Chapter 10. Transformation Theory

10.1. General Theory
- **§1** Introduction
- **§2** The properties of the operator $\hat{U}$ and of the matrix U
- **§3** The connection between ψ_i' and ψ_i
- **§4** The connection between the matrix elements A_{ij} and A_{ij}'
- **§5** The physics must be the same
- **§6** Summary

10.2. Two Applications
- **§7** Introduction
- **§8** $2\mathrm{p_x}$, $2\mathrm{p_y}$, and $2\mathrm{p_z}$
- **§9** sp^3 orbitals

Section 10.1. General Theory

§ 1 *Introduction.* We assume in what follows that the system of interest can be described by a finite-dimensional $\boldsymbol{\ell}^2$ space generated by the basis set

$$\{|\phi_i\rangle\}_{i=1}^{N} \tag{1}$$

As a consequence we have

$$|\psi\rangle = \sum_{i=1}^{N} |\phi_i\rangle \langle\phi_i \mid \psi\rangle \equiv \sum_{i=1}^{N} \psi_i \, |\phi_i\rangle \tag{2}$$

Through this equation, the ket $|\psi\rangle$ is represented by an N-dimensional vector in $\boldsymbol{\ell}^2$ denoted by ψ and given by

$$\psi = \{\psi_1, \psi_2, \ldots, \psi_N\} \tag{3}$$

The components ψ_i of this vector are given by

$$\psi_i = \langle \phi_i \,|\, \psi \rangle \tag{4}$$

An operator $\hat{A}$ is represented, in the $\boldsymbol{\ell}^2$ space, by the matrix A, whose elements are

$$A_{ij} = \langle \phi_i \,|\, \hat{A} \,|\, \phi_j \rangle \tag{5}$$

Sometimes it is advantageous to use, in the same calculation, besides the "old" basis set Eq. 1, a "new", complete, orthonormal basis set

$$\{|\phi_i'\rangle\}_{i=1}^{N} \tag{6}$$

Why would I want to use two basis sets in the same calculation? Perhaps the problem is easy to formulate by using the old basis set but later I decided that the numerical evaluations are easier to perform with the new basis set. Or it might be easy to perform numerical calculations by using the old basis set but the physical interpretation is more transparent when the new basis set is used.

Using two basis sets in the same calculation raises a number of questions.

Question 1. The conversion from the old basis set to the new one defines an operator $\hat{U}$ through

$$|\phi_i'\rangle \equiv \hat{U}|\phi_i\rangle \tag{7}$$

Obviously this operator will be the key for converting expressions given in the old basis set into expressions given in the new basis set. What are the properties of $\hat{U}$, how do we calculate the elements of the matrix that represents it, and how is $\hat{U}$ involved in relating ψ_i to ψ'_i and relating A_{ij} to A'_{ij}?

Question 2. When different basis sets are used, the same ket $|\psi\rangle$ is *represented by different vectors* in $\boldsymbol{\ell}^2$. When we use the old basis set the vector representing ψ is given by Eq. 3. When we use the new basis set the same ket $|\psi\rangle$ is represented by

$$|\psi\rangle = \sum_{i=1}^{N} |\phi'_i\rangle \langle\phi'_i \,|\, \psi\rangle \equiv \sum_{i=1}^{N} \psi'_i \,|\phi_i\rangle \tag{8}$$

This means that if we use the new basis set, $|\psi\rangle$ is represented by the new vector ψ'

$$\psi' = \{\psi'_1, \psi'_2, \ldots, \psi'_N\} \tag{9}$$

with components ψ'_i given by

$$\psi'_i \equiv \langle\phi'_i \,|\, \psi\rangle \tag{10}$$

When performing calculations, we need to switch between these two representations of $|\psi\rangle$. The question is: how is ψ'_i connected to ψ_i? If I know ψ, how do I calculate ψ' and vice versa?

Question 3. In the old basis set the operator $\hat{A}$ is represented by a matrix A having the elements defined by Eq. 5. If we use the new basis set, $\hat{A}$ will be represented by a "new" matrix A' whose elements are

$$A'_{ij} \equiv \langle\phi'_i \,|\, \hat{A} \,|\, \phi'_j\rangle \tag{11}$$

How are the matrix elements A'_{ij} related to A_{ij}? If I know A, how do I calculate A' and vice versa?

Question 4. We expect that all measurable properties of a system are independent of the basis set used to represent the kets and the operators. Is that true?

§ **2** *The properties of the operator $\hat{U}$ and of the matrix U.* The definition of $\hat{U}$ given by Eq. 7 shows that $\hat{U}$ converts an orthonormal basis set into another orthonormal basis set. In Chapter 4 we have seen that an operator having this property is *unitary.* I remind you below some of the properties of unitary operators that will be used in what follows.

1. If $\hat{O}$ is unitary, then
$$\hat{O}^{\dagger} = \hat{O}^{-1} \tag{12}$$
where $\hat{O}^{\dagger}$ is the Hermitian conjugate of $\hat{O}$ and $\hat{O}^{-1}$ is the inverse of $\hat{O}$ (a unitary operator always has an inverse).

2. The Hermitian conjugate $\hat{A}^{\dagger}$ of an operator $\hat{A}$ satisfies the equations
$$\langle \hat{A}^{\dagger}\psi \mid \phi\rangle = \langle \psi \mid \hat{A}\phi\rangle \tag{13}$$
and
$$\langle \hat{A}\psi \mid \phi\rangle = \langle \psi \mid \hat{A}^{\dagger}\phi\rangle \tag{14}$$
for any kets $|\psi\rangle$ and $|\phi\rangle$.

An important fact is that
$$\hat{U} = \sum_{i=1}^{N} |\phi_i'\rangle\,\langle \phi_i| \tag{15}$$
The proof is straightforward. Act with $\hat{U}$, as given by Eq. 15, on $|\phi_j\rangle$:
$$\hat{U}|\phi_j\rangle = \sum_{i=1}^{N} |\phi_i'\rangle\,\langle \phi_i \mid \phi_j\rangle = \sum_{i=1}^{N} |\phi_i'\rangle\,\delta_{ij} = |\phi_j'\rangle \tag{16}$$

This shows that $\hat{U}$ given by Eq. 15 has the property (Eq. 7) that defines $\hat{U}$.

It is easy to calculate the matrix elements of $\hat{U}$ in the old basis set (use Eq. 15):

$$\begin{aligned} U_{ij} &\equiv \langle \phi_i \,|\, \hat{U} \,|\, \phi_j \rangle = \langle \phi_i \,|\, \sum_{k=1}^{N} |\phi'_k\rangle \; \langle \phi_k | \phi_j \rangle \\ &= \sum_{k=1}^{N} \langle \phi_i \,|\, \phi'_k \rangle \, \delta_{kj} = \langle \phi_i \,|\, \phi'_j \rangle \end{aligned} \tag{17}$$

This formula gives us a simple recipe for calculating the matrix U representing $\hat{U}$ in the old basis set. You can derive, if you need it, the elements of the matrix U' representing $\hat{U}$ in the new basis set:

$$\langle \phi'_k \,|\, \hat{U} \,|\, \phi'_j \rangle = \sum_{i=1}^{N} \langle \phi'_k \,|\, \phi'_i \rangle \, \langle \phi_i \,|\, \phi'_j \rangle = \sum_{i=1}^{N} \delta_{ki} \, \langle \phi_i \,|\, \phi'_j \rangle = \langle \phi_k \,|\, \phi'_j \rangle \tag{18}$$

Exercise 1 An orthonormal set $\{|\phi_i\rangle\}_{i=1}^{N}$ is complete if and only if

$$\langle \phi_i \,|\, y \rangle = 0 \text{ for all } i \text{ implies that } |y\rangle = 0 \tag{19}$$

Taking Eq. 19 to be the definition of completeness, show that the set $|\phi'_i\rangle = \hat{U}|\phi_i\rangle$, $i = 1, 2, \ldots, N$, is complete and orthonormal if $\{|\phi_i\rangle\}_{i=1}^{N}$ is complete and $\hat{U}$ is unitary.

Exercise 2 Prove that the following statements are equivalent for an orthonormal set $\{|\phi_i\rangle\}_{i=1}^{N}$:

(a) $\{|\phi_i\rangle\}_{i=1}^{N}$ is complete.

(b) For any ket $|y\rangle$, $\sum_{i=1}^{N} \langle \phi_i \,|\, y \rangle \, |\phi_i\rangle = |y\rangle$.

(c) For any ket $|y\rangle$, if $\langle \phi_i \,|\, y\rangle = 0$, $i = 1, 2, \ldots, N$, then $|y\rangle = 0$.

Exercise 3 Consider the space $\mathcal{L}$ consisting of functions $f(x)$ taking real values for $x \in [0, L]$. The scalar product in $\mathcal{L}$ is

$$\langle f \,|\, \phi\rangle \equiv \int_0^L f(x)\,\phi(x)\,\mathrm{d}x$$

Use the Gram-Schmidt procedure to construct a set of orthonormal polynomials $P_1(x)$, $\ldots$, $P_N(x)$ in $\mathcal{L}$ starting from the set $1, x, x^2, \ldots, x^{N-1}$. You also know that the set $\phi_n(x) = \sqrt{2/L}\,\sin(n\pi x/L)$, $n = 1, 2, \ldots, N$, is orthonormal.

1. Define $\hat{U}$ by

   $$P_i(x) = \hat{U}\phi_i(x)$$

   Show that the operator $\hat{U}$ is unitary.

2. Find the eigenvalues of U.

---

§ 3 *The connection between ψ_i' and ψ_i.* The vectors ψ and ψ' represent the same ket $|\psi\rangle$ in two different basis sets. Clearly, to connect ψ to ψ', I need to use the operator $\hat{U}$ which converts one basis set into another. Here is how the calculation goes:

$$\begin{aligned}
\psi_j \equiv\ & \langle \phi_j \,|\, \psi\rangle && \text{(used Eq. 4)} \\
=\ & \textstyle\sum_{i=1}^{N} \psi_i' \langle \phi_j \,|\, \phi_i'\rangle && \text{(used Eq. 8 for } |\psi\rangle\text{)} \\
=\ & \textstyle\sum_{i=1}^{N} \psi_i' \,\langle \phi_j \,|\, \hat{U}\phi_i\rangle && \text{(used Eq. 7 for } \phi_i'\text{)} \\
=\ & \textstyle\sum_{i=1}^{N} U_{ji}\,\psi_i'
\end{aligned} \tag{20}$$

Here

$$U_{ji} \equiv \langle \phi_j \,|\, \hat{U} \,|\, \phi_i \rangle \tag{21}$$

is the matrix element of the operator $\hat{U}$ in the old basis set $\{|\phi_i\rangle\}_{i=1}^{N}$. Using matrix and vector notation, we can write Eq. 20 as

$$\psi = U\psi' \tag{22}$$

The elements of the matrix U can be calculated from Eq. 17, which I reproduce below

$$U_{ij} = \langle \phi_i \,|\, \phi'_j \rangle \tag{23}$$

We know the two basis sets and therefore we can calculate the matrix elements U_{ij} from Eq. 23. This allows us to use Eq. 20 to calculate ψ from ψ'.

If we know ψ' and want to calculate ψ, we act with U^{-1} on Eq. 22 and obtain

$$\psi' = U^{-1}\psi = U^{\dagger}\psi \tag{24}$$

Even though $U^{\dagger} = U^{-1}$, it is much easier to calculate the adjoint matrix than the inverse matrix. This is why I gave ψ' in Eq. 24 in terms of $U^{\dagger}$.

Exercise 4 Show that

$$\psi'_j = \sum_i \langle \phi'_j \,|\, \hat{U}^{\dagger} \,|\, \phi'_i \rangle \psi_i$$

---

§ 4 *The connection between the matrix elements A_{ij} and A'_{ij}.* An operator $\hat{A}$ can be represented in the old basis by the matrix A having the elements

$$A_{ij} \equiv \langle \phi_i \,|\, \hat{A} \,|\, \phi_j \rangle \tag{25}$$

or in the new basis by a matrix A' having the elements

$$A'_{ij} \equiv \langle \phi'_i \,|\, \hat{A} \,|\, \phi'_j \rangle \tag{26}$$

The connection between A_{ij} and A'_{ij} is easily calculated:

$$\begin{aligned} A'_{ij} &= \langle \phi'_i \,|\, \hat{A} \,|\, \phi'_j \rangle \\ &= \langle \hat{U}\phi_i \,|\, \hat{A}\hat{U} \,|\, \phi_j \rangle \quad \text{(used Eq. 16)} \\ &= \langle \phi_i \,|\, \hat{U}^\dagger \hat{A}\hat{U} \,|\, \phi_j \rangle \quad \text{(used Eq. 14)} \end{aligned} \tag{27}$$

The matrix elements A'_{ij} of $\hat{A}$ in the new basis set $\{|\phi'_i\rangle\}_{i=1}^N$ are equal to the matrix elements of $\hat{U}^\dagger\hat{A}\hat{U}$ in the old basis set $\{|\phi_i\rangle\}_{i=1}^N$.

By inserting $\hat{I} = \sum_{\alpha=1}^N |\phi_\alpha\rangle\langle\phi_\alpha|$ between the operators, we can turn Eq. 27 into a matrix equation:

$$\begin{aligned} A'_{ij} &= \sum_{\alpha=1}^N \sum_{\beta=1}^N \langle \phi_i \,|\, \hat{U}^\dagger \,|\, \phi_\alpha \rangle \langle \phi_\alpha \,|\, \hat{A} \,|\, \phi_\beta \rangle \langle \phi_\beta \,|\, \hat{U} \,|\, \phi_j \rangle \\ &= \sum_{\alpha=1}^N \sum_{\beta=1}^N (U^\dagger)_{i\alpha} \, A_{\alpha\beta} \, U_{\beta j} \end{aligned} \tag{28}$$

We can write Eq. 28 as

$$A' = U^\dagger A U \tag{29}$$

Note that U, $U^\dagger$, and A are matrices calculated with the old basis set. Because $U^\dagger = U^{-1}$ and $U^{-1}U = I$, Eq. 29 is equivalent to

$$U A' U^\dagger = A \tag{30}$$

Exercise 5 Show that

$$A'_{ij} = \sum_{\alpha=1}^N \sum_{\beta=1}^N (U^*)_{\alpha i} \, A_{\alpha\beta} \, U_{\beta j}$$

---

§ 5 *The physics must be the same.* We can take an infinite number of basis sets and obtain an infinite number of representations of a ket or an operator. This is a bit disconcerting but it is not fatal: kets and operators are not observables and we have no reason to demand that they have a unique representation. However, the magnitude of quantities that are measurable must be independent of representation. Let us show that this is the case.

We start with the probability amplitude $\langle a_n \,|\, \psi\rangle$ where $|\psi\rangle$ is the state of the system and $|a_n\rangle$ is a pure state of an observable A. The probability that a measurement of A, when the system is in the state $|\psi\rangle$, gives the result a_n is

$$P_\psi(a_n) = |\langle a_n \,|\, \psi\rangle|^2 \tag{31}$$

The magnitude of $P_\psi(a_n)$ must be the same regardless of how we represent $|\psi\rangle$.

Suppose $|\psi\rangle$ is represented by

$$|\psi\rangle = \sum_{i=1}^{N} \psi'_i |\phi'_i\rangle \tag{32}$$

and by

$$|\psi\rangle = \sum_{i=1}^{N} \psi_i |\phi_i\rangle \tag{33}$$

If we use the basis set $\{|\phi_i\rangle\}_{i=1}^{N}$ to calculate $\langle a_n \,|\, \psi\rangle$, we have

$$\langle a_n \,|\, \psi\rangle = \sum_{i=1}^{N} \psi_i \langle a_n \,|\, \phi_i\rangle \quad \text{(used Eq. 33)} \tag{34}$$

If we use the basis set

$$\{|\phi'_i\rangle \equiv \hat{U}|\phi_i\rangle\}_{i=1}^{N}, \tag{35}$$

then the same amplitude is written as

$$\langle a_n \,|\, \psi\rangle' = \sum_{i=1}^{N} \psi_i' \langle a_n \,|\, \phi_i'\rangle \quad \text{(used Eq. 32)} \tag{36}$$

The prime in $\langle a_n \,|\, \psi\rangle'$ reminds me that this is $\langle a_n \,|\, \psi\rangle$ calculated with the basis set $\{|\phi_i'\rangle\}_{i=1}^{N}$. I will show that $\langle a_n \,|\, \psi\rangle'$ is identical to $\langle a_n \,|\, \psi\rangle$: changing the basis set does not change the probability amplitude.

I can write $\langle a_n \,|\, \psi\rangle'$ as

$$\begin{aligned}
\langle a_n \,|\, \psi\rangle' &= \textstyle\sum_{i=1}^{N} \langle a_n \,|\, \phi_i'\rangle \,\langle \phi_i' \,|\, \psi\rangle && \text{(used Eq. 36 and } \psi_i' \equiv \langle \phi_i' \,|\, \psi\rangle\text{)} \\
&= \textstyle\sum_{i=1}^{N} \langle a_n \,|\, \hat{U}\phi_i\rangle \,\langle \hat{U}\phi_i \,|\, \psi\rangle && \text{(used Eq. 35)} \\
&= \textstyle\sum_{i=1}^{N} \langle \hat{U}^\dagger a_n \,|\, \phi_i\rangle \,\langle \phi_i \,|\, \hat{U}^\dagger \psi\rangle && \text{(used } \textstyle\sum_{i=1}^{N} |\phi_i\rangle\,\langle\phi_i| = \hat{I}\text{)} \\
&= \langle \hat{U}^\dagger a_n \,|\, \hat{U}^\dagger \psi\rangle = \langle a_n \,|\, \hat{U}\hat{U}^\dagger \psi\rangle = \langle a_n \,|\, \psi\rangle && \text{(used } \hat{U}^{\dagger\dagger} = \hat{U} \text{ and Eq. 12)}
\end{aligned}$$

Thus we see that $\langle a_n \,|\, \psi\rangle$ and $\langle a_n \,|\, \psi\rangle'$ are equal: the probability amplitude is independent of the basis set used to represent $|\psi\rangle$. That is the way it should be!

The eigenvalues of an observable must also be independent of which basis set we use. In Chapter 9 we established a theorem that says that $\hat{A}$ and $\hat{U}^{-1}\hat{A}\hat{U}$ have the same eigenvalues. Since the matrix A' in the new basis set is equal to $U^{-1}AU$, that theorem tells us that A' has the same eigenvalues as A, which is what we wanted to know.

Another quantity to consider is the expectation value of an observable A, when the system is in a state $|\psi\rangle$:

$$\langle \psi \,|\, \hat{A} \,|\, \psi\rangle = \sum_{i=1}^{N} a_i P_\psi(a_i) \tag{37}$$

We have already shown that both $P_\psi(a_i)$ and a_i are unchanged when we change the basis set; therefore $\langle \psi \,|\, \hat{A} \,|\, \psi\rangle$ is also unchanged.

No measurable quantity changes when we change the basis set.

§ 6 *Summary.* We can use two orthonormal, complete basis sets

$$\{\, |\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_N\rangle \,\} \tag{38}$$

and

$$\{\, |\phi'_1\rangle, |\phi'_2\rangle, \ldots, |\phi'_N\rangle \,\} \tag{39}$$

to represent kets or operators. The operator $\hat{U}$ defined by

$$|\phi'_i\rangle = \hat{U}|\phi_i\rangle,\ i = 1, 2, \ldots, N \tag{40}$$

is unitary. We can represent a ket by

$$|\psi\rangle = \sum_{i=1}^{N} \psi_i |\phi_i\rangle \tag{41}$$

or by

$$|\psi\rangle = \sum_{i=1}^{N} \psi'_i |\phi'_i\rangle \tag{42}$$

and an operator $\hat{A}$ by

$$\hat{A} = \sum_{i=1}^{N} \sum_{j=1}^{N} |\phi_i\rangle\, A_{ij}\, \langle\phi_j| \tag{43}$$

or by

$$\hat{A} = \sum_{i=1}^{N} \sum_{j=1}^{N} |\phi'_i\rangle\, A'_{ij}\, \langle\phi'_j| \tag{44}$$

This means that in $\boldsymbol{\ell}^2$, a ket $|\psi\rangle$ can be represented either by the vector

$$\psi = \{\psi_1, \psi_2, ..., \psi_N\} \tag{45}$$

or by

$$\psi' = \{\psi'_1, \psi'_2, ..., \psi'_N\} \tag{46}$$

where

$$\psi_i = \langle \phi_i \,|\, \psi \rangle \tag{47}$$

and

$$\psi'_i = \langle \phi'_i \,|\, \psi \rangle \tag{48}$$

The two vectors are connected by

$$\psi'_i = \sum_{j=1}^{N} U_{ij} \psi_j \tag{49}$$

where

$$U_{ij} = \langle \phi_i \,|\, \hat{U} \,|\, \phi_j \rangle = \langle \phi_i \,|\, \phi'_j \rangle \tag{50}$$

The two matrices A and A', representing the same operator $\hat{A}$ in the different basis sets, are related through

$$\langle \phi'_i \,|\, \hat{A} \,|\, \phi'_j \rangle \equiv A'_{ij} = (U^{\dagger} A U)_{ij} = \sum_{\alpha=1}^{N} \sum_{\beta=1}^{N} (U^{\dagger})_{i\alpha}\, A_{\alpha\beta}\, U_{\beta j} \tag{51}$$

The matrices in the right-hand side are calculated in the basis set $\{|\phi_i\rangle\}_{i=1}^{N}$.

The value of any observable is the same whether we use the "old" or the "new" basis set. In particular, the eigenvalues, the probability amplitude, and the mean value of any observable are unchanged when we change the basis set.

Section 10.2. Two Applications

§ 7 *Introduction.* Transformation theory points out that there are many equivalent descriptions of a quantum process. They are interesting either because they illuminate different aspects in the behavior of the system, or because some representations make certain calculations simpler than other representations. In this section I show that you have already encountered applications of this theory.

When you studied the hydrogen atom, you learned that the angular dependence of the wave function is given by the spherical harmonics $Y_\ell^m(\theta, \phi)$ (see Metiu, *Quantum Mechanics*, p. 317). You were also told that for the purpose of describing the chemical bond, it is more useful to use the orbitals $n\mathrm{p_x}$, $n\mathrm{p_y}$, and $n\mathrm{p_z}$ instead of Y_ℓ^m. I will show here that the set $2\mathrm{p_x}$, $2\mathrm{p_y}$, $2\mathrm{p_z}$ is obtained from $Y_1^m(\theta, \phi)$ according to transformation theory: the matrix converting Y_1^1, Y_1^0, Y_1^{-1} into $2\mathrm{p_x}$, $2\mathrm{p_y}$, $2\mathrm{p_z}$ is unitary, and therefore $\{2\mathrm{p_x}, 2\mathrm{p_y}, 2\mathrm{p_z}\}$ and $\{Y_1^1, Y_1^0, Y_1^{-1}\}$ are equivalent orthonormal basis sets. For example, the eigenvalues of $\hat{L}_z$ for $j = 1$ can be calculated with either set to obtain the same results. The calculation is much easier in the basis set $\{Y_1^1, Y_1^0, Y_1^{-1}\}$. Moreover, that set is more convenient when we examine the spectroscopy of the atom. However, the basis $\{2\mathrm{p_x}, 2\mathrm{p_y}, 2\mathrm{p_z}\}$ is more convenient when we discuss the chemical bond in a diatomic molecule.

You encountered another example of transformation theory in general chemistry. You were told that the sp^3 hybrid orbitals are more appropriate for describing the valence electrons of the carbon atom than are $\{2\mathrm{s}, 2\mathrm{p_x}, 2\mathrm{p_y}, 2\mathrm{p_z}\}$. The reason: they point towards the corners of a tetrahedron, just like the bonds in $\mathrm{CH_4}$ (or other saturated bonds of the carbon atom). I will show here that the sp^3 orbitals are obtained by a unitary transformation of the basis set $\{2\mathrm{s}, 2\mathrm{p_x}, 2\mathrm{p_y}, 2\mathrm{p_z}\}$. In accordance with transformation theory, these two sets are equivalent.

§ 8 $2\mathrm{p_x}$, $2\mathrm{p_y}$, *and* $2\mathrm{p_z}$. The orbitals $2\mathrm{p_x}$, $2\mathrm{p_y}$, and $2\mathrm{p_z}$ are defined by (see Metiu, *Quantum Mechanics*)

$$2\mathrm{p_x}(\theta, \phi) = \frac{1}{\sqrt{2}}[Y_1^{-1}(\theta, \phi) - Y_1^1(\theta, \phi)] \tag{52}$$

$$\begin{aligned} 2\mathrm{p_y}(\theta,\phi) &= \frac{\mathrm{i}}{\sqrt{2}}[Y_1^{-1}(\theta,\phi) + Y_1^1(\theta,\phi)] &(53)\\ 2\mathrm{p_z}(\theta,\phi) &= Y_1^0(\theta,\phi) &(54)\end{aligned}$$

You can think of $Y_j^m(\theta,\phi)$ as the spherical-coordinate representation of the abstract ket $|j,m\rangle$. Similarly, $2\mathrm{p_x}(\theta,\phi)$ is the spherical-coordinate representation of the abstract ket $|2\mathrm{p_x}\rangle$. The kets $|2\mathrm{p_y}\rangle$ and $|2\mathrm{p_z}\rangle$ have similar interpretations. Because Eqs. 52–54 contain only spherical harmonics having $j = 1$, the presence of j in the notation $|j,m\rangle$ is superfluous — I'll use instead $|m\rangle$, with $m = -1$, 0, or 1.

In this abstract ket space, Eqs. 52–54 become

$$\begin{aligned} |2\mathrm{p_x}\rangle &= \frac{1}{\sqrt{2}}\left(|-1\rangle - |1\rangle\right) &(55)\\ |2\mathrm{p_y}\rangle &= \frac{\mathrm{i}}{\sqrt{2}}\left(|-1\rangle + |1\rangle\right) &(56)\\ |2\mathrm{p_z}\rangle &= |0\rangle &(57)\end{aligned}$$

I can write this as

$$\begin{pmatrix} |2\mathrm{p_x}\rangle \\ |2\mathrm{p_y}\rangle \\ |2\mathrm{p_z}\rangle \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{\mathrm{i}}{\sqrt{2}} & 0 & \frac{\mathrm{i}}{\sqrt{2}} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} |1\rangle \\ |0\rangle \\ |-1\rangle \end{pmatrix} \equiv U \begin{pmatrix} |1\rangle \\ |0\rangle \\ |-1\rangle \end{pmatrix} \qquad (58)$$

In Cell 1 of `WorkBook10_TransformationTheory.nb`, I verified that indeed the matrix U defined by Eq. 58 does the conversion of $\{|1\rangle, |0\rangle, |-1\rangle\}$ to $\{|2\mathrm{p_x}\rangle, |2\mathrm{p_y}\rangle, |2\mathrm{p_z}\rangle\}$ correctly and that the absolute value of each of its eigenvalues is equal to 1. This proves that the matrix U is unitary.

According to the theory developed in this chapter, any measurable quantity can be calculated by using $\{|-1\rangle, |0\rangle, |1\rangle\}$ as a basis set or by using

$\{|2\mathrm{p_x}\rangle, |2\mathrm{p_y}\rangle, |2\mathrm{p_z}\rangle\}$. As an example, let us calculate the eigenvalues of $\hat{L}_z$, by using these two basis sets.

For this calculation, the basis set $\{|-1\rangle, |0\rangle, |1\rangle\}$ is particularly advantageous because

$$\hat{L}_z|m\rangle = \hbar m|m\rangle, \quad m = -1, 0, 1 \tag{59}$$

The states $|m\rangle$ are pure states of $\hat{L}_z$, hence they are eigenstates of $\hat{L}_z$. The off-diagonal matrix elements of $\hat{L}_z$ in this basis are all equal to zero. For example,

$$\begin{aligned} \langle 1\,|\,\hat{L}_z\,|-1\rangle &= -\hbar\langle 1\,|-1\rangle && \text{(used Eq. 59)} \\ &= 0 && \text{(used orthogonality, } \langle 1\,|-1\rangle = 0\text{)} \end{aligned}$$

The diagonal elements are given by

$$\langle m\,|\,\hat{L}_z\,|\,m\rangle = \hbar m, \quad m = -1, 0, 1 \tag{60}$$

The matrix of $\hat{L}_z$ in this basis is therefore

$$L_z = \begin{pmatrix} \hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\hbar \end{pmatrix} \tag{61}$$

The eigenvalues are $\hbar$, 0, and $-\hbar$.

The matrix of $\hat{L}_z$ in the basis $\{|2\mathrm{p_x}\rangle, |2\mathrm{p_y}\rangle, |2\mathrm{p_z}\rangle\}$ can be calculated by using Eqs. 59 and 55–57. For example,

$$\begin{aligned} \langle 2\mathrm{p_x}\,|\,\hat{L}_z\,|\,2\mathrm{p_y}\rangle &= \tfrac{1}{\sqrt{2}}(\langle -1| - \langle 1|)\hat{L}_z(\tfrac{\mathrm{i}}{\sqrt{2}})(|-1\rangle + |1\rangle) && \text{(used Eqs. 55–57)} \\ &= \tfrac{\mathrm{i}}{2}(\langle -1\,|\,\hat{L}_z\,|-1\rangle - \langle 1\,|\,\hat{L}_z\,|\,1\rangle) && \text{(used } \langle m\,|\,\hat{L}_z\,|\,m'\rangle = \delta_{mm'}\text{)} \\ &= \tfrac{\mathrm{i}}{2}(-\hbar - \hbar) && \text{(used Eq. 59 and } \langle m\,|\,\hat{L}_z\,|\,m\rangle = 1\text{)} \\ &= -\hbar\mathrm{i} \end{aligned}$$

I calculated the matrix elements $\langle 2\mathrm{p_x} \,|\, \hat{L}_z \,|\, 2\mathrm{p_x}\rangle$, $\langle 2\mathrm{p_x} \,|\, \hat{L}_z \,|\, 2\mathrm{p_y}\rangle$, etc., in Cell 2b of `WorkBook10_TransformationTheory.nb`. The matrix L_z, corresponding to the operator $\hat{L}_z$, in the basis set $\{|2\mathrm{p_x}\rangle, |2\mathrm{p_y}\rangle, |2\mathrm{p_z}\rangle\}$ is

$$L_z = \begin{array}{c} \\ 2\mathrm{p_x} \\ 2\mathrm{p_y} \\ 2\mathrm{p_z} \end{array} \begin{array}{c} \begin{array}{ccc} 2\mathrm{p_x} & 2\mathrm{p_y} & 2\mathrm{p_z} \end{array} \\ \begin{pmatrix} 0 & -\mathrm{i}\hbar & 0 \\ \mathrm{i}\hbar & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}$$

The eigenvalues of this matrix are (see Cell 2b of `WorkBook10`) $\hbar$, $-\hbar$, and 0. They are exactly the same as the eigenvalues calculated in the basis $\{|-1\rangle, |0\rangle, |1\rangle\}$, as theory says they should be.

§ 9 *sp^3 orbitals.* The carbon atom has four valence electrons in the orbitals $\{2\mathrm{s}, 2\mathrm{p_x}, 2\mathrm{p_y}, 2\mathrm{p_z}\}$. These have the same energy as the orbitals $\{Y_0^0, Y_1^0, Y_1^1, Y_1^{-1}\}$ but they are not eigenstates of $\hat{L}_z$. Which one should we use if we want to construct a simple theory of the chemical bonds in methane? It turns out that neither is as suggestive of the structure of the molecule as the set $\{\mathrm{h_1}, \mathrm{h_2}, \mathrm{h_3}, \mathrm{h_4}\}$ of *hybrid orbitals*. These are defined by

$$\mathrm{h_1} = \frac{1}{2}(2\mathrm{s} + 2\mathrm{p_x} + 2\mathrm{p_y} + 2\mathrm{p_z}) \tag{62}$$

$$\mathrm{h_2} = \frac{1}{2}(2\mathrm{s} + 2\mathrm{p_x} - 2\mathrm{p_y} - 2\mathrm{p_z}) \tag{63}$$

$$\mathrm{h_3} = \frac{1}{2}(2\mathrm{s} - 2\mathrm{p_x} + 2\mathrm{p_y} - 2\mathrm{p_z}) \tag{64}$$

$$\mathrm{h_4} = \frac{1}{2}(2\mathrm{s} - 2\mathrm{p_x} - 2\mathrm{p_y} + 2\mathrm{p_z}) \tag{65}$$

The conversion from $\{2\mathrm{s}, 2\mathrm{p_x}, 2\mathrm{p_y}, 2\mathrm{p_z}\}$ to $\{\mathrm{h_1}, \mathrm{h_2}, \mathrm{h_3}, \mathrm{h_4}\}$ is performed by the matrix U defined by

$$\begin{pmatrix} \mathrm{h}_1 \\ \mathrm{h}_2 \\ \mathrm{h}_3 \\ \mathrm{h}_4 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2\mathrm{s} \\ 2\mathrm{p}_\mathrm{x} \\ 2\mathrm{p}_\mathrm{y} \\ 2\mathrm{p}_\mathrm{z} \end{pmatrix} \equiv U \begin{pmatrix} 2\mathrm{s} \\ 2\mathrm{p}_\mathrm{x} \\ 2\mathrm{p}_\mathrm{y} \\ 2\mathrm{p}_\mathrm{z} \end{pmatrix}$$

I show in Cell 3 of `WorkBook10_TransformationTheory.nb` that U does the job (i.e. gives Eqs. 62–65) and is unitary (the absolute values of its eigenvalues are equal to 1). Since the set $\{2\mathrm{s}, 2\mathrm{p}_\mathrm{x}, 2\mathrm{p}_\mathrm{y}, 2\mathrm{p}_\mathrm{z}\}$ is orthonormal and U is unitary, $\{\mathrm{h}_1, \mathrm{h}_2, \mathrm{h}_3, \mathrm{h}_4\}$ is orthonormal. Either set can be used for calculating the properties of the atom. However, h_1, h_2, h_3, and h_4 have a special property: they point towards the corners of a tetrahedron that has the carbon atom at its center. It so happens that the H atoms in CH_4 are also located at the corners of a tetrahedron with the C atom in the center. The orbitals h_1, h_2, h_3, h_4 are much better for building a theory of bonding in methane than are $2\mathrm{s}$, $2\mathrm{p}_\mathrm{x}$, $2\mathrm{p}_\mathrm{y}$, $2\mathrm{p}_\mathrm{z}$.

This statement should puzzle you. We spent quite a bit of time proving that two representations using two basis sets connected by a unitary transformation are completely equivalent. How, then, can one set be better than the other? The difference is qualitative. The hybrid orbitals "explain" why carbon atoms form tetrahedral bonds. The other, equivalent, basis sets do not. However, if the structure of methane is calculated with any of the equivalent basis sets, the geometry of the molecule and its energy will not change.

While hybrid orbitals are convenient for explaining the coordination of the carbon atoms, they are very cumbersome to use to explain spectroscopy.

Exercise 6 Find the matrix that converts $\{Y_0^0, Y_1^1, Y_1^0, Y_1^{-1}\}$ into $\{\mathrm{h}_1$, h_2, h_3, $\mathrm{h}_4\}$. Prove that it is unitary. Explain the consequences of this fact.

Exercise 7 Find the eigenvalues of $\hat{L}_z$ by using the basis set $\{\mathrm{h}_1$, h_2, h_3, $\mathrm{h}_4\}$.