

## Hermitian operators for observables

Given any observable  $\mathcal{O}$ , we can define a Hermitian operator  $\hat{\mathcal{O}}$  associated with it, by defining

$$\hat{\mathcal{O}}|\lambda_n\rangle = \lambda_n|\lambda_n\rangle \quad (1)$$

for the eigenstates of  $\mathcal{O}$  and using linearity to define the action on any other state. By our definition, the eigenvectors of  $\hat{\mathcal{O}}$  are orthonormal and the eigenvalues are real, so the operator is Hermitian.

We will see below what the physical interpretation of the state  $\hat{\mathcal{O}}|\Psi\rangle$  is; for now we point out the useful result that the expectation value of  $\mathcal{O}$  defined above can be calculated for the state  $\Psi$  as

$$\bar{\mathcal{O}} = \langle\Psi|\hat{\mathcal{O}}|\Psi\rangle. \quad (2)$$

This can be checked using the definition (??) to reproduce the result (??).

## Commuting vs noncommuting observables

For two different observables  $\mathcal{O}_1$  and  $\mathcal{O}_2$ , it is typically the case that the eigenvectors of one are not eigenvectors of the other. Physically, this means that a state with a definite value for  $\mathcal{O}_1$  will not have a definite value for  $\mathcal{O}_2$  and vice versa.

On the other hand, if the two observables do share a common basis of eigenstates, it is possible to know the value of both observables at the same time.

It is not hard to show that two Hermitian operators will share a common basis of eigenstates if and only if the corresponding operators *commute* with each other, i.e. that in acting with the two operators in succession on any state, we get the same result regardless of the order. This will be true if and only if the *commutator*  $[\hat{\mathcal{O}}_1, \hat{\mathcal{O}}_2] \equiv \hat{\mathcal{O}}_1\hat{\mathcal{O}}_2 - \hat{\mathcal{O}}_2\hat{\mathcal{O}}_1$  is equal to zero.

In the case where the operators share a common basis, both  $\mathcal{O}_1$  and  $\mathcal{O}_2$  will have matrix representations that are diagonal in this basis, so it is clear that the corresponding matrices also commute with each other.

The fact that some observables cannot have definite values simultaneously leads directly to the idea of an *uncertainty principle* in quantum mechanics. Starting from the definition (??) of uncertainty and making use of the Cauchy-Schwarz inequality (see Griffiths for details) it is possible to prove the *generalized uncertainty principle*

$$\Delta\mathcal{O}_1\Delta\mathcal{O}_2 \geq \frac{1}{2}|\langle i[\mathcal{O}_1, \mathcal{O}_2]\rangle|. \quad (3)$$

We will see that this leads to the more familiar Heisenberg Uncertainty Principle when applied to the position and momentum operators.

## Unitary operators

Another type of operator that will be important in quantum mechanics is the *unitary operator*. These are operators that preserve the inner product (and therefore all lengths and distances). Preserving the inner product means that if  $|v\rangle$  and  $|w\rangle$  are two vectors, and if  $\hat{U}$  is an operator, then the inner product between  $\hat{U}|v\rangle$  and  $\hat{U}|w\rangle$  is the same as the inner product between  $|v\rangle$  and  $|w\rangle$ . This gives the condition

$$\langle v|\mathcal{U}^\dagger\mathcal{U}|w\rangle = \langle v|w\rangle . \quad (4)$$

If this is true for all states, it means that the matrix representation of  $\mathcal{U}^\dagger\mathcal{U}$  in any basis is the same as the matrix representation of the identity operator  $\mathbf{1}$ , so they must be the same operator:

$$\mathcal{U}^\dagger\mathcal{U} = \mathbf{1} . \quad (5)$$

We can think of a unitary operator as a complex generalization of a rotation: unitary operators take an orthonormal basis to another orthonormal basis, and any operator with this property is unitary.

### Relation between unitary operators and Hermitian operators

The simplest unitary operator is the identity operator  $\mathbf{1}$ . If we consider a family of unitary operators  $\hat{U}(\epsilon)$  with  $\hat{U}(0) = \mathbf{1}$ , then for very small  $\epsilon$ , we can write

$$\hat{U}(\epsilon) = \mathbf{1} + \epsilon\hat{\mathcal{B}} + \dots \quad (6)$$

where the dots indicate terms at higher orders in epsilon.<sup>1</sup>

If the relation

$$\hat{U}^\dagger(\epsilon)\hat{U}(\epsilon) = \mathbf{1} \quad (7)$$

holds for all  $\epsilon$ , the Taylor expansion of the left side must equal the Taylor expansion of the right side. Inserting the expansion (6) and looking at the terms of order  $\epsilon$ , we find that

$$\hat{\mathcal{B}}^\dagger + \hat{\mathcal{B}} = 0 . \quad (8)$$

Defining  $\mathcal{A} = i\hat{\mathcal{B}}$ , we see that the operator  $\mathcal{A}$  satisfies  $-\mathcal{A}^\dagger + \mathcal{A} = 0$ . Thus, if  $\mathcal{U}(\epsilon)$  represents a family of unitary operators near with

$$\hat{U}(\epsilon) = \mathbf{1} - i\epsilon\hat{\mathcal{A}} + \dots \quad (9)$$

---

<sup>1</sup>If you are not comfortable with Taylor expanding an operator, think of the matrix elements of that operator in some basis. Each of these is an ordinary function of  $\epsilon$  which can be Taylor expanded. The operator  $\mathcal{B}$  is the one whose matrix elements are the terms at order  $\epsilon$  in the Taylor expansion of the matrix elements of  $\hat{U}(\epsilon)$ .

then  $\mathcal{A}$  must be Hermitian.

We can say that  $-i\hat{\mathcal{A}}|v\rangle$  gives the small change in the vector under an *infinitesimal* unitary transformation.

Given any Hermitian operator, we can also go the other direction and define a family of unitary operators “generated” by this operator. From the discussion above, the operator

$$(1 - i\epsilon\mathcal{A}) \tag{10}$$

should be almost a unitary operator when  $\epsilon$  is very small. It’s not quite, because we need to add certain higher order terms in  $\epsilon$  for (??) to be satisfied. However, we can generate a family of unitary operators by the following trick. We consider a limit where  $\epsilon \rightarrow 0$  so that we can ignore any higher order terms, but we also apply the operator multiple times, so that we don’t just end up with  $\mathbf{1}$ . Specifically, we define

$$\hat{\mathcal{U}}_{\mathcal{A}}(a) = \lim_{N \rightarrow \infty} (1 - i\frac{a}{N}\hat{\mathcal{A}})^N \tag{11}$$

This limit is well known as one definition of the exponential function, so we can write simply

$$\hat{\mathcal{U}}_{\mathcal{A}}(a) = e^{-ia\hat{\mathcal{A}}}. \tag{12}$$

We can think of the exponential of an operator as being defined by the power series

$$e^x = 1 + x + \frac{1}{2}x^2 + \dots \tag{13}$$

or we can work in a basis where the matrix  $\mathcal{A}$  is diagonal with eigenvalues  $\lambda_n$  and say that in this basis, the matrix for  $\hat{\mathcal{U}}_{\mathcal{A}}(a)$  is diagonal with matrix elements  $e^{-ia\lambda_n}$ .

In summary, we can associate a Hermitian operator to any one-parameter family of unitary operators near the identity operator, and we can associate a family of unitary operators to any Hermitian operator by exponentiation.

## Physical transformations as unitary operators

In quantum mechanics, physical transformations such as rotations, translations, and time evolution correspond to maps that take states to other states. These maps should be *linear* and *preserve the inner product*, that is they should act as unitary operators.

The linearity property means that if we have a superposition of two states, the transformation of this state will just be the superposition of the transformed states. For example, if we have the state  $\frac{1}{2}|\uparrow\rangle + \frac{1}{2}|\downarrow\rangle$  for a spin half particle, and we rotate by  $\pi/2$  about the  $y$  axis so that  $|\uparrow\rangle$  goes to  $|\rightarrow\rangle$  and  $|\downarrow\rangle$  goes to  $|\leftarrow\rangle$ , the state  $\frac{1}{2}|\uparrow\rangle + \frac{1}{2}|\downarrow\rangle$  should be transformed to  $\frac{1}{2}|\rightarrow\rangle + \frac{1}{2}|\leftarrow\rangle$ . This seems very plausible, but in the end, it is just a basic

assumption of quantum mechanics. Mathematically, it means that physical transformations are represented by linear operators.

Preserving the inner product is equivalent to the requirement that a normalized state  $|\Psi\rangle$  should transform to another normalized state. This is essentially the statement that probabilities must be conserved: given any orthonormal basis corresponding to a physical observable, if the norm of a state decreased after some physical transformation, the probabilities for the various possible outcomes of a measurement would no longer sum to 1.

## Infinitesimal physical transformations

Physical transformations such as rotations or translations come in continuous families. For example, we can rotate about an axis by any angle  $\theta$  or translate in the  $x$  direction by any distance  $a$ . For a quantum system, these families of transformations correspond to families of unitary operators. If  $\hat{\mathcal{T}}(a)$  is such a family of operators depending on some parameter  $a$ , with  $\hat{\mathcal{T}}(0) = \mathbf{1}$ , i.e.  $a = 0$  corresponds to doing nothing, then it is useful to consider the possible *infinitesimal transformations*, that is, the transformations for infinitesimal values of  $a$ . For example, we can consider a rotation by an infinitesimal angle or translation by an infinitesimal distance. These are important, because many more general transformations can be built up by doing a series of these infinitesimal transformations.<sup>2</sup>

According the general discussion of unitary operators above, any infinitesimal transformation can be expressed as

$$\hat{\mathcal{T}}(\epsilon) = \mathbf{1} - i\epsilon\hat{\mathcal{O}} + \dots \quad (14)$$

where  $\hat{\mathcal{O}}$  is a hermitian operator. This can be associated some physical observable  $\mathcal{O}$ . Conversely, given any physical observable  $\mathcal{O}$ , there is an associated Hermitian operator  $\hat{\mathcal{O}}$ . From this, we can define an infinitesimal physical transformation as in (??). Applying this transformation many times, we generate a family of physical transformations described by the unitary operators (see discussion in section 1)

$$\hat{\mathcal{T}}(a) = e^{-ia\hat{\mathcal{O}}} . \quad (15)$$

Thus we see that there is a one-to-one correspondence between infinitesimal physical transformations (or the families of transformations that they generate) and physical observables. As we will see below, the connection is the familiar(?) one from classical mechanics:

- energy  $\leftrightarrow$  time translation/evolution
- momentum in direction  $\hat{n} \leftrightarrow$  translation in direction  $\hat{n}$

---

<sup>2</sup>There are some physical transformations, such as mirror reflection, that cannot be built up from a series of infinitesimal transformations. These transformations are often referred to as *discrete*.

- angular momentum around the axis in direction  $\hat{n} \leftrightarrow$  rotations about the axis in direction  $\hat{n}$

Normally this connection is discussed as the connection between symmetries and conserved quantities. But only some physical quantities will be conserved and only some transformations are symmetries. We will now see that in quantum mechanics, the physical observables that are conserved quantities are exactly the ones associated with physical transformations that are symmetries.