Hermitian and unitary operators *[still Sec. 18]*

Let $\mathcal{V}$ be a Hilbert space. (This includes finite-dimensional inner product spaces!) I shall phrase the initial definitions in sufficient generality to cover the case of "unbounded" operators.

Definitions: Let $\underline{A}$ be a linear operator with codomain $\mathcal{V}$ whose domain is a subspace of $\mathcal{V}$. $\underline{A}$ is *Hermitian* [or *symmetric*] if

$$\underline{A}\vec{v}_1 \cdot \vec{v}_2 = \vec{v}_1 \cdot \underline{A}\vec{v}_2, \qquad \forall \vec{v}_1, \vec{v}_2 \in \operatorname{dom} \underline{A}.$$

$\underline{A}$ is *self-adjoint* if $\underline{A}^* = \underline{A}$.

Remark: The usual terminological correlation is

Hermitian $\leftrightarrow$ complex space

symmetric $\leftrightarrow$ real space.

If $\dim \mathcal{V} = \infty$, we may need to consider $\underline{A}$ such that $\operatorname{dom} \underline{A}$ is not equal to all of $\mathcal{V}$. Example: As we have seen earlier, if $\underline{A}$ is a differential operator, then $\operatorname{dom} \underline{A}$ must contain only sufficiently differentiable functions $\subset \mathcal{L}^2$. However, as previously remarked without proof, if $\underline{A}$ *can* be defined everywhere in the Hilbert space $\mathcal{V}$, and if it satisfies the additional technical assumption of boundedness (equivalent to continuity), then $\underline{A}^*$ will be defined everywhere in $\mathcal{V}$. In that situation, as in finite dimensions, the distinction between Hermiticity and self-adjointness is unnecessary:

Theorem 18.5$'$. *If* $\operatorname{dom} \underline{A} = \mathcal{V} = \operatorname{dom} \underline{A}^*$, *then*

$$\underline{A} \text{ Hermitian} \iff \underline{A} \text{ self-adjoint.}$$

(If $\operatorname{dom} \underline{A} \neq \mathcal{V}$ and $\underline{A}$ is Hermitian, then $\operatorname{dom} \underline{A}$ will be a *subset* of $\operatorname{dom} \underline{A}^*$ but may not be all of it; in the latter case, $\underline{A}$ is not self-adjoint and $\underline{A}^*$ may not even be Hermitian.)

Proof: Obvious from definitions.

Remark: Most of the special classes of operators about which we need to prove theorems in this course, such as projections and isometries (see below), are in fact always bounded and everywhere defined. Therefore, we can deal freely with their adjoints without worrying about domain technicalities. The rest of the present discussion is confined to that case.

Definition: $\underline{A}$ is *anti-Hermitian* [or *skew-symmetric*, etc.] if $\underline{A}\vec{v}_1 \cdot \vec{v}_2 = -\vec{v}_1 \cdot \underline{A}\vec{v}_2$; in other words, $\underline{A}^* = -\underline{A}$.

THEOREM. *With respect to an* **ON** *basis,* $\underline{A}$ *is Hermitian iff* $\underline{A}$ *is represented by an Hermitian matrix.* (See next definition.) *(Similarly for (real) symmetric, anti-Hermitian, etc.)*

DEFINITIONS: $\{A^j{}_k\}$ is *Hermitian* if $A^k{}_j = \overline{A^j{}_k}$,
symmetric if $A^k{}_j = A^j{}_k$, ...

PROOF OF THEOREM: Recall that $(\underline{A}\vec{v})^k = \sum_j A^k{}_j v^j$ defines A. Let $\{\hat{e}_j\}_{j=1}^n$ be the ON basis. Then $A^k{}_j = (\underline{A}\hat{e}_j)\cdot\hat{e}_k$. [Check: $(\underline{A}\hat{e}_j)\cdot\hat{e}_k = (\underline{A}\hat{e}_j)^k = \sum_l A^k{}_l(\hat{e}_j)^l = A^k{}_j$.] On the other hand,

$$\overline{A^j{}_k} = \overline{(\underline{A}\hat{e}_k)\cdot\hat{e}_j} = \hat{e}_j\cdot(\underline{A}\hat{e}_k) = (\underline{A}\hat{e}_j)\cdot\hat{e}_k,$$

QED.

QUICKER PROOF: This is a corollary of the theorem that

$$\text{matrix of adjoint} = \overline{\text{transpose of matrix}}.$$

THEOREM 18.6.

(a) $\mathcal{V}$ *real* $\Rightarrow$ $\mathcal{L}(\mathcal{V};\mathcal{V}) = \mathcal{S}(\mathcal{V};\mathcal{V}) \oplus \mathcal{A}(\mathcal{V};\mathcal{V})$, *where* $\mathcal{S}$ *and* $\mathcal{A}$ *are the symmetric and antisymmetric operators. That is, any* $\underline{L}\in\mathcal{L}$ *can be uniquely decomposed as*

$$\underline{L} = \underline{S} + \underline{A}, \qquad \underline{S}\in\mathcal{S}, \quad \underline{A}\in\mathcal{A}.$$

(b) $\mathcal{V}$ *complex* $\Rightarrow$ $\mathcal{S}$ *(the Hermitian operators) and* $\mathcal{A}$ *are not subspaces: indeed,* $\mathcal{A} = i\mathcal{S}$. *The decomposition still exists, but it is not a direct sum.*

PROOF: $\underline{S} = \frac{1}{2}(\underline{L} + \underline{L}^*)$, $\qquad \underline{A} = \frac{1}{2}(\underline{L} - \underline{L}^*)$.

Uniqueness: If $\underline{B}$ is both symmetric and antisymmetric, then $\underline{B} = \underline{0}$. Thus the sum is direct.

Definitions: $\underline{A}:\mathcal{V}\to\mathcal{V}$ is *isometric* (or is an *isometry*) if

$$(\underline{A}\vec{v}_1)\cdot(\underline{A}\vec{v}_2) = \vec{v}_1\cdot\vec{v}_2, \qquad \forall \vec{v}_1,\vec{v}_2\in\mathcal{V}.$$

$\underline{A}$ is *unitary* if $\underline{A}^{-1} = \underline{A}^*$ (complex case).

$\underline{A}$ is *orthogonal* if $\underline{A}^{-1} = \underline{A}^* \equiv$ transpose of $\underline{A}$ (real case).

REMARKS:

1. The definitions of "unitary" and "orthogonal" carry over to *matrices* with respect to ON bases.

2. For $\underline{A} \in \mathcal{L}(\mathcal{V};\mathcal{U})$, $\quad \mathcal{V} \neq \mathcal{U}$, the definitions of "isometric" and "unitary" still make sense. (However, instead of "unitary" one usually says "isometric isomorphism" in this case.) Part (e) of the next theorem doesn't hold if $\dim \mathcal{V} \neq \dim \mathcal{U}$. Part (a) of the theorem shows that no isometries exist if $\dim \mathcal{U} < \dim \mathcal{V}$.

3. Bowen & Wang do not make a distinction between isometries and unitary operators. This is partly explained by their finite-dimensional focus. It is partly justified by the observation that an isometry becomes a unitary operator if its range is taken to be the codomain (see part (d) of the theorem).

THEOREM 18.7′,9′.

(a) *Isometric* $\Rightarrow$ *injective.*

(b) *Isometric* $\iff$ $\underline{A}^*$ *is a left inverse (for* $\underline{A}$*).*

(c) *Unitary* $\Rightarrow$ *isometric.*

(d) *Isometric & surjective* $\iff$ *unitary.*

(e) $\dim \mathcal{V} < \infty,\ \underline{A} \in \mathcal{L}(\mathcal{V};\mathcal{V}) \ \Rightarrow :$ *Isometric* $\iff$ *unitary.*

PROOF:

(a) It suffices to show $\ker \underline{A} = \{\vec{0}\}$.

$$\underline{A}\vec{v} = \vec{0} \ \Rightarrow \ 0 = \underline{A}\vec{v} \cdot \underline{A}\vec{v} = \vec{v} \cdot \vec{v} \ \Rightarrow \ \vec{v} = \vec{0}.$$

(b) Isometric $\iff$ $(\underline{A}\vec{v}_1) \cdot (\underline{A}\vec{v}_2) = \vec{v}_1 \cdot \vec{v}_2$

$$\begin{aligned}
&\iff \vec{v}_1 \cdot (\underline{A}^*\underline{A}\vec{v}_2) = \vec{v}_1 \cdot \vec{v}_2 \quad (\forall \vec{v}_1) \\
&\iff \underline{A}^*\underline{A}\vec{v}_2 = \vec{v}_2 \quad (\forall \vec{v}_2) \\
&\iff \underline{A}^*\underline{A} = \underline{1}.
\end{aligned}$$

(dom $\underline{A}^* = \mathcal{V}$ since $\underline{A}$ is bounded.)

(c) Unitary $\iff$ $\underline{A}^{-1} = \underline{A}^*$ $\iff$ $\underline{A}^*\,\underline{A} = \underline{1}$ **and** $\underline{A}\,\underline{A}^* = \underline{1}$.
Therefore $\underline{A}$ is isometric by (b).

(d) Injective & surjective $\iff$ invertible $\iff$ left inverse = inverse.
But isometric $\iff$ the left inverse is $\underline{A}^*$. So conclusion follows from proof of (c).

(e) For a finite dimensional space, injective $\iff$ surjective. Conclusion follows from (b) or (d).

Remark: $\dim \mathcal{V} = \infty \;\Rightarrow\;$ (e) not true: Let $\underline{A}$ = right shift in ℓ^2 (the space of square-summable sequences):

$$\underline{A}(x^1, x^2, \dots) \equiv (0, x^1, x^2, \dots).$$

Clearly $\underline{A}$ is isometric: $\sum x^j \overline{y^j} < \infty$ is unchanged. But $\underline{A}$ is not surjective, hence not unitary. ($\underline{A}^{-1}$ doesn't exist.)

Note: $\underline{A}^*$ = left shift:

$$\underline{A}^*(y^1, y^2, y^3, \dots) = (y^2, y^3, \dots).$$

$$\begin{aligned}
(y^1, y^2, \dots) \cdot \underline{A}(x^1, x^2, \dots) &= (y^1, y^2, \dots) \cdot (0, x^1, x^2, \dots) \\
&= \sum_{j=1}^{\infty} y^{j+1} \overline{x^j} \\
&= (y^2, y^3, \dots) \cdot (x^1, x^2, \dots).
\end{aligned}$$

QED

One can easily verify that $\underline{A}^* \underline{A} = \underline{1}$. What is $\underline{A}\,\underline{A}^*$?

$$\underline{A}\,\underline{A}^*(x^1, x^2, x^3, \dots) = (0, x^2, x^3, \dots).$$

Therefore, $\underline{A}\,\underline{A}^* = \underline{P} \equiv$ projection onto $\operatorname{ran} \underline{A}$ along $\ker \underline{A}^*$. This is a general property:

Theorem. *$\underline{A}$ isometric $\Rightarrow$ $\underline{A}^* \underline{A} = \underline{1}$ and $\underline{A}\,\underline{A}^*$ = the orthogonal projection onto* $\operatorname{ran} \underline{A}$. (For the moment, "orthogonal" for projections means $\operatorname{ran} \underline{P} \perp \ker \underline{P}$. An alternative definition will appear soon.)

Proof: The only thing not already proved is that $\underline{A}\,\underline{A}^* = \underline{P}$ in case $\operatorname{ran} \underline{A} \neq \mathcal{V}$. Well, $(\underline{A}\,\underline{A}^*)(\underline{A}\,\underline{A}^*) = \underline{A}(\underline{A}^*\underline{A})\underline{A}^* = \underline{A}\,\underline{A}^*$, so it's a projection, $\underline{P}$. Also,

$$\vec{v} \in \operatorname{ran} \underline{A} \;\Rightarrow\; \underline{A}\,\underline{A}^*\vec{v} = \underline{A}\,\underline{A}^*\underline{A}\vec{u} = \underline{A}\vec{u} = \vec{v} \;\Rightarrow\; \vec{v} \in \operatorname{ran} \underline{P},$$

and

$$\vec{v} \in (\operatorname{ran} \underline{A})^\perp = \ker \underline{A}^* \;\Rightarrow\; \underline{A}^*\vec{v} = \vec{0} \;\Rightarrow\; \underline{A}\,\underline{A}^*\vec{v} = \vec{0} \;\Rightarrow\; \vec{v} \in \ker \underline{P}.$$

It follows that $\overline{\operatorname{ran} \underline{A}} = \operatorname{ran} \underline{P}$ and $(\operatorname{ran} \underline{A})^\perp = \ker \underline{P}$. Finally, $\underline{P}\vec{v} = \underline{A}(\underline{A}^*\vec{v}) \;\Rightarrow\; \operatorname{ran} \underline{P} \subseteq \operatorname{ran} \underline{A}$, so $\operatorname{ran} \underline{P} = \operatorname{ran} \underline{A}$.

Theorem 18.8. *$\underline{A}$ isometric $\iff$ $\|\underline{A}\vec{v}\| = \|\vec{v}\|, \quad \forall \vec{v}$. (I.e., an operator preserves the inner product if and only if it preserves the norm.)*

Proof: $\Rightarrow$ is trivial; $\Leftarrow$ follows from the polarization identity (see Example 2A in the section of these notes introducing inner products).

Theorem. $\{\hat{e}_j\}$ *an ON basis* $\Rightarrow$

$$\{\underline{A}\hat{e}_j\} \text{ is an } \begin{cases} \text{ON set if } \underline{A} \text{ is isometric,} \\ \text{ON basis if } \underline{A} \text{ is unitary} \end{cases}$$

PROOF: Write out $\underline{A}^*\underline{A} = \underline{1}$ as a matrix equation:

$$\sum_l \overline{A^l{}_j} A^l{}_k = \delta^j{}_k .$$

That is,

$$(\underline{A}\hat{e}_k) \cdot (\underline{A}\hat{e}_j) = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

Moreover, if $\underline{A}$ is surjective (the unitary case), then the image of the basis spans $\mathcal{V}$, hence is a basis. (This theorem also follows more abstractly from the definitions, but this matrix observation is really more instructive.)

THEOREM. *Let* $\dim \mathcal{V} < \infty$ *and let* A *be a matrix representing the unitary or orthogonal operator* $\underline{A}$ *(with respect to the same basis for* $\mathcal{V}$ *as both domain and codomain). Then* $|\det A| = 1$.

PROOF: Later we will see that the determinant is independent of the basis chosen. For now consider a matrix with respect to an ON basis, and note that $\det A^* = \overline{\det A}$. Therefore,

$$|\det A|^2 = \det (A^* A) = \det \underline{1} = 1.$$

In the complex case, $\det A = e^{i\theta}$ for some $\theta \in \mathbf{R}$. In the real case, $\det \underline{A} = \pm 1$. An orthogonal transformation with $\det A = +1$ is a *rotation.* If $\det A = -1$, a rotation (possibly trivial) is combined with a *reflection*, such as

$$(x^1, x^2, \dots) \mapsto (-x^1, x^2, \dots).$$

MOTIVATIONAL REMARKS: Isometric isomorphisms preserve the structure of inner product spaces, so they are natural things to study. Self-adjoint operators lead to nice and powerful theorems and also appear in many applications. In particular, they "generate" unitary groups: $\underline{A}$ self-adjoint $\Rightarrow$ $e^{i\underline{A}t}$ unitary and $e^{i\underline{A}t_1} e^{i\underline{A}t_2} = e^{i\underline{A}(t_1+t_2)}$; equivalently, $\underline{A} = -i\frac{d}{dt} e^{i\underline{A}t}\Big|_{t=0}$. (The definition of the exponential of an operator will come later.)

Now the unfinished business about projections:

Definition: An *orthogonal projection* is a $\underline{P} \in \mathcal{L}(\mathcal{V};\mathcal{V})$ satisfying

$$\underline{P}^2 = \underline{P} = \underline{P}^*.$$

REMARK: An orthogonal projection is *not* an orthogonal operator! ($\underline{P}^{-1}$ doesn't even exist unless $\underline{P} = \underline{1}$.) It *is* a self-adjoint operator.

To justify this definition we need:

Theorem 18.10. *If $\underline{P}$ is a projection, then*

$$\underline{P} = \underline{P}^* \iff \operatorname{ran} \underline{P} = (\ker\, \underline{P})^\perp.$$

That is, orthogonal projections correspond to orthogonal direct-sum decompositions. $\underline{P}$ is the projection onto $\operatorname{ran} \underline{P}$ along its orthogonal complement.

PROOF:

$\Rightarrow$: By Theorem 18.3,

$$\operatorname{ran} \underline{P} = (\ker\, \underline{P}^*)^\perp = (\ker\, \underline{P})^\perp.$$

(We don't have to worry about proving $\operatorname{ran} \underline{P}$ closed in the infinite-dimensional case, because in Theorem 17.3 we already established that $\operatorname{ran} \underline{P}$ is a *direct* complement of $\ker\, \underline{P}$.)

$\Leftarrow$: By hypothesis, $(\underline{P}\vec{u}) \cdot \vec{v} = 0$ for all $\vec{v} \in \ker\, \underline{P}$ and all $\vec{u} \in \mathcal{V}$. Every $\vec{w} \in \mathcal{V}$ has the form $\vec{w} = \underline{P}\vec{w} + (\underline{1} - \underline{P})\vec{w}$, where the second term is in $\ker\, \underline{P}$. Therefore,

$$(\underline{P}\vec{u}) \cdot \vec{w} = (\underline{P}\vec{u}) \cdot (\underline{P}\vec{w}).$$

By symmetry,

$$\vec{u} \cdot (\underline{P}\vec{w}) = (\underline{P}\vec{u}) \cdot (\underline{P}\vec{w}).$$

Equality of the left sides of these equations says that $\underline{P} = \underline{P}^*$, since $\vec{u}$ and $\vec{w}$ are arbitrary.

EXAMPLE: In the context of the theorem about the projection $\underline{P} \equiv \underline{A}\,\underline{A}^*$ onto the range of an isometry, we have $(\underline{A}\,\underline{A}^*)^* = \underline{A}^{**}\underline{A}^* = \underline{A}\,\underline{A}^*$. This confirms that $\underline{P}$ is an orthogonal projection. (Since isometries and their adjoints are bounded operators, $\underline{A}^{**}$ is always equal to $\underline{A}$, even in infinite dimensions.)