Notation: \( V = \text{vector space} \) (real or complex) 
\( x, y, \ldots \in V = \text{vectors} \)
\( a, b, \ldots \in \mathbb{C} = \text{coefficients (numbers, real or complex).} \)

Let \( T : V \to V \) be a \underline{linear} operator. (Here mapping a vector space into itself.)

Linear means, \( T(ax + by) = aT(x) + bT(y) \).

Now introduce a basis and look at representation of \( T \) in the basis. Let \( \{e_i, i = 1, \ldots, n\} \) be a basis in \( V \) where \( n = \dim(V) \).

(Assume you know basic linear algebra, linear independence, etc.)

Let \( y = Tx \), expand both \( x, y \) in the basis \( \{e_i\} : \)

\[
\begin{align*}
  x &= \sum_i x_i e_i, \\
  y &= \sum_j y_j e_j,
\end{align*}
\]

The components \( x_i \) represent the vector \( x \) in the basis \( \{e_i\} \).

Now look at \( y = Tx = T \left( \sum_i x_i e_i \right) = \sum_i x_i T(e_i) \) by linearity. (Note \( Tx \) means same as \( T(\mathbf{x}) \).)

This makes it clear that if you know the action of \( T \) on the basis vectors, you know what \( T \) does to any vector. So look at \( T(e_i) \). This is a vector, so can be represented as a linear combination of the basis vectors,
say,

\[ T(e_i) = \sum_j e_j T_{ji} \]  \((*)\)

This is a definition of the numbers (matrix of numbers) \(T_{ji}\), also called the components of \(T\) in the basis \(\{e_i\}\).

Ok, now plug this back into \(y = Tx\),

\[ y = \sum_j y_j e_j = \sum_i x_i \sum_j e_j T_{ji} \]

or, since \(\{e_j\}\) are lin. indep.,

\[ y_j = \sum_i T_{ji} x_i \]

Swap \(i, j\) indices,

\[ y_i = \sum_j T_{ij} x_j \]  \((i*)\)

Contrast this with (*) above. This tells how the components of a vector transform under \(T\) (it's a matrix times a component vector), whereas (*) tells how the basis vectors transform under \(T\). Notice that the component matrices are \(T_{ij}\) in one case, \(T_{ji}\) in the other (transposes of each other).

Don't confuse these two. (Transforming basis vectors is usually more profound.)
Notice that in deriving \((\ast)\) and \((\ast\ast)\), we did not use a metric
(no scalar products), in fact the relations depend only on linear
independence arguments.

If a vector space does not have a metric, then there is no meaning

\[\text{to "orthogonal vectors", "unit vectors", "the angle between vectors" etc.}\]

But if the space does have a metric, then we can take scalar

products, for example, between the basis vectors \((\mathbf{e}_i, \mathbf{e}_j)\).

Then an orthonormal basis is one that satisfies,

\[ (\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij} \]

If we have an orthonormal basis, then we can take the scalar

product of \((\ast)\) with \(\mathbf{e}_k\),

\[ (\mathbf{e}_k,\mathbf{T}\mathbf{e}_i) = \sum_j (\mathbf{e}_k,\mathbf{e}_j) \mathbf{T}_{ji} = \sum_j \delta_{kj} \mathbf{T}_{ji} = \mathbf{T}_{ki}, \]

or

\[ T_{ij} = (\mathbf{e}_i, \mathbf{T}\mathbf{e}_j) \quad \text{< easy to remember.} \]

In this case, we say the components of \(\mathbf{T}\) in the (orthonormal)

basis \(\{\mathbf{e}_i\}\) are given by the matrix elements \((\mathbf{e}_i, \mathbf{T}\mathbf{e}_j)\).
In the following assume we have a metric (complex, to be general).

Some definitions:

**Def.** If $T$ is a linear operator, then the **adjoint** $T^+$ is the operator that satisfies

$$(T^+x, y) = (x, Ty) \quad \text{for all } x, y \in V.$$ 

**Notes.**

1. This is an implicit definition. To see that it actually specifies $T^+$, given $T$, introduce a basis \{${e_i}$\} (orthonormal), so that

$$(T^+)_{ij} = (e_i, T^+e_j)$$

$$= (T^+e_j, e_i)^*$$

$$= (e_j, Te_i)^*$$

$$= T_{ji}^*.$$ 

Thus, the component matrix of $T^+$ (in an orthonormal basis) is the transpose complex conjugate of the component matrix of $T$.

2. Note that this implies $T^{++} = T$.

3. $T^+$ is also called the **Hermitean conjugate** of $T$. To be precise, the adjoint and Hermitean conjugate are the same on finite dimensional vector spaces.

4. In mathematic's literature, the adjoint is represented by $\ast$ (not $+$), and the complex conjugate by overbar ($\overline{\cdot}$), (not $\ast$).
Def. An operator is **self-adjoint** or **Hermitian** if \( T = T^* \), so that \( T_{ij} = T_{ji}^* \). On a real vector space, it is called **symmetric** (since \( T_{ij} = T_{ji} \)).

A Hermitian operator satisfies \( (Tx, y) = (x, Ty) \) for all \( x, y \in V \). Hermitian operators are important in quantum mechanics. It is believed that every physical observable is represented by a Hermitian operator.

Def. An operator is **unitary** if \( U^{-1} = U^* \), so that \( (U^{-1})_{ij} = U_{ji}^* \).

On a real vector space, the matrix of a unitary operator is orthogonal, \( (U^{-1})_{ij} = U_{ji} \) or \( U^{-1} = U^T \) (\( T = \) transpose).

Unitary operators have the property that they preserve the scalar product, i.e. \( U \) is unitary if

\[
(Ux, Uy) = (x, y) \quad \text{for all} \quad x, y \in V.
\]

This is equivalent to the def. above because \( (Ux, Uy) = (x, U^*Uy) = (x, y) \) for all \( x, y \) only if \( U^*U = I \) (the identity).

---

Notes on the eigenvalue problem. If \( A \) is an operator, then if there exists some number \( \lambda \) and nonzero vector \( x \) such that \( Ax = \lambda x \), then we say that \( \lambda \) is an eigenvalue of \( A \) and \( x \) is an *eigenvector* corresponding to that eigenvalue.
The number $a$ is (in general) complex. The set of $a$'s such that $a$ are eigenvalues of $A$ is called the spectrum of the operator $A$. This is some set of points in the complex plane.

So any multiple of $x$ is also an eigenvector.

Suppose $a$ is an eigenvalue of $A$. Then there exists some vector $x \neq 0$ such that $Ax = ax$. It's easy to show the set of all $x$ such that $Ax = ax$ (for fixed $a$) is a vector space; call it the eigenspace of $A$ corresponding to eigenvalue $a$.

Eigenspace of $A$ with corresp to $a = \{x \in V \mid Ax = ax\}$.

The eigenspace has dimensionality at least $1$, but it may be $>1$. This dimensionality is called the order of degeneracy of the eigenvalue $a$. It is the number of linearly independent $x$ such that $Ax = ax$.

The eigenvalues of Hermitian operators are real. Proof:

Let $A = A^*$, and assume $Ax = ax$ (for $x \neq 0$). Then

$$(x, Ax) = a(x, x)$$

$$= (A^*x, x) \quad \text{(defn of $A^*$)}$$

$$= (Ax, x) \quad \text{(since $A^* = A$)}$$

$$= (x, Ax)^* \quad \text{(properties of scalar product)}$$

$$= \alpha^* (x, x).$$

So, $(a - \alpha^*)(x, x) = 0$. 

6
Another important property of Hermitian operators is that their eigenspaces correspond to distinct eigenvalues are orthogonal. Proof:

Let

\[ \begin{align*}
A x_1 &= a_1 x_1 \quad \text{assume } x_1 \neq x_2 \quad \text{Assume } A = A^+. \\
A x_2 &= a_2 x_2
\end{align*} \]

Then

\[ \begin{align*}
(x_2, Ax_1) &= a_1 (x_2, x_1) \\
and (x_1, Ax_2) &= a_2 (x_1, x_2).
\end{align*} \]

2nd eqn \( \Rightarrow \) \( (Ax_1, x_2) = a_2 (x_1, x_2) \) since \( A = A^+ \), which implies

\[ \Rightarrow (x_2, Ax_1) = a_2 (x_2, x_1) \quad \text{taking } * \quad \text{and using fact } a_2^* = a_2. \]

So, subtr., get \( 0 = (a_1 - a_2) (x_2, x_1) \Rightarrow \) either \( a_1 = a_2 \) or \( (x_2, x_1) = 0. \)

An important issue is that of completeness. An operator is complete if its eigenvectors span the whole vector space \( V \) on which the operator acts.

The issue is easy on a finite-dimensional vector space. All Hermitian operators on a finite-dimensional vector space are complete. For this course we will pretend the same is true on an \( \infty \)-dimensional space. Completeness is necessary for the usual physical interpretation of quantum mechanics (otherwise the measurement of an observable would annihilate probability).
Notation about orthogonal vector spaces. Let $E_1, E_2$ be orthogonal subspaces of $V$.

\[ \dim E_1 = 2 \]
\[ \dim E_2 = 1 \]

Then $E_1 \oplus E_2$, the direct sum of $E_1$ and $E_2$, is the set of all vectors that are linear combinations of vectors from $E_1$ and $E_2$.

For a Hermitian operator $A = A^*$, let the eigenvalues (the spectrum) be $(\lambda_1, \lambda_2, \ldots)$, and let $E_{\lambda_n}$ be the eigenspace corresponding to $\lambda_n$. Then:

\[ V = \bigoplus_{\lambda_n} E_{\lambda_n} = E_1 \oplus E_2 \oplus \ldots \]

This is a way of stating the completeness property of Hermitian operators.

Now induced transformation of functions, try to motivate this better than the book. Consider a wave packet in 1D, centered about $x_0$.

We'd like to define a translation operator that acts on wave functions and that moves the wave packet from $x_0$ to $x_0 + \xi$.
Let $\Psi'(x) =$ the new (translated) wave function, $\Psi = T(\xi)\Psi$.

What is $\Psi'(x)$ in terms of $\Psi(x)$? Not $\Psi'(x) = \Psi(x + \xi)$, but rather

$$\Psi'(x) = (T(\xi)\Psi)(x) = \Psi(x - \xi)$$

An important sign.

For example, at the center of the wave packet (old and new) we must have equal values,

$$\Psi'(x_0 + \xi) = \Psi(x_0)$$

more generally,

$$\Psi'(x + \xi) = \Psi(x) \text{ any } x$$

which is the same as $\Psi'(x) = \Psi(x - \xi)$.

Easy to remember in words: The value of the new wave fn at the new point is equal to the value of the old wave function at the old point.

There really are 2 translation operators, one that acts on points and one that acts on functions.

$$T(\xi)x = x + \xi \quad \text{(points)} \quad (A)$$

$$(T(\xi)\Psi)(x) = \Psi(x - \xi) = \Psi(T(\xi)x) \text{ (fnos)} \quad (B)$$

Notice the notation: in 2nd eqn., $T(\xi)$ acts on functions to produce new functions, hence the parentheses around $(T(\xi)\Psi)$. The book (many other books, too) writes $T(\xi)\Psi(x)$, confusing since $\Psi(x)$ is a number and $T(\xi)$ does not act on numbers.

Since the action of $T(\xi)$ on points, eqn. (A), the action of $T(\xi)$ on functions, eqn. (B), is called the induced transformation.
Similarly, consider rotations in 3D space. Let $R$ be such a rotation (mapping from pts of 3D space to other such points),

$$ R \mathbf{x} = \mathbf{x}' $$

Then the transformation of functions is

$$ (R \psi)(\mathbf{x}) = \psi(R^{-1} \mathbf{x}) $$

note the inverse.

More generally, let $S$ be any space, let $\psi$ be a fnc on $S$, that is $\psi : S \to \mathbb{R}$ (or $\mathbb{C}$), and let $T : S \to S$ be a transformation on $S$. Then we define the induced transformation on $\psi$ to be

$$ (T \psi)(x) = \psi(T^{-1} x) \quad \text{for } x \in S. $$

Suppose we have a group action on $S$. This means a set of transformations $T_a, T_b$ etc., acting on $S$, for $a, b \in \mathcal{G}$, such that $T_a T_b = T_{ab}$.

Do the set of induced transformations, acting on the space of functions on $S$, also a group action (does $T_a T_b = T_{ab}$ when the $T$'s are interpreted as acting on fns, not points)? Answer is yes (proof left as exercise).
About rotations. These begin life as operators that map 3D space to itself, in such a way that all distances and angles are preserved. If also handedness is preserved (a right-handed triad is mapped to another right-handed triad) then the rotation is proper.

In any case, these are linear operators, so if \{e_i\} is a basis, we have the components of the rotation defined by

\[
R e_i = \sum_j e_j R_{ij}
\]

like for any linear operator on any space. If \{e_i\} is orthonormal (usually is), then \( R_{ij} = (e_i, Re_j) \).

For example, in 2D let \( R(\theta) \) be a rotation by \( \theta \),

![Diagram of rotation](image)

Then can see geometrically,

- \( R e_1 = \cos \theta \ e_1 + \sin \theta \ e_2 \) (let \( e_1 = \hat{x} \) )
- \( R e_2 = -\sin \theta \ e_1 + \cos \theta \ e_2 \).

So

\[
R_{ij} = \begin{pmatrix}
R_{11} & R_{12} \\
R_{21} & R_{22}
\end{pmatrix}
= \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}.
\]

be careful about signs!

And

\[
\vec{r}' = \sum_j R_{ij} \vec{r}_j \quad \text{component version of } \vec{F}' = R \vec{F}.
\]
This is the active point of view of rotations: when we apply the rotation, old points get up and move to new points. The eqn \( r'_i = \sum_j R_{ij} r_j \) gives the coordinates of the old and new points in the same coordinate system.

Beware! There is also the passive point of view, used (almost always) in courses like 105, where the eqn \( r'_i = \sum_j R_{ij} r_j \) usually means, the relation between the coordinates of one and only one point with respect to old and new coordinate systems. In the passive point of view you rotate the coordinate axes, not the points of space. The \( R_{ij} \) matrix in the active point of view and that in the passive are not the same. (Inverses of each other.)

In this course we will always use the active point of view.

Now discuss an example in the book. Consider space of functions,

\[
\begin{align*}
\psi_1 &= x^2 \\
\psi_2 &= y^2 \\
\psi_3 &= z^2 \\
\psi_4 &= xy \\
\psi_5 &= xz \\
\psi_6 &= xy
\end{align*}
\]

These span the 6-dimensional space of all quadratic polynomials in \((x, y, z)\). So let's talk about the vector space that is the 6-dimensional span of these functions.
These functions are not orthonormal with respect to the usual scalar product for wave functions,

\[ (\psi, \phi) = \int d^3\mathbf{r} \, \psi(\mathbf{r})^* \phi(\mathbf{r}), \] (usual scalar product)

in fact, these integrals diverge for \((\psi_1, \ldots, \psi_6)\) so the scalar products don't even exist. Therefore the book suggests using another scalar product,

\[ (\psi, \phi) = \int d^3\mathbf{r} \, \psi(\mathbf{r})^* \phi(\mathbf{r}) \] (unit sphere)

This has no physical motivation, it's just used to get a finite scalar product. With this scalar product you find that \((\psi_1, \ldots, \psi_6)\) are not orthonormal, but by forming linear combinations you can make an orthonormal basis.

Another possibility for fixing up the metric (defining a scalar product that does not diverge) is to use

\[ (\psi, \phi) = \int d^3\mathbf{r} \, e^{-r} \psi(\mathbf{r})^* \phi(\mathbf{r}) \] (all space)

where the convergence factor \(e^{-r}\) guarantees convergence of the integrals. Obviously this is the same as replacing \((\psi_1, \ldots, \psi_6)\) by \((e^{-r_1/2} \psi_1, \ldots, e^{-r_6/2} \psi_6)\) and using the usual scalar product. The new wave functions are similar to hydrogen atom wave fns.
Anyway, now let's consider what happens when we apply a rotation operator to one of these functions. Choose

$$R = R_1 = R_x\left(\frac{2\pi}{3}\right)$$

which is the rotation $R_1$ in the group $D_3$.

Note, \[
\sin \frac{2\pi}{3} = \frac{\sqrt{3}}{2} \quad \cos \frac{2\pi}{3} = -\frac{1}{2} \]

So, the matrix for $R_1$ is

$$R_1 = \begin{pmatrix}
-\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\
\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\
0 & 0 & 1
\end{pmatrix}$$

Let's apply this to $\psi_1$, where $\psi_1(\vec{r}) = \psi_1(x,y,z) = x^2$.

Defn says,

$$(R_1 \psi_1)(\vec{r}) = \psi_1(R_1^{-1} \vec{r})$$

so we need $R_1^{-1}$, (obtained by $\theta \rightarrow -\theta$),

$$R_1^{-1} = \begin{pmatrix}
-\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\
-\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\
0 & 0 & 1
\end{pmatrix}$$

(3 x 3 matrix)
\[ R_i^{-1} \Phi = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}x + \frac{\sqrt{3}}{2}y \\ -\frac{\sqrt{3}}{2}x - \frac{1}{2}y \\ z \end{pmatrix} \]

so,

\[ (R_i \Psi_1)(\Phi) = \Psi_1 \left( -\frac{1}{2}x + \frac{\sqrt{3}}{2}y, -\frac{\sqrt{3}}{2}x - \frac{1}{2}y, z \right) \]

\[ = \Psi_1 \left( \frac{x^2}{4} - \frac{\sqrt{3}}{2}xy + \frac{3}{4}y^2 \right) \]

or,

\[ R_i \Psi_1 = \frac{1}{4} \Psi_1 - \frac{\sqrt{3}}{2} \Psi_6 + \frac{3}{4} \Psi_2. \quad (\text{see eq. } 3.8.6 \text{ of text}) \]

So the rotated fn. belongs to the same 6-dimensional space.

If we do this to the other 5 basis fns, we find

\[ R_i \Psi_2 = \frac{3}{4} \Psi_1 + \frac{1}{4} \Psi_2 - \frac{\sqrt{3}}{2} \Psi_6 \]
\[ R_i \Psi_3 = \Psi_3 \]
\[ R_i \Psi_4 = -\frac{1}{2} \Psi_4 + \frac{\sqrt{3}}{2} \Psi_5 \]
\[ R_i \Psi_5 = -\frac{\sqrt{3}}{2} \Psi_4 - \frac{1}{2} \Psi_5 \]
\[ R_i \Psi_6 = -\frac{\sqrt{3}}{2} \Psi_1 + \frac{\sqrt{3}}{2} \Psi_2 - \frac{1}{2} \Psi_6 \]

C.f. general formula,

\[ T e_i = \sum_j e_j \tau_{ji}, \]

and we can read off the components of \( R_i \) in the 6-dimensional basis \((\psi_1, \ldots, \psi_6)\). For example, \((R_1)_{21} = \frac{3}{4}, (R_1)_{61} = -\frac{\sqrt{3}}{2}, \) etc.
So we get the matrix,

\[
(R_1)_{ij} = \begin{pmatrix}
\frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & \sqrt{3}/2 \\
\frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 & -\sqrt{3}/2 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{2} & \sqrt{2}/2 & 0 \\
0 & 0 & 0 & -\sqrt{3}/2 & -\frac{1}{2} & 0 \\
-\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{2} & 0 & 0 & 0 & -\frac{1}{2}
\end{pmatrix}
\] (6x6 matrix)

Above we wrote a 3x3 matrix for $R_1$, now we are writing a 6x6 matrix. What's the difference? The 3x3 matrix refers to the action of $R_1$ on 3D space $(x,y,z)$, whereas the 6x6 matrix refers to the induced action of $R_1$ on the 6-dim space of functions $(\psi_1, \ldots, \psi_6)$. ^maybe should write $\vec{\psi}$

(it's really abuse of notation to use the same symbol $R_1$ for these different actions, but $R_1$, $R_1$, $\bar{R}_1$, etc. gets awkward.)