SYSTEMS IN CONTINUOUS-TIME: THE THEORY OF SEMIGROUPS

LIAM HODGKINSON

There are a number of fundamental ideas in the realm of dynamical systems and applied measure theory that I find absolutely wonderful. An obvious choice is the class of ergodic theorems, with their incredible applications not only in physical systems, but in statistics and experimental design, as well as number theory. There is also the Monge and Kantorovich duality formulae, which led to the founding of descriptive geometry, with some awesome applications in operations research (OR is basically all just duality theory to be honest). Today's presentation will touch upon one of my favourite ideas in all of mathematics.

Often in mathematical modelling we consider memoryless, time-homogeneous systems: excluding all other factors, based on where we are right now, where will we go? Examples of this are ODEs, PDEs, time-homogeneous Markov processes. Please note: extensions to time-inhomogeneous systems are also available, but are rarely considered because almost all time-inhomogeneous systems can be embedded in time-homogeneous systems. To imbue these ideas into a mathematical definition, we introduce the idea of a semidynamical system.

Definition. A *semidynamical system* $\{S_t\}_{t\geq 0}$ on X is a family of measurable transformations $S_t : X \to X$ for $t \in \mathbb{R}_+$ satisfying

(1) $S_0(x) = x$ for all $x \in X$,
(2) $S_t \circ S_{t'}(x) = S_{t+t'}(x)$ for all $x \in X$ and $t, t' \in \mathbb{R}_+$ (future locations determined only based on current position; *semigroup/group property*), and
(3) the mapping $(t, x) \to S_t(x)$ is continuous from $X \times \mathbb{R}_+$ into X.

Example. Let $S_t(x) = y(t)$ where $y(t)$ satisfies the initial value problem $y' = F(y)$ and $y(0) = x$ where $F : \mathbb{R}^d \to \mathbb{R}^d$ is sufficiently smooth to ensure solutions exist and are unique (suppose $F \in \mathcal{C}^1$). Then $\{S_t\}_{t\geq 0}$ is a dynamical system. But delay-differential equations and other time-inhomogeneous systems can also be expressed in terms of a semidynamical system.

The construction of a semidynamical system is very useful: many ideas of ergodic theory for discrete-time systems extend to semidynamical systems almost directly since $S_{nt_0} = S_{t_0}^n$ for any $t_0 > 0$.
To develop some ergodic theory on these semidynamical systems, we need the notion of an invariant measure.

Definition. A measure μ is called *invariant* under a family $\{S_t\}_{t\geq 0}$ of measurable transformations $S_t : X \to X$ if $\mu\left(S_t^{-1}(A)\right) = \mu(A)$ for all measurable sets A.

Theorem. *Let μ be a finite invariant measure with respect to the semidynamical system $\{S_t\}_{t\geq 0}$ and let $f \in L^1(X)$. Then*

$$f^*(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(S_t(x))\, dt$$

exists for μ-almost every $x \in X$.

Proof. For any integer $n \in \mathbb{N}$, there is $\int_0^n f(S_t(x))\,dt = \sum_{k=0}^{n-1} \int_0^1 f(S_t \circ S_k(x))\,dt$ and so by defining $g_f(x) = \int_0^1 f(S_t(x))\,dt$, there is

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(S_t(x))\,dt = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g_f\left(S_1^k(x)\right),$$

and so the result follows by the pointwise ergodic theorem. □

Using this same idea, many asymptotic results in discrete time systems also hold for continuous time systems as well. Thus, the definitions of ergodicity, mixing, exactness, and asymptotic stability have obvious counterparts for the continuous-time system corresponding to an invariant measure. Most importantly, however, the idea of this group property for time-homogeneous systems extends to transfer operators as well. For this purpose, we define *semigroups* as follows.

Definition (Semigroups)**.** A continuous semigroup on a Banach space X is a family of linear operators $\{T_t\}_{t\geq 0}$, with $T_t : X \to X$ such that $T_0 x = x$ for every $x \in X$, $T_{t+s} = T_t \circ T_s$ for every $s, t \geq 0$, and $T_t x \to x$ strongly in X for every $x \in X$ as $t \to 0^+$. A contraction semigroup is a continuous semigroup $\{T_t\}_{t\geq 0}$ which also satisfies $\|T_t x\| \leq \|x\|$ for every $x \in X$. For a Polish space S, a C_0-semigroup is a contraction semigroup on the Banach space $C_0(S)$ (closure of $C_c^\infty(S)$ functions).

Application to Probability: If X_t is a continuous-time Markov process on the state space S with sufficiently local transitions, then the collection of operators $\{T_t\}_{t\geq 0}$ defined by $T_t f(x) = \mathbb{E}_x f(X_t)$ is a C_0-semigroup on S (follows from the *Chapman-Kolmogorov relations*). Alternatively, for any C_0-semigroup T_t on S, there exists a unique continuous-time Markov process X_t on the one-point compactification $\hat{S}$ such that $T_t f(x) = \mathbb{E}_x f(X_t)$ for every $f \in C_0\left(\hat{S}\right)$ and $x \in \hat{S}$ [Kallenberg, Proposition 19.14]. In this way, semigroup theory is *fundamental* in the study of continuous-time stochastic processes.

Application to Dynamical Systems: Suppose that a semidynamical system $\{S_t\}_{t\geq 0}$ is non-singular with respect to a measure μ, that is, $\mu \circ S_t^{-1} \ll \mu$. We can define the Frobenius-Perron and Koopman operators $P_t : L^1(X) \to L^1(X)$ and $U_t : L^\infty(X) \to L^\infty(X)$ by

$$\int_{S_t(A)} P_t f(x)\,\mu(dx) = \int_A f(x)\,\mu(dx) \text{ for } f \in L^1$$

$$U_t f(x) = f(S_t(x)) \text{ for } f \in L^\infty.$$

For each t, P_t is a linear Markov operator, and there is

$$\int_A P_{t+t'} f(x)\,\mu(dx) = \int_{S_{t'}^{-1} \circ S_t^{-1}(A)} f(x)\,\mu(dx) = \int_A P_t \circ P_{t'} f(x)\,\mu(dx),$$

and so $P_{t+t'} = P_t \circ P_{t'}$ in L^1 for any $t, t' \geq 0$. Additionally, since $S_0(x) = x$, $P_0 f = f$. Similarly, $U_{t+t'} = U_t \circ U_{t'}$ and $U_0 f = f$. Thus, both P_t and U_t are examples of contraction semigroups.

Application to PDEs: To come.

Infinitesimal Generators

The greatest advantage with working with semigroups is the construction of the so-called *infinitesimal generator*. This terminology is familiar to anyone who has studied probability to any reasonable extent. For the rest of this talk, I will look at these objects and the essential theorems surrounding them. I believe the following quote explains it best:

> "The problems associated with the study of continuous-time processes are more difficult than those encountered in discrete time systems. This is partially due to concerns over continuity of processes with respect to time. Also, equivalent formulations of discrete and continuous time properties may appear more complicated in the continuous case because of the use of integrals rather than summations, for example, in the Birkhoff ergodic theorem. However, there is one great advantage in the study of continuous time problems over discrete time dynamics, and this is the existence of a new tool - the infinitesimal generator." [Lasota & Mackey, Chapter 7].

Definition. Let $\{T_t\}_{t\geq 0}$ be a semigroup on a Banach space X. Let $\mathcal{D}(A)$ be the linear subspace of all $x \in X$ such that the limit

$$Ax := \lim_{t\to 0^+} \frac{T_t x - x}{t}$$

exists, where the limit is in the strong (norm) topology on X. The resulting operator $A : \mathcal{D}(A) \to X$ is called the *infinitesimal generator*. A dense subset of $\mathcal{D}(A)$ is called a *core* of the generator: for $\mathcal{D}(A) \subset C_0(S)$, the subspace $\mathcal{D}(A) \cap C_c^\infty(S)$ is a classically chosen core.

Remark. $\mathcal{D}(A)$ is a linear subspace of L. It will be shown later that the subspace $\mathcal{D}(A)$ is dense and so A uniquely determines a semigroup (consequence of the Hille-Yosida Theorem).

It is immediately clear from the linearity of T_t that A is a linear operator. The boundedness of A has a very important interpretation as well. If A is bounded, then the above limit is occuring at a rate which is more or less independent of t.

Lemma. *A is a bounded operator if and only if T_t is uniformly continuous, that is,* $\lim_{t\to 0^+} \|T_t - I\| = 0$*. In this case,* $\mathcal{D}(A) = X$*.*

The proof of this is surprisingly non-trivial, but can be found in [Dunford & Schwarz, VIII.1.9]. With this in mind, we can properly introduce the infinitesimal generator as a 'derivative' of the semigroup as a whole, linking our original time-homogeneous ODE idea with semigroups.

Theorem (Kolmogorov Equations)**.** *Let $\{T_t\}_{t\geq 0}$ be a continuous semigroup on X with corresponding infinitesimal generator $A : \mathcal{D}(A) \to X$. Suppose that A is a bounded operator. For any $x \in \mathcal{D}(A)$, $T_t x \in \mathcal{D}(A)$ for every $t \geq 0$, and*

$$\frac{d}{dt}(T_t x) = \overset{\text{(forward)}}{A T_t x} = \overset{\text{(backward)}}{T_t A x}.$$

Proof:

- From above, since A is bounded, T_t is uniformly continuous.
- Consider any $t_0 > 0$.
- Taking the limit from above ($t > t_0$), for any $x \in \mathcal{D}(A)$, by the continuity of T_{t_0},

$$\lim_{t\to t_0^+} \frac{T_t x - T_{t_0} x}{t - t_0} = \lim_{t\to t_0^+} T_{t_0}\left(\frac{T_{t-t_0} x - x}{t - t_0}\right) = T_{t_0} A x.$$

- For $t < t_0$, there is

$$\left\| \frac{T_t x - T_{t_0} x}{t - t_0} - T_{t_0} A x \right\| = \left\| T_t \left(\frac{T_{t_0 - t} x - x}{t_0 - t} - Ax \right) \right\| + \|T_t A x - T_{t_0} A x\|.$$

- The right term goes to zero as $t \to t_0^-$ by definition of continuous semigroup.

- By the uniform continuity of T_t, the family of $\{\|T_t\|\}$ is uniformly bounded for t sufficiently close to t_0 (to do this rigorously is actually really damn hard) and so the left term also converges to zero as $t \to t_0^-$. This proves the backward equations hold.
- The forward equations are found on observation that

$$\frac{T_t x - T_{t_0} x}{t - t_0} = \frac{T_{t-t_0}\left(T_{t_0} x\right) - T_{t_0} x}{t - t_0} \qquad \text{for } t > t_0.$$

In light of the Kolmogorov equations, it is common to use the notation $T_t = e^{At}$, as it can be shown that the solution is given by

$$T_t x = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k x.$$

Defining $u(t) = T_t x$, we find that u is the solution to $u' = Au$ and $u(0) = x$. Often, we might want to know conditions for when such solutions exist for some arbitrary operator A (which may even be an integro-differential operator or a partial differential operator). The link to semigroup theory is made rigorous in the following theorem. For this, we remind the readers about a bit of spectral theory:

Definition. The resolvent set of a linear operator A is the collection of all $\lambda \geq 0$ such that $\lambda I - A$ is an invertible bounded linear operator. The resolvent operator $R_\lambda(A)$ is defined for all λ in the resolvent set of A by $R_\lambda(A) = (\lambda I - A)^{-1}$.

We will be dropping the (A) notation, since it will be understood that the resolvent operator always applies to the infinitesimal generator.

Theorem (Mild Solutions to Abstract Cauchy Problems). *Suppose that A is a closed linear operator (if $x_n \to x$ and $Ax_n \to y$ then $Ax = y$) on X with non-empty resolvent set. Then for every $x \in X$, there exists a unique continuously differentiable function $u : \mathbb{R}_+ \to X$ with $u(t) \in \mathcal{D}(A)$ for all $t > 0$ satisfying*

$$\begin{aligned} u'(t) &= Au(t), \qquad t \geq 0 \\ u(0) &= x, \end{aligned}$$

if and only if A is the infinitesimal generator of a continuous semigroup T_t, in which case $u(t) = T_t x$.

Proof (Outline):

- Sufficiency, that is, the existence of a corresponding semigroup implies solutions, follows from the Kolmogorov equations.
- Choose a λ in the resolvent set. For any $x \in X$, define $u_x(t) = R_\lambda^{-1} v(t)$ where $v(t)$ satisfies $v'(t) = Av(t)$ and $v(0) = R_\lambda x$, so that u_x uniquely satisfies

$$\int_0^t u_x(s) - \lambda v(s)\, ds = (\lambda - A)^{-1}(x - u_x(t))$$

and since $\int_0^t Av(s)\, ds = A \int_0^t v(s)\, ds$,

$$u_x(t) = x + A \int_0^t u_x(s)\, ds.$$

- Since $\int_0^t u_x(s)\, ds$ is the solution to the problem for initial value 0, u_x is the unique solution to this above equation.

- From uniqueness, $x \mapsto u_x$ is linear and so for every $t \geq 0$, the mapping $T_t : X \to X$ defined by $T_t x = u_x(t)$ is linear.
- It can be shown that T_t is also a bounded linear operator (this is a little tedious; it makes use of the closed graph theorem for A) for every $t \geq 0$.
- By uniqueness, $t \mapsto u_x(t+s)$ is equivalent to $t \mapsto u_{u_x(s)}(t)$ and so $T_{t+s} = T_t \circ T_s$.
- From the Kolmogorov equations, the resulting infinitesimal generator of T_t must be equivalent to A.

Alternatively, we might wish to construct a Markov process based on how we expect the process to evolve through time (unlike Brownian motion, which is constructed based on its finite-dimensional distributions). Both problems can be answered using the Hille-Yosida Theorem which provides necessary and sufficient conditions for when a given operator A is the infinitesimal generator of a semigroup. This theorem is highly celebrated, and there are many different versions available for many different applications.

Theorem (Hille-Yosida Theorem)**.** *The following are necessary and sufficient conditions for a closed linear operator $A : \mathcal{D}(A) \to X$ where $\mathcal{D}(A) \subset X$ is dense in X, to be the infinitesimal generator of a corresponding semigroup:*

- **Hille-Yosida-Phillips (Continuous Semigroup):** There exist constants $M > 0$ and $\omega \geq 0$ such that $\|R_\lambda^n\| \leq M(\lambda - \omega)^{-n}$ for all $n \in \mathbb{N}$ and $\lambda > \omega$, in which case $\|T_t\| \leq Me^{\omega t}$ for all $t \geq 0$.
- **Lumer-Phillips (Contraction Semigroup):** A is dissipative, that is, $\|\lambda x - Ax\| \geq \lambda \|x\|$ for every $x \in \mathcal{D}(A)$ and $\lambda > 0$, and R_{λ_0} is defined on all of X for some λ_0.
- **Hille-Yosida-Ray (C_0-Semigroup):** For every function $f \in C_0$ with a global maximum x and $f(x) \geq 0$, $Af(x) \leq 0$ (the *positive-maximum principle*), and the range of $\lambda_0 - A$ is dense in C_0 for some $\lambda_0 > 0$.

Remark. If X is a reflexive space, the conditions that $\mathcal{D}(A) \subset X$ and A is closed are automatically implied by any of the other conditions.

The proof of this result is *always very* challenging. It relies on two very deep fundamental concepts in semigroup theory; these ideas are absolutely inescapable when proving any low-level result in the theory. Rigorous proofs of all of the previous results are usually done in the context of these ideas. The first is an explicit description of the resolvent operators.

Lemma. *The resolvent operators are the Laplace transform of the semigroup, that is, $R_\lambda x = \int_0^\infty e^{-\lambda t} T_t x dt$.*

Proof:

- Observe that $R_\mu^{-1} - R_\lambda^{-1} = (\mu - \lambda) I$ and so $R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu$.
- This same equation can be found to hold for the integral expression, and so the result follows by uniqueness of Laplace transforms.
- For a simpler proof in the case when A is bounded, we can use the fact that $e^{-\lambda t} T_t$ is a continuous semigroup with infinitesimal generator $-R_\lambda^{-1}$ from which the result follows by the Kolmogorov equations.

The second technique that is used is called the *Yosida semigroup*, which acts as an approximation to the semigroup, much like mollified functions do in PDEs. Observe that for any $x \in \mathbb{R}$,

$$(\lambda - x)^{-1} = \frac{1}{\lambda}\left[1 + \frac{x}{\lambda} + \frac{x^2}{\lambda^2} + o\left(\frac{1}{\lambda^2}\right)\right], \qquad \text{and so} \qquad \lambda\left(\frac{\lambda}{\lambda - x} - 1\right) = x + \frac{x^2}{\lambda} + o\left(\frac{1}{\lambda}\right),$$

so by taking the limit of this object as $\lambda \to \infty$, we get x back. Similarly, defining the *Yosida semigroup approximation to* A by

$$A^\lambda = -\lambda\left(I - \lambda R_\lambda\right),$$

this can be easily found to induce a corresponding approximative semigroup. Indeed, $A^\lambda \to A$ as $\lambda \to \infty$ uniformly.

Proof of Hille-Yosida Theorem:

- An infinitesimal generator satisfies the conditions of Hille-Yosida-Phillips and Lumer-Phillips from the formula for the resolvents.
- If A is an infinitesimal generator for a C_0-semigroup, then if $f \in \mathcal{D}(A)$ has $x \in S$ such that $f_+ \leq f(x)$, then
$$T_t f(x) \leq T_t f_+(x) \leq \|T_t f_+\|_\infty \leq \|f_+\|_\infty = f(x),$$
which immediately implies that $Af(x) \leq 0$, thus, the conditions for Hille-Yosida-Ray are satisfied.
- Lumer-Phillips implies Hille-Yosida-Phillips: for contraction semigroups, since $\|T_t\| \leq 1$ then $\|R_\lambda\| \leq \lambda^{-1}$. By repeated integration, we have that $\|R_\lambda^n\| \leq \lambda^{-n}$.
- Hille-Yosida-Ray implies Lumer-Phillips: let $f \in \mathcal{D}(A)$ be arbitrary and $x \in S$ satisfy $|f(x)| = \|f\|$. Let $g(x) = f \cdot \operatorname{sign} f(x)$ so that $g_+ \leq g(x)$ and $Ag(x) \leq 0$. Then for any $\lambda > 0$,
$$\|(\lambda - A) f\| \geq \lambda g(x) - Ag(x) \geq \lambda g(x) = \lambda \|f\|.$$
- To prove Hille-Yosida-Phillips:
  - Let $T_t^\lambda = e^{tA_\lambda}$: from the series expansion of e^{tA_λ} and the conditions, we can show that
  $$\left\|T_t^\lambda\right\| \leq M \cdot \exp\left(t\omega \cdot \frac{\lambda}{\lambda - \omega}\right),$$
  and so for any $\omega^* > \omega$, by taking λ sufficiently large, $\left\|T_t^\lambda\right\| < Me^{t\omega^*}$. Thus, $\left\|T_t^\lambda\right\| \leq Me^{t\omega}$.
  - There is for any $x \in X$,
  $$T_t^\lambda x - T_t^\mu x = \int_0^t \frac{d}{ds}\left[T_{t-s}^\mu T_s^\lambda x\right] ds = \int_0^t T_{t-s}^\mu T_s^\lambda \left(A_\lambda - A_\mu\right) x ds,$$
  and so we obtain the estimate
  $$\left\|T_t^\lambda x - T_t^\mu x\right\| \leq M^2 t e^{t\omega} \left\|A_\lambda x - A_\mu x\right\|.$$
  - Since $A_\lambda x \to Ax$, $\left\{T_t^\lambda\right\}_{\lambda \geq 0}$ is Cauchy with a limit T_t. Since T_t^λ is a semigroup, T_t is a semigroup and A is its infinitesimal generator.

The last result to be presented is a remarkable general convergence theorem for semigroups. While this is certainly useful in the numerical side of PDE theory (you know, the side of PDE theory that is actually valuable; trust Min, not me), it has even greater consequences in probability.

Consequence of Lax's Theorem. Suppose that $A_n x \to Ax$ for all $x \in X$. Then for the corresponding semigroups $T_{n,t}$ and T_t, there is $T_{n,t}x \to T_t x$ for all $x \in X$ as $n \to \infty$ if and only if the corresponding semigroups $T_{n,t}$ are uniformly bounded.

As it turns out, the sequence of semigroups $T_{n,t}$ must also be uniformly bounded whenever $A_n \to A$. This follows from two facts: the first is that the Yosida semigroup approximations $A_n^\lambda \to A^\lambda$ strongly, that is, $\left\|A_n^\lambda - A^\lambda\right\| \to 0$. The second, is that uniform convergence of A_n implies uniform convergence of $T_{n,t}$ through the following estimate:

Lemma. *Let T_t and T_t' be semigroups with corresponding generators A and A'. Assume A' is bounded. Then*

$$\left\|T_t x - T_t' x\right\| \leq \int_0^t \left\|T_{t-s}'\right\| \left\|\left(A - A'\right) T_s x\right\| ds, \qquad x \in \mathcal{D}(A), \, t \geq 0.$$

Proof. Observe that by the commutativity of A and T_t, there is

$$\begin{aligned} T_t x - T_t' x &= \int_0^t \frac{d}{ds}\left[T_{t-s}' T_s x\right] ds = \int_0^t T_{t-s}' T_s \left(A - A'\right) x ds \\ &= \int_0^t T_{t-s}' \left(A - A'\right) T_s x ds, \end{aligned}$$

from which the result follows. □

Theorem (Trotter-Kurtz Theorem). *For a sequence of semigroups $\{T_{n,t}\}_{n=1}^\infty$ and semigroup T_t with corresponding generators A_n and A, $T_{n,t} \to T_t$ uniformly for every $t \geq 0$ if and only if $A_n x \to Ax$ for every $x \in X$.*

Applications. The above theorem is instrumental in proving the physical relevance of linear ODEs and parabolic linear PDEs as it immediately implies convergence of Markov processes to their corresponding deterministic limits. Consider this: the infinitesimal generator of the scaled random walk $\frac{1}{\sqrt{n}} X_{nt}$ (where X_t is the random walk on $\mathbb{Z}$) is given by

$$\begin{aligned} A_n f(x) &= \frac{n}{2}\left[f\left(x + \frac{1}{\sqrt{n}}\right) - 2f(x) + f\left(x - \frac{1}{\sqrt{n}}\right)\right] \\ &\to \frac{1}{2} f''(x), \end{aligned}$$

which is the infinitesimal generator for Brownian motion. Thus, the Trotter-Kurtz Theorem implies that the distributions of the scaled random walk converge to those of Brownian motion.

Resources

- "Vector-Valued Laplace Transforms and Cauchy Problems" by Arendt et al.
- "Linear Operators, Part I General Theory" by Dunford & Schwarz
- "Fundamentals of Modern Probability" by Kallenberg
- "Functional Analysis" by Yosida
- "Chaos, Fractals, and Noise" by Lasota and Mackey