This chapter is devoted to a review of standard topics from the theory of semigroups which forms a functional analytic background for the proof of Theorems 1.2, 1.3, 1.4 and 1.5.

2.1 Analytic Semigroups

This section provides a brief description of the basic results of the theory of analytic semigroups which forms a functional analytic background for the proof of Theorems 1.2 and 1.3. Moreover, Subsection 2.1.3 is devoted to the semigroup approach to a class of initial-boundary value problems for semilinear parabolic differential equations (Theorem 2.8). Theorem 1.5 follows by verifying all the conditions of Theorem 2.8. For more leisurely treatments of analytic semigroups, the reader is referred to Friedman [Fr1], Pazy [Pa], Tanabe [Tn], Yosida [Yo] and also Taira [Ta4].

2.1.1 Generation of Analytic Semigroups

Let \( E \) be a Banach space over the real or complex number field, and let \( A : E \to E \) be a densely defined, closed linear operator with domain \( \mathcal{D}(A) \).

Assume that the operator \( A \) satisfies the following two conditions (see Figure 2.1 below):

1. The resolvent set of \( A \) contains the region
   \[
   \Sigma_\omega = \{ \lambda \in \mathbb{C} : \lambda \neq 0, \ |\arg \lambda| < \pi/2 + \omega \}, \quad 0 < \omega < \pi/2.
   \]

2. For each \( \varepsilon > 0 \), there exists a positive constant \( M_\varepsilon \) such that the resolvent
   \[
   R(\lambda) = (A - \lambda I)^{-1}
   \]
   satisfies the estimate
   \[
   \|R(\lambda)\| \leq \frac{M_\varepsilon}{|\lambda|}
   \]
   for all \( \lambda \in \Sigma^\varepsilon := \{ \lambda \in \mathbb{C} : \lambda \neq 0, \ |\arg \lambda| \leq \pi/2 + \omega - \varepsilon \} \).

\[\text{(2.1)}\]
Then we let
\[ U(t) = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda t} R(\lambda) d\lambda. \] (2.2)

Here \( \Gamma \) is a path in the set \( \Sigma^\epsilon_\omega \) consisting of the following three curves (see Figure 2.2):

\begin{align*}
\Gamma^{(1)} &= \{ re^{-i(\pi/2 + \omega - \epsilon)} : 1 \leq r < \infty \}, \\
\Gamma^{(2)} &= \{ e^{i\theta} : -(\pi/2 + \omega - \epsilon) \leq \theta \leq \pi/2 + \omega - \epsilon \}, \\
\Gamma^{(3)} &= \{ re^{i(\pi/2 + \omega - \epsilon)} : 1 \leq r < \infty \}.
\end{align*}

It is easy to see that the integral
\[ U(t) = -\frac{1}{2\pi i} \sum_{k=1}^{3} \int_{\Gamma^{(k)}} e^{\lambda t} R(\lambda) d\lambda \]
converges in the uniform operator topology of the Banach space $L(E,E)$ for all $t > 0$, and thus defines a bounded linear operator on $E$. Here $L(E,E)$ denotes the space of bounded linear operators on $E$.

Furthermore, we have the following:

**Proposition 2.1.** The operators $U(t)$, defined by formula (2.2), form a semigroup on $E$, that is, they enjoy the semigroup property

$$U(t + s) = U(t) \cdot U(s) \quad \text{for all } t, s > 0.$$  

*Proof.* By Cauchy’s theorem, we may assume that

$$U(s) = -\frac{1}{2\pi i} \int_{\Gamma'} e^{\mu s} R(\mu) d\mu, \quad s > 0.$$  

Here $\Gamma'$ is a path obtained from the path $\Gamma$ by translating each point of $\Gamma$ to the right by a fixed small positive distance (see Figure 2.3).

![Fig. 2.3.](image)

Then we have, by Fubini’s theorem,  

$$U(t) \cdot U(s) = \frac{1}{(2\pi i)^2} \int_{\Gamma} \int_{\Gamma'} e^{\lambda t} e^{\mu s} R(\lambda) R(\mu) d\lambda d\mu$$  

$$= \frac{1}{(2\pi i)^2} \int_{\Gamma} \int_{\Gamma'} e^{\lambda t} e^{\mu s} \frac{R(\lambda) - R(\mu)}{\lambda - \mu} d\lambda d\mu$$  

$$= \frac{1}{2\pi i} \int_{\Gamma} e^{\lambda t} R(\lambda) \left[ \frac{1}{2\pi i} \int_{\Gamma'} e^{\mu s} \frac{d\mu}{\lambda - \mu} \right] d\lambda$$  

$$- \frac{1}{2\pi i} \int_{\Gamma'} e^{\mu s} R(\mu) \left[ \frac{1}{2\pi i} \int_{\Gamma} e^{\lambda t} \frac{d\lambda}{\lambda - \mu} \right] d\mu.$$  

We calculate the two terms in the last part.

(a) We let

$$f(\mu) = \frac{e^{\mu s}}{\lambda - \mu}, \quad \mu \in \mathbb{C}.$$
Then, by applying the residue theorem we obtain that (see Figure 2.4)

$$\int_{\Gamma_1} f(\mu) \, d\mu + \int_{\Gamma_2} f(\mu) \, d\mu + \int_{\Gamma_3} f(\mu) \, d\mu + \int_{\Gamma_4} f(re^{i\theta}) rie^{i\theta} \, d\theta$$

$$= 2\pi i \operatorname{Res}[f(\mu)]_{\mu=\lambda}$$

$$= -2\pi i e^{\lambda s}.$$

However, we have, as \( r \to \infty \),

$$\int_{\Gamma_1} f(\mu) \, d\mu \to \int_{\Gamma_1} f(\mu) \, d\mu,$$

$$\int_{\Gamma_3} f(\mu) \, d\mu \to \int_{\Gamma_3} f(\mu) \, d\mu,$$

and

$$\left| \int_{-\pi/2+\omega-\epsilon}^{\pi/2+\omega-\epsilon} f(re^{i\theta}) rie^{i\theta} \, d\theta \right| \leq e^{-rs \sin(\omega-\epsilon)} \int_{-\pi/2+\omega-\epsilon}^{\pi/2+\omega-\epsilon} \frac{d\theta}{\frac{\lambda}{r} - e^{i\theta}} \to 0.$$

Therefore, we find that

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{e^{\lambda s}}{\lambda - \mu} \, d\mu = -e^{\lambda s}.$$

(b) Similarly, since the path \( \Gamma \) lies to the left of the path \( \Gamma' \), we find that

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{e^{\lambda M}}{\lambda - \mu} \, d\lambda = 0.$$
Summing up, we obtain that
\[ U(t) \cdot U(s) = -\frac{1}{2\pi i} \int_{\gamma} e^{\lambda(t+s)} R(\lambda) d\lambda = U(t+s) \quad \text{for all } t, s > 0. \]

The proof of Proposition 2.1 is complete. \( \square \)

The next theorem states that the semigroup \( U(t) \) can be extended to an analytic semigroup in some sector containing the positive real axis.

**Theorem 2.2.** The semigroup \( U(t) \), defined by formula (2.2), can be extended to a semigroup \( U(z) \) which is analytic in the sector
\[ \Delta_\omega = \{ z = t + is : z \neq 0, |\arg z| < \omega \}, \]
and enjoys the following properties:

(a) The operators \( AU(z) \) and \( \frac{dU}{dz}(z) \) are bounded operators on \( E \) for each \( z \in \Delta_\omega \), and satisfy the relation
\[ \frac{dU}{dz}(z) = AU(z) \quad \text{for all } z \in \Delta_\omega. \] (2.3)

(b) For each \( 0 < \varepsilon < \omega / 2 \), there exist positive constants \( \bar{M}_0(\varepsilon) \) and \( \bar{M}_1(\varepsilon) \) such that
\[ \|U(z)\| \leq \bar{M}_0(\varepsilon) \quad \text{for all } z \in \Delta^{2\varepsilon}_\omega, \] (2.4)
\[ \|AU(z)\| \leq \frac{\bar{M}_1(\varepsilon)}{|z|} \quad \text{for all } z \in \Delta^{2\varepsilon}_\omega, \] (2.5)
where (see Figure 2.5)
\[ \Delta^{2\varepsilon}_\omega = \{ z \in \mathbb{C} : z \neq 0, |\arg z| \leq \omega - 2\varepsilon \}. \]

(c) For each \( x \in E \), we have, as \( z \to 0, z \in \Delta^{2\varepsilon}_\omega \),
\[ U(z)x \longrightarrow x \quad \text{in } E. \]

**Proof.** (i) The analyticity of \( U(z) \): If \( \lambda \in T^{(3)} \) and \( z \in \Delta^{2\varepsilon}_\omega \), that is, if we have the formulas
\[ \lambda = |\lambda|e^{i\theta}, \quad \theta = \pi/2 + \omega - \varepsilon, \]
\[ z = |z|e^{i\varphi}, \quad |\varphi| \leq \omega - 2\varepsilon, \]
then it follows that
\[ \lambda z = |\lambda||z|e^{i(\theta+\varphi)}, \]
with
\[ \pi/2 + \varepsilon \leq \theta + \varphi \leq \pi/2 + 2\omega - 3\varepsilon < 3\pi/2 - 3\varepsilon. \]
Note that
\[ \cos(\theta + \varphi) \leq \cos(\pi/2 + \varepsilon) = -\sin \varepsilon. \]

Hence we have the inequality
\[ |e^{\lambda z}| \leq e^{-|\lambda||z|\sin \varepsilon} \quad \text{for all } \lambda \in \Gamma^{(3)} \text{ and } z \in \Delta_{\omega}^{2\varepsilon}. \] (2.6)

Similarly, we have the inequality
\[ |e^{\lambda z}| \leq e^{-|\lambda||z|\sin \varepsilon} \quad \text{for all } \lambda \in \Gamma^{(1)} \text{ and } z \in \Delta_{\omega}^{2\varepsilon}. \] (2.7)

For each small \( \varepsilon > 0 \), we let
\[ K_{\varepsilon} = \Delta_{\omega}^{2\varepsilon} \cap \{ z \in C : |z| \geq \varepsilon \} = \{ z \in C : |z| \geq \varepsilon, |\arg z| \leq \omega - 2\varepsilon \}. \]

Then, by combining estimates (2.1), (2.6) and (2.7) we obtain that
\[ \|e^{\lambda z} R(\lambda)\| \leq \frac{M}{|\lambda|} e^{-\varepsilon \sin \varepsilon |\lambda|} \quad \text{for all } \lambda \in \Gamma^{(1)} \bigcup \Gamma^{(3)} \text{ and } z \in K_{\varepsilon}. \] (2.8)

On the other hand, we have the estimate
\[ \|e^{\lambda z} R(\lambda)\| \leq M e^{\varepsilon |z|} \quad \text{for all } \lambda \in \Gamma^{(2)} \text{ and } z \in K_{\varepsilon}. \] (2.9)

Therefore, we find that the integral
\[ U(z) = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} R(\lambda) d\lambda = -\frac{1}{2\pi i} \sum_{k=1}^{3} \int_{\Gamma^{(k)}} e^{\lambda z} R(\lambda) d\lambda \] (2.10)

converges in the Banach space \( L(E, E) \), uniformly in \( z \in K_{\varepsilon} \), for every \( \varepsilon > 0 \). This proves that the operator \( U(z) \) is analytic in the domain \( \Delta_{\omega} = \bigcup_{\varepsilon > 0} K_{\varepsilon} \).

By the analyticity of \( U(z) \), it follows that the operators \( U(z) \) also enjoy the semigroup property
2.1 Analytic Semigroups

\[ U(z + w) = U(z) \cdot U(w) \quad \text{for all } z, w \in \Delta_\omega. \]

(ii) We prove that the operators \( U(z) \) enjoy properties (a), (b) and (c).

(b) First, by using Cauchy’s theorem we obtain that

\[ U(z) = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} R(\lambda) d\lambda = -\frac{1}{2\pi i} \int_{\Gamma_{|z|}} e^{\lambda z} R(\lambda) d\lambda, \]

where \( \Gamma_{|z|} \) is a path consisting of the following three curves (see Figure 2.6):

\[ \Gamma_{|z|}^{(1)} = \left\{ re^{-i(\pi/2 + \omega - \epsilon)} : \frac{1}{|z|} \leq r < \infty \right\}, \]

\[ \Gamma_{|z|}^{(2)} = \left\{ \frac{1}{|z|} e^{i\theta} : -(\pi/2 + \omega - \epsilon) \leq \theta \leq \pi/2 + \omega - \epsilon \right\}, \]

\[ \Gamma_{|z|}^{(3)} = \left\{ re^{i(\pi/2 + \omega - \epsilon)} : \frac{1}{|z|} \leq r < \infty \right\}. \]

Fig. 2.6.

However, by estimates (2.1), (2.6) and (2.7), it follows that

\[ \|e^{\lambda z} R(\lambda)\| \leq \frac{M_\epsilon}{|\lambda|} e^{-|\lambda||z| \sin \epsilon} \quad \text{for all } \lambda \in \Gamma_{|z|}^{(1)} \cup \Gamma_{|z|}^{(3)} \text{ and } z \in \Delta_\omega^{2\epsilon}. \]

Hence we have, for \( k = 1, 3 \),

\[ \int_{\Gamma_{|z|}^{(k)}} \|e^{\lambda z} R(\lambda)\| \ d\lambda \leq M_\epsilon \int_1^\infty e^{-|\lambda| \sin \epsilon \rho^{-1}} d\rho \]

\[ = M_\epsilon \int_1^\infty e^{-\sin \epsilon s s^{-1}} ds. \]
We have also, for $k = 2$, 
\[
\int_{r^{(2)}_{|z|}} \|e^{\lambda z} R(\lambda)\| \, d\lambda \leq M_\varepsilon \int_{-\left(\frac{\pi}{2} + \omega - \varepsilon\right)}^{\frac{\pi}{2} + \omega - \varepsilon} e^{\theta} \mathrm{d}\theta \\
= 2\varepsilon M_\varepsilon \left(\frac{\pi}{2} + \omega - \varepsilon\right) \\
\leq 2\pi \varepsilon M_\varepsilon.
\]

Summing up, we obtain the following estimate:
\[
\|U(z)\| \leq \frac{1}{2\pi} \sum_{k=1}^{3} \int_{r^{(k)}_{|z|}} \|e^{\lambda z} R(\lambda)\| \, d\lambda \\
\leq \frac{1}{2\pi} \left(2\varepsilon M_\varepsilon \int_{1}^{\infty} s^{-1} e^{-\sin \varepsilon s} \, ds + 2\pi \varepsilon M_\varepsilon\right) \\
= \frac{M_\varepsilon}{\pi} \left(\int_{1}^{\infty} s^{-1} e^{-\sin \varepsilon s} \, ds + \pi \varepsilon\right).
\]

This proves the desired estimate (2.4), with
\[
\widetilde{M}_0(\varepsilon) = \frac{M_\varepsilon}{\pi} \left(\int_{1}^{\infty} s^{-1} e^{-\sin \varepsilon s} \, ds + \pi \varepsilon\right).
\]

To prove estimate (2.5), note that
\[
AR(\lambda) = (A - \lambda I + \lambda I)R(\lambda) = I + \lambda R(\lambda),
\]
so that
\[
\|AR(\lambda)\| \leq 1 + M_\varepsilon \quad \text{for all } \lambda \in \Sigma^\varepsilon_\omega.
\]
Hence, by arguing just as in the proof of estimate (2.4) we obtain that
\[
\left\|\int_{\Gamma} e^{\lambda z} AR(\lambda) \, d\lambda\right\| \leq 2 \int_{1}^{\infty} e^{-\rho |z| \sin \varepsilon} (1 + M_\varepsilon) \, d\rho \\
+ \int_{-\left(\frac{\pi}{2} + \omega - \varepsilon\right)}^{\frac{\pi}{2} + \omega - \varepsilon} e^{\theta} \mathrm{d}\theta \\
\leq 2 (1 + M_\varepsilon) \left(\int_{1}^{\infty} e^{-\sin \varepsilon s} \, ds + \pi \varepsilon\right) \frac{1}{|z|},
\]

for all $z \in \Delta^{2\varepsilon}_\omega$. (2.11)

This proves that the integral
\[
\int_{\Gamma} e^{\lambda z} AR(\lambda) \, d\lambda
\]
is convergent in the Banach space $L(E, E)$, for every $z \in \Delta^{2\varepsilon}_\omega$. By the closedness of $A$, it follows that
\[
U(z) \in \mathcal{D}(A) \quad \text{for all } z \in \Delta^{2\varepsilon}_\omega.
\]
and
\[ AU(z) = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} AR(\lambda) \, d\lambda \quad \text{for all } z \in \Delta_\omega^2. \] (2.12)

Therefore, the desired estimate (2.5) follows from estimate (2.11), with
\[ \tilde{M}_1(\varepsilon) = \frac{1 + M_\varepsilon}{\pi} \left( \int_1^\infty e^{-\sin \varepsilon s} \, ds + \pi \varepsilon \right). \]

We remark that formula (2.12) remains valid for all \( z \in \Delta_\omega \), since \( \Delta_\omega = \bigcup_{\varepsilon > 0} \Delta_\omega^{2\varepsilon} \).

(a) By estimates (2.8) and (2.9), we can differentiate formula (2.10) under the integral sign to obtain that
\[ \frac{dU}{dz}(z) = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} \lambda R(\lambda) \, d\lambda \quad \text{for all } z \in \Delta_\omega. \] (2.13)

On the other hand, it follows from formula (2.12) that
\[ AU(z) = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} AR(\lambda) \, d\lambda = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z}(I + \lambda R(\lambda)) \, d\lambda = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} \lambda R(\lambda) \, d\lambda \quad \text{for all } z \in \Delta_\omega, \] (2.14)

since we have, by Cauchy’s theorem,
\[ \int_{\Gamma} e^{\lambda z} \, d\lambda = 0. \]

Therefore, the desired formula (2.3) follows immediately from formulas (2.13) and (2.14).

(c) Now let \( x_0 \) be an arbitrary element of \( \mathcal{D}(A) \). By the residue theorem, it follows that
\[ x_0 = \frac{1}{2\pi i} \int_{\Gamma} \frac{e^{\lambda z}}{\lambda} x_0 \, d\lambda. \]

Hence we have the formula
\[ U(z)x_0 - x_0 = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} \left( R(\lambda) + \frac{1}{\lambda} \right) x_0 \, d\lambda = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} R(\lambda) A x_0 \, d\lambda. \]

Here we remark that
\[ \left\| \frac{1}{\lambda} R(\lambda) \right\| \leq \frac{M_\varepsilon}{|\lambda|^2} \quad \text{for all } \lambda \in \Gamma, \]
\[ |e^{\lambda z}| \leq 2e^{-|\lambda||z|\sin \varepsilon} + e^{|z|} \quad \text{for all } z \in \Delta_\omega^{2\varepsilon} \text{ and } \lambda \in \Gamma. \]
Thus it follows from an application of the Lebesgue dominated convergence theorem that, as \( z \to 0, \ z \in \Delta^{2\varepsilon}_\omega \),

\[
U(z)x_0 - x_0 \longrightarrow -\frac{1}{2\pi i} \int_{\Gamma} \frac{1}{\lambda} R(\lambda)Ax_0 \ d\lambda.
\]

However, we have the assertion

\[
\int_{\Gamma} \frac{1}{\lambda} R(\lambda)Ax_0 \ d\lambda = 0.
\]

Indeed, by Cauchy’s theorem it follows that

\[
\int_{\Gamma} \frac{1}{\lambda} R(\lambda)Ax_0 \ d\lambda = \lim_{r \to \infty} \int_{\Gamma \cap \{|\lambda| \leq r\}} \frac{1}{\lambda} R(\lambda)Ax_0 \ d\lambda
\]

\[
= -\lim_{r \to \infty} \int_{C_r} \frac{1}{\lambda} R(\lambda)Ax_0 \ d\lambda
\]

\[
= 0,
\]

where \( C_r \) is a closed path shown in Figure 2.7.

Summing up, we have proved that

\[
U(z)x_0 \longrightarrow x_0 \quad \text{as} \ z \to 0, \ z \in \Delta^{2\varepsilon}_\omega,
\]

for each \( x_0 \in D(A) \).

Since the domain \( D(A) \) is dense in \( E \) and \( \|U(z)\| \leq M_0(\varepsilon) \) for all \( z \in \Delta^{2\varepsilon}_\omega \), it follows that, for each \( x \in E \),

\[
U(z)x \longrightarrow x \quad \text{as} \ z \to 0, \ z \in \Delta^{2\varepsilon}_\omega.
\]

The proof of Theorem 2.2 is now complete. \( \square \)
Remark 2.1. Assume that the operator $A$ satisfies a stronger condition than condition (2.1):
\[
\|R(\lambda)\| \leq \frac{M_\varepsilon}{|\lambda| + 1} \quad \text{for all } \lambda \in \Sigma_\varepsilon. \tag{2.15}
\]
Then we have the estimates
\[
\|U(z)\| \leq \tilde{M}_0(\varepsilon)e^{-\delta \Re z} \quad \text{for all } z \in \Delta_\omega^2, \tag{2.16}
\]
\[
\|AU(z)\| \leq \frac{\tilde{M}_1(\varepsilon)}{|z|}e^{-\delta \Re z} \quad \text{for all } z \in \Delta_\omega^2, \tag{2.17}
\]
with some positive constant $\delta$.

Proof. Take a real number $\delta$ such that
\[
0 < \delta < \frac{1}{M_\varepsilon}.
\]
Then we have, by estimate (2.15),
\[
\delta \| (A - \lambda I)^{-1} \| \leq \frac{\delta M_\varepsilon}{|\lambda| + 1} \leq \delta M_\varepsilon < 1 \quad \text{for all } \lambda \in \Sigma_\varepsilon.
\]
Hence it follows that the operator $(A + \delta I) - \lambda I$ has the inverse
\[
((A + \delta I) - \lambda I)^{-1} = (I + \delta(A - \lambda I)^{-1})^{-1}(A - \lambda I)^{-1},
\]
and
\[
\|((A + \delta I) - \lambda I)^{-1}\| \leq \|(I + \delta(A - \lambda I)^{-1})^{-1}\| \cdot \|(A - \lambda I)^{-1}\|
\]
\[
\leq \frac{M_\varepsilon}{|\lambda| + 1 - \delta M_\varepsilon} \leq \frac{1}{1 - \delta M_\varepsilon} \cdot \frac{1}{|\lambda|}.
\]
This proves that the operator $A + \delta I$ satisfies condition (2.1), so that estimates (2.4) and (2.5) remain valid for the operator $A + \delta I$:
\[
\|V(z)\| \leq \tilde{M}_0(\varepsilon) \quad \text{for all } z \in \Delta_\omega^2, \tag{2.18}
\]
\[
\|(A + \delta I)V(z)\| \leq \frac{\tilde{M}_1(\varepsilon)}{|z|} \quad \text{for all } z \in \Delta_\omega^2, \tag{2.19}
\]
where
\[
V(z) = -\frac{1}{2\pi i} \int_R e^{\lambda z}(A + \delta I - \lambda I)^{-1} \, d\lambda.
\]
However, we have, by Cauchy’s theorem,

\[
V(z) = -\frac{1}{2\pi i} \int_{\Gamma} e^{\lambda z} (A + \delta I - \lambda I)^{-1} d\lambda
\]

\[
= -\frac{1}{2\pi i} \int_{\Gamma+\delta} e^{\lambda z} (A + \delta I - \lambda I)^{-1} d\lambda
\]

\[
= -\frac{1}{2\pi i} \int_{\Gamma} e^{\mu z} e^{\delta z} (A - \mu I)^{-1} d\mu = e^{\delta z} U(z)
\]

for all \( z \in \Delta_{2\varepsilon}^{2\varepsilon} \). (2.20)

In view of formula (2.16), the desired estimates (2.16) and (2.17) follow from estimates (2.18) and (2.19). \( \square \)

### 2.1.2 Fractional Powers

Assume that the operator \( A \) satisfies a stronger condition than condition (2.1):

1. The resolvent set of \( A \) contains the region \( \Sigma \) as in Figure 2.8.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{Fig2.8.png}
\caption{Fig. 2.8.}
\end{figure}

2. There exists a positive constant \( M \) such that the resolvent \( R(\lambda) = (A - \lambda I)^{-1} \) satisfies the estimate

\[
\| R(\lambda) \| \leq \frac{M}{(1 + |\lambda|)} \quad \text{for all } \lambda \in \Sigma. \quad (2.21)
\]

If \( \alpha > 0 \), we can define the fractional power \( (-A)^{-\alpha} \) of \( -A \) by the following formula:

\[
(-A)^{-\alpha} = -\frac{1}{2\pi i} \int_{\Gamma} (-\lambda)^{-\alpha} R(\lambda) \, d\lambda. \quad (2.22)
\]

Here the path \( \Gamma \) runs in the set \( \Sigma \) from \( \infty e^{-\omega} \) to \( \infty e^{i\omega} \), avoiding the positive real axis and the origin (see Figure 2.9), and for the function \( (-\lambda)^{-\alpha} = \)
\[ e^{-\alpha \log(-\lambda)}, \] we choose the branch whose argument lies between \(-\alpha \pi\) and \(\alpha \pi\); it is analytic in the region obtained by omitting the positive real axis.

The integral (2.22) converges in the uniform operator topology of the Banach space \(L(E, E)\) for all \(\alpha > 0\), and thus defines a bounded linear operator on \(E\).

![Figure 2.9](image)

Some basic properties of the fractional power \((-A)^{-\alpha}\) are summarized in the following:

**Proposition 2.3.** (i) We have, for all \(\alpha, \beta > 0\),
\[ (-A)^{-\alpha}(-A)^{-\beta} = (-A)^{-(\alpha+\beta)}. \]

(ii) If \(\alpha\) is a positive integer \(n\), then we have the formula
\[ (-A)^{-\alpha} = ((-A)^{-1})^n. \]

(iii) The fractional power \((-A)^{-\alpha}\) is invertible for all \(\alpha > 0\).

If \(0 < \alpha < 1\), we have the following useful formula for the fractional power \((-A)^{-\alpha}\):

**Theorem 2.4.** We have, for all \(0 < \alpha < 1\),
\[ (-A)^{-\alpha} = -\frac{\sin \alpha \pi}{\pi} \int_0^\infty s^{-\alpha} R(s) ds. \] (2.23)

By Remark 2.1, we may assume that there exist positive constants \(M_0\), \(M_1\) and \(a\) such that
\[
\|U(t)\| \leq M_0 e^{-at} \quad \text{for all } t > 0.
\]
\[
\|AU(t)\| \leq M_1 e^{-at} \frac{1}{t} \quad \text{for all } t > 0.
\]

Then we can prove still another useful formula for the fractional power \((-A)^{-\alpha}\) for all \(0 < \alpha < 1\).
Theorem 2.5. We have, for all $0 < \alpha < 1$,

$$(-A)^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} U(t) \, dt.$$ 

In view of part (iii) of Proposition 2.3, we can define the fractional power $(-A)^{\alpha}$ for all $\alpha > 0$ as follows:

$$(-A)^{\alpha} = \text{the inverse of } (-A)^{-\alpha}, \quad \alpha > 0.$$ 

The next theorem states that the domain $\mathcal{D}((-A)^{\alpha})$ of $(-A)^{\alpha}$ is bigger than the domain $\mathcal{D}(A)$ of $A$ when $0 < \alpha < 1$.

Theorem 2.6. We have, for all $0 < \alpha < 1$,

$$\mathcal{D}(A) \subset \mathcal{D}((-A)^{\alpha}).$$

We can give an explicit formula for the fractional power $(-A)^{\alpha}$ ($0 < \alpha < 1$) on the domain $\mathcal{D}(A)$:

Theorem 2.7. Let $0 < \alpha < 1$. Then we have, for any $x \in \mathcal{D}(A)$,

$$(-A)^{\alpha}x = \frac{\sin \alpha \pi}{\pi} \int_0^\infty s^{\alpha-1} R(s) Ax \, ds.$$

2.1.3 The Semilinear Cauchy Problem

This subsection is devoted to the semigroup approach to a class of initial-boundary value problems for semilinear parabolic differential equations. By making good use of fractional powers of analytic semigroups, we formulate a local existence and uniqueness theorem for semilinear initial-boundary value problems (Theorem 2.8). Our semigroup approach can be traced back to the pioneering work of Fujita–Kato [FK] for the Navier–Stokes equation in fluid dynamics.

Assume that the operator $A$ satisfies condition (2.21). Then we can define the fractional power $(-A)^{\alpha}$ for all $0 < \alpha < 1$ (see Subsection 2.1.2). The operator $(-A)^{\alpha}$ is a closed linear, invertible operator with domain $\mathcal{D}((-A)^{\alpha}) \supset \mathcal{D}(A)$. We let

$$E_\alpha = \text{the space } \mathcal{D}((-A)^{\alpha}) \text{ endowed with the graph norm } \| \cdot \|_\alpha \text{ of } (-A)^{\alpha},$$

where

$$\|x\|_\alpha = \left(\|x\|^2 + \|(-A)^{\alpha}x\|^2\right)^{1/2}, \quad x \in \mathcal{D}((-A)^{\alpha}).$$

Then we have the following three assertions:

(i) The space $E_\alpha$ is a Banach space.
The graph norm \( \|x\|_\alpha \) is equivalent to the norm \( \|(-A)^\alpha x\| \).

(ii) If \( 0 < \alpha < \beta < 1 \), then we have \( E_\beta \subset E_\alpha \) with continuous injection.

Now we consider the following semilinear Cauchy problem:

\[
\begin{cases}
\frac{du}{dt} = Au(t) + f(t, u(t)), & t_0 < t < t_1, \\
u(t_0) = x_0.
\end{cases}
\tag{2.24}
\]

Here \( f(t, x) \) is a function defined on an open set \( U \) of \( [0, \infty) \times E_\alpha \) (\( 0 < \alpha < 1 \)), taking values in \( E \). We assume that \( f(t, x) \) is locally Hölder continuous in \( t \) and locally Lipschitz continuous in \( x \). That is, for each point \((t, x)\) of \( U \) there exist a neighborhood \( V \subset U \), constants \( L = L(t, x, V) > 0 \) and \( 0 < \gamma \leq 1 \) such that

\[
\|f(s_1, y_1) - f(s_2, y_2)\| \leq L (|s_1 - s_2|^\gamma + \|y_1 - y_2\|_\alpha), \quad (s_1, y_1, (s_2, y_2) \in V.
\]

A function \( u(t) : [t_0, t_1) \rightarrow E \) is called a solution of problem (2.24) if it satisfies the following three conditions:

1. \( u(t) \in C([t_0, t_1); E) \cap C^1((t_0, t_1); E) \) and \( u(t_0) = x_0 \).
2. \( u(t) \in D(A) \) and \( (t, u(t)) \in U \) for all \( t_0 < t < t_1 \).
3. \( \frac{du}{dt} = Au(t) + f(t, u(t)) \) for all \( t_0 < t < t_1 \).

Here \( C([t_0, t_1); E) \) denotes the space of continuous functions on \([t_0, t_1)\) taking values in \( E \), and \( C^1((t_0, t_1); E) \) denotes the space of continuously differentiable functions on \((t_0, t_1)\) taking values in \( E \), respectively.

Our main result is the following local existence and uniqueness theorem for problem (2.24):

**Theorem 2.8.** Let \( f(t, x) \) be a function defined on an open subset \( U \) of \([0, \infty) \times E_\alpha \) (\( 0 < \alpha < 1 \)), taking values in \( E \). Assume that \( f(t, x) \) is locally Hölder continuous in \( t \) and locally Lipschitz continuous in \( x \). Then, for every \((t_0, x_0) \in U\), problem (2.24) has a unique local solution \( u(t) \in C([t_0, t_1); E) \cap C^1((t_0, t_1); E) \) where \( t_1 = t_1(t_0, x_0) > t_0 \).

For a proof of Theorem 2.8, the reader is referred to [He, Theorem 3.3.3], Pazy [Pa, Chapter 6, Theorem 3.1] and also Taira [Ta4, Theorem 1.18].

### 2.2 Markov Processes and Feller Semigroups

This section provides a brief description of basic definitions and results about Markov processes and a class of semigroups (Feller semigroups) associated with Markov processes. In Subsection 2.2.6 we prove various generation theorems of Feller semigroups by using the Hille–Yosida theory of semigroups (Theorems 2.16 and 2.18) which form a functional analytic background for the proof of Theorem 1.4. The results discussed here are adapted from
Blumenthal–Getoor [BG], Dynkin [Dy2], Lamperti [La], Revuz–Yor [RY] and also Taira [Ta2, Chapter 9]. The semigroup approach to Markov processes can be traced back to the pioneering work of Feller [Fe1] and [Fe2] in early 1950s (cf. [BCP], [SU], [Ta3]).

2.2.1 Markov Processes

In 1828 the English botanist R. Brown observed that pollen grains suspended in water move chaotically, incessantly changing their direction of motion. The physical explanation of this phenomenon is that a single grain suffers innumerable collisions with the randomly moving molecules of the surrounding water. A mathematical theory for Brownian motion was put forward by A. Einstein in 1905 (cf. [Ei]). Let \( p(t, x, y) \) be the probability density function that a one-dimensional Brownian particle starting at position \( x \) will be found at position \( y \) at time \( t \). Einstein derived the following formula from statistical mechanical considerations:

\[
p(t, x, y) = \frac{1}{\sqrt{2\pi Dt}} \exp \left[ -\frac{(y-x)^2}{2Dt} \right].
\]

Here \( D \) is a positive constant determined by the radius of the particle, the interaction of the particle with surrounding molecules, temperature and the Boltzmann constant. This gives an accurate method of measuring Avogadro’s number by observing particles. Einstein’s theory was experimentally tested by J. Perrin between 1906 and 1909.

Brownian motion was put on a firm mathematical foundation for the first time by N. Wiener in 1923 ([Wi]). Let \( \Omega \) be the space of continuous functions \( \omega : [0, \infty) \rightarrow \mathbb{R} \) with coordinates \( x_t(\omega) = \omega(t) \) and let \( \mathcal{F} \) be the smallest \( \sigma \)-algebra in \( \Omega \) which contains all sets of the form \( \{ \omega \in \Omega : a \leq x_t(\omega) < b \} \), \( t \geq 0, a < b \). Wiener constructed probability measures \( P_x, x \in \mathbb{R}, \) on \( \mathcal{F} \) for which the following formula holds:

\[
P_x \{ \omega \in \Omega : a_1 \leq x_{t_1}(\omega) < b_1, a_2 \leq x_{t_2}(\omega) < b_2, \ldots, a_n \leq x_{t_n}(\omega) < b_n \} = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} p(t_1, x, y_1)p(t_2 - t_1, y_1, y_2)\cdots p(t_n - t_{n-1}, y_{n-1}, y_n) dy_1 dy_2 \ldots dy_n,
\]

\[
0 < t_1 < t_2 < \ldots < t_n < \infty.
\]

This formula expresses the “starting afresh” property of Brownian motion that if a Brownian particle reaches a position, then it behaves subsequently as though that position had been its initial position. The measure \( P_x \) is called the Wiener measure starting at \( x \).

P. Lévy found another construction of Brownian motion, and gave a profound description of qualitative properties of the individual Brownian path in his book: Processus stochastiques et mouvement brownien (1948).

Markov processes are an abstraction of the idea of Brownian motion. Let \( K \) be a locally compact, separable metric space and let \( \mathcal{B} \) be the \( \sigma \)-algebra
of all Borel sets in $K$, that is, the smallest $\sigma$-algebra containing all open sets in $K$. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A function $X(\omega)$ defined on $\Omega$ taking values in $K$ is called a random variable if it satisfies the condition

$$X^{-1}(E) = \{\omega \in \Omega : X(\omega) \in E\} \in \mathcal{F} \text{ for all } E \in \mathcal{B}.$$ 

We express this by saying that $X$ is $\mathcal{F}/\mathcal{B}$-measurable. A family $\{x_t\}_{t \geq 0}$ of random variables is called a stochastic process, and it may be thought of as the motion in time of a physical particle. The space $K$ is called the state space and $\Omega$ the sample space. For a fixed $\omega \in \Omega$, the function $x_t(\omega)$, $t \geq 0$, defines in the state space $K$ a trajectory or path of the process corresponding to the sample point $\omega$.

In this generality the notion of a stochastic process is of course not so interesting. The most important class of stochastic processes is the class of Markov processes which is characterized by the Markov property. Intuitively, this is the principle of the lack of any “memory” in the system. More precisely, (temporally homogeneous) Markov property is that the prediction of subsequent motion of a particle, knowing its position at time $t$, depends neither on the value of $t$ nor on what has been observed during the time interval $[0,t)$; that is, a particle “starts afresh”.

Now we introduce a class of Markov processes which we will deal with in this book (cf. [Dy2], [BG], [RY]).

Assume that we are given the following:

1. A locally compact, separable metric space $K$ and the $\sigma$-algebra $\mathcal{B}$ of all Borel sets in $K$. A point $\partial$ is adjoined to $K$ as the point at infinity if $K$ is not compact, and as an isolated point if $K$ is compact (see Figure 2.10). We let

$$K_\partial = K \cup \{\partial\},$$

$$\mathcal{B}_\partial = \text{the } \sigma\text{-algebra in } K_\partial \text{ generated by } \mathcal{B}.$$ 

(Fig. 2.10.)

2. The space $\Omega$ of all mappings $\omega : [0, \infty] \to K_\partial$ such that $\omega(\infty) = \partial$ and that if $\omega(t) = \partial$ then $\omega(s) = \partial$ for all $s \geq t$. Let $\omega_\partial$ be the constant map $\omega_\partial(t) = \partial$ for all $t \in [0, \infty]$. 

(3) For each $t \in [0, \infty]$, the coordinate map $x_t$ defined by $x_t(\omega) = \omega(t)$, $\omega \in \Omega$.

(4) For each $t \in [0, \infty)$, a mapping $\varphi_t : \Omega \rightarrow \Omega$ defined by $(\varphi_t \omega)(s) = \omega(t+s)$, $\omega \in \Omega$. Note that $\varphi_\infty \omega = \omega_\partial$ and $x_t \circ \varphi_s = x_{t+s}$ for all $t, s \in [0, \infty]$.

(5) A $\sigma$-algebra $F$ in $\Omega$ and an increasing family $\{F_t\}_{0 \leq t \leq \infty}$ of sub-$\sigma$-algebras of $F$.

(6) For each $x \in K_\partial$, a probability measure $P_x$ on $(\Omega, F)$. We say that these elements define a (temporally homogeneous) Markov process $X = (x_t, F, F_t, P_x)$ if the following four conditions are satisfied:

(i) For each $0 \leq t < \infty$, the function $x_t$ is $F_t/\mathcal{B}_\partial$-measurable, that is,

$$\{x_t \in E\} = \{\omega \in \Omega : x_t(\omega) \in E\} \in F_t \quad \text{for all } E \in \mathcal{B}_\partial.$$

(ii) For all $0 \leq t < \infty$ and $E \in \mathcal{B}$, the function

$$p_t(x, E) = P_x\{x_t \in E\} \quad (2.25)$$

is a Borel measurable function of $x \in K$.

(iii) $P_x\{\omega \in \Omega : x_0(\omega) = x\} = 1$ for each $x \in K_\partial$.

(iv) For all $t, h \in [0, \infty)$, $x \in K_\partial$ and $E \in \mathcal{B}_\partial$, we have the formula

$$P_x\{x_{t+h} \in E \mid F_t\} = p_h(x_t, E) \quad \text{a. e.,}$$

or equivalently,

$$P_x(A \cap \{x_{t+h} \in E\}) = \int_A p_h(x_t(\omega), E) \, dP_x(\omega) \quad \text{for all } A \in F_t.$$

Here is an intuitive way of thinking about the above definition of a Markov process. The sub-$\sigma$-algebra $F_t$ may be interpreted as the collection of events which are observed during the time interval $[0, t]$. The value $P_x(A)$, $A \in F$, may be interpreted as the probability of the event $A$ under the condition that a particle starts at position $x$; hence the value $p_t(x, E)$ expresses the transition probability that a particle starting at position $x$ will be found in the set $E$ at time $t$ (see Figure 2.11). The function $p_t(x, \cdot)$ is called the transition function of the process $X$. The transition function $p_t(x, \cdot)$ specifies the probability structure of the process. The intuitive meaning of the crucial condition (iv) is that the future behavior of a particle, knowing its history up to time $t$, is the same as the behavior of a particle starting at $x_t(\omega)$, that is, a particle starts afresh. A particle moves in the space $K$ until it “dies” at the time when it reaches the point $\partial$; hence the point $\partial$ is called the terminal point.

With this interpretation in mind, we let

$$\zeta(\omega) = \inf\{t \in [0, \infty] : x_t(\omega) = \partial\}.$$

The random variable $\zeta$ is called the lifetime of the process $X$. 
2.2.2 Markov Transition Functions

In the first works devoted to Markov processes, the most fundamental was A. N. Kolmogorov's work ([Ko]) where the general concept of a Markov transition function was introduced for the first time and an analytic method of describing Markov transition functions was proposed. From the point of view of analysis, the transition function is something more convenient than the Markov process itself. In fact, it can be shown that the transition functions of Markov processes generate solutions of certain parabolic partial differential equations such as the classical diffusion equation; and, conversely, these differential equations can be used to construct and study the transition functions and the Markov processes themselves.

In the 1950s, the theory of Markov processes entered a new period of intensive development. We can associate with each transition function in a natural way a family of bounded linear operators acting on the space of continuous functions on the state space, and the Markov property implies that this family forms a semigroup. The Hille–Yosida theory of semigroups in functional analysis made possible further progress in the study of Markov processes, as will be shown in Subsection 2.2.5.

Our first job is thus to give the precise definition of a transition function adapted to the theory of semigroups:

Definition 2.1. Let $(K, \rho)$ be a locally compact, separable metric space and let $\mathcal{B}$ be the $\sigma$-algebra of all Borel sets in $K$. A function $p_t(x, E)$, defined for all $t \geq 0$, $x \in K$ and $E \in \mathcal{B}$, is called a (temporally homogeneous) Markov transition function on $K$ if it satisfies the following four conditions:

(a) $p_t(x, \cdot)$ is a non-negative measure on $\mathcal{B}$ and $p_t(x, K) \leq 1$ for all $t \geq 0$ and $x \in K$.

(b) $p_t(\cdot, E)$ is a Borel measurable function for all $t \geq 0$ and $E \in \mathcal{B}$.

(c) $p_0(x, \{x\}) = 1$ for all $x \in K$. 
(d) (The Chapman–Kolmogorov equation) For all \( t, s \geq 0, x \in K \) and \( E \in \mathcal{B} \), we have the equation

\[
p_{t+s}(x, E) = \int_K p_t(x, dy) p_s(y, E).
\]

(2.26)

Here is an intuitive way of thinking about the above definition of a Markov transition function. The value \( p_t(x, E) \) expresses the transition probability that a physical particle starting at position \( x \) will be found in the set \( E \) at time \( t \). The Chapman–Kolmogorov equation (2.26) expresses the idea that a transition from the position \( x \) to the set \( E \) in time \( t + s \) is composed of a transition from \( x \) to some position \( y \) in time \( t \), followed by a transition from \( y \) to the set \( E \) in the remaining time \( s \); the latter transition has probability \( p_s(y, E) \) which depends only on \( y \) (see Figure 2.12). Thus a particle “starts afresh”; this property is called the Markov property.

The Chapman–Kolmogorov equation (2.26) asserts that \( p_t(x, K) \) is monotonically increasing as \( t \downarrow 0 \), so that the limit

\[
p_{+0}(x, K) = \lim_{t \downarrow 0} p_t(x, K)
\]

exists.

A transition function \( p_t(x, \cdot) \) is said to be normal if it satisfies the condition

\[
p_{+0}(x, K) = 1 \quad \text{for all } x \in K.
\]

The next theorem, due to Dynkin [Dy1, Chapter 4, Section 2], justifies the definition of a transition function, and hence it will be fundamental for our further study of Markov processes:
Theorem 2.9. For every Markov process, the function $p_t(x, \cdot)$, defined by formula (2.25), is a Markov transition function. Conversely, every normal Markov transition function corresponds to some Markov process.

Here are some important examples of normal transition functions on the line $\mathbb{R} = (-\infty, \infty)$:

Example 2.1 (Uniform motion). If $t \geq 0$, $x \in \mathbb{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \chi_E(x + vt),$$

where $v$ is a constant, and $\chi_E(y) = 1$ if $y \in E$ and $= 0$ if $y \notin E$.

This process, starting at $x$, moves deterministically with constant velocity $v$.

Example 2.2 (Poisson process). If $t \geq 0$, $x \in \mathbb{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \chi_E(x + n),$$

where $\lambda$ is a positive constant.

This process, starting at $x$, advances one unit by jumps, and the probability of $n$ jumps during the time $0$ and $t$ is equal to $e^{-\lambda t} (\lambda t)^n / n!$.

Example 2.3 (Brownian motion). If $t > 0$, $x \in \mathbb{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \int_E \exp \left[ -\frac{(y - x)^2}{2t} \right] dy,$$

and

$$p_0(x, E) = \chi_E(x).$$

This is a mathematical model of one-dimensional Brownian motion. Its character is quite different from that of the Poisson process; the transition function $p_t(x, E)$ satisfies the condition

$$p_t(x, [x - \varepsilon, x + \varepsilon]) = 1 - o(t) \quad \text{as } t \downarrow 0,$$

for all $\varepsilon > 0$ and $x \in \mathbb{R}$. This means that the process never stands still, as does the Poisson process. Indeed, this process changes state not by jumps but by continuous motion. A Markov process with this property is called a diffusion process.

Example 2.4 (Brownian motion with constant drift). If $t > 0$, $x \in \mathbb{R}$ and $E \in \mathcal{B}$, we let

$$p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \int_E \exp \left[ -\frac{(y - mt - x)^2}{2t} \right] dy,$$

and

$$p_0(x, E) = \chi_E(x),$$

where $m$ is a constant.
This represents Brownian motion with a constant drift of magnitude \( m \) superimposed; the process can be represented as \( \{ x_t + mt \} \), where \( \{ x_t \} \) is Brownian motion on \( \mathbb{R} \).

**Example 2.5 (Cauchy process).** If \( t > 0 \), \( x \in \mathbb{R} \) and \( E \in \mathcal{B} \), we let

\[
p_t(x, E) = \frac{1}{\pi} \int_E \frac{t}{t^2 + (y - x)^2} \, dy,
\]
and

\[
p_0(x, E) = \chi_E(x).
\]

This process can be thought of as the “trace” on the real line of trajectories of two-dimensional Brownian motion, and it moves by jumps (see [Kn, Lemma 2.12]). More precisely, if \( B_1(t) \) and \( B_2(t) \) are two independent Brownian motions and if \( T \) is the first passage time of \( B_1(t) \) to \( x \), then \( B_2(T) \) has the Cauchy density

\[
\frac{1}{\pi} \frac{|x|}{x^2 + y^2}, \quad -\infty < y < \infty.
\]

Here are two more examples of diffusion processes on the half line \( \mathbb{R}^+ = [0, \infty) \) in which we must take account of the effect of the boundary point 0:

**Example 2.6 (Reflecting barrier Brownian motion).** If \( t > 0 \), \( x \in \mathbb{R}^+ \) and \( E \in \mathcal{B} \), we let

\[
p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \left( \int_E \exp \left[ \frac{(y - x)^2}{2t} \right] \, dy + \int_E \exp \left[ \frac{(y + x)^2}{2t} \right] \, dy \right),
\]
and

\[
p_0(x, E) = \chi_E(x).
\]

This represents Brownian motion with a reflecting barrier at \( x = 0 \); the process may be represented as \( \{|x_t|\} \), where \( \{x_t\} \) is Brownian motion on \( \mathbb{R} \). Indeed, since \( \{|x_t|\} \) goes from \( x \) to \( y \) if \( \{x_t\} \) goes from \( x \) to \( \pm y \) due to the symmetry of the transition function in Example 2.3 about \( x = 0 \), it follows that

\[
p_t(x, E) = P_x(\{|x_t| \in E\})
= \frac{1}{\sqrt{2\pi t}} \left( \int_E \exp \left[ \frac{(y - x)^2}{2t} \right] \, dy + \int_E \exp \left[ \frac{(y + x)^2}{2t} \right] \, dy \right).
\]

**Example 2.7 (Sticking barrier Brownian motion).** If \( t > 0 \), \( x \in \mathbb{R}^+ \) and \( E \in \mathcal{B} \), we let

\[
p_t(x, E) = \frac{1}{\sqrt{2\pi t}} \left( \int_E \exp \left[ -\frac{(y - x)^2}{2t} \right] \, dy - \int_E \exp \left[ -\frac{(y + x)^2}{2t} \right] \, dy \right) + \left( 1 - \frac{1}{\sqrt{2\pi t}} \int_{-x}^{x} \exp \left[ \frac{-z^2}{2t} \right] \, dz \right) \chi_E(0),
\]

\[

\]
and
\[ p_0(x, E) = \chi_E(x). \]

This represents Brownian motion with a sticking barrier at \( x = 0 \). When a Brownian particle reaches the boundary point 0 for the first time, instead of reflecting it sticks there forever; in this case the state 0 is called a trap.

### 2.2.3 Path Functions of Markov Processes

It is naturally interesting and important to ask the following problem:

**Problem.** Given a Markov transition function \( p_t(x, \cdot) \), under which conditions on \( p_t(x, \cdot) \) does there exist a Markov process with transition function \( p_t(x, \cdot) \) whose paths are almost surely continuous?

A Markov process \( X = (x_t, \mathcal{F}, \mathcal{F}_t, P_x) \) is said to be **right continuous** provided that we have, for each \( x \in K \),
\[
P_x\{\omega \in \Omega : \text{the mapping } t \mapsto x_t(\omega) \text{ is a right continuous function from } [0, \infty) \text{ into } K_{\partial}\} = 1.
\]

Furthermore, we say that \( X \) is **continuous** provided that we have, for each \( x \in K \),
\[
P_x\{\omega \in \Omega : \text{the mapping } t \mapsto x_t(\omega) \text{ is a continuous function from } [0, \zeta(\omega)) \text{ into } K_{\partial}\} = 1,
\]
where \( \zeta \) is the lifetime of the process \( X \).

Now we give some useful criteria for path continuity in terms of Markov transition functions (see Dynkin [Dy1, Chapter 6], [Dy2, Chapter 3, Section 2]):

**Theorem 2.10.** Let \((K, \rho)\) be a locally compact, separable metric space and let \( p_t(x, \cdot) \) be a normal Markov transition function on \( K \). Assume that the following two conditions are satisfied:

1. **(L)** For each \( s > 0 \) and each compact \( E \subset K \), we have the condition
   \[
   \lim_{x \to \partial} \sup_{0 \leq t \leq s} p_t(x, E) = 0.
   \]
2. **(M)** For each \( \varepsilon > 0 \) and each compact \( E \subset K \), we have the condition
   \[
   \lim_{t \downarrow 0} \sup_{x \in E} p_t(x, K \setminus U_\varepsilon(x)) = 0,
   \]
   where \( U_\varepsilon(x) = \{ y \in K : \rho(y, x) < \varepsilon \} \) is an \( \varepsilon \)-neighborhood of \( x \).
Then there exists a Markov process \( X \) with transition function \( p_t(x, \cdot) \) whose paths are right continuous on \([0, \infty)\) and have left-hand limits on \([0, \zeta)\) almost surely.

(ii) Assume that condition (L) and the following condition (N) (replacing condition (M)) are satisfied:

\[ \text{(N) For each } \varepsilon > 0 \text{ and each compact } E \subset K, \text{ we have the condition} \]
\[
\lim_{t \downarrow 0} \frac{1}{t} \sup_{x \in E} p_t(x, K \setminus U_\varepsilon(x)) = 0,
\]

or equivalently
\[
\sup_{x \in E} p_t(x, K \setminus U_\varepsilon(x)) = o(t) \text{ as } t \downarrow 0.
\]

Then there exists a Markov process \( X \) with transition function \( p_t(x, \cdot) \) whose paths are almost surely continuous on \([0, \zeta)\).

Remark 2.2. It is known (see Dynkin [Dy1, Lemma 6.2]) that if the paths of a Markov process are right continuous, then the transition function \( p_t(x, \cdot) \) satisfies the condition
\[
\lim_{t \downarrow 0} p_t(x, U_\varepsilon(x)) = 1 \text{ for all } x \in K.
\]

2.2.4 Strong Markov Processes and Transition Functions

A Markov process is called a strong Markov process if the “starting afresh” property holds not only for every fixed moment but also for suitable random times.

We shall formulate precisely this “strong” Markov property. Let \( \mathcal{X} = (x_t, \mathcal{F}, \mathcal{F}_t, P_x) \) be a Markov process. A mapping \( \tau: \Omega \to [0, \infty) \) is called a stopping time or Markov time with respect to \( \{ \mathcal{F}_t \} \) if it satisfies the condition
\[
\{ \tau \leq t \} = \{ \omega \in \Omega : \tau(\omega) \leq t \} \in \mathcal{F}_t \text{ for all } t \in [0, \infty).
\]

Intuitively, this means that the events \( \{ \tau \leq t \} \) depend on the process only up to time \( t \), but not on the “future” after time \( t \). It should be noticed that any non-negative constant mapping is a stopping time.

If \( \tau \) is a stopping time with respect to \( \{ \mathcal{F}_t \} \), we let
\[
\mathcal{F}_\tau = \{ A \in \mathcal{F} : A \cap \{ \tau \leq t \} \in \mathcal{F}_t \text{ for all } t \in [0, \infty) \}.
\]

Intuitively, we may think of \( \mathcal{F}_\tau \) as the “past” up to the random time \( \tau \). It is easy to verify that \( \mathcal{F}_\tau \) is a \( \sigma \)-algebra. If \( \tau \equiv t_0 \) for some constant \( t_0 \geq 0 \), then \( \mathcal{F}_\tau \) reduces to \( \mathcal{F}_{t_0} \).

For each \( t \in [0, \infty] \), we define a mapping
\[
\Phi_t: [0, t] \times \Omega \to K_0
\]
by the formula
\[ \Phi_t(s, \omega) = x_s(\omega). \]

A Markov process \( X = (x_t, \mathcal{F}, \mathcal{F}_t, P_x) \) is said to be \emph{progressively measurable} with respect to \( \{\mathcal{F}_t\} \) if the mapping \( \Phi_t \) is \( \mathcal{B}_{[0,t]} \times \mathcal{F}_t / \mathcal{B}_\partial \)-measurable for each \( t \in [0, \infty) \), that is, if we have the condition
\[ \Phi_t^{-1}(E) = \{ \Phi_t \in E \} \in \mathcal{B}_{[0,t]} \times \mathcal{F}_t \text{ for all } E \in \mathcal{B}_\partial. \]

Here \( \mathcal{B}_{[0,t]} \) is the \( \sigma \)-algebra of all Borel sets in the interval \([0, t]\) and \( \mathcal{B}_\partial \) is the \( \sigma \)-algebra in \( K_\partial \) generated by \( \mathcal{B} \). It should be noticed that if \( X \) is progressively measurable and if \( \tau \) is a stopping time, then the mapping \( x_\tau: \omega \mapsto x_\tau(\omega)(\omega) \) is \( \mathcal{F}_\tau / \mathcal{B}_\partial \)-measurable.

The next definition expresses the idea of “starting afresh” at random times:

\begin{definition}
We say that a progressively measurable Markov process \( X = (x_t, \mathcal{F}, \mathcal{F}_t, P_x) \) has the \emph{strong Markov property} with respect to \( \{\mathcal{F}_t\} \) if the following condition is satisfied: For all \( h \geq 0 \), \( x \in K_\partial \), \( E \in \mathcal{B}_\partial \) and all stopping times \( \tau \), we have the formula
\[ P_x \{ x_\tau + h \in E \mid \mathcal{F}_\tau \} = p_h(x_\tau, E), \]
or equivalently,
\[ P_x (A \cap \{ x_\tau + h \in E \}) = \int_A p_h(x_\tau(\omega), E) \, dP_x(\omega) \text{ for all } A \in \mathcal{F}_\tau. \]
\end{definition}

We shall state a simple criterion for the strong Markov property in terms of transition functions.

Let \((K, \rho)\) be a locally compact, separable metric space. We add a point \( \partial \) to the metric space \( K \) as the point at infinity if \( K \) is not compact, and as an isolated point if \( K \) is compact; so the space \( K_\partial = K \cup \{ \partial \} \) is compact (see Figure 2.10). Let \( C(K) \) be the space of real-valued, bounded continuous functions \( f(x) \) on \( K \); the space \( C(K) \) is a Banach space with the supremum norm
\[ \| f \|_\infty = \sup_{x \in K} |f(x)|. \]

We say that a function \( f \in C(K) \) converges to zero as \( x \to \partial \) if, for each \( \varepsilon > 0 \), there exists a compact subset \( E \) of \( K \) such that
\[ |f(x)| < \varepsilon \text{ for all } x \in K \setminus E, \]
and we then write \( \lim_{x \to \partial} f(x) = 0 \). We let
\[ C_0(K) = \left\{ f \in C(K) : \lim_{x \to \partial} f(x) = 0 \right\}. \]

The space \( C_0(K) \) is a closed subspace of \( C(K) \); hence it is a Banach space. Note that \( C_0(K) \) may be identified with \( C(K) \) if \( K \) is compact.
Now we introduce a useful convention as follows:

Any real-valued function \( f(x) \) on \( K \) is extended to the space \( K_\partial = K \cup \{ \partial \} \) by setting \( f(\partial) = 0 \).

From this point of view, the space \( C_0(K) \) is identified with the subspace of \( C(K_\partial) \) which consists of all functions \( f(x) \) satisfying the condition \( f(\partial) = 0 \):

\[
C_0(K) = \{ f \in C(K_\partial) : f(\partial) = 0 \}.
\]

Furthermore, we can extend a Markov transition function \( p_t(x, \cdot) \) on \( K \) to a Markov transition function \( p'_t(x, \cdot) \) on \( K_\partial \) by the formulas:

\[
\begin{aligned}
p'_t(x, E) &= p_t(x, E) \quad \text{for all } x \in K \text{ and } E \in \mathcal{B}, \\
p'_t(x, \{ \partial \}) &= 1 - p_t(x, K) \quad \text{for all } x \in K, \\
p'_t(\partial, K) &= 0, \quad p'_t(\partial, \{ \partial \}) = 1.
\end{aligned}
\]

Intuitively this means that a Markovian particle moves in the space \( K \) until it “dies” at the time when it reaches the point \( \partial \); hence the point \( \partial \) is called the terminal point.

Now we introduce some conditions on the measures \( p_t(x, \cdot) \) related to continuity in \( x \in K \), for fixed \( t \geq 0 \):

**Definition 2.3.** (i) A Markov transition function \( p_t(x, \cdot) \) is called a Feller function if the function

\[
T_t f(x) = \int_K p_t(x, dy) f(y)
\]

is a continuous function of \( x \in K \) whenever \( f \) is in \( C(K) \), that is, if we have the condition

\[
f \in C(K) \implies T_t f \in C(K).
\]

(ii) We say that \( p_t(x, \cdot) \) is a \( C_0 \)-function if the space \( C_0(K) \) is an invariant subspace of \( C(K) \) for the operators \( T_t \):

\[
f \in C_0(K) \implies T_t f \in C_0(K).
\]

**Remark 2.3.** The Feller property is equivalent to saying that the measures \( p_t(x, \cdot) \) depend continuously on \( x \in K \) in the usual weak topology, for every fixed \( t \geq 0 \).

The next theorem gives a useful criterion for the strong Markov property (see [Dy1, Theorem 5.10]):

**Theorem 2.11.** If the transition function \( p_t(x, \cdot) \) of a right continuous Markov process \( \mathcal{X} \) has the \( C_0 \)-property, then \( \mathcal{X} \) is a strong Markov process.
Furthermore, we state a simple criterion for the strong Markov property in terms of Markov transition functions. To do this, we introduce the following definition:

**Definition 2.4.** A Markov transition function \( p_t(x, \cdot) \) on \( K \) is said to be **uniformly stochastically continuous** on \( K \) if the following condition is satisfied: For each \( \varepsilon > 0 \) and each compact \( E \subset K \), we have the condition

\[
\lim_{t \downarrow 0} \sup_{x \in E} [1 - p_t(x, U_\varepsilon(x))] = 0,
\]

where \( U_\varepsilon(x) = \{ y \in K : \rho(y, x) < \varepsilon \} \) is an \( \varepsilon \)-neighborhood of \( x \).

It should be emphasized that every uniformly stochastically continuous transition function is normal and satisfies condition (M) in Theorem 2.10. By combining part (i) of Theorem 2.10 and Theorem 2.11, we obtain the following result (see [Dy1, Theorem 6.3]):

**Theorem 2.12.** Assume that a uniformly stochastically continuous, \( C_0 \)-transition function \( p_t(x, \cdot) \) satisfies condition (L). Then it is the transition function of some strong Markov process \( X \) whose paths are right continuous and have no discontinuities other than jumps.

A continuous strong Markov process is called a **diffusion process**.

The next theorem states a sufficient condition for the existence of a diffusion process with a prescribed Markov transition function:

**Theorem 2.13.** Assume that a uniformly stochastically continuous, \( C_0 \)-transition function \( p_t(x, \cdot) \) satisfies conditions (L) and (N). Then it is the transition function of some diffusion process \( X \).

This is an immediate consequence of part (ii) of Theorem 2.10 and Theorem 2.12.

### 2.2.5 Markov Transition Functions and Feller Semigroups

The Feller or \( C_0 \)-property deals with continuity of a Markov transition function \( p_t(x, E) \) in \( x \), and does not, by itself, have no concern with continuity in \( t \). We give a necessary and sufficient condition on \( p_t(x, E) \) in order that its associated operators \( \{ T_t \}_{t \geq 0} \), defined by the formula

\[
T_tf(x) = \int_K p_t(x, dy)f(y), \quad f \in C_0(K),
\]

is **strongly continuous** in \( t \) on the space \( C_0(K) \):

\[
\lim_{s \downarrow t} ||T_{t+s}f - T_tf||_\infty = 0, \quad f \in C_0(K).
\]
Then we have the following (cf. [Ta2, Theorem 9.2.3]):

**Theorem 2.14.** Let \( \mu_t(x, \cdot) \) be a \( C^0 \)-transition function on \( K \). Then the associated operators \( \{T_t\}_{t \geq 0} \), defined by formula (2.28) is strongly continuous in \( t \) on \( C_0(K) \) if and only if \( \mu_t(x, \cdot) \) is uniformly stochastically continuous on \( K \) and satisfies condition (L).

**Proof.** (i) The “if” part: Since continuous functions with compact support are dense in \( C_0(K) \), it suffices to prove the strong continuity of \( \{T_t\} \) at \( t = 0 \):

\[
\lim_{t \downarrow 0} \|T_t f - f\|_{\infty} = 0 \tag{2.30}
\]

for all such functions \( f \).

For any compact subset \( E \) of \( K \) containing the support \( \text{supp} f \) of \( f \), we have the inequality

\[
\|T_t f - f\|_{\infty} \leq \sup_{x \in E} |T_t f(x) - f(x)| + \sup_{x \in K \setminus E} |T_t f(x)|
\]

\[
\leq \sup_{x \in E} |T_t f(x) - f(x)| + \|f\|_{\infty} \cdot \sup_{x \in K \setminus E} \mu_t(x, \text{supp} f). \tag{2.31}
\]

However, condition (L) implies that, for each \( \epsilon > 0 \), we can find a compact subset \( E \) of \( K \) such that, for all sufficiently small \( t > 0 \),

\[
\sup_{x \in K \setminus E} \mu_t(x, \text{supp} f) < \epsilon. \tag{2.32}
\]

On the other hand, we have, for each \( \delta > 0 \),

\[
T_t f(x) - f(x) = \int_{U_\delta(x)} \mu_t(x, dy)(f(y) - f(x))
\]

\[
+ \int_{K \setminus U_\delta(x)} \mu_t(x, dy)(f(y) - f(x)) - f(x)(1 - \mu_t(x, K)),
\]

and hence

\[
\sup_{x \in E} |T_t f(x) - f(x)|
\]

\[
\leq \sup_{\rho(x, y) < \delta} |f(y) - f(x)| + 3\|f\|_{\infty} \cdot \sup_{x \in E} [1 - \mu_t(x, U_\delta(x))].
\]

Since the function \( f(x) \) is uniformly continuous, we can choose a positive constant \( \delta \) such that

\[
\sup_{\rho(x, y) < \delta} |f(y) - f(x)| < \epsilon.
\]

Furthermore, it follows from condition (2.27) with \( \epsilon := \delta \) (the uniform stochastic continuity of \( \mu_t(x, \cdot) \)) that, for all sufficiently small \( t > 0 \),
Hence we have, for all sufficiently small \( t > 0 \),
\[
\sup_{x \in E} |T_t f(x) - f(x)| < \varepsilon (1 + 3\|f\|_{\infty}).
\] (2.33)

Therefore, by carrying inequalities (2.32) and (2.33) into inequality (2.31) we obtain that, for all sufficiently small \( t > 0 \),
\[
\|T_t f - f\|_{\infty} < \varepsilon (1 + 4\|f\|_{\infty}).
\]

This proves the desired formula (2.30), that is, the strong continuity (2.29) of \( \{T_t\} \).

(ii) The “only if” part: For any \( x \in K \) and \( \varepsilon > 0 \), we define a continuous function \( f_x(y) \) by the formula (see Figure 2.13)
\[
f_x(y) = \begin{cases} 1 - \frac{1}{\varepsilon} \rho(x, y) & \text{if } \rho(x, y) \leq \varepsilon, \\ 0 & \text{if } \rho(x, y) > \varepsilon. \end{cases}
\] (2.34)

Let \( E \) be an arbitrary compact subset of \( K \). Then, for all sufficiently small \( \varepsilon > 0 \), the functions \( f_x, x \in E \), are in \( C_0(K) \) and satisfy the condition
\[
\|f_x - f_z\|_{\infty} \leq \frac{1}{\varepsilon} \rho(x, z) \text{ for all } x, z \in E.
\] (2.35)

However, for any \( \delta > 0 \), by the compactness of \( E \) we can find a finite number of points \( x_1, x_2, \ldots, x_n \) of \( E \) such that
\[
E = \bigcup_{k=1}^{n} U_{\delta \varepsilon/4}(x_k),
\]
and hence
\[
\min_{1 \leq k \leq n} \rho(x, x_k) \leq \frac{\delta \varepsilon}{4} \text{ for all } x \in E.
\]

Thus, by combining this inequality with inequality (2.35) with \( z := x_k \) we obtain that
\[
\min_{1 \leq k \leq n} \| f_x - f_{x_k} \|_\infty \leq \frac{\delta}{4} \quad \text{for all } x \in E.
\]

(2.36)

Now we have, by formula (2.34),

\[
0 \leq 1 - p_t(x, U_\varepsilon(x)) \leq 1 - \int_{K_0} p_t(x, dy) f_x(y)
\]

\[
= f_x(x) - T_t f_x(x)
\]

\[
\leq \| f_x - T_t f_x \|_\infty
\]

\[
\leq \| f_x - f_{x_k} \|_\infty + \| f_{x_k} - T_t f_{x_k} \|_\infty
\]

\[
+ \| T_t f_{x_k} - T_t f_x \|_\infty
\]

\[
\leq 2 \| f_x - f_{x_k} \|_\infty + \| f_{x_k} - T_t f_{x_k} \|_\infty
\]

for all \( x \in E \).

In view of inequality (2.36), the first term on the last inequality is bounded by \( \delta/2 \) for the right choice of \( k \). Furthermore, it follows from the strong continuity (2.30) of \( \{ T_t \} \) that the second term tends to zero as \( t \downarrow 0 \) for each \( k = 1, 2, \ldots, n \).

Consequently, we have, for all sufficiently small \( t > 0 \),

\[
\sup_{x \in E} [1 - p_t(x, U_\varepsilon(x))] \leq \delta.
\]

This proves the desired condition (2.27), that is, the uniform stochastic continuity of \( p_t(x, \cdot) \).

Finally, it remains to verify condition (L). Assume, to the contrary, that:

For some \( s > 0 \) and some compact \( E \subset K \), there exist a positive constant \( \varepsilon_0 \), a sequence \( \{ t_k \} \), \( t_k \downarrow t \) (\( 0 \leq t \leq s \)) and a sequence \( \{ x_k \} \), \( x_k \to \partial \), such that

\[
p_{t_k}(x_k, E) \geq \varepsilon_0.
\]

(2.37)

Now we take a relatively compact subset \( U \) of \( K \) containing \( E \), and let (see Figure 2.14)

\[
f(x) = \frac{\rho(x, K \setminus U)}{\rho(x, E) + \rho(x, K \setminus U)}.
\]

Then it follows that the function \( f(x) \) is in \( C_0(K) \) and satisfies the condition

\[
T_t f(x) = \int_K p_t(x, dy) f(y) \geq p_t(x, E) \geq 0.
\]

Therefore, by combining this inequality with inequality (2.37) we obtain that

\[
T_{t_k} f(x_k) \geq p_{t_k}(x_k, E) \geq \varepsilon_0.
\]

(2.38)

However, we have the inequality

\[
T_{t_k} f(x_k) \leq \| T_{t_k} f - T_t f \|_\infty + T_t f(x_k).
\]

(2.39)
Since the semigroup \( \{T_t\} \) is strongly continuous and since we have the assertion
\[
\lim_{k \to \infty} T_t f(x_k) = T_t f(\partial) = 0,
\]
we can let \( k \to \infty \) in inequality (2.39) to obtain that
\[
\limsup_{k \to \infty} T_{tk} f(x_k) = 0.
\]
This contradicts inequality (2.38).

The proof of Theorem 2.14 is now complete. \( \square \)

A family \( \{T_t\}_{t \geq 0} \) of bounded linear operators acting on the space \( C_0(K) \) is called a Feller semigroup on \( K \) if it satisfies the following three conditions:

(i) \( T_{t+s} = T_t \cdot T_s, t, s \geq 0 \) (the semigroup property); \( T_0 = I \).

(ii) The family \( \{T_t\} \) is strongly continuous in \( t \) for all \( t \geq 0 \):
\[
\lim_{s \downarrow 0} \|T_{t+s} f - T_t f\|_\infty = 0, \quad f \in C_0(K).
\]

(iii) The family \( \{T_t\} \) is non-negative and contractive on \( C_0(K) \):
\[
f \in C_0(K), \ 0 \leq f(x) \leq 1 \quad \text{on } K \implies 0 \leq T_t f(x) \leq 1 \quad \text{on } K.
\]

Rephrased, Theorem 2.14 gives a characterization of Feller semigroups in terms of Markov transition functions:

**Theorem 2.15.** If \( p_t(x, \cdot) \) is a uniformly stochastically continuous \( C_0 \)-transition function on \( K \) and satisfies condition (L), then its associated operators \( \{T_t\}_{t \geq 0} \), defined by formula (2.28), form a Feller semigroup on \( K \).

Conversely, if \( \{T_t\}_{t \geq 0} \) is a Feller semigroup on \( K \), then there exists a uniformly stochastically continuous \( C_0 \)-transition \( p_t(x, \cdot) \) on \( K \), satisfying condition (L), such that formula (2.28) holds.

The most important applications of Theorem 2.15 are of course in the second statement.
2.2.6 Generation Theorems of Feller Semigroups

In this subsection we prove various generation theorems of Feller semigroups by using the Hille–Yosida theory of semigroups.

If \( \{T_t\}_{t \geq 0} \) is a Feller semigroup on \( K \), we define its \textit{infinitesimal generator} \( A \) by the formula
\[
Au = \lim_{t \downarrow 0} \frac{T_t u - u}{t}, \quad u \in C_0(K),
\]
provided that the limit (2.40) exists in the space \( C_0(K) \). More precisely, the generator \( A \) is a linear operator from \( C_0(K) \) into itself defined as follows:

(1) The domain \( \mathcal{D}(A) \) of \( A \) is the set
\[
\mathcal{D}(A) = \{ u \in C_0(K) : \text{the limit (2.40) exists} \}.
\]

(2) \( Au = \lim_{t \downarrow 0} \frac{T_t u - u}{t}, \quad u \in \mathcal{D}(A) \).

The next theorem is a version of the Hille–Yosida theorem adapted to the present context (cf. [Ta2, Theorem 9.3.1 and Corollary 9.3.2]):

**Theorem 2.16 (Hille–Yosida).** (i) Let \( \{T_t\}_{t \geq 0} \) be a Feller semigroup on \( K \) and let \( A \) be its infinitesimal generator. Then we have the following four assertions:

(a) The domain \( \mathcal{D}(A) \) is dense in the space \( C_0(K) \).

(b) For each \( \alpha > 0 \), the equation \( (\alpha I - A)u = f \) has a unique solution \( u \) in \( \mathcal{D}(A) \) for any \( f \in C_0(K) \). Hence, for each \( \alpha > 0 \) the Green operator \( (\alpha I - A)^{-1} : C_0(K) \to C_0(K) \) can be defined by the formula
\[
u = (\alpha I - A)^{-1}f, \quad f \in C_0(K).
\]

(c) For each \( \alpha > 0 \), the operator \( (\alpha I - A)^{-1} \) is non-negative on \( C_0(K) \):
\[
f \in C_0(K), \quad f(x) \geq 0 \quad \text{on } K \implies (\alpha I - A)^{-1}f(x) \geq 0 \quad \text{on } K.
\]

(d) For each \( \alpha > 0 \), the operator \( (\alpha I - A)^{-1} \) is bounded on \( C_0(K) \) with norm
\[
\| (\alpha I - A)^{-1} \| \leq \frac{1}{\alpha}.
\]

(ii) Conversely, if \( A \) is a linear operator from \( C_0(K) \) into itself satisfying condition (a) and if there is a non-negative constant \( \alpha_0 \) such that, for all \( \alpha > \alpha_0 \), conditions (b) through (d) are satisfied, then \( A \) is the infinitesimal generator of some Feller semigroup \( \{T_t\}_{t \geq 0} \) on \( K \).

**Proof.** In view of the Hille–Yosida theory (see [Yo, Chapter IX, Section 7]), it suffices to show that the semigroup \( \{T_t\}_{t \geq 0} \) is non-negative if and only if its resolvents (Green operators) \( \{(\alpha I - A)^{-1}\}_{\alpha > \alpha_0} \) are non-negative.
The “only if” part is an immediate consequence of the following expression of \((\alpha I - A)^{-1}\) in terms of the semigroup \(\{T_t\}\):

\[
(\alpha I - A)^{-1} = \int_0^\infty \exp[-\alpha t] T_t \, dt, \quad \alpha > 0.
\]

On the other hand, the “if” part follows from the expression of the semigroup \(T_t(\alpha)\) in terms of the Yosida approximation \(J_\alpha = \alpha(\alpha I - A)^{-1}\):

\[
T_t(\alpha) = \exp[-\alpha t] \exp[\alpha t J_\alpha] = \exp[-\alpha t] \sum_{n=0}^{\infty} \frac{(\alpha t)^n}{n!} J_\alpha^n,
\]

and the definition of the semigroup \(T_t\):

\[
T_t = \lim_{\alpha \to \infty} T_t(\alpha).
\]

The proof of Theorem 2.16 is complete. □

**Corollary 2.17.** Let \(K\) be a compact metric space and let \(A\) be the infinitesimal generator of a Feller semigroup on \(K\). Assume that the constant function \(1\) belongs to the domain \(D(A)\) of \(A\) and that we have, for some constant \(c\),

\[
A1(x) \leq -c \quad \text{on} \ K. \tag{2.41}
\]

Then the operator \(A’ = A + cI\) is the infinitesimal generator of some Feller semigroup on \(K\).

**Proof.** It follows from an application of part (i) of Theorem 2.16 that, for all \(\alpha > c\), the operators

\[
(\alpha I - A’)^{-1} = ((\alpha - c)I - A)^{-1}
\]

are defined and non-negative on the whole space \(C(K)\). However, in view of inequality (2.41) we obtain that

\[
\alpha \leq \alpha - (A1 + c) = (\alpha I - A’)1 \quad \text{on} \ K,
\]

so that

\[
\alpha(\alpha I - A’)^{-1}1 \leq (\alpha I - A’)^{-1}(\alpha I - A’)1 = 1 \quad \text{on} \ K.
\]

Hence we have, for all \(\alpha > c\),

\[
\|\alpha I - A’\|^{-1} = \|\alpha I - A’\|^{-1} \leq \frac{1}{\alpha}.
\]

Therefore, by applying part (ii) of Theorem 2.16 to the operator \(A’\) we find that \(A’\) is the infinitesimal generator of some Feller semigroup on \(K\).

The proof of Corollary 2.17 is complete. □
Now we write down explicitly the infinitesimal generators of Feller semi-
groups associated with the transition functions in Examples 2.1 through 2.7
(cf. [DY]).

**Example 2.8 (Uniform motion).** $K = \mathbb{R}$ and

\[
\begin{align*}
\mathcal{D}(A) &= \{ f \in C_0(K) : f' \in C_0(K) \}, \\
Af &= vf', \quad f \in \mathcal{D}(A).
\end{align*}
\]

The operator $A$ is not “local”; the value $Af(x)$ depends on the values $f(x)$
and $f(x + 1)$. This reflects the fact that the Poisson process changes state by
jumps.

**Example 2.9 (Poisson process).** $K = \mathbb{R}$ and

\[
\begin{align*}
\mathcal{D}(A) &= C_0(K), \\
Af(x) &= \lambda(f(x + 1) - f(x)), \quad f \in \mathcal{D}(A).
\end{align*}
\]

The operator $A$ is not “local”; the value $Af(x)$ depends on the values $f(x)$
and $f(x + 1)$. This reflects the fact that the Poisson process changes state by
jumps.

**Example 2.10 (Brownian motion).** $K = \mathbb{R}$ and

\[
\begin{align*}
\mathcal{D}(A) &= \{ f \in C_0(K) : f' \in C_0(K), f'' \in C_0(K) \}, \\
Af &= \frac{1}{2}f'', \quad f \in \mathcal{D}(A).
\end{align*}
\]

The operator $A$ is “local”, that is, the value $Af(x)$ is determined by the
values of $f$ in an arbitrary small neighborhood of $x$. This reflects the fact that
Brownian motion changes state by continuous motion.

**Example 2.11 (Brownian motion with constant drift).** $K = \mathbb{R}$ and

\[
\begin{align*}
\mathcal{D}(A) &= \{ f \in C_0(K) : f' \in C_0(K), f'' \in C_0(K) \}, \\
Af &= \frac{1}{2}f'' + mf', \quad f \in \mathcal{D}(A).
\end{align*}
\]

**Example 2.12 (Cauchy process).** $K = \mathbb{R}$ and, the domain $\mathcal{D}(A)$ contains $C^2$
functions on $K$ with compact support, and the infinitesimal generator $A$ is of
the form

\[
Af(x) = \frac{1}{\pi} \int_{-\infty}^{+\infty} (f(x + y) - f(x)) \frac{dy}{y^2}.
\]

The operator $A$ is not “local”, which reflects the fact that the Cauchy
process changes state by jumps.

**Example 2.13 (Reflecting barrier Brownian motion).** $K = [0, \infty)$ and

\[
\begin{align*}
\mathcal{D}(A) &= \{ f \in C_0(K) : f' \in C_0(K), f'' \in C_0(K), f'(0) = 0 \}, \\
Af &= \frac{1}{2}f'', \quad f \in \mathcal{D}(A).
\end{align*}
\]
Example 2.14 (Sticking barrier Brownian motion). $K = [0, \infty)$ and
\[
\begin{align*}
\mathcal{D}(A) &= \{f \in C_0(K) : f' \in C_0(K), f'' \in C_0(K), f''(0) = 0\}, \\
Af &= \frac{1}{2}f'', \quad f \in \mathcal{D}(A).
\end{align*}
\]

Finally, here are two more examples where it is difficult to begin with a transition function and the infinitesimal generator is the basic tool of describing the process.

Example 2.15 (Sticky barrier Brownian motion). $K = [0, \infty)$ and
\[
\begin{align*}
\mathcal{D}(A) &= \{f \in C_0(K) : f' \in C_0(K), f'' \in C_0(K), f'(0) - \alpha f''(0) = 0\}, \\
Af &= \frac{1}{2}f'', \quad f \in \mathcal{D}(A).
\end{align*}
\]
Here $\alpha$ is a positive constant.

This process may be thought of as a “combination” of the reflecting and sticking Brownian motions. The reflecting and sticking cases are obtained by letting $\alpha \to 0$ and $\alpha \to \infty$, respectively.

Example 2.16 (Absorbing barrier Brownian motion). $K = [0, \infty)$ where the boundary point $0$ is identified with the point at infinity $\partial$.
\[
\begin{align*}
\mathcal{D}(A) &= \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K), f'' \in C_0(K), f(0) = 0\}, \\
Af &= \frac{1}{2}f'', \quad f \in \mathcal{D}(A).
\end{align*}
\]
This represents Brownian motion with an absorbing barrier at $x = 0$; a Brownian particle “dies” at the first moment when it hits the boundary $x = 0$. Namely, the point 0 is the terminal point.

It is worth pointing out here that a strong Markov process cannot stay at a single position for a positive length of time and then leave that position by continuous motion; it must either jump away or leave instantaneously.

We give a simple example of a strong Markov process which changes state not by continuous motion but by jumps when the motion reaches the boundary.

Example 2.17. $K = [0, \infty)$.
\[
\begin{align*}
\mathcal{D}(A) &= \{f \in C_0(K) \cap C^2(K) : f' \in C_0(K), f'' \in C_0(K), f''(0) = 2c \int_0^\infty (f(y) - f(0))dF(y), \\
Af &= \frac{1}{2}f'', \quad f \in \mathcal{D}(A).
\end{align*}
\]
Here $c$ is a positive constant and $F$ is a distribution function on the interval $(0, \infty)$.

This process may be interpreted as follows. When a Brownian particle reaches the boundary $x = 0$, it stays there for a positive length of time and then jumps back to a random point, chosen with the function $F$, in the interior
(0, ∞). The constant $c$ is the parameter in the “waiting time” distribution at the boundary $x = 0$. We remark that the boundary condition

$$f''(0) = 2c \int_0^\infty (f(y) - f(0)) dF(y)$$

depends on the values of $f$ far away from the boundary $x = 0$, unlike the boundary conditions in Examples 2.13 through 2.16.

Although Theorem 2.16 asserts precisely when a linear operator $A$ is the infinitesimal generator of some Feller semigroup, it is usually difficult to verify conditions (b) through (d). So we give useful criteria in terms of the maximum principle (see [BCP], [SU], [Ra], [Ta2, Theorem 9.3.3 and Corollary 9.3.4]):

**Theorem 2.18 (Hille–Yosida–Ray).** Let $K$ be a compact metric space. Then we have the following two assertions:

(i) Let $B$ be a linear operator from $C(K) = C_0(K)$ into itself, and assume that:

(a) The domain $\mathcal{D}(B)$ of $B$ is dense in the space $C(K)$.

(b) There exists an open and dense subset $K_0$ of $K$ such that if a function $u \in \mathcal{D}(B)$ takes a positive maximum at a point $x_0$ of $K_0$, then we have the inequality

$$Bu(x_0) \leq 0.$$  

Then the operator $B$ is closable in the space $C(K)$.

(ii) Let $B$ be as in part (i), and further assume that:

(b′) If a function $u \in \mathcal{D}(B)$ takes a positive maximum at a point $x'$ of $K$, then we have the inequality

$$Bu(x') \leq 0.$$  

(γ) For some $\alpha_0 \geq 0$, the range $\mathcal{R}(\alpha_0 I - B)$ of $\alpha_0 I - B$ is dense in the space $C(K)$.

Then the minimal closed extension $\overline{B}$ of $B$ is the infinitesimal generator of some Feller semigroup on $K$.

**Proof.** (i) It suffices to show that:

$$\{u_n\} \subset \mathcal{D}(B), \ u_n \to 0 \quad \text{and} \quad Bu_n \to v \quad \text{in} \ C(K) \implies v = 0.$$  

By replacing $v$ by $-v$ if necessary, we assume, to the contrary, that:

The function $v(x)$ takes a positive value at some point of $K$.

Then, since $K_0$ is open and dense in $K$, we can find a point $x_0$ of $K_0$, a neighborhood $U$ of $x_0$ contained in $K_0$ and a positive constant $\varepsilon$ such that we have, for all sufficiently large $n$,

$$Bu_n(x) > \varepsilon \quad \text{for all} \ x \in U.$$  

(2.42)
On the other hand, by condition \((\alpha)\) there exists a function \(h \in \mathcal{D}(B)\) such that
\[
\begin{cases}
h(x_0) > 1, \\
h(x) < 0 \quad \text{for all } x \in K \setminus U.
\end{cases}
\]
Therefore, since \(u_n \to 0\) in \(C(K)\), it follows that the function
\[
u'_n(x) = u_n(x) + \frac{\varepsilon h(x)}{1 + \|Bh\|_{\infty}}
\]
satisfies the conditions
\[
\begin{align*}
u'_n(x_0) &= u_n(x_0) + \frac{\varepsilon h(x_0)}{1 + \|Bh\|_{\infty}} > 0, \\
u'_n(x) &= u_n(x) + \frac{\varepsilon h(x)}{1 + \|Bh\|_{\infty}} < 0 \quad \text{for all } x \in K \setminus U,
\end{align*}
\]
if \(n\) is sufficiently large. This implies that the function \(u'_n \in \mathcal{D}(B)\) takes its positive maximum at a point \(x'_n\) of \(U \subset K_0\). Hence we have, by condition \((\beta)'\),
\[
Bu'_n(x'_n) \leq 0.
\]
However, it follows from inequality (2.42) that
\[
Bu'_n(x'_n) = Bu_n(x'_n) + \varepsilon \frac{Bh(x'_n)}{1 + \|Bh\|_{\infty}} > Bu_n(x'_n) - \varepsilon > 0.
\]
This is a contradiction.

(ii) We apply part (ii) of Theorem 2.16 to the minimal closed extension \(\overline{B}\) of \(B\). The proof is divided into six steps.

**Step 1**: First, we show that
\[
u \in \mathcal{D}(B), \quad (\alpha_0 I - B)u \geq 0 \quad \text{on } K \implies u \geq 0 \quad \text{on } K. \tag{2.43}
\]
By condition \((\gamma)\), we can find a function \(v \in \mathcal{D}(B)\) such that
\[
(\alpha_0 I - B)v \geq 1 \quad \text{on } K. \tag{2.44}
\]
Then we have, for any \(\varepsilon > 0\),
\[
\begin{cases}
u + \varepsilon v \in \mathcal{D}(B), \\
(\alpha_0 I - B)(\nu + \varepsilon v) \geq \varepsilon \quad \text{on } K.
\end{cases}
\]
In view of condition \((\beta')\), this implies that the function \(-(\nu(x) + \varepsilon v(x))\) does not take any positive maximum on \(K\), so that
\[
u(x) + \varepsilon v(x) \geq 0 \quad \text{on } K.
\]
Thus, by letting $\varepsilon \downarrow 0$ in this inequality we obtain that
\[ u(x) \geq 0 \quad \text{on } K. \]

This proves the desired assertion (2.43).

**Step 2**: It follows from assertion (2.43) that the inverse $(\alpha_0 I - B)^{-1}$ of $\alpha_0 I - B$ is defined and non-negative on the range $R(\alpha_0 I - B)$. Moreover, it is bounded with norm
\[ \|(\alpha_0 I - B)^{-1}\| \leq \|v\|_\infty. \quad (2.45) \]

Here $v(x)$ is the function which satisfies condition (2.44).

Indeed, since $g = (\alpha_0 I - B)v \geq 1$ on $K$, it follows that, for all $f \in C(K)$,
\[ -\|f\|_\infty g \leq f \leq \|f\|_\infty g \quad \text{on } K. \]

Hence, by the non-negativity of $(\alpha_0 I - B)^{-1}$ we have, for all $f \in R(\alpha_0 I - B)$,
\[ -\|f\|_\infty v \leq (\alpha_0 I - B)^{-1}f \leq \|f\|_\infty v \quad \text{on } K. \]

This proves the desired inequality (2.45).

**Step 3**: Next we show that $R(\alpha_0 I - B) = C(K)$. (2.46)

Let $f(x)$ be an arbitrary element of $C(K)$. By condition (γ), we can find a sequence $\{u_n\}$ in $D(B)$ such that $f_n = (\alpha_0 I - B)u_n \to f$ in $C(K)$. Since the inverse $(\alpha_0 I - B)^{-1}$ is bounded, it follows that $u_n = (\alpha_0 I - B)^{-1}f_n$ converges to some function $u \in C(K)$, and hence $Bu_n = \alpha_0 u_n - f_n$ converges to $\alpha_0 u - f$ in $C(K)$. Thus we have, by the closedness of $B$,
\[ \left\{ \begin{array}{l} u \in D(B), \\ Bu = \alpha_0 u - f, \end{array} \right. \]
so that
\[ (\alpha_0 I - B)u = f. \]

This proves the desired assertion (2.46).

**Step 4**: Furthermore, we show that
\[ u \in D(B), \quad (\alpha_0 I - B)u \geq 0 \quad \text{on } K \implies u \geq 0 \quad \text{on } K. \quad (2.47) \]

Since $R(\alpha_0 I - B) = C(K)$, in view of the proof of assertion (2.47) it suffices to show the following:

If a function $u \in D(B)$ takes a positive maximum at a point $x'$ of $K$, then we have the inequality
\[ Bu(x') \leq 0. \quad (2.48) \]

Assume, to the contrary, that
\[ Bu(x') > 0. \]
Since there exists a sequence \( \{u_n\} \) in \( \mathcal{D}(B) \) such that \( u_n \to u \) and \( Bu_n \to Bu \) in \( C(K) \), we can find a neighborhood \( U \) of \( x' \) and a positive constant \( \varepsilon \) such that, for all sufficiently large \( n \),

\[
Bu_n(x) > \varepsilon \quad \text{for all } x \in U. \tag{2.49}
\]

Furthermore, by condition (\( \alpha \)) we can find a function \( h \in \mathcal{D}(B) \) such that

\[
\begin{cases}
h(x') > 1, \\ h(x) < 0 \quad \text{for all } x \in K \setminus U.
\end{cases}
\]

Then it follows that the function

\[
u'_n(x) = u_n(x) + \frac{\varepsilon h(x)}{1 + \|Bh\|_{\infty}}
\]

satisfies the condition

\[
\begin{cases}
u'_n(x') > u(x') > 0, \\ u'_n(x) < u(x') \quad \text{for all } x \in K \setminus U,
\end{cases}
\]

if \( n \) is sufficiently large. This implies that the function \( u'_n \in \mathcal{D}(B) \) takes its positive maximum at a point \( x'_n \) of \( U \). Hence we have, by condition (\( \beta' \)),

\[
Bu'_n(x'_n) \leq 0, \quad x'_n \in U.
\]

However, it follows from inequality (2.49) that

\[
Bu'_n(x'_n) = Bu_n(x'_n) + \varepsilon \frac{Bh(x'_n)}{1 + \|Bh\|_{\infty}} > Bu_n(x'_n) - \varepsilon > 0.
\]

This is a contradiction.

**Step 5:** In view of Steps 3 and 4, we obtain that the inverse \((\alpha_0 I - B)^{-1}\) of \( \alpha_0 I - B \) is defined on the whole space \( C(K) \), and is bounded with norm

\[
\|((\alpha_0 I - B)^{-1})\| = \|((\alpha_0 I - B)^{-1})1\|_{\infty}.
\]

**Step 6:** Finally, we show that:

For all \( \alpha > \alpha_0 \), the inverse \((\alpha I - B)^{-1}\) of \( \alpha I - B \) is defined on the whole space \( C(K) \), and is non-negative and bounded with norm

\[
\|((\alpha I - B)^{-1})\| \leq \frac{1}{\alpha}. \tag{2.50}
\]

We let

\[
G_{\alpha_0} = (\alpha_0 I - B)^{-1}.
\]

First, we choose a constant \( \alpha_1 > \alpha_0 \) such that

\[
(\alpha_1 - \alpha_0)\|G_{\alpha_0}\| < 1,
\]
and let
\[ \alpha_0 < \alpha \leq \alpha_1. \]
Then, for any \( f \in C(K) \), the Neumann series
\[
u = \left( I + \sum_{n=1}^{\infty} (\alpha_0 - \alpha)^n G_{\alpha_0} \right) G_{\alpha_0} f
\]
converges in \( C(K) \), and is a solution of the equation
\[ u - (\alpha_0 - \alpha) G_{\alpha_0} u = G_{\alpha_0} f. \]
Hence we have the assertions
\[
\left\{ \begin{array}{l}
u \in D(B), \\
(\alpha I - B) \nu \geq 0 \text{ on } K \iff \nu \geq 0 \text{ on } K.
\end{array} \right.
\]
This proves that
\[
R(\alpha I - B) = C(K), \quad \alpha_0 < \alpha \leq \alpha_1.
\] (2.51)
Thus, by arguing just as in the proof of Step 1 we obtain that, for any \( \alpha_0 < \alpha \leq \alpha_1 \),
\[ u \in D(B), \quad (\alpha I - B) u \geq 0 \text{ on } K \implies u \geq 0 \text{ on } K. \] (2.52)
By combining assertions (2.51) and (2.52), we find that, for any \( \alpha_0 < \alpha \leq \alpha_1 \),
the inverse \((\alpha I - B)^{-1}\) is defined and non-negative on the whole space \( C(K) \).
We let
\[ G_{\alpha} = (\alpha I - B)^{-1}, \quad \alpha_0 < \alpha \leq \alpha_1. \]
Then it follows that the operator \( G_{\alpha} \) is bounded with norm
\[
\|G_{\alpha}\| \leq \frac{1}{\alpha}.
\] (2.53)
Indeed, in view of assertion (2.48) it follows that if a function \( u \in D(B) \) takes a positive maximum at a point \( x' \) of \( K \), then we have the inequality
\[ Bu(x') \leq 0, \]
so that
\[
\max_{x \in K} u(x) = u(x') \leq \frac{1}{\alpha} (\alpha I - B) u(x') \leq \frac{1}{\alpha} \| (\alpha I - B) u \|_\infty.
\] (2.54)
Similarly, if the function \( u \in D(B) \) takes a negative minimum at a point of \( K \), then (replacing \( u(x) \) by \(-u(x))\), we have the inequality
\[
-\min_{x \in K} u(x) = \max_{x \in K} (-u(x)) \leq \frac{1}{\alpha} \| (\alpha I - B) u \|_\infty.
\] (2.55)
The desired inequality (2.53) follows from inequalities (2.54) and (2.55).

Summing up, we have proved assertion (2.50) for all \(0 < \alpha \leq \alpha_1\).

Now we assume that assertion (2.50) is proved for all \(0 < \alpha_0 < \alpha \leq \alpha_{n-1}\), \(n = 2, 3, \ldots\). Then, by taking

\[
\alpha_n = 2\alpha_{n-1} - \frac{\alpha_1}{2}, \quad n \geq 2,
\]

or equivalently

\[
\alpha_n = \left(2^{n-2} + \frac{1}{2}\right)\alpha_1, \quad n \geq 2,
\]

we have, for all \(\alpha_{n-1} < \alpha \leq \alpha_n\),

\[
(\alpha - \alpha_{n-1})\|G_{\alpha_{n-1}}\| \leq \frac{\alpha - \alpha_{n-1}}{\alpha_{n-1}} \leq \frac{\alpha_n - \alpha_{n-1}}{\alpha_{n-1}} = \frac{1}{1 + 2^{n-2}} < 1.
\]

Hence assertion (2.50) for \(\alpha_{n-1} < \alpha \leq \alpha_n\) is proved just as in the proof of assertion (2.50) for \(\alpha_0 < \alpha \leq \alpha_1\). This proves the desired assertion (2.50).

Consequently, by applying part (ii) of Theorem 2.16 to the operator \(B\) we obtain that \(B\) is the infinitesimal generator of some Feller semigroup on \(K\).

The proof of Theorem 2.18 is now complete. \(\square\)

**Corollary 2.19.** Let \(A\) be the infinitesimal generator of a Feller semigroup on a compact metric space \(K\) and let \(M\) be a bounded linear operator on \(C(K)\) into itself. Assume that either \(M\) or \(A' = A + M\) satisfies condition \((\beta')\). Then the operator \(A'\) is the infinitesimal generator of some Feller semigroup on \(K\).

**Proof.** We apply part (ii) of Theorem 2.18 to the operator \(A'\).

First, we remark that \(A' = A + M\) is a densely defined, closed linear operator from \(C(K)\) into itself. Since the semigroup \(\{T_t\}_{t \geq 0}\) is non-negative and contractive on \(C(K)\), it follows that if a function \(u \in \mathcal{D}(A)\) takes a positive maximum at a point \(x'\) of \(K\), then we have the inequality

\[
Au(x') = \lim_{t \downarrow 0} \frac{T_t u(x') - u(x')}{t} \leq 0.
\]

This implies that if \(M\) satisfies condition \((\beta')\), so does \(A' = A + M\).

We let

\[
G_{\alpha_0} = (\alpha_0 I - A)^{-1}, \quad \alpha_0 > 0.
\]

If \(\alpha_0\) is so large that

\[
\|G_{\alpha_0}M\| \leq \|G_{\alpha_0}\| \cdot \|M\| \leq \frac{\|M\|}{\alpha_0} < 1,
\]

...
then the Neumann series

\[ u = \left( I + \sum_{n=1}^{\infty} (G_{\alpha_0}M)^n \right) G_{\alpha_0}f \]

converges in \( C(K) \) for any \( f \in C(K) \), and is a solution of the equation

\[ u - G_{\alpha_0}Mu = G_{\alpha_0}f. \]

Hence we have the assertions

\[ \begin{cases} 
  u \in D(A) = D(A'), \\
  (\alpha_0 I - A')u = f.
\end{cases} \]

This proves that

\[ \mathcal{R}(\alpha_0 I - A') = C(K). \]

Therefore, by applying part (ii) of Theorem 2.18 to the operator \( A' \) we obtain that \( A' \) is the infinitesimal generator of some Feller semigroup on \( K \).

The proof of Corollary 2.19 is complete. \( \square \)
Boundary Value Problems and Markov Processes
Taira, K.
2009, XII, 192 p. 41 illus., Softcover
ISBN: 978-3-642-01676-9