omega_1 = 0$ and we have shown the assertion. □

In a general Banach space X the boundedness of the resolvent $R(\,\cdot\,, A)$ on $\mathbb{C}_+$ only implies the existence of some constants $M, \varepsilon > 0$ such that we have

$$\|T(t)x\| \le M \mathrm{e}^{-\varepsilon t}\|x\|_A$$

for all $t \ge 0$ and $x \in \mathrm{D}(A)$. See [**WW96**] for this and related results and examples indicating their optimality.

Lemma 7.13 (Datko, 1970). *Let $T(\cdot)$ be a C_0–semigroup on X and $1 \le p < \infty$. If $T(\cdot)x \in L^p(\mathbb{R}_+, X)$ for all $x \in X$, then $T(\cdot)$ is exponentially stable.*

Proof. Define the bounded operator

$$\Phi_n : X \to L^p(\mathbb{R}_+, X), \quad x \mapsto \mathbb{1}_{[0,n]} T(\cdot)x,$$

for each $n \in \mathbb{N}$. The assumption shows that $\sup_{n\in\mathbb{N}} \|\Phi_n(x)\| < \infty$ for all $x \in X$, and hence $C := \sup_{n\in\mathbb{N}} \|\Phi_n\|$ is finite thanks to the principle of uniform boundedness. As a result, $\int_0^t \|T(s)x\|^p \,\mathrm{d}s \le C^p\|x\|^p$ for all $t \ge 0$ and $x \in X$. Fix constants $M \ge 1$ and $\omega > 0$ such that $\|T(t)\| \le M\mathrm{e}^{\omega t}$ for all $t \ge 0$. We then calculate

$$\begin{aligned}
\frac{1 - \mathrm{e}^{-p\omega t}}{p\omega}\|T(t)x\|^p &= \int_0^t \mathrm{e}^{-p\omega s}\|T(s)T(t-s)x\|^p \,\mathrm{d}s \\
&\le \int_0^t M^p \mathrm{e}^{\omega s p}\mathrm{e}^{-\omega s p}\|T(t-s)x\|^p \,\mathrm{d}s \\
&= M^p \int_0^t \|T(\tau)x\|^p \,\mathrm{d}\tau \le (CM)^p\|x\|^p
\end{aligned}$$

for all $t \ge 0$ and $x \in X$. Therefore $\|T(t)\| \le N$ for all $t \ge 0$, where $N := \max\{M\mathrm{e}^{\omega}, (p\omega)^{\frac{1}{p}} CM(1 - \mathrm{e}^{-p\omega})^{-\frac{1}{p}}\}$. It follows that

$$t\|T(t)x\|^p = \int_0^t \|T(t-s)T(s)x\|^p \,\mathrm{d}s \le N^p \int_0^t \|T(s)x\|^p \,\mathrm{d}s \le (CN)^p\|x\|^p,$$

and hence $\|T(t)\| \le CNt^{-\frac{1}{p}}$. Proposition 7.10 now implies the assertion. □

We conclude with an application to the damped wave equation.

Example 7.14 (damped wave equation).
Let $\emptyset \neq U \subseteq \mathbb{R}^d$ be open and bounded and let $b \in L^\infty(U)$ satisfy $b(x) \geq \beta$ for almost every $x \in U$ and some $\beta > 0$. Example 7.7 yields that

$$A = \begin{pmatrix} 0 & I \\ \Delta_D & -b \end{pmatrix} \quad \text{with } \mathrm{D}(A) = \mathrm{D}(\Delta_D) \times \mathring{H}^1(U)$$

generates a contraction C_0–semigroup $T(\cdot)$ on $X = \mathring{H}^1(U) \times L^2(U)$ and the function $(w(t), w'(t)) = T(t)(w_0, w_1)$ solves (7.6) for $g = 0$ and $(w_0, w_1) \in \mathrm{D}(A)$. We assert that $T(\cdot)$ is exponentially stable, and thus the "energy"

$$\|T(t)(w_0, w_1)\|_X^2 = \||\nabla w|\|_2^2 + \|w'(t)\|_2^2$$

of the solution decays as $c\mathrm{e}^{-2\varepsilon t}\|(w_0, w_1)\|_X^2$ for some $c, \varepsilon > 0$. In order to show this fact, we use Theorem 7.12. We first note that

$$R\begin{pmatrix} \varphi \\ \psi \end{pmatrix} = \begin{pmatrix} \Delta_D^{-1}(b\varphi + \psi) \\ \varphi \end{pmatrix}, \quad (\varphi, \psi) \in X,$$

defines the bounded inverse of A (Δ_D is invertible by Example 5.12). Next, we show that

$$\mathrm{i}\mathbb{R} \subseteq \rho(A) \quad \text{and} \quad \sup_{\mathrm{i}\tau \in \mathbb{R}} \|R(\mathrm{i}\tau, A)\| =: C < \infty. \tag{7.14}$$

If (7.14) holds, it will follow that $\lambda \in \rho(A)$ and $\|R(\lambda, A)\| \leq 2C$ whenever $|\operatorname{Re}\lambda| \in [0, \frac{1}{2C}]$, due to Lemma 7.4 (with λ replaced by $\mathrm{i}\tau$ and B by $\pm\operatorname{Re}\lambda I$). Combining this inequality with the Hille-Yosida estimate, we then deduce the asserted exponential stability from Theorem 7.12.

We show (7.14). Because of $s(A) \leq 0$, any $\mathrm{i}\tau \in \sigma(A)$ would belong to $\partial\sigma(A)$ and it would thus follow that

$$m(\tau) := \inf\{\|\mathrm{i}\tau u - Au\|_X \mid u \in \mathrm{D}(A), \|u\|_X = 1\} = 0$$

due to Proposition C.2. Hence, if $\inf_{\tau \in \mathbb{R}} m(\tau) =: m_0 > 0$, then $\mathrm{i}\mathbb{R} \subseteq \rho(A)$ and moreover (7.14) holds with $C = \frac{1}{m_0}$. Since $0 \in \rho(A)$ and $\rho(A)$ is open, there exists $\tau_0 > 0$ such that $[-\mathrm{i}\tau_0, \mathrm{i}\tau_0] \subseteq \rho(A)$, and so $m(\tau) \geq \delta := (\max_{|s| \leq \tau_0} \|R(\mathrm{i}s, A)\|)^{-1} > 0$ for all $\tau \in [-\tau_0, \tau_0]$. Fix $\varepsilon \in (0, \frac{\beta}{2})$ such that $0 < \frac{3\varepsilon\beta}{\beta - 2\varepsilon} < \tau_0$. Suppose there were $|\tau| \geq \tau_0$ and $u = (\varphi, \psi) \in \mathrm{D}(A)$ such that

$$\|u\|_X^2 = \||\nabla\varphi|\|_2^2 + \|\psi\|_2^2 = 1 \quad \text{and} \quad \|\mathrm{i}\tau u - Au\|_X \leq \varepsilon. \tag{7.15}$$

We then compute

$$\begin{aligned}
\varepsilon &\geq \left|\left((\mathrm{i}\tau I - A)\begin{pmatrix} \varphi \\ \psi \end{pmatrix} \middle| \begin{pmatrix} \varphi \\ \psi \end{pmatrix}\right)\right| \\
&= \left|\int_U \nabla(\mathrm{i}\tau\varphi - \psi) \cdot \nabla\overline{\varphi}\,\mathrm{d}x + \int_U (-\Delta_D\varphi\,\overline{\psi} + (\mathrm{i}\tau + b)\psi\overline{\psi})\,\mathrm{d}x\right| \\
&= \left|\mathrm{i}\tau(\||\nabla\varphi|\|_2^2 + \|\psi\|_2^2) - \overline{\int_U \nabla\varphi \cdot \nabla\overline{\psi}\,\mathrm{d}x} + \int_U \nabla\varphi \cdot \nabla\overline{\psi}\,\mathrm{d}x + \int_U b|\psi|^2\,\mathrm{d}x\right| \\
&= \left|\mathrm{i}(\tau + 2\operatorname{Im}\int_U \nabla\varphi \cdot \nabla\overline{\psi}\,\mathrm{d}x) + \int_U b|\psi|^2\,\mathrm{d}x\right|,
\end{aligned}$$

using the definition of Δ_D. Considering imaginary and real part, we infer that

$$\varepsilon \geq \left|\tau + 2\operatorname{Im}\int_U \nabla\varphi \cdot \nabla\overline{\psi}\,\mathrm{d}x\right| \quad \text{and} \quad \varepsilon \geq \int_U b|\psi|^2\,\mathrm{d}x \geq \beta\|\psi\|_2^2,$$

The second estimate yields $\| |\nabla\varphi| \|_2^2 = 1 - \|\psi\|_2^2 \geq 1 - \frac{\varepsilon}{\beta}$ and hence $1 - 2\| |\nabla\varphi| \|_2^2 \leq \frac{2\varepsilon}{\beta} - 1 < 0$, because $\varepsilon < \frac{\beta}{2}$. We conclude that

$$\begin{aligned}
|\tau| \left(1 - \frac{2\varepsilon}{\beta}\right) &\leq |\tau| \left|1 - 2\| |\nabla\varphi| \|_2^2\right| = \left|\tau + 2\,\mathrm{Im} \int_U \nabla\varphi \cdot \overline{\mathrm{i}\tau\nabla\varphi}\,\mathrm{d}x\right| \\
&\leq \left|\tau + 2\,\mathrm{Im} \int_U \nabla\varphi \cdot \nabla\overline{\psi}\,\mathrm{d}x\right| + \left|2\,\mathrm{Im} \int_U \nabla\varphi \cdot \left(\overline{\mathrm{i}\tau\nabla\varphi - \nabla\psi}\right)\mathrm{d}x\right| \\
&\leq \varepsilon + 2\| |\nabla\varphi| \|_2 \cdot \| |\nabla(\mathrm{i}\tau\varphi - \psi)| \|_2 \leq \varepsilon + 2\|(\mathrm{i}\tau I - A)u\|_X \leq 3\varepsilon,
\end{aligned}$$

by means of (7.15) and the definition of A. As a result, $|\tau| \leq \frac{3\varepsilon\beta}{\beta - 2\varepsilon} < \tau_0$ which is impossible. We have shown that $m(\tau) \geq \varepsilon > 0$ for all $|\tau| \geq \tau_0$, as needed. $\diamondsuit$

We refer to Theorem VI.3.18 in [**EN99**] for a generalization of the above example. For a detailed study of the asymptotic behavior of C_0–semigroups we recommend the monograph [**vN96**] as well as the relevant chapters in [**ABHN11**] and [**EN99**].

Exercises

Exercise 7.1. In the context of the bounded perturbation Theorem 7.6, show that a strongly continuous function $R(\cdot) : \mathbb{R}_+ \to \mathcal{B}(X)$, which satisfies

$$R(t)x = T(t)x + \int_0^t T(t-s)BR(s)x \,\mathrm{d}s$$

for all $x \in X$ and $t \geq 0$, already coincides with the semigroup $S(\cdot)$ generated by $A + B$. Further show that

$$S(t) = \sum_{n=0}^{\infty} S_n(t) \quad \text{with } S_0(t) = T(t) \text{ and}$$
$$S_{n+1}(t)x = \int_0^t T(t-s)BS_n(s)x \,\mathrm{d}s$$

for $t \geq 0$, $n \in \mathbb{N}_0$ and $x \in X$, where the series converges in $\mathcal{B}(X)$ uniformly for t in compact subsets of $\mathbb{R}_+$.

Exercise 7.2. In Example 7.7 let $d = 3$ and replace the condition $0 \leq b \in L^\infty(U)$ by $0 \leq b \in L^q(U)$ for any $q \in (3, \infty)$. Show that then the operator A defined in Example 7.7 still generates a contraction semigroup.

Exercise 7.3. Let $U \subseteq \mathbb{R}^d$ be open, $X = C_0(U)$ and $m \in C(U)$ be such that $\sup_{x \in U} \operatorname{Re} m(x) =: a < \infty$. We set $T(t)f = \mathrm{e}^{tm} f$ for $t \geq 0$ and $f \in X$. From Exercise 2.4 we know that $T(\cdot)$ is a C_0–semigroup with generator A given by

$$Af = mf, \quad \mathrm{D}(A) = \{ f \in X \mid mf \in X \}.$$

Show that $T(\cdot)$ is exponentially stable if and only if $s(A) = a < 0$.

Let $a = 0$ and $\overline{m(U)} \subseteq \{\lambda \in \mathbb{C} \mid \operatorname{Re} \lambda < 0\}$. Show that the following assertions are equivalent for every $\alpha > 0$.

(a) There is $c_1 > 0$ such that $\|T(t)A^{-1}\| \leq c_1 t^{-\alpha}$ for all $t \geq 1$.
(b) There is $c_2 > 0$ such that $\|T(t)f\| \leq c_2 t^{-\alpha} \|f\|_A$ for all $t \geq 1$ and $f \in \mathrm{D}(A)$.
(c) There is $c_3 > 0$ such that $\|(\mathrm{i}\tau I - A)^{-1}\| \leq c_3 |\tau|^{\frac{1}{\alpha}}$ for all $\tau \in \mathbb{R}$ with $|\tau| \geq 1$.
(d) There are $c_4, \delta > 0$ such that $|\operatorname{Im} \lambda| \geq c_4 |\operatorname{Re} \lambda|^{-\alpha}$ for all $\lambda \in \sigma(A)$ with $\operatorname{Re} \lambda \geq -\delta$.

LECTURE 8

Local wellposedness of semilinear evolution equations

In the remaining lectures we investigate nonlinear problems. We begin with a relatively simple setting so that the new features and methods become more transparent. We study semilinear equations governed by a (linear) generator A and a (nonlinear) map $F : X \to X$ with certain Lipschitz properties.

In the first main result of the present lecture we prove local wellposedness of such evolution equations and give a general criterion for global existence. This result basically relies on the contraction mapping principle. But the proof involves several additional arguments and techniques which are typical for the field. Actually, this result is concerned with "mild solutions" and we will construct "classical solutions" in the second theorem of this lecture.

Our main application, the nonlinear wave equation (1.4), will be studied in the next lecture. There and in many other cases nonlinear equations arise if one replaces linear material laws (which often are valid only for "small" states) by nonlinear ones.

Let A generate a C_0–semigroup $T(\cdot)$ on X and let $F : X \to X$ be Lipschitz (continuous) on bounded sets, i.e.,

$$\forall\, r > 0\ \exists\, L(r) > 0\ \forall\, x, y \in \overline{B}(0, r) : \quad \|F(x) - F(y)\| \le L(r)\|x - y\|. \tag{8.1}$$

Here we set $\overline{B}(x_0, r) = \{x \in X \mid \|x - x_0\| \le r\}$. Throughout, let $J = [0, b]$ for some $b \in (0, \infty)$ or $J = [0, b)$ for some $b \in (0, \infty]$. For a given initial value $u_0 \in X$, we study the semilinear evolution equation

$$u'(t) = Au(t) + F(u(t)), \quad t \in J, \qquad u(0) = u_0. \tag{8.2}$$

If A generates a C_0–group, one could also consider intervals J with $0 \in J$ and $\inf J < 0$. For simplicity we do not treat this case. In view of the inhomogeneous equations in Lecture 6, it is also reasonable to allow for nonlinearities $F : J \times X \to X$ depending on time. To streamline the exposition, also this generalization is not studied below (but see Exercise 8.2).

We point out that in (8.2) the time interval J is not given a priori. Below we introduce a "maximal existence interval" of u which may depend on u_0. In this context we recall a simple example from ordinary differential equations: Let $X = \mathbb{C}$, $A = 0$, $F(u) = u^2$ and $u_0 > 0$. Then the problem

$$u'(t) = u(t)^2, \quad t \in J, \qquad u(0) = u_0,$$

is solved by $u(t) = (u_0^{-1} - t)^{-1}$, $t \in [0, u_0^{-1}) = J$, which "explodes" as $t \to u_0^{-1}$.

We first extend our solution concept to the new setting in a straight forward way. In the literature our "solutions" are often called "classical solutions".

Definition 8.1. *Let A generate the C_0–semigroup $T(\cdot)$, $u_0 \in \mathrm{D}(A)$, and $F : X \to X$ be continuous. A* solution *u of (8.2) is a function $u \in C^1(J, X)$ such that $u(t) \in \mathrm{D}(A)$ for all $t \in J$ and (8.2) holds.*

Clearly, if (8.2) has a solution u, then u_0 must belong to $\mathrm{D}(A)$ and $u \in C(J, [\mathrm{D}(A)])$. As for inhomogeneous linear problems, solutions are given by Duhamel's formula. In fact, if u solves (8.2), then the function $f = F \circ u : J \to X$ is continuous provided that $F : X \to X$ is continuous. Hence, Proposition 6.5 directly implies the next result.

Proposition 8.2. *Let A generate the C_0–semigroup $T(\cdot)$ on X, let $F : X \to X$ be continuous, and $u_0 \in \mathrm{D}(A)$. If $u \in C^1(J, X)$ solves (8.2), then it satisfies*

$$u(t) = T(t)u_0 + \int_0^t T(t-s)F(u(s))\,\mathrm{d}s, \quad t \in J. \tag{8.3}$$

For the development of the theory it is crucial to note that (8.3) is a fixed point equation for any given $u_0 \in X$, which makes sense for $u \in C(J, X)$. This observation leads us to the next definition.

Definition 8.3. *Let A generate the C_0-semigroup $T(\cdot)$ on X, $u_0 \in X$, and $F : X \to X$ be continuous. A* mild solution *of (8.2) is a function $u \in C(J, X)$ satisfying (8.3).*

Guided by the Picard-Lindelöf theorem for ordinary differential equations, we will solve the problem (8.3) by means of Banach's fixed point theorem, where we assume that (8.1) holds. For any given $u_0 \in X$, the right-hand side of (8.3) will define the nonlinear map Φ whose fixed point we want to construct. As the space E for the fixed point problem we will choose the closed ball with center 0 and radius r in $C([0, b], X)$ for suitable $b, r > 0$. By choosing a set E of uniformly bounded functions here, we achieve that F acts *globally* Lipschitz on E (with constant $L(r)$) due to (8.1). The operator Φ becomes strictly contractive if we take a sufficiently small $b = b(u_0) > 0$. It turns out that one can choose a time b which only depends on $\|u_0\|$. In view of more complicated equations, we stress that one should be careful with the constants here. They must be under control as b tends to 0, and one should specify how they depend on u_0.

Lemma 8.4. *Let A generate the C_0–semigroup $T(\cdot)$ on X and let $F : X \to X$ satisfy (8.1). Set $M_0 = \sup_{0 \le t \le 1} \|T(t)\| \in [1, \infty)$. Take any $\rho > 0$. Then there is a number $b_0(\rho) > 0$ (see (8.7) below) such that for each $u_0 \in \overline{B}(0, \rho)$ there is a mild solution $u \in C([0, b_0(\rho)], X)$ of (8.2). Moreover, for every $b \in (0, b_0(\rho)]$ the restriction $u|_{[0,b]}$ is the unique mild solution of (8.2) on $[0, b]$ with the initial value u_0 that satisfies $\|u(t)\| \le 1 + M_0\rho$ for all $0 \le t \le b$.*

Proof. Let $\rho > 0$ and take $u_0 \in X$ with $\|u_0\| \le \rho$. Fix $r := 1 + M_0\rho$. Take $0 < b \le 1$ to be specified below. Define the closed ball

$$E(b) := \{u \in C([0, b], X) \mid \|u\|_\infty \le r\}.$$

We note that $E(b)$ is a complete metric space for the metric induced by the sup-norm $\|\cdot\|_\infty$ on $C([0, b], X)$. We further introduce the map

$$[\Phi_{u_0}(u)](t) := \Phi(u)(t) := T(t)u_0 + \int_0^t T(t-s)F(u(s))\,\mathrm{d}s \tag{8.4}$$

for $t \in [0,b]$ and $u \in E(b)$. Clearly, $\Phi(u) \in C([0,b],X)$. We point out that each mild solution $v \in E(b)$ of (8.2) is a fixed point of Φ on $E(b)$ and vice versa.

Let $u, v \in E(b)$. Using (8.1) and that $u(s), v(s) \in \overline{B}(0,r)$, we estimate

$$\begin{aligned} \|\Phi(u)(t)\| &\le M_0\|u_0\| + \int_0^t M_0(\|F(u(s)) - F(0)\| + \|F(0)\|)\,\mathrm{d}s \\ &\le M_0\rho + bM_0(L(r)r + \|F(0)\|). \end{aligned} \tag{8.5}$$

Moreover,

$$\|\Phi(u)(t) - \Phi(v)(t)\| \le \int_0^t M_0\|F(u(s)) - F(v(s))\|\,\mathrm{d}s \le bM_0L(r)\|u-v\|_\infty \tag{8.6}$$

for all $0 \le t \le b \le 1$. We define

$$b_0(\rho) = \min\left\{1, \frac{1}{M_0(L(r)r + \|F(0)\|)}, \frac{1}{2M_0L(r)}\right\} \in (0,1]. \tag{8.7}$$

For every $b \in (0, b_0(\rho)]$, it follows that $\Phi(u) \in E(b)$ and that Φ is Lipschitz on $E(b)$ with Lipschitz constant smaller than or equal to $\frac{1}{2}$. Banach's fixed point theorem then gives a unique fixed point $u_b = \Phi(u_b) \in E(b)$ for each $0 < b \le b_0(\rho)$. Hence, u_b is the unique mild solution of (8.2) belonging to $E(b)$. We set $u = u_{b_0(\rho)}$ and note that u_b is the restriction $u|_{[0,b]} \in E(b)$ of u, due to uniqueness. □

In the following proofs we will often shift or "glue together" mild solutions. These procedures are justified in our second lemma.

LEMMA 8.5. *Let A generate the C_0–semigroup $T(\cdot)$ on X and let $F : X \to X$ be continuous. Assume that $u \in C([0,b_1],X)$ is a mild solution of (8.2) on $[0,b_1]$ with the initial value u_0. Then the following assertions hold.*

(a) If $v \in C([0,b_2],X)$ is a mild solution of (8.2) on $[0,b_2]$ with the initial value $u(b_1)$, then the function $w \in C([0,b_1+b_2],X)$ given by

$$w(t) = \begin{cases} u(t), & 0 \le t < b_1, \\ v(t-b_1), & b_1 \le t \le b_1 + b_2, \end{cases}$$

is a mild solution of (8.2) on $[0,b_1+b_2]$ with the initial value u_0.

(b) Let $\beta \in (0,b_1)$. Then the function $u(\,\cdot\,+\beta) \in C([0,b_1-\beta],X)$ is a mild solution of (8.2) with the initial value $u(\beta)$.

PROOF. (a) By its definition, w is continuous and it is a mild solution of (8.2) for $t \in [0,b_1]$. For $t \in (b_1, b_1+b_2]$ we calculate

$$\begin{aligned} w(t) = v(t-b_1) &= T(t-b_1)u(b_1) + \int_0^{t-b_1} T(t-b_1-s)F(v(s))\,\mathrm{d}s \\ &= T(t-b_1)T(b_1)u_0 + T(t-b_1)\int_0^{b_1} T(b_1-s)F(u(s))\,\mathrm{d}s \\ &\quad + \int_{b_1}^t T(t-r)F(v(r-b_1))\,\mathrm{d}r \\ &= T(t)u_0 + \int_0^t T(t-s)F(w(s))\,\mathrm{d}s, \end{aligned}$$

where we used the representation (8.3) of $v(t-b_1)$ and $u(b_1)$.

(b) Set $\varphi(t) = u(t+\beta)$ for $t \in [0, b_1 - \beta]$. As above, we obtain

$$\begin{aligned}\varphi(t) = u(t+\beta) &= T(t+\beta)u_0 + \int_0^{t+\beta} T(t+\beta-s)F(u(s))\,\mathrm{d}s \\ &= T(t)\Big(T(\beta)u_0 + \int_0^{\beta} T(\beta-s)F(u(s))\,\mathrm{d}s\Big) + \int_0^t T(t-r)F(u(r+\beta))\,\mathrm{d}r \\ &= T(t)u(\beta) + \int_0^t T(t-s)F(\varphi(s))\,\mathrm{d}s. \qquad \square\end{aligned}$$

We next upgrade the above basic existence lemma to a full local wellposedness theorem. To that purpose, for each initial value $u_0 \in X$ we define its *maximal existence time*

$$t^+(u_0) = \sup\left\{b > 0 \,\middle|\, \exists \text{ mild solution } u \in C([0,b], X) \text{ of } (8.2)\right\}.$$

Lemma 8.4 implies that the above set is non-empty and that $t^+(u_0) \in (0, \infty]$. A mild solution $u \in C([0, t^+(u_0)), X)$ of (8.2) with initial value u_0 is called *maximal solution*. Note that the existence interval of this solution has to be right-open due to Theorem 8.6 (b) below.

Our wellposedness theorem below says that the maximal solution is unique and that it depends locally Lipschitz on u_0. Moreover, we characterize the case that $t^+(u_0) = \infty$ and describe how $t^+(u_0)$ depends on u_0.

We point out that in all arguments one has to make sure that one uses the solution $u(t)$ only for $t < t^+(u_0)$ (unless one knows that $t^+(u_0) = \infty$). To obtain uniform bounds, one often restricts the solution to a compact time interval $[0, b]$ for a fixed $b \in (0, t^+(u_0))$.

THEOREM 8.6. *Let A generate the C_0–semigroup $T(\cdot)$ on X and let $F : X \to X$ be Lipschitz on bounded sets. Let $u_0 \in X$ and let $b_0(\|u_0\|) > 0$ be defined by* (8.7). *Then the following assertions hold.*

(a) There is a unique maximal mild solution $u = u(\cdot\,; u_0) \in C([0, t^+(u_0)), X)$ of (8.2)*, where $t^+(u_0) \in (b_0(\|u_0\|), \infty]$.*

(b) If $t^+(u_0) < \infty$, then $\lim_{t \to t^+(u_0)^-} \|u(t)\| = \infty$.

(c) Take any $b \in (0, t^+(u_0))$. Then there exists a radius $\delta > 0$ such that $t^+(v_0) > b$ for all $v_0 \in \overline{B}(u_0, \delta)$. Moreover, the map

$$\overline{B}(u_0, \delta) \to C([0,b], X), \quad v_0 \mapsto u(\cdot\,; v_0),$$

is Lipschitz continuous.

PROOF. (a) Using Lemmas 8.4 and 8.5, we can extend any solution u from $[0, b_0(\|u_0\|)]$ to a larger interval, so that $t^+(u_0) > b_0(\|u_0\|)$. Let u and v be mild solutions of (8.2) on intervals $J_v, J_u \subseteq [0, t^+(u_0))$ containing 0. We claim that either $J_u \subseteq J_v$ and $u = v|_{J_u}$, or vice versa. If this were not true, then there is a time $\tau \in J_u \cap J_v$ such that $u(t) = v(t)$ for all $t \in [0, \tau]$ and there are $t_n \in J_u \cap J_v$ converging to τ as $n \to \infty$ such that $u(t_n) \neq v(t_n)$ for all $n \in \mathbb{N}$. As in Lemma 8.4, we set $M_0 = \sup_{0 \le t \le 1} \|T(t)\| \ge 1$, $\rho = \|u(\tau)\|$ and $r = 1 + M_0\rho$. Since $u(t) \to u(\tau)$ and $v(t) \to u(\tau)$ as $t \to \tau^+$, we can find a time $b \in (0, b_0(\rho)]$ such that both functions $u(\cdot + \tau)$ and $v(\cdot + \tau)$ are mild solutions of (8.2) on $[0, b]$ with the initial value $u(\tau)$ and such that

$$\|u(s+\tau)\| \le r \quad \text{and} \quad \|v(s+\tau)\| \le r$$

for all $s \in [0,b]$. The uniqueness part of Lemma 8.4 now yields that $u(t) = v(t)$ for all $t \in [\tau, \tau + b]$ which contradicts $u(t_n) \neq v(t_n)$ for sufficiently large $n \in \mathbb{N}$. Hence, the uniqueness assertion holds.

To define a maximal solution, we take $b_n \to t^+(u_0)^-$ with corresponding mild solutions $u_n \in C([0,b_n],X)$ of (8.2). We then set $u(t;u_0) = u_n(t)$ for $t \in [0,b_n]$. This function is well defined on $[0,t^+(u_0))$ by what we have just shown. It is thus the unique mild solution of (8.2) on $[0,t^+(u_0))$.

(b) Let $t^+(u_0) < \infty$ and $u = u(\,\cdot\,;u_0)$. Assume that there were $b_n < t^+(u_0)$, $n \in \mathbb{N}$ such that $b_n \to t^+(u_0)^-$ as $n \to \infty$ and $C := \sup_{n\in\mathbb{N}} \|u(b_n)\| < \infty$. Fix an index $n \in \mathbb{N}$ with $b_n + b_0(C) > t^+(u_0)$. Lemma 8.4 gives a mild solution v of (8.2) on $[0,b_0(C)]$ the with initial value $u(b_n)$. By means of Lemma 8.5, we thus obtain a mild solution of (8.2) on $[0,b_n+b_0(C)]$, contradicting the defining property of $t^+(u_0)$. As a result, $\|u(t)\| \to \infty$ as $t \to t^+(u_0)^-$.

(c) Let $b \in (0,t^+(u_0))$ and set $u = u(\,\cdot\,;u_0)$. We consider the radius $\overline{\rho} = 1 + \sup_{0\le t\le b} \|u(t)\|$. Further let $M_0 \ge 1$ and Φ be defined as in (the proof of) Lemma 8.4. This lemma then yields $b_0(\overline{\rho}) =: \overline{b}$ such that for every $v_0, w_0 \in \overline{B}(0,\overline{\rho})$ the maximal mild solutions $v := u(\,\cdot\,;v_0) = \Phi_{v_0}(v)$ and $w = u(\,\cdot\,;w_0) = \Phi_{w_0}(w)$ exist at least on $[0,\overline{b}]$. In view of (8.4), (8.6) and (8.7), we can estimate

$$\begin{aligned}\|v-w\|_\infty &\le \|\Phi_{v_0}(v) - \Phi_{v_0}(w)\|_\infty + \|\Phi_{v_0}(w) - \Phi_{w_0}(w)\|_\infty \\ &\le \tfrac{1}{2}\|v-w\|_\infty + \|T(\cdot)(v_0-w_0)\|_\infty \\ &\le \tfrac{1}{2}\|v-w\|_\infty + M_0\|v_0-w_0\|,\end{aligned}$$

where $\|\cdot\|_\infty$ denotes the sup-norm on $[0,\overline{b}]$. It follows

$$\|v-w\|_\infty \le 2M_0\|v_0-w_0\|. \tag{8.8}$$

For $j \in \mathbb{N}_0$ we set $b_j = j\overline{b}$. Then there exists a minimal $N \in \mathbb{N}$ with $b_N \ge b$. If $b_N > t^+(u_0)$, we replace b_N by any number in $(b,t^+(u_0))$.

We choose $\delta := (2M_0)^{-N}$. We inductively show that for every $v_0 \in \overline{B}(u_0,\delta)$ and $j \in \{0,\dots,N-1\}$ the maximal mild solution $v = u(\,\cdot\,;v_0)$ exists at least on $[0,b_j]$ and that $v(b_j) \in \overline{B}(0,\overline{\rho})$. We then infer that v even exists on $[0,b_{j+1}]$, due to Lemmas 8.4 and 8.5. For $j = 0$ the claim holds since $\delta < 1$. Next, assume it also holds for all $k \in \{0,\dots,j-1\}$. Again from (8.8) and Lemma 8.5 we deduce

$$\begin{aligned}\|v(b_j)\| &\le \|v(b_j) - u(b_j)\| + \|u(b_j)\| \\ &\le \|u(\,\cdot\,;v(b_{j-1})) - u(\,\cdot\,;u(b_{j-1}))\|_\infty + \|u(b_j)\| \\ &\le 2M_0\|v(b_{j-1}) - u(b_{j-1})\| + \|u(b_j)\| \\ &\le \ldots \le (2M_0)^j\|v_0-u_0\| + \|u(b_j)\| \le (2M_0)^{j-N} + \sup_{t\in[0,b]}\|u(t)\| < \overline{\rho}.\end{aligned}$$

So we have $t^+(v_0) > b_N \ge b$ which is the first assertion in (c).

To show the asserted Lipschitz continuity, we note that for $v_0, w_0 \in \overline{B}(u_0,\delta)$ the vectors $v(b_j)$ and $w(b_j)$ with $j = 0,\dots,N-1$ stay in $\overline{B}(0,\overline{\rho})$. We can thus apply (8.8) repeatedly and derive (c) in this way. □

We add a simple example for Theorem 8.6. In Lecture 9 we discuss a more sophisticated application.

Example 8.7. Let $\varphi : \mathbb{C} \to \mathbb{C}$ be Lipschitz with constant L and let $\varphi(0) = 0$. Observe that then $|\varphi(z)| \le L|z|$ for all $z \in \mathbb{C}$. We define the substitution operator F by $F(u) = \varphi(u)$ for $u \in L^2(\mathbb{R}^d)$. Then $F(u) \in L^2(\mathbb{R}^d)$ and

$$\begin{aligned}\|F(u) - F(v)\|_2^2 &= \int_{\mathbb{R}^d} |\varphi(u(x)) - \varphi(v(x))|^2 \,\mathrm{d}x \\ &\le L^2 \int_{\mathbb{R}^d} |u(x) - v(x)|^2 \,\mathrm{d}x = L^2 \|u - v\|_2^2\end{aligned}$$

for all $u, v \in L^2(\mathbb{R}^d)$ so that $F : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is Lipschitz with constant L. We can thus apply Theorem 8.6 to the nonlinear Schrödinger equation

$$u'(t) = \mathrm{i}\Delta u(t) + \mathrm{i}F(u(t)), \quad t \in J, \qquad u(0) = u_0,$$

for $u_0 \in L^2(\mathbb{R}^d)$ and derive the local wellposedness of mild solutions in the sense of Theorem 8.6. Since $\|F(u)\|_2 \le L\|u\|_2$, we even obtain $t^+(u_0) = \infty$ for all $u_0 \in L^2(\mathbb{R}^d)$, due to Exercise 8.1. We point out that the setting of this lecture does not allow to take locally Lipschitz φ. For instance, for $\varphi(z) = |z|^2$ the operator $F : u \mapsto |u|^2$ does *not* map $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^d)$. ♢

If $u_0 \in \mathrm{D}(A)$, we can hope that our mild solution u of (8.2) in fact solves (8.2). In particular, the initial regularity $u_0 \in \mathrm{D}(A)$ should be preserved by the solution $u(\cdot\,; u_0)$. To show such a result, we need some preparations.

Lemma 8.8. *Let A generate a C_0–semigroup on X, $u_0 \in \mathrm{D}(A)$, and $F : X \to X$ be Lipschitz on bounded sets. Then the maximal mild solution $u = u(\cdot\,; u_0) : [0, t^+(u_0)) \to X$ of (8.2) is locally Lipschitz continuous.*

Proof. Take $b \in [0, t^+(u_0))$ and $0 \le t \le t + h \le b$. Equation (8.3) leads to

$$\begin{aligned} u(t+h) - u(t) &= T(t)(T(h)u_0 - u_0) + \int_0^h T(t+h-s)F(u(s))\,\mathrm{d}s \\ &\quad + \int_0^t T(t-s)F(u(s+h))\,\mathrm{d}s - \int_0^t T(t-s)F(u(s))\,\mathrm{d}s \\ &= \int_0^h T(t+s)Au_0\,\mathrm{d}s + \int_0^h T(t+h-s)F(u(s))\,\mathrm{d}s \\ &\quad + \int_0^t T(t-s)(F(u(s+h)) - F(u(s)))\,\mathrm{d}s. \end{aligned} \tag{8.9}$$

Observe that the numbers $r = \sup_{0\le s\le b} \|u(s)\|$, $M_0 = \sup_{0\le s\le b} \|T(s)\|$ and $C = \sup_{0\le s\le b} \|F(u(s))\|$ are all finite. Formula (8.9) combined with (8.1) yield

$$\|u(t+h) - u(t)\| \le M_0\|Au_0\|h + M_0Ch + M_0L(r)\int_0^t \|u(s+h) - u(s)\|\,\mathrm{d}s.$$

Gronwall's inequality then implies that

$$\|u(t+h) - u(t)\| \le M_0(\|Au_0\| + C)\mathrm{e}^{M_0L(r)b}h,$$ □

In our regularity theorem we will require that F is real continuously differentiable. To that purpose, we define

$$\mathcal{B}_{\mathbb{R}}(X,Y) := \Big\{T : X \to Y \,\big|\, T \text{ is } \mathbb{R}\text{-linear and } \|T\|_{\mathcal{B}_{\mathbb{R}}(X,Y)} := \sup_{\|x\|\le 1} \|Tx\| < \infty\Big\},$$

recalling that the Banach spaces X and Y are complex. As for $\mathcal{B}(X,Y)$ (the space of bounded $\mathbb{C}$-linear operators) one shows that $\mathcal{B}_\mathbb{R}(X,Y)$ is a Banach space when endowed with $\|\cdot\|_{\mathcal{B}_\mathbb{R}(X,Y)}$. Each $T\in\mathcal{B}_\mathbb{R}(X,Y)$ is Lipschitz continuous. We clearly have $\mathcal{B}(X,Y)\subset\mathcal{B}_\mathbb{R}(X,Y)$, but the converse inclusion is false even for $X=Y=\mathbb{C}$. As usual we write $\mathcal{B}_\mathbb{R}(X):=\mathcal{B}_\mathbb{R}(X,X)$.

Let $\emptyset\neq D\subseteq X$ be open. A map $F:D\to Y$ is called *real (Fréchet) differentiable* at $x_0\in D$ if there is an operator $S\in\mathcal{B}_\mathbb{R}(X,Y)$ such that the limit

$$\lim_{\substack{h\to 0,\, h\neq 0,\\ x_0+h\in D}} \frac{1}{\|h\|}\|F(x_0+h)-F(x_0)-Sh\|=0$$

exists. We then set $F'(x_0):=S$ and call $F'(x_0)$ the *(Fréchet) derivative* of F at x_0.We say that F is *real continuously differentiable* on D if F is real differentiable at each point of D and the function

$$F':D\to\mathcal{B}_\mathbb{R}(X,Y),\quad x\mapsto F'(x),$$

is continuous. In this case we write $F\in C^1_\mathbb{R}(D,Y)$. The usual rules of calculus (including the chain rule) hold in this setting with analogous proofs and straightforward modifications. (See, e.g., Chapter XIII in [**Lan93**].) If D is convex and $F\in C^1_\mathbb{R}(D,Y)$, we have

$$F(z)-F(x)=\int_0^1\frac{\mathrm{d}}{\mathrm{d}t}F(x+t(z-x))\,\mathrm{d}t=\int_0^1 F'(x+t(z-x))(z-x)\,\mathrm{d}t. \quad (8.10)$$

In this situation we thus obtain

$$\|F(z)-F(x)\|\le\max_{0\le t\le 1}\|F'(x+t(z-x))\|\,\|z-x\| \quad (8.11)$$

for all $z,x\in D$. As a result, a function $F\in C^1_\mathbb{R}(X,X)$ is Lipschitz on bounded sets provided that its derivative is bounded on bounded sets. (Observe that a continuous function on a Banach space does not need to be bounded on a closed ball.) We establish a final prerequisite.

LEMMA 8.9. *Let $u\in C([a,b),X)$ be differentiable from the right with right-hand side derivative $v\in C([a,b),X)$. Then $u\in C^1([a,b),X)$ and $u'=v$.*

PROOF. Let $x^*\in X^*$. Then $\varphi(t)=\langle u(t),x^*\rangle$ satisfies the assumptions for $X=\mathbb{C}$ (with right-hand side derivative $x^*\circ v$). Corollary 2.1.2 of [**Paz83**] implies that $\varphi\in C^1([a,b))$ with $\varphi'=x^*\circ v$. Fix $h\in(0,b-a)$ and take $t\in[a+h,b)$. Due to the Hahn-Banach theorem, there exists a functional $x_h^*\in X^*$ such that $\|x_h^*\|=1$ and

$$\big|\langle\tfrac{1}{h}(u(t)-u(t-h))-v(t),x_h^*\rangle\big|=\|\tfrac{1}{h}(u(t)-u(t-h))-v(t)\|=:D_h(t).$$

Setting $\varphi_h(t):=\langle u(t),x_h^*\rangle$, we then compute

$$\begin{aligned}D_h(t)&=\big|\tfrac{1}{h}(\varphi_h(t)-\varphi_h(t-h))-\varphi_h'(t)\big|=\Big|\frac{1}{h}\int_{t-h}^t(\varphi_h'(\tau)-\varphi_h'(t))\,\mathrm{d}\tau\Big|\\&=\Big|\frac{1}{h}\int_{t-h}^t\langle v(\tau)-v(t),x_h^*\rangle\,\mathrm{d}\tau\Big|\le\tfrac{h}{h}\max_{t-h\le\tau\le t}\|v(\tau)-v(t)\|.\end{aligned}$$

Since the right-hand side tends to 0 as $h\to 0$, we obtain that u is differentiable at each $t\in[a,b)$ with the (continuous) derivative v. □

In the next proof we follow a standard strategy to prove additional regularity of a given (mild) solution. Assume for a moment that our mild solution u were in fact a solution of (8.2) in $C^1(J,X)$. One can then differentiate (8.2) with respect to t and obtain a linear (non-autonomous) evolution equation for $v := u'$ with the initial value $u'(0) = Au_0 + F(u_0)$. Assuming $u_0 \in \mathrm{D}(A)$, we can now pass to the integrated version of this equation (see (8.12) below), which is easy to solve in our case. The resulting solution v is a candidate for the derivative of u. To verify the differentiability of u, we then rewrite the difference quotient of u by means of (8.3) and subtract the equation (8.12) for v. A Gronwall type estimate finally yields the assertion.

THEOREM 8.10. *Let A generate the C_0–semigroup $T(\cdot)$ on X, $u_0 \in \mathrm{D}(A)$, $F \in C^1_{\mathbb{R}}(X,X)$ and assume that F' is bounded on bounded sets. Then the maximal mild solution $u = u(\,\cdot\,; u_0)$ of* (8.2) *in fact solves* (8.2) *on* $[0, t^+(u_0))$.

PROOF. Let $u_0 \in \mathrm{D}(A)$ and $F \in C^1_{\mathbb{R}}(X,X)$. Let $b \in (0, t^+(u_0))$ be arbitrary. We have to show that $u \in C^1([0,b), X)$ since then $F \circ u \in C^1([0,b), X)$ and thus the assertion will follow from Theorem 6.9 and (8.3). Set $M_0 = \sup_{0 \le s \le b} \|T(s)\|$.

1) We first prove a preliminary result. The operators $B(s) := F'(u(s)) \in \mathcal{B}_{\mathbb{R}}(X)$ depend continuously on $s \in [0,b]$ and $L := \sup_{0 \le s \le b} \|B(s)\|$ is finite. The ($\mathbb{R}$-linear non-autonomous) problem

$$v(t) = T(t)(F(u_0) + Au_0) + \int_0^t T(t-s)B(s)v(s)\,\mathrm{d}s \tag{8.12}$$

can be solved as in Lemma 8.4 for $t \in [0, b_0]$ and a sufficiently small $b_0 > 0$ by a fixed point argument on $C([0,b_0], X)$ (using that each $B(s)$ is Lipschitz on X with Lipschitz constant less than or equal to L). Since (8.5) is not needed here, the analogue of equation (8.6) allows to choose $b_0 = \min\{1, \frac{1}{2M_0 L}\}$ independently of the initial value. As a result, we can solve (8.12) on $[0, b_0]$ with $F(u_0) + Av_0$ replaced by $v(b_0)$ and thus obtain a solution of (8.12) on $[0, 2b_0]$ as in Lemma 8.5. In finitely many steps we then construct a solution $v \in C([0,b], X)$ of (8.12). (See also Exercise 8.2.)

2) We now show that the function v of step 1) is the derivative of u. Let $0 \le t \le t+h \le b$ for some $h > 0$. Equations (8.3) and (8.12) imply that

$$\begin{aligned}
w_h(t) &:= \tfrac{1}{h}(u(t+h) - u(t)) - v(t)\\
&= T(t)\tfrac{1}{h}(T(h) - I)u_0 - T(t)Au_0\\
&\qquad + \frac{1}{h}\int_0^h T(t+h-s)F(u(s))\,\mathrm{d}s - T(t)F(u_0)\\
&\qquad + \int_0^t T(t-s)\big[\tfrac{1}{h}(F(u(s+h)) - F(u(s))) - F'(u(s))v(s)\big]\,\mathrm{d}s\\
&=: S_1(h,t) + S_2(h,t) + S_3(h,t).
\end{aligned}$$

We first observe that

$$\|S_1(h,t)\| \le M_0 \|\tfrac{1}{h}(T(h) - I)u_0 - Au_0\| =: \alpha_1(h) \;\longrightarrow\; 0,$$

$$\|S_2(h,t)\| = \Big\|T(t)\frac{1}{h}\int_0^h (T(h-s)F(u(s)) - F(u_0))\,\mathrm{d}s\Big\|$$
$$\le M_0\tfrac{h}{h}\sup_{0\le s\le h}\|T(h-s)F(u(s)) - F(u_0)\| =: \alpha_2(h) \;\longrightarrow\; 0$$

as $h \to 0^+$. Here we use $u_0 \in \mathrm{D}(A)$ in the first limit and Lemma 2.9 for the second one. We then write

$$S_3(h,t) = \int_0^t T(t-s)\tfrac{1}{h}[F(u(s+h)) - F(u(s)) - F'(u(s))(u(s+h)-u(s))]\,\mathrm{d}s$$
$$+ \int_0^t T(t-s)F'(u(s))w_h(s)\,\mathrm{d}s =: S_{3,1}(h,t) + S_{3,2}(h,t).$$

By Lemma 8.8 the function u is Lipschitz on $[0,b]$. Denote its Lipschitz constant by ℓ. Employing this fact and (8.10), we estimate $\|S_{3,1}(h,t)\|$ by

$$M_0 b \sup_{\substack{0\le s\le b\\ 0\le s+h\le b}} \frac{1}{h}\Big\|\int_0^1 [F'(u(s)+\tau(u(s+h)-u(s))) - F'(u(s))](u(s+h)-u(s))\,\mathrm{d}\tau\Big\|$$
$$\le M_0 b\,\ell\,\tfrac{h}{h} \sup_{\substack{0\le s\le b\\ 0\le s+h\le b\\ 0\le\tau\le 1}} \|F'(u(s)+\tau(u(s+h)-u(s))) - F'(u(s))\| =: \alpha_3(h).$$

Here $\alpha_3(h)\to 0$ as $h\to 0^+$ since F' is uniformly continuous on the compact set

$$\big\{u(s)+\tau(u(r)-u(s)) \,\big|\, 0\le\tau\le 1, 0\le r,s\le b\big\}\,.$$

Altogether we have shown

$$\|w_h(t)\| \le \alpha_1(h) + \alpha_2(h) + \alpha_3(h) + M_0 L\int_0^t \|w_h(s)\|\,\mathrm{d}s.$$

Gronwall's inequality thus yields

$$\|w_h(t)\| \le (\alpha_1(h)+\alpha_2(h)+\alpha_3(h))\mathrm{e}^{tM_0L}$$

for all $t\in[0,b]$. Letting $h\to 0^+$, we then derive that u is differentiable from the right and that the right-hand side derivative coincides with v. Since v is continuous on $[0,b]$, Lemma 8.9 implies $u\in C^1([0,b),X)$. □

Exercises

Exercise 8.1. In the setting of Theorem 8.6 in addition assume that F is "linearly bounded", i.e., that there exists a $c > 0$ such that $\|F(x)\| \le c(1+\|x\|)$ for all $x \in X$. Show that then $t^+(u_0) = \infty$ for each $u_0 \in X$.

Exercise 8.2. Let A generate the C_0–semigroup $T(\cdot)$, $J = [0,T]$ and $F : J \times X \to X$ be continuous. Assume that there is an $L > 0$ such that

$$\|F(t,x) - F(t,y)\| \le L\|x-y\|$$

for all $x, y \in X$ and $t \in J$. Let $u_0 \in X$ and $s \in J$. Show that there is a unique solution $u = u(\,\cdot\,; s, u_0) \in C([s,T], X)$ of the equation

$$u(t) = T(t-s)u_0 + \int_s^t T(t-\tau)F(\tau, u(\tau))\,\mathrm{d}\tau, \quad t \in [s,T].$$

In addition, let $x \mapsto F(s,x)$ be linear for each $s \in [0,T]$. We set $U(t,s)u_0 = u(t; s, u_0)$ for $0 \le s \le t \le T$ and $u_0 \in X$. Show that

i) $U(t,s) \in \mathcal{B}(X)$ and $\sup_{0\le s\le t\le T} \|U(t,s)\| < \infty$,
ii) $U(t,t) = I$ and $U(t,r)U(r,s) = U(t,s)$ for $0 \le s \le r \le t \le T$,
iii) the map $\{(t,s) \,|\, 0 \le s \le t \le T\} \to X, (t,s) \mapsto U(t,s)x$, is continuous for all $x \in X$.

Exercise 8.3. We know from Example 4.13 that the operator A given by $Au = \partial_x^2 u$ on $\mathrm{D}(A) = \{u \in C^2([0,1]) \,|\, \partial_x u(0) = \partial_x u(1) = 0\}$ generates a C_0–semigroup on $X = C([0,1])$. Let $\varphi \in C^1(\mathbb{R}, \mathbb{R})$ and set $F(u) := \varphi(\operatorname{Re} u)$ for all $u \in X$. Show the following assertions.

(a) $F \in C^1_{\mathbb{R}}(X, X)$ with derivative given by
$$F'(u)v = \varphi'(\operatorname{Re} u)\operatorname{Re} v, \qquad u, v \in X.$$
(b) The "reaction-diffusion equation"
$$u'(t) = Au(t) + F(u(t)), \quad t \in J, \qquad u(0) = u_0,$$
has for all $u_0 \in \mathrm{D}(A)$ a unique maximal solution $u \in C^1([0, t^+(u_0)), X)$ with $u(t) \in \mathrm{D}(A)$ for all $t \in [0, t^+(u_0))$. If u_0 is real valued, then also u is real valued.
(c) Let $\varphi(s) = s^2$. Find an initial function $u_0 \in X$ such that $t^+(u_0) < \infty$.
(d) Let $\varphi(s) = s(1-s)$ and $u_0 \in \mathrm{D}(A)$ with $0 \le u_0 \le 1$. Show that $t^+(u_0) = \infty$ and $0 \le u(t) \le 1$ for all $t \ge 0$. [Hint: Try a contradiction argument such as "Assume that there were $t_0 > 0$ and $x_0 \in [0,1]$ with $u(t_0, x_0) < 0 \ldots$". But the proof requires certain tricks.]

LECTURE 9

The nonlinear wave equation with cubic forcing term

This lecture is devoted to the investigation of the nonlinear wave equation with Dirichlet boundary conditions and a cubic forcing term on a bounded domain in $\mathbb{R}^3$. We first show that the theory of the previous lecture can be applied to this problem and thus derive local wellposedness and regularity of mild solutions. We then focus on the qualitative behavior of the solutions. Depending on the sign of the forcing term, we obtain either global existence or blow-up. As another important feature we establish the finite speed of propagation of the solutions to the wave equation. Finally, we construct a so-called "standing wave solution" of the nonlinear wave equation.

We start with the real differentiability of the nonlinearities F arising in this and later lectures. As often in partial differential equations, our applications lead to superposition operators of the form $F(u) = \phi(u)$ for a real differentiable function $\phi : \mathbb{C} \to \mathbb{C}$, where we have $F(u) = \mathrm{i}u|u|^{\alpha-1}$ for the nonlinear Schrödinger equation. These operators act on L^p-spaces of complex-valued functions. However, when differentiating F it is convenient to identify $\mathbb{C}$ with $\mathbb{R}^2$ and to consider ϕ as a function from $\mathbb{R}^2$ to $\mathbb{R}^2$.

To this aim, we introduce the following notations. Let $z \in \mathbb{C}$. For $\varphi : \mathbb{R}^2 \to \mathbb{R}$ and $\phi = (\phi_1, \phi_2) : \mathbb{R}^2 \to \mathbb{R}^2$, we define

$$\varphi(z) = \varphi(\operatorname{Re} z, \operatorname{Im} z) \in \mathbb{R} \quad \text{ and } \quad \phi(z) = \phi_1(z) + \mathrm{i}\phi_2(z) \in \mathbb{C}.$$

Moreover, for $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$ and $M = \left(\begin{smallmatrix} \xi \\ \eta \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$, we set

$$\xi \cdot z = \xi_1 \operatorname{Re} z + \xi_2 \operatorname{Im} z \in \mathbb{R} \quad \text{ and } \quad Mz = \xi \cdot z + \mathrm{i}\eta \cdot z \in \mathbb{C}.$$

Our first lemma will allow us to treat the part of the "energy" arising from the nonlinearity in our applications, whereas the second lemma is concerned with the nonlinearity itself.

LEMMA 9.1. *Let $\emptyset \neq U \subseteq \mathbb{R}^d$ be open. Let $\varphi \in C^1(\mathbb{R}^2, \mathbb{R})$ satisfy $|\varphi(z)| \leq c_0|z|^{1+\alpha}$ and $|\nabla\varphi(z)| \leq c_0|z|^\alpha$ for all $z \in \mathbb{C}$ and some constants $c_0 > 0$ and $\alpha \geq 1$. Then the map*

$$\Phi : L^{1+\alpha}(U) \to \mathbb{R}, \quad \Phi(u) = \int_U \varphi(u)\,\mathrm{d}x,$$

is real continuously differentiable. Its derivative $\Phi'(u) \in \mathcal{B}_\mathbb{R}(L^{1+\alpha}(U), \mathbb{C})$ at $u \in L^{1+\alpha}(U)$ is given by

$$\Phi'(u)v = \int_U \nabla\varphi(u) \cdot v\,\mathrm{d}x, \qquad v \in L^{1+\alpha}(U). \tag{9.1}$$

Moreover, $\|\Phi'(u)\|_{\mathcal{B}_\mathbb{R}(L^{1+\alpha}(U),\mathbb{C})} \leq c_0\|u\|_{1+\alpha}^\alpha$ for all $u \in L^{1+\alpha}(U)$ and thus Φ' is bounded on bounded sets.

Proof. Let $u, v \in L^{1+\alpha}(U)$. Set $p = \frac{1+\alpha}{\alpha}$. The growth assumptions on φ yield $\varphi(u) \in L^1(U)$ so that Φ maps $L^{1+\alpha}(U)$ into $\mathbb{R}$. Next, the map $v \mapsto \int_U \nabla\varphi(u) \cdot v \,\mathrm{d}x$ belongs to $\mathcal{B}_{\mathbb{R}}(L^{1+\alpha}(U), \mathbb{C})$ since it is $\mathbb{R}$-linear and

$$\Big|\int_U \nabla\varphi(u) \cdot v \,\mathrm{d}x\Big| \le \|\nabla\varphi(u)\|_p \|v\|_{1+\alpha} \le c_0 \|u\|_{1+\alpha}^{\alpha} \|v\|_{1+\alpha}, \tag{9.2}$$

due to Hölder's inequality and the growth assumption on $\nabla\varphi$. To check the asserted differentiability of Φ at u, we compute

$$\begin{aligned} &\varphi(u(x) + v(x)) - \varphi(u(x)) - \nabla\varphi(u(x)) \cdot v(x) \\ &\qquad = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\tau} \varphi(u(x) + \tau v(x)) \,\mathrm{d}\tau - \nabla\varphi(u(x)) \cdot v(x) \\ &\qquad = \int_0^1 (\nabla\varphi(u(x) + \tau v(x)) - \nabla\varphi(u(x))) \cdot v(x) \,\mathrm{d}\tau \end{aligned}$$

for a.e. $x \in U$. Integrating over U and using Fubini's theorem, we infer

$$D_v := \Phi(u+v) - \Phi(u) - \int_U \nabla\varphi(u) \cdot v \,\mathrm{d}x = \int_0^1 \int_U (\nabla\varphi(u+\tau v) - \nabla\varphi(u)) \cdot v \,\mathrm{d}x \,\mathrm{d}\tau.$$

As above, Hölder's inequality then implies that

$$|D_v| \le \|v\|_{1+\alpha} \int_0^1 \|\nabla\varphi(u + \tau v) - \nabla\varphi(u)\|_p \,\mathrm{d}\tau =: \|v\|_{1+\alpha} I(v).$$

We claim that $I(v) \to 0$ as $v \to 0$ in $L^{1+\alpha}(U)$, which means that Φ is real differentiable at u and (9.1) holds. Moreover, the asserted estimate for Φ' then follows from (9.2). The claim holds if for each null sequence $(v_n)_n$ in $L^{1+\alpha}(U)$ there is a subsequence $(v_{n_j})_j$ such that $I(v_{n_j}) \to 0$ as $j \to \infty$.

So let $v_n \in L^{1+\alpha}(U)$ converge to 0 in $L^{1+\alpha}(U)$ as $n \to \infty$. By (the proof of) the Riesz-Fischer theorem there are a subsequence $(v_{n_j})_j$ and a function $g \in L^{1+\alpha}(U)$ such that $v_{n_j} \to 0$ a.e. as $j \to \infty$ and $|v_{n_j}| \le g$ a.e. for all $j \in \mathbb{N}$. For $\tau \in [0, 1]$, the growth assumption on $\nabla\varphi$ thus implies the pointwise estimate

$$|\nabla\varphi(u + \tau v_{n_j}) - \nabla\varphi(u)| \le c_0(|u| + g)^{\alpha} + c_0|u|^{\alpha} =: f \in L^p(U).$$

The theorem of dominated convergence then shows that the integrand of $I(v_{n_j})$ tends to 0 as $j \to \infty$ for each $\tau \in [0, 1]$. Since this integrand is bounded by the constant $\|f\|_p$, we further derive that $I(v_{n_j}) \to 0$ as $j \to \infty$, as asserted.

In the same way one proves that $\Phi'(w) \to \Phi'(u)$ in $\mathcal{B}_{\mathbb{R}}(L^{1+\alpha}(U), \mathbb{C})$ as $w \to u$ in $L^{1+\alpha}(U)$, i.e., Φ' is continuous. $\square$

Lemma 9.2. *Let $\phi = (\phi_1, \phi_2) \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ satisfy $|\phi(z)| \le c_0|z|^{\alpha}$ and $|\phi'(z)| \le c_0|z|^{\alpha-1}$ for all $z \in \mathbb{C}$ and some constants $c_0 > 0$ and $\alpha > 1$, where $\phi'(z) = \phi'(\operatorname{Re} z, \operatorname{Im} z)$. Let $p \in [\alpha, \infty)$ and $\emptyset \ne U \subseteq \mathbb{R}^d$ be open. Then the map*

$$F : L^p(U) \to L^{\frac{p}{\alpha}}(U), \qquad F(u) = \phi(u) = \phi_1(u) + \mathrm{i}\phi_2(u),$$

is real continuously differentiable and its derivative at $u \in L^p(U)$ is given by

$$F'(u)v = \phi'(u)v = \nabla\phi_1(u) \cdot v + \mathrm{i}\nabla\phi_2(u) \cdot v, \qquad v \in L^p(U).$$

We further have $\|F'(u)\|_{\mathcal{B}_{\mathbb{R}}(L^p, L^{p/\alpha})} \le c_0 \|u\|_p^{\alpha-1}$, so that the derivative is bounded on bounded sets.

Proof. Let $u, v \in L^p(U)$. As in the previous lemma, the growth assumptions imply that $F(u) \in L^{\frac{p}{\alpha}}(U)$ and that $|\phi'(u)| \in L^q(U)$ with $q = \frac{p}{\alpha-1}$. Hölder's inequality with exponents $\frac{\alpha}{p} = \frac{1}{p} + \frac{\alpha-1}{p}$ yields that $v \mapsto \phi'(u)v$ belongs to $\mathcal{B}_{\mathbb{R}}(L^p(U), L^{\frac{p}{\alpha}}(U))$. We further obtain the pointwise identity

$$F(u+v) - F(u) - \phi'(u)v = \int_0^1 (\phi'(u+\tau v) - \phi'(u))v \,\mathrm{d}\tau.$$

Minkowski's inequality for integrals and Hölder's inequality then imply

$$\begin{aligned}\|F(u+v) - F(u) - \phi'(u)\cdot v\|_{\frac{p}{\alpha}} &\le \int_0^1 \|(\phi'(u+\tau v) - \phi'(u))v\|_{\frac{p}{\alpha}} \,\mathrm{d}\tau \\ &\le \|v\|_p \int_0^1 \|\phi'(u+\tau v) - \phi'(u)\|_{\frac{p}{\alpha-1}} \,\mathrm{d}\tau.\end{aligned}$$

The same arguments as in the previous proof show that the above integral converges to zero as $v \to 0$ in $L^p(U)$, where one has to use the growth conditions on ϕ'. Therefore $F : L^p(U) \to L^{\frac{p}{\alpha}}(U)$ is real differentiable and its derivative $F'(u)$ can be represented as asserted. The continuity of $u \mapsto F'(u)$ and the norm bound for $F'(u)$ follow as before. □

We now apply the above lemmas to the nonlinear maps used below. To use complex notation, for $z, w \in \mathbb{C}$ we set $z \cdot w = \operatorname{Re} z \operatorname{Re} w + \operatorname{Im} z \operatorname{Im} w \in \mathbb{R}$ and note that $\operatorname{Re}(z\overline{w}) = z \cdot w$.

Corollary 9.3. *Let $\alpha > 1$, $\beta \ge 1$, $p \in [\alpha, \infty)$ and let $U \subseteq \mathbb{R}^d$ be open. Then the maps*

$$\begin{aligned} &\Phi : L^{1+\beta}(U) \to \mathbb{R}, \qquad \Phi(u) = \frac{1}{1+\beta}\int_U |u|^{1+\beta} \,\mathrm{d}x, \\ &F : L^p(U) \to L^{\frac{p}{\alpha}}(U), \qquad F(u) = |u|^{\alpha-1}u, \end{aligned}$$

are real continuously differentiable, Lipschitz on bounded sets and their derivatives are given by

$$\begin{aligned} \Phi'(u)v &= \int_U |u|^{\beta-1}\operatorname{Re}(u\overline{v}) \,\mathrm{d}x \qquad &&\text{for } u, v \in L^{1+\beta}(U), \\ F'(u)v &= |u|^{\alpha-1}v + (\alpha-1)|u|^{\alpha-3}u\operatorname{Re}(u\overline{v}) &&\text{for } u, v \in L^p(U). \end{aligned}$$

Proof. Set $\varphi(z) = \frac{1}{1+\beta}|z|^{1+\beta}$ and $\phi(z) = |z|^{\alpha-1}z$ for $z \in \mathbb{C} \cong \mathbb{R}^2$. The growth conditions for φ and ϕ from Lemma 9.1 and 9.2 clearly hold. Writing $r = \operatorname{Re} z$ and $s = \operatorname{Im} z$ for $z \in \mathbb{C}$ and identifying z and (r, s), we obtain

$$\begin{aligned} \nabla\varphi(z) &= \frac{1}{1+\beta}\Big(\partial_r (r^2+s^2)^{\frac{\beta+1}{2}}, \partial_s (r^2+s^2)^{\frac{\beta+1}{2}}\Big)^\top = |z|^{\beta-1}z, \\ \phi'(z) &= \big(\partial_r\big((r^2+s^2)^{\frac{\alpha-1}{2}}\left(\begin{smallmatrix} r\\ s\end{smallmatrix}\right)\big), \partial_s\big((r^2+s^2)^{\frac{\alpha-1}{2}}\left(\begin{smallmatrix} r\\ s\end{smallmatrix}\right)\big)\big) \\ &= |z|^{\alpha-3}\begin{pmatrix} (\alpha-1)r^2 + |z|^2 & (\alpha-1)rs \\ (\alpha-1)rs & (\alpha-1)s^2 + |z|^2 \end{pmatrix} \quad \text{for } z \neq 0, \end{aligned}$$

and $\phi'(0) = 0$. As a result, also $\nabla\varphi$ and ϕ' satisfy the growth assumptions of the two previous lemmas. Moreover, for $w \in \mathbb{C}$ with $\rho = \operatorname{Re} w$ and $\sigma = \operatorname{Im} w$, we compute $\nabla\varphi(z)\cdot w = |z|^{\beta-1}\,\mathrm{Re}(z\overline{w})$, $(r,s)(r\rho + s\sigma) = z\,\mathrm{Re}(z\overline{w})$ and

$$\phi'(z)w = |z|^{\alpha-1}w + (\alpha-1)|z|^{\alpha-3}z\,\mathrm{Re}(z\overline{w}), \qquad z \neq 0.$$

The assertions now follow from Lemmas 9.1 and 9.2. $\square$

We can now treat the nonlinear wave equation (1.4) with Dirichlet boundary conditions on a bounded open set $\emptyset \neq U \subseteq \mathbb{R}^3$. In the same way as for the linear wave equation (6.6) we rewrite (1.4) as an evolution equation in $L^2(U)$ having second order in time, where we use again the Dirichlet Laplacian from Example 5.12. We thus obtain the problem

$$\begin{aligned} &w''(t) = \Delta_D w(t) - a w(t)|w(t)|^2, \qquad t \in J,\\ &w(0) = w_0, \qquad w'(0) = w_1, \end{aligned} \tag{9.3}$$

where $w_0 \in \mathrm{D}(\Delta_D), w_1 \in \mathring{H}^1(U)$ and $a \in \mathbb{R}$ are given. We look for solutions

$$w \in C^2(J, L^2(U)) \cap C^1(J, \mathring{H}^1(U)) \cap C(J, [\mathrm{D}(\Delta_D)])$$

of (9.3). To solve (9.3), we proceed as in Lecture 6 and reformulate (9.3) as the semilinear evolution equation

$$u'(t) = Au(t) + F(u(t)), \quad t \in J, \qquad u(0) = u_0, \tag{9.4}$$

on the Hilbert space $X = \mathring{H}^1(U) \times L^2(U)$ endowed with the norm given by $\|(u_1,u_2)\|^2 = \||\nabla u_1|\|_2^2 + \|u_2\|_2^2$. Here we set $u_0 = (w_0, w_1)$, and

$$\begin{aligned} A &= \begin{pmatrix} 0 & I \\ \Delta_D & 0 \end{pmatrix} \qquad \text{with } \mathrm{D}(A) = \mathrm{D}(\Delta_D) \times \mathring{H}^1(U),\\ F(u) &= (0, -a u_1 |u_1|^2) =: (0, F_0(u_1)) \quad \text{for } u = (u_1,u_2) \in X. \end{aligned} \tag{9.5}$$

Recall that A is skewadjoint by Example 5.13. Moreover, Corollary 9.3 implies that $F_0 : L^6(U) \to L^2(U)$ is real continuously differentiable and Lipschitz on bounded sets. Since $\mathring{H}^1(U) \hookrightarrow L^6(U)$ by Sobolev's embedding (5.7), we obtain that $F : X \to X$ has the same properties. In particular, F satisfies (8.1). We point out that this reasoning crucially depends on the fact that $U \subseteq \mathbb{R}^3$ and that the nonlinearity is cubic (i.e., $\alpha = 3$ in Corollary 9.3).

We can now apply Theorems 8.6 and 8.10 to (9.4) with A and F as above. Exactly as in Lemma 6.10, one shows that solutions of (9.3) and (9.4) are in unique correspondence, where $u = (w, w')$.

To define *mild* solutions for (9.3), we note that we can write the unitary group $T(\cdot)$ generated by A as

$$T(t) = \begin{pmatrix} T_{11}(t) & T_{12}(t) \\ T_{21}(t) & T_{22}(t) \end{pmatrix}$$

for operators $T_{11}(t) \in \mathcal{B}(\mathring{H}^1(U))$, $T_{12}(t) \in \mathcal{B}(L^2(U), \mathring{H}^1(U))$, $T_{21}(t) \in \mathcal{B}(\mathring{H}^1(U), L^2(U))$ and $T_{22}(t) \in \mathcal{B}(L^2(U))$, where $t \in \mathbb{R}$. Duhamel's formula (8.3) thus leads to the integral equation

$$w(t) = T_{11}(t)w_0 + T_{12}(t)w_1 + \int_0^t T_{12}(t-s)F_0(w(s))\,\mathrm{d}s, \quad t \in J, \tag{9.6}$$

on $\mathring{H}^1(U)$. We call a function $w \in C(J, \mathring{H}^1(U))$ satisfying (9.6) a *mild solution* of (9.3), where $w_0 \in \mathring{H}^1(U)$ and $w_1 \in L^2(U)$ are given.

After these preparations we can now establish a local wellposedness result for the nonlinear wave equation (9.3).

Proposition 9.4. *Let $U \subseteq \mathbb{R}^3$ be open and bounded, and let $a \in \mathbb{R}$. Then for each $u_0 = (w_0, w_1) \in X = \mathring{H}^1(U) \times L^2(U)$ there is a unique maximal mild solution $w \in C([0, t^+(u_0)), \mathring{H}^1(U))$ of (9.3). Moreover, let $J^+ = [0, t^+(u_0))$. Then the following assertions hold.*

(a) The maximal mild solution w of (9.3) belongs to $C^1(J^+, L^2(U))$ and $u = (w, w')$ is the mild solution of (9.4).

(b) If w_0 and w_1 are real-valued, then w is real-valued.

(c) If $(w_0, w_1) \in \mathrm{D}(A) = \mathrm{D}(\Delta_D) \times \mathring{H}^1(U)$, then the corresponding mild solution w belongs to $C^2(J^+, L^2(U)) \cap C^1(J^+, \mathring{H}^1(U)) \cap C(J^+, [\mathrm{D}(\Delta_D)])$ and it solves (9.3).

(d) If $t^+(u_0) < \infty$, then $\lim_{t \to t^+(u_0)} (\||\nabla w(t)|\|_2^2 + \|w'(t)\|_2^2) = \infty$.

(e) For any given $b \in (0, t^+(u_0))$ there is a radius $r > 0$ such that for $\tilde{u}_0 = (\tilde{w}_0, \tilde{w}_1) \in \overline{B}_X(u_0, r)$, we have $t^+(\tilde{u}_0) > b$ and the map $\overline{B}_X(u_0, r) \to C([0, b], X), \tilde{u}_0 \mapsto (\tilde{w}, \tilde{w}')$, is Lipschitz. Here, $\tilde{w}$ is the mild solution of (9.3) with the initial values $(\tilde{w}_0, \tilde{w}_1)$.

Proof. 1) We first note that due to the above observations, Theorem 8.10 yields a unique solution u of (9.4) on $J^+ = [0, t^+(u_0))$ for A and F given by (9.5) if $u_0 = (w_0, w_1) \in \mathrm{D}(A)$. We have $u = (w, w')$, where $w \in C^2(J^+, L^2(U)) \cap C^1(J^+, \mathring{H}^1(U)) \cap C(J^+, [\mathrm{D}(\Delta_D)])$ solves (9.3), so that (c) holds.

2) We next show that mild solutions of (9.3) are unique. Let $w, \tilde{w} \in C([0, b], \mathring{H}^1(U)) =: E$ be mild solutions of (9.3) for some $b > 0$ with the same initial values $(w_0, w_1) \in X$. Let r be larger than the norms of w and $\tilde{w}$ in E. Since $\|T(t)\| = 1$, it is easy to see that $\|T_{12}(t)\|_{\mathcal{B}(L^2, \mathring{H}^1)} \leq 1$ for all $t \in \mathbb{R}$. From (9.6), (8.1) and (5.7), we then deduce

$$\begin{aligned}\|w(t) - \tilde{w}(t)\|_{1,2} &= \Big\| \int_0^t T_{12}(t-s)(F_0(w(s)) - F_0(\tilde{w}(s)))\,\mathrm{d}s \Big\|_{1,2} \\ &\leq \int_0^t \|F_0(w(s)) - F_0(\tilde{w}(s))\|_2 \,\mathrm{d}s \leq cL(r) \int_0^t \|w(s) - \tilde{w}(s)\|_{1,2}\,\mathrm{d}s\end{aligned}$$

for all $t \in [0, b]$. Gronwall's inequality thus implies that $w = \tilde{w}$.

3) To obtain a mild solution for (9.3), let $u_0 = (w_0, w_1) \in X$. We then find $u_{0n} = (w_{0n}, w_{1n}) \in \mathrm{D}(A)$ converging to u_0 in X as $n \to \infty$. Theorem 8.6 shows that the corresponding solutions $u_n = (w_n, w_n')$ tend in $C([0, b], X)$ to the mild solution u of (9.4) for the initial value u_0, as $n \to \infty$, where $b \in (0, t^+(u_0))$ is arbitrary. By definition, the first component $w = [u]_1$ is a mild solution of (9.3). Since w_n and w_n' converge in $C([0, b], L^2(U))$, the function w belongs to $C^1([0, b], L^2(U))$ and $u = (w, w')$. This also proves (a).

4) The assertions (d) and (e) now directly follow from (a) and Theorem 8.6. To show (b), let w_0 and w_1 be real-valued. Due to part (e), we may assume that $(w_0, w_1) \in \mathrm{D}(A)$. We set $g = \operatorname{Re} w$ and $v = \operatorname{Im} w$. Then $v \in C^2(J^+, L^2(U)) \cap C^1(J^+, \mathring{H}^1(U)) \cap C(J^+, [\mathrm{D}(\Delta_D)])$ solves

$$v''(t) = \Delta_D v(t) - ag(t)^2 v(t) - av(t)^3, \quad t \in J^+, \qquad v(0) = 0, \quad v'(0) = 0,$$

where g is considered as a given function. Let $b \in J^+$. Since $\mathring{H}^1(U) \hookrightarrow L^6(U)$, the number $c(b) = \sup_{t\in[0,b]} \|w(t)\|_6^2$ is finite. Hölder's inequality yields $\|g(t)^2 v(t)\|_2 \le \|g(t)\|_6^2 \|v(t)\|_6 \le c(b)\|v(t)\|_6$ and $\|v(t)^3\|_2 \le c(b)\|v(t)\|_6$ for $t \in [0, b]$. One can now estimate as in step 2)

$$\|v(t)\|_{1,2} \le \int_0^t (\|ag(s)^2 v(s)\|_2 + \|av(s)^3\|_2)\, \mathrm{d}s \le 2|a|c(b)c \int_0^t \|v(s)\|_{1,2}\, \mathrm{d}s,$$

for $t \in [0, b]$, using once more Sobolev's embedding (5.7). Gronwall's inequality then shows that $v = 0$, and hence w is real-valued. □

Let $J^+ = [0, t^+(w_0, w_1))$ for $(w_0, w_1) \in X$. Since the mild solution w of (9.3) belongs to $C(J^+, \mathring{H}^1(U)) \cap C^1(J^+, L^2(U))$, we can define its "energy"

$$E_w(t) = E(w(t), w'(t)) = \int_U (\tfrac{1}{2}|w'(t)|^2 + \tfrac{1}{2}|\nabla w(t)|^2 + \tfrac{a}{4}|w(t)|^4)\, \mathrm{d}x \qquad (9.7)$$
$$= \tfrac{1}{2}\|(w(t), w'(t))\|_X^2 + \tfrac{a}{4}\|w(t)\|_4^4$$

for $t \in J^+$ since $\mathring{H}^1(U) \hookrightarrow L^4(U)$ by Sobolev's embedding (5.7). Moreover, the map $X \to \mathbb{R}, (w_0, w_1) \mapsto E(w_0, w_1)$, is continuous.

We next show that E is constant along mild solutions of (9.3) so that it is a natural quantity for the nonlinear wave equation. Further observe that $E_w(t)^{\frac{1}{2}}$ controls the $\mathring{H}^1(U) \times L^2(U)$ norm of $(w(t), w'(t))$, *provided* $a \ge 0$. This fact leads to global existence of all mild solutions in this case.

Proposition 9.5. *Let $U \subseteq \mathbb{R}^3$ be open and bounded, $a \in \mathbb{R}$, and $(w_0, w_1) \in X$. The following assertions hold for each mild solution w of (9.3).*

(a) $E_w(t) = E_w(0) = \frac{1}{2}\|(w_0, w_1)\|_X^2 + \frac{a}{4}\|w_0\|_4^4$ for $t \in [0, t^+(w_0, w_1))$.
(b) If $a \ge 0$, then $t^+(w_0, w_1) = \infty$ for each initial value $(w_0, w_1) \in X$.

Proof. (a) First let $(w_0, w_1) \in \mathrm{D}(A)$ and denote by w the solution of (9.3) on $J^+ = [0, t^+(u_0))$. Employing the regularity of w, Corollary 9.3 and the chain rule, we infer that $E_w \in C^1(J^+)$ and

$$E_w'(t) = \int_U \mathrm{Re}\big(w'(t)\overline{w''(t)} + \nabla w(t) \cdot \overline{\nabla w'(t)} + a|w(t)|^2 w(t)\overline{w'(t)}\big)\, \mathrm{d}x$$
$$= \mathrm{Re} \int_U \overline{w'(t)}(w''(t) - \Delta_D w(t) + aw(t)|w(t)|^2)\, \mathrm{d}x = 0$$

for $t \in J^+$. Here we used the definition of Δ_D, $\mathrm{Re}(z\overline{w}) = \mathrm{Re}(\overline{z}w)$, and equation (9.3). As a result, $E_w(t) = E_w(0)$ for all $t \in J^+$. Since the map $(w_0, w_1) \mapsto E_w(t)$ is continuous from X to $\mathbb{R}$ by Proposition 9.4 (e) and $\mathring{H}^1(U) \hookrightarrow L^4(U)$, we obtain (a) by approximation.

(b) If $a \ge 0$, then $\|(w(t), w'(t))\|_X^2 \le 2E_w(t) = 2E_w(0)$ for all $t \in J^+$ so that the blow-up criterion from Proposition 9.4 (d) implies $t^+(w_0, w_1) = \infty$. □

We want to repeat the main features of the above prototypical proof of global solvability. One first shows that the time derivative of the energy vanishes along (sufficiently smooth) solutions.[1] Here it is crucial that the energy fits to the equation. By approximation, we then deduce that the energy stays constant

[1] Of course, the argument also works if $E_w' \le 0$.

along all mild solutions using the continuous dependence on initial data and the continuity of the map $X \to \mathbb{R}, (w_0, w_1) \mapsto E(w_0, w_1)$, defined in (9.7). If the energy dominates the norm, the energy equality combined with the blow-up criterion in Proposition 9.4 (d) finally yield global existence.

Now, what is happening if $a < 0$? It is useful to neglect spatial derivatives for a moment and to consider the corresponding ordinary differential equation $\phi'' = |a|\phi^3$, with $\phi(t) \in \mathbb{R}$. This equation has the blow-up solutions $\phi_c(t) = c\Big(1 - \sqrt{\frac{|a|}{2}}ct\Big)^{-1}$ for each $c > 0$. Therefore, if we consider the cubic wave equation on U with Neumann boundary conditions $\partial_\nu w(t) = 0$ instead of Dirichlet conditions on ∂U, then we obtain the exploding solutions $w(t) = \phi_c(t)\mathbb{1}$, for $0 \le t < \sqrt{\frac{2}{|a|}}\frac{1}{c}$.

In the Dirichlet case one can also derive blow-up for certain initial values, but here one has to argue in a different way. In our proof we derive a differential inequality for the scalar function $\phi(t) = \|w(t)\|_2^2$ which forces ϕ to explode in finite time. Besides the wave equation (9.3), the crucial ingredient of the proof is the energy equality from Proposition 9.5 and the assumption that the energy at time $t = 0$ is negative.

Proposition 9.6. *Let $U \subseteq \mathbb{R}^3$ be open and bounded, $a < 0$ and $(w_0, w_1) \in \mathrm{D}(\Delta_D) \times \mathring{H}^1(U)$. We assume that*

$$\tfrac{|a|}{4}\|w_0\|_4^4 \ge \tfrac{1}{2}\||\nabla w_0|\|_2^2 + \tfrac{1}{2}\|w_1\|_2^2, \tag{9.8}$$

i.e., the initial energy $E(w_0, w_1)$ is negative (see (9.7)). We further require that

$$\int_U w_0 w_1 \,\mathrm{d}x > 0. \tag{9.9}$$

Then there is a blow-up time $t^+ > 0$ such that $\lim_{t\to t^+} \|w(t)\|_2 = \infty$.

We note that for each $a < 0$ there are $w_0, w_1 \in \mathrm{D}(\Delta_D)$ such that (9.8) and (9.9) hold. In fact, take $w_0 = \tau w_1$ for a given $0 \ne w_1 \in \mathrm{D}(\Delta_D)$ and choose a sufficiently large $\tau > 0$.

Proof. We have a solution

$$w \in C(J^+, [\mathrm{D}(\Delta_D)]) \cap C^1(J^+, \mathring{H}^1(U)) \cap C^2(J^+, L^2(U))$$

of (9.3) on the maximal existence interval $J^+ = [0, t^+(w_0, w_1))$. We set $\phi(t) = \frac{1}{4}\|w(t)\|_2^2$ for $t \in J^+$. Observe that $\phi'(0) > 0$ due to (9.9). Equation (9.7) and the definition of Δ_D then imply

$$\begin{aligned}\phi''(t) &= \frac{1}{2}\int_U (|w'(t)|^2 + w(t)w''(t))\,\mathrm{d}x \\ &= \frac{1}{2}\int_U |w'(t)|^2\,\mathrm{d}x - \frac{1}{2}\int_U |\nabla w(t)|^2\,\mathrm{d}x + \frac{|a|}{2}\int_U |w(t)|^4\,\mathrm{d}x\end{aligned}$$

for $t \in J^+$. Using the conservation of energy shown in Proposition 9.5 (a), we deduce

$$\phi''(t) = \int_U |w'(t)|^2\,\mathrm{d}x + \frac{|a|}{4}\int_U |w(t)|^4\,\mathrm{d}x - E(w_0, w_1)$$

Hölder's inequality further yields

$$\phi(t)^2 = \tfrac{1}{16}\|w(t)\|_2^4 \le \frac{\operatorname{vol}(U)}{16}\|w(t)\|_4^4.$$

We set $C = \frac{4|a|}{\operatorname{vol}(U)}$. Since $E(w_0, w_1) \le 0$ by (9.8), we arrive at

$$\phi''(t) \ge \frac{4|a|}{\operatorname{vol}(U)}\phi(t)^2 = C\phi(t)^2$$

for $t \in J^+$. Hence, $\phi'(t) \ge \phi'(0) + C\int_0^t \phi(s)^2\,\mathrm{d}s \ge \phi'(0) > 0$ and therefore ϕ strictly increases with $\phi(t) \ge \phi(0) + t\phi'(0)$ for $t \in J^+$. We can now estimate

$$\tfrac{\mathrm{d}}{\mathrm{dt}}\,\tfrac{1}{2}(\phi'(t))^2 = \phi''(t)\phi'(t) \ge C\phi(t)^2\phi'(t)$$

on J^+. Integrating once more, we infer

$$\phi'(t)^2 \ge \phi'(0)^2 + 2C\int_0^t \phi'(s)\phi(s)^2\,\mathrm{d}s = \phi'(0)^2 + \tfrac{2C}{3}\phi(t)^3 - \tfrac{2C}{3}\phi(0)^3$$

for $t \in J^+$. We suppose that $J^+ = \mathbb{R}_+$. Since $\phi'(0) > 0$, we can fix $t_0 \ge 0$ such that $\phi(0) + t_0\phi'(0) \ge \max\{0, (2\phi(0)^3 - \frac{3}{C}\phi'(0)^2)^{\frac{1}{3}}\}$. Let $t \ge t_0$. Because of $\phi(t)^3 \ge \phi(t_0)^3 \ge (\phi(0) + t_0\phi'(0))^3$, it follows

$$\phi'(t)^2 \ge \tfrac{C}{3}\phi(t)^3 + \tfrac{C}{3}(\phi(0) + t_0\phi'(0))^3 + \phi'(0)^2 - \tfrac{2C}{3}\phi(0)^3 \ge \tfrac{C}{3}\phi(t)^3.$$

As a result,

$$\phi'(t) \ge \sqrt{\tfrac{C}{3}}\,\phi(t)^{\frac{3}{2}}, \quad t \ge t_0, \quad \phi(t_0) > 0.$$

Observe that the solution of the equation

$$\psi'(t) = \sqrt{\tfrac{C}{3}}\,\psi(t)^{\frac{3}{2}}, \quad t \ge t_0, \quad \psi(t_0) = \phi(t_0),$$

has blow-up in finite time. The comparison principle for scalar differential equations now implies the assertion. □

An important feature of wave equations is the finite speed of propagation of their solutions. Roughly speaking, this means that if an initial value is compactly supported, then the support of $w(t)$ moves with finite velocity. This behavior is in accordance with the theory of relativity, in contrast to the diffusion equation $u' = \Delta_D u$ whose solutions u are strictly positive for each $t > 0$ if $u(0) \ge 0$ is nonzero. To describe this behavior, we consider the cone

$$C(x_0, t_0) = \Big\{(x,t) \in \mathbb{R}^3 \times \mathbb{R}_+ \,\Big|\, 0 \le t \le t_0, |x - x_0| \le t_0 - t\Big\}$$

with base $\overline{B}(x_0, t_0) \subseteq \mathbb{R}^3$ and vertex $(x_0, t_0) \in \mathbb{R}^4$, where $t_0 > 0$ and $x_0 \in \mathbb{R}^3$. Recall the comment after Example 5.12 stating that $[\mathrm{D}(\Delta_D)]$ is isomorphic to $H^2(U) \cap \mathring{H}^1(U)$ if $\partial U \in C^2$ so that $\Delta_D u = \Delta u$ for all $u \in \mathrm{D}(\Delta_D)$.

PROPOSITION 9.7. *Let $U \subseteq \mathbb{R}^3$ be open and bounded with $\partial U \in C^2$ and let $a \ge 0$. Take $x_0 \in U$ and $t_0 > 0$ such that $\overline{B}(x_0, t_0) \subseteq U$. Let $(w_0, w_1) \in \mathrm{D}(\Delta_D) \times \mathring{H}^1(U)$ and assume that $w_0 = w_1 = 0$ on $B(x_0, t_0)$. Then the corresponding solution w of (9.3) vanishes on $C(x_0, t_0)$.*

Proof. The result is proved by a local energy estimate. Let $t_0 > 0$ and $x_0 \in U$ with $\overline{B}(x_0, t_0) \subseteq U$. Take $(w_0, w_1) \in \mathrm{D}(\Delta_D) \times \mathring{H}^1(U)$ and let w solve (9.3). Since $[\mathrm{D}(\Delta_D)] \subseteq H^2(U)$, we have $w \in C^2(\mathbb{R}_+, L^2(U)) \cap C^1(\mathbb{R}_+, H^1(U)) \cap C(\mathbb{R}_+, H^2(U))$. For $0 \le t < t_0$, we set $B_t = B(x_0, t_0 - t)$ and define

$$e(t) = \int_{B_t} (\tfrac{1}{2}|\partial_t w(t)|^2 + \tfrac{1}{2}|\nabla w(t)|^2 + \tfrac{a}{4}|w(t)|^4)\,\mathrm{d}x.$$

We will show that $e' \le 0$ so that $0 \le e(t) \le e(0) = 0$ for all $t \in [0, t_0)$, and hence $w = 0$ on $C(x_0, t_0)$. (Actually, in case $a = 0$ one deduces $w = 0$ on $C(x_0, t_0)$ from $\partial_t w = 0$ on $C(x_0, t_0)$ and $w(0) = 0$ on $B(x_0, t_0)$.) Using Reynolds' transport theorem (see Exercise 9.1), we compute

$$\begin{aligned} e'(t) = \mathrm{Re} \int_{B_t} (\partial_{tt} w(t) \partial_t \overline{w}(t) + \nabla w(t) \cdot \nabla \partial_t \overline{w}(t) + a|w(t)|^2 w(t) \partial_t \overline{w}(t))\,\mathrm{d}x \\ - \int_{\partial B_t} (\tfrac{1}{2}|\partial_t w(t)|^2 + \tfrac{1}{2}|\nabla w(t)|^2 + \tfrac{a}{4}|w(t)|^4)\,\mathrm{d}\sigma, \end{aligned}$$

cf. the proof of Proposition 9.5. We note that the traces on ∂B_t exists since $\partial_t w(t), \partial_k w(t) \in H^1(U)$ and $w(t) \in H^2(U) \hookrightarrow C(\overline{U})$ (see Corollary D.21 and Theorem D.23). Because of $w(t) \in H^2(U)$, we can use Gauß' formula from Theorem D.28 on the ball B_t leading to

$$\int_{B_t} \nabla w(t) \cdot \nabla \partial_t \overline{w}(t) = -\int_{B_t} \Delta w(t)\, \partial_t \overline{w}(t)\,\mathrm{d}x + \int_{\partial B_t} \partial_\nu w(t)\, \partial_t \overline{w}(t)\,\mathrm{d}\sigma,$$

where $\partial_\nu \varphi = \nu \cdot \mathrm{tr}\, \nabla \varphi$ for $\varphi \in H^2(B_t)$ and the outer unit normal ν on ∂B_t. Employing equation (9.3) and Hölder's inequality, we thus obtain

$$\begin{aligned} e'(t) &= \mathrm{Re} \int_{\partial B_t} \partial_\nu w(t) \partial_t \overline{w}(t)\,\mathrm{d}\sigma - \int_{\partial B_t} (\tfrac{1}{2}|\partial_t w(t)|^2 + \tfrac{1}{2}|\nabla w(t)|^2 + \tfrac{a}{4}|w(t)|^4)\,\mathrm{d}\sigma \\ &\le -\int_{\partial B_t} \tfrac{a}{4}|w(t)|^4\,\mathrm{d}\sigma \le 0. \end{aligned}$$

□

It would be nice to see a wave type solution of our wave equation. We thus construct "standing waves" w for (9.3), i.e., solutions w of (9.3) given by

$$w(t,x) = \mathrm{e}^{\mathrm{i}\omega t} \varphi(x), \quad t \ge 0,\ x \in U, \tag{9.10}$$

where $\omega \in \mathbb{R}$ and $\varphi \in \mathrm{D}(\Delta_D)$. Inserting this ansatz into (9.3), we infer that w in (9.10) solves (9.3) (with $w_0 = \varphi$ and $w_1 = \mathrm{i}\omega\varphi$) if $\omega^2 \ge 0$ and $\varphi \in \mathrm{D}(\Delta_D)$ solve the "nonlinear eigenvalue problem"

$$\Delta_D \varphi + \omega^2 \varphi - a\varphi|\varphi|^2 = 0. \tag{9.11}$$

This observation leads us to the theory of semilinear elliptic equations which is a bit off the main track of the Internet Seminar. Here we focus on a rather simple result.

We want to solve (9.11) via bifurcation theory. It allows to construct solutions to (9.11) as perturbations of the maximal eigenvalue of the linearization. We only need one of most basic results in this direction, the Crandall-Rabinowitz theorem, see e.g. Theorem I.5.1 in [**Kie04**]. To this aim, we set

$$\Phi(\lambda, \varphi) = \Delta_D \varphi + \lambda\varphi - a\varphi|\varphi|^2$$

for $\lambda \in \mathbb{R}$ and $\varphi \in \mathrm{D}(\Delta_D)$. Clearly, $\Phi(\lambda, 0) = 0$ for all $\lambda \in \mathbb{R}$. We are looking for $0 \neq \varphi_\omega \in \mathrm{D}(\Delta_D)$ and $\omega \in \mathbb{R}$ with $\Phi(\omega^2, \varphi_\omega) = 0$, i.e., (ω, φ_ω) solve (9.11).

To check the assumptions of the Crandall-Rabinowitz theorem, let U be connected and let all spaces be real. We claim that, Δ_D has a largest eigenvalue $-\lambda_0 < 0$ with corresponding eigenfunction φ_0, where the kernel of $\Delta_D + \lambda_0 I$ is one dimensional and its range has codimension 1.

To establish these facts, we recall a few facts about the spectrum and the eigenfunctions of Δ_D. We first note that $\mathring{H}^1(U)$ is compactly embedded into $L^2(U)$. In fact, one can extend each $f \in \mathring{H}^1(U)$ by 0 to a function $\widetilde{f} \in H^1(V)$ for an open ball V containing $\overline{U}$. Due to Theorem D.24 in Appendix D, the space $H^1(V)$ is compactly embedded in $L^2(V)$ which implies the claim. (See also Remark 20 in Section 9.4 in [**Bre11**].) Observe that $[\mathrm{D}(\Delta_D)]$ is continuously embedded into $\mathring{H}^1(U)$ since for $u \in \mathrm{D}(\Delta_D)$ we have $u \in \mathring{H}^1(U)$ and

$$\int_U |\nabla u|^2 \,\mathrm{d}x = \int_U \Delta_D u \,\overline{u} \,\mathrm{d}x \leq \|\Delta_D u\|_2 \|u\|_2 \leq \tfrac{1}{2}\|\Delta_D u\|_2^2 + \tfrac{1}{2}\|u\|_2^2, \tag{9.12}$$

by the definition of Δ_D and Hölder's inequality. In particular, $[\mathrm{D}(\Delta_D)]$ is compactly embedded into $L^2(U)$ and thus Δ_D has a compact resolvent. Therefore $\sigma(\Delta_D)$ only consists of isolated eigenvalues of finite multiplicity, see e.g. Theorem III.6.9 in [**Kat95**]. Example 5.12 tells us that $\sigma(\Delta_D) \subseteq (-\infty, 0)$. Hence there is a maximal eigenvalue $-\lambda_0 < 0$ of Δ_D. It is further known that the kernel of $\lambda_0 I + \Delta_D$ is one dimensional (and so the range of $\lambda_0 I + \Delta_D$ has codimension 1 by compactness). See e.g. Theorem 6.5.2 in [**Eva10**].

Due to Corollary 9.3 and $[\mathrm{D}(\Delta_D)] \hookrightarrow L^6(U)$, we obtain $\Phi \in C^1(\mathbb{R} \times [\mathrm{D}(\Delta_D)], L^2(U))$ and $L := \partial_2\Phi(\lambda_0, 0) = \Delta_D + \lambda_0 I$, where I is the embedding of $[\mathrm{D}(\Delta_D)]$ into $L^2(U)$. One can further check that $\Phi \in C^2(\mathbb{R} \times [\mathrm{D}(\Delta_D)], L^2(U))$ and $\Lambda := \partial_{12}\Phi(\lambda_0, 0) = I$. Finally, $\varphi_0 = \Lambda\varphi_0$ is orthogonal to the range of L since

$$(\Delta_D \varphi + \lambda_0 \varphi | \varphi_0) = (\varphi | \Delta_D \varphi_0 + \lambda_0 \varphi_0) = 0$$

for all $\varphi \in \mathrm{D}(\Delta_D)$. We have now verified the assumptions Theorem I.5.1 in [**Kie04**] which gives us $0 \neq \varphi_\omega \in \mathrm{D}(\Delta_D)$ and $\omega \in \mathbb{R}$ such that $\Phi(\omega^2, \varphi_\omega) = 0$. More precisely, here ω^2 is close to λ_0 and φ_ω is small in the norm of $[D(\Delta_D)]$.

PROPOSITION 9.8. *Let $a \in \mathbb{R}$ and $U \subseteq \mathbb{R}^3$ be open, bounded and connected. Then for ω^2 close to λ_0 there are real-valued functions $0 \neq \varphi_\omega \in \mathrm{D}(\Delta_D)$ solving* (9.11). *Hence, the standing waves given by $w(t) = \mathrm{e}^{\mathrm{i}\omega t}\varphi_\omega$, $t \geq 0$, solve* (9.3).

Exercises

Exercise 9.1. Let $U \subseteq \mathbb{R}^d$ be open. Let $x_0 \in \mathbb{R}^d$ and $t_0 > 0$ be such that $\overline{B}(x_0, t_0) \subseteq U$. Set $J = [0, t_0]$ and $B_t = B(x_0, t_0 - t)$. Prove for $f \in C^1(J, L^1(U)) \cap C(J, W_1^1(U))$ and each $t \in J$ the identity

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_{B_t} f(x,t)\,\mathrm{d}x = \int_{B_t} (\partial_t f)(x,t)\,\mathrm{d}x - \int_{\partial B_t} f(x,t)\,\mathrm{d}\sigma.$$

This is a special case of *Reynolds' transport theorem*.

Exercise 9.2. Let A generate the C_0–semigroup $T(\cdot)$ on X and that $F : X \to X$ is Lipschitz on bounded subsets of X. Consider the semilinear problem

$$u'(t) = Au(t) + F(u(t)), \quad t \geq 0, \qquad u(0) = u_0,$$

and suppose that $t^+(u_0) = \infty$ for all $u_0 \in X$. Define $\Psi : \mathbb{R}_+ \times X \to X$ by $\Psi(t, u_0) = u(t; u_0)$, where $u(\cdot\,; u_0)$ is the unique mild solution with initial value u_0. Prove that Ψ is a *nonlinear semiflow*, i.e., we have $\Psi \in C(\mathbb{R}_+ \times X, X)$ and for all $u_0 \in X$ and $t_1, t_2 \geq 0$ it holds that

$$\Psi(0, u_0) = u_0, \qquad \Psi(t_2 + t_1, u_0) = \Psi(t_2, \Psi(t_1, u_0)).$$

Exercise 9.3. Let A generate the C_0–semigroup $T(\cdot)$ on X. Let $F, F_n : X \to X$ for $n \in \mathbb{N}$ be Lipschitz on bounded sets and suppose that $F_n \to F$ locally uniformly as $n \to \infty$. Fix $u_0 \in X$. Let u be the unique mild solution of

$$u'(t) = Au(t) + F(u(t)), \quad t \geq 0, \qquad u(0) = u_0,$$

with maximal existence time $t^+(u_0, F)$, and let u_n be the unique mild solution of

$$u_n'(t) = Au_n(t) + F_n(u_n(t)), \quad t \geq 0, \qquad u_n(0) = u_0,$$

with maximal existence time $t^+(u_0, F_n)$. Take any $0 < b < t^+(u_0, F)$. Show that $\liminf_{n\to\infty} t^+(u_0, F_n) > b$ and that $u_n \to u$ uniformly on $[0, b]$ as $n \to \infty$.

LECTURE 10

The nonlinear Schrödinger equation

We now start with the last major part of the Internet Seminar which is devoted to the basic version of the nonlinear Schrödinger equation. In this lecture we first discuss a bit its derivation and its role in mathematical physics. We then study several (more or less) explicit solutions of the equation and derive two important conservation laws for sufficiently regular solutions. In view of these results, we will then reconsider the free Schrödinger group $T(\cdot)$ generated by $\mathrm{i}\Delta$ on $L^2(\mathbb{R}^d)$ and extend these operators to a larger space. Finally, we establish a representation formula for $T(t)$ which implies that $T(t)$ maps $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ into $L^\infty(\mathbb{R}^d)$. This property expresses the *dispersive* behavior of the linear Schrödinger equation and it is a first step towards Strichartz' estimates shown in the next lecture.

We investigate the nonlinear Schrödinger equation

$$\begin{aligned} \mathrm{i}\partial_t u(t,x) &= -\Delta u(t,x) + \mu |u(t,x)|^{\alpha-1} u(t,x), \quad x \in \mathbb{R}^d, \quad t \in J, \\ u(0,x) &= u_0(x), \qquad\qquad\qquad\qquad\qquad\qquad x \in \mathbb{R}^d, \end{aligned} \tag{10.1}$$

where J is a nontrivial interval with $0 \in J$. Equivalently, one can write

$$\begin{aligned} \partial_t u(t,x) &= \mathrm{i}\Delta u(t,x) - \mathrm{i}\mu |u(t,x)|^{\alpha-1} u(t,x), \quad x \in \mathbb{R}^d, \quad t \in J, \\ u(0,x) &= u_0(x), \qquad\qquad\qquad\qquad\qquad\qquad x \in \mathbb{R}^d. \end{aligned} \tag{10.2}$$

We put $a_+ = \max\{a, 0\}$ for $a \in \mathbb{R}$. Throughout it is assumed that

$$\mu \in \{-1, 1\} \qquad \text{and} \qquad 1 < \alpha < \frac{d+2}{(d-2)_+} =: \alpha_c. \tag{10.3}$$

Right away we point out that (10.1) or (10.2) cannot be treated with the methods of the two previous lectures. The operator $\mathrm{i}\Delta$ is a generator on $L^2(\mathbb{R}^d)$, but the nonlinearity $F(u) = -\mathrm{i}\mu|u|^{\alpha-1}u$ does not map $L^2(\mathbb{R}^d)$ into itself.

The case $\mu = 1$ is called *defocusing*, and one calls the case $\mu = -1$ *focusing*. We later explain this terminology a bit and describe the differences between the two cases. Actually, one can consider more general nonlinearities (under appropriate growth assumptions) and one could replace $-\Delta$ by a more general Schrödinger operator, e.g., of the form $-\Delta + V$ for a potential V. We refer to the monograph [**Caz03**] for a systematic study of nonlinear Schrödinger equations. To avoid technical difficulties, we restrict ourselves to the model nonlinearity in (10.1) which already gives a very good insight into the field.

The borderline "critical" case $\alpha = \alpha_c$ for $d \geq 3$ (see (10.3)) is quite demanding. One could tackle the corresponding local wellposedness results at or after the end of our course, but for this case the main theorems on global existence are far beyond our scope. An extended and detailed survey is given in [**Tao06**].

Mathematically, the nonlinear Schrödinger equation serves as a model problem for the large class of so-called *dispersive* partial differential equations, see [**Tao06**]. It further arises in various areas of physics. We briefly describe two of the main fields of applications.

The first one is nonlinear optics in a so-called *Kerr medium.* Here one considers electromagnetic waves in a material, say a glass fiber. The time evolution of the waves is governed by Maxwell's equations. The effects of the material enter into the polarization P which depends on the electric field E. One expresses the relation $P = P(E)$ by Taylor's formula up to third order, say. If the material has a centrosymmetric and isotropic structure, the second order term in this expansion vanishes. Also the third order term simplifies and involves only the expression $(E \cdot E)E$. One calls the resulting material law a *Kerr nonlinearity.* Still, the nonlinear Maxwell equation is a quite demanding problem and thus one looks for further simplifications. To this end, one writes the electric field as a Taylor series whose leading term is a *wave packet* of the form

$$E_0(t,x) = A(t,x)\mathrm{e}^{\mathrm{i}(\xi_0\cdot x-\omega_0 t)} + \overline{A(t,x)}\mathrm{e}^{-\mathrm{i}(\xi_0\cdot x-\omega_0 t)}$$

with the wave vector $\xi_0 \in \mathbb{R}^3$, the frequency $\omega_0 \in \mathbb{R}$ and the amplitude function A. One then inserts this ansatz into the nonlinear Maxwell equation. After long calculations, further approximations and transformations, one ends up with an equation for the transformed amplitude $\widetilde{A}$, which only depends on the (transformed) space variable from $\mathbb{R}^3$. This equation is a nonlinear Schrödinger equation with $d = 2$ and $\alpha = 3$, where the t-variable corresponds to the coordinate of the direction of propagation of the wave (along the glass fiber). Moreover, in the equation appear additional coefficients determined by the material law. We refer to Chapters 2 and 3 of [**MN04**] for a detailed treatment of this sophisticated derivation. It can also be seen that a negative coefficient in front of $|\widetilde{A}|^2\widetilde{A}$ corresponds to an increased refraction index so that light is focused here, in accordance with the above terminology.

We note that the above reasoning can also be used for other problems, and quite often it turns out that the nonlinear Schrödinger equation (approximately) describes envelopes of wave packets (see e.g. [**SW11**]).

A second source for nonlinear Schrödinger equations are *Bose-Einstein* condensates. Such a condensate is the state of matter of a gas of N bosons which are trapped by an external potential V and kept at an absolute temperature very close to 0. All bosons occupy the same quantum state and can thus be described by a single wave function $u : \mathbb{R}^3 \to \mathbb{C}$ giving the particle density $|u|^2$ (so that $\|u\|_2^2 = N$). The interactions between the bosons lead to nonlinear contributions to the Schrödinger equation for this quantum system. Considering only binary collisions between the bosons, one sees that u satisfies a version of (10.1) with $\alpha = 3$, where one adds in (10.1) the term Vu on the right-hand side and multiplies the other terms by suitable coefficients. The resulting equation is often called *Gross-Pitaevski equation.*

We observe that in the above applications the exponent $\alpha = 3$ occurs naturally. This case is admitted in (10.3) for the space dimensions $d \in \{1,2,3\}$, whereas it is critical for $d = 4$. In view of the applications also $\alpha = 5$ would be interesting. This number is allowed in (10.3) only for $d \in \{1, 2\}$ and it is critical for $d = 3$. We also consider non-integer α mainly for mathematical reasons.

We first derive several more or less explicit solutions of (10.1) to get an insight to this equation, where (10.1) is understood in a pointwise sense.

Example 10.1 (Plane waves).
Let $\xi \in \mathbb{R}^d \setminus \{0\}$ and $a \in \mathbb{C}$. We claim that the smooth function w_ξ given by

$$w_\xi(t, x) = a e^{i\xi \cdot x} e^{-i|\xi|^2 t} e^{-i\mu|a|^{\alpha-1} t}, \quad t \in \mathbb{R},\ x \in \mathbb{R}^d, \tag{10.4}$$

satisfies (10.2) with $w_\xi(0, x) = a e^{i\xi \cdot x}$. In fact, consider functions of the form $v(t, x) = \phi(t) e^{i\xi \cdot x}$ for some $\phi \in C^1(\mathbb{R})$ and for $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$. We then have $\partial_t v(t, x) = \phi'(t) e^{i\xi \cdot x}$ and $\Delta v(t, x) = -|\xi|^2 v(t, x)$ for $(t, x) \in \mathbb{R}^{d+1}$. Since $|v| = |\phi|$, the function v satisfies (10.2) if

$$\phi'(t) = -i(|\xi|^2 + \mu|\phi(t)|^{\alpha-1})\phi(t), \quad t \in \mathbb{R}.$$

For an initial value $\phi(0) = a \in \mathbb{C}$, this scalar differential equation has the solution $\phi(t) = a \exp(-i(|\xi|^2 + \mu|a|^{\alpha-1})t)$. ♢

One calls functions as w_ξ in (10.4) *plane waves.* Here a is the amplitude, ξ is the wave vector and $\omega = |\xi|^2 + \mu|a|^{\alpha-1}$ is the frequency. Observe that the summand $|\xi|^2$ in ω comes from $-\Delta$ whereas $\mu|a|^{\alpha-1}$ is the contribution of the nonlinear part which depends on the absolute value of w_ξ. For $\mu = 1$ these two terms add up and so the nonlinearity increases the frequency and thus the time oscillation, whereas for $\mu = -1$ the oscillations partly cancel.

The wave w_ξ equals the amplitude a on the plane given by

$$\frac{1}{|\xi|} \xi \cdot x = \Big(|\xi| + \mu \frac{|a|^{\alpha-1}}{|\xi|}\Big) t =: v(\xi) t.$$

This plane moves along its unit normal vector $\frac{1}{|\xi|}\xi$ with the *phase velocity* $v(\xi)$ which depends on the length of the wave vector. This behavior is called *dispersion.* Dispersion causes plane waves with different wave vectors ξ_j (say, having the same direction $\frac{1}{|\xi_j|}\xi_j = \eta$) to spread out in space as time evolves. This effect will be stronger in the defocusing case $\mu = 1$, since again the nonlinear effect adds to the linear one. In the case $\mu = -1$ the waves exhibit less dispersion, they longer stay *focused.*

We next consider standing waves which are solutions of (10.1) where a time independent profile φ_ω oscillates in time due to a factor $e^{i\omega t}$.

Example 10.2 (Standing waves in the focusing case).
For $\mu = -1$ and each $\omega > 0$ and $1 < \alpha < \alpha_c$ there is a function $\varphi_\omega \in H^2(\mathbb{R}^d)$ such that the function $u_\omega \in C^\infty(\mathbb{R}, H^2(\mathbb{R}^d))$ given by $u_\omega(t) = e^{i\omega t}\varphi_\omega$, $t \in \mathbb{R}$, satisfies (10.1). In fact, we first note that u_ω satisfies (10.1) if $\varphi_\omega \in H^2(\mathbb{R}^d)$ solves the semilinear elliptic equation

$$-\Delta\varphi_\omega + \omega\varphi_\omega = |\varphi_\omega|^{\alpha-1}\varphi_\omega \tag{10.5}$$

on $\mathbb{R}^d$. For $d = 1$ and $\omega = 1$ one finds an explicit solution, namely

$$\varphi_1(x) = \Big(\frac{\sqrt{q+1}}{\cosh(qx)}\Big)^{\frac{1}{q}}, \quad x \in \mathbb{R},$$

where we set $q = \frac{\alpha-1}{2}$. For $\alpha = 3$, one has $q = 1$ and $\varphi_1 = \frac{\sqrt{2}}{\cosh}$. In general, one can solve (10.5) by means of the calculus of variations.[1] This is done in Section 8.1 of [**Caz03**]. The combination of several theorems in this section even gives a strictly positive solution $\varphi_\omega \in H^2(\mathbb{R}^d) \cap C^2(\mathbb{R}^d)$ of (10.5). Moreover, φ_ω is spherically symmetric and φ_ω, $|\nabla\varphi_\omega|$ decay exponentially as $|x| \to \infty$. $\diamondsuit$

The cases $\mu = 1$ and $\mu = -1$ in (10.1) with $1 < \alpha < \alpha_c$ crucially differ with respect to global existence. For the defocusing case we will show global existence in Lecture 13. In the focusing case one has blow-up if $1 + \frac{4}{d} \le \alpha < \alpha_c$, see Corollary 6.5.14 in [**Caz03**]. For the case $\alpha = 1 + \frac{4}{d}$ one can find an exploding solution based on Example 10.2. We stress that the blow-up occurs in $H^1(\mathbb{R}^d)$, whereas the norm in $L^2(\mathbb{R}^d)$ stays constant.

Example 10.3 (Blow-up solution).
Let $\mu = -1$, $\alpha = 1 + \frac{4}{d}$ and $\omega > 0$. Take the standing wave solution $u_\omega(t) = \mathrm{e}^{\mathrm{i}\omega t}\varphi_\omega$, $t \in \mathbb{R}$, from Example 10.2, where $0 < \varphi_\omega \in H^2(\mathbb{R}^d) \cap C^2(\mathbb{R}^d)$ solves (10.5). We define

$$u(t,x) = (\mathrm{i}(t-1))^{-\frac{d}{2}} \mathrm{e}^{\mathrm{i}\frac{|x|^2}{4(t-1)}} \mathrm{e}^{-\mathrm{i}\frac{\omega}{(t-1)}} \varphi_\omega\big(\tfrac{1}{t-1}x\big), \quad t \in \mathbb{R} \setminus \{1\},\ x \in \mathbb{R}^d.$$

(If you wonder how one can guess such a formula, you may look at p. 115/116 of [**Tao06**].) A direct, but somewhat tedious computation shows that u satisfies (10.1) for $t \ne 1$ with $\mu = -1$ and $\alpha = 1 + \frac{4}{d}$. Moreover, for $t \ne 1$ we have

$$\begin{aligned}
\|u(t)\|_2^2 &= \int_{\mathbb{R}^d} |t-1|^{-d} |\varphi_\omega\big(\tfrac{1}{t-1}x\big)|^2 \,\mathrm{d}x = \|\varphi_\omega\|_2^2, \\
\||\nabla u(t)|\|_2^2 &= \int_{\mathbb{R}^d} |t-1|^{-d} \Big|\tfrac{\mathrm{i}}{2(t-1)}\varphi_\omega\big(\tfrac{1}{t-1}x\big)x + \tfrac{1}{t-1}\nabla\varphi_\omega\big(\tfrac{1}{t-1}x\big)\Big|^2 \,\mathrm{d}x \\
&\ge \int_{\mathbb{R}^d} |t-1|^{-d-2} \big|\nabla\varphi_\omega\big(\tfrac{1}{t-1}x\big)\big|^2 \,\mathrm{d}x = |t-1|^{-2} \||\nabla\varphi_\omega|\|_2^2.
\end{aligned}$$

As a result, this solution explodes as $t \to 1^-$ in $H^1(\mathbb{R}^d)$ though it stays bounded in $L^2(\mathbb{R}^d)$. Observe that the initial value $u(0)$ belongs to $H^2(\mathbb{R}^d)$ since $\varphi_\omega \in H^2(\mathbb{R}^d)$ and φ_ω, $|\nabla\varphi_\omega|$ decay exponentially as $|x| \to \infty$, see Example 10.2. $\diamondsuit$

To understand the different behavior of the L^2– and H^1–norm in the above example, we next discuss preliminary versions of the two fundamental conservation laws for the L^2-norm and for the energy. To that purpose, we recall that Sobolev's embedding (5.5) yields $H^1(\mathbb{R}^d) \hookrightarrow L^{1+\alpha}(\mathbb{R}^d)$ for $1 < \alpha < \alpha_c$ since $1 - \frac{d}{2} > -\frac{d}{1+\alpha}$.[2] As a result, on $H^1(\mathbb{R}^d)$ the "energy"

$$E(v) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla v|^2 \,\mathrm{d}x + \frac{\mu}{\alpha+1}\int_{\mathbb{R}^d} |v|^{\alpha+1} \,\mathrm{d}x = \tfrac{1}{2}\||\nabla v|\|_2^2 + \tfrac{\mu}{\alpha+1}\|v\|_{1+\alpha}^{1+\alpha} \quad (10.6)$$

is defined, and $E : H^1(\mathbb{R}^d) \to \mathbb{R}$ is real continuously differentiable by Corollary 9.3. If $\mu = 1$, we have $\|v\|_{1,2}^2 \le 2E(v) + \|v\|_2^2$ so that the H^1–norm is controlled by the energy and the L^2–norm in the defocusing case.

[1] We note that the corresponding result proved in Proposition 9.8 is much easier because of the existence of an isolated simple eigenvalue for the (Dirichlet) Laplacian.

[2] Here one can also allow for $\alpha = \alpha_c$ if $d \ne 2$. Since the critical case $\alpha = \alpha_c$ is not treated below, mostly we will not mention such facts further on.

Remark 10.4 (Conservation of the L^2–norm). Let $u \in C^1(J, L^2(\mathbb{R}^d)) \cap C(J, H^2(\mathbb{R}^d))$ satisfy (10.1). Integration by parts then yields

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{dt}} \|u(t)\|_2^2 &= \tfrac{\mathrm{d}}{\mathrm{dt}} \int_{\mathbb{R}^d} u(t)\overline{u(t)}\,\mathrm{d}x = 2\operatorname{Re} \int_{\mathbb{R}^d} u'(t)\overline{u(t)}\,\mathrm{d}x \\ &= 2\operatorname{Re}\mathrm{i} \int_{\mathbb{R}^d} (\Delta u(t) - \mu|u(t)|^{\alpha-1}u(t))\overline{u(t)}\,\mathrm{d}x \\ &= 2\operatorname{Im} \int_{\mathbb{R}^d} (|\nabla u(t)|^2 + \mu|u(t)|^{\alpha+1})\,\mathrm{d}x = 0\end{aligned}$$

for $t \in J$. As a result, $\|u(t)\|_2 = \|u_0\|_2$ for all $t \in J$.

Recall that in the linear case $\mu = 0$ one has the same conservation property, which extends to (10.1) due to the structure of the nonlinearity. ♢

Remark 10.5 (Conservation of energy). Let $u \in C^1(J, H^1(\mathbb{R}^d)) \cap C(J, H^2(\mathbb{R}^d))$ satisfy (10.1). The chain rule, Corollary 9.3 and an integration by parts then imply

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{dt}} E(u(t)) &= \int_{\mathbb{R}^d} \operatorname{Re}(\nabla u(t) \cdot \nabla\overline{u'(t)})\,\mathrm{d}x + \mu \int_{\mathbb{R}^d} \operatorname{Re}(|u(t)|^{\alpha-1}u(t)\overline{u'(t)})\,\mathrm{d}x \\ &= \operatorname{Re} \int_{\mathbb{R}^d} (-\Delta u(t) + \mu|u(t)|^{\alpha-1}u(t))\overline{u'(t)}\,\mathrm{d}x \\ &= \operatorname{Re} \int_{\mathbb{R}^d} \mathrm{i}u'(t)\overline{u'(t)}\,\mathrm{d}x = 0\end{aligned}$$

for $t \in J$. Therefore $E(u(t)) = E(u_0)$ for all $t \in J$. ♢

The above results indicate that there is some hope to control the H^1–norm of solutions at least in the defocusing case. Aiming at global existence, we are thus looking for a framework for the nonlinear Schrödinger equation (10.2) in which we can derive a blow-up condition involving only the H^1–norm. Looking at the proof of Theorem 8.6, one sees that the norm for the blow-up condition is closely tied to the required regularity for the initial value. So one would like to take $u_0 \in H^1(\mathbb{R}^d)$ and thus looks for solutions u in $C(J, H^1(\mathbb{R}^d))$.

But how to extend the Laplacian to $H^1(\mathbb{R}^d)$? Actually, there are several ways to do this, where we use an approach motivated by the Lax-Milgram theorem. To this aim, we define the Sobolev spaces of negative order by duality, i.e.,

$$W_p^{-k}(\mathbb{R}^d) = W_{p'}^k(\mathbb{R}^d)^* \quad \text{for } k \in \mathbb{N},\ 1 < p < \infty,$$

endowed with the norm

$$\|\varphi\|_{-k,p} = \sup_{\|v\|_{k,p'} \le 1} |\varphi(v)| = \sup_{\|v\|_{k,p'} \le 1} \langle v, \varphi \rangle_{W_{p'}^k},$$

where $p' = \frac{p}{p-1}$ and $\langle \cdot\,, \cdot \rangle_{W_{p'}^k}$ denotes the duality pairing between $W_{p'}^k(\mathbb{R}^d)$ and $W_p^{-k}(\mathbb{R}^d)$. As before, we set

$$H^{-k}(\mathbb{R}^d) = H^k(\mathbb{R}^d)^*$$

for $p = 2$. Observe that the space $W_p^{-k}(\mathbb{R}^d)$ is reflexive and that its dual can be identified with $W_{p'}^k(\mathbb{R}^d)$.

We focus on $H^{-1}(\mathbb{R}^d)$ and first discuss its relation to the L^p spaces. The inclusion map $J : H^1(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is injective and has dense range. We identify $L^2(\mathbb{R}^d)$ with $L^2(\mathbb{R}^d)^*$. Standard results from functional analysis then imply that also $J^* : L^2(\mathbb{R}^d) \to H^{-1}(\mathbb{R}^d)$ is injective and has dense range. (See the proof of Theorem C.5 and note that one can identify J^{**} with J.) We thus obtain the dense embeddings

$$H^1(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d) \cong L^2(\mathbb{R}^d)^* \hookrightarrow H^{-1}(\mathbb{R}^d). \tag{10.7}$$

Here a function $f \in L^2(\mathbb{R}^d)$ induces the continuous linear functional $\varphi_f = J^* f$ on $H^1(\mathbb{R}^d)$ given by

$$\varphi_f(v) = \langle v, \varphi_f \rangle_{H^1} = \int_{\mathbb{R}^d} v f \, \mathrm{d}x, \qquad v \in H^1(\mathbb{R}^d). \tag{10.8}$$

Sobolev's embedding (5.5) further shows that

$$H^1(\mathbb{R}^d) \hookrightarrow L^{1+\alpha}(\mathbb{R}^d) \qquad \text{for } 1 < \alpha < \frac{d+2}{(d-2)_+} = \alpha_c. \tag{10.9}$$

As above we deduce that

$$L^{1+\alpha}(\mathbb{R}^d)^* \cong L^{\frac{1+\alpha}{\alpha}}(\mathbb{R}^d) \hookrightarrow H^{-1}(\mathbb{R}^d), \tag{10.10}$$

where the embedding is given as in (10.8). We thus identify $L^q(\mathbb{R}^d)$ with a subspace of $H^{-1}(\mathbb{R}^d)$ for $1 + \frac{1}{\alpha_c} < q \le 2$ and write f instead of φ_f in (10.8). We stress that the duality between $H^1(\mathbb{R}^d)$ and $H^{-1}(\mathbb{R}^d)$ extends that of $L^{q'}(\mathbb{R}^d)$ and $L^q(\mathbb{R}^d)$ for these q.

Remark 10.6. (a) Corollary 9.3 shows that the nonlinearity $F(u) = -\mathrm{i}\mu|u|^{\alpha-1}u$ is real continuously differentiable and Lipschitz on bounded sets as a map $F : L^{1+\alpha}(\mathbb{R}^d) \to L^{\frac{1+\alpha}{\alpha}}(\mathbb{R}^d)$, and it satisfies $\|F'(u)\| \le c\|u\|_{1+\alpha}^{1+\alpha}$. In view of the embeddings (10.9) and (10.10), the analogous results hold if we consider F as a map $F : H^1(\mathbb{R}^d) \to H^{-1}(\mathbb{R}^d)$. $\diamondsuit$

(b) Sobolev's embedding (5.5) further yields that $H^2(\mathbb{R}^d) \hookrightarrow L^{2\alpha}(\mathbb{R}^d)$ since $2 - \frac{d}{2} \ge -\frac{d}{2\alpha}$ by (10.3). From Corollary 9.3 we thus deduce that $F : H^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is real continuously differentiable and Lipschitz on bounded sets. $\diamondsuit$

We next extend the weak partial derivative $\partial_k : H^1(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ to a bounded linear map from $L^2(\mathbb{R}^d)$ to $H^{-1}(\mathbb{R}^d)$, again denoted by ∂_k. For $u \in L^2(\mathbb{R}^d)$ we define $\partial_k u$ by

$$\langle v, \partial_k u \rangle_{H^1} = -\int_{\mathbb{R}^d} (\partial_k v) u \, \mathrm{d}x, \qquad v \in H^1(\mathbb{R}^d). \tag{10.11}$$

Clearly, $|\langle v, \partial_k u \rangle_{H^1}| \le \|u\|_2 \|v\|_{1,2}$ so that $\partial_k u$ belongs to $H^{-1}(\mathbb{R}^d)$. As a consequence, we can extend the second derivatives ∂_{kl} and the Laplacian Δ to maps in $\mathcal{B}(H^1(\mathbb{R}^d), H^{-1}(\mathbb{R}^d))$, which we also denote by ∂_{kl} and Δ, respectively. By (10.11), for $u \in H^1(\mathbb{R}^d)$ the functional Δu acts on $H^1(\mathbb{R}^d)$ as

$$\langle v, \Delta u \rangle_{H^1} = -\int_{\mathbb{R}^d} \nabla u \cdot \nabla v \, \mathrm{d}x, \qquad v \in H^1(\mathbb{R}^d). \tag{10.12}$$

One often calls this operator *weak Laplacian.*

We will need a scalar product on $H^{-1}(\mathbb{R}^d)$. To this aim, we first claim that $I - \Delta : H^1(\mathbb{R}^d) \to H^{-1}(\mathbb{R}^d)$ is an invertible isometry. In fact, for $u \in H^1(\mathbb{R}^d)$ the definitions (10.8) and (10.12) imply

$$\|u - \Delta u\|_{-1,2} = \sup_{\|v\|_{1,2}=1} |\langle \overline{v}, u - \Delta u\rangle_{H^1}| = \sup_{\|v\|_{1,2}=1} \Big| \int_{\mathbb{R}^d} \overline{v} u \,\mathrm{d}x + \int_{\mathbb{R}^d} \nabla \overline{v} \cdot \nabla u \,\mathrm{d}x \Big|$$
$$= \sup_{\|v\|_{1,2}=1} |(u|v)_{H^1}| = \|u\|_{1,2}. \tag{10.13}$$

Applying the Lax-Milgram Theorem 5.11 to the bounded and strictly accretive sesquilinear form given by $a(v,w) = (v|w)_{H^1}$ on $H^1(\mathbb{R}^d)$, we further see that $I - \Delta : H^1(\mathbb{R}^d) \to H^{-1}(\mathbb{R}^d)$ is bijective, as claimed. We then introduce

$$(\varphi|\psi)_{H^{-1}} := ((I - \Delta)^{-1}\varphi | (I - \Delta)^{-1}\psi)_{H^1} \quad \text{for } \varphi, \psi \in H^{-1}(\mathbb{R}^d). \tag{10.14}$$

It is clear that (10.14) defines a scalar product on $H^{-1}(\mathbb{R}^d)$. Setting $u - \Delta u = \varphi \in H^{-1}(\mathbb{R}^d)$ in (10.13), we see that this scalar product induces the norm $\|\cdot\|_{-1,2}$ on $H^{-1}(\mathbb{R}^d)$. We stress that one should *not* identify the Hilbert space $H^{-1}(\mathbb{R}^d)$ with its dual $H^1(\mathbb{R}^d)$ in the present context and that the scalar product in (10.14) does *not* coincide with that of $L^2(\mathbb{R}^d)$ if $\phi, \psi \in L^2(\mathbb{R}^d)$. We can now establish the relevant properties of Δ on $H^{-1}(\mathbb{R}^d)$.

LEMMA 10.7. *The operator $I - \Delta$ is an invertible isometry from $H^1(\mathbb{R}^d)$ to $H^{-1}(\mathbb{R}^d)$, and Δ with domain $H^1(\mathbb{R}^d)$ is selfadjoint and dissipative in $H^{-1}(\mathbb{R}^d)$. Thus $\mathrm{i}\Delta$ with domain $H^1(\mathbb{R}^d)$ generates a unitary C_0–group on $H^{-1}(\mathbb{R}^d)$ which extends the unitary C_0–group in $L^2(\mathbb{R}^d)$ generated by $\mathrm{i}\Delta$ with domain $H^2(\mathbb{R}^d)$.*

PROOF. We start with preparations. In this proof, we distinguish between the Laplacian Δ_{-1} in $H^{-1}(\mathbb{R}^d)$ with domain $H^1(\mathbb{R}^d)$ and the Laplacian Δ_1 in $H^1(\mathbb{R}^d)$ with domain $H^3(\mathbb{R}^d)$. The operator $I - \Delta_1$ is invertible in $H^1(\mathbb{R}^d)$, which can be seen as the corresponding result in $L^2(\mathbb{R}^d)$, cf. Example 5.10. It is straight forward to check that Δ_1 is symmetric. Moreover, $(I - \Delta_1)^{-1}$ is the restriction of $(I - \Delta_{-1})^{-1}$ to $H^1(\mathbb{R}^d)$. In fact, for $f \in H^1(\mathbb{R}^d)$ we have $u = (I - \Delta_{-1})^{-1} f \in H^1(\mathbb{R}^d)$ and $v = (I - \Delta_1)^{-1} f \in H^3(\mathbb{R}^d)$. Hence, $(I - \Delta_{-1})u = (I - \Delta_1)v = (I - \Delta_{-1})v$ and so $u = v$.

The first assertion was shown above. To verify the symmetry of Δ_{-1}, we take $u, v \in H^1(\mathbb{R}^d)$ and compute

$$\begin{aligned}
(\Delta_{-1} u \,|\, v)_{H^{-1}} &= ((I - \Delta_{-1})^{-1} \Delta_{-1} u \,|\, (I - \Delta_{-1})^{-1} v)_{H^1} \\
&= (\Delta_{-1} (I - \Delta_{-1})^{-1} u \,|\, (I - \Delta_{-1})^{-1} v)_{H^1} \\
&= (\Delta_1 (I - \Delta_1)^{-1} u \,|\, (I - \Delta_1)^{-1} v)_{H^1} \\
&= ((I - \Delta_1)^{-1} u \,|\, \Delta_1 (I - \Delta_1)^{-1} v)_{H^1} \\
&= (u \,|\, \Delta_{-1} v)_{H^{-1}} .
\end{aligned}$$

Similarly, one sees that Δ_{-1} is dissipative. Remark 5.6 (f) now implies the selfadjointness of Δ_{-1} since $I - \Delta_{-1}$ is invertible.

By Stone's Theorem 5.7, $\mathrm{i}\Delta_{-1}$ generates a unitary C_0–group on $H^{-1}(\mathbb{R}^d)$. For $u_0 \in H^2(\mathbb{R}^d)$ the corresponding Cauchy problem is also solved by the group generated by $\mathrm{i}\Delta$ on $L^2(\mathbb{R}^d)$ (because Δ_{-1} extends Δ with domain $H^2(\mathbb{R}^d)$ and

$L^2(\mathbb{R}^d) \hookrightarrow H^{-1}(\mathbb{R}^d)$). Since $H^2(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, it is densely embedded into $H^{-1}(\mathbb{R}^d)$. The uniqueness of the solutions to the Cauchy problem then yields the last assertion. □

By $T(\cdot)$ we denote both the unitary C_0–group in $H^{-1}(\mathbb{R}^d)$ constructed in Lemma 10.7 and the unitary C_0–group generated by $\mathrm{i}\Delta$ on $L^2(\mathbb{R}^d)$, which is the restriction of the first one. These groups are called *free Schrödinger groups.*

Using the above observations, we rewrite the Schrödinger equation (10.2) as the evolution equation

$$u'(t) = \mathrm{i}\Delta u(t) - \mathrm{i}\mu |u(t)|^{\alpha-1} u(t), \quad t \in J, \qquad u(0) = u_0, \tag{10.15}$$

in $H^{-1}(\mathbb{R}^d)$ or $L^2(\mathbb{R}^d) =: H^0(\mathbb{R}^d)$. We next introduce our solution concepts.

Definition 10.8. *Let* $1 < \alpha < \frac{d+2}{(d-2)_+} = \alpha_c$, $\mu \in \{-1, 1\}$ *and* $k \in \{1, 2\}$. *Let* $u_0 \in H^k(\mathbb{R}^d)$. *An* H^k*–solution of* (10.15) *is a function* $u \in C(J, H^k(\mathbb{R}^d)) \cap C^1(J, H^{k-2}(\mathbb{R}^d))$ *that satisfies* (10.15)*.*

If u is an H^1-solution of (10.15), then $F \circ u : J \to H^{-1}(\mathbb{R}^d)$ is continuous by Remark 10.6. By Proposition 6.5, u thus satisfies the integral equation

$$u(t) = T(t)u_0 + \int_0^t T(t-s)F(u(s))\,\mathrm{d}s, \quad t \in J, \tag{10.16}$$

in $H^{-1}(\mathbb{R}^d)$, where $u_0 \in H^1(\mathbb{R}^d)$ and the integral exists in the space $H^{-1}(\mathbb{R}^d)$.

As in Lecture 8, we will establish the local wellposedness theory for the nonlinear Schrödinger equation by solving the fixed point problem (10.16). But one sees right away that this is a much harder task than before since F only maps $L^{1+\alpha}(\mathbb{R}^d)$ into $L^{\frac{1+\alpha}{\alpha}}(\mathbb{R}^d)$ and thus loses integrability. At first glance, the convolution with $T(\cdot)$ does not help since $T(t)$ is unitary on $H^{-1}(\mathbb{R}^d)$ or $L^2(\mathbb{R}^d)$ and hence does not map into a "better" space. However, if one leaves the L^2 framework, one can show that $T(t)$ in fact improves integrability to some extent. This property is encoded in Strichartz' estimates proved in the next lecture.

As a first step in this direction, we derive a representation formula for $T(t)$ on $L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. It says that on this space $T(t)$ looks like the diffusion semigroup with "imaginary time" $\mathrm{i}t$. The formula implies that $T(t)v \in C^\infty(\mathbb{R}^d)$ if $v \in L^2(\mathbb{R}^d)$ has compact support.

Lemma 10.9. *The free Schrödinger group is given by*

$$T(t)v(x) = \frac{1}{(4\pi \mathrm{i}t)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\frac{|x-y|^2}{4t}} v(y)\,\mathrm{d}y \tag{10.17}$$

for all $v \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, $t \in \mathbb{R} \setminus \{0\}$, *and a.e.* $x \in \mathbb{R}^d$.

Proof. 1) Observe that the right-hand side of (10.17) defines a bounded map from $L^1(\mathbb{R}^d)$ to $L^\infty(\mathbb{R}^d)$ for $t \neq 0$. Moreover, $C_c^\infty(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ with respect to the sum norm $\|\cdot\|_1 + \|\cdot\|_2$. (This fact can be seen by an inspection of part 1) in the proof of Theorem D.13.) It thus suffices to show (10.17) for $v \in C_c^\infty(\mathbb{R}^d)$.

2) To derive (10.17), we will use the Fourier transform $\mathcal{F} : \varphi \mapsto \widehat{\varphi}$. We recall that $\mathcal{F}$ is a unitary operator on $L^2(\mathbb{R}^d)$ and that for $\varphi \in H^2(\mathbb{R}^d)$ we

have $\mathcal{F}(\Delta\varphi) = -|\xi|^2\mathcal{F}\varphi$, see Theorem 5.9 or Appendix E. For $v \in C_c^\infty(\mathbb{R}^d)$, the function $u = T(\cdot)v$ belongs to $C(\mathbb{R}, H^2(\mathbb{R}^d)) \cap C^1(\mathbb{R}, L^2(\mathbb{R}^d))$ and satisfies $u'(t) = \mathrm{i}\Delta u(t)$ for $t \in \mathbb{R}$. Since

$$\tfrac{1}{t'-t}(\mathcal{F}u(t') - \mathcal{F}u(t)) = \mathcal{F}\big(\tfrac{1}{t'-t}(u(t') - u(t))\big)$$

tends to $\mathcal{F}u'(t)$ in $L^2(\mathbb{R}^d)$ as $t' \to t$, it follows that $\widehat{u} \in C^1(\mathbb{R}, L^2(\mathbb{R}^d))$ and $\widehat{u}'(t) = \mathcal{F}u'(t)$. We thus obtain

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\,\widehat{u}(t) = \mathcal{F}(\mathrm{i}\Delta u(t)) = -\mathrm{i}|\xi|^2\widehat{u}(t), \quad t \in \mathbb{R}, \qquad \widehat{u}(0) = \widehat{v}.$$

Solving this ordinary differential equation for fixed $\xi \in \mathbb{R}^d$, we arrive at

$$\widehat{u}(t,\xi) = \mathrm{e}^{-\mathrm{i}t|\xi|^2}\widehat{v}(\xi) = \gamma_{\mathrm{i}t}(\xi)\widehat{v}(\xi)$$

for all $t \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$, where $\gamma_z(\xi) := \mathrm{e}^{-z|\xi|^2}$ for $\xi \in \mathbb{R}^d$ and $z \in \mathbb{C}$, $\gamma_{\mathrm{i}t} \in C_b(\mathbb{R}^d)$ and $\widehat{v} \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. As a result, $u(t) = \mathcal{F}^{-1}(\gamma_{\mathrm{i}t}\widehat{v})$. Since $\gamma_{\mathrm{i}t}$ is not the Fourier transform of an L^1-function, we cannot directly apply the convolution formula in Theorem 5.9 (b). Instead we consider the regularization $m_\varepsilon(t) = \gamma_{\mathrm{i}t+\varepsilon} \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ for $\varepsilon > 0$. We compute below $\mathcal{F}^{-1}m_\varepsilon$ and obtain $\mathcal{F}^{-1}m_\varepsilon \in L^1(\mathbb{R}^d)$. Hence, we can apply Theorem 5.9 (b) to $m_\varepsilon\widehat{v}$.

Since $|m_\varepsilon| \le \mathbb{1}$ and $m_\varepsilon(t)$ converges pointwise to $\gamma_{\mathrm{i}t}$, Lebesgue's theorem and Theorem 5.9 thus imply that

$$u(t) = \mathcal{F}^{-1}(\gamma_{\mathrm{i}t}\widehat{v}) = \lim_{\varepsilon\to0}\mathcal{F}^{-1}(m_\varepsilon(t)\widehat{v}) = \lim_{\varepsilon\to0}(2\pi)^{-\frac{d}{2}}(\mathcal{F}^{-1}m_\varepsilon(t)) * v$$

for all $t \in \mathbb{R}$. Using $m_\varepsilon \in L^1(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)$ and Theorem 5.9, we further compute

$$[\mathcal{F}^{-1}m_\varepsilon(t)](x) = (2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d}\mathrm{e}^{\mathrm{i}x\cdot\xi}\mathrm{e}^{-\mathrm{i}t|\xi|^2}\mathrm{e}^{-\varepsilon|\xi|^2}\,\mathrm{d}\xi = \prod_{k=1}^d\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\mathrm{e}^{\mathrm{i}x_k\xi_k-(\mathrm{i}t+\varepsilon)\xi_k^2}\,\mathrm{d}\xi_k$$

for all $t \neq 0$ and $x \in \mathbb{R}^d$. By means of complex contour integrals we show in step 3) that

$$\int_{\mathbb{R}}\mathrm{e}^{-(\mathrm{i}t+\varepsilon)s^2}\mathrm{e}^{\mathrm{i}x_ks}\,\mathrm{d}s = \sqrt{\frac{\pi}{\mathrm{i}t+\varepsilon}}\,\mathrm{e}^{\frac{-x_k^2}{4(\mathrm{i}t+\varepsilon)}}. \tag{10.18}$$

Hence, $\mathcal{F}^{-1}m_\varepsilon \in L^1(\mathbb{R}^d)$ as needed above, and we arrive at

$$u(t,x) = \frac{1}{(4\pi)^{\frac{d}{2}}}\lim_{\varepsilon\to0}\int_{\mathbb{R}^d}\frac{1}{(\mathrm{i}t+\varepsilon)^{\frac{d}{2}}}\mathrm{e}^{-\frac{|x-y|^2}{4(\mathrm{i}t+\varepsilon)}}v(y)\,\mathrm{d}y.$$

For fixed $t \neq 0$ and $x \in \mathbb{R}^d$, Lebesgue's theorem allows to let $\varepsilon \to 0$ in the integral since $v \in C_c^\infty(\mathbb{R}^d)$, and hence (10.9) holds.

3) It remains to check (10.18), where $x := x_k \in \mathbb{R}$, $t \in \mathbb{R}$ and $\varepsilon > 0$ are fixed. We set $\zeta = \frac{\mathrm{i}x}{2(\mathrm{i}t+\varepsilon)}$ and compute

$$\int_{\mathbb{R}}\mathrm{e}^{-(\mathrm{i}t+\varepsilon)s^2}\mathrm{e}^{\mathrm{i}xs}\,\mathrm{d}s = \mathrm{e}^{-\frac{x^2}{4(\mathrm{i}t+\varepsilon)}}\int_{\mathbb{R}}\mathrm{e}^{-(\mathrm{i}t+\varepsilon)(s-\zeta)^2}\,\mathrm{d}s = \mathrm{e}^{-\frac{x^2}{4(\mathrm{i}t+\varepsilon)}}\int_{\mathbb{R}-\zeta}\mathrm{e}^{-(\mathrm{i}t+\varepsilon)z^2}\,\mathrm{d}z.$$

We set $f(z) = \mathrm{e}^{-(\mathrm{i}t+\varepsilon)z^2}$ and $I = \int_{\mathbb{R}-\zeta} f \,\mathrm{d}z$. We have to show that $I = \sqrt{\frac{\pi}{\mathrm{i}t+\varepsilon}}$. To this purpose, we consider the contour $\Gamma_n = \Gamma_n^{\mathrm{b}} \cup \Gamma_n^{\mathrm{r}} \cup \Gamma_n^{\mathrm{t}} \cup \Gamma_n^{\mathrm{l}}$, where

$$\Gamma_n^{\mathrm{b}} = [-n, n], \qquad \Gamma_n^{\mathrm{r}} = \{z = n + \tau\zeta \mid -1 \le \tau \le 0\},$$
$$\Gamma_n^{\mathrm{t}} = \{z = \tau n - \zeta \mid -1 \le \tau \le 1\}, \qquad \Gamma_n^{\mathrm{l}} = \{z = -n + \tau\zeta \mid -1 \le \tau \le 0\}$$

for $n \in \mathbb{N}$, where Γ_n is oriented counterclockwise. Cauchy's theorem yields $\int_{\Gamma_n} f \,\mathrm{d}z = 0$. We can find a constant $c > 0$ such that

$$\sup_{z \in \Gamma_n^{\mathrm{r}} \cup \Gamma_n^{\mathrm{l}}} |\mathrm{e}^{-(\mathrm{i}t+\varepsilon)z^2}| \le \mathrm{e}^{-\varepsilon n^2} \mathrm{e}^{cn}$$

for all $n \in \mathbb{N}$, and hence $\int_{\Gamma_n^j} f \,\mathrm{d}z \to 0$ as $n \to \infty$, for $j = \mathrm{l}, \mathrm{r}$. By a similar estimate one sees that $\int_{\Gamma_n^{\mathrm{t}}} f \,\mathrm{d}z$ tends to I as $n \to \infty$. Letting $n \to \infty$, we thus deduce that $I = 2\int_0^\infty f(s) \,\mathrm{d}s$.

Let $t > 0$ and set $\beta = \frac{1}{2} \arg(\mathrm{i}t + \varepsilon) \in (0, \frac{\pi}{4})$. Since $|\sqrt{\mathrm{i}t+\varepsilon}|\, \mathrm{e}^{\mathrm{i}\beta} = \sqrt{\mathrm{i}t+\varepsilon}$, we have

$$I = \frac{2}{\sqrt{\mathrm{i}t+\varepsilon}} \int_{\mathrm{e}^{\mathrm{i}\beta}\mathbb{R}_+} \mathrm{e}^{-z^2} \,\mathrm{d}z.$$

To evaluate this integral, we use the contour $\Gamma_n = \mathrm{e}^{\mathrm{i}\beta}[0, n] \cup \{n\mathrm{e}^{\mathrm{i}\sigma} \mid 0 \le \sigma \le \beta\} \cup [0, n]$ (with positive orientation). Since $|\mathrm{e}^{-n^2 \mathrm{e}^{2\mathrm{i}\sigma}}| \le \mathrm{e}^{-n^2 \cos 2\beta}$ for $\sigma \in [0, \beta]$, Cauchy's theorem now yields

$$I = \frac{2}{\sqrt{\mathrm{i}t+\varepsilon}} \int_0^\infty \mathrm{e}^{-s^2} \,\mathrm{d}s = \frac{\sqrt{\pi}}{\sqrt{\mathrm{i}t+\varepsilon}},$$

as asserted. The case $t \le 0$ is treated in the same way. □

This representation formula allows to establish the *dispersive* behavior of $T(t)$. Indeed, the next corollary says that $T(t)$ flattens initial data in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, in the sense that they become immediately bounded and then tend to 0 in all L^q–norms for $q > 2$ as $t \to \infty$. Since the L^2–norm is preserved, local concentrations of $T(t)v$ must be pushed towards infinity in $\mathbb{R}^d$.

Corollary 10.10. *The operators $T(t)$ from Lemma 10.9 extend from $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ to operators in $\mathcal{B}(L^{p'}(\mathbb{R}^d), L^p(\mathbb{R}^d))$ for all $p \in [2, \infty]$ and $t \in \mathbb{R} \setminus \{0\}$, with norm bounded by $(4\pi|t|)^{d(\frac{1}{p}-\frac{1}{2})}$.*

Proof. By Lemma 10.9, $T(t)$ maps $(L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d), \|\cdot\|_1)$ into $L^\infty(\mathbb{R}^d)$ with norm bounded by $(4\pi|t|)^{-\frac{d}{2}}$. Moreover, it has norm 1 as an operator on $L^2(\mathbb{R}^d)$. Let $p \in [2, \infty]$. The Riesz-Thorin interpolation theorem now shows that we can extend $T(t)$ to an operator from $L^{p'}(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d)$ with norm less or equal $(4\pi|t|)^{\frac{d}{p}-\frac{d}{2}}$, see e.g. Theorem IX.17 in [**RS75**] and use $t = \frac{2}{p}$ there. □

LECTURE 11

Strichartz' estimates

In this lecture we establish Strichartz' estimates for the linear Schrödinger equation $u' = \mathrm{i}\Delta u$ in the subcritical case. These estimates will be the crucial ingredient for our further analysis of the nonlinear Schrödinger equation. For the proof of these estimates we need Corollary 10.10 of the previous lecture, the Hardy-Littlewood-Sobolev inequality and a few other preparations. We state, discuss and show Strichartz' estimates at the end of this lecture.

To motivate the need for Strichartz' estimates, we sketch the way to solve the nonlinear Schrödinger equation

$$u'(t) = \mathrm{i}\Delta u(t) - \mathrm{i}\mu|u(t)|^{\alpha-1}u(t), \quad t \in J, \qquad u(0) = u_0. \tag{11.1}$$

Here one cannot apply the results of Lecture 8 since the nonlinearity does not map $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^d)$. We still want to use the *methods* of Lecture 8. In this spirit we look at the integrated version of (11.1),

$$u(t) = T(t)u_0 - \mathrm{i}\mu \int_0^t T(t-s)u(s)|u(s)|^{\alpha-1}\,\mathrm{d}s, \quad t \in J, \tag{11.2}$$

see (10.16), where $T(\cdot)$ is the unitary C_0–group generated by $\mathrm{i}\Delta$. The right-hand side of (11.2) defines an operator Φ whose fixed point will solve (11.1). To apply the contraction mapping principle, we need a function space which is mapped into itself by Φ. Because of $\alpha > 1$, we lose integrability in the nonlinearity. However, the one-sided convolution with $T(\cdot)$, given by

$$T *_+ f(t) = \int_0^t T(t-s)f(s)\,\mathrm{d}s, \quad t \in J,$$

could regain some integrability. There is hope for such a result since we have seen in Corollary 10.10 that $T(t)$ defined on $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ has a bounded extension $T(t) : L^{p'}(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ with norm

$$\|T(t)\|_{\mathcal{B}(L^{p'},L^p)} \le (4\pi|t|)^{\frac{d}{p}-\frac{d}{2}} \tag{11.3}$$

for all $p \in [2,\infty]$ and $t \in \mathbb{R} \setminus \{0\}$. One calls (11.3) a *dispersive estimate* since it says that initial values in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ are mapped to bounded functions by the free Schrödinger group $T(\cdot)$. Because the L^2–norm is preserved, the solution must spread out in space.

To treat the fixed point problem (11.2), we have to bound $T *_+ f$ and the orbit $T(\cdot)u_0$. The needed inequalities will follow from (11.3), but this derivation fills the entire lecture. The resulting estimates are named after R. Strichartz who proved such estimates for the wave equation in 1977 (see [**Str77**]). Various versions were then shown in the following years, see Section 2.3 in [**Caz03**]. The development culminated in the paper [**KT98**] from 1998 by M. Keel and T. Tao who established the result in a certain borderline case (described below) which is needed for the critical case $\alpha = \alpha_c$ in (11.1), cf. (10.3). By now, Strichartz type estimates are known for a wide range of dispersive equations.

We will show a basic version of Strichartz' estimates in the subcritical case for the free Schrödinger group. This result fits to our problem (11.1) with $1 < \alpha < \alpha_c$. To exploit the dispersive estimate (11.3), one needs the Hardy-Littlewood-Sobolev inequality for the space dimension $n = 1$, see (11.4), which is of independent interest.

To derive this inequality, we first represent the L^p–norm of a function $f : \mathbb{R}^n \to \mathbb{C}$ in terms of its *distribution function* d_f defined by

$$d_f(s) = \int_{\mathbb{R}^n} \mathbb{1}_{\{|f|>s\}}(x)\,\mathrm{d}x, \qquad s \geq 0.$$

Here we use the notation $\{|f| > s\} = \{\xi \in \mathbb{R}^n \mid |f(\xi)| > s\}$.

LEMMA 11.1. *For all $p \in [1,\infty)$ and $f \in L^p(\mathbb{R}^n)$ we have*

$$\|f\|_p^p = p\int_0^\infty s^{p-1} d_f(s)\,\mathrm{d}s.$$

PROOF. Using Fubini's theorem, we compute

$$\begin{aligned}\|f\|_p^p &= \int_{\mathbb{R}^n} |f(x)|^p\,\mathrm{d}x = \int_{\mathbb{R}^n}\Big(\int_0^{|f(x)|} ps^{p-1}\,\mathrm{d}s\Big)\,\mathrm{d}x\\ &= p\int_0^\infty s^{p-1}\Big(\int_{\mathbb{R}^n} \mathbb{1}_{(0,|f(x)|)}(s)\,\mathrm{d}x\Big)\,\mathrm{d}s.\end{aligned}$$

The assertion now follows from the fact that $\mathbb{1}_{(0,|f(x)|)}(s) = \mathbb{1}_{\{|f|>s\}}(x)$. $\square$

We next show the Hardy-Littlewood-Sobolev inequality. Its proof is elementary and laborious, but quite entertaining.

THEOREM 11.2 (Hardy-Littlewood-Sobolev inequality). *Let $\beta, \gamma \in (1,\infty)$ and $0 < \lambda < n$ satisfy $\frac{1}{\beta} + \frac{\lambda}{n} + \frac{1}{\gamma} = 2$. Then there is a constant $C > 0$ such that*

$$\int_{\mathbb{R}^n} |f(x)| \int_{\mathbb{R}^n} \frac{|g(y)|}{|x-y|^\lambda}\,\mathrm{d}y\,\mathrm{d}x \leq C\|f\|_\beta \|g\|_\gamma \tag{11.4}$$

for all $f \in L^\beta(\mathbb{R}^n)$ and all $g \in L^\gamma(\mathbb{R}^n)$. As a result,

$$\Big(\int_{\mathbb{R}^n}\Big[\int_{\mathbb{R}^n} \frac{|g(y)|}{|x-y|^\lambda}\,\mathrm{d}y\Big]^{\beta'}\,\mathrm{d}x\Big)^{\frac{1}{\beta'}} \leq C\|g\|_\gamma\,. \tag{11.5}$$

The inequality (11.5) resembles Young's convolution estimate. (See Theorem 1.2.12 in [**Gra08**].) To see this relation, let $\varphi_\lambda(x) = |x|^{-\lambda}$ for $x \in \mathbb{R}^n \setminus \{0\}$ and $\varphi_\lambda(0) = 0$, where $\lambda \in (0,n)$. If we had $\varphi_\lambda \in L^r(\mathbb{R}^n)$ with $\frac{1}{r} = 1 + \frac{1}{\beta'} - \frac{1}{\gamma}$ and $r, \beta, \gamma \in [1,\infty]$, then Young's inequality would give

$$\|\varphi_\lambda * g\|_{\beta'} \leq \|\varphi_\lambda\|_r \|g\|_\gamma \qquad \text{for all } g \in L^\gamma(\mathbb{R}^n). \tag{11.6}$$

To show (11.5), we would need here $\lambda = n(2 - \frac{1}{\beta} - \frac{1}{\gamma})$, i.e., $\lambda r = n$. So $\varphi_\lambda \in L^r(\mathbb{R}^n)$ would require that $\int_{\mathbb{R}^n} |x|^{-n}\,\mathrm{d}x < \infty$ which is not quite true. In (11.5) we have shown (11.6) (at least for positive g) with $\|\varphi_\lambda\|_r$ replaced by C.

Proof. 1) Assertion (11.5) follows from (11.4) by a duality argument. For (11.4) it suffices to prove the inequality

$$I(f,g) := \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{f(x)g(y)}{|x-y|^\lambda}\,\mathrm{d}x\,\mathrm{d}y \le C \tag{11.7}$$

under the assumption that $f, g \ge 0$ and $\|f\|_\beta = \|g\|_\gamma = 1$. The derivation of (11.7) is based on Lemma 11.1 and the distribution functions d_f and d_g. As in the proof of Lemma 11.1, for $x \in \mathbb{R}^n$ we have

$$f(x) = \int_0^\infty \mathbb{1}_{(0,f(x))}(r)\,\mathrm{d}r = \int_0^\infty \mathbb{1}_{\{f>r\}}(x)\,\mathrm{d}r,$$

and in the same way $g(y) = \int_0^\infty \mathbb{1}_{\{g>s\}}(y)\,\mathrm{d}s$ for $y \in \mathbb{R}^n$. Moreover,

$$|x|^{-\lambda} = \lambda\int_{|x|}^\infty t^{-\lambda-1}\,\mathrm{d}t = \lambda\int_0^\infty t^{-\lambda-1}\mathbb{1}_{(|x|,\infty)}(t)\,\mathrm{d}t = \lambda\int_0^\infty t^{-\lambda-1}\mathbb{1}_{B(0,t)}(x)\,\mathrm{d}t.$$

These facts and Fubini's theorem yield

$$I(f,g) = \lambda\int_0^\infty\!\!\int_0^\infty\!\!\int_0^\infty t^{-\lambda-1}J(r,s,t)\,\mathrm{d}t\,\mathrm{d}s\,\mathrm{d}r, \tag{11.8}$$

where we have set

$$J(r,s,t) := \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \mathbb{1}_{\{f>r\}}(x)\mathbb{1}_{\{g>s\}}(y)\mathbb{1}_{B(0,t)}(x-y)\,\mathrm{d}x\,\mathrm{d}y.$$

2) Let $r, s, t > 0$. We estimate $J(r,s,t)$ in three ways. We first derive

$$\begin{aligned} J(r,s,t) &\le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \mathbb{1}_{\{g>s\}}(y)\mathbb{1}_{B(0,t)}(x-y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \mathbb{1}_{\{g>s\}}(y)\int_{B(y,t)} 1\,\mathrm{d}x\,\mathrm{d}y = C_1 d_g(s)t^n, \end{aligned} \tag{11.9}$$

where $C_1 = \mathrm{vol}(B(0,1))$. Similarly, one obtains

$$J(r,s,t) \le C_1 d_f(r)t^n, \qquad J(r,s,t) \le d_f(r)\,d_g(s). \tag{11.10}$$

3) We fix $r, s > 0$ and estimate the inner integral in (11.8) in two ways. We first split the integral at $t = d_f(r)^{\frac{1}{n}}$ and use (11.9) for the first summand and the second part of (11.10) for the second summand, arriving at

$$\begin{aligned} \int_0^\infty t^{-\lambda-1}J(r,s,t)\,\mathrm{d}t &\le C_1\int_0^{d_f(r)^{\frac{1}{n}}} d_g(s)t^{-\lambda-1+n}\,\mathrm{d}t + \int_{d_f(r)^{\frac{1}{n}}}^\infty d_f(r)d_g(s)t^{-\lambda-1}\,\mathrm{d}t \\ &= \frac{C_1}{n-\lambda}d_g(s)d_f(r)^{\frac{n-\lambda}{n}} + \frac{1}{\lambda}d_f(r)d_g(s)d_f(r)^{-\frac{\lambda}{n}} \\ &= \frac{\lambda C_1 + n - \lambda}{\lambda(n-\lambda)}d_f(r)^{1-\frac{\lambda}{n}}d_g(s). \end{aligned}$$

Second, splitting at $t = d_g(s)^{\frac{1}{n}}$ and applying both parts of (11.10), we deduce

$$\int_0^\infty t^{-\lambda-1}J(r,s,t)\,\mathrm{d}t \le \frac{\lambda C_1 + n - \lambda}{\lambda(n-\lambda)}d_f(r)d_g(s)^{1-\frac{\lambda}{n}}.$$

Inserting these inequalities into (11.8), we conclude

$$I(f,g) \le \frac{n(C_1+1)}{n-\lambda} \int_0^\infty \int_0^\infty \min\{d_f(r)^{1-\frac{\lambda}{n}} d_g(s), d_f(r) d_g(s)^{1-\frac{\lambda}{n}}\} \, \mathrm{d}s \, \mathrm{d}r. \tag{11.11}$$

4) We next split the inner integral in (11.11) at $s = r^{\frac{\beta}{\gamma}}$ and obtain

$$\begin{aligned} I(f,g) \le \frac{n(C_1+1)}{n-\lambda} & \int_0^\infty d_f(r) \Big(\int_0^{r^{\beta/\gamma}} d_g(s)^{1-\frac{\lambda}{n}} \, \mathrm{d}s \Big) \, \mathrm{d}r \\ & + \frac{n(C_1+1)}{n-\lambda} \int_0^\infty d_f(r)^{1-\frac{\lambda}{n}} \Big(\int_{r^{\beta/\gamma}}^\infty d_g(s) \, \mathrm{d}s \Big) \, \mathrm{d}r. \end{aligned} \tag{11.12}$$

In the first summand we apply Hölder's inequality with the exponents $(1 - \frac{\lambda}{n})^{-1} > 1$ and $\frac{n}{\lambda} > 1$. Lemma 11.1 and $\|g\|_\gamma = 1$ then imply

$$\begin{aligned} \int_0^{r^{\beta/\gamma}} d_g(s)^{1-\frac{\lambda}{n}} \, \mathrm{d}s &= \int_0^{r^{\beta/\gamma}} \big(d_g(s)^{1-\frac{\lambda}{n}} s^{(\gamma-1)(1-\frac{\lambda}{n})} \big) s^{-(\gamma-1)(1-\frac{\lambda}{n})} \, \mathrm{d}s \\ &\le \Big(\int_0^{r^{\beta/\gamma}} d_g(s) s^{\gamma-1} \, \mathrm{d}s \Big)^{1-\frac{\lambda}{n}} \Big(\int_0^{r^{\beta/\gamma}} s^{-(\gamma-1)(\frac{n}{\lambda}-1)} \, \mathrm{d}s \Big)^{\frac{\lambda}{n}} \\ &\le \big(\tfrac{1}{\gamma} \|g\|_\gamma^\gamma \big)^{1-\frac{\lambda}{n}} \Big(\int_0^{r^{\beta/\gamma}} s^{-(\gamma-1)(\frac{n}{\lambda}-1)} \, \mathrm{d}s \Big)^{\frac{\lambda}{n}} \le C_2 r^{\beta-1} \end{aligned}$$

for a constant C_2 only depending on β, λ and n, where the assumption $\frac{1}{\beta} + \frac{\lambda}{n} + \frac{1}{\gamma} = 2$ yields $(\gamma-1)(\frac{n}{\lambda}-1) < 1$ as well as $\frac{\beta}{\gamma}(1-(\gamma-1)(\frac{n}{\lambda}-1))\frac{\lambda}{n} = \beta - 1$. Employing again Lemma 11.1, we infer

$$\int_0^\infty d_f(r) \Big(\int_0^{r^{\beta/\gamma}} d_g(s)^{1-\frac{\lambda}{n}} \, \mathrm{d}s \Big) \, \mathrm{d}r \le C_2 \int_0^\infty d_f(r) r^{\beta-1} \, \mathrm{d}r = \tfrac{C_2}{\beta} \|f\|_\beta^\beta = \tfrac{C_2}{\beta}.$$

The second double integral in (11.12) can be rewritten as

$$\begin{aligned} \int_0^\infty d_f(r)^{1-\frac{\lambda}{n}} \Big(\int_{r^{\beta/\gamma}}^\infty d_g(s) \, \mathrm{d}s \Big) \, \mathrm{d}r &= \int_0^\infty \int_0^\infty \mathbb{1}_{(r^{\beta/\gamma},\infty)}(s) d_f(r)^{1-\frac{\lambda}{n}} d_g(s) \, \mathrm{d}s \, \mathrm{d}r \\ &= \int_0^\infty \int_0^\infty \mathbb{1}_{(0,s^{\gamma/\beta})}(r) d_f(r)^{1-\frac{\lambda}{n}} d_g(s) \, \mathrm{d}r \, \mathrm{d}s \\ &= \int_0^\infty d_g(s) \Big(\int_0^{s^{\gamma/\beta}} d_f(r)^{1-\frac{\lambda}{n}} \, \mathrm{d}r \Big) \, \mathrm{d}s, \end{aligned}$$

due to Fubini's theorem. This term can be bounded in the same way as above by a constant C_3. We have thus shown (11.7). □

In the proof of Strichartz' estimates we have to commute partial derivatives with the free Schrödinger group $T(\cdot)$, which is justified by the next lemma.

LEMMA 11.3. *For each $k \in \mathbb{N}$, the free Schrödinger group $T(\cdot)$ leaves $H^k(\mathbb{R}^d)$ invariant and induces a unitary C_0–group on $H^k(\mathbb{R}^d)$. Moreover, $\partial_j T(t) v = T(t) \partial_j v$ for all $t \in \mathbb{R}$, $j \in \{1, \dots, d\}$ and $v \in H^1(\mathbb{R}^d)$. An analogous result holds for higher derivatives.*

Proof. Let $t \in \mathbb{R} \setminus \{0\}$ and $k = 1$. We first show $T(t)v \in H^1(\mathbb{R}^d)$ for $v \in C_c^\infty(\mathbb{R}^d)$. Due to Lemma 10.9, there is a kernel $K_t \in C_b(\mathbb{R}^d)$ such that

$$T(t)v(x) = (K_t * v)(x) = \int_{\mathbb{R}^d} K_t(y)v(x-y)\,\mathrm{d}y$$

for $x \in \mathbb{R}^d$. Thus $T(t)v \in C^1(\mathbb{R}^d)$ and

$$\partial_j T(t)v = \int_{\mathbb{R}^d} K_t(y)(\partial_j v)(\cdot - y)\,\mathrm{d}y = T(t)(\partial_j v)$$

on $\mathbb{R}^d$. Next, let $v \in H^1(\mathbb{R}^d)$. By Remark 5.8 (c), there are functions $v_n \in C_c^\infty(\mathbb{R}^d)$ that converge to v in $H^1(\mathbb{R}^d)$ as $n \to \infty$. Hence, $\partial_j T(t)v_n = T(t)\partial_j v_n$ tend to $T(t)\partial_j v$ in $L^2(\mathbb{R}^d)$. Remark 5.8 (b) thus yields that $T(t)v \in H^1(\mathbb{R}^d)$ and $\partial_j T(t)v = T(t)\partial_j v$ for $j \in \{1, \dots, d\}$. It follows that $T(t)$, $t \in \mathbb{R}$, can be restricted to an isometry on $H^1(\mathbb{R}^d)$. Clearly, the restriction still satisfies the group property. Since $\|\partial_j(T(t)v - v)\|_2 = \|T(t)\partial_j v - \partial_j v\|_2 \to 0$ as $t \to 0$, the group $T(\cdot)$ is strongly continuous on $H^1(\mathbb{R}^d)$. By the group law the operators $T(t) : H^1(\mathbb{R}^d) \to H^1(\mathbb{R}^d)$ are bijective and hence unitary due to Theorem C.7. The case $k \geq 2$ is treated similarly. $\square$

Strichartz' estimate will be formulated in the Banach space-valued Lebesgue space $L^q(J, W_p^k(\mathbb{R}^d))$ with $p, q \in [1, \infty]$, $k \in \mathbb{N}_0$ and an interval $J \subseteq \mathbb{R}$, where we recall our notations $W_p^0 = L^p$ and $H^0 = L^2$. In Appendix F we have collected the definitions and some of the basic properties of these spaces. We further need the following results.

Remark 11.4. (a) Let X be reflexive and $1 \leq q < \infty$. Then $L^q(J, X)^*$ is isometrically isomorphic to $L^{q'}(J, X^*)$, where $g \in L^{q'}(J, X^*)$ acts via

$$\langle f, g \rangle_{L^q(J,X)} = \int_J \langle f(t), g(t) \rangle_X \,\mathrm{d}t$$

on $f \in L^q(J, X)$. Moreover, $L^q(J, X)$ is reflexive if $1 < q < \infty$. (See Theorems 8.20.3 and 8.20.5 in [**Edw65**].) $\diamondsuit$

(b) Let X be separable and $1 \leq q < \infty$. Then the space $L^q(J, X)$ is separable. In fact, in Lemma F.13 we have seen that the set of X-valued simple functions is dense in $L^q(J, X)$ if $1 \leq q < \infty$. Standard properties of the Lebesgue measure allow to approximate a simple function by sums of functions of the form $\mathbb{1}_Q x$ for $x \in X$ and an interval $Q \subseteq \mathbb{R}$ with rational vertices. Inserting here a dense sequence $(x_n)_n$ in X, we obtain a dense countable subset of $L^q(J, X)$. $\diamondsuit$

In some of our arguments we have to approximate a function in $L^q(J, W_p^k(\mathbb{R}^d))$ by smoother ones. In the next lemma we show a corresponding density result.

Lemma 11.5. *Let $J \subseteq \mathbb{R}$ be an open interval, $k \in \mathbb{N}_0$ and $1 \leq p, q < \infty$. For each $f \in L^q(J, W_p^k(\mathbb{R}^d)) =: E$ there are $\varphi_n \in \bigcap_{m \in \mathbb{N}_0, r \in [1,\infty]} C_c^\infty(J, W_r^m(\mathbb{R}^d))$ converging to f in E as $n \to \infty$.*

Proof. Let $f \in E$, $m \in \mathbb{N}_0$ and $r \in [1, \infty]$. Let $\varepsilon > 0$. We proceed in four steps involving cutoff and mollification in space and time, cf. Appendix D.

1) Let $\phi \in C_c^\infty(\mathbb{R}^d)$ be equal to 1 on $\overline{B}(0,1)$ and have support in $B(0,2)$. Set $\phi_n(x) = \phi(\frac{1}{n}x)$. Then $\phi_n = 1$ on $\overline{B}(0,n)$, $\operatorname{supp}\phi_n \subseteq B(0,2n)$ and all derivatives

of ϕ_n of order $k \in \mathbb{N}_0$ are uniformly bounded by $n^{-k}\|\phi\|_{C^k}$. As a result, $\phi_n f(t)$ belongs to $W_p^k(\mathbb{R}^d)$ and has support in $B(0,2n)$ for each $n \in \mathbb{N}$ and a.e. $t \in J$. Moreover, the functions $\phi_n f(t)$ tend to $f(t)$ in $W_p^k(\mathbb{R}^d)$ as $n \to \infty$, for a.e. $t \in J$. (See part 1) of the proof of Theorem D.13.) Lebesgue's theorem (with majorant $c_k\|f(t)\|_{k,p}$) thus shows that $\phi_n f \to f$ in E. We can now fix N such that $g = \phi_N f \in E$ satisfies $\|f - g\|_E \le \varepsilon$ and $\operatorname{supp} g(t) \subseteq B(0,2N)$ for a.e. $t \in J$.

2) We next use the mollifiers G_n on $W_p^k(\mathbb{R}^d)$ given by $G_n v = \rho_n * v$, where $\rho_n(x) = n^d\|\chi\|_1^{-1}\chi(nx)$ for $n \in \mathbb{N}$ and $x \in \mathbb{R}$ as well as $\chi(x) = \exp\frac{1}{|x|^2-1}$ for $|x| < 1$ and $\chi(x) = 0$ for $|x| \ge 1$, see (D.3). Let $m \in \mathbb{N}_0$ and $r \in [1,\infty]$. If $v \in L^1(\mathbb{R}^d)$ has support in $B(0,R)$, one easily checks that $\operatorname{supp} G_n v \subseteq B(0,R+1)$, $G_n v \in C^\infty(\mathbb{R}^d)$ and $\|G_n v\|_{m,r} \le C(m,r,R)\|v\|_1$ for all $n \in \mathbb{N}$ and a constant only depending on m, r, R and χ. Hence, $G_n g \in L^q(J, W_r^m(\mathbb{R}^d))$. The standard properties of G_n stated in (D.6) – (D.7) and Lemma D.6 (a) imply that $\|G_n g(t)\|_{k,p} \le \|g(t)\|_{k,p}$ for all $n \in \mathbb{N}$ and $G_n g(t) \to g(t)$ in $W_p^k(\mathbb{R}^d)$ as $n \to \infty$ for a.e. $t \in J$. As in step 1), we can find an index N such that $h = G_N g \in L^q(J, W_r^m(\mathbb{R}^d)) \cap E$ satisfies $\|g - h\|_E \le \varepsilon$.

3) Let $J_n \subseteq \overline{J_n} \subseteq J_{n+1} \subseteq J$ be open bounded intervals with $\bigcup_{n\in\mathbb{N}} J_n = J$. There are functions $\phi_n \in C_c(J)$ with $0 \le \phi_n \le 1$, $\phi_n = 1$ on J_n and $\operatorname{supp}\phi_n \subseteq J_{n+1}$. Lebesgue's theorem gives an index N such that $\psi = \phi_N h \in L^q(J, W_r^m(\mathbb{R}^d)) \cap E$, $\|\psi - h\|_E \le \varepsilon$ and ψ has compact support in J.

4) Finally, we apply the mollifier $G_n^1\psi = \rho_n^1 * \widetilde{\psi}$ from step 2) with $d = 1$, where $\widetilde{\psi}$ is the 0-extension of ψ to $\mathbb{R}$ and we restrict $G_n^1\psi$ to J. It is straight forward to check that $G_n^1\psi \in C_c^\infty(J, W_r^m(\mathbb{R}^d)) \cap E$. The usual properties of mollifiers also work in the Banach space-valued case and we obtain an index N and $\varphi = \rho_N^1 * \psi \in C_c^\infty(J, W_r^m(\mathbb{R}^d))$ such that $\|\psi - \varphi\|_E \le \varepsilon$. □

Strichartz' estimates involve the spaces $L^q(J, W_p^k(\mathbb{R}^d))$ for certain pairs of *admissible* exponents (q,p). This means that

$$\frac{2}{q} + \frac{d}{p} = \frac{d}{2}, \quad \text{where } 2 \le q \le \infty \text{ and } \begin{cases} 1 \le p < \infty, & d \ge 2, \\ 1 \le p \le \infty, & d = 1. \end{cases} \tag{11.13}$$

We see in Remark 11.7 that this relation between p and q is determined by scaling properties of the linear Schrödinger equation. For other dispersive equations (e.g. the wave equation) one obtains a different concept of admissibility (see e.g. Section 2.3 in [**Tao06**]).

We can visualize the reciprocals $(\frac{1}{q}, \frac{1}{p})$ of an admissible pair as the line segment between the reciprocals of $(\infty, 2)$ and $(2, \frac{2d}{d-2})$ if $d \ge 2$, where $(2,\infty) = (2, \frac{2d}{d-2})$ is excluded for $d = 2$. In the case $d = 1$ one has the line between the inverses of $(\infty, 2)$ and $(4, \infty)$.

We remark that for admissible (q,p) we have $H^1(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$ and thus $L^{p'}(\mathbb{R}^d) \hookrightarrow H^{-1}(\mathbb{R}^d)$, since $1 - \frac{d}{2} \ge \frac{2}{q} - \frac{d}{2} = -\frac{d}{p}$, see Sobolev's embedding (5.5). As a result, for admissible (q,p) and $f \in L^{q'}(J, L^{p'}(\mathbb{R}^d))$ the one-sided convolution

$$T *_+ f(t) = \int_0^t T(t-s)f(s)\,\mathrm{d}s, \quad t \in J,$$

is defined as a Bochner integral in $H^{-1}(\mathbb{R}^d)$. Recall that the free Schrödinger group $T(\cdot)$ is strongly continuous on $H^{-1}(\mathbb{R}^d)$ by Lemma 10.7.

We now state our main theorem. Part (a) is called the homogeneous and (b) the inhomogeneous Strichartz' estimate.

THEOREM 11.6 (Strichartz' estimates for iΔ). *Let (q,p) and $(\overline{q},\overline{p})$ be admissible, $k \in \mathbb{N}_0$, $\varphi \in H^k(\mathbb{R}^d)$, $J \subseteq \mathbb{R}$ be an interval with $0 \in J$, and $f \in L^{\overline{q}'}(J, W^k_{\overline{p}'}(\mathbb{R}^d))$. Then $T *_+ f(t)$ exists in $W^k_p(\mathbb{R}^d)$ for a.e. $t \in J$, $T(\cdot)\varphi$ and $T *_+ f$ belong to $L^q(J, W^k_p(\mathbb{R}^d))$ and*

(a) $\|T(\cdot)\varphi\|_{L^q(J,W^k_p(\mathbb{R}^d))} \le c\,\|\varphi\|_{k,2}\,,$

(b) $\|T *_+ f\|_{L^q(J,W^k_p(\mathbb{R}^d))} \le c\,\|f\|_{L^{\overline{q}'}(J,W^k_{\overline{p}'}(\mathbb{R}^d))}$

for a constant $c > 0$ (independent of φ, f and J). If $q = \infty$ and $p = 2$, we can replace L^∞ by C_b in (a) and (b).

Compared to the L^2–setting, in the above estimates one gains space integrability from $p = 2$ to $p > 2$, but one loses time integrability from $q = \infty$ to $q < \infty$. Similarly, in (b) the exponents on the right–hand side are smaller than 2, whereas they are larger than 2 on the left–hand side. We point out that $(\overline{q},\overline{p})$ can be chosen independently of (q,p) in Theorem 11.6 (b).

We will prove Theorem 11.6 only for $q,\overline{q} > 2$ and either for $(q,p) = (\overline{q},\overline{p})$ or for $(q,p) = (\infty,2)$ and any admissible $(\overline{q},\overline{p})$, since we only work with these cases later on. Exponents $(q,p) \neq (\overline{q},\overline{p})$ are needed for certain more general nonlinearities in (11.1). This case requires another tool, the Christ-Kiselev lemma, see Section 2.3 in [**Tao06**]. The endpoint case $(2, \frac{2d}{d-2})$ for $d \ge 3$ is much more difficult, see [**KT98**]. It is needed to study (11.1) for the critical case $\alpha = \alpha_c$.

PROOF OF 11.6 (FOR $q,\overline{q} > 2$ AND EITHER $(q,p) = (\overline{q},\overline{p})$ OR $(q,p) = (\infty,2)$). We first consider $J = \mathbb{R}$ and $k = 0$. The other cases and the final assertion are treated afterwards.

1) Let $(q,p) = (\overline{q},\overline{p})$ be admissible, $2 < q < \infty$, $\varphi \in L^2(\mathbb{R}^d)$ and $f \in L^{q'}(\mathbb{R}, L^{p'}(\mathbb{R}^d)) =: E'$. Set $E = L^q(\mathbb{R}, L^p(\mathbb{R}^d))$. We first prove (b). Inequality (11.3) and Theorem 11.2 (with $\lambda = \frac{d}{2} - \frac{d}{p}$, $n = 1$, $\beta = \gamma = q'$) imply the crucial estimate

$$I_1 := \Big[\int_{\mathbb{R}} \Big[\int_0^t \|T(t-s)f(s)\|_p \,\mathrm{d}s\Big]^q \mathrm{d}t\Big]^{\frac{1}{q}} \le \Big[\int_{\mathbb{R}} \Big[\int_{\mathbb{R}} (4\pi|t-s|)^{\frac{d}{p}-\frac{d}{2}} \|f(s)\|_{p'} \,\mathrm{d}s\Big]^q \mathrm{d}t\Big]^{\frac{1}{q}}$$
$$\le C_0 \Big[\int_{\mathbb{R}} \|f(s)\|_{p'}^{q'} \,\mathrm{d}s\Big]^{\frac{1}{q'}} \qquad (11.14)$$

where C_0 only depends on d, p and q. The conditions of Theorem 11.2 hold since (q,p) is admissible and $2 < q < \infty$. (The measurability of the integrand of I_1 is verified below.)

From this estimate assertion (b) will follow by means of Fubini's theorem, but the details concerning integrability are a bit tricky. To this aim, take $m \in \mathbb{N}$ with $m \ge \frac{d}{2} - \frac{d}{p}$ so that $H^m(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$ by Sobolev's embedding (5.5). Lemma 11.5 yields $g_n \in C_c(\mathbb{R}, H^m(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d))$ that converge to f in E'.

(Observe that $H^m(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d)$ is a Banach space when endowed with the norm given by $\|v\|_{m,2} + \|v\|_{p'}$.) The function

$$\mathbb{R}^2 \to L^p(\mathbb{R}^d), \quad (t,s) \mapsto T(t-s)g_n(s),$$

is continuous for each $n \in \mathbb{N}$ since it is continuous in $H^m(\mathbb{R}^d)$ by Lemma 11.3. There is a subsequence such that the functions $g_{n_j}(s)$ converge in $L^{p'}(\mathbb{R}^d)$ to $f(s)$ as $j \to \infty$ for a.e. $s \in \mathbb{R}$. Moreover, $T(t-s)$ maps $L^{p'}(\mathbb{R}^d)$ continuously into $L^p(\mathbb{R}^d)$ for $t \neq s$, see (11.3). Therefore $(t,s) \mapsto T(t-s)f(s)$ is strongly measurable with values in $L^p(\mathbb{R}^d)$, outside a set of measure 0. In particular the integral I_1 is defined. Similarly, one sees that $T(\cdot)\varphi : \mathbb{R} \to L^p(\mathbb{R}^d)$ is strongly measurable if $\varphi \in L^2(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d)$.

It now follows from Fubini's theorem and (11.14) that the integral $(T *_+ f)(t)$ exists in $L^p(\mathbb{R}^d)$ for a.e. $t \in \mathbb{R}$ and that $T *_+ f : \mathbb{R} \to L^p(\mathbb{R}^d)$ is strongly measurable. Since $\|T *_+ f\|_E \le I_1$, assertion (b) holds. In the same way one derives $\|T * f\|_E \le C_0 \|f\|_{E'}$ for the usual convolution.

2) We show (a) by a duality argument in the framework of step 1). We first consider $g \in C_c(\mathbb{R}, L^2(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d))$. Remark 11.4 and step 1) imply

$$I_2 := \int_{\mathbb{R}} \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}} T(t-s)g(s)\,\mathrm{d}s \Big) \overline{g(t)}\,\mathrm{d}x\,\mathrm{d}t = \langle T * g, \overline{g} \rangle_{L^q(\mathbb{R}, L^p)},$$
$$|I_2| \le \|T * g\|_{L^q(\mathbb{R},L^p)} \, \|g\|_{L^{q'}(\mathbb{R},L^{p'})} \le C_0 \, \|g\|^2_{L^{q'}(\mathbb{R},L^{p'})}.$$

From Fubini's theorem we further deduce

$$\begin{aligned} I_2 &= \int_{\mathbb{R}} \int_{\mathbb{R}} (T(t-s)g(s)|g(t))_{L^2}\,\mathrm{d}s\,\mathrm{d}t = \int_{\mathbb{R}} \int_{\mathbb{R}} (T(-s)g(s)|T(-t)g(t))_{L^2}\,\mathrm{d}s\,\mathrm{d}t \\ &= \Big(\int_{\mathbb{R}} T(-s)g(s)\,\mathrm{d}s \Big| \int_{\mathbb{R}} T(-s)g(s)\,\mathrm{d}s \Big)_{L^2} = \Big\| \int_{\mathbb{R}} T(-s)g(s)\,\mathrm{d}s \Big\|_2^2, \end{aligned}$$

where all integrals are $\mathbb{C}$– or L^2–valued Riemann integrals. We have thus shown

$$\Big\| \int_{\mathbb{R}} T(-t)g(t)\,\mathrm{d}t \Big\|_2 \le \sqrt{C_0}\, \|g\|_{L^{q'}(\mathbb{R},L^{p'})} \tag{11.15}$$

for $g \in C_c(\mathbb{R}, L^2(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d))$. Let $\varphi \in C_c(\mathbb{R}^d)$. Observe that the scalar function $\langle T(\cdot)\varphi, \overline{g} \rangle_{L^p} = (T(\cdot)\varphi|g)_{L^2}$ is measurable. Estimate (11.15) leads to

$$\begin{aligned} \Big| \int_{\mathbb{R}} \langle T(t)\varphi, \overline{g(t)} \rangle_{L^p}\,\mathrm{d}t \Big| &= \Big| \int_{\mathbb{R}} (T(t)\varphi|g(t))_{L^2}\,\mathrm{d}t \Big| = \Big| \int_{\mathbb{R}} (\varphi|T(-t)g(t))_{L^2}\,\mathrm{d}t \Big| \\ &= \Big| \Big(\varphi \Big| \int_{\mathbb{R}} T(-t)g(t)\,\mathrm{d}t \Big)_{L^2} \Big| \le \sqrt{C_0}\, \|\varphi\|_2 \, \|g\|_{L^{q'}(\mathbb{R},L^{p'})}. \end{aligned}$$

Since $C_c(\mathbb{R}, L^2(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d))$ is dense in $L^{q'}(\mathbb{R}, L^{p'}(\mathbb{R}^d))$ by Lemma 11.5, Remark 11.4 yields that $T(\cdot)\varphi \in L^q(\mathbb{R}, L^p(\mathbb{R}^d))$ and

$$\|T(\cdot)\varphi\|_{L^q(\mathbb{R},L^p)} = \sup_{\|g\|_{E'} \le 1} |\langle T(\cdot)\varphi, \overline{g} \rangle_{L^q(\mathbb{R},L^p)}| \le \sqrt{C_0}\, \|\varphi\|_2.$$

By approximation, we derive (a) for $\varphi \in L^2(\mathbb{R}^d)$ and $q \in (2,\infty)$.

3) Let $k = 0$, $(q,p) = (\infty, 2)$ and $(\overline{q}, \overline{p})$ be admissible with $\overline{q} > 2$. Then (a) holds with C_b instead of L^∞ since $T(\cdot)$ is a unitary C_0-group on $L^2(\mathbb{R}^d)$. To prove (b), we set $f_t = \mathbb{1}_{[0,t]} f$ for $t > 0$ and $f_t = \mathbb{1}_{[t,0]} f$ for $t < 0$. We write (q,p)

instead of $(\overline{q},\overline{p})$. First let $f \in C_c(\mathbb{R}, L^2(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d))$. Using also (11.15), we obtain $T *_+ f \in C_b(\mathbb{R}, L^2(\mathbb{R}^d))$ and

$$\|T *_+ f\|_{C_b(\mathbb{R},L^2)} = \sup_{t\in\mathbb{R}} \Big\| \int_0^t T(t)T(-s)f_t(s)\,\mathrm{d}s \Big\|_2 = \sup_{t\in\mathbb{R}} \Big\| \int_{\mathbb{R}} T(-s)f_t(s)\,\mathrm{d}s \Big\|_2$$
$$\leq \sup_{t\in\mathbb{R}} \sqrt{C_0}\, \|f_t\|_{L^{q'}(\mathbb{R},L^{p'})} = \sqrt{C_0}\, \|f\|_{L^{q'}(\mathbb{R},L^{p'})} .$$

If we approximate the given f in $L^{q'}(\mathbb{R}, L^{p'}(\mathbb{R}^d))$ by $f_n \in C_c(\mathbb{R}, L^2(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d))$, then the above estimate shows that $(T *_+ f_n)_n$ converges to a function u in $C_b(\mathbb{R}, L^2(\mathbb{R}^d))$. On the other hand, step 1) implies that for a subsequence the functions $(T *_+ f_{n_j})(t)$ tend to $(T *_+ f)(t)$ in $L^p(\mathbb{R}^d)$ for a.e. $t \in \mathbb{R}$. Hence, $T *_+ f = u$ belongs to $C_b(\mathbb{R}, L^2(\mathbb{R}^d))$ and (b) is true in the present case.

4) Let $k = 1$. By Lemma 11.3 the spatial derivatives ∂_j $(j = 1, \ldots, d)$ commute with $T(t)$ on $H^1(\mathbb{R}^d)$. This fact easily implies that (a) holds with $k = 1$ and that $T(\cdot)\varphi \in C_b(\mathbb{R}, H^1(\mathbb{R}^d))$ if $\varphi \in H^1(\mathbb{R}^d)$.

For (b), take $f \in L^{q'}(\mathbb{R}, W^1_{p'}(\mathbb{R}^d))$. As in step 1) we approximate f in $L^{q'}(\mathbb{R}, W^1_{p'}(\mathbb{R}^d))$ by $g_n \in C_c(\mathbb{R}, H^m(\mathbb{R}^d) \cap L^{p'}(\mathbb{R}^d))$, where $m > 1 + \frac{d}{2} - \frac{d}{p}$ so that $H^m(\mathbb{R}^d) \hookrightarrow W^1_p(\mathbb{R}^d)$ by Corollary D.16. Step 1) shows that the functions $\partial_j T *_+ g_n = T *_+ \partial_j g_n$ tend to $T *_+ \partial_j f$ in $L^q(\mathbb{R}, L^p(\mathbb{R}^d))$ as $n \to \infty$. As a result, $T *_+ f$ belongs to $L^q(\mathbb{R}, W^1_p(\mathbb{R}^d))$ and (b) holds for $k = 1$. The case $k \geq 2$ is treated similarly.

Let $J \subseteq \mathbb{R}$ be an interval with $0 \in J$. Part (a) for J then follows from the assertion for $J = \mathbb{R}$. To derive (b), we extend $f \in L^q(J, W^k_p(\mathbb{R}^d))$ by 0 to $\widetilde{f} \in L^q(\mathbb{R}, W^k_p(\mathbb{R}^d))$. Then $T*_+ f = T*_+ \widetilde{f}$ on J and so (b) is also true for J. □

Part (a) of Theorem 11.6 is wrong for non–admissible (q, p) as we see in the next remark, whereas part (b) is true for some non–admissible exponents, see §2.4 of [**Caz03**] and Exercise 11.3. Strichartz' estimates fail for $(2, \infty)$ if $d = 2$, see [**MS98**].

Remark 11.7. A scaling argument shows that Theorem 11.6 (a) can only be valid for admissible exponents. In fact, let $\varphi \in H^2(\mathbb{R}^d) \setminus \{0\}$ and $u = T(\cdot)\varphi$. For $\lambda > 0$, we define $\varphi_\lambda(x) = \varphi(\lambda x)$, $x \in \mathbb{R}^d$. Clearly $\varphi_\lambda \in H^2(\mathbb{R}^d)$ and the solution $u_\lambda = T(\cdot)\varphi_\lambda$ of $\mathrm{i}u' = -\Delta u$ with initial value φ_λ is given by $(u_\lambda(t))(x) = u(\lambda^2 t, \lambda x)$. Let $E = L^q(\mathbb{R}, L^p(\mathbb{R}^d))$ for some $1 \leq p, q < \infty$. Observe that

$$\|\varphi_\lambda\|_2 = \Big(\int_{\mathbb{R}^d} |\varphi(\lambda x)|^2\,\mathrm{d}x\Big)^{\frac{1}{2}} = \Big(\int_{\mathbb{R}^d} \lambda^{-d}|\varphi(y)|^2\,\mathrm{d}y\Big)^{\frac{1}{2}} = \lambda^{-\frac{d}{2}}\|\varphi\|_2\,,$$
$$\|u_\lambda\|_E = \Big(\int_{\mathbb{R}}\Big(\int_{\mathbb{R}^d} |u(\lambda^2 t, \lambda x)|^p\,\mathrm{d}x\Big)^{\frac{q}{p}}\mathrm{d}t\Big)^{\frac{1}{q}}$$
$$= \Big(\int_{\mathbb{R}}\Big(\int_{\mathbb{R}^d} |u(s,y)|^p \lambda^{-d}\,\mathrm{d}y\Big)^{\frac{q}{p}}\lambda^{-2}\,\mathrm{d}s\Big)^{\frac{1}{q}} = \lambda^{-\frac{d}{p}}\lambda^{-\frac{2}{q}}\|u\|_E.$$

Suppose that Theorem 11.6 (a) holds for (q, p). Then

$$\lambda^{-\frac{d}{p}-\frac{2}{q}}\|u\|_E = \|u_\lambda\|_E \leq c\|\varphi_\lambda\|_2 = \lambda^{-\frac{d}{2}}c\|\varphi\|_2$$

for a constant $c > 0$ and all $\lambda > 0$. Hence, $\lambda^{\frac{d}{2}-\frac{d}{p}-\frac{2}{q}}$ is uniformly bounded for $\lambda > 0$ and thus $\frac{d}{2} = \frac{d}{p} + \frac{2}{q}$. ♢

Exercises

EXERCISE 11.1 (Scaling). Let $\varphi \in H^2(\mathbb{R}^d)$ and let u be an H^2–solution of (11.1) on J with $u(0) = \varphi$. For $\lambda > 0$ and $\kappa \in \mathbb{R}$ define the function $(u_\lambda(t))(x) = \lambda^\kappa (u(\lambda^2 t))(\lambda x)$, $t \in J$, $x \in \mathbb{R}^d$. For which exponent $\kappa = \kappa(\alpha)$ does the map u_λ solve (11.1) for the initial value φ_λ given by $\varphi_\lambda(x) = \lambda^\kappa \varphi(\lambda x)$? Fix this exponent $\kappa(\alpha) = \overline{\kappa}$ in the definition of u_λ. For which $\alpha > 1$ we then have $\|u_\lambda(t)\|_2 = \|u(\lambda^2 t)\|_2$ or $\|\partial_k u_\lambda(t)\|_2 = \|\partial_k u(\lambda^2 t)\|_2$ for all $t \in J$ and $k \in \{1, \ldots, d\}$?

EXERCISE 11.2 (Symmetries). Let u be an H^2–solution of (11.1) on $J = \mathbb{R}$ (see Definition 10.8). For $h \in \mathbb{R}^d$, $Q \in \mathbb{R}^{d\times d}$ and $\varphi : \mathbb{R}^d \to \mathbb{C}$ we set $(S_h\varphi)(x) = \varphi(x - h)$ and $(R_Q\varphi)(x) = \varphi(Qx)$, $x \in \mathbb{R}^d$. We define

$$
\begin{aligned}
w_1(t) &= \mathrm{e}^{\mathrm{i}\theta} S_{x_0} u(t - t_0) && \text{for fixed } t_0 \in \mathbb{R}, x_0 \in \mathbb{R}^d \text{ and } \theta \in \mathbb{R},\\
w_2(t) &= \overline{u(-t)},\\
w_3(t) &= R_Q u(t) && \text{for a fixed orthogonal } Q \in \mathbb{R}^{d\times d},\\
w_4(t) &= \mathrm{e}_{\mathrm{i}v} \mathrm{e}^{-\mathrm{i}|v|^2 t} S_{2vt} u(t) && \text{for a fixed } v \in \mathbb{R}^d,
\end{aligned}
$$

where $t \in \mathbb{R}$ and $\mathrm{e}_{\mathrm{i}v}(x) = \mathrm{e}^{\mathrm{i}v\cdot x}$. Show that the functions w_j $(j = 1, 2, 3, 4)$ satisfy (11.1) for the appropriate initial values. Further show that $u(t)$ is spherically symmetric for all $t \in \mathbb{R}$ if $u(0)$ is spherically symmetric.

EXERCISE 11.3. Let $p \in (2, \frac{2d}{d-2})$ and $r, s \in (1, \infty)$ such that

$$\frac{1}{r} + \frac{1}{s} = \frac{d}{2} - \frac{d}{p}.$$

Let $f \in L^{s'}(\mathbb{R}, L^{p'}(\mathbb{R}^d))$. Show that $T *_+ f \in L^r(\mathbb{R}, L^p(\mathbb{R}^d))$ and $\|T *_+ f\|_{L^r(\mathbb{R},L^p)} \le c\|f\|_{L^{s'}(\mathbb{R},L^{p'})}$ for a constant $c \ge 0$.

LECTURE 12

Local wellposedness of the nonlinear Schrödinger equation

In this lecture we establish the local wellposedness theory for the nonlinear Schrödinger equation

$$u'(t) = \mathrm{i}\Delta u(t) - \mathrm{i}\mu|u(t)|^{\alpha-1}u(t), \quad t \in J, \qquad u(0) = u_0, \tag{12.1}$$

where $u_0 \in H^1(\mathbb{R}^d)$, $\mu \in \{-1, 1\}$ and $\alpha \in (1, \alpha_c)$ with $\alpha_c = \frac{d+2}{(d-2)_+}$. We will follow the strategy of the proofs in Lecture 8. This approach will be combined with the properties of the nonlinearity

$$F(u) = -\mathrm{i}\mu|u|^{\alpha-1}u$$

shown in Lecture 9 and with Strichartz' estimates for the free Schrödinger group $T(\cdot)$ generated by $\mathrm{i}\Delta$ that we derived in Theorem 11.6. In particular, we prove existence and uniqueness of a maximal H^1–solution $u \in C(J(u_0), H^1(\mathbb{R}^d)) \cap C^1(J(u_0), H^{-1}(\mathbb{R}^d))$ of (12.1), a blow-up condition in terms of the norm of $H^1(\mathbb{R}^d)$ and the continuous dependence on initial data.

Next week we will show that for an initial value $u_0 \in H^2(\mathbb{R}^d)$ the H^1–solution u of (12.1) is in fact an H^2–solution on its full existence interval $J(u_0)$. At the end of this lecture, we discuss a few results on Banach space valued Sobolev spaces which are needed for this regularity theorem.

As in Lecture 8 we reformulate (12.1) as the integrated fixed point problem

$$u(t) = T(t)u_0 + \int_0^t T(t-s)F(u(s))\,\mathrm{d}s =: \Phi_{u_0}(u)(t) = \Phi u(t), \quad t \in J, \tag{12.2}$$

for $u_0 \in H^1(\mathbb{R}^d)$ and $u \in C(J, H^1(\mathbb{R}^d))$. In contrast to Lecture 8, now F does not map from, say, $L^2(\mathbb{R}^d)$ into itself. As stated in Remark 10.6, this nonlinearity is real continuously differentiable from $H^1(\mathbb{R}^d)$ to $H^{-1}(\mathbb{R}^d)$ and from $L^p(\mathbb{R}^d)$ to $L^{p'}(\mathbb{R}^d)$. Here and throughout, we set

$$p = 1 + \alpha \in (2, 1 + \alpha_c), \qquad p' = \frac{1+\alpha}{\alpha}, \qquad \frac{2}{q} = \frac{d}{2} - \frac{d}{p}, \tag{12.3}$$

so that $q \in (2, \infty)$ and (q, p) is admissible in the sense of (11.13). Sobolev's embedding (5.5) yields that

$$\|v\|_p \le C_{\mathrm{So}}\|v\|_{1,2} \tag{12.4}$$

for all $v \in H^1(\mathbb{R}^d)$ and the Sobolev constant C_{So}. A uniform bound on the H^1–norm of $u(s)$ in (12.2) thus allows to control the nonlinearity $F : L^p(\mathbb{R}^d) \to$

$L^{p'}(\mathbb{R}^d)$. But then $F(u(s))$ only belongs to $L^{p'}(\mathbb{R}^d)$. Here Strichartz' estimates from Theorem 11.6 come into play. They yield

$$|||T(\cdot)\varphi|||_{k,J} \leq C_{\mathrm{St}} \|\varphi\|_{k,2}, \tag{12.5}$$
$$|||T *_+ f|||_{k,J} \leq C_{\mathrm{St}} \|f\|_{L^{q'}(J, W^k_{p'}(\mathbb{R}^d))}, \tag{12.6}$$

for every $\varphi \in H^k(\mathbb{R}^d)$, $f \in L^{q'}(J, W^k_{p'}(\mathbb{R}^d))$, $k \in \{0,1\}$, every interval $J \subseteq \mathbb{R}$ with $0 \in J$ and a constant $C_{\mathrm{St}} \geq 1$ (independent of φ, f, J). Moreover, the functions $T(\cdot)\varphi$ and $T *_+ f$ are continuous with values in $H^k(\mathbb{R}^d)$, see Theorem 11.6. Here and throughout, we set

$$E_k(J) = L^q(J, W^k_p(\mathbb{R}^d)), \quad E'_k(J) = L^{q'}(J, W^k_{p'}(\mathbb{R}^d)), \quad G_k(J) = L^\infty(J, H^k(\mathbb{R}^d))$$
$$\mathcal{E}_k(J) = L^\infty(J, H^k(\mathbb{R}^d)) \cap L^q(J, W^k_p(\mathbb{R}^d)) = G_k(J) \cap E_k(J),$$
$$|||u|||_{k,J} = \max\{\|u\|_{G_k(J)}, \|u\|_{E_k(J)}\} \quad \text{for } u \in \mathcal{E}_k(J),$$

for $k \in \mathbb{N}_0$. If $J = [-b, b]$ we replace J by b in the above notation. We point out that in (12.6) both parts of the $\mathcal{E}_k(J)$–norm are controlled by the norm in E'_k (where we note that E'_k is not the dual of E_k if $k \geq 1$). We next consider the mapping properties of F on these spaces.

Lemma 12.1. *Let $J = [-b, b]$, $u, v \in G_1(b) \hookrightarrow E_0(b)$, $w \in \mathcal{E}_1(b)$, $\varphi, \psi \in L^p(\mathbb{R}^d)$ and $\chi \in W^1_p(\mathbb{R}^d)$. Set $r = \operatorname{ess\,sup}\{\|u(t)\|_{1,2}, \|v(t)\|_{1,2}, \|w(t)\|_{1,2} \,|\, t \in J\}$. Then $F(\varphi) \in L^{p'}(\mathbb{R}^d)$, $F(\chi) \in W^1_{p'}(\mathbb{R}^d)$, $F(u) \in E'_0(J)$, $F(w) \in E'_1(J)$ and the inequalities*

$$\|F(\varphi) - F(\psi)\|_{p'} \leq c\big(\|\varphi\|_p^{\alpha-1} + \|\psi\|_p^{\alpha-1}\big)\|\varphi - \psi\|_p, \tag{12.7}$$
$$\|F(u) - F(v)\|_{L^{q'}(J, L^{p'})} \leq c r^{\alpha-1} b^{\frac{1}{q'} - \frac{1}{q}} \|u - v\|_{L^q(J, L^p)}, \tag{12.8}$$
$$\|\nabla F(\chi)\|_{p'} \leq c\|\chi\|_p^{\alpha-1} \||\nabla\chi|\|_p, \tag{12.9}$$
$$\|\nabla F(w)\|_{L^{q'}(J, L^{p'})} \leq c r^{\alpha-1} b^{\frac{1}{q'} - \frac{1}{q}} \||\nabla w|\|_{L^q(J, L^p)}, \tag{12.10}$$
$$\|F(w)\|_{L^{q'}(J, W^1_{p'})} \leq c r^{\alpha-1} b^{\frac{1}{q'} - \frac{1}{q}} \|w\|_{L^q(J, W^1_p)} \tag{12.11}$$

hold for a constant c only depending on $\alpha \in (1, \alpha_c)$ and d.

Proof. Since $p > \alpha$ and $p' = \frac{p}{\alpha}$, Lemma 9.2 and Corollary 9.3 yield

$$\begin{aligned}
\|F(\varphi) - F(\psi)\|_{p'} &= \Big\| \int_0^1 F'(\varphi + \tau(\psi - \varphi))(\psi - \varphi)\,\mathrm{d}\tau \Big\|_{\frac{p}{\alpha}} \\
&\leq \alpha \sup_{\tau \in [0,1]} \|(1-\tau)\varphi + \tau\psi\|_p^{\alpha-1} \|\varphi - \psi\|_p \\
&\leq c_1\big(\|\varphi\|_p^{\alpha-1} + \|\psi\|_p^{\alpha-1}\big)\|\varphi - \psi\|_p.
\end{aligned}$$

Here and below c_j denote constants which only depend on α and d. In this estimate we insert $u(t)$ and $v(t)$ and take the q–norm in time. Using Sobolev's embedding (12.4) and $\|u(t)\|_{1,2}, \|v(t)\|_{1,2} \leq r$, we arrive at

$$\|F(u) - F(v)\|_{L^q(J, L^{p'})} \leq 2c_1 C_{\mathrm{So}}^{\alpha-1} r^{\alpha-1} \|u - v\|_{L^q(J, L^p)}. \tag{12.12}$$

Since $q' < 2 < q$, Hölder's inequality then yields

$$\|F(u) - F(v)\|_{L^{q'}(J,L^{p'})} \le c_2 r^{\alpha-1} b^{\frac{1}{q'}-\frac{1}{q}} \|u - v\|_{L^q(J,L^p)}.$$

It remains to check (12.9) – (12.11). Let $\phi(z) = -\mathrm{i}\mu|z|^{\alpha-1}z$ for $z \in \mathbb{R}^2$ and $j \in \{1, \dots, d\}$. As in Lemma 9.2 and Corollary 9.3, we derive

$$\|\partial_j F(\chi)\|_{p'} = \|\phi'(\chi)\partial_j\chi\|_{p'} \le \alpha\||\chi|^{\alpha-1}\partial_j\chi\|_{p'} \le \alpha\|\chi\|_p^{\alpha-1}\|\partial_j\chi\|_p \tag{12.13}$$

from Hölder's inequality with exponents $\frac{1}{p'} = \frac{1}{p} + \frac{\alpha-1}{p}$. The estimates (12.9) – (12.11) then follow as above (using also (12.8) with $v = 0$ for (12.11)). □

If we combine the above result with Strichartz' estimates (12.5) and (12.6), we see that Φ from (12.2) maps $\mathcal{E}_1(b)$ into itself and that it is Lipschitz on bounded sets of $\mathcal{E}_1(b)$, but only with respect to the metric of $\mathcal{E}_0(b)$. Fortunately, these properties still allow to apply Banach's fixed point theorem, thanks to the next lemma. It relies on the Banach-Alaoglu theorem.

Lemma 12.2. *Let $r > 0$ and $J \subseteq \mathbb{R}$ be an interval. Then the ball $\Sigma(r, J) = \{u \in \mathcal{E}_1(J) \mid \|\!|u|\!\|_{1,J} \le r\}$ is a complete metric space when endowed with the metric induced by $\|\!|\cdot|\!\|_{0,J}$.*

Proof. Let $(u_n)_n$ be a Cauchy sequence in $\Sigma(r, J)$ for $\|\!|\cdot|\!\|_{0,J}$. Since $\mathcal{E}_0(b)$ is a Banach space, $(u_n)_n$ converges in $\mathcal{E}_0(J)$ to a function $u \in \mathcal{E}_0(J)$ as $n \to \infty$. We have to show that $u \in \Sigma(r, J)$. As indicated in Remark 11.4 the space $L^\infty(J, H^1(\mathbb{R}^d)) = G_1(J)$ is the dual of $L^1(J, H^{-1}(\mathbb{R}^d))$ and $L^q(J, W^1_p(\mathbb{R}^d)) = E_1(J)$ is reflexive with dual $L^{q'}(J, W^{-1}_{p'}(\mathbb{R}^d))$. Observe that $H^{-1}(\mathbb{R}^d)$ is separable since its dual $H^1(\mathbb{R}^d)$ is separable. Remark 11.4 thus yields the separability of $L^1(J, H^{-1}(\mathbb{R}^d))$. The Banach-Alaoglu theorem then gives a subsequence $(u_{n_j})_j$ which converges weakly in $E_1(J)$ to some v with $\|v\|_{E_1(J)} \le r$ and $(u_{n_j})_j$ converges weak* in $G_1(J)$ to some w with $\|w\|_{G_1(J)} \le r$ as $j \to \infty$. Since $E_0(J)^* \hookrightarrow E_1(J)^*$ and $L^1(J, L^2(\mathbb{R}^d)) \hookrightarrow L^1(J, H^{-1}(\mathbb{R}^d))$, the functions u_{n_j} also tend weakly in $E_0(J)$ to v and weak* in $G_0(J)$ to w. On the other hand, $(u_{n_j})_j$ converges to u in $E_0(J)$ and $G_0(J)$ so that $u = v$ and $u = w$ by the uniqueness of weak and weak* limits, i.e., $u \in \Sigma(r, J)$. □

We further note how one can concatenate H^k–solutions.

Lemma 12.3. *Let $k \in \{1, 2\}$ and let u and v be H^k–solutions of (12.1) on $[a, b]$ and $[b, c]$, respectively. Assume that $u(b) = v(b)$. Then the function w given by $w(t) = u(t)$ for $t \in [a, b]$ and $w(t) = v(t)$ for $t \in (b, c]$ is an H^k–solution of (12.1) with $w(a) = u(a)$.*

Proof. It is clear that $w \in C([a, c], H^k(\mathbb{R}^d))$. Because of

$$\begin{aligned}\left(\tfrac{\mathrm{d}}{\mathrm{d}t}\right)^{-} w(b) &= u'(b) = \mathrm{i}\Delta u(b) + F(u(b)) = \mathrm{i}\Delta v(b) + F(v(b)) \\ &= v'(b) = \left(\tfrac{\mathrm{d}}{\mathrm{d}t}\right)^{+} w(b),\end{aligned}$$

it further follows that $w \in C^1([a, c], H^{k-2}(\mathbb{R}^d))$ solves (12.1) on $[a, c]$. □

As a first part of the local wellposedness result, we establish the uniqueness of H^1–solutions.

Lemma 12.4. *Let $u_0 \in H^1(\mathbb{R}^d)$. Let u and v be H^1–solutions of (12.1) on intervals J_u and J_v containing 0, respectively. Then $u = v$ on $J_u \cap J_v$.*

Proof. If the assumption were not true, there would exist $\tau \in J_u \cap J_v$ such that $u = v$ on $[-\tau, \tau]$ and $u(t_n) \neq v(t_n)$ for certain $t_n \in J_u \cap J_v$ with, say, $t_n \to \tau^+$ as $n \to \infty$. The case that $t_n \to (-\tau)^-$ is treated similarly. Take $\delta_0 > 0$ such that $J_0 := [\tau, \tau + \delta_0] \subseteq J_u \cap J_v$. We may assume that all t_n belong to J_0. Let $r = \max\{\|u(t)\|_{1,2}, \|v(t)\|_{1,2} \,|\, t \in J_0\}$. Since u, v are H^1–solutions of (12.1), Remark 10.6 yields $F(u), F(v) \in C(J_0, H^{-1}(\mathbb{R}^d))$. Proposition 6.5 thus implies that

$$
\begin{aligned}
u(t+\tau) &= T(t)u(\tau) + \int_0^t T(t-s)F(u(s+\tau))\,\mathrm{d}s,\\
v(t+\tau) &= T(t)u(\tau) + \int_0^t T(t-s)F(v(s+\tau))\,\mathrm{d}s
\end{aligned}
$$

for all $t \in [0, \delta_0]$, see (10.10). Take $J := [\tau, \tau + \delta] \subseteq J_0$. Strichartz' inequality (12.6) and the estimate (12.8) then yield

$$
|||u - v|||_{\mathcal{E}_0(J)} \le cC_{\mathrm{St}} r^{\alpha-1} \delta^{\frac{1}{q'} - \frac{1}{q}} |||u - v|||_{\mathcal{E}_0(J)} \,.
$$

Choosing a sufficiently small $\delta \in (0, \delta_0)$, we deduce that $u = v$ on $[\tau, \tau + \delta]$ which contradicts $t_n \to \tau^+$. □

As in Lecture 8 we state the core existence result as a separate lemma.

Lemma 12.5. *Let $\rho > 0$. Then there is a number $b_0(\rho) > 0$ (see (12.16) below) such that for each $u_0 \in \overline{B}_{H^1}(0, \rho)$ there is a unique H^1–solution $u \in \mathcal{E}_1(b_0(\rho))$ of (12.1) on the time interval $[-b_0(\rho), b_0(\rho)] =: J_0$. Moreover, $|||u|||_{1,b} \le r := 1 + C_{\mathrm{St}}\rho$, where $C_{\mathrm{St}} \ge 1$ is taken from (12.5) and (12.6). It further holds $u = \Phi_{u_0} u$ on J_0, cf. (12.2).*

Proof. Let $\rho > 0$ and take $u_0 \in H^1(\mathbb{R}^d)$ with $\|u_0\|_{1,2} \le \rho$. Take $b > 0$ to be specified below. Fix $r = 1 + C_{\mathrm{St}}\rho$. Lemma 12.2 provides us with the complete metric space $\Sigma(r, b)$ where the metric is given by $|||u - v|||_{0,b}$. Let $\Phi u = \Phi_{u_0} u$ be defined by (12.2) for $u \in \Sigma(r, b)$. Combining (12.5), (12.6), (12.11) and (12.8), we estimate

$$
|||\Phi u|||_{1,b} \le C_{\mathrm{St}}(\|u_0\|_{1,2} + \|F(u)\|_{E_1'(b)}) \le C_{\mathrm{St}}\rho + cC_{\mathrm{St}} r^{\alpha} b^{\frac{1}{q'} - \frac{1}{q}}, \tag{12.14}
$$

$$
|||\Phi u - \Phi v|||_{0,b} \le C_{\mathrm{St}} \|F(u) - F(v)\|_{E_0'(b)} \le cC_{\mathrm{St}} r^{\alpha-1} b^{\frac{1}{q'} - \frac{1}{q}} \|u - v\|_{E_0(b)} \tag{12.15}
$$

for $u, v \in \Sigma(r, b)$, where c is a constant only depending on α and d. We now set

$$
b_0(\rho) := \min\Big\{ (cC_{\mathrm{St}} r^{\alpha})^{\frac{q'q}{q'-q}}, (2cC_{\mathrm{St}} r^{\alpha-1})^{\frac{q'q}{q'-q}} \Big\}. \tag{12.16}
$$

Let $J_0 = [-b_0(\rho), b_0(\rho)]$. It follows that $\Phi u \in \Sigma(r, b_0(\rho))$ and $|||\Phi u - \Phi v|||_{0,J_0} \le \frac{1}{2}|||u - v|||_{0,J_0}$. The contraction mapping principle then yields a unique fixed point $u = \Phi u$ in $\Sigma(r, b_0(\rho))$.

Theorem 11.6 further shows that u belongs to $C(J_0, H^1(\mathbb{R}^d))$, and hence $f := F(u) \in C(J_0, H^{-1}(\mathbb{R}^d))$ by Remark 10.6. Since u is a mild solution of $u' = \mathrm{i}\Delta u + f$ in $H^{-1}(\mathbb{R}^d)$ with $u(0) = u_0$, Lemma 6.8 implies that u is an H^1–solution of (12.1) on J_0. Uniqueness follows from Lemma 12.4. □

Before coming to the local wellposedness theorem, we define the maximal existence times

$$t^+(u_0) = \sup\{b > 0 \,|\, \exists H^1\text{–solution } u \in C([0,b], H^1(\mathbb{R}^d)) \text{ of } (12.1)\},$$
$$t^-(u_0) = \inf\{b < 0 \,|\, \exists H^1\text{–solution } u \in C([b,0], H^1(\mathbb{R}^d)) \text{ of } (12.1)\}.$$

Lemma 12.5 implies that the above sets are non-empty and that $-t^-(u_0), t^+(u_0) \in (0,\infty]$. We set $J(u_0) = (t^-(u_0), t^+(u_0))$ and call an H^1–solution u of (12.1) on $J(u_0)$ *maximal.* (In view of Theorem 12.6 (c), we have to take an open time interval here.)

Our local wellposedness theorem follows the pattern of Theorem 8.6 though the underlying function spaces are adapted to Strichartz' estimates. Moreover, we only show the continuity of the solution map $u_0 \mapsto u(\,\cdot\,; u_0)$ and not its Lipschitz continuity as in Theorem 8.6. One obtains the Lipschitz continuity only under stronger conditions, see Exercise 12.2, or in weaker norms, cf. (12.18).

There is an analogous result for the critical case $\alpha = \frac{d+2}{d-2}$ for $d \geq 3$, see e.g. Theorem 4.5.1 in [**Caz03**]. Here one obtains a much less convenient blow-up condition and one needs the endpoint Strichartz estimates.

THEOREM 12.6. *Let $1 < \alpha < \alpha_c = \frac{d+2}{(d-2)_+}$, $\mu \in \{-1, 1\}$. Let $u_0 \in H^1(\mathbb{R}^d)$ and $b_0(\|u_0\|_{1,2})$ be defined by (12.16). Set $p = 1 + \alpha$ and $\frac{2}{q} = \frac{d}{2} - \frac{d}{p}$. Then the following assertions hold.*

(a) There is a unique maximal H^1–solution $u = u(\,\cdot\,; u_0)$ of (12.1) on $J(u_0) = (t^-(u_0), t^+(u_0))$ where $-t^-(u_0), t^+(u_0) \in (b_0(\|u_0\|_{1,2}), \infty]$.

(b) For every compact interval $J \subseteq J(u_0)$ we have $u \in L^q(J, W^1_p(\mathbb{R}^d))$.

(c) If $t^\pm(u_0) < \infty$, then $\lim_{t \to t^\pm(u_0)} \|u(t)\|_{1,2} = \infty$.

(d) Let $J \subseteq J(u_0)$ be a compact interval with $0 \in J$. Then there is a radius $\delta > 0$ such that for $v_0 \in \overline{B}_{H^1}(u_0, \delta)$ we have $J \subseteq J(v_0)$ and the map

$$\overline{B}_{H^1}(u_0, \delta) \to C(J, H^1) \cap L^q(J, W^1_p), \quad v_0 \mapsto u(\,\cdot\,; v_0),$$

is continuous.

PROOF. (a) Let $u_0 \in H^1(\mathbb{R}^d)$. Lemma 12.5 gives an H^1–solution u of (12.1) on $[-b_0(\|u_0\|_{1,2}), b_0(\|u_0\|_{1,2})]$. Using Lemmas 12.3 and 12.5, we can extend u to a solution on a larger time interval, so that $-t^-(u_0), t^+(u_0) > b_0(\|u_0\|_{1,2})$.

Next, we take $b_n^\pm \to t^\pm(u_0)$ as $n \to \infty$ with corresponding H^1–solutions u_n of (12.1) on $[b_n^-, b_n^+]$. Lemma 12.4 allows to introduce a unique maximal H^1–solution $u = u(\,\cdot\,; u_0)$ on $J(u_0)$ by setting $u(t) = u_n(t)$ for $t \in J(u_0)$ and $n \in \mathbb{N}$ with $t \in [b_n^-, b_n^+]$.

(b) Let $\tau \in J(u_0)$. Applying Lemma 12.5 to the initial value $u(\tau) \in H^1(\mathbb{R}^d)$, we find a time $\beta(\tau) > 0$ and an H^1–solution v of (12.1) with $v(0) = u(\tau)$ belonging to $\mathcal{E}_1(\beta(\tau))$. By the uniqueness Lemma 12.4, v is a restriction of $u(\tau + \cdot\,)$ and thus $u \in L^q([\tau - \beta(\tau), \tau + \beta(\tau)], W^1_p(\mathbb{R}^d))$. A compactness argument then yields assertion (b).

(c) Let $u = u(\,\cdot\,; u_0)$ and $t^+(u_0) < \infty$. Suppose there were $b_n \to t^+(u_0)^-$ such that $\sup_n \|u(b_n)\|_{1,2} =: C < \infty$. Take a time b_n with $b_n + b_0(C) > t^+(u_0)$. Using Lemmas 12.3 and 12.5 we can extend the given H^1–solution to $[0, b_n + b_0(C)]$ by considering (12.1) with initial value $u(b_n)$. This fact contradicts the definition of $t^+(u_0)$. One treats $t^-(u_0)$ in the same way. Hence, (c) holds.

(d) Fix $J = [a, b] \subseteq J(u_0)$ with $0 \in J$. We show that every sequence $(\varphi_n)_n$ converging in $H^1(\mathbb{R}^d)$ to u_0 as $n \to \infty$ has a subsequence $(\varphi_{n_j})_j$ such that $J \subseteq J(\varphi_{n_j})$ for all j and the solutions $u_{n_j} = u(\,\cdot\,; \varphi_{n_j})$ tend to u in $\mathcal{E}_1(J)$ as $j \to \infty$. This fact implies assertion (d) by a straight forward contradiction argument. Observe that for H^1–solutions v one has $|||v|||_{1,J} = \max\{\|v\|_{C(J,H^1)}, \|v\|_{E_1(J)}\}$.

So let $(\varphi_n)_n$ converge to u_0 in $H^1(\mathbb{R}^d)$ and set $u_n = u(\,\cdot\,; \varphi_n)$. Set $\overline{\rho} := 1 + \max_{t\in J} \|u(t)\|_{1,2} < \infty$. There is an index $n_0 \in \mathbb{N}$ such that $\|\varphi_n\|_{1,2} \le \overline{\rho}$ for all $n \ge n_0$. Lemma 12.5 thus yields $-t^-(\varphi_n), t^+(\varphi_n) > b_0(\overline{\rho}) =: b_0$ for all $n \ge n_0$. Lemma 12.4 shows that the restrictions of u and u_n to $J_0 := [-b_0, b_0]$ coincide with the solutions obtained in Lemma 12.5 for the initial value u_0 and φ_n, respectively. From Lemma 12.5 we then deduce that $u = \Phi_{u_0}(u)$ and $u_n = \Phi_{\varphi_n}(u_n)$ on J_0 and that

$$|||u|||_{1,b_0}, |||u_n|||_{1,b_0} \le \overline{r} := 1 + C_{\mathrm{St}}\overline{\rho}. \tag{12.17}$$

Due to (12.15) and the choice of b_0 in (12.16), the operator Φ_{u_0} is Lipschitz on $\Sigma(\overline{r}, b_0) = \overline{B}_{\mathcal{E}_1(b_0)}(0, \overline{r})$ for the metric induced by $|||\cdot|||_{0,b_0}$ with a constant bounded by $\frac{1}{2}$. Combining this fact with Strichartz' estimate (12.5), we conclude

$$\begin{aligned}
|||u - u_n|||_{0,b_0} &\le |||\Phi_{u_0}(u) - \Phi_{u_0}(u_n)|||_{0,b_0} + |||\Phi_{u_0}(u_n) - \Phi_{\varphi_n}(u_n)|||_{0,b_0}\\
&\le \tfrac{1}{2}|||u - u_n|||_{0,b_0} + |||T(\cdot)(u_0 - \varphi_n)|||_{0,b_0}\\
&\le \tfrac{1}{2}|||u - u_n|||_{0,b_0} + C_{\mathrm{St}}\|u_0 - \varphi_n\|_2,\\
|||u - u_n|||_{0,b_0} &\le 2C_{\mathrm{St}}\|u_0 - \varphi_n\|_2 \qquad (12.18)
\end{aligned}$$

for all $n \ge n_0$. Unfortunately, this argument does not give the desired continuity of $v_0 \mapsto u(\,\cdot\,; v_0)$ from $H^1(\mathbb{R}^d)$ to $\mathcal{E}_1(b_0)$. Also in this sense the situation is much more difficult than in Theorem 8.6 and requires a more sophisticated analysis.

The estimate (12.18) shows that, after passing to a subsequence $(u_m)_m$, the functions $u_m(t)$ tend to $u(t)$ in $L^p(\mathbb{R}^d)$ as $m \to \infty$, for a.e. $t \in J$. Recall that

$$u_m - u = T(\cdot)(\varphi_m - u_0) + T *_+ (F(u_m) - F(u)) \quad \text{on } J_0.$$

Let $b \in (0, b_0]$. Strichartz' estimates (12.5) and (12.6) now yield

$$|||u_m - u|||_{1,b} \le C_{\mathrm{St}}\big(\|\varphi_m - u_0\|_{1,2} + \|F(u_m) - F(u)\|_{L^{q'}([-b,b],W^1_{p'})}\big) \tag{12.19}$$

for $m \in \mathbb{N}$. (Observe that we may assume $m \ge n_0$ and that $F(u_m)$ and $F(u)$ belong to $E_1'(b) = L^{q'}([-b, b], W^1_{p'}(\mathbb{R}^d))$ by Lemma 12.1.) Let $\phi(z) = -\mathrm{i}\mu|z|^{\alpha-1}z$ for $z \in \mathbb{R}^2$ and $j \in \{1, \dots, d\}$. As in (12.13) we compute

$$\begin{aligned}
&\|\partial_j F(u_m(t)) - \partial_j F(u(t))\|_{p'}\\
&\quad\le \|\phi'(u_m(t))[\partial_j u_m(t) - \partial_j u(t)]\|_{p'} + \|[\phi'(u_m(t)) - \phi'(u(t))]\partial_j u(t)\|_{p'}\\
&\quad\le \alpha\|u_m(t)\|_p^{\alpha-1}\|\partial_j u_m(t) - \partial_j u(t)\|_p + \|[\phi'(u_m(t)) - \phi'(u(t))]\partial_j u(t)\|_{p'}\\
&\quad\le \alpha C_{\mathrm{So}}^{\alpha-1}\overline{r}^{\alpha-1}\|u_m(t) - u(t)\|_{1,p} + \|[\phi'(u_m(t)) - \phi'(u(t))]\partial_j u(t)\|_{p'},
\end{aligned}$$

for $t \in J_0$, where we use Sobolev's embedding (12.4) and inequality (12.17). We next take the norm in $L^{q'}([-b, b])$. Employing also (12.8) and Hölder's inequality, it follows

$$\|F(u_m) - F(u)\|_{E_1'(b)} \le c_1 \overline{r}^{\alpha-1} b^{\frac{1}{q'}-\frac{1}{q}} \|u_m - u\|_{E_1(b)} + c_2 \|(\phi'(u_m) - \phi'(u))|\nabla u|\|_{L^{q'}([-b,b],L^{p'})}.$$

Here and below, c_j is a constant only depending on d and α. We fix

$$b = \min\{b_0, (2c_1 C_{\mathrm{St}} \overline{r}^{\alpha-1})^{\frac{q'q}{q'-q}}\}$$

and insert the above inequality into (12.19), arriving at

$$\begin{aligned} |\!|\!| u_m - u |\!|\!|_{1,b} &\le C_{\mathrm{St}} \|\varphi_m - u_0\|_{1,2} + \tfrac{1}{2}\|u_m - u\|_{E_1(b)} \\ &\quad + c_2 C_{\mathrm{St}} \|(\phi'(u_m) - \phi'(u))|\nabla u|\|_{E_0'(b)}, \\ |\!|\!| u_m - u |\!|\!|_{1,b} &\le 2C_{\mathrm{St}} \|\varphi_m - u_0\|_{1,2} + 2c_2 C_{\mathrm{St}} \|(\phi'(u_m) - \phi'(u))|\nabla u|\|_{E_0'(b)}. \end{aligned} \tag{12.20}$$

Corollary 9.3 implies that the functions $\phi'(u_m(t))|\nabla u(t)|$ tend to $\phi'(u(t))|\nabla u(t)|$ in $L^{p'}(\mathbb{R}^d)$ as $m \to \infty$ for a.e. $t \in [-b,b]$, since $u_m(t) \to u(t)$ in $L^p(\mathbb{R}^d)$. Combining Corollary 9.3 with (12.17) and (12.4), we further estimate

$$\begin{aligned} \|(\phi'(u_m(t)) - \phi'(u(t)))|\nabla u(t)|\|_{p'} &\le c_3(\|u_m(t)\|_p^{\alpha-1} + \|u(t)\|_p^{\alpha-1})\||\nabla u(t)|\|_p \\ &\le 2c_3 C_{\mathrm{So}}^{\alpha-1} \overline{r}^{\alpha-1} \|u(t)\|_{1,p}. \end{aligned}$$

Since $q' < q$, the function u belongs to $L^{q'}([-b,b], W_p^1(\mathbb{R}^d))$. Due to dominated convergence, the second term on the right-hand side in (12.20) thus tends to 0 as $m \to \infty$. As a result, $u_m \to u$ in $\mathcal{E}_1(b)$ and we can fix an index m_1 such that $\|u_m(\pm b)\|_{1,2} \le \overline{\rho}$ for all $m \ge m_1$.

We repeat the above argument on the intervals $[-2b, 0]$ and $[0, 2b]$, passing to further subsequences. In finitely many steps, we thus construct a subsequence $(\varphi_{n_j})_j$ with $J \subseteq J(\varphi_{n_j})$ for all $j \in \mathbb{N}$ and $u_{n_j} \to u$ in $\mathcal{E}_1(J)$ as $j \to \infty$. □

Our derivation of the conservation laws in Lecture 10 was valid only for H^2–solutions. Of course, such solutions can only exist if $u_0 \in H^2(\mathbb{R}^d)$. By a refinement of the above proof we show in the next lecture that the solution of (12.1) preserves this initial regularity on its full existence interval $J(u_0)$.

To this aim, we will employ the X-valued Sobolev space $W_r^1(J,X)$ for $r \in [1,\infty]$, an open interval $J \subseteq \mathbb{R}$ and a Banach space X (see Appendix F). One says that $u \in W_r^1(J,X)$ if $u \in L^r(J,X)$ and there exists $v \in L^r(J,X)$ such that

$$u(t) = u(a) + \int_a^t v(s)\,\mathrm{d}s$$

for all $t, a \in J$ (where Proposition F.15 allows to choose a representative of u belonging to $C(\overline{J}, X)$). We set $u' = v$ (see Definition F.14).

We discuss a few properties of these spaces needed later on. The space $W_r^1(J,X)$ is a Banach space when equipped with the norm given by

$$\|u\|_{1,r} = \begin{cases} (\|u\|_r^r + \|u'\|_r^r)^{\frac{1}{r}}, & \text{if } 1 \le r < \infty, \\ \max\{\|u\|_\infty, \|u'\|_\infty\}, & \text{if } r = \infty, \end{cases}$$

where $\|\cdot\|_r$ is the norm on $L^r(J,X)$. Moreover $W^1_r(J,X)$ is isometrically isomorphic to a closed subspace of $L^r(J,X)^2$ via the map $u \mapsto (u,u')$. The observations in Remark 11.4 then lead to the following properties.

REMARK 12.7. (a) If $1 \le r < \infty$ and X is separable, then $W^1_r(J,X)$ is separable.

(b) If $1 < r < \infty$ and X is reflexive, then $W^1_r(J,X)$ is reflexive.

(c) Let X be reflexive. Then $W^1_\infty(J,X)$ is isometrically isomorphic to the space of bounded Lipschitz functions $u : \overline{J} \to X$. (See §1.2 in [**ABHN11**].)

(d) Let $a < b < c$ and $r \in [1,\infty)$. Let $u \in W^1_r((a,b),X)$ and $v \in W^1_r((b,c),X)$ satisfy $u(b) = v(b)$. Define $w(t) = u(t)$ for $t \in (a,b)$, $w(b) = u(b)$ and $w(t) = v(t)$ for $t \in (b,c)$. Set $g(t) = u'(t)$ for $t \in (a,b)$ and $g(t) = v'(t)$ for $t \in (b,c)$. It is then straight forward to check that $g \in L^r((a,c),X)$ is the derivative of w. The concatenation w thus belongs to $W^1_r((a,c),X)$. ♢

We further need a simple density and embedding result for $W^1_p(J,X)$.

LEMMA 12.8. *Let $J \subseteq \mathbb{R}$ be an open and bounded interval. Then the following assertions hold.*

(a) If $1 \le r < \infty$ then $C^1(\overline{J},X)$ is dense in $W^1_r(J,X)$.
(b) If $1 \le r \le \infty$, then $W^1_r(J,X) \hookrightarrow C(\overline{J},X)$.

PROOF. Let $u \in W^1_r(J,X)$. We choose a representative of u which belongs to $C(\overline{J},X)$, see Proposition F.15.

(a) Let $1 \le r < \infty$ and $\overline{J} = [a,b]$. We can approximate u' in $L^r(J,X)$ by $v_n \in C_c(J,X)$, $n \in \mathbb{N}$, see steps 3) and 4) of the proof of Lemma 11.5. Setting

$$u_n(t) = u(a) + \int_a^t v_n(s)\,\mathrm{d}s, \qquad t \in \overline{J},\ n \in \mathbb{N},$$

we obtain functions $u_n \in C^1(\overline{J},X)$ with $u_n' = v_n \to u'$ in $L^r(J,X)$. Hölder's inequality further yields

$$\|u_n(t) - u(t)\|_X^r \le \Big(\int_a^t \|v_n(s) - u'(s)\|_X\,\mathrm{d}s\Big)^r \le |J|^{\frac{r}{r'}} \int_J \|v_n(s) - u'(s)\|_X^r\,\mathrm{d}s,$$
$$\|u_n - u\|_r \le |J| \|v_n - u'\|_r \longrightarrow 0$$

as $n \to \infty$, where $|J|$ is the length of J. Thus, (a) is true.

(b) First let $J = (0,1)$. We set $\widetilde{u}(t) = u(t)$ for $t \in [0,1)$ and $\widetilde{u}(t) = u(-t)$ for $t \in (-1,0)$. The function $\widetilde{u}$ belongs to $W^1_r((-1,1),X)$ with derivative given by $\widetilde{u}'(t) = u'(t)$ for $t \in (0,1)$ and $\widetilde{u}'(t) = -u'(-t)$ for $t \in (-1,0)$, since

$$u(-t) = u(0) - \int_0^t u'(-s)\,\mathrm{d}s$$

for $t \in (-1,0)$. Let $\widetilde{J} = (-1,1)$. Take a function $\phi \in C^1(\widetilde{J})$ with $\phi = 1$ on $[0,1)$ and $\phi = 0$ on $(-1,-\frac{1}{2})$. Set $v = \phi\widetilde{u}$. We claim that $v \in W^1_r(\widetilde{J},X)$ with $v' = \phi\widetilde{u}' + \phi'\widetilde{u}$. In fact, take $\widetilde{u}_n \in C^1(\widetilde{J},X)$ converging to $\widetilde{u}$ in $W^1_r(\widetilde{J},X)$. Passing to a subsequence, we may assume that $\widetilde{u}_n(t) \to \widetilde{u}(t)$ as $n \to \infty$ for

$t \in \widetilde{J} \setminus N$ and a set N of measure 0. Taking $t, a \in \widetilde{J} \setminus N$, we thus obtain

$$\begin{aligned} v(t) - v(a) &= \lim_{n\to\infty} (\phi(t)\widetilde{u}_n(t) - \phi(a)\widetilde{u}_n(a)) = \lim_{n\to\infty} \int_a^t (\phi(s)\widetilde{u}_n'(s) + \phi'(s)\widetilde{u}_n(s)\,\mathrm{d}s \\ &= \int_a^t (\phi(s)\widetilde{u}'(s) + \phi'(s)\widetilde{u}(s))\,\mathrm{d}s. \end{aligned}$$

Since v is continuous, this equation holds for all $t, a \in \widetilde{J}$ and the claim is shown.

For $t \in [0,1]$ we thus derive

$$u(t) = v(t) = \int_{-1}^t v'(s)\,\mathrm{d}s, \qquad \|u(t)\| \le \|v'\|_1 \le c\|u\|_{W_r^1((0,1),X)},$$

where c only depends on ϕ. Hence, (b) holds in this case.

The assertion for a general interval $J = (a,b)$ follows by means of the transformation $t \mapsto \frac{t-a}{b-a}$ from J to $(0,1)$. □

Exercises

EXERCISE 12.1. Let $d = 3$ and $\phi = (\phi_1, \phi_2) \in C^3(\mathbb{R}^2, \mathbb{R}^2)$. For $u : \mathbb{R}^3 \to \mathbb{C}$ set $F(u) = \phi_1(\operatorname{Re} u, \operatorname{Im} u) + \mathrm{i}\phi_2(\operatorname{Re} u, \operatorname{Im} u) =: \phi(u)$. Show that $F : H^2(\mathbb{R}^3) \to H^2(\mathbb{R}^3)$ is Lipschitz on bounded subsets of $X = H^2(\mathbb{R}^3)$. Show that one can apply Theorem 8.6 to the nonlinear Schrödinger equation

$$u'(t) = \mathrm{i}\Delta u(t) + \mathrm{i}F(u(t)), \quad t \ge 0, \qquad u(0) = u_0,$$

on X. What is the difference between this result and Theorem 12.6?

EXERCISE 12.2. Let $d \le 5$ and $\alpha \in (2, \alpha_c)$. Set $\phi(z) = z|z|^{\alpha-1}$ for $z \in \mathbb{R}^2$.

(a) Show that $\phi \in C^2(\mathbb{R}^2, \mathbb{R}^2)$ with $|\phi''(z)| \le c_0|z|^{\alpha-2}$ for all $z \in \mathbb{R}^2$ and some constant $c_0 > 0$.

(b) We use the notations of Lecture 12. Let $b > 0$ and $u, v \in \mathcal{E}_1(b)$ with $|\!|\!|u|\!|\!|_{1,b}, |\!|\!|v|\!|\!|_{1,b} \le r$ for some $r > 0$. Let $p = 1 + \alpha$ and $\frac{2}{q} + \frac{d}{p} = \frac{d}{2}$. Show that

$$\|F(u) - F(v)\|_{L^{q'}(J, W^1_{p'})} \le c r^{\alpha-1} b^{\frac{1}{q'} - \frac{1}{q}} |\!|\!|u - v|\!|\!|_{1,b}$$

for some constant $c > 0$ only depending on c_0, α and d.

(c) Let $u_0 \in H^1(\mathbb{R}^d)$. Show that there is a radius $\delta > 0$ and a time $b_0 > 0$ such that $[-b_0, b_0] \subseteq J(v_0)$ and

$$\overline{B}_{H^1}(u_0, \delta) \to \mathcal{E}_1(b_0), \quad v_0 \mapsto u(\,\cdot\,; v_0),$$

is Lipschitz continuous.

EXERCISE 12.3. Let $V, \Gamma \in W^1_\infty(\mathbb{R}^d)$ be real valued. Set $F_1(u) = -\mathrm{i}Vu$ and $F_2(u) = -\mathrm{i}\Gamma u|u|^{\alpha-1}$ for $u \in H^1(\mathbb{R}^d)$, as well as $F = F_1 + F_2$. Let $\rho > 0$ and $u_0 \in H^1(\mathbb{R}^d)$ with $\|u_0\|_{1,2} \le \rho$. Show that there is a time $b_0 > 0$ such that the nonlinear Schrödinger equation

$$u'(t) = \mathrm{i}\Delta u(t) + F(u(t)), \quad t \in [-b_0, b_0], \qquad u(0) = u_0,$$

has an H^1–solution. (Hint: One can use Theorem 11.6 with $(\overline{q}, \overline{p}) = (\infty, 2)$.)

LECTURE 13

Regularity and global existence for the nonlinear Schrödinger equation

In Theorem 12.6 we have constructed H^1–solutions $u \in C^1(J(u_0), H^{-1}(\mathbb{R}^d)) \cap C(J(u_0), H^1(\mathbb{R}^d))$ of the nonlinear Schrödinger equation

$$u'(t) = \mathrm{i}\Delta u(t) - \mathrm{i}\mu |u(t)|^{\alpha-1} u(t), \quad t \in J(u_0), \qquad u(0) = u_0, \tag{13.1}$$

for $u_0 \in H^1(\mathbb{R}^d)$, the maximal existence interval $J(u_0)$, $\mu \in \{-1, 1\}$ and $1 < \alpha < \alpha_c = \frac{d+2}{(d-2)_+}$. As we saw in Example 10.3, there is a blow-up solution in the focusing case $\mu = -1$ for $\alpha = 1 + \frac{4}{d}$. In this lecture we present three positive results on global existence for (13.1), namely:

(a) *Defocusing case*: If $\mu = 1$, then $J(u_0) = \mathbb{R}$ for all $u_0 \in H^1(\mathbb{R}^d)$ and $\alpha \in (1, \alpha_c)$.
(b) *Small* α: If $\alpha \in (1, 1 + \frac{4}{d})$, then $J(u_0) = \mathbb{R}$ for all $u_0 \in H^1(\mathbb{R}^d)$.
(c) *Small* u_0: There is a radius $\rho > 0$ such that $J(u_0) = \mathbb{R}$ for all $u_0 \in H^1(\mathbb{R}^d)$ with $\|u_0\|_{1,2} \le \rho$, where $\alpha \in (1, \alpha_c)$.

In all three cases the nonlinearity is relatively tame so that it does not destroy the global existence that we have in the linear case: In (a) we have the right sign $\mu = 1$. In (b) the nonlinearity does not grow too much. In (c) the solution u is small initially which leads to an even smaller nonlinearity $|u|^{\alpha-1}u$. These results are shown in Theorem 13.3. Moreover, the proof of (c) leads to the Lyapunov stability of the solution $u_* = 0$ of (13.1), see Corollary 13.4.

The basic criteria (a) – (c) for global existence rely on the conservation laws for the energy and for the L^2–norm of H^1–solutions. In Remarks 10.4 and 10.5 we derived them only for $u \in C^1(J(u_0), H^1(\mathbb{R}^d)) \cap C(J(u_0), H^2(\mathbb{R}^d))$. We show the conservation laws for H^1–solutions in Theorem 13.2 by various approximation arguments. In this proof we work with H^2–solutions. For our reasoning it is crucial to know that the maximal H^1–solution of (13.1) is even an H^2–solution on its *full* existence interval $J(u_0)$ provided that $u_0 \in H^2(\mathbb{R}^d)$. Thus, the solution $u(t)$ preserves the initial regularity $u_0 \in H^2(\mathbb{R}^d)$ for all $t \in J(u_0)$, which is of independent interest.

To prove this fact, as in Theorem 12.6 we look for a fixed point v satisfying

$$v = T(\cdot)u_0 + T *_+ F(v) =: \Phi_{u_0}(v),$$

where $F(v) = -\mathrm{i}\mu|v|^{\alpha-1}v$ and $T(\cdot)$ is the free Schrödinger group. In contrast to Lemma 12.5 we now consider Φ on a space involving differentiability of v in time. The resulting fixed point then turns out to be an H^2–solution and it coincides with the known H^1–solution u from Theorem 12.6.

At the end of the lecture we briefly discuss further results on blow-up and global existence.

We now come to the announced regularity theorem. As in the previous lectures, we set $p = 1+\alpha \in (2, 1+\alpha_c)$ and $\frac{2}{q} = \frac{d}{2} - \frac{d}{p}$. It follows that $p, q \in (2,\infty)$, $p' = \frac{p}{\alpha}$ and that (q,p) is admissible in the sense of (11.13). In the next proof we use properties of Banach space valued Sobolev spaces stated in Remark 12.7 and Lemma 12.8.

THEOREM 13.1. *Let $1 < \alpha < \frac{d+2}{(d-2)_+}$, $\mu \in \{-1,1\}$, and $u_0 \in H^2(\mathbb{R}^d)$. Then the maximal H^1–solution of (13.1) is even an H^2–solution on $J(u_0)$, i.e., $u \in C^1(J(u_0), L^2(\mathbb{R}^d)) \cap C(J(u_0), H^2(\mathbb{R}^d))$. It further satisfies $u \in W^1_q(I, L^p(\mathbb{R}^d))$ for all open bounded intervals I with $\overline{I} \subseteq J(u_0)$.*

PROOF. Let $u_0 \in H^2(\mathbb{R}^d)$ and let u be the maximal H^1–solution of (13.1) obtained in Theorem 12.6. Take any compact interval $J_0 \subseteq J(u_0)$ containing 0. We have to show that u is an H^2–solution on J_0 with $u \in W^1_q(\mathring{J}_0, L^p(\mathbb{R}^d))$. By a refinement of the fixed point argument in the proof of Theorem 12.6, we first prove this claim on an interval $J_1 = [-b_1, b_1]$. It turns out that this time $b_1 > 0$ only depends on α, d and

$$\overline{\rho} := \max_{t \in J_0} \|u(t)\|_{1,2},$$

and in particular not on $\|u_0\|_{2,2}$. We can thus repeat the argument for the initial values $u(\pm b_1)$ with the same time step size b_1 and deduce the assertion in finitely many iterations. Throughout we use the setting and the notation of the proof of Theorem 12.6.

1) We fix $r = 1 + C_{\mathrm{St}}\overline{\rho}$. Let $b > 0$. We set $J = (-b, b)$ and use the spaces

$$\mathcal{E}_k(b) = L^\infty(J, H^k(\mathbb{R}^d)) \cap L^q(J, W^k_p(\mathbb{R}^d)),$$

endowed with the norms given by

$$|||v|||_{k,b} = \max\{\|v\|_{L^\infty(J,H^k)}, \|v\|_{L^q(J,W^k_p)}\},$$

where $k \in \{0,1\}$. Lemma 12.5 says that the H^1–solution u is a fixed point of the operator Φ (see (13.2)) in the set $\Sigma(r,b) = \overline{B}_{\mathcal{E}_1(b)}(0,r)$, where $0 < b \le b_0(\|u_0\|_{1,2})$. We will show that it is also a fixed point in a subset of more regular functions, using that $u_0 \in H^2(\mathbb{R}^d)$. To this aim, for $R \ge C_{\mathrm{St}}\|\Delta u_0\|_2 =: R_0$ (to be fixed below) we define the spaces

$$\begin{aligned} \mathcal{F}(b) &= \mathcal{E}_1(b) \cap W^1_q(J, L^p(\mathbb{R}^d)) \cap W^1_\infty(J, L^2(\mathbb{R}^d)), \\ \Theta = \Theta(R,b) &= \{v \in \mathcal{F}(b) \,|\, v(0) = u_0, |||v|||_{1,b} \le r, |||v'|||_{0,b} \le R\}. \end{aligned}$$

Strichartz' estimate (12.5) yields that $|||T(\cdot)u_0|||_{1,b} \le C_{\mathrm{St}}\overline{\rho} \le r$ and $|||\frac{\mathrm{d}}{\mathrm{dt}}T(\cdot)u_0|||_{0,b} = |||T(\cdot)\Delta u_0|||_{0,b} \le C_{\mathrm{St}}\|\Delta u_0\|_2 \le R$. Hence, $\Theta(R,b) \ne \emptyset$.

We endow Θ with the metric given by $|||v-w|||_{0,b}$. We recall from Remark 12.7 (c) that $W^1_\infty(J, L^2(\mathbb{R}^d))$ is isomorphic to the space of Lipschitz functions $f : \overline{J} \to L^2(\mathbb{R}^d)$. Moreover, for v in these isomorphic spaces the Lipschitz constant of v coincides with $\|v'\|_{L^\infty(J,L^2)}$. We first claim that Θ is complete.

In fact, take a Cauchy sequence $(v_n)_n$ in Θ. In Lemma 12.2 we have seen that $(v_n)_n$ converges in $\mathcal{E}_0(b)$ to a function $v \in \mathcal{E}_1(b)$ with $|||v|||_{1,b} \le r$ as $n \to \infty$. Since the maps $v_n : J \to L^2(\mathbb{R}^d)$ converge in $L^\infty(J, L^2(\mathbb{R}^d))$ to v and are uniformly Lipschitz with bound R, we conclude that $(v_n)_n$ tends to v in $C(\overline{J}, L^2(\mathbb{R}^d))$ as $n \to \infty$, that $v(0) = u_0$ and that $v : \overline{J} \to L^2(\mathbb{R}^d)$ is Lipschitz with bound R. Using again Remark 12.7 (c), we see that the function v belongs to $W^1_\infty(J, L^2(\mathbb{R}^d))$ with $\|v'\|_{L^\infty(J,L^2)} \le R$. Further, after passing to a subsequence, $(v_{n_j})_j$ tends weakly in $W^1_q(J, L^p(\mathbb{R}^d))$ to a function $w \in W^1_q(J, L^p(\mathbb{R}^d))$ as $j \to \infty$ with $\|w'\|_{L^q(J,L^p)} \le R$. Since $L^q(J, L^p(\mathbb{R}^d))^* \hookrightarrow W^1_q(J, L^p(\mathbb{R}^d))^*$ and $\frac{\mathrm{d}}{\mathrm{d}t} : W^1_q(J, L^p(\mathbb{R}^d)) \to L^q(J, L^p(\mathbb{R}^d))$ is linear and bounded, v_{n_j} and v'_{n_j} converge weakly in $L^q(J, L^p(\mathbb{R}^d))$ to w and w', respectively. We thus obtain $v = w$, $v \in \mathcal{F}(b)$ and $|||v'|||_{0,b} \le R$. Summing up, $v \in \Theta$ and Θ is complete.

2) Let $v, w \in \Theta(R, b)$ for $R \ge R_0$ and $b \in (0, b_0(\overline{\rho})]$, see (12.16). We define again

$$\Phi(v)(t) = T(t)u_0 + \int_0^t T(t-s)F(v(s))\,\mathrm{d}s = T(t)u_0 + \int_0^t T(s)F(v(t-s))\,\mathrm{d}s \tag{13.2}$$

for $t \in \overline{J} = [-b, b]$. We fix below R and b so that $\Phi : \Theta(R, b) \to \Theta(R, b)$ is a strict contraction. In the proof of Lemma 12.5 we have already shown that $\Phi(v) \in \mathcal{E}_1(b) \cap C(\overline{J}, H^1(\mathbb{R}^d))$,

$$|||\Phi(v) - \Phi(w)|||_{0,b} \le \frac{1}{2}|||v - w|||_{0,b} \quad \text{and} \quad |||\Phi(v)|||_{1,b} \le r, \tag{13.3}$$

see (12.14) – (12.16).

3) We next prove that $\frac{\mathrm{d}}{\mathrm{d}t}\Phi(v) \in \mathcal{E}_0(b)$ with $|||\frac{\mathrm{d}}{\mathrm{d}t}\Phi(v)|||_{0,b} \le R$ for all $v \in \Theta(R, b)$. To this aim, we first differentiate the integral in (13.2) with respect to t. This is done via an approximation argument. Corollary 9.3 shows that $F : L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d)$ is real continuously differentiable with derivative given by $F'(\varphi)\psi = \phi'(\varphi)\psi$ for $\varphi, \psi \in L^p(\mathbb{R}^d)$ and $\phi(z) = -\mathrm{i}\mu|z|^{\alpha-1}z$ for $z \in \mathbb{R}^2$. Moreover,

$$\|F'(\varphi)\psi\|_{p'} \le c_1\|\varphi\|_p^{\alpha-1}\|\psi\|_p. \tag{13.4}$$

Here and below c_j is a constant only depending on α and d. Lemma 12.8 (a) allows to approximate v in $W^1_q(J, L^p(\mathbb{R}^d))$ by $w_n \in C^1(\overline{J}, L^p(\mathbb{R}^d))$. Passing to a subsequence if necessary, we may further assume that $w'_n(t)$ converges in $L^p(\mathbb{R}^d)$ as $n \to \infty$ and $\|w'_n(t)\|_p \le h(t)$ for all $n \in \mathbb{N}$, a.e. $t \in J$ and a function $h \in L^q(J) \hookrightarrow L^{q'}(J)$, where we note that $q' < 2 < q$. Finally, taking $a = 0$ and $J = (-b, b)$ in the proof of Lemma 12.8 (b), we see that $w_n(0) = v(0) = u_0$. Due to Lemma 12.8 (b), the sequence $(w_n)_n$ converges to v in $C(\overline{J}, L^p(\mathbb{R}^d))$. It is thus bounded by a constant $\overline{c}$ in this space. The properties of F yield that the functions $F'(w_n(t))w'_n(t)$ tend to $F'(v(t))v'(t)$ in $L^{p'}(\mathbb{R}^d)$ as $n \to \infty$ for a.e. $t \in J$ and that

$$\sup_{n\in\mathbb{N}} \|F'(w_n(t))w'_n(t)\|_{p'} \le \sup_{s\in\overline{J}, n\in\mathbb{N}} c_1\|w_n(s)\|_p^{\alpha-1}\|w'_n(t)\|_p \le c_1\overline{c}^{\,\alpha-1}h(t)$$

for a.e. $t \in J$. From dominated convergence we deduce that $F'(w_n)w'_n \to F'(v)v'$ in $L^{q'}(J, L^{p'}(\mathbb{R}^d))$ as $n \to \infty$.

Since $L^{p'}(\mathbb{R}^d) \hookrightarrow H^{-1}(\mathbb{R}^d)$, we have $F(w_n) \in C^1(\overline{J}, H^{-1}(\mathbb{R}^d))$ and so the derivative

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_0^t T(s)F(w_n(t-s))\,\mathrm{d}s = T(t)F(w_n(0)) + \int_0^t T(s)F'(w_n(t-s))w_n'(t-s)\,\mathrm{d}s$$
$$= T(t)F(u_0) + \int_0^t T(t-s)F'(w_n(s))w_n'(s)\,\mathrm{d}s$$

exists in $H^{-1}(\mathbb{R}^d)$. (In this calculation we identify $\mathbb{C}$ with $\mathbb{R}^2$.) Due to Strichartz' estimate (12.6), the right-hand side of the above identity is continuous in $L^2(\mathbb{R}^d)$ and converges to

$$T(t)F(u_0) + \int_0^t T(t-s)F'(v(s))v'(s)\,\mathrm{d}s$$

in $L^2(\mathbb{R}^d)$ uniformly in t as $n \to \infty$. Similarly, the integral on the left-hand side tends to $T *_+ F(v)(t)$ in $L^2(\mathbb{R}^d)$ uniformly in t. We can thus differentiate the integral in (13.2) in $L^2(\mathbb{R}^d)$ and obtain

$$\frac{\mathrm{d}}{\mathrm{dt}} \int_0^t T(s)F(v(t-s))\,\mathrm{d}s = T(t)F(u_0) + \int_0^t T(t-s)F'(v(s))v'(s)\,\mathrm{d}s,$$
$$\frac{\mathrm{d}}{\mathrm{dt}} \Phi(v)(t) = T(t)(\mathrm{i}\Delta u_0 + F(u_0)) + \int_0^t T(t-s)F'(v(s))v'(s)\,\mathrm{d}s \qquad (13.5)$$

for all $t \in J$.

4) In this step we establish that $\frac{\mathrm{d}}{\mathrm{dt}}\Phi(v) \in \mathcal{E}_0(b)$ and estimate its norm. It is crucial that R will enter only linearly. Using inequality (13.4), Sobolev's embedding (12.4) and $\|v(s)\|_{1,2} \le r$, we derive

$$\|F'(v(s))v'(s)\|_{p'} \le c_1 C_{\mathrm{So}}^{\alpha-1} r^{\alpha-1} \|v'(s)\|_p$$

for all $s \in J$. Strichartz' estimate (12.6) and Hölder's inequality now allow to bound the $\mathcal{E}_0(b)$–norm of the integral term in (13.5) by

$$C_{\mathrm{St}}\|F'(v)v'\|_{L^{q'}(J,L^{p'})} \le C_{\mathrm{St}} c_1 C_{\mathrm{So}}^{\alpha-1} r^{\alpha-1} \|v'\|_{L^{q'}(J,L^p)} \le c_2 r^{\alpha-1} b^{\frac{1}{q'}-\frac{1}{q}} \|v'\|_{L^q(J,L^p)}$$
$$\le c_2 r^{\alpha-1} b^{\frac{1}{q'}-\frac{1}{q}} R \qquad (13.6)$$

with $c_2 := 2^{\frac{1}{q'}-\frac{1}{q}} c_1 C_{\mathrm{St}} C_{\mathrm{So}}^{\alpha-1}$, using $v \in \Theta(R,b)$. We further recall that $H^2(\mathbb{R}^d) \hookrightarrow L^{2\alpha}(\mathbb{R}^d)$ by Sobolev's embedding (5.5) since $\alpha < \frac{d}{(d-4)_+}$. Hence, $\|F(u_0)\|_2 = \|u_0\|_{2\alpha}^\alpha \le c_3 \|u_0\|_{2,2}^\alpha$. Strichartz' estimates, (13.5) and (13.6) thus yield that

$$\|\!|\tfrac{\mathrm{d}}{\mathrm{dt}}\Phi(v)\|\!|_{0,b} \le C_{\mathrm{St}}(\|\Delta u_0\|_2 + c_3\|u_0\|_{2,2}^\alpha) + c_2 r^{\alpha-1} b^{\frac{1}{q'}-\frac{1}{q}} R \qquad (13.7)$$

and that $\frac{\mathrm{d}}{\mathrm{dt}}\Phi(v)$ belongs to $C(\overline{J}, L^2(\mathbb{R}^d))$. We fix $R_1 = 2C_{\mathrm{St}}(\|\Delta u_0\|_2 + c_3\|u_0\|_{2,2}^\alpha) \ge R_0$ and choose

$$b_1 = \min\{b_0(\overline{\rho}), (2c_2 r^{\alpha-1})^{\frac{q'-q}{q'q}}\}.$$

Since $r = 1 + C_{\mathrm{St}}\overline{\rho}$, the number b_1 only depends on $\overline{\rho}$, α and d. The inequalities (13.3) and (13.7) now show that $\Phi : \Theta(R_1, b_1) \to \Theta(R_1, b_1)$ is a strict contraction. We thus obtain a fixed point $v_* = \Phi(v_*) \in \Theta(R_1, b_1) \subseteq \mathcal{F}(b_1)$ with $v_* \in C(J_1, H^1(\mathbb{R}^d)) \cap C^1(J_1, L^2(\mathbb{R}^d)) \cap W_q^1(\mathring{J}_1, L^p(\mathbb{R}^d))$, where $J_1 := [-b_1, b_1]$.

By Lemma 12.5, the H^1–solution u is the only fixed point of Φ in $\Sigma(r, b_1) \supseteq \Theta(R_1, b_1)$. Hence, $u = v_* \in C^1(J_1, L^2(\mathbb{R}^d)) \cap C(J_1, H^1(\mathbb{R}^d)) \cap W^1_q(\mathring{J}_1, L^p(\mathbb{R}^d))$.

5) We still have to show that $u \in C(J_1, H^2(\mathbb{R}^d))$. To prove this, we use a "boot-strapping" argument based on the following fact, see Theorems 4.3.8 (ii) and 4.3.10 (ii) in [**Kry08**].

> Let $r, s \in (1, \infty)$. Then the operator $I - \Delta : W^2_r(\mathbb{R}^d) \to L^r(\mathbb{R}^d)$ is invertible with bounded inverse R_r. If $g \in L^r(\mathbb{R}^d) \cap L^s(\mathbb{R}^d)$, then $R_r g = R_s g$. We thus write $(I - \Delta)^{-1}$ instead of R_r.

As a starting point we note that (13.1) yields

$$u - \Delta u = u + \mathrm{i}u' - \mathrm{i}F(u) = f + g,$$

where $f := u + \mathrm{i}u' \in C(J_1, L^2(\mathbb{R}^d))$ and $g := -\mathrm{i}F(u) \in C(J_1, L^{\frac{p}{\alpha}}(\mathbb{R}^d))$ since $u \in C^1(J_1, L^2(\mathbb{R}^d)) \cap C(J_1, L^p(\mathbb{R}^d))$ and $p' = \frac{p}{\alpha}$. Note that $\frac{p}{\alpha} \in (1, 2)$. The above stated regularity result for Δ shows that $(I - \Delta)^{-1} f \in C(J_1, H^2(\mathbb{R}^d))$ and $(I - \Delta)^{-1} g \in C(J_1, W^2_{\frac{p}{\alpha}}(\mathbb{R}^d))$. Sobolev's embedding (5.5) thus yields

$$u = (I - \Delta)^{-1}(f + g) \in C(J_1, L^{r_1}(\mathbb{R}^d))$$

for $r_1 = \frac{d}{d\alpha - 2p} p =: \gamma p > 2$ if $d\alpha > 2p$ and for any $r_1 \in (2, \infty)$ otherwise. Note that $\gamma > 1$ if $d\alpha > 2p$ since $p = \alpha + 1$ and $\alpha < \frac{d+2}{(d-2)_+}$.

This extra integrability of u implies that $g \in C(J_1, L^{\frac{r_1}{\alpha}}(\mathbb{R}^d))$. If $r_1 \geq 2\alpha$, we obtain that g belongs to $C(J_1, L^2(\mathbb{R}^d))$ since then $L^{\frac{p}{\alpha}}(\mathbb{R}^d) \cap L^{\frac{r_1}{\alpha}}(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d)$ by (D.13). As a result, $u = (I - \Delta)^{-1}(f + g) \in C(J_1, H^2(\mathbb{R}^d))$.

If $r_1 < 2\alpha$, as above we infer that $u \in C(J_1, L^{r_2}(\mathbb{R}^d))$ for $r_2 = \frac{d r_1}{d\alpha - 2r_1} \geq \gamma r_1 = \gamma^2 p$ if $d\alpha > 2r_1$ and for any $r_2 \in (2, \infty)$ if $d\alpha \leq 2r_1$. Since $\gamma > 1$, in finitely many steps we arrive at $r_m \geq \gamma^m p \geq 2\alpha$, and hence $u \in C(J_1, H^2(\mathbb{R}^d))$.

6) We can now finish the proof. If $J_0 \subseteq J_1$ we are done. If not, assume that $\max J_0 > b_1$. Since $u(b_1) \in H^2(\mathbb{R}^d)$, we can repeat steps 2) – 5) with initial value $u(b_1)$ and the same time step b_1. We then obtain an H^2–solution u_1 of (13.1) on $[b_1, 2b_1]$ with $u_1(b_1) = u(b_1)$. By Lemma 12.3, we can glue together these functions to an H^2–solution v on $[-b_1, 2b_1]$ with $v \in W^1_q((-b_1, 2b_1), L^p(\mathbb{R}^d))$. The uniqueness of H^1–solutions yields that $v = u$ on $[-b_1, 2b_1]$. We can iterate this procedure and derive in finitely many steps that u is an H^2–solution of (12.1) on J_0 with $u \in W^1_q(\mathring{J}_0, L^p(\mathbb{R}^d))$, where we use Remark 12.7 (d). □

Before discussing global existence, we next derive the conservation laws for the L^2–norm and the energy of H^1–solutions u. We had shown these laws in Remarks 10.4 and 10.5 for more regular solutions. Theorems 12.6 and 13.1 now allow to extend these results to H^1–solutions by approximation. We recall that

$$E(v) = \tfrac{1}{2}\|\nabla v\|_2^2 + \tfrac{\mu}{\alpha+1}\|v\|_{\alpha+1}^{\alpha+1} = \int_{\mathbb{R}^d} \big(\tfrac{1}{2}|\nabla v|^2 + \tfrac{\mu}{\alpha+1}|v|^{\alpha+1}\big)\,\mathrm{d}x \tag{13.8}$$

for $v \in H^1(\mathbb{R}^d)$ is the energy and that $H^1(\mathbb{R}^d) \hookrightarrow L^{1+\alpha}(\mathbb{R}^d)$ because of $1 < \alpha < \alpha_c$, see (12.4). In particular, $E : H^1(\mathbb{R}^d) \to \mathbb{R}$ is real continuously differentiable by Corollary 9.3.

THEOREM 13.2. *Let $1 < \alpha < \alpha_c$, $\mu \in \{-1, 1\}$, $u_0 \in H^1(\mathbb{R}^d)$ and let u be the corresponding maximal H^1–solution of (13.1) on $J(u_0)$. We then have*

$$\|u(t)\|_2 = \|u_0\|_2 \quad \textit{and} \quad E(u(t)) = E(u_0) \quad \textit{for all } t \in J(u_0).$$

PROOF. 1) Take $u_{0,n} \in H^2(\mathbb{R}^d)$ that converge to u_0 in $H^1(\mathbb{R}^d)$. Let u_n be the H^2–solution of (13.1) with the initial value $u_{0,n}$, see Theorem 13.1. Fix a compact interval $J \subseteq J(u_0)$ with $0 \in J$. Theorem 12.6 yields that $J \subseteq J(u_{0,n})$ for all n larger than some n_0 and that $u_n(t) \to u(t)$ in $H^1(\mathbb{R}^d)$ as $n \to \infty$ for all $t \in J$. Remark 10.4 now says that $\|u_n(t)\|_2 = \|u_{0,n}\|_2$ for all $t \in J$ and $n \geq n_0$. Letting $n \to \infty$, we derive the first assertion.

2) Since the energy E is continuous on $H^1(\mathbb{R}^d)$, the second assertion will follow as the first one if we can show that $E(u(t)) = E(u_0)$ for all $t \in J(u_0)$ and $u_0 \in H^2(\mathbb{R}^d)$. Unfortunately, in Remark 10.5 we needed $u \in C^1(J(u_0), H^1(\mathbb{R}^d)) \cap C(J(u_0), H^2(\mathbb{R}^d))$ instead of the available regularity $u \in C^1(J(u_0), L^2(\mathbb{R}^d)) \cap C(J(u_0), H^2(\mathbb{R}^d))$. We handle the two terms of the energy in different ways, assuming that $u_0 \in H^2(\mathbb{R}^d)$ with corresponding H^2–solution u $J(u_0)$.

i) We show directly that the function $t \mapsto \frac{1}{2}\||\nabla u(t)|\|_2^2$ is continuously differentiable on $J(u_0)$ with derivative $-\operatorname{Re}(\Delta u|\overline{u}')_{L^2}$. In fact, let $t, t+h \in [a,b] \subseteq J(u_0)$. Integrating by parts, we compute

$$\begin{aligned} D_h &:= \tfrac{1}{2}\||\nabla u(t+h)|\|_2^2 - \tfrac{1}{2}\||\nabla u(t)|\|_2^2 + \operatorname{Re}(\Delta u(t)|\overline{u}'(t))_{L^2} h = \operatorname{Re} D_h \\ &= -\frac{1}{2}\operatorname{Re}\int_{\mathbb{R}^d} \Delta u(t+h)\overline{u}(t+h)\,\mathrm{d}x + \frac{1}{2}\operatorname{Re}\int_{\mathbb{R}^d} \Delta u(t)\overline{u}(t)\,\mathrm{d}x \\ &\quad + \operatorname{Re}\int_{\mathbb{R}^d} \Delta u(t)\overline{u}'(t)h\,\mathrm{d}x \\ &= -\frac{1}{2}\operatorname{Re}\int_{\mathbb{R}^d} (\Delta u(t+h) - \Delta u(t))(\overline{u}(t+h) - \overline{u}(t))\,\mathrm{d}x \\ &\quad - \operatorname{Re}\int_{\mathbb{R}^d} \Delta u(t)(\overline{u}(t+h) - \overline{u}(t) - \overline{u}'(t)h)\,\mathrm{d}x \end{aligned}$$

using also the symmetry of Δ on $H^2(\mathbb{R}^d)$ and $\operatorname{Re} z = \operatorname{Re}\overline{z}$ for $z \in \mathbb{C}$. Observe that $\overline{u} : [a,b] \to L^2(\mathbb{R}^d)$ is Lipschitz with constant $c := \sup_{a \leq \tau \leq b} \|u'(\tau)\|_2$. For $h \neq 0$ we thus derive

$$\begin{aligned} |\tfrac{1}{h} D_h| &\leq \tfrac{c}{2}\|\Delta u(t+h) - \Delta u(t)\|_2 + \|\Delta u(t)\|_2 \|\tfrac{1}{h}(u(t+h) - u(t)) - u'(t)\|_2 \\ &\longrightarrow 0 \quad \text{as } h \to 0. \end{aligned}$$

ii) The second summand in the energy is differentiated by means of an approximation of the nonlinearity. To this aim, we fix a function $\psi \in C_b^1(\mathbb{R})$ such that $\psi(r) = \frac{1}{1+\alpha} r^{1+\alpha}$ for $0 \leq r \leq 1$ as well as $0 \leq \psi(r) \leq \frac{1}{1+\alpha} r^{1+\alpha}$ and $0 \leq \psi'(r) \leq r^\alpha$ for all $r \geq 1$. We then set $\phi_n(z) = n^{1+\alpha}\psi(\frac{1}{n}|z|)$ for $z \in \mathbb{R}^2$ and $G_n(v) = \mu \int_{\mathbb{R}^d} \phi_n(v)\,\mathrm{d}x$ for $v \in L^2(\mathbb{R}^d)$. Lemma 9.1 shows that $G_n : L^2(\mathbb{R}^d) \to \mathbb{R}$ is real continuously differentiable with

$$G_n'(v)w = \mu \int_{\mathbb{R}^d} \nabla\phi_n(v) \cdot w\,\mathrm{d}x$$

for $v, w \in L^2(\mathbb{R}^d)$. As a result,

$$\frac{\mathrm{d}}{\mathrm{dt}}\, G_n(u(t)) = \mu \int_{\mathbb{R}^d} \nabla\phi_n(u(t)) \cdot u'(t)\, \mathrm{d}x$$

for $t \in J(u_0)$. We next note that $\phi_n(u(t))$ tends to $\frac{1}{1+\alpha}|u(t)|^{\alpha+1}$ and $\nabla\phi_n(u(t)) \cdot u'(t)$ tends to $|u(t)|^{\alpha-1}\operatorname{Re}(u(t)\overline{u}'(t))$, pointwise as $n \to \infty$. Moreover,

$$|\phi_n(z)| \le \frac{1}{1+\alpha}\frac{n^{1+\alpha}}{n^{1+\alpha}}|z|^{1+\alpha} = \frac{1}{1+\alpha}|z|^{1+\alpha}, \qquad |\nabla\phi_n(z)| \le \frac{n^{1+\alpha}}{n^\alpha}\frac{|z|}{n|z|}|z|^\alpha = |z|^\alpha$$

for $z \in \mathbb{R}^2$. Recall that $u(t) \in H^2(\mathbb{R}^d) \hookrightarrow L^{1+\alpha}(\mathbb{R}^d) \cap L^{2\alpha}(\mathbb{R}^d)$ by Sobolev's embedding (5.5) and that $u'(t) \in L^2(\mathbb{R}^d)$. With the majorants $\frac{1}{1+\alpha}|u(t)|^{\alpha+1}$ and $|u(t)|^\alpha|u'(t)|$, dominated convergence yields

$$\begin{aligned} G_n(u(t)) \;&\longrightarrow\; \frac{\mu}{1+\alpha}\int_{\mathbb{R}^d} |u(t)|^{\alpha+1}\,\mathrm{d}x, \\ \frac{\mathrm{d}}{\mathrm{dt}}\, G_n(u(t)) = \mu \int_{\mathbb{R}^d} \nabla\phi_n(u(t)) \cdot u'(t)\, \mathrm{d}x \;&\longrightarrow\; \mu \operatorname{Re}\int_{\mathbb{R}^d} |u(t)|^{\alpha-1}u(t)\overline{u}'(t)\,\mathrm{d}x \end{aligned}$$

as $n \to \infty$ for each $t \in \mathbb{R}$. Moreover, $G_n(u(t))$ and $\frac{\mathrm{d}}{\mathrm{dt}}\, G_n(u(t))$ are uniformly bounded for $n \in \mathbb{N}$ and t in compact subsets of $J(u_0)$.

iii) We next define the "approximative energy"

$$E_n(v) = \tfrac{1}{2}\||\nabla v|\|_2^2 + G_n(v)$$

for $n \in \mathbb{N}$ and $v \in H^1(\mathbb{R}^d)$. The results in i) and ii) then show that

$$\begin{aligned} E_n(u(t)) \;&\longrightarrow\; E(u(t)), \\ \frac{\mathrm{d}}{\mathrm{dt}}\, E_n(u(t)) \;&\longrightarrow\; \operatorname{Re}\int_{\mathbb{R}^d} (-\Delta u(t) + \mu|u(t)|^{\alpha-1}u(t))\overline{u}'(t)\,\mathrm{d}x \\ &= \operatorname{Re}\int_{\mathbb{R}^d} \mathrm{i}u'(t)\overline{u}'(t)\,\mathrm{d}x = 0 \end{aligned}$$

as $n \to \infty$ for all $t \in J(u_0)$, because u solves (13.1). Since $E_n(u)$ and $\frac{\mathrm{d}}{\mathrm{dt}}\, E_n(u)$ are locally bounded, the above limits also hold in $L^1(J)$ for each open bounded interval with $\overline{J} \subseteq J(u_0)$. Hence, $E(u) \in W_1^1(J)$ with vanishing derivative and $E(u)$ is constant. □

In several cases the above conservation laws allow to bound the H^1–norm of a solution. Theorem 12.6 shows that this norm must explode in finite time if we do not have global existence. This line of arguments leads to our final theorem. In assertion (a) we only use that in the defocusing case the energy plus the L^2–norm dominate the norm in $H^1(\mathbb{R}^d)$. In the focusing case the second summand of $E(u(t))$ is negative and has to be controlled by the first part of $E(u(t))$ and $\|u(t)\|_2^2$. This can be done if either α or u_0 is small.

THEOREM 13.3. *Let $\mu \in \{-1, 1\}$, $1 < \alpha < \frac{d+2}{(d-2)_+} = \alpha_c$, $u_0 \in H^1(\mathbb{R}^d)$ and let u be the corresponding maximal H^1–solution of (13.1) on $J(u_0)$. Then the following assertions hold.*

(a) If $\mu = 1$, then $J(u_0) = \mathbb{R}$ for all $u_0 \in H^1(\mathbb{R}^d)$.
(b) If $\mu = -1$ and $1 < \alpha < 1 + \frac{4}{d}$, then $J(u_0) = \mathbb{R}$ for all $u_0 \in H^1(\mathbb{R}^d)$.

(c) *There are numbers $\rho, \kappa > 0$ such that $J(u_0) = \mathbb{R}$ and $\sup_{t\in\mathbb{R}} \|u(t)\|_{1,2} \le \kappa$ if $\|u_0\|_{1,2} \le \rho$.*

PROOF. (a) If $\mu = 1$, Theorem 13.2 yields that

$$\|u(t)\|_{1,2}^2 = \|u(t)\|_2^2 + \||\nabla u(t)|\|_2^2 \le \|u(t)\|_2^2 + 2E(u(t)) = \|u_0\|_2^2 + 2E(u_0)$$

for all $t \in J(u_0)$. From the blow-up criterion in Theorem 12.6 (c) we thus deduce $J(u_0) = \mathbb{R}$.

(b) Let $1 < \alpha < 1 + \frac{4}{d}$ and $\mu = -1$. We consider $d \ge 3$, the proof for $d = 1, 2$ is similar. The proof of Sobolev's embedding even yields that

$$\|v\|_{\frac{2d}{d-2}} \le c_d \||\nabla v|\|_2 \tag{13.9}$$

for $v \in H^1(\mathbb{R}^d)$ and a constant c_d only depending on d, see (D.15) in Appendix D. To use this extra integrability of $u(t)$, we note that

$$\frac{1}{\alpha+1} = \frac{1-\theta}{2} + \frac{\theta}{\frac{2d}{d-2}} \qquad \text{for } \theta = \frac{d}{2} - \frac{d}{\alpha+1} \in (0,1).$$

The interpolation inequality (D.13) and (13.9) then imply[1]

$$\|v\|_{\alpha+1}^{\alpha+1} \le \|v\|_2^{(1-\theta)(\alpha+1)} \|v\|_{\frac{2d}{d-2}}^{\theta(\alpha+1)} \le c_d^{\theta(\alpha+1)} \|v\|_2^{\alpha+1-d\frac{\alpha-1}{2}} \||\nabla v|\|_2^{d\frac{\alpha-1}{2}}$$

for all $v \in H^1(\mathbb{R}^d)$. We have $\beta := \frac{4}{d(\alpha-1)} > 1$ by the assumption on α. Young's inequality with exponents β and β' now leads to

$$\tfrac{1}{\alpha+1}\|v\|_{\alpha+1}^{\alpha+1} \le \tfrac{1}{4}\||\nabla v|\|_2^2 + c\|v\|_2^{\beta'(\alpha+1-d\frac{\alpha-1}{2})}$$

for a constant c only depending on α and d. Denoting the last summand by $k(\|v\|_2)$, we infer from Theorem 13.2 that $\|u(t)\|_2^2 = \|u_0\|_2^2$ and

$$\begin{aligned} E(u_0) = E(u(t)) &= \tfrac{1}{2}\||\nabla u(t)|\|_2^2 - \tfrac{1}{\alpha+1}\|u(t)\|_{\alpha+1}^{\alpha+1} \\ &\ge \tfrac{1}{4}\||\nabla u(t)|\|_2^2 - k(\|u(t)\|_2) = \tfrac{1}{4}\||\nabla u(t)|\|_2^2 - k(\|u_0\|_2) \end{aligned}$$

for all $t \in J(u_0)$. Therefore, $\|u(t)\|_{1,2}^2 \le 4E(u_0) + 4k(\|u_0\|_2) + \|u_0\|_2^2$ for all $t \in J(u_0)$, and as before it follows that $J(u_0) = \mathbb{R}$.

(c) Let $\mu \in \{-1, 1\}$. This part relies on the observation that by Sobolev's embedding the "nonlinear" part of the energy can be bounded by

$$\tfrac{1}{1+\alpha}\|u(t)\|_{\alpha+1}^{\alpha+1} \le c\|u(t)\|_{1,2}^{\alpha-1}\|u(t)\|_{1,2}^2.$$

So if $\|u_0\|_{1,2}$ is small, one can absorb this term by the other part of the energy and $\|u(t)\|_2^2$ as long as $\|u(t)\|_{1,2}$ stays under a certain constant γ. Choosing a suitable $\gamma > \|u_0\|_{1,2}$, one then sees by a contradiction argument that actually $\|u(t)\|_{1,2} \le \gamma$ for all t and the assertion will follow.

To make this precise, we note that the conservation laws and Sobolev's embedding yield

$$\begin{aligned} \tfrac{1}{2}\|u(t)\|_{1,2}^2 &= \tfrac{1}{2}\|u_0\|_2^2 + E(u(t)) - \tfrac{\mu}{\alpha+1}\|u(t)\|_{\alpha+1}^{\alpha+1} \\ &\le \tfrac{1}{2}\|u_0\|_2^2 + E(u_0) + c_0\|u(t)\|_{1,2}^{\alpha-1}\|u(t)\|_{1,2}^2, \end{aligned} \tag{13.10}$$

[1] Such estimates are called Gagliardo-Nirenberg inequalities.

where $c_0 = C_{\mathrm{So}}^{\alpha+1}\frac{1}{1+\alpha}$ with Sobolev's constant C_{So} from (12.4). We set $\gamma = (4c_0)^{\frac{1}{1-\alpha}}$ and take any $\rho \in (0,\gamma)$. Let $\|u_0\|_{1,2} \le \rho$. We now define

$$t_0 = \sup\left\{t \in (0, t^+(u_0)) \,\middle|\, \|u(s)\|_{1,2} \le \gamma \text{ for all } s \in [0,t]\right\}. \tag{13.11}$$

Observe that $t_0 \in (0, t^+(u_0)]$. Let $0 \le t < t_0$. The estimate (13.10), the choice of γ and Sobolev's embedding (12.4) next imply that

$$\begin{aligned}\tfrac{1}{2}\|u(t)\|_{1,2}^2 &\le \tfrac{1}{2}\|u_0\|_2^2 + E(u_0) + \tfrac{1}{4}\|u(t)\|_{1,2}^2,\\ \|u(t)\|_{1,2}^2 &\le 2\|u_0\|_2^2 + 2\||\nabla u_0|\|_2^2 + \tfrac{4}{\alpha+1}\|u_0\|_{\alpha+1}^{\alpha+1} \le c_1(\rho^2 + \rho^{\alpha+1}),\end{aligned} \tag{13.12}$$

where $c_1 = \max\{2, \frac{4}{1+\alpha}C_{\mathrm{So}}^{1+\alpha}\}$. We now fix $\rho \in (0,\gamma)$ such that $c_1(\rho^2 + \rho^{\alpha+1}) \le \frac{\gamma^2}{4}$. If $t_0 < t^+(u_0)$, then $u(t) \to u(t_0)$ in $H^1(\mathbb{R}^d)$ as $t \to t_0^-$ so that (13.11) yields $\|u(t_0)\|_{1,2} = \gamma$, but (13.12) leads to the contradiction $\|u(t_0)\|_{1,2} \le \frac{\gamma}{2}$. Hence, $t_0 = t^+(u_0)$ and from (13.12) we derive that $\|u(t)\|_{1,2} \le \frac{\gamma}{2} =: \kappa$ for all $t \in [0, t^+(u_0))$. Theorem 12.6 (c) now implies $t^+(u_0) = \infty$, as asserted. Similarly one treats negative times. □

The proof of Theorem 13.3 (c) leads to a corollary concerning stability.

Corollary 13.4. *Let $\mu \in \{-1,1\}$ and $1 < \alpha < \alpha_c$. Then the solution $u_* = 0$ of* (13.3) *is* Lyapunov stable *in $H^1(\mathbb{R}^d)$, i.e.:*

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, u_0 \in \overline{B}_{H^1}(0,\delta):\ J(u_0) = \mathbb{R} \text{ and } \|u(t;u_0)\|_{1,2} \le \varepsilon \text{ for all } t \in \mathbb{R}.$$

Proof. Let c_0 and c_1 be given as in the proof of Theorem 13.3 (c). Set $\varepsilon_0 = \frac{1}{2}(4c_0)^{\frac{1}{1-\alpha}}$ and take any $\varepsilon \in (0, \varepsilon_0]$. Choose $\delta \in (0, 2\varepsilon)$ such that $c_1(\delta^2 + \delta^{\rho+1}) \le \varepsilon^2$. Take $u_0 \in \overline{B}_{H^1}(0,\delta)$. Define t_0 as in (13.11) with γ replaced by 2ε. The estimate (13.12) still holds and implies the assertion. □

We conclude this lecture with a few remarks about further results on global existence and blow-up.

Global existence holds in the defocusing case $\mu = 1$ also if $\alpha = \frac{d+2}{d-2}$ and $d \ge 3$. This result is far beyond the scope of these lectures, see Chapter 5 of [**Tao06**] for an extended survey.

For $\alpha \in [1 + \frac{4}{d}, \alpha_c)$ and $\mu = -1$, Theorem 6.5.4 in [**Caz03**] establishes blow-up in (13.1) if $E(u_0) < 0$ and $|x|u_0 \in L^2(\mathbb{R}^d)$.[2] (This additional integrability is not needed if u_0 is spherically symmetric by Theorem 6.5.10 in [**Caz03**].) One could guess that a negative initial energy is necessary for blow-up. This is not the case as Remark 6.5.8 in [**Caz03**] gives a blow-up solution with $E(u_0) > 0$.

In Example 10.3 we have seen a blow-up solution for $\alpha = 1 + \frac{4}{d}$ and $\mu = -1$. Denote its initial value by φ and consider (13.1) with $\mu = -1$ and $\alpha = 1 + \frac{4}{d}$. If $\|u_0\|_2 < \|\varphi\|_2$, then $J(u_0) = \mathbb{R}$ by Theorem 6.6.1 in [**Caz03**]. So in the borderline case $\alpha = 1 + \frac{4}{d}$, where global existence starts to fail for $\mu = -1$, one has a precise threshold for the occurrence of blow-up solutions.

Let $\mu = -1$ and $\max\{1, \alpha_0\} < \alpha < \alpha_c$, where $\alpha_0 > 0$ satisfies $d\alpha_0^2 + (d-2)\alpha_0 = 4$. Let $u_0 \in H^1(\mathbb{R}^d)$ with $|x|u_0 \in L^2(\mathbb{R}^d)$. Set $\varphi_b(x) = \mathrm{e}^{\mathrm{i}b|x|^2}u_0(x)$ for $b > 0$ and

[2] Recall that in Proposition 9.6 we showed blow-up for a certain nonlinear wave equation if the initial energy is negative and another condition holds.

$x \in \mathbb{R}^d$. Then $\varphi_b \in H^1(\mathbb{R}^d)$ and there is a number $b_0 > 0$ such that $t^+(\varphi_b) = \infty$ for all $b \geq b_0$. (See Theorem 6.3.4 of [**Caz03**].) Hence, one has global existence (to the right) if the initial value is rapidly oscillating.

Exercises

EXERCISE 13.1. Let $u_0 \in H^1(\mathbb{R}^d)$ and u be the corresponding H^1–solution of (13.1). Define its "momentum" by

$$p_j(t) = \operatorname{Im} \int_{\mathbb{R}^d} \overline{u}(t) \partial_j u(t) \, \mathrm{d}x, \quad t \in J(u_0), \ j \in \{1, \dots, d\}.$$

Show that $p_j(t) = p_j(0)$ for all $t \in J(u_0)$. (Hint: Consider first $u_0 \in H^2(\mathbb{R}^d)$.)

EXERCISE 13.2. Let u be an H^2–solution of (13.1) on $J = [0, b)$, where $b \in (0, \infty]$ and $\alpha = 1 + \frac{4}{d}$. Let $\gamma \geq \frac{1}{b}$ and set

$$u_\gamma(t, x) = (1 + \gamma t)^{-\frac{d}{2}} \mathrm{e}^{\mathrm{i}\frac{\gamma |x|^2}{4(1+\gamma t)}} u\Big(\tfrac{t}{1+\gamma t}, \tfrac{1}{1+\gamma t} x\Big)$$

for $t \in \mathbb{R}_+$ and $x \in \mathbb{R}^d$. Show that u_γ satisfies (13.1) with $u_\gamma(0) = \mathrm{e}^{\mathrm{i}\frac{\gamma}{4}|\cdot|^2} u(0)$. Further show that

(i) $\|u_\gamma(t)\|_2 = \|u(s)\|_2, \ \ \|u_\gamma(t)\|_{1+\alpha} = (1 + \gamma t)^{-\frac{2}{1+\alpha}} \|u(s)\|_{1+\alpha}$,
(ii) $\||\nabla u_\gamma(t)|\|_2 = \frac{1}{2+2\gamma t} \||(\mathrm{i}\gamma x + 2\nabla) u(s)|\|_2$

for all $t \geq 0$, where $s = s(t) = t(1 + \gamma t)^{-1}$ and we assume that $|x| u(s) \in L^2(\mathbb{R}^d)$ in (ii).

EXERCISE 13.3. Let $d \geq 3$, $\alpha = \frac{d+2}{d-2}$ and $\mu \in \{-1, 1\}$. Show that there is a radius $\rho > 0$ such that (13.1) has an H^1–solution in $\mathcal{E}_1(\mathbb{R})$ for all $u_0 \in \overline{B}_{H^1}(0, \rho)$. (Hint: Use Strichartz' estimates for the endpoint case $(q, p) = (2, \frac{2d}{d-2})$.)

LECTURE 14

The asymptotic behavior in the defocusing case

In the previous lecture we have seen that the defocusing nonlinear Schrödinger equation

$$u'(t) = \mathrm{i}\Delta u(t) - \mathrm{i}|u(t)|^{\alpha-1}u(t), \quad t \in J, \qquad u(0) = u_0, \tag{14.1}$$

is globally solvable in the subcritical case $1 < \alpha < \frac{d+2}{(d-2)_+} = \alpha_c$. So for each $u_0 \in H^1(\mathbb{R}^d)$ there is a unique solution $u \in C^1(\mathbb{R}, H^{-1}(\mathbb{R}^d)) \cap C(\mathbb{R}, H^1(\mathbb{R}^d))$ of (14.1). We next inquire how $u(t)$ behaves as t tends to $\pm\infty$.

It turns out that the problem (14.1) has a similar long-term behavior as the free linear Schrödinger equation $v' = \mathrm{i}\Delta v$. From Corollary 10.10 we know that the free Schrödinger group $T(\cdot)$ satisfies

$$\|T(t)v_0\|_r \le c|t|^{\frac{d}{r}-\frac{d}{2}}\|v_0\|_{r'}$$

for all $r \in (2, \infty]$, $t \neq 0$ and $v_0 \in L^{r'}(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. In our main Theorem 14.4 we show that the solution of (14.1) decays in a similar way as $t \to \infty$ if $u_0 \in H^1(\mathbb{R}^d)$ satisfies $|x|u_0 \in L^2(\mathbb{R}^d)$ and $r > 2$ is not too large.

Recall that for $r = 2$ there is no decay since $\|u(t)\|_2 = \|u_0\|_2$ for all $t \in \mathbb{R}$ by Theorem 13.2. In view of this conservation law in $L^2(\mathbb{R}^d)$ and the decay in $L^r(\mathbb{R}^d)$ for $r > 2$, the solution of (14.1) should spread out in space as time evolves. This behavior is in accordance with the dispersive phenomena we discussed a bit in Lecture 10.

Actually, for $r = 2$ one can show an even closer relationship between (14.1) and the free linear Schrödinger equation. In scattering theory one constructs $u_\pm \in H^1(\mathbb{R}^d)$ such that $u(t) - T(t)u_\pm \to 0$ in $H^1(\mathbb{R}^d)$ as $t \to \pm\infty$. We mention a few scattering results at the end of the lecture, where we also talk a bit about the long-term behavior in the focusing case.

The proof of the convergence result in L^r depends on explicit formulas for the first and second derivatives of the quantity $\||x|u(t)\|_2^2$ for an H^1–solution u of (13.1) whose initial value $u_0 \in H^1(\mathbb{R}^d)$ satisfies $|x|u_0 \in L^2(\mathbb{R}^d)$. We note that these formulas hold in the focusing case, too. They are established in Proposition 14.1. We remark that they are also crucial for the blow-up results stated at the end of the last lecture. The proof of Proposition 14.1 depends on lengthy calculations and delicate approximation arguments. As shown in Corollary 14.3, these formulas imply an expression for the energy of the functions $v(t) = \mathrm{e}^{-\frac{\mathrm{i}|\cdot|^2}{4t}}u(t)$ for $t \neq 0$. In the defocusing case this expression leads to the desired decay estimate when combined with the conservation laws, the Gagliardo-Nirenberg inequality and the Gronwall lemma, see Theorem 14.4.

The results of this lecture were established by J. Ginibre and G. Velo in e.g. [**GV79**] for more general nonlinearities. Our presentation follows (parts of) Sections 6.5, 7.2 and 7.3 of [**Caz03**].

Throughout we use the spaces

$$L^2_\ell(\mathbb{R}^d) = \Big\{\varphi \in L^2(\mathbb{R}^d) \,\big|\, |\ell|\varphi \in L^2(\mathbb{R}^d)\Big\} \quad \text{and} \quad H^1_\ell(\mathbb{R}^d) = L^2_\ell(\mathbb{R}^d) \cap H^1(\mathbb{R}^d)$$

where we set $\ell(x) = x$ for $x \in \mathbb{R}^d$. They are Hilbert spaces when endowed with the norms given by

$$\|\varphi\|_{2,\ell}^2 = \||\ell|\varphi\|_2^2 + \|\varphi\|_2^2 \quad \text{and} \quad \|\varphi\|_{1,2,\ell}^2 = \|\varphi\|_{2,\ell}^2 + \|\varphi\|_{1,2}^2\,,$$

respectively. Observe that $H^1(\mathbb{R}^d)$ is not contained in $L^2_\ell(\mathbb{R}^d)$.

In our first proposition we show that for $u_0 \in H^1_\ell(\mathbb{R}^d)$ the solution u of (13.1) is continuous in $L^2_\ell(\mathbb{R}^d)$ and we calculate the derivatives of $t \mapsto \|u(t)\|_{2,\ell}^2$.

PROPOSITION 14.1. *Let $u_0 \in H^1_\ell(\mathbb{R}^d)$, $\alpha \in (1, \alpha_c)$ and $\mu \in \{-1, 1\}$. Then the maximal H^1–solution u of the nonlinear Schrödinger equation (13.1) belongs to $C(J(u_0), L^2_\ell(\mathbb{R}^d))$. Moreover, the function ϕ_u given by $\phi_u(t) = \||\ell|u(t)\|_2^2$, $t \in J(u_0)$, is twice continuously differentiable and satisfies*

$$\phi_u'(t) = 4\,\mathrm{Im}\int_{\mathbb{R}^d} \overline{u}(t)\ell \cdot \nabla u(t)\,\mathrm{d}x, \tag{14.2}$$

$$\phi_u''(t) = 16E(u_0) + \big(4d - 8\tfrac{d+2}{\alpha+1}\big)\mu\|u(t)\|_{\alpha+1}^{\alpha+1} \tag{14.3}$$

for all $t \in J(u_0)$ and the initial energy $E(u_0) = \frac{1}{2}\||\nabla u_0|\|_2^2 + \frac{\mu}{\alpha+1}\|u_0\|_{\alpha+1}^{\alpha+1}$.

Observe that the expression for ϕ_u'' is rather simple and depends on $u(t)$ continuously with respect to the H^1–norm. If the solution u is sufficiently smooth and decays rapidly enough, one can deduce the assertions of Proposition 14.1 in a direct but tedious way, using (13.1) and integration by parts. To perform such computations in a rigorous way, we need several approximation arguments. Already the weight $|\ell|$ leads to integrability problems at infinity. To overcome them, we use the function $\gamma_\varepsilon(x) = \mathrm{e}^{-\varepsilon|x|^2}$ for $x \in \mathbb{R}^d$ and $\varepsilon > 0$.

PROOF OF PROPOSITION 14.1. 1) We first show that $u \in C(J(u_0), L^2_\ell(\mathbb{R}^d))$, $\phi_u \in C^1(J(u_0))$ and that (14.2) holds. We consider any time interval $J = [0, b] \subseteq J(u_0)$. Negative times are treated in the same way. For $\varepsilon > 0$ we define

$$\phi_{u,\varepsilon}(t) = \|\gamma_\varepsilon|\ell|u(t)\|_2^2 = \int_{\mathbb{R}^d} \gamma_\varepsilon^2|\ell|^2|u(t)|^2\,\mathrm{d}x, \quad t \in J.$$

We first consider $u_0 \in H^2(\mathbb{R}^d)$ so that $u \in C^1(J, L^2(\mathbb{R}^d)) \cap C(J, H^2(\mathbb{R}^d))$ by Theorem 13.1. It is then clear that $\phi_{u,\varepsilon} \in C^1(J)$. Using (13.1) and integrating by parts, we derive

$$\begin{aligned}\phi'_{u,\varepsilon}(t) &= 2\,\mathrm{Re}\int_{\mathbb{R}^d} \gamma_\varepsilon^2|\ell|^2\overline{u}(t)u'(t)\,\mathrm{d}x = 2\,\mathrm{Re}\,\mathrm{i}\int_{\mathbb{R}^d}\gamma_\varepsilon^2|\ell|^2\overline{u}(t)(\Delta u(t)-\mu|u(t)|^{\alpha-1}u(t))\,\mathrm{d}x\\ &= -2\,\mathrm{Re}\,\mathrm{i}\int_{\mathbb{R}^d}\Big(\nabla(\gamma_\varepsilon^2|\ell|^2\overline{u}(t))\cdot\nabla u(t)+\mu\gamma_\varepsilon^2|\ell|^2|u(t)|^{\alpha+1}\Big)\,\mathrm{d}x\\ &= -2\,\mathrm{Re}\,\mathrm{i}\int_{\mathbb{R}^d}\Big(-4\varepsilon\ell\cdot\nabla u(t)\gamma_\varepsilon^2|\ell|^2\overline{u}(t)+2\ell\cdot\nabla u(t)\gamma_\varepsilon^2\overline{u}(t)+\gamma_\varepsilon^2|\ell|^2|\nabla u(t)|^2\Big)\,\mathrm{d}x\\ &= 4\,\mathrm{Im}\int_{\mathbb{R}^d}\gamma_\varepsilon^2(1-2\varepsilon|\ell|^2)\overline{u}(t)\ell\cdot\nabla u(t)\,\mathrm{d}x\end{aligned}$$

for $t\in J$. By integration it follows

$$\phi_{u,\varepsilon}(t) = \|\gamma_\varepsilon|\ell|u_0\|_2^2 + 4\,\mathrm{Im}\int_0^t\int_{\mathbb{R}^d}\gamma_\varepsilon(1-2\varepsilon|\ell|^2)\gamma_\varepsilon\overline{u}(s)\ell\cdot\nabla u(s)\,\mathrm{d}x\,\mathrm{d}s. \quad (14.4)$$

Consider now $u_0\in H^1_\ell(\mathbb{R}^d)$ and approximate it by $u_{0,n}\in H^2(\mathbb{R}^d)$ in $H^1(\mathbb{R}^d)$ with corresponding solutions u_n. Theorem 12.6 shows that $J(u_{0,n})\subseteq J$ for all sufficiently large n and that $u_n\to u$ in $C(J,H^1(\mathbb{R}^d))$ as $n\to\infty$. Hence, (14.4) holds for $u_0\in H^1_\ell(\mathbb{R}^d)$ by approximation. Observe that $|\gamma_\varepsilon(1-2\varepsilon|\ell|^2)|\le 1$ for all $\varepsilon>0$ and $x\in\mathbb{R}^d$. Due to Hölder's and Young's inequalities, (14.4) leads to

$$\begin{aligned}\phi_{u,\varepsilon}(t) &\le \||\ell|u_0\|_2^2 + 4\int_0^t\phi_{u,\varepsilon}(s)^{\frac{1}{2}}\||\nabla u(s)|\|_2\,\mathrm{d}s\\ &\le \||\ell|u_0\|_2^2 + 2\int_0^t\||\nabla u(s)|\|_2^2\,\mathrm{d}s + 2\int_0^t\phi_{u,\varepsilon}(s)\,\mathrm{d}s\end{aligned}$$

for $t\in J$. From $u\in C(J,H^1(\mathbb{R}^d))$ and Gronwall's lemma we deduce that $\sup_{\varepsilon>0}\|\phi_{u,\varepsilon}\|_\infty =: C < \infty$. As $\varepsilon\to 0$, Fatou's lemma yields $\phi_u(t) = \||\ell|u(t)\|_2^2 \le C$ for all $t\in J$. For later use we note that C only depends on $\|u_0\|_{2,\ell}$, $\||\nabla u|\|_{C(J,L^2)}$ and b. By means of the majorants $|\ell u(t)|^2$ and $4|\ell u||\nabla u|$ we can now let $\varepsilon\to 0$ in (14.4) to obtain

$$\phi_u(t) = \||\ell|u(t)\|_2^2 = \||\ell|u_0\|_2^2 + 4\,\mathrm{Im}\int_0^t\int_{\mathbb{R}^d}\overline{u}(s)\ell\cdot\nabla u(s)\,\mathrm{d}x\,\mathrm{d}s \quad (14.5)$$

for $t\in J$ and $u_0\in H^1_\ell(\mathbb{R}^d)$. Using again the second majorant, one sees that the right-hand side of (14.5) depends continuously on t, so that the norms $\||\ell|u(t)\|_2$ converge to $\||\ell|u(t_0)\|_2$ as $t\to t_0$ in J. Moreover, the functions $|\ell|u(t)$ tend pointwise a.e. on $\mathbb{R}^d$ to $|\ell|u(t_0)$ as $t\to t_0$. Hence, a result by Riesz implies that the map $t\mapsto|\ell|u(t)\in L^2(\mathbb{R}^d)$ is continuous, see Lemma 1.32 in [**Kal02**]. We can thus differentiate (14.5) with respect to $t\in J$ and deduce (14.2). This formula implies the continuity of ϕ'_u.

2) We still have to show that $\phi_u\in C^2(J(u_0))$ and (14.3) holds. As a first step, we consider functions $v\in C(J,H^1(\mathbb{R}^d))\cap C^1(J,L^2(\mathbb{R}^d))$ and define

$$\psi_{v,\varepsilon}(t) = \mathrm{Im}\int_{\mathbb{R}^d}\gamma_\varepsilon\overline{v}(t)\ell\cdot\nabla v(t)\,\mathrm{d}x \quad (14.6)$$

for $\varepsilon>0$ and $t\in J$, where $J\subseteq\mathbb{R}$ is a compact interval containing 0. Later, the functions $4\psi_{u,\varepsilon}$ shall approximate ϕ'_u as $\varepsilon\to 0$. We want to show that

$$\psi_{v,\varepsilon}(t) = \psi_{v,\varepsilon}(0) - \mathrm{Im}\int_0^t\int_{\mathbb{R}^d}v'(s)\Big[2\gamma_\varepsilon\ell\cdot\nabla\overline{v}(s)+(d\gamma_\varepsilon+\ell\cdot\nabla\gamma_\varepsilon)\overline{v}(s)\Big]\,\mathrm{d}x\,\mathrm{d}s \quad (14.7)$$

for $t \in J$ and each fixed $\varepsilon > 0$. We see in Lemma 14.2 below that the space $C^1(J, H^1(\mathbb{R}^d))$ is dense in $C(J, H^1(\mathbb{R}^d)) \cap C^1(J, L^2(\mathbb{R}^d))$ with respect to the norm given by $\max\{\|v\|_{C(J,H^1)}, \|v\|_{C^1(J,L^2)}\}$. By continuity, it thus suffices to show (14.7) for $v \in C^1(J, H^1(\mathbb{R}^d))$. For such functions, $\psi_{v,\varepsilon}$ is continuously differentiable and

$$\begin{aligned} \psi'_{v,\varepsilon}(t) &= \operatorname{Im} \int_{\mathbb{R}^d} \Big(\gamma_\varepsilon \overline{v}'(t) \ell \cdot \nabla v(t) + \gamma_\varepsilon \overline{v}(t) \ell \cdot \nabla v'(t)\Big) \,\mathrm{d}x \\ &= -\operatorname{Im} \int_{\mathbb{R}^d} v'(t) \Big[\gamma_\varepsilon \ell \cdot \nabla \overline{v}(t) + \operatorname{div}(\overline{v}(t)\gamma_\varepsilon \ell)\Big] \,\mathrm{d}x \\ &= -\operatorname{Im} \int_{\mathbb{R}^d} v'(t) \Big[2\gamma_\varepsilon \ell \cdot \nabla \overline{v}(t) + (d\gamma_\varepsilon + \ell \cdot \nabla \gamma_\varepsilon)\overline{v}(t)\Big] \,\mathrm{d}x, \end{aligned} \tag{14.8}$$

where we used $\operatorname{Im} \overline{z} = -\operatorname{Im} z$ for $z \in \mathbb{C}$ and Gauß' formula (5.4). By integration we arrive at (14.7).

3) Temporarily we assume that $u_0 \in H^2(\mathbb{R}^d) \cap L^2_\ell(\mathbb{R}^d)$. The solution u then belongs to $C^1(J(u_0), L^2(\mathbb{R}^d)) \cap C(J(u_0), H^2(\mathbb{R}^d))$ by Theorem 13.1. We define $\psi_{u,\varepsilon}$ by (14.6). Equation (14.7) yields that $\psi_{u,\varepsilon} \in C^1(J(u_0))$ and that $\psi'_{u,\varepsilon}$ is given by (14.8) with v replaced by u. Inserting (13.1) into (14.8) for u, we infer

$$\psi'_{u,\varepsilon}(t) = -\operatorname{Re} \int_{\mathbb{R}^d} (\Delta u(t) - \mu |u(t)|^{\alpha-1} u(t)) \Big[2\gamma_\varepsilon \ell \cdot \nabla \overline{u}(t) + (d\gamma_\varepsilon + \ell \cdot \nabla \gamma_\varepsilon)\overline{u}(t)\Big] \,\mathrm{d}x. \tag{14.9}$$

We first consider the term in (14.9) involving Δu. Integrating by parts twice, we compute

$$\begin{aligned} &-\operatorname{Re} \int_{\mathbb{R}^d} \Delta u(t) \Big[2\gamma_\varepsilon \ell \cdot \nabla \overline{u}(t) + (d\gamma_\varepsilon + \ell \cdot \nabla \gamma_\varepsilon)\overline{u}(t)\Big] \,\mathrm{d}x \\ &\quad = \operatorname{Re} \int_{\mathbb{R}^d} 2(\nabla u(t) \cdot \nabla \gamma_\varepsilon)(\ell \cdot \nabla \overline{u}(t)) \,\mathrm{d}x \\ &\qquad + \int_{\mathbb{R}^d} (2\gamma_\varepsilon |\nabla u(t)|^2 + \gamma_\varepsilon \ell \cdot \nabla |\nabla u(t)|^2) \,\mathrm{d}x \\ &\qquad + \operatorname{Re} \int_{\mathbb{R}^d} \Big(d \nabla u(t) \cdot \nabla \gamma_\varepsilon + \nabla u(t) \cdot \nabla \gamma_\varepsilon + \nabla u(t) \cdot (D^2 \gamma_\varepsilon)\ell\Big) \overline{u}(t) \,\mathrm{d}x \\ &\qquad + \int_{\mathbb{R}^d} (d\gamma_\varepsilon + \ell \cdot \nabla \gamma_\varepsilon) |\nabla u(t)|^2 \,\mathrm{d}x \\ &\quad = \operatorname{Re} 2 \int_{\mathbb{R}^d} (\nabla u(t) \cdot \nabla \gamma_\varepsilon)(\ell \cdot \nabla \overline{u}(t)) \,\mathrm{d}x + 2 \int_{\mathbb{R}^d} \gamma_\varepsilon |\nabla u(t)|^2 \,\mathrm{d}x \qquad (14.10) \\ &\qquad + (d+1) \operatorname{Re} \int_{\mathbb{R}^d} \overline{u}(t) \nabla u(t) \cdot \nabla \gamma_\varepsilon \,\mathrm{d}x + \operatorname{Re} \int_{\mathbb{R}^d} \overline{u}(t) \nabla u(t) \cdot (D^2 \gamma_\varepsilon)\ell \,\mathrm{d}x, \end{aligned}$$

where $D^2 \gamma_\varepsilon$ is the Hessian matrix. The other terms in (14.9) can be written as

$$\begin{aligned} &\operatorname{Re} \int_{\mathbb{R}^d} \mu |u(t)|^{\alpha-1} u(t) \Big[2\gamma_\varepsilon \ell \cdot \nabla \overline{u}(t) + (d\gamma_\varepsilon + \ell \cdot \nabla \gamma_\varepsilon)\overline{u}(t)\Big] \,\mathrm{d}x \\ &= 2\mu \int_{\mathbb{R}^d} |u(t)|^{\alpha-1} \gamma_\varepsilon \operatorname{Re}(\ell \cdot \nabla u(t) \overline{u}(t)) \,\mathrm{d}x + \mu \int_{\mathbb{R}^d} |u(t)|^{\alpha+1} (d\gamma_\varepsilon + \ell \cdot \nabla \gamma_\varepsilon) \,\mathrm{d}x. \\ &= \frac{2\mu}{\alpha+1} \int_{\mathbb{R}^d} \gamma_\varepsilon \ell \cdot \nabla |u(t)|^{\alpha+1} \,\mathrm{d}x + \mu \int_{\mathbb{R}^d} |u(t)|^{\alpha+1} (d\gamma_\varepsilon + \ell \cdot \nabla \gamma_\varepsilon) \,\mathrm{d}x. \end{aligned}$$

$$= -\frac{2\mu}{\alpha+1}\int_{\mathbb{R}^d}(\ell\cdot\nabla\gamma_\varepsilon + d\gamma_\varepsilon)|u(t)|^{\alpha+1}\,\mathrm{d}x + \mu\int_{\mathbb{R}^d}|u(t)|^{\alpha+1}(d\gamma_\varepsilon + \ell\cdot\nabla\gamma_\varepsilon)\,\mathrm{d}x.$$
$$= \frac{\alpha-1}{\alpha+1}\mu\int_{\mathbb{R}^d}(\ell\cdot\nabla\gamma_\varepsilon + d\gamma_\varepsilon)|u(t)|^{\alpha+1}\,\mathrm{d}x, \tag{14.11}$$

where we integrated by parts. Equations (14.9), (14.10) and (14.11) now imply

$$\begin{aligned}\psi'_{u,\varepsilon}(t) = \int_{\mathbb{R}^d}&\gamma_\varepsilon(2|\nabla u(t)|^2 + \mu d\tfrac{\alpha-1}{\alpha+1}|u(t)|^{\alpha+1})\,\mathrm{d}x\\ &+ \mathrm{Re}\int_{\mathbb{R}^d}((d+1)\overline{u}(t) + 2\ell\cdot\nabla\overline{u}(t))\nabla u(t)\cdot\nabla\gamma_\varepsilon\,\mathrm{d}x\\ &+ \mathrm{Re}\int_{\mathbb{R}^d}\overline{u}(t)\nabla u(t)\cdot(D^2\gamma_\varepsilon)\ell\,\mathrm{d}x\\ &+ \frac{\alpha-1}{\alpha+1}\mu\int_{\mathbb{R}^d}|u(t)|^{\alpha+1}\ell\cdot\nabla\gamma_\varepsilon\,\mathrm{d}x\\ &=: S_{1,\varepsilon}(t) + S_{2,\varepsilon}(t) + S_{3,\varepsilon}(t) + S_{4,\varepsilon}(t), \qquad t\in J(u_0).\end{aligned} \tag{14.12}$$

4) We next take the limit $\varepsilon\to 0$ in (14.12). Dominated convergence yields

$$\begin{aligned}S_{1,\varepsilon}(t) \;\longrightarrow\; &2\int_{\mathbb{R}^d}|\nabla u(t)|^2\,\mathrm{d}x + \frac{d(\alpha-1)}{\alpha+1}\mu\int_{\mathbb{R}^d}|u(t)|^{\alpha+1}\,\mathrm{d}x\\ &= 4E(u(t)) + \frac{d(\alpha-1)-4}{\alpha+1}\mu\int_{\mathbb{R}^d}|u(t)|^{\alpha+1}\,\mathrm{d}x\\ &= 4E(u_0) + \Big(d - \frac{2d+4}{\alpha+1}\Big)\mu\int_{\mathbb{R}^d}|u(t)|^{\alpha+1}\,\mathrm{d}x =: \chi_u(t)\end{aligned} \tag{14.13}$$

as $\varepsilon\to 0$, where we use that $E(u(t)) = E(u_0)$ by Theorem 13.2. Since $4\psi'_{u,\varepsilon}$ should tend to ϕ''_u, we expect that (14.3) follows from (14.12) and (14.13) if we can show that $S_{2,\varepsilon}$, $S_{3,\varepsilon}$ and $S_{4,\varepsilon}$ tend to 0 as $\varepsilon\to 0$. To this aim, observe that

$$\nabla\gamma_\varepsilon(x) = -2\varepsilon\mathrm{e}^{-\varepsilon|x|^2}x, \qquad |D^2\gamma_\varepsilon(x)| \le c(\varepsilon+\varepsilon^2|x|^2)\mathrm{e}^{-\varepsilon|x|^2} \le c(1+\mathrm{e}^{-1})\varepsilon \le 2c\varepsilon$$

for $\varepsilon>0$, $x\in\mathbb{R}^d$ and a constant c. Hölder's inequality now implies that

$$\begin{aligned}|S_{2,\varepsilon}(t)| &\le 2(d+1)\varepsilon\||\ell|u(t)\|_2\||\nabla u(t)|\|_2 + 4\int_{\mathbb{R}^d}\varepsilon|x|^2\mathrm{e}^{-\varepsilon|x|^2}|\nabla u(t)|^2\,\mathrm{d}x,\\ |S_{3,\varepsilon}(t)| &\le 2c\varepsilon\||\ell|u(t)\|_2\||\nabla u(t)|\|_2,\\ |S_{4,\varepsilon}(t)| &\le 2\int_{\mathbb{R}^d}\varepsilon|x|^2\mathrm{e}^{-\varepsilon|x|^2}|u(t)|^{\alpha+1}\,\mathrm{d}x.\end{aligned} \tag{14.14}$$

All terms tend to 0 as $\varepsilon\to 0$, by dominated convergence with majorants $4|\nabla u(t)|^2$ and $2|u(t)|^{\alpha+1}$. If we integrate the terms $S_{j,\varepsilon}$ over bounded open intervals J with $\overline{J}\subseteq J(u_0)$, we obtain the convergence in (14.13) and (14.14) in $L^1(J)$ (and not only pointwise) since $u\in C(\overline{J}, L^2_\ell(\mathbb{R}^d))$ and $u\in C(\overline{J}, H^1(\mathbb{R}^d))\hookrightarrow C(\overline{J}, L^{1+\alpha}(\mathbb{R}^d))$. Similarly, one sees that $4\psi_{u,\varepsilon}$ tends to ϕ'_u in $L^1(J)$ as $\varepsilon\to 0$ (recall (14.6)). As a result, $\phi'_u\in W^1_1(J)$ and $\phi''_u = 4\chi_u$. Since χ_u is continuous, it follows that $\phi_u\in C^2(J(u_0))$ and ϕ''_u is given by (14.3), provided that $u_0\in H^2(\mathbb{R}^d)\cap L^2_\ell(\mathbb{R}^d)$.

5) It remains to extend the results from step 4) from $u_0\in H^2(\mathbb{R}^d)\cap L^2_\ell(\mathbb{R}^d)$ to $u_0\in H^1_\ell(\mathbb{R}^d)$. Let $u_0\in H^1_\ell(\mathbb{R}^d)$. We first approximate u_0 in $H^1_\ell(\mathbb{R}^d)$ by $\varphi_n\in C_c^\infty(\mathbb{R}^d)$.

To this aim, as in step 1) of the proof of Theorem D.13, we choose functions $\phi_n \in C_c^\infty(\mathbb{R}^d)$ with $0 \le \phi_n \le 1$ and $\|\partial_j \phi_n\|_\infty \le \frac{c}{n}$ for all $n \in \mathbb{N}$ and $\phi_n \to 1$ pointwise on $\mathbb{R}^d$ as $n \to \infty$. Hence, $\phi_n u_0 \in H_\ell^1(\mathbb{R}^d)$ has compact support and $\phi_n u_0 \to u_0$ in $H_\ell^1(\mathbb{R}^d)$ by dominated convergence. Let $\varepsilon > 0$. We can thus find a function $v \in H_\ell^1(\mathbb{R}^d)$ with compact support such that $\|u_0 - v\|_{1,2,\ell} \le \frac{\varepsilon}{2}$. We next set $v_n = \psi_{\frac{1}{n}} * v$ for the mollifiers $\psi_{\frac{1}{n}}$ from formulas (D.2) – (D.7) with $n \in \mathbb{N}$. From (D.7) and Lemma D.6 we deduce that v_n tends to v in $H^1(\mathbb{R}^d)$ as $n \to \infty$. Since $\operatorname{supp} v_n \subseteq \operatorname{supp} v + \overline{B}(0,1)$, this fact also yields that $|\ell| v_n \to |\ell| v$ in $L^2(\mathbb{R}^d)$. As a result, we find a function $w \in C_c^\infty(\mathbb{R}^d)$ such that $\|u_0 - w\|_{1,2,\ell} \le \varepsilon$.

So there exist initial values $\varphi_n \in C_c^\infty(\mathbb{R}^d)$ converging to u_0 in $H_\ell^1(\mathbb{R}^d)$ as $n \to \infty$. Let $J \subseteq J(u_0)$ be a compact interval with $0 \in J$. Due to Theorem 12.6 (d), there is an index $n_0 \in \mathbb{N}$ such that $J \subseteq J(\varphi_n)$ for all $n \ge n_0$ and the solutions $u_n = u(\,\cdot\,; \varphi_n)$ tend to u in $C(J, H^1(\mathbb{R}^d))$ as $n \to \infty$. Sobolev's embedding (5.5) then yields that $\|u_n(t)\|_{\alpha+1}^{\alpha+1} \to \|u(t)\|_{\alpha+1}^{\alpha+1}$ as $n \to \infty$, uniformly for $t \in J$. Due to step 4), ϕ''_{u_n} is given by (14.3) and hence

$$\phi''_{u_n}(t) \to 16E(u_0) + (4d - 8\tfrac{d+2}{\alpha+1})\mu \|u(t)\|_{\alpha+1}^{\alpha+1}$$

as $n \to \infty$, uniformly for $t \in J$. It remains to show that for a subsequence $\phi'_{u_{n_j}}(t) \to \phi'_u(t)$ for each $t \in J$ as $j \to \infty$, where these derivatives are given by (14.2) due to step 1). Fix $t \in J$. We know that $u_n(t) \to u(t)$ in $L^2(\mathbb{R}^d)$ and $\nabla u_n(t) \to \nabla u(t)$ in $L^2(\mathbb{R}^d)^d$ as $n \to \infty$. As we have seen before (14.5), the functions $\ell u_n(t)$, $n \in \mathbb{N}$, are uniformly bounded in $L^2(\mathbb{R}^d)^d$. Hence they posses a weakly converging subsequence $(\ell u_{n_j}(t))_j$. On the other hand, for $v \in C_c(\mathbb{R}^d)^d$ we know that $(\ell u_{n_j}(t)|v)_{L^2} \to (\ell u(t)|v)_{L^2}$ as $j \to \infty$. As a result, $\ell u_{n_j}(t)$ tends weakly to $\ell u(t)$ and hence

$$\phi'_{u_{n_j}}(t) = 4 \operatorname{Im}(\nabla u_{n_j}(t)|\ell u_{n_j}(t))_{L^2} \to 4 \operatorname{Im}(\nabla u(t)|\ell u(t))_{L^2}$$

as $j \to \infty$, as needed. □

In the above proof we used the following density result to approximate a given function in two norms simultaneously. These norms involve differentiability in time and in space, respectively.

Lemma 14.2. *Let $J = [a,b]$ be a compact interval and let $\mathcal{G}(J) = C^1(J, L^2(\mathbb{R}^d)) \cap C(J, H^1(\mathbb{R}^d))$ be endowed with the norm given by $\|u\|_{\mathcal{G}} = \max\{\|u\|_{C^1(J,L^2)}, \|u\|_{C(J,H^1)}\}$. Then the space $C^1(J, H^1(\mathbb{R}^d))$ is dense in $\mathcal{G}(J)$.*

Proof. Let $u \in \mathcal{G}(J)$. There are functions $\varphi_1 \in C^1([a-1,a])$ and $\varphi_2 \in C^1([b,b+1])$ such that the extension $\widetilde{u}$ of u given by

$$\widetilde{u}(t) = \begin{cases} \varphi_1(t)u(a), & t \in [a-1,a), \\ u(t), & t \in [a,b], \\ \varphi_2(t)u(b), & t \in (b,b+1], \end{cases}$$

belongs to $\mathcal{G}([a-1,b+1])$ and has compact support in $(a-1,b+1)$. Extend $\widetilde{u}$ by zero to $\mathbb{R}$. Let $\psi_{\frac{1}{n}}$, $n \in \mathbb{N}$, be a one dimensional mollifier as in formula

(D.2). For $f \in C(\mathbb{R}, L^2(\mathbb{R}^d))$ we define

$$\psi_{\frac{1}{n}} * f(t) = \int_{\mathbb{R}} \psi_{\frac{1}{n}}(t-s) f(s)\,\mathrm{d}s = \int_{\mathbb{R}} \psi_{\frac{1}{n}}(s) f(t-s)\,\mathrm{d}s, \qquad t \in \mathbb{R}.$$

It is straight forward to show that $\psi_{\frac{1}{n}} * f(t) \to f(t)$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$, locally uniformly for $t \in \mathbb{R}$. We set $u_n = \psi_{\frac{1}{n}} * \widetilde{u}$. One easily sees that $u_n \in C^1(\mathbb{R}, H^1(\mathbb{R}^d))$, $\nabla u_n = \psi_{\frac{1}{n}} * \nabla \widetilde{u}$ and $u_n' = \psi_{\frac{1}{n}} * \widetilde{u}'$. These facts imply that $u_n \to u$ in $C([a,b], H^1(\mathbb{R}^d))$ and $C^1([a,b], L^2(\mathbb{R}^d))$. □

In our main result we will need a somewhat differently formulated version of the above proposition which is stated in the next corollary. We set $\theta_t(x) = \mathrm{e}^{-\frac{\mathrm{i}|x|^2}{4t}}$ for $x \in \mathbb{R}^d$ and $t \neq 0$. We recall that $E(\varphi) = \frac{1}{2}\||\nabla\varphi|\|_2^2 + \frac{\mu}{\alpha+1}\|\varphi\|_{\alpha+1}^{\alpha+1}$ is the energy of $\varphi \in H^1(\mathbb{R}^d)$.

Corollary 14.3. *Let $u_0 \in H^1_\ell(\mathbb{R}^d)$, $\mu \in \{-1, 1\}$, $\alpha \in (1, \alpha_c)$ and let $u \in C^1(J(u_0), H^{-1}(\mathbb{R}^d)) \cap C(J(u_0), H^1_\ell(\mathbb{R}^d))$ be the solution of (13.1). Set $v(t) = \theta_t u(t) = \mathrm{e}^{-\frac{\mathrm{i}|\cdot|^2}{4t}} u(t)$ for $t \in J(u_0) \setminus \{0\}$. We then have*

$$\begin{aligned} h_u(t) &:= \||(\ell + 2\mathrm{i}t\nabla)u(t)|\|_2^2 + \frac{8t^2}{\alpha+1}\mu\|u(t)\|_{\alpha+1}^{\alpha+1} \\ &= \||\ell|u_0\|_2^2 + \int_0^t s\big(8\tfrac{d+2}{\alpha+1} - 4d\big)\mu\|u(s)\|_{\alpha+1}^{\alpha+1}\,\mathrm{d}s, \end{aligned} \tag{14.15}$$

$$8t^2 E(v(t)) = \||\ell|u_0\|_2^2 + \int_0^t s\big(8\tfrac{d+2}{\alpha+1} - 4d\big)\mu\|u(s)\|_{\alpha+1}^{\alpha+1}\,\mathrm{d}s \tag{14.16}$$

for all $t \in J(u_0)$, where $t \neq 0$ in (14.16).

Proof. Since $E(u(t)) = E(u_0)$ by Theorem 13.2, we can compute

$$\begin{aligned} h_u(t) &= \||\ell|u(t)\|_2^2 + 4t^2\||\nabla u(t)|\|_2^2 + 4t\,\mathrm{Re}\,\mathrm{i}\int_{\mathbb{R}^d} \overline{u}(t)\ell\cdot\nabla u(t)\,\mathrm{d}x \\ &\quad + \frac{8t^2}{\alpha+1}\mu\|u(t)\|_{\alpha+1}^{\alpha+1} \\ &= \||\ell|u(t)\|_2^2 + 8t^2 E(u_0) - 4t\,\mathrm{Im}\int_{\mathbb{R}^d} \overline{u}(t)\ell\cdot\nabla u(t)\,\mathrm{d}x \end{aligned}$$

for $t \in J(u_0)$. In view of Proposition 14.1, each term of the right-hand side is continuously differentiable in time so that $h_u \in C^1(J(u_0))$. Moreover, (14.2) and (14.3) yield

$$\begin{aligned} h_u'(t) &= 4\,\mathrm{Im}\int_{\mathbb{R}^d} \overline{u}(t)\ell\cdot\nabla u(t)\,\mathrm{d}x + 16tE(u_0) - 4\,\mathrm{Im}\int_{\mathbb{R}^d} \overline{u}(t)\ell\cdot\nabla u(s)\,\mathrm{d}x \\ &\quad - 16tE(u_0) - t\big(4d - 8\tfrac{d+2}{\alpha+1}\big)\mu\|u(t)\|_{\alpha+1}^{\alpha+1}, \\ &= t\big(8\tfrac{d+2}{\alpha+1} - 4d\big)\mu\|u(t)\|_{\alpha+1}^{\alpha+1}. \end{aligned}$$

Equation (14.15) follows by integration. To derive (14.16), we note that

$$(\ell + 2\mathrm{i}t\nabla)u(t) = 2\mathrm{i}t\theta_{-t}\nabla v(t), \qquad t \in J(u_0) \setminus \{0\}.$$

This identity and (14.15) lead to

$$\begin{aligned}8t^2 E(v(t)) &= \||2t|\nabla v(t)|\|_2^2 + \frac{8t^2}{\alpha+1}\mu \int_{\mathbb{R}^d} |v(t)|^{\alpha+1}\,\mathrm{d}x \\ &= \|(\ell + 2\mathrm{i}t\nabla)u(t)\|_2^2 + \frac{8t^2}{\alpha+1}\mu \int_{\mathbb{R}^d} |u(t)|^{\alpha+1}\,\mathrm{d}x \\ &= \||\ell|u_0\|_2^2 + \int_0^t s(8\tfrac{d+2}{\alpha+1} - 4d)\mu \|u(s)\|_{\alpha+1}^{\alpha+1}\,\mathrm{d}s\end{aligned}$$

for $t \in J(u_0) \setminus \{0\}$. □

We can finally establish the announced convergence result in the defocusing case $\mu = 1$. Recall from Theorem 13.3 that in this case we have $J(u_0) = \mathbb{R}$ for each $u_0 \in H^1(\mathbb{R}^d)$. For simplicity we restrict ourselves to the case $d \geq 3$, cf. Theorem 7.3.1 in [**Caz03**].

THEOREM 14.4. *Let $u_0 \in H^1_\ell(\mathbb{R}^d)$ and let u be the solution of the defocusing nonlinear Schrödinger equation (14.1) with $1 < \alpha < \alpha_c = \frac{d+2}{d-2}$ and $d \geq 3$. Then there is a constant c only depending on α, d, r and $\|u_0\|_{1,2,\ell}$ such that*

(a) if $1 + \frac{4}{d} \leq \alpha < \alpha_c$ and $2 < r \leq \alpha_c$, then $\|u(t)\|_r \leq c|t|^{-d(\frac12 - \frac1r)}$;
(b) if $1 < \alpha < 1 + \frac{4}{d}$ and $2 < r \leq \alpha + 1$, then $\|u(t)\|_r \leq c|t|^{-d(\frac12-\frac1r)}$;
(c) if $1 < \alpha < 1+\frac{4}{d}$ and $\alpha+1 < r \leq \alpha_c$, then $\|u(t)\|_r \leq c|t|^{-d(\frac12-\frac1r)(1-\theta(r))}$;

where $\theta(r) = \dfrac{(r-\alpha-1)(4+d-d\alpha)}{(r-2)(2+d-\alpha(d-2))} \in (0,1)$ and $|t| \geq 1$ in (b) and (c).

PROOF. We set $v(t) = \theta_t u(t)$ as in Corollary 14.3. Observe that $\|v(t)\|_p = \|u(t)\|_p$ for all $p \in [1,\infty]$.

1) Let $\alpha \in [1 + \frac{4}{d}, \alpha_c)$. Since $\alpha \geq 1 + \frac{4}{d}$, the integral in (14.16) is non positive so that

$$2|t|\|\nabla v(t)\|_2 \leq (8t^2 E(v(t)))^{\frac12} \leq \||\ell|u_0\|_2. \tag{14.17}$$

As in the proof of Theorem 13.3 (b), we obtain the Gagliardo-Nirenberg inequality

$$\|\varphi\|_r \leq c_d \|\varphi\|_2^{1-d(\frac12-\frac1r)} \||\nabla\varphi|\|_2^{d(\frac12-\frac1r)} \tag{14.18}$$

for $\varphi \in H^1(\mathbb{R}^d)$ and a constant c_d only depending on d. Estimates (14.18) and (14.17) yield

$$\begin{aligned}\|u(t)\|_r = \|v(t)\|_r &\leq c_d \|v(t)\|_2^{1-d(\frac12-\frac1r)} \big(\tfrac{1}{2|t|}\||\ell|u_0\|_2\big)^{d(\frac12-\frac1r)} \\ &= c_d (2|t|)^{\frac{d}{r}-\frac{d}{2}} \|u(t)\|_2^{1-\frac{d}{2}+\frac{d}{r}} \||\ell|u_0\|_2^{\frac{d}{2}-\frac{d}{r}} \\ &\leq 2^{\frac{d}{r}-\frac{d}{2}} c_d |t|^{\frac{d}{r}-\frac{d}{2}} \|u_0\|_{2,\ell}\end{aligned} \tag{14.19}$$

for $t \neq 0$, because of $\|u(t)\|_2 = \|u_0\|_2$ by Theorem 13.2. So (a) holds.

2) Let $1 < \alpha < 1 + \frac{4}{d}$. We consider the case $t \geq 1$. Negative times are treated in the same way. We first establish the decay of $\|v(t)\|_{\alpha+1}$ and $\||\nabla v(t)|\|_2$. Formula (14.16) implies that

$$8t^2 E(v(t)) = 8E(v(1)) + \big(8\tfrac{d+2}{\alpha+1} - 4d\big) \int_1^t s\|u(s)\|_{\alpha+1}^{\alpha+1}\,\mathrm{d}s$$

for $t \geq 1$. We set $g_u(t) = t^2\|u(t)\|_{\alpha+1}^{\alpha+1} = t^2\|v(t)\|_{\alpha+1}^{\alpha+1}$ for $t \in \mathbb{R}$. It follows

$$g_u(t) \leq t^2(1+\alpha)E(v(t)) = (1+\alpha)E(v(1)) + \tfrac{1}{2}(d+4-d\alpha)\int_1^t \tfrac{1}{s} g_u(s)\,\mathrm{d}s.$$

Since $E(u(s)) = E(u_0)$ by Theorem 13.2, equation (14.16) leads to

$$\begin{aligned} 8E(v(1)) &\leq \|u_0\|_{2,\ell}^2 + c \sup_{0\leq s\leq 1} \|u(s)\|_{\alpha+1}^{\alpha+1} \leq \|u_0\|_{2,\ell}^2 + c \sup_{0\leq s\leq 1} E(u(s)) \\ &= \|u_0\|_{2,\ell}^2 + cE(u_0) \leq \|u_0\|_{2,\ell}^2 + c(\|u_0\|_{1,2}^2 + \|u_0\|_{1,2}^{\alpha+1}). \end{aligned}$$

We thus arrive at

$$g_u(t) \leq c + \tfrac{1}{2}(d+4-d\alpha)\int_1^t \tfrac{1}{s} g_u(s)\,\mathrm{d}s, \quad t \geq 1.$$

Here and below c denotes differing constants only depending on α, d, r and $\|u_0\|_{1,2,\ell}$. Set $\beta = \frac{1}{2}(d+4-d\alpha) > 0$. Gronwall's inequality then yields

$$g_u(t) \leq c \exp \int_1^t \tfrac{\beta}{s}\,\mathrm{d}s = ct^\beta, \tag{14.20}$$

$$\|v(t)\|_{\alpha+1} = \big(t^{-2} g_u(t)\big)^{\frac{1}{\alpha+1}} \leq ct^{\frac{\beta-2}{\alpha+1}} = ct^{\frac{d}{\alpha+1}-\frac{d}{2}}. \tag{14.21}$$

for $t \geq 1$. Combining (14.16) and (14.20) we further infer

$$\begin{aligned} 4t^2 \int_{\mathbb{R}^d} |\nabla v(t)|^2\,\mathrm{d}x &\leq c + c\int_0^t \frac{1}{s} g_u(s)\,\mathrm{d}s \\ &\leq c + c\int_0^t s^{\beta-1}\,\mathrm{d}s = c(1 + t^2 t^{\frac{d}{2}-\frac{d\alpha}{2}}), \end{aligned}$$

$$\||\nabla v(t)|\|_2 \leq c(t^{-1} + t^{\frac{d}{4}-\frac{d\alpha}{4}}) \leq ct^{\frac{d}{4}-\frac{d\alpha}{4}}, \quad t \geq 1. \tag{14.22}$$

3) We can now show the assertions (b) and (c). We start with $r \in (2, \alpha+1]$. Using the interpolation inequality (D.13), estimate (14.21) and $\|u(t)\|_2 = \|u_0\|_2$, we conclude

$$\|u(t)\|_r = \|v(t)\|_r \leq \|v(t)\|_{\alpha+1}^\theta \|v(t)\|_2^{1-\theta} \leq ct^{\theta(\frac{d}{1+\alpha}-\frac{d}{2})}\|u(t)\|_2^{1-\theta} = ct^{\frac{d}{r}-\frac{d}{2}}$$

for $t \geq 1$, where $\theta = (\frac{1}{2}-\frac{1}{r})(\frac{1}{2}-\frac{1}{1+\alpha})^{-1} \in (0,1]$. Hence, (b) is valid.

To prove (c), we take $r \in (\alpha+1, \alpha_c]$. We now employ the interpolation inequality (D.13), Sobolev's inequality (13.9) and the estimates (14.21) and (14.22). These results lead to

$$\begin{aligned} \|u(t)\|_r = \|v(t)\|_r &\leq \|v(t)\|_{\frac{2d}{d-2}}^\vartheta \|v(t)\|_{\alpha+1}^{1-\vartheta} \leq c\||\nabla v(t)|\|_2^\vartheta\, t^{(1-\vartheta)(\frac{d}{1+\alpha}-\frac{d}{2})} \\ &\leq ct^{\vartheta(\frac{d}{4}-\frac{d\alpha}{4})} t^{(1-\vartheta)(\frac{d}{1+\alpha}-\frac{d}{2})} = ct^{-d(\frac{1}{2}-\frac{1}{r})(1-\theta(r))} \end{aligned}$$

for $\vartheta = (\frac{1}{1+\alpha}-\frac{1}{r})(\frac{1}{1+\alpha}-\frac{d-2}{2d})^{-1} \in (0,1]$ and $t \geq 1$. □

We note that there are (weaker) decay results for $u_0 \in H^1(\mathbb{R}^d)$, see Theorem 7.7.1 of [**Caz03**]. The above convergence Theorem 14.4 concerns the norm in $L^r(\mathbb{R}^d)$ for $r > 2$. What happens for $r = 2$? Here the behavior of the defocusing nonlinear Schrödinger equation (14.1) is even closer to the free linear Schrödinger equation. It is described by the following *scattering* results:

Let $d \geq 3$, $\alpha \in (1+\frac{4}{d}, \alpha_c)$ and $u_0 \in H^1(\mathbb{R}^d)$. Then there are unique $u_\pm \in H^1(\mathbb{R}^d)$ such that

$$\|u(t) - T(t)u_\pm\|_{1,2} = \|T(-t)u(t) - u_\pm\|_{1,2} \ \longrightarrow\ 0 \quad \text{as } t \to \pm\infty.$$

Moreover, $\|u_\pm\|_2 = \|u_0\|_2$ and $\||\nabla u_\pm|\|_2^2 = 2E(u_0)$. Finally, the maps

$$U_\pm : H^1(\mathbb{R}^d) \to H^1(\mathbb{R}^d), \quad u_0 \mapsto u_\pm,$$

are continuous, bijective and have continuous inverses $\Omega_\pm : H^1(\mathbb{R}^d) \to H^1(\mathbb{R}^d)$ which are called "wave operators". We refer to Section 7.8 of [**Caz03**] for these and related results and for references to the original papers. See also Exercise 14.3. There is an analogous scattering theory in the space $H^1_\ell(\mathbb{R}^d)$, see Section 7.4 in [**Caz03**].

In the focusing case $\mu = -1$, the asymptotic behavior of solutions is completely different. We have noted before that blow-up occurs if $\alpha \geq 1 + \frac{4}{d}$. (See Example 10.3 and Sections 6.5 and 6.6 of [**Caz03**].) Moreover, for $1 < \alpha < \alpha_c$ and $\omega > 0$ there are functions $0 \neq \varphi_\omega \in H^2(\mathbb{R}^d)$ such that

$$-\Delta\varphi_\omega + \omega\varphi_\omega = |\varphi_\omega|^{\alpha-1}\varphi_\omega, \tag{14.23}$$

see Section 8.1 of [**Caz03**] and Example 10.2. As a result, (13.1) with $\mu = -1$ admits a standing wave solution u_ω given by $u_\omega(t) = \mathrm{e}^{\mathrm{i}\omega t}\varphi_\omega$ for $t \in \mathbb{R}$. Actually, also the functions given by

$$u_\omega^{\theta,y}(t) = \mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{\mathrm{i}\theta}\varphi_\omega(\,\cdot\, - y), \qquad \theta \in \mathbb{R},\ y \in \mathbb{R}^d,$$

solve (13.1). Due to Section 8.1 of [**Caz03**], there also exist solutions of (14.23) minimizing the functional $v \mapsto E(v) + \frac{\omega}{2}\|v\|_2^2$ among all solutions of (14.23). Fix such a minimizing solution φ_ω of (14.23). If $1 < \alpha < 1 + \frac{4}{d}$, then the corresponding standing wave u_ω is *orbitally stable*, i.e.,

$$\forall\,\varepsilon > 0\ \exists\,\delta > 0\ \forall u_0 \in \overline{B}_{H^1}(\varphi_\omega, \delta) :\ \operatorname{dist}(u(t; u_0), S) \leq \varepsilon \text{ for all } t \in \mathbb{R},$$

where $S = \left\{\mathrm{e}^{\mathrm{i}\theta}\varphi_\omega(\,\cdot\, - y) \,\middle|\, \theta \in \mathbb{R}, y \in \mathbb{R}^d\right\}$. See Theorem 8.3.1 in [**Caz03**]. If $1 + \frac{4}{d} \leq \alpha < \alpha_c$, then there are $\varphi_n \in H^1(\mathbb{R}^d)$ converging to φ_ω in $H^1(\mathbb{R}^d)$ as $n \to \infty$ which have a bounded existence interval $J(\varphi_n)$. We refer to Section 8.2 of [**Caz03**] for this and related instability results.

Recently, the description of the blow-up case $\alpha \geq 1 + \frac{4}{d}$ and $\mu = -1$ was much refined by W. Schlag and M. Beceanu for the model case $\alpha = d = 3$. In somewhat differing settings, they constructed a manifold of finite codimension consisting of solutions of (13.1) which converge to a variant of the manifold S above, see [**Bec12**], [**Sch09**].

Exercises

Exercise 14.1. Let $1+\frac{4}{d} \le \alpha < \alpha_c$. Let $u_0 \in H^1(\mathbb{R}^d)$ satisfy $|x|u_0 \in L^2(\mathbb{R}^d)$ and $E(u_0) < 0$. Let u be the solution of the focusing nonlinear Schrödinger equation (13.1) with $\mu = -1$. Show that the maximal existence interval $J(u_0)$ is bounded.

Exercise 14.2. Let $\alpha_0 > 0$ satisfy $d\alpha_0^2 - (d+2)\alpha_0 - 2 = 0$, $\max\{\alpha_0, 1\} < \alpha < \alpha_c$ and $d \ge 3$. Let $u_0 \in H^1(\mathbb{R}^d)$ with $|x|u_0 \in L^2(\mathbb{R}^d)$ and let u be the solution of the defocusing nonlinear Schrödinger equation (14.1) on $\mathbb{R}$. Let $p = 1+\alpha$ and q with $\frac{2}{q} + \frac{d}{p} = \frac{d}{2}$. Show that $u \in L^q(\mathbb{R}, W_q^1(\mathbb{R}^d))$.

Exercise 14.3. In the setting of Exercise 14.2 set $v(t) = T(-t)u(t)$ for $t \in \mathbb{R}$, where $T(\cdot)$ is the free Schrödinger group. Show that $v(t)$ converges in $H^1(\mathbb{R}^d)$ as $t \to +\infty$ and as $t \to -\infty$. (Hint: You may use the equation $v(t) - v(\tau) = \int_\tau^t v'(s)\,\mathrm{d}s$ as a starting point.)

APPENDIX A

Closed operators

There are important unbounded linear operators in analysis. Fortunately they often have certain weaker continuity properties. We start with a basic example.

EXAMPLE A.1. Let $X = C([0,1])$ be endowed with the supremum norm and let $Af = f'$ with $\mathrm{D}(A) = C^1([0,1])$. Then A is linear, but not bounded. Indeed, consider the functions $u_n \in \mathrm{D}(A)$ given by $u_n(x) = (\frac{1}{\sqrt{n}})\sin(nx)$ for $n \in \mathbb{N}$, which satisfy $\|u_n\|_\infty \to 0$ and

$$\|Au_n\|_\infty \geq |u_n'(0)| = \sqrt{n} \to \infty$$

as $n \to \infty$. However, if $(f_n)_n \subseteq \mathrm{D}(A)$ satisfies $f_n \to f$ **and** $Af_n \to g$ in $C([0,1])$ as $n \to \infty$, then $f \in C^1([0,1]) = \mathrm{D}(A)$ and $g = f' = Af$ (which is shown in introductory courses in analysis). ♢

The above example motivates the next definition.

DEFINITION A.2. *Let A be a linear operator from X to Y, where X and Y are Banach spaces. The operator A is called* closed *if for all $x_n \in \mathrm{D}(A)$, $n \in \mathbb{N}$, such that there exists $x = \lim_{n\to\infty} x_n$ in X and $y = \lim_{n\to\infty} Ax_n$ in Y it holds that $x \in \mathrm{D}(A)$ and $Ax = y$.*

Hence, $\lim_{n\to\infty} Ax_n = A(\lim_{n\to\infty} x_n)$ if both $(x_n)_n$ and $(Ax_n)_n$ converge.

REMARK A.3. It is clear that every operator $A \in \mathcal{B}(X,Y)$ is closed (where $\mathrm{D}(A) = X$). The operator A from Example A.1 is closed.

EXAMPLE A.4. (a) Let $X = C([0,1])$ and $Af = f'$ with

$$\mathrm{D}(A) = \Big\{ f \in C^1([0,1]) \,\big|\, f(0) = 0 \Big\}.$$

Let $(f_n)_n \subseteq \mathrm{D}(A)$ and $f, g \in X$ be such that $f_n \to f$ and $Af_n = f_n' \to g$ in X as $n \to \infty$. As observed above it follows $f \in C^1([0,1])$ and $f' = g$. Since $0 = f_n(0) \to f(0)$ as $n \to \infty$, we obtain $f \in \mathrm{D}(A)$. This means that A is closed on X. In the same way we see that $A_1 f = f'$ with

$$\mathrm{D}(A_1) = \Big\{ f \in C^1([0,1]) \,\big|\, f(0) = f'(1) = 0 \Big\}$$

is closed. There are many more variants of this result. ♢

(b) Let $X = C([0,1])$ and $Af = f'$ with $\mathrm{D}(A) = C_c^1((0,1])$. This operator is not closed. In fact, consider the functions $f_n \in \mathrm{D}(A)$ given by

$$f_n(t) = \begin{cases} 0, & 0 \leq t < \frac{1}{n}, \\ (t - \frac{1}{n})^2, & \frac{1}{n} \leq t \leq 1, \end{cases}$$

for every $n \in \mathbb{N}$. Then, $f_n \to f$ and $f_n' \to f'$ in X as $n \to \infty$, where $f(t) = t^2$. However, $\operatorname{supp} f = [0,1]$ and so $f \notin \mathrm{D}(A)$. $\diamondsuit$

(c) Let $X = L^p(\mathbb{R}^d)$, $1 \le p \le \infty$, and $m : \mathbb{R}^d \to \mathbb{C}$ be measurable. Define $Af = mf$ with its *maximal domain*

$$\mathrm{D}(A) = \{f \in X \,|\, mf \in X\}\,.$$

Then A is closed. Indeed, let $f_n \to f$ and $Af_n = mf_n \to g$ in X as $n \to \infty$. Then there is a subsequence $(n_j)_j \subseteq \mathbb{N}$ such that $f_{n_j}(x) \to f(x)$ and $m(x)f_{n_j}(x) \to g(x)$ for a.e. $x \in \mathbb{R}^d$, as $j \to \infty$. Hence, $mf = g$ in $L^p(\mathbb{R}^d)$ and we thus obtain $f \in \mathrm{D}(A)$ and $Af = g$. $\diamondsuit$

(d) Let $X = L^1([0,1])$, $Y = \mathbb{C}$, and $Af = f(0)$ with $\mathrm{D}(A) = C([0,1])$. Then A is not closed. In fact, consider the functions $f_n \in \mathrm{D}(A)$ given by

$$f_n(t) = \begin{cases} 1 - nt, & 0 \le t \le \frac{1}{n}, \\ 0, & \frac{1}{n} < t \le 1, \end{cases}$$

for every $n \in \mathbb{N}$. Then $\|f_n\|_1 = \frac{1}{2n} \to 0$ as $n \to \infty$, but $Af_n = f_n(0) = 1$. $\diamondsuit$

Definition A.5. *Let A be a linear operator from X to Y. The* graph *of A is given by*

$$\operatorname{gr}(A) = \{(x, Ax) \in X \times Y \,|\, x \in \mathrm{D}(A)\}\,.$$

The graph norm *of A is defined by $\|x\|_A = \|x\|_X + \|Ax\|_Y$. We write $[\mathrm{D}(A)]$ if we equip $\mathrm{D}(A)$ with $\|\cdot\|_A$.*

Of course, $\|\cdot\|_A$ is equivalent to $\|\cdot\|_X$ if $A \in \mathcal{B}(X,Y)$. We endow $X \times Y$ with the norm $\|(x,y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$.

Lemma A.6. *For a linear operator A from X to Y the following assertions hold.*

(a) $\operatorname{gr}(A) \subseteq X \times Y$ is a linear subspace.
(b) $[\mathrm{D}(A)]$ is a normed vector space and $A \in \mathcal{B}([\mathrm{D}(A)], Y)$.
(c) A is closed if and only if $\operatorname{gr}(A)$ is closed in $X \times Y$ if and only if $[\mathrm{D}(A)]$ is a Banach space.
(d) Let A be injective and put $\mathrm{D}(A^{-1}) := \mathrm{R}(A)$. Then, A is closed from X to Y if and only if A^{-1} is closed from Y to X.

Proof. Assertions (a) and (b) are straightforward to check.

(c) The operator A is closed if and only if for all $x_n \in \mathrm{D}(A)$, $n \in \mathbb{N}$, and $(x,y) \in X \times Y$ with $(x_n, Ax_n) \to (x,y)$ in $X \times Y$ as $n \to \infty$, we have $(x,y) \in \operatorname{gr}(A)$. This property is equivalent to the closedness of $\operatorname{gr}(A)$. Since $\|(x,Ax)\|_{X\times Y} = \|x\|_X + \|Ax\|_Y$, a Cauchy sequence or a converging sequence in $\operatorname{gr}(A)$ corresponds to a Cauchy or a converging sequence in $[\mathrm{D}(A)]$, respectively. Hence, $[\mathrm{D}(A)]$ is complete if and only if $(\operatorname{gr}(A), \|\cdot\|_{X\times Y})$ is complete if and only if $\operatorname{gr}(A) \subseteq X \times Y$ is closed.

(d) Assertion (d) follows from (c) since

$$\operatorname{gr}(A^{-1}) = \{(y, A^{-1}y) \,|\, y \in \mathrm{R}(A)\} = \{(Ax, x) \,|\, x \in \mathrm{D}(A)\}$$

is closed in $Y \times X$ if and only if $\operatorname{gr}(A)$ is closed in $X \times Y$. $\square$

Theorem A.7 (Closed Graph Theorem). *Let X and Y be Banach spaces and A be a closed operator from X to Y. Then A is* bounded *(i.e., $\|Ax\| \le c\|x\|$ for some $c \ge 0$ and all $x \in \mathrm{D}(A)$) if and only if $\mathrm{D}(A)$ is closed in X.*

In particular, a closed operator with $\mathrm{D}(A) = X$ already belongs to $\mathcal{B}(X,Y)$.

Proof. "$\Leftarrow$": Let $\mathrm{D}(A)$ be closed in X. Then $\mathrm{D}(A)$ is a Banach space for $\|\cdot\|_X$ and $\|\cdot\|_A$. Since $\|x\|_X \le \|x\|_A$ for all $x \in \mathrm{D}(A)$, the open mapping theorem gives a constant $c > 0$ such that $\|Ax\|_Y \le \|x\|_A \le c\|x\|_X$ for all $x \in \mathrm{D}(A)$.

"$\Rightarrow$": Let A be bounded and let $(x_n)_{n\in\mathbb{N}} \subseteq \mathrm{D}(A)$ converge to $x \in X$ with respect to $\|\cdot\|_X$. Then $\|Ax_n - Ax_m\|_Y \le c\,\|x_n - x_m\|_X$, and so the sequence $(Ax_n)_{n\in\mathbb{N}}$ is Cauchy in Y. There thus exists $y := \lim_{n\to\infty} Ax_n$ in Y. The closedness of A shows that $x \in \mathrm{D}(A)$; i.e., $\mathrm{D}(A)$ is closed in X. □

Remark A.8. (a) Theorem A.7 is wrong without completeness. Consider for instance the operator T given by $(Tf)(t) = tf(t)$, $t \in \mathbb{R}$, on $C_c(\mathbb{R})$ with supremum norm. This linear operator is everywhere defined, unbounded and closed. To check the closedness, take $f_n, f, g \in C_c(\mathbb{R})$ such that $f_n(t) \to f(t)$ and $(Tf_n)(t) = tf_n(t) \to g(t)$ uniformly for $t \in \mathbb{R}$ as $n \to \infty$. Then $g(t) = tf(t)$ for all $t \in \mathbb{R}$ and so $g = Tf$. ♢

(b) On an infinite dimensional Banach space X there are non-closed operators. In fact, let $\mathcal{B}$ be an algebraic basis of X (i.e., for each $x \in X$ there exist $n \in \mathbb{N}$ and unique $\alpha_1, \dots, \alpha_n \in \mathbb{C}$, $b_1, \dots, b_n \in \mathcal{B}$ such that $x = \alpha_1 b_1 + \ldots + \alpha_n b_n$). We may assume that $\|b\| = 1$ for all $b \in \mathcal{B}$. Choose a countable subset $\mathcal{B}_0 = \{b_k \mid k \in \mathbb{N}\}$ of $\mathcal{B}$ and set

$$Tb_k = kb_k \text{ for } b_k \in \mathcal{B}_0 \quad \text{and} \quad Tb = 0 \text{ for } b \in \mathcal{B} \setminus \mathcal{B}_0.$$

Then T can be extended to a linear operator on X which is unbounded, since $\|Tb_k\| = k$ and $\|b_k\| = 1$. Thus T is not closed. ♢

It is a delicate mater to add or multiply closed operators. The situation is simpler if one operator is bounded.

Proposition A.9. *Let A be closed from X to Y, $T \in \mathcal{B}(X,Y)$, and $S \in \mathcal{B}(Z,X)$. Then the following operators are closed.*

(a) $B = A + T$ with $\mathrm{D}(B) = \mathrm{D}(A)$,

(b) $C = AS$ with $\mathrm{D}(C) = \{z \in Z \mid Sz \in \mathrm{D}(A)\}$.

In particular for $\lambda \in \mathbb{C}$ the operator $\lambda I - A$ with domain $\mathrm{D}(A)$ is closed in X.

Proof. (a) Let $x_n \in \mathrm{D}(B)$, $n \in \mathbb{N}$, and $x \in X$, $y \in Y$ such that $x_n \to x$ in X and $Bx_n = Ax_n + Tx_n \to y$ in Y as $n \to \infty$. Since T is bounded, there exists $Tx = \lim_{n\to\infty} Tx_n$ and so $Ax_n \to y - Tx$ as $n \to \infty$. The closedness of A implies that $x \in \mathrm{D}(A) = \mathrm{D}(B)$ and $Ax = y - Tx$; i.e., $Bx = Ax + Tx = y$.

(b) Let $z_n \in \mathrm{D}(C)$, $n \in \mathbb{N}$, and $z \in Z$, $y \in Y$ such that $z_n \to z$ in Z and $ASz_n \to y$ in Y as $n \to \infty$. Since S is bounded, $x_n := Sz_n$ converges to Sz. Moreover, $Ax_n \to y$. We thus deduce $Sz \in \mathrm{D}(A)$ and $ASz = y$ from the closedness of A; i.e., $z \in \mathrm{D}(C)$ and $Cz = y$. □

Corollary A.10. *Let A be a linear operator on X and $\lambda \in \mathbb{C}$. Then the following assertions hold.*

(a) If $\lambda I - A$ (or $\lambda I + A$) is closed, then A is closed.

(b) If $\lambda I - A$ is bijective with $(\lambda I - A)^{-1} \in \mathcal{B}(X)$, then A is closed.

PROOF. Assertion (a) is a consequence of Proposition A.9 since $A = \pm((\lambda I \pm A) - \lambda I)$. For (b), Lemma A.6 shows that $\lambda I - A$ is closed, and then assertion (a) yields (b). □

The following basic examples show that closedness can be lost when taking sums or products of closed operators.

EXAMPLE A.11. (a) Let $E = C_b(\mathbb{R}^2)$ and $A_k = \partial_k$ with

$$\mathrm{D}(A_k) = \{ f \in E \,|\, \text{the partial derivative } \partial_k f \text{ exists and belongs to } E \},$$

for $k = 1, 2$. Set $B = \partial_1 + \partial_2$ on

$$\mathrm{D}(B) := \mathrm{D}(A_1) \cap \mathrm{D}(A_2) = C_b^1(\mathbb{R}^2) = \Big\{ f \in C^1(\mathbb{R}^2) \,|\, f, \partial_1 f, \partial_2 f \in E \Big\}.$$

It is an exercise to show that A_1 and A_2 are closed.

However, B is not closed. Take $\phi_n \in C_b^1(\mathbb{R})$ converging uniformly to some $\phi \in C_b(\mathbb{R}) \backslash C^1(\mathbb{R})$. Set $f_n(x, y) = \phi_n(x - y)$ and $f(x, y) = \phi(x - y)$ for $(x, y) \in \mathbb{R}^2$ and $n \in \mathbb{N}$. We then obtain $f \in E$, $f_n \in \mathrm{D}(B)$, $\|f_n - f\|_\infty = \|\phi_n - \phi\|_\infty \to 0$ and $Bf_n = \phi_n' - \phi_n' = 0 \to 0$ as $n \to \infty$, but $f \notin \mathrm{D}(B)$. ♢

(b) Let $X = C([0, 1])$, $Af = f'$ with $\mathrm{D}(A) = C^1([0, 1])$, and $m \in C([0, 1])$ such that $m = 0$ on $[0, \frac{1}{2}]$. Define $T \in \mathcal{B}(X)$ by $Tf = mf$ for all $f \in X$. Then the operator TA with $\mathrm{D}(TA) = \mathrm{D}(A)$ is not closed.

To see this, take functions $f_n \in \mathrm{D}(A)$ such that $f_n = 1$ on $[\frac{1}{2}, 1]$ and f_n converges in X to some $f \notin C^1([0, 1])$. Then, $TAf_n = mf_n' = 0$ converges to 0, but $f \notin \mathrm{D}(A)$. ♢

APPENDIX B

The spectrum

The spectrum and the resolvent of a closed operator A are used in many parts of analysis to study A.

Definition B.1. *Let A be a closed operator on X. The* resolvent set *of A is given by*

$$\rho(A) = \{\lambda \in \mathbb{C} \mid \lambda I - A : \mathrm{D}(A) \to X \text{ is bijective}\},$$

and its spectrum *by*

$$\sigma(A) = \mathbb{C} \setminus \rho(A).$$

We further define the point spectrum *of A by*

$$\sigma_p(A) = \{\lambda \in \mathbb{C} \mid \text{there exists some } v \in \mathrm{D}(A) \setminus \{0\} \text{ with } \lambda v = Av\} \subseteq \sigma(A),$$

where we call $\lambda \in \sigma_p(A)$ an eigenvalue *of A and the corresponding v an* eigenvector *or* eigenfunction *of A. For $\lambda \in \rho(A)$ the operator*

$$R(\lambda, A) := (\lambda I - A)^{-1} : X \to X$$

is called the resolvent *(of A at λ).*

We will introduce further subdivisions of $\sigma(A)$ in Appendix C.

Remark B.2. (a) Let A be closed in X and $\lambda \in \rho(A)$. The resolvent $R(\lambda, A)$ has the range $\mathrm{D}(A)$. Proposition A.9 and Lemma A.6 further show that $R(\lambda, A)$ is closed and thus it belongs to $\mathcal{B}(X)$ by Theorem A.7. $\diamondsuit$

(b) Let A be a linear operator such that $\lambda I - A : \mathrm{D}(A) \to X$ has a bounded inverse for some $\lambda \in \mathbb{C}$. Then A is closed by Corollary A.10. In this case, the closedness assumption in Definition B.1 is redundant. $\diamondsuit$

We set $\mathrm{e}_\lambda(t) = \mathrm{e}^{\lambda t}$ for $\lambda \in \mathbb{C}$, $t \in J$, and any interval $J \subseteq \mathbb{R}$.

Example B.3. (a) Let $X = \mathbb{C}^d$ and $T \in \mathcal{B}(X)$. Then $\sigma(T)$ only consists of the eigenvalues $\lambda_1, \dots, \lambda_m$ of T, where $m \le d$. $\diamondsuit$

(b) Let $X = C([0,1])$, and $Au = u'$ with $\mathrm{D}(A) = C^1([0,1])$. Then $\mathrm{e}_\lambda \in \mathrm{D}(A)$ and $A\mathrm{e}_\lambda = \lambda \mathrm{e}_\lambda$ for each $\lambda \in \mathbb{C}$. Hence, $\lambda \in \sigma_p(A)$ and so $\sigma(A) = \sigma_p(A) = \mathbb{C}$. $\diamondsuit$

(c) Let $X = C([0,1])$, and $Au = u'$ with $\mathrm{D}(A) = \{u \in C^1([0,1]) \mid u(0) = 0\}$. Let $\lambda \in \mathbb{C}$ and $f \in X$. We then have $u \in \mathrm{D}(A)$ and $(\lambda I - A)u = f$ if and only if $u \in C^1([0,1])$, $u'(t) = \lambda u(t) - f(t)$ for all $t \in [0,1]$, and $u(0) = 0$, which is equivalent to

$$u(t) = -\int_0^t \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s,$$

for all $0 \le t \le 1$. Hence, $\sigma(A) = \emptyset$ and the resolvent is given by

$$R(\lambda, A)f(t) = -\int_0^t \mathrm{e}^{\lambda(t-s)} f(s)\,\mathrm{d}s,$$

for all $0 \le t \le 1$, $f \in X$, and $\lambda \in \mathbb{C}$. $\diamondsuit$

Let $U \subseteq \mathbb{C}$ be open. The *derivative* of a map $f : U \to Y$ at $\lambda \in U$ is given by

$$f'(\lambda) = \lim_{\mu \to \lambda} \frac{1}{\mu - \lambda}(f(\mu) - f(\lambda)) \in Y,$$

if the limit exists in Y.

The next theorem collects the basic properties of the resolvent and the spectrum of a closed operator.

Theorem B.4. *Let A be a closed operator on X and let $\lambda \in \rho(A)$. Then the following assertions hold.*

(a) $AR(\lambda, A) = \lambda R(\lambda, A) - I$, $AR(\lambda, A)x = R(\lambda, A)Ax$ *for all* $x \in \mathrm{D}(A)$,

$$\frac{1}{\mu - \lambda}(R(\lambda, A) - R(\mu, A)) = R(\lambda, A)R(\mu, A) = R(\mu, A)R(\lambda, A)$$

if $\mu \in \rho(A) \setminus \{\lambda\}$. *The latter identity is called the* resolvent equation.

(b) The spectrum $\sigma(A)$ *is closed, where* $B(\lambda, \frac{1}{\|R(\lambda,A)\|}) \subseteq \rho(A)$ *and*

$$R(\mu, A) = \sum_{n=0}^{\infty} (\lambda - \mu)^n R(\lambda, A)^{n+1} =: R_\mu,$$

if $|\lambda - \mu| < \frac{1}{\|R(\lambda,A)\|}$. *The series converges absolutely in* $\mathcal{B}(X, [\mathrm{D}(A)])$, *uniformly on* $\overline{B}(\lambda, \frac{\delta}{\|R(\lambda,A)\|})$ *for each* $\delta \in (0,1)$. *Moreover,*

$$\|R(\mu, A)\|_{\mathcal{B}(X,[\mathrm{D}(A)])} \le \frac{c(\lambda)}{1 - \delta}$$

for all $\mu \in \overline{B}(\lambda, \frac{\delta}{\|R(\lambda,A)\|})$ *and a constant* $c(\lambda)$ *depending only on* λ.

(c) The function $\rho(A) \to \mathcal{B}(X, [\mathrm{D}(A)])$, $\lambda \mapsto R(\lambda, A)$, *is infinitely often differentiable with*

$$\Big(\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big)^n R(\lambda, A) = (-1)^n \, n! \, R(\lambda, A)^{n+1} \qquad \text{for every } \; n \in \mathbb{N}.$$

(d) $\|R(\lambda, A)\| \ge \operatorname{dist}(\lambda, \sigma(A))^{-1}$.

Proof. (a) The first assertions are a consequence of

$$x = (\lambda I - A)R(\lambda, A)x = R(\lambda, A)(\lambda I - A)x,$$

where $x \in X$ in the first equality and $x \in \mathrm{D}(A)$ in the second one. For $\mu \in \rho(A)$ we further have

$$(\lambda R(\lambda, A) - AR(\lambda, A))R(\mu, A) = R(\mu, A),$$
$$R(\lambda, A)(\mu R(\mu, A) - AR(\mu, A)) = R(\lambda, A).$$

The resolvent equation then follows by subtraction and interchanging λ and μ.

(b) Let $|\mu - \lambda| \le \frac{\delta}{\|R(\lambda,A)\|}$ for some $\delta \in (0,1)$ and $x \in X$ with $\|x\| \le 1$. Then,

$$\begin{aligned}\|(\lambda - \mu)^n R(\lambda, A)^{n+1} x\|_A &\le \tfrac{\delta^n}{\|R(\lambda,A)\|^n}(\|AR(\lambda, A)R(\lambda, A)^n x\| + \|R(\lambda, A)^{n+1} x\|) \\ &\le \delta^n(\|\lambda R(\lambda, A)\| + 1 + \|R(\lambda, A)\|) =: \delta^n c(\lambda),\end{aligned}$$

where we used (a). So the series in (b) converges absolutely in $\mathcal{B}(X,[\mathrm{D}(A)])$ uniformly on $\overline{B}(\lambda, \frac{\delta}{\|R(\lambda,A)\|})$, and it can be estimated in norm by $c(\lambda)(1-\delta)^{-1}$. Using also $(\mu I-A)R(\lambda,A) = (\mu-\lambda)R(\lambda,A)+I$, we obtain

$$(\mu I-A)R_\mu = -\sum_{n=0}^{\infty}(\lambda-\mu)^{n+1}R(\lambda,A)^{n+1} + \sum_{n=0}^{\infty}(\lambda-\mu)^n R(\lambda,A)^n = I,$$

and similarly $R_\mu(\mu I-A)x = x$ for each $x\in \mathrm{D}(A)$. Hence, $\mu\in\rho(A)$ and $R_\mu = R(\mu,A)$. Assertion (d) follows from (b).

(c) Since $\lambda\mapsto R(\lambda,A)\in\mathcal{B}(X,[\mathrm{D}(A)])$ is locally bounded, due to the estimate in (b), the resolvent equation implies that the map $\lambda\mapsto R(\lambda,A)\in\mathcal{B}(X,[\mathrm{D}(A)])$ is continuous. Assertion (c) for $n=1$ thus follows from the resolvent equation. Assume that (c) holds for some $n\in\mathbb{N}$. We then obtain

$$\Big(\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big)^{n+1}R(\lambda,A) = (-1)^n\, n!\,\frac{\mathrm{d}}{\mathrm{d}\lambda}\,(R(\lambda,A)^{n+1}).$$

Using the formula

$$R(\mu,A)^{n+1}-R(\lambda,A)^{n+1} = \sum_{j=0}^{n} R(\mu,A)^{n-j}(R(\mu,A)-R(\lambda,A))R(\lambda,A)^j,$$

the continuity of $R(\cdot,A)$ and the assertion for $n=1$, we then conclude that (c) holds for $n+1$. □

Proposition B.5. *Let $\Omega\subseteq\mathbb{R}^d$, $m\in C(\Omega)$, $X=C_b(\Omega)$, and $Af=mf$ with* $\mathrm{D}(A)=\{f\in X\,|\,mf\in X\}$. *Then A is closed,*

$$\sigma(A)=\overline{m(\Omega)},$$

and $R(\lambda,A)f=\frac{1}{\lambda-m}f$ for all $\lambda\in\rho(A)$ and $f\in X$.

In particular, for every closed subset $S\subseteq\mathbb{C}$ there is a closed operator B on a Banach space with $\sigma(B)=S$. If $S\neq\emptyset$ is compact, then B can be chosen to be bounded.

Proof. The closedness of A can be shown as in Remark A.8. Let $\lambda\notin\overline{m(\Omega)}$ and $g\in C_b(\Omega)$. The function $f:=\frac{1}{\lambda-m}g$ then belongs to $C_b(\Omega)$ and $\lambda f-mf=g$ so that $mf=\lambda f-g\in C_b(\Omega)$. As a result, $f\in\mathrm{D}(A)$ and f is the unique solution in $\mathrm{D}(A)$ of the equation $\lambda f-Af=g$. This means that $\lambda\in\rho(A)$ and $R(\lambda,A)g=\frac{1}{\lambda-m}g$. In the case that $\lambda=m(z)$ for some $z\in\Omega$, we obtain

$$((\lambda I-A)f)(z)=\lambda f(z)-m(z)f(z)=0$$

for every $f\in\mathrm{D}(A)$. Consequently, $\lambda I-A$ is not surjective and so $\lambda\in\sigma(A)$. We now conclude that $\sigma(A)=\overline{m(\Omega)}$ since the spectrum is closed.

The final assertion follows from Example B.3 if $S=\emptyset$. Otherwise, consider $\Omega=S$ and $m(z)=z$. Define A and X as above. Then, $\sigma(A)=S$ and A is bounded if S is compact (where $C_b(S)=C(S)$). □

A similar result holds on L^p-spaces, see e.g. Example IX.2.6 in [**Con90**]. The next theorem gives additional properties of the spectrum of bounded operators.

THEOREM B.6. *Let $T \in \mathcal{B}(X)$. Then $\sigma(T)$ is a non-empty compact set. The* spectral radius $r(T) := \max\{|\lambda| \mid \lambda \in \sigma(T)\}$ *is given by*

$$r(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \inf_{n\in\mathbb{N}} \|T^n\|^{\frac{1}{n}} \le \|T\|,$$

and for $\lambda \in \mathbb{C}$ with $|\lambda| > r(T)$ we have

$$R(\lambda, T) = \sum_{n=0}^{\infty} \lambda^{-n-1} T^n =: R_\lambda.$$

PROOF. 1) Since $\|T^{n+m}\| \le \|T^n\| \, \|T^m\|$ for all $n, m \in \mathbb{N}$, an elementary lemma (see Lemma VI.1.4 in [**Wer07**]) yields that there exists

$$\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \inf_{n\in\mathbb{N}} \|T^n\|^{\frac{1}{n}} =: r \le \|T\|.$$

If $|\lambda| > r$, then

$$\limsup_{n\to\infty} \|\lambda^{-n} T^n\|^{\frac{1}{n}} = \frac{1}{|\lambda|} \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \frac{r}{|\lambda|} < 1.$$

Therefore the series R_λ converges absolutely in $\mathcal{B}(X)$, and uniformly for λ in all compact subsets of $\mathbb{C} \setminus B(0, r)$ (proof as in the one dimensional case). Moreover,

$$(\lambda I - T) R_\lambda = \sum_{n=0}^{\infty} \lambda^{-n} T^n - \sum_{n=0}^{\infty} \lambda^{-n-1} T^{n+1} = I,$$

and similarly $R_\lambda(\lambda I - T) = I$. Hence, $\lambda \in \rho(T)$ and $R_\lambda = R(\lambda, T)$. Due to its closedness, the spectrum $\sigma(T) \subseteq \overline{B}(0, r)$ is compact. Therefore, $r(T)$ exists as the maximum of a compact subset of $\mathbb{R}$, and $r(T) \le r$.

2) Take $s > r(T)$, $\Phi \in \mathcal{B}(X)^*$, and $m \in \mathbb{N}$. We define $f_\Phi(\lambda) = \Phi(R(\lambda, T))$ for $\lambda \in \rho(T)$. Note that $f_\Phi : \rho(T) \to \mathbb{C}$ is complex differentiable. We set

$$C_m(\Phi) = \frac{1}{2\pi i} \int_{|\lambda|=s} \lambda^m \Phi(R(\lambda, T)) \, \mathrm{d}\lambda.$$

Since f_Φ is holomorphic, this integral does not depend on $s > r(T)$ due to complex analysis. So we may choose for a moment $s > r$ and use the uniformly convergent series of step 1) to deduce

$$\begin{aligned} C_m(\Phi) &= \sum_{n=0}^{\infty} \frac{1}{2\pi i} \int_{|\lambda|=s} \lambda^{m-n-1} \, \mathrm{d}\lambda \; \Phi(T^n) \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi i} \int_0^{2\pi} (s\mathrm{e}^{it})^{m-n-1} i s\mathrm{e}^{it} \, \mathrm{d}t \; \Phi(T^n) = \Phi(T^m). \end{aligned}$$

The Hahn-Banach theorem yields a functional $\Phi_m \in \mathcal{B}(X)^*$ with $\|\Phi_m\| = 1$ and $\Phi_m(T^m) = \|T^m\|$. Again for any $s > r(T)$, we can then estimate

$$\begin{aligned} \|T^m\| = \Phi_m(T^m) = C_m(\Phi_m) &\le \frac{1}{2\pi} \int_0^{2\pi} |s\mathrm{e}^{it}|^m \, \|\Phi_m\| \, \|R(s\mathrm{e}^{it}, T)\| |s\mathrm{e}^{it}| \, \mathrm{d}t \\ &\le s^m s \max_{|\lambda|=s} \|R(\lambda, T)\| =: c(s) s^m. \end{aligned}$$

Thus, $\|T^m\|^{\frac{1}{m}} \le s c(s)^{\frac{1}{m}}$ and so $r \le s$. This means that $r(T) = r$.

Finally, suppose that $\sigma(T) = \emptyset$. Then the functions f_Φ are entire for every $\Phi \in \mathcal{B}(X)^*$. Moreover, step 1) yields

$$|f_\Phi(\lambda)| \leq \|\Phi\||\lambda|^{-1} \sum_{n=0}^{\infty} \frac{\|T\|^n}{|\lambda|^n} \leq \frac{2\|\Phi\|}{|\lambda|},$$

for all $\lambda \in \mathbb{C}$ with $|\lambda| \geq 2\|T\|$. Hence, f_Φ is bounded and thus constant by Liouville's theorem from complex analysis. The above estimate then shows that $\Phi(R(\lambda, T)) = 0$ for all $\lambda \in \mathbb{C}$ and $\Phi \in \mathcal{B}(X)^*$. Using again the Hahn-Banach theorem, we obtain $R(\lambda, T) = 0$, which is impossible since $R(\lambda, T)$ is injective and $X \neq \{0\}$. □

It can happen that $r(T) < \|T\|$. Take for instance a Jordan block $T = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ on $X = \mathbb{C}^2$, where $r(T) = 0$ and $\|T\| = 1$.

In finite dimensions we have $\sigma(T) = \sigma_p(T)$. The following examples indicate that in infinite dimensions $\sigma_p(T)$ can be empty or much smaller than $\sigma(T)$.

EXAMPLE B.7. (a) Let $X = C([0, 1])$ and define the Volterra operator V by

$$(Vf)(t) = \int_0^t f(s) \, \mathrm{d}s$$

for $t \in [0, 1]$ and $f \in X$. Then $V \in \mathcal{B}(X)$ and

$$|V^n f(t)| \leq \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} \|f\|_\infty \, \mathrm{d}s_n \ldots \mathrm{d}s_1 \leq \frac{1}{n!} \|f\|_\infty,$$

for all $n \in \mathbb{N}$, $t \in [0, 1]$, and $f \in X$. Hence, $\|V^n\| \leq \frac{1}{n!}$. For $f = \mathbb{1}$ we obtain $\|V^n\| \geq \|V^n \mathbb{1}\|_\infty = \frac{1}{n!}$ and so $\|V^n\| = \frac{1}{n!}$. This gives

$$r(V) = \lim_{n \to \infty} (\tfrac{1}{n!})^{\frac{1}{n}} = 0 \quad \text{and} \quad \sigma(V) = \{0\}.$$

Observe that $\sigma_p(V) = \emptyset$ since $Vf = 0$ implies that $f = (Vf)' = 0$. ♢

(b) The right shift R given by $R(x_n)_n = (0, x_1, x_2, \ldots)$ on $X \in \{c_0, \ell^p \,|\, p \in [1, \infty]\}$ has the spectrum and point spectrum

$$\sigma(R) = \overline{B(0, 1)} \quad \text{and} \quad \sigma_p(R) = \emptyset.$$

In fact, $R \in \mathcal{B}(X)$ has norm 1 and so $\sigma(R) \subseteq \overline{B(0, 1)}$. If there are $x \in X$ and $\mu \in \mathbb{C}$ such that $\mu x = Rx$, then $\mu x_1 = 0$, $\mu x_2 = x_1$, $\mu x_3 = x_2$, ... Hence, $x = 0$ and we conclude that $\sigma_p(R) = \emptyset$. Let $\lambda \in B(0, 1) \setminus \{0\}$ and assume that $\lambda I - A$ is surjective. Then there would exist a vector $x \in X$ such that $\lambda x - Rx = e_1$, i.e., $\lambda x_1 = 1$, $\lambda x_2 = x_1$, $\lambda x_3 = x_2$, ... It follows that x is equal to $(\lambda^{-n})_n$ which does not belong to X. This means that $\lambda I - A$ is not surjective and therefore $B(0, 1) \setminus \{0\} \subseteq \sigma(R)$. The closedness of the spectrum thus implies $\sigma(R) = \overline{B(0, 1)}$. ♢

(c) The left shift L given by $L(x_n)_n = (x_{n+1})_n$ on $X \in \{c_0, \ell^p \,|\, p \in [1, \infty]\}$ has the spectrum $\sigma(L) = \overline{B(0, 1)}$. In fact, $L \in \mathcal{B}(X)$ has norm 1, and so $\sigma(L) \subset \overline{B(0, 1)}$. Moreover, $L(1, 0, \cdots) = 0$, and for $|\lambda| < 1$ the sequence $v = (\lambda^n)_{n \in \mathbb{N}}$ belongs to X and satisfies $\lambda v = Lv$ so that $B(0, 1) \subseteq \sigma_p(L) \subseteq \sigma(L)$. The closedness of $\sigma(L)$ then yields $\sigma(L) = \overline{B(0, 1)}$. Note that $\sigma_p(L) = \overline{B(0, 1)}$ if $X = \ell^\infty$, but $\sigma_p(L) = B(0, 1)$ in the other cases. ♢

APPENDIX C

More spectral theory and selfadjoint operators

In courses on linear algebra it is shown that Hermitian matrices can be diagonalized, which makes their analysis rather simple. Selfadjoint operators are the natural extension of the concept of Hermitian matrices to infinite dimensional Hilbert spaces. We will not discuss the generalization of the diagonalization theorem to infinite dimensions, the so–called spectral theorem. Instead we establish several strong spectral theoretic properties of selfadjoint operators needed in the lectures. To this aim, we need a bit more general spectral theory on Banach spaces which is presented first.

1. Subdivisions of the spectrum and adjoint operators

We recall that a number $\lambda \in \mathbb{C}$ belongs to the spectrum of a closed operator A if $\lambda I - A$ is not bijective. This can happen in various ways which leads to several possibilities to subdivide the spectrum. We use one that links the spectrum to eigenvectors as far as possible. We start with the relevant definition.

Definition C.1. *Let A be a closed operator on X. Then we call*

$$\sigma_{ap}(A) = \{\lambda \in \mathbb{C} \,|\, \text{there exist } x_n \in \mathrm{D}(A) \text{ with } \|x_n\| = 1 \text{ for all } n \in \mathbb{N} \\ \text{and } \lambda x_n - A x_n \to 0 \text{ as } n \to \infty\}$$

the approximate point spectrum *of A and*

$$\sigma_r(A) = \{\lambda \in \mathbb{C} \,|\, (\lambda I - A)\,\mathrm{D}(A) \text{ is not dense in } X\}$$

the residual spectrum *of A.*

Proposition C.2. *For a closed operator A the following assertions hold.*

(a) $\sigma_{ap}(A) = \sigma_p(A) \cup \{\lambda \in \mathbb{C} \,|\, (\lambda I - A)\,\mathrm{D}(A)$ is not closed in $X\}$.
(b) $\sigma(A) = \sigma_{ap}(A) \cup \sigma_r(A)$.
(c) $\partial\sigma(A) \subset \sigma_{ap}(A)$.

(Note that the unions need not be disjoint.)

Proof. (a) We have $\lambda \notin \sigma_{\mathrm{ap}}(A)$ if and only if there is a constant $c > 0$ such that $\|(\lambda I - A)x\| \geq c\|x\|$ for all $x \in \mathrm{D}(A)$. This lower estimate implies that $\lambda \notin \sigma_p(A)$. Moreover, if $y_n := \lambda x_n - A x_n \to y$ in X as $n \to \infty$ for a sequence $(x_n)_n$ in $\mathrm{D}(A)$, then the lower estimate shows that $(x_n)_n$ is Cauchy in X, and thus x_n converges to some $x \in X$. Hence, $A x_n = \lambda x_n - y_n \to \lambda x - y$ and the closedness of A yields $x \in \mathrm{D}(A)$ and $\lambda x - A x = y$. Consequently, $(\lambda I - A)\,\mathrm{D}(A)$ is closed.

Conversely, if $(\lambda I - A)\,\mathrm{D}(A)$ is closed and $\lambda \notin \sigma_p(A)$, then the inverse $(\lambda I - A)^{-1}$ exists and is closed on its closed domain $(\lambda I - A)\,\mathrm{D}(A)$. The closed graph theorem A.7 then yields the boundedness of $(\lambda I - A)^{-1}$. Thus,

$$\|x\| = \|(\lambda I - A)^{-1}(\lambda I - A)x\| \leq C\|(\lambda I - A)x\|$$

for all $x \in \mathrm{D}(A)$ and a constant $C > 0$. This means that $\lambda \notin \sigma_{\mathrm{ap}}(A)$.

(b) Assertion (b) follows from (a).

(c) Let $\lambda \in \partial\sigma(A)$. Then there exist $\lambda_n \in \rho(A)$ with $\lambda_n \to \lambda$ as $n \to \infty$. Theorem B.4 yields $\|R(\lambda_n, A)\| \to \infty$ as $n \to \infty$, and thus there are $y_n \in X$ such that $\|y_n\| = 1$ for all $n \in \mathbb{N}$ and $a_n := \|R(\lambda_n, A)y_n\| \to \infty$ as $n \to \infty$, where we can assume that $a_n > 0$ for all $n \in \mathbb{N}$. Set $x_n = \frac{1}{a_n}R(\lambda_n, A)y_n \in \mathrm{D}(A)$. We then have $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $\lambda x_n - Ax_n = (\lambda - \lambda_n)x_n + \frac{1}{a_n}y_n \to 0$ as $n \to \infty$. As a result, $\lambda \in \sigma_{\mathrm{ap}}(A)$. □

We already used the above proposition implicitly in Example 3.5 (a). To describe the residual spectrum , we will use the adjoint of A.

Definition C.3. *Let X and Y be Banach spaces. Let A be a linear operator from X to Y with dense domain. We define its* adjoint A^* *on Y^* by setting*

$$\begin{aligned} \mathrm{D}(A^*) &= \{y^* \in Y^* \,|\, \exists\, z^* \in X^* \;\forall\, x \in \mathrm{D}(A) : \langle Ax, y^*\rangle = \langle x, z^*\rangle\}\,, \\ A^*y^* &= z^*. \end{aligned}$$

Observe that $\langle Ax, y^*\rangle = \langle x, A^*y^*\rangle$ holds for all $x \in \mathrm{D}(A)$ and $y^* \in \mathrm{D}(A^*)$. Several properties of the adjoint are collected in the next remark.

Remark C.4. Let A be a densely defined linear operator from X to Y.

(a) Since $\mathrm{D}(A)$ is dense, there is at most one vector z^* as in the definition of $\mathrm{D}(A^*)$, so that $A^* : \mathrm{D}(A^*) \to X^*$ is a map. It is clear that A^* is linear. If $A \in \mathcal{B}(X, Y)$, then $\mathrm{D}(A^*) = Y^*$ and Definition C.3 coincides with the definition of the adjoint usually given in courses of functional analysis. ◊

(b) The operator A^* is closed from Y^* to X^*. In fact, let $y_n^* \in \mathrm{D}(A^*)$, $y^* \in Y^*$, and $z^* \in X^*$ such that $y_n^* \to y^*$ in Y^* and $z_n^* := A^*y_n^* \to z^*$ in X^* as $n \to \infty$. Then for every $x \in \mathrm{D}(A)$ the equation,

$$\langle x, z^*\rangle = \lim_{n\to\infty}\langle x, z_n^*\rangle = \lim_{n\to\infty}\langle Ax, y_n^*\rangle = \langle Ax, y^*\rangle$$

hold. As a result, $y^* \in \mathrm{D}(A^*)$ and $A^*y^* = z^*$. ◊

(c) Let $X = Y$ for simplicity. If $T \in \mathcal{B}(X)$, then the sum $A + T$ and the product TA with domains $\mathrm{D}(A + T) = \mathrm{D}(A)$ and $\mathrm{D}(TA) = \mathrm{D}(A)$ have the adjoints $(A + T)^* = A^* + T^*$ and A^*T^* with $\mathrm{D}(A^* + T^*) = \mathrm{D}(A^*)$ and $\mathrm{D}(A^*T^*) = \{x^* \in X^* \,|\, T^*x^* \in \mathrm{D}(A^*)\}$, respectively. To verify the first claim, let $x \in \mathrm{D}(A)$ and $x^* \in X^*$. We calculate

$$\langle (A + T)x, x^*\rangle = \langle Ax, x^*\rangle + \langle x, T^*x^*\rangle.$$

Hence, $x^* \in \mathrm{D}((A + T)^*)$ if and only if $x^* \in \mathrm{D}(A^*)$, and then $(A + T)^*x^* = A^*x^* + T^*x^*$. The second fact can be shown analogously. ◊

The next result shows in particular that the eigenvalues of A^* give the residual spectrum of A, cf. the proof of Corollary 5.2.

Theorem C.5. *Let A be a closed operator on X with dense domain. Then the following assertions hold.*

(a) $\sigma_r(A) = \sigma_p(A^*)$.

(b) $\sigma(A) = \sigma(A^*)$ *and* $R(\lambda, A)^* = R(\lambda, A^*)$ *for every* $\lambda \in \rho(A)$.

Proof. (a) Due to a corollary of the Hahn-Banach theorem, the set $(\lambda I - A)\,\mathrm{D}(A)$ is not dense in X if and only if there is a vector $y^* \in X^* \setminus \{0\}$ such that $\langle \lambda x - Ax, y^* \rangle = 0 = \langle x, 0 \rangle$ for every $x \in \mathrm{D}(A)$. This fact is equivalent to $y^* \in \mathrm{D}(A^*) \setminus \{0\}$ and $\lambda y^* - A^* y^* = 0$, i.e., $\lambda \in \sigma_p(A^*)$.

(b) Let $\lambda \in \rho(A)$. Take $x^* \in X^*$, and set $y^* = R(\lambda, A)^* x^*$. We then obtain

$$\langle (\lambda I - A)x, y^* \rangle = \langle R(\lambda, A)(\lambda I - A)x, x^* \rangle = \langle x, x^* \rangle,$$

for all $x \in \mathrm{D}(A)$. Thus, $y^* \in \mathrm{D}(A^*)$ and $x^* = (\lambda I - A)^* y^* = (\lambda I - A^*) y^*$, where we use Remark C.4. This means that $\lambda I - A^*$ is surjective. Further, let $x^* \in \mathrm{D}(A^*)$ and $x \in X$. We compute

$$\begin{aligned} \langle x, R(\lambda, A)^*(\lambda I - A^*)x^* \rangle &= \langle R(\lambda, A)x, (\lambda I - A^*)x^* \rangle \\ &= \langle (\lambda I - A)R(\lambda, A)x, x^* \rangle = \langle x, x^* \rangle, \end{aligned}$$

using Definition C.3 and that $R(\lambda, A)x$ belongs to $\mathrm{D}(A)$. Hence, $R(\lambda, A)^*(\lambda I - A^*)x^* = x^*$, so that $\lambda I - A^*$ is also injective. It thus exists $R(\lambda, A^*) = R(\lambda, A)^*$.

Conversely, let $\lambda \in \rho(A^*)$. Take $x \in \mathrm{D}(A)$. For every $x^* \in X^*$, we calculate as above

$$\langle (\lambda I - A)x, R(\lambda, A^*)x^* \rangle = \langle x, (\lambda I - A^*)R(\lambda, A^*)x^* \rangle = \langle x, x^* \rangle.$$

Due to a corollary of the Hahn-Banach theorem, there is a functional $y^* \in X^*$ such that $\|y^*\| = 1$ and $\langle x, y^* \rangle = \|x\|$. Hence,

$$\|x\| = \langle (\lambda I - A)x, R(\lambda, A^*)y^* \rangle \le \|R(\lambda, A^*)\| \|\lambda x - Ax\|.$$

This estimate implies that $\lambda \notin \sigma_{\mathrm{ap}}(A)$. Further, $\lambda \notin \sigma_p(A^*) = \sigma_r(A)$ by part (a), and so Proposition C.2 shows that $\lambda \notin \sigma(A)$. ☐

Example C.6. Let $X = \ell^p$, $1 \le p < \infty$, or $X = c_0$. Let $Rx = (0, x_1, x_2, \dots)$ be the right shift on X. Then $R^* = L$, where the left shift L acts on ℓ^1 if $X = c_0$ and on $\ell^{p'}$ otherwise. Since $\sigma(L) = \overline{B(0,1)}$ by Example B.7 (c), Theorem C.5 yields

$$\sigma(R) = \sigma(L) = \overline{B(0,1)}.$$

In this way we obtain a faster proof of a part of Example B.7 (b). Moreover, $\sigma_r(R) = \sigma_p(L) = B(0,1)$ if $X = c_0$ or $X = \ell^p$ with $1 < p < \infty$ and $\sigma_r(R) = \sigma_p(L) = \overline{B(0,1)}$ if $X = \ell^1$.

If $X = \ell^\infty$, then we consider L on ℓ^1 and R on $\ell^\infty = (\ell^1)^*$. Since $R = L^*$, we again deduce $\sigma(R) = \overline{B(0,1)}$.

2. Selfadjoint operators and their spectra

Let X and Y be Hilbert spaces with scalar product $(\cdot|\cdot)_X$ and $(\cdot|\cdot)_Y$. Let A be a linear densely defined operator from X to Y. We define the *Hilbert space adjoint* A' as in the Banach space case by

$$\mathrm{D}(A') = \{ y \in Y \mid \exists\, z \in X \ \forall\, x \in \mathrm{D}(A) : (Ax|y)_Y = (x|z)_X \}, \quad A'y := z.$$

As in Remark C.4 one sees that A' is well-defined, linear and closed. Also Remark C.4 (c) holds analogously.

We say that $T \in \mathcal{B}(X,Y)$ is *unitary* if T is invertible with $T^{-1} = T'$. A densely defined linear operator A is called *selfadjoint* if $A = A'$ and *skewadjoint* if $A = -A'$ (in particular, $\mathrm{D}(A) = \mathrm{D}(A')$ and A must be closed). Finally, A is *normal* if $AA' = A'A$. Note that if A is selfadjoint or skewadjoint, then $\lambda I - A$ is normal for $\lambda \in \mathbb{C}$.

A linear operator A on X is called *symmetric* if we have $(Ax|y) = (x|Ay)$ for all $x, y \in \mathrm{D}(A)$. We note that a densely defined linear operator is symmetric if and only if $A \subseteq A'$.

Let $\Phi : X \to X^*$ be the antilinear Riesz isomorphism given by $(\Phi(x))(z) = (z|x)$ for all $x, z \in X$. For $\lambda \in \mathbb{C}$ and $X = Y$, we then obtain

$$\lambda I_{X^*} - A^* = \Phi(\overline{\lambda} I_X - A')\Phi^{-1}, \tag{C.1}$$

with $\mathrm{D}(A^*) = \Phi\,\mathrm{D}(A')$. So Theorem C.5 implies that

$$\sigma(A) = \sigma(A^*) = \overline{\sigma}(A') \quad \text{and} \quad \sigma_r(A) = \sigma_p(A^*) = \overline{\sigma}_p(A'), \tag{C.2}$$

where the bars mean complex conjugation. Moreover,

$$R(\overline{\lambda}, A') = \Phi^{-1}R(\lambda, A^*)\Phi = \Phi^{-1}R(\lambda, A)^*\Phi = R(\lambda, A)',$$

for all $\lambda \in \rho(A)$.

Recall from functional analysis that $\|T'\| = \|T\|$ for every $T \in \mathcal{B}(X,Y)$, $T'' = T$ and $(ST)' = T'S'$ for $S \in \mathcal{B}(Y,Z)$ where Z is another Hilbert space. We first establish an important characterization of unitary operators, used e.g. in Lecture 5.

Proposition C.7. *Let X and Y be Hilbert spaces and $T \in \mathcal{B}(X,Y)$. Then the following assertions hold.*

(a) T is a isometry if and only if we have $(Tx|Tz)_Y = (x|z)_X$ for all $x, z \in X$.

(b) T is unitary if and only if T is bijective and isometric if and only if T is bijective and preserves the scalar product.

Proof. (a) The implication "$\Leftarrow$" is shown by setting $x = z$. To verify "$\Rightarrow$", take $\alpha \in \mathbb{C}$ and $x, z \in X$. Using that T is isometric, we calculate

$$\begin{aligned} (T(x+\alpha z)|T(x+\alpha z)) &= \|Tx\|^2 + 2\,\mathrm{Re}(Tx|\alpha Tz) + \|\alpha Tz\|^2 \\ &= \|x\|^2 + 2\,\mathrm{Re}(\overline{\alpha}(Tx|Tz)) + |\alpha|^2\|z\|^2, \\ \|T(x+\alpha z)\|^2 &= (x+\alpha z|x+\alpha z) = \|x\|^2 + 2\,\mathrm{Re}(\overline{\alpha}(x|z)) + |\alpha|^2\|z\|^2. \end{aligned}$$

Since $(T(x+\alpha z)|T(x+\alpha z)) = \|T(x+\alpha z)\|^2$, it follows $\mathrm{Re}(\overline{\alpha}(Tx|Tz)) = \mathrm{Re}(\overline{\alpha}(x|z))$. Choosing $\alpha = 1$ and $\alpha = i$, we deduce assertion (a).

(b) The second equivalence is a consequence of part (a). To show the first equivalence, take $x, z \in X$. If T is unitary, we obtain $(Tx|Tz) = (x|T^{-1}Tz) = (x|z)$, so that T is isometric by (a). If T is isometric, (a) yields $(T'Tx|z) = (Tx|Tz) = (x|z)$. Since $z \in X$ is arbitrary, we conclude that $T'Tx = x$ for all $x \in X$ and hence $T'T = I$. Now, the bijectivity of T implies that $T' = T^{-1}$. $\square$

The next lemma gives the crucial estimates for the proof of the announced spectral properties of selfadjoint operators stated below.

Lemma C.8. *Let A be symmetric. Set $\lambda = \alpha + \mathrm{i}\beta$ for any $\alpha, \beta \in \mathbb{R}$. Let $x \in \mathrm{D}(A)$. It then holds $(Ax|x) \in \mathbb{R}$ and*

$$\|\lambda x - Ax\|^2 = \|\alpha x - Ax\|^2 + |\beta|^2\|x\|^2 \geq |\beta|^2\|x\|^2.$$

Let A also be closed. Then $\sigma_{ap}(A) \subseteq \mathbb{R}$. If $\lambda \in \rho(A)\backslash\mathbb{R}$, then $\|R(\lambda, A)\| \leq \frac{1}{|\operatorname{Im}\lambda|}$.

Proof. For $x \in \mathrm{D}(A)$ we have $(Ax|x) = (x|Ax) = \overline{(Ax|x)}$ so that $(Ax|x) = (x|Ax)$ is real. From this fact we deduce that

$$\begin{aligned}
\|\lambda x - Ax\|^2 &= (\alpha x - Ax + \mathrm{i}\beta x|\alpha x - Ax + \mathrm{i}\beta x) \\
&= \|\alpha x - Ax\|^2 + 2\operatorname{Re}(\mathrm{i}\beta x|\alpha x - Ax) + \|\mathrm{i}\beta x\|^2 \\
&= \|\alpha x - Ax\|^2 + 2\operatorname{Re}\big(\mathrm{i}\beta\alpha\|x\|^2 - \mathrm{i}\beta(x|Ax)\big) + |\beta|^2\|x\|^2 \\
&= \|\alpha x - Ax\|^2 + |\beta|^2\|x\|^2 \geq |\beta|^2\|x\|^2.
\end{aligned}$$

In particular, $\lambda \notin \sigma_{ap}(A)$ if $\operatorname{Im}\lambda = \beta \neq 0$.

If $\lambda \in \rho(A) \setminus \mathbb{R}$ and $y \in X$, write $x = R(\lambda, A)y \in \mathrm{D}(A)$. We then calculate

$$\|y\|^2 = \|\lambda x - Ax\|^2 \geq |\operatorname{Im}\lambda|^2\|x\|^2 = |\operatorname{Im}\lambda|^2\|R(\lambda, A)y\|^2$$

which yields the final inequality in the lemma. □

Theorem C.9. *Let X be a Hilbert space and A be densely defined, closed and symmetric. Then the following assertions hold.*

- *(a) The spectrum $\sigma(A)$ of A is either a subset of $\mathbb{R}$ or $\sigma(A) = \mathbb{C}$ or $\sigma(A) = \{\lambda \in \mathbb{C} \mid \operatorname{Im}\lambda \geq 0\}$ or $\sigma(A) = \{\lambda \in \mathbb{C} \mid \operatorname{Im}\lambda \leq 0\}$.*
- *(b) The following assertions are equivalent.*
  - *(1) $A = A'$,*
  - *(2) $\sigma(A) \subseteq \mathbb{R}$,*
  - *(3) $\mathrm{i}I - A'$ and $\mathrm{i}I + A'$ are injective,*
  - *(4) $(\mathrm{i}I - A)\,\mathrm{D}(A)$ and $(\mathrm{i}I + A)\,\mathrm{D}(A)$ are dense.*
- *(c) Let A be selfadjoint. Then we have*

  $$\|R(\lambda, A)\| \leq \frac{1}{|\operatorname{Im}\lambda|}$$

  for $\lambda \notin \mathbb{R}$. Further, $\sigma(A) = \sigma_{ap}(A)$ and A has no selfadjoint extension $B \neq A$.

Proof. (a) Suppose that there would exists $\lambda \in \sigma(A)$ and $\mu \in \rho(A)$ with $\operatorname{Im}\lambda \cdot \operatorname{Im}\mu > 0$. The line segment from λ to μ must contain a point $\gamma \in \partial\sigma(A)$. Then, $\operatorname{Im}\gamma \neq 0$ and $\gamma \in \sigma_{ap}(A)$ by Proposition C.2, which contradicts Lemma C.8 since A is symmetric. Assertion (a) thus follows from the closedness of the spectrum.

(b) Let A be selfadjoint. Lemma C.8 yields $\sigma_p(A) \subseteq \sigma_{ap}(A) \subseteq \mathbb{R}$. Due to (C.2) we also have $\sigma_r(A) = \overline{\sigma}_p(A') = \overline{\sigma}_p(A) = \sigma_p(A) \subseteq \mathbb{R}$. From Proposition C.2 we thus deduce $\sigma(A) \subseteq \mathbb{R}$, i.e., (1) implies (2). The implication "(2) ⇒ (3)" is obvious. Equation (C.2) also shows that $\pm\mathrm{i} \in \sigma_p(A')$ if and only if $\mp\mathrm{i} \in \sigma_r(A)$ so that (3) and (4) are equivalent. Let (4) (and thus (3)) hold. Due to Lemma C.8 and Proposition C.2, the range of $\mathrm{i}I \pm A$ is closed. In view of (4), $\mathrm{i}I \pm A$ is then surjective. Due to (3), $\mathrm{i}I \pm A'$ is injective, and hence $A = A'$ thanks to $A \subseteq A'$ and Lemma 3.6. So the implication "(4) $\Rightarrow$ (1)" holds.

(c) Let $A = A'$. Then $\sigma(A) = \sigma(A') \subseteq \mathbb{R}$ so that assertion (c) follows from Lemma C.8, Proposition C.2 (c) and Lemma 3.6. $\square$

We illustrate the above result by a simple example. For the Sobolev spaces we refer to Intermezzo 3 in Lecture 5 or Appendix D.

EXAMPLE C.10. (a) Let $X = L^2(0,\infty)$ and $Au = \mathrm{i}u'$ for $u \in \mathrm{D}(A) = \mathring{W}_2^1(0,\infty)$. We first show that A is closed. Let $u_n \in \mathrm{D}(A)$ such that $u_n \to u$ and $\mathrm{i}u_n' \to v$ in X. Lemma 3.6 (c) then yields that $u \in W_2^1(0,\infty)$ and $\mathrm{i}u' = v$. Since $\mathring{W}_2^1(0,\infty)$ is closed in $W_2^1(0,\infty)$, we also have $u \in \mathrm{D}(A)$.

To see the symmetry, take $u \in \mathrm{D}(A)$ and $v \in C_c^\infty(0,\infty)$. We have

$$(Au|v) = \mathrm{i}\int_0^\infty u'\overline{v}\,\mathrm{d}x = -\mathrm{i}\int_0^\infty u\overline{v}'\,\mathrm{d}x = \int_0^\infty u\overline{\mathrm{i}v}'\,\mathrm{d}x = (u|Av).$$

For $v \in \mathrm{D}(A)$ there are $v_n \in C_c^\infty(0,\infty)$ converging to v in $W_2^1(0,\infty)$, so that $v_n \to v$ and $Av_n \to Av$ in X as $n \to \infty$. Hence, $(Au|v) = (u|Av)$ for all $u, v \in \mathrm{D}(A)$, i.e., A is symmetric. By Exercise 5.4, $\sigma(\mathrm{i}A) = \{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda \le 0\}$ so that $\sigma(A) = -\mathrm{i}\sigma(\mathrm{i}A) = \{\lambda \in \mathbb{C} \mid \operatorname{Im}\lambda \ge 0\}$ (use Exercise 2.5). Consequently, A is not selfadjoint. $\diamondsuit$

(b) Let $X = L^2(\mathbb{R})$ and $Au = \mathrm{i}u'$ for $u \in \mathrm{D}(A) = W_2^1(\mathbb{R})$. As in part (a) one shows that A is symmetric. Exercises 2.5 and 3.2 imply that $\sigma(A) = \mathrm{i}\sigma(-\mathrm{i}A) = \mathrm{i}^2\mathbb{R} = \mathbb{R}$. Hence, A is selfadjoint. $\diamondsuit$

APPENDIX D

Sobolev spaces and weak derivatives

Throughout, $U \subseteq \mathbb{R}^d$ is open and non-empty.

1. Basic properties

We are looking for properties of C^1 function and their derivatives which can be generalized to a concept of derivatives suited to L^p spaces, which is in particular not based on pointwise limits. To that purpose, take $f \in C^1(U)$ and $\varphi \in C_c^\infty(U)$. There is an open set V in $\mathbb{R}^d$ such that

$$\operatorname{supp}\varphi \subseteq V \subseteq \overline{V} \subseteq U.$$

We write $x \in V$ as $x = (x_1, y)$ for $x_1 \in V_y := \{t \in \mathbb{R} \,|\, (t, y) \in V\}$ and $y \in \mathbb{R}^{d-1}$ with $V_y \neq \emptyset$. Observe that $V_y \subseteq \mathbb{R}$ is open and thus a disjoint union of open intervals. Integrating by parts, we deduce

$$\begin{aligned}\int_U (\partial_1 f)\varphi \,\mathrm{d}x &= \int_{y: V_y \neq \emptyset} \int_{V_y} \partial_1 f(x_1, y)\varphi(x_1, y)\,\mathrm{d}x_1\,\mathrm{d}y \\ &= \int_{y: V_y \neq \emptyset} \Big(- \int_{V_y} f(x_1, y)\partial_1\varphi(x_1, y)\,\mathrm{d}x_1 + [f(x_1, y)\varphi(x_1, y)]|_{x_1 \in \partial V_y} \Big)\,\mathrm{d}y \\ &= -\int_{y: V_y \neq \emptyset} \int_{V_y} f(x_1, y)\partial_1\varphi(x_1, y)\,\mathrm{d}x_1\,\mathrm{d}y = -\int_U f\partial_1\varphi\,\mathrm{d}x,\end{aligned}$$

since $\varphi = 0$ on ∂V_y for each $y \in \mathbb{R}^{d-1}$ with $V_y \neq \emptyset$. Inductively one shows that

$$\int_U (\partial^\alpha f)\varphi\,\mathrm{d}x = (-1)^{|\alpha|}\int_U f\partial^\alpha\varphi\,\mathrm{d}x, \tag{D.1}$$

for all $f \in C^k(U)$, $\varphi \in C_c^\infty(U)$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$.

Here and below we use the *multi index notation*: Let $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ and $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. We then set

$$|\alpha| := \alpha_1 + \ldots + \alpha_d, \quad x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_d^{\alpha_d}, \quad \partial^\alpha := \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}\,.$$

For any f mapping $\mathbb{R}^d$ to $\mathbb{C}$ we further denote the function $\mathbb{R}^d \to \mathbb{C}$, $x \mapsto x^\alpha f(x)$, by $x^\alpha f$. We set

$$L^1_{\mathrm{loc}}(U) = \Big\{ f : U \to \mathbb{C} \,|\, f \text{ is measurable, } f|_K \in L^1(K) \text{ for all compact } K \subseteq U \Big\}.$$

We extend $f \in L^1_{\mathrm{loc}}(U)$ by 0 to a measurable function $f : \mathbb{R}^d \to \mathbb{C}$ without notice. Convergence in $L^1_{\mathrm{loc}}(U)$ means that $(f_n|_K)_n$ converges in $L^1(K)$ for each compact subset K of U. Observe that $L^p(U) \subseteq L^1_{\mathrm{loc}}(U)$ for all $1 \leq p \leq \infty$.

Definition D.1. *Let $f \in L^1_{loc}(U)$ and $\alpha \in \mathbb{N}_0^d$. If there is a function $g \in L^1_{loc}(U)$ such that*

$$\int_U g\varphi \, \mathrm{d}x = (-1)^{|\alpha|} \int_U f \partial^\alpha \varphi \, \mathrm{d}x$$

for all $\varphi \in C_c^\infty(U)$, then g is called weak derivative *of f. We then use the symbol $\partial^\alpha f := g$. Let $W^k(U)$ be the space of all $f \in L^1_{loc}(U)$ which possess weak derivatives for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$. Moreover, one defines the* Sobolev spaces *by*

$$W_p^k(U) = \Big\{ f \in L^p(U) \cap W^k(U) \,\Big|\, \partial^\alpha f \in L^p(U) \text{ for all } |\alpha| \le k \Big\}$$

for $k \in \mathbb{N}$ and $1 \le p \le \infty$ and endows them with

$$\|f\|_{k,p} = \begin{cases} \Big(\sum\limits_{0 \le |\alpha| \le k} \|\partial^\alpha f\|_p^p \Big)^{1/p}, & 1 \le p < \infty, \\ \max_{0 \le |\alpha| \le k} \|\partial^\alpha f\|_\infty, & p = \infty, \end{cases}$$

where $\partial^0 f := f$. We write $W_p^0(U) = L^p(U)$ and $H^k(U) = W_2^k(U)$.

As usual, the spaces $L^1_{\mathrm{loc}}(U)$, $W^k(U)$ and $W_p^k(U)$ are spaces of equivalence classes modulo the subspace $\mathcal{N} = \{ f : U \to \mathbb{R} \,|\, f \text{ is measurable}, f = 0 \text{ a.e.} \}$ of null functions. We note that most of the following calculations and results become much simpler if one restricts to functions in $L^p(U)$ or even in $L^p(\mathbb{R}^d)$. But the present generality is needed in many applications.

Remark D.2. (a) We will see in Lemma D.5 that $\partial^\alpha f$ is uniquely defined. It is then also clear that for $|\alpha| \le k$ the map

$$\partial^\alpha : W^k(U) \to L^1_{\mathrm{loc}}(U)$$

is linear.

(b) Formula (D.1) implies that $C^k(U) + \mathcal{N} \subseteq W^k(U)$ and that weak and classical derivatives coincide for $f \in C^k(U)$.

(c) Let $1 \le p \le \infty$ and $k \in \mathbb{N}$. Clearly, $(W_p^k(U), \|\cdot\|_{k,p})$ is a normed vector space and

$$J : W_p^k(U) \to L^p(U)^m, \quad f \mapsto (\partial^\alpha f)_{|\alpha| \le k},$$

is an isometry, where m is the number of all multiindices α with $|\alpha| \le k$. We will see in the proof of the next proposition that $W_p^k(U)$ is isometrically isomorphic to a closed subspace of $L^p(U)^m$.

(d) Let $1 \le p \le \infty$ and $k \in \mathbb{N}$. Since the p–norm and the 1–norm on $\mathbb{R}^m$ are equivalent, there are constants $C_k, c_k > 0$ such that

$$c_k \sum_{0 \le |\alpha| \le k} \|\partial^\alpha f\|_p \le \|f\|_{k,p} \le C_k \sum_{0 \le |\alpha| \le k} \|\partial^\alpha f\|_p$$

for all $f \in W_p^k(U)$.

(e) Observe that $\|f\|_{1,p}^p = \|f\|_p^p + \| |\nabla f|_p \|_p^p$ for all $1 \le p < \infty$, where $|\cdot|_p$ is the p–norm on $\mathbb{C}^d$.

Proposition D.3. *For all $1 \le p \le \infty$ and $k \in \mathbb{N}$, $W^k_p(U)$ is a Banach space. It is separable if $p < \infty$ and reflexive if $1 < p < \infty$. Moreover, $W^k_2(U)$ is a Hilbert space endowed with the scalar product*

$$(f|g)_{W_2^k} = \sum_{|\alpha|\le k} \int_U (\partial^\alpha f)\overline{\partial^\alpha g}\,\mathrm{d}x.$$

Proof. Let $(f_n)_n$ be a Cauchy sequence in $W^k_p(U)$. Then $(\partial^\alpha f_n)_n$ is a Cauchy sequence in $L^p(U)$ for every $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$. Thus $\partial^\alpha f_n \to g_\alpha$ in $L^p(U)$ for some $g_\alpha \in L^p(U)$ as $n \to \infty$. We set $f := g_0$. Let $\varphi \in C_c^\infty(U)$ and $|\alpha| \le k$. Since $f_n \in W^k_p(U)$, we deduce

$$\begin{aligned}\int_U f\partial^\alpha\varphi\,\mathrm{d}x &= \lim_{n\to\infty}\int_U f_n\partial^\alpha\varphi\,\mathrm{d}x = \lim_{n\to\infty}(-1)^{|\alpha|}\int_U(\partial^\alpha f_n)\varphi\,\mathrm{d}x\\ &= (-1)^{|\alpha|}\int_U g_\alpha\varphi\,\mathrm{d}x.\end{aligned}$$

This means that g_α is the weak derivative $\partial^\alpha f$ so that $f \in W^k_p(U)$ and $f_n \to f$ in $W^k_p(U)$. Hence, $W^k_p(U)$ is a Banach space. We then deduce from Remark D.2 (c) that $W^k_p(U)$ is isometrically isomorphic to a closed subspace of $L^p(U)$. The remaining assertions now follow by isomorphy from known results of functional analysis. □

Example D.4. (a) Let $f \in C_c(\mathbb{R})$ be such that $f_\pm := f|_{\mathbb{R}_\pm}$ belong to $C^1(\mathbb{R}_\pm)$. We then have $f \in W^1(\mathbb{R})$ with

$$f' := \partial_1 f = \left\{\begin{matrix} f'_+ & \text{on } [0,\infty) \\ f'_- & \text{on } (-\infty,0) \end{matrix}\right\} =: g.$$

For $f(x) = |x|$, we thus obtain $f' = \mathbb{1}_{\mathbb{R}_+} - \mathbb{1}_{(-\infty,0)}$.

Proof. For every $\varphi \in C_c^\infty(\mathbb{R})$, we compute

$$\begin{aligned}\int_{\mathbb{R}} f\varphi'\,\mathrm{d}t &= \int_{-\infty}^0 f_-\varphi'\,\mathrm{d}t + \int_0^\infty f_+\varphi'\,\mathrm{d}t\\ &= -\int_{-\infty}^0 f'_-\varphi\,\mathrm{d}t + f_-\varphi|_{-\infty}^0 - \int_0^\infty f'_+\varphi\,\mathrm{d}t + f_+\varphi|_0^\infty\\ &= -\int_{\mathbb{R}} g\varphi\,\mathrm{d}t,\end{aligned}$$

since $f_+(0) = f_-(0)$ by the continuity of f. □

(b) The function $f = \mathbb{1}_{\mathbb{R}_+}$ does not belong to $W^1(\mathbb{R})$. Assume there would exist a function $g = f' \in L^1_{\mathrm{loc}}(\mathbb{R})$. Let $\varphi \in C_c^\infty(\mathbb{R})$. We compute

$$\int_{\mathbb{R}} g\varphi\,\mathrm{d}t = -\int_{\mathbb{R}} \mathbb{1}_{\mathbb{R}_+}\varphi'\,\mathrm{d}t = -\int_0^\infty \varphi'(t)\,\mathrm{d}t = \varphi(0).$$

Taking φ with $\operatorname{supp}\varphi \subseteq (0,\infty)$, we deduce from Lemma D.5 below that $g = 0$ on $(0,\infty)$. Similarly, it follows that $g = 0$ on $(-\infty,0)$. Hence, $g = 0$ and so $\varphi(0) = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$, which is false. ◇

(c) Let $d \geq 2$, $U = B_{\mathbb{R}^d}(0,1)$, $1 \leq p < d$, and $f(x) = \log |x|_2$ for $x \in U \setminus \{0\}$. Then we have $f \in W^1_p(U)$ with

$$\partial_j f(x) = \frac{x_j}{|x|_2^2} =: g_j(x), \quad u \in U \setminus \{0\}.$$

Observe that f is unbounded and has no continuous extension at $x = 0$.

PROOF. Using polar coordinates, we obtain

$$\|f\|_p^p = c \int_0^1 |\log r|^p r^{d-1} \, \mathrm{d}r < \infty,$$

for all $p \in [1, \infty)$. Estimating $|x_j| \leq r$, we further compute

$$\|g_j\|_p^p \leq c \int_0^1 \frac{r^p}{r^{2p}} r^{d-1} \, \mathrm{d}r = c \int_0^1 r^{d-p-1} \, \mathrm{d}r < \infty,$$

and thus $g_j \in L^p(U)$, for all $p \in [1, d)$ and $j \in \{1, \dots, d\}$. Take $j = 1$, $\varepsilon \in (0,1)$ and $\varphi \in C_c^\infty(U)$. We set $J_\varepsilon = (-1, -\varepsilon] \cup [\varepsilon, 1)$ and write $y = (x_2, \dots, x_d)$. We then obtain

$$\begin{aligned}
\int_U f \partial_1 \varphi \, \mathrm{d}x &= \int_{[-1,1]^d} f(x_1, y) \partial_1 \varphi(x_1, y) \, \mathrm{d}x_1 \, \mathrm{d}y \\
&= \lim_{\varepsilon \to 0} \int_{[-1,1]^{d-1}} \int_{J_\varepsilon} \Big(\log \sqrt{x_1^2 + |y|_2^2} \Big) \partial_1 \varphi(x_1, y) \, \mathrm{d}x_1 \, \mathrm{d}y \\
&= \lim_{\varepsilon \to 0} \int_{(-1,1)^{d-1}} \Big[- \int_{J_\varepsilon} \frac{x_1}{x_1^2 + |y|_2^2} \varphi(x_1, y) \, \mathrm{d}x_1 - \Big(\log \sqrt{\varepsilon^2 + |y|_2^2} \Big) \varphi(\varepsilon, y) \\
&\qquad + \Big(\log \sqrt{\varepsilon^2 + |y|_2^2} \Big) \varphi(-\varepsilon, y) \Big] \, \mathrm{d}y \\
&= - \int_U g_1 \varphi \, \mathrm{d}x.
\end{aligned}$$

Here we used the theorem of dominated convergence with majorants $\|\partial_1 \varphi\|_\infty |f|$ and $\|\varphi\|_\infty |g_1|$ and that

$$\Big| \Big(\log \sqrt{\varepsilon^2 + |y|_2^2} \Big) (\varphi(\varepsilon, y) - \varphi(-\varepsilon, y)) \Big| \leq 2\varepsilon \Big(|\log \varepsilon| + \log \sqrt{d} \Big) \|\partial_1 \varphi\|_\infty$$

converges to 0 as $\varepsilon \to 0$ uniformly in $y \in (-1,1)^{d-1}$. One similarly sees that $\partial_j f = g_j$ for $j = 2, \dots, d$. $\square$

We next investigate the properties of *mollifiers*. Besides duality, they are the basic tool in the study of Sobolev spaces. Set $\chi(x) = \exp \frac{-1}{1 - |x|_2^2}$ for $|x|_2 < 1$ and $\chi(x) = 0$ for $|x|_2 \geq 1$, where $x \in \mathbb{R}^d$. Observe that $\chi \in C^\infty(\mathbb{R}^d)$. We define

$$\chi_0(x) = \frac{1}{\|\chi\|_1} \chi(x) \quad \text{and} \quad \psi_\varepsilon(x) = \varepsilon^{-d} \chi_0(\tfrac{1}{\varepsilon} x), \tag{D.2}$$

for $x \in \mathbb{R}^d$ and $\varepsilon > 0$. We then have $0 \leq \psi_\varepsilon \in C^\infty(\mathbb{R}^d)$, $\psi_\varepsilon(x) > 0$ if and only if $|x|_2 < \varepsilon$, $\psi_\varepsilon = 0$ on $\mathbb{R}^d \setminus B(0, \varepsilon)$ and $\|\psi_\varepsilon\|_1 = 1$. For $f \in L^1_{\mathrm{loc}}(U)$ and $\varepsilon > 0$, we now introduce the mollifier G_ε by

$$G_\varepsilon f(x) = (\psi_\varepsilon * f)(x) = \int_{B(x,\varepsilon)} \psi_\varepsilon(x-y) f(y)\,\mathrm{d}y \qquad \text{(D.3)}$$
$$= \int_{B(0,\varepsilon)} \psi_\varepsilon(z) f(x-z)\,\mathrm{d}z,$$

for $x \in \mathbb{R}^d$. We will consider $G_\varepsilon f$ as a function on $\mathbb{R}^d$, U or $\overline{U}$ depending on our needs. From e.g. §4.4 of [**Bre11**], we recall that

$$G_\varepsilon f \in C^\infty(\mathbb{R}^d), \qquad \text{(D.4)}$$

$$\operatorname{supp} G_\varepsilon f \subseteq S_\varepsilon := S + \overline{B(0,\varepsilon)}, \quad \text{if } \operatorname{supp} f = S, \qquad \text{(D.5)}$$

$$\|G_\varepsilon f\|_{L^p(U)} \le \|G_\varepsilon f\|_{L^p(\mathbb{R}^d)} \le \|\psi_\varepsilon\|_1 \, \|f\|_p = \|f\|_p, \quad \text{if } f \in L^p(U),\ 1 \le p \le \infty, \qquad \text{(D.6)}$$

$$G_\varepsilon f \to f \text{ in } L^p(U) \text{ as } \varepsilon \to 0, \text{ if } f \in L^p(U) \text{ and } 1 \le p < \infty, \qquad \text{(D.7)}$$

$$G_\varepsilon f \to f \text{ as } \varepsilon \to 0 \text{ uniformly on compact subsets of } \mathbb{R}^d \text{ if } f \in C(\mathbb{R}^d). \qquad \text{(D.8)}$$

Observe that S_ε is compact if S is compact. In the following two core lemmas we improve the above properties of mollifiers and we apply them to the study of weak derivatives.

Lemma D.5. *Let $K \subseteq U$ be compact. Then there is a function $\varphi \in C_c^\infty(U)$ such that $0 \le \varphi \le 1$ on U and $\varphi = 1$ on K. Let $f \in L^1_{loc}(U)$ satisfy*

$$\int_U f\psi\,\mathrm{d}x = 0$$

for all $\psi \in C_c^\infty(U)$. Then $f = 0$ a.e.. In particular, weak derivatives are uniquely defined.

Proof. Assume that $f \neq 0$ on a Borel set $B \subseteq U$ with $\lambda(B) > 0$. Because of the regularity of the Lebesgue measure, there is a compact set $K \subseteq B \subseteq U$ with $\lambda(K) > 0$. Fix $0 < \delta < \frac{1}{2}\operatorname{dist}(\partial K, \partial U)$. Thus $K_{2\delta} \subseteq U$. The function $\varphi := G_\delta \mathbb{1}_{K_\delta}$ belongs to $C_c^\infty(U)$ by (D.4) and (D.5), where $\operatorname{supp}\varphi \subset K_{2\delta}$. Moreover, (D.6) and (D.3) yield that $0 \le \varphi(x) \le \|\varphi\|_\infty \le \|\mathbb{1}_{K_\delta}\|_\infty = 1$ for all $x \in U$ and that

$$\varphi(x) = \int_{B(x,\delta)} \psi_\delta(x-y) \mathbb{1}_{K_\delta}(y)\,\mathrm{d}y = \|\psi_\delta\|_1 = 1$$

for all $x \in K$. This construction shows the first claim.

Since $\varphi f \in L^1(U)$, the functions $G_\varepsilon(\varphi f)$ converge to φf in $L^1(U)$ as $\varepsilon \to 0$, due to (D.7). Hence, there is a nullset N and a subsequence $\varepsilon_j \to 0$ with $\varepsilon_j \le \delta$, such that $(G_{\varepsilon_j}(\varphi f))(x) \to f(x) \neq 0$ as $j \to \infty$ for each $x \in K \setminus N$. For every $x \in K \setminus N$ and $j \in \mathbb{N}$, we also deduce

$$(G_{\varepsilon_j}(\varphi f))(x) = \int_U \psi_{\varepsilon_j}(x-y)\varphi(y) f(y)\,\mathrm{d}y = 0$$

from the assumption, since the function $y \mapsto \psi_{\varepsilon_j}(x-y)\varphi(y)$ belongs to $C_c^\infty(U)$. This is a contradiction. □

Recall that Hölder's inequality implies that the map

$$L^p(B) \times L^{p'}(B) \to \mathbb{C}, \quad (f,g) \mapsto \int_B fg\,\mathrm{d}x, \tag{D.9}$$

is continuous for all $1 \le p \le \infty$ and Borel sets $B \subseteq \mathbb{R}^d$, where $\frac{1}{p} + \frac{1}{p'} = 1$.

Lemma D.6. *(Approximation lemma)*

(a) Let $f \in L^p_{loc}(U)$ possess the weak derivative $\partial^\alpha f \in L^p_{loc}(U)$ for some $\alpha \in \mathbb{N}_0^d$ and $p \in [1,\infty)$. Then the functions $G_\varepsilon f \in C^\infty(U)$ converge to f and $\partial^\alpha(G_\varepsilon f)$ converge to $\partial^\alpha f$ in $L^p_{loc}(U)$ as $\varepsilon \to 0$. We further have $\partial^\alpha(G_\varepsilon f)(x) = G_\varepsilon(\partial^\alpha f)(x)$ for all $x \in U$ and $\varepsilon < \operatorname{dist}(x, \partial U)$.

(b) Let f be given as in (a). Then there is a null sequence $(\varepsilon_n)_n$ such that $G_{\varepsilon_n} f \to f$ and $\partial^\alpha(G_{\varepsilon_n} f) \to \partial^\alpha f$ a.e. on U as $n \to \infty$. If $f \in W^k(U)$, we can take the same f_n for all $|\alpha| \le k$.

(c) If $f, g \in L^1_{loc}(U)$ and there are $f_n \in W^{|\alpha|}(U)$ such that $f_n \to f$ and $\partial^\alpha f_n \to g$ in $L^1_{loc}(U)$ as $n \to \infty$, then g is the weak derivative $\partial^\alpha f$. If this convergence holds in $L^p(U)$ for some $p \in [1,\infty]$ and all α with $|\alpha| \le k$, then $f \in W^k_p(U)$.

Proof. (a) Let $\varepsilon > 0$ and $x \in U$. If $\varepsilon < \operatorname{dist}(x, \partial U)$, then the function $y \mapsto \varphi_{\varepsilon,x}(y) := \psi_\varepsilon(x-y)$ belongs to $C_c^\infty(U)$. Using a corollary to Lebesgue's theorem and Definition D.1, we can thus deduce

$$\begin{aligned}(\partial^\alpha G_\varepsilon f)(x) &= \int_U \partial_x^\alpha \psi_\varepsilon(x-y) f(y)\,\mathrm{d}y = (-1)^{|\alpha|} \int_U (\partial^\alpha \varphi_{\varepsilon,x})(y) f(y)\,\mathrm{d}y \\ &= \int_U \varphi_{\varepsilon,x}(y)(\partial^\alpha f)(y)\,\mathrm{d}y = (G_\varepsilon \partial^\alpha f)(x).\end{aligned}$$

Choose a compact subset $K \subseteq U$ and fix $\delta > 0$ with $K_\delta \subseteq U$. Take $\varepsilon \in (0,\delta]$. Note that the integrand of $(G_\varepsilon g)(x)$ is then supported in K_δ for every $x \in K$ and $g \in L^1_{\mathrm{loc}}(U)$, see (D.3). Due to (D.7), the functions

$$\mathbb{1}_K \partial^\alpha(G_\varepsilon f) = \mathbb{1}_K G_\varepsilon(\partial^\alpha f) = \mathbb{1}_K G_\varepsilon(\mathbb{1}_{K_\delta} \partial^\alpha f)$$

converge in $L^p(K)$ to $\mathbb{1}_K \mathbb{1}_{K_\delta} \partial^\alpha f = \mathbb{1}_K \partial^\alpha f$ as $\varepsilon \to 0$. So we have shown the asserted convergence in $L^p_{\mathrm{loc}}(U)$.

(b) To derive (b) from (a) we note that the sets

$$K_m = \left\{ x \in U \,\middle|\, \operatorname{dist}(x, \partial U) \ge \tfrac{1}{m} \text{ and } |x|_2 \le m \right\}$$

are compact and $\bigcup_{m \in \mathbb{N}} K_m = U$. Let $\varepsilon_k \to 0$. Then, for each $m \in \mathbb{N}$ there is a null set $N_m \subseteq K_m$ and a subsequence $\nu_m(j)$ such that $\partial^\alpha G_{\varepsilon_{\nu_m(j)}} f(x)$ converges to $\partial^\alpha f(x)$ and $G_{\varepsilon_{\nu_m(j)}} f(x)$ converges to $f(x)$ for all $x \in K_m \setminus N_m$ as $j \to \infty$. By means of a diagonal sequence, one obtains a null sequence $(\varepsilon_n)_n$ such that $\partial^\alpha G_{\varepsilon_n} f(x) \to \partial^\alpha f(x)$ and $G_{\varepsilon_n} f(x) \to f(x)$ for $x \in U \setminus (\bigcup_{m \in \mathbb{N}} N_m)$, as $n \to \infty$, where $\bigcup_{m \in \mathbb{N}} N_m$ is a null set. This procedure can also be done for finitely many $\partial^\alpha f$ at the same time.

(c) Let $f_n \in W^{|\alpha|}(U)$ be given such that $f_n \to f$ and $\partial^\alpha f_n \to g$ in $L^1_{\mathrm{loc}}(U)$ as $n \to \infty$. From (D.9) on $\operatorname{supp} \varphi$ we deduce that

$$\int_U f \partial^\alpha \varphi\,\mathrm{d}x = \lim_{n\to\infty} \int_U f_n \partial^\alpha \varphi\,\mathrm{d}x = (-1)^\alpha \lim_{n\to\infty} \int_U (\partial^\alpha f_n) \varphi\,\mathrm{d}x = (-1)^\alpha \int_U g\varphi\,\mathrm{d}x$$

for all $\varphi \in C_c^\infty(U)$. Hence, $g = \partial^\alpha f$. In the setting of the last assertion we thus obtain $\partial^\alpha f \in L^p(U)$ for all $|\alpha| \le k$, and hence $f \in W_p^k(U)$. □

Proposition D.7. *(Product rules)*
(a) Let $f, g \in W^1(U) \cap L^\infty(U)$. Then, $fg \in W^1(U) \cap L^\infty(U)$ and

$$\partial_j(fg) = (\partial_j f)g + f(\partial_j g) \tag{D.10}$$

for all $j \in \{1, \dots, d\}$.
(b) Let $1 \le p \le \infty$, $f \in W_p^1(U)$ and $g \in W_{p'}^1(U)$. Then, $fg \in W_1^1(U)$ and (D.10) holds.

Proof. 1) Let $f, g \in W^1(U)$. Set $f_n = G_{\varepsilon_n} f \in C^\infty(U)$ and $g_n = G_{\varepsilon_n} g \in C^\infty(U)$ with $\varepsilon_n \to 0$ as in Lemma D.6 (b). Fix $m \in \mathbb{N}$ and take $\varphi \in C_c^\infty(U)$ and $j \in \{1, \dots, d\}$. Choose an open and bounded set V such that $\operatorname{supp} \varphi \subseteq V \subseteq \overline{V} \subseteq U$. Since $f_n \to f$ and $\partial_j f_n \to \partial_j f$ on $L^1(\overline{V})$ by Lemma D.6 (b), the formulas (D.9) and (D.1) yield

$$\begin{aligned}\int_U f g_m \partial_j \varphi \, \mathrm{d}x &= \lim_{n\to\infty} \int_{\overline{V}} f_n g_m \partial_j \varphi \, \mathrm{d}x = -\lim_{n\to\infty} \int_{\overline{V}} ((\partial_j f_n) g_m + f_n(\partial_j g_m))\varphi \, \mathrm{d}x \\ &= -\int_U ((\partial_j f) g_m + f(\partial_j g_m))\varphi \, \mathrm{d}x\end{aligned}$$

so that the weak derivative $\partial_j(fg_m) = (\partial_j f)g_m + f(\partial_j g_m) \in L^1_{\mathrm{loc}}(U)$ exists.

2) Let $f, g \in W^1(U) \cap L^\infty(U)$ and g_m as in 1). Note that $g_m \to g$ and $\partial_j g_m \to \partial_j g$ in $L^1_{\mathrm{loc}}(U)$ as $m \to \infty$. Since f is bounded, we obtain

$$\int_U fg \partial_j \varphi \, \mathrm{d}x = \lim_{m\to\infty} \int_U f g_m \partial_j \varphi \, \mathrm{d}x = \lim_{m\to\infty} -\left[\int_U (\partial_j f) g_m \varphi \, \mathrm{d}x + \int_U f(\partial_j g_m)\varphi \, \mathrm{d}x\right],$$

using step 1). On the right-hand side, the second integral converges to $\int_U f(\partial_j g)\varphi \, \mathrm{d}x$, again because of $f \in L^\infty(U)$. For the first integral we use that $g_m \to g$ a.e. by Lemma D.6 (b) and that $\|g_m\|_\infty \le \|g\|_\infty$ by (D.6). The theorem of dominated convergence (with the majorant $|\partial_j f| \|g\|_\infty \|\varphi\|_\infty \mathbb{1}_{\operatorname{supp}\varphi}$) now yields

$$\int_U fg \partial_j \varphi \, \mathrm{d}x = -\int_U ((\partial_j f)g + f(\partial_j g))\varphi \, \mathrm{d}x.$$

Hence, (a) holds since also $(\partial_j f)g + f(\partial_j g) \in L^1_{\mathrm{loc}}(U)$ by our assumptions.

3) Let $f \in W_p^1(U)$ and $g \in W_{p'}^1(U)$. If $p \in (1, \infty]$, we show (D.10) as in step 2), using (D.9) and that g_m, $\partial_j g_m$ converge in $L^{p'}(U)$ by (D.7) (a). If $p = 1$, we replace the roles of f and g to derive (D.10). Hölder's inequality and (D.10) finally yield $\partial_j(fg) \in L^1(U)$. As a result, (b) is true. □

Proposition D.8. *(Chain rules)*
(a) Let $f \in W^1(U)$ be real valued and $h \in C^1(\mathbb{R})$ with $h' \in C_b(\mathbb{R})$. We then have $h \circ f \in W^1(U)$ and

$$\partial_j(h \circ f) = (h' \circ f)\partial_j f$$

for all $j \in \{1, \dots, d\}$.

(b) Let $f \in W^1(U)$, $V \subseteq \mathbb{R}^d$ be open and $\Phi : V \to U$ be a diffeomorphism such that Φ' and $(\Phi^{-1})'$ are bounded. We then have $f \circ \Phi \in W^1(V)$ and

$$\partial_j(f \circ \Phi) = \sum_{m=1}^{d} ((\partial_m f) \circ \Phi)\partial_j \Phi_m$$

for all $j = 1, \dots, d$.

In both results we can replace $W^1(U)$ by $W^1_p(U)$ for $1 \le p \le \infty$, if in (a) also $h(0) = 0$ holds in the case that $\lambda(U) = \infty$ and $p < \infty$.

Proof. (a) By Lemma D.6, there are $f_n \in C^\infty(U)$ such that $f_n \to f$ and $\partial_j f_n \to \partial_j f$ in $L^1_{\mathrm{loc}}(U)$ and a.e. as $n \to \infty$, for every $j \in \{1, \dots, d\}$. Since

$$|h(f(x))| \le |h(f(x)) - h(0)| + |h(0)| \le \|h'\|_\infty |f(x)| + |h(0)|$$

for all $x \in U$, the function $h \circ f$ belongs to $L^1_{\mathrm{loc}}(U)$ (and to $L^p(U)$ if $f \in L^p(U)$ and if $h(0) = 0$ in the case that $\lambda(U) = \infty$ and $p \neq \infty$). Let $K \subseteq U$ be compact. We obtain that

$$\begin{aligned}
&\int_K |h(f_n(x)) - h(f(x))| \,\mathrm{d}x \le \|h'\|_\infty \int_K |f_n(x) - f(x)| \,\mathrm{d}x \;\longrightarrow\; 0,\\
&\int_K |h'(f_n(x))\partial_j f_n(x) - h'(f(x))\partial_j f(x)| \,\mathrm{d}x\\
&\quad\le \|h'\|_\infty \int_K |\partial_j f_n(x) - \partial_j f(x)| \,\mathrm{d}x + \int_K |h'(f_n(x)) - h'(f(x))||\partial_j f(x)| \,\mathrm{d}x \to 0
\end{aligned}$$

as $n \to \infty$ where we also used Lebesgue's theorem and the majorant $2\|h'\|_\infty |\partial_j f|$ in the last integral. Since $h \circ f_n \in C^1(U)$, $(h' \circ f)\partial_j f \in L^1_{\mathrm{loc}}(U)$ and $\partial_j(h \circ f_n) = (h' \circ f_n)\partial_j f_n$, Lemma D.6 (c) yields assertion (a). If $f \in W^1_p(U)$ then $(h' \circ f)\partial_j f \in L^p(U)$ and so $h \circ f \in W^1_p(U)$.

Assertion (b) can similarly be shown using the transformation rule. □

Corollary D.9. *Let $f \in W^1(U)$ be real valued. Then $f_+, f_-, |f|$ belong to $W^1(U)$ with derivatives*

$$\partial_j f_\pm = \pm \mathbb{1}_{\{f \gtrless 0\}} \partial_j f \quad \textit{and} \quad \partial_j |f| = (\mathbb{1}_{\{f>0\}} - \mathbb{1}_{\{f<0\}})\partial_j f$$

for all $j \in \{1, \dots, d\}$. Here one can replace W^1 by W^1_p for all $1 \le p \le \infty$.

Proof. We use the function $h_\varepsilon \in C^1(\mathbb{R})$ given by $h_\varepsilon(t) := \sqrt{t^2 + \varepsilon^2} - \varepsilon$ for $t \ge 0$ and $h_\varepsilon(t) := 0$ for $t < 0$, where $\varepsilon > 0$. Observe that $\|h'_\varepsilon\|_\infty = 1$, $h_\varepsilon(t) \le t$ for $t \ge 0$ as well as $h_\varepsilon(t) \to t$ for $t > 0$ and $h_\varepsilon(t) \to 0$ for $t \le 0$, as $\varepsilon \to 0$. Proposition D.8 shows that $h_\varepsilon \circ f \in W^1(U)$ and

$$\int_U h_\varepsilon(f)\partial_j \varphi \,\mathrm{d}x = -\int_U h'_\varepsilon(f)(\partial_j f)\varphi \,\mathrm{d}x = -\int_{\{f>0\}} \frac{f}{\sqrt{f^2 + \varepsilon^2}}(\partial_j f)\varphi \,\mathrm{d}x$$

for each $\varphi \in C^\infty_c(U)$. Thanks to the majorants $\|\partial_j \varphi\|_\infty \mathbb{1}_B |f|$ and $\|\varphi\|_\infty |\partial_j f| \mathbb{1}_B$ with $B = \operatorname{supp} \varphi$, Lebesgue's convergence theorem shows that

$$\int_U f_+ \partial_j \varphi \,\mathrm{d}x = -\int_{\{f>0\}} \frac{f}{|f|}(\partial_j f)\varphi \,\mathrm{d}x = -\int_U \mathbb{1}_{\{f>0\}}(\partial_j f)\varphi \,\mathrm{d}x.$$

There thus exists $\partial_j f_+ = \mathbb{1}_{\{f>0\}} \partial_j f \in L^1_{\mathrm{loc}}(U)$. The other assertions follow from $f_- = (-f)_+$ and $|f| = f_+ + f_-$. □

We describe Sobolev spaces for the simpler cases $d = 1$ and $p = \infty$.

THEOREM D.10. *Let $J \subseteq \mathbb{R}$ be an open interval and $f \in L^1_{loc}(J)$. We then have $f \in W^1(J)$ if and only if f has a continuous representative (also denoted by f) and there is a function $g \in L^1_{loc}(J)$ such that*

$$f(t) = f(s) + \int_s^t g(\tau)\,\mathrm{d}\tau \tag{D.11}$$

holds for all $s, t \in J$. In this case, $g = f'$ a.e.. (Compare the Exercises of Lecture 3.)

PROOF. 1) Let $f \in W^1(J)$. Take $f_n = G_{\varepsilon_n} f \in C^\infty(J)$ from Lemma D.6 (b). Then for a.e. $t \in J$ and for a.e. $t_0 \in J$ we have

$$f(t) - f(t_0) = \lim_{n\to\infty} (f_n(t) - f_n(t_0)) = \lim_{n\to\infty} \int_{t_0}^t f_n'(\tau)\,\mathrm{d}\tau = \int_{t_0}^t f'(\tau)\,\mathrm{d}\tau.$$

Fixing one such t_0 and noting that $t \mapsto \int_{t_0}^t f'(\tau)\,\mathrm{d}\tau$ is continuous, we obtain a continuous representative of f which satisfies (D.11) for $s = t_0$ and $g = f'$. Subtracting two such equations for any given $t, s \in J$ and the fixed t_0, we deduce (D.11) with $g = f'$ for all $t, s \in J$.

2) If (D.11) holds for some $f \in C(J)$ and $g \in L^1_{\mathrm{loc}}(J)$, take $g_n \in C^\infty(J)$ such that $g_n \to g$ in $L^1_{\mathrm{loc}}(J)$ as $n \to \infty$. For any $s \in J$ and $n \in \mathbb{N}$, the function $f_n(t) := f(s) + \int_s^t g_n(\tau)\,\mathrm{d}\tau$, $t \in J$, belongs to $C^\infty(J)$ with $f_n' = g_n$. Moreover, for $[a, b] \subseteq J$ with $s \in [a, b]$, we estimate

$$\|f_n - f\|_{L^1([a,b])} \le \int_a^b \int_s^t |g_n(\tau) - g(\tau)|\,\mathrm{d}\tau\,\mathrm{d}t \le (b - a)\|g_n - g\|_{L^1([a,b])},$$

using (D.11), so that $f_n \to f$ in $L^1_{\mathrm{loc}}(J)$ as $n \to \infty$. Lemma D.6 (b) then yields $f \in W^1(J)$ and $f' = g$. □

PROPOSITION D.11. *Let $U \subseteq \mathbb{R}^d$ be open and convex. Then $W^1_\infty(U)$ is isomorphic to*

$$C_b^{1-}(U) := \{ f \in C_b(U) \mid f \text{ is Lipschitz} \},$$

and the norm of $W^1_\infty(U)$ is equivalent to

$$\|f\|_{C_b^{1-}} = \|f\|_\infty + [f]_{Lip},$$

where $[f]_{Lip}$ is the Lipschitz constant of f.

PROOF. Let $f \in W^1_\infty(U)$. Take the null sequence $(\varepsilon_n)_n$ from Lemma D.6 (b). Let $K \subseteq U$ be compact. For sufficiently large $n \in \mathbb{N}$, Lemma D.6 and (D.6) yield

$$|\partial_j G_{\varepsilon_n} f(z)| = |G_{\varepsilon_n} \partial_j f(z)| \le \|\partial_j f\|_\infty \le \|f\|_{1,\infty},$$

for all $j \in \{1, \dots, d\}$ and $z \in K$. Using that $G_{\varepsilon_n} f(x) \to f(x)$ as $n \to \infty$ for all $x \in U \setminus N$ and a null set N, we then estimate

$$\begin{aligned} |f(x) - f(y)| &= \lim_{n\to\infty} |G_{\varepsilon_n} f(x) - G_{\varepsilon_n} f(y)| \qquad \text{(D.12)} \\ &= \lim_{n\to\infty} \Big| \int_0^1 \nabla G_{\varepsilon_n} f(y + \tau(x - y)) \cdot (x - y)\,\mathrm{d}\tau \Big| \le \sqrt{d}\|f\|_{1,\infty}\,|x - y|_2 \end{aligned}$$

for all $x, y \in U \setminus N$. By continuous extension we obtain a Lipschitz continuous representative of f with Lipschitz constant $\sqrt{d}\|f\|_{1,\infty}$.

Let $f \in C_b^{1-}(U)$. Take $\varphi \in C_c^\infty(U)$, $j \in \{1, \dots, d\}$, and $\delta > 0$ such that $(\operatorname{supp}\varphi)_\delta \subseteq U$. For $\varepsilon \in (0, \delta]$ the difference quotient $\frac{1}{\varepsilon}(\varphi(x + \varepsilon e_j) - \varphi(x))$ converges uniformly on $\operatorname{supp}\varphi$ as $\varepsilon \to 0$, and hence

$$\begin{aligned}
\Big|\int_U f \partial_j \varphi \,\mathrm{d}x\Big| &= \lim_{\varepsilon\to 0} \Big|\int_{\operatorname{supp}\varphi} f(x)\tfrac{1}{\varepsilon}(\varphi(x+\varepsilon e_j) - \varphi(x))\,\mathrm{d}x\Big| \\
&\le \limsup_{\varepsilon\to 0} \int_{\mathbb{R}^d} \tfrac{1}{\varepsilon}|f(y-\varepsilon e_j) - f(y)|\,|\varphi(y)|\,\mathrm{d}y \\
&\le [f]_{\mathrm{Lip}}\|\varphi\|_1.
\end{aligned}$$

Taking into account that $C_c^\infty(U)$ is dense in $L^1(U)$, we see that the map $\varphi \mapsto -\int_U f \partial_j \varphi \,\mathrm{d}x$ has a continuous linear extension $F_j : L^1(U) \to \mathbb{C}$. Therefore there is a function $g_j \in L^\infty(U) = L^1(U)^*$ with $\|g_j\|_\infty = \|F_j\| \le [f]_{\mathrm{Lip}}$ such that

$$-\int_U f \partial_j \varphi \,\mathrm{d}x = F_j(\varphi) = \int_U g_j \varphi \,\mathrm{d}x$$

for all $\varphi \in C_c^\infty(U)$. This means that f has the weak derivative $\partial_j f = g_j \in L^\infty(U)$. As a result, $f \in W^1_\infty(U)$ and $\|f\|_{W^1_\infty(U)} \le \|f\|_\infty + [f]_{\mathrm{Lip}}$. □

Without convexity of U, the above proof still gives $C_b^{1-}(U) \hookrightarrow W^1_\infty(U)$ and that $f \in W^1_\infty(U)$ has a locally Lipschitz representative. If U is bounded and $\partial U \in C^1$, then Proposition D.11 still holds, see Theorem 5.8.4 in [**Eva10**]. The proof given there is based on an extension argument, cf. Theorem D.23.

2. Density and embedding theorems

In this section we prove some of the most important theorems on Sobolev spaces. We now restrict ourselves to the case $p < \infty$ since one partly has different results for $p = \infty$ and since this case is essentially settled by Proposition D.11.

Definition D.12. *For $k \in \mathbb{N}$ and $1 \le p < \infty$, the closure of $C_c^\infty(U)$ in $W_p^k(U)$ is denoted by $\mathring{W}_p^k(U)$. We set $\mathring{H}^k(U) = \mathring{W}_2^k(U)$.*

Theorem D.13. *Let $k \in \mathbb{N}$ and $p \in [1, \infty)$. We then have*

$$\mathring{W}_p^k(\mathbb{R}^d) = W_p^k(\mathbb{R}^d).$$

Moreover, the set $C^\infty(U) \cap W_p^k(U)$ is dense in $W_p^k(U)$.

Proof. We prove the theorem only for $k = 1$, the general case can be treated similarly.

1) Let $f \in W_p^1(\mathbb{R}^d)$. Take any $\phi \in C^\infty(\mathbb{R}_+)$ such that $0 \le \phi \le 1$, $\phi = 1$ on $[0,1]$ and $\phi = 0$ on $[2, \infty)$. Set

$$\varphi_n(x) = \phi(\tfrac{1}{n}|x|_2) \qquad (\text{“cut-off function”})$$

for $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$. We then have $\varphi_n \in C_c^\infty(\mathbb{R}^d)$, $0 \le \varphi_n \le 1$ and $\|\partial_j \varphi_n\|_\infty \le \|\phi'\|_\infty \frac{1}{n}$ for all $n \in \mathbb{N}$, as well as $\varphi_n(x) \to 1$ for all $x \in \mathbb{R}^d$ as

$n \to \infty$. Hence, $\|\varphi_n f - f\|_p \to 0$ as $n \to \infty$ by Lebesgue's convergence theorem. Moreover, Proposition D.7 implies that

$$\begin{aligned}\|\partial_j(\varphi_n f - f)\|_p &= \|(\varphi_n \partial_j f - \partial_j f) + (\partial_j \varphi_n) f\|_p \\ &\le \|\varphi_n \partial_j f - \partial_j f\|_p + \tfrac{1}{n}\|\phi'\|_\infty \|f\|_p\,,\end{aligned}$$

and the right-hand side tends to 0 as $n \to \infty$ for each $j \in \{1, \dots, d\}$. Given $\varepsilon > 0$, we can thus fix $m \in \mathbb{N}$ such that $\|\varphi_m f - f\|_{1,p} \le \varepsilon$. Due to (D.4) and (D.5), the functions $G_{\frac{1}{n}}(\varphi_m f)$ belong to $C_c^\infty(\mathbb{R}^d)$ for all $n \in \mathbb{N}$. Equation (D.7) and Lemma D.6 further yield that

$$G_{\frac{1}{n}}(\varphi_m f) \to \varphi_m f \quad \text{and} \quad \partial_j G_{\frac{1}{n}}(\varphi_m f) = G_{\frac{1}{n}} \partial_j(\varphi_m f) \to \partial_j(\varphi_m f)$$

in $L^p(\mathbb{R}^d)$ as $n \to \infty$, for each $j \in \{1, \dots, d\}$. So there is an $n \in \mathbb{N}$ such that

$$\|G_{\frac{1}{n}}(\varphi_m f) - \varphi_m f\|_{1,p} \le \varepsilon,$$

and thus

$$\|G_{\frac{1}{n}}(\varphi_m f) - f\|_{1,p} \le 2\varepsilon.$$

2) For the second assertion, we can assume that $\partial U \neq \emptyset$. Let $f \in W_p^1(U)$. Define the open sets

$$U_n = \Big\{x \in U \,\Big|\, |x|_2 < n \text{ and } \operatorname{dist}(x, \partial U) > \tfrac{1}{n}\Big\}$$

for each $n \in \mathbb{N}$. Then $U_n \subseteq \overline{U}_n \subseteq U_{n+1} \subset U$, $\overline{U}_n$ is compact and $\bigcup_{n=1}^\infty U_n = U$. Observe that $U = \bigcup_{n=1}^\infty U_{n+1} \setminus \overline{U}_{n-1}$, where $U_0, U_{-1} := \emptyset$. As shown in analysis courses, there are functions $\varphi_n \in C_c^\infty(U)$ such that $\operatorname{supp} \varphi_n \subseteq U_{n+1} \setminus \overline{U}_{n-1}$, $\varphi_n \ge 0$, and $\sum_{n=1}^\infty \varphi_n(x) = 1$ for all $x \in U$.

Fix $\varepsilon > 0$. As in Step 1), for each $n \in \mathbb{N}$ there is a $\delta_n > 0$ such that $g_n := G_{\delta_n}(\varphi_n f) \in C_c^\infty(U)$, $\operatorname{supp} g_n \subseteq (\operatorname{supp} \varphi_n f)_{\delta_n} \subseteq U_{n+1} \setminus \overline{U}_{n-1}$ and $\|g_n - \varphi_n f\|_{1,p} \le 2^{-n}\varepsilon$. Define $g(x) = \sum_{n=1}^\infty g_n(x)$ for all $x \in U$. Observe that on each ball $\overline{B(x,r)} \subseteq U$ this sum is finite. Hence, $g \in C^\infty(U)$. Since $f = \sum_{n=1}^\infty \varphi_n f$, we further have

$$g(x) - f(x) = \sum_{n=1}^\infty (g_n(x) - \varphi_n(x) f(x)),$$

for all $x \in U$ and $n \in \mathbb{N}$. Due to $\|g_n - \varphi_n f\|_{1,p} \le 2^{-n}\varepsilon$, this series converges absolutely in $W_p^1(U)$, and

$$\|f - g\|_{1,p} \le \sum_{n=1}^\infty \|g_n - \varphi_n f\|_{1,p} \le \varepsilon. \qquad \square$$

Remark D.14. (a) If U is bounded, then $\mathring{W}_p^k(U) \neq W_p^k(U)$, see Lemma 6.67 in [**RR04**]. ♢

(b) For "good" ∂U one can replace in $C^\infty(U)$ by $C^\infty(\overline{U})$ in Theorem D.13, see Corollary D.21 below. ♢

We now want to study embeddings of Sobolev spaces. We clearly have

$$W_p^k(U) \hookrightarrow W_p^j(U) \quad \text{if } k \ge j \ge 0$$

and

$$W_p^k(U) \hookrightarrow W_q^j(U) \quad \text{if } k \geq j \geq 0,\ 1 \leq q \leq p \leq \infty \text{ and } \lambda(U) < \infty.$$

(Here we put $W_p^0(U) = L^p(U)$ for $1 \leq p \leq \infty$.) The embedding $X \hookrightarrow Y$ means that there is an injective map $J \in \mathcal{B}(X,Y)$. Above it holds $Jf = f$, and below we also use $Jf = f + \mathcal{N}$. Writing $c = \|J\|$, one obtains $\|f\|_Y \leq c\|f\|_X$ if $X \hookrightarrow Y$ and one identifies Jf with f.

Theorem D.15 (Sobolev, Morrey). *Let $k \in \mathbb{N}$ and $1 \leq p < \infty$. We have the following embeddings.*

(a) If $kp < d$, then

$$p^* := \frac{pd}{d - kp} \in (p, \infty) \quad \text{and} \quad W_p^k(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$$

for all $q \in [p, p^*]$.

(b) If $kp = d$, then

$$W_p^k(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d)$$

for all $q \in [p, \infty)$.

(c) If $kp > d$, then there are either $j \in \mathbb{N}_0$ and $\beta \in (0,1)$ such that $k - \frac{d}{p} = j + \beta$ or $k - \frac{d}{p} \in \mathbb{N}$. In the latter case we set $j := k - \frac{d}{p} - 1 \in \mathbb{N}_0$ and take any $\beta \in (0,1)$. Then

$$W_p^k(\mathbb{R}^d) \hookrightarrow C_0^j(\mathbb{R}^d) \quad \text{and} \quad |\partial^\alpha f(x) - \partial^\alpha f(y)| \leq c|x - y|_2^\beta$$

for all $x, y \in \mathbb{R}^d$, $|\alpha| \leq j$, $f \in W_p^k(\mathbb{R}^d)$ and a constant $c > 0$, where we take a representative $f \in C_0^j(\mathbb{R}^d)$ of f and

$$C_0^j(\mathbb{R}^d) = \left\{ u \in C^j(\mathbb{R}^d) \,\middle|\, \partial^\alpha u(x) \to 0 \text{ as } |x|_2 \to \infty \text{ for all } 0 \leq |\alpha| \leq j \right\}.$$

Corollary D.16. *Let $k \in \mathbb{N}$ and $p \in [1, \infty)$. If there are $j \in \mathbb{N}_0$ and $q \in [p, \infty)$ with $k - \frac{d}{p} = j - \frac{d}{q}$, then*

$$W_p^k(\mathbb{R}^d) \hookrightarrow W_q^j(\mathbb{R}^d).$$

Proof of Corollary D.16. By assumption, we have $(k-j)p = d - \frac{dp}{q} \in [0, d)$. Theorem D.15 (a) thus yields

$$W_p^{k-j}(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d) \quad \text{since } q = \frac{pd}{d - kp + jp}.$$

Applying this embedding to $\partial^\alpha f \in W_p^{k-j}(\mathbb{R}^d)$ for all $|\alpha| \leq j$ and $f \in W_p^k(\mathbb{R}^d)$, we deduce that

$$\|\partial^\alpha f\|_q \leq c\|\partial^\alpha f\|_{k-j,p} \leq c\|f\|_{k,p},$$

as asserted. □

Example D.17. There is an unbounded function $f \in W_d^1(\mathbb{R}^d)$ for $d \geq 2$, showing that Theorem D.15 (b) is sharp. In fact, for any $\alpha \in (0, 1 - \frac{1}{d})$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi \subseteq B(0, \frac{3}{4})$ and $\varphi = 1$ on $B(0, \frac{1}{2})$, we define

$$f(x) := \begin{cases} \varphi(x)(-\log|x|_2)^\alpha, & 0 < |x|_2 \leq \frac{3}{4}, \\ 0, & |x|_2 > \frac{3}{4} \text{ or } x = 0. \end{cases}$$

Arguing as in Example D.4 (c), one sees that $f \in L^p(\mathbb{R}^d)$ for all $p \in [1,\infty)$, $f \notin L^\infty(\mathbb{R}^d)$ and that, for all $j \in \{1,\dots,d\}$, we have

$$\partial_j f(x) = (\partial_j \varphi(x))(-\log |x|_2)^\alpha - \alpha\varphi(x)(-\log |x|_2)^{\alpha-1} \frac{x_j}{|x|_2^2}$$

for $0 < |x|_2 < \frac{3}{4}$ and $\partial_j f(x) = 0$ otherwise. Using polar coordinates, we further estimate

$$\begin{aligned}\Big(\int_{\mathbb{R}^d} |\partial_j f|^d \,\mathrm{d}x\Big)^{1/d} &\le c\|\partial_j\varphi\|_\infty \Big(\int_0^{\frac{3}{4}} |\log r|^{\alpha d} r^{d-1}\,\mathrm{d}r\Big)^{1/d} \\ &\quad + c\|\varphi\|_\infty \Big(\int_0^{\frac{3}{4}} \frac{|\log r|^{(\alpha-1)d}}{r^d} r^{d-1}\,\mathrm{d}r\Big)^{1/d} \\ &\le c + c\Big(\int_0^{\frac{3}{4}} \frac{1}{r|\log r|^{(1-\alpha)d}}\,\mathrm{d}r\Big)^{1/d} < \infty\end{aligned}$$

for some constants $c > 0$, since $(1-\alpha)d > 1$. Hence, $f \in W_d^1(\mathbb{R}^d) \setminus L^\infty(\mathbb{R}^d)$. $\Diamond$

For the proof of Theorem D.15 we set $\hat{x}^j = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_d) \in \mathbb{R}^{d-1}$ for all $x \in \mathbb{R}^d$, $j \in \{1,\dots,d\}$ and $d \ge 2$. We start with a lemma.

LEMMA D.18. *Let $d \ge 2$ and $f_1, \dots, f_d \in L^{d-1}(\mathbb{R}^{d-1}) \cap C(\mathbb{R}^{d-1})$. Set $f(x) = f_1(\hat{x}^1) \cdot \dots \cdot f_d(\hat{x}^d)$ for $x \in \mathbb{R}^d$. We then have $f \in L^1(\mathbb{R}^d)$ and*

$$\|f\|_{L^1(\mathbb{R}^d)} \le \|f_1\|_{L^{d-1}(\mathbb{R}^{d-1})} \cdot \dots \cdot \|f_d\|_{L^{d-1}(\mathbb{R}^{d-1})}.$$

PROOF. If $d = 2$, then Fubini's theorem shows that

$$\int_{\mathbb{R}^2} |f(x)|\,\mathrm{d}x = \int_{\mathbb{R}}\int_{\mathbb{R}} |f_1(x_2)||f_2(x_1)|\,\mathrm{d}x_1\,\mathrm{d}x_2 = \|f_1\|_1\|f_2\|_1,$$

as asserted.

Assume that the assertion holds for some $d \in \mathbb{N}$ with $d \ge 2$. Take $f_1, \dots, f_{d+1} \in L^d(\mathbb{R}^d) \cap C(\mathbb{R}^d)$. Write $y = (x_1, \dots, x_d) \in \mathbb{R}^d$ and $x = (y, x_{d+1}) \in \mathbb{R}^{d+1}$. For a.e. $x_{d+1} \in \mathbb{R}$, the maps $\hat{y}^j \mapsto |f_j(\hat{y}^j, x_{d+1})|^d$ are integrable on $\mathbb{R}^{d-1}$ for every $j \in \{1,\dots,d\}$ due to Fubini's theorem. Fix such an $x_{d+1} \in \mathbb{R}$ and write

$$\tilde{f}(y, x_{d+1}) := \prod_{j=1}^d f_j(\hat{x}^j).$$

Using Hölder's inequality and $d' = \frac{d}{d-1}$, we obtain

$$\begin{aligned}\int_{\mathbb{R}^d} |f(y, x_{d+1})|\,\mathrm{d}y &= \int_{\mathbb{R}^d} |\tilde{f}(y, x_{d+1})||f_{d+1}(y)|\,\mathrm{d}y \\ &\le \|f_{d+1}\|_{L^d(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} |\tilde{f}(y, x_{d+1})|^{d'}\,\mathrm{d}y\Big)^{1/d'}.\end{aligned}$$

We set $g_j(\hat{y}^j) = |f_j(\hat{y}^j, x_{d+1})|^{d'}$ for $j \in \{1,\dots,d\}$ and $x \in \mathbb{R}^{d+1}$. Since $d'(d-1) = d$, we have $g_j \in L^{d-1}(\mathbb{R}^{d-1})$, and the induction hypothesis yields

$$\begin{aligned}\int_{\mathbb{R}^d} |\tilde{f}(y, x_{d+1})|^{d'}\,\mathrm{d}y &= \int_{\mathbb{R}^d} g_1(\hat{y}^1) \cdot \dots \cdot g_d(\hat{y}^d)\,\mathrm{d}y \le \|g_1\|_{d-1} \cdot \dots \cdot \|g_d\|_{d-1} \\ &= \prod_{j=1}^d \Big(\int_{\mathbb{R}^{d-1}} |f_j(\hat{y}^j, x_{d+1})|^d\,\mathrm{d}y\Big)^{\frac{1}{d-1}}.\end{aligned}$$

Integrating over $x_{d+1} \in \mathbb{R}$, we thus arrive at

$$\int_{\mathbb{R}^{d+1}} |f| \,\mathrm{d}x \le \|f_{d+1}\|_d \int_{\mathbb{R}} \prod_{j=1}^{d} \Big(\int_{\mathbb{R}^{d-1}} |f_j(\hat{x}^j)|^d \,\mathrm{d}y\Big)^{\frac{1}{d-1}\frac{d-1}{d}} \,\mathrm{d}x_{d+1}.$$

Applying the d-fold Hölder inequality to the x_{d+1}-integral, we conclude that

$$\begin{aligned}\int_{\mathbb{R}^{d+1}} |f| \,\mathrm{d}x &\le \|f_{d+1}\|_d \prod_{j=1}^{d} \Big(\int_{\mathbb{R}} \Big(\int_{\mathbb{R}^{d-1}} |f_j(\hat{x}^j)|^d \,\mathrm{d}y\Big)^{\frac{1}{d}d} \,\mathrm{d}x_{d+1}\Big)^{\frac{1}{d}} \\ &= \|f_1\|_d \cdot \ldots \cdot \|f_{d+1}\|_d.\end{aligned}$$

□

Recall from integration theory that for $f \in L^p(U) \cap L^q(U)$ and $r \in [p,q]$ with $1 \le p < q \le \infty$, we have

$$\|f\|_r \le \|f\|_p^\theta \|f\|_q^{1-\theta} \le \theta\|f\|_p + (1-\theta)\|f\|_q, \tag{D.13}$$

where $\theta \in [0,1]$ is given by $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}$ and we also used Young's inequality from undergraduate courses.

Proof of Theorem D.15. We only prove the case $k = 1$, the rest can be done by induction, see e.g. §5.6.3 in [**Eva10**]. Since $W_p^1(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$, estimate (D.13) implies that for assertion (a) it suffices to show

$$W_p^1(\mathbb{R}^d) \hookrightarrow L^{p^*}(\mathbb{R}^d).$$

1) Let $f \in C_c^1(\mathbb{R}^d)$. Let first $p = 1 < d$, whence $p^* = \frac{d}{d-1}$. For $x \in \mathbb{R}^d$ and $j \in \{1,\dots,d\}$, we then obtain

$$\begin{aligned}|f(x)| &= \Big|\int_{-\infty}^{x_j} \partial_j f(x_1,\dots,x_{j-1},t,x_{j+1},\dots,x_d)\,\mathrm{d}t\Big| \le \int_{\mathbb{R}} |\partial_j f(x)| \,\mathrm{d}x_j, \\ |f(x)|^d &\le \prod_{j=1}^{d} \int_{\mathbb{R}} |\partial_j f(x)| \,\mathrm{d}x_j.\end{aligned}$$

Setting $g_j(\hat{x}^j) = (\int_{\mathbb{R}} |\partial_j f(x)| \,\mathrm{d}x_j)^{\frac{1}{d-1}}$, we deduce

$$|f(x)|^{\frac{d}{d-1}} \le \prod_{j=1}^{d} g_j(\hat{x}^j).$$

After integration over $x \in \mathbb{R}^d$, Lemma D.18 yields

$$\begin{aligned}\|f\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d)}^{\frac{d}{d-1}} &\le \int_{\mathbb{R}^d} g_1(\hat{x}^1) \cdot \ldots \cdot g_d(\hat{x}^d) \,\mathrm{d}x \le \prod_{j=1}^{d} \|g_j\|_{L^{d-1}(\mathbb{R}^{d-1})} \\ &= \prod_{j=1}^{d} \Big(\int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} |\partial_j f(x)| \,\mathrm{d}x_j \,\mathrm{d}\hat{x}^j\Big)^{\frac{1}{d-1}}, \\ \|f\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} &\le \prod_{j=1}^{d} \|\partial_j f\|_{L^1(\mathbb{R}^d)}^{\frac{1}{d}} \le \||\nabla f|_1\|_1 \le \|f\|_{1,1}.\end{aligned} \tag{D.14}$$

2) Next, let $p \in (1,d)$ and $p^* = \frac{pd}{d-p}$. Set $t_* := \frac{d-1}{d} p^* = \frac{d-1}{d-p} p > 1$. An elementary calculation shows that $(t_* - 1)p' = p^* = t_* \frac{d}{d-1}$. Set

$$g = f|f|^{t-1} = f(f\overline{f})^{\frac{t-1}{2}}$$

for $t > 1$. We compute

$$\begin{aligned}
\partial_j g &= |f|^{t-1}\partial_j f + f\tfrac{t-1}{2}(f\overline{f})^{\frac{t-1}{2}-1}((\partial_j f)\overline{f} + f(\partial_j \overline{f})) \\
&= |f|^{t-1}\partial_j f + (t-1)f|f|^{t-3}\operatorname{Re}(f\partial_j\overline{f}), \\
|g| &= |f|^t, \quad |\partial_j g| \le t|\partial_j f||f|^{t-1}.
\end{aligned}$$

Applying (D.14) to g, we estimate

$$\begin{aligned}
\|f\|_{\frac{td}{d-1}}^t &= \Big(\int_{\mathbb{R}^d} |f|^{t\frac{d}{d-1}} \,\mathrm{d}x\Big)^{\frac{d-1}{d}} = \Big(\int_{\mathbb{R}^d} |g|^{\frac{d}{d-1}} \,\mathrm{d}x\Big)^{\frac{d-1}{d}} \\
&\le \prod_{j=1}^d \|\partial_j g\|_1^{\frac{1}{d}} \le \prod_{j=1}^d t^{\frac{1}{d}} \Big(\int_{\mathbb{R}^d} |\partial_j f||f|^{t-1} \,\mathrm{d}x\Big)^{\frac{1}{d}} \\
&\le t \prod_{j=1}^d \Big(\int_{\mathbb{R}^d} |\partial_j f|^p \,\mathrm{d}x\Big)^{\frac{1}{dp}} \Big(\int_{\mathbb{R}^d} |f|^{(t-1)p'} \,\mathrm{d}x\Big)^{\frac{1}{p'd}} \\
&= t \prod_{j=1}^d \||\nabla f|_p\|_p^{\frac{1}{d}} \; \|f\|_{(t-1)p'}^{\frac{t-1}{d}} = t\||\nabla f|_p\|_p \; \|f\|_{(t-1)p'}^{t-1},
\end{aligned}$$

where we used Hölder's inequality. For $t = t^*$, we use the properties of t stated above and obtain

$$\|f\|_{p^*} \le p\frac{d-1}{d-p}\||\nabla f|_p\|_p \le p\frac{d-1}{d-p}\|f\|_{1,p}. \tag{D.15}$$

By density (see Theorem D.13), this estimate can be extended to all $f \in W_p^1(\mathbb{R}^d)$. Then the identity map on $W_p^1(\mathbb{R}^d)$ is the required embedding in (a).

3) Let $f \in C_c^1(\mathbb{R}^d)$, $p = d$, and $t > 1$. Then $p' = \frac{d}{d-1}$, and Step 2) yields

$$\|f\|_{t\frac{d}{d-1}} \le t^{\frac{1}{t}}\|f\|_{(t-1)\frac{d}{d-1}}^{1-\frac{1}{t}}\||\nabla f|_p\|_d^{\frac{1}{t}} \le c\Big(\|f\|_{(t-1)\frac{d}{d-1}} + \||\nabla f|_p\|_d\Big) \tag{D.16}$$

using Young's inequality from undergraduate courses. For $t = d$, this estimate gives $f \in L^{\frac{d^2}{d-1}}(\mathbb{R}^d)$ and

$$\|f\|_{\frac{d^2}{d-1}} \le c\|f\|_{1,d}.$$

Here and below the constant $c > 0$ does not depend on f. For $q \in (d, d\frac{d}{d-1})$, inequality (D.13) further yields

$$\|f\|_q \le c(\|f\|_d + \|f\|_{\frac{d^2}{d-1}}) \le c\|f\|_{1,d}.$$

Now, we can apply (D.16) with $t = d+1$ and obtain

$$\|f\|_{\frac{d^2+d}{d-1}} \le c(\|f\|_{\frac{d^2}{d-1}} + \||\nabla f|_p\|_d) \le c\|f\|_{1,d}.$$

As above, we see that $f \in L^q(\mathbb{R}^d)$ for $d \le q \le d\frac{d+1}{d-1}$. We can now iterate this procedure with $t_n = d + n$ and obtain

$$\|f\|_q \le c(q)\|f\|_{1,p}$$

for all $q < \infty$. Assertion (b) then follows by density.

4) Let $f \in C_c^1(\mathbb{R}^d)$, $p > d$, $Q(r) = [-\frac{r}{2}, \frac{r}{2}]^d$ for $r > 0$, and $x_0 \in Q(r)$. Set $M(r) = r^{-d} \int_{Q(r)} f \, \mathrm{d}x$. We further put $\beta := 1 - \frac{d}{p} \in (0,1)$. Using $|x - x_0|_\infty \le r$ for $x \in Q(r)$, the transformation $y = t(x - x_0)$ and Hölder's inequality, we compute

$$\begin{aligned}
|f(x_0) - M(r)| &= \Big| r^{-d} \int_{Q(r)} (f(x_0) - f(x)) \, \mathrm{d}x \Big| \\
&= r^{-d} \Big| \int_{Q(r)} \int_1^0 \frac{\mathrm{d}}{\mathrm{d}t} f(x_0 + t(x - x_0)) \, \mathrm{d}t \, \mathrm{d}x \Big| \\
&\le r^{-d} \int_{Q(r)} \int_0^1 |\nabla f(x_0 + t(x - x_0)) \cdot (x - x_0)| \, \mathrm{d}t \, \mathrm{d}x \\
&\le r^{1-d} \int_0^1 \int_{Q(r)} |\nabla f(x_0 + t(x - x_0))|_1 \, \mathrm{d}x \, \mathrm{d}t \\
&= r^{1-d} \int_0^1 \int_{t(Q(r)-x_0)} |\nabla f(x_0 + y)|_1 \, \mathrm{d}y \; t^{-d} \, \mathrm{d}t \\
&\le r^{1-d} \int_0^1 \Big(\int_{t(Q(r)-x_0)} |\nabla f(x_0 + y)|_1^p \, \mathrm{d}y \Big)^{\frac{1}{p}} \operatorname{vol}(t(Q(r) - x_0))^{\frac{1}{p'}} t^{-d} \, \mathrm{d}t \\
&\le c r^{1-d} \| |\nabla f|_p \|_{L^p(\mathbb{R}^d)} \int_0^1 r^{\frac{d}{p'}} t^{\frac{d}{p'} - d} \, \mathrm{d}t \\
&= C r^{1 - \frac{d}{p}} \| |\nabla f|_p \|_{L^p(\mathbb{R}^d)}
\end{aligned}$$

for constants $C, c > 0$ only depending on d and p, using also that $\frac{d}{p'} - d > -1$ due to $p > d$. A translation then gives

$$\Big| f(x_0 + z) - r^{-d} \int_{z + Q(r)} f(y) \, \mathrm{d}y \Big| \le C r^\beta \| |\nabla f|_p \|_{L^p(\mathbb{R}^d)} \tag{D.17}$$

for all $z \in \mathbb{R}^d$. Taking $x = z$, $x_0 = 0$, $r = 1$, and using Hölder's inequality, we thus obtain

$$\begin{aligned}
|f(x)| &\le \Big| f(x) - \int_{x+Q(1)} f \, \mathrm{d}y \Big| + \Big| \int_{x+Q(1)} f \, \mathrm{d}y \Big| \\
&\le C \| |\nabla f|_p \|_p + \|f\|_p \le c \|f\|_{1,p}
\end{aligned} \tag{D.18}$$

for all $x \in \mathbb{R}^d$, where c only depends on d and p. Given $x, y \in \mathbb{R}^d$, we find a closed cube Q of side length $|x - y|_\infty =: r$ such that $x, y \in Q$ and Q is parallel to the axes. Estimate (D.17) now yields

$$\begin{aligned}
|f(x) - f(y)| &\le \Big| f(x) - r^{-d} \int_Q f \, \mathrm{d}y \Big| + \Big| r^{-d} \int_Q f \, \mathrm{d}y - f(y) \Big| \\
&\le 2C \| |\nabla f|_p \|_p \, |x - y|_\infty^\beta \le 2C \| |\nabla f|_p \|_p \, |x - y|_2^\beta.
\end{aligned}$$

If $f \in W_p^1(\mathbb{R}^d)$, then there are $f_n \in C_c^1(\mathbb{R}^d)$ converging to f in $W_p^1(\mathbb{R}^d)$. By (D.18), f_n is a Cauchy sequence in $C_0(\mathbb{R}^d)$. Hence, f has a representative

$\tilde f \in C_0(\mathbb{R}^d)$ such that $f_n \to \tilde f$ uniformly as $n \to \infty$. So the above estimates imply that

$$\|\tilde f\|_\infty + \sup_{x\neq y} \frac{|\tilde f(x) - \tilde f(y)|}{|x-y|_2^\beta} \le c\|f\|_{1,p}.$$

The map $f \mapsto \tilde f$ is the required embedding. □

Applying Theorem D.15 and (D.15) to the 0-extension of a function in $C_c^\infty(U)$, we obtain the following result by approximation.

Corollary D.19. *Let $U \subseteq \mathbb{R}$ be open. Then the assertion of Theorem D.15 holds for all $u \in \mathring{W}_p^k(U)$ instead of $W_p^k(U)$, where one replaces $\mathbb{R}^d$ by U. If U is also bounded, we have* Poincare's inequality

$$\int_U |\nabla u|_p^p \,\mathrm{d}x \ge \delta \int_U |u|^p \,\mathrm{d}x \tag{D.19}$$

for some $\delta > 0$ and all $u \in \mathring{W}_p^1(U)$ and $p \in [1,\infty)$.

Proof. We only have to show (D.19). For $p \in [1,d)$, the estimate (D.19) follows from (D.15) since $L^{p^*}(U) \hookrightarrow L^p(U)$. Let $p \in [d,\infty)$. The case $p = d = 1$ is easy to check. For the other cases, fix $r \in (p,\infty)$ and $u \in \mathring{W}_p^1(U)$. Then (D.13), (D.14) and $\mathring{W}_p^1(U) \hookrightarrow L^r(U)$ imply

$$\|u\|_p \le c_\varepsilon\|u\|_1 + \varepsilon\|u\|_r \le c_\varepsilon\||\nabla u|_1\|_1 + c\varepsilon\|u\|_{1,p} \le c_\varepsilon\||\nabla u|_p\|_p + c\varepsilon\|u\|_p + c\varepsilon\||\nabla u|_p\|_p$$

for all $\varepsilon > 0$ and some constants $c_\varepsilon, c > 0$ independent of u, where c does not depend on $\varepsilon > 0$. Choosing a small ε, we derive (D.19). □

Next we want to extend the Sobolev embedding to $W_p^k(U)$ (although this is not really needed in the lectures). This can easily be done for "nice" domains in the following sense.

Definition D.20. *Let $k \in \mathbb{N}$. The open set $U \subseteq \mathbb{R}^d$ has the k-*extension property *if for all $m \in \{0,1,\dots,k\}$ and $p \in [1,\infty)$, there is an operator*

$$E_{m,p} \in \mathcal{B}(W_p^m(U), W_p^m(\mathbb{R}^d))$$

with $E_{m,p}f = f$ on U and $E_{m,p}f = E_{l,q}f$ for all $f \in W_p^m(U) \cap W_q^l(U)$, $1 \le m$, $l \le k$ and $1 \le p,q < \infty$. We write E_U instead of $E_{m,p}$.

Observe that $E_0 f(x) = f(x)$ for $x \in U$ and $E_0 f(x) = 0$ for $x \in \mathbb{R}^d \setminus U$ defines an isometry $E_0 : L^p(U) \to L^p(\mathbb{R}^d)$. Moreover, $R_U f = f|_U$ defines a contractive map in all spaces $\mathcal{B}(W_p^k(\mathbb{R}^d), W_p^k(U))$.

Corollary D.21. *If U has the m-extension property for some $m \in \mathbb{N}$, then Theorem D.15 and Corollary D.16 hold for $k \in \{1,\dots,m\}$ with $\mathbb{R}^d$ replaced by U and $C_0^j(\mathbb{R}^d)$ replaced by*

$$\begin{aligned} C_0^j(\overline{U}) = \{ f \in C^j(U) \,|\, &\partial^\alpha f \text{ has a continuous extension to } \partial U \\ &\text{and } \partial^\alpha f(x) \to 0 \text{ as } |x|_2 \to \infty \text{ if } U \text{ is unbounded, for all } 0 \le |\alpha| \le j \}. \end{aligned}$$

Proof. Consider e.g. Theorem D.15 (a). We have the embedding

$$J : W_p^k(\mathbb{R}^d) \hookrightarrow L^{p^*}(\mathbb{R}^d)$$

given by the identity. Then,

$$R_U J E_U : W_p^k(U) \to L^{p^*}(U)$$

is continuous and injective. The other assertions are proved in the same way. $\square$

Corollary D.22. *Let U possess the k-extension property for some $k \in \mathbb{N}$. Then $W_p^k(U) \cap C^\infty(\overline{U})$ is dense in $W_p^k(U)$ for all $1 \le p < \infty$. Here $C^\infty(\overline{U}) := \{f|_{\overline{U}} \mid f \in C^\infty(\mathbb{R}^d)\}$.*

Proof. If $f \in W_p^k(U)$, then $E_U f \in W_p^k(\mathbb{R}^d)$. By Theorem D.13, there are $g_n \in C_c^\infty(\mathbb{R}^d)$ converging to $E_U f$ in $W_p^k(\mathbb{R}^d)$. Hence, $R_U g_n \in W_p^k(U) \cap C^\infty(\overline{U})$ converges to $f = R_U E_U f$ in $W_p^k(U)$ as $n \to \infty$. $\square$

We now give a sufficient condition for the k-extension property, see e.g. [**AF03**] for improvements and more details. The definition of $\partial U \in C^1$ can be found in the proof.

Theorem D.23. *Let $U \subseteq \mathbb{R}^d$ be bounded and open with $\partial U \in C^k$. Then U has the k-extension property.*

Sketch of the proof for $k = 1$. (See also Theorem 5.22 in [**AF03**].)

1) Let

$$H_\pm = \left\{(y,t) \in \mathbb{R}^d \,\middle|\, y \in \mathbb{R}^{d-1}, t \gtrless 0\right\}$$

and $f \in W_p^1(H_-) \cap C^1(\overline{H}_-)$. Define

$$E_- f(y,r) = \begin{cases} f(y,t), & (y,t) \in \overline{H}_-, \\ 4f(y,-\frac{t}{2}) - 3f(y,-t), & (y,t) \in H_+. \end{cases}$$

Note that $E_- f \in C^1(\mathbb{R}^d)$. One can check that $\|E_- f\|_{W_p^1(\mathbb{R}^d)} \le c\|f\|_{W_p^1(H_-)}$ for a constant $c > 0$.

2) We show that $W_p^1(H_-) \cap C^1(\overline{H}_-)$ is dense in $W_p^1(H_-)$, so that E_- can be extended to a 1-extension operator on $W_p^1(H_-)$. In fact, let $f \in W_p^1(H_-)$ and $\varepsilon > 0$. By Theorem D.13, there is a function $g \in C^\infty(H_-) \cap W_p^1(H_-)$ with $\|f - g\|_{1,p} \le \varepsilon$. Setting $g_n(y,t) = g(y, t - \frac{1}{n})$ for $t \le 0$, $y \in \mathbb{R}^{d-1}$ and $n \in \mathbb{N}$, we define the functions $g_n \in C^1(\overline{H}_-) \cap W_p^1(H_-)$. Observe that

$$\partial^\alpha g_n = R_{\overline{H}_-} T_n E_0 \partial^\alpha g$$

for $0 \le |\alpha| \le 1$, where $T_n \in \mathcal{B}(L^p(\mathbb{R}^d))$ is given by $T_n h(y,t) = h(y, t - \frac{1}{n})$ for $h \in L^p(\mathbb{R}^d)$. One can see that $T_n h \to h$ in $L^p(\mathbb{R}^d)$ as in Example 2.6. Hence, g_n converges to g in $W_p^1(H_-)$ implying the claim.

3) Since $\partial U \in C^1$, there are bounded open sets $U_0, U_1, \dots, U_m \subseteq \mathbb{R}^d$ such that $U \subseteq U_0 \cup \dots \cup U_m$, $\overline{U}_0 \subseteq U$ and $\partial U \subseteq U_1 \cup \dots \cup U_m$, as well as a diffeomorphism $\Psi_j : U_j \to V_j$ such that Ψ_j' and $(\Psi_j^{-1})'$ are bounded and $\Psi_j(U_j \cap U) \subseteq H_-$ and $\Psi_j'(U_j \cap \partial U) \subseteq \mathbb{R}^{d-1} \times \{0\}$, for each $j \in \{1, \dots, m\}$. Moreover, there are

functions $0 \le \varphi_j \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi_j \subseteq U_j$ for all $j = 0, 1, \dots, m$ and $\sum_{j=0}^m \varphi_j(x) = 1$ for all $x \in \overline{U}$.

Let $j \in \{1, \dots, m\}$. Set $S_j g(y) = g(\Psi_j^{-1}(y))$ for $y \in H_- \cap V_j$ and $S_j g(y) = 0$ for $y \in H_- \setminus V_j$, where $g \in W_p^1(U_j \cap U)$. For $h \in W_p^1(\mathbb{R}^d)$, set $\hat{S}_j h(x) = h(\Psi_j(x))$ for $x \in U_j$ and $\hat{S}_j h(x) = 0$ for $x \in \mathbb{R}^d \setminus U_j$. Take any $\tilde{\varphi}_j \in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\tilde{\varphi}_j \subseteq U_j$ and $\tilde{\varphi}_j = 1$ on $\operatorname{supp}\varphi_j$ (see Lemma D.5). For $f \in W_p^1(U)$, we define

$$E_U f = E_0 \varphi_0 f + \sum_{j=1}^m \tilde{\varphi}_j \hat{S}_j E_- S_j \big(R_{(U_j \cap U)}(\varphi_j f)\big).$$

Note that $\operatorname{supp} E_U f \subseteq U_0 \cup U_1 \cup \dots \cup U_m$. Using part 2) and Proposition D.7 and D.8, we see that $E_U \in \mathcal{B}(W_p^1(U), W_p^1(\mathbb{R}^d))$. Let $x \in U$. If $x \in U_k$ for some $k \in \{1, \dots, m\}$, we have $\Psi_k(x) \in H_-$. If $x \notin U_j$, then $\tilde{\varphi}_j(x) = 0$. Thus

$$\begin{aligned} E_U f(x) &= \varphi_0(x) f(x) + \sum_{\substack{j=1,\dots,m \\ x \in U_j}} \tilde{\varphi}_j(x)(\varphi_j f)(\Psi_j^{-1}(\Psi_j(x))) \\ &= \sum_{j=0}^m \varphi_j(x) f(x) = f(x). \end{aligned}$$

If $x \in U_0 \setminus (U_1 \cup \dots \cup U_m)$, we also have $E_U f(x) = \varphi_0(x) f(x) = f(x)$. □

We add three more important results which are used in the lectures only to give additional information.

Theorem D.24 (Rellich-Kondrachov). *Let $U \subseteq \mathbb{R}^d$ be bounded and open with $\partial U \in C^k$, $k \in \mathbb{N}$, and $1 \le p < \infty$. Then the following assertions hold.*

(a) If $kp \le d$ and $1 \le q < p^* = \frac{dp}{d-kp} \in (p, \infty]$, then the embedding

$$J : W_p^k(U) \hookrightarrow L^q(U)$$

is compact. (For instance, $q = p$.)

(b) If $k - \frac{d}{p} > j \in \mathbb{N}_0$, then the embedding

$$J : W_p^k(U) \hookrightarrow C^j(\overline{U})$$

is compact.

Note that a compact embedding $J : Y \hookrightarrow X$ means that any bounded sequence $(y_n)_n$ in Y has a subsequence such that $(Jy_{n_j})_j$ converges in X. In Theorem D.24 (a), J is given by the identity, in (b) it is given by choosing the representative in C^j.

Proof of Theorem D.24. We prove the result only for $k = 1$ (and thus $j = 0$). Part (b) follows from the Arzelà-Ascoli theorem since Corollary D.21 and Theorem D.15 (c) give constants $\beta, c > 0$ such that $|f(x) - f(y)| \le c|x-y|^\beta$ and $|f(x)| \le c$ for all $x, y \in U$ and $f \in W_p^1(U)$ with $\|f\|_{1,p} \le 1$, where $p > d$.

In the case $p < d$, take $f_n \in W_p^1(U)$ with $\|f_n\|_{1,p} \le M$ for all $n \in \mathbb{N}$. Set $g_n = E_U f_n \in W_p^1(\mathbb{R}^d)$. From the proof of Theorem D.23 we know that the support of g_n belongs to a fixed open bounded set $\widetilde{U} \subseteq \mathbb{R}^d$ containing $\overline{U}$.

Moreover, $\|g_n\|_{1,p} \le \|E_U\| M =: M_1$ for all $n \in \mathbb{N}$. Take $q \in [1, p^*)$ and $\theta \in (0,1]$ with $\frac{1}{q} = \frac{\theta}{1} + \frac{1-\theta}{p^*}$. Inequality (D.13) and Theorem D.15 yield that

$$\begin{aligned}
\|f_n - f_m\|_{L^q(U)} &\le \|g_n - g_m\|_{L^q(\widetilde{U})} \le \|g_n - g_m\|^{\theta}_{L^1(\widetilde{U})} \|g_n - g_m\|^{1-\theta}_{L^{p^*}(\widetilde{U})} \\
&\le \|g_n - g_m\|^{\theta}_{L^1(\widetilde{U})} (\|g_n\|^{1-\theta}_{1,p} + \|g_m\|^{1-\theta}_{1,p}) \\
&\le 2M_1^{1-\theta} \|g_n - g_m\|^{\theta}_{L^1(\widetilde{U})}
\end{aligned}$$

for all $n, m \in \mathbb{N}$. So it suffices to construct a subsequence of g_n which converges in $L^1(\widetilde{U})$. For $x \in \widetilde{U}$, $n \in \mathbb{N}$ and $\varepsilon > 0$, we compute

$$\begin{aligned}
|g_n(x) - G_\varepsilon g_n(x)| &= \Big| \int_{\mathbb{R}^d} \psi_\varepsilon(x-y)(g_n(x) - g_n(y)) \,\mathrm{d}y \Big| \\
&\le \varepsilon^{-d} \int_{B(x,\varepsilon)} \chi_0(\tfrac{1}{\varepsilon}(x-y)) |g_n(x) - g_n(y)| \,\mathrm{d}y \\
&= \int_{B(0,1)} \chi_0(z) |g_n(x) - g_n(x - \varepsilon z)| \,\mathrm{d}z \\
&= \int_{B(0,1)} \chi_0(z) \Big| \int_\varepsilon^0 \frac{\mathrm{d}}{\mathrm{d}t} g_n(x - tz) \,\mathrm{d}t \Big| \,\mathrm{d}z \\
&\le \int_{B(0,1)} \chi_0(z) \int_0^\varepsilon |\nabla g_n(x - tz) \cdot z| \,\mathrm{d}t \,\mathrm{d}z \\
&\le \int_0^\varepsilon \int_{B(0,1)} \chi_0(z) |\nabla g_n(x - tz)|_2 \,\mathrm{d}z \,\mathrm{d}t \\
&= \int_0^\varepsilon \int_{B(x,t)} \chi_0(\tfrac{1}{t}(x-y)) |\nabla g_n(y)|_2 t^{-d} \,\mathrm{d}y \,\mathrm{d}t \\
&= \int_0^\varepsilon \|\psi_t * |\nabla g_n|_2\|_1 \,\mathrm{d}t \le \varepsilon \sup_{0 \le t \le \varepsilon} \|\psi_t\|_1 \||\nabla g_n|_2\|_1 \\
&\le c\varepsilon \||\nabla g_n|_p\|_p \le c M_1 \varepsilon,
\end{aligned}$$

where we have used the transformations $z = \frac{1}{\varepsilon}(x-y)$ and $y = x - tz$, as well as Fubini's theorem, Young's inequality (E.4), (D.6) and $L^p(\widetilde{U}) \hookrightarrow L^1(\widetilde{U})$. We thus obtain

$$\|g_n - G_\varepsilon g_n\|_{L^1(\widetilde{U})} \le c\lambda(\widetilde{U}) M_1 \varepsilon =: C\varepsilon \tag{D.20}$$

for all $n \in \mathbb{N}$ and $\varepsilon > 0$, where C only depends on $\lambda(\widetilde{U})$, $\|E_U\|$, M and d. On the other hand, the definition of $G_\varepsilon g_n$ yields

$$|G_\varepsilon g_n(x)| \le \|\psi_\varepsilon\|_\infty \|g_n\|_{L^1(\widetilde{U})} \quad \text{and} \quad |\nabla G_\varepsilon g_n(x)| \le \|\nabla \psi_\varepsilon\|_\infty \|g_n\|_{L^1(\widetilde{U})}$$

for all $x \in \widetilde{U}$ and $n \in \mathbb{N}$ and each fixed $\varepsilon > 0$. The Arzelà-Ascoli theorem now implies that the set $F_\varepsilon := \{G_\varepsilon g_n \,|\, n \in \mathbb{N}\}$ is relatively compact in $C\Big(\overline{\widetilde{U}}\Big)$ for each $\varepsilon > 0$, and thus in $L^1(\widetilde{U})$ since $C\Big(\overline{\widetilde{U}}\Big) \hookrightarrow L^1(\widetilde{U})$. Let $\delta > 0$ be given and fix $\varepsilon = \frac{\delta}{2C}$. Then there are $n_1, \dots, n_l \in \mathbb{N}$ such that

$$F_\varepsilon \subseteq \bigcup_{j=1}^{l} B_{L^1(\widetilde{U})}(G_\varepsilon g_{n_j}, \tfrac{\delta}{2}) =: \bigcup_{j=1}^{l} B_j.$$

Hence, given $n \in \mathbb{N}$, there is an index n_j such that $G_\varepsilon g_n \in B_j$. The estimates (D.20) and (D.6) then yield

$$\|g_n - G_\varepsilon g_{n_j}\|_{L^1(\widetilde{U})} \le \|g_n - G_\varepsilon g_n\|_{L^1(\widetilde{U})} + \|G_\varepsilon(g_n - g_{n_j})\|_{L^1(\widetilde{U})} \le C\varepsilon + \tfrac{\delta}{2} = \delta.$$

We have shown that for each $\delta > 0$, the set $\Gamma := \{g_n \,|\, n \in \mathbb{N}\}$ is covered by finitely many open balls B_j of radius $\frac{\delta}{2}$, i.e., Γ is totally bounded in $L^1(\widetilde{U})$. Thus Γ contains a subsequence converging in $L^1(\widetilde{U})$. In the case $p = d$ one simply replaces p^* by any $r \in (q, \infty)$. □

REMARK D.25. (a) Theorem D.24 is wrong for unbounded domains. In fact, let $d = 1$, $k \in \mathbb{N}$, $p \in [1, \infty)$ and define $f_n = f(\cdot - n)$ in $W_p^k(\mathbb{R})$ for any function $0 \ne f \in C^\infty(\mathbb{R})$ with $\operatorname{supp} f \subseteq (-\frac{1}{2}, \frac{1}{2})$. Then $\|f_n\|_{k,p}$ and $\|f_n - f_m\|_q > 0$ do not depend on $n \ne m$ in $\mathbb{N}$ so that $(f_n)_n$ is bounded in $W_p^k(\mathbb{R})$ and has no subsequence which converges in L^q with $1 \le q < p^*$.

(b) The embedding $W_p^1(U) \hookrightarrow L^{p^*}(U)$ is never compact, see Example 6.12 in [**AF03**].

PROPOSITION D.26. *Let $1 \le p < \infty$ and either $f \in \mathring{W}_p^2(U)$ or U be bounded with $\partial U \in C^2$ and $f \in W_p^2(U)$. Then there are constants $C, \varepsilon_0 > 0$ such that*

$$\Big(\sum_{j=1}^d \|\partial_j f\|_p^p\Big)^{1/p} \le \varepsilon \Big(\sum_{i,j=1}^d \|\partial_{ij} f\|_p^p\Big)^{1/p} + \tfrac{C}{\varepsilon}\|f\|_p,$$

for all $\varepsilon > 0$ if $f \in \mathring{W}_p^2(U)$ and for all $0 < \varepsilon \le \varepsilon_0$ if $f \in W_p^2(U)$.

PROOF. Let $f \in C_c^2(U)$ and extend it to $\mathbb{R}^d$ by 0. Take $j = 1$. Write $x = (t, y) \in \mathbb{R} \times \mathbb{R}^{d-1}$ for $x \in \mathbb{R}^d$. Fix $y \in \mathbb{R}^{d-1}$ and set $g(t) = f(t, y)$ for $t \in \mathbb{R}$. Let $\varepsilon > 0$ and $a, b \in \mathbb{R}$ with $b - a = \varepsilon$. Take any $r \in (a, a + \frac{\varepsilon}{3})$ and $t \in (b - \frac{\varepsilon}{3}, b)$. There there is a $\overline{s} = \overline{s}(r, t) \in (a, b)$ such that

$$|g'(\overline{s})| = \left|\frac{g(t) - g(r)}{t - r}\right| \le \tfrac{3}{\varepsilon}(|g(t)| + |g(r)|).$$

For every $s \in (a, b)$ we thus obtain

$$|g'(s)| = \Big|g'(\overline{s}) + \int_{\overline{s}}^s g''(\tau)\,\mathrm{d}\tau\Big| \le \tfrac{3}{\varepsilon}(|g(r)| + |g(t)|) + \int_a^b |g''(\tau)|\,\mathrm{d}\tau.$$

Integrating first over r and then over t, we conclude

$$\begin{aligned}
\tfrac{\varepsilon}{3}|g'(s)| &\le \frac{3}{\varepsilon}\int_a^{a+\frac{\varepsilon}{3}} |g(r)|\,\mathrm{d}r + |g(t)| + \frac{\varepsilon}{3}\int_a^b |g''(\tau)|\,\mathrm{d}\tau,\\
\tfrac{\varepsilon^2}{9}|g'(s)| &\le \int_a^{a+\frac{\varepsilon}{3}} |g(r)|\,\mathrm{d}r + \int_{b-\frac{\varepsilon}{3}}^b |g(t)|\,\mathrm{d}t + \frac{\varepsilon^2}{9}\int_a^b |g''(\tau)|\,\mathrm{d}\tau,\\
|g'(s)| &\le \frac{9}{\varepsilon^2}\int_a^b |g(\tau)|\,\mathrm{d}\tau + \int_a^b |g''(\tau)|\,\mathrm{d}\tau\\
&\le \varepsilon^{\frac{1}{p'}}\frac{9}{\varepsilon^2}\Big(\int_a^b |g(\tau)|^p\,\mathrm{d}\tau\Big)^{1/p} + \varepsilon^{\frac{1}{p'}}\Big(\int_a^b |g''(\tau)|^p\,\mathrm{d}\tau\Big)^{1/p}\\
&\le \varepsilon^{\frac{p-1}{p}} 2^{\frac{p-1}{p}}\bigg(\Big(\frac{9}{\varepsilon^2}\Big)^p \int_a^b |g(\tau)|^p\,\mathrm{d}\tau + \int_a^b |g''(\tau)|\,\mathrm{d}\tau\bigg)^{1/p},
\end{aligned}$$

where we used Hölder's inequality first for the integrals and then in $\mathbb{R}^2$. We take now the p-th power and then integrate over s arriving at

$$\int_a^b |g'(s)|^p \,\mathrm{d}s \le \varepsilon\varepsilon^{p-1}2^{p-1}\Big(\frac{9^p}{\varepsilon^{2p}}\int_a^b |g(\tau)|^p \,\mathrm{d}\tau + \int_a^b |g''(\tau)|^p \,\mathrm{d}\tau\Big).$$

Now choose $a = a_k = k\varepsilon$ and $b = b_k = (k+1)\varepsilon$ for $k \in \mathbb{Z}$. Summing the integrals on $[k\varepsilon, (k+1)\varepsilon)$ for $k \in \mathbb{Z}$ and then integrating over $y \in \mathbb{R}^{d-1}$, it follows that

$$\int_{\mathbb{R}} |g'(\tau)|^p \,\mathrm{d}\tau \le \varepsilon^p 2^{p-1}\Big(\frac{9^p}{\varepsilon^{2p}}\int_{\mathbb{R}} |g(\tau)|^p \,\mathrm{d}\tau + \int_{\mathbb{R}} |g''(\tau)|^p \,\mathrm{d}\tau\Big),$$

$$\int_U |\partial_1 f|^p \,\mathrm{d}x \le (2\varepsilon)^p \int_U |\partial_{11} f|^p \,\mathrm{d}x + \frac{36^p}{(2\varepsilon)^p}\int_U |f|^p \,\mathrm{d}x. \tag{D.21}$$

By approximation, (D.21) can be established for all $f \in \mathring{W}_p^2(U)$. The same result holds for $\partial_j f$ and $\partial_{jj} f$ with $j \in \{2, \dots, d\}$. We now replace 2ε by ε, sum over j and take the p-th root to arrive at

$$\begin{aligned}\Big(\sum_{j=1}^d \|\partial_j f\|_p^p\Big)^{1/p} &\le \Big(\varepsilon^p \sum_{j=1}^d \|\partial_{jj} f\|_p^p + \tfrac{36^p}{\varepsilon^p}\|f\|_p^p\Big)^{1/p} \\ &\le \varepsilon\Big(\sum_{j=1}^d \|\partial_{jj} f\|_p^p\Big)^{1/p} + \tfrac{36}{\varepsilon}\|f\|_p,\end{aligned} \tag{D.22}$$

for all $f \in \mathring{W}_p^2(U)$, as asserted. If $u \in W_p^2(U)$ and U is bounded with $\partial U \in C^2$, we use the extension operator $E_U \in \mathcal{B}(W_p^2(U), W_p^2(\mathbb{R}^d))$ from Theorem D.23 to deduce from (D.22) with $U = \mathbb{R}^d$ that

$$\begin{aligned}\Big(\sum_{j=1}^d \|\partial_j f\|_{L^p(U)}^p\Big)^{1/p} &\le \Big(\sum_{j=1}^d \|\partial_j E_U f\|_{L^p(\mathbb{R}^d)}^p\Big)^{1/p} \\ &\le \varepsilon\Big(\sum_{j=1}^d \|\partial_{jj} E_U f\|_{L^p(\mathbb{R}^d)}^p\Big)^{1/p} + \tfrac{36}{\varepsilon}\|E_U f\|_{L^p(\mathbb{R}^d)} \\ &\le \varepsilon\|E_U f\|_{W_p^2(\mathbb{R}^d)} + \tfrac{36}{\varepsilon}\|E_U f\|_{L^p(\mathbb{R}^d)} \\ &\le c\varepsilon\|f\|_{W_p^2(U)} + \tfrac{c}{\varepsilon}\|f\|_{L^p(U)} \\ &\le c_0\varepsilon\Big(\sum_{i,j=1}^d \|\partial_{ij} f\|_p^p\Big)^{1/p} + c_1\varepsilon\Big(\sum_{j=1}^d \|\partial_j f\|_p^p\Big)^{1/p} + \tfrac{c}{\varepsilon}\|f\|_p\end{aligned}$$

where we assume that $\varepsilon \in (0, 1]$. The constants c, c_0, c_1 do not depend on ε or f. Choosing $\varepsilon_1 = \min\{\frac{1}{2c_1}, 1\}$ we arrive at

$$\frac{1}{2}\Big(\sum_{j=1}^d \|\partial_j f\|_p^p\Big)^{1/p} \le c_0\varepsilon\Big(\sum_{i,j=1}^d \|\partial_{ij} f\|_p^p\Big)^{1/p} + \tfrac{c}{\varepsilon}\|f\|_p$$

if $0 < \varepsilon \le \varepsilon_1$. This inequality implies the assertion for U with $\partial U \in C^2$. □

To study boundary values of functions in $W_p^1(U)$, we need the following result.

THEOREM D.27 (Trace theorem). *Let $U \subseteq \mathbb{R}^d$ be bounded and open with $\partial U \in C^1$, and let $p \in [1,\infty)$. Then the trace map $f \mapsto f|_{\partial U}$ from $W_p^1(U) \cap C(\overline{U})$ to $L^p(\partial U, \sigma)$ has a bounded linear extension* $\operatorname{tr} : W_p^1(U) \to L^p(\partial U, \sigma)$ *whose kernel is $\mathring{W}_p^1(U)$. Here σ is the surface measure on ∂U.*

PROOF. 1) Let $u \in C^1(\overline{U})$. By the definition of the surface integral, there are finitely many diffeomorphisms $\Psi_j : \widetilde{U}_j \to \widetilde{V}_j$ and $\varphi_j \in C_c^1(\widetilde{U}_j)$ with $0 \le \varphi_j \le 1$ such that $\|u\|_{L^p(\partial U,\sigma)}^p$ is dominated by

$$c \sum_{j=1}^{m} \int_{V_j} \varphi_j \circ \Psi_j^{-1} \, |u \circ \Psi_j^{-1}|^p \, \mathrm{d}y'$$

where $\widetilde{U}_j$ and $\widetilde{V}_j$ are open subsets of $\mathbb{R}^d$, the sets $\widetilde{U}_j$ cover ∂U, the functions φ_j form a partition of unity subordinated to $\widetilde{U}_j$, $V_j := \Big\{(y', y_d) \in \widetilde{V_j} \,\big|\, y_d = 0\Big\}$, $V_{j+} := \Big\{(y', y_d) \in \widetilde{V_j} \,\big|\, y_d > 0\Big\}$, $\Psi_j(\widetilde{U}_j \cap \partial U) = V_j$, and $\Psi_j(\widetilde{U}_j \cap U) = V_{j+}$. We set $v = u \circ \Psi_j^{-1}$ and $\psi = \varphi_j \circ \Psi_j^{-1} \in C_c^1(\widetilde{V}_j)$ and drop the indices j below. By means of Fubini's theorem and the fundamental theorem of calculus, we compute

$$\begin{aligned} \int_V \psi \, |v(y')|^p \, \mathrm{d}y' &= -\int_{V_+} \partial_d(\psi \, |v|^p) \, \mathrm{d}y = -\int_{V_+} [(\partial_d \psi) \, |v|^p + p\psi |v|^{p-2} \operatorname{Re}(\overline{v} \partial_d v)] \, \mathrm{d}y \\ &\le c \int_{V_+} [|v|^p + |v|^{p-1} \, |\partial_d v|] \, \mathrm{d}y \le c \, \|v\|_p^p + c \, \|v\|_p^{p-1} \, \|\partial_d v\|_p \\ &\le c \, (\|v\|_p^p + \|\partial_d v\|_p^p) \le c \, \|v\|_{W_p^1(V_+)}^p \le c \, \|u\|_{W_p^1(U)}^p \end{aligned}$$

for constants c independent of v. Here we also used Hölder's and Young's inequality, Proposition D.8 and the transformation rule. As a result, the map $\operatorname{tr} : (C^1(\overline{U}), \|\cdot\|_{1,p}) \to L^p(\partial U, \sigma)$, $\operatorname{tr} u = u|_{\partial U}$, is continuous. Corollary D.22 and Theorem D.23 allow to extend tr to an operator in $\mathcal{B}(W_p^1(U), L^p(\partial U, \sigma))$. If we start with $u \in W_p^1(U) \cap C(\overline{U})$, then we can define an approximating sequence of $u_n \in C^1(\overline{U})$ which converge to u in $W_p^1(U)$ and in $C(\overline{U})$, see the proof of Theorem 5.3.3 in [**Eva10**]. Hence, $\operatorname{tr} u_n = u_n|_{\partial U}$ tends to $u|_{\partial U}$ uniformly on ∂U and to $\operatorname{tr} u$ in $L^p(\partial U, \sigma)$, so that $\operatorname{tr} u = u|_{\partial U}$.

2a) We next observe that the inclusion $\mathring{W}_p^1(U) \subseteq \mathrm{N}(\operatorname{tr})$ is a consequence of the continuity of tr since tr vanishes on $C_c^\infty(U)$ and this space is dense in $\mathring{W}_p^1(U)$ by definition. To prove the converse, we start with the model case that $v \in W_p^1(V_+)$ has a compact support in V_+ and $\operatorname{tr} v = 0$. Our density results yield $v_n \in C^1(\overline{V_+})$ converging to v in $W_p^1(V_+)$, and hence $\operatorname{tr} v_n = v_n|_V \to 0$ in $L^p(V)$, as $n \to \infty$. Observe that

$$\begin{aligned} |v_n(y', y_d)| &\le |v_n(y', 0)| + \int_0^{y_d} |\partial_d v_n(y', s)| \, \mathrm{d}s, \\ |v_n(y', y_d)|^p &\le 2 \, |v_n(y', 0)|^p + 2 \Big(\int_0^{y_d} |\partial_d v_n(y', s)| \, \mathrm{d}s \Big)^p \end{aligned}$$

for $y' \in V$ and $y_d \geq 0$. Integrating over y' and employing Hölder's inequality, we obtain

$$\int_V |v_n(y', y_d)|^p \,\mathrm{d}y' \leq 2\int_V |v_n(y',0)|^p \,\mathrm{d}y' + 2y_d^{p-1}\int_V \int_0^{y_d} |\partial_d v_n(y', s)|^p \,\mathrm{d}s \,\mathrm{d}y'.$$

We can now let $n \to \infty$ and arrive at

$$\int_V |v(y', y_d)|^p \,\mathrm{d}y' \leq 2y_d^{p-1}\int_V \int_0^{y_d} |\partial_d v(y', s)|^p \,\mathrm{d}s \,\mathrm{d}y', \tag{D.23}$$

for a.e. $y_d > 0$. We next use a cutoff argument to obtain a support in the interior of V_+. Choose a function $\chi \in C^\infty(\mathbb{R}_+)$ such that $\chi = 0$ on $[0,1]$ and $\chi = 1$ on $[2,\infty)$. Set $\chi_n(s) = \chi(ns)$ for $s \geq 0$ and $n \in \mathbb{N}$, and define $w_n = \chi_n v$ on V_+. Observe that $w_n \to v$ in $L^p(V_+)$ as $n \to \infty$. It holds $\partial_j w_n = \chi_n \partial_j v$ for $j = 1,\dots,d-1$ and $\partial_d w_n = \chi_n \partial_d v + n\chi'(n\,\cdot\,)v$. Estimate (D.23) further implies

$$\begin{aligned}
\int_{V_+} |\nabla w_n - \nabla v|_p^p \,\mathrm{d}y &\leq c\int_0^{\frac{2}{n}}\int_V |1-\chi_n|^p |\nabla v|_p^p \,\mathrm{d}y' \,\mathrm{d}s + cn^p \int_0^{\frac{2}{n}}\int_V |v(y',s)|^p \,\mathrm{d}y' \,\mathrm{d}s \\
&\leq c\int_0^{\frac{2}{n}}\int_V |\nabla v|_p^p \,\mathrm{d}y' \,\mathrm{d}s + cn^p \int_0^{\frac{2}{n}} s^{p-1}\int_0^s \int_V |\partial_d v(y',\tau)|^p \,\mathrm{d}y' \,\mathrm{d}\tau \,\mathrm{d}s \\
&\leq c\int_0^{\frac{2}{n}}\int_V |\nabla v|_p^p \,\mathrm{d}y' \,\mathrm{d}s + c\int_0^{\frac{2}{n}}\int_V |\partial_d v(y',\tau)|^p \,\mathrm{d}y' \,\mathrm{d}\tau'
\end{aligned}$$

for some constants $c > 0$. Because of $v \in W_p^1(V_+)$, the above integrals tend to 0 as $n \to \infty$, and so $w_n \to v$ in $W_p^1(V_+)$ as $n \to \infty$. Since $w_n = 0$ for $y_d \in (0, \frac{1}{n}]$, we can mollify w_n to obtain a function $\widehat{w}_n \in C_c^\infty(V_+)$ such that $\|\widehat{w}_n - w_n\|_{1,p} \leq \frac{1}{n}$. This means that $\widehat{w}_n \to v$ in $W_p^1(V_+)$ as $n \to \infty$.

2b) We come back to $u \in W_p^1(U)$ and consider the sets $\widetilde{U}_j$ and $\widetilde{V}_j$ and the functions Ψ_j and φ_j from step 1). Let $v_j = (\varphi_j u) \circ (\Psi_j^{-1})$. First, observe that the trace of v_j to the set V_j is given by $(\operatorname{tr} \varphi_j \operatorname{tr} u) \circ \Psi_j^{-1}$ if $u \in C(\overline{U})$, in addition. By continuity one can extend this identity to all $u \in W_p^1(U)$. Let $\operatorname{tr} u = 0$. Then we can apply part 2a) to v_j and obtain $\widehat{w}_n^j \in C_c^1(V_{j+})$ converging to v_j in $W_p^1(V_{j+})$. The function

$$\widehat{u}_n = \sum_{j=1}^m \widehat{w}_n^j \circ (\Psi_j|_{U\cap\widetilde{U}_j})$$

thus belongs to $C_c^1(U)$ and converges to u in $W_p^1(U)$ as $n \to \infty$. Since $\widehat{u}_n$ has compact support, we can mollify $\widehat{u}_n$ to a function $u_n \in C_c^\infty(U)$ with $\|\widehat{u}_n - u_n\|_{1,p} \leq \frac{1}{n}$. This means that $u_n \to u$ in $W_p^1(U)$ as $n \to \infty$, and hence $u \in \mathring{W}_p^1(U)$. □

We can thus say that $u \in \mathring{W}_p^1(U)$ is 0 at the boundary in the "sense of trace". Observe that for a Borel set $\Gamma \in \partial U$ we can define the trace of $u \in W_p^1(U)$ on Γ by $\operatorname{tr}_\Gamma u = R_\Gamma \operatorname{tr} u$, where $R_\Gamma : L^p(\partial U, \sigma) \to L^p(\Gamma, \sigma)$ is the contractive restriction map. Finally, for $p > d$ we have $W_p^1(U) \hookrightarrow C(\overline{U})$ by Theorem D.23 and Corollary D.21 so that Theorem D.27 is trivial in this case. Theorem D.10 further implies that $W_1^1(a,b) \hookrightarrow C([a,b])$ for $d = 1$.

Theorem D.28 (Gauß' formula). *Let $U \subseteq \mathbb{R}^d$ be bounded and open with $\partial U \in C^2$. Let $p \in [1, \infty]$, $F \in W^1_p(U)^d$ and $\varphi \in W^1_{p'}(U)$. We then have*

$$\int_U \operatorname{div}(F)\varphi \,\mathrm{d}x = -\int_U F \cdot \nabla\varphi \,\mathrm{d}x + \int_{\partial U} \nu \cdot F\varphi \,\mathrm{d}\sigma \tag{D.24}$$

where ν is the outer unit normal of ∂U. If $u \in W^2_p(U)$ and $v \in W^2_{p'}(U)$, with $F = \nabla u$ we deduce Green's formula

$$\int_U (\Delta u)v \,\mathrm{d}x = \int_U u\Delta v \,\mathrm{d}x + \int_{\partial U} (\partial_\nu u)v \,\mathrm{d}\sigma - \int_{\partial U} u\partial_\nu v \,\mathrm{d}\sigma. \tag{D.25}$$

Here we omit the trace map in the boundary integrals and set $\partial_\nu u = \sum_{j=1}^d \nu_j \operatorname{tr} \partial_j u$.

Proof. Gauß' formula (D.24) holds for $F \in C^1(\overline{U})^d$ and $\varphi \in C^1(\overline{U})$, as shown in analysis courses. For $F \in W^1_p(U)^d$ and $\varphi \in W^1_{p'}(U)$ there are $F_n \in C^1(\overline{U})^d$ and $\varphi_n \in C^1(\overline{U})$ converging to F and φ in $W^1_p(U)^d$ and $W^1_{p'}(U)$ respectively, as long as $p, p' < \infty$, due to Corollary D.22 and Theorem D.23. If say $p' = \infty$, then we observe that $\varphi \in C(\overline{U})$ by Corollary D.21 and Theorem D.23. Hence φ can be extended to a function $\varphi \in C_c(\mathbb{R}^d)$. Set $\varphi_n = G_{\frac{1}{n}}\varphi \in C_c^\infty(\mathbb{R}^d)$. Properties (D.6) and (D.8) and Lemma D.6 then imply that $\varphi_n \to \varphi$ in $C(\overline{U})$, $\|\nabla\varphi_n\|_\infty \le \|\nabla\varphi\|_\infty$ and $\nabla\varphi_n \to \nabla\varphi$ pointwise a.e., as $n \to \infty$, where we possibly pass to a subsequence also denoted by $(\varphi_n)_n$. The case $p = \infty$ is treated similarly.

First, let $p \in (1, \infty)$ and hence also $p' \in (1, \infty)$. Theorem D.27 yields that $F_n|_{\partial U} \to \operatorname{tr} F$ in $L^p(U, \sigma)^d$ and $\varphi_n|_{\partial U} \to \operatorname{tr}\varphi$ in $L^{p'}(U, \sigma)$ as $n \to \infty$. Since $\partial_j : W^1_q(U) \to L^q(U)$ is continuous for $q \in \{p, p'\}$, we further obtain that the terms with derivatives converge in L^p, respectively in $L^{p'}$. Formula (D.24) now follows by approximation.

For the other cases, let $p = 1$, say. Then $\operatorname{div} F_n$ and F_n converge to $\operatorname{div} F$ and F in $L^1(U)$ and $L^1(U)^d$, respectively, as well as $F_n|_{\partial U} \to \operatorname{tr} F$ in $L^1(\partial U, \sigma)^d$. Due to the above listed properties of φ_n for $p' = \infty$, the assertion now follows, where we use

$$\Big|\int_U F_n \cdot \nabla\varphi_n \,\mathrm{d}x - \int_U F \cdot \nabla\varphi \,\mathrm{d}x\Big| \le \|F_n - F\|_1 \|\nabla\varphi_n\|_\infty + \int_U |F||\nabla\varphi_n - \nabla\varphi| \,\mathrm{d}x$$

and Lebesgue's theorem for the integral on the right-hand side.

Green's formula (D.25) is a straightforward consequence of (D.24). □

APPENDIX E

The Fourier transform

The Fourier transform is a fundamental tool in many branches of mathematics and its applications. We treat its basic properties in an L^2 context and then establish the important links between the Fourier transform and the Sobolev spaces $W_2^k(\mathbb{R}^d)$.

DEFINITION E.1. *Let* $f \in L^1(\mathbb{R}^d)$ *and* $\xi \in \mathbb{R}^d$. *Then*

$$\widehat{f}(\xi) = (\mathcal{F}f)(\xi) := (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\,\xi\cdot x} f(x)\,\mathrm{d}x \tag{E.1}$$

is the Fourier transform *of* f, *where* $\xi \cdot x := \sum_{k=1}^d \xi_k x_k$ *for* $\xi = (\xi_k)_k \in \mathbb{C}^d$ *and* $x = (x_k)_k \in \mathbb{R}^d$, *hence* $|x|_2^2 = x \cdot x$ *for* $x \in \mathbb{R}^d$.

We set $\varphi(x,\xi) = \mathrm{e}^{-\mathrm{i}\,\xi\cdot x} f(x)$. Observe that $|\varphi(x,\xi)| = |f(x)|$ is integrable in $x \in \mathbb{R}^d$ for every $\xi \in \mathbb{R}^d$ and that $\mathbb{R}^d \ni \xi \mapsto \varphi(\xi, x)$ is continuous for a.e. $x \in \mathbb{R}^d$. Using a corollary of the theorem of dominated convergence, we thus conclude that $\widehat{f}$ is continuous on $\mathbb{R}^d$ and

$$\|\widehat{f}\|_\infty \le (2\pi)^{-\frac{d}{2}} \|f\|_1 \quad \text{for each } f \in L^1(\mathbb{R}^d). \tag{E.2}$$

EXAMPLE E.2. (a) Let $d = 1$ and $f = \mathbb{1}_{[a,b]}$. We then have $\widehat{f}(0) = \frac{(b-a)}{\sqrt{2\pi}}$ and

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-\mathrm{i}\xi x}\,\mathrm{d}x = \frac{\mathrm{i}(\mathrm{e}^{-\mathrm{i}b\xi} - \mathrm{e}^{-\mathrm{i}a\xi})}{\sqrt{2\pi}\xi}, \quad \xi \ne 0. \qquad \diamondsuit$$

(b) Let $\gamma(x) = \exp(-\frac{1}{2}|x|_2^2)$ for $x \in \mathbb{R}^d$ be the standard Gaussian. We show that γ is a fixed point of the Fourier transform, i.e., $\widehat{\gamma} = \gamma$. Let $\xi \in \mathbb{R}^d$. Observe that $\frac{1}{2}(x + \mathrm{i}\xi)\cdot(x + \mathrm{i}\xi) = \frac{1}{2}|x|_2^2 + \mathrm{i}\xi\cdot x - \frac{1}{2}|\xi|_2^2$. We then obtain

$$\begin{aligned}
\widehat{\gamma}(\xi) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-(\mathrm{i}\xi\cdot x + \frac{1}{2}|x|_2^2)}\,\mathrm{d}x = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\frac{1}{2}|\xi|_2^2} \mathrm{e}^{-\frac{1}{2}(x+\mathrm{i}\xi)\cdot(x+\mathrm{i}\xi)}\,\mathrm{d}x \\
&= \mathrm{e}^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^d \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{2}(x_k+\mathrm{i}\xi_k)^2}\,\mathrm{d}x_k = \mathrm{e}^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^d \frac{1}{\sqrt{2\pi}} \int_{\mathrm{i}\xi_k+\mathbb{R}} \mathrm{e}^{-\frac{1}{2}z^2}\,\mathrm{d}z \\
&= \mathrm{e}^{-\frac{1}{2}|\xi|_2^2} \prod_{k=1}^d \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{2}t^2}\,\mathrm{d}t = \gamma(\xi),
\end{aligned}$$

using the formula $\int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{2}t^2}\,\mathrm{d}t = \sqrt{2\pi}$ from undergraduate courses. In the penultimate equality we shifted the path of integration in the complex plane. To justify this shift, we fix $\xi \in \mathbb{R}\setminus\{0\}$ and use the rectangular path Γ_n with vertices $-n$, n, $n+\mathrm{i}\xi$ and $-n+\mathrm{i}\xi$. Cauchy's theorem yields $\int_{\Gamma_n} \mathrm{e}^{-\frac{1}{2}z^2}\,\mathrm{d}z = 0$. The two vertical lines $S_n^\pm$ in Γ_n have length $|\xi|$, and on $S_n^\pm$ it holds

$$\big|\mathrm{e}^{-\frac{1}{2}z^2}\big| = \mathrm{e}^{-\frac{1}{2}\operatorname{Re}(\pm n+\mathrm{i}\tau)^2} \le \mathrm{e}^{-\frac{1}{2}n^2}\mathrm{e}^{\frac{1}{2}|\xi|^2},$$

where $0 \le |\tau| \le |\xi|$. Hence, $\int_{S_n^\pm} \mathrm{e}^{-\frac{1}{2}z^2}\,\mathrm{d}z$ tends to 0 as $n \to \infty$, and the above used equality is true. $\diamondsuit$

Let $f \in L^p(\mathbb{R}^d)$, $x \in \mathbb{R}^d$, and $1 \le p < \infty$. For any $t \in \mathbb{R}^d$, we introduce the *translation operator* T_t by $(T_t f)(x) = f(x+t)$. As in Exercise 2.2 (cf. Example 2.6) one sees that $T_t : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ is an isometric isomorphism with inverse T_{-t}. For $a > 0$ we further define the *dilation operator* D_a by $(D_a f)(x) = f(ax)$. Observe that $D_{1/a}D_a = D_a D_{1/a} = I$ and that the substitution $y = ax$ yields

$$\|D_a f\|_p^p = \int_{\mathbb{R}^d} |f(ax)|^p\,\mathrm{d}x = \int_{\mathbb{R}^d} a^{-d}|f(y)|^p\,\mathrm{d}y = a^{-d}\|f\|_p^p\,.$$

As a result, $D_a : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ is isomorphic with $\|D_a\| = a^{-\frac{d}{p}}$. We further set $\mathrm{e}_t(x) = \mathrm{e}^{\mathrm{i}t\cdot x}$ for all $t, x \in \mathbb{R}^d$.

Let $p \in [1,\infty]$. For $f \in L^1(\mathbb{R}^d)$ and $g \in L^p(\mathbb{R}^d)$ we define the *convolution*

$$f * g(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\,\mathrm{d}y, \quad x \in \mathbb{R}^d. \tag{E.3}$$

In Theorem 4.15 of [**Bre11**] it is shown that $f * g(x)$ indeed exists for a.e. $x \in \mathbb{R}^d$ and belongs to $L^p(\mathbb{R}^d)$ satisfying *Young's inequality*

$$\|f * g\|_p \le \|f\|_1 \|g\|_p. \tag{E.4}$$

In the proof of Theorem 4.15 in [**Bre11**] it is also seen that the map $\mathbb{R}^{2d} \ni (x,y) \mapsto f(x-y)g(y)$ is integrable if $p = 1$.

PROPOSITION E.3. *Let $f, g \in L^1(\mathbb{R}^d)$, $t \in \mathbb{R}^d$ and $a > 0$. The following formulas hold.*

(a) $\mathcal{F}(T_t f) = \mathrm{e}_t \widehat{f}$,
(b) $\mathcal{F}(\mathrm{e}_t f) = T_{-t}\widehat{f}$,
(c) $\mathcal{F}(D_a f) = a^{-d} D_{1/a}\widehat{f}$,
(d) $\mathcal{F}(f * g) = (2\pi)^{\frac{d}{2}} \widehat{f}\widehat{g}$.

PROOF. Let $f, g \in L^1(\mathbb{R}^d)$, $t \in \mathbb{R}^d$, $a > 0$, and $\xi \in \mathbb{R}^d$. Using the substitutions $y = x + t$ and $z = ax$, we check the assertions (a), (b), (c) by the calculations

$$\begin{aligned}
\mathcal{F}(T_t f)(\xi) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x+t)\,\mathrm{d}x = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot(y-t)} f(y)\,\mathrm{d}y \\
&= \mathrm{e}^{\mathrm{i}\xi\cdot t}\widehat{f}(\xi), \\
\mathcal{F}(\mathrm{e}_t f)(\xi) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot x}\mathrm{e}^{\mathrm{i}t\cdot x} f(x)\,\mathrm{d}x = \widehat{f}(\xi - t), \\
\mathcal{F}(D_a f)(\xi) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(ax)\,\mathrm{d}x = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} a^{-d}\mathrm{e}^{-\mathrm{i}\frac{1}{a}\xi\cdot z} f(z)\,\mathrm{d}z \\
&= a^{-d}\widehat{f}(\tfrac{1}{a}\xi).
\end{aligned}$$

To prove (d), we first recall that $f * g \in L^1(\mathbb{R}^d)$ and that the map $\mathbb{R}^{2d} \ni (x, y) \mapsto f(y-x)g(x)$ is integrable. Hence, the function mapping (x, y) to $\mathrm{e}^{-\mathrm{i}\xi\cdot y} f(y-x)g(x)$ belongs to $L^1(\mathbb{R}^{2d})$. Fubini's theorem thus yields

$$\begin{aligned}
\mathcal{F}(f * g)(\xi) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot y} f(y-x)g(x)\,\mathrm{d}x\,\mathrm{d}y \\
&= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot(y-x)} f(y-x)\mathrm{e}^{-\mathrm{i}\xi\cdot x} g(x)\,\mathrm{d}y\,\mathrm{d}x \\
&= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot z} f(z)\,\mathrm{d}z \Big) \mathrm{e}^{-\mathrm{i}\xi\cdot x} g(x)\,\mathrm{d}x = (2\pi)^{\frac{d}{2}} \widehat{f}(\xi)\widehat{g}(\xi),
\end{aligned}$$

where we also employed the substitution $z = y - x$. □

The following example indicates the usefulness of the above basic properties of the Fourier transform.

EXAMPLE E.4. We set $f(x) = \exp(-\frac{a}{2}|x-v|^2)$ for all $x \in \mathbb{R}^d$ and some $a > 0$ and $v \in \mathbb{R}^d$. The Fourier transform of this Gaussian function is given by $\widehat{f}(\xi) = a^{-\frac{d}{2}} \exp(-\mathrm{i}v\cdot\xi) \exp(-\frac{1}{2a}|\xi|_2^2)$ for all $\xi \in \mathbb{R}^d$. In fact, we have $f = T_{-v} D_{\sqrt{a}} \gamma$ with γ from Example E.2. Proposition E.3 and Example E.2 thus yield

$$\widehat{f} = \mathrm{e}_{-v} \mathcal{F}(D_{\sqrt{a}}\gamma) = \mathrm{e}_{-v} a^{-\frac{d}{2}} D_{\frac{1}{\sqrt{a}}} \widehat{\gamma} = a^{-\frac{d}{2}} \mathrm{e}_{-v} D_{\frac{1}{\sqrt{a}}} \gamma,$$

as asserted. ♢

One of the main properties of the Fourier transform is that it transforms derivatives into multiplication by polynomials, and vice versa. To formulate this fact concisely, we use the *multi index notation*: Let $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ and $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. We then set

$$|\alpha| := \alpha_1 + \ldots + \alpha_d, \quad x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_d^{\alpha_d}, \quad \partial^\alpha := \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}\,.$$

We further denote the function $\mathbb{R}^d \ni x \to x^\alpha f(x)$ by $x^\alpha f$. Observe that

$$|x^\alpha| = |x_1|^{\alpha_1} \cdot \ldots \cdot |x_d|^{\alpha_d} \le |x|_2^{|\alpha|} \le 1 + |x|_2^m \tag{E.5}$$

for $x \in \mathbb{R}^d$ and $|\alpha| \le m$.

To relate the Fourier transform with derivatives we need a space of smooth functions. Unfortunately, the space $C_c^\infty(\mathbb{R}^d)$ is not invariant under the Fourier transform. Instead one uses the (somewhat less convenient) “Schwartz space” on which $\mathcal{F}$ becomes a bijection, as seen below.

DEFINITION E.5. *For $f \in C^\infty(\mathbb{R}^d)$, $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$, we set*

$$p_{m,\alpha}(f) = \sup_{x\in\mathbb{R}^d} |x|_2^{2m} |\partial^\alpha f(x)|.$$

We define the Schwartz space $\mathcal{S}_d$ *by*

$$\mathcal{S}_d = \Big\{ f \in C^\infty(\mathbb{R}^d) \,\Big|\, p_{m,\alpha}(f) < \infty \text{ for all } m \in \mathbb{N}_0,\ \alpha \in \mathbb{N}_0^d \Big\}.$$

Notice that $\mathcal{S}_d$ is a vector space and that all derivatives of $f \in \mathcal{S}_d$ decay faster than $|x|_2^{-2m}$ for any $m \in \mathbb{N}$, as $|x|_2 \to \infty$. One thus calls $f \in \mathcal{S}_d$ *rapidly decreasing*. It is straightforward to check that the function $\gamma = \mathrm{e}^{-\frac{1}{2}|\cdot|_2^2}$ belongs to $\mathcal{S}_d$.

REMARK E.6. (a) Let $f \in \mathcal{S}_d$, $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$. It then holds

$$\begin{aligned}|x|_2^{2m}\,|\partial^\alpha f(x)| &= (1+|x|_2^{2d})^{-1}\,(|x|_2^{2m}+|x|_2^{2(d+m)})\,|\partial^\alpha f(x)|\\ &\le (1+|x|_2^{2d})^{-1}\,(p_{m,\alpha}(f)+p_{d+m,\alpha}(f))\end{aligned}$$

for all $x \in \mathbb{R}^d$. Since the function $x \mapsto (1+|x|_2^{2d})^{-1}$ is integrable on $\mathbb{R}^d$, we deduce $|x|_2^{2m}\partial^\alpha f \in L^1(\mathbb{R}^d) \cap C_0(\mathbb{R}^d)$, and hence $|x|_2^{2m}\partial^\alpha f \in L^p(\mathbb{R}^d)$ for all $p \in [1,\infty]$ by (D.13). $\diamondsuit$

(b) Because $C_c^\infty(\mathbb{R}^d) \subseteq \mathcal{S}_d \subseteq L^p(\mathbb{R}^d)$ the space $\mathcal{S}_d$ is dense in $L^p(\mathbb{R}^d)$ for every $p \in [1,\infty)$. $\diamondsuit$

(c) Observe that $p_{m,\alpha}$ is a seminorm on $\mathcal{S}_d$ for all $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$, where $p_{0,0}$ is the supnorm. We order these seminorms as a sequence $(p_j)_{j\in\mathbb{N}}$. It can be seen that $\mathcal{S}_d$ has the metric

$$d(f,g) = \sum_{j=1}^\infty 2^{-j}\,\frac{p_j(f-g)}{1+p_j(f-g)},$$

and that $d(f_n,f) \to 0$ as $n \to \infty$ if and only if $p_{m,\alpha}(f-f_n) \to 0$ as $n \to \infty$ for all $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^d$. One can verify that $\mathcal{S}_d$ is complete for this metric. $\diamondsuit$

Below we use the *Laplace operator* given by $\Delta = \partial_1^2 + \ldots + \partial_d^2$. We now come to the announced relation between Fourier transform and derivatives.

LEMMA E.7. *Let $f, g \in \mathcal{S}_d$ and $\alpha \in \mathbb{N}_0^d$. Then the following assertions hold.*

(a) $\widehat{f} \in C^\infty(\mathbb{R}^d)$, $\partial^\alpha \widehat{f} = (-\mathrm{i})^{|\alpha|}\mathcal{F}(x^\alpha f)$, $\mathcal{F}(\partial^\alpha f) = \mathrm{i}^{|\alpha|}\xi^\alpha \widehat{f}$.

(b) $\mathcal{F}\Delta f = \mathcal{F}\partial_1^2 f + \ldots + \mathcal{F}\partial_d^2 f = \mathrm{i}^2(\xi_1^2 + \ldots \xi_d^2)\mathcal{F}f = -|\xi|_2^2\mathcal{F}f$.

(c) The mappings $f \mapsto fg$, $f \mapsto x^\alpha f$ and $f \mapsto \partial^\alpha f$ are continuous from $\mathcal{S}_d$ to $\mathcal{S}_d$.

(d) The Fourier transform is continuous from $\mathcal{S}_d$ to $\mathcal{S}_d$.

PROOF. Let $\xi, x \in \mathbb{R}^d$, $f, g \in \mathcal{S}_d$, $\alpha \in \mathbb{N}_0^d$ and $m \in \mathbb{N}_0$.

We show (a) for $\alpha = e_k$, the assertion then follows by induction. We have

$$\frac{\partial}{\partial \xi_k}\mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x) = -\mathrm{i}x_k \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x) =: \varphi_k(\xi,x),$$

and $\mathbb{R}^d \ni x \mapsto |\varphi_k(\xi,x)| = |x_k f(x)|$ is integrable by Remark E.6. A corollary of the theorem of dominated convergence thus shows that there exists

$$\frac{\partial}{\partial \xi_k}\widehat{f}(\xi) = (2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d} -\mathrm{i}x_k \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x = -\mathrm{i}\mathcal{F}(x_k f)(\xi).$$

For the second part of (a) we write $[-n,n]^k = C_n^k$ and $x = (x', x_k)$ with $x' = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_d) \in \mathbb{R}^{d-1}$. Using that $\partial_k f \in L^1(\mathbb{R}^d)$ and integrating by parts in x_k, we compute

$$\begin{aligned}\mathcal{F}(\partial_k f)(\xi) &= (2\pi)^{-\frac{d}{2}} \lim_{n\to\infty} \int_{C_n^{d-1}} \int_{-n}^{n} \mathrm{e}^{-\mathrm{i}\xi\cdot(x',x_k)} \partial_k f(x',x_k)\,\mathrm{d}x_k\,\mathrm{d}x' \\ &= (2\pi)^{-\frac{d}{2}} \lim_{n\to\infty} \Big[\int_{C_n^d} \mathrm{i}\xi_k \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x + \int_{C_n^{d-1}} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x',x_k)\Big|_{x_k=-n}^{n}\,\mathrm{d}x'\Big] \\ &= \mathrm{i}\xi_k \widehat{f}(\xi).\end{aligned}$$

Here the second integral J_n in the second line tends to 0 as $n\to\infty$ since

$$|J_n| \le 2\int_{C_n^{d-1}} |(x',n)|_2^{-2d}\, |(x',n)|_2^{2d} |f(x',n)|\,\mathrm{d}x' \le 2^d n^{d-1} n^{-2d} p_{0,d}(f).$$

Assertion (b) is a consequence of (a). In (c) and (d) we do not show the continuity of the maps since this is not needed below, see Theorem 7.4 in [**Rud91**]. Thanks to Leibniz' rule, the function $|x|_2^{2m}\partial^\alpha(fg)$ is a linear combination of terms $|x|_2^{2m}(\partial^\beta f)\partial^\gamma g$ which are all bounded since $f,g\in\mathcal{S}_d$. Thus, $fg\in\mathcal{S}_d$. Similarly, one sees that $x^\alpha f, \partial^\alpha f\in\mathcal{S}_d$. Hence, assertion (c) holds. Using (a) and (b), we further compute

$$|\xi|_2^{2m}\partial^\alpha\widehat{f} = (-\mathrm{i})^{|\alpha|}(\xi_1^2+\ldots+\xi_d^2)^m\mathcal{F}(x^\alpha f) = (-\mathrm{i})^{|\alpha|}(-\mathrm{i})^{2m}\mathcal{F}(\Delta^m(x^\alpha f)).$$

Due to (c), the function $\Delta^m(x^\alpha f)$ belongs to $\mathcal{S}_d\subseteq L^1(\mathbb{R}^d)$ so that its Fourier transform is bounded. As a result, $\widehat{f}$ is contained in $\mathcal{S}_d$. □

Corollary E.8 (Riemann-Lebesgue lemma). *If $f\in L^1(\mathbb{R}^d)$, then $\widehat{f}\in C_0(\mathbb{R}^d)$. Hence, $\mathcal{F}\in\mathcal{B}(L^1(\mathbb{R}^d), C_0(\mathbb{R}^d))$.*

Proof. Let $f\in L^1(\mathbb{R}^d)$. Due to Remark E.6, there are $f_n\in\mathcal{S}_d$ converging to f in $L^1(\mathbb{R}^d)$. Lemma E.7 yields $\widehat{f_n}\in\mathcal{S}_d\subseteq C_0(\mathbb{R}^d)$. By (E.2) the functions $\widehat{f_n}$ converge to $\widehat{f}$ in supnorm, so that $\widehat{f}\in C_0(\mathbb{R}^d)$. The second assertion then follows from (E.2). □

For $f\in L^p(\mathbb{R}^d)$ and $1\le p\le\infty$, we define the *reflection operator* R by $(Rf)(x)=f(-x)$, $x\in\mathbb{R}^d$. Clearly, $R^2=I$ and R is an isometric isomorphism on $L^p(\mathbb{R}^d)$. The next lemma is the crucial step towards the main results of this appendix.

Lemma E.9. *The following assertions hold.*

(a) $\int_{\mathbb{R}^d}\widehat{f}g\,\mathrm{d}x = \int_{\mathbb{R}^d} f\widehat{g}\,\mathrm{d}x$ for all $f,g\in\mathcal{S}_d$.
(b) $\mathcal{F}^2=R$, i.e., $(\mathcal{F}\mathcal{F}f)(x)=f(-x)$ for all $f\in\mathcal{S}_d$ and $x\in\mathbb{R}^d$.

Proof. Let $f,g\in\mathcal{S}_d$. Since the map $(x,y)\mapsto\mathrm{e}^{-\mathrm{i}y\cdot x}f(x)g(y)$ is integrable on $\mathbb{R}^{2d}$, Fubini's theorem yields

$$\begin{aligned}\int_{\mathbb{R}^d}\widehat{f}(y)g(y)\,\mathrm{d}y &= \int_{\mathbb{R}^d}(2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d}\mathrm{e}^{-\mathrm{i}y\cdot x}f(x)g(y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^d} f(x)(2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d}\mathrm{e}^{-\mathrm{i}y\cdot x}g(y)\,\mathrm{d}y\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\widehat{g}(x)\,\mathrm{d}x.\end{aligned}$$

In the second assertion one is led to the integrand $\mathrm{e}^{-\mathrm{i}\xi\cdot y}\mathrm{e}^{-\mathrm{i}y\cdot x}f(x)$ which is not integrable for $(x,y)\in\mathbb{R}^{2d}$. So Fubini's theorem does not apply directly, and one has to use a regularization. To that purpose, fix $\xi\in\mathbb{R}^d$ and let $a>0$.

Set $h = \mathrm{e}_{-\xi} D_a \gamma \in \mathcal{S}_d$, i.e., $h(y) = \mathrm{e}^{-\mathrm{i}\xi\cdot y} \exp(-\frac{a^2}{2}|y|_2^2)$ for $y \in \mathbb{R}^d$. Due to the theorem of dominated convergence with the majorant $|\widehat{f}|$, the integral

$$J_a := \int_{\mathbb{R}^d} \widehat{f}(x) h(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \widehat{f}(y) \mathrm{e}^{-\mathrm{i}\xi\cdot y} \gamma(ay)\,\mathrm{d}y$$

converges to $(2\pi)^{\frac{d}{2}} (\mathcal{F}\widehat{f})(\xi)$ as $a \to 0$. On the other hand, part (a), Proposition E.3 and Example E.2 imply that

$$\begin{aligned} J_a &= \int_{\mathbb{R}^d} f(x) \mathcal{F}(\mathrm{e}_{-\xi} D_a \gamma)(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x) a^{-d} (T_\xi D_{1/a} \gamma)(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} f(x) a^{-d} \gamma(\tfrac{1}{a}(x+\xi))\,\mathrm{d}x = \int_{\mathbb{R}^d} f(az-\xi)\gamma(z)\,\mathrm{d}z, \end{aligned}$$

where we also use the substitution $z = \frac{1}{a}(x+\xi)$. We can now apply the theorem of dominated convergence with the majorant $\|f\|_\infty \gamma$ to conclude that $J_a \to f(-\xi)\|\gamma\|_1 = (2\pi)^{\frac{d}{2}} f(-\xi)$ as $a \to 0$, which shows assertion (b). □

PROPOSITION E.10. *The Fourier transform $\mathcal{F} : \mathcal{S}_d \to \mathcal{S}_d$ is bijective with $\mathcal{F}^4 = I$. For all $f, g \in \mathcal{S}_d$ and $x \in \mathbb{R}^d$, we have*

$$\mathcal{F}^{-1} g(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}x\cdot\xi} g(\xi)\,\mathrm{d}\xi, \tag{E.6}$$

$$(\mathcal{F}f|\mathcal{F}g)_{L^2} = (f|g)_{L^2}, \tag{E.7}$$

and $f * g \in \mathcal{S}_d$.

PROOF. Lemma E.9 shows that $I = R^2 = \mathcal{F}^4 = \mathcal{F}\mathcal{F}^3 = \mathcal{F}^3\mathcal{F}$ on $\mathcal{S}_d$ so that $\mathcal{F}$ has the inverse $\mathcal{F}^3 = R\mathcal{F}$ on $\mathcal{S}_d$. This fact already gives (E.6). Let $f, g \in \mathcal{S}_d$ and $x \in \mathbb{R}^d$. Equation (E.6) further yields

$$\begin{aligned} \mathcal{F}(\overline{\mathcal{F}g})(x) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}x\cdot\xi}\, \overline{\widehat{g}(\xi)}\,\mathrm{d}\xi = (2\pi)^{-\frac{d}{2}} \overline{\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}x\cdot\xi} \widehat{g}(\xi)\,\mathrm{d}\xi} \\ &= \overline{(\mathcal{F}^{-1}\mathcal{F}g)(x)} = \overline{g(x)}. \end{aligned}$$

So we can deduce from Lemma E.9 (a) that

$$(\mathcal{F}f|\mathcal{F}g)_{L^2} = \int_{\mathbb{R}^d} \widehat{f}\overline{\widehat{g}}\,\mathrm{d}\xi = \int_{\mathbb{R}^d} f \mathcal{F}(\overline{\mathcal{F}g})\,\mathrm{d}x = \int_{\mathbb{R}^d} f\overline{g}\,\mathrm{d}x = (f|g)_{L^2}.$$

Finally, Proposition E.3 and Lemma E.7 imply that $\mathcal{F}(f*g) = (2\pi)^{\frac{d}{2}} \widehat{f}\widehat{g} =: \varphi$ belongs to $\mathcal{S}_d$. Hence, $f * g = \mathcal{F}^{-1}\varphi \in \mathcal{S}_d$. □

The equality (E.7) shows that $\|\mathcal{F}f\|_2 = \|f\|_2$ for all $f \in \mathcal{S}_d$. Since $\mathcal{S}_d$ is dense in $L^2(\mathbb{R}^d)$ by Remark E.6, we can extend $\mathcal{F}$ to an isometry $\mathcal{F}_2 : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ which is also called *Fourier transform*. Let $f \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. By Theorem D.13 in Appendix D (and it's poof) there are functions $f_n \in C_c^\infty(\mathbb{R}^d) \subseteq \mathcal{S}_d$ which converge to f both in $L^2(\mathbb{R}^d)$ and in $L^1(\mathbb{R}^d)$. Since $\mathcal{F}f_n \to \mathcal{F}_2 f$ in $L^2(\mathbb{R}^d)$, there is a subsequence $\mathcal{F}f_{n_j}$ converging to $\mathcal{F}_2 f$ a.e. as $j \to \infty$ due to the theorem of Riesz-Fischer. On the other hand, $\mathcal{F}f_{n_j}$ converges uniformly to $\mathcal{F}f$ by (E.2). Thus, $\mathcal{F}_2 f = \mathcal{F}f$ a.e.. We now write $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ instead of $\mathcal{F}_2$, and also $\mathcal{F}_2 f = \widehat{f}$.
Warning: $\mathcal{F}f$ is **not** given by the formula (E.1) if $f \in L^2(\mathbb{R}^d) \setminus L^1(\mathbb{R}^d)$.

THEOREM E.11. *The Fourier transform on $\mathcal{S}_d$ extends to a unitary operator $\mathcal{F}: L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ which is given by (E.1) on $L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. Let $f, g \in L^2(\mathbb{R}^d)$, $h \in L^1(\mathbb{R}^d)$, $t \in \mathbb{R}^d$ and $a > 0$. We then have $\mathcal{F}^2 = R$, $\mathcal{F}^4 = I$, $\mathcal{F}^{-1} = R\mathcal{F}$, and*

(a) $(\mathcal{F}f|\mathcal{F}g)_{L^2} = (f|g)_{L^2}$ *(Plancherel),*
(b) $\int_{\mathbb{R}^d} \widehat{f} g \,\mathrm{d}x = \int_{\mathbb{R}^d} f \widehat{g} \,\mathrm{d}x$,
(c) $f(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}x\cdot\xi} \widehat{f}(\xi)\,\mathrm{d}\xi$ *for a.e. $x \in \mathbb{R}^d$ if $\widehat{f} \in L^1(\mathbb{R}^d)$,*
(d) $\mathcal{F}(T_t f) = \mathrm{e}_t \widehat{f}$, $\mathcal{F}(\mathrm{e}_t f) = T_{-t}\widehat{f}$, $\mathcal{F}(D_a f) = a^{-d} D_{1/a} \widehat{f}$,
(e) $\mathcal{F}(h * f) = (2\pi)^{\frac{d}{2}} \widehat{h}\widehat{f}$, $\mathcal{F}^{-1}(\widehat{h}\widehat{f}) = (2\pi)^{-\frac{d}{2}} h * f$,

PROOF. Recall that $\mathcal{F}$ is an isometry on $L^2(\mathbb{R}^d)$. The equations $\mathcal{F}^2 = R$, $\mathcal{F}^4 = I$, and those in (d) hold on the dense subspace $\mathcal{S}_d$ as shown in Proposition E.3, Lemma E.9 and Proposition E.10. Since the maps $\mathcal{F}$, R, T_t, D_a and $f \mapsto \mathrm{e}_t f$ are continuous on $L^2(\mathbb{R}^d)$, these identities can be extended to $L^2(\mathbb{R}^d)$ by approximation. Similarly, (a) and (b) hold since the real and complex scalar products are continuous on $L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. The equation $\mathcal{F}^4 = I$ yields $I = \mathcal{F}\mathcal{F}^3 = \mathcal{F}^3\mathcal{F}$ so that $\mathcal{F}$ has the inverse $\mathcal{F}^{-1} = \mathcal{F}^3 = R\mathcal{F}$. Proposition C.7 in Appendix C says that a bijective isometry on a Hilbert space is unitary. Hence, $\mathcal{F}$ is unitary. From $\mathcal{F}^{-1} = R\mathcal{F}$ we infer assertion (c). For (e), note that $f \mapsto h * f$ is continuous on $L^2(\mathbb{R}^d)$ by Young's inequality (E.4), using $h \in L^1(\mathbb{R}^d)$. The first part of (e) can thus be deduced from Proposition E.3 by approximation. Since $\widehat{h}$ is bounded by (E.2), the function $\widehat{h}\widehat{f}$ belongs to $L^2(\mathbb{R}^d)$. The second part of (e) then follows from the first one by applying $\mathcal{F}^{-1}$. $\square$

REMARK E.12. Let $f \in L^2(\mathbb{R}^d)$. Set $f_n = \mathbb{1}_{B(0,n)} f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ for $n \in \mathbb{N}$. Since $|f_n| \le |f| \in L^2(\mathbb{R}^d)$, we have $f_n \to f$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$ by the theorem of dominated convergence. Hence, $\widehat{f_n} \to \widehat{f}$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$, where

$$\widehat{f_n}(\xi) = (2\pi)^{-\frac{d}{2}} \int_{B(0,n)} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x, \quad \xi \in \mathbb{R}^d,$$

is the *truncated Fourier transform.* $\diamondsuit$

EXAMPLE E.13. We consider the diffusion equation

$$\begin{aligned} \partial_t u(t,x) &= \Delta u(t,x), \quad t \ge 0,\ x \in \mathbb{R}^d, \\ u(0,x) &= u_0(x), \qquad x \in \mathbb{R}^d, \end{aligned} \tag{E.8}$$

for a given initial value $u_0 \in \mathcal{S}_d$. Let us assume for moment that we have a function $u : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{C}$ such that the function $u(t) : x \mapsto u(t,x)$ belongs to $\mathcal{S}_d$ for each $t \ge 0$, $u \in C^1(\mathbb{R}_+, L^2(\mathbb{R}^d))$ and u satisfies (E.8). We set $\widehat{u}(t,\xi) = (\mathcal{F}u(t))(\xi)$ for all $t \ge 0$ and $\xi \in \mathbb{R}^d$. Since $\mathcal{F}$ is continuous on $L^2(\mathbb{R}^d)$, we have

$$\mathcal{F}\partial_t u(t) = \mathcal{F} \lim_{h\to 0} \tfrac{1}{h}(u(t+h) - u(t)) = \lim_{h\to 0} \tfrac{1}{h}(\widehat{u}(t+h) - \widehat{u}(t))$$

for all $t \ge 0$, so that $\widehat{u} \in C^1(\mathbb{R}_+, L^2(\mathbb{R}^d))$ and $\partial_t \widehat{u} = \mathcal{F}\partial_t u$. Applying $\mathcal{F}$ to (E.8), we then deduce from Lemma E.7 that

$$\partial_t \widehat{u}(t) = \mathcal{F}\partial_t u(t) = \mathcal{F}\Delta u(t) = -|\xi|_2^2\, \widehat{u}(t).$$

For each fixed $\xi \in \mathbb{R}^d$ we thus arrive at the ordinary differential equation

$$\begin{aligned} \partial_t \widehat{u}(t,\xi) &= -|\xi|_2^2\, \widehat{u}(t,\xi), && t \geq 0,\ \xi \in \mathbb{R}^d, \\ \widehat{u}(0,\xi) &= \widehat{u}_0(\xi), && \xi \in \mathbb{R}^d, \end{aligned}$$

which has the solution

$$\widehat{u}(t,\xi) = \mathrm{e}^{-t|\xi|_2^2}\, \widehat{u}_0(\xi) = (D_{\sqrt{2t}}\gamma)(\xi)\widehat{u}_0(\xi).$$

Using the properties of $\mathcal{F}$, we compute

$$\begin{aligned} u(t) &= \mathcal{F}^{-1}[(\mathcal{F}\mathcal{F}^{-1}D_{\sqrt{2t}}\gamma)\widehat{u}_0] = (2\pi)^{-\frac{d}{2}}\,(\mathcal{F}^{-1}D_{\sqrt{2t}}\gamma) * u_0 \\ &= (4\pi t)^{-\frac{d}{2}}(D_{1/\sqrt{2t}}\gamma) * u_0 \end{aligned}$$

for all $t > 0$. We define $\gamma_t = (4\pi t)^{-\frac{d}{2}}(D_{1/\sqrt{2t}}\gamma)$. This gives

$$u(t,x) = (\gamma_t * u_0)(x) = \int_{\mathbb{R}^d} (4\pi t)^{-\frac{d}{2}}\mathrm{e}^{-\frac{|x-y|_2^2}{4t}} u_0(x)\,\mathrm{d}x$$

for all $t > 0$ and $x \in \mathbb{R}^d$. Define u by this equation. Since $\gamma_t \in \mathcal{S}_d$, we can reverse the above arguments (or check directly) that this function u solves (E.8) and has the asserted regularity. Moreover, $\|u(t)\| = \|\widehat{u}(t)\|_2 \leq \|\widehat{u}_0\|_2 = \|u_0\|_2$ since the Fourier transform is an isometry and $|D_{\sqrt{2t}}\gamma| \leq 1$. $\diamondsuit$

In the above example we have restricted ourselves to initial values u_0 in $\mathcal{S}_d$. We now want to solve the diffusion equation on $\mathbb{R}^d$ for $u_0 \in L^2(\mathbb{R}^d)$ still using the Fourier transform. To that purpose, we extend the differentiation formulas in Lemma E.7 (a) to the Sobolev spaces and describe the space $W_2^k(\mathbb{R}^d)$ via the Fourier transform in a very convenient way. Recall that $\|u\|_2 = \|\widehat{u}\|_2$ for $u \in L^2(\mathbb{R}^d)$ by Theorem E.11.

Theorem E.14. *Let $k \in \mathbb{N}$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$. We then have*

$$W_2^k(\mathbb{R}^d) = \Big\{ u \in L^2(\mathbb{R}^d) \;\big|\; |\xi|_2^k \widehat{u} \in L^2(\mathbb{R}^d) \Big\} =: H^k$$

and the norm of $W_2^k(\mathbb{R}^d)$ is equivalent to $(\|u\|_2^2 + \||\xi|_2^k \widehat{u}\|_2^2)^{\frac{1}{2}}$. For $u \in W_2^k(\mathbb{R}^d)$ it further holds

$$\mathcal{F}(\partial^\alpha u) = \mathrm{i}^{|\alpha|}\xi^\alpha \widehat{u}. \tag{E.9}$$

Proof. Due to Lemma E.7, the equations (E.9) and $\partial^\alpha \widehat{u} = (-\mathrm{i})^{|\alpha|}\,\mathcal{F}(x^\alpha u)$ holds for $u \in \mathcal{S}_d$. We thus obtain

$$\|u\|_{k,2}^2 = \sum_{|\alpha|\leq k} \|\partial^\alpha u\|_2^2 = \sum_{|\alpha|\leq k} \|\mathcal{F}\partial^\alpha u\|_2^2 = \sum_{|\alpha|\leq k} \|\xi^\alpha \widehat{u}\|_2^2 \tag{E.10}$$

$$\begin{cases} \leq c_1\,(\|u\|_2^2 + \||\xi|_2^k\,\widehat{u}\|_2^2), \\ \geq c_2\,(\|u\|_2^2 + \||\xi|_2^k\,\widehat{u}\|_2^2) \end{cases}$$

for $u \in \mathcal{S}_d$ and constants $c_j > 0$ using (E.9) and (E.5).

Now, let $u \in W_2^k(\mathbb{R}^d)$. Theorem D.13 in Appendix D gives $u_n \in C_c^\infty(\mathbb{R}^d) \subseteq \mathcal{S}_d$ which converge to u in $W_2^k(\mathbb{R}^d)$ as $n \to \infty$. Since $\mathcal{F}$ is continuous on $L^2(\mathbb{R}^d)$, the functions $\widehat{u}_n$ tend to $\widehat{u}$ in $L^2(\mathbb{R}^d)$ and (possibly after passing to a subsequence) pointwise a.e., as $n \to \infty$. Hence, the functions $\xi^\alpha \widehat{u}_n$ converge pointwise a.e. to $\xi^\alpha \widehat{u}$. On the other hand, equation (E.10) yields that $(\xi^\alpha \widehat{u}_n)_n$

is Cauchy in $L^2(\mathbb{R}^d)$ for each $|\alpha| \le k$, and thus $\xi^\alpha \widehat{u}_n$ converge to $\xi^\alpha \widehat{u}$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$. This fact implies that (E.10) also holds for every $u \in W_2^k(\mathbb{R}^d)$ and that $W_2^k(\mathbb{R}^d) \subseteq H^k$.

Conversely, take $u \in H^k$. Let $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $|\alpha| \le k$. From Theorem E.11 and Lemma E.7 we deduce

$$\begin{aligned}\int_{\mathbb{R}^d} u \partial^\alpha \varphi \,\mathrm{d}x &= (u|\partial^\alpha \overline{\varphi}) = (\mathcal{F}u|\mathcal{F}\partial^\alpha \overline{\varphi}) = (\widehat{u}\,|\,\mathrm{i}^{|\alpha|}\,\xi^\alpha\,\widehat{\overline{\varphi}}) = ((-\mathrm{i})^{|\alpha|}\,\xi^\alpha\,\widehat{u}|\mathcal{F}\overline{\varphi}) \\ &= (\mathcal{F}'((-\mathrm{i})^{|\alpha|}\,\xi^\alpha\,\widehat{u})|\overline{\varphi}) = (-1)^{|\alpha|}\int_{\mathbb{R}^d} \varphi\,\mathcal{F}^{-1}(\mathrm{i}^{|\alpha|}\,\xi^\alpha\,\widehat{u})\,\mathrm{d}x.\end{aligned}$$

Therefore u has the weak derivative $\partial^\alpha u = \mathcal{F}^{-1}(\mathrm{i}^{|\alpha|}\,\xi^\alpha\,\widehat{u}) \in L^2(\mathbb{R}^d)$, and hence $u \in W_2^k(\mathbb{R}^d)$, i.e., $W_2^k(\mathbb{R}^d) = H^k$. Applying $\mathcal{F}$ to the equation in the previous sentence, we also derive (E.9) for all $u \in W_2^k(\mathbb{R}^d)$. □

Example E.15. We consider the diffusion equation

$$\begin{aligned}\partial_t u(t,x) &= \Delta u(t,x), \quad t \ge 0,\ x \in \mathbb{R}^d, \\ u(0,x) &= u_0(x), \qquad x \in \mathbb{R}^d,\end{aligned} \tag{E.11}$$

for a given initial value $u_0 \in L^2(\mathbb{R}^d)$. We look for a solution $u \in C([0,\infty), L^2(\mathbb{R}^d)) \cap C^1((0,\infty), L^2(\mathbb{R}^d))$ such that $u(t) \in W_2^2(\mathbb{R}^d)$ for all $t > 0$ and (E.11) holds as equations in $L^2(\mathbb{R}^d)$ for every $t \ge 0$. To that purpose, we define

$$u(t) = \mathcal{F}^{-1}(m_t \widehat{u}_0), \quad \text{where} \quad m_t(\xi) = \mathrm{e}^{-t|\xi|_2^2}$$

for $t > 0$ and $\xi \in \mathbb{R}^d$ as in Example E.13. Since $|\xi|_2^k\,\widehat{u}(t) = |\xi|_2^k\,m_t \widehat{u}_0$ belongs to $L^2(\mathbb{R}^d)$, Theorem E.14 implies that $u(t) \in W_2^k(\mathbb{R}^d)$ for all $k \in \mathbb{N}$ and $t > 0$. From (E.9) we then infer

$$\mathcal{F}\Delta u(t) = -|\xi|_2^2\,\widehat{u}(t) = -|\xi|_2^2\,m_t \widehat{u}_0, \qquad \Delta u(t) = -\mathcal{F}^{-1}(|\xi|_2^2\,m_t \widehat{u}_0).$$

Let $v(t) = m_t \widehat{u}_0$ for $t > 0$. Then $\frac{1}{h}(v(t+h) - v(t))$ converges pointwise to $-|\xi|_2^2\,m_t \widehat{u}_0$ as $h \to 0$ and $|\frac{1}{h}(v(t+h) - v(t))| \le |\xi|_2^2\,m_t|\widehat{u}_0| \in L^2(\mathbb{R}^d)$. Dominated convergence then implies that v has the (continuous) derivative $v'(t) = -|\xi|_2^2\,m_t \widehat{u}_0$ in $L^2(\mathbb{R}^d)$ for $t > 0$. The continuity of $\mathcal{F}^{-1}$ on $L^2(\mathbb{R}^d)$ thus yields that $u \in C^1((0,\infty), L^2(\mathbb{R}^d))$ and

$$u'(t) = -\mathcal{F}^{-1}(|\xi|_2^2\,m_t \widehat{u}_0) = \Delta u(t)$$

for $t > 0$. Finally, $m_t \widehat{u}_0$ tends to $\widehat{u}_0$ in $L^2(\mathbb{R}^d)$ as $t \to 0$ by Lebesgue's theorem with majorant $|\widehat{u}_0|$. Hence, $u \in C([0,\infty), L^2(\mathbb{R}^d))$ with $u(0) = u_0$.

APPENDIX F

The Bochner integral and vector-valued L^p-spaces

In this appendix we introduce the integral of Banach space valued functions, the so-called Bochner integral, define the corresponding Lebesgue and Sobolev spaces and consider the Fourier transform on a Hilbert space valued L^2–space. The construction of the integral is analogous to the scalar-valued case. However, at certain points one has to be more careful with separability issues.

It is assumed that the reader is familiar with the Lebesgue integral and its properties. To prove the corresponding properties of the Bochner integral, we mostly reduce to the scalar situation.

1. The Bochner integral

The presented construction of the Bochner integral follows lecture notes by R. Denk (Konstanz).

If Y is a normed vector space and $M \subseteq Y$, then the *Borel σ-algebra* $\mathcal{B}(M)$ over M is the σ-algebra generated by the system of relatively open subsets of M. We write $\mathcal{B}_d := \mathcal{B}(\mathbb{R}^d)$ for the Borel σ-algebra over $\mathbb{R}^d$. The *d-dimensional Lebesgue measure* is denoted by $\mathrm{d}x$. In one dimension we often write $\mathrm{d}t$. The Lebesgue measure of $A \in \mathcal{B}_d$ is denoted by $|A|$. We define

$$\mathcal{N}_d := \{N \in \mathcal{B}_d \,|\, |N| = 0\}$$

as the set of Borel measurable sets of measure zero. For $A \in \mathcal{B}_d$, a function $f : A \to Y$ is called (Borel-)*measurable* if $f^{-1}(B) \in \mathcal{B}(A)$ for all $B \in \mathcal{B}(Y)$. If $f : A \to Y$ is measurable, then $\|f\|$ is measurable as well, where $\|f\|(x) := \|f(x)\|$ for $x \in A$.

Throughout, let E be a complex Banach space. A function $f : \mathbb{R}^d \to E$ is called *simple* if there are $N \in \mathbb{N}$, $A_n \in \mathcal{B}_d$ and $x_n \in E$ for $n = 1, \dots, N$ with

$$f = \sum_{n=1}^{N} \mathbb{1}_{A_n} x_n.$$

Observe that simple functions are measurable. We start with the integral over simple functions.

Definition F.1. *Let $f = \sum_{n=1}^{N} \mathbb{1}_{A_n} x_n$ be a simple function with $|A_n| < \infty$ for $n = 1, \dots, N$. Then f is called* Bochner integrable *and the* Bochner integral *of f is defined by*

$$\int_{\mathbb{R}^d} f \,\mathrm{d}x = \int_{\mathbb{R}^d} f(x) \,\mathrm{d}x := \sum_{n=1}^{N} |A_n|\, x_n \in E.$$

We note that the above integral is independent of the representation of the simple function f. It is further clear that the Bochner integral is linear on the vector space of simple functions whose support has finite measure. Moreover, as a consequence of the triangle inequality, for each simple function f we have the estimate

$$\Big\| \int_{\mathbb{R}^d} f \, \mathrm{d}x \Big\| \le \int_{\mathbb{R}^d} \|f\| \, \mathrm{d}x, \tag{F.1}$$

where the integral on the right-hand side is now the usual scalar-valued Lebesgue integral.

As in the scalar case, we extend the Bochner integral to a larger class of function by taking limits of simple functions. As it turns out, besides measurability for this procedure a separability condition is necessary.

LEMMA F.2. *For $f : \mathbb{R}^d \to E$ the following assertions are equivalent.*

(a) There is a sequence $(f_k)_{k\in\mathbb{N}}$ of simple functions $f_k : \mathbb{R}^d \to E$ such that $f_k(x) \to f(x)$ as $k \to \infty$ for all $x \in \mathbb{R}^d$.

(b) f is measurable and $f(\mathbb{R}^d) \subseteq E$ is separable.

If one of the assertions is true, then in (a) one can choose $(f_k)_{k\in\mathbb{N}}$ such that $\|f_k(x)\| \le 2\|f(x)\|$ for all $x \in \mathbb{R}^d$.

PROOF. Assume that (a) holds. Then f is measurable as a pointwise limit of measurable functions. Write $f_k = \sum_{n=1}^{N_k} \mathbb{1}_{A_{n,k}} \, x_{n,k}$. Then the rational linear hull of $\{x_{n,k} \,|\, k \in \mathbb{N}, n = 1, \dots, N_k\}$ is countable and dense in $f(\mathbb{R}^d)$. Hence $f(\mathbb{R}^d)$ is separable and (b) follows.

Now suppose that (b) is true. Let $\{y_k \,|\, k \in \mathbb{N}\}$ be a dense subset of $f(\mathbb{R}^d)$. Enlarging this set by an at most countable number of vectors if necessary, we may assume that $y_k \neq 0$ for all k. For $k, N \in \mathbb{N}$ define

$$\tilde{A}_k^N := \{x \in \mathbb{R}^d \,|\, \|f(x)\| \ge 1/N, \;\; \|f(x) - y_k\| < 1/N\}.$$

As an intersection of Borel sets the $\tilde{A}_k^N$ are Borel sets as well. Fixing N, we obtain a disjoint decomposition $(A_k^N)_{k\in\mathbb{N}}$ of $\bigcup_{k\in\mathbb{N}} \tilde{A}_k^N$ by setting

$$A_1^N := \tilde{A}_1^N, \qquad A_k^N := \tilde{A}_k^N \setminus \bigcup_{j=1}^{k-1} A_j^N, \qquad k \in \mathbb{N}.$$

Indeed, we have $x \in A_{k_0}^N$ for some $x \in \bigcup_{k\in\mathbb{N}} \tilde{A}_k^N$ if and only if k_0 is the smallest number such that $x \in \tilde{A}_{k_0}^N$. Since $\{y_k \,|\, k \in \mathbb{N}\}$ is dense in $f(\mathbb{R}^d)$, for each N we have

$$\bigcup_{k\in\mathbb{N}} A_k^N = \bigcup_{k\in\mathbb{N}} \tilde{A}_k^N = \{x \in \mathbb{R}^d \,|\, \|f(x)\| \ge 1/N\}. \tag{F.2}$$

Now we can define the desired sequence $(f_N)_{N\in\mathbb{N}}$ of simple functions. For $N \in \mathbb{N}$ we set

$$f_N(x) := 0 \quad \text{if } x \notin \bigcup_{M,k=1}^{N} A_k^M, \qquad f_N(x) := y_{k_{N,x}} \quad \text{if } x \in \bigcup_{M,k=1}^{N} A_k^M,$$

where the number $k_{N,x} \in \{1, \dots, N\}$ is determined as follows. Take the largest integer $M_{N,x} \le N$ such that $x \in \bigcup_{k=1}^{N} A_k^{M_{N,x}}$. Employing the disjointness of this union, $k_{N,x}$ is defined as the unique number such that $x \in A_{k_{N,x}}^{M_{N,x}}$. Note that this implies $k_{N,x} \leq N$. Hence f_N takes only values in $\{0, y_1, \ldots, y_N\}$. Moreover, if $x \in A_{k_{N,x}}^{M_{N,x}}$ then

$$\|f_N(x)\| = \|y_{k_{N,x}}\| \leq \|y_{k_{N,x}} - f(x)\| + \|f(x)\| \leq 1/M_{N,x} + \|f(x)\| \leq 2\|f(x)\|,$$

so that $\|f_N\| \leq 2\|f\|$.

Let us check that f_N is measurable. First we have $f_N^{-1}(\{0\}) = \mathbb{R}^d \setminus \bigcup_{M,k=1}^N A_k^M \in \mathcal{B}_d$. Next, let $k_* \in \{1, \ldots, N\}$. Then $f_N(x) = y_{k_*}$ if either $x \in A_{k_*}^N$ or if, for some $l \in \{1, \ldots, N-1\}$, we have that x belongs to $A_{k_*}^{N-l}$ but not to $\bigcup_{k=1}^N A_k^{N-l+m}$ for all $m \in \{1, \ldots, l\}$. In other words,

$$f_N^{-1}(\{y_{k_*}\}) = A_{k_*}^N \cup \bigcup_{l=1}^{N-1} \Big(A_{k_*}^{N-l} \setminus \bigcup_{m=1}^{l} \bigcup_{k=1}^{N} A_k^{N-l+m} \Big) \in \mathcal{B}_d.$$

Therefore, each f_N is a simple function.

We show that $f_N \to f$ pointwise as $N \to \infty$. Let $x \in \mathbb{R}^d$. First suppose that $f(x) \neq 0$. Given $\varepsilon > 0$, we choose a natural number N_0 such that $\frac{1}{N_0} < \min\{\|f(x)\|, \varepsilon\}$. Using (F.2) and that the sets $A_k^{N_0}$, $k \in \mathbb{N}$, are disjoint, we find a unique k_0 such that $x \in A_{k_0}^{N_0}$. Now consider an arbitrary number $N \geq \max\{N_0, k_0\}$. Then we have $x \in \bigcup_{M,k=1}^N A_k^M$, since $A_{k_0}^{N_0}$ appears in this union. Further $N_0 \leq M_{N,x} \leq N$, since $M_{N,x}$ is defined as the largest number smaller than N such that $x \in \bigcup_{k=1}^N A_k^{M_{N,x}}$ and N_0 has this property. Since $x \in A_{k_{N,x}}^{M_{N,x}}$, it follows that

$$\|f(x) - f_N(x)\| = \|f(x) - y_{k_{N,x}}^{M_{N,x}}\| \leq 1/M_{N,x} \leq 1/N_0 \leq \varepsilon.$$

Finally, for $x \in \mathbb{R}^d$ with $f(x) = 0$ we have $x \notin A_k^N$ for all k and N. Therefore $f_N(x) = 0$ and $\|f(x) - f_N(x)\| \leq \varepsilon$ is trivially satisfied. We thus conclude the pointwise convergence of f_N as $N \to \infty$, and (a) follows. ☐

The lemma suggest the following notion.

DEFINITION F.3. *A map $f : \mathbb{R}^d \to E$ is called* strongly measurable *if there is a sequence $(f_k)_{k\in\mathbb{N}}$ of simple functions $f_k : \mathbb{R}^d \to E$ such that $f_k(x) \to f(x)$ as $k \to \infty$ for all $x \in \mathbb{R}^d$.*

Another characterization of strong measurability involving continuous functionals is given in Lemma F.9 below.

For a strongly measurable f one would like to define the Bochner integral as a limit of Bochner integrals of simple functions. Fortunately, there is a simple criterion when this is possible.

LEMMA F.4. *Let $f : \mathbb{R}^d \to E$ be strongly measurable. Then the following assertions are equivalent.*

(a) There is a sequence of simple integrable functions $(f_k)_{k\in\mathbb{N}}$ such that $f_k(x) \to f(x)$ as $k \to \infty$ for each $x \in \mathbb{R}^d$ and

$$\lim_{k\to\infty} \int_{\mathbb{R}^d} \|f_k - f\| \,\mathrm{d}x = 0.$$

(b) It holds that $\int_{\mathbb{R}^d} \|f(x)\| \,\mathrm{d}x < \infty$.

If one of the assertions is true, then the limit $\lim_{k\to\infty} \int_{\mathbb{R}^d} f_k \,\mathrm{d}x$ *exists in* E *and is independent of the sequence of simple functions* $(f_k)_{k\in\mathbb{N}}$ *as in (a).*

PROOF. Assume that (a) is true. First note that $\|f_k - f\|, \|f\| : \mathbb{R}^d \to \mathbb{R}_+$ are measurable as compositions of measurable maps. Hence $\int_{\mathbb{R}^d} \|f_k - f\| \,\mathrm{d}x$ and $\int_{\mathbb{R}^d} \|f\| \,\mathrm{d}x$ are well-defined numbers, the second one possibly equal to ∞. But it indeed is finite since

$$\int_{\mathbb{R}^d} \|f\| \,\mathrm{d}x \leq \int_{\mathbb{R}^d} \|f - f_k\| \,\mathrm{d}x + \int_{\mathbb{R}^d} \|f_k\| \,\mathrm{d}x < \infty$$

for all k such that e.g. $\int_{\mathbb{R}^d} \|f - f_k\| \,\mathrm{d}x \leq 1$.

Conversely, suppose that $\int_{\mathbb{R}^d} \|f\| \,\mathrm{d}x < \infty$. Lemma F.2 gives a sequence of simple functions $(f_k)_{k\in\mathbb{N}}$ converging pointwise to f and $\|f_k\| \leq 2\|f\|$. Now $g_k = \|f_k - f\|$ defines a sequence of measurable functions converging pointwise to zero as $k \to \infty$, dominated by the (Lebesgue-)integrable function $3\|f\|$. Thus $\lim_{k\to\infty} \int_{\mathbb{R}^d} g_k \,\mathrm{d}x = 0$ by the (scalar) dominated convergence theorem, which shows (a).

Finally, take a sequence $(f_k)_{k\in\mathbb{N}}$ of simple functions as in (a). Using linearity and (F.1), for $k, l \in \mathbb{N}$ we obtain

$$\begin{aligned}\Big\| \int_{\mathbb{R}^d} f_k \,\mathrm{d}x - \int_{\mathbb{R}^d} f_l \,\mathrm{d}x \Big\| &\leq \int_{\mathbb{R}^d} \|f_k - f_l\| \,\mathrm{d}x \\ &\leq \int_{\mathbb{R}^d} \|f_k - f\| \,\mathrm{d}x + \int_{\mathbb{R}} \|f - f_l\| \,\mathrm{d}x,\end{aligned}$$

which shows that $\big(\int_{\mathbb{R}^d} f_k \,\mathrm{d}x\big)_{k\geq 0}$ is a Cauchy sequence in E. Hence $\int_{\mathbb{R}^d} f_k \,\mathrm{d}x$ converges in E as $k \to \infty$. Let $(g_k)_{k\in\mathbb{N}}$ be another sequence of simple functions as in (a). If one replaces f_l by g_k in the above estimates and takes the limit $k \to \infty$, then one obtains that $\lim_{k\to\infty} \int_{\mathbb{R}^d} f_k \,\mathrm{d}x = \lim_{k\to\infty} \int_{\mathbb{R}^d} g_k \,\mathrm{d}x$. □

Now we can define integrability and the Bochner integral for a large class of functions.

DEFINITION F.5. *A function* $f : \mathbb{R}^d \to E$ *is called* Bochner integrable *if it is strongly measurable and if* $\int_{\mathbb{R}^d} \|f\| \,\mathrm{d}x < \infty$. *In this case one sets*

$$\int_{\mathbb{R}^d} f \,\mathrm{d}x := \lim_{k\to\infty} \int_{\mathbb{R}^d} f_k \,\mathrm{d}x,$$

where $(f_k)_{k\in\mathbb{N}}$ *is any sequence of simple functions as in Lemma F.4 (a). Furthermore, for* $A \in \mathcal{B}_d$ *a function* $f : A \to E$ *is called* Bochner integrable *if its extension* f_0 *by zero to* $\mathbb{R}^d$ *is Bochner integrable, and in this case one defines*

$$\int_A f \mathrm{d}x := \int_{\mathbb{R}^d} f_0 \mathrm{d}x.$$

For $A \in \mathcal{B}_d$ *one finally sets*

$$\mathcal{L}(A, E) := \{ f : A \to E \,|\, f \text{ is integrable} \}.$$

Given $A \in \mathcal{B}_d$, it follows from an approximation argument and the corresponding properties of simple functions that $\mathcal{L}(A, E)$ is a vector space, that the Bochner integral is linear and that

$$\Big\| \int_A f \,\mathrm{d}x \Big\| \leq \int_A \|f\| \,\mathrm{d}x \tag{F.3}$$

for all $f \in \mathcal{L}(A, E)$.

2. Properties of the Bochner integral

We collect some further basic properties of the Bochner integral that are analogous to the Lebesgue integral. Essentially, there are two strategies for the proofs. The first is to reduce to the case of simple functions by approximation, and the second is to reduce to the scalar case.

LEMMA F.6. *Let $A, A_1, A_2 \in \mathcal{B}_d$ such that $A = A_1 \dot\cup A_2$ and suppose that $f \in \mathcal{L}(A, E)$. Then $f|_{A_1} \in \mathcal{L}(A_1, E)$, $f|_{A_2} \in \mathcal{L}(A_2, E)$ and*

$$\int_A f \,\mathrm{d}x = \int_{A_1} f|_{A_1} \,\mathrm{d}x + \int_{A_2} f|_{A_2} \,\mathrm{d}x.$$

Moreover, if $N \in \mathcal{B}_d$ is such that $|N| = 0$, then $\int_N f \,\mathrm{d}x = 0$ for all $f \in \mathcal{L}(N, E)$.

PROOF. Let $f_0, f_1, f_2 \in \mathcal{L}(\mathbb{R}^d, E)$ be the trivial extensions of f, $f|_{A_1}$ and $f|_{A_2}$ to $\mathbb{R}^d$. Take a sequence of measurable simple functions such that $f_k \to f_0$ pointwise as $k \to \infty$. Then $\mathbb{1}_{A_1} f_k$ and $\mathbb{1}_{A_2} f_k$ are sequences of simple functions converging pointwise to $\mathbb{1}_{A_1} f_0 = f_1$ and $\mathbb{1}_{A_2} f_0 = f_2$, respectively. We thus obtain

$$\begin{aligned}\int_{\mathbb{R}^d} f_0 \,\mathrm{d}x = \lim_{k\to\infty} \int_{\mathbb{R}^d} f_k \,\mathrm{d}x &= \lim_{k\to\infty} \Big(\int_{\mathbb{R}^d} \mathbb{1}_{A_1} f_k \,\mathrm{d}x + \int_{\mathbb{R}^d} \mathbb{1}_{A_2} f_k \,\mathrm{d}x \Big) \\ &= \int_{\mathbb{R}^d} f_1 \,\mathrm{d}x + \int_{\mathbb{R}^d} f_2 \,\mathrm{d}x,\end{aligned}$$

from which the first assertion follows. Now let $|N| = 0$ and $f \in \mathcal{L}(N, E)$. Let f_0 be the trivial extension of f to $\mathbb{R}^d$ and let $(f_k)_{k\in\mathbb{N}}$ be simple functions such that $\int_{\mathbb{R}^d} \|f_0 - f_k\| \,\mathrm{d}x \to 0$ as $k \to \infty$. Since $\|f_0 - f_k\| = \|f_k\|$ almost everywhere on $\mathbb{R}^d$, we have that $\int_{\mathbb{R}^d} \|f_k\| \,\mathrm{d}x \to 0$ as $k \to \infty$, and thus

$$\Big\| \int_{\mathbb{R}^d} f_0 \,\mathrm{d}x \Big\| = \lim_{k\to\infty} \Big\| \int_{\mathbb{R}^d} f_k \,\mathrm{d}x \Big\| \leq \lim_{k\to\infty} \int_{\mathbb{R}^d} \|f_k\| \,\mathrm{d}x = 0,$$

where we used (F.1). □

LEMMA F.7. *Let $f \in \mathcal{L}(A, E)$ and let $T \in \mathcal{B}(E, Y)$ for another Banach space Y. Then Tf, defined by $(Tf)(x) := Tf(x)$ for $x \in A$, belongs to $\mathcal{L}(A, Y)$ and we have*

$$T \int_A f \,\mathrm{d}x = \int_A Tf \,\mathrm{d}x.$$

PROOF. Take simple functions f_k converging to f pointwise on $\mathbb{R}^d$ as $k \to \infty$ such that $\lim_{k\to\infty} \int_{\mathbb{R}^d} \|f - f_k\| \,\mathrm{d}x = 0$. Then Tf_k is simple for each k, $Tf_k \to Tf$ pointwise as $k \to \infty$ and

$$\int_{\mathbb{R}^d} \|Tf - Tf_k\| \,\mathrm{d}x \leq \|T\| \int_{\mathbb{R}^d} \|f - f_k\| \,\mathrm{d}x \to 0, \qquad k \to \infty.$$

Moreover, $\int_A Tf_k \, \mathrm{d}x = T \int_A f_k \, \mathrm{d}x$ in view of Definition F.1. The assertions now follow from the Lemmas F.2 and F.4. $\square$

Proposition F.8 (Dominated convergence theorem). *Let $f_k \in \mathcal{L}(A, E)$ for $k \in \mathbb{N}$ and let $f : A \to E$ be strongly measurable such that $f_k \to f$ pointwise almost everywhere on A as $k \to \infty$. Suppose there is $g \in \mathcal{L}(A, \mathbb{C})$ with $\|f_k\| \le g$ for all k. Then $f \in \mathcal{L}(A, E)$ and*

$$\lim_{k\to\infty} \int_A f_k \, \mathrm{d}x = \int_A f \, \mathrm{d}x.$$

Proof. It suffices to consider the case $A = \mathbb{R}^d$. There is $N \in \mathcal{N}_d$ such that $g_k := \|f_k - f\| \to 0$ pointwise on $\mathbb{R}^d \setminus N$ as $k \to \infty$. Since $\|f\| \le g$ on $\mathbb{R}^d \setminus N$, we have $f \in \mathcal{L}(\mathbb{R}^d, E)$ and $\|g_k\| \le 2\|g\|$ on $\mathbb{R}^d \setminus N$. Thus, using (F.3) and the scalar dominated convergence theorem, we get

$$\Big\| \int_{\mathbb{R}^d} f_k \, \mathrm{d}x - \int_{\mathbb{R}^d} f \, \mathrm{d}x \Big\| \le \int_{\mathbb{R}^d} g_k \, \mathrm{d}x \to 0$$

as $k \to \infty$. $\square$

We finish this section with Fubini's theorem for Bochner integrals. To this end we need another characterization of strong measurability.

Lemma F.9 (Pettis). *For $f : \mathbb{R}^d \to E$, the following assertions are equivalent.*

(a) f is strongly measurable.
(b) For each $x^* \in E^*$ the map $\langle f, x^* \rangle : \mathbb{R}^d \to \mathbb{C}$ is measurable, and $f(\mathbb{R}^d) \subseteq E$ is separable.

Proof. If f is strongly measurable, then $\langle f, x^* \rangle$ is measurable for each $x^* \in E^*$ since x^* is continuous. The separability of $f(\mathbb{R}^d)$ holds by definition. Let us prove that (b) implies (a). Define $E_0 := \overline{\operatorname{span} f(\mathbb{R}^d)}$ and let $\{x_k \,|\, k \in \mathbb{N}\}$ be dense in E_0. Given $N \in \mathbb{N}$, define $s_N : E_0 \to \{x_1, \dots, x_N\}$ by $s_N(y) = x_{k_{N,y}}$, where $k_{N,y}$ is the smallest number $1 \le k \le N$ such that

$$\|y - x_k\| = \min_{1\le j\le N} \|y - x_j\|.$$

By density we have $\min_{1\le j\le N} \|y - x_j\| \to 0$ as $N \to \infty$, hence $s_N(y) \to y$ for each $y \in E_0$ as $N \to \infty$. Now define the function $f_N : A \to E$ by

$$f_N(\xi) = s_N(f(\xi)), \qquad \xi \in A.$$

Then $f_N \to f$ poinwise on A as $N \to \infty$. Moreover, for $1 \le k \le N$,

$$\begin{aligned} f_N^{-1}(x_k) = \{\xi \in A \,|\, &\|f(\xi) - x_k\| = \min_{1\le j\le N} \|f(\xi) - x_j\|\} \\ &\cap \{\xi \in A \,|\, \|f(\xi) - x_l\| > \min_{1\le j\le N} \|f(\xi) - x_j\| \text{ for } l = 1, \dots, k-1\}. \end{aligned}$$

To conclude the strong measurability of f, we have to show that the sets on the right-hand side are measurable. To this end we prove that for each $x \in E_0$ the functions $\xi \mapsto \|f(\xi) - x\|$ are measurable. Then (a) follows and we are finished.

We claim that there is a sequence $(x_k^*)_{k\in\mathbb{N}}$ of unit vectors in E^* such that

$$\|y\| = \sup_{k\in\mathbb{N}} |\langle y, x_k^* \rangle|, \qquad y \in E_0.$$

To see this, we note that for each x_k the Hahn-Banach theorem gives a unit vector $x_k^* \in E^*$ such that $\langle x_k, x_k^* \rangle = \|x_k\|$. Given $y \in E_0$ and $\varepsilon > 0$, take x_{k_0} such that $\|y - x_{k_0}\| \leq \varepsilon$. Then

$$\|y\| \leq \varepsilon + |\langle x_{k_0}, x_{k_0}^* \rangle| \leq 2\varepsilon + \sup_{k \in \mathbb{N}} |\langle y, x_k^* \rangle|.$$

Since ε is arbitrary, the claim follows.

Now we can write $\|f(\xi) - x\| = \sup_{k\in\mathbb{N}} |\langle f(\xi) - x, x_k^* \rangle|$ for $\xi \in A$ and $x \in E_0$. By assumption in (b), $\xi \mapsto |\langle f(\xi) - x, x_k^* \rangle|$ is measurable for each k. Hence $\xi \mapsto \|f(\xi) - x\|$ is measurable, which finishes the proof. □

For $d = m + n$ with $m, n \in \mathbb{N}$ we write $x = (y, z) \in \mathbb{R}^d$ with $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$. For $f : \mathbb{R}^d \to E$, $y_* \in \mathbb{R}^m$ and $z_* \in \mathbb{R}^n$ we define the functions $f^{y_*} : \mathbb{R}^n \to E$ and $f^{z_*} : \mathbb{R}^m \to E$ by

$$f^{y_*}(z) := f(y_*, z), \quad z \in \mathbb{R}^n, \qquad f^{z_*} = f(y, z_*), \quad y \in \mathbb{R}^m.$$

Proposition F.10 (Fubini's theorem). *Let $f \in \mathcal{L}(\mathbb{R}^d, E)$. Then there are $M \in \mathcal{N}_m$ and $N \in \mathcal{N}_n$ such that $f^y \in \mathcal{L}(\mathbb{R}^n, E)$ for all $y \in \mathbb{R}^m \setminus M$ and $f^z \in \mathcal{L}(\mathbb{R}^m, E)$ for all $z \in \mathbb{R}^n \setminus N$. Moreover, define the maps $F : \mathbb{R}^m \to E$ and $G : \mathbb{R}^n \to E$ by*

$$F(y) := \int_{\mathbb{R}^n} f^y(z)\,\mathrm{d}z, \quad y \in \mathbb{R}^m \setminus M, \qquad G(z) := \int_{\mathbb{R}^m} f^z(y)\,\mathrm{d}y, \quad z \in \mathbb{R}^n \setminus N,$$

and equal to zero otherwise, respectively. Then $F \in \mathcal{L}(\mathbb{R}^m, E)$ and $G \in \mathcal{L}(\mathbb{R}^n, E)$, and further

$$\int_{\mathbb{R}^d} f\,\mathrm{d}x = \int_{\mathbb{R}^m} F(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} G(z)\,\mathrm{d}z. \tag{F.4}$$

As usual, with abuse of notation, one writes equation (F.4) as

$$\int_{\mathbb{R}^d} f\,\mathrm{d}x = \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^m} f(y,z)\,\mathrm{d}y \Big)\,\mathrm{d}z = \int_{\mathbb{R}^m} \Big(\int_{\mathbb{R}^n} f(y,z)\,\mathrm{d}z \Big)\,\mathrm{d}y.$$

Proof. We show that $f^y \in \mathcal{L}(\mathbb{R}^n, E)$ for all $y \in \mathbb{R}^m \setminus M$, where $M \in \mathcal{N}_m$. Let $x^* \in E^*$ be arbitrary. Then $\langle f, x^* \rangle : \mathbb{R}^d \to \mathbb{C}$ is measurable by Lemma F.9. Thus for each $y \in \mathbb{R}^m$ the scalar function $\langle f^y, x^* \rangle$ is measurable. Since the image of f^y is separable, the strong measurability of f^y follows from Lemma F.9. By Tonelli's theorem, we further have that

$$\int_{\mathbb{R}^m} \Big(\int_{\mathbb{R}^n} \|f(y,z)\|\,\mathrm{d}z \Big)\,\mathrm{d}y = \int_{\mathbb{R}^d} \|f\|\,\mathrm{d}x < \infty.$$

Hence $\int_{\mathbb{R}^n} \|f^y(z)\|\,\mathrm{d}z < \infty$ for almost all y. We conclude $f^y \in \mathcal{L}(\mathbb{R}^n, E)$ for $y \in \mathbb{R}^m \setminus M$ with $M \in \mathcal{N}_m$ from Lemma F.4.

We prove that $F \in \mathcal{L}(\mathbb{R}^m, E)$. Take simple functions f_k converging pointwise to f as $k \to \infty$. For $y \in \mathbb{R}^m \setminus M$, each $\int_{\mathbb{R}^n} f_k^y(z)\,\mathrm{d}z$ is simple as well, and $\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k^y(z)\,\mathrm{d}z = \int_{\mathbb{R}^n} f^y(z)\,\mathrm{d}z$ by dominated convergence. Hence F is strongly measurable. Since

$$\int_{\mathbb{R}^m} \|F\|\,\mathrm{d}y \leq \int_{\mathbb{R}^d} \|f\|\,\mathrm{d}x < \infty$$

by (F.3) and Tonelli's theorem, we obtain $F \in \mathcal{L}(\mathbb{R}^m, E)$. Finally, the equality (F.4) follows from the application of continuous functionals, Lemma F.7, Fubini's theorem in the scalar case and the Hahn-Banach theorem. The assertions for G are shown in the same way. $\square$

3. Vector-valued L^p and Sobolev spaces

The definition of the vector-valued Lebesgue spaces is analogous to the scalar case.

Definition F.11. *Let E be a Banach space and $A \in \mathcal{B}_d$.*

(a) For $p \in [1, \infty)$ we define $\mathcal{L}^p(A, E)$ as the set of all $f : A \to E$ such that $f|_{A\setminus N}$ is strongly measurable for some $N \in \mathcal{N}_d$ and $\int_{A\setminus N} \|f\|^p \, \mathrm{d}x < \infty$. Moreover, for $f \in \mathcal{L}^p(A, E)$ we set

$$\|f\|_{\mathcal{L}^p(A,E)} := \Big(\int_{A\setminus N} \|f\|^p \, \mathrm{d}x \Big)^{1/p}.$$

(b) For $p = \infty$, $\mathcal{L}^\infty(A, E)$ is defined as the set of all $f : A \to E$ such that $f|_{A\setminus N}$ is strongly measurable for some $N \in \mathcal{N}_d$ and

$$\|f\|_{\mathcal{L}^\infty(A,E)} := \inf\{c \in [0, \infty] \,|\, |\{\|f(x)\| > c\}| = 0\} < \infty.$$

We observe that $\|f\|_{\mathcal{L}^p(A,E)}$ is independent of $N \in \mathcal{N}_d$ as above. In the same way as for $E = \mathbb{C}$ one can prove the following facts. Part (c) below is the vector-valued version of the Fischer-Riesz theorem.

Proposition F.12. *For $p \in [1, \infty]$ the following holds true.*

(a) For $f, g \in \mathcal{L}^p(A, E)$ we have Minkowski's inequality

$$\|f + g\|_{\mathcal{L}^p(A,E)} \leq \|f\|_{\mathcal{L}^p(A,E)} + \|g\|_{\mathcal{L}^p(A,E)}.$$

(b) For $f \in \mathcal{L}^p(A, \mathbb{C})$ and $g \in \mathcal{L}^{p'}(A, E)$, where $p' := \frac{p}{p-1}$ for $p \neq 1$ and $p' := \infty$ for $p = 1$, we have Hölder's inequality

$$\|fg\|_{\mathcal{L}^1(A,E)} \leq \|f\|_{\mathcal{L}^p(A,\mathbb{C})} \|g\|_{\mathcal{L}^{p'}(A,E)}.$$

(c) $\mathcal{L}^p(A, E)$ endowed with $\|\cdot\|_{\mathcal{L}^p(A,E)}$ becomes a complete semi-normed vector space under pointwise addition and scalar multiplication.

(d) If $f_k \to f$ in $\mathcal{L}^p(A, E)$, then there is a subsequence $(f_{k_l})_{l\in\mathbb{N}}$ such that $f_{k_l} \to f$ pointwise almost everywhere on A.

In the usual way one obtains from $\mathcal{L}^p(A, E)$ a normed space by considering the factor space

$$L^p(A, E) := \mathcal{L}^p(A, E) / \{ f : A \to E \,|\, f(x) = 0 \text{ for } \mathrm{d}x\text{-almost every } x \in A \},$$

i.e., by identifying functions that are equal outside a set of measure zero. Setting

$$\|[f]\|_{L^p(A,E)} := \|f\|_{\mathcal{L}^p(A,E)}$$

for an equivalence class $[f] \in L^p(A, E)$ and any $f \in [f]$, we get that $\|\cdot\|_{L^p(A,E)}$ is a norm on $L^p(A, E)$. It is clear that $\|[f]\|_{L^p(A,E)}$ is independent of the chosen $f \in [f]$. The completeness of $L^p(A, E)$ carries over from $\mathcal{L}^p(A, E)$, so that $L^p(A, E)$ is a Banach space with respect to $\|\cdot\|_{L^p(A,E)}$. As usual, with a slight abuse of notation, in the sequel we drop the equivalence brackets $[\cdot]$ and just write $f \in L^p(A, E)$.

We finish our considerations on L^p-spaces by observing a density result.

LEMMA F.13. *Let $p \in [1, \infty)$. Then the set of simple functions is dense in $L^p(A, E)$. In particular, for any $p, q \in [1, \infty)$ we have that $L^p(A, E) \cap L^q(A, E)$ is dense in $L^p(A, E)$.*

PROOF. Let $f \in L^p(A, E)$. Since f is measurable and we may assume that its image is separable, it follows from Lemma F.2 that there is a sequence of simple functions $(f_k)_{k\in\mathbb{N}}$ converging pointwise to f which satisfies $\|f_k\| \le 2\|f\|$ for all k. Using the dominated convergence theorem, it follows as in the proof of Lemma F.4 that $f_k \to f$ in $L^p(A, E)$. □

We continue with vector-valued Sobolev spaces over open intervals $(a, b) \subseteq \mathbb{R}$, which is sufficient for our purposes in the course.

DEFINITION F.14. *Let $J = (a, b)$ for $-\infty \le a < b \le +\infty$, let E be a Banach space and let $p \in [1, \infty]$. Then the* Sobolev space $W_p^1(J, E)$ *is defined as the set of all $f \in L^p(J, E)$ such that there is $g \in L^p(J, E)$ satisfying*

$$f(y) - f(x) = \int_x^y g(t)\mathrm{d}t \quad \text{for almost every } x, y \in J.$$

In this case we call g the weak derivative *of f and write $f' := g$. For $f \in W_p^1(J, E)$ we further set*

$$\|f\|_{W_p^1(J,E)} := \begin{cases} \left(\|f\|_{L^p(J,E)}^p + \|f'\|_{L^p(J,E)}^p\right)^{1/p} & \text{if } p \in [1, \infty), \\ \max\{\|f\|_{L^\infty(J,E)}, \|f'\|_{L^\infty(J,E)}\} & \text{if } p = \infty. \end{cases}$$

We observe that a weak derivative is uniquely determined in $L^p(J, E)$, so that f' is well-defined. It is clear that $\|\cdot\|_{W_p^1(J,E)}$ is a norm on $W_p^1(J, E)$. As in the scalar case one shows that $W_p^1(J, E)$ is a Banach space with respect to $\|\cdot\|_{W_p^1(J,E)}$. The weak derivative is a bounded linear map from $W_p^1(J, E)$ to $L^p(J, E)$.

PROPOSITION F.15. *Let $p \in [1, \infty]$. Then each $f \in W_p^1(J, E)$ has a representative belonging to $C(\overline{J}, E)$.*

PROOF. There is $x_0 \in J$ such that $f(y) = g(y) := f(x_0) + \int_{x_0}^y f'(t)\,\mathrm{d}t$ for almost every $y \in J$. This defines a representative $g : \overline{J} \to E$ of f. By the dominated convergence theorem, g is continuous. □

We remark that for $f \in W_p^1(J, E)$ the above result in particular allows to give a meaning to $f(x) \in E$ for all $x \in \overline{J}$.

4. The Fourier transform on a Hilbert space valued L^2-space

Let E be a *Hilbert space* with scalar product $(\cdot\,|\,\cdot)_E$. For $f \in L^1(\mathbb{R}^d, E)$ one defines the Fourier transform as in the scalar case by

$$\widehat{f}(\xi) = (\mathcal{F}f)(\xi) := (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\,\mathrm{d}x, \qquad \xi \in \mathbb{R}^d.$$

We extend the Fourier transform to an isometric isomorphism on $L^2(\mathbb{R}^d, E)$. Note here that $L^2(\mathbb{R}^d, E)$ is a Hilbert space with respect to the scalar product

$$(f|g)_{L^2(\mathbb{R}^d,E)} := \int_{\mathbb{R}^d} (f(x)|g(x))_E \,\mathrm{d}x.$$

THEOREM F.16. *Let E be a Hilbert space. Then $\mathcal{F}$ extends to an isometric isomorphism on $L^2(\mathbb{R}^d, E)$, which is denoted by $\mathcal{F}$ again, with inverse $\mathcal{F}^{-1}$ given by $(\mathcal{F}^{-1}g)(y) = (\mathcal{F}g)(-y)$ for $y \in \mathbb{R}^d$. Moreover, for $f, g \in L^2(\mathbb{R}^d, E)$ we have* Plancherel's formula

$$(\widehat{f}|\widehat{g})_{L^2(\mathbb{R}^d,E)} = (f|g)_{L^2(\mathbb{R}^d,E)}.$$

PROOF. Let $f = \sum_n \mathbb{1}_{A_n} x_n$ and $g = \sum_k \mathbb{1}_{B_k} y_k$ be simple functions. Using Plancherel's theorem for the scalar-valued case, we obtain

$$\begin{aligned}
(\widehat{f}|\widehat{g})_{L^2(\mathbb{R}^d,E)} &= \sum_{n,k} (\widehat{\mathbb{1}}_{A_n} x_n | \widehat{\mathbb{1}}_{B_k} y_k)_{L^2(\mathbb{R}^d,E)} \\
&= \sum_{n,k} (\widehat{\mathbb{1}}_{A_n} | \widehat{\mathbb{1}}_{B_k})_{L^2(\mathbb{R}^d)} (x_n|y_k) \\
&= \sum_{n,k} (\mathbb{1}_{A_n} | \mathbb{1}_{B_k})_{L^2(\mathbb{R}^d)} (x_n|y_k) = (f|g)_{L^2(\mathbb{R}^d,E)}.
\end{aligned}$$

In particular, $\|\widehat{f}\|_{L^2(\mathbb{R}^d,E)} = \|f\|_{L^2(\mathbb{R}^d,E)}$ for simple functions f. By density (see Lemma F.13), $\mathcal{F}$ extends continuously to an isometry on $L^2(\mathbb{R}^d, E)$, and Plancherel's formula continues to hold for the extension. Similarly, using the inversion formula in the scalar case, we obtain that $(\mathcal{F}^2 f)(y) = f(-y)$ for a simple function f and $y \in \mathbb{R}^d$, and this equation continues to hold on $L^2(\mathbb{R}^d, E)$. □

Bibliography

[ABHN11] Wolfgang Arendt, Charles J.K. Batty, Matthias Hieber, and Frank Neubrander. *Vector-valued Laplace Transforms and Cauchy Problems*. Monographs in Mathematics ; 96. Birkhäuser/Springer Basel AG, Basel, second edition, 2011.

[AF03] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[Bec12] Marius Beceanu. A critical center-stable manifold for Schrödinger's equation in three dimensions. *Comm. Pure Appl. Math.*, 65(4):431–507, 2012.

[Bre11] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[Caz03] Thierry Cazenave. *Semilinear Schrödinger equations*, volume 10 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 2003.

[Con90] John B. Conway. *A course in functional analysis*, volume 96 of *Graduate texts in mathematics*. Springer, New York, 1990. 2. ed.

[Edw65] Robert E. Edwards. *Functional analysis. Theory and applications*. Holt, Rinehart and Winston, New York, 1965.

[EN99] Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations*, volume 194 of *Graduate texts in mathematics*. Springer-Verlag, New York, Berlin, Heidelberg, 1999.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[GT01] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[GV79] J. Ginibre and G. Velo. On a class of nonlinear Schrödinger equations. *J. Funct. Anal.*, 32(1):1–71, 1979.

[Gra08] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, 2008.

[HP57] Einar Hille and Ralph S. Phillips. *Functional analysis and semi-groups*. American Mathematical Society Colloquium Publications, vol. 31. American Mathematical Society, Providence, R. I., 1957. rev. ed.

[HS65] Edwin Hewitt and Karl Stromberg. *Real and abstract analysis. A modern treatment of the theory of functions of a real variable*. Springer-Verlag, New York, 1965.

[Kal02] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[Kat95] Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[KT98] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.

[Kie04] Hansjörg Kielhöfer. *Bifurcation theory, An introduction with applications to PDEs*, volume 156 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.

[Kry08] Nicolai V. Krylov. *Lectures on elliptic and parabolic equations in Sobolev spaces*, volume 96 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.

[Lan93] Serge Lang. *Real and functional analysis*, volume 142 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 1993.

[Lun95] Alessandra Lunardi. *Analytic semigroups and optimal regularity in parabolic problems*. Progress in Nonlinear Differential Equations and their Applications, 16. Birkhäuser Verlag, Basel, 1995.

[MN04] Jerome Moloney and Alan Newell. *Nonlinear optics*. Westview Press. Advanced Book Program, Boulder, CO, 2004.

[MS98] Stephen Montgomery-Smith. Time decay for the bounded mean oscillation of solutions of the Schrödinger and wave equations. *Duke Math. J.*, 91(2):393–408, 1998.

[vN96] Jan van Neerven. *The asymptotic behaviour of semigroups of linear operators*, volume 88 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1996.

[Ouh05] El Maati Ouhabaz. *Analysis of heat equations on domains*, volume 31 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2005.

[Paz83] Amnon Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied mathematical sciences*. Springer, New York, 1983.

[Ren94] Michael Renardy. On the linear stability of hyperbolic PDEs and viscoelastic flows. *Z. Angew. Math. Phys.*, 45(6):854–865, 1994.

[RR04] Michael Renardy and Robert C. Rogers. *An introduction to partial differential equations*, volume 13 of *Texts in Applied Mathematics*. Springer, New York, second edition, 2004.

[RS72] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I. Functional analysis*. Academic Press, New York, 1972.

[RS75] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press, New York, 1975.

[Rud87] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[Rud91] Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.

[Sch09] Wilhelm Schlag. Stable manifolds for an orbitally unstable nonlinear Schrödinger equation. *Ann. of Math. (2)*, 169(1):139–227, 2009.

[Str77] Robert S. Strichartz. Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations. *Duke Math. J.*, 44(3):705–714, 1977.

[SW11] Guido Schneider and C. Eugene Wayne. Justification of the NLS approximation for a quasilinear water wave model. *J. Differential Equations*, 251(2):238–269, 2011.

[Tan97] Hiroki Tanabe. *Functional analytic methods for partial differential equations*, volume 204 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1997.

[Tao06] Terence Tao. *Nonlinear dispersive equations. Local and global analysis*, volume 106 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006.

[TL80] Angus Ellis Taylor and David C. Lay. *Introduction to functional analysis*. John Wiley & Sons, New York-Chichester-Brisbane, second edition, 1980.

[WW96] Lutz Weis and Volker Wrobel. Asymptotic behavior of C_0-semigroups in Banach spaces. *Proc. Amer. Math. Soc.*, 124(12):3663–3671, 1996.

[Wer07] Dirk Werner. *Funktionalanalysis*. Springer-Lehrbuch. Springer Verlag, Berlin, 2007. 6., korr. Aufl.

[Yos80] Kôsaku Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, sixth edition, 1980.

[Zab75] Jerzy Zabczyk. A note on C_0-semigroups. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 23(8):895–898, 1975.