1 Wigner's theorem

Let $\mathcal{H}$ be an arbitrary Hilbert space. We define a relation $\sim$ on $\mathcal{H}$ by $x \sim y$ if there's a $c \in \mathbb{C}$ such that $|c| = 1$ and $x = cy$. This is clearly an equivalence relation. For each $x \in \mathcal{H}$, the equivalence class that x belongs to will be denoted by $[x]$. The set of equivalence classes will be denoted by $\mathcal{S}$. For each $a \in \mathbb{C}$ and each $x, y \in \mathcal{H}$, we define

$$a[x] = [ax] \tag{1}$$

$$[x] \cdot [y] = |\langle x, y \rangle|. \tag{2}$$

Note that the right-hand sides don't depend on the representatives x, y from the equivalence classes $[x], [y]$. We will be particularly interested in the equivalence classes $[x]$ such that $\|x\| = 1$. These classes are called the *unit rays* of $\mathcal{H}$. (A *ray* of $\mathcal{H}$ is a 1-dimensional subspace of $\mathcal{H}$). The set of unit rays of $\mathcal{H}$ will be denoted by $\mathcal{R}$.

In this section, the symbols θ and η will denote automorphisms of $\mathbb{C}$. For all $a \in \mathbb{C}$, we will write a^θ and a^η instead of $\theta(a)$ and $\eta(a)$.

Definition 1.1 (θ-unitary)**.** Suppose that θ is an automorphism of $\mathbb{C}$. An operator $U : \mathcal{H} \to \mathcal{H}$ is said to be *θ-linear* if for all $a, b \in \mathbb{C}$ and all $x, y \in \mathcal{H}$,

$$U(ax + by) = a^\theta Ux + b^\theta Uy.$$

A θ-linear operator $U : \mathcal{H} \to \mathcal{H}$ is said to be *θ-unitary* if for all $x, y \in \mathcal{H}$,

$$\langle Ux, Uy \rangle = \langle x, y \rangle^\theta.$$

Let I be the identity map on $\mathbb{C}$. Denote the complex conjugation map $\lambda \mapsto \lambda^*$ on $\mathbb{C}$ by I^*.

Theorem 1.2 (Wigner's theorem)**.** *If T is a permutation of $\mathcal{R}$ such that $T[x] \cdot T[y] = [x] \cdot [y]$ for all $x, y \in \mathcal{H} - \{0\}$, then there's a $\theta \in \{I, I^*\}$ and a θ-unitary $U : \mathcal{H} \to \mathcal{H}$ such that $Ux \in T[x]$ for all $x \in \mathcal{H} - \{0\}$. If $\dim \mathcal{H} \geq 2$, then θ is uniquely determined by T, and U is unique up to multiplication by a complex number of absolute value* 1.

The proof is very long, so instead of trying to prove it all at once, we're going to state and prove a number of lemmas that lead up this result. Lemma 1.15 will be the final step.

Lemma 1.3 (Extension of T from $\mathcal{R}$ to $\mathcal{S}$)**.** *For each $x \in \mathcal{H}$, we define $T[x] = \|x\| T[e]$, where e is the unit vector in the direction of x. The map $T : \mathcal{S} \to \mathcal{S}$ defined this way has the following properties.*

(a) $T[ax] = aT[x]$ *for all* $a \in \mathbb{C}$ *and all* $x \in \mathcal{H}$.

(b) $T[x] \cdot T[y] = [x] \cdot [y]$ *for all* $x, y \in \mathcal{H}$.

(c) For all $x \in \mathcal{H}$, *and all* $x' \in T[x]$, *we have* $\|x'\| = \|x\|$.

Proof. Let $a \in \mathbb{C}$ and $x, y \in \mathcal{H}$ be arbitrary. Define $e = \frac{x}{\|x\|}$ and $f = \frac{y}{\|y\|}$.

(a):

$$T[ax] = T\big[a\|x\|e\big] = a\|x\|T[e] = aT\big[\|x\|e\big] = aT[x]. \tag{3}$$

(b): Let $e' \in T[e]$ and $f' \in T[f]$ be arbitrary. Since

$$\begin{aligned} T[x] &= T\big[\|x\|e\big] = \|x\|T[e] = \|x\|\,[e'] = \big[\|x\|e'\big] \\ T[y] &= \big[\|y\|f'\big], \end{aligned} \tag{4}$$

we have

$$\begin{aligned} T[x] \cdot T[y] &= \big[\|x\|e'\big] \cdot \big[\|x\|f'\big] = \big|\big\langle \|x\|e', \|y\|f' \big\rangle\big| \\ &= \|x\|\,\|y\|\,\underbrace{|\langle e', f'\rangle|}_{= [e'] \cdot [f'] = T[e] \cdot T[f] = [e] \cdot [f] = |\langle e, f\rangle|} = \big|\big\langle \|x\|e, \|y\|f \big\rangle\big| = |\langle x, y\rangle| = [x] \cdot [y]. \end{aligned} \tag{5}$$

(c): For all $x' \in T[x]$,

$$\|x'\|^2 = |\langle x', x'\rangle| = [x'] \cdot [x'] = T[x] \cdot T[x] = [x] \cdot [x] = |\langle x, x\rangle| = \|x\|^2. \tag{6}$$

$\square$

The following lemma is the only theorem in this section that's not a part of the proof of Wigner's theorem. We're proving it because it answers a question suggested by the previous theorem.

Lemma 1.4 (The extended T is a bijection)**.** *The* $T : \mathcal{S} \to \mathcal{S}$ *defined above is a permutation of* $\mathcal{S}$.

Proof. Injectivity: Let x and y be arbitrary members of $\mathcal{H}$ such that $T[x] = T[y]$. Let $x' \in T[x]$ be arbitrary. Since $x' \in T[y]$, lemma 1.3(c) tells us that $\|x'\| = \|y\|$. Similarly, $\|y'\| = \|x\|$. Since x' and y' belong to the same equivalence class, we also have $\|x'\| = \|y'\|$. So $\|x\| = \|y'\| = \|x'\| = \|y\|$. Define $e = \frac{x}{\|x\|}$ and $f = \frac{y}{\|y\|}$. Let $e' \in T[e]$ and $f' \in T[f]$ be arbitrary.

$$\begin{aligned} T[x] = T[y] \;&\Rightarrow\; T\big[\|x\|e\big] = T\big[\|y\|f\big] \;\Rightarrow\; \|x\|\,T[e] = \|x\|\,T[f] \\ &\Rightarrow\; \|x\|\,[e'] = \|x\|\,[f'] \;\Rightarrow\; \big[\|x\|e'\big] = \big[\|x\|f'\big]. \end{aligned} \tag{7}$$

This result implies that there's a $c \in \mathbb{C}$ such that $|c| = 1$ and $\|x\|e = c\|x\|f$. Clearly, any such c also satisfies $e' = cf'$. So $[e'] = [f']$. This means that $T[e] = T[f]$. Since the original T is a permutation of $\mathcal{R}$, this implies that $[e] = [f]$. Let c be a complex number such that $e = cf$. Clearly, $\|x\|e = c\|x\|f = c\|y\|f$. So $\big[\|x\|e\big] = \big[\|y\|f\big]$. This means that $[x] = [y]$.

Surjectivity: Let $y \in \mathcal{H}$ be arbitrary. Define $f = \frac{y}{\|y\|}$. Let $e \in \mathcal{H}$ be unit vector such that $T[e] = [f]$. We have

$$[y] = \big[\|y\|f\big] = \|y\|\,[f] = \|y\|\,T[e] = T\big[\|y\|e\big]. \tag{8}$$

□

Lemma 1.5 (Properties of any partially defined U)**.** *Let $\mathcal{D}$ be an arbitrary subset of $\mathcal{H}$. Let $U : \mathcal{D} \to \mathcal{H}$ be an arbitrary map such that $U0 = 0$ if $0 \in \mathcal{D}$, and $Ux \in T[x]$ for all $x \in \mathcal{D}$ such that $x \neq 0$.*

(a) For all $x, y \in \mathcal{D}$, $|\langle Ux, Uy\rangle| = |\langle x, y\rangle|$.

(b) For all $x \in \mathcal{D}$, $\|Ux\| = \|x\|$.

(c) For each $x \in \mathcal{D}$ such that $x \neq 0$, there's a unique function $p_x : \mathbb{C} \to \mathbb{C}$ such that $U(ax) = p_x(a)Ux$ and $|p_x(a)| = |a|$ for all $a \in \mathbb{C}$.

Proof. Let $x, y \in \mathcal{D}$ and $a \in \mathbb{C}$ be arbitrary.

(a): If $x = 0$ or $y = 0$, we have $|\langle Ux, Uy\rangle| = 0 = |\langle x, y\rangle|$. If $x \neq 0$ and $y \neq 0$, we have $|\langle Ux, Uy\rangle| = T[x] \cdot T[y] = [x] \cdot [y] = |\langle x, y\rangle|$.

(b): Part (a) implies that $\|Ux\|^2 = |\langle Ux, Ux\rangle| = |\langle x, x\rangle| = \|x\|^2$.

(c): Suppose that $x \neq 0$. We have $[U(ax)] = T[ax] = aT[x] = a[Ux] = [aUx]$. So there's a unique $c \in \mathbb{C}$ such that $|c| = 1$ and $U(ax) = caUx$. Define $p_x(a) = ca$.

Suppose that $q_x : \mathbb{C} \to \mathbb{C}$ is such that $U(bx) = q_x(b)Ux$ for all $b \in \mathbb{C}$. Then $q_x(a)Ux = U(ax) = p_x(a)Ux$, and therefore $(q_x(a) - p_x(a))Ux = 0$. If $q_x(a) \neq p_x(a)$, we can multiply this by $1/(q_x(a) - p_x(a))$ to get $Ux = 0$. This contradicts part (b) or the assumption that $x \neq 0$. So $q_x(a) = p_x(a)$. Since a is an arbitrary member of $\mathbb{C}$, this implies that $q_x = p_x$. □

Lemma 1.6 (Linear combinations of orthonormal vectors)**.** *Let $\langle e_k\rangle_{k=1}^n$ be an arbitrary orthonormal finite sequence in $\mathcal{H}$. Define $I = \{1, \dots, n\}$. For each $k \in I$, let $e'_k \in T[e_k]$ be arbitrary. If $x = \sum_{k=1}^n a_k e_k$, then for each $x' \in T[x]$, there are complex numbers $a'_1, \dots, a'_n \in \mathbb{C}$ such that $x' = \sum_{k=1}^n a'_k e'_k$ and $|a'_k| = |a_k|$ for all $k \in I$.*

Proof. For all $i, j \in I$, we have

$$|\langle e'_i, e'_j \rangle| = [e'_i] \cdot [e'_j] = T[e_i] \cdot T[e_j] = [e_i] \cdot [e_j] = |\langle e_i, e_j \rangle| = \delta_{ij}.$$

This implies that $\langle e'_i, e'_j \rangle = \delta_{ij}$ for all $i, j \in I$. Note that for all $k \in I$, we have $a_k = \langle e_k, x \rangle$. This follows from $x = \sum_{k=1}^n a_k e_k$ and the fact that $\{e_k\}_{k \in I}$ is an orthonormal set. We're going to define a'_k for each $k \in I$. Since $\{e'_k\}_{k \in I}$ is an orthonormal set and $x' = \sum_{k=1}^n a'_k e'_k$, there's only one definition that can possibly work: For each $k \in I$, we define $a'_k = \langle e'_k, x' \rangle$. For all $k \in I$, we have

$$|a'_k| = |\langle e'_k, x' \rangle| = [e'_k] \cdot [x'] = T[e_k] \cdot T[x] = [e_k] \cdot [x] = |\langle e_k, x \rangle| = |a_k|. \tag{9}$$

We will prove that $x' = \sum_{k=1}^n a'_k e'_k$. First note that

$$\left\| x' - \sum_{k=1}^n a'_k e'_k \right\|^2 = \|x'\|^2 - 2 \operatorname{Re} \left\langle x', \sum_{k=1}^n a'_k e'_k \right\rangle + \left\| \sum_{k=1}^n a'_k e'_k \right\|^2 \tag{10}$$

Since $\langle a'_k e'_k \rangle_{k=1}^n$ is an orthogonal finite sequence in $\mathcal{H}$, the Pythagorean theorem tells us that the third term is equal to $\sum_{k=1}^n \|a'_k e'_k\|^2 = \sum_{k=1}^n |a'_k|^2$. To evaluate the second term, we note that

$$\left\langle x', \sum_{k=1}^n a'_k e'_k \right\rangle = \sum_{k=1}^n a'_k \langle x', e'_k \rangle = \sum_{k=1}^n |a'_k|^2 \in \mathbb{R}. \tag{11}$$

These results imply that

$$\left\| x' - \sum_{k=1}^n a'_k e'_k \right\|^2 = \|x'\|^2 - \sum_{k=1}^n |a'_k|^2 = \|x\|^2 - \sum_{k=1}^n |a_k|^2 = \left\| x - \sum_{k=1}^n a_k e_k \right\|^2 = 0. \tag{12}$$

□

Let e be an arbitrary unit vector in $\mathcal{H}$. e will denote the same vector until the end of the section.

Definition 1.7 (Definition of Ux for all $x \in \mathcal{H}$ with $\langle e, x \rangle \in \{0, 1\}$)**.** We will define Ux for all x in the subset $\{e + y | y \perp e\}$, and then define Ux for all x in the Hilbert subspace $\{e\}^\perp$.

Let $y \in \{e\}^\perp$ be arbitrary. Define f by $f = y/\|y\|$. Let $e' \in T[e]$ and $f' \in T[f]$ be arbitrary. Lemma 1.6 tells us that since $\{e, f\}$ is an orthonormal set and $e + y = e + \|y\| f$, there exist $a, b \in \mathbb{C}$ such that $ae' + bf' \in T[e + y]$, and $|a| = 1$, $|b| = \|y\|$. Since $T[e + y]$ is an equivalence class whose members differ only by complex factors of absolute value 1, this means that there's a unique member of $T[e+y]$ that can be expressed as $e' + bf'$, where $b \in \mathbb{C}$. Let b be the unique member of $\mathbb{C}$ such that $e' + bf' \in T[e+y]$. Define $U(e+y) = e' + bf'$. Since y is an arbitrary member of $\{e\}^\perp$, this defines Ux for all x in $\{e+y | y \perp e\}$. Since $|b| = \|y\|$, the definition implies that $Ue = e'$.

For each $y \in \{e\}^\perp$, we define $Uy = U(e+y) - Ue$. This defines Ux for all x in $\{e\}^\perp$. So Ux is now defined for all x in $\big\{x \in \mathcal{H} | \langle e, x\rangle \in \{0,1\}\big\}$. This set will be denoted by $\mathcal{D}$.

Lemma 1.8 (Useful facts about complex numbers)**.** *Let $a, b \in \mathbb{C}$ be arbitrary.*

(a) If $\operatorname{Re} a = \operatorname{Re} b$ and $|a| = |b|$, then $b = a$ or $b = a^*$.

(b) If $\operatorname{Re}(ab^*) = \operatorname{Re}(ab)$, then $(\operatorname{Im} a)(\operatorname{Im} b) = 0$.

(c) If $\operatorname{Re}(|a|^2 b) = \operatorname{Re}(a^2 b)$ and $\operatorname{Im} a \neq 0$, then $(\operatorname{Re} a)(\operatorname{Im} b) = -(\operatorname{Im} a)(\operatorname{Re} b)$.

Proof. Let (p, q, r, s) be the unique 4-tuple of real numbers such that $a = p + iq$ and $b = r + is$.

(a): By assumption, $p = r$ and $|a| = |b|$. So

$$p^2 + q^2 = |a|^2 = |b|^2 = r^2 + s^2 = p^2 + s^2.$$

This implies that $s = \pm q$. So either $b = r + is = p + iq$, or $b = r + is = p - iq$.

(b): Since

$$\begin{aligned} \operatorname{Re}(ab) &= \operatorname{Re}\big((p+iq)(r+is)\big) = pr - qs, \\ \operatorname{Re}(ab^*) &= \operatorname{Re}\big((p+iq)(r-is)^*\big) = pr + qs, \end{aligned} \tag{13}$$

the assumption implies that that $0 = qs = (\operatorname{Im} a)(\operatorname{Im} b)$.

(c): Since

$$\begin{aligned} \operatorname{Re}(|a|^2 b) &= \operatorname{Re}\big((p^2+q^2)(r+is)\big) = (p^2+q^2)r \\ \operatorname{Re}(a^2 b) &= \operatorname{Re}\big((p^2 - q^2 + 2ipq)(r+is)\big) = (p^2-q^2)r - 2pqs \end{aligned} \tag{14}$$

the assumption that $\operatorname{Re}(|a|^2 b) = \operatorname{Re}(a^2 b)$ implies that

$$-2pqs = (p^2+q^2)r - (p^2-q^2)r = 2q^2 r. \tag{15}$$

Since $q = \operatorname{Im} a \neq 0$, we can cancel q from this equality. So $(\operatorname{Re} a)(\operatorname{Im} b) = ps = -qr = -(\operatorname{Im} a)(\operatorname{Re} b)$. $\square$

Let $U : \mathcal{D} \to \mathcal{H}$ be the map defined by definition 1.7.

Lemma 1.9 (Properties of U)**.** *(a) $U0=0$*

(b) For all $x \in \mathcal{D}$ such that $x \neq 0$, $Ux \in T[x]$.

(c) For all $y, z \in \{e\}^\perp$, $\mathrm{Re}\langle Uy, Uz\rangle = \mathrm{Re}\langle y, z\rangle$.

(d) For all $y, z \in \{e\}^\perp$, $|\langle Uy, Uz\rangle| = |\langle y, z\rangle|$.

(e) For all $y, z \in \{e\}^\perp$, we have $\langle Uy, Uz\rangle = \langle y, z\rangle$ or $\langle Uy, Uz\rangle = \langle y, z\rangle^*$.

Note that part (e) implies that when $\langle y, z\rangle \in \mathbb{R}$, we have $\langle Uy, Uz\rangle = \langle y, z\rangle$.

Proof. Let $y, z \in \{e\}^\perp$ and $a \in \mathbb{C}$ be arbitrary.

(a): $U0 = U(e + 0) - U(e) = 0$.

(b): $U(e+y)$ is by definition a member of $T[e+y]$. We will prove that if $y \neq 0$, then $Uy \in T[y]$. Suppose that $y \neq 0$. Define $f = y/\|y\|$. We saw in definition 1.7 that there's an $f' \in T[f]$ and a $b \in \mathbb{C}$ such that $|b| = \|y\|$ and $Uy = bf'$. So

$$[Uy] = [bf'] = b[f'] = bT[f] = T[bf] = T[\|y\|f] = T[y]. \tag{16}$$

(c): Lemma 1.5 tells us that $|\langle U(e+y), U(e+z)\rangle| = |\langle e+y, e+z\rangle|$, and that

$$\begin{aligned} |\langle U(e+y), U(e+z)\rangle| &= |\langle Ue + Uy, Ue + Uy\rangle| \\ &= \left|\|Ue\|^2 + \langle Ue, Uz\rangle + \langle Ue, Uy\rangle + \langle Uy, Uz\rangle\right| \\ &= |1 + \langle Uy, Uz\rangle|^2 = 1 + 2\,\mathrm{Re}\langle Uy, Uz\rangle + |\langle Uy, Uz\rangle|^2 \\ &= 1 + 2\,\mathrm{Re}\langle Uy, Uz\rangle + |\langle y, z\rangle|^2, \\ |\langle e+y, e+z\rangle| &= \left|\|e\|^2 + \langle e, z\rangle + \langle e, y\rangle + \langle y, z\rangle\right| = \left|1 + \langle y, z\rangle\right|^2 \\ &= 1 + 2\,\mathrm{Re}\langle y, z\rangle + |\langle y, z\rangle|^2. \end{aligned} \tag{17}$$

(d): This follows from parts (a) and (b), and lemma 1.5(a).

(e): This follows from parts (c) and (d), and lemma 1.8(a). □

For each $x \in \mathcal{D}$ such that $x \neq 0$, let p_x be the unique function such that $U(ax) = p_x(a)Ux$ for all $a \in \mathbb{C}$.

Lemma 1.10 (p_f and p_g when $\{e, f, g\}$ is an orthonormal set)**.** *For all $a, b \in \mathbb{C}$ and all $f, g \in \{e^\perp\}$ such that $\|f\| = \|g\| = 1$ and $\langle f, g\rangle = 0$, we have $U(af + bg) = p_f(a)Uf + p_g(b)Ug$.*

Proof. Let $a, b \in \mathbb{C}$ be arbitrary. If $a = 0$ or $b = 0$, then we clearly have $U(af + bg) = p_f(a)Uf + p_g(b)Ug$. (If $a = b = 0$, the equality follows the result $U0 = 0$. If only one of a and b is zero, the equality follows from the definitions of p_f and p_g).

Suppose that $a, b \neq 0$. Define $x = af + bg$. Lemma 1.6 tells us that since $\{f, g\}$ is an orthonormal set, there exist $a', b' \in \mathbb{C}$ such that $Ux = a'Uf + b'Ug$, and $|a'| = 1$, $|b'| = 1$. We will prove that $a' = p_f(a)$ and $b' = p_g(b)$. First note that since

$$\begin{aligned} \langle f, f\rangle &= \|f\|^2 = 1 \in \mathbb{R}, \\ \langle x, af\rangle &= \langle af + bg, af\rangle = \langle af, af\rangle = |a|^2 \in \mathbb{R}, \end{aligned} \tag{18}$$

we have $\langle Uf, Uf\rangle = \langle f, f\rangle$ and $\langle Ux, U(af)\rangle = \langle x, af\rangle$.

$$\begin{aligned} \langle Ux, U(af)\rangle &= \langle x, af\rangle = \langle af + bg, af\rangle = \langle af, af\rangle = |a|^2 = |a'|^2, \\ \langle Ux, U(af)\rangle &= p_f(a)\langle Ux, Uf\rangle = p_f(a)\langle a'Uf + b'Ug, Uf\rangle \\ &= p_f(a)a'^*\langle Uf, Uf\rangle = p_f(a)a'^*\langle f, f\rangle = p_f(a)a'^*. \end{aligned} \tag{19}$$

So $a'^*a' = |a'|^2 = p_f(a)a'^*$. Since $|a'| = |a|$, the assumption that $a \neq 0$ implies that we can cancel a'^* from this equality to get $p_f(a) = a'$. A similar argument shows that $p_g(b) = b'$. So

$$U(af + bg) = p_f(a)Uf + p_g(b)Ug. \tag{20}$$

□

Lemma 1.11 (All the p_y such that $y \perp e$ are the same)**.** *There's a $\theta \in \{I, I^*\}$ such that $p_y = \theta$ for all $y \in \{e\}^\perp$ such that $y \neq 0$.*

Proof. We will prove the theorem by proving the following statements:

(a) For all $f \in \{e\}^\perp$ such that $\|f\| = 1$, and all $a \in \mathbb{C}$, we have $p_y(a) \in \{a, a^*\}$.

(b) For all $f \in \{e\}^\perp$ such that $\|f\| = 1$, $p_f \in \{I, I^*\}$.

(c) For all $f, g \in \{e^\perp\}$ such that $\|f\| = \|g\| = 1$ and $\langle f, g\rangle = 0$, $p_f = p_g$.

(d) For all $f, g \in \{e^\perp\}$ such that $\|f\| = \|g\| = 1$ and $\langle f, g\rangle \neq 0$, $p_f = p_g$.

(e) For all $y \in \{e\}^\perp - \{0\}$, we have $p_y = p_f$, where f is defined by $f = y/\|y\|$.

These results imply that p_y is the same function for all $y \in \{e\}^\perp$ such that $y \neq 0$, and that this function is either I or I^*. If this function is denoted by θ, we have $U(ay) = a^\theta Uy$ for all $a \in \mathbb{C}$ and all $y \in \{e\}^\perp$.

(a): Lemma 1.9(e) tells us that $\langle Uf, U(af)\rangle = \langle f, af\rangle = a$ or $\langle Uf, U(af)\rangle = \langle f, af\rangle^* = a^*$. But we also have $\langle Uf, U(af)\rangle = p_f(a)\langle Uf, Uf\rangle = p_f(a)$.

(That last equality follows from lemma 1.9(d), because $\langle f, f\rangle = \|f\| = 1 \in \mathbb{R}$).

(b): Define $A_f = \{a \in \mathbb{C} | p_f(a) = a\}$ and $B_f = \{a \in \mathbb{C} | p_f(a) = a^*\}$. We will prove that one of these sets is equal to $\mathbb{C}$. Note that $A_f \cap B_f = \mathbb{R}$. Let $a \in A_f$ and $b_f \in B$ be arbitrary. 1.9(c) tells us that

$$\begin{aligned} \operatorname{Re}\langle U(af), U(bf)\rangle &= \operatorname{Re}\langle af, bf\rangle = \operatorname{Re}(a^* b), \\ \operatorname{Re}\langle U(af), U(bf)\rangle &= \operatorname{Re}\left(p_f(a)^* p_f(b)\langle Uf, Uf\rangle\right) = \operatorname{Re}(a^* b^*). \end{aligned} \tag{21}$$

So $\operatorname{Re}(a^*b) = \operatorname{Re}(a^*b^*)$. Lemma 1.8(b) tells us that this implies that $(\operatorname{Im} a)(\operatorname{Im} b) = 0$. So one of a and b is real. If a is real, then $a \in \mathbb{R} = A_f \cap B_f \subset B_f$. Since a is an arbitrary member of A_f, this implies that $A_f \subset B_f$. Similarly, if b is real, then $B_f \subset A_f$. So one of these sets is a subset of the other. Since their union is $\mathbb{C}$, this implies that the larger of the two sets is $= \mathbb{C}$. If $A_f = \mathbb{C}$, then $p_f = I$. If $B_f = \mathbb{C}$, then $p_f = I^*$.

(c): Since $0 = U0 = U(0f) = p_f(0)Uf$, and $\|Uf\| = \|f\| = 1$, we have $p_f(0) = 0$. Similarly, $p_g(0) = 0$. So $p_f(0) = p_g(0)$.

Let $a \in C - \{0\}$ be arbitrary. Lemma 1.10 (applied to the linear combinations $f + g$ and $af + ag$) implies that

$$\begin{aligned} U(a(f+g)) &= p_{f+g}(a)U(f+g) = p_{f+g}(a)Uf + p_{f+g}(a)Ug, \\ U(a(f+g)) &= U(af + ag) = p_f(a)Uf + p_g(a)Ug. \end{aligned} \tag{22}$$

This implies that

$$(p_{f+g}(a) - p_f(a))Uf + (p_{f+g}(a) - p_f(a))Ug = 0. \tag{23}$$

Uf and Ug are orthogonal (lemma 1.9), and therefore linearly independent. So this implies that $p_f(a) = p_{f+g}(a) = p_g(a)$. Since a is an arbitrary non-zero complex number, and we have already proved that $p_f(0) = p_g(0)$, this implies that $p_f = p_g$.

(d): Suppose that $p_f \neq p_g$. Then either $p_f = I$ and $p_g = I^*$, or $p_f = I^*$ and $p_g = I$. Suppose that $p_f = I^*$ and $p_g = I$. (The other possibility can be dealt with by swapping f for g and g for f in the argument we're about to make). Let a be an arbitrary complex number such that $\operatorname{Im} a \neq 0$.

$$\begin{aligned} \operatorname{Re}\langle U(af), U(ag)\rangle &= \operatorname{Re}\langle af, ag\rangle = |a|^2 \operatorname{Re}\langle f, g\rangle = |a|^2 \operatorname{Re}\langle Uf, Ug\rangle \\ &= \operatorname{Re}\left(|a|^2\langle Uf, Ug\rangle\right), \\ \operatorname{Re}\langle U(af), U(ag)\rangle &= \operatorname{Re}\langle p_f(a)Uf, p_g(a)Ug\rangle = \operatorname{Re}\left(a^2\langle Uf, Ug\rangle\right) \end{aligned} \tag{24}$$

Since $\operatorname{Im} a \neq 0$, lemma 1.8(c) tells us that this implies that

$$(\operatorname{Re} a)\big(\operatorname{Im}\langle Uf, Ug\rangle\big) = -(\operatorname{Im} a)\big(\operatorname{Re}\langle Uf, Ug\rangle\big). \tag{25}$$

We have proved that this equality holds for all $a \in \mathbb{C}$ such that $\operatorname{Im} a \neq 0$. If we can prove that this result is actually false, this will disprove the assumption that $p_f \neq p_g$. It's sufficient to prove that there's an $a \in \mathbb{C}$ that doesn't satisfy (25) and has a non-zero imaginary part.

Lemma 1.9(d) and the assumption that $\langle f, g\rangle \neq 0$ imply that $\langle Uf, Ug\rangle \neq 0$. If $\operatorname{Re}\langle Uf, Ug\rangle \neq 0$, then (25) implies that

$$\operatorname{Im} a = -\frac{\operatorname{Im}\langle Uf, Ug\rangle}{\operatorname{Re}\langle Uf, Ug\rangle} \operatorname{Re} a. \tag{26}$$

So (25) is false for all a that don't satsify this condition, for example $1 - it$, where t is any real number such that $t \neq \frac{\operatorname{Im}\langle Uf,Ug\rangle}{\operatorname{Re}\langle Uf,Ug\rangle}$ and $t \neq 0$. If $\operatorname{Im}\langle Uf, Ug\rangle \neq 0$, then (25) implies that

$$\operatorname{Re} a = -\frac{\operatorname{Re}\langle Uf, Ug\rangle}{\operatorname{Im}\langle Uf, Ug\rangle} \operatorname{Im} a. \tag{27}$$

So (25) is false for all a that don't satisfy this condition, for example $t - i$, where t is any real number such that $t \neq \frac{\operatorname{Im}\langle Uf,Ug\rangle}{\operatorname{Re}\langle Uf,Ug\rangle}$. So for all possible values of $\langle Uf, Ug\rangle$, there's an $a \in \mathbb{C}$ that doesn't satisfy (25) and has a non-zero imaginary part. This contradicts the assumption $p_f \neq p_g$, so $p_f = p_g$.

(e): We will prove that $p_f(a) = p_f(a)p_f(\|y\|)$. If $p_f = I$, then $p_f(a\|y\|) = a\|y\| = p_f(a)p_f(\|y\|)$. If $p_f = I^*$, then $p_f(a\|y\|) = a^*\|y\| = p_f(a)p_f(\|y\|)$.

$$\begin{aligned}
U(ay) &= p_y(a)Uy,\\
U(ay) &= U(a\|y\|f) = p_f(a\|y\|)Uf = p_f(a)p_f(\|y\|)Uf\\
&= p_f(a)U(\|y\|f) = p_f(a)Uy.
\end{aligned} \tag{28}$$

This implies that $(p_y(a) - p_f(a))Uy = 0$. If $p_y \neq p(a)$, we can multiply this by $1/(p_y(a) - p_f(a))$ to get $Uy = 0$. This contradicts the result $\|Uy\| = \|y\|$ (which is implied by lemma 1.8(c)) or the assumption that $y \neq 0$. So $p_y(a) = p_f(a)$. Since a is an arbitrary complex number, this implies that $p_y = p_f$. □

Let θ be the unique member of $\{I, I^*\}$ such that $p_y = \theta$ for all $y \in \{e\}^\perp$ such that $y \neq 0$.

Lemma 1.12 (U is θ-linear and θ-unitary on $\{e\}^\perp$)**.** *(a) For all $a, b \in \mathbb{C}$ and all $y, z \in \{e\}^\perp$, $U(ay + bz) = a^\theta Uy + b^\theta Uz$.*

(b) For all $y, z \in \{e\}^\perp$, $\langle Uy, Uz\rangle = \langle y, z\rangle^\theta$.

Proof. Let $a, b \in \mathbb{C}$ be arbitrary. Let $y, z \in \{e\}^\perp$ be arbitrary. If $y = 0$ or $z = 0$, we clearly have $U(ay + bz) = a^\theta Uy + b^\theta Uz$. So suppose that $y, z \neq 0$.

These two vectors span a subspace of $\{e\}^\perp$ that's either 1-dimensional or 2-dimensional. Suppose that it's 1-dimensional. Define f by $f = \frac{y}{\|y\|}$. We have $y = \|y\|f$, and $z = \|z\|f$ or $z = -\|z\|f$. We will use the $\pm$ notation to deal with both cases at once.

$$\begin{aligned} U(ay+bz) &= U\big((a\|y\| \pm b\|z\|)f\big) = (a\|y\| \pm b\|z\|)^\theta Uf \\ &= a^\theta \|y\| Uf \pm b^\theta \|z\| Uf = a^\theta U(\|y\|f) + b^\theta U(\pm\|z\|f) \\ &= a^\theta Uy + b^\theta Uz. \end{aligned} \tag{29}$$

Since $\langle y,z\rangle = \big\langle \|y\|f, \pm\|z\|f\big\rangle = \pm\|y\|\,\|z\| \in \mathbb{R}$, lemma 1.9(e) tells us that $\langle Uy,Uz\rangle = \langle y,z\rangle \in \mathbb{R}$. Since $r^\theta = r$ for all $r \in \mathbb{R}$, this implies that $\langle Uy,Uz\rangle = \langle y,z\rangle^\theta$.

Now suppose that the subspace spanned by $\{y,z\}$ is 2-dimensional. Define f by $f = \frac{f}{\|f\|}$, and let g be an arbitrary member of $\{e\}^\perp$ such that $\{f,g\}$ is an orthonormal basis for that subspace. Let (c,d) be the unique pair of complex numbers such that $z = cf + dg$. Lemma 1.10 tells us that

$$\begin{aligned} U(ay+bz) &= U\big(a\|y\|f + b(cf+dg)\big) = U\big((a\|y\|+bc)f + dg\big) \\ &= (a\|y\|+bc)^\theta Uf + (bd)^\theta Ug = a^\theta\|y\|Uf + b^\theta(c^\theta Uf + d^\theta Ug) \\ &= a^\theta U(\|y\|f) + b^\theta U(cf+dg) = a^\theta Uy + b^\theta Uz. \end{aligned} \tag{30}$$

Since $\langle y,z\rangle = \langle \|y\|f, cf+dg\rangle = \|y\|c$, lemma 1.10 tells us that

$$\begin{aligned} \langle Uy,Uz\rangle &= \big\langle U(\|y\|f), U(cf+dg)\big\rangle = \big\langle \|y\|Uf, c^\theta Uf + d^\theta Ug\big\rangle \\ &= \|y\|c^\theta = (\|y\|c)^\theta = \langle y,z\rangle^\theta. \end{aligned} \tag{31}$$

□

Definition 1.13 (Extension of U to all of $\mathcal{H}$)**.** Let $U : \mathcal{D} \to \mathcal{H}$ be the map defined by definition 1.7. Let θ be the unique member of $\{I, I^*\}$ such that $U(ay) = a^\theta Uy$ for all $a \in \mathbb{C}$ and all $y \in \{e\}^\perp$. For each $x \in \mathcal{H} - \mathcal{D}$, define Ux by

$$Ux = \langle e,x\rangle^\theta U\left(\frac{x}{\langle e,x\rangle}\right). \tag{32}$$

Lemma 1.14 (U is θ-linear and θ-unitary)**.** *Let $U : \mathcal{H} \to \mathcal{H}$ be the map defined by definition 1.13. Let θ be the unique member of $\{I, I^*\}$ such that $U(ay) = a^\theta Uy$ for all $a \in \mathbb{C}$ and all $y \in \{e\}^\perp$.*

(a) For all $x \in \mathcal{H}$ such that $x \neq 0$, $Ux \in T[x]$.

(b) For all $a \in \mathbb{C}$ and all $x \in \mathcal{H}$, $U(ax) = a^\theta Ux$.

(c) For all $x, y \in \mathcal{H}$, $U(x + y) = Ux + Uy$.

(d) For all $x, y \in \mathcal{H}$, $\langle Ux, Uy\rangle = \langle x, y\rangle^\theta$.

Proof. (a): Lemma 1.9 tells us that $Ux \in T[x]$ for all $x \in \mathcal{D}$. Definition 1.13 and lemma 1.3(a) tell us that for all $x \in \mathcal{H} - \mathcal{D}$,

$$\begin{aligned}
[Ux] &= \left[\langle e, x\rangle^\theta U\left(\frac{x}{\langle e, x\rangle}\right)\right] = \langle e, x\rangle^\theta \left[U\left(\frac{x}{\langle e, x\rangle}\right)\right] \\
&= \langle e, x\rangle^\theta T\left[\frac{x}{\langle e, x\rangle}\right] = T\left[\frac{\langle e, x\rangle^\theta}{\langle e, x\rangle}x\right] = T[x]. \qquad (33)
\end{aligned}$$

(b): Let $a \in \mathbb{C}$ and $x, y \in \mathcal{H}$ be arbitrary. We will prove that $U(ax) = a^\theta Ux$. If $a = 0$ or $x \in \{e\}^\perp$, we have already done that. So suppose that $a \neq 0$ and $x \notin \{e\}^\perp$. The latter assumption implies that $\langle e, x\rangle \neq 0$. Since $\langle e, ax\rangle = a\langle e, x\rangle$ and $a \neq 0$, this implies that $\langle e, ax\rangle \neq 0$. So

$$U(ax) = \langle e, ax\rangle^\theta U\left(\frac{ax}{\langle e, ax\rangle}\right) = a^\theta \langle e, x\rangle^\theta U\left(\frac{x}{\langle e, x\rangle}\right) = a^\theta Ux. \qquad (34)$$

(c) and (d): Let $x, y \in \mathcal{H}$ be arbitrary. The projection theorem tells us that there's a unique pair (a, b) of complex numbers and a unique pair $(x_\perp, y_\perp)$ of vectors in $\{e\}^\perp$, such that $x = ae + x_\perp$ and $y = be + y_\perp$. Define f and g respectively by $f = \frac{f}{\|f\|}$ and $g = \frac{g}{\|g\|}$. We have

$$\begin{aligned}
x &= ae + \|x_\perp\| f, \\
y &= be + \|y_\perp\| g. \qquad (35)
\end{aligned}$$

Recall that for all $z \in \{e\}^\perp$, Uz is defined by $Uz = U(e + z) - Ue$. So for all $z \in \{e\}^\perp$, $U(e + z) = Ue + Uz$. This result, part (a), lemma 1.10 and lemma 1.12 imply that

$$\begin{aligned}
U(x + y) &= U\big((a + b)e + \|x_\perp\| f + \|y_\perp\| g\big) = U\left((a + b)\left(e + \frac{\|x_\perp\| f + \|y_\perp\| g}{a + b}\right)\right) \\
&= (a + b)^\theta U\left(e + \frac{\|x_\perp\| f + \|y_\perp\| g}{a + b}\right) = (a + b)^\theta\left(Ue + U\left(\frac{\|x_\perp\| f + \|y_\perp\| g}{a + b}\right)\right) \\
&= (a + b)^\theta Ue + U\big(\|x_\perp\| f + \|y_\perp\| g\big) = a^\theta Ue + b^\theta Ue + \|x_\perp\| Uf + \|y_\perp\| Ug \\
&= a^\theta\left(Ue + \frac{\|x_\perp\|}{a^\theta} Uf\right) + b^\theta\left(Ue + \frac{\|y_\perp\|}{b^\theta} Ug\right) \\
&= a^\theta U\left(e + \frac{\|x_\perp\|}{a^\theta} f\right) + b^\theta U\left(e + \frac{\|y_\perp\|}{b^\theta} Ug\right) \\
&= U(e + \|x_\perp\| f) + U(e + \|y_\perp\| g) = Ux + Uy. \qquad (36)
\end{aligned}$$

We have proved (c). To prove (d), first note that $\langle x,y\rangle = \langle ae+\|x_\perp\|f, be+\|y_\perp\|g\rangle = a^*b$, and that lemma 1.5 implies that $\{Ue,Uf,Ug\}$ is an orthonormal set. These results and part (c) imply that

$$\begin{aligned}\langle Ux,Uy\rangle &= \big\langle U(ae+\|x_\perp\|f), U(be+\|y_\perp\|g)\big\rangle\\ &= \big\langle a^\theta Ue+\|x_\perp\|Uf, b^\theta Ue+\|y_\perp\|Ug\big\rangle\\ &= (a^*b)^\theta = \langle x,y\rangle^\theta.\end{aligned}\tag{37}$$

□

Lemma 1.15 (U is unique up to a phase factor)**.** *Let $U:\mathcal{H}\to\mathcal{H}$ be the map defined by definition 1.13. Let θ be the unique member of $\{I,I^*\}$ such that U is θ-unitary. Let $V:\mathcal{H}\to\mathcal{H}$ and $\eta\in\{I,I^*\}$ be arbitrary.*

(a) If $\dim\mathcal{H}\geq 2$, $\eta\in\{I,I^*\}$, V is η-unitary, and $Vx\in T[x]$ for all $x\in\mathcal{H}-\{0\}$, then $\eta=\theta$ and there's a $\lambda\in\mathbb{C}$ such that $|\lambda|=1$ and $V=\lambda U$.

(b) For all $\lambda\in\mathbb{C}$ such that $|\lambda|=1$, λU is θ-unitary and $(\lambda U)x\in T[x]$ for all $x\in\mathcal{H}-\{0\}$.

Proof. (a): The assumption that V is η-unitary implies that $V0=V(0\cdot 0)=0^\eta V(0)=0$. Since we also assumed that $Vx\in T[x]$ for all $x\in\mathcal{H}$ such that $x\neq 0$, V satisfies the assumptions of lemma 1.5.

We will prove that V is injective. Let x and y be arbitrary vectors in $\mathcal{H}$ such that $Vx=Vy$. We have $V(x-y)=0$, and therefore

$$0=\langle V(x-y),V(x-y)\rangle=\langle x-y,x-y\rangle^\eta=\|x-y\|^2.\tag{38}$$

This implies that $x=y$. So V is injective.

Let $\{x,y\}$ be an arbitrary linearly independent set in $\mathcal{H}$. We will prove that $\{Vx,Vy\}$ is linearly independent. Suppose that a and b are complex numbers such that $aVx+bVy=0$. Then $V(a^\eta x+b^\eta y)=0$. Since V is injective, this implies that $a^\eta x+b^\eta y=0$. Since $\{x,y\}$ is linearly independent, this implies that $a=b=0$. So $\{Vx,Vy\}$ is linearly independent.

For all $x\in\mathcal{H}-\{0\}$, we have $[Vx]=T[x]=[Ux]$. This implies that there exists a function $c:\mathcal{H}\to\mathbb{C}$ such that for all $x\in\mathcal{H}$, $|c(x)|=1$ and $Vx=c(x)Ux$. We will prove that c is a constant function.

$$\begin{aligned}V(x+y)&=c(x+y)Ux+Uy=c(x+y)Ux+c(x+y)Uy,\\ V(x+y)&=Vx+Vy=c(x)Ux+c(y)Vy.\end{aligned}\tag{39}$$

This implies that $\big(c(x+y)-c(x)\big)Ux+\big(c(x+y)-c(y)\big)Uy$. Since $\{Ux,Uy\}$ is linearly independent, this implies that $c(x)=c(x+y)=c(y)$. Since the right-hand side is independent of x, c must be a constant function.

We will prove that $\eta = \theta$ by deriving a contradiction from the assumption that this is false. So suppose that $\eta \neq \theta$. This means that either $\theta = I$ and $\eta = I^*$ or $\theta = I^*$ and $\eta = I$. Suppose that $\theta = I$ and $\eta = I^*$. (The other possibility can be dealt with by making a few obvious changes in the argument we're about to make). Let $x \in \mathcal{H} - \{0\}$ be arbitrary.

$$V(ix) = c(ix)U(ix) = c(ix)iUx, \tag{40}$$

$$V(ix) = i^* Vx = -ic(x)Ux. \tag{41}$$

Since $\|Ux\|^2 = \|x\|^2 \neq 0$, $Ux \neq 0$. So the above implies that for all $x \in \mathcal{H} - \{0\}$, $c(ix) = -c(x)$. Since c is a constant function, this implies that $c(x) = 0$. So $Vx = c(x)Ux = 0$. This implies that $0 = \|Vx\|^2 = \langle Vx, Vx\rangle = \langle x, x\rangle^\eta = \|x\|^2 \neq 0$, which is obviously false.

(b): Define $V = \lambda U$. For all $x \in \mathcal{H}$ such that $x \neq 0$,

$$[Vx] = [(\lambda U)x] = [\lambda(Ux)] = \lambda[Ux] = \lambda T[x] = T[\lambda x] = T[x]. \tag{42}$$

The last equality follows from the assumption that $|\lambda| = 1$. This result implies that $Vx \in T[x]$ for all $x \in \mathcal{H} - \{0\}$.

For all $a, b \in \mathbb{C}$ and all $x, y \in \mathcal{H}$,

$$\begin{aligned} V(ax + by) &= \lambda U(ax + by) = \lambda(a^\theta Ux + b^\theta Uy) = a^\theta Vx + b^\theta Vy, \\ \langle Vx, Vy\rangle &= \langle \lambda Ux, \lambda Uy\rangle = |\lambda|^2 \langle Ux, Uy\rangle = \langle x, y\rangle^\theta. \end{aligned} \tag{43}$$

□