SOME NOTES ON THE SPECTRAL THEOREM

1. Introduction

These are notes for an introductory graduate algebra class explaining a spectral theorem: a self adjoint operator on a real inner product space or a normal operator on a complex Hermitian space is always diagonalizable and the eigenspaces are orthogonal.

2. Inner Product Spaces and Hermitian Spaces

The theory has one formulation in terms of real inner product spaces and one in terms of complex Hermitian spaces. In fact the real case is naturally encompassed by the complex case. But since the real case is so often used, we will give both formulations starting with the real case.

Let V be a real vector space. A real pairing $\langle \cdot, \cdot \rangle$ on V,

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}, \quad (\vec{v}, \vec{w}) \mapsto \langle \vec{v}, \vec{w} \rangle$$

is *symmetric* if for every $\vec{v}, \vec{w}$ in V,

$$\langle \vec{w}, \vec{v} \rangle = \langle \vec{v}, \vec{w} \rangle.$$

A pairing which is symmetric turns out to be *bilinear* if and only if for every $\vec{v}_1, \vec{v}_2, \vec{w}$ in V and for every scalar c in $\mathbb{R}$,

$$\langle c\vec{v}_1 + \vec{v}_2, \vec{w} \rangle = c \cdot \langle \vec{v}_1, \vec{w} \rangle + \langle \vec{v}_2, \vec{w} \rangle.$$

For every $\vec{v}$ in V, define $\|\vec{v}\|^2$ by

$$\|\vec{v}\|^2 := \langle \vec{v}, \vec{v} \rangle.$$

A symmetric, bilinear pairing is *positive definite* if for every nonzero $\vec{v}$ in V, $\|\vec{v}\|^2$ is a positive real number. In this case define $\|\vec{v}\|$ by

$$\|\vec{v}\| := \sqrt{\|\vec{v}\|^2} = \sqrt{\langle \vec{v}, \vec{v} \rangle}.$$

Definition 2.1. A *real inner product space* is a pair $(V, \langle \cdot, \cdot \rangle)$ of a real vector space V and a symmetric, bilinear, positive definite pairing $\langle \cdot, \cdot \rangle$ on V.

Next let V be a $\mathbb{C}$-vector space. A complex pairing $\langle \cdot, \cdot \rangle$ on V,

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}, \quad (\vec{v}, \vec{w}) \mapsto \langle \vec{v}, \vec{w} \rangle$$

is *conjugate symmetric* if for every $\vec{v}, \vec{w}$ in V,

$$\overline{\langle \vec{w}, \vec{v} \rangle} = \langle \vec{v}, \vec{w} \rangle.$$

A pairing which is conjugate symmetric turns out to be *sesquilinear* (or "half-linear") if and only if for every $\vec{v}_1, \vec{v}_2, \vec{w}$ in V and for every scalar c in $\mathbb{R}$,

$$\langle c\vec{v}_1 + \vec{v}_2, \vec{w} \rangle = c \cdot \langle \vec{v}_1, \vec{w} \rangle + \langle \vec{v}_2, \vec{w} \rangle.$$

Date: November 5, 2009.

As above define $\|\vec{v}\|^2$ by

$$\|\vec{v}\|^2 := \langle \vec{v}, \vec{v} \rangle.$$

A conjugate symmetric, sesquilinear pairing is *positive definite* if for every nonzero $\vec{v}$ in V, $\|\vec{v}\|^2$ is a positive real number. In this case define $\|\vec{v}\|$ by

$$\|\vec{v}\| := \sqrt{\|\vec{v}\|^2} = \sqrt{\langle \vec{v}, \vec{v} \rangle}.$$

Definition 2.2. A *complex Hermitian space* is a pair $(V, \langle \cdot, \cdot \rangle)$ of a complex vector space V and a conjugate symmetric, sesquilinear, positive definite pairing $\langle \cdot, \cdot \rangle$ on V.

In both the case of a real inner product space and a complex Hermitian space, the (conjugate) symmetry of the pairing and the additivity of the first argument imply the additivity of the second argument,

$$\langle \vec{v}, \vec{w}_1 + \vec{w}_2 \rangle = \langle \vec{v}, \vec{w}_1 \rangle + \langle \vec{v}, \vec{w}_2 \rangle.$$

And in the case of a real inner product space, symmetry and scaling in the first argument imply scaling in the second argument,

$$\langle \vec{v}, c \cdot \vec{w} \rangle = c \cdot \langle \vec{v}, \vec{w} \rangle.$$

Thus a real symmetric pairing which satisfies the condition above is $\mathbb{R}$-bilinear in the usual sense. However, in the case of a complex Hermitian space, complex conjugation intervenes,

$$\langle \vec{v}, c \cdot \vec{w} \rangle = \overline{c} \langle \vec{v}, \vec{w} \rangle.$$

A pairing which is linear in the first argument but "conjugate linear" in the second is *sesquilinear*. For a conjugate symmetric pairing, linearity in the first argument implies sesquilinearity.

Proposition 2.3 (Cauchy-Bunyakovsky-Schwarz inequality)**.** *For every pair of elements $\vec{v}, \vec{w}$ of V,*

$$|\langle \vec{v}, \vec{w} \rangle|^2 \leq \|\vec{v}\|^2 \cdot \|\vec{w}\|^2$$

and equality holds if and only if $\vec{v}, \vec{w}$ are linearly dependent.

Proof. If $\vec{v}$ and $\vec{w}$ are linearly dependent, i.e., if one is a scalar multiple of the other, then it is straightforward to verify that the inequality is an equality. Similarly, if $\langle \vec{v}, \vec{w} \rangle$ equals 0, then the proposition trivially follows. Therefore assume that $\vec{v}$, $\vec{w}$ are linearly independent and assume that $\langle \vec{v}, \vec{w} \rangle$ is nonzero.

Define c to the be the scalar (complex in the Hermitian case) of unit length defined by

$$c = \frac{|\langle \vec{v}, \vec{w} \rangle|}{\langle \vec{v}, \vec{w} \rangle}.$$

In particular, the scalar $c\langle \vec{v}, \vec{w} \rangle$ is a positive real number. Consider the quadratic polynomial in the real variable t,

$$q(t) = \langle c\vec{v} + t\vec{w}, c\vec{v} + t\vec{w} \rangle.$$

On the one hand, expanding out gives

$$q(t) = \|\vec{w}\|^2 t^2 + 2t|\langle \vec{v}, \vec{w} \rangle| + \|\vec{v}\|^2.$$

Since $\vec{v}$ and $\vec{w}$ are linearly independent, $c\vec{v} + t\vec{w}$ is never the zero vector. Thus, by positive definiteness, $q(t)$ is never 0. Since $q(t)$ has no real roots, the discriminant of the quadratic polynomial is negative, i.e.,

$$(2|\langle \vec{v}, \vec{w}\rangle|)^2 < 4\|\vec{v}\|^2 \cdot \|\vec{w}\|^2.$$

Dividing by 4 gives the inequality. □

Corollary 2.4. *The function* $\|\cdot\| : V \to \mathbb{R}$ *is a* norm, *i.e., it satisfies each of the following.*

(i). Homogeneity. *For every* $\vec{v}$ *in* V *and every scalar* c, $\|c \cdot \vec{v}\|$ *equals* $c \cdot \|\vec{v}\|$.

(ii). Positive definite. *For every nonzero* $\vec{v}$ *in* V, $\|\vec{v}\}$ *is positive.*

(iii). Triangle inequality. *For every* $\vec{v}, \vec{w}$ *in* V,

$$\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|.$$

Proof. The first two items are obvious. The triangle inequality follows from the previous result. By definition,

$$\|\vec{v} + \vec{w}\|^2 = \langle \vec{v}, \vec{v}\rangle + 2\mathrm{Re}(\langle \vec{v}, \vec{w}\rangle) + \langle \vec{w}, \vec{w}\rangle.$$

Applying the Cauchy-Bunyakovsky-Schwarz inequality, this is bounded by

$$\|\vec{v} + \vec{w}\|^2 \leq \|\vec{v}\|^2 + 2\|\vec{v}\|\|\vec{w}\| + \|\vec{w}\|^2 = (|\vec{v}\| + \|\vec{w}\|)^2.$$

Taking square roots gives the result. □

Examples. (1). Let $n \geq 0$ be an integer. Let V be the real vector space $\mathbb{R}^n$ with its standard ordered basis $(\mathbf{e}_1, \dots, \mathbf{e}_n)$. Let $\langle \cdot, \cdot \rangle$ be the Euclidean inner product,

$$\langle \sum_{i=1}^{n} x_i \mathbf{e}_i, \sum_{j=1}^{n} y_j \mathbf{e}_j \rangle = \sum_{k=1}^{n} x_k y_k.$$

This is a real inner product space. It is usually called the real *Euclidean n-space.*

Examples. (2). Similarly, let V be the complex vector space $\mathbb{C}^n$ with its standard ordered basis $(\mathbf{e}_1, \dots, \mathbf{e}_n)$. Let $\langle \cdot, \cdot \rangle$ be the pairing

$$\langle \sum_{i=1}^{n} x_i \mathbf{e}_i, \sum_{j=1}^{n} y_j \mathbf{e}_j \rangle = \sum_{k=1}^{n} x_k \overline{y_k}.$$

This is a complex Hermitian space. It is sometimes also called a Euclidean n-space.

Examples. (3). Let $\ell^2_{\mathbb{R}}$, respectively $\ell^2_{\mathbb{C}}$, denote the real vector space, resp. complex vector space, of all sequences $(a_n)_{n\geq 0}$ of real numbers, resp. complex numbers, such that the series

$$\sum_{n=0}^{\infty} |a_n|^2$$

is convergent. Define a pairing by the series

$$\langle (a_n)_{n\geq 0}, (b_n)_{n\geq 0}\rangle = \sum_{n=0}^{\infty} a_n \overline{b_n}.$$

Applying the Cauchy-Bunyakovsky-Scwarz inequality to the finite sequences $(a_n)_{n\leq N}$, $(b_n)_{n\leq N}$ considered as elements of the Euclidean N-space, it follows that the series above is bounded by the product of the convergent sequences $\|(a_n)\| \cdot \|(b_n)\|$.

Therefore the series is absolutely convergent. It is straightforward to verify that this defines a real inner product space, resp. a complex Hermitian space.

Examples. (4). Let I be a closed interval in $\mathbb{R}$ (possibly unbounded), and let $C^2_{\mathbb{R}}(I)$, resp. $C^2_{\mathbb{C}}(I)$, denote the real vector space, resp. complex vector space, of continuous functions f on I which are real-valued, resp. complex-valued, and such that the integral

$$\int_I |f(t)|^2 dt$$

is convergent. Notice the convergence is automatic when I is a bounded interval. Define an inner product by

$$\langle f, g \rangle = \int_I f(t)\overline{g(t)}dt.$$

For every bounded interval J, the restrictions of f and g to J satisfy the Cauchy-Bunyakovsky-Schwarz inequality, i.e.,

$$\left| \int_J f(t)\overline{g(t)}dt \right|^2 \leq \left| \int_I |f(t)|^2 dt \right| \cdot \left| \int_I |g(t)|^2 dt \right|.$$

Since the product on the right is convergent, it follows the supremum over all J of the integeral on the left is also defined. Thus the inner product of f and g is well-defined. This pairing defines a real inner product space, resp. a complex Hermitian space. (However, it is usually not complete with respect to the corresponding norm.)

Examples. (5). Let $(V, \langle \cdot, \cdot \rangle)$ be a real inner product space. Denote by $V_{\mathbb{C}}$ the associated complex vector space,

$$V_{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C}.$$

Every element in this space has a unique decomposition of the form $\vec{v}_R + i\vec{v}_I$ where $\vec{v}_R, \vec{v}_I$ are elements of V. Define a pairing

$$\langle \cdot, \cdot \rangle : V_{\mathbb{C}} \times V_{\mathbb{C}} \to \mathbb{C},$$

by the rule

$$\langle \vec{v}_R + i\vec{v}_I, \vec{w}_R + i\vec{w}_I \rangle := (\langle \vec{v}_R, \vec{w}_R \rangle + \langle \vec{v}_I, \vec{w}_I \rangle) + i(\langle \vec{v}_I, \vec{w}_R \rangle - \langle \vec{v}_R, \vec{w}_I \rangle).$$

It is straightforward to verify that this is a Hermitian product. Thus associated to every real inner product space there is a complex Hermitian space. This is the beginning of many parallels between the theory of real inner product spaces and complex Hermitian spaces.

3. Orthonormal Bases and Gram-Schmidt

Let $(V, \langle \cdot, \cdot \rangle)$ be a real inner product space or complex Hermitian space. An element $\vec{v}$ of V is *normal* if $\|v\|$ equals 1. A subset S of elements in V is *orthogonal* if for every pair $\vec{v}, \vec{w}$ of distinct elements in S, $\langle \vec{v}, \vec{w} \rangle$ equals 0. An orthogonal subset of normal vectors is an *orthonormal subset*. When $(V, \langle \cdot, \cdot \rangle)$ is a complex Hermitian space, the word "unitary" is sometimes also used to mean "orthonormal". Every orthonormal set is linearly independent.

A collection $(U_i)_{i \in I}$ of subspaces of V is *orthogonal* if for every pair i, j of distinct elements of I, every element in U_i is orthogonal to every element in U_j. Given any subspace U of V, the set of all vectors $\vec{v}$ in V which are orthogonal to every element of U is called the *orthogonal complement* of U and denote $U^{\perp}$.

Lemma 3.1. *Every collection of orthogonal subspaces is linearly independent. In particular, for every linear subspace U of V, U and $U^{\perp}$ are linearly independent.*

Proof. The definition of linear independence involves finite linear combinations. Thus it suffices to verify the first assertion for finite collections of orthogonal subspaces. This is proved by induction on the number of subspaces in the collection. The main case is when there are two subspaces, say U' and U''. And U', U'' are linearly independent if and only if $U' \cap U''$ is $\{0\}$. Let $\vec{u}$ be an element in $U' \cap U''$. Since $\vec{u}$ is in U' it is orthogonal to every element of U''. Since also $\vec{u}$ is in U'', $\vec{u}$ is orthogonal to itself, i.e., $\|\vec{u}\|^2$ equals 0. By positive definiteness, $\vec{u}$ equals 0. Therefore U', U'' are linearly independent.

Now suppose that $n \geq 3$ and let $(U_1, \ldots, U_n)$ be an orthogonal collection of vectors. Denote by U' the sum $U_1 + \cdots + U_{n-1}$ and denote by U'' the space U_n. By the induction hypothesis, $(U_1, \ldots, U_n)$ is linearly independent. Thus to prove that $(U_1, \ldots, U_{n-1}, U_n)$ is linearly independent, it suffices to prove that U' and U'' are linearly independent. Since $(U_1, \ldots, U_n)$ is orthogonal, U' and U'' are orthogonal. Therefore, by the last paragraph, U' and U'' are linearly independent. $\square$

Proposition 3.2 (Gram-Schmidt)**.** *Let $(\vec{v}_1, \ldots, \vec{v}_n)$ be a linearly independent set of elements in V. There exists a unique orthonormal set $(\hat{u}_1, \ldots, \hat{u}_n)$ such that for every $r = 1, \ldots, n$, span$(\vec{v}_1, \ldots, \vec{v}_r)$ equals span$(\hat{u}_1, \ldots, \hat{u}_r)$ and $\langle \vec{v}_r, \hat{u}_r \rangle$ is a positive real number.*

Proof. This is proved by induction on n. For $n = 1$, $\hat{u}_1 := \vec{v}_1 / \|\vec{v}_1\|$ is the unique normal vector in $\text{span}(\vec{v}_1)$ such that $\langle \vec{v}_1, \hat{u}_1 \rangle$ is a positive real number. Thus, by way of induction, assume that $n > 1$ and assume the result is proved for $n - 1$.

Denote $\text{span}(\vec{v}_1, \ldots, \vec{v}_n)$ by U. Denote by U' the subspace $\text{span}(\vec{v}_1, \ldots, \vec{v}_{n-1})$. And denote by U'' the orthogonal complement of U' in U. Since $\vec{v}_n$ is linearly independent from $\vec{v}_1, \ldots, \vec{v}_{n-1}$, $\vec{v}_n$ is not contained in U'.

By the induction hypothesis, there exists an orthonormal set $(\hat{u}_1, \ldots, \hat{u}_{n-1})$ satisfying the conditions of the proposition for $r = 1, \ldots, n-1$. Denote

$$\vec{v}'_n = \sum_{m=1}^{n-1} \langle \vec{v}_n, \hat{u}_m \rangle \hat{u}_m$$

and denote $\vec{v}''_n := \vec{v}_n - \vec{v}'_n$. By construction $\vec{v}'_n$ is in U'. For every $i = 1, \ldots, n-1$,

$$\langle \vec{v}'_n, \hat{u}_i \rangle = \langle \vec{v}_n, \hat{u}_i \rangle.$$

Therefore $\langle \vec{v}''_n, \hat{u}_i \rangle$ equals 0, i.e., $\vec{v}''_n$ is in U''. By the previous lemma U' and U'' are linearly independent. Thus $\vec{v}'_n$ and $\vec{v}''_n$ are the unique elements in U' and U'' such that $\vec{v}_n$ equals $\vec{v}'_n + \vec{v}''_n$. In particular, note that

$$\langle \vec{v}_n, \vec{v}''_n \rangle = \|\vec{v}''_n\|^2$$

which is a nonnegative real number.

In fact, since $\vec{v}_n$ is not contained in U', also $\vec{v}''_n$ is not contained in U', i.e., $\vec{v}''_n$ is nonzero. So the nonnegative real number above is a positive real number. Also, since U' and U'' are linearly independent and since $U' + U''$ is contained in U,

$\dim(U') + \dim(U'')$ is at most $\dim(U)$. Thus $\dim(U'')$ is at most 1. Therefore $\vec{v}''_n$ is a basis for U''. Since any vector $\hat{u}_n$ satisfying the conditions of the proposition lies in U'', $\hat{u}_n$ must be a scalar multiple of $\vec{v}''_n$. But since $\hat{u}_n$ must be normal and since $lr\vec{v}_n, \hat{u}_n$ must be a positive real number, there is no choice about the scalar multiple. The only scalar multiple satisfying these conditions is

$$\hat{u}_n := \frac{\vec{v}''_n}{\|\vec{v}''_n\|}.$$

Since $(\hat{u}_1, \dots, \hat{u}_{n-1})$ is a basis for U' and since $\hat{u}_n$ is a basis for U'', $(\hat{u}_1, \dots, \hat{u}_{n-1}, \hat{u}_n)$ is a linearly independent set of n elements in the n-dimensional space U. Therefore this is a basis for U.

Since $\vec{v}''_n$ is a normal vector which is orthogonal to U', and since $(\hat{u}_1, \dots, \hat{u}_{n-1})$ are orthonormal by the induction hypothesis, also the set $(\hat{u}_1, \dots, \hat{u}_{n-1}, \hat{u}_n)$ is orthonormal. By the induction hypothesis $(\hat{u}_1, \dots, \hat{u}_{n-1})$ is the unique set satisfying the conditions for $i = 1, \dots, n-1$. And, as shown, $\hat{u}_n$ is the unique normal vector in $\mathrm{span}(\vec{v}_1, \dots, \vec{v}_n)$ which is orthogonal to $(\hat{u}_1, \dots, \hat{u}_{n-1})$ and such that $\langle \vec{v}_r, \hat{u}_r \rangle$ is a positive real number. Thus $(\hat{u}_1, \dots, \hat{u}_n)$ is the unique orthonormal set satisfying the conditions of the proposition. Thus the proposition is proved by induction on n. □

In particular, if V is finite dimensional, then it follows that V has a basis which is orthonormal. When V is complex Hermitian, an orthonormal basis is sometimes also called a *unitary frame*.

It also follows from Gram-Schmidt that if V is a finite dimensional vector space, then for every linear subspace U, $(U, U^\perp)$ gives a direct sum decomposition of V. Since by the lemma $(U, U^\perp)$ is linearly independent, it suffices to prove that every vector $\vec{v}$ in V is a sum of elements in U and $U^\perp$. If $\vec{v}$ is in U, this is obvious. If $\vec{v}$ is not in U, first use Gram-Schmidt to produce an orthnormal basis $(\hat{u}_1, \dots, \hat{u}_m)$ for U. Next apply Gram-Schmidt to $(\hat{u}_1, \dots, \hat{u}_m, \vec{v})$ to produce $(\hat{u}_1, \dots, \hat{u}_m, \hat{u}'')$. Then $\hat{u}''$ is in $U^\perp$ and $\vec{v}$ is a sum of a linear combination of $\hat{u}_1, \dots, \hat{u}_m$, which is in U, and a scalar multiple of $\hat{u}''$, which is in $U^\perp$. This proves that $V = U \oplus U^\perp$. Incidentally, this fails when V is infinite dimensional. One characterization of *Hilbert spaces*, i.e., real inner product spaces or complex Hermitian spaces where the metric induced by $\|\cdot\|$ is complete, is that for every *closed* subspace U, the pair $(U, U^\perp)$ does give an orthogonal decomposition of V.

4. Adjoint Operators and Normal Operators

For a real vector space V, resp. a complex vector space V, a *linear functional* on V is a linear transformation from V to $\mathbb{R}$, resp. $\mathbb{C}$. The set of linear functionals is denoted $V^\vee$ and is naturally a real vector space, resp. complex vector space, as discussed in lecture. A linear functional χ is *bounded* (with respect to $\|\cdot\|$) if there exist a positive real number M such that for every $\vec{v}$ in V,

$$|\chi(\vec{v})| \leq M\|\vec{v}\|.$$

In this case, for every nonzero $\vec{v}$ in V, $|\chi(\vec{v})|/\|\vec{v}\|$ is defined and bounded by M. Therefore the supremum of $|\chi(\vec{v})|/\|\vec{v}\|$ over nonzero $\vec{v}$ is a well-defined nonnegative real number which is positive if χ is nonzero. This is the *operator norm* of χ,

$$\|\chi\| = \sup_{\vec{v}\neq 0} \frac{|\chi(\vec{v})|}{\|\vec{v}\|}.$$

In particular, for every vector $\vec{w}$ in V, define $\phi_{\vec{w}}$ to be the linear functional

$$\phi_{\vec{w}}(\vec{v}) = \langle \vec{v}, \vec{w} \rangle.$$

Then the Cauchy-Schwarz inequality implies that $\phi_{\vec{w}}$ is bounded and

$$\|\phi_{\vec{w}}\| = \|\vec{w}\|.$$

In all that follows we will assume that V is finite dimensional. In fact there are infinite dimensional analogues which are even more important than the finite dimensional results. But this is more the topic of an analysis course rather than an algebra course.

Lemma 4.1 (Riesz Representation Theorem)**.** *Assume that V is finite dimensional. Every linear functional on V is bounded. The map $\Phi : V \to V^\vee$ defined by $\Phi(\vec{w}) = \phi_{\vec{w}}$ is an $\mathbb{R}$-linear isomorphism of $\mathbb{R}$-vector spaces. Moreover, $\|\Phi(\vec{w})\|$ equals $\|\vec{w}\|$ so that $\|\cdot\|$ is a norm on $V^\vee$ and Φ is an isometry.*

Proof. As we have seen, when V is finite dimensional then $V^\vee$ is also finite dimensional and has the same dimension as V. If V and $V^\vee$ are complex vector spaces, then they are also both finite dimensional real vector spaces whose real dimension is twice the complex dimension, via the inclusion $\mathbb{R} \subset \mathbb{C}$. It is straightforward that Φ is $\mathbb{R}$-linear. Thus Φ is an $\mathbb{R}$-linear transformation between $\mathbb{R}$-vector spaces of the same finite dimension. It follows from the rank-nullity theorem that Φ is surjective if and only if Φ is an $\mathbb{R}$-linear isomorphism if and only if Φ is injective.

Let $\vec{w}$ be any nonzero element of V. Since $\phi_{\vec{w}}(\vec{w}) = \|\vec{w}\|^2$ and since $\langle \cdot, \cdot \rangle$ is positive definite, $\phi_{\vec{w}}(\vec{w})$ is nonzero. Thus $\phi_{\vec{w}}$ is nonzero. Thus Φ is injective. Therefore Φ is an $\mathbb{R}$-linear isomorphism. A posteriori it follows that every element in $V^\vee$ is bounded (since they are each of the form $\phi_{\vec{w}}$) and that $\|\cdot\|$ is a norm on $V^\vee$ (since $\|\cdot\|$ is a norm on V). ☐

Remarks. **(1).** It is important to remark that when V is a complex Hermitian space, Φ is *not* $\mathbb{C}$-linear. In fact $\Phi(c \cdot \vec{w}) = \overline{c} \cdot \Phi(w)$.

(2). In analysis classes it is proved that if V is complete with respect to the norm $\|\cdot\|$ (which is automatic if V is finite dimensional), then Φ is an isometric $\mathbb{R}$-linear isomorphism from V onto the subspace of bounded linear functionals. This is the usual statement of the Riesz representation theorem.

Now let $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ be finite dimensional real inner product spaces, resp. finite dimensional complex Hermitian spaces. Let $T : V \to W$ be a linear transformation. Recall that the linear transformation $T^\dagger : W^\vee \to V^\vee$ is defined by

$$(T^\dagger(\vec{\chi}))(\vec{v}) = \chi(T(\vec{v})).$$

By the Riesz Representation Theorem there exists a unique linear transformation

$$T^* : W \to V, \text{ such that } \Phi_V \circ T^* = T^\dagger \circ \Phi_W.$$

In other words, for every $\vec{v}$ in V and for every $\vec{w}$ in W,

$$\langle T(\vec{v}), \vec{w} \rangle_W = \langle \vec{v}, T^*(\vec{w}) \rangle_V.$$

The linear transformation T^* is called the *adjoint* of T.

The most important case is when $(W, \langle \cdot, \cdot \rangle_W)$ equals $(V, \langle \cdot, \cdot \rangle_V)$, i.e., T is a linear operator on V. In this case T^* is also a linear operator on V.

Definition 4.2. The linear transformation T is *self-adjoint* if T^* equals T. The linear transformation T is *anti-self-adjoint* if T^* equals $-T$. The linear transformation is *normal* if T^*T equals TT^*, i.e., T commutes with its adjoint. And the linear transformation T is *orthogonal*, resp. *unitary* in the complex case, if T^* and T are inverse operators.

Examples. (1). Let $(V, \langle \cdot, \cdot \rangle)$ be the real Euclidean space $\mathbb{R}^n$. Let T be a linear operator on V. And let M be the matrix representative of T. Then the matrix representative of T^* is $M^\dagger$, the transpose matrix of M, $(M^\dagger)_{i,j} = M_{j,i}$. Thus T is self-adjoint if and only if the associated matrix M is symmetric, $M_{i,j} = M_{j,i}$. Also T is anti-self-adjoint if and only if M is skew-symmetric, $-M_{i,j} = M_{j,i}$. And T is orthogonal if and only if the column vectors of M are an orthonormal basis for $\mathbb{R}^n$. In fact, for every finite dimensional real inner product space V, for every linear operator T and for every ordered orthonormal basis $\mathcal{B} = (\hat{u}_1, \dots, \hat{u}_n)$, T is self-adjoint, resp. anti-self-adjoint, orthogonal, if and only if the matrix representative $[T]_{\mathcal{B},\mathcal{B}}$ is symmetric, resp. skew-symmetric, orthogonal. This is because $\mathcal{B}$ determines a linear, isometric isomorphism between $(V, \langle \cdot, \cdot \rangle)$ and the real Euclidean space.

Examples. (2). Let $(V, \langle \cdot, \cdot \rangle)$ be the complex Euclidean space $\mathbb{C}^n$. Let T be a linear operator on V. And let M be the matrix representative of T. Then the matrix representative of T^* is $\overline{M^\dagger}$, the conjugate transpose matrix of M, $(\overline{M^\dagger})_{i,j} = \overline{M_{j,i}}$. Thus T is self-adjoint if and only if M is "conjugate symmetric", $\overline{M_{i,j}} = M_{j,i}$. Also T is anti-self-adjoing if and only if M is "conjugate skew-symmetric", $-\overline{M_{i,j}} = M_{j,i}$. And T is unitary if and only if the column vectors of M are a unitary frame (i.e., orthonormal basis) for $\mathbb{C}^n$. For the same reason as above, for every finite dimensional complex Hermitian space V, for every linear operator T and for every ordered orthonormal basis $\mathcal{B} = (\hat{u}_1, \dots, \hat{u}_n)$, T is self-adjoing, resp. anti-self-adjoing, unitary, if and only if the matrix representative $[T]_{\mathcal{B},\mathcal{B}}$ is conjugate symmetric, resp. conjugate skew-symmetric, unitary.

The point of normal operators is that self-adjoint, anti-self-adjoint and orthogonal, resp. unitary, operators are all examples of normal operators. So the class of normal operators is the correct one in which to consider all three types of operators. The basic fact about normal operators is the following.

Lemma 4.3. *For a normal operator T, the kernel of T^* equals the kernel of T. And the kernel and image of T are orthogonal subspaces of V.*

Proof. Let $\vec{v}$ be in the kernel of T. Then $T^*T(\vec{v})$ equals 0. Since T is normal, i.e., TT^* equals T^*T, also $TT^*(\vec{v})$ equals 0. Thus the square norm of $T^*(\vec{v})$ equals 0 by

$$\|T^*(\vec{v})\|^2 = \langle T^*(\vec{v}), T^*(\vec{v}) \rangle = \langle TT^*(\vec{v}), \vec{v} \rangle = \langle 0, \vec{v} \rangle = 0.$$

Since $\langle \cdot, \cdot \rangle$ is positive definite, $T^*(\vec{v})$ equals 0. Thus v is in the kernel of T^*. Since $(T^*)^*$ equals T, the same argument proves that the kernel of T is contained in the kernel of T^*. Therefore the kernel of T equals the kernel of T^*.

Now let $\vec{v}$ be in the kernel of T. Then $T^*(\vec{v})$ equals 0. Therefore, for every element $\vec{w}$ of V,

$$\langle \vec{v}, T(\vec{w})\rangle = \langle T^*(\vec{v}), \vec{w}\rangle = \langle 0, \vec{w}\rangle = 0.$$

Thus $\vec{v}$ is orthogonal to every element $T(\vec{w})$ in the image of T. So the kernel of T is orthogonal to the image of T. $\square$

A basic fact about commutativity of operators is that it extends to polynomial expressions in those operators (and in analysis, even to holomorphic expressions on open neighborhoods of the spectra of the operators). For every linear operator T and for every polynomial with coefficients in the scalar field,

$$p(t) = c_n t^n + \cdots + c_1 t + c_0,$$

the linear operator $p(T)$ is defined by

$$p(T) = c_n T^n + \cdots + c_1 T + c_0 \mathrm{Id}_V.$$

Exercise. Let S and T be linear operators that commute, $ST = TS$. Prove that for all nonnegative integers m, n, S^m and T^n commute. Conclude that for all polynomials $p(s)$ and $q(t)$ with coefficients in the scalar field, $p(S)$ and $q(T)$ commute. In fact, for every pair of 2-variable operators $u(s,t)$ and $v(s,t)$, the operators $u(S,T)$ and $v(S,T)$ are well-defined and commute.

Lemma 4.4. *Let T be a normal operator on a complex Hermitian space V. Then for every polynomial $u(s,t)$ with complex coefficients, $u(T*, T)$ is also normal. Also a vector $\vec{v}$ is an eigenvector of T with eigenvalue λ if and only if $\vec{v}$ is an eigenvector of T^* with eigenvalue $\overline{\lambda}$.*

Proof. The first assertion follows immediately from the exercise since $u(T*, T)$ is well-defined and $(u(T*, T))^*$ equals $v(T^*, T)$ where $v(s,t)$ is the complex polynomial $v(s,t) = \overline{v(\overline{t}, \overline{s})}$. Because T is normal, and using the previous lemma, also $T - \lambda\mathrm{Id}$ is normal. And $(T - \lambda\mathrm{Id})^*$ equals $T^* - \overline{\lambda}\mathrm{Id}$. By Lemma 4.3, the kernel of $T - \lambda\mathrm{Id}$ equals the kernel of $(T - \lambda\mathrm{Id})^*$. Thus the λ-eigenspace of T equals the $\overline{\lambda}$-eigenspace of T^*. $\square$

Theorem 4.5 (The Spectral Theorem)**.** *Let T be a normal operator on a finite dimensional complex Hermitian space. Then T is diagonalizable and the eigenspaces are pairwise orthogonal. If T is self-adjoint, then the eigenvalues are all real. Similarly, every self-adjoint operator on a finite dimensional real inner product space is diagonalizable with orthogonal eigenspaces.*

Proof. First we prove that a normal operator on a complex Hermitian space is diagonalizable with orthogonal eigenspaces. For every complex number λ, denote by $E_{T,\lambda}$ the λ-eigenspace, i.e.,

$$E_{T,\lambda} := \ker(\lambda\mathrm{Id}_V - T).$$

By Lemma 4.4, $E_{T,\lambda}$ equals $E_{T^*,\overline{\lambda}}$ and $\lambda\mathrm{Id}_V - T$ is normal. And for every complex number μ with $\mu \neq \lambda$, $\lambda\mathrm{Id}_V - T$ maps $E_{T,\mu}$ isomorphically back to itself via the scaling $(\lambda - \mu)$, i.e., for $\vec{v}$ in $E_{T,\mu}$,

$$(\lambda\mathrm{Id}_V - T)(\vec{v}) = \lambda\vec{v} - T(\vec{v}) = (\lambda - \mu)\vec{v}).$$

Thus $E_{T,\mu}$ is in the image of $\lambda\mathrm{Id}_V - T$. Therefore, by Lemma 4.3, the eigenspaces $(E_{T,\lambda_1}, \dots, E_{T,\lambda_r})$ are orthogonal. So, denoting by $\lambda_1, \dots, \lambda_r$ the distinct eigenvalues of T, $(E_{T,\lambda_1}, \dots, E_{T,\lambda_r})$ is an orthogonal set of subspaces of V. It remains only to prove that the subspace $U'_T = E_{T,\lambda_1} + \dots + E_{T,\lambda_r}$ equals all of V.

Denote by U''_T the orthogonal complement of U'_T. Of course T maps each eigenspace $E_{T,\lambda}$ back into itself via the scaling λ. And by Lemma 4.4, for every $\vec{v}$ in V and every $\vec{w}$ in $E_{T,\lambda}$,

$$\langle T(\vec{v}), \vec{w}\rangle = \langle \vec{v}, T^*(\vec{w})\rangle = \langle \vec{v}, \overline{\lambda}\vec{w}\rangle = \lambda\langle \vec{v}, \vec{w}\rangle.$$

In particular, if $T(\vec{v})$ is orthogonal to $\vec{w}$, then also $T(\vec{v})$ is orthogonal to $\vec{w}$. Therefore T maps U''_T back into itself. Also by Lemma 4.4 once more, U'_{T^*} equals U'_T, and thus also U''_{T^*} equals U''_T. Therefore, by the same argument as above, T^* also maps U''_T back into itself.

But this means that the restriction

$$T|_{U''} : U''_T \to U''_T$$

is a normal operator on the complex Hermitian space U''_T with the restricted Hermitian product. Every linear operator S on a nonzero, finite dimensional complex vector space U has a nonzero eigenvector by the Fundamental Theorem of Algebra applied to the characteristic polynomial

$$\chi_S(x) = \det(x\mathrm{Id}_U - S).$$

Therefore, if U''_T is nonzero, then there exists an eigenvalue λ and a nonzero λ-eigenvector $\vec{v}$ for $T|_{U''}$. Then $\vec{v}$ is in $E_{T,\lambda}$ which is contained in U'_T. Since U'_T and U''_T are orthgonal, their intersection equals $\{0\}$ by Lemma 3.1. So there cannot exist a nonzero vector $\vec{v}$ in U''_T and in $E_{T,\lambda}$. This proves that U''_T does equal $\{0\}$ (by contradiction). Therefore V equals the sum of the T-eigenspaces, i.e., T is diagonalizable on V.

Next assume that T is self-adjoint. Let λ be an eigenvalue with a nonzero λ-eigenvector $\vec{v}$. By Lemma 4.4, $\vec{v}$ is also a $\overline{\lambda}$-eigenvector of T^*. But since $T = T^*$, it follows that $\lambda = \overline{\lambda}$, i.e., λ is real. Therefore every eigenvalue of T is real.

Finally, consider the real case. By Gram-Schmidt, every real inner product space has an orthonormal basis. Thus we may reduce to the case when the real inner product space is the Euclidean space $\mathbb{R}^n$. Let

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

be an $\mathbb{R}$-linear operator. Define a $\mathbb{C}$-linear operator $T_{\mathbb{C}}$ on $\mathbb{C}^n$ by

$$T_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n, \quad T_{\mathbb{C}}(\vec{v}_R + i\vec{v}_I) := T(\vec{v}_R) + iT(\vec{v}_I).$$

It is straightforward to verify that $(T_{\mathbb{C}})^*$ equals $(T^*)_{\mathbb{C}}$. In particular, if T is self-adjoint then also $T_{\mathbb{C}}$ is self-adjoint. Then by the previous case, the $\mathbb{C}^n$ is a direct sum of eigenspace $E_{T_{\mathbb{C}},\lambda}$ for real numbers λ. But now let $\vec{v}_R + i\vec{v}_I$ be any element in $E_{T_{\mathbb{C}},\lambda}$. Then we have the identity

$$T(\vec{v}_R) + iT(\vec{v}_I) = T_{\mathbb{C}}(\vec{v}_R + i\vec{v}_I) = \lambda(\vec{v}_R + i\vec{v}_I) = (\lambda\vec{v}) + i(\lambda\vec{v}_I).$$

It follows that both $\vec{v}_R$ and $\vec{v}_I$ are in $E_{T,\lambda}$, i.e., $E_{T_{\mathbb{C}},\lambda} = E_{T,\lambda} \otimes_{\mathbb{R}} \mathbb{C}$. Therefore the λ-eigenspaces $E_{T,\lambda}$ of T on $\mathbb{R}^n$ also span $\mathbb{R}^n$. Orthogonality of the distinct eigenspaces of $T_{\mathbb{C}}$ implies orthogonality of the corresponding eigenspaces of T. Thus the eigenspaces of T are orthogonal and give a direct sum decomposition of $\mathbb{R}^n$. Therefore T is diagonalizable on $\mathbb{R}^n$. $\square$