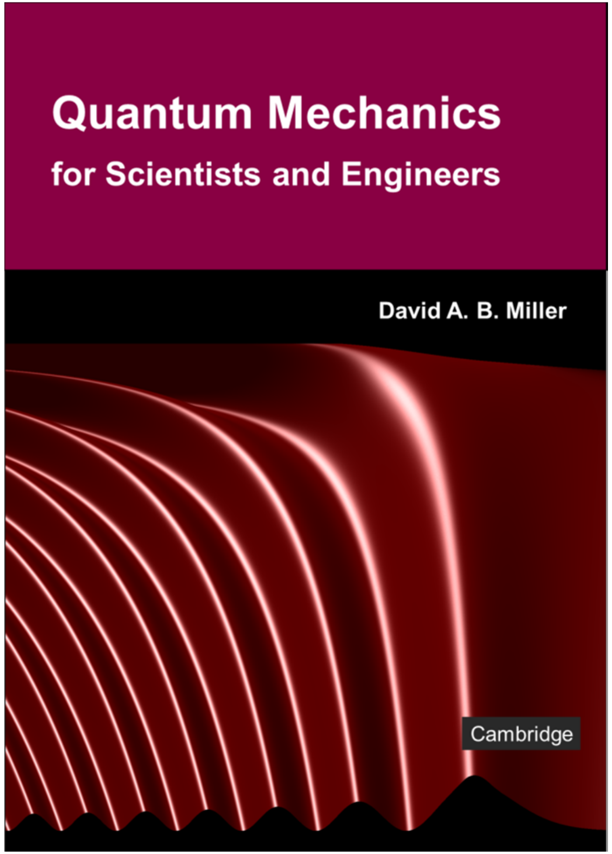
5.3 Vector spaces, operators and matrices

Slides: Video 5.3.1 Vector space

Text reference: Quantum Mechanics for Scientists and Engineers

Section 4.2



Vector spaces, operators and matrices

Vector space

Quantum mechanics for scientists and engineers

David Miller

Vector space

We need a “space” in which our vectors exist

- For a vector with three components $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$
  - we imagine a three dimensional Cartesian space

The vector can be visualized as a line

- starting from the origin
  - with projected lengths a_1, a_2, and a_3 along the x, y, and z axes respectively
    - with each of these axes being at right angles

Vector space

For a function expressed as its value at a set of points

- instead of 3 axes labeled x, y, and z
  - we may have an infinite number of orthogonal axes
    - labeled with their associated basis function
      - e.g., ψ_n

Just as we label axes in conventional space with unit vectors

- one notation is $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ for the unit vectors
- so also here we label the axes with the kets $|\psi_n\rangle$
  - Either notation is acceptable

Mathematical properties – existence of inner product

Geometrical space has a vector dot product

that defines both the orthogonality of the axes

$$\hat{\mathbf{x}} \cdot \hat{\mathbf{y}} = 0$$

and the components of a vector along those axes

$\mathbf{f} = f_x \hat{\mathbf{x}} + f_y \hat{\mathbf{y}} + f_z \hat{\mathbf{z}}$ with $f_x = \mathbf{f} \cdot \hat{\mathbf{x}}$

and similarly for the other components

Our vector space has an inner product that defines both

the orthogonality of the basis functions

$$\langle \psi_m | \psi_n \rangle = \delta_{nm}$$

as well as the components $c_m = \langle \psi_m | f \rangle$

Mathematical properties – addition of vectors

With respect to addition of vectors

both geometrical space and our vector space are

commutative

$$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$$

$$|f\rangle+|g\rangle=|g\rangle+|f\rangle$$

and associative

$$\mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c}$$

$$|f\rangle+(|g\rangle+|h\rangle)=(|f\rangle+|g\rangle)+|h\rangle$$

Mathematical properties - linearity

Both the geometrical space and our vector space are

linear in multiplying by constants

$$c(\mathbf{a}+\mathbf{b}) = c\mathbf{a} + c\mathbf{b}$$

our constants may be complex

$$c\left(|f\rangle + |g\rangle\right) = c|f\rangle + c|g\rangle$$

And the inner product is linear

both in multiplying by constants

$$\mathbf{a}\cdot(c\mathbf{b}) = c(\mathbf{a}\cdot\mathbf{b})$$

$$\langle f|cg\rangle = c\langle f|g\rangle$$

and in superposition of vectors

$$\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c}$$

$$\langle f|\left(|g\rangle + |h\rangle\right) = \langle f|g\rangle + \langle f|h\rangle$$

Mathematical properties – norm of a vector

There is a well-defined "length" to a vector

formally a "norm"

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$$

$$\| f \| = \sqrt{\langle f \mid f \rangle}$$

Mathematical properties – completeness

- In both cases
  - any vector in the space
    - can be represented to an arbitrary degree of accuracy
      - as a linear combination of the basis vectors
        - This is the completeness requirement on the basis set
- In vector spaces
  - this property of the vector space itself is sometimes described as “compactness”

Mathematical properties – commutation and inner product

In geometrical space, the lengths a_x, a_y, and a_z of a vector's components are real

so the inner product (vector dot product) is commutative

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$

But with complex coefficients rather than real lengths

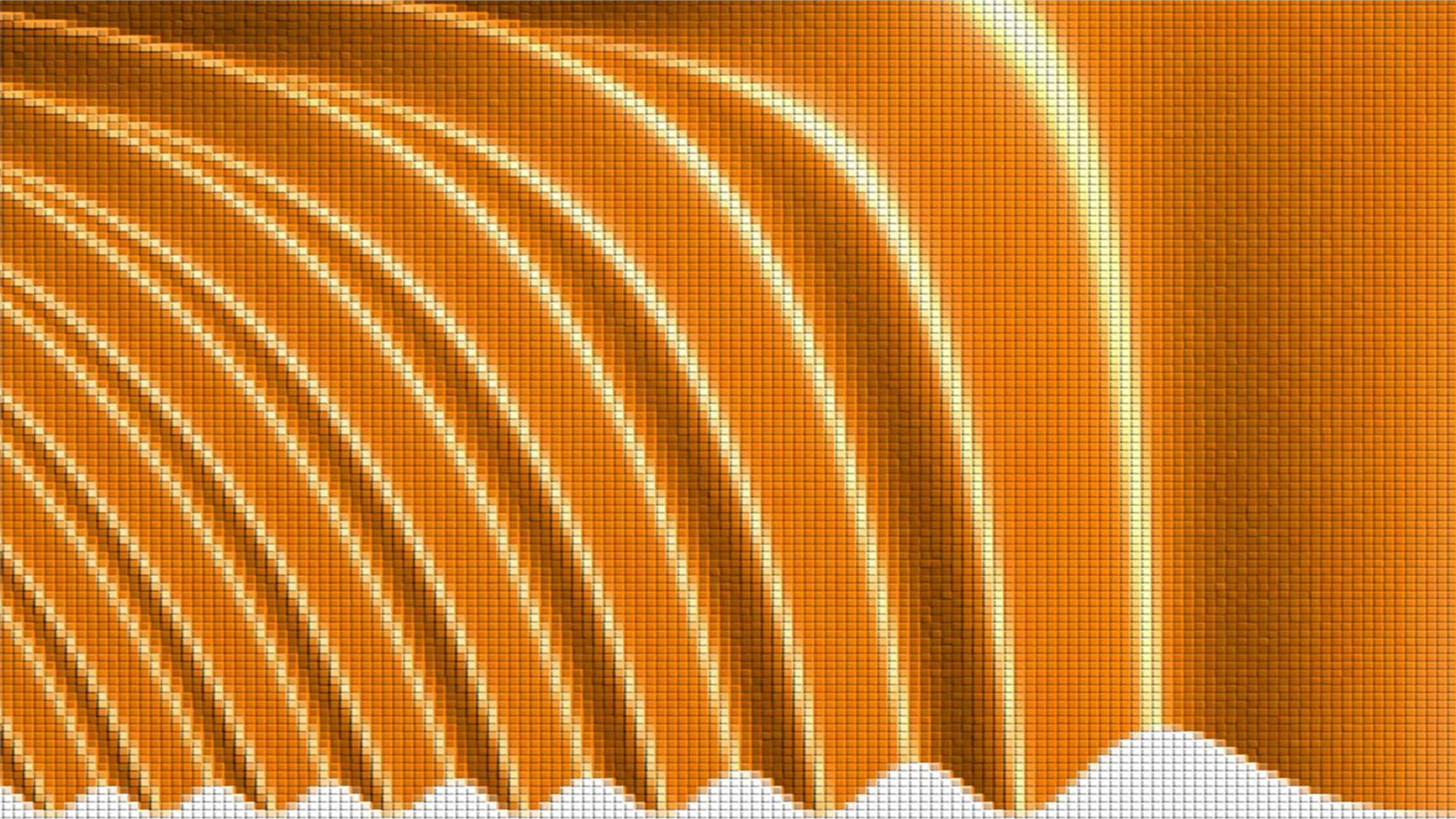
we choose a non-commutative inner product of the form

$$\langle f | g \rangle = \left(\langle g | f \rangle \right)^*$$

This ensures that $\langle f | f \rangle$ is real

even if we work with complex numbers

as required for it to form a useful norm

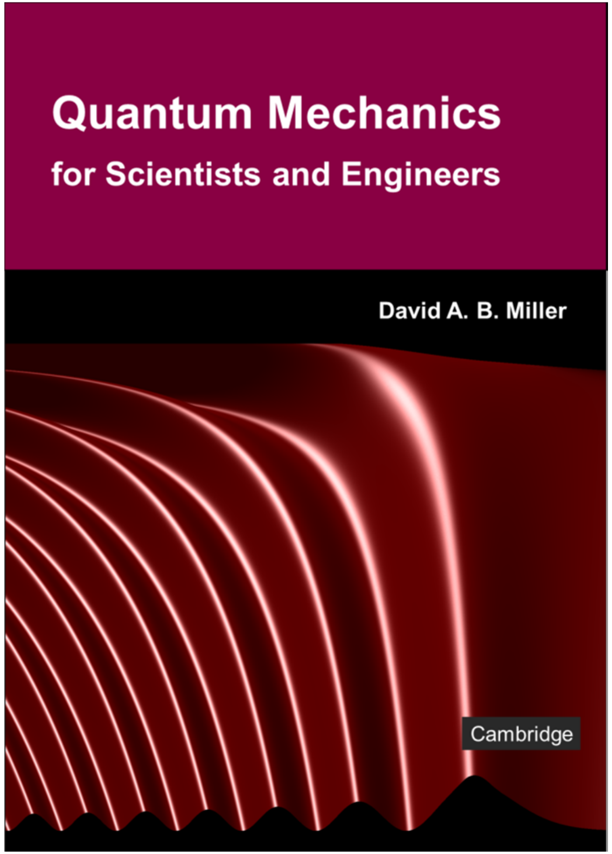


5.3 Vector spaces, operators and matrices

Slides: Video 5.3.3 Operators

Text reference: Quantum Mechanics for Scientists and Engineers

Sections 4.3 – 4.4



Vector spaces, operators and matrices

Operators

Quantum mechanics for scientists and engineers

David Miller

Operators

A function turns one number
- the argument
  - into another
    - the result

An operator turns one function into another
- In the vector space representation of a function
  - an operator turns one vector into another

Operators

Suppose that we are constructing the new function $g(y)$

from the function $f(x)$

by acting on $f(x)$

with the operator $\hat{A}$

The variables x and y might be the same kind of variable

as in the case where the operator corresponds to differentiation of the function

$$g(x) = \left(\frac{d}{dx}\right) f(x)$$

Operators

The variables x and y might be quite different

as in the case of a Fourier transform operation where

x might represent time and

y might represent frequency

$$g(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-iyx)\, dx$$

A standard notation for writing any such operation on a function is

$$g(y) = \hat{A} f(x)$$

This should be read as $\hat{A}$ operating on $f(x)$

Operators

For $\hat{A}$ to be the most general operation possible

it should be possible for the value of $g(y)$

for example, at some particular value of $y = y_1$

to depend on the values of $f(x)$

for all values of the argument x

This is the case, for example, in the Fourier transform operation

$$g(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-iyx)\, dx$$

Linear operators

We are interested here solely in linear operators

They are the only ones we will use in quantum mechanics

because of the fundamental linearity of quantum mechanics

A linear operator has the following characteristics

$$\hat{A}\left[f(x)+h(x)\right]=\hat{A}f(x)+\hat{A}h(x)$$

$$\hat{A}\left[cf(x)\right]=c\hat{A}f(x)$$

for any complex number c

Consequences of linearity for operators

Let us consider the most general way we could have the function $g(y)$

at some specific value y_1 of its argument

that is, $g(y_1)$

be related to the values of $f(x)$

for possibly all values of x

and still retain the linearity properties for this relation

Consequences of linearity for operators

Think of the function $f(x)$

as being represented by a list of values

$f(x_1)$, $f(x_2)$, $f(x_3)$, … ,

just as we did when considering $f(x)$ as a vector

We can take the values of x to be as closely spaced as we want

We believe that this representation can give us as accurate a representation of $f(x)$

for any calculation we need to perform

Consequences of linearity for operators

Then we propose that

for a linear operation

the value of $g(y_1)$

might be related to the values of $f(x)$

by a relation of the form

$$g(y_1) = a_{11} f(x_1) + a_{12} f(x_2) + a_{13} f(x_3) + \dots$$

where the a_{ij} are complex constants

Consequences of linearity for operators

This form $g(y_1) = a_{11} f(x_1) + a_{12} f(x_2) + a_{13} f(x_3) + \dots$

shows the linearity behavior we want

If we replaced $f(x)$ by $f(x) + h(x)$

then we would have

$$g(y_1) = a_{11}\left[f(x_1) + h(x_1) \right] + a_{12}\left[f(x_2) + h(x_2) \right] + a_{13}\left[f(x_3) + h(x_3) \right] + \dots$$

$$= a_{11} f(x_1) + a_{12} f(x_2) + a_{13} f(x_3) + \dots$$

$$+ a_{11} h(x_1) + a_{12} h(x_2) + a_{13} h(x_3) + \dots$$

as required for a linear operator relation from

$$\hat{A}\left[f(x) + h(x) \right] = \hat{A} f(x) + \hat{A} h(x)$$

Consequences of linearity for operators

And, in this form $g(y_1) = a_{11} f(x_1) + a_{12} f(x_2) + a_{13} f(x_3) + \ldots$

if we replaced $f(x)$ by $cf(x)$

then we would have

$$g(y_1) = a_{11} cf(x_1) + a_{12} cf(x_2) + a_{13} cf(x_3) + \ldots$$

$$= c\left[a_{11} f(x_1) + a_{12} f(x_2) + a_{13} f(x_3) + \ldots \right]$$

as required for a linear operator relation from

$$\hat{A}\left[cf(x) \right] = c\hat{A} f(x)$$

Consequences of linearity for operators

Now consider whether this form

$$g(y_1) = a_{11} f(x_1) + a_{12} f(x_2) + a_{13} f(x_3) + \dots$$

is as general as it could be and still be a linear relation

We can see this by trying to add other powers and "cross terms" of $f(x)$

Any more complicated relation of $g(y_1)$ to $f(x)$

could presumably be written as a power series in $f(x)$

possibly involving $f(x)$

for different values of x

that is, "cross terms"

Consequences of linearity for operators

If we were to add higher powers of $f(x)$

such as $\left[f(x)\right]^2$

or cross terms such as $f(x_1)f(x_2)$

into the series $g(y_1) = a_{11}f(x_1) + a_{12}f(x_2) + a_{13}f(x_3) + \ldots$

it would no longer have the required linear behavior of

$$\hat{A}\left[f(x) + h(x)\right] = \hat{A}f(x) + \hat{A}h(x)$$

We also cannot add a constant term to this series

That would violate the second linearity condition

$$\hat{A}\left[cf(x)\right] = c\hat{A}f(x)$$

The additive constant would not be multiplied by c

Generality of the proposed linear operation

Hence we conclude

$$g\left(y_1\right)=a_{11}f\left(x_1\right)+a_{12}f\left(x_2\right)+a_{13}f\left(x_3\right)+\ldots$$

is the most general form possible

for the relation between $g\left(y_1\right)$

and $f\left(x\right)$

if this relation is to correspond to a linear operator

Construction of the entire operator

To construct the entire function $g(y)$

we should construct series like

$$g(y_1) = a_{11} f(x_1) + a_{12} f(x_2) + a_{13} f(x_3) + \ldots$$

for each value of y

If we write $f(x)$ and $g(y)$ as vectors

then we can write all these series at once

$$\begin{bmatrix} g(y_1) \\ g(y_2) \\ g(y_3) \\ \vdots \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \end{bmatrix}$$

Construction of the entire operator

We see that $\begin{bmatrix} g(y_1) \\ g(y_2) \\ g(y_3) \\ \vdots \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \end{bmatrix}$

can be written as $g(y) = \hat{A} f(x)$

where the operator $\hat{A}$ can be written as a matrix

$$\hat{A} \equiv \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Bra-ket notation and operators

Presuming functions can be represented as vectors

then linear operators can be represented by matrices

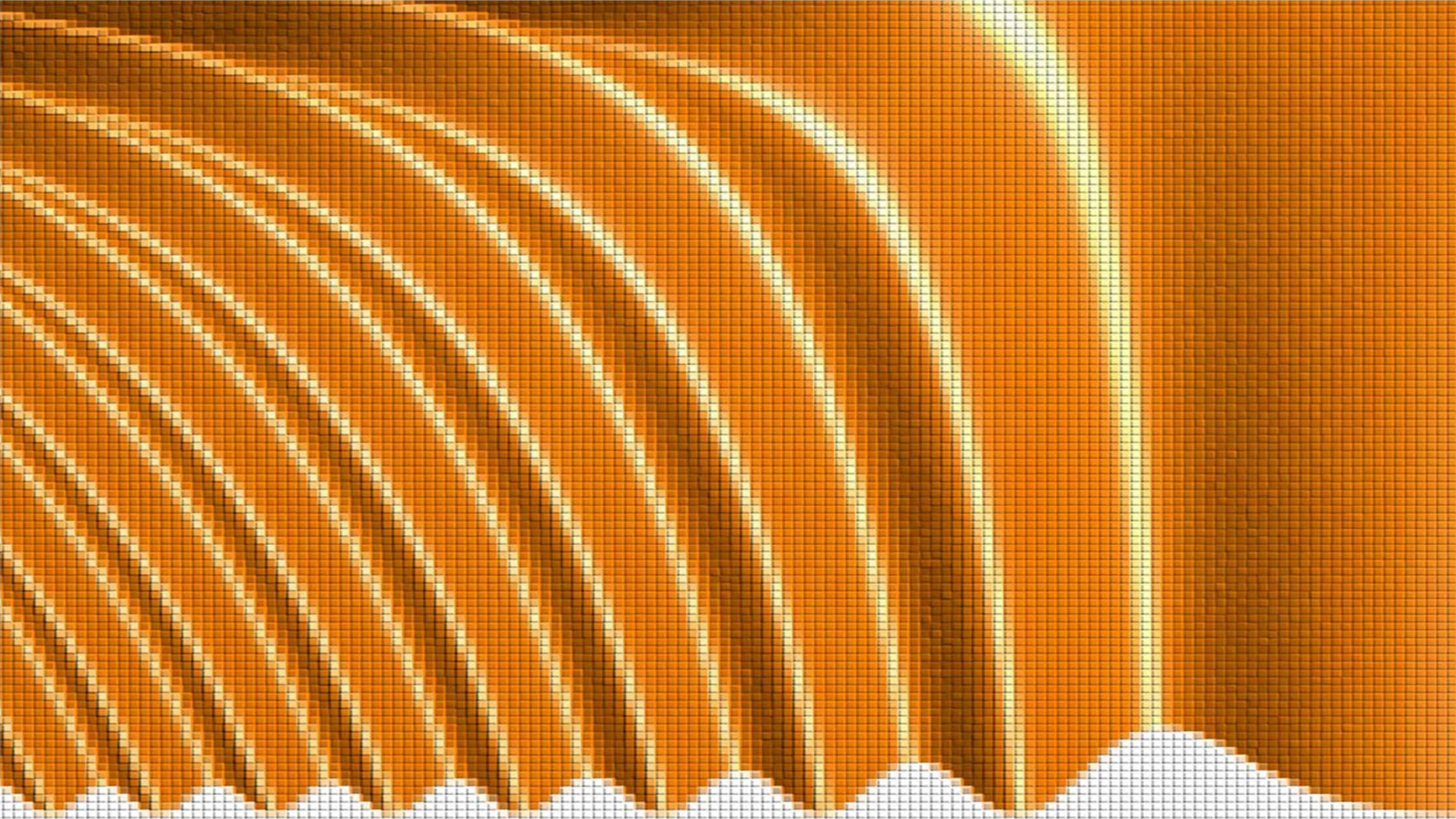
In bra-ket notation, we can write $g(y) = \hat{A} f(x)$ as

$$|g\rangle = \hat{A} | f \rangle$$

If we regard the ket as a vector

we now regard the (linear) operator $\hat{A}$

as a matrix

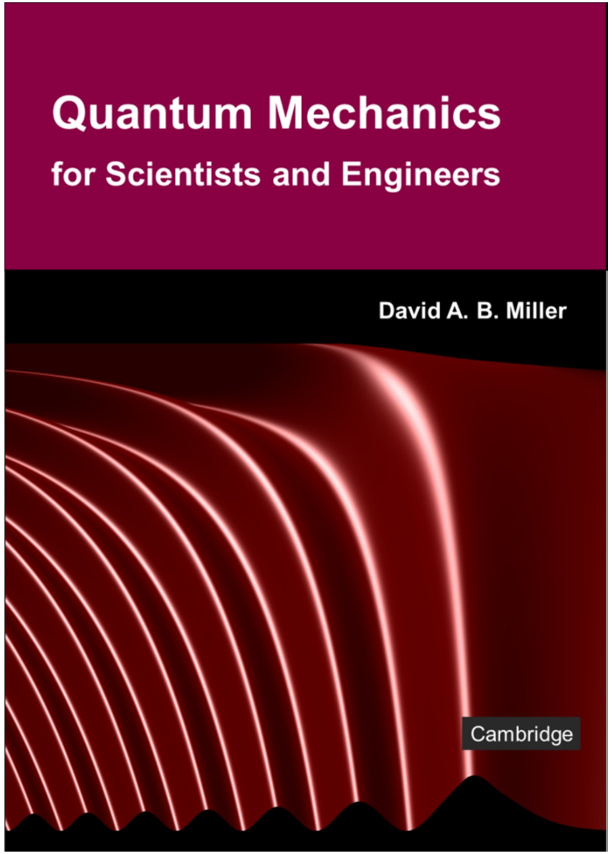


5.3 Vector spaces, operators and matrices

Slides: Video 5.3.5 Linear operators and their algebra

Text reference: Quantum Mechanics for Scientists and Engineers

Sections 4.4 – 4.5



Vector spaces, operators and matrices

Linear operators and their algebra

Quantum mechanics for scientists and engineers

David Miller

Consequences of linear operator algebra

Because of the mathematical equivalence of matrices and linear operators

the algebra for such operators

is identical to that of matrices

In particular

operators do not in general commute

$\hat{A}\hat{B}|f\rangle$ is not in general equal to $\hat{B}\hat{A}|f\rangle$

for any arbitrary $|f\rangle$

Whether or not operators commute

is very important in quantum mechanics

Generalization to expansion coefficients

We discussed operators

for the case of functions of position (e.g., x)

but we can also use expansion coefficients on basis sets

We expanded $f(x)=\sum_n c_n \psi_n(x)$ and $g(x)=\sum_n d_n \psi_n(x)$

We could have followed a similar argument

requiring each expansion coefficient d_i

depends linearly on all the expansion coefficients c_n

Generalization to expansion coefficients

By similar arguments

we would deduce the most general linear relation

between the vectors of expansion coefficients

could be represented as a matrix

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix}$$

The bra-ket statement of the relation between *f*, *g*, and $\hat{A}$ remains unchanged as $|g\rangle = \hat{A}|f\rangle$

Evaluating the matrix elements of an operator

Now we will find out how we can write some operator

as a matrix

That is, we will deduce how to calculate all the elements of the matrix

if we know the operator

Suppose we choose our function $f(x)$

to be the jth basis function $\psi_j(x)$

so $f(x) = \psi_j(x)$ or equivalently $|f\rangle = |\psi_j\rangle$

Evaluating the matrix elements of an operator

Then, in the expansion $f(x) = \sum_n c_n \psi_n(x)$

we are choosing $c_j = 1$

with all the other c's being 0

Now we operate on this $|f\rangle$ with $\hat{A}$

in $|g\rangle = \hat{A}|f\rangle$

to get $|g\rangle$

Suppose specifically

we want to know the resulting coefficient d_i

in the expansion $g(x) = \sum_n d_n \psi_n(x)$

Evaluating the matrix elements of an operator

From the matrix form of $|g\rangle = \hat{A}|f\rangle$

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix}$$

with our choice $c_j = 1$ and all other c's 0 then

we would have

$$d_i = A_{ij}$$

Evaluating the matrix elements of an operator

For example, for $j = 2$

that is, $c_2 = 1$ and all other c's 0 then

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{12} \\ A_{22} \\ A_{32} \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

so in this example

$$d_3 = A_{32}$$

Evaluating the matrix elements of an operator

But, from the expansions for $|f\rangle$ and $|g\rangle$

for the specific case of $|f\rangle = |\psi_j\rangle$

$$|g\rangle = \sum_n d_n |\psi_n\rangle = \hat{A}|f\rangle = \hat{A}|\psi_j\rangle$$

To extract d_i from this expression

we multiply by $\langle\psi_i|$ on both sides to obtain

$$d_i = \langle\psi_i|\hat{A}|\psi_j\rangle$$

But we already concluded for this case that $d_i = A_{ij}$

So $$A_{ij} = \langle\psi_i|\hat{A}|\psi_j\rangle$$

Evaluating the matrix elements of an operator

But our choices of i and j here were arbitrary

- So quite generally
  - when writing an operator $\hat{A}$ as a matrix
    - when using a basis set $|\psi_n\rangle$
      - the matrix elements of that operator are

$$\boxed{A_{ij} = \langle \psi_i | \hat{A} | \psi_j \rangle}$$

We can now turn any linear operator into a matrix

- For example, for a simple one-dimensional spatial case

$$A_{ij} = \int \psi_i^*(x) \hat{A} \psi_j(x) dx$$

Visualization of a matrix element

Operator $\hat{A}$

acting on the unit vector $|\psi_j\rangle$

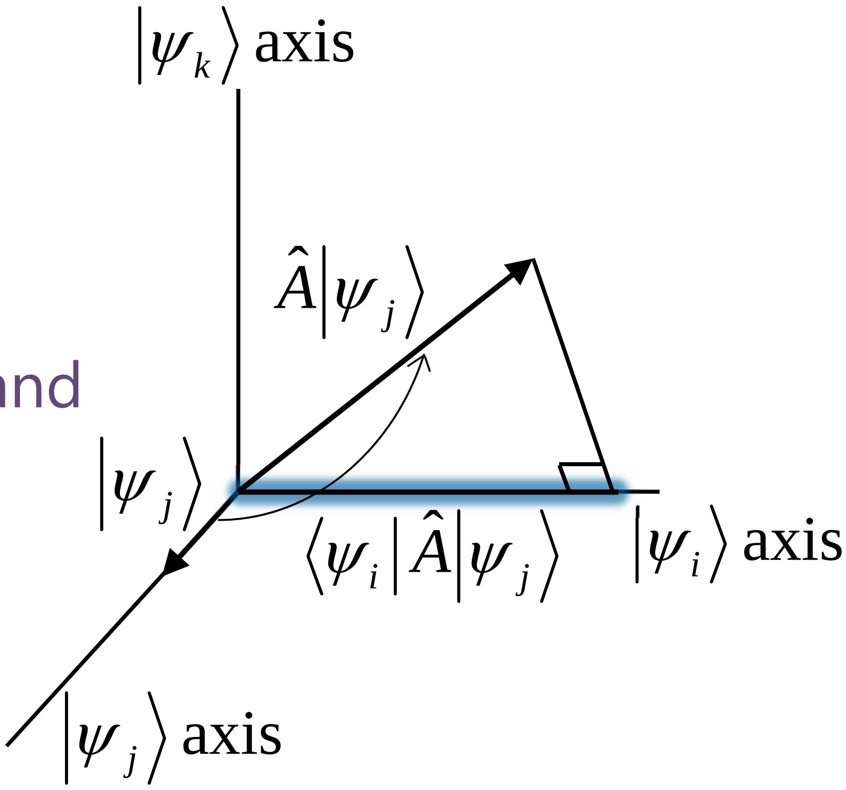
generates the vector $\hat{A}|\psi_j\rangle$

with generally a new length and direction

The matrix element $\langle\psi_i|\hat{A}|\psi_j\rangle$

is the projection of $\hat{A}|\psi_j\rangle$

onto the $|\psi_i\rangle$ axis



Evaluating the matrix elements

We can write the matrix for the operator $\hat{A}$

$$\hat{A} \equiv \begin{bmatrix} \langle \psi_1 | \hat{A} | \psi_1 \rangle & \langle \psi_1 | \hat{A} | \psi_2 \rangle & \langle \psi_1 | \hat{A} | \psi_3 \rangle & \cdots \\ \langle \psi_2 | \hat{A} | \psi_1 \rangle & \langle \psi_2 | \hat{A} | \psi_2 \rangle & \langle \psi_2 | \hat{A} | \psi_3 \rangle & \cdots \\ \langle \psi_3 | \hat{A} | \psi_1 \rangle & \langle \psi_3 | \hat{A} | \psi_2 \rangle & \langle \psi_3 | \hat{A} | \psi_3 \rangle & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

We have now deduced how to set up

a function as a vector and

a linear operator as a matrix

which can operate on the vectors