Let $p$ be the smallest prime dividing the order of a finite group $G$. Show that a subgroup $H$ of $G$ of index $p$ is necessarily normal.

Let $G$ act on cosets $gH$ of $H$ by left multiplication. This gives a homomorphism $f$ of $G$ to the group of permutations of $[G : H] = p$ things. The kernel $\ker f$ certainly lies inside $H$, since $gH = H$ only for $g \in H$. Thus, $p|[G : \ker f]$. On the other hand,

$$|f(G)| = [G : \ker f] = |G|/|\ker f|$$

and $|f(G)|$ divides the order $p!$ of the symmetric group on $p$ things, by Lagrange. But $p$ is the smallest prime dividing $|G|$, so $f(G)$ can only have order 1 or $p$. Since $p$ divides the order of $f(G)$ and $|f(G)|$ divides $p$, we have equality. That is, $H$ is the kernel of $f$. Every kernel is normal, so $H$ is normal.

Let $T \in \text{Hom}_k(V)$ for a finite-dimensional $k$-vectorspace $V$, with $k$ a field. Let $W$ be a $T$-stable subspace. Prove that the minimal polynomial of $T$ on $W$ is a divisor of the minimal polynomial of $T$ on $V$. Define a natural action of $T$ on the quotient $V/W$, and prove that the minimal polynomial of $T$ on $V/W$ is a divisor of the minimal polynomial of $T$ on $V$.

Let $f(x)$ be the minimal polynomial of $T$ on $V$, and $g(x)$ the minimal polynomial of $T$ on $W$. (We need the $T$-stability of $W$ for this to make sense at all.) Since $f(T) = 0$ on $V$, and since the restriction map

$$\text{End}_k(V) \to \text{End}_k(W)$$

is a ring homomorphism,

$$\text{(restriction of) } f(t) = f(\text{restriction of } T)$$

Thus, $f(T) = 0$ on $W$. That is, by definition of $g(x)$ and the PID-ness of $k[x]$, $f(x)$ is a multiple of $g(x)$, as desired.

Define $\overline{T}(v + W) = Tv + W$. Since $TW \subset W$, this is well-defined. Note that we cannot assert, and do not need, an equality $TW = W$, but only containment. Let $h(x)$ be the minimal polynomial of $T$ (on $V/W$). Any polynomial $p(T)$ stabilizes $W$, so gives a well-defined map $p(T)$ on $V/W$. Further, since the natural map

$$\text{End}_k(V) \to \text{End}_k(V/W)$$

is a ring homomorphism, we have

$$p(\overline{T})(v + W) = p(T)(v) + W = p(T)(v + W) + W = p(\overline{T})(v + W)$$

Since $f(T) = 0$ on $V$, $f(\overline{T}) = 0$. By definition of minimal polynomial, $h(x)|f(x)$.

Let $T \in \text{Hom}_k(V)$ for a finite-dimensional $k$-vectorspace $V$, with $k$ a field. Suppose that $T$ is diagonalizable on $V$. Let $W$ be a $T$-stable subspace of $V$. Show that $T$ is diagonalizable on $W$.

Since $T$ is diagonalizable, its minimal polynomial $f(x)$ on $V$ factors into linear factors in $k[x]$ (with zeros exactly the eigenvalues), and no factor is repeated. By the previous example, the minimal polynomial $g(x)$ of $T$ on $W$ divides $f(x)$, so (by unique factorization in $k[x]$) factors into linear factors without repeats. And this implies that $T$ is diagonalizable when restricted to $W$.

Let $T \in \text{Hom}_k(V)$ for a finite-dimensional $k$-vectorspace $V$, with $k$ a field. Suppose that $T$ is diagonalizable on $V$, with distinct eigenvalues. Let $S \in \text{Hom}_k(V)$ commute with $T$, in the natural sense that $ST = TS$. Show that $S$ is diagonalizable on $V$.

The hypothesis of distinct eigenvalues means that each eigenspace is one-dimensional. We have seen that commuting operators stabilize each other’s eigenspaces. Thus, $S$ stabilizes each one-dimensional $\lambda$-eigenspaces $V_\lambda$ for $T$. By the one-dimensionality of $V_\lambda$, $S$ is a scalar $\mu_\lambda$ on $V_\lambda$. That is, the basis of eigenvectors for $T$ is unavoidably a basis of eigenvectors for $S$, too, so $S$ is diagonalizable.
[16.5] Let $T \in \text{Hom}_k(V)$ for a finite-dimensional $k$-vectorspace $V$, with $k$ a field. Suppose that $T$ is diagonalizable on $V$. Show that $k[T]$ contains the projectors to the eigenspaces of $T$.

Though it is only implicit, we only want projectors $P$ which commute with $T$.

Since $T$ is diagonalizable, its minimal polynomial $f(x)$ factors into linear factors and has no repeated factors. For each eigenvalue $\lambda$, let $f_\lambda(x) = f(x)/(x-\lambda)$. The hypothesis that no factor is repeated implies that the \textit{gcd} of all these $f_\lambda(x)$ is 1, so there are polynomials $a_\lambda(x)$ in $k[x]$ such that

$$1 = \sum \lambda a_\lambda(x) f_\lambda(x)$$

For $\mu \neq \lambda$, the product $f_\lambda(x)f_\mu(x)$ picks up all the linear factors in $f(x)$, so

$$f_\lambda(T)f_\mu(T) = 0$$

Then for each eigenvalue $\mu$

$$(a_\mu(T)f_\mu(T))^2 = (a_\mu(T)f_\mu(T))(1 - \sum_{\lambda \neq \mu} a_\lambda(T)f_\lambda(T)) = (a_\mu(T)f_\mu(T))$$

Thus, $P_\mu = a_\mu(T)f_\mu(T)$ has $P_\mu^2 = P_\mu$. Since $f_\lambda(T)f_\mu(T) = 0$ for $\lambda \neq \mu$, we have $P_\mu P_\lambda = 0$ for $\lambda \neq \mu$. Thus, these are projectors to the eigenspaces of $T$, and, being polynomials in $T$, commute with $T$.

For uniqueness, observe that the diagonalizability of $T$ implies that $V$ is the sum of the $\lambda$-eigenspaces $V_\lambda$ of $T$. We know that any endomorphism (such as a projector) commuting with $T$ stabilizes the eigenspaces of $T$. Thus, given an eigenvalue $\lambda$ of $T$, an endomorphism $P$ commuting with $T$ and such that $P(V) = V_\lambda$ must be 0 on $T$-eigenspaces $V_\mu$ with $\mu \neq \lambda$, since

$$P(V_\mu) \subset V_\mu \cap V_\lambda = 0$$

And when restricted to $V_\lambda$ the operator $P$ is required to be the identity. Since $V$ is the sum of the eigenspaces and $P$ is determined completely on each one, there is only one such $P$ (for each $\lambda$). ///

[16.6] Let $V$ be a complex vector space with a (positive definite) inner product. Show that $T \in \text{Hom}_k(V)$ cannot be a normal operator if it has any non-trivial Jordan block.

The spectral theorem for normal operators asserts, among other things, that normal operators are diagonalizable, in the sense that there is a basis of eigenvectors. We know that this implies that the minimal polynomial has no repeated factors. Presence of a non-trivial Jordan block exactly means that the minimal polynomial \textit{does} have a repeated factor, so this cannot happen for normal operators. ///

[16.7] Show that a positive-definite hermitian $n$-by-$n$ matrix $A$ has a unique positive-definite square root $B$ (that is, $B^2 = A$).

Even though the question explicitly mentions matrices, it is just as easy to discuss endomorphisms of the vector space $V = \mathbb{C}^n$.

By the spectral theorem, $A$ is diagonalizable, so $V = \mathbb{C}^n$ is the sum of the eigenspaces $V_\lambda$ of $A$. By hermitian-ness these eigenspaces are mutually orthogonal. By positive-definiteness $A$ has \textit{positive} real eigenvalues $\lambda$, which therefore have real square roots. Define $B$ on each orthogonal summand $V_\lambda$ to be the scalar $\sqrt{\lambda}$. Since these eigenspaces are mutually orthogonal, the operator $B$ so defined really is hermitian, as we now verify. Let $v = \sum \lambda v_\lambda$ and $w = \sum \mu w_\mu$ be \textit{orthogonal} decompositions of two vectors into eigenvectors $v_\lambda$ with eigenvalues $\lambda$ and $w_\mu$ with eigenvalues $\mu$. Then, using the orthogonality of eigenvectors with distinct eigenvalues,

$$\langle Bv, w \rangle = \langle B \sum \lambda v_\lambda, \sum \mu w_\mu \rangle = \langle \sum \lambda v_\lambda, \sum \mu w_\mu \rangle = \sum \lambda \langle v_\lambda, w_\mu \rangle$$

2
Uniqueness is slightly subtler. Since we do not know \textit{a priori} that two positive-definite square roots $B$ and $C$ of $A$ commute, we cannot immediately say that $B^2 = C^2$ gives $(B + C)(B - C) = 0$, etc. If we could do that, then since $B$ and $C$ are both positive-definite, we could say
\[
\langle (B + C)v, v \rangle = \langle Bv, v \rangle + \langle Cv, v \rangle > 0
\]
so $B + C$ is positive-definite and, hence invertible. Thus, $B - C = 0$. But we cannot directly do this. We must be more circumspect.

Let $B$ be a positive-definite square root of $A$. Then $B$ commutes with $A$. Thus, $B$ stabilizes each eigenspace of $A$. Since $B$ is diagonalizable on $V$, it is diagonalizable on each eigenspace of $A$ (from an earlier example). Thus, since all eigenvalues of $B$ are positive, and $B^2 = \lambda$ on the $\lambda$-eigenspace $V_\lambda$ of $A$, it must be that $B$ is the scalar $\sqrt{\lambda}$ on $V_\lambda$. That is, $B$ is uniquely determined.

[16.8] Given a square $n$-by-$n$ complex matrix $M$, show that there are unitary matrices $A$ and $B$ such that $AMB$ is diagonal.

We prove this for \textit{not-necessarily} square $M$, with the unitary matrices of appropriate sizes.

This asserted expression
\[
M = \text{unitary} \cdot \text{diagonal} \cdot \text{unitary}
\]
is called a \textbf{Cartan decomposition} of $M$.

First, if $M$ is \textit{(square) invertible}, then $T = MM^*$ is self-adjoint and invertible. From an earlier example, the spectral theorem implies that there is a self-adjoint (necessarily invertible) square root $S$ of $T$. Then
\[
1 = S^{-1}TS^{-1} = (S^{-1}M)(-S^{-1}SM)^*
\]
so $k_1 = S^{-1}M$ is unitary. Let $k_2$ be unitary such that $D = k_2Sk_2^*$ is diagonal, by the spectral theorem. Then
\[
M = Sk_1 = (k_2Dk_2^*)k_1 = k_2 \cdot D \cdot (k_2^*k_1)
\]
expresses $M$ as
\[
M = \text{unitary} \cdot \text{diagonal} \cdot \text{unitary}
\]
as desired.

In the case of \textit{m-by-n} (not necessarily invertible) $M$, we want to reduce to the invertible case by showing that there are \textit{m-by-m} unitary $A_1$ and \textit{n-by-n} unitary $B_1$ such that
\[
A_1MB_1 = \begin{pmatrix} M' & 0 \\ 0 & 0 \end{pmatrix}
\]
where $M'$ is \textit{square} and invertible. That is, we can (in effect) do column and row reduction with \textit{unitary} matrices.

Nearly half of the issue is showing that by left (or right) multiplication by a suitable unitary matrix $A$ an arbitrary matrix $M$ may be put in the form
\[
AM = \begin{pmatrix} M_{11} & M_{12} \\ 0 & 0 \end{pmatrix}
\]
with 0’s below the $r^{th}$ row, where the column space of $M$ has dimension $r$. To this end, let $f_1, \ldots, f_r$ be an orthonormal basis for the \textit{column space} of $M$, and extend it to an orthonormal basis $f_1, \ldots, f_m$ for the
whole \( \mathbb{C}^m \). Let \( e_1, \ldots, e_m \) be the standard orthonormal basis for \( \mathbb{C}^m \). Let \( A \) be the linear endomorphism of \( \mathbb{C}^m \) defined by \( Af_i = e_i \) for all indices \( i \). We claim that this \( A \) is unitary, and has the desired effect on \( M \). That is has the desired effect on \( M \) is by design, since any column of the original \( M \) will be mapped by \( A \) to the span of \( e_1, \ldots, e_r \), so will have all 0’s below the \( r^{th} \) row. A linear endomorphism is determined exactly by where it sends a basis, so all that needs to be checked is the unitariness, which will result from the orthonormality of the bases, as follows. For \( v = \sum_i a_i f_i \) and \( w = \sum_i b_i f_i \),

\[
\langle Av, Aw \rangle = \langle \sum_i a_i Af_i, \sum_j b_j Af_j \rangle = \langle \sum_i a_i e_i, \sum_j b_j e_j \rangle = \sum_i a_i \bar{b}_i
\]

by orthonormality. And, similarly,

\[
\sum_i a_i \bar{b}_i = \langle \sum_i a_i f_i, \sum_j b_j f_j \rangle = \langle v, w \rangle
\]

Thus, \( \langle Av, Aw \rangle = \langle v, w \rangle \). To be completely scrupulous, we want to see that the latter condition implies that \( A^*A = 1 \). We have \( \langle A^*Av, w \rangle = \langle v, w \rangle \) for all \( v \) and \( w \). If \( A^*A \neq 1 \), then for some \( v \) we would have \( A^*Av \neq v \), and for that \( v \) take \( w = (A^*A - 1)v \), so

\[
\langle (A^*A - 1)v, w \rangle = \langle (A^*A - 1)v, (A^*A - 1)v \rangle > 0
\]

contradiction. That is, \( A \) is certainly unitary.

If we had had the foresight to prove that row rank is always equal to column rank, then we would know that a combination of the previous left multiplication by unitary and a corresponding right multiplication by unitary would leave us with

\[
\begin{pmatrix}
M' \\
0 \\
0
\end{pmatrix}
\]

with \( M' \) square and invertible, as desired.

[16.9] Given a square \( n \)-by-\( n \) complex matrix \( M \), show that there is a unitary matrix \( A \) such that \( AM \) is upper triangular.

Let \( \{e_i\} \) be the standard basis for \( \mathbb{C}^n \). To say that a matrix is upper triangular is to assert that (with left multiplication of column vectors) each of the maximal family of nested subspaces (called a maximal flag)

\[
V_0 = 0 \subset V_1 = \mathbb{C}e_1 \subset \mathbb{C}e_1 + \mathbb{C}e_2 \subset \ldots \subset \mathbb{C}e_1 + \ldots + \mathbb{C}e_{n-1} \subset V_n = \mathbb{C}^n
\]

is stabilized by the matrix. Of course

\[
MV_0 \subset MV_1 \subset MV_2 \subset \ldots \subset MV_{n-1} \subset V_n
\]

is another maximal flag. Let \( f_{i+1} \) be a unit-length vector in the orthogonal complement to \( MV_i \) inside \( MV_{i+1} \). Thus, these \( f_i \) are an orthonormal basis for \( V \), and, in fact, \( f_1, \ldots, f_i \) is an orthonormal basis for \( MV_i \). Then let \( A \) be the unitary endomorphism such that \( Af_i = e_i \). (In an earlier example and in class we checked that, indeed, a linear map which sends one orthonormal basis to another is unitary.) Then

\[
AMV_i = V_i
\]

so \( AM \) is upper-triangular.

[16.10] Let \( Z \) be an \( m \)-by-\( n \) complex matrix. Let \( Z^* \) be its conjugate-transpose. Show that

\[
\det(1_m - ZZ^*) = \det(1_n - Z^*Z)
\]
Write $Z$ in the (rectangular) Cartan decomposition

$$Z = ADB$$

with $A$ and $B$ unitary and $D$ is $m$-by-$n$ of the form

$$D = \begin{pmatrix}
    d_1 & & & \\
    & d_2 & & \\
    & & \ddots & \\
    & & & d_r
\end{pmatrix}$$

where the diagonal $d_i$ are the only non-zero entries. We grant ourselves that $\det(xy) = \det(x) \cdot \det(y)$ for square matrices $x, y$ of the same size. Then

$$\det(1_m - ZZ^*) = \det(1_m - ADB^*D^*A^*) = \det(1_m - ADD^*A^*) = \det(A \cdot (1_m - DD^*) \cdot A^*)$$

$$= \det(AA^*) \cdot \det(1_m - DD^*) = \det(1_m - DD^*) = \prod_i (1 - d_i \bar{d_i})$$

Similarly,

$$\det(1_n - Z^*Z) = \det(1_n - B^*D^*A^*ADB) = \det(1_n - B^*D^*DB) = \det(B^* \cdot (1_n - D^*D) \cdot B)$$

$$= \det(B^*B) \cdot \det(1_n - D^*D) = \det(1_n - D^*D) = \prod_i (1 - d_i \bar{d_i})$$

which is the same as the first computation.