21-740 **Functional Analysis II** **Fall 2013**

Lecture Notes for Week 1 (First Draft)

This course will be a continuation of 21-640 (*Introduction to Functional Analysis*) from last semester. The main topics that I plan to cover are

I. More on Hilbert Spaces (including Spectral Theory)

II. Spectral Theory in Banach Spaces

III. Unbounded Linear Operators

IV. Semigroups of Linear Operators

V. Fourier Transforms & Applications

VI. Nonlinear Operators

In order to make the course accessible to students who took 21-640 from another instructor, I will begin by reviewing a few notational conventions and basic results concerning Hilbert spaces.

Inner Products and Hilbert Spaces

We shall use $\mathbb{K}$ to denote a field that is either $\mathbb{R}$ or $\mathbb{C}$. Let X be a linear space over $\mathbb{K}$. (Unless stated otherwise, all linear spaces are over $\mathbb{K}$.) By an *inner product* on X we mean a mapping $(\cdot,\cdot) : X \times X \to \mathbb{K}$ satisfying the following 5 conditions:

(i) $\forall\ x, y, z \in X,\ (x + y, z) = (x, z) + (y, z)$,

(ii) $\forall x, y \in X,\ \alpha \in \mathbb{K},\ (\alpha x, y) = \alpha(x, y)$,

(iii) $\forall\ x, y \in X,\ (x, y) = \overline{(x, y)}$,

(iv) $\forall\ x \in X,\ (x, x) \geq 0$,

(v) $\forall x \in X,\ (x, x) = 0 \Rightarrow x = 0$.

A linear space equipped with an inner product is called an *inner product space*. If $(X, (\cdot,\cdot))$ is an inner product space, then the function $\|\cdot\| : X \to \mathbb{R}$ defined by

$$\|x\| = \sqrt{(x, x)} \ \text{ for all } x \in X$$

is a norm on X; this norm obeys the *Cauchy-Schwartz inequality*

$$|(x, y)| \leq \|x\| \|y\| \text{ for all } x, y \in X. \tag{1}$$

When we speak of the norm on an inner product space, we always mean the norm associated with the inner product via (1).

One of the most important features of Hilbert spaces is the notion of orthogonal projections. We record two very important results that were proved in 21-640.

Theorem 1.1 (Projection onto Closed Convex Sets): Let X be a Hilbert space and assume that $K \subset X$ is nonempty, closed, and convex. Then there is exactly one $y_0 \in K$ such that

$$\|x - y_0\| = \inf\{\|x - y_0\| : y \in K\}. \tag{2}$$

Remark 1.2:

(a) If X is a reflexive Banach space and K is a nonempty, closed, convex subset of X, then for every $x \in X$ there is at least one $y_o \in K$ satisfying (2).

(b) If X is a uniformly convex Banach space and K is a nonempty, closed, convex subset of X, then for every $x \in X$ there is exactly one $y_o \in K$ satisfying (2).

(c) It is not difficult to give an example of a nonempty, closed, convex subset K of a nonreflexive Banach space and a point $x \in X$ such that there is no point $y_0 \in K$ satisfying (2).

Recall that if S is a subset of a Hilbert space X, then the *orthogonal complement* $S^\perp$ of S is defined by

$$S^\perp = \{y \in X : (x, y) = 0 \text{ for all } x \in S\}.$$

Moreover $S^\perp$ is always a closed subspace of X.

Theorem 1.3 (Projection Theorem): Let X be a Hilbert space and M be a closed subspace of X. Let $x \in X$ be given. Then there is exactly one pair $(y, z) \in M \times M^\perp$ such that $x = y + z$.

Let M be a closed subspace of a Hilbert space X. For each $x \in X$ let $P_M x$ denote the unique element of M such that

$$\|x - P_m x\| \leq \|x - y\| \text{ for all } y \in M.$$

Then $P_M : X \to X$ is linear and

$$x - P_M x \in M^\perp \text{ for all } x \in X.$$

Notice that
$$x = P_M x + (x - P_M x) \text{ for all} x \in X$$
which implies
$$\|x\|^2 = \|P_M x\|^2 + \|x - P_M x\|^2 \text{ for all } x \in X$$
since $P_M x$ is orthogonal to $x - P_M x$. It follows that
$$\|P_M x\| \leq \|x\| \text{ for all } x \in X,$$
and consequently P_M is continuous with $\|P_M\| \leq 1$. Observe that
$$P_{M^\perp} = I - P_M.$$

We refer to P_M as the *orthogonal projection* onto M. We refer to a linear operator that is the orthogonal projection onto some closed subspace as an *orthogonal projection.*

Recall that a list $(e_j | j \in J)$ is said to be *orthonormal* provided that
$$\forall\ i, j \in J \text{ we have } (e_j, e_j) = 1 \text{ and } (e_i, e_j) = 0 \text{ whenever } i \neq j.$$
(Here J can be any index set.) An orthonormal list $(e_j | j \in J)$ is said to be an *orthonormal basis* provided that it is *maximal* in the sense that
$$\forall x \in X\ (x \perp e_j \text{ for all } j \in J) \Rightarrow x = 0.$$
If $(e_j | j \in J)$ is an orthonormal basis then for every $x \in X$ we have
$$x = \sum_{j \in J} (x, e_j) e_j.$$
(Convergence of the above sum to x means that for every $\epsilon > 0$ there is a finite set $F \subset J$ such that for every finite set G with $F \subset G \subset J$ we have
$$\|x - \sum_{j \in G} (x, e_j) e_j\| < \epsilon.)$$

A very convenient feature of Hilbert spaces is that they can be identified with their own dual spaces via the inner product.

Theorem 1.4 (Riesz Representation Theorem): Let X be a Hilbert space and let $x^* \in X^*$ be given. Then there exists exactly one $y \in X$ such that
$$x^*(x) = (x, y) \text{ for all } x \in X;$$
moreover $\|x^*\|_* = \|y\|$.

It is convenient to introduce the *Riesz Operator* $\hat{R} : X \to X^*$ defined by

$$(\hat{R}(y))(x) = (x, y) \ \text{ for all } x, y \in X.$$

Recall that $\hat{R}$ is conjugate linear and isometric.

Let X be a Hilbert space and $A \in \mathcal{L}(X; X)$ be given. Since X is also a Banach space, A has an adjoint $A_B^* \in \mathcal{L}(X^*; X^*)$ in the sense of Banach spaces defined by

$$(A_B^* x^*)(x) = x^*(Ax) \ \text{ for all } x^* \in X^*, \ x \in X.$$

The Hilbert space adjoint $A_H^* \in \mathcal{L}(X; X)$ of A is defined by

$$(Ax, y) = (x, A_H^* y) \ \text{ for all } x, y \in X.$$

Notice that the Banach and Hilbert adjoints are related by the formula

$$A_H^* = \hat{R}^{-1} A_B^* \hat{R}.$$

When X is a Hilbert space, the Hilbert adjoint is generally more convenient than the Banach Adjoint. Unless stated otherwise, when X is a Hilbert space and $A \in \mathcal{L}(X; X)$ we use A^* to denote the Hilbert adjoint.

Recall that if X is a Hilbert space, then

$$(A^*)^* = A \ \text{ and } \ \mathcal{N}(A) = \mathcal{R}(A^*) \ \text{ for all } A \in \mathcal{L}(X; X).$$

Self-Adjoint and Normal Operators

Definition 1.5: Let X be a Hilbert space and let $A \in \mathcal{L}(X; X)$ be given. We say that A is

(a) *self-adjoint* if $A = A^*$.

(b) *normal* if $AA^* = A^*A$.

Clearly, every self-adjoint operator is normal, but not conversely.

Proposition 1.6: Assume that X is a complex Hilbert space and let $A \in \mathcal{L}(X; X)$ be given. Then A is self-adjoint if and only if $(Ax, x) \in \mathbb{R}$ for all $x \in X$.

Proof: Assume first that $A^* = A$ and let $x \in X$ be given. Then we have

$$(Ax, x) = (A^*x, x) = (x, Ax) = \overline{(Ax, x)},$$

so $(Ax, x) \in \mathbb{R}$.

Conversely, assume that $(Az, z) \in \mathbb{R}$ for all $z \in X$. Let $x, y \in X$ and $\alpha \in \mathbb{C}$ be given. Then

$$(A(x + \alpha y), x + \alpha y) \in \mathbb{R}.$$

Expanding the above expression and using the fact that $(Ax, x), (A(\alpha y), (\alpha y)) \in \mathbb{R}$ we find that

$$\overline{\alpha}(Ax, y) + \alpha(Ay, x) \in \mathbb{R}. \tag{3}$$

Putting $\alpha = 1$ and $\alpha = i$ in (3) we find that

$$(Ax, y) + (Ay, x) = (y, Ax) + (x, Ay), \tag{4}$$

$$-(Ax, y) + (Ay, x) = (y, Ax) - (x, Ay). \tag{5}$$

Adding (4) and (**??**) we find that

$$(AY, x) = (y, Ax) \ \text{ for all} x, y \in X,$$

which implies that $A = A^*$. $\square$

Proposition 1.7: Let X be a Hilbert space and let $A \in \mathcal{L}(X; X)$ be given. Assume that A is self-adjoint and that $X \neq \{0\}$. Then

$$\|A\| = \sup\{|(Ax, x)| : x \in X, \|x\| = 1\}.$$

Proof: Put

$$M = \sup\{|(Ax, x)| : x \in X, \|x\| = 1\}.$$

We shall show $\|A\| \leq M$ and $M \leq \|A\|$. (Notice that the definition of M implies $|(Az, z)| \leq M\|z\|^2$ for all $z \in X$.)

Let $x \in X$ with $\|x\| = 1$ be given. Then we have

$$|(Ax, x)| \leq \|Ax\|\|x\| \leq \|A\|,$$

and consequently

$$M \leq \|A\|.$$

(Notice that self-adjointness of A is not needed to conclude that $M \leq \|A\|$.

To establish the reverse inequality, let $x, y \in X$ with $\|x\| = \|y\| = 1$ be given. Then we have

$$\begin{aligned} (A(x + y), (x + y)) &= (Ax, x) + (Ax, y) + (Ay, x) + (Ay, y) \\ &= (Ax, x) + (Ax, y) + (y, Ax) + (Ay, y) \\ &= (Ax, x) + 2\mathrm{Re}(Ax, y) + (Ay, y). \end{aligned}$$

Replacing y with $-y$ in the above, we find that

$$(A(x - y), (x - y)) = (Ax, x) - 2\mathrm{Re}(Ax, y) + (Ay, y).$$

Subtracting the expressions for $(A(x+y),(x+y))$ and $(A(x-y),(x-y))$ we find that

$$4\mathrm{Re}(Ax,y) = (A(x+y),(x+y)) - (A(x-y),(x-y)).$$

Taking the absolute value and using the fact that $|(Az,z)| \le M\|z\|^2$ we find that

$$4|\mathrm{Re}(Ax,y)| \le M(\|x+y\|^2 + \|x-y\|^2) = 2M(\|x\|^2 + \|y\|^2).$$

Since $\|x\| = \|y\| = 1$ we conclude that

$$|\mathrm{Re}(Ax,y)| \le M. \tag{6}$$

Choose $\theta \in [0, 2\pi)$ such that

$$(Ax,y) = e^{i\theta}\|(Ax,y)|$$

and put $z = e^{-i\theta}x$. Observe that $\|z\| = 1$. Then we have

$$|(Ax,y)| = e^{-i\theta}(Ax,y) = (Az,y).$$

It follows that (Az,y) is real and nonnegative. Since $\|z\| = \|y\| = 1$, it follows from

$$|(Ax,y)| = |\mathrm{Re}(Az,y)| \le M.$$

We conclude that

$$\|Ax\| = \sup\{(Ax,y) : y \in X, \|y\| = 1\} \le M.$$

Taking the supremum over $x \in X$ with $\|x\| = 1$ we find that

$$\|A\| \le M,$$

and we are done. $\square$

Corollary 1.8: Let X be a Hilbert space and $A \in \mathcal{L}(X;X)$ be given and assume that A is self-adjoint. Assume further that $(Ax,x) = 0$ for all $x \in X$. Then $A = 0$.

Remark 1.9: If $\mathbb{K} = \mathbb{C}$ then self-adjointness of A is not needed in the above corollary.

Proposition 1.10: Let X be a Hilbert space and $A \in \mathcal{L}(X;X)$ be given. Then A is normal if and only if $\|Ax\| = \|A^*x\|$ for all $x \in X$.

Proof: Let $x \in X$ be given. Then we have

$$\begin{aligned} \|Ax\|^2 - \|A^*x\|^2 &= (Ax,Ax) - (A^*x,A^*x) \\ &= (A^*Ax,x) - (AA^*x,x) \\ &= ((A^*A - AA^*)x,x). \end{aligned}$$

Since $A^*A - AA^*$ is self-adjoint, the result follows easily from Corollary . $\square$

Corollary 1.11: Let X be a Hilbert space and assume that $A \in \mathcal{L}(X;X)$ is normal. Then $\mathcal{N}(A) = \mathcal{N}(A^*)$.

Isometric and Unitary Operators

Definition 1.12: Let X be a Hilbert space and $A \in \mathcal{L}(X;X)$ be given. We say that A is

(a) an *isometry* provided that $\|Ax\| = \|x\|$ for all $x \in X$.

(b) *unitary* if it is a surjective isometry.

Notice that every isometry is injective and consequently every unitary operator is bijective.

Example 1.13: Let $X = l^2$ and define the right and left shift operators $R, L \in \mathcal{L}(L^2; l^2)$ by

$$Rx = (0, x_1, x_2, x_3, \cdots) \ \text{ for all } x \in L^2,$$
$$Lx = (x_2, x_3, x_4, \cdots) \ \text{ for all } x \in l^2.$$

Recall that $R^* = L$ and $L^* = R$. Notice that R is an isometry, but fails to be surjective and L is surjective, but fails to be an isometry. Notice also that $LR = I$, but

$$RLx = (0, x_2, x_3, x_4, \cdots) \ \text{ for all } x \in X,$$

so neither R nor L is normal.

Proposition 1.14: Let X be a Hilbert space and $A \in \mathcal{L}(X;Y)$ be given. Then A is an isometry if and only if

$$(Ax, Ay) = (x, y) \ \text{ for all } x, y \in X. \tag{7}$$

Proof: If (7) holds, then putting $y = x$ gives

$$\|Ax\|^2 = \|x\|^2 \ \text{ for all } x \in X$$

and A is an isometry.

Conversely, assume that A is an isometry and let $x, y \in X$ and $\alpha \in \mathbb{K}$ be given. Then we have

$$\begin{aligned} \|A(x + \alpha y)\|^2 &= (Ax + \alpha Ay, Ax + \alpha Ay) \\ &= \|Ax\|^2 + 2\mathrm{Re}[\alpha(Ay, Ax)] + |\alpha|^2\|Ay\|^2 \qquad (8) \\ &= \|x\|^2 + 2\mathrm{Re}[\alpha(Ay, Ax)] + |\alpha|^2\|y\|^2. \end{aligned}$$

On the other hand, we also have

$$\begin{aligned} \|A(x+\alpha y)\|^2 &= \|x+\alpha y\|^2 \\ &= \|x\|^2 + 2\mathrm{Re}[\alpha(y,x)] + |\alpha|^2\|y\|^2. \end{aligned} \tag{9}$$

Combining (8) and (9) we obtain

$$\mathrm{Re}[\alpha(Ay,Ax) = \mathrm{Re}[\alpha(y,x)]. \tag{10}$$

If $\mathbb{K}=\mathbb{R}$ we are done. If $\mathbb{K}=\mathbb{C}$ then we put $\alpha=1$ and $\alpha=i$ in (10). $\square$

Proposition 1.15 Let X be a Hilbert space and $A \in \mathcal{L}(X;X)$ be given. Then A is an isometry if and only if $A^*A=I$.

Proof: Observe that

$$\begin{aligned} A \text{ is isometric } &\Leftrightarrow (Ax,Ay)=(x,y) \text{ for all } x,y\in X \\ &\Leftrightarrow (A^*Ax,y)=(x,y) \text{ for all } x,y\in X \\ &\Leftrightarrow A^*Ax=x \text{ for all } x\in X. \end{aligned}$$

Remark 1.16: The order of the product A^*A in Proposition is important. Observe that R is isometric, but $RR^*\neq I$ and $LL^*=I$, but L is not isometric, where R and L are as in Example 1.13.

Proposition 1.17: Let X be a Hilbert space and $A\in\mathcal{L}(X;X)$ be given. Assume that A is an isometry. Then A is normal if and only if A is surjective.

Proof: Assume first that A is normal. Then we have

$$I = A^*A = AA^*,$$

which implies that A is surjective.

Assume now that A is surjective. Then A is bijective and A^{-1} is an isometry. Using , we see that

$$(A^{-1})^*A = I.$$

Since the adjoint of A^{-1} is the inverse of A^* we have

$$(A^*)^{-1}A^{-1} = I,$$

and this implies that

$$(AA^*)^{-1} = I$$

and consequently

$$AA^* = I.$$

Since A is isometric, we know that $A^*A = I$ and consequently $AA^* = A^*A$. $\square$

Idempotent Operators and Orthogonal Projections

For the definition of idempotent operator (and the subsequent remark), we let X be a normed linear space.

Definition 1.18: Let X be a normed linear space and $E \in \mathcal{L}(X;X)$ be given. We say that E is *idempotent* provided that $E^2 = E$.

Observe that for any linear operator E we have

$$(I-E)^2 - (I-E) = E^2 - E. labelEq:3.idem \tag{11}$$

Observe also that if E is idempotent the $y \in \mathcal{R}(E)$ if and only if $y + Ey$.

Remark 1.19: Let X be a normed linear space and $E \in \mathcal{L}(X;X)$ be given. Then

(a) E is idempotent if and only if $I - E$ is idempotent.

(b) If E is idempotent then $\mathcal{R}(E) = \mathcal{N}(I-E)$ and $\mathcal{R}(I-E) = \mathcal{N}(E)$. In particular, every idempotent operator has closed range.

If X is a Hilbert space, then every orthogonal projection is idempotent, but not conversely. It is instructive to look at a simple example in $\mathbb{R}^2$.

Example 1.20: Let $X = \mathbb{R}^2$ and put

$$E = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

It is straightforward to check that E is idempotent, E is not normal, and $\|E\| = \sqrt{2}$.

Lecture Notes for Week 2 (First Draft)

Idempotent Operators and Projections (Continued)

Proposition 2.1: Let X be a Hilbert space, $E \in \mathcal{L}(X;X)$ be given and assume that E is idempotent. Let $M = \mathcal{R}(E)$. Then $E = P_M$ if and only if

$$\mathcal{N}(E) = \mathcal{R}(E)^{\perp}.$$

Proof: Assume first that $E = P_M$. Then $I - E = P_{M^{\perp}}$, so $\mathcal{R}(I-E) = M^{\perp}$. By Remark 1.19, we know that $\mathcal{N} = \mathcal{R}(I-E) = M^{\perp}$. Since M is a closed subspace, we know that $(M^{\perp})^{\perp} = M$, and consequently $\mathcal{N}(E) = \mathcal{R}(E)^{\perp}$.

Assume now that $\mathcal{N}(E) = \mathcal{R}(E)^{\perp} = M^{\perp}$. Then, by Remark 1.19, $\mathcal{R}(I-E) = M^{\perp}$. Let $x \in X$ be given and observe that

$$x = Ex + (I-E)x.$$

Since $Ex \in M$ and $(I-E)x \in M^{\perp}$, it follows from the Projection Theorem that $Ex = P_M x$. $\square$

Remark 2.2: Let X be a normed linear space and assume that $E \in \mathcal{L}(X;X)$ is idempotent. Then

$$\|E\| = \|E^2\| \leq \|E\|^2,$$

and consequently either $E = 0$ or $\|E\| \geq 1$. Notice that in Example 1.20, the idempotent operator E has norm equal to $\sqrt{2}$.

Proposition 2.3: Let X be a Hilbert space and let $E \in \mathcal{L}(X;X)$ be given. Assume that E is idempotent and $E \neq 0$. Let $M = \mathcal{R}(E)$. Then $E = P_M$ if and only if $\|E\| = 1$.

Proof: Assume first that $E = P_M$. Let $x \in X$ be given. Then

$$x = P_M x + (I - P_M)x,$$

and $(P_M x, (I-P_M)x) = 0$. It follows from the Pythagorean Theorem that

$$\|x\|^2 = \|P_M x\|^2 + \|(I-P_M)x\|^2 \geq \|P_M x\|^2 = \|Ex\|^2.$$

Consequently, we have $\|E\| \leq 1$. By Remark 2.2, we also have $\|E\| \geq 1$.

Assume now that $\|E\| = 1$. Since $\mathcal{R}(E)$ is closed, it suffices to show that $\mathcal{R}(E) = \mathcal{N}(E)^{\perp}$. Let $x \in \mathcal{N}(E)^{\perp}$ be given. Since $\mathcal{R}(I-E) = \mathcal{N}(E)$ we have

$$\begin{aligned} 0 &= (x - Ex, x) = \|x\|^2 - (Ex, x) \\ &\geq \|x\|^2 - \|Ex\|\|x\| = \|x\|(\|x\| - \|Ex\|). \end{aligned} \tag{1}$$

If $x = 0$ then certainly $x \in \mathcal{R}(E)$, so we may assume that $x \neq 0$. It follows from (1) that $\|Ex\| \geq \|x\|$. Since $\|E\| = 1$ we know that $\|Ex\| \leq \|x\|$ and consequently $\|Ex\| = \|x\|$. Using (1) again, we see that

$$\|Ex\|^2 = \|x\|^2 = (Ex, x),$$

and consequently

$$\|x - Ex\|^2 = \|x\|^2 + \|Ex\|^2 - 2(Ex, x) = 0.$$

It follows that $x = Ex$ and $x \in \mathcal{R}(E)$.

Now let $y \in \mathcal{R}(E)$ be given. Let us write

$$y = x + z$$

with $x \in \mathcal{N}(E)$ and $z \in \mathcal{N}(E)^{\perp}$. Since $\mathcal{N}(E)^{\perp} \subset \mathcal{R}(E)$ we know that $Ez = z$. Consequently, we have

$$y = Ey = E(x + z) = Ez = z,$$

and $y \in \mathcal{N}(E)^{\perp}$. $\square$

Theorem 2.4: Let X be a Hilbert space and $E \in \mathcal{L}(X; X)$ be given. Assume that E is idempotent and that $E \neq 0$. Put $M = \mathcal{R}(E)$. The following four statements are equivalent:

(i) $E = P_M$.

(ii) $\|E\| = 1$.

(iii) E is self-adjoint.

(iv) E is normal.

Proof: We have already shown that (i) holds if and only if (ii) holds. Moreover the implication (iii) $\Rightarrow$ (iv) is immediate. It suffices to prove (i) $\Rightarrow$ (iii) and (iv) $\Rightarrow$ (i).

Assume that (i) holds and let $x, y \in X$ be given. We write

$$x = x_1 + x_2, \quad y = y_1 + y_2, \quad x_1, y_1 \in M, \ x_2, y_2 \in M^{\perp}.$$

Then we have

$$(Ex, y) = (Ex_1 + Ex_2, y_1 + y_2) = (x_1, y_1 + y_2) = (x_1, y_1),$$

and also

$$(x, Ey) = (x_1 + x_2, Ey_1 + Ey_2) = (x_1 + x_2, y_1) = (x_1, y_1).$$

We conclude that

$$(Ex, y) = (x, Ey) \text{ for all } x, y \in X$$

and E is self-adjoint.

Assume now that (iv) holds. Then

$$\|Ex\| = \|E^*x\| \text{ for all } x \in X$$

and $\mathcal{N}(E) = \mathcal{N}(E^*)$. Since $\mathcal{N}(A^*) = \mathcal{R}(A)^\perp$ for every $A \in \mathcal{L}(X; X)$ we conclude that $\mathcal{N}(E) = \mathcal{R}(E)^\perp$ and $E = P_M$ by virtue of . $\square$

Proposition 2.5: Let X be a Hilbert space and $E \in \mathcal{L}(X; X)$ be given. Assume that E is idempotent and put $M = \mathcal{R}(E)$. Then $E = P_M$ if and only if

$$(Ex, x) \geq 0 \text{ for all } x \in X. \tag{2}$$

(The condition (2) means that for every $x \in X$, the quantity (Ex, x) is real and nonnegative.)

Proof: Assume first that $E = P_M$. Let $x \in X$ be given and write

$$x = x_1 + x_2, \quad x_1 \in M, \ x_2 \in M^\perp.$$

Then we have

$$(Ex, x) = (Ex_1 + Ex_2, x_1 + x_2) = (x_1, x_1) \geq 0.$$

Conversely, assume that (2) holds. If $\mathbb{K} = \mathbb{C}$ then E is self-adjoint by Proposition 1.6 and using Thoerem 2.4, we conclude that $E = P_M$.

Assume that $\mathbb{K} = \mathbb{R}$ and let $y \in \mathcal{R}(E)$ and $z \in \mathcal{R}(I - E)$ be given. Then, since $\mathcal{R}(I - E) = \mathcal{N}(E)$, we have

$$0 \leq (Ey + Ez, y + z) = (y, y + z) = \|y\|^2 + (y, z).$$

We conclude that

$$\|y\|^2 + (y, z) \geq 0 \text{ for all } y \in \mathcal{R}(E), \ z \in \mathcal{R}(I - E). \tag{3}$$

I claim that (3) implies that $(y, z) = 0$ for all $y \in \mathcal{R}(E)$, $z \in \mathcal{R}(I - E)$. To see why the claim is true, let $y \in \mathcal{R}(E)$ and $z \in \mathcal{R}(I - E)$ be given. Then we have

$$\|y\|^2 + t(y, z) \geq 0 \text{ for all } t \in \mathbb{R}. \tag{4}$$

If $(y, z) \neq 0$, then we can choose t to have the opposite sign of (y, z) and be large in magnitude. This will contradict (4). $\square$

Lecture Notes for Week 3 (First Draft)

Invariant and Reducing Subspaces

Let X be a Hilbert space, M be a closed subspace of X, and $A \in \mathcal{L}(X;X)$. The Projection Theorem says that

$$X = M \oplus M^{\perp}.$$

In order to understand the behavior of the operator A it is useful break x and Ax down into their components in M and $M^{\perp}$:

$$\begin{pmatrix} P_M(Ax) \\ \\ P_{M^{\perp}}(Ax) \end{pmatrix} = \begin{pmatrix} B & C \\ \\ D & F \end{pmatrix} \begin{pmatrix} P_M x \\ \\ P_{M^{\perp}} x \end{pmatrix}, \tag{1}$$

where $B \in \mathcal{L}(M;M)$,$C \in \mathcal{L}(M^{\perp};M)$, $D \in \mathcal{L}(M;M^{\perp})$, and $F \in \mathcal{L}(M^{\perp};M^{\perp})$.

Definition 3.1 Let X be a Hilbert space, M be a closed subspace of X, and $A \in \mathcal{L}(X;X)$ be given. We say that

(a) M is invariant under A provided $A[M] \subset M$.

(b) M reduces A provided that M and $M^{\perp}$ both are invariant under A.

Notice that with regard to the decomposition (1) we have

- M is invariant under $A \Leftrightarrow D = 0$,
- $M^{\perp}$ is invariant under $A \Leftrightarrow C = 0$,
- M reduces $A \Leftrightarrow (C = 0$ and $D = 0)$.

Proposition 3.2: Let X be a Hilbert space, M be a closed subspace of X, and $A \in \mathcal{L}(X;X)$ be given. Then M is invariant under A if and only if $M^{\perp}$ is invariant under A^*.

Proof: Let $x \in M,\ y \in M^{\perp}$ be given. Since

$$(x, A^*y) = (Ax, y),$$

we have

$$(x, A^*y) = 0 \ \Leftrightarrow\ (Ax, y) = 0,$$

and the result follows. $\square$

Proposition 3.3: Let X be a Hilbert space, M be a closed subspace of X, and $A \in \mathcal{L}(X;X)$ be given. Then M is invariant under A if and only if $P_M A P_M = A P_M$.

Proof: Assume first that M is invariant under A and let $x \in X$ be given. Then $P_M x \in M$. Since M is invariant under A we have $A P_M x \in M$ so $P_M A P_M x = A P_M x$. Assume now that $P_M A P_M = A P_M$. Let $x \in M$ be given. Then

$$Ax = AP_M x = P_M A P_M x \in M,$$

so that M is invariant under A. $\square$

Theorem 3.4: Let X be a Hilbert space, M be a closed subspace of X, and $A \in \mathcal{L}(X;X)$ be given. The following three statements are equivalent.

(i) M reduces A.

(ii) $P_M A = A P_M$.

(iii) M is invariant under both A and A^*

Proof: It follows from Proposition 3.2 that (i) holds if and (iii) holds, so we only need to prove (i) $\Leftrightarrow$ (ii). Assume that (i) holds and let $x \in X$ be given. Then we have

$$x = P_M x + (I - P_M)x,$$

and consequently

$$Ax = AP_M x + A(I - P_M)x. \tag{2}$$

Since M and $M^\perp$ are both invariant under A, we have

$$AP_M x \in M, \quad A(I - P_M)x \in M^\perp.$$

On the other hand, we also have

$$Ax = P_M(Ax) + (I - P_M)(Ax). \tag{3}$$

Since there is exactly one decomposition of Ax into a sum of an element of M and an element of $M^\perp$ we conclude that

$$AP_M x = P_M Ax \ \text{ for all } x \in X,$$

and consequently (ii) holds.

Assume now that (ii) holds. Then we have

$$P_M A P_M = A P_M^2 = A P_M.$$

Since $(P_M)^* = P_M$ and $(AP_M)^* = (P_M)^* A^*$ and $(P_M A)^* = A^*(P_M)^*$ it follows from (ii) that

$$P_M A^* = A^* P_M.$$

Consequently we have

$$P_M A^* P_M = A^* P_M.$$

It follows from Proposition 3.3 that M is invariant under A and A^* Using Proposition 3.2, we conclude that M reduces A. $\square$

Spectral Theory

Definition 3.5: Let X be a Hilbert space over $\mathbb{K}$ and $A \in \mathcal{L}(X; X)$ be given.

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

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

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

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

The definition of resolvent set given above is appropriate only for bounded linear operators from a complete space to itself. (It is appropriate for Banach spaces as well as Hilbert spaces.) We will give a more general definition of resolvent set later on in the course that applies to incomplete normed linear spaces and to linear operators that need not be continuous.

Observe that by the Bounded Inverse Theorem, if $\lambda \in \rho(A)$ then $(\lambda I - A)^{-1}$ is bounded.

Definition 3.6: Let X be a Hilbert space over $\mathbb{K}$ and $A \in \mathcal{L}(X; X)$ be given.

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

(b) If λ is an eigenvalue for A, then the nonzero elements of $\mathcal{N}(\lambda I - A)$ are called *eigenvectors* associated with λ.

(c) The set of all eigenvalues of A is called the *point spectrum* of A and is denoted by $\sigma_p(A)$.

Notice that

$$\sigma_p(A) \subset \sigma(A).$$

If X is finite dimensional then $\sigma_p(A) = \sigma(A)$. However, it is easy to give examples in infinite-dimensional Hilbert spaces where the spectrum of a bounded linear operator contains elements that are not eigenvalues.

Definition 3.7: Let X be a Hilbert space over $\mathbb{K}$ and $A \in \mathcal{L}(X;X)$ be given. A number $\lambda \in \mathbb{K}$ is called a generalized eigenvalue for A provided that

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

The same definitions of eigenvalue, eigenvector, point spectrum, and generalized eigenvector apply to bounded linear operators from a Banach space X to itself.

Remark 3.8: Notice that $\lambda \in \mathbb{K}$ is a generalized eigenvalue for $A \in \mathcal{L}(X;X)$ if and only if there is a sequence $\{x_n\}_{n=1}^{\infty}$ such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $(\lambda I - A)x_n \to 0$ as $n \to \infty$.

Every generalized eigenvalue for A belongs to $\sigma(A)$. If X is finite dimensional then every generalized eigenvalue is an eigenvalue. However, it is easy to give examples in infinite dimensions of a generalized eigenvalue that is not an eigenvalue.

Example 3.9: Let $X = l^2$ and define $A \in \mathcal{L}(L^2; l^2)$ by

$$Ax = (x_1, \frac{x_2}{2}, \frac{x_3}{3}, \cdots, \frac{x_k}{k}, \cdots) \ \text{ for all } x \in X.$$

It is straightforward to check that

$$\sigma_p(A) = \{\frac{1}{k} : k \in \mathbb{N}\}.$$

If $\{e^{(n)}\}_{n=1}^{\infty}$ denotes the standard orthonormal basis for l^2, then

$$\|Ae^{(n)}\| = \frac{1}{n} \to 0 \text{ as } n \to \infty,$$

and consequently 0 is a generalized eigenvalue, but not an eigenvalue.

Compact Operators

Compact operators have special properties with regard to spectral analysis.

Proposition 3.10: Let X be a Hilbert space, $\lambda \in \mathbb{K}\backslash\{0\}$, and $A \in \mathcal{C}(X;X)$ be given. Then $\mathcal{N}(\lambda I - A)$ is finite dimensional.

Proof: Suppose that $\mathcal{N}(\lambda I - A)$ is infinite dimensional. Then we may choose an orthonormal sequence $\{z_n\}_{n=1}^{\infty}$ such that $z_n \in \mathcal{N}(\lambda I - A)$ for all $n \in \mathbb{N}$. Then $z_n \rightharpoonup 0$ (weakly) as $n \to \infty$. Since A is compact, we can conclude that $Az_n \to 0$ (strongly) as $n \to \infty$. Since $\lambda z_n = Az_n$ for all $n \in \mathbb{N}$ and $\lambda \neq 0$ we see that $z_n \to 0$ (strongly) as $n \to \infty$, which is impossible since $\|z_n\| = 1$ for all $n \in \mathbb{N}$. $\square$

Proposition 3.11: Let X be a Hilbert space, $\lambda \in \mathbb{K}\backslash\{0\}$, and $A \in \mathcal{C}(X;X)$ be given. Then $\mathcal{R}(\lambda I - A)$ is closed.

Proof: Let $y \in X$ and a sequence $\{x_n\}_{n=1}^{\infty}$ in X be given such that

$$(\lambda I - A)x_n \to y \ \text{ as } n \to \infty.$$

Put $M = \mathcal{N}(\lambda I - A)^{\perp}$ and

$$z_n = P_M x_n \ \text{ for all } n \in \mathbb{N}.$$

Since $x_n - z_n \in \mathcal{N}(\lambda I - A) = M^{\perp}$ for all $n \in \mathbb{N}$ we see that

$$(\lambda I - A)z_n \to y \ \text{ as } n \to \infty.$$

I claim that $\{z_n\}_{n=1}^{\infty}$ is bounded. Indeed, suppose $\{z_n\}_{n=1}^{\infty}$ is unbounded and choose a subsequence $\{z_{n_k}\}_{k=1}^{\infty}$ such that

$$\|z_{n_k}\| > k \ \text{ for all } k \in \mathbb{N}.$$

Put

$$w_k = \frac{z_{n_k}}{\|z_{n_k}\}|} \ \text{ for all } k \in \mathbb{N},$$

and observe that

$$\|w_k\| = 1 \text{ for all } k \in \mathbb{N} \ \text{ and } (\lambda I - A)w_k \to 0 \text{ (strongly) as } n \to \infty.$$

Since A is compact, we may choose a sequence $\{w_{k_j}\}_{j=1}^{\infty}$ converges strongly as $k \to \infty$. Since $\lambda w_{k_j} = A w_{k_j}$ for all $j \in \mathbb{N}$ and $\lambda \neq 0$ we conclude that $\{w_{k_j}\}_{j=1}^{\infty}$ is strongly convergent; put

$$w = \lim_{j \to \infty} w_{k_j}.$$

Then $\|w\| = 1$, $w \in \mathcal{N}(\lambda I - A)^{\perp}$, and $(\lambda I - A)w = 0$, which is impossible. We conclude that $\{z_n\}_{n=1}^{\infty}$ is bounded.

Since $\{z_n\}_{n=1}^{\infty}$ is bounded, we may choose a subsequence $\{z_{n_j}\}_{j=1}^{\infty}$ such that $\{Az_{n_j}\}_{j=1}^{\infty}$ is strongly convergent. Since $\lambda \neq 0$ we conclude that $\{z_{n_j}\}_{j=1}^{\infty}$ is strongly convergent to some $z \in X$ and consequenty $(\lambda I - A)z = y$ and $y \in \mathcal{R}(\lambda I - A)$. $\square$

Proposition 3.12: Let X be a Hilbert space, $\lambda \in \mathbb{K}\backslash\{0\}$, and $A \in \mathcal{C}(X; X)$ be given and assume that λ is a generalized eigenvalue for A. Then $\lambda \in \sigma_p(A)$.

Proof: Choose a sequence $\{x_n\}_{n=1}^{\infty}$ such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $(\lambda I - A)x_n \to 0$ as $n \to \infty$. Since A is compact, we may choose a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ such that $\{Ax_{n_k}\}_{k=1}^{\infty}$ is strongly convergent. Since $\lambda \neq 0$ and $\{(\lambda I - A)x_{n_k}\}_{k=1}^{\infty}$ is strongly convergent, we conclude that $\{x_{n_k}\}_{k=1}^{\infty}$ is strongly convergent; put

$$x = \lim_{k \to \infty} x_{n_k}.$$

Then, we have $\|x\| = 1$ and $(\lambda I - A)x = 0$, so that $\lambda \in \sigma_p(A)$. $\square$

Corollary 3.13: Let X be a Hilbert space, $\lambda \in \mathbb{K}\backslash\{0\}$, and $A \in \mathcal{C}(X;X)$ be given. Assume that $\lambda \notin \sigma_p(A)$ and $\overline{\lambda} \notin \sigma_p(A^*)$. Then $\lambda \in \rho(A)$.

Proof: Since $\mathcal{R}(\lambda I - A)$ is closed, we have

$$\mathcal{R}(\lambda I - A) = \mathcal{N}(\overline{\lambda} I - A^*)^\perp = \{0\}^\perp = X,$$

so that $(\lambda I - A)$ is surjective. Since $\lambda \notin \sigma_p(A)$, we also have $(\lambda I - A)$ is also injective. $\square$

The assumption that $\overline{\lambda} \notin \sigma_p(A^*)$ in Corollary 3.13 is actually redundant. We shall prove this later on. However, for developing spectral prroperties of compact normal operators, Corollary 3.13 in the form given above will be sufficient.

Some Spectral Properties of Normal Operators

Proposition 3.14: Let X be a Hilbert space, $A \in \mathcal{L}(X;X)$ and $\lambda \in \mathbb{K}$ be given. Assume that A is normal. Then

$$\mathcal{N}(\lambda I - A) = \mathcal{N}(\overline{\lambda} I - A^*) \tag{4}$$

and $\mathcal{N}(\lambda I - A)$ reduces A.

Proof: Since $\lambda I - A$ is normal and $(\lambda I - A)^* = \overline{\lambda} I - A^*$, Corollary 1.11 implies that (4) holds. We want to show that $\mathcal{N}(\lambda I - A)$ is invariant under A and A^*. Let

$$x \in \mathcal{N}(\lambda I - A) = \mathcal{N}(\overline{\lambda} I - A^*)$$

be given.Then we have

$$Ax = \lambda x \in \mathcal{N}(\lambda I - A), \quad A^* x = \overline{\lambda} x \in \mathcal{N}(\lambda I - A),$$

and $\mathcal{N}(\lambda I - A)$ is invariant under both A and A^*. $\square$

Proposition 3.15: Let X be a Hilbert space, $A \in \mathcal{L}(X;X)$ be given and assume that A is normal. Let $\lambda, \mu \in \sigma_p(A)$ with $\lambda \neq \mu$ be given. Then

$$\mathcal{N}(\lambda I - A) \perp \mathcal{N}(\mu I - A).$$

Proof:Let $x \in \mathcal{N}(\lambda I - A)$, $y \in \mathcal{N}(\mu I - A)$ be given. Then $y \in \mathcal{N}(\overline{\mu} I - A^*)$ and consequently

$$\begin{aligned} \lambda(x,y) &= (\lambda x, y) = (Ax, y) \\ &= (x, A^* y) = (x, \overline{\mu} y) \\ &= \mu(x,y). \end{aligned}$$

Since $\lambda \neq \mu$ we conclude that $(x, y) = 0$. $\square$

Corollary 3.16: Let X be a Hilbert space and assume that $A \in \mathcal{L}(X; X)$ is self-adjoint. Then $\sigma_p(A) \subset \mathbb{R}$.

Proof: Let $\lambda \in \sigma_p(A)$ be given and choose a unit vector $x \in \mathcal{N}(\lambda I - A) = \mathcal{N}(\overline{\lambda} I - A)$. Then we have

$$\lambda x = Ax = \overline{\lambda} x,$$

which gives

$$\lambda(x, x) = \overline{\lambda}(x, x).$$

Since $\|x\| = 1$ we have $\lambda = \overline{\lambda}$. $\square$

Spectral Decomposition of Compact Self-Adjoint Operators

We are going to prove that every compact self-adjoint operator can be diagonalized. The key step is to show that every compact self-adjoint operator (on a nontrivial Hilbert space) has at least one eigenvalue.

Proposition 3.17: Let X be a Hilbert space and assume that $X \neq \{0\}$. Let $A \in \mathcal{C}(X; X)$ be given and assume that $A = A^*$. Then at least one of $\|A\|$, $-\|A\|$ is an eigenvalue of A.

Proof: If $A = 0$ we are done, so assume that $A \neq 0$. Recall that

$$\|A\| = \sup\{|(Ax, x)| : x \in X, \|x\| = 1\}.$$

Choose a sequence $\{x_n\}_{n=1}^{\infty}$ in X such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and

$$|(Ax_n, x_n)| \to \|A\| \text{ as } n \to \infty.$$

Notice that $(Ax_n, x_n) \in \mathbb{R}$ for all $n \in \mathbb{N}$ since A is self-adjoint. We may choose $\lambda \in \mathbb{R}$ such that $|\lambda| = \|A\|$ and a subsequence $\{x_{n_k}\}_{n=1}^{\infty}$ such that

$$(Ax_{n_k}, x_{n_k}) \to \lambda \text{ as } n \to \infty.$$

Observe that

$$\begin{aligned} \|(\lambda I - A)x_{n_k}\|^2 &= \lambda^2 \|x_{n_k}\|^2 - 2\lambda(Ax_{n_k}, x_{n_k}) = \|Ax_{n_k}\|^2 \\ &\leq \lambda^2 - 2\lambda(Ax_{n_k}, x_{n_k}) + \lambda^2 \to 0 \text{as } k \to \infty. \end{aligned}$$

We conclude that λ is a generalized eigenvalue for A. Since $\lambda \neq 0$, it follows from Proposition that λ is an eigenvalue for A. $\square$

Lemma 3.18: Let X be a Hilbert space and $A \in \mathcal{C}(X; X)$. Assume that A is self-adjoint. Then there is an orthonormal basis $(e_i | i \in J)$ for X such that

$$\forall j \in I, \quad e_i \text{ is an eigenvector for } A.$$

Remark 3.19: Sometimes unnecessary assumptions regarding separability of X are made when discussing the spectral decomposition of compact self-adjoint opeartors. Notice that we care only about

$$A\Big|_{\mathcal{N}(A)^\perp}.$$

If A is compact and self-adjoint then $\mathrm{cl}(\mathcal{R}(A))$ is separable, and consequently $\mathcal{N}(A)^\perp = \mathcal{N}(A^*)^\perp$ is separable.

Proof of Lemma 3.19: If $A = 0$ the result is immediate, so assume that $A \neq 0$. By Proposition 3.17, we may choose an eigenvector x with $\|x\| = 1$ corresponding to an eigenvalue $\pm\|A\|$. Let $\mathcal{E}$ be the collection of all orthonormal sets of eigenvectors for A, partially ordered by set inclusion. Since every chain has an upper bound (just take the union), Zorn's Lemma implies that $\mathcal{E}$ has a maximal element $\mathcal{O}$. Let us put

$$W = \mathrm{cl}(\mathrm{span}(\mathcal{O})).$$

We want to show that $W = X$, i.e. that $W^\perp = \{0\}$. Notice that $A[W] \subset W$ and $A[W^\perp] \subset W^\perp$. Furthermore

$$A\Big|_{W^\perp}$$

is self-adjoint. Suppose $W^\perp \neq \{0\}$. Then there is an eigenvector $y \in W^\perp$ with $\|y\| = 1$ for $A\Big|_{W^\perp}$. Since y is also an eigenvector for A and y is orthogonal to $\mathcal{O}$, this contradicts the maximality of $\mathcal{O}$. It follows that $W^\perp = \{0\}$. $\square$

Combining results above with a few additional observations, we obtain the following very important theorem.

Theorem 3.20 (Spectral Theorem for Compact Self-Adjoint Operators): Let X be a (real or complex) Hilbert space and let $A \in \mathcal{C}(X;X)$ be given. Assume that $A^* = A$. Then we have

(i) $\sigma(A) \subset \sigma_p(A) \cup \{0\}$.

(ii) $\sigma_p(A) \subset \mathbb{R}$, $\sigma_p(A) \subset [-\|A\|, \|A\|]$, at least one $-\|A\|, \|A\|$ belongs to $\sigma_p(A)$.

(iii) $\sigma_p(A)$ is countable and 0 is the only possible accumulation point.

(iv) A has finite rank if and only if $\sigma_p(A)$ is finite.

(v) $\forall \lambda, \mu \in \mathbb{K}$ with $\lambda \neq \mu$ we have $\mathcal{N}(\lambda I - A) \perp \mathcal{N}(\mu I - A)$.

(vi) $\forall \lambda \in \mathbb{K}\backslash\{0\}$, $\mathcal{N}(\lambda I - A)$ is finite dimensional and $\mathcal{R}(\lambda I - A)$ is closed.

(vii) There is an orthonormal basis $(e_i | i \in J)$ for X such that for every $i \in J$, e_i is an eigenvector for A.

Proof: We have already established (i), (v), (vi), (vii).

To establish (ii), it remains only to show that $|\lambda| \le \|A\|$ for all $\lambda \in \sigma_P(A)$. Let $\lambda \in \sigma_p(A)$ be given and choose $x \in \mathcal{N}(\lambda I - A)$ with $\|x\| = 1$. Then we have

$$|\lambda| = \|\lambda x\| = \|Ax\| \le \|A\|.$$

In order to prove (iii), we shall show that for every $\epsilon > 0$, the set

$$\{\lambda \in \sigma_p(A) : |\lambda| \ge \epsilon\}$$

is finite. Suppose that there is some $\epsilon_0 > 0$ such that

$$\{\lambda \in \sigma_p(A) : |\lambda| \ge \epsilon_0\}$$

is infinite. Then we may choose an injective sequence $\{\lambda_i\}_{i=1}^{\infty}$ such that

$$\lambda_i \in \sigma_p(A), \quad |\lambda_i| \ge \epsilon_0 \text{ for all } i \in \mathbb{N}.$$

Using the fact that eigenvectors corresponding to distinct eigenvalues are orthogonal, we may choose an orthonormal sequence $\{e_i\}_{i=1}^{\infty}$ such that

$$Ae_i = \lambda_i e_i \text{ for all } i \in \mathbb{N}.$$

Notice that for $i \ne j$ we have

$$\|Ae_i - Ae_j\|^2 \ge |\lambda_i|^2 + |\lambda_j|^2 \ge 2\epsilon_0^2.$$

This is impossible since A is compact which implies that $\{Ae_i\}_{i=1}^{\infty}$ must have a strongly convergent subsequence.

Finally, (iv) follows easily from (vii) and the fact that the eigenspaces corresponding to distinct nonzero eigenvalues are finite dimensional and orthogonal. $\square$

Lecture Notes for Week 4 (First Draft)

Simultaneous Diagonalization of Compact Self-Adjoint Operators

Proposition 4.1: Let X be a (real or complex) Hilbert space and let $A, B \in \mathcal{C}(X;X)$ with $A^* = A$ and $B^* = B$ be given. Then $AB = BA$ if and only if there is an orthonormal basis $(e_j | j \in J)$ for X such that for every $j \in J$, e_j is an eigenvector for A and an eigenvector for B.

Proof: Assume first that $AB = BA$. For every $\lambda \in \sigma_p(A)$, put

$$M_\lambda = \mathcal{N}(\lambda I - B).$$

Let $\lambda \in \sigma_p(A)$ and $x \in M_\lambda$ be given. Notice that

$$\begin{aligned} B(Ax) &= A(Bx) = A(\lambda x) \\ &= \lambda(Ax), \end{aligned}$$

and consequently $Ax \in M_\lambda$. In other words, M_λ is invariant under A. Since

$$A\Big|_{M_\lambda}$$

is compact and self-adjoint, we may choose an orthonormal basis $\mathcal{O}_\lambda$ for M_λ such that every $e \in \mathcal{O}_\lambda$ is an eigenvalue for

$$A\Big|_{M_\lambda}$$

and hence also an eigenvalue for A. (Each such e is, of course, an eigenvalue for B. The desired orthonormal basis for X is given by

$$\mathcal{U} = \bigcup_{\lambda \in \sigma_p(A)} \mathcal{O}_\lambda.$$

The converse implication is immediate. $\square$

Spectral Decomposition of Compact Normal Operators

A compact normal operator need not have any eigenvalues if $\mathbb{K} = \mathbb{R}$. A simple example in $\mathbb{R}^2$ is given by

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Fact: If $X \neq \{0\}$, $\mathbb{K} = \mathbb{C}$, and $A \in \mathcal{C}(X;X)$ is normal, then A has an eigenvalue λ with $|\lambda| = \|A\|$.

Using this fact (which I plan to prove later), we could repeat the proof of Theorem 3.20 and obtain an analogous spectral decomposition theorem for compact normal operators on Complex Hilbert spaces. We can also obtain the desired spectral theorem for compact normal operators by applying Proposition 4.1 to ensure that a compact normal operator on a nontrivial complex Hilbert space has an eigenvalue. I shall adopt the latter approach.

Let X be a complex Hilbert space and let $A \in \mathcal{L}(X;X)$ be given. Assume that A is normal. Then we have

$$A = \frac{1}{2}(A + A^*) + i\left(\frac{1}{2i}(A - A^*)\right).$$

Let us put

$$B = \frac{1}{2}(A + A^*), \quad C = \frac{1}{2i}(A - A^*).$$

It is immediate that

$$A = B + iC, \quad B^* = B, \quad C^* = C.$$

Moreover, since A is normal, we have

$$BC = CB.$$

Assume now that A is compact. Then A^* is compact, so B and C are both compact. Thus we can choose an orthormal basis $(e_j | j \in J)$ for X such that for every $j \in J$, e_j is an eigenvector for B and for C, i.e. for every $j \in J$ we may choose $\lambda_j, \mu_j \in \mathbb{R}$ such that

$$Ae_j = \lambda_j e_j + i\mu_j e_j.$$

Making minor adjustments to the proof of Theorem 3.20, we obtain

Theorem 4.2: Let X be a complex Hilbert space and let $A \in \mathcal{C}(X;X)$ be given and assume that A is normal. Then

(i) $\sigma(A) \subset \sigma_p(A) \cup \{0\}$

(ii) $|\lambda| \leq \|A\|$ for all $\lambda \in \sigma_p(A)$

(iii) $\sigma_p(A)$ is countable and the only possible accumulation point is 0

(iv) A has finite rank if and only if $\sigma_p(A)$ is finite

(v) $\forall \lambda, \mu \in \sigma_p(A)$ with $\lambda \neq \mu$ we have $\mathcal{N}(\lambda I - A) \perp \mathcal{N}(\mu I - A)$

(vi) For all $\lambda \in \sigma_p(A) \backslash \{0\}$, $\mathcal{N}(\lambda I - A)$ is finite dimensional and $\mathcal{R}(\lambda I - A)$ is closed

(vii) There is an orthonormal basis $(e_j | j \in J)$ for X such that for every $j \in J$, e_j is an eigenvector for A.

Spectral Resolution of Bounded Self-Adjoint Operators

Theorem 4.3: Let X be a (real or complex) Hilbert space and assume that $X \neq \{0\}$. Let $A \in \mathcal{L}(X;X)$ be given and assume that $A^* = A$. Put

$$m = \inf\{(Ax, x) : x \in X, \|x\| = 1\}, \quad M = \sup\{(Ax, x) : x \in X, \|x\| = 1\}.$$

Then there is a family $(E(\lambda) | \lambda \in \mathbb{R})$ of orthogonal projections [i.e., for every $\lambda \in \mathbb{R}$ we have $E(\lambda) \in \mathcal{L}(X;X)$, $E(\lambda)^2 = E(\lambda)$ and $E(\lambda)^* = E(\lambda)$] having the following properties.

(i) $\mathcal{R}(\lambda_1) \subset \mathcal{R}(\lambda_2)$ for $\lambda_1 \leq \lambda_2$,

(ii) For all $\lambda_0 \in \mathbb{R}$ and all $x \in X$ we have $E(\lambda)x \to E(\lambda_0)$ as $\lambda \downarrow \lambda_0$

(iii) $E(\lambda) = 0$ for all $\lambda < m$ and $E(\mu) = I$ for all $\mu \geq M$

(iv) $AE(\lambda) = E(\lambda)A$ for all $\lambda \in \mathbb{R}$

(v) For all $a, b \in \mathbb{R}$ with $a < m$ and $b \geq M$ we have

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

Before proving Theorem 4.3, we need to develop some preliminary material concerning order on self-adjoint operators and decomposition of a self-adjoint operator into a difference of two positive self-adjoint operators.

Order on Self-adjoint Operators

Definition 4.4: Let X be a Hilbert space and $A, B \in \mathcal{L}(X;X)$ with $A^* = A$ and $B^* = B$ be given. We write $A \geq B$ [or, equivalently $B \leq A$] provided that

$$\forall x \in X, \text{ we have } (Ax, x) \geq (Bx, x).$$

[Recall that $(Ax, x), (Bx, x) \in \mathbb{R}$ for all $x \in X$ since A and B are self-adjoint.] We say that A is *positive* provived that $A \geq 0$.

You should verify the following two remarks for yourself as a simple exercise.

Remark 4.5: Let $A, B \in \mathcal{L}(X;X)$ with $A^* = A$ and $B^* = B$ be given. Assume that $AB = BA$ and $A \geq 0$, $B \geq 0$. Then we have $AB \geq 0$.

Remark 4.6: Let W and Z be closed subspaces of X. Then we have

$$W \subset Z \;\Leftrightarrow\; P_W \leq P_Z.$$

Lemma 4.7: Let X be a Hilbert space and $A \in \mathcal{L}(X;X)$ be given. Assume that $A^* = A$ and $A \geq 0$. Then there is exactly one $B \in \mathcal{L}(X;X)$ with $B^* = B$ and $B \geq 0$ such that $B^2 = A$. We call B the *square root* of A and we write $B = \sqrt{A}$. Moreover, if $C \in \mathcal{L}(X;X)$ and satisfies $AC = CA$ then C also satisfies $\sqrt{A}C = C\sqrt{A}$.

The proof of Lemma 4.7 will be a homework exercise. The idea is to use a suitable iteration scheme.

Decomposition of Self-Adjoint Operators

Let $A \in \mathcal{L}(X;X)$ with $A^* = A$ be given. Observe that

$$(A^2x, x) = (Ax, Ax) \geq 0 \;\text{ for all } x \in X,$$

so that $A^2 \geq 0$ (and, of course, A^2 is self-adjoint). We put

$$\begin{aligned} |A| &= \sqrt{A^2}, \\ A^+ &= \frac{1}{2}(|A| + A), \\ A^- &= \frac{1}{2}(|A| - A). \end{aligned}$$

Observe that

$$A = A^+ - A^-, \quad |A| = A^+ + A^-.$$

The following remark summarizes some basic properties of the above decomposition that will be used in the proof of Theorem 4.3.

Remark 4.8: Let X be a Hilbert space and let $A, B \in \mathcal{L}(X;X)$ with $A^* = A$ be given.

(i) $A^+ \geq 0$, $A^- \geq 0$, $A^+ \geq A$, $A^- \geq -A$

(ii) $A^+A^- = A^-A^+ = 0$

(iii) If $A \geq 0$ then $A = |A| = A^+$

(iv) If $A \leq 0$ then $A = -|A| = -A^-$

(v) If $AB = BA$ then $A^+B = BA^+$ and $A^-B = BA^-$.

(vi) If $B^* = B$ and $B \geq A$ then $B \geq A^+$.

You should prove this remark for yourself as a simple exercise. (All parts are reasonably straightforward.)

Proof of Theorem 4.3: Assume that $X \neq \{0\}$, $A \in \mathcal{L}(X;X)$, and $A^* = A$. Put

$$m = \inf\{(Ax,x) : x \in X, \|x\| = 1\}, \quad M = \sup\{(Ax,x) : x \in X, \|x\| = 1\}. \tag{1}$$

For all $\lambda \in \mathbb{R}$, let

$$L(\lambda) = A - \lambda I,$$

and observe that

$$L(\lambda)L(\mu) = L(\mu)L(\lambda) \ \text{ for all } \lambda, \mu \in \mathbb{R}.$$

Since $\lambda \in \mathbb{R}$ and $A^* = A$ we know that $L(\lambda)$ is self-adjoint for all $\lambda \in \mathbb{R}$. A simple computation shows that

$$L(\lambda_1) \geq L(\lambda_2) \ \text{ for } \lambda_1 \leq \lambda_2. \tag{2}$$

It follows from part (i) of Remark 4.8 and equation (2) that

$$L(\lambda_1)^+ \geq L(\lambda_1) \geq L(\lambda_2) \ \text{ for } \lambda_1 \leq \lambda_2. \tag{3}$$

Using part (v) of Remark 4.8 we conclude that

$$L(\lambda_1)^+ \geq L(\lambda_2)^+ \ \text{ for } \lambda_1 \leq \lambda_2. \tag{4}$$

It follows immediately from (1) and parts (iii) and (iv) of Remark 4.8 that

$$L(\lambda) = |L(\lambda)| = L(\lambda)^+ \ \text{ for } \lambda \leq m \tag{5}$$

and

$$L(\lambda) = -|L(\lambda)| = -L(\lambda)^- \ \text{ for } \lambda \geq M. \tag{6}$$

It follows from (2) and Remark4.5 that

$$L(\lambda_2)^+[L(\lambda_1)^+ - L(\lambda_2)^+] \geq 0. \tag{7}$$

Rearranging (7) we find that

$$L(\lambda_2)^+L(\lambda_1)^+ \geq (L(\lambda_2)^+)^2 \ \text{ for} \lambda_1 \leq \lambda_2. \tag{8}$$

It follows from (8) that

$$\mathcal{N}(L(\lambda_1)^+) \subset \mathcal{N}(L(\lambda_2)^+) \ \text{ for } \lambda_1 \leq \lambda_2. \tag{9}$$

To see why (9) follows from (8), let $x \in \mathcal{N}(L(\lambda_1)^+)$ be given. Using (8) we find that

$$\|L(\lambda_2)^+x\|^2 = (L(\lambda_2)^+x, L(\lambda_2)^+x) = ((L(\lambda_2)^+)^2x, x) = 0.$$

Now, for all $\lambda \in \mathbb{R}$, let $E(\lambda)$ denote the orthogonal projection onto $\mathcal{N}(L(\lambda)^+)$. It follows easily from Remark 4.6, equation (9) and the definition of $E(\lambda)$ that

$$E(\lambda_1) \leq E(\lambda_2) \ \text{ for } \lambda_1 \leq \lambda_2. \tag{10}$$

A simple computation gives

$$((A-\lambda I)x,x) \geq (m-\lambda)\|x\|^2 \ \text{ for all } \lambda \in \mathbb{R}, x \in X. \tag{11}$$

Let $\lambda < m$ be given. It follows from (11) that $L(\lambda) \geq 0$ and $\mathcal{N}(L(\lambda)) = \{0\}$. We conclude that

$$\mathcal{N}(L(\lambda)^+) = \{0\} \text{ for } \lambda < m, \tag{12}$$

and consequently

$$E(\lambda) = 0 \ \text{ for } \lambda < m. \tag{13}$$

Observe also that

$$((A-\lambda I)x,x) \leq 0 \ \text{ for } \lambda \geq M,$$

and consequently

$$L(\lambda)^+ = 0, \text{ i.e., } \mathcal{N}(L(\lambda))^+ = X \text{ for } \lambda \geq M. \tag{14}$$

It follows from (14) that

$$E(\lambda) = I \ \text{ for} \lambda \geq M. \tag{15}$$

Lecture Notes for Week 5 (First Draft)

Continuation of the Proof of Theorem 4.3

It follows from Remark 4.6 that

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

Put

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

and observe that

$$\mathcal{E}(\lambda_1, \lambda_2) \geq 0 \quad \text{for } \lambda_1 \leq \lambda_2.$$

Observe further that

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

and also that

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

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

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

Using (2) and (4) we find that

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

Combining (5) and (6) we arrive at

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

Let $a < m$ and $b \ge M$ 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

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

Since

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

it follows from (8) that

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

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

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

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

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

and notice that

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

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

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

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
$$z^n - \mu = \prod_{j=1}^{n} (z - \alpha_j) \quad \text{for all } z \in \mathbb{C}.$$
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

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

It follows that

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

It remains to show that

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

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

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

Lecture Notes for Week 6 (First Draft)

A General Spectral Representation Theorem

Let X be a complex Banach space and $T \in \mathcal{L}(X;X)$ be given. For $|\lambda| > r_\sigma(T)$ we have

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

Let $k \in \mathbb{N}$ be given. It follows from (1) that

$$\lambda^k R(\lambda;T) = \sum_{n=0}^{\infty}\lambda^{k-n-1}T^n \quad \text{for } |\lambda| > r_\sigma(T). \tag{2}$$

Let C be a circle in the complex plane with center at 0 and radius strictly greater than $r_\sigma(T)$, oriented counterclockwise. Recall that if m is an integer, then

$$\int_C \lambda^m \, d\lambda = \begin{cases} 2\pi i & \text{if } m = -1, \\ 0 & \text{if } m \neq -1. \end{cases} \tag{3}$$

It follows from (2) and (3) that

$$\int_C \lambda^k R(\lambda;T)\, d\lambda = 2\pi i A^k.$$

We have just proved the following important result.

Theorem 6.1: Let X be a complex Banach space (with $X \neq \{0\}$) and $T \in \mathcal{L}(X;X)$ and $k \in \mathbb{N}$ be given. Let C be a circle in the complex plane with center at 0, radius strictly greater than $r_\sigma(T)$, and oriented in the counterclockwise direction. Then

$$T^k = \frac{1}{2\pi i}\int_C \lambda^k R(\lambda;T)\, d\lambda;$$

in particluar

$$T = \frac{1}{2\pi i}\int_C \lambda R(\lambda;T)\, d\lambda.$$

This gives us a way to "extend" ordinary scalar analytic functions to analytic operator-valued functions. Let $T \in \mathcal{L}(X;X)$ and $\rho > r_\sigma(T)$ be given. Assume that

$$f : \{z \in \mathbb{C} : |z| < \rho\} \to \mathbb{C}$$

is analytic and let C be a circle in $\mathbb{C}$ centered at 0, having radius γ, with $r_\sigma(T) < \gamma < \rho$, and oriented in the counterclockwise direction. Then we can define

$$f(T) = \frac{1}{2\pi i} \int_C f(\lambda) R(\lambda; T)\, d\lambda.$$

One can prove many nice properties of $f(T)$; in particular

$$\mu \in \sigma(f(T)) \;\Leftrightarrow\; \mu = f(\lambda) \text{ for some } \lambda \in \sigma(T).$$

Moreover if

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad \text{for all } z \in \{w \in \mathbb{C} : |w| < \rho\},$$

then

$$f(T) = \sum_{n=0}^{\infty} a_n T^n.$$

Remark 6.2: In the integrals above, the circle C can be deformed to other "rectifiable" Jordan curves that lie within the domain of analyticity of f, contain $\sigma(T)$ in the interior region, and have counterclockwise orientation.

Some Simple Remarks about Spectra

Recall that an operator $B \in \mathcal{L}(X; X)$ is bijective if and only if the adjoint B^* is bijective and in this case we have

$$(B^{-1})^* = (B^*)^{-1}.$$

Moreover if $T \in \mathcal{L}(X; X)$ and $\lambda \in \mathbb{K}$ then

$$(\lambda I - T)^* = \lambda I - T^*.$$

(Here * indicates the Banach-space adjoint.)

Observe also that if $\lambda \neq 0$ and $T \in \mathcal{L}(X; X)$ is bijective then

$$\frac{1}{\lambda} I - T^{-1} \text{ is bijective } \;\Leftrightarrow\; \lambda I - T \text{ is bijective.}$$

The above observations yield the following simple, but very useful, results.

Remark 6.3: Let X be a Banach space and let $T \in \mathcal{L}(X; X)$ be given.

(a) Then $\sigma(T) = \sigma(T^*)$. Moreover, for all $\lambda \in \rho(T)$ we have $R(\lambda; T^*) = (R(\lambda; T))^*$.

(b) Assume that $0 \in \rho(T)$ and let $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then $\lambda \in \sigma(T^{-1})$ if and only if $\lambda^{-1} \in \sigma(T)$.

Applying Remark 6.3 to the special case when X is a Hilbert space we obtain the following useful observations.

Remark 6.4: Let X be a Hilbert space and let $A \in \mathcal{L}(X;X)$ be given. Let A^* denote the Hilbert space adjoint of A.

(a) Then $\sigma(A^*) = \{\lambda \in \mathbb{K} : \overline{\lambda} \in \sigma(A)\}$. Moreover, for all $\lambda \in \rho(A)$ we have $R(\overline{\lambda}; A^*) = (R(\lambda; A))^*$.

(b) Assume that $0 \in \rho(A)$ and let $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then $\lambda \in \sigma(A^{-1})$ if and only if $\lambda^{-1} \in \sigma(A)$.

(c) If A is self adjoint then $\sigma(A) \subset \mathbb{R}$.

(d) If A is unitary then $|\lambda| = 1$ for all $\lambda \in \sigma(A)$.

Spectral Theory of Compact Operators

Proposition 6.5: Let X be a Banach space and let $T \in \mathcal{C}(X;X)$ be given. Then $\sigma_p(T)$ is countable and 0 is the only possible accumulation point.

Proof: For $\epsilon > 0$ put
$$\Lambda_\epsilon = \{\lambda \in \sigma_p(T) : |\lambda| \geq \epsilon.\}.$$
It suffices to show that Λ_ϵ is a finite set for every $\epsilon > 0$. Let $\epsilon_0 > 0$ be given and suppose that Λ_{ϵ_0} is infinite. Then we may choose an injective sequence $\{\lambda_n\}_{n=1}^\infty$ in Λ_{ϵ_0} and a sequence $\{x_n\}_{n=1}^\infty$ of corresponding eigenvectors. For each $n \in \mathbb{N}$, put

$$M_n = \mathrm{span}\{x_1, x_2, \cdots, x_n\},$$

and observe that M_n is invariant under T, i.e. $T[M_n] \subset M_n$, and that $M_n \subset M_{n+1}$. Observe also that each M_n is a closed subspace of X.

Let $n \in \mathbb{N}$ and $x \in M_n$ be given. Then we may choose $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{K}$ such that
$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n.$$
Therefore we have

$$(\lambda_n I - T)x = (\lambda_n - \lambda_1)x_1 + (\lambda_n - \lambda_2)x_2 + \cdots + (\lambda_n - \lambda_{n-1})x_{n-1} + 0.$$

It follows that
$$(\lambda_n I - T)[M_n] \subset M_{n-1}. \text{ for } n \geq 2.$$

Using the Riesz Lemma, we may choose a sequence $\{y_n\}_{n=1}^{\infty}$ such that

$$y_n \in M_n, \quad \|y_n\| = 1 \text{ for all } n \in \mathbb{N},$$

$$\forall n \geq 2, \text{ we have } \|y_n - x\| \geq \frac{1}{2} \text{ for all } x \in M_{n-1}.$$

Now let $m, n \in \mathbb{N}$ with $m < n$ be given and put

$$x = Ty_m + (\lambda_n I - T)y_n.$$

Notice that $x \in M_{n-1}$ since $Ty_m \in M_m \subset M_{n-1}$ and $(\lambda_n I - T)y_n \in M_{n-1}$. We have

$$Ty_n - Ty_m = \lambda_n y_n - (\lambda_n I - T)y_n - Ty_m = \lambda_n \left(y_n - \frac{1}{\lambda_n} x \right). \tag{4}$$

It follows from (4) that

$$\|Ty_n - Ty_m\| \geq \frac{1}{2}|\lambda_n| \geq \frac{1}{2}\epsilon_0.$$

We conclude that $\{Ty_n\}_{n=1}^{\infty}$ has no convergent sequence, which is impossible because $\{y_n\}_{n=1}^{\infty}$ is bounded and T is compact. $\square$

Proposition 6.6: Let X be a Banach space and let $T \in \mathcal{C}(X;X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then $\mathcal{N}(\lambda I - T)$ is finite dimensional.

Proof: Let

$$K = \{x \in \mathcal{N}(\lambda I - T) : \|x\| \leq 1\},$$

i.e., the closed unit ball in $\mathcal{N}(\lambda I - T)$ equipped with the norm of X. Let $\{x_n\}_{n=1}^{\infty}$ be any sequence in K. Then

$$Tx_n = \lambda x_n \text{ for all } n \in \mathbb{N}.$$

Since T is compact and $\{x_n\}_{n=1}^{\infty}$ is bounded, we can extract a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ such that $\{Tx_{n_k}\}_{k=1}^{\infty}$ is strongly convergent. Since $\lambda \neq 0$, we see that $\{x_{n_k}\}_{k=1}^{\infty}$ is also strongly convergent. Let

$$x = \lim_{k\to\infty} x_{n_k},$$

and observe that $x \in K$ since K is closed. We conclude that K is compact and consequently $\mathcal{N}(\lambda I - T)$ is finite dimensional. $\square$

Proposition 6.7: Let X be a Banach space and let $T \in \mathcal{C}(X;X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then $\mathcal{R}(\lambda I - T)$ is closed.

Proof: Put $W = \mathcal{N}(\lambda I - T)$ and observe that W is finite dimensional by Proposition 6.6. Therefore, we may choose a closed subspace Z of X such that

$$X = W \oplus Z.$$

Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X and put

$$y_n = (\lambda I - T)x_n \quad \text{for all } n \in \mathbb{N}.$$

Let $y \in X$ be given and assume that $y_n \to y$ as $n \to \infty$. We need to show that $y \in \mathcal{R}(\lambda I - T)$.

We may choose sequences $\{w_n\}_{n=1}^{\infty}$ in W and $\{z_n\}_{n=1}^{\infty}$ in Z such that

$$x_n = w_n + z_n \quad \text{for all } n \in \mathbb{N}.$$

Since $(\lambda I - T)w_n = 0$ for all $n \in \mathbb{N}$ we see that

$$(\lambda I - T)z_n \to y \text{ as } n \to \infty. \tag{5}$$

I claim that $\{z_n\}_{n=1}^{\infty}$ is bounded. To verify the claim, suppose that $\{z_n\}_{n=1}^{\infty}$ is unbounded. Then we may choose a subsequence $\{z_{n_k}\}_{k=1}^{\infty}$ such that

$$\|z_{n_k}\| > k \quad \text{for all } k \in \mathbb{N}.$$

Put

$$v_k = \frac{z_k}{\|z_k\|} \quad \text{for all } k \in \mathbb{N},$$

and observe that $\|v_k\| = 1$ for all $k \in \mathbb{N}$ and that

$$(\lambda I - T)v_k \to 0 \text{ as } k \to \infty. \tag{6}$$

Since T is compact, we may choose a subsequence $\{v_{k_j}\}_{j=1}^{\infty}$ such that $\{Tv_{k_j}\}_{j=1}^{\infty}$ is strongly convergent. Since $\lambda \neq 0$, we conclude from (6) that $\{v_{k_j}\}_{j=1}^{\infty}$ is also strongly convergent; put

$$v = \lim_{j\to\infty} v_{k_j}.$$

We have $\|v\| = 1$ (since the convergence is strong), $v \in Z$ (since Z is closed), and $(\lambda I - T)v = 0$ (which means that $v \in W$). This is impossible, because $W \cap Z = \{0\}$. We conclude that $\{z_n\}_{n=1}^{\infty}$ is bounded.

Since $\{z_n\}_{n=1}^{\infty}$ is bounded, we may extract a subsequence $\{z_{n_j}\}_{j=1}^{\infty}$ such that $\{Tz_{n_j}\}_{j=1}^{\infty}$ is strongly convergent. Since $\lambda \neq 0$, it follows from (5) that $\{z_{n_j}\}_{j=1}^{\infty}$ is strongly convergent; put

$$z = \lim_{j\to\infty} z_{n_j}.$$

We have $(\lambda I - T)z = y$ and $y \in \mathcal{R}(\lambda I - T)$. $\square$

Corollary 6.8: Let X be a Banach space and let $T \in \mathcal{C}(X;X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then

$$\mathcal{R}(\lambda I - T) = {}^{\perp}\mathcal{N}(\lambda I - T^*), \text{ and}$$
$$\mathcal{R}(\lambda I - T^*) = \mathcal{N}(\lambda I - T)^{\perp}.$$

In spectral analysy in finite dimensions, when there are not enough eigenvectors corresponding to an eigenvalue λ of an $N \times N$ matrix B, then we look at the null spaces of higher powers of $(\lambda I - B)$. The same idea is useful for compact operators in infinite-dimensional space.

We now want to look at powers of the form $(\lambda I - T)^n$ where T is a compact operator, $\lambda \neq 0$, and n is a nonnegative integer. For $n \geq 1$ we introduce

$$L = \sum_{k=1}^{n} \binom{n}{k} (-1)^k \lambda^{n-k} T^k, \tag{7}$$

and observe that L is compact since T is compact.

Proposition 6.9: Let X be a Banach space, $T \in \mathcal{L}(X;X)$, $\lambda \in \mathbb{K}\backslash\{0\}$, and n be a nonnegative integer. Then $\mathcal{N}((\lambda I - T)^n)$ is finite dimensional and $\mathcal{R}((\lambda I - T)^n)$ is closed.

Proof: If $n = 0$ then $(\lambda I - T)^n = I$ and the conclusion is immediate. If $n > 0$ then

$$(\lambda I - T)^n = \mu I - L,$$

where $\mu = \lambda^n \neq 0$ and L is given by (7) and is therefore compact. The conclusion follows from Propositions 6.6 and 6.7. $\square$

Lemma 6.10: Let X be a Banach space, $T \in \mathcal{C}(X;X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then there exists a nonnegative integer p such that $\mathcal{N}((\lambda I - T)^n)) = \mathcal{N}((\lambda I - T)^{n+1})$ for all integers $n \geq p$. Let $p_*(\lambda;T)$ be the smallest such nonnegative integer. If $p_*(\lambda;T) > 0$ then

$$\mathcal{N}((\lambda I - T)^0) \subset \mathcal{N}((\lambda I - T)^1) \subset \cdots \subset \mathcal{N}((\lambda I - T)^{p_*(\lambda;T)},$$

with all inclusions strict.

Proof: Put $N_m = \mathcal{N}((\lambda I - T)^m)$ and notice that

$$N_m \subset N_n \text{ for all nonnegative integers } m, n \text{ with } m \leq n.$$

Suppose that $N_n \neq N_{n+1}$ for all nonnegative integers n. Then, by the Riesz Lemma, we may choose a sequence $\{y_n\}_{n=1}^{\infty}$ such that

$$y_n \in N_{n+1}\backslash N_n, \quad \|y_n\| = 1, \quad \operatorname{dist}(y_n, N_n) \geq \frac{1}{2} \text{ for all } n \in \mathbb{N}. \tag{8}$$

Let $m, n \in \mathbb{N}$ with $m < n$ be given and observe that

$$(\lambda I - T)^{n+1} y_n = 0, \quad (\lambda I - T)^n y_m = 0.$$

It follows that

$$\begin{aligned} (\lambda I - T)^n[(\lambda I - T)y_n + Ty_m] &= (\lambda I - T)^{n+1} y_n + T(\lambda I - T)^n y_m \\ &= 0, \end{aligned}$$

and we conclude that

$$(\lambda I - T)y_n + Ty_m \in N_n. \tag{9}$$

Observe that

$$\begin{aligned} Ty_n - Ty_m &= \lambda y_n - (\lambda y_n - Ty_n + Ty_m) \\ &= \lambda[y_n - \lambda^{-1}(\lambda y_n - Ty_n + Ty_m)] \end{aligned} \tag{10}$$

It follows from (8), (9), and (10) that

$$\|Ty_n - Ty_m\| \geq \lambda\|y_n - y_m\| \geq \frac{|\lambda|}{2}.$$

This is impossible because $\{Ty_n\}_{n=1}^{\infty}$ must have a strongly convergent subsequence by virtute of the fact that T is compact and $\{y_n\}_{n=1}^{\infty}$ is bounded. It follows that there is some nonnegative integer k such that $N_k = N_{k+1}$.

Let a nonnegative integer n be given and assume that $N_n = N_{n+1}$. We shall show that $N_{n+1} = N_{n+2}$. It suffices to show that $N_{n+2} \subset N_{n+1}$. Let $x \in N_{n+2}$ be given. Then

$$0 = (\lambda I - T)^{n+2}x = (\lambda I - T)^{n+1}(\lambda I - T)x,$$

and consequently $(\lambda I - T)x \in N_{n+1}$. Since $N_{n+1} = N_n$, we see that $(\lambda I - T)x \in N_n$, which gives

$$(\lambda I - T)^{n+1}x = (\lambda I - T)^n(\lambda I - T)x = 0. \quad \square$$

Let $T \in \mathcal{L}(X; X)$ and $\lambda \in \backslash\{0\}$ be given and make the following observations:

- $T^* \in \mathcal{L}(X^*; X^*)$,
- $\mathcal{R}((\lambda I - T)^n) = {}^{\perp}\mathcal{N}(((\lambda I - T)^n)^*)$ for every nonnegative integer n,
- $((\lambda I - T)^n)^* = (\lambda I - T^*)^n$ for every nonnegative integer n,

In view of these observations and Lemma 6.10, we have

Lemma 6.11: Let X be a Banach space and $T \in \mathcal{C}(X; X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then there exists a nonnegative integer q such that $\mathcal{R}((\lambda I - T)^n) = \mathcal{R}((\lambda I - T)^{n+1})$ for all integers n with $n \geq q$. Let $q_*(\lambda; T)$ be the smallest such integer. If $q_*(\lambda; T) > 0$ then

$$\mathcal{R}((\lambda I - T)^0) \supset \mathcal{R}((\lambda I - T)^1) \supset \cdots \supset \mathcal{R}((\lambda I - T)^{q_*(\lambda; T)},$$

with all inclusions strict.

Lemma 6.12: Let X be a Banach space and $T \in \mathcal{C}(X; X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be gievn. Let $p_*(\lambda; T)$ and $q_*(\lambda; T)$ be as in Lemmas 6.10 and 6.11. Then we have $p_*(\lambda; T) = q_*(\lambda; T)$.

Proof: For each nonnegative integer n, let us put

$$N_n = \mathcal{N}((\lambda I - T)^n), \quad R_n = \mathcal{R}((\lambda I - T)^n).$$

For ease of notation let us write

$$p = p_*(\lambda; T), \quad q = q_*(\lambda; T).$$

We shall show first that $q \geq p$. Since $R_q = R_{q+1}$ we have

$$(\lambda I - T)[R_q] = R_q. \tag{11}$$

To establish the desired inequality we want to show that $N_{q+1} = N_q$. For this purpose, it is convenientshall show that

$$(\lambda I - T)\Big|_{R_q} \text{ is injective.} \tag{12}$$

Suppose that $x_1 \in R_q \backslash \{0\}$ satisfies $(\lambda I - T)x_1 = 0$. By virtue of (11), we may choose $x_2 \in R_q \backslash \{0\}$ such that $(\lambda I - T)x_2 = x_1$. Proceedingly inductively, we can construct a sequence $\{x_n\}_{n=1}^{\infty}$ satisfying

$$(\lambda I - T)^{n-1}x_n = x_1 \neq 0, \quad (\lambda I - T)^n x_n = (\lambda I - T)x_1 = 0 \quad \text{for all } n \in \mathbb{N}.$$

In other words, we have $x_n \in N_n \backslash N_{n-1}$ for all $n \in \mathbb{N}$. This contradicts Lemma 6.10 and we conclude that (12) holds.

We now prove that $N_{q+1} = N_q$. We already know that $N_q \subset N_{q+1}$. Suppose that $N_q \neq N_{q+1}$. Then we may choose $x \in N_{q+1} \backslash N_q$. Put $y = (\lambda I - T)^q x \in R_q$ and notice that $y \neq 0$. However

$$(\lambda I - T)y = (\lambda I - T)^{q+1}x = 0,$$

which contradicts (12). We conclude that $N_q = N_{q+1}$ and consequently $q \geq p$.

To establish the reverse inequality, we shall show that $R_{p+1} = R_p$. For this purpose it is convenient to show that

$$N_q + \mathcal{R}(\lambda I - T) = X. \tag{13}$$

To establish (13), let $x \in X$ be given. Since $R_q = R_{q+1}$, we may choose $y \in X$ such that

$$(\lambda I - T)^q x = (\lambda I - T)^{q+1} y.$$

Let us put

$$x_1 = x - (\lambda I - T)y, \quad x_2 = (\lambda I - T)y.$$

It is immediate that $x_2 \in \mathcal{R}(\lambda I - T)$. Since

$$(\lambda I - T)^q x_1 = (\lambda I - T)^q x - (\lambda I - T)^{q+1} y = 0,$$

we see that $x_1 \in N_q$ and (13) is established.

Since $N_q \subset N_p$, it follows from (13) that

$$N_p + \mathcal{R}(\lambda I - T) = X. \tag{14}$$

We already know that $R_{p+1} \subset R_p$. To establish the reverse inclusion, let $x \in R_p$ be given. Then we may choose $y \in X$ such that

$$x = (\lambda I - T)^p y.$$

By (14) we may choose $y_1 \in N_p$ and $y_2 \in \mathcal{R}(\lambda I - T)$ such that $y = y_1 + y_2$. Now, we may choose $y_3 \in X$ such that $y_2 = (\lambda I - T)y_3$. It follows that

$$x = (\lambda I - T)^p y_1 + (\lambda I - T)^{p+1} y_3 = (\lambda I - T)^{p+1} y_3.$$

We conclude that $x \in R_{p+1}$. This implies that $R_p = R_{p+1}$ which implies that $p \geq q$ and the proof is complete. $\square$

Proposition 6.13: Let X be a Banach space and $T \in \mathcal{C}(X;X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$. be given. Assume that $\lambda \in \sigma(T)$. Then $\lambda \in \sigma_p(T)$.

Proof: Suppose that $\lambda \notin \sigma_p(T)$. Then $\mathcal{N}(\lambda I - T) = \{0\}$, so $p_*(\lambda;T) = 0$. By Lemma 6.12, we also have $q_*(\lambda;T) = 0$ which implies that $\mathcal{R}(\lambda I - T) = X$ and consequently $\lambda I - T$ is bijective, which contradicts the assumption that $\lambda \in \sigma(T)$. $\square$

Theorem 6.14: Let X be a Banach space and $T \in \mathcal{C}(X;X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Let $p = p_*(\lambda;T)$ where $p_*(\lambda;T)$ is as in Lemma 6.10. Then

$$\mathcal{N}((\lambda I - T)^p) \oplus \mathcal{R}((\lambda I - T)^p) = X.$$

Proof: Let N_n and R_n be as in the proof of Lemma 6.12 and observe that $R_{2p} = R_p$. Let $x \in X$ be given. Then we may choose $z \in X$ such that $(\lambda I - T)^{2p} z = (\lambda I - T)^p x$. Now put

$$y = (\lambda I - T)^p z \in R_p,$$

and observe that

$$(\lambda I - T)^p y = (\lambda I - T)^{2p} z = (\lambda I - T)^p x.$$

It follows that $x - y \in N_p$ and we have

$$x = (x - y) + y.$$

To see that the decomposition is unique, let $\tilde{y} \in R_p$ be given with $x - \tilde{y} \in N_p$. Put

$$z = y - \tilde{y} \in R_p.$$

We want to show that $z = 0$. We may choose $w \in X$ such that

$$z = (\lambda I - T)^p w.$$

Observe that

$$z = (x - \tilde{y}) - (x - y) \in N_p,$$

and consequently

$$(\lambda I - T)^{2p} w = (\lambda I - T)^p z = 0.$$

Since $N_{2p} = N_p$, we find that $w \in N_p$, and this gives

$$0 = (\lambda I - T)^p w = z. \quad \square$$

Theorem 6.15: Let X be a Banach space and let $T \in \mathcal{C}(X;X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then

$$\dim(\mathcal{N}(\lambda I - T)) = \dim(\mathcal{N}(\lambda I - T^*)).$$

To establish Theorem 6.15, we shall make use of the following lemma, whose proof is left as an exercise.

Lemma 6.16: Let X be a Banach space and let $\{x_1^*, x_2^*, \cdots, x_n^*\}$ be a linearly independent subset of X^*. Then there exist $x_1, x_2, \cdots, x_n \in X$ such that for all $i, j \in \{1, 2, \cdots, n\}$ we have

$$x_i^*(x_j) = \begin{cases} 1 \text{ if } i = j, \\ 0 \text{ if } i \neq j. \end{cases}$$

To prove Theorem 6.15, it is convenient to show that

$$\dim(\mathcal{N}(\lambda I - T)) \geq \dim(\mathcal{N}(\lambda - T^*)) \tag{15}$$

and then look at $\dim(\mathcal{N}(\lambda I - T^{**}))$.

Lemma 6.17: Let X be a Banach space and $T \in \mathcal{C}(X, X)$ and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then (15) holds.

Proof: We know that both of the null spaces in question have finite dimension. Let

$$n = \dim(\mathcal{N}(\lambda I - T)), \quad m = \dim(\mathcal{N}(\lambda I - T^*)).$$

Let $\{x_1, x_2, \cdots, x_n\}$ be a basis for $\mathcal{N}(\lambda I - T)$ and $\{y_1^*, y_2^*, \cdots, y_m^*\}$ be a basis for $\mathcal{N}(\lambda I - T^*)$. By a straightforward application of the Hahn-Banach Theorem we may choose $x_1^*, x_2^*, \cdots, x_n^* \in X^*$ such that

$$x_i^*(x_j) = \begin{cases} 1 \text{ if } i = j, \\ 0 \text{ if } i \neq j. \end{cases}$$

Using Lemma 6.16, we may choose $y_1, y_2, \cdots, y_m \in X$ such that

$$y_i^*(y_j) = \begin{cases} 1 \text{ if } i = j, \\ 0 \text{ if } i \neq j. \end{cases}$$

Suppose that $n < m$ and define $L \in \mathcal{L}(X;X)$ by

$$Lx = \sum_{i=1}^{n} x_i^*(x) y_i.$$

Observe that $L \in \mathcal{C}(X;X)$ since L has finite rank. Now put

$$S = T + L,$$

and observe that S is also compact.

We shall show that $\mathcal{N}(\lambda I - S) = \{0\}$. To this end, let $x \in \mathcal{N}(\lambda I - S)$ be given. Then $Sx = \lambda x$ so that

$$(\lambda I - T)x = Lx.$$

It follows that for all $j \in \{1, 2, \cdots, n\}$ we have

$$0 = (\lambda y_j^* - T^* y_j^*)(x),$$

and consequently $x_j^*(x) = 0$ for all $j \in \{1, 2, \cdots, m\}$. It follows that $(\lambda I - T)x = 0$ and therefore we may choose $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{K}$ such that

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n. \tag{16}$$

For each $j \in \{1, 2, \cdots, n\}$, we apply x_j^* to (16) to conclude that $x_j^*(x) = \alpha_j$. It follows that $x = 0$ and consequently $\mathcal{N}(\lambda I - S) = \{0\}$.

Since $\lambda \neq 0$, S is compact and $\lambda \notin \sigma_p(S)$, it follows from Proposition 6.13 that $\lambda \in \rho(T)$ and consequently that $\mathcal{R}(\lambda I - S) = X$. Thus we may choose $z \in X$ such that

$$(\lambda I - S)z = y_{n+1}.$$

Then by the choice of $y_1, y_2, \cdots, y_m$ we see that

$$1 = y_{n+1}^*(y_{n+1}) = y_{n+1}^*(x) = y_{n+1}^* \left((\lambda I - T)x - \sum_{i=1}^{n} x_i^*(x) y_i \right) = 0.$$

This is, of course, a contradiction and consequently it is not possible to have $n < m$. □

Proof of Theorem 6.15: By Lemma 6.16 we have

$$\dim(\mathcal{N}(\lambda I - T) \geq \dim(\mathcal{N}(\lambda I - T^*)).$$

Applying Lemma 6.16 to T^* we see that

$$\dim(\mathcal{N}(\lambda I - T^*) \geq \dim(\mathcal{N}(\lambda I - T^{**})).$$

On the other hand, since $\lambda I - T^{**}$ is an extension of $\lambda I - T$ we have

$$\dim(\mathcal{N}(\lambda I - T^{**}) \geq \dim(\mathcal{N}(\lambda I - T)),$$

and we are done. $\square$

The Fredholm Alternative

There is an important principle known as the *Fredholm Alternative* which expresses some of the main conclusions concerning spectral theory of compact operators in terms of solutions of equations of the form $(\lambda I - T)x = y$.

Theorem 6.18 (The Fredholm Alternative): Let X be a Banach space and let $T \in \mathcal{C}(X;X)$, $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then exactly one of (a) or (b) below must hold:

(a) For every $y \in X$ the equation $(\lambda I - T)x = y$ has exactly one solution $x \in X$. has only the trivial solution $x = 0$.

(b) There exists $x \in X\backslash\{0\}$ such that $(\lambda I - T)x = 0$

Moreover, if (a) holds then for every $y^* \in X^*$ the equation $(\lambda I - T^*)x^* = y^*$ has exactly one solution $x^* \in X^*$. If (b) holds then (i), (ii), and (iii) below hold.

(i) The number of linearly independent solutions of $(\lambda I - T)x = 0$ in X is finite and is the same as the number of linearly independent solutions of $(\lambda I - T^*)x^* = 0$ in X^*.

(ii) For a given $y \in X$, the equation $(\lambda I - T)x = y$ has a solution $x \in X$ if and only if $y \in {}^{\perp}\mathcal{N}(\lambda I - T)$.

(iii) For a given $y^* \in X^*$, the equation $(\lambda I - T^*)x^* = y^*$ has a solution $x^* \in X^*$ if and only if $y^* \in \mathcal{N}(\lambda I - T)^{\perp}$.

Fredholm Operators

TO BE FILLED IN

General Linear Operators Between Normed Linear Spaces

We now turn our attention to linear operators that need not be bounded.

Let X and Y be normed linear spaces over $\mathbb{K}$.

Definition 6.19: Let $\mathcal{D}(A) \subset X$. When we say that $A : \mathcal{D}(A) \to Y$ is a linear operator we mean that $\mathcal{D}(A)$ is a linear manifold and

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay \;\text{ for all } x, y \in \mathcal{D}(A),\; \alpha, \beta \in \mathbb{K}.$$

When dealing with a linear operator A whose domain could be properly included in X we shall write $A : \mathcal{D}(A) \to Y$ rather than try to introduce special terminology to indicaate that the domain might be properly included in X because there does not seem to be any universally accepted terminology for this and it is more important to avoid potential confusion rather than to have a more elegant way of phrasing results.

Definition 6.20: Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear. We say that

(i) A is *bounded* provided there exists $K \in \mathbb{R}$ such that

$$\|Ax\|_Y \leq K\|x\|_X \text{ forall } x \in X.$$

(ii) A is *unbounded* if it is not bounded.

(iii) A is closed provided that $\mathrm{Gr}(A)$ is closed in $X \times Y$.

Theorem 6.21 (Closed Graph Theorem): Assume that X and Y are Banach spaces that $\mathcal{D}(A) \subset X$ and that $A : \mathcal{D}(A) \to Y$ is linear. Assume further that A is closed and that $\mathcal{D}(A)$ is closed. Then A is bounded.

Remark 6.22: Let X and Y be normed linear spaces, $\mathcal{D}(A) \subset X$, and assume that $A : \mathcal{D}(A) \to Y$ is linear. Then A is closed if and only if for every $x \in X$, $y \in Y$ and every sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathcal{D}(A)$ such that $x_n \to x$ and $Ax_n \to y$ as $n \to \infty$ we have $x \in \mathcal{D}(A)$ and $Ax = y$.

Example 6.23 Let $X = Y = C[0, 1]$ equipped with the supremum norm. Let us put $\mathcal{D}(A) = C^1[0, 1]$, $\mathcal{D}(B) = C^{\infty}[0, 1]$, and define

$$Au = u' \text{ for all } u \in \mathcal{D}(A), \quad Bu = u' \text{ for all } u \in \mathcal{D}(B).$$

Notice that A is an extension of B. Using Remark we see that A is closed, but B is not. [Indeed if $\{u_n\}_{n=1}^{\infty}$ is a sequence in $C^1[0, 1]$ and $u_n \to u$ uniformly and $u_n' \to v$ uniformly as $n \to \infty$ then $v \in C^1[0, 1]$ and $u' = v$. To see that B is not closed, let any function $w \in C[0, 1] \backslash C^1[0, 1]$ be given. Then we may choose a sequence $\{w_n\}_{n=1}^{\infty}$ of polynomials such that $w_n \to w$ uniformly as $n \to \infty$. Then we can define

$$u_n(x) = \int_0^x w_n(t)\, dt \text{ for all } n \in \mathbb{N},\ x \in [0, 1.$$

Then we have $u_n \to u$ uniformly and $u_n' \to w$ uniformly as $n \to \infty$, but $w \notin \mathcal{D}(B)$.]

Lecture Notes for Week 7 (First Draft)

General Linear Operators (Continued)

Proposition 7.1: Let X and Y be normed linear spaces. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear. Assume further that $\mathcal{D}(A)$ is closed and that A is bounded. Then A is closed.

Proof: Since A is bounded, we may choose $K \in \mathbb{R}$ such that $\|Ax\| \leq K\|x\|$ for all $x \in X$. Let $x \in X$, $y \in Y$, and a sequence $\{x_n\}_{n=1}^{\infty}$ be given such that $x_n \in \mathcal{D}(A)$ for all $n \in \mathbb{N}$. Assume that $x_n \to x$ and $Ax_n \to y$ as $n \to \infty$. Since $\mathcal{D}(A)$ is closed, we know that $x \in \mathcal{D}(A)$. We need to show that $y = Ax$. Since

$$\|Ax_n - Ax\| \leq K\|x_n - x\| \quad \text{for all } n \in \mathbb{N},$$

we see that $Ax_n \to Ax$ as $n \to \infty$. By uniqueness of limits, we have $Ax_n \to y$ as $n \to \infty$ and A is closed. $\square$

Proposition 7.2: Let X and Y be normed linear spaces. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear. Then $\mathcal{N}(A)$ is closed.

Proof: Let $x \in X$ be given and let $\{x_n\}_{n=1}^{\infty}$ be a sequence such that $x_n \in \mathcal{N}(A)$ for all $n \in \mathbb{N}$ and $x_n \to x$ as $n \to \infty$. Since $x_n \in \mathcal{N}(A)$ for all $n \in \mathbb{N}$ we have $Ax_n \to 0$ as $n \to \infty$. Since A is closed, we have $x \in \mathcal{D}(A)$ and $Ax = 0$, i.e. $x \in \mathcal{N}(A)$ and $\mathcal{N}(A)$ is closed. $\square$

Proposition 7.3: Let X and Y be linear spaces. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear, closed, and injective. Then $A^{-1} : \mathcal{R}(A) \to X$ is closed.

Proof: Let $y \in Y$, $x \in X$ and a sequence $\{y_n\}_{n=1}^{\infty}$ be given such that $y_n \in \mathcal{R}(A)$ for all $n \in \mathbb{N}$. Assume that $y_n \to y$ and $A^{-1}y_n \to x$ as $n \to \infty$.

Notice that $A^{-1}y_n \in \mathcal{D}(A)$ for all $n \in \mathbb{N}$ and $A(A^{-1}y_n) \to y$ as $n \to \infty$. Since A is closed, we conclude that $x \in \mathcal{D}(A)$ and $Ax = y$. This implies that $y \in \mathcal{R}(A)$ and $A^{-1}y = x$ and consequently A^{-1} is closed. $\square$

Definition 7.4: Let X and Y be normed linear spaces. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear. A is said to be *closable* provided that there exists a linear manifold $\mathcal{D}(\tilde{A}) \subset X$ and a closed linear operator linear operator $\tilde{A} : \mathcal{D}(\tilde{A}) \to Y$ such that $\mathrm{Gr}(A) \subset \mathrm{Gr}(\tilde{A})$, i.e. $\tilde{A}$ is a closed linear extension of A.

Remark 7.5: Let X and Y be normed linear spaces. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear. It is straightforward to show that A is closable if and only if

$$\forall y \in Y \backslash \{0\}, \quad \text{we have } (0, y) \notin \mathrm{cl}(\mathrm{Gr}(A)).$$

Moreover, if A is closable then the minimal closed linear extension of A is called the *closure* of A and is denoted $\overline{A}$.

Adjoints

We have seen that for an (everywhere-defined) bounded linear operator, it is frequently useful to work with the adjoint operator. Let $A : \mathcal{D}(A) \to Y$ be a linear operator with $\mathcal{D}(A) \subset X$, where X and Y are normed linear spaces. We want a linear operator $A^* : \mathcal{D}(A^*) \to X^*$ with $\mathcal{D}(A^*) \subset Y^*$ such that

$$(A^*(y^*))(x) = y^*(Ax) \ \text{ for all } y^* \in \mathcal{D}(A^*),\ x \in \mathcal{D}(A). \tag{1}$$

In order to ensure that A^*y^* is uniquely determined by (1) we need to have $\mathcal{D}(A)$ dense in X. The following simple remark will be employed in the definition of adjoint.

Remark 7.6: Let X be a normed linear space and $S \subset X$ be a dense linear manifold. Assume that $l : S \to \mathbb{K}$ is linear and bounded. Then there exists exactly one $x^* \in X^*$ such that $l(x) = x^*(x)$ for all $x \in S$.

Definition 7.7: Let X and Y be normed linear spaces. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear. Assume further that $\mathcal{D}(A)$ is dense in X. The *adjoint* of A, denoted A^* is the linear operator defined by

$$\mathcal{D}(A^*) = \{y^* \in Y^* : y^*A : \mathcal{D}(A) \to \mathbb{K} \text{ is bounded}\} \tag{2}$$

$$A^*y^* = z^*_{y^*} \ \text{ for all } y^* \in \mathcal{D}(A^*), \tag{3}$$

where for each $y^* \in \mathcal{D}(A^*)$, $z^*_{y^*}$ is the unique element of X^* such that

$$y^*(x) = z^*_{y^*}(x) \ \text{ for all } x \in \mathcal{D}(A). \tag{4}$$

Warning: Even though we have assumed that $\mathcal{D}(A)$ is dense in Definition 7.7, it can happen that $\mathcal{D}(A^*) = \{0\}$.

Definition 7.8: Let X be a Hilbert space. Let $\mathcal{D}(\mathcal{A}) \subset X$ and assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to X$ is linear. The *Hilbert adjoint* A^*_H of A is the linear operator defined by

$$\mathcal{D}(A^*) = \{y \in X : \exists \text{ exactly one } z \in X,\ (Ax, y) = (x, z) \text{ for all } x \in \mathcal{D}(A)\},$$

$$(Ax, y) = (x, A^*y) \ \text{ for all } x \in \mathcal{D}(A),\ y \in \mathcal{D}(A^*).$$

Example 7.9: Let $p \in [1, \infty)$ be given and choose $q \in (1, \infty]$ such that $p^{-1} + q^{-1} = 1$. Put $X = Y = l^p$ and identify X^* and Y^* with l^q in the usual way. Put $\mathcal{D}(A) = \mathbb{K}^{(\mathbb{N})}$ and

$$Ax = \left(\sum_{k=1}^{\infty} kx_k, x_2, x_3, x_4, \cdots\right) \ \text{ for all } x \in \mathcal{D}(A).$$

In order to determine the adjoint of A we need to find $z \in l^q$ satisfying

$$y_1 \sum_{k=1}^{\infty} k x_k + \sum_{k=2}^{\infty} x_k y_k = z_1 x_1 + \sum_{k=2}^{\infty} z_k x_k \quad \text{for all } x \in \mathbb{K}^{(\mathbb{N})}. \tag{5}$$

Substituting $x = e^{(n)}$ (the vector with 1 in the n^{th} slot and zeros elsewhere) we arrive at

$$y_1 = z_1 \quad \text{using } e^{(1)},$$

$$n y_1 + y_n = z_n \quad \text{using } e^{(n)} \text{ for } n \geq 2.$$

Since the sequences $\{y_n\}_{n=1}^{\infty}$ and $\{z_n\}_{n=1}^{\infty}$ must be bounded we must have $y_1 = 0$ and we must also have $z_k = y_k$ for all $k \geq 2$. We conclude that

$$\mathcal{D}(A^*) = \{y \in l^q : y_1 = 0\},$$

$$A^* y^* = (0, y_2, y_3, y_4, \cdots) \quad \text{for all } y \in \mathcal{D}(A^*).$$

Notice that $\mathcal{D}(A^*)$ fails to be dense in Y^*.

Example 7.10: Let $\mathbb{K} = \mathbb{R}$ and $X = Y = L^2[0,1]$. We identify X^* and Y^* with $L^2[0,1]$ in the usual way. Consider the linear operator A defined by

$$\mathcal{D}(A) = \{f \in AC[0,1] : f' \in L^2[0,1],\ f(0) = f(1) = 0\},$$

$$Af = f' \quad \text{for all } f \in \mathcal{D}(A).$$

Here $AC[0,1]$ is the set of all absolutely continuous functions $f : [0,1] \to \mathbb{R}$ and f' is the derivative of f. The domain of A is the same as the Sobolev space $W_0^{1,2}(0,1)$ which is also frequently denoted by $H_0^1(0,1)$. Notice that $\mathcal{D}(A)$ is dense. We want to determine the adjoint of A. Let $g \in L^2[0,1]$ be given. In order to have $g \in \mathcal{D}(A^*)$ it is necessary and sufficient for there exist to $h \in L^2[0,1]$ such that

$$\int_0^1 f' g = \int_0^1 f h \quad \text{for all } f \in \mathcal{D}(A). \tag{6}$$

Notice that if $g \in \mathcal{D}(A)$, then $g' \in L^2[0,1]$ and using integration by parts we have

$$\int_0^1 f' g = -\int_0^1 f g' \quad \text{for all } f \in \mathcal{D}(A). \tag{7}$$

It follows that $\mathcal{D}(A) \subset \mathcal{D}(A^*)$. The condition $g(0) = g(1) = 0$ played no role in the derivation of (7). In fact if $g \in AC[0,1]$ and $g' \in L^2[0,1]$ then (7) holds. It can be shown (without too much difficulty) that if there exists $h \in L^2[0,1]$ such that (6) holds then $g \in AC[0,1]$ and $h = -g'$. Consequently

$$\mathcal{D}(A^*) = \{g \in AC[0,1] : g' \in L^2[0,1]\},$$

$$A^* g = -g' \quad \text{for all } g \in \mathcal{D}(A^*).$$

The domain of A^* is the same as the Soloblev space $W^{1,2}(0,1)$ which is also frequently denoted $H^1(0,1)$.

For differential operators, boundary conditions play crucial roles in the specification of domains of operators and their adjoints. One can show that with A as above we have $(A^*)^* = A$ (with equality of domains, i.e. $\mathcal{D}(A^{**}) = \mathcal{D}(A)$.)

You should convince yourself as an exercise that if we define the linear operator $B : \mathcal{D}(B) \to Y$ by

$$\mathcal{D}(B) = \{f \in AC[0,1] : f' \in L^2[0,1],\ f(0) = 0\},$$

$$Bf = f' \text{ for all } f \in \mathcal{D}(B),$$

then

$$\mathcal{D}(B^*) = \{g \in AC[0,1] : g' \in L^2,\ g(1) = 0\},$$
$$B^*g = -g' \text{ for all } g \in \mathcal{D}(B^*).$$

An important feature of adjoints is that they are always closed operators.

Theorem 7.11: Let X and Y be normed linear spaces. Let $\mathcal{D}(A) \subset X$ and assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to y$ is linear. Then A^* is closed.

Proof: Let $x^* \in X^*$, $y^* \in Y^*$, and a sequence $\{y_n^*\}_{n=1}^\infty$ in $\mathcal{D}(A^*)$ be given. Assume that $y_n^* \to y^*$ and $A^*y_n^* \to x^*$ as $n \to \infty$. We need to show that $y^* \in \mathcal{D}(A^*)$ and $A^*y^* = x^*$.

Let $x \in \mathcal{D}(A)$ be given. Then we have

$$y_n^*(Ax) \to y^*(Ax) \text{ as } n \to \infty, \tag{8}$$

$$y_n^*(Ax) = (A^*y_n^*)(x) \to x^*(x). \tag{9}$$

It follows from (8), (9), and uniqueness of limits that

$$y^*(Ax) = x^*(x) \text{ for all } x \in \mathcal{D}(A). \tag{10}$$

We conclude from (10) that $y^* \in \mathcal{D}(A^*)$ and $A^*y^* = x^*$ and therefore A^*. It follows that A^* is closed. $\square$

Proposition 7.12: Let X and Y be normed linear spaces. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear. Assume further that $\mathcal{D}(A)$ is dense in X. Then $\mathcal{D}(A^*) = Y^*$ if and only if A is bounded. Moreover if $\mathcal{D}(A^*) = Y^*$ then $A^* \in \mathcal{L}(Y^*; X^*)$ and

$$\|A^*\| = \sup\{\|Ax\| : x \in X,\ \|x\| \leq 1\}.$$

Proof: Assume first that A is bounded. Then $y^*A : \mathcal{D}(A) \to \mathbb{K}$ is bounded for every $y^* \in Y^*$ and consequently $\mathcal{D}(A^*) = Y^*$.

Assume that $\mathcal{D}(A^*) = Y^*$. Then, by the Closed Graph Theorem, we have $A^* \in \mathcal{L}(Y^*; X^*)$. Put

$$\mathcal{B} = \{x \in \mathcal{D}(A) : \|x\| \leq 1\},$$

$$|||A||| = \sup\{\|Ax\| : x \in \mathcal{B}\} \leq \infty.$$

Let $y^* \in Y^*$ be given. Then we have

$$\sup\{|y^*(Ax)| : x \in \mathcal{B}\} = \sup\{|(A^*y^*)(x)| : x \in \mathcal{B}\} \leq \|A^*\|\|y^*\|.$$

I claim that

$$|||A||| < \infty.$$

To see why the claim is true, let J denote the canonical injection of Y into Y^{**} and put $\mathcal{S} = \{Ax : x \in \mathcal{B}\}$. For every $y^* \in Y^*$ we have

$$\sup\{|J(y)y^*| : y \in \mathcal{S}\} < \infty,$$

so the Principle of Uniform Boundedness tells us that

$$\sup\{\|J(y)\| : y \in \mathcal{S}\} < \infty.$$

We conclude that $J[\mathcal{S}]$ is bounded in Y^{**} and consequently $\mathcal{S}$ is bounded in Y.

For all $y^* \in Y^*$ we have

$$\begin{aligned}
\|A^*y^*\| &\leq \sup\{|(A^*y^*)(x)| : x \in X, \|x\| \leq 1\} \\
&\leq \sup\{|(A^*y^*)(x)| : x \in \mathcal{B}\} \quad \text{since } B \text{ is dense} \\
&\leq \sup\{|y^*(Ax)| : x \in \mathcal{B}\} \\
&\leq \|y^*\| \cdot |||A|||.
\end{aligned}$$

It follows that $\|A^*\| \leq |||A|||$.

To establish the reverse inequality, observe that for all $x \in \mathcal{D}(A)$ we have

$$\begin{aligned}
\|Ax\| &= \sup\{|y^*(Ax)| : y^* \in Y^*, \|y^*\| \leq 1\} \\
&= \sup\{|(A^*y^*)(x)| : y^* \in Y^*, \|y^*\| \leq 1\} \\
&\leq \|A^*\| \cdot \|x\|.
\end{aligned}$$

It follows that $|||A||| \leq \|A^*\|$. $\square$

Since adjoints are always closed, a natural way to look for a closed extension of an operator A is to try to construct the second adjoint A^{**}. This, of course, will require that $\mathcal{D}(A)$ is dense in X and that $\mathcal{D}(A^*)$ is dense in Y^*.

Let X be a normed linear space and Y be a reflexive Banach space. Assume that $\mathcal{D}(A)$ is dense in X and that $A : \mathcal{D}(A) \to Y$ is linear. Assume further that $\mathcal{D}(A^*)$ is dense in Y^*. Then it is relatively straightforward to show that the operator $\tilde{A}$ defined by

$$\mathcal{D}(\tilde{A}) = \{x \in X : J_X(x) \in \mathcal{D}(A^{**})\},$$

$$\tilde{A}x = (J_Y)^{-1}(A^{**}J_X(x)) \quad \text{for all } x \in \mathcal{D}(\tilde{A})$$

is a closed extension of A. It is in fact the minimal closed extension of A.

Proposition 7.13: Let X be a normed linear space and Y be a reflexive Banach space. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to Y$ is linear. Then A is closable if and only if $\mathcal{D}(A^*)$ is dense in Y^*. Moreover, in this case the minimal closed extension of A is given by

$$\overline{A} = (J_Y)^{-1}A^{**}J_X.$$

See Section II.2 of *Unbounded Linear Operators* by Seymour Goldberg for the details.

Spectral Theory for General Linear Operators

We now extend some ideas from spectral theory to general linear operators. We assume here that X is a Banach space.

Definition 7.14: Let X be a Banach space. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to X$ is linear. The *resolvent set* for A, denoted by $\rho(A)$ is the set of all $\lambda \in \mathbb{K}$ satisfying

(i) $\lambda I - A$ is injective,

(ii) $\mathcal{R}(\lambda I - A)$ is dense,

(iii) $(\lambda I - A)^{-1} : \mathcal{R}(\lambda I - A) \to X$ is bounded.

Definition 7.15: Let X be a Banach space. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to X$ is linear. The *spectrum* of A, denoted by $\sigma(A)$ is defined by

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

It is customary to partition the spectrum into several pairwise disjoint pieces. The most common way of doing this is contained in the next definition.

Definition 7.16: Let X be a Banach space. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to X$ is linear.

(a) The *point spectrum* of A, denoted $\sigma_p(A)$ is the set of all $\lambda \in \mathbb{K}$ such that (i) in Definition 7.14 fails to hold.

(b) The *continuous spectrum* of A, denoted $\sigma_c(A)$, is the set of all $\lambda \in \mathbb{K}$ such that (i) and (ii) of Definition 7.14 hold, but (iii) fails.

(c) The *residual spectrum* of A, denoted $\sigma_r(A)$, is the set of all $\lambda \in \mathbb{K}$ such that (i) in Definition 7.14 holds, but (ii) fails.

Remark 7.17: The sets $\sigma_p(A)$, $\sigma_c(A)$, and $\sigma_r(A)$ are pairwise disjoint and we have

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A).$$

The notion of continuous spectrum and residual spectrum are relevant even for compact linear operators as the following example shows.

Example 7.18: Let $X = l^2$ and define $A, B, C \in \mathcal{C}(X; X)$ by

(a) $Ax = \left(0, \frac{1}{2}x_2, \frac{1}{3}x_3, \frac{1}{4}x_4, \cdots\right)$ for all $x \in X$,

(b) $Bx = \left(x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \cdots\right)$ for all $x \in X$,

(c) $Cx = \left(0, x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \cdots\right)$ for all $x \in X$.

Observe that $0 \in \sigma_p(A)$, $0 \in \sigma_c(B)$, and $0 \in \sigma_r(C)$. Observe also that A, B, C all are compact, that A and B are self-adjoint, and that C fails to be normal. (In fact $C = RB$, where R is the right shift operator.)

it is not an accident that C is Example 7.18 is not normal. In fact, we have the following useful observation.

Remark 7.19 Let X be a Hilbert space and let $A \in \mathcal{L}(X; X)$ be given. Assume that A is normal. Let $\lambda \in \mathbb{K}$ be given. It follows from a homework problem that $\lambda I - A$ is injective if and only if $\mathcal{R}(\lambda I - A)$ is dense. This implies that $\sigma_r(A) = \emptyset$.

Remark 7.20: The elements of $\sigma_p(A)$ are called *eigenvalues* and for $\lambda \in \sigma_p(A)$, the nonzero elements of $\mathcal{N}(\lambda I - A)$ are called *eigenvectors* associated with λ.

Definition 7.21: Let X be a Banach space, $\mathcal{D}(A) \subset X$, and assume that $A : \mathcal{D}(A) \to X$ is linear. The approximate point spectrum of A, denoted $\sigma_{ap}(A)$ is the set of all $\lambda \in \mathbb{K}$ such that

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

The elements of $\sigma_{ap}(A)$ are called *generalized eigenvalues.*

An important feature of closed linear operators A is that if $\lambda \in \rho(A)$ then $\mathcal{R}(\lambda I - A) = X$.

Proposition 7.22: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D} \to X$ is linear and closed and let $\lambda \in \rho(A)$ be given. Then $(\lambda I - A)^{-1} \in \mathcal{L}(X; X)$.

Proof: Since A is closed, we see that $\lambda I - A$ is closed and consequently $(\lambda I - A)^{-1}$ is closed. Since $(\lambda I - A)^{-1}$ is bounded, we must have $\mathcal{D}((\lambda I - A)^{-1}) = \mathcal{R}(\lambda I - A)$ is closed. Since $\mathcal{R}(\lambda I - A)$ is dense, we conclude that $\mathcal{R}(\lambda I - A) = X$. $\square$

For unbounded closed operators with dense domains, it can happen that the resolvent set is empty or (even if $\mathbb{K} = \mathbb{C}$) that the spectrum is empty (or that neither of these sets is empty).

Example 7.23: Let $\mathbb{K} = \mathbb{C}$ and $X = L^2[0,1]$. We want to find closed, densely defined linear operators $A : \mathcal{D}(A) :\to X$, $B : \mathcal{D}(B) \to X$ such that $\rho(A) = \emptyset$ and $\sigma(B) = \emptyset$. The key will be to look at the differential equation

$$f' - \lambda f = g.$$

We shall insist that

$$\mathcal{D}(A), \mathcal{D}(B) \subset \{f \in AC[0,1] : f' \in L^2[0,1]\}.$$

For $\lambda \in \mathbb{C}$, put

$$u_\lambda(x) = e^{\lambda x},$$

and observe that

$$u_\lambda' = \lambda u_\lambda,$$

so that

$$\lambda u_\lambda - u_\lambda' = 0 \quad \text{for all} \lambda \in \mathbb{C}.$$

We put

$$\mathcal{D}(A) = \{f \in AC[0,1] : f' \in L^2[0,1]\},$$

and define $A : \mathcal{D}(A) \to X$ by

$$Au = u' \quad \text{for all } u \in \mathcal{D}(A).$$

It is immediate that $\sigma(A) = \sigma_p(A) = \mathbb{C}$.

Given $g \in L^2[0,1]$ and $\lambda \in \mathbb{C}$, put

$$u_{\lambda,g}(x) = -\int_0^x e^{\lambda(x-t)} g(t)\, dt \quad \text{for all } x \in [0,1].$$

Then we see that $u_{\lambda,g}(0) = 0$ and

$$(u_{\lambda,g})'(x) = -g(x) + \lambda u_{\lambda,g}(x) \quad \text{for a.e. } x \in [0,1].$$

It is straightforward to check that

$$\|u_{\lambda,g}\|_X \leq \|g\|_X \quad \text{for all} \lambda \in \mathbb{C}.$$

We put

$$\mathcal{D}(B) = \{f \in AC[0,1] : f' \in L^2[0,1], \; f(0) = 0\},$$

and define $B : \mathcal{D}(B) \to X$ by

$$Bu = u' \quad \text{for all } u \in \mathcal{D}(B).$$

Then $\mathcal{D}(B)$ is dense, B is closed, and $\rho(B) = \mathbb{C}$.

The Resolvent Operator

Definition 7.24: Let X be a Banach space and $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to X$ is linear and closed. For $\lambda \in \rho(A)$ the *resolvent operator* of A at λ is defined by

$$R(\lambda; A) = (\lambda I - A)^{-1} \in \mathcal{L}(X; X).$$

Notice that $R(\lambda; A) : X \to \mathcal{D}(A)$ and for all $x \in \mathcal{D}(A)$ we have

$$x = (\lambda I - A)R(\lambda; A)x = R(\lambda; A)(\lambda I - A)x,$$

and consequently

$$AR(\lambda; A)x = R(\lambda; A)Ax.$$

Making minor modifications in the proofs of Propositions 5.6 and 5.7, we easily establish the following generaliztions for closed linear operators.

Proposition 7.25: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A) \to X$ is linear and closed. Then

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

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

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

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

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

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

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

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

Proposition 7.26: Let X be a Banach space and let $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A) \to X$ is linear and closed. Let $\mu, \lambda \in \rho(A)$ be given. Then we have

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

(ii) $R(\lambda; A) R(\mu; A) = R(\mu; A) R(\lambda; A)$.

Lecture Notes for Week 8 (First Draft)

Semigroups of Linear Operators

Many linear evolution problems can be cast in the form

$$\dot{u}(t) = Au(t),\ t \geq 0,\ u(0) = u_0, \tag{IVP}$$

where X is a Banach space, $\mathcal{D}(A) \subset X$, $A : \mathcal{D}(A) \to X$ is a linear operator and $u_0 \in X$ is given. We "know" that the solution to (IVP) should be given by

$$u(t) = e^{tA}u_0,\ t \geq 0.$$

The crucial questions are: "What do we mean by e^{tA} if A is unbounded?"; and "How can we "construct" e^{tA} if A is unbounded?". To address these questions, let us recall some approaches to constructing e^{tA} when A is an $N \times N$ matrix.

Matrix Exponentials

Let A be an $N \times N$ (real or complex) matrix. The following standard methods can be used to produce e^{tA}.

I. *Series*: We can represent e^{tA} as a power series:

$$e^{tA} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}. \tag{1}$$

This construction will work fine in an infinite-dimensional Banach space X provided that $A \in \mathcal{L}(X; X)$. However, it seems doomed to failure if A is unbounded.

II. *Explicit Euler Scheme*: Let $n \in \mathbb{N}$ and $t > 0$ be given. We can use the explicit Euler method to approximate $u(t)$ where u is the solution to (IVP). For $k \in \{0, 1, 2, \cdots, n-1\}$, we have

$$\begin{aligned} u\left(\frac{(k+1)t}{n}\right) &\approx u\left(\frac{kt}{n}\right) + \frac{t}{n}Au\left(\frac{kt}{n}\right) \\ &\approx \left(I + \frac{t}{n}A\right)u\left(\frac{kt}{n}\right). \end{aligned}$$

If we start with $u(0) = u_0$ and iterate the scheme above we find that

$$u(t) \approx \left(I + \frac{tA}{n}\right)^n u_0.$$

This leads to the formula

$$e^{tA} = \lim_{n\to\infty} \left(I + \frac{tA}{n} \right)^n .$$

As before, this will work in an infinite-dimensional Banach space X if $A \in \mathcal{L}(X;X)$, but runs into serious difficulties if A is unbounded.

III. *Implicit Euler Method*: Let $t > 0$ and $n \in \mathbb{N}$. The difference between this approach and the explicit Euler method is that the derivative is approximated by using the right endpoint of each time interval rather than the left endpoint. For $k \in \{0, 1, 2, \cdots, n-1\}$ we have

$$u\left(\frac{(k+1)t}{n}\right) \approx u\left(\frac{kt}{n}\right) + \frac{t}{n} Au\left(\frac{(k+1)t}{n}\right),$$

which gives

$$\left(I - \frac{t}{n}A\right) u\left(\frac{(k+1)t}{n}\right) \approx u\left(\frac{kt}{n}\right).$$

Starting with $u(0) = u_0$ and iterating the formula above gives

$$u(t) \approx \left(I - \frac{t}{n}\right)^{-n} u_0.$$

This leads to the formula

$$e^{tA} = \lim_{n\to\infty} \left(I - \frac{t}{n}A\right)^{-n},$$

where

$$\begin{aligned}
\left(I - \frac{t}{n}A\right)^{-n} &= \left(\left(I - \frac{t}{n}A\right)^{-1}\right)^n \\
&= \left(\frac{n}{t}\right)^n R\left(\frac{n}{t}; A\right)^n .
\end{aligned}$$

This approach appears promising for unbounded operators A because it involves high powers of the resolvent operator, and the resolvent operator is an everywhere-defined bounded operator.

IV. *Laplace Transforms*: The Laplace transform is a powerful tool for studying (IVP). Recall that for a scalar-valued function $x : [0, \infty) \to \mathbb{K}$, the Laplace transform $\hat{x}$ of x is defined by

$$\hat{x}(\lambda) = \int_0^\infty e^{-\lambda t} x(t)\, dt.$$

In particular, for the scalar exponential function

$$y(t) = e^{at},$$

we have

$$\hat{y}(\lambda) = \frac{1}{\lambda - a}, \quad \mathrm{Re}(\lambda) > \mathrm{Re}(a).$$

For a matrix exponential

$$X(t) = e^{tA},$$

we have

$$\hat{X}(\lambda) = \int_0^\infty e^{-\lambda t} e^{tA}\, dt = (\lambda I - A)^{-1} = R(\lambda; A).$$

We can recover e^{tA} form its Laplace transform via the inversion formula

$$e^{tA} = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{\lambda t} R(\lambda; A)\, d\lambda.$$

Here γ is a sufficiently large positive number. Once again, this looks promising for unbounded operators A because it involves the resolvent operator.

V. *Spectral Decomposition*: Suppose that there is an orthonormal basis $(e^{(k)}|k = 1, 2, \cdots n)$ for N-space such each $e^{(k)}$ is an eigenvector for A, i.e.

$$Ae^{(k)} = \lambda_k e^{(k)}.$$

Then for every vector v we have

$$e^{tA} v = \sum_{k=1}^{n} e^{\lambda_k t} (v, e^{(k)}) e^{(k)}.$$

This approach will work for some unbounded operators of interest. In particular, if A is an unbounded self-adjoint operator on a Hilbert space X then A has a spectral decomposition

$$A = \int_{-\infty}^{\infty} dP(\lambda).$$

If A is bounded above, then we can define

$$e^{tA} = \int_{-\infty}^{\infty} e^{\lambda t} dP(\lambda), \quad t \geq 0.$$

Properties of Matrix Exponentials

Let A be an $N \times N$ matrix and put

$$S(t) = e^{tA},\ t \geq 0.$$

Then the mapping $S : [0, \infty) \to \mathbb{K}^{N \times N}$ satisfies

(i) $S(0) = I$.

(ii) $S(t+s) = S(t)S(s)$ for all $s, t \geq 0$,

(iii) $\lim_{t \downarrow 0} S(t) = I$.

Moreover, we can recover A from S via the formula

$$A = \lim_{h \downarrow 0} \frac{S(h) - I}{h}. \tag{2}$$

Linear C_0-Semigroups

In order to extend the notion of matrix exponential to general linear operators, we shall start with mappings $T : [0, \infty) \to \mathcal{L}(X; X)$ satisfying conditions analogous to (i) through (iii) above and construct A through a formula analogous to (2).

Definition 8.1: Let X be a Banach space. A *linear C_0-semigroup* is a mapping $T : [0, \infty) \to \mathcal{L}(X; X)$ satisfying (i) thru (iii) below.

(i) $T(0) = I$,

(ii) $T(t+s) = T(t)T(s)$ for all $s, t \in [0, \infty)$.

(iii) $\forall x \in X$, we have $\lim_{t \downarrow 0} T(t)x = x$.

Remark 8.2: It follows immediately from (ii) that

$$T(t)T(s) = T(s)T(t) \quad \text{for all } s, t \in [0, \infty).$$

Definition 8.3: Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup. The *infinitesimal generator* of A is the linear operator $A : \mathcal{D}(A) \to X$ defined by

$$\mathcal{D}(A) = \left\{ x \in X : \lim_{h \downarrow 0} \frac{T(h)x - x}{h} \text{ exists} \right\},$$

$$Ax = \lim_{h \downarrow 0} \frac{T(h)x - x}{h} \quad \text{for all } x \in \mathcal{D}(A).$$

If T is a linear C_0-semigroup with infinitesimal generator A then in some very broad sense we can "think of $T(t)$ as e^{tA}".

Remark 8.4: Let X be a Banach space and assume $T : [0, \infty) \to \mathcal{L}(X; X)$ satisfies (i) and (ii) of Definition 8.1.

(a) If
$$\lim_{t\downarrow 0} x^*(T(t)x) = x^*(x) \ \text{ for all } x \in X,\ x^* \in X^*,$$
then T is a linear C_0-semigroup. In other words, if (i) and (ii) of Definition 8.1 hold then T is continuous from the right at 0 in the weak operator topology if and only if it is continuous from the right at 0 in the strong operator topology. [This result will be a homework problem.] However, continuity from the right at 0 in the uniform operator topology implies that the infinitesimal generator is bounded.

(b) If
$$\lim_{t\downarrow 0} \|T(t) - I\| = 0$$
then there exists $B \in \mathcal{L}(X;X)$ such that
$$T(t) = \sum_{n=0}^{\infty} \frac{(tB)^n}{n!} \quad t \geq 0.$$
In this case B is the infinitesimal generator of T.

Example 8.5: Let $X = BUC(\mathbb{R})$, the space of all bounded uniformly continuous functions $f : \mathbb{R} \to \mathbb{K}$, equipped with the supremum norm, i.e.
$$\|f\| = \sup\{|f(x)| : x \in \mathbb{R}\}.$$
Consider the mapping $T : [0,\infty) \to \mathcal{L}(X;X)$ given by
$$(T(t)f)(x) = f(x+t) \ \text{ for all } u \in X,\ x \in \mathbb{R},\ t \geq 0.$$
It is straightforward to check that T is a linear C_0-semigroup. [Uniform continuity of the functions in X is essential for this.] It is called a *translation semigroup*. In order for a function $f \in X$ to be in the domain of the infinitesimal generator A, it is necessary (but not sufficient) that
$$\forall x \in \mathbb{R},\ \lim_{h\downarrow 0} \frac{f(x+h) - f(x)}{h} \ \text{ exists.}$$
Thus in order for a function to belong to $\mathcal{D}(A)$ it must be continuously right differentiable, and consequently it must be everywhere differentiable. Therefore we must have $Af = f'$ for all $f \in \mathcal{D}(A)$. It follows that
$$\mathcal{D}(A) = BUC^1(\mathbb{R}),$$
the set of all functions in $BUC(\mathbb{R})$ having first derivatives that belong to $BUC(\mathbb{R})$.

Let $u_0 \in \mathcal{D}(A)$ be given and define $u : [0,\infty) \times \mathbb{R} \to \mathbb{K}$ by
$$u(t,x) = (T(t)u_0)(x) = u_0(x+t) \ \text{ for all } t \geq 0,\ x \in \mathbb{R}.$$

Then u is a solution of an initial-value problem for a first-order partial differential equation:

$$u_t(t,x) = u_x(t,x), \quad t \geq 0, \; x \in \mathbb{R}; \quad u(0,x) = u_0(x), \quad x \in \mathbb{R}.$$

[Here u_t and u_x are the partial derivatives of u with respect to the first and second arguments.]

Lemma 8.6: Let X be a Banach space and assume that $T : [0,\infty) \to \mathcal{L}(X;X)$ is a linear C_0-semigroup. Then there exist $M, \omega > 0$ such that

$$\|T(t)\| \leq Me^{\omega t} \text{ for all } t \geq 0.$$

Proof: I claim that we may choose $\eta > 0$ such that

$$\sup\{\|T(t)\| : t \in [0,\eta]\} < \infty.$$

To validate the claim, suppose that no such η exists. Then we may choose a sequence $\{t_n\}_{n=1}^{\infty}$ such that $t_n > 0$ for all $n \in \mathbb{N}$, $t_n \to 0$ as $n \to \infty$, and $\{\|T(t_n)\|\}_{n=1}^{\infty}$ is unbounded. On the other hand

$$\forall x \in X, \; \lim_{n\to\infty} \|T(t_n)x\| \text{ exists},$$

so we have

$$\forall x \in X, \; \sup\{\|T(t_n)x\| : n \in \mathbb{N}\} < \infty.$$

The Principle of Uniform Boundedness implies that

$$\sup\{\|T(t_n)\| : n \in \mathbb{N}\} < \infty,$$

and this is a contradiction.

Put

$$M = \sup\{\|T(t)\| : t \in [0,\eta]\},$$

and

$$\omega = \frac{1}{\eta} \log M.$$

Let $t > 0$ be given. Then we may choose $n \in \mathbb{N} \cup \{0\}$ and $\alpha \in [0,\eta)$ such that

$$t = n\eta + \alpha.$$

Then we have

$$\begin{aligned} T(t) &= T(n\eta + \alpha) = T(n\eta)T(\alpha) \\ &= (T(\eta))^n T(\alpha). \end{aligned}$$

We see that

$$\|T(t)\| \leq \|T(\eta)\|^n \|T(\alpha)\| = M^{n+1}.$$

Since $\omega t \geq n \log M$, we have

$$e^{\omega t} \geq M^n,$$

and we conclude that

$$\|T(t)\| \leq Me^{\omega t}. \quad \square$$

Uniformly Bounded and Contraction Semigroups

Definition 8.7: Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup. We say that T is

(i) *uniformly bounded* if there exists $M \in \mathbb{R}$ such that $\|T(t)\| \leq M$ for all $t \geq 0$.

(ii) a *contraction semigroup* provided $\|T(t)\| \leq 1$ for all $t \geq 0$.

(iii) a *quasicontraction semigroup* provided there exists $\omega \in \mathbb{R}$ such that $\|T(t)\| \leq e^{\omega t}$ for all $t \geq 0$.

Let $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup and choose $M, \omega \in \mathbb{R}$ such that $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$. Put

$$S(t) = e^{-\omega t} T(t) \quad \text{for all } t \geq 0.$$

It is easy to see that S is a linear C_0-semigroup satisfying

$$\|S(t)\| \leq M \quad \text{for all } t \geq 0.$$

Let B be the infinitesimal generator of S. It is straightforward to check that $\mathcal{D}(B) = \mathcal{D}(A)$ and

$$Bx = -\omega x + Ax \quad \text{for all } x \in \mathcal{D}(B).$$

Moreover, we can construct an equivalent norm $|||\cdot|||$ on X such that S becomes a contraction semigroup. Indeed, if we put

$$|||x||| = \sup\{\|T(s)\| : s \in [0, \infty)\} \quad \text{for all } x \in X,$$

then

$$\|x\| \leq |||x||| \leq M\|x\| \quad \text{for all } x \in X,$$

and

$$|||S(t)x||| \leq |||x||| \quad \text{for all } x \in X,\ t \geq 0.$$

Observe that T becomes a quasicontraction semigroup under $|||\cdot|||$.

Remark 8.8: The equivalent norm $|||\cdot|||$ constructed above need not preserve important geometric properties of the original norm on X. In fact, if X is a Hilbert space and $T : [0, \infty) \to \mathcal{L}(X; X)$ is a linear C_0-semigroup, it might happen that there is no norm that is equivalent to the original one, satisfies the parallelogram law, and has the property that the semigroup is quasicontractive.

Lemma 8.9: Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup. Let $x \in X$ be given. Then the mapping $t \to T(t)x$ is continuous on $[0, \infty)$.

Proof: For right continuity, let $t \geq 0$ and $h > 0$ be given. Then $T(t+h)x = T(h)T(t)x$, so we have

$$\lim_{h \downarrow 0} T(t+h)x = \lim_{h \downarrow 0} T(h)(T(t)x) = T(t)x.$$

To establish left continuity, we first choose $M, \omega \in \mathbb{R}$ such that $\|T(s)\| \leq Me^{\omega s}$ for all $s \geq 0$. Let $t > 0$ and $h \in (0, t)$ be given. Then we have

$$\begin{aligned}
\|(T(t-h)x - T(t)x\| &= \|T(t-h)[I - T(h)]x\| \\
&\leq \|T(t-h)\| \cdot \|T(h)x - x\| \\
&\leq Me^{\omega(t-h)}\|T(h)x - x\| \\
&\to 0 \text{ as } h \downarrow 0. \quad \square
\end{aligned}$$

Lecture Notes for Week 9 (First Draft)

Infinitesimal Generators

Lemma 9.1: Let X be a Banach space and $T : [0,\infty) \to \mathcal{L}(X;X)$ be a linear C_0-semigroup with infinitesimal generator A. Let $x \in X$ be given.

(i) $\forall t \geq 0$, we have $\displaystyle\lim_{h\to 0}\frac{1}{h}\int_t^{t+h} T(s)x\,ds = T(t)x$ (right limit if $t = 0$).

(ii) $\forall t \geq 0$, we have $\displaystyle\int_0^t T(s)x\,ds \in \mathcal{D}(A)$ and $\displaystyle A\int_0^t T(s)x\,ds = T(t)x - x$.

Proof: Part (i) is a standard result from calculus.

If $t = 0$ then (ii) is immediate, so assume $t > 0$. For $h > 0$ we have

$$(T(h) - I)\int_0^t T(s)x\,ds = \int_0^t T(s+h)x\,ds - \int_0^t T(s)x\,ds. \tag{1}$$

Putting $\tau = s + h$ we see that

$$\int_0^t T(s+h)x\,ds = \int_h^{t+h} T(\tau)x\,d\tau. \tag{2}$$

Moreover, we have

$$\int_0^t T(s)x\,ds = \int_0^h T(s)x\,ds + \int_h^t T(s)x\,ds. \tag{3}$$

We also have

$$\int_h^{t+h} T(\tau)x\,d\tau - \int_h^t T(s)x\,ds = \int_t^{t+h} T(s)x\,ds. \tag{4}$$

Combining (1), (2), (3), and (4) we find that

$$\left(\frac{T(h) - I}{h}\right)\int_0^t T(s)x\,ds = \frac{1}{h}\int_t^{t+h} T(s)x\,ds - \frac{1}{h}\int_0^h T(s)x\,ds. \tag{5}$$

Using part (i), we can take the limit as $h \downarrow 0$ of the right side of (5) to conclude that

$$\lim_{h\downarrow 0}\left(\frac{T(h) - I}{h}\right)\int_0^t T(s)x\,ds = T(t)x - x. \quad \Box$$

Lemma 9.2: Let X be a Banach space and let $T : [0,\infty) \to \mathcal{L}(X;X)$ be a linear C_0-semigroup with infinitesimal generator A. Let $x \in \mathcal{D}(A)$ be given. Then

(a) $T(t)x \in \mathcal{D}(A)$ for all $t \geq 0$.

(b) $T(t)Ax = AT(t)x$ for all $t \geq 0$.

(c) Put $u(t) = T(t)x$ for all $t \geq 0$. Then $u \in C^1([0,\infty); X)$ and

$$\dot{u}(t) = Au(t), \quad \text{for all } t \geq 0.$$

Proof: Let $t \geq 0$ and $h > 0$ be given. Then we have

$$\begin{aligned} \left(\frac{T(h)-I}{h}\right) T(t)x &= T(t)\left(\frac{T(h)-I}{h}\right) x \\ &\to T(t)Ax \text{ as } h \downarrow 0. \end{aligned}$$

We conclude that $T(t)x \in \mathcal{D}(A)$ and $AT(t)x = T(t)Ax$.

This proves (a) and (b) and shows that u is right differentiable with

$$(D^+u)(t) = Au(t) \;\text{ for all } t \geq 0.$$

To establish left differentiability, let $t > 0$ and $h \in (0, t)$ be given. Then we have

$$\begin{aligned} \frac{u(t-h)-u(t)}{h} &= T(t-h)\left(\frac{I-T(h)}{h}\right) x \\ &\to T(t)(-Ax) \text{ as } h \downarrow 0. \end{aligned}$$

We see that

$$(D^-u)(t) = (D^+u)(t) = Au(t) = T(t)Ax.$$

Since the mapping $t \to T(t)Ax$ is continuous, we are done. $\square$

The following result is an immediate consequence of Lemma 9.2 and the fundamental theorem of calculus.

Lemma 9.3: Let X be a Banach space and $T : [0,\infty) \to \mathcal{L}(X;X)$ be a linear C_0-semigroup with infinitesimal generator A. Let $x \in \mathcal{D}(A)$ be given. Then

$$\forall \tau, t \geq 0, \quad T(t)x - T(\tau)x = \int_\tau^t AT(s)x\,ds = \int_\tau^t T(s)Ax\,ds.$$

Theorem 9.4: Let X be a Banach space and $T : [0,\infty) \to \mathcal{L}(X;X)$ be a linear C_0-semigroup with infinitesimal generator A. Then $\mathcal{D}(A)$ is dense in X and A is closed.

Proof: Let $x \in X$ be given. For $h > 0$ put

$$x_h = \frac{1}{h}\int_0^h T(s)x\,ds.$$

Then $x_h \in \mathcal{D}(A)$ for all $h > 0$ and $x_h \to x$ as $h \downarrow 0$ by Lemma 9.1.

To show that A is closed, let $x, y \in X$ and a sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathcal{D}(A)$ be given. Assume that $x_n \to x$ and $Ax_n \to y$ as $n \to \infty$. Let $h > 0$ be given. Then, by Lemma 9.3, we have

$$T(h)x_n - x_n = \int_0^h T(s)Ax_n\, ds \tag{6}$$

Letting $n \to \infty$ in (6) we obtain

$$T(h)x - x = \int_0^h T(s)y\, ds. \tag{7}$$

(We can pass to the limit under the integral because the integrand in (6) converges uniformly on $[0, h]$ to the integrand in (7). It follows immediately from (7) that

$$\frac{T(h)x - x}{h} = \frac{1}{h}\int_0^h T(s)y\, ds. \tag{8}$$

The right-hand side of (8) converges to y as $h \downarrow 0$. It follows that $x \in \mathcal{D}(A)$ and $Ax = y$ and consequently A is closed. $\square$

Lemma 9.5: Let X be a Banach space and $S, T : [0, \infty) \to \mathcal{L}(X; X)$ be linear C_0-semigroups having the same infinitesimal generator A. Then $S(t) = T(t)$ for all $t \geq 0$.

Proof: Let $x \in \mathcal{D}(A)$ and $t > 0$ be given. Define $u : [0, t] \to X$ by

$$u(s) = T(t-s)S(s)x \quad \text{for all } s \in [0, t].$$

Let $s \in [0, t]$ and $h \in \mathbb{R}\backslash\{0\}$ be given with $s + h \in [0, t]$. Then we have

$$\begin{aligned}
\frac{u(s+h) - u(s)}{h} &= \frac{1}{h}[T(t-s-h)S(s+h)x - T(t-s)S(s)x] \\
&= \frac{1}{h}[T(t-s-h)(S(s+h) - S(s))x + (T(t-s-h) - T(t-s))S(s)x] \\
&= T(t-s-h)\left[\frac{S(s+h)x - S(S)x}{h}\right]x + \left[\frac{T(t-s-h) - T(t-s)}{h}\right]S(s)x \\
&\to T(t-s)AS(s)x - T(t-s)AS(s)x = 0 \text{ as } h \to 0
\end{aligned}$$

It follows that u is constant on $[0, t]$; in particular

$$T(t)x = u(0) = u(t) = S(t)x.$$

Since $\mathcal{D}(A)$ is dense, we can conclude that $T(t)x = S(t)x$ for all $x \in X$. $\square$

Let $a \in \mathbb{K}$ and $n \in \mathbb{N}$ be given and define $f : [0, \infty) \to \mathbb{K}$ by

$$f(t) = t^{n-1}e^{at}, \quad \text{for all } t \geq 0. \tag{9}$$

The Laplace transform $\hat{f}$ of f is given by

$$\hat{f}(\lambda) = \int_0^\infty e^{-\lambda t} t^{n-1} e^{at} \, dt = \frac{(n-1)!}{(\lambda - a)^n} \quad \text{for} \lambda \in \mathbb{K} \text{ with } \mathrm{Re}(\lambda) > \mathrm{Re}(a). \tag{10}$$

Evaluation of the integral in (10) is discussed in most elementary textbooks on differential equations.

If A is an $N \times N$ (real or complex) matrix and we put

$$F(t) = t^{n-1}e^{tA}, \quad \text{for all } t \geq 0, \tag{11}$$

then we have

$$\hat{F}(\lambda) = \int_0^\infty e^{-\lambda t} t^{n-1} e^{tA} \, dt = (n-1)! R(\lambda; A)^n, \tag{12}$$

for $\mathrm{Re}(\lambda)$ sufficiently large. Here, $R(\lambda; A) = (\lambda I - A)^{-1}$.

It is natural to conjecture that if $T : [0, \infty) \to \mathcal{L}(X; X)$ is a linear C_0-semigroup and n is a positive integer then

$$R(\lambda; A)^n = \frac{1}{(n-1)!} \int_0^\infty e^{-\lambda t} t^{n-1} T(t) \, dt$$

for $\mathrm{Re}(\lambda)$ suitably large. This is, in fact, correct. We begin by providing a proof when $n = 1$.

Lemma 9.6: Let X be a Banach space. Let $M, \omega \in \mathbb{R}$ and $\lambda \in \mathbb{K}$ with $\mathrm{Re}(\lambda) > \omega$ be given. Assume that $T : [0, \infty) \to \mathcal{L}(X; X)$ is a linear C_0-semigroup satisfying $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$ and having infinitesimal generator A. Then $\lambda \in \rho(A)$ and

$$R(\lambda; A)x = \int_0^\infty e^{-\lambda t} T(t) \, dt \quad \text{for all } x \in X. \tag{13}$$

Proof: Define $\Phi(\lambda) \in \mathcal{L}(X; X)$ by

$$\Phi(\lambda)x = \int_0^\infty e^{-\lambda t} T(t)x \, dt \quad \text{for all } x \in X.$$

We need to show that $\lambda \in \rho(A)$ and $R(\lambda; A) = \Phi(\lambda)$. Let $x \in \mathcal{D}(A)$ be given.

Then we have

$$
\begin{aligned}
\Phi(\lambda)Ax &= \int_0^\infty e^{-\lambda t} T(t)Ax\,dt \\
&= \int_0^\infty e^{-\lambda t} \frac{d}{dt}(T(t)x)\,dt \quad \text{(integration by parts)} \\
&= -x + \lambda \int_0^\infty e^{-\lambda t} T(t)x\,dt \\
&= \lambda\Phi(\lambda)x - x.
\end{aligned}
$$

Now let $x \in X$ be given. We need to show that $\Phi(\lambda)x \in \mathcal{D}(A)$ and that

$$A\Phi(\lambda)x = \lambda\Phi(\lambda)x - x.$$

For this purpose, let $h > 0$ be given. Then we have

$$
\begin{aligned}
\left(\frac{T(h)-I}{h}\right)\Phi(\lambda)x &= \frac{1}{h}\int_0^\infty e^{-\lambda t}(T(t+h)x - T(t)x)\,dt \\
&= \frac{1}{h}\int_0^\infty e^{-\lambda t} T(t+h)x\,dt - \frac{1}{h}\int_0^\infty e^{-\lambda t} T(t)x\,dt \\
&= \frac{1}{h}\int_0^\infty e^{-\lambda(s-h)} T(s)x\,ds - \frac{1}{h}\int_0^\infty e^{-\lambda t} T(t)x\,dt \\
&= \frac{1}{h}\int_0^\infty e^{-\lambda(s-h)} T(s)x\,ds - \frac{1}{h}\int_0^h e^{-\lambda(s-h)} T(s)x\,ds \\
&\quad -\frac{1}{h}\int_0^\infty e^{-\lambda t} T(t)x\,dt \\
&= \int_0^\infty \left[\frac{e^{-\lambda(t-h)} - e^{-\lambda t}}{h}\right] T(t)x\,dt - \frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda(t-h)} T(t)x\,dt \\
&\to \lambda\Phi(\lambda)x - x \quad \text{as } h \downarrow 0.
\end{aligned}
$$

We conclude that $\Phi(\lambda)x \in \mathcal{D}(A)$ and $A\Phi(\lambda)x = \lambda\Phi(\lambda)x - x$. It follows that $\lambda \in \rho(A)$ and $R(\lambda; A) = \Phi(\lambda)$. $\square$

Lemma 9.7: Let $M, \omega \in \mathbb{R}$ and $\lambda \in \mathbb{K}$ with $\mathrm{Re}(\lambda) > \omega$ and $n \in \mathbb{N}$ be given. Assume that $T : [0, \infty) \to \mathcal{L}(X; X)$ is a linear C_0-semigroup satisfying $\|T(t)\| \le Me^{\omega t}$ for all $t \ge 0$ and having infinitesimal generator A. Then

$$R(\lambda; A)^n x = \frac{1}{(n-1)!}\int_0^\infty e^{-\lambda t} t^{n-1} T(t)\,dt \quad \text{for all } x \in X.$$

Proof: We know that the mapping $\mu \to R(\mu; A)$ is analytic and that

$$\frac{R^{(n-1)}(\lambda; A)}{(n-1)!} = (-1)^{n-1} R(\lambda; A)^n, \tag{14}$$

where $R^{(n-1)}$ is the $(n-1)^{st}$ derivative of R with respect to the first argument. By Lemma 9.8, we have

$$R(\lambda; A)x = \int_0^\infty e^{-\lambda t} T(t)x\, dt. \tag{15}$$

Combining (14) and (15) we arrive at

$$\begin{aligned} R(\lambda; A)^n &= \frac{(-1)^{n-1}}{(n-1)!} \int_0^\infty e^{-\lambda t}(-t)^{n-1} T(t)x\, dt \\ &= \frac{1}{(n-1)!} \int_0^\infty e^{-\lambda t} t^{n-1} T(t)x\, dt. \quad \square \end{aligned}$$

Theorem 9.8 (Hille-Yosida, 1948): Let X be a Banach space and $M, \omega \in \mathbb{R}$ be given. Let $\mathcal{D}(A) \subset X$ and assume that $A : \mathcal{D}(A) \to X$ is linear. Then A is the infinitesimal generator of a linear C_0-semigroup satisfying $\|T(t)\| \leq M e^{\omega t}$ for all $t \geq 0$ if and only if (i) and (ii) below hold:

(i) $\mathcal{D}(A)$ is dense in X and A is closed.

(ii) $\rho(A) \supset \{\lambda \in \mathbb{R} : \lambda > \omega\}$ and

$$\|R(\lambda; A)^n\| \leq \frac{M}{(\lambda - \omega)^n} \quad \text{for all } n \in \mathbb{N} \text{ and all } \lambda \in \mathbb{R} \text{ with } \lambda > \omega.$$

Remark 9.9: The inequality in (ii) of Theorem 9.8 can be quite complicated to check in practice. Observe that if

$$\|R(\lambda; A)\| \leq \frac{M}{\lambda - \omega},$$

then

$$\|R(\lambda; A)^n\| \leq \frac{M^n}{(\lambda - \omega)^n}.$$

Consequently, if $M = 1$ and the inequality in (i) holds when $n = 1$ then it automatically holds for all $n \in \mathbb{N}$.

Proof of the Hille-Yosida Theorem: (Necessity) It follows from Theorem 9.4 that (i) holds. Also, it follows from Lemma 9.6 that

$$\rho(A) \supset \{\lambda \in \mathbb{R} : \lambda > \omega\}.$$

Let $x \in X$, $n \in \mathbb{N}$ and $\lambda \in \mathbb{R}$ with $\lambda > \omega$ be given. By Lemma 9.7, we have

$$\|R(\lambda; A)^n\| x = \frac{1}{(n-1)!} \int_0^\infty e^{-\lambda t} t^{n-1} T(t) x \, dt.$$

It follows that

$$\|R(\lambda; A)^n\| \leq \frac{M}{(n-1)!} \int_0^\infty e^{-\lambda t} t^{n-1} e^{\omega t} \|x\| \, dt. \tag{16}$$

Combining (15) with (16) we obtain

$$\|R(\lambda; A)^n x\| \leq \frac{M\|x\|}{(\lambda - \omega)^n}.$$

(Sufficiency) We shall approximate A by bounded linear operators, construct semigroups generated by the approximating operators by using the standard exponential series, and then pass to the limit to obtain a semigroup generated by A.

Let $x \in \mathcal{D}(A)$ be given. Then for all $\lambda > \omega$ we have

$$(\lambda I - A) R(\lambda; A) x = x$$

and consequently

$$\begin{aligned} \lambda R(\lambda; A) x - x &= A R(\lambda; A) x \\ &= R(\lambda; A) A x. \end{aligned} \tag{17}$$

It follows that

$$\|\lambda R(\lambda; A) x - x\| \leq \frac{M\|Ax\|}{(\lambda - \omega)} \to 0 \quad \text{as } \lambda \to \infty.$$

Since $\mathcal{D}(A)$ is dense in X we have established that

$$\forall x \in X, \; \lambda R(\lambda; A) x \to x \quad \text{as } \lambda \to \infty.$$

Lecture Notes for Week 10 (First Draft)

Proof of the Hille-Yosida Theorem (Sufficiency Continued)

We assume that (i) and (ii) of Theorem 9.8 hold.

Last time we showed that

$$\forall x \in X, \quad \lambda R(\lambda; A)x \to x \quad \text{as } \lambda \to \infty. \tag{1}$$

This suggests a nice way of approximating A by bounded linear operators. For (real) $\lambda > \omega$ put

$$\begin{aligned} A_\lambda x &= \lambda A R(\lambda; A)x \\ &= \lambda^2 R(\lambda; A)x - \lambda x \quad \text{for all } x \in X. \end{aligned} \tag{2}$$

Observe that $A_\lambda \in \mathcal{L}(X; X)$ for all $\lambda > \omega$.

The operators A_λ are called the *Yosida approximates* of A. Using (1) and (2) the fact that $AR(\lambda; A)x = R(\lambda; A)Ax$ for $x \in \mathcal{D}(A)$ we see that

$$\forall x \in \mathcal{D}(A), \quad A_\lambda x \to Ax \quad \text{as } \lambda \to \infty. \tag{3}$$

In order to complete the proof, we need some basic results concerning semigroups generated by bounded linear operators. These will be stated in the form of a claim. Proof of the claim will be a homework exercise.

Claim: Let $B \in \mathcal{L}(X; X)$ be given. For each $t \geq 0$, put

$$e^{tB} = \sum_{n=0}^{\infty} \frac{(tB)^n}{n!}.$$

[The above definition is also valid for negative values of t as well.] Then we have

(i) The mapping $t \to e^{tB}$ is a linear C_0-semigroup with infinitesimal generator B.

(ii) If $C \in \mathcal{L}(X; X)$ and $CB = BC$ then $Ce^{tB} = e^{tB}C$ for all $t \geq 0$.

(iii) $\lim_{t \downarrow 0} \|e^{tB} - I\| = 0$.

(iv) For all $\lambda \in \mathbb{K}$ we have

$$e^{t(B-\lambda I)} = e^{-\lambda t} e^{tB} \quad \text{for all } t \geq 0.$$

Using (iv) of the claim and (2) we see that for all $\lambda > \omega$ we have

$$e^{tA_\lambda} = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{\lambda^{2n} t^n R(\lambda; A)^n}{n!} \quad \text{for all } t \geq 0. \tag{4}$$

Using (ii) from Theorem 9.8 and (4), we find that for all $\lambda > \omega$

$$\begin{aligned} \|e^{tA_\lambda}\| &\leq Me^{-\lambda t} \sum_{n=1}^{\infty} \frac{\lambda^{2n} t^n}{(\lambda - \omega) n!} \quad \text{for all } t \geq 0, \\ &\leq Me^{-\lambda t} \exp\left(\frac{\lambda^2 t}{\lambda - \omega}\right) \quad \text{for all } t \geq 0, \qquad (5) \\ &\leq M \exp\left(\frac{\lambda \omega t}{\lambda - \omega}\right) \quad \text{for all } t \geq 0. \end{aligned}$$

Let $\omega_1 > \omega$ be given. Then, in view of (5) we may choose $\lambda_1 > \omega$ such that

$$\|e^{tA_\lambda}\| \leq Me^{\omega_1 t} \quad \text{for all } t \geq 0, \ \lambda > \lambda_1. \tag{6}$$

Put

$$T_\lambda(t) = e^{tA_\lambda} \quad \text{for all } \lambda > \omega, \ t \geq 0.$$

Observe that

$$A_\mu A_\lambda = A_\lambda A_\mu, \quad A_\lambda T_\mu(t) = T(t) A_\mu \quad \text{for all } \lambda, \mu > \omega, \ t \geq 0. \tag{7}$$

Let $x \in \mathcal{D}(A)$ and $\lambda, \mu > \omega$ be given. Then we have

$$\begin{aligned} T_\lambda(t)x - T_\mu(t)x &= \int_0^t \frac{d}{ds}(T_\mu(t-s)T_\lambda(s)x)\, ds \\ &= \int_0^t T_\mu(t-s)A_\lambda T_\lambda(s) - T_\mu(t-s)A_\mu T_\lambda(s)x\, ds \qquad (8) \\ &= \int_0^t (T_\mu(t-s)T_\lambda(s))(A_\lambda x - A_\mu x)\, ds. \end{aligned}$$

Using (6) and (8) we find that

$$\begin{aligned} \|T_\lambda(t)x - T_\mu(t)x\| &\leq M^2 e^{\omega_1 t} t \|A_\lambda x - A_\mu x\| \\ &\leq tM^2 e^{\omega_1 t}(\|A_\lambda x - Ax\| + \|A_\mu x - Ax\|) \quad \text{for all } \lambda, \mu > \lambda_1. \end{aligned} \tag{9}$$

For $x \in \mathcal{D}(A)$, it follows from (9) that $\{T_\lambda(t)x\}_{\lambda > \lambda_1}$ has the Cauchy property in λ, uniformly for t in bounded sets. Using (6), (9), and the fact that $\mathcal{D}(A)$ is dense in

X, we see that for all $x \in X$, $\{T_\lambda(t)x\}_{\lambda>\lambda_1}$ has the Cauchy property in λ, uniformly for t in bounded sets. Now define

$$T(t)x = \lim_{\lambda\to\infty} T_\lambda(t)x \quad \text{for all } x \in X,\ t \geq 0.$$

It is immediate that

- $T(0) = I$,
- $T(t+s) = T(t)T(s)$ for all $s, t \geq 0$,
- $\|T(t)\| \leq Me^{\omega_1 t}$ for all $t \geq 0$.

Since the last inequality holds (with the same M) for every $\omega_1 > \omega$ we conclude that

- $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$.

Since the convergence of $T_\lambda(t)x$ to $T(t)x$ is uniform for t in bounded sets, we conclude that for every $x \in X$ the mapping $t \to T(t)x$ is continuous. It follows that T is a linear C_0-semigroup.

It remains to show that the infinitesimal generator is A. For this purpose, let us denote the infinitesimal generator of T by $\hat{A}$. We shall first show that $\hat{A}$ is an extension of A and then use a resolvent argument to show that $\mathcal{D}(\hat{A}) = \mathcal{D}(A)$.

Let $x \in \mathcal{D}(A)$ be given and observe that

$$\begin{aligned} \|T_\lambda(t)A_\lambda x - T(t)Ax\| &\leq \|T_\lambda(t)(A_\lambda x - Ax)\| + \|(T_\lambda(t) - T(t)Ax\| \\ &\leq Me^{\omega_1 t}\|A_\lambda x - Ax\| + \|(T_\lambda(t) - T(t))Ax\| \qquad (10) \\ &\to 0 \text{ as } \lambda \to \infty. \end{aligned}$$

Since the convergence of $T_\lambda(t)A_\lambda x$ to $T(t)Ax$ is uniform for t in bounded intervals we have

$$\begin{aligned} T(t)x - x &= \lim_{\lambda\to\infty} T_\lambda(t)x - x \\ &= \lim_{\lambda\to\infty} \int_0^t T_\lambda(s)A_\lambda x\, ds = \int_0^t T(s)Ax\, ds. \end{aligned} \qquad (11)$$

For $h > 0$ we have

$$\frac{T(h)x - x}{h} = \frac{1}{h}\int_0^h T(s)Ax\, ds \to Ax \quad \text{as } h \downarrow 0.$$

We conclude that $x \in \mathcal{D}(\hat{A})$ and $\hat{A}x = Ax$. In other words, $\hat{A}$ is an extension of A. In order to complete the proof, it suffices to show that $\mathcal{D}(\hat{A}) \subset \mathcal{D}(A)$.

Since $\hat{A}$ is an infinitesimal generator, it is a closed operator. Moreover, A is closed by assumption. Recall that for a closed linear operator $C : \mathcal{D}(C) \to X$ and $\lambda \in \rho(C)$ the operator $\lambda I - C$ is surjective, i.e. $(\lambda I - C)[\mathcal{D}(C)] = X$. By Lemma 9.6, $\rho(\hat{A}) \supset (\omega, \infty)$ and by assumption we have $\rho(A) \supset (\omega, \infty)$. Therefore we may choose $\lambda \in \rho(A) \cap \rho(\hat{A})$.

Since $\hat{A}$ and A are closed and $\lambda \in \rho(A) \cap \rho(\hat{A})$

$$(\lambda I - \hat{A})[\mathcal{D}(\hat{A})] = X,$$

$$(\lambda I - A)\mathcal{D}(A)] = X,$$

and since $\hat{A}$ extends A, we also have

$$(\lambda I - \hat{A})[\mathcal{D}(A)] = X.$$

It follows that

$$\mathcal{D}(A) = R(\lambda; \hat{A})[X] = \mathcal{D}(\hat{A}). \quad \square$$

Corollary 10.1: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $\mathcal{D}(A)$ is dense in X and that $A : \mathcal{D}(A) \to X$ is linear and closed. Then A generates a linear C_0 contraction semigroup. if and only if $\rho(A) \supset (0, \infty)$ and

$$\|R(\lambda; A)\| \leq \frac{1}{\lambda} \text{ for all } \lambda > 0.$$

Remark 10.2: *Historical Comments* TO BE FILLED IN

Contraction Semigroups

Assume that $T : [0, \infty) \to \mathcal{L}(X; X)$ is a linear C_0 contraction semigroup with infinitesimal generator A. Then for all $t, h \geq 0$ and all $x \in X$ we have

$$\|T(t+h)x\| \leq \|T(h)\| \cdot \|T(t)x\| \leq \|T(t)x\|$$

and

$$\|T(t+h)\| \leq \|T(h)\| \cdot \|T(t)\| \leq \|T(t)\|.$$

In other words, the mappings $t \to \|T(t)x\|$ and $t \to \|T(t)\|$ both are nonincreasing on $[0, \infty)$.

Suppose that X is a Hilbert space. Let $x \in \mathcal{D}(A)$ be given and put

$$u(t) = \|T(t)x\|^2 = (T(t)x, T(t)x) \text{ for all } t \geq 0.$$

Then u is nonincreasing and differentiable, so we have

$$\dot{u}(t) = (T(t)x, T(t)Ax) + (T(t)Ax, T(t)x) = 2\mathrm{Re}(T(t)Ax, x) \leq 0 \text{ for all } t \geq 0.$$

Putting $t = 0$ we find that

$$\mathrm{Re}(Ax, x) \leq 0 \quad \text{for all } x \in \mathcal{D}(A).$$

We will prove the following remarkable result: If X is a Hilbert space, $\mathcal{D}(A) \subset X$, and $A : \mathcal{D}(A) \to X$ is linear then A generates a linear C_0 contraction C_0-semigroup if and only if (i) and (ii) below hold:

(i) $\mathrm{Re}(Ax, x) \leq 0$ for all $x \in \mathcal{D}(A)$,

(ii) There exists $\lambda_0 > 0$ such that $\lambda_0 I - A$ is surjective.

[Density of $\mathcal{D}(A)$, closedness of A, and the required resolvent properties of A all follow from (i) and (ii) above!]

We shall actually prove some extensions of this type of result for semigroups on Banach spaces. For this purpose we shall introduce the notion of semi-inner product on a Banach space. We shall present only those concepts that we need to prove a few basic results concerning contraction semigroups.

Semi-Inner Products on Banach Spaces

Definition 10.3: Let X be a linear space over $\mathbb{K}$. By a *semi-inner product* on X, we mean a mapping $[\cdot, \cdot] : X \times X \to \mathbb{K}$ such that

(i) $[x + y, z] = [x, z] + [y, z]$ for all $x, y, z \in X$,

(ii) $[\alpha x, y] = \alpha[x, y]$ for all $x, y \in X$, $\alpha \in \mathbb{K}$,

(iii) $[x, x] > 0$ for all $x \in X \backslash \{0\}$,

(iv) $|[x, y]|^2 \leq [x, x] \cdot [y, y]$ for all $x, y \in X$.

Observe that every inner product is a semi-inner product. It follows easily from Definition 10.3 that if $[\cdot, \cdot]$ is a semi-inner product on X then the function $x \to \sqrt{[x, x]}$ is a norm on X. It turns out that every norm on X is induced by a semi-inner product. This, of course, means that there are semi-inner products that are not inner products. The notion of semi-inner product was introduced by Lumer in 1961 and developed further by Giles in 1967. Giles showed that one can include the homogoneity property $[x, \alpha y] = \overline{\alpha}[x, y]$ for all $x, y \in X$, $\alpha \in \mathbb{K}$ to the definition of semi-inner product without introducing any significant complications.

Definition 10.4: Let $(X, \| \cdot \|)$ be a normed linear space and $[\cdot, \cdot]$ be a semi-inner product on X. We say that $[\cdot, \cdot]$ is *compatible* with $\| \cdot \|$ provided that $[x, x] = \|x\|^2$ for all $x \in X$.

Proposition 10.5: Let $(X, \|\cdot\|)$ be a normed linear space. Then there is at least one semi-inner product $[\cdot,\cdot]$ on X compatible with $\|\cdot\|$.

Proof: For every $x \in X$, put

$$\mathcal{F}(x) = \{x^* \in X^* : x^*(x) = \|x\|^2 = \|x^*\|^2\}. \tag{12}$$

By the Hahn-Banach Theorem, $\mathcal{F}(x) \neq \emptyset$ for every $x \in X$, so we may choose $F(x) \in \mathcal{F}(x)$. Define $[\cdot,\cdot] : X \times X \to \mathbb{K}$ by

$$[x, y] = (F(y))(x) \text{ for all } x, y \in X. \tag{13}$$

It is straightforward to check that $[\cdot,\cdot]$ is a semi-inner product compatible with $\|\cdot\|$. □

Remark 10.6: There is exactly one semi-inner product on X compatible with $\|\cdot\|$ if and only if for every $x \in X$ the set $\mathcal{F}(x)$ in (12) is a singleton. A convenient sufficient condition for this property to hold is that X^* be strictly convex.

Remark 10.7: If $[\cdot,\cdot]$ is a semi-inner product compatible with the norm on a normed linear space $(X, \|\cdot\|)$ then for each $y \in X$, the mapping $x \to [x, y]$ is a continuous linear functional. As an aside, it is interesting to note that if X is reflexive and $x^* \in X^*$, then there exists a semi-inner product on X and $y \in Y$ such that $x^*(x) = [x, y]$ for all $x \in X$.

Example 10.8: Let $p \in (1, \infty)$ be given and let $X = L^p[0, 1]$ with the usual norm

$$\|f\|_p = \left(\int_0^1 |f(x)|^p \, dx\right)^{\frac{1}{p}}.$$

It is straightforward to check that the mapping $[\cdot,\cdot] : X \times X \to \mathbb{R}$ defined by $[f, 0] = 0$ and

$$[f, g] = \frac{1}{\|g\|_p^{p-2}} \int_0^1 f(x)\overline{g(x)}|g(x)|^{p-2}\mathrm{sgn}(g(x)) \, dx \text{ for } \|g\|_p \neq 0$$

is a semi-inner product compatible with $\|\cdot\|_p$.

Definition 10.9: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A) \to X$ is linear. We say that A is *dissipative* provided that there exists a semi-inner product $[\cdot,\cdot]$ compatible with the norm on X such that

$$\mathrm{Re}[Ax, x] \leq 0 \text{ for all } x \in \mathcal{D}(A).$$

Lemma 10.10: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A) \to X$ is linear. Then A is dissipative if and only if

$$\|(\lambda I - A)x\| \geq \lambda\|x\| \text{ for all } x \in \mathcal{D}(A), \ \lambda > 0.$$

Proof: Assume that A is dissipative and choose a semi-inner product $[\cdot,\cdot]$ compatible with the norm on X such that $\mathrm{Re}[x,x] \le 0$ for all $x \in \mathcal{D}(A)$. Let $x \in \mathcal{D}(A)$ and $\lambda > 0$ be given. Then we have

$$\mathrm{Re}[(\lambda I - A)x, x] = \lambda\|x\|^2 - \mathrm{Re}[Ax, x] \ge \lambda\|x\|^2. \tag{14}$$

On the other hand we have

$$\mathrm{Re}[(\lambda I - A)x, x] \le |[(\lambda I - A)x, x]| \le \|(\lambda I - A)x\| \cdot \|x\|. \tag{15}$$

Combining (14) with (15) we get

$$\lambda\|x\|^2 \le \|(\lambda I - A)x\| \cdot \|x\|,$$

which yields the desired conclusion.

Assume now that

$$\|(\lambda I - A)x\| \ge \lambda\|x\| \;\; \text{for all } x \in calD(A),\; \lambda > 0. \tag{16}$$

In order to avoid some technical complications we assume here that X is either separable or reflexive. (To handle the general case, one can proceed as below using nets instead of sequences.) As before, for all $z \in X$, put

$$\mathcal{F}(z) = \{x^* \in X^* : x^*(z) = \|z\|^2 = \|x^*\|^2\}.$$

Let $x \in \mathcal{D}(A)\backslash\{0\}$ be given and observe that

$$\|(nI - A)x\| \ge n\|x\| > 0 \;\; \text{for all } n \in \mathbb{N}.$$

For every $n \in \mathbb{N}$ we choose $y_n^* \in \mathcal{F}(nx - Ax)$ and notice that

$$y_n^*(nx - Ax) = \|nx - Ax\|^2 = \|y_n^*\|^2 > 0.$$

Now put

$$z_n^* = \frac{1}{\|y_n^*\|} y_n^* \;\; \text{for all } n \in \mathbb{N}.$$

Then we have

$$\begin{aligned} \|nx - Ax\| &= \frac{\|y_n^*\|}{y}{}^*_n (nx - Ax) \\ &= z_n^*(nx - Ax) \\ &= n\mathrm{Re} z_n^*(x) - \mathrm{Re} z_n^*(Ax). \end{aligned}$$

Since $\|z_n^*\| = 1$ for all $n \in \mathbb{N}$ we find that

$$n\|x\| \le \|nx - Ax\| = n\mathrm{Re} z_n^*(x) - \mathrm{Re} z_n^*(Ax) \le n\|x\| - \mathrm{Re} z_n^*(Ax) \tag{17}$$

We choose a subsequence $\{z^*_{n_k}\}_{k=1}^\infty$ and $x^* \in X^*$ such that

$$z^*_{n_k} \overset{*}{\rightharpoonup} z^* \text{ (weakly}^*\text{) as } k \to \infty.$$

Then we have $\|z^*\| \leq 1$, $\mathrm{Re} z^*(Ax) \leq 0$, and $\mathrm{Re} z^*(x) \geq \|x\|$. It follows that $z^*(x) = \|x\|$.

Define $F : X \to X^*$ by

$$F(x) = \begin{cases} 0 & x = 0 \\ z^*\|x\| & x \in \mathcal{D}(A)\backslash\{0\} \\ \text{any element of } \mathcal{F}(x) & x \in X\backslash\mathcal{D}(A). \end{cases}$$

If we define the semi-inner product $[\cdot,\cdot]$ by

$$[x, y] = (F(y))(x) \quad \text{for all } x, y \in X,$$

then $\mathrm{Re}[Ax, x] \leq 0$ for all $x \in \mathcal{D}(A)$. $\square$

Lemma 10.11: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A) \to X$ is linear and dissipative. Let $\lambda_0 > 0$ be given and assume that $\lambda_0 I - A$ is surjective. Then A is closed, $\rho(A) \supset (0, \infty)$ and

$$\|R(\lambda; A)\| \leq \frac{1}{\lambda} \quad \text{for all } \lambda > 0.$$

Proof: The key points to prove are that A is closed and that $\lambda I - A$ is surjective for every $\lambda > 0$. (Injectivity of $\lambda I - A$ and the inequality for $\|R(\lambda; A)\|$ follow easily from Lemma 10.9.)

Since $\|(\lambda_0 - A)x\| \geq \lambda_0\|x\|$ for all $x \in \mathcal{D}(A)$ by Lemma 10.9, we conclude that $\lambda_0 I - A$ is injective and

$$\|(\lambda_0 I - A)^{-1} y\| \frac{1}{\lambda_0}\|y\| \quad \text{for all } y \in X.$$

It follows that $(\lambda_0 I - A)^{-1} \in \mathcal{L}(X; X)$ and consequently this operator is closed. We conclude that $\lambda_0 I - A$ is closed and this implies that A is closed.

To show that $\rho(A) \supset (0, \infty)$, put

$$\Lambda = \{\lambda \in (0, \infty) : \lambda \in \rho(A)\}.$$

We know that $\Lambda \neq \emptyset$ (because it contains λ_0). Observe that Λ is open in $(0, \infty)$. We shall show that Λ is also closed in $(0, \infty)$ (with the relative topology). Since $(0, \infty)$ is connected, this will ensure that $\Lambda = (0, \infty)$.

Let $\{\lambda_n\}_{n=1}^\infty$ be a sequence in Λ converging to $\lambda_* \in (0, \infty)$. To show that $\lambda_* \in \Lambda$ it suffices to show that $\lambda_* I - A$ is surjective. Let $y \in X$ be given and for every $n \in \mathbb{N}$, put

$$x_n = R(\lambda_n; A)y.$$

We shall show that that the sequence $\{x_n\}_{n=1}^{\infty}$ is convergent to some $x \in X$ and that $(\lambda_* I - A)x = y$. Before developing the relevant inequalities, we observe that the sequence $\{1/\lambda_n\}_{n=1}^{\infty}$ is bounded because it converges to $1/\lambda_*$.

Let $m, n \in \mathbb{N}$ be given. Then we have

$$\begin{aligned} \|x_n - x_m\| &= \|R(\lambda_n; A)y - R(\lambda_m; A)y\| \\ &= \|(\lambda_n - \lambda_m)R(\lambda_n; A)R(\lambda_m; A)\| \\ &= |\lambda_n - \lambda_m| \cdot \|R(\lambda_n; A)\| \cdot \|R(\lambda_m; A)\| \\ &\leq \|\lambda_n - \lambda_m\| \frac{y}{\lambda_n \lambda_m}. \end{aligned}$$

It follows that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Put

$$x = \lim_{n \to \infty} x_n.$$

Since $x_n \in \mathcal{D}(A)$ for all $n \in \mathbb{N}$ and $Ax_n \to \lambda_* x - y$ and the operator A is closed we infer that $x \in \mathcal{D}(A)$ and $Ax = \lambda_* x - x$. This completes the proof. $\square$

Lecture Notes for Week 11 (First Draft)

Dissipative Operators (Continued)

Theorem 11.1 (Lumer, Philips, 1961): Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to X$ is linear.

(a) If A is dissipative and there exists $\lambda_0 > 0$ such that $\lambda_0 I - A$ is surjective then A generates a linear C_0-contraction semigroup.

(b) If A generates a linear C_0-contraction semigroup then $\lambda I - A$ is surjective for every $\lambda > 0$ and for every semi-inner product $[\cdot, \cdot]$ compatible with the norm on X we have

$$\mathrm{Re}[Ax, x] \leq 0 \quad \text{for all } x \in \mathcal{D}(A).$$

In particular, A is dissipative.

Proof: Let $\lambda_0 > 0$ be given. Assume that A is dissipative and that $\lambda_0 I - A$ is surjective. Then, by Lemma 10.11, A is closed, $\rho(A) \supset (0, \infty)$, and

$$\|R(\lambda; A)\| \leq \frac{1}{\lambda} \quad \text{for all } \lambda > 0.$$

Consequently, for all $n \in \mathbb{N}$ we have

$$\|R(\lambda; A)^n\| \leq \|R(\lambda; A)\|^n \leq \frac{1}{\lambda^n} \quad \text{for all } \lambda > 0.$$

It follows from the Hille-Yosida Theorem that A generates a linear C_0-semigroup satisfying $\|T(t)\| \leq 1$ for all $t \geq 0$ (i.e., a contraction semigroup).

Assume now that A generates a linear C_0-contraction semigroup. Them, by the Hille-Yosida Theorem, A is closed and $\rho(A) \supset (0, \infty)$. It follows that $\lambda I - A$ is surjective for all $\lambda > 0$. Let $[\cdot, \cdot]$ be a semmi-inner product compatible with the norm on X. Let $x \in \mathcal{D}(A)$ and $h > 0$ be given. then we have

$$\begin{aligned} \mathrm{Re}[T(h)x - x, x] &= \mathrm{Re}[T(h)x, x] - \|x\|^2 \\ &\leq \|T(h)\| \cdot \|x\| - \|x\|^2 \\ &\leq 0. \end{aligned} \tag{1}$$

Using (1) we see that

$$\mathrm{Re}[Ax, x] = \lim_{h \downarrow 0} \mathrm{Re}\left[\frac{T(h)x - x}{h}, x\right] \leq 0. \quad \square$$

Corollary 11.2: Let X be a Banach space and $\mathcal{D}(B) \subset X$. Assume that $\mathcal{D}(B)$ is dense and that $B : \mathcal{D}(B) \to X$ is linear. Let $\omega, \lambda_0 \in \mathbb{R}$ be given with $\lambda_0 > \omega$. Assume that $\lambda_0 I - A$ is surjective and that there exists a semi-inner product compatible with the norm on X such that

$$\mathrm{Re}[Bx, x] \leq \omega \|x\|^2 \text{ for all } x \in \mathcal{D}(B).$$

Then B generates a linear C_0-semigroup satisfying

$$\|T(t)\| \leq e^{\omega t} \text{ for all } t \geq 0.$$

Proof: Put $\mathcal{D}(A) = \mathcal{D}(B)$, $A = B - \omega I$, and use Theorem 11.1. $\square$

Lemma 11.3: Let X be a reflexive Banach space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A) \to X$ is linear and dissipative. Let $\lambda_0 > 0$ be given and assume that $\lambda_0 I - A$ is surjective. Then $\mathcal{D}(A)$ is dense.

Remark 11.4: Let X be a Banach space (not necessarily reflexive) and $Z \subset X$ be a linear manifold. Then Z is dense if and only if for every $y \in X$ there is a sequence $\{x_n\}_{n=1}^{\infty}$ in Z such that $x_n \rightharpoonup y$ (weakly) as $n \to \infty$. [Indeed, if Z is dense, then for every $y \in Z$ we can find a sequence of elements of Z that converges strongly to y. To see that the converse is true, if $y \notin \mathrm{cl}(Z)$ then $\mathrm{dist}(Z, y) > 0$ and by the Hahn-Banach Theorem we may choose a linear functional $x^* \in X^*$ such that $x^*(y) \neq 0$ and and $x^*(x) = 0$ for all $x \in Z$.]

Proof of Lemma 11.13: Let $y \in X$ be given. We shall construct a sequence $\{x_n\}_{n=1}^{\infty}$ such that $x_n \in \mathcal{D}(A)$ for all $n \in \mathbb{N}$ and $x_n \rightharpoonup y$ (weakly) as $n \to \infty$.

By Lemma 10.11, we know that A is closed and $\rho(A) \supset (0, \infty)$. For all $n \in \mathbb{N}$, put

$$x_n = \left(I - \frac{1}{n}A\right)^{-1} y = nR(n; A)y \in \mathcal{D}(A). \tag{2}$$

A simple computation shows that

$$A\left(\frac{x_n}{n}\right) = x_n - y \text{ for all } n \in \mathbb{N}. \tag{3}$$

Lemma 10.11 also ensures that $\|R(n; A)\| \leq n^{-1}$ for all $n \in \mathbb{N}$, and consequently we have

$$\|x_n\| \leq n\|R(n; A)\| \cdot \|y\| \leq \|y\| \text{ for all } n \in \mathbb{N}.$$

Since X is reflexive and $\{x_n\}_{n=1}^{\infty}$ is bounded, we may choose a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ and $z \in X$ such that

$$x_{n_k} \rightharpoonup z \text{ (weakly) as } k \to \infty.$$

We want to show that $z = y$. Using (3) we see that

$$A\left(\frac{x_{n_k}}{n_k}\right) = x_{n_k} - y \rightharpoonup z - y \text{ (weakly) as } k \to \infty.$$

We also know that
$$\frac{x_{n_k}}{n_k} \rightharpoonup 0 \text{ (weakly) as } k \to \infty$$
(in fact; it converges strongly to 0). Since $\mathrm{Gr}(A)$ is closed and convex, it is weakly closed and we deduce that $(0, z-y) \in \mathrm{Gr}(A)$. This implies that $z = y$. $\square$

Theorem 11.5 (Lumer-Philips Theorem for Hilbert Spaces): Let X be a Hilbert space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A) \to X$ is linear. Assume further that $\mathrm{Re}(Ax, x) \le 0$ for all $x \in \mathcal{D}(A)$ and that there exists $\lambda_0 > 0$ such that $\lambda_0 I - A$ is surjective. Then A generates a linear C_0-contraction semigroup.

Example 11.6 (Heat Equation): Let $X = L^2[0,1]$ and put
$$\mathcal{D}(A) = \{u \in AC[0,1] : u' \in AC[0,1],\ u'' \in L^2[0,1],\ u(0) = u(1) = 0\},$$
$$Au = u'' \text{ for all } u \in \mathcal{D}(A).$$

We shall apply Theorem 11.5 to show that A generates a linear C_0-contraction semigroup. Let $u \in \mathcal{D}(A)$ be given. Using integration by parts, we find that
$$\begin{aligned}(Au, u) &= \int_0^1 u''(x)\overline{u(x)}\,dx \\ &= u'(x)\overline{u(x)}\Big|_{x=0}^1 - \int_0^1 u'(x)\overline{u'(x)}\,dx \\ &= -\int_0^1 |u'(x)|^2\,dx \le 0.\end{aligned}$$
We shall show that $I - A$ is surjective. Let $g \in L^2[0,1]$ be given. We want to find $v \in \mathcal{D}(A)$ such that
$$\begin{cases} v(x) - v''(x) = g(x) \quad \text{a.e. } x \in [0,1] \\ v(0) = v(1) = 0. \end{cases} \tag{4}$$
It is possible to appeal to (or to prove) general theorems that ensure the existence of a suitable solution to (4); however using techniques from elementary differential equations, we can simply exhibit a suitable solution of (4), namely
$$v(x) = k \sinh x + \int_0^x \sinh(y - x) g(y)\,dy,$$
where the constant k is given by
$$k = \frac{1}{\sinh 1}\int_0^1 \sinh(1-y)\,dy.$$
(This solution is obtained by using variation of parameters to find the general solution of $v - v'' = g$ and then choosing the constants so that $v(0) = v(1) = 0$.)

It follows from Theorem 11.5 that A generates a linear C_0 contraction semigroup $T : [0,\infty) \to \mathcal{L}(X;X)$. For $u_0 \in \mathcal{D}(A)$ let us put

$$u(t,x) = (T(t)u_0)(x) \quad \text{for all } x \in [0,1],\ t \geq 0. \tag{5}$$

Then u is a solution of the initial-boundary value problem

$$\begin{cases} u_t(t,x) = u_{xx}(t,x) \quad x \in [0,1],\ t \geq 0 \\ u(t,0) = u(t,1) = 0 \quad t \geq 0 \\ u(0,x) = u_0(x) \quad x \in [0,1] \end{cases}$$

for the *heat equation* $u_t = u_{xx}$. Here u_t and u_x indicate partial derivatives of u with respect to the first and second argument, respectively.

The semigroup T of this example has important regularizing properties that will be addressed later. Even if $u_0 \notin \mathcal{D}(A)$, the function u produced by (5) is very smooth on $(0,\infty) \times [0,1]$.

Nonhomogeneous Differential Equations

Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to X$ is linear and closed. Let $\tau > 0$, $f \in C([0,\tau];X)$ and $u_0 \in X$ be given.

Consider the nonhomogeneous initial value problem

$$\begin{cases} \dot{u}(t) = Au(t) + f(t), \quad t \in (0,\tau] \\ u(0) = u_0. \end{cases} \tag{NHIVP}$$

Suppose that A generates a linear C_0-semigroup $T : [0,\infty) \to \mathcal{L}(X;X)$. We “know” that in this case, the solution to (NHIVP) should be given by

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\,ds. \tag{6}$$

Equation (6) is known as the *variation of parameters formula.* We have seen that if $u_0 \in \mathcal{D}(A)$ then the mapping $t \to T(t)u_0$ is differentiable on $[0,\tau]$ and takes values in $\mathcal{D}(A)$. What about the integral term in (6)? Let us put

$$v(t) = \int_0^t T(t-s)f(s)\,ds \quad \text{for all } t \in [0,\tau].$$

It can easily happen that

$$\forall t \in (0,\tau],\ v(t) \notin \mathcal{D}(A),$$

and

$$\forall t \in (0,\tau],\ v \text{ is not differentiable at } t.$$

We shall illustrate how things can go wrong with a simple example.

Example 11.7: (See Example 8.5) Let $X = BUC(\mathbb{R})$ and define $T : [0,\infty) \to \mathcal{L}(X;X)$ by

$$(T(t)w)(x) = w(x+t) \text{ for all } w \in X,\ x \in \mathbb{R},\ t \geq 0.$$

The infinitesimal generator A is given by

$$\mathcal{D}(A) = BUC^1(\mathbb{R}), \quad Aw = w' \text{ for all } w \in \mathcal{D}(A).$$

Let $z \in X \backslash \mathcal{D}(A)$ be given and put

$$f(t) = T(t)z \text{ for all } t \in [0,\tau],$$

so that

$$(f(t))(x) = z(x+t) \text{ for all } x \in \mathbb{R},\ t \geq 0.$$

Observe that for all $t \in [0,\tau]$ we have $f(t) \notin \mathcal{D}(A)$. [This is because a function in X belongs to $\mathcal{D}(A)$ if and only if all of its translates belong to $\mathcal{D}(A)$.] Observe further that

$$\begin{aligned} v(t) &= \int_0^t T(t-s)f(s)\,ds = \int_0^t T(t-s)T(s)z\,ds \\ &= \int_0^t T(t)z\,ds = tT(t)z. \end{aligned}$$

It follows immediately that for all $t \in (0,\tau]$, $v(t) \notin \mathcal{D}(A)$. We also see that v is not differentiable.

The following lemma (whose proof will be a homework exercise) gives simple conditions which ensure that the integral term from the variation of parameters formula is differentiable and takes values in the domain of A.

Lemma 11.8: Let X be a Banach space and $T : [0,\infty) \to \mathcal{L}(X;X)$ be a linear C_0-semigroup with infinitesimal generator A. Let $X_A = \mathcal{D}(A)$ equipped with the graph norm

$$\|x\|_A = \|x\| + \|Ax\| \text{ for all } x \in \mathcal{D}(A).$$

Let $\tau > 0$ and

$$F \in C^1([0,\tau];X), \quad G \in C([0,\tau];X_A)$$

be given. Put

$$f(t) = F(t) + G(t), \quad v(t) = \int_0^t T(t-s)f(s)\,ds \text{ for all } t \in [0,\tau].$$

Then

$$v \in C^1([0,\tau];X) \cap C([0,\tau];X_A)$$

and

$$\dot{v}(t) = Av(t) + f(t) \quad \text{for all } t \in [0, \tau].$$

Weak Solutions

Many authors define a "mild solution" of (NHIVP) via the variation of parameters formula (6). This approach is convenient, but not completely satisfactory, because one needs to know in advance that A generates a linear C_0-semigroup. It is desirable to have a notion of weak solution of (NHIVP) that makes no appeal to any semigroup, and then prove that if A generates a linear C_0-semigroup T the initial-value problem (NHIVP) has a unique weak solution and this solution is given by (6). The definition given here, as well as the theorem, is due to John Ball. Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to X$ is linear and closed. We shall employ the adjoint A^* of A. Let X^* denote the dual space of X and $\langle \cdot, \cdot \rangle : X^* \times X \to \mathbb{K}$ denote the duality pairing.

Suppose that $u : [0, \tau] \to X$ is differentiable on $(0, \tau]$, and that $u(t) \in \mathcal{D}(A)$ for all $t \in (0, \tau]$ and that u satisfies

$$\dot{u}(t) = Au(t) + f(t), \quad t \in (0, \tau]. \tag{NHODE}$$

Then for all $x^* \in \mathcal{D}(A^*)$ we have

$$\langle x^*, \dot{u}(t) \rangle = \langle x^*, Au(t) \rangle + \langle x^*, f(t) \rangle \quad \text{for all } t \in (0, \tau].$$

We can rewrite this equation as

$$\frac{d}{dt} \langle x^*, u(t) \rangle = \langle A^* x^*, u(t) \rangle + \langle x^*, f(t) \rangle, \quad t \in (0, \tau]. \tag{7}$$

Equation (7) makes sense for a much broader class of functions u. Moreover, if u is differentiable on $(0, \tau]$, $u(t) \in \mathcal{D}(A)$ for all $t \in (0, \tau]$, and u satisfies (7) for all $x^* \in \mathcal{D}(A^*)$ then u also satisfies (NHODE). This motivates the following definition.

Definition 11.9: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Let $\tau > 0$ and $f \in C([0, \tau]; X)$ be given. Assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to X$ is linear and closed. By a weak solution of (NHODE) we mean a function $u \in C([0, \tau]; X)$ such that for every $x^* \in \mathcal{D}(A^*)$ the function $t \to \langle x^*, u(t) \rangle$ is absolutely continuous on $[0, \tau]$ and satisfies

$$\frac{d}{dt} \langle x^*, u(t) \rangle = \langle A^* x^*, u(t) \rangle + \langle x^*, f(t) \rangle \quad \text{a.e. } t \in [0, \tau].$$

Remark 11.10: Notice that with $f \in C([0, \tau]; X)$, if u is a weak solution of (NHODE) then for every $x^* \in \mathcal{D}(A^*)$ the mapping $t \to \langle x^*, u(t) \rangle$ will actually belong to $C^1[0, \tau]$ (rather than just to $AC[0, 1]$). Definition 11.9 is still appropriate under the weaker assumption that $f \in L^1([0,\tau];X)$. Moreover, Theorem 11.11 below remains valid for $f \in L^1([0,\tau];X)$. [Ball gave the definition and proved the theorem for $f \in L^1([0,\tau];X)$.]

Theorem 11.11: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to X$ is linear and closed. Let $\tau > 0$ and $f \in C([0,\tau];X)$ be given. Then (i) and (ii) below are equivalent.

(i) For every $u_0 \in X$, (NHODE) has exactly one weak solution $u \in C([0,\tau];X)$ such that $u(0) = u_0$.

(ii) A generates a linear C_0-semigroup $T : [0,\infty) \to \mathcal{L}(X;X)$.

Moreover, if (ii) [and hence also (i)] holds, then for each $u_0 \in X$, the unique weak solution u of (NHODE) satisfying $u(0) = u_0$ is given by

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\,ds \quad \text{for all } t \in [0,\tau].$$

In order to prove Theorem 11.11, we shall make use of the following lemma, which is of interest in its own right.

Lemma 11.12: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to X$ is linear and closed. Let $x, z \in X$ be given and assume that

$$\langle v^*, z \rangle = \langle A^* v^*, x \rangle \quad \text{for all } v^* \in \mathcal{D}(A^*). \tag{8}$$

Then $x \in \mathcal{D}(A)$ and $Ax = z$.

Proof: Suppose that $(x,z) \notin \mathrm{Gr}(A)$. Since $\mathrm{Gr}(A)$ is a closed subspace of $X \times X$ the Hahn-Banach Theorem implies that we may choose $x^*, y^* \in X^*$ satisfying

$$\langle x^*, x \rangle + \langle y^*, z \rangle \neq 0 \tag{9}$$

and

$$\langle x^*, y \rangle + \langle y^*, Ay \rangle = 0 \quad \text{for all } y \in \mathcal{D}(A). \tag{10}$$

It follows from (10) that

$$y^* \in \mathcal{D}(A^*), \quad \text{and} \quad A^* y^* = -x^*. \tag{11}$$

Using (8) and (11) we obtain

$$\langle x^*, x \rangle = -\langle y^*, z \rangle,$$

which contradicts (9). $\square$

Proof of Theorem 11.11: Assume first that A generates a linear C_0-semigroup $T : [0, \infty) \to \mathcal{L}(X; X)$. Let $u_0 \in X$ be given and put

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\, ds, \quad \text{for all } t \in [0, \tau]. \tag{12}$$

Let $x^* \in \mathcal{D}(A^*)$ be given.

Claim: Given $x \in X$ let us put $w(t) = \langle x^*, T(t)x\rangle$ for all $t \in [0, \tau]$. Then $w \in C^1[0, \tau]$ and $\dot{w}(t) = \langle A^*x^*, T(t)x\rangle$ for all $t \in [0, \tau]$.

The claim is immediate if $x \in \mathcal{D}(A)$. If $x \notin \mathcal{D}(A)$, we may choose a sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathcal{D}(A)$ such that $x_n \to x$ as $n \to \infty$ and put $w_n(t) = \langle A^*x^*, T(t)x_n\rangle$. Then $w_n \to w$ uniformly on $[0, \tau]$ and the sequence $\{\dot{w}_n\}_{n=1}^{\infty}$ of derivatives also converges uniformly on $[0, \tau]$. If follows that $w \in C^1[0, \tau]$ and $\dot{w}(t) = \langle A^*x^*, T(t)x\rangle$ for all $t \in [0, \tau]$.

To handle the integral term in (12), observe that by the claim

$$\frac{\partial}{\partial t}\langle x^*, T(t-s)f(s)\rangle = \langle A^*x^*, T(t-s)f(s)\rangle. \tag{13}$$

The right-hand side of (13) is jointly continuous in s and t and consequently we have

$$\frac{d}{dt}\int_0^t \langle x^*, T(t-s)f(s)\rangle\, ds = \langle x^*, f(t)\rangle + \int_0^t \langle A^*x^*, T(t-s)f(s)\rangle\, ds. \tag{14}$$

It follows that the mapping $t \to \langle x^*, u(t)\rangle$ is continuously differentiable and

$$\begin{aligned}\frac{d}{dt}\langle x^*, u(t)\rangle &= \langle A^*x^*, T(t)u_0\rangle + \int_0^t \langle A^*x^*, T(t-s)f(s)\rangle\, ds + \langle x^*, f(t)\rangle \\ &= \langle A^*x^*, u(t)\rangle + \langle x^*, f(t)\rangle \quad \text{for all } t \in [0, t].\end{aligned}$$

Suppose that $\tilde{u}$ is another weak solution with $\tilde{u}(0) = u_0$. Put

$$w(t) = u(t) - \tilde{u}(t) \quad \text{for all } t \in [0, \tau],$$

and

$$W(t) = \int_0^t w(s)\, ds \quad \text{for all } t \in [0, \tau].$$

Observe that $W \in C^1([0, \tau]; X)$. Using the definition of weak solution and integrating with respect to time, we find that for all $x^* \in \mathcal{D}(A^*)$ and $t \in [0, \tau]$

$$\langle x^*, w(t)\rangle = \langle A^*x^*, \int_0^t w(s)\, ds\rangle,$$

and consequently

$$\langle x^*, \dot{W}(t)\rangle = \langle A^*x^*, W(t)\rangle.$$

It follows from Lemma 11.12 that

$$W(t) \in \mathcal{D}(A) \text{ and } AW(t) = \dot{W}(t) \text{ for all } t \in [0, \tau].$$

Now fix $t \in (0, \tau]$ and define $z : [0, t] \to X$ by

$$z(s) = T(t-s)W(s) \text{ for all } s \in [0, t].$$

Then z is differentiable, $z(0) = 0$ (since $W(0) = 0$), and

$$\dot{z}(s) = T(t-s)AW(s) - T(t-s)AW(s) = 0 \text{ for all } s \in [0, t].$$

We conclude that z is constant on $[0, t]$; in particular

$$0 = z(0) = z(t) = W(t).$$

Since $W(t) = 0$ for all $t \in [0, \tau]$, we have $w(t) = 0$ for all $t \in [0, \tau]$ and $u(t) = \tilde{u}(t)$ for all $t \in [0, \tau]$.

Assume now that for every $u_0 \in X$, (NHODE) has exactly one weak solution u satisfying $u(0) = u_0$. Let us denote the value of this solution at time t by $u(t; u_0)$. Now for each $u_0 \in X$ we can define

$$T(t)u_0 = u(t; t_0) - u(t; 0) \text{ for all } t \in [0, \tau]. \tag{15}$$

Observe that for all $x^* \in \mathcal{D}(A)$ we have

$$\frac{d}{dt}\langle x^*, T(t)u_0\rangle = \langle A^*x^*, T(t)u_0\rangle \text{ for all } t \in [0, \tau].$$

(In other words, $t \to T(t)u_0$ is a weak solution of $\dot{u} = Au$.) For $t > \tau$, we choose $n \in \mathbb{N}$ and $s \in [0, \tau)$ such that $t = n\tau + s$ and we define

$$T(t)u_0 = T(s)T(\tau)^n u_0.$$

It is not too difficult to show that T is a linear C_0-semigroup. This is left as an exercise. [Notice that the only weak solution of $\dot{w} = Aw$ on $[0, \tau]$ satisfying $w(0) = 0$ is identically 0 on $[0, \tau]$; otherwise there would be multiple weak solutions of (NHODE) on $[0, \tau]$ satisfying the same initial condition. This observation is useful for establishing the semigroup property. To establish continuity of T in the strong operator topology, consider the graph of the mapping $x \to T(\cdot)x$ from X to $C([0, \tau]; X)$.]

Let $\hat{A}$ be the infinitesimal generator of T. We need to show that $\hat{A} = A$. To accomplish this we shall first show that A is an extension of $\hat{A}$. Let $x \in \mathcal{D}(\hat{A})$ and $x^* \in \mathcal{D}(A^*)$ be given. Then we have

$$\frac{d}{dt}\langle x^*, T(t)x\rangle = \langle A^*x^*, T(t)x\rangle,$$

and evaluating this expression at $t = 0$ we find that

$$\frac{d}{dt}\langle x^*, T(t)x\rangle\Big|_{t=0} = \langle A^*x^*, x\rangle. \tag{16}$$

Since $\hat{A}$ is the infinitesimal generator of T we also have

$$\frac{d}{dt}\langle x^*, T(t)x\rangle\Big|_{t=0} = \langle x^*, \hat{A}x\rangle. \tag{17}$$

It follows from (16) and (17) that

$$\langle x^*, \hat{A}x\rangle = \langle A^*x^*, x\rangle \ \text{ for all } x^* \in \mathcal{D}(A^*).$$

Lemma 11.12 implies that $x \in \mathcal{D}(A)$ and $Ax = \hat{A}x$.

Now let $x \in X$ and $x^* \in \mathcal{D}(A^*)$ be given. By the definition of weak solution and the construction of T we have

$$\frac{d}{dt}\langle x^*, T(t)x\rangle = \langle A^*x^*, T(t)x\rangle \ \text{ for all } t \in [0, \tau]. \tag{18}$$

Integration of (18) gives

$$\langle x^*, T(t)x\rangle - \langle x^*, x\rangle = \langle A^*x^*, \int_0^t T(s)x\, ds\rangle \ \text{ for all } x \in X,\ x^* \in \mathcal{D}(A^*). \tag{19}$$

Consequently we also have

$$\langle x^*, T(t)Ax\rangle - \langle x^*, Ax\rangle = \langle A^*x^*, \int_0^t T(s)Ax\, ds\rangle \ \text{ for all } x \in \mathcal{D}(A),\ x^* \in \mathcal{D}(A^*). \tag{20}$$

Let $x \in \mathcal{D}(A)$ and $t \in [0, \tau]$ be given. Applying Lemma 11.12 to (19) and (20) we conclude that

$$\int_0^t T(s)x\, ds \in \mathcal{D}(A), \quad \int_0^t T(s)Ax\, ds \in \mathcal{D}(A),$$

and

$$T(t)x = x + A\int_0^t T(s)x\, ds, \tag{21}$$

$$T(t)Ax = Ax + A\int_0^t T(s)Ax\, ds. \tag{22}$$

Now put

$$V(t) = \int_0^t T(s)Ax\, ds - A\int_0^t T(s)x\, ds \ \text{ for all } t \in [0, \tau],$$

and observe that $V \in C([0, \tau]; X)$ by virtue of (21). Clearly $V(0) = 0$. Let $x^* \in \mathcal{D}(A)$ be given. Using (21) and (22) and some straightforward computations we find that

$$\frac{d}{dt}\langle x^*, V(t)\rangle = \langle A^*x^*, V(t)\rangle, \quad t \in [0, \tau].$$

As noted above, the only weak solution of $\dot{w} = Aw$, $w(0) = 0$ on $[0, \tau]$ is the zero solution so we can conclude that $V(t) = 0$ for all $t \in [0, \tau]$ which yields

$$\int_0^t T(s)Ax\,ds = A\int_0^t T(s)x\,ds \quad \text{for all } t \in [0, \tau]. \tag{23}$$

Using (21) and (23) we find that

$$T(h)x - x = \int_0^h T(s)Ax\,ds \quad \text{for all } h \in (0, \tau].$$

We can conclude that

$$\lim_{h \downarrow 0} \frac{T(h)x - x}{h} = \lim_{h \downarrow 0} \frac{1}{h}\int_0^h T(s)Ax\,ds = Ax.$$

It follows that $x \in \mathcal{D}(\hat{A})$ and the proof is complete. $\square$

Compact Semigroups

Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup. Let $t_0 > 0$ be given. If $T(t_0)$ is compact and $t > t_0$ then $T(t)$ is also compact because

$$T(t) = (T(t - t_0))T(t_0)$$

and the product of a bounded operator with a compact one is compact.

It follows that the set of all $t \in [0, \infty)$ such that $T(t)$ is compact is an interval of the form $[\tau, \infty)$ or (τ, ∞) with $\tau \geq 0$. Recall that the identity operator is compact if and only if X is finite-dimensional, so we want to be careful about making any assumptions that might imply $T(0)$ is compact. Semigroups having the property that $\{t \in [0, \infty) : T(t) \in \mathcal{C}(X; X)\}$ is a nonempty proper subset of $(0, \infty)$ are referred to as *eventually compact*. In 1953, Phillips gave an example which showed that the class of eventually compact semigroups is not stable under bounded perturbations of the infinitesimal generator. We shall focus here on linear C_0-semigroups having the property that $T(t)$ is compact for every $t > 0$. (This class of semigroups is, in fact, stable under bounded perturbations of the infinitesimal generator.)

Definition 11.13: Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup. We say that T is compact on $(0, \infty)$ provided that $T(t) \in \mathcal{C}(X; X)$ for every $t > 0$.

Our first result says that semigroups that are compact on $(0, \infty)$ are continuous in the uniform operator topology on $(0, \infty)$.

Lemma 11.14: Let X be a Banach space and $T : [0, \infty) \to \infty)$ be a linear C_0-semigroup. Assume that T is compact on $(0, \infty)$. Then T is continuous in the uniform operator topology on $(0, \infty)$.

Proof: Put

$$M = \sup\{\|T(s)\| : s \in [0,1]\}, \tag{24}$$

and notice that $M \geq 1$. Let $t, \epsilon > 0$ be given and put

$$U_t = \{T(t)x : x \in X,\ \|x\| \leq 1\},$$

$$\eta = \frac{\epsilon}{2(M+1)}. \tag{25}$$

The set U_t is totally bounded since its closure is compact. Therefore we may choose $x_1, x_2, \cdots, x_N \in \{x \in X : \|x\| \leq 1\}$ such that

$$U_t \subset \bigcup_{k=1}^{N} B_\eta(T(t)x_k) \tag{26}$$

Since each of the mappings $s \to T(s)x_k,\ k = 1, 2, \cdots, N$ is continuous at t we may choose $\delta \in (0,1]$ such that

$$\|T(t+h)x_k - T(t)x_k\| < \frac{\epsilon}{2} \text{ for all } h \in (0,\delta),\ k = 1, 2, \cdots, N. \tag{27}$$

Let $x \in X$ with $\|x\| \leq 1$ be given. In view of (26) we may choose $k \in \{1, 2, \cdots, N\}$ such that

$$\|T(t)x - T(t)x_k\| < \eta. \tag{28}$$

For $h \geq 0$ we have

$$\begin{aligned} T(t+h)x - T(t)x &= T(t+h)x - T(t+h)x_k + T(t+h)x_k - T(t)x_k \\ &\quad + T(t)x_k - T(t)x \\ &= T(h)[T(t)x - T(t)x_k] + (T(t+h)x_k - T(t)x_k) \\ &\quad + (T(t)x_k - T(t)x). \end{aligned} \tag{29}$$

Taking norms in (29) and using (24), (25), (27), (28) we find that for all $h \in [0,\delta)$

$$\|T(t+h)x - T(t)x\| < M\eta + \frac{\epsilon}{2} + \eta \leq \epsilon.$$

This establishes right continuity in the uniform operator topology at t.

Left continuity in the uniform operator topology at t follows easily from right continuity at $\frac{t}{2}$ and the observation

$$\|T(t-h) - T(t)\| \leq \left\|T\left(\frac{t}{2} - h\right)\right\| \cdot \left\|T\left(\frac{t}{2}\right) - T\left(\frac{t}{2} + h\right)\right\|. \quad \square$$

Before stating our next result about compact semigroups we make a simple, but useful, observation concerning compactness of resolvents of closed operators.

Proposition 11.15: Let X be a Banach space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A) \to X$ is linear and closed. Let $\lambda, \mu \in \rho(A)$ be given and assume that $R(\mu; A) \in \mathcal{C}(X; X)$. Then $R(\lambda; A) \in \mathcal{C}(X; X)$.

Proof: By Proposition 7.26, we have

$$R(\lambda; A) = R(\mu; A) + (\mu - \lambda)R(\lambda; A)R(\mu; A).$$

The conclusion now follows from the facts the product of a bounded operator with a compact one is compact and linear combinations of compact operators are compact. $\square$

Theorem 11.16: Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup with infinitesimal generator A. Assume that T is continuous in the uniform operator topology on $(0, \infty)$. The following three statements are equivalent.

(i) T is compact on $(0, \infty)$.

(ii) There exists $\lambda_0 \in \rho(A)$ such that $R(\lambda_0; A)$ is compact.

(iii) $R(\lambda; A)$ is compact for every $\lambda \in \rho(A)$.

Proof: In view of Proposition 11.15, we already know that (ii) $\Leftrightarrow$ (iii). Choose $M, \omega \in \mathbb{R}$ such that

$$\|T(t)\| \leq Me^{\omega t} \text{ for all } t \geq 0.$$

Assume first that (i) holds. Let $\lambda > \omega$ be given. By Lemma 9.6, we know that $\lambda \in \rho(A)$ and

$$R(\lambda; A)x = \int_0^\infty e^{-\lambda t} T(t)x \, dt \text{ for all } x \in X.$$

Since T is continuous in the uniform operator topology on $(0, \infty)$ (and because of our bound for $\|T(t)\|$), we know that the integral

$$\int_0^\infty e^{-\lambda t} T(t) \, dt \tag{30}$$

converges in the uniform operator topology. (The lack of continuity at the left end-point does not cause any difficulties.) It follows that

$$R(\lambda; A) = \int_0^\infty e^{-\lambda t} T(t) \, dt.$$

For each $\epsilon > 0$, put

$$R_\epsilon(\lambda; A) = \int_\epsilon^\infty e^{-\lambda t} T(t) \, dt \tag{31}$$

and observe that the integral converges in the uniform operator topology. Since $T(t)$ is compact for every $t > 0$ we can conclude that $R_\epsilon(\lambda; A)$ is compact for every $\epsilon > 0$. On the other hand, for every $x \in X$ with $\|x\| \leq 1$ and every $\epsilon > 0$ we have

$$\|R(\lambda; A)x - R_\epsilon(\lambda; A)x\| \leq \| \int_0^\epsilon e^{-\lambda t} T(t)x \, dt\| \leq \epsilon M. \tag{32}$$

It follows that $R(\lambda; A)$ is the uniform limit of compact operators and is therefore compact.

Assume now that (iii) holds. We have $\rho(A) \supset (\omega, \infty)$ and

$$R(\lambda; A) = \int_0^\infty e^{-\lambda s} T(s) \, ds \quad \text{for all } \lambda > \omega. \tag{33}$$

The integral in (33) converges in the uniform operator topology.

Let $t > 0$ and $\lambda > \omega$ be given. It follows from (33) that

$$\lambda R(\lambda; A)T(t) = \lambda \int_0^\infty e^{-\lambda s} T(t+s) \, ds,$$

and consequently

$$\lambda R(\lambda; A)T(t) - T(t) = \lambda \int_0^\infty e^{-\lambda s} [T(t+s) - T(t)] \, ds. \tag{34}$$

We therefore find that for every $\delta > 0$

$$\begin{aligned} \|\lambda R(\lambda; A)T(t) - T(t)\| &\leq \int_0^\delta \lambda e^{-\lambda s} \|T(t+s) - T(t)\| \, ds \\ &\quad + \int_\delta^\infty \lambda e^{-\lambda s} \|T(t+s) - T(t)\| \, ds \\ &\leq \sup_{0 \leq s \leq \delta} \|T(t+s) - T(t)\| \\ &\quad + \frac{2\lambda M e^{\omega(t+\delta)} e^{-\lambda\delta}}{\lambda - \omega}. \end{aligned} \tag{35}$$

Let $\epsilon > 0$ be given. We may choose $\delta > 0$ such that

$$\sup_{0 \leq s \leq \delta} \|T(t+s) - T(t)\| < \frac{\epsilon}{2} \tag{36}$$

Then we choose Λ such that

$$\frac{2\lambda M e^{\omega(t+\delta)} e^{-\lambda\delta}}{\lambda - \omega} < \frac{\epsilon}{2} \quad \text{for all } \lambda > \Lambda. \tag{37}$$

It follows that (35), (36), and (37) that

$$\|\lambda R(\lambda; A)T(t) - T(t)\| \to 0 \quad \text{as } \lambda \to \infty.$$

Since $\lambda R(\lambda; A)T(t)$ is compact for every $\lambda > \omega$ we conclude that $T(t)$ is compact. $\square$

Remark 11.17: Theorem 11.16 is not completely satisfactory because it does not characterize semigroups that are compact on $(0, \infty)$ solely in terms of properties of the infinitesimal generators A. The difficulty is in ensuring continuity in the uniform operator topology on $(0, \infty)$. I do not know of a nice characterization of semigroups that are continuous in the uniform operator topology on $(0, \infty)$ solely in terms of the generators of such semigroups.

Differentiable Semigroups

Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup with infinitesimal generator A. We know that for $x \in \mathcal{D}(A)$, the mapping $t \to T(t)X$ is differentiable on $[0, \infty)$. We are interested here in semigroups having the property that for every $x \in X$, the mapping $t \to T(t)x$ is differentiable on $(0, \infty)$. Such semigroups will be called *differentiable semigroups*. There are two (equivalent) basic approaches to making a definition of differentiable semigroup – one can require differentiability on $(0, \infty)$ of the mapping $t \to T(t)x$ for every $x \in X$ or one can require that $T(t) : X \to \mathcal{D}(A)$ for all $t > 0$. (It is also possible to study semigroups that are *eventually differentiable*, i.e. semigroups for which there exists $t_0 \geq 0$ such that for every $x \in X$, the mapping $t \to T(t)x$ is differentiable on (t_0, ∞). We shall not do so here. The interested reader is referred to Section 2.4 of Pazy.)

Definition 11.18: Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup. We say that T is differentiable on $(0, \infty)$ provided that for all $x \in X$, the mapping $t \to T(t)x$ is differentiable on $(0, \infty)$.

Proposition 11.19: Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup with infinitesimal generator A. The following two statements are equivalent.

(i) T is differentiable on $(0, \infty)$.

(ii) $T(t)[X] \subset \mathcal{D}(A)$ for all $t > 0$.

Moreover if (i) [and hence also (ii)] holds we have $T'(t)x = AT(t)x$ for every $x \in X$ and $t > 0$.

Proof: Assume that (i) holds and let $x \in X$ and $t > 0$ be given. Then we have

$$\begin{aligned} T'(t)x &= \lim_{h \to 0} \frac{T(t+h)x - T(t)x}{h} \\ &= \lim_{h \downarrow 0} \frac{T(t+h)x - T(t)x}{h} \\ &= \lim_{h \downarrow 0} \left(\frac{T(h) - I}{h} \right) T(t)x. \end{aligned}$$

It follows that $T(t)x \in \mathcal{D}(A)$ and $AT(t)x = T'(t)x$.

Assume now that (ii) holds and let $t > 0$ be given. Then we know that $T(\cdot)x$ is right differentiable at t with right derivative equal to $AT(t)x$. We need to show that the left derivative exists and also equals $AT(t)x$. Let $h \in (0, t)$ be given and observe

$$T(t)x - T(t-h)x = T\left(\frac{t}{2}\right)[T(h) - I]T\left(\frac{t}{2}\right)x.$$

It follows that

$$\lim_{h \downarrow 0} \frac{T(t-h)x - T(t)x}{-h} = T\left(\frac{t}{2}\right) AT\left(\frac{t}{2}\right) x = AT(t)x. \quad \square$$

Theorem 11.20: Let X be a Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup. Assume that T is differentiable on $(0, \infty)$. Then

(a) The mapping $t \to T(t)$ is of class C^∞ in the uniform operator topology on $(0, \infty)$.

(b) For every $t > 0$, $T(t) : X \to \bigcap_{n=1}^{\infty} \mathcal{D}(A^n)$.

(c) For every $t > 0$ and every $n \in \mathbb{N}$, $A^n T(t) \in \mathcal{L}(X; X)$ and

$$T^{(n)}(t) = A^n T(t) = \left(AT\left(\frac{t}{n}\right) \right)^n.$$

Before proving the theorem we observe that if an operator-valued function V is differentiable in the strong operator topology and the derivative is continuous in the uniform operator topology then V is differentiable in the uniform operator topology.

Lemma 11.21: Let X be a Banach space and $J \subset \mathbb{R}$ be an open interval. Assume that $V : J \to \mathcal{L}(X; X)$ is differentiable on J in the strong operator topology with derivative V', i.e.

$$\forall t \in J,\ x \in X, \quad V'(t)x = \lim_{h \to 0} \frac{V(t+h)x - V(t)x}{h}.$$

Assume further that the mapping $t \to V'(t)$ is continuous in the uniform operator topology. Then V is differentiable on J in the uniform operator topology, i.e.

$$\forall t \in J, \quad \lim_{h \to 0} \left\| \frac{V(t+h) - V(t)}{h} - V'(t) \right\| = 0. \tag{38}$$

Proof: For every $x \in X$ and every $t_1, t_2 \in J$ we have

$$V(t_2)x - V(t_1)x = \int_{t_1}^{t_2} V'(s)x\, ds.$$

Since V' is continuous in the uniform operator topology, the integral

$$\int_{t_1}^{t_2} V'(s)\, ds$$

exists in the uniform operator topology, so we must have

$$V(t_1) - V(t_2) = \int_{t-1}^{t_2} V'(s)\, ds \quad \text{for all } t_1, t_2 \in J. \tag{39}$$

It follows from (39) (and continuity of V') that (38) holds. $\square$

Proof of Theorem 11.20: Let $x \in X$ be given. For each $n \in \mathbb{N}$, consider the statement

(S_n) For every $t > 0$, $T(t)x \in \mathcal{D}(A^n)$, $T(\cdot)x$ is n-times differentiable at t,

$$T^{(n)}(t)x = A^n T(t)x = \left(AT\left(\frac{t}{n}\right)\right)^n x,$$

and $A^n T(t) \in \mathcal{L}(X; X)$.

We use induction to show that (S_n) holds for all $n \in \mathbb{N}$.

Base Case: Let $t > 0$ be given. By the definition of differentiable semigroup, we know that $T(\cdot)x$ is differentiable at t, and by Proposition 11.19, we know that $T(t)x \in \mathcal{D}(A)$. We also have

$$T'(t)x = \lim_{h \to 0} \frac{T(t+h)x - T(t)x}{h} = \lim_{h \downarrow 0} \left(\frac{T(h) - I}{h} \right) T(t)x = AT(t)x.$$

Since A is closed and $T(t) \in \mathcal{L}(X; X)$ we conclude that $AT(t)$ is closed. By the Closed Graph Theorem, we have $AT(t) \in \mathcal{L}(X; X)$.

Inductive Step: Let $n \in \mathbb{N}$ be given and assume that (S_n) holds. Let $t > 0$ and $h \in \mathbb{R}$ with $|h| < \dfrac{t}{n+1}$ be given. Then we have

$$\begin{aligned} T^{(n)}(t+h)x - T^{(n)}x &= A^n[T(t+h)x - T(t)x] \\ &= A^n T\left(\frac{nt}{n+1}\right)\left[T\left(\frac{t}{n+1} + h\right)x - T\left(\frac{t}{n+1}\right)x\right] \end{aligned} \tag{40}$$

Dividing by h, letting $h \to 0$, and using the fact A commutes with $T(\frac{t}{2})$ on $\mathcal{D}(A)$ we find that

$$\begin{aligned} T^{(n+1)}(t)x &= A^n T\left(\frac{nt}{n+1}\right) T'\left(\frac{t}{n+1}\right) x \\ &= A^n T\left(\frac{nt}{n+1}\right) AT\left(\frac{t}{n+1}\right) x \end{aligned} \tag{41}$$

Using the semigroup property and the fact that A commutes with $T(\cdot)$ on the domain of A we infer from (41) that

$$T^{(n+1)}(t)x = A^{n+1}T(t)x.$$

On the other hand, we know that

$$A^n T\left(\frac{nt}{n+1}\right) = \left(AT\left(\frac{t}{n+1}\right)\right)^n,$$

so we also conclude from (41) that

$$T^{(n+1)}(t) = \left(AT\left(\frac{t}{n+1}\right)\right)^{n+1}.$$

The Closed Graph Theorem implies that $A^{n+1}T(t) \in \mathcal{L}(X;X)$.

We conclude that (S_n) holds for all $n \in \mathbb{N}$.

We have shown that $T(\cdot)$ is of class C^∞ in the strong operator topology. It remains to establish infinite differentiability in the uniform operator topology. In view of Lemma 11.21, it suffices to show that for every $n \in \mathbb{N}$, $T^{(n)}$ is continuous in the uniform operator topology on $(0, \infty)$.

We choose $M \geq 1$ and $\omega \geq 0$ such that

$$\|T(t)\| \leq Me^{\omega t} \text{ for all } t \geq 0.$$

Let $n \in \{0\} \cup \mathbb{N}$, $x \in X$ with $\|x\| \leq 1$, and $t_1, t_2 \in (0, \infty)$ with $t_1 \leq t_2$ be given. Then we have

$$\begin{aligned} T^{(n)}(t_2)x - T^{(n)}(t_1)x &= \int_{t_1}^{t_2} A^{n+1}T(s)x\,ds \\ &= \int_{t_1}^{t_2} A^{n+1}T(t_1)T(s-t_1)x\,ds. \end{aligned} \tag{42}$$

Taking norms in (42) and then taking the supremum over all $x \in X$ with $\|x\| \leq 1$ we find that

$$\|T^{(n)}(t_2) - T^{(n)}(t_1)\| \leq (t_2 - t_1)Me^{\omega(t_2-t_1)}\|A^{(n+1)}T(t_1)\|.$$

It follows that $T^{(n)}$ is continuous in the uniform operator topology for every $n \in \{0\} \cup \mathbb{N}$.

This completes the proof $\square$.

If the generator A of a differentiable semigroup T is unbounded then $\|AT(t)\|$ blows up as $t \downarrow 0$. We shall show that if A is unbounded then

$$\limsup_{t\downarrow 0} t^{-1}\|AT(t)\| \geq e^{-1}.$$

There is no upper limit to how fast $\|AT(t)\|$ can blow up as $t \downarrow 0$.

In 1995, Renardy gave an example (in Hilbert space) of a linear operator A generating a linear C_0-semigroup that is differentiable on $(0,\infty)$ and an everywhere-defined bounded linear operator L such that the semigroup generated by $A+L$ fails to be differentiable. (In fact, it is not even eventually differentiable.) Subsequently, Doytchinov, Hrusa, and Watson gave a sharp growth condition on $\|AT(t)\|$ as $t \downarrow 0$ for a differentiable semigroup T with infinitesimal generator A which guarantees that the semigroup generated by $A+L$ will be differentiable on $(0,\infty)$ for every $L \in \mathcal{L}(X;X)$.

We close this section by stating a theorem of Pazy (and a corollary) that characterizes generators of differentiable semigroups in terms of spectral properties. In order to state these results, it is convenient to introduce a family of subsets of $\mathbb{C}$. Given $a \in \mathbb{R}$ and $b > 0$, put

$$\Sigma_{b,a} = \{\lambda \in \mathbb{C} : \mathrm{Re}(\lambda) > a - b\log|\mathrm{Im}(\lambda)|\}. \tag{43}$$

Theorem 11.22: Let X be a complex Banach space and let $M, \omega \in \mathbb{R}$ be given. Let $T : [0,\infty) \to \mathcal{L}(X;X)$ be a linear C_0-semigroup satisfying $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$ and having infinitesimal generator A. Then T is differentiable on $(0,\infty)$ if and only if for every $b > 0$ there exist constants $a \in \mathbb{R}$ and $C > 0$ such that $\rho(A) \supset \Sigma_{b,a}$ and

$$\|R(\lambda; A)\| \leq C|\mathrm{Im}(\lambda)| \quad \text{for all } \lambda \in \Sigma_{b,a} \text{ with } \mathrm{Re}(\lambda) \leq \omega.$$

Corollary 11.23: Let X be a complex Banach space and let $M, \omega \in \mathbb{R}$ be given. Let $T : [0,\infty) \to \mathcal{L}(X;X)$ be a linear C_0-semigroup satisfying $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$ and having infinitesimal generator A. Let $\mu > \omega$ be given and assume

$$\limsup_{|\tau|\to\infty} \log|\tau| \cdot \|R(\mu + i\tau; A)\| = 0.$$

Then T is differentiable on $(0,\infty)$. [The variable τ in the limit above is real.]

Proofs of Theorem 11.22 and Corollary 11.23 are given in Section 2.4 of Pazy.

Lecture Notes for Week 12 (First Draft)

Analytic Semigroups

Definition 12.1 (Complex Sector): For each $\phi \in (0, \pi)$ and $\omega \in \mathbb{R}$, the *open sector* of angle ϕ at ω is the subset of $\mathbb{C}$ defined by

$$\Delta_{\phi,\omega} = \{\lambda \in \mathbb{C} : \lambda \neq \omega,\ |\arg(\lambda - \omega)| < \phi\}.$$

Theorem 12.2: Let X be a complex Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup with infinitesimal generator A. Let $K > 0$ be given. Assume that T is differentiable on $(0, \infty)$ and that

$$\|AT(t)\| \leq \frac{K}{t} \quad \text{for all } t \in (0, 1].$$

Then there exists $\phi \in (0, \frac{\pi}{2})$ and an analytic mapping $\widetilde{T} : \Delta_{\phi,0} \to \mathcal{L}(X; X)$ such that (a), (b), and (c) below hold:

(a) $\widetilde{T}(t) = T(t)$ for all $t \in (0, \infty)$,

(b) $\widetilde{T}(z_1 + z_2) = \widetilde{T}(z_1)\widetilde{T}(z_2)$ for all $z_1, z_2 \in \Delta_{\phi,0}$,

(c) For every $x \in X$, we have $\widetilde{T}(z)x \to x$ as $z \to 0$, $z \in \Delta_{\phi,0}$.

Proof: Let t_0 and $N \in \mathbb{N}$ be given. Then, by Taylor's Theorem, for $t > t_0$ we have

$$T(t) = \sum_{n=0}^{N-1} \frac{(t - t_0)^n}{n!} A^n T(t_0) + R_{N-1}(t_0; t),$$

where

$$R_{N-1}(t_0; t) = \frac{1}{(N-1)!} \int_{t_0}^{t} (t - s)^{N-1} A^N T(s)\, ds. \tag{1}$$

We want to estimate the remainder. To this end, observe that

$$\begin{aligned} \|A^N T(s)\| &= \|AT\left(\frac{s}{N}\right)^N\| \\ &\leq \|AT\left(\frac{s}{N}\right)\|^N \qquad (2) \\ &\leq K^N \left(\frac{N}{s}\right) \quad \text{for } \frac{s}{N} \leq 1. \end{aligned}$$

Recall from calculus that
$$N!e^N \geq N^N. \tag{3}$$
It follows from (2) and (3) that
$$\begin{aligned}\|R_N(t_0;t)\| &\leq \frac{N^N K^N}{N!}\left(\frac{t-t_0}{t_0}\right)^N \\ &\leq \left(eK\left(\frac{t-t_0}{t_0}\right)\right)^N.\end{aligned}$$
We see that if
$$eK\left(\frac{t-t_0}{t_0}\right) < 1, \tag{4}$$
then
$$R_{N-1}(t_0,t) \to 0 \ \text{ as } N \to 0,$$
so that
$$T(t) = \sum_{n=0}^{\infty} \frac{(t-t_0)^n}{n!} A^n T(t_0). \tag{5}$$
For $z \in \mathbb{C}$ with
$$|z-t_0| < \frac{t_0}{eK} \tag{6}$$
the power series
$$\sum_{n=0}^{\infty} \frac{(z-t_0)^n}{n!} A^n T(t_0)$$
converges absolutely and we put
$$\widetilde{T}(z) = \sum_{n=0}^{\infty} \frac{(z-t_0)^n}{n!} A^n T(t_0). \tag{7}$$
Now choose $\psi \in (0, \frac{\pi}{2})$ such that
$$\psi < \frac{1}{eK}.$$
Then for every $z \in \Delta_{\psi,0}$ there exists $t_0 > 0$ such that (6) holds. Therefore we can use the series in (7) to define $\widetilde{T} : \Delta_{\psi,0} \to \mathcal{L}(X;X)$.

Notice that
$$\widetilde{T}(t) = T(t) \ \text{ for all } t > 0,$$
and
$$\begin{aligned}\frac{d}{dz}\widetilde{T}(z) &= \sum_{n=1}^{\infty} \frac{n(z-t_0)^{n-1}}{n!} A^n(t_0) = \sum_{n=0}^{\infty} \frac{(z-t_0)^n}{n!} A^{n+1} T(t_0) \\ &= A\widetilde{T}(z).\end{aligned}$$

To verify the semigroup property, we observe that for fixed $t > 0$ and $z \in \Delta_{\psi,0}$ satisfying (6) we have

$$\begin{aligned}
\widetilde{T}(t)\widetilde{T}(z) &= \sum_{n=0}^{\infty} \frac{(z-t_0)^n}{n!} A^n T(t_0+t) \\
&= \sum_{n=0} \infty \frac{[(z+t)-(t_0+t)]^n}{n!} A^n T(t_0+t) \\
&= \widetilde{T}(z+t).
\end{aligned}$$

The semigroup property on the full sector $\Delta_{\psi,0}$ follows from a standard stepping argument.

Now choose $\phi \in (0,\psi)$ and put

$$\Omega = \{z \in \mathbb{C} : 0 < \mathrm{Re}(z) \leq 1,\ |\arg(z)| \leq \phi\},$$

$$M = \sup\{|T(s)\| : s \in [0,1]\},$$

and

$$\epsilon = \frac{\phi}{\psi}.$$

Notice that

$$0 < \epsilon < 1.$$

Let $z \in \Omega$ be given and set

$$t_0 = \mathrm{Re}(z).$$

Then we have

$$\frac{|z-t_0|}{t_0} \leq \frac{\epsilon}{eK},$$

and consequently

$$\begin{aligned}
\|\widetilde{T}(z)\| &\leq \|T(t)\| + \sum_{n=1}^{\infty} \frac{|z-t_0|^n}{n!} \|A^n T(t_0)\| \\
&\leq \|T(t)\| + \sum_{n=1}^{\infty} \epsilon^n \\
&\leq M + \frac{\epsilon}{1-\epsilon} \quad \text{for all } z \in \Omega.
\end{aligned}$$

Now we are ready to establish (c). Let $x \in X$ and $t > 0$ be given. Then we have

$$\widetilde{T}(z)T(t)x = \widetilde{T}(z+t)x \to T(t)x \quad \text{as } z \to 0. \tag{8}$$

Using (8) together with the facts that $\|\widetilde{T}\|$ is bounded on Ω and $T(t)x \to x$ as $t \downarrow 0$ we find that (c) holds. This completes the proof. $\square$

Theorem 12.2: Let X be a complex Banach space and $\phi \in (0, \frac{\pi}{2})$ be given. Assume that $\widetilde{T} : \Delta_{\phi,0} \to \mathcal{L}(X;X)$ is analytic and that (i) and (ii) below hold:

(i) $\widetilde{T}(z_1 + z_2) = \widetilde{T}(z_1)\widetilde{T}(z_2)$ for all $z_1, z_2 \in \Delta_{\phi,0}$,

(ii) For every $x \in X$, we have $\widetilde{T}(z)x \to x$ as $z \to 0$, $z \in \Delta_{\phi,0}$.

Put $T(0) = I$ and $T(t) = \widetilde{T}(t)$ for all $t > 0$. Then T is a linear C_0 semigroup that is differentiable on $(0, \infty)$ and there is a constant $K > 0$ such that

$$\|AT(t)\| \leq \frac{K}{t} \quad \text{for all } t \in (0, 1],$$

where A is the infinitesimal generator of A.

Proof: That T is a linear C_0-semigroup and is differentiable on $(0.\infty)$ is immediate. Let $\alpha \in (0, \phi)$ be given and put

$$\Lambda_\alpha = \{z \in \mathbb{C} : 0 < \mathrm{Re}(z) \leq 2,\ |\arg(z)| \leq \alpha\}.$$

Using the Principle of Uniform Boundedness (as in the proof of Lemma x.x) we may choose $M_\alpha > 0$ such that

$$\|\widetilde{T}(z)\| \leq M \quad \text{for all } z \in \Lambda_\alpha. \tag{9}$$

Let $t \in (0, 1]$ be given and let C be the circle

$$C = \{z \in \mathbb{C} : |z - t| = (\sin \alpha)t\}$$

oriented in the counterclockwise direction. Then, by Cauchy's Theorem

$$AT(t) = T'(t) = \widetilde{T}'(t) = \frac{1}{2\pi i} \int_C \frac{\widetilde{T}(z)}{(z - t)^2}\, dz. \tag{10}$$

Taking norms in (10) and using (9) we see that

$$\|AT(t)\| \leq \frac{M_\alpha 2\pi t \sin \alpha}{2\pi((\sin \alpha)t)^2} = \frac{M_\alpha (\sin \alpha)^{-1}}{t}.$$

This completes the proof. $\square$

Definition 12.3: Let X be a real or complex Banach space. By an *analytic semigroup* we mean a linear C_0-semigroup $T : \mathcal{L}(X;X) \to \mathcal{L}(X;X)$ such that T is differentiable on $(0, \infty)$ and there exists a constant K such that

$$\|AT(t)\| \leq Kt^{-1} \quad \text{for all } t \in (0, 1].$$

Theorem 12.4: Let X be a complex Banach space and $\mathcal{D}(A) \subset X$. Assume that $\mathcal{D}(A)$ is dense and that $A : \mathcal{D}(A) \to x$ is linear and closed.

(a) Assume that there are constants $C, \omega \in \mathbb{R}$ such that

$$\rho(A) \supset \{\lambda \in \mathbb{C} : \mathrm{Re}(\lambda) > \omega\} = \Delta_{\frac{\pi}{2},\omega}$$

and

$$\|R(\lambda; A)\| \leq \frac{C}{|\lambda - \omega|} \quad \text{for all } \lambda \in \Delta_{\frac{\pi}{2},\omega}.$$

Then A generates an analytic semigroup.

(b) Assume that A generates an analytic semigroup T. Then there exist constants $C', \omega' \in \mathbb{R}$ and $\delta \in (0, \frac{\pi}{2})$ such that

$$\rho(A) \supset \{\lambda \in \mathbb{C} : |\arg(\lambda - \omega')| < \frac{\pi}{2} + \delta\} = \Delta_{\frac{\pi}{2}+\delta,\omega'}$$

and

$$\|R(\lambda; A)\| \leq \frac{C'}{|\lambda - \omega'|} \quad \text{for all } \lambda \in \Delta_{\frac{\pi}{2}+\delta,\omega'}.$$

Moreover, for $\phi \in (\frac{\pi}{2}, \frac{\pi}{2} + \delta)$ we have

$$T(t) = \frac{1}{2\pi i} \int_\Gamma e^{\lambda t} R(\lambda; A)\, d\lambda \quad \text{for all } t > 0,$$

where Γ is any curve from $\infty e^{-i\phi}$ to $\infty e^{i\phi}$ lying entirely in $\{\lambda \in \mathbb{C}\backslash\{\omega'\} : |\arg(\lambda - \omega')| \leq \phi\}$.

Theorem 12.5: Let X be a real or complex Banach space and $T : [0, \infty) \to \mathcal{L}(X; X)$ be a linear C_0-semigroup with infinitesimal generator A. Then T is an analytic semigroup if and only if there exist constants $C > 0$ and $\Lambda \geq 0$ such that for every $n \in \mathbb{N}$ we have

$$\|AR(\lambda; A)^{n+1}\| \leq \frac{C}{n\lambda^n} \quad \text{for all } \lambda > n\Lambda.$$

Some Additional Comments on Analytic Semigroups

We close our discussion of analytic semigroups with a few additional remarks.

- The class of analytic semigroups is stable under bounded perturbations of the infinitesimal generator. (This is a homework exercise.) In fact, one can perturb the generator of an analytic semigroup by a certain class of unbounded operators and still have a generator of an analytic semigroup. See, for example, Theorem 2.1 in Pazy.

- When A generates an analytic semigroup, there are improved regularity results for solutions of

$$\dot{u}(t) = Au(t) + f(t).$$

There will be a homework exercise on this topic. See the monograph of Yagi for an extensive discussion of such results.

- If A generates an analytic semigroup on a complex Banach space and $\mathrm{Re}(\lambda) < 0$ for all $\lambda \in \sigma(A)$ then one can define fractional powers of $-A$. See, for example, Section 2.6 of Pazy.

- There are some simplified generation theorems for analytic semigroups involving semi-inner products. There will be a homework exercise concerning this topic. We also give a result below for the Hilbert space case.

Proposition 12.6: Let X be a complex Hilbert space and $\mathcal{D}(A) \subset X$. Assume that $A : \mathcal{D}(A)$ is self-adjoint. Assume further that there exists $\beta \in \mathbb{R}$ such that

$$(Ax, x) \leq \beta \|x\|^2 \text{ for all } x \in \mathcal{D}(A).$$

Then A generates an analytic semigroup.

Proof: Without loss of generality, we assume that $\beta = 0$. Then we have

$$\rho(A) \supset \{\lambda \in \mathbb{C} : \mathrm{Re}(\lambda) > 0\} = \Delta_{\frac{\pi}{2},0}.$$

Let $\mu > 0$ and $\sigma \in \mathbb{R}$ be given and put

$$\lambda = \mu + i\sigma.$$

Then we have

$$((\lambda I - A)x, x) = \mu \|x\|^2 - (Ax, x) + i\sigma \|x\|^2 \text{ for all } x \in \mathcal{D}(A),$$

from which we can conclude that

$$|((\lambda - A)x, x)| \geq |\lambda| \cdot \|x\|^2 \text{ for all } x \in \mathcal{D}(A). \tag{11}$$

It follows from (11) that

$$\|R(\lambda; A)\| \leq \frac{1}{|\lambda|} \text{ for all } \lambda \in \mathbb{C} \text{ with } \mathrm{Re}(\lambda) > 0.$$

The desired conclusion follows from part (a) of Theorem 12.4. $\square$

Example 12.7 (Heat Equation): Let $X = L^2[0, 1]$ and put

$$\mathcal{D}(A) = \{u \in AC[0, 1] : u' \in AC[0, 1],\ u'' \in L^2[0, 1],\ u(0) = u(1) = 0\},$$

and define $A : \mathcal{D}(A) \to X$ by

$$Au = u'' \text{ for all } u \in \mathcal{D}(A).$$

It is left as an exercise to check that $\mathcal{D}(A^*) = \mathcal{D}(A)$, where A^* is the Hilbert adjoint of A. Since

$$\begin{aligned}\int_0^1 u''(x)\overline{v}(x)\,dx &= u'(x)\overline{v}(x)\Big|_0^1 - \int_0^1 u'(x)\overline{v}'(x)\,dx \\ &= -u(x)\overline{v}'(x)\Big|_0^1 + \int_0^1 u(x)\overline{v}''(x)\,dx \qquad (12) \\ &= \int_0^1 u(x)\overline{v}''(x)\,dx \quad \text{for all } u, v \in \mathcal{D}(A),\end{aligned}$$

we conclude that A is self-adjoint.

Using (13) with $u = v$ we find that

$$(Au, u) = -\int_0^1 |u'(x)|^2\,dx \leq 0 \quad \text{for all } u \in \mathcal{D}(A).$$

It follows from Proposition 12.6 that A generates an analytic semigroup.

Fourier Transforms

We begin with some useful notation.

Definition 12.8 (Multi-indices): Let $M_n = (\mathbb{N} \cup \{0\})^n$. The elements of M_n are called multi-indices

Definition 12.9: Let $\alpha, \beta \in M_n$ be given. We say that $\alpha \leq \beta$ provided that

$$\forall j \in \{1, 2, \cdots, n\}, \quad \text{we have } \alpha_j \leq \beta_j.$$

Definition 12.10: Let $\alpha, \beta \in M_n$ and $x \in \mathbb{R}^n$ be given. We define

- $|\alpha| = \sum_{j=1}^{n} \alpha_j,$
- $\alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!,$
- $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$

Definition 12.11: For $\alpha, \beta \in M_n$ with $\alpha \leq \beta$ we put

$$\binom{\beta}{\alpha} = \frac{\beta!}{(\beta - \alpha)!}.$$

Definition 12.12: Let $\alpha \in M_n$ and $f \in C^{|\alpha|}(\mathbb{R}^n)$ be given. We define

$$(D^\alpha f)(x) = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}} f(x).$$

Observe that with the above definitions, the Binomial Theorem takes the form

$$(x+y)^\beta = \sum_{\alpha \leq \beta} \binom{\beta}{\alpha} x^\alpha y^{\beta-\alpha} \quad \text{for all } x, y \in \mathbb{R}^n,\ \beta \in M_n,$$

and, for sufficiently smooth functions $f, g : \mathbb{R}^n \to \mathbb{C}$, Leibniz's product rule takes the form

$$D^\beta(fg) = \sum_{\alpha \leq \beta} (D^\alpha f)(D^{\beta-\alpha} g).$$

In order that we can write formulas involving Fourier transforms in a clean way, it is convenient to give a name to the mapping $x \to x^\alpha$.

Definition 12.13: Let $\alpha \in M_n$ be given. We define $P_\alpha : \mathbb{R}^n \to \mathbb{C}$ by

$$P_\alpha(x) = x^\alpha \quad \text{for all } x \in \mathbb{R}^n.$$

[Actually P_α takes values in $\mathbb{R}$, but I have taken the codomain to be $\mathbb{C}$ because the scalar field will always be $\mathbb{C}$ when we use this mapping.]

Definition 12.14: Let $f \in L^1(\mathbb{R}^n)$ be given. We define the *Fourier transform* $\widehat{f} : \mathbb{R}^n \to \mathbb{C}$ of f by

$$\widehat{f}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, dx \quad \text{for all } x \in \mathbb{R}^n. \tag{13}$$

Remark 12.15: There are several other conventions used in defining the Fourier transform. These involve placing factors of 2π in different places and/or changing the sign in the exponential. Also some authors (e.g., Rudin) normalize Lebesgue measure to absorb the factor $(2\pi)^{-\frac{n}{2}}$ in front of the integral. The bottom line is that one must be very careful to check what convention is being employed in situations where the numerical values of constants could be important.

Proposition 12.16: Let $f \in L^1(\mathbb{R}^n)$ be given. Then $\widehat{f}$ is bounded and continuous on $\mathbb{R}^n$.

Proof: Let $\xi \in \mathbb{R}^n$ be given and let $\{\xi^{(k)}\}_{k=1}^\infty$ be a sequence in $\mathbb{R}^n$ such that $\xi^{(k)} \to \xi$ as $k \to \infty$. Since

$$|e^{-ix\cdot\xi^{(k)}} f(x)| = |e^{-ix\cdot\xi} f(x)| = |f(x)| \quad \text{for all } x \in \mathbb{R}^n$$

(and since $f \in L^1(\mathbb{R}^n)$) the Lebesgue Dominated Convergence Theorem implies that

$$\widehat{f}(\xi^{(k)}) \to \widehat{f}(\xi).$$

It follows immediately from (13) that

$$|\widehat{f}(\xi)| \leq \frac{1}{(2\pi)^{\frac{n}{2}}} \|f\|_1 \text{ for all } \xi \in \mathbb{R}^n,$$

and consequently

$$\|\widehat{f}\|_\infty \leq \frac{1}{(2\pi)^{\frac{n}{2}}} \|f\|_1 \text{ for all } f \in L^1(\mathbb{R}^n). \tag{14}$$

This completes the proof. $\square$

An extremely useful feature of the Fourier transform is that it converts differentiation into multiplication by polynomials. This will allow us to use elementary facts from basic algebra to deduce powerful results concerning differential operators. To get an idea of how this will work, we shall perform some formal computations first and then make assumptions that justify these computations later on.

Suppose that $f : \mathbb{R}^n$ is "really nice" (meaning that is very smooth and that f and derivatives of f go to zero rapidly at infinity). Let $\alpha \in M_n$ be given. Then we may write

$$\widehat{f}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, dx.$$

Applying D^α and differentiating under the integral sign, we find that

$$(D^\alpha \widehat{f})(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} (-i)^{|\alpha|} x^\alpha e^{-ix\cdot\xi} f(x)\, dx.$$

We can rewrite this expression (without using variables) as

$$D^\alpha \widehat{f} = (-i)^{|\alpha|} (P_\alpha f)\widehat{}.$$

Using integration by parts, we find that

$$\begin{aligned} (D^\alpha f)\widehat{}(\xi) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} D^\alpha f(x)\, dx \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} (-1)^{|\alpha|} (-i)^{|\alpha|} \xi^\alpha e^{-ix\cdot\xi} f(x)\, dx \end{aligned}$$

We can rewrite this as

$$(D^\alpha f)\widehat{} = i^{|\alpha|} P_\alpha \widehat{f}.$$

Lecture Notes for Week 13 (First Draft)

Schwartz Space

In order to validate the computations involving Fourier transforms and derivatives, we want a function space that is invariant under differentiation and multiplication by polynomials, and has the property that its members are integrable.

Definition 13.1: Let

$$\mathcal{S}(\mathbb{R}^n) = \{\phi \in C^\infty(\mathbb{R}^n) : \forall \alpha, \beta \in M_n,\ \sup_{x\in\mathbb{R}^n} |x^\beta D^\alpha \phi(x)| < \infty\}.$$

$\mathcal{S}(\mathbb{R}^n)$ is known as *Schwartz space* or the space of *rapidly decreasing functions.*

Observe that if $\phi \in C^\infty(\mathbb{R}^n)$ then $\phi \in \mathcal{S}(\mathbb{R}^n)$ if and only if

$$P_\beta D^\alpha \phi \in L^\infty(\mathbb{R}^n) \ \text{ for all } \alpha, \beta \in M_n.$$

Moreover, it is clear that

$$\forall \phi \in \mathcal{S}(\mathbb{R}^n),\ \alpha, \beta \in M_n, \quad P_\beta D^\alpha \phi \in \mathcal{S}(\mathbb{R}^n).$$

Observe also that for every $p \in [1, \infty]$ we have

$$C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n),$$

where $C_c^\infty(\mathbb{R}^n)$ is the space of C^∞-functions with compact support. A simple example of a function ψ that belongs to $\mathcal{S}(\mathbb{R}^n)$, but not to $C_c^\infty(\mathbb{R}^n)$ is provided by

$$\psi(x) = e^{-|x|^2} \ \text{ for all } x \in \mathbb{R}^n.$$

Since $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $p \in [1, \infty)$ we see that $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for each $p \in [1, \infty)$. We shall topologize $\mathcal{S}(\mathbb{R}^n)$ a bit later on.

Lemma 13.2: Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $\alpha \in M_n$ be given. Then we have

(a) $\widehat{\phi} \in \mathcal{S}(\mathbb{R}^n)$,

(b) $D^\alpha \widehat{\phi} = (-i)^{|\alpha|} (P_\alpha \phi)\widehat{\ }$,

(c) $(D^\alpha \phi)\widehat{\ } = i^{|\alpha|} P_\alpha \widehat{\phi}$.

Lemma 13.3 (Riemann-Lebesgue): Let $f \in L^1(\mathbb{R}^n)$ be given. Then $\widehat{f}$ vanishes at infinity, i.e.

$$\widehat{f}(\xi) \to 0 \quad \text{as } |\xi| \to \infty.$$

Proof: Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$ we may choose a sequence $\{\phi_k\}_{k=1}^{\infty}$ in $\mathcal{S}(\mathbb{R}^n)$ such that $\|\phi_k - f\|_1 \to 0$ as $k \to \infty$. By (14) from Week 12, we have

$$\|\widehat{\phi_k} - \widehat{f}\|_\infty \to 0 \quad \text{as } k \to \infty.$$

The conclusion thus follows from the facts that $\widehat{\phi_k} \in \mathcal{S}(\mathbb{R}^n)$ and functions in $\mathcal{S}(\mathbb{R}^n)$ vanish at infinity. $\Box$

Lemma 13.4: Let $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$ be given. Then

$$\int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{\phi}(\xi)\psi(\xi)\,d\xi = \int_{\mathbb{R}^n} \phi(x+y)\widehat{\psi}(y)\,dy.$$

Proof: Using the definition of Fourier transform we find that

$$\begin{aligned}
\int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{\phi}(\xi)\psi(\xi)\,d\xi &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}e^{-iz\cdot\xi}\phi(z)\psi(\xi)\,dz\,d\xi \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(z-x)\cdot\xi}\phi(z)\psi(\xi)\,d\xi\,dz \\
&= \int_{\mathbb{R}^n} \widehat{\psi}(x-z)\phi(z)\,dz \\
&= \int_{\mathbb{R}^n} \phi(x+y)\widehat{\psi}(y)\,dy. \quad \Box
\end{aligned}$$

Theorem 13.5 (Fourier Inversion): Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ be given. Then

$$\phi(x) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{\phi}(\xi)\,d\xi \quad \text{for all } x \in \mathbb{R}^n.$$

Proof: Let $\psi \in \mathcal{S}(\mathbb{R}^n)$ be given. For every $\epsilon > 0$ define $\psi_\epsilon : \mathbb{R}^n \to \mathbb{C}$ by

$$\psi_\epsilon(x) = \psi(\epsilon x) \quad \text{for all } x \in \mathbb{R}^n.$$

Let us compute $\widehat{\psi_\epsilon}$:

$$\begin{aligned}
\widehat{\psi_\epsilon}(y) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot y}\psi(\epsilon x)\,dx \\
&= \frac{\epsilon^{-n}}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i\frac{z}{\epsilon}\cdot y}\psi(z)\,dz \qquad (1) \\
&= \epsilon^{-n}\widehat{\psi}\left(\frac{y}{\epsilon}\right) \quad \text{for all } y \in \mathbb{R}^n.
\end{aligned}$$

Let $x \in \mathbb{R}^n$ be given. Using Lemma 13.4 and (1) we find that

$$\begin{aligned}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{\phi}(\xi)\psi(\epsilon\xi)\,d\xi &= \epsilon^{-n}\int_{\mathbb{R}^n}\phi(x+y)\widehat{\psi}\left(\frac{y}{\epsilon}\right)\,dy \\ &= \int_{\mathbb{R}^n}\phi(x+\epsilon z)\widehat{\psi}(z)\,dz.\end{aligned} \tag{2}$$

Letting $\epsilon \downarrow 0$ in (2), we obtain

$$\psi(0)\int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{\phi}(\xi)\,d\xi = \phi(x)\int_{\mathbb{R}^n}\widehat{\psi}(z)\,dz. \tag{3}$$

In order to complete the proof, we simply need to substitute one cleverly chosen function $\psi \in \mathcal{S}(\mathbb{R}^n)$ into (3). We shall make use of the following standard facts that are not too difficult to verify. (A proof of the claim below will be a homework exercise.)

Claim: Let

$$\psi(x) = e^{-\frac{1}{2}|x|^2} \quad \text{for all } x \in \mathbb{R}^n.$$

Then

$$\widehat{\psi} = \psi \quad \text{and} \quad \int_{\mathbb{R}^n}\widehat{\psi} = (2\pi)^{\frac{n}{2}}.$$

Substituting the function ψ from the claim (and observing that $\psi(0) = 1$) into (3) we obtain

$$\int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{\phi}(\xi)\,d\xi = (2\pi)^{\frac{n}{2}}\phi(x). \quad \square$$

Taking the Fourier transform of a function twice almost returns the original function, but not quite. The result is a "reflection" of the original function. For this reason, it is helpful to introduce a reflection operator $\check{\ }$.

Definition 13.6 (Reflection): Given $\phi \in \mathcal{S}(\mathbb{R}^n)$, define $\check{\phi} \in \mathcal{S}(\mathbb{R}^n)$ by

$$\check{\phi}(x) = \phi(-x) \quad \text{for all } x \in \mathbb{R}^n.$$

Remark 13.7:

(a) The Fourier inversion formula can be rewritten as

$$\check{\phi} = \widehat{\widehat{\phi}} \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

(b) It is also useful to observe that

$$\check{\widehat{\phi}} = \widehat{\check{\phi}} \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Lemma 13.8: Let $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ be given. Then

$$\int_{\mathbb{R}^n} \phi\widehat{\psi} = \int_{\mathbb{R}^n} \widehat{\phi}\psi.$$

Proof: Put $x = 0$ in Lemma 13.4. $\square$

Proposition 13.9: Let $\psi \in \mathcal{S}(\mathbb{R}^n)$ be given. Then

$$\overline{\psi} = \widehat{\widehat{\overline{\psi}}}.$$

Lemma 13.10 (Parseval's Relation): Let $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ be given. Then

$$\int_{\mathbb{R}^n} \phi\overline{\psi} = \int_{\mathbb{R}^n} \widehat{\phi}\overline{\widehat{\psi}}.$$

Corollary 13.11: Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ be given. Then

$$\|\phi\|_2 = \|\widehat{\phi}\|_2,$$

where $\|\cdot\|_2$ is the norm on $L^2(\mathbb{R}^n)$.

Example 13.12: Let $f \in \mathcal{S}(\mathbb{R}^n)$ be given. Consider the equation

$$-\Delta u(x) + u(x) = f(x) \ \text{ for all } x \in \mathbb{R}^n. \tag{4}$$

Here Δu denotes the Laplacian of u, i.e.

$$\Delta u(x) = \sum_{k=1}^{n} \frac{\partial^2 u}{\partial x_k^2}(x).$$

Let us look for solutions $u \in \mathcal{S}(\mathbb{R}^n)$ of (4). Observe that for $u \in \mathcal{S}(\mathbb{R}^n)$ we have

$$(\Delta u)\widehat{\ }(\xi) = -|\xi|^2 \ \text{ for all } \xi \in \mathbb{R}^n.$$

It is convenient to put

$$P(\xi) = 1 + |\xi|^2 \ \text{ for all } \xi \in \mathbb{R}^n.$$

Then $u \in \mathcal{S}(\mathbb{R}^n)$ satisfies (4) if and only if

$$P\widehat{u} = \widehat{f},$$

or

$$\widehat{u} = \frac{\widehat{f}}{P}.$$

It is easy to see that

$$\frac{\widehat{f}}{P} \in \mathcal{S}(\mathbb{R}^n).$$

We conclude that (4) has exactly one solution $u \in \mathcal{S}(\mathbb{R}^n)$. Moreover the solution is given by

$$\check{u} = \left(\frac{\widehat{f}}{p}\right)^{\widehat{}}.$$

Tempered Distributions

We now endow $\mathcal{S}(\mathbb{R}^n)$ with a natural topology. For each $N \in \mathbb{N} \cup \{0\}$ put

$$|||\phi|||_N = \sum_{|\alpha|,|\beta| \le N} \|P_\beta D^\alpha \phi\|_\infty \ \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Then $(|||\cdot|||_N | N \in \{0\} \cup \mathbb{N})$ is a separating family of seminorms on $\mathcal{S}(\mathbb{R}^n)$. (In fact each $|||\cdot|||_N$ is actually a norm.) This family of seminorms induces a topology that turns $\mathcal{S}(\mathbb{R}^n)$ into a locally convex topological vector space. The topology is generated by the translation invariant metric

$$\rho(\phi, \psi) = \sum_{k=0}^{\infty} \frac{2^{-k}|||\phi - \psi|||_k}{1 + |||\phi - \psi|||_k} \ \text{ for all } \phi, \psi \in \mathcal{S}(\mathbb{R}^n).$$

Remark 13.13: Let a sequence $\{\phi_k\}_{k=1}^\infty$ in $\mathcal{S}(\mathbb{R}^n)$ and $\phi \in \mathcal{S}(\mathbb{R}^n)$ be given. Then $\phi_k \to \phi$ as $k \to \infty$ in $\mathcal{S}(\mathbb{R}^n)$ if and only if

$$\forall \alpha, \beta \in M_n, \quad P_\beta D^\alpha \phi_k \to P_\beta D^\alpha \phi \text{ uniformly on } \mathbb{R}^n \text{ as } k \to \infty.$$

It is straightforward to show that $(\mathcal{S}(\mathbb{R}^n), \rho)$ is complete.

Lemma 13.14: Let X be a Banach space and assume that $L : \mathcal{S}(\mathbb{R}^n) \to X$ is linear. Then L is continuous if and only if there exist $N \in \mathbb{N} \cup \{0\}$ and $K \in \mathbb{R}$ such that

$$\|L\phi\|_X \le K|||\phi|||_N \ \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Definition 13.15: By a *tempered distribution*, we mean a continuous linear functional $u : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$. The set of all tempered distributions will be denoted by $\mathcal{S}'(\mathbb{R}^n)$.

In the context of tempered distributions, the elements of $\mathcal{S}(\mathbb{R}^n)$ are often referred to as *test functions*. We write $\langle \cdot, \cdot \rangle : \mathcal{S}'(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ for the duality pairing. Given $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\phi \in \mathcal{S}(\mathbb{R}^n)$ we write either $u(\phi)$ or $\langle u, \phi \rangle$ for the value of u at ϕ.

Example 13.16: A nonnegative Borel measure μ on $\mathbb{R}^n$ is said to be of *slow growth* provided that there exists $N \in \mathbb{N}$ such that

$$\int_{\mathbb{R}^n} \frac{d\mu(x)}{(1+|x|^2)^N} < \infty.$$

Every such measure induces a tempered distribution l_μ through the formula

$$\langle l_\mu, \phi \rangle = \int_{\mathbb{R}^n} \phi d\mu \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

When there is no danger of confusion, it is customary to use the same symbol "μ" for both the measure and the associated tempered distribution.

In order to define operations on tempered distributions, we are going to need continuity of basic operations on functions in $\mathcal{S}(\mathbb{R}^n)$. It is useful to observe that $\mathcal{S}(\mathbb{R}^n)$ is invariant under multiplication by by a class of functions that is somewhat larger than the class of polynomials. The following definition is convenient.

Definition 13.17: Let

$$\mathcal{PGD}(\mathbb{R}^n) = \{\psi \in C^\infty(\mathbb{R}^n) : \forall \alpha \in M_n,\ \exists N \in \mathbb{N},\ \lim_{|x|\to\infty} |x|^{-N}|D^\alpha \psi(x)| = 0\}.$$

Functions in $\mathcal{PGD}(\mathbb{R}^n)$ are said to be of *slow growth at infinity.*

It is immediate that $P_\alpha \in \mathcal{PGD}(\mathbb{R}^n)$ for every $\alpha \in M_n$. There does not seem to be a "standard notation" for this space of functions. I chose $\mathcal{PGD}$ with *Polynomial Growth of Derivatives* in mind. Using Leibniz product rule it is easy to see that if $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $\psi \in \mathcal{PGD}(\mathbb{R}^n)$ then $\psi\phi \in \mathcal{S}(\mathbb{R}^n)$.

Proposition 13.18: Let $\alpha \in M_n$ and $\psi \in \mathcal{PGD}(\mathbb{R}^n)$ be given. Then the mappings

- $\phi \to \widehat{\phi}$
- $\phi \to \check{\phi}$
- $\phi \to \psi\phi$
- $\phi \to D^\alpha \phi$

are continuous from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$.

Let $p \in [1,\infty]$ and $f \in L^p(\mathbb{R}^n)$ be given. Then the linear functional $L_f : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ defined by

$$L_f(\phi) = \int_{\mathbb{R}^n} f\phi \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n),$$

is a tempered distribution. Moreover the mapping $f \to L_f$ is linear and injective from $L^p(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$. [Of course, $f \to L_f$ is linear and injective from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$.]

We want to extend the notions of differentiation, Fourier transform, etc. to tempered distributions. The philosophy behind taking an operation that is defined for smooth functions and extending the definition of this operation to tempered distributions is that operation should commute with the natural injection $f \to L_f$. For example to define the Fourier transform of a tempered distribution, we observe that if $f \in \mathcal{S}(\mathbb{R}^n)$ then (using Lemma 13.8) we have

$$\begin{aligned} \langle L_{\widehat{f}}, \phi\rangle &= \int_{\mathbb{R}^n} \widehat{f}\phi = \int_{\mathbb{R}^n} f\widehat{\phi} \\ &= \langle L_f, \widehat{\phi}\rangle \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n). \end{aligned}$$

It is therefore natural to define the Fourier transform $\widehat{u} \in \mathcal{S}'(\mathbb{R}^n)$ of a tempered distribution $u \in \mathcal{S}'(\mathbb{R}^n)$ by

$$\langle \widehat{u}, \phi\rangle = \langle u, \widehat{\phi}\rangle \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

To define the reflection $\breve{u}$ of a tempered distribution u, we observe that for $f \in \mathcal{S}(\mathbb{R}^n)$ we have

$$\langle L_{\check{f}}, \phi\rangle = \int_{\mathbb{R}^n} \check{f}\phi = \int_{\mathbb{R}^n} f\check{\phi} \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

(Notice that there is no minus sign associated with changing variables in the integrals above because when putting $y_k = -x_k$ we get $dy_k = -dx_k$ but then we also must interchange the upper and lower limits of integration for the k^{th} variable.) This makes it natural to define

$$\langle \breve{u}, \phi\rangle = \langle u, \check{\phi}\rangle \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Let $\alpha \in M_n$, and $\psi \in \mathcal{S}(\mathbb{R}^n)$ be given In order to define $D^\alpha u$, $P_\alpha u$ and ψu, we observe that

$$\langle D^\alpha f, \phi\rangle = \int_{\mathbb{R}^n} (D^\alpha \phi)\phi = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(D^\alpha \phi) \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n). \tag{5}$$

$$\int_{\mathbb{R}^n} (P_\alpha f)\phi = \int_{\mathbb{R}^n} f(P_\alpha \phi) \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n),$$

$$\int_{\mathbb{R}^n} (\psi f)\phi = \int_{\mathbb{R}^n} f(\psi\phi) \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Integration by parts was used to obtain (5). We summarize these ideas in the following definition.

Definition 13.19: Let $\alpha \in M_n$, $\psi \in \mathcal{PGD}(\mathbb{R}^n)$, and $u \in \mathcal{S}'(\mathbb{R}^n)$ be given. We define $\psi u, D^\alpha u$, $\widehat{u}$, $\breve{u} \in \mathcal{S}'(\mathbb{R}^n)$ by

(a) $\langle \psi u, \phi\rangle = \langle u, \psi\phi\rangle$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$,

(b) $\langle D^\alpha u, \phi\rangle = (-1)^{|\alpha|}\langle u, D^\alpha \phi\rangle$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$,

(c) $\langle \widehat{u}, \phi \rangle = \langle u, \widehat{\phi} \rangle$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$,

(d) $\langle \check{u}, \phi \rangle = \langle u, \check{\phi} \rangle$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$.

We refer to $D^\alpha u$ as a *distributional derivative* (or *weak derivative*) of u and to $\widehat{u}$ as the *Fourier transform* of u.

Theorem 13.20: Let $\alpha \in M_n$ and $u \in \mathcal{S}'(\mathbb{R}^n)$ be given. Then

(a) $(\widehat{u})\widehat{\ } = \check{u}$,

(b) $(\widehat{u})\check{\ } = (\check{u})\widehat{\ }$,

(c) $D^\alpha \widehat{u} = (-i)^{|\alpha|} (P_\alpha u)\widehat{\ }$,

(d) $(D^\alpha u)\widehat{\ } = i^{|\alpha|} P_\alpha \widehat{u}$.

Example 13.21 (Dirac Delta Function): Consider the tempered distribution δ_0 defined by

$$\langle \delta_0, \phi \rangle = \phi(0) \ \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Let's find the Fourier transform of δ_0. Using Definitions 13.19 and 12.14 we see that

$$\langle \widehat{\delta_0}, \phi \rangle = \langle \delta_0, \widehat{\phi} \rangle = \widehat{\phi}(0) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \phi \ \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

It is customary to write this formula as

$$\widehat{\delta_0} = \frac{1}{(2\pi)^{\frac{n}{2}}}.$$

Observe that

$$\check{\delta_0} = \delta_0,$$

so using Theorem 13.20, we see that

$$\delta_0 = (\widehat{\delta_0})\widehat{\ } = \frac{1}{(2\pi)^{\frac{n}{2}}} \widehat{1},$$

which can be rewritten as

$$\widehat{1} = (2\pi)^{\frac{n}{2}} \delta_0.$$

[Here I have made a slight (and very common) abuse of notation by using the same symbol "1" to indicate the number 1 and the constant function whose only value is 1.]

Example 13.22 (Heaviside Function): Let $n = 1$ and define $H : \mathbb{R} \to \mathbb{C}$ by

$$H(x) = \begin{cases} 0 \text{ for } x < 0, \\ 1 \text{ for } x \geq 0. \end{cases}$$

Let us compute the distributional derivative H' of H. For $\phi \in \mathcal{S}(\mathbb{R})$ we have

$$\begin{aligned}\langle H', \phi\rangle &= -\langle H, \phi'\rangle \\ &= -\int_{-\infty}^{\infty} H(x)\phi'(x)\,dx = -\int_{0}^{\infty} \phi'(x)\,dx \\ &= \phi(0) = \langle \delta_0, \phi\rangle,\end{aligned}$$

where δ_0 is the one-dimensional version of the Dirac Delta from Example 13.21. In other words, the distributional derivative of the Heaviside step function is the Dirac delta, i.e. $H' = \delta_0$.

Example 13.23 (Fourier Transform on $L^2(\mathbb{R}^n)$.) Let $f \in L^2(\mathbb{R}^n)$ be given and consider the associated tempered distribution $L_f$defined by

$$\langle L_f, \phi\rangle = \int_{\mathbb{R}^n} f\phi \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Then we have

$$\langle \widehat{L_f}, \phi\rangle = \langle L_f, \widehat{\phi}\rangle = \int_{\mathbb{R}^n} f\widehat{\phi} \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Using Holder's Inequality and Corollary 13.11, we find that

$$|\langle \widehat{L_f}, \phi\rangle| \leq \|f\|_2\|\widehat{\phi}\|_2 = \|f\|_2\|\phi\|_2 \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

By the Riesz Representation Theorem (and the fact that $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$), there is exactly one function $g \in L^2(\mathbb{R}^n)$ such that

$$\langle \widehat{L_f}, \phi\rangle = \int_{\mathbb{R}^n} g\phi \quad \text{for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

Moreover this function g satisfies $\|g\|_2 = \|f\|_2$. (It is customary to write $\widehat{f} = g$.) In other words, the Fourier transform can be defined in a natural way as an isometric linear mapping from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$. This mapping is surjective and consequently it is unitary. The content of this example is known as Plancherel's Theorem.

Remark 13.24: Since (for every $p \in [1, \infty]$) $L^p(\mathbb{R}^n)$ can be identified with a subset of $\mathcal{S}'(\mathbb{R}^n)$ we can talk about the behavior of the Fourier transform on $L^p(\mathbb{R}^n)$. As before, let us write $\mathcal{F}(f) = \widehat{f}$. It is a classical result of Titchmarsh that for

$$p \in [1, 2], \quad q = \frac{p}{p-1}$$

the Fourier transform $\mathcal{F}$ maps $L^p(\mathbb{R}^n)$ into $L^q(\mathbb{R}^n)$ continuously and injectively. Moreover this mapping is surjective if and only if $p = q = 2$.

Convolution

Definition 13.25: Let $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ be given. The convolution of ϕ with ψ is the function $\phi * \psi : \mathbb{R}^n \to \mathbb{C}$ defined by

$$(\phi * \psi)(x) = \int_{\mathbb{R}^n} \phi(y)\psi(x-y)\, dy \quad \text{for all } x \in \mathbb{R}^n.$$

Lemma 13.26: Let $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ and $\alpha \in M_N$ be given. Then $\phi * \psi = \psi * \phi$, $\phi * \psi \in \mathcal{S}(\mathbb{R}^n)$, and

$$D^\alpha(\phi * \psi) = (D^\alpha \phi) * \psi = \phi * (D^\alpha \psi). \tag{6}$$

Proof: Observe first that

$$(\phi * \psi)(x) = \int_{\mathbb{R}^n} \phi(x-z)\psi(z)\, dz = (\psi * \phi)(x) \quad \text{for all } x \in \mathbb{R}^n.$$

It follows easily that $\phi * \psi \in C^\infty(\mathbb{R}^n)$ and that (6) holds. We need to verify that $\phi * \psi \in \mathcal{S}(\mathbb{R}^n)$. The definition of $\mathcal{S}(\mathbb{R}^n)$ implies that

$$\forall \alpha \in M_n, \quad P_\alpha \psi \in L^1(\mathbb{R}^n).$$

Let us put

$$C_\alpha(\psi) = \|P_\alpha \psi\|_1 \quad \text{for all } \alpha \in M_n$$

and let $\beta \in M_n$ and $x \in \mathbb{R}^n$ be given. Then we have

$$\begin{aligned}
|x^\beta (\phi * \psi)(x)| &\leq \int_{\mathbb{R}^n} |x^\beta D^\alpha \phi(x-y)\psi(y)|\, dy \\
&\leq \int_{\mathbb{R}^n} |((x-y)+y)^\beta D^\alpha \phi(x-y)\psi(y)|\, dy \\
&\leq \sum_{\gamma \leq \beta} \int_{\mathbb{R}^n} \binom{\beta}{\gamma} |(x-y)^\gamma D\alpha\phi(x-y)\psi(y)|\, dy \\
&\leq \sum_{\gamma \leq \beta} \|P_\gamma D^\alpha\|_\infty C_{\beta-\gamma}(\psi)
\end{aligned} \tag{7}$$

It follows from (7) that

$$\sup_{x \in \mathbb{R}^n} |x^\beta D^\alpha (\phi * \psi)(x)| < \infty,$$

and consequently $\phi * \psi \in \mathcal{S}(\mathbb{R}^n)$. $\square$

Lemma 13.27: Let $\phi, \psi, \chi \in \mathcal{S}(\mathbb{R}^n)$ be given. Then we have

$$\phi * (\psi * \chi) = (\phi * \psi) * \chi.$$

Proof: For every $x \in \mathbb{R}^n$ we have

$$\begin{aligned}
((\phi * \psi) * \chi)(x) &= \int_{\mathbb{R}^n} (\phi * \psi)(y)\chi(x-y)\,dy \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \phi(z)\psi(y-z)\chi(x-y)\,dz\,dy \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \phi(z)\psi(y-z)\chi(x-y)\,dy\,dz \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \phi(z)\psi(s)\chi(x-s-z)\,ds\,dz \\
&= \int_{\mathbb{R}^n} \phi(z)(\phi * \chi)(x-z)\,dz \\
&= (\phi * (\psi * \chi))(x). \quad \square
\end{aligned}$$

Lemma 13.28: Let $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ be given. Then we have

(a) $(\phi * \psi)\widehat{\ } = (2\pi)^{\frac{n}{2}}\widehat{\phi}\widehat{\psi}$,

(b) $\widehat{\phi} * \widehat{\psi} = (2\pi)^{\frac{n}{2}}\widehat{\phi\psi}$.

Proof: (a) Let $\xi \in \mathbb{R}^n$ be given and observe that

$$\begin{aligned}
(\phi * \psi)\widehat{\ }(\xi) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\left(\int_{\mathbb{R}^n} \phi(y)\psi(x-y)\,dy\right)dx \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-i(x-y)\cdot\xi}e^{-iy\cdot\xi}\phi(y)\psi(x-y)\,dx\,dy \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-iz\cdot\xi}e^{-iy\cdot\xi}\phi(y)\psi(z)\,dz\,dy \\
&= (2\pi)^{\frac{n}{2}}\widehat{\phi}\widehat{\psi}.
\end{aligned}$$

For part (b), we apply (a) replacing ϕ with $\widehat{\phi}$ and ψ with $\widehat{\psi}$. This gives

$$\begin{aligned}(\widehat{\phi} * \widehat{\psi})\widehat{\ } &= (2\pi)^{\frac{n}{2}} \widehat{\widehat{\phi}} \widehat{\widehat{\psi}} \\ &= (2\pi)^{\frac{n}{2}} \check{\phi} \check{\psi} \\ &= (2\pi)^{\frac{n}{2}} (\phi\psi)\check{\ } \\ &= \widehat{\widehat{\phi\psi}}.\end{aligned}$$

The conclusion now follows by applying the inverse transform. $\square$

Translation

In order to extend the definition of convolution of two functions in $\mathcal{S}(\mathbb{R}^n)$ to convolution of a tempered distribution with a function in $\mathcal{S}(\mathbb{R}^n)$ it is convenient to introduce a translation operator.

Definition 13.29: Given $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$ define $\tau_h \phi \in \mathcal{S}(\mathbb{R}^n)$ by

$$(\tau_h \phi)(x) = \phi(x - h) \text{ for all } x \in \mathbb{R}^n.$$

Since we have

$$\int_{\mathbb{R}^n} \phi(x - h)\psi(x)\, dx = \int_{\mathbb{R}^n} \phi(y)\psi(y + h)\, dy$$

for all $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ and all $h \in \mathbb{R}^n$, it is appropriate to make the following definition.

Definition 13.30: Given $u \in \mathcal{S}'(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$ define $\tau_h u \in \mathcal{S}'(\mathbb{R}^n)$ by

$$\langle \tau_h u, \phi \rangle = \langle u, \tau_{-h} \phi \rangle \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

For the sake of completeness, we record below some simple results concerning the interaction of translations with Fourier transforms. We begin with a simple definition.

Definition 13.31: For each $z \in \mathbb{R}$ define $e_z : \mathbb{R}^n \to \mathbb{C}$ by

$$e_z(x) = e^{iz \cdot x} \text{ for all } x \in \mathbb{R}^n. \tag{8}$$

Observe that for each $z \in \mathbb{R}^n$ we have $e_z \in \mathcal{PGD}(\mathbb{R}^n)$ so we can talk about the product $e_z u$ when u is a tempered distribution.

Proposition 13.32: Let $h \in \mathbb{R}^n$ and $\phi \in \mathcal{S}(\mathbb{R}^n)$, $\psi \in \mathcal{S}'(\mathbb{R}^n)$ be given. Then we have

(a) $\tau_h\widehat{\phi} = (e_h f)\widehat{}$

(b) $(\tau_h\phi)\widehat{} = e_{-h}\widehat{\phi}$

(c) $\tau_h\widehat{u} = (e_h f)\widehat{}$

(d) $(\tau_h u)\widehat{} = e_{-h}\widehat{u}$.

The proof of Proposition 13.32 is left as an elementary exercise.

Convolution of a Tempered Distribution with a Test Function

Observe that for $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ we have

$$\begin{aligned}
(\phi * \psi)(x) &= \int_{\mathbb{R}^n} \phi(y)\psi(x-y)\,dy \\
&= \int_{\mathbb{R}^n} \phi(y)\check{\phi}(y-x)\,dy \\
&= \int_{\mathbb{R}^n} \phi(y)(\tau_x\check{\psi})(y)\,dy \quad \text{for all } x \in \mathbb{R}^n.
\end{aligned}$$

It is therefore natural to make the following definition:

Definition 13.33: Given $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\phi \in \mathcal{S}(\mathbb{R}^n)$ define $u * \phi : \mathbb{R}^n \to \mathbb{C}$ by

$$(u * \phi)(x) = \langle u, \tau_x\check{\phi}\rangle \quad \text{for all } x \in \mathbb{R}^n.$$

Theorem 13.34: Let $u \in \mathcal{S}(\mathbb{R}^n)$, $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$, and $\alpha \in M_n$ be given. Then we have

(a) $u * \phi \in \mathcal{PGD}(\mathbb{R}^n)$,

(b) $D^\alpha(u * \phi) = (D^\alpha u) * \phi = u * (D^\alpha \phi)$,

(c) $(u * \phi)\widehat{} = (2\pi)^{\frac{n}{2}}\widehat{\phi}\widehat{u}$,

(d) $\widehat{u} * \widehat{\phi} = (2\pi)^{\frac{n}{2}}\phi u$,

(e) $u * (\phi * \psi) = (u * \phi) * \psi$.

Sobolev Spaces

Definition 13.35 Let $m \in \mathbb{N} \cup \{0\}$, $p \in [1, \infty]$ be given and put

$$W^{m,p}(\mathbb{R}^n) = \{u \in \mathcal{S}'(\mathbb{R}^n) : D^\alpha u \in L^p(\mathbb{R}^n),\ |\alpha| \leq m\}.$$

We equip $W^{m,p}(\mathbb{R}^n)$ with the norm given by

$$|||u|||_{m,p} = \begin{cases} \left(\sum_{|\alpha| leqm} \|u\|_p^p \right)^{\frac{1}{p}} & \text{if } p \in [1, \infty), \\ \sum_{|\alpha| \leq m} \|D^\alpha u\|_\infty & \text{if } p = \infty. \end{cases}$$

Here $\| \cdot \|_p$ is the norm on $L^p(\mathbb{R}^n)$.

Remark 13.36: Using the fact the all norms on a finite-dimensional linear space are equivalent, it is easy to construct equivalent norms on $W^{m,p}(\mathbb{R}^n)$ simply by applying different norms to the finite list

$$(\|D^\alpha u\|_p | \alpha \in M_n,\ |\alpha| \leq m)$$

of real numbers. However, one must be aware that changing to one of these equivalent norms can destroy geometric properties such as uniform convexity. With the norm $|||\cdot|||_{m,p}$ as above, $W^{m,p}(\mathbb{R}^n$ is uniformly convex when $1 < p < \infty$.

Proposition 13.37: Let $m \in \mathbb{N} \cup \{0\}$ and p $[1, \infty]$ be given. Then $W^{m,p}(\mathbb{R}^n)$ is a Banach space.

Proof: That $W^{m,p}(\mathbb{R}^n)$ is a linear space and that $|||\cdot|||_{m,p}$ is a norm is clear. We need to check completeness. Let $\{u_k\}_{k=1}^\infty$ be a Cauchy sequence in $W^{m,p}(\mathbb{R}^n)$. Since $L^p(\mathbb{R}^n)$ is complete, we may choose functions $v_\alpha \in L^p(\mathbb{R}^n)$, $\alpha \in M_n$, $|\alpha| \leq m$ such that

$$\forall \alpha \in M_n,\ |\alpha| \leq m,\ \ \|D^\alpha u_k - v^\alpha\|_p \to 0 \text{ as } k \to \infty.$$

Let us put $u = v_0$. We need to show that $D^\alpha u = v_\alpha$ for all multi-indices α with $|\alpha| \leq m$.

Let $\alpha \in M_n$ with $|\alpha| \leq m$ be given. Since strong convergence in $L^p(\mathbb{R}^n)$ implies weak convergence (when $1 \leq p < \infty$) and weak* convergence (when $p = \infty$) and because $\mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$, we know that for every $\phi \in \mathcal{S}(\mathbb{R}^n)$

$$\langle D^\alpha u_k, \phi \rangle = \int_{\mathbb{R}^n} (D^\alpha u_k)\phi \to \int_{\mathbb{R}^n} v_\alpha \phi = \langle v_\alpha, \phi \rangle \text{ as } k \to \infty.$$

We also know that

$$\begin{aligned} \langle D^\alpha u_k, \phi \rangle &= (-1)^{|\alpha|} \langle u_k, D^\alpha \phi \rangle \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} u_k D^\alpha \phi \\ &\to (-1)^{|\alpha|} \int_{\mathbb{R}^n} u D^\alpha \phi = \langle D^\alpha u, \phi \rangle \text{ as } k \to \infty. \end{aligned}$$

It follows that $v_\alpha = D^\alpha u$. $\square$

Using the fact that $W^{m,p}(\mathbb{R}^n)$ can be identified with a closed subspace of a finite product of $L^p(\mathbb{R}^n)$ spaces, we can infer that

- $W^{m,p}(\mathbb{R}^n)$ is separable if $1 \leq p < \infty$, and
- $W^{m,p}(\mathbb{R}^n)$ is reflexive if $1 < p < \infty$.

Notice that $W^{0,p}(\mathbb{R}^n) = L^p(\mathbb{R}^n)$ and that $W^{m,2}(\mathbb{R}^n)$ is a Hilbert space with inner product

$$(\!(u, v)\!)_m = \sum_{|\alpha| \leq m} \int_{\mathbb{R}^n} D^\alpha u \overline{D^\alpha v} \quad \text{for all } u, v \in W^{m,2}(\mathbb{R}^n).$$

The spaces $W^{m,p}(\mathbb{R}^n)$ are special instances of a class of spaces known as *Sobolev spaces*. Such spaces are extremely useful in the study of partial differential equations and calculus of variations. There are entire books (e.g. Adams) devoted to Sobolev spaces and we can only begin to scratch the surface here.

We shall prove a couple of representative results concerning Sobolev spaces in the special case $p = 2$ using Fourier transforms. For each $s \in \mathbb{R}$ put

$$Q_s(\xi) = (1 + |\xi|^2)^s \quad \text{for all } \xi \in \mathbb{R}^n. \tag{9}$$

Let $m \in \mathbb{N} \cup \{0\}$ and $u \in \mathcal{S}'(\mathbb{R}^n)$ be given. Observe that (i) through (iv) below are equivalent

(i) $u \in W^{m,2}(\mathbb{R}^n)$

(ii) $D^\alpha u \in L^2(\mathbb{R}^n)$ for all multi-indices α with $|\alpha| \leq m$.

(iii) $(D^\alpha u)\widehat{\ }= i^{|\alpha|} P_\alpha \widehat{u} \in L^2(\mathbb{R}^n)$ for all multi-indices α with $|\alpha| \leq m$.

By elementary algebra, (iii) holds if and only if (iv) below holds:

(iv) $Q_{\frac{m}{2}} \widehat{u} \in L^2(\mathbb{R}^n)$.

The idea is that all powers ξ^α for $|\alpha| \leq m$ can be controlled by $Q_{\frac{m}{2}}(\xi)$. It is extremely interesting to observe that (iv) makes sense for arbitrary real numbers m (positive or negative) and not just integers.

Definition 13.38: For each $s \in \mathbb{R}$ let

$$H^s(\mathbb{R}^n) = \{u \in \mathcal{S}'(\mathbb{R}^n) : Q_{\frac{s}{2}} \widehat{u} \in L^2(\mathbb{R}^n\},$$

equipped with the inner product defined by

$$(u, v)_s = \int_{\mathbb{R}^n} Q_s \widehat{u}\overline{\widehat{v}} \;\text{ for all } u, v \in H^s(\mathbb{R}^n),$$

and the associated norm

$$\|u\|_{s,2} = \sqrt{(u, u)_s} \;\text{ for all } u \in H^s(\mathbb{R}^n).$$

Remark 13.39: Let $m \in \mathbb{N} \cup \{0\}$. Then $H^m(\mathbb{R}^n) = W^{m,2}(\mathbb{R}^n)$ and the norm $\|\cdot\|_{m,2}$ is equivalent to $|||\cdot|||_{m.2}$.

When $s \geq 0$ the elements of $H^s(\mathbb{R}^n)$ can be identified with functions. The functions become more regular as s increases. For $s < 0$, the elements of $H^s(\mathbb{R}^n)$ are tempered distributions, but they need not be associated with functions. It is instructive to look at an example.

Example 13.40 (Dirac Delta): Consider the tempered distribution δ_0 defined by

$$\langle \delta_0, \phi \rangle = \phi(0) \;\text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n).$$

In Example x.x we saw that

$$\widehat{\delta_0} = \frac{1}{(2\pi)^{\frac{n}{2}}}.$$

Let $s \in \mathbb{R}$ be given. In order to have $\delta_0 \in H^s(\mathbb{R}^n)$ it is necessary and sufficient that $Q_{\frac{s}{2}} \in L^2(\mathbb{R}^n)$, i.e. that

$$\int_{\mathbb{R}^n} (1 + |\xi|^2)^s \, d\xi < \infty.$$

Switching to polar coordinates, we find that the integral above is equal to

$$C_n \int_0^\infty (1 + r^2)^s r^{n-1} \, dr, \tag{10}$$

where C_n is the "surface area" of the unit sphere in $\mathbb{R}^n$. The integral in (10) converges if and only if $s < -\frac{n}{2}$. We conclude that

$$\delta_0 \in H^s(\mathbb{R}^n) \;\Leftrightarrow\; s < -\frac{n}{2}.$$

Theorem 13.41 (Sobolev Embedding Theorem – Special Case): Let $s > \frac{n}{2}$ be given. Then

$$H^s(\mathbb{R}^n) \hookrightarrow C_0(\mathbb{R}^n). \tag{11}$$

Proof: Let $u \in H^s(\mathbb{R}^n)$ be given and observe that

$$Q_{-\frac{s}{2}} \widehat{u} \in L^2(\mathbb{R}^n). \tag{12}$$

Observe further that
$$\widehat{u} = (Q_{\frac{s}{2}}\widehat{u})Q_{-\frac{s}{2}}. \tag{13}$$
We want to show that $Q_{-\frac{s}{2}} \in L^2(\mathbb{R}^n)$, because then it will follow from (21), (13) that $\widehat{u} \in L^1(\mathbb{R}^n)$ and we will also get a useful bound for $\|\widehat{u}\|_1$. Using polar coordinates, we have
$$\int_{\mathbb{R}^n} Q_{-s} = C_n \int_0^\infty (1+r^2)^{-s} r^{n-1} \tag{14}$$
where C_n is the "surface area" of the unit sphere in $\mathbb{R}^n$. [The integrand is behaving like r^{n-2s-1} as $r \to \infty$.] Since $s > \dfrac{n}{2}$ the integral in (14) is finite. Let us put
$$K = \|Q_{-\frac{s}{2}}\|_2.$$
(Here $\|\cdot\|_2$ is the L^2-norm.) Then, by (13) and Holder's inequality, we have
$$\|\widehat{u}\|_1 \leq K\|u\|_{s,2}. \tag{15}$$
By Theorem 13.20 we have
$$\breve{u} = \widehat{\widehat{u}} \tag{16}$$
and consequently we have $\breve{u} \in C_0(\mathbb{R}^n)$ and
$$\|\breve{u}\|_\infty \leq \frac{1}{(2\pi)^{\frac{n}{2}}}\|\widehat{\widehat{u}}\|_1 \text{ for all } u \in H^S(\mathbb{R}^n). \tag{17}$$
It follows that $u \in C_0$ and
$$\|u\|_\infty \leq \frac{K}{(2\pi)^{\frac{n}{2}}}\|u\|_{s,2} \text{ for all } u \in H^S(\mathbb{R}^n)$$
and the proof is complete. $\square$

Theorem 13.42 (Interpolation Inequality): Let $s, t \in \mathbb{R}$ with $0 < s < t$ be given. Then for every $\epsilon > 0$ there exists $C(\epsilon) > 0$ such that
$$\|u\|_{s,2}^2 \leq \epsilon\|u\|_{t,2}^2 + C(\epsilon)\|u\|_{0,2}^2 \text{ for all } u \in H^t(\mathbb{R}^n). \tag{18}$$

Proof: Recall Young's Inequality which says that for
$$p \in (1,\infty), \quad q = \frac{p}{p-1},$$
we have
$$AB \leq \frac{A^p}{p} + \frac{B^q}{q} \text{ for all } A, B \geq 0. \tag{19}$$
Put
$$p = \frac{t}{s}, \quad q = \frac{t}{t-s}. \tag{20}$$

Let $\epsilon > 0$, $a \geq 0$ be given. Then using (19) with p, q as in (20) we have

$$a = \epsilon^{\frac{s}{t}} a \cdot \epsilon^{-\frac{s}{t}} 1 \leq \frac{s}{t} \epsilon a^{\frac{t}{s}} + \left(\frac{t-s}{t} \right) \epsilon^{-\frac{s}{t-s}}. \tag{21}$$

If we put

$$C(\epsilon) = \frac{t-s}{t} \epsilon^{-\frac{s}{t-s}},$$

and notice that $\frac{s}{t} < 1$ we infer from (**??**) that

$$a \leq \epsilon^{\frac{t}{s}} + C(\epsilon) \quad \text{for all } a \geq 0. \tag{22}$$

It follows from (22) that

$$Q_s(\xi) \leq \epsilon Q_t(\xi) + C(\epsilon) \quad \text{for all } \xi \in \mathbb{R}^n. \tag{23}$$

Given $u \in H^t(\mathbb{R}^n)$, we multiply (23) by $|\widehat{u}(\xi)|^2$ to obtain

$$Q_s(\xi)|\widehat{u}(\xi)|^2 \leq \epsilon Q_t(\xi)|\widehat{u}(\xi)|^2 + C(\epsilon)|\widehat{u}(\xi)|^2 \quad \text{for all } \xi \in \mathbb{R}^n. \tag{24}$$

Integration of (24) over $\mathbb{R}^n$ produces the desired conclusion. $\square$

Example 13.43: Let $f \in L^2(\mathbb{R}^n)$ be given. We seek $u \in \mathcal{S}'(\mathbb{R}^n)$ such that

$$-\Delta u + u = f. \tag{25}$$

For $u \in \mathcal{S}'(\mathbb{R}^n)$ we have

$$(-\Delta u + u)\widehat{\ } = Q_1 \widehat{u}.$$

Consequently u satisfies (25) if and only if

$$\widehat{u} = \frac{\widehat{f}}{Q_1} = Q_{-1}\widehat{f}. \tag{26}$$

It is easy to see that $Q_{-1} \in \mathcal{PGD}(\mathbb{R}^n)$ which implies that $Q_{-1}\widehat{f} \in \mathcal{S}'(\mathbb{R}^n)$. It follows that (25) has exactly one solution $u \in \mathcal{S}'(\mathbb{R}^n)$ and this solution is given by

$$\breve{u} = (Q_{-1}\widehat{f})\widehat{\ }.$$

Since $f \in L^2(\mathbb{R}^n)$ we can conclude from (26) that the solution u of (25) actually satisfies $u \in H^2(\mathbb{R}^n)$.

The Wave Equation

Let us consider that initial-value problem

$$\begin{cases} u_{tt}(t,x) = \Delta u(t,x), \quad t \geq 0, \ x \in \mathbb{R}^n, \\ u(0,x) = u_0(x), \ u_t(0,x) = v_0(x), \quad x \in \mathbb{R}^n, \end{cases} \tag{27}$$

where $u_0, v_0 : \mathbb{R}^n \to \mathbb{C}$ are given functions. Here Δ is the Laplacian with respect to the spatial variables x. The partial differential equation $u_{tt} = \Delta u$ is called the *wave equation.*

Before choosing function spaces for this problem, we shall derive a formal "energy equation". Such energy equations can reveal a great deal information. Before proceeding, we observe that if $J \subset \mathbb{R}$ is an interval and $z : J \to \mathbb{C}$ is differentiable then we have

$$\begin{aligned} \frac{d}{dt}|z(t)|^2 &= 2z(t)\dot{\overline{z}}(t) \\ &= 2\overline{z}(t)\dot{z}(t). \end{aligned}$$

(The derivative of the function $t \to |z(t)|^2$ is real-valued and $\mathrm{Re}(\alpha\overline{\beta}) = \mathrm{Re}(\overline{\alpha}\beta)$ for all complex numbers α, β.)

Assuming that (27) has a sufficiently regular solution u, we multiply the equation by $\overline{u}_t$ and integrate (with respect to x) over $\mathbb{R}^n$, using integration by parts to obtain:

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}^n} |u_t(t,x)|^2\, dx = -\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}^n} |\nabla u(t,x)|^2\, dx.$$

Integrating the above expression with respect to time and using the initial conditions, we find that

$$\frac{1}{2}\int_{\mathbb{R}^n} (|u_t(t,x)|^2 + |\nabla u(t,x)|^2)\, dx = \frac{1}{2}\int_{\mathbb{R}^n} (|v_0(x)|^2 + |\nabla u_0(x)|^2)\, dx \ \text{ for all } t \geq 0. \quad (28)$$

It is natural to employ function spaces such $u_t(t,\cdot)$ lives in $L^2(\mathbb{R}^n)$ and $\nabla u(t,\cdot)$ lives in $L^2(\mathbb{R}^n;\mathbb{C}^n)$ (vector-valued L^2.) It is reasonable (but not essential) to require $u(t,\cdot)$ to live in $L^2(\mathbb{R}^n)$ also (and this can be ensured by assuming that $u_0 \in L^2(\mathbb{R}^n)$.)

We shall rewrite the wave equation as a first-order system by letting $v = \dot{u}$:

$$(\dot{u}(t), \dot{v}(t)) = (v(t), \Delta u(t)).$$

Let

$$X = H^1(\mathbb{R}^n) \times L^2(\mathbb{R}^n),$$

equipped with the inner product $\langle\!\langle \cdot,\cdot \rangle\!\rangle$ defined by

$$\langle\!\langle (\phi_1,\psi_1),(\phi_2,\psi_2) \rangle\!\rangle = \int_{\mathbb{R}^n} [\phi_1\overline{\phi_2} + \nabla\phi_1 \cdot \overline{\nabla\phi_2} + \psi_1\overline{\psi_2}].$$

Formally, the action of the operator A will be given by

$$A(\phi,\psi) = (\psi, \Delta\phi).$$

The relationship between the energy integral and the inner product will allow us to show that A is quasidissipative, so we will be able to use the Lumer-Phillips Theorem for Hilbert spaces. Before choosing a domain for A we look at the surjectivity of $I - A$. Let $(f, g) \in X$ be given. We need to find $(\phi, \psi) \in \mathcal{D}(A)$ such that

$$\begin{aligned} \phi - \psi &= f \\ \psi - \Delta\phi &= g. \end{aligned} \tag{29}$$

Adding the equations in (29) we obtain

$$\phi - \Delta\phi = f + g. \tag{30}$$

Since $f + g \in L^2(\mathbb{R}^n)$ it follows from Example 13.43 that (30) has a solution $\phi \in H^2(\mathbb{R}^n)$. The corresponding ψ is given by

$$\psi = \phi - f \in H^1(\mathbb{R}^n). \tag{31}$$

Consequently, a natural choice is

$$\mathcal{D}(A) = H^2(\mathbb{R}^n) \times H^1(\mathbb{R}^n) \tag{32}$$

and we define $A : \mathcal{D}(A) \to X$ by

$$A(\phi, \psi) = (\psi, \Delta\phi) \ \text{ for all } (\phi, \psi) \in \mathcal{D}(A). \tag{33}$$

Let $(\phi, \psi) \in \mathcal{D}(A)$ be given. Then we have

$$\begin{aligned} \mathrm{Re}\langle\langle A(\phi, \psi), (\phi, \psi)\rangle\rangle &= \mathrm{Re} \int_{\mathbb{R}^n} [\psi\overline{\phi} + \nabla\phi \cdot \nabla\overline{\psi} + (\Delta\phi)\overline{\psi}] \\ &= \mathrm{Re} \int_{\mathbb{R}^n} \psi\overline{\phi} \\ &\leq \frac{1}{2} \int_{\mathbb{R}^n} [|\psi|^2 + |\phi|^2] \leq \frac{1}{2} \|(\phi, \psi)\|_X^2. \end{aligned}$$

We conclude that A generates a linear C_0-semigroup $T : [0, \infty) \to \mathcal{L}(X; X)$ satisfying

$$\|T(t)\| \leq e^{\frac{1}{2}t} \ \text{ for all } t \geq 0. \tag{34}$$

The growth estimate (34) is not optimal. In view of (28) one should expect

$$\int_{\mathbb{R}^n} (|u_t(t, x)|^2 + |\nabla u(t, x)|^2)\, dx$$

to remain bounded in t; however, in general, there is no reason to believe that

$$\int_{\mathbb{R}^n} |u(t, x)|^2\, dx \tag{35}$$

will be bounded in t. The growth of the integral in (35) is at most algebraic in t and depends on the spatial dimension. See, for example, for a discussion of such results.

The wave equation is reversible in time, so we cannot expect any smoothing of the initial data. Moreover, Littman has shown that if $n \geq 2$ then the solution operator for the wave equation does not form a linear C_0-semigroup on $W^{1,p}(\mathbb{R}^n) \times L^p(\mathbb{R}^n)$ unless $p = 2$. This means that the result of Problem 5 on Assignment 3 gives a rather sharp regularity result for the initial value problem (27): For each $m \in \mathbb{N}$ and $u_0 \in H^{m+1}(\mathbb{R}^n)$, $v_0 \in H^m(\mathbb{R}^n)$ the solution of (27) satisfies

$$u \in C^k([0, \infty); H^{m+1-k}(\mathbb{R}^n) \quad \text{for all } k = 0, 1, 2, \cdots m + 1.$$

Lecture Notes for Week 14 (First Draft)

Monotone Mappings on $\mathbb{R}$ and $\mathbb{R}^n$

Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is said to be

- *increasing* if $f(x_2) \geq f(x_1)$ for all $x_1, x_2 \in \mathbb{R}$ with $x_2 \geq x_1$,
- *decreasing* if $-f$ is increasing, and
- *monotone* provided that f is either increasing or decreasing.

Monotone functions from $\mathbb{R}$ to $\mathbb{R}$ have many special properties. For example, it is a simple exercise in real analysis to prove that if $f : \mathbb{R} \to \mathbb{R}$ is monotone, continuous and satisfies

$$|f(x)| \to \infty \quad \text{as } |x| \to \infty. \tag{1}$$

then f is surjective. We shall generalize this result to mappings on infinite-dimensional spaces.

Since it is completely straightforward to "adjust" results obtained for increasing functions so that they apply to decreasing functions, we shall focus on generalizing notions associated with increasing functions. We shall refer to such mappings as "monotone" (rather than increasing).

The condition

$$f(x_2) \geq f(x_1) \quad \text{for all } x_1, x_2 \in \mathbb{R} \text{ with } x_2 \geq x_1 \tag{2}$$

is not particularly well-suited to generalization. Observe that a function $f : \mathbb{R} \to \mathbb{R}$ satisfies (2) if and only if

$$(f(x_2) - f(x_1)) \cdot (x_2 - x_1) \geq 0 \quad \text{for all } x_1, x_2 \in \mathbb{R}. \tag{3}$$

Observe also that if $f : \mathbb{R} \to \mathbb{R}$ satisfies (3) then (1) holds if and only if

$$\frac{f(x) \cdot x}{|x|} \to \infty \quad \text{as } |x| \to \infty. \tag{4}$$

A function $g : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *monotone* provided that

$$(g(x) - g(y)) \cdot (x - y) \geq 0 \quad \text{for all } x, y \in \mathbb{R}^n. \tag{5}$$

It is a standard result in convex analysis (and a special case of a result that we will prove later) that a C^1-function $\phi : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if $\nabla\phi : \mathbb{R}^n \to \mathbb{R}^n$ is monotone as defined above.

In extending (5) to an infinite-dimensional setting we could use inner products (or semi-inner products) but it is actually very convenient to consider mappings from a Banach space to its dual space.

Definition 14.1: Let X be a real Banach space with dual space X^*. A mapping $F : X \to X^*$ is said to be

(a) *monotone* provided that $\langle F(u) - F(v), u - v\rangle \geq 0$ for all $u, v \in X$,

(b) *strictly monotone* provided that $\langle F(u) - F(v), u - v\rangle > 0$ for all $u, v \in X$ with $u \neq v$,

(c) *bounded* provided $F[B]$ is bounded in X^* for every bounded set $B \subset X$,

(d) *coercive* provided that

$$\frac{\langle F(u), u\rangle}{\|u\|} \to \infty \quad \text{as } \|u\| \to \infty.$$

Remark 14.2: It is an immediate consequence of the definition that if $F : X \to X^*$ is strictly monotone then F is injective.

Theorem 14.3 (Browder-Minty): Let X be a real Banach space with dual space X^*. Assume that X is separable and reflexive. Let $F : X \to X^*$ be given and assume that F is monotone, bounded, continuous, and coercive. Then F is surjective.

The proof of Theorem 14.3 will be based on the Galerkin Method. (Here the idea is to replace a problem in an infinite-dimesional space with a sequence of approximating, finite-dimensional, problems, and pass to the limit.) In order to obtain solutions to the finite-dimensional problems we shall make use of a result that follows fairly easily from Brouwer's Fixed-Point Theorem.

Lemma 14.4: Let $m \in \mathbb{N}$ and $\rho > 0$ be given. Assume that $\Phi : \mathbb{R}^m \to \mathbb{R}^m$ is continuous and satisfies

$$\xi \cdot \Phi(\xi) \geq 0 \quad \text{for all } \xi \in \mathbb{R}^m \text{ with } |\xi| = \rho. \tag{6}$$

Then there exists $\eta \in \mathbb{R}^m$ with $|\eta| \leq \rho$ such that $\Phi(\eta) = 0$.

Proof: Suppose that there is no such η. Define

$$\Psi(\xi) = -\rho\frac{\Phi(\xi)}{|\Phi(\xi)|} \quad \text{for all } \xi \in \overline{B}_\rho(0), \tag{7}$$

and observe that $\Psi : \overline{B}_\rho(0) \to \overline{B}_\rho(0)$ continuously. It follows from Brouwer's Fixed-Point Theorem that there exists $\xi^* \in \overline{B}_\rho(0)$ such that $\xi^* = \Psi(\xi^*)$. Notice that

$$|\xi^*| = |\Psi(\xi^*)| = \rho. \tag{8}$$

Appealing to (6) and using (8) we see that

$$\begin{aligned} 0 \leq \xi^* \cdot \Phi(\xi^*) &= -\frac{|\Phi(\xi^*)|}{\rho}\Psi(\xi^*)\cdot\xi^* \\ &= -\rho|\Phi(\xi^*)| < 0, \end{aligned}$$

and this is a contradiction. □

Proof of Theorem 14.3: Let $g \in X^*$ be given. We want to find $u \in X$ such that $F(u) = g$.

Since X is separable, we may choose a linearly dependent sequence $\{x_j\}_{j=1}^\infty$ such that

$$\mathrm{span}(x_j | j \in \mathbb{N}) \text{ is dense in } X.$$

For each $m \in \mathbb{N}$, put

$$V_m = \mathrm{span}(x_1, x_2, \cdots, x_m).$$

Let $m \in \mathbb{N}$ be given. We seek $u_m \in V_m$ such that

$$\langle F(u_m), v\rangle = \langle g, v\rangle \quad \text{for all } v \in V_m. \tag{9}$$

To construct a solution of (9) we define $\Phi : \mathbb{R}^m \to \mathbb{R}^m$ by

$$\Phi_i(\xi) = \langle F\left(\sum_{j=1}^m \xi_j v_j\right) - g, v_i\rangle \quad \text{for all } \xi \in \mathbb{R}^m. \tag{10}$$

For ease of notation, it is convenient to put

$$v_m^\xi = \sum_{j=1}^m \xi_j x_j \quad \text{for all } \xi \in \mathbb{R}^m. \tag{11}$$

Continuity of F implies continuity of Φ. Moreover, we have

$$\begin{aligned} \xi\cdot\Phi(\xi) &= \sum_{i=1}^m \xi_i \Phi_i(\xi) \\ &= \langle F(v_m^\xi) - g, v_m^\xi\rangle \\ &= \langle F(v_m^\xi), v^\xi\rangle - \langle g, v_m^\xi\rangle \\ &\geq \left(\frac{\langle F(v_m^\xi), v_m^\xi\rangle}{\|v_m^\xi\|} - \|g\|\right). \end{aligned} \tag{12}$$

Since all norms on V_m are equivalent (and m is fixed), we know that

$$\|v_m^{\xi}\| \to \infty \quad \text{as } |\xi| \to \infty. \tag{13}$$

Using the continuity of Φ, (12), (13) and Lemma 14.4, we may choose $\xi^* \in \mathbb{R}^m$ such that $\Phi(\xi^*) = 0$. If we put

$$u_m = v_m^{\xi^*} \tag{14}$$

then u_m satisfies (9). In this way we generate a sequence $\{u_m\}_{m=1}^{\infty}$ such that u_m satisfies (9) for every $m \in \mathbb{N}$.

Since $u_m \in V_m$ for every $m \in \mathbb{N}$, it follows from (9) and a simple computation that

$$\frac{\langle F(u_m), u_m\rangle}{\|u_m\|} \leq \|g\| \quad \text{for all } m \in \mathbb{N}. \tag{15}$$

Since F is coercive, it follows from (15) that the sequence $\{u_m\}_{m=1}^{\infty}$ is bounded.

Since X is reflexive and separable, and since $\{u_m\}_{m=1}^{\infty}$ and $\{F(u_m)\}_{m=1}^{\infty}$ are bounded we may choose a subsequence $\{u_{m_j}\}_{j=1}^{\infty}$ and $u \in X$, $\phi \in X^*$ such that

$$u_{m_j} \rightharpoonup u \text{ (weakly)}, \quad F(u_{m_j}) \overset{*}{\rightharpoonup} \phi \text{ (weakly}*\text{) as } j \to \infty.$$

We shall show that $\phi = g$ and $F(u) = g$. We know that

$$\langle F(u_{m_j}), v\rangle = \langle g, v\rangle \quad \text{for all } v \in V_{m_j}.$$

Since the spaces V_m are "nested", we see that

$$\langle F(u_{m_j}), v\rangle \to \langle g, v\rangle \text{ as } j \to \infty \quad \text{for all } v \in \bigcup_{m=1}^{\infty} V_m.$$

It follows that

$$\langle \phi, v\rangle = \langle g, v\rangle \quad \text{for all } v \in \bigcup_{m=1}^{\infty} V_m.$$

Since $\operatorname{span}(x_k | k \in \mathbb{N})$ is dense in X we see that

$$\phi = g. \tag{16}$$

Now let $v \in X$ be given. Since F is monotone we have

$$\begin{aligned}
0 \quad &\leq \quad \langle F(v) - F(u_{m_j}), v - u_{m_j}\rangle \\
&\leq \quad \langle F(v), v\rangle - \langle F(v), u_{m_j}\rangle - \langle F(u_{m_j}), v\rangle + \langle F(u_{m_j}, u_{m_j}\rangle \\
&\leq \quad \langle F(v), v\rangle - \langle F(v), u_{m_j}\rangle - \langle F(u_{m_j}, v\rangle + \langle g, u_{m_j}\rangle \\
&\to \quad \langle F(v), v\rangle - \langle F(v), u\rangle - \langle g, v\rangle + \langle g, u\rangle \quad \text{as } j \to \infty.
\end{aligned} \tag{17}$$

It follows that

$$\langle F(v) - g, v - u\rangle \geq 0 \ \text{ for all } v \in X. \tag{18}$$

We can use (18) to conclude that $F(u) = g$. To this end, let $w \in X$, $t > 0$ be given and put

$$v = u + tw.$$

Substituting this choice of v into (18) and using the fact that $t > 0$ we find that

$$\langle F(u + tw) - g, w\rangle \geq 0 \ \text{ for all } w \in X,\ t > 0. \tag{19}$$

Letting $t \downarrow 0$ in (19) we see that

$$\langle F(u) - g, w\rangle \geq 0 \ \text{ for all } w \in X. \tag{20}$$

Replacing w with $-w$ in (20) we see that

$$\langle F(u) - g, w\rangle = 0 \ \text{ for all } w \in X,$$

which implies that $F(u) = g$. $\square$

Remark 14.5: Examination of the proof of Theorem 14.3 reveals that the assumption of continuity of F can be weakened. In particular, it would be enough to know that the restriction of F to finite-dimensional subspaces is continuous and that for every $u, v, w \in X$ the mapping

$$t \to \langle F(u + tw), v\rangle$$

is continuous. This last condition is sometimes called *hemicontinuity*. See Kato for a discussion of this issue.

Fréchet and Gâteaux Differentiability

Definition 14.6: Let X and Y be Banach spaces and $F : X \to Y$ and $x_0 \in X$ be given. We say that F is *Fréchet differentiable* at x_0 provided there exists $L \in \mathcal{L}(X; Y)$ such that

$$\frac{F(x_0 + h) - F(x_0) - Lh}{\|h\|} \to 0 \ \text{ as } h \to 0,\ h \in X.$$

The linear operator L is called the *Fréchet derivative* of F at x_0 and we write

$$F'(x_0) = \nabla F(x_0) = DF(x_0) = L.$$

We say that F is *Fréchet differentiable on* X if F is Fréchet differentiable at each $x_0 \in X$.

Remark 14.7: If F is Fréchet differentiable at x_0 then F is continuous at x_0. [This follows easily from the definition.]

Remark 14.8: The definition of Fréchet differentiability (and most of the basic results) can easily be adapted to mappings $F : U \to Y$, where U is an open subset of X.

Remark 14.9 (Higher Differentiability):

Proposition 14.10 (Chain Rule): Let X, Y, and Z be Banach spaces, $G : X \to Y$, $F : Y \to Z$, and $x_0 \in X$ be given. Assume that G is Fréchet differentiable at x_0 and that F is Fréchet differentiable at $G(x_0)$. Then $F \circ G$ is Fréchet differentiable at x_0 and

$$(F \circ G)'(x_0) = F'(G(x_0))G'(x_0).$$

Proof: We employ the standard "little oh" notation; in particular, if we write

$$\phi(h) = o(\|h\|) \ \text{ as } h \to 0,$$

for a function $\phi : X \to Y$ what we really mean is that

$$\lim_{h \to 0} \frac{\phi(h)}{\|h\|} = 0.$$

For $h \in X \backslash \{0\}$ we have

$$G(x_0 + h) = G(x_0) + G'(X_0)h + o(\|h\|) \ \text{ as } h \to 0,$$

and consequently

$$\begin{aligned} F(G(x_0 + h)) &= F(G(x_0) + G'(x_0)h + o(\|h\|)) \ \text{ as } h \to 0 \\ &= F(G(x_0)) + F'(G(x_0))[G'(x_0)h + o(\|h\|)] + o(\|h\|) \ \text{ as } h \to 0 \\ &= F(G(x_0)) + F'(G(x_0))G'(x_0)h + o(\|h\|) \ \text{ as } h \to 0, \end{aligned}$$

and the result follows. $\square$

Example 14.11: Let $\mathbb{K} = \mathbb{R}$.

(a) Let $X = L^2[0, 1]$ and define $F : X \to X$ by

$$(F(u))(x) = \sin u(x) \ \text{ for all } x \in [0, 1].$$

Notice that

$$\|F(u) - F(v)\|_2 \leq \|u - v\|_2 \ \text{ for all } u, v \in X,$$

so that F is (globally) Lipschitz continuous. You should check as an exercise for yourself that F is nowhere Fréchet differentiable. However, changing the space a bit, dramatically alters the result

(b) Let $Y = C[0,1]$ and define $G : Y \to Y$ by

$$G(u(x)) = \sin u(x) \quad \text{for all } x \in [0,1].$$

You should check for yourself as an exercise that G is everywhere Fréchet differentiable and

$$(G'(u)h)(x) = (\cos u(x))h(x) \quad \text{for all } u, h \in Y,\ x \in [0,1].$$

(In fact, $G : Y \to Y$ is actually analytic!)

Definition 14.12: Let X and Y be Banach spaces and $F : X \to Y$, $x_0, v \in X$ be given. We say that F has a Gâteaux variation at x_0 in the direction v provided that

$$\lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t} \quad \text{exists;}$$

in this case we write

$$\delta F(x_0; v) = \lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t}.$$

(The variable t is real in the limits above.) $\delta F(x_0; v)$ is called the *Gâteaux variation* of F at x_0 in the direction v. We say that F is *Gâteaux differentiable* at x_0 provided F has a Gâteaux variation $\delta F(x_0; u)$ at x_0 in every direction $u \in X$.

Remark 14.13: Assume that F is Gâteaux differentiable at x_0. The mapping $v \to \delta F(x_0; v)$ need not be linear (even in finite dimensions) as the example below shows. Many authors require that that the mapping $v \to \delta F(x_0; v)$ be linear and continuous as part of the definition of Gâteaux differentiability at x_0.

Example 14.14: Assume that $F : \mathbb{R}^2 \to \mathbb{R}^2$ is homogeneous of degree 1, i.e.

$$F(tx) = tF(x) \quad \text{for all } x \in \mathbb{R}^2,\ t \in \mathbb{R}.$$

Let $v \in \mathbb{R}^2$ be given. Then we have

$$\delta F(0; v) = F(v).$$

It is straightforward to construct a function $F : \mathbb{R}^2 \to \mathbb{R}^2$ that is homogeneous of degree 1, but fails to be linear.

Remark 14.15: Gâteaux differentiability at x_0 does not imply continuity at x_0 – even in finite dimensions – (and even if one assumes that the mapping $v \to \delta F(x_0; v)$ is linear and continuous) as the following example shows.

Example 14.16: Let $X = \mathbb{R}^2$, put

$$s = \{x \in \mathbb{R}^2 : x_2^2 < x_2 < 2x_2^2\},$$

and

$$F(x) = \begin{cases} 1 \text{ if } x \in S, \\ 0 \text{ if } x \notin S. \end{cases}$$

Then we have

$$\delta F(0; v) = 0 \quad \text{for all } v \in \mathbb{R}^2,$$

but F is not continuous at 0.

Proposition 14.17: Let X and Y be Banach spaces and $F : X \to Y$ and $x_0 \in X$ be given. Assume that F is Fréchet differentiable at x_0. Then F is Gâteaux differentiable at x_0 and

$$\delta F(x_0; v) = F'(x_0)v \quad \text{for all } v \in X.$$