Lecture Notes for Week 5 (First Draft)

Continuation of the Proof of Theorem 4.3

It follows from Remark 4.6 that

$$E(\lambda_1)E(\lambda_2) = E(\lambda_2)E(\lambda_1) = E(\lambda_1) \ \text{ for } \lambda_1 \le \lambda_2.$$

Put

$$\mathcal{E}(\lambda_1, \lambda_2) = E(\lambda_2) - E(\lambda_1),$$

and observe that

$$\mathcal{E}(\lambda_1, \lambda_2) \ge 0 \ \text{ for } \lambda_1 \le \lambda_2.$$

Observe further that

$$\begin{aligned} E(\lambda_2)\mathcal{E}(\lambda_1, \lambda_2) &= E(\lambda_2)^2 - E(\lambda_2)E(\lambda_1) \\ &= E(\lambda_2) - E(\lambda_1) = \mathcal{E}(\lambda_1, \lambda_2) \ \text{ for } \lambda_1 \le \lambda_2, \end{aligned} \tag{1}$$

and also that

$$\begin{aligned} E(\lambda_1)\mathcal{E}(\lambda_1, \lambda_2) &= E(\lambda_1)E(\lambda_2) - E(\lambda_1)^2 \\ &= E(\lambda_1) - E(\lambda_1) = 0 \ \text{ for } \lambda_1 \le \lambda_2. \end{aligned} \tag{2}$$

Moreover, we have

$$L(\lambda)E(\lambda) = E(\lambda)L(\lambda) = -L(\lambda)^- \tag{3}$$

and

$$L(\lambda)[I - E(\lambda)] = [I - E(\lambda)]L(\lambda) = L(\lambda)^+. \tag{4}$$

Using (1) and (3) we find that

$$\begin{aligned} L(\lambda_2)\mathcal{E}(\lambda_1, \lambda_2) &= L(\lambda_2)E(\lambda_2)\mathcal{E}(\lambda_1, \lambda_2) \\ &= -L(\lambda_2)^-\mathcal{E}(\lambda_1, \lambda_2) \\ &\le 0 \ \text{ for } \lambda_1 \le \lambda_2. \end{aligned} \tag{5}$$

Using (2) and (4) we find that

$$\begin{aligned} L(\lambda_1)\mathcal{E}(\lambda_1, \lambda_2) &= L(\lambda_1)[I - E(\lambda_1)]\mathcal{E}(\lambda_1, \lambda_2) \\ &= L(\lambda_1)^+\mathcal{E}(\lambda_1, \lambda_2) \\ &\ge 0 \ \text{ for } \lambda_1 \le \lambda_2. \end{aligned} \tag{6}$$

Combining (5) and (6) we arrive at

$$\lambda_1 \mathcal{E}(\lambda_1, \lambda_2) \leq A\mathcal{E}(\lambda_1, \lambda_2) \leq \lambda_2 \mathcal{E}(\lambda_1, \lambda_2) \quad \text{for } \lambda_1 \leq \lambda_2. \tag{7}$$

Let $a < m$ and $b \geq M$ be given and take any partition

$$a = \lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n = b.$$

of $[a, b]$. Put

$$\delta = \max\{\lambda_k - \lambda_{k-1}, k = 1, 2, \cdots, n\}.$$

Using (7) we find that

$$\sum_{k=1}^{n} \lambda_{k-1}[E(\lambda_k) - E(\lambda_{k-1})] \leq A \sum_{k=1}^{n} [E(\lambda_k - E(\lambda_{k-1})] \leq \sum_{k=1}^{n} \lambda_k [E(\lambda_k) - E(\lambda_{k-1})] \tag{8}$$

Since

$$\sum_{k=1}^{n} [E(\lambda_k - E(\lambda_{k-1})] = I,$$

it follows from (8) that

$$\sum_{k=1}^{n} \lambda_{k-1}[E(\lambda_k) - E(\lambda_{k-1})] \leq A \leq \sum_{k=1}^{n} \lambda_k [E(\lambda_k - E(\lambda_{k-1})]. \tag{9}$$

For any choice of

$$\lambda_k^* \in [\lambda_{k-1}, \lambda_k], \quad k = 1, 2, \cdots, n,$$

it follows from (9) and a simple computation that

$$\|A - \sum_{k=1}^{n} \lambda_k^* [E(\lambda_k - E(\lambda_{k-1})]\| \leq \delta,$$

which implies that

$$A = \int_a^b \lambda \, dE(\lambda).$$

To prove right continuity of the mapping $\lambda \to E(\lambda)$ in the strong operator topology, we make use of Problem 1 from Assignment 2 concerning bounded monotonic sequence of self-adjoint operators.

Fix $\lambda \in \mathbb{R}$ and notice that $\mathcal{E}(\lambda_1, \lambda_2)$ is nondecreasing in λ_2 and is bounded above. (In addition $\mathcal{E}(\lambda_1, \lambda_2)\mathcal{E}(\lambda_1, \hat{\lambda}_2) = \mathcal{E}(\lambda_1, \hat{\lambda}_2)\mathcal{E}(\lambda_1, \lambda_2)$.) Consequently there is a bounded self-adjoint operator $G(\lambda_1)$ such that

$$\forall x \in X, \quad \text{we have} \quad \lim_{\lambda_2 \downarrow \lambda_1} \mathcal{E}(\lambda_1, \lambda_2)x = G(\lambda_1)x.$$

We need to show that $G(\lambda_1) = 0$. Letting $\lambda_2 \downarrow \lambda_1$ in (7) we find that

$$\lambda_1(G(\lambda_1)x, x) \leq (AG(\lambda_1)x, x) \leq \lambda_1(G(\lambda_1)x, x) \quad \text{for all } x \in X. \tag{10}$$

It follows from (10) that

$$0 \leq L(\lambda_1)G(\lambda_1) \leq 0,$$

and since $L(\lambda_1)G(\lambda_1)$ is self-adjoint, Corollary 1.8 implies that

$$L(\lambda_1)G(\lambda_1) = 0. \tag{11}$$

Using (4) we have

$$L(\lambda_1)^+G(\lambda_1) = [I - E(\lambda_1)]L(\lambda_1)G(\lambda_1) = 0,$$

which implies that

$$\mathcal{R}(G(\lambda_1)) \subset \mathcal{N}(L(\lambda_1)^+),$$

and consequently

$$E(\lambda_1)G(\lambda_1) = G(\lambda_1).$$

Letting $\lambda_2 \downarrow \lambda_1$ in (2), we obtain

$$E(\lambda_1)G(\lambda_1) = 0,$$

and consequently

$$G(\lambda_1) = 0. \quad \square$$

Remark 5.1:

(a) The family $(E(\lambda)|\lambda \in \mathbb{R})$ of projections in Theorem 4.3 is called the *spectral resolution of the identity corresponding to* A or the *family of spectral projections for* A.

(b) Under the assumptions of Theorem 4.1, we also have

$$\text{For all } \lambda_0 \in \mathbb{R} \text{ and all } x \in X, \text{ we have } \lim_{\lambda \uparrow \lambda_0} E(\lambda)x = E(\lambda_0)x.$$

(c) Under the assumptions of Theorem 4.3, we also have

$$A^n = \int_a^b \lambda^n \, dE(\lambda) \quad \text{for all } n \in \mathbb{N}, \; a < m, \; b \geq M.$$

(d) Some authors construct the spectral family in such a way that the mapping $\lambda \to E(\lambda)$ is continuous from the left in the strong operator topology. In this case, the integral representations are valid for $a \leq m$, $b > M$.

Definition 5.2: Let X be a Banach space and $T \in \mathcal{L}(X;X)$ be given.

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

$$\rho(T) = \{\lambda \in \mathbb{K} : \lambda I - T \text{ is bijective}\}.$$

(b) The *spectrum* of T, denoted $\sigma(T)$ is defined by

$$\sigma(T) = \mathbb{K} \backslash \rho(T).$$

(c) A number $\lambda \in \mathbb{K}$ is called an *eigenvalue* for T provided $\mathcal{N}(\lambda I - T) \neq \{0\}$.

(d) If λ is an eigenvalue for T, the nonzero elements of $\mathcal{N}(\lambda I - T)$ are called eigenvectors corresponding to λ.

(e) The set of all eigenvalues of T is called the point spectrum of T and is denoted by $\sigma_p(T)$.

(f) A number $\lambda \in \mathbb{K}$ is called a *generalized eigenvalue* or *approximate eigenvalue* provided that

$$\inf\{\|(\lambda I - T)x) : x \in X, \|x\| = 1\} = 0.$$

(g) The set of all generalized eigenvalues is called the *approximate point spectrum* of T and is denoted by $\sigma_{ap}(T)$.

Proposition 5.3: Let X be a Banach space and $T \in \mathcal{L}(X;X)$ be given. Assume that $\lambda_1, \lambda_2, \cdots, \lambda_m$ are eigenvalues for T and let $x_1, x_2, \cdots, x_m$ be corresponding eigenvectors. Assume further that $\lambda_j \neq \lambda_k$ for $j \neq k$ (i.e. that the eigenvalues are distinct). Then $\{x_1, x_2, \cdots, x_m\}$ is a linearly independent set.

The proof is the same as in the finite dimensional setting and will therefore be omitted.

Proposition 5.4: Let X be a Banach space and let $T \in \mathcal{L}(X;X)$ with $\|T\| < 1$ be given. Then $1 \in \rho(T)$ and

$$(I - T)^{-1} = \sum_{k=0}^{\infty} T^k.$$

Proof: Observe that

$$\|T^k\| \leq \|T\|^k \text{ for all } k \in \mathbb{N}.$$

Since $\|T\| < 1$, the series

$$\sum_{k=0}^{\infty} \|T\|^k$$

converges, and consequently the series

$$\sum_{k=0}^{\infty} \|T^k\|$$

converges. Since $\mathcal{L}(X;X)$ is complete, absolute summability implies summability, so the series

$$\sum_{k=0}^{\infty} T^k$$

converges in the uniform operator topology.

Put

$$S_n = \sum_{k=0}^{n} T^k \ \text{ for all } n \in \mathbb{N},$$

and notice that

$$(I-T)S_n = I - T^{n+1} = S_n(I-T) \ \text{ for all } n \in \mathbb{N}. \tag{12}$$

Since $\|T^{n+1}\| \to 0$ as $n \to \infty$, we conclude from (12) that

$$(I-T)\sum_{k=0}^{\infty} T^k = I = \left(\sum_{k=0}^{\infty} T^k\right)(I-T). \quad \square$$

Corollary 5.5: Let $T \in \mathcal{L}(X;X)$ and $\lambda \in \mathbb{K}$ with $|\lambda| > \|T\|$ be given. Then $\lambda \in \rho(T)$.

Let $\lambda_0 \in \rho(T)$ and $\lambda \in \mathbb{K}$ be given. Observe that

$$\begin{aligned}(\lambda I - T) &= (\lambda_0 I - T) + (\lambda - \lambda_0)I \\ &= (\lambda_0 I - T)[I + (\lambda - \lambda_0)(\lambda_0 I - T)^{-1}] \\ &= (\lambda_0 I - T)[I - (\lambda_0 - \lambda)R(\lambda_0;T)].\end{aligned}$$

If $|\lambda - \lambda_0| \cdot \|R(\lambda_0;T)\| < 1$ then we can apply Proposition 5.4 to conclude that $\lambda I - T$ is bijective and

$$\begin{aligned}R(\lambda;T) &= \left(\sum_{k=0}^{\infty}(\lambda_0 - \lambda)^k R(\lambda_0;T)^k\right) R(\lambda_0;T) \\ &= \sum_{k=0}^{\infty}(-1)^k(\lambda - \lambda_0)^k R(\lambda_0;)^{k+1}.\end{aligned}$$

We have just proved the following result.

Proposition 5.6: Let X be a Banach space and $T \in \mathcal{L}(X;X)$ be given. Then

(i) $\rho(T)$ is open,

(ii) $\sigma(T)$ is closed,

(iii) for all $\lambda \in \rho(T)$ and all $\lambda \in \mathbb{K}$ with $|\lambda - \lambda_0| \cdot \|R(\lambda_0; T)\| < 1$ we have $\lambda \in \rho(T)$ and

$$R(\lambda; T) = \sum_{n=1}^{\infty} (-1)^n (\lambda - \lambda_0)^n R(\lambda_0; T)^{n+1},$$

(iv) the mapping $\lambda \to R(\lambda; T)$ is analytic on $\rho(T)$,

(v) for all $\lambda_0 \in \rho(T)$ and all $n \in \mathbb{N}$ we have

$$R^{(n)}(\lambda_0; T) = (-1)^n n! R(\lambda_0; T)^{n+1}.$$

Here $R^{(n)}$ is the n^{th} derivative of R with respect to the first argument.

Proposition 5.7: Let X be a Banach space and $S, T \in \mathcal{L}(X; X)$. Let $\lambda, \mu \in \rho(T)$ be given. Then

(i) $R(\lambda; T) - R(\mu; T) = (\mu - \lambda) R(\lambda; T) R(\mu; T)$,

(ii) $R(\lambda; T) R(\mu; T) = R(\mu; T) R(\lambda; T)$

(iii) If $ST = TS$ then $SR(\lambda; T) = R(\lambda; T)S$.

Proof: Let $\lambda, \mu \in \rho(T)$ be given. Then we have

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

which establishes (i). Part (ii) follows from part (i) by interchanging λ and μ. To prove part (iii), observe that

$$S(\lambda I - T) = (\lambda I - T)S.$$

Multiplying on the right by $R(\lambda; T)$ we find that

$$S = (\lambda I - T) S R(\lambda; T).$$

Multiplying this last expression on the left by $R(\lambda; T)$ we obtain

$$SR(\lambda; T) = R(\lambda; T)S. \quad \square$$

Theorem 5.8: Let X be a complex Banach space and $T \in \mathcal{L}(X;X)$ be given. Assume that $X \neq \{0\}$. Then $\sigma(T) \neq \emptyset$.

Proof: Suppose $\sigma(T) = \emptyset$. Then $\rho(T) = \mathbb{C}$. Put

$$D = \{\lambda \in \mathbb{C} : |\lambda| \leq 2\|T\|\},$$

and observe that D is nonempty and compact. Let

$$M = \max\{\|R(\lambda;T)\| : \lambda \in D\} < \infty. \tag{13}$$

For all $\lambda \in \mathbb{C} \backslash D$ we have

$$R(\lambda;T) = \frac{1}{\lambda}\left(I - \frac{T}{\lambda}\right) = \frac{1}{\lambda}\sum_{k=0}^{\infty}\left(\frac{T}{\lambda}\right)^k \tag{14}$$

and

$$\left\|\frac{T}{\lambda}\right\| \leq \frac{1}{2} \tag{15}$$

Combining (13) and (15), we find that

$$\|R(\lambda;T)\| \leq \max\{M, \|T\|\} \text{ for all} \lambda \in \mathbb{C} \tag{16}$$

and

$$\|R(\lambda;T)\| \to 0 \text{ as } |\lambda| \to \infty. \tag{17}$$

Now let $x \in X$ and $x^* \in X^*$ be given and define $f : \mathbb{C} \to \mathbb{C}$ by

$$f(\lambda) = x^* R(\lambda;T)x \text{ for all } \lambda \in \mathbb{C}.$$

Then f is an entire function and it is bounded by virtue of (16). Liouville's Theorem implies that f is constant. We see from (17) that

$$f(\lambda) \to 0 \text{ as } |\lambda| \to \infty,$$

and consequently

$$f(\lambda) = 0 \text{ for all} \lambda \in \mathbb{C}.$$

In other words, we have

$$x^*(R(\lambda;T)x) = 0 \text{ for all } x \in X, x^* \in X^*, \lambda \in \mathbb{C}.$$

This is impossible, because for $R(\lambda;T)$ is invertible, so we may choose $x \in X$ such that $R(\lambda;T)x \neq 0$ and then (by the Hahn Banach Theorem), we may choose $x^* \in X^*$ such that $x^*(R(\lambda;T)x \neq 0$. $\square$

Spectral Mapping Theorem for Polynomials

Lemma 5.9: Let X be a Banach space, $T \in \mathcal{L}(X;X)$, $\lambda \in \sigma(T)$ and $n \in \mathbb{N}$ be given. Then $\lambda^n \in \sigma(T^n)$.

Proof: Put
$$B = T^{n-1} + \lambda T^{n-2} + \lambda^2 T^{n-3} + \cdots + \lambda^{n-1} I,$$
and observe that
$$T^n - \lambda^n I = (T - \lambda I)B = B(T - \lambda I). \tag{18}$$
Suppose that $\lambda^n \in \rho(T^n)$. Then $T^n - \lambda I$ is bijective. It follows from (18) that $(T - \lambda I)B$ is surjective and consequently $T - \lambda I$ is surjective. It also follows from (18) that $B(T - \lambda I)$ is injective and consequently $T - \lambda I$ is injective. This contradicts the fact that $\lambda \in \sigma(T)$. $\square$

Lemma 5.10: Let X be a complex Banach space, $T \in \mathcal{L}(X;X)$, $n \in \mathbb{N}$, and $\mu \in \sigma(T^n)$ be given. Then there exists $\lambda \in \sigma(T)$ such that $\mu = \lambda^n$.

Proof: We may choose $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{C}$ such that
$$z^n - \mu = \prod_{j=1}^{n} (z - \alpha_j) \ \text{ for all } z \in \mathbb{C}.$$
Then we also have
$$T^n - \mu I = \prod_{j=1}^{n} (T - \alpha_j I).$$
Since $T^n - \mu I$ fails to be bijective, we may choose $k \in \{1, 2, \cdots, n\}$ such that $T - \alpha_k I$ fails to be bijective. It follows that $\alpha_k \in \sigma(T)$ and $\alpha_k^n = \mu$. $\square$

Theorem 5.11 (Spectral Mapping Theorem for Polynomials): Let X be a complex Banach space, $T \in \mathcal{L}(X;X)$ and $p : \mathbb{C} \to \mathbb{C}$ be a polynomial. Let $\mu \in \mathbb{C}$ be given. Then $\mu \in \sigma(p(T)$ if and only if there exists $\lambda \in \sigma(T)$ such $\mu = p(\lambda)$.

Spectral Radius

Definition 5.12: Let X be a Banach space and $T \in \mathcal{L}(X;X)$ be given. Assume that $\sigma(T) \neq \emptyset$. The *spectral radius* of T, denoted by $r_\sigma(T)$, is defined by
$$r_\sigma(T) = \max\{|\lambda| : \lambda \in \sigma(T)\}.$$

Observe that
$$0 \leq r_\sigma(T) \leq \|T\|. \tag{19}$$

Theorem 5.13: Let X be a complex Banach space (with $X \neq \{0\}$) and $T \in \mathcal{L}(X;X)$ be given. Then
$$r_\sigma(T) = \lim_{n\to\infty} \sqrt[n]{\|T^n\|},$$

and the limit above exists.

Proof: Let $n \in \mathbb{N}$ be given. By Lemmas 5.9 and 5.10, and (19), we have

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

It follows that

$$r_\sigma(T) \leq \liminf_{n\to\infty} \sqrt[n]{\|T^n\|}. \tag{20}$$

It remains to show that

$$\limsup_{n\to\infty} \sqrt[n]{\|T^n\|} \leq r_\sigma(T).$$

For $|\lambda|$ large, we have

$$R(\lambda;T) = \frac{1}{\lambda}\sum_{n=0}^{\infty}\left(\frac{T}{\lambda}\right)^n. \tag{21}$$

For z near zero, let us consider the power series

$$F(z) = z\sum_{n=0}^{\infty} z^n T^n, \tag{22}$$

(which is obtained from the series in (21) by putting $z = \lambda^{-1}$.) This power series has radius of convergence $r \in [0,\infty]$ satisfying

$$\frac{1}{r} = \limsup_{n\to\infty} \sqrt[n]{\|T^n\|}.$$

To understand the relationship between r and r_σ, we put

$$\Omega = \{z \in \mathbb{C}\backslash\{0\} : z^{-1} \in \rho(T)\} \cup \{0\}.$$

Consider the function $G : \Omega \to \mathcal{L}(X;X)$ defined by

$$G(z) = \begin{cases} R(z^{-1};T) & \text{for } z \neq 0 \\ \\ 0 & \text{for } z = 0. \end{cases}$$

Observe that G is analytic on Ω. (For z near zero, analyticity of G follows from the series representation for F. For z away from 0, analyticity of G follows from analyticity of the resolvent.) Moreover, $F(z) = G(z)$ for all z for which the series for F converges. The radius of convergence r of the series for F will therefore be the supremum of all radii ρ such the disc of radius ρ centered at 0 is included in Ω. In other words, we have

$$r_\sigma = \frac{1}{r}.$$

We conclude that

$$\limsup_{n\to\infty} \sqrt[n]{\|T^n\|} \leq r_\sigma(T). \quad \square$$