CHAPTER 5

The Heat and Schrödinger Equations

The heat, or diffusion, equation is

$$u_t = \Delta u. \tag{5.1}$$

Section 4.A derives (5.1) as a model of heat flow.

Steady solutions of the heat equation satisfy Laplace's equation. Using (2.4), we have for smooth functions that

$$\begin{aligned}\Delta u(x) &= \lim_{r\to 0^+} \fint_{B_r(x)} \Delta u \, dx \\ &= \lim_{r\to 0^+} \frac{n}{r}\frac{\partial}{\partial r}\left[\fint_{\partial B_r(x)} u \, dS\right] \\ &= \lim_{r\to 0^+} \frac{2n}{r^2}\left[\fint_{\partial B_r(x)} u \, dS - u(x)\right].\end{aligned}$$

Thus, if u is a solution of the heat equation, then the rate of change of $u(x,t)$ with respect to t at a point x is proportional to the difference between the value of u at x and the average of u over nearby spheres centered at x. The solution decreases in time if its value at a point is greater than the nearby mean and increases if its value is less than the nearby averages. The heat equation therefore describes the evolution of a function towards its mean. As $t \to \infty$ solutions of the heat equation typically approach functions with the mean value property, which are solutions of Laplace's equation.

We will also consider the Schrödinger equation

$$iu_t = -\Delta u.$$

This PDE is a dispersive wave equation, which describes a complex wave-field that oscillates with a frequency proportional to the difference between the value of the function and its nearby means.

5.1. The initial value problem for the heat equation

Consider the initial value problem for $u(x,t)$ where $x \in \mathbb{R}^n$

$$\begin{aligned} u_t &= \Delta u \qquad \text{for } x \in \mathbb{R}^n \text{ and } t > 0, \\ u(x,0) &= f(x) \qquad \text{for } x \in \mathbb{R}^n. \end{aligned} \tag{5.2}$$

We will solve (5.2) explicitly for smooth initial data by use of the Fourier transform, following the presentation in [**34**]. Some of the main qualitative features illustrated by this solution are the smoothing effect of the heat equation, the irreversibility of its semiflow, and the need to impose a growth condition as $|x| \to \infty$ in order to pick out a unique solution.

5.1.1. Schwartz solutions. Assume first that the initial data $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth, rapidly decreasing, real-valued Schwartz function $f \in \mathcal{S}$ (see Section 5.6.2). The solution we construct is also a Schwartz function of x at later times $t > 0$, and we will regard it as a function of time with values in $\mathcal{S}$. This is analogous to the geometrical interpretation of a first-order system of ODEs, in which the finite-dimensional phase space of the ODE is replaced by the infinite-dimensional function space $\mathcal{S}$; we then think of a solution of the heat equation as a parametrized curve in the vector space $\mathcal{S}$. A similar viewpoint is useful for many evolutionary PDEs, where the Schwartz space may be replaced other function spaces (for example, Sobolev spaces).

By a convenient abuse of notation, we use the same symbol u to denote the scalar-valued function $u(x,t)$, where $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$, and the associated vector-valued function $u(t)$, where $u : [0,\infty) \to \mathcal{S}$. We write the vector-valued function corresponding to the associated scalar-valued function as $u(t) = u(\cdot,t)$.

Definition 5.1. Suppose that (a,b) is an open interval in $\mathbb{R}$. A function $u : (a,b) \to \mathcal{S}$ is continuous at $t \in (a,b)$ if

$$u(t+h) \to u(t) \qquad \text{in } \mathcal{S} \text{ as } h \to 0,$$

and differentiable at $t \in (a,b)$ if there exists a function $v \in \mathcal{S}$ such that

$$\frac{u(t+h)-u(t)}{h} \to v \qquad \text{in } \mathcal{S} \text{ as } h \to 0.$$

The derivative v of u at t is denoted by $u_t(t)$, and if u is differentiable for every $t \in (a,b)$, then $u_t : (a,b) \to \mathcal{S}$ denotes the map $u_t : t \mapsto u_t(t)$.

In other words, u is continuous at t if

$$u(t) = \mathcal{S}\text{-}\lim_{h\to 0} u(t+h),$$

and u is differentiable at t with derivative $u_t(t)$ if

$$u_t(t) = \mathcal{S}\text{-}\lim_{h\to 0} \frac{u(t+h)-u(t)}{h}.$$

We will refer to this derivative as a strong derivative if it is understood that we are considering $\mathcal{S}$-valued functions and we want to emphasize that the derivative is defined as the limit of difference quotients in $\mathcal{S}$.

We define spaces of differentiable Schwartz-valued functions in the natural way. For half-open or closed intervals, we make the obvious modifications to left or right limits at an endpoint.

Definition 5.2. The space $C([a,b];\mathcal{S})$ consists of the continuous functions

$$u : [a,b] \to \mathcal{S}.$$

The space $C^k(a,b;\mathcal{S})$ consists of functions $u : (a,b) \to \mathcal{S}$ that are k-times strongly differentiable in (a,b) with continuous strong derivatives $\partial_t^j u \in C(a,b;\mathcal{S})$ for $0 \le j \le k$, and $C^\infty(a,b;\mathcal{S})$ is the space of functions with continuous strong derivatives of all orders.

Here we write $C(a,b;\mathcal{S})$ rather than $C((a,b);\mathcal{S})$ when we consider functions defined on the open interval (a,b). The next proposition describes the relationship between the C^1-strong derivative and the pointwise time-derivative.

PROPOSITION 5.3. *Suppose that* $u \in C(a, b; \mathcal{S})$ *where* $u(t) = u(\cdot, t)$. *Then* $u \in C^1(a, b; \mathcal{S})$ *if and only if:*

(1) *the pointwise partial derivative* $\partial_t u(x, t)$ *exists for every* $x \in \mathbb{R}^n$ *and* $t \in (a, b)$*;*
(2) $\partial_t u(\cdot, t) \in \mathcal{S}$ *for every* $t \in (a, b)$*;*
(3) *the map* $t \mapsto \partial_t u(\cdot, t)$ *belongs* $C(a, b; \mathcal{S})$.

PROOF. The convergence of functions in $\mathcal{S}$ implies uniform pointwise convergence. Thus, if $u(t) = u(\cdot, t)$ is strongly continuously differentiable, then the pointwise partial derivative $\partial_t u(x, t)$ exists for every $x \in \mathbb{R}^n$ and $\partial_t u(\cdot, t) = u_t(t) \in \mathcal{S}$, so $\partial_t u \in C(a, b; \mathcal{S})$.

Conversely, if a pointwise partial derivative with the given properties exist, then for each $x \in \mathbb{R}^n$

$$\frac{u(x, t+h) - u(x, t)}{h} - \partial_t u(x, t) = \frac{1}{h}\int_t^{t+h} \left[\partial_s u(x, s) - \partial_t u(x, t)\right] ds.$$

Since the integrand is a smooth rapidly decreasing function, it follows from the dominated convergence theorem that we may differentiate under the integral sign with respect to x, to get

$$x^\alpha \partial^\beta \left[\frac{u(x, t+h) - u(x, t)}{h}\right] = \frac{1}{h}\int_t^{t+h} x^\alpha \partial^\beta \left[\partial_s u(x, s) - \partial_t u(x, t)\right] ds.$$

Hence, if $\|\cdot\|_{\alpha,\beta}$ is a Schwartz seminorm (5.72), we have

$$\begin{aligned}\left\|\frac{u(t+h) - u(t)}{h} - \partial_t u(\cdot, t)\right\|_{\alpha,\beta} &\le \frac{1}{|h|}\left|\int_t^{t+h} \|\partial_s u(\cdot, s) - \partial_t u(\cdot, t)\|_{\alpha,\beta}\, ds\right| \\ &\le \max_{t \le s \le t+h} \|\partial_s u(\cdot, s) - \partial_t u(\cdot, t)\|_{\alpha,\beta},\end{aligned}$$

and since $\partial_t u \in C(a, b; \mathcal{S})$

$$\lim_{h\to 0} \left\|\frac{u(t+h) - u(t)}{h} - \partial_t u(\cdot, t)\right\|_{\alpha,\beta} = 0.$$

It follows that

$$\mathcal{S}\text{-}\lim_{h\to 0} \left[\frac{u(t+h) - u(t)}{h}\right] = \partial_t u(\cdot, t),$$

so u is strongly differentiable and $u_t = \partial_t u \in C(a, b; \mathcal{S})$. □

We interpret the initial value problem (5.2) for the heat equation as follows: A solution is a function $u : [0, \infty) \to \mathcal{S}$ that is continuous for $t \ge 0$, so that it makes sense to impose the initial condition at $t = 0$, and continuously differentiable for $t > 0$, so that it makes sense to impose the PDE pointwise in t. That is, for every $t > 0$, the strong derivative $u_t(t)$ is required to exist and equal $\Delta u(t)$ where $\Delta : \mathcal{S} \to \mathcal{S}$ is the Laplacian operator.

THEOREM 5.4. *If* $f \in \mathcal{S}$, *there is a unique solution*

$$u \in C([0, \infty); \mathcal{S}) \cap C^1(0, \infty; \mathcal{S}) \tag{5.3}$$

of (5.2). Furthermore, $u \in C^\infty([0, \infty); \mathcal{S})$. *The spatial Fourier transform of the solution is given by*

$$\hat{u}(k, t) = \hat{f}(k) e^{-t|k|^2}, \tag{5.4}$$

and for $t > 0$ the solution is given by

$$u(x,t) = \int_{\mathbb{R}^n} \Gamma(x-y,t) f(y)\, dy \tag{5.5}$$

where

$$\Gamma(x,t) = \frac{1}{(4\pi t)^{n/2}}\, e^{-|x|^2/4t}. \tag{5.6}$$

PROOF. Since the spatial Fourier transform $\mathcal{F}$ is a continuous linear map on $\mathcal{S}$ with continuous inverse, the time-derivative of u exists if and only if the time derivative of $\hat{u} = \mathcal{F}u$ exists, and

$$\mathcal{F}(u_t) = (\mathcal{F}u)_t\,.$$

Moreover, $u \in C([0,\infty);\mathcal{S})$ if and only if $\hat{u} \in C([0,\infty);\mathcal{S})$, and $u \in C^k(0,\infty;\mathcal{S})$ if and only if $\hat{u} \in C^k(0,\infty;\mathcal{S})$.

Taking the Fourier transform of (5.2) with respect to x, we find that $u(x,t)$ is a solution with the regularity in (5.3) if and only if $\hat{u}(k,t)$ satisfies

$$\hat{u}_t = -|k|^2\hat{u}, \qquad \hat{u}(0) = \hat{f}, \qquad \hat{u} \in C([0,\infty);\mathcal{S}) \cap C^1(0,\infty;\mathcal{S})\,. \tag{5.7}$$

Equation (5.7) has the unique solution (5.4).

To show this in detail, suppose first that $\hat{u}$ satisfies (5.7). Then, from Proposition 5.3, the scalar-valued function $\hat{u}(k,t)$ is pointwise-differentiable with respect to t in $t > 0$ and continuous in $t \geq 0$ for each fixed $k \in \mathbb{R}^n$. Solving the ODE (5.7) with k as a parameter, we find that $\hat{u}$ must be given by (5.4).

Conversely, we claim that the function defined by (5.4) is strongly differentiable with derivative

$$\hat{u}_t(k,t) = -|k|^2\hat{f}(k)e^{-t|k|^2}. \tag{5.8}$$

To prove this claim, note that if $\alpha, \beta \in \mathbb{N}_0^n$ are any multi-indices, the function

$$k^\alpha \partial^\beta \left[\hat{u}(k,t+h) - \hat{u}(k,t)\right]$$

has the form

$$\hat{a}(k,t)\left[e^{-h|k|^2} - 1\right]e^{-t|k|^2} + h\sum_{i=0}^{|\beta|-1} h^i \hat{b}_i(k,t)e^{-(t+h)|k|^2}$$

where $\hat{a}(\cdot,t), \hat{b}_i(\cdot,t) \in \mathcal{S}$, so taking the supremum of this expression we see that

$$\|\hat{u}(t+h) - \hat{u}(t)\|_{\alpha,\beta} \to 0 \qquad \text{as } h \to 0.$$

Thus, $\hat{u}(\cdot,t)$ is a continuous $\mathcal{S}$-valued function in $t \geq 0$ for every $\hat{f} \in \mathcal{S}$. By a similar argument, the pointwise partial derivative $\hat{u}_t(\cdot,t)$ in (5.8) is a continuous $\mathcal{S}$-valued function. Thus, Proposition 5.3 implies that $\hat{u}$ is a strongly continuously differentiable function that satisfies (5.7). Hence $u = \mathcal{F}^{-1}[\hat{u}]$ satisfies (5.3) and is a solution of (5.2). Moreover, using induction and Proposition 5.3 we see in a similar way that $u \in C^\infty([0,\infty);\mathcal{S})$.

Finally, from Example 5.65, we have

$$\mathcal{F}^{-1}\left[e^{-t|k|^2}\right] = \left(\frac{\pi}{t}\right)^{n/2} e^{-|x|^2/4t}.$$

Taking the inverse Fourier transform of (5.4) and using the convolution theorem, Theorem 5.67, we get (5.5)–(5.6). $\square$

The function $\Gamma(x,t)$ in (5.6) is called the Green's function or fundamental solution of the heat equation in $\mathbb{R}^n$. It is a C^∞-function of (x,t) in $\mathbb{R}^n \times (0,\infty)$, and one can verify by direct computation that

$$\Gamma_t = \Delta\Gamma \qquad \text{if } t > 0. \tag{5.9}$$

Also, since $\Gamma(\cdot,t)$ is a family of Gaussian mollifiers, we have

$$\Gamma(\cdot,t) \rightharpoonup \delta \qquad \text{in } \mathcal{S}' \text{ as } t \to 0^+.$$

Thus, we can interpret $\Gamma(x,t)$ as the solution of the heat equation due to an initial point source located at $x = 0$. The solution is a spherically symmetric Gaussian with spatial integral equal to one which spreads out and decays as t increases; its width is of the order $\sqrt{t}$ and its height is of the order $t^{-n/2}$.

The solution at time t is given by convolution of the initial data with $\Gamma(\cdot,t)$. For any $f \in \mathcal{S}$, this gives a smooth classical solution $u \in C^\infty(\mathbb{R}^n \times [0,\infty))$ of the heat equation which satisfies it pointwise in $t \geq 0$.

5.1.2. Smoothing. Equation (5.5) also gives solutions of (5.2) for initial data that is not smooth. To be specific, we suppose that $f \in L^p$, although one can also consider more general data that does not grow too rapidly at infinity.

THEOREM 5.5. *Suppose that $1 \leq p \leq \infty$ and $f \in L^p(\mathbb{R}^n)$. Define*

$$u : \mathbb{R}^n \times (0,\infty) \to \mathbb{R}$$

by (5.5) where Γ is given in (5.6). Then $u \in C_0^\infty(\mathbb{R}^n \times (0,\infty))$ and $u_t = \Delta u$ in $t > 0$. If $1 \leq p < \infty$, then $u(\cdot,t) \to f$ in L^p as $t \to 0^+$.

PROOF. The Green's function Γ in (5.6) satisfies (5.9), and $\Gamma(\cdot,t) \in L^q$ for every $1 \leq q \leq \infty$, together with all of its derivatives. The dominated convergence theorem and Hölder's inequality imply that if $f \in L^p$ and $t > 0$, we can differentiate under the integral sign in (5.5) arbitrarily often with respect to (x,t) and that all of these derivatives approach zero as $|x| \to \infty$. Thus, u is a smooth, decaying solution of the heat equation in $t > 0$. Moreover, $\Gamma^t(x) = \Gamma(x,t)$ is a family of Gaussian mollifiers and therefore for $1 \leq p < \infty$ we have from Theorem 1.28 that $u(\cdot,t) = \Gamma^t * f \to f$ in L^p as $t \to 0^+$. □

The heat equation therefore immediately smooths any initial data $f \in L^p(\mathbb{R}^n)$ to a function $u(\cdot,t) \in C_0^\infty(\mathbb{R}^n)$. From the Fourier perspective, the smoothing is a consequence of the very rapid damping of the high-wavenumber modes at a rate proportional to $e^{-t|k|^2}$ for wavenumbers $|k|$, which physically is caused by the diffusion of thermal energy from hot to cold parts of spatial oscillations.

Once the solution becomes smooth in space it also becomes smooth in time. In general, however, the solution is not (right) differentiable with respect to t at $t = 0$, and for rough initial data it satisfies the initial condition in an L^p-sense, but not necessarily pointwise.

5.1.3. Irreversibility. For general 'final' data $f \in \mathcal{S}$, we cannot solve the heat equation backward in time to obtain a solution $u : [-T,0] \to \mathcal{S}$, however small we choose $T > 0$. The same argument as the one in the proof of Theorem 5.4 implies that any such solution would be given by (5.4). If, for example, we take $f \in \mathcal{S}$ such that

$$\hat{f}(k) = e^{-\sqrt{1+|k|^2}}$$

then the corresponding solution

$$\hat{u}(k,t) = e^{-t|k|^2 - \sqrt{1+|k|^2}}$$

grows exponentially as $|k| \to \infty$ for every $t < 0$, and therefore $u(t)$ does not belong to $\mathcal{S}$ (or even $\mathcal{S}'$). Physically, this means that the temperature distribution f cannot arise by thermal diffusion from any previous temperature distribution in $\mathcal{S}$ (or $\mathcal{S}'$). The heat equation does, however, have a backward uniqueness property, meaning that if f arises from a previous temperature distribution, then (under appropriate assumptions) that distribution is unique [**9**].

Equivalently, making the time-reversal $t \mapsto -t$, we see that Schwartz-valued solutions of the initial value problem for the backward heat equation

$$u_t = -\Delta u \quad t > 0, \qquad u(x,0) = f(x)$$

do not exist for every $f \in \mathcal{S}$. Moreover, there is a loss of continuous dependence of the solution on the data.

Example 5.6. Consider the one-dimensional heat equation $u_t = u_{xx}$ with initial data

$$f_n(x) = e^{-n}\sin(nx)$$

and corresponding solution

$$u_n(x,t) = e^{-n}\sin(nx)e^{n^2 t}.$$

Then $f_n \to 0$ uniformly together with of all its spatial derivatives as $n \to \infty$, but

$$\sup_{x\in\mathbb{R}} |u_n(x,t)| \to \infty$$

as $n \to \infty$ for any $t > 0$. Thus, the solution does not depend continuously on the initial data in $C_b^\infty(\mathbb{R}^n)$. Multiplying the initial data f_n by e^{-x^2}, we can get an example of the loss of continuous dependence in $\mathcal{S}$.

It is possible to obtain a well-posed initial value problem for the backward heat equation by restricting the initial data to a small enough space with a strong enough norm — for example, to a suitable Gevrey space of C^∞-functions whose spatial derivatives decay at a sufficiently fast rate as their order tends to infinity. These restrictions, however, limit the size of derivatives of all orders, and they are too severe to be useful in applications.

Nevertheless, the backward heat equation is of interest as an inverse problem, namely: Find the temperature distribution at a previous time that gives rise to an observed temperature distribution at the present time. There is a loss of continuous dependence in any reasonable function space for applications, because thermal diffusion damps out large, rapid variations in a previous temperature distribution leading to an imperceptible effect on an observed distribution. Special methods — such as Tychonoff regularization — must be used to formulate such ill-posed inverse problems and develop numerical schemes to solve them.[1]

[1] J. B. Keller, Inverse Problems, *Amer. Math. Month.* **83** (1976) illustrates the difficulty of inverse problems in comparison with the corresponding direct problems by the example of guessing the question to which the answer is "Nine W." The solution is given at the end of this section.

5.1.4. Nonuniqueness. A solution $u(x,t)$ of the initial value problem for the heat equation on $\mathbb{R}^n$ is not unique without the imposition of a suitable growth condition as $|x| \to \infty$. In the above analysis, this was provided by the requirement that $u(\cdot,t) \in \mathcal{S}$, but the much weaker condition that u grows more slowly than $Ce^{a|x|^2}$ as $|x| \to \infty$ for some constants C, a is sufficient to imply uniqueness [**9**].

EXAMPLE 5.7. Consider, for simplicity, the one-dimensional heat equation

$$u_t = u_{xx}.$$

As observed by Tychonoff (*c.f.* [**21**]), a formal power series expansion with respect to x gives the solution

$$u(x,t) = \sum_{n=0}^{\infty} \frac{g^{(n)}(t) x^{2n}}{(2n)!}$$

for some function $g \in C^\infty(\mathbb{R}^+)$. We can construct a nonzero solution with zero initial data by choosing $g(t)$ to be a nonzero C^∞-function all of whose derivatives vanish at $t = 0$ in such a way that this series converges uniformly for x in compact subsets of $\mathbb{R}$ and $t > 0$ to a solution of the heat equation. This is the case, for example, if

$$g(t) = \exp\left(-\frac{1}{t^2}\right).$$

The resulting solution, however, grows very rapidly as $|x| \to \infty$.

A physical interpretation of this nonuniqueness it is that heat can diffuse from infinity into an unbounded region of initially zero temperature if the solution grows sufficiently quickly. Mathematically, the nonuniqueness is a consequence of the the fact that the initial condition is imposed on a characteristic surface $t = 0$ of the heat equation, meaning that the heat equation does not determine the second-order normal (time) derivative u_{tt} on $t = 0$ in terms of the second-order tangential (spatial) derivatives u, Du, D^2u.

According to the Cauchy-Kowalewski theorem [**14**], any non-characteristic Cauchy problem with analytic initial data has a unique local analytic solution. If $t \in \mathbb{R}$ denotes the normal variable and $x \in \mathbb{R}^n$ the transverse variable, then in solving the PDE by a power series expansion in t we exchange one t-derivative for one x-derivative and the convergence of the Taylor series in x for the analytic initial data implies the convergence of the series for the solution in t. This existence and uniqueness fails for a characteristic initial value problem, such as the one for the heat equation.

The Cauchy-Kowalewski theorem is not as useful as its apparent generality suggests because it does not imply anything about the stability or existence of solutions under non-analytic perturbations, even arbitrarily smooth ones. For example, the Cauchy-Kowalewski theorem is equally applicable to the initial value problem for the wave equation

$$u_{tt} = u_{xx}, \qquad u(x,0) = f(x),$$

which is well-posed in every Sobolev space $H^s(\mathbb{R})$, and the initial value problem for the Laplace equation

$$u_{tt} = -u_{xx}, \qquad u(x,0) = f(x),$$

which is ill-posed in every Sobolev space $H^s(\mathbb{R})$.[2]

5.2. Generalized solutions

In this section we obtain generalized solutions of the initial value problem of the heat equation as a limit of the smooth solutions constructed above. In order to do this, we require estimates on the smooth solutions which ensure that the convergence of initial data in suitable norms implies the convergence of the corresponding solution.

5.2.1. Estimates for the Heat equation. Solutions of the heat equation satisfy two basic spatial estimates, one in L^2 and the L^∞. The L^2 estimate follows from the Fourier representation, and the L^1 estimate follows from the spatial representation. For $1 \leq p < \infty$, we let

$$\|f\|_{L^p} = \left(\int_{\mathbb{R}^n} |f|^p \, dx \right)^{1/p}$$

denote the spatial L^p-norm of a function f; also $\|f\|_{L^\infty}$ denotes the maximum or essential supremum of $|f|$.

THEOREM 5.8. *Let $u : [0,\infty) \to \mathcal{S}(\mathbb{R}^n)$ be the solution of (5.2) constructed in Theorem 5.4 and $t > 0$. Then*

$$\|u(t)\|_{L^2} \leq \|f\|_{L^2}, \qquad \|u(t)\|_{L^\infty} \leq \frac{1}{(4\pi t)^{n/2}} \|f\|_{L^1}.$$

PROOF. By Parseval's inequality and (5.4),

$$\|u(t)\|_{L^2} = (2\pi)^n \|\hat{u}(t)\|_{L^2} = (2\pi)^n \left\| e^{-t|k|^2} \hat{f} \right\|_{L^2} \leq (2\pi)^n \|\hat{f}\|_{L^2} = \|f\|_{L^2},$$

which gives the first inequality. From (5.5),

$$|u(x,t)| \leq \left(\sup_{x \in \mathbb{R}^n} |\Gamma(x,t)| \right) \int_{\mathbb{R}^n} |f(y)| \, dy,$$

and from (5.6)

$$|\Gamma(x,t)| = \frac{1}{(4\pi t)^{n/2}}.$$

The second inequality then follows. □

Using the Riesz-Thorin theorem, Theorem 5.72, it follows by interpolation between $(p,p') = (2,2)$ and $(p,p') = (\infty,1)$, that for $2 \leq p \leq \infty$

$$\|u(t)\|_{L^p} \leq \frac{1}{(4\pi t)^{n(1/2-1/p)}} \|f\|_{L^{p'}}. \tag{5.10}$$

This estimate is not particularly useful for the heat equation, because we can derive stronger parabolic estimates for $\|Du\|_{L^2}$, but the analogous estimate for the Schrödinger equation is very useful.

A generalization of the L^2-estimate holds in any Sobolev space H^s of functions with s spatial L^2-derivatives (see Section 5.C for their definition). Such estimates of L^2-norms of solutions or their derivative are typically referred to as energy estimates, although the corresponding L^2-norms may not correspond to a physical

[2]Finally, here is the question to the answer posed above: Do you spell your name with a "V," Herr Wagner?

energy. In the case of the heat equation, the thermal energy (measured from a zero-point energy at $u = 0$) is proportional to the integral of u.

THEOREM 5.9. *Suppose that $f \in \mathcal{S}$ and $u \in C^\infty([0,\infty);\mathcal{S})$ is the solution of (5.2). Then for any $s \in \mathbb{R}$ and $t \geq 0$*

$$\|u(t)\|_{H^s} \leq \|f\|_{H^s}.$$

PROOF. Using (5.4) and Parseval's identity, and writing $\langle k\rangle = (1+|k|^2)^{1/2}$, we find that

$$\|u(t)\|_{H^s} = (2\pi)^n \left\|\langle k\rangle^s e^{-t|k|^2}\hat{f}\right\|_{L^2} \leq (2\pi)^n \left\|\langle k\rangle^s \hat{f}\right\|_{L^2} = \|f\|_{H^s}.$$

□

We can also derive this H^s-estimate, together with an additional a space-time estimate for Du, directly from the equation without using the explicit solution. We will use this estimate later to construct solutions of a general parabolic PDE by the Galerkin method, so we derive it here directly.

For $1 \leq p < \infty$ and $T > 0$, the L^p-in-time-H^s-in-space norm of a function $u \in C([0,T];\mathcal{S})$ is given by

$$\|u\|_{L^p([0,T];H^s)} = \left(\int_0^T \|u(t)\|_{H^s}^p \, dt\right)^{1/p}.$$

The maximum-in-time-H^s-in-space norm of u is

$$\|u\|_{C([0,T];H^s)} = \max_{t\in[0,T]} \|u(t)\|_{H^s}. \tag{5.11}$$

In particular, if $\Lambda = (I-\Delta)^{1/2}$ is the spatial operator defined in (5.75), then

$$\|u\|_{L^2([0,T];H^s)} = \left(\int_0^T \int_{\mathbb{R}^n} |\Lambda^s u(x,t)|^2 \, dxdt\right)^{1/2}.$$

THEOREM 5.10. *Suppose that $f \in \mathcal{S}$ and $u \in C^\infty([0,T];\mathcal{S})$ is the solution of (5.2). Then for any $s \in \mathbb{R}$*

$$\|u\|_{C([0,T];H^s)} \leq \|f\|_{H^s}, \qquad \|Du\|_{L^2([0,T];H^s)} \leq \frac{1}{\sqrt{2}}\|f\|_{H^s}.$$

PROOF. Let $v = \Lambda^s u$. Then, since $\Lambda^s : \mathcal{S} \to \mathcal{S}$ is continuous and commutes with Δ,

$$v_t = \Delta v, \qquad v(0) = g$$

where $g = \Lambda^s f$. Multiplying this equation by v, integrating the result over $\mathbb{R}^n$, and using the divergence theorem (justified by the continuous differentiability in time and the smoothness and decay in space of v), we get

$$\frac{1}{2}\frac{d}{dt}\int v^2 \, dx = -\int |Dv|^2 \, dx.$$

Integrating this equation with respect to t, we obtain for any $T > 0$ that

$$\frac{1}{2}\int v^2(T)\, dx + \int_0^T \int |Dv(t)|^2 \, dxdt = \frac{1}{2}\int g^2 \, dx. \tag{5.12}$$

Thus,

$$\max_{t\in[0,T]} \int v^2(t)\,dx \le \int g^2\,dx, \qquad \int_0^T \int |Dv(t)|^2\,dxdt \le \frac{1}{2}\int g^2\,dx,$$

and the result follows. $\square$

5.2.2. H^s-solutions. In this section we use the above estimates to obtain generalized solutions of the heat equation as a limit of smooth solutions (5.5). In defining generalized solutions, it is convenient to restrict attention to a finite, but arbitrary, time-interval $[0,T]$ where $T > 0$. For $s \in \mathbb{R}$, let $C([0,T];H^s)$ denote the Banach space of continuous H^s-valued functions $u : [0,T] \to H^s$ equipped with the norm (5.11).

Definition 5.11. Suppose that $T > 0$, $s \in \mathbb{R}$ and $f \in H^s$. A function

$$u \in C([0,T];H^s)$$

is a generalized solution of (5.2) if there exists a sequence of Schwartz-solutions $u_n : [0,T] \to \mathcal{S}$ such that $u_n \to u$ in $C([0,T];H^s)$ as $n \to \infty$.

According to the next theorem, there is a unique generalized solution defined on any time interval $[0,T]$ and therefore on $[0,\infty)$.

Theorem 5.12. *Suppose that $T > 0$, $s \in \mathbb{R}$ and $f \in H^s(\mathbb{R}^n)$. Then there is a unique generalized solution $u \in C([0,T];H^s)$ of (5.2). The solution is given by (5.4).*

Proof. Since $\mathcal{S}$ is dense in H^s, there is a sequence of functions $f_n \in \mathcal{S}$ such that $f_n \to f$ in H^s. Let $u_n \in C([0,T];\mathcal{S})$ be the solution of (5.2) with initial data f_n. Then, by linearity, $u_n - u_m$ is the solution with initial data $f_n - f_m$, and Theorem 5.9 implies that

$$\sup_{t\in[0,T]} \|u_n(t) - u_m(t)\|_{H^s} \le \|f_n - f_m\|_{H^s}.$$

Hence, $\{u_n\}$ is a Cauchy sequence in $C([0,T];H^s)$ and therefore there exists a generalized solution $u \in C([0,T];H^s)$ such that $u_n \to u$ as $n \to \infty$.

Suppose that $f, g \in H^s$ and $u, v \in C([0,T];H^s)$ are generalized solutions with $u(0) = f$, $v(0) = g$. If $u_n, v_n \in C([0,T];\mathcal{S})$ are approximate solutions with $u_n(0) = f_n$, $v_n(0) = g_n$, then

$$\begin{aligned}\|u(t) - v(t)\|_{H^s} &\le \|u(t) - u_n(t)\|_{H^s} + \|u_n(t) - v_n(t)\|_{H^s} + \|v_n(t) - v(t)\|_{H^s}\\ &\le \|u(t) - u_n(t)\|_{H^s} + \|f_n - g_n\|_{H^s} + \|v_n(t) - v(t)\|_{H^s}\end{aligned}$$

Taking the limit of this inequality as $n \to \infty$, we find that

$$\|u(t) - v(t)\|_{H^s} \le \|f - g\|_{H^s}.$$

In particular, if $f = g$ then $u = v$, so a generalized solution is unique.

Finally, from (5.4) we have

$$\hat{u}_n(k,t) = e^{-t|k|^2}\hat{f}_n(k).$$

Taking the limit of this expression in $C([0,T];\hat{H}^s)$ as $n \to \infty$, where $\hat{H}^s$ is the weighted L^2-space (5.74), we get the same expression for $\hat{u}$. $\square$

We may obtain additional regularity of generalized solutions in time by use of the equation; roughly speaking, we can trade two space-derivatives for one time-derivative.

PROPOSITION 5.13. *Suppose that* $T > 0$, $s \in \mathbb{R}$ *and* $f \in H^s(\mathbb{R}^n)$. *If* $u \in C([0,T];H^s)$ *is a generalized solution of (5.2), then* $u \in C^1([0,T];H^{s-2})$ *and*

$$u_t = \Delta u \qquad \text{in } C([0,T];H^{s-2}).$$

PROOF. Since u is a generalized solution, there is a sequence of smooth solutions $u_n \in C^\infty([0,T];\mathcal{S})$ such that $u_n \to u$ in $C([0,T];H^s)$ as $n \to \infty$. These solutions satisfy $u_{nt} = \Delta u_n$. Since $\Delta : H^s \to H^{s-2}$ is bounded and $\{u_n\}$ is Cauchy in H^s, we see that $\{u_{nt}\}$ is Cauchy in $C([0,T];H^{s-2})$. Hence there exists $v \in C([0,T];H^{s-2})$ such that $u_{nt} \to v$ in $C([0,T];H^{s-2})$. We claim that $v = u_t$. For each $n \in \mathbb{N}$ and $h \neq 0$ we have

$$\frac{u_n(t+h) - u_n(t)}{h} = \frac{1}{h}\int_t^{t+h} u_{ns}(s)\, ds \qquad \text{in } C([0,T];\mathcal{S}),$$

and in the limit $n \to \infty$, we get that

$$\frac{u(t+h) - u(t)}{h} = \frac{1}{h}\int_t^{t+h} v(s)\, ds \qquad \text{in } C([0,T];H^{s-2}).$$

Taking the limit as $h \to 0$ of this equation we find that $u_t = v$ and

$$u \in C([0,T];H^s) \cap C^1([0,T];H^{s-2}).$$

Moreover, taking the limit of $u_{nt} = \Delta u_n$ we get $u_t = \Delta u$ in $C([0,T];H^{s-2})$. □

More generally, a similar argument shows that $u \in C^k([0,T];H^{s-2k})$ for any $k \in \mathbb{N}$. In contrast with the case of ODEs, the time derivative of the solution lies in a different space than the solution itself: u takes values in H^s, but u_t takes values in H^{s-2}. This feature is typical for PDEs when — as is usually the case — one considers solutions that take values in Banach spaces whose norms depend on only finitely many derivatives. It did not arise for Schwartz-valued solutions, since differentiation is a continuous operation on $\mathcal{S}$.

The above proposition did not use any special properties of the heat equation. For $t > 0$, solutions have greatly improved regularity as a result of the smoothing effect of the evolution.

PROPOSITION 5.14. *If* $u \in C([0,T];H^s)$ *is a generalized solution of (5.2), where* $f \in H^s$ *for some* $s \in \mathbb{R}$, *then* $u \in C^\infty((0,T];H^\infty)$ *where* H^∞ *is defined in (5.76).*

PROOF. If $s \in \mathbb{R}$, $f \in H^s$, and $t > 0$, then (5.4) implies that $\hat{u}(t) \in \hat{H}^r$ for every $r \in \mathbb{R}$, and therefore $u(t) \in H^\infty$. It follows from the equation that $u \in C^\infty(0,\infty;H^\infty)$. □

For general H^s-initial data, however, we cannot expect any improved regularity in time at $t = 0$ beyond $u \in C^k([0,T);H^{s-2k})$. The H^∞ spatial regularity stated here is not optimal; for example, one can prove [**9**] that the solution is a real-analytic function of x for $t > 0$, although it is not necessarily a real-analytic function of t.

5.3. The Schrödinger equation

The initial value problem for the Schrödinger equation is

$$\begin{aligned} iu_t &= -\Delta u \qquad \text{for } x \in \mathbb{R}^n \text{ and } t \in \mathbb{R}, \\ u(x,0) &= f(x) \qquad \text{for } x \in \mathbb{R}^n, \end{aligned} \tag{5.13}$$

where $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{C}$ is a complex-valued function. A solution of the Schrödinger equation is the amplitude function of a quantum mechanical particle moving freely in $\mathbb{R}^n$. The function $|u(\cdot,t)|^2$ is proportional to the spatial probability density of the particle.

More generally, a particle moving in a potential $V : \mathbb{R}^n \to \mathbb{R}$ satisfies the Schrödinger equation

$$iu_t = -\Delta u + V(x)u. \tag{5.14}$$

Unlike the free Schrödinger equation (5.13), this equation has variable coefficients and it cannot be solved explicitly for general potentials V.

Formally, the Schrödinger equation (5.13) is obtained by the transformation $t \mapsto -it$ of the heat equation to 'imaginary time.' The analytical properties of the heat and Schrödinger equations are, however, completely different and it is interesting to compare them. The proofs are similar, and we leave them as an exercise (or see [**34**]).

The Fourier solution of (5.13) is

$$\hat{u}(k,t) = e^{-it|k|^2}\hat{f}(k). \tag{5.15}$$

The key difference from the heat equation is that these Fourier modes oscillate instead of decay in time, and higher wavenumber modes oscillate faster in time. As a result, there is no smoothing of the initial data (measuring smoothness in the L^2-scale of Sobolev spaces H^s) and we can solve the Schrödinger equation both forward and backward in time.

Theorem 5.15. *For any $f \in \mathcal{S}$ there is a unique solution $u \in C^\infty(\mathbb{R};\mathcal{S})$ of (5.13). The spatial Fourier transform of the solution is given by (5.15), and*

$$u(x,t) = \int \Gamma(x-y,t) f(y)\, dy$$

where

$$\Gamma(x,t) = \frac{1}{(4\pi i t)^{n/2}}\, e^{-i|x|^2/4t}.$$

We get analogous L^p estimates for the Schrödinger equation to the ones for the heat equation.

Theorem 5.16. *Suppose that $f \in \mathcal{S}$ and $u \in C^\infty(\mathbb{R};\mathcal{S})$ is the solution of (5.13). Then for all $t \in \mathbb{R}$,*

$$\|u(t)\|_{L^2} \le \|f\|_{L^2}, \qquad \|u(t)\|_{L^\infty} \le \frac{1}{(4\pi|t|)^{n/2}}\|f\|_{L^1},$$

and for $2 < p < \infty$,

$$\|u(t)\|_{L^p} \le \frac{1}{(4\pi|t|)^{n(1/2-1/p)}}\|f\|_{L^{p'}}. \tag{5.16}$$

Solutions of the Schrödinger equation do not satisfy a space-time estimate analogous to the parabolic estimate (5.12) in which we 'gain' a spatial derivative. Instead, we get only that the H^s-norm is conserved. Solutions do satisfy a weaker space-time estimate, called a Strichartz estimate, which we derive in Section 5.6.1.

The conservation of the H^s-norm follows from the Fourier representation (5.15), but let us prove it directly from the equation.

THEOREM 5.17. *Suppose that $f \in \mathcal{S}$ and $u \in C^\infty(\mathbb{R}; \mathcal{S})$ is the solution of (5.13). Then for any $s \in \mathbb{R}$*

$$\|u(t)\|_{H^s} = \|f\|_{H^s} \qquad \text{for every } t \in \mathbb{R}.$$

PROOF. Let $v = \Lambda^s u$, so that $\|u(t)\|_{H^s} = \|v(t)\|_{L^2}$. Then

$$iv_t = -\Delta v$$

and $v(0) = \Lambda^s f$. Multiplying this PDE by the conjugate $\bar{v}$ and subtracting the complex conjugate of the result, we get

$$i\left(\bar{v}v_t + v\bar{v}_t\right) = v\Delta\bar{v} - \bar{v}\Delta v.$$

We may rewrite this equation as

$$\partial_t |v|^2 + \nabla \cdot \left[i\left(vD\bar{v} - \bar{v}Dv\right)\right] = 0.$$

If $v = u$, this is the equation of conservation of probability where $|u|^2$ is the probability density and $i\left(uD\bar{u} - \bar{u}Du\right)$ is the probability flux. Integrating the equation over $\mathbb{R}^n$ and using the spatial decay of v, we get

$$\frac{d}{dt}\int |v|^2\, dx = 0,$$

and the result follows. □

We say that a function $u \in C\left(\mathbb{R}; H^s\right)$ is a generalized solution of (5.13) if it is the limit of smooth Schwartz-valued solutions uniformly on compact time intervals. The existence of such solutions follows from the preceding H^s-estimates for smooth solutions.

THEOREM 5.18. *Suppose that $s \in \mathbb{R}$ and $f \in H^s(\mathbb{R}^n)$. Then there is a unique generalized solution $u \in C\left(\mathbb{R}; H^s\right)$ of (5.13) given by*

$$\hat{u}(k) = e^{-it|k|^2}\hat{f}(k).$$

Moreover, for any $k \in \mathbb{N}$, we have $u \in C^k\left(\mathbb{R}; H^{s-2k}\right)$.

Unlike the heat equation, there is no smoothing of the solution and there is no H^s-regularity for $t \neq 0$ beyond what is stated in this theorem.

5.4. Semigroups and groups

The solution of an $n \times n$ linear first-order system of ODEs for $\vec{u}(t) \in \mathbb{R}^n$,

$$\vec{u}_t = A\vec{u},$$

may be written as

$$\vec{u}(t) = e^{tA}\vec{u}(0) \qquad -\infty < t < \infty$$

where $e^{tA} : \mathbb{R}^n \to \mathbb{R}^n$ is the matrix exponential of tA. The finite-dimensionality of the phase space $\mathbb{R}^n$ is not crucial here. As we discuss next, similar results hold for any linear ODE in a Banach space generated by a bounded linear operator.

5.4.1. Uniformly continuous groups. Suppose that X is a Banach space. We denote by $\mathcal{L}(X)$ the Banach space of bounded linear operators $A : X \to X$ equipped with the operator norm

$$\|A\|_{\mathcal{L}(X)} = \sup_{u \in X \setminus \{0\}} \frac{\|Au\|_X}{\|u\|_X}.$$

We say that a sequence of bounded linear operators converges uniformly if it converges with respect to the operator norm.

For $A \in \mathcal{L}(X)$ and $t \in \mathbb{R}$, we define the operator exponential by the series

$$e^{tA} = I + tA + \frac{1}{2!}t^2A^2 + \cdots + \frac{1}{n!}A^n + \dots. \tag{5.17}$$

This operator is well-defined. Its properties are similar to those of the real-valued exponential function e^{at} for $a \in \mathbb{R}$ and are proved in the same way.

THEOREM 5.19. *If $A \in \mathcal{L}(X)$ and $t \in \mathbb{R}$, then the series in (5.17) converges uniformly in $\mathcal{L}(X)$. Moreover, the function $t \mapsto e^{tA}$ belongs to $C^\infty(\mathbb{R}; \mathcal{L}(X))$ and for every $s, t \in \mathbb{R}$*

$$e^{sA}e^{tA} = e^{(s+t)A}, \qquad \frac{d}{dt}e^{tA} = Ae^{tA}.$$

Consider a linear homogeneous initial value problem

$$u_t = Au, \qquad u(0) = f \in X, \qquad u \in C^1(\mathbb{R}; X). \tag{5.18}$$

The solution is given by the operator exponential.

THEOREM 5.20. *The unique solution $u \in C^\infty(\mathbb{R}; X)$ of (5.18) is given by*

$$u(t) = e^{tA}f.$$

EXAMPLE 5.21. For $1 \le p < \infty$, let $A : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ be the bounded translation operator

$$Af(x) = f(x+1).$$

The solution $u \in C^\infty(\mathbb{R}; L^p)$ of the differential-difference equation

$$u_t(x,t) = u(x+1,t), \qquad u(x,0) = f(x)$$

is given by

$$u(x,t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} f(x+n).$$

EXAMPLE 5.22. Suppose that $a \in L^1(\mathbb{R}^n)$ and define the bounded convolution operator $A : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ by $Af = a * f$. Consider the IVP

$$u_t(x,t) = \int_{\mathbb{R}^n} a(x-y)u(y)\,dy, \qquad u(x,0) = f(x) \in L^2(\mathbb{R}^n).$$

Taking the Fourier transform of this equation and using the convolution theorem, we get

$$\hat{u}_t(k,t) = (2\pi)^n \hat{a}(k)\hat{u}(k,t), \qquad \hat{u}(k,0) = \hat{f}(k).$$

The solution is

$$\hat{u}(k,t) = e^{(2\pi)^n \hat{a}(k)t}\hat{f}(k).$$

It follows that

$$u(x,t) = \int g(x-y,t)f(y)\,dy$$

where the Fourier transform of $g(x,t)$ is given by

$$\hat{g}(k,t) = \frac{1}{(2\pi)^n} e^{(2\pi)^n \hat{a}(k) t}.$$

Since $a \in L^1(\mathbb{R}^n)$, the Riemann-Lebesgue lemma implies that $\hat{a} \in C_0(\mathbb{R}^n)$, and therefore $\hat{g}(\cdot,t) \in C_b(\mathbb{R}^n)$ for every $t \in \mathbb{R}$. Since convolution with g corresponds to multiplication of the Fourier transform by a bounded multiplier, it defines a bounded linear map on $L^2(\mathbb{R}^n)$.

The solution operators $\mathrm{T}(t) = e^{tA}$ of (5.18) form a uniformly continuous one-parameter group. Conversely, any uniformly continuous one-parameter group of transformations on a Banach space is generated by a bounded linear operator.

Definition 5.23. Let X be a Banach space. A one-parameter, uniformly continuous group on X is a family $\{\mathrm{T}(t) : t \in \mathbb{R}\}$ of bounded linear operators $\mathrm{T}(t) : X \to X$ such that:

(1) $\mathrm{T}(0) = I$;
(2) $\mathrm{T}(s)\mathrm{T}(t) = \mathrm{T}(s+t)$ for all $s,t \in \mathbb{R}$;
(3) $\mathrm{T}(h) \to I$ uniformly in $\mathcal{L}(X)$ as $h \to 0$.

Theorem 5.24. *If $\{\mathrm{T}(t) : t \in \mathbb{R}\}$ is a uniformly continuous group on a Banach space X, then:*

(1) $\mathrm{T} \in C^\infty(\mathbb{R}; \mathcal{L}(X))$;
(2) $A = \mathrm{T}_t(0)$ *is a bounded linear operator on* X;
(3) $\mathrm{T}(t) = e^{tA}$ *for every* $t \in \mathbb{R}$.

Note that the differentiability (and, in fact, the analyticity) of $\mathrm{T}(t)$ with respect to t is implied by its continuity and the group property $\mathrm{T}(s)\mathrm{T}(t) = \mathrm{T}(s+t)$. This is analogous to what happens for the real exponential function: The only continuous functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy the functional equation

$$f(0) = 1, \qquad f(s)f(t) = f(s+t) \quad \text{for all } s,t \in \mathbb{R} \tag{5.19}$$

are the exponential functions $f(t) = e^{at}$ for $a \in \mathbb{R}$, and these functions are analytic.

Some regularity assumption on f is required in order for (5.19) to imply that f is an exponential function. If we drop the continuity assumption, then the function defined by $f(0) = 1$ and $f(t) = 0$ for $t \neq 0$ also satisfies (5.19). This function and the exponential functions are the only Lebesgue measurable solutions of (5.19). If we drop the measurability requirement, then we get many other solutions.

Example 5.25. If $f = e^g$ where $g : \mathbb{R} \to \mathbb{R}$ satisfies

$$g(0) = 0, \qquad g(s) + g(t) = g(s+t),$$

then f satisfied (5.19). The linear functions $g(t) = at$ satisfy this functional equation for any $a \in \mathbb{R}$, but there are many other non-measurable solutions. To “construct” examples, consider $\mathbb{R}$ as a vector space over the field $\mathbb{Q}$ of rational numbers, and let $\{e_\alpha \in \mathbb{R} : \alpha \in I\}$ denote an algebraic basis. Given any values $\{c_\alpha \in \mathbb{R} : \alpha \in I\}$ define $g : \mathbb{R} \to \mathbb{R}$ such that $g(e_\alpha) = c_\alpha$ for each $\alpha \in I$, and if $x = \sum x_\alpha e_\alpha$ is the finite expansion of $x \in \mathbb{R}$ with respect to the basis, then

$$g\left(\sum x_\alpha e_\alpha\right) = \sum x_\alpha c_\alpha.$$

5.4.2. Strongly continuous semigroups. We may consider the heat equation and other linear evolution equations from a similar perspective to the Banach space ODEs discussed above. Significant differences arise, however, as a result of the fact that the Laplacian and other spatial differential operators are unbounded maps of a Banach space into itself. In particular, the solution operators associated with unbounded operators are strongly but not uniformly continuous functions of time, and we get solutions that are, in general, continuous but not continuously differentiable. Moreover, as in the case of the heat equation, we may only be able to solve the equation forward in time, which gives us a semigroup of solution operators instead of a group.

Abstracting the notion of a family of solution operators with continuous trajectories forward in time, we are led to the following definition.

Definition 5.26. Let X be a Banach space. A one-parameter, strongly continuous (or C_0) semigroup on X is a family $\{\mathrm{T}(t) : t \geq 0\}$ of bounded linear operators $\mathrm{T}(t) : X \to X$ such that

(1) $\mathrm{T}(0) = I$;
(2) $\mathrm{T}(s)\mathrm{T}(t) = \mathrm{T}(s + t)$ for all $s, t \geq 0$;
(3) $\mathrm{T}(h)f \to f$ strongly in X as $h \to 0^+$ for every $f \in X$.

The semigroup is said to be a contraction semigroup if $\|\mathrm{T}(t)\| \leq 1$ for all $t \geq 0$, where $\|\cdot\|$ denotes the operator norm.

The semigroup property (2) holds for the solution maps of any well-posed autonomous evolution equation: it says simply that we can solve for time $s + t$ by solving for time t and then for time s. Condition (3) means explicitly that

$$\|\mathrm{T}(t)f - f\|_X \to 0 \qquad \text{as } t \to 0^+.$$

If this holds, then the semigroup property (2) implies that $\mathrm{T}(t + h)f \to \mathrm{T}(t)f$ in X as $h \to 0$ for every $t > 0$, not only for $t = 0$ [**8**]. The term 'contraction' in Definition 5.26 is not used in a strict sense, and the norm of the solution of a contraction semigroup is not required to be strictly decreasing in time; it may for example, remain constant.

The heat equation

$$u_t = \Delta u, \qquad u(x, 0) = f(x) \tag{5.20}$$

is one of the primary motivating examples for the theory of semigroups. For definiteness, we suppose that $f \in L^2$, but we could also define a heat-equation semigroup on other Hilbert or Banach spaces, such as H^s or L^p for $1 < p < \infty$.

From Theorem 5.12 with $s = 0$, for every $f \in L^2$ there is a unique generalized solution $u : [0, \infty) \to L^2$ of (5.20), and therefore for each $t \geq 0$ we may define a bounded linear map $\mathrm{T}(t) : L^2 \to L^2$ by $\mathrm{T}(t) : f \mapsto u(t)$. The operator $\mathrm{T}(t)$ is defined explicitly by

$$\begin{aligned} &\mathrm{T}(0) = I, \quad \mathrm{T}(t)f = \Gamma^t * f \qquad \text{for } t > 0, \\ &\widehat{\mathrm{T}(t)f}(k) = e^{-t|k|^2}\hat{f}(k). \end{aligned} \tag{5.21}$$

where the $*$ denotes spatial convolution with the Green's function $\Gamma^t(x) = \Gamma(x, t)$ given in (5.6).

We also use the notation

$$\mathrm{T}(t) = e^{t\Delta}$$

and interpret $\mathrm{T}(t)$ as the operator exponential of $t\Delta$. The semigroup property then becomes the usual exponential formula

$$e^{(s+t)\Delta} = e^{s\Delta}e^{t\Delta}.$$

THEOREM 5.27. *The solution operators* $\{\mathrm{T}(t) : t \geq 0\}$ *of the heat equation defined in (5.21) form a strongly continuous contraction semigroup on* $L^2(\mathbb{R}^n)$.

PROOF. This theorem is a restatement of results that we have already proved, but let us verify it explicitly. The semigroup property follows from the Fourier representation, since

$$e^{-(s+t)|k|^2} = e^{-s|k|^2}e^{-t|k|^2}.$$

It also follows from the spatial representation, since

$$\Gamma^{s+t} = \Gamma^s * \Gamma^t.$$

The probabilistic interpretation of this identity is that the sum of independent Gaussian random variables is a Gaussian random variable, and the variance of the sum is the sum of the variances.

Theorem 5.12, with $s = 0$, implies that the semigroup is strongly continuous since $t \mapsto \mathrm{T}(t)f$ belongs to $C\left([0,\infty); L^2\right)$ for every $f \in L^2$. Finally, it is immediate from (5.21) and Parseval's theorem that $\|\mathrm{T}(t)\| \leq 1$ for every $t \geq 0$, so the semigroup is a contraction semigroup. $\square$

An alternative way to view this result is that the solution maps

$$\mathrm{T}(t) : \mathcal{S} \subset L^2 \to \mathcal{S} \subset L^2$$

constructed in Theorem 5.4 are defined on a dense subspace $\mathcal{S}$ of L^2, and are bounded on L^2, so they extend to bounded linear maps $\mathrm{T}(t) : L^2 \to L^2$, which form a strongly continuous semigroup.

Although for every $f \in L^2$ the trajectory $t \mapsto \mathrm{T}(t)f$ is a continuous function from $[0,\infty)$ into L^2, it is *not* true that $t \mapsto \mathrm{T}(t)$ is a continuous map from $[0,\infty)$ into the space $\mathcal{L}(L^2)$ of bounded linear maps on L^2 since $\mathrm{T}(t+h)$ does not converge to $\mathrm{T}(t)$ as $h \to 0$ uniformly with respect to the operator norm.

Proposition 5.13 implies a solution $t \mapsto \mathrm{T}(t)f$ belongs to $C^1\left([0,\infty); L^2\right)$ if $f \in H^2$, but for $f \in L^2 \setminus H^2$ the solution is not differentiable with respect to t in L^2 at $t = 0$. For every $t > 0$, however, we have from Proposition 5.14 that the solution belongs to $C^\infty(0,\infty; H^\infty)$. Thus, the the heat equation semiflow maps the entire phase space L^2 forward in time into a dense subspace H^∞ of smooth functions. As a result of this smoothing, we cannot reverse the flow to obtain a map backward in time of L^2 into itself.

5.4.3. Strongly continuous groups. Conservative wave equations do not smooth solutions in the same way as parabolic equations like the heat equation, and they typically define a group of solution maps both forward and backward in time.

DEFINITION 5.28. Let X be a Banach space. A one-parameter, strongly continuous (or C_0) group on X is a family $\{\mathrm{T}(t) : t \in \mathbb{R}\}$ of bounded linear operators $\mathrm{T}(t) : X \to X$ such that

(1) $\mathrm{T}(0) = I$;
(2) $\mathrm{T}(s)\mathrm{T}(t) = \mathrm{T}(s+t)$ for all $s, t \in \mathbb{R}$;
(3) $\mathrm{T}(h)f \to f$ strongly in X as $h \to 0$ for every $f \in X$.

If X is a Hilbert space and each $\mathrm{T}(t)$ is a unitary operator on X, then the group is said to be a unitary group.

Thus $\{\mathrm{T}(t) : t \in \mathbb{R}\}$ is a strongly continuous group if and only if $\{\mathrm{T}(t) : t \geq 0\}$ is a strongly continuous semigroup of invertible operators and $\mathrm{T}(-t) = \mathrm{T}^{-1}(t)$.

THEOREM 5.29. *Suppose that $s \in \mathbb{R}$. The solution operators $\{\mathrm{T}(t) : t \in \mathbb{R}\}$ of the Schrödinger equation (5.13) defined by*

$$(\widehat{\mathrm{T}(t)f})(k) = e^{-it|k|^2}\hat{f}(k). \tag{5.22}$$

form a strongly continuous, unitary group on $H^s(\mathbb{R}^n)$.

Unlike the heat equation semigroup, the Schrödinger equation is a dispersive wave equation which does not smooth solutions. The solution maps $\{\mathrm{T}(t) : t \in \mathbb{R}\}$ form a group of unitary operators on L^2 which map H^s onto itself (*c.f.* Theorem 5.17). A trajectory $u(t)$ belongs to $C^1(\mathbb{R}; L^2)$ if and only if $u(0) \in H^2$, and $u \in C^k(\mathbb{R}; L^2)$ if and only if $u(0) \in H^{1+k}$. If $u(0) \in L^2 \setminus H^2$, then $u \in C(\mathbb{R}; L^2)$ but u is nowhere strongly differentiable in L^2 with respect to time. Nevertheless, the continuous non-differentiable trajectories remain close in L^2 to the differentiable trajectories. This dense intertwining of smooth trajectories and continuous, non-differentiable trajectories in an infinite-dimensional phase space is not easy to imagine and has no analog for ODEs.

The Schrödinger operators $\mathrm{T}(t) = e^{it\Delta}$ do *not* form a strongly continuous group on $L^p(\mathbb{R}^n)$ when $p \neq 2$. Suppose, for contradiction, that $\mathrm{T}(t) : L^p \to L^p$ is bounded for some $1 \leq p < \infty$, $p \neq 2$ and $t \in \mathbb{R} \setminus \{0\}$. Then since $\mathrm{T}(-t) = T^*(t)$, duality implies that $\mathrm{T}(-t) : L^{p'} \to L^{p'}$ is bounded, and we can assume that $1 \leq p < 2$ without loss of generality. From Theorem 5.16, $\mathrm{T}(t) : L^p \to L^{p'}$ is bounded, and thus for every $f \in L^p \cap L^{p'} \subset L^2$

$$\|f\|_{L^p} = \|\mathrm{T}(t)\mathrm{T}(-t)f\|_{L^p} \leq C_1 \|\mathrm{T}(-t)f\|_{L^{p'}} \leq C_1 C_2 \|f\|_{L^{p'}}.$$

This estimate is false if $p \neq 2$, so $\mathrm{T}(t)$ cannot be bounded on L^p.

5.4.4. Generators. Given an operator A that generates a semigroup, we may define the semigroup $\mathrm{T}(t) = e^{tA}$ as the collection of solution operators of the equation $u_t = Au$. Alternatively, given a semigroup, we may ask for an operator A that generates it.

DEFINITION 5.30. Suppose that $\{\mathrm{T}(t) : t \geq 0\}$ is a strongly continuous semigroup on a Banach space X. The generator A of the semigroup is the linear operator in X with domain $\mathcal{D}(A)$,

$$A : \mathcal{D}(A) \subset X \to X,$$

defined as follows:

(1) $f \in \mathcal{D}(A)$ if and only if the limit

$$\lim_{h \to 0^+} \left[\frac{\mathrm{T}(h)f - f}{h}\right]$$

exists with respect to the strong (norm) topology of X;

(2) if $f \in \mathcal{D}(A)$, then

$$Af = \lim_{h \to 0^+} \left[\frac{\mathrm{T}(h)f - f}{h}\right].$$

To describe which operators are generators of a semigroup, we recall some definitions and results from functional analysis. See [**8**] for further discussion and proofs of the results.

Definition 5.31. An operator $A : \mathcal{D}(A) \subset X \to X$ in a Banach space X is closed if whenever $\{f_n\}$ is a sequence of points in $\mathcal{D}(A)$ such that $f_n \to f$ and $Af_n \to g$ in X as $n \to \infty$, then $f \in \mathcal{D}(A)$ and $Af = g$.

Equivalently, A is closed if its graph

$$\mathcal{G}(A) = \{(f, g) \in X \times X : f \in \mathcal{D}(A) \text{ and } Af = g\}$$

is a closed subset of $X \times X$.

Theorem 5.32. *If A is the generator of a strongly continuous semigroup $\{\mathrm{T}(t)\}$ on a Banach space X, then A is closed and its domain $\mathcal{D}(A)$ is dense in X.*

Example 5.33. If $\mathrm{T}(t)$ is the heat-equation semigroup on L^2, then the L^2-limit

$$\lim_{h \to 0^+} \left[\frac{\mathrm{T}(h)f - f}{h}\right]$$

exists if and only if $f \in H^2$, and then it is equal to Δf. The generator of the heat equation semigroup on L^2 is therefore the unbounded Laplacian operator with domain H^2,

$$\Delta : H^2(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n).$$

If $f_n \to f$ in L^2 and $\Delta f_n \to g$ in L^2, then the continuity of distributional derivatives implies that $\Delta f = g$ and elliptic regularity theory (or the explicit Fourier representation) implies that $f \in H^2$. Thus, the Laplacian with domain $H^2(\mathbb{R}^n)$ is a closed operator in $L^2(\mathbb{R}^n)$. It is also self-adjoint.

Not every closed, densely defined operator generates a semigroup: the powers of its resolvent must satisfy suitable estimates.

Definition 5.34. Suppose that $A : \mathcal{D}(A) \subset X \to X$ is a closed linear operator in a Banach space X and $\mathcal{D}(A)$ is dense in X. A complex number $\lambda \in \mathbb{C}$ is in the resolvent set $\rho(A)$ of A if $\lambda I - A : \mathcal{D}(A) \subset X \to X$ is one-to-one and onto. If $\lambda \in \rho(A)$, the inverse

$$R(\lambda, A) = (\lambda I - A)^{-1} : X \to X \tag{5.23}$$

is called the resolvent of A.

The open mapping (or closed graph) theorem implies that if A is closed, then the resolvent $R(\lambda, A)$ is a bounded linear operator on X whenever it is defined. This is because $(f, Af) \mapsto \lambda f - Af$ is a one-to-one, onto map from the graph $\mathcal{G}(A)$ of A to X, and $\mathcal{G}(A)$ is a Banach space since it is a closed subset of the Banach space $X \times X$.

The resolvent of an operator A may be interpreted as the Laplace transform of the corresponding semigroup. Formally, if

$$\tilde{u}(\lambda) = \int_0^\infty u(t) e^{-\lambda t}\, dt$$

is the Laplace transform of $u(t)$, then taking the Laplace transform with respect to t of the equation

$$u_t = Au \qquad u(0) = f,$$

we get

$$\lambda \tilde{u} - f = A\tilde{u}.$$

For $\lambda \in \rho(A)$, the solution of this equation is

$$\tilde{u}(\lambda) = R(\lambda, A)f.$$

This solution is the Laplace transform of the time-domain solution

$$u(t) = \mathrm{T}(t)f$$

with $R(\lambda, A) = \widetilde{\mathrm{T}(t)}$, or

$$(\lambda I - A)^{-1} = \int_0^\infty e^{-\lambda t} e^{tA}\, dt.$$

This identity can be given a rigorous sense for the generators A of a semigroup, and it explains the connection between semigroups and resolvents. The Hille-Yoshida theorem provides a necessary and sufficient condition on the resolvents for an operator to generate a strongly continuous semigroup.

THEOREM 5.35. *A linear operator $A : \mathcal{D}(A) \subset X \to X$ in a Banach space X is the generator of a strongly continuous semigroup $\{\mathrm{T}(t); t \geq 0\}$ on X if and only if there exist constants $M \geq 1$ and $a \in \mathbb{R}$ such that the following conditions are satisfied:*

(1) *the domain $\mathcal{D}(A)$ is dense in X and A is closed;*
(2) *every $\lambda \in \mathbb{R}$ such that $\lambda > a$ belongs to the resolvent set of A;*
(3) *if $\lambda > a$ and $n \in \mathbb{N}$, then*

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\lambda - a)^n} \tag{5.24}$$

where the resolvent $R(\lambda, A)$ is defined in (5.23).

In that case,

$$\|\mathrm{T}(t)\| \leq M e^{at} \qquad \textit{for all } t \geq 0. \tag{5.25}$$

This theorem is often not useful in practice because the condition on arbitrary powers of the resolvent is difficult to check. For contraction semigroups, we have the following simpler version.

COROLLARY 5.36. *A linear operator $A : \mathcal{D}(A) \subset X \to X$ in a Banach space X is the generator of a strongly continuous contraction semigroup $\{\mathrm{T}(t); t \geq 0\}$ on X if and only if:*

(1) *the domain $\mathcal{D}(A)$ is dense in X and A is closed;*
(2) *every $\lambda \in \mathbb{R}$ such that $\lambda > 0$ belongs to the resolvent set of A;*
(3) *if $\lambda > 0$, then*

$$\|R(\lambda, A)\| \leq \frac{1}{\lambda}. \tag{5.26}$$

This theorem follows from the previous one since

$$\|R(\lambda, A)^n\| \leq \|R(\lambda, A)\|^n \leq \frac{1}{\lambda^n}.$$

The crucial condition here is that $M = 1$. We can always normalize $a = 0$, since if A satisfies Theorem 5.35 with $a = \alpha$, then $A - \alpha I$ satisfies Theorem 5.35 with $a =$

0. Correspondingly, the substitution $u = e^{\alpha t} v$ transforms the evolution equation $u_t = Au$ to $v_t = (A - \alpha I)v$.

The Lumer-Phillips theorem provides a more easily checked condition (that A is 'm-dissipative') for A to generate a contraction semigroup. This condition often follows for PDEs from a suitable energy estimate.

DEFINITION 5.37. A closed, densely defined operator $A : \mathcal{D}(A) \subset X \to X$ in a Banach space X is dissipative if for every $\lambda > 0$

$$\lambda \|f\| \leq \|(\lambda I - A) f\| \qquad \text{for all } f \in \mathcal{D}(A). \tag{5.27}$$

The operator A is maximally dissipative, or m-dissipative for short, if it is dissipative and the range of $\lambda I - A$ is equal to X for some $\lambda > 0$.

The estimate (5.27) implies immediately that $\lambda I - A$ is one-to-one. It also implies that the range of $\lambda I - A : \mathcal{D}(A) \subset X \to X$ is closed. To see this, suppose that g_n belongs to the range of $\lambda I - A$ and $g_n \to g$ in X. If $g_n = (\lambda I - A) f_n$, then (5.27) implies that $\{f_n\}$ is Cauchy since $\{g_n\}$ is Cauchy, and therefore $f_n \to f$ for some $f \in X$. Since A is closed, it follows that $f \in \mathcal{D}(A)$ and $(\lambda I - A)f = g$. Hence, g belongs to the range of $\lambda I - A$.

The range of $\lambda I - A$ may be a proper closed subspace of X for every $\lambda > 0$; if, however, A is m-dissipative, so that $\lambda I - A$ is onto X for some $\lambda > 0$, then one can prove that $\lambda I - A$ is onto for every $\lambda > 0$, meaning that the resolvent set of A contains the positive real axis $\{\lambda > 0\}$. The estimate (5.27) is then equivalent to (5.26). We therefore get the following result, called the Lumer-Phillips theorem.

THEOREM 5.38. *An operator $A : \mathcal{D}(A) \subset X \to X$ in a Banach space X is the generator of a contraction semigroup on X if and only if:*

(1) *A is closed and densely defined;*
(2) *A is m-dissipative.*

EXAMPLE 5.39. Consider $\Delta : H^2(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. If $f \in H^2$, then using the integration-by-parts property of the weak derivative on H^2 we have for $\lambda > 0$ that

$$\begin{aligned}
\|(\lambda I - \Delta) f\|_{L^2}^2 &= \int (\lambda f - \Delta f)^2 \, dx \\
&= \int \left[\lambda^2 f^2 - 2\lambda f \Delta f + (\Delta f)^2\right] dx \\
&= \int \left[\lambda^2 f^2 + 2\lambda Df \cdot Df + (\Delta f)^2\right] dx \\
&\geq \lambda^2 \int f^2 \, dx.
\end{aligned}$$

Hence,

$$\lambda \|f\|_{L^2} \leq \|(\lambda I - \Delta) f\|_{L^2}$$

and Δ is dissipative. The range of $\lambda I - \Delta$ is equal to L^2 for any $\lambda > 0$, as one can see by use of the Fourier transform (in fact, $I - \Delta$ is an isometry of H^2 onto L^2). Thus, Δ is m-dissipative. The Lumer-Phillips theorem therefore implies that $\Delta : H^2 \subset L^2 \to L^2$ generates a strongly continuous semigroup on $L^2(\mathbb{R}^n)$, as we have seen explicitly by use of the Fourier transform.

Thus, in order to show that an evolution equation

$$u_t = Au \qquad u(0) = f$$

in a Banach space X generates a strongly continuous contraction semigroup, it is sufficient to check that $A : \mathcal{D}(A) \subset X \to X$ is a closed, densely defined, dissipative operator and that for some $\lambda > 0$ the resolvent equation

$$\lambda f - Af = g$$

has a solution $f \in X$ for every $g \in X$.

EXAMPLE 5.40. The linearized Kuramoto-Sivashinsky (KS) equation is

$$u_t = -\Delta u - \Delta^2 u.$$

This equation models a system with long-wave instability, described by the backward heat-equation term $-\Delta u$, and short wave stability, described by the forth-order diffusive term $-\Delta^2 u$. The operator

$$A : H^4(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n), \qquad Au = -\Delta u - \Delta^2 u$$

generates a strongly continuous semigroup on $L^2(\mathbb{R}^n)$, or $H^s(\mathbb{R}^n)$. One can verify this directly from the Fourier representation,

$$\widehat{[e^{tA} f]}(k) = e^{t(|k|^2 - |k|^4)} \hat{f}(k),$$

but let us check the hypotheses of the Lumer-Phillips theorem instead. Note that

$$|k|^2 - |k|^4 \leq \frac{3}{16} \qquad \text{for all } |k| \geq 0. \tag{5.28}$$

We claim that $\tilde{A} = A - \alpha I$ is m-dissipative for $\alpha \leq 3/16$. First, $\tilde{A}$ is densely defined and closed, since if $f_n \in H^4$ and $f_n \to f$, $\tilde{A} f_n \to g$ in L^2, the Fourier representation implies that $f \in H^4$ and $\tilde{A} f = g$. If $f \in H^4$, then using (5.28), we have

$$\begin{aligned}
\left\| \lambda f - \tilde{A} f \right\|^2 &= \int_{\mathbb{R}^n} \left(\lambda + \alpha - |k|^2 + |k|^4 \right)^2 \left| \hat{f}(k) \right|^2 \, dk \\
&\geq \lambda \int_{\mathbb{R}^n} \left| \hat{f}(k) \right|^2 \, dk \\
&\geq \lambda \|f\|_{L^2}^2,
\end{aligned}$$

which means that $\tilde{A}$ is dissipative. Moreover, $\lambda I - \tilde{A} : H^4 \to L^2$ is one-to-one and onto for any $\lambda > 0$, since $(\lambda I - \tilde{A}) f = g$ if and only if

$$\hat{f}(k) = \frac{\hat{g}(k)}{\lambda + \alpha - |k|^2 + |k|^4}.$$

Thus, $\tilde{A}$ is m-dissipative, so it generates a contraction semigroup on L^2. It follows that A generates a semigroup on $L^2(\mathbb{R}^n)$ such that

$$\left\| e^{tA} \right\|_{\mathcal{L}(L^2)} \leq e^{3t/16},$$

corresponding to $M = 1$ and $a = 3/16$ in (5.25).

Finally, we state Stone's theorem, which gives an equivalence between self-adjoint operators acting in a Hilbert space and strongly continuous unitary groups. Before stating the theorem, we give the definition of an unbounded self-adjoint operator. For definiteness, we consider complex Hilbert spaces.

DEFINITION 5.41. Let $\mathcal{H}$ be a complex Hilbert space with inner-product

$$(\cdot,\cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}.$$

An operator $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ is self-adjoint if:

(1) the domain $\mathcal{D}(A)$ is dense in $\mathcal{H}$;
(2) $x \in \mathcal{D}(A)$ if and only if there exists $z \in \mathcal{H}$ such that $(x, Ay) = (z, y)$ for every $y \in \mathcal{D}(A)$;
(3) $(x, Ay) = (Ax, y)$ for all $x, y \in \mathcal{D}(\mathcal{A})$.

Condition (2) states that $\mathcal{D}(A) = \mathcal{D}(A^*)$ where A^* is the Hilbert space adjoint of A, in which case $z = Ax$, while (3) states that A is symmetric on its domain. A precise characterization of the domain of a self-adjoint operator is essential; for differential operators acting in L^p-spaces, the domain can often be described by the use of Sobolev spaces. The next result is Stone's theorem (see *e.g.* [**44**] for a proof).

THEOREM 5.42. *An operator $iA : \mathcal{D}(iA) \subset \mathcal{H} \to \mathcal{H}$ in a complex Hilbert space $\mathcal{H}$ is the generator of a strongly continuous unitary group on $\mathcal{H}$ if and only if A is self-adjoint.*

EXAMPLE 5.43. The generator of the Schrödinger group on $H^s(\mathbb{R}^n)$ is the self-adjoint operator

$$i\Delta : \mathcal{D}(i\Delta) \subset H^s(\mathbb{R}^n) \to H^s(\mathbb{R}^n), \qquad \mathcal{D}(i\Delta) = H^{s+2}(\mathbb{R}^n).$$

EXAMPLE 5.44. Consider the Klein-Gordon equation

$$u_{tt} - \Delta u + u = 0$$

in $\mathbb{R}^n$. We rewrite this as a first-order system

$$u_t = v, \qquad v_t = \Delta u,$$

which has the form $w_t = Aw$ where

$$w = \begin{pmatrix} u \\ v \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & I \\ \Delta - I & 0 \end{pmatrix}.$$

We let

$$\mathcal{H} = H^1(\mathbb{R}^n) \oplus L^2(\mathbb{R}^n)$$

with the inner product of $w_1 = (u_1, v_1)$, $w_2 = (u_2, v_2)$ defined by

$$(w_1, w_2)_{\mathcal{H}} = (u_1, u_2)_{H^1} + (v_1, v_2)_{L^2}, \quad (u_1, u_2)_{H^1} = \int (u_1 u_2 + Du_1 \cdot Du_2)\, dx.$$

Then the operator

$$A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}, \qquad \mathcal{D}(A) = H^2(\mathbb{R}^n) \oplus H^1(\mathbb{R}^n)$$

is self-adjoint and generates a unitary group on $\mathcal{H}$.

We can instead take

$$\mathcal{H} = L^2(\mathbb{R}^n) \oplus H^{-1}(\mathbb{R}^n), \qquad \mathcal{D}(A) = H^1(\mathbb{R}^n) \oplus L^2(\mathbb{R}^n)$$

and get a unitary group on this larger space.

5.4.5. Nonhomogeneous equations. The solution of a linear nonhomogeneous ODE

$$u_t = Au + g, \qquad u(0) = f \tag{5.29}$$

may be expressed in terms of the solution operators of the homogeneous equation by the variation of parameters, or Duhamel, formula.

THEOREM 5.45. *Suppose that $A : X \to X$ is a bounded linear operator on a Banach space X and $\mathrm{T}(t) = e^{tA}$ is the associated uniformly continuous group. If $f \in X$ and $g \in C(\mathbb{R}; X)$, then the solution $u \in C^1(\mathbb{R}; X)$ of (5.29) is given by*

$$u(t) = \mathrm{T}(t)f + \int_0^t \mathrm{T}(t-s)g(s)\, ds. \tag{5.30}$$

This solution is continuously strongly differentiable and satisfies the ODE (5.29) pointwise in t for every $t \in \mathbb{R}$. We refer to such a solution as a classical solution. For a strongly continuous group with an unbounded generator, however, the Duhamel formula (5.30) need not define a function $u(t)$ that is differentiable at any time t even if $g \in C(\mathbb{R}; X)$.

EXAMPLE 5.46. Let $\{\mathrm{T}(t) : t \in \mathbb{R}\}$ be a strongly continuous group on a Banach space X with generator $A : \mathcal{D}(A) \subset X \to X$, and suppose that there exists $g_0 \in X$ such that $\mathrm{T}(t)g_0 \notin \mathcal{D}(A)$ for every $t \in \mathbb{R}$. For example, if $\mathrm{T}(t) = e^{it\Delta}$ is the Schrödinger group on $L^2(\mathbb{R}^n)$ and $g_0 \notin H^2(\mathbb{R}^n)$, then $\mathrm{T}(t)g_0 \notin H^2(\mathbb{R}^n)$ for every $t \in \mathbb{R}$. Taking $g(t) = \mathrm{T}(t)g_0$ and $f = 0$ in (5.30) and using the semigroup property, we get

$$u(t) = \int_0^t \mathrm{T}(t-s)\mathrm{T}(s)g_0\, ds = \int_0^t \mathrm{T}(t)g_0\, ds = t\mathrm{T}(t)g_0.$$

This function is continuous but not differentiable with respect to t, since $\mathrm{T}(t)f$ is differentiable at t_0 if and only if $\mathrm{T}(t_0)f \in \mathcal{D}(A)$.

It may happen that the function $u(t)$ defined in (5.30) is is differentiable with respect to t in a distributional sense and satisfies (5.29) pointwise almost everywhere in time. We therefore introduce two other notions of solution that are weaker than that of a classical solution.

DEFINITION 5.47. Suppose that A be the generator of a strongly continuous semigroup $\{\mathrm{T}(t) : t \geq 0\}$, $f \in X$ and $g \in L^1([0,T]; X)$. A function $u : [0,T] \to X$ is a strong solution of (5.29) on $[0,T]$ if:

1. u is absolutely continuous on $[0,T]$ with distributional derivative $u_t \in L^1(0,T; X)$;
2. $u(t) \in \mathcal{D}(A)$ pointwise almost everywhere for $t \in (0,T)$;
3. $u_t(t) = Au(t) + g(t)$ pointwise almost everywhere for $t \in (0,T)$;
4. $u(0) = f$.

A function $u : [0,T] \to X$ is a mild solution of (5.29) on $[0,T]$ if u is given by (5.30) for $t \in [0,T]$.

Every classical solution is a strong solution and every strong solution is a mild solution. As Example 5.46 shows, however, a mild solution need not be a strong solution.

The Duhamel formula provides a useful way to study semilinear evolution equations of the form

$$u_t = Au + g(u) \tag{5.31}$$

where the linear operator A generates a semigroup on a Banach space X and

$$g : \mathcal{D}(F) \subset X \to X$$

is a nonlinear function. For semilinear PDEs, $g(u)$ typically depends on u but none of its spatial derivatives and then (5.31) consists of a linear PDE perturbed by a zeroth-order nonlinear term.

If $\{\mathrm{T}(t)\}$ is the semigroup generated by A, we may replace (5.31) by an integral equation for $u : [0, T] \to X$

$$u(t) = \mathrm{T}(t)u(0) + \int_0^t \mathrm{T}(t-s) g\left(u(s)\right)\, ds. \tag{5.32}$$

We then try to show that solutions of this integral equation exist. If these solutions have sufficient regularity, then they also satisfy (5.31).

In the standard Picard approach to ODEs, we would write (5.31) as

$$u(t) = u(0) + \int_0^t \left[Au(s) + g\left(u(s)\right)\right]\, ds. \tag{5.33}$$

The advantage of (5.32) over (5.33) is that we have replaced the unbounded operator A by the bounded solution operators $\{\mathrm{T}(t)\}$. Moreover, since $\mathrm{T}(t-s)$ acts on $g(u(s))$ it is possible for the regularizing properties of the linear operators T to compensate for the destabilizing effects of the nonlinearity F. For example, in Section 5.5 we study a semilinear heat equation, and in Section 5.6 to prove the existence of solutions of a nonlinear Schrödinger equation.

5.4.6. Non-autonomous equations. The semigroup property $\mathrm{T}(s)\mathrm{T}(t) = \mathrm{T}(s+t)$ holds for autonomous evolution equations that do not depend explicitly on time. One can also consider time-dependent linear evolution equations in a Banach space X of the form

$$u_t = A(t)u$$

where $A(t) : \mathcal{D}\left(A(t)\right) \subset X \to X$. The solution operators $\mathrm{T}(t; s)$ from time s to time t of a well-posed nonautonomous equation depend separately on the initial and final times, not just on the time difference; they satisfy

$$\mathrm{T}(t; s)\mathrm{T}(s; r) = \mathrm{T}(t; r) \qquad \text{for } r \leq s \leq t.$$

The time-dependence of A makes such equations more difficult to analyze from the semigroup viewpoint than autonomous equations. First, since the domain of $A(t)$ depends in general on t, one must understand how these domains are related and for what times a solution belongs to the domain. Second, the operators $A(s)$, $A(t)$ may not commute for $s \neq t$, meaning that one must order them correctly with respect to time when constructing solution operators $\mathrm{T}(t; s)$.

Similar issues arise in using semigroup theory to study quasi-linear evolution equations of the form

$$u_t = A(u)u$$

in which, for example, $A(u)$ is a differential operator acting on u whose coefficients depend on u (see *e.g.* [**44**] for further discussion). Thus, while semigroup theory is an effective approach to the analysis of autonomous semilinear problems, its application to nonautonomous or quasilinear problems often leads to considerable technical difficulties.

5.5. A semilinear heat equation

Consider the following initial value problem for $u : \mathbb{R}^n \times [0, T] \to \mathbb{R}$:

$$u_t = \Delta u + \lambda u - \gamma u^m, \qquad u(x, 0) = g(x) \tag{5.34}$$

where $\lambda, \gamma \in \mathbb{R}$ and $m \in \mathbb{N}$ are parameters. This PDE is a scalar, semilinear reaction diffusion equation. The solution $u = 0$ is linearly stable when $\lambda < 0$ and linearly unstable when $\lambda > 0$. The nonlinear reaction term is potentially stabilizing if $\gamma > 0$ and m is odd or m is even and solutions are nonnegative (they remain nonegative by the maximum principle). For example, if $m = 3$ and $\gamma > 0$, then the spatially-independent reaction ODE $u_t = \lambda u - \gamma u^3$ has a supercritical pitchfork bifurcation at $u = 0$ as λ passes through 0. Thus, (5.34) provides a model equation for the study of bifurcation and loss of stability of equilbria in PDEs.

We consider (5.34) on $\mathbb{R}^n$ since this allows us to apply the results obtained earlier in the Chapter for the heat equation on $\mathbb{R}^n$. In some respects, the behavior this IBVP on a bounded domain is simpler to analyze. The negative Laplacian on $\mathbb{R}^n$ does not have a compact resolvent and has a purely continuous spectrum $[0, \infty)$. By contrast, negative Laplacian on a bounded domain, with say homogeneous Dirichlet boundary conditions, has compact resolvent and a discrete set of eigenvalues $\lambda_1 < \lambda_2 \le \lambda_3 \le \dots$. As a result, only finitely many modes become unstable as λ increases, and the long time dynamics of (5.34) is essentially finite-dimensional in nature.

Equations of the form

$$u_t = \Delta u + f(u)$$

on a bounded one-dimensional domain were studied by Chafee and Infante (1974), so this equation is sometimes called the Chafee-Infante equation. We consider here the special case with

$$f(u) = \lambda u - \gamma u^m \tag{5.35}$$

so that we can focus on the essential ideas. We do not attempt to obtain an optimal result; our aim is simply to illustrate how one can use semigroup theory to prove the existence of solutions of semilinear parabolic equations such as (5.34). Moreover, semigroup theory is not the only possible approach to such problems. For example, one can also use a Galerkin method.

5.5.1. Motivation.

5.5.1. Motivation. We will use the linear heat equation semigroup to reformulate (5.34) as a nonlinear integral equation in an appropriate function space and apply a contraction mapping argument.

To motivate the following analysis, we proceed formally at first. Suppose that $A = -\Delta$ generates a semigroup e^{-tA} on some space X, and let F be the nonlinear operator $F(u) = f(u)$, meaning that F is composition with f regarded as an operator on functions. Then (5.34) maybe written as the abstract evolution equation

$$u_t = -Au + F(u), \qquad u(0) = g.$$

Using Duhamel's formula, we get

$$u(t) = e^{-tA}g + \int_0^t e^{-(t-s)A} F\left(u(s)\right) \, ds.$$

We use this integral equation to *define* mild solutions of the equation.

We want to formulate the integral equation as a fixed point problem $u = \Phi(u)$ on a space of Y-valued functions $u : [0,T] \to Y$. There are many ways to achieve this. In the framework we use here, we choose spaces $Y \subset X$ such that: (a) $F : Y \to X$ is locally Lipschitz continuous; (b) $e^{-tA} : X \to Y$ for $t > 0$ with integrable operator norm as $t \to 0^+$. This allows the smoothing of the semigroup to compensate for a loss of regularity in the nonlinearity.

As we will show, one appropriate choice in $1 \le n \le 3$ space dimensions is $X = L^2(\mathbb{R}^n)$ and $Y = H^{2\alpha}(\mathbb{R}^n)$ for $n/4 < \alpha < 1$. Here $H^{2\alpha}(\mathbb{R}^n)$ is the L^2-Sobolev space of fractional order 2α defined in Section 5.C. We write the order of the Sobolev space as 2α because $H^{2\alpha}(\mathbb{R}^n) = \mathcal{D}\left(A^\alpha\right)$ is the domain of the αth-power of the generator of the semigroup.

5.5.2. Mild solutions. Let A denote the negative Laplacian operator in L^2,

$$A : \mathcal{D}(A) \subset L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n), \qquad A = -\Delta, \qquad \mathcal{D}(A) = H^2(\mathbb{R}^n). \tag{5.36}$$

We define A as an operator acting in L^2 because we can study it explicitly by use of the Fourier transform.

As discussed in Section 5.4.2, A is a closed, densely defined positive operator, and $-A$ is the generator of a strongly continuous contraction semigroup

$$\{e^{-tA} : t \ge 0\}$$

on $L^2(\mathbb{R}^n)$. The Fourier representation of the semigroup operators is

$$e^{-tA} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n), \qquad (\widehat{e^{-tA}h})(k) = e^{-t|k|^2}\hat{h}(k). \tag{5.37}$$

If $t > 0$ we have for any $\alpha > 0$ that

$$e^{-tA} : L^2(\mathbb{R}^n) \to H^{2\alpha}(\mathbb{R}^n).$$

This property expresses the instantaneous smoothing of solutions of the heat equation *c.f.* Proposition 5.14.

We define the nonlinear operator

$$F : H^{2\alpha}(\mathbb{R}^n) \to L^2(\mathbb{R}^n), \qquad F(h)(x) = \lambda h(x) - \gamma h^m(x). \tag{5.38}$$

In order to ensure that F takes values in L^2 and has good continuity properties, we assume that $\alpha > n/4$. The Sobolev embedding theorem (Theorem 5.79) implies that $H^{2\alpha}(\mathbb{R}^n) \hookrightarrow C_0(\mathbb{R}^n)$. Hence, if $h \in H^{2\alpha}$, then $h \in L^2 \cap C_0$, so $h \in L^p$ for every $2 \le p \le \infty$, and $F(h) \in L^2 \cap C_0$. We then define mild $H^{2\alpha}$-valued solutions of (5.34) as follows.

Definition 5.48. Suppose that $T > 0$, $\alpha > n/4$, and $g \in H^{2\alpha}(\mathbb{R}^n)$. A mild $H^{2\alpha}$-valued solution of (5.34) on $[0,T]$ is a function

$$u \in C\left([0,T]; H^{2\alpha}(\mathbb{R}^n)\right)$$

such that

$$u(t) = e^{-tA}g + \int_0^t e^{-(t-s)A}F\left(u(s)\right)\,ds \qquad \text{for every } 0 \le t \le T, \tag{5.39}$$

where e^{-tA} is given by (5.37), and F is given by (5.38).

5.5.3. Existence. In order to prove a local existence result, we choose α large enough that the nonlinear term is well-behaved by Sobolev embedding, but small enough that the norm of the semigroup maps from L^2 into $H^{2\alpha}$ is integrable as $t \to 0^+$. As we will see, this is the case if $n/4 < \alpha < 1$, so we restrict attention to $1 \le n \le 3$ space dimensions.

THEOREM 5.49. *Suppose that $1 \le n \le 3$ and $n/4 < \alpha < 1$. Then there exists $T > 0$, depending only on α, n, $\|g\|_{H^{2\alpha}}$, and the coefficients of f, such that (5.34) has a unique mild solution $u \in C\left([0,T]; H^{2\alpha}\right)$ in the sense of Definition 5.48.*

PROOF. We write (5.39) as

$$
\begin{aligned}
&u = \Phi(u),\\
&\Phi : C\left([0,T]; H^{2\alpha}\right) \to C\left([0,T]; H^{2\alpha}\right),\\
&\Phi(u)(t) = e^{-tA}g + \int_0^t e^{-(t-s)A} F\left(u(s)\right)\, ds.
\end{aligned}
\tag{5.40}
$$

We will show that Φ defined in (5.40) is a contraction mapping on a suitable ball in $C\left([0,T]; H^{2\alpha}\right)$. We do this in a series of Lemmas. The first Lemma is an estimate of the norm of the semigroup operators on the domain of a fractional power of the generator.

LEMMA 5.50. *Let e^{-tA} be the semigroup operator defined in (5.37) and $\alpha > 0$. If $t > 0$, then*

$$e^{-tA} : L^2(\mathbb{R}^n) \to H^{2\alpha}(\mathbb{R}^n)$$

and there is a constant $C = C(\alpha, n)$ such that

$$\left\|e^{-tA}\right\|_{\mathcal{L}(L^2, H^{2\alpha})} \le \frac{Ce^t}{t^\alpha}.$$

PROOF. Suppose that $h \in L^2(\mathbb{R}^n)$. Using the Fourier representation (5.37) of e^{-tA} as multiplication by $e^{-t|k|^2}$ and the definition of the $H^{2\alpha}$-norm, we get that

$$
\begin{aligned}
\left\|e^{-tA}h\right\|_{H^{2\alpha}}^2 &= (2\pi)^n \int_{\mathbb{R}^n} \left(1+|k|^2\right)^{2\alpha} e^{-2t|k|^2} \left|\hat{h}(k)\right|^2\, dk\\
&\le (2\pi)^n \sup_{k\in\mathbb{R}^n} \left[\left(1+|k|^2\right)^{2\alpha} e^{-2t|k|^2}\right] \int_{\mathbb{R}^n} \left|\hat{h}(k)\right|^2\, dk.
\end{aligned}
$$

Hence, by Parseval's theorem,

$$\left\|e^{-tA}h\right\|_{H^{2\alpha}} \le M\|h\|_{L^2}$$

where

$$M = (2\pi)^{n/2} \sup_{k\in\mathbb{R}^n} \left[\left(1+|k|^2\right)^{2\alpha} e^{-2t|k|^2}\right]^{1/2}.$$

Writing $1 + |k|^2 = x$, we have

$$M = (2\pi)^{n/2} e^t \sup_{x\ge 1} \left[x^\alpha e^{-tx}\right] \le \frac{Ce^t}{t^\alpha}.$$

and the result follows. □

Next, we show that Φ is a locally Lipschitz continuous map on the space $C\left([0,T]; H^{2\alpha}(\mathbb{R}^n)\right)$.

LEMMA 5.51. *Suppose that $\alpha > n/4$. Let Φ be the map defined in (5.40) where F is given by (5.38), A is given by (5.36) and $g \in H^{2\alpha}(\mathbb{R}^n)$. Then*

$$\Phi : C\left([0,T]; H^{2\alpha}(\mathbb{R}^n)\right) \to C\left([0,T]; H^{2\alpha}(\mathbb{R}^n)\right) \tag{5.41}$$

and there exists a constant $C = C(\alpha, m, n)$ such that

$$\begin{aligned} &\|\Phi(u) - \Phi(v)\|_{C([0,T];H^{2\alpha})} \\ &\leq CT^{1-\alpha}\left(1 + \|u\|^{m-1}_{C([0,T];H^{2\alpha})} + \|v\|^{m-1}_{C([0,T];H^{2\alpha})}\right)\|u - v\|_{C([0,T];H^{2\alpha})} \end{aligned}$$

for every $u, v \in C\left([0,T]; H^{2\alpha}\right)$.

PROOF. We write Φ in (5.40) as

$$\Phi(u)(t) = e^{-tA}g + \Psi(u)(t), \qquad \Psi(u)(t) = \int_0^t e^{-(t-s)A}F\left(u(s)\right)\,ds.$$

Since $g \in H^{2\alpha}$ and $\{e^{-tA} : t \geq 0\}$ is a strongly continuous semigroup on $H^{2\alpha}$, the map $t \mapsto e^{-tA}g$ belongs to $C\left([0,T]; H^{2\alpha}\right)$. Thus, we only need to prove the result for Ψ.

The fact that $\Psi(u) \in C\left([0,T]; H^{2\alpha}\right)$ if $u \in C\left([0,T]; H^{2\alpha}\right)$ follows from the Lipschitz continuity of Ψ and a density argument. Thus, we only need to prove the Lipschitz estimate.

If $u, v \in C\left([0,T]; H^{2\alpha}\right)$, then using Lemma 5.50 we find that

$$\begin{aligned} \|\Psi(u)(t) - \Psi(v)(t)\|_{H^{2\alpha}} &\leq C\int_0^t \frac{e^{(t-s)}}{|t-s|^\alpha}\|F\left(u(s)\right) - F\left(v(s)\right)\|_{L^2}\,ds \\ &\leq C\sup_{0\leq s\leq T}\|F\left(u(s)\right) - F\left(v(s)\right)\|_{L^2}\int_0^t \frac{1}{|t-s|^\alpha}\,ds. \end{aligned}$$

Evaluating the s-integral, with $\alpha < 1$, and taking the supremum of the result over $0 \leq t \leq T$, we get

$$\|\Psi(u) - \Psi(v)\|_{L^\infty(0,T;H^{2\alpha})} \leq CT^{1-\alpha}\|F\left(u\right) - F\left(v\right)\|_{L^\infty(0,T;L^2)}. \tag{5.42}$$

From (5.35), if $g, h \in C_0 \subset H^{2\alpha}$ we have

$$\|F\left(g\right) - F\left(h\right)\|_{L^2} \leq |\lambda|\,\|g - h\|_{L^2} + |\gamma|\,\|g^m - h^m\|_{L^2}$$

and

$$\|g^m - h^m\|_{L^2} \leq C\left(\|g\|^{m-1}_{L^\infty} + \|h\|^{m-1}_{L^\infty}\right)\|g - h\|_{L^2}.$$

Hence, using the Sobolev inequality $\|g\|_{L^\infty} \leq C\|g\|_{H^{2\alpha}}$ for $\alpha > n/4$ and the fact that $\|g\|_{L^2} \leq \|g\|_{H^{2\alpha}}$, we get that

$$\|F(g) - F(h)\|_{L^2} \leq C\left(1 + \|g\|^{m-1}_{H^{2\alpha}} + \|h\|^{m-1}_{H^{2\alpha}}\right)\|g - h\|_{H^{2\alpha}},$$

which means that $F : H^{2\alpha} \to L^2$ is locally Lipschitz continuous.[3] The use of this result in (5.42) proves the Lemma. □

[3]Actually, under the assumptions we make here, $F : H^{2\alpha} \to H^{2\alpha}$ is locally Lipschitz continuous as a map from $H^{2\alpha}$ into itself, and we don't need to use the smoothing properties of the heat equation semigroup to obtain a fixed point problem in $C([0,T]; H^{2\alpha})$, so perhaps this wasn't the best example to choose! For stronger nonlinearities, however, it would be necessary to use the smoothing.

The existence theorem now follows by a standard contraction mapping argument. If $\|g\|_{H^{2\alpha}} = R$, then

$$\left\|e^{-tA}g\right\|_{H^{2\alpha}} \leq R \qquad \text{for every } 0 \leq t \leq T$$

since $\{e^{-tA}\}$ is a contraction semigroup on $H^{2\alpha}$. Therefore, if we choose

$$E = \left\{u \in C([0,T]; H^{2\alpha} : \|u\|_{C([0,T];H^{2\alpha})} \leq 2R\right\}$$

we see from Lemma 5.51 that $\Phi : E \to E$ if we choose $T > 0$ such that

$$CT^{1-\alpha}\left(1 + 2R^{m-1}\right) = \theta R$$

where $0 < \theta < 1$. Moreover, in that case

$$\|\Phi(u) - \Phi(v)\|_{C([0,T];H^{2\alpha})} \leq \theta \|u - v\|_{C([0,T];H^{2\alpha})} \qquad \text{for every } u, v \in E.$$

The contraction mapping theorem then implies the existence of a unique solution $u \in E$. □

This result can be extended and improved in many directions. In particular, if A is the negative Laplacian acting in $L^p(\mathbb{R}^n)$,

$$A : W^{2,p}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n), \qquad A = -\Delta.$$

then one can prove that $-A$ is the generator of a strongly continuous semigroup on L^p for every $1 < p < \infty$. Moreover, we can define fractional powers of A

$$A^\alpha : \mathcal{D}(A^\alpha) \subset L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n).$$

If we choose $2p > n$ and $n/2p < \alpha < 1$, then Sobolev embedding implies that $\mathcal{D}(A^\alpha) \hookrightarrow C_0$ and the same argument as the one above applies. This gives the existence of local mild solutions with values in $\mathcal{D}(A^\alpha)$ in any number of space dimensions. The proof of the necessary estimates and embedding theorems is more involved that the proofs above if $p \neq 2$, since we cannot use the Fourier transform to obtain out explicit solutions.

More generally, this local existence proof extends to evolution equations of the form ([**41**], §15.1)

$$u_t + Au = F(u),$$

where we look for mild solutions $u \in C([0,T]; X)$ taking values in a Banach space X and there is a second Banach spaces Y such that:

(1) $e^{-tA} : X \to X$ is a strongly continuous semigroup for $t \geq 0$;
(2) $F : X \to Y$ is locally Lipschitz continuous;
(3) $e^{-tA} : Y \to X$ for $t > 0$ and for some $\alpha < 1$

$$\left\|e^{-tA}\right\|_{\mathcal{L}(X,Y)} \leq \frac{C}{t^\alpha} \qquad \text{for } 0 < t \leq T.$$

In the above example, we used $X = H^{2\alpha}$ and $Y = L^2$. If A is a sectorial operator that generates an analytic semigroup on Y, then one can define fractional powers A^α of A, and the semigroup $\{e^{-tA}\}$ satisfies the above properties with $X = \mathcal{D}(A^\alpha)$ for $0 \leq \alpha < 1$ [**36**]. Thus, one gets a local existence result provided that $F : \mathcal{D}(A^\alpha) \to L^2$ is locally Lipschitz, with an existence-time that depends on the X-norm of the initial data.

In general, the X-norm of the solution may blow up in finite time, and one gets only a local solution. If, however, one has an *a priori* estimate for $\|u(t)\|_X$ that is global in time, then global existence follows from the local existence result.

5.6. The nonlinear Schrödinger equation

The nonlinear Schrödinger (NLS) equation is

$$iu_t = -\Delta u - \lambda |u|^\alpha u \tag{5.43}$$

where $\lambda \in \mathbb{R}$ and $\alpha > 0$ are constants. In many applications, such as the asymptotic description of weakly nonlinear dispersive waves, we get $\alpha = 2$, leading to the cubically nonlinear NLS equation.

A physical interpretation of (5.43) is that it describes the motion of a quantum mechanical particle in a potential $V = -\lambda|u|^\alpha$ which depends on the probability density $|u|^2$ of the particle *c.f.* (5.14). If $\lambda \neq 0$, we can normalize $\lambda = \pm 1$ so the magnitude of λ is not important; the sign of λ is, however, crucial.

If $\lambda > 0$, then the potential becomes large and negative when $|u|^2$ becomes large, so the particle 'digs' its own potential well; this tends to trap the particle and further concentrate is probability density, possibly leading to the formation of singularities in finite time if $n \geq 2$ and $\alpha \geq 4/n$. The resulting equation is called the focusing NLS equation.

If $\lambda < 0$, then the potential becomes large and positive when $|u|^2$ becomes large; this has a repulsive effect and tends to make the probability density spread out. The resulting equation is called the defocusing NLS equation. The local L^2-existence result that we obtain here for subcritical nonlinearities $0 < \alpha < 4/n$ is, however, not sensitive to the sign of λ.

The one-dimensional cubic NLS equation

$$iu_t + u_{xx} + \lambda |u|^2 u = 0$$

is completely integrable. If $\lambda > 0$, this equation has localized traveling wave solutions called solitons in which the effects of nonlinear self-focusing balance the tendency of linear dispersion to spread out the the wave. Moreover, these solitons preserve their identity under nonlinear interactions with other solitons. Such localized solutions exist for the focusing NLS equation in higher dimensions, but the NLS equation is not integrable if $n \geq 2$, and in that case the soliton solutions are not preserved under nonlinear interactions.

In this section, we obtain an existence result for the NLS equation. The linear Schrödinger equation group is not smoothing, so we cannot use it to compensate for the nonlinearity at a fixed time as we did in Section 5.5 for the semilinear equation. Instead, we use some rather delicate space-time estimates for the linear Schrödinger equation, called Strichartz estimates, to recover the powers lost by the nonlinearity. We derive these estimates first.

5.6.1. Strichartz estimates.

The Strichartz estimates for the Schrödinger equation (5.13) may be derived by use of the interpolation estimate in Theorem 5.16 and the Hardy-Littlewood-Sobolev inequality in Theorem 5.77. The space-time norm in the Strichartz estimate is $L^q(\mathbb{R})$ in time and $L^r(\mathbb{R}^n)$ in space for suitable exponents (q, r), which we call an admissible pair.

Definition 5.52. The pair of exponents (q, r) is an admissible pair if

$$\frac{2}{q} = \frac{n}{2} - \frac{n}{r} \tag{5.44}$$

where $2 < q < \infty$ and

$$2 < r < \frac{2n}{n-2} \qquad \text{if } n \geq 3 \tag{5.45}$$

or $2 < r < \infty$ if $n = 1, 2$.

The Strichartz estimates continue to hold for some endpoints with $q = 2$ or $q = \infty$, but we will not consider these cases here.

THEOREM 5.53. *Suppose that $\{\mathrm{T}(t) : t \in \mathbb{R}\}$ is the unitary group of solution operators of the Schrödinger equation on $\mathbb{R}^n$ defined in (5.22) and (q, r) is an admissible pair as in Definition 5.52.*

(1) *For $f \in L^2(\mathbb{R}^n)$, let $u(t) = \mathrm{T}(t)f$. Then $u \in L^q(\mathbb{R}; L^r)$, and there is a constant $C(n, r)$ such that*

$$\|u\|_{L^q(\mathbb{R};L^r)} \leq C\|f\|_{L^2}. \tag{5.46}$$

(2) *For $g \in L^{q'}(\mathbb{R}; L^{r'})$, let*

$$v(t) = \int_{-\infty}^{\infty} \mathrm{T}(t-s)g(s)\, ds.$$

Then $v \in L^{q'}(\mathbb{R}; L^{r'}) \cap C(\mathbb{R}; L^2)$ and there is a constant $C(n, r)$ such that

$$\|v\|_{L^\infty(\mathbb{R};L^2)} \leq C\, \|g\|_{L^{q'}(\mathbb{R};L^{r'})}, \tag{5.47}$$

$$\|v\|_{L^q(\mathbb{R};L^r)} \leq C\, \|g\|_{L^{q'}(\mathbb{R};L^{r'})}. \tag{5.48}$$

PROOF. By a density argument, it is sufficient to prove the result for smooth functions. We therefore assume that $g \in C_c^\infty(\mathbb{R}; \mathcal{S})$ is a smooth Schwartz-valued function with compact support in time and $f \in \mathcal{S}$. We prove the inequalities in reverse order.

Using the interpolation estimate Theorem 5.16, we have for $2 < r < \infty$ that

$$\|v(t)\|_{L^r} \leq \int_{-\infty}^{\infty} \frac{\|g(s)\|_{L^{r'}}}{(4\pi|t-s|)^{n(1/2-1/r)}}\, ds.$$

If r is admissible, then $0 < n(1/2 - 1/r) < 1$. Thus, taking the L^q-norm of this inequality with respect to t and using the Hardy-Littlewood-Sobolev inequality (Theorem 5.77) in the result, we find that

$$\|v\|_{L^q(\mathbb{R};L^r)} \leq C\, \|g\|_{L^p(\mathbb{R};L^{r'})}$$

where p is given by

$$\frac{1}{p} = 1 + \frac{1}{q} + \frac{n}{r} - \frac{n}{2}.$$

If q, r satisfy (5.44), then $p = q'$, and we get (5.48).

Using Fubini's theorem and the unitary group property of $\mathrm{T}(t)$, we have

$$\begin{aligned}
(v(t), v(t))_{L^2(\mathbb{R}^n)} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\mathrm{T}(t-r)g(r), \mathrm{T}(t-s)g(s))_{L^2(\mathbb{R}^n)}\, drds \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\mathrm{T}(s-r)g(r), g(s))_{L^2(\mathbb{R}^n)}\, drds \\
&= \int_{-\infty}^{\infty} (v(s), g(s))_{L^2(\mathbb{R}^n)}\, ds.
\end{aligned}$$

Using Hölder's inequality and (5.48) in this equation, we get

$$\|v(t)\|^2_{L^2(\mathbb{R}^n)} \leq \|v\|_{L^q(\mathbb{R};L^r)} \|g\|_{L^{q'}(\mathbb{R};L^{r'})} \leq C \|g\|^2_{L^{q'}(\mathbb{R};L^{r'})}.$$

Taking the supremum of this inequality over $t \in \mathbb{R}$, we obtain (5.47). In fact, since

$$v(t) = \mathrm{T}(t) \int_{-\infty}^{\infty} \mathrm{T}(-s)g(s)\, ds$$

$v \in C(\mathbb{R}; L^2)$ is an L^2-solution of the homogeneous Schrödinger equation and $\|v(t)\|_{L^2(\mathbb{R}^n)}$ is independent of t.

If $f \in \mathcal{S}$, $u(t) = \mathrm{T}(t)f$, and $g \in C_c^\infty(\mathbb{R}; \mathcal{S})$, then using (5.48) we get

$$\begin{aligned}
\int_{-\infty}^{\infty} (u(t), g(t))_{L^2}\ dt &= \int_{-\infty}^{\infty} (\mathrm{T}(t)f, g(t))_{L^2}\ dt \\
&= \left(f, \int_{-\infty}^{\infty} \mathrm{T}(-t)g(t)\, dt\right)_{L^2} \\
&\leq \|f\|_{L^2} \left\| \int_{-\infty}^{\infty} \mathrm{T}(-t)g(t)\, dt \right\|_{L^2} \\
&\leq C \|f\|_{L^2} \|g\|_{L^{q'}(\mathbb{R};L^{r'})}.
\end{aligned}$$

It then follows by duality and density that

$$\|u\|_{L^q(\mathbb{R};L^r)} = \sup_{g \in C_c^\infty(\mathbb{R};\mathcal{S})} \frac{\int_{-\infty}^{\infty} (u(t), g(t))_{L^2}\ dt}{\|g\|_{L^{q'}(\mathbb{R};L^{r'})}} \leq C \|f\|_{L^2},$$

which proves (5.46). □

This estimate describes a dispersive smoothing effect of the Schrödinger equation. For example, the L^r-spatial norm of the solution may blow up at some time, but it must be finite almost everywhere in t. Intuitively, this is because if the Fourier modes of the solution are sufficiently in phase at some point in space and time that they combine to form a singularity, then dispersion pulls them apart at later times.

Although the above proof of the Schrödinger equation Strichartz estimates is elementary, in the sense that given the interpolation estimate for the Schrödinger equation and the one-dimensional Hardy-Littlewood-Sobolev inequality it uses only Hölder's inequality, it does not explicitly clarify the role of dispersion (beyond the dispersive decay of solutions in time). An alternative point of view is in terms of restriction theorems for the Fourier transform.

The Fourier solution of the Schrödinger equation (5.13) is

$$u(x,t) = \int_{\mathbb{R}^n} \hat{f}(k) e^{ik\cdot x + i|k|^2 t}\, dk.$$

Thus, the space-time Fourier transform $\hat{u}(k,\tau)$ of $u(x,t)$,

$$\hat{u}(k,\tau) = \frac{1}{(2\pi)^{n+1}} \int u(x,t) e^{ik\cdot x + i\tau t}\, dxdt,$$

is a measure supported on the paraboloid $\tau + |k|^2 = 0$. This surface has non-singular curvature, which is a geometrical expression of the dispersive nature of the Schrödinger equation. The Strichartz estimates describe a boundedness property of the restriction of the Fourier transform to curved surfaces.

As an illustration of this phenomenon, we state the Tomas-Stein theorem on the restriction of the Fourier transform in $\mathbb{R}^{n+1}$ to the unit sphere $\mathbb{S}^n$.

THEOREM 5.54. *Suppose that $f \in L^p(\mathbb{R}^{n+1})$ with*

$$1 \leq p \leq \frac{2n+4}{n+4}$$

and let $\hat{g} = \hat{f}\Big|_{\mathbb{S}^n}$. Then there is a constant $C(p,n)$ such that

$$\|\hat{g}\|_{L^2(\mathbb{S}^n)} \leq C\, \|f\|_{L^p(\mathbb{R}^{n+1})}.$$

5.6.2. Local L^2-solutions. In this section, we use the Strichartz estimates for the linear Schrödinger equation to obtain a local existence result for solutions of the nonlinear Schrödinger equation with initial data in L^2.

If X is a Banach space and $T > 0$, we say that $u \in C([0,T];X)$ is a mild X-valued solution of (5.43) if it satisfies the Duhamel-type integral equation

$$u = \mathrm{T}(t)f + i\lambda \int_0^t \mathrm{T}(t-s)\left\{|u|^\alpha(s)u(s)\right\}\, ds \qquad \text{for } t \in [0,T] \tag{5.49}$$

where $\mathrm{T}(t) = e^{it\Delta}$ is the solution operator of the linear Schrödinger equation defined by (5.22). If a solution of (5.49) has sufficient regularity then it is also a solution of (5.43), but here we simply take (5.49) as our definition of a solution. We suppose that $t \geq 0$ for definiteness; the same arguments apply for $t \leq 0$.

Before stating an existence theorem, we explain the idea of the proof, which is based on the contraction mapping theorem. We write (5.49) as a fixed-point equation

$$u = \Phi(u) \qquad \Phi(u)(t) = \mathrm{T}(t)f + i\lambda\Psi(u)(t), \tag{5.50}$$

$$\Psi(u)(t) = \int_0^t \mathrm{T}(t-s)\left\{|u|^\alpha(s)u(s)\right\}\, ds. \tag{5.51}$$

We want to find a Banach space E of functions $u : [0,T] \to L^r$ and a closed ball $B \subset E$ such that $\Phi : B \to B$ is a contraction mapping when $T > 0$ is sufficiently small.

As discussed in Section 5.4.3, the Schrödinger operators $\mathrm{T}(t)$ form a strongly continuous group on L^p only if $p = 2$. Thus if $f \in L^2$, then

$$\Phi : C\left([0,T]; L^{2/(\alpha+1)}\right) \to C\left([0,T]; L^2\right),$$

but Φ does not map the space $C\left([0,T]; L^r\right)$ into itself for any exponent $1 \leq r \leq \infty$.

If α is not too large, however, there are exponents q, r such that

$$\Phi : L^q\left(0,T; L^r\right) \to L^q\left(0,T; L^r\right). \tag{5.52}$$

This happens because, as shown by the Strichartz estimates, the linear solution operator T can regain the space-time regularity lost by the nonlinearity. (For a brief discussion of vector-valued L^p-spaces, see Section 6.A.)

To determine values of q, r for which (5.52) holds, we write

$$L^q\left(0,T; L^r\right) = L^q_t L^r_x$$

for short, and consider the action of Φ defined in (5.50)–(5.51) on such a space.

First, consider the term $\mathrm{T}f$ in (5.50) which is independent of u. Theorem 5.53 implies that $\mathrm{T}f \in L^q_t L^r_x$ if $f \in L^2$ for any admissible pair (q,r).

Second, consider the nonlinear term $\Psi(u)$ in (5.51). We have

$$\begin{aligned}\| |u|^\alpha u \|_{L_t^q L_x^r} &= \left[\int_0^T \left(\int_{\mathbb{R}^n} |u|^{r(\alpha+1)}\, dx\right)^{q/r} dt\right]^{1/q} \\ &= \left[\int_0^T \left(\int_{\mathbb{R}^n} |u|^{r(\alpha+1)}\, dx\right)^{q(\alpha+1)/r(\alpha+1)} dt\right]^{(\alpha+1)/q(\alpha+1)} \\ &= \|u\|^{\alpha+1}_{L_t^{q(\alpha+1)} L_x^{r(\alpha+1)}}.\end{aligned}$$

Thus, if $u \in L_t^{q_1} L_x^{r_1}$ then $|u|^\alpha u \in L_t^{q_2'} L_x^{r_2'}$ where

$$q_1 = q_2'(\alpha+1), \qquad r_1 = r_2'(\alpha+1). \tag{5.53}$$

If (q_2, r_2) is an admissible pair, then the Strichartz estimate (5.48) implies that

$$\Psi(u) \in L_t^{q_2} L_x^{r_2}.$$

In order to ensure that Ψ preserves the L_x^r-norm of u, we need to choose $r = r_1 = r_2$, which implies that $r = r'(\alpha+1)$, or

$$r = \alpha + 2. \tag{5.54}$$

If r is given by (5.54), then it follows from Definition 5.52 that

$$(q_2, r_2) = (q, \alpha + 2)$$

is an admissible pair if

$$q = \frac{4(\alpha+2)}{n\alpha} \tag{5.55}$$

and $0 < \alpha < 4/(n-2)$, or $0 < \alpha < \infty$ if $n = 1, 2$. In that case, we have

$$\Psi : L_t^{q_1} L_x^{\alpha+2} \to L_t^q L_x^{\alpha+2}$$

where

$$q_1 = q'(\alpha+1). \tag{5.56}$$

In order for Ψ to map $L_t^q L_x^{\alpha+2}$ into itself, we need $L_t^{q_1} \supset L_t^q$ or $q_1 \le q$. This condition holds if $\alpha + 2 \le q$ or $\alpha \le 4/n$. In order to prove that Φ is a contraction we will interpolate in time from $L_t^{q_1}$ to L_t^q, which requires that $q_1 < q$ or $\alpha < 4/n$. A similar existence result holds in the critical case $\alpha = 4/n$ but the proof requires a more refined argument which we do not describe here.

Thus according to this discussion,

$$\Phi : L_t^q L_x^{\alpha+2} \to L_t^q L_x^{\alpha+2}$$

if q is given by (5.55) and $0 < \alpha < 4/n$. This motivates the hypotheses in the following theorem.

THEOREM 5.55. *Suppose that* $0 < \alpha < 4/n$ *and*

$$q = \frac{4(\alpha+2)}{n\alpha}.$$

For every $f \in L^2(\mathbb{R}^n)$*, there exists*

$$T = T\left(\|f\|_{L^2}, n, \alpha, \lambda\right) > 0$$

and a unique solution u of (5.49) with

$$u \in C\left([0,T]; L^2(\mathbb{R}^n)\right) \cap L^q\left(0,T; L^{\alpha+2}(\mathbb{R}^n)\right).$$

Moreover, the solution map $f \mapsto u$ is locally Lipschitz continuous.

Proof. For $T > 0$, let E be the Banach space

$$E = C\left([0,T]; L^2\right) \cap L^q\left(0,T; L^{\alpha+2}\right)$$

with norm

$$\|u\|_E = \max_{[0,T]} \|u(t)\|_{L^2} + \left(\int_0^T \|u(t)\|_{L^{\alpha+2}}^q \, dt\right)^{1/q} \tag{5.57}$$

and let Φ be the map in (5.50)–(5.51). We claim that $\Phi(u)$ is well-defined for $u \in E$ and $\Phi : E \to E$.

The preceding discussion shows that $\Phi(u) \in L_t^q L_x^{\alpha+2}$ if $u \in L_t^q L_x^{\alpha+2}$. Writing $C_t L_x^2 = C\left([0,T]; L^2\right)$, we see that $\mathrm{T}(\cdot)f \in C_t L_x^2$ since $f \in L^2$ and T is a strongly continuous group on L^2. Moreover, (5.47) implies that $\Psi(u) \in C_t L_x^2$ since $\Psi(u)$ is the uniform limit of smooth functions $\Psi(u_k)$ such that $u_k \to u$ in $L_t^q L_x^{\alpha+2}$ *c.f.* (5.71). Thus, $\Phi : E \to E$.

Next, we estimate $\|\Phi(u)\|_E$ and show that there exist positive numbers

$$T = T\left(\|f\|_{L^2}, n, \alpha, \lambda\right), \qquad a = a\left(\|f\|_{L^2}, n, \alpha\right)$$

such that Φ maps the ball

$$B = \{u \in E : \|u\|_E \leq a\} \tag{5.58}$$

into itself.

First, we estimate $\|\mathrm{T}f\|_E$. Since T is a unitary group, we have

$$\|\mathrm{T}f\|_{C_t L_x^2} = \|f\|_{L^2} \tag{5.59}$$

while the Strichartz estimate (5.46) implies that

$$\|\mathrm{T}f\|_{L_t^q L_x^{\alpha+2}} \leq C\|f\|_{L^2}. \tag{5.60}$$

Thus, there is a constant $C = C(n, \alpha)$ such that

$$\|\mathrm{T}f\|_E \leq C\|f\|_{L^2}. \tag{5.61}$$

In the rest of the proof, we use C to denote a generic constant depending on n and α.

Second, we estimate $\|\Psi(u)\|_E$ where Ψ is given by (5.51). The Strichartz estimate (5.47) gives

$$\begin{aligned} \|\Psi(u)\|_{C_t L_x^2} &\leq C\| \, |u|^{\alpha+1} \, \|_{L_t^{q'} L_x^{(\alpha+2)'}} \\ &\leq C\|u\|_{L_t^{q'(\alpha+1)} L_x^{(\alpha+2)'(\alpha+1)}}^{\alpha+1} \\ &\leq C\|u\|_{L_t^{q_1} L_x^{\alpha+2}}^{\alpha+1} \end{aligned} \tag{5.62}$$

where q_1 is given by (5.56). If $\phi \in L^p(0,T)$ and $1 \le p \le q$, then Hölder's inequality with $r = q/p \ge 1$ gives

$$\begin{aligned} \|\phi\|_{L^p(0,T)} &= \left(\int_0^T 1 \cdot |\phi(t)|^p \, dt\right)^{1/p} \\ &\le \left[\left(\int_0^T 1^{r'} \, dt\right)^{1/r'} \left(\int_0^T |\phi(t)|^{pr} \, dt\right)^{1/r}\right]^{1/p} \\ &\le T^{1/p-1/q} \|\phi\|_{L^q(0,T)}. \end{aligned} \tag{5.63}$$

Using this inequality with $p = q_1$ in (5.62), we get

$$\|\Psi(u)\|_{C_t L^2_x} \le C T^\theta \|u\|^{\alpha+1}_{L^q_t L^{\alpha+2}_x} \tag{5.64}$$

where $\theta = (\alpha+1)(1/q_1 - 1/q) > 0$ is given by

$$\theta = 1 - \frac{n\alpha}{4}. \tag{5.65}$$

We estimate $\|\Psi(u)\|_{L^q_t L^{\alpha+2}_x}$ in a similar way. The Strichartz estimate (5.48) and the Hölder estimate (5.63) imply that

$$\|\Psi(u)\|_{L^q_t L^{\alpha+2}_x} \le C \|u\|^{\alpha+1}_{L^{q_1}_t L^{\alpha+2}_x} \le C T^\theta \|u\|^{\alpha+1}_{L^q_t L^{\alpha+2}_x}. \tag{5.66}$$

Thus, from (5.64) and (5.66), we have

$$\|\Psi(u)\|_E \le C T^\theta \|u\|^{\alpha+1}_{L^q_t L^{\alpha+2}_x}. \tag{5.67}$$

Using (5.61) and (5.67), we find that there is a constant $C = C(n,\alpha)$ such that

$$\|\Phi(u)\|_E \le \|\mathrm{T}f\|_E + |\lambda| \, \|\Psi(u)\|_E \le C\|f\|_{L^2} + C|\lambda| T^\theta \|u\|^{\alpha+1}_{L^q_t L^{\alpha+2}_x} \tag{5.68}$$

for all $u \in E$. We choose positive constants a, T such that

$$a \ge 2C\|f\|_{L^2}, \qquad 0 < 2C|\lambda| T^\theta a^\alpha \le 1.$$

Then (5.68) implies that $\Phi : B \to B$ where $B \subset E$ is the ball (5.58).

Next, we show that Φ is a contraction on B. From (5.50) we have

$$\Phi(u) - \Phi(v) = i\lambda \left[\Psi(u) - \Psi(v)\right]. \tag{5.69}$$

Using the Strichartz estimates (5.47)–(5.48) in (5.51) as before, we get

$$\|\Psi(u) - \Psi(v)\|_E \le C \left\| \, |u|^\alpha u - |v|^\alpha v \, \right\|_{L^{q'}_t L^{(\alpha+2)'}_x}. \tag{5.70}$$

For any $\alpha > 0$ there is a constant $C(\alpha)$ such that

$$\left| \, |w|^\alpha w - |z|^\alpha z \, \right| \le C \left(|w|^\alpha + |z|^\alpha\right) |w - z| \qquad \text{for all } w, z \in \mathbb{C}.$$

Using the identity

$$(\alpha+2)' = \frac{\alpha+2}{\alpha+1}$$

and Hölder's inequality with $r = \alpha + 1$, $r' = (\alpha + 1)/\alpha$, we get that

$$\begin{aligned}
\| \, |u|^\alpha u - |v|^\alpha v \, \|_{L_x^{(\alpha+2)'}} &= \left(\int | \, |u|^\alpha u - |v|^\alpha v \, |^{(\alpha+2)'} \, dx \right)^{1/(\alpha+2)'} \\
&\leq C \left(\int (|u|^\alpha + |v|^\alpha)^{(\alpha+2)'} \, |u - v|^{(\alpha+2)'} \, dx \right)^{1/(\alpha+2)'} \\
&\leq C \left(\int (|u|^\alpha + |v|^\alpha)^{r'(\alpha+2)'} \, dx \right)^{1/r'(\alpha+2)'} \\
&\qquad \left(\int |u - v|^{r(\alpha+2)'} \, dx \right)^{1/r(\alpha+2)'} \\
&\leq C \left(\|u\|_{L_x^{\alpha+2}}^\alpha + \|v\|_{L_x^{\alpha+2}}^\alpha \right) \|u - v\|_{L_x^{\alpha+2}}
\end{aligned}$$

We use this inequality in (5.70) followed by Hölder's inequality in time to get

$$\begin{aligned}
\|\Psi(u) - \Psi(v)\|_E &\leq C \left(\int_0^T \left[\|u\|_{L_x^{\alpha+2}}^\alpha + \|v\|_{L_x^{\alpha+2}}^\alpha \right]^{q'} \|u - v\|_{L_x^{\alpha+2}}^{q'} \, dt \right)^{1/q'} \\
&\leq C \left(\int_0^T \left[\|u\|_{L_x^{\alpha+2}}^\alpha + \|v\|_{L_x^{\alpha+2}}^\alpha \right]^{p'q'} \, dt \right)^{1/p'q'} \\
&\qquad \left(\int_0^T \|u - v\|_{L_x^{\alpha+2}}^{pq'} \, dt \right)^{1/pq'} .
\end{aligned}$$

Taking $p = q/q' > 1$ we get

$$\|\Psi(u) - \Psi(v)\|_E \leq C \left(\int_0^T \left[\|u(t)\|_{L_x^{\alpha+2}}^{\alpha p'q'} + \|v(t)\|_{L_x^{\alpha+2}}^{\alpha p'q'} \right] dt \right)^{1/pq'} \|u - v\|_{L_t^q L_x^{\alpha+2}} .$$

Interpolating in time as in (5.63), we have

$$\int_0^T \|u(t)\|_{L_x^{\alpha+2}}^{\alpha p'q'} \, dt \leq \left(\int_0^T 1^{\alpha p'q'r'} \, dt \right)^{1/r'} \left(\int_0^T \|u(t)\|_{L_x^{\alpha+2}}^{\alpha p'q'r} \, dt \right)^{1/r}$$

and taking $\alpha p'q'r = q$, which implies that $1/p'q'r' = \theta$ where θ is given by (5.65), we get

$$\left(\int_0^T \|u(t)\|_{L_x^{\alpha+2}}^{\alpha p'q'} \, dt \right)^{1/r} \leq T^\theta \|u - v\|_{L_t^q L_x^{\alpha+2}} .$$

It therefore follows that

$$\|\Psi(u) - \Psi(v)\|_E \leq C T^\theta \left(\|u\|_{L_t^q L_x^{\alpha+2}}^\alpha + \|v\|_{L_t^q L_x^{\alpha+2}}^\alpha \right) \|u - v\|_{L_t^q L_x^{\alpha+2}} . \tag{5.71}$$

Using this result in (5.69), we get

$$\|\Phi(u) - \Phi(v)\|_E \leq C|\lambda| T^\theta \left(\|u\|_E^\alpha + \|v\|_E^\alpha \right) \|u - v\|_E .$$

Thus if $u, v \in B$,

$$\|\Phi(u) - \Phi(v)\|_E \leq 2C|\lambda| T^\theta a^\alpha \|u - v\|_E .$$

Choosing $T > 0$ such that $2C|\lambda|T^\theta a^\alpha < 1$, we get that $\Phi : B \to B$ is a contraction, so it has a unique fixed point in B. Since we can choose the radius a of B as large as we wish by taking T small enough, the solution is unique in E.

The Lipshitz continuity of the solution map follows from the contraction mapping theorem. If Φ_f denotes the map in (5.50), $\Phi_{f_1}, \Phi_{f_2} : B \to B$ are contractions, and u_1, u_2 are the fixed points of Φ_{f_1}, Φ_{f_2}, then

$$\|u_1 - u_2\|_E \leq C \|f_1 - f_2\|_{L^2} + K \|u_1 - u_2\|_E$$

where $K < 1$. Thus

$$\|u_1 - u_2\|_E \leq \frac{C}{1-K} \|f_1 - f_2\|_{L^2}.$$

□

This local existence theorem implies the global existence of L^2-solutions for subcritical nonlinearities $0 < \alpha < 4/n$ because the existence time depends only the L^2-norm of the initial data and one can show that the L^2-norm of the solution is constant in time.

For more about the extensive theory of the nonlinear Schrödinger equation and other nonlinear dispersive PDEs see, for example, [**6, 29, 39, 40**].