2

Spectrum

Let $T: \mathcal{D}(T) \to \mathcal{X}$ be a linear transformation, where $\mathcal{X}$ is a nonzero normed space and $\mathcal{D}(T)$, the domain of $T$, is a linear manifold of $\mathcal{X}$. The general notion of spectrum, which applies to bounded or unbounded transformations, goes as follows. Let $F$ denote either the real field $\mathbb{R}$ or the complex field $\mathbb{C}$, and let $I$ be the identity on $\mathcal{X}$. The resolvent set $\rho(T)$ of $T$ is the set of all scalars $\lambda$ in $F$ for which the linear transformation $\lambda I - T: \mathcal{D}(T) \to \mathcal{X}$ has a densely defined continuous inverse. That is,

$$\rho(T) = \{ \lambda \in F : (\lambda I - T)^{-1} \in B[\mathcal{R}(\lambda I - T), \mathcal{D}(T)] \text{ and } \mathcal{R}(\lambda I - T)^{-} = \mathcal{X} \}$$

(see, e.g., [8, Definition 18.2] — there are different definitions of the resolvent set for unbounded linear transformations, but they all coincide for the bounded case). The spectrum $\sigma(T)$ of $T$ is the complement of set $\rho(T)$ in $F$. We shall however restrict the theory to operators on a complex Banach space (i.e., to bounded linear transformations of a complex Banach space to itself).

2.1 Basic Spectral Properties

Throughout this chapter $T: \mathcal{X} \to \mathcal{X}$ will be a bounded linear transformation of $\mathcal{X}$ into itself (i.e., an operator on $\mathcal{X}$), so that $\mathcal{D}(T) = \mathcal{X}$, where $\mathcal{X} \neq \{0\}$ is a complex Banach space. That is, $T \in B[\mathcal{X}]$, where $\mathcal{X}$ is a nonzero complex Banach space. In such a case (i.e., in the unital complex Banach algebra $B[\mathcal{X}]$), Theorem 1.2 ensures that the resolvent set $\rho(T)$ is precisely the set of all complex numbers $\lambda$ for which $\lambda I - T \in B[\mathcal{X}]$ is invertible (i.e., has a bounded inverse on $\mathcal{X}$). Therefore (cf. Theorem 1.1),

$$\rho(T) = \{ \lambda \in \mathbb{C} : \lambda I - T \in \mathcal{G}[\mathcal{X}] \}$$

$$= \{ \lambda \in \mathbb{C} : \lambda I - T \text{ has an inverse in } B[\mathcal{X}] \}$$

$$= \{ \lambda \in \mathbb{C} : \mathcal{N}(\lambda I - T) = \{0\} \text{ and } \mathcal{R}(\lambda I - T) = \mathcal{X} \},$$

and so

$$\sigma(T) = \mathbb{C} \setminus \rho(T) = \{ \lambda \in \mathbb{C} : \lambda I - T \text{ has no inverse in } B[\mathcal{X}] \}$$

$$= \{ \lambda \in \mathbb{C} : \mathcal{N}(\lambda I - T) \neq \{0\} \text{ or } \mathcal{R}(\lambda I - T) \neq \mathcal{X} \}.$$
Theorem 2.1. The resolvent set $\rho(T)$ is nonempty and open, and the spectrum $\sigma(T)$ is compact.

Proof. Take any $T \in \mathcal{B}[\mathcal{X}]$. By the Neumann expansion (Theorem 1.3), if $\|T\| < |\lambda|$, then $\lambda \in \rho(T)$. Equivalently, since $\sigma(T) = \mathbb{C} \setminus \rho(T)$,

$$|\lambda| \leq \|T\| \quad \text{for every } \lambda \in \sigma(T).$$

Thus $\sigma(T)$ is bounded, and therefore $\rho(T) \neq \emptyset$.

Claim. If $\lambda \in \rho(T)$, then the open ball $B_\delta(\lambda)$ with center at $\lambda$ and (positive) radius $\delta = \|(\lambda I - T)^{-1}\|^{-1}$ is included in $\rho(T)$.

Proof. If $\lambda \in \rho(T)$, then $\lambda I - T \in \mathcal{G}[\mathcal{X}]$ so that $(\lambda I - T)^{-1}$ is nonzero and bounded, and hence $0 \leq \|(\lambda I - T)^{-1}\|^{-1} < \infty$. Set $\delta = \|(\lambda I - T)^{-1}\|^{-1}$, let $B_\delta(0)$ be the nonempty open ball of radius $\delta$ about the origin of the complex plane $\mathbb{C}$, and take any $\nu$ in $B_\delta(0)$. Since $|\nu| \leq \|(\lambda I - T)^{-1}\|^{-1}$, it follows that $\|\nu(\lambda I - T)^{-1}\| < 1$. Then $[I - \nu(\lambda I - T)^{-1}] \in \mathcal{G}[\mathcal{X}]$ by Theorem 1.3, and so $(\lambda - \nu)I - T = (\lambda I - T)[I - \nu(\lambda I - T)^{-1}] \in \mathcal{G}[X]$. Thus $\lambda - \nu \in \rho(T)$ so that

$$B_\delta(\lambda) = B_\delta(0) + \lambda = \{\nu \in \mathbb{C} : \nu = \nu + \lambda \text{ for some } \nu \in B_\delta(0)\} \subseteq \rho(T),$$

which completes the proof of the claimed result.

Thus $\rho(T)$ is open (it includes a nonempty open ball centered at each of its points) and so $\sigma(T)$ is closed. Compact in $\mathbb{C}$ means closed and bounded. \qed

Remark. Since $B_\delta(\lambda) \subseteq \rho(T)$, it follows that the distance of any $\lambda$ in $\rho(T)$ to the spectrum $\sigma(T)$ is greater than $\delta$; that is (compare with Proposition 2.E),

$$\lambda \in \rho(T) \quad \text{implies} \quad \|(\lambda I - T)^{-1}\|^{-1} \leq d(\lambda, \sigma(T)).$$

The resolvent function $R_T: \rho(T) \rightarrow \mathcal{G}[\mathcal{X}]$ of an operator $T \in \mathcal{B}[\mathcal{X}]$ is the mapping of the resolvent set $\rho(T)$ of $T$ into the group $\mathcal{G}[\mathcal{X}]$ of all invertible operators from $\mathcal{B}[\mathcal{X}]$ defined by

$$R_T(\lambda) = (\lambda I - T)^{-1} \quad \text{for every } \lambda \in \rho(T).$$

Since $R_T(\lambda) - R_T(\nu) = R_T(\lambda)[R_T(\nu)^{-1} - R_T(\lambda)^{-1}]R_T(\nu)$, it follows that

$$R_T(\lambda) - R_T(\nu) = (\nu - \lambda)R_T(\lambda)R_T(\nu)$$

for every $\lambda, \nu \in \rho(T)$ (because $R_T(\nu)^{-1} - R_T(\lambda)^{-1} = (\nu - \lambda)I$). This is the resolvent identity. Swapping $\lambda$ and $\nu$ in the resolvent identity, it follows that $R_T(\lambda)$ and $R_T(\nu)$ commute for every $\lambda, \nu \in \rho(T)$. Also, $T R_T(\lambda) = R_T(\lambda)T$ for every $\lambda \in \rho(T)$ (since $R_T(\lambda)^{-1} R_T(\lambda) = R_T(\lambda) R_T(\lambda)^{-1}$ trivially).

Let $\Lambda$ be a nonempty open subset of the complex plane $\mathbb{C}$. Take a function $f: \Lambda \rightarrow \mathbb{C}$ and a point $\nu \in \Lambda$. Suppose there exists a complex number
2.1 Basic Spectral Properties

$f'(\nu)$ with the following property. For every $\varepsilon > 0$ there is a $\delta > 0$ such that $\left| \frac{f(\lambda) - f(\nu)}{\lambda - \nu} - f'(\nu) \right| < \varepsilon$ for all $\lambda$ in $\Lambda$ for which $0 < |\lambda - \nu| < \delta$. If there exists such an $f'(\nu) \in \mathbb{C}$, then it is called the derivative of $f$ at $\nu$. If $f'(\nu)$ exists for every $\nu$ in $\Lambda$, then $f: \Lambda \to \mathbb{C}$ is analytic (or holomorphic) on $\Lambda$. A function $f: \mathbb{C} \to \mathbb{C}$ is entire if it is analytic on the whole complex plane $\mathbb{C}$. To prove the next result we need the Liouville Theorem, which says that every bounded entire function is constant.

**Theorem 2.2.** If $\mathcal{X}$ is nonzero, then the spectrum $\sigma(T)$ is nonempty.

**Proof.** Let $T \in \mathcal{B}[\mathcal{X}]$ be an operator on a nonzero complex Banach space $\mathcal{X}$, and let $\mathcal{B}[\mathcal{X}]^*$ stand for the dual of $\mathcal{B}[\mathcal{X}]$. That is, $\mathcal{B}[\mathcal{X}]^* = \mathcal{B}{}^{*}[\mathcal{X}]$, $\mathcal{X}$ is the Banach space of all bounded linear functionals on $\mathcal{B}[\mathcal{X}]$. Since $\mathcal{X} \neq \{0\}$, it follows that $\mathcal{B}[\mathcal{X}] \neq \{O\}$, and so $\mathcal{B}[\mathcal{X}]^* \neq \{0\}$, which is a consequence of the Hahn–Banach Theorem (see, e.g., [66, Corollary 4.64]). Take any nonzero $\eta$ in $\mathcal{B}[\mathcal{X}]^*$ (i.e., a nonzero bounded linear functional $\eta: \mathcal{B}[\mathcal{X}] \to \mathbb{C}$), and consider the composition of it with the resolvent function, $\eta \circ R_T: \rho(T) \to \mathbb{C}$. Recall that $\rho(T) = \mathbb{C} \setminus \sigma(T)$ is nonempty and open in $\mathbb{C}$.

**Claim 1.** If $\sigma(T)$ is empty, then $\eta \circ R_T: \rho(T) \to \mathbb{C}$ is bounded.

**Proof.** The resolvent function $R_T: \rho(T) \to \mathcal{G}[\mathcal{X}]$ is continuous (since scalar multiplication and addition are continuous mappings, and inversion is also a continuous mapping; see, e.g., [66, Problem 4.48]). Thus $\|R_T(\cdot)\|: \rho(T) \to \mathbb{R}$ is continuous. Then $\sup_{|\lambda| \leq \|T\|} \|R_T(\lambda)\| < \infty$ if $\sigma(T) = \emptyset$. Indeed, if $\sigma(T)$ is empty, then $\rho(T) \cap \mathcal{B}[\mathcal{X}]^* = \mathcal{B}[\mathcal{X}]^* = \{\lambda \in \mathbb{C}: |\lambda| \leq \|T\|\}$ is a compact set in $\mathbb{C}$, so that the continuous function $\|R_T(\cdot)\|$ attains its maximum on it by the Weierstrass Theorem: a continuous real-valued function attains its maximum and minimum on any compact set in a metric space. On the other hand, if $\|T\| < |\lambda|$, then $\|R_T(\lambda)\| \leq (|\lambda| - \|T\|)^{-1}$ (cf. Remark that follows Theorem 1.3), so that $\|R_T(\lambda)\| \to 0$ as $|\lambda| \to \infty$. Thus, since $\|R_T(\cdot)\|: \rho(T) \to \mathbb{R}$ is continuous, $\sup_{\|T\| \leq \lambda} \|R_T(\lambda)\| < \infty$. Hence $\sup_{\lambda \in \rho(T)} \|R_T(\lambda)\| < \infty$, and so

$$\sup_{\lambda \in \rho(T)} \|\eta \circ R_T(\lambda)\| = \|\eta\| \sup_{\lambda \in \rho(T)} \|R_T(\lambda)\| < \infty,$$

which completes the proof of Claim 1.

**Claim 2.** $\eta \circ R_T: \rho(T) \to \mathbb{C}$ is analytic.

**Proof.** If $\lambda$ and $\nu$ are distinct points in $\rho(T)$, then

$$\frac{R_T(\lambda) - R_T(\nu)}{\lambda - \nu} + R_T(\nu)^2 = (R_T(\nu) - R_T(\lambda)) R_T(\nu)$$

by the resolvent identity. Set $f = \eta \circ R_T: \rho(T) \to \mathbb{C}$. Let $f': \rho(T) \to \mathbb{C}$ be defined by $f'(\lambda) = -\eta(R_T(\lambda)^2)$ for each $\lambda \in \rho(T)$. Therefore,

$$\left| \frac{f(\lambda) - f(\nu)}{\lambda - \nu} - f'(\nu) \right| = \left| \eta[(R_T(\nu) - R_T(\lambda)) R_T(\nu)] \right| \leq \|\eta\| \|R_T(\nu)\| \|R_T(\nu) - R_T(\lambda)\|.$$
so that \( f: \rho(T) \to \mathbb{C} \) is analytic because \( R_T: \rho(T) \to \mathcal{G}[\mathcal{X}] \) is continuous, which completes the proof of Claim 2.

Thus, by Claims 1 and 2, if \( \sigma(T) = \emptyset \) (i.e., if \( \rho(T) = \mathbb{C} \)), then \( \eta \circ R_T: \mathbb{C} \to \mathbb{C} \) is a bounded entire function, and so a constant function by the Liouville Theorem. But (see proof of Claim 1) \( \| R_T(\lambda) \| \to 0 \) as \( |\lambda| \to \infty \), and hence \( \eta(R_T(\lambda)) \to 0 \) as \( |\lambda| \to \infty \) (since \( \eta \) is continuous). Then \( \eta \circ R_T = 0 \) for all \( \eta \) in \( \mathcal{B}[\mathcal{H}]^* \neq \{0\} \) so that \( R_T = O \) (by the Hahn–Banach Theorem). That is, \( (\lambda I - T)^{-1} = O \) for \( \lambda \in \mathbb{C} \), which is a contradiction. Thus \( \sigma(T) \neq \emptyset \). \( \square \)

Remark. \( \sigma(T) \) is compact and nonempty, and so is its boundary \( \partial \sigma(T) \). Hence, \( \partial \sigma(T) = \partial \rho(T) \neq \emptyset \).

### 2.2 A Classical Partition of the Spectrum

The spectrum \( \sigma(T) \) of an operator \( T \) in \( \mathcal{B}[\mathcal{X}] \) is the set of all scalars \( \lambda \) in \( \mathbb{C} \) for which the operator \( \lambda I - T \) fails to be an invertible element of the algebra \( \mathcal{B}[\mathcal{X}] \) (i.e., fails to have a bounded inverse on \( \mathcal{R}(\lambda I - T) = \mathcal{X} \)). According to the nature of such a failure, \( \sigma(T) \) can be split into many disjoint parts. A classical partition comprises three parts. The set \( \sigma_P(T) \) of those \( \lambda \) for which \( \lambda I - T \) has no inverse (i.e., such that the operator \( \lambda I - T \) is not injective) is the point spectrum of \( T \),

\[
\sigma_P(T) = \{ \lambda \in \mathbb{C}: \mathcal{N}(\lambda I - T) \neq \{0\} \}.
\]

A scalar \( \lambda \in \mathbb{C} \) is an eigenvalue of \( T \) if there exists a nonzero vector \( x \) in \( \mathcal{X} \) such that \( Tx = \lambda x \). Equivalently, \( \lambda \) is an eigenvalue of \( T \) if \( \mathcal{N}(\lambda I - T) \neq \{0\} \).

If \( \lambda \in \mathbb{C} \) is an eigenvalue of \( T \), then the nonzero vectors in \( \mathcal{N}(\lambda I - T) \) are the eigenvectors of \( T \), and \( \mathcal{N}(\lambda I - T) \) is the eigenspace (which is a subspace of \( \mathcal{X} \)), associated with the eigenvalue \( \lambda \). The multiplicity of an eigenvalue is the dimension of the respective eigenspace. Thus the point spectrum of \( T \) is precisely the set of all eigenvalues of \( T \). Now consider the set \( \sigma_C(T) \) of those \( \lambda \) for which \( \lambda I - T \) has a densely defined but unbounded inverse on its range,

\[
\sigma_C(T) = \{ \lambda \in \mathbb{C}: \mathcal{N}(\lambda I - T) = \{0\}, \mathcal{R}(\lambda I - T)^{-} = \mathcal{X} \text{ and } \mathcal{R}(\lambda I - T) \neq \mathcal{X}\}
\]

(see Theorem 1.3), which is referred to as the continuous spectrum of \( T \). The residual spectrum of \( T \) is the set \( \sigma_R(T) \) of all scalars \( \lambda \) such that \( \lambda I - T \) has an inverse on its range that is not densely defined:

\[
\sigma_R(T) = \{ \lambda \in \mathbb{C}: \mathcal{N}(\lambda I - T) = \{0\} \text{ and } \mathcal{R}(\lambda I - T)^{-} \neq \mathcal{X}\}.
\]

The collection \( \{\sigma_P(T), \sigma_C(T), \sigma_R(T)\} \) is a partition (i.e., it is a disjoint covering) of \( \sigma(T) \), which means that they are pairwise disjoint and

\[
\sigma(T) = \sigma_P(T) \cup \sigma_C(T) \cup \sigma_R(T).
\]
2.2 A Classical Partition of the Spectrum

The diagram below, borrowed from [62], summarizes such a partition of the spectrum. The residual spectrum is split into two disjoint parts, \( \sigma_R(T) = \sigma_{R_1}(T) \cup \sigma_{R_2}(T) \), and the point spectrum into four disjoint parts, \( \sigma_P(T) = \bigcup_{n=1}^{N} \sigma_{P_n}(T) \). We adopt the following abbreviated notation: \( T_\lambda = \lambda I - T \), \( \mathcal{N}_\lambda = \mathcal{N}(T_\lambda) \), and \( \mathcal{R}_\lambda = \mathcal{R}(T_\lambda) \). Recall that if \( \mathcal{N}(T_\lambda) = \{0\} \), then its linear inverse \( T_\lambda^{-1} \) on \( \mathcal{R}_\lambda \) is continuous if and only if \( \mathcal{R}_\lambda \) is closed (Theorem 1.2).

<table>
<thead>
<tr>
<th>( \mathcal{N}_\lambda = {0} )</th>
<th>( T_\lambda^{-1} \in \mathcal{B}[\mathcal{R}_\lambda, X] )</th>
<th>( \rho(T) )</th>
<th>( \varnothing )</th>
<th>( \varnothing )</th>
<th>( \sigma_{R_1}(T) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{N}_\lambda \neq {0} )</td>
<td>( T_\lambda^{-1} \notin \mathcal{B}[\mathcal{R}_\lambda, X] )</td>
<td>( \varnothing )</td>
<td>( \sigma_{C}(T) )</td>
<td>( \sigma_{R_2}(T) )</td>
<td>( \varnothing )</td>
</tr>
<tr>
<td>( \sigma_{P_1}(T) )</td>
<td>( \sigma_{P_2}(T) )</td>
<td>( \sigma_{P_3}(T) )</td>
<td>( \sigma_{P_4}(T) )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. § 2.2. A classical partition of the spectrum

Theorem 2.2 says that \( \sigma(T) \neq \varnothing \), but any of the above disjoint parts of the spectrum may be empty (Section 2.7). However, if \( \sigma_P(T) \neq \varnothing \), then a set of eigenvectors associated with distinct eigenvalues is linearly independent.

**Theorem 2.3.** Let \( \{\lambda_\gamma\}_{\gamma \in \Gamma} \) be a family of distinct eigenvalues of \( T \). For each \( \gamma \in \Gamma \) let \( x_\gamma \) be an eigenvector associated with \( \lambda_\gamma \). The set \( \{x_\gamma\}_{\gamma \in \Gamma} \) is linearly independent.

**Proof.** For each \( \gamma \in \Gamma \) take \( 0 \neq x_\gamma \in \mathcal{N}(\lambda_\gamma I - T) \neq \{0\} \), and consider the set \( \{x_\gamma\}_{\gamma \in \Gamma} \) (whose existence is ensured by the Axiom of Choice).

**Claim.** Every finite subset of \( \{x_\gamma\}_{\gamma \in \Gamma} \) is linearly independent.

**Proof.** Take an arbitrary finite subset of \( \{x_\gamma\}_{\gamma \in \Gamma} \), say \( \{x_i\}_{i=1}^{n} \). If \( n = 1 \), then linear independence is trivial. Suppose \( \{x_i\}_{i=1}^{n+1} \) is linearly independent for some \( n \geq 1 \). If \( \{x_i\}_{i=1}^{n+1} \) is not linearly independent, then \( x_{n+1} = \sum_{i=1}^{n} \alpha_i x_i \), where the family \( \{\alpha_i\}_{i=1}^{n} \) of complex numbers has at least one nonzero number. Thus

\[
\lambda_{n+1} x_{n+1} = T x_{n+1} = \sum_{i=1}^{n} \alpha_i T x_i = \sum_{i=1}^{n} \alpha_i \lambda_i x_i.
\]

If \( \lambda_{n+1} = 0 \), then \( \lambda_i \neq 0 \) for every \( i \neq n + 1 \) (because the eigenvalues are distinct) and \( \sum_{i=1}^{n} \alpha_i \lambda_i x_i = 0 \) so that \( \{x_i\}_{i=1}^{n} \) is not linearly independent, which is a contradiction. If \( \lambda_{n+1} \neq 0 \), then \( x_{n+1} = \sum_{i=1}^{n} \alpha_i \lambda_{n+1}^{-1} \lambda_i x_i \), and therefore \( \sum_{i=1}^{n} \alpha_i (1 - \lambda_i^{-1} \lambda_{n+1}) x_i = 0 \) so that \( \{x_i\}_{i=1}^{n} \) is not linearly independent (since \( \lambda_i \neq \lambda_{n+1} \) for every \( i \neq n + 1 \) and \( \alpha_i \neq 0 \) for some \( i \)), which is again a contradiction. This completes the proof by induction: \( \{x_i\}_{i=1}^{n+1} \) is linearly independent.

However, if every finite subset of \( \{x_\gamma\}_{\gamma \in \Gamma} \) is linearly independent, then so is the set \( \{x_\gamma\}_{\gamma \in \Gamma} \) itself (see, e.g., [66, Proposition 2.3]). □
There are some overlapping parts of the spectrum which are commonly used. For instance, the compression spectrum $\sigma_{CP}(T)$ and the approximate point spectrum (or approximation spectrum) $\sigma_{AP}(T)$, which are defined by

$$\sigma_{CP}(T) = \{ \lambda \in \mathbb{C}: \mathcal{R}(\lambda I - T) \text{ is not dense in } \mathcal{X} \}$$

$$= \sigma_{P_3}(T) \cup \sigma_{P_4}(T) \cup \sigma_R(T),$$

$$\sigma_{AP}(T) = \{ \lambda \in \mathbb{C}: \lambda I - T \text{ is not bounded below} \}$$

$$= \sigma_P(T) \cup \sigma_C(T) \cup \sigma_{R_2}(T) = \sigma(T) \setminus \sigma_{R_1}(T).$$

The points of $\sigma_{AP}(T)$ are referred to as the approximate eigenvalues of $T$.

**Theorem 2.4.** The following assertions are pairwise equivalent.

(a) For every $\varepsilon > 0$ there is a unit vector $x_\varepsilon$ in $\mathcal{X}$ such that $\| (\lambda I - T)x_\varepsilon \| < \varepsilon$.

(b) There is a sequence $\{x_n\}$ of unit vectors in $\mathcal{X}$ such that $\| (\lambda I - T)x_n \| \to 0$.

(c) $\lambda \in \sigma_{AP}(T)$.

**Proof.** Clearly (a) implies (b). If (b) holds, then there is no constant $\alpha > 0$ such that $\alpha = \alpha \| x_n \| \leq \| (\lambda I - T)x_n \|$ for all $n$. Thus $\lambda I - T$ is not bounded below, and so (b) implies (c). Conversely, if $\lambda I - T$ is not bounded below, then there is no constant $\alpha > 0$ such that $\alpha \| x \| \leq \| (\lambda I - T)x \|$ for all $x \in \mathcal{X}$ or, equivalently, for every $\varepsilon > 0$ there exists a nonzero $y_\varepsilon$ in $\mathcal{X}$ such that $\| (\lambda I - T)y_\varepsilon \| < \varepsilon \| y_\varepsilon \|$. Set $x_\varepsilon = \| y_\varepsilon \|^{-1} y_\varepsilon$, and hence (c) implies (a). $\square$

**Theorem 2.5.** The approximate point spectrum $\sigma_{AP}(T)$ is a nonempty closed subset of $\mathbb{C}$ that includes the boundary $\partial \sigma(T)$ of the spectrum $\sigma(T)$.

**Proof.** Take an arbitrary $\lambda$ in $\partial \sigma(T)$. Recall that $\rho(T) \neq \emptyset$, $\sigma(T)$ is closed (Theorem 2.1), and $\partial \sigma(T) = \partial \rho(T) = \rho(T)^- \cap \sigma(T)$. Thus $\lambda$ is a point of adherence of $\rho(T)$, and so there exists a sequence $\{\lambda_n\}$ with each $\lambda_n$ in $\rho(T)$ such that $\lambda_n \to \lambda$. Since

$$(\lambda_n I - T) - (\lambda I - T) = (\lambda_n - \lambda)I$$

for every $n$, it follows that $\lambda_n I - T \to \lambda I - T$ in $\mathcal{B}[\mathcal{X}]$.

**Claim.** $\sup_n \| (\lambda_n I - T)^{-1} \| = \infty$.

**Proof.** Since each $\lambda_n$ lies in $\rho(T)$ and $\lambda \in \partial \sigma(T)$ does not lie in $\rho(T)$ (because $\sigma(T)$ is closed), it follows that the sequence $\{\lambda_n I - T\}$ of operators in $\mathcal{G}[\mathcal{X}]$ converges in $\mathcal{B}[\mathcal{X}]$ to $\lambda I - T \in \mathcal{B}[\mathcal{X}] \setminus \mathcal{G}[\mathcal{X}]$. If $\lambda \in \sigma_P(T)$, then there exists $x \neq 0$ in $\mathcal{X}$ such that, if $\sup_n \| (\lambda_n I - T)^{-1} \| < \infty$,

$$0 \neq \| x \| = \| (\lambda_n I - T)^{-1}(\lambda_n I - T)x \|$$

$$\leq \sup_n \| (\lambda_n I - T)^{-1} \| \limsup_n \| (\lambda_n I - T)x \|$$

$$\leq \sup_n \| (\lambda_n I - T)^{-1} \| \| (\lambda I - T)x \| = 0,$$
which is a contradiction. If \( \lambda \not\in \sigma_P(T) \), then \( N(\lambda I - T) = \{0\} \), and hence there exists the inverse \((\lambda I - T)^{-1}\) on \( R(\lambda I - T) \) so that, for each \( n \),
\[
(\lambda_n I - T)^{-1} - (\lambda I - T)^{-1} = (\lambda_n I - T)^{-1}[(\lambda I - T) - (\lambda_n I - T)](\lambda I - T)^{-1}.
\]

If \( \sup_n \| (\lambda_n I - T)^{-1} \| < \infty \), then \( (\lambda_n I - T)^{-1} \to (\lambda I - T)^{-1} \) in \( B[\mathcal{X}] \), because \((\lambda_n I - T) \to (\lambda I - T) \) in \( B[\mathcal{X}] \), and so \((\lambda I - T)^{-1} \in B[\mathcal{X}] \). That is, 
\((\lambda I - T) \in \mathcal{G}[\mathcal{X}] \), which is again a contradiction (since \( \lambda \in \sigma(T) \)). This completes the proof of the claimed result.

Since \( \| (\lambda_n I - T)^{-1} \| = \sup_{\|x\|=1} \| (\lambda_n I - T)^{-1} y \| \), take a unit vector \( y_n \) in \( \mathcal{X} \) for which \( \| (\lambda_n I - T)^{-1} \| - \frac{1}{n} \leq \| (\lambda_n I - T)^{-1} y_n \| \leq \| (\lambda_n I - T)^{-1} \| \) for each \( n \). The above claim ensures that \( \sup_n \| (\lambda_n I - T)^{-1} y_n \| = \infty \), and hence \( \inf_n \| (\lambda_n I - T)^{-1} y_n \|^{-1} = 0 \), so that there exist subsequences of \( \{\lambda_n\} \) and \( \{y_n\} \), say \( \{\lambda_k\} \) and \( \{y_k\} \), for which
\[
\| (\lambda_k I - T)^{-1} y_k \|^{-1} \to 0.
\]

Set \( x_k = (\lambda_k I - T)^{-1} y_k \|^{-1} (\lambda_k I - T)^{-1} y_k \) and get a sequence \( \{x_k\} \) of unit vectors in \( \mathcal{X} \) such that \( \| (\lambda_k I - T) x_k \| = \| (\lambda_k I - T)^{-1} y_k \|^{-1} \). Hence
\[
\| (\lambda I - T) x_k \| = \| (\lambda_k I - T) x_k - (\lambda_k - \lambda) x_k \| \leq \| (\lambda_k I - T)^{-1} y_k \|^{-1} + |\lambda_k - \lambda|.
\]

Since \( \lambda_k \to \lambda \), it then follows that \( \| (\lambda I - T) x_k \| \to 0 \), and so \( \lambda \in \sigma_{AP}(T) \) according to Theorem 2.4. Therefore,
\[
\partial \sigma(T) \subseteq \sigma_{AP}(T).
\]

This inclusion clearly implies that \( \sigma_{AP}(T) \neq \emptyset \) (for \( \sigma(T) \) is closed and non-empty). Finally, take an arbitrary \( \lambda \in \mathbb{C}\setminus \sigma_{AP}(T) \) so that \( \lambda I - T \) is bounded below. Therefore, there exists an \( \alpha > 0 \) for which
\[
\alpha \|x\| \leq \| (\lambda I - T) x \| \leq \| (\nu I - T) x \| + \| (\lambda - \nu) x \|,
\]
and hence \((\alpha - |\lambda - \nu|) \|x\| \leq \| (\nu I - T) x \| \), for all \( x \in \mathcal{X} \) and \( \nu \in \mathbb{C} \). Then \( \nu I - T \) is bounded below for every \( \nu \) such that \( 0 < \alpha - |\lambda - \nu| \). Equivalently, \( \nu \in \mathbb{C}\setminus \sigma_{AP}(T) \) for every \( \nu \) sufficiently close to \( \lambda \) (i.e., if \( |\lambda - \nu| < \alpha \)). Thus the nonempty open ball \( B_\alpha(\lambda) \) centered at \( \lambda \) is included in \( \mathbb{C}\setminus \sigma_{AP}(T) \). Therefore \( \mathbb{C}\setminus \sigma_{AP}(T) \) is open, and so \( \sigma_{AP}(T) \) is closed. \( \square \)

**Remark.** \( \sigma_{AP}(T) = \sigma(T)\setminus \sigma_{R_1}(T) \) is closed in \( \mathbb{C} \) and includes \( \partial \sigma(T) = \partial \rho(T) \). So \( \mathbb{C}\setminus \sigma_{R_1}(T) = \rho(T) \cup \sigma_{AP}(T) = \rho(T) \cup \partial \rho(T) \cup \sigma_{AP}(T) = \rho(T) \setminus \cup \sigma_{AP}(T) \) is closed in \( \mathbb{C} \). Outcome:

\( \sigma_{R_1}(T) \) is an open subset of \( \mathbb{C} \).

For the remainder of this section we assume that \( T \) lies in \( B[H] \), where \( H \) is a nonzero complex Hilbert space. This will bring forth some important
simplifications. A particularly useful instance of such simplifications is the formula for the residual spectrum in the next theorem. First we need the following piece of notation. If $A$ is any subset of $\mathbb{C}$, then set

$$A^* = \{ \overline{\lambda} \in \mathbb{C} : \lambda \in A \}$$

so that $A^{**} = A$, $(\mathbb{C} \setminus A)^* = \mathbb{C} \setminus A^*$, and $(A_1 \cup A_2)^* = A_1^* \cup A_2^*$.

**Theorem 2.6.** If $T^* \in \mathcal{B}[\mathcal{H}]$ is the adjoint of $T \in \mathcal{B}[\mathcal{H}]$, then

$$\rho(T) = \rho(T^*)^*, \quad \sigma(T) = \sigma(T^*)^*,$$

and the residual spectrum of $T$ is given by the formula

$$\sigma_R(T) = \sigma_P(T^*)^* \setminus \sigma_P(T).$$

As for the subparts of the point and residual spectra,

$$\sigma_{P_1}(T) = \sigma_{R_1}(T^*)^*, \quad \sigma_{P_2}(T) = \sigma_{R_2}(T^*)^*,$$

$$\sigma_{P_3}(T) = \sigma_{P_3}(T^*)^*, \quad \sigma_{P_4}(T) = \sigma_{P_4}(T^*)^*.$$

For the compression and approximate point spectra we get

$$\sigma_{CP}(T) = \sigma_P(T^*)^* \quad \sigma_{AP}(T^*)^* = \sigma(T) \setminus \sigma_{P_1}(T),$$

$$\partial \sigma(T) \subseteq \sigma_{AP}(T) \cap \sigma_{AP}(T^*)^* = \sigma(T) \setminus (\sigma_{P_1}(T) \cup \sigma_{R_1}(T)).$$

**Proof.** Since $S \in \mathcal{G}[\mathcal{H}]$ if and only if $S^* \in \mathcal{G}[\mathcal{H}]$, we get $\rho(T) = \rho(T^*)^*$. Hence $\sigma(T^*)^* = (\mathbb{C} \setminus \rho(T))^* = \mathbb{C} \setminus \rho(T^*) = \sigma(T^*)$. Recall that $\mathcal{R}(S)^{-} = \mathcal{R}(S)$ if and only if $\mathcal{R}(S^*)^{-} = \mathcal{R}(S^*)$, and $\mathcal{N}(S) = \{0\}$ if and only if $\mathcal{R}(S^*)^{-} = \{0\}$ (cf. Lemmas 1.4 and 1.5), which means that $\mathcal{R}(S^*)^{-} = \mathcal{H}$. Thus $\sigma_{P_1}(T) = \sigma_{R_1}(T^*)^*$, $\sigma_{P_2}(T) = \sigma_{P_2}(T^*)^*$, $\sigma_{P_3}(T) = \sigma_{P_3}(T^*)^*$, and $\sigma_{P_4}(T) = \sigma_{P_4}(T^*)^*$. Applying the same argument, $\sigma_{C}(T) = \sigma_{C}(T^*)^*$ and $\sigma_{CP}(T) = \sigma_{P}(T^*)^*$. Hence,

$$\sigma_{R}(T) = \sigma_{CP}(T)^* \setminus \sigma_{P}(T) \quad \text{implies} \quad \sigma_{R}(T) = \sigma_{P}(T^*)^* \setminus \sigma_{P}(T).$$

Moreover, by using the above properties and the definition of $\sigma_{AP}(T^*) = \sigma_{P}(T^*) \cup \sigma_{C}(T^*) \cup \sigma_{R_2}(T^*) = \sigma_{CP}(T)^* \cup \sigma_{C}(T^*) \cup \sigma_{P_2}(T)^*$, we get

$$\sigma_{AP}(T^*)^* = \sigma_{CP}(T)^* \cup \sigma_{C}(T) \cup \sigma_{P_2}(T) = \sigma(T) \setminus \sigma_{P_1}(T).$$

Therefore, $\sigma_{AP}(T^*)^* \cap \sigma_{AP}(T) = \sigma(T) \setminus (\sigma_{P_1}(T) \cup \sigma_{R_1}(T))$. But $\sigma(T)$ is closed and $\sigma_{R_1}(T)$ is open (and so is $\sigma_{P_1}(T) = \sigma_{R_1}(T^*)^*$) in $\mathbb{C}$. This implies that $\sigma_{P_1}(T) \cup \sigma_{R_1}(T) \subseteq \sigma(T)^{\circ}$, where $\sigma(T)^{\circ}$ denotes the interior of $\sigma(T)$, and $\partial \sigma(T) \subseteq \sigma(T) \setminus (\sigma_{P_1}(T) \cup \sigma_{R_1}(T))$. \qed

**Remark.** We have just shown that $\sigma_{P_1}(T)$ is an open subset of $\mathbb{C}$. 

### 2.3 Spectral Mapping

Spectral Mapping Theorems are pivotal results in spectral theory. Here we focus on the important particular case for polynomials. Further versions will be considered in Chapter 4. Let \( p: \mathbb{C} \rightarrow \mathbb{C} \) be a polynomial with complex coefficients, take any subset \( A \) of \( \mathbb{C} \), and consider its image under \( p \), viz.,

\[ p(A) = \{ p(\lambda) \in \mathbb{C} : \lambda \in A \}. \]

**Theorem 2.7.** (Spectral Mapping Theorem for Polynomials). Take an operator \( T \in B[X] \) on a complex Banach space \( X \). If \( p \) is any polynomial with complex coefficients, then

\[ \sigma(p(T)) = p(\sigma(T)). \]

**Proof.** To avoid trivialities, let \( p: \mathbb{C} \rightarrow \mathbb{C} \) be an arbitrary nonconstant polynomial with complex coefficients,

\[ p(\lambda) = \sum_{i=0}^{n} \alpha_i \lambda^i, \quad \text{with } n \geq 1 \text{ and } \alpha_n \neq 0, \]

for every \( \lambda \in \mathbb{C} \). Take an arbitrary \( \nu \in \mathbb{C} \) and consider the factorization

\[ \nu - p(\lambda) = \beta_n \prod_{i=1}^{n} (\lambda_i - \lambda), \]

with \( \beta_n = -1^{n+1} \alpha_n \) where \( \{\lambda_i\}_{i=1}^{n} \) are the roots of \( \nu - p(\lambda) \), so that

\[ \nu I - p(T) = \nu I - \sum_{i=0}^{n} \alpha_i T^i = \beta_n \prod_{i=1}^{n} (\lambda_i I - T). \]

If \( \lambda_i \in \rho(T) \) for every \( i = 1, \ldots, n \), then \( \beta_n \prod_{i=1}^{n} (\lambda_i I - T) \in G[X] \) so that \( \nu \in \rho(p(T)) \). Thus if \( \nu \in \sigma(p(T)) \), then there exists \( \lambda_j \in \sigma(T) \) for some \( j = 1, \ldots, n \). However, \( \lambda_j \) is a root of \( \nu - p(\lambda) \), that is,

\[ \nu - p(\lambda_j) = \beta_n \prod_{i=1}^{n} (\lambda_i - \lambda_j) = 0, \]

and so \( p(\lambda_j) = \nu \). Hence, if \( \nu \in \sigma(p(T)) \), then

\[ \nu = p(\lambda_j) \in \{ p(\lambda) \in \mathbb{C} : \lambda \in \sigma(T) \} = p(\sigma(T)) \]

because \( \lambda_j \in \sigma(T) \), and therefore

\[ \sigma(p(T)) \subseteq p(\sigma(T)). \]

Conversely, if \( \nu \in p(\sigma(T)) = \{ p(\lambda) \in \mathbb{C} : \lambda \in \sigma(T) \} \), then \( \nu = p(\lambda) \) for some \( \lambda \in \sigma(T) \). Thus \( \nu - p(\lambda) = 0 \) so that \( \lambda = \lambda_j \) for some \( j = 1, \ldots, n \), and hence
\[
\nu I - p(T) = \beta_n \prod_{i=1}^{n} (\lambda_i I - T) \\
= (\lambda_j I - T) \beta_n \prod_{j \neq i=1}^{n} (\lambda_i I - T) = \beta_n \prod_{j \neq i=1}^{n} (\lambda_i I - T) (\lambda_j I - T)
\]
since \((\lambda_j I - T)\) commutes with \((\lambda_i I - T)\) for every \(i\). If \(\nu \in p(p(T))\), then 
\((\nu I - p(T)) \in \mathcal{G}[\mathcal{X}]\) so that
\[
(\lambda_j I - T) \left( \beta_n \prod_{j \neq i=1}^{n} (\lambda_i I - T) (\nu I - p(T))^{-1} \right) \\
= (\nu I - p(T)) (\nu I - p(T))^{-1} = (\nu I - p(T))^{-1} (\nu I - p(T)) \\
= \left( (\nu I - p(T))^{-1} \beta_n \prod_{j \neq i=1}^{n} (\lambda_i I - T) \right) (\lambda_j I - T).
\]

Then \((\lambda_j I - T)\) has a right and a left inverse (i.e., it is invertible), and so \((\lambda_j I - T) \in \mathcal{G}[\mathcal{X}]\) by Theorem 1.1. Hence, \(\lambda = \lambda_j \in \rho(T)\), which contradicts the fact that \(\lambda \in \sigma(T)\). Conclusion: if \(\nu \in p(\sigma(T))\), then \(\nu \notin \rho(p(T))\), that is, \(\nu \in \sigma(p(T))\). Thus
\[
p(\sigma(T)) \subseteq \sigma(p(T)).
\]

In particular,
\[
\sigma(T^n) = \sigma(T)^n \quad \text{for every} \quad n \geq 0,
\]
that is, \(\nu \in \sigma(T)^n = \{ \lambda^n \in \mathbb{C} : \lambda \in \sigma(T) \}\) if and only if \(\nu \in \sigma(T^n)\), and
\[
\sigma(\alpha T) = \alpha \sigma(T) \quad \text{for every} \quad \alpha \in \mathbb{C},
\]
that is, \(\nu \in \alpha \sigma(T) = \{ \alpha \lambda \in \mathbb{C} : \lambda \in \sigma(T) \}\) if and only if \(\nu \in \sigma(\alpha T)\). Also notice (even though this is not a particular case of the Spectral Mapping Theorem for polynomials) that if \(T \in \mathcal{G}[\mathcal{X}]\), then
\[
\sigma(T^{-1}) = \sigma(T)^{-1},
\]
which means that \(\nu \in \sigma(T)^{-1} = \{ \lambda^{-1} \in \mathbb{C} : 0 \neq \lambda \in \sigma(T) \}\) if and only if \(\nu \in \sigma(T^{-1})\). Indeed, if \(T \in \mathcal{G}[\mathcal{X}]\) (so that \(0 \in \rho(T)\)) and if \(\nu \neq 0\), then 
\(-\nu T^{-1}(\nu^{-1} I - T) = \nu I - T^{-1}\). Thus \(\nu^{-1} \in \rho(T)\) if and only if \(\nu \in \rho(T^{-1})\). Also note that if \(T \in \mathcal{B}[\mathcal{H}]\), then
\[
\sigma(T^*) = \sigma(T)^*
\]
by Theorem 2.6, where \(\mathcal{H}\) is a complex Hilbert space.

The next result is an extension of the Spectral Mapping Theorem for polynomials which holds for normal operators in a Hilbert space \(\mathcal{H}\). If \(\Lambda_1\) and \(\Lambda_2\) are arbitrary subsets of \(\mathbb{C}\) and \(p : \mathbb{C} \times \mathbb{C} \to \mathbb{C}\) is any polynomial in two variables (with complex coefficients), then set
2.3 Spectral Mapping

\[ p(A_1, A_2) = \{ p(\lambda_1, \lambda_2) \in \mathbb{C} : \lambda_1 \in A_1, \lambda_2 \in A_2 \} \]

in particular, with \( A^* = \{ \overline{\lambda} \in \mathbb{C} : \lambda \in A \} \),

\[ p(A, A^*) = \{ p(\lambda, \overline{\lambda}) \in \mathbb{C} : \lambda \in A \}. \]

**Theorem 2.8. (Spectral Mapping Theorem for Normal Operators).** If \( T \in \mathcal{B}[\mathcal{H}] \) is normal and \( p(\cdot, \cdot) \) is a polynomial in two variables, then

\[ \sigma(p(T, T^*)) = p(\sigma(T), \sigma(T^*)) = \{ p(\lambda, \overline{\lambda}) \in \mathbb{C} : \lambda \in \sigma(T) \}. \]

**Proof.** Take any normal operator \( T \in \mathcal{B}[\mathcal{H}] \). If \( p(\lambda, \overline{\lambda}) = \sum_{i,j=0}^{m,n} \alpha_{i,j} \lambda^i \overline{\lambda}^j \), then set \( p(T, T^*) = \sum_{i,j=0}^{m,n} \alpha_{i,j} T^i T^{*j} = p(T^*, T) \). Let \( \mathcal{P}(T, T^*) \) be the collection of all those polynomials \( p(T, T^*) \), which is a commutative subalgebra of \( \mathcal{B}[\mathcal{H}] \) since \( T \) commutes with \( T^* \). Consider the collection \( \mathcal{T} \) of all commutative subalgebras of \( \mathcal{B}[\mathcal{H}] \) containing \( T \) and \( T^* \), which is partially ordered (in the inclusion ordering) and nonempty (e.g., \( \mathcal{P}(T, T^*) \in \mathcal{T} \)). Moreover, every chain in \( \mathcal{T} \) has an upper bound in \( \mathcal{T} \) (the union of all subalgebras in a given chain of subalgebras in \( \mathcal{T} \) is again a subalgebra in \( \mathcal{T} \)). Thus Zorn’s Lemma says that \( \mathcal{T} \) has a maximal element, say \( \mathcal{A}(T) \). Outcome: If \( T \) is normal, then there is a maximal (thus closed) commutative subalgebra \( \mathcal{A}(T) \) of \( \mathcal{B}[\mathcal{H}] \) containing \( T \) and \( T^* \). Since \( \mathcal{P}(T, T^*) \subseteq \mathcal{A}(T) \subseteq \mathcal{T} \), and every \( p(T, T^*) \in \mathcal{P}(T, T^*) \) is normal, \( \mathcal{A}(p(T, T^*)) = \mathcal{A}(T) \) for every nonconstant \( p(T, T^*) \). Furthermore,

\[ \Phi(p(T, T^*)) = p(\Phi(T), \Phi(T^*)) \]

for every homomorphism \( \Phi : \mathcal{A}(T) \to \mathbb{C} \). Thus, by Proposition 2.Q(b),

\[ \sigma(p(T, T^*)) = \{ p(\Phi(T), \Phi(T^*)) \in \mathbb{C} : \Phi \in \hat{\mathcal{A}}(T) \}. \]

Take a surjective homomorphism \( \Phi : \mathcal{A}(T) \to \mathbb{C} \) (i.e., take any \( \Phi \in \hat{\mathcal{A}}(T) \)). Consider the Cartesian decomposition \( T = A + iB \), where \( A, B \in \mathcal{B}[\mathcal{H}] \) are self-adjoint, and so \( T^* = A - iB \) (Proposition 1.O). Thus \( \Phi(T) = \Phi(A) + i\Phi(B) \) and \( \Phi(T^*) = \Phi(A) - i\Phi(B) \). Since \( A = \frac{1}{2}(T + T^*) \) and \( B = -\frac{i}{2}(T - T^*) \) lie in \( \mathcal{P}(T, T^*) \), we get \( \mathcal{A}(A) = \mathcal{A}(B) = \mathcal{A}(T) \). Moreover, since they are self-adjoint, \{ \( \Phi(A) \in \mathbb{C} : \Phi \in \hat{\mathcal{A}}(T) \} = \sigma(A) \subset \mathbb{R} \) and \{ \( \Phi(B) \in \mathbb{C} : \Phi \in \hat{\mathcal{A}}(T) \} = \sigma(B) \subset \mathbb{R} \) (Propositions 2.A and 2.Q(b)), and so \( \Phi(A) \in \mathbb{R} \) and \( \Phi(B) \in \mathbb{R} \). Hence

\[ \Phi(T^*) = \overline{\Phi(T)}. \]

Therefore, since \( \sigma(T^*) = \sigma(T)^* \) for every \( T \in \mathcal{B}[\mathcal{H}] \) by Theorem 2.6, and according to Proposition 2.Q(b),

\[ \sigma(p(T, T^*)) = \{ p(\Phi(T), \overline{\Phi(T)}) \in \mathbb{C} : \Phi \in \hat{\mathcal{A}}(T) \} = \{ p(\lambda, \overline{\lambda}) \in \mathbb{C} : \lambda \in \{ \Phi(T) \in \mathbb{C} : \Phi \in \hat{\mathcal{A}}(T) \} \} = \{ p(\lambda, \overline{\lambda}) \in \mathbb{C} : \lambda \in \sigma(T) \} = p(\sigma(T), \sigma(T^*)) = p(\sigma(T), \sigma(T^*)). \]
2.4 Spectral Radius

The spectral radius of an operator $T \in B[\mathcal{X}]$ on a nonzero complex Banach space $\mathcal{X}$ is the nonnegative number

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

The first identity in the above expression defines the spectral radius $r_\sigma(T)$, and the second one is a consequence of the Weierstrass Theorem (cf. proof of Theorem 2.2) since $\sigma(T) \neq \emptyset$ is compact in $\mathbb{C}$ and the function $|\cdot| : \mathbb{C} \to \mathbb{R}$ is continuous. A straightforward consequence of the Spectral Mapping Theorem for polynomials reads as follows.

**Corollary 2.9.** $r_\sigma(T^n) = r_\sigma(T)^n$ for every $n \geq 0$.

**Proof.** Take an arbitrary nonnegative integer $n$. Theorem 2.7 ensures that $\sigma(T^n) = \sigma(T)^n$. Hence $\nu \in \sigma(T^n)$ if and only if $\nu = \lambda^n$ for some $\lambda \in \sigma(T)$, and so $\sup_{\nu \in \sigma(T^n)} |\nu| = \sup_{\lambda \in \sigma(T)} |\lambda^n| = \sup_{\lambda \in \sigma(T)} |\lambda|^n = (\sup_{\lambda \in \sigma(T)} |\lambda|)^n$. □

If $\lambda \in \sigma(T)$, then $|\lambda| \leq \|T\|$. This follows by the Neumann expansion of Theorem 1.3 (cf. proof of Theorem 2.1). Thus $r_\sigma(T) \leq \|T\|$. Therefore, for every operator $T \in B[\mathcal{X}]$, and for each nonnegative integer $n$,

$$0 \leq r_\sigma(T^n) = r_\sigma(T)^n \leq \|T^n\| \leq \|T\|^n.$$  

Thus $r_\sigma(T) \leq 1$ if $T$ is power bounded (i.e., if $\sup_n \|T^n\| < \infty$). Indeed, in this case, $r_\sigma(T)^n = r_\sigma(T^n) \leq \sup_k \|T^k\|$ and $\lim_n (\sup_k \|T^k\|)^{\frac{n}{k}} = 1$, and so

$$\sup_n \|T^n\| < \infty \quad \text{implies} \quad r_\sigma(T) \leq 1.$$  

**Remark.** If $T$ is a nilpotent operator (i.e., if $T^n = O$ for some $n \geq 1$), then $r_\sigma(T) = 0$, and so $\sigma(T) = \sigma_p(T) = \{0\}$ (cf. Proposition 2.1). An operator $T \in B[\mathcal{X}]$ is quasinilpotent if $r_\sigma(T) = 0$ (i.e., if $\sigma(T) = \{0\}$). Thus every nilpotent is quasinilpotent. Since $\sigma_p(T)$ may be empty for a quasinilpotent operator (cf. Proposition 2.2), these classes are related by proper inclusion:

$$\text{Nilpotent} \subset \text{Quasinilpotent}.$$  

The next result is the well-known Gelfand–Beurling formula for the spectral radius. Its proof requires another piece of elementary complex analysis, viz., every analytic function has a power series representation. That is, if $f : \Lambda \to \mathbb{C}$ is analytic, and if $B_{\alpha,\beta}(\nu) = \{\lambda \in \mathbb{C} : 0 \leq \alpha < |\lambda - \nu| < \beta\}$ lies in the open set $\Lambda \subseteq \mathbb{C}$, then $f$ has a unique Laurent expansion about the point $\nu$, namely, $f(\lambda) = \sum_{k=-\infty}^{\infty} \gamma_k (\lambda - \nu)^k$ for every $\lambda \in B_{\alpha,\beta}(\nu)$.

**Theorem 2.10.** (Gelfand–Beurling Formula).

$$r_\sigma(T) = \lim_n \|T^n\|^{\frac{1}{n}}.$$  


2.4 Spectral Radius 39

Proof. Since \( r_\sigma(T)^n \leq \|T^n\| \) for every positive integer \( n \), and since the limit of the sequence \( \{\|T^n\|^{\frac{1}{n}}\} \) exists by Lemma 1.10, we get

\[
r_\sigma(T) \leq \lim_{n} \|T^n\|^\frac{1}{n}.
\]

For the reverse inequality, proceed as follows. Consider the Neumann expansion (Theorem 1.3) for the resolvent function \( R_T: \rho(T) \rightarrow \mathcal{G}[\mathcal{X}] \),

\[
R_T(\lambda) = (\lambda I - T)^{-1} = \lambda^{-1} \sum_{k=0}^{\infty} T^k \lambda^{-k}
\]
for every \( \lambda \in \rho(T) \) such that \( \|T\| < |\lambda| \), where the above series converges in the (uniform) topology of \( \mathcal{B}[\mathcal{X}] \). Take an arbitrary bounded linear functional \( \eta: \mathcal{B}[\mathcal{X}] \rightarrow \mathbb{C} \) in \( \mathcal{B}[\mathcal{X}]^* \) (cf. proof of Theorem 2.2). Since \( \eta \) is continuous,

\[
\eta(R_T(\lambda)) = \lambda^{-1} \sum_{k=0}^{\infty} \eta(T^k) \lambda^{-k}
\]
for every \( \lambda \in \rho(T) \) such that \( \|T\| < |\lambda| \).

Claim. The above displayed identity holds whenever \( r_\sigma(T) < |\lambda| \).

Proof. \( \lambda^{-1} \sum_{k=0}^{\infty} \eta(T^k) \lambda^{-k} \) is a Laurent expansion of \( \eta(R_T(\lambda)) \) about the origin for every \( \lambda \in \rho(T) \) such that \( \|T\| < |\lambda| \). But \( \eta \circ R_T \) is analytic on \( \rho(T) \) (cf. Claim 2 in Theorem 2.2) so that \( \eta(R_T(\lambda)) \) has a unique Laurent expansion about the origin for every \( \lambda \in \rho(T) \), and hence for every \( \lambda \in \mathbb{C} \) such that \( r_\sigma(T) < |\lambda| \). Then \( \eta(R_T(\lambda)) = \lambda^{-1} \sum_{k=0}^{\infty} \eta(T^k) \lambda^{-k} \), which holds whenever \( r_\sigma(T) \leq \|T\| < |\lambda| \), must be the Laurent expansion about the origin for every \( \lambda \in \mathbb{C} \) such that \( r_\sigma(T) < |\lambda| \), thus proving the claimed result.

Hence, if \( r_\sigma(T) < |\lambda| \), then the series of complex numbers \( \sum_{k=0}^{\infty} \eta(T^k) \lambda^{-k} \) converges, and so \( \eta((\lambda^{-1}T)^k) = \eta(T^k) \lambda^{-k} \rightarrow 0 \), for every \( \eta \) in the dual space \( \mathcal{B}[\mathcal{X}]^* \). This means that the \( \mathcal{B}[\mathcal{X}] \)-valued sequence \( \{(\lambda^{-1}T)^k\} \) converges weakly. Then it is bounded (in the uniform topology of \( \mathcal{B}[\mathcal{X}] \) as a consequence of the Banach–Steinhaus Theorem). That is, the operator \( \lambda^{-1}T \) is power bounded. Thus \( |\lambda|^{-n} \|T^n\| \leq \sup_k \|((\lambda^{-1}T)^k)\| < \infty \), so that

\[
|\lambda|^{-1} \|T^n\|^\frac{1}{n} \leq \left( \sup_k \|((\lambda^{-1}T)^k)\| \right)^\frac{1}{n},
\]
for every \( n \). Therefore, \( |\lambda|^{-1} \lim_n \|T^n\|^\frac{1}{n} \leq 1 \), and so \( \lim_n \|T^n\|^\frac{1}{n} \leq |\lambda| \) for every \( \lambda \in \mathbb{C} \) such that \( r_\sigma(T) < |\lambda| \). That is, \( \lim_n \|T^n\|^\frac{1}{n} \leq r_\sigma(T) + \varepsilon \) for every \( \varepsilon > 0 \). Outcome:

\[
\lim_n \|T^n\|^\frac{1}{n} \leq r_\sigma(T).
\]

What Theorem 2.10 says is that \( r_\sigma(T) = r(T) \), where \( r_\sigma(T) \) is the spectral radius of \( T \) and \( r(T) \) is the limit of the sequence \( \{\|T^n\|^\frac{1}{n}\} \) (whose existence was proved in Lemma 1.10). We shall then adopt one and the same notation (the simplest, of course) for both of them. Thus, from now on, we write
\[ r(T) = \sup_{\lambda \in \sigma(T)} |\lambda| = \max_{\lambda \in \sigma(T)} |\lambda| = \lim_n \|T^n\|^\frac{1}{n}. \]

A normaloid was defined in Section 1.6 as an operator \( T \) for which \( r(T) = \|T\| \). Since \( r_\sigma(T) = r(T) \), it follows that a normaloid operator acting on a complex Banach space is precisely an operator whose norm coincides with the spectral radius. Moreover, since on a complex Hilbert space \( \mathcal{H} \) every normal operator is normaloid, and so is every nonnegative operator, and since \( T^*T \) is always nonnegative, it follows that, for every \( T \in B[\mathcal{H}] \),

\[ r(T^*T) = r(TT^*) = \|T^*T\| = \|TT^*\| = \|T\|^2 = \|T^*\|^2. \]

Further useful properties of the spectral radius follow from Theorem 2.10:

\[ r(\alpha T) = |\alpha| r(T) \quad \text{for every} \quad \alpha \in \mathbb{C} \]

and, if \( \mathcal{H} \) is a complex Hilbert space and \( T \in B[\mathcal{H}] \), then

\[ r(T^*) = r(T). \]

An important application of the Gelfand–Beurling formula is the characterization of uniform stability in terms of the spectral radius. An operator \( T \) in \( B[\mathcal{X}] \) is uniformly stable if the power sequence \( \{T^n\} \) converges uniformly to the null operator (i.e., if \( \|T^n\| \to 0 \)). Notation: \( T^n \uarrow O \).

**Corollary 2.11.** If \( T \in B[\mathcal{X}] \) is an operator on a complex Banach space \( \mathcal{X} \), then the following assertions are pairwise equivalent.

(a) \( T^n \uarrow O \).

(b) \( r(T) < 1 \).

(c) \( \|T^n\| \leq \beta \alpha^n \) for every \( n \geq 0 \), for some \( \beta \geq 1 \) and some \( \alpha \in (0,1) \).

**Proof.** Since \( r(T)^n = r(T^n) \leq \|T^n\| \) for each \( n \geq 1 \), it follows that \( \|T^n\| \to 0 \) implies \( r(T) < 1 \). Now suppose \( r(T) < 1 \) and take any \( \alpha \) in \( (r(T),1) \). Since \( r(T) = \lim_n \|T^n\|^\frac{1}{n} \) (Gelfand–Beurling formula), there is an integer \( n_\alpha \geq 1 \) such that \( \|T^n\| \leq \alpha^n \) for every \( n \geq n_\alpha \). Thus \( \|T^n\| \leq \beta \alpha^n \) for every \( n \geq 0 \) with \( \beta = \max_{0 \leq n \leq n_\alpha} \|T^n\|^{-n_\alpha} \), which clearly implies \( \|T^n\| \to 0 \). \( \square \)

An operator \( T \in B[\mathcal{H}] \) on a complex Hilbert space \( \mathcal{H} \) is strongly stable or weakly stable if the power sequence \( \{T^n\} \) converges strongly or weakly to the null operator (i.e., if \( \|T^n x\| \to 0 \) for every \( x \) in \( \mathcal{H} \), or \( \langle T^n x, y \rangle \to 0 \) for every \( x \) and \( y \) in \( \mathcal{H} \) — equivalently, \( \langle T^n x, x \rangle \to 0 \) for every \( x \) in \( \mathcal{H} \) — cf. Section 1.1). These are denoted by \( T^n \stackrel{s}{\to} O \) and \( T^n \stackrel{w}{\to} O \), respectively. Therefore, from what we have considered so far,

\[ r(T) < 1 \iff T^n \uarrow O \iff T^n \warrow O \iff \sup_n \|T^n\| < \infty \implies r(T) \leq 1. \]
The converses to the above one-way implications fail in general. The next result applies the preceding characterization of uniform stability to extend the Neumann expansion of Theorem 1.3.

**Corollary 2.12.** Let $T \in \mathcal{B}[\mathcal{X}]$ be an operator on a complex Banach space, and let $\lambda \in \mathbb{C}$ be any nonzero complex number.

(a) $r(T) < |\lambda|$ if and only if $\left\{ \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} \right\}$ converges uniformly. In this case we get $\lambda \in \rho(T)$ and $R_{T}(\lambda) = (\lambda I - T)^{-1} = \frac{1}{\lambda} \sum_{k=0}^{\infty} \left( \frac{T}{\lambda} \right)^{k}$ where $\sum_{k=0}^{\infty} \left( \frac{T}{\lambda} \right)^{k}$ denotes the uniform limit of $\left\{ \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} \right\}$.

(b) If $r(T) = |\lambda|$ and $\left\{ \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} \right\}$ converges strongly, then $\lambda \in \rho(T)$ and $R_{T}(\lambda) = (\lambda I - T)^{-1} = \frac{1}{\lambda} \sum_{k=0}^{\infty} \left( \frac{T}{\lambda} \right)^{k}$ where $\sum_{k=0}^{\infty} \left( \frac{T}{\lambda} \right)^{k}$ denotes the strong limit of $\left\{ \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} \right\}$.

(c) If $|\lambda| < r(T)$, then $\left\{ \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} \right\}$ does not converge strongly.

**Proof.** If $\left\{ \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} \right\}$ converges uniformly, then $\left( \frac{T}{\lambda} \right)^{n} \xrightarrow{u} O$, and therefore $|\lambda|^{-1}r(T) = r\left( \frac{T}{\lambda} \right) < 1$ by Corollary 2.11. On the other hand, if $r(T) < |\lambda|$, then $\lambda \in \rho(T)$ so that $\lambda I - T \in \mathcal{G}[\mathcal{X}]$, and also $r\left( \frac{T}{\lambda} \right) = |\lambda|^{-1}r(T) < 1$. Hence, $\left\| \left( \frac{T}{\lambda} \right)^{n} \right\| \leq \beta n$ for every $n \geq 0$, for some $\beta \geq 1$ and $\alpha \in (0, 1)$, according to Corollary 2.11, and so $\sum_{k=0}^{\infty} \left\| \left( \frac{T}{\lambda} \right)^{k} \right\| < \infty$, which means that $\left\{ \left( \frac{T}{\lambda} \right)^{n} \right\}$ is an absolutely summable sequence in $\mathcal{B}[\mathcal{X}]$. Now follow the steps in the proof of Theorem 1.3 to conclude the results in (a). If $\left\{ \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} \right\}$ converges strongly, then $\left( \frac{T}{\lambda} \right)^{n} x \to 0$ in $\mathcal{X}$ for every $x \in \mathcal{X}$ so that $\sup_{n} \left\| \left( \frac{T}{\lambda} \right)^{n} x \right\| < \infty$ for every $x \in \mathcal{X}$, and hence $\sup_{n} \left\| \left( \frac{T}{\lambda} \right)^{n} \right\| < \infty$ (by the Banach–Steinhaus Theorem). Thus $|\lambda|^{-1}r(T) = r\left( \frac{T}{\lambda} \right) \leq 1$, which proves (c). Moreover,

$$
(\lambda I - T)^{\frac{1}{\lambda}} \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} = \frac{1}{\lambda} \sum_{k=0}^{\infty} \left( \frac{T}{\lambda} \right)^{k} (\lambda I - T) = I - \left( \frac{T}{\lambda} \right)^{n+1} \xrightarrow{s} I.
$$

Therefore, $(\lambda I - T)^{-1} = \frac{1}{\lambda} \sum_{k=0}^{\infty} \left( \frac{T}{\lambda} \right)^{k}$, where $\sum_{k=0}^{\infty} \left( \frac{T}{\lambda} \right)^{k} \in \mathcal{B}[\mathcal{X}]$ is the strong limit of $\left\{ \sum_{k=0}^{n} \left( \frac{T}{\lambda} \right)^{k} \right\}$, which concludes the proof of (b). \qed

### 2.5 Numerical Radius

The **numerical range** $W(T)$ of an operator $T \in \mathcal{B}[\mathcal{H}]$ acting on a nonzero complex Hilbert space $\mathcal{H}$ is the (nonempty) set consisting of the inner products $\langle Tx; x \rangle$ for unit vectors $x \in \mathcal{H}$; that is,

$$
W(T) = \{ \lambda \in \mathbb{C} : \lambda = \langle Tx; x \rangle \text{ for some } \|x\| = 1 \}.
$$

It can be shown that $W(T)$ is a convex set in $\mathbb{C}$ (see, e.g., [50, Problem 210]), and it is clear that

$$
W(T^*) = W(T)^*.
$$
Theorem 2.13. \( \sigma_P(T) \cup \sigma_R(T) \subseteq \text{W}(T) \) and \( \sigma(T) \subseteq \text{W}(T)^{-} \).

Proof. Take any operator \( T \in \mathcal{B}[\mathcal{H}] \) on a nonzero complex Hilbert space \( \mathcal{H} \).
If \( \lambda \in \sigma_P(T) \), then there is a unit vector \( x \in \mathcal{H} \) such that \( Tx = \lambda x \).
Hence \( \langle Tx; x \rangle = \lambda \|x\|^2 = \lambda \) so that \( \lambda \in \text{W}(T) \).
If \( \lambda \in \sigma_R(T) \), then \( \lambda \in \text{W}(T) \).
Hence \( \sigma_P(T) \cup \sigma_R(T) \subseteq \text{W}(T) \).

If \( \lambda \in \sigma_{AP}(T) \), then there is a sequence \( \{x_n\} \) of unit vectors in \( \mathcal{H} \) such that \( \|\lambda I - T\| x_n \| \to 0 \) by Theorem 2.4. Hence
\[
0 \leq |\lambda - \langle Tx_n; x_n \rangle| = |\langle (\lambda I - T) x_n; x_n \rangle| \leq \|\lambda I - T\| x_n \| \to 0
\]
so that \( \langle Tx_n; x_n \rangle \to \lambda \).
Since each \( \langle Tx_n; x_n \rangle \) lies in \( \text{W}(T) \), the classical Closed Set Theorem says that \( \lambda \in \text{W}(T)^{-} \).
Thus \( \sigma_{AP}(T) \subseteq \text{W}(T)^{-} \), and so
\[
\sigma(T) = \sigma_R(T) \cup \sigma_{AP}(T) \subseteq \text{W}(T)^{-}.
\]

The numerical radius of an operator \( T \in \mathcal{B}[\mathcal{H}] \) on a nonzero complex Hilbert space \( \mathcal{H} \) is the nonnegative number
\[
w(T) = \sup_{\lambda \in \text{W}(T)} |\lambda| = \sup_{\|x\|=1} |\langle Tx; x \rangle|.
\]
It is readily verified that
\[
w(T^*) = w(T) \quad \text{and} \quad w(T^*T) = \|T\|^2.
\]
Unlike the spectral radius, the numerical radius is a norm on \( \mathcal{B}[\mathcal{H}] \). That is,
\( 0 \leq w(T) \) for every \( T \in \mathcal{B}[\mathcal{H}] \) and \( 0 < w(T) \) if \( T \neq 0 \), \( w(\alpha T) = |\alpha| w(T) \), and \( w(T + S) \leq w(T) + w(S) \) for every \( \alpha \in \mathbb{C} \) and every \( S, T \in \mathcal{B}[\mathcal{H}] \).
However, the numerical radius does not have the operator norm property in the sense that the inequality \( w(ST) \leq w(S) w(T) \) is not true for all operators \( S, T \in \mathcal{B}[\mathcal{H}] \).
Nevertheless, the power inequality holds: \( w(T^n) \leq w(T)^n \) for all \( T \in \mathcal{B}[\mathcal{H}] \) and every positive integer \( n \) (see, e.g., [50, p. 118 and Problem 221]).
Moreover, the numerical radius is a norm equivalent to the (induced uniform) operator norm of \( \mathcal{B}[\mathcal{H}] \) and dominates the spectral radius, as in the next theorem.

Theorem 2.14. \( 0 \leq r(T) \leq w(T) \leq \|T\| \leq 2w(T) \).

Proof. Take any \( T \in \mathcal{B}[\mathcal{H}] \). Since \( \sigma(T) \subseteq \text{W}(T)^{-} \), we get \( r(T) \leq w(T) \). Moreover, \( w(T) = \sup_{\|x\|=1} |\langle Tx; x \rangle| \leq \sup_{\|x\|=1} \|Tx\| = \|T\| \).
Now recall that, by the polarization identity (cf. Proposition 1.A),
\[
\langle Tx; y \rangle = \frac{1}{4} (\langle T(x + y); (x + y) \rangle - \langle T(x - y); (x - y) \rangle + i \langle T(x + iy); (x + iy) \rangle - i \langle T(x - iy); (x - iy) \rangle)
\]
for every $x, y$ in $\mathcal{H}$. Therefore, since $|\langle Tz ; z \rangle| \leq \sup_{||u||=1} |\langle Tu ; u \rangle||z||^2 = w(T)||z||^2$ for every $z \in \mathcal{H}$, it follows that, for every $x, y \in \mathcal{H}$,

$$|\langle Tx ; y \rangle| \leq \frac{1}{4} (|\langle T(x + y) ; (x + y) \rangle| + |\langle T(x - y) ; (x - y) \rangle| + |\langle T(x + iy) ; (x + iy) \rangle| + |\langle T(x - iy) ; (x - iy) \rangle|)$$

$$\leq \frac{1}{4} w(T)(||x + y||^2 + ||x - y||^2 + ||x + iy||^2 + ||x - iy||^2).$$

So it follows by the parallelogram law (cf. Proposition 1.1) that

$$|\langle Tx ; y \rangle| \leq w(T)(||x||^2 + ||y||^2) \leq 2w(T)$$

whenever $||x|| = ||y|| = 1$. Thus, since $||T|| = \sup_{||x||=||y||=1} |\langle Tx ; y \rangle|$ (see, e.g., [66, Corollary 5.71]), it follows that $||T|| \leq 2w(T)$. \hfill \Box

An operator $T \in \mathcal{B}[\mathcal{H}]$ is **spectraloid** if $r(T) = w(T)$. Recall that an operator is normaloid if $r(T) = ||T||$ or, equivalently, if $||T^n|| = ||T||^n$ for every $n \geq 1$ (see Theorem 1.11). The next result is a straightforward application of the previous theorem.

**Corollary 2.15.** Every normaloid operator is spectraloid.

Indeed, $r(T) = ||T||$ implies $r(T) = w(T)$, but $r(T) = ||T||$ also implies $w(T) = ||T||$ (according to Theorem 2.14). Thus $w(T) = ||T||$ is a property of every normaloid operator on $\mathcal{H}$. Actually, this can be viewed as a third definition of a normaloid operator on a complex Hilbert space.

**Theorem 2.16.** $T \in \mathcal{B}[\mathcal{H}]$ is normaloid if and only if $w(T) = ||T||$.

**Proof.** Half of the proof was presented above. It remains to prove that

$$w(T) = ||T|| \quad \text{implies} \quad r(T) = ||T||.$$

Suppose $w(T) = ||T||$ (and $T \neq O$ to avoid trivialities). Recall that $W(T)^-\lambda$ is compact in $\mathbb{C}$ (for $W(T)$ is clearly bounded). Thus $\max_{\lambda \in W(T)^-\lambda} |\lambda| = \sup_{\lambda \in W(T)} |\lambda| = w(T) = ||T||$, and therefore there exists $\lambda \in W(T)^-\lambda$ such that $\lambda = ||T||$. Since $W(T)$ is nonempty, $\lambda$ is a point of adherence of $W(T)$, and so there is a sequence $\{\lambda_n\}$ with each $\lambda_n$ in $W(T)$ such that $\lambda_n \rightarrow \lambda$. This means that there is a sequence $\{x_n\}$ of unit vectors in $\mathcal{H}$ (i.e., $||x_n|| = 1$) such that $\lambda_n = \langle Tx_n ; x_n \rangle \rightarrow \lambda$, where $|\lambda| = ||T|| \neq 0$. Hence, if $S = \lambda^{-1}T \in \mathcal{B}[\mathcal{H}]$, then

$$\langle Sx_n ; x_n \rangle \rightarrow 1.$$

**Claim.** $||Sx_n|| \rightarrow 1$ and $\text{Re} \langle Sx_n ; x_n \rangle \rightarrow 1$.

**Proof.** $|\langle Sx_n ; x_n \rangle| \leq ||Sx_n|| \leq ||S|| = 1$ for each $n$. But $\langle Sx_n ; x_n \rangle \rightarrow 1$ implies $||Sx_n|| \rightarrow 1$ and also $\text{Re} \langle Sx_n ; x_n \rangle \rightarrow 1$ (and so $||Sx_n|| \rightarrow 1$ and also $\text{Re} \langle Sx_n ; x_n \rangle \rightarrow 1$ (since $|\cdot|$ and $\text{Re}(\cdot)$ are continuous functions), which concludes the proof.
Then \( \| (I - S)x_n \|^2 = \| Sx_n - x_n \|^2 = \| Sx_n \|^2 - 2\Re \langle Sx_n; x_n \rangle + \| x_n \|^2 \to 0 \) so that \( 1 \in \sigma_{AP}(S) \subseteq \sigma(S) \) (cf. Theorem 2.4). Hence \( r(S) \geq 1 \), and so \( r(T) = r(\lambda S) = |\lambda| r(S) \geq |\lambda| = \| T \| \), which implies that \( r(T) = \| T \| \) (because \( r(T) \leq \| T \| \) for every operator \( T \)).

Remark. If \( T \in B[\mathcal{H}] \) is spectraloid and quasinilpotent, then \( T = O \). In fact, if \( w(T) = 0 \), then \( T = O \) (since the numerical radius is a norm — also see Theorem 2.14); in particular, if \( w(T) = r(T) = 0 \), then \( T = O \). Therefore, the unique normal (or hyponormal, or normoid, or spectraloid) quasinilpotent operator is the null operator.

**Corollary 2.17.** If there exists \( \lambda \in W(T) \) such that \( |\lambda| = \| T \| \), then \( T \) is normaloid and \( \lambda \in \sigma_P(T) \). In other words, if there exists a unit vector \( x \) such that \( \| T \| = |\langle T x; x \rangle| \), then \( r(T) = w(T) = \| T \| \) and \( \langle T x; x \rangle \in \sigma_P(T) \).

**Proof.** If \( \lambda \in W(T) \) is such that \( |\lambda| = \| T \| \), then \( w(T) = \| T \| \) (Theorem 2.14) so that \( T \) is normaloid (Theorem 2.16). Moreover, since \( \lambda = \langle T x; x \rangle \) for some unit vector \( x \), it follows that \( \| T \| = |\lambda| = |\langle T x; x \rangle| \leq \| T \| \| x \| \leq \| T \| \), and hence \( |\langle T x; x \rangle| = \| T \| \| x \| \). That is, the Schwarz inequality becomes an identity, which implies that \( T x = \alpha x \) for some \( \alpha \in \mathbb{C} \) (see, e.g., [66, Problem 5.2]). Thus \( \alpha \in \sigma_P(T) \). But \( \alpha = \alpha \| x \|^2 = \langle \alpha x; x \rangle = \langle T x; x \rangle = \lambda \). \( \square \)

## 2.6 Spectrum of Compact Operators

The spectral theory of compact operators plays a central role in the Spectral Theorem for compact normal operators of the next chapter. Normal operators were defined on Hilbert spaces; thus we keep on working with compact operators on Hilbert spaces, as we did in Section 1.8, although the spectral theory of compact operators can be equally developed on nonzero complex Banach spaces. So we assume that all operators in this section act on a nonzero complex Hilbert space \( \mathcal{H} \). The main result for characterizing the spectrum of compact operators is the Fredholm Alternative of Corollary 1.20, which can be restated as follows.

**Theorem 2.18.** (Fredholm Alternative). Take \( T \in B_{\infty}[\mathcal{H}] \). If \( \lambda \in \mathbb{C} \setminus \{0\} \), then \( \lambda \in \rho(T) \cup \sigma_P(T) \). Equivalently,

\[
\sigma(T) \setminus \{0\} = \sigma_P(T) \setminus \{0\}.
\]

Moreover, if \( \lambda \in \mathbb{C} \setminus \{0\} \), then \( \dim \mathcal{N}(\lambda I - T) = \dim \mathcal{N}(\overline{\lambda}I - T^*) < \infty \) so that \( \lambda \in \rho(T) \cup \sigma_{P_1}(T) \). Equivalently,

\[
\sigma(T) \setminus \{0\} = \sigma_{P_1}(T) \setminus \{0\}.
\]

**Proof.** Take a compact operator \( T \) on a Hilbert space \( \mathcal{H} \) and a nonzero scalar \( \lambda \) in \( \mathbb{C} \). Corollary 1.20 and the diagram of Section 2.2 ensure that
2.6 Spectrum of Compact Operators

\[ \lambda \in \rho(T) \cup \sigma_P(T) \cup \sigma_R(T) \cup \sigma_{P_4}(T). \]

Also by Corollary 1.20, \( \mathcal{N}(\lambda I - T) = \{0\} \) if and only if \( \mathcal{N}(\lambda I - T^*) = \{0\} \), so that \( \lambda \in \sigma_P(T) \) if and only if \( \lambda \in \sigma_P(T) \) by Theorem 2.6, so that \( \lambda \in \rho(T) \cup \sigma_P(T) \) or, equivalently, \( \lambda \in \rho(T) \cup \sigma_P(T) \) (since \( \lambda \in \rho(T) \cup \sigma_P(T) \)). Therefore,

\[ \sigma(T) \setminus \{0\} = \sigma_P(T) \setminus \{0\} = \sigma_{P_4}(T) \setminus \{0\}. \]

The scalar 0 may be anywhere. That is, if \( T \in \mathcal{B}_{\infty}[\mathcal{H}] \), then \( \lambda = 0 \) may lie in \( \sigma_P(T) \), \( \sigma_R(T) \), \( \sigma_C(T) \), or \( \rho(T) \). However, if \( T \) is a compact operator on a nonzero space \( \mathcal{H} \) and \( 0 \in \rho(T) \), then \( \mathcal{H} \) must be finite dimensional. Indeed, if \( 0 \in \rho(T) \), then \( T^{-1} \in \mathcal{B}[\mathcal{H}] \) so that \( I = T^{-1}T \) is compact (since \( \mathcal{B}_{\infty}[\mathcal{H}] \) is an ideal of \( \mathcal{B}[\mathcal{H}] \)), which implies that \( \mathcal{H} \) is finite dimensional (cf. Proposition 1.4). The preceding theorem in fact is a rewriting of the Fredholm Alternative (and it is also referred to as the Fredholm Alternative). It will be applied often from now on. Here is a first application. Let \( \mathcal{B}_0[\mathcal{H}] \) denote the class of all \textit{finite-rank operators} on \( \mathcal{H} \) (i.e., the class of all operators from \( \mathcal{B}[\mathcal{H}] \) with a finite-dimensional range). Recall that \( \mathcal{B}_0[\mathcal{H}] \subseteq \mathcal{B}_{\infty}[\mathcal{H}] \) (finite-rank operators are compact — cf. Proposition 1.4). Let \( \#A \) denote the cardinality of a set \( A \), so that \( \#A < \infty \) means “\( A \) is a finite set”.

**Corollary 2.19.** If \( T \in \mathcal{B}_0[\mathcal{H}] \), then

\[ \sigma(T) = \sigma_P(T) = \sigma_{P_4}(T) \quad \text{and} \quad \#\sigma(T) < \infty. \]

**Proof.** If \( \dim \mathcal{H} < \infty \), then an injective operator is surjective, and linear manifolds are closed (see, e.g., [66, Problem 2.18 and Corollary 4.29]), and so the diagram of Section 2.2 says that \( \sigma(T) = \sigma_P(T) = \sigma_{P_4}(T) \) (for \( \sigma_{P_4}(T) = \sigma_{R_1}(T^*) \)) according to Theorem 2.6. On the other hand, suppose \( \dim \mathcal{H} = \infty \). Since \( \mathcal{B}_0[\mathcal{H}] \subseteq \mathcal{B}_{\infty}[\mathcal{H}] \), Theorem 2.18 says that

\[ \sigma(T) \setminus \{0\} = \sigma_P(T) \setminus \{0\} = \sigma_{P_4}(T) \setminus \{0\}. \]

Since \( \dim \mathcal{R}(T) < \infty \) and \( \dim \mathcal{H} = \infty \), it follows that \( \mathcal{R}(T)^{-} = \mathcal{R}(T) \neq \mathcal{H} \) and \( \mathcal{N}(T) \neq \{0\} \) (because \( \dim \mathcal{N}(T) + \dim \mathcal{R}(T) = \dim \mathcal{H} \); see, e.g., [66, Problem 2.17]). Then \( 0 \in \sigma_{P_4}(T) \) (cf. diagram of Section 2.2), and therefore

\[ \sigma(T) = \sigma_P(T) = \sigma_{P_4}(T). \]

If \( \sigma_P(T) \) is infinite, then there exists an infinite set of linearly independent eigenvectors of \( T \) (Theorem 2.3). Since every eigenvector of \( T \) lies in \( \mathcal{R}(T) \), this implies that \( \dim \mathcal{R}(T) = \infty \) (because every linearly independent subset of a linear space is included in some Hamel bases — see, e.g., [66, Theorem 2.5]), which is a contradiction. Conclusion: \( \sigma_P(T) \) must be finite.

In particular, the above result clearly holds if \( \mathcal{H} \) is finite dimensional since, as we saw above, \( \dim \mathcal{H} < \infty \) implies \( \mathcal{B}[\mathcal{H}] = \mathcal{B}_0[\mathcal{H}] \).
Corollary 2.20. Take an arbitrary compact operator $T \in \mathcal{B}_\infty[\mathcal{H}]$.

(a) $0$ is the only possible accumulation point of $\sigma(T)$.

(b) If $\lambda \in \sigma(T) \setminus \{0\}$, then $\lambda$ is an isolated point of $\sigma(T)$.

(c) $\sigma(T) \setminus \{0\}$ is a discrete subset of $\mathbb{C}$.

(d) $\sigma(T)$ is countable.

Proof. Let $T$ be a compact operator on $\mathcal{H}$.

Claim. An infinite sequence of distinct points of $\sigma(T)$ converges to zero.

Proof. Let $\{\lambda_n\}_{n=1}^\infty$ be an infinite sequence of distinct points of $\sigma(T)$. Without loss of generality, suppose that every $\lambda_n$ is nonzero. Since $T$ is compact and $0 \neq \lambda_n \in \sigma(T)$, it follows by Theorem 2.18 that $\lambda_n \in \sigma_p(T)$. Let $\{x_n\}_{n=1}^\infty$ be a sequence of eigenvectors associated with $\{\lambda_n\}_{n=1}^\infty$ (i.e., $Tx_n = \lambda_n x_n$ with each $x_n \neq 0$), which is a sequence of linearly independent vectors by Theorem 2.3. For each $n \geq 1$, set

$$\mathcal{M}_n = \text{span}\{x_i\}_{i=1}^n,$$

which is a subspace of $\mathcal{H}$ with $\dim \mathcal{M}_n = n$, and

$$\mathcal{M}_n \subset \mathcal{M}_{n+1}$$

for every $n \geq 1$ (because $\{x_i\}_{i=1}^{n+1}$ is linearly independent and so $x_{n+1}$ lies in $\mathcal{M}_{n+1} \setminus \mathcal{M}_n$). From now on the argument is similar to that in the proof of Theorem 1.18. Since each $\mathcal{M}_n$ is a proper subspace of the Hilbert space $\mathcal{M}_{n+1}$, it follows that there exists $y_{n+1}$ in $\mathcal{M}_{n+1}$ with $\|y_{n+1}\| = 1$ for which

$$\frac{1}{2} < \inf_{u \in \mathcal{M}_n} \|y_{n+1} - u\|.$$ 

Write $y_{n+1} = \sum_{i=1}^{n+1} \alpha_i x_i$ in $\mathcal{M}_{n+1}$ so that

$$(\lambda_{n+1} I - T) y_{n+1} = \sum_{i=1}^{n+1} \alpha_i (\lambda_{n+1} - \lambda_i) x_i = \sum_{i=1}^n \alpha_i (\lambda_{n+1} - \lambda_i) x_i \in \mathcal{M}_n.$$ 

Recall that $\lambda_n \neq 0$ for all $n$, take any pair of integers $1 \leq m < n$, and set

$$y = y_m - \lambda_m^{-1} (\lambda_m I - T) y_m + \lambda_n^{-1} (\lambda_n I - T) y_n$$

so that $T(\lambda_m^{-1} y_m) - T(\lambda_n^{-1} y_n) = y - y_n$. Since $y$ lies in $\mathcal{M}_{n-1}$,

$$\frac{1}{2} < \|y - y_n\| = \|T(\lambda_m^{-1} y_m) - T(\lambda_n^{-1} y_n)\|,$$

which implies that the sequence $\{T(\lambda_n^{-1} y_n)\}$ has no convergent subsequence. Thus, since $T$ is compact, Proposition 1.5 ensures that $\{\lambda_n^{-1} y_n\}$ has no bounded subsequence. That is, $\sup_k |\lambda_k|^{-1} = \sup_k \|\lambda_k^{-1} y_k\| = \infty$, and so $\inf_k |\lambda_k| = 0$ for every subsequence $\{\lambda_k\}_{k=1}^\infty$ of $\{\lambda_n\}_{n=1}^\infty$. Thus $\lambda_n \to 0$, which concludes the proof of the claimed result.

(a) Thus, if $\lambda \neq 0$, then there is no sequence of distinct points in $\sigma(T)$ that converges to $\lambda$; that is, $\lambda \neq 0$ is not an accumulation point of $\sigma(T)$. 

(b) Therefore, every $\lambda$ in $\sigma(T) \setminus \{0\}$ is not an accumulation point of $\sigma(T)$; equivalently, every $\lambda$ in $\sigma(T) \setminus \{0\}$ is an isolated point of $\sigma(T)$.

(c) Hence $\sigma(T) \setminus \{0\}$ consists entirely of isolated points, which means that $\sigma(T) \setminus \{0\}$ is a discrete subset of $\mathbb{C}$.

(d) Since a discrete subset of a separable metric space is countable (see, e.g., [66, Example 3.8]), and since $\mathbb{C}$ is separable, $\sigma(T) \setminus \{0\}$ is countable. □

Corollary 2.21. If an operator $T \in \mathcal{B}[\mathcal{H}]$ is compact and normaloid, then $\sigma_P(T) \neq \emptyset$ and there exists $\lambda \in \sigma_P(T)$ such that $|\lambda| = \|T\|$.

Proof. Suppose $T$ is normaloid (i.e., $r(T) = \|T\|$). Thus $\sigma(T) = \{0\}$ only if $T = O$. If $T = O$ and $\mathcal{H} \neq \{0\}$, then $0 \in \sigma_P(T)$ and $\|T\| = 0$. If $T \neq O$, then $\sigma(T) \neq \{0\}$ and $\|T\| = r(T) = \max_{\lambda \in \sigma(T)} |\lambda|$, so that there exists $\lambda \neq 0$ in $\sigma(T)$ such that $|\lambda| = \|T\|$. Moreover, if $T$ is compact and $\sigma(T) \neq \{0\}$, then $\emptyset \neq \sigma(T) \setminus \{0\} \subseteq \sigma_P(T)$ by Theorem 2.18. Hence $r(T) = \max_{\lambda \in \sigma(T)} |\lambda| = \max_{\lambda \in \sigma_P(T)} |\lambda| = \|T\|$. Thus there exists $\lambda \in \sigma_P(T)$ such that $|\lambda| = \|T\|$.

Corollary 2.22. Every compact hyponormal operator is normal.

Proof. Suppose $T \in \mathcal{B}[\mathcal{H}]$ is a compact hyponormal operator on a nonzero complex Hilbert space $\mathcal{H}$. Corollary 2.21 says that $\sigma_P(T) \neq \emptyset$. Consider the subspace $\mathcal{M} = \left( \sum_{\lambda \in \sigma_P(T)} N(\lambda I - T) \right)^{\perp}$ of Theorem 1.16 with $\{\lambda_{\gamma}\}_{\gamma \in \Gamma} = \sigma_P(T)$. Observe that $\sigma_P(T|_{\mathcal{M}^{\perp}}) = \emptyset$. Indeed, if there is a $\lambda \in \sigma_P(T|_{\mathcal{M}^{\perp}})$, then there exists $0 \neq x \in \mathcal{M}^{\perp}$ such that $\lambda x = T|_{\mathcal{M}^{\perp}} x = T x$, and so $x \in N(\lambda I - T) \subseteq \mathcal{M}$, which is a contradiction. Moreover, recall that $T|_{\mathcal{M}^{\perp}}$ is compact and hyponormal (Propositions 1.0 and 1.1). Thus, if $\mathcal{M}^{\perp} \neq \{0\}$, then Corollary 2.21 says that $\sigma_P(T|_{\mathcal{M}^{\perp}}) \neq \emptyset$, which is another contradiction. Therefore, $\mathcal{M}^{\perp} = \{0\}$ so that $\mathcal{M} = \mathcal{H}$ (see Section 1.3), and hence $T = T|_{\mathcal{H}} = T|_{\mathcal{M}}$ is normal according to Theorem 1.16. □

2.7 Additional Propositions

Proposition 2.A. Let $\mathcal{H} \neq \{0\}$ be a complex Hilbert space and let $\mathbb{T}$ denote the unit circle about the origin of the complex plane.

(a) If $H \in \mathcal{B}[\mathcal{H}]$ is hyponormal, then $\sigma_P(H)^* \subseteq \sigma_P(H^*)$ and $\sigma_R(H^*) = \emptyset$.

(b) If $N \in \mathcal{B}[\mathcal{H}]$ is normal, then $\sigma_P(N^*) = \sigma_P(N)^*$ and $\sigma_R(N) = \emptyset$.

(c) If $U \in \mathcal{B}[\mathcal{H}]$ is unitary, then $\sigma(U) \subseteq \mathbb{T}$.

(d) If $A \in \mathcal{B}[\mathcal{H}]$ is self-adjoint, then $\sigma(A) \subseteq \mathbb{R}$.

(e) If $Q \in \mathcal{B}[\mathcal{H}]$ is nonnegative, then $\sigma(Q) \subseteq [0, \infty)$.

(f) If $R \in \mathcal{B}[\mathcal{H}]$ is strictly positive, then $\sigma(R) \subseteq [\alpha, \infty)$ for some $\alpha > 0$.

(g) If $E \in \mathcal{B}[\mathcal{H}]$ is a nontrivial projection, then $\sigma(E) = \sigma_P(E) = \{0, 1\}$.
Proposition 2.B. Similarity preserves the spectrum and its parts, and so it preserves the spectral radius. That is, let $\mathcal{H}$ and $\mathcal{K}$ be nonzero complex Hilbert spaces. For every $T \in \mathcal{B}[\mathcal{H}]$ and $W \in \mathcal{G}[\mathcal{H}, \mathcal{K}]$,

(a) $\sigma_P(T) = \sigma_P(WTW^{-1})$,
(b) $\sigma_R(T) = \sigma_R(WTW^{-1})$,
(c) $\sigma_C(T) = \sigma_C(WTW^{-1})$.

Hence $\sigma(T) = \sigma(WTW^{-1})$, $\rho(T) = \rho(WTW^{-1})$, and $r(T) = r(WTW^{-1})$. Unitary equivalence also preserves the norm: if $W$ is a unitary transformation, then, in addition, $\|T\| = \|WTW^{-1}\|$.

Proposition 2.C. $\sigma(ST) \setminus \{0\} = \sigma(TS) \setminus \{0\}$ for every $S, T \in \mathcal{B}[\mathcal{H}]$.

Proposition 2.D. If $Q \in \mathcal{B}[\mathcal{H}]$ is nonnegative, then $\sigma(Q^{1/2}) = \sigma(Q)^{1/2}$.

Proposition 2.E. Take any operator $T \in \mathcal{B}[\mathcal{H}]$. Let $d$ denote the usual distance in $\mathbb{C}$. If $\lambda \in \rho(T)$, then

$$r((\lambda I - T)^{-1}) = [d(\lambda, \sigma(T))]^{-1}.$$ 

If $T$ is hyponormal and $\lambda \in \rho(T)$, then

$$\| (\lambda I - T)^{-1} \| = [d(\lambda, \sigma(T))]^{-1}.$$ 

Proposition 2.F. Let $\{\mathcal{H}_k\}$ be a collection of Hilbert spaces, let $\{T_k\}$ be a (similarly indexed) collection of operators with each $T_k$ in $\mathcal{B}[\mathcal{H}_k]$, and consider the (orthogonal) direct sum $\bigoplus_k T_k$ in $\mathcal{B}[\bigoplus_k \mathcal{H}_k]$. Then

(a) $\sigma_P(\bigoplus_k T_k) = \bigcup_k \sigma_P(T_k)$,
(b) $\sigma(\bigoplus_k T_k) = \bigcup_k \sigma(T_k)$ if the collection $\{T_k\}$ is finite.

In general (if the collection $\{T_k\}$ is not finite), then

(c) $(\bigcup_k \sigma(T_k))^{-1} \subseteq \sigma(\bigoplus_k T_k)$ and the inclusion may be proper.

However, if $\| (\lambda I - T_k)^{-1} \| = [d(\lambda, \sigma(T_k))]^{-1}$ for each $k$ and every $\lambda \in \rho(T_k)$,

(d) $(\bigcup_k \sigma(T_k))^{-1} = \sigma(\bigoplus_k T_k)$, which happens whenever each $T_k$ is hyponormal.

Proposition 2.G. An operator $T \in \mathcal{B}[\mathcal{X}]$ on a complex Banach space is normaloid if and only if there is a $\lambda \in \sigma(T)$ such that $|\lambda| = \|T\|$.

Proposition 2.H. For every operator $T \in \mathcal{B}[\mathcal{H}]$ on a complex Hilbert space, $\sigma_R(T) \subseteq \{ \lambda \in \mathbb{C} : |\lambda| < \|T\| \}$.

Proposition 2.I. If $\mathcal{H}$ and $\mathcal{K}$ are complex Hilbert spaces and $T \in \mathcal{B}[\mathcal{H}]$, then
\[ r(T) = \inf_{W \in \mathcal{G}[H,K]} \|WTW^{-1}\|. \]

The spectral radius expression in the preceding proposition ensures that an operator is uniformly stable if and only if it is similar to a strict contraction.

**Proposition 2.J.** If \( T \in \mathcal{B}[\mathcal{X}] \) is a nilpotent operator on a complex Banach space, then \( \sigma(T) = \sigma_p(T) = \{0\} \).

**Proposition 2.K.** An operator \( T \in \mathcal{B}[\mathcal{H}] \) on a complex Hilbert space is spectraloid if and only if \( w(T^n) = w(T)^n \) for every \( n \geq 0 \).

An operator \( T \in \mathcal{B}[\mathcal{H}] \) on a separable infinite-dimensional Hilbert space \( \mathcal{H} \) is diagonalizable if \( Tx = \sum_{k=0}^{\infty} \alpha_k \langle x; e_k \rangle e_k \) for every \( x \in \mathcal{H} \), for some orthonormal basis \( \{e_k\}_{k=0}^{\infty} \) for \( \mathcal{H} \) and some bounded sequence \( \{\alpha_k\}_{k=0}^{\infty} \) of scalars.

**Proposition 2.L.** If \( T \in \mathcal{B}[\mathcal{H}] \) is diagonalizable and \( \mathcal{H} \) is complex, then
\[
\sigma_p(T) = \{ \lambda \in \mathbb{C} : \lambda = \alpha_k \text{ for some } k \geq 1 \}, \quad \sigma_R(T) = \varnothing, \quad \text{and} \quad \sigma_C(T) = \{ \lambda \in \mathbb{C} : \inf_{k} |\lambda - \alpha_k| = 0 \text{ and } \lambda \neq \alpha_k \text{ for every } k \geq 1 \}.
\]

An operator \( S_+ \in \mathcal{B}[\mathcal{K}_+] \) on a Hilbert space \( \mathcal{K}_+ \) is a unilateral shift, and an operator \( S \in \mathcal{B}[\mathcal{K}] \) on a Hilbert space \( \mathcal{K} \) is a bilateral shift, if there exists an infinite sequence \( \{H_k\}_{k=0}^{\infty} \) and an infinite family \( \{H_k\}_{k=-\infty}^{\infty} \) of nonzero pairwise orthogonal subspaces of \( \mathcal{K}_+ \) and \( \mathcal{K} \) such that \( \mathcal{K}_+ = \bigoplus_{k=0}^{\infty} H_k \) and \( \mathcal{K} = \bigoplus_{k=-\infty}^{\infty} H_k \) (cf. Section 1.3), and both \( S_+ \) and \( S \) map each \( H_k \) isometrically onto \( H_{k+1} \), so that each transformation \( U_+(k+1) = S_+|_{H_k} : H_k \to H_{k+1} \), and each transformation \( U_{k+1} = S|_{H_k} : H_k \to H_{k+1} \), is unitary, and therefore \( \dim H_{k+1} = \dim H_k \). Such a common dimension is the multiplicity of \( S_+ \) and \( S \). If \( \mathcal{H}_k = \mathcal{H} \) for all \( k \), then \( \mathcal{K}_+ = \ell^2(\mathcal{H}) = \bigoplus_{k=0}^{\infty} H_k \) and \( \mathcal{K} = \ell^2(\mathcal{H}) = \bigoplus_{k=-\infty}^{\infty} H_k \) are the direct orthogonal sums of countably infinite copies of a single nonzero Hilbert space \( \mathcal{H} \), indexed either by the nonnegative integers or by all integers, which are precisely the Hilbert spaces consisting of all square-summable \( \mathcal{H} \)-valued sequences \( \{x_k\}_{k=0}^{\infty} \) and of all square-summable \( \mathcal{H} \)-valued families \( \{x_k\}_{k=-\infty}^{\infty} \). In this case (if \( \mathcal{H}_k = \mathcal{H} \) for all \( k \)), \( U_+(k+1) = S_+|_{H_k} = U_+ \) and \( U_{k+1} = S|_{H_k} = U \) for all \( k \), where \( U_+ \) and \( U \) are any unitary operators on \( \mathcal{H} \). In particular, if \( U_+ = U = I \), the identity on \( \mathcal{H} \), then \( \mathcal{K}_+ \) and \( \mathcal{K} \) are referred to as the canonical unilateral and bilateral shifts on \( \ell^2(\mathcal{H}) \) and on \( \ell^2(\mathcal{H}) \). The adjoint \( S_+^* \in \mathcal{B}[\mathcal{K}_+] \) of \( S_+ \in \mathcal{B}[\mathcal{K}_+] \) and the adjoint of \( S^* \in \mathcal{B}[\mathcal{K}] \) of \( S \in \mathcal{B}[\mathcal{K}] \) are referred to as a backward unilateral shift and as a backward bilateral shift.

Writing \( \bigoplus_{k=0}^{\infty} x_k \) for \( \{x_k\}_{k=0}^{\infty} \) in \( \bigoplus_{k=0}^{\infty} H_k \), and \( \bigoplus_{k=-\infty}^{\infty} x_k \) for \( \{x_k\}_{k=-\infty}^{\infty} \) in \( \bigoplus_{k=0}^{\infty} H_k \), it follows that \( S_+: \mathcal{K}_+ \to \mathcal{K}_+ \) and \( S^*: \mathcal{K}_+ \to \mathcal{K}_+ \), and \( S: \mathcal{K} \to \mathcal{K} \) and \( S^*: \mathcal{K} \to \mathcal{K} \), are given by the formulas
\[
S_+ x = 0 \oplus \bigoplus_{k=1}^{\infty} U_+(k) x_{k-1} \quad \text{and} \quad S^* x = \bigoplus_{k=0}^{\infty} U_+(k+1)^* x_{k+1}
\]
for all \( x = \bigoplus_{k=0}^{\infty} x_k \) in \( \mathcal{K}_+ = \bigoplus_{k=0}^{\infty} H_k \), with 0 being the origin of \( H_0 \), where \( U_+(k+1) \) is any unitary transformation of \( H_k \) onto \( H_{k+1} \) for each \( k \geq 0 \), and...
Sx = \bigoplus_{k=-\infty}^{\infty} U_k x_{k-1} \quad \text{and} \quad S^* x = \bigoplus_{k=\infty}^{\infty} U_{k+1}^* x_{k+1}

for all \( x = \bigoplus_{k=-\infty}^{\infty} x_k \) in \( K = \bigoplus_{k=-\infty}^{\infty} H_k \), where, for each integer \( k \), \( U_{k+1} \) is any unitary transformation of \( H_k \) onto \( H_{k+1} \). The spectrum of a bilateral shift is simpler than that of a unilateral shift, for bilateral shifts are unitary operators (i.e., besides being isometries as unilateral shifts are, bilateral shifts are normal too). For a full treatment on shifts on Hilbert spaces see [49].

Proposition 2.M. Let \( \mathbb{D} \) and \( T = \partial \mathbb{D} \) denote the open unit disk and the unit circle about the origin of the complex plane, respectively. If \( S_+ \in \mathcal{B}[K] \) is a unilateral shift and \( S \in \mathcal{B}[K] \) is a bilateral shift on complex spaces, then

(a) \( \sigma_P(S_+) = \sigma_R(S_+) = \emptyset \), \( \sigma_R(S_+) = \sigma_P(S_+) = \mathbb{D} \), \( \sigma_C(S_+) = \sigma_C(S_+) = \mathbb{T} \).

(b) \( \sigma(S) = \sigma(S^*) = \sigma_C(S^*) = \sigma_C(S) = \mathbb{T} \).

A unilateral weighted shift \( T_+ = S_+ D_+ \) in \( \mathcal{B}[\ell_2^+(\mathcal{H})] \) is the product of a canonical unilateral shift \( S_+ \) and a diagonal operator \( D_+ = \bigoplus_{k=0}^{\infty} \alpha_k I, \) both in \( \mathcal{B}[\ell_2^+(\mathcal{H})] \), where \( \{ \alpha_k \}_{k=0}^{\infty} \) is a bounded sequence of scalars. A bilateral weighted shift \( T = SD \) in \( \mathcal{B}[\ell_2^2(\mathcal{H})] \) is the product of a canonical bilateral shift \( S \) and a diagonal operator \( D = \bigoplus_{k=-\infty}^{\infty} \alpha_k I, \) both in \( \mathcal{B}[\ell_2^2(\mathcal{H})] \), where \( \{ \alpha_k \}_{k=-\infty}^{\infty} \) is a bounded family of scalars.

Proposition 2.N. Let \( T_+ \in \mathcal{B}[\ell_2^+(\mathcal{H})] \) be a unilateral weighted shift, and let \( T \in \mathcal{B}[\ell_2^2(\mathcal{H})] \) be a bilateral weighted shift, where \( \mathcal{H} \) is complex.

(a) If \( \alpha_k \to 0 \) as \( |k| \to \infty \), then \( T_+ \) and \( T \) are compact and quasinilpotent.

If, in addition, \( \alpha_k \neq 0 \) for all \( k \), then

(b) \( \sigma(T_+) = \sigma_R(T_+) = \sigma_R(T_+) = \{0\} \) and \( \sigma(T_+^*) = \sigma_P(T_+^*) = \sigma_P(T_+^*) = \{0\} \),

(c) \( \sigma(T) = \sigma_C(T) = \sigma_C(T^*) = \sigma(T^*) = \{0\} \).

Proposition 2.O. Let \( T \in \mathcal{B}[\mathcal{H}] \) be an operator on a complex Hilbert space and let \( \mathbb{D} \) be the open unit disk about the origin of the complex plane.

(a) If \( T^n \overset{w}{\to} O \), then \( \sigma_P(T) \subseteq \mathbb{D} \).

(b) If \( T \) is compact and \( \sigma_P(T) \subseteq \mathbb{D} \), then \( T^n \overset{w}{\to} O \).

(c) The concepts of weak, strong, and uniform stabilities coincide for a compact operator on a complex Hilbert space.

Proposition 2.P. Take an operator \( T \in \mathcal{B}[\mathcal{H}] \) on a complex Hilbert space and let \( \mathbb{T} = \partial \mathbb{D} \) be the unit circle about the origin of the complex plane.

(a) \( T^n \overset{w}{\to} O \) if and only if \( T^n \overset{w}{\to} O \) and \( \sigma_C(T) \cap \mathbb{T} = \emptyset \).

(b) If the continuous spectrum does not intersect the unit circle, then the concepts of weak, strong, and uniform stabilities coincide.
The concepts of resolvent set $\rho(T)$ and spectrum $\sigma(T)$ of an operator $T$ in the unital complex Banach algebra $\mathcal{B}[\mathcal{X}]$ as in Section 2.1, namely, $\rho(T) = \{ \lambda \in \mathbb{C} : \lambda I - T \text{ has an inverse in } \mathcal{B}[\mathcal{X}] \}$ and $\sigma(T) = \mathbb{C} \setminus \rho(T)$, of course, hold in any unital complex Banach algebra $\mathcal{A}$, and so does the concept of spectral radius $r(T) = \sup_{\lambda \in \sigma(T)} |\lambda|$, where the Gelfand–Beurling formula of Theorem 2.10, namely, $r(T) = \lim_{n} \|T^n\|^{1/n}$, holds in any unital complex Banach algebra (whose proof is essentially the same as the proof of Theorem 2.10). Recall that a component of a set in a topological space is any maximal (in the inclusion ordering) connected subset of it. A hole of a compact set is any bounded component of its complement. Thus the holes of the spectrum $\sigma(T)$ are the bounded components of the resolvent set $\rho(T) = \mathbb{C} \setminus \sigma(T)$.

If $\mathcal{A}'$ is a closed unital subalgebra of a unital complex Banach algebra $\mathcal{A}$ (for instance, $\mathcal{A} = \mathcal{B}[\mathcal{X}]$ where $\mathcal{X}$ is a Banach space), and if $T \in \mathcal{A}'$, then let $\rho'(T)$ be the resolvent set of $T$ with respect to $\mathcal{A}'$, let $\sigma'(T) = \mathbb{C} \setminus \rho'(T)$ be the spectrum of $T$ with respect to $\mathcal{A}'$, and set $r'(T) = \sup_{\lambda \in \sigma'(T)} |\lambda|$. Recall that a homomorphism (or an algebra homomorphism) between two algebras is a linear transformation between them that also preserves product. Let $\mathcal{A}'$ be a maximal (in the inclusion ordering) commutative subalgebra of a unital complex Banach algebra $\mathcal{A}$ (i.e., a commutative subalgebra of $\mathcal{A}$ that is not included in any other commutative subalgebra of $\mathcal{A}$). Note that $\mathcal{A}'$ is trivially unital, and closed in $\mathcal{A}$ because the closure of a commutative subalgebra of a Banach algebra is again commutative since multiplication is continuous in $\mathcal{A}$. Consider the (unital complex commutative) Banach algebra $\mathbb{C}$ (of all complex numbers). Let $\hat{\mathcal{A}}' = \{ \Phi : \mathcal{A}' \to \mathbb{C} : \Phi \text{ is an homomorphism} \}$ stand for the collection of all algebra homomorphisms of $\mathcal{A}'$ onto $\mathbb{C}$.

**Proposition 2.Q.** Let $\mathcal{A}$ be any unital complex Banach algebra (for instance, $\mathcal{A} = \mathcal{B}[\mathcal{X}]$). If $T \in \mathcal{A}'$, where $\mathcal{A}'$ is any closed unital subalgebra of $\mathcal{A}$, then

(a) $\rho'(T) \subseteq \rho(T)$ and $r'(T) = r(T)$ (invariance of the spectral radius).

Hence $\partial \sigma'(T) \subseteq \partial \sigma(T)$ and $\sigma(T) \subseteq \sigma'(T)$.

Thus $\sigma'(T)$ is obtained by adding to $\sigma(T)$ some holes of $\sigma(T)$.

(b) If $\mathcal{A}'$ is a maximal commutative subalgebra of $\mathcal{A}$, then

$$\sigma(T) = \{ \Phi(T) \in \mathbb{C} : \Phi \in \hat{\mathcal{A}}' \} \quad \text{for each} \quad T \in \mathcal{A}'.$$  

Moreover, in this case,

$$\sigma(T) = \sigma'(T).$$

**Proposition 2.R.** If $\mathcal{A}$ is a unital complex Banach algebra, and if $S,T$ in $\mathcal{A}$ commute (i.e., if $S,T \in \mathcal{A}$ and $ST - TS$), then

$$\sigma(S + T) \subseteq \sigma(S) + \sigma(T) \quad \text{and} \quad \sigma(ST) \subseteq \sigma(S) \cdot \sigma(T).$$
If $M$ is an invariant subspace for $T$, then it may happen that $\sigma(T|M) \not\subseteq \sigma(T)$. Sample: every unilateral shift is the restriction of a bilateral shift to an invariant subspace (see, e.g., [62, Lemma 2.14]). However, if $M$ reduces $T$, then $\sigma(T|M) \subseteq \sigma(T)$ by Proposition 2.F(b). The full spectrum of $T \in B[H]$ (notation: $\sigma(T)^\#$) is the union of $\sigma(T)$ and all bounded components of $\rho(T)$ (i.e., $\sigma(T)^\#$ is the union of $\sigma(T)$ and all holes of $\sigma(T)$).

**Proposition 2.S.** If $M$ is $T$-invariant, then $\sigma(T|M) \subseteq \sigma(T)^\#$. 

**Proposition 2.T.** Let $T \in B[H]$ and $S \in B[K]$ be operators on Hilbert spaces $H$ and $K$. If $\sigma(T) \cap \sigma(S) = \emptyset$, then for every bounded linear transformation $Y \in B[H,K]$ there exists a unique bounded linear transformation $X \in B[H,K]$ such that $XT - SX = Y$. In particular,

$$\sigma(T) \cap \sigma(S) = \emptyset \quad \text{and} \quad XT = SX \quad \Rightarrow \quad X = O.$$ 

This is the Rosenblum Corollary, which will be used to prove the Fuglede Theorems in Chapter 3.

**Notes:** Again, as in the previous chapter, these are basic results that will be needed throughout the text. We will not point out here the original sources but, instead, well-known secondary sources where the reader can find the proofs for those propositions, as well as some deeper discussions on them. Proposition 2.A holds independently of the forthcoming Spectral Theorem (Theorem 3.11) — see, e.g., [66, Corollary 6.18]. A partial converse, however, needs the Spectral Theorem (see Proposition 3.D in Chapter 3). Proposition 2.B is a standard result (see, e.g., [50, Problem 75] and [66, Problem 6.10]), as is Proposition 2.C (see, e.g., [50, Problem 76]). Proposition 2.D also dismisses the Spectral Theorem of the next chapter; it is obtained by the Square Root Theorem (Proposition 1.M) and by the Spectral Mapping Theorem (Theorem 2.7). Proposition 2.E is a rather useful technical result (see, e.g., [63, Problem 6.14]). Proposition 2.F is a synthesis of some sparse results (cf. [28, Proposition I.5.1], [50, Solution 98], [55, Theorem 5.42], [66, Problem 6.37], and Proposition 2.E). For Propositions 2.G and 2.H see [66, Sections 6.3 and 6.4]. The spectral radius formula in Proposition 2.I (see, e.g., [42, p. 22]) ensures that an operator is uniformly stable if and only if it is similar to a strict contraction (hint: use the equivalence between (a) and (b) in Corollary 2.11). For Propositions 2.J and 2.K see [66, Sections 6.3 and 6.4]. The examples of spectra in Propositions 2.L, 2.M, and 2.N are widely known (see, e.g., [66, Examples 6.C, 6.D, 6.E, 6.F, and 6.G]). Proposition 2.O deals with the equivalence between uniform and weak stabilities for compact operators, and it is extended to a wider class of operators in Proposition 2.P (see [63, Problems 8.8. and 8.9]). Proposition 2.Q(a) is readily verified, where the invariance of the spectral radius follows by the Gelfand–Beurling formula. Proposition 2.Q(b) is a key result for proving both Theorem 2.8 (the Spectral Mapping Theorem for normal operators) and also Lemma 5.43 of Chapter 5 (for the characterization
of the Browder spectrum in Theorem 5.44) — see, e.g., [76, Theorems 0.3 and 0.4]. For Propositions 2.R, 2.S, and 2.T see [80, Theorem 11.22], [76, Theorem 0.8], and [76, Corollary 0.13], respectively. Proposition 2.T will play a central role in the proof of the Fuglede Theorems of the next chapter (precisely, in Corollary 3.5 and Theorem 3.17).

Suggested Reading

Bachman and Narici [8]  
Berberian [17]  
Conway [27, 28]  
Douglas [34]  
Dowson [35]  
Dunford and Schwartz [39]  
Fillmore [42]  
Gustafson and Rao [45]  
Halmos [50]  
Herrero [55]  
Istrăţescu [58]  
Kato [60]  
Kubrusly [62, 63, 66]  
Radjavi and Rosenthal [76]  
Rudin [80]  
Taylor and Lay [87]
Spectral Theory of Operators on Hilbert Spaces
Kubrusly, C.S.
2012, X, 197 p. 2 illus., Hardcover
ISBN: 978-0-8176-8327-6
A product of Birkhäuser Basel