Functional Analysis Notes

泛函分析笔记

何憾[1]

February 28, 2021

[1]访问我的个人主页, 可查看并下载其他笔记.

Contents

Chapter 1

Spectral Theory

1.1 Spectrum of Closed Operators

Let T be a linear operator whose domain $D(T)$ and range $R(T)$ both lie in the same complex linear topological space X. We consider the linear operator

$$T_\lambda = \lambda I - T$$

where λ is a complex number and I the identity operator. The distribution of the values of λ for which T_λ has an inverse and the properties of the inverse when it exists, are called the *spectral theory* for the operator T. We shall thus discuss the general theory of the inverse of T_λ.

Definition 1.1. If λ is such that the range $R(T_\lambda)$ is dense in X and T_λ has a continuous inverse $T_\lambda^{-1} = (\lambda I - T)^{-1}$, we say that λ is in the *resolvent set* $\varrho(T)$ of T, and we denote this inverse $(\lambda I - T)^{-1}$ by $R(\lambda; T)$ and call it the resolvent (at λ) of T.

Definition 1.2. All complex numbers λ not in $\varrho(T)$ form a set $\sigma(T)$ called the *spectru*m of T. The spectrum $\sigma(T)$ is decomposed into disjoint sets $\sigma_p(T), \sigma_c(T)$ and $\sigma_r(T)$ with the following properties:

- $\sigma_p(T)$ is the totality of complex numbers λ for which T_λ does not have an inverse ($N(\lambda I - T) \neq \{0\}$); $\sigma_p(T)$ is called the *point spectrum* of T.
- $\sigma_c(T)$ is the totality of complex numbers λ for which T_λ has a discontinuous inverse with domain dense in X ($\overline{R(\lambda I - T)} = X$); $\sigma_c(T)$ is called the *continuous spectrum* of T.
- $\sigma_r(T)$ is the totality of complex numbers λ for which T_λ has an inverse whose domain is not dense in X ($\overline{R(\lambda I - T)} \neq X$); $\sigma_r(T)$ is called the *residual spectrum* of T.

Remark 1.1. Why we need the scalar field be $\mathbb{C}$? The answer can be found in Theorem 1.3 and Theorem 1.8: The spectrum is not empty when X is a complex Banach space and $T \in \mathcal{B}(X)$, and $\lambda \to R(\lambda, T)$ is a operator-valued holomorphic function on the open subset $\varrho(T)$ of $\mathbb{C}$. However, if X is a real Banach space, all the definitions about spectrum can be done similarly.

From these definitions and the linearity of T we have : A necessary and sufficient condition for $\lambda \in \sigma_p(T)$ is that the equation

$$Tx = \lambda x$$

has a nonzero solution $x \in D(T)$. In this case λ is called an *eigenvalue* of T, and x the corresponding *eigenvector*. The null space $N(\lambda I - T)$ is called the *eigenspace* of T corresponding to the eigenvalue λ. It consists of the vector 0 and the totality of eigenvectors corresponding to λ. The dimension of the eigenspace corresponding to λ is called the *multiplicity* of the eigenvalue λ.

In practice, when discussing the spectrum, we always work on a closed (or closable) linear operator T defined on a complex Banach space X.

Lemma 1.1. *Let X be a complex Banach space, and T a closed linear operator with its domain $D(T)$ and range $R(T)$ both in X. Then, the following statements are equivalent.*

(a) $\lambda \in \varrho(T)$, *i.e.,* T_λ^{-1} *exists,* T_λ^{-1} *is densely defined and continuous.*

(b) T_λ *is a bijection of* $D(T)$ *onto* X.

(c) T_λ^{-1} *exists and* $T_\lambda^{-1} \in B(X)$.

Proof. Recall that the inverse of a closed linear operator is also closed. Then part (c) follows from part (b) by using the closed graph theorem. Part (c) implies part (a) trivially. It remains to show that part (a) implies part (b). By assumption, T_λ^{-1} is a densely defined continuous operator, and is clsoed. From Remark **??**, $D(T_\lambda^{-1})$ is closed in X. Since $D(T_\lambda^{-1})$ is dense in X, thus $D(T_\lambda^{-1}) = X$ as desired. □

Lemma 1.2. *Let* X *be a complex Banach space, and* $(T, D(T))$ *a closable linear operator in* X. *Then we have*

$$\sigma(\overline{T}) = \sigma(T) \quad \text{and} \quad \varrho(\overline{T}) = \varrho(T)\,.$$

Proof. It suffices to show that $\varrho(\overline{T}) \subset \varrho(T)$ and $\varrho(T) \subset \varrho(\overline{T})$. To prove the first one, assume $\lambda \in \varrho(\overline{T})$. Then clearly $\lambda I - T$ is injective and $(\lambda I - T)^{-1}$ is continuous. Moreover,

$$X = R(\lambda I - \overline{T}) \subset \overline{R(\lambda I - T)}$$

thus $(\lambda I - T)^{-1}$ is densely defined. So $\lambda \in \varrho(T)$.

On the other hand, assume $\lambda \in \varrho(T)$. Then clearly $R(\lambda I - \overline{T})$ is dense in X. Since $T_\lambda^{-1} = (\lambda I - T)^{-1}$ is continuos, there exists a constant $c_\lambda > 0$ so that

$$\|(\lambda I - T)x\| \geq c_\lambda \|x\| \quad \text{whenever} \quad x \in D(T)\,.$$

Then it's easy to see that

$$\left\|(\lambda I - \bar{T})x\right\| \geq c_\lambda \|x\| \quad \text{whenever} \quad x \in D(\bar{T})\,.$$

Thus $(\lambda I - \bar{T})$ is injective and $(\lambda I - \bar{T})^{-1}$ is continuous. So $\lambda \in \varrho(\bar{T})$ as claimed. □

As a consequence of Lemma 1.1, the partition of the spectrum can be restated as following:

$$\sigma_p(T) = \{\lambda \in \mathbb{C} : N(T_\lambda) \neq \{0\}\}\,;$$
$$\sigma_c(T) = \{\lambda \in \mathbb{C} : N(T_\lambda) = \{0\}\ ,\ \overline{R(T_\lambda)} = X\ ,\ R(T_\lambda) \neq X\}\,;$$
$$\sigma_r(T) = \{\lambda \in \mathbb{C} : N(T_\lambda) = \{0\}\ ,\ \overline{R(T_\lambda)} \neq X\}\,.$$

Henceforth, without specific statement, we always suppose that X is a complex Banach space and T a closed linear operator of $D(T) \subset X$ into X.

The most important cases is that the domain of T is the whole space X. In this case,

$$\varrho(T) = \{\lambda \in \mathbb{C} : (\lambda I - T)^{-1} \in \mathcal{B}(X)\}\,.$$

We note that the resolvent set $\varrho(T)$ consists of all those $\lambda \in \mathbb{C}$ for which the equation

$$\lambda x - Tx = y$$

has a unique solution for each y which furthermore depends continuously on the right hand side y.

Remark 1.2. Let X be a complex Banach space, and T a closed linear operator with its domain $D(T)$ and range $R(T)$ both in X. We call $\lambda \in \mathbb{C}$ an *approximate eigenvalue* for T, if there exists a sequence $x_n \in D(T)$ with $\|x_n\| = 1$ so that $\|Tx_n - \lambda x_n\| \to 0$. We call the set

$$\sigma_{ap}(T) := \{\lambda : \lambda \text{ approximate eigenvalue of } T\}$$

the *approximate point spectrum* of T.

We note that every eigenvalue λ is also an approximate eigenvalue (as we may simply choose $x_n = x$ for an element of $N(\lambda I - T)$ that is normalised to $\|x\| = 1$), so we have

$$\sigma_p(T) \subset \sigma_{ap}(T) \subset \sigma(T)\,.$$

We also remark that λ is an approximate eigenvalue if and only if there exists NO $\delta > 0$ so that

$$\|Tx - \lambda x\| \geq \delta \|x\| \quad \text{for all } x \in X \,, \tag{1.1}$$

as is equivalent to

$$\inf_{x \in X, \|x\|=1} \|Tx - \lambda x\| = 0 \,.$$

Moreover, we get directly that $\sigma_c(T) \subset \sigma_{ap}(T)$ from (1.1).

1.1.1 Basic Properties

Our first main result about the spectrum of closed linear operators is

Theorem 1.3. *Let T be a closed linear operator with domain $D(T)$ and range $R(T)$ both in a complex Banach space X. Then the following assertions hold.*

(a) *For each $\lambda_0 \in \varrho(T)$, $B(\lambda_0; \frac{1}{\|R(\lambda_0;T)\|}) \subset \varrho(T)$ and*

$$R(\lambda;T) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0;T)^{n+1} \quad \text{for all} \quad |\lambda - \lambda_0| < \frac{1}{\|R(\lambda_0;T)\|} \,.$$

Particularly, the resolvent set $\varrho(T)$ is an open set of $\mathbb{C}$;

(b) *$R(\lambda;T)$ is a holomorphic function of λ in $\varrho(T)$, and*

$$\frac{d^n}{d\lambda^n} R(\lambda;T) = (-1)^n n! R(\lambda;T)^{n+1} \qquad \text{for all } n \in \mathbb{N} \,.$$

(c) *Let $\lambda_n \in \varrho(T)$ with $\lim_{n\to\infty} \lambda_n = \lambda_0$. Then $\lambda_0 \in \partial\varrho(T) \subset \sigma(T)$ if and only if $\lim_{n\to\infty} \|R(\lambda_n, T)\| = \infty$.*

Proof. By the theorem of the preceding section, $R(\lambda;T)$ for $\lambda \in \varrho(T)$ is an everywhere defined continuous operator. Let $\lambda_0 \in \varrho(T)$ and consider

$$S(\lambda) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, A)^{n+1} \,.$$

The series is convergent in the operator norm whenever

$$|\lambda_0 - \lambda| \, \|R\,(\lambda_0; T)\| < 1\,,$$

and within this circle of the complex plane, the series defines a holomorphic function of λ. Multiplication by $(\lambda I - T) = (\lambda - \lambda_0)\, I + (\lambda_0 I - T)$ on the left or right gives I so that the series $S(\lambda)$ actually represents the resolvent $R(\lambda; T)$. Thus we have proved that a circular neighbourhood of λ_0 belongs to $\varrho(T)$ and $R(\lambda; T)$ is holomorphic in this neighbourhood. Assertion (b) follows immediately from the series representation for the resolvent.

To show (c) we use (a), which implies $\|R(\mu, T)\| \geq 1/\operatorname{dist}(\mu, \sigma(T))$ for all $\mu \in \varrho(T)$. This already proves one implication. For the converse, assume that $\lambda_0 \in \varrho(T)$. Then the continuous resolvent map remains bounded on the compact set $\{\lambda_n : n \geq 0\}$. This contradicts the assumption that $\lim_{n\to\infty} \|R\,(\lambda_n, T)\| = \infty$; hence $\lambda_0 \in \sigma(T)$. $\square$

Theorem 1.4 (Resolvent Equation)**.** *Let T be a closed linear operator with domain and range both in a complex Banach space X. If λ and μ both belong to $\varrho(T)$, then the resolvent equation holds:*

$$\begin{aligned} R(\lambda; T) - R(\mu; T) &= (\mu - \lambda) R(\lambda; T) R(\mu; T) \\ &= (\mu - \lambda) R(\mu; T) R(\lambda; T)\,. \end{aligned}$$

particularly, $R(\lambda; T)$ and $R(\mu; T)$ commute.

Proof. We have

$$\begin{aligned} R(\lambda; T) &= R(\lambda; T)(\mu I - T) R(\mu; T) \\ &= R(\lambda; T)\{(\mu - \lambda) I + (\lambda I - T)\} R(\mu; T) \\ &= (\mu - \lambda) R(\lambda; T) R(\mu; T) + R(\mu; T)\,. \end{aligned}$$

Then the desired result follows. $\square$

Exercise 1.1. Let $\Omega \subset \mathbb{C}$. Suppose $\{R_\lambda\}_{\lambda\in\Omega}$, is a family of bounded linear operators on X satisfying the resolvent equation

$$R_\lambda = R_\mu + (\mu - \lambda)R_\lambda R_\mu \quad \text{for all } \lambda, \mu \in \Omega\,.$$

Then the following statements hold. (i) $\{R_\lambda\}_{\lambda\in\Omega}$ commutes. (ii) The kernel and the range of R_λ is independent of $\lambda \in \Omega$. (iii) If R_λ is injective for some, hence all, $\lambda \in \Omega$, then let $T := \lambda I - R_\lambda^{-1}$. The defined of T is independent of λ and T is a closed linear operator.

Proposition 1.5. *Let T be a closed linear operator with domain and range both in a complex Banach space X. Let $\lambda_1, \cdots, \lambda_n$ be distinct eigenvalues of T. If $x_1, \cdots, x_n$ are corresponding non-trivial eigenvectors, then $\{x_1, \cdots, x_n\}$ are linear independent. In particular, the (internal) direct sum*

$$N(\lambda_1 I - T) \oplus \cdots \oplus N(\lambda_n I - T)$$

is well-defined.

Proof. We prove by induction. When $n = 1$, the propostion is trivial. Assume the propostion holds for $n-1$. If $\alpha_j \in \mathbb{C}$, $j = 1, \cdots, n$ so that

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0\,.$$

Let T act on the equality, then we get

$$\alpha_1 \lambda_1 x_1 + \cdots + \alpha_n \lambda_n x_n = 0$$

Thus we get

$$\alpha_1(\lambda_n - \lambda_1)x_1 + \cdots + \alpha_{n-1}(\lambda_n - \lambda_{n-1})x_{n-1} = 0\,.$$

Thus $\alpha_1 = \cdots = \alpha_{n-1} = 0$ since $\{x_1, \cdots, x_{n-1}\}$ are linear independent by induction and $\lambda_1, \cdots, \lambda_n$ are distinct. Then clearly $\alpha_n = 0$ and the desired result follows. $\square$

1.1.2 Examples

In the end of this section we discuss some concrete examples.

Example 1.1. If the space X is of finite dimension, then any bounded linear operator T is represented by a matrix (t_{ij}). It is known that the eigenvalues of T are obtained as the roots of the algebraic equation, the so-called characteristic equation of the matrix (t_{ij}),

$$\det(\lambda\delta_{ij} - t_{ij}) = 0 .$$

Moreover, $\sigma(T) = \sigma_p(T)$ is the set of eigenvalues and $|\sigma(T)| \leq n$.

Example 1.2 (Multiplication Operators on $C_0(\Omega)$)**.** We start from a locally compact Hausdorff space Ω and define the Banach space (endowed with the sup-norm) $\mathrm{C}_0(\Omega) = \mathrm{C}_0(\Omega; \mathbb{F})$ of all continuous, $\mathbb{F}$-valued functions on Ω that vanish at infinity. With any continuous function $q : \Omega \to \mathbb{F}$ we associate a linear operator M_q on $\mathrm{C}_0(\Omega)$ defined on its "maximal domain" $D(M_q)$ in $\mathrm{C}_0(\Omega)$. Specifically, let $M_q f \coloneqq q \cdot f$, for all f in the domain

$$D(M_q) \coloneqq \{f \in \mathrm{C}_0(\Omega) : q \cdot f \in \mathrm{C}_0(\Omega)\} .$$

The main feature of these multiplication operators is that most operator-theoretic properties of M_q can be characterized by analogous properties of the function q. In the following proposition we give some examples for this correspondence:

(i) $(M_q, D(M_q))$ is closed and densely defined.

(ii) $M_q \in \mathcal{B}(\mathrm{C}_0(\Omega))$ if and only if the function q is bounded. In that case, one has $\|M_q\| = \|q\| := \sup\{|q(s)| : s \in \Omega\}$.

(iii) The spectrum of M_q is the closed range of q; i.e., $\sigma(M_q) = \overline{\mathrm{im}(q)}$.

Example 1.3 (Multiplication Operators on L^p). Multiplication operators arise in a natural way in various instances. For example, if one applies the Fourier transform to a linear differential operator on $L^2\left(\mathbb{R}^d\right)$, this operator becomes a multiplication operator on $L^2\left(\mathbb{R}^d\right)$.

Let $(\Omega, \mathcal{F}, \mu)$ be a σ-finite measure space. For fixed $1 \leq p < \infty$, we consider the Banach space $L^p(\mu) = L^p(\Omega, \mathcal{F}, \mu; \mathbb{C})$ of all (equivalence classes of) p-integrable complex functions on Ω. For a measurable function q, we associate a linear operator M_q on $L^p(\mu)$ defined on its "maximal domain" $D\left(M_q\right)$ in $L^p(\mu)$. Specifically, let $M_q f \coloneqq q \cdot f$, for all f in the domain

$$D\left(M_q\right) \coloneqq \left\{f \in L^p(\mu) : q \cdot f \in L^p(\mu)\right\} .$$

We define the *essential range* of q by

$$\text{ess.im}(q) := \left\{z \in \mathbb{C} : \text{ for all } r > 0 : \mu(q \in B(z, r)) \neq 0\right\}.$$

In other words $\text{ess.im}(q) = \operatorname{supp}(\mu \circ q^{-1})$. One can easily check that the complement of $\text{ess.im}(q)$, given by

$$\left\{z \in \mathbb{C} : \text{ there exists } r > 0, \mu(q \in B(z, r)) = 0\right\},$$

is open, and since $\mathbb{C}$ is second countable, $\mu(q \in \text{ess.im}(q)^c) = 0$. Hence the essential image $\text{ess.im}(q)$ is always closed. If B is a Borel set in $\mathbb{C}$ disjoint with $\text{ess.im}(q)$, then $\mu(q \in B) = 0$. This fact characterises the essential image: It is the smallest closed subset of $\mathbb{C}$ with this property.

In analogy to the preceding example, we now have the following result.

(i) $\left(M_q, D\left(M_q\right)\right)$ is closed and densely defined.

(ii) $M_q \in \mathcal{B}(\mathrm{C}_0(\Omega))$ if and only if the function q is essentially bounded. In that case, one has $\|M_q\| = \|q\|_\infty := \sup\{\lambda : \lambda \in \text{ess.im}(q)\}$.

(iii) The spectrum of M_q is the essential range of q; i.e., $\sigma\left(M_q\right) = \text{ess.im}(q)$.

Example 1.4. Let $X = L^2(\mathbb{R}, \mathbb{C})$ and let T be defined by

$$(Tx)(t) = tx(t) \text{ for } t \in \mathbb{R}\,;$$

where, $D(T) = \left\{x(t) : x(t) \text{ and } tx(t) \in L^2(\mathbb{R}, \mathbb{C})\right\}$.Then by the example above, $\sigma(T) = \mathbb{R}$. We claim that indeed $\sigma_c(T) = \mathbb{R}$.

To see this, let $\lambda_0 \in \mathbb{R}$. Note that $(\lambda_0 I - T)$ is injective. Thus $(\lambda_0 I - T)^{-1}$ exists. Its' domain $R\left(\lambda_0 I - T\right)$ comprises those $y(t) \in L^2(\mathbb{R}, \mathbb{C})$ which vanish identically in the neighbourhood of $t = \lambda_0$; the neighbourhood may vary with $y(t)$. Hence $R\left(\lambda_0 I - T\right)$ is dense in $L^2(\mathbb{R}, \mathbb{C})$. It is easy to see that the operator $(\lambda_0 I - T)^{-1}$ is not bounded on the totality of such $y(t)$'s.

1.2 Spectrum of Bounded Linear Operators

In this section, we always set X a complex Banach space and T a continuous linear operator from X into X.

Theorem 1.6. *Let X be a complex Banach space and $T \in \mathcal{B}(X)$. Then the following limit exists:*

$$\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \inf_{n\geq 1} \|T^n\|^{\frac{1}{n}} =: r_\sigma(T)\,.$$

It is called the spectral radius of T. If $|\lambda| > r_\sigma(T)$ (for example $\lambda > \|T\|$), then the resolvent $R(\lambda; T)$ exists and is given by the series

$$R(\lambda; T) = \sum_{n=1}^{\infty} \frac{1}{\lambda^n} T^{n-1}\,, \tag{1.2}$$

which converges in the norm of operators.

Proof. Set $r = \inf_{n\geq 1} \|T^n\|^{\frac{1}{n}} \geq 0$. It suffices to show that $\limsup_n \|T^n\|^{\frac{1}{n}} \leq r$. For any $\epsilon > 0$, choose k such that $\left\|T^k\right\|^{\frac{1}{k}} \leq r + \epsilon$. For arbitrary n write $n = pk + q$ where $0 \leq q \leq (k-1)$. Then, by $\|AB\| \leq \|A\|\|B\|$, we obtain

$$\|T^n\|^{\frac{1}{n}} \leq \left\|T^k\right\|^{\frac{p}{n}} \cdot \|T\|^{\frac{q}{n}} \leq (r+\epsilon)^{\frac{kp}{n}} \|T\|^{\frac{q}{n}}\,.$$

Since $(kp)/n \to 1$ and $q/n \to 0$ as $n \to \infty$, we have $\limsup_n \|T^n\|^{\frac{1}{n}} \le r + \epsilon$. Since e was arbitrary, $\limsup_n \|T^n\|^{\frac{1}{n}} \le r$.

The series is convergent in the norm of operators when $|\lambda| > r_\sigma(T)$. For, if $|\lambda| \ge r_\sigma(T) + \epsilon$, then

$$\left\| \frac{1}{\lambda^{n+1}} T^n \right\| \le \frac{(r_\sigma(T) + \epsilon/2)^n}{(r_\sigma(T) + \epsilon)^{n+1}} .$$

Thus the series in (1.2) converges. Multiplication by $(\lambda I - T)$ on the left or right of this series gives I so that the series actually represents the resolvent $R(\lambda; T)$. $\square$

Corollary 1.7. *Let X be a complex Banach space and $T \in \mathcal{B}(X)$. Then resolvent set $\varrho(T)$ is not empty.*

One of the most important aspects of the following theorem is that every bounded operator has non-empty spectrum. Here we crucially use that the vector space is over $\mathbb{C}$. The claim is not true if we were to only consider the real spectrum as you already know from Linear Algebra.

Theorem 1.8 (Spectral Radius)**.** *Let X be a complex Banach space and $T \in \mathcal{B}(X)$. Then the spectrum $\sigma(T)$ is non-empty and compact subset of $\mathbb{C}$. In fact,*

$$r_\sigma(T) = \sup_{\lambda \in \sigma(T)} |\lambda| ,$$

and this is why we call $r_\sigma(T)$ the spectral radius of T.

Proof. Step 1. We show that spectrum $\sigma(T)$ is non-empty.

If $\sigma(T) = \emptyset$, then $\varrho(T) = \mathbb{C}$ and $\lambda \mapsto R(\lambda, T)$ is holomorphic on the plane $\mathbb{C}$. Since for $\lambda > \|T\|$, $R(\lambda, T) = \sum_{n=1}^{\infty} \lambda^{-n} T^{n-1}$, we have

$$\|R(\lambda, T)\| \le \frac{1}{\lambda - \|T\|} .$$

Thus $\lambda \mapsto R(\lambda, T)$ is a bounded holomorphic vector-valued function on $\mathbb{C}$. It follows from Theorem **??** that $R(\lambda, T)$ is a constant, which is a contradiction!

Step 2. We show that $r_\sigma(T) = \sup_{\lambda\in\sigma(T)} |\lambda|$. Combine this and Theorem 1.3, $\sigma(T)$ is closed and bounded, hence compact.

By Theorem 1.6, we know that $r_\sigma(T) \geq \sup_{\lambda\in\sigma(T)} |\lambda|$. Hence we have only to show that $r_\sigma(T) \leq \sup_{\lambda\in\sigma(T)} |\lambda|$.

By Theorem 1.3, $R(\lambda;T)$ is holomorphic in λ when $|\lambda| > \sup_{\lambda\in\sigma(T)} |\lambda|$. Thus it admits a uniquely determined Laurent expansion in positive and non-positive powers of λ convergent in the operator norm for $|\lambda| > \sup_{\lambda\in\sigma(T)} |\lambda|$. By Theorem 1.6, this Laurent series must coincide with $\sum_{n=1}^\infty \lambda^{-n}T^{n-1}$. Hence $\lim_{n\to\infty} \|\lambda^{-n}T^n\| = 0$ if $|\lambda| > \sup_{\lambda\in\sigma(T)} |\lambda|$, and so for any $\epsilon > 0$,

$$\|T^n\| \leq \left(\epsilon + \sup_{\lambda\in\sigma(T)} |\lambda|\right)^n \qquad \text{for large } n\,.$$

This proves that

$$r_\sigma(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq \sup_{\lambda\in\sigma(T)} |\lambda|\,,$$

as desired. □

Remark 1.3. Since $\sigma(T)$ is a compact subset in $\mathbb{C}$, it's not hard to see that indeed

$$r_\sigma(T) = \sup_{\lambda\in\sigma(T)} |\lambda| = \max_{\lambda\in\sigma(T)} |\lambda|\,.$$

Corollary 1.9. *The series $\sum_{n=1}^\infty \lambda^{-n}T^{n-1}$ diverges if $|\lambda| < r_\sigma(T)$.*

Proof. Let r be the smallest non-negative real number such that the series

$$\sum_{n=1}^\infty \lambda^{-n}T^{n-1}$$

converges in the operator norm for $|\lambda| > r$. The existence of such an r is proved as for ordinary power series in λ^{-1}. Then, for $|\lambda| > r$, we have $\lim_{n\to\infty} \|\lambda^{-n}T^n\| = 0$ and so, as in the proof of $r_\sigma(T) \leq \sup_{\lambda\in\sigma(T)} |\lambda|$, we must have $\lim_{n\to\infty} \|T^n\|^{1/n} \leq r$. This proves that $r_\sigma(T) \leq r$. In fact, $r_\sigma(T) = r$. □

We furthermore record the following useful lemma. This lemma has been shown in Exercise **??**.

Lemma 1.10. *Let X be a vector space, Let $S, T \in \mathcal{L}(X)$. Suppose that $ST = TS$. Then ST is bijective if and only if S and T are bijective.*

Theorem 1.11. *Let X be a complex Banach space, $T \in \mathcal{B}(X)$ and let p be a complex polynomial. Then*

$$\sigma(p(T)) = p(\sigma(T)) := \{p(\lambda) : \lambda \in \sigma(T)\}$$

Here we set $p(T) := \sum_{j=0}^{n} a_j T^j$ if the polynomial p is given by $p(z) = \sum_{j=0}^{n} a_j z^j$, with the usual convention that $T^0 = I$.

Proof. We first remark that if p is constant, say $p = c \in \mathbb{C}$, then the spectrum of $p(T) = cI$ is simply $\{c\}$. while the fact that $\sigma(T)$ is non-empty implies that also $p(\sigma(T)) = \{c\}$. So suppose that p has degree $n \geq 1$, let $\mu \in \mathbb{C}$ be any given number. As we are working in $\mathbb{C}$ we can factorise $p(\cdot) - \mu$ and write it as $p(z) - \mu = \alpha\,(z - \beta_1(\mu)) \dots (z - \beta_n(\mu))$ for some $\alpha \neq 0$ and equally factorise

$$p(T) - \mu I = \alpha\,(T - \beta_1(\mu)I) \dots (T - \beta_n(\mu)I) \tag{1.3}$$

where we note that all operators on the right hand side commute which will allow us to apply Lemma 1.10. Thus, $\mu \in \varrho(p(T)) \Leftrightarrow \beta_j(\mu) \in \varrho(T)$ for all j. In other words,

$$\mu \in \sigma(p(T)) \Leftrightarrow \exists\ j \text{ so that } \beta_j(\mu) \in \sigma(T)\,.$$

We now note that $\mu \in p(\sigma(T))$ if and only if the equation $p(z) - \mu = 0$ has a root in $\sigma(T)$, in other words,

$$\mu \in p(\sigma(T)) \Leftrightarrow\ \ j \text{ so that } \beta_j(\mu) \in \sigma(T)\,.$$

Then the desired result follows. ☐

This theorem can in particular be applied if a given operator can be written as a polynomial of a simpler operator. Indeed we will generalize this result in Theorem 2.14.

As a final result of this section, we prove that there is the following close connection between the spectrum of an operator and the spectrum of its dual operator.

Theorem 1.12. *Let X be a complex Banach space, let $T \in \mathcal{B}(X)$ and let $T^* \in \mathcal{B}(X^*)$ be the corresponding dual operator of T. Then*

(a) $\sigma(T) = \sigma(T^*)$.

(b) The point, residual, and continuous spectra of T and T^* are related by

$$\begin{array}{ll} \sigma_p(T^*) \subset \sigma_p(T) \cup \sigma_r(T), & \sigma_p(T) \subset \sigma_p(T^*) \cup \sigma_r(T^*)\ ; \\ \sigma_r(T^*) \subset \sigma_p(T) \cup \sigma_c(T), & \sigma_r(T) \subset \sigma_p(T^*)\ ; \\ \sigma_c(T^*) \subset \sigma_c(T), & \sigma_c(T) \subset \sigma_r(T^*) \cup \sigma_c(T^*)\ . \end{array}$$

(c) If X is reflexive, then $\sigma_c(A^*) = \sigma_c(A)$ and

$$\begin{array}{ll} \sigma_p(A^*) \subset \sigma_p(A) \cup \sigma_r(A), & \sigma_p(A) \subset \sigma_p(A^*) \cup \sigma_r(A^*)\ ; \\ \sigma_r(A^*) \subset \sigma_p(A), & \sigma_r(A) \subset \sigma_p(A^*)\ . \end{array}$$

(d) $\sigma(T) = \sigma_{ap}(T) \cup \sigma_p(T^*)$.

Proof. To prove part (a), notice that for any $\lambda \in \mathbb{C}$, $(\lambda I_X - T)^* = \lambda I_{X^*} - T^*$, and then (a) follows from Theorem **??**.

To prove part (b), note tice that by Theorem **??**, we have

$$\overline{R(\lambda I_X - T)} = {}^{\perp}N(\lambda I_{X^*} - T^*)\,,\ \overline{R(\lambda I_{X^*} - T^*)}^{\sigma(X^*,X)} = N(\lambda I_X - T)^{\perp}\,.$$

Assume first that $\lambda \in \sigma_p(T^*)$. Then $N(\lambda I_{X^*} - T^*) \neq \{0\}$, so $\lambda I_X - T$ does not have a dense image, and hence $\lambda \in \sigma_p(T) \cup \sigma_r(T)$. Next assume $\lambda \in \sigma_r(T^*)$. Then $N(\lambda I_{X^*} - T^*) = \{0\}$, hence $\lambda I_X - T$ has a dense image, and hence $\lambda \in \sigma_p(T) \cup \sigma_c(T)$. Third, assume $\lambda \in \sigma_c(T^*)$. Then

$\lambda I_{X^*} - T^*$ is injective and has a dense image and therefore also has a weak-star dense image. Thus $\lambda I_X - T$ is injective and has a dense image, so $\lambda \in \sigma_c(T)$. It follows from these three inclusions that $\sigma_p(T)$ is disjoint from $\sigma_c(T^*)$, that $\sigma_c(T)$ is disjoint from $\sigma_p(T^*)$, and that $\sigma_r(T)$ is disjoint from $\sigma_r(T^*) \cup \sigma_c(T^*)$. This proves part (b).

To prove part (c) observe that in the reflexive case a linear subspace of X^* is weak-star dense if and only if it is dense (Theorem **??**). Hence $\sigma_c(T) = \sigma_c(T^*)$ whenever X is reflexive. With this understood, the remaining assertions of part (c) follow directly from part (b).

To prove part (d), observe that $\sigma(T)\backslash\sigma_{ap}(T) \subset \sigma_r(T)$. In fact, by (1.1), if $\lambda \in \sigma(T)\backslash\sigma_{ap}(T)$, then there exists $\delta > 0$ with

$$\|(\lambda I - T)x\| \geq \delta\|x\| \quad \text{for all } x \in X\,.$$

Then it follows from Exercise **??** that $\lambda \in \sigma_r(T)$, and the desired result follows from (b). □

We now give some examples.

Example 1.5 (Spectrum of the Shift Operators)**.** Let X be the Hilbert space $\ell^2 = \ell^2(\mathbb{N}, \mathbb{C})$, and define the operators $A, B : \ell^2 \to \ell^2$ by

$$Ax := (x_2, x_3, x_4, \ldots)\,, \quad Bx := (0, x_1, x_2, x_3, \ldots)$$

for $x = (x_i)_{i\in\mathbb{N}} \in \ell^2$. Then

$$\sigma(A) = \sigma(B) = D$$

is the closed unit disc in $\mathbb{C}$ and

$$\begin{array}{lll} \sigma_p(A) = \operatorname{int}(D), & \sigma_r(A) = \emptyset, & \sigma_c(A) = S^1\,. \\ \sigma_p(B) = \emptyset, & \sigma_r(B) = \operatorname{int}(D), & \sigma_c(B) = S^1\,. \end{array}$$

Example 1.6. Let $X = \ell^2(\mathbb{N}, \mathbb{C})$ and let $(\lambda_i)_{i\in\mathbb{N}}$ be a bounded sequence of complex numbers. Define the bounded linear operator $A : \ell^2 \to \ell^2$ by

$$Ax := (\lambda_i x_i)_{i\in\mathbb{N}} \qquad \text{for } x = (x_i)_{i\in\mathbb{N}} \in \ell^2 .$$

Then

$$\sigma(A) = \overline{\{\lambda_i \mid i \in \mathbb{N}\}}, \quad \sigma_p(A) = \{\lambda_i \mid i \in \mathbb{N}\}, \quad \sigma_r(A) = \emptyset .$$

Thus *every nonempty compact subset of* $\mathbb{C}$ *is the spectrum of a bounded linear operator on an infinite-dimensional Hilbert space.*

Example 1.7. Let X be a complex Banach space and $P \in \mathcal{B}(X)$ is a projection, i.e., $P^2 = P$. Suppose P is not trivial, that is $P \neq 0$ and $P \neq I$. Then

$$\sigma(P) = \sigma_p(P) = \{0, 1\} .$$

Clearly $\{0, 1\} \subset \sigma_p(P)$. It suffices to show that for $\lambda \in \mathbb{C}\backslash\{0, 1\}$, $\lambda \in \varrho(P)$. Firstly, we claim that $N(\lambda I - P) = \{0\}$. If not, there exists $x \neq 0$ so that $Px = \lambda x \neq 0$. Hence $P^2 x = Px = \lambda Px$, and we deduce that $\lambda x = 1$, which is a contradiction. Secondly, we show that $R(\lambda I - P) = X$, then the desired result follows from the Banach inverse operator theorem. Take any $y \in X$, if $\lambda x - Px = y$, then $\lambda Px - Px = Py$. So $Px = (\lambda - 1)^{-1} Py$ and

$$x = \frac{1}{\lambda}(y - Px) = \frac{1}{\lambda}\left(y - \frac{1}{\lambda - 1}Py\right) .$$

Example 1.8. Let A be a self-adjoint operator in a Hilbert space H (See Example **??**). Then the resolvent set $\varrho(A)$ of A comprises all the complex numbers λ with $\mathrm{Im}(\lambda) \neq 0$, and the resolvent $R(\lambda; A)$ is a bounded linear operator with the estimate

$$\|R(\lambda; A)\| \leq \frac{1}{|\,\mathrm{Im}(\lambda)|} . \tag{1.4}$$

We now prove this. If $x \in H$ then $\langle Ax, x\rangle$ is real since $\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$. Therefore, we have

$$\|(\lambda I - A)x\|^2 = \|(\mathrm{Re}(\lambda)I - A)x\|^2 + |\,\mathrm{Im}(\lambda)|^2\|x\|^2\,.$$

As a consequence,

$$\|(\lambda I - A)x\| \geq |\,\mathrm{Im}(\lambda)| \cdot \|x\|, x \in H\,.$$

Hence the inverse $(\lambda I - A)^{-1}$ exists and is continuous if $\mathrm{Im}(\lambda) \neq 0$.

Moreover, the range $R(\lambda I - A)$ is dense in X if $\mathrm{Im}(\lambda) \neq 0$. If otherwise, there would exist a $y \neq 0$ orthogonal to $R(\lambda I - A)$, i.e., $\langle(\lambda I - A)x, y\rangle = 0$ for all $x \in H$ and so $\langle x, (\bar{\lambda} I - A)y\rangle = 0$ for all $x \in H$. We must have $(\bar{\lambda} I - A)y = 0$, that is, $Ay = \bar{\lambda} y$, contrary to the reality of the value $\langle Ay, y\rangle$. Therefore, by Theorem 1.1, we see that, for any complex number λ with $\mathrm{Im}(\lambda) \neq 0$, the resolvent $R(\lambda; A)$ is a bounded linear operator with the estimate (1.4).

1.3 Regular Points and Defect Numbers

In this section we always assume that H is a complex Hilbert space. Let $(T, D(T))$ be a linear operator in H.

Definition 1.3. $\lambda \in \mathbb{C}$ is called a *regular point* for T if there exists a number $c_\lambda > 0$ such that

$$\|(T - \lambda I)x\| \geq c_\lambda \|x\| \quad \text{for all } x \in D(T)\,. \tag{1.5}$$

The *regularity domain* of T, denoted by $\pi(T)$[1], consists of all the regular points of T.

Clearly, $\lambda \in \pi(T)$ if and only if $T - \lambda I$ has a bounded inverse $(T - \lambda I)^{-1}$ defined on $R(T - \lambda I)$. In this case inequality (1.5) holds with $c_\lambda = \left\|(T - \lambda I)^{-1}\right\|^{-1}$. Moreover, one can see that

$$\varrho(T) \subset \pi(T) \subset \varrho(T) \cup \sigma_r(T)\,.$$

[1]There is no unique symbol for the regularity domain of an operator in the literature.

Proposition 1.13. *Show that* $\lambda_0 \in \pi(T), \lambda \in \mathbb{C}$, *and* $|\lambda - \lambda_0| < c_{\lambda_0}$, *where* c_{λ_0} *is a constant satisfying* (1.5) *for* λ_0, *then* $\lambda \in \pi(T)$. *Particularly,* $\pi(T)$ *is an open subset of* $\mathbb{C}$.

Proof. In fact for $x \in D(T)$,

$$\begin{aligned}
&\|Tx - \lambda x\| = \|Tx - \lambda_0 x + \lambda_0 x - \lambda x\| \\
&\geq \|Tx - \lambda_0 x\| - |\lambda - \lambda_0|\|x\| \geq (c_{\lambda_0} - |\lambda - \lambda_0|)\|x\| \,.
\end{aligned}$$

Thus $\lambda \in \pi(T)$ as desired. □

Recall that the dimension of a Hilbert space $\mathcal{H}$, denoted by $\dim \mathcal{H}$, is defined by the cardinality of an orthonormal basis of $\mathcal{H}$.

Definition 1.4. For $\lambda \in \pi(T)$, we call the linear subspace $R(T-\lambda I)^\perp$ of $\mathcal{H}$ the *deficiency subspace* of T at λ and its dimension $d_\lambda(T) := \dim R(T-\lambda I)^\perp$ the *defect number* of T at λ.

Exercise 1.2. Let $(T, D(T))$ be a closed operator in H. Show that (i) $\lambda \in \varrho(T)$ if and only if $\lambda \in \pi(T)$ and $d_\lambda(T) = 0$; (ii) if $\lambda \in \pi(T)$ and $d_\lambda(T) > 0$, then $\lambda \in \sigma_r(T)$.

Deficiency spaces and defect numbers will play a crucial role in the theory of self-adjoint extensions of symmetric operators developed in Chapter **??**.

Proposition 1.14. *Let* $(T, D(T))$ *be a closable linear operator in* $\mathcal{H}$. *Then we have* $\pi(\bar{T}) = \pi(T)$, *and for each* $\lambda \in \pi(T)$, $R(\bar{T} - \lambda I) = \overline{R(T - \lambda I)}$ *and* $d_\lambda(\bar{T}) = d_\lambda(T)$.

Proof. For each $x \in D(\overline{T})$, there exists a sequence (x_n) in $D(T)$ so that $x_n \to x$ and $Tx_n \to \overline{T}x$. Thus $\overline{T}x - \lambda x = \lim_{n\to\infty} Tx_n - \lambda x_n$ belongs to $\overline{R(T - \lambda I)}$. Hence we have $R(\bar{T} - \lambda I) \subset \overline{R(T - \lambda I)}$. Besides,

$$\|\overline{T}x - \lambda x\| = \lim_{n\to\infty} \|\overline{T}x_n - \lambda x_n\| \geq \lim_{n\to\infty} c_\lambda \|x_n\| = c_\lambda \|x\| \,.$$

Thus $\lambda \in \pi(T)$.

Next we show that $\overline{R(T-\lambda I)} \subset R(\bar{T}-\lambda I)$. Take y in $\overline{R(T-\lambda I)}$, then there exists (x_n) in $D(T)$ so that $Tx_n - \lambda x_n$ converges to y. By (1.5) thus (x_n) is a cauchy sequence in H. Let $x_n \to x$ then $Tx_n \to y + \lambda x$. Since T is closable, we have $x \in D(\overline{T})$ and $\overline{T}x = y + \lambda x$. Thus $y \in R(\bar{T} - \lambda I)$.

Observe that

$$\begin{aligned} d_\lambda(\bar{T}) &= \dim R(\bar{T}-\lambda I)^\perp = \dim \overline{R(T-\lambda I)}^\perp \\ &= \dim R(T-\lambda I)^\perp = d_\lambda(T)\,. \end{aligned}$$

We are done. □

The following technical lemma is needed in the proof of the next proposition.

Lemma 1.15. *If $\mathcal{F}$ and $\mathcal{G}$ are closed linear subspaces of a Hilbert space $\mathcal{H}$ such that* $\dim \mathcal{F} < \dim \mathcal{G}$, *then there exists a nonzero vector $y \in \mathcal{G} \cap \mathcal{F}^\perp$*

Proof. In this proof we denote by $|M|$ the cardinality of a set M. First, we suppose that $k = \dim \mathcal{F}$ is finite. We take a $(k+1)$ -dimensional subspace $\mathcal{G}_0$ of $\mathcal{G}$ and define the mapping $\Phi : \mathcal{G}_0 \to \mathcal{F}$ by $\Phi(x) = Px$, where P is the projection of $\mathcal{H}$ onto $\mathcal{F}$. If Φ would be injective, then $k+1 = \dim \mathcal{G}_0 = \dim \Phi(\mathcal{G}_0) \leq \dim \mathcal{F} = k$, which is a contradiction. Hence, there is a nonzero vector $y \in \mathcal{N}(\Phi)$. Clearly, $y \in \mathcal{G} \cap \mathcal{F}^\perp$. Now suppose that $\dim \mathcal{F}$ is infinite. Let $\{f_k : k \in K\}$ and $\{g_l : l \in L\}$ be orthonormal bases of $\mathcal{F}$ and $\mathcal{G}$, respectively. Set $L_k := \{l \in L : \langle f_k, g_l \rangle \neq 0\}$ for $k \in K$ and $L' = \bigcup_{k\in K} L_k$. Since each set L_k is at most countable and $\dim \mathcal{F} = |K|$ is infinite we have $|L'| \leq |K||\mathbb{N}| = |K|$. Since $|K| = \dim \mathcal{F} < \dim \mathcal{G} = |L|$ by assumption, we deduce that $L' \neq L$. Each vector g_l with $l \in L \backslash L'$ is orthogonal to all $f_k, k \in K$ and hence, it belongs to $\mathcal{G} \cap \mathcal{F}^\perp$. □

The next theorem is a classical result of M.A. Krasnosel'skii and M.G. Krein.

Theorem 1.16. *Let $(T, D(T))$ be a closable linear operator in H. Then the defect number $d_\lambda(T)$ is constant on each connected component of the open set $\pi(T)$.*

Proof. By Proposition 1.14, we can assume without loss of generality that T is closed and then $R(T - \lambda I)$ is closed for all $\lambda \in \pi(T)$.

Take $\lambda_0 \in \pi(T)$. We claim that $d_\lambda(T) = d_{\lambda_0}(T)$ for all $\lambda \in \mathbb{C}$ so that $|\lambda - \lambda_0| < c_{\lambda_0}$. (Note that by Exercise 1.13 we have $\lambda \in \pi(T)$.) Then the mapping $\lambda \mapsto d_\lambda(T)$ is locally constant which implies the desired result.

Assume to the contrary that $d_\lambda(T) \neq d_{\lambda_0}(T)$. First suppose that $d_\lambda(T) < d_{\lambda_0}(T)$. By Lemma 1.15, there exists a nonzero vector

$$y \in R(T - \lambda_0 I)^\perp \cap R(T - \lambda I)$$

Then $y = (T - \lambda I)x$ for some nonzero $x \in \mathcal{D}(T)$ and

$$\langle (T - \lambda I)x, (T - \lambda_0 I)\, x\rangle = 0 \tag{1.6}$$

Equation (1.6) is symmetric in λ and λ_0, so it holds also when $d_{\lambda_0}(T) < d_\lambda(T)$.Using (1.6) we derive

$$\begin{aligned} \|(T - \lambda_0 I)\, x\|^2 &= \langle (T - \lambda I)x + (\lambda - \lambda_0)\, x, (T - \lambda_0 I)\, x\rangle \\ &\leq |\lambda - \lambda_0|\, \|x\|\, \|(T - \lambda_0 I)\, x\| \end{aligned}$$

Thus, $\|(T - \lambda_0 I)\, x\| \leq |\lambda - \lambda_0|\, \|x\|$. Since $x \neq 0$ and $|\lambda - \lambda_0| < c_{\lambda_0}$, we obtain

$$|\lambda - \lambda_0|\, \|x\| < c_{\lambda_0} \|x\| \leq \|(T - \lambda_0 I)\, x\| \leq |\lambda - \lambda_0|\, \|x\|$$

by (1.5) which is a contradiction. Thus, we have proved that $d_\lambda(T) = d_{\lambda_0}(T)$. $\square$

Exercise 1.3. The numerical range of a linear operator T in $\mathcal{H}$ is defined by

$$\Theta(T) = \{\langle Tx, x\rangle : x \in \mathcal{D}(T), \|x\| = 1\}$$

A classical result of F. Hausdorff says that $\Theta(T)$ is a convex set. In general, the set $\Theta(T)$ is neither closed nor open for a bounded or closed operator.

Show that if $\lambda \in \mathbb{C}$ is not in the closure of $\Theta(T)$, then $\lambda \in \pi(T)$.

Chapter 2

Banach Algebra and C* Algebra

2.1 Definition and Examples

Recall that we say $(\mathscr{A}, +, \cdot)$ is an (*associative*) *algebra* over the filed $\mathbb{F}$, if the following statements holds

(i) $(\mathscr{A}, +)$ is a vector space over $\mathbb{F}$.

(ii) $(\mathscr{A}, +, \cdot)$ is a ring.

(iii) The scalar multiplication and vector multiplication satisfy that

$$\lambda(x \cdot y) = (\lambda x) \cdot y = x \cdot (\lambda y)\,,$$

for all $x, y \in \mathscr{A}$ and $\lambda \in \mathbb{F}$.

If in addition, $(\mathscr{A}, +, \cdot)$ is a commutative ring, we say $(\mathscr{A}, +, \cdot)$ is a *commutative algebra*. An element $e \in \mathscr{A}$ is a *unit element* (or *identity*) if

$$e \cdot x = x \cdot e = x \quad \text{for all } x \in \mathscr{A}\,.$$

In this case, we say $\mathscr{A}$ is *unital*. It's easy to see that the unit element must be unique in $\mathscr{A}$. In this note, we always assume that a unital algebra is nonzero, i.e., the unital element is not equal to zero.

Let $\mathscr{A}$ be an algebra with a unit e. Then $x \in \mathscr{A}$ is called *invertible* if there exists some $y \in \mathscr{A}$, called the *inverse* of x, so that

$$x \cdot y = y \cdot x = e\,.$$

It's easy to see that if x is invertible, then the inverse of x is unique. If in addition, every nonzero element in $\mathscr{A}$ is invertible, then $\mathscr{A}$ is called a *division algebra*.

Let $\mathscr{A}, \mathscr{B}$ be two unital algebra over $\mathbb{F}$. We say ψ is a (algebra) *homomorphism* of $\mathscr{A}$ into $\mathscr{B}$, if ψ protects the algebraic operations. In other words for all $x, y \in \mathscr{A}$ and $\alpha, \beta \in \mathbb{F}$,

$$\psi(\alpha x + \beta y) = \alpha\psi(x) + \beta\psi(y)\ ;\ \psi(xy) = \psi(x)\psi(y)\,.$$

If in addition ψ is bijective, we call it an *isomorphism*. In this case, we can see that ψ maps the identity in $\mathscr{A}$ into the identity in $\mathscr{B}$ if ψ is nonzero.

Let $\mathscr{A}$ be a algebra over $\mathbb{F}$. $\mathscr{B} \subset \mathscr{A}$ is called a *subalgebra*, if $\mathscr{B}$ equipped with the operations inherited form $\mathscr{A}$ is an algebra. In other words, a subalgebra of an algebra is a subset of elements that is closed under addition, multiplication, and scalar multiplication.

Definition 2.1. Let $\mathscr{A}$ be an algebra over $\mathbb{F}$, $\|\cdot\|$ a norm on $\mathscr{A}$. We say that $(\mathscr{A}, \|\cdot\|)$ is a *normed algebra* over $\mathbb{F}$, if the norm satisfies the multiplicative inequality:

$$\|x \cdot y\| \leq \|x\|\|y\| \tag{2.1}$$

for all $x, y \in \mathscr{A}$. If in addition, $\mathscr{A}$ is complete with respect to the norm, we call $\mathscr{A}$ a *Banach algebra*.

Remark 2.1. If $\mathscr{A}$ is a normed algebra with a unit e, then clearly $\|e\| \geq 1$. We can always assume that

$$\|e\| = 1$$

(without change the topology on $\mathscr{A}$). We will explain this in Example 2.4.

Remark 2.2. The inequality (2.1) makes multiplication a continuous operation in $\mathscr{A}$. This means that if $x_n \to x$ and $y_n \to y$ then $x_n y_n \to xy$.

On the contrary, if $\mathscr{A}$ is an algebra endowed with a norm $\|\cdot\|$ so that the multiplication is continuous, then there exists a constant $C > 0$ so that

$$\|xy\| \leq C\|x\|\|y\|$$

for all $x, y \in \mathscr{A}$. In fact if not, for each $n \in \mathbb{N}$, there exists x_n, y_n so that $\|x_n\| = 1$, $\|y_n\| = 1$ and $\|x_n y_n\| \geq n^2$. Then $\frac{x_n}{n} \to 0$, $\frac{y_n}{n} \to 0$ but $\|\frac{x_n y_n}{n^2}\| \geq 1$, which contradicts to that the multiplication is continuous.

Remark 2.3. Every normed algebra $\mathscr{A}$ can be regarded as an subalgebra of some unital algebra. Indeed, let $\mathscr{A}_1 = \mathscr{A} \times \mathbb{F}$. Define algebraic operations on $\mathscr{A}_1$ by

(i) $(x, \alpha) + (y, \beta) = (x + y, \alpha + \beta)$;

(ii) $\beta(x, \alpha) = (\beta x, \beta\alpha)$;

(iii) $(x, \alpha)(y, \beta) = (xy + \alpha y + \beta x, \alpha\beta)$.

Define

$$\|(x, \alpha)\| = \|x\| + |\alpha| \,.$$

Then $\mathscr{A}_1$ with this norm and the algebraic operations defined in (i), (ii), and (iii) is a normed algebra with identity $(0, 1)$ and $\|(0, 1)\| = 1$. Moreover, $\mathscr{A}$ is a Banach algebra, then so is $\mathscr{A}_1$; also $\mathscr{A}_1$ if commutative if $\mathscr{A}$ is.

The mapping $x \mapsto (x, 0)$ is an *isometric isomorphism* of $\mathscr{A}$ onto a subspace of $\mathscr{A}_1$ (in fact, onto a closed two-sided ideal of $\mathscr{A}_1$) whose codimension is 1. If x is identified with $(x, 0)$ then $\mathscr{A}_1$ is simply $\mathscr{A}$ plus the one-dimensional vector space generated by e. For concrete examples, see Example 2.2 and 2.5.

We now give some examples for Banach algebra.

Example 2.1. Let X be a compact Hausdorff space. Then the Banach space $C(X, \mathbb{F})$ of all complex continuous functions on X, with the supremum norm and multiplication defined in the usual way: $(fg)(x) = f(x)g(x)$, is a unital commutative Banach algebra. The constant function 1 is the unit element.

If X is a finite set with the discrete topology, consisting of, say, d points, then $C(X, \mathbb{F})$ is simply $\mathbb{F}^d$, with coordinatewise multiplication. In particular, when $d = 1$, we obtain the simplest Banach algebra, namely $\mathbb{F}$, with the absolute value as norm.

Example 2.2. If X is a LCH space, then the Banach space $\mathscr{A} = C_0(X; \mathbb{F})$ of all complex continuous functions on X vanishing at infinity, with the supremum norm, is a Banach algebra when the multiplication is defined pointwise as in the preceding example. $\mathscr{A}$ is commutative, but if X is not compact, $\mathscr{A}$ does not have an identity.

One can adjoin a uint element by the abstract procedure outlined in Remark 2.3 or one can do it more concretely by enlarging $C_0(X; \mathbb{F})$ to $C(X_\infty; \mathbb{F})$, where X_∞ is the one-point compactification of X. Indeed, for each $f \in C_0(X; \mathbb{F})$, we can extend f on X_∞ by defined $f(\infty) = 0$ and then the extended function f belongs to $C(X_\infty; \mathbb{F})$. Note that for each $f \in C(X_\infty; \mathbb{F})$, $g_f = f - f(\infty) \in C_0(X; \mathbb{F})$, and then

$$f = g + f(\infty) \cdot 1$$

where 1 is the constant 1 function.

Example 2.3. Let (X, Ω, μ) be a σ-finite measure space and $\mathscr{A} = L^\infty(\mu; \mathbb{F})$. Then $\mathscr{A}$ is an commutative Banach algebra with identity if the operations are defined pointwise.

Example 2.4. Let X be a Banach space over $\mathbb{F}$. Then $\mathcal{B}(X)$, the algebra of all bounded linear operators on X, is a unital Banach algebra, with respect to the usual operator norm. The identity operator I is its unit element. If X has finite dimension, then $\mathcal{B}(X)$ is (isomorphic to) the algebra of all complex $n \times n$-matrices. If $\dim X \geq 2$, then $\mathcal{B}(X)$ is not commutative.

Every closed subalgebra of $\mathcal{B}(X)$ that contains I is also a unital Banach algebra. In fact, that every unital Banach algebra is isomorphic and homeomorphism to one of these:

Let $\mathscr{A}$ be a Banach algebra. Assign to each $x \in \mathscr{A}$ the left-multiplication operator M_x defined by

$$M_x(a) = xa \quad \text{for} \quad a \in \mathscr{A} \,.$$

Clearly $M_x \in \mathcal{B}(\mathscr{A})$, the Banach space of all continuous linear operators on $\mathscr{A}$, with $\|M_x\| \leq \|x\|$.

Let $\tilde{\mathscr{A}}$ be the set of all M_x. It is clear that $x \to M_x$ is linear. The associative law implies that $M_{xy} = M_x M_y$. On the other hand, note that

$$\|x\| = \|xe\| = \|M_x e\| \leq \|M_x\| \, \|e\| \,,$$

we get

$$\|M_x\| \geq \frac{1}{\|e\|} \|x\| \,.$$

Thus the mapping

$$x \mapsto M_x \ ; \ \mathscr{A} \to \tilde{\mathscr{A}}$$

is both a isomorphism and a homeomorphism of $\mathscr{A}$ onto $\tilde{\mathscr{A}} \subset \mathcal{B}(\mathscr{A})$. Then it follows that $\tilde{\mathscr{A}}$ is a closed unital subalgebra of $\mathcal{B}(\mathscr{A})$.

Moreover, the unit element in $\tilde{a}$ is $M_e = I$, and hence $\|M_e\| = \|I\| = 1$. As a consequence, we can always assume that the unit element in a Banach algebra has norm 1 (without change the topology on it).

Example 2.5. $L^1\left(\mathbb{R}^d;\mathbb{F}\right)$, with convolution as multiplication, is a Banach algebra has no unit element. One can adjoin one by the abstract procedure outlined in Remark 2.3 or one can do it more concretely by enlarging $L^1\left(\mathbb{R}^d;\mathbb{F}\right)$ to the algebra of all complex Borel measures μ on $\mathbb{R}^d$ of the form

$$d\mu = f dm + \lambda d\delta\,,$$

where $f \in L^1\left(\mathbb{R}^d;\mathbb{F}\right)$, m is the Lebesgue measure $\mathbb{R}^d$, δ is the Dirac measure on $\mathbb{R}^d$, and $\lambda \in \mathbb{F}$ is a scalar.

Example 2.6. Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the one-dimensional tours. We denote by $A(\mathbb{T})$ the space of complex-valued continuous functions on $\mathbb{T}$ having an absolutely convergent Fourier series. That is,

$$A(\mathbb{T}) := \left\{ f \in C(\mathbb{T};\mathbb{C}) : f(t) = \sum_{k\in\mathbb{Z}} \hat{f}(k)e^{2\pi ikt}\,, \sum_{k\in\mathbb{Z}} |\hat{f}(k)| < \infty \right\}.$$

under the usual addition of multiplication. The mapping $f \mapsto \hat{f} = \{\hat{f}(k)\}_{k\in\mathbb{Z}}$ of $A(\mathbb{T})$ into ℓ^1 is clearly linear and one-to-one. If $\sum |a_k| < \infty$ the series $\sum a_k e^{ikt}$ converges uniformly on $\mathbb{T}$ and, denoting its sum by g, we have $a_k = \hat{g}(k)$. It follows that the mapping above is an algebra isomorphism of $A(\mathbb{T})$ onto ℓ^1. We introduce a norm to $A(\mathbb{T})$ by

$$\|f\|_{A(\mathbb{T})} = \|\hat{f}\|_{\ell^1} = \sum_{-\infty}^{\infty} |\hat{f}(k)|\,.$$

We emphasize that the norm on $A(\mathbb{T})$ is not the supremum norm! With this norm $A(\mathbb{T})$ is a Banach space isometric to ℓ^1. We now claim it is a Banach algebra.

Assume that $f, g \in A(\mathbb{T})$, we are going to show that $fg \in A(\mathbb{T})$ and

$$\|fg\|_{A(\mathbb{T})} \le \|f\|_{A(\mathbb{T})}\|g\|_{A(\mathbb{T})}\,.$$

We have $f(t) = \sum \hat{f}(k)e^{ikt}$, $\quad g(t) = \sum \hat{g}(k)e^{ikt}$ and since both series converge absolutely:

$$f(t)g(t) = \sum_k \sum_m \hat{f}(k)\hat{g}(m)e^{i(k+m)t}$$

Collecting the terms for which $k + m = l$ we obtain

$$f(t)g(t) = \sum_l \sum_k \hat{f}(k)\hat{g}(l-k)e^{ilt}$$

so that $\widehat{fg}(l) = \sum_k \hat{f}(k)\hat{g}(l-k)$; hence

$$\sum_l |\widehat{fg}(l)| \leq \sum_l \sum_k |\hat{f}(k)||\hat{g}(l-k)| = \sum_k |\hat{f}(k)| \sum_l |\hat{g}(l)| \, .$$

Example 2.7. Let $\mathbb{D}$ be the unit open disc in the complex plane. Define the disc algebra $A(\mathbb{D})$ by

$$A(\mathbb{D}) \coloneqq \left\{ f \in C(\overline{\mathbb{D}}) : f \in H(\mathbb{D}) \right\}$$

equipped with the supremum norm, under the usual addition and multiplication. It's easy to see that $A(\mathbb{D})$ is a subalgebra of $C(\overline{\mathbb{D}}; \mathbb{C})$, and one can easily verify that $A(\mathbb{D})$ is closed in $C(\overline{\mathbb{D}}; \mathbb{C})$, thus $A(\mathbb{D})$ is a unital commutative Banach algebra.

2.2 Basic Properties of Spectra

Let $\mathscr{A}$ be a Banach algebra over $\mathbb{F}$ with identity e. Let $G(\mathscr{A})$ be the set of all invertible elements of $\mathscr{A}$. It's easy to verify that $(G(\mathscr{A}), \cdot)$ is a group, where $\cdot$ is the multiplication inherited form $\mathscr{A}$.

The following lemma is easy but very useful, and the proof is omitted.

Lemma 2.1. *Let $\mathscr{A}$ be a Banach algebra over $\mathbb{F}$ with identity e. Then for each $\|x\| < 1$, $e - x$ is invertible, and*

$$(e - x)^{-1} = \sum_{n=0}^{\infty} x^n \, .$$

Proposition 2.2. *Let $\mathscr{A}$ be a Banach algebra over $\mathbb{F}$ with identity e. Then* *(a)* *For fixed* $x \in G(\mathscr{A})$, *if* $\|h\| < \frac{1}{\|x^{-1}\|}$ *then* $x - h \subset G(\mathscr{A})$.

(b) $G(\mathscr{A})$ *is an open subset of* $\mathscr{A}$, *and the mapping* $x \mapsto x^{-1}$ *is a homeomorphism of* $G(\mathscr{A})$ *onto* $G(\mathscr{A})$.

Proof. For each $x \in G(\mathscr{A})$ and $h \in \mathscr{A}$, we have

$$x - h = x(e - x^{-1}h)\,.$$

Then use the last lemma, we can see that $B(x, \frac{1}{\|x^{-1}\|}) \subset G(\mathscr{A})$. Thus $G(\mathscr{A})$ is open. Moreover,

$$\begin{aligned}
&\|(x-h)^{-1} - x^{-1}\| = \|(e - x^{-1}h)^{-1}x^{-1} - x^{-1}\| \\
&\leq \left\| \sum_{n=0}^{\infty} (x^{-1}h)^n - e \right\| \|x^{-1}\| \\
&\leq \sum_{n=1}^{\infty} \|x^{-1}h\|^n \, \|x^{-1}\| = \frac{\|x^{-1}h\|}{1 - \|x^{-1}h\|} \, \|x^{-1}\|\,.
\end{aligned}$$

Thus for each $x \in G(\mathscr{A})$,

$$\lim_{h \to 0} \|(x-h)^{-1} - x^{-1}\| = 0$$

Then clearly $x \mapsto x^{-1}$ is continuous. Since $x \to x^{-1}$ maps $G(\mathscr{A})$ onto $G(\mathscr{A})$ and since it is its own inverse, it is a homeomorphism. □

Let $\mathscr{A}$ be a Banach algebra over $\mathbb{F}$ with identity e. The *spectrum* of $a \in \mathscr{A}$, denoted by $\sigma(a)$, is defined by

$$\sigma(a) = \{\lambda \in \mathbb{F} : \lambda e - a \text{ is not invertible}\}\,.$$

The *resolvent set* of a, denoted by $\varrho(a)$, is defined by

$$\varrho(a) = \{\lambda \in \mathbb{F} : \lambda e - a \text{ is invertible}\} = \mathbb{F} \backslash \sigma(a)\,.$$

The following proposition is similar to Theorem 1.6 and the proofs are basically the same.

Proposition 2.3. *Let $\mathscr{A}$ be a Banach algebra over $\mathbb{F}$ and $x \in \mathscr{A}$. Then the following limit exists:*

$$\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} = \inf_{n\geq 1} \|x^n\|^{\frac{1}{n}} =: r_\sigma(x)\,,$$

called the spectral radius of x. For each $|\lambda| > r_\sigma(x)$ (for example $\lambda > \|x\|$), $\lambda e - x$ is invertible and its converse is given by the series

$$(\lambda e - x)^{-1} = \sum_{n=1}^{\infty} \frac{1}{\lambda^n} x^{n-1}\,.$$

Example 2.8. Let X be a compact Hausdorff space. For each $f \in C(X;\mathbb{F})$, we have

$$\sigma(f) = f(X) = \{f(x) : x \in X\}\ .$$

In fact, if $\lambda = f(x_0)$, then $\lambda - f$ has a zero and cannot be invertible. So $f(X) \subset \sigma(f)$. On the other hand, if $\lambda \notin f(X), \lambda - f$ is a non-vanishing continuous function on X. Hence $\frac{1}{\lambda - f} \in C(X;\mathbb{F})$ and so $\lambda - f$ is invertible. Thus $\lambda \in \varrho(f)$.

Example 2.9. If X is a Banach space over $\mathbb{F}$, and $T \in \mathcal{B}(X)$, then

$$\begin{aligned} \rho(T) &= \{\lambda \in \mathbb{F} : \lambda I - T \text{ is bijective}\}\ ; \\ \sigma(T) &= \{\lambda \in \mathbb{F} :\ \text{either } N(\lambda I - T) \neq \{0\} \text{ or } R(T-\lambda) \neq X\}\,. \end{aligned}$$

In fact, if $\lambda \in \rho(T)$, clearly, $\lambda I - T$ is bijective. On the other hand, if $\lambda I - T$ is bijective, $(\lambda I - T)^{-1} \in \mathcal{B}(X)$ by the inverse mapping theorem.

Example 2.10. If $\mathscr{A} = M_2(\mathbb{R})$ and $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then

$$\sigma(A) = \emptyset$$

In fact, $\lambda I - A$ is not invertible if and only if $0 = \det(\lambda I - A) = \lambda^2 + 1$, which is impossible in $\mathbb{R}$.

There are several reasons for restricting our attention to Banach algebras over the complex field $\mathbb{C}$.

- The first reason is that, the phenomenon of the last example does not occur if $\mathscr{A}$ is a Banach algebra over $\mathbb{C}$.
- Another reason is that certain elementary facts about holomorphic functions play an important role in the foundations of the subject. This will be observed in the next section when we discussing the spectrum, and becomes even more obvious in the Holomorphic functional calculus.
- The last reason - one whose implications are not quite so obvious - is that $\mathbb{C}$ has a natural nontrivial *involution* (see Definition 2.6), namely, *conjugation*, and that many of the deeper properties of certain types of Banach algebras depend on the presence of an involution. For the same reason, the theory of complex Hilbert spaces is richer than that of real ones.

By the same argument of Theorem 1.8, we can get the following result.

Theorem 2.4. *Let $\mathscr{A}$ be a complex unital Banach algebra. Let $x \in \mathscr{A}$. Then the spectrum $\sigma(x)$ is a non-empty compact subset of $\mathbb{C}$. In fact,*

$$\sup_{\lambda \in \sigma(x)} |\lambda| = r_\sigma(x)\,,$$

and this is why we call $r_\sigma(x)$ the spectral radius of x.

Thus form now on, without special announcement, we always suppose that all Banach algebras are over the complex field $\mathbb{C}$.

Remark 2.4. Whether an element of $\mathscr{A}$ is or is not invertible in $\mathscr{A}$ is a purely *algebraic property*. The spectrum and the spectral radius of an $x \in \mathscr{A}$ are thus defined in terms of the *algebraic structure* of $\mathscr{A}$, regardless of any metric (or topological) considerations. On the other hand, $\lim \|x^n\|^{\frac{1}{n}}$

depends obviously on the *norm* of $\mathscr{A}$. This is one of the remarkable features of the spectral radius formula: It asserts the equality of certain quantities which arise in entirely different ways.

Remark 2.5. Our algebra $\mathscr{A}$ may have a subalgebra $\mathscr{B}$ containing the unit element, and it may then very well happen that some $x \in \mathscr{B}$ is not invertible in $\mathscr{B}$ but is invertible in $\mathscr{A}$. The spectrum of x depends on the algebra, and the inclusion $\sigma_{\mathscr{A}}(x) \subset \sigma_{\mathscr{B}}(x)$ holds (the notation is self-explanatory). The two spectra can be different (see Example 2.12). The spectral radius is, however, unaffected by the passage from $\mathscr{A}$ to $\mathscr{B}$, since the spectral radius formula expresses it in terms of metric properties of powers of x, and these are independent of anything that happens outside $\mathscr{A}$. Later we will describe the relation between $\sigma_{\mathscr{A}}(x)$ and $\sigma_{\mathscr{B}}(x)$ in greater detail.

Example 2.11. Let $\mathscr{A} = M_2(\mathbb{C})$ and let $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Then $A^2 = 0$ and $\sigma(A) = \{0\}$; so $r_\sigma(A) = 0$. So it is possible to have $r(A) = 0$ with $A \neq 0$.

The non-emptiness of spectrum leads to an easy characterization of those Banach algebras that are division algebras.

Theorem 2.5 (Gelfand-Mazur)**.** *If $\mathscr{A}$ is a complex unital Banach algebra in which every nonzero element is invertible, then $\mathscr{A}$ is (isometrically isomorphic to) the complex field $\mathbb{C}$.*

Proof. If $x \in \mathscr{A}$ and $\lambda_1 \neq \lambda_2$, then at most one of the elements $\lambda_1 e - x$ and $\lambda_2 e - x$ is 0; hence at least one of them is invertible. Since $\sigma(x)$ is not empty, it follows that $\sigma(x)$ consists of exactly one point, say $\lambda(x)$, for each $x \in \mathscr{A}$. since $\lambda(x)e - x$ is not invertible, it is 0. Hence $x = \lambda(x)e$.

The mapping $x \to \lambda(x)$ is therefore an isomorphism of $\mathscr{A}$ onto $\mathbb{C}$, which is also an isometry, since $|\lambda(x)| = \|\lambda(x)e\| = \|x\|$ for every $x \in \mathscr{A}$. ☐

It is natural to ask whether the spectra of two elements x and y of $\mathscr{A}$

are close together, in some suitably defined sense, if x and y are close to each other. The next proposition gives a very simple answer.

Proposition 2.6 (Perturbation of Spectrum)**.** *Let $\mathscr{A}$ be a complex unital Banach algebra, $x \in \mathscr{A}$, Ω is an open set in $\mathbb{C}$ containing $\sigma(x)$. Then there exists $\delta > 0$ so that for ever $y \in \mathscr{A}$ with $\|y\| < \delta$, we have $\sigma(x+y) \subset \Omega$.*

Proof. Since $\left\|(\lambda e - x)^{-1}\right\|$ is a continuous function of λ in the complement of $\sigma(x)$, and since this norm tends to 0 as $\lambda \to \infty$, there is a number $M < \infty$ such that

$$\left\|(\lambda e - x)^{-1}\right\| < M \text{ for all } \lambda \in \Omega^c\,.$$

If $y \in \mathscr{A}$, $\|y\| < 1/M$, and $\lambda \in \Omega^c$, it follows that

$$\lambda e - (x+y) = (\lambda e - x)\left[e - (\lambda e - x)^{-1} y\right]$$

is invertible in $\mathscr{A}$, since $\left\|(\lambda e - x)^{-1} y\right\| < 1$; hence $\lambda \notin \sigma(x+y)$. This gives the desired conclusion, with $\delta = 1/M$. □

Exercise 2.1. Let $\mathscr{A}$ be a complex Banach algebra with identity e. Let x, y in $\mathscr{A}$. Then the following statements hold.

(a) $e - xy$ is invertible if and only if $e - yx$ is.

(b) $\sigma(xy)\backslash\{0\} = \sigma(yx)\backslash\{0\}$.

(c) $r_\sigma(xy) = r_\sigma(yx)$.

(d) If $xy = yx$, then $r_\sigma(x+y) \le r_\sigma(x) + r_\sigma(y)$ and $r_\sigma(xy) \le r_\sigma(x) r_\sigma(y)$.

2.2.1 Dependence of the Spectrum on the Algebra

We begin with an example.

Example 2.12. Let $\partial\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$ and $\mathscr{A} = C(\partial\mathbb{D}; \mathbb{C})$. In Example 2.8, we have computed the spectrum of (the identity mapping) z as an element of $\mathscr{A}$:

$$\sigma_{\mathscr{A}}(z) = \partial\mathbb{D}\,.$$

Let $\mathscr{B}$ = the uniform closure of the polynomials in $\mathscr{A}$. (Here "polynomial" means a polynomial in z.) Now $z \in \mathscr{B}$ and so it has a spectrum as an element of this algebra; denoted by $\sigma_{\mathscr{B}}(z)$. As we have pointed, $\sigma_{\mathscr{A}}(z) \subset \sigma_{\mathscr{B}}(z)$. But there is no reason to believe that $\sigma_{\mathscr{B}}(z) = \sigma_{\mathscr{A}}(z)$. In fact, they are not equal:

$$\sigma_{\mathscr{B}}(z) = \overline{\mathbb{D}}\,.$$

To see this first note that $\|z\| = 1$, so that $\sigma_{\mathscr{B}}(z) \subset \overline{\mathbb{D}}$.

If $|\lambda| \leq 1$ and $\lambda \notin \sigma_{\mathscr{B}}(z)$, there is an f in $\mathscr{B}$ such that $(\lambda - z)f(z) = 1$ for $z \in \partial\mathbb{D}$. Note that this implies that $|\lambda| < 1$. Because $f \in \mathscr{B}$, there is a sequence of polynomials $\{p_n\}$ such that $p_n \to f$ uniformly on $\partial\mathbb{D}$. Thus for every $\epsilon > 0$ there is a $N = N_\epsilon$ such that for $m, n \geq N$,

$$\sup\{|p_n(z) - p_n(z)| : z \in \partial\mathbb{D}\} \leq \epsilon\,.$$

By the maximum principle,

$$\sup\{|p_n(z) - p_n(z)| : z \in \mathbb{D}\} \leq \epsilon\,.$$

Thus $g(z) = \lim p_n(z)$ is analytic on $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$; also, $g|_{\partial\mathbb{D}} = f$. By the same argument, since $p_n(z)(\lambda - z) \to 1$ uniformly on $\partial\mathbb{D}$, $p_n(z)(\lambda - z) \to 1$ uniformly on $\mathbb{D}$. Thus $g(z)(\lambda - z) = 1$ on $\mathbb{D}$. But $g(\lambda)(\lambda - \lambda) = 0$, a contradiction. Thus, $\overline{\mathbb{D}} \subset \sigma_{\mathscr{B}}(z)$.

Thus the spectrum not only depends on the element of the algebra, but also on the algebra. Precisely how this dependence occurs is given below. We begin with the following lemma.

Lemma 2.7. *Suppose V and W are open sets in some topological space X, $V \subset W$, and W contains no boundary point of V. Then V is a union of maximal connected components of W.*

Proof. Let Ω be a maximal connected component of W that intersects V. Since W contains no boundary point of V, Ω is the union of the two disjoint open sets $\Omega \cap V$ and $\Omega \backslash \overline{V}^c$. Since Ω is connected, $\Omega \backslash \overline{V}^c$ is empty. Thus $\Omega \subset V$. $\square$

Lemma 2.8. *Let $\mathscr{A}$ be a complex unital Banach algebra. Let x be a boundary point of $G(\mathscr{A})$ with $x_n \in G(\mathscr{A})$ for $n \geq 1$ so that $x_n \to x$ as $n \to \infty$. Then $\|x_n^{-1}\| \to \infty$ as $n \to \infty$.*

Proof. If the conclusion is false, there exists $M < \infty$ such that $\|x_n^{-1}\| < M$ for infinitely many n. For one of these, $\|x_n - x\| < 1/M$. For this n

$$\left\|e - x_n^{-1}x\right\| = \left\|x_n^{-1}(x_n - x)\right\| < 1$$

so that $x_n^{-1}x \in G(\mathscr{A})$. Since $x = x_n\left(x_n^{-1}x\right)$ and $G(\mathscr{A})$ is a group, it follows that $x \in G(\mathscr{A})$. This contradicts the hypothesis, since $G(\mathscr{A})$ is open. $\square$

Let $\mathscr{A}$ be a unital Banach algebra and $\mathscr{B}$ a closed subalgebra in $\mathscr{A}$ contains the unit element of $\mathscr{A}$. Then $G(\mathscr{B})$ and $\mathscr{B} \cap G(\mathscr{A})$ are two open sets in $\mathscr{B}$. By the previous lemmas, it's easy to see that $G(\mathscr{B})$ contains no boundary point of $\mathscr{B} \cap G(\mathscr{A})$ and hence $G(\mathscr{B})$ is a union of maximal connected components of $\mathscr{B} \cap G(\mathscr{A})$. Moreover, we have :

Theorem 2.9. *Let $\mathscr{A}$ be a complex unital Banach algebra and $\mathscr{B}$ a closed subalgebra in $\mathscr{A}$ contains the unit element of $\mathscr{A}$. Let x be an element in $\mathscr{B}$. Then the following statements hold.*

(a) $\partial\varrho_{\mathscr{B}}(x) = \partial\sigma_{\mathscr{B}}(x) \subset \sigma_{\mathscr{A}}(x)$.

(b) $\varrho_{\mathscr{B}}(x)$ *is a union of maximal connected components of* $\varrho_{\mathscr{A}}(x)$.

(c) $\sigma_{\mathscr{B}}(x)$ *is the union of* $\sigma_{\mathscr{A}}(x)$ *and a (possibly empty) collection of bounded maximal connected components of* $\varrho_{\mathscr{A}}(x)$.

Proof. Note that by Lemma 2.8, $\varrho_{\mathscr{B}}(x)$ contains no boundary point of $\varrho_{\mathscr{A}}(x)$. Thus $\partial\varrho_{\mathscr{B}}(x) = \partial\sigma_{\mathscr{B}}(x) \subset \mathbb{C}\backslash\sigma_{\mathscr{A}}(x) = \varrho_{\mathscr{A}}(x)$. Then by Lemma 2.7, part (b) and part (c) follows. $\square$

Remark 2.6. Intuitively, if $\sigma_{\mathscr{B}}(x)$ is larger than $\sigma_{\mathscr{A}}(x)$, then $\sigma_{\mathscr{B}}(x)$ is obtained from $\sigma_{\mathscr{A}}(x)$ by "*filling in some holes*" in $\sigma_{\mathscr{A}}(x)$. In particular, if $\varrho_{\mathscr{A}}(x)$ is connected (for example $\sigma_{\mathscr{A}}(x)$ contains only real numbers) or if $\sigma_{\mathscr{B}}(x)$ has no interior point, then we have $\sigma_{\mathscr{A}}(x) = \sigma_{\mathscr{B}}(x)$.

As another application of Lemma 2.8, we now give an exercise whose conclusion is the same as that of the Gelfand-Mazur theorem, although its consequences are not nearly so important.

Exercise 2.2. If $\mathscr{A}$ is a complex unital Banach algebra and if there exists $M < \infty$ such that

$$\|x\|\|y\| \leq M\|xy\| \quad (x \in \mathscr{A}, y \in \mathscr{A})$$

then $\mathscr{A}$ is isometrically isomorphic to $\mathbb{C}$. (Hint: Show that $\partial G(\mathscr{A}) = \{0\}$, as a consequence, for each $x \in \mathscr{A}$, $\sigma(x)$ consists of a single point.)

2.2.2 Complex Homomorphisms

Among the important mappings from one Banach algebra into another are the *homomorphisms*.

Definition 2.2. Suppose $\mathscr{A}$, $\mathscr{B}$ be two complex algebra and $\psi : \mathscr{A} \to \mathscr{B}$ is a linear mapping. If

$$\psi(xy) = \psi(x)\psi(y)$$

for all $x, y \in \mathscr{A}$, then ψ is called a *homomorphism*. If in addition ψ is a bijection, then we say it is an *isomorphism*. We say a homomorphism ψ is *nonzero* or *nontrivial*, if ψ is not identically zero.

Of particular interest is the case in which the range is the simplest of all Banach algebras, namely, $\mathbb{C}$ itself. Many of the significant features of the commutative theory depend crucially on a sufficient supply of homomorphisms onto $\mathbb{C}$.

A homomorphism ϕ taking values in $\mathbb{C}$ is called a *complex homomorphism.* For every complex algebra $\mathscr{A}$, trivially, $\phi \equiv 0$ is of course a complex homomorphism on $\mathscr{A}$.

Proposition 2.10. *If ϕ is a nonzero complex homomorphism on a complex algebra $\mathscr{A}$ with unit e, then $\phi(e) = 1$, and $\phi(x) \neq 0$ for every invertible $x \in \mathscr{A}$.*

Proof. For some $y \in \mathscr{A}, \phi(y) \neq 0$. Since

$$\phi(y) = \phi(ye) = \phi(y)\phi(e)$$

it follows that $\phi(e) = 1$. If x is invertible, then

$$\phi(x)\phi\left(x^{-1}\right) = \phi\left(xx^{-1}\right) = \phi(e) = 1$$

so that $\phi(x) \neq 0$. $\square$

Theorem 2.11. *Let $\mathscr{A}$ be a complex unital Banach algebra. Then every complex homomorphism on $\mathscr{A}$ is a contraction; that is,*

$$|\phi(x)| \leq \|x\| \quad \text{for all } x \in \mathscr{A}\,.$$

Proof. Without loss of generality, let $\phi(e) = 1$ where e is the unit element in $\mathscr{A}$. It suffices to show that $|\phi(x)| \leq 1$ for $\|x\| = 1$. In other words, for any $\lambda \in \mathbb{C}$ with $|\lambda| > 1$,

$$\lambda - \phi(x) = \phi(\lambda e - x) \neq 0\,.$$

Observe that since $\|x\| = 1$, $\lambda e - x = \lambda(e - \frac{x}{\lambda})$ is invertible by Lemma 2.1, and then by Proposition 2.10 we have $\phi(\lambda e - x) =\neq 0$ as desired. $\square$

We now interrupt the main line of development and insert a theorem which shows, for Banach algebras, that Theorem 2.20 actually characterizes the complex homomorphisms among the linear functionals. This striking result has apparently found no interesting applications as yet.

Theorem (Gleason, Kahane, Zelazko)**.** *If ϕ is a linear functional on the complex unital Banach algebra $\mathscr{A}$, such that $\phi(e) = 1$ and $\phi(x) \neq 0$ for every invertible $x \in \mathscr{A}$, then ϕ is a complex homomorphism:*

$$\phi(xy) = \phi(x)\phi(y) \quad \textit{for all } x, y \in \mathscr{A} \,.$$

The proof is not easy so we omit it, whereas it can be found in 10.9 Theorem, *Functional Analysis* by W.Rudin.

2.3 Holomorphic Functional Calculus

2.3.1 Introduction

Let $\mathscr{A}$ be a complex Banach algebra with an unit element e. If $x \in \mathscr{A}$ and if $f(\lambda) = \alpha_0 + \alpha_1\lambda + \cdots + \alpha_n\lambda^n$ is a polynomial with complex coefficients α_i, there can be no doubt about the meaning of the symbol $f(x)$; it obviously denotes the element of $\mathscr{A}$ defined by

$$f(x) = \alpha_0 e + \alpha_1 x + \cdots + \alpha_n x^n \,.$$

The defines a mapping from the algebra of polynomials into the algebra $\mathscr{A}$, that is, clearly, a homomorphism. This homomorphism can be extended to a larger class of functions than polynomials; for instance, we can define

$$\exp\{x\} \coloneqq \sum_{n=0}^{\infty} \frac{1}{n!}\, x^n \,.$$

More generally, if $f(\lambda) = \sum_{n=0}^{\infty} \alpha_n\lambda^n$ is any *entire function* in $\mathbb{C}$, it is natural to define $f(x) \in \mathscr{A}$ by

$$f(x) = \sum_{n=0}^{\infty} \alpha_n x^n \,,$$

this series always (absolutely) converges. Another example is given by the meromorphic functions

$$f(\lambda) = \frac{1}{\alpha - \lambda} \,,$$

In this case, the natural definition of $f(x)$ is

$$f(x) = (\alpha e - x)^{-1}$$

which makes sense for all x whose spectrum does not contain α.

One is thus led to the conjecture that $f(x)$ should be definable, within $\mathscr{A}$, whenever f is holomorphic in an open set that contains $\sigma(x)$. This turns out to be correct and can be accomplished by a version of the Cauchy formula that converts complex functions defined in open subsets of $\mathbb{C}$ to $\mathscr{A}$-valued ones defined in certain open subsets of $\mathscr{A}$. (Just as in classical analysis, the Cauchy formula is a much more adaptable tool than the power series representation.)

In certain algebras one can go further. For instance, if $\mathscr{A} = \mathcal{B}(H)$ and $\mathscr{A}$ is a bounded normal operator on a complex Hilbert space H, the symbol $f(\mathscr{A})$ can be interpreted as a bounded normal operator on H when f is any continuous complex function on $\sigma(\mathscr{A})$, and even when f is any complex bounded Borel function on $\sigma(\mathscr{A})$. Later we shall see how this leads to an efficient proof of a very general form of the spectral theorem.

2.3.2 The Riesz-Dunford Integral

Let K be a compact subset of an open $\Omega \subset \mathbb{C}$. It is well known that exists a cycle (see Section **??**) $\gamma = (\gamma_1, \cdots, \gamma_n)$ in $\Omega \backslash K$ so that

$$\mathrm{Ind}_\gamma(z) = \frac{1}{2\pi i} \int_\gamma \frac{d\zeta}{\zeta - z} = \begin{cases} 1\,, & \text{if } z \in K\,; \\ 0\,, & \text{if } z \notin \Omega\,. \end{cases}$$

In other words, γ winds once around every point in K but winds zero time around any point of the complement of Ω. We shall describe this situation briefly by saying that *the cycle γ surrounds K in Ω*. As we know in this case, the Cauchy formula

$$f(z) = \frac{1}{2\pi i} \int_\gamma (\zeta - z)^{-1} f(\zeta) d\zeta$$

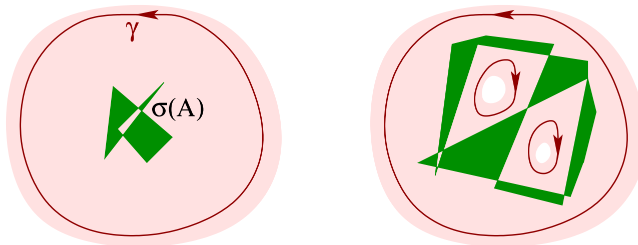

Figure 2.1: A cycle surrounding the spectrum

then holds for every holomorphic function f in Ω and for every $z \in K$.

Note that neither K nor Ω nor the union of the intervals γ_i has been assumed to be connected.

Definition 2.3. Let $\mathscr{A}$ be a complex unital Banach algebra and $x \in \mathscr{A}$. Let Ω be an open set in $\mathbb{C}$ containing $\sigma(x)$. For every holomorphic function f on Ω, i.e., $f \in H(\Omega)$, we define

$$\tilde{f}(x) := \frac{1}{2\pi i}\int_{\gamma} f(\zeta)(\zeta e - x)^{-1} d\zeta\,.$$

where γ is a cycle surrounding $\sigma(x)$ in Ω.

Remark 2.7. The existence of the integral is provided by Theorem **??**. The vector $f(x)$ is independent of the choice of the cycle γ in $\Omega \backslash \sigma(x)$ surrounding $\sigma(x)$ in Ω. In fact, let γ and $\tilde{\gamma}$ be two cycles in $\Omega\backslash\sigma(x)$ that surrounding $\sigma(x)$ in Ω. Then since $\zeta \mapsto f(\zeta)(\zeta e - x)^{-1}$; $\Omega\backslash\sigma(x) \to \mathscr{A}$ is holomorphic, by Theorem **??**, we get

$$\int_{\gamma} f(\zeta)(\zeta e - x)^{-1} d\zeta = \int_{\tilde{\gamma}} f(\zeta)(\zeta e - x)^{-1} d\zeta\,.$$

Lemma 2.12. *Let $\mathscr{A}$ be a complex unital Banach algebra and $x \in \mathscr{A}$. Then for any cycle γ surrounding $\sigma(x)$ in $\mathbb{C}$ and for any $n \geq 0$, we have*

$$\frac{1}{2\pi i}\int_{\gamma} \zeta^n (\zeta e - x)^{-1} d\zeta = x^n\,.$$

Proof. By Remark 2.7, without loss of generality, we can suppose γ is the positive oriented circle with radius strictly lager than $\|x\|$. Then, for each $n \geq 0$,

$$\frac{1}{2\pi i}\int_\gamma \zeta^n(\zeta e - x)^{-1}d\zeta = \frac{1}{2\pi i}\int_\gamma \zeta^n \sum_{k=0}^{\infty} \frac{1}{\zeta^{k+1}}\, x^k d\zeta$$
$$= \sum_{k=0}^{\infty} \frac{1}{2\pi i}\int_\gamma \zeta^n\, \frac{1}{\zeta^{k+1}}\, x^k d\zeta = \sum_{k=0}^{\infty} \frac{1}{2\pi i}\int_\gamma \frac{1}{\zeta^{k+1-n}}\, d\zeta\; x^k = x^n\,,$$

as desired. □

Definition 2.4. Suppose $\mathscr{A}$ is a Banach algebra, Ω is an open set in $\mathbb{C}$, and $H(\Omega)$ is the algebra of all complex holomorphic functions in Ω. By Proposition 2.6,

$$\mathscr{A}_\Omega \coloneqq \{x \in \mathscr{A} : \sigma(x) \subset \Omega\}$$

is an open subset of $\mathscr{A}$.

We define $\tilde{H}\left(\mathscr{A}_\Omega\right)$ to be the set of all $\mathscr{A}$-valued functions $\tilde{f}$, with domain $\mathscr{A}_\Omega$, that arise from an $f \in H(\Omega)$ by the formula

$$\tilde{f}(x) = \frac{1}{2\pi i}\int_\gamma f(\zeta)(\zeta e - x)^{-1}d\zeta\,, \tag{2.2}$$

where γ is any cycle that surrounds $\sigma(x)$ in Ω.

This definition calls for some comments.

Remark 2.8. If $x = \lambda e$ and $\lambda \in \Omega$, (2.2) becomes

$$\tilde{f}(\lambda e) = f(\lambda) e\,. \tag{2.3}$$

Note that $\sigma(\lambda e) = \lambda$, so $\lambda e \in \mathscr{A}_\Omega$ if and only if $\lambda \in \Omega$. If we identity $\lambda \in \mathbb{C}$ with $\lambda e \in \mathscr{A}$, every $f \in H(\Omega)$ may be regarded as mapping a certain subset of $\mathscr{A}_\Omega$ (namely, the intersection of $\mathscr{A}_\Omega$ with the one-dimensional subspace of $\mathscr{A}$ generated by e) into $\mathscr{A}$, and then (2.3) shows that $\tilde{f}$ may be regarded as an extension of f. In most treatments of this topic, $f(x)$ is written in place of our $\tilde{f}(x)$. The notation $\tilde{f}$ is used here because it avoids certain ambiguities that might cause misunderstandings.

Remark 2.9. If S is any set and $\mathscr{A}$ is any algebra, the collection of all $\mathscr{A}$-valued functions on S is an algebra, if scalar multiplication, addition, and multiplication are defined pointwise. For instance, if u and v map S into $\mathscr{A}$, then

$$(uv)(s) = u(s)v(s) \quad (s \in S)\,.$$

This will be applied to $\mathscr{A}$-valued functions defined in $\mathscr{A}_\Omega$.

Theorem 2.13. *Suppose $\mathscr{A}$, $H(\Omega)$, and $\tilde{H}(\mathscr{A}_\Omega)$ are as in Definition 2.4. Then clearly $H(\Omega)$, $\tilde{H}(\mathscr{A}_\Omega)$ both are complex algebras.*

(a) The mapping $f \to \tilde{f}$ is an isomorphism of $H(\Omega)$ onto $\tilde{H}(\mathscr{A}_\Omega)$.

(b) The mapping $f \to \tilde{f}$ is continuous in the following sense: if $f_n \in H(\Omega)$ and $f_n \to f$ uniformly on compact subsets of Ω, then

$$\tilde{f}(x) = \lim_{n\to\infty} \tilde{f}_n(x) \quad \textit{for} \quad x \in \mathscr{A}_\Omega\,.$$

Remark 2.10. Since $H(\Omega)$ is obviously a commutative algebra, Theorem 2.13 implies that $\tilde{H}(\mathscr{A}_\Omega)$ is also commutative. This may be surprising, because $\tilde{f}(x)$ and $\tilde{f}(y)$ need not commute. However, $\tilde{f}(x)$ and $\tilde{g}(x)$ do commute in $\mathscr{A}$ for every $x \in \mathscr{A}_\Omega$.

Proof. Take $x \in \mathscr{A}_\Omega$. The assertion that $\widetilde{f+g}(x) = \tilde{f}(x) + \tilde{g}(x)$ follows directly from the definition.

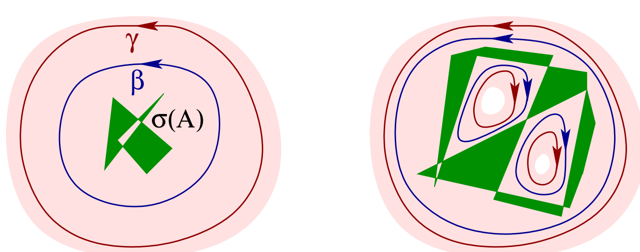

Figure 2.2: Two cycles encircling the spectrum

To prove $\widetilde{fg}(x) = \tilde{f}(x)\tilde{g}(x)$, choose two cycles β and γ in $\Omega\backslash\sigma(x)$ that both surround $\sigma(x)$ in Ω have disjoint images so that

$$\mathrm{im}(\beta) \cap \mathrm{im}(\gamma) = \emptyset$$

and such that the image of β is encircled by γ, i.e.

$$\begin{aligned} \mathrm{Ind}_\gamma(w) &= 1 \quad \text{for all } w \in \mathrm{im}(\beta) \\ \mathrm{Ind}_\gamma(z) &= 0 \quad \text{for all } z \in \mathrm{im}(\gamma) \end{aligned}$$

Then, by the resolvent identity,

$$\begin{aligned} f(x)g(x) =& \frac{1}{2\pi \mathrm{i}} \int_\beta f(w)(we-x)^{-1}dw \; \frac{1}{2\pi \mathrm{i}} \int_\gamma g(z)(ze-x)^{-1}dz \\ =& \frac{1}{2\pi \mathrm{i}} \frac{1}{2\pi \mathrm{i}} \int_\beta \int_\gamma f(w)g(z) \frac{(we-x)^{-1} - (ze-x)^{-1}}{z-w} dzdw \\ =& \frac{1}{2\pi \mathrm{i}} \int_\beta f(w) \left(\frac{1}{2\pi \mathrm{i}} \int_\gamma \frac{g(z)dz}{z-w} \right) (we-x)^{-1}dw \\ &+ \frac{1}{2\pi \mathrm{i}} \int_\gamma g(z) \left(\frac{1}{2\pi \mathrm{i}} \int_\beta \frac{f(w)dw}{w-z} \right) (ze-x)^{-1}dz \\ =& \frac{1}{2\pi \mathrm{i}} \int_\beta f(w)g(w)(we-x)^{-1}dw = (fg)(x)\,. \end{aligned}$$

Therefore, $f \mapsto \tilde{f}$ is a homomorphism.

If $\tilde{f} = 0$, then $f(\lambda)e = \tilde{f}(\lambda e) = 0$ for all $\lambda \in \Omega$ so that $f = 0$. Thus $f \to \tilde{f}$ is one-to-one. By the definition on $\tilde{H}(\mathscr{A}_\Omega)$, $f \to \tilde{f}$ is a surjection. Thus $f \to \tilde{f}$ is an algebra isomorphism of $H(\Omega)$ onto $\tilde{H}(\Omega)$.

The asserted continuity follows directly from the integral (2.2). Since for fixed x, $\left\|(\zeta e - x)^{-1}\right\|$ is bounded on $\zeta \in \gamma$. Use the same cycle γ in $\Omega\backslash\sigma(x)$ for all representation of $\tilde{f}_n$, and apply Proposition **??**, we get the desired result. □

Theorem 2.14 (Spectral Mapping Theorem)**.** *Suppose $x \in \mathscr{A}_\Omega$ and $f \in H(\Omega)$. Then $\tilde{f}(x)$ is invertible in $\mathscr{A}$ if and only if $f(\lambda) \neq 0$ for every $\lambda \in \sigma(x)$. Moreover,*

$$\sigma(\tilde{f}(x)) = f(\sigma(x))\,.$$

Proof. If f has no zero on $\sigma(x)$, then $g = 1/f$ is holomorphic in an open set U such that $\sigma(x) \subset U \subset \Omega$. Since $fg = 1$ in U, Theorem 2.13 (with U in place of Ω) shows that $\tilde{f}(x)\tilde{g}(x) = e$, and thus $\tilde{f}(x)$ is invertible.

Conversely, if $f(\lambda) = 0$ for some $\lambda \in \sigma(x)$, then there exists $h \in H(\Omega)$ such that

$$f(z) = (\lambda - z)h(z) \text{ for } z \in \Omega\,;$$

which, by Theorem 2.13, implies that

$$\tilde{f}(x) = (\lambda e - x)\tilde{h}(x) = \tilde{h}(x)(x - \lambda e) \tag{2.4}$$

Since $\lambda e - x$ is not invertible in A, neither is $\tilde{f}(x)$ by (2.4).

Now fix $\lambda \in \mathbb{C}$. By definition, $\lambda \in \sigma(\tilde{f}(x))$ if and only if $\lambda e - \tilde{f}(x)$ is not invertible in $\mathscr{A}$. By the previous argument, applied to $f - \lambda$ is place of f, this happens if and only if $f - \lambda$ has a zero in $\sigma(x)$, that is, if and only if $\lambda \in f(\sigma(x))$. □

The spectral mapping theorem makes it possible to include composition of functions among the operations of the symbolic calculus.

Theorem 2.15. *Suppose $x \in \mathscr{A}_\Omega$, $f \in H(\Omega)$, U is an open set containing $f(\Omega)$, $g \in H\,(U)\,$, and $(g \circ f)(\lambda) = g(f(\lambda))$ for $\lambda \in \Omega$. Then*

$$\widetilde{(g \circ f)}(x) = \tilde{g}(\tilde{f}(x))\,.$$

Proof. Note first that the operator $g(f(x))$ is well defined, because $\sigma(f(x)) = f(\sigma(x)) \subset f(\Omega) \subset U$ by the spectrum mapping theorem. Fix a contour Γ_1 that surrounds $f(\sigma(x))$ in U. Then, by definition,

$$\tilde{g}(\tilde{f}(x)) = \frac{1}{2\pi i}\int_{\Gamma_1} g(\zeta)[\zeta e - \tilde{f}(x)]^{-1} d\zeta$$

For fixed $\zeta \in \operatorname{im}(\Gamma_1)$, we want to represent $(\zeta e - f(x))^{-1}$ by Cauchy integral and hence we need the mapping $\lambda \to [\zeta e - \tilde{f}(x)]^{-1}$ is holomorphic on some open neighborhood of $\sigma(x)$. It suffices to request that $f(\lambda) \cap \operatorname{im}(\Gamma_1) =$

$\emptyset$. Note that Γ_1 surrounds $f(\sigma(x))$ in U, we can choose an open set W, with $f(\sigma(x)) \subset W \subset U$, so small that

$$\mathrm{Ind}_{\Gamma_1}(z) = 1 \text{ for } z \in W\,.$$

Thus $W \cap \mathrm{im}(\Gamma_1) = \emptyset$ and $\sigma(x) \subset f^{-1}(W) \subset \Omega$. Now, since $\lambda \to (\zeta - f(\lambda))^{-1}$ is holomorphic on $f^{-1}(W)$, fix a contour Γ_0 that surrounds $\sigma(x)$ in $f^{-1}(W)$, we have

$$[\zeta e - \tilde{f}(x)]^{-1} = \frac{1}{2\pi i}\int_{\Gamma_0} [\zeta - f(z)]^{-1}(ze - x)^{-1} dz \quad (\zeta \in \Gamma_1)$$

Thus

$$\begin{aligned}
\tilde{g}(\tilde{f}(x)) &= \frac{1}{2\pi i}\int_{\Gamma_1} g(\zeta)[\zeta e - \tilde{f}(x)]^{-1} d\zeta \\
&= \frac{1}{2\pi i}\int_{\Gamma_0} \frac{1}{2\pi i}\int_{\Gamma_1} g(\zeta)[\zeta - f(z)]^{-1} d\zeta (ze - x)^{-1} dz \\
&= \frac{1}{2\pi i}\int_{\Gamma_0} g(f(z))(ze - x)^{-1} dz = \widetilde{(g \circ f)}(x)\,,
\end{aligned}$$

as desired. $\square$

2.3.3 Spectral Projection

Now we consider the complex Banach algebra $\mathscr{A} = \mathcal{B}(X)$, the Banach algebra of all bounded linear operators on the complex Banach space X.

Suppose that the spectrum of $A \in \mathcal{B}(X)$ can be decomposed as the union of N pairwise disjoint closed components:

$$\sigma(A) = \Sigma_1 \cup \cdots \cup \Sigma_N, \quad \Sigma_j \cap \Sigma_k = \emptyset \;\text{ if } k \neq j\,.$$

Let $U_j \subset \mathbb{C}$ be disjoint open sets such that $\Sigma_j \subset U_j$ for $j = 0, 1, \cdots, N$. Let $U = \cup_{j=1}^{N} U_j$ and define the function $f_j : U \to \mathbb{C}$ by

$$f|_{U_j} \coloneqq 1 \;\text{ and }\; f|_{U_k} \coloneqq 0 \;\text{ for } k \neq j\,.$$

Trivially, f_j is a holomorphic function on $U \supset \sigma(A)$. Define

$$P_j \coloneqq f_j(A)\,.$$

In other words, if a cycle γ surrounds $\sigma(A)$ in U, there must exists cycles γ_j surrounding Σ_j in U_j so that $\gamma = \gamma_1 + \cdots + \gamma_N$. Then it's easy to see that

$$P_j = \int_{\gamma_j} (\zeta I - A)^{-1} d\zeta \, .$$

Theorem 2.16. *Let $\{P_j\}_{j=1}^N$ be defined above. Then*

(a) The $\{P_j\}_{j=1}^N$ are disjoint projections that is,

$$P_j^2 = P_j \quad \textit{and} \quad P_j P_k = 0 \quad \textit{for} \quad j \neq k \, ;$$

and $\sum_{j=1}^N P_j = I$.

(b) For each j, P_j and A commute. Thus $X_j \coloneqq R(P_j)$ are closed A-invariant subspaces of X such that $X = X_1 \oplus \cdots \oplus X_N$. The spectrum of the operator $A_j \coloneqq A|_{X_j} : X_j \to X_j$ is given by $\sigma(A_j) = \Sigma_j$.

Proof. Since $f_j f_k = \delta_{jk} f_j$ it follows from that $P_j P_k = \delta_{jk} P_j$. Moreover P_j commutes with A by definition.

Let $\alpha \in \mathbb{C}$. Consider the operator

$$\tilde{A}_j \coloneqq \begin{bmatrix} \alpha I_{X_1} & & & & \\ & \ddots & & & \\ & & A_j & & \\ & & & \ddots & \\ & & & & \alpha I_{X_N} \end{bmatrix}$$

Define $g : U \to \mathbb{C}$ by $g(z) = z$ for $z \in U$ and let $c \in \mathbb{C}$. Then, we can see that

$$\tilde{A} = (\alpha(1-f) + gf)(A) \; , \quad \sigma(\tilde{A}) = \{\alpha\} \cup \Sigma_j \, .$$

If $\lambda \in \mathbb{C} \backslash \Sigma_j$, it follows that the operator $\lambda I - \tilde{A}$ is bijective for $\alpha \neq \lambda$ and so $\lambda I_{X_j} - A_j$ is bijective. Conversely, suppose $\lambda \in \Sigma_j$. Then $\lambda I - \tilde{A}$ is not bijective and, for $\alpha \neq \lambda$, this implies that $\lambda I_{X_j} - A_j$ is not bijective. Thus $\sigma(A_j) = \Sigma_j$. $\square$

Example 2.13 (Spectral Projection)**.** Let X be a complex Banach space, let $A \in \mathcal{C}(X)$ be a compact linear operator on X, let $\lambda \in \sigma(A)$ be a nonzero eigenvalue of A, and choose $r \in \mathbb{N}$ such that

$$E_\lambda \coloneqq N(\lambda I - A)^r = N(\lambda I - A)^{r+1} .$$

The proof of Lemma **??** and Theorem **??** shows that such an integer $r \geq 1$ exists, that E_λ is a finite-dimensional linear subspace of X, that the operator $(\lambda I - A)^r$ has a closed image, and that

$$X = N(\lambda I - A)^r \oplus R(\lambda I - A)^r .$$

Hence the formula

$$P_\lambda\left(x_0 + x_1\right) = x_0 \text{ for } x_0 \in N(\lambda I - A)^r \quad \text{and} \quad x_1 \in R(\lambda I - A)^r \tag{2.5}$$

defines a bounded linear operator $P_\lambda : X \to X$ which is an A-invariant projection onto E_λ, i.e.

$$P_\lambda^2 = P_\lambda, \quad P_\lambda A = AP_\lambda, \quad R\left(P_\lambda\right) = E_\lambda .$$

The operator P_λ is uniquely determined by (2.5) and is called the *spectral projection* associated to the eigenvalue λ.

It can also be written in the form

$$P_\lambda = \frac{1}{2\pi i}\int_\gamma (\zeta I - A)^{-1} d\zeta ,$$

where $\gamma(t)$ is the positive oriented circle centering at λ with radius r so that $D(\lambda, r) \cap \sigma(A) = \{\lambda\}$.

2.4 The Gelfand Theory

This chapter deals primarily with the Gelfand theory of commutative Banach algebras, although some of the results of this theory will be applied to noncommutative situations.

The main result of this section is that, for each commutative Banach algebra $\mathscr{A}$, there is a compact Hausdorff space Δ, and a continuous homomorphism $\Gamma : \mathscr{A} \to C(\Delta)$. Indeed, Δ consists of all the nonzero complex homomorphisms on $\mathscr{A}$.

We begin with some preliminary works.

2.4.1 Ideals and Homomorphisms

Definition 2.5. A subset J of a complex algebra $\mathscr{A}$ is said to be an *ideal* if

(a) J is a subspace of $\mathscr{A}$ (in the vector space sense), and

(b) xy and yx in J whenever $x \in \mathscr{A}$ and $y \in J$.

If $J \neq \mathscr{A}$, we say J is a *proper ideal.* If J is a proper ideal which is not contained in any larger proper ideal, we say J is a *maximal ideal.*

Given a ideal of $\mathscr{A}$, we can define the *quotient algebra* $\mathscr{A}/J$ as follows. Clearly $\mathscr{A}/J$ has a natural linear structure. We define the multiplication on $\mathscr{A}/J$ by

$$\tilde{x} \cdot \tilde{y} \coloneqq \widetilde{xy} \qquad \text{for all} \;\; \tilde{x}, \tilde{y} \in \mathscr{A}/J\,.$$

The it's easy to see that $\mathscr{A}/J$ is an algebra.

The following proposition is trivial, so we omit the proof.

Proposition 2.17. *Let $\mathscr{A}$ be a complex algebra with unit e. Let J be a proper ideal of $\mathscr{A}$. Then*

(a) $e \notin J$, and

(b) if a is invertible, then $a \notin J$.

By Zorn's lemma, we have the following assertion.

Proposition 2.18. *Every proper ideal is contained in a maximal ideal.*

Proof. Let J be a proper ideal of $\mathscr{A}$. Let $\mathscr{P}_J$ be the collection of all proper ideals of $\mathscr{A}$ that contain J. Partially order $\mathscr{P}$ by set inclusion, let $\mathscr{I}$ be a totally ordered subset of $\mathscr{P}$, and let M be the union of all members of $\mathscr{I}$. Being the union of a totally ordered collection of ideals, M is an ideal. Obviously $J \subset M$, and $M \neq \mathscr{A}$. Thus $\mathscr{I}$ has an upper bounded. By zorn's lemma, $\mathscr{P}_J$ has a maximal element, as required. □

Proposition 2.19. *Let $\mathscr{A}$ be a complex Banach algebra with unit e. Then the following statements hold.*

(a) Let J be a ideal of $\mathscr{A}$, then so is $\overline{J}$.

(b) If J is a maximal ideal of $\mathscr{A}$, then J is closed.

Proof. We begin with part (a). Clearly $\overline{J}$ is a linear subspace of $\mathscr{A}$, it suffices to show the absorption law. Take $x \in \overline{J}$ and $y \in \mathscr{A}$, then there exists (x_n) in J so that $x_n \to x$. Then $x_n y \to xy$ and $yx_n \to yx$. Since $x_n y$ and yx_n belong to J, thus xy and yx belong to $\overline{J}$. Thus $\overline{J}$ is an ideal.

Now show that part (b). If J is a maximal ideal, to show $J = \overline{J}$, it suffices to show that $\overline{J} \neq \mathscr{A}$. Observe that an ideal I of $\mathscr{A}$ equals $\mathscr{A}$ iff $I \cap G(\mathscr{A}) \neq \emptyset$. So we have only to show that $\overline{J} \cap G(\mathscr{A}) = \emptyset$. However, since J is maximal,

$$J \cap G(\mathscr{A}) = \emptyset\,.$$

Since $G(\mathscr{A})$ is open, we get $\overline{J} \cap G(\mathscr{A}) = \emptyset$, as desired. □

The following result highly depends on the commutativity of $\mathscr{A}$ and the fact that $\mathscr{A}$ has a unit element.

Theorem 2.20. *Let $\mathscr{A}$ be a complex commutative algebra with unit element e. Then the following statement hold.*

(a) An element $x \in \mathscr{A}$ is invertible if and only if x is not contained in any maximal ideal of $\mathscr{A}$.

(b) *A proper ideal J of $\mathscr{A}$ is maximal if and only if $\mathscr{A}/J$ is a division algebra.*

Proof. (a). Clearly if x is invertible, then any proper ideal does not contain x. If x is not invertible, then consider

$$x\mathscr{A}\,,$$

the ideal generated by $\mathscr{A}$. Since x is not invertible, so $x\mathscr{A} \neq \mathscr{A}$. Thus $x\mathscr{A}$ is a proper ideal containing x.

(b). Note that a proper ideal J is maximal iff for each $x \notin J$, the ideal generated by $J \cup \{x\}$ equals $\mathscr{A}$. Since $\mathscr{A}$ is unital and commutative, the ideal generated by $J \cup \{x\}$ is given by

$$x\mathscr{A} + J\,.$$

On the other hand, $x\mathscr{A} + J = \mathscr{A}$ if and only if $\tilde{x}$ is invertible in $\mathscr{A}/J$, and hence the desired result follows. □

Homomorphisms and quotient algebras If $\mathscr{A}$ and $\mathscr{B}$ are Banach algebras and ψ is a nonzero continuous homomorphism of $\mathscr{A}$ into $\mathscr{B}$, then

$$\ker(\psi)$$

the null space or kernel of ψ, is obviously an *closed proper ideal* in $\mathscr{A}$.

Conversely, dose every closed proper ideal J in $\mathscr{A}$ is the kernel of some homomorphism of $\mathscr{A}$? First of all, $\mathscr{A}/J$ is a Banach space, with respect to the quotient norm, and also an algebra. Indeed it's a Banach algebra, since

$$\begin{aligned}\|\tilde{x}\tilde{y}\| = \|\widetilde{xy}\| &= \inf\{\|xy + j\| : j \in J\} \\ &\leq \inf\{\|(x + j_1)(y + j_2)\| : j_1, j_2 \in J\} \\ &\leq \inf\{\|(x + j_1)\|\|(y + j_2)\| : j_1, j_2 \in J\} = \|\tilde{x}\|\|\tilde{y}\|\,.\end{aligned}$$

Then clearly the quotient map $Q : \mathscr{A} \to \mathscr{A}/J$ is a continuous homomorphism so that $\ker(Q) = J$.

Complex homomorphisms and maximal ideals Let $\mathscr{A}$ be a complex Banach algebra with unit e. If ϕ is a nonzero complex homomorphism of $\mathscr{A}$ onto $\mathbb{C}$, then $\ker(\phi)$ must be a maximal ideal of $\mathscr{A}$ since

$$\mathscr{A}/\ker(\phi) \text{ ad } \operatorname{Im}(\phi) = \mathbb{C} \text{ are isomorphism } .$$

We denote by Δ all the nonzero complex homomorphism on $\mathscr{A}$ an by $\mathfrak{M}$ all the maximal ideals of $\mathscr{A}$. Then the mapping

$$\Delta \to \mathfrak{M} \; ; \; \phi \to \ker(\phi) \tag{2.6}$$

is a injection. Indeed, if $\phi_1, \phi_2 \in \Delta$ so that $\ker(\phi_1) = \ker(\phi_2)$, then by Lemma **??**, there exists constant $c \in \mathbb{C}$ so that $\phi_1 = c\phi_2$. Since $\phi_1(e) = \phi_2(e) = 1$, thus $c = 1$ and $\phi_1 = \phi_2$.

It's natural to ask in which case the mapping (2.6) is a bijection? We will give an affirmative answer in the case that $\mathscr{A}$ is a complex commutative unital Banach algebra. This ono-on-one correspondence (2.6) is one of the key facts of the whole theory.

Theorem 2.21. *Let $\mathscr{A}$ be a complex commutative Banach algebra with unit element, and let Δ be the set of all complex homomorphisms of $\mathscr{A}$.*

(a) The mapping $\phi \mapsto \ker(\phi)$; $\Delta \to \mathfrak{M}$ is a bijection.

(b) $x \in \mathscr{A}$ is invertible if and only if $\phi(x) \neq 0$ for all $\phi \in \Delta$.

(c) The spectrum of x is given by $\sigma(x) = \{\phi(x) : \phi \in \Delta\}$.

Proof. (a). It suffices to show that the mapping $\phi \mapsto \ker(\phi)$; $\Delta \to \mathfrak{M}$ is a surjection. Let $J \in \mathfrak{M}$ be a maximal ideal of $\mathscr{A}$. Then by Theorem 2.19, J is closed and $\mathscr{A}/J$ is therefore a Banach algebra. It follows from Theorem 2.20 that $\mathscr{A}/J$ is division algebra. By the Gelfand-Mazur theorem, there is an isometrical (algebra) isomorphism h of $\mathscr{A}/J$ onto $\mathbb{C}$. Put $\phi_J = h \circ \pi$. Then $\phi \in \Delta$, and J is the kernel of ϕ_J, as desired.

(b). If x is invertible in $\mathscr{A}$, then by Proposition 2.10, $\phi(x) \neq 0$ for all $\phi \in \Delta$. If x is not invertible in $\mathscr{A}$, by Theorem 2.20, there is a maximal ideal J of $\mathscr{A}$ containing x. Then ϕ_J defined in the proof (a) satisfies $\phi_J(x) = 0$ and hence part (b) follows.

(c). Apply part (b) to $\lambda e - x$ in place of x. $\square$

Now we are ready for the Gelfand representation.

2.4.2 The Gelfand Representation

In this subsection, unless otherwise specified, we suppose that $\mathscr{A}$ is a complex commutative unital Banach algebra. Let Δ be the set of all nonzero complex homomorphisms of $\mathscr{A}$. Since every complex homomorphism is a contraction linear operator (see Theorem 2.11), $\Delta \subset \mathscr{A}^*$.

The formula

$$\hat{x}(\phi) = \phi(x) \quad (\phi \in \Delta)$$

assigns to each $x \in \mathscr{A}$ a function $\hat{x} : \Delta \to \mathbb{C}$. We call $\hat{x}$ the *Gelfand representation* of x and also denote it by Γx. Indeed, $\hat{x}$ is the restriction of $J_{\mathscr{A}} x$ on $\Delta \subset \mathscr{A}^*$, where $J_{\mathscr{A}}$ is the natural embedding of $\mathscr{A}$ into $\mathscr{A}^{**}$.

The *Gelfand topology* of Δ is the weakest topology that makes every $\hat{x} \in \hat{\mathscr{A}}$ continuous. In other words, regarding Δ as a subset of $\mathscr{A}^*$ the Gelfand topology is the subspace topology inherited form $(\mathscr{A}^*, \tau_{w^*})$.

Let $\hat{\mathscr{A}}$ be the set of all $\hat{x}$, for $x \in \mathscr{A}$. Since Δ is equipped with the Gelfand topology, we can see that $\hat{\mathscr{A}} \subset C(\Delta)$, the algebra of all complex continuous functions on Δ. The mapping

$$\Gamma : \mathscr{A} \to \hat{\mathscr{A}} \subset C(\Delta)$$

is called the *Gelfand representation* of $\mathscr{A}$.

Lemma 2.22. *Endowed with the Gelfand topology, Δ is a compact Hausdorff space.*

Proof. As we have pointed, the Gelfand topology on Δ is the subspace topology inherited form $(\mathscr{A}^*, \tau_{w^*})$. Note that by Theorem 2.11, $\Delta \subset B_{\mathscr{A}^*}$, the closed unit ball in $\mathscr{A}^*$. By Theorem **??**, it suffices to show that Δ is a weak-star closed subset of $B_{\mathscr{A}^*}$.

So it suffices to show that if (ϕ_α) is a net in Δ converging to $\phi \in \mathscr{A}^*$ relative to the weak star topology, then $\phi \in \Delta$. To this end, since $\phi \in \mathscr{A}^*$, we have only to show that ϕ preserve the multiplication. This is trivial. Indeed, for each x, y in $\mathscr{A}$, since $\phi_\alpha \to \phi$ in $\mathscr{A}^*$, we have $\phi_\alpha(z) \to \phi(z)$ for all z, and hence

$$\begin{aligned} \phi(xy) &= \lim_\alpha \phi_\alpha(xy) = \lim_\alpha \phi_\alpha(x)\phi_\alpha(y) \\ &= \lim_\alpha \phi_\alpha(x) \lim_\alpha \phi_\alpha(y) = \phi(x)\phi(y)\,, \end{aligned}$$

as desired. □

Remark 2.11. Since there is a one-to-one correspondence between Δ and $\mathfrak{M}$, all the maximal ideals of $\mathscr{A}$, we can also define the Gelfand topology on $\mathfrak{M}$. Equipped with its Gelfand topology, $\mathfrak{M}$ is a compact Hausdorff space, and is usually called the *maximal ideal space* of $\mathscr{A}$.

Theorem 2.23. *Let Δ be the all the nonzero complex homomorphisms on $\mathscr{A}$ equipped with the Gelfand topology.*

(a) The Gelfand representation is a continuous homomorphism of $\mathscr{A}$ onto a subalgebra $\hat{\mathscr{A}}$ of $C(\Delta)$.

(b) For each $x \in \mathscr{A}$, the range of $\hat{x} = \Gamma x$ is the spectrum $\sigma(x)$. Hence

$$\|\Gamma x\|_{C(\Delta)} = r_\sigma(x) \le \|x\|\,.$$

Proof. It's trivial to check that the mapping $x \to \hat{x}$ is a homomorphism and $\hat{\mathscr{A}}$ is a subalgebra of $C(\Delta)$. Observe that by Theorem 2.21, the range of $\hat{x}$

$$\{\hat{x}(\phi) : \phi \in \Delta\} = \{\phi(x) : \phi \in \Delta\} = \sigma(x)\,;$$

and hence

$$\|\hat{x}\|_{C(\Delta)} = \sup_{\phi\in\Delta} |\phi(x)| = \sup_{\lambda\in\sigma(x)} |\lambda| = r_\sigma(x) \le \|x\| \,.$$

Thus Γ is continuous. $\square$

Next, we are going to discuss the following questions.

- When does the Gelfand representation is an isomorphism of $\mathscr{A}$ onto $\hat{\mathscr{A}}$?
- When does the Gelfand representation is an isometry of $\mathscr{A}$ into $\hat{\mathscr{A}}$?
- When does the Gelfand representation is an isometric isomorphism of $\mathscr{A}$ onto $\hat{\mathscr{A}}$ and $\mathscr{A} = C(\Delta)$?

The third question is rather difficult, we will encounter it when talking about the C^* algebra. Let's see the first two questions.

When does the Gelfand representation is an isomorphism of $\mathscr{A}$ onto $\hat{\mathscr{A}}$? It's easy to see that

$$\begin{aligned}\hat{x} = 0 &\Leftrightarrow \hat{x}(\phi) = 0 \text{ for all } \phi \in \Delta \\ &\Leftrightarrow \phi(x) = 0 \text{ for all } \phi \in \Delta \\ &\Leftrightarrow x \in \bigcap_{\phi\in\Delta} \ker(\phi) \Leftrightarrow x \in \bigcap_{J\in\mathfrak{M}} J \,.\end{aligned}$$

Thus

$$\ker(\Gamma) = \bigcap_{\phi\in\Delta} \ker(\phi) = \bigcap_{J\in\mathfrak{M}} J \,.$$

In the study of algebra, for any algebra $\mathscr{A}$,

$$\operatorname{rad}(\mathscr{A}) \coloneqq \bigcap_{J\in\mathfrak{M}} J \,,$$

the intersection of all maximal ideals of $\mathscr{A}$, is called the *radical* of $\mathscr{A}$. If $\operatorname{rad}(\mathscr{A}) = \{0\}$, then $\mathscr{A}$ is called *semisimple*. Clearly, $\Gamma : \mathscr{A} \to \hat{\mathscr{A}}$ is an isomorphism if and only if $\mathscr{A}$ is semisimple.

Exercise 2.3. Let $\mathscr{A}$ be a complex commutative unital Banach algebra. Then the following statements are equivalent.

(a) $\mathscr{A}$ is semisimple.

(b) Δ separates the points on $\mathscr{A}$, i.e., $\cap_{\phi\in\Delta}\ker(\phi) = \{0\}$.

(c) For each $x \in \mathscr{A}$, $r_\sigma(x) = \{0\}$ if and only if $x = 0$.

(d) For each $x \in \mathscr{A}$, $\sigma(x) = \{0\}$ if and only if $x = 0$.

Semisimple algebras have an important property which was earlier proved for the complex field $\mathbb{C}$ in Theorem 2.11:

Proposition 2.24. *Let $\mathscr{A}$, $\mathscr{B}$ be two complex unital Banach algebra. Suppose $\mathscr{B}$ is commutative and semisimple. Then each homomorphism $\psi : \mathscr{A} \to \mathscr{B}$ is continuous.*

Proof. By the closed graph theorem, it suffices to show that ψ is a closed operator. Suppose $x_n \to x$ in $\mathscr{A}$ and $\psi(x_n) \to y$ in $\mathscr{B}$. In order to show $y = \psi(x)$, since $\mathscr{B}$ is commutative and semisimple, it suffices to show that $\phi(y) = \phi(\psi(x))$ for each homomorphism $\phi : \mathscr{B} \to \mathbb{C}$.

By Theorem 2.20, since ϕ is continuous ,$\phi(\psi(x_n)) \to \phi(y)$. On the other hand, note that $\phi \circ \psi$ is also a complex homomorphism (of $\mathscr{A}$), $\phi \circ \psi$ is continuous, hence $\phi(\psi(x_n)) \to \phi(\psi(x))$. Thus $\phi(y) = \phi(\psi(x))$. $\square$

Remark 2.12. Every isomorphism between two semisimple commutative unital Banach algebras is a homeomorphism. In particular, this is true of every automorphism of a semisimple commutative unital Banach algebra. *The topology of such an algebra is therefore completely determined by its algebraic structure.*

When does the Gelfand representation is an isometry of $\mathscr{A}$ into $\hat{\mathscr{A}}$? If $\|\Gamma x\| \equiv r_\sigma(x) = \|x\|$ for all $x \in \mathscr{A}$, then we have $\|x^2\|^{1/2} \geq r_\sigma(x) = \|x\|$, thus $\|x^2\| \geq \|x\|^2$. Recall that $\left\|x^2\right\| \leq \|x\|^2$ is always true, then we get $\|x^2\| = \|x\|^2$.

On the contrary, if $\|x^2\| = \|x\|^2$ for all $x \in \mathscr{A}$, then by induction we have $\|x^{2^n}\| = \|x\|^{2^n}$, thus

$$\|\Gamma x\| \equiv r_\sigma(x) = \lim_{n\to\infty} \|x^{2^n}\|^{\frac{1}{2^n}} = \|x\|\,.$$

Therefore we get

Proposition 2.25. *The Gelfand transform is an isometry iff*

$$\left\|x^2\right\| = \|x\|^2$$

for all $x \in \mathscr{A}$.

In the end of this subsection, we point out that whether the Gelfand representation Γ is a homeomorphism of $\mathscr{A}$ onto $\hat{\mathscr{A}}$ can be decided by comparing $\left\|x^2\right\|$ with $\|x\|^2$, for all $x \in \mathscr{A}$.

Lemma 2.26. *Let $\mathscr{A}$ be a complex commutative unital Banach algebra. Then*

$$\inf_{\|x\|=1} \left\|x^2\right\| \leq \inf_{\|x\|=1} \|\Gamma x\| \leq \left(\inf_{\|x\|=1} \left\|x^2\right\|\right)^{1/2}.$$

Proof. Set $r = \inf_{\|x\|=1} \left\|x^2\right\|$. Then we have $\|x^2\| \geq r\|x\|^2$ for all $x \in \mathscr{A}$. By induction

$$\|x^{2^n}\| \geq r^{2^n-1}\|x\|^{2^n}$$

for all $n \geq 1$. Thus for each x,

$$\|\Gamma x\| = \lim_{n\to\infty} \|x^{2^n}\|^{\frac{1}{2^n}} \geq \lim_{n\to\infty} r^{1-\frac{1}{2^n}} \|x\| = r\|x\|\,.$$

To show another inequality, set $s = \inf_{\|x\|=1} \|\hat{x}\|$, then $\|\Gamma x\| \geq s\|x\|$ for all $x \in \mathscr{A}$. Since $\|\Gamma x\| = \inf_{n\geq 1} \|x^n\|^{\frac{1}{n}}$, we have $\|x^2\| \geq s^2\|x\|^2$. and hence $s^2 \leq \inf_{\|x\|=1} \left\|x^2\right\|$. $\square$

Theorem 2.27. *Let $\mathscr{A}$ be a commutative unital Banach algebra. $\mathscr{A}$ is semisimple and $\hat{\mathscr{A}}$ is closed in $C(\Delta)$ iff $\inf_{\|x\|=1} \|x\|^2 > 0$.*

Proof. If $\inf_{\|x\|=1} \|x\|^2 > 0$, then by the previous lemma $\inf_{\|x\|=1} \|\Gamma x\| > 0$ thus $\Gamma : \mathscr{A} \to \hat{\mathscr{A}}$ is a homeomorphism. Thus $\mathscr{A}$ is semisimple and $\hat{\mathscr{A}}$ is closed. Conversely, if $\mathscr{A}$ is semisimple, then Γ is injective, if $\hat{\mathscr{A}}$ is closed, then $\hat{\mathscr{A}}$ is a Banach space. By the conversing mapping theorem, Γ^{-1} is continuous and hence $\inf_{\|x\|=1} \|x\|^2 > 0$. $\square$

2.4.3 Noncommutative Algebras

Noncommutative algebras always contain commutative ones. Their presence can sometimes be exploited to extend certain theorems from the commutative situation to the noncommutative one. On a trivial level, we have already done this: In the elementary discussion of spectra, our attention was usually fixed on one element $x \in \mathscr{A}$; the (closed) subalgebra $\mathscr{A}_x$ of $\mathscr{A}$ that x generates is commutative, and much of the discussion took place within $\mathscr{A}_x$. One possible difficulty was that x might have different spectra with respect to $\mathscr{A}$ and $\mathscr{A}_x$. There is a simple construction (Theorem 2.29) that circumvents this. Another device (Theorem 2.38) can be used when $\mathscr{A}$ has an involution.

If S is a subset of a complex unital Banach algebra $\mathscr{A}$, the *centralizer*, also called the *commutator* of S is the set

$$S' = C(S) \coloneqq \{x \in \mathscr{A} : xs = sx \text{ for every } s \in S\}\,.$$

Clearly S' is not empty since it always contains the unit element. We say that S *commutes* if any two elements of S commute with each other. Clearly, S commutes if and only if

$$S \subset S'\,.$$

We shall use the following simple properties of centralizers.

Proposition 2.28. *Let S be a subset of a complex unital Banach algebra $\mathscr{A}$. Then the following statements holds:*

(a) S' *is a closed subalgebra of* $\mathscr{A}$*;*

(b) $S \subset S''$*, and if* S *commutes, then* S'' *commutes.*

Proof. Indeed, if x and y commute with every $s \in S$, so do $\lambda x, x+y$, and xy. Since multiplication is continuous in $\mathscr{A}$, S' is closed. This proves (a).

Since every $s \in S$ commutes with every $x \in S'$, we have $S \subset S''$.

If S commutes, then $S \subset S'$. Since S is a subset of S', $S'' \subset S'$. By the same reason, $S'' \subset S'''$, thus S'' commutes. (b) holds. $\square$

Theorem 2.29. *Suppose $\mathscr{A}$ is a complex unital Banach algebra, $S \subset \mathscr{A}$, S commutes, and $\mathscr{B} = S''$. Then $\mathscr{B}$ is a commutative unital Banach algebra, $S \subset \mathscr{B}$, and $\sigma_{\mathscr{B}}(x) = \sigma_{\mathscr{A}}(x)$ for ever $x \in \mathscr{B}$.*

Proof. Since $e \in \mathscr{B}$, the last proposition shows that $\mathscr{B}$ is a commutative unital Banach algebra that contains S. To show that $\sigma_{\mathscr{B}}(x) = \sigma_{\mathscr{A}}(x)$ for ever $x \in \mathscr{B}$, it suffices to show that for $\mathscr{B} \cap G(\mathscr{A}) \subset G(\mathscr{B})$.

Suppose $x \in \mathscr{B}$ and x is invertible in $\mathscr{A}$. We have to show that $x^{-1} \in \mathscr{B}$. Since $x \in \mathscr{B}$, $xy = yx$ for every $y \in S'$; hence $y = x^{-1}yx$, $yx^{-1} = x^{-1}y$. This says that $x^{-1} \in S'' = \mathscr{B}$. $\square$

We give an application of Theorem 2.29.

Proposition 2.30. *Suppose $\mathscr{A}$ is a complex unital Banach algebra, $x \in \mathscr{A}, y \in \mathscr{A}$, and $xy = yx$. Then*

$$\sigma(x+y) \subset \sigma(x) + \sigma(y) \quad \textit{and} \quad \sigma(xy) \subset \sigma(x)\sigma(y)\,.$$

Proof. To see this, put $S = \{x, y\}$; put $\mathscr{B} = S''$. Then $x+y \in \mathscr{B}, xy \in \mathscr{B}$, and Theorem 2.29 shows that we have to prove that

$$\sigma_{\mathscr{B}}(x+y) \subset \sigma_{\mathscr{B}}(x) + \sigma_{\mathscr{B}}(y) \quad \text{and} \quad \sigma_{\mathscr{B}}(xy) \subset \sigma_{\mathscr{B}}(x)\sigma_{\mathscr{B}}(y)\,.$$

Since $\mathscr{B}$ is commutative, $\sigma_{\mathscr{B}}(z)$ is the range of the Gelfand transform $\hat{z}$, for every $z \in \mathscr{B}$. (The Gelfand transforms are now functions on the maximal ideal space of $\mathscr{B}$.) Since

$$(x+y)\hat{} = \hat{x} + \hat{y} \quad \text{and} \quad (xy)\hat{} = \hat{x}\hat{y}\,.$$

we have the desired conclusion. ☐

2.5 Examples

In some cases, the maximal ideal space of a given commutative Banach algebra can easily be described explicitly. In others, extreme pathologies occur. We shall now give some examples to illustrate this. However, we begin with the following useful lemma.

Lemma 2.31. *Let X be a compact Hausdorff space and Y be a Hausdorff space. If $\Phi : X \to Y$ is a continuous bijection, then Φ is a homeomorphism.*

Proof. It remains to show that Φ^{-1} is continuous. Observe that for each F closed in X, F is compact and hence $(\Phi^{-1})^{-1}(F) = \Phi(F)$ is compact in Y hence closed. Then the desired result follows. ☐

Example 2.14 (Continuous Functions on Compact Space)**.** Let X be a compact Hausdorff space, put $C(X) = C(X;\mathbb{C})$, with the supremum norm. Then $C(X)$ becomes a commutative unital Banach algebra. Then the maximal ideal space Δ of $C(X)$ is X itself. Specifically, associate each $x \in X$ the evaluation mapping $\phi_x : \mathscr{A} \to \mathscr{A}; f \mapsto f(x)$, then clearly $\phi_x \in \Delta$. The mapping

$$\Phi : x \mapsto \phi_x$$

gives a homeomorphism of X onto Δ. As a consequence, we can regard X as the maximal ideal space of $C(X)$; and then the Gelfand representation is the identity mapping on $C(X)$.

Firstly, since $C(X)$ separates points on X (Urysohn's lemma), $x \neq y$ implies $\phi_x \neq \psi_y$ if $x \neq y$, thus Φ is injective.

Secondly, we show that Φ is surjective. If this is false, there exists a $\psi \in \Delta$ so that for each $x \in X$, there exists $f_x \in C(X)$ so that

$$\phi_x(f_x) = f_x(x) \neq \psi(f_x)\,.$$

By the continuity of f_x, we can choose an open neighborhood U_x of x so that $|f_x(y) - \psi(f_x)| > 0$ for all $y \in U_x$. Then $\{U_x : x \in X\}$ is an open cover of the compact space X. Suppose $\{U_{x_1}, \ldots, U_{x_k}\}$ is a finite subcover, then

$$f(y) = \sum_{j=1}^{k} |f_{x_j}(y) - \psi(f_{x_j})|^2 > 0 \quad \text{for all} \quad y \in X\,.$$

On the one hand, $f \in C(X)$ and hence $1/f \in C(X)$, thus f is invertible in $C(X)$. On the other hand, since

$$f = \sum_{j=1}^{k} |f_{x_j} - \psi(f_{x_j})|^2 = \sum_{j=1}^{k} \left[f_{x_j} - \psi(f_{x_j})\right] \left[\overline{f}_{x_j} - \overline{\psi(f_{x_j})}\right]$$

we have $\psi(f) = 0$, which contradicts to the fact that f is invertible. Thus Φ must be surjective.

Finally we show that Φ is a homeomorphism. To show that continuity of Φ, let (x_α) be a net in X converging to x. Then $f(x_\alpha) \to f(x)$ for each $f \in C(X)$; in other words, $\phi_{x_\alpha}(f) \to \phi_x(f)$ for each $f \in C(X)$. Thus (ϕ_{x_α}) converges to ϕ_x in Δ and hence Φ is continuous. By Lemma 2.31, Φ is a homeomorphism as claimed.

Example 2.15 (Absolutely Convergent Fourier Series)**.** Let $A(\mathbb{T})$ be the Banach algebra of all absolutely convergent trigonometric series, as in Example 2.6. Then the maximal ideal space Δ of $A(\mathbb{T})$ is $\mathbb{T}$. Specifically, for each $t \in \mathbb{T}$ let $\phi_t(f) = f(t)$ for all $f \in A(\mathbb{T})$, then clearly $\phi \in \Delta$. The mapping

$$\Phi : t \mapsto \phi_t$$

gives a homeomorphism of $\mathbb{T}$ onto Δ. As a consequence, we can regard $\mathbb{T}$ itself as its the maximal ideal space; and the Gelfand representation is the embedding mapping of $A(\mathbb{T})$ into $C(\mathbb{T})$.

As an consequence, we have

Theorem (Wiener)**.** *Suppose $f \in A(\mathbb{T})$ so that $f(t) \neq 0$ for all $t \in \mathbb{T}$. Then $1/f \in A(\mathbb{T})$, i.e., $1/f$ has absolutely convergent Fourier series.*

Example 2.16. Let $A(\mathbb{D})$ be the disc algebra of all holomorphic functions on the open disk $\mathbb{D}$ that can be extended on $\overline{\mathbb{D}}$ continuously, as in Example 2.7. Then the maximal ideal space Δ of $A(\mathbb{D})$ is $\overline{\mathbb{D}}$. Specifically, for each $z \in \overline{\mathbb{D}}$, let $\phi_z(f) = f(z)$ for all $f \in A(\mathbb{D})$, then clearly $\phi \in \Delta$. The mapping

$$\Phi : z \mapsto \phi_z$$

gives a homeomorphism of $\overline{\mathbb{D}}$ onto Δ. As a consequence, we can regard $\overline{\mathbb{D}}$ itself as its the maximal ideal space; and the Gelfand representation is the embedding mapping of $A(\mathbb{D})$ into $C(\overline{\mathbb{D}})$.

As an consequence, we have

Theorem. *Suppose $\{f_1, \cdots, f_n\} \subset A(\mathbb{D})$ is non-vanishing on $\overline{\mathbb{D}}$, that is, for each $x \in \overline{\mathbb{D}}$, al least one of $\{f_1(x), \cdots, f_n(x)\}$ is nonzero. Then there exists $\{g_1, \cdots, g_n\} \subset A(\mathbb{D})$ so that*

$$f_1 g_1 + \cdots + f_n g_n \equiv 1 .$$

Exercise 2.4. Let

$$\mathscr{A} = \left\{ f : \mathbb{Z} \to \mathbb{C} : \|f\| = \sum_{k=-\infty}^{\infty} |f(k)| 2^{|k|} < \infty \right\}$$

under the usual addition of scalar multiplication and the following multiplication

$$f * g(k) = \sum_{l=-\infty}^{\infty} f(k-l) g(l)$$

Show that

(a) $\mathscr{A}$ is a commutative Banach algebra;

(b) Let $K = \left\{z \in \mathbb{C} : \frac{1}{2} \leq |z| \leq 2\right\}$ then K is one-to-one correspondent with Δ and the Gelfand representation of $\mathscr{A}$ is the Laurent series that are absolutely convergent on K.

2.6 Involution and C*-Algebra

Definition 2.6. A mapping $x \mapsto x^*$ of a complex (not necessarily commutative) algebra $\mathscr{A}$ into $\mathscr{A}$ is called an *involution* if it is a *conjugate-linear anti-automorphism of period 2*. In other words, it has the following properties: for all $x, y \in \mathscr{A}$, and $\lambda \in \mathbb{C}$:

$$(x+y)^* = x^* + y^* \ , \ (\lambda x)^* = \bar{\lambda} x^* \, ;$$
$$(xy)^* = y^* x^* \ ; \ x^{**} = x \, .$$

We say $x \in \mathscr{A}$ is *hermitian*, or *self-adjoint* if$x^* = x$.

There are two classical example of the Banach algebra with an involution.

Example 2.17. Let X be a compact Hausdorff space and $\mathscr{A} = C(X, \mathbb{C})$, then

$$f \to \bar{f}$$

is an involution. Clearly f is hermitian iff f is real-valued.

Example 2.18. Let H be a Hilbert space and $\mathscr{A} = \mathcal{B}(H)$, then

$$A \mapsto A^*$$

where A^* is the adjoint operator of $\mathscr{A}$, is an involution on $\mathcal{B}(H)$. Clearly A is hermitian if and only if A is a self-adjoint operator. In practice, we will be most concerned with this example.

The following lemma is easy so we omit the proof.

Lemma 2.32. *If $\mathscr{A}$ is a Banach algebra with an involution, and if $x \in \mathscr{A}$, then*

(a) $x + x^*, i(x - x^*)$, *and xx^* are hermitian;*

(b) *x has a unique representation $x = u + iv$, with $u, v \in \mathscr{A}$ and both u and v hermitian;*

(c) *the unit e is hermitian;*

(d) *x is invertible in $\mathscr{A}$ if and only if x^* is, in which case $(x^*)^{-1} = \left(x^{-1}\right)^*$; and*

(e) *$\lambda \in \sigma(x)$ if and only if $\bar{\lambda} \in \sigma(x^*)$.*

Theorem 2.33. *Let $\mathscr{A}$ be a semisimple commutative unital Banach algebra. Then every involution on $\mathscr{A}$ is continuous.*

Proof. Although the involution is conjugate-linear, the closed graph theorem holds in this case. So it suffices to show that if $x_n \to x$ and $x_n^* \to y$ in $\mathscr{A}$, then $y = x^*$. Since $\mathscr{A}$ is commutative and semisimple, it suffices to show that

$$\phi(x^*) = \phi(y) \quad \text{for all} \quad \phi \in \Delta\,.$$

Clearly $\phi(y) = \lim_n \phi(x_n^*)$. Note that if we define φ by $\varphi(x) = \overline{\phi(x^*)}$, then $\varphi \in \Delta$. Thus $\varphi(x_n) = \overline{\phi(x_n^*)} \to \varphi(x) = \overline{\phi(x^*)}$, and hence $\phi(x^*) = \lim_n \phi(x_n^*)$. Then we get $\phi(x^*) = \phi(y)$ as desired. □

Definition 2.7. A complex unital Banach algebra $\mathscr{A}$ with an involution $x \mapsto x^*$ that satisfies

$$\|xx^*\| = \|x\|^2 \quad \text{for all } x \in \mathscr{A} \tag{2.7}$$

is called a B^*-*algebra.* The identity (2.7) is called the B^*-condition.

If $x \in \mathscr{A}$ is hermitian, then we have $\|x^2\| = \|x\|^2$.

Note that $\|x\|^2 = \|xx^*\| \leq \|x\| \, \|x^*\|$ implies $\|x\| \leq \|x^*\|$, hence also

$$\|x^*\| \leq \|x^{**}\| = \|x\|$$

Thus

$$\|x^*\| = \|x\| \tag{2.8}$$

in every B^* -algebra. It also follows that

$$\|xx^*\| = \|x\| \, \|x^*\| \qquad \text{for all } x \in \mathscr{A} \, . \tag{2.9}$$

Identity (2.9) is often called the *C^*-condition*. Complex unital Banach algebras with an involution that satisfy the C^*-condition (2.9) is called a *C^*-algebra*. It's easy to see that (2.8) and (2.9) implies (2.7).

Remark 2.13. The term B^*-algebra was introduced by C. E. Rickart in 1946 to describe Banach algebras with an involution. Clearly a B*-algebra is also a C*-algebra. Conversely, the C^*-condition implies the B^*-condition. This is nontrivial, and can be proved without using the condition (2.8). For these reasons, the term B^*-algebra is rarely used in current terminology, and has been replaced by the term 'C^*-algebra'. The term C*-algebra was introduced by I. E. Segal in 1947 to describe norm-closed subalgebras of $\mathcal{B}(H)$, namely, the space of bounded operators on some Hilbert space H. 'C' stood for 'closed'. In his paper Segal defines a C^*-algebra as a "uniformly closed, self-adjoint algebra of bounded operators on a Hilbert space".

Definition 2.8. Let $\mathscr{A}$ and $\mathscr{B}$ be two complex algebra with involutions. Then $\psi : \mathscr{A} \to \mathscr{B}$ is called a *$*$-isomorphism*, if ψ is an algebra isomorphism preserving involution, i.e., $\psi(x^*) = \psi(x)^*$ for all $x \in \mathscr{A}$.

2.6.1 Commutative C*-algebra

It's easy to see that the algebra $C(X)$, consisting of complex-valued continuous functions on compact Hausdorff space X, is a commutative C^*

algebra. Surprisingly, the converse is true. This is the key to the proof of the spectral theorem that will be given later.

We need the following lemma, which does not depend on the commutativity conditions.

Lemma 2.34. *Let $\mathscr{A}$ be a C^*-algebra and $x \in \mathscr{A}$ is hermitian. Then*

$$\phi(x) \in \mathbb{R} \quad \textit{for all} \quad \phi \in \Delta\,.$$

Proof. We shall show that $\operatorname{Im}(\phi(x)) = 0$. Consider $x + ite$, where $t \in \mathbb{R}$. Since

$$|\phi(x+ite)| = |\phi(x) + it| \leq \|x + ite\|\,,$$

and $(x+ite)^* = x - ite$, by C^*-condition we have

$$\begin{aligned} |\phi(x) + it|^2 &= |\operatorname{Re}(\phi(x))|^2 + |\operatorname{Im}(\phi(x)) + t|^2 \\ &= |\operatorname{Re}(\phi(x))|^2 + |\operatorname{Im}(\phi(x))|^2 + t^2 + 2\operatorname{Im}(\phi(x))t \\ &\leq \|x+ite\|^2 = \|(x+ite)(x-ite)\| \leq \|x^2\| + t^2\,. \end{aligned}$$

Then

$$|\operatorname{Re}(\phi(x))|^2 + |\operatorname{Im}(\phi(x))|^2 + 2\operatorname{Im}(\phi(x))t \leq \|x^2\|\,.$$

Since $t \in \mathbb{R}$ is arbitrary, it must be $\operatorname{Im}(\phi(x)) = 0$ as required. ☐

Theorem 2.35 (Gelfand-Naimark)**.** *Let $\mathscr{A}$ be a commutative C^*-algebra. Let Δ be all the nonzero complex homomorphisms on X. Then the Gelfand representation Γ is an isometric $*$-isomorphism of $\mathscr{A}$ onto $C(\Delta)$.*

Proof. It suffices to show that (i) Γ preserves the involution; (ii) Γ is an isometry, and; (iii) $\Gamma : \mathscr{A} \to C(\Delta)$ is surjective.

(i). Since the involution on $C(\Delta)$ is the conjugation, to show that $\Gamma(x^*) = \overline{\Gamma(x)}$, we have to show

$$\phi(x^*) = \overline{\phi(x)}\,.$$

Let $x = u + iv$ with u and v hermitian, then since $\phi(u)$ and $\phi(v)$ are real number

$$\begin{aligned}\phi(x^*) &= \phi(u - iv) = \phi(u) - i\phi(v)\\ &= \overline{\phi(u) + i\phi(v)} = \overline{\phi(x)}\,,\end{aligned}$$

as desired.

(ii). In order that Γ is an isometry, we shall show that for each $x \in \mathscr{A}$,

$$\|x^2\| = \|x\|^2\,.$$

Since x and x^* commutes, and x^*x is hermitian, we have

$$\begin{aligned}\|x^2\|^2 &= \|(x^2)^*x^2\| = \|(x^*)^2x^2\|\\ &= \|(x^*x)^2\| = \|x^*x\|^2 = \|x\|^4\,,\end{aligned}$$

and then the desired result follows.

(iii). Since we have shown Γ is an isometry, then $\hat{\mathscr{A}}$ is a closed subalgebra in $C(\Delta)$. So it suffices to show that $\hat{\mathscr{A}}$ is dense in $C(\Delta)$. We will prove this by Theorem **??**. Clearly the constant function $1 = \Gamma(e) \in \hat{\mathscr{A}}$, thus $\hat{\mathscr{A}}$ vanishes nowhere. Since $\overline{\Gamma(x)} = \Gamma(x^*)$, $\hat{\mathscr{A}}$ is closed under complex conjugation. For $\phi_1 \neq \phi_2$, there exists $x \in \mathscr{A}$ so that $\phi(x_1) \neq \phi(x_2)$, thus $\Gamma(x_1)(\phi) \neq \Gamma(x_2)(\phi)$ and hence $\hat{\mathscr{A}}$ separates the points. Thus by the Stone-Weierstrass theorem (see Theorem **??**), $\hat{\mathscr{A}}$ is dense in $C(\Delta)$. $\square$

2.6.2 Noncommutative C*-algebra

Clearly, $\mathcal{B}(H)$ is a C^*-algebra. Besides, every closed subalgebra of $\mathcal{B}(H)$ which is closed under involution is also a C^*-algebra. Surprisingly, the converse is true. This is a deep result and we will not show it.

Theorem (Gelfand-Naimark)**.** *Let $\mathscr{A}$ be a C^*-algebra. Then there exists a closed unital subalgebra $\mathscr{B}$ of $\mathcal{B}(H)$, which is also closed under $*$ operation, i.e., $\mathscr{B}^* = \mathscr{B}$, so that $\mathscr{A}$ isometrically $*$-isomorphic to $\mathscr{B}$.*

Let $\mathscr{A}$ be a C^*-algebra. Now we concern the following question as before. To use the result for the commutative case, we want to find some commutative sub-C^*-algebra $\mathscr{B}$ of $\mathscr{A}$ so that for $x \in \mathscr{B}$ there holds $\sigma_{\mathscr{B}}(x) = \sigma_{\mathscr{A}}(x)$.

Definition 2.9. Let $\mathscr{A}$ be a C^*-algebra. $x \in \mathscr{A}$ is said to be *normal* if $xx^* = x^*x$. A set $S \subset \mathscr{A}$ is said to be *normal* if S commutes and if $x^* \in S$ whenever $x \in S$.

Clearly, a closed unital subalgebra $\mathscr{B}$ of $\mathscr{A}$ is a commutative C^*-algebra if and only if it is normal.

Since the union of a family of totally ordered normal subsets is still normal, by Zorn's lemma, every normal subset S of $\mathscr{A}$ is contained in a maximal one. The following theorem asserts that the maximal normal subset of $\mathscr{A}$ satisfies our requests.

Theorem 2.36. *Let $\mathscr{A}$ be a C^*-algebra, and $\mathscr{B}$ is a maximal normal subset of $\mathscr{A}$. Then $\mathscr{B}$ is a commutative sub-C^*-algebra of $\mathscr{A}$, and*

$$\sigma_{\mathscr{B}}(x) = \sigma_{\mathscr{A}}(x) \quad \text{for every } x \in \mathscr{B}\,.$$

Proof. We begin with a simple criterion for membership in $\mathscr{B}$:

If $x \in \mathscr{A}$ is normal and if x commutes with $\mathscr{B}$, i.e., $xy = yx$ for all $y \in \mathscr{B}$, then $x \in \mathscr{B}$.

Indeed, if x satisfies these conditions, we also have $xy^* = y^*x$ for all $y \in \mathscr{B}$, since $\mathscr{B}$ is normal, and therefore $x^*y = yx^*$. It follows that $\mathscr{B} \cup \{x, x^*\}$ is normal. Hence $x \in \mathscr{B}$, since $\mathscr{B}$ is maximal.

This criterion makes it clear that sums and products of members of $\mathscr{B}$ are in $\mathscr{B}$, the unit element is in $\mathscr{B}$. Thus $\mathscr{B}$ is a commutative unital algebra.

Suppose $x_n \in \mathscr{B}$ and $x_n \to x$. since $x_n y = y x_n$ for all $y \in \mathscr{B}$, and multiplication is continuous, we have $xy = yx$ and therefore also

$$x^*y = (y^*x)^* = (xy^*)^* = yx^*$$

In particular, $x^*x_n = x_nx^*$ for all n, which leads to $x^*x = xx^*$. Hence $x \in \mathscr{B}$, by the above criterion. This proves that $\mathscr{B}$ is closed.

To prove $\sigma_{\mathscr{B}}(x) = \sigma_{\mathscr{A}}(x)$ for every $x \in \mathscr{B}$, it suffices to show that $G(\mathscr{B}) = \mathscr{B} \cap G(\mathscr{A})$. Assume $x \in \mathscr{B}, x^{-1} \in \mathscr{A}$. since x is normal, so is x^{-1}, and since x commutes with every $y \in \mathscr{B}$, so does x^{-1}. Hence $x^{-1} \in \mathscr{B}$. $\square$

Our next application of Theorem 2.36 will extend some consequences of Theorem 2.35 to arbitrary (not necessarily commutative) C^* algebras.

In a C^*-algebra, the statement "$x \geq 0$" means that $x = x^*$ and that $\sigma(x) \subset [0, \infty)$.

Proposition 2.37. *Let $\mathscr{A}$ be a C^* algebra. Then the following holds.*

(a) *If $x \in \mathscr{A}$ is Hermitian, then $\sigma_{\mathscr{A}}(x) \subset \mathbb{R}$.*

(b) *If $x \in \mathscr{A}$ is normal, then $r_\sigma(x) = \|x\|$. In particular, $r_\sigma(y^*y) = \|y\|^2$ for all $y \in \mathscr{A}$.*

(c) *If $u, v \in \mathscr{A}$, $u, v \geq 0$, then $u + v \geq 0$.*

(d) *If $y \in \mathscr{A}$, then $y^*y \geq 0$. In particularly, $e + y^*y$ is invertible in $\mathscr{A}$.*

Proof of (a), (b). Every normal $x \in \mathscr{A}$ lies in a maximal normal set $\mathscr{B} \subset \mathscr{A}$. By the preceding theorem, $\mathscr{B}$ is a commutative C^*-algebra which is isometrically $*$-isomorphic to its Gelfand transform $\hat{\mathscr{B}} = C(\Delta)$ and which has the property that

$$\sigma(z) = \hat{z}(\Delta) \text{ for } z \in \mathscr{B}. \tag{2.10}$$

Here $\sigma(z)$ is the spectrum of z relative to $\mathscr{A}$, Δ is the maximal ideal space of $\mathscr{B}$, and $\hat{z}(\Delta)$ is the range of the Gelfand transform of z, regarded as an element of $\mathscr{B}$.

If $x = x^*$, Lemma 2.34 shows that $\hat{x}$ is a real-valued function on Δ. Hence (2.38) implies (a).

For any normal x, (2.38) implies $r_\sigma(x) = \|\hat{x}\|_{C(\Delta)}$. Also, $\|\hat{x}\|_{C(\Delta)} = \|x\|$, since $\mathscr{B}$ and $\hat{\mathscr{B}}$ are isometric. Thus $r_\sigma(x) = \|x\|$. If $y \in \mathscr{A}$, then yy^* is hermitian, and hence $r_\sigma(yy^*) = \|yy^*\| = \|y\|^2$. This proves (b). $\square$

We can not use the previous argument to show (c), (d) since we can not say that u and v commute or y and y^* commute.

Proof of (c). Since $\sigma(u) \subset [0, \|u\|]$, so that

$$\sigma(\|u\|e - u) \subset [0, \|u\|]$$

and this implies therefore that $\big\|\|u\|e - u\big\| \le \|u\|$. For the same reason, $\big\|\|v\|e - v\big\| \le \|v\|$. Hence

$$\big\|\|u\|e + \|v\|e - (u+v)\big\| \le \|u\| + \|v\|\,.$$

since $\|u\|e + \|v\|e - (u+v)$ is hermitian, (a) implies that its spectrum is a subset of $\mathbb{R}$. Thus

$$\sigma(\|u\|e + \|v\|e - (u+v)) \subset [-\|u\| - \|v\|, \|u\| + \|v\|]\,.$$

Therefore,

$$\sigma(u+v) \subset [0, 2\|u\| + 2\|v\|]\,,$$

which deduce that $u + v \ge 0$. $\square$

Proof of (d). We turn to the proof of (d). Put $x = y^*y$. Then x is hermitian, and if $\mathscr{B}$ is chosen as in the proof of (a),(b), then $\hat{x}$ is a real-valued function on Δ. By (2.38), we have to show that $\hat{x} \ge 0$ on Δ.

We begin with a classic mistake: For each $\phi \in \Delta$,

$$\hat{x}(\phi) = \phi(x) = \phi(y^*y) = \phi(y^*)\phi(y) = \overline{\phi(y)}\phi(y) = |\phi(y)|^2 \ge 0\,.$$

This is wrong since ϕ is a complex homomorphism on $\mathscr{B}$, and we can not say $y, y^* \in \mathscr{B}$. Thus $\phi(y)$ and $\phi(y^*)$ may not make sense and $\phi(y^*y) = \phi(y^*)\phi(y)$, $\phi(y^*) = \overline{\phi(y)}$ may not hold.

Next, we give the proof. Since $\hat{\mathscr{B}} = C(\Delta)$, there exists $z \in \mathscr{B}$ such that

$$\hat{z} = |\hat{x}| - \hat{x} \quad \text{on } \Delta\,. \tag{2.11}$$

Note that $\hat{z}$ is real, then $z = z^*$. Put

$$yz = w = u + iv\,,$$

where u and v are hermitian elements of $\mathscr{A}$. Then

$$w^*w = z^*y^*yz = zxz = z^2x \tag{2.12}$$

and therefore

$$ww^* = 2u^2 + 2v^2 - w^*w = 2u^2 + 2v^2 - z^2x\,. \tag{2.13}$$

Since $u = u^*$, $\sigma(u)$ is real, by (a), hence $u^2 \geq 0$, by the spectral mapping theorem. Likewise $v^2 \geq 0$. By (2.11), $\hat{z}^2\hat{x} \leq 0$ on Δ. Since $z^2x \in \mathscr{B}$, it follows from that $-z^2x \geq 0$. Now (2.13) and (c) imply that

$$ww^* \geq 0\,.$$

But $\sigma(w^*w) \subset \sigma(ww^*) \cup \{0\}$ (Exercise 2.1). Hence $w^*w \geq 0$. By (2.12), this means that $\hat{z}^2\hat{x} \geq 0$ on Δ. By (2.11), this last inequality holds only when $\hat{x} = |\hat{x}|$. Thus $\hat{x} \geq 0$ as desired. □

Remark 2.14. In fact, (d) is a key step in the proof of the Gelfand-Naimark theorem. On the other hand, if we admit the Gelfand-Naimark theorem, (d) is a trivial consequence.

Equality of spectra can now be proved in yet another situation, in which commutativity plays no role.

Theorem 2.38. *Let $\mathscr{A}$ be a C^*-algebra. Let $\mathscr{B}$ be a sub-C^*-algebra of $\mathscr{A}$. Then*

$$\sigma_{\mathscr{B}}(x) = \sigma_{\mathscr{A}}(x) \quad \textit{for every } x \in \mathscr{B}\,.$$

Proof. We shall show that $G(\mathscr{B}) = \mathscr{B} \cap G(\mathscr{A})$. Suppose $x \in \mathscr{B}$ and x has an inverse in $\mathscr{A}$. We have to show that $x^{-1} \in \mathscr{B}$.

Since x is invertible in $\mathscr{A}$, so is x^*, hence also x^*x, and therefore $0 \notin \sigma_{\mathscr{A}}(x^*x)$. Observe that x^*x is hermitian, by Proposition 2.37, $\sigma_{\mathscr{A}}(x^*x) \subset \mathbb{R}$, so that $\varrho_{\mathscr{A}}(x^*x)$ is connected in $\mathbb{C}$. Then Theorem 2.9 shows now that $\sigma_{\mathscr{B}}(x^*x) = \sigma_{\mathscr{A}}(x^*x)$. Hence $(x^*x)^{-1} \in \mathscr{B}$. Finally since $x^* \in \mathscr{B}$,

$$x^{-1} = (x^*x)^{-1}\, x^* \in \mathscr{B}$$

as desired. $\square$

2.6.3 Continuous Functional Calculus for Normal Operators

The construction of the continuous functional calculus for normal operators is based on several lemmas. Assume throughout that H is a complex Hilbert space and that $N \in \mathcal{B}(H)$ is a normal operator. Let $\mathscr{A}_N \subset \mathcal{B}(H)$ be the smallest sub-C^*-algebra that contains N; i.e.,

$$\mathscr{A}_N = \mathrm{cl}\left\{p(N, N^*) : p \text{ is a polynomial in two variables}\right\}. \tag{2.14}$$

Clearly, $\mathscr{A}_N$ is a commutative sub-C^*-algebra of $\mathcal{B}(H)$, the spectrum of N relative to $\mathcal{B}(H)$ and $\mathscr{A}_N$ coincide by Theorem 2.38. Therefore we always denote it by $\sigma(N)$.

Theorem 2.39. *Let the commutative C^*-algebra $\mathscr{A}_N$ be given above. Let Δ be all the nonzero complex homomorphisms of $\mathscr{A}_N$. Let $\hat{N}$ be the Gelfand representation of N. Then the following hold.*

(a) *$\hat{N}$ is a homeomorphism of Δ onto $\sigma(N)$.*

(b) *For each $f \in C(\sigma(N))$, let $\tilde{f}(N) \in \mathscr{A}_N \subset \mathcal{B}(H)$ satisfy $\phi(\tilde{f}(N)) = f(\phi(N))$ for all $\phi \in \Delta$. Then the mapping*

$$f \mapsto \tilde{f}(N)\ ;\ C(\sigma(N)) \to \mathscr{A}_N$$

is an isometric $*$-isomorphism.

Since $\mathscr{A}_N$ is a commutative C^*-algebra, the Gelfand representation $\Gamma : \mathscr{A}_N \to \Delta$ is a isomorphic $*$-isomorphism. Thus $\tilde{f}(N)$ is well defined. Indeed,

$$\tilde{f}(N) = \Gamma^{-1}(f(\Gamma(N)))$$

Proof. As we know, $\hat{N}$ is a continuous function on Δ whose range is $\sigma(N)$, i.e., $\hat{N}$ is a continuous surjection of Δ onto $\sigma(N)$.

To see that $\hat{N}$ is injective, suppose $\phi_1 \in \Delta, \phi_2 \in \Delta$, and $\hat{N}(\phi_1) = \hat{N}(\phi_2)$, that is, $\phi_1(N) = \phi_2(N)$. Then that $\phi_1(N^*) = \overline{\phi_1(N)} = \overline{\phi_2(N)} = \phi_2(N^*)$. If p is any polynomial in two variables, it follows that

$$\phi_1(p(N, N^*)) = \phi_2(p(N, N^*)) ,$$

since ϕ_1 and ϕ_2 are homomorphisms. Since elements of the form $p(N, N^*)$ are dense in $\mathscr{A}_N$, the continuity of ϕ_1 and ϕ_2 implies therefore $\phi_1 = \phi_2$. We have proved that $\hat{N}$ is one-to-one.

Since Δ and $\sigma(N)$ both are compact, it follows that $\hat{N}$ is a homeomorphism of Δ onto $\sigma(N)$, and (a) follows.

The mapping $f \to f \circ \hat{N}$ is therefore an isometric $*$-isomorphism of $C(\sigma(N))$ onto $C(\Delta)$. Since the inverse of the Gelfand representation Γ^{-1} is also an isometric $*$-isomorphism of $C(\Delta)$ onto $\mathscr{A}_N$, then the mapping $f \to \Gamma^{-1} \circ f \circ \hat{N}$ is therefore an isometric $*$-isomorphism of $C(\sigma(N))$ onto $\mathscr{A}_N$, as desired. □

With these preparations in place we are ready to establish the continuous functional calculus for normal operators on Hilbert spaces.

Theorem 2.40 (Continuous Functional Calculus)**.** *Let H be a complex Hilbert space, let $N \in \mathcal{B}(H)$ be a bounded normal operator. Then the isometric $*$-isomorphism*

$$f \mapsto \tilde{f}(N) \; ; \; C(\sigma(N)) \to \mathscr{A}_N \subset \mathcal{B}(H) \tag{2.15}$$

satisfies the following axioms.

(a) *(Normalization) If $f(\lambda) = \lambda$ for all $\lambda \in \sigma(N)$ then $\tilde{f}(N) = N$.*

(b) *(Commutative) If $A \in \mathcal{B}(H)$ satisfies $NA = AN$ and $N^*A = AN^*$ then $\tilde{f}(N)A = A\tilde{f}(N)$ for all $f \in C(\sigma(N))$.*

(c) *(Eigenvector) If $\lambda \in \sigma_p(N)$ and $x \in H$ satisfy $Nx = \lambda x$ then*

$$\tilde{f}(N)x = f(\lambda)x \quad \textit{for all } f \in C(\sigma(N)) .$$

(d) *(Spectrum) For every $f \in C(\sigma(N))$ the operator $\tilde{f}(N)$ is normal and*

$$\sigma(\tilde{f}(N)) = f(\sigma(N)) .$$

(e) *(Positive) If $f \in C(\sigma(N), \mathbb{R})$ and $f \geq 0$ then $\tilde{f}(N) = \tilde{f}(N)^* \geq 0$.*

(f) *(Composition) If $f \in C(\sigma(N))$ and $g \in C(f(\sigma(N)))$ then*

$$\widetilde{(g \circ f)}(N) = \tilde{g}(\tilde{f}(N)) .$$

Proof. We prove part (a). If $f(\lambda) = \lambda$ for all $\lambda \in \sigma(N)$, then $f \circ \hat{N} = \hat{N}$. Thus $\tilde{f}(N) = N$.

We prove part (b). In fact, by the definition of $\mathscr{A}_N$, if $AN = NA$ then A commutes with $p(N, N^*)$ for each polynomial p in two variables. By the continuity of multiplication, A commutes with $\mathscr{A}_N$ and hence (b) follows.

We prove part (c). Let $\lambda \in \sigma_p(N)$ and $x \in H$ such that $Nx = \lambda x$. Then $N^*x = \bar{\lambda}x$ by Theorem **??**. Hence $p(N, N^*)\, x = p(\lambda, \overline{\lambda})x$ for every polynomial p in z and $\bar{z}$. By the Stone-Weierstrass theorem, $\{p(z, \bar{z}) : p \text{ is a polynomial in two variables}\}$ is dense in $C(\sigma(N))$, using $f \mapsto \tilde{f}(N)$ is isometric we have $\tilde{f}(N)x = f(\lambda)x$, as required.

We prove part (d). By the Gelfand representation, letting Δ be all the nonzero complex homomorphism on $\mathscr{A}_N$, then

$$\begin{aligned}\sigma(\tilde{f}(N)) &= \left\{ \widehat{\tilde{f}(N)}(\phi) : \phi \in \Delta \right\} = \left\{ f(\hat{N}(\phi)) : \phi \in \Delta \right\} \\ &= \{f(\phi(N)) : \phi \in \Delta\} = \{f(\lambda) : \lambda \in \sigma(N)\} = f(\sigma(N)) .\end{aligned}$$

(e) follows directly from (d) and the fact that $f \mapsto f(N)$ preserve the involution.

We prove part (f). By the definition

$$\Big(\widetilde{(g \circ f)}(N)\Big)\,\hat{} = g \circ f \circ \hat{N}\,;$$

and

$$\Big(\tilde{g}(\tilde{f}(N))\Big)\,\hat{} = g \circ (\widehat{\tilde{f}(N)}) = g \circ f \circ \hat{N}\,;$$

thus (f) holds. □

Remark 2.15. Let H be an infinite-dimensional complex Hilbert space. It is useful to examine the special case of Theorem 2.40 where the normal operator $N \in \mathcal{B}(H)$ is compact, which we now assume. By Theorem **??**, there are complex numbers $\{\lambda_n\}$ projections $\{P_n\}$ with $P_n P_m = 0$ if $n \neq m$ so that

$$A = \text{s-}\sum_n \lambda_n P_n\,.$$

Then it's easy to see that for a polynomial p in two variables

$$p(A, A^*) = \text{s-}\sum_n p(\lambda_n, \bar{\lambda}_n) P_n\,.$$

Then by the Stone-Weierstrass theorem, $\{p(z, \bar{z}) : p$ is a polynomial in two variables$\}$ is dense in $C(\sigma(N))$, using $f \mapsto \tilde{f}(N)$ is isometric we have

$$f(A) = \text{s-}\sum_n f(\lambda_n) P_n\,.$$

We end this section with a commutativity theorem, as an application of the continuous functional calculus.

Theorem 2.41. *Let $N \in \mathscr{B}(H)$ be normal. Then for each $T \in \mathscr{B}(H)$ commutes with N, T must commutes with N^*.*

Proof. Since $NT = TN$ holds, then by induction we have $N^k T = TN^k$ for integer $k \geq 0$. Hence

$$\exp(N)T = T\exp(N)\,.$$

Thanks to the continuous functional calculus $\exp(N)$ has inverse $\exp(-N)$. Thus

$$T = \exp(-N)T\exp(N)\,.$$

By the continuous functional calculus, we have

$$\exp\left(N^*\right)T\exp\left(-N^*\right) = \exp(N^* - N)T\exp(N - N^*)$$

Since $|e^{z-\bar{z}}| = 1$ for all $z \in \sigma(N)$, by the continuous functional calculus, we have

$$\|\exp(N^* - N)\| = \|\exp(N^* - N)\| = 1\,.$$

thus

$$\|\exp\left(N^*\right)T\exp\left(-N^*\right)\| \leq \|T\|\,.$$

We now define

$$f(\lambda) = \exp\left(\lambda N^*\right)T\exp\left(-\lambda N^*\right) \text{ for } \lambda \in \mathbb{C}\,.$$

Then f is an analytic $\mathcal{B}(H)$-valued function, since $\lambda \mapsto \exp\left(\lambda N^*\right)$ and $\lambda \mapsto \exp\left(-\lambda N^*\right)$ are. The hypotheses of the theorem hold with $\bar{\lambda}N$ in place of N. Therefore we have

$$\|f(\lambda)\| \leq \|T\| \text{ for all } \lambda \in \mathbb{C}\,.$$

Then f is a bounded entire function. By Liouville's theorem, f is a constant function, so $f(\lambda) = f(0) = T$, for every $\lambda \in \mathbb{C}$. Hence

$$\exp\left(\lambda N^*\right)T = T\exp\left(\lambda N^*\right) \text{ for all } \lambda \in \mathbb{C}\,.$$

Compute the derivative of the both side, we obtain $N^*T = TN^*$. □

Remark 2.16. Following this result, commutative axiom (part (b)) of Theorem 2.40 can be reduced as following: If $A \in \mathcal{B}(H)$ satisfies $NA = AN$, then $\tilde{f}(N)A = A\tilde{f}(N)$ for all $f \in C(\sigma(N))$.

In fact, a more general result is true:

Exercise 2.5 (Fuglede-Putnam-Rosenblum). Assume that $M, N, T \in \mathscr{B}(H)$, M and N are normal, and $MT = TN$, then $M^*T = TN^*$.

2.7 Spectral Decomposition of Normal Operators on Hilbert Space

[Spectral Decomposition of Normal Operators]

Let H be a complex Hilbert space. The main results of this section are the following. Firstly, if $\mathscr{A}$ is a commutative sub-C^*-algebra of $\mathcal{B}(H)$, we can give a specific formula of the inverse of the Gelfand transform.

The second is a important consequence of the first one. If $N \in \mathcal{B}(H)$ is normal, then there exists complex Borel measures $\{E_{x,y} : x, y \in H\}$ so that

$$\langle Nx, y\rangle = \int_{\sigma(N)} \lambda \, E_{x,y}(d\lambda)\,.$$

Moreover, $\{E_{x,y} : x, y \in H\}$ is very special. To make it clear, we begin some preliminaries.

2.7.1 Projection Valued Measures

Denote by $\mathcal{P}(H)$ all the orthogonal projections on H. Let $(\Omega, \mathcal{F})$ be a measurable space. Indeed we always assume that Ω is a Hausdorff space and $\mathcal{F}$ is the Borel algebra on Ω.

Definition 2.10. A *projection valued measure* P on $(\Omega, \mathcal{F})$ is a map $\mathcal{F} \to \mathcal{P}(H)$ which assigns to every measurable set A a projection $P(A)$ on H and satisfies the following axioms.

(a) (Normalization) $P(\emptyset) = 0,\ P(\Omega) = I$.

(b) (Intersection) If $A_1,\ A_2$ are in $\mathcal{F}$ then

$$P(A_1 \cap A_2) = P(A_1)P(A_2)\,.$$

(c) (σ-Additive) If $\{A_n\}$ is a sequence of pairwise disjoint sets in $\mathcal{F}$ then

$$P\big(\bigcup_{n=1}^{\infty} A_n\big) = \text{s-}\sum_{n=1}^{\infty} P(A_n)\,.$$

Remark 2.17. One should note that (b) implies that $P(A_1)$ and $P(A_2)$ commute for every $A_1,\ A_2$ in $\mathcal{F}$. Moreover, in fact (a) and (c) implies (b).

First we note that $P\,(A_1)\,P\,(A_2) = 0$ if $A_1, A_2 \in \mathcal{F}$ are disjoint. Indeed since $P\,(A_1) + P\,(A_2) = P(A_1 \cup A_2)$ is a projection. By Proposition **??** we have $P\,(A_1)\,P\,(A_2) = 0$. Now for each A_1, A_2 in $\mathcal{F}$, we have

$$\begin{aligned} P(A_1)P(A_2) &= [P(A_1 \cap A_2) + P(A_1 \backslash A_2)]\,[P(A_1 \cap A_2) + P(A_2 \backslash A_1)] \\ &= P(A_1 \cap A_2)^2 = P(A_1 \cap A_2)\,. \end{aligned}$$

Remark 2.18. In fact, if the mapping $P : \mathcal{F} \to \mathcal{P}(H)$ satisfies (a) and (b), then (c) can be replaced by

(c$'$) For all $x, y \in H,\ A \mapsto \langle P(A)x, y\rangle$ is a *complex measure* on $(\Omega, \mathcal{F})$.

To see this, trivially, (a) and (c) implies (c$'$). On the contrary, by part (b), if $\{A_n\} \subset \mathcal{F}$ is pairwise disjoint, then $\{P(A_n)x\}$ is pairwise orthogonal. Thus the series

$$\sum_{n=1}^{\infty} P(A_n)x$$

converges absolutely. By part (c$'$),

$$\langle\sum_{n=1}^{\infty} P(A_n)x, y\rangle = \sum_{n=1}^{\infty}\langle P(A_n)x, y\rangle = \langle P\big(\bigcup_{n=1}^{\infty} A_n\big)x, y\rangle$$

for all $y \in H$. Thus

$$P(\bigcup_{n=1}^{\infty} A_n)x = \sum_{n=1}^{\infty} P(A_n)x\,.$$

Since $x \in H$ is arbitrary, (c) follows.

Henceforth, for a projection valued measure P, we denote by $P_{x,y}$ the complex measure $A \mapsto \langle P(A)x, y\rangle$. Then clearly the following holds.

(a) For each $x \in H$, $P_{x,x}$ is a finite positive measure on $(\Omega, \mathcal{F})$; for $A \in \mathcal{F}$, we have $P_{x,x}(A) = \|P(A)x\|^2$. So if $\|x\| = 1$, then $P_{x,x}$ is a probability measure on $(\Omega, \mathcal{F})$.

(b) For each $x, y \in H$, and $A \in \mathcal{F}$, we have

$$|P_{x,y}(A)| \leq \|y\||P_{x,x}(A)|^{1/2}\,.$$

Exercise 2.6. Let P be a projection valued measure on $(\Omega, \mathcal{F})$. Let $A \in \mathcal{F}$. Then the following statements are equivalent.

(a) $P(A) = 0$.

(b) $P_{x,y}(A) = 0$ for all x, y in H.

(c) $|P_{x,y}|(A) = 0$ for all x, y in H, where $|P_{x,y}|$ is the total variance of $P_{x,y}$.

(d) $P_{x,x}(A) = 0$ for all x in H.

The Essential Bounded Functions Next, we will define the integrals of measurable functions on the projection-valued measure space $(\Omega, \mathcal{F}, P)$. To this end, we need some preliminaries.

Let f be a measurable, complex-valued function on $(\Omega, \mathcal{F}, P)$. Then the *essential range* of f is defined to be the set:

$$\text{ess.\,im}(f) = \{z \in \mathbb{C} : \text{ for all } r > 0 : P(f \in B(z, r)) \neq 0\}\,.$$

In other words $\text{ess.\,im}(f) = \text{supp}(P \circ f^{-1})$. Thus, one can easily check that:

(a) The complement of $\operatorname{ess.im}(f)$, given by

$$\{z \in \mathbb{C} : \text{ there exists } r > 0, P(f \in B(z,r)) = 0\},$$

is open, and since $\mathbb{C}$ is second countable, $P(f \in \operatorname{ess.im}(f)^c) = 0$.

(b) The essential image $\operatorname{ess.im}(f)$ is always closed.

(c) If B is a Borel set in $\mathbb{C}$ disjoint with $\operatorname{ess.im}(f)$, then $P(f \in B) = 0$. This fact characterises the essential image: It is the smallest closed subset of $\mathbb{C}$ with this property.

(d) The essential image cannot be used to distinguish functions that are almost everywhere equal: If $f = g$ holds P-almost everywhere, then $\operatorname{ess.im}(f) = \operatorname{ess.im}(g)$.

(e) If $\operatorname{ess.im}(f) = \{0\}$, then $f = 0$ P-almost everywhere.

(f) The essential image of f the biggest set contained in the closures of $\operatorname{im}(g)$ for all g that are a.e. equal to f:

$$\operatorname{ess.im}(f) = \bigcap_{f=g \text{ a.e.}} \overline{g(\Omega)} \subset \overline{f(\Omega)}.$$

(g) If f is continuous and $\operatorname{supp}(P) = \Omega$, then $\operatorname{ess.im}(f) = \overline{f(\Omega)}$.

We say that f is *essentially bounded* if its essential range is bounded, hence compact. In that case, the largest value of $|\lambda|$, as λ runs through the essential range of f, is called the *essential supremum* $\|f\|_\infty$ of f:

$$\|f\|_\infty \coloneqq \sup_{\lambda \in \operatorname{ess.im}(f)} |\lambda|.$$

Denote by $L^\infty(P) = L^\infty(\Omega, \mathcal{F}, P; \mathbb{C})$ all the essentially bounded complex valued measurable functions on $(\Omega, \mathcal{F}, P)$. We agree that in $L^\infty(P)$ two functions f, g are the same if $\|f - g\|_\infty = 0$, i.e., $f = g$ P-a.e.. Then $L^\infty(P)$ is a commutative C^*-algebra with the essential supremum norm and with the involution given by complex conjugation.

If easy to see that $f \in L^\infty(P)$ is invertible if and only if $0 \notin \operatorname{ess.im}(f)$. As a consequence, the spectrum of f is exactly the essential range of f:

$$\sigma(f) = \operatorname{ess.im}(f)\,.$$

Remark 2.19. Denote by $L^\infty(\Omega, \mathcal{F})$ all the bounded complex valued measurable functions on $(\Omega, \mathcal{F})$. Then $L^\infty(\Omega, \mathcal{F})$ is also a commutative C^*-algebra with the uniform norm (supremum norm)

$$\|f\|_u \coloneqq \sup_{x\in\Omega} |f(x)| \quad \text{for} \quad f \in L^\infty(\Omega, \mathcal{F})\,.$$

and with the involution given by complex conjugation. However, the essential supremum $\|\cdot\|_\infty$ is only a seminorm on $L^\infty(\Omega, \mathcal{F})$. Fixing a projection valued measure P, one sees easily that

$$N = \{f \in L^\infty(\Omega, \mathcal{F}) : \|f\|_\infty = 0\}$$

is an ideal of $L^\infty(\Omega, \mathcal{F})$ which is closed. Then one can check that $L^\infty(\Omega, \mathcal{F})/N$, equipped with the quotient norm, is exactly $L^\infty(P)$. As is usually done in measure theory, the distinction between f and its equivalence class $f + N$ will be ignored.

Integrals For each $f \in L^\infty(P)$, we shall define $\int_\Omega f\,\mathrm{d}P$, the *integral* of f with respect to P, be the bounded linear operator on H satisfies that

$$\langle \int_\Omega f\,\mathrm{d}P\,x, y\rangle = \int_\Omega f\,\mathrm{d}P_{x,y} \quad \text{for all } x, y \in H\,. \tag{2.16}$$

First of all, $\int_\Omega f\,\mathrm{d}P$ is well-defined, in the since that if $f = g$ P-a.e., then $\int_\Omega f\,\mathrm{d}P = \int_\Omega g\,\mathrm{d}P$, since $\int_\Omega f\,\mathrm{d}P_{x,y} = \int_\Omega g\,\mathrm{d}P_{x,y}$ for all x, y in H. Secondly, the integral must be unique if it exists. The existence is provided by the next theorem.

Theorem 2.42. *Let P be a projection valued measure on $(\Omega, \mathcal{F})$. Then for all $f \in L^\infty(P)$, the integral $\int_\Omega f dP \in \mathcal{B}(H)$, satisfying* (2.16)*, does exist. Moreover,*

$$\mathscr{A}_P \coloneqq \overline{\operatorname{span}}\,\{P(A) : A \in \mathcal{F}\} \subset \mathcal{B}(H)\,.$$

is a sub-C^*-algebra of $\mathcal{B}(H)$, and the mapping

$$L^\infty(P) \to \mathscr{A}_P \ ; \ f \mapsto \int_\Omega f dP \tag{2.17}$$

is an isometric $*$-isomorphism.

Proof. Step 1. Let $S(\Omega, \mathcal{F})$ be the collection of all the the simple measurable functions on $(\Omega, \mathcal{F})$, as a subspace of $L^\infty(P)$. We shall show for each $\phi \in S(\Omega, \mathcal{F})$, $\int_\Omega \phi \, \mathrm{d}P$ exists, and the mapping

$$\phi \mapsto \int_\Omega \phi \, \mathrm{d}P$$

is a isometric $*$-isomorphism of $S(\Omega, \mathcal{F})$ onto span $\{P(A) : A \in \mathcal{F}\}$.

Suppose $\phi = \sum_{k=1}^n \alpha_k 1_{A_k}$ so that $A_k \cap A_m = \emptyset$ for $k \neq m$ and A_k's are measurable. Then, following the intuition, we define

$$\int_\Omega \phi \, \mathrm{d}P \coloneqq \sum_{k=1}^n \alpha_k P(A_k) \in \operatorname{span}\{P(A) : A \in \mathcal{F}\} \tag{2.18}$$

Thus, clearly for x, y in H,

$$\begin{aligned} \langle \int_\Omega \phi \, \mathrm{d}P \, x, y \rangle &= \langle \sum_{k=1}^n \alpha_k P(A_k) \, x, y \rangle = \sum_{k=1}^n \alpha_k \langle P(A_k) \, x, y \rangle \\ &= \sum_{k=1}^n \alpha_k P_{x,y}(A_k) = \int_\Omega \phi \, \mathrm{d}P_{x,y} \, . \end{aligned}$$

Thus our definition satisfies (2.16). This also show that the definition (2.18) does not depend on the particular representation of ϕ.

- Since $\{A_k\}$ are pairwise disjoint, $\{P(A_k)\}$ are pairwise orthogonal, then

$$\begin{aligned} \| \int_\Omega \phi \, \mathrm{d}P \| &= \| \sum_k \alpha_k P(A_k) \| \\ &= \max\{|\alpha_k| : P(A_k) \neq 0\} = \|\phi\|_\infty \, . \end{aligned} \tag{2.19}$$

- Trivially, for simple function ϕ, ψ an complex numbers α, β we have
$$\int_\Omega \alpha\phi + \beta\psi \, \mathrm{d}P = \alpha \int_\Omega \phi \, \mathrm{d}P + \beta \int_\Omega \psi \, \mathrm{d}P \, .$$
- If ψ is also a simple function and $\psi = \sum_{m=1}^{n'} \beta_m 1_{B_m}$ so that $B_k \cap B_m = \emptyset$ for $k \neq m$ and B_m's are measurable, then $\phi\psi = \sum_{k,m} \alpha_k \beta_m 1_{A_k \cap B_m}$. By (2.18), we have
$$\begin{aligned}\int_\Omega \phi\psi \, \mathrm{d}P &= \sum_{k,m} \alpha_k \beta_m P(A_k \cap B_m) = \sum_{k,m} \alpha_k \beta_m P(A_k) P(B_m) \\ &= \sum_k \alpha_k P(A_k) \ \sum_{k,m} \alpha_k \beta_m P(A_k) P(B_m) = \int_\Omega \phi \, \mathrm{d}P \int_\Omega \psi \, \mathrm{d}P \, .\end{aligned}$$
- Since $\bar{\phi} = \sum_{k=1}^n \overline{\alpha}_k 1_{A_k}$, we have
$$\begin{aligned}\int_\Omega \bar{\phi} \, \mathrm{d}P &= \sum_{k=1}^n \overline{\alpha}_k P(A_k) = \sum_{k=1}^n (\alpha_k P(A_k))^* \\ &= \sum_{k=1}^n (\alpha_k P(A_k))^* = \left(\sum_{k=1}^n \alpha_k P(A_k) \right)^* = (\int_\Omega \phi \, \mathrm{d}P)^* \, .\end{aligned}$$

Step 2. Now suppose $f \in L^\infty(P)$. Then there is a sequence of simple measurable functions ϕ_n that converges to f uniformly, i.e., in the norm of $L^\infty(P)$. By (2.19) the corresponding operators $\int_\Omega \phi_n \, \mathrm{d}P$ form a Cauchy sequence in the Banach space $\mathcal{B}(H)$. Set
$$\int_\Omega f \, \mathrm{d}P \coloneqq \lim_{n\to\infty} \int_\Omega \phi_n \, \mathrm{d}P \in \mathscr{A}_P \, .$$
Note that this definition does not depend on the particular choice of $\{\phi_n\}$. To check (2.16), given x, y in H, we have
$$\begin{aligned}\langle \int_\Omega f \, \mathrm{d}P \, x, y \rangle &= \lim_{n\to\infty} \langle \int_\Omega \phi_n \, \mathrm{d}P \, x, y \rangle \\ &= \lim_{n\to\infty} \int_\Omega \phi_n \, \mathrm{d}P_{x,y} = \int_\Omega f \, \mathrm{d}P_{x,y} \, .\end{aligned}$$
We are going to show that the mapping (2.17) is a isometric $*$-isomorphism of $L^\infty(P)$ onto $\mathscr{A}_P$.

Obviously (2.19) leads to

$$\| \int_\Omega f \, \mathrm{d}P \| = \lim_{n\to\infty} \| \int_\Omega \phi_n \, \mathrm{d}P \| = \lim_{n\to\infty} \|\phi_n\|_\infty = \|f\|_\infty \,.$$

Thus $f \mapsto \int_\Omega f \, \mathrm{d}P$ is a isometric. It's easy to see that the mapping (2.17) is a $*$-homomorphism. Then the rangle of the mapping (2.17) is a closed subspace of $\mathscr{A}_P$ containing $\operatorname{span}\{P(A) : A \in \mathcal{F}\}$. Thus the mapping (2.17) is surjective. $\square$

Corollary 2.43. *If $T \in \mathcal{B}(H)$ commutes with $\{P(A) : A \in \mathcal{F}\}$, then T commutes with $\{\int_\Omega f \, \mathrm{d}P : f \in L^\infty(P)\}$.*

Corollary 2.44. *Let P be a projection valued measure on $(\Omega, \mathcal{F})$. Let $f \in L^\infty(P)$. Then the following statements holds.*

(a) *For each $x \in H$,*

$$\left\| \Big(\int_\Omega f \, \mathrm{d}P \Big) \, x \right\|^2 = \int_\Omega |f|^2 \, \mathrm{d}P_{x,x} \,;$$

and hence

$$\|f\|_\infty = \left\| \Big(\int_\Omega f \, \mathrm{d}P \Big) \, x \right\| = \sup_{\|x\| \le 1} \left[\int_\Omega |f|^2 \, \mathrm{d}P_{x,x} \right]^{1/2} \,.$$

(b) *If $\{f_n\}$ is a bounded sequence in $L^\infty(P)$, and $f_n(\omega) \to f(\omega)$ P-a.e., we have $f \in L^\infty(P)$, and*

$$\int f \, \mathrm{d}P = s\text{-}\lim_{n\to\infty} \int f_n \, \mathrm{d}P \,.$$

Proof. To show (a), note that

$$\begin{aligned} \left\| \Big(\int_\Omega f \, \mathrm{d}P \Big) \, x \right\|^2 &= \langle \int_\Omega f \, \mathrm{d}P \, x, \int_\Omega f \, \mathrm{d}P, x \rangle = \langle \int_\Omega \bar{f} \, \mathrm{d}P \int_\Omega f \, \mathrm{d}P \, x, x \rangle \\ &= \langle \int_\Omega |f|^2 \, \mathrm{d}P \, x, x \rangle = \int_\Omega |f|^2 \, \mathrm{d}P_{x,x} \,. \end{aligned}$$

(b) is a direct consequence of (a) and the dominated convergence theorem. $\square$

Spectrum of the Integrals In this paragraph we will discuss the spectrum of the integrals and its classification.

Theorem 2.45. *Let P be a projection valued measure on $(\Omega, \mathcal{F})$. Let $f \in L^\infty(P)$. Then*

$$\sigma\left(\int_\Omega f \,\mathrm{d}P\right) = \operatorname{ess.im}(f) \subset \overline{f(\Omega)}\,.$$

In particular, if f is continuous and $\operatorname{supp}(P) = \Omega$*, then*

$$\sigma\left(\int_\Omega f \,\mathrm{d}P\right) = \overline{f(\Omega)}\,.$$

Proof. Note that since $\mathscr{A}_P$ is a commutative sub-C^*-algebra of $\mathcal{B}(H)$, by Theorem 2.38,

$$\sigma\left(\int_\Omega f \,\mathrm{d}P\right) = \sigma_{\mathscr{A}_P}\left(\int_\Omega f \,\mathrm{d}P\right)\,.$$

Since the mapping (2.17) is a $*$-isomorphism, we have

$$\sigma\left(\int_\Omega f \,\mathrm{d}P\right) = \sigma_{L^\infty(P)}(f) = \operatorname{ess.im}(f) \subset \overline{f(\Omega)}\,.$$

If f is continuous and $\operatorname{supp}(P) = \Omega$, then $\operatorname{ess.im}(f) = \overline{f(\Omega)}$ and the desired result holds. □

Lemma 2.46. *Let P be a projection valued measure on $(\Omega, \mathcal{F})$. Let $f \in L^\infty(P)$. Then*

$$\operatorname{Ker}(\int_\Omega f \,\mathrm{d}P) = \operatorname{Ran}(P(f = 0))\,.$$

Proof. Firstly, for $x \in \operatorname{Ran}(P(f = 0))$ we have $P(f = 0)x = x$, then

$$(\int_\Omega f \,\mathrm{d}P)x = (\int_\Omega f \,\mathrm{d}P)P(f = 0)x = \left(\int_\Omega f 1_{\{f=0\}} \,\mathrm{d}P\right) x = 0.$$

On the other hand, if $x \in \operatorname{Ker}(\int_\Omega f \,\mathrm{d}P)$, then

$$\left\|(\int_\Omega f \,\mathrm{d}P)x\right\|^2 = \int_\Omega |f|^2 \,\mathrm{d}P_{x,x} = 0\,.$$

Thus $P_{x,x}(|f| > 0) = 0$, i.e, $P(|f| > 0)x = 0$. Therefore we have

$$x = P(|f| > 0)x + P(f = 0)x = P(|f| > 0)x \in \mathrm{Ran}(P(f = 0))\,.$$

We are done. □

Theorem 2.47. *Let P be a projection valued measure on $(\Omega, \mathcal{F})$. Let $f \in L^\infty(P)$. Then*

$$\begin{aligned}
\sigma\left(\int_\Omega f\,\mathrm{d}P\right) &= \mathrm{ess.\,im}(f)\ ;\ \sigma_r\left(\int_\Omega f\,\mathrm{d}P\right) = \emptyset\,; \\
\sigma_p\left(\int_\Omega f\,\mathrm{d}P\right) &= \{\lambda \in \mathrm{ess.\,im}(f) : P(f = \lambda) \neq 0\}\ ; \\
\sigma_c\left(\int_\Omega f\,\mathrm{d}P\right) &= \{\lambda \in \mathrm{ess.\,im}(f) : P(f = \lambda) = 0\}\ .
\end{aligned}$$

Proof. It has been shown in Theorem 2.45 that $\sigma\left(\int_\Omega f\,\mathrm{d}P\right) = \mathrm{ess.\,im}(f)$.

Since $\int_\Omega f\,\mathrm{d}P$ is normal, we have $\sigma_r\left(\int_\Omega f\,\mathrm{d}P\right) = \emptyset$ by Theorem **??**.

It remains to show that $\lambda \in \sigma_p(\int_\Omega f\,\mathrm{d}P)$ if and only if $\lambda \in \mathrm{ess.\,im}(f)$ and $P(f = \lambda) \neq 0$. Observe that by Lemma 2.46,

$$\mathrm{Ker}(\lambda I - \int_\Omega f\,\mathrm{d}P) = \mathrm{Ker}(\int_\Omega \lambda - f\,\mathrm{d}P) = P(f = \lambda)$$

then the desired result follows. □

2.7.2 The Spectral Theorem

The principal assertion of the spectral theorem is that every bounded normal operator N on a Hilbert space induces (in a canonical way) a projection valued measure E on the Borel subsets of its spectrum $\sigma(N)$ and that N can be reconstructed from E by the integral of the type discussed above. A large part of the theory of normal operators depends on this fact.

It should perhaps be stated explicitly that the spectrum $\sigma(T)$ of an operator $T \in \mathcal{B}(H)$ will always refer to the full algebra $\mathcal{B}(H)$. In other words, $\lambda \in \sigma(T)$ if and only if $\lambda I - T$ has no inverse in $\mathcal{B}(H)$. Sometimes we shall also be concerned with the sub-C^*-algebra $\mathscr{A}$ of $\mathcal{B}(H)$. Suppose that $N \in \mathscr{A}$, then by Theorem 2.38, $\sigma_{\mathscr{A}}(N) = \sigma(N)$. Thus N has the same spectrum relative to all sub-C^*-algebras in $\mathcal{B}(\mathrm{H})$ that contain N.

Recall that $\mathscr{A}$ is a commutative sub-C^*-algebra of $\mathcal{B}(H)$ iff $\mathscr{A}$ is a closed normal subalgebra of $\mathcal{B}(H)$ containing the identity operator I.

Theorem 2.48. *Let $\mathscr{A}$ be a commutative sub-C^*-algebra of $\mathcal{B}(H)$. Let Δ be all the nonzero complex homomorphisms of $\mathscr{A}$. Then the inverse of the Gelfand representation $\Gamma^{-1} : \hat{\mathscr{A}} = C(\Delta) \to \mathscr{A}$ has a specific formular. In fact, there exists a unique projection valued measure P on $(\Delta, \mathcal{B}(\Delta))$, so that*

$$N = \int_{\Delta} \hat{N} \, \mathrm{d}P \,, \quad \textit{for all } N \in \mathscr{A} \,, \tag{2.20}$$

where $\hat{N}$ is the Gelfand transform of N; and $\operatorname{supp}(P) = \Delta$.

Recall that by the Gelfand-Naimark theorem, the Gelfand representation Γ is an isometric $*$-isomorphism of $\mathscr{A}$ onto $\hat{\mathscr{A}} = C(\Delta)$. Thus $\Gamma^{-1} : \hat{\mathscr{A}} = C(\Delta) \to \mathscr{A}$ is well-defined.

Proof. Recall that (2.20) is an abbreviation for

$$\langle Nx, y \rangle = \int_{\Delta} \hat{N} \, \mathrm{d}P_{x,y} \quad \text{for all} \quad x, y \in H \,.$$

Step 1. Our first task is to find the complex measures $\{P_{x,y} : x, y \in H\}$. Fix now $x, y \in H$, define

$$\ell(\hat{N}) = \langle Nx, y \rangle \quad \text{for all} \quad \hat{N} \in \hat{\mathscr{A}} = C(\Delta) \,.$$

Then

$$|\ell(\hat{N})| \leq \|N\| \|x\| \|y\| = \|\hat{N}\| \|x\| \|y\|$$

since the $\Gamma : \mathscr{A} \to C(\Delta)$ is isometric. Thus ℓ is a bounded continuous linear functional on $C(\Delta)$. By the Riesz-Markov-Kakutani representation theorem, there exists a unique complex Radon measure, namely $\mu_{x,y}$, on $(\Delta, \mathcal{B}(\Delta))$, so that

$$|\ell(\hat{N})| = \langle Nx, y\rangle = \int_\Delta \hat{N}\,\mathrm{d}\mu_{x,y} \tag{2.21}$$

for all $\hat{N} \in C(\Delta)$; and the total variance

$$\|\mu_{x,y}\|_{TV} = \|\ell\| \le \|x\|\|y\|\,.$$

It suffices now to show that there exists a projection valued measure P on $(\Delta, \mathcal{B}(\Delta))$ so that

$$\mu_{x,y}(A) = \langle P(A)x, y\rangle \tag{2.22}$$

for all A in $\mathcal{B}(\Delta)$ and x, y in H and $\mathrm{supp}(P) = \Delta$.

Step 2. We shall show that for each fixed $f \in L^\infty(\Delta, \mathcal{B}(\Delta))$, there exists a self-adjoint operator $P(f)$ so that $\int_\Delta f\,\mathrm{d}\mu_{x,y} = \langle P(f)x, y\rangle$. Then $P(A) \coloneqq P(1_A)$ satisfies (2.22). By Theorem **??**, we have only to show the mapping

$$(x, y) \mapsto \int_\Delta f\,\mathrm{d}\mu_{x,y}$$

is a bounded sesquilinear hermitian form on H. For given α, β in $\mathbb{C}$ and x_1, x_2, y in H, since

$$\begin{aligned}\int_\Delta \hat{N}\,\mathrm{d}\mu_{\alpha x_1+\beta x_2,y} &= \langle N(\alpha x_1 + \beta x_2), y\rangle \\ &= \alpha\langle Nx_1, y\rangle + \beta\langle Nx_2, y\rangle \\ &= \int_\Delta \hat{N}\,\mathrm{d}(\alpha\mu_{x_1,y} + \beta\mu_{x_2,y})\end{aligned}$$

for all $\hat{N} \in C(\Delta)$, we get

$$\mu_{\alpha x_1+\beta x_2,y} = \alpha\mu_{x_1,y} + \beta\mu_{x_2,y}\,.$$

Similarly, for given x, y in H, since

$$\begin{aligned}\int_\Delta \hat{N}\,\mathrm{d}\mu_{x,y} &= \langle x, N^*y\rangle = \overline{\langle N^*y, x\rangle} \\ &= \int_\Delta \widehat{N^*}\,\mathrm{d}\mu_{y,x} = \int_\Delta \overline{\hat{N}}\,\mathrm{d}\mu_{y,x} = \overline{\int_\Delta \hat{N}\,\mathrm{d}\mu_{y,x}}\end{aligned}$$

for all $\hat{N} \in C(\Delta)$, this implies that $\mu_{x,y} = \overline{\mu_{y,x}}$. Moreover,

$$\left\| \int_\Delta f \, \mathrm{d}\mu_{x,y} \right\| \leq \|f\|_u \|\mu_{x,y}\|_{TV} \leq \|f\|_u \|x\| \|y\| \, ;$$

so we get the existence of $\|P(f)\|$, with

$$\|P(f)\| \leq \|f\|_u \, .$$

Besides, note that by (2.21)

$$P(\hat{N}) = N \text{ for all } \hat{N} \in C(\Delta) \, .$$

Step 3. Trivially $P(\emptyset) = 0$. We also have $P(\Delta) = I$, since for all x, y in H there holds

$$\langle P(\Delta)x, y\rangle = \int_\Delta 1 \, \mathrm{d}\mu_{x,y} = \int_\Delta \hat{I} \, \mathrm{d}\mu_{x,y} = \langle x, y\rangle \, .$$

In order that P is a projection valued measure, we have only to show that

$$P(A \cap B) = P(A)P(B) \quad \text{for all} \quad A, B \in \mathcal{B}(\Delta) \, . \tag{2.23}$$

Indeed, this implies $P(A) = P(A)^2$, then combine this with that $P(A)$ is self-adjoint, we get that $P(A)$ is a (orthogonal) projection. Set

$$P : \mathcal{B}(\Delta) \to \mathcal{P}(H) \ ; \ A \mapsto P(A) \, ,$$

then the mapping P satisfies part (a), (b) of Definition 2.10. Note that for all x, y in H, $\langle P(A)x, y\rangle = \mu_{x,y}(A)$ is a complex measure on $\mathcal{B}(\Delta)$. Thus (c$'$) is true for P. So P is a projection valued measure with $P_{x,y} = \mu_{x,y}$.

Thus for $f \in L^\infty(\Delta, \mathcal{B}(\Delta))$,

$$\langle P(f)x, y\rangle = \int_\Delta f \, \mathrm{d}\mu_{x,y} = \int_\Delta f \, \mathrm{d}P_{x,y} = \langle \int_\Delta f \, \mathrm{d}P \, x, y\rangle \, ;$$

and hence

$$P(f) = \int_\Delta f \, \mathrm{d}P \, .$$

Let $f = \hat{N} \in C(\Delta)$, since $P(\hat{N}) = N$, we get

$$N = \int_\Delta \hat{N} \, \mathrm{d}P \, .$$

Step 4. We enhance (2.23) to

$$P(fg) = P(f)P(g) \quad \text{for all } f, g \in L^\infty(\Delta, \mathcal{B}(\Delta)) \, .$$

To this end, observe that

$$P(\hat{N}\hat{M}) = P(\widehat{NM}) = NM = P(\hat{N})P(\hat{M})$$

for all $\hat{N}$, $\hat{M}$ in $C(\Delta)$. For fixed x, y in H, thus

$$\int \hat{N}\hat{M} \, \mathrm{d}\mu_{x,y} = \int \hat{N} \, \mathrm{d}\mu_{Mx,y} \, .$$

By Lusin's theorem, and the dominated convergence theorem, we get

$$\begin{aligned} \int f\hat{M} \, \mathrm{d}\mu_{x,y} &= \int f \, \mathrm{d}\mu_{Mx,y} = \langle P(f)Mx, y\rangle \\ &= \langle Mx, P(f)^* y\rangle = \int \hat{M} \, \mathrm{d}\mu_{x,P(f)^* y} \, , \end{aligned}$$

for all $f \in L^\infty(\Delta, \mathcal{B}(\Delta))$. Again, by Lusin's theorem and the dominated convergence theorem,

$$\int fg \, \mathrm{d}\mu_{x,y} = \int g \, \mathrm{d}\mu_{x,P(f)^* y} = \langle P(g)x, P(f)^* y\rangle = \langle P(f)P(g)x, y\rangle \, ,$$

that is

$$\langle P(fg)x, y\rangle = \langle P(f)P(g)x, y\rangle \, .$$

Since x, y is arbitrary, we get $P(fg) = P(f)P(g)$ as desired.

Step 5. Finally, we show that $\operatorname{supp}(P) = \Omega$. To this end, we have to show that for every open sets U in Δ, $P(U) \neq 0$. Suppose for contradiction that there exists an open set U in Δ with $P(U) = 0$. Then, by Urysohn's lemma, for fixed $\phi \in U$, the exists $\hat{N} \in C(\Delta)$ with $\operatorname{supp}(\hat{N}) \subset U$ and $\hat{N}(\phi) = \phi(N) = 1$. However, then

$$N = \int_\Delta \hat{N} \, \mathrm{d}P = \int_\Delta \hat{N} 1_U \, \mathrm{d}P = 0$$

and hence $\phi(N) = 0$, which is a contradiction! $\square$

We now specialize this theorem to a single operator.

Theorem 2.49 (The Spectrum Theorem)**.** *If* $N \in \mathcal{B}(H)$ *and* N *is normal, then there exists a unique projection-valued measure* E *on* $(\sigma(N), \mathcal{B}(\sigma(N)))$ *so that* $\operatorname{supp}(E) = \sigma(N)$ *and*

$$N = \int_{\sigma(N)} \lambda \, E(\mathrm{d}\lambda) \, .$$

We shall refer to this E *as the spectral decomposition of* N*.*

Proof. Let $\mathscr{A}_N \subset \mathcal{B}(H)$ be the smallest sub-C^*-algebra that contains N; i.e.,

$$\mathscr{A}_N = \operatorname{cl}\left\{p(N, N^*) : p \text{ is a polynomial in two variables}\right\} \, .$$

Clearly, $\mathscr{A}_N$ is a commutative sub-C^*-algebra of $\mathcal{B}(H)$. Let Δ be all the nonzero complex homomorphisms of $\mathscr{A}_N$. Then by Theorem 2.48, there is a unique projection valued measure P on $(\Delta, \mathcal{B}(\Delta))$ so that $\operatorname{supp}(P) = \Delta$, and

$$N = \int_{\Delta} \hat{N} \, \mathrm{d}P \, .$$

By (2.24), $\hat{N}$ is a homeomorphism of Δ onto $\sigma(N)$. Let $E = P \circ \hat{N}^{-1}$ be the image measure on $(\sigma(N), \mathcal{B}(\sigma(N)))$, then we get $\operatorname{supp}(E) = \sigma(N)$, and

$$N = \int_{\Delta} \lambda E(\mathrm{d}\lambda) \, .$$

In fact E is unique is guaranteed by the fact that P is unique. To give a direct proof, note that for any polynomial in two variables p,

$$p(N, N^*) = \int_{\Delta} p(\lambda, \bar{\lambda}) E(\mathrm{d}\lambda) \, . \tag{2.24}$$

By the Stone-Weierstrass theorem, these polynomials form a dense subalgebra in $C(\sigma(T))$. Thus for each given x, y, $E_{x,y}$ are therefore uniquely determined by the integrals (2.24), hence by N. $\square$

Remark 2.20. For normal operator $N \in \mathcal{B}(H)$, we sometimes assume it's spectrum measure E is defined on the complex plane $\mathbb{C}$ equipped with the Borel algebra by setting $E(A) = E(A \cap \sigma(N))$ for all Borel sets A in $\mathbb{C}$. If N is self-adjoint, as we know then $\sigma(N) \subset \mathbb{R}$, so we shall think that E is defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and if N is unitary then $\sigma(N) \subset S^1$, we can think that E is defined on $(S^1, \mathcal{B}(S^1)$.

Bounded Measurable Functional Calculus If E is the spectral decomposition of a normal operator $N \in \mathcal{B}(H)$, and if f is a bounded Borel function on $\sigma(N)$, it is customary to define

$$\tilde{f}(N) := \int_{\sigma(N)} f \, \mathrm{d}E \,.$$

Using this notation, part of the content of Theorems 2.42 to 2.49 can be summarized as follows:

Exercise 2.7 (Bounded Measurable Functional Calculus). Let E be the spectral decomposition of a normal operator N in $\mathcal{B}(H)$. Recall that

$$\mathscr{A}_E := \overline{\operatorname{span}} \left\{ E(A) : A \in \mathcal{B}(\sigma(N)) \right\} \,.$$

Then, the mapping

$$L^\infty(E) \to \mathscr{A}_E \;;\; f \mapsto \tilde{f}(N)$$

is an isometric $*$-isomorphism satisfying the following statements.

(a) (Normalization) If $f(\lambda) = \lambda$ for all $\lambda \in \sigma(N)$ then $\tilde{f}(N) = N$.

(b) (Commutative) If $T \in \mathcal{B}(H)$ satisfies $NT = TN$, then $\tilde{f}(N)T = T\tilde{f}(N)$ for all $f \in L^\infty(E)$.

(c) (Convergence) If $f_n \in L^\infty(E)$ be a bounded sequence in $L^\infty(E)$, and $f_n \to f$ P-a.e., then $f \in L^\infty(P)$ and

$$\tilde{f}(N) = \text{s-}\lim_{n\to\infty} \tilde{f}_n(N) \,.$$

(d) (Image) $\mathscr{A}_E$ is the closure of $\mathscr{A}_N$ relative to the strong operator topology, where $\mathscr{A}_N$ is defined in (2.14).

(e) (Eigenvalues) If $\lambda \in \sigma_p(N)$ and $x \in H$ satisfy $Nx = \lambda x$ then

$$\tilde{f}(N)x = f(\lambda)x \quad \text{for all } f \in L^\infty(E)\,.$$

(f) (Spectrum) For every $f \in L^\infty(E)$ the operator $\tilde{f}(N)$ is normal and

$$\sigma(\tilde{f}(N)) = \operatorname{ess.im}(f) \subset \overline{f(\sigma(N))}\,.$$

In particular, if f is continuous, then $\sigma(\tilde{f}(N)) = f(\sigma(N))$.

(g) (Positive) If $f \in L^\infty(E)$ and $f \geq 0$ then $\tilde{f}(N) = \tilde{f}(N)^* \geq 0$.

(h) (Composition) If $f \in C(\sigma(N))$ and $g \in L^\infty(f(\sigma(N)))$ then

$$\widetilde{(g \circ f)}(N) = \tilde{g}(\tilde{f}(N))\,.$$

2.7.3 Discrete Spectrum and Essential Spectrum

If $N \in \mathcal{B}(H)$ is normal, its eigenvalues bear a simple relation to its spectral decomposition E. This will be derived from the fact that for each $f \in L^\infty(E)$,

$$\operatorname{Ker}(f(N)) = \operatorname{Ran}\left(E\left(f = 0\right)\right)\,.$$

Clearly this is a restatement of Lemma 2.46.

Theorem 2.50. *Suppose $N \in \mathcal{B}(H)$ is a normal operator with spectral decomposition E.*

(a) $E(\{\lambda\})$ is the projection of H onto $\operatorname{Ker}(\lambda I - N)$*, and hence λ is an eigenvalue of N if and only if $E(\{\lambda\}) \neq 0$.*

(b) Every isolated point of $\sigma(N)$ is an eigenvalue of N.

(c) *If* $\sigma(N) = \{\lambda_1, \lambda_2, \lambda_3, \ldots\}$ *is at most countable, then* $E(\lambda_k)$ *is the projection of* H *onto* $\mathrm{Ker}(\lambda_k I - N)$ *and*

$$\tilde{f}(N) = \sum_k f(\lambda_k) E(\{\lambda_k\}) \quad \textit{for all} \quad f \in L^\infty(E)\,.$$

Statements (a) and (b) explain the term *point spectrum* of N for the set of all eigenvalues of N.

Definition 2.11. For a normal operator $N \in \mathcal{B}(H)$, the *discrete spectrum* $\sigma_d(N)$ of N is the set of all eigenvalues of N of finite multiplicities which are isolated points of the spectrum $\sigma(N)$. The complement set $\sigma_{\text{ess}}(N) := \sigma(N) \backslash \sigma_d(N)$ is called the *essential spectrum* of N.

By the Theorem 2.50 an isolated point of the spectrum $\sigma(N)$ is always an eigenvalue of N. Therefore, a number belongs to $\sigma_{\text{ess}}(N)$ of N if and only if it is an accumulation point of $\sigma(N)$ or an eigenvalue of infinite multiplicity.

Theorem 2.51. *Let* $N \in \mathcal{B}(H)$ *be a normal operator with spectrum decomposition* E. *Then* $\lambda_0 \in \sigma_d(N)$ *if and only if there is an open neighborhood* U *of* λ_0,

$$\dim(\mathrm{Ran} E(U)) < \infty\,;$$

and $\lambda_0 \in \sigma_{ess}(N)$ *if and only if for each open neighborhood* U *of* λ_0,

$$\dim(\mathrm{Ran} E(U)) = \infty\,.$$

Proof. It suffices to show that $\lambda_0 \in \sigma_d(N)$ if and only if there is an open neighborhood U of λ_0,

$$\dim(\mathrm{Ran} E(U)) < \infty\,;$$

The necessity is trivial, since if $\lambda_0 \in \sigma_d(N)$, then take an open neighborhood U of λ_0 so that $U \cap \sigma(N) = \{\lambda_0\}$. Thanks to $\mathrm{supp}(E) = \sigma(N)$, we have $E(U) = E(\{\lambda_0\})$. By part (a) of Theorem 2.50, then the desired result follows.

To prove the sufficiency, assume that $\lambda_0 \in \sigma_d(N)$ and there exists an open neighborhood U of λ_0 so that $\dim(\mathrm{Ran}E(U)) < \infty$. If λ_0 is not an isolated point of $\sigma(N)$, then there is a sequence $\{\lambda_n\}$ of distinct spectral points of N so that $\lambda_n \to \lambda_0$. We get take a sequence $\{U_n\}$ of pairwise disjoint open sets so that $\lambda_n \in U_n \subset U$. Since $\mathrm{supp}(E) = \sigma(N)$, $E(U_n) \neq 0$. Thus

$$\dim R(E(U)) \geq \sum_n \dim R(E(U_n)) = \infty$$

which is absurd. Thus λ_0 is an isolated point of $\sigma(N)$ and hence $E(U) = E(\{\lambda_0\})$. By Theorem 2.50, then $\dim\mathrm{Ker}(\lambda_0 - I) = \dim\mathrm{Ker}(E(\{\lambda_0\})) < \infty$ follows as desired. □

2.7.4 Applications of Functional Calculus

The following proof, containing our first application of this symbolic calculus, generalizes Theorem 2.52.

Theorem 2.52. *Let $N \in \mathcal{B}(H)$ be normal. Then*

$$\|N\| = \sup_{\|x\|=1} |\langle Nx, x\rangle| \,.$$

Proof. Choose $\epsilon > 0$. It is clearly enough to show that

$$|\langle Nx_0, x_0\rangle| > \|N\| - \epsilon$$

for some $x_0 \in H$ with $\|x_0\| = 1$. Indeed, since $\|N\| = r_\sigma(N)$, there is $\lambda_0 \in \sigma(N)$ so that $|\lambda_0| = \|N\|$. Then since $\lambda_0 \in \mathrm{supp}(E)$, $E(B(\lambda_0, \epsilon)) \neq 0$. Take $x_0 \in \mathrm{Ran}(E(B(\lambda_0, \epsilon)))$ and $\|x_0\| = 1$, then

$$\begin{aligned}\|Nx_0 - \lambda_0 x_0\|^2 &= \int_{\sigma(N)} |\lambda - \lambda_0|^2 E_{x_0,x_0}(\mathrm{d}\lambda) \\ &= \int_{\sigma(N)} |\lambda - \lambda_0|^2 1_{B(\lambda_0,\epsilon)} E_{x_0,x_0}(\mathrm{d}\lambda) \leq \epsilon^2 \,;\end{aligned}$$

and hence

$$|\langle Nx_0, x_0\rangle - \lambda_0| \leq \epsilon \,.$$

Thus $|\langle Nx_0, x_0\rangle| > |\lambda_0| - \epsilon = \|N\| - \epsilon$ as desired. □

Remark 2.21. To see that normality is needed here, let T be the linear operator on $\mathbb{C}^2$ (with basis $\{e_1, e_2\}$) given by $Te_1 = 0, Te_2 = e_1$. It has $\|T\| = 1$, but $|(Tx, x)| \leq \frac{1}{2}$ if $\|x\| \leq 1$.

We give a application of the functional calculus for normal operators.

Theorem 2.53. *Let N be a normal operator on H.*

(a) N is self-adjoint if and only if $\sigma(N) \subset \mathbb{R}$.

(b) N is non-negative (that is N is self-adjoint and $\langle Nx, x\rangle \geq 0$ for all $x \in H$) if and only if $\sigma(N) \subset [0, \infty)$, i.e., $N \geq 0$.

(c) N is unitary if and only if $\sigma(N) \subset S^1$.

Proof. We show now part (a). If N is self-adjoint, then by Example 1.8 we get $\sigma(N) \subset \mathbb{R}$. If $N \in \mathcal{B}(H)$ is normal and $\sigma(N) \subset \mathbb{R}$, using the continuous functional calculus, let $f(z) = \bar{z}$ for $z \in \sigma(N)$, then $N^* = \tilde{f}(N)$. Since $\sigma(N) \subset \mathbb{R}$, so that $f(z) = z$, we get $N^* = N$.

We show now part (b). If N is nonnegative, then by the same method in Example 1.8, or by Theorem **??**, we get $(-\infty, 0) \subset \varrho(N)$. Conversely, if $\sigma(N) \subset [0, \infty)$, by part (a) N is self-adjoint. By Theorem **??** or by the spectrum decomposition theorem, $\langle Nx, x\rangle \geq 0$ for all $x \in H$.

If N is unitary, then by Theorem **??**, $\sigma(N) \subset S^1$. On the contrary if $\sigma(N) \subset S^1$, then

$$NN^* = N^*N = \int_{\sigma(N)} \lambda E(\mathrm{d}\lambda) \int_{\sigma(N)} \overline{\lambda} E(\mathrm{d}\lambda) = \int_{\sigma(N)} |\lambda|^2 E(\mathrm{d}\lambda) = I\,.$$

Thus N is unitary. □

Square Roots

Theorem 2.54. *Every nonnegative $T \in \mathcal{B}(H)$ has a unique nonnegative square root $\sqrt{T} \in \mathcal{B}(H)$. If T is invertible in $\mathcal{B}(H)$, so is $\sqrt{T}$.*

Proof. Let $f(z) = \sqrt{z}$ for $z \in \sigma(T)$. By the continuous functional calculus, we denote $\tilde{f}(T)$ by $\sqrt{T}$. Then $\sqrt{T} \in \mathscr{A}_T \subset \mathcal{B}(H)$ is self-adjoint with $\sigma(\sqrt{T}) \subset [0, \infty)$ and $\sqrt{T}\sqrt{T} = T$. (Recall that for a normal operator N, $\mathscr{A}_N$ is the smallest sub-C^*-algebra of $\mathcal{B}(H)$ containing N given by (2.14).) Clearly if $T = S^2$ is invertible, then $\mathrm{Ker}(S) = \{0\}$ and $\mathrm{Ran}(S) = H$ and hence S is invertible.

If $S \in \mathcal{B}(H)$ is also a nonnegative square root of T. Then $T = S^2$ and hence $\mathscr{A}_T \subset \mathscr{A}_S$. Thus there is some $g \in C(\sigma(S))$ is nonnegative so that

$$\sqrt{T} = \int_{\sigma(S)} g(\lambda) E^S(\mathrm{d}\lambda)$$

Then

$$T = \int_{\sigma(S)} g(\lambda)^2 E^S(\mathrm{d}\lambda) = \int_{\sigma(S)} \lambda^2 E^S(\mathrm{d}\lambda)$$

Since the functional calculus is a isometric $*$-isomorphism, we get $g(\lambda)^2 = \lambda$ a.e.-E^S. Thus $g(\lambda) = \lambda$ a.e.-E^S and $\sqrt{T} = S$. □

Proposition 2.55. *If $T \in \mathcal{B}(H)$, then the nonnegative square root of T^*T is the only nonnegative operator $S \in \mathcal{B}(H)$ that satisfies $\|Sx\| = \|Tx\|$ for ever $x \in H$.*

Proof. Note first that T^*T is self-adjoint and

$$\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 \geq 0$$

so that $T^*T \geq 0$. (In the more abstract setting of part (d) of Proposition 2.37 this was much harder to prove!) Next, if $S \in \mathcal{B}(H)$ and $S = S^*$, then

$$\langle S^2 x, x\rangle = \|Sx\|^2 \quad (x \in H)$$

It follows that $\|Sx\| = \|Tx\|$ for every $x \in H$ if and only if

$$\langle (S^2 - T^*T)x, x\rangle = 0 \;\text{ for all } x \in H\,.$$

By Corollary **??** the desired result follows. □

Polar Decomposition The fact that every complex number λ can be factored in the form $\lambda = \alpha|\lambda|$, where $|\alpha| = 1$, suggests the problem of trying to factor $T \in \mathcal{B}(H)$ in the form $T = US$, with U unitary and $S \geq 0$. When this is possible, we call US a *polar decomposition* of T.

Note that U, being unitary, is an isometry. Proposition 2.55 shows therefore that S is uniquely determined by T.

Theorem 2.56. *Let $T \in \mathcal{B}(H)$. Then the following holds.*

(a) If $T \in \mathcal{B}(H)$ is invertible, then T has a unique polar decomposition $T = US$

(b) If $T \in \mathcal{B}(H)$ is normal, then T has a polar decomposition $T = US$ in which U and S commute with each other and with T.

Proof. If T is invertible, so are T^* and T^*T, and the nonnegative square root S of T^*T is also invertible. Put $U = TS^{-1}$. Then U is invertible, and

$$U^*U = S^{-1}T^*TS^{-1} = S^{-1}S^2S^{-1} = I$$

so that U is unitary. Since S is invertible, it is obvious that TS^{-1} is the only possible choice for U.

If $T \in \mathcal{B}(H)$ is normal, Put $s(\lambda) = |\lambda|, u(\lambda) = \lambda/|\lambda|$ if $\lambda \neq 0, u(0) = 1$. Then s and u are bounded Borel functions on $\sigma(T)$. Put $S = s(T), U = u(T)$. Since $s \geq 0$, we have $S \geq 0$. Since $u\bar{u} = 1, UU^* = U^*U = I$. Since $\lambda = u(\lambda)s(\lambda)$, the relation $T = US$ follows from the functional calculus. $\square$

Remark 2.22. In (a), no two of T, U, S need to commute. For example,

$$\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Remark 2.23. It is NOT true that every $T \in \mathcal{B}(H)$ has a polar decomposition. However, if S is the positive square root of T^*T, then $\|Sx\| = \|Tx\|$ for every $x \in H$; hence $Tx = Ty$ if $Sx = Sy$ by linearity. The formula

$$VSx = Tx$$

defines a linear isometry V of $\mathrm{Ran}(S)$ onto $\mathrm{Ran}(T)$, which has a continuous extension to a linear isometry of the closure of $\mathrm{Ran}(S)$ onto the closure of $\mathrm{Ran}(T)$. If there is a linear isometry of $\mathrm{Ran}(S)^\perp$ onto $\mathrm{Ran}(T)^\perp$, then V can be extended to a unitary operator on H, and then T has a polar decomposition. This always happens when $\dim H < \infty$, since $\mathrm{Ran}(S)$ and $\mathrm{Ran}(T)$ have then the same codimension.

If V is extended to a member of $\mathcal{B}(H)$ by defining $Vy = 0$ for all $y \in \mathrm{Ran}(S)^\perp$, then V is called a partial isometry.

Every $T \in \mathcal{B}(H)$ thus has a factorization $T = VS$ in which S is positive and V is a partial isometry.

In combination with Exercise 2.5 the polar decomposition leads to an interesting result concerning similarity of normal operators. The following theorem thus asserts that similar normal operators are actually unitarily equivalent.

Theorem 2.57. *Suppose $M, N, T \in \mathcal{B}(H), M$ and N are normal, T is invertible, and*

$$M = TNT^{-1}$$

If $T = US$ is the polar decomposition of T, then

$$M = UNU^{-1}$$

Proof. By hypotheses, we have $M = (US)N(US)^{-1} = USNS^{-1}U^{-1}$. It suffices to show that N and $S = \sqrt{T^*T}$ commutes. To this end, we show that N and T^*T commutes, then by Remark 2.5 the desired result holds.

Since $MT = TN$ we have $T^*TN = T^*MT$. On the other hand, by Exercise 2.5, we have $M^*T = TN^*$, Consequently,

$$T^*M = (M^*T)^* = (TN^*)^* = NT^*$$

and hence $NT^*T = T^*M$ as required. $\square$

Chapter 3

Operator Semigroups

Introduction

Generally speaking, a *dynamical system* is a family $(T(t))_{t\geq 0}$ of mappings on a set X satisfying

$$\begin{cases} T(t+s) = T(t)T(s) \text{ for all } t,s \geq 0 \\ T(0) = I \end{cases}$$

where I is the identity mapping on X. Here X is viewed as the set of all states of a system, $t \in \mathbb{R}_+ := [0,\infty)$ as time and $T(t)$ as the map describing the change of a state $x \in X$ at time 0 into the state $T(t)x$ at time t. In the linear context, the state space X is a vector space, each $T(t)$ is a linear operator on X, and $(T(t))_{t\geq 0}$ is called a (one-parameter) *operator semigroup.*

The standard situation in which such operator semigroups naturally appear are so-called *Abstract Cauchy Problems* (ACP).

$$\begin{cases} u'(t) = Au(t) \quad \text{for } t \geq 0 \\ u(0) = x \end{cases},$$

where $(A, D(A))$ is a linear operator on a Banach space X. Here, the problem consists in finding a differentiable function $u : \mathbb{R}_+ \to X$ such that (ACP) holds. If for each initial value $x \in X$ a unique solution $u(\cdot, x)$ exists, then

$$T(t)x := u(t, x), \quad t \geq 0, x \in X$$

defines an operator semigroup.

For the "working mathematician," (ACP) is the problem, and $(T(t))_{t\geq 0}$ the solution to be found. The opposite point of view also makes sense: given an operator semigroup (i.e., a dynamical system) $(T(t))_{t\geq 0}$, under what conditions can it be "described" by a differential equation (ACP), and how can the operator A be found?

In some simple and concrete situations the relation between $(T(t))_{t\geq 0}$ and A is given by the formulas

$$T(t) = \mathrm{e}^{tA} \quad \text{and} \quad A = \frac{d}{dt}T(t)\bigg|_{t=0} .$$

In general, a comparably simple relation seems to be out of reach. However, miraculously as it may seem, a simple continuity assumption on the semigroup produces, in the usual Banach space setting, a rich and beautiful theory with a broad and almost universal field of applications. It is the aim of this chapter to develop this theory.

3.1 Strongly Continuous Semigroups

In this chapter we always assume that X is a Banach space over the scalar field $\mathbb{F}$. The following is our basic definition.

Definition 3.1. A family $(T(t))_{t\geq 0}$ of bounded linear operators on X is called a *strongly continuous (one-parameter) semigroup* if it satisfies the functional equation

$$\begin{cases} T(t+s) = T(t)T(s) & \text{for all } t, s \geq 0 \\ T(0) = I \end{cases} \tag{FE}$$

and is strongly continuous in the following sense: For every $x \in X$ the orbit maps

$$\xi_x : t \mapsto \xi_x(t) := T(t)x \tag{SC}$$

are continuous from $\mathbb{R}_+$ into X.

Remark 3.1. The property (SC) can also be expressed by saying that the map $t \mapsto T(t)$ is continuous from $\mathbb{R}_+$ into the space $(\mathcal{B}(X), \mathcal{T}_s)$ of all bounded operators on X endowed with the strong operator topology $\mathcal{T}_s$.

If these properties hold for $\mathbb{R}$ instead of $\mathbb{R}_+$, we call $(T(t))_{t\in\mathbb{R}}$ a *strongly continuous (one-parameter) group* on X.

3.1.1 Basic Properties

Our first goal is to facilitate the verification of the strong continuity (SC) required in Definition 3.1. This is possible thanks to the uniform boundedness principle.

As an easy consequence of this lemma, in combination with the functional equation (FE), we obtain that the continuity of the orbit maps

$$\xi_x : t \mapsto T(t)x$$

at each $t \geq 0$ and for each $x \in X$ is already implied by much weaker properties.

Proposition 3.1. *For a semigroup $(T(t))_{t\geq 0}$ on a Banach space X, the following assertions are equivalent.*

(a) $(T(t))_{t\geq 0}$ *is strongly continuous.*

(b) $(T(t))_{t\geq 0}$ *is strongly continuous at* 0*, in other words,* $\lim_{t\downarrow 0} T(t)x = x$ *for all* $x \in X$*.*

(c) There exist $t_0 > 0$*,* $M \geq 1$*, and a dense subset* $D \subset X$ *such that (ci)* $\|T(t)\| \leq M$ *for all* $t \in [0, t_0]$*; (cii)* $\lim_{t\downarrow 0} T(t)x = x$ *for all* $x \in D$*.*

Proof. Clearly (a) implies (b).

We claim that (b) and (c) are equivalent. Indeed if (c) holds, then by the Banach-Steinhaus theorem, we conclude that $T(t)$ converges strongly to I as $t \downarrow 0$. Thus (b) follows. Now if (b) holds, we claim that for each $t_0 > 0$, $\{T(t)\}_{0\leq t\leq t_0}$ is uniformly bounded. To see this, fix $x \in X$. Since $T(t) \to x$ as $t \downarrow 0$, there is $\delta_x > 0$ so that

$$\|T(t)x\| \leq \|x\| + 1 \text{ for all } 0 \leq t \leq \delta_x\,.$$

For each $t \in [0, t_0]$, there are $n \in \mathbb{N}$ and $0 \leq r < \delta_x$ so that $t = n\delta_x + r$. Then

$$\|T(t)x\| \leq \|T(\delta_x)\|^n \|T(r)x\| \leq \|T(\delta_x)\|^n (\|x\| + 1)\,.$$

Therefore we conclude that $\sup_{t\in[0,t_0]} \|T(t)x\| < \infty$. By the PUB, $\{T(t)\}_{0\leq t\leq 1}$ is uniformly bounded.

It remains to show that (b) implies (a). Assume (b) holds, fix $x \in X$ and $t_0 > 0$. We will show that $t \to T(t)x$ is continuous at t_0. Then (a) follows. The right continuity is easy: we compute

$$\lim_{h\downarrow 0} \|T(t_0 + h)\,x - T(t_0)\,x\| \leq \|T(t_0)\| \cdot \lim_{h\downarrow 0} \|T(h)x - x\| = 0\,.$$

For the right continuity, we have the estimate

$$\|T(t_0 - h)\,x - T(t_0)\,x\| \leq \|T(t_0 - h)\| \cdot \|x - T(h)x\|, \text{ for } h \geq 0\,.$$

This implies the left continuity since $\{T(t)\}_{0\leq t\leq t_0}$ remains uniformly bounded for $t \in [0, t_0]$ by the preceding argument. □

Remark 3.2. Because in many cases the uniform boundedness of the operators $T(t)$ for $t \in [0, t_0]$ is obvious, one obtains strong continuity by checking (right) continuity of the orbit maps ξ_x at $t = 0$ for a dense set of "nice" elements $x \in X$ only. We demonstrate the advantage of this procedure in the examples discussed below.

The following surprising result, asserts that if we use the weak operator topology instead of the strong operator topology in Remark 3.1 will not change our class of semigroups.

Theorem 3.2. *A semigroup $(T(t))_{t\geq 0}$ on a Banach space X is strongly continuous if and only if it is weakly continuous, i.e., if the mappings*

$$t \mapsto \langle T(t)x, x^* \rangle \;\; ; \;\; [0,\infty) \to \mathbb{F}$$

are continuous for each $x \in X, x^* \in X^*$.

Proof. We have only to show that weak continuity implies strong continuity. As a first step, for any $t_0 > 0$, $(T(t) : t \in [0,t_0])$ is bounded for the weak operator topology. By PUB, it is uniformly bounded. Using Proposition 3.1 (c), it suffices to show that

$$E := \left\{ x \in X : \lim_{t \downarrow 0} \|T(t)x - x\| = 0 \right\}$$

is a (strongly) dense subspace of X.

Fix $x \in X$ and $r > 0$. By Remark 3.1, since $t \mapsto T(t)x$ is a continuous mapping $[0,r]$ into (X, τ_w),the integral

$$x_r = \frac{1}{r} \int_0^r T(s)x \,\mathrm{d}s \in X$$

is well-defined in the sense that for each $x^* \in X^*$,

$$\langle x_r, x^* \rangle = \frac{1}{r} \int_0^r \langle T(s)x, x^* \rangle \,\mathrm{d}s\,.$$

Then by Exercise **??**, for each $t \geq 0$ we have

$$T(t)x_r = \frac{1}{r} \int_0^r T(s+t)x \,\mathrm{d}s = \frac{1}{r} \int_t^{r+t} T(s)x \,\mathrm{d}s\,.$$

We compute

$$
\begin{aligned}
\|T(t)x_r - x_r\| &= \sup_{\|x^*\|\le 1} |\langle T(t)x_r, x^*\rangle - \langle x_r, x^*\rangle| \\
&= \sup_{\|x^*\|\le 1} \left| \frac{1}{r}\int_t^{r+t} \langle T(s)x, x^*\rangle \,\mathrm{d}s - \frac{1}{r}\int_0^r \langle T(s)x, x^*\rangle \,\mathrm{d}s \right| \\
&\le \sup_{\|x^*\|\le 1} \left(\left| \frac{1}{r}\int_r^{r+t} \langle T(s)x, x^*\rangle \,\mathrm{d}s \right| + \left| \frac{1}{r}\int_0^t \langle T(s)x, x^*\rangle \,\mathrm{d}s \right| \right) \\
&\le \frac{2t}{r}\|x\| \sup_{0\le s\le r+t} \|T(s)\| \to 0 \quad \text{as } t\downarrow 0\,.
\end{aligned}
$$

In other words $\lim_{t\downarrow 0} T(t)x_r = x_r$. Now let $D = \{x_r : x \in X, r > 0\}$. Then $D \subset E$.

It remains to show that E is dense in X. Suppose for contradiction that E is not dense in X. Note that E is a linear subspace of X, then by Corollary **??**, there is a nonzero $x^* \in X^*$ so that $x^* \in E^\perp$. Since $D \subset E$, then for each fixed $x \in X$, we have

$$
\langle x_r, x^*\rangle = \frac{1}{r}\int_0^r \langle T(s)x, x^*\rangle \,\mathrm{d}s = 0 \qquad \text{for all} \ \ r > 0\,.
$$

Letting $r \downarrow 0$, since $s \mapsto \langle T(s)x, x^*\rangle$ is continuous, we conclude that $\langle x, x^*\rangle = 0$. As x is arbitrary, there must be the case that $x^* = 0$, which is a contradiction. Thus E is (strongly) dense in X as required.

Indeed there is an much easier proof for the (strong) density of E. Observe that D is weakly dense in X, we conclude that E is weakly dense in X. Since E is convex, by Theorem **??**, E is strongly dense in X. □

Exercise 3.1. Show that the semigroup $\{T(t)\}_{t\ge 0}$ is strongly continuous if and only if $t \mapsto T(t)$ from $\mathbb{R}_+$ into $(\mathcal{B}(X), \mathcal{T}_w)$ is weakly continuous at 0.

Exercise 3.2. Let X be a Banach space and let $(T(t))_{t\ge 0}$ be a strongly continuous semigroup. Then the following holds.

(a) The operator $T(t)$ is injective for some $t > 0$ if and only if it is injective for all $t > 0$.

(b) The operator $T(t)$ is surjective for some $t > 0$ if and only if it is surjective for all $t > 0$.

(c) The operator $T(t)$ has a dense image for some $t > 0$ if and only if it has a dense image for all $t > 0$.

Exercise 3.3. Let $(T(t))_{t\geq 0}$ be a semigroup on X. Suppose that $T(t)$ is bijective for some, and hence all, $t > 0$. Let $T(-t) = T(t)^{-1}$ for $t \geq 0$. Then $(T(t))_{t\in\mathbb{R}}$ is a group on X.

3.1.2 Growth Bound

We repeat that for a strongly continuous semigroup $(T(t))_{t\geq 0}$ the finite orbits

$$\{T(t)x : t \in [0, t_0]\}$$

are continuous images of a compact interval, hence compact and therefore bounded for each $x \in X$. So by the uniform boundedness principle each strongly continuous semigroup is uniformly bounded on each compact interval, a fact that implies exponential boundedness on $\mathbb{R}_+$.

Proposition 3.3. *For every strongly continuous semigroup* $(T(t))_{t\geq 0}$,

$$\lim_{t\downarrow 0} \frac{\log \|T(t)\|}{t} = \inf_{t>0} \frac{\log \|T(t)\|}{t} := \omega_0 \in [-\infty, \infty)\,. \tag{3.1}$$

Moreover, we have

$$\omega_0 = \inf\left\{ w \in \mathbb{R} : \begin{array}{c} \textit{there exists } M_w \geq 1 \textit{ such that} \\ \|T(t)\| \leq M_w \mathrm{e}^{wt} \textit{ for all } t \geq 0 \end{array} \right\} \tag{3.2}$$

is called the growth bound (or type) of $\{T(t)\}_{t\geq 0}$.

Proof. Observe that for each $t, s \geq 0$, we have

$$\log \|T(s+t)\| \leq \log \|T(s)\| + \log \|T(t)\|\,.$$

Now fix $t_0 > 0$, for each $t > 0$, there is unique $n \in \mathbb{N}$ and $0 \le r < t_0$ so that $t = nt_0 + r$. Then

$$\frac{\log \|T(t)\|}{t} \le \frac{n}{t} \log \|T(t_0)\| + \frac{\log \|T(r)\|}{t} \,.$$

Letting $t \to \infty$ and noting that $t/n \to t_0$, we get

$$\limsup_{t \uparrow \infty} \frac{\log \|T(t)\|}{t} \le \frac{\log \|T(t_0)\|}{t_0} \,.$$

Since t_0 is arbitrary, we get

$$\limsup_{t \uparrow \infty} \frac{\log \|T(t)\|}{t} \le \inf_{t_0 > 0} \frac{\log \|T(t_0)\|}{t_0} \le \liminf_{t \uparrow \infty} \frac{\log \|T(t)\|}{t} \,,$$

and then (3.1) follows. (3.2) is an easy consequence of (3.1). □

It becomes clear in the discussion below, but is presently left as a challenge to the reader that

- $\omega_0 = -\infty$ may occur,
- the infimum in (3.2) may not be attained; i.e, it might happen that no constant M exists such that $\|T(t)\| \le M \mathrm{e}^{\omega_0 t}$ for all $t \ge 0$, and
- Constants $M > 1$ may be necessary; i.e., no matter how large $w \ge \omega_0$ is chosen, $\|T(t)\|$ will not be dominated by e^{wt} for all $t \ge 0$.

Definition 3.2. A semigroup $(T(t))_{t \ge 0}$ is called *bounded* if there is $M > 0$ so that

$$\|T(t)\| \le M \qquad \text{for all } \; t \ge 0\,;$$

and *contractive* if $M = 1$ is possible. Finally, $(T(t))_{t \ge 0}$ is called *isometric* if $\|T(t)x\| = \|x\|$ for all $t \ge 0$ and $x \in X$.

3.2 Uniformly Continuous Semigroups

In order to create a feeling for the concepts introduced so far, we discuss first the case in which the semigroup $(T(t))_{t\geq 0}$ can be represented as an operator-valued exponential function $\left(\mathrm{e}^{tA}\right)_{t\geq 0}$. Due to this representation, we later consider this case as rather trivial.

For $A \in \mathcal{B}(X)$ we define

$$\mathrm{e}^{tA} := \sum_{n=0}^{\infty} \frac{t^n A^n}{n!} \tag{3.3}$$

for each $t \geq 0$ or $t \in \mathbb{R}$. It follows from the completeness of X that e^{tA} is a well-defined bounded operator on X.

Proposition 3.4. *$\left(\mathrm{e}^{tA}\right)_{t\geq 0}$ is a semigroup on X such that $t \mapsto \mathrm{e}^{tA}$; $\mathbb{R}_+ \to (\mathcal{B}(X), \|\cdot\|)$ is continuous.*

Proof. Because the series $\sum_{k=0}^{\infty} t^k \|A\|^k / k!$ converges, one can show, as for the Cauchy product of scalar series, that

$$\begin{aligned}
\sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \cdot \sum_{k=0}^{\infty} \frac{s^k A^k}{k!} &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{t^{n-k} A^{n-k}}{(n-k)!} \cdot \frac{s^k A^k}{k!} \\
&= \sum_{n=0}^{\infty} \frac{(t+s)^n A^n}{n!} .
\end{aligned}$$

This proves that $(\mathrm{e}^{tA})_{t\geq 0}$ is a semigroup. In order to show that $t \mapsto \mathrm{e}^{tA}$ is continuous, we first observe that

$$\mathrm{e}^{(t+h)A} - \mathrm{e}^{tA} = \mathrm{e}^{tA}\left(\mathrm{e}^{hA} - I\right)$$

for all $t, h \in \mathbb{R}$. Therefore, it suffices to show that $\lim_{h\to 0} \mathrm{e}^{hA} = I$. This follows from the estimate

$$\begin{aligned}
\left\|\mathrm{e}^{hA} - I\right\| &= \left\| \sum_{k=1}^{\infty} \frac{h^k A^k}{k!} \right\| \\
&\leq \sum_{k=1}^{\infty} \frac{|h|^k \cdot \|A\|^k}{k!} = \mathrm{e}^{|h|\cdot\|A\|} - 1 .
\end{aligned}$$

$\square$

Remark 3.3. In fact, there is no need to restrict the (time) parameter t to $\mathbb{R}_+$. The definition, the continuity, and the functional equation hold for any real and even complex t. Then the map

$$T(\cdot) : t \mapsto \mathrm{e}^{tA}$$

extends to a continuous homomorphism from the additive group $(\mathbb{R}, +)$ into the multiplicative group $\mathscr{G}$ of all invertible elements in the Banach algebra $\mathcal{B}(X)$. We call $\left(\mathrm{e}^{tA}\right)_{t\in\mathbb{R}}$ the (one-parameter) group generated by A.

Semigroups having the continuity property stated in the preceding Proposition are called uniformly continuous. Specifically:

Definition 3.3. Semigroup $\{T(t)\}_{t\geq 0}$ on X is called uniformly continuous if $t \mapsto T(t)$ is a continuous mapping from $\mathbb{R}_+$ into $(\mathcal{B}(X), \|\cdot\|)$.

Exercise 3.4. Show that the semigroup $\{T(t)\}_{t\geq 0}$ is uniformly continuous if and only if $\|T(t) - I\| \to 0$ as $t \downarrow 0$.

Proposition 3.5. *Let $A \in \mathcal{B}(X)$. Then the map $t \mapsto T(t) \coloneqq \mathrm{e}^{tA}$; $\mathbb{R}_+ \to (\mathcal{B}(X), \|\cdot\|)$ is differentiable and satisfies the differential equation*

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}T(t) = AT(t) \text{ for } t \geq 0\,, \\ T(0) = I\,. \end{cases} \tag{3.4}$$

Conversely, every differentiable function $T(\cdot) : \mathbb{R}_+ \to (\mathcal{B}(X), \|\cdot\|)$ satisfying (3.4) *is already of the form $T(t) = \mathrm{e}^{tA}$ for $A = \frac{\mathrm{d}}{\mathrm{d}t}T(0) \in \mathcal{B}(X)$.*

Proof. We only show that $T(\cdot)$ satisfies (3.4). Because the functional equation implies that, for all $t, h \in \mathbb{R}$,

$$\frac{T(t+h) - T(t)}{h} = \frac{T(h) - I}{h} \cdot T(t)\,.$$

(3.4) is proved if $\lim_{h\to 0}\frac{T(h)-I}{h}=A$. This, however, follows because

$$\begin{aligned}\left\|\frac{T(h)-I}{h}-A\right\| &\leq \sum_{k=2}^{\infty}\frac{|h|^{k-1}\cdot\|A\|^k}{k!}\\ &= \frac{\mathrm{e}^{|h|\cdot\|A\|}-1}{|h|}-\|A\|\to 0 \quad \text{as } h\to 0\end{aligned}$$

Conversely, if $\{T(t)\}_{t\geq 0}$ solves the differential equation, then we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(T(t)e^{-tA}\right) &= \left(\frac{\mathrm{d}}{\mathrm{d}t}T(t)\right)e^{-tA}+T(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}e^{-tA}\right)\\ &= T(t)Ae^{-tA}+T(t)(-A)e^{-tA}=0\,.\end{aligned}$$

Combine this with the initial value condition, we then get that

$$T(t)e^{-tA}=I \quad \text{for all } t\geq 0\,.$$

Then the desired result follows. □

Theorem 3.6. *Let $(T(t))_{t\geq 0}$ be a uniformly continuous semigroup on X. Then $t\mapsto T(t)$ is a differentiable mapping from $\mathbb{R}_+$ into $(\mathcal{B}(X),\|\cdot\|)$, with the form*

$$T(t)=\mathrm{e}^{tA}, \quad t\geq 0\,,$$

where $A=\frac{\mathrm{d}}{\mathrm{d}t}T(0)\in\mathcal{B}(X)$.

Proof. By the uniform continuity of $s\mapsto T(s)$, for each $h>0$, the integral

$$V(t):=\int_0^t T(s)\,\mathrm{d}s\in\mathcal{B}(X)$$

is well-defined by Theorem **??**. Besides, we have

$$\frac{1}{t}V(t)\to I \ \text{ in } (\mathcal{B}(X),\|\cdot\|) \quad \text{as } t\downarrow 0\,.$$

Thus for sufficiently small $h>0$, $V(h)$ is invertible in $\mathcal{B}(X)$.

By Exercise **??**, we have

$$
\begin{aligned}
(T(t)-I)V(h) &= \int_0^h T(t)T(s)\,\mathrm{d}s - \int_0^h T(s)\,\mathrm{d}s \\
&= \int_t^{h+t} T(s)\,\mathrm{d}s - \int_0^h T(s)\,\mathrm{d}s = \int_h^{h+t} T(s)\,\mathrm{d}s - \int_0^t T(s)\,\mathrm{d}s \\
&= \int_0^t (T(h)-I)T(s)\,\mathrm{d}s\,.
\end{aligned}
$$

Since $V(h)$ is invertible, by Exercise **??**, we have

$$
T(t)-I = \int_0^t T(s)(T(h)-I)V(h)^{-1}\,\mathrm{d}s\,.
$$

Thus we can see that $t \mapsto T(t)$ is a differentiable mapping from $\mathbb{R}_+$ into $\mathcal{B}(X)$. Let $A = (T(h)-I)V(h)^{-1} \in \mathcal{B}(X)$. By Exercise **??** again, one can see that A and $(T(t))_{t\geq 0}$ commutes. Then we conclude that

$$
\frac{\mathrm{d}}{\mathrm{d}t}T(t) = AT(t) = AT(t)\,.
$$

By Proposition 3.5, the desired result follows. □

Remark 3.4. Because definition for e^{tA} works also for $t \in \mathbb{R}$, it follows that each uniformly continuous semigroup can be extended to a uniformly continuous group $\left(\mathrm{e}^{tA}\right)_{t\in\mathbb{R}}$.

From the differentiability of $t \mapsto T(t)$ it follows that for each $x \in X$ the orbit map $\mathbb{R}_+ \ni t \mapsto T(t)x \in X$ is differentiable as well. Therefore, the map $\xi_x(t) := T(t)x$ is the unique solution of the abstract Cauchy problem

$$
\begin{cases}
u'(t) = Au(t)\ , \text{ for } t \geq 0\,; \\
u(0) = x\,.
\end{cases}
$$

If we set $X = \mathbb{C}^n$ and A is a $n \times n$ complex matrix, this is what we have learned in the course of ODE.

3.3 Generators and Resolvents

We recall that for a one-parameter semigroup $(T(t))_{t\geq 0}$ on a Banach space X uniform continuity implies differentiability of the map $t \mapsto T(t); \mathbb{R}_+ \to (\mathcal{B}(X), \|\cdot\|)$. The derivative of $T(\cdot)$ at $t = 0$ then yields a bounded operator A for which $T(t) = \mathrm{e}^{tA}$ for all $t \geq 0$.

We now hope that strong continuity of a semigroup $(T(t))_{t\geq 0}$ still implies some differentiability of the orbit maps

$$\xi_x : t \mapsto T(t)x \ ; \ \mathbb{R}_+ \to X \, .$$

In order to pursue this idea we first show, in analogy to Proposition 3.1 and Exercise 3.1 that differentiability of ξ_x is already implied by the differentiability at $t = 0$.

Lemma 3.7. *Let $(T(t))_{t\geq 0}$ be a strongly continuous semigroup on X and fix $x \in X$. Then the orbit map ξ_x differentiable on $\mathbb{R}_+$ if and only if it is differentiable at $t = 0$.*

Proof. If ξ_x is differentiable at 0, then for fixed $t_0 > 0$,

$$\lim_{h\to 0} \frac{T(t_0+h)x - T(t_0)x}{h} = T(t_0) \lim_{h\to 0} \frac{T(h)x - x}{h} = T(t_0)\xi_x'(0).$$

Thus ξ_x is right differentiable at t_0. On the other hand, by that the uniform boundedness of $\{T(t)\}_{0\leq t\leq t_0}$, we compute

$$\begin{aligned}
&\lim_{h\to 0} \frac{T(t_0)x - T(t_0-h)x}{h} = T(t_0-h) \lim_{h\to 0} \frac{T(h)x - x}{h} \\
&= T(t_0)\xi_x'(0) + [T(t_0-h) - T(t_0)] \lim_{h\to 0} \frac{T(h)x - x}{h} \\
&= T(t_0)\xi_x'(0) + [T(t_0-h) - T(t_0)] \lim_{h\to 0} [\frac{T(h)x - x}{h} - \xi_x'(0)] \\
&= T(t_0)\xi_x'(0) \, .
\end{aligned}$$

Then the desired result holds. □

On the subspace of X consisting of all those x for which the orbit maps ξ_x are differentiable, the right derivative at $t = 0$ then yields an operator A from which we obtain, in a sense to be specified later, the operators $T(t)$ as the "exponentials e^{tA}". This is already expressed in the choice of the term "generator" in the following definition.

In the following, we set $A_h := \frac{T(h)-I}{h}$ for $h > 0$. Then clearly $A_h \in \mathcal{B}(X)$.

Definition 3.4. The *generator* $(A, D(A))$ for strongly continuous semigroup $(T(t))_{t\geq 0}$ is the operator given by

$$Ax = \lim_{h\downarrow 0} A_h x = \lim_{h\downarrow 0} \frac{T(h)x - x}{h} = \xi_x'(0)\,;$$
$$D(A) = \{x : \xi_x \text{ is differentiable}\} = \{x : \lim_{h\downarrow 0} A_h x \text{ exists }\}\,.$$

It's easy to see that $(A, D(A))$ is a linear operator. To ensure that the operator $(A, D(A))$ has reasonable properties, we proceed as in Theorem 3.2 and Theorem 3.6, we need to look at "average" elements of the orbit map:

$$x_t := \frac{1}{t}\int_0^t \xi_x(s)\,\mathrm{d}s = \frac{1}{t}\int_0^t T(s)x\,\mathrm{d}s \qquad \text{for } x \in X, t > 0\,.$$

Thus first of all we discuss the integral $\int_0^t T(s)x\,\mathrm{d}s$.

Lemma 3.8. *For the generator $(A, D(A))$ of a strongly continuous semigroup $(T(t))_{t\geq 0}$, the following properties hold.*

(i) For every $t \geq 0$ and $x \in X$, one has $\int_0^t T(s)x\,\mathrm{d}s \in D(A)$ and

$$T(t)x - x = A\int_0^t T(s)x\,\mathrm{d}s\,. \tag{3.5}$$

(ii) If $x \in D(A)$, then $T(t)x \in D(A)$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}T(t)x = T(t)Ax = AT(t)x \qquad \textit{for all } t \geq 0.$$

Particularly,

$$T(t)x - x = \int_0^t T(s)Ax\,\mathrm{d}s \qquad \textit{for all } t \geq 0. \tag{3.6}$$

(iii) Let x, y in X. If

$$T(t)x - x = \int_0^t T(s)y \,\mathrm{d}s \quad \textit{for all } t \geq 0,$$

then we have $x \in D(A)$ and $y = Ax$.

Proof. To show part (i), note that

$$\begin{aligned} A_h \int_0^t T(s)x \,\mathrm{d}s &= \frac{1}{h}\left(\int_0^t T(s+h)x \,\mathrm{d}s - \int_0^t T(s)x \,\mathrm{d}s\right) \\ &= \frac{1}{h}\int_h^{t+h} T(s)x \,\mathrm{d}s - \frac{1}{h}\int_0^t T(s)x \,\mathrm{d}s \\ &= \frac{1}{h}\int_t^{t+h} T(s)x \,\mathrm{d}s - \frac{1}{h}\int_0^h T(s)x \,\mathrm{d}s\,. \end{aligned}$$

Thus letting $h \downarrow 0$, by Exercise **??** and the strong continuity, we get

$$\lim_{h \downarrow 0} A_h \int_0^t T(s)x \,\mathrm{d}s = T(t)x - x\,.$$

Hence part (i) follows.

To show part (ii), note that $A_h T(t) = T(t) A_h$ for all $h, t > 0$. Thus for $x \in D(A)$, we have

$$\lim_{h \downarrow 0} A_h T(t)x = T(t) \lim_{h \downarrow 0} A_h x = T(t)Ax\,.$$

By the definition of $(A, D(A))$, $T(t)x \in D(A)$ and $AT(t)x = T(t)Ax$. Since $x \in D(A)$, $t \mapsto T(t)x$ is differentiable on $\mathbb{R}_+$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} T(t)x = \lim_{h \downarrow 0} \frac{T(t+h)x - T(t)x}{h} = \lim_{h \downarrow 0} T(t)A_h x = T(t)Ax\,.$$

Then by Newton-Leibniz formula formula,

$$T(t)x - x = \int_0^t T(s)Ax \,\mathrm{d}s\,.$$

Indeed we can use (3.5) to show (3.6). Since A and $(T(t))_{t \geq 0}$ commute on $D(A)$, it's easy to see that for $x \in D(A)$, $t \mapsto T(t)x$ is a continuous mapping form $\mathbb{R}_+$ into $(D(A), \|\cdot\|_A)$. Thus by Exercise **??**, since A is a bounded linear operator on $(D(A), \|\cdot\|_A)$, we conclude that

$$\int_0^t T(t)Ax \,\mathrm{d}s = A\int_0^t T(t)x \,\mathrm{d}s \,.$$

as desired.

To show part (iii), since $s \mapsto T(s)y$ is continuous, we have

$$\lim_{h\downarrow 0} A_h x = \lim_{h\downarrow 0} \frac{1}{h}\int_0^h T(s)y \,\mathrm{d}s = y \,.$$

Thus by definition $x \in D(A)$ and $Ax = y$. $\square$

Part (iii) is an easier way to show some element x in X belongs to $D(A)$. For example, we shall use it to prove that for a semigroup, weak differentiability is equivalent to the strong differentiability:

Proposition 3.9. *Let $(T(t))_{t\geq 0}$ be a strongly continuous semigroup on X with generator $(A, D(A))$. Let x, y in X. If*

$$w\text{-}\lim_{h\downarrow 0} \frac{T(h)x - x}{h} = y \,.$$

Then $x \in D(A)$ and $Ax = y$.

Proof. By the same argument in the proof of Lemma 3.7, we conclude that the orbits map $\xi_x : t \mapsto T(t)x$ is weakly differentiable, i.e., for each $x^* \in X^*$, $t \mapsto \langle T(t)x, x^*\rangle$ is a differentiable mapping from $\mathbb{R}_+$ into the scalar field of X, with

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle T(t)x, x^*\rangle = \langle T(t)y, x^*\rangle \qquad \text{for all } t \geq 0 \,.$$

By Lemma 3.8, it suffices to show that

$$T(t)x - x = \int_0^t T(s)y \,\mathrm{d}s$$

for all $t \geq 0$. Observe that for all $x^* \in X^*$

$$\begin{aligned}\langle T(t)x - x, x^*\rangle = \langle T(t)x, x^*\rangle - \langle x, x^*\rangle &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\langle T(s)x, x^*\rangle \,\mathrm{d}s \\ &= \int_0^t \langle T(s)y, x^*\rangle \,\mathrm{d}s = \langle \int_0^t T(s)y \,\mathrm{d}s, x^*\rangle \,.\end{aligned}$$

then the desired result follows. $\square$

Lemma 3.10. *Let $J \subset \mathbb{R}$ be an interval and $P, Q : J \to \mathcal{B}(X)$ be two strongly continuous operator-valued functions. If $P(\cdot)x : J \to X$ and $Q(\cdot)x : J \to X$ are differentiable for all $x \in D$ for some subspace D of X and D invariant under Q. Then $(PQ)(\cdot)x : J \to X; t \mapsto P(t)Q(t)x$ is differentiable for every $x \in D$ and for $t_0 \in J$,*

$$\frac{\mathrm{d}}{\mathrm{d}t}(P(\cdot)Q(\cdot)x)(t_0) = \frac{\mathrm{d}}{\mathrm{d}t}\left(P(\cdot)Q\left(t_0\right)x\right)\left(t_0\right) + P\left(t_0\right)\left(\frac{\mathrm{d}}{\mathrm{d}t}Q(\cdot)x\right)\left(t_0\right)$$

Proof of Lemma 3.10. Fix $x \in D$ and $t_0 \in J$. For $h \in \mathbb{R}$ with $t_0 + h \in J$, note that

$$\begin{aligned}
&\frac{P\left(t_0+h\right)Q\left(t_0+h\right)x - P\left(t_0\right)Q\left(t_0\right)x}{h} \\
&= P\left(t_0+h\right)\frac{Q\left(t_0+h\right)x - Q\left(t_0\right)x}{h} + \frac{P\left(t_0+h\right) - P\left(t_0\right)}{h}Q\left(t_0\right)x
\end{aligned}$$

On the one hand, since D is invariant under Q, we have $Q(t_0)x \in D$ and hence

$$\frac{P\left(t_0+h\right) - P\left(t_0\right)}{h}Q\left(t_0\right)x \to \frac{\mathrm{d}}{\mathrm{d}t}\left(P(\cdot)Q\left(t_0\right)x\right)\left(t_0\right).$$

as $h \to 0$ with $t_0 + h \in J$. On the other hand, we compute

$$\begin{aligned}
&P\left(t_0+h\right)\frac{Q\left(t_0+h\right)x - Q\left(t_0\right)x}{h} \\
&= P\left(t_0\right)\frac{Q\left(t_0+h\right)x - Q\left(t_0\right)x}{h} + \left[P\left(t_0+h\right) - P(t_0)\right]\left(\frac{\mathrm{d}}{\mathrm{d}t}Q(\cdot)x\right)\left(t_0\right) \\
&+ \left[P\left(t_0+h\right) - P(t_0)\right]\left[\frac{Q\left(t_0+h\right)x - Q\left(t_0\right)x}{h} - \left(\frac{\mathrm{d}}{\mathrm{d}t}Q(\cdot)x\right)\left(t_0\right)\right];
\end{aligned}$$

As $h \to 0$ with $t_0 + h \in J$, since $x \in D$ we have

$$P\left(t_0\right)\frac{Q\left(t_0+h\right)x - Q\left(t_0\right)x}{h} \to P\left(t_0\right)\left(\frac{\mathrm{d}}{\mathrm{d}t}Q(\cdot)x\right)\left(t_0\right);$$

since $P(\cdot) : J \mapsto \mathcal{B}(X)$ are strongly continuous, we have

$$\left[P\left(t_0+h\right) - P(t_0)\right]\left(\frac{\mathrm{d}}{\mathrm{d}t}Q(\cdot)x\right)\left(t_0\right) \to 0;$$

since the strong continuity of $P(\cdot)$ implies the uniform boundeness of $P(\cdot)$ on compact intervals and $x \in D$ we have

$$[P(t_0+h) - P(t_0)]\left[\frac{Q(t_0+h)\,x - Q(t_0)\,x}{h} - \left(\frac{\mathrm{d}}{\mathrm{d}t}Q(\cdot)x\right)(t_0)\right] \to 0\,.$$

Then the desired result follows. □

With the help of Lemma 3.8 and Lemma 3.10, we now show that the generator, although unbounded in general, has nice properties.

Theorem 3.11. *The generator $(A, D(A))$ of a strongly continuous semigroup $(T(t))_{t\geq 0}$ is a densely defined closed linear operator that determines the semigroup uniquely.*

Proof. Firstly, we show that $D(A)$ is dense in X. For each fixed $x \in X$, consider

$$x_t := \frac{1}{t}\int_0^t T(s)x\,\mathrm{d}s \qquad \text{for } t > 0\,.$$

Then by Lemma 3.8, $x_t \in D(A)$. By the strong continuity of $(T(t))_{t\geq 0}$, $x_t \to= x$ in X as $t \downarrow 0$. Then the density of $D(A)$ follows.

To show that A is closed, assume that (x_n) is a sequence in $D(A)$ and $x_n \to x$, $Ax_n \to y$. By Lemma 3.8, for each n we have

$$T(t)x_n - x_n = \int_0^t T(s)Ax_n\,\mathrm{d}s\,.$$

Then letting $n \to \infty$ we get,

$$T(t)x - x = \int_0^t T(s)y\,\mathrm{d}s\,.$$

By Exercise 3.8 we conclude that $x \in D(A)$ and $y = Ax$ as required.

Finally we show that the generator determines the semigroup uniquely. Suppose that $(S(t))_{t\geq 0}$ is also a strongly continuous semigroup with generator $(A, D(A))$. We show that $T(t) = S(t)$ for all $t > 0$. To this end, since $T(t), S(t) \in \mathcal{B}(X)$ and $D(A)$ is dense in X, it suffices to show that

$$T(t)x = S(t)x \qquad \text{for all } x \in D(A)\,.$$

Observe that, formally we have

$$T(t)x - S(t)x = T(t-s)S(s)x\,\Big|_{s=0}^{t} = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} T(t-s)S(s)x\,\mathrm{d}s\,, \tag{3.7}$$

Thus we shall consider the mapping $s \mapsto T(t-s)S(s)$ form $[0,t]$ into $\mathcal{B}(X)$. By Lemma 3.10, since $D(A)$ is invariant under $S(s)$, for all $x \in D(A)$, the orbits map $s \mapsto T(t-s)S(s)x$ is differentiable with

$$\frac{\mathrm{d}}{\mathrm{d}s} T(t-s)S(s)x = -T(t-s)AS(s)x + T(t-s)AS(s)x = 0 \qquad \text{for all } 0 \le s \le t\,.$$

Thus by (3.7) we get $T(t)x = S(t)x$ as required. $\square$

Combining these properties of the generator with the closed graph theorem gives a new characterization of uniformly continuous semigroups, thus complementing Theorem 3.6.

Corollary 3.12. *For a strongly continuous semigroup $(T(t))_{t\ge0}$ with generator $(A, D(A))$, the following assertions are equivalent.*

(a) The generator A is bounded; i.e., there exists $M > 0$ such that $\|Ax\| \le M\|x\|$ for all $x \in D(A)$.

(b) The domain $D(A)$ is all of X.

(c) The semigroup $(T(t))_{t\ge0}$ is uniformly continuous.

In each case, the semigroup is given by

$$T(t) = \mathrm{e}^{tA} := \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}, \quad t \ge 0\,.$$

3.3.1 The Core for the Generator

Property (b) of the Corollary 3.12 indicates that the domain of the generator contains important information about the semigroup and therefore has to be taken into account carefully. However, in many examples it is often routine to compute the expression Ax for some or even many elements in the domain $D(A)$, although it is difficult to identify $D(A)$ precisely. In these situations, the concept of core helps to distinguish between "small" and "large" subspaces of $D(A)$.

Recall that a subspace D of the domain $D(A)$ of is called a *core* for A if D is dense in $D(A)$ for the graph norm $\|x\|_A := \|x\| + \|Ax\|$. In this case, $(A|_D, D)$ is closable and its closure is exactly $(A, D(A))$. We now state a useful criterion for subspaces to be a core for the generator.

Lemma 3.13. *Let $(A, D(A))$ be the generator of a strongly continuous semigroup $(T(t))_{t\geq 0}$ on X. A subspace D of $D(A)$ that is dense in X and invariant under the semigroup $(T(t))_{t\geq 0}$ is always a core for A.*

Proof. Since A is closed, the $\|\cdot\|_A$-closure of D, namely $\overline{D}^{\|\cdot\|_A}$, is contained in $D(A)$. It remains to show that $D(A) \subset \overline{D}^{\|\cdot\|_A}$.

Fix $x \in D(A)$. As we have pointed in the proof of Lemma 3.8, since $AT(t)x = T(t)Ax$ for all $t \geq 0$, the orbit map $t \mapsto T(t)x \in (D(A), \|\cdot\|_A)$ is continuous. Thus $\int_0^t T(s)x \,\mathrm{d}s \in D(A)$ for each $t \geq 0$ and

$$\left\| x - \frac{1}{t}\int_0^t T(s)x \,\mathrm{d}s \right\|_A \to 0 \quad \text{ as } t \downarrow 0\,.$$

It follows that we have only to show that $\int_0^t T(s)x \,\mathrm{d}s \in \overline{D}^{\|\cdot\|_A}$ for all $t \geq 0$.

Since D is dense in X, we can find a sequence (x_n) in D such that $\|x_n - x\| \to 0$. By the hypothesis that D is invariant under $(T(t))_{t\geq 0}$, we have $T(t)x_n \in D$ for all $t \geq 0$. Then we deduce that the map $s \mapsto T(s)x_n \in (D, \|\cdot\|_A)$ is continuous and hence for each $t \geq 0$,

$$\int_0^t T(s)x_n \,\mathrm{d}s \in \overline{D}^{\|\cdot\|_A}\,.$$

Fix $t > 0$. We claim that

$$\int_0^t T(s)x_n \,\mathrm{d}s \to \int_0^t T(s)x \,\mathrm{d}s \;\text{ in }\; (D(A), \|\cdot\|_A)\,,$$

then the desired result follows.

Let $M = \sup\{\|T(s)\| : 0 \le s \le t\}$. By Lemma 3.8, we compute

$$
\begin{aligned}
&\left\| \int_0^t T(s)x_n \,\mathrm{d}s - \int_0^t T(s)x \,\mathrm{d}s \right\|_A \\
&\le tM\|x_n - x\| + \left\| A\int_0^t T(s)x_n \,\mathrm{d}s - A\int_0^t T(s)x \,\mathrm{d}s \right\| \\
&\le tM\|x_n - x\| + \|T(t)x_n - x_n - T(t)x + x\| \\
&\le (tM + M + 1)\|x_n - x\| \to 0 \quad \text{as} \quad n \to \infty\,.
\end{aligned}
$$

We are done. □

Important examples of cores are given by the domains $D\left(A^n\right)$ of the powers A^n of a generator A.

Theorem 3.14. *For the generator $(A, D(A))$ of a strongly continuous semi-group $(T(t))_{t\ge 0}$ the space*

$$
D\left(A^\infty\right) := \bigcap_{n\in\mathbb{N}} D\left(A^n\right)\,,
$$

hence each $D\left(A^n\right) := \left\{x \in D\left(A^{n-1}\right) : A^{n-1}x \in D(A)\right\}$, is a core for A.

Proof. Because the space $D\left(A^\infty\right)$ is a $(T(t))_{t\ge 0}$ -invariant subspace of $D(A)$, by the preceding lemma, it remains to show that it is dense in X. Let

$$
D := \left\{ x_\varphi := \int_0^\infty \varphi(s)T(s)x \,\mathrm{d}s : x \in X, \varphi \in C_c^\infty(\mathbb{R}) \text{ with } \operatorname{supp}(\varphi) \subset (0,\infty) \right\}.
$$

Note that since ϕ has compact support, x_φ is well-defined by Theorem **??**. We claim that (i) $D \subset D(A^\infty)$; (ii) D is dense in X. Then the desired result follows.

To demonstrate (i), note that for $h > 0$, by Exercise **??** we have

$$
\begin{aligned}
\frac{T(h) - I}{h} x_\varphi &= \frac{1}{h}\int_0^\infty \varphi(s)(T(s+h) - T(s))x \,\mathrm{d}s \\
&= \frac{1}{h}\int_h^\infty \varphi(s-h)T(s)x \,\mathrm{d}s - \frac{1}{h}\int_0^\infty \varphi(s)T(s)x \,\mathrm{d}s \\
&= \int_0^\infty \frac{1}{h}(\varphi(s-h) - \varphi(s))T(s)x \,\mathrm{d}s\,.
\end{aligned}
$$

Letting $h \downarrow 0$, by the dominated convergence theorem, we get $x_\varphi \in D(A)$ and

$$Ax_\varphi = -\int_0^\infty \varphi'(s)T(s)x\,\mathrm{d}s = x_{\varphi'}$$

Thus we can see that $x_\varphi \in D(A^\infty)$ and claim (i) follows.

To show (ii), fix $x \in X$. For each n, choose $\varphi_n \in C_c^\infty(\mathbb{R})$ with $\operatorname{supp}(\varphi_n) \subset (0, \frac{1}{n})$, $\varphi_n \geq 0$ and $\int_{\mathbb{R}} \varphi_n = 1$. Then by the dominated convergence theorem,

$$\begin{aligned}\|x - x_{\varphi_n}\| &= \left\| \int_0^\infty \varphi_n(s)[T(s)x - x]\,\mathrm{d}s \right\| \\ &\leq \int_0^\infty \varphi_n(s)\,\|T(s)x - x\|\,\mathrm{d}s \to 0 \qquad \text{as } n \to \infty\,.\end{aligned}$$

Thus D is dense in X. We are done. □

3.3.2 Resolvent of the Generator

Our starting points are the following two identities, which are easily derived from their predecessors in Theorem 3.8. We stress that these identities will be used very frequently throughout these notes.

Lemma 3.15. *Let $(A, D(A))$ be the generator of a strongly continuous semigroup $(T(t))_{t\geq 0}$. Then, for every $\lambda \in \mathbb{C}$ and $t > 0$,*

$$\begin{aligned}\mathrm{e}^{-\lambda t}T(t)x - x = (A - \lambda I)&\int_0^t \mathrm{e}^{-\lambda s}T(s)x\,\mathrm{d}s \qquad \textit{if } x \in X\,; \\ &= \int_0^t \mathrm{e}^{-\lambda s}T(s)(A - \lambda I)x\,\mathrm{d}s \qquad \textit{if } x \in D(A)\,.\end{aligned}$$

Proof. Observe that the rescaled semigroup $(\mathrm{e}^{-\lambda t}T(t))_{t\geq 0}$ has generator $(A-\lambda I, D(A))$. Then the desired result follows from Lemma 3.8. □

Next, we give an important formula relating the semigroup to the resolvent of its generator.

Theorem 3.16. *Let $(T(t))_{t\geq 0}$ be a strongly continuous semigroup on X and take constants $\omega \in \mathbb{R}, M \geq 1$ such that $\|T(t)\| \leq M\mathrm{e}^{\omega t}$ for $t \geq 0$. For the* *generator $(A, D(A))$ of $(T(t))_{t\geq 0}$ and scalar λ with $\operatorname{Re}\lambda > \omega$, the following properties hold.*

(i) For each $x \in X$, the integral $\int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t$ is well-defined.

(ii) $\lambda \in \varrho(A)$, and

$$R(\lambda; A)x = \int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t\,.$$

for each $x \in X$. Moreover,

$$\|R(\lambda, A)\| \leq \frac{M}{\operatorname{Re}\lambda - \omega}\,.$$

(iii) Let $n \in \mathbb{N}$. Then

$$R(\lambda, A)^n x = \frac{1}{(n-1)!} \int_0^\infty t^{n-1} \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t$$

for all $x \in X$. In particular, the estimates

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\operatorname{Re}\lambda - w)^n}\,.$$

The formula for $R(\lambda, A)$ in (ii) is called the *integral representation* of the resolvent. Indeed, by Theorem **??**, we can deduce that

$$R(\lambda, A) = \int_0^\infty \mathrm{e}^{-\lambda t} T(t) \,\mathrm{d}t \in (\mathcal{B}(X), \mathcal{T}_t)\,.$$

Proof. To prove (i), note that for $n \in \mathbb{N}$, $\int_0^n \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t$ is well-defined, since $t \mapsto \mathrm{e}^{-\lambda t} T(t)x$ is a continuous mapping from $[0, n]$ into X. Observe that

$$\left\| \int_n^{n+p} e^{-\lambda t} T(t)x \,\mathrm{d}t \right\| \leq M\|x\| \int_n^{n+p} e^{-(\operatorname{Re}\lambda - \omega)t} \,\mathrm{d}t$$

for all $n, p \geq 0$. Thus $\left\{\int_0^n \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t\right\}_{n\geq 1}$ is a Cauchy sequence in X. Let

$$\int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t := \lim_{n\to\infty} \int_0^n \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t\,. \tag{3.8}$$

Then for each $x^* \in X^*$,

$$\begin{aligned}\langle \int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t, x^* \rangle &= \lim_{n\to\infty} \langle \int_0^n \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t, x^* \rangle \\ &= \lim_{n\to\infty} \int_0^n \mathrm{e}^{-\lambda t} \langle T(t)x, x^* \rangle \,\mathrm{d}t = \int_0^\infty \mathrm{e}^{-\lambda t} \langle T(t)x, x^* \rangle \,\mathrm{d}t\end{aligned}$$

as desired. Thus assertion (i) follows.

To prove (ii), we need to that $\lambda I - A : D(A) \to X$ is bijective, which implies $\lambda \in \varrho(A)$ since A is closed. We have only to check that the inverse of $\lambda I - A$ is given by $x \mapsto \int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t$. That is

- for each $x \in X$, $\int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t \in D(A)$ and
$$(\lambda I - A) \int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t = x\,; \tag{3.9}$$
- for each $x \in D(A)$,
$$\int_0^\infty \mathrm{e}^{-\lambda t} T(t)(\lambda I - A)x \,\mathrm{d}t = x\,. \tag{3.10}$$

By (3.8) and Lemma 3.15, we have

$$\begin{aligned}(\lambda I - A) \int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t &= \lim_{n\to\infty} (\lambda I - A) \int_0^n \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t \\ &= x - \lim_{n\to\infty} \mathrm{e}^{-\lambda n} T(n)x = x\,; \\ \int_0^\infty \mathrm{e}^{-\lambda t} T(t)(\lambda I - A)x \,\mathrm{d}t &= \lim_{n\to\infty} \int_0^n \mathrm{e}^{-\lambda t} T(t)(\lambda I - A)x \,\mathrm{d}t \\ &= x - \lim_{n\to\infty} \mathrm{e}^{-\lambda n} T(n)x = x\,,\end{aligned}$$

since $\|\mathrm{e}^{-\lambda n} T(n)x\| \le M\mathrm{e}^{-(\operatorname{Re}\lambda - \omega)n}\|x\|$, as required. For the norm of the resolvent, we compute

$$\|R(\lambda; A)\| \le M \int_0^\infty \mathrm{e}^{-(\operatorname{Re}\lambda - \omega)t} \,\mathrm{d}t = \frac{M}{\operatorname{Re}\lambda - \omega}\,.$$

To show (iii), by part (b) of Theorem 1.3, and the dominated convergence theorem, we have

$$
\begin{aligned}
R(\lambda, A)^n x &= \frac{(-1)^{n-1}}{(n-1)!} \cdot \frac{\mathrm{d}^{n-1}}{\mathrm{d}\lambda^{n-1}} R(\lambda, A)x \\
&= \frac{(-1)^{n-1}}{(n-1)!} \cdot \frac{\mathrm{d}^{n-1}}{\mathrm{d}\lambda^{n-1}} \int_0^\infty \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t \\
&= \frac{1}{(n-1)!} \int_0^\infty t^{n-1} \mathrm{e}^{-\lambda t} T(t)x \,\mathrm{d}t \,.
\end{aligned}
$$

Finally, the estimate follows from

$$
\|R(\lambda, A)^n\| \leq \frac{M}{(n-1)!} \int_0^\infty t^{n-1} \mathrm{e}^{-(\operatorname{Re}\lambda - \omega)t} \,\mathrm{d}t = \frac{M}{|\operatorname{Re}\lambda - \omega|^n} \,.
$$

We are done. □

Remark 3.5. Property (ii) in Theorem 3.16 says that the spectrum of a semigroup generator is always contained in a left half-plane. The number determining the smallest such half-plane is an important characteristic of any linear operator and is called the *spectral bound*, defined by

$$
\mathrm{s}(A) := \sup\{\operatorname{Re}\lambda : \lambda \in \sigma(A)\}
$$

Then for a strongly continuous semigroup $(T(t))_{t\geq 0}$ with generator A, one has $-\infty \leq \mathrm{s}(A) \leq \omega_0 < +\infty$, where ω_0 is the growth bound.

The following exercise, though trivial, is very useful when we conclude that an operator is the generator of some semigroup.

Exercise 3.5. Let $(A, D(A))$, $(B, D(B))$ be two closed linear operator and $A \subset B$. If $\varrho(A) \cap \varrho(B) \neq \emptyset$, then $A = B$. (Hint: Use Exercise **??**.)

Proposition 3.17. *Let $(T(t))_{t\geq 0}$ be a strongly continuous semigroup on X with generator $(A, D(A))$. Then for each $t > 0$ one has*

$$
T(t) = s\text{-}\lim_{n\to\infty} \left(I - \frac{t}{n}A\right)^{-n} = s\text{-}\lim_{n\to\infty} \left[\frac{n}{t} R(\frac{n}{t}, A)\right]^n
$$

Proof. First of all, we show that the operators $\left[\frac{n}{t}R(\frac{n}{t},A)\right]^n$ are uniformly bounded. Take $\omega \in \mathbb{R}$ and $M > 0$ so that $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$. Then by Theorem 3.16, for large n,

$$\left\|\left[\frac{n}{t}R(\frac{n}{t},A)\right]^n\right\| \leq M\left(\frac{n/t}{(n/t-\omega)}\right)^n = M\left(1-\frac{\omega t}{n}\right)^{-n}$$

Since $\left(1-\frac{\omega t}{n}\right)^{-n}$ converges as $n \to \infty$, we conclude that $\left[\frac{n}{t}R(\frac{n}{t},A)\right]^n$ are uniformly bounded. Then by the Banach-Steinhaus theorem, it suffices to show that

$$T(t)x = \lim_{n\to\infty}\left(I-\frac{t}{n}A\right)^{-n}x \quad \text{for } x \in D\,,$$

where D is a dense subspace of X.

Note that formally we have

$$\begin{aligned}\left(I-\frac{t}{n}A\right)^{-n}x - T(t)x &= T(t-s)\left(I-\frac{s}{n}A\right)^{-n}x\Big|_{s=0}^{t} \\ &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\left[T(t-s)\left(I-\frac{s}{n}A\right)^{-n}x\right]\mathrm{d}s\,.\end{aligned}$$

By Lemma 3.10 and part (ii) of Theorem 1.3, since $D(A)$ is invariant under $R(n/t;A)$, for fixed $x \in D(A)$, then the mapping $s \to T(t-s)\left(I-\frac{s}{n}A\right)^{-n}x$ from $[0,t]$ into X is differentiable with

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}s}\left[T(t-s)\left(I-\frac{s}{n}A\right)^{-n}x\right] \\ &= -T(t-s)A\left(I-\frac{s}{n}A\right)^{-n}x + T(t-s)\left(I-\frac{s}{n}A\right)^{-(n+1)}Ax \\ &= T(t-s)\left(I-\frac{s}{n}A\right)^{-(n+1)}\left[\left(I-\frac{s}{n}A\right)-I\right]Ax \\ &= -\frac{s}{n}T(t-s)\left(I-\frac{s}{n}A\right)^{-(n+1)}A^2x\,,\end{aligned}$$

if we assume in addition that $x \in D(A^2)$. Therefore,

$$\begin{aligned}&\left\|\left(I-\frac{t}{n}A\right)^{-n}x - T(t)x\right\| \\ &\leq \int_0^t \frac{s}{n}\left\|T(t-s)\left(I-\frac{s}{n}A\right)^{-(n+1)}A^2x\right\|\mathrm{d}s \\ &\leq \frac{1}{n}\cdot t\cdot t\cdot Me^{|\omega|t}\cdot K\cdot\|A^2x\|\,,\end{aligned}$$

where K is a real number given by

$$\begin{aligned}\sup_{s\in[0,t];n\geq 1}\left\|\left(I-\frac{s}{n}A\right)^{-(n+1)}\right\| &\leq M\sup_{s\in[0,t];n\geq 1}\left(1-\frac{\omega s}{n}\right)^{-(n+1)}\\ &\leq M\sup_{n\geq 1}\left(1-\frac{|\omega|t}{n}\right)^{-(n+1)} := K<\infty\,.\end{aligned}$$

Letting $n\to\infty$ we get hat

$$T(t)x=\lim_{n\to\infty}\left(I-\frac{t}{n}A\right)^{-n}x\qquad\text{for } x\in D(A^2)\,,$$

Since $D(A^2)$ is dense in X by Theorem 3.14, the desired result follows. $\square$

Remark 3.6. Indeed one can show that for $x\in X$, $\left(I-\frac{t}{n}A\right)^{-n}x$ converges to $T(t)x$ as $h\downarrow 0$ uniformly for t in compact intervals.

To conclude this section, we collect in a diagram the information obtained so far on the relations between a semigroup, its generator, and its resolvent.

$$\begin{tikzcd}
& \big(T(t)\big)_{t\geq 0} \arrow[dl, "Ax=\lim\limits_{t\downarrow 0}\frac{T(t)x-x}{t}"'] \arrow[dr, shift left, "R(\lambda,A)=\int_0^\infty e^{-\lambda t}T(t)\,dt"] & \\
\big(A,D(A)\big) \arrow[rr, shift left, "R(\lambda,A)=(\lambda-A)^{-1}"] & & \big(R(\lambda,A)\big)_{\lambda\in\rho(A)} \arrow[ll, shift left, "A=\lambda-R(\lambda,A)^{-1}"] \arrow[ul, shift left, "T(t)=\lim\limits_{n\to\infty}[{}^n/_t R({}^n/_t,A)]^n"]
\end{tikzcd}$$

Figure 3.1: Semigroup, generator, and resolvent

3.4 Standard Constructions

In this section, we explain how one can construct in various ways new strongly continuous semigroups from a given one. In each case, we try to identify the corresponding generator, its spectrum and resolvent, so that our abstract definitions gain a more concrete meaning.

Example 3.1 (Rescaled Semigroups)**.** Let $(T(t))_{t\geq 0}$ to be a strongly continuous semigroup on X. For any scalar λ and any $\alpha > 0$, we define the *rescaled semigroup* $(S(t))_{t\geq 0}$ by

$$S(t) := \mathrm{e}^{\lambda t} T(\alpha t) \quad \text{ for } \ t \geq 0\,.$$

It's easy to check that $(S(t))_{t\geq 0}$ a strongly continuous semigroup on X. Denote by $(B, D(B))$ the generator of $(S(t))_{t\geq 0}$. Then we asserts that

$$B = \alpha A + \lambda I \quad \text{ with domain } \quad D(A) = D(B)\,.$$

Moreover, $\sigma(B) = \alpha\sigma(A) + \lambda$ and

$$R(\mu; B) = \frac{1}{\alpha} R(\frac{\mu - \lambda}{\alpha}; A) \quad \text{ for } \mu \in \varrho(B)\,.$$

This shows that we can switch quite easily between the original and the rescaled objects.

To prove the assertion, by the symmetry it suffices to show that $D(A) \subset D(B)$ and $Bx = \alpha Ax + \lambda x$ for $x \in D(A)$. We compute

$$\begin{aligned}
&\lim_{h\downarrow 0} \frac{e^{\lambda h} T(\alpha h)x - x}{h} = \lim_{h\downarrow 0} e^{\lambda h} \frac{[T(\alpha h)x - x]}{h} + \lim_{h\downarrow 0} \frac{e^{\lambda h} - 1}{h} x \\
&= \alpha Ax + \lambda x\,, \quad \text{for all } x \in D(A)
\end{aligned}$$

as required.

Taking $\lambda = -\omega_0$ (or $\lambda < -\omega_0$) and $\alpha = 1$ the rescaled semigroup will have growth bound equal to (or less than) zero. This is an assumption we make without loss of generality in many situations.

Example 3.2 (Product Semigroups.)**.** Let $(T(t))_{t\geq 0}$, $(S(t))_{t\geq 0}$ be two strongly continuous semigroup on X with generators $(A, D(A))$ and $(B, D(B))$ respectively. If $(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$ commutes; i.e., $S(t)T(t) = T(t)S(t)$ for all $t \geq 0$, then

(i) The operators $U(t) := S(t)T(t)$ form a strongly continuous semigroup $(U(t))_{t\geq 0}$, called the *product semigroup* of $(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$.

(ii) Denote by $(C, D(C))$ the generator of $(U(t))_{t\geq 0}$. Then $D(A) \cap D(B)$ is a core for C and

$$Cx = Ax + Bx \quad \text{for all} \quad x \in D(A) \cap D(B)\,.$$

In other words, the generator of $(U(t))_{t\geq 0}$ is the closure of $(D(A) \cap D(B), A + B)$.

To prove (i), it suffices to show that $T(s)$ and $S(r)$ commute for all $s, r \geq 0$. To this end, we first take $r = p_1/q$ and $s = p_2/q \in \mathbb{Q}_+$. Then

$$\begin{aligned} S(r)T(s) &= S(1/q)^{p_1} \cdot T(1/q)^{p_2} \\ &= T(1/q)^{p_2} \cdot S(1/q)^{p_1} = T(s)S(r)\,. \end{aligned}$$

Now fix $x \in X$. Since $(s, r) \mapsto S(r)T(s)x$ and $(s, r) \mapsto T(s)S(r)x$ both are continuous mapping from $\mathbb{R}_+^2$ in to X, and they coincide on $\mathbb{Q}_+^2$, we conclude that

$$S(r)T(s)x = T(s)S(r)x \quad \text{ for all } \quad s, r \geq 0$$

as desired.

To prove (ii), firstly we show if $x \in D(A) \cap D(B)$ then $x \in D(C)$ and $Cx = Ax + Bx$. We compute

$$\begin{aligned} \lim_{h\downarrow 0} \frac{T(h)S(h)x - x}{h} &= \lim_{h\downarrow 0} T(h)\frac{S(h)x - x}{h} + \lim_{h\downarrow 0} \frac{T(h)x - x}{h} \\ &= Bx + Ax \end{aligned}$$

as desired. It remains to show that $D(A) \cap D(B)$ is a core for C.

By Lemma 3.13, we have only to prove that $D(A) \cap D(B)$ is invariant under $(U(t))_{t\geq 0}$ and is dense in X. The former is trivial since $T(t)$ and $S(t)$ commutes. To prove the latter, take λ large enough, then we have the representations: $R(\lambda; A) = \int_0^\infty \mathrm{e}^{-\lambda s} T(t)\,\mathrm{d}t$ and $R(\lambda; B) = \int_0^\infty \mathrm{e}^{-\lambda s} S(t)\,\mathrm{d}t$. From these we deduce that

$$R(\lambda; A)R(\lambda; B) = R(\lambda; B)R(\lambda; A)\,.$$

Therefore, $R(\lambda; B)R(\lambda; A)X$ is contained in $D(A) \cap D(B)$. Because both $R(\lambda; A)$ and $R(\lambda; B)$ are continuous and have dense range, we conclude that $D(A) \cap D(B)$ is dense in X as desired.

Example 3.3 (Adjoint Semigroups)**.** Let $(T(t))_{t\geq 0}$ to be a strongly continuous semigroup on X. Then it's easy to see that $(T(t)^*)_{t\geq 0}$ consisting of all adjoint operators $T(t)^*$ on the dual space X^* is a semigroup.In general, it is NOT strongly continuous. An example is provided by the (left) translation group on $L^1(\mathbb{R})$. Its adjoint is the (right) translation group on $L^\infty(\mathbb{R})$, which is not strongly continuous.

However, it is easy to see that $(T(t)^*)_{t\geq 0}$ is always weak*-continuous in the sense that the maps

$$t \mapsto \langle x, T(t)^* x^* \rangle = \langle T(t)x, x^* \rangle$$

are continuous for all $x \in X$ and $x^* \in X^*$. Because on the dual of a reflexive Banach space weak and weak* topology coincide, if we assume X is reflexive, then the adjoint semigroup $(T(t)^*)_{t\geq 0}$ is weakly, and hence by Theorem 3.2 strongly, continuous.

We claim that if X is a Hilbert space and the generator of $(T(t))_{t\geq 0}$ is $(A, D(A))$, then the generator of the adjoint semigroup $(T(t)^*)_{t\geq 0}$ is $(A^*, D(A^*))$.

Denote by $(B, D(B))$ the generator of $(T(t)^*)_{t\geq 0}$. On the one hand, for each $y \in D(B)$,

$$\langle Ax, y \rangle = \lim_{h\downarrow 0} \langle \frac{T(h) - I}{h} x, y \rangle = \lim_{h\downarrow 0} \langle x, \frac{T(h)^* - I}{h} y \rangle = \langle x, By \rangle$$

for all $x \in D(A)$. Thus we have $B \subset A^*$. On the other hand, for $y \in D(A^*)$, we asserts that

$$T(t)^* y - y = \int_0^t T(s)^* A^* y \, \mathrm{d}s \quad \text{for all } t \geq 0 \,. \tag{3.11}$$

Then by Lemma 3.8, we conclude that $y \in D(B)$ and $By = A^* y$ as desired. To show (3.11), note that for all $x \in D(A)$,

$$\begin{aligned}
\langle x, T(t)^* y - y \rangle = \langle T(t)x - x, y \rangle &= \langle \int_0^t AT(s)x \, \mathrm{d}s, y \rangle \\
&= \int_0^t \langle AT(s)x, y \rangle \, \mathrm{d}s = \int_0^t \langle x, T(s)^* A^* y \rangle \, \mathrm{d}s \\
&= \langle x, \int_0^t T(s)^* A^* y \, \mathrm{d}s \rangle \,.
\end{aligned}$$

Since $D(A)$ is dense, (3.11) holds. We are done.

Example 3.4 (Similar Semigroups)**.** Let X, Y be two Banach spaces with a linear homeomorphism V from Y onto X. Let $(T(t))_{t\geq 0}$ to be a strongly continuous semigroup on X with generator $(A, D(A))$. Then we obtain a new strongly continuous semigroup $(S(t))_{t\geq 0}$ on Y by defining

$$S(t) := V^{-1} T(t) V \quad \text{for } t \geq 0 \,.$$

Its generator is

$$B = V^{-1} A V \quad \text{with domain} \quad D(B) = \{ y \in Y : Vy \in D(A) \}$$

Equality of the spectra

$$\sigma(A) = \sigma(B)$$

is clear, and the resolvent of B is $R(\lambda; B) = V^{-1} R(\lambda; A) V$ for $\lambda \in \rho(A)$.

Without explicit reference to the linear homeomorphism V, we call the two semigroups $(T(t))_{t\geq 0}$ and$(S(t))_{t\geq 0}$ *similar*. Two such semigroups have the same topological properties; e.g., they have the same growth bound.

3.5 The Stone Theorem

This section is devoted to describe the characteristic strongly continuous unitary group, and give an application of it. First of all, we describe which operators are generators of strongly continuous groups. In order to make this more precise we first adapt the definition in Section to this situation.

3.5.1 Generators of Groups

Let $(T(t))_{t\in\mathbb{R}}$ be a strongly continuous group on the Banach space X. Recall that the orbit maps is defined by

$$\xi_x : t \mapsto T(t)x \ ; \ \mathbb{R}_+ \to X$$

for each given $x \in X$. Similar to Lemma 3.7 we have:

Exercise 3.6. Then the orbit map ξ_x differentiable on $\mathbb{R}$ if and only if it is differentiable at $t = 0$.

Definition 3.5. The *generator* $(A, D(A))$ for the strongly continuous semigroup $(T(t))_{t\in\mathbb{R}}$ is the operator given by

$$\begin{aligned} Ax &= \lim_{h\to 0} A_h x = \lim_{h\to 0} \frac{T(h)x - x}{h} = \xi_x'(0)\,; \\ D(A) &= \{x : \xi_x \text{ is differentiable}\} = \{x : \lim_{h\to 0} A_h x \text{ exists }\}\,. \end{aligned}$$

Given a strongly continuous group $(T(t))_{t\in\mathbb{R}}$ with generator $(A, D(A))$ we can define

$$T_+(t) := T(t) \ \text{ and } \ T_-(t) := T(-t) \ \text{ for } \ t \geq 0\,.$$

Then, from the previous definition, it's clear that $(T_+(t))_{t\geq 0}$ and $(T_-(t))_{t\geq 0}$ are strongly continuous semigroups with generators $(A, D(A))$, $(-A, D(A))$ respectively. Therefore, if A is the generator of a group, then both A and $-A$ generate strongly continuous semigroups. The next result shows that the converse of this statement is also true.

Proposition 3.18. *Let $w \in \mathbb{R}$ and $M \geq 1$ be constants. For a linear operator $(A, D(A))$ on a Banach space X the following properties are equivalent.*

(a) $(A, D(A))$ *generates a strongly continuous group* $(T(t))_{t\in\mathbb{R}}$ *satisfying the growth estimate*

$$\|T(t)\| \leq M\mathrm{e}^{w|t|} \quad \text{for } t \in \mathbb{R}\,. \tag{3.12}$$

(b) $(A, D(A))$ *and* $(-A, D(A))$ *are the generators of strongly continuous semigroups* $(T_+(t))_{t\geq 0}$ *and* $(T_-(t))_{t\geq 0}$, *respectively, which satisfy*

$$\|T_+(t)\|\,, \|T_-(t)\| \leq M\mathrm{e}^{wt} \quad \text{for all } t \geq 0\,. \tag{3.13}$$

Proof. It remains to show that (b) $\Rightarrow$ (a). Let

$$T(t) = \begin{cases} T_+(t) & \text{if } t \geq 0\,; \\ T_-(-t) & \text{if } t \leq 0\,. \end{cases}$$

Then we asserts that $(T(t))_{t\in\mathbb{R}}$ forms a operator group. By Exercise 3.3, it suffices to show that

$$T_+(t)T_-(t) = T_-(t)T_+(t) = I \quad \text{for } t \geq 0\,.$$

To this end, firstly we show that $T_+(t)$ and $T_-(t)$ commutes. Observe that the Yosida approximants $A_{+,n}$ and $A_{-,n}$ of A and $-A$, respectively, commute; and we have

$$T_+(t)x = \lim_{n\to\infty} \exp\{tA_{+,n}\}x \quad \text{and} \quad T_-(t)x = \lim_{n\to\infty} \exp\{tA_{-,n}\}x$$

for all $x \in X$, we see that $T_+(t)$ and $T_-(t)$ commute. We have only to show that $T_+(t)T_-(t) = I$ for all $t \geq 0$.

Note that, formally, there holds

$$T_+(t)T_-(t)x - x = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} T_+(s)T_-(s)x\,\mathrm{d}s\,.$$

By Lemma 3.10, since $D(A)$ is invariant under $(T_-(t))_{t\geq 0}$, for $x \in D(A)$, the map $s \mapsto T_+(s)T_-(s)x$ is differentiable form $\mathbb{R}_+$ into X with

$$\frac{\mathrm{d}}{\mathrm{d}s}T_+(s)T_-(s)x = -T_+(s)AT_-(s)x + T_+(s)AT_-(s)x = 0\,.$$

Thus $T_+(t)T_-(t)x = x$ for all $x \in D(A)$. Since $T_+(t)T_-(t) \in B(X)$ and $D(A)$ is dense in X, we get $T_+(t)T_-(t) = I$.

Finally, the estimation (3.12) follows from (3.13); the strong continuity of $(T(t))_{t\in\mathbb{R}}$ follows form the strong continuity of $(T_+(t))_{t\geq 0}$ and $(T_-(t))_{t\geq 0}$; by the definition $(T(t))_{t\in\mathbb{R}}$ has generator $(A, D(A))$. We are done. $\square$

3.5.2 Stone's Theorem

In the following, we always assume that H is a complex Hilbert space.

Definition 3.6. An operator group $\{U(t)\}_{t\in\mathbb{R}}$ on H is called a *unitary group* if for each t, $U(t)$ is a unitary operator on H.

Remark 3.7. If $\{U(t)\}_{t\geq 0}$ is a strongly continuous unitary semigroup, since $U(t)^{-1} = U(t)^* \in \mathcal{B}(X)$ for each t, define $U(-t) = U(t)^* = U(t)^{-1}$ for $t \geq 0$. Then $\{U(t)\}_{t\in\mathbb{R}}$ forms a strongly continuous unitary group by Exercise 3.3 and the fact that

$$\|U(-t)x - x\| = \|U(t)^*x - x\| = \|U(t)x - x\| \text{ for all } t \geq 0, x \in H\,.$$

Remark 3.8. The weak continuity of a unitary group $\{U(t)\}_{t\in\mathbb{R}}$ implies the continuity. There is an proof much easier than Theorem 3.2. Indeed if $U(h)x \xrightarrow{w} x$ as $h \to 0$, since $\|U(h)x\| = \|x\|$ for all h, we conclude that $U(h)x \to x$ as $h \to 0$ by Theorem **??**.

Example 3.5. Let $(A, D(A))$ be a self-adjoint operator on the complex Hilbert H. Let E^A be the spectrum decomposition for A. For each $t \in \mathbb{R}$, set

$$U(t) \coloneqq e^{itA} \equiv \int_{\mathbb{R}} e^{it\lambda}E^A(\mathrm{d}\lambda)\,.$$

By the functional calculus, it's easy to see that $\{U(t)\}_{t\in\mathbb{R}}$ is a unitary group. We show that it is weakly continuous and hence, strongly continuous. Indeed, for each x, y in H,

$$t \mapsto \langle U(t)x, y\rangle = \int_{\mathbb{R}} e^{it\lambda} E^A_{x,y}(\mathrm{d}\lambda)$$

is a continuous function by the dominated convergence theorem as required.

We claim that the generator of $\{U(t)\}_{t\in\mathbb{R}}$ is $(iA, D(A))$.

Denote by $(B, D(B))$ the generator of $\{U(t)\}_{t\in\mathbb{R}}$. Then fix $x \in D(A)$, there holds

$$\left\| \frac{U(h)x - x}{h} x - iAx \right\|^2 = \int_{\mathbb{R}} \left| \frac{e^{ih\lambda} - 1}{h} - i\lambda \right|^2 E^A_{x,x}(\mathrm{d}\lambda) \to 0$$

as $h \to 0$, $h \neq 0$. Thus $iA \subset B$. On the other hand note that

$$\varrho(iA) \cap \varrho(B) \neq \emptyset$$

we conclude that $B = iA$ by Exercise 3.5.

Theorem 3.19 (Stone)**.** *Let $(B, D(B))$ be a densely defined operator on H. Then $(B, D(B))$ generates a strongly continuous unitary group $\{U(t)\}_{t\in\mathbb{R}}$ if and only if B is skew-adjoint, i.e., $B^* = -B$. In this case, $A = -iB$ is self-adjoint and*

$$U(t) = e^{itA} \quad \text{for all } t \in \mathbb{R}\,.$$

Proof. If B is skew-adjoint, let $A = -iB$ then $A^* = iB^* = -iB = A$, so A is self-adjoint. By the example above, the desired result follows.

If $(B, D(B))$ generates a strongly continuous unitary group $\{U(t)\}_{t\in\mathbb{R}}$, then by Proposition 3.18, $(-B, D(B))$ generator the semigroup $\{U_(t) = U(t)^*\}_{t\geq 0}$. However, by Example 3.3, the generator of the adjoint semigroup $\{U(t)^*\}_{t\geq 0}$ is $(B^*\ D(B^*)$. Thus we conclude that B is skew-adjoint

$$B^* = -B\,.$$

By the example above, since $\{U(t)\}_{t\geq 0}$ and $\{e^{itA}\}_{t\geq 0}$ has the same generator, by Theorem 3.11 there must be

$$U(t) = e^{itA}$$

for all $t \in \mathbb{R}$. We are done. $\square$

Finally we give an application of Stone's theorem. We shall answer the following question.

> A characteristic function is defined to be the Fourier transform of a probability measure. Can it be characterized by some other properties?

Several answers are known, but the following characterization, due to Bochner and Herglotz, is most useful. It plays a basic role in harmonic analysis and in the theory of second-order stationary processes.

Definition 3.7. A complex-valued function f defined on $\mathbb{R}$ is called *positive definite* iff for any finite set of real numbers t_j and complex numbers z_j (with conjugate complex $\bar{z}_j$), $1 \leq j \leq n$, we have

$$\sum_{j=1}^{n} \sum_{k=1}^{n} f(t_j - t_k) z_j \bar{z}_k \geq 0 \,. \tag{3.14}$$

Theorem 3.20 (Bochner)**.** *$f : \mathbb{R} \to \mathbb{C}$ is a characteristic function if and only if it is positive definite and uniformly continuous on $\mathbb{R}$ with $f(0) = 1$.*

Remark 3.9. Indeed if f is positive definite then there holds $f(-t) = \overline{f(t)}$, $|f(t)| \leq f(0)$ for all $t \in \mathbb{R}$ and; if f is continuous at 0 then f is uniformly continuous on $\mathbb{R}$. See Theorem 6.5.1 in *A Course in Probability Theory* by KaiLai Chung.

Proof. If f is the characteristic function of the probability measure μ, then we need only verify that it is positive definite. This is immediate, since

$$\begin{aligned}\sum_{1\le j,k\le n} f\left(t_j-t_k\right) z_j \bar{z}_k &= \int \sum_{1\le j,k\le n} e^{i(t_j-t_k)x} z_j \bar{z}_k \mu(\mathrm{d}x) \\ &= \int \big(\sum_{j=1}^{n} e^{it_j x} z_j\big)^2 \mu(\mathrm{d}x) \ge 0\,.\end{aligned}$$

On the contrary, assume that f is positive definite and uniformly continuous on $\mathbb{R}$. Let

$$L = \{x(t) : \mathbb{R} \to \mathbb{C} : x(t) = 0 \text{ for all except finitely many } t \in \mathbb{R}\}$$

and define

$$\langle x, y\rangle_f \coloneqq \sum_{s,t\in\mathbb{R}} f\left(t-s\right) x(t)\overline{y}(s) \quad \text{ for } \; x, y \in L\,.$$

Here for convenience we assume f is strictly positive definite, that is the equality in (3.14) holds if and only if $z_j = 0$ for all $1 \le j \le n$. Then $(L, \langle\cdot,\cdot\rangle_f)$ is a complex inner product space. (If f is only positive definite, then $\langle\cdot,\cdot\rangle_f$ is only a semi-inner product, we need to consider the quotient space L modulo $\langle\cdot,\cdot\rangle_f$.) Denote by $(H, \langle\cdot,\cdot\rangle_f)$ the completion of $(L, \langle\cdot,\cdot\rangle_f)$. Then L is a dense subspace in H.

For $\tau \in \mathbb{R}$, let $e_\tau \coloneqq 1_{\{\tau\}}(t) \in L$. Then

$$f(\tau) = \sum_{s,t\in\mathbb{R}} f\left(t-s\right) 1_{\{\tau\}}(t) 1_{\{0\}}(s) = \langle e_\tau, e_0\rangle_f\,.$$

Define $U_\tau \in B(H)$ by

$$(U_\tau x)(t) \coloneqq x(t-\tau) \;\; \text{ for } t \in \mathbb{R}, x \in L\,.$$

Since L is dense in H, U_τ is well-defined. Observe that $U_\tau e_0 = e_\tau$. Then

$$f(\tau) = \langle U_\tau e_0, e_0\rangle_f\,.$$

In fact $\{U_\tau\}_{\tau\in\mathbb{R}}$ is a strongly continuous unitary group on H. It's clear that $\{U_\tau\}_{\tau\in\mathbb{R}}$ is an operator group. Moreover, U_τ protects the norm

$$\|U_\tau x\|_f = \|x\|_f \quad \text{for } x \in L\,;$$

and by the uniform continuity of f,

$$\lim_{h\to 0} \|U_h x - x\|_f^2 = \lim_{h\to 0} \sum_{s,t\in\mathbb{R}} \left[f\left(t-s\right) - f\left(t-s+h\right)\right] x(t)\overline{x}(s) = 0\,.$$

for $x \in L$. Therefore by Stone's theorem, there is self-adjoint operator $(A, D(A))$ on H with spectrum decomposition E^A so that

$$f(\tau) = \langle U_\tau e_0, e_0\rangle_f = \int_{\mathbb{R}} \mathrm{e}^{it\tau} E^A_{e_0,e_0}(\mathrm{d}t)\,.$$

Since $\|e_0\|_f = 1$, $E^A_{e_0,e_0}$ is a probability measure on $\mathbb{R}$ as desired. $\square$

3.6 Examples (Needs to be modified)

In order to convince us that new and interesting phenomena appear for semigroups on infinite-dimensional Banach spaces, we first discuss several classes of semigroups on concrete spaces. These semigroups are not uniformly continuous anymore and hence not of the form $\left(\mathrm{e}^{tA}\right)_{t\geq 0}$ for some bounded operator A. On the other hand, they are not "pathological" in the sense of being completely unrelated to any analytic structure. In addition, these semigroups accompany us through the further development of the theory and provide a source of illuminating examples and counterexamples.

3.6.1 Multiplication Semigroups on $\mathrm{C}_0(\Omega)$

Multiplication operators can be considered as an infinite-dimensional generalization of *diagonal matrices*. They are extremely simple to construct, and most of their properties are evident. Nevertheless, their value should not be underestimated. They appear, for example, naturally in the context of Fourier analysis or when one applies the spectral theorem for self-adjoint operators on Hilbert spaces. We therefore strongly recommend that any first attempt to illustrate a result or disprove a conjecture on semigroups should be made using multiplication semigroups.

3.7 Hille-Yosida Theorems

We now turn to the fundamental problem of semigroup theory, which is to find arrows in Figure 3.1 leading from the generator to the semigroup. This means that we discuss the following problem:

> Characterize those linear operators that are generators of some strongly continuous semigroup, and describe how the semigroup is generated.

3.7.1 Exponential Formulas

To tackle the above problem, it is helpful to recall the results from Section 3.2 and to think of the semigroup generated by an operator A as an "exponential function"

$$t \mapsto \mathrm{e}^{tA} .$$

This is also implied by the following proposition.

Proposition 3.21. *Let* $(T(t))_{t\geq 0}$ *be a strongly continuous semigroup with generator* $(A, D(A))$. *Then*

$$T(t) = s\text{-}\lim_{h\downarrow 0} \mathrm{e}^{tA_h} .$$

Proof. By the Banach-Steinhaus theorem, it suffices to show that e^{tA_h} are uniformly bounded for $h \in (0,1]$; and $\mathrm{e}^{tA_h}x \to T(t)x$ as $h \downarrow 0$ for each $x \in D(A)$. To prove the first assertion, note that

$$\mathrm{e}^{tA_h} = e^{-\frac{t}{h}I} e^{\frac{t}{h}T(h)} = e^{-\frac{t}{h}} e^{\frac{t}{h}T(h)} = e^{-\frac{t}{h}} \sum_{n=0}^{\infty} \frac{1}{n!} \frac{t^n}{h^n} T(nh) .$$

By Proposition 3.3, there are $\omega, M > 0$ so that $\|T(t)\| \leq M\mathrm{e}^{\omega t}$ for all $t \geq 0$. Hence, set $K := \sup_{0<h\leq 1} \frac{1}{h}\left(e^{\omega h} - 1\right)$, and we compute for all $0 < h \leq 1$,

$$\|\mathrm{e}^{tA_h}\| \leq e^{-\frac{t}{h}} \sum_{n=0}^{\infty} \frac{1}{n!} \frac{t^n}{h^n} M\mathrm{e}^{\omega n h} = M \exp\left\{\frac{t}{h}\left(e^{\omega h} - 1\right)\right\} \leq M \exp\left\{tK\right\} .$$

Thus the uniform boundedness follows.

Observe that, formally we have

$$\mathrm{e}^{tA_h}x - T(t)x = \mathrm{e}^{(t-s)A_h}T(s)x\,\Big|_{s=0}^{t} = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{(t-s)A_h}T(s)x\,\mathrm{d}s\,.$$

By Lemma 3.10, since $D(A)$ is invariant under $T(s)$, then $s \to \mathrm{e}^{(t-s)A_h}T(s)x$ is differentiable mapping from $[0,t]$ into X with

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{(t-s)A_h}T(s)x &= -\mathrm{e}^{(t-s)A_h}A_hT(s)x + \mathrm{e}^{(t-s)A_h}AT(s)x \\ &= \mathrm{e}^{(t-s)A_h}T(s)[Ax - A_hx]\,.\end{aligned}$$

Therefore we get for all $0 < h \leq 1$,

$$\begin{aligned}\|\mathrm{e}^{tA_h}x - T(t)x\| &\leq \int_0^t \|\mathrm{e}^{(t-s)A_h}\|\|T(s)\|\|Ax - A_hx\|\,\mathrm{d}s \\ &\leq tM\exp\{tK\}M\exp\{\omega t\}\|Ax - A_hx\|\,.\end{aligned}$$

Letting $h \downarrow 0$, the desired result follows. $\square$

We pursue this idea by recalling the various ways by which we can define "exponential functions." Each of these formulas and each method is then checked for a possible generalization to infinite-dimensional Banach spaces and, in particular, to unbounded operators. Here are some more or less promising formulas for "e^{tA}".

(i) We might use the power series and define

$$\mathrm{e}^{tA} := \sum_{n=0}^{\infty} \frac{t^n}{n!}A^n$$

However, for unbounded A, it is unrealistic to expect convergence of this series.

(ii) We might use the Cauchy integral formula and define

$$\mathrm{e}^{tA} := \frac{1}{2\pi\mathrm{i}} \int_\gamma \mathrm{e}^{\lambda t} R(\lambda; A)\,\mathrm{d}\lambda$$

As already noted, the generator A, hence also its spectrum $\sigma(A)$, may be unbounded. Therefore, the path γ surrounding $\sigma(A)$ will be unbounded, and so we need extra conditions to make the integral converge. For a class of semigroups, this approach does work. See Section II.4 in *A Short Course on Operator Semigroups* by K.Engel and R.Nagel.

(iii) At least in the one-dimensional case, the formulas

$$\mathrm{e}^{tA} = \lim_{n\to\infty}\left(1 + \frac{t}{n}A\right)^n = \lim_{n\to\infty}\left(1 - \frac{t}{n}A\right)^{-n}$$

are well known. Whereas the first formula again involves powers of the unbounded operator A and therefore will rarely converge, we can rewrite the second. This yields a formula involving only powers of bounded operators. It was Hille's idea (in 1948) to use this formula and to prove that under appropriate conditions, the limit exists and defines a strongly continuous semigroup.

(iv) Because it is well understood how to define the exponential function for bounded operators, one can try to approximate A by a sequence $(A_n)_{n\geq 1}$ of bounded operators and hope that

$$\mathrm{e}^{tA} := \lim_{n\to\infty} \mathrm{e}^{tA_n}$$

exists and is a strongly continuous semigroup. This was Yosida's idea (also in 1948) and is now examined in detail in order to obtain strongly continuous semigroups.

We start with an important convergence property for the resolvent under the assumption that $\|\lambda R(\lambda; A)\|$ remains bounded as $\lambda \to \infty$. It suggests immediately which bounded operators A_n should be chosen to approximate the unbounded operator A.

Lemma 3.22. *Let $(A, D(A))$ be a closed, densely defined operator. Suppose there exist $w \in \mathbb{R}$ and $M > 0$ such that $[w, \infty) \subset \rho(A)$ and $\|\lambda R(\lambda; A)\| \leq M$ for all $\lambda \geq w$. Then the following convergence statements hold for $\lambda \to \infty$.*

(a) $\lambda R(\lambda; A)x \to x$ *for all* $x \in X$

(b) $\lambda AR(\lambda; A)x = \lambda R(\lambda; A)Ax \to Ax$ *for all* $x \in D(A)$

Proof. The second statement is an immediate consequence of the first one. Note that indeed the first one assets that $\lambda R(\lambda; A) \xrightarrow{s} I$ as $\lambda \to \infty$. Since $\|\lambda R(\lambda; A)\| \leq M$ for all $\lambda \geq w$, by the Banach-Steinhaus theorem, it suffices to show that $\lambda R(\lambda; A)x \to x$ for all $x \in D(A)$. Observe that

$$\|\lambda R(\lambda; A)x - x\| = \|R(\lambda; A)Ax\| \leq \frac{M}{\lambda}\|Ax\|\,,$$

the desired result follows. □

3.7.2 Generation Theorem : Contraction Case

Because for contraction semigroups the technical details of the subsequent proof become much easier (and because the general case can then be deduced from this one), we first give the characterization theorem for generators in this special case.

Theorem 3.23 (Contraction Case, Hille, Yosida)**.** *For a linear operator $(A, D(A))$ on a Banach space X, the following properties are all equivalent.*

(a) $(A, D(A))$ *generates a strongly continuous contraction semigroup.*

(b) $(A, D(A))$ *is closed, densely defined, and for every $\lambda > 0$ one has $\lambda \in \rho(A)$ and*

$$\|R(\lambda, A)\| \leq \frac{1}{\lambda}\,.$$

Proof. In view of Theorem 3.16 it suffices to show (b) $\Rightarrow$ (a). To that purpose, we define the so-called *Yosida approximants*

$$A_n := nAR(n, A) = n^2 R(n, A) - nI \in \mathcal{B}(X)$$

which are mutually commuting bounded operators for each $n \in \mathbb{N}$. Consider then the uniformly continuous semigroups given by

$$T_n(t) := \mathrm{e}^{tA_n}, \quad t \geq 0$$

Because A_n converges to A pointwise on $D(A)$ by Lemma 3.22, we anticipate that the following properties hold. By establishing these statements we complete the proof.

(i) $T(t)x := \lim_n T_n(t)x$ exists for each $x \in X$.

(ii) $(T(t))_{t\geq 0}$ is a strongly continuous contraction semigroup on X.

(iii) This semigroup has generator $(A, D(A))$.

Step 1. We shall show that for each fixed t, $\{T_n(t)\}_n$ converges strongly. Observe that each $(T_n(t))_{t\geq 0}$ is a contraction semigroup, because

$$\|T_n(t)\| \leq \mathrm{e}^{-nt}\mathrm{e}^{\|n^2 R(n,A)\| t} \leq \mathrm{e}^{-nt}\mathrm{e}^{nt} = 1\,.$$

So by the Banach-Steinhaus theorem, it suffices to prove that for each $x \in D(A)$,

$$\{T_n(t)x\}_{n=1}^{\infty} \quad \text{is a Cauchy sequence.}$$

Observe that, formally we have

$$T_n(t)x - T_m(t)x = T_m(t-s)T_n(s)x\,\Big|_{s=0}^{t} = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} T_m(t-s)T_n(s)x\,\mathrm{d}s\,.$$

By Lemma 3.10, the map $s \to T_m(t-s)T_n(s)x$ from $[0,t]$ into X is differentiable with

$$\frac{\mathrm{d}}{\mathrm{d}s} T_m(t-s)T_n(s)x = T_m(t-s)T_n(s)\left(A_n x - A_m x\right)\,.$$

Therefore we get

$$\begin{aligned}\|T_n(t)x - T_m(t)x\| &\leq \int_0^t \|T_m(t-s)T_n(s)\left(A_n x - A_m x\right)\| \,\mathrm{d}s \\ &\leq t\,\|A_n x - A_m x\| \;.\end{aligned}$$

By Lemma 3.22, $\{A_n x\}$ is a Cauchy sequence for each $x \in D(A)$. Therefore, $\{T_n(t)x\}$ converges for each $x \in D(A)$ as required.

Moreover, one can see that for fixed $x \in X$, $\{T_n(t)x\}$ converges uniformly for t in each compact interval $[0, t_0]$ as $n \to \infty$.

Step 2. The pointwise convergence of $(T_n(t)x)_n$ implies that the limit family $(T(t))_{t\geq 0}$ satisfies the functional equation, hence is a semigroup, and consists of contractions. Moreover, for each $x \in X$, the corresponding orbit map

$$\xi : t \mapsto T(t)x, \quad 0 \leq t \leq t_0$$

is the uniform limit of continuous functions

$$\xi_n : t \mapsto T_n(t)x, \quad 0 \leq t \leq t_0$$

and so is continuous itself. The strong continuity of $(T(t))_{t\geq 0}$ follows.

Step 3. Denote by $(B, D(B))$ the generator of $(T(t))_{t\geq 0}$ and fix $x \in D(A)$. We first show that $A \subset B$. Fix $x \in D(A)$, then by Lemma 3.8,

$$T_n(t)x - x = \int_0^t T_n(s)A_n x \,\mathrm{d}s \quad \text{ for } t \geq 0\,.$$

Letting $n \to \infty$, by the fact that $T_n(t)x \to T(t)x$ and $T_n(s)A_n x \to T(s)Ax$ and the dominated convergence theorem, we conclude that

$$T(t)x - x = \int_0^t T(s)Ax \,\mathrm{d}s \quad \text{ for } t \geq 0\,.$$

By Lemma 3.8 again, $x \in D(B)$ and $Bx = Ax$. On the other hand, since

$$\varrho(A) \cap \varrho(B) \supset (0, \infty) \neq \emptyset$$

there must be $A = B$ by Exercise 3.5. We are done. $\square$

If a strongly continuous semigroup $(T(t))_{t\geq 0}$ with generator A satisfies, for some $\omega \in \mathbb{R}$, an estimate

$$\|T(t)\| \leq \mathrm{e}^{\omega t} \quad \text{for } t \geq 0\,,$$

then we say $(T(t))_{t\geq 0}$ is *quasi-contractive*, and we can apply the above characterization to the rescaled contraction semigroup given by

$$S(t) := \mathrm{e}^{-\omega t} T(t) \quad \text{for } t \geq 0\,.$$

Because the generator of $(S(t))_{t>0}$ is $B = A - \omega I$.

Corollary 3.24. *Let $\omega \in \mathbb{R}$. For a linear operator $(A, D(A))$ on a Banach space X the following conditions are equivalent.*

(a) *$(A, D(A))$ generates a strongly continuous semigroup $(T(t))_{t\geq 0}$ satisfying*

$$\|T(t)\| \leq \mathrm{e}^{\omega t} \quad \textit{for } t \geq 0\,.$$

(a) *$(A, D(A))$ is closed, densely defined, and for each $\lambda > \omega$ one has*

$$\|R(\lambda, A)\| \leq \frac{1}{\lambda - \omega}\,.$$

In the end of this section we characterize the resolvents of the generator of a contraction semigroup. This is also an easy application of the Hille-Yosida theorem and Exercise 1.1.

Exercise 3.7. Let $\{R_\lambda\}_{\lambda>0}$ be a family of bounded linear operators on X. Then $\{R_\lambda\}_{\lambda>0}$ is the resolvents of $(A, D(A))$ which generates a strongly continuous contraction semigroup, iff the following statements hold.

(a) The resolvent equation holds: $R_\lambda - R_\mu + (\lambda - \mu) R_\lambda R_\mu = 0$ for each $\lambda, \mu > 0$.

(b) $\{R_\lambda\}_{\lambda>0}$ is strongly continuous in the sense that $\lambda R_\lambda \xrightarrow{s} I$ as $\lambda \to \infty$.

(c) $\{R_\lambda\}_{\lambda>0}$ is contractive, that is, $\|\lambda R_\lambda\| \leq 1$ for all $\lambda > 0$.

3.7.3 Generation Theorem : Contraction Case

The characterization of generators of arbitrary strongly continuous semigroups can be deduced from the above method for contraction semigroups. However, norm estimates for all powers of the resolvent are needed.

Proposition 3.25 (General Case, Feller, Miyadera, Phillips)**.** *Let* $(A, D(A))$ *be a linear operator on a Banach space* X *and let* $w \in \mathbb{R}$*,* $M \geq 1$ *be constants. Then the following properties are equivalent.*

(a) $(A, D(A))$ *generates a strongly continuous semigroup* $(T(t))_{t\geq 0}$ *satisfying*

$$\|T(t)\| \leq M\mathrm{e}^{\omega t} \quad \textit{for } t \geq 0\,.$$

(b) $(A, D(A))$ *is closed, densely defined, and for every* $\lambda > \omega$ *one has* $\lambda \in \rho(A)$ *and*

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\lambda - \omega)^n} \quad \textit{for all } n \in \mathbb{N}\,.$$

Proof. It suffices to show (b) $\Rightarrow$ (a). Without loss of generality we assume $\omega = 0$. Define the Yosida approximants

$$A_n := nAR(n, A) = n^2R(n, A) - nI \quad \text{and} \quad T_n(t) := \mathrm{e}^{tA_n}\,.$$

We shall establish these statements.

(i) $T(t)x := \lim_n T_n(t)x$ exists for each $x \in X$.

(ii) $(T(t))_{t\geq 0}$ is a strongly continuous contraction semigroup on X.

(iii) This semigroup has generator $(A, D(A))$.

We only show (i) since (ii) and (iii) can be proved by the same argument as before. Fix t, wo show the uniformly boundedness of $(T_n(t))_n$, note that

$$T_n(t) = \mathrm{e}^{-nt}\mathrm{e}^{n^2R(n,A)t} = \mathrm{e}^{-nt}\sum_{m=0}^{\infty}\frac{(n^2t)^m}{m!}R(n, A)^m$$

and hence for all n,

$$\|T_n(t)\| \le e^{-nt} \sum_{m=0}^{\infty} \frac{(n^2 t)^m}{m!} \frac{M}{n^m} = e^{-nt} M e^{nt} = M \, .$$

So by the Banach-Steinhaus theorem, it suffices to prove that for each $x \in D(A)$,

$$\{T_n(t)x\}_{n=1}^{\infty} \text{ is a Cauchy sequence.}$$

As before we have

$$\begin{aligned}
\|T_n(t)x - T_m(t)x\| &\le \int_0^t \left\| \frac{\mathrm{d}}{\mathrm{d}s} T_m(t-s) T_n(s) x \right\| \mathrm{d}s \\
&\le \int_0^t \|T_m(t-s)T_n(s)\left(A_n x - A_m x\right)\| \, \mathrm{d}s \\
&\le M^2 t \, \|A_n x - A_m x\| \ .
\end{aligned}$$

By Lemma 3.22, $\{A_n x\}$ is a Cauchy sequence for each $x \in D(A)$. Therefore, $\{T_n(t)x\}$ converges for each $x \in D(A)$ as required. □

Remark 3.10. As a general rule, we point out that for an operator $(A, D(A))$ to be a generator one needs

- Conditions on the location of $\sigma(A)$ in some left half-plane and
- Growth estimates of the form

  $$\|R(\lambda, A)^n\| \le \frac{M}{(\operatorname{Re}\lambda - \omega)^n}$$

  for all powers of the resolvent $R(\lambda, A)$ in some right half-plane or on some semiaxis (ω, ∞). We show emphasize that the estimate with only $n = 1$ does not suffice (if $M \neq 1$).

This last condition is rather complicated and can be checked for nontrivial examples only in the (quasi) contraction case, i.e., only if $n = 1$ is sufficient in the case $M = 1$.

3.8 The Lumer–Phillips Theorem

Due to their importance, we now return to the study of contraction semigroups and look for a characterization of their generator that does not require explicit knowledge of the resolvent. The following is a key notion towards this goal.

Definition 3.8. A linear operator $(A, D(A))$ on a Banach space X is called *dissipative* if

$$\|(\lambda I - A)x\| \geq \lambda \|x\| \tag{3.15}$$

for all $\lambda > 0$ and $x \in D(A)$.

To familiarize ourselves with these operators we state some of their basic properties.

Proposition 3.26. *For a dissipative operator* $(A, D(A))$ *the following properties hold.*

(i) $\lambda I - A$ *is injective for all* $\lambda > 0$ *and*

$$\big\|(\lambda I - A)^{-1} y\big\| \leq \frac{1}{\lambda}\|y\| \quad \textit{for all} \quad y \in \operatorname{Ran}(\lambda I - A)\,.$$

(ii) $\lambda I - A$ *is surjective for some* $\lambda > 0$ *if and only if it is surjective for each* $\lambda > 0$. *In that case, one has* $(0, \infty) \subset \rho(A)$.

(iii) A *is closed if and only if the range* $\operatorname{Ran}(\lambda I - A)$ *is closed for some (hence all)* $\lambda > 0$.

(iv) If A *is densely defined, then* A *is closable. Its closure* $\bar{A}$ *is again dissipative and satisfies* $\operatorname{Ran}(\lambda I - \bar{A}) = \overline{\operatorname{Ran}(\lambda I - A)}$ *for all* $\lambda > 0$.

Proof. (i) is just a reformulation of estimate (3.15).

Tp prove (ii), assume that $\lambda_0 I - A$ is surjective for some $\lambda_0 > 0$. Then $\lambda_0 \in \varrho(A)$ and $\|R(\lambda_0; A)\| \leq \frac{1}{\lambda}$ by (i). Note that for $\lambda > 0$,

$$\begin{aligned}\lambda I - A &= \lambda_0 I - A - (\lambda_0 - \lambda) I \\ &= [I - (\lambda_0 - \lambda) R(\lambda_0; A)] (\lambda_0 I - A) \ .\end{aligned}$$

Thus for all $|\lambda - \lambda_0| < \lambda_0$ we have $|\lambda - \lambda_0| \| R(\lambda_0; A) \| < 1$, which implies $\lambda \in \varrho(A)$. From this we can see that indeed $(0, \infty) \subset \varrho(A)$.

To prove property (iii) observe that

$$\lambda \|x\| \leq \|(\lambda I - A)x\| \leq \|Ax\| + \lambda \|x\|$$

for all $x \in D(A)$. Thus $R(\lambda I - A)$ is closed if and only if A is closed.

To prove property (iv), take a sequence (x_n) in $D(A)$ satisfying $x_n \to 0$ and $Ax_n \to y$. By Lemma **??**, we have to show that $y = 0$. For each $\epsilon > 0$, let $x \in D(A)$ so that $\|x - y\| < \epsilon$. Then for fixe $\lambda > 0$,

$$\|(\lambda I - A)x_n - (\lambda I - A)\frac{x}{\lambda}\| \geq \lambda \|x_n - \frac{x}{\lambda}\|$$

Passing to the limit as $n \to \infty$ yields

$$\left\| -y + x - \frac{1}{\lambda} Ax \right\| \geq \|x\|$$

For $\lambda \to \infty$ we obtain that $\|x\| \leq \|x - y\|$ Thus

$$\|y\| \leq \|y - x\| + \|x\| \leq 2\|y - x\| \leq 2\epsilon \, .$$

Since ϵ is arbitrary, we conclude that $y = 0$.

In order to verify that $\bar{A}$ is dissipative, take $x \in D(\bar{A})$. By definition of the closure of a linear operator, there exists a sequence (x_n) in $D(A)$ satisfying $x_n \to x$ and $Ax_n \to \bar{A}x$. Because A is dissipative and the norm is continuous, this implies that $\|(\lambda I - \bar{A})x\| \geq \lambda \|x\|$ for all $\lambda > 0$. Hence $\bar{A}$ is dissipative. Finally, observe that $\operatorname{Ran}(\lambda I - A)$ is dense in $\operatorname{Ran}(\lambda I - \bar{A})$. Because by assertion (iii) $\operatorname{Ran}(\lambda I - \bar{A})$ is closed in X, we obtain the final assertion in (iv). $\square$

From the resolvent estimate in Theorem 3.23, it is evident that the generator of a contraction semigroup satisfies the estimate (3.15), and hence is dissipative. On the other hand, many operators can be shown directly to be dissipative and densely defined. We therefore reformulate Generation Theorem 3.23 in such a way as to single out the property that ensures that a densely defined, dissipative operator is a generator.

Theorem 3.27 (Lumer, Phillips)**.** *For a densely defined, dissipative operator $(A, D(A))$ on a Banach space X the following statements are equivalent.*

(a) *The closure $\bar{A}$ of A generates a contraction semigroup.*

(b) $\mathrm{Ran}(\lambda I - A)$ *is dense in X for some (hence all) $\lambda > 0$.*

Proof. (a) $\Rightarrow$ (b). Theorem 3.23 implies that $\mathrm{Ran}(\lambda I - \bar{A}) = X$ for all $\lambda > 0$. By Proposition 3.26, $\mathrm{Ran}(\lambda - \bar{A}) = \overline{\mathrm{Ran}(\lambda - A)}$, then we obtain (b).

(b) $\Rightarrow$ (a). By the same argument, the density of the range $\mathrm{Ran}(\lambda - A)$ implies that $(\lambda - \bar{A})$ is surjective. Property (ii) of Proposition 3.26 shows that $(0, \infty) \subset \rho(\bar{A})$, and dissipativity of A implies the estimate

$$\|R(\lambda, \bar{A})\| \le \frac{1}{\lambda} \quad \text{for } \lambda > 0\,.$$

This was required in Theorem 3.23 to assure that $\bar{A}$ generated a contraction semigroup. □

Remark 3.11. The above theorem gains its significance when viewed in the context of the abstract Cauchy problem associated with an operator A. Assume that the operator A is known to be closed, densely defined, and dissipative. Then in order to solve the (time-dependent) initial value problem

$$x'(t) = Ax(t), x(0) = x$$

for all $x \in D(A)$, it is sufficient to solve the (stationary) resolvent equation

$$x - Ax = y$$

for all y in some dense subset in the Banach space X.

There is a simpler method that works particularly well in concrete function spaces such as $\mathrm{C}_0(\Omega)$ or $\mathrm{L}^p(\mu)$.

To introduce this method we start with a Banach space X and its dual space X^*. By the Hahn-Banach theorem, for every $x \in X$ there exists $x^* \in X^*$ such that

$$\langle x, x^* \rangle = \|x\|^2 = \|x^*\|^2 .$$

Hence, for every $x \in X$ the following set, called its *duality set*,

$$\mathcal{J}(x) := \left\{ x^* \in X^* : \langle x, x^* \rangle = \|x\|^2 = \|x^*\|^2 \right\}$$

is nonempty. Particular if X is a Hilbert space, then the duality set is always a singleton $\mathcal{J}(x) = \{y \in X : \langle x, y \rangle = \|x\|^2 = \|y\|^2\} = \{x\}$. Such sets allow a new characterization of dissipativity.

Theorem 3.28. *An operator $(A, D(A))$ is dissipative if and only if for every $x \in D(A)$ there exists $j(x) \in \mathcal{J}(x)$ such that*

$$\mathrm{Re}\langle Ax, j(x) \rangle \leq 0 . \tag{3.16}$$

If A is the generator of a strongly continuous contraction semigroup, then (3.16) *holds for all $x \in D(A)$ and arbitrary $x^* \in \mathcal{J}(x)$.*

Proof. Assume (3.16) is satisfied for $x \in D(A), \|x\| = 1$, and some $j(x) \in \mathcal{J}(x)$. Then $\langle x, j(x) \rangle = \|j(x)\|^2 = 1$ and

$$\begin{aligned} \|\lambda x - Ax\| &\geq |\langle \lambda x - Ax, j(x) \rangle| \\ &\geq \mathrm{Re}\langle \lambda x - Ax, j(x) \rangle \geq \lambda \end{aligned}$$

for all $\lambda > 0$. This proves one implication.

To show the converse, we take $x \in D(A), \|x\| = 1$, and assume that $\|\lambda x - Ax\| \geq \lambda$ for all $\lambda > 0$. Choose $y_\lambda^* \in \mathcal{J}(\lambda x - Ax)$ and consider the normalized elements $z_\lambda^* := y_\lambda^* / \|y_\lambda^*\|$.

Then the inequalities

$$\lambda \leq \|\lambda x - Ax\| = \langle \lambda x - Ax, z_\lambda^* \rangle = \lambda \,\mathrm{Re}\, \langle x, z_\lambda^* \rangle - \mathrm{Re}\, \langle Ax, z_\lambda^* \rangle$$

are valid for each $\lambda > 0$. Since $\|x\| \leq 1$ and $\|z_\lambda^*\| = 1$, we have $\operatorname{Re}\langle x, z_\lambda^*\rangle \leq 1$. Then this yields

$$\operatorname{Re}\langle Ax, z_\lambda^*\rangle \leq 0\,.$$

Moreover,

$$\operatorname{Re}\langle x, z_\lambda^*\rangle \geq 1 + \frac{1}{\lambda}\operatorname{Re}\langle Ax, z_\lambda^*\rangle \geq 1 - \frac{1}{\lambda}\|Ax\|\,.$$

Regarding $\{z_\lambda\}_{\lambda>0}$ as a net in weak*-compact set B_{X^*}, the closed unit ball in X^*, we can choose $z \in B_{X^*}$ a weak-star accumulation point of z_λ^* as $\lambda \to \infty$. Then

$$\|z^*\| \leq 1, \quad \operatorname{Re}\langle Ax, z^*\rangle \leq 0, \qquad \text{and} \qquad \operatorname{Re}\langle x, z^*\rangle \geq 1$$

Combining these facts, it follows that z^* belongs to $\mathcal{J}(x)$ and satisfies (3.16).

Finally, assume that A generates a contraction semigroup $(T(t))_{t\geq 0}$ on X. Then, for every $x \in D(A)$ and arbitrary $x^* \in \mathcal{J}(x)$, we have

$$\begin{aligned}\operatorname{Re}\langle Ax, x^*\rangle &= \lim_{h\downarrow 0}\left(\frac{\operatorname{Re}\langle T(h)x, x^*\rangle}{h} - \frac{\operatorname{Re}\langle x, x^*\rangle}{h}\right)\\ &\leq \overline{\lim_{h\downarrow 0}}\left(\frac{\|T(h)x\|\cdot\|x^*\|}{h} - \frac{\|x\|^2}{h}\right) \leq 0.\end{aligned}$$

This completes the proof. $\square$

3.9 Evolution Equations

We turn our attention to what could have been, in a certain perspective, our starting point: We want to solve the abstract Cauchy problem (ACP) associated with $(A, D(A))$ and the initial value x:

$$\begin{cases} u'(t) = Au(t)\,, \text{ for } t \geq 0\,; \\ u(0) = x\,, \end{cases} \tag{ACP}$$

where the independent variable t represents time, $u(\cdot)$ is a function on $\mathbb{R}_+$ with values in a Banach space X, $(A, D(A))$ a linear operator on X, and $x \in X$ the initial value.

Definition 3.9. A function $u : \mathbb{R}_+ \to X$ is called a *(classical) solution* of (ACP) if u is continuously differentiable, $u(t) \in D(A)$ for all $t \geq 0$, and (ACP) holds.

If u is a classical solution of (ACP), combine the fact that $Au(t) = u'(t)$ and $u : \mathbb{R}_+ \to X$ is continuous differentiable, there holds

$$u \in C^1(\mathbb{R}_+; X) \cap C(\mathbb{R}_+; (D(A), \|\cdot\|_A)) .$$

Theorem 3.29. *Let $(A, D(A))$ be the generator of the strongly continuous semigroup $(T(t))_{t\geq 0}$. Then, for every $x \in D(A)$, the function*

$$u : t \mapsto u(t) := T(t)x$$

is the unique classical solution of (ACP).

Proof. It suffices to show the uniqueness of the solution. Let $v(t)$ be another classical solution of (ACP). Fix $t > 0$, then

$$v(t) - u(t) = v(t) - T(t)x = T(t-s)v(s)\,\Big|_{s=0}^{t} .$$

It's easy to show that the mapping $s \mapsto T(t-s)v(s)$ form $\mathbb{R}_+$ into X is differentiable with

$$\frac{\mathrm{d}}{\mathrm{d}s} T(t-s)v(s) = -T(t-s)Av(s) + T(t-s)Av(s) = 0 .$$

By the Newton-Leibniz formula we get $v(t) = u(t)$. We are done. □

The important point is that (classical) solutions exist if (and, by the definition of $D(A)$, only if) the initial value x belongs to $D(A)$. However, one might *substitute the differential equation by an integral equation*, thereby obtaining a more general concept of "solution".

Definition 3.10. A continuous function $u : \mathbb{R}_+ \to X$ is called a *mild solution* of (ACP) if $\int_0^t u(s)\,\mathrm{d}s \in D(A)$ for all $t \geq 0$ and

$$u(t) = A\int_0^t u(s)\,\mathrm{d}s + x .$$

Every classical solution u of (ACP) is a mild solution. Indeed since $u \in C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$ and $A : (D(A), \|\cdot\|_A) \to X$ is continuous, using the Newton-Leibniz formula we have

$$u(t) = \int_0^t u'(s)\,\mathrm{d}s + x = \int_0^t Au(s)\,\mathrm{d}s + x = A\int_0^t u(s)\,\mathrm{d}s + x$$

for all t. Conversely, if u is a mild solution of (ACP) and if $u \in C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$ then u must be a classical solution since

$$u(t) = A\int_0^t u(s)\,\mathrm{d}s + x = \int_0^t Au(s)\,\mathrm{d}s + x\,.$$

Then we conclude that $u \in C^1(\mathbb{R}_+; X)$ with $u'(t) = Au(t)$ and $u(0) = x$.

It follows from our previous results (Lemma 3.8) that for A being the generator of a strongly continuous semigroup, mild solutions exist for every initial value $x \in X$ and are again given by the semigroup.

Proposition 3.30. *Let $(A, D(A))$ be the generator of the strongly continuous semigroup $(T(t))_{t\geq 0}$. Then, for every $x \in X$, the orbit map*

$$u : t \mapsto u(t) := T(t)x$$

is the unique mild solution of the associated abstract Cauchy problem (ACP).

Proof. We only have to show the uniqueness of the zero solution for the initial value 0. To this end, assume u to be a mild solution of (ACP) for $x = 0$. Then

$$u(t) = A\int_0^t u(s)\,\mathrm{d}s \quad \text{for} \quad t \geq 0\,.$$

Let $v(t) = \int_0^t u(s)\,\mathrm{d}s$ for $t \geq 0$. Since $u \in C(\mathbb{R}_+; X)$ we have $v \in C^1(\mathbb{R}_+; X)$ with

$$v'(t) = u(t) = A\int_0^t u(s)\,\mathrm{d}s = Av(t) \quad \text{for} \quad t \geq 0\,.$$

Thus v is a classical solution of (ACP) with initial value $x = 0$. By Theorem 3.29, $v(t) = 0$ for all $t \geq 0$ and hence $u(t) = 0$ for all $t \geq 0$ as desired. □

The above two propositions are just reformulations of results on strongly continuous semigroups. They might suggest that the converse holds. The following example shows that this is not true.

Example 3.6. Let $(B, D(B))$ be a closed and unbounded operator on X. On the product space $X \times X$, consider the operator $(A, D(A))$ written in matrix form as

$$A := \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \qquad \text{with domain} \qquad D(A) = X \times D(B)\,.$$

Then $t \mapsto u(t) := \binom{x+tBy}{y}$ is the unique solution of (ACP) associated with A for every $\binom{x}{y} \in D(A)$. However, the operator A does not generate a strongly continuous semigroup, because for every $\lambda \in \mathbb{C}$, one has

$$\operatorname{Ran}(\lambda I - A) = \left\{ \begin{pmatrix} \lambda x - By \\ \lambda y \end{pmatrix} : x \in X, y \in D(B) \right\} \subset X \times D(B) \neq X \times X\,,$$

and hence $\sigma(A) = \mathbb{C}$ and $\varrho(A) = \emptyset$.

3.9.1 Well-Posedness

We now show which properties of the solutions $u(\cdot, x)$ or of the operator $(A, D(A))$ have to be added in order to characterize semigroup generators.

Definition 3.11. The abstract Cauchy problem (ACP) is called *well-posed* if the following statements hold.

- (Existence and Uniqueness; EU) There exists a unique classical solution $u(\cdot, x)$ of (ACP) for every initial value $x \in D(A)$.

- (Continuous Dependence; CD) For every $T > 0$, there is an $M > 0$ such that

$$\|u(t, x)\| \leq M\|x\| \text{ for all } t \in [0, T] \text{ and } x \in D(A)\,.$$

Exercise 3.8. Assume the property for existence and uniqueness holds. Show that the following statement is equivalent to "Continuous Dependence".

- (Continuous Dependence') For every sequence $(x_n)_{n\geq 1}$ in $D(A)$ so that $x_n \to 0$, one has $u(t, x_n) \to 0$ uniformly in $t \in [0, T]$ for each $T > 0$.

(Hint: Note that by the uniqueness we have $u(t, \frac{x}{n}) = \frac{1}{n}u(t, x)$.)

Theorem 3.31. *Let* $(A, D(A))$ *be a densely defined closed operator on* X*. Then for the associated abstract Cauchy problem* (ACP) *the following properties are equivalent.*

(a) $(A, D(A))$ *generates a strongly continuous semigroup.*

(b) *(EU) holds and* $\rho(A) \neq \emptyset$*.*

(c) *The abstract Cauchy problem is well-posed.*

Proof. From the basic properties of semigroup generators, it follows that (a) implies (b) and (c).

Step 1. For (b) $\Rightarrow$ (c), we first show that for all $x \in X$ there exists a unique mild solution of (ACP) with initial value x.

By the hypothesis there is $\lambda \in \varrho(A)$. We set $y = R(\lambda; A)x \in D(A)$. Then there is a unique classical solution $u(\cdot, y)$ of (ACP) with initial value y. Define

$$v(t, x) := (\lambda I - A)u(t, y) \quad \text{for} \quad t \geq 0\,.$$

We asserts that $v(\cdot, x)$ defines a mild solution for the initial value x.

In fact, since $u(\cdot, y) \in C^1(\mathbb{R}_+; X) \cap C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$, we conclude that $v(\cdot, x) : \mathbb{R}_+ \to X$ is continuous. We compute for each $t \geq 0$

$$\begin{aligned}\int_0^t v(s, x)\,\mathrm{d}s &= \int_0^t (\lambda I - A)u(s, y)\,\mathrm{d}s = \lambda \int_0^t u(s, y)\,\mathrm{d}s - \int_0^t Au(s, y)\,\mathrm{d}s \\ &= \lambda \int_0^t u(s, y)\,\mathrm{d}s - (u(t, y) - y)\,.\end{aligned}$$

Thus

$$\begin{aligned}
A\int_0^t v(s,x)\,\mathrm{d}s &= \lambda A\int_0^t u(s,y)\,\mathrm{d}s - A(u(t,y)-y)\\
&= \lambda\int_0^t Au(s,y)\,\mathrm{d}s - A(u(t,y)-y)\\
&= (\lambda I - A)(u(t,y)-y) = v(t,x)-x\,.
\end{aligned}$$

In order to prove uniqueness let $u(\cdot,0)$ be a mild solution to the initial value 0. Then $v(t,0) := \int_0^t u(s,0)\,\mathrm{d}s$ is the classical solution for the initial value 0, (see the proof of Proposition 3.30), hence $v(\cdot,0)=0$ and consequently $u(\cdot,0)=0$ as well.

Step 2. For (b) $\Rightarrow$ (c), We now show the continuous dependence upon the initial data.

We consider for fixed $T>0$ the linear map

$$\Phi : X \to \mathrm{C}\left([0,T]\,,X\right), \quad x \mapsto u(\cdot,x)$$

where $u(\cdot,x)$ is the mild solution of (ACP) with initial data $x\in X$. By the existence and uniqueness of mild solution Φ is a well-defined linear operator. It remains to show that Φ is bounded.

By the closed graph theorem, it's sufficient to prove that Φ is closed. In fact, if $x_n\to x$ and $u\left(\cdot,x_n\right)\to v(\cdot)\in\mathrm{C}\left([0,T]\,,X\right)$. Then $v(0)=x$. We have only to show that $v(t)=u(t,x)$ for all $0\le t\le T$.

Observe that $t\in[0,T]$

$$D(A)\ni\int_0^t u\left(s,x_n\right)\mathrm{d}s \to \int_0^t v(s)\,\mathrm{d}s\,,$$

and

$$A\int_0^t u\left(s,x_n\right)\mathrm{d}s = u\left(t,x_n\right)-x_n \to v(t)-x\,.$$

Hence, by the closedness of A we conclude that $\int_0^t v(s)\,\mathrm{d}s\in D(A)$ and

$$A\int_0^t v(s)\,\mathrm{d}s = v(t)-x.$$

for all $0 \le t \le T$. Consequently $v(\cdot)$ is the unique mild solution of (ACP) with initial value x if we define $v(t) = u(t - T, v(T))$ for $t > T$. Therefore $v(t) = u(t, x)$ for all $0 \le t \le T$ as required.

Step 3. Finally we show that (c) $\Rightarrow$ (a).

Firstly for each t, since $x \mapsto u(t, x)$, $D(A) \to X$ is bounded and $D(A)$ is dense in X, the operator $T(t) \in \mathcal{B}(X)$ is well-defined by

$$T(t)x := u(t, x) \quad \text{for all } x \in D(A)\,.$$

Moreover, $\sup_{0 \le t \le T} \|T(t)\| < \infty$, since by the hypothesis

$$\sup_{0 \le t \le T} \|T(t)\| = \sup_{0 \le t \le T} \sup_{x \in D(A), \|x\| \le 1} \|u(t, x)\| < \infty\,.$$

The uniqueness of the solutions implies

$$T(t + s)x = T(t)T(s)x$$

for each $x \in D(A)$ and all $t, s \ge 0$. Thus $(T(t))_{t \ge 0}$ is a semigroup on X. Since $T(t)x = u(t, x) \to x$ as $t \downarrow 0$ for all $x \in D(A)$, by Proposition 3.1 $(T(t))_{t \ge 0}$ is strongly continuous.

Now it suffices to show that the generator of $(T(t))_{t \ge 0}$ is exactly $(A, D(A))$. We denote it by $(B, D(B))$ temporarily. Since for $x \in D(A)$,

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t, x) = Au(t, x)\,.$$

We get $x \in D(B)$ and $Bx = Ax$, that is $A \subset B$. On the other hand note that $D(A)$ is invariant under $(T(t))_{t \ge 0}$, by Lemma 3.13, $D(A)$ is a core for B. Since A is closed, we get $A = B$ as desired. $\square$

3.9.2 Regularity of the Solution

We now assume that $(A, D(A))$ is the generator of a strongly continuous semigroup $\{T(t)\}_{t \ge 0}$ on the Banach space X. We shall prove here that the classical solution u of the homogeneous equation

$$\begin{cases} u'(t) = Au(t)\,, \ \text{for } t \geq 0\,; \\ u(0) = x\,, \end{cases} \tag{ACP}$$

is more regular than just $C^1(\mathbb{R}_+; X) \cap C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$, provided one makes additional assumptions on the initial data x.

For each $k \geq 1$ and we define the *k-norm* on $D\left(A^k\right)$ by

$$\|x\|_k = \Big[\|x\|^2 + \|Ax\|^2 + \ldots + \|A^k x\|^2\Big]^{1/2} \ \text{ for } x \in D(A^k)\,.$$

We set $D(A^0) = X$ and $A^0 = I$. Then $\big(D(A^k), \|\cdot\|_k\big)$ is a Banach space. Indeed if (x_n) is a Cauchy sequence in $\big(D(A^k), \|\cdot\|_k\big)$, then there are x, y_j, $1 \leq j \leq k$ in X so that $x_n \to x$ and $A^j x_n \to y_j$. Then since A is closed, by induction, we conclude that $x \in D(A^k) = \big(D(A^k), \|\cdot\|_k\big)$ and $y_j = A^j x$ for $1 \leq j \leq k$. Thus $x_n \to x$ in $\big(D(A^k), \|\cdot\|_k\big)$.

$(T(t)|_{D(A^k)})_{t\geq 0}$ is a strongly continuous semigroup on $\big(D(A^k), \|\cdot\|_k\big)$. By Lemma 3.8, $T(t)$ maps $D(A^k)$ into $D(A^k)$. $T(t)$ is bounded since

$$\begin{aligned} \|T(t)x\|_k^2 &= \sum_{j=0}^{k} \|A^j T(t)x\|^2 = \sum_{j=0}^{k} \|T(t)A^j x\|^2 \\ &\leq \|T(t)\|^2 \sum_{j=0}^{k} \|A^j x\|^2 = \|T(t)\|^2 \|x\|_k^2\,. \end{aligned}$$

for all $x \in D(A^k)$. The strong continuity follows from the fact that

$$\|T(h)x - x\|_k^2 = \sum_{j=0}^{k} \|T(h)A^j x - A^j x\|^2 \to 0 \ \text{ as } \ h \downarrow 0$$

for each $x \in D(A^k)$.

The generator of $(T(t)|_{D(A^k)})_{t\geq 0}$ is exactly $A|_{D(A^{k+1})}$. Indeed, for $x, y \in D(A^k)$, there holds

$$\begin{aligned} \lim_{h\downarrow 0} \left\| \frac{T(h)x - x}{h} - y \right\|_k = 0 &\Longleftrightarrow \lim_{h\downarrow 0} \sum_{j=0}^{k} \left\| \frac{T(h)A^j x - A^j x}{h} - A^j y \right\| = 0 \\ &\Longleftrightarrow x \in D(A^{k+1}) \text{ and } y = Ax\,. \end{aligned}$$

Moreover, it's clear that the graph norm of the generator $A|_{D(A^{k+1})}$, namely $\|\cdot\|_{k,A|_{D(A^{k+1})}}$ on $D(A^{k+1})$ is equivalent to the norm $\|\cdot\|_{k+1}$.

Proposition 3.32. *Assume $x \in D(A^k)$ for some integer $k \geq 1$. Then the classical solution of* (ACP) *$u(t) = T(t)x$ satisfies*

$$u \in C^j\Big(\mathbb{R}_+; (D(A^{k-j}), \|\cdot\|_{k-j})\Big) \quad \textit{for} \ \ 0 \leq j \leq k\,.$$

In particular, if $x \in D(A^\infty)$, then

$$u \in C^\infty\Big(\mathbb{R}_+; (D(A^l), \|\cdot\|_l)\Big) \quad \textit{for all} \ \ l \geq 0\,. \tag{3.17}$$

Proof. The assumption $x \in D(A^k)$ means that the initial value x belongs to the domain of the generator of the semigroup $(T(t)|_{D(A^{k-1})})_{t\geq 0}$. Since $u(t) = T(t)x = T(t)|_{D(A^{k-1})}x$ for all $t \geq 0$, by Theorem 3.29,

$$\begin{aligned} u &\in C^1(\mathbb{R}_+; (D(A^{k-1}), \|\cdot\|_{k-1})) \cap C(\mathbb{R}_+; (D(A^k), \|\cdot\|_{k-1,A|_{D(A^k)}})) \\ &= C^1(\mathbb{R}_+; (D(A^{k-1}), \|\cdot\|_{k-1})) \cap C(\mathbb{R}_+; (D(A^k), \|\cdot\|_k))\,. \end{aligned}$$

Also, we have

$$u'(t) = T(t)|_{D(A^{k-1})}A|_{D(A^k)}x = T(t)Ax$$

where u' is the derivative of $u : \mathbb{R}_+ \to C^1(\mathbb{R}_+; (D(A^{k-1}), \|\cdot\|_{k-1}))$.

By induction, suppose that for some $1 \leq j < k$ we have

$$u \in C^j\Big(\mathbb{R}_+; (D(A^{k-j}), \|\cdot\|_{k-j})\Big)\,, u^{(j)}(t) = T(t)A^jx\,.$$

Then $A^jx \in D(A^{k-j})$, i.e., A^jx belongs to the domain of the generator of the semigroup $(T(t)|_{D(A^{k-j-1})})_{t\geq 0}$. Since $u^{(j)}(t) = T(t)A^jx = T(t)|_{D(A^{k-j-1})}A^jx$, by Theorem 3.29 again,

$$\begin{aligned} u^{(j)} &\in C^1(\mathbb{R}_+; (D(A^{k-j-1}), \|\cdot\|_{k-j-1})) \cap C(\mathbb{R}_+; (D(A^k), \|\cdot\|_{k-j-1,A|_{D(A^{k-j})}})) \\ &= C^1(\mathbb{R}_+; (D(A^{k-j-1}), \|\cdot\|_{k-j-1})) \cap C(\mathbb{R}_+; (D(A^{k-j}), \|\cdot\|_{k-j}))\,. \end{aligned}$$

Therefore

$$u \in C^{(j+1)}(\mathbb{R}_+; (D(A^{k-j-1}), \|\cdot\|_{k-j-1}))$$

and $u^{(j+1)}(t) = T(t)|_{D(A^{k-j-1})}A|_{D(A^{k-j})}A^jx = T(t)A^{j+1}x$. Now the desired result follows. $\square$

The Self-Adjoint Case Let H be a complex Hilbert space and $(A, D(A))$ be a self-adjoint operator on H so that $A \leq 0$, i.e., $\langle Ax, x\rangle \leq 0$ for all $x \in D(A)$. It's also easy to show that A is dissipative. For $\lambda > 0$, note that

$$\operatorname{Ran}(\lambda I - A)^{\perp} = \operatorname{Ker}(\lambda I - A^*) = \operatorname{Ker}(\lambda I - A) = \{0\}\,.$$

Thus $\operatorname{Ran}(\lambda I - A)$ is dense in H. By the Lumer–Phillips theorem, $(A, D(A))$ generates a strongly continuous contraction semigroup $(T(t))_{t\geq 0}$ on H. Moreover, using the same argument in Example 3.5, we have

$$T(t) = \mathrm{e}^{tA} := \int_{\sigma(A)} \mathrm{e}^{t\lambda}\; E^A(\mathrm{d}\lambda) \qquad \text{for } t \geq 0\,, \tag{3.18}$$

where E^A is the spectrum decomposition of A.

Theorem 3.33. *For every $x \in H$, the abstract Cauchy equation*

$$\begin{cases} u'(t) = Au(t) & \text{for} \;\; t > 0\,; \\ u(0) = x\,. \end{cases} \tag{ACP$'$}$$

has a unique solution u in

$$C(\mathbb{R}_+; H) \cap C^1((0,+\infty); H) \cap C((0,+\infty); (D(A), \|\cdot\|_A))\,, \tag{3.19}$$

given by $u(t) = T(t)x = \mathrm{e}^{tA}x$ for all $t \geq 0$. Moreover,

(i) $u \in C^{\infty}\big((0,+\infty); (D(A^l), \|\cdot\|_l)\big)$*; and*

(ii) for each $k \geq 0$ there is a constant C_k so that

$$\left\|u^{(k)}(t)\right\| = \left\|A^k u(t)\right\| \leq \frac{C_k}{t^k}\|x\| \;\; \text{for all} \;\; t > 0\,.$$

Proof. Step 1. We show the uniqueness for the solution of (ACP$'$) in (3.19). It suffices to show that the solution u in (3.19) of (ACP$'$) with initial value 0 is $u(t) = 0$ for all $t \geq 0$. To this end we compute

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)\|^2 = \frac{\mathrm{d}}{\mathrm{d}t}\langle u(t), u(t)\rangle = 2\langle Au(t), u(t)\rangle \leq 0 \;\; \text{for all} \;\; t > 0\,.$$

Thus $t \mapsto \|u(t)\|^2$ is decreasing on $(0, \infty)$. Combine this with the fact that $t \mapsto \|u(t)\|^2$ is continuous on $[0, \infty)$, we get $u(t) = 0$ as desired.

Step 2. We now that the theorem holds provided that the initial value $x \in D(A^\infty)$. By Theorem 3.29, $u(t) = T(t)x = \mathrm{e}^{tA}x$ is the solution of (ACP) and hence must be the solution of (ACP′). Part (i) follows form by Proposition 3.32. It remains to show part (ii).

The case for $k = 0$ is trivial. Let $k \geq 1$, by part (ii) of Lemma 3.8, $u^{(k)}(t) = T(t)A^k x = \mathrm{e}^{tA}A^k x$. Thus by (3.18),Theorem **??**, Exercise **??** and Theorem **??**, we have

$$\begin{aligned}
\|u^{(k)}(t)\|^2 &= \int_{\sigma(A)} |\lambda^k \mathrm{e}^{t\lambda}|^2 \, E^A_{x,x}(\mathrm{d}\lambda) \\
&= \int_{\sigma(A)} \big(|\lambda|^k \mathrm{e}^{-t|\lambda|}\big)^2 \, E^A_{x,x}(\mathrm{d}\lambda) \\
&\leq \int_{\sigma(A)} \big(\frac{C_k}{t^k}\big)^2 \, E^A_{x,x}(\mathrm{d}\lambda) = \big(\frac{C_k}{t^k}\|x\|\big)^2 ,
\end{aligned}$$

where $C_k = \max\{\mathrm{e}^{-s}s^k : s > 0\} < \infty$. We are done.

Step 3. We now handle the general case and assume that $x \in H$. Let $u(t) = T(t)x$. By Theorem 3.14, $D(A^\infty)$ is dense in H, thus we can choose a sequence (x_n) in $D(A^\infty)$ so that $\|x_n - x\| \to 0$. Let $u_n(t) \coloneqq T(t)x_n$. Since $\{T(t)\}_{t\geq 0}$ is a contraction semigroup

$$\|u_n(t) - u(t)\| = \|T(t)x - T(t)x_n\| \leq \|x - x_n\|$$

for all $t \geq 0$. Thus (u_n) converges uniformly to u on $\mathbb{R}_+$.

By Step 1, fix $k \geq 0$ and $t > 0$, we have

$$\begin{aligned}
&\left\|u_n^{(k)}(t) - u_m^{(k)}(t)\right\|_l = \sum_{j=0}^{l} \left\|A^j u_n^{(k)}(t) - A u_m^{(k)}(t)\right\| \\
&= \sum_{j=0}^{l} \left\|u_n^{(k+j)}(t) - u_m^{(k+j)}(t)\right\| \leq \sum_{j=0}^{l} \frac{C_{k+j}}{t^{k+j}} \|x_n - x_m\|
\end{aligned}$$

for all $m, n \geq 1$ and $l \geq 1$. It's clear that $\{u_n^{(k)}(t)\}$ is a Cauchy sequence in $(D(A^l), \|\cdot\|_l)$.

Suppose that $u_n^{(k)}(t) \to v_k(t)$ in $(D(A^l), \|\cdot\|_l)$. (The limit $v_k(t)$ does not depend on l since $(D(A^l), \|\cdot\|_l)$ can be embedded into $(D(A^{l-1}), \|\cdot\|_{l-1})$ continuously.) Thn it must be $v_0(t) = u(t)$, and hence $u_n(t) \to u(t)$ in $(D(A^l), \|\cdot\|_l)$ for each l. As a consequence,

$$A^k u(t) = \lim_{n\to\infty} A^k u_n(t) = \lim_{n\to\infty} u_n^{(k)}(t) = v_k(t)\,.$$

Fix $l \geq 1$. Observe that $\|u_n^{(k)}(t) - v_k(t)\|_l \to 0$ uniformly for $t \in [\delta, \infty)$ for every $\delta > 0$. We can conclude that the map $u : (0,\infty) \to (D(A^l), \|\cdot\|_l); t \mapsto u(t)$ is C^∞-differentiable with derivatives $u^{(k)}(t) = v_k(t) = A^k u(t)$; that is

$$u \in C^\infty\Big((0,+\infty); (D(A^l), \|\cdot\|_l)\Big)\,.$$

Trivially u is the solution of (ACP′) and part (ii) holds. □

3.9.3 Nonhomogeneous Evolution Equations

We are going to handle the nonhomogeneous evolution equations

$$\begin{cases} u'(t) = Au(t) + f(t)\,, \text{ for } t \geq 0\,; \\ u(0) = x\,, \end{cases} \tag{3.20}$$

where $(A, D(A))$ generates a strongly continuous semigroup $(T(t))_{t\geq 0}$ on the Banach space X, $f : \mathbb{R}_+ \to X$ and $x \in X$. A function $u : \mathbb{R}_+ \to X$ is called a *(classical) solution* of (3.20) if u is continuously differentiable, $u(t) \in D(A)$ for all $t \geq 0$, and (3.20) holds. Clearly any classical solution belongs to

$$C^1(\mathbb{R}_+; X) \cap C(\mathbb{R}_+; (D(A), \|\cdot\|_A))\,.$$

By the Duhamel's principle or the variation of constant method, we guess the solution has the form

$$u(t) = T(t)x + \int_0^t T(t-s)f(s)\,\mathrm{d}s\,.$$

To verify this, we need some assumption on the function f.

Theorem 3.34. *The nonhomogeneous evolution equations* (3.20) *has a unique classical solution* $u : \mathbb{R}_+ \to X$ *given by*

$$u(t) = T(t)x + \int_0^t T(t-s)f(s)\,\mathrm{d}s \quad \textit{for all} \quad t \geq 0\,,$$

provided that

$$f \in C^1(\mathbb{R}_+; X) \quad \textit{or} \quad f \in C(\mathbb{R}_+; (D(A), \|\cdot\|_A))\,.$$

Proof. The uniqueness of the solution follows form Theorem 3.29. It remains to show that u is a classical solution. To this end, let

$$v(t) := \int_0^t T(t-s)f(s)\,\mathrm{d}s = \int_0^t T(s)f(t-s)\,\mathrm{d}s \quad \text{for} \quad t \geq 0\,.$$

It suffices to show that v is a classical solution of (3.20) with the initial value $x = 0$.

Step 1. We show that $v \in C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$.

If $f \in C^1(\mathbb{R}_+; X)$, then we compute, for $h > 0$,

$$\begin{aligned}
\frac{T(h) - I}{h} v(t) &= \frac{1}{h}\left[\int_0^t T(t+h-s)f(s)\,\mathrm{d}s - \int_0^t T(t-s)f(s)\,\mathrm{d}s\right] \\
&= \frac{1}{h}\left[\int_h^{t+h} T(t)f(t-s+h)\,\mathrm{d}s - \int_0^t T(s)f(t-s)\,\mathrm{d}s\right] \\
&= \frac{1}{h}\int_t^{t+h} T(s)f(t-s+h)\,\mathrm{d}s - \frac{1}{h}\int_0^h T(s)f(t-s)\,\mathrm{d}s \\
&\quad + \int_h^t T(s)\frac{f(t-s+h) - f(t-s)}{h}\,\mathrm{d}s\,.
\end{aligned}$$

Letting $h \downarrow 0$, by the dominated convergence theorem, we get $v(t) \in D(A)$ and

$$Av(t) = T(t)f(0) - f(t) + \int_0^t T(s)f'(t-s)\,\mathrm{d}s\,. \tag{3.21}$$

From this we can see that $v \in C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$.

If $f \in C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$, then the mapping $s \mapsto T(t-s)f(s); [0,t] \to (D(A), \|\cdot\|_A)$ is continuous. By the definition of the integral $v(t) \in D(A)$

for all $t \geq 0$. Moreover,

$$Av(t) = \int_0^t T(t-s)Af(s)\,\mathrm{d}s\,. \tag{3.22}$$

From this we can see that $v \in C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$.

Step 2. We show that $v \in C^1(\mathbb{R}_+; X)$ and $v'(t) = f(t) + Av(t)$ for $t \geq 0$.

We compute, for $t \geq 0$ and $h > 0$,

$$\begin{aligned}
&\frac{v(t+h)-v(t)}{h} \\
&= \frac{1}{h}\left[\int_0^{t+h} T(t+h-s)f(s)\,\mathrm{d}s - \int_0^t T(t-s)f(s)\,\mathrm{d}s\right] \\
&= \frac{1}{h}\left[\int_t^{t+h} T(t+h-s)f(s)\,\mathrm{d}s + \int_0^t [T(t+h)-T(t)]f(s)\,\mathrm{d}s\right] \\
&= \frac{1}{h}\int_0^h T(s)f(t+h-s)\,\mathrm{d}s + \frac{T(h)-I}{h}v(t)\,.
\end{aligned}$$

Letting $h \downarrow 0$, by the dominated convergence theorem and the hypothesis on f, we get

$$\lim_{h\downarrow 0} \frac{v(t+h)-v(t)}{h} = f(t) + Av(t)\,.$$

That is v is right-differentiable on $\mathbb{R}_+$. We then show the left-differentiability.

If $f \in C^1(\mathbb{R}_+; X)$, we compute, for $t > 0$ and $h > 0$,

$$\begin{aligned}
&\frac{v(t)-v(t-h)}{h} \\
&= \frac{1}{h}\left[\int_0^t T(s)f(t-s)\,\mathrm{d}s - \int_0^{t-h} T(s)f(t-s-h)\,\mathrm{d}s\right] \\
&= \frac{1}{h}\int_{t-h}^t T(s)f(t-s)\,\mathrm{d}s + \int_0^{t-h} T(s)\frac{f(t-s)-f(t-s-h)}{h}\,\mathrm{d}s\,.
\end{aligned}$$

Letting $h \downarrow 0$, by the dominated convergence theorem and (3.21), we get

$$\lim_{h\downarrow 0} \frac{v(t)-v(t-h)}{h} = T(t)f(0) + \int_0^t T(s)f'(t-s)\,\mathrm{d}s = f(t) + Av(t)\,.$$

If $f \in C(\mathbb{R}_+; (D(A), \|\cdot\|_A))$, we compute, for $t > 0$ and $h > 0$,

$$\begin{aligned}
&\frac{v(t) - v(t-h)}{h} \\
&= \frac{1}{h}\left[\int_0^t T(t-s)f(s)\,\mathrm{d}s - \int_0^{t-h} T(t-s-h)f(s)\,\mathrm{d}s\right] \\
&= \frac{1}{h}\int_{t-h}^t T(t-s)f(s)\,\mathrm{d}s + \int_0^{t-h} \frac{T(t-s) - T(t-s-h)}{h} f(s)\,\mathrm{d}s
\end{aligned}$$

Letting $h \downarrow 0$, by the dominated convergence theorem and (3.22), we get

$$\lim_{h\downarrow 0} \frac{v(t) - v(t-h)}{h} = f(t) + \int_0^t T(t-s)Af(s)\,\mathrm{d}s = f(t) + Av(t)\,.$$

We are done. □

In the ned, we consider the semi-linear equation

$$\begin{cases} u'(t) = Au(t) + f(t, u(t))\,, \text{ for } 0 \le t \le T\,; \\ u(0) = x\,, \end{cases} \tag{3.23}$$

where $(A, D(A))$ generates a strongly continuous semigroup $(T(t))_{t\ge 0}$ on the Banach space X, $f : [0,T] \times X \to X$ is continuous and $x \in X$.

A function $u : [0,T] \to X$ is called a *(classical) solution* of (3.23) if u is continuously differentiable, $u(t) \in D(A)$ for all $t \in [0,T]$ and (3.20) holds. Clearly any classical solution belongs to

$$C^1([0,T]; X) \cap C([0,T]; (D(A), \|\cdot\|_A))\,.$$

A function $u : [0,T] \to X$ is called a *mild solution* of (3.23) if u satisfies the integral equation

$$u(t) = T(t)x + \int_0^t T(t-s)f(s, u(s))\,\mathrm{d}s\,.$$

Lemma 3.35. *If $f(t,x) \in C([0,T] \times X; X)$ satisfies the Lipschitz condition for x, that is, there is some constant $L > 0$ so that*

$$\|f(t, x_1) - f(t, x_2)\| \le L\,\|x_1 - x_2\| \text{ for all } t \in [0,T], x_1, x_2 \in X\,.$$

Then for any initial value x, there exists a unique mild solution for (3.23) *.*

Proof. We shall use the Banach fixed point theorem. Define $F : C([0,T];X) \to C([0,T];X)$ by

$$(Fu)(t) = T(t)x + \int_0^t T(t-s)f(s,u(s))\,\mathrm{d}s \quad \text{for} \quad 0 \le t \le T\,.$$

Then for u, v in $C([0,T];X)$ and $t \in [0,T]$ we have

$$\begin{aligned}\|(Fu)(t) - (Fv)(t)\| &\le \int_0^t \|T(t-s)\|\|f(s,u(s)) - f(s,v(s))\|\,\mathrm{d}s \\ &\le MtL\|u-v\|_{C([0,T];X)}\end{aligned}$$

where $M \coloneqq \sup_{t\in[0,T]} \|T(t)\|$. Then by induction,

$$\|(F^n u)(t) - (F^n v)(t)\| \le \frac{(MLt)^n}{n!}\|u-v\|_{C([0,T];X)}\,.$$

Thus we get

$$\|F^n u - F^n v\| \le \frac{(MLT)^n}{n!}\|u-v\|_{C([0,T];X)}\,.$$

Since the series $\sum_n \frac{(MLT)^n}{n!} < \infty$, by the Banach fixed point theorem, there exists a unique fixed point $u \in C^1([0,T])$ for F. We are done. $\square$

Remark 3.12. Indeed, the mill solution depends continuously on the integral value. To see this, let $u(t,x)$ be the mild solution with initial value $x \in X$. Then

$$\begin{aligned}&\|u(t,x) - u(t,y)\| \\ &\le \|T(t)x - T(t)y\| + \int_0^t \|T(t-s)\|\|f(s,u(s,x)) - f(s,u(s,y))\|\,\mathrm{d}s \\ &\le M\|x-y\| + ML\int_0^t \|u(s,x) - u(s,y)\|\,\mathrm{d}s\end{aligned}$$

By Grönwall's inequality

$$\|u(t,x) - u(t,y)\| \le Me^{MLt}\|x-y\|\,.$$

Thus $\|u(\cdot,x) - u(\cdot,y)\| \le Me^{MLT}\|x-y\|$ as desired.

Exercise 3.9. Suppose the assumption in the preceding lemma holds. Let $g \in C([0,T],X)$. Then show that the integral equation

$$y(t) = g(t) + \int_0^t T(t-s)f(s,y(s))\,\mathrm{d}s$$

has a unique solution $y \in C([0,T],X)$.

Theorem 3.36. *Suppose that $f \in C^1([0,T]\times X;X)$. Then for each initial data $x \in D(A)$, the mill solution u a classical solution.*

Proof. Let $u(t)$ be the unique mild solution with initial data $x \in D(A)$. Then

$$u(t) = T(t)x + \int_0^t T(s)f(t-s,u(t-s))\,\mathrm{d}s\,.$$

Step 1. We shall compute the derivative of u formally. Then we conclude that u' should satisfy the integral equation

$$\begin{aligned} u'(t) =& T(t)Ax + T(t)f(0,x) + \int_0^t T(t-s)\partial_1 f(s,u(s))\,\mathrm{d}s \\ &+ \int_0^t T(t-s)\partial_2 f(s,u(s))u'(s)\,\mathrm{d}s\,. \end{aligned}$$

Note that $\partial_2 f(s,u(s))$ is a bounded linear operator in $\mathcal{B}(X)$. To make the equation look simple, since u is given, we set

$$g(t) \coloneqq T(t)Ax + T(t)f(0,x) + \int_0^t T(t-s)\partial_1 f(s,u(s))\,\mathrm{d}s\,.$$

Then we guess that u' is the (unique) solution of the integral equation

$$y(t) = g(t) + \int_0^t T(t-s)\partial_2 f(s,u(s))y(s)\,\mathrm{d}s\,. \tag{3.24}$$

Step 2. We show that (3.24) has a unique solution $y(t) \in C([0,T];X)$.

Since $f \in C^1([0,T]\times X;X)$ we get $g \in C([0,T];X)$. On the other hand, by the hypothesis $s \mapsto \partial_2 f(s,u(s)); [0,T] \to \mathcal{B}(X)$ is continuous, thus $(st,y) \mapsto \partial_2 f(s,u(s))y$ is continuous and satisfies the Lipschitz condition for y.Thus, by Exercise 3.9, there is unique $y(t) \in C([0,T];X)$ so that

$$y(t) = g(t) + \int_0^t T(t-s)\partial_2 f(s,u(s))y'(s)\,\mathrm{d}s\,.$$

Step 3. We show that $u \in C^1([0,T;X])$ with $u'(t) = y(t)$ for $t \in [0,T]$.

Fix $t \in [0,T]$, for $h > 0$, we have

$$\begin{aligned}\frac{u(t+h)-u(t)}{h} &= \frac{T(t+h)x - T(t)x}{h} \\ &\quad + \frac{1}{h}\int_0^{t+h} T(t+h-s)f(s,u(s))\,\mathrm{d}s - \frac{1}{h}\int_0^t T(t-s)f(s,u(s))\,\mathrm{d}s\,.\end{aligned}$$

The second line can be rewritten as

$$\begin{aligned}&\frac{1}{h}\int_0^h T(t+h-s)f(s,u(s))\,\mathrm{d}s \\ &\quad + \frac{1}{h}\left\{\int_h^{t+h} T(t+h-s)f(s,u(s))\,\mathrm{d}s - \int_0^t T(t-s)f(s,u(s))\,\mathrm{d}s\right\} \\ &= \frac{1}{h}\int_0^h T(t+h-s)f(s,u(s))\,\mathrm{d}s \\ &\quad + \frac{1}{h}\left\{\int_0^t T(t-s)[f(s+h,u(s+h)) - f(s,u(s))]\,\mathrm{d}s\right\}\end{aligned}$$

Since $f \in C^1([0,T] \times X; X)$,

$$\begin{aligned}&f(s+h,u(s+h)) - f(s,u(s)) \\ &= \partial_1 f(s,u(s))h + \partial_2 f(s,u(s))[u(s+h)-u(s)] + r(h; u(s+h)-u(s))\,.\end{aligned}$$

By the finite-increment theorem (Section 10.4, Theorem 1 in *Mathematical Analysis II* by Zorich),

$$\|r(h; u(s+h)-u(s))\| \to 0 \quad \text{as} \quad h \downarrow 0 \quad \text{uniformly for} \quad s \in [0,T]\,.$$

We define $\Delta_h(t) \coloneqq \frac{u(t+h)-u(t)}{h} - y(t)$ for $h > 0$. Then

$$\Delta_h(t) = \frac{T(t+h)x - T(t)x}{h} - T(t)Ax \tag{3.25}$$

$$+ \frac{1}{h}\int_0^h T(t+h-s)f(s,u(s))\,\mathrm{d}s - T(t)f(0,x) \tag{3.26}$$

$$+ \int_0^t T(t-s)\partial_2 f(s,u(s))\Delta_h(t)\,\mathrm{d}s \tag{3.27}$$

$$+ \frac{1}{h}\int_0^t T(t-s)r(h;u(s+h)-u(s))\,\mathrm{d}s\,. \tag{3.28}$$

Note that, by the dominated convergence theorem, (3.25), (3.26) and (3.27) converges to zero as $h \downarrow 0$, we get

$$\|\Delta_h(t)\| \leq \epsilon(h) + M'\int_0^t \|\Delta_h(s)\|\,\mathrm{d}s\,.$$

where

$$\lim_{h\downarrow 0}\epsilon(h) = 0 \quad \text{and} \quad M' \coloneqq M\sup_{t\in[0,T]}\|\partial_2 f(s,u(s))\|\,.$$

Thus again by Grönwall's inequality, we get

$$\|\Delta_h(t)\| \leq \epsilon(h)e^{M't} \text{ for all } t \in [0,T]\,.$$

Letting $h \downarrow 0$, we conclude that

$$\lim_{h\downarrow 0}\|\Delta_h(t)\| = \lim_{h\downarrow 0}\left\|\frac{u(t+h)-u(t)}{h} - y(t)\right\| = 0\,.$$

By the same argument we can show that

$$\lim_{h\downarrow 0}\left\|\frac{u(t)-u(t-h)}{h} - y(t)\right\| = 0\,.$$

Then the desired result follows.

Step 4. Finally we show that u is a classical solution.Note that for $h > 0$ and $t \in [0,T]$,

$$\frac{T(h)-I}{h}u(t) = \frac{u(t+h)-u(t)}{h} - \frac{1}{h}\int_t^{t+h} T(t+h-s)f(s,u(s))\,\mathrm{d}s\,.$$

Since we have shown $u \in C^1([0,T];X)$, then $f(s,u(s)) \in C^1([0,T];X)$. Let $h \downarrow 0$, we get $u(t) \in D(A)$ and

$$Au(t) = u'(t) - f(t,u(t))\,.$$

We are done. $\square$