THE COMPLEX SPECTRAL THEOREM

MUHIIM ALI

Abstract. The goal of this paper is to present a proof of the Complex Spectral Theorem and its applications.

1. Introduction

The topic of this paper is a fundamental theorem of mathematics: The Spectral Theorem. Because the conclusion of The Spectral Theorem depends on F, it can be broken down into two pieces: The Complex Spectral Theorem and the Real Spectral Theorem. In this paper, we're only going to discuss the Complex Spectral Theorem, present its proof and show areas in which it is applicable.

Theorem A (As stated in [1]). *Complex Spectral Theorem: S Suppose F = C and T ∈ L(V). Then the following are equivalent:*

(a) T is normal
(b) v has an orthonormal basis consisting of eigenvectors of T
(c) T has a diagonal matrix with respect to some orthonormal basis of v.

In a study that was published some time between 1900 and 1910 that looked at integral equations in infinite-dimensional spaces, David Hilbert initially provided the fundamental definition and theorem of spectral theory. Hilbert's original spectral theorem applied to just real quadratic forms (or infinite matrices) that were bounded and symmetric. Erchard Schmidt (1876-1956) and others then extended the theorem to bounded complex matrices A= (a_{ij}), where $a_{ij} = \overline{a}_{ji}$. These matrices are called Hermition named after the french mathematician Charles Hermite (1822-1901). But the major breakthrough occured in 1927-1929, when John Von Neumann (1903-1957), a 25-year old Hungarian introduced linear operators on Hilbert Space. In [1927], he represented the transformation theory of quantum mechanics in terms of operators on a Hilbert space. From this, he noticed the need to extend the theorem from the bounded to the unbouned. (Steen, L.A [3])

The structure of the paper is as follows. In Section 2 we give the definitions of the diagonal matrix, orthonormal bases, orthonormal basis, adjoints and normal operators with examples. We also write important propositions for the theorem and show their proofs. In Section 3, we give the proof of Theorem A. Finally, in Section 4, we give two applications of Theorem A, showing that it can be used in spectral graph theory and in the the second derivative test in multivariable calculus.

Acknowledgements. I want to thank Maddie for helping me understand and write the proofs of the propositions and the main theorem

2. Background

We assume the reader is familiar with the fundamentals of fields, vector spaces, linear maps, Eigenvalues and Eigenvectors, and Inner Product Spaces as in [1, Chapters 1, 2, and 3, 5, 6]. Throughout, F denotes R or C and V is a finite-dimensional, nonzero, inner-product space over F. Before we move to proving the Spectral theorem, we need to familiarize ourselves with the following concepts: Diagonal matrices, Orthonormal bases, Adjoints, and Normal operators.

Definition 2.1. A diagonal matrix is a square that is 0 everywhere except possibly along the diagonal.

It is important to keep in mind that diagonal matrices compute.
We know give an example to illustrate that diagonal matrices compute

Example 2.2. Suppose

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$$

To show the diagonal matrices compute, we need to show that AB = BA.

$$AB = \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix}$$

$$BA = \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix}$$

Thus, we've shown that diagonal matrices compute.

Another concept, we need to understand is the orthonormal bases.

Definition 2.3. A list of vectors is called orthonormal if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list. In other words, a list $e_1, \ldots, e_m$ of vectors in V is orthonormal if

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}$$

Remember orthogonal is just a fancy word meaning perpendicular. Two vectors u, v $\in$ V are called orthogonal if $\langle$u v$\rangle$ = 0

Here is an example of the orthonormal lists.

Example 2.4.

$$(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}), (\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$$

We're going to show that each vector has norm 1 and it is orthogonal to the other vector.

$$\sqrt{(\frac{1}{\sqrt{3}})^2 + (\frac{1}{\sqrt{3}})^2 + (\frac{1}{\sqrt{3}})^2} = 1$$

$$\sqrt{(\frac{-1}{\sqrt{2}})^2 + (\frac{1}{\sqrt{2}})^2 + 0} = 1$$

we've shown that the normal of each vector is 1. Now, we're going to show that the vectors are orthogonal to each other. In other words, we must show that their inner product is zero.

Using Wolfram Alpha calculator, we got that

$$\langle (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}), (\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0) \rangle = 0$$

and

$$\langle (\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}) \rangle = 0$$

Thus the two vectors are orthogonal to each other. Hence, the two vectors are examples of orthonormal lists.

It is also important to define the orthonormal basis.

Definition 2.5. An orthonormal basis of V is an orthonormal list of vectors in V that is also a basis of V.

Example 2.6. The standard basis is an orthonormal basis of F^n

Here is an important theorem we're going to use in our proof of the main theorem.

Proposition 2.7. *Schur's theorem. Suppose V is a finite-dimensional complex vector space and T ∈ L(V). Then T has an upper-triangular matrix with respect to some ortthonormal basis of V.*

Proof. The proof of this theorem is beyond the scope of this paper. The full proof can be found on page 186 of Linear Algebra Done Right [1]. □

Now, we're going to define the adjoint of a matrix.

Definition 2.8. suppose T ∈ L(v, w). The adjoint of T is the function $T^*: W \to V$ such that

$$\langle Tv, w \rangle = \langle v, T^* w \rangle$$

for every v ∈ V and every w ∈ W

To see why the definition above makes sense, we're going to provide an example.

Example 2.9. Define T: $R^3 \to R^2 by T(x_1, x_2, x_3) = (x_2, 3x_3, 2x_1)$. Find a formula for T^*

To compute T^*, fix a point $(y_1, y_2) \in R^2$.

Then for every $(x_1, x_2, x_3) \in R^3$, we have

$$\begin{aligned}
&\langle (x_1, x_2, x_3), T^*(y_1, y_2) \rangle = \langle T(x_1, x_2, x_3), (y_1, y_2) \rangle && \text{by the definition of adjoint} \\
&= \langle (x_2, 3x_3, 2x_1), (y_1, y_2) \rangle && \text{by the definition of T} \\
&= x_2 y_1, +3x_3 y_1 + 2x_1 y_2 && \text{by computing the inner product} \\
&= \langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1) \rangle && \text{by rearranging the inner product}
\end{aligned}$$

Thus, $T^*(y_1, y_2) = (2y_2, y_1, 3y_1)$

To find the adjoint T^*, we compute the conjugate transpose of the matrix T.

Definition 2.10. The conjugate transpose of an m-b-n matrix is the n-by-m matrix obtained by interchanging the rows and columns and then taking the complex conjugate of each entry

Example 2.11.

The conjugate transpose of the matrix

$$\begin{bmatrix} 2 & 3+4i & 7 \\ 6 & 5 & 8i \end{bmatrix}$$

is the matrix

$$\begin{bmatrix} 2 & 6 \\ 3-4i & 5 \\ 7 & -8i \end{bmatrix}$$

The final concept we need to understand before proving the main theorem is normal operators

Definition 2.12. An operator on an inner product space is called normal if it commutes with its adjoint. In other words, $T \in L(v)$ is normal if $TT^* = T^*T$

Example 2.13. Let T be an operator on F^n whose matrix is

$$T = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$$

$$TT^* = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 13 & 0 \\ 0 & 13 \end{bmatrix}$$

$$T^*T = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 13 & 0 \\ 0 & 13 \end{bmatrix}$$

Before we move to the second most important proposition for our main theorem, we're going to explain the relationship between the inner product and norms.

Definition 2.14. For v ∈ V, the norm of v, denoted $||v||$, is defined by

$$||v|| = \sqrt{\langle v, v \rangle}$$

if $||Z_1, \ldots, z_n|| \in F$, then

$$||z_1, \ldots, z_n|| = \sqrt{||z_1||^2 + \ldots |||z_n||^2}$$

Proposition 2.15. *An operator T is normal if and only if* $||TV||$ = $||T^*v||$, *for all v.*

We follow the proof given in [1, 7.20]

Proof. Let $T \in L(V)$. We're going to prove both directions of this proof at the same time.

T is normal

$$\begin{aligned}
&\Leftrightarrow T^*T = TT^* && \text{by the definition of normal operators}\\
&\Leftrightarrow \langle (T^*Tv, v\rangle = \langle TT^*v, v\rangle \forall v \in V && \text{taking the inner product of each side and applying v to it}\\
&\Leftrightarrow \langle TV, TV\rangle = \langle T^*V, T^*V\rangle \forall v \in V && \text{in each of these inner products, flip to the otherside}\\
&\Leftrightarrow ||Tv||^2 = ||T^*v||^2\ \forall v \in V && \text{by Definition 2.14}\\
&\Leftrightarrow ||Tv|| = ||T^*v||\ \forall v \in V && \text{by taking the square root from both sides}
\end{aligned}$$

□

3. Main Theorem

Now that we have introduced diagonal matrices, orthonormal bases, adjoints, and normal operators, we are ready to state and prove Theorem A. This theorem says that if V is a complex inner-product space and $T \in L(V)$. Then V has an orthonormal basis consisting of eigenvectors of T if and only if T is normal.

Theorem A (the Complex Spectral Theorem, [1, 7.24])**.** Suppose F = C and $T \in L(V)$. Then the following are equivalent:

(a) T is normal
(b) v has an orthonormal basis consisting of eigenvectors of T
(c) T has a diagonal matrix with respect to some orthonormal basis of v.

We follow the proof given in [1, 7.25].

Proof. First, we assume that (c) holds. In other words, T has a diagonal matrix with respect to some orthonormal basis of v. We can compute the matrix of the adjoint T^* (with respect to the same basis) by taking the conjugate transpose of the matrix T; hence T also has a diagonal matrix. Recall that any two diagonal matrices compute; thus T computes with T. Thus T is normal. In other words, (a) holds.

We've shown that (c) implies (a).

Now let's show that (a) implies (c).

By Schur's theorem, there is an orthonormal basis $(e1, \ldots, en)$ of V with respect to which T has an upper-triangular matrix. Thus,

$$A = M(T, (e1, \ldots, en)) = \begin{bmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ 0 & \ldots & a_{n,n} \end{bmatrix}$$

$$B = M(T^*, (e1, \ldots, en)) = \begin{bmatrix} \overline{a_{1,1}} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ \overline{a_{1,n}} & \ldots & \overline{a_{n,n}} \end{bmatrix}$$

We will show that the matrix A is actually a diagonal matrix

We see from the matrix above that

$$T(e_1) = a_{1,1}e_1$$

and

$$T^*(e_1) = \overline{a_{1,1}}e_1 + \ldots + \overline{a_{1,n}}e_n$$

Since the basis are orthonormal, we can easily calculate their squared norm.

$$||T(e_1)||^2 = |a_{1,1}|^2$$

and

$$||T^*(e_1)||^2 = |a_{1,1}|^2 + |a_{1,2}|^2 + \ldots + |a_{1,n}|^2$$

Because T is normal, $||Te_1||$ = $||T^*e_1||$ by proposition 2

Thus the two equations above imply that all entries in the first row of the matrix except possibly the diagonal entry $a_{1,1}$ equal 0.

Now we see that

$$||T(e_2)||^2 = |a_{2,2}|^2$$

and

$$||T(e_2)||^2 = |a_{2,2}|^2 + |a_{2,2}|^2 + \ldots + |a_{2,n}|^2$$

Because T is normal, $||Te_2||$ = $||T^*e_2||$ by proposition 2

Thus the two equations above imply that all entries in the second row of the matrix except possibly the diagonal entry $a_{2,2}$ equal 0.

Continuing in this fashion, we see that all the nondiagonal entries in the matrix equal 0. Thus (c) holds.

We're now going to show that (b) implies (c).

Suppose (b) holds. In other words, V has an orthonormal basis B consisting of eigenvectors of T. Then by 5.41 ((b)-> (a)), T is diagonalizable. Looking at the proof of 5.41, T is digonalized by the basis B. Thus is B is orthonormal. Therefore, T has a diagonal matrix with respect to some orthonormal basis of V.

We're now going to show that (c) implies (b).

Suppose c holds, which means that T has a diagonal matrix with respect to some orthonormal basis of v. Then by 5.41 ((a)->(b))(and looking again at the first paragraph of the proof), this means tha the vectors of B are eigenvectors. Therefore, V has a basis of eigenvectors of T and these vectors are orthonormal.

Hence, this proves theorem A.

□

The following example illustrates this theorem:

Consider the normal operator $T \in L(C^2)$ whose matrix with respect to the standard basis is

$$\begin{bmatrix} 2 & -3 \\ 3 & 3 \end{bmatrix}$$

$\frac{(i,1)}{\sqrt{2}}, \frac{(-i,1)}{\sqrt{2}}$ is an orthonormal basis of C^2 consisting of eigenvectors of T. With respect to the above basis, The matrix T is the diagonal matrix

$$\begin{bmatrix} 2+3i & 0 \\ 0 & 2-3i \end{bmatrix}$$

4. Applications

We now give one application that demonstrate the utility of Theorem A.

4.1. **Raising a matrix to a Higher Power.** Suppose we have a matrix A and we want to find A^{50}. One can compute it by multiplying A with itself 50 times. But this computation is a lot. However, with diagonalization, we can easily compute high powers of a matrix relatively easily.

We follow the example in [2]

Suppose A is a diagonalizable matrix. In other words, suppose A is similar to a diagonal matrix D.

$$D = P^{-1}AP$$

We can rearrange this equation to write:

$$A = PDP^{-1}$$

let's find A^2

since A= PDP^{-1}, then

$$A^2 = (PDP^{-1})^2$$

Similarly,

$$A^3 = (PDP^{-1})^3 = PD^3P^{-1}$$

In general,

$$A^n = (PDP^{-1})^N = PD^NP^{-1}$$

As you can see above, we've reduced the problem to just finding D^n.

Because D is diagonal, we can just multiply all of the entries on the main diagonal. (Kuttler, Ken [2])

References

[1] S. Axler. *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Springer International Publishing, Cham, 3rd ed. 2015 edition.

[2] K. Ken. 7.3: Applications of spectral theory. *Mathematics LibreTexts, Libretexts*.

[3] S. L.A. Highlights in the history of spectral theory. *The American Mathematical Monthly*.

Department of Mathematics, Brown University, Providence, RI 02912

Email address: muhiim_ali@brown.edu