21-238, Math Studies Algebra 2, Department of Mathematical Sciences, Carnegie Mellon University
Spring 2012: Monday, Wednesday, Friday, 10:30 am, Doherty Hall 1211.
Luc TARTAR, University Professor of Mathematics, Wean Hall 6212, tartar@cmu.edu

9- Friday February 3, 2012.

Definition 9.1: If V, W are Hermitian spaces, and A is a linear mapping from V into W, then the *adjoint* A^* of A is defined by $(A\,v, w) = (v, A^*w)$ for all $v \in V, w \in W$.[1]

If $W = V$, then A is said to be *normal* if A^* commutes with A, A is said to be *self-adjoint* or *Hermitian* if $A^* = A$, A is said to be *skew Hermitian* if $A^* = -A$, A is said to be *unitary* if $A^*A = A\,A^* = I$.[2]

Remark 9.2: If V, W, X are Hermitian spaces, one has $(\lambda\,I)^* = \overline{\lambda}\,I$ for all $\lambda \in \mathbb{C}$; one has $(A^*)^* = A$ for all $A \in L(V, W)$; one has $(B\,A)^* = A^*B^*$ for all $A \in L(V, W), B \in L(W, X)$; if $A \in L(V, W)$ is invertible, then A^* is invertible and $(A^*)^{-1} = (A^{-1})^*$.

If $e_i, i \in I$, is an orthonormal basis of V, and $f_j, j \in J$, is an orthonormal basis of W, then $A^*_{i,j} = (A^*f_j, e_i) = (f_j, A\,e_i) = \overline{(A\,e_i, f_j)} = \overline{A_{j,i}}$ for all $i \in I, j \in J$; in particular, if $A \in L(V, V)$ is diagonal on an orthonormal basis, then A is normal (since both A and A^* are diagonal on such a basis).

If M is an Hermitian operator from V into itself, it can be recovered from the function $v \mapsto (M\,v, v)$: for $v, w \in V$, one has $(M(v \pm w), v \pm w) = (M\,v, v) + (M\,w, w) \pm 2\Re(M\,w, v)$, so that $(M(v \pm i\,w), v \pm i\,w) = (M\,v, v) + (M\,w, w) \mp 2\Im(M\,w, v)$.

Remark 9.3: If $V_\mathbb{R}$ is an Euclidean space (so that the field of scalars is $\mathbb{R}$), it is useful to embed it into a Hermitian space $V_\mathbb{C}$ by extending the field of scalars from $\mathbb{R}$ to $\mathbb{C}$, and since $[\mathbb{C}:\mathbb{R}] = 2$, $V_\mathbb{C}$ is an $\mathbb{R}$-vector space isomorphic to $V_\mathbb{R} \times V_\mathbb{R}$, and the multiplication by $\lambda + i\,\mu$ (with $\lambda, \mu \in \mathbb{R}$) is $(\lambda + i\,\mu)\,(u, v) = \lambda\,u - \mu\,v, \mu\,u + \lambda\,v)$, so that $(u, v) \in V_\mathbb{R} \times V_\mathbb{R}$ is thought of as $u + i\,v \in V_\mathbb{C}$.

If A is a symmetric operator on $V_\mathbb{R}$, there is no need to extend the scalars to $\mathbb{C}$ since there exists an orthonormal ($\mathbb{R}$-) basis of $V_\mathbb{R}$ on which A is diagonal,[3] but if A is either skew symmetric or orthogonal,[4] it may have complex eigenvalues, and it is useful to consider the Hermitian space $V_\mathbb{C}$, in order to have A either skew Hermitian or unitary, and apply the general result for normal operators in a Hermitian space below; then, it has implication on what kind of block-diagonal structure one may find for A on an adapted orthonormal ($\mathbb{R}$-) basis of $V_\mathbb{R}$.

Lemma 9.4: Let V be a Hermitian space, and $A \in L(V, V)$. A is normal if and only if $||A\,v|| = ||A^*v||$ for all $v \in V$. If A is normal and $A\,e = \lambda\,e$, then $A^*e = \overline{\lambda}\,e$, and both A and A^* map $e^\perp$ into itself. If A is normal and V is finite dimensional, there exists an orthonormal basis of eigenvectors (of both A and A^*).[5]

---

[1] Of course, the scalar product in W is used in $(A\,v, w)$, and the scalar product in V is used in (v, A^*w).

[2] If V is finite dimensional, $A^*A = I$ is equivalent to $A\,A^* = I$, since $A^*A = I$ implies that A is injective and A^* is surjective, hence they are invertible. It is not so if V is infinite dimensional: for example, if $V = \ell^2$ and A is the right shift (where $A\,x = y$ means $y_1 = 0$ and $y_n = x_{n-1}$ for $n \geq 2$), then A^* is the left shift (where $A^*x = z$ means $x_n = x_{n+1}$ for all $n \geq 1$); A is an isometry (i.e. $||Av|| = ||v||$ for all $v \in V$) which is not surjective (its image is $e_1^\perp$) and $A^*A = I$; A^* is a contraction (i.e. $||A^*v|| \leq ||v||$ for all $v \in V$) but since $A\,e_1 = 0$ it is not injective (hence not an isometry), and it is surjective (since $A\,A^* = I$), and because $A\,A^*$ is the orthogonal projection on $e_1^\perp$, one has $A\,A^* \neq I$.

[3] Because a proof was already given for the existence of an orthonormal ($\mathbb{R}$-) basis of $V_\mathbb{R}$ made of eigenvectors of A, but Lemma 9.4 actually provides a different proof of the existence of a real eigenvalue for a symmetric operator: since $A\,e = \lambda\,e$ implies $A^*e = \overline{\lambda}\,e$, and $A^* = A$ and $e \neq 0$ imply $\overline{\lambda} = \lambda$, i.e. $\lambda \in \mathbb{R}$.

[4] If the dimension of $V_\mathbb{R}$ is even, an operator may be both skew symmetric and orthogonal, in which case there exists an orthonormal ($\mathbb{R}$-) basis of $V_\mathbb{R}$ on which A is block-diagonal, with blocks of size 2 being either a rotation of $+\frac{\pi}{2}$ or a rotation of $-\frac{\pi}{2}$.

[5] It is not true for an infinite dimensional Hermitian space. For continuous functions on $[0, 1]$ with the scalar product $(f, g) = \int_0^1 f(x)\,\overline{g(x)}\,dx$, although the operator A of multiplication by x is Hermitian, it has no eigenvalues, i.e. $A - \lambda\,I$ is injective for all $\lambda \in \mathbb{C}$, but $A - \lambda\,I$ is not surjective for λ real $\in [0, 1]$. Assuming that one has a Hilbert space (i.e. the space is complete for the norm), it is true that if besides being normal A is also compact (i.e. it sends the closed unit ball inside a compact set) then there is a Hilbert basis $e_i, i \in I$, (i.e. $(e_i, e_j) = \delta_{i,j}$ for $i, j \in I$ and finite linear combinations are a dense set) of eigenvectors of A.

Proof: Since both A^*A and $A\,A^*$ are Hermitian, $A^*A = A\,A^*$ is equivalent to $(A^*A\,v, v) = (A\,A^*v, v)$ for all $v \in V$, i.e. $||A\,v||^2 = ||A^*v||^2$ for all $v \in V$. In particular, if A is normal, $A\,e = 0$ is equivalent to $A^*e = 0$; that $(A - \lambda\,I)\,e = 0$ is equivalent to $(A^* - \overline{\lambda}\,I)\,e = 0$ follows then from the fact that $(\lambda\,I)^* = \overline{\lambda}\,I$ and that $A - \lambda\,I$ and $A^* - \overline{\lambda}\,I$ commute, so that $A - \lambda\,I$ is normal. If $f \in e^\perp$, i.e. $(f, e) = 0$, one has $(A\,f, e) = (f, A^*e) = (f, \overline{\lambda}\,e) = \lambda\,(f, e) = 0$, so that $A\,f \in e^\perp$, and similarly, $(A^*f, e) = (f, A\,e) = (f, \lambda\,e) = \overline{\lambda}\,(f, e) = 0$, so that $A^*f \in e^\perp$. If V has dimension n, then the characteristic polynomial $P_{char}(\lambda) = det(A - \lambda\,I)$ has degree n and has at least one root (since $\mathbb{C}$ is algebraically closed), hence there is an eigenvector e for an eigenvalue λ, and then the problem is to repeat the operation of $e^\perp$, which has dimension $n-1$, and one concludes by induction on n (the case $n = 1$ being trivial).

Remark 9.5: If A is skew symmetric on an Euclidean space $V_\mathbb{R}$, then one extends the scalars to be complex, which creates a Hermitian space $V_\mathbb{C}$, where the natural extension of A is skew Hermitian, so that each eigenvalue λ satisfies $\overline{\lambda} = -\lambda$, i.e. besides 0 the eigenvalues are purely imaginary. One starts by working on $\big(ker(A)\big)^\perp$ in $V_\mathbb{R}$, i.e. one is led to the case where 0 is not an eigenvalue; an eigenvalue $\lambda \in \mathbb{C}$ of A in $V_\mathbb{C}$ is then $\lambda = i\,\mu$ for a non-zero $\mu \in \mathbb{R}$, and one writes an eigenvector as $v + i\,w$ with $v, w \in V_\mathbb{R}$, so that $A\,(v + i\,w) = i\,\mu\,(v + i\,w)$ means $A\,v = -\mu\,w$ and $A\,w = \mu\,v$; one deduces that neither v nor w is 0, and since $(A\,v, v) = (A\,w, w) = 0$ one has $(v, w) = 0$, and by normalizing v and w in $V_\mathbb{R}$, it corresponds to a diagonal block $\begin{pmatrix} 0 & \mu \\ -\mu & 0 \end{pmatrix}$, which has eigenvalues $\pm i\,\mu$.

Remark 9.6: If A is orthogonal on $V_\mathbb{R}$, then the natural extension of A to $V_\mathbb{C}$ is unitary, so that every eigenvalue $\lambda \in \mathbb{C}$ must satisfy $\overline{\lambda}\,\lambda = 1$, i.e. λ has modulus 1. After taking care of the eigenvalues equal to $+1$ or -1, one is led to work on the orthogonal, where neither $+1$ nor -1 is an eigenvalue, so that $\lambda = \cos\theta + i\,\sin\theta$ with $\theta \neq k\,\pi$, and $A\,(v + i\,w) = (\cos\theta + i\,\sin\theta)\,(v + i\,w)$ means $A\,v = \cos\theta\,v - \sin\theta\,w$ and $A\,w = \sin\theta\,v + \cos\theta\,w$, which imply that $v - i\,w$ is an eigenvector of A for the eigenvalue $\cos\theta - i\,\sin\theta \neq \cos\theta + i\,\sin\theta$ so that $v - i\,w$ is orthogonal to $v + i\,w$; since $(v - i\,w, v + i\,w) = ||v||^2 - ||w||^2 - 2i\,(v, w)$, and $||v + i\,w||^2 = ||v||^2 + ||w||^2$, one deduces that $||v|| = ||w|| \neq 0$ and $(v, w) = 0$; by normalizing v and w in $V_\mathbb{R}$, it corresponds to a diagonal block $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, which is a rotation of angle $-\theta$ and has eigenvalues $\cos\theta \pm i\,\sin\theta$.

If $A \in SO(n)$ (i.e. V has dimension n, and $det(A) = +1$), then the multiplicity of the eigenvalue -1 is even, and one may create diagonal blocks $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ which correspond to rotation of π.

Remark 9.7: If A is normal on a Hermitian space, it has an orthonormal basis of eigenvectors $e_1, \ldots, e_n$ with $A\,e_i = \lambda_i e_i$ and $A^*e_i = \overline{\lambda_i}\,e_i$ for $i = 1, \ldots, n$. If there are m distinct eigenvalues of A, let $P \in \mathbb{C}[x]$ be the interpolation polynomial of degree $\leq m-1$ such that $P(\lambda_i) = \overline{\lambda_i}\,e_i$ for $i = 1, \ldots, n$; then, one has $A^* = P(A)$ on this basis, hence in every basis: a normal operator is then any operator such that $A^* = P(A)$ for some polynomial P (which implies that A^* commutes with A).

If A commutes with A^T on $V_\mathbb{R}$, then its extension to $V_\mathbb{C}$ is normal, and since $\overline{\lambda_i}$ is also an eigenvalue of A, the interpolation polynomial satisfies $P(\overline{\lambda_i}) = \lambda_i$, and it is easy to check on the explicit formula giving the Lagrange interpolation polynomial that $P \in \mathbb{R}[x]$.

Remark 9.8: If V is an n-dimensional E-vector space, $A \in L(V, V)$ and its characteristic polynomial $P_{char}(x) = det(A - x\,I)$ splits over E (for example if E is algebraically closed), then if the distinct eigenvalues are $\lambda_1, \ldots, \lambda_r$ with algebraic multiplicity $m_1, \ldots, m_r$, there is a unique decomposition $V = V_1 \oplus \ldots \oplus V_r$, where for $j = 1, \ldots, r$, $dim(V_j) = m_j$, A maps V_j into itself and has only the eigenvalue λ_j, and $V_j = ker\big((A - \lambda_i)^{k_j}\big)$ for some smallest k_j. Selecting $j \in \{1, \ldots, r\}$, and restricting attention to $W = V_j$ (with $dim(W) = m$) and writing $A = \lambda_j I + B$ with $B \in L(W, W)$ satisfying $B^k = 0$ and $B^{k-1} \neq 0$ for some $k \geq 1$, one looks for a (non uniquely defined) basis where the matrix of B has a simple form, and then by adding λ in the diagonal one recovers the form of A.

Let $Y_i = ker(B^i)$, and $d_i = dim(Y_i)$ for $i = 1, \ldots, k$, so that $0 < d_1 < \ldots < d_k = m$ (d_1 is the geometric multiplicity of the eigenvalue 0, $Y_k = W$ has dimension the algebraic multiplicity of the eigenvalue 0). There are other inequalities satisfied by $d_1, \ldots, d_k$, namely $d_2 - d_1 \geq d_3 - d_2 \geq \ldots \geq d_{k-1} - d_k$, which follow from the construction of a particular basis, made of Jordan blocks, of maximum size k: for an eigenvalue λ, a Jordan block of size d is the $d \times d$ matrix of the mapping M defined by $M\,e_1 = \lambda\,e_1$ and $M\,e_j = \lambda\,e_j + e_{j-1}$ for $j = 2, \ldots, d$ (i.e. with λs in the diagonal, 1s in the diagonal above it, and 0s elsewhere).