MATH 247/Winter 2010

Notes on the adjoint and on normal operators.

In these notes, V is a finite dimensional inner product space over $\mathbb{C}$, with given inner product $\langle u,v\rangle$. T, S, T^*, … are linear operators on V. U, W are subspaces of V. When we say "subspace", we mean one of the fixed space V. $a,b,...,\mu,\lambda,...$ denote complex numbers. u,v,w denote elements of V.

U is T*-invariant*, or U is *invariant under* T, if for all $u \in U$, we have $T(u) \in U$. If U is T-invariant, we have the operator $T \restriction U$ on U, called the *restriction* of T to U, which is defined by $(T \restriction U)(u) = T(u)$. Indeed, $T \restriction U : U \to U$, since $T(u) \in U$ whenever $u \in U$, and the linearity of $T \restriction U$ follows from the linearity of T.

Note that the total space V and the trivial subspace $\{0\}$ are always T-invariant.

A particular type of T-invariant subspace is an *eigenspace* of T. For any $\mu \in \mathbb{C}$, U = $E_\mu[T] \overset{def}{=} \mathrm{Ker}(T-\mu I) = \{u \in V : T(u) = \mu u\}$ is a T-invariant subspace: if $u \in U$, then $T(u) \in U$, since $T(T(u)) = T(\mu u) = \mu T(u)$. $E_\mu[T] \neq \{0\}$ precisely when μ is an *eigenvalue* of T; the non-zero elements of $E_\mu[T]$, if any (when μ is an of T), are the *eigenvectors* of T *for* the eigenvalue μ. $E_\mu[T]$ is called an *eigenspace* of T if $E_\mu[T] \neq \{0\}$.

We say that T and S *commute* if $T \cdot S = S \cdot T$.

A basis $\mathcal{P} = (P_1,...,P_n)$ of V is a *diagonalizing basis* for T (it could also be called an "eigenbasis" for T) if the matrix $[T]_{\mathcal{P}}$ is diagonal; equivalently, if each basis element P_i is an eigenvector of T.

Theorem 1 For every T, there is a unique T^* such that $\langle T(u),v\rangle = \langle u, T^*(v)\rangle$ identically for all $u,v \in V$.

The proof of this theorem is omitted here. It can be found in our textbook: Theorem 13.1; proved in Problem 13.4.

(The reason why Theorem 1. true is simple. To determine T^*, we have to determine each value $T^*(v)$. Fix v, and seek $w = T^*(v)$. It has to satisfy $\langle T(u),v\rangle = \langle u,w\rangle$ for all u. The left-hand-side, $\langle T(u), v\rangle$, is a function of u; in fact, it is a linear function $F: V \to \mathbb{C}$, a *linear functional*, meaning just that the codomain in the field $\mathbb{C}$ itself (the vector space of dimension =1). Now, it turns out, that just the linearity of F ensures the unique existence of w such that $F(u) = \langle u, w\rangle$ for all u: all linear functionals can be represented in the form $u \mapsto \langle u, w\rangle$. If you follow this, and so determine $w = T^*(v)$, the second and final step is to show that $T^*: V \to V$ is linear, which is easy.)

A very easy but very important consequence of the definition of the adjoint is that the adjoint of the adjoit is the original: $(T^*)^* = T^{**} = T$.

Note also the following simple consequence of Theorem 1. If U is invariant under both T and T^*, then the adjoint $(T\restriction U)^*$ of the restricted operator $T\restriction U$ equals $T^*\restriction U$. Indeed, first of all, $T^*\restriction U$ is a well-defined operator on the space U by the assumption that U is T^*-invariant. The equality $\langle (T\restriction U)(u), v\rangle = \langle u, (T^*\restriction U)(v)\rangle$ for $u, v \in U$ follows from the equality $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$, since $(T\restriction U)(u) = T(u)$ and $(T^*\restriction U)(v) = T^*(v)$. Thus, $T^*\restriction U$ satisfies all the characteristic properties of the adjoint of $T\restriction U$; since, by the theorem, there is only one adjoint to $T\restriction U$, $T^*\restriction U$ must be *the* adjoint of $T\restriction U$.

Lemma 2 Suppose T and S commute. Then any eigenspace of T is $S-$invariant.

Proof Let $U = E_\mu[T]$. Let $u \in U$; we want to show that $S(u) \in U$ (?). $u \in U$ means that $T(u) = \mu u$. Therefore, $(ST)(u) = S(T(u)) = S(\mu u) = \mu S(u)$. But also $(TS)(u) = (ST)(u)$. Thus, $T(S(u)) = (TS)(u) = (ST)(u) = \mu S(u)$, which means that $S(u) \in U$ as desired.

Lemma 3 Suppose that U is T-invariant. Then $U^\perp$ is T^*-invariant.

Proof Let $w \in U^\perp$, to show that $T^*(w) \in U^\perp$, or in other words, that $\langle u, T^*(w)\rangle = 0$ for all $u \in U$. Let $u \in U$. We have $\langle u, T^*(w)\rangle = \langle T(u), w\rangle$ by the definition of the adjoint T^*. Since U is T-invariant, we have $T(u) \in U$, and since $w \in U^\perp$, we have $\langle T(u), w\rangle = 0$. Therefore, $\langle u, T^*(w)\rangle = 0$ as desired.

Proposition 4 Suppose $\mathcal{S}$ is a set of linear operators on V such that, for any $S, T \in \mathcal{S}$, we have that S and T commute, as well as S^* and T commute. Then there is an orthonormal basis of V which is a diagonalizing basis for *every* S in $\mathcal{S}$: a single common orthonormal diagonalizing basis for all operators in $\mathcal{S}$ at once.

Proof Note that the assumption implies that each $S \in \mathcal{S}$ is normal: take $S = T$ in the assumption.

The proof is by induction on the dimension n of V. When $n = 1$, then the assertion is trivial: any basis (P_1) is a diagonalizing basis for any operator on V.

Suppose $n > 1$, and suppose that the assertion is true for all inner product spaces V' of dimension less than n. Let V be an inner product space of dimension equal to n.

There are two cases. In Case 1, we assume that every $T \in \mathcal{S}$ is a scalar multiple of the identity operator I on V: $T = \mu \cdot I$ for some $\mu \in \mathbb{C}$; equivalently, $E_\mu[T] = V$. Then, again, the assertion is trivial, for the same reason as before: every basis of V is a diagonalizing basis for all $T \in \mathcal{S}$.

Case 2 is when Case 1 does not hold. Then: there is $T \in \mathcal{S}$ which is *not* of the form $T = \mu \cdot I$. *Let* $\mu \in \mathbb{C}$ *be an eigenvalue of* T. There is such μ by the Fundamental Theorem of Algebra: the polynomial $\mathrm{char}_T(\lambda)$ has at least one root. (This is the point in this proof where we use complex scalars in an essential way.) We have assumed that $E_\mu[T] \neq V$. Let $U = E_\mu[T]$, and $W = U^\perp$. Then $U \oplus W = V$, and $U \neq \{0\}$, and $W \neq \{0\}$; the first inequality because μ is an eigenvalue, the second because $U \neq V$. Therefore, $0 < \dim(U) < n$, and $0 < \dim(W) < n$.

I claim that both U and W are invariant under both S and S^*, for any $S \in \mathcal{S}$. Indeed, since U is an eigenspaces of (the chosen) T, and both S and S^* commute with T, by Lemma 2, we have that U is S-invariant and S^*-invariant. Next, by Lemma 3, $W = U^\perp$ is S^*-invariant, since U is S-invariant; and $W = U^\perp$ is $S = (S^*)^*$-invariant, since U is S^*-invariant.

We therefore have the situation on both of the subspaces $V' = U$ and $V' = W$ that we have a set of operators $\mathcal{S}' = \{S \restriction V' : S \in \mathcal{S}\}$ with the property that for any $S' = S \restriction V'$ (with $S \in \mathcal{S}$) and $T' = T \restriction V'$ (with $T \in \mathcal{S}$), both in $\mathcal{S}'$, $S' = S \restriction V'$ and $T' = T \restriction V'$ commute, as well as $(S')^* = S^* \restriction V'$ and $T' = T \restriction V'$ commute, directly following from the facts that S and T commute as well as S^* and T commute.

Since, for both $V' = U$ and $V' = W$, we have that $\dim(V') < n$, we can apply the induction hypothesis, to conclude that there is a single common orthogonal diagonalizing basis $\mathcal{P}_1$ for all operators in the set $\mathcal{S}_1 = \{S \restriction U : S \in \mathcal{S}\}$, and another one, $\mathcal{P}_2$, for $\mathcal{S}_2 = \{S \restriction W : S \in \mathcal{S}\}$.

Let $k = \dim(U)$, and $l = \dim(W)$; let $\mathcal{P}_1 = (P_1, ..., P_k)$ and $\mathcal{P}_2 = (P_{k+1}, ..., P_{k+l=n})$. Since $U \oplus W = V$, we have that $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2 = (P_1, ..., P_n)$ is a basis of V. It is an orthogonal basis: any two of the first k elements of $\mathcal{P}$ are orthogonal since $\mathcal{P}_1$ is an orthogonal system; any two of the last l elements of $\mathcal{P}$ are orthogonal since $\mathcal{P}_2$ is an orthogonal system; and any element among the first k and any one among the last l are orthogonal since the first is in U, the second is in W, and $U \perp W$.

Let $S \in \mathcal{S}$. Every basis vector P_i is an eigenvector of S: if $i = 1, ..., k$, then P_i is an eigenvector of $S \upharpoonright U$, $(S \upharpoonright U)(P_i) = \lambda_i P_i$ for some λ_i, and thus $S(P_i) = \lambda_i P_i$, and similarly for $i = k+1, ..., k+l = n$. Incidentally, the eigenvalues of S are thus seen to be $\lambda_1, ..., \lambda_k, \lambda_{k+1}, ..., \lambda_{k+l=n}$, where $\lambda_1, ..., \lambda_k$ are the eigenvalues of $S \upharpoonright U$, and $\lambda_{k+1}, ..., \lambda_{k+l=n}$ are the eigenvalues of $S \upharpoonright W$.

This completes the proof of the Proposition.

Proposition 4 already contains the

Theorem 5 (***Spectral Theorem for finite dimensional Hilbert spaces***) Every normal operator (on a finite dimensional Hilbert space) has an orthonormal diagonalizing basis.

Proof Apply Proposition 4 to the set $\mathcal{S} = \{T\}$. The hypothesis of Proposition 4 holds since T commutes with T (obviously), and T^* commutes with T (by the normality of T).

However, we can also use Proposition 4 to prove a stronger result. First, another proposition, one that is interesting in itself -- whose proof uses Theorem 5. (This is interesting because the statement has nothing to do with diagonalization.)

Proposition 6 Suppose S is a normal operator, T is *any* operator (on V, of course). If S and T commute, then S^* and T commute as well.

Proof Let, by Theorem 5, $\mathcal{P}$ be an orthonormal diagonalizing basis for S. Therefore, for $D = [S]_{\mathcal{P}}$, $A = [T]_{\mathcal{P}}$, we have that $D = (d_{ij})^{n \times n}$ is diagonal, $d_{ij} = 0$ whenever $i \neq j$, and for $A = (a_{ij})^{n \times n}$, we have $D \cdot A = A \cdot D$. This means that, for any i, j from 1 to n, we have $\sum_{k=1}^{n} d_{ik} \cdot a_{kj} = \sum_{k=1}^{n} a_{ik} \cdot d_{kj}$, which, by $d_{ij} = 0$ whenever $i \neq j$, reduces to

$d_{ii} \cdot a_{ij} = a_{ij} \cdot d_{jj}$. In other words, $a_{ij}(d_{ii} - d_{jj}) = 0$. Therefore, either $a_{ij} = 0$ (Case 1), or $d_{ii} - d_{jj} = 0$ (Case 2).

Since $\mathcal{P}$ is orthonormal, $[S^*]_{\mathcal{P}} = \overline{D}^{tr} = (\overline{d}_{ji})^{n\times n}$. I claim that $\overline{D}^{tr} \cdot A = A \cdot \overline{D}^{tr}$. Indeed, since $\overline{D}^{tr}$ is diagonal, this reduces, just as before, to the question whether we can see that $a_{ij}(\overline{d}_{ii} - \overline{d}_{jj}) = 0$. If Case1 holds, this is true. But if Case 2 holds, then the complex conjugate of $d_{ii} - d_{jj}$ being $\overline{d}_{ii} - \overline{d}_{jj}$, $d_{ii} - d_{jj}$ equals zero implies that $\overline{d}_{ii} - \overline{d}_{jj}$ equals zero, hence again $a_{ij}(\overline{d}_{ii} - \overline{d}_{jj}) = 0$.

$\overline{D}^{tr} \cdot A = A \cdot \overline{D}^{tr}$, $A = [T]_{\mathcal{P}}$ and $[S^*]_{\mathcal{P}} = \overline{D}^{tr}$ together imply that $S^* \cdot T = T \cdot S^*$ as desired.

Theorem 7 ***(Generalized Spectral Theorem for finite dimensional Hilbert spaces*****)**

1) Suppose that $\mathcal{S}$ is a set of *commuting normal* linear operators on V : every S in $\mathcal{S}$ is normal, and, for any $S, T \in \mathcal{S}$, S and T commute. Then there is an orthonormal basis of V which is a diagonalizing basis for *every* S in $\mathcal{S}$: there is a single common orthonormal diagonalizing basis for all operators in $\mathcal{S}$ at once.

2) In particular, if S and T are commuting normal operators, then there is an orthonormal basis of V which is a diagonalizing basis for both S and T .

Proof This follows from Proposition 4 and Proposition 6. Indeed, the conditions of Proposition 4 are fulfilled. Let S and T be both from the set $\mathcal{S}$. Then S and T commute, directly by the assumption of the theorem. But also, S^* and T commute, since, by assumption, S is normal, and thus, by Proposition 6, the fact that S and T commute implies that S^* and T commute.

Proposition 8 **1)** $(S \cdot T)^* = T^* \cdot S^*$

2) $(S + T)^* = S^* + T^*$

3) $(a \cdot T)^* = \overline{a} \cdot T^*$

4) If S and T are normal and commute with each other, then

4.1) $S \cdot T$ is normal, and

4.2) $S + T$ is normal.

5) If T is normal, so is $a \cdot T$.

6) Let $f(x_1, x_2, ...)$ be a polynomial with complex coefficients in any number of variables $x_1, x_2, ...$. Assume that $T_1, T_2, ...$ are commuting normal operators

($T_i \cdot T_j = T_j \cdot T_i$ for all $i, j = 1, 2, \ldots$). Then $T = f(T_1, T_2, \ldots)$ is a normal operator. Moreover, if $\wp$ is a common orthonormal diagonalizing basis for $T_1, T_2, \ldots$, then $\wp$ also diagonalizes T; if $\wp = (P_1, \ldots, P_n)$ and $T_k(P_i) = \lambda_i^{(k)} \cdot P_i$, then $T(P_i) = \lambda_i \cdot P_i$ where $\lambda_i = f(\lambda_i^{(1)}, \lambda_i^{(2)}, \ldots)$.

Proof 1): We have the identity

$$\langle (ST)(u), v \rangle = \langle S(T(u)), v \rangle = \langle T(u), S^*(v) \rangle = \langle u, T^*(S^*(v)) \rangle = \langle u, (T^*S^*)(v) \rangle .$$

This implies the assertion.

2): $\langle (S+T)(u), v \rangle = \langle S(u) + T(u), v \rangle = \langle S(u), v \rangle + \langle T(u), v \rangle =$

$$= \langle u, S^*(v) \rangle + \langle u, T^*(v) \rangle = \langle u, S^*(v) + T^*(v) \rangle = \langle u, (S^* + T^*)(v) \rangle .$$

Again, this implies the assertion.

3): $\langle (a \cdot T)(u), v \rangle = \langle a \cdot (T(u)), v \rangle = a \cdot \langle T(u), v \rangle = a \cdot \langle u, T^*(v) \rangle = \langle u, \overline{a} \cdot T^*(v) \rangle$, which is sufficient.

4.2): We need to show that $(ST) \cdot ((ST)^*) = ((ST)^*) \cdot (ST)$. We have, using 1):

$$(ST) \cdot ((ST)^*) = (ST) \cdot (T^* \cdot S^*) = STT^*S^*$$

and

$$((ST)^*) \cdot (ST) = (T^*S^*) \cdot (ST) = T^*S^*ST .$$

But since every one of S, S^*, T, T^* commutes with every other, *as a consequence of Proposition 6*, the two products are equal.

4.3): Using 2) we have

$$(S+T) \cdot (S+T)^* = (S+T) \cdot (S^* + T^*) = S \cdot S^* + S \cdot T^* + T \cdot S^* + T \cdot T^*,$$

and similarly,

$$(S+T)^* \cdot (S+T) = (S^* + T^*) \cdot (S+T) = S^* \cdot S + S^* \cdot T + T^* \cdot S + T^* \cdot T .$$

For the same reason as in 4.2), these values are the same.

5): $(aT) \cdot ((aT)^*) = (aT) \cdot (\overline{a}(T^*)) = a \cdot \overline{a} \cdot T \cdot T^*$

and

$$((aT)^*)\cdot(aT)=(\bar{a}(T^*))\cdot(aT)=\bar{a}\cdot a\cdot T^*\cdot T$$

Since T is normal, these are equal.

6): This follows from Theorem 7, by applying 4) and 5) repeatedly, to build up the polynomial $f(x_1,x_2,...)$.