Chapter I

Hilbert space

1 INNER-PRODUCTS AND NORMS

1.1 DEFINITION: An inner-product on a complex vector space \( \mathcal{V} \) is a Hermitian, complex-valued, positive definite sesquilinear form on \( \mathcal{V} \). That is a form satisfying

a. \( \langle u, v \rangle = \overline{\langle v, u \rangle} \)

b. \( \langle u, v \rangle \) is sesquilinear, that is linear in \( u \) and skew linear in \( v \):
\[ \langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle. \]

c. \( \langle u, u \rangle \geq 0 \), with \( \langle u, u \rangle = 0 \) if and only if \( u = 0 \).

An inner-product on a real vector space is defined similarly. The values of \( \langle v, u \rangle \) are real, \( \langle v, u \rangle \) is symmetric and linear in either variable.

An Inner-product space is a vector space endowed with an inner-product

1.2 DEFINITION: The norm of a vector \( v \) in an inner-product space is defined by:
\[ \|v\| = \sqrt{\langle v, v \rangle}. \]

Lemma (Cauchy–Schwarz).

(1.1) \[ |\langle u, v \rangle| \leq \|u\| \|v\|. \]

PROOF: If \( v \) is a scalar multiple of \( u \) we have equality. If \( v, u \) are not proportional, then for \( \lambda \in \mathbb{R} \),
\[ 0 < \langle u + \lambda v, u + \lambda v \rangle = \|u\|^2 + 2\lambda \Re \langle u, v \rangle + \lambda^2 \|v\|^2. \]
A quadratic polynomial with real coefficients and no real roots has negative discriminant, here \((\Re \langle u, v \rangle)^2 - \|u\|^2 \|v\|^2\).

For every \(\tau\) with \(|\tau| = 1\) we have \(|\Re \langle \tau u, v \rangle| \leq \|u\| \|v\|\); take \(\tau\) such that \(\langle \tau u, v \rangle = |\langle u, v \rangle|\).

The norm has the following properties:

a. Positivity: If \(v \neq 0\) then \(\|v\| > 0\); \(\|0\| = 0\).

b. Homogeneity: \(\|av\| = |a|\|v\|\) for scalars \(a\) and vectors \(v\).

c. The triangle inequality: \(\|v + u\| \leq \|v\| + \|u\|\).

d. The parallelogram law: \(\|v + u\|^2 + \|v - u\|^2 = 2(\|v\|^2 + \|u\|^2)\).

Properties a. and b. are obvious. Property c. is equivalent to

\[\|v\|^2 + \|u\|^2 + 2\Re \langle v, u \rangle \leq \|v\|^2 + \|u\|^2 + 2\|v\|\|u\|,\]

which reduces to (1.1). The parallelogram law is obtained by “opening brackets” in the inner-products that correspond to the various \(\| \|^2\).

The first three properties are common to all norms, whether defined by an inner-product or not. They imply that the norm can be viewed as length, and \(d(u, v) = \|u - v\|\) has the properties of a metric.

The parallelogram law, on the other hand, is specific to, and in fact characterizes, the norms defined by an inner-product.

**1.3 Definition:** The vectors \(v, u\) in an inner-product space are **mutually orthogonal**, denoted \(v \perp u\), if \(\langle v, u \rangle = 0\).

The vector \(v\) is orthogonal to a set \(A\), denoted \(v \perp A\), if it is orthogonal to every vector in \(A\). For any set \(A\), the set \(A^\perp = \{v : v \perp A\}\) is a subspace. Observe that \(A^\perp = (\text{span}[A])^\perp\).

**1.4 Definition:** A **Hilbert space** is an inner-product space which is **complete** with respect to the metric defined by the norm. Equivalently, it is a Banach space whose norm is given by an inner-product.

\(^*\)Observe that \(v \perp u \iff u \perp v\).

March 28, 2009
I. Hilbert space

Recall that a Banach space $B$ is uniformly convex if for every $\delta > 0$ there exists $\eta > 0$ such that for $x, y \in B$ such that $\|x\| = \|y\| = 1$ and $\|x - y\| \geq 2\delta$, we have $\|(x + y)/2\|^2 \leq 1 - \eta$.

The parallelogram law makes it obvious that a Hilbert space $\mathcal{H}$ is uniformly convex: if $u, v \in \mathcal{H}$, $\|u\| = \|v\| = 1$ and $\|u - v\| = 2\delta$ then

$$\|(u + v)/2\|^2 = 1 - \delta^2.$$ 

**Proposition.** Let $B$ be a uniformly convex Banach space, and let $E \subset B$ be closed and convex. For every $x \in B$ there exist a unique $y \in E$ such that

$$\|x - y\| = \text{dist}(x, E) = \inf_{z \in E} \|x - z\|.$$ 

**Proof:** There is no loss of generality in assuming dist$(x, E) = \inf_{z \in E} \|z\| = 1$, and $x = 0$.

The uniqueness: If $y_1, y_2 \in E$ are such that $\|y_1\| = \|y_2\| = 1$ then (uniform convexity) $(y_1 + y_2)/2 \in E$ will have a smaller norm unless $y_1 = y_2$.

The same argument gives the existence: Let $z_n \in E$ be such that $\|z_n\| \to 1$. Claim: $\{z_n\}$ is a Cauchy sequence. Otherwise it would contain two subsequences $\{z'_n\}, \{z''_n\}$ such that $\|z'_n - z''_n\| > 2\delta$ for some $\delta > 0$. For the appropriate $\eta > 0$, we would have (uniform convexity) $\lim\|(z'_n + z''_n)/2\| \leq (1 - \eta)$, a contradiction since $(z'_n + z''_n)/2 \in E$.

The (unique) closest point is $y = \lim z_n$. $\blacksquare$

The point $y$ given by the proposition is sometimes referred to as the projection of $x$ on $E$ and denoted $\pi_x x$. If $E$ is a closed subspace, $\pi_x$ is called the orthogonal projection onto $E$. This terminology is justified by the following lemma.

**1.5 Lemma.** If $\mathcal{H}$ is a Hilbert space and $E \subset \mathcal{H}$ is a closed subspace, then for every vector $x$, the projection $\pi_x x$ is the unique vector in $E$ such that $x - \pi_x x \perp E$.

**Proof:** Denote $y = \pi_x x$ and let $z \in E$. Since for $\lambda \in \mathbb{R}$,

$$\|x - y - \lambda z\|^2 = \|x - y\|^2 - 2\lambda \Re\langle z, x - y \rangle + \lambda^2 \|z\|^2$$

is minimal at $\lambda = 0$, $\Re\langle z, x - y \rangle = 0$. Replacing $z$ by $-iz$ we have $\Im\langle z, x - y \rangle = \Re\langle -iz, x - y \rangle = 0$, and $x - y \perp E$.

If $y_1 \in E$ and $x - y_1 \perp E$, then $y_1 - y$ is perpendicular to both $x - y$ and $x - y_1$, hence to their difference $y_1 - y$, and $y_1 = y$. $\blacksquare$

March 28, 2009
Corollary. If $E$ is a proper subspace then $E^\perp = \{ v \in \mathcal{H} : v \perp E \}$ is nontrivial, in fact $\mathcal{H} = E \oplus E^\perp$.

Theorem. If $\mathcal{H}$ is a Hilbert space and $E \subset \mathcal{H}$ is a closed subspace, $E \neq \{0\}$, then the map $\pi_E : x \mapsto \pi_E x$ given by proposition 1.4 is a linear operator of norm 1.

Proof: Let $x_1, x_2 \in \mathcal{H}$, $a_1, a_2 \in \mathbb{C}$, and write $y_j = \pi_E x_j$. Then

$$a_1y_1 + a_2y_2 - (a_1x_1 + a_2x_2) = a_1(y_1 - x_1) + a_2(y_2 - x_2) \perp E$$

which, by the lemma, proves that $\pi_E (a_1x_1 + a_2x_2) = a_1y_1 + a_2y_2$.

The fact that for all $x$, $\|\pi_E x\| \leq \|x\|$ follows from the fact that $0 \in E$ and is a contender for $\pi_E x$. Finally, $\pi_E$ is the identity on the non-trivial $E$, and $\|\pi_E\| = 1$.

1.6 A sequence $\{u_j\}$ is orthonormal if the vectors $v_j$ are “normalized”, $\|v_j\| = 1$, and are pairwise orthogonal, that is if

$$\langle u_j, u_k \rangle = \delta_{j,k} = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k. \end{cases}$$

Observe that if $\{u_1, \ldots, u_m\}$ is orthonormal and $a_j$ are complex numbers, $j = 1, \ldots, m$, then

$$\| \sum_{j=1}^m a_j u_j \|^2 = \sum_{j=1}^m |a_j|^2.$$ 

since $\langle \sum_{j=1}^m a_j u_j, \sum_{k=1}^m a_k u_k \rangle = \sum_{j,k} a_j \bar{a}_k \langle u_j, u_k \rangle = \sum |a_j|^2$. This clearly implies the following proposition.

Proposition. Let $\mathcal{H}$ be a Hilbert space and $\{u_j\}$ an orthonormal sequence in it. Then, if $a_j \in \mathbb{C}$ and $\sum |a_j|^2 < \infty$, the series $\sum a_j u_j$ converges in norm to a vector $v \in \mathcal{H}$ such that $\|v\|^2 = \sum |a_j|^2$ and $a_j = \langle v, u_j \rangle$.\[\dagger\]

\[\dagger\]This is the Pythagorean theorem.

March 28, 2009
Observe further that if $v \in \mathcal{H}$ is arbitrary and $\{u_1, \ldots, u_m\}$ is orthonormal, then the vector $v - \sum_1^m \langle v, u_j \rangle u_j$ is orthogonal to $\text{span}[\{u_j\}_{j=1}^m]$ and hence to $\sum_1^m \langle v, u_j \rangle u_j$. It follows that

\begin{equation}
\|v\|^2 = \|v - \sum_1^m \langle v, u_j \rangle u_j\|^2 + \sum_1^m |\langle v, u_j \rangle|^2,
\end{equation}

and, in particular, $\sum_1^m |\langle v, u_j \rangle|^2 \leq \|v\|^2$, with equality if, and only if, $v = \sum_1^m \langle v, u_j \rangle u_j$.

If $\{u_j\}$ is an infinite orthonormal sequence and $\mathcal{U}$ is the closed linear span of $\{u_j\}$, the previous remark applied to any finite subsequence of $\{u_j\}$, shows that, for any $v \in \mathcal{H}$,

\begin{equation}
\sum_1^\infty |\langle v, u_j \rangle|^2 \leq \|v\|^2.
\end{equation}

The proposition guarantees that the series $\sum_1^\infty \langle v, u_j \rangle u_j$ converges in norm to a vector $\tilde{v} \in \mathcal{U}$. Since $v - \tilde{v}$ is orthogonal to every $u_j$, and hence to $\mathcal{U}$, we have $\tilde{v} = \pi_\mathcal{U} v$.

**Theorem.** Let $\mathcal{H}$ be a Hilbert space, $\{u_j\}$ an orthonormal sequence in it, and $\mathcal{U} = \text{span}[\{u_j\}]$. For every $v \in \mathcal{H}$ the series $\sum \langle v, u_j \rangle u_j$ converges in norm to $\pi_\mathcal{U} v$.

Moreover, $v - \pi_\mathcal{U} v \perp \mathcal{U}$, so that if $\mathcal{U} = \mathcal{H}$ we have for every $v \in \mathcal{H}$, $v = \sum \langle v, u_j \rangle u_j$.

**DEFINITION:** A sequence $\{u_j\} \subset \mathcal{H}$ is an orthonormal basis for $\mathcal{H}$ if it is orthonormal and spans $\mathcal{H}$. The last part of the theorem can be stated as:

If $\{u_j\}$ is an orthonormal basis for $\mathcal{H}$ then $v = \sum \langle v, u_j \rangle u_j$ for every $v \in \mathcal{H}$.

1.7 A set $A \subset \mathcal{H}$ is independent if every finite subset thereof is independent.

**Proposition (Gram-Schmidt).** Let $\{v_j\}_{j=1}^m$, $m \leq \infty$, be independent. Then there exists an orthonormal $\{u_j\}_{j=1}^m$, such that for all $k \in [1, m]$, (if $m = \infty$, then for all finite $k$).

\begin{equation}
\text{span}[u_1, \ldots, u_k] = \text{span}[v_1, \ldots, v_k].
\end{equation}
Proof: (By induction). The independence of \( \{v_j\} \) implies that \( v_1 \neq 0 \). Write \( u_1 = v_1/\|v_1\| \). Then \( u_1 \) is normal and (1.7) is satisfied for \( k = 1 \).

Assume that \( \{u_1, \ldots, u_l\} \) is orthonormal and that (1.7) is satisfied for \( k \leq l \). Since \( v_{l+1} \not\in \text{span}\{v_1, \ldots, v_l\} \) the vector

\[
\tilde{v}_{l+1} = v_{l+1} - \sum_{j=1}^l \langle v_{l+1}, u_j \rangle u_j
\]

is non-zero and we set \( u_{l+1} = \tilde{v}_{l+1}/\|\tilde{v}_{l+1}\| \).

One immediate corollary is that a separable Hilbert space \( \mathcal{H} \) has an orthonormal basis. All we need is an independent sequence \( \{v_j\} \) that spans \( \mathcal{H} \). For that we can take a sequence \( \{w_j\} \) which is dense in \( \mathcal{H} \) and discard from it every element that is a linear combination of ones that precede it.

1.8 Proposition (Parseval’s identity). If \( \{u_j\} \) is an orthonormal basis for \( \mathcal{H} \) then, for all \( v, w \in \mathcal{H} \),

\[
\langle v, w \rangle = \sum \langle v, u_j \rangle \langle w, u_j \rangle;
\]

in particular,

\[
\|v\|^2 = \sum |\langle v, u_j \rangle|^2.
\]

Proof:

\[
\langle v, w \rangle = \langle \sum \langle v, u_j \rangle u_j, \sum \langle w, u_j \rangle u_j \rangle = \sum_{j,l} \langle v, u_j \rangle \langle w, u_l \rangle \langle u_j, u_l \rangle = \sum_j \langle v, u_j \rangle \langle w, u_j \rangle
\]

\[\Box\]

Bessel’s inequality states that if \( \{u_j\} \) is orthonormal, (but not necessarily a basis) then for every \( v \in \mathcal{H} \), \( \sum |\langle v, u_j \rangle|^2 \leq \|v\|^2 \). This is clearly weaker than (1.9).

\[\uparrow\]Topologically, that is, such that the finite linear combinations thereof form a dense subspace of \( \mathcal{H} \).

March 28, 2009
1.9 Polarization. “Polarization” is an expression of a bilinear form on $\mathcal{H}$ in terms of the corresponding quadratic form. The underlying assumption is that $\mathcal{H}$ is a complex vector space.

One main example is: In a complex inner-product space, the inner-product is determined by the norm, in fact, (polarization formula)

$$\langle u, v \rangle = \frac{1}{4} \left( \|u + v\|^2 - \|u - v\|^2 + i\|u + iv\|^2 - i\|u - iv\|^2 \right).$$

More generally, if $T$ is a linear operator on a complex inner-product space then, for $z \in \mathbb{C}$,

$$\langle T(u + zv, u + zv) = \langle Tu, u \rangle + |z|^2 \langle Tv, v \rangle + \langle Tzv, u \rangle + \langle Tu, zv \rangle$$

Using this with $z = 1$, and then with $z = i$ gives

$$\langle Tu, v \rangle + \langle Tv, u \rangle = \langle Tu + v, u + v \rangle - \langle Tu, u \rangle - \langle Tv, v \rangle$$

$$\langle Tu, v \rangle - \langle Tv, u \rangle = i \left( \langle T(u + iv, u + iv) - \langle Tu, u \rangle - \langle Tv, v \rangle \right)$$

which implies

$$\langle Tu, v \rangle = \frac{1}{2} \left( \langle T(u + v, u + v) - \langle Tu, u \rangle - \langle Tv, v \rangle \right)$$

$$+ \frac{i}{2} \left( \langle T(u + iv, u + iv) - \langle Tu, u \rangle - \langle Tv, v \rangle \right).$$

An immediate consequence of (1.13) is the following lemma

**Lemma.** If $S, T \in \mathcal{LH}$ and $\langle Tv, v \rangle = \langle Sv, v \rangle$ for all $v \in \mathcal{H}$, then $T = S$.

If $T \in \mathcal{LH}$ and the quadratic form $\langle Tu, u \rangle$ is real-valued, then the first row of (1.13) gives the real-part of $\langle Tu, v \rangle$, and the second row gives the imaginary part. In particular we have in this case

$$\Re \langle Tu, v \rangle = \Re \langle Tv, u \rangle \quad \text{and} \quad \Re \langle Tiu, v \rangle = \Re \langle Tv, iu \rangle,$$

that is

$$\langle Tu, v \rangle = \overline{\langle Tv, u \rangle} = \langle u, Tv \rangle.$$
EXERCISES FOR SECTION 1.

1.1 — BASIC ORTHOGONALITY. —

a. Let $\mathcal{H}$ be a real inner-product space. $v, u$ are orthogonal if, and only if, $\|v + u\|^2 = \|v\|^2 + \|u\|^2$.

b. If $\mathcal{H}$ is a complex inner-product space, and $v, u \in \mathcal{H}$, the condition $\|v + u\|^2 = \|v\|^2 + \|u\|^2$ is necessary, but not sufficient, for $v \perp u$.

Hint: Connect to the condition $\langle u, v \rangle$ purely imaginary.

c. If $\mathcal{H}$ is a complex inner-product space, and $v, u \in \mathcal{H}$, the condition: For all $a, b \in \mathbb{C}$, $\|av + bu\|^2 = |a|^2\|v\|^2 + |b|^2\|u\|^2$ is necessary and sufficient for $v \perp u$.

d. Let $\mathcal{V}$ and $\mathcal{W}$ be subspaces of $\mathcal{H}$. Prove that $\mathcal{V} \perp \mathcal{W}$ if, and only if, for $v \in \mathcal{V}$ and $u \in \mathcal{W}$, $\|v + u\|^2 = \|v\|^2 + \|u\|^2$.

e. The set $\{v_1, \ldots, v_m\}$ is orthonormal if, and only if $\|\sum a_j v_j\|^2 = \sum |a_j|^2$ for all choices of scalars $a_j, j = 1, \ldots, m$. (Here $\mathcal{H}$ is either real or complex.)

1.2 Show that the map $\pi_\mathcal{W}$ defined in 1.4 is an idempotent linear operator$^\dagger$ and is independent of the particular basis used in its definition.

1.3 Show that the sequence $\{u_1, \ldots, u_m\}$ obtained by the Gram-Schmidt procedure is essentially unique: each $u_j$ is unique up to multiplication by a number of modulus 1.

Hint: If the set $\{v_1, \ldots, v_m\}$ is independent, $\mathcal{W}_0 = \{0\}$, and $\mathcal{W}_k = \text{span}\{v_1, \ldots, v_k\}$, $k = 1, \ldots, m - 1$, then $u_j$ is $c \pi_{\mathcal{W}_{j-1}^\perp} v_j$, with $|c| = \|\pi_{\mathcal{W}_{j-1}^\perp} v_j\|^{-1}$.

1.4 — COMPACT SETS IN HILBERT SPACE. —

a. Prove that a bounded closed subset $E \subset \mathcal{H}$ is compact if, and only if, it can be approximated in norm by finite dimensional subsets.

b. Let $\{e_n\}$ be an orthonormal basis for $\mathcal{H}$. The (standard) cube defined by a sequence $\{c_n\}$ of positive numbers is the set

$$C(\{c_n\}) = \{v \in \mathcal{H} : v = \sum a_n e_n, |a_n| \leq c_n\}.$$  

(The set $C(\{1\})$ is known as “the Hilbert cube”). Prove that $C(\{c_n\})$ is compact if, and only if, $\sum c_n^2 < \infty$.

c. Let $\mathcal{H} = \bigoplus \mathcal{H}_j$ be an orthogonal decomposition of $\mathcal{H}$ into finite dimensional subspaces $\mathcal{H}_j$. and let $\{c_n\}$ be a sequence of positive numbers. The generalized cube determined by this is the set

$$C(\{\mathcal{H}_j\}, \{c_n\}) = \{v \in \mathcal{H} : v = \sum v_j \text{ with } v_j \in \mathcal{H}_j \text{ and } \|v_j\| \leq c_j\}.$$  

$^\dagger$An operator $T$ is idempotent if $T^2 = T$.

March 28, 2009
Prove that the condition $\sum c_n^2 < \infty$ is again necessary and sufficient for the compactness of $C(\{H_j\}, \{c_n\})$

d. Prove that any compact $E \subset H$ is contained in a compact generalized cube.

e. Show that there exist compact sets in $H$ (assumed infinite dimensional) which are not contained in any compact (standard) cube.

f. Given a compact $E \subset H$ there exists a compact $F \subset H$ which contains uncountably many disjoint translates of $E$.

Hint: $C(\{3c_n\})$ contains uncountably many disjoint translates of $C(\{c_n\})$.

1.5 Let $\{e_n\}$ be an orthonormal basis for $H$. Let $\{c_n\}$ be positive numbers and assume $\sum c_n^2 = 1$.

Prove that if the cube $C(\{c_n\}) = \{v: v = \sum a_n e_n, |a_n| \leq c_n\}$ contains an $n$ dimensional ball of radius $r$, then $Nr^2 \leq 1$.

Hint: Consider the model $H = L^2([0,1], dx), e_n = c_n^{-1} I_{I_n}$, where $|I_n| = c_n^2$, and $\mathcal{I} = \{I_n\}$ is a partition of $[0,1]$. Check that $C(\{c_n\})$ is the set of all the functions which are bounded by 1 and are $\mathcal{I}$-measurable. In the span of $\{u_n\}$ the $L^\infty$ norm is bounded by $r^{-1}$ times the $L^1$ norm.

2 DUALITY AND THE ADJOINT

2.1 $H$ as its own dual. The inner-product defined in $H$ associates with every vector $u \in H$ the linear functional $\varphi_u : v \mapsto \langle v, u \rangle$.

Theorem. Let $\varphi$ be a linear functional on a Hilbert space $H$. Then there exist a unique $u^* \in H$ such that for all $v \in H$,

$$\varphi(v) = \varphi_{u^*}(v) = \langle v, u^* \rangle. \quad (2.1)$$

Proof: If $\varphi = 0$ we can take $u^* = 0$. If $\varphi \neq 0$, then the kernel of $\varphi$, $\ker(\varphi) = \{v \in H : \varphi(v) = 0\}$ is a closed subspace of codimension 1 in $H$. Let $u \in \ker(\varphi)^\perp, \|u\| = 1$. Define $c = \varphi(u)$ and set $u^* = cu$. Then the functionals $\varphi$ and $\varphi_{u^*}$ agree both on $u$ (and its multiples) and on $\ker(\varphi)$; it follows that $\varphi = \varphi_{u^*}$. \hfill $\blacksquare$

2.2 The weak topology on $H$. Since $H$ is its own dual, the weak topology and the weak* topology on $H$ are the same. In particular, the unit ball $B(0,1) = \{v : \|v\| = 1\}$, is weakly compact.

\[ B = \{v : v = \sum a_n u_n, \sum |a_n|^2 \leq r^2\}. \]

March 28, 2009
2.3 The adjoint of an operator. Once we identify $\mathcal{H}$ with its dual space, the adjoint of an operator $T \in \mathcal{L}(\mathcal{H})$ is again an operator on $\mathcal{H}$. Given $u \in \mathcal{H}$, the map $v \mapsto \langle Tv, u \rangle$ is a linear functional and therefore equal to $v \mapsto \langle v, w \rangle$ for some $w \in \mathcal{H}$. We write $w = T^* u$ and check that $u \mapsto w$ is linear. In other words $T^*$ is a linear operator on $\mathcal{H}$, characterized by

$$\langle Tv, u \rangle = \langle v, T^* u \rangle. \tag{2.2}$$

**Lemma.** For $T \in \mathcal{L}(\mathcal{H})$, $(T^*)^* = T$.

**Proof:** \[ \langle T^* v, u \rangle = \overline{\langle u, T^* v \rangle} = \overline{\langle Tu, v \rangle} = \langle v, Tu \rangle. \]

**Proposition.** For $T \in \mathcal{L}(\mathcal{H})$, $\text{range}(T)^\perp = \ker(T^*)$.

**Proof:** $\langle Tx, y \rangle = \langle x, T^* y \rangle$ so that $y \perp \text{range}(T)$ if, and only if $T^* y \perp \mathcal{H}$, that is, $y \in \ker(T^*)$.

The subspace $\mathcal{D}(T) = (\ker(T))^\perp$ is called the essential domain of $T$. Thus, $\mathcal{D}(T) = \text{range}(T^*)$ and $\mathcal{D}(T^*) = \text{range}(T)$.

2.4 **Proposition.** If $T \in \mathcal{L}(\mathcal{H})$ and $\langle Tu, u \rangle$ is real-valued for $u \in \mathcal{H}$, then $T$ is self-adjoint, that is $T^* = T$.

**Proof:** See (1.15).
Linear operators

1 TOPOLOGIES AND CONTINUITY

1.1 The topologies used for \( \mathcal{H} \) are the norm topology and the weak topology. As \( \mathcal{H} \) is self dual, the weak topology is the same as the weak* topology. In particular, the unit ball \( B(0,1) = \{ v : \| v \| \leq 1 \} \) of \( \mathcal{H} \) is weakly compact.

1.2 We denote by \( \mathcal{LH} \) the algebra of bounded linear operators on \( \mathcal{H} \). Recall that a linear operator on \( \mathcal{H} \) is bounded if and only if it is (norm) continuous.

The three most commonly used topologies on \( \mathcal{LH} \) are:

a. The norm topology (or uniform topology) on \( \mathcal{LH} \) is the metric topology determined by the norm \( \| T \| = \| T \|_{\mathcal{LH}} = \sup_{v \in \mathcal{H}, \| v \| \leq 1} \| Tv \| \).

b. The Strong operator topology defined by the family of seminorms \( \| T \|_v = \| Tv \|_{\mathcal{H}}, v \in \mathcal{H} \).

c. The Weak operator topology defined by the family of seminorms \( \| T \|_{v,u} = | \langle Tv, u \rangle |, v, u \in \mathcal{H} \).

1.3 Lemma. If \( T \in \mathcal{LH} \) then it is continuous in the weak topology of \( \mathcal{H} \).

Proof: Sets of the form \( O = \cap_{j=1}^k \{ y : | \langle y - Tv, u_j \rangle | < \varepsilon \} \), given by a finite collection \( \{ u_j \}_1^k \) and some \( \varepsilon > 0 \), form a basis of neighborhoods of \( Tv \) in the weak topology of \( \mathcal{H} \). Each weak neighborhoods of \( Tv \) contains such sets.

The set \( A = \cap_{j=1}^k \{ x : | \langle x - v, T^* u_j \rangle | < \varepsilon \} \) is a weak neighborhood of \( v \), and is mapped by \( T \) into \( O \).

Corollary. If \( T \in \mathcal{LH} \) then \( TB(0,1) \) is norm closed.
PROOF: $TB(0, 1)$ is weakly compact, (a continuous image of the weakly compact $B(0, 1)$ in $\mathcal{H}$). Since it is also convex, it is norm closed (Hahn–Banach).

2 SPECTRUM

2.1 DEFINITION: We say that an operator $T \in \mathcal{L} \mathcal{H}$ is bounded below if $\inf_{\|x\|=1} \|T x\| > 0$.

An operator $T \in \mathcal{L} \mathcal{H}$ that is injective (1 − 1 into), has dense range, and is bounded below is invertible. It follows that $T$ fails to be invertible for one of the following mutually exclusive reasons:

s-1 $T$ is not injective, that is: $\ker(T) = \{x : T x = 0\} \neq \{0\}$.

s-2 $T$ is injective and $T \mathcal{H}$ is dense in $\mathcal{H}$, but $T$ is not “bounded below”.

s-3 $T$ is injective but $T \mathcal{H}$ is not dense in $\mathcal{H}$.

Condition s-2 is equivalent to

s-2* $T$ is injective and $T \mathcal{H}$ is a dense proper subspace of $\mathcal{H}$.

The spectrum of an operator $T$ is, by definition, the set

$$\sigma(T) = \{\lambda : \not\exists (T - \lambda)^{-1}\}$$

The point-spectrum $\sigma_p(T)$ is the set of $\lambda$ such that $T - \lambda$ satisfies s-1 above; in other words, $\sigma_p(T)$ consists of the eigenvalues of $T$.

The continuous-spectrum $\sigma_c(T)$ is, by definition, the set of $\lambda$ such that $T - \lambda$ satisfies s-2 above. $\sigma_c(T)$ is also referred to as the approximate-point-spectrum and its elements as approximate eigenvalues. The residual-spectrum $\sigma_r(T)$ is the set of $\lambda$ such that $T - \lambda$ satisfies s-3 above.

2.2 DEFINITION: The resolvent set of $T$, denoted $\rho(T)$, is the complement of $\sigma(T)$ in $\mathbb{C}$.

Proposition. The resolvent $\rho(T)$ is open in $\mathbb{C}$, $\rho(T) \supset \{z : |z| > \|T\|\}$, and the (operator valued) function $F(\lambda) = (T - \lambda)^{-1}$ is holomorphic in $\rho(T)$.

March 28, 2009
PROOF: For $|\lambda| > \|T\|$, the series $\sum_0^\infty T^n \lambda^{-n-1}$ converges in norm and its sum is $-(T - \lambda)^{-1}$. If $\lambda_0 \in \rho(T)$ and $|\lambda - \lambda_0| < \|(T - \lambda_0 I)^{-1}\|^{-1}$, we can write

$$(T - \lambda) = (T - \lambda_0) - (\lambda - \lambda_0) = (T - \lambda_0)(1 - (\lambda - \lambda_0)(T - \lambda_0)^{-1})$$

so that

$$(T - \lambda)^{-1} = (T - \lambda_0)^{-1}(1 - (\lambda - \lambda_0)(T - \lambda_0)^{-1})^{-1}$$

$$= \sum_0^\infty (T - \lambda_0)^{-n-1}(\lambda - \lambda_0)^n,$$

This shows that the disk $|\lambda - \lambda_0| < \|(T - \lambda_0 I)^{-1}\|^{-1}$ is contained in $\rho(T)$, and that $(T - \lambda)^{-1}$ is holomorphic in it.

**Corollary (of the proof).** Let $\lambda \in \rho(T)$, then

$$\|(T - \lambda)^{-1}\| \geq \text{dist}(\lambda, \sigma(T))^{-1}.$$  

In particular, if $\lambda \in \text{bdry}(\sigma(T))$, then $T - \lambda$ is not bounded below, and it follows that

$$\text{bdry}(\sigma(T)) \subset \sigma_p(T) \cup \sigma_c(T).$$

**2.3 Definition:** The spectral norm of $T$, denoted $\|T\|_{sp}$, is defined by

$$\|T\|_{sp} = \max_{\lambda \in \sigma(T)} |\lambda|.$$  

**Proposition.** $\|T\|_{sp} = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}$.

**Proof:** There are two issues: the existence of the limit and its relation to the spectral norm.

For the existence of the limit notice that $a_n = \log \|T^n\|$ is subadditive: $a_{n+m} \leq a_n + a_m$. This implies $a_{kn} \leq ka_n$, or $\frac{1}{kn}a_{kn} \leq \frac{1}{n}a_n$, for all $k \in \mathbb{N}$, and in turn, implies $\lim \frac{1}{n}a_n = \lim \frac{1}{n}a_n$.

The series $\sum T^n \lambda^{-n}$ converges in norm provided $|\lambda| > \lim \|T^n\|^\frac{1}{n}$. This gives $\|T\|_{sp} \leq \lim_{n\to\infty} \|T^n\|^\frac{1}{n}$. On the other hand, for all $x, y \in \mathcal{H}$, the Laurent expansion $(T - \lambda I)^{-1}x, y) = \sum \langle T^n x, y \rangle \lambda^{-n}$ converges for $|\lambda| > \|T\|_{sp}$, and for every such $\lambda$,

$$\langle T^n x, y \rangle \lambda^{-n} = O(1) \quad \text{for all } x, y \in \mathcal{H}.$$  

March 28, 2009
For every \( x \in \mathcal{H} \), the linear functionals \( y \mapsto \langle T^n x, y \rangle \lambda^{-n} \) are bounded for every \( y \in \mathcal{H} \) and by the uniform boundedness principle we obtain \( \lambda^{-n} \| T^n x \| = O(1) \). Since this holds for every \( x \in \mathcal{H} \), the uniform boundedness principle applies again and we have \( \lambda^{-n} \| T^n \| = O(1) \).

**2.4 Lemma** (Spectral mapping theorem). Let \( T \in \mathcal{L} \mathcal{H} \) and let \( P(T) = \sum_n a_n T^n \) with \( a_n \in \mathbb{C} \), then

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

**PROOF:** For \( \lambda \in \mathbb{C} \) let \( \{ \lambda_j \} \) be the roots of \( P(z) - \lambda \) so that (assuming \( a_N \neq 0 \)) \( P(z) - \lambda = a_N \prod_j (z - \lambda_j) \). Then \( P(T) - \lambda = a_N \prod_j (T - \lambda_j) \) is invertible unless one of the factors isn’t, which happens precisely if \( \lambda_j \in \sigma(T) \) for some \( j \), that is, if \( \lambda \in P(\sigma(T)) \).

2.5 Since the adjoint of an invertible operator is invertible, and the adjoint of \( T - \lambda \) is \( T^* - \bar{\lambda} \), we have

\[
\sigma(T^*) = \overline{\sigma(T)}, \quad \text{and} \quad \rho(T^*) = \overline{\rho(T)}.
\]

**2.6 Lemma.** For any \( T \in \mathcal{L} \mathcal{H} \), \( \ker(T^* T) = \ker(T) \).

**PROOF:** For any \( x \in \mathcal{H} \),

\[
\langle T^* T x, x \rangle = \langle T x, T x \rangle = \| T x \|^2,
\]

so that if \( T^* T x = 0 \) then \( T x = 0 \). The other implication is obvious.

EXERCISES FOR SECTION 2.

2.1 Prove that the spectrum \( \sigma(T) \) depends continuously on \( T \): For every \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that if \( \| T - T_1 \| < \delta \), then every point in \( \sigma(T_1) \) lies within \( \varepsilon \) from \( \sigma(T) \).

**Hint:** Check that, for some constant \( K \), \( \| (T - \lambda)^{-1} \| \leq K \) on the set \( A = \{ \lambda : \text{dist}(\lambda, \sigma(T)) \geq \varepsilon \} \). Take \( \delta < K^{-1} \), then

\[
T_1 - \lambda = (T - \lambda)(1 + (T - \lambda)^{-1}(T_1 - T))
\]

and for \( \lambda \in A \) both factors are invertible.

MARCH 28, 2009
3 SELF-ADJOINT AND NORMAL OPERATORS

3.1 Self-adjoint operators.

**Definition:** An operator \( S \in \mathcal{L}(\mathcal{H}) \) is self-adjoint or Hermitian if \( S^* = S \). An equivalent condition is that the associated bilinear form, \( \langle Sx, y \rangle \), be Hermitian, i.e., for all \( x, y \in \mathcal{H} \) we have
\[
\langle Sx, y \rangle = \overline{\langle Sy, x \rangle} = \langle x, Sy \rangle.
\]

**Definition:** An operator \( S \in \mathcal{L}(\mathcal{H}) \) is positive if \( \langle Sx, x \rangle \geq 0 \) for all \( x \in \mathcal{H} \). Notice that a positive operator is self-adjoint.

If \( T \in \mathcal{L}(\mathcal{H}) \) is arbitrary, the operators
\[
\Re T = \frac{T + T^*}{2}, \quad \Im T = \frac{T - T^*}{2i}, \quad \text{and} \quad |T|^2 = T^*T
\]
are self-adjoint, and \( T = \Re T + i\Im T \). Since \( \langle T^*Tx, x \rangle = \|Tx\|^2 \), the operator \( |T|^2 \) is positive.

Observe that for any \( T \in \mathcal{L}(\mathcal{H}) \), range(\( T \))\(^{\perp} = \ker(\overline{T}) \), the range of an injective self-adjoint operator \( S \) is dense in \( \mathcal{H} \). It follows that if \( S \) is self-adjoint, the residual spectrum of \( S \) is void and \( S \) is invertible if, and only if, it is bounded below.

**Lemma.** If \( S \in \mathcal{L}(\mathcal{H}) \) is self-adjoint then \( \sigma(S) \subset \mathbb{R} \). If \( S \) is positive then \( \sigma(S) \subset [0, \infty) \).

**Proof:** For \( a, c \in \mathbb{R} \), \( \| (S - a + ic)x \|^2 = \| (S - a)x \|^2 + c^2\|x\|^2 \), as can be seen by expanding \( \langle (S - a + ic)x, (S - a + ic)x \rangle \), and \( S - a + ic \) is bounded below. Since \( (S - a + ic)^* = S - a - ic \) has the same form and is bounded below for the same reason, the observation above implies that both have dense range and are therefore invertible.

**Theorem.** Let \( S \) be self-adjoint on \( \mathcal{H} \). Then,

a. \( \ker(S^2) = \ker(S) \).

**Proof:** Lemma 2.6.

b. \( \|S^2\| = \|S\|^2 \), hence \( \|S\|_{sp} = \|S\| \).

**Proof:** \( \|S\|^2 = \sup_{\|x\| = 1} \|\langle Sx, Sx \rangle\| = \sup_{\|x\| = 1} |\langle S^2x, x \rangle| \leq \|S^2\| \).

The reverse inequality is universal.

c. The quadratic form \( \langle Sx, x \rangle \) is real-valued.

March 28, 2009
Proof: \( \langle Sx, x \rangle = \langle x, Sx \rangle = \langle Sx, x \rangle \).

d. \( \|S\| = \sup_{\|x\| = 1} |\langle Sx, x \rangle| \).

Proof: With no loss of generality assume \( \|S\| = 1 \). If \( \|x\| = 1 \) and \( \|Sx\| \sim 1 \) then
\[
\langle S^2 x, x \rangle = \langle Sx, Sx \rangle \sim 1
\]
so that \( S^2 x \sim x \). If \( x + Sx \sim 0 \), then \( \langle Sx, x \rangle \sim -1 \); otherwise let \( y = \|x + Sx\|^{-1} (x + Sx) \), and observe that \( \langle Sy, y \rangle \sim 1 \).

e. If \( Sv_1 = \lambda_1 v_1 \) and \( Sv_2 = \lambda_2 v_2 \), \( \lambda_2 \neq \lambda_1 \), then \( \langle v_1, v_2 \rangle = 0 \).

Proof: \( \lambda_1 \langle v_1, v_2 \rangle = \langle Sv_1, v_2 \rangle = \langle v_1, Sv_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle \) which is possible only if \( \langle v_1, v_2 \rangle = 0 \).

3.2 Normal operators. An operator \( T \in \mathcal{L} \mathcal{H} \) is normal if \( T^* T = T T^* \). An equivalent condition is that \( \Re T = \frac{T + T^*}{2} \) and \( \Im T = \frac{T - T^*}{2i} \) commute; another is that \( T \) is a member of a commutative self-adjoint subalgebra of \( \mathcal{L} \mathcal{H} \).

Lemma. If \( T \) is normal, then \( \ker(T) = \ker(T^2) \).

Proof: By Lemma 2.6, \( \ker(T) = \ker(T^* T) = \ker((T^* T)^2) = \ker(T^2) \). (The second equality because \( T^* T \) is self-adjoint, the third because \( T \) is assumed normal).

Notice that this implies \( \ker(T) = \ker(T^n) \) for any positive integer \( n \).

Proposition. If \( T \) is normal, then

- a. For all \( v \in \mathcal{H} \), \( \|T^* v\| = \|Tv\| \).
- b. \( \|T^*\| = \|T\| \).
- c. \( \|T^* T\| = \|T^2\| \).
- d. \( \|T^2\| = \|T\|^2 \), hence \( \|T\|_{sp} = \|T\| \).

Proof:

- a. \( \|T^* v\|^2 = \langle T^* v, T^* v \rangle = \langle TT^* v, v \rangle = \langle T^* T v, v \rangle = \langle T v, T v \rangle = \|Tv\|^2 \).

b. Follows immediately from a.

March 28, 2009
c. $T^*T$ is self-adjoint so that
\[ \|T^*T\|^2 = \|(T^*T)^2\| = \|T^2(T^*)^2\| \leq \|T^2\|\|(T^*)^2\| = \|T^2\|^2. \]

d. $\|T\|^2 = \sup_{\|x\|=1} |\langle Tx, Tx \rangle| = \sup_{\|x\|=1} |\langle T^*Tx, x \rangle| = \|T^*T\| = \|T^2\|.$

The proof of part a. can be used to show the converse: if for all $v \in \mathcal{H}$, $\|T^*v\| = \|Tv\|$ then $\langle TT^*v, v \rangle = \langle T^*Tv, v \rangle$. Polarization (Lemma 1.9) gives $TT^* = T^*T$.

3.3 Proposition. Assume $T$ normal. Then if $Tv = \lambda v$ then $T^*v = \overline{\lambda}v$. Also, $\Re T \cdot v = \Re \lambda v$, and $\Im T \cdot v = \Im \lambda v$.

Proof: With no loss of generality assume $\|v\| = 1$. Then $\langle T^*v, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \overline{\lambda}$. Since $|\langle T^*v, v \rangle| = \|T^*v\|\|v\|$, $T^*v$ is a scalar multiple of $v$, and since $\langle T^*v, v \rangle = \overline{\lambda}$, the scalar is $\overline{\lambda}$.

$\Re T \cdot v = \frac{Tv + T^*v}{2} = \frac{\lambda + \overline{\lambda}}{2}v = \Re \lambda v$; similarly for $\Im T \cdot v$.

3.4 Partial isometries. An operator $O \in \mathcal{L} \mathcal{H}$ is a partial isometry if for all $v \in \mathcal{D}(O)$ we have $\|OV\| = \|v\|$. $(\mathcal{D}(O) = \ker(O)^\perp$ is the essential domain of $O)$.

Lemma. Let $O$ be a partial isometry with essential domain $\mathcal{D}(O)$ and range $\text{range}(O)$. Then $O^*$ is a partial isometry with essential domain $\mathcal{D}(O^*) = \text{range}(O)$, $O^*O$ is the orthogonal projection on $\mathcal{D}(O)$ and $OO^*$ is the orthogonal projection on $\mathcal{D}(O^*)$.

Proof: If $u, v \in \mathcal{D}(O)$ then
\[ \langle u, v \rangle = \langle Ou, Ov \rangle = \langle u, O^*Ov \rangle \]
which means that $v - O^*Ov$ is orthogonal to $\mathcal{D}(O)$. But $\text{range}(O^*) \subset \mathcal{D}(O)$ and hence $v - O^*Ov \in \mathcal{D}(O)$. This gives $O^*Ov = v$. If $v \in \mathcal{D}(O)^\perp$ then $Ov = 0$ and $O^*Ov = 0$. If also shows that $O^*$ is a partial isometry and $\mathcal{D}(O^*) = \text{range}(O)$.

Now the roles of $O$ and $O^*$ are exchangeable and the beginning of the proof shows that $OO^*$ is the orthogonal projection on $\mathcal{D}(O^*)$.
4 FUNCTIONAL CALCULUS FOR SELF-ADJOINT OPERATORS.

4.1 Theorem. Let $S \in \mathcal{L} \mathcal{H}$ be self-adjoint. If $P(S) = \sum_{n=0}^{N} a_n S^n$ is a polynomial with coefficients $a_n \in \mathbb{C}$, then $(P(S))^* = P(S)$, and $\|P(S)\| = \sup_{z \in \sigma(S)} |P(z)|$.

Proof: Clearly $(aS^i)^* = \bar{a}S^i$, so that $(P(S))^* = \overline{P(S)}$. By the Spectral Mapping Theorem, (Theorem 2.4), $\sigma(P(S)) = P(\sigma(S))$, and by part b. of Theorem 3.1, $\|P(S)\| = \sup \{|z| : z \in \sigma(P(S))\} = \sup_{z \in \sigma(S)} |P(z)|$.

We extend the map $P \mapsto P(S)$ by continuity to an isometry $\Phi_S$ of the sup-normed algebra $C(\sigma(S))$ into $\mathcal{L} \mathcal{V}$. $\Phi_S$ is a *-isomorphism\(^\dagger\), and introduces the "functional calculus": for each $F \in C(\sigma(S))$ we define $F(S) = \Phi_S(F)$. Notice that if $F$ is real-valued then $F(S)$ is self-adjoint.

4.2 Polar decomposition. One useful example: the absolute value $|T|$, of an (arbitrary) operator $T \in \mathcal{L} \mathcal{V}$, is defined by $|T| = \sqrt{T^* T}$. We observed in 3.1 that $T^* T$ is a positive operator and, by Lemma 3.1 the spectrum of the positive $T^* T$ is contained in $\mathbb{R}_+ = [0, \infty)$. The (restriction to $\sigma(T^* T)$) of the function $F(x) = \sqrt{x}$ is continuous and $|T| = F(T)$ is a well defined positive operator whose square is $T^* T$ (which explains the notation introduced in (3.1)). Observe that for every vector $v \in \mathcal{H}$

\begin{equation}
\| |T| v \|^2 = \langle |T| v, v \rangle = \langle T^* T v, v \rangle = \langle T v, T v \rangle = \| T v \|^2.
\end{equation}

Write $\mathcal{H}_0 = \ker(|T|) = \ker(T)$, and $\mathcal{H}_1 = \mathcal{H}_0^\perp$. The restrictions of both $T$ and $|T|$ to $\mathcal{H}_1$ are injective, so that the map $O : |T| v \mapsto T v$ is well defined for $v \in \mathcal{H}_1$, is linear, and by (4.1), is an isometry from $\text{range}(|T|)$ onto $\text{range}(T)$ and so can be extended to a linear isometry from $\text{range}(|T|)$ onto $\text{range}(T)$, the corresponding closures. We extend the definition of $O$ to the entire space by defining it to be zero\(^\dagger\) on $\text{range}(|T|)^\perp$ and extend by linearity.

By Lemma 3.4 we have $O^* O = \pi_{\text{range}(|T|)}$, the orthogonal projection on $\text{range}(|T|)$. This proves the following theorem:

\(^\dagger\)The *-operation in $C(X)$—where $X$ is an arbitrary compact Hausdorff space—is complex conjugation; in $\mathcal{L} \mathcal{V}$ it is taking the adjoint.

\(^\dagger\)Notice that if $\text{range}(|T|)^\perp$ and $\text{range}(T)^\perp$ are isomorphic, i.e., have the same dimension, $O$ can be extended to a unitary operator $U$ on $\mathcal{H}$ and we would have $T = U |T|$.

March 28, 2009
**Theorem.** Let $T \in \mathcal{L}_V$. There is a partial isometry $O$ such that $\mathcal{D}(O) = \text{range}(|T|)$ and

$$T = O|T|, \quad \text{and} \quad |T| = O^*T.$$  

### 4.3 Spectral Theorem for self-adjoint operators.

Let $S \in \mathcal{L}_H$ be self-adjoint. For $x, y \in \mathcal{H}$, the map $F \mapsto \langle F(S)x, y \rangle$ is a bounded linear functional on $C(\sigma(S))$. Hence there exists a measure $\mu_{x,y}$ on $\sigma(S)$ such that

$$\langle F(S)x, y \rangle = \int F(\lambda) \, d\mu_{x,y}.$$  

For $x = y$ we shorten the notation $\mu_{x,x}$ to $\mu_x$.

If $f \in C(\sigma(S))$ is nonnegative, then $g = f^{\frac{1}{2}} \in C(\sigma(S))$, is real-valued (can be taken nonnegative) and $f^{\frac{1}{2}}(S)$ is self-adjoint. It follows that

$$\int f(\lambda) \, d\mu_x = \langle f(S)x, x \rangle = \|f^{\frac{1}{2}}(S)x\|^2 \geq 0,$$

so that for every $x \in \mathcal{H}$, $\mu_x$ is a positive measure.

If $y = P(S)x$ and $g \in C(\sigma(S))$ then

$$\langle g(S)y, y \rangle = \langle g(S)P(S)x, P(S)x \rangle = \langle g(S)|P|^2(S)x, x \rangle,$$

so that $\mu_y = |P|^2 \mu_x$.

Denote by $\mathcal{H}_x^0$ the linear span in $\mathcal{H}$ of $\{S^n x\}_{n=0}^{\infty}$, and by $\overline{\mathcal{H}_x}$ its closure in $\mathcal{H}$. Since $\|P(S)x\|^2 = \langle P(S)x, P(S)x \rangle = \langle |P|^2(S)x, x \rangle = \int |P|^2 \, d\mu_x$, the map $\Psi_x : P(S)x \mapsto P$ from $\mathcal{H}_x^0$ into $L^2(\mu_x)$ is an isometry.

Since the image—the set of all polynomials on $\sigma(S)$—is dense in $L^2(\mu_x)$, and $\mathcal{H}_x^0$ is dense in $\mathcal{H}_x$, the map can be extended to an isometry, still denoted $\Psi_x$, of $\mathcal{H}_x$ onto $L^2(\mu_x)$.

The isometry conjugates $S$ to the operator of “multiplication by the variable $\lambda$” on $L^2(\mu_x)$. This is the basic Spectral Theorem for self-adjoint operators.

The inverse of $\Psi_x$ assigns to every $f \in L^2(\mu_x)$ a vector in $\mathcal{H}_x$ which is denoted $f(S)x$.

The multiplication by a function $F \in L^\infty(\mu_x)$ defines a bounded linear operator on $L^2(\mu_x)$. Conjugation by $\Psi_x$ gives a linear operator on $\mathcal{H}_x$ which we denote by $F(S)$. If $F$ happens to be continuous on $\sigma(S)$ we obtain nothing new, but the extension to $L^\infty(\mu_x)$ typically gives a much richer subalgebra of $\mathcal{L}(\mathcal{H}_x)$.

March 28, 2009
4.4 If $y \in \mathcal{H}_x$, it is limit in norm of a sequence $P_n(S)x$ where $P_n$ are polynomials which converge, in the $L^2(\mu_x)$ norm, to a function $\varphi = \Psi(S)y.$ Passing to a limit in (4.3) we obtain $\mu_y = |\varphi|^2 \mu_x$. In particular $^\dagger \mu_y < \mu_x$.

4.5 Lemma. The following statements are equivalent:

a. $y \perp \mathcal{H}_x$;

b. $\mathcal{H}_x \perp \mathcal{H}_y$;

c. $\mu_{x,y} = 0$.

PROOF: Both a. and c. can be written as: $\langle F(S)x, y \rangle = 0$ for all $F \in C(\sigma(S))$. Statement b. is equivalent to: $\langle F(S)x, G(S)y \rangle = \langle G(S)F(S)x, y \rangle = 0$ for all $F, G \in C(\sigma(S))$. ▲

4.6 Lemma. Assume that $\mathcal{H}_x \perp \mathcal{H}_y$. Then $\mu_{x+y} = \mu_x + \mu_y$.

PROOF: If $f \in C(\sigma(S))$, then $f(S)x \perp f(S)y$ so that

$$\langle f(S)(x+y), (x+y) \rangle = \langle f(S)x, x \rangle + \langle f(S)y, y \rangle.$$ ▲

4.7 Lemma. Assume $\mu_x$ and $\mu_y$ mutually singular; then $\mathcal{H}_x \perp \mathcal{H}_y$ and

$$\mathcal{H}_{x+y} = \mathcal{H}_x \oplus \mathcal{H}_y.$$ (4.4)

PROOF: Let $E$ be Borel measurable such that $\mu_x(E) = \|x\|^2$ and $\mu_y(E) = 0$. Then $1_E(S)$ is the projection of $\mathcal{H}$ onto $\mathcal{H}_E$ of all vectors $z$ such that $\mu_z$ is carried by $E$, an $S$-invariant subspace. $I - 1_E(S)$ is the orthogonal projection on the orthogonal complement $\mathcal{H}_E^\perp$, equally $S$-invariant. $x \in \mathcal{H}_E$ and $y \in \mathcal{H}_E^\perp$. ▲

4.8 Theorem. Assume $\mathcal{H}$ separable, $S \in \mathcal{L}\mathcal{H}$ self-adjoint. There exists $x \in \mathcal{H}$ such that for every $y \in \mathcal{H}$, $\mu_y < \mu_x$.

PROOF: Let $\{x_n\}$ be a maximal family such that $^\dagger \mathcal{H} = \bigoplus \mathcal{H}_{x_n}$. Assume, with no loss of generality, that $\|x_n\| = 1$ for all $n$, and take $x = \sum n^{-1}x_n$. ▲

$^\dagger \nu < \mu$ means: $\nu$ is absolutely continuous with respect to $\mu$.

$^\dagger$Orthogonal direct sum.

March 28, 2009
Neither the vector $x$ nor the corresponding $\mu_x$ are unique. The equivalence class of $\mu_x$ (for mutual absolute continuity) is clearly unique, as is $L^\infty(\mu_x)$.

Let $x \in \mathcal{H}$ be such that for every $y \in \mathcal{H}$, $\mu_y \prec \mu_x$, and let $y \in \mathcal{H}$ be arbitrary. Denote $h = \frac{d\mu_y}{d\mu_x}$, the Radon-Nikodym derivative, and let $E = \{ \lambda : h(\lambda) \neq 0 \}$. Then $\tilde{x} = y + (1 - 1_E)x$ satisfies $\mu_{\tilde{x}} \sim \mu_x$, and $y \in \mathcal{H}_\tilde{x}$.

We denote by $\mu_S$ an arbitrary measure in the equivalence class, and by $L^\infty_S$ the (common) $L^\infty$ space corresponding to the class. $L^\infty_S$ is contained in $L^2(\mu_S)$ for every $x \in \mathcal{H}$, and contains the first class of Baire, $B_1(\sigma(S))$, the space of all bounded functions on $\sigma(S)$ which are pointwise limits of functions in $C(\sigma(S))$.

For $F \in L^\infty_S$ and every $x \in \mathcal{H}$, $F(S)$ is well defined linear operator on $\mathcal{H}_S$. In particular, $F(S)x$ is well defined and is is linear and bounded in $x$, so that $F(S)$ is a well defined operator in $\mathcal{L}\mathcal{H}$. As opposed to the case of continuous $F$, where $F(S)$ is defined as a limit in norm of $P_n(S)$, $P_n$ polynomials, the definition of $F(S)$ now uses the strong topology.

If $E$ is $\mu_S$ measurable, and in particular if $E$ is an interval, then $\mathbb{1}_E(S)$ is well defined. It is the orthogonal projection $\pi_{S,E}$ onto the subspace $\mathcal{H}_{S,E} = \{ x : \mu_x(E) = \| \mu_x \| \}$ (i.e., $\mu_x$ is carried by $E$). More generally, for any $x \in \mathcal{H}$

\[ \| \pi_{S,E} x \|^2 = \int \mathbb{1}_E d\mu_x = \mu_x(E). \]

An equivalent definition of $\mathcal{H}_{S,E}$ is

\[ \mathcal{H}_{S,E} = \ker(I - \pi_{S,E}) = \ker(\pi_{S,E}'), \]

where $E'$ is the complement of $E$.

Spaces of the form $\mathcal{H}_{S,E}$ are clearly $S$-invariant, and (4.6) makes it clear that they are $T$ invariant for every $T \in \mathcal{L} \mathcal{H}$ which commutes with $S$, see exercise 4.4 below. They are called $S$-spectral subspaces of $\mathcal{H}$.

4.9 If $P$ is a polynomial then $\pi_{S,E} P(S)$ is the operator which agrees with $P(S)$ on $\mathcal{H}_{S,E}$ and vanishes on its orthogonal complement, $\mathcal{H}_{S,E}^\perp$.

We have $\| \pi_{S,E} P(S) \| = \sup_{\lambda \in E} |P(\lambda)|$ and, in particular, if $E = [a,b)$ and $\lambda \in E$ then $\| \pi_{S,E} (S - \lambda) \| \leq (b - a)$.

\[ \text{March 28, 2009} \]
Proposition. Let \( \{ E_j = [a_j, a_{j+1}) \} \) be a partition of an interval carrying \( \sigma(S) \), and assume \( \max |a_{j+1} - a_j| < \varepsilon \). Then \( \| S - \sum a_j \pi_{E_j} \| < \varepsilon \).

Proof: \( \mathcal{H} = \bigoplus \mathcal{H}_{[a_j, a_{j+1})} \) is an orthogonal decomposition into \( S \)-invariant subspaces. It follows that \( \| S - \sum a_j \pi_{E_j} \| = \max \| \pi_{E_j} (S - a_j) \| < \varepsilon \).

EXERCISES FOR SECTION 4.

In the exercises below \( S \) denotes a self-adjoint operator on \( \mathcal{H} \). Prove:

4.1 The spectrum of the restriction of \( S \) to \( \mathcal{H}_x \) is the support of \( \mu_x \).

4.2 If \( \lambda \) is an isolated point in \( \sigma(S) \), then \( \lambda \) is an eigenvalue.

4.3 \( \sigma(S) \) is the closed support of \( \mu_x \).

4.4 \( S \)-spectral subspaces are invariant for any operator \( T \) which commutes with \( S \).

Hint: \( T \) commutes with \( \pi_{E_j} \).

4.5 Prove that the map \( f \mapsto f(S) \) is an isometric algebra isomorphism of \( L^\infty(\mu_S) \) into \( L^\infty(\mathcal{H}) \).

5 COMMUTING SELF-ADJOINT OPERATORS.

5.1 Let \( \mathcal{S} = \{ S_j \}_{j=1}^k \subset L^\infty(\mathcal{H}) \) be commuting self-adjoint operators. Every \( S_j \)-spectral subspace \( \mathcal{H}_j \) of \( \mathcal{H} \) is invariant under any operator that commutes with \( S_j \), and in particular under \( S_j \in \mathcal{S} \). The pairwise commutation of the \( S_j \)'s extends to the algebras they span, all the way to their strong topology closures, and in particular to the orthogonal projections \( \{ \pi_{E_j} \}_{j=1}^k \), \( E_j \) measurable \( \mu_S \).

We denote \( \pi(\lambda, \varepsilon) = \prod_{j=1}^k \pi_{S_j(\lambda_j - \varepsilon, \lambda_j + \varepsilon)} \), the orthogonal projection onto the intersection of the \( S_j \)-spectral spaces \( \mathcal{H}(\lambda, \varepsilon) = \bigcap_{j=1}^k \mathcal{H}_j(\lambda_j - \varepsilon, \lambda_j + \varepsilon) \). Notice that \( \mathcal{H}(\lambda, \varepsilon) \) is \( S_j \)-invariant for all \( j \).

Definition: The joint spectrum, \( \sigma(S_1, \ldots, S_k) \), of \( \{ S_i \} \) is the set of points \( \lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k \) having the property that for all \( \varepsilon > 0 \) there exists vectors \( x \in \mathcal{H} \) such that \( \| x \| = 1 \), and \( \| (S_l - \lambda_l)x \| \leq \varepsilon \) for all \( l = 1, \ldots, k \). \( \sigma(S_1, \ldots, S_k) \) is clearly a closed subset of \( \mathbb{R}^k \).

Lemma. \( \lambda \in \sigma(S_1, \ldots, S_k) \) if, and only if, \( \pi(\lambda, \varepsilon) \neq 0 \) for all \( \varepsilon > 0 \).

March 28, 2009
II. Linear operators

PROOF: As observed in 4.9, \( \| \pi_{\lambda_j-e_{\lambda_j+\varepsilon}} (S_l - \lambda_l) \| \leq \varepsilon \). If \( \pi(\lambda, \varepsilon) \neq 0 \), any vector of norm 1 in its range satisfies \( \| (S_l - \lambda_l) x \| \leq \varepsilon \) for all \( l = 1, \ldots, k \).

Conversely, if \( \| x \| = 1 \) and \( \| (S_l - \lambda_l) x \| \leq \delta \), then \( \int (\lambda - \lambda_l)^2 d\mu_{t, x} \leq \delta^2 \), where \( \mu_{t, x} \) denotes the spectral measure of \( x \) under \( S_l \). This implies that
\[
\mu_{t, x}((\lambda_l - \varepsilon, \lambda_l + \varepsilon)) \geq 1 - \delta^2 \varepsilon^{-2},
\]
so that \( \| x - \pi_{\lambda_j - e_{\lambda_j + \varepsilon}} x \| \leq \delta \varepsilon^{-1} \). If \( \delta \) is small enough (compared to \( \varepsilon \)) we obtain \( \pi(\lambda, \varepsilon) x \neq 0 \).

Let \( \varepsilon > 0 \) be given and let \( \{ \lambda_j \} \subset \mathbb{R}^k \) be such that the cubes \( Q_j \) of sides parallel to the axes and length \( 2\varepsilon \) centered at \( \lambda_j \) form a partition of a cube which covers \( \sigma(S_1, S_k) \). Since \( \sum \pi(\lambda_j, \varepsilon) = I \) we have, just as in 4.9, an orthogonal decomposition \( \mathcal{H} = \bigoplus \mathcal{H}_{Q_j} \) where \( \mathcal{H}_{Q_j} = \pi(\lambda_j, \varepsilon) \mathcal{H} \). The subspaces \( \mathcal{H}_{Q_j} \) are \( S_j \) invariant for all \( l \) and for each \( l \) and the appropriate \( \lambda_j \), \( \| (S_l - \lambda) \| \leq \varepsilon \). Taking into account the fact that \( \mathcal{H}_{Q_j} \) is trivial unless \( Q_j \) intersects the joint spectrum \( \sigma(S_1, S_k) \), and that \( \varepsilon \) can be taken arbitrarily small, we obtain

**Proposition.** For any polynomial \( P \) in \( k \) variables with complex coefficients,

\[
\| P(S_1, \ldots, S_k) \|_{\mathcal{L} \mathcal{H}} = \| P \|_{C(\sigma(S_1, \ldots, S_k))}.
\]

The picture now evolves just as in the case of a single self-adjoint operator and we obtain

**Theorem.** Let \( \mathcal{S} = \{ S_j \}_{j=1}^k \subset \mathcal{L} \mathcal{H} \) be commuting self-adjoint operators.

a. For any \( F \in C(\sigma(S_1, \ldots, S_k)) \) define \( F(S_1, \ldots, S_k) \) as the limit (in \( \mathcal{L} \mathcal{H} \) norm) of \( P_n(S_1, \ldots, S_k) \), \( \{ P_n \} \) an arbitrary sequence of polynomials which converges to \( F \) uniformly on \( \sigma(S_1, S_k) \). Then the map
\[
\Psi : F \mapsto F(S_1, \ldots, S_k)
\]
is an algebra isometric isomorphism of \( C(\sigma(S_1, \ldots, S_k)) \) into \( \mathcal{L} \mathcal{H} \).

b. For every \( x \in \mathcal{H} \), the mapping \( F \mapsto \langle F(S_1, \ldots, S_k) x, x \rangle \) is a linear functional of norm \( \| x \|^2 \) on \( C(\sigma(S_1, \ldots, S_k)) \), positive for all \( F \) of the form \( |g|^2 \), so that there exists a positive measure \( \mu_{\mathcal{S}, x} \) such that

\[
\int F d\mu_{\mathcal{S}, x} = \langle F(S_1, \ldots, S_k) x, x \rangle.
\]

March 28, 2009
c. There exist vectors $x \in \mathcal{H}$ such that $\mu_{\mathcal{F}, y} \prec \mu_{\mathcal{F}, x}$ for all $y \in \mathcal{H}$. The equivalence class of such $\mu_{x,}$ along with the multiplicity function, see ??, is called the spectral type of $\{S_1, \ldots, S_k\}$.

EXERCISES FOR SECTION 5.

5.1 If $T \in \mathcal{L}(\mathcal{H})$ is normal, the joint spectrum of $\Re T$ and $\Im T$ is $\{(x, y) : x + iy \in \sigma(T)\}$.

6 THE SPECTRAL THEOREM FOR UNITARY OPERATORS.

**Theorem.** Let $\mathcal{H}$ be a separable Hilbert space, and let $U$ be a unitary operator on $\mathcal{H}$. Then there exists a sequence of mutually singular measures $\mu_n \in M(\mathbb{T})$ such that $U$ is isometrically conjugate to the operator of multiplication by $e^{it}$ on the space $\bigoplus L^2(\mu_n, \mathbb{R}^n)$. The sequence \{equivalence class of $\mu_n$\} is a complete invariant.

The proof will consist of a sequence of simple observations.

6.1 For $f \in \mathcal{H}$, we denote by $\mathcal{H}_f$ the closed subspace of $\mathcal{H}$ spanned by $\{U^n f\}_{n \in \mathbb{Z}}$. Since the sequence $\{(f, U^n f)\}$ is positive definite, there exists, by Herglotz' theorem, (see [1] page 41) a positive measure $\mu_f \in M(\mathbb{T})$ such that $\hat{\mu}_f(n) = \langle f, U^n f \rangle$. $\mu_f$ is called the spectral measure of $f$.

\[
\|\mu_f\|_{M(\mathbb{T})} = \hat{\mu}_f(0) = \langle f, f \rangle = \|f\|^2.
\]

(6.1)

6.2 The map (for finite sums) $\sum a_n U^n f \mapsto \sum a_n e^{int}$ is an isometry into $L^2(\mu_f)$ and has a unique extension by continuity to an isometry $\mathcal{H}_f \hookrightarrow L^2(\mu_f)$. This isometry conjugates $U|_{\mathcal{H}_f}$ with the operator of multiplication by $e^{it}$ on $L^2(\mu_f)$. In particular we have

\[
\|\sum a_n U^n\|_{\mathcal{L}(\mathcal{H}_f)} = \|\sum a_n e^{int}\|_{L^2(\mu_f)}.
\]

(6.2)

6.3 If $g \in \mathcal{H}_f$ and $\mu_g \sim \mu_f$, (each absolutely continuous with respect to the other,) then $\mathcal{H}_g = \mathcal{H}_f$. Conversely, if $\mathcal{H}_g = \mathcal{H}_f$ then $\mu_g \sim \mu_f$.

MARCH 28, 2009
PROOF: There is no loss of generality in assuming that $\mathcal{H} = L^2(\mu_f)$, $f = 1$ and $U$ the multiplication by $e^{it}$ (so that $F(U)$ is the multiplication by $F$). For $g \in L^2(\mu_f)$ we have
\[ \hat{\mu}_g(n) = \langle g, e^{int} \rangle = \int g\bar{g}e^{-int}d\mu_f \]
which means that $\mu_g = |g|^2\mu_f$. The assumption that the two measures are equivalent means $g \neq 0$ a.e. and we need to show that the functions of the form $Fg$ where $F$ is a continuous function (or a trigonometric polynomial) are dense in $L^2(\mu_f)$. This can be done directly or via duality as follows: If $\psi$ is orthogonal to $\{Fg : F \in C(\mathbb{T})\}$ then $\int Fg\bar{\psi}d\mu_f = 0$ for all $F$ so that $g\bar{\psi} = 0$ a.e. and since $g \neq 0$, $\psi = 0$.

6.4 There exist functions $f \in \mathcal{H}$ such that $\mu_g \prec \mu_f$ for every $g \in \mathcal{H}$. The (closed) support of such $\mu_f$ is equal to $\sigma(U)$, the spectrum of $U$.

PROOF: There exist a (finite or countable) sequence $\{f_n\}$ such that $\mathcal{H} = \bigoplus \mathcal{H}_{f_n}$. In fact, let $\{u_n\}$ be a dense subset of $\mathcal{H}$, take $f_1 = u_1$. Take $f_2$ in the orthogonal complement of $\mathcal{H}_{f_1}$ such that $u_2 \in \mathcal{H}_{f_1} \oplus \mathcal{H}_{f_2}$, etc.

If $a_n > 0$ are such that $\sum a_n^2\|f_n\|^2 < \infty$ and if we put $f = \sum a_n f_n$, then $\mu_f = \sum a_n^2 \mu_{f_n}$, and $\mu_g \prec \mu_f$ for every $g \in \mathcal{H}$.

The spectrum of the operator of multiplication by $e^{it}$ on $L^2(\mathbb{T}, B, \mu)$ is precisely the closed support of $\mu$.

6.5 Identity (6.2) now reads:

\begin{equation}
\left\| \sum a_n U^n \right\|_{L^p(\mathcal{H})} = \sup_{t \in \sigma(U)} \left| \sum a_n e^{int} \right|.
\end{equation}

6.6 For $f, g \in \mathcal{H}$, the expression $\langle P(U)f, g \rangle$ is a bounded linear functional on $C(\sigma(U))$ hence there exists a (signed) measure $\mu_{f,g}$ such that
\[ \langle P(U)f, g \rangle = \int P(e^{it})d\mu_{f,g}. \]
If \( g = g_1 + g_2 \) with \( g_1 \in \mathcal{H}_f \) and \( g_2 \in \mathcal{H}_f^\perp \), and we denote by \( \varphi_1 \) the image of \( g_1 \) in \( L^2(\mu_f) \), then
\[
\langle P(U)f, g \rangle = \langle P(U)f, g_1 \rangle = \int P(e^{it})\varphi_1 d\mu_f.
\]
Thus \( \mu_{f,g} = \overline{\varphi_1}\mu_f \) and \( \mu_{f,g} \) is absolutely continuous with respect to \( \mu_f \). Reversing the roles of \( f \) and \( g \) in the argument shows that \( \mu_{f,g} \) is absolutely continuous with respect to \( \mu_g \). The argument also shows that \( \mu_{f,g} = 0 \) if, and only if, \( \mathcal{H}_f \perp \mathcal{H}_g \).

6.7 There exist a (possibly finite) sequence \( \{f_n\} \subset \mathcal{H} \) such that \( \mathcal{H}_{f_n} \) are mutually orthogonal, \( \mu_{f_{n+1}} \prec \mu_{f_n} \), and \( \mathcal{H} = \bigoplus \mathcal{H}_{f_n} \).

**PROOF:** Take an \( f \) as described in 6.4 and call it \( f_1 \). If \( \mathcal{H} = \mathcal{H}_{f_1} \) we are done. Otherwise, let \( \mathcal{H}^1 \) be the orthogonal complement of \( \mathcal{H}_{f_1} \) in \( \mathcal{H} \). \( \mathcal{H}^1 \) is \( U \)-invariant and we can take for \( f_2 \) any vector in \( \mathcal{H}^1 \) that plays the role that \( f_1 \) played in \( \mathcal{H} \). If \( \mathcal{H}_{f_2} = \mathcal{H}^1 \) we are done, otherwise repeat the step in the orthogonal complement of \( \mathcal{H}_{f_2} \) in \( \mathcal{H}^1 \), etc. As in the previous part, use a dense subsequence to guarantee that the union of \( \mathcal{H}_{f_n} \) spans \( \mathcal{H} \). \( \blacklozenge \)

6.8 The condition \( \mu_{f_{n+1}} \prec \mu_{f_n} \) in 6.7 can be replaced by \( \mu_{f_{n+1}} = \mathbb{1}_{E_{n+1}} \mu_{f_n} \). In other words one can guarantee that \( \mu_{f_{n+1}} \) is just the restriction of \( \mu_n \) to a measurable set \( E_{n+1} \). This follows from part 6.3. Define the **spectral type** of \( U \) to be the equivalence class of \( \mu = \mu_{f_1} \), and the **multiplicity function** of \( U \) to be \( \sum_1^\infty \mathbb{1}_{E_n} \).

The sets \( E_n \) as defined in 6.8 above, and \( E_1 = \sigma(U) \) (all defined modulo sets of \( \mu_{f_1} \) measure 0), and the sum may take the value \(+\infty\) on a set of positive measure.

**Remark:** One should avoid the temptation to say that since \( E_2 \) comes from the spectrum of \( U \) restricted to \( \mathcal{H}_1 \) one may take it closed. The restriction of \( \mu_{f_1} \) to \( \sigma(U|\mathcal{H}_1) \) may not be \( < \mu_{f_2} \).

We need to show that the sets \( E_j \), and the multiplicity function are canonically determined, and prove that together they are a complete invariant for \( U \): if \( U \) is unitary on \( \mathcal{H} \) and \( U_1 \) is unitary on \( \mathcal{H}_1 \) and they have the same spectral type and multiplicity function, then there exists an isometry of \( \mathcal{H} \) onto \( \mathcal{H}_1 \) which conjugates \( U \) to \( U_1 \).

March 28, 2009