CHAPTER 5

The Heat Equation

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 averages 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.

The properties of the heat equation and more general parabolic PDEs parallel those of Laplace's equation and elliptic PDEs. For example, there are parabolic versions of maximum principles, Harnack inequalities, Schauder theory, and Sobolev solutions.

5.1. The initial value problem

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

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

We will solve (5.2) explicitly by use of the Fourier transform, following the presentation in [**15**]. Before doing this, we describe the sense in which we define a solution.

5.1.1. Schwartz solutions. Assume first that the initial data $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth, rapidly decreasing Schwartz function $f \in \mathcal{S}$ (see Section 5.A). 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 the strong derivative of u if we want to emphasize that it is defined as the limit of difference quotients in $\mathcal{S}$.

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

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 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.

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 equal $\Delta u(t)$ where $\Delta : \mathcal{S} \to \mathcal{S}$ is the Laplacian operator.

THEOREM 5.3. *If $f \in \mathcal{S}$, there is a unique solution*

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

of (5.2). Furthermore, $u \in C^\infty\left((0,\infty);\mathcal{S}\right)$ and for $t > 0$ it is given by

$$u(x,t) = \int_{\mathbb{R}^n} \Gamma(x-y,t) f(y)\, dy \tag{5.4}$$

where

$$\Gamma(x,t) = \frac{1}{(4\pi t)^{n/2}}\, e^{-|x|^2/4t}. \tag{5.5}$$

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

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

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

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

$$\hat{u}_t = -|k|^2 \hat{u}, \qquad \hat{u}(0) = \hat{f}. \tag{5.6}$$

This ODE has a unique solution $\hat{u} \in C\left([0,\infty);\mathcal{S}\right) \cap C^\infty\left((0,\infty);\mathcal{S}\right)$ given by

$$\hat{u}(k,t) = \hat{f}(k) e^{-t|k|^2}. \tag{5.7}$$

To prove this in detail, suppose first that u satisfies (5.3). Then

$$\hat{u} \in C\left([0,\infty);\mathcal{S}\right) \cap C^1\left((0,\infty);\mathcal{S}\right),$$

which implies that for each fixed $k \in \mathbb{R}^n$ the scalar-valued function $\hat{u}(k,t)$ is pointwise-differentiable with respect to t in $t > 0$ and continuous in $t \geq 0$. Solving the ODE (5.6) with k as a parameter, we find that $\hat{u}$ must be given by (5.7). Conversely, we claim that the function defined by (5.7) is strongly differentiable with derivative

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

To prove this claim, note that for $h > 0$ we have

$$\frac{\hat{u}(k,t+h) - \hat{u}(k,t)}{h} - u_t(k,t) = \hat{f}(k) e^{-t|k|^2} \left(\frac{e^{-h|k|^2} - 1 + h|k|^2}{h} \right).$$

and

$$\frac{e^{-h|k|^2} - 1 + h|k|^2}{h} \to 0 \qquad \text{in } \mathcal{S} \text{ as } h \to 0^+;$$

while for $h < 0$ we have

$$\frac{\hat{u}(k,t+h) - \hat{u}(k,t)}{h} - u_t(k,t) = \hat{f}(k) e^{-(t+h)|k|^2} \left(\frac{1 - h|k|^2 - e^{h|k|^2}}{h} \right),$$

and a similar conclusion follows. Thus, (5.2) has a unique solution that satisfies (5.3). Moreover, using induction, we see that $u \in C^\infty\left((0,\infty);\mathcal{S}\right)$.

From Example 5.24, 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 (5.7) and using the convolution theorem, we get (5.4)–(5.5). □

This solution of the heat equation satisfies two basic estimates, one in L^2 and the other in L^∞; the L^2 estimate follows from the Fourier representation, and the L^1 estimate follows from the spatial representation. We let $\|\cdot\|_{L^p}$ denote the spatial L^p-norm,

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

for $1 \le p < \infty$ and the essential supremum for $p = \infty$.

Corollary 5.4. *If $u : [0,\infty) \to \mathcal{S}(\mathbb{R}^n)$ is the solution of (5.2) constructed in Theorem 5.3, then for $t > 0$*

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

Proof. By Parseval's inequality and (5.7),

$$\|u(t)\|_{L^2} = (2\pi)^n \|\hat{u}(t)\|_{L^2} \le (2\pi)^n \|\hat{f}\|_{L^2} = \|f\|_{L^2},$$

which gives the first inequality. From (5.4),

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

and from (5.5)

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

The second inequality then follows. □

Using Theorem 5.31, it follows by interpolation between $(p,p') = (2,2)$ and $(p,p') = (\infty,1)$, that for $2 \le p \le \infty$

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

The requirement that $u(t) \in \mathcal{S}$ imposes a condition on the behavior of the solution at infinity. A solution of the initial value problem for the heat equation is not unique without the imposition of some kind of growth condition at infinity. A physical interpretation of this nonuniqueness it is that heat can diffuse from infinity into a 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.

We cannot solve the heat equation backward in time to obtain a solution $u : [-T,0] \to \mathcal{S}$ for general final data $f \in \mathcal{S}$, even if $T > 0$ is small. The same argument as the one in the proof of Theorem 5.3 implies that any such solution would be given by (5.7). 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, in fact, in $\mathcal{S}'$).

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

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

need not exist, so that this problem is not well-posed in $\mathcal{S}$. It is possible to obtain a well-posed initial value problem for the backward heat equation by restricting the initial data, 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, but these restrictions are typically too strong to be useful in applications.

5.1.2. Sobolev solutions. For any initial data $f \in \mathcal{S}$, the solutions constructed above satisfy an estimate of the form $\|u(t)\|_{L^2} \le \|f\|_{L^2}$ and we may therefore extend them by continuity and density to arbitrary initial data $f \in L^2$. More generally, similar estimates hold in any Sobolev space H^s (see Section 5.A.8), which allows us to define generalized solutions for $f \in H^s$.

PROPOSITION 5.5. *Suppose that $u : [0,\infty) \to \mathcal{S}$ is the solution of (5.2) constructed in Theorem 5.3. Then for any $s \in \mathbb{R}$*

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

PROOF. Using (5.7) and Parseval's identity, we find that

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

□

For $T > 0$ and $s \in \mathbb{R}$, let $C([0,T]; H^s)$ denote the Banach space of continuous functions $u : [0,T] \to H^s$ equipped with the norm

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

DEFINITION 5.6. Suppose that $T > 0$, $s \in \mathbb{R}$ and $f \in H^s$. A function $u : [0,T] \to H^s$ is a generalized solution of (5.2) if there exists a sequence of solutions $u_n : [0,T] \to \mathcal{S}$ such that $u_n \to u$ in $C([0,T]; H^s)$ as $n \to \infty$.

THEOREM 5.7. *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*

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

PROOF. Fix $T > 0$. 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 Proposition 5.5 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, we have

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

Taking the limit of this expression in $C([0,T]; H^s)$, we get the solution for $\hat{u}$. □

Since a unique generalized solution is defined on any time interval $[0,T]$, there is a unique generalized solution $u \in C_{\text{loc}}([0,\infty); H^s)$. 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 5.8. *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 $u_t = \Delta u$ in $C^1([0,T]; H^{s-2})$.*

PROOF. Suppose that $u_n \in C([0,T]; \mathcal{S})$ and $u_n \to u$ in $C([0,T]; H^s)$. Then $u_n \in C^1([0,T]; \mathcal{S})$ and $u_{nt} = \Delta u_n$, so $\{u_{nt}\}$ is Cauchy in $C([0,T]; H^{s-2})$ since $\{u_n\}$ is Cauchy in H^s and $\Delta : H^s \to H^{s-2}$ is bounded. Hence there exists $v \in C([0,T]; H^{s-2})$ such that $u_{nt} \to v$ in $C([0,T]; H^{s-2})$. It follows that

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

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

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 which 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, and solutions have much greater regularity as a result of the spatially smoothing effect of the evolution; in fact,

$$u \in C([0,\infty); H^s) \cap C^\infty((0,\infty); H^\infty).$$

5.1.3. The heat-equation semigroup. 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 solution operators $T(t) = e^{tA}$ form a uniformly continuous one-parameter group. We may consider the heat equation and other linear evolution equations from a similar perspective. There are, however, significant new issues that arise as a result of the fact that the Laplacian and other spatial differential operators are unbounded maps of a Banach space into itself.

Consider the heat equation

$$u_t = \Delta u, \qquad u(x,0) = f(x)$$

and suppose, for definiteness, that $f \in L^2(\mathbb{R}^n)$. We could equally well consider initial data that lies in other Banach or Hilbert spaces, such as L^1 or H^s. From Theorem 5.7, with $s = 0$, there is a unique generalized solution $u : [0,\infty) \to L^2$ of the heat equation. For each $t \geq 0$ we may therefore define a bounded linear map $T(t) : L^2 \to L^2$ by $T(t) : f \mapsto u(t)$. Thus, $T(t)$ is the flow or solution operator for the heat equation that maps the initial data at time 0 to the solution at time t. In particular, $T(0) = I$ is the identity.

Since the PDE does not depend explicitly on time, we have

$$T(s+t) = T(s)T(t) \qquad \text{for all } s, t \geq 0, \tag{5.8}$$

so the operators $\{T(t) : t \geq 0\}$ form a one-parameter semigroup. They do not form a group because $T(-t)$ is undefined for $t < 0$ and the operators $T(t)$ are not invertible. This irreversibilty does not arise in the case of ODEs.

The semigroup property in (5.8) is obvious from the explicit Fourier representation (5.7) since

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

It is less obvious from the spatial representation (5.4), but follows from the fact that

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

where the $*$ denotes the spatial convolution and $\Gamma^t(x) = \Gamma(x,t)$.

This semigroup is strongly continuous, meaning that for each $f \in L^2$, the map $t \mapsto T(t)f$ from $[0,\infty)$ into L^2 is continuous; equivalently $T(t+h) \to T(t)$ as $h \to 0$ (or $h \to 0^+$ if $t = 0$) with respect to the strong operator topology. It is *not* true, however, that $T(t+h) \to T(t)$ as $h \to 0$ uniformly with respect to the operator norm, as is the case for ODEs.

We also use the notation

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

and interpret $T(t)$ as the operator exponential of $t\Delta$. Equation (5.8) then becomes the usual exponential formula

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

It is remarkable that although the Laplacian is an unbounded linear operator

$$\Delta : H^2 \subset L^2 \to L^2$$

on L^2, the forward-in-time solution operators $T(t) = e^{t\Delta}$ that it generates are bounded.

In this discussion, we began with the heat equation and the Laplacian and derived the corresponding semigroup. We can instead begin with a semigroup and determine the operator that generates it. A key question is then to characterize the operators that generate a semigroup. We will briefly describe some basic results of semigroup theory without proof. For a detailed discussion see, for example, [**4**].

DEFINITION 5.9. Let X be a Banach space. A one-parameter, strongly continuous (or C_0) semigroup on X is a family $\{T(t) : t \geq 0\}$ of bounded linear operators $T(t) : X \to X$ such that

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

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

Explicitly, (3) means that

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

If this condition holds, then the semigroup property implies that $T(t+h)f \to T(t)f$ in X as $h \to 0$ for every $t > 0$, not only for $t = 0$.

The heat equation semigroup on $X = L^2(\mathbb{R}^n)$ is an example of a contraction semigroup. The term 'contraction' is not used here in a strict sense. The wave equation and Schrödinger equation also generate contraction semigroups (and, in fact, groups since their evolution is time-reversible). Thus, the norm of the solution of a contraction semigroup is not required to be strictly decreasing in time and it may, for example, remain constant.

DEFINITION 5.10. Suppose that $\{T(t) : t \geq 0\}$ 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

$$\lim_{h\to 0^+} \frac{T(h)f - f}{h}$$

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

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

$$Af = \lim_{h\to 0^+} \frac{T(h)f - f}{h}.$$

DEFINITION 5.11. 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 $Af = g$.

A bounded operator with dense domain $\mathcal{D}(A)$ is closed if and only if $\mathcal{D}(A) = X$ are closed. Differential operators defined in terms of weak derivatives give typical examples of unbounded closed operators.

THEOREM 5.12. *If A is the generator of a strongly continuous semigroup $\{T(t)\}$ on a Banach space X, then A is closed and its domain $\mathcal{D}(A)$ is dense in X.*

The semigroup $T(t)$ may be recovered from its generator in various ways, many of which generalize ways of defining the standard exponential function in a manner that is appropriate for an operator that is unbounded.

Finally, we state some conditions for an operator to generate a semigroup.

DEFINITION 5.13. 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 of A if $\lambda I - A : \mathcal{D}(A) \subset X \to X$ is one-to-one and onto and with bounded inverse

$$R(\lambda, A) = (\lambda I - A)^{-1} : X \to X. \tag{5.9}$$

called the resolvent of A.

The Hille-Yoshida theorem, provides a necessary and sufficient condition for an operator A to generate a strongly continuous semigroup

THEOREM 5.14. *A linear operator $A : \mathcal{D}(A) \subset X \to X$ is the generator of a strongly continuous semigroup $\{T(t); t \geq 0\}$ in 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*

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

where the resolvent $R(\lambda, A)$ is defined in (5.9).

In that case,

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

The Lummer-Phillips theorem provides a more easily checked condition (that A is 'm-dissipative') for A to generate a contraction semigroup on a Hilbert space.