Math 951 Lecture Notes
Chapter 6 – Introduction to Semigroup Methods

Mathew A. Johnson [1]
Department of Mathematics, University of Kansas
matjohn@ku.edu

Contents

1 Introduction

In this chapter, we consider a set of techniques referred to as "semigroup methods" to study the existence of solutions of both linear and nonlinear evolution equations. The basic idea is to essentially view an evolution equation as an infinite dimensional dynamical system and to try to use the basic ideas and techniques from ODE theory (i.e. finite dimensional dynamical systems) in the PDE setting. The advantages of this viewpoint can not be understated, and it is evidence by the fact that this methodology us used throughout both pure and applied analysis of PDE including the development of well-posedness theories, the existence of solutions, the analysis of nonlinear smoothing effects, and in understanding the local stability and global dynamics of nonlinear PDE. Given its wide applicability, we clearly can't cover all those topics here. For these notes, we will primarily concern ourselves with the development of existence and well-posedness theories.

To this end, we first motivate some of the basic ideas by first reviewing the easiest case possible, which is when what we know from finite dimensional dynamical systems actually works without modification.

1.1 Motivation: Uniformly Continuous Groups of Opertaors

To begin, suppose X is a Banach space and recall the set

$$\mathcal{L}(X) = \{A : X \to X \ : \ A \text{ is linear and bounded}\}$$

is a Banach space when equipped with the operator norm

$$\|A\|_{\mathcal{L}(X)} = \sup_{u \in X \setminus \{0\}} \frac{\|Au\|_X}{\|u\|_X} = \sup_{\|u\|_X = 1} \|Au\|_X.$$

Now, fix $A \in \mathcal{L}(X)$ and note that for each $t \in \mathbb{R}$ we can define the operator $e^{At} : X \to X$ by

$$e^{At} = \sum_{j=0}^{\infty} \frac{A^j t^j}{j!}. \tag{1}$$

Note that, since $\|A\|_{\mathcal{L}(X)} < \infty$, the above series clearly converges uniformly (i.e. in the operator norm) on compact intervals of $t \in \mathbb{R}$. Indeed, note for any $k < m$ we have

$$\left\| \sum_{j=1}^{k} \frac{A^j t^j}{j!} - \sum_{j=1}^{m} \frac{A^j t^j}{j!} \right\|_{\mathcal{L}(X)} \leq \sum_{j=k+1}^{m} \frac{\|A\|_{\mathcal{L}(X)}^j t^j}{j!}$$

with the latter sum being uniformly small on compact intervals of $t \in \mathbb{R}$ for all k, m sufficiently large. Extending this calculation, the following facts can be easily established.

Lemma 1. *Given $A \in \mathcal{L}(X)$, the following are true.*

(a) The map

$$\mathbb{R} \ni t \mapsto e^{At} \in \mathcal{L}(X)$$

is C^∞.

(b) For all $t, s \in \mathbb{R}$ we have

$$e^{As} e^{At} = e^{A(s+t)} = e^{At} e^{As}.$$

(c) $\frac{d}{dt} e^{At} = A e^{At} = e^{At} A$.

The proof of this lemma is left as an exercise. A consequence is the following fundamental existence and uniqueness result from elementary ODE theory.

Theorem 1. *Let $A \in \mathcal{L}(X)$ and consider the homogeneous IVP*

$$\begin{cases} u_t = Au, \quad t \in \mathbb{R} \\ u(0) = f \in X. \end{cases}$$

Then the unique solution $u \in C^\infty(\mathbb{R}; X)$ is

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

We now illustrate this result with some basic examples.

Example: If $A \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ is an $n \times n$ matrix, then the unique solution of the linear IVP

$$\begin{cases} u_t = Au, \quad t \in \mathbb{R} \\ u(0) = u_0 \in \mathbb{R}^n \end{cases}$$

is given by the matrix exponential $u(t) = e^{At} u_0$. This is just basic linear ODE theory! Of particular note, observe that for each fixed $u_0 \in \mathbb{R}^n$ we have

$$\left| e^{At} u_0 - u_0 \right| = \left| \sum_{j=1}^\infty \frac{t^j}{j!} A^j u_0 \right| \leq \left(\sum_{j=1}^\infty \frac{|t|^j}{j!} \|A\|^j_{\mathcal{L}(\mathbb{R}^n)} \right) |u_0| = \left(e^{t\|A\|_{\mathcal{L}(\mathbb{R}^n)}} - 1 \right) |u_0|$$

so that, in particular,

$$\left\| e^{At} - I \right\|_{\mathcal{L}(\mathbb{R}^n)} \leq e^{t\|A\|_{\mathcal{L}(\mathbb{R}^n)}} - 1.$$

It follows in this case that the operator e^{At} converges uniformly on $\mathbb{R}^n$ (i.e. in the operator norm) to the identity operator as $t \to 0$. Note that since

$$e^{A(t+h)} - e^{At} = e^{At} \left(e^{Ah} - I \right),$$

it actually follows from above that for any fixed $t \in \mathbb{R}$ we have $e^{A(t+h)} \to e^{At}$ uniformly in $\mathcal{L}(\mathbb{R}^n)$ as $h \to 0$.

Example: For a fixed $p \in [1, \infty]$ define $A : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ be the translation operator

$$Af(x) = f(x+1), \quad x \in \mathbb{R}.$$

Since we clearly have $A \in \mathcal{L}(L^p(\mathbb{R}))$, it follows that the unique solution $u \in C^\infty(\mathbb{R}; L^p(\mathbb{R}))$ of the differential-difference equation

$$\begin{cases} u_t(x,t) = u(x+1,t), \quad x, t \in \mathbb{R} \\ u(\cdot, 0) = f \in L^p(\mathbb{R}) \end{cases}$$

is given by $u(x,t) = e^{At}f(x)$. More explicitly, the unique solution is given by

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

In particular, note that for each fixed $f \in L^p(\mathbb{R}^n)$ we have

$$\left\| e^{At} f - f \right\|_{L^p(\mathbb{R}^n)} \leq \left(e^{|t|} - 1 \right) \|f\|_{L^p(\mathbb{R}^n)}$$

and hence that

$$\left\| e^{At} - I \right\|_{\mathcal{L}(L^p(\mathbb{R}^n))} \leq e^{|t|} - 1.$$

It follows that, as with the previous example, the operator e^{At} converges uniformly on $L^p(\mathbb{R}^n)$, i.e. in $\mathcal{L}(L^p(\mathbb{R}^n))$, to the identity operator as $t \to 0$.

Example: Given $a \in L^1(\mathbb{R}^n)$, define the convolution operator $A : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ by

$$Af(x) = a * f(x) = \int_{\mathbb{R}^n} a(x-y) f(y) dy$$

and consider the integral equation IVP

$$\begin{cases} u_t = Au, \quad t \in \mathbb{R} \\ u(0) = f \in L^2(\mathbb{R}^n) \end{cases} \tag{2}$$

Now, $A \in \mathcal{L}(L^2(\mathbb{R}^n))$ by Young's convolution inequality[2], which states that

$$\|g * h\|_{L^r(\mathbb{R}^n)} \leq \|g\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}$$

where the $r, p, q \geq 1$ satisfy

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

Applying this inequality with $p = 1$ and $q = r = 2$ we get

$$\|Af\|_{L^2(\mathbb{R}^n)} \leq \|a\|_{L^1(\mathbb{R}^n)} \|f\|_{L^2(\mathbb{R}^n)},$$

as desired. By the above theorem, it follows that for each $f \in L^2(\mathbb{R}^n)$ the unique solution to (2) is given by

$$u(x,t) = e^{At} f(x).$$

Note that using the Fourier transform we find an alternative representation for the solution of (2) as

$$u(x,t) = \int_{\mathbb{R}^n} g(x-y,t) f(x) dx$$

[2]An elementary, yet complicated, proof can be given with Hölder's inequality. Alternatively, this maybe be established as a consequence of Riesz-Thorin interpolation.

where here $g \in L^1(\mathbb{R}^n)$ is such that

$$\widehat{g}(\xi, t) = e^{(2\pi)^n \widehat{a}(\xi)t}.$$

By uniqueness, it follows that

$$e^{At} f(x) = \int_{\mathbb{R}^n} g(x-y, t) f(x) dx.$$

which, Young's convolution inequality, clearly shows that $e^{At} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. In particular, observe by Plancerel's theorem[3] that for each fixed $f \in L^2(\mathbb{R}^n)$ we have

$$\left\| e^{At} f - f \right\|_{L^2(\mathbb{R}^n)} \leq \left\| e^{(2\pi)^n \widehat{a}(\xi)t} - 1 \right\|_{L^\infty_\xi(\mathbb{R}^n)} \|f\|_{L^2(\mathbb{R}^n)}$$

which, since $a \in L^1(\mathbb{R}^n)$ implies $\widehat{a} \in C(\mathbb{R}^n)$ with $\widehat{a}(\xi) \to 0$ as $\xi \to \infty$, again implies that $e^{At} \to I$ in $\mathcal{L}(L^2(\mathbb{R}^n))$ as $t \to 0$.

In each of the examples above, the solution operators $T(t) = e^{At}$ can be shown to satisfy the following properties:

(1) $T(0) = I$.

(2) $T(s)T(t) = T(s+t)$ for all $s, t \in \mathbb{R}$.

(3) $T(h) \to I$ uniformly in $\mathcal{L}(X)$ as $h \to 0$.

Such a one-parameter family of operators $\{T(t)\}_{t \in \mathbb{R}}$ defines a *uniformly continuous group of operators*. While such a family of operators is certainly nice to have[4], it turns out that they practically never occur in the study of PDE due to the following result.

Lemma 2. *Given a Banach space X, a family $\{T(t)\}_{t \in \mathbb{R}}$ is a uniformly continuous group of operators on X if and only if*

$$T_t(0) \in \mathcal{L}(X).$$

In each of the above examples, $T_t(0) = A$, which was always a bounded linear operator on the associated Banach space X. In PDE applications, however, we are working with differential operators which, as we have sen before, are not continuous from natural Sobolev spaces into themselves. It follows that if we want the the above methodologies to be applicable to the study of PDE, we need to find a suitable replacement for our definition of e^{At} when $A \notin \mathcal{L}(X)$. This is precisely the basic goal of semigroup theory. In order to motivate what such an extension might look like in the PDE setting, lets consider an example.

[3]This says that, up to a scaling factor, the Fourier transform is an isometry from $L^2(\mathbb{R}^n)$ into itself.

[4]In fact, its all you ever see in finite dimensional dynamical systems!

1.2 Semigroups: A Motivating Example

Let $\Omega \subset \mathbb{R}^n$ be open and bounded, and consider the following IVBVP for the heat equation:

$$\begin{cases} u_t = \Delta u, \quad x \in \Omega,\ t > 0 \\ u = 0 \quad \text{on } \partial\Omega \\ u(0) = g \end{cases} \tag{3}$$

where here $g \in H^2(\Omega) \cap H_0^1(\Omega)$. In our previous work, we saw that the unique solution of (3) is given by

$$u(x,t) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k t} \phi_k(x).$$

where the (λ_k, ϕ_k) are the eigenvalue/eigenfunction pairs for $-\Delta$ on $H_0^1(\Omega)$, with the ϕ_k chosen to be orthonormal in $L^2(\Omega)$, and $a_n := \langle \phi_n, g \rangle_{L^2(\mathbb{R}^n)}$. The above naturally defines a one-parameter family of operators

$$T(t) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n), \quad T(t)g(x) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k t} \phi_k(x) \tag{4}$$

defined for all $t \geq 0$ such that for each $g \in H^2(\Omega) \cap H_0^1(\Omega)$ the solution to (3) is given by $u(t) = T(t)g$. Note that in addition to only being defined for $t \geq 0$, and hence having no possibility of forming an actual group of operators, we have

$$T_t(0)g(x) = -\sum_{k=1}^{\infty} a_k e^{-\lambda_k t} \lambda_k \phi_k(x) = -\Delta g(x)$$

for all[5] $g \in H^2(\Omega) \cap H_0^1(\Omega)$ so that $T_t(0) \notin \mathcal{L}(L^2(\mathbb{R}^n))$. In particular, note the domain of $T_t(0)$ is not even all of $L^2(\Omega)$. Nevertheless, lets study what structure, if any, from the previous section we still retain in this setting.

First, note that, by construction, we clearly have $T(0)g = g$ for all $g \in L^2(\Omega)$. Moreover, for a fixed $g \in L^2(\Omega)$ and $s, t \geq 0$ we have

$$\begin{aligned} T(t)T(s)g(x) &= T(t)\left(\sum_{k=1}^{\infty} a_k e^{-\lambda_k s} \phi_k(x)\right) \\ &= \sum_{j=1}^{\infty} \underbrace{\left\langle \sum_{k=1}^{\infty} a_k e^{-\lambda_k s} \phi_k(\cdot), \phi_j(\cdot) \right\rangle_{L^2(\Omega)}}_{a_k e^{-\lambda_k s} \delta_{j,k}} e^{-\lambda_j t} \phi_j(x) \\ &= \sum_{k=1}^{\infty} a_k e^{-\lambda_k (t+s)} \phi_k(x) \\ &= T(t+s)g(x) \end{aligned}$$

[5]Specifically, not e that $T_t(0)$ is note even *defined* on all of $L^2(\mathbb{R}^n)$. As we will see, this is a reflection of the fact that $T_t(0)$ is an unbounded operator on $L^2(\mathbb{R}^n)$.

so that $T(s)T(t) = T(s+t)$ for all $s, t \geq 0$. In particular, it follows that the family of operators $\{T(t)\}_{t\geq 0}$ forms a *semigroup* of operators[6].

Finally, we characterize the continuity of T. Note that for all $g \in L^2(\Omega)$ and $h > 0$ we have

$$\begin{aligned} \|T(h)g - g\|_{L^2(\Omega)}^2 &= \left\| \sum_{k=1}^{\infty} \langle g, \phi_k \rangle_{L^2(\Omega)} \left(e^{-\lambda_k t} - 1 \right) \phi_k \right\|_{L^2(\Omega)}^2 \\ &= \sum_{k=1}^{\infty} |\langle g, \phi_k \rangle|^2 \left(e^{-\lambda_k h} - 1 \right)^2, \end{aligned}$$

where the last equality follows by the orthonormality of the eigenfunctions $\{\phi_k\}$. In particular, it follows that for each $g \in L^2(\Omega)$ we certainly have the strong convergence

$$T(h)g \to g \;\text{ in } L^2(\Omega) \text{ as } h \to 0^+.$$

However, we note that this convergence is *not uniform in* $\mathcal{L}(L^2(\Omega))$ since we can't bound the sequence

$$\left(e^{-\lambda_k h} - 1 \right)^2$$

above uniformly in k by something that decays to zero as $h \to 0^+$.

So, in this linear diffusion example, it appears that the family $\{T(t)\}_{t\geq 0}$ forms a semigroup of operators which is strongly continuous, but not uniformly continuous. It turns out that such families of operators are common in the study of PDE, and in the next section we begin our analysis of such families.

2 Semigroups of Operators

2.1 Basic Definitions & Properties

Motivated by the examples in the previous section, we make the following definition.

Definition 1. *Let X be a Banach space. A one-parameter strongly continuous (i.e. C_0) semigroup on X is a family of operators*

$$\{T(t)\}_{t\geq 0} \subset \mathcal{L}(X)$$

such that the following hold:

(i) $T(0) = I$.

[6]Recall that a semigroup is an algebraic structure consisting of a set and an associative binary operation. In particular, in the strictly algebraic sense, semigroups need not have multiplicative inverses, nor are they required to have identity elements. So, technically speaking, the family $\{T(t)\}_{t\geq 0}$ forms a semigroup with identity, but we will never make this distinction.

(ii) $T(t)T(s) = T(t+s)$ *for all* $s, t \geq 0$.

(iii) For all $f \in X$, $T(h)f \to f$ *strongly in* X *as* $h \to 0^+$.

Further, such a C_0 *semigroup is said to be* ω*-contractive for some* $\omega \in \mathbb{R}$ *if, additionally,*

$$\|T(t)\|_{\mathcal{L}(X)} \leq e^{\omega t}$$

for all $t \geq 0$.

Note that condition (iii) above is equivalent to requiring that for each $f \in X$ the mapping

$$[0, \infty) \ni t \mapsto T(t)f \in X$$

is continuous. Indeed, fixing $f \in X$, the continuity of $T(\cdot)f$ at $t = 0$ implies that $\delta > 0$ sufficiently small there exists a constant $M_{\delta,f} > 0$ such that

$$\|T(t)f\|_X \leq M_{\delta,f} \text{ for all } t \in [0, \delta].$$

By the Uniform Boundedness Principle[7], it follows that the operator norm $\|T(t)\|$ is bounded for $t \in [0, \delta]$ and hence, by the semigroup property (ii) above, it follows that $\|T(t)\|$ is bounded on any finite time time interval. For a given $f \in X$ and $t > 0$ fixed, the continuity of $T(t)f$ from the right follows directly from the semigroup property and (iii), while the continuity from the left follows from the identity

$$T(t-h)f - T(t)f = T(t-h)\left(I - T(h)\right)f, \quad h > 0.$$

Another important observation is that every C_0-semigroup is ω-contractive for some $\omega \in \mathbb{R}$. For instance, in the heat equation example considered in Section 1.2 we have that

$$\|T(t)g\|_{L^2(\Omega)} \leq e^{-\lambda_1 t}\|g\|_{L^2(\Omega)}$$

for all $g \in L^2(\Omega)$ so that, in that case, the C_0 semigroup $\{T(t)\}_{t\geq 0}$ constructed there is actually $-\lambda_1$-contractive.

Next, we see that every C_0-semigroup is essentially of the form "e^{At}" for some operator A.

Definition 2. *Let* $\{T(t)\}_{t\geq 0}$ *be a* C_0*-semigroup on a Banach space* X*. Define an operator* A *as*

$$Au := \lim_{t\to 0^+} \frac{T(t)u - u}{t} = \frac{d}{dt}T(t)u\Big|_{t=0},$$

with domain

$$D(A) = \{u \in X : \text{ above limit exists in } X\}.$$

The operator A *is called the (infinitesimal) generator of the semigroup.*

[7]Take Math 960.

Note that if A is the generator for a semigroup $\{T(t)\}_{t\geq 0}$ then we always have (at least) $0 \in D(A)$. Furthermore, it follows from the definition that generators are always linear operators on $D(A)$. Note in the heat equation example in Section 1.2 we had

$$T_t(0) = -\Delta =: A$$

which is well-defined on the domain $D(A) = H^2(\Omega) \cap H^1_0(\Omega)$. In particular, in this case the domain $D(A)$ was dense in $L^2(\Omega)$. As we will see later, this is a general feature of the generators for C_0-semigroups.

The usefulness of the above concepts is made clear by the following result.

Theorem 2. *Suppose A is the generator of a C_0-semigroup $\{T(t)\}_{t>0}$ on a Banach space X. Then for all $u \in D(A)$, we have the following:*

(i) $T(t)u \in D(A)$ for all $t \geq 0$, i.e. $D(A)$ is an invariant set for the semigroup.

(ii) $AT(t)u = T(t)Au$ for each $t > 0$, i.e. generators and semigroups commute.

(iii) The mapping $t \mapsto T(t)u \in X$ is $C^1(0,\infty;X)$.

(iv) We have $\frac{d}{dt}T(t)u = AT(t)u$ for all $t > 0$.

Proof. To prove (i) and (ii), let $u \in D(A)$ be fixed and note that by the semigroup property we have

$$\frac{T(s)T(t)u - T(t)u}{s} = T(t)\left(\frac{T(s)u - u}{s}\right)$$

valid for all $t \geq 0$ and $s > 0$. Fixing $t \geq 0$ and noting that

$$\lim_{s\to 0^+} T(t)\left(\frac{T(s)u - u}{s}\right) = T(t)\left(\lim_{s\to 0^+}\frac{T(s)u - u}{s}\right) = T(t)Au,$$

it follows that

$$\lim_{s\to 0^+}\frac{T(s)T(t)u - T(t)u}{s} = T(t)Au.$$

In particular, we have that $T(t)u \in D(A)$ with $AT(t)u = T(t)Au$, as claimed.

Next, note that if $t > 0$ is fixed then for all $h > 0$ we have by the semigroup property that

$$\frac{T(t+h)u - T(t)u}{h} = T(t)\left(\frac{T(h)u - u}{h}\right)$$

so that

$$\lim_{h\to 0^+}\frac{T(t+h)u - T(t)u}{h} = T(t)Au.$$

For the left hand limit, note that for $t > 0$ fixed and $h > 0$ sufficiently small we have by the semigroup property that

$$T(t) = T(t-h)T(h)$$

and hence

$$\begin{aligned}\frac{T(t)u - T(t-h)u}{h} &= T(t-h)\left(\frac{T(h)u-u}{h}\right)\\ &= T(t-h)\left(\frac{T(h)u-u}{h} - Au\right) + T(t-h)Au.\end{aligned}$$

Since $\|T(t-h)\| \leq 1$ for all $0 < h < t$ and since $\frac{T(h)u-u}{h} \to Au$ in X as $h \to 0^+$ it follows that

$$\lim_{h\to 0^+} \frac{T(t)u - T(t-h)u}{h} = T(t)Au.$$

Therefore, the mapping $t \mapsto T(t)u$ is differentiable for each $t > 0$ with $\frac{d}{dt}T(t)u = T(t)Au = AT(t)u$. Finally, note that since the map $t \mapsto T(t)Au$ is continuous, we have that the mapping $t \mapsto T(t)u$ is C^1 on $t > 0$, as claimed. □

The main utility of the above technical result is the following. Let X be a Banach space and consider the following abstract IVP

$$\begin{cases} \dfrac{du}{dt} = Au, \quad 0 < t < \infty \\ u(0) = g, \end{cases} \tag{5}$$

where here A is a linear operator on X. By Theorem 2, it follows that if A is the generator for a C_0-semigroup $\{T(t)\}_{t\geq 0}$ on X, and if $g \in D(A)$, then the function

$$u(t) = T(t)g$$

is a *classical solution* of (5), in the sense that

$$u \in C([0,\infty); D(A)) \cap C^1((0,\infty); X)$$

and $u(t) \in D(A)$ for all $t > 0$ with $u(0) = g$ and u satisfies (5) pointwise for all $t > 0$. Consequently, we can use semigroup theory to solve linear IVP of the form (5), provided that the operator A is the generator for some C_0-semigroup on the Banach space X.

Example: Returning to our linear heat example in Section 1.2, consider the linear IVBVP

$$\begin{cases} u_t = \Delta u, \quad x \in \Omega,\ t > 0 \\ u = 0 \ \text{ on } \partial\Omega \\ u(0) = g \end{cases} \tag{6}$$

Recall that in Section 1.2 we proved that the family of operators $\{T(t)\}_{t\geq 0}$ defined in (4) formed a semigroup on $L^2(\Omega)$ with generator

$$A = -\Delta, \quad D(A) = H^2(\Omega) \cap H_0^1(\Omega).$$

According to Theorem 2, it follows for each $g \in D(A)$ the function $u(t) = T(t)g$ is a classical (in the above sense) solution of the linear IVBVP (6).

Of course, in the above example one would be correct in saying that we "cheated", in the sense that we only proved that $-\Delta$ generated a C_0-semigroup on $L^2(\mathbb{R}^n)$ by actually analyzing the semigroup directly. If we already had a way to study the semigroup directly, there would be little need for abstract results like Theorem 2! Consequently, to make the above methodology useful in applications, we need to come up with a way of determining if a given linear operator A is the generator for some C_0-semigroup on X that does not rely on having an explicit form of the semigroup itself. This is addressed in the following section.

2.2 Classification of Generators

In this section, our goal is to find necessary and sufficient conditions for a given linear operator A on a Banach space X to be the generator for some C_0-semigroup on X. As seen in Section 1.1, it is clear that a sufficient condition is for $A \in \mathcal{L}(X)$, in which case A is actually the generator of a uniformly continuous group of operators. In PDE applications however, we are typically considering linear evolution equations on X of the form

$$\frac{du}{dt} = Au, \quad 0 < t < \infty$$

where A is an unbounded operator on X. While such operators are clearly not continuous on X (by definition), it turns out that they often belong to a special class of "closed" operators.

Definition 3. *Let X and Y be Banach spaces, and let $A : D(A) \subset X \to Y$ be a linear operator with domain $D(A)$. Then A is said to be closed if for every sequence $\{x_n\}$ in $D(A)$ such that $x_n \to x$ in X and $Ax_n \to y$ in Y, we have $x \in D(A)$ and $Ax = y$.*

Said more abstractly, an operator A is said to be closed if its graph

$$\text{gra}(A) = \{x \oplus Ax : x \in D(A)\}$$

is a closed subset of $X \oplus Y$. Notice, in particular, that all bounded operators are closed. Indeed, if $A \in \mathcal{L}(X, Y)$ and if $x_n \to x$ in X then clearly $Ax_n \to Ax$ in Y However, in the definition of being a closed operator we only require that the property

$$\lim_{n\to\infty} Ax_n = Ax$$

holds for sequences $\{x_n\}$ in $D(A)$ that converge to $x \in X$ that have the additional property that the image sequence $\{Ax_n\}$ converges in Y. As the next example shows, the property of being closed requires an appropriate selection of the domain of the operator.

Example: Let $\Omega \subset \mathbb{R}^n$ be an open and bounded domain and consider the operator $A := -\Delta$ as a linear map from $L^2(\Omega)$ to $L^2(\Omega)$ with domain $D(A) = C_c^\infty(\Omega)$. Then A is not closed. Indeed, fix $f \in H^2(\Omega) \cap H_0^1(\Omega)$ and recall there exists a sequence $\{f_k\} \subset C_c^\infty(\Omega)$ such that $f_k \to f$ in $H^2(\Omega) \subset L^2(\Omega)$. Then in this case we know that $\{Af_k\}$ converges in $L^2(\Omega)$, but $f \notin D(A)$. It follows that the operator

$$-\Delta : C_c^\infty(\Omega) \subset L^2(\Omega) \to L^2(\Omega)$$

is not closed, as claimed.

Nevertheless, we can obtain a closed operator by extending A in the following way. Let $D(\widetilde{A})$ be the set of all functions $f \in L^2(\Omega)$ such that there exists a sequence $f_n \in D(A)$ and an element $g \in L^2(\mathbb{R})$ such that $f_n \to f$ in $L^2(\Omega)$ and $A(f_n) \to g$ in $L^2(\Omega)$. Thus, we can define an operator

$$\widetilde{A} : D(\widetilde{A}) \subset L^2(\Omega) \to L^2(\Omega)$$

by requiring that $\widetilde{A}(f) = g$. The operator $\widetilde{A}$ is clearly closed an is an *extension* of the operator A, in the sense that $D(A) \subset D(\widetilde{A})$ and $\widetilde{A}(f) = A(f)$ for all $f \in D(A)$. Notice that the domain of the extension can be recognized as

$$D(\widetilde{A}) = H^2(\Omega) \cap H_0^1(\Omega).$$

Following the above example, it seems that a natural way to try to obtain a closed extension of a given linear operator $A : D(A) \subset X \to Y$ is to simply take the closure of its graph in $X \oplus Y$. The problem is that the set $\overline{\text{gra}}^{X\oplus Y}$ may not be the graph of an operator, and hence it is not true that every linear operator has a closed extension. We can, however, characterize those operators which admit closed extensions.

Lemma 3. *A linear operator $A : D(A) \subset X \to Y$ has a closed extension, i.e. is closable, if for every seqeunce $\{x_n\}$ in $D(A)$ with $x_n \to 0$ in X, we either have $Ax_n \to 0$ in Y or else $\lim_{n\to\infty} Ax_n$ does not exist.*

Example: Let $X = L^2(\mathbb{R})$ and $Y = \mathbb{R}$, and consider the linear operators $Au = \int_{\mathbb{R}} u(x)dx$ with densely defined domain $D(A) = C_c^\infty(\mathbb{R})$. Then A is not closable. Indeed, fix $f \in D(A)$ and for each $k \in \mathbb{N}$ set

$$f_k(x) = k^{-1} f(x/k)$$

and note that

$$\|f_k\|_{L^2(\mathbb{R})}^2 = k^{-1} \int_{\mathbb{R}} f(y)^2 dy \to 0$$

as $k \to \infty$, but that

$$Af_k = \int_{\mathbb{R}} f(x)dx$$

for all k. By the above Lemma, it follows that $(D(A), A)$ does not have a closed extension.

Thankfully for us, it turns out that nearly every operator encounters in practical applications is closable. Neverthtless, the above example should serve as a warning that closability is still something one has to check.

With the above notions in mind, we now provide a necessary condition for a given operator A to generate a C_0 semigroup.

Proposition 1. *Suppose that A is the generator of a C_0-semigroup on a Banach space X. Then the operator A is closed and the domain $D(A)$ is dense in X.*

Proof. Suppose that A is the infinitesimal generator of a C_0-semigroup $\{T(t)\}_{t\geq 0}$. We first verify that the its domain $D(A)$, as defined in Definition 2, is dense in X. To this end, I claim that for every $f \in X$ and $t > 0$ that

$$\int_0^t T(s)f\ dx \in D(A).$$

To see this, observe that if $h > 0$ then

$$\begin{aligned}\left(\frac{T(h)-I}{h}\right)\int_0^t T(s)f\ ds &= \frac{1}{h}\int_0^t \left(T(s+h)-T(s)\right)f\ ds\\ &= \frac{1}{h}\left(\int_h^{t+h} T(s)f\ ds - \int_0^t T(s)f\ ds\right)\\ &= \frac{1}{h}\int_t^{t+h} T(s)f\ ds - \frac{1}{h}\int_0^h T(s)f\ ds.\end{aligned}$$

Since Theorem 2 implies the function $t \mapsto T(t)f \in X$ is a continuous function on $(0,\infty)$, it follows that

$$\lim_{h\to 0^+}\left(\frac{T(h)-I}{h}\right)\int_0^t T(s)f\ ds = T(t)f - f$$

which, since the limit exists, verifies the claim. Using again that the continuity of the semigroup we now see that

$$f = \lim_{h\to 0^+}\frac{1}{h}\int_0^h T(t)f\ dt$$

for every $f \in X$, from which the density of $D(A)$ in X follows.

To see that A is closed, let $f_n \in D(A)$ be a sequence with $f_n \to f \in D(A)$ and $Af_n \to g$ in X. Recalling that Theorem 2 implies

$$\frac{d}{dt}\left(T(t)f_n\right) = T(t)Af_n$$

for each n, it follows by the Fundamental Theorem of Calculus that

$$T(h)f_n - f_n = \int_0^h T(s)Af_n\ ds$$

which, taking $n \to \infty$, implies

$$T(h)f - f = \int_0^h T(s)g \, ds.$$

Multiplying by $\frac{1}{h}$ and again taking $h \to 0^+$, it follows as above that $f \in D(A)$ and that

$$Af = \lim_{h \to 0^+} \frac{T(h)f - f}{h} = \lim_{h \to 0^+} \frac{1}{h} \int_0^h T(s)g \, ds = g,$$

as desired. □

Next, we wish to complement the necessary conditions in Proposition 1 with sufficient conditions. For this, we need to consider the invertibility of a given linear operator.

Definition 4. *Let $A : D(A) \subset X \to X$ be a closed linear operator on a Banach space X with dense domain $D(A)$. We define the resolvent set of A to be*

$$\rho(A) = \{\lambda \in \mathbb{C} : (\lambda I - A) : D(A) \to X \quad \text{is a bijection}\}.$$

Furthermore, for each $\lambda \in \rho(A)$ we define the resolvent operator

$$R_\lambda(A) = (\lambda I - A)^{-1} : X \to X.$$

With the above, we are now able to give a necessary and sufficient condition for a given operator A to be generate a C_0-semigroup on X. This is provided by the famous HIlle-Yosida Theorem.

Theorem 3 (Hille-Yosida)**.** *Let A be a linear operator on a Banach space X. Then A is the generator of a C_0-semigroup on X if and only if the following conditions hold:*

(i) $D(A)$ is dense in X and A is a closed operator.

(ii) There exists an $\omega \in \mathbb{R}$ such that $(\omega, \infty) \subset \rho(A)$, and

$$\|R_\lambda(A)\| \leq \frac{1}{\lambda - \omega}, \quad \text{for all} \quad \lambda > \omega.$$

In this case the associated semigroup $\{T(t)\}_{t \geq 0}$ is ω-contractive, i.e. it satisfies

$$\|T(t)\| \leq e^{\omega t}$$

for all $t \geq 0$.

A proof of Theorem 3 is given in the Appendix. Note the necessity of (i) follows from Proposition 1. To motivate the necessity of (ii), suppose that A generates a C_0-semigroup $\{T(t)\}_{t\geq 0}$ on X which is ω-contractive and note for each $g \in X$ that the unique solution of

$$\begin{cases} \dfrac{du}{dt} = Au \\ u(0) = g \end{cases}$$

is given by $u(t) = T(t)g$. Taking the Laplace transform in t of the above IVP gives

$$\lambda\tilde{u}(\lambda) - g = A\tilde{u}(\lambda),$$

where $\tilde{u}(\lambda) = \int_0^\infty u(t)e^{-\lambda t}dt$ denotes the Laplace transform of u, which, recalling that $\|T(t)\| \leq e^{\omega t}$, is well-defined for all $\lambda > \omega$. It follows that **IF** $\lambda > \omega$ and **IF** $\lambda \in \rho(A)$ then we have

$$(\lambda I - A)\tilde{u}(\lambda) = g, \quad \text{i.e.} \quad \tilde{u}(\lambda) = R_\lambda(A)g,$$

which, by uniqueness, implies

$$R_\lambda(A)g = \int_0^\infty e^{-\lambda t}T(t)g\ dt. \tag{7}$$

Using that $T(t)$ is ω-contractive, it would then follow that

$$\|R_\lambda(A)g\|_X \leq \|g\|_X \int_0^\infty e^{(\omega-\lambda)t}dt = \frac{\|g\|_X}{\lambda - \omega},$$

for all $\lambda > \omega$ and $g \in X$. Of course, to make the above rigorous we must show that $(\omega, \infty) \subset \rho(A)$ and that the identity (7) holds. This, along with the sufficiency of conditions (i)-(ii) in Theorem 3, will be established in the Appendix.

Equipped with Theorem 3, we can now use semigroup methods to solve linear evolution equations. This is demonstrated in the next section, where we consider a generalized wave equation.

2.3 Application: Linear Hyperbolic PDE

Let $\Omega \subset \mathbb{R}^n$ be open and bounded with smooth boundary, and let

$$Lu = -\sum_{i,j=1}^n \left(a^{i,j}u_{x_i}\right)_{x_j}$$

be uniformly elliptic on Ω with $a^{i,j} = a^{j,i} \in C^1(\overline{\Omega})$ being time independent. Now, fix $T > 0$ and consider the linear hyperbolic IVBVP

$$\begin{cases} u_{tt} + Lu = 0, \quad x \in \Omega,\ t \in (0,T) \\ u = 0 \ \text{ on } \partial\Omega \times [0,T) \\ u = g, \quad u_t = h \ \text{ on } \Omega \times \{t = 0\}, \end{cases} \tag{8}$$

for $g, h \in L^2(\Omega)$ fixed. In order to approach this via semigroup theory, we first rewrite (8) as the equivalent first order system

$$\begin{pmatrix} u \\ v \end{pmatrix}_t = \begin{pmatrix} v \\ -Lu \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -L & 0 \end{pmatrix}}_{=:A} \begin{pmatrix} u \\ v \end{pmatrix} \tag{9}$$

with boundary conditions

$$\begin{cases} u = 0 & \text{on } \partial\Omega \times [0, T) \\ u = g, \quad v = h & \text{on } \Omega \times \{t = 0\}, \end{cases}$$

and we consider the above dynamical system on the infinite dimensional phase space

$$X = H_0^1(\Omega) \times L^2(\Omega)$$

with norm

$$\|(u, v)\| := \left(B[u, u] + \|v\|_{L^2(\Omega)}^2 \right)^{1/2},$$

where B is the bilinaer form associated to L. In particular, we consider the operator $A : X \to X$ as an densely defiend linaer operator with domain

$$D(A) = \left(H^2(\Omega) \times H_0^1(\Omega) \right) \times H_0^1(\Omega).$$

Using Hille-Yosida, we can establish the existence of classical (in time) solutions of (9) when $(g, h) \in D(A)$.

Theorem 4. *The operator A defined above generates a C_0-semigroup on X.*

Proof. The strategy is to verify the hypotheses of Hille-Yosida. Note that $D(A)$ is clearly dense in X. To see that A is closed, suppose $\{(u_k, v_k)\}$ is a sequence in $D(A)$ such that $(u_k, v_k) \to (u, v) \in X$ and $A \begin{pmatrix} u_k \\ v_k \end{pmatrix} \to \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ in X. Clearly then we have $v_k \to v$ in $L^2(\Omega)$ and, by the definition of A, $v_k \to f_1$ in $H_0^1(\Omega)$. By uniqueness of limits, it follows that $v = f_1$ so that, in particular, $v \in H_0^1(\Omega)$. Further, observe by elliptic (boundary) regularity theory[8] that

$$\|u_k - u_\ell\|_{H^2(\Omega)} \leq C \left(\|Lu_k - Lu_\ell\|_{L^2(\Omega)} + \|u_k - u_\ell\|_{L^2(\Omega)} \right).$$

Since $u_k \to u$ in $H_0^1(\Omega)$ and $Lu_k \to -f_2$ in $L^2(\Omega)$, it follows from above that the sequence $\{u_k\}$ is Cauchy in $H^2(\Omega)$ and hence, by uniqueness of limits, we have $u_k \to u$ in $H^2(\Omega) \cap H_0^1(\Omega)$. Thus, $(u, v) \in D(A)$ and

$$A \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v \\ -Lu \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}.$$

[8] See, for example, Theorem 4 in Section 6.3.2 in Evans.

It follows that A is a closed operator, as desired.

It remains to study the resolvent of A. To this end, let $\lambda > 0$ and $(f, g) \in X$ be given and consider the equation

$$\lambda \begin{pmatrix} u \\ v \end{pmatrix} - A \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}. \tag{10}$$

Note the above is equivalent to the system

$$\begin{cases} \lambda u - v = f \\ \lambda v + Lu = g \end{cases} \tag{11}$$

with $(u, v) \in D(A)$ which, in turn, is equivalent to the scalar equation

$$Lu + \lambda^2 u = \lambda f + g, \quad u \in H^2(\Omega) \cap H_0^1(\Omega). \tag{12}$$

Now, recall from elliptic existence theory that the operator L has strictly positive spectrum. Since $\lambda^2 > 0$, it follows that (12) has a unique solution $u \in H^2(\Omega) \cap H^1(\Omega)$ which, setting

$$v = \lambda u - f \in H_0^1(\Omega), \tag{13}$$

it follows that $(u, v) \in D(A)$ is the unique weak solution of (10). Since $(f, g) \in X$ and $\lambda > 0$ was arbitrary, it follows that $(0, \infty) \in \rho(A)$. Furthermore, note that from the second equation in (11) that

$$\lambda \|v\|_{L^2(\Omega)}^2 + B[u, v] = \langle g, v \rangle_{L^2(\Omega)},$$

where here $\begin{pmatrix} u \\ v \end{pmatrix} = R_\lambda(A) \begin{pmatrix} f \\ g \end{pmatrix}$. It follows that

$$\begin{aligned} \lambda \|(u, v)\|_X^2 &= \lambda \left(\|v\|_{L^2(\Omega)}^2 + B[u, u] \right) \\ &= \langle g, v \rangle_{L^2(\Omega)}^2 + B[u, \lambda u - v] \\ &= \langle g, v \rangle_{L^2(\Omega)}^2 + B[u, f], \end{aligned}$$

where the last equality follows from (13). Recalling the inner-product structure on X and using Cauchy-Schwartz, it follows that

$$\lambda \|(u, v)\|_X^2 = \left\langle \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} f \\ g \end{pmatrix} \right\rangle_X \leq \|(u, v)\|_X \|(f, g)\|_X$$

and hence

$$\|(u, v)\|_X \leq \frac{1}{\lambda} \|(f, g)\|_X, \quad \text{i.e. } \|R_\lambda(A)\| \leq \frac{1}{\lambda} \text{ for all } \lambda > 0.$$

The result now follows from the Hille-Yosida theorem. $\square$

By the above theorem, it follows that for all $(g,h) \in D(A)$ the linear hyperbolic IVBVP (9) admits a classical solution, i.e. it admits a solution

$$(u,v) \in C([0,\infty); D(A)) \cap C^1((0,\infty); X)$$

such that $(u(t), v(t)) \in D(A)$ for all $t \geq 0$ and it satisfies (9) pointwise for all $t > 0$. Furthermore, note that this solution is unique. Indeed, define the energy

$$E(t) = \frac{1}{2} \int_\Omega \left(v^2 + \sum_{i,j=1}^n a^{i,j} u_{x_i} u_{x_j} \right) dx$$

and note that for (u,v) as above we have $E \in C^1(0,\infty)$ with

$$E'(t) = \int_\Omega v\,(v_t + Lu)\,dx = 0 \quad \text{for all } t > 0.$$

It follows that if $(u,v)(0) = (0,0)$ then $E(t) = 0$ for all $t > 0$, and hence the unique solution of (9) is $(u,v) = (0,0)$. By linearity, it follows that classical (in the above sense) solutions of (9) are unique.

Before continuing, we point out that there are other sets of necessary and sufficient conditions for a given linear operator to be the generator of a C_0-semigroup. One commonly used result on Hilbert spaces is the following.

Theorem 5 (Lumer-Phillips)**.** *Let H be Hilbert space and let A be a linear operator on H that satisfies the following conditions:*

(i) $D(A)$ is dense in X.

(ii) There exists a constant $\omega \in \mathbb{R}$ such that $\operatorname{Re}\langle x, Ax\rangle \leq \omega \|x\|^2$ for all $x \in H$.

(iii) There exists a $\lambda_0 > \omega$ such that $\lambda_0 I - A$ is surjective.

Then A is the generator of a C_0-semigroup on X which is ω-contractive.

From Lumer-Phillips, it is evident that any self-adjoint operator whose spectrum is bounded from above generates a C_0-semigroup. Similarly, any skew-adjoint operator generates a C_0-semigroup of contractions. As an exercise, students should attempt to reprove Theorem 4 above using Lumer-Phillips.

3 Nonlinear Evolution Equations

So far, the application of semigroup methods has been restricted to the case of linear operators. This is of course natural since, by construction, semigroups are designed as a tool to solve linear equations. However, recalling our general methodologies from Chapter 4 (on fixed point methods), it makes sense that such techniques may be combined with appropriate fixed point theorems in order to provide tools for studying nonlinear equations. This is precisely the goal of this section! As a first step, however, we must study how semigruop methods can be extended to nonhomogeneous linear problems.

3.1 Nonhomogeneous Linear Problems & Mild Solutions

Let X be a Banach space and suppose $A : D(A) \subset X \to X$ is a linear operators. Our goal is to consider nonhomogeneous linear IVPs of the form

$$\begin{cases} u_t = Au + f(t), & t > 0 \\ u(0) = u_0 \end{cases} \tag{14}$$

where here $u_0 \in X$ and $f \in C(\mathbb{R}; X)$ are given. As in previous sections, given $u_0 \in D(A)$ we say u is a classical solution of (14) if

$$u \in C([0,\infty); D(A)) \cap C^1((0,\infty); X)$$

and u satisfies (14) pointwise for all $t \in \mathbb{R}$.

To begin, note that if A generates a C_0-semigroup $\{T(t)\}_{t\geq 0}$ on X, then any classical solution of (14) can be represented as

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)ds. \tag{15}$$

Indeed, for fixed $t > 0$ define for $s \in [0,t]$ the function $g(s) = T(t-s)u(s)$ and note that

$$\begin{aligned} \frac{dg}{ds} &= -AT(t-s)u(s) + T(t-s)\frac{du}{ds}(s) \\ &= -AT(t-s)u(s) + T(t-s)\left(Au(s) + f(s)\right) \\ &= T(t-s)f(s) \end{aligned}$$

so that, by the Fundamental Theorem of Calculus,

$$g(t) = g(0) + \int_0^t T(t-s)f(s)ds,$$

which is equivalent to (15). Note that (15) is simply an abstract version of the well-known Duhamel (or variational of constants) formula from elementary ODE.

Now, one can easily check that if $A \in \mathcal{L}(X)$, so that $\{T(t)\}_{t\geq 0}$ is a uniformly continuous semigroup of operators on X, then u defined by (15) is a classical solution of the nonhomogeneous IVBVP (14). Unfortunately, however, if A only generates a C_0-semigroup, then (15) may not define a function $u(t)$ that is differentiable at any $t > 0$, even if $f \in C(\mathbb{R}; X)$. This is illustrated in the following example.

Example: Let $\{T(t)\}_{t\geq 0}$ be a C_0-semigroup on a Banach space X with generator $A : D(A) \subset X \to X$, and suppose there exists[9] a $g_0 \in X$ such that $T(t)g_0 \notin D(A)$ for all $t > 0$. Taking $f(t) = T(t)g_0$ and $u_0 = 0$ in (15) we obtain the function

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

[9] Think of the wave equation, for example. The hyperbolic nature of the equation implies propagation of singularities, implying solutions can only be as smooth as the initial data.

By the semigroup property, the above is equivalent to

$$u(t) = \int_0^t T(t)g_0 \, ds = tT(t)g_0$$

which is clearly continuous in t, but not differentiable at any $t > 0$ since $T(t)g_0$ is differentiable at t_0 if and only if $T(t_0)g_0 \in D(A)$. Indeed, note that for all $h > 0$ we have

$$\frac{T(t_0 + h)g_0 - T(t_0)g_0}{h} = \frac{T(h)T(t_0)g_0 - T(t_0)g_0}{h}$$

so that the limit as $h \to 0^+$ above exists if and only if $T(t_0)g_0 \in D(A)$, as claimed.

Remark 1. *Note that for parabolic PDE, the associated semigroups satisfy $T(t)g \in D(A)$ for all $t > 0$ and or any $g \in X$. These are examples of so-called "analytic" semigroups, which are* extremely *important in applications but not be discussed here.*

While the above example shows that the representation formula (15) may not always define a classical solution of the IVP (14), it is true that every classical solution must be represented in the form (15). Furthermore, we note that the representation formula (15) makes sense under the much weaker hypotheses that $u_0 \in X$ and $f \in L^1(0, \infty); X)$, and that, even under these weaker assumptions, $u(t)$ defined in (15) is still in $C([0, \infty); X)$ and satisfies the initial condition $u(0) = u_0$. With this in mind, we make the following definition.

Definition 5. *Suppose that A generates a C_0-semigroup on a Banach space X. Given $u_0 \in X$ and $f \in L^1((0, \infty); X)$ the function $u(t)$ defined in* (15) *is called a mild solution of the IVP* (14)*.*

Following our general methodology of this class, we will henceforth consider mild solutions as actual "solutions" of (14). Of course, it is natural to ask when a mild solution actually corresponds to a classical solution of (14). Such a regularity result is the content of the following result.

Theorem 6. *Suppose that A generates a C_0-semigroup on a Banach space X. Assume that $u_0 \in D(A)$ $f \in C([0, \infty); X)$ and that, additionally, f satisfies either*

(i) $f \in W^{1,1}((0, \infty); X)$

or

(ii) $f \in L^1((0, \infty); D(A))$.

Then (15) *defines a function $u \in C^1((0, \infty); X)$ which a classical solution of* (14)*.*

For a proof of this result, see Theorem 2 in Section 9.2(c) in McOwen or Theorem 12.16 in Section 12.1.3 in Renardy and Rogers. In summary, we have that if A generates a C_0-semigroup $\{T(t)\}_{t \geq 0}$ on a Banach space X, then given any $f \in C([0, \infty); X)$ and $u_0 \in X$ the formula (15) provides a mild solution of the nonhomogeneous IVP (14) which, under appropriate additional assumptions on u_0 and f, actually solves the IVP classically.

3.2 Mild Solutions of Nonlinear Problems

Our next goal is to use the theory developed in the previous section in conjunction with the Contraction Mapping Theorem to study the existence of solutions to nonlinear evolution equations. To this end, let X be a Banach space and suppose $A : D(A) \subset X \to X$ is a closed, densely defined linear operator on X. Consider the nonlinear IVP

$$\begin{cases} u_t = Au + f(u), \quad t > 0 \\ u(0) = u_0, \end{cases} \tag{16}$$

where here $u_0 \in X$ and $f : X \to X$ is a continuous, possibly nonlinear, map. In the study of such equations, there are at least four key issues to address.

(i) Local Existence: Show there exists a unique solution for $t \in (0, \tau)$, provided that $\tau > 0$ is sufficiently small.

(ii) Global Existence vs. Blow Up: Does the local solution exist for all $t > 0$, or does it have some finite time of existence $T > 0$? If $T < \infty$, what happens when $t \to T^-$?

(iii) Continuous Dependence: Does the solution depend continuously on u_0?

(iv) Asymptotic Properties: What happens to global solutions when $t \to \infty$?

Note that each of the above issues typically requires analysis specifically designed for the particular nonlinear dynamics involved. In particular, people working on different classes of PDEs (such as Hamiltonian, dissipative, hyperbolic, etc.) will likely have different techniques for addressing each of the above issues. Nevertheless, we can derive some general results for (i) and (ii) above that can be fine-tuned for specific applications.

To begin, suppose A generates a C_0-semigroup $\{T(t)\}_{t\geq 0}$ on X. From our work in the previous section, we know that if (16) admits a classical solution u then it must satisfy the implicit integral equation

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(u(s))ds \tag{17}$$

for the entire time of existence. With this in mind, we say that any solution $u \in C([0, \tau); X)$ of (17) is a mild solution of (16) for $t \in [0, \tau]$. We now establish the following fundamental local existence and uniqueness result.

Theorem 7. *Let X be a Banach space and let $f : X \to X$ be locally Lipschitz, i.e. for all $R > 0$ there exists a constant $M = M(R) > 0$ such that*

$$\|f(u) - f(v)\|_X \leq M\|u - v\|_X \quad \text{for all } \|u\|_X,\ \|v\|_X \leq R.$$

If A is the generator of a C_0-semigroup on X, then for all $u_0 \in X$ there exists $t_{\max} \in (0, \infty]$ such that the IVP

$$\begin{cases} u_t = Au + f(u), \quad t > 0 \\ u(0) = u_0 \end{cases} \tag{18}$$

has a unique mild solution $u \in C([0, t_{\max}); X)$*. Moreover, if* $t_{\max} < \infty$ *then we necessarily have*

$$\lim_{t \to t_{\max}^-} \|u(t)\|_X = \infty.$$

Proof. Without loss of generality, assume that A generates a C_0-semigroup of contractions $\{T(t)\}_{t \geq 0}$. For each finite $\tau > 0$, define

$$Y = C([0, \tau]; X)$$

equipped with the natural norm

$$\|u\|_Y = \max_{0 \leq t \leq \tau} \|u(t)\|_X$$

and note that $(Y, \|\cdot\|_Y)$ is a Banach space. Now, fix $u_0 \in X$ and consider the nonlinear map $F : Y \to Y$ defined by

$$F(u)(t) = T(t)u_0 + \int_0^t T(t-s) f(u(s)) ds$$

and note that a mild solution of the IVP (18) on $[0, \tau]$ corresponds to a fixed point of F in Y.

To prove that F has a fixed point in Y, we aim to use the Contraction Mapping Theorem. To this end, let $R = 2\|u_0\|_X$ and define the ball

$$W_R := \{u \in Y : \|u\|_Y < R\}.$$

Given $u \in W_R$, it follows that for all $t \in [0, \tau]$ we have by the triangle inequality and the assumption that $\|T(t)\| \leq 1$ the bound

$$\begin{aligned}
\|F(u)(t)\|_X &\leq \|T(t)u_0\|_X + \int_0^t \|T(t-s) f(u(s))\|_X ds \\
&\leq \|u_0\|_X + \int_0^t \|f(u(s))\|_X ds
\end{aligned}$$

Since f is locally Lipschitz, we know that

$$\|f(z)\|_X \leq \|f(z) - f(0)\|_X + \|f(0)\|_X \leq M\|z\|_X + \|f(0)\|_X$$

for all $z \in X$, and hence, combining with the above estimate, we have that

$$\|F(u)(t)\|_X \leq \|u_0\|_X + \tau \left(MR + \|f(0)\|_X\right)$$

for all $u \in W_R$ and $t \in [0, \tau]$. Similarly, for all $u, v \in W_R$ and $t \in [0, \tau]$ we have

$$\begin{aligned}
\|F(u)(t) - F(v)(t)\|_X &\leq \int_0^t \left\|T(t-s)\left(f(u(s)) - f(v(s))\right)\right\|_X \, ds \\
&\leq M \int_0^t \|u(s) - v(s)\|_X ds \\
&\leq M\tau \|u - v\|_Y.
\end{aligned}$$

Setting

$$\tau_1 := \min\left\{\frac{R-\|u_0\|_X}{MR+\|f(0)\|_X}, \frac{1}{2M}\right\}$$

it follows that taking $\tau = \tau_1$ above we have that $F : W_R \to W_R$ and that F is a strict contraction on W_R. By the Contraction Mapping Theorem, it follows there exists a unique $u \in W_R = W_R(\tau_1)$ such that $F(u) = u$. By construction, this fixed point corresponds to a mild solution of the IVP (18) on the time interval $[0, \tau_1]$.

Note by above that the mild solution u on $[0, \tau_1]$ can be extended to a mild solution on $[0, \tau_1 + \tau_2]$ for some $\tau_2 > 0$ by defining $u(t) = w(t)$ on $[\tau_1, \tau_1 + \tau_2]$ where $w(t)$ solves

$$w(t) = T(t-\tau_1)u(\tau_1) + \int_{\tau_1}^{t} T(t-s)f(w(s))ds$$

for $\tau_1 \le t \le \tau_1 + \tau_2$. Notice that $\tau_2 = \tau_2(\|u(\tau_2)\|_X, \|f(u(\tau_2))\|_X)$. Continuing to extend the mild solution u as above, can define $[0, t_{\max})$ to be the maximal interval of existence of the mild solution u of the IVP.

Next, I claim that if $t_{\max} < \infty$, then $\lim_{t\to t_{\max}^-} \|u(t)\|_X = \infty$. If this were false, then there would exist a sequence of times $\{t_n\}_{n=1}^{\infty}$ with $t_n \nearrow t_{\max}$ and $\|u(t_n)\|_X \le C$ for all $n \in \mathbb{N}$. By above, it would follow that for all n sufficiently large then u defined on $[0, t_n]$ could be extended to $[0, t_n + \delta]$ where now, since the $C > 0$ above is uniform in n, the $\delta > 0$ can be chosen independent of n. Thus, u could be extended beyond $t_{\max}$, which is a contradiction. This verifies the "blow-up" alternative.

Finally, it remains to establish uniqueness of the mild solution. Suppose that u and v are both mild solutions of the IVP with initial data $u_0, v_0 \in X$, respectively. Then so long as both $u(t)$ and $v(t)$ are defined we have

$$\begin{aligned}\|u(t)-v(t)\|_X &\le \|T(t)(u_0-v_0)\|_X + \int_0^t \|T(t-s)\left(f(u(s))-f(v(s))\right)\|_X ds \\ &\le \|u_0-v_0\|_X + \widetilde{M}\int_0^t \|u(s)-v(s)\|_X ds\end{aligned}$$

for some constant $\widetilde{M} > 0$. By Gronwall's inequality[10] it follows that if both $u(t)$ and $v(t)$ exist for $t \in [0, T]$ then we have

$$\|u(t)-v(t)\|_X \le e^{\widetilde{M}(T-t)}\|u_0-v_0\|_X \quad \text{for all } t \in [0, T].$$

Uniquenss now follows on any closed subset $[0, T]$ of times for which both $u(t)$ and $v(t)$ exist. Thus, if $u_0 = v_0$ then u and v have the same $t_{\max}$ and $u(t) = v(t)$ for all $t \in [0, t_{\max})$, completing the proof. □

Before we continue, we establish some important remarks. First, note by the blow-up alternative that if one can establish an a-priori bound of the form

$$\|u(t)\|_X \le K$$

[10] See, for example, Appendix B in Evans.

for the time of existence of a solution u, it follows from above that the mild solutions will necessarily be global, i.e. they will have $t_{\max} = \infty$. The development of such a-priori bounds can be quite delicate, and often relies on subtle energy estimates. Secondly, the above proof can be modified to show that if in fact the function f were *globally* Lipschitz on X, then the mild solutions will again exist globally in time: see the exercises. We further note that if the nonlinearity f in Theorem 7 is $C^1(X;X)$, then mild solutions with $u_0 \in D(A)$ are in fact classical solutions.

Finally, I wish to emphasize that in most PDE applications, even the local Lipschitz property for f fails to hold on all of X. In the next section, we apply the general local existence and uniqueness result in Theorem 7 to one of the most commonly studied PDEs in mathematical physics and, in particular, we will see how the lack of Lipschitz continuity of f on the entire Banach space X can be handled.

3.3 Application: Cubic Nonlinear Schrödinger Equation on $\mathbb{R}^3$

Consider the following initial value problem for the cubic Nonlinear Schrödinger equation posed on $\mathbb{R}^3$:

$$\begin{cases} iu_t = -\Delta u + k|u|^2 u, \quad u > 0, \quad x \in \mathbb{R}^3 \\ u(0) = u_0 \end{cases} \tag{19}$$

where here $u : \mathbb{R}^3 \times [0, \infty) \to \mathbb{C}$, $k = \pm 1$ is a constant, and u_0 is some initial data. Note that the PDE in (19) can be rewritten in the abstract form

$$u_t = Au + f(u), \quad t > 0$$

where here $A : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the linear operator $A = i\Delta$ with densely defined domain $D(A) = H^2(\mathbb{R}^3)$ and $f(u) = -ik|u|^2 u$.

Lemma 4. *The operator A generates a C_0-semigroup of contractions on $L^2(\mathbb{R}^3)$.*

"Proof". We will prove this by applying the Hille-Yosida theorem. To this end, first note that since

$$\Delta : H^2(\mathbb{R}^3) \subset L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$$

is closed and densely defined, so is A. Furthermore, using the Fourier transform one can show that[11]

$$\sigma(A) = \left\{ik^2 : k \in \mathbb{R}\right\} \subset \mathbb{R}i$$

so that, specifically, $(0, \infty) \subset \rho(A)$. Furthermore, for all $\lambda > 0$ and $g \in L^2(\mathbb{R}^3)$, note that if

$$u = R_\lambda(A)g, \quad \text{i.e. if} \quad (A - \lambda I)\, u = g,$$

[11] Indeed, the Fourier transform maps differential operator in the spatial domain to multiplication operators in the frequency domain. In this context, we have $\widehat{\Delta f}(k) = -k^2 \hat{f}(k)$. Since the Fourier transform is an isometry of $L^2(\mathbb{R}^3)$, it follows that the spectrum of Δ is given by $\sigma(-k^2\cdot)$, where here the operator $-k^2\cdot$ is a multiplication operator on $L^2(\mathbb{R}^3)$. Since the spectrum of a multiplication operator with a piecewise strictly monotone "symbol" (in this case, the symbol is $-k^2$) is simply the range of the symbol, the result follows. For more details, take Math 890 and Math 960.

then multiplying by $\bar{u}$ and integrating gives

$$\lambda \int_{\mathbb{R}^3} |u|^2 dx = -i \int_{\mathbb{R}^3} |Du|^2 dx - \int_{\mathbb{R}^3} \bar{u} g \; dx.$$

Since $\lambda \in \mathbb{R}$, taking the real parts of the above identity and using Cauchy-Schwartz yields

$$\lambda \int_{\mathbb{R}^3} |u|^2 dx \leq \|u\|_{L^2(\mathbb{R}^3)} \|g\|_{L^2(\mathbb{R}^3)}$$

which is clearly equivalent to

$$\|R_\lambda(A) g\|_{L^2(\mathbb{R}^3)} \leq \frac{\|g\|_{L^2(\mathbb{R}^3)}}{\lambda}.$$

The Lemma now follows from Hille-Yosida. □

Remark 2. *Alternatively, one can prove the above result using Stone's Theorem, which states that if $\mathcal{A}$ is closed, densely defined and self-adjoint operator on a Hilbert space H, then the operator $i\mathcal{A}$ is the generator of a unitary group $\{e^{i\mathcal{A}t}\}_{t \in \mathbb{R}}$ of operators on $L^2(\mathbb{R}^3)$. For details, see Math 960.*

To be able to invoke our general local existence result in Theorem 7, it remains to study the nonlinearity $f(u) = -ik|u|^2 u$ on $L^2(\mathbb{R}^3)$. First, note that for a given $u \in L^2(\mathbb{R}^3)$ the function $f(u)$ does not even make sense in $L^2(\mathbb{R}^3)$. Indeed, recalling $k = \pm 1$ we clearly have

$$\|f(u)\|_{L^2(\mathbb{R}^3)} = \|u\|_{L^6(\mathbb{R}^3)},$$

so that f is note even defined, let alone Lipschitz continuous, on $L^2(\mathbb{R}^3)$. Note, however, that since $2 > \frac{3}{2}$ we know from Sobolev Embedding that

$$D(A) = H^2(\mathbb{R}^3) \subset L^\infty(\mathbb{R}^3)$$

and that, in particular, this embedding is continuous. It follows that for $u \in D(A)$ we have

$$\|f(u)\|^2_{L^2(\mathbb{R}^3)} = \int_{\mathbb{R}^3} |u|^6 dx \leq \|u\|^4_{L^\infty(\mathbb{R}^3)} \int_{\mathbb{R}^3} |u|^2 dx \leq \|u\|^6_{H^2(\mathbb{R}^3)}$$

so that the mapping $f : D(A) \to L^2(\mathbb{R}^3)$ is well-defined. The next result shows that, in fact, f is locally Lipschitz continuous as a map from $D(A)$ into itself.

Lemma 5. *The function $f : H^2(\mathbb{R}^3) \to H^2(\mathbb{R}^3)$ given by $f(u) = -ik|u|^2 u$ is well-defined. Furthermore, there exists a constant $C > 0$ such that*

$$\|f(u) - f(v)\|_{H^2(\mathbb{R}^3)} \leq C \left(\|u\|^2_{H^2(\mathbb{R}^3)} + \|v\|^2_{H^2(\mathbb{R}^3)} \right) \|u - v\|_{H^2(\mathbb{R}^3)}.$$

for all $u, v \in H^2(\mathbb{R}^3)$.

Proof. We have already seen that $f : H^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ is well-defined. Given $u \in H^2(\mathbb{R}^3)$, we now refine this result to show that $f(u) \in H^2(\mathbb{R}^3)$. To this end, observe that

$$\begin{aligned}
\|Df(u)\|^2_{L^2(\mathbb{R}^3)} &\leq \int_{\mathbb{R}^3} |D(\bar{u}u)|^2 \, |u|^2 dx + \int_{\mathbb{R}^3} |u|^4 |Du|^2 dx \\
&\leq \left(\int_{\mathbb{R}^3} |D(\bar{u}u)|^4 dx\right)^{1/2} \left(\int_{\mathbb{R}^3} |u|^4 dx\right)^{1/2} + \|u\|^4_{L^\infty(\mathbb{R}^3)} \int_{\mathbb{R}^3} |Du|^2 dx.
\end{aligned}$$

Now, since $H^2(\mathbb{R}^3)$ is continuously embedded in $W^{1,4}(\mathbb{R}^3)$ by Sobolev embedding, we know there exists a constant $C > 0$ such that

$$\|u\|_{L^4(\mathbb{R}^3)} \leq C\|u\|_{H^2(\mathbb{R}^3)}$$

and similarly, using the product rule,

$$\|D(\bar{u}u)\|_{L^4(\mathbb{R}^3)} \leq 2\|u\|_{L^\infty(\mathbb{R}^3)} \|Du\|_{L^4(\mathbb{R}^3)} \leq C\|u\|^2_{H^2(\mathbb{R}^3)}.$$

All together it follows that if $u \in H^2(\mathbb{R}^3)$ then

$$\|Df(u)\|^2_{L^2(\mathbb{R}^3)} \leq C\|u\|^6_{H^2(\mathbb{R}^3)}$$

for some constant $C > 0$. Finally, note that

$$\|D^2 f(u)\|^2_{L^2(\mathbb{R}^3)} \leq \int_{\mathbb{R}^3)} \left(|D^2(\bar{u}u)|^2 |u|^2 + 2|D(\bar{u}u)|^2 |Du|^2 + |u|^4 |D^2 u|^2\right) dx.$$

Using similar calculations to above, we can estimate the last two terms as

$$\int_{\mathbb{R}^3} |u|^2 |D^2 u|^2 dx \leq \|u\|^4_{L^\infty(\mathbb{R}^3)} \|D^2 u\|^2_{L^2(\mathbb{R}^3)} \leq C\|u\|^6_{H^2(\mathbb{R}^3)}$$

and

$$\int_{\mathbb{R}^3} |D(\bar{u}u)|^2 |Du|^2 dx \leq \left(\int_{\mathbb{R}^3} |D(\bar{u}u)|^4\right)^{1/2} \left(\int_{\mathbb{R}^3} |Du|^4 dx\right)^{1/2} \leq C\|u\|^6_{H^2(\mathbb{R}^3)}.$$

For the final term observe that

$$\begin{aligned}
\int_{\mathbb{R}^3} |D^2(\bar{u}u)|^2 |u|^2 dx &\leq \int_{\mathbb{R}^3)} \left(2|D^2 u|^2 |u|^4 + |Du|^4 |u|^2\right) dx \\
&\leq C\left(\|u\|^4_{L^\infty(\mathbb{R}^3)} \|D^2 u\|^2_{L^2(\mathbb{R}^3)} + \|u\|^2_{L^\infty(\mathbb{R}^3)} \|Du\|^4_{L^4(\mathbb{R}^3)}\right) \\
&\leq C\|u\|^6_{H^2(\mathbb{R}^3)}.
\end{aligned}$$

All together, the above calculations show that

$$\|f(u)\|_{H^2(\mathbb{R}^3)} \leq C\|u\|^3_{H^2(\mathbb{R}^3)} \quad \text{for all } u \in H^2(\mathbb{R}^3).$$

so that, in particular, the map

$$f : H^2(\mathbb{R}^3) \to H^2(\mathbb{R}^3)$$

is well-defined.

The calculations to show that f is in fact Lipschitz continuous from $D(A)$ into itself are similar. For example, given $u, v \in H^2(\mathbb{R}^3)$ we have

$$\begin{aligned}
\|f(u) - f(v)\|_{L^2(\mathbb{R}^3)} &= \left\| |u|^2 u - |v|^2 v \right\|_{L^2(\mathbb{R}^3)} \\
&= \left\| |u|^2 (u - v) + (|u|^2 - |v|^2) v \right\|_{L^2(\mathbb{R}^3)} \\
&= \left\| |u|^2 (u - v) + |u - v|(|u| + |v|) v \right\|_{L^2(\mathbb{R}^3)}
\end{aligned}$$

which, by Cauchy-Schwartz, gives

$$\begin{aligned}
\|f(u) - f(v)\|_{L^2(\mathbb{R}^3)} &\leq \|u\|^2_{L^4(\mathbb{R}^3)} \|u - v\|_{L^2(\mathbb{R}^3)} \\
&\quad + \|u - v\|_{L^2(\mathbb{R}^3)} \left(\|u\|_{L^4(\mathbb{R}^3)} \|v\|_{L^4(\mathbb{R}^3)} + \|v\|^2_{L^4(\mathbb{R}^3)} \right) \\
&\leq C \left(\|u\|^2_{H^2(\mathbb{R}^3)} + \|v\|^2_{H^2(\mathbb{R}^3)} \right) \|u - v\|_{L^2(\mathbb{R}^3)},
\end{aligned}$$

where the last inequality follows by Sobolev embedding. The remaining inequalities are left as an exercise. □

Unfortunately, since the map f is not locally Lipschitz continuous on all of $L^2(\mathbb{R}^2)$, the local existence and uniquness result in Theorem 7 does not apply to the IVP (19) posed on $L^2(\mathbb{R}^3)$. However, observe that the domain $D(A) = H^2(\mathbb{R}^3)$ equipped with the standard H^2 norm is clearly a Banach space. Furthermore, since $D(A) \subset L^2(\mathbb{R}^3)$ we clearly have from Lemma 4 that

$$T(t) : D(A) \to D(A)$$

and that $\{T(t)\}_{t \geq 0}$ defines a C_0-semigroup on the Banach space $D(A)$. Since Lemma 5 implies f is Locally lipschitz continuous on $D(A)$, we immediately obtain the following local existence result.

Theorem 8 (Local Existence for Cubic NLS)**.** *For every $u_0 \in H^2(\mathbb{R}^3)$ there exists a unique classical solution u of the IVP* (19) *defined for all $t \in [0, T_{\max})$ with the property that either $T_{\max} = \infty$ or else*

$$\lim_{t \to T_{\max}^-} \|u(t)\|_{H^2(\mathbb{R}^3)} = \infty.$$

Consequently, for either the focusing ($k < 0$) or defocusing ($k > 0$) NLS, the IVP (19) has a unique local solution for all initial data in $H^2(\mathbb{R}^3)$. We now attempt to find conditions under which local solutions exists globally in time. In the defocusing case, we have the following result guaranteeing global existence.

Proposition 2. *Let $u_0 \in H^2(\mathbb{R}^3)$ and let $u(t)$ be the unique solution of the IVP* (19) *defined on $[0,T)$. If $k > 0$, i.e. in the ase of the defocusing NLS, then $\|u(t)\|_{H^2(\mathbb{R}^3)}$ is bounded uniformly in time on $[0,T)$.*

The proof of Proposition 2 is primarily based on studying the conservation laws of the (19). However, we also require the following technical result.

Lemma 6. *Let $\{T(t)\}_{t\geq 0}$ be the C_0-semigroup on $L^2(\mathbb{R}^3)$ generated by the operator $A = i\Delta$ in* (19)*. If $2 \leq p \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, then $T(t)$ can be uniquely extended to a bounded linear operator*

$$T(t) \in \mathcal{L}\left(L^q(\mathbb{R}^3); L^p(\mathbb{R}^3)\right)$$

with

$$\|T(t)u_0\|_{L^p(\mathbb{R}^3)} \leq (4\pi t)^{-(2/q-1)}\|u_0\|_{L^q(\mathbb{R}^3)}.$$

Proof. Since we already know $T(t)$ is a C_0-semigroup of contractions on $L^2(\mathbb{R}^3)$, we clearly have the bound

$$\|T(t)u_0\|_{L^2(\mathbb{R}^3)} \leq \|u_0\|_{L^2(\mathbb{R}^3)}$$

for all $u_0 \in L^2(\mathbb{R}^3)$. Further, using the Fourier transform, we obtain an explicit formula for $T(t)$ acting on $L^2(\mathbb{R}^3)$:

$$\left(T(t)u\right)(x) = \frac{1}{4\pi i t}\int_{\mathbb{R}^3} e^{i|x-y|^2/4t}u(y)dy$$

and hence, for $u \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$ we find

$$\|T(t)u\|_{L^\infty(\mathbb{R}^3)} \leq (4\pi t)^{-1}\|u\|_{L^1(\mathbb{R}^3)}.$$

The result now follows by L^p-interpolation and the fact that $L^2(\mathbb{R}^3) \cap L^p(\mathbb{R}^3)$ is dense in $L^p(\mathbb{R}^3)$ for all $p \in (2,\infty]$. □

"Proof" of Of Proposition 2. To see that the H^2-norm of $u(t)$ is bounded, note that since the solution u satisfies the integral equation (15) it follows that

$$\begin{aligned}\|u(t)\|_{H^2(\mathbb{R}^3)} &\leq \|T(t)u_0\|_{H^2(\mathbb{R}^3)} + \int_0^t \|T(t-s)f(u(s))\|_{H^2(\mathbb{R}^3)}\,ds \\ &\leq \|u_0\|_{H^2(\mathbb{R}^3)} + \int_0^t \|u(s)\|^2_{L^\infty(\mathbb{R}^3)}\|u(s)\|_{H^2(\mathbb{R}^2)}ds,\end{aligned}$$

where the last inequality follows since $T(t)$ is a semigroup of contractions on $L^2(\mathbb{R}^3)$. It follows that if we can prove a uniform bound like

$$\|u(t)\|_{L^\infty(\mathbb{R}^3)} \leq K \tag{20}$$

for all $t \in [0,T)$ then by Gronwall's inequality[12] we would have

$$\|u(t)\|_{H^2(\mathbb{R}^3)} \leq \|u_0\|_{H^2(\mathbb{R}^3)}\left(1 + K^2te^{k^2t}\right)$$

[12] See Appendix B.k in Evans.

valid for all $t \in [0,T)$, providing the desired uniform bound on $\|u(t)\|_{H^2(\mathbb{R}^3)}$.

It thus remains to obtain a uniform L^∞-bound of the for (20). To this end, I first claim that $\|u(t)\|_{H^1(\mathbb{R}^3)}$ is bounded on $[0,T)$. Indeed, multiply the PDE

$$iu_t = -\Delta u + k|u|^2 u$$

by $\bar{u}$ and integrating gives

$$\frac{i}{2}\int_{\mathbb{R}^3} \frac{d}{dt}|u|^2 dx = \int_{\mathbb{R}^3} |Du|^2 dx + k\int_{\mathbb{R}^3} |u|^4 dx.$$

Taking real parts gives[13]

$$\frac{d}{dt}\|u(t)\|_{L^2(\mathbb{R}^3)} = 0$$

so that $\|u(t)\|_{L^2(\mathbb{R}^3)} = \|u_0\|_{L^2(\mathbb{R}^3)}$ for all $t \in [0,T)$. Similarly, multiplying the PDE by $\frac{d\bar{u}}{dt}$ and integrating gives[14]

$$\frac{d}{dt}\left(\frac{1}{2}\int_{\mathbb{R}^3} |Du|^2 dx + \frac{k}{4}\int_{\mathbb{R}^3} |u|^4 dx\right) = 0.$$

Since $k > 0$, it follows that

$$\frac{d}{dt}\int_{\mathbb{R}^3} |Du|^2 dx \leq 0$$

for all $t \in (0,T)$ so that, in particular,

$$0 \leq \int_{\mathbb{R}^3} |Du|^2 dx \leq \int_{\mathbb{R}^3} |Du_0|^2 dx$$

for all $t \in [0,T)$. Together with the L^2-conservation above, it follows that $\|u(t)\|_{H^1(\mathbb{R}^3)}$ is uniformly bounded on $[0,T)$, as claimed.

Next, recall by the Gagliardo-Nirenberg-Sobolev inequality that for all $p \in [2,\infty)$ there exists a constant $C = C_p > 0$ such that

$$\|v\|_{L^p(\mathbb{R}^3)} \leq C\|v\|_{H^1(\mathbb{R}^3)}$$

for all $v \in H^1(\mathbb{R}^3)$. In particular, it follows by above that $\|u(t)\|_{L^p(\mathbb{R}^3)}$ is uniformly bounded on $[0,T)$ for every $p \in [2,\infty)$. Fixing $p \in (2,\infty)$ and using again that u satisfies the integral equation (15) it follows that

$$Du(t) = T(t)Du_0 - \int_0^t T(t-s)Df(u(s))ds$$

[13] This is just conservation of mass.

[14] This is actually equivalent to conservation of energy in this case.

and hence, using the GNS inequality and Lemma 6, we have

$$\begin{aligned}\|Du\|_{L^p(\mathbb{R}^3)} &\leq C\|T(t)Du_0\|_{H^2(\mathbb{R}^3)} + \int_0^t \|T(t-s)Df(u(s))\|_{L^p(\mathbb{R}^3)}\,ds \\ &\leq C\|Du_0\|_{H^2(\mathbb{R}^3)} + C\int_0^t (t-s)^{1-2/q}\left\||u(s)|^2|Du(s)|\right\|_{L^q(\mathbb{R}^3)}\,ds,\end{aligned}$$

where here $q = \frac{p}{p-1}$. Now, notice that by Hölder's inequality that

$$\left\||u|^2|Du|\right\|_{L^q(\mathbb{R}^3)} \leq \|u\|^2_{L^r(\mathbb{R}^3)}\|Du\|_{L^2(\mathbb{R}^3)}, \quad r = \frac{4p}{p-2}$$

and that both factors on the right hand side above have already been shown to be uniformly bounded on $[0,T)$. Thus, for all $t \in [0,T)$ we have

$$\|Du(t)\|_{L^p(\mathbb{R}^3)} \leq C\left(\|Du_0\|_{H^2(\mathbb{R}^3)} + \int_0^t (t-s)^{1-2/q}ds\right).$$

Now, note that $\int_0^t (t-s)^{1-2/q}ds$ is finite provided that $1-\frac{2}{q} > -1$, i.e. provided $q = \frac{p}{p-1} > 1$, which clearly holds here since $p > 2$. It follows that $\|u(t)\|_{W^{1,p}(\mathbb{R}^3)}$ is uniformly bounded on $[0,T)$ and hence, since $W^{1,p}(\mathbb{R}^3)$ is continuously embedded in $L^\infty(\mathbb{R}^3)$ for all $p \in [2,\infty)$ by Sobolev embedding, it follows that there exist a constant $K > 0$ such that

$$\|u(t)\|_{L^\infty(\mathbb{R}^3)} \leq K$$

for all $t \in [0,T)$. By the above work, this completes the proof. □

Equipped with Proposition 2, we immediately obtain the following result establishing the global existence of solutions of the defocusing NLS.

Theorem 9 (Global Existence for Cubic Defocusing NLS)**.** *Consider the defocusing NLS, given by*

$$iu_t = -\Delta u + |u|^2 u, \quad u > 0, \quad x \in \mathbb{R}^3 \tag{21}$$

For each $u_0 \in H^2(\mathbb{R}^3)$, there exists a unique solution

$$u \in C([0,\infty); H^2(\mathbb{R}^3)) \cap C^1((0,\infty); L^2(\mathbb{R}^2))$$

of (21) *satisfying $u(0) = u_0$.*

With more work, one can in fact prove that the Cauchy problem for the cubic defocusing NLS (21) is globally well-posed on $H^2(\mathbb{R}^3)$ which, among other things, implies continuous dependence of solutions on the initial data. It should also be noted that *huge* amounts of research have been devoted to showing well-posedness properties of NLS Cauchy problems in $H^s(\mathbb{R}^3)$ for as as small as possible (typically one tries to get $s = -|\alpha|$ for $|\alpha|$ as large as possible). Notice the above arguments, which are based on conservation of the L^2-norm and the "energy" immediately fail in this case since such quantities are no longer even well-defined. In such cases, more advanced tools are needed, which we will not touch upon here.

Finally, we note that while Theorem 8 local existence of $H^2(\mathbb{R}^3)$ solutions holds for both $k = \pm 1$, global existence generally fails for the focusing case $k = -1$. However, recall that the L^2-norm and the "energy"

$$\mathcal{E}(t) = \frac{1}{2}\int_{\mathbb{R}^3} |Du|^2 dx + \frac{k}{4}\int_{\mathbb{R}^3} |u|^4 dx$$

are both conserved quantities (for the time of existence) for H^2 solutions of the NLS. Thus, using the Garliardo-Nirenberg-Sobolev inequality

$$\int_{\mathbb{R}^3} |u|^4 dx \leq \widetilde{C}\left(\int_{\mathbb{R}^3} |u|^2 dx\right)\left(\int_{\mathbb{R}^3} |Du|^2 dx\right) \tag{22}$$

we find that for any local H^2-solution $u(t)$ of (19) with $k = -1$ and $u(0) = u_0 \in H^2(\mathbb{R}^3)$ satisfies

$$\begin{aligned}\frac{1}{2}\|u(t)\|^2_{H^1(\mathbb{R}^3)} &= \frac{1}{2}\|u(t)\|^2_{L^2(\mathbb{R}^3)} + \mathcal{E}(u(t)) + \frac{1}{4}\int_{\mathbb{R}^3} |u(t)|^4 dx \\ &\leq \frac{1}{2}\|u_0\|^2_{L^2(\mathbb{R}^3)} + \mathcal{E}(u_0) + \frac{\widetilde{C}}{4}\|u_0\|^2_{L^2(\mathbb{R}^3)}\|u(t)\|^2_{H^1(\mathbb{R}^3)},\end{aligned}$$

and hence

$$\left(1 - \frac{\widetilde{C}}{2}\|u_0\|^2_{L^2(\mathbb{R}^3)}\right)\|u(t)\|^2_{H^1(\mathbb{R}^3)} \leq \|u_0\|^2_{L^2(\mathbb{R}^3)} + 2\mathcal{E}(u_0)$$

for all $t \in [0, t_{\max})$. Thus, if the initial data $u_0 \in H^2(\mathbb{R}^3)$ is chosen with sufficiently small L^2-norm, then the above calculations yield a uniform bound on $\|u(t)\|_{H^1(\mathbb{R}^3)}$ valid for the entire time of existence of the solution. Following the proof of Theorem 2, one can now prove that $\|u(t)\|_{H^2(\mathbb{R}^3)}$ is also uniformly bounded for all $t \in [0, t_{\max})$, which leads us to the following result.

Theorem 10 (Global Existence for Focusing Cubic NLS)**.** *If $u_0 \in H^2(\mathbb{R}^3)$, then the IVP*

$$\begin{cases} iu_t = -\Delta u - |u|^2 u, \quad u > 0, \quad x \in \mathbb{R}^3 \\ u(0) = u_0 \end{cases} \tag{23}$$

for the focusing NLS has a unique global solution

$$u \in C([0,\infty); H^2(\mathbb{R}^3)) \cap C^1((0,\infty); L^2(\mathbb{R}^3))$$

provided that

$$\|u_0\|_{L^2(\mathbb{R}^3)} < \frac{2}{\widetilde{C}},$$

where $\widetilde{C} > 0$ is the sharp constant in the Gagliardo-Nirenberg-Sobolev inequality (22)*.*

Of course, it is now natural to ask what happens for larger initial data. Does blow-up necessarialy happen? Are there sufficient conditions that either guarantee or prevent blow up? Also, one can attempt to redo the above theory with more general classes of nonlineartiies, such as power-nonlinearities of the form

$$f(u) = k|u|^{p-1}|u|, \quad p > 1$$

or even for more general nonlineartiies which are not homogeneous. While these are all interesting and fun questions to consider, this seems as good of a place as any to stop.

4 Exercises

Complete the following exercises.

1. Consider the operator

$$A : C^1(0,1) \subset L^2(0,1) \to L^2(0,1)$$

   defined by $Af = f'$. Prove that A is not closed. Furthermore, find an appropriate domain $D(A)$ for which the operator

$$A : D(A) \subset L^2(0,1) \to L^2(0,1)$$

   is indeed a closed operator. Prove your claim.

2. For each $t \geq 0$ define the operator $T(t) : L^2(\mathbb{R}) \to \mathbb{R}$ by $T(t)u(x) = u(x+t)$. Prove that $\{T(t) : t \geq 0\}$ defines a C_0-semigroup on $L^2(\mathbb{R})$, and determine its generator A and $D(A)$.

3. Justify equation (24) in the proof of Hille-Yosida. That is, using the fact that A is a closed operator, show that

$$A\widetilde{R}(\lambda)f = \int_0^\infty e^{-\lambda t} AT(t)f \; dt.$$

4. Use the Lumer Phillips Theorem to provide an alternative proof of Theorem 4.

5. Show that if the function f in Theorem 7 is globally Lipschitz on X, then the unique mild solutions constructed there exist globally in time, i.e. they have $t_{\max} = \infty$.

6. Complete the proof of Lemm 5.

5 Appendix: Proof of Hille-Yosida Theorem

In this appendix, we give a proof of the HIlle-Yosida theorem Theorem 3, which we state again here for completeness.

Theorem 11 (Hille-Yosida)**.** *Let A be a linear operator on a Banach space X. Then A is the generator of a C_0-semigroup on X if and only if the following conditions hold:*

(i) $D(A)$ is dense in X and A is a closed operator.

(ii) There exists an $\omega > 0$ such that $(\omega, \infty) \subset \rho(A)$, and

$$\|R_\lambda(A)\| \leq \frac{1}{\lambda - \omega}, \quad \text{for all} \quad \lambda > \omega.$$

Proof. We begin by proving that the conditions (i)-(ii) are necessary. To this end, suppose that A is the infinitesimal generator of a C_0-semigroup $\{T(t)\}_{t\geq 0}$ which is ω-contractive for some $\omega > 0$, and recall from Proposition 1 that we have already seen that (i) holds for A. We now show that $(\omega, \infty) \subset \rho(A)$. Recall from the discussion directly after Theorem 3 that once this is established it we immediately obtain the desired resolvent bound.

To this end, observe that since $T(t)$ is ω-contractive, given any $f \in X$ the Laplace transform of $T(t)f$, defined as

$$\widetilde{R}(\lambda)f := \int_0^\infty e^{-\lambda t} T(t)f \; dt$$

is well-defined for all $\lambda > \omega$. I claim that $\widetilde{R}(\lambda)f \in D(A)$ for all $\lambda > \omega$ and $f \in X$. To see this, note that for $h > 0$, using similar considerations as in the proof of Proposition 1, we have

$$\begin{aligned}
\left(\frac{T(h) - 1}{h}\right) \widetilde{R}(\lambda)f &= \frac{1}{h}\int_h^\infty e^{-\lambda(t-h)} T(t)f \; dt - \frac{1}{h}\int_0^\infty e^{-\lambda t} T(t)f \; dt \\
&= \frac{1}{h}\int_0^\infty \left(e^{-\lambda(t-h)} - e^{-\lambda t}\right) T(t)f \; dt - \frac{1}{h}\int_0^h e^{-\lambda(t-h)} T(t)f \; dt \\
&= \left(\frac{e^{\lambda h} - 1}{h}\right) \int_0^\infty e^{-\lambda t} T(t)f \; dx - \frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda t} T(t)f \; dt
\end{aligned}$$

Taking $h \to 0^+$ it follows that $\widetilde{R}(\lambda)f \in D(A)$ and, in fact,

$$A\widetilde{R}(\lambda)f = \lim_{h\to 0^+}\left(\frac{T(h) - 1}{h}\right)\widetilde{R}(\lambda)f = \lambda\widetilde{R}(\lambda)f - f.$$

Rearranging, we have

$$(\lambda I - A)\,\widetilde{R}(\lambda)f = f$$

for all $f \in X$ and $\lambda > \omega$. It follows that $(\lambda I - A)$ is surjective for all $\lambda > \omega$. To see that it is injective, note that if $f \in D(A)$ then[15]

$$A\widetilde{R}(\lambda)f = \int_0^\infty e^{-\lambda t} AT(t)f\ dt = \int_0^\infty e^{-\lambda t} T(t)Af\ dt = \widetilde{R}(\lambda)Af \tag{24}$$

so that, in particular,

$$\widetilde{R}(\lambda)\,(\lambda I - A)\,f = f$$

for all $f \in D(A)$. Taking $f = 0$, which clearly lies in $D(A)$, it follows that

$$(\lambda I - A)f = 0 \ \Rightarrow\ \widetilde{R}(\lambda)(\lambda I - A)f = 0 \ \Rightarrow\ f = 0$$

and hence that $(\lambda I - A)$ is injective. Recalling the above holds for all $\lambda > \omega$, it follows that $(0, \infty) \subset \rho(A)$ and, in particular,

$$R_\lambda(f) = \widetilde{R}(\lambda)f$$

for all $f \in X$ and $\lambda > \omega$. As in the discussion immediately following the statement of Theorem 3, it immediately follows that for all $f \in X$ we have

$$\|R_\lambda(f)f\|_X \le \|f\|_X \int_0^\infty e^{(\omega-\lambda)t} dt = \frac{\|f\|_X}{\lambda - \omega},$$

as desired.

To finish the proof of Theorem 3, it remains to prove that the sufficiency of (i)-(ii). To this end, note we can assume $\omega = 0$ since, if A satisfies (i)-(ii) above, then $A - \omega I$ satisfies (i)-(ii) with $\omega = 0$. Equivalently, simply observe that substitution $u = e^{\omega t} v$ transforms the evolution equation $u_t = Au$ to $v_t = (A - \omega I)v$. In any case, we now suppose that A is a linear operator on X satisfying (i)-(ii) with $\omega = 0$ and prove that A is the generator for some C_0-semigroup on X. Of course, it is tempting to simply define $T(t) = e^{At}$ via the series definition (1) and be done. We know this is impossible, however, since the series (1) does not converge in $\mathcal{L}(X)$ unless $A \in \mathcal{L}(X)$. To get around this, we first regularize the operator A, i.e. we define a sequence of approximating bounded linear bounded linear operators, for which the series formula (1) makes sense, that converge to A in some appropriate limiting sense.

To this end, for each $\lambda > 0$ define the operator A_λ on X by

$$A_\lambda := \lambda A R_\lambda(A)$$

and note that $A_\lambda \in \mathcal{L}(X)$ since

$$(\lambda I - A)R_\lambda(A) = I \quad \Rightarrow \quad AR_\lambda(A) = \lambda R_\lambda(A) - I. \tag{25}$$

[15] Here, moving A through the integral is justified since A is a closed operator. See the exercises.

Note that, formally, we have $A_\lambda = \frac{\lambda A}{\lambda I - A}$ which, as $\lambda \to \infty$, should converge to A. To make this rigorous, let $f \in D(A)$ and note that since $\|R_\lambda\| \le \frac{1}{\lambda}$ we have that

$$AR_\lambda(A)f = R_\lambda(A)Af \to 0 \quad \text{as } \lambda \to \infty.$$

Recalling (25), it follows that for all $f \in D(A)$ we have

$$\lim_{\lambda\to\infty} \lambda R_\lambda(A)f \to f$$

and hence for all $f \in D(A)$ we have

$$A_\lambda f = \lambda A R_\lambda(A) f = \lambda R_\lambda(A) A f \to Af \tag{26}$$

as $\lambda \to \infty$, as desired.

Now, fix $\lambda > 0$ and define for all $t \ge 0$ the linear linear operator

$$T_\lambda(t) := e^{A_\lambda t}$$

which, recalling (25), can be represented as

$$T_\lambda(t) = e^{-\lambda t} e^{\lambda^2 t R_\lambda(A)} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda^2 t)^k}{k!} R_\lambda(A)^k.$$

In particular, note that since $\|R_\lambda\| \le \frac{1}{\lambda}$ this implies that

$$\|T_\lambda(t)\| \le e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = 1$$

so that for each $\lambda > 0$ the family $\{T_\lambda(t)\}_{t\ge 0}$ is a C_0-semigroup of contractions on X with generator A_λ and $D(A_\lambda) = X$.

I now claim that for each fixed $t \ge 0$ and $f \in D(A)$, the sequence $\{T_\lambda(t)f\}_{\lambda>0}$ is Cauchy in X. To see this, note that for all $\lambda, \mu > 0$ and $f \in X$ we have by the Fundamental Theorem of Calculus that

$$\begin{aligned} T_\lambda(t)f - T_\mu(t)f &= \int_0^t \frac{d}{ds}\left[T_\mu(t-s)T_\lambda(s)f\right] ds \\ &= \int_0^t T_\mu(t-s)T_\lambda(s)\left[A_\lambda f - A_\mu f\right] ds, \end{aligned}$$

where the last equaltiy follows by the product rule and the fact that the operators A_μ and A_λ commute. In particular, using that $\|T_\lambda\| \le 1$ for all $\lambda > 0$ and (26), it follows that

$$\|T_\lambda(t)f - T_\mu(t)f\|_X \le t \left\|A_\lambda f - A_\mu f\right\|_X \to 0$$

as $\mu, \lambda \to \infty$ for all $f \in D(A)$. Therefore, for all $t \ge 0$ we can define the operator

$$T(t) : D(A) \to X$$

by

$$T(t)f := \lim_{\lambda\to\infty} T_\lambda(t)f.$$

Note that since $\|T_\lambda(t)\| \leq 1$, the density of $D(A)$ in X implies that the above limit in fact exists for all $u \in X$, uniformly in t on compact subsets of $[0,\infty)$. Indeed, if $g \in X$ and f_n is a sequence in $D(A)$ with $f_n \to g$ in X as $n \to \infty$, note for all $\lambda, \mu > 0$ that

$$\begin{aligned}\|T_\lambda(t)g - T_\mu(t)g\|_X &\leq \|T_\lambda(t)g - T_\lambda(t)f_n\|_X + \|T_\lambda(t)f_n - T_\mu(t)f_n\|_X \\ &\quad + \|T_\mu(t)f_n - T_\mu(t)g\|_X \\ &\leq 2\|g - f_n\|_X + t\|A_\lambda f_n - A_\mu f_n\|_X.\end{aligned}$$

Thus, the domain of the operator $T(t)$ defined above can be extended to all of X, as claimed. Similarly, it is readily checked that $\{T(t)\}_{t\geq 0}$ is a C_0-semigroup of contractions on X.

It remains to show that A is the generator for $\{T(t)\}_{t\geq 0}$. To this end, let

$$B := \frac{d}{dt}T(t)\big|_{t=0}$$

be the generator for $\{T(t)\}_{t\geq 0}$ and note for all $f \in D(A)$, $t > 0$ and $\lambda > 0$ that

$$\frac{T_\lambda(t)f - f}{t} = \frac{1}{t}\int_0^t \frac{d}{ds}T_\lambda(s)f\ ds = \frac{1}{t}\int_0^t T_\lambda(s)A_\lambda f\ ds.$$

Taking $\lambda \to \infty$ above gives

$$\frac{T(t)f - f}{t} = \frac{1}{t}\int_0^t T(s)Af\ ds$$

for all $t > 0$ and $f \in D(A)$, and now taking $t \to 0^+$ we find that

$$Bf = \lim_{t\to 0^+} \frac{1}{t}\int_0^t T(s)Af\ ds = Af.$$

It follows that $Bf = Af$ for all $f \in D(A)$, and hence that $D(A) \subset D(B)$. For the reverse inclusion, note that[16]

$$(0,\infty) \subset \rho(A) \cap \rho(B)$$

so that for all $\lambda > 0$ we have

$$D(B) = (\lambda I - B)^{-1}(X) = (\lambda I - B)^{-1}(\lambda I - A)(D(A)) = D(A),$$

where the last equality follows since $A = B$ on $D(A)$. Thus, $A = B$ as claimed, which completes the proof. □

[16] This follows by assumption on A and the fact that B is known to be the generator for a C_0-semigroup of contractions.

We end by noting that from above we the semigroup generated by A can be written as

$$T(t) = \lim_{\lambda\to\infty} e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda^2 t)^k}{k!} R_\lambda(A)^k,$$

which in some sense gives us a way of approximating the semigroup through a series of powers of the resolvent. It should be pointed out that there are several other ways of representing the semigroup generated by A, each of which has specific merit and uses. For example, one can show that if A generates a C_0-semigroup on X then

$$T(t) = \lim_{n\to\infty} \left(1 - \frac{t}{n}A\right)^{-n},$$

which corresponds using the implicit Euler scheme

$$\frac{u(t+h) - u(t)}{h} = Au(t+h)$$

to try to solve the evolution equation $u_t = Au$. Additional analytical methods, including using Laplace transforms or Cauchy's integral formula, provide more alternative ways of constructing semigroups and which are useful in various contexts.