1. The Theorem of Hille and Yosida Concerning Semi-groups

From now we consider $X$ to be a Banach space.

**Definition 1.1.** A family of bounded operators $P_t : X \to X, t \geq 0$ is a strongly continuous semi-group if

a) $P_0 = I$.  

b) For any $s, t \geq 0$ we have that $P_{s+t} = P_s P_t$.  

c) For any $f \in X$, $\lim_{t \to 0, t > 0} \| P_t f - f \| = 0$.

Note that $t \to P_t f$ is continuous for all $t \geq 0$, since

$$\lim_{\varepsilon \to 0} \| P_{t+\varepsilon} f - P_tf \| = \lim_{\varepsilon \to 0} \| P_{t} P_t f - P_tf \| = 0.$$  

Such semi groups are natural in the context of linear evolution equations. An important sub-class are the contraction semi-groups.

**Definition 1.2.** A family of bounded operators $P_t : X \to X, t \geq 0$ is a contraction semi-group if is a strongly continuous semi-group and for all $t \geq 0$

$$\| P_t \| \leq 1.$$  

One would like to think of a semi group as an operator of the form $e^{At}$ for some operator which one would call the generator of the semi-group. Let $P_t$ be a contraction semi-group. Consider the set

$$D(A) = \{ f \in X : \lim_{t \to 0} P_t f - f \text{ exists} \}.$$  

On $D(A)$ define

$$A f = \lim_{t \to 0} \frac{P_t f - f}{t}.$$  

Note that apriori $D(A)$ might consist only of the zero vector. We have, however, the following theorem.

**Theorem 1.3.** The set $D(A)$ is dense in $X$ and the operator $A$ define above is a linear closed operator.

**Proof.** Consider

$$U_t f = \frac{1}{t} \int_0^t P_s f ds$$  

which exists as a Riemann integral, since the function $s \to P_s f$ is continous. $U_t$ is a bounded operator since

$$\| U_t f \| \leq \frac{1}{t} \int_0^t \| P_s f \| ds \leq \| f \|$$  

since $P_t$ is a contraction. By the definition of the Riemann integral we also have that

$$\| U_t f - f \| \leq \frac{1}{t} \int_0^t \| P_s f - f \| ds$$  

from which we see that

$$\lim_{t \to 0} \| U_t f - f \| = 0.$$  

In other words, the set

$$\cup_{t > 0} U_t(X)$$
is dense in $X$. For $t > 0$ we also find that
\[ P_\varepsilon U_t f - U_t f = \frac{1}{t} \int_0^t P_{\varepsilon+s} f \, ds - \frac{1}{t} \int_0^t P_s f \, ds = \frac{1}{t} \int_\varepsilon^{t+\varepsilon} P_s f \, ds - \frac{1}{t} \int_0^t P_s f \, ds \]
so that
\[ \frac{P_\varepsilon U_t f - U_t f}{\varepsilon} = \frac{1}{t} \left[ U_\varepsilon U_t f - U_\varepsilon f \right] \] (1)
which converges to $U_t f - f$ as $\varepsilon \to 0$. Hence $U_t f \in D(A)$ and
\[ AU_t f = \frac{1}{t} [U_t f - f] . \]
Likewise, from (1)
\[ U_t \frac{P_\varepsilon f - f}{\varepsilon} = \frac{1}{t} \left[ U_t - I \right] U_\varepsilon f \]
and for $f \in D(A)$ we find that
\[ U_t Af = \frac{1}{t} [U_t - I] f \]
and in particular,
\[ AU_t f = U_t Af . \] (2)
To see that $A$ is closed, let $f_n \in D(A)$ be such that $f_n \to f$ and $A f_n \to v$. We have that
\[ \lim_{n \to \infty} U_t A f_n = U_t v , \]
since $U_t$ is continuous. Further,
\[ \lim_{n \to \infty} AU_t f_n = \lim_{n \to \infty} \frac{1}{t} [U_t f_n - f_n] = \frac{1}{t} [U_t f - f] \]
so that using (2)
\[ U_t v = \frac{1}{t} [U_t f - f] \]
for all $t > 0$. As $t \to 0$ the left side converges to $v$ and hence the right side converges too. This implies that $f \in D(A)$ and $A f = v$. $\square$

We call $A$ the **infinitesimal generator** of $P_t$.

Since $P_t$ is a contraction, one can define the integral
\[ R_\lambda(A)f := \int_0^\infty e^{-\lambda t} P_t f \, dt \]
for all $f \in X$ and $\lambda \in \mathbb{C}$ with $\text{Re}\lambda > 0$ as a Riemann integral.

**Theorem 1.4.** The operator $R_\lambda(A)$ maps $X$ into $D(A)$ and obeys the bound
\[ \|R_\lambda(A)\| \leq \frac{1}{\text{Re}\lambda} . \] (3)
Moreover, for all $f \in X$
\[ (\lambda I - A)R_\lambda(A)f = f \] (4)
and for all $f \in D(A)$
\[ R_\lambda(A)(\lambda I - A)f = f . \] (5)
Thus, the resolvent set of $A$ contains the right half plane and $R_\lambda(A) = (A - \lambda I)^{-1}$. 
Proof. For Reλ > 0 we find
\[ \|R_λ(A)\| \leq \int_0^∞ e^{-\Reλt}\|Pf\|dt \leq \frac{1}{\Reλ}\|f\|. \]

Again we compute
\[ [P_ε-I]R_λ(A)f = e^{ελ} \int_ε^∞ e^{-λt}Pf dt - \int_0^∞ e^{-λt}Pf dt = (e^{ελ} - 1) \int_ε^∞ e^{-λt}Pf dt - \int_0^ε e^{-λt}Pf dt \]
so that
\[ \frac{P_ε-I}{ε}R_λ(A)f = \frac{(e^{ελ} - 1) - 1}{ε} \int_ε^∞ e^{-λt}Pf dt - \frac{1}{ε} \int_0^ε e^{-λt}Pf dt . \]

As ε → 0 we see that the right side converges and hence so does the left side and
\[ AR_λ(A)f = λR_λ(A)f - f \]
which proves (4). To see (5) we assume that f ∈ D(AR) and write
\[ R_λ(A)[P_ε-I]f = e^{ελ} \int_ε^∞ e^{-λt}Pf dt - \int_0^∞ e^{-λt}Pf dt = (e^{ελ} - 1) \int_ε^∞ e^{-λt}Pf dt - \int_0^ε e^{-λt}Pf dt \]
so that upon dividing by ε and taking the limit as ε → 0 we get that
\[ R_λ(A)Af = λR_λ(A)f - f \]
which proves (5). The last statement follows from (3). □

Remark 1.5. Note that we defined the resolvent to be \((λI-A)^{-1}\) which differs from our usual definition by a minus sign.

Lemma 1.6. Let A be a closed densely defined operator and assume that \((0,∞) ⊂ ρ(A)\) and that
\[ \| (λI-A)^{-1} \| ≤ \frac{1}{λ}, \ λ > 0 . \]

Then
\[ λ(λI-A)^{-1}f → f \]
as λ → ∞.

Proof. Let f ∈ D(A). Then
\[ λ(λI-A)^{-1}f - f = (λI-A)^{-1}[λI-λI+A]f = (λI-A)^{-1}Af \]
and therefore
\[ \|λ(λI-A)^{-1}f - f\| ≤ \frac{1}{λ}\|Af\| , \]
which tends to 0 as λ → ∞. If f ∈ X for any ε > 0 we can find g ∈ D(A) so that \|f - g\| < ε. Now
\[ λ(λI-A)^{-1}f - f = λ(λI-A)^{-1}(f - g) - (f - g) + λ(λI-A)^{-1}g - g . \]
The term
\[ λ(λI-A)^{-1}(f - g) - (f - g) \]
can be estimated
\[ \|λ(λI-A)^{-1}(f - g) - (f - g)\| ≤ \|λ(λI-A)^{-1}(f - g)\| + \|f - g\| ≤ 2\|f - g\| = 2ε \]
and the second term tends to zero as λ → ∞ which proves the lemma. □
Lemma 1.7. With the same assumptions as in the previous lemma the operator
\[
\lambda(\lambda I - A)^{-1}A
\]
is bounded and for any \( f \in D(A) \)
\[
\|\lambda(\lambda I - A)^{-1}Af - Af\| \to 0
\]
as \( \lambda \to \infty \).

Proof. \( \lambda(\lambda I - A)^{-1}A = \lambda(\lambda I - A)^{-1}(A - \lambda I) + \lambda^2(\lambda I - A)^{-1} \)
\[
= \lambda^2(\lambda I - A)^{-1} - \lambda
\]
and therefore for any \( f \in D(A) \)
\[
\|\lambda(\lambda I - A)^{-1}Af\| = \|\lambda^2(\lambda I - A)^{-1}f - \lambda f\| \leq 2\lambda\|f\|.
\]
Since \( D(A) \) is dense, this proves that \( \lambda(\lambda I - A)^{-1}A \) is bounded. The other statement follows from the previous lemma. \( \square \)

The idea is now to replace the operator \( A \) by the operator
\[
A_\lambda := \lambda(\lambda I - A)^{-1}A
\]
which is bounded. The semigroup
\[
e^{A_\lambda t}
\]
is now simply defined by the power series, which is norm convergent.

Lemma 1.8. The operator
\[
e^{A_\lambda t} := \sum_{k=0}^{\infty} \frac{(A_\lambda t)^k}{k!}
\]
is norm convergent and is a contraction semi group.

Proof. That it is norm convergent and a semi group is standard and the proof is left to the reader. Now,
\[
\|e^{A_\lambda t}\| = \|e^{t(\lambda^2(\lambda I - A)^{-1} - \lambda)}\| \leq e^{-\lambda t} \sum_{k=0}^{\infty} \frac{t^k}{k!} \|\lambda^2(\lambda I - A)^{-1}\| \leq e^{-\lambda t} \sum_{k=0}^{\infty} \frac{t^k \lambda^k}{k!} = 1
\]
since
\[
\|\lambda^2(\lambda I - A)^{-1}\| \leq \lambda.
\]
\( \square \)

Theorem 1.9. A closed operator \( A \) is the generator of a contraction semi group if and only if
\[
(0, \infty) \subset \rho(A)
\]
and
\[
\|R_\lambda(A)\| \leq \frac{1}{\lambda}, \ \lambda > 0.
\]
Proof. We have to show that

\[ \lim_{\lambda \to \infty} e^{A\lambda t} \]

exists and defines a contraction semi group with infinitesimal generator \( A \). Fix \( \lambda > 0 \) and \( \mu > 0 \) and write

\[ e^{A\lambda t} - e^{A\mu t} = e^{A\mu t + (A\lambda - A\mu) s} \mid _0^t = \int_0^t \frac{d}{ds} e^{A\mu t + (A\lambda - A\mu) s} ds \]

which equals

\[ \int_0^t e^{A\mu (t-s)} (A\lambda - A\mu) e^{A\lambda s} ds = \int_0^t e^{A\mu (t-s)} e^{A\lambda s} (A\lambda - A\mu) ds \]

where we have used that \( A\lambda A\mu = A\mu A\lambda \) and that

\[ e^{A\mu t + (A\lambda - A\mu) s} = e^{A\mu t} e^{(A\lambda - A\mu) s} . \]

Now for \( f \in X \)

\[ \| [e^{A\lambda t} - e^{A\mu t}] f \| \leq \int_0^t \| e^{A\mu (t-s)} e^{A\lambda s} (A\lambda - A\mu) f \| ds \]

so that

\[ \| [e^{A\lambda t} - e^{A\mu t}] f \| \leq t \| (A\lambda - A\mu) f \| . \] (6)

If \( f \in D(A) \) then

\[ \| A\lambda f - Af \| \to 0 \]

as \( \lambda \to \infty \) and hence \( e^{A\lambda t} f \) is a Cauchy sequence and hence converges. Since \( D(A) \) is dense, by standard arguments, \( e^{A\lambda t} f \) converges to \( P_t f \) for all \( f \in X \) and the linear operator \( P_t \) is a contraction. We have to show that it is a semi group.

\[ P_{t+s} f = \lim_{\lambda \to \infty} e^{A\lambda (t+s)} f = \lim_{\lambda \to \infty} e^{A\lambda t} e^{A\lambda s} f \]

\[ = \lim_{\lambda \to \infty} e^{A\lambda t} [e^{A\lambda s} - P_s] f + \lim_{\lambda \to \infty} e^{A\lambda t} P_s f \]

Now note that

\[ \| e^{A\lambda t} [e^{A\lambda s} - P_s] f \| \leq \| [e^{A\lambda s} - P_s] f \| \]

which tends to zero as \( \lambda \to \infty \) and

\[ \lim_{\lambda \to \infty} e^{A\lambda t} P_s f = P_t P_s f . \]

Since

\[ \| [P_t - I] f \| \leq 2 \| f \| \]

it suffices to show that

\[ \lim_{t \to 0} \| [P_t - I] f \| = 0 \]

for a dense set of vectors \( f \). Pick \( f \in D(A) \). Then by (6) we have that

\[ \| [P_t - e^{A\lambda t}] f \| \leq t \| (A - A\mu) f \| \]

and hence

\[ \| [P_t - I] f \| \leq \| [P_t - e^{A\mu t}] f \| + \| [e^{A\mu t} - I] f \| \leq t \| (A - A\mu) f \| + \| [e^{A\mu t} - I] f \| \to 0 \]

(7)
as $t \to 0$. Thus, we have shown that $P_t$ is a contraction semi group and therefore it has a generator $B$. We have shown that necessarily $\rho(B)$ contains the complex numbers with positive real part and, moreover,

$$\| (\lambda I - B)^{-1} \| \leq \frac{1}{\text{Re}\lambda}.$$ 

For $f \in X$ we find

$$e^{A\lambda t} f - f = \int_0^t e^{A\lambda s} A f ds$$

and find that for $f \in D(A)$

$$P_t f - f = \int_0^t P_s A f ds.$$ 

From this we find that for $f \in D(A)$

$$\lim_{t \to 0} \frac{[P_t - I] f}{t} = A f$$

and hence $A \subset B$. But for arbitrary $g \in X$

$$(\lambda I - B)(\lambda I - A)^{-1} g = (\lambda I - A)(\lambda I - A)^{-1} g = g$$

and hence

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

so

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

and therefore $D(B) = D(A)$. 

\qed