An Intuitive Introduction to Operator Semi-Groups

Martin Keller-Ressel

January 17, 2006

This article aims to give an intuitive introduction to operator semi-groups and their generators from a probabilistic perspective. By 'intuitive' it is meant that the article relies mainly on heuristics and analogies to make its points. For example the definition of the generator is motivated from an analogy to the Cauchy functional equation $g(t+s) = g(t)g(s)$ for real-valued functions. No proofs are given throughout the text. They can be found in either Kallenberg [1997], Yosida [1980] or Butzer and Berens [1967]. Other recommended literature is the article Jacob and Schilling [2001], on which section 5 is based.

1 Operator semi-groups and their generator

The theory of operator semi-groups originates from the study of the equation

$$T(t+s) = T(t)T(s) \qquad T(0) = I \tag{1.1}$$

where $T(t)$ is an operator-valued function taking values in the set of bounded linear operators acting on a suitable function space. This problem was independently studied by Hille and Yosida around 1948. They introduced the notion of a operator semi-group, which is defined as follows:

> Let $\mathcal{X}$ be a Banach-space. Then a family $\{T_t\}_{t\geq 0}$ of bounded linear operators $\in L(\mathcal{X}, \mathcal{X})$ is called an operator semi-group if it satisfies Operator Semi-group
>
> $$T_{t+s} = T_t T_s \qquad T_0 = I$$

A semi-group is called equi-bounded if $\|T_t\| \leq M$ for some constant M and all $t \geq 0$ and is called a contraction semi-group if $\|T_t\| \leq 1$ for all $t \geq 0$.

From a stochastic point of view operator semi-groups originate from the study of Markov processes. Recall the following definition: Markov process

> Given a stochastic base $(\Omega, \mathcal{F}, \mathbb{F}, P)$ a process X is called a Markov process with respect to $\mathbb{F}$ if $\mathcal{F}_t$ and $\sigma\{X_s; s \geq t\}$ are conditionally independent given X_t for all $t \geq 0$.

Informally one can say that 'given the present state of a Markov process its future is independent of the past'.
For a Markov process taking values in $\mathbb{R}^d$ we can define its transition operator $T_{s,t}$ on a suitable space of measurable functions from $\mathbb{R}^d$ to $\mathbb{R}$ by

$$T_{s,t}f(x) = \mathbb{E}[f(X_t)|X_s = x]$$

If the process X is time-homogenous – which will be assumed from now on – we can write

$$T_{t-s} = T_{s,t}$$

and the transition operators $\{T_t\}_{t\geq 0}$ form a semi-group due to the Markov property. The corresponding semi-group satisfies $T_t 1 = 1$; this property is called conservative. If $\mathcal{X}$ is equipped with the supremum-norm this also implies that $\{T_t\}_{t\geq 0}$ is a contraction semi-group.

Going back to equation (1.1) it is apparent that it bears a resemblance to the (exponential) Cauchy functional equation

Cauchy Functional Equation

$$g(s+t) = g(s)g(t) \qquad g(0) = 1 \tag{1.2}$$

where $g(.)$ is a non-negative function from $\mathbb{R}$ to $\mathbb{R}$. The solution to the exponential Cauchy functional equation is well known: It is the family of exponential functions $g(t) = e^{\alpha t}\,; \alpha \in \mathbb{R}$. However this family represents all possible solutions only if an additional assumption of continuity is made. In fact the assumption that $g(t)$ is continuous from the right *in the origin* is already sufficient to make the functions $g(t) = e^{\alpha t}\,; \alpha \in \mathbb{R}$ the only solutions of (1.2) (cf. Aczél [1966]).
We now introduce a similar assumption for the operator semi-group defined by (1.1):

Continuity at the origin

> An operator semi-group is **continuous at the origin** or **of class** C_0 if
>
> $$\lim_{t\to 0} T_t f = f \quad \forall\, f \in \mathcal{X}$$

One would now hope that in analogy to the Cauchy equation an operator semi-group of class C_0 can be represented in the form $T_t = e^{tA}$ for a suitable operator A. Let us for a moment assume that this is possible[1] and we have a representation of the form $T_t = e^{tA}$. How can the operator A be derived from the semigroup T_t? Again looking at the Cauchy functional equation proofs worthwhile. We know the solution is of type $g(t) = e^{\alpha t}$ and to retrieve the parameter α without resorting to the logarithm we can use differentiation:

$$\alpha = \lim_{t\downarrow 0} \frac{g(t)-1}{t}$$

The analogous approach for the operator semi-group yields

Infinitesimal Generator

[1] We will see in section 3 in how far this assumption is justified.

$$(Af)(x) = \lim_{t \downarrow 0} \frac{((T_t - I)f)(x)}{t} \tag{1.3}$$

Without checking the validity of the assumption $T_t = e^{tA}$ and the formal derivation of (1.3) we proclaim (1.3) as the definition of the **infinitesimal generator** of a semi-group. The domain $D(A)$ of the operator A is the set of f where the limit in (1.3) exists for all x. It can be shown that the domain $D(A)$ is always dense in $\mathcal{X}$.

To do a first 'reality check' of our construction so far we calculate the generator of a compound Poisson process:

Generator of a compound Poisson process

> Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of i.i.d random variables with distribution function $F(x)$ and N_t a Poisson process with intensity λ. Denote by $S_n := X_1 + \cdots + X_n$ the partial sums over the X_i. Now the compound Poisson process Y is defined by
>
> $$Y_t := \sum_{n \geq 1} S_n \, \mathbf{1}_{\{N_t = n\}}$$
>
> The transition operator T_t of Y is given by $\mathbb{E}_x[f(Y_t)]$ where we assume that $f \in C_b(\mathbb{R})$. To simplify further calculations define an operator L by
>
> $$Lf(x) := \mathbb{E}[f(x + X_1)] = \int_{\mathbb{R}} f(x+y) \, F(dy)$$
>
> and note that
>
> $$L^n f(x) = \mathbb{E}[f(x + X_1 + \cdots + X_n)] = \mathbb{E}[f(x + S_n)]$$
>
> Now it holds that
>
> $$\begin{aligned} (T_t f)(x) = \mathbb{E}^x[f(Y_t)] &= \sum_{n \geq 0} \mathbb{E}[f(x+S_n)] \cdot P(N_t = n) = \\ = \sum_{n \geq 0} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \mathbb{E}[f(x + S_n)] &= \sum_{n \geq 0} e^{-\lambda t} \frac{(\lambda t)^n}{n!} L^n f(x) = \\ &= \left(e^{\lambda t (L - I)} f \right)(x) \end{aligned} \tag{1.4}$$
>
> and indeed the transition semi-group can be written as e^{tA} where A is given by
>
> $$Af(x) = \lambda (L - I) f(x) = \lambda \int_{\mathbb{R}} (f(x+y) - f(x)) \; F(dy)$$

2 Kolmogorov's forward and backward equation and the resolvent

While in (1.3) A was defined essentially as the derivative of T_t at 0 also the derivative of T_t at any point can be calculated by

$$(D_t\ T_t f)(x) = \lim_{h \downarrow 0} \frac{((T_{t+h} - T_t)f)(x)}{h}$$

This derivative is well defined on A's domain $D(A)$ and it can be shown that

Kolmogorov's Backward Equation

$$D_t T_t f = A T_t f = T_t A f \qquad \forall\ f \in D(A), t \geq 0 \tag{2.1}$$

which again parallels the case of $g(t) = e^{t\alpha}$, the solution of the scalar Cauchy equation $g(t+s) = g(t)g(s)$. In particular (2.1) shows that A and T_t commute. The equation $D_t T_t f = A T_t f$ is called Kolmogorov's backward equation and $D_t T_t f = T_t A f$ is called Kolmogorov's forward equation. Both will be familiar to the probabilist (for example from Øksendal [2003]) in the following form:

> Let f be sufficiently regular function (e.g. twice differentiable with compact support) and let X_t be an Ito diffusion with infinitesimal generator A. Define
>
> $$u(t,x) := \mathbb{E}_x[f(X_t)]$$
>
> then $u(t,x)$ is differentiable with respect to t and
>
> $$\frac{\partial u}{\partial t} = \mathbb{E}_x[Af(X_t)] = Au$$

which is nothing else than (2.1).

In the definition of the infinitesimal generator in (1.3) we have pointed out the analogy to the exponential function $g(t) = e^{\alpha t}$ where the parameter α can be retrieved by calculating the derivative at 0. We now present another method to calculate α from $g(t)$, which has a useful analogue for linear operators. The Laplace transform of $g(t)$ is given by

$$\mathcal{L}(g)(\lambda) = \int_0^\infty e^{-\lambda t} g(t)\ dt = \frac{1}{\lambda - \alpha} \qquad \lambda > 0 \tag{2.2}$$

For the expression (2.2) to make sense when α is replaced by A, the operator $(\lambda I - A)$ has to be invertible for $\lambda > 0$. Recall that a complex number $\lambda \in \mathbb{C}$ is said to be in the spectrum of a bounded linear operator L if $(\lambda I - L)$ has no (bounded linear) inverse and is said to be in the resolvent set if $(\lambda I - L)$ does have a bounded linear inverse. The corresponding inverse operator $R_\lambda := (\lambda I - L)^{-1}$ is then called resolvent. For a semi-group generator A it can be shown that the resolvent set always contains the positive real halfline and the resolvent R_λ is indeed given by

Resolvent

$$R_\lambda f = \int_0^\infty e^{-\lambda t}(T_t f)\,dt = (\lambda I - A)^{-1} f \qquad \lambda > 0 \tag{2.3}$$

Given an arbitrary operator its resolvent can be used to answer the question whether this operator is the generator of a semigroup. The following theorem is a version of the Hille-Yosida theorem for contraction semi-groups:

> A closed linear operator $A \in L(\mathcal{X}, \mathcal{X})$ with dense domain is the generator of a contraction semi-group of class C_0 if and only if its resolvent set contains the positive real halfline and $\|\lambda R_\lambda\| \leq 1$ for all $\lambda > 0$.

3 Approximation by Pseudo-Poisson processes

At the end of section 1 we calculated the transition operator semi-group of a compound Poisson process and saw that it could be represented in the form $T_t = e^{tA}$. In fact a good part of the heuristic argumentation in section 1 has revolved around the assumption that it makes sense to think of any semi-group in terms of such an exponential representation. We explore this assumption now more thoroughly. First we define the Feller process, a type of process that is essentially a Markov process satisfying some additional mild regularity assumptions. Then we consider the class of pseudo-Poisson processes, which are a generalization of compound Poisson processes and have the property that their transition semi-group can still be represented in exponential form. Finally we show that any Feller process can be approximated by a pseudo-Poisson process.

Feller process

Let X be a Markov process, and $\{T_t\}_{t\geq 0}$ its transition semi-group acting on $\mathcal{X} = C_0(\mathbb{R}^d)$, the space of all functions from $\mathbb{R}^d$ to $\mathbb{R}$ vanishing at infinity[2]. Equipped with the supremum-norm $C_0(\mathbb{R}^d)$ is a Banach space and the semi-group is a contraction semi-group. If the semi-group fulfills the regularity conditions

$$T_t C_0(\mathbb{R}^d) \subseteq C_0(\mathbb{R}^d) \tag{3.1}$$

$$\lim_{t\to 0} T_t f(x) = f(x) \qquad \forall f \in C_0(\mathbb{R}^d), x \in \mathbb{R}^d \tag{3.2}$$

it is called **Feller semi-group** and X a **Feller process**. The regularity conditions also imply continuity at the origin.

Pseudo-Poisson process

Pseudo-Poisson processes can be derived from compound Poisson processes in the following way: A compound Poisson process Y_t can be defined, as in section 1, by

$$Y_t = \sum_{n\geq 1} S_n \mathbf{1}_{\{N_t = n\}} \tag{3.3}$$

[2]The reason why the space $C_0(\mathbb{R}^d)$ is convenient to work with here is that its functions are continuous on a one-point compactification of $\mathbb{R}^d$.

where N_t is a Poisson process with intensity λ and S_n the partial sums of i.i.d random variables. A **Pseudo-Poisson process** is given by the same defining equation (3.3) but now S_n are the successive states of a (time-homogenous) Markov chain. We denote the transition operator of the Markov chain by L:

$$(Lf)(x) = \mathbb{E}[f(S_{n+1})|S_n = x]$$

and note that

$$(L^k f)(x) = \mathbb{E}[f(S_{n+k})|S_n = x]$$

We also define a probability kernel F by

$$F(A, x) := (L\mathbf{1}_{\{A\}})(x) = P\left(f(S_{n+1}) \in A|S_n = x\right) \qquad \forall\, A \in \mathcal{B}(\mathbb{R})$$

Now the same calculation as in (1.4) yields a representation of the transition semigroup of Y:

$$(T_t f)(x) = \left(e^{\lambda t(L-I)} f\right)(x)$$

The generator A is given by

Generator of a pseudo-Poisson process

$$(Af)(x) = \lambda(L - I)f(x) = \int_{\mathbb{R}} \left(f(x+y) - f(x)\right)\, F(dy, x)$$

Now let A be the generator of a Feller semi-group. As we have seen in section 2 the resolvent R_λ of A is well-defined for all $\lambda > 0$. Define an operator A^λ by

$$A^\lambda = \lambda A R_\lambda = \lambda(\lambda R_\lambda - I)) \qquad \lambda > 0$$

where the second equality is due to the fact that $R_\lambda = (\lambda I - A)^{-1}$. Clearly A^λ is the generator of a Pseudo-Poisson process with $L = \lambda R_\lambda$ being the transition operator of the implied Markov chain. The following theorem is due to Yosida and justifies the statement that Pseudo-Poisson processes are approximators of Feller processes:

> Let A be the generator of a Feller semi-group. Then for any $f \in D(A)$ it holds that
>
> $$\lim_{\lambda\to\infty} A^\lambda f = Af \tag{3.4}$$
>
> $$\lim_{\lambda\to\infty} e^{tA^\lambda} f = T_t f \qquad \forall\, t \geq 0 \tag{3.5}$$
>
> $$\left\| T_t f - e^{tA^\lambda} f \right\| \leq t \left\| Af - A^\lambda f \right\| \qquad \forall\, t, \lambda \geq 0 \tag{3.6}$$

We can now answer the question if Feller semi-groups can be represented in the form $T_t = e^{tA}$. The answer is in general no, because the generator A may be an unbounded operator for which the exponential is not defined. But we can approximate the semi-group by the semi-group of a Pseudo-Poisson process which has an exponential representation. This shows that while the motivation used in section 1 is not correct in a formal sense, it makes sense at least under the regularity assumptions of a Feller process.

4 Maximum principles and elliptic operators

It is known from elementary calculus that a function $f \in C^2([a,b])$ taking a local maximum at x_0 satisfies $f'(x_0) = 0$ and $f''(x_0) \leq 0$. When this statement is inverted we get the following theorem:

Maximum principle

> Let $f \in C^2([a,b])$ satisfy $f'' \geq 0$ in (a,b). If f attains a maximum in (a,b) then f is constant.

Such a statement is called a **maximum principle**. Note that it still holds true when '$f'' \geq 0$' is replaced by '$c(x)f'' + g(x)f' \geq 0$' where g is a bounded function and c is positive and bounded away from 0. Extending this idea to $\mathbb{R}^d$ yields the maximum principle for **elliptic operators**. Recall that a second order differential operator

Elliptic operator

$$L(u) = \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \frac{\partial u}{\partial x_i} - c(x)u \tag{4.1}$$

is called elliptic at a point x if the matrix $A(x) = (a_{i,j}(x))$ is positive semidefinite or equivalently

$$z^T A(x) z \geq \mu(x)|z|^2 \qquad \forall\, z \in \mathbb{R}^d \tag{4.2}$$

and is called uniformly elliptic if in (4.2) μ is independent of x.
Many variations of a 'maximum principle for elliptic operators' can be found in the literature. The following version is due to Hopf (1927):

> Let L be an elliptic operator (as in (4.1)) with $c \geq 0$ and $Lu \geq 0$ in Ω. If u takes a nonnegative maximum in Ω then u is constant.

The connection to transition semi-groups is made by fact that generators of Feller processes satisfy the same maximum principle. For a Feller semi-group the **positive maximum principle** holds:

Positive maximum principle

> Let A be the generator of a Feller semi-group with $Af \geq 0$ in $S \subseteq \mathbb{R}^d$. If f takes a non-negative maximum at in S then f is constant.

The fact that generators of Feller semi-groups and elliptic operators satisfy the same maximum principle suggests a deeper connection between these objects. Denote by C_c^∞ the class of functions on $\mathbb{R}^d$ that are differentiable infinitely often and have compact support. It holds that

> If X is a continuous Feller-process on $[0,T]$ with generator A and $C_c^\infty \subseteq D(A)$ then A is elliptic.

5 Pseudo-Differential Operators and Levy processes

In the preceding section we have seen that if X is a continuous Feller process its generator is a second order elliptic differential operator. Consider a simple operator of this kind:

$$L : C_c^\infty(\mathbb{R}) \to C_c^\infty(\mathbb{R}), \quad f \mapsto af'' + bf' - c \qquad a > 0$$

Under the Fourier transform the action of the operator is mapped to multiplication with a polynomial of degree 2:

$$\widehat{Lf}(u) = (-au^2 + ibu - c)\widehat{f}(u) = p(u)\widehat{f}(u)$$

Using the inverse Fourier transform L can be written as

$$Lf(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iux} p(u)\widehat{f}(u)\, du \tag{5.1}$$

Similarly for a general k-th order differential operator, $p(u)$ is a polynomial in u of degree k, where the coefficients of the polynomial might well depend on x. The point of equation (5.1) is that even for non-differential operators a similar representation can be obtained. As an example take a $\mathbb{R}$-valued Levy process X and its transition operator T_t. Due to independence and stationarity of increments T_t can be written as a convolution integral: Levy process

$$T_t f(x) = \mathbb{E}[f(X_t)|X_0 = x] = \int_{\mathbb{R}} f(x+y)\mu_t(dy)$$

where μ_t is a probability measure on the real line. We can calculate the Fourier transformation of μ_t as

$$\widehat{\mu_t}(u) = \int_{\mathbb{R}} e^{ixu}\, \mu_t(dx) = \mathbb{E}[e^{iuX_t}|X_0 = 0] = e^{t\psi(u)}$$

where ψ is the characteristic exponent of the Levy process X, given by the Levy-Khintchine formula. Using the convolution theorem we get

$$\widehat{T_t f}(u) = \widehat{f \star \mu_t}(u) = \widehat{f}(u) \cdot \widehat{\mu_t}(u) = \widehat{f}(u)e^{t\psi(u)}$$

and Fourier inversion yields

$$T_t f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iux} e^{t\psi(u)} \widehat{f}(u)\, du \tag{5.2}$$

This means we have just represented T_t in the form of (5.1) with the polynomial $p(u)$ replaced by $e^{t\psi(u)}$. But it gets even better: The generator A of X is given by

$$\begin{aligned} Af(x) = \lim_{t\to 0} \frac{T_t f(x) - f(x)}{t} &= \frac{1}{\sqrt{2\pi}} \lim_{t\to 0} \int_{\mathbb{R}} e^{-iux} \frac{e^{t\psi(u)} - 1}{t} \widehat{f}(u)\, du = \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iux} \psi(u)\widehat{f}(u)\, du \quad (5.3) \end{aligned}$$

and again we have a representation of type (5.1). In general operators with such a representation are called **pseudo-differential operators**. We give a formal definition:

Pseudo-differential operator

Let $\mathcal{S}(\mathbb{R}^d)$ be the space of rapidly decreasing functions and $q : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ a measurable function such that $u \mapsto q(x,u)$ is continuous and polynomially bounded. Then the operator L given by

$$Lf(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-ix.u} q(x,u) \widehat{f}(u)\, du) \qquad f \in \mathcal{S}(\mathbb{R}^d)$$

is called a pseudo-differential operator with **symbol** $q(x,u)$.

The conclusions of (5.2) and (5.3) can be summarized in this new terminology as follows:

Let X be a Levy process with characteristic exponent $\psi(u)$. Then its transition operator is a pseudo-differential operator with symbol $\lambda_t(x,u) = e^{t\psi(u)}$ and its generator is a pseudo-differential operator with symbol $\psi(x,u) = \psi(u)$.

Note that in this case the symbols do not depend on x. Writing (5.2) and (5.3) in a probabilistic way gives:

$$\lambda_t(x,u) = \mathbb{E}^x[e^{i(X_t - x)u}] \tag{5.4}$$

$$\psi(x,u) = \frac{d}{dt}\lambda_t(u)|_{t=0} = \lim_{t\to 0} \frac{\mathbb{E}^x[e^{i(X_t-x)u}] - 1}{t} \tag{5.5}$$

Both expressions can be calculated not only for Levy processes, but for arbitrary Markov processes. And indeed it can be shown that for a Markov process X

- The transition operator T_t acting on $\mathcal{S}(\mathbb{R}^d)$ is a pseudo-differential operator with symbol $\lambda_t(x,u)$ given by (5.4)
- Under mild additional conditions also the generator A is a pseudo-differential operator with symbol $\psi(x,u)$ given by (5.5)

References

J. Aczél. *Functional Equations and their Applications.* Academic Press, 1966.

Paul Butzer and Hubert Berens. *Semi-Groups of Operators and Approximation.* Number 145 in Grundlehren. Springer, 1967.

Niels Jacob and René L. Schilling. Lévy-type processes and pseudodifferential operators. In *Barndorff-Nielsen, Ole E. (ed.) et al., Lévy processes. Theory and applications. Boston: Birkhäuser. 139-168* . 2001.

Olav Kallenberg. *Foundations of Modern Probability.* Springer, 1997.

Berndt Øksendal. *Stochastic Differential Equations.* Springer, 2003.

Kosaku Yosida. *Functional Analysis.* Springer, 6th edition, 1980.