Appendix H

Basic Functional Analysis

H.1 Finite Hilbert Space.

For a finite-dimensional space every Hermitian or unitary operator can be described completely by its eigenvalues and eigenvectors.

For a finite-dimensional space the eigenvectors span the whole space.

This comprises the so-called spectral theorem.

H.1.1 The Hamilton-Cayley Theorem.

Let \( A \) be a square \( N \times N \) matrix representing in some basis the operator \( A \), and let \( \lambda \) be a parameter. The equation

\[
\varphi(\lambda) := \det(\lambda E - A) = 0
\]  

(H.1)

is called the characteristic equation of the operator \( A \) (or of the matrix \( A \)). It is evident that \( \psi(\lambda) \) is a polynomial of the \( N \)th degree in \( \lambda \) with numerical coefficients the leading (that of \( \lambda^N \) being equal to 1)

\[
\varphi(\lambda) = \varphi_0 + \varphi_1 \lambda + \varphi_2 \lambda^2 + \cdots + \varphi_{N-1} \lambda^{N-1} + \varphi_N \lambda^N
\]  

(H.2)

The form of the characteristic (and thus the numerical values of the coefficients, \( \psi_i \)) does not depend on the choice of basis, since the determinant of a matrix, in our case, the matrix \( \det(\lambda E - A) \), is a scalar.

Replacing \( \lambda \) by the operator \( A \) we get the operator
\( A(\lambda) = A_0 + A_1 \lambda + A_2 \lambda^2 + \cdots + A_{N-1} \lambda^{N-1} + A_N \lambda^N \) \hfill (H.3)

### H.1.2 Projection Operators

In chapter 3 we saw that a projection operator is equivalent to operators that are self-adjoint and satisfy \( P^2 = P \).

We say that a subspace \( \mathcal{M} \) reduces a linear operator \( T \) if \( T \psi \) is in \( \mathcal{M} \) for every \( \psi \in \mathcal{M} \) and \( T \phi \) is in \( \mathcal{M}^\perp \) for every \( \phi \) in \( \mathcal{M}^\perp \). Let \( P \) be the projection operator onto \( \mathcal{M} \). We say that a subspace \( \mathcal{M} \) is invariant under an operator if \( T \psi \) is in \( \mathcal{M} \) for every vector in \( \mathcal{M} \).

**Theorem H.1.1** Let \( P \) be the projection operator onto the subspace \( \mathcal{M} \). The following statements are equivalent:

1. \( \mathcal{M} \) reduces \( T \);
2. \( PT = TP \)

**Proof:** First we show i) implies ii). For any vector

\[ \psi = \psi_M + \psi_{M^\perp} \]

we have

\[ T\psi = T\psi_M + T\psi_{M^\perp}. \]

If \( \mathcal{M} \) reduces \( T \), then \( T\psi_M \) is in \( \mathcal{M} \) and \( T\psi_{M^\perp} \) is in \( \mathcal{M}^\perp \). Therefore

\[
PT\psi = PT(\psi_M + \psi_{M^\perp}) = PT\psi_M = T\psi_M = TP\psi_M = TP\psi.
\]

Thus i) implies ii). We now show the converse. If \( \psi \) is in \( \mathcal{M} \) and \( PT = TP \), then \( P\psi = \psi \) and

\[ T\psi = TP\psi = PT\psi \]

so \( T\psi \) is in \( \mathcal{M} \). It is easy to see that \( PT = TP \) and \( (1 - P)T = T(1 - P) \) are equivalent. If \( \phi \) is in \( \mathcal{M}^\perp \) and \( (1 - P)T = T(1 - P) \), then
\[ T\phi = T(1 - P)\phi = (1 - P)T\phi \]

so \( T\phi \) is in \( \mathcal{M}^\perp \).

\[ \square \]

**Theorem H.1.2** If a subspace is invariant under \( T \) and \( T^\dagger \) then it reduces \( T \).

**Proof:** Let \( \mathcal{M} \) be a subspace that is invariant under \( T \) and \( T^\dagger \) and let \( P \) be the projection operator onto \( \mathcal{M} \). Then \( TP\psi \) is in \( \mathcal{M} \), so

\[ TP\psi = PTP\psi \]

for every vector \( \psi \), or

\[ TP = PTP. \]

Because \( \mathcal{M} \) is also invariant under \( T^\dagger \), we have

\[ T^\dagger P = PT^\dagger P. \]

Taking the adjoint, we have

\[ PT = PTP. \]

Therefore

\[ PT = TP. \]

This implies that \( \mathcal{M} \) reduces \( T \) by theorem H.1.1.

\[ \square \]

**Theorem H.1.3** If \( P \) and \( Q \) are the projections on closed linear subspaces \( M \) and \( N \) of \( H \), then \( M \perp N \) if and only if \( PQ = 0 \) or equivalently \( QP = 0 \).

**Proof:** First note that \( PQ = 0 \) is equivalent to \( QP = 0 \) through taking adjoints. If \( M \perp N \), so that \( N \subseteq M^\perp \), as \( Q\psi \) is in \( N \) for every \( \psi \) we have \( PQ\psi = 0 \), or \( PQ = 0 \). Conversely, if \( PQ = 0 \) then for every \( \psi \) in \( N \) we have \( P\psi = PQ\psi = 0 \), so \( N \subseteq M^\perp \) and \( M \perp N \).

\[ \square \]

We say that two projections \( P \) and \( Q \) are orthogonal if \( PQ = 0 \).
Theorem H.1.4  If \( P_1, P_2, \ldots, P_n \) are the projections on closed linear subspaces \( M_1, M_2, \ldots, M_n \) of \( H \), then \( P = P_1 + P_2 + \cdots + P_n \) is a projection if and only if the \( P_i \)'s are pairwise orthogonal (\( P_i P_j = 0 \) whenever \( i \neq j \)). In this case \( P \) is the projection onto

\[ M = M_1 + M_2 + \cdots + M_n. \]

**Proof:** Since \( P \) is self-adjoint, to prove it is a projection, we need only show it is idempotent, that is, \( P^2 = P \). As the \( P_i \)'s are pairwise orthogonal,

\[
P^2 = (P_1 + P_2 + \cdots + P_n)(P_1 + P_2 + \cdots + P_n) = P_1^2 + P_2^2 + \cdots + P_n^2 = P_1 + P_2 + \cdots + P_n = P. \tag{H.5}
\]

Therefore \( P \) is a projection. Conversely, assume that \( P \) is idempotent. Let \( \psi \) be a vector in the range of \( P_i \) then

\[
\|\psi\|^2 = \|P_i \psi\|^2 \leq \sum_{j=1}^{n} \|P_j \psi\|^2 = \sum_{j=1}^{n} (P_j \psi, \psi) = (P \psi, \psi) = \|P \psi\|^2 \leq \|\psi\|^2
\]

Given we started with \( \|\psi\|^2 \) and ended with \( \|\psi\|^2 \), equality must hold through, so

\[
\sum_{j=1}^{n} \|P_j \psi\| = \|P_i \psi\|
\]

and

\[
\|P_j \psi\| = 0 \quad \text{for } j \neq i.
\]

Therefore the range of \( P_i \) is contained in the null space of \( P_j \), that is, \( M_i \subseteq M_j^\perp \), for every \( j \neq i \). So \( M_i \perp M_j \) whenever \( j \neq i \), and by the previous theorem we conclude that the \( P_i \)'s are pairwise orthogonal. We now prove the final statement. Denote the range of \( P \) by \( \text{Ran}(P) \). First note that since \( \|P \psi\| = \|\psi\| \) for every \( \psi \) in \( M_i \), each \( M_i \) is contained in the range of \( P \), and therefore

\[ M \subseteq \text{Ran}(P). \tag{H.6} \]
Next, if $\psi$ is in the range of $P$, then

\[
\psi = P\psi = P_1\psi + P_2\psi + \cdots P_n\psi
\]

is obviously in $M$, and so $\psi$ is in $M$, hence

\[
\text{Ran}(P) \subseteq M. \quad (H.7)
\]

Comparing (H.6) and (H.7) implies $M = \text{Ran}(P)$.

\hfill \Box

### H.1.3 Spectral Theorem for Finite Spaces

An operator on $H$ is said to be **normal** if it commutes with its adjoint, that is, $NN^\dagger = N^\dagger N$. Self-adjoint and unitary operators are examples of normal operators. The spectral theorem states that for each normal operator $N$ on $H$ has a spectral resolution, that is, there exist distinct complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_m$ and non-zero pairwise orthogonal projections $P_1, P_2, \ldots, P_m$ such that $\sum_{i=1}^m P_i = I$, and

\[
N = \sum_{i=1}^m \lambda_i P_i.
\]

Before coming on to the spectral theorem, we prove results for normal operators.

**Lemma H.1.5** If $T$ is normal, then $\psi$ is an eigenvector of $T$ with eigenvalue $\lambda$ if and only if $\psi$ is an eigenvector of $T^\dagger$ with eigenvalue $\lambda^*$.  

**Proof:**

First we show that an operator $T$ is normal if and only if $\|T^\dagger\psi\| = \|T\psi\|$ for every $\psi$. Obviously $\|T^\dagger\psi\| = \|T\psi\|$ is equivalent to $\|T^\dagger\psi\|^2 = \|T\psi\|^2$, that is, $(T^\dagger, \psi T^\dagger \psi) = (T\psi, T\psi)$. This in turn is equivalent to $(T T^\dagger \psi, \psi) = (T^\dagger T \psi, \psi)$. Finally this is equivalent to $(T T^\dagger - T^\dagger T)\psi, \psi) = 0$.

Since $T$ is normal, it is obvious that also is $T - \lambda I$ for any scalar $\lambda$,

\[
(T - \lambda I)(T^\dagger - \lambda^* I) = TT^\dagger - \lambda^* T - \lambda T^\dagger + |\lambda|^2 I = (T^\dagger - \lambda^* I)(T - \lambda I).
\]

Hence we have
\[ \|T - \lambda I\| = \|T^\dagger - \lambda^* I\| \]

for all \( \psi \), and the lemma follows at once.

\[ \square \]

**Lemma H.1.6** If \( T \) is normal, then each \( M_i \) reduces \( T \).

**Proof:**

It is obvious that each \( M_i \) is mapped onto itself under \( T \). By theorem H.1.2 it suffices to show that each \( M_i \) is invariant under \( T^\dagger \). But this follows from lemma H.1.5, for if \( \psi_i \) is a vector in \( M_i \), so \( T\psi_i = \lambda_i \psi_i \), then \( T^\dagger \psi_i = \lambda_i^* \psi_i \) is also in \( M_i \).

\[ \square \]

We are now in a position to prove the spectral theorem.

**Theorem H.1.7** Let \( T \) be an arbitrary operator on \( H \). By ... we know that the distinct eigenvalues of \( T \) form a non-empty finite set of complex numbers. Let \( \lambda_1, \lambda_2, \ldots, \lambda_m \) be these eigenvalues, let \( M_1, M_2, \ldots, M_n \) be their corresponding eigenspaces, and let \( P_1, P_2, \ldots, P_n \) be the projections onto these eigenspaces. The following statements are equivalent:

i) The \( M_i \)'s are pairwise orthogonal and span \( H \).

ii) The \( P_i \)'s orthogonal, \( I = \sum_i^m P_i \) and \( T = \sum_i^m \lambda_i P_i \).

iii) \( T \) is normal.

**Proof:**

First we prove that i) implies ii). By i) every vector \( \psi \) in \( H \) can be expressed uniquely in the form

\[ \psi = \psi_1 + \psi_2 + \cdots + \psi_m, \quad (H.8) \]

where \( \psi_i \) is in \( M_i \) for each \( i \) and the \( \psi_i \)'s are pairwise orthogonal. It follows that

\[ T\psi = T\psi_1 + T\psi_2 + \cdots + T\psi_m = \lambda_1 \psi_1 + \lambda_2 \psi_2 + \cdots + \lambda_m \psi_m. \quad (H.9) \]

By theorem (H.1.3) the \( M_i \)'s being pairwise orthogonal is equivalent to the projection operators \( P_i \) being pairwise orthogonal.
\[ Ix = x = x_1 + x_2 + \cdots + x_m = P_1x + P_2x + \cdots + P_mx = (P_1 + P_2 + \cdots + P_m)x \]  
\[ \text{(H.10)} \]

for every \( x \) in \( H \), so

\[ I = P_1 + P_2 + \cdots + P_m. \]  
\[ \text{(H.11)} \]

Equation (H.9)

\[ T\psi = \lambda_1\psi_1 + \lambda_2\psi_2 + \cdots + \lambda_m\psi_m = \lambda_1P_1\psi + \lambda_2P_2\psi + \cdots + \lambda_mP_m\psi = (\lambda_1P_1 + \lambda_2P_2 + \cdots + \lambda_mP_m)\psi. \]  
\[ \text{(H.12)} \]

for every \( \psi \), therefore

\[ T = \lambda_1P_1 + \lambda_2P_2 + \cdots + \lambda_mP_m. \]  
\[ \text{(H.13)} \]

We now show that ii) implies iii). By (H.13) we have

\[ T^\dagger = \lambda_1^*P_1 + \lambda_2^*P_2 + \cdots + \lambda_m^*P_m, \]

and by using that the \( P_i \)'s are pairwise orthogonal we obtain

\[ TT^\dagger = (\lambda_1P_1 + \lambda_2P_2 + \cdots + \lambda_mP_m)(\lambda_1^*P_1 + \lambda_2^*P_2 + \cdots + \lambda_m^*P_m) = |\lambda_1|^2P_1 + |\lambda_2|^2P_2 + \cdots + |\lambda_m|^2P_m \]  
\[ \text{(H.14)} \]

and, similarly,

\[ T^\dagger T = |\lambda_1|^2P_1 + |\lambda_2|^2P_2 + \cdots + |\lambda_m|^2P_m. \]

so that

\[ TT^\dagger = T^\dagger T. \]
Lastly we show that iii) implies i), completing the proof that all conditions of the theorem imply each other. First we show that if $T$ is normal, then the $M_i$’s are pairwise orthogonal. Let $\psi_i$ and $\psi_j$ be vectors in $M_i$ and $M_j$ for $i \neq j$, so that $T\psi_i = \lambda_i \psi_i$ and $T\psi_j = \lambda_j \psi_j$. We have

$$
\lambda_i(\psi_i, \psi_j) = (\lambda_i^* \psi_i, \psi_j)
= (T^\dagger \psi_i, \psi_j)
= (\psi_i, T\psi_j)
= (\psi_i, \lambda_j \psi_j) = \lambda_j(\psi_i, \psi_j) \quad \text{(H.15)}
$$

where we used lemma H.1.5 in going from the first line to second line, and since $\lambda_i \neq \lambda_j$, it is obvious that $(\psi_i, \psi_j) = 0$, which is the desired result.

We now prove that the $M_i$’s span $H$ when $T$ is normal. By theorem H.1.3, the fact that the $M_i$’s are pairwise orthogonal implies the $P_i$’s are pairwise orthogonal. In turn by theorem H.1.4 we have that $M = M_1 + M_2 + \cdots + M_m$ is a closed linear subspace of $H$, and that its associated projection is

$$
P = P_1 + P_2 + \cdots + P_m.
$$

Since by lemma H.1.6 each $M_i$ reduces $T$, from theorem H.1.1 we have that $TP_i = P_i T$ for each $P_i$. It follows from this that $TP = PT$, so, from H.1.1 again, $M$ also reduces $T$, and consequently $M^\perp$ is invariant under $T$. Say $M^\perp \neq \{\emptyset\}$. All the eigenvectors of $T$ reside in $M$, the restriction of $T$ to $M^\perp$ represents an operator on the space $M^\perp$ but one which has no eigenvectors, and hence no eigenvalues. It follows from the fact that every operator must have at least one eigenvalue that we have a contradiction. We conclude that $M^\perp = \{\emptyset\}$ and hence the $M_i$’s span $H$.

\[\square\]

The expression

$$
T = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m. \quad \text{(H.16)}
$$

for $T$, when it exists, is called the spectral resolution of $T$. The spectral theorem in particular proves that if $T$ is normal, then it has the spectral resolution (H.16).