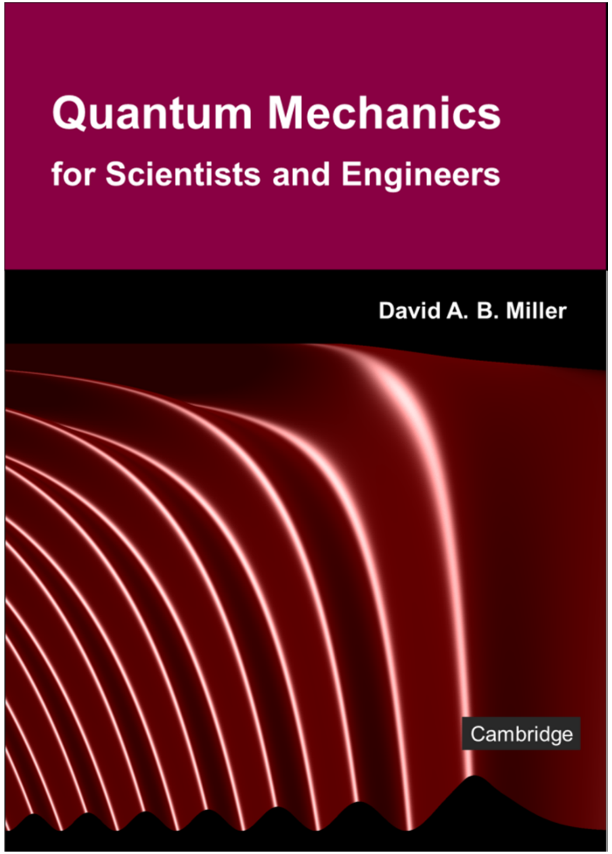
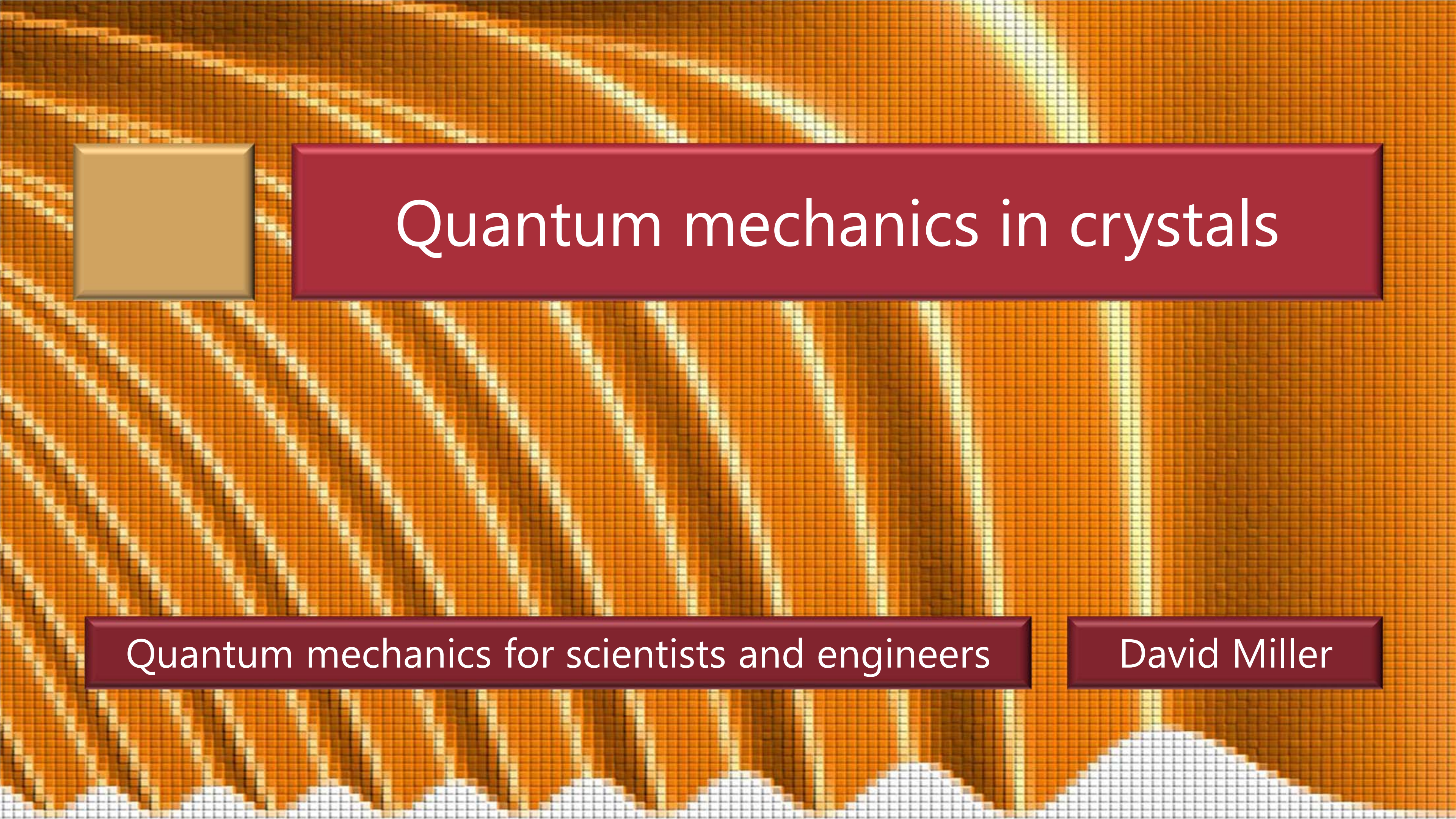
10.1 Quantum mechanics in crystals

Slides: Video 10.1.1 Introduction to quantum mechanics in crystals

Text reference: Quantum Mechanics for Scientists and Engineers

Chapter 8 Introduction

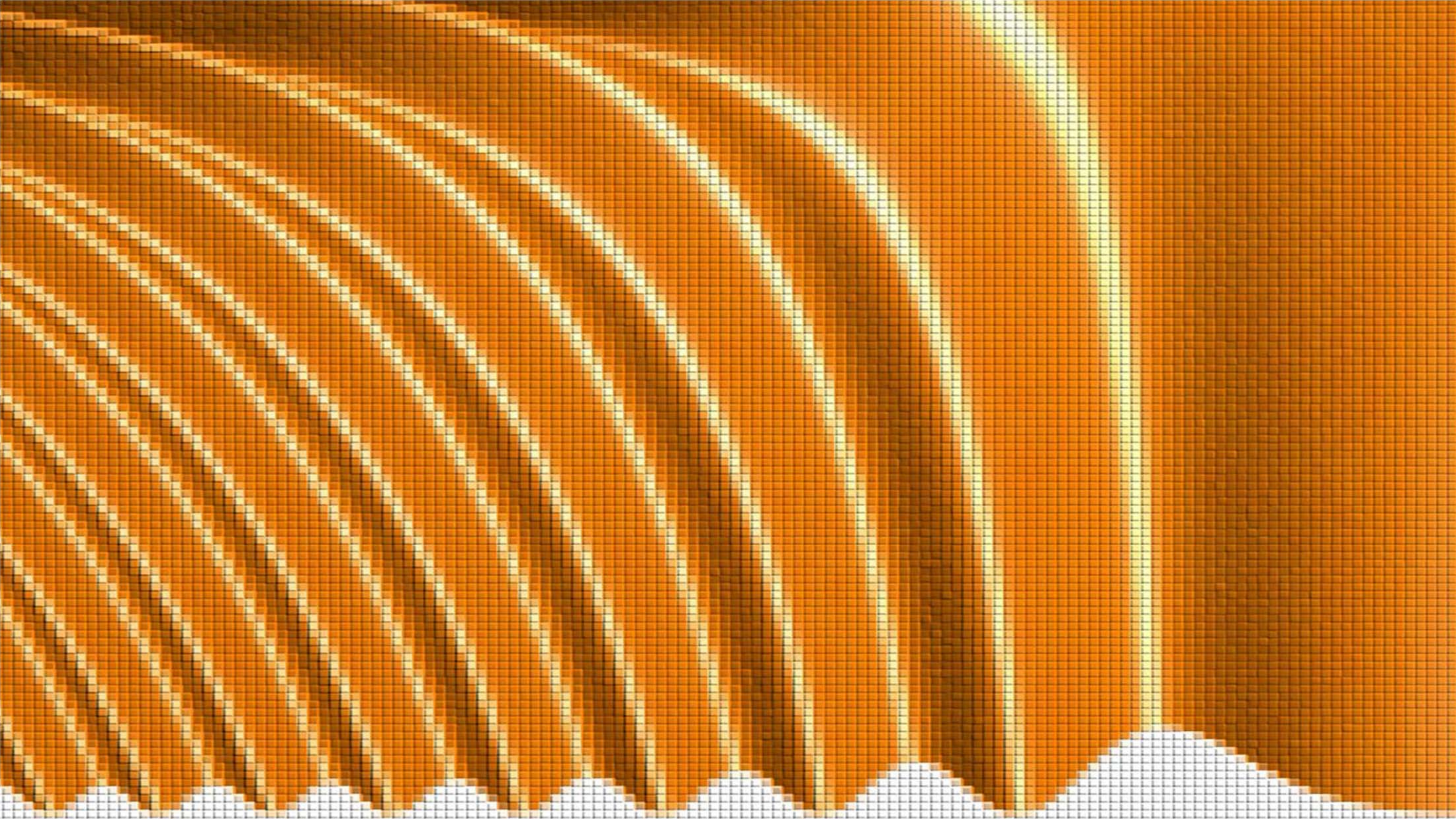




Quantum mechanics in crystals

Quantum mechanics for scientists and engineers

David Miller

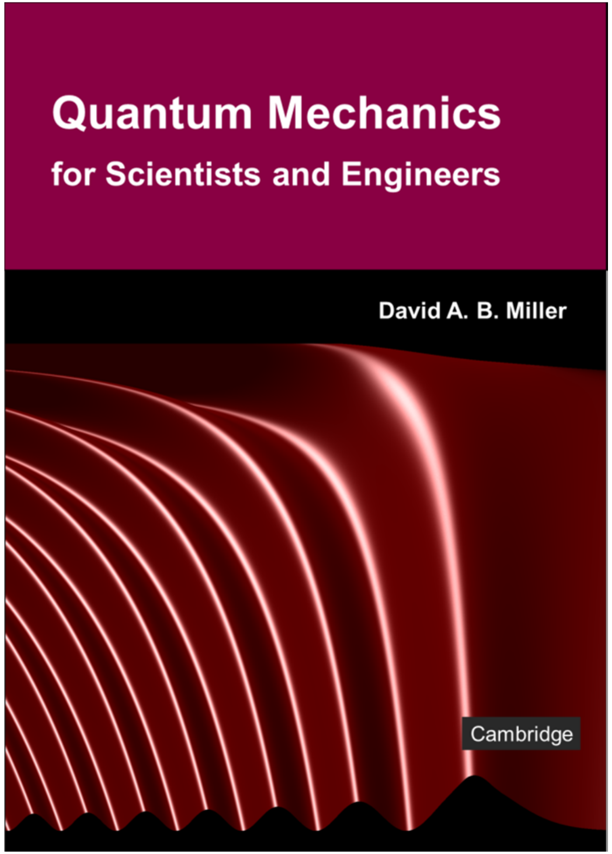


10.1 Quantum mechanics in crystals

Slides: Video 10.1.2 Crystal structures

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.1



Quantum mechanics in crystals

Crystal structures

Quantum mechanics for scientists and engineers

David Miller

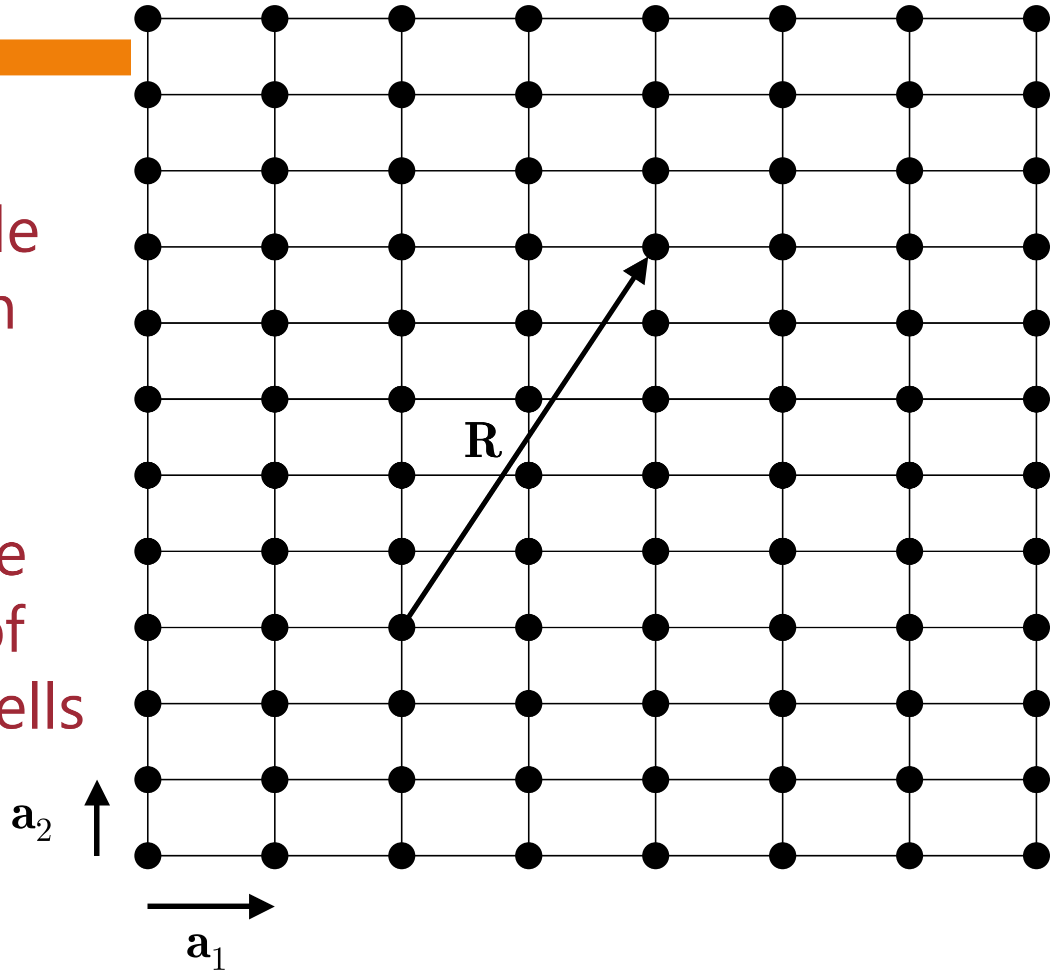
Crystal structures

Crystal

- material whose measurable properties are periodic in space

Crystal structure

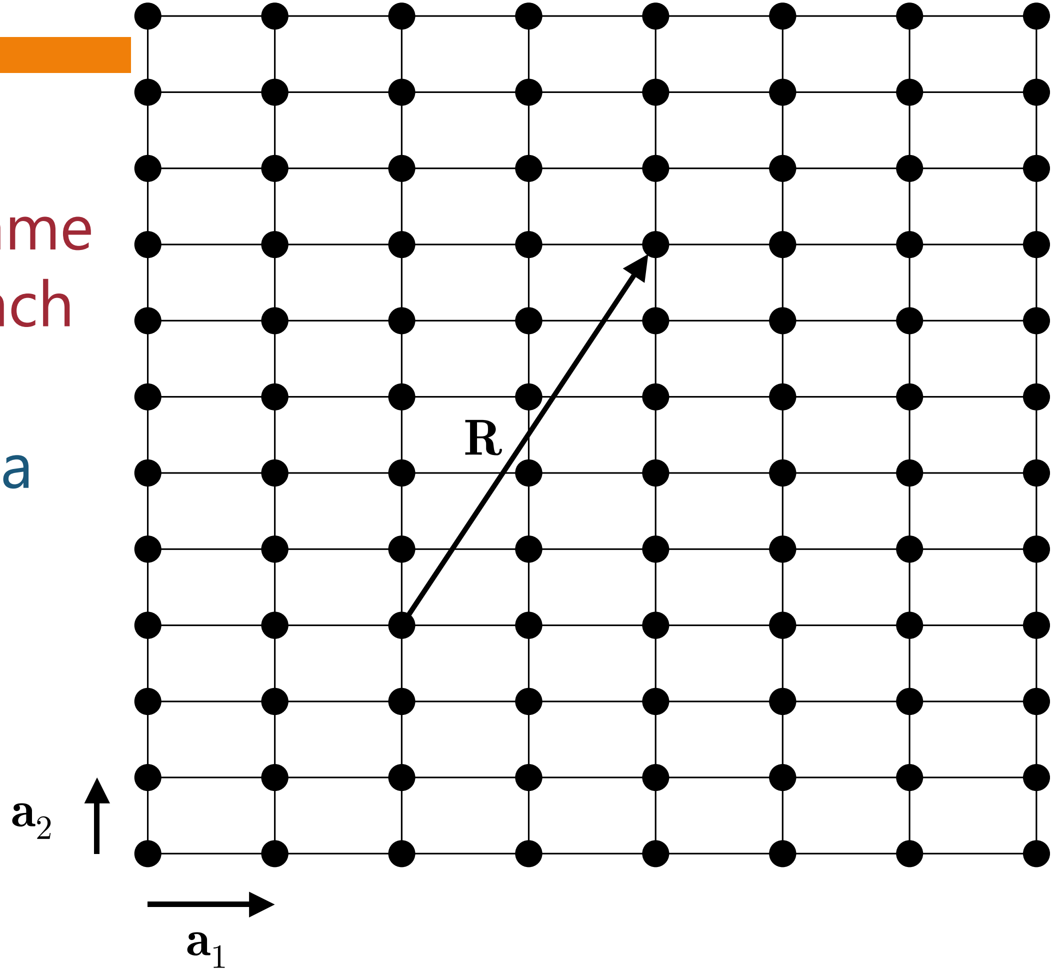
- is one that can fill all space by the regular stacking of identical blocks or unit cells



Crystal structures

Crystal lattice

- If we put a mark on the same spot on the surface of each block
  - these spots would form a crystal lattice



Lattice vectors

The set of lattice vectors

consists of all of the vectors

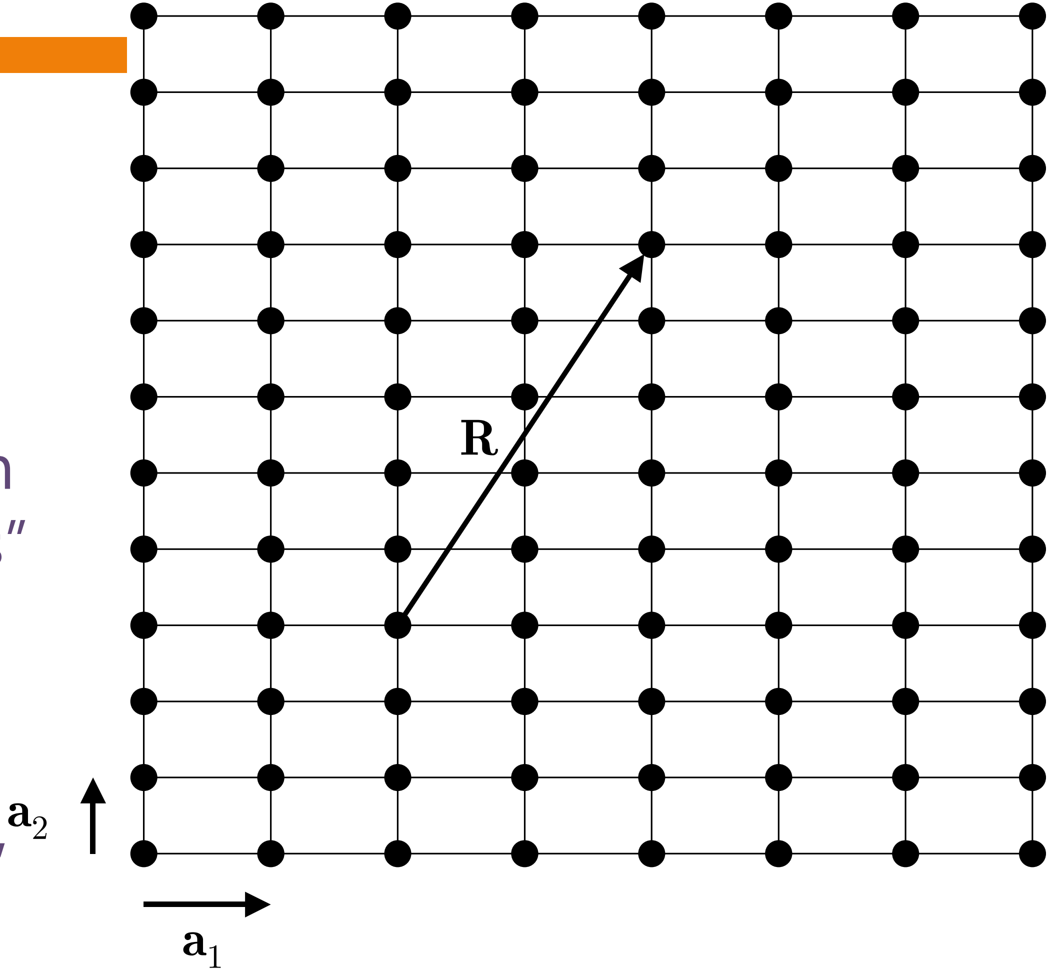
$$\mathbf{R}_L = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$$

with integer n_1, n_2, and n_3

linking the same point in two different "unit cells"

"Translating" (moving) by a lattice vector makes no difference

"translational symmetry"

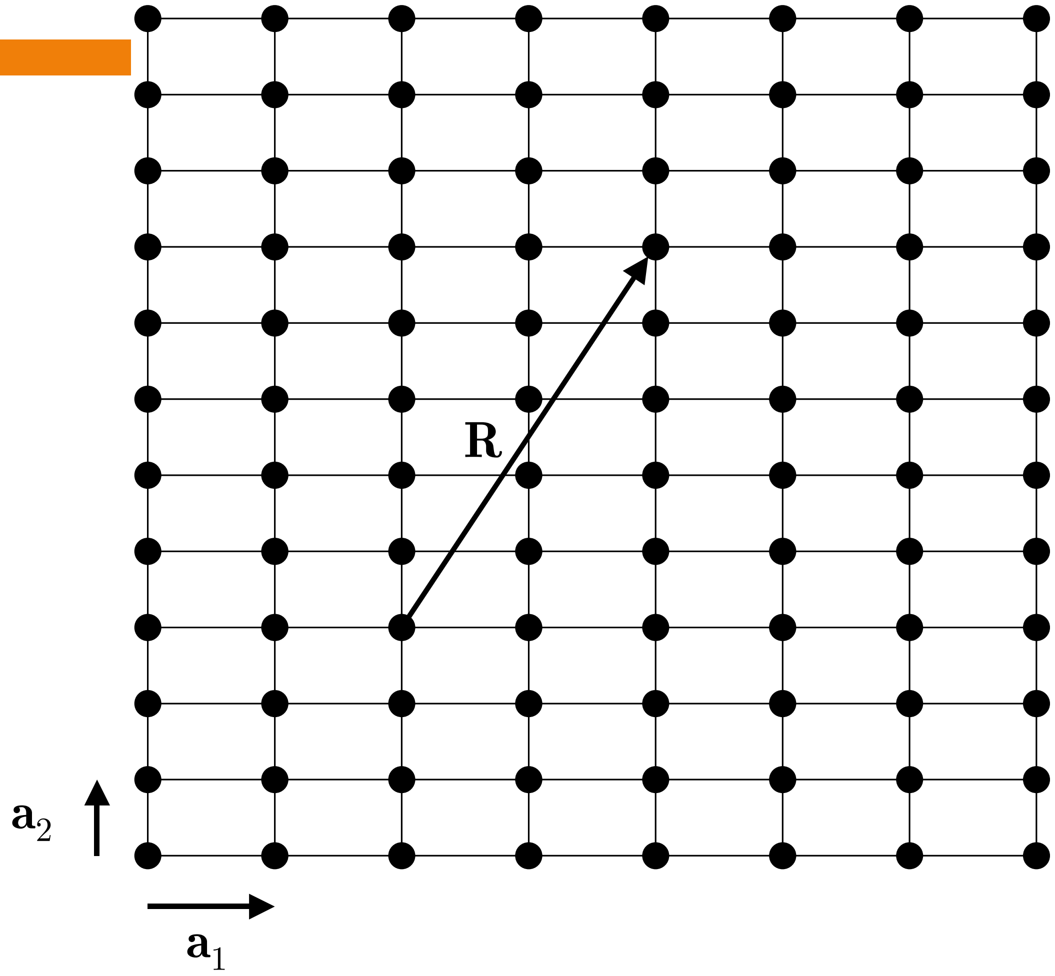


Lattice vectors

In

$$\mathbf{R}_L = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$$

$\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are the three linearly independent vectors that take us from a point in one unit cell to the equivalent point in the adjacent unit cell



Bravais lattices

In a given dimensionality, e.g., 1D, 2D, 3D

- there are only specific finite numbers of different types or symmetries of lattices of points
  - Bravais lattices
    - 1D – only one kind of lattice is possible
      - equally spaced points
    - 2D – 5 are possible
      - note no 5-sided object can be repeated to fill all space in a plane, for example
    - 3D – 14 are possible

Diamond and zinc-blende lattices

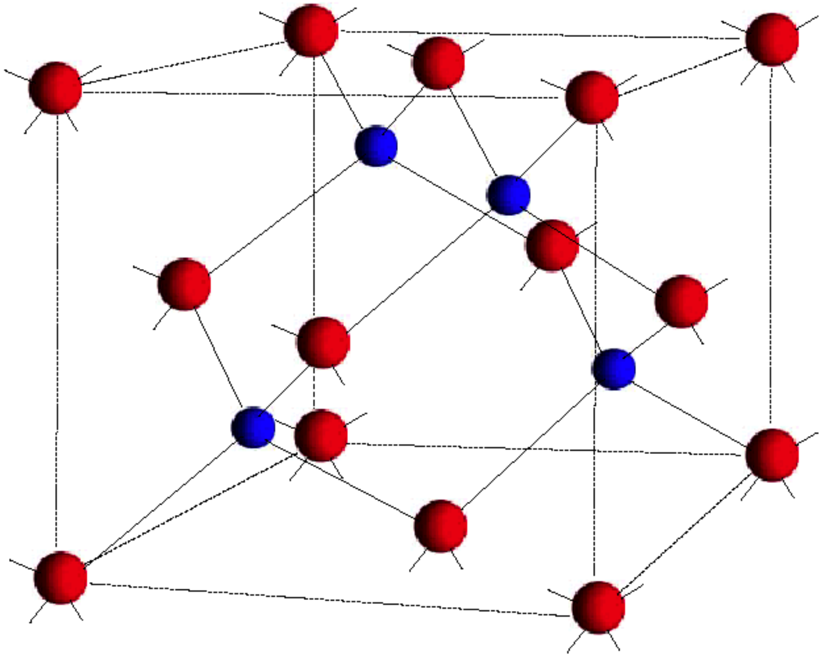
A large fraction of the semiconductor materials of practical interest

- such as silicon, germanium, and most of the III-V (e.g., GaAs) and II-VI (e.g., ZnSe) materials

have a specific form of cubic lattice

Diamond and zinc-blende lattices

- The Bravais lattice for these
  - is “face-centered cubic”
- The actual physical lattice
  - has two interlocking face-centered cubic lattices
    - which is the Bravais lattice with a pair of atoms associated with each lattice point



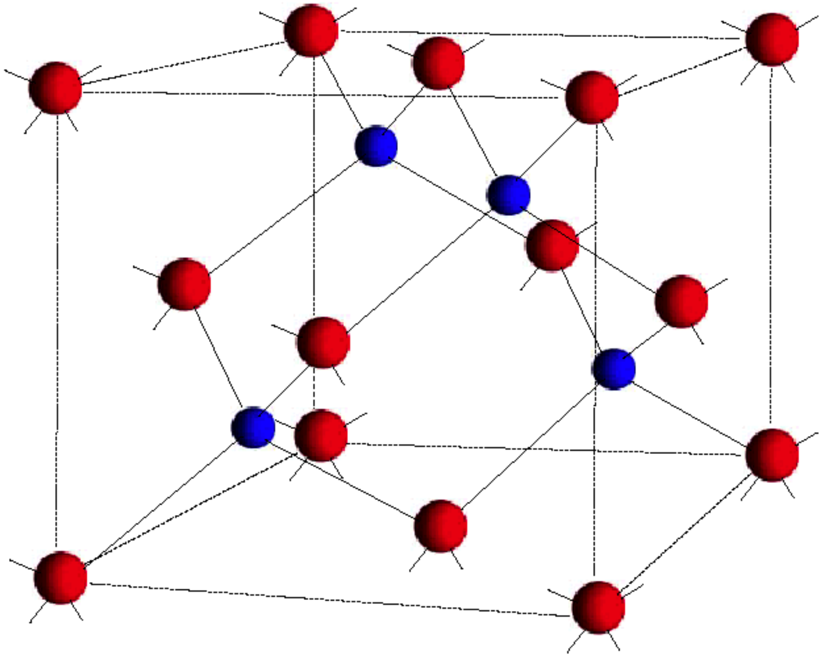
Diamond and zinc-blende lattices

"Zinc-blende" is the crystal structure for

most III-V and II-VI materials

The group III (or II) atoms lie on one such face-centered cubic lattice

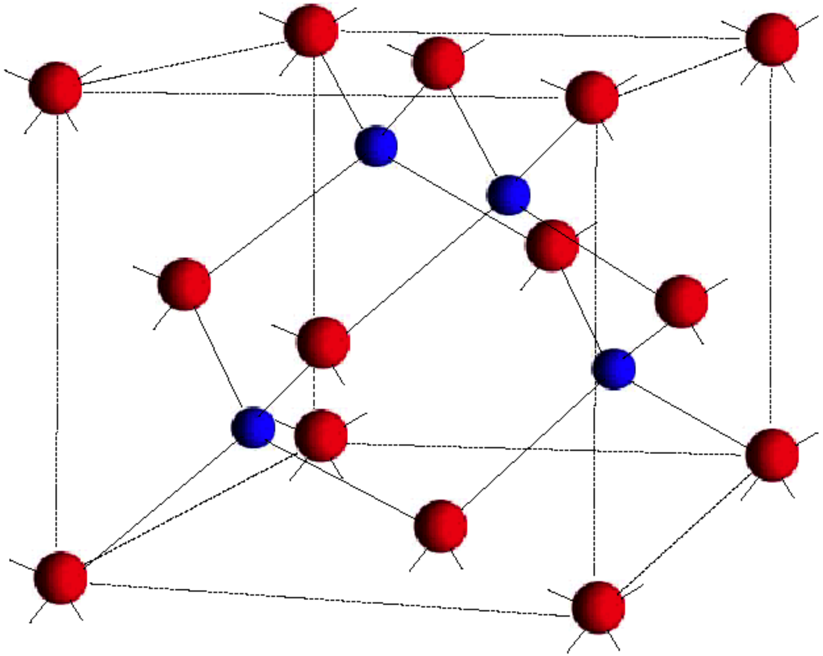
and the group V (or VI) lie on the interlocking face-centered cubic lattice



Diamond and zinc-blende lattices

"Diamond" is the lattice for
- some group IV materials
  - e.g., silicon, germanium
  - some forms of carbon (diamond itself) and tin

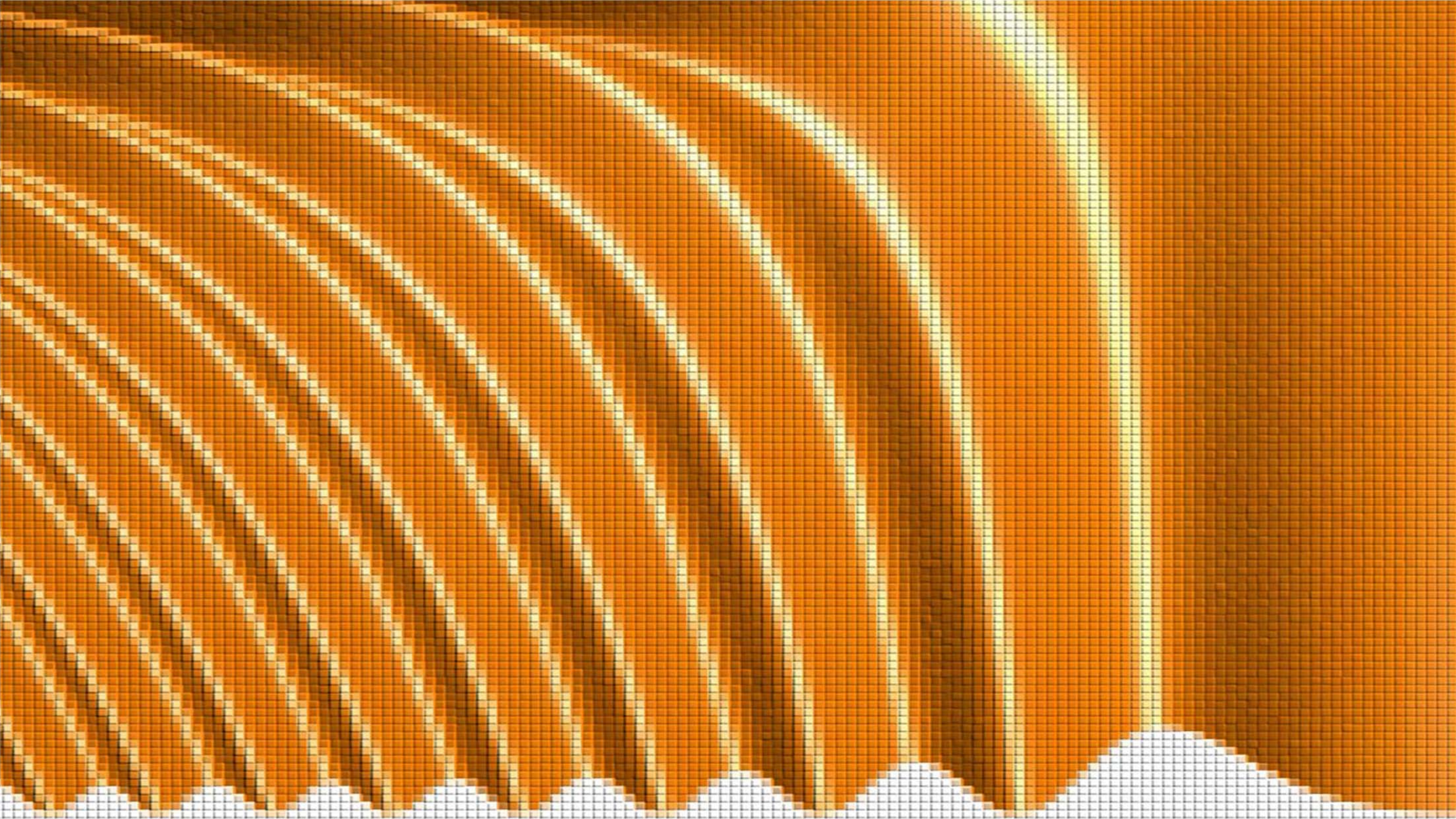
Both interlocking lattices
- have the same kinds of atoms on them



Other important semiconductor lattice structures

Hexagonal
- as in the graphite form of carbon
  - also graphene
    - a single sheet of hexagonal carbon atoms
  - and the basis of carbon nanotubes
    - rolled up sheets of hexagonal carbon atoms

Wurtzite
- a form of hexagonal lattice with two atoms per lattice point

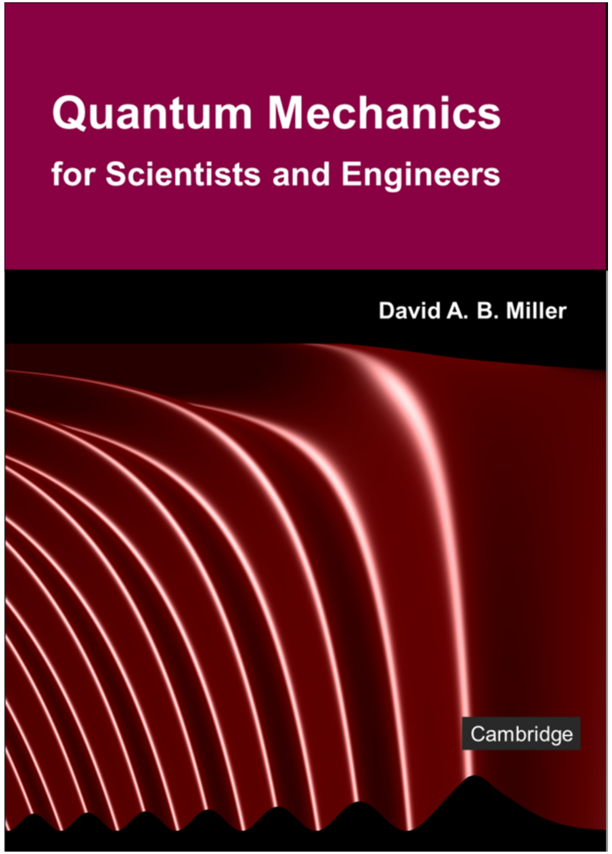


10.2 Bloch theorem

Slides: Video 10.2.1 Periodic boundary conditions

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.2 – 8.3 (through Eq. 8.6)



The Bloch theorem

Periodic boundary conditions

Quantum mechanics for scientists and engineers

David Miller

One electron approximation

In this approximation, we presume that we can write

an effective periodic potential

$$V_P(\mathbf{r}+\mathbf{R}_L)=V_P(\mathbf{r})$$

periodic with the crystal lattice periodicity

and therefore

an effective, approximate Schrödinger equation

for the one electron in which we are interested

$$-\frac{\hbar^2}{2m_e}\nabla^2\psi(\mathbf{r})+V_P(\mathbf{r})\psi(\mathbf{r})=E\psi(\mathbf{r})$$

giving our one-electron approximation

Periodicity of $|\psi|^2$

In one dimension

- the crystal is periodic with “repeat length” a
  - having the same potential at $x + sa$
    - as it has at x
      - Here
        - s is an integer

Similarly, any observable quantity must also have the same periodicity

- because the crystal must look the same in every unit cell

Consequences of periodicity of $|\psi|^2$

For example charge density $\rho \propto |\psi|^2$

must be periodic in the same way

Hence $|\psi(x)|^2 = |\psi(x+a)|^2$

which means

$$\psi(x+a) = C\psi(x)$$

where C is a unit amplitude complex number

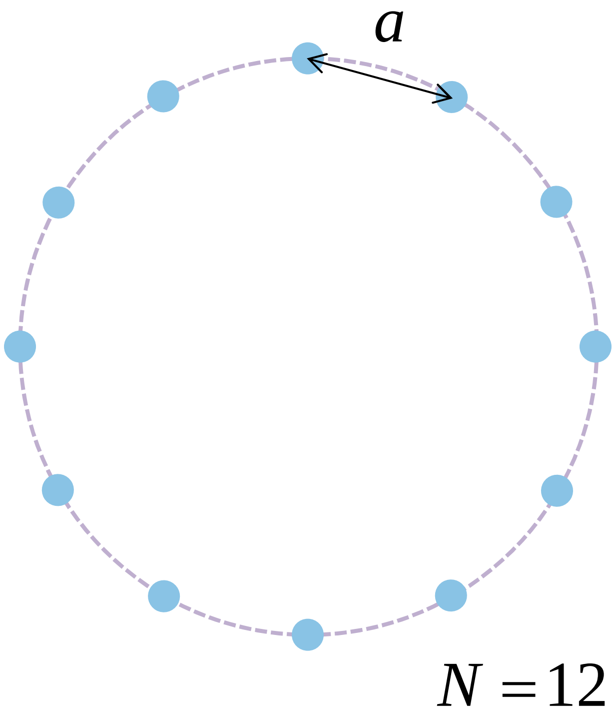
Note that there is no requirement that the wavefunction itself is periodic with the crystal periodicity

since it is not apparently an observable or measurable quantity

Periodic boundary conditions

In one dimension, we could argue as follows

- Suppose we have a long chain of N equally spaced atoms
- and that we join the two ends of the chain together



Periodic boundary conditions

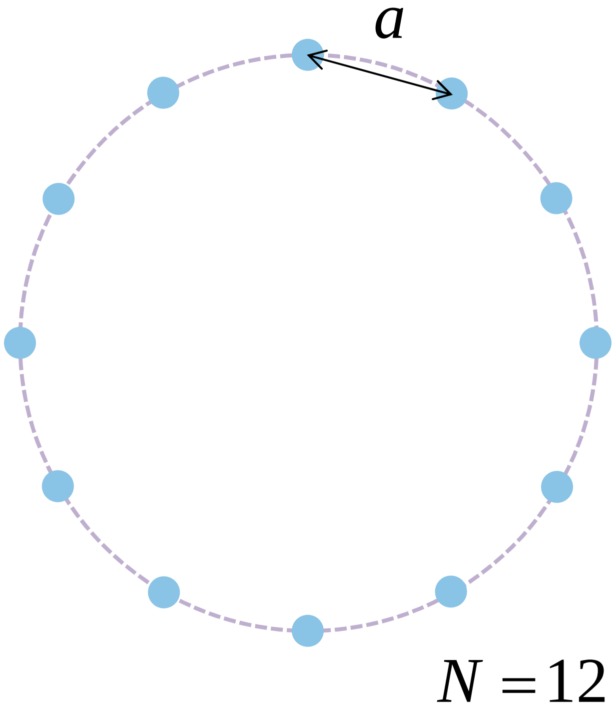
With x as the distance along this loop

then on this loop, the potential can be written

$$V_P(x+ma)=V_P(x)$$

where m is any integer

even possibly an integer much larger than N



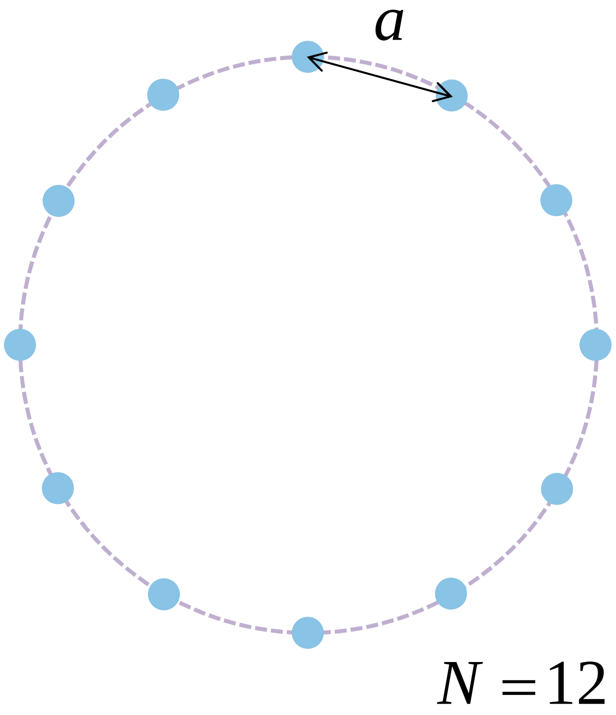
Periodic boundary conditions

This expression $V_P(x+ma)=V_P(x)$

- is just like the one for the infinite crystal

If this chain is very long

- its internal properties will not be substantially different from an infinitely long chain
  - so this is a good model
    - that gives us a finite system
      - while keeping it periodic



Periodic boundary conditions

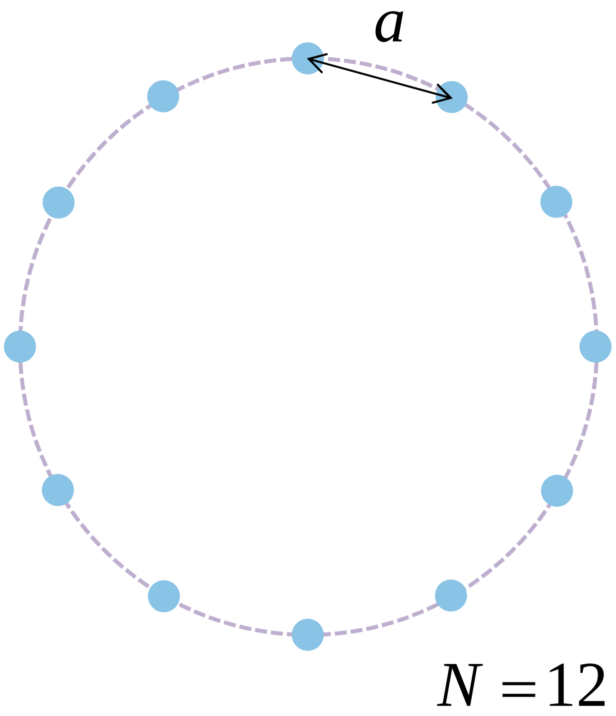
This loop gives a boundary condition

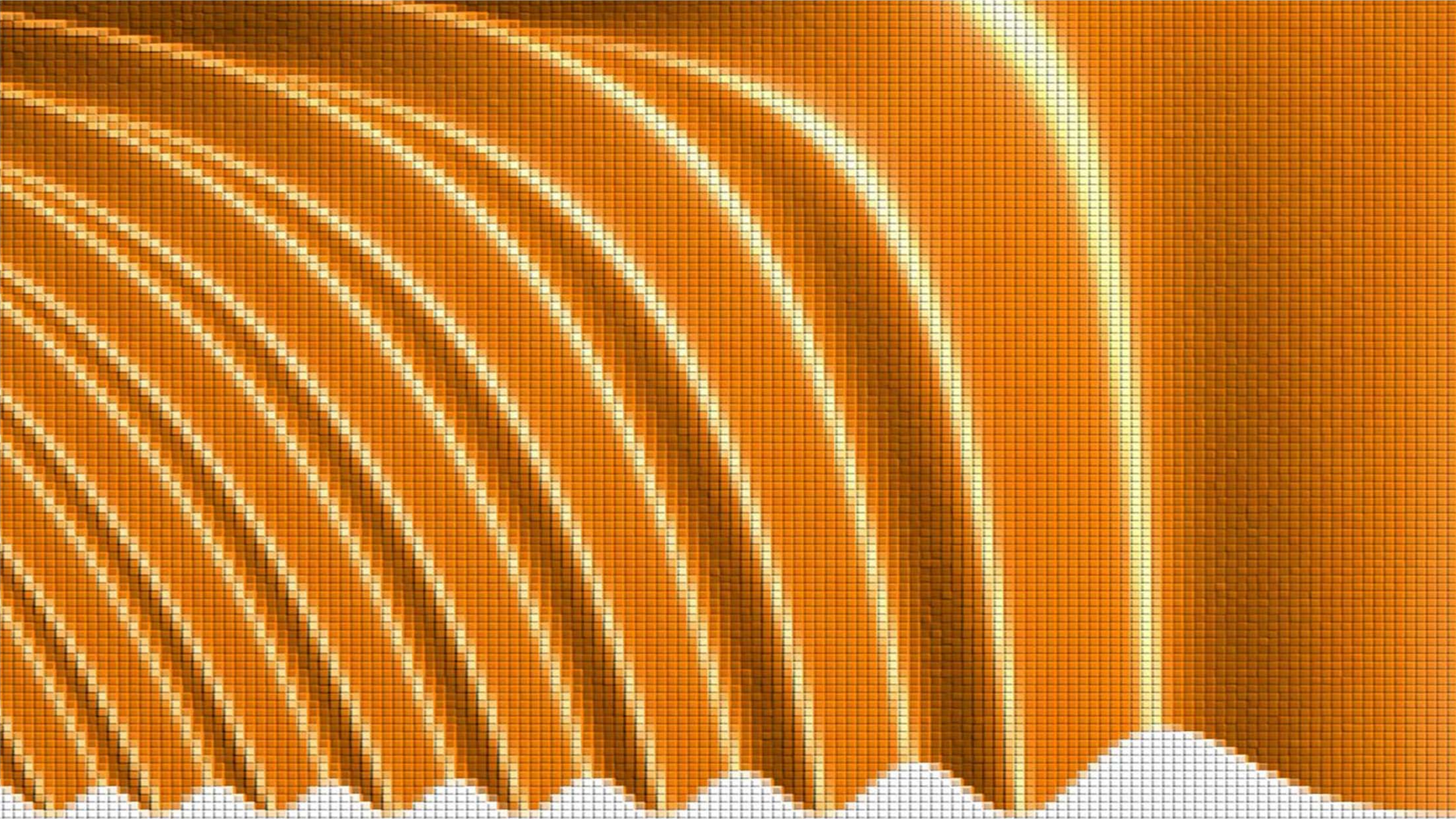
- We do want the wavefunction to be single-valued
  - otherwise how could we differentiate it, evaluate its squared modulus, etc.

So, going round the loop, we must get back to where we started

$$\psi(x)=\psi(x+Na)$$

- a "periodic boundary condition"



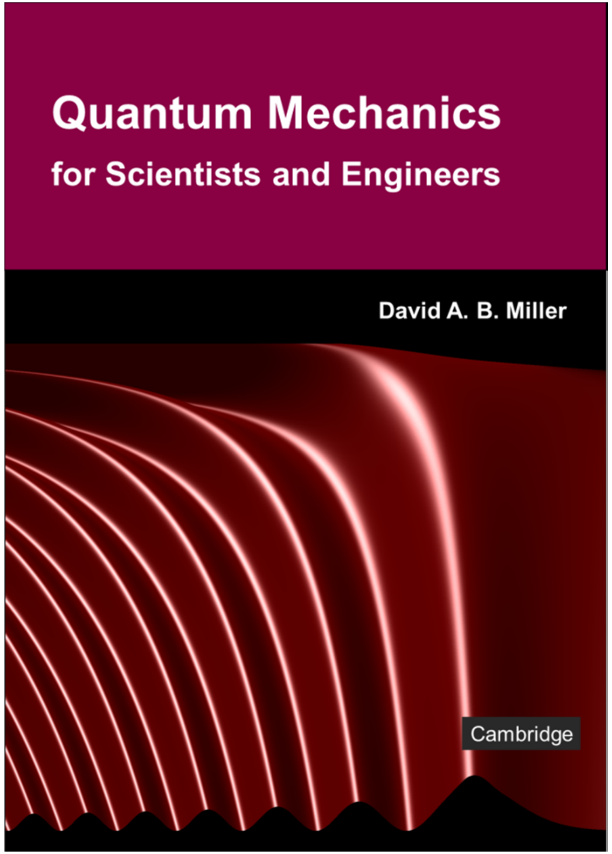


10.2 Bloch theorem

Slides: Video 10.2.3 Bloch theorem derivation

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.3 (from Eq. 8.7)



The Bloch theorem

Bloch theorem derivation

Quantum mechanics for scientists and engineers

David Miller

Bloch theorem derivation

If we take this "single value" requirement $\psi(x) = \psi(x + Na)$

and combine it with the required periodicity of a measureable quantity like probability density

$$|\psi(x)|^2 = |\psi(x+a)|^2$$

which we deduced implied that $\psi(x+a) = C\psi(x)$

where C is a unit complex number

then $\psi(x) = \psi(x + Na) = C^N \psi(x)$

so $C^N = 1$

Hence, C is one of the N "Nth roots of unity", e.g.,

$$C = \exp(2\pi is / N);\ s = 0, 1, 2, \ldots N - 1$$

Bloch theorem derivation

Substituting C from

$$C = \exp\left(2\pi is / N\right);\ s = 0, 1, 2, \ldots N-1$$

in

$$\psi(x) = \psi(x + Na) = C^N \psi(x)$$

gives

$$\psi(x+a) = \exp\left(ika\right)\psi(x)$$

where

$$k = \frac{2\pi s}{Na};\ s = 0, 1, 2, \ldots N-1$$

Bloch theorem derivation

Though the form $C = \exp(2\pi is / N);\ s = 0, 1, 2, \ldots N-1$

for C is mathematically common, it is not unique

We can choose any consecutive set of N values of the integer s

and end up with the same set of possible values for C, just in a different order

Remember, for any integer m

$$\exp(0) = \exp(2\pi) = \exp(2m\pi) = 1$$

so the values for C just keep cycling round as we keep increasing s

Bloch theorem derivation

We can therefore end up with correspondingly different sets of values for k

all of which are physically equivalent

Instead of $k = \frac{2\pi s}{Na}; s = 0, 1, 2, \ldots N-1$

we more conventionally use a symmetrical version

$$k = \frac{2\pi n}{Na} \quad \ldots n = 0, \pm 1, \pm 2, \ldots \pm N/2$$

which strictly has one too many values

We should omit one of the "end values" here

Bloch theorem derivation

Note also that it makes no difference in our expression

$$\psi(x+a) = \exp(ika)\psi(x)$$

if we add $2\pi m / a$ (where m is any integer) to k

The set of allowed values of $\exp(ika)$

remains the same

So we can use $k = \dfrac{2\pi n}{Na}$ $\quad \ldots n = 0, \pm 1, \pm 2, \ldots \pm N/2$

or $k = \dfrac{2\pi n}{Na} + \dfrac{2\pi m}{a}$ $\quad \ldots n = 0, \pm 1, \pm 2, \ldots \pm N/2$

This point will have a specific significance later in "extended zone" schemes

Bloch theorem – one statement

The wavefunction in a (one-dimensional) crystal with N unit cells of length a can be written in the form

$$\psi(x+a) = \exp(ika)\psi(x)$$

subject to the condition

$$k = \frac{2\pi n}{Na} \quad \ldots n = 0, \pm 1, \pm 2, \ldots \pm N/2$$

Note the allowed k values are evenly spaced by $2\pi / L$

where $L = Na$ is the length of the crystal (loop)

regardless of the detailed form of the periodic potential

Bloch theorem – alternative (equivalent) statement

Multiply $\psi(x+a)=\exp(ika)\psi(x)$

by $\exp(-ik(x+a))$

to obtain $\psi(x+a)\exp(-ik(x+a))=\psi(x)\exp(-ikx)$

Hence if we define a function

$$u(x)=\psi(x)\exp(-ikx)$$

we have $u(x+a)=u(x)$

Hence $u(x)$ is periodic with the lattice periodicity

Equivalently, $u(x)$ is a function that is the same in every unit cell

Rearranging gives $\psi(x)=u(x)\exp(ikx)$

Bloch theorem – equivalent statement

The wavefunction in a (one-dimensional) crystal with N unit cells of length a can be written in the form

$$\psi(x) = u(x)\exp(ikx)$$

where $u(x)$ is the same in every unit cell

subject to the condition

$$k = \frac{2\pi n}{Na} \quad \ldots n = 0, \pm 1, \pm 2, \ldots \pm N/2$$

Bloch theorem –equivalence of statements

Note that the two forms

$\psi(x) = u(x)\exp(ikx)$ and $\psi(x+a) = \exp(ika)\psi(x)$

are entirely equivalent

We derived the "left" from the "right" one

and we can derive the "right" one from the "left" one

From the "left" form, we have

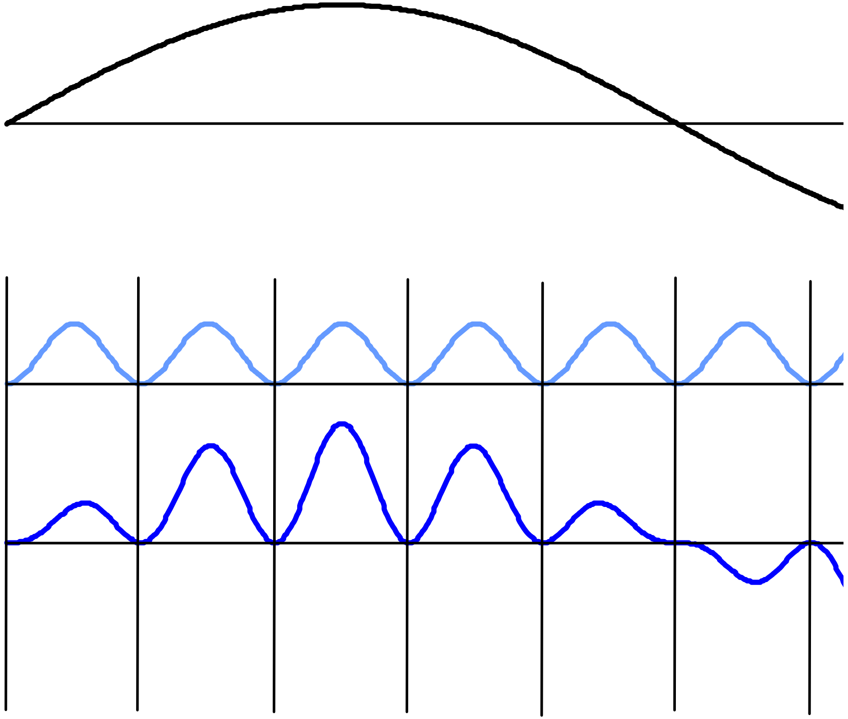
$$\psi(x+a) = u(x+a)\exp\left[ik(x+a)\right] = u(x)\exp\left[ik(x+a)\right]$$

$$= \exp(ika)u(x)\exp(ikx) = \exp(ika)\psi(x)$$

which is the "right" form

Bloch theorem visualization

Visualization of the real part of the wavefunction

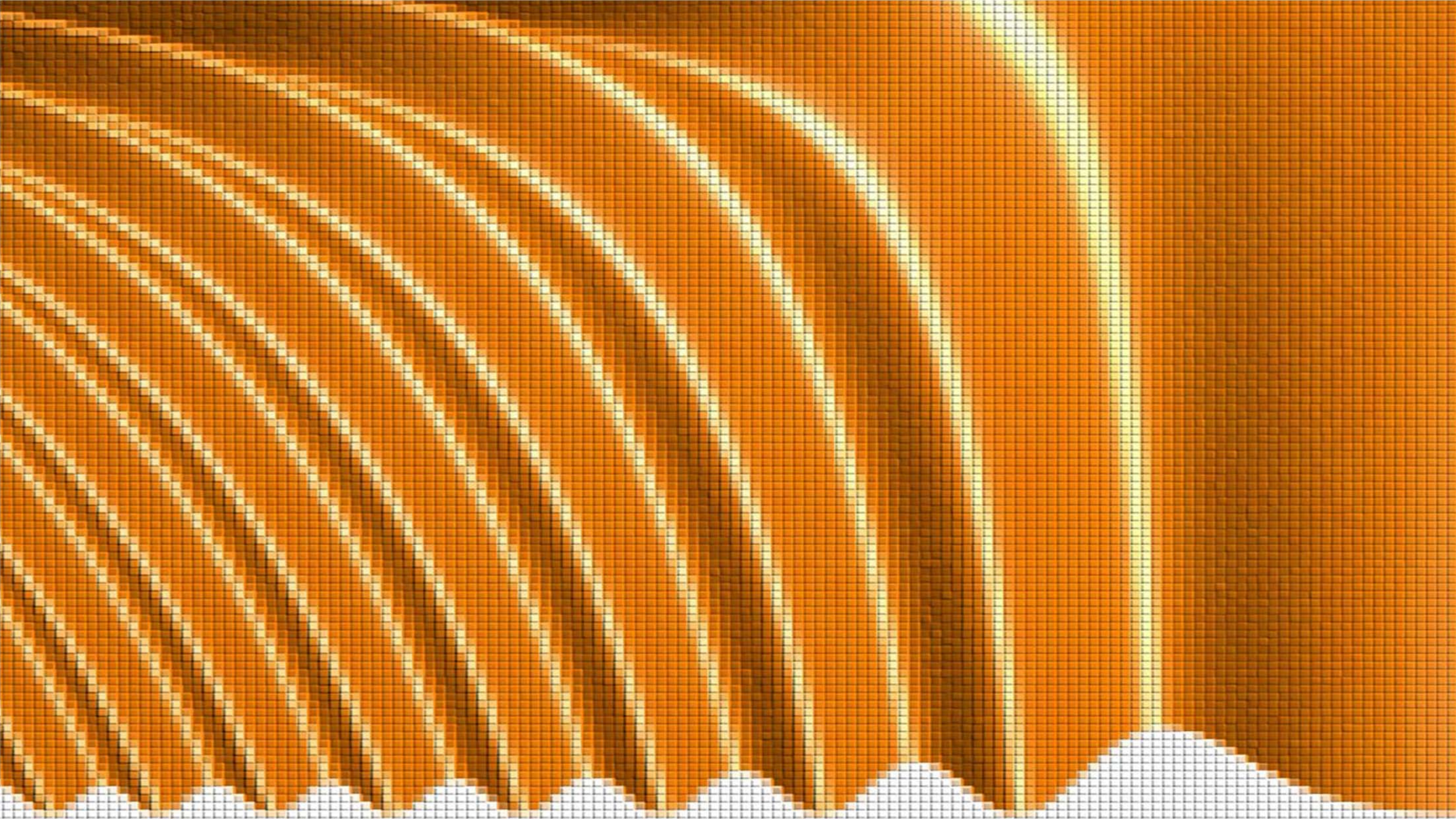


envelope

unit cell function

Bloch function

We can think of the $\exp(ikx)$ as an "envelope" function multiplying the unit cell function $u(x)$

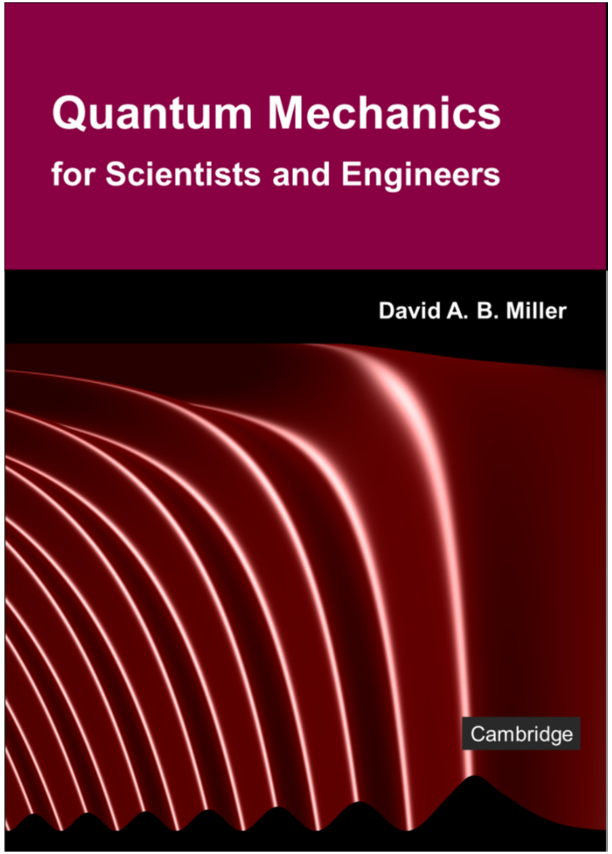


10.2 Bloch theorem

Slides: Video 10.2.5 Density of states in k-space

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.4



The Bloch theorem

Density of states in k-space

Quantum mechanics for scientists and engineers

David Miller

Bloch theorem in three dimensions

To construct the Bloch theorem in three dimensions

we propose a straightforward extension from 1-D

We have

$$\psi(\mathbf{r}+\mathbf{a}) = \exp(i\mathbf{k}\cdot\mathbf{a})\psi(\mathbf{r})$$

where $\mathbf{a}$ is any crystal lattice vector

or equivalently

$$\psi(\mathbf{r}) = u(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{r})$$

and $u(\mathbf{r})$ is the same in every unit cell, i.e.,

$$u(\mathbf{r}+\mathbf{a}) = u(\mathbf{r})$$

Bloch theorem in three dimensions

With the three crystal basis vector directions 1, 2, and 3

with lattice constants (repeat distances) a_1, a_2, and a_3

and numbers of atoms N_1, N_2, and N_3

$$k_1 = \frac{2\pi n_1}{N_1 a_1} \quad \ldots \quad n_1 = 0, \pm 1, \pm 2, \ldots \pm N_1 / 2$$

and similarly for the other two components of $\mathbf{k}$ in the other two crystal basis vector directions

Note that the number of possible values of $\mathbf{k}$ is the same as the number of unit cells in the crystal

(formally dropping the k values at one end or the other)

Reciprocal lattice

We see that the allowed values of k_1, k_2, and k_3 are each equally spaced, with separations

$$\delta k_1 = \frac{2\pi}{N_1 a_1} = \frac{2\pi}{L_1}, \quad \delta k_2 = \frac{2\pi}{N_2 a_2} = \frac{2\pi}{L_2}, \text{ and } \delta k_3 = \frac{2\pi}{N_3 a_3} = \frac{2\pi}{L_3}$$

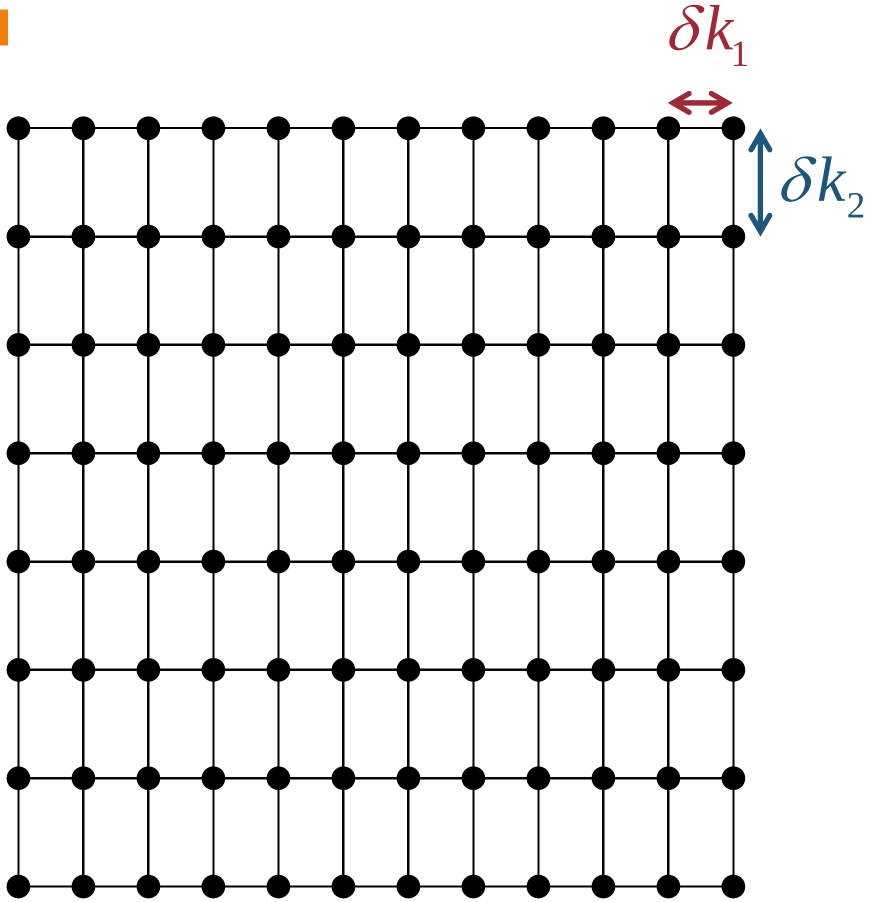
respectively along the three axes

where the lengths of the crystal along the three axes are respectively

$$L_1 = N_1 a_1, \; L_2 = N_2 a_2, \; L_3 = N_3 a_3$$

Reciprocal lattice

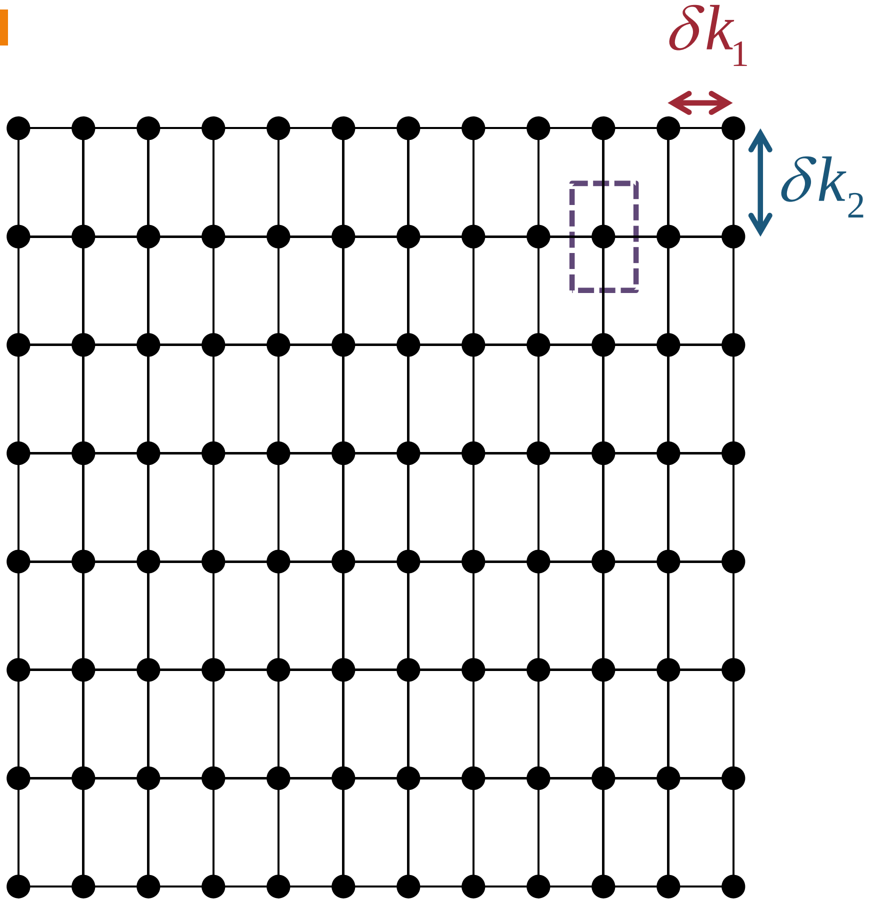
- We could draw a three-dimensional diagram
  - with axes k_1, k_2, and k_3
    - and mark the allowed values of $\mathbf{k}$
- This set of dots themselves constitutes a mathematical lattice
  - This kind of lattice is one kind of “reciprocal lattice”



Density of states in k-space

We imagine each point has a volume surrounding it

with these volumes touching one another to completely fill all the space

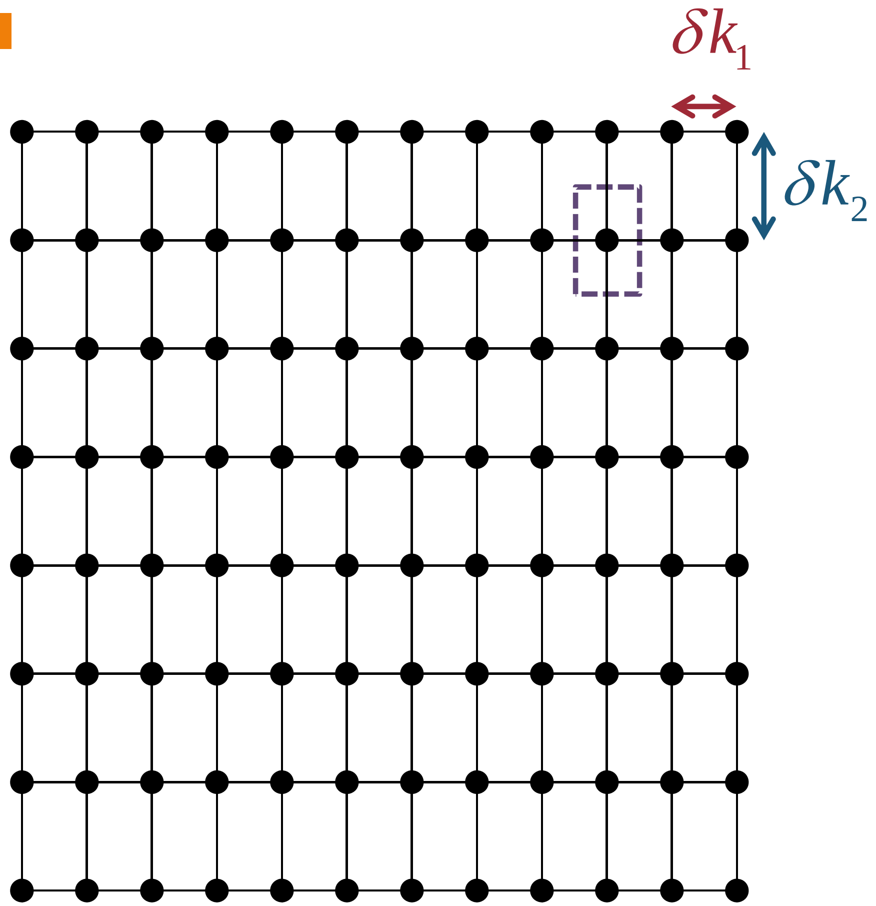


Density of states in k-space

We imagine each point has a volume surrounding it

with these volumes touching one another to completely fill all the space

For our cubic lattices, we can define



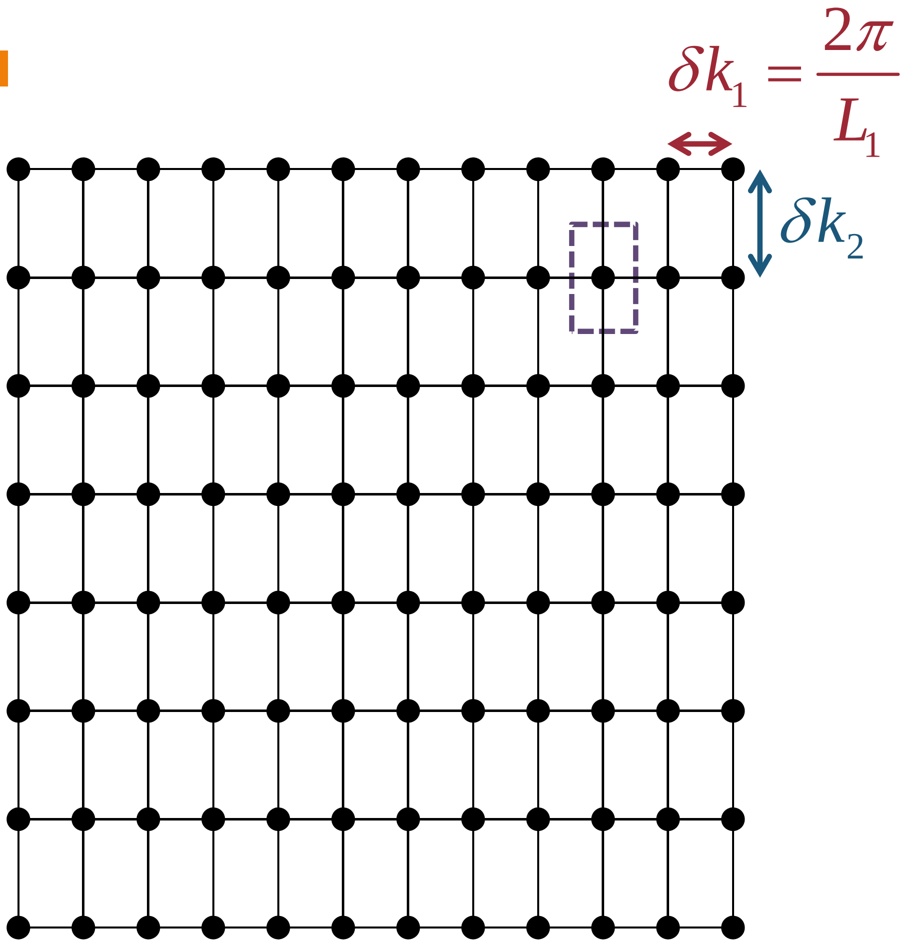
Density of states in *k*-space

We imagine each point has a volume surrounding it

with these volumes touching one another to completely fill all the space

For our cubic lattices, we can define

$$\delta k_1 = \frac{2\pi}{L_1}$$



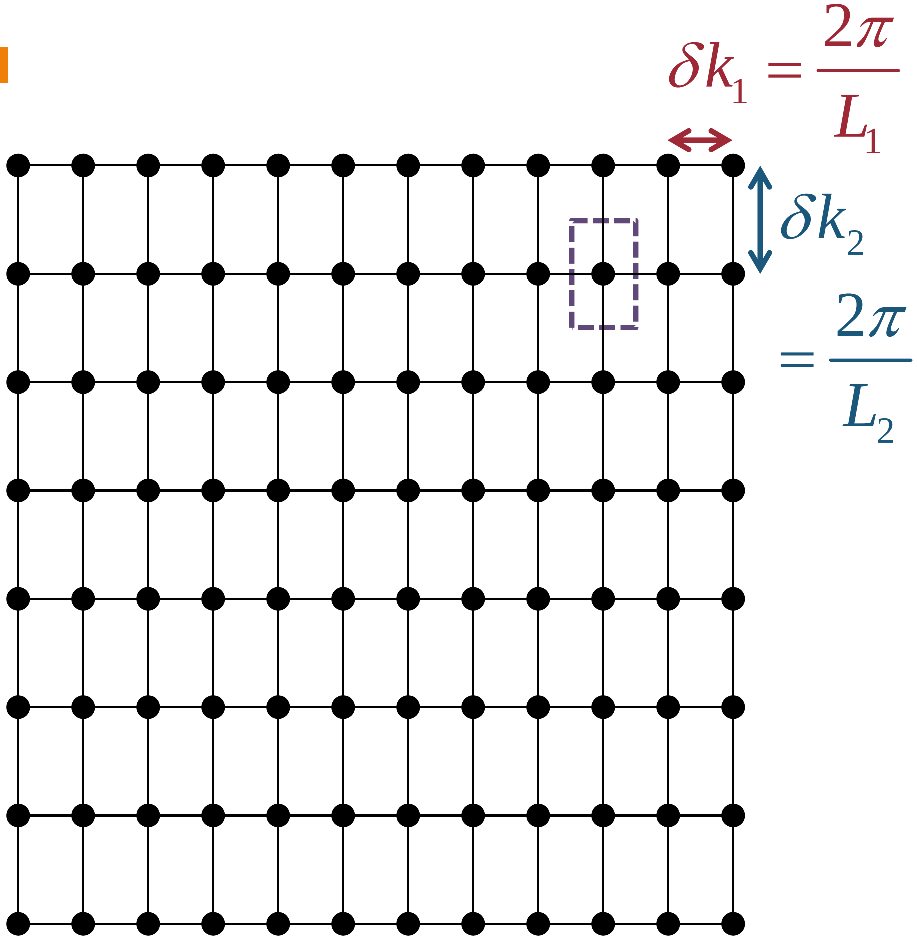
Density of states in *k*-space

We imagine each point has a volume surrounding it

with these volumes touching one another to completely fill all the space

For our cubic lattices, we can define

$$\delta k_1 = \frac{2\pi}{L_1} \quad \delta k_2 = \frac{2\pi}{L_2} \quad \delta k_3 = \frac{2\pi}{L_3}$$



$$\delta k_1 = \frac{2\pi}{L_1}$$

$$\delta k_2 = \frac{2\pi}{L_2}$$

Density of states in *k*-space

For our cuboidal lattices

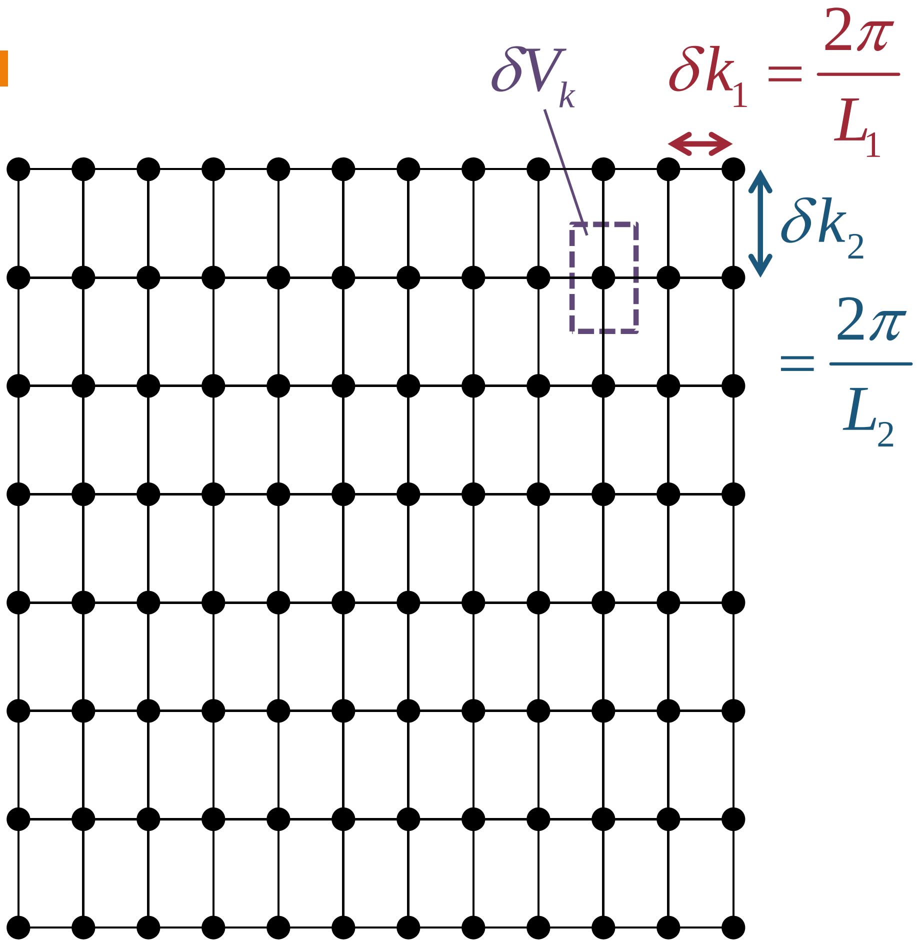
these volumes in *k*-space will be of size $\delta V_k = \delta k_1 \delta k_2 \delta k_3$

i.e., $\delta V_k = \frac{2\pi}{L_1}\frac{2\pi}{L_2}\frac{2\pi}{L_3}$

Since the crystal is $V = L_1 L_2 L_3$

the *k*-space "volume" round each point is

$$\delta V_k = \frac{(2\pi)^3}{V}$$



Density of states in *k*-space

With this specific *k*-space

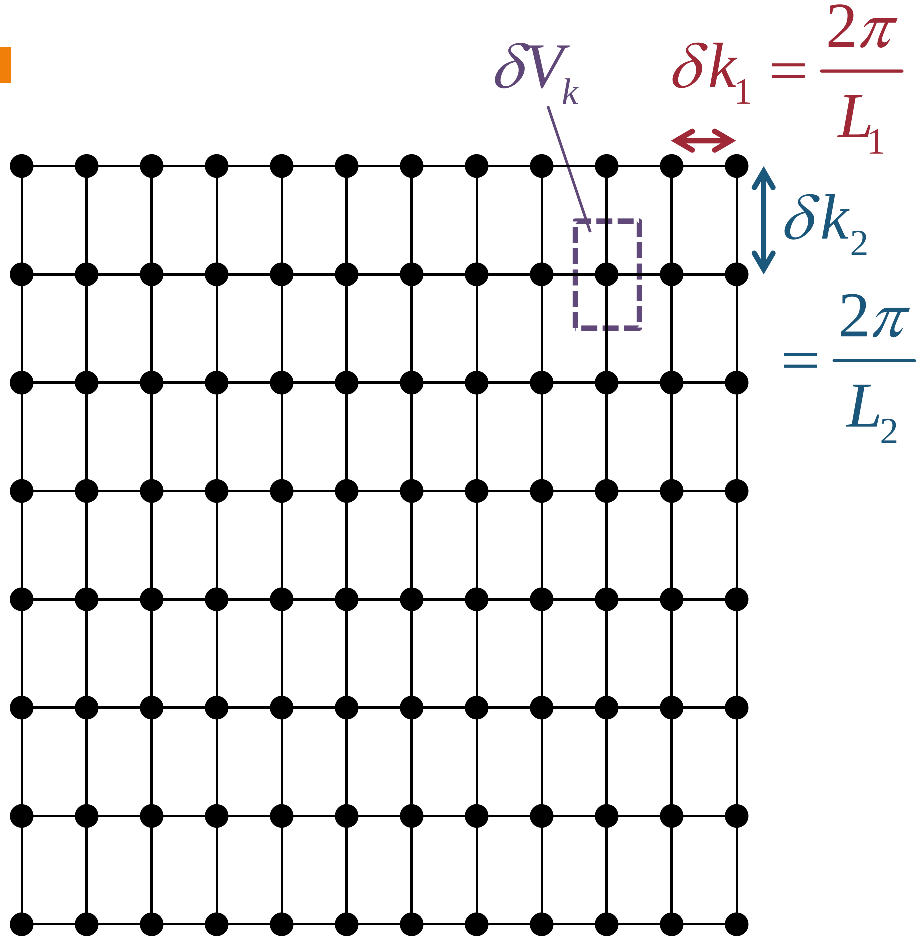
"volume" $\delta V_k = (2\pi)^3 / V$

round each point in *k*-space

we could define

a "density of states in k-space"

$$\frac{1}{\delta V_k} = \frac{V}{(2\pi)^3}$$



Density of states in *k*-space

This density of states in k-space

$$1/\delta V_k = V/(2\pi)^3$$

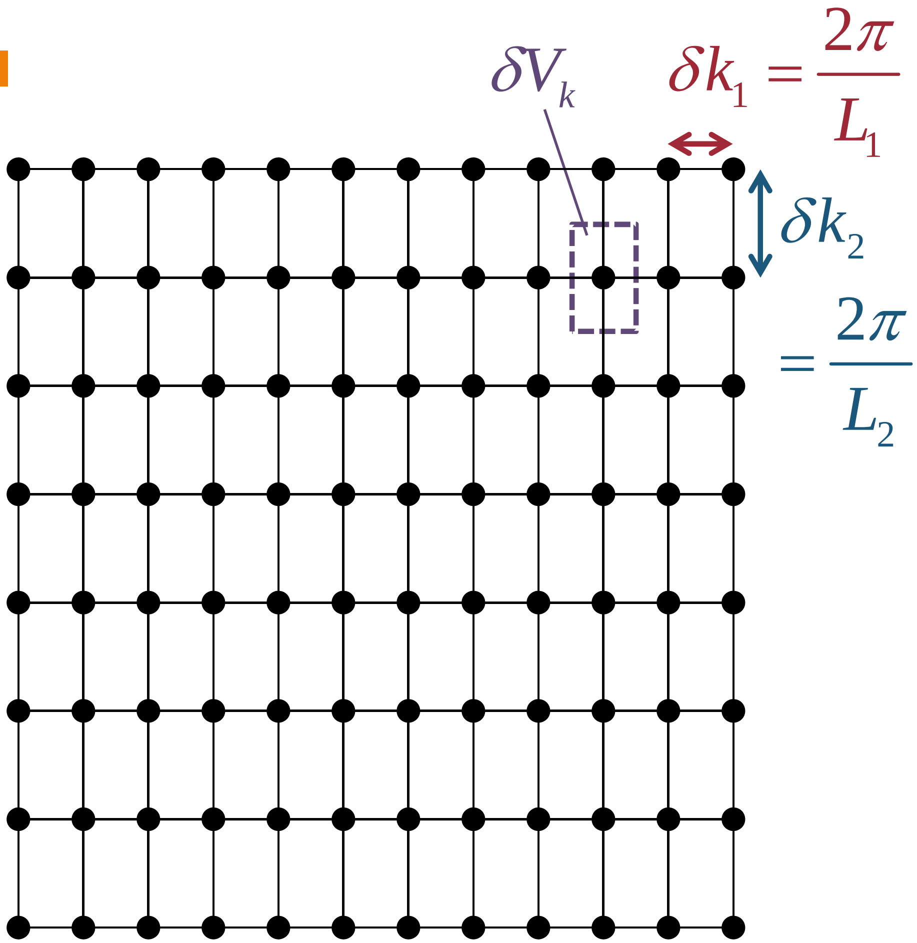
is ∝ crystal volume V

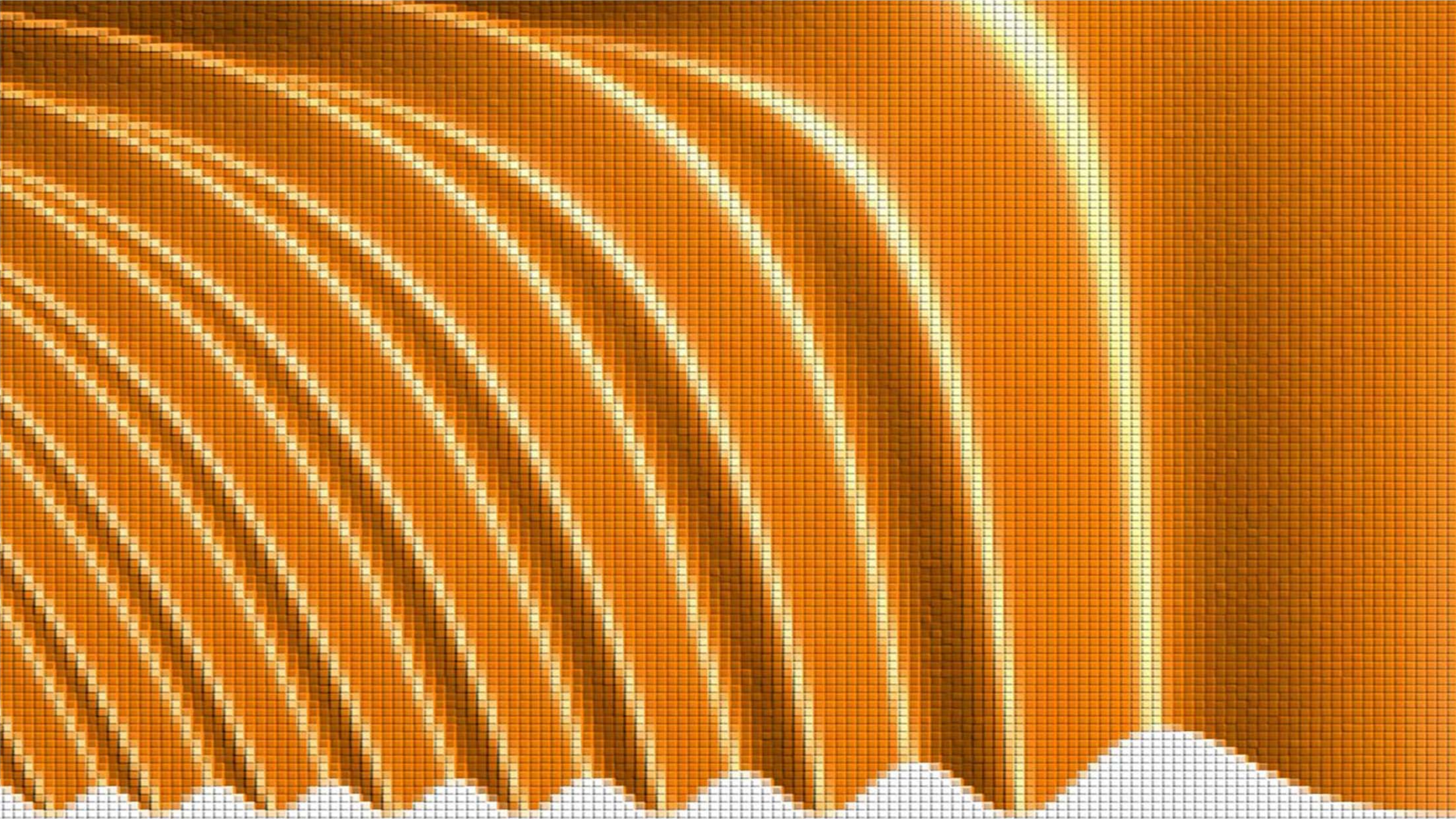
So, more commonly, we define

a "density of states in *k*-space per unit (real space) volume"

$$g(\mathbf{k}) = \frac{1}{(2\pi)^3}$$

for quantum mechanical calculations in crystals





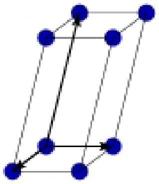
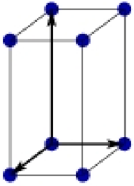
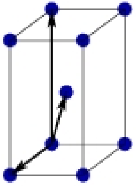
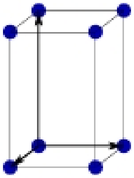
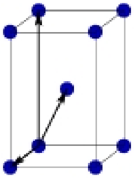
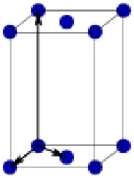
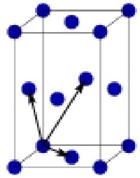
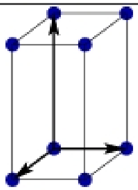
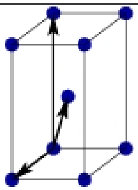
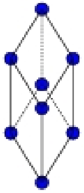
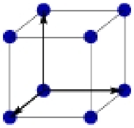
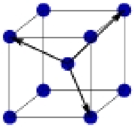
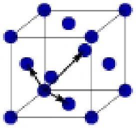
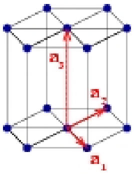
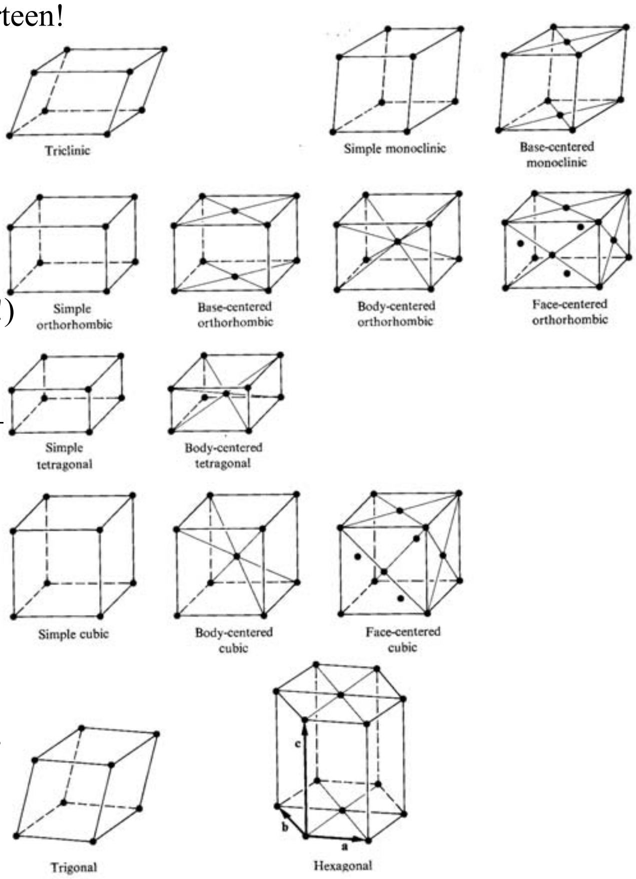
Bravais lattice	Parameters	Simple (P)	Volume centered (I)	Base centered (C)	Face centered (F)
Triclinic	$a_1 \neq a_2 \neq a_3$ $\alpha_{12} \neq \alpha_{23} \neq \alpha_{31}$				
Monoclinic	$a_1 \neq a_2 \neq a_3$ $\alpha_{23} = \alpha_{31} = 90°$ $\alpha_{12} \neq 90°$				
Orthorhombic	$a_1 \neq a_2 \neq a_3$ $\alpha_{12} = \alpha_{23} = \alpha_{31} = 90°$				
Tetragonal	$a_1 = a_2 \neq a_3$ $\alpha_{12} = \alpha_{23} = \alpha_{31} = 90°$				
Trigonal	$a_1 = a_2 = a_3$ $\alpha_{12} = \alpha_{23} = \alpha_{31} < 120°$				
Cubic	$a_1 = a_2 = a_3$ $\alpha_{12} = \alpha_{23} = \alpha_{31} = 90°$				
Hexagonal	$a_1 = a_2 \neq a_3$ $\alpha_{12} = 120°$ $\alpha_{23} = \alpha_{31} = 90°$				

Table 1.1: Bravais lattices in three-dimensions.

The Bravais Lattices Song
Walter F. Smith 1-22-02

If you have to fill a volume with a structure that's repetitive,
Just keep your wits about you, you don't need to take a sedative!
Don't freeze with indecision, there's no need for you to bust a seam!
Although the options may seem endless, really there are just fourteen!
There's cubic, orthorhombic, monoclinic, and tetragonal,
There's trigonal, triclinic, and then finally hexagonal!
There's only seven families, but kindly set your mind at ease—
'Cause four have sub-varieties, so there's no improprieties!
(Chorus:
'Cause four have sub-varieties, so there's no improprieties.
'Cause four have sub-varieties, so there's no improprieties.
'Cause four have sub-varieties, so there's no impropri-e, prieties!)
These seven crystal systems form the fourteen Bravais lattices.
They've hardly anything to do with artichokes or radishes –
They're great for metals, minerals, conductors of the semi-kind –
The Bravais lattices describe all objects that are crystalline!

The cubic is the most important one in my "experience",
It comes in simple and in face- and body-centered variants.
And next in line's tetragonal, it's not at all diagonal,
Just squished in one dimension, so it's really quite rectagonal!
The orthorhombic system has one less degree of symmetry
Because an extra squish ensures that *a* not equals *b* or *c*.
If angle gamma isn't square, the side lengths give the "sig-o-nal"
For monoclinic if they're different, or, if equal, trigonal!
(Chorus (reprovingly):
Of course for trigonal, recall that alpha, beta, gamma all
Are angles that are equal but don't equal ninety, tut, tut, tut!
Are angles that are equal but don't equal ninety, tut, tut, tut, tut tut!)
If you squish the lattice up in every way that is conceivable,
You'll get the least amount of symmetry that is achievable –
It's called triclinic, then remains the one that really self explains –
Hexagonal gives us no pains, and so we now may rest our brains!



The 14 Bravais lattices grouped into the 7 crystal systems.

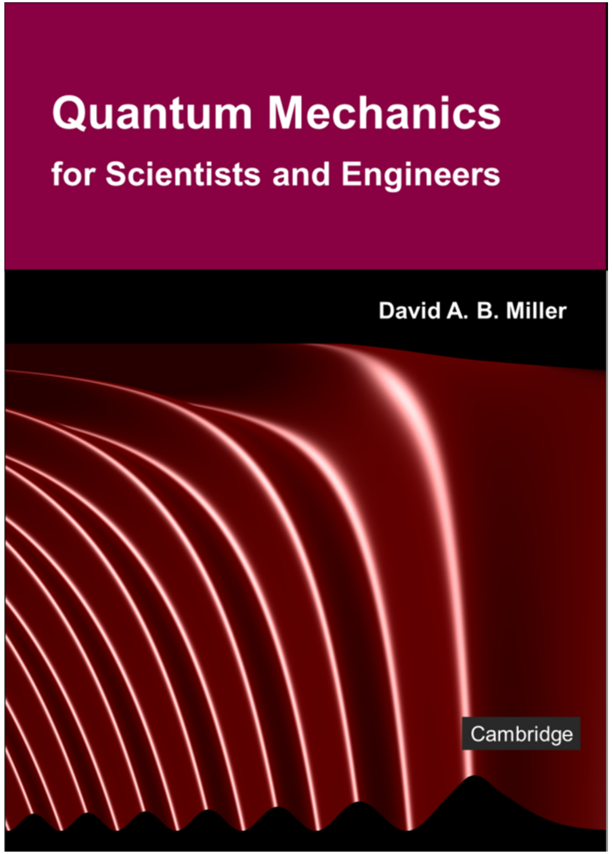
Figure from
Elementary Solid State Physics,
by M. Ali Omar (Addison Wesley, 1993)

11.1 Band structures

Slides: Video 11.1.1 Band structures

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.5 introduction



Band structures

Quantum mechanics for scientists and engineers

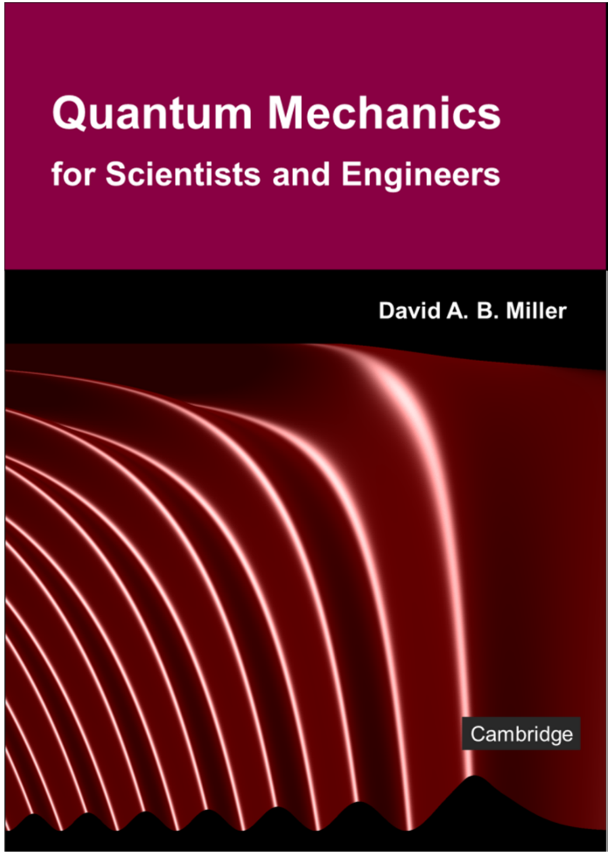
David Miller

11.1 Band structures

Slides: Video 11.1.2 Band structure diagrams

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.5



Band structures

Band structure diagrams

Quantum mechanics for scientists and engineers

David Miller

Band structure

If we knew the periodic potential $V_P(\mathbf{r})$

we could solve the resulting one-electron Schrödinger equation

$$-\frac{\hbar^2}{2m_e}\nabla^2\psi(\mathbf{r})+V_P(\mathbf{r})\psi(\mathbf{r})=E\psi(\mathbf{r})$$

using the Bloch function form $\psi(\mathbf{r})=u(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{r})$

calculating the energies E of all the possible states

These calculations give a “band structure”

There are various band structure calculation methods

Many methods “guess” $V_P(\mathbf{r})$, adjusting it to fit data

Band structure diagrams

To construct a band structure

presuming we know $V_P(\mathbf{r})$

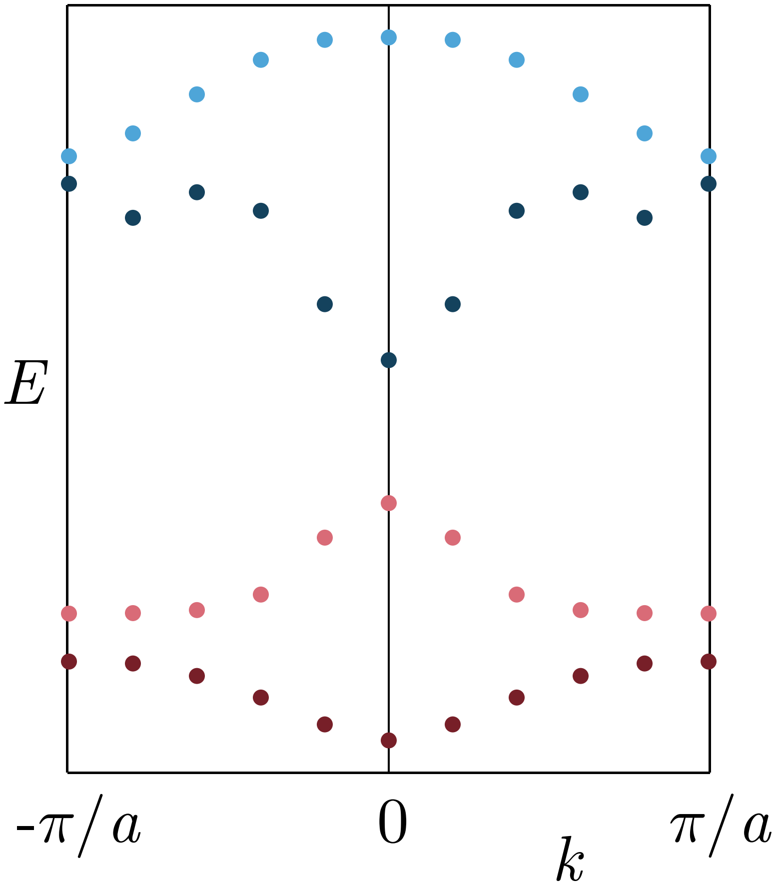
we choose one of the allowed values of k

For simplicity we restrict to one dimension for the moment

Solving the equation

gives energy eigensolutions

We continue with the other allowed values of k



Band structure diagrams

A larger crystal gives more allowed values of k

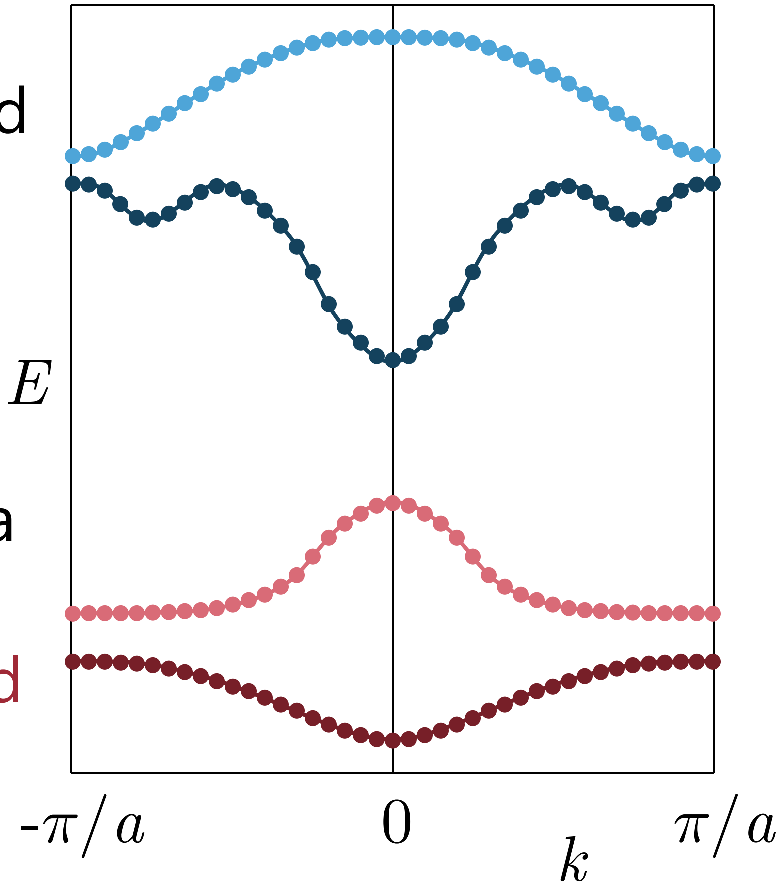
For a large crystal

- the sets of "dots" effectively become like lines

We refer to the group of dots on a line as a "band"

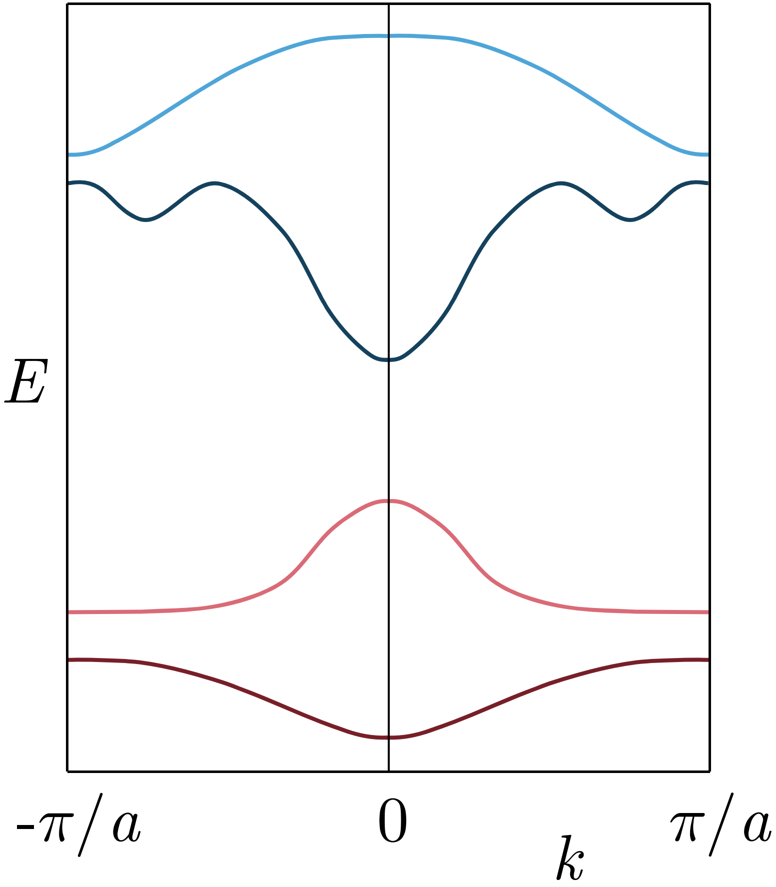
- The number of k-states in a band
  - is the number of unit cells in the crystal

In practice, we just show the lines



Band structure diagrams

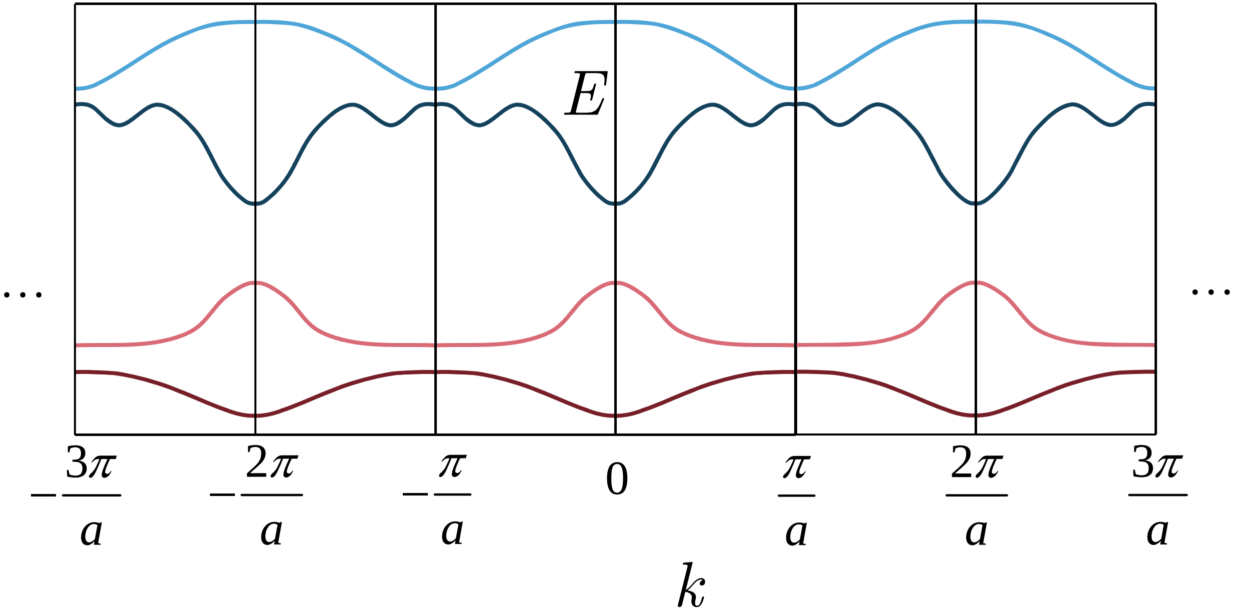
- There are multiple bands in a band structure
  - in fact an infinite number
    - but usually only a few are important for the properties of a material
- In each band, we only have to plot k-values from $-\pi/a$ to π/a
  - This range is known as
    - the (first) Brillouin zone



Extended Brillouin zone scheme

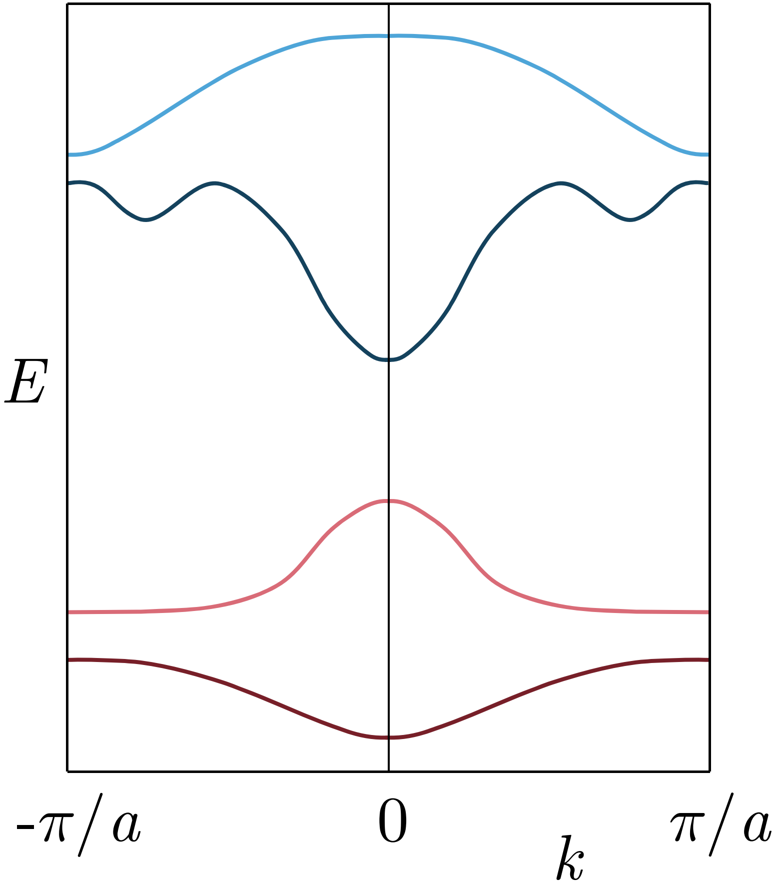
If we continue to larger k

- the band structure just repeats
  - in multiple Brillouin zones
  - an "extended zone scheme"
- so we only need to plot one Brillouin zone



Band structure diagrams

- Each band loosely corresponds to a different atomic state in the constituent atoms
  - or at least orthogonal combinations of atomic states
- The bands are formed from the atomic states
  - as the atoms are "pushed together" to make the crystal



Kramers degeneracy

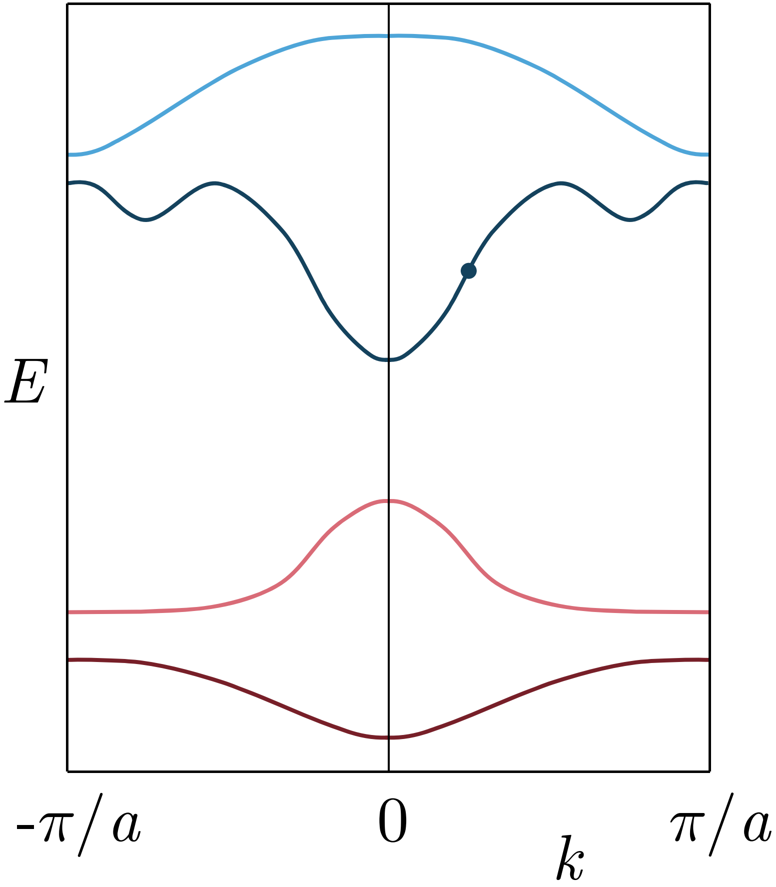
The band structure is drawn to be symmetric about $k = 0$

This common symmetry is easily proved

Suppose that the Bloch function $\psi(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{r})$

satisfies the Schrödinger equation for a specific $\mathbf{k}$

Note the unit cell function $u_{\mathbf{k}}(\mathbf{r})$ may be different for different $\mathbf{k}$



Kramers degeneracy

Hence we have

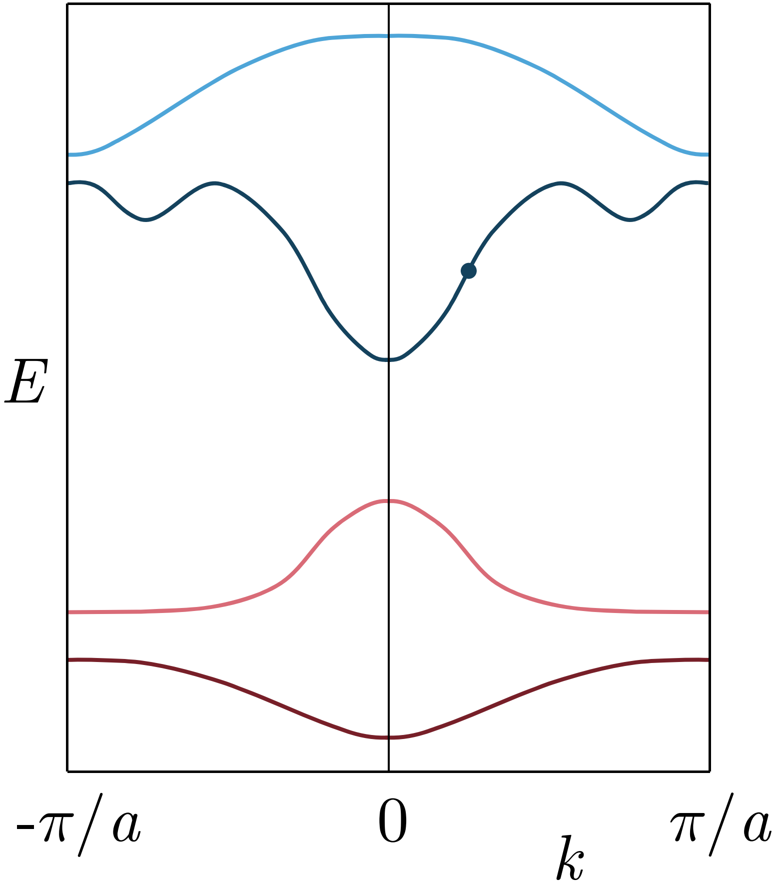
$$H\psi(\mathbf{k},\mathbf{r}) = E_{\mathbf{k}}\psi(\mathbf{k},\mathbf{r})$$

where $E_{\mathbf{k}}$ is the eigenenergy associated with this specific $\mathbf{k}$

in this specific band

and

$$H = -\left(\hbar^2 / 2m_e\right)\nabla^2 + V_P(\mathbf{r})$$



Kramers degeneracy

Taking the complex conjugate of both sides of

$$H\psi(\mathbf{k},\mathbf{r}) = E_{\mathbf{k}}\psi(\mathbf{k},\mathbf{r})$$

noting that $H = H^*$

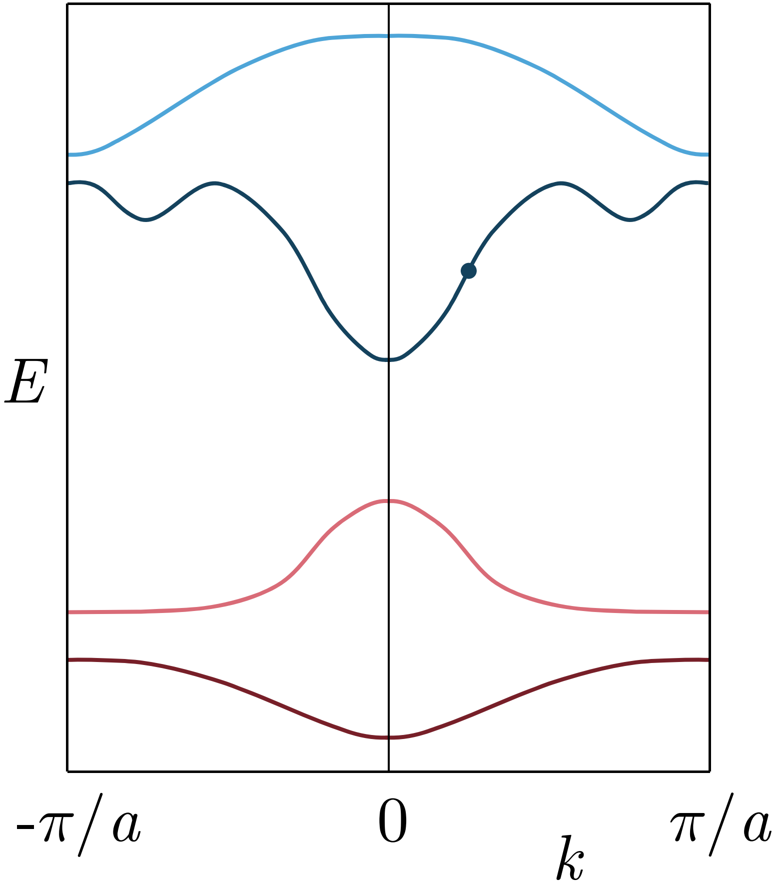
and that $E_{\mathbf{k}}$ is necessarily real

$$H\psi^*(\mathbf{k},\mathbf{r}) = E_{\mathbf{k}}\psi^*(\mathbf{k},\mathbf{r})$$

But $\psi^*(\mathbf{k},\mathbf{r}) = u_{\mathbf{k}}^*(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r})$

which is also a wavefunction in Bloch form

but for -$\mathbf{k}$



Kramers degeneracy

Hence we are saying that

for every solution with wavevector $\mathbf{k}$ and energy $E_{\mathbf{k}}$

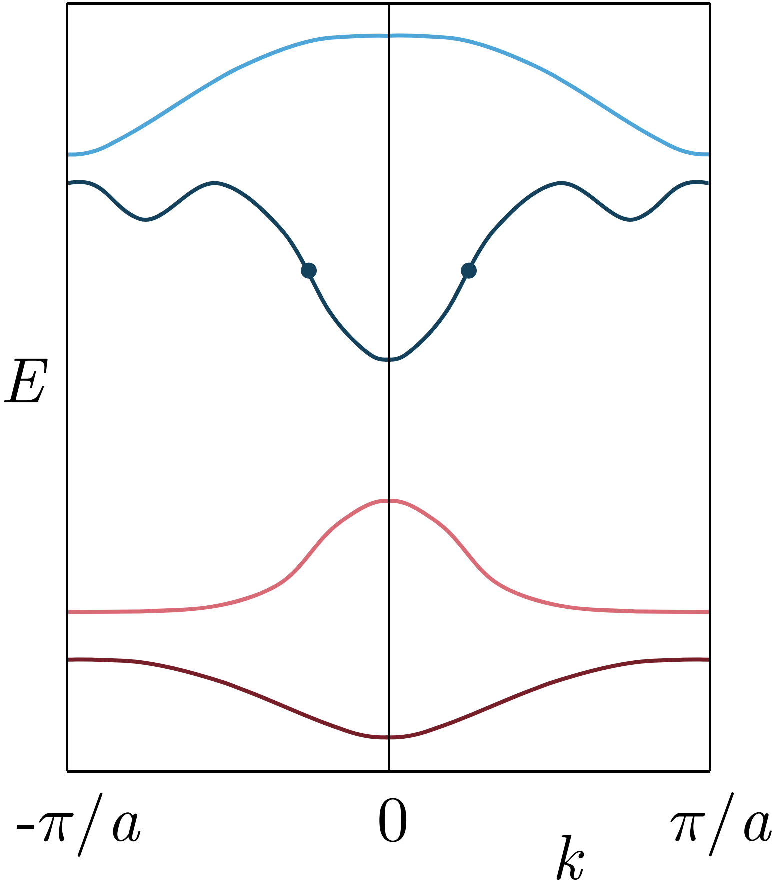
there is one with wavevector $-\mathbf{k}$ with the same energy

Hence the band structure is symmetric about $k = 0$

We can choose to write

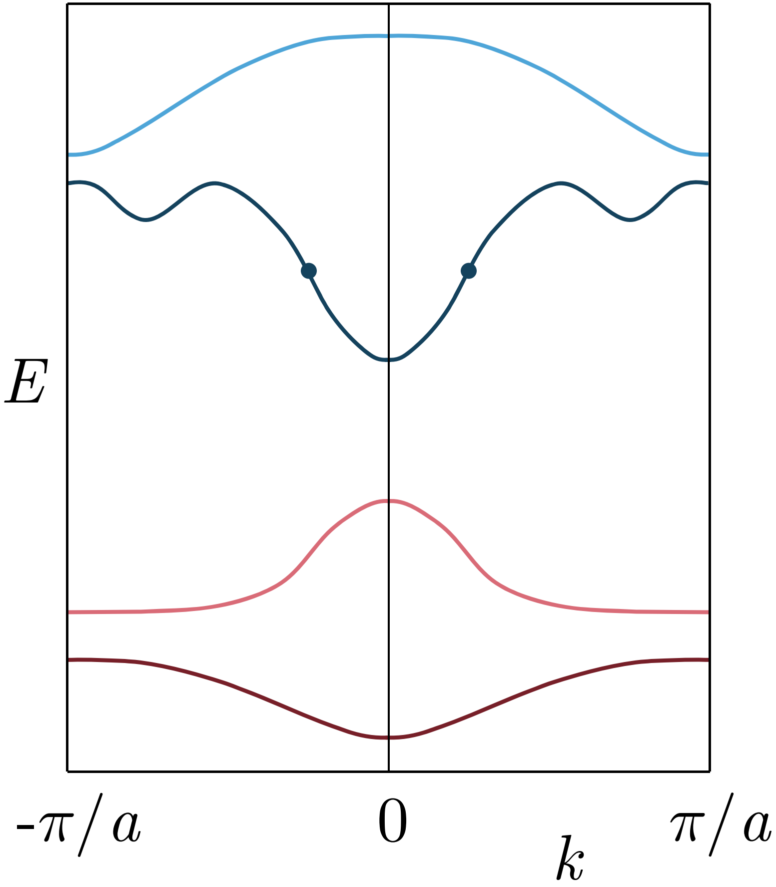
$$\psi^*(\mathbf{k},\mathbf{r}) = u_{\mathbf{k}}^*(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r})$$

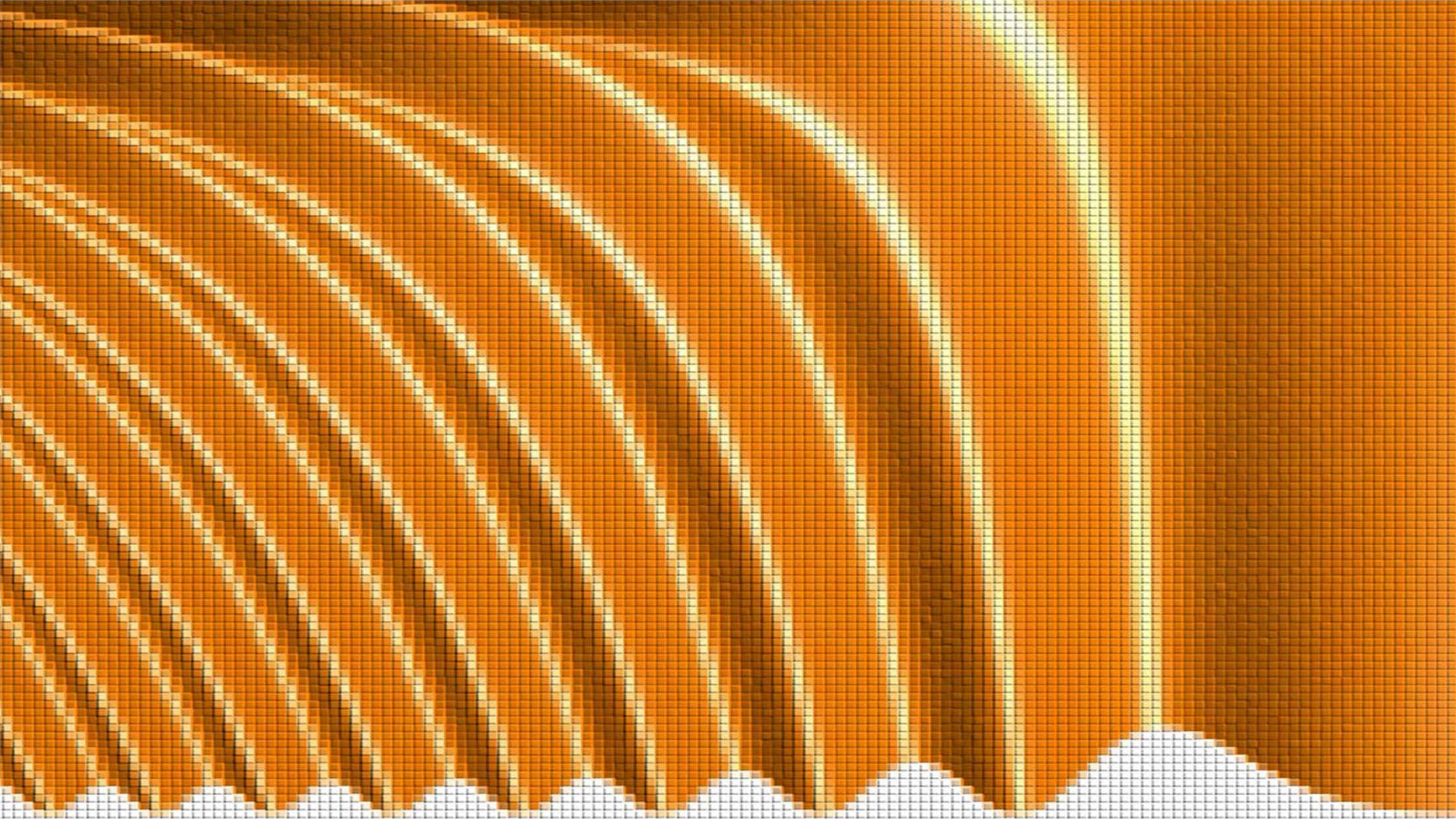
$$\equiv u_{-\mathbf{k}}(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r}) = \psi(-\mathbf{k},\mathbf{r})$$



Kramers degeneracy

- This equivalence of the energies for **k** and -**k** is known as
  - Kramers degeneracy
- Note that, once we include spin
  - these two states will have opposite spin
    - but often the spin makes no difference to the energy
- Hence bands often have minima or maxima at $k = 0$



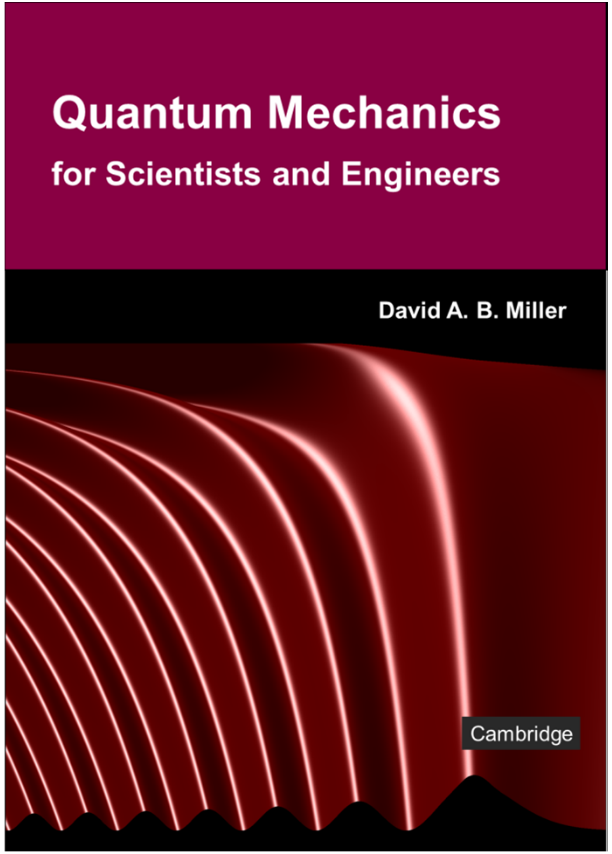


11.1 Band structures

Slides: Video 11.1.4 Semiconductors, insulators and metals

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.5



Band structures

Semiconductors, insulators and metals

Quantum mechanics for scientists and engineers

David Miller

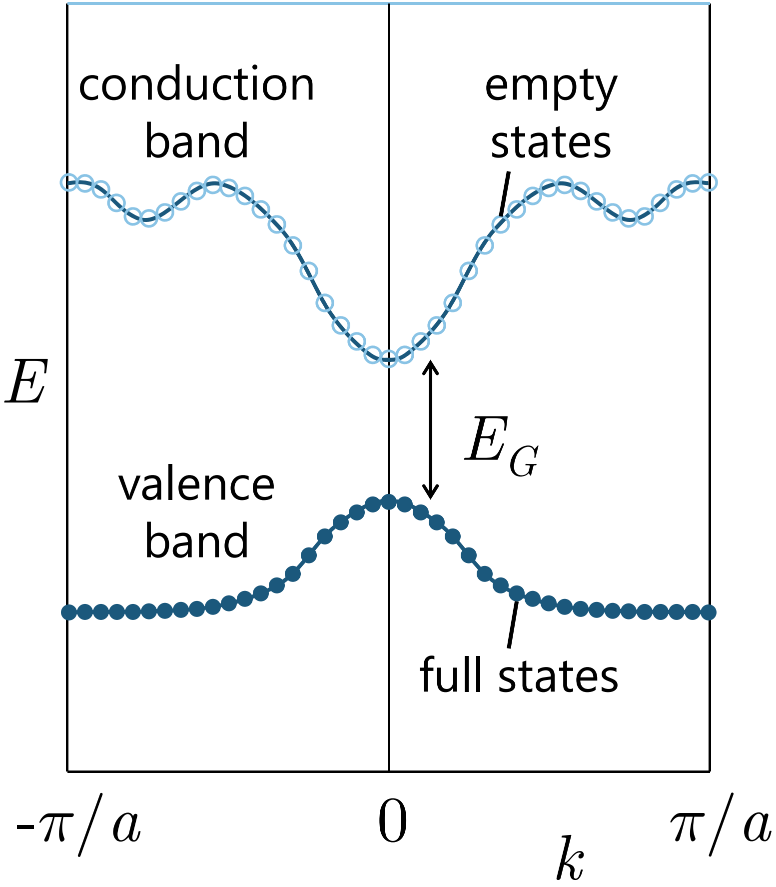
Semiconductors and insulators

Semiconductors and insulators

- have an (almost) completely full band
  - the valence band

separated by a "bandgap" energy E_G

- from an (almost) completely empty band
  - the conduction band



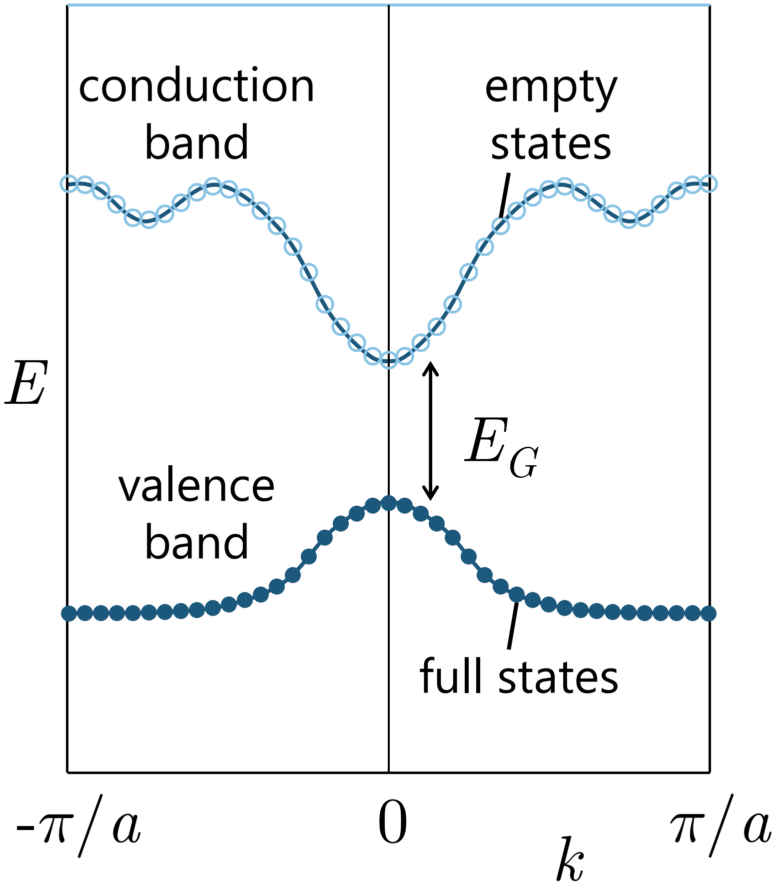
Semiconductors and insulators

Note that

- an empty band does not conduct electricity
  - There are no mobile electrons

Also

- a full band does not conduct electricity
  - The electrons cannot change states within the band
    - because all the states are full



Semiconductors and insulators

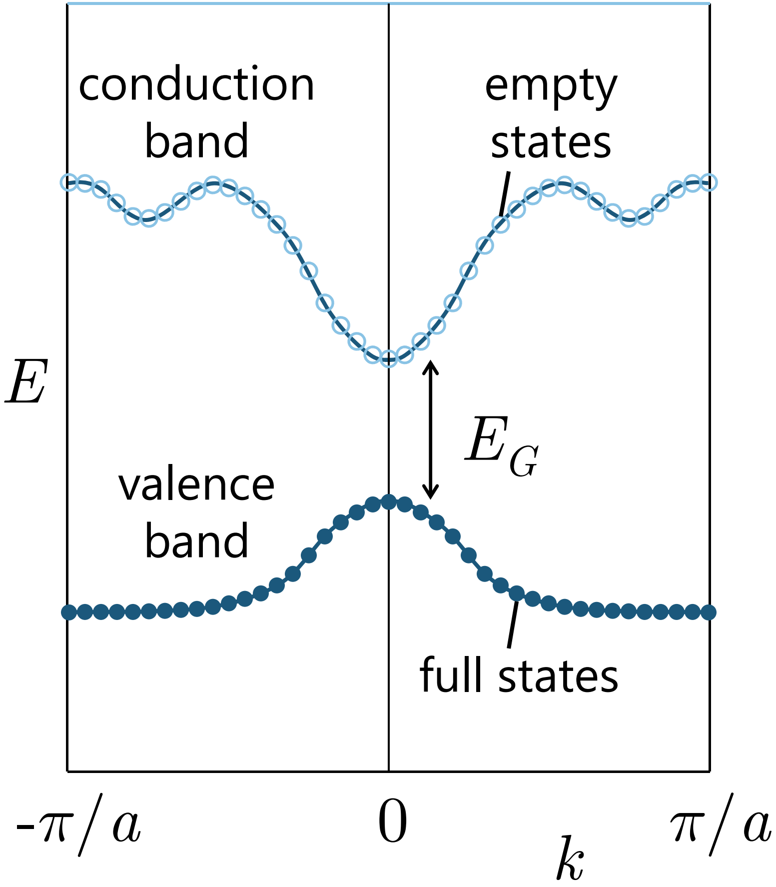
The difference between semiconductors and insulators is primarily that

insulators have such a large bandgap energy

that there is negligible thermal excitation of electrons

from the valence band

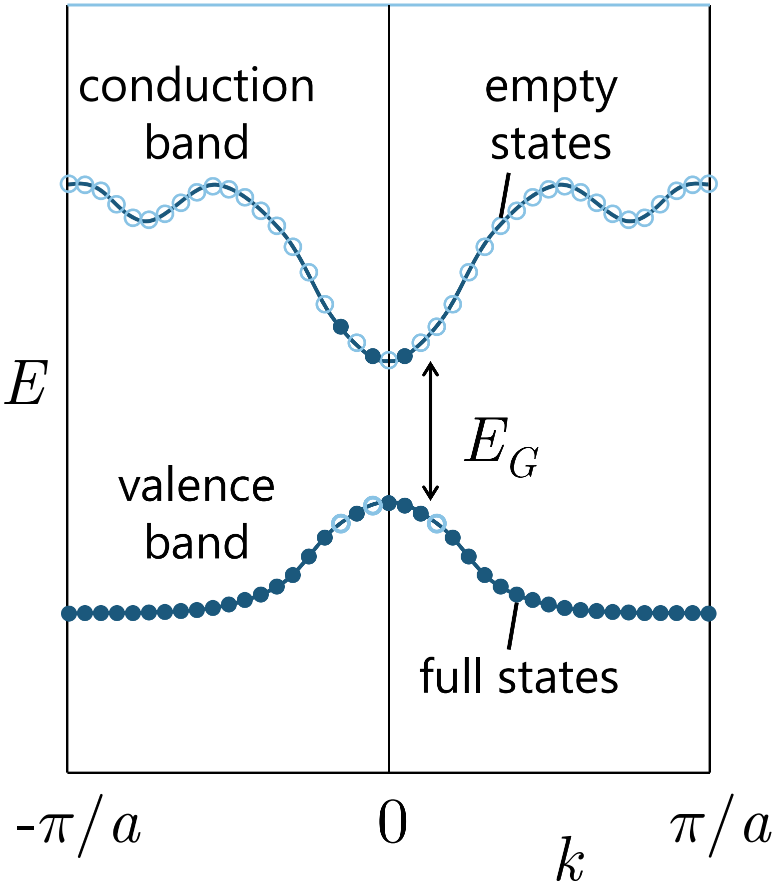
to the conduction band



Semiconductors

At finite temperatures in a semiconductor

- a small number of electrons are excited
  - from the valence band
    - to the conduction band



Semiconductors

These electrons in the conduction band

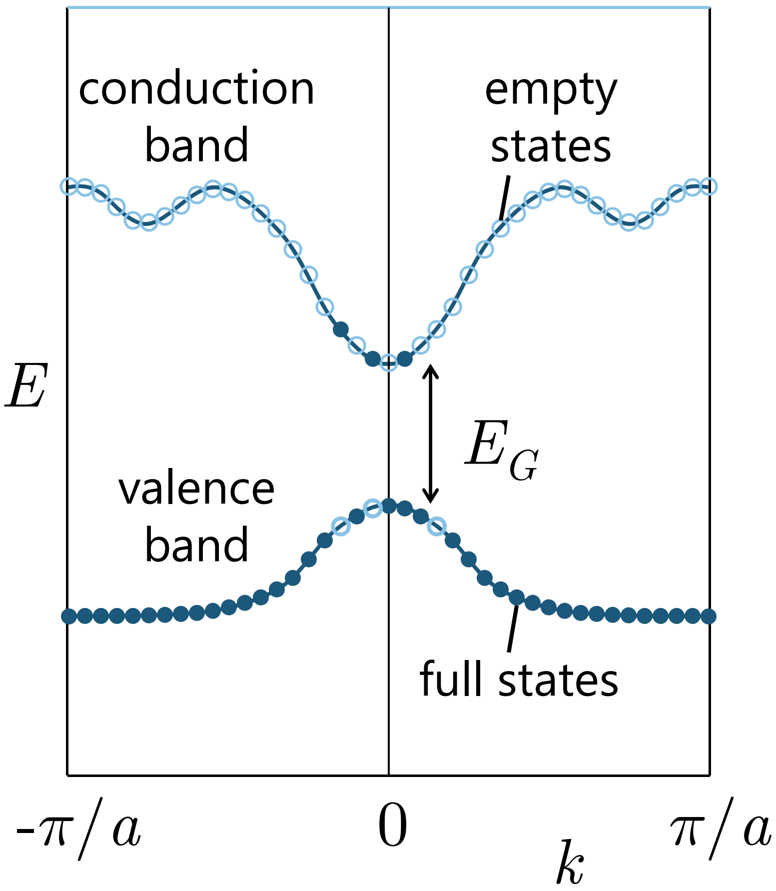
and

absences of electrons or “holes” in the valence band

can conduct electricity within their bands

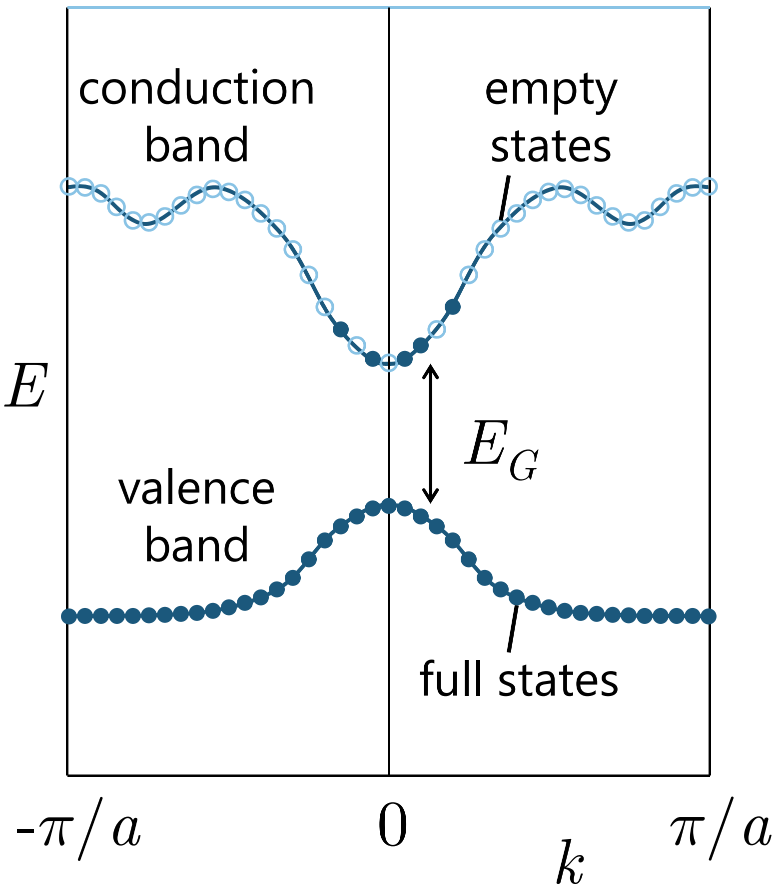
So semiconductor materials conduct electricity weakly

hence the name



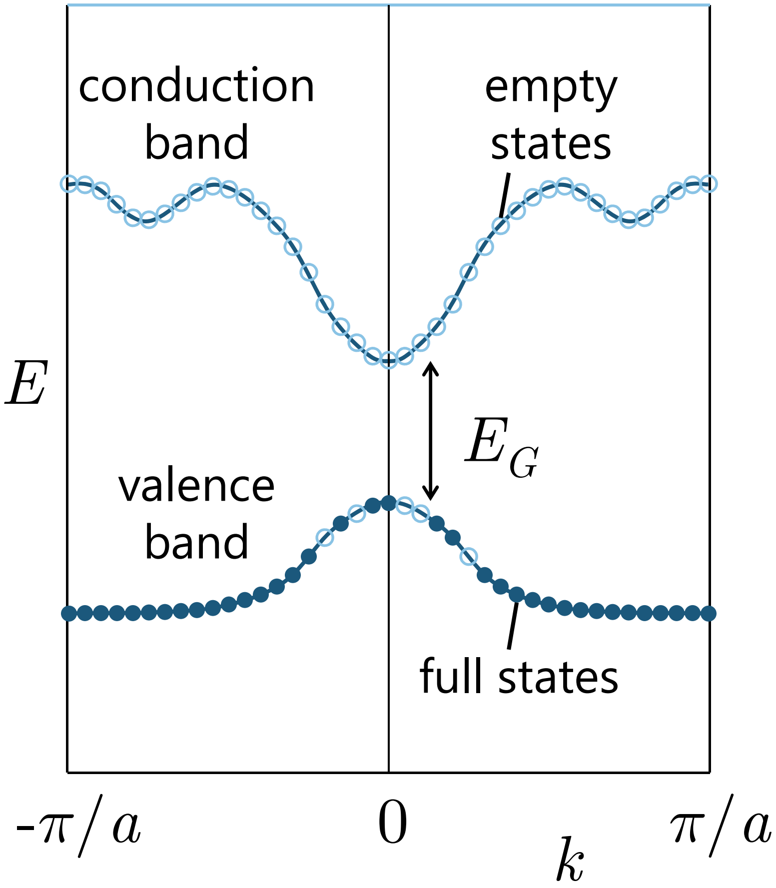
Doping semiconductors

- Substituting a few atoms with more electrons
  - e.g., a Group V element like phosphorus in a Group IV semiconductor like silicon
    - known as n-type doping
  - makes the material conduct more
    - using these additional electrons



Doping semiconductors

- Substituting a few atoms with fewer electrons
  - e.g., a Group III element like boron in a Group IV semiconductor like silicon
    - known as p-type doping
  - makes the material conduct more
    - using these additional “holes”



Direct gap semiconductor

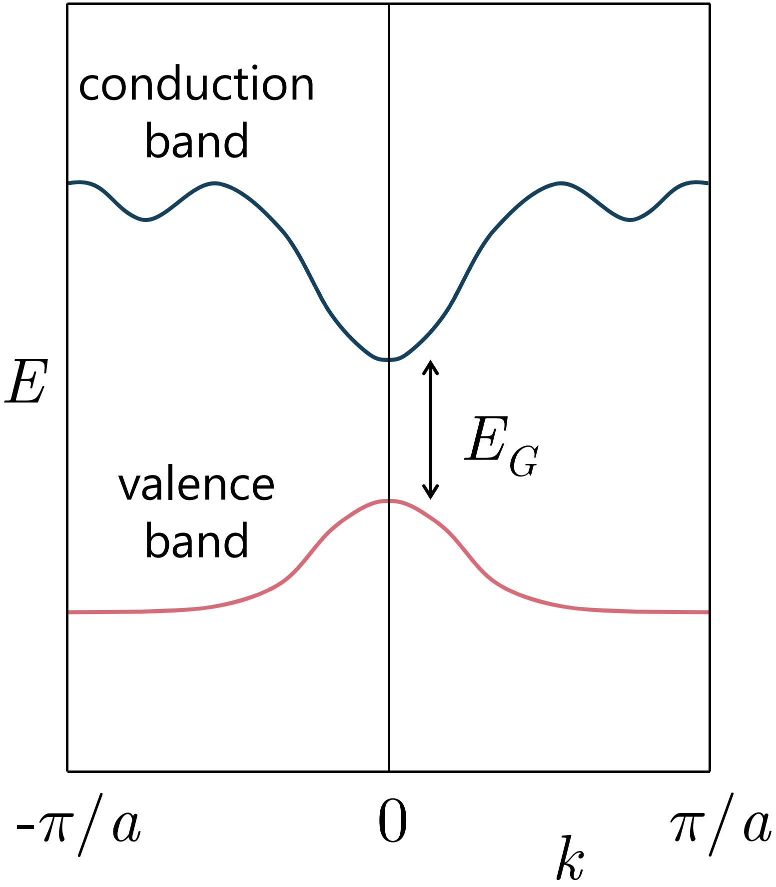
If the lowest minimum in the conduction band

lies directly above

the highest maximum in the valence band

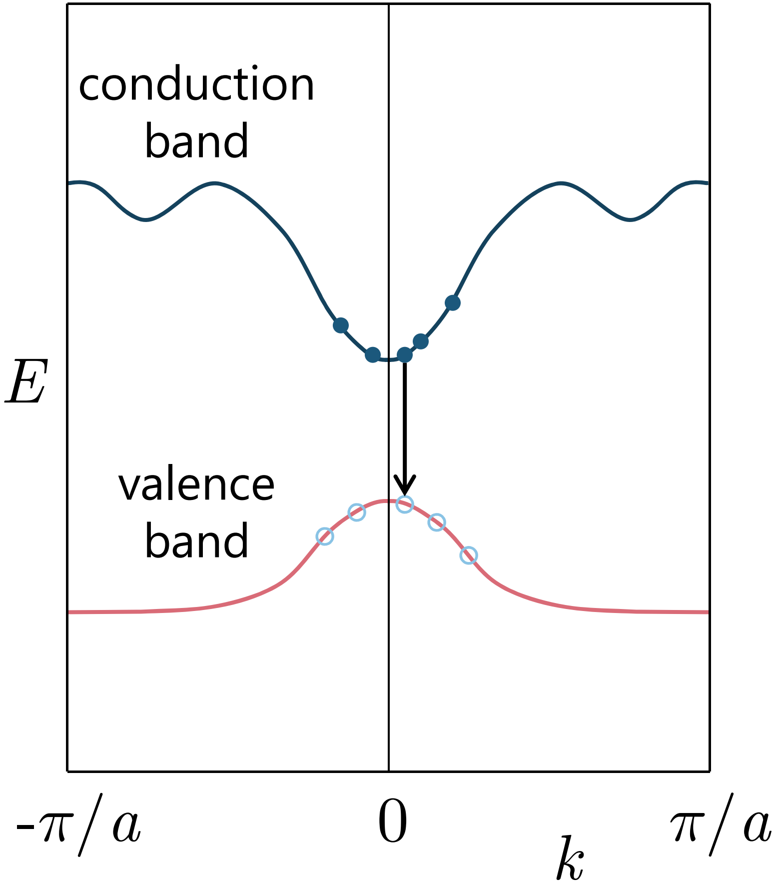
the semiconductor is said to have a

"direct gap"



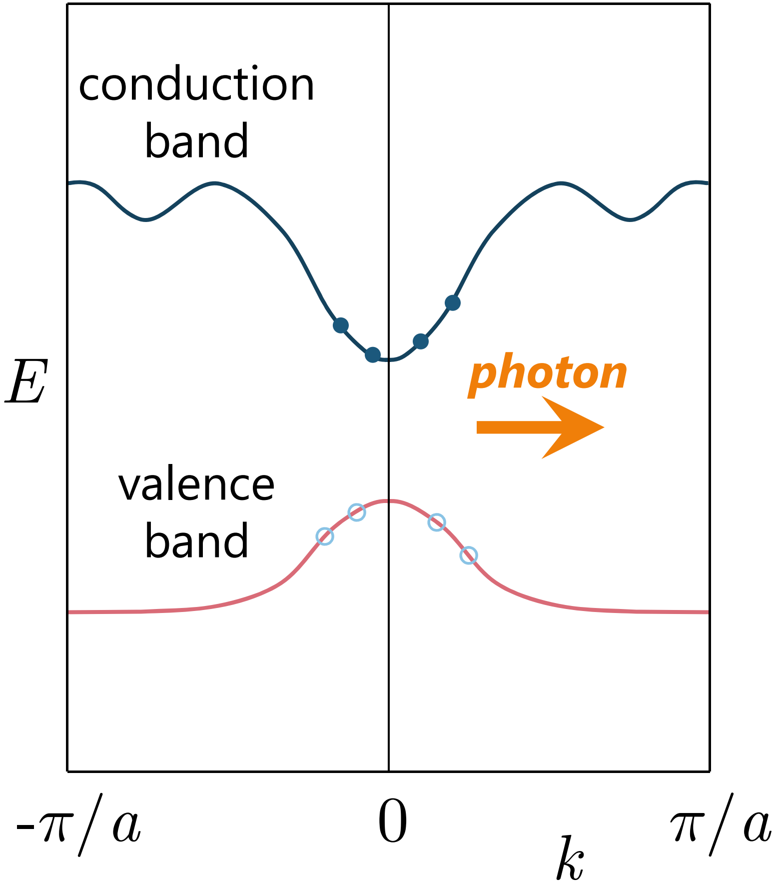
Direct gap semiconductor

- Direct gaps are important for light emitters
  - Electrons "pumped" into the conduction band gather in the lowest minimum
  - "Holes" pumped into the valence band gather in the highest maximum
- An electron can fall "vertically" to fill in a hole beneath it



Direct gap semiconductor

- Direct gaps are important for light emitters
  - Electrons "pumped" into the conduction band gather in the lowest minimum
  - "Holes" pumped into the valence band gather in the highest maximum
- An electron can fall "vertically" to fill in a hole beneath it
  - emitting light



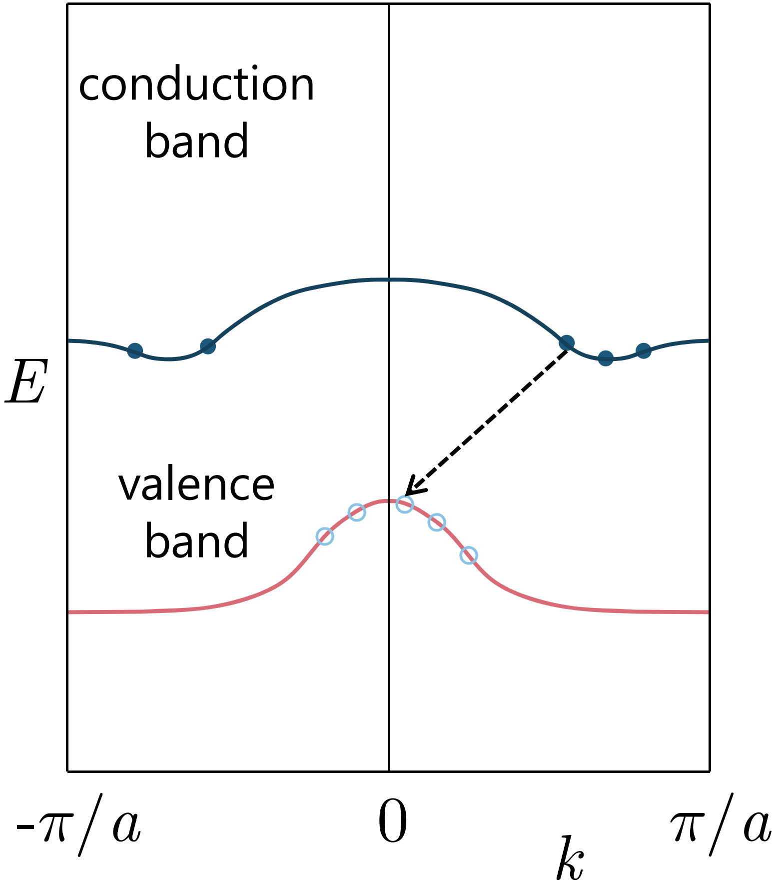
Indirect gap semiconductor

In an indirect gap semiconductor

- e.g., silicon, germanium
  - the lowest conduction band minimum (or minima)
  - is not directly above the highest valence band maximum

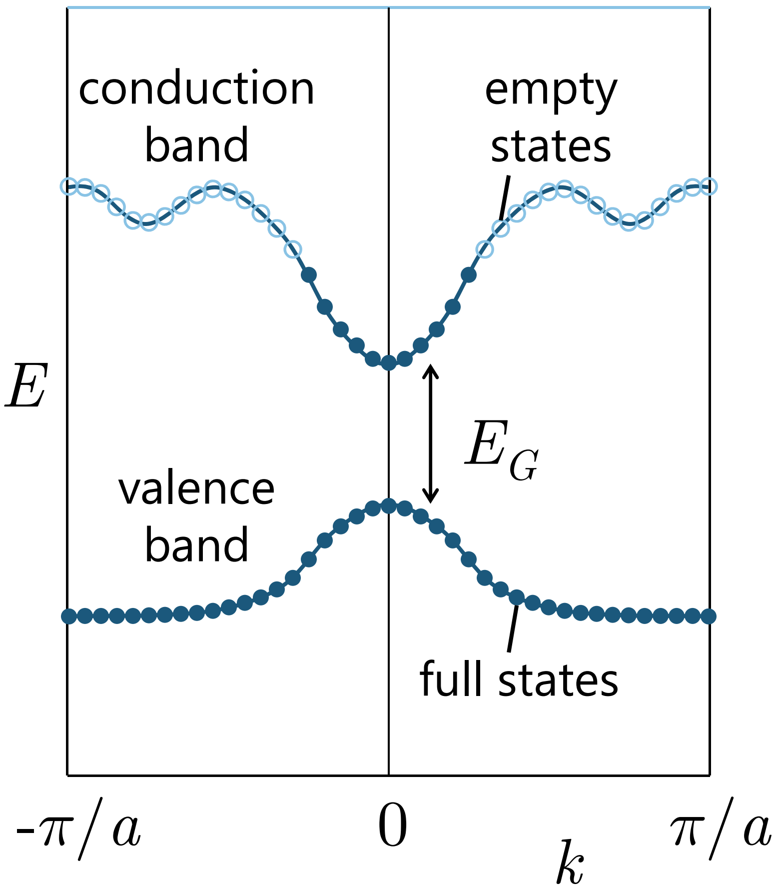
Light emission is weak

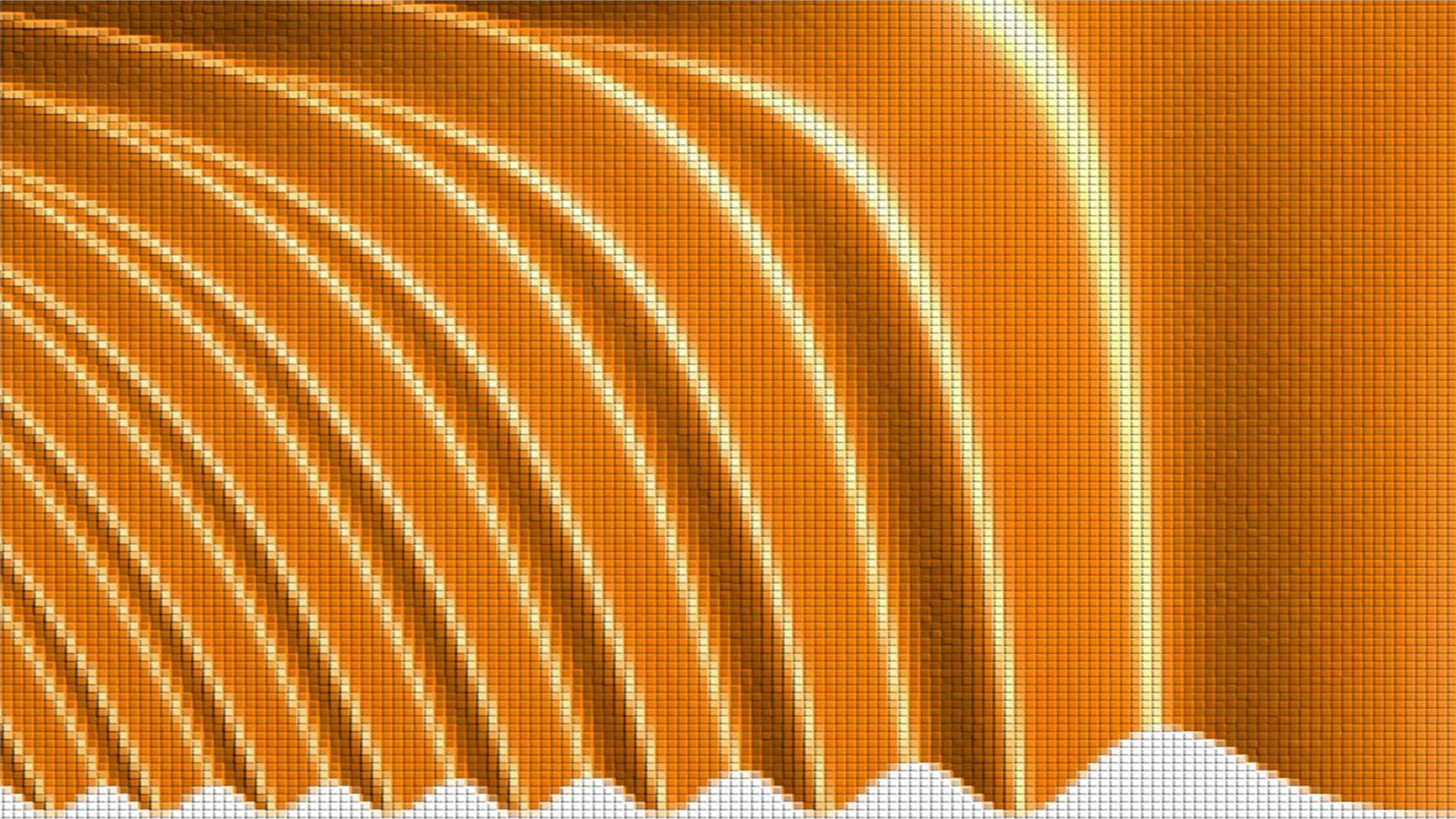
- "non-vertical" transitions by emission of photons are weak



Metals

- Because of the number of electrons in the metal atoms
  - the lowest conduction band is partially full of electrons
    - e.g., half-full
      - even at zero temperature
- Hence metals conduct electricity well



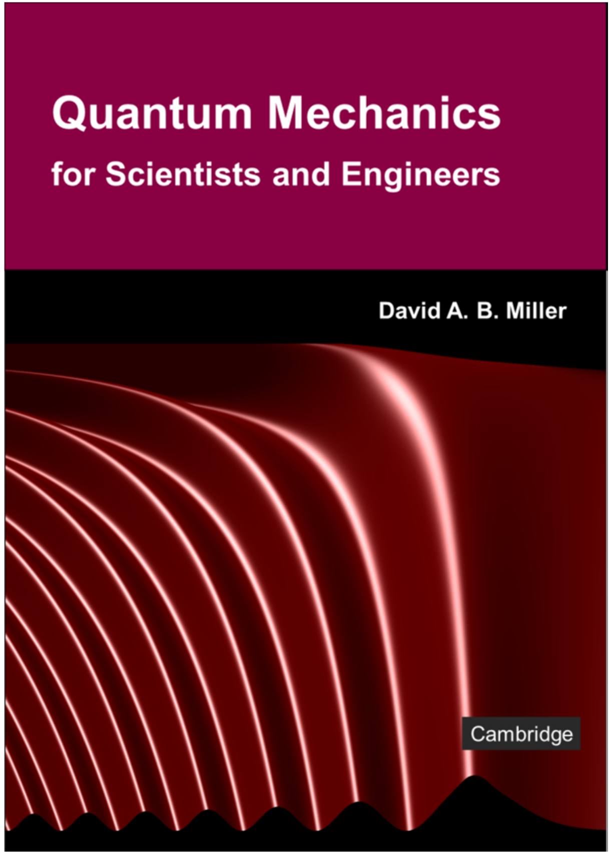


11.1 Band structures

Slides: Video 11.1.6 Band structures in 3D

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.5



Band structures

Band structures in 3D

Quantum mechanics for scientists and engineers

David Miller

Brillouin zone in 3D

This Brillouin zone for the diamond or zinc-blende lattice

is itself a 3D object

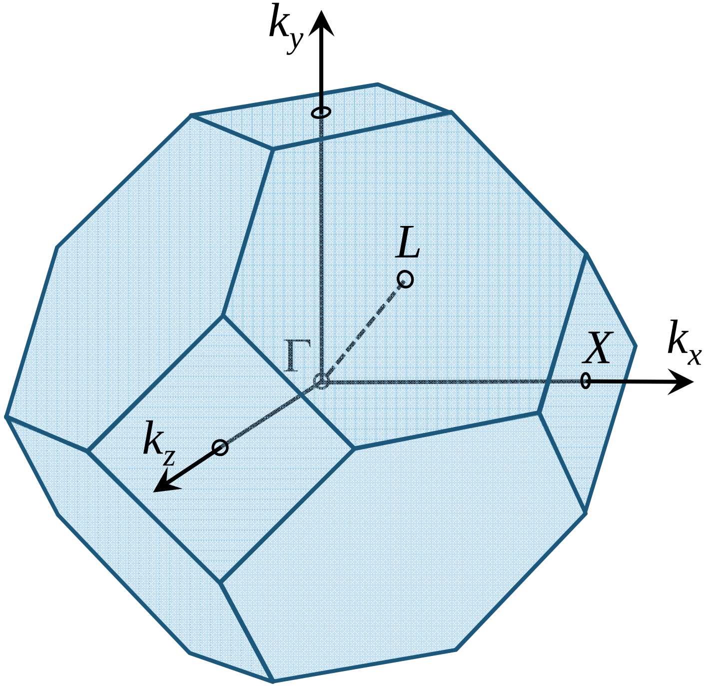
in k-space (or reciprocal space)

Two important directions are

X – along one of the x, y, or z coordinate directions

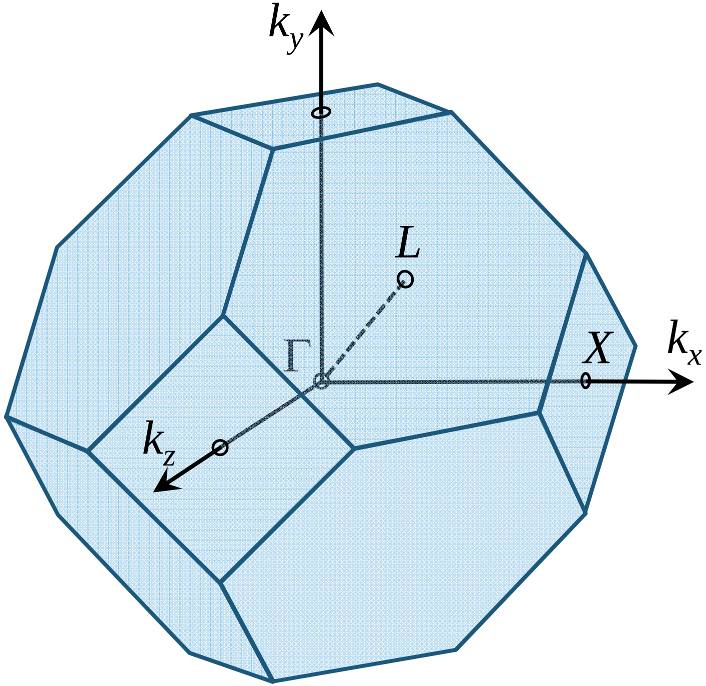
L – along one of the cube space diagonals

The center is the Γ (gamma) point



Band structures for 3D crystals

- At least as a first useful representation of band structure
  - typically the band structure is calculated only along a few directions
    - such as along the lines from the Γ point (at the center of the Brillouin zone)
      - to the X point and the L point



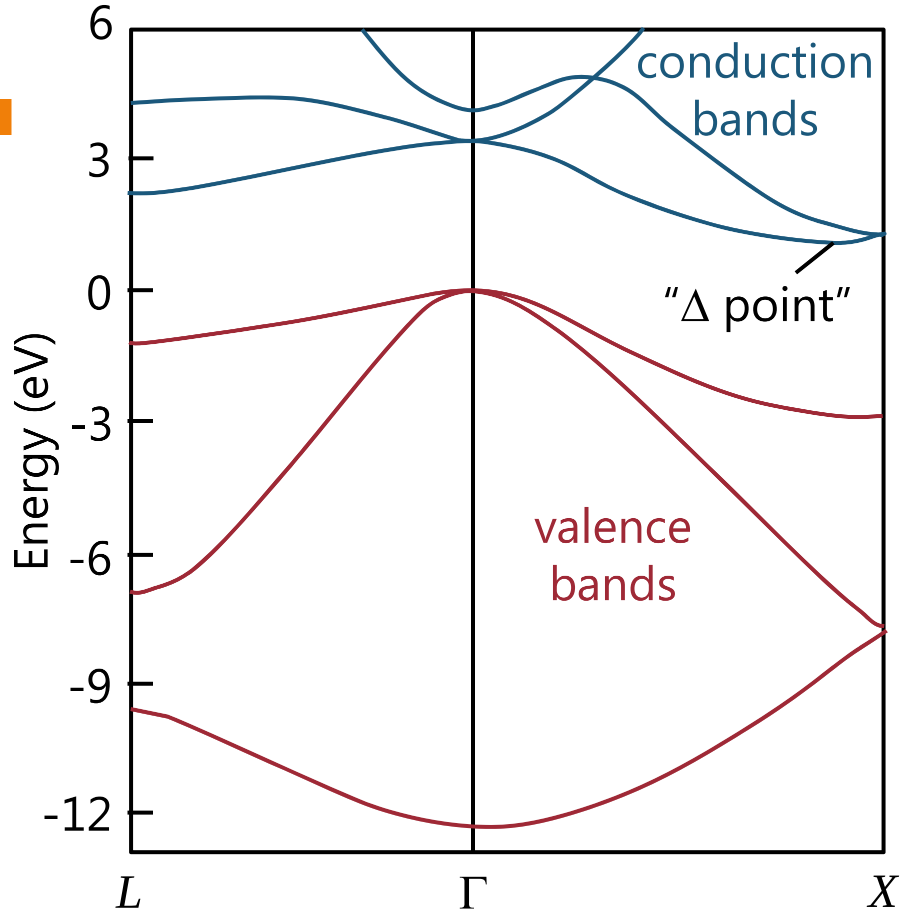
Si band structure

Sketch of major valence and conduction bands

- with the conduction band minimum at the Δ point

By Kramers degeneracy

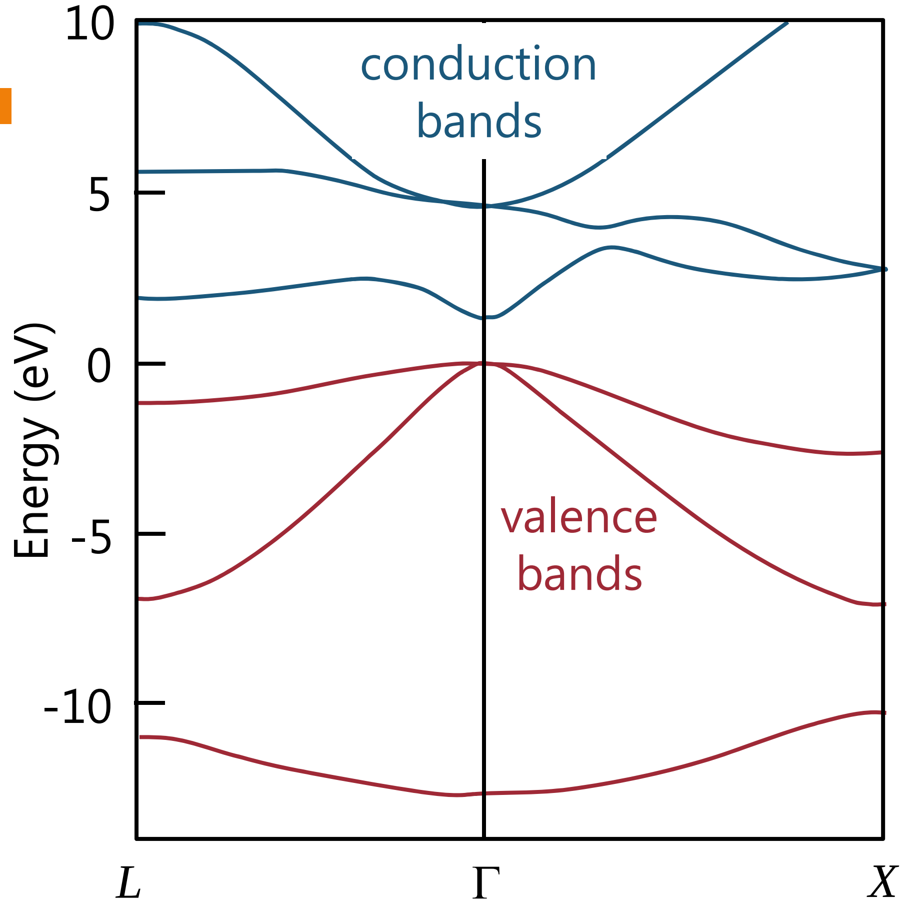
- we need only show one half of the band structure
  - so we can use the other half of the figure
    - for the band structure in another direction



after K. S. Sieh and P. V. Smith, Phys. Status Solidi (b) **129**, 259 (1985)

GaAs band structure

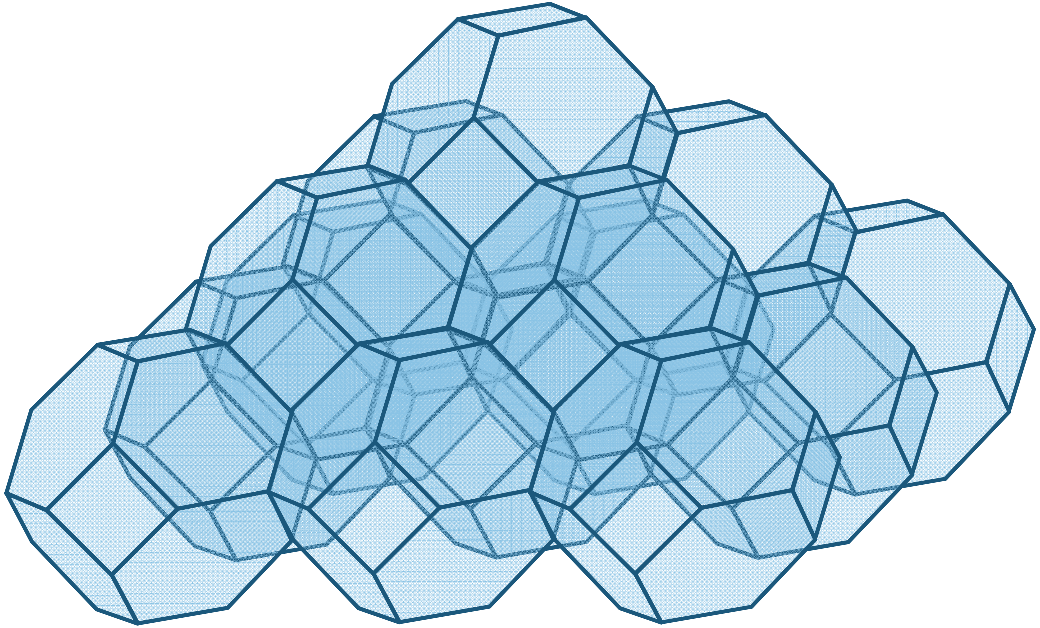
- Sketch of major valence and conduction bands
  - with the conduction band minimum at the Γ point
- Note that GaAs is a direct gap semiconductor
  - unlike Si
    - which is indirect



after M. Rohlfing, P. Krüger and J. Pollmann, Phys. Rev. B **48**, 17791 (1993)

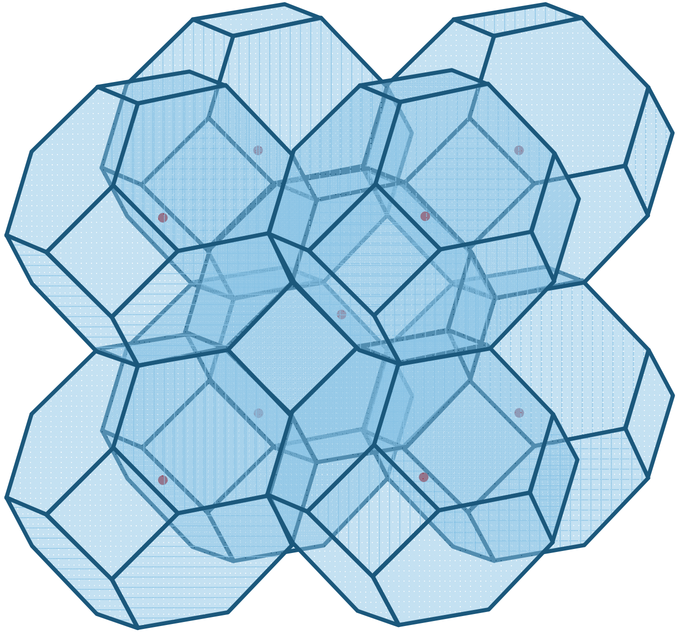
Extended zones in 3D

- In 3D
  - additional Brillouin zones repeat the same band structure
- These zones form “unit cells” in k-space
  - filling all k-space (reciprocal space)



Extended zones in 3D

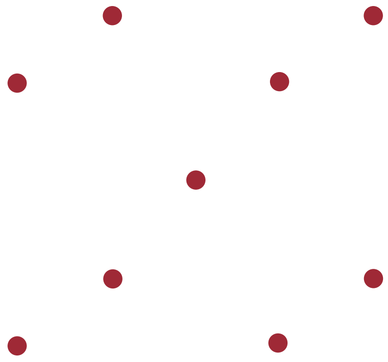
Marking the same “reciprocal lattice point” in each cell



Extended zones in 3D

Marking the same "reciprocal lattice point" in each cell

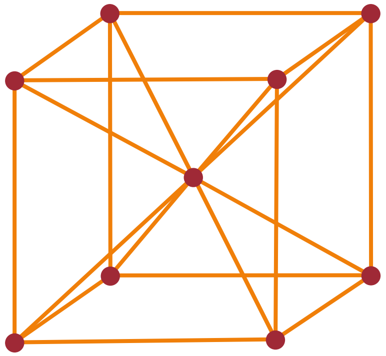
and erasing the "unit cells" (Brillouin zone shapes) themselves for clarity



Extended zones in 3D

Marking the same "reciprocal lattice point" in each cell

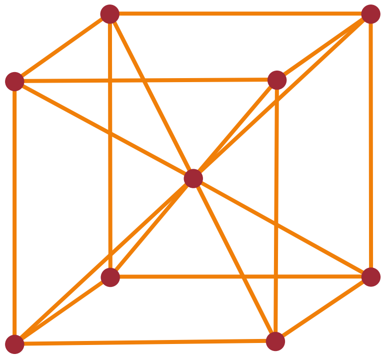
- and erasing the "unit cells" (Brillouin zone shapes) themselves for clarity
  - and adding guide lines
    - shows these extended Brillouin zones give a body centered cubic "reciprocal lattice"

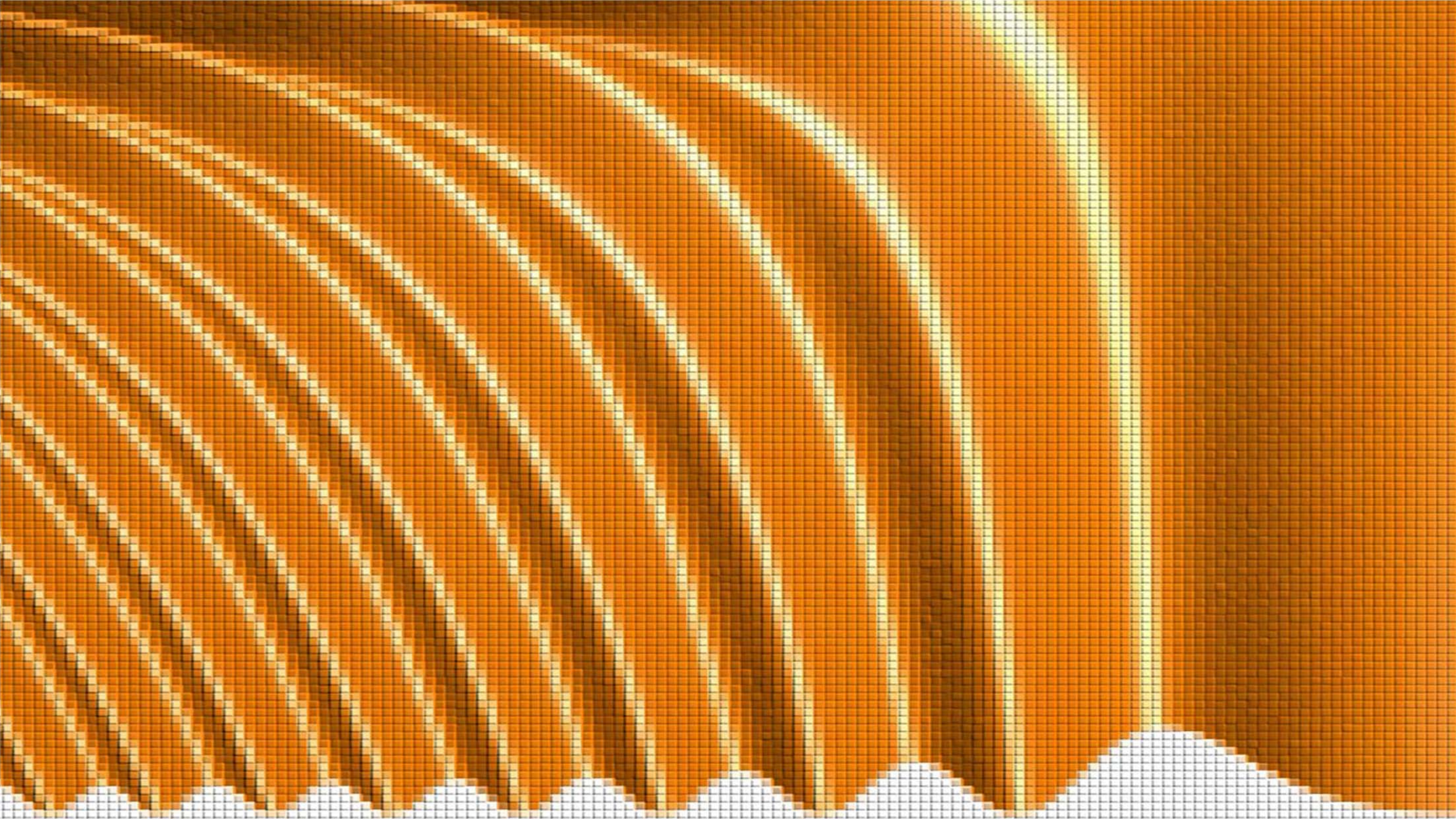


Extended zones in 3D

This particular reciprocal lattice with one mathematical lattice point for each Brillouin zone is typically the one meant when talking about "the reciprocal lattice"

The vectors in k-space between these lattice points are called reciprocal lattice vectors



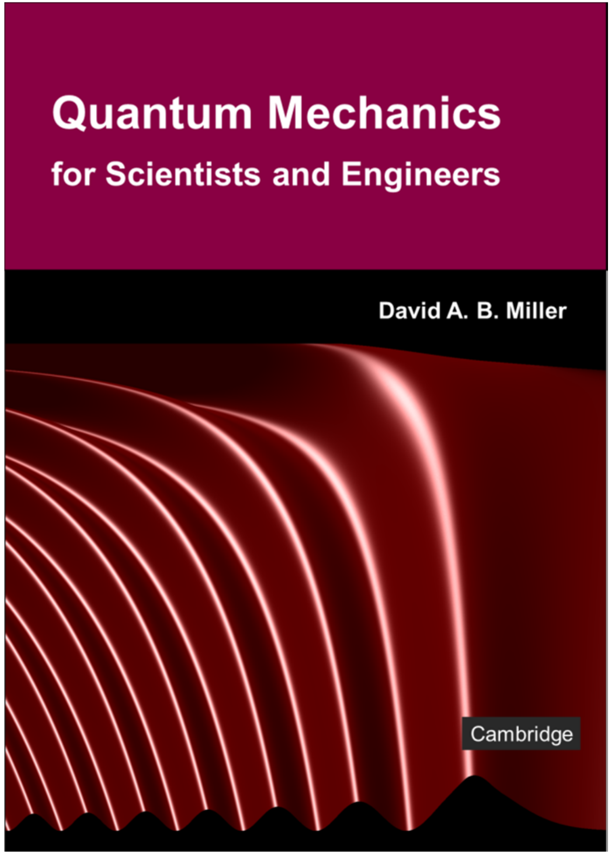


11.2 Effective mass theory

Slides: Video 11.2.1 Effective mass approximation

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.6 (up to ~ Eq. 8.29)



Effective mass theory

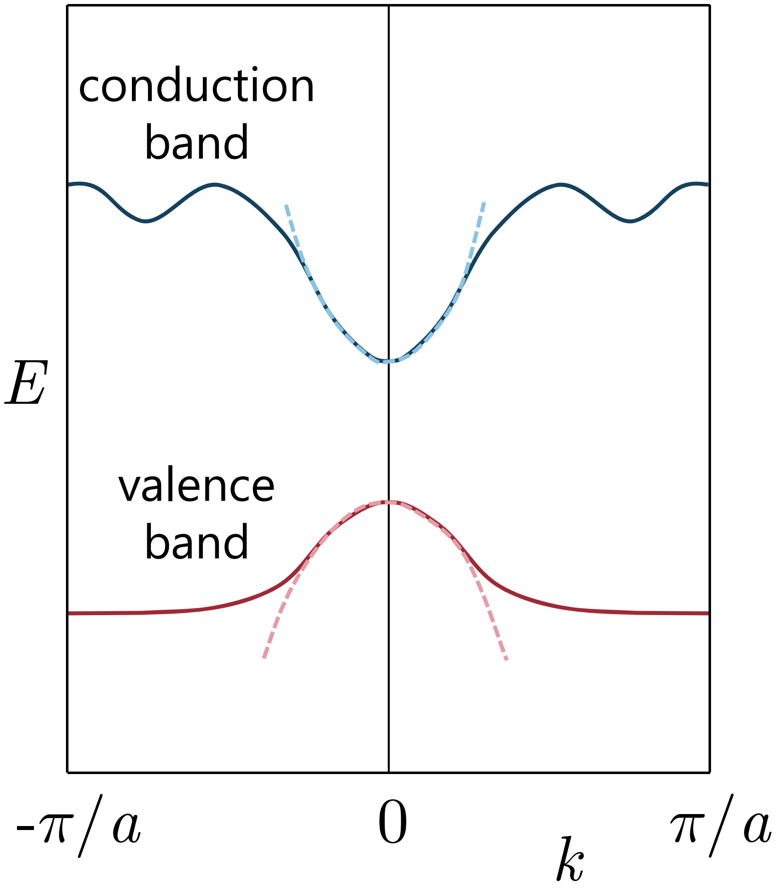
Effective mass approximation

Quantum mechanics for scientists and engineers

David Miller

Effective mass approximation

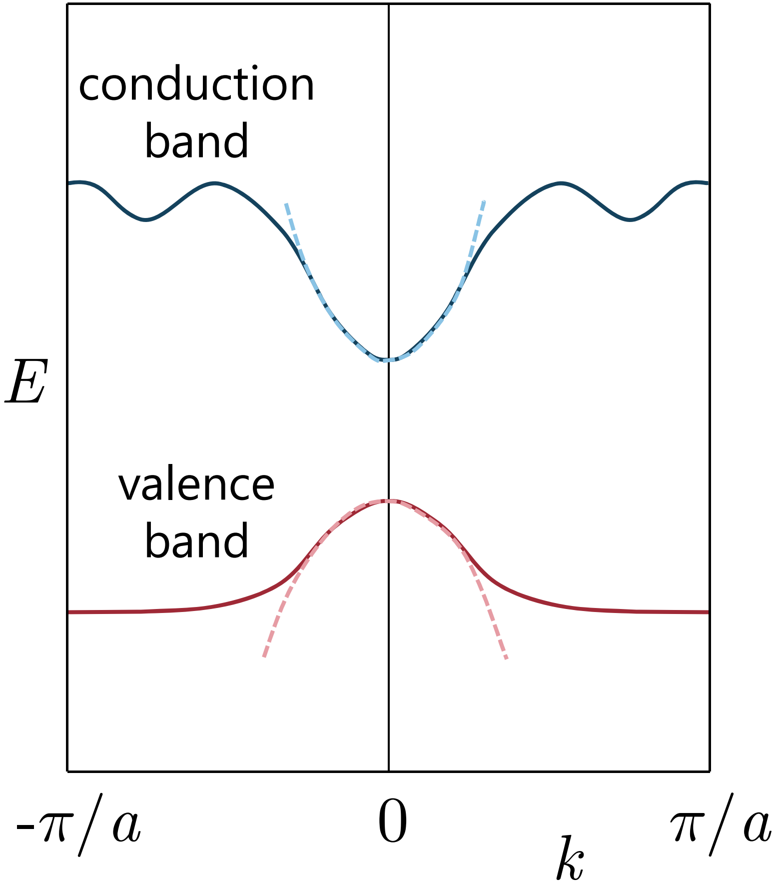
- Near a minimum or maximum, energy E should vary $\propto k^2$ in a given direction
- For simplicity we presume
  - the variation is isotropic and
  - the minimum or maximum of interest is located at $k = 0$
- Neither of these simplifications is necessary for this effective mass approach



Effective mass approximation

This isotropic $k = 0$ minimum or maximum is a good first approximation for

- the lowest conduction band
  - and
- the highest valence bands
  - in the direct gap
    - semiconductors important in optoelectronics
    - e.g., GaAs, InGaAs



Effective mass approximation

For the lowest conduction bands in indirect gap semiconductors

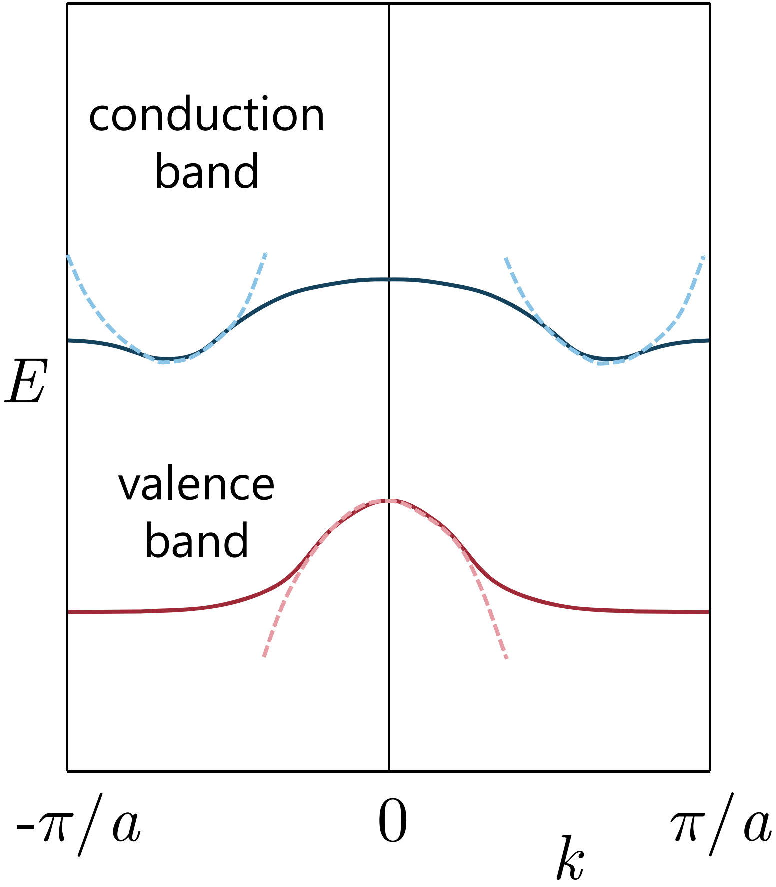
like silicon, germanium, AlAs

the minima are not at $k = 0$

and they are not isotropic

The theory is easily extended to cover these cases

though we will not do this here



Effective mass approximation

If the energy at the minimum or maximum itself is some amount V

then, by assumption, the energy $E_{\mathbf{k}}$ of the state in the band at wavevector $\mathbf{k}$ is

$$E_{\mathbf{k}} - V \propto k^2$$

For reasons that will become obvious, we choose to write this as

$$E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m_{eff}} + V$$

where the "effective mass" m_{eff} is a parameter that sets the appropriate proportionality

Effective mass approximation

A relation such as

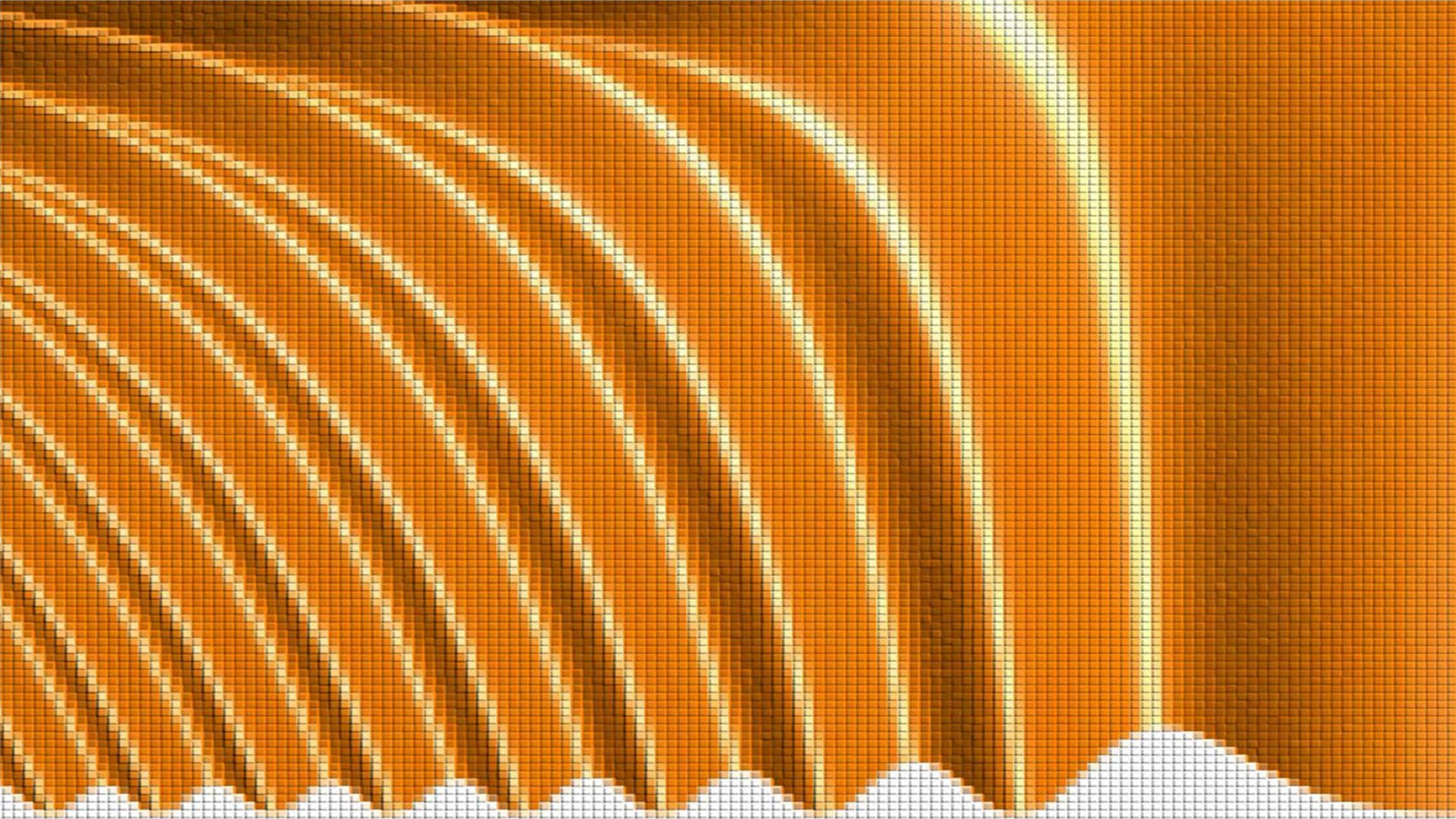
$$E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m_{eff}} + V$$

between energy and k-value is called a

is called a dispersion relation

This particular approximation for the behavior of the energies in a band is called

an isotropic parabolic band

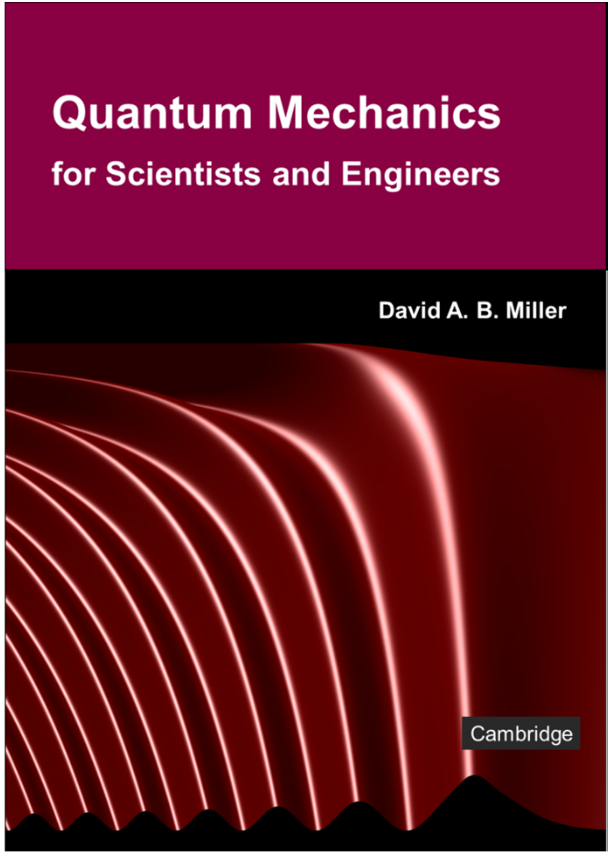


11.2 Effective mass theory

Slides: Video 11.2.3 Wavepackets and effective mass theory

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.6 (from ~ Eq. 8.29 to "Effective mass approximation …" subsection)



Effective mass theory

Wavepackets and effective mass theory

Quantum mechanics for scientists and engineers

David Miller

Behavior of a wavepacket

Consider a wave packet

a linear superposition of different Bloch states

Since we are going to consider the time evolution

we will also include the time-varying factor $\exp\left(-iE_{\mathbf{k}}t/\hbar\right)$ for each component in the superposition

Hence we consider a wavefunction

$$\Psi(\mathbf{r},t)=\sum_{\mathbf{k}}c_{\mathbf{k}}u_{\mathbf{k}}(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{r})\exp\left(-iE_{\mathbf{k}}t/\hbar\right)$$

where c_k are the coefficients of the different Bloch states in this superposition

Behavior of a wavepacket

We have restricted this superposition

$$\Psi(\mathbf{r},t)=\sum_{\mathbf{k}} c_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k}\cdot\mathbf{r}) \exp(-iE_{\mathbf{k}}t/\hbar)$$

to states within only one band

We will make the further assumption that

this superposition is only from a small range of k-states (near $k = 0$)

This is what can be called

a slowly varying envelope approximation

since it means that the resulting wavepacket does not vary rapidly in space

Behavior of a wavepacket

Because of this restriction to a small range of k

we can presume that, for all the k of interest to us

all of the unit cell functions $u_{\mathbf{k}}(\mathbf{r})$

are approximately the same

Hence we presume $u_{\mathbf{k}}(\mathbf{r}) \cong u_0(\mathbf{r})$

for the range of interest to us

Hence we can factor out this unit cell part, writing

$$\Psi(\mathbf{r},t) = u_0(\mathbf{r})\Psi_{env}(\mathbf{r},t)$$

where the envelope function can be written

$$\Psi_{env}(\mathbf{r},t) = \sum_{\mathbf{k}} c_{\mathbf{k}} \exp(i\mathbf{k}\cdot\mathbf{r})\exp(-iE_{\mathbf{k}}t/\hbar)$$

Behavior of a wavepacket

Now, differentiating with respect to time, we construct

a Schrödinger equation for this envelope function

$$i\hbar\frac{\partial\Psi_{env}}{\partial t}=i\hbar\frac{\partial}{\partial t}\sum_{\mathbf{k}}c_{\mathbf{k}}\exp\left(i\mathbf{k}\cdot\mathbf{r}\right)\exp\left(-iE_{\mathbf{k}}t/\hbar\right)=\sum_{\mathbf{k}}c_{\mathbf{k}}E_{\mathbf{k}}\exp\left(i\mathbf{k}\cdot\mathbf{r}\right)\exp\left(-iE_{\mathbf{k}}t/\hbar\right)$$

$$=\frac{\hbar^2}{2m_{eff}}\sum_{\mathbf{k}}c_{\mathbf{k}}k^2\exp\left(i\mathbf{k}\cdot\mathbf{r}\right)\exp\left(-iE_{\mathbf{k}}t/\hbar\right)+V\sum_{\mathbf{k}}c_{\mathbf{k}}\exp\left(i\mathbf{k}\cdot\mathbf{r}\right)\exp\left(-iE_{\mathbf{k}}t/\hbar\right)$$

$$=\frac{\hbar^2}{2m_{eff}}\sum_{\mathbf{k}}\left[-c_{\mathbf{k}}\nabla^2\exp\left(i\mathbf{k}\cdot\mathbf{r}\right)\right]\exp\left(-iE_{\mathbf{k}}t/\hbar\right)+V\Psi_{env}$$

since $\nabla^2\exp\left(i\mathbf{k}\cdot\mathbf{r}\right)=-k^2\exp\left(i\mathbf{k}\cdot\mathbf{r}\right)$

Effective mass approximation

Taking $-\nabla^2$ outside the sum, we have

$$i\hbar\frac{\partial\Psi_{env}}{\partial t}=-\frac{\hbar^2}{2m_{eff}}\nabla^2\sum_{\mathbf{k}}c_{\mathbf{k}}\exp\left(i\mathbf{k}\cdot\mathbf{r}\right)\exp\left(-iE_{\mathbf{k}}t/\hbar\right)+V\Psi_{env}$$

Hence, we have managed to construct a Schrödinger equation for this envelope function

$$-\frac{\hbar^2}{2m_{eff}}\nabla^2\Psi_{env}(\mathbf{r},t)+V\left(\mathbf{r}\right)\Psi_{env}\left(\mathbf{r},t\right)=i\hbar\frac{\partial}{\partial t}\Psi_{env}\left(\mathbf{r},t\right)$$

So we can approximately treat the electron

as a particle with an effective mass m_{eff} and a wavefunction given by the envelope function $\Psi_{env}\left(\mathbf{r},t\right)$

Effective mass approximation

In $-\frac{\hbar^2}{2m_{eff}}\nabla^2\Psi_{env}(\mathbf{r},t)+V(\mathbf{r})\Psi_{env}(\mathbf{r},t)=i\hbar\frac{\partial}{\partial t}\Psi_{env}(\mathbf{r},t)$

all of the details of the periodic potential and the unit cell wavefunction

have been suppressed

Their consequences are all contained in the single parameter, the effective mass m_{eff}

This effective mass model is a very powerful for modeling processes in semiconductors

Use of V(r)

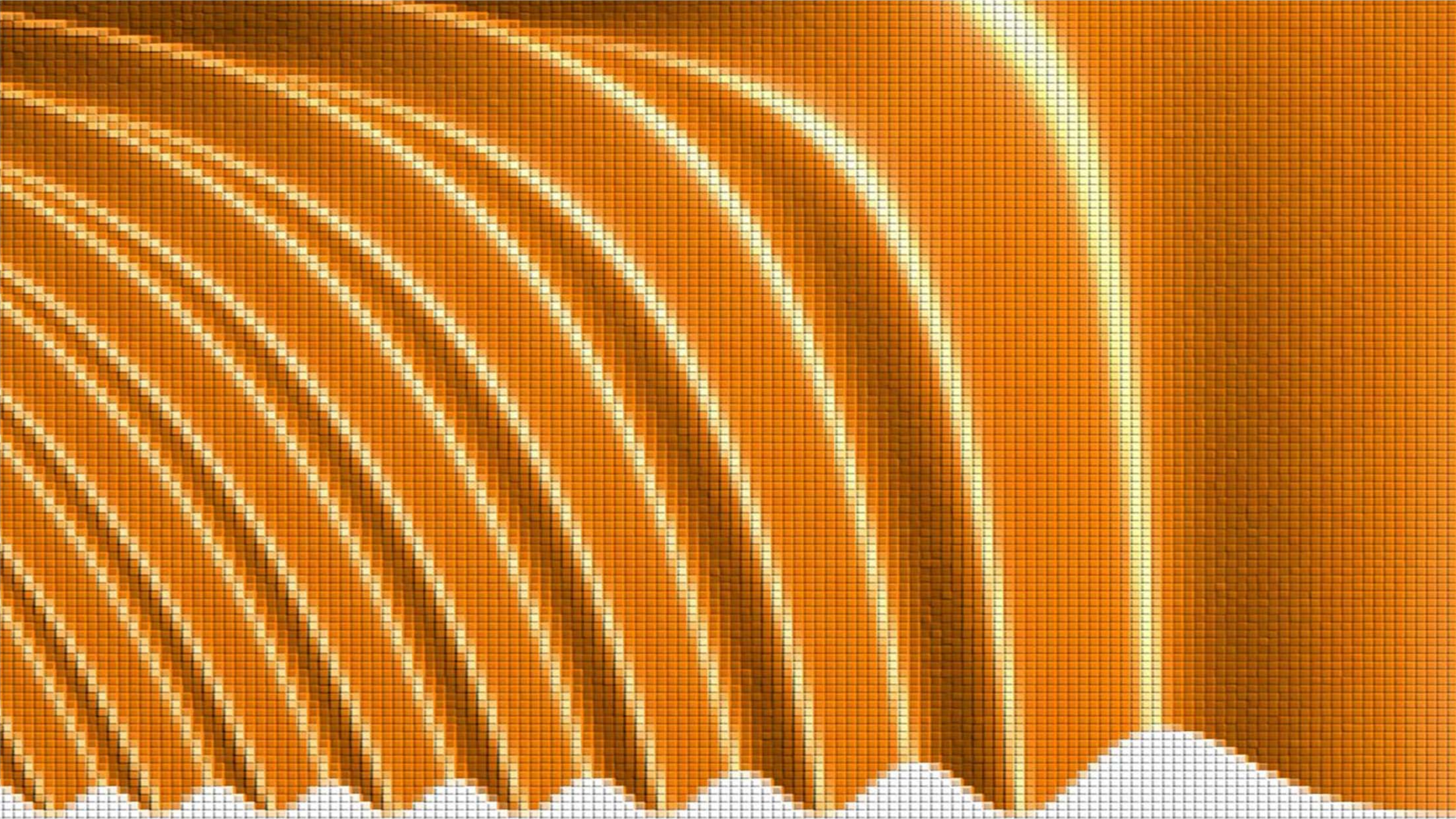
Note in the Schrödinger equation we are now writing for the envelope function

$$-\frac{\hbar^2}{2m_{eff}}\nabla^2\Psi_{env}(\mathbf{r},t)+V\left(\mathbf{r}\right)\Psi_{env}\left(\mathbf{r},t\right)=i\hbar\frac{\partial}{\partial t}\Psi_{env}\left(\mathbf{r},t\right)$$

we have allowed the energy of the band at $k=0$ i.e., the potential $V(\mathbf{r})$, to vary with position $\mathbf{r}$

We can argue this is allowable if the changes in that potential are very small compared to $\hbar^2 k^2 / 2m_{eff}$

over the scale of a unit cell and over the wavelength $2\pi / k$

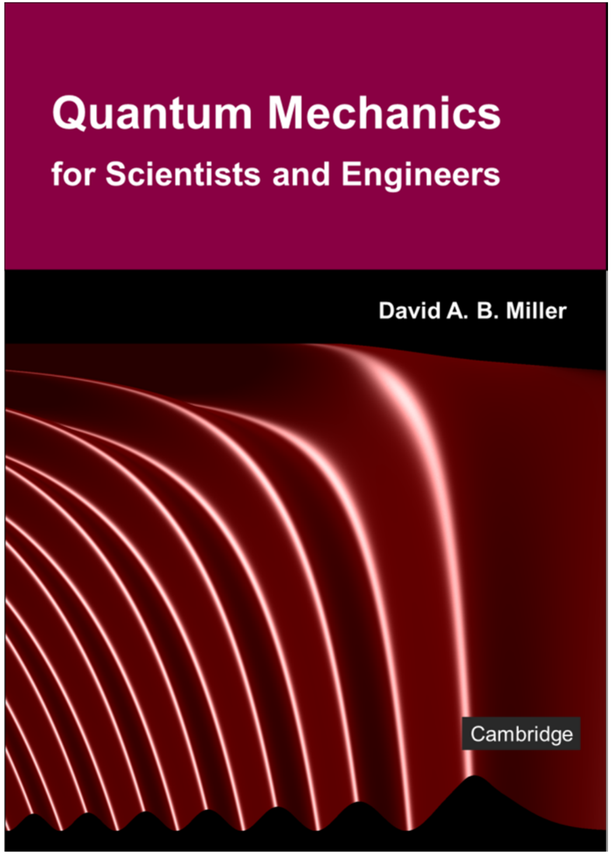


11.2 Effective mass theory

Slides: Video 11.2.5 Semiconductor heterostructures

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.6 ("Effective mass approximation …" subsection)



Effective mass theory

Semiconductor heterostructures

Quantum mechanics for scientists and engineers

David Miller

Semiconductor heterostructures

Structures involving more than one kind of material are called heterostructures

- e.g., changing x in the semiconductor $Al_xGa_{1-x}As$.

Such changes are made routinely in modern semiconductor structures

- especially abrupt changes in material concentration
  - e.g., the interface between $GaAs$ and $Al_{0.3}Ga_{0.7}As$ in laser diodes
  - quantum well structures involving very thin layers (e.g., 10 nm)

Analyzing heterostructures with effective mass theory

Note that the effective mass is in general different in different materials

It is then better to write the envelope function equation as

$$-\frac{\hbar^2}{2}\nabla\cdot\left[\frac{1}{m_{eff}}\nabla\Psi_{env}(\mathbf{r},t)\right]+V(\mathbf{r})\Psi_{env}(\mathbf{r},t)=i\hbar\frac{\partial}{\partial t}\Psi_{env}(\mathbf{r},t)$$

and to use boundary conditions such as

Ψ_{env} continuous

and $\frac{1}{m_{eff}}\nabla\Psi_{env}$ continuous

to handle abrupt changes in material and/or potential

Semiconductor heterostructures

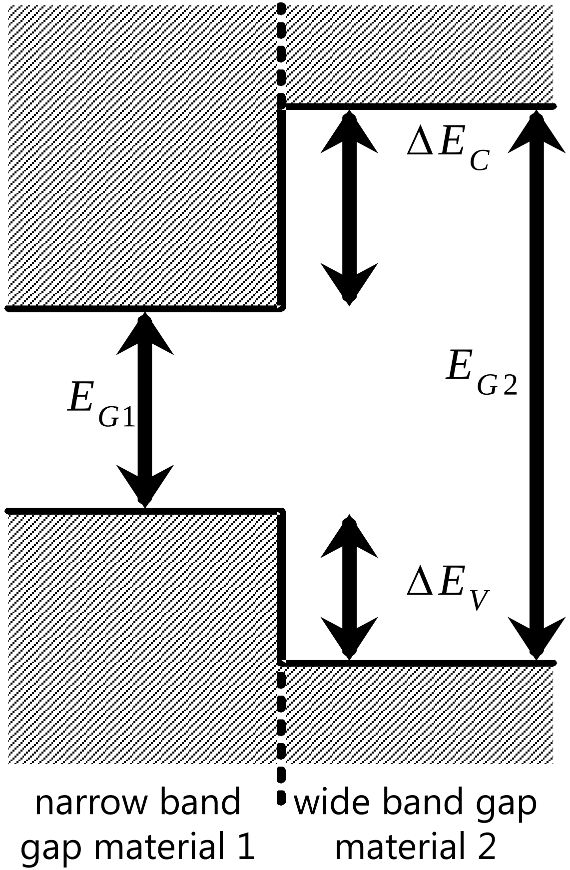
We treat the "band offsets" between the different materials as

abrupt changes in potential energy V as we go from one material to the other

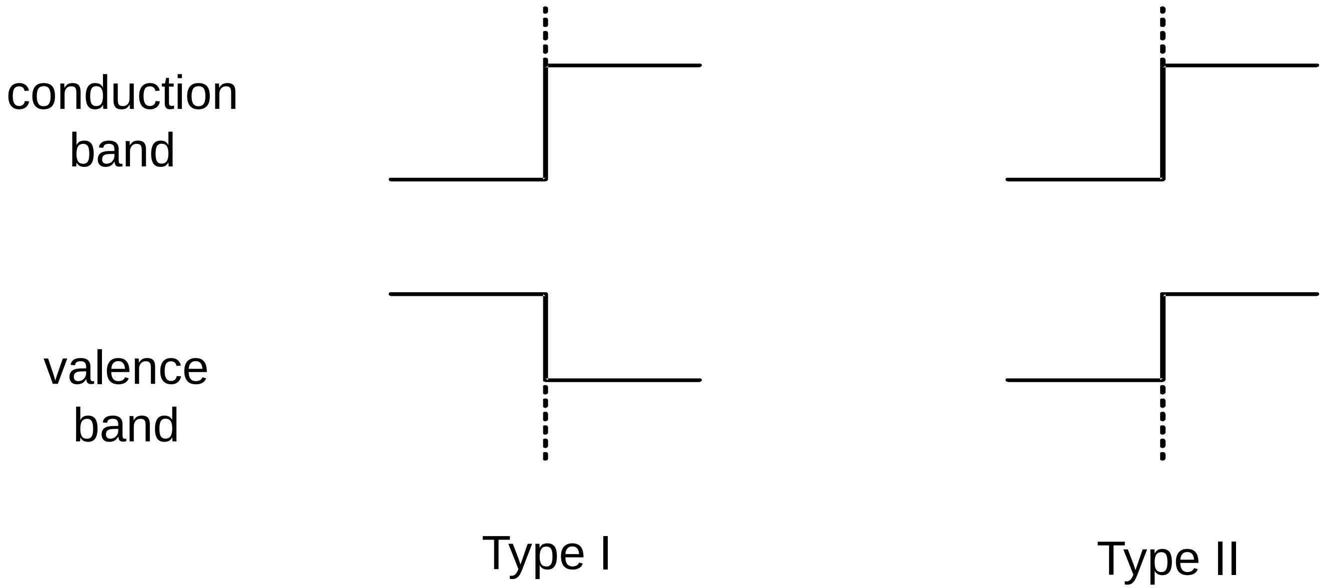
Offsets are usually determined empirically for given materials

conduction band offset ΔE_C

valence band offset ΔE_V



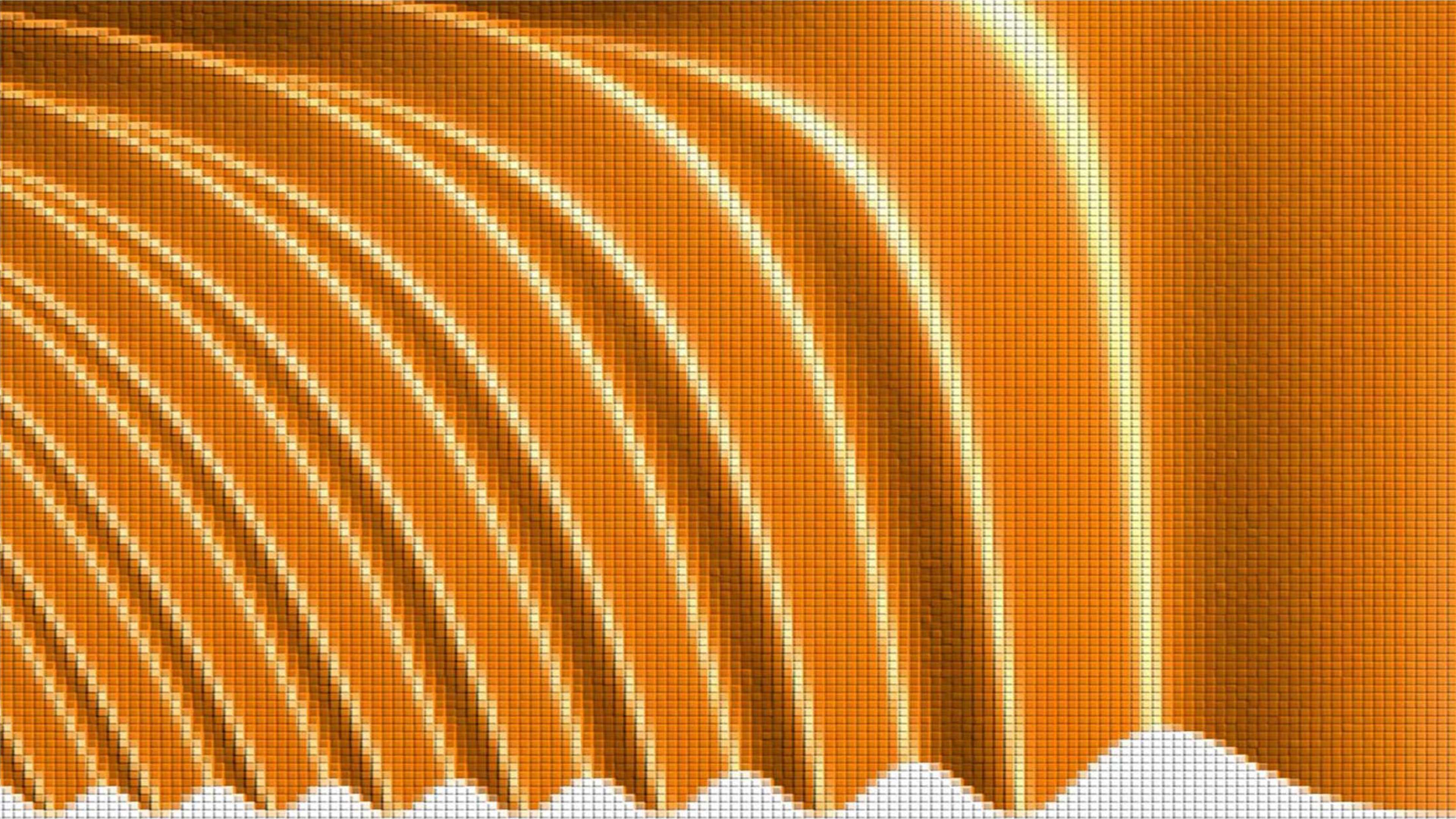
Classes of band line-up in heterostructures



Nearly all devices use "Type I" heterostructures

Electrons and holes have "lowest" energy in the same material

Remember holes "bubble up" to the top, where they have "lowest" energy

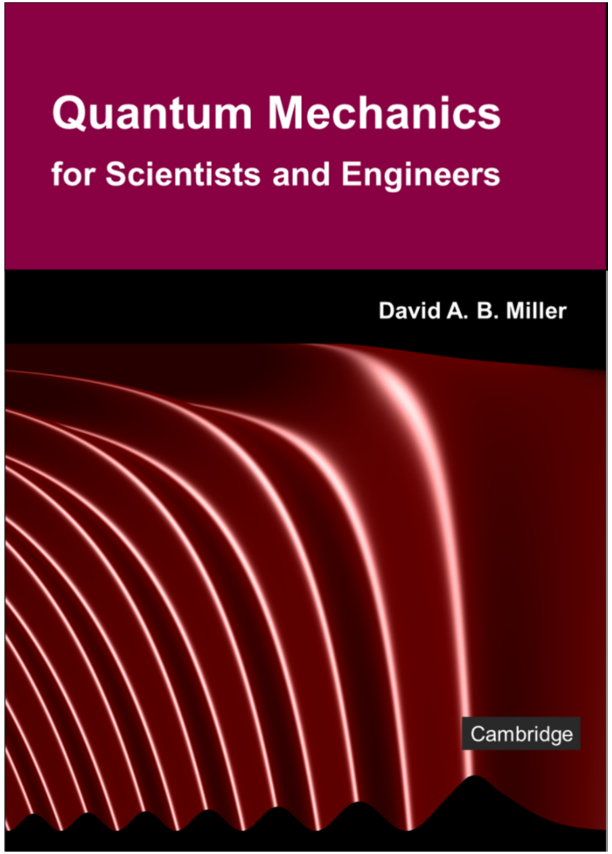


11.3 Density of states

Slides: Video 11.3.1 Energy density of states in bulk crystals

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.7



Density of states

Energy density of states in bulk crystals

Quantum mechanics for scientists and engineers

David Miller

Density of states in energy

The density of states in k-space (per unit real volume) is

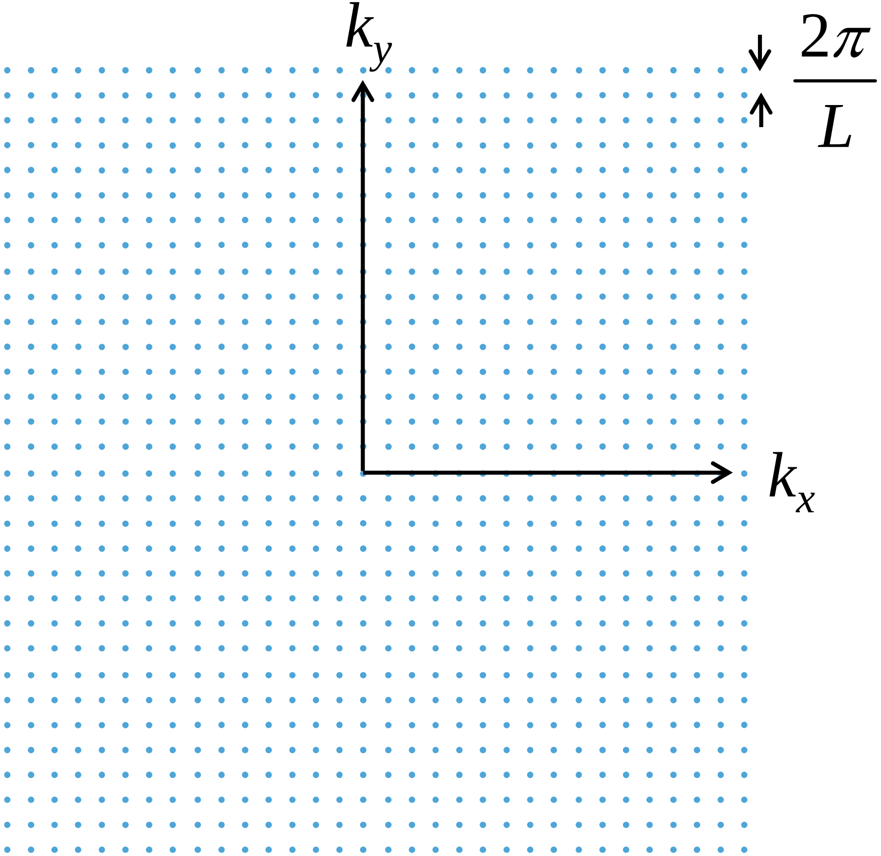
$$g(\mathbf{k}) = \frac{1}{(2\pi)^3}$$

which is constant

independent of the form of the band structure

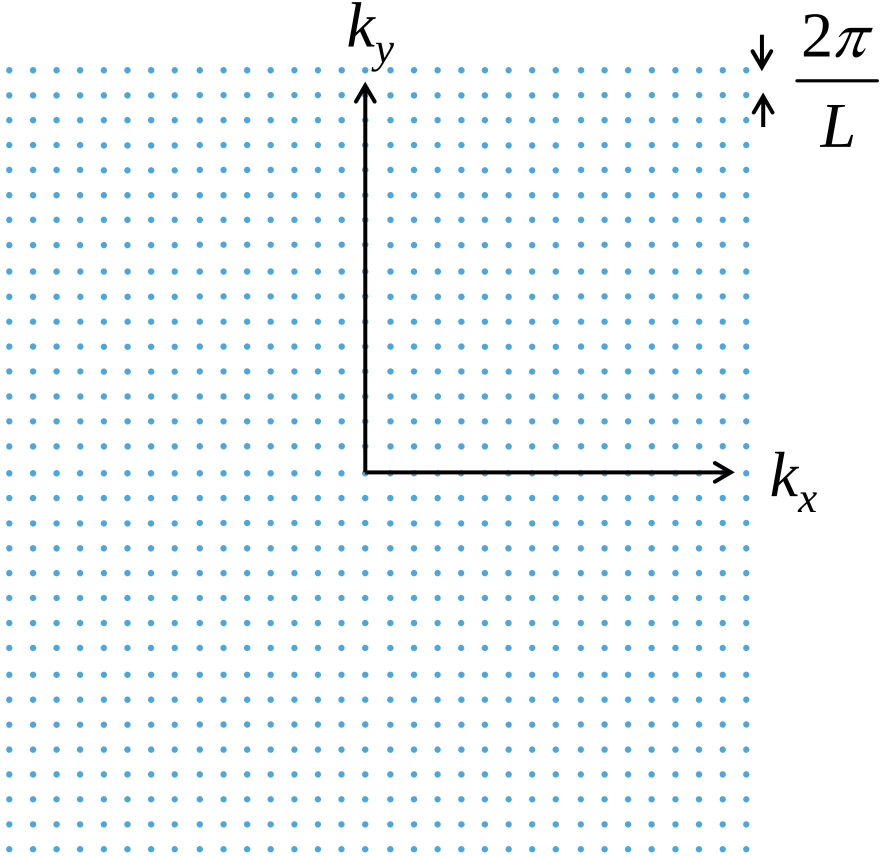
Here for simplicity we presume the crystal is a cube

of side L



Density of states in energy

- Now we want the density of states in energy per unit real volume $g(E)$
  - which depends on band structure
    - so we need the relation between the electron energy, E, and $\mathbf{k}$
- Now we work out that density of states for an isotropic parabolic band



Density of states in energy

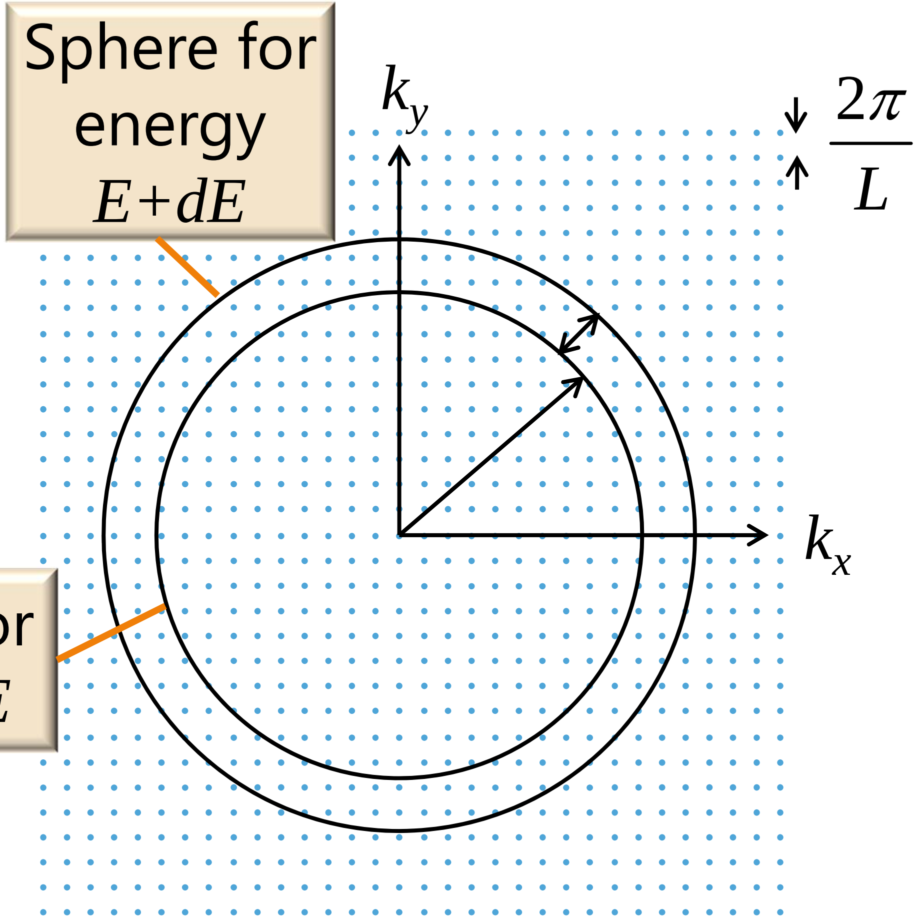
Since by assumption in our isotropic parabolic band

$$E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m_{eff}} + V$$

the number of states between energies E and $E+dE$

i.e., $g(E)dE$

is the number of states in k-space in the "shell" between the two spherical surfaces



Sphere for energy $E+dE$

Sphere for energy E

Density of states in energy

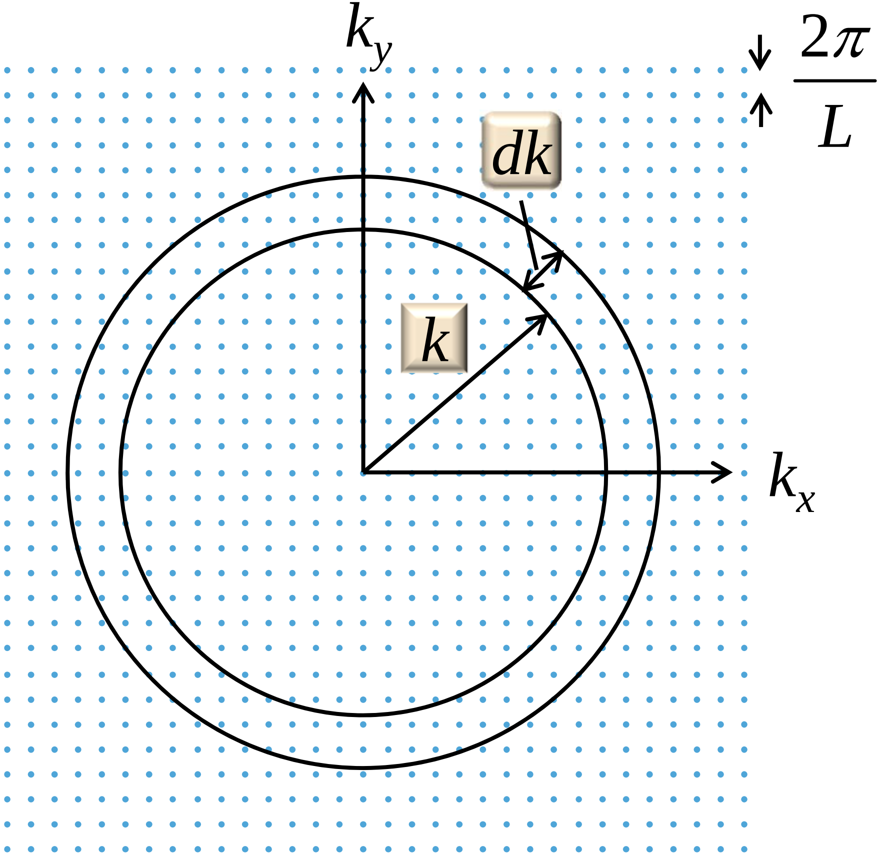
The radius of the inner sphere

from $$E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m_{eff}} + V$$

is $$k = \sqrt{\frac{2m_{eff}}{\hbar^2}\left(E - V\right)}$$

The radius of the outer sphere is $k + dk$, where

$$dk = \left(\frac{dk}{dE}\right) dE$$



Density of states in energy

From $k = \sqrt{\frac{2m_{eff}}{\hbar^2}\left(E - V\right)}$

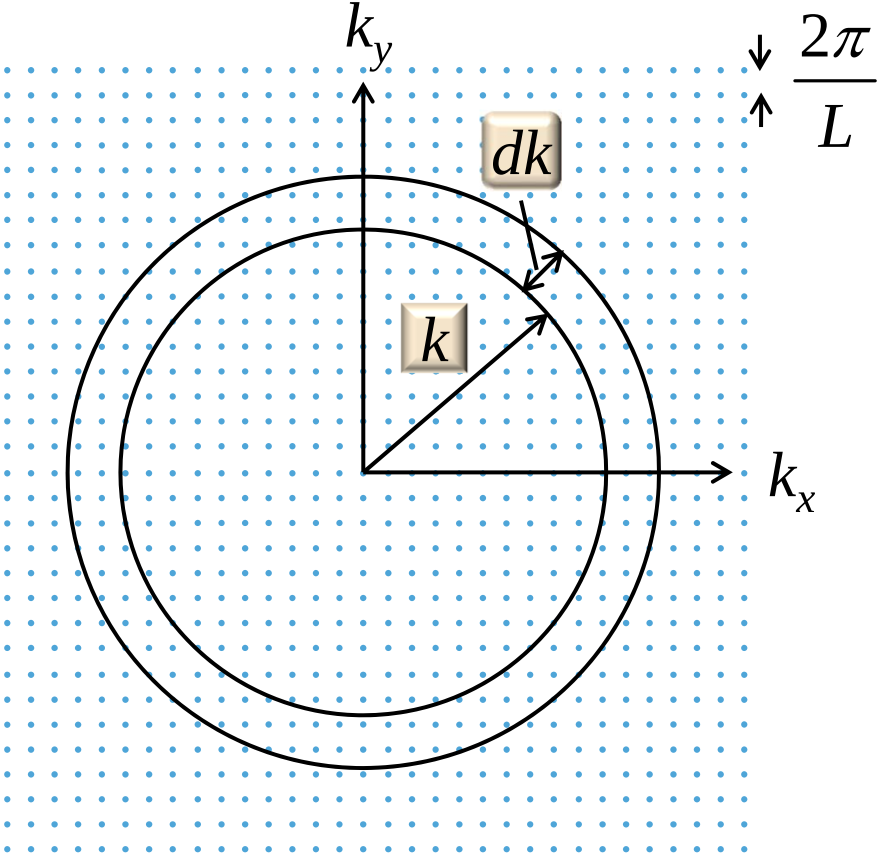
we deduce

$$\frac{dk}{dE} = \frac{1}{2}\sqrt{\frac{2m_{eff}}{\hbar^2}}\frac{1}{\sqrt{E-V}}$$

so the volume of the spherical shell in *k*-space is

$$4\pi k^2 dk = 4\pi k^2 \frac{dk}{dE} dE$$

$$= 2\pi\left(2m_{eff} / \hbar^2\right)^{3/2}\left(E - V\right)^{1/2}$$



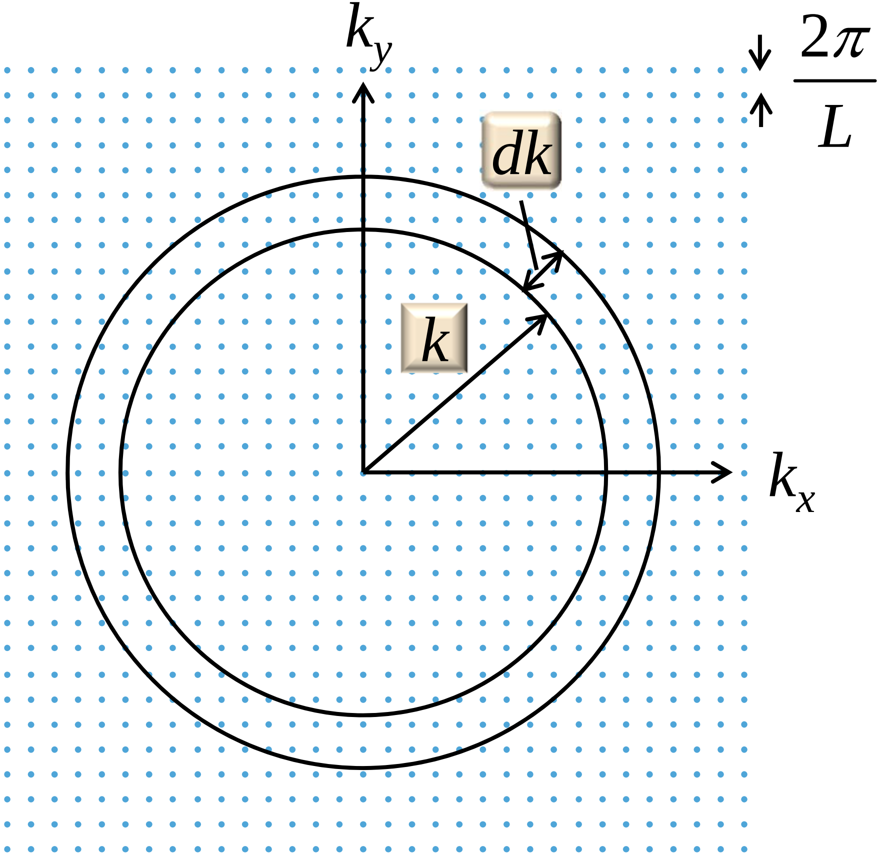
Density of states in energy

Since the density of states in k-space is

$$g(\mathbf{k}) = 1/(2\pi)^3$$

the total number of k-states in this spherical shell is

$$\frac{1}{(2\pi)^3} 2\pi \left(\frac{2m_{eff}}{\hbar^2}\right)^{3/2} (E-V)^{1/2}$$



Density of states in energy

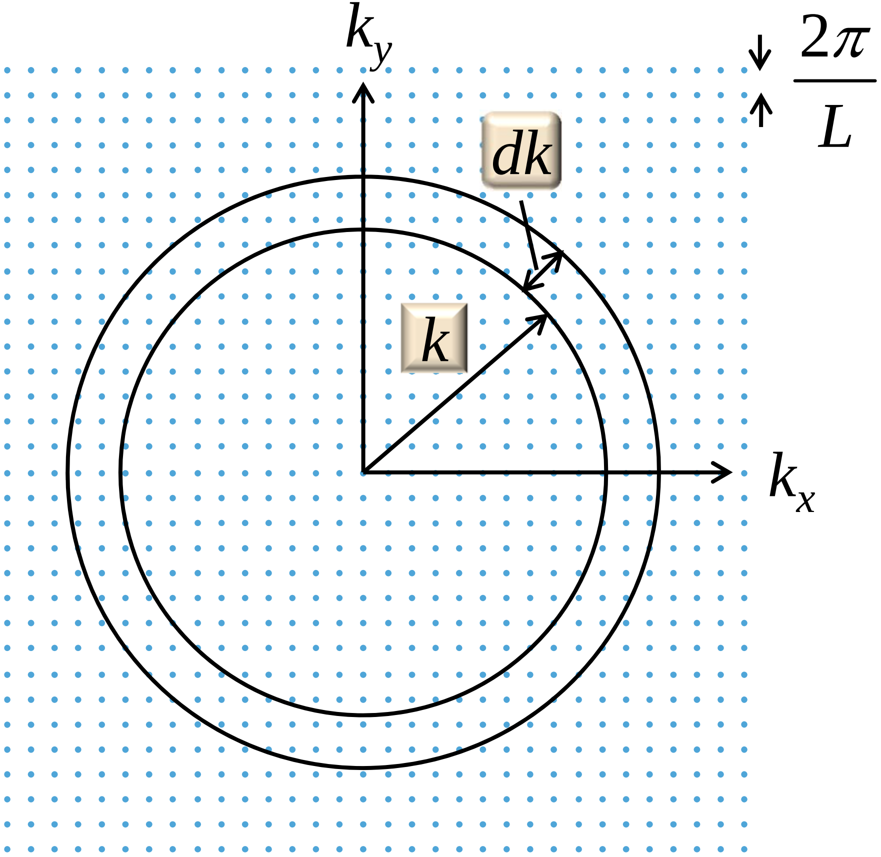
We now introduce the idea that the electron has two possible spin states, so

the number of states between energy E and $E+dE$ is

$$g(E) = 2\frac{1}{(2\pi)^3}2\pi\left(\frac{2m_{eff}}{\hbar^2}\right)^{3/2}(E-V)^{1/2}$$

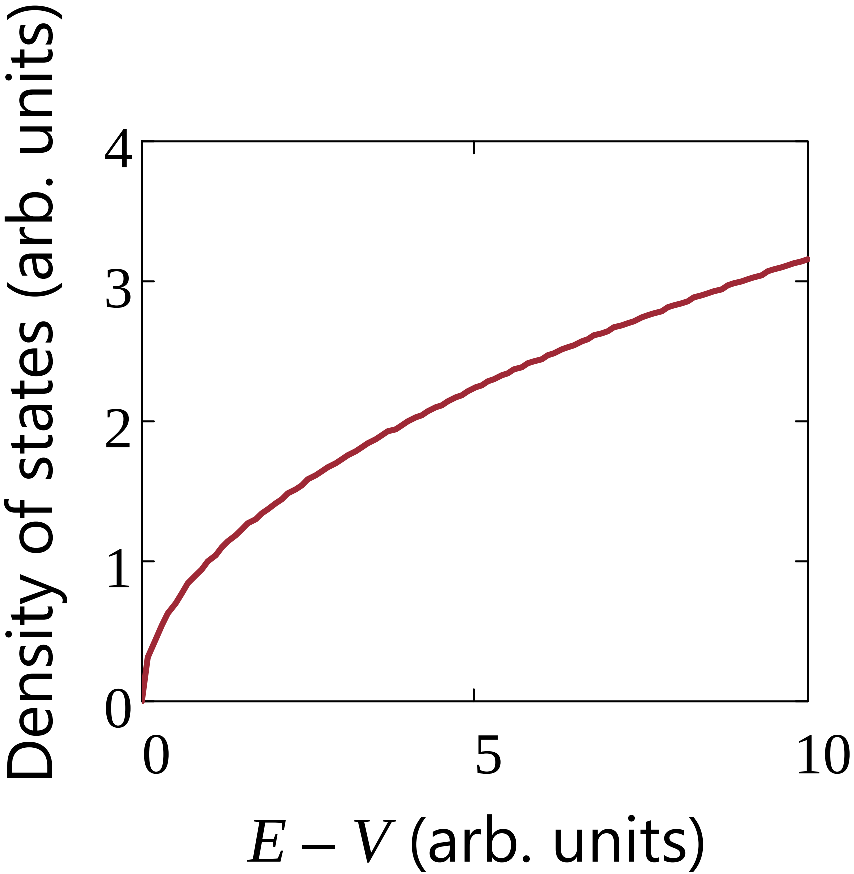
i.e.,

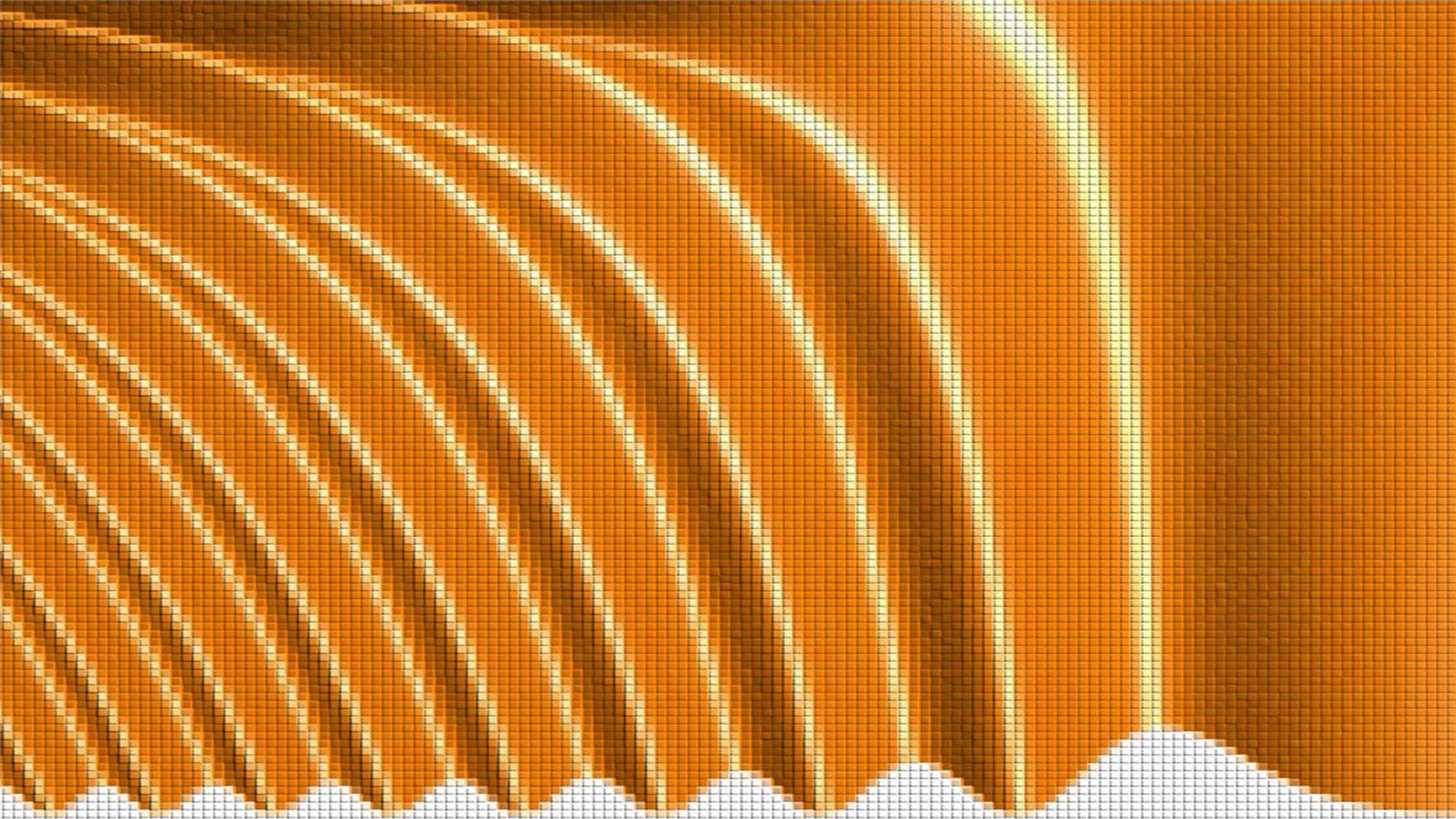
$$\boxed{g(E) = \frac{1}{2\pi^2}\left(\frac{2m_{eff}}{\hbar^2}\right)^{3/2}(E-V)^{1/2}}$$



Density of states in energy

- This gives the classic "$E^{1/2}$" density of states
- As the energy E rises above the energy of the bottom of the "parabola"
  - the density of states rises as the square root of the extra energy



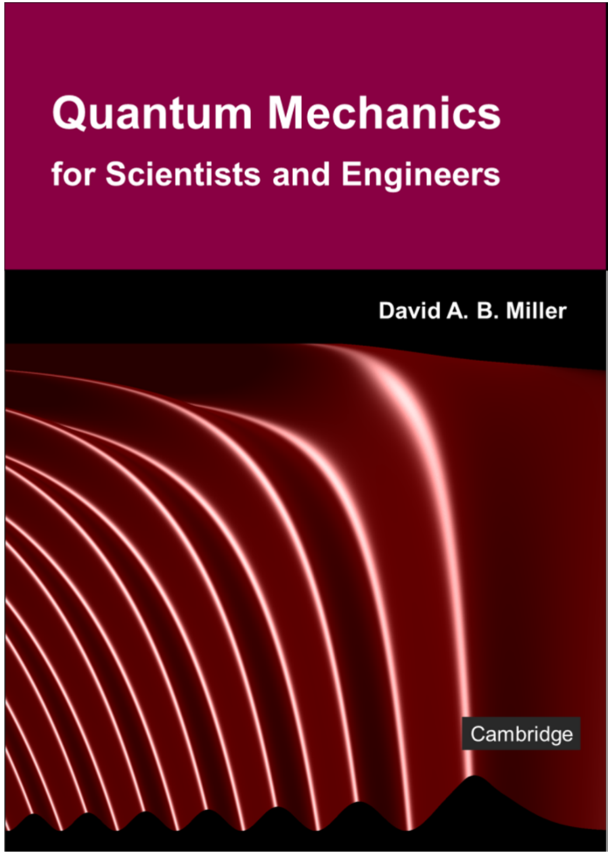


11.3 Density of states

Slides: Video 11.3.3 Quantum wells

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.8 through end of "Formal separation …" subsection



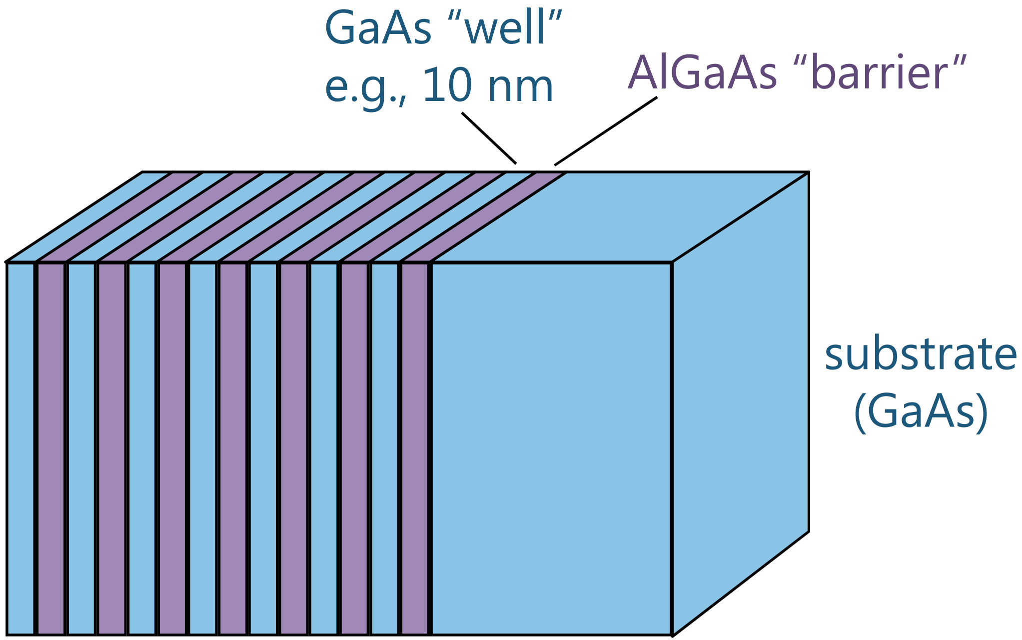
Density of states

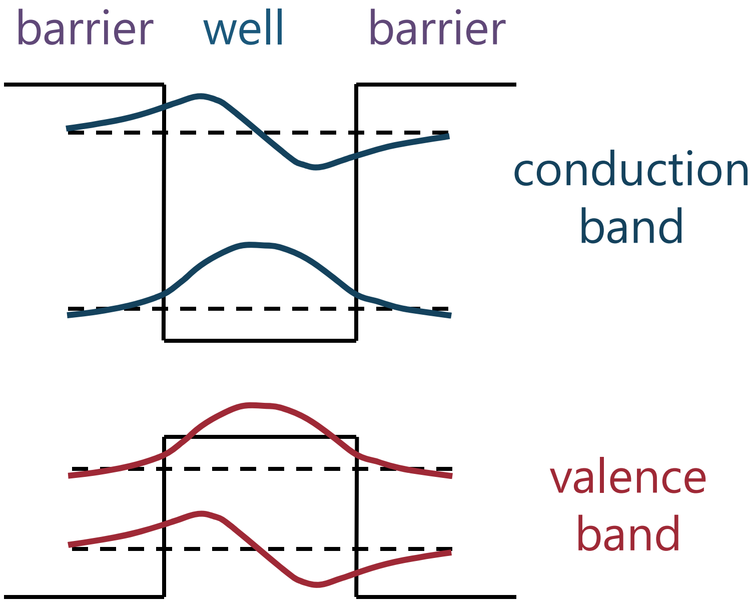
Quantum wells

Quantum mechanics for scientists and engineers

David Miller

Quantum well structures





Electrons and holes both see “lower” energies in the well

“particle in a box” quantum confinement in one direction

Separation of the quantum well problem

The eigenstates of a particle (electron or hole)

- will be “particle in a box” states in the z direction
  - with envelope wavefunction $\psi_n(z)$
    - and unconstrained "free" plane-wave motion
      - in the two directions in the plane of the quantum well layer
        - with wavevector $\mathbf{k}_{xy}$

We see this by formal separation of the problem

Separation of the quantum well problem

The Schrödinger equation for the envelope function is

$$-\frac{\hbar^2}{2m_{eff}}\nabla^2\psi(\mathbf{r})+V(z)\psi(\mathbf{r})=E\psi(\mathbf{r})$$

where $V(z)$ is only a function of z

For quantum-confined structures such as

quantum wires or

quantum boxes or "dots"

the potential would be a function of

two directions or

three directions, respectively

Separation of the quantum well problem

We formally rewrite the envelope function equation as

$$-\frac{\hbar^2}{2m_{eff}}\nabla_{xy}^2\psi(\mathbf{r})-\frac{\hbar^2}{2m_{eff}}\frac{\partial^2}{\partial z^2}\psi(\mathbf{r})+V(z)\psi(\mathbf{r})=E\psi(\mathbf{r})$$

where $\nabla_{xy}^2 \equiv \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}$

We postulate a separation

$$\psi(\mathbf{r})=\psi_n(z)\psi_{xy}(\mathbf{r}_{xy})$$

where $\mathbf{r}_{xy} \equiv x\mathbf{i}+y\mathbf{j}$ is the electron position in the quantum well plane

Separation of the quantum well problem

Substituting this form $\psi(\mathbf{r}) = \psi_n(z)\psi_{xy}(\mathbf{r}_{xy})$

into the envelope function Schrödinger equation

$$-\frac{\hbar^2}{2m_{eff}}\nabla_{xy}^2\psi(\mathbf{r}) - \frac{\hbar^2}{2m_{eff}}\frac{\partial^2}{\partial z^2}\psi(\mathbf{r}) + V(z)\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

and dividing by this form throughout, leads to

$$-\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_{xy}(\mathbf{r}_{xy})}\nabla_{xy}^2\psi_{xy}(\mathbf{r}_{xy}) - \frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_n(z)}\frac{\partial^2}{\partial z^2}\psi_n(z) + V(z) = E$$

Separation of the quantum well problem

We can formally separate this equation

$$-\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_{xy}\left(\mathbf{r}_{xy}\right)}\nabla_{xy}^2\psi_{xy}\left(\mathbf{r}_{xy}\right)-\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_n\left(z\right)}\frac{\partial^2}{\partial z^2}\psi_n\left(z\right)+V\left(z\right)=E$$

as

$$-\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_{xy}\left(\mathbf{r}_{xy}\right)}\nabla_{xy}^2\psi_{xy}\left(\mathbf{r}_{xy}\right)=E+\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_n\left(z\right)}\frac{\partial^2}{\partial z^2}\psi_n\left(z\right)-V\left(z\right)=E_{xy}$$

with a separation constant we have chosen as E_{xy}

Separation of the quantum well problem

The left part of

$$-\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_{xy}\left(\mathbf{r}_{xy}\right)}\nabla^2_{xy}\psi_{xy}\left(\mathbf{r}_{xy}\right)=E+\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_n\left(z\right)}\frac{\partial^2}{\partial z^2}\psi_n\left(z\right)-V\left(z\right)=E_{xy}$$

gives $-\frac{\hbar^2}{2m_{eff}}\nabla^2_{xy}\psi_{xy}\left(\mathbf{r}_{xy}\right)=E_{xy}\psi_{xy}\left(\mathbf{r}_{xy}\right)$

which is simply solved for the “in-plane” motion

$$\psi_{xy}\left(\mathbf{r}_{xy}\right)\propto\exp\left(i\mathbf{k}_{xy}.\mathbf{r}_{xy}\right)$$

with $E_{xy}=\frac{\hbar^2k_{xy}^2}{2m_{eff}}$

Separation of the quantum well problem

The right part of

$$-\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_{xy}\left(\mathbf{r}_{xy}\right)}\nabla^2_{xy}\psi_{xy}\left(\mathbf{r}_{xy}\right)=E+\frac{\hbar^2}{2m_{eff}}\frac{1}{\psi_n\left(z\right)}\frac{\partial^2}{\partial z^2}\psi_n\left(z\right)-V\left(z\right)=E_{xy}$$

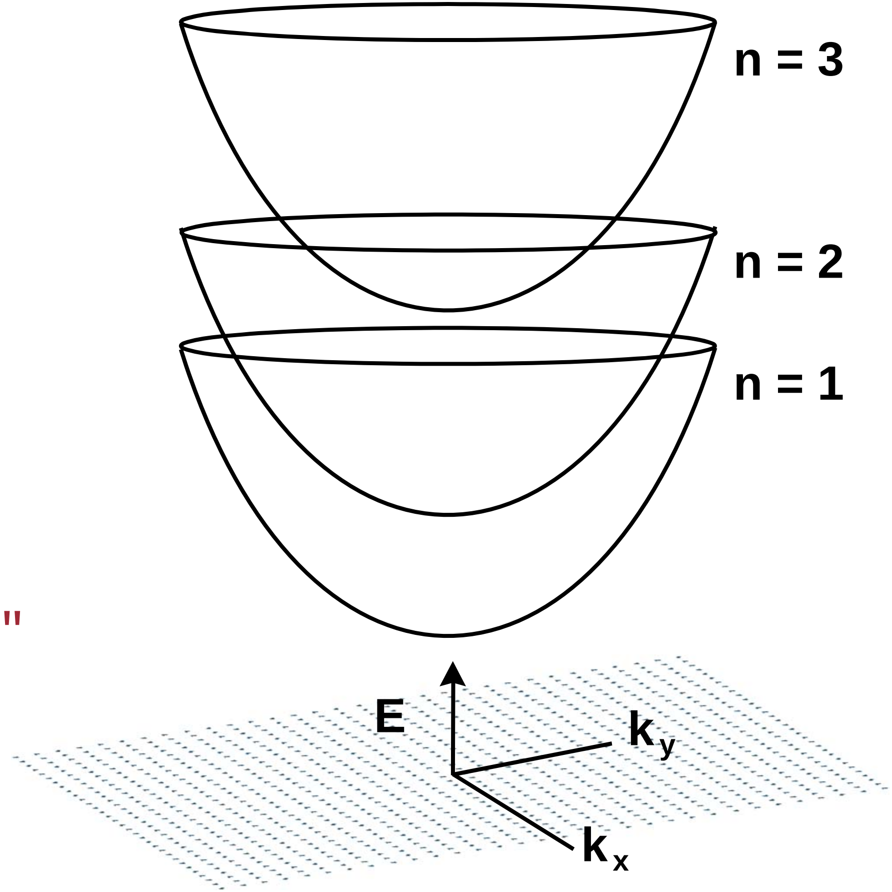
with the choice $E_n = E - E_{xy}$

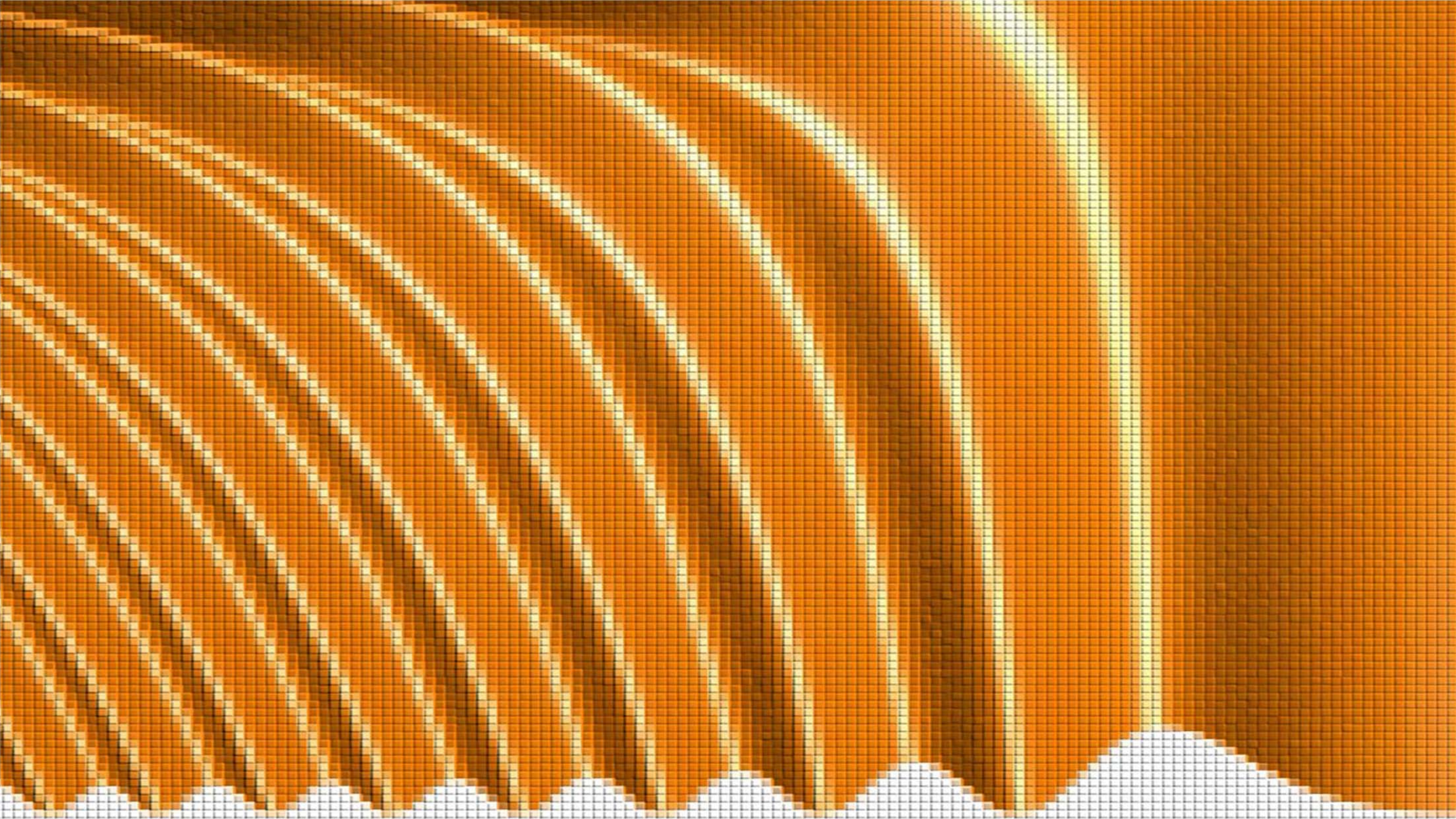
gives $-\frac{\hbar^2}{2m_{eff}}\frac{d^2}{dz^2}\psi_n\left(z\right)+V\left(z\right)\psi_n\left(z\right)=E_n\psi_n\left(z\right)$

which is a simple “particle in a box” equation for a particle of effective mass m_{eff} in an effective potential $V(z)$

Subbands

- The total allowed energies are
  - the energies E_n for the particle in a box energies
    - plus the additional energy associated with the in-plane motion
- Instead of discrete energy levels
  - we have so called "subbands"
- Note that the bottom of each subband has the energy E_n



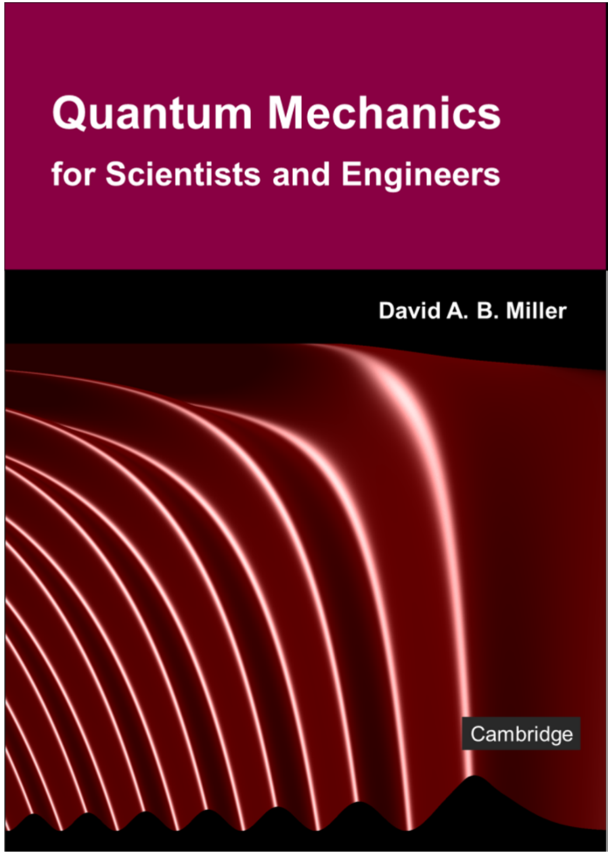


11.3 Density of states

Slides: Video 11.3.5 Density of states in quantum wells

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.8 from "Quantum well …" subsection



Density of states

Density of states in quantum wells

Quantum mechanics for scientists and engineers

David Miller

Quantum well density of states

Just as for the bulk case

- we formally impose periodic boundary conditions in the x and y directions

This gives us allowed values of the wavevector

- in the x direction, k_x, spaced by $2\pi/L_x$
- in the y direction, k_y, spaced by $2\pi/L_y$

Each $\mathbf{k}_{xy}$ state occupies a $\mathbf{k}_{xy}$ space “area” of $(2\pi)^2/A_{qw}$

- where $A_{qw} = L_x L_y$
  - and there is one allowed value of k_{xy} for each unit cell in the x-y plane of the quantum well

Quantum well density of states

Since each $\mathbf{k}_{xy}$ state occupies a $\mathbf{k}_{xy}$ space "area" of $(2\pi)^2/A_{qw}$

- the number of states in a small area $d^2\mathbf{k}_{xy}$ of $\mathbf{k}_{xy}$ space is $(A_{qw}/(2\pi)^2)\, d^2\mathbf{k}_{xy}$

Hence we can usefully define a ($\mathbf{k}_{xy}$ space) density of states per unit (real) area, $g_{2D}(\mathbf{k}_{xy})$, given by

$$g_{2D}\left(\mathbf{k}_{xy}\right) = \frac{1}{\left(2\pi\right)^2}$$

Quantum well density of states

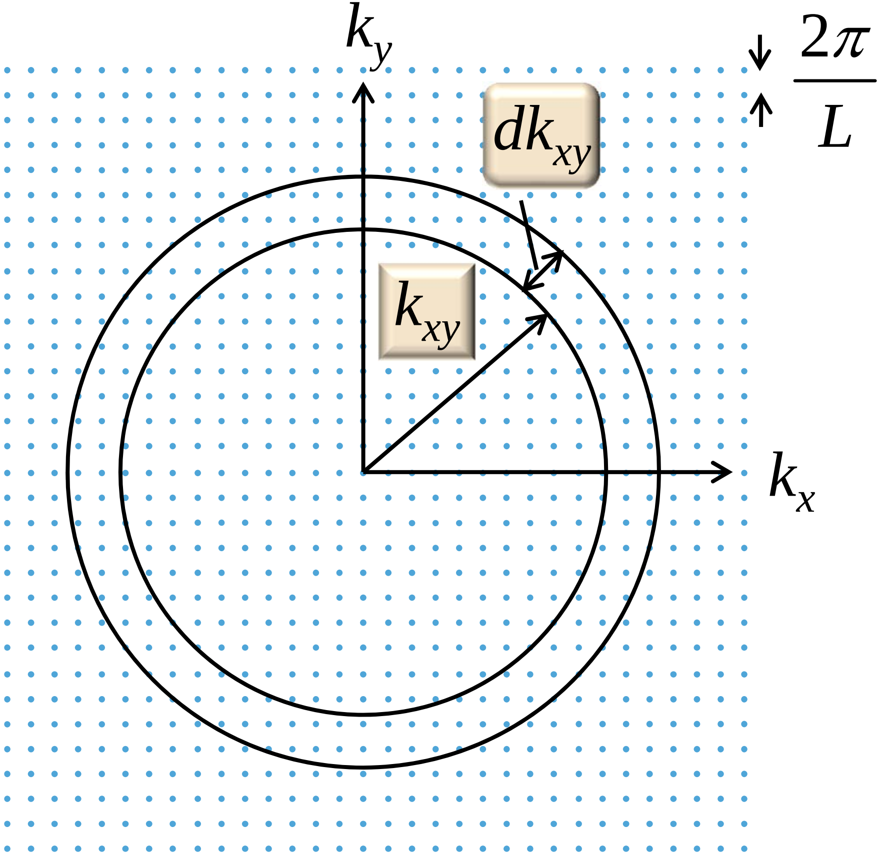
The number of $\mathbf{k}$ states between energies E_{xy} and $E_{xy} + dE_{xy}$, i.e., $g_{2D}(E_{xy})dE_{xy}$

is the number of states in $\mathbf{k}_{xy}$-space in the annular ring

of area $2\pi k_{xy} dk_{xy}$

between k_{xy} and $k_{xy} + dk_{xy}$

where $dk_{xy} = \left(\frac{dk_{xy}}{dE_{xy}}\right) dE_{xy}$



Quantum well density of states

Using the assumed parabolic relation between E_{xy} and k_{xy}, we have therefore,

now multiplying by 2 to include the different spins

$$g_{2D}\left(E_{xy}\right)dE_{xy} = 2g_{2D}\left(\mathbf{k}_{xy}\right)2\pi k_{xy}\frac{dk_{xy}}{dE_{xy}}dE_{xy}$$

$$= \frac{2}{\left(2\pi\right)^2}2\pi\sqrt{\frac{2m_{eff}}{\hbar^2}}\sqrt{E_{xy}}\frac{1}{2}\sqrt{\frac{2m_{eff}}{\hbar^2}}\frac{1}{\sqrt{E_{xy}}}dE_{xy}$$

i.e., $g_{2D}\left(E_{xy}\right) = \dfrac{m_{eff}}{\pi\hbar^2}$

Quantum well density of states

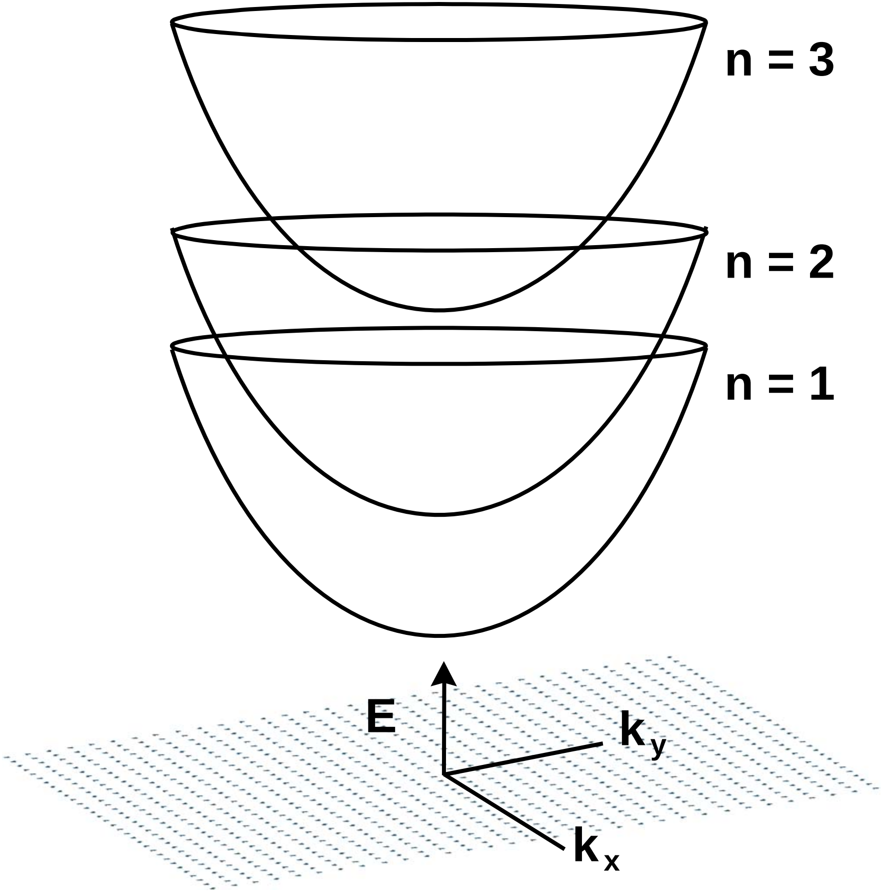
This relation $g_{2D}\left(E_{xy}\right)=\frac{m_{eff}}{\pi\hbar^2}$

means that the density of states in energy within a given subband

is a constant, independent of energy

within that subband

for all $E_{xy} > 0$



Quantum well density of states

It is therefore a "step" density of states

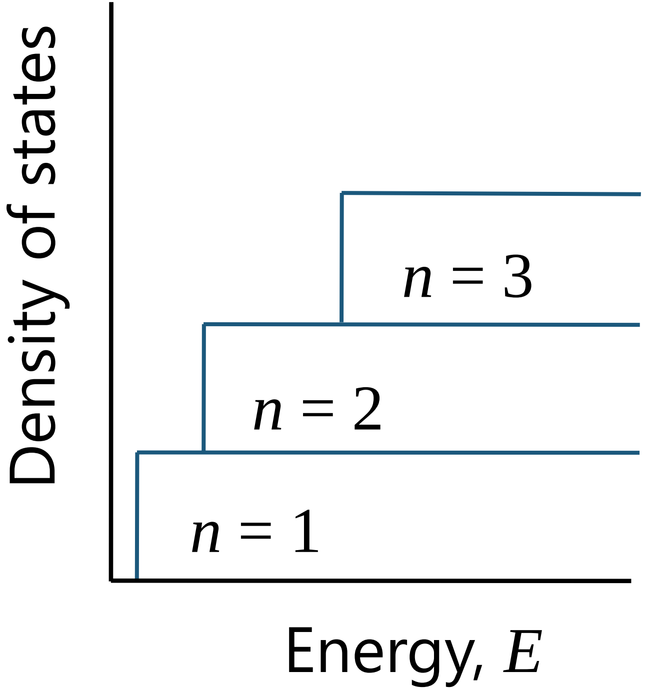
starting at $E_{xy} = 0$

i.e., starting at $E = E_n$

Hence, the total density of states as a function of the energy E

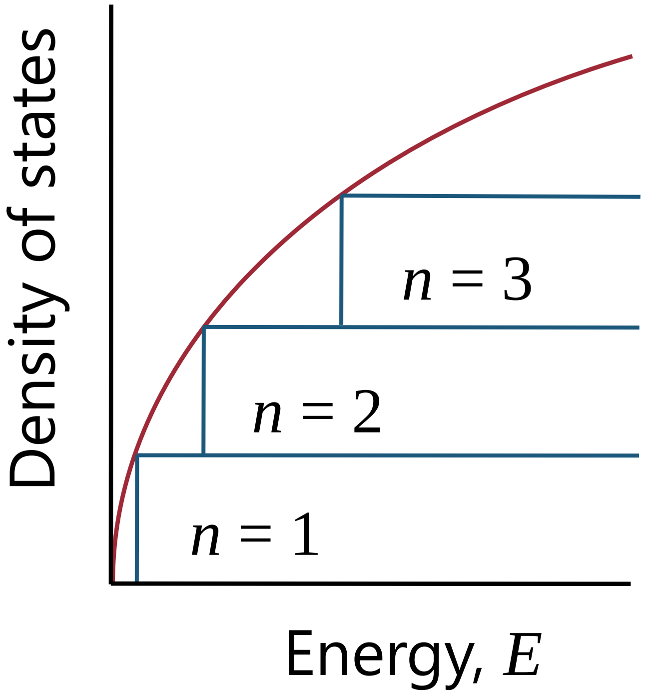
rises as a series of steps

with a new step starting as we reach each E_n



Quantum well density of states

- Plotting the quantum well density of states (per unit volume)
  - for a quantum well with infinitely high walls
    - it "touches" the bulk density of states (per unit volume)
      - at the edge of the first step



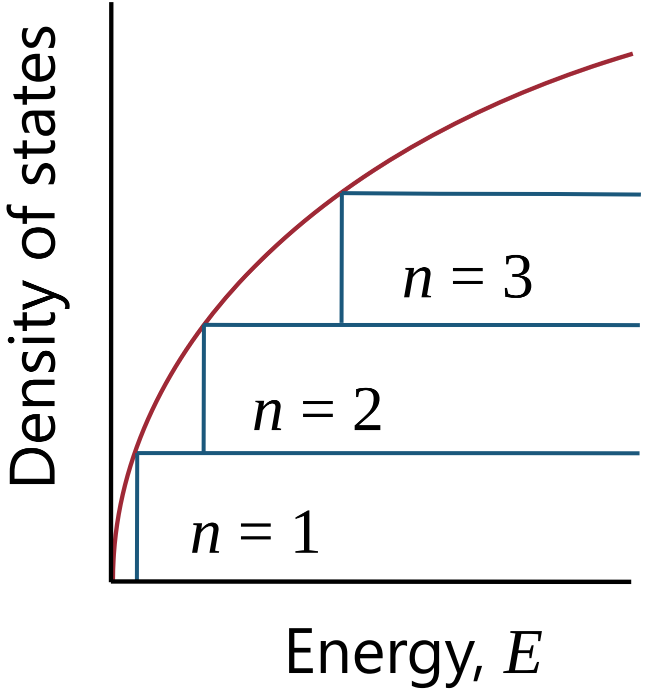
Quantum well density of states

Furthermore, since the steps are spaced quadratically in energy

and the bulk density of states is a "parabola on its side"

the quantum well (volume) density of states touches the bulk (volume) density of states

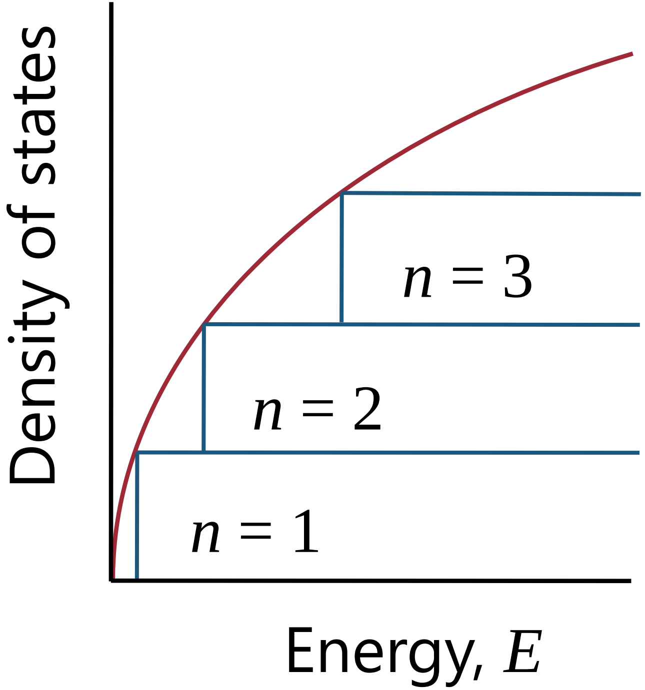
at the corner of each step

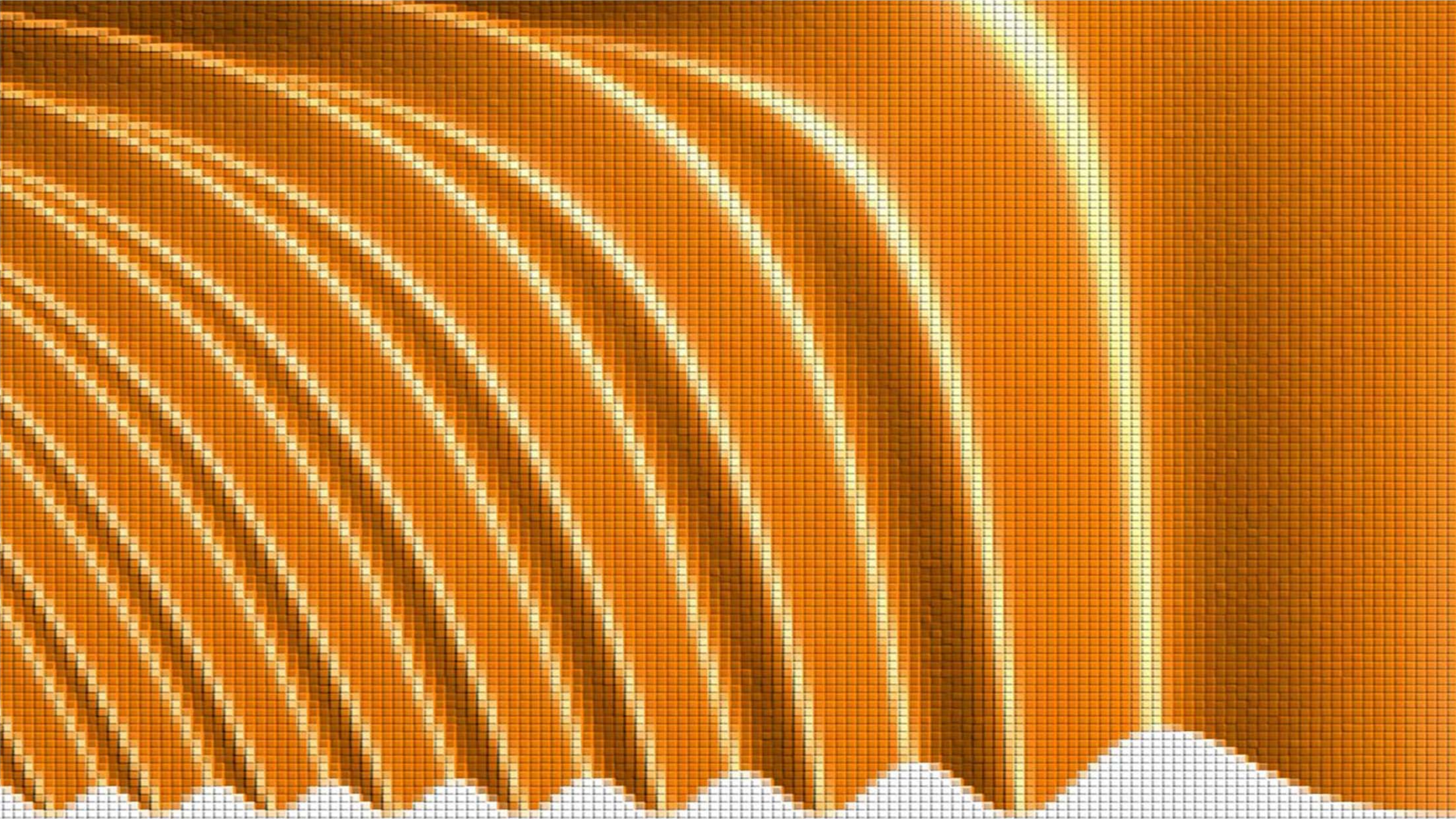


Quantum well density of states

If we started to increase the thickness of the quantum well

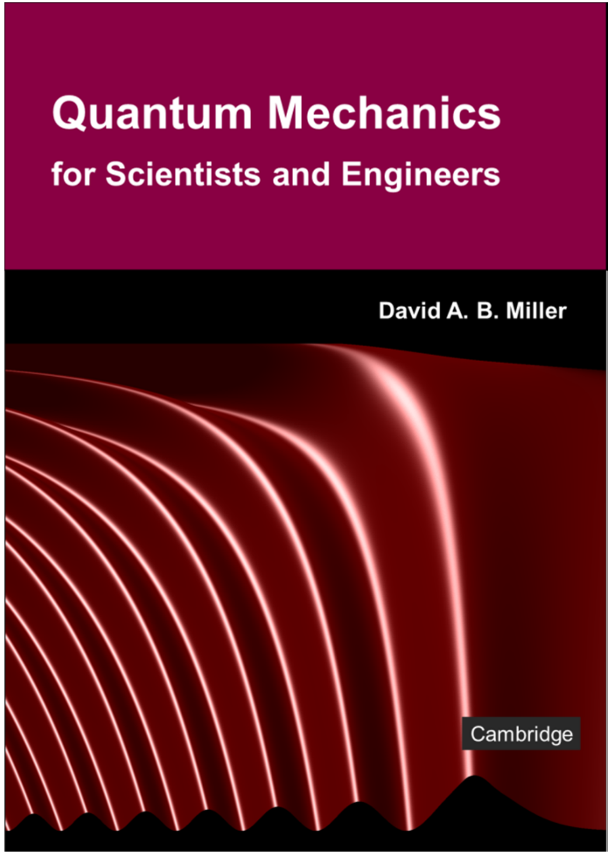
- the steps would get closer and closer together
  - but their corners would still touch the bulk curve
    - so that, as the quantum well became very thick
      - its density of states would tend to that of the bulk material





12.1 Optical absorption in semiconductors

Slides: Video 12.1.1 Introduction to optical absorption in semiconductors



Optical absorption in semiconductors

Quantum mechanics for scientists and engineers

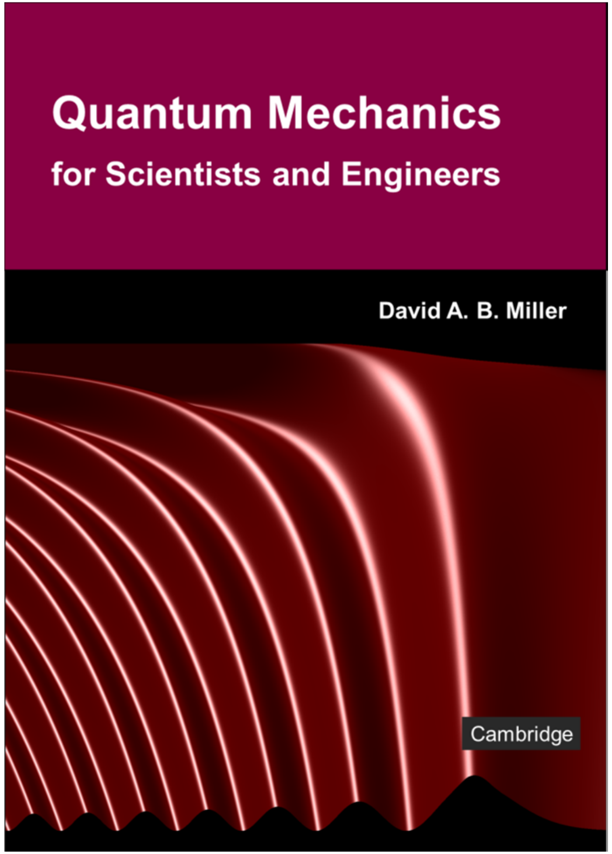
David Miller

12.1 Optical absorption in semiconductors

Slides: Video 12.1.2 Perturbing Hamiltonian

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.10 through "Form of the perturbing …" (with Appendix E for optional background)



Optical absorption in semiconductors

Perturbing Hamiltonian

Quantum mechanics for scientists and engineers

David Miller

Direct optical absorption

The transition rate for absorption

from an initial electron state $|\psi_i\rangle$

with energy E_i

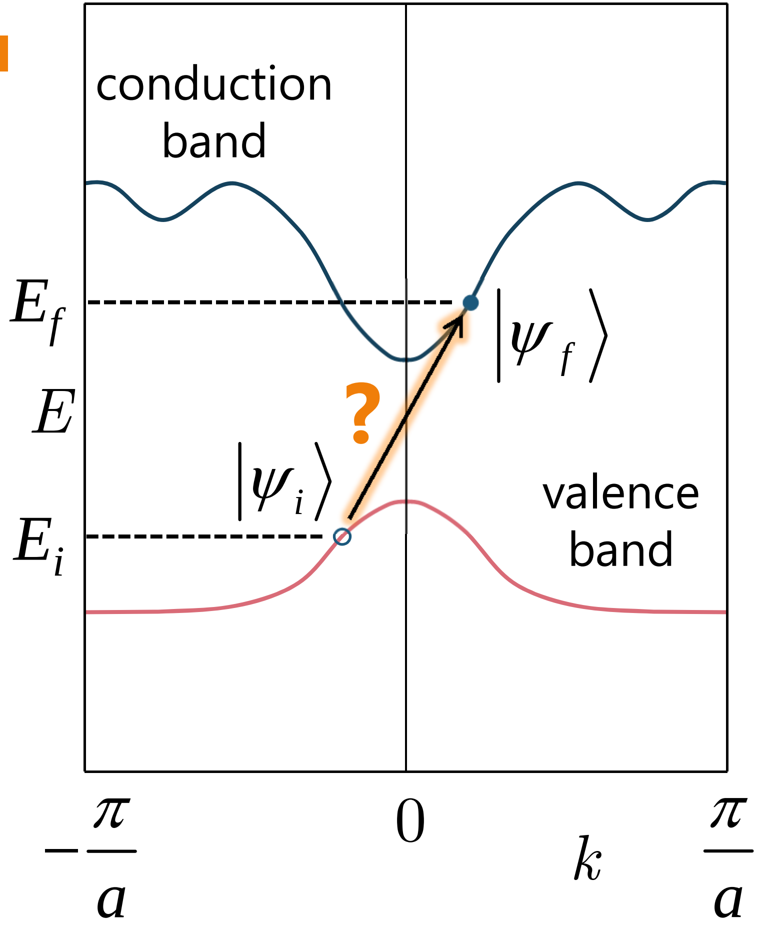
to a final state $|\psi_f\rangle$ with energy E_f

with an oscillating perturbation of

angular frequency ω

is, from Fermi's Golden Rule

$$w_{abs} = \frac{2\pi}{\hbar}\left|\left\langle \psi_f \left| \hat{H}_{po} \right| \psi_i \right\rangle\right|^2 \delta\left(E_f - E_i - \hbar\omega\right)$$



Direct optical absorption

In Fermi's Golden Rule $w_{abs} = \frac{2\pi}{\hbar}\left|\left\langle \psi_f \middle| \hat{H}_{po} \middle| \psi_i \right\rangle\right|^2 \delta\left(E_f - E_i - \hbar\omega\right)$

$\hat{H}_{po}$ is the amplitude of a perturbation oscillating in time at (angular) frequency ω, such as is defined in

$$\hat{H}_p(t) = e\mathrm{E}(t)z = \hat{H}_{po}\left[\exp(-i\omega t) + \exp(i\omega t)\right]$$

where E is the electric field in the z direction

We are also now interested in the spatial variation of $\hat{H}_{po}$ through the spatial dependence of the electromagnetic wave amplitude, as explicitly in

$$\hat{H}_p(\mathbf{r},t) = \hat{H}_{po}(\mathbf{r})\left[\exp(-i\omega t) + \exp(i\omega t)\right]$$

Direct optical absorption

The matrix element in Fermi’s Golden Rule, $\left\langle \psi_f \middle| \hat{H}_{po} \middle| \psi_i \right\rangle$

can now be written explicitly as

$$\left\langle \psi_f \middle| \hat{H}_{po} \middle| \psi_i \right\rangle = \int \psi_f^*(\mathbf{r}) \hat{H}_{po}(\mathbf{r}) \psi_i(\mathbf{r}) d^3\mathbf{r}$$

where $\psi_i(\mathbf{r})$ and $\psi_f(\mathbf{r})$ are, respectively,

the wave functions of the initial and final states

and we explicitly allow $\hat{H}_{po}(\mathbf{r})$ to vary in space

because the electromagnetic field at

(angular) frequency ω will vary in space

Perturbing Hamiltonian for electromagnetic field

Though we could continue with the “electric dipole” version of the perturbing Hamiltonian $\hat{H}_p(\mathbf{r},t) = e\mathrm{E}(\mathbf{r},t)z$

we can connect to other semiconductor phenomena

if we switch to using the magnetic vector potential $\mathbf{A}$ instead of the electric field, and write

$$\hat{H}_p(\mathbf{r},t) \cong -\frac{e}{m_o}\mathbf{A}\cdot\hat{\mathbf{p}}$$

where m_o is the usual free electron mass and

$\hat{\mathbf{p}} = -i\hbar\nabla$ is the momentum operator

(Note: both the electric and magnetic fields can be deduced from the vector potential $\mathbf{A}$)

Perturbing Hamiltonian for electromagnetic field

Now we write the vector potential of a wave at frequency ω

$$\mathbf{A} = \mathbf{e}\left\{\frac{A_0}{2}\exp\left[i\left(\mathbf{k}_{op}\cdot\mathbf{r}-\omega t\right)\right]+\frac{A_0}{2}\exp\left[-i\left(\mathbf{k}_{op}\cdot\mathbf{r}-\omega t\right)\right]\right\}$$

Here $\mathbf{k}_{op}$ is the wave vector of the optical field inside the material, and

we take the field to be linearly polarized

with its electric vector in the direction of the unit vector $\mathbf{e}$

Perturbing Hamiltonian for electromagnetic field

Now in $\mathbf{A} = \mathbf{e}\left\{ \frac{A_0}{2} \exp\left[i\left(\mathbf{k}_{op} \cdot \mathbf{r} - \omega t \right) \right] + \frac{A_0}{2} \exp\left[-i\left(\mathbf{k}_{op} \cdot \mathbf{r} - \omega t \right) \right] \right\}$

we are going to keep only the term in $\exp\left(-i\omega t\right)$

because we know from our previous discussion of Fermi's Golden Rule that this term

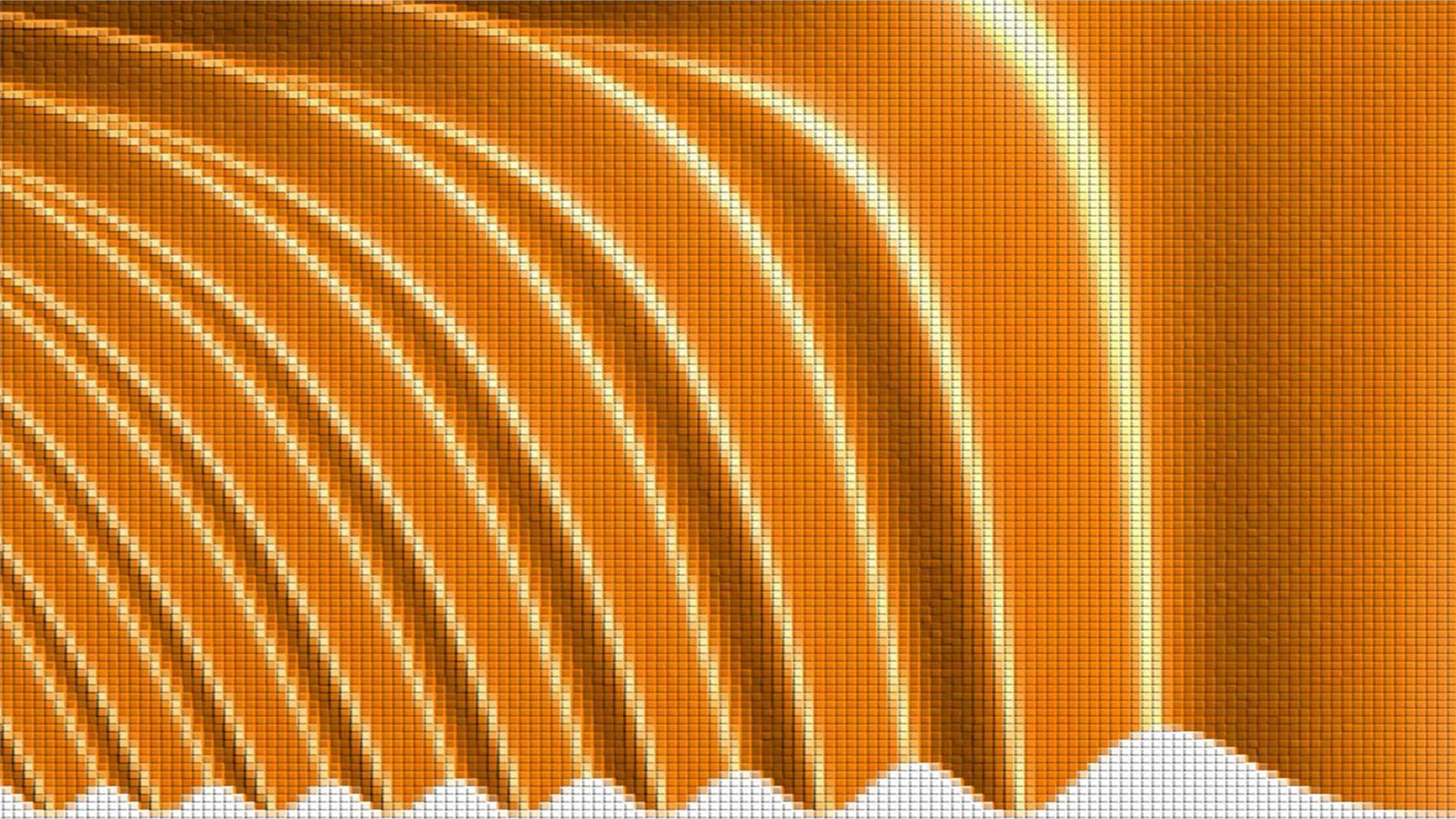
corresponds to absorption rather than emission

So the spatial part of our perturbing Hamiltonian is

$$\hat{H}_{po}\left(\mathbf{r}\right) = -\left(e / 2m_o\right) A_0 \exp\left(i\mathbf{k}_{op} \cdot \mathbf{r}\right) \mathbf{e} \cdot \hat{\mathbf{p}}$$

and the total perturbing Hamiltonian is

$$\hat{H}_p(\mathbf{r}, t) = \hat{H}_{po}\left(\mathbf{r}\right) \exp(-i\omega t)$$

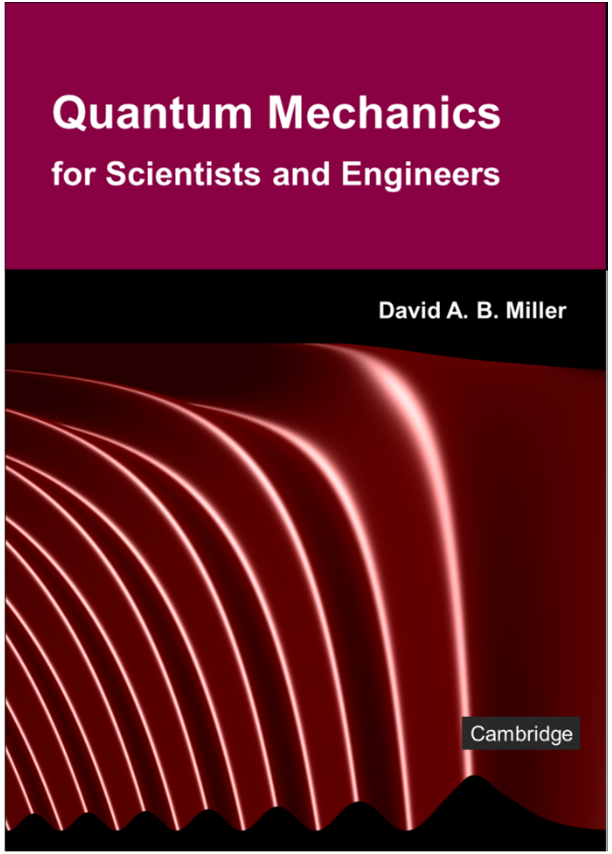


12.1 Optical absorption in semiconductors

Slides: Video 12.1.4 Direct transitions

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.10 subsection "Direct valence …" through paragraph after Eq. 8.98



Optical absorption in semiconductors

Direct transitions

Quantum mechanics for scientists and engineers

David Miller

Direct optical absorption

To proceed, we need to know

the initial and final wavefunctions

We are most interested in the transitions between

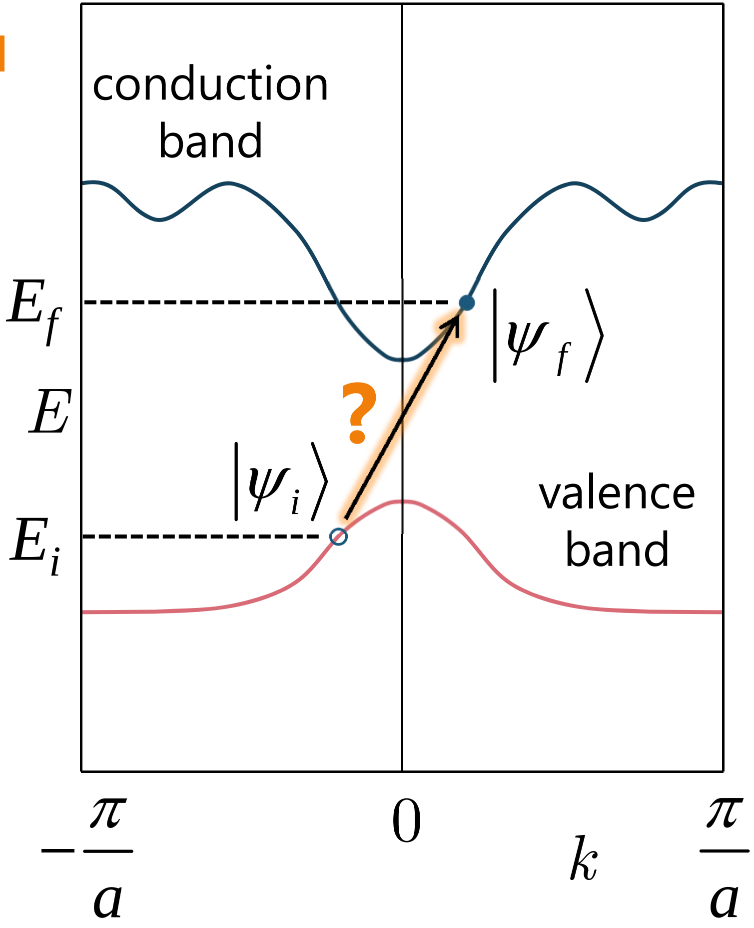
an initial valence band state $|\psi_i\rangle$ and

a final conduction band state $|\psi_f\rangle$

We presume that we can write them as “single-electron” Bloch states

and to calculate matrix elements

we must normalize them



Normalizing Bloch functions

With normalization constants B_i and B_f, we have

$$\psi_i(\mathbf{r}) = B_i u_v(\mathbf{r}) \exp(i\mathbf{k}_v \cdot \mathbf{r})$$

and

$$\psi_f(\mathbf{r}) = B_f u_c(\mathbf{r}) \exp(i\mathbf{k}_c \cdot \mathbf{r})$$

Here, and below, for simplicity

we presume u_v and u_c do not depend on $\mathbf{k}$

a good approximation for an "allowed" process

so we omit the "$\mathbf{k}$" subscripts on u_v and u_c

We do, however, now explicitly allow the conduction (u_v) and valence (u_c) unit cell functions to be different

Normalizing Bloch functions

We choose $u_v(\mathbf{r})$ and $u_c(\mathbf{r})$ to be normalized over a unit cell

$$\int_{unit\ cell} u_v^*(\mathbf{r})u_v(\mathbf{r})d^3\mathbf{r} = 1$$ and similarly for $u_c(\mathbf{r})$

Hence, normalizing $\psi_i(\mathbf{r})$ and $\psi_f(\mathbf{r})$, we have, e.g., for $\psi_i(\mathbf{r})$

$$\int_V \psi_i^*(\mathbf{r})\psi_i(\mathbf{r})d^3\mathbf{r} = 1 = B_i^2 \int_V u_v^*(\mathbf{r})\exp(-i\mathbf{k}_v\cdot\mathbf{r})u_v(\mathbf{r})\exp(i\mathbf{k}_v\cdot\mathbf{r})d^3\mathbf{r}$$

$$= B_i^2 \int_V u_v^*(\mathbf{r})u_v(\mathbf{r})d^3\mathbf{r} = B_i^2 N \int_{unit\,cell} u_v^*(\mathbf{r})u_v(\mathbf{r})d^3\mathbf{r} = B_i^2 N$$

where V is the volume of the crystal

where N is the number of unit cells and

Hence we have $B_i = B_f = 1/\sqrt{N}$

Matrix element for Fermi's Golden Rule

With $\hat{H}_{po}(\mathbf{r}) = -(e/2m_o)A_o \exp\left(i\mathbf{k}_{op}\cdot\mathbf{r}\right)\mathbf{e}\cdot\hat{\mathbf{p}}$

$\psi_i(\mathbf{r}) = \frac{1}{\sqrt{N}}u_v(\mathbf{r})\exp\left(i\mathbf{k}_v\cdot\mathbf{r}\right)$ and $\psi_f(\mathbf{r}) = \frac{1}{\sqrt{N}}u_c(\mathbf{r})\exp\left(i\mathbf{k}_c\cdot\mathbf{r}\right)$

the matrix element for Fermi's Golden Rule is

$$\left\langle \psi_f \middle| \hat{H}_{po}(\mathbf{r}) \middle| \psi_i \right\rangle =$$

$$-\frac{eA_o}{2m_oN}\int_V \left[u_c^*(\mathbf{r})\exp\left(-i\mathbf{k}_c\cdot\mathbf{r}\right)\right]\exp\left(i\mathbf{k}_{op}\cdot\mathbf{r}\right)\mathbf{e}\cdot\hat{\mathbf{p}}\left[u_v(\mathbf{r})\exp\left(i\mathbf{k}_v\cdot\mathbf{r}\right)\right]d^3\mathbf{r}$$

Matrix element for Fermi's Golden Rule

We are interested in transitions involving states near the center of the Brillouin zone, so

$|\mathbf{k}_v|$ and $|\mathbf{k}_c|$ are both $\ll \pi / a$

Though strictly $\hat{\mathbf{p}} = -i\hbar\nabla$ operates on all of $u_v(\mathbf{r})\exp(i\mathbf{k}_v \cdot \mathbf{r})$

with these small values of $\mathbf{k}_v$

$\exp(i\mathbf{k}_v \cdot \mathbf{r})$ changes very slowly compared to the rate of change of $u_v(\mathbf{r})$

so, at least as a first approximation, we take

$$\hat{\mathbf{p}}\left[u_v(\mathbf{r})\exp(i\mathbf{k}_v \cdot \mathbf{r})\right] = \exp(i\mathbf{k}_v \cdot \mathbf{r})\hat{\mathbf{p}}u_v(\mathbf{r}) + u_v(\mathbf{r})\hat{\mathbf{p}}\exp(i\mathbf{k}_v \cdot \mathbf{r})$$

$$\simeq \exp(i\mathbf{k}_v \cdot \mathbf{r})\hat{\mathbf{p}}u_v(\mathbf{r})$$

Matrix element for Fermi's Golden Rule

For definiteness, we choose the polarization direction (i.e., unit vector $\mathbf{e}$) in the x direction

x will be one of the directions perpendicular to the propagation of the electromagnetic wave

With this choice and our approximations so far

$$\left\langle \psi_f \middle| \hat{H}_{po}(\mathbf{r}) \middle| \psi_i \right\rangle =$$

$$-\frac{eA_0}{2m_o N}\int_V \left[u_c^*(\mathbf{r})\exp(-i\mathbf{k}_c\cdot\mathbf{r})\right]\exp\left(i\mathbf{k}_{op}\cdot\mathbf{r}\right)\mathbf{e}\cdot\hat{\mathbf{p}}\left[u_v(\mathbf{r})\exp(i\mathbf{k}_v\cdot\mathbf{r})\right]d^3\mathbf{r}$$

$$\simeq -\frac{eA_0}{2m_o N}\int_V \exp\left[i\left(\mathbf{k}_v-\mathbf{k}_c+\mathbf{k}_{op}\right)\cdot\mathbf{r}\right]\left[u_c^*(\mathbf{r})\,\hat{p}_x u_v(\mathbf{r})\right]d^3\mathbf{r}$$

Matrix element for Fermi's Golden Rule

The optical wave vector $\mathbf{k}_{op}$ corresponds to wavelengths

of 100's of nm or more (even inside the material)

A typical crystal unit cell size is, e.g., $a \sim 0.5\ \mathrm{nm}$

so $\exp\left(i\mathbf{k}_{op}\cdot\mathbf{r}\right)$ is slowly varying over a unit cell

Hence, since we have already restricted $\left|\mathbf{k}_v\right|$ and $\left|\mathbf{k}_c\right|$ to being small by assumption

then the entire factor

$$\exp\left[i\left(\mathbf{k}_v-\mathbf{k}_c+\mathbf{k}_{op}\right)\cdot\mathbf{r}\right]$$

varies slowly over the length scale, a, of a unit cell

Matrix element for Fermi's Golden Rule

As a result, we can approximately separate the integral

$$\left\langle \psi_f \middle| \hat{H}_{po}(\mathbf{r}) \middle| \psi_i \right\rangle \simeq -\frac{eA_o}{2m_o N}\int_V \exp\left[i\left(\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op}\right)\cdot\mathbf{r}\right]\left[u_c^*(\mathbf{r})\,\hat{p}_x u_v(\mathbf{r})\right]d^3\mathbf{r}$$

into a sum of integrals over a unit cell

treating the value of $\exp\left[i\left(\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op}\right)\cdot\mathbf{r}\right]$

as approximately constant within a unit cell, i.e.,

$$\left\langle \psi_f \middle| \hat{H}_{po}(\mathbf{r}) \middle| \psi_i \right\rangle = -\frac{eA_o}{2m_o N}\langle c|\hat{p}_x|v\rangle \sum_{m\,(i.e.,\, unit\, cells)} \exp\left[i\left(\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op}\right)\cdot\mathbf{R}_m\right]$$

where $\mathbf{R}_m$ is the position of (the center of) the *m*th unit cell

and $\langle c|\hat{p}_x|v\rangle \equiv \int_{\text{unit cell}} u_c^*(\mathbf{r})\,\hat{p}_x u_v(\mathbf{r})\,d^3\mathbf{r} \equiv p_{cv}$

Matrix element for Fermi's Golden Rule

The summation in

$$\left\langle \psi_f \middle| \hat{H}_{po}(\mathbf{r}) \middle| \psi_i \right\rangle = -\frac{eA_o}{2m_o N} \langle c | \hat{p}_x | v \rangle \sum_{m(i.e.,\, unit\, cells)} \exp\left[i \left(\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op} \right) \cdot \mathbf{R}_m \right]$$

will average approximately to zero unless $\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op} = 0$

because otherwise

the function $\exp\left[i \left(\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op} \right) \cdot \mathbf{r} \right]$ is oscillatory

(Note this condition $\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op} = 0$ can be seen to correspond to conservation of crystal momentum $\hbar \mathbf{k}$)

In this case, the sum becomes

$$\sum_m \exp\left[i \left(\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op} \right) \cdot \mathbf{R}_m \right] = \sum_m \exp(0) = N$$

Fermi's Golden Rule transition rate

Hence

$$\left\langle \psi_f \left| \hat{H}_{po}(\mathbf{r}) \right| \psi_i \right\rangle = -\frac{eA_o}{2m_o N} \left\langle c \left| \hat{p}_x \right| v \right\rangle \sum_{m\,(i.e.,\,unit\,cells)} \exp\left[i\left(\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op}\right) \cdot \mathbf{R}_m \right]$$

becomes $\left\langle \psi_f \left| \hat{H}_{po}(\mathbf{r}) \right| \psi_i \right\rangle = -\frac{eA_o}{2m_o} p_{cv}$

and Fermi's Golden Rule gives a transition rate

$$w_{abs} = \frac{2\pi}{\hbar} \left| \left\langle \psi_f \left| \hat{H}_{po} \right| \psi_i \right\rangle \right|^2 \delta\left(E_f - E_i - \hbar\omega\right)$$

$$= \frac{2\pi}{\hbar} \frac{e^2 A_o^2}{4m_o^2} \left| p_{cv} \right|^2 \delta\left(E_f - E_i - \hbar\omega\right)$$

Direct valence to conduction band absorption

Hence, for an incident optical wave of frequency ω, wavevector $\mathbf{k}_{op}$ and magnetic vector potential amplitude A_o

- for an initial state of energy E_i and Bloch wavevector $\mathbf{k}_v$ in the valence band
- and a final state of energy E_f and Bloch wavevector $\mathbf{k}_c$ in the conduction band
  - if and only if $\mathbf{k}_v - \mathbf{k}_c + \mathbf{k}_{op} = 0$, we have a "direct transition rate"

$$w_{abs} = \frac{2\pi}{\hbar}\frac{e^2 A_o^2}{4m_o^2}\left|p_{cv}\right|^2 \delta\left(E_f - E_i - \hbar\omega\right)$$

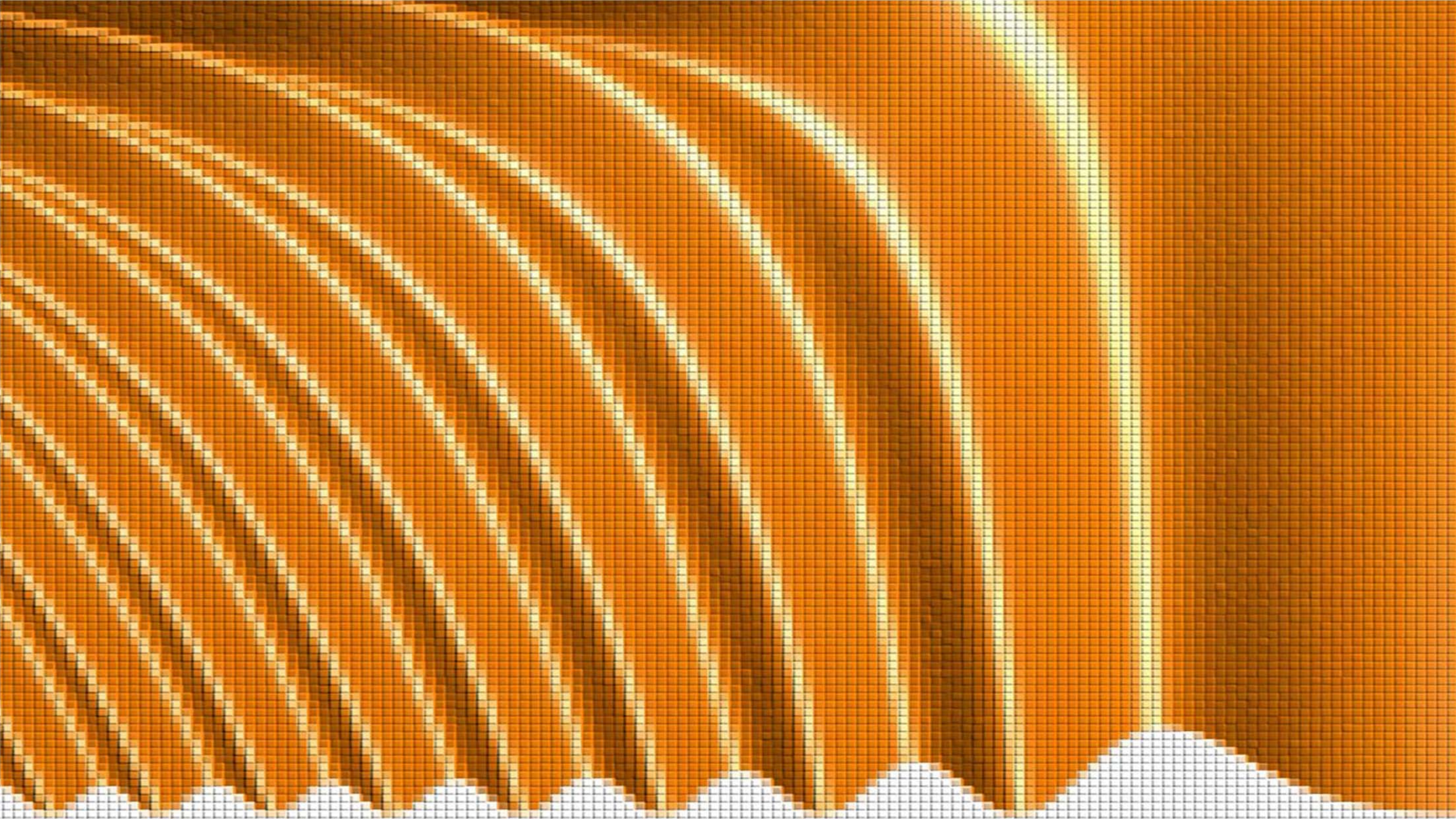
Direct valence to conduction band absorption

Note also that the transition rate

$$w_{abs} = \frac{2\pi}{\hbar}\frac{e^2 A_o^2}{4m_o^2}\left|p_{cv}\right|^2 \delta\left(E_f - E_i - \hbar\omega\right)$$

is proportional to

- a (squared) matrix element, $\left|p_{cv}\right|^2$
- the optical intensity (which is proportional to A_o^2)
  - and hence to the average arrival rate of photons in the semiconductor (per unit area)

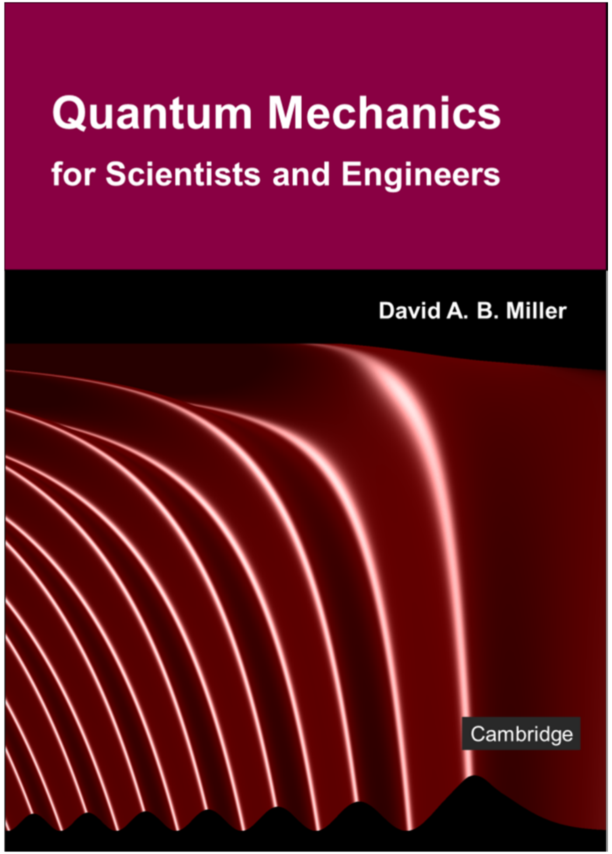


12.1 Optical absorption in semiconductors

Slides: Video 12.1.6 Transitioning from sums to integrals

Text reference: Quantum Mechanics for Scientists and Engineers

Section 5.3



Optical absorption in semiconductors

Transitioning from sums to integrals

Quantum mechanics for scientists and engineers

David Miller

Transitioning from sums to integrals

Suppose we have some states

indexed by an integer q

For each of these states

some quantity

such as energy

has a value f_q

Hence, summing all of these quantities would give a result

$$S = \sum_q f_q$$

Transitioning from sums to integrals

It could be that f_q can also equivalently be written as a function of some parameter u

such as momentum

that takes on some value u_q for each q

i.e., $f_q \equiv f(u_q)$

e.g., f_q might be the energy associated with momentum u_q

and q might be indexing all the allowed values of $\mathbf{k}$

So, we could write $S = \sum_q f(u_q)$

Transitioning from sums to integrals

Suppose now that the u_q and the f_q

are very closely spaced as we change q

and vary relatively smoothly with q

We suppose that this smooth change of u_q with q is such that

we can approximately represent u as some smooth and differentiable function of a continuous variable q

i.e., $u(q)$

that is the same as u_q for all integer values of q

Transitioning from sums to integrals

We can define the difference between two specific adjacent values as $u_{q+1} - u_q \equiv \delta u$

With $\delta q = (q+1) - q$

then, trivially
$$\delta u = \frac{\delta u}{\delta q} \delta q$$

Because of the presumed "smoothness" of $u(q)$

$$u_{q+1} - u_q \equiv \delta u = \frac{\delta u}{\delta q} \delta q \simeq \frac{du}{dq} \delta q = \frac{du}{dq}$$

where in the last step we use the fact
$$\delta q = (q+1) - q = 1$$

Transitioning from sums to integrals

So, considering some small range Δu

within which the separation δu between adjacent values of u

was approximately constant

the number of different terms in the sum that would lie within that range is

$$\Delta u / \delta u \simeq \Delta u / \left(du / dq \right)$$

Transitioning from sums to integrals

Equivalently, defining a density of states $g(u) = \frac{1}{(du/dq)}$

we could say, equivalently, that the number of terms

in the sum that lie within Δu is $\frac{\Delta u}{\delta u} \simeq \frac{\Delta u}{du/dq} = g(u)\Delta u$

Hence, instead of summing over q we could instead

consider a set of values of u each separated by Δu

and spanning the same range

and write the sum over all those values

i.e., $S = \sum_q f_q \equiv \sum_q f(u_q) \simeq \sum_u f(u)g(u)\Delta u$

Transitioning from sums to integrals

Finally we formally let Δu in $S \simeq \sum_u f(u)g(u)\Delta u$

become very small

so we can approximate the sum by an integral

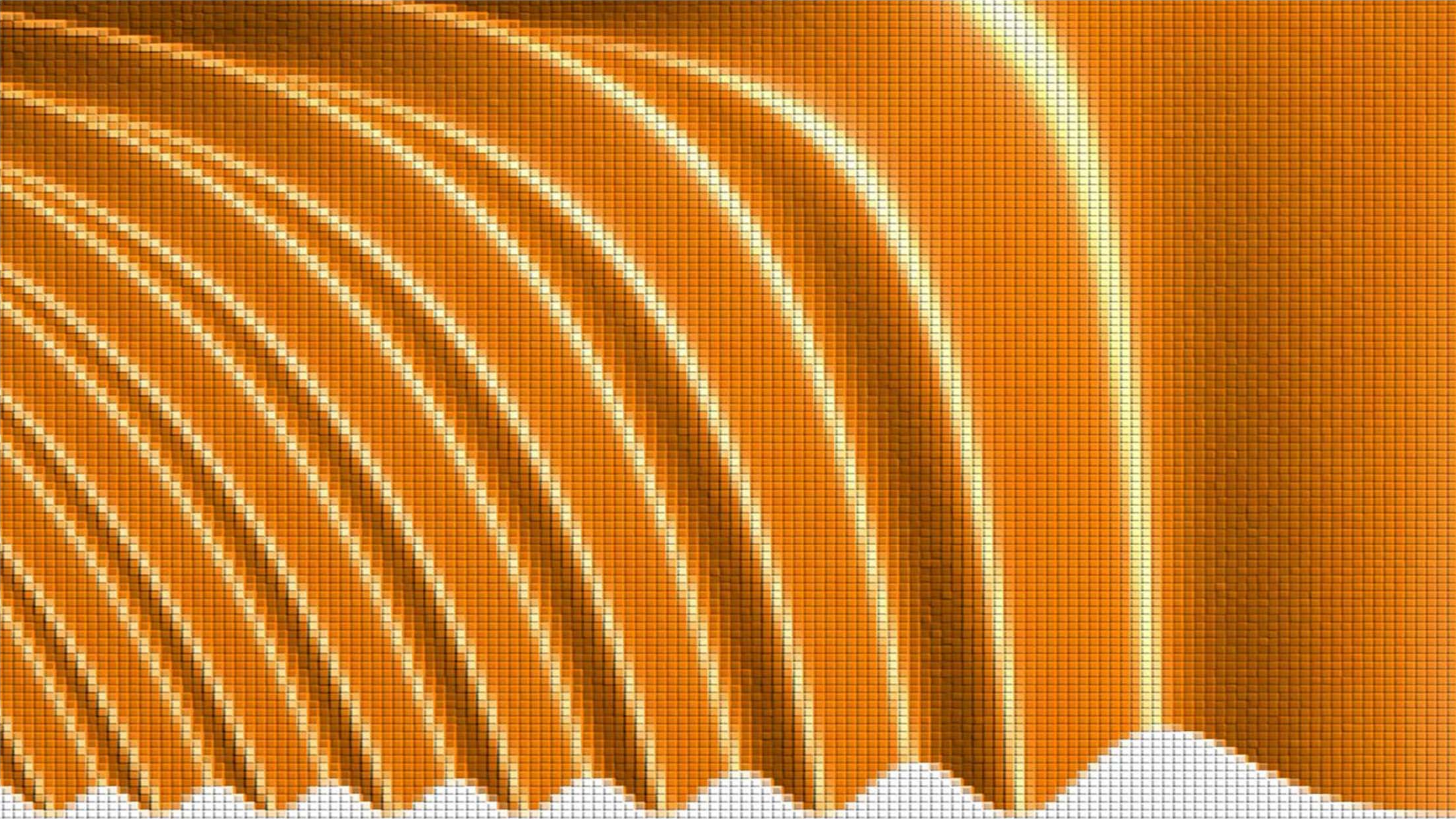
to obtain $S \simeq \int f(u)g(u)du$

So in going from a sum to an integral, we

insert the density of states in the integration variable into the integrand

i.e., $\boxed{\sum_q \ldots \rightarrow \int \ldots g(u)du}$

(The integral limits must correspond to the limits in the sum)

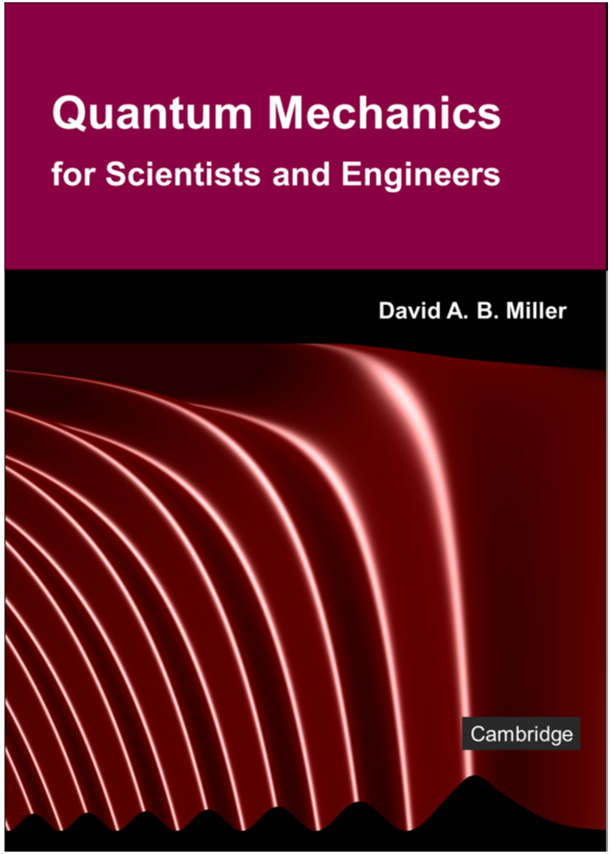


12.1 Optical absorption in semiconductors

Slides: Video 12.1.8 Total transition rate

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.10 subsection "Direct valence …" starting above Eq. 8.99 through Eq. 8.107



Optical absorption in semiconductors

Total transition rate

Quantum mechanics for scientists and engineers

David Miller

Total transition rate

$$w_{abs} = \frac{2\pi}{\hbar} \frac{e^2 A_o^2}{4m_o^2} |p_{cv}|^2 \delta\left(E_f - E_i - \hbar\omega\right)$$

is transition rate from

an initial valence band state with Bloch wavevector $\mathbf{k}_v$

to the conduction band state with wave vector

$\mathbf{k}_c = \mathbf{k}_v + \mathbf{k}_{op}$

To get the total transition rate

and hence the optical absorption rate of photons

we need to sum over all the possible initial and final states

Total transition rate

The total transition rate W_{TOT} is formally

$$W_{TOT} = \sum_i \sum_f w_{abs} = \frac{2\pi}{\hbar} \frac{e^2 A_o^2}{4m_o^2} \left| p_{cv} \right|^2 \sum_i \sum_f \delta\left(E_f - E_i - \hbar\omega \right)$$

where we presume $\left| p_{cv} \right|^2$ is ~ independent of $\mathbf{k}$

We have shown above that

for a given initial state with wave vector $\mathbf{k}_v$

the only final state possible is the conduction band state with $\mathbf{k}_c = \mathbf{k}_v + \mathbf{k}_{op}$

With this knowledge

we can drop the separate sum over final states

Total transition rate

Since $\left|\mathbf{k}_{op}\right|$ is generally a very small fraction of the size of the Brillouin zone

- we will now neglect it for simplicity, taking $\mathbf{k}_c = \mathbf{k}_v$

This negligible size of the optical wavevector means that

- the direct optical transitions are essentially "vertical" on the energy-momentum diagram

Hence, for a given initial state $\mathbf{k}_v$

- only one term remains in the sum over the final states
  - namely, the one with $\mathbf{k}_c = \mathbf{k}_v$

From now on, we drop the suffixes "v" and "c", using just $\mathbf{k}$

Total transition rate

Hence the total transition rate becomes

$$W_{TOT} = \frac{2\pi}{\hbar}\frac{e^2 A_o^2}{4m_o^2}|p_{cv}|^2 \sum_{\mathbf{k},\,\text{spin}} \delta\left[E_c(\mathbf{k}) - E_v(\mathbf{k}) - \hbar\omega\right]$$

now including a sum over the two possible spin states

Now we formally rewrite (considering unit volume)

$$\sum_{\mathbf{k},\,\text{spin}} = 2\sum_{\mathbf{k}} \approx 2\int_{\mathbf{k}} g(\mathbf{k})\, d^3\mathbf{k}$$

where $g(\mathbf{k})$ is the density of states in k space

We will next change variables in the integral

to the energy $E_c(\mathbf{k}) - E_v(\mathbf{k})$

Direct optical absorption

Assuming parabolic bands

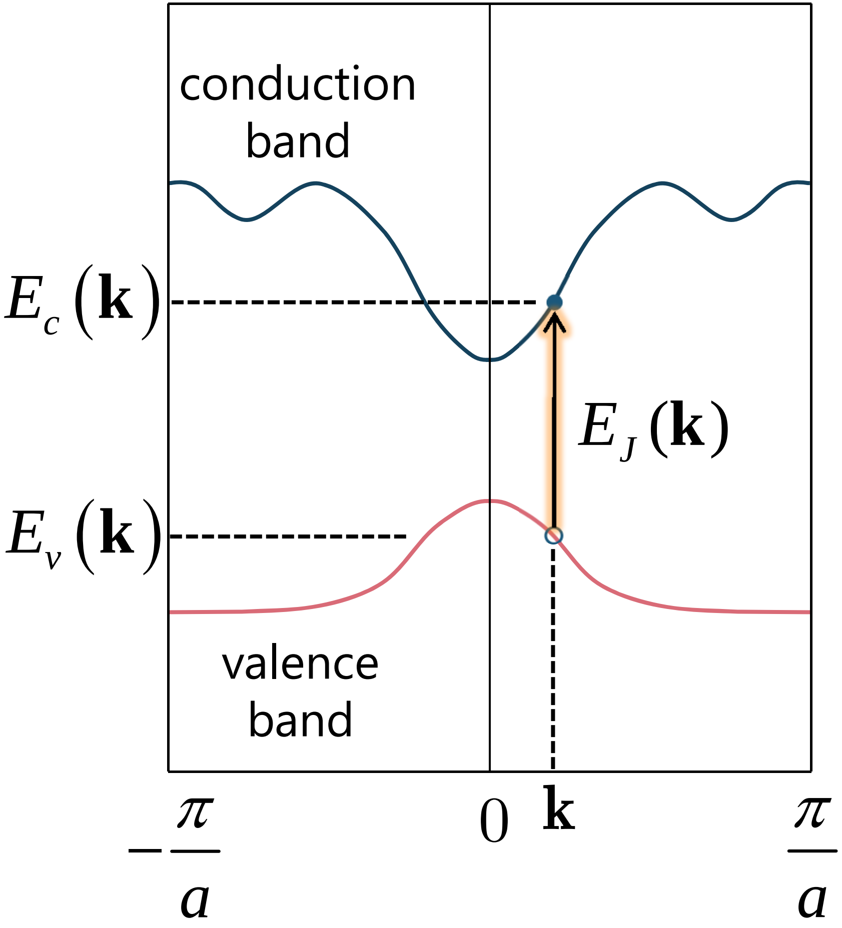
with m_{effe} and m_{effh} as the hole and electron effective masses

both positive by convention

we can define the transition energy

$$E_J(\mathbf{k}) = E_c(\mathbf{k}) - E_v(\mathbf{k})$$

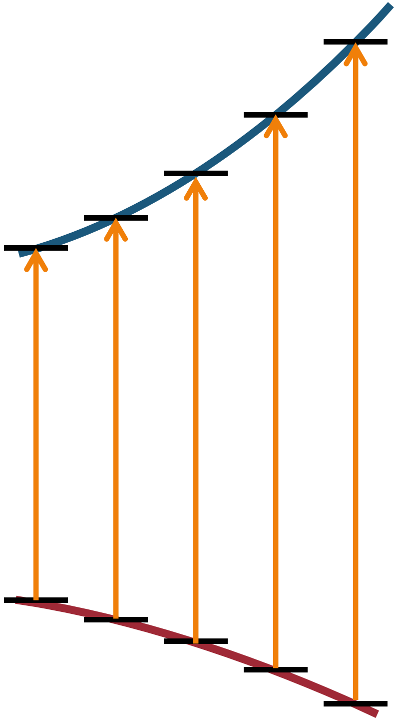
$$= \frac{\hbar^2 k^2}{2}\left(\frac{1}{m_{effe}} + \frac{1}{m_{effh}}\right) + E_g$$



Dense sets of possible transitions

Now we can think of a set of possible transitions

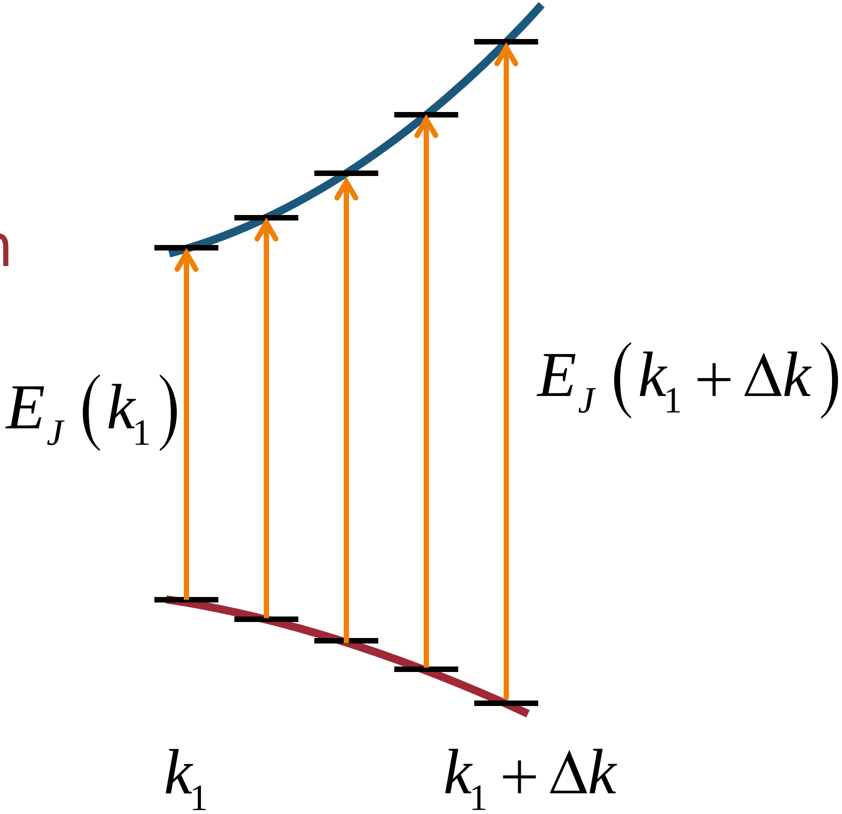
- corresponding to transitions between a specific k-state in the valence band
  - and the corresponding k-state in the conduction band
    - as sketched here for some portion of the Brillouin zone



Dense sets of possible transitions

Here, in a one-dimensional visualization

- the transition energy E_J goes from
  - $E_J(k_1)$ at one end, to
  - $E_J(k_1 + \Delta k)$ at the other



Dense sets of possible transitions

In what follows, we are interested in

the number of different transitions possible

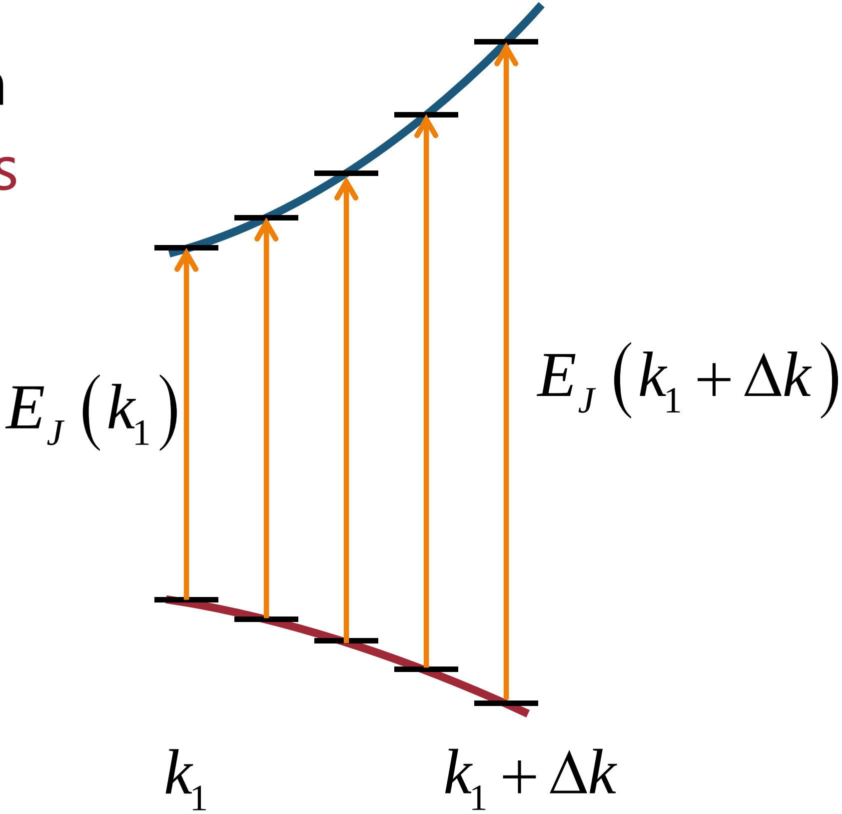
in a range of photon energies

from $E_J(k_1)$ to $E_J(k_1+\Delta k)$

that is, within a range

$$\Delta E_J = E_J(k_1+\Delta k) - E_J(k_1)$$

in the vicinity of $E_J(k_1)$



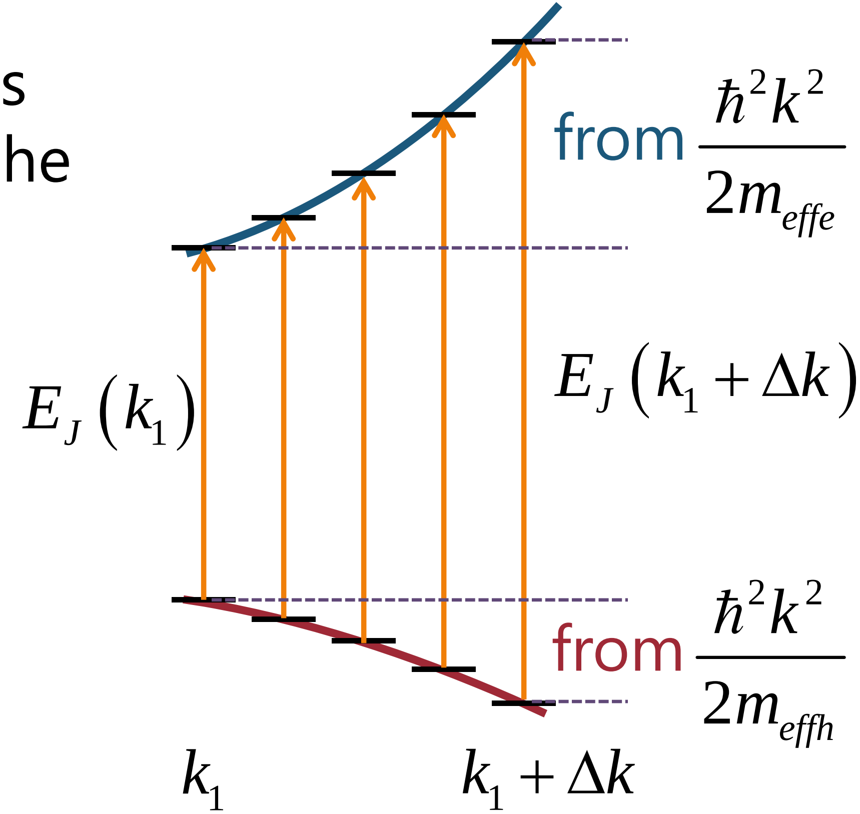
Dense sets of possible transitions

The difference in transition energies from one side of this diagram to the other

is the sum of two parts

one from $\dfrac{\hbar^2 k^2}{2m_{effe}}$

the other from $\dfrac{\hbar^2 k^2}{2m_{effh}}$



Dense sets of possible transitions

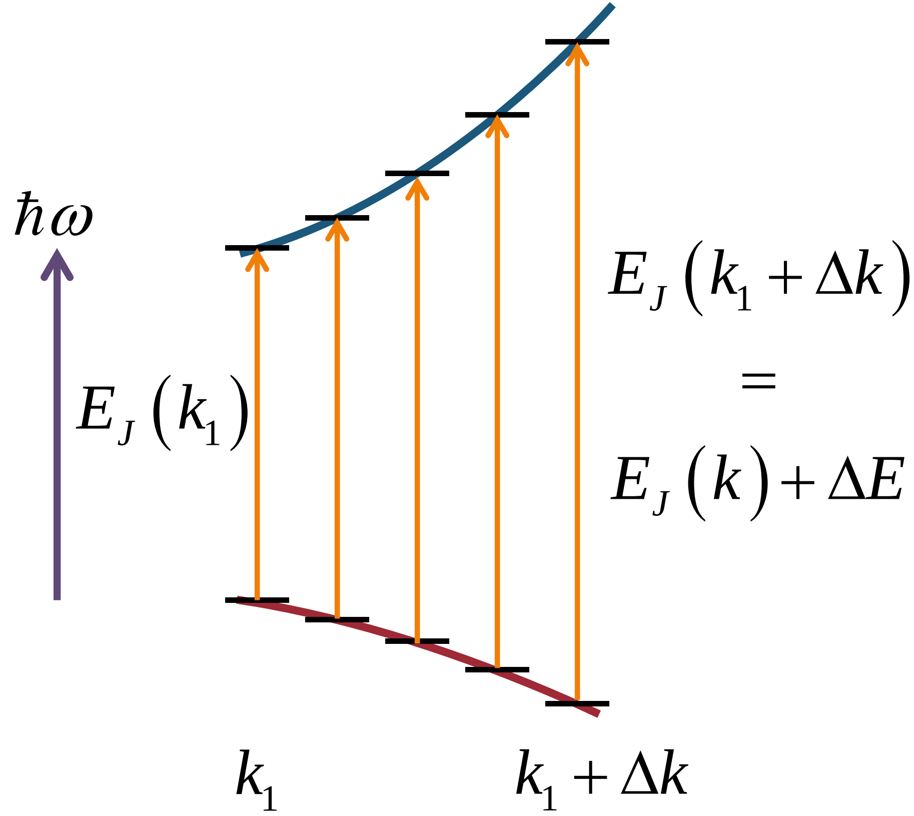
We presume this set of possible transitions is very dense

with density $g_J(\hbar\omega)$ per unit energy near photon energy $\hbar\omega$

giving $g_J(\hbar\omega)\Delta E$ transitions within energy range ΔE

$g_J(\hbar\omega)$ is known as a "joint density of states"

since it refers to transitions between states



Joint density of states

So, we can write the transition energies as

$$E_J(\mathbf{k}) = E_c(\mathbf{k}) - E_v(\mathbf{k}) = \frac{\hbar^2 k^2}{2}\left(\frac{1}{m_{effe}} + \frac{1}{m_{effh}}\right) + E_g = \frac{\hbar^2 k^2}{2\mu_{eff}} + E_g$$

where we have introduced a convenient parameter

a “reduced effective mass”

$$\frac{1}{\mu_{eff}} = \frac{1}{m_{effe}} + \frac{1}{m_{effh}}$$

Joint density of states

We can then define a "joint density of states" $g_J\left(E_J\right)$

the number of transitions per unit (photon) energy

where we view E_J as being a function of $\boldsymbol{k}$

Like the energy density of states in a band, we can write

$$g_J(E_J)dE_J = 2g(\mathbf{k})d^3\mathbf{k}$$

including a factor of 2 for spin

where $g\left(\mathbf{k}\right)$ is the density of states in k-space

just as before

Joint density of states

Because our expression for the transition energy

$$E_J(\mathbf{k}) = \frac{\hbar^2 k^2}{2\mu_{eff}} + E_g$$

has exactly the same form as the energy of k-states in a given parabolic band

we now follow mathematically identical arguments to those used to deduce

the density of states in energy

from the density of states in k-space

Joint density of states

Now we deduce

- the density of transitions per unit photon energy
  - now called the "joint density of states"
    - obtaining, for photon energy greater than the bandgap energy
      - i.e., $E_J \geq E_g$
        - the "joint density of states"

$$g_J(E_J) = \frac{1}{2\pi^2}\left(\frac{2\mu_{eff}}{\hbar^2}\right)^{3/2}\left(E_J - E_g\right)^{1/2}$$

Total transition rate

So, from $W_{TOT} = \frac{2\pi}{\hbar}\frac{e^2 A_o^2}{4m_o^2}\left|p_{cv}\right|^2 \sum_{\mathbf{k},\,\text{spin}} \delta\left[E_c\left(\mathbf{k}\right) - E_v\left(\mathbf{k}\right) - \hbar\omega\right]$

using $\sum_{\mathbf{k},\,\text{spin}} = 2\sum_{\mathbf{k}} \approx 2\int_{\mathbf{k}} g(\mathbf{k}) d^3\mathbf{k} = \int_{E_J \geq E_g} g_J\left(E_J - E_g\right) dE_J$

with $g_J(E_J) = \frac{1}{2\pi^2}\left(\frac{2\mu_{eff}}{\hbar^2}\right)^{3/2}\left(E_J - E_g\right)^{1/2}$

then

$$W_{TOT} = \frac{2\pi}{\hbar}\frac{e^2 A_0^2}{4m_0^2}\left|p_{cv}\right|^2 \int_{E_J \geq E_g} \frac{1}{2\pi^2}\left(\frac{2\mu_{eff}}{\hbar^2}\right)^{3/2}\left(E_J - E_g\right)^{1/2} \delta\left(E_J - \hbar\omega\right) dE_J$$

Total transition rate

Integrating to eliminate the delta function in

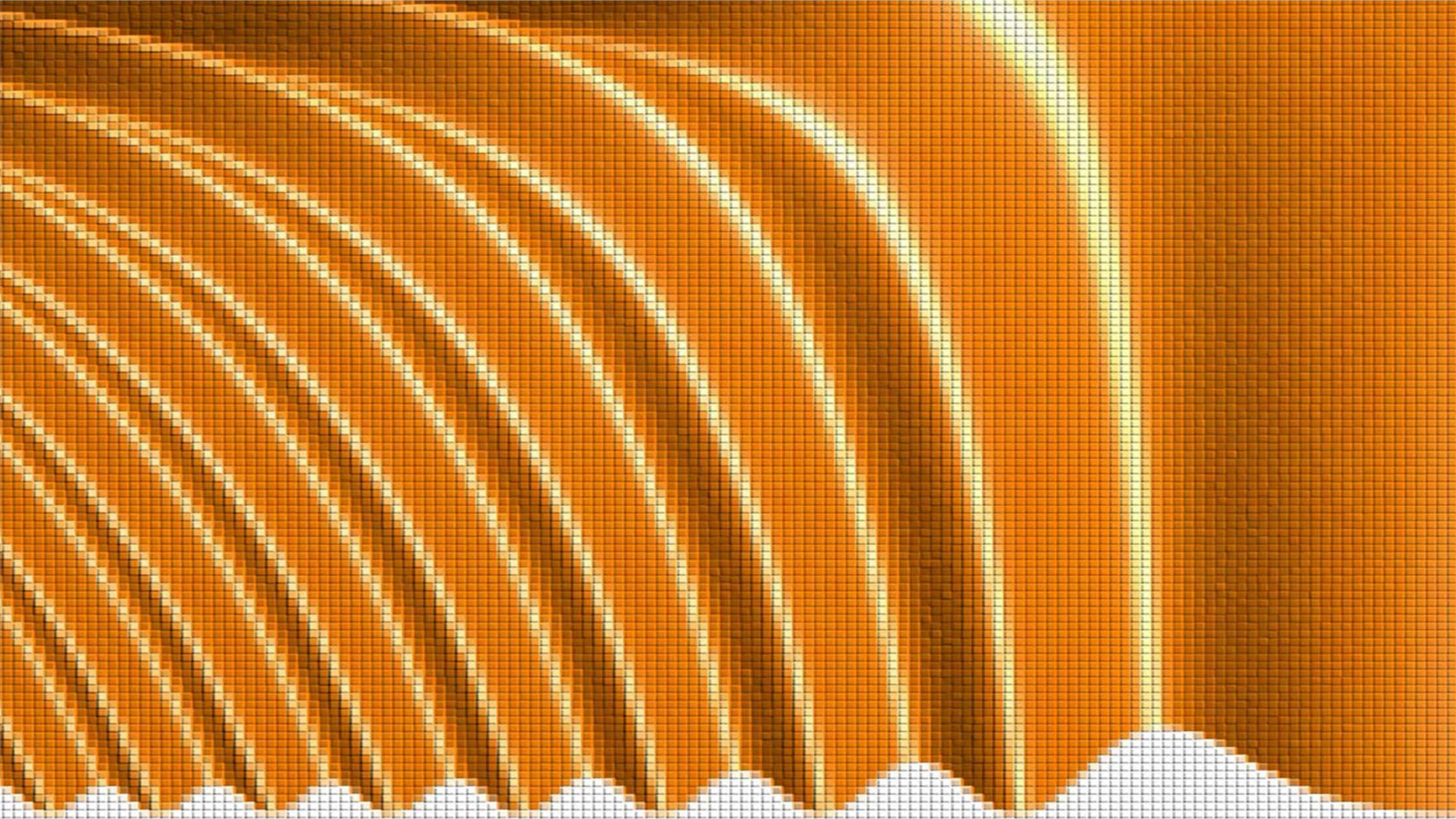
$$W_{TOT} = \frac{2\pi}{\hbar}\frac{e^2 A_o^2}{4m_o^2}\left|p_{cv}\right|^2 \int_{E_J \geq E_g} \frac{1}{2\pi^2}\left(\frac{2\mu_{eff}}{\hbar^2}\right)^{3/2}\left(E_J - E_g\right)^{1/2}\delta\left(E_J - \hbar\omega\right)dE_J$$

we finally obtain, for $\hbar\omega \geq E_g$

$$W_{TOT}(\hbar\omega) = \frac{2\pi}{\hbar}\frac{e^2 A_o^2}{4m_o^2}\left|p_{cv}\right|^2 \frac{1}{2\pi^2}\left(\frac{2\mu_{eff}}{\hbar^2}\right)^{3/2}\left(\hbar\omega - E_g\right)^{1/2}$$

Hence we finally see how Fermi's Golden Rule can be used with an appropriate density of states

to give a total transition rate

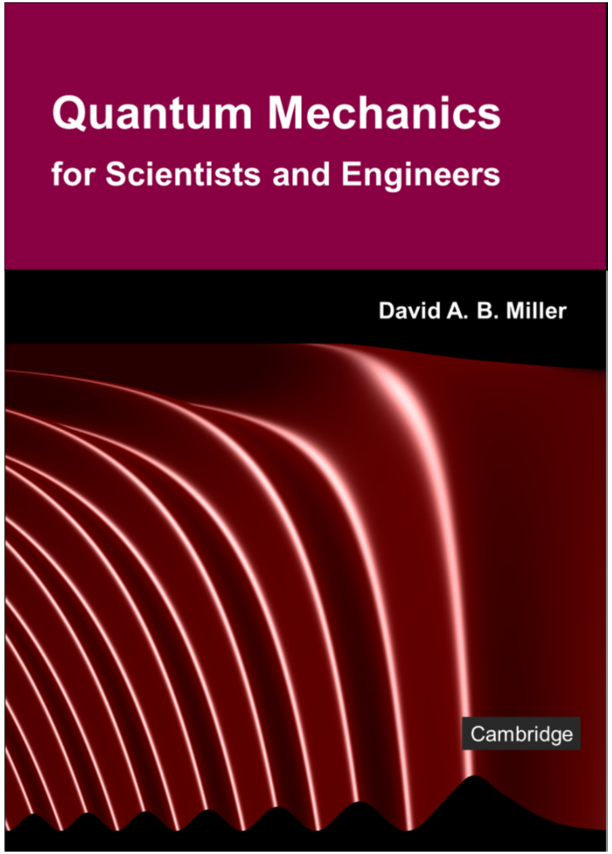


12.1 Optical absorption in semiconductors

Slides: Video 12.1.10 Absorption coefficient

Text reference: Quantum Mechanics for Scientists and Engineers

Section 8.10 starting just above Eq. 8.108



Optical absorption in semiconductors

Absorption coefficient

Quantum mechanics for scientists and engineers

David Miller

Absorption coefficient

The final step is to relate the absorption coefficient, α

the probability of absorption of a photon per unit length

to the total transition rate per unit volume W_{TOT}

With photon energy $\hbar\omega$

the number of photons incident per unit area per second

is, by definition $n_p = I / \hbar\omega$

where I is the optical intensity (power per unit area)

So the probability a photon is absorbed per unit length is

$$\alpha = \frac{W_{TOT}}{n_p} = \frac{\hbar\omega W_{TOT}}{I}$$

Absorption coefficient

From electromagnetism, we can deduce the intensity from the vector potential amplitude through

$$I = \frac{n_r c \varepsilon_0 \omega^2 A_o^2}{2}$$

where

n_r is the refractive index

c is the velocity of light, and

ε_o is the permittivity of free space (electric constant)

Absorption coefficient

Hence

$$\alpha(\hbar\omega) = \frac{\hbar\omega}{I} W_{TOT} = \frac{2\hbar\omega}{n_r c\varepsilon_0 \omega^2 A_o^2} \frac{2\pi}{\hbar} \frac{e^2 A_o^2}{4m_o^2} \frac{1}{2\pi^2} \left(\frac{2\mu_{eff}}{\hbar^2} \right)^{3/2} \left| p_{cv} \right|^2 \left(\hbar\omega - E_g \right)^{1/2}$$

$$= \frac{e^2}{2\pi m_o^2 c\varepsilon_o} \frac{1}{n_r \omega} \left| p_{cv} \right|^2 \left(\frac{2\mu_{eff}}{\hbar^2} \right)^{3/2} \left(\hbar\omega - E_g \right)^{1/2}$$

The parameter $E_p = \left(2/m_o\right)\left|p_{cv}\right|^2$ (~ 20 eV in many semiconductors) is often used in calculations

in which case we can rewrite

$$\alpha\left(\hbar\omega\right) = \frac{\hbar e^2}{4\pi m_o c\varepsilon_o} \frac{1}{n_r} \frac{E_p}{\hbar\omega} \left(\frac{2\mu_{eff}}{\hbar^2} \right)^{3/2} \left(\hbar\omega - E_g \right)^{1/2}$$

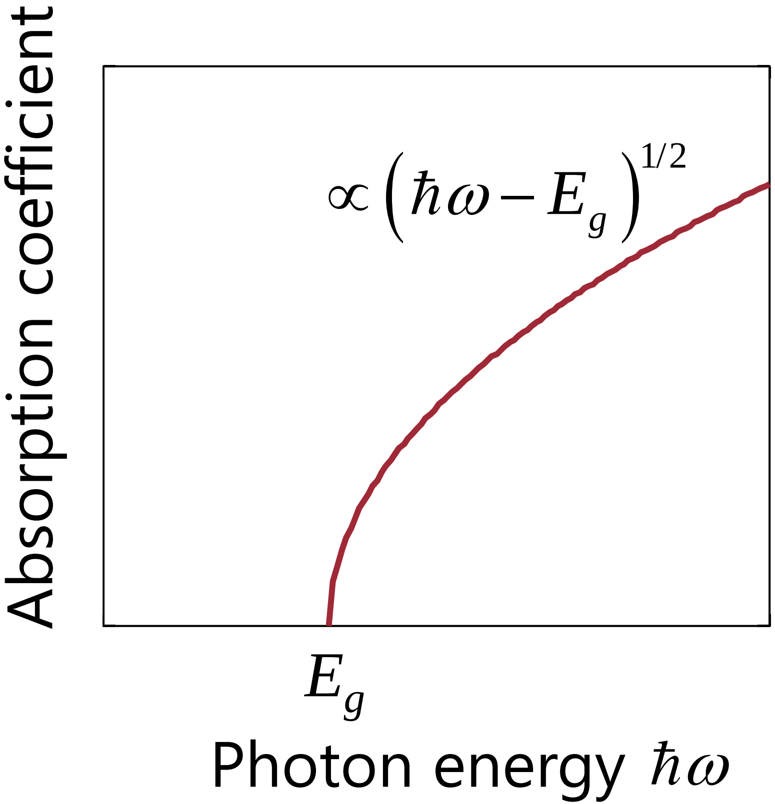
Optical absorption

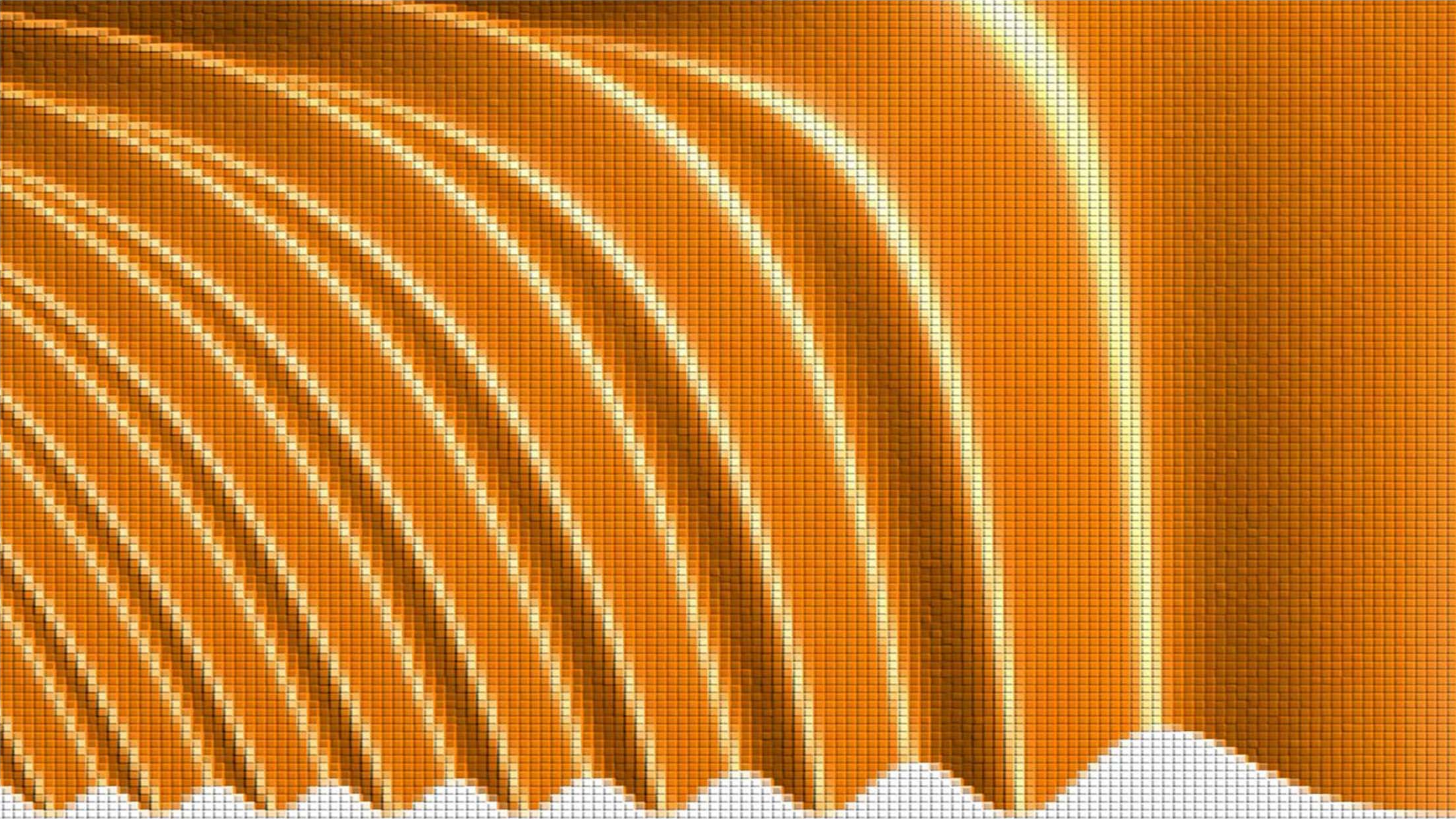
In direct gap semiconductors

the optical absorption in this model

rises as $\left(\hbar\omega - E_g\right)^{1/2}$ above the bandgap energy E_g

This model is used as the starting point for optical calculations in semiconductors



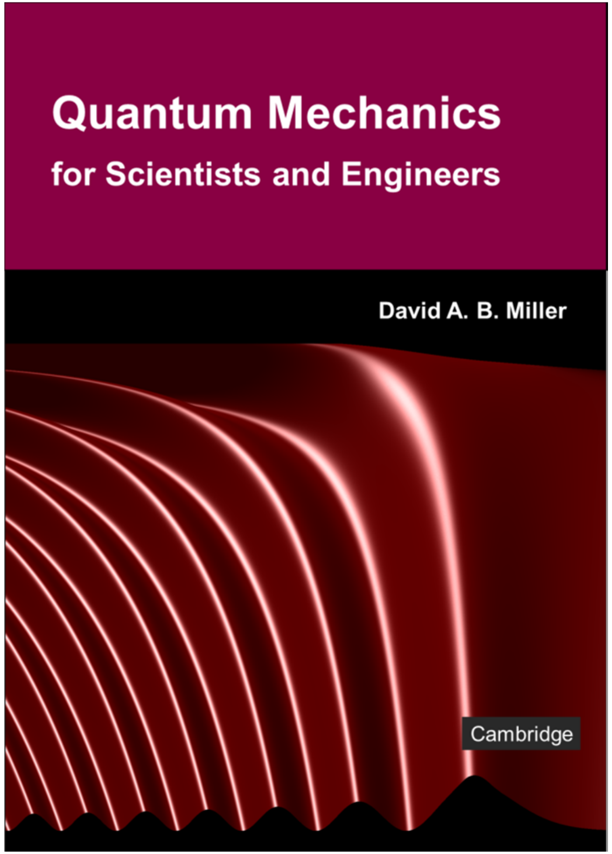


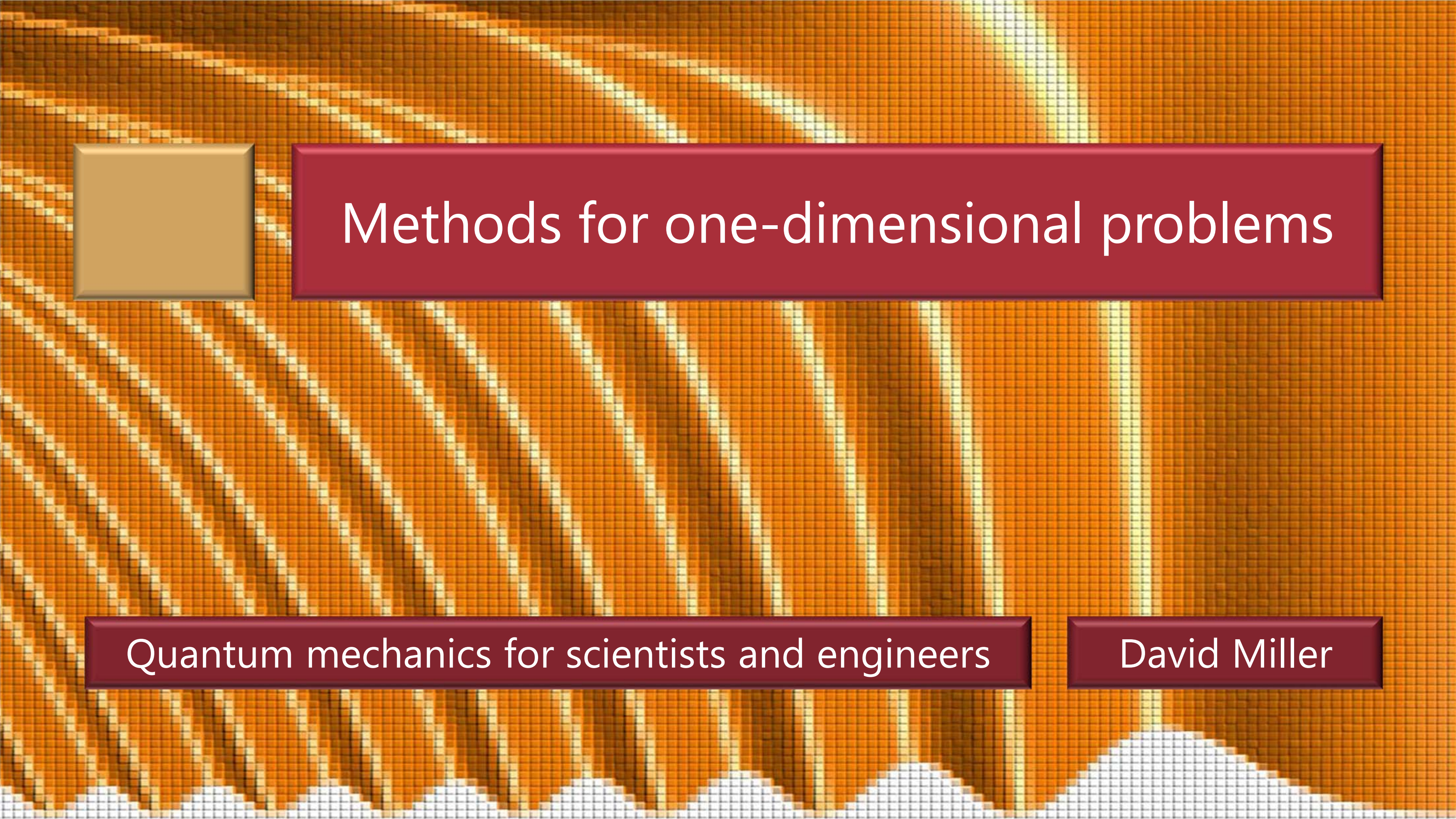
12.2 Methods for one-dimensional problems

Slides: Video 12.2.1 Introduction to methods for one-dimensional problems

Text reference: Quantum Mechanics for Scientists and Engineers

Section Chapter 11 introduction





Methods for one-dimensional problems

Quantum mechanics for scientists and engineers

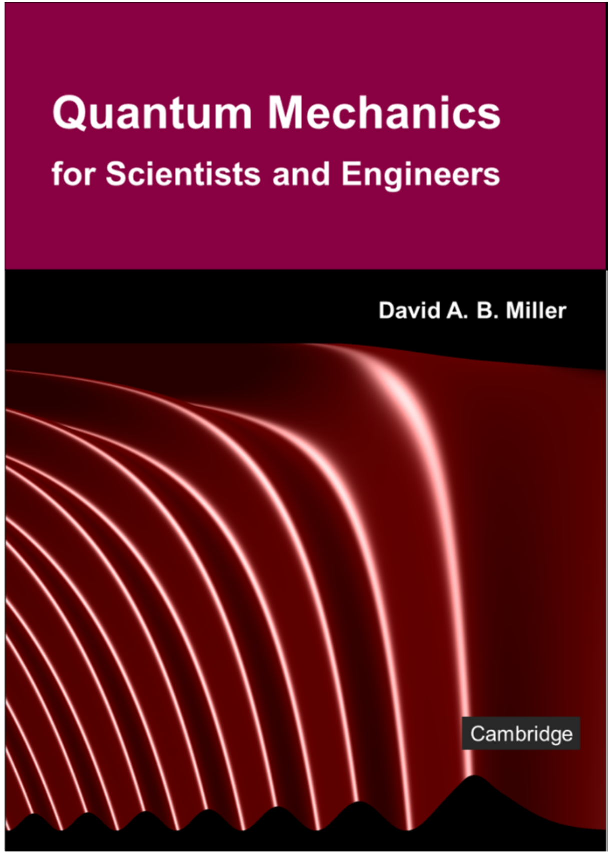
David Miller

12.2 Methods for one-dimensional problems

Slides: Video 12.2.2 Tunneling currents

Text reference: Quantum Mechanics for Scientists and Engineers

Section 11.1



Methods for one-dimensional problems

Tunneling currents

Quantum mechanics for scientists and engineers

David Miller

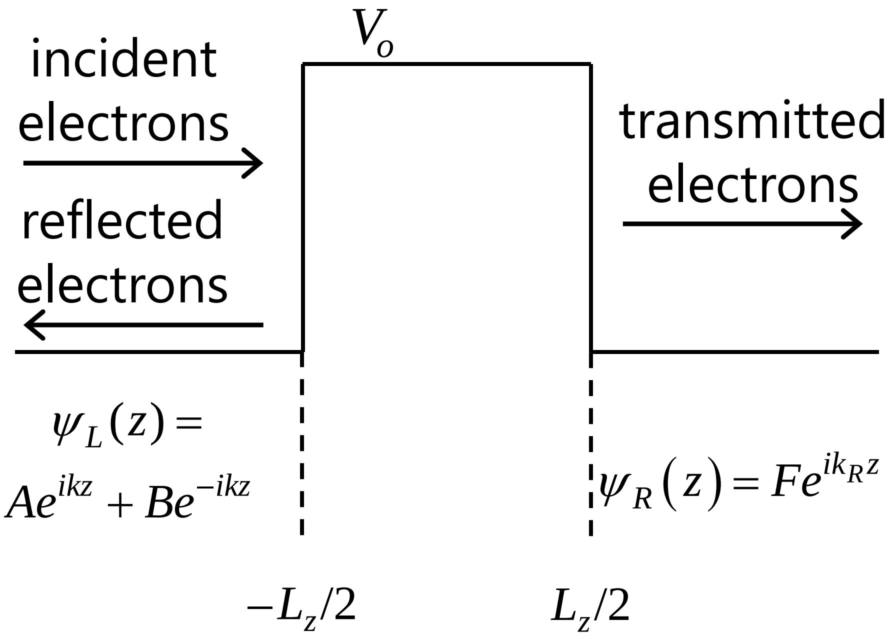
Tunneling rates

Consider a simple rectangular barrier

- with incident electron energy E
- below the (peak) barrier height V_o
  - i.e., $E < V_o$

We also presume no electrons incident from the right

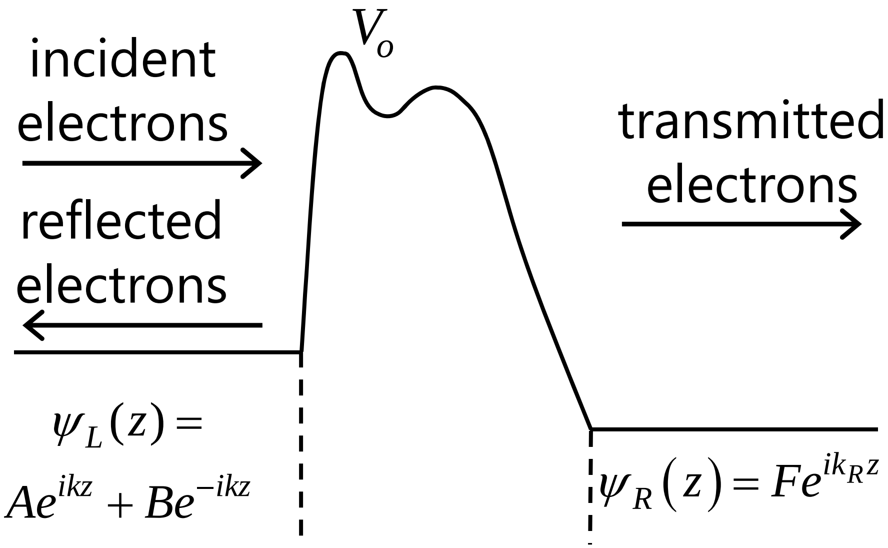
- so there are only transmitted electrons there



Tunneling rates

We could have a more complicated barrier

- still with $E < V_o$
  - but possibly with a different potential on the right
    - and different wavevector k_R



V_o

incident electrons

reflected electrons

transmitted electrons

$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$

$\psi_R(z) = Fe^{ik_R z}$

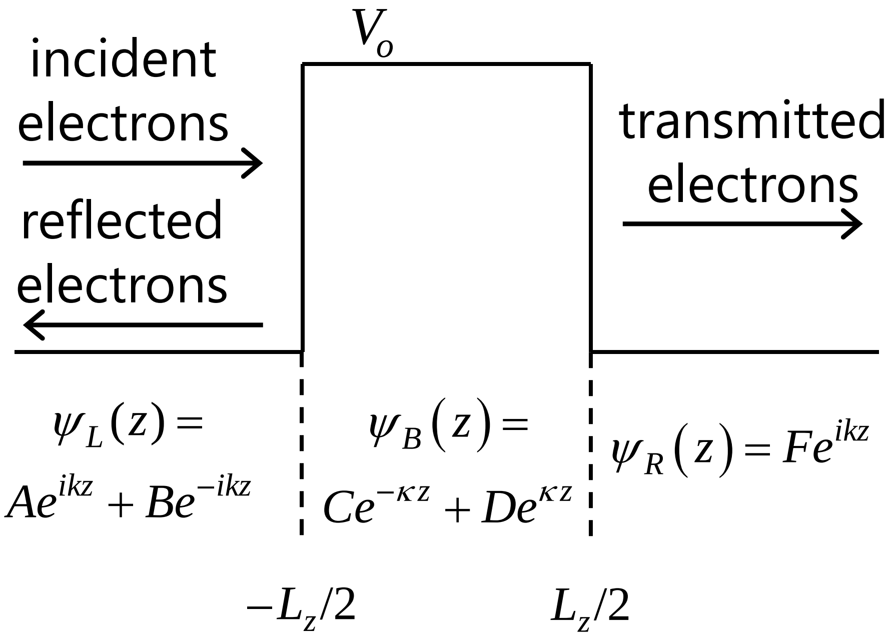
Tunneling rates

Suppose we have found the relations between the amplitudes of the

incident, A,

reflected, B, and

transmitted, F, waves



$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$

$\psi_B(z) = Ce^{-\kappa z} + De^{\kappa z}$

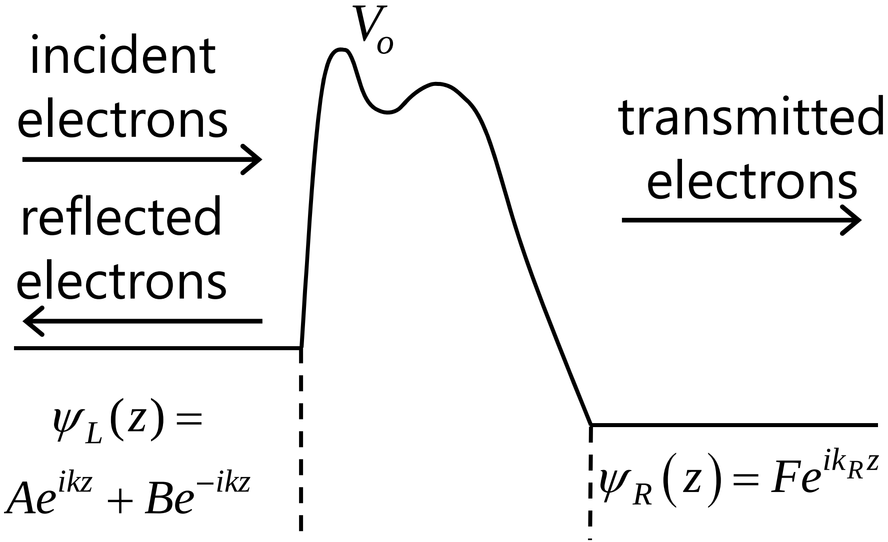
$\psi_R(z) = Fe^{ikz}$

Tunneling rates

Suppose we have found the relations between the amplitudes of the

- incident, A,
- reflected, B, and
- transmitted, F, waves
  - in either case

How do we relate these to actual electron currents?



incident electrons

reflected electrons

transmitted electrons

V_o

$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$

$\psi_R(z) = Fe^{ik_R z}$

Evaluation of tunneling current

The particle current density in quantum mechanics is

$$\mathbf{j}_p = \frac{i\hbar}{2m}\left(\Psi\nabla\Psi^* - \Psi^*\nabla\Psi\right)$$

where $\Psi = \Psi(\mathbf{r},t)$ is the time-dependent wavefunction

If we consider particles of well-defined energy E

the wavefunction is of the form $\Psi(\mathbf{r},t) = \psi(\mathbf{r})\exp(-iEt/\hbar)$

In the products $\Psi\nabla\Psi^*$ and $\Psi^*\nabla\Psi$,

the term $\exp(-iEt/\hbar)$ is multiplied by its complex conjugate to give 1, so we then have

$$\mathbf{j}_p = \frac{i\hbar}{2m}\left(\psi\nabla\psi^* - \psi^*\nabla\psi\right)$$

Evaluation of tunneling current

If we consider only a one-dimensional problem

e.g., for a potential varying only in the z direction

we only need to calculate the current in the z direction

which we can call j_p

so we simplify $\mathbf{j}_p = \frac{i\hbar}{2m}\left(\psi \nabla \psi^* - \psi^* \nabla \psi\right)$

to just

$$j_p = \frac{i\hbar}{2m}\left(\psi \nabla \psi^* - \psi^* \nabla \psi\right)$$

where $\psi = \psi(z)$ is now just a spatial wavefunction varying in z

Evaluation of tunneling current

For a simple barrier

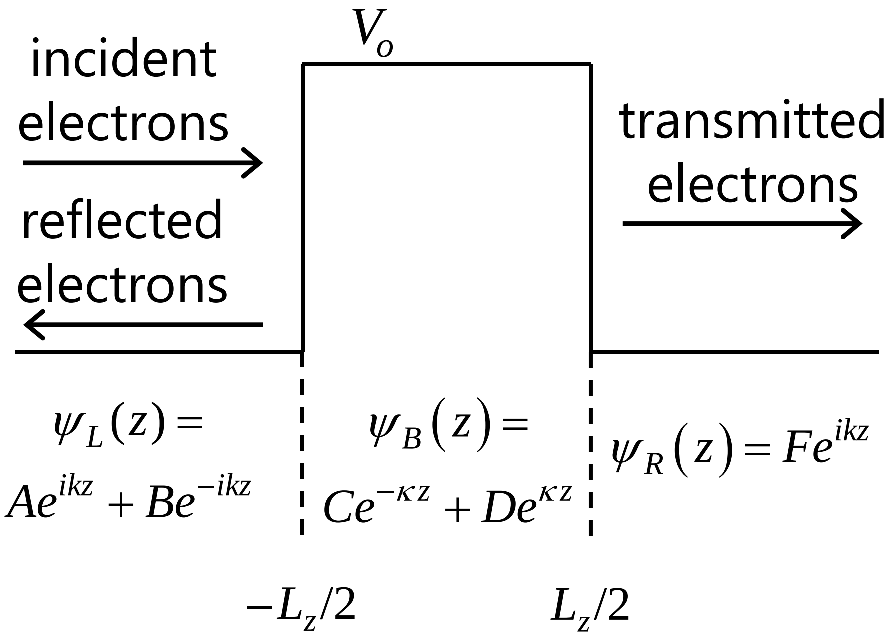
with the same potential on the left and the right

which we take to be zero for simplicity

for a particle of mass m and energy E

we have, as usual

$$E = \frac{\hbar^2 k^2}{2m}$$



V_o

incident electrons

reflected electrons

transmitted electrons

$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$

$\psi_B(z) = Ce^{-\kappa z} + De^{\kappa z}$

$\psi_R(z) = Fe^{ikz}$

$-L_z/2$

$L_z/2$

Evaluation of tunneling current

With the wave on the right in the form $\psi_R(z) = Fe^{ikz}$

from $j_p = \frac{i\hbar}{2m}\left(\psi\nabla\psi^* - \psi^*\nabla\psi\right)$

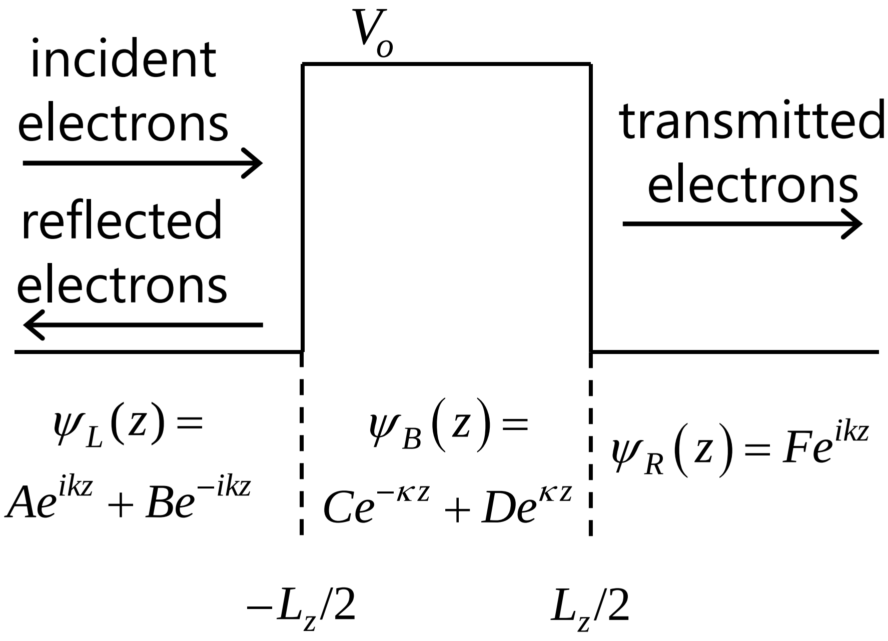
noting that

$$\psi\nabla\psi^* = |F|^2 \exp(ikz)\frac{d}{dz}\exp(-ikz)$$

$$= -ik|F|^2$$

and similarly $\psi^*\nabla\psi = ik|F|^2$

we have $j_p = |F|^2 \frac{\hbar k}{m}$



Evaluation of tunneling current

On the left, with $\psi(z) = A\exp(ikz) + B\exp(-ikz)$

from $j_p = (i\hbar / 2m)(\psi \nabla \psi^* - \psi^* \nabla \psi)$

we have

$$j_p = \frac{i\hbar}{2m}\left\{ \begin{array}{l} \left[A\exp(ikz) + B\exp(-ikz)\right]\left[-ikA^* \exp(-ikz) + ikB^* \exp(ikz)\right] \\ -\left[A^* \exp(-ikz) + B^* \exp(ikz)\right]\left[ikA\exp(ikz) - ikB\exp(-ikz)\right] \end{array} \right\}$$

$$= \frac{\hbar k}{m}\left(|A|^2 - |B|^2\right)$$

Note that all the spatially oscillating terms cancel

The net current is not varying spatially on the left

Evaluation of tunneling current

Since we have deduced

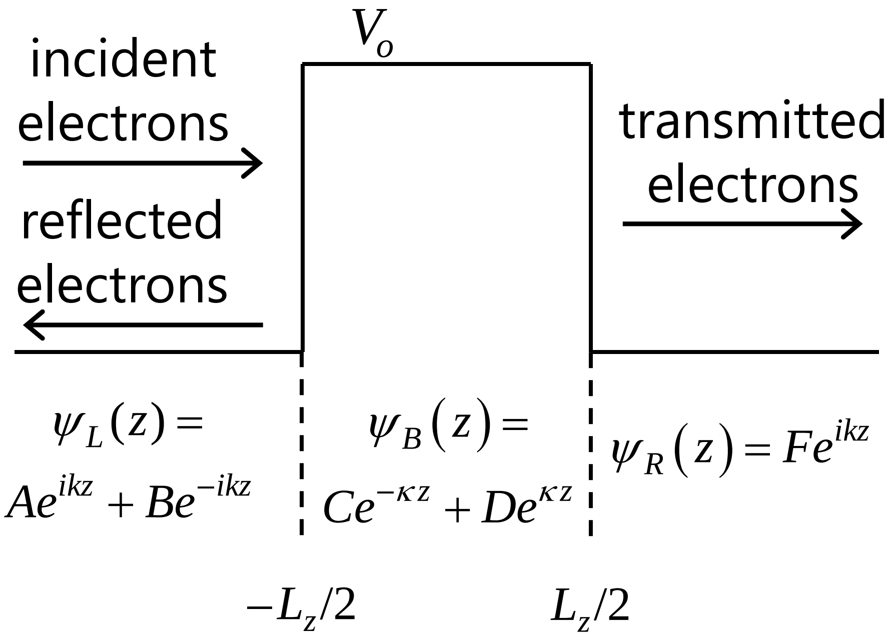
$$j_p = (\hbar k / m)\left(|A|^2 - |B|^2\right)$$

we can therefore consider

$\hbar k |A|^2 / m$ as the forward current on the left and

$\hbar k |B|^2 / m$ as the reflected or backward current

adding the two to get the net current



V_o

incident electrons

reflected electrons

transmitted electrons

$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$

$\psi_B(z) = Ce^{-\kappa z} + De^{\kappa z}$

$\psi_R(z) = Fe^{ikz}$

$-L_z/2$

$L_z/2$

Relation to group velocity

For $E = \hbar^2 k^2 / 2m \equiv \hbar\omega$

where ω is the frequency associated with E, gives

group velocity $v_g = \frac{d\omega}{dk} = \frac{\hbar k}{m}$

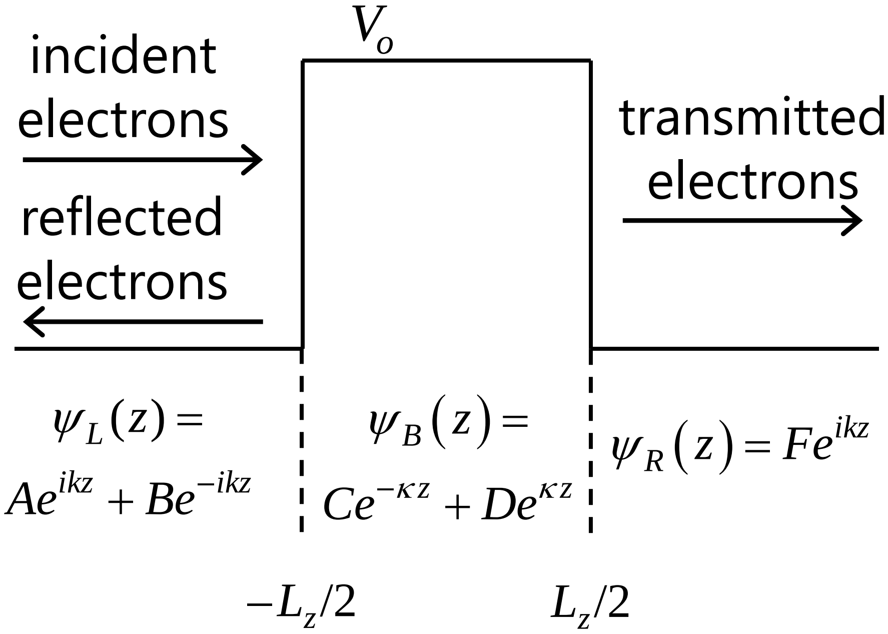
So the currents can be written

forward $\hbar k |A|^2 / m \equiv |A|^2 v_g$

backward $\hbar k |B|^2 / m \equiv |B|^2 v_g$

transmitted $\hbar k |F|^2 / m \equiv |F|^2 v_g$

though group velocity is not required for our argument



V_o

incident electrons

reflected electrons

transmitted electrons

$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$

$\psi_B(z) = Ce^{-\kappa z} + De^{\kappa z}$

$\psi_R(z) = Fe^{ikz}$

$-L_z/2$

$L_z/2$

Evaluation of tunneling current

With these currents

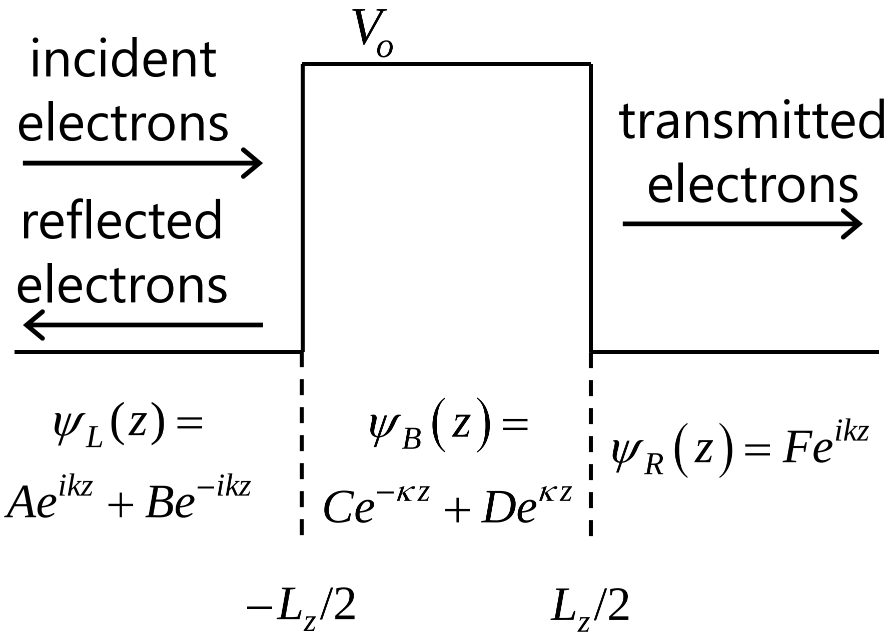
forward $\hbar k|A|^2/m$

backward $\hbar k|B|^2/m$

the fraction transmitted by the barrier can be written as

$$\eta = \frac{\text{current not reflected}}{\text{incident forward current}}$$

$$= \frac{|A|^2 - |B|^2}{|A|^2}$$



Evaluation of tunneling current

It might seem more obvious to write, with currents

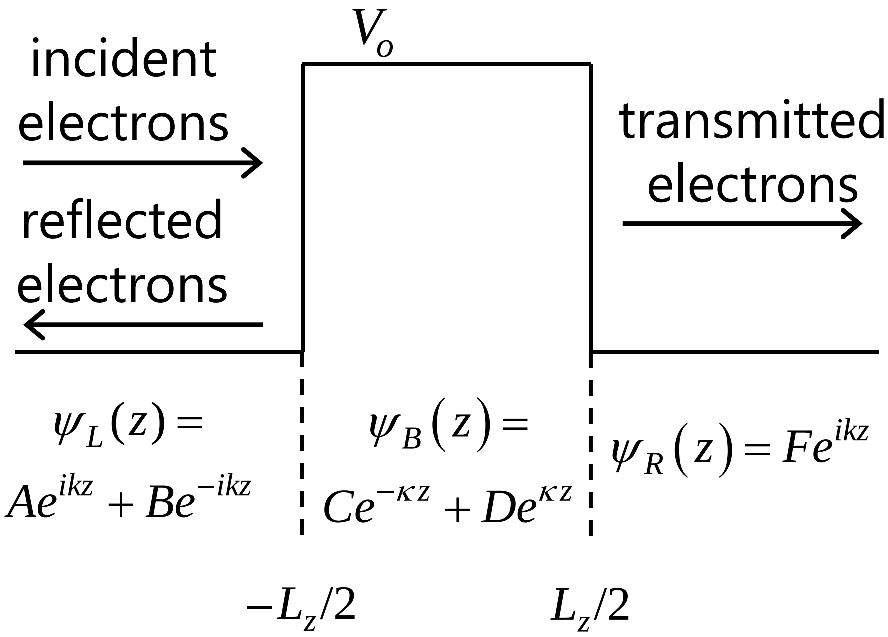
forward $\hbar k|A|^2 / m$

transmitted $\hbar k|F|^2 / m$

that the fraction transmitted is

$$\eta = \frac{\text{current transmitted}}{\text{incident forward current}}$$

$$= \frac{|F|^2}{|A|^2}$$



V_o

incident electrons

reflected electrons

transmitted electrons

$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$

$\psi_B(z) = Ce^{-\kappa z} + De^{\kappa z}$

$\psi_R(z) = Fe^{ikz}$

$-L_z/2$

$L_z/2$

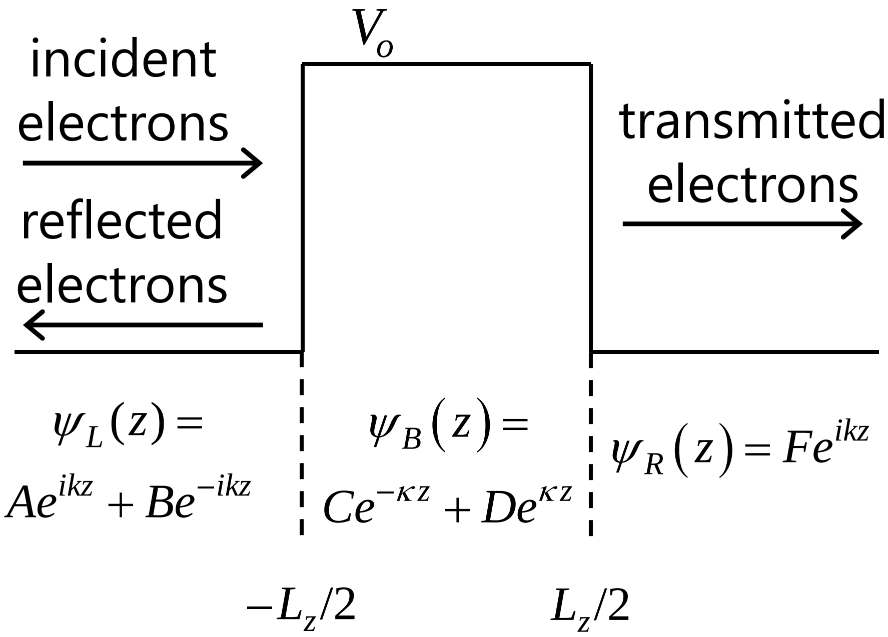
Evaluation of tunneling current

For a barrier with the same potential and material on both sides

these two expressions

$$\eta = \frac{|A|^2 - |B|^2}{|A|^2} \quad \text{and} \quad \eta = \frac{|F|^2}{|A|^2}$$

give the same answers



V_o

incident electrons

reflected electrons

transmitted electrons

$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$

$\psi_B(z) = Ce^{-\kappa z} + De^{\kappa z}$

$\psi_R(z) = Fe^{ikz}$

$-L_z/2$

$L_z/2$

Tunneling rates

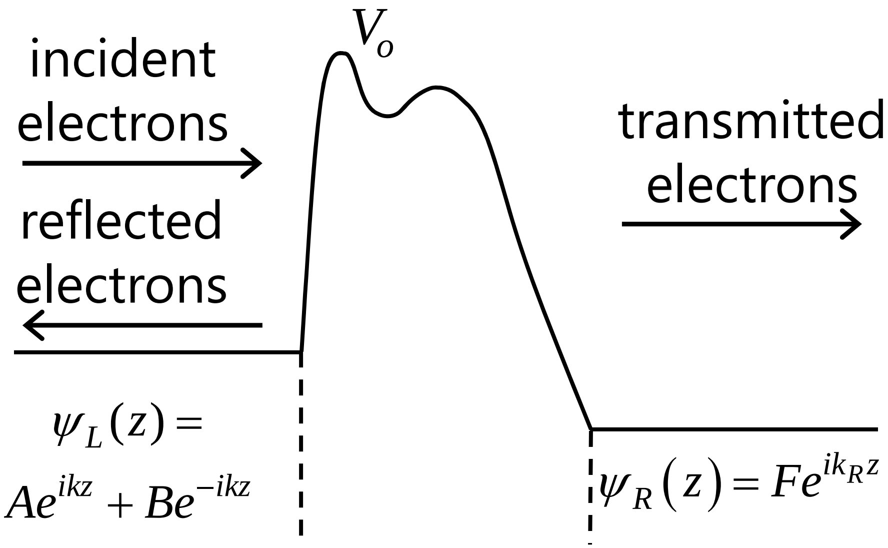
But with different potential or material on the right

- the group velocity may be different on the right
  - or might not be defined in some complicated case

While we might handle that

$$\eta = \left(|A|^2 - |B|^2 \right) / |A|^2$$

- avoids these issues
  - and is often otherwise just as easy to calculate



Evaluation of tunneling current

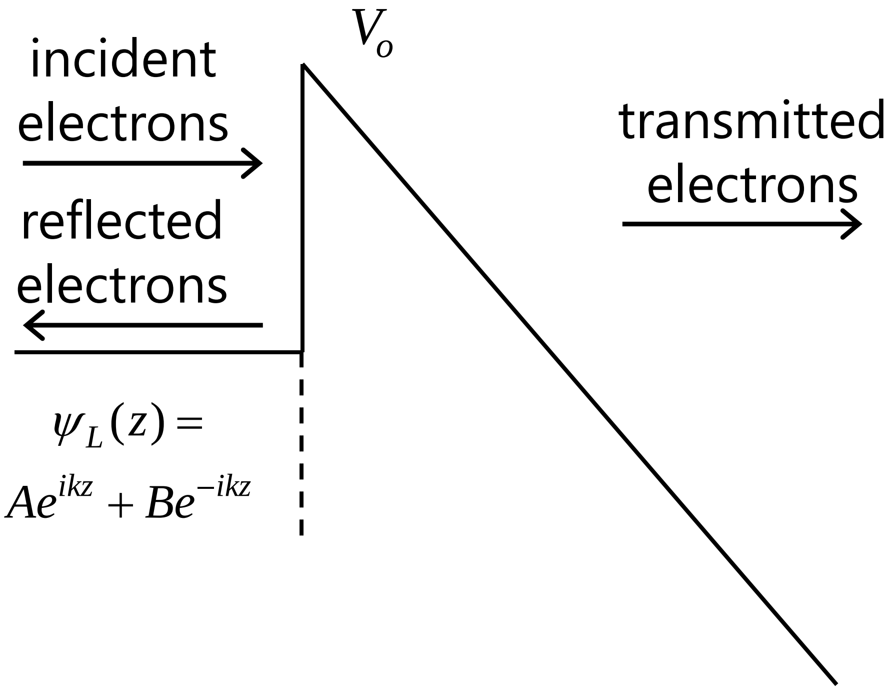
For example

- in field-emission tunneling
  - the barrier may continue its slope
    - giving no constant group velocity on the right

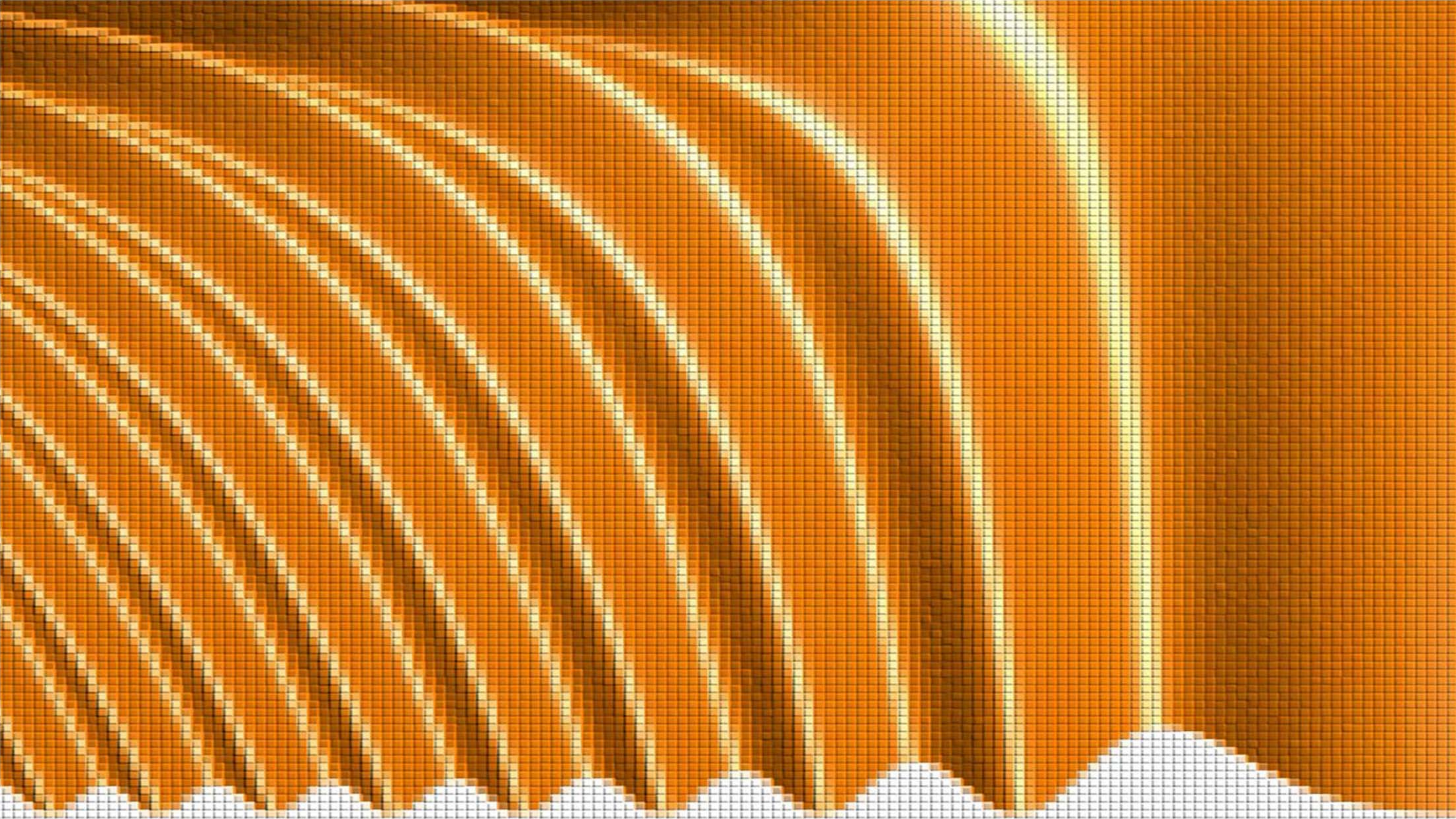
Here, however,

$$\eta = \frac{|A|^2 - |B|^2}{|A|^2}$$

still works



$$\psi_L(z) = Ae^{ikz} + Be^{-ikz}$$

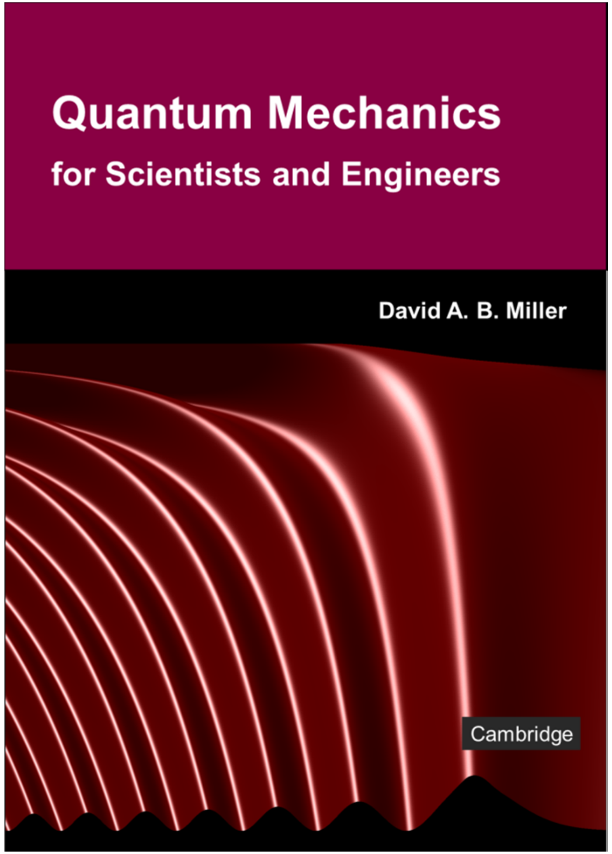


12.2 Methods for one-dimensional problems

Slides: Video 12.2.4 Transfer matrix method

Text reference: Quantum Mechanics for Scientists and Engineers

Section 11.2 up to “Calculation of eigenenergies …”



Methods for one-dimensional problems

Transfer matrix method

Quantum mechanics for scientists and engineers

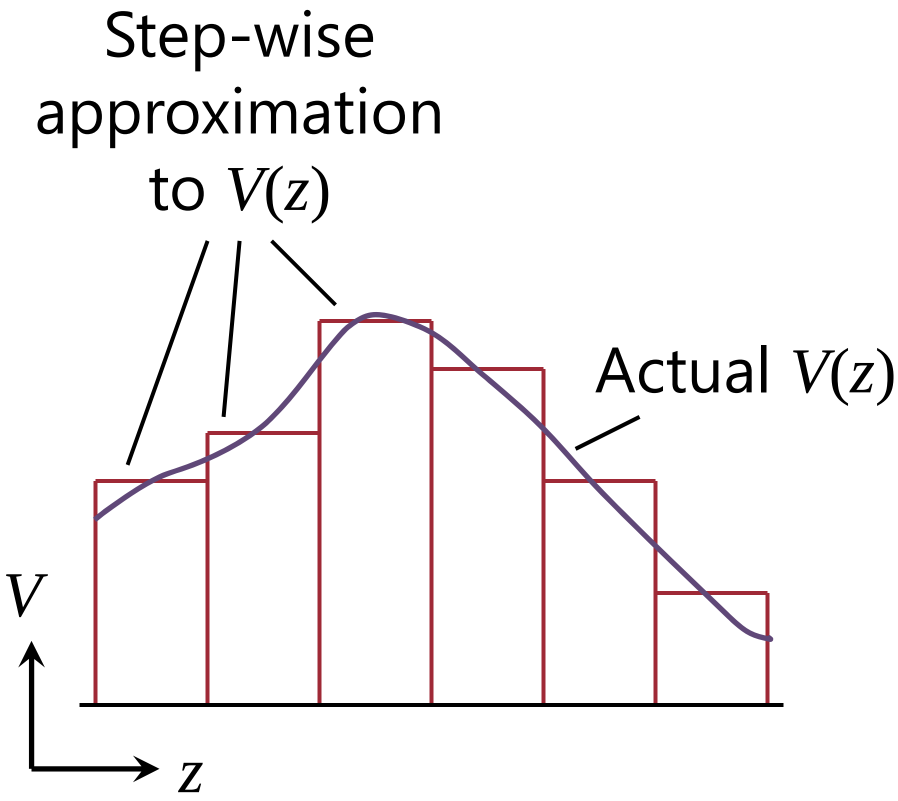
David Miller

Transfer matrix method

We presume that the potential is a series of steps

This could be an actual step-like potential

or we could be approximating some continuously varying potential



Transfer matrix method

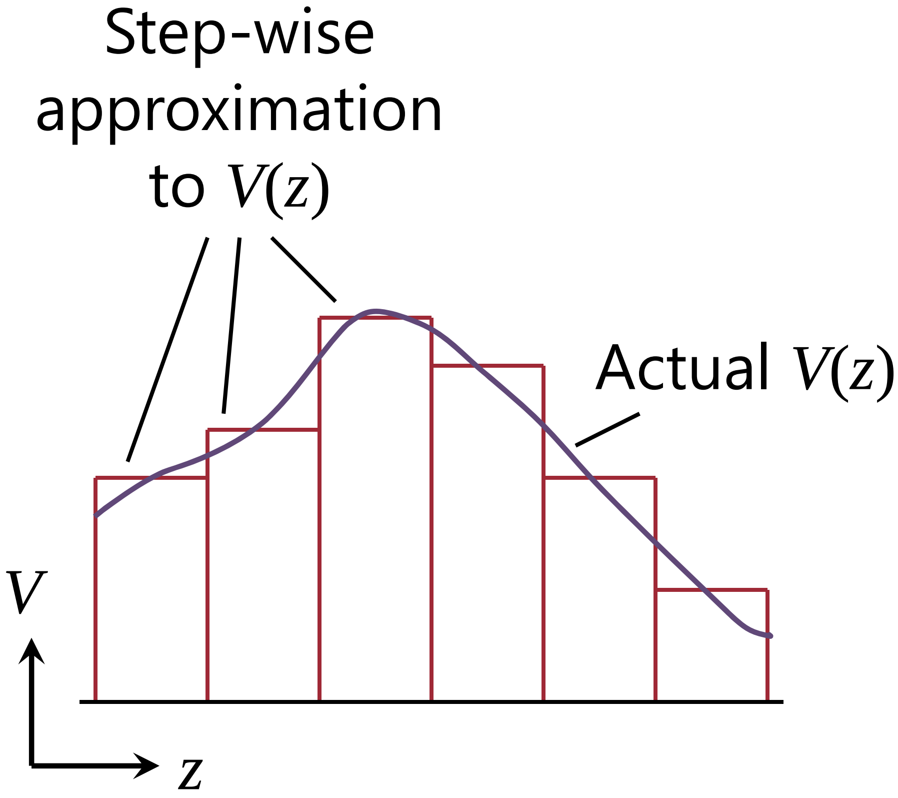
We therefore reduce the problem to that of

waves within a simple constant potential

which are either sinusoidal or exponential

together with appropriate boundary conditions

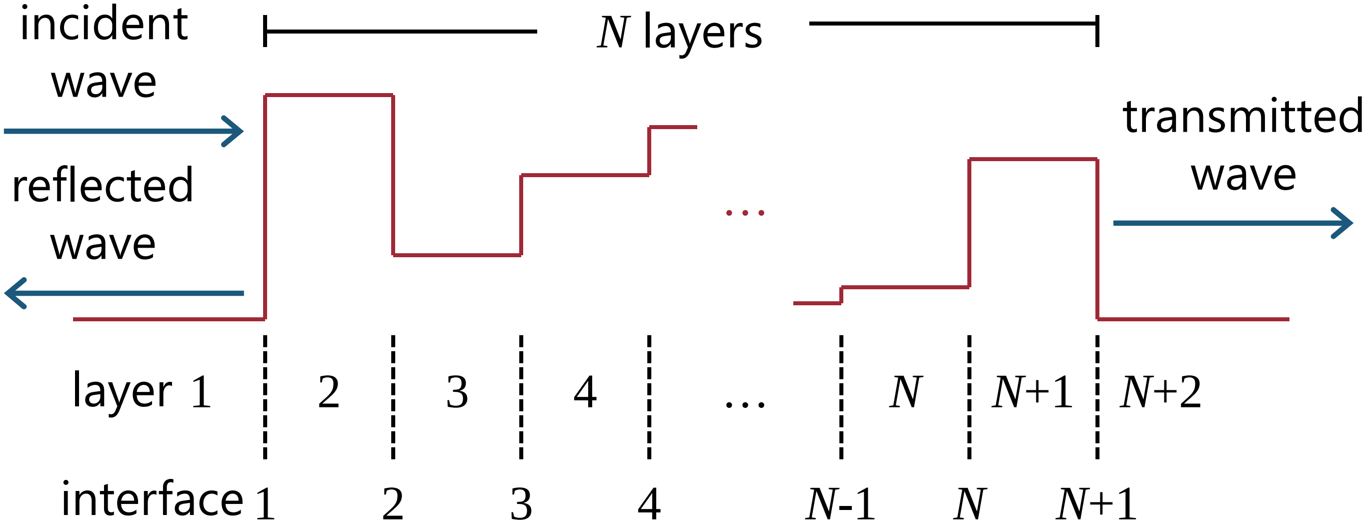
to link the solutions in adjacent layers



Transfer matrix method

Consider an electron wave incident on the structure

- from one side, with a particular energy, E
  - There will be reflected waves
    - and transmitted waves



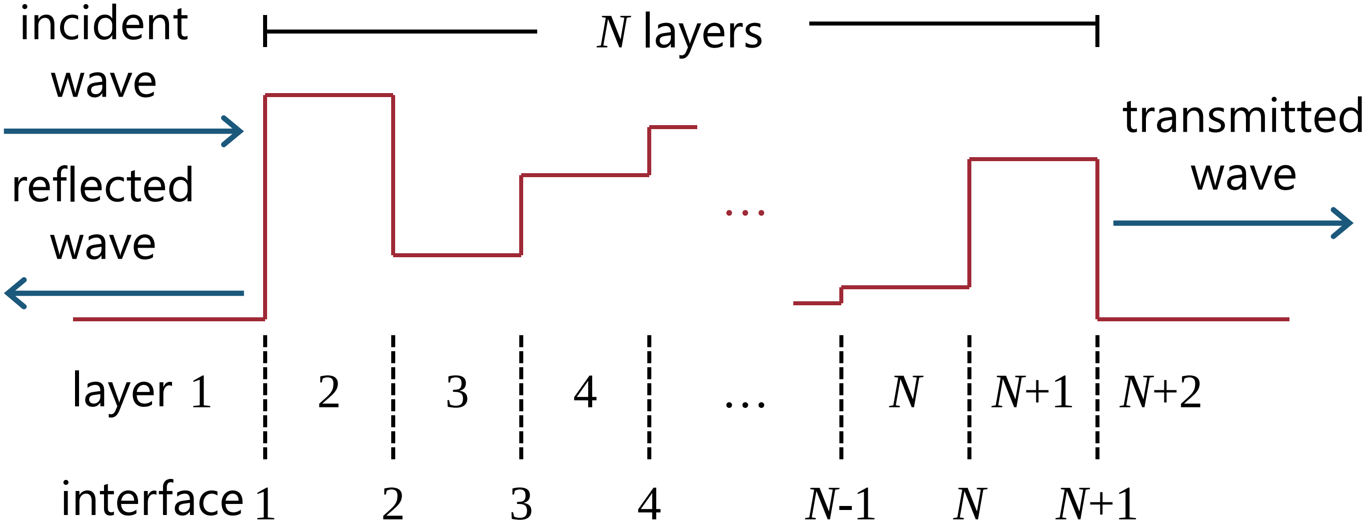
Transfer matrix method

For each layer in the structure, we derive a matrix that

relates the forward and backward amplitudes

A_m and B_m, just to the right of the (m-1)th interface, to

A_{m+1} and B_{m+1}, just to the right of the mth interface

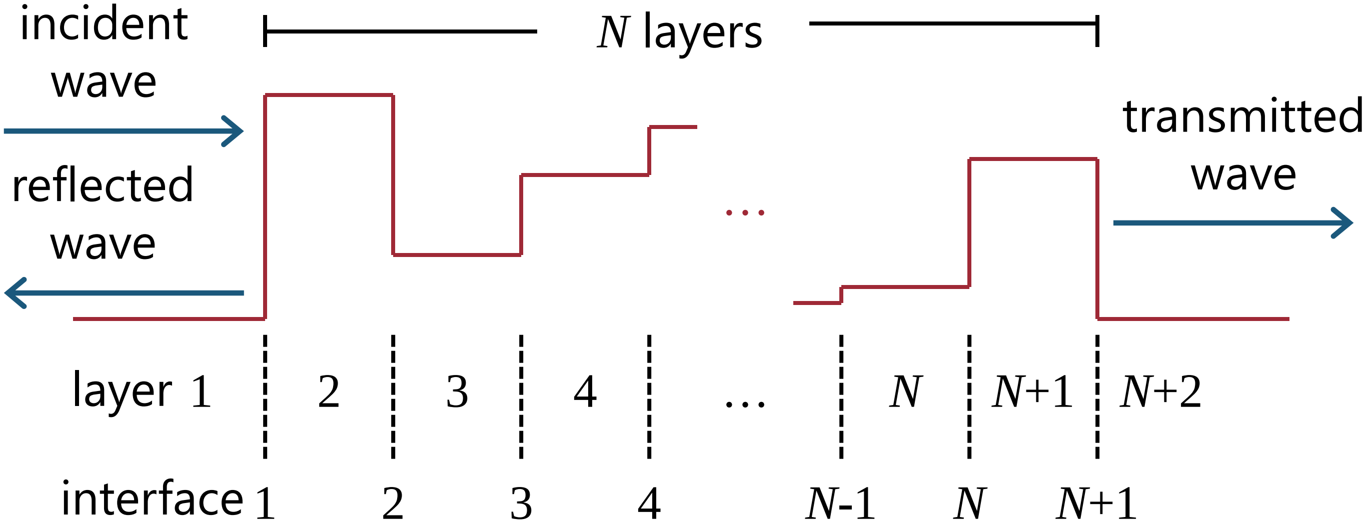


Transfer matrix method

By multiplying those matrices together for all of the layers

we will construct a single "transfer matrix" for the whole structure

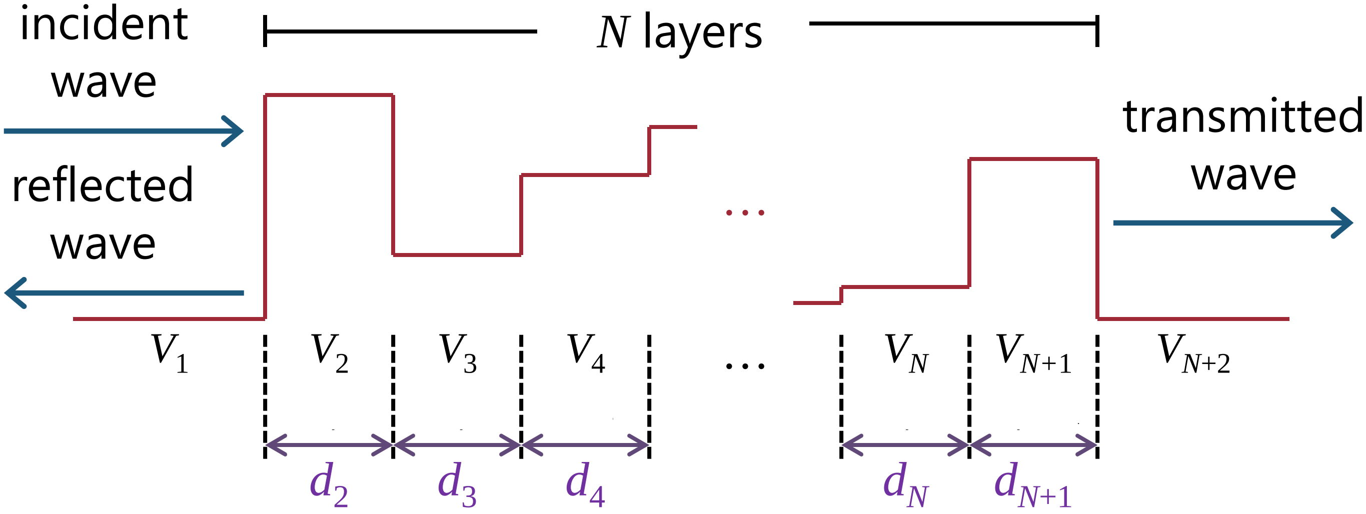
enabling us to analyze the entire multilayer structure



Transfer matrix method

In this formalism, each layer m will have

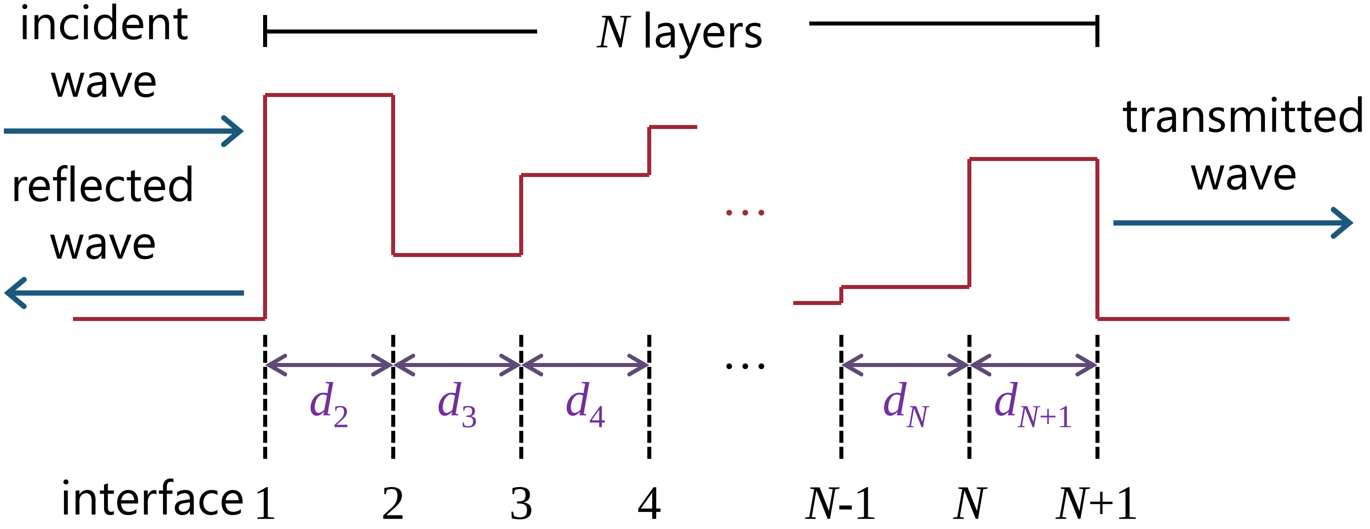
- a potential energy V_m
- a thickness d_m
- and possibly a mass or effective mass m_{fm}



Transfer matrix method

For interfaces 2 and higher, the position of the mth interface relative to interface 1, e.g., $z_2=d_2$, $z_3=d_2+d_3$, etc., is

$$z_m = \sum_{q=2}^{m} d_q$$



Transfer matrix method

In any given layer, if the particle energy $E > V_m$

we know we will have in general both

a "forward" propagating wave $A = A_o \exp\left[ik_m\left(z - z_{m-1}\right)\right]$

and

a "backward" propagating wave $B = B_o \exp\left[-ik_m\left(z - z_{m-1}\right)\right]$

where A and B are complex numbers for

the forward and backward wave amplitudes

In this case

$$k_m = \sqrt{\frac{2m_{fm}}{\hbar^2}\left(E - V_m\right)}$$

where m_{fm} is the mass of the particle in that layer

Transfer matrix method

Similarly, if the particle energy $E < V_m$

we know we will have in general both

a "forward" decaying "wave" $A = A_o \exp\left[-\kappa_m \left(z - z_{m-1}\right)\right]$

and

a "backward" decaying "wave" $B = B_o \exp\left[\kappa_m \left(z - z_{m-1}\right)\right]$

where A and B are complex numbers for

the forward and backward "wave" amplitudes

In this case

$$\kappa_m = \sqrt{\frac{2m_{fm}}{\hbar^2}\left(V_m - E\right)}$$

where m_{fm} is the mass of the particle in that layer

Transfer matrix method

Note that for $E<V_o$ if we use the form $k_m = \sqrt{\frac{2m_{fm}}{\hbar^2}\left(E - V_m\right)}$

we obtain an imaginary k_m

As long as we choose the positive square root (either real or imaginary) in both cases

we can work with only this form

For example, a forward propagating "wave" can then be written in the form $\exp\left[ik_m\left(z - z_{m-1}\right)\right]$

for both cases

$E<V_o$ and $E>V_o$

and similarly for the backward propagating "wave"

Transfer matrix method

Now in any layer we have a wave that we can write as

$$\psi(z) = A_m \exp\left[ik_m\left(z - z_{m-1}\right)\right] + B_m \exp\left[-ik_m\left(z - z_{m-1}\right)\right]$$

where k_m can be either real or imaginary

and is given by

$$k_m = \sqrt{\frac{2m_{fm}}{\hbar^2}\left(E - V_m\right)}$$

This can greatly simplify the algebra for this method

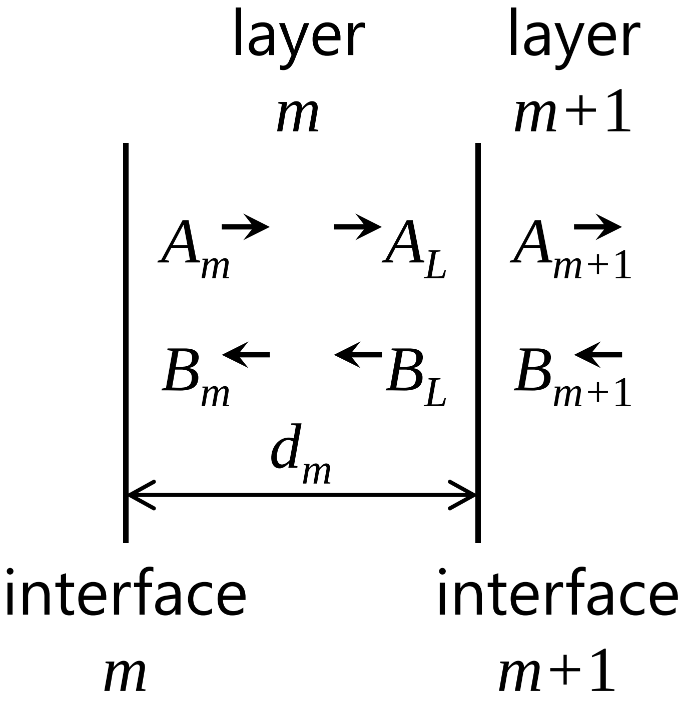
Boundary conditions

Now let us look at the boundary conditions in going from

- just inside one layer to the right of the boundary to
- just inside the adjacent layer on the left of the boundary

For a reason that will become apparent later

- we will work from right to left in setting these up



Boundary conditions

Using the notation of the figure

for continuity of the wavefunction

$$\psi = A_L + B_L = A_{m+1} + B_{m+1}$$

for the continuity of $d\psi / dz$

on either side of the boundary

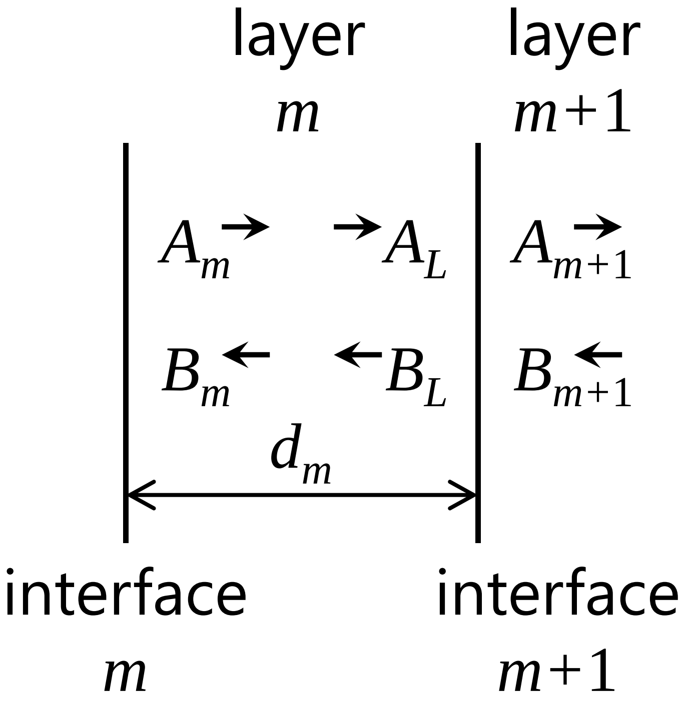
$$\frac{d\psi}{dz} = ik\left(A - B\right)$$

so at the right interface

$$A_L - B_L = \Delta_m \left(A_{m+1} - B_{m+1}\right)$$

where

$$\Delta_m = k_{m+1} / k_m$$



Boundary conditions

In a layered semiconductor structure

we might use continuity of

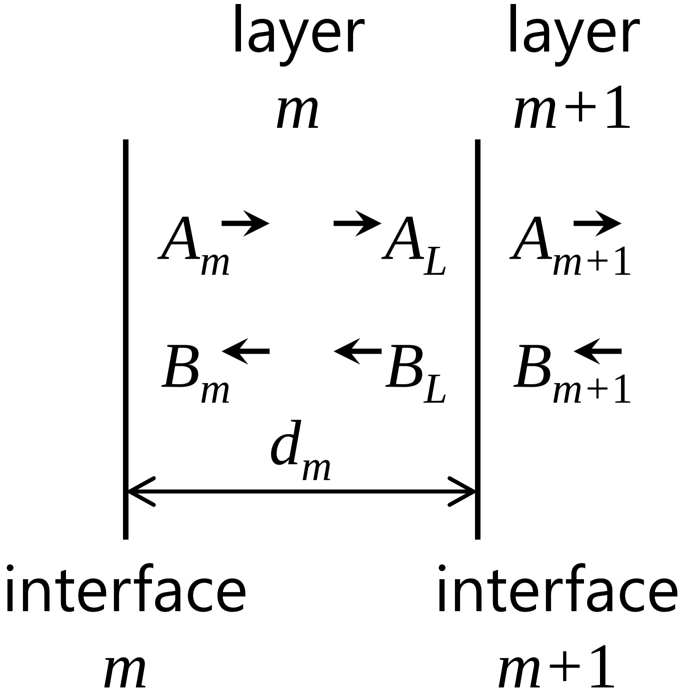
$$\left(1/m_f\right)d\psi/dz$$

for the second boundary condition

in which case we would obtain

$$\Delta_m = \frac{k_{m+1}}{k_m}\frac{m_{fm}}{m_{fm+1}}$$

and we would use this Δ_m in all subsequent algebra here



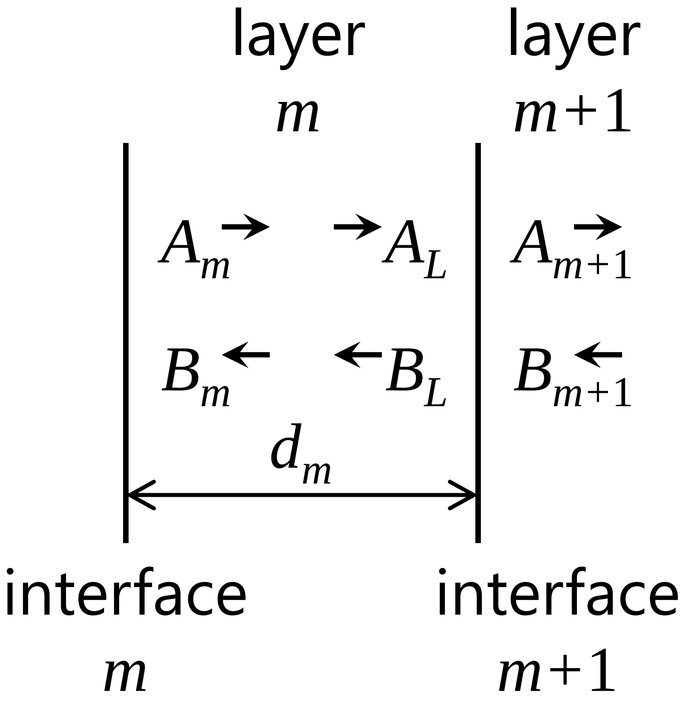
Boundary conditions

Using $A_L + B_L = A_{m+1} + B_{m+1}$ and

$$A_L - B_L = \Delta_m \left(A_{m+1} - B_{m+1} \right)$$

gives $A_L = A_{m+1}\left(\dfrac{1+\Delta_m}{2}\right) + B_{m+1}\left(\dfrac{1-\Delta_m}{2}\right)$

and $B_L = A_{m+1}\left(\dfrac{1-\Delta_m}{2}\right) + B_{m+1}\left(\dfrac{1+\Delta_m}{2}\right)$



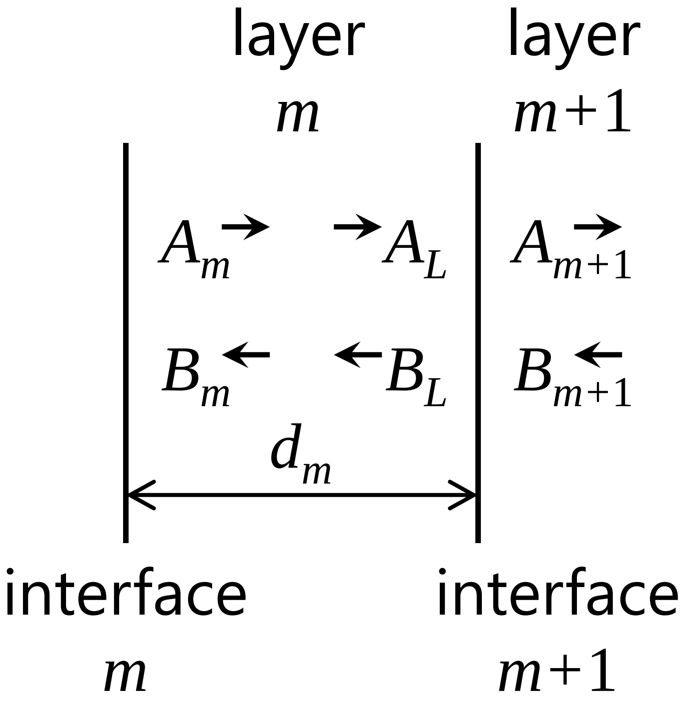
Boundary conditions

$$A_L = A_{m+1}\left(\frac{1+\Delta_m}{2}\right) + B_{m+1}\left(\frac{1-\Delta_m}{2}\right)$$

and $B_L = A_{m+1}\left(\frac{1-\Delta_m}{2}\right) + B_{m+1}\left(\frac{1+\Delta_m}{2}\right)$

can be written in matrix form as

$$\begin{bmatrix} A_L \\ B_L \end{bmatrix} = \mathbf{D}_m \begin{bmatrix} A_{m+1} \\ B_{m+1} \end{bmatrix} \qquad \mathbf{D}_m = \begin{bmatrix} \frac{1+\Delta_m}{2} & \frac{1-\Delta_m}{2} \\ \frac{1-\Delta_m}{2} & \frac{1+\Delta_m}{2} \end{bmatrix}$$



layer m

layer $m+1$

$A_m \rightarrow$ $\rightarrow A_L$ $A_{m+1} \rightarrow$

$B_m \leftarrow$ $\leftarrow B_L$ $B_{m+1} \leftarrow$

d_m

interface m

interface $m+1$

Propagation matrix

Now we treat the propagation that relates A_m and B_m to A_L and B_L

For a minor formal reason

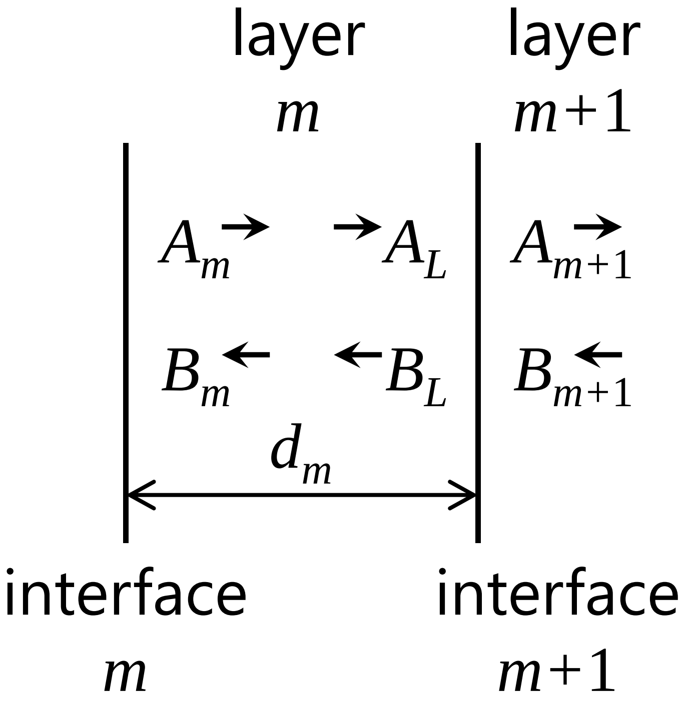
we calculate the matrices for going "backwards" through the structure

For the propagation in layer m

with thickness d_m, we have

$$A_m = A_L \exp\left(-ik_m d_m\right)$$

$$B_m = B_L \exp\left(ik_m d_m\right)$$



Propagation matrix

These relations $A_m = A_L \exp\left(-ik_m d_m\right)$

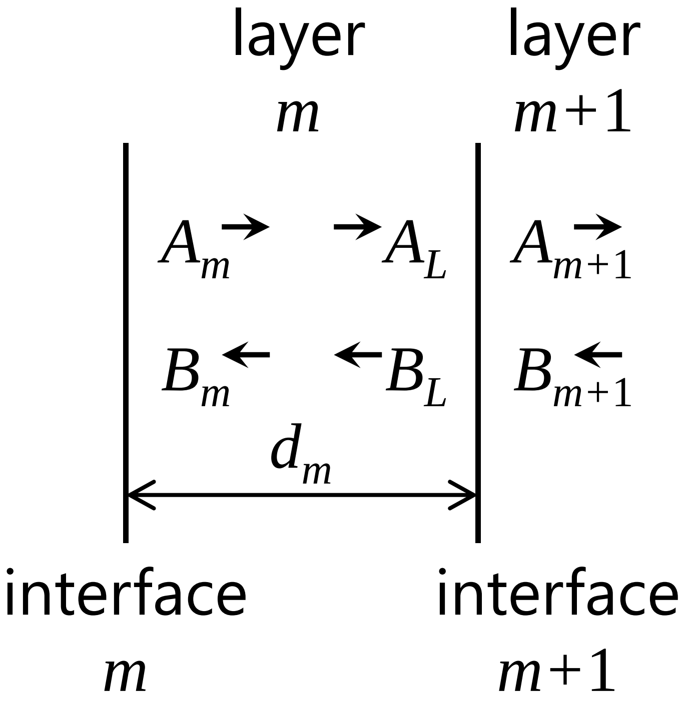
$$B_m = B_L \exp\left(ik_m d_m\right)$$

can be written in matrix form as

$$\begin{bmatrix} A_m \\ B_m \end{bmatrix} = \mathbf{P}_m \begin{bmatrix} A_L \\ B_L \end{bmatrix}$$

with

$$\mathbf{P}_m = \begin{bmatrix} \exp\left(-ik_m d_m\right) & 0 \\ 0 & \exp\left(ik_m d_m\right) \end{bmatrix}$$



Full transfer matrix

The full transfer matrix, $\mathbf{T}$, for the structure relates

forward, A_1, and backward, B_1, "entrance" amplitudes

i.e., just to the left of the first interface

to forward, A_{N+2}, and backward, B_{N+2}, "exit" amplitudes

i.e., just to the right of the last interface

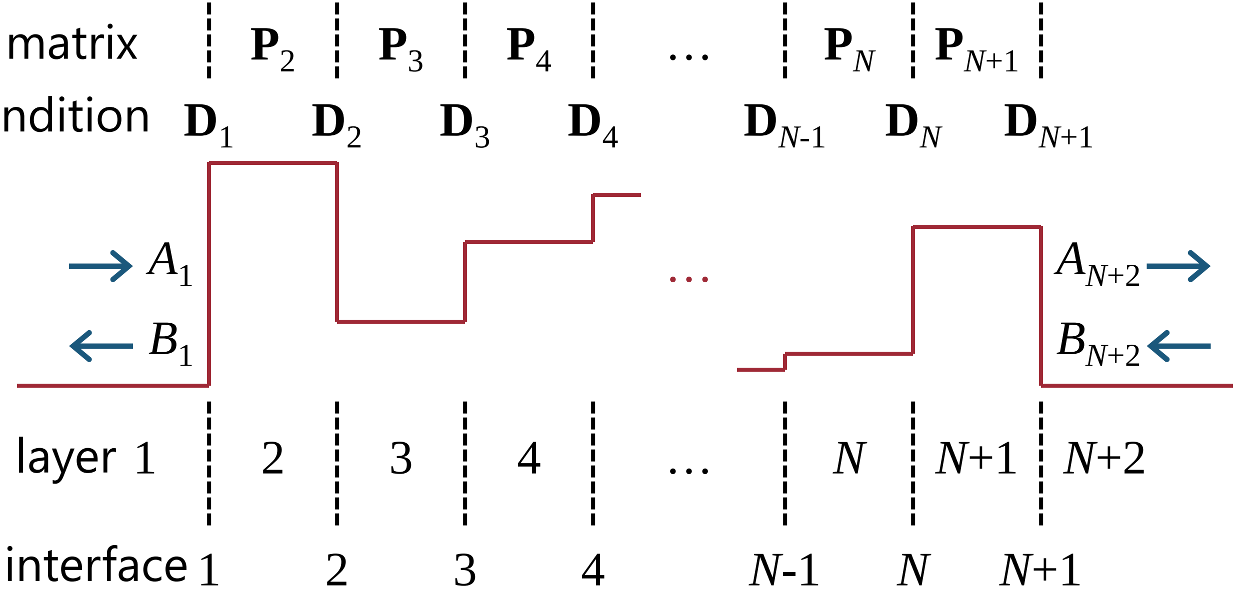
$$\begin{bmatrix} A_1 \\ B_1 \end{bmatrix} = \mathbf{T} \begin{bmatrix} A_{N+2} \\ B_{N+2} \end{bmatrix} \quad \text{where} \quad \mathbf{T} = \mathbf{D}_1 \mathbf{P}_2 \mathbf{D}_2 \mathbf{P}_3 \mathbf{D}_3 \cdots \mathbf{P}_{N+1} \mathbf{D}_{N+1}$$

Note that this transfer matrix depends on the energy E we chose for the calculation of the k's in each layer

Full transfer matrix

In $\begin{bmatrix}A_1\\B_1\end{bmatrix} = \mathbf{T}\begin{bmatrix}A_{N+2}\\B_{N+2}\end{bmatrix}$ and the product $\mathbf{T} = \mathbf{D}_1\mathbf{P}_2\mathbf{D}_2\mathbf{P}_3\mathbf{D}_3\cdots\mathbf{P}_{N+1}\mathbf{D}_{N+1}$

we move progressively from right to left



propagation matrix $\mathbf{P}_2$ $\mathbf{P}_3$ $\mathbf{P}_4$ … $\mathbf{P}_N$ $\mathbf{P}_{N+1}$

boundary condition matrix $\mathbf{D}_1$ $\mathbf{D}_2$ $\mathbf{D}_3$ $\mathbf{D}_4$ $\mathbf{D}_{N\text{-}1}$ $\mathbf{D}_N$ $\mathbf{D}_{N+1}$

layer 1 2 3 4 … N N+1 N+2

interface 1 2 3 4 N-1 N N+1

Calculation of tunneling rates

Having calculated the transfer matrix $\mathbf{T} \equiv \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$

for some structure and energy E

we now deduce the fraction of incident particles at that energy that are transmitted by the barrier

We presume no wave incident from the right, so

there is no backward wave amplitude on the right

Hence, for incident forward and backward amplitudes A and B respectively,

and a transmitted amplitude F

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} F \\ 0 \end{bmatrix}$$

Calculation of tunneling rates

From $\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} F \\ 0 \end{bmatrix}$ we see that $A = T_{11}F$ and $B = T_{21}F$

and hence the fraction of particles transmitted by this barrier is

$$\eta = \frac{|A|^2 - |B|^2}{|A|^2} = 1 - \frac{|T_{21}|^2}{|T_{11}|^2}$$

This approach is well suited for numerical calculations

being straightforward to program

It is a very useful practical technique for investigating one-dimensional potentials and their behavior

Calculation of wavefunctions

Note that this method also enables us to calculate the wavefunction at any point in the structure

We can readily calculate forward and backward amplitudes, A_m and B_m, at the left of each layer

Obviously, we have $\begin{bmatrix} A_{N+1} \\ B_{N+1} \end{bmatrix} = \mathbf{P}_{N+1}\mathbf{D}_{N+1}\begin{bmatrix} A_{N+2} \\ B_{N+2} \end{bmatrix}$

and similarly, we have in general for any layer within the structure

$$\begin{bmatrix} A_m \\ B_m \end{bmatrix} = \mathbf{P}_m\mathbf{D}_m \ldots \mathbf{P}_N\mathbf{D}_N\mathbf{P}_{N+1}\mathbf{D}_{N+1}\begin{bmatrix} A_{N+2} \\ B_{N+2} \end{bmatrix}$$

Calculation of wavefunctions

Given that we know the forward and backward amplitudes at the left of layer m from

$$\begin{bmatrix} A_m \\ B_m \end{bmatrix} = \mathbf{P}_m \mathbf{D}_m \ldots \mathbf{P}_N \mathbf{D}_N \mathbf{P}_{N+1} \mathbf{D}_{N+1} \begin{bmatrix} A_{N+2} \\ B_{N+2} \end{bmatrix}$$

then the wavefunction at some point z in that layer

is the sum of the forward and backward wavefunctions as in

$$\psi(z) = A_m \exp\left[ik_m\left(z - z_m\right)\right] + B_m \exp\left[-ik_m\left(z - z_m\right)\right]$$

Calculation of wavefunctions

Note that we could calculate these forward and backward amplitudes as intermediate results

if we progressively evaluate the forward and backward amplitudes for each successive layer as in

$$\begin{bmatrix} A_m \\ B_m \end{bmatrix} = \mathbf{P}_m \mathbf{D}_m \begin{bmatrix} A_{m+1} \\ B_{m+1} \end{bmatrix}$$

rather than evaluating the transfer matrix $\mathbf{T}$ itself

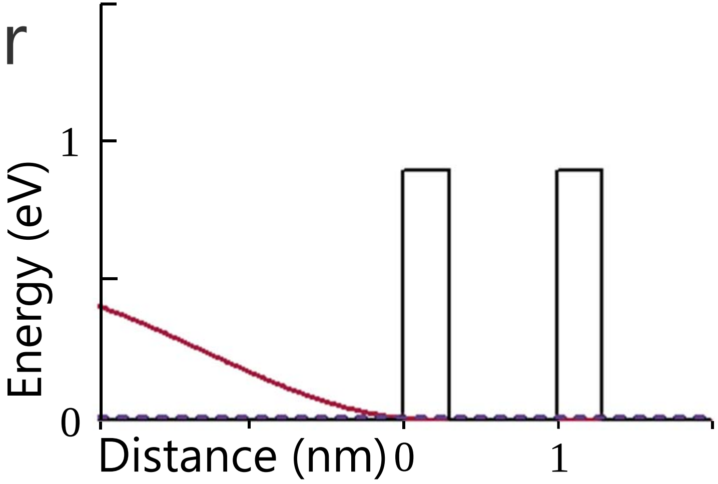
By choosing no inward wave on the right

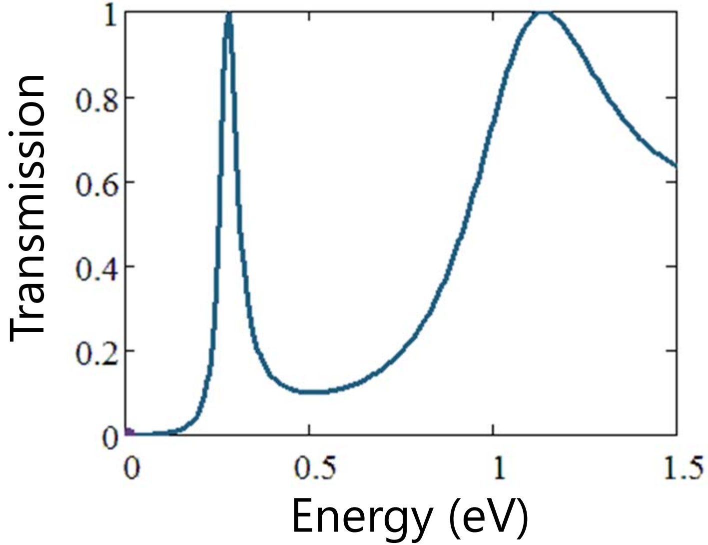
we can still calculate the transmission probability from

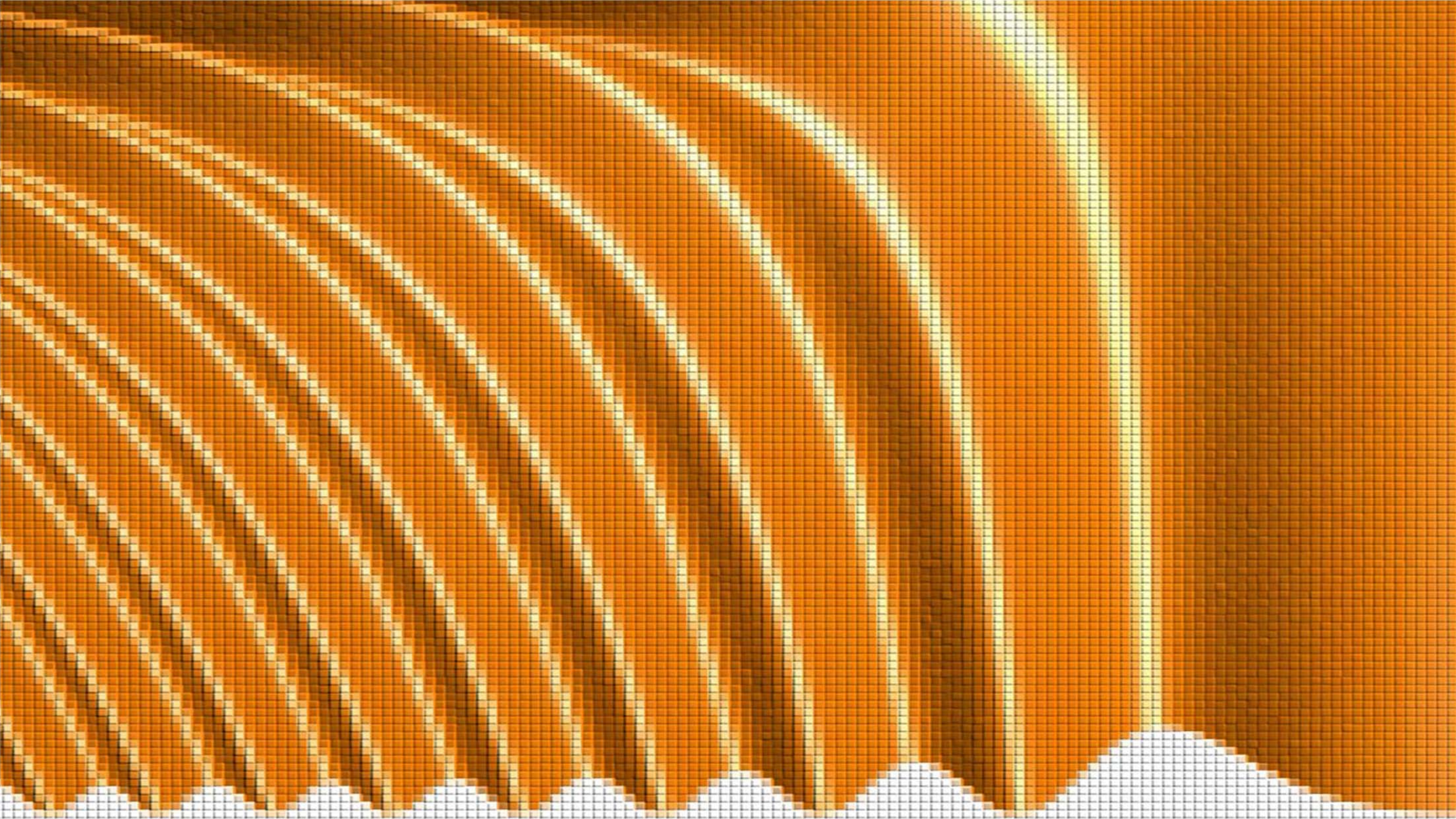
$$\eta = \left(|A_1|^2 - |B_1|^2 \right) / |A_1|^2$$

Tunneling through a double barrier

- This structure shows a resonance in the tunneling probability (or transmission)
  - where the incident energy coincides with the energy of a resonance in the structure
- If the barriers were infinitely thick
  - there would be an eigenstate approximately at the energy where the first resonance occurs





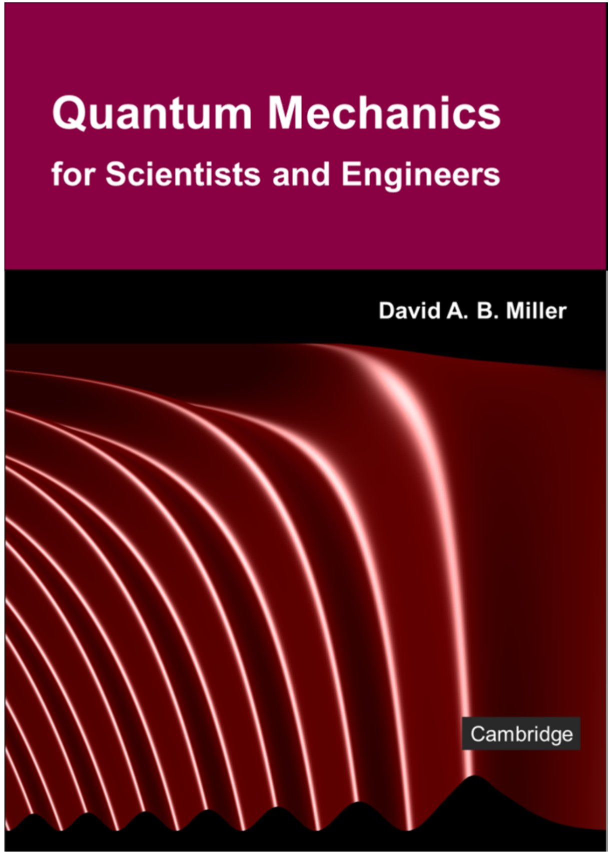


12.2 Methods for one-dimensional problems

Slides: Video 12.2.6 Transfer matrix and bound states

Text reference: Quantum Mechanics for Scientists and Engineers

Section 11.2 from "Calculation of eigenenergies …"



Methods for one-dimensional problems

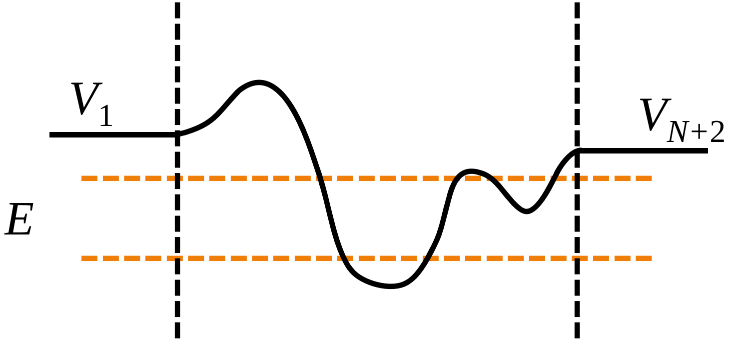
Transfer matrix and bound states

Quantum mechanics for scientists and engineers

David Miller

Eigenenergies of bound states

- It is possible to use the transfer matrix itself to find eigenstates in cases of truly bound states
- For example, if
  - the first layer (layer 1) and last layer (layer N+2) are infinitely thick, and
    - have potentials V_1 and V_{N+2}
- there may be values of $E < V_1, V_{N+2}$
  - for which there are bound eigenstates



Eigenenergies of bound states

The wavefunctions would be

exponentially decaying

into the first and last layers

So the forward amplitude on the left

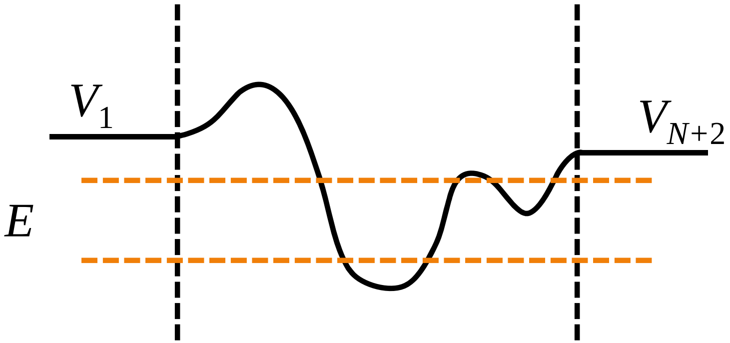
$$A_1 = 0$$

i.e., no exponentially growing wave to the left of the structure and

the backward amplitude on the right

$$B_{N+2} = 0$$

i.e., no exponentially growing wave to the right of the structure



Eigenenergies of bound states

So for a bound eigenstate, we have

$$\begin{bmatrix} 0 \\ B_1 \end{bmatrix} = \mathbf{T} \begin{bmatrix} A_{N+2} \\ 0 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} A_{N+2} \\ 0 \end{bmatrix}$$

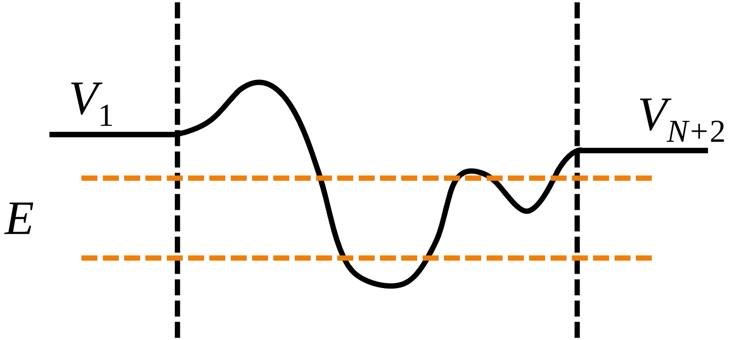
This can only be the case if

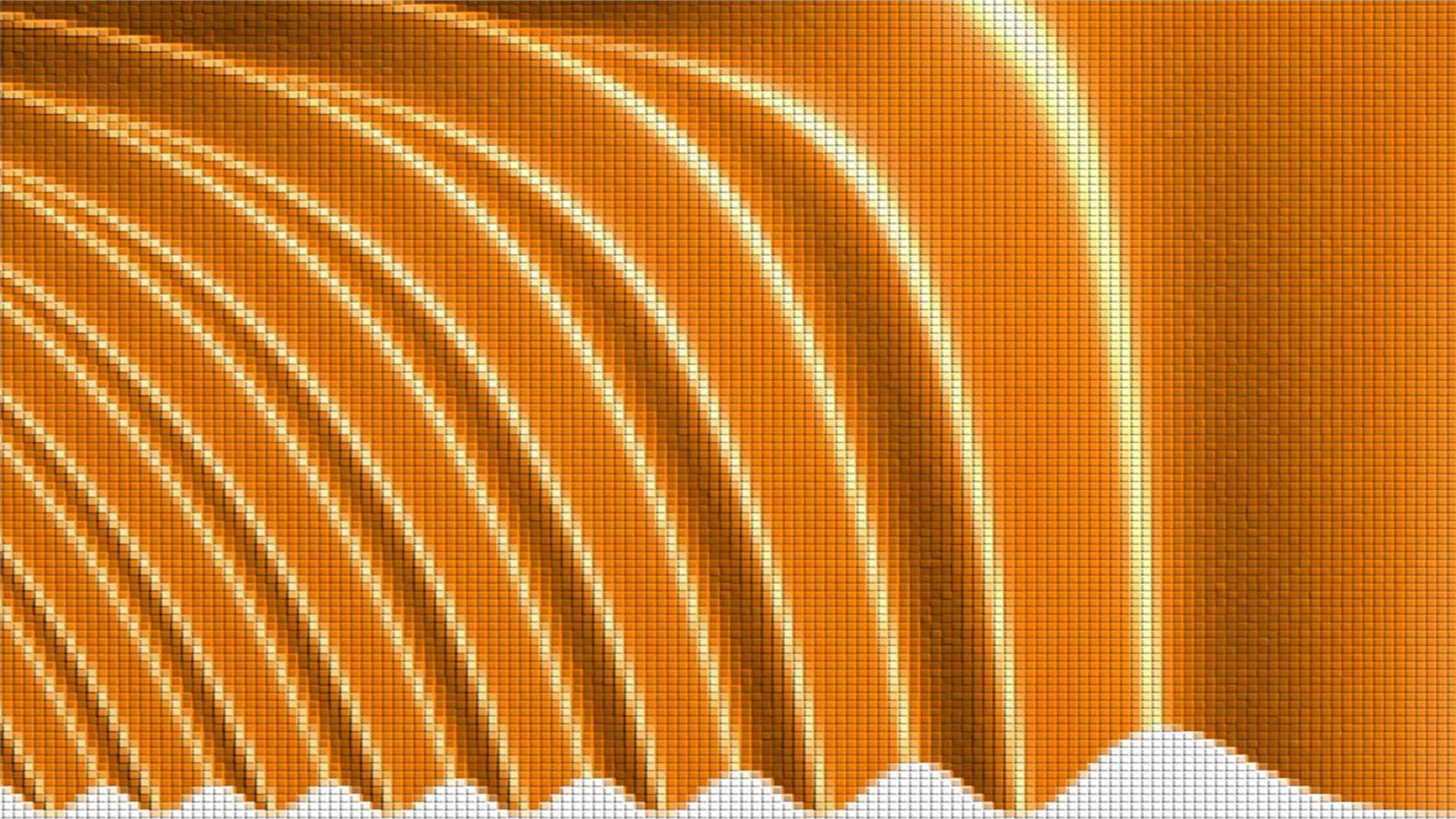
$$T_{11} = 0$$

This condition can be used

to solve analytically for eigenenergies in simple structures

or in a numerical search for eigenenergies by varying E



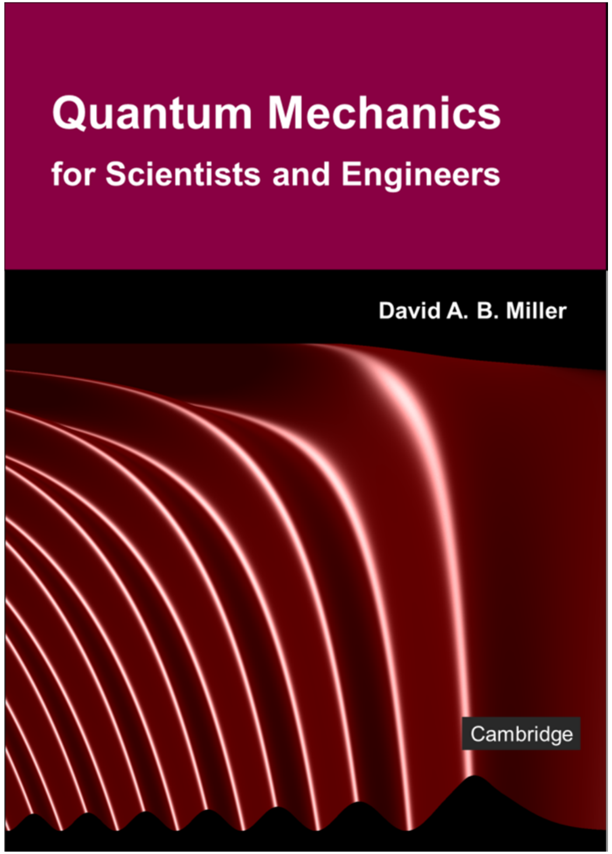


12.2 Methods for one-dimensional problems

Slides: Video 12.2.8 Penetration factor for slowly varying barriers

Text reference: Quantum Mechanics for Scientists and Engineers

Section 11.3



Methods for one-dimensional problems

Penetration factor for slowly varying barriers

Quantum mechanics for scientists and engineers

David Miller

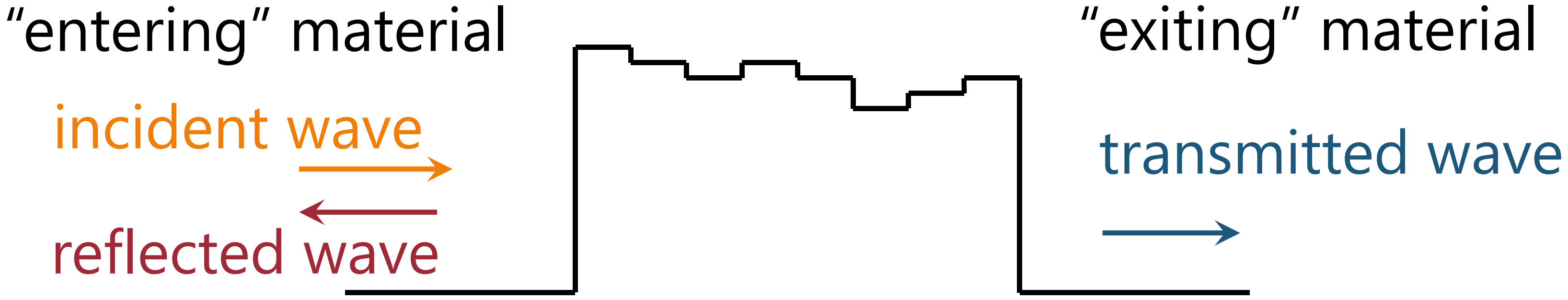
Penetration factor for slowly varying barriers

Consider a slowly varying potential

approximated as a series of steps

For simplicity

we choose the "entering" and "exiting" materials as having the same energy



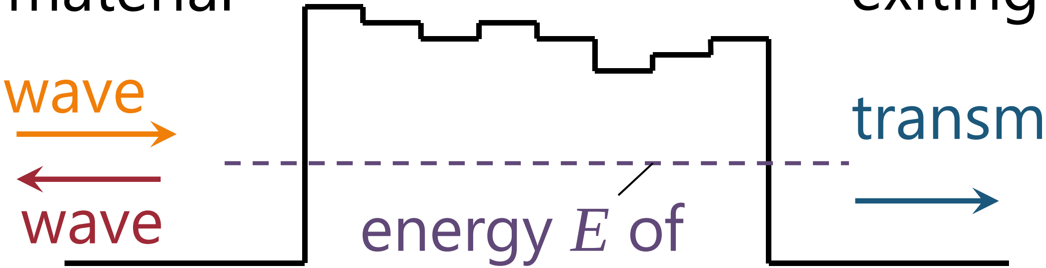
Penetration factor for slowly varying barriers

We presume for an energy E of interest

$E \ll V_m$ for each layer inside the structure, and that

we have chosen the layers sufficiently thin in our calculation so that

at least for interfaces within the structure, $k_m \simeq k_{m+1}$



“entering” material

incident wave

reflected wave

“exiting” material

transmitted wave

energy E of interest

Penetration factor for slowly varying barriers

Then, for interfaces within the structure

the boundary condition matrix

$$\mathbf{D}_m = \begin{bmatrix} \dfrac{1+\Delta_m}{2} & \dfrac{1-\Delta_m}{2} \\ \dfrac{1-\Delta_m}{2} & \dfrac{1+\Delta_m}{2} \end{bmatrix}$$ with $\Delta_m = \dfrac{k_{m+1}}{k_m} \simeq 1$ by assumption

can be approximated as the identity matrix

$$D_m \cong \begin{bmatrix} \sim 1 & \sim 0 \\ \sim 0 & \sim 1 \end{bmatrix}$$

Penetration factor for slowly varying barriers

With internal boundary condition matrices therefore approximated by identity matrices

we can omit them, so the transfer matrix becomes

$$\mathbf{T} = \mathbf{D}_1 \mathbf{P}_2 \mathbf{P}_3 \cdots \mathbf{P}_N \mathbf{P}_{N+1} \mathbf{D}_{N+1}$$

We have left in the boundary condition matrices for the beginning and end of the structure

where the potential may be quite discontinuous

Product of diagonal matrices

Note that the product of two diagonal matrices

is simply a diagonal matrix

whose elements are the products of the corresponding diagonal elements

For example

$$\begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix} \begin{bmatrix} b & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} ab & 0 \\ 0 & cd \end{bmatrix}$$

Penetration factor for slowly varying barriers

Since the propagation matrices are all diagonal

and previously shown to be of the form

$$\mathbf{P}_m = \begin{bmatrix} \exp(-ik_m d_m) & 0 \\ 0 & \exp(ik_m d_m) \end{bmatrix}$$

then

$$\mathbf{P}_2\mathbf{P}_3 \cdots \mathbf{P}_N\mathbf{P}_{N+1} = \begin{bmatrix} 1/G & \sim 0 \\ \sim 0 & G \end{bmatrix}$$

where

$$G = \prod_{q=2}^{N+1} \exp(ik_q d_q) = \prod_{q=2}^{N+1} \exp(-\kappa_q d_q) = \exp\left(-\sum_{q=2}^{N+1} \kappa_q d_q \right)$$

Penetration factor for slowly varying barriers

Now, if we have chosen the layers to be sufficiently thin

we may take the summation to be approximately equal

to an integral, i.e., $\sum_{q=2}^{N+1} \kappa_q d_q \cong \int_0^{z_{tot}} \kappa(z)\, dz$

where z_{tot} $(= z_{N+1})$ is the total structure thickness

which is taken to start on the left at $z = 0$

Hence

$$G \simeq \exp\left(-\int_0^{z_{tot}} \kappa(z)\, dz \right) = \exp\left(-\int_0^{z_{tot}} \sqrt{\frac{2m_f}{\hbar^2}\left(V(z) - E\right)}\, dz \right)$$

where $V(z)$ is the potential as a function of position z

Penetration factor for slowly varying barriers

With first and last boundary condition matrices

$$\mathbf{D}_1 = \begin{bmatrix} \frac{1+\Delta_1}{2} & \frac{1-\Delta_1}{2} \\ \frac{1-\Delta_1}{2} & \frac{1+\Delta_1}{2} \end{bmatrix} \quad \text{and} \quad \mathbf{D}_{N+1} = \begin{bmatrix} \frac{1+\Delta_{N+1}}{2} & \frac{1-\Delta_{N+1}}{2} \\ \frac{1-\Delta_{N+1}}{2} & \frac{1+\Delta_{N+1}}{2} \end{bmatrix}$$

then $\mathbf{T} = \mathbf{D}_1 \mathbf{P}_2 \mathbf{P}_3 \cdots \mathbf{P}_N \mathbf{P}_{N+1} \mathbf{D}_{N+1}$

$$= \begin{bmatrix} \frac{1+\Delta_1}{2} & \frac{1-\Delta_1}{2} \\ \frac{1-\Delta_1}{2} & \frac{1+\Delta_1}{2} \end{bmatrix} \begin{bmatrix} 1/G & \sim 0 \\ \sim 0 & G \end{bmatrix} \begin{bmatrix} \frac{1+\Delta_{N+1}}{2} & \frac{1-\Delta_{N+1}}{2} \\ \frac{1-\Delta_{N+1}}{2} & \frac{1+\Delta_{N+1}}{2} \end{bmatrix}$$

Penetration factor for slowly varying barriers

So $$\mathbf{T}=\begin{bmatrix}\frac{1+\Delta_1}{2} & \frac{1-\Delta_1}{2}\\ \frac{1-\Delta_1}{2} & \frac{1+\Delta_1}{2}\end{bmatrix}\begin{bmatrix}1/G & \sim 0\\ \sim 0 & G\end{bmatrix}\begin{bmatrix}\frac{1+\Delta_{N+1}}{2} & \frac{1-\Delta_{N+1}}{2}\\ \frac{1-\Delta_{N+1}}{2} & \frac{1+\Delta_{N+1}}{2}\end{bmatrix}$$

$$\mathbf{T}\simeq\begin{bmatrix}\frac{1+\Delta_1}{2} & \frac{1-\Delta_1}{2}\\ \frac{1-\Delta_1}{2} & \frac{1+\Delta_1}{2}\end{bmatrix}\begin{bmatrix}\left(\frac{1+\Delta_{N+1}}{2}\right)\frac{1}{G} & \left(\frac{1-\Delta_{N+1}}{2}\right)\frac{1}{G}\\ \left(\frac{1-\Delta_{N+1}}{2}\right)G & \left(\frac{1+\Delta_{N+1}}{2}\right)G\end{bmatrix}$$

$$T_{11}\simeq\left(\frac{1+\Delta_1}{2}\right)\left(\frac{1+\Delta_{N+1}}{2}\right)\frac{1}{G}+\left(\frac{1-\Delta_1}{2}\right)\left(\frac{1-\Delta_{N+1}}{2}\right)G$$

Penetration factor for slowly varying barriers

Since the barrier is presumed thick $G \simeq \exp\left(-\int_0^{z_{tot}} \kappa(z)\,dz\right)$

is presumed small, so

$$T_{11} \simeq \left(\frac{1+\Delta_1}{2}\right)\left(\frac{1+\Delta_{N+1}}{2}\right)\frac{1}{G} + \left(\frac{1-\Delta_1}{2}\right)\left(\frac{1-\Delta_{N+1}}{2}\right)G \simeq \frac{(1+\Delta_1)(1+\Delta_{N+1})}{4G}$$

From $\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}\begin{bmatrix} F \\ 0 \end{bmatrix}$, $A = T_{11}F$ so $\eta = |F|^2 / |A|^2 = 1/|T_{11}|^2$

so

$$\boxed{\eta \simeq \frac{16}{(1+\Delta_1)(1+\Delta_{N+1})}\exp\left(-2\int_0^{z_{tot}} \sqrt{\frac{2m_f}{\hbar^2}(V(z)-E)}\,dz\right)}$$

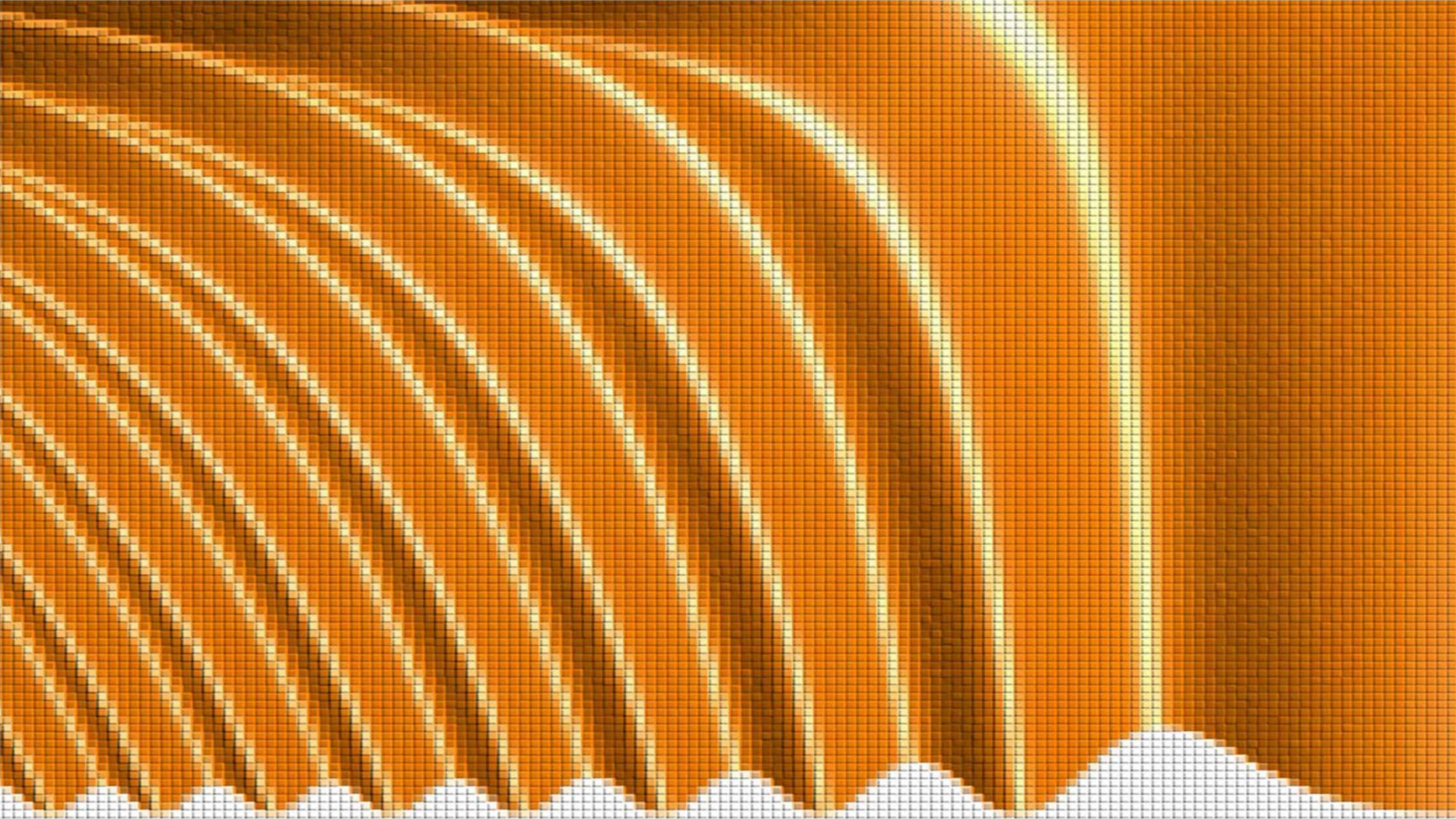
Penetration factor for slowly varying barriers

In this frequently used expression for tunneling probability or "penetration factor"

$$\eta \simeq \frac{16}{\left(1+\Delta_1\right)\left(1+\Delta_{N+1}\right)} \exp\left(-2\int_0^{z_{tot}} \sqrt{\frac{2m_f}{\hbar^2}\left(V\left(z\right)-E\right)}dz\right)$$

the prefactor contains the input and output boundary conditions and

the exponential approximately expresses the "penetration" within the barrier



Transfer Matrix Calculation

David A. B. Miller, January 2015

The following sheet is technically a program written for the Mathcad software. In its original form, it is a live sheet that will recalculate if you change the numbers, for example. The Mathcad software is particularly good for creating active sheets that can be laid out in relatively conventional and intelligible document-like fashions. Blocks such as this one are simply "inactive" text that are ignored from the point of view of calculation. Simple non-executable graphics objects can also be inserted.

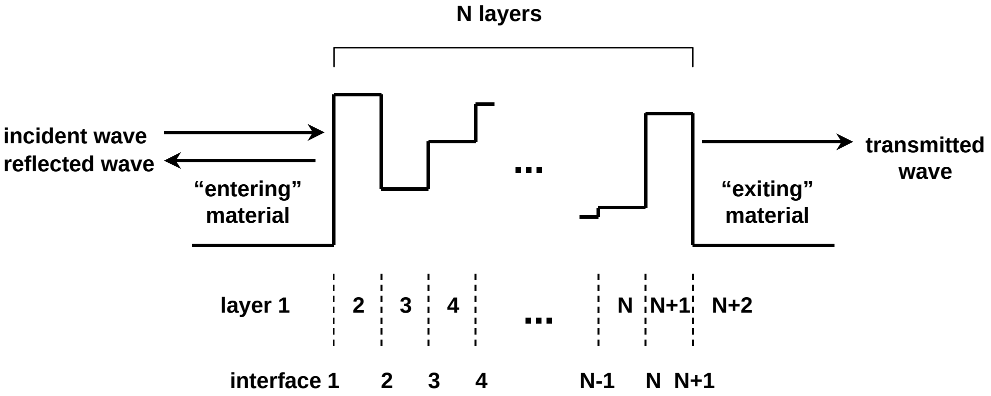
The syntax of Mathcad is mostly fairly self-evident. The characters ":=" mean that what is on the left is assigned to be what is on the right. The "=" character on its own causes the program to write out on the right the value or values of what is on the left.

Now we can start the program proper, which is designed to plot out the behavior of a double barrier "resonant tunneling" structure as a function of the incident electron energy.

This first "ORIGIN:=1" statement below formally sets the "origin" of all arrays to start indexing from 1 rather than 0.

ORIGIN := 1

We wish to calculate the transfer matrix for a structure with a series of steps of potential as shown in the figure.



Formal construction of matrices

We first define the necessary fundamental constants.

$\mathrm{hbar} := 1.055\cdot 10^{-34}$ $\mathrm{mo} := 9.1095\cdot 10^{-31}$ $q := 1.602\cdot 10^{-19}$

For a given layer *m* of potential energy V_m, mass m_{fm}, and thickness d_m, we can define the necessary quantities required by the algebra. To allow the use of mass units of the free electron mass m_o, thickness units of nanometers, and energy units of electron volts, we define a units scaling parameter *s* by

$s := \dfrac{2q\cdot mo\cdot 10^{-18}}{hbar^2}$ which leads to the numerical value $s = 26.223$

so we can define a function for calculating the wavevector (which may be real or imaginary) for a given layer, formally as a function of the energy E of interest (in elecron volts), the potential Vm in a layer (in electron volts), and the effective mass mfm (in units of the free electron mass)

$$k(E, Vm, mfm) := \sqrt{s\cdot mfm\cdot(E - Vm)}$$

We can also formally define a function to calculate the quantity Δ that comes from the algebra to set up the transfer matrix method

$$\Delta(E, Vm, mfm, Vm1, mfm1) := k(E, Vm1, mfm1)\cdot\frac{mfm}{k(E, Vm, mfm)\cdot mfm1}$$

where by mfm we mean m_{fm} and by mfm1 we mean m_{fm+1}, i.e., the quantity in the layer m+1, and similarly for *Vm* and *Vm1*.

This leads to a boundary condition matrix, also formally defined as a function of these parameters

$$D(E, Vm, mfm, Vm1, mfm1) := \frac{\begin{pmatrix} 1 + \Delta(E, Vm, mfm, Vm1, mfm1) & 1 - \Delta(E, Vm, mfm, Vm1, mfm1) \\ 1 - \Delta(E, Vm, mfm, Vm1, mfm1) & 1 + \Delta(E, Vm, mfm, Vm1, mfm1) \end{pmatrix}}{2}$$

relating the forward and backward amplitudes just inside the right side of layer *m* to those just inside the layer *m*+1.

We can also define the propagation matrix in layer *m* that also comes from the transfer matrix algebra, again as a function of the various parameters, including now the thickness dm of the layer of interest.

$$P(E, Vm, mfm, dm) := \begin{pmatrix} \exp(-i\cdot k(E, Vm, mfm)\cdot dm) & 0 \\ 0 & \exp(i\cdot k(E, Vm, mfm)\cdot dm) \end{pmatrix}$$

Having set up all these functions, for a given structure, we will have to choose appropriate parameters and proceed to calculate the tranfer matrix for a given energy E of interest.

Choice of parameters

First, we choose the number *N* of layers in the structure (not including the "entering" and "exiting" layers)

$N := 3$

Now we explicitly input the values of the parameters. For each layer, we have to choose the potential V_m, the mass m_{fm}, and the thickness d_m. As stated before, we will use mass units of the free electron mass m_o, thickness units of nanometers, and energy units of electron volts for inputting the parameters.

$mf_1 := 1 \qquad Vm_1 := 0$

$mf_2 := 1 \qquad Vm_2 := 0.9 \qquad dm_2 := 0.3$

$mf_3 := 1 \qquad Vm_3 := 0 \qquad dm_3 := .7$

$mf_4 := 1 \qquad Vm_4 := 0.9 \qquad dm_4 := 0.3$

$mf_5 := 1 \qquad Vm_5 := 0$

Now we can formally construct the overall transfer matrix by multiplying the various constituent matrices. Incidentally, when setting such a matrix multiplication up in some software, do make sure you understand what order the matrices are being multiplied in. Checking this with some explicit examples may be worthwhile. It is not always obvious otherwise what a specific piece of software will actually do.

$$T(E) := D\left(E, Vm_1, mf_1, Vm_2, mf_2\right) \cdot \prod_{q=2}^{N+1} \left(P\left(E, Vm_q, mf_q, dm_q\right) \cdot D\left(E, Vm_q, mf_q, Vm_{q+1}, mf_{q+1}\right)\right)$$

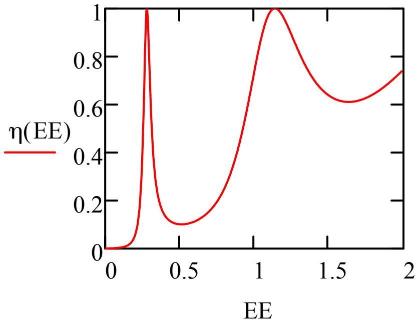
and we can define the transmission fraction by

$$\eta(E) := 1 - \frac{\left(\left|T(E)_{2,1}\right|\right)^2}{\left(\left|T(E)_{1,1}\right|\right)^2}$$

To plot up the resulting transmission for a range of values of energy, we can define a "range variable" EE

$EE := .0025, .0075 \, .. \, 1.9875$

Then using built-in capabilities of the Mathcad program, we can have it plot up for us the transmission fraction as a function of that range variable.



Graphing the probability density

Now that we have completed the core calculation method for the transfer matrix, we can go on to plot out the resulting probability densities. Mathcad handles animations like this using a built-in variable FRAME. When creating an animation, at "run time" the user selects an area of the sheet that they want to appear in the video animation, and then chooses a range of integer values for this FRAME variable. The program will then recalculate the sheet for each value of FRAME, saving the selected area each time as a frame in the final video. So this FRAME variable appears here in the formula that sets the energy value E to be used for the various frames. (Until "run time", FRAME = 0)

Here we avoid setting E equal to zero, which can sometimes cause problems with functions, starting at the small value 0.01. Then we increment the energy by 0.0125 for every successive frame of the final video. With 120 frames, this set of values below would run E up to about 1.5 eV.

We first formally choose the energy for which we wish to graph the probability density.

$$EG := 0.01 + FRAME\cdot 0.0125$$

We next evaluate the forward amplitude in layer 1 (the "entering" layer) on the assumption of unit forward amplitude (and no backward amplitude) in the exiting layer.

$$FB1 := T(EG)\cdot\begin{pmatrix}1\\0\end{pmatrix}$$

Just for interest, we can see what this pair of amplitudes is for the current value of EG. The upper value is the forward amplitude and the lower value is the backward amplitude.

$$FB1 = \begin{pmatrix}23.326 + 100.991i\\-103.645i\end{pmatrix}$$

We will capture the forward amplitude to use it as a number for renormalizing the results.

$$Renorm := FB1_1$$

For example, here the value is

$$Renorm = 23.326 + 100.991i$$

We will use this to normalize all other results to a unit forward amplitude in layer 1. For future use, we start by constructing a normalized version of this first amplitude itself.

$$FB1norm := \frac{FB1}{Renorm} \qquad FB1norm = \begin{pmatrix}1\\-0.974 - 0.225i\end{pmatrix}$$

Now we construct a function that gives us the forward and backward amplitudes on the left of any layer with index greater than 1. Formally, this function is multiplying progressively from the right "back" to the interface of interest, renormlaizing the result to correspond to unit input amplitude.

$$FB(p) := \frac{\left[\left[\prod_{q=p}^{N+1}\left(P\left(EG, Vm_q, mf_q, dm_q\right)\cdot D\left(EG, Vm_q, mf_q, Vm_{q+1}, mf_{q+1}\right)\right)\right]\cdot\begin{pmatrix}1\\0\end{pmatrix}\right]}{Renorm}$$

Now we construct a vector that gives us the distance from the left of the structure (i.e., from the position of the first interface on the left) at the left of any layer of the structure. For reasons of mathematical convenience in this formula, we first define dm1=0, though we do not in fact restrict this layer to zero thickness.

$$dm_1 := 0$$

$$dp(p) := \sum_{q=1}^{p-1} dm_q$$

Now we construct a function that tells us, for any given position z what layer of the structure we are in. (In Mathcad, functions based on programs require a specific form of syntax, though again, the should be relatively self-evident to read if not to program.)

```
pp(z) := | 1  if  z < 0
         | otherwise
         |   | (N + 2)  if  z ≥ dp(N + 2)
         |   | otherwise
         |   |   | qq ← 2
         |   |   | while  dp(qq) < z
         |   |   |   qq ← qq + 1
         |   |   | qq − 1
```

Now we construct a function that gives us the wavefunction at any point z in the structure. Note that, at any given point, the wavefunction is the sum of the forward and backward propagating parts.

$$\psi(z) := \left| \begin{array}{l} \text{FB1norm}_1 \cdot \exp\left(i \cdot k\left(EG, Vm_1, mf_1\right) \cdot z\right) + \text{FB1norm}_2 \cdot \exp\left[-i \cdot \left(k\left(EG, Vm_1, mf_1\right) \cdot z\right)\right] \text{ if } z < 0 \\ \text{otherwise} \\ \quad \left| \begin{array}{l} \left(\dfrac{1}{\text{Renorm}}\right) \cdot \exp\left[i \cdot k\left(EG, Vm_{N+2}, mf_{N+2}\right) \cdot (z - dp(N+2))\right] \text{ if } z > dp(N+2) \\ \text{otherwise} \\ \quad \left| \begin{array}{l} jj \leftarrow pp(z) \\ kz \leftarrow k\left(EG, Vm_{jj}, mf_{jj}\right) \cdot (z - dp(jj)) \\ FB(jj)_1 \cdot \exp(i \cdot kz) + FB(jj)_2 \cdot \exp(-i \cdot kz) \end{array} \right. \end{array} \right. \end{array} \right.$$

Hence we have for the probability density

$$P(z) := \overline{\psi(z)} \cdot \psi(z)$$

Now we can graph the probability density. We define minimum and maximum values of position z for plotting, and an appropriate range variable. We extend z to a negative range so we can see the standing waves to the left of the input.

$zmin := -2$ $\qquad$ $zmax := 2$

$zz := zmin, zmin + 0.01 .. zmax$

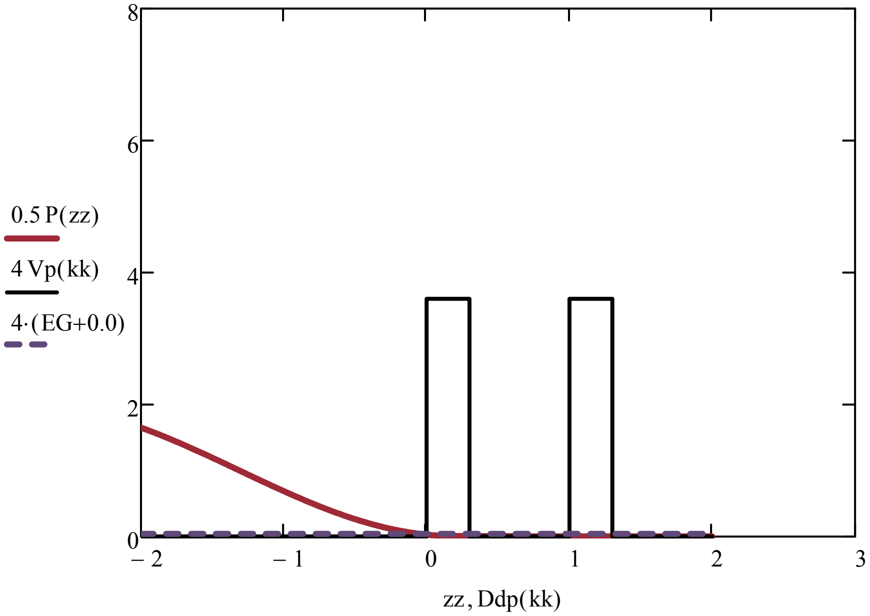
We also prepare mathematical vectors containing the potential values and the positions of the interfaces in a form suitable for plotting the structure itself. These employ a simple mathematical trick involving the "ceiling" function and a "doubled" range of the variable k so that we can directly plot the potential structure for graphics purposes.

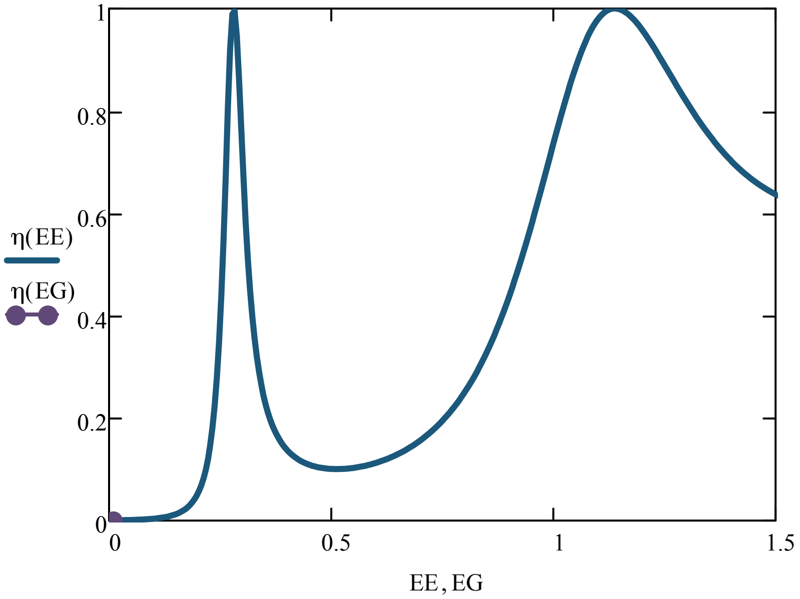
$$Vp(kk) := \left| \begin{array}{l} Vm_1 \text{ if } kk < 1 \\ \text{otherwise} \\ \quad \left| \begin{array}{l} Vm_{N+2} \text{ if } kk > 2 \cdot N + 2 \\ Vm_{\text{ceil}\left(\frac{kk}{2}\right)} \text{ otherwise} \end{array} \right. \end{array} \right.$$

$$Ddp(kk) := \left| \begin{array}{l} zmin \text{ if } kk < 2 \\ \text{otherwise} \\ \quad \left| \begin{array}{l} zmax \text{ if } kk > 2 \cdot N + 3 \\ dp\left(\text{ceil}\left(\dfrac{kk+1}{2}\right)\right) \text{ otherwise} \end{array} \right. \end{array} \right.$$

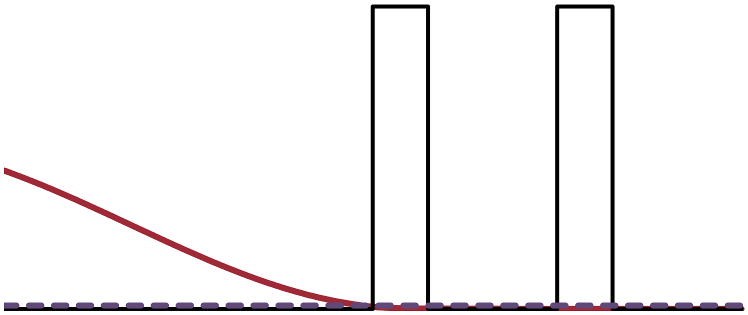
$kk := 1 .. 2N + 4$

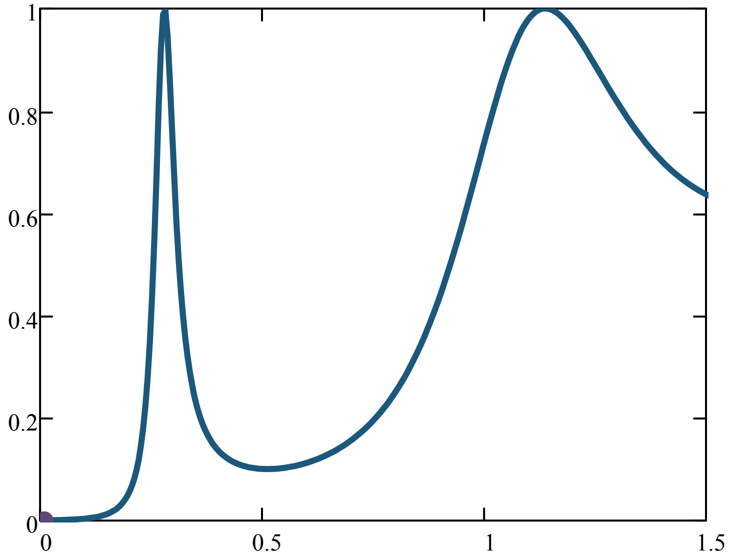
In the graphs below, we have shown the axes and the arguments to show exactly what is being plotted. For final animations, these can be suppressed (see the graphics on the final page). When running the actual animation at "run time" the FRAME variable is incremented to generate a progression of different graphs below, each of which is recorded as a frame in the overall animation.





The graphics versions below have the some of the explicit information suppressed for clearer graphic presentation for animations.





1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5

Contents

- Define Constants
- Outermost loop through all "m" energy values
- Inner Loop, Step 1: Initialize transmission boundary condition matrix for first layer
- Inner Loop through all remaining layers "n" of the structure (2 through final-1)
- Calculate transmission fraction given the resulting transmission matrix for energy "m"
- Create transmission-vs-energy plot
- Function to compute the boundary condition matrix between layers "n" and "n+1", using "delta" for layer n
- Function to compute the propagation matrix for layer "n" using the potential, effective mass, and thickness of this layer

```
function TransmissionVsEnergyPlot(MF, VM, DM, energy)
```

```
if nargin < 4   %Define defaults if no inputs to the function given
    MF = [1,1,1,1,1];              %Effective masses in each layer
    VM = [0,0.9,0,0.9,0];          %Potential in each layer
    DM = [0,0.3,0.7,0.3,0];        %Thickness of each layer
                                   %(note: first and last values
                                   %for infinite layers are not used)
    energy = linspace(0,2,50);     %Vector of energies of interest
end

tfrac = zeros(1,length(energy));  %initialize transmission fraction
```

Define Constants

```
hbar = 1.055*10^-34;               %reduced Planck constant
mo = 9.1095*10^-31;                %electron mass
q = 1.602*10^-19;                  %electron charge
s = (2*q*mo*10^-18)/(hbar^2);      %useful parameter
```

Outermost loop through all "m" energy values

```
for m = 1:length(energy)
```

Inner Loop, Step 1: Initialize transmission boundary condition matrix for first layer

```
    k1 = sqrt(s*MF(1)*(energy(m)-VM(1)));
    %Calculate "k" for layer "1"
    k2 = sqrt(s*MF(2)*(energy(m)-VM(2)));
```

```
    %Calculate "k" for layer "2"
    delta1 = (k2/k1)*(MF(1)/MF(2));
    %Calculate "delta" for layer "1"
    trans = Calculate_Boundary_Condition_Matrix(delta1);
    %Begin to construct transmission matrix by creating boundary
    %condition matrix between layer 1 and 2
```

Inner Loop through all remaining layers "n" of the structure (2 through final-1)

```
    for n = 2:length(MF)-1
    %Create the propagation matrix and boundary condition matrix
    %for each layer and multiply to update the transmission matrix
        kn = sqrt(s*MF(n)*(energy(m)-VM(n)));
        %Calculate "k" for layer "n"
        kn1 = sqrt(s*MF(n+1)*(energy(m)-VM(n+1)));
        %Calculate "k" for next layer
        deltan = (kn1/kn)*(MF(n)/MF(n+1));
        %Calculate "delta" for layer "n"
        LayerMatrixn = Calculate_Propagation_Matrix(kn,DM(n))  ...
            *Calculate_Boundary_Condition_Matrix(deltan);
        trans = trans*LayerMatrixn;
        %Multiply to the running matrix product "trans"
        %the propagation and boundary condition matrix
        %for layer "n"
    end
```

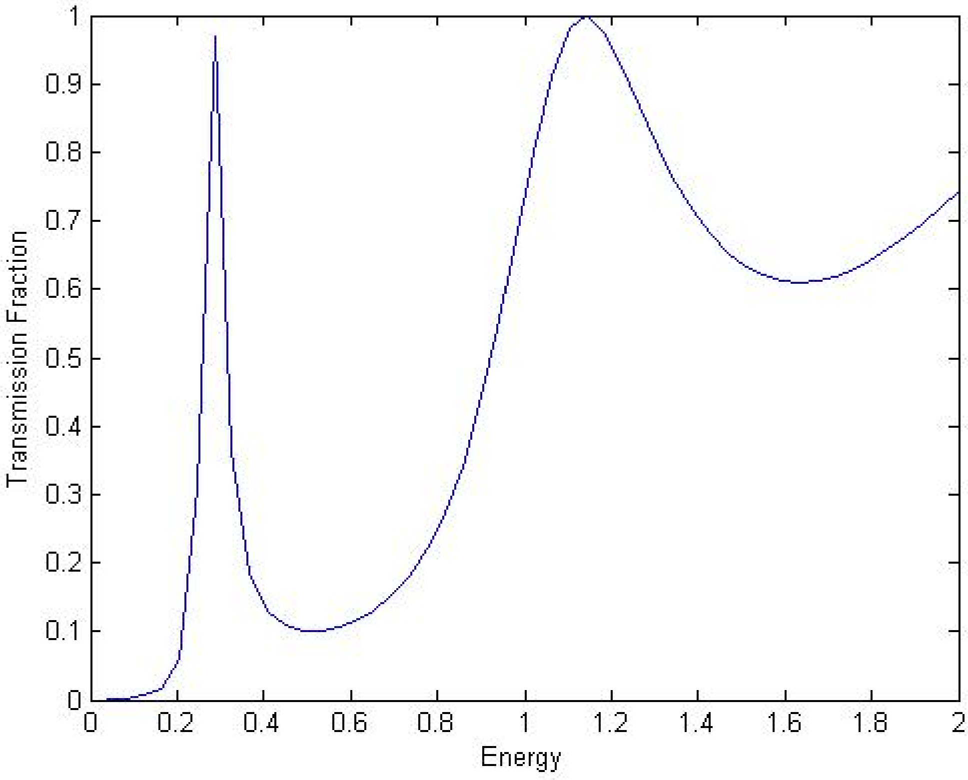
Calculate transmission fraction given the resulting transmission matrix for energy "m"

```
    tfrac(m) = 1 - abs(trans(2,1))^2/abs(trans(1,1))^2;
```

```
end
```

Create transmission-vs-energy plot

```
figure;plot(energy,tfrac);
xlabel('Energy');ylabel('Transmission Fraction');
```



Function to compute the boundary condition matrix between layers "n" and "n+1", using "delta" for layer n

```
function [bcMatrix]=Calculate_Boundary_Condition_Matrix(deltan)

bcMatrix = (1/2)*[1+deltan, 1-deltan; 1-deltan, 1+deltan];

end
```

Function to compute the propagation matrix for layer "n" using the potential, effective mass, and thickness of this layer

```
function [propMatrix]=Calculate_Propagation_Matrix(kn,dmn)

propMatrix = [exp(-1i*kn*dmn), 0; 0, exp(1i*kn*dmn)];
%"1i" is syntax for imaginary unit in MATLAB

end
```

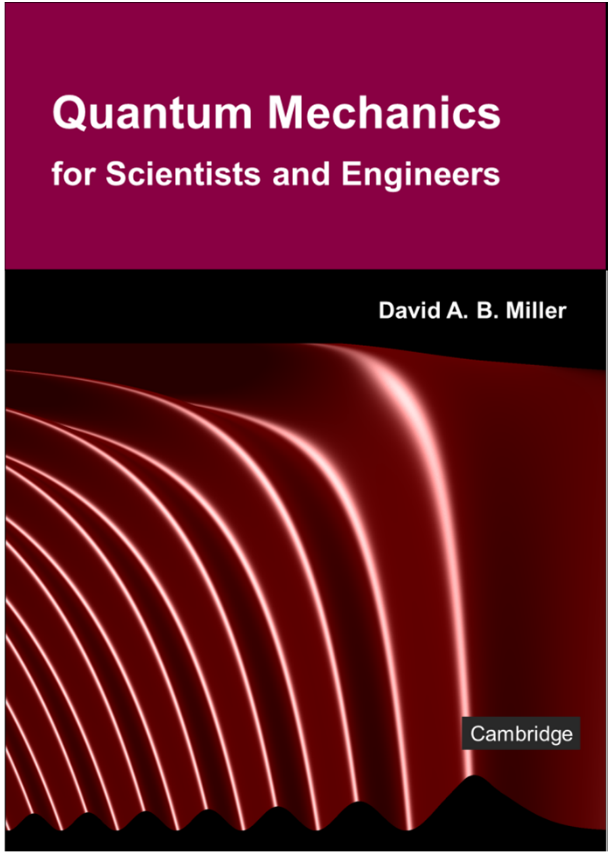
```
end
```

13.1 Electron spin

Slides: Video 13.1.1 Introduction to electron spin

Text reference: Quantum Mechanics for Scientists and Engineers

Chapter 12 introduction



Electron spin

Quantum mechanics for scientists and engineers

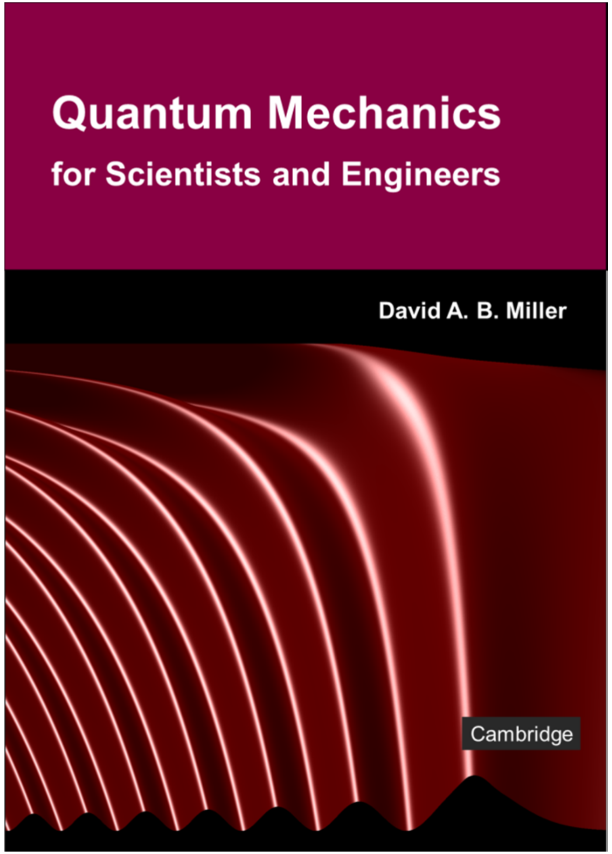
David Miller

13.1 Electron spin

Slides: Video 13.1.2 Angular momentum and magnetic moments

Text reference: Quantum Mechanics for Scientists and Engineers

Section 12.1



Electron spin

Angular momentum and magnetic moments

Quantum mechanics for scientists and engineers

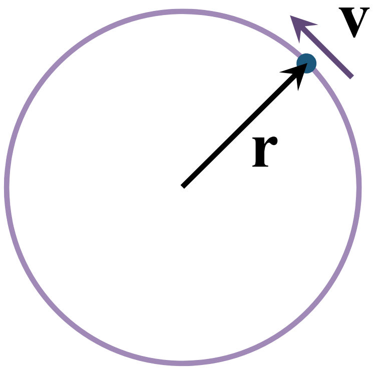
David Miller

Magnetic moments

Charged particles with angular momentum have magnetic moments

- Classically, an electron orbiting
  - with velocity v
  - in a circular orbit of radius r
    - as in the Bohr model of the hydrogen atom
  - has an angular momentum of magnitude

$$L = m_o v r$$



Magnetic moments

We can also write angular momentum as a vector

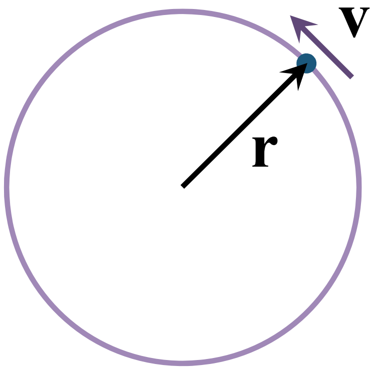
i.e., classically $\mathbf{L} = \mathbf{r} \times \mathbf{p} = \mathbf{r} \times m_o \mathbf{v}$

The electron takes a time $2\pi r / v$ to complete an orbit

so it completes $v / 2\pi r$ orbits/s

so the amount of charge passing a point on the loop per second

i.e., the current, is $I = -ev / 2\pi r$



Magnetic moments

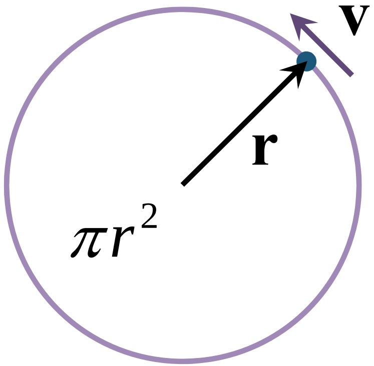
We define the magnetic dipole or magnetic dipole moment μ_d

a quantity that is essentially the strength of a magnet

For any closed current loop

$$\mu_d = \text{current} \times \text{area}$$

The current loop corresponding to the orbit has an area πr^2



Magnetic moments

So, with current $I = -ev/2\pi r$

and area πr^2

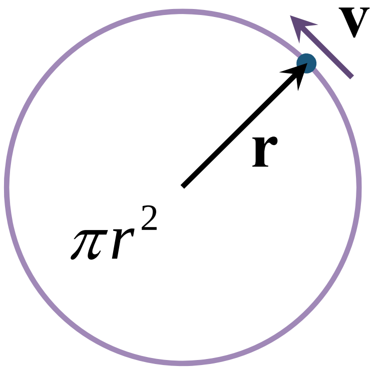
an orbiting electron classically has

$$\mu_d = \text{current} \times \text{area} = -evr/2 = -eL/2m_o$$

or in vector form

$$\boldsymbol{\mu}_e = -\frac{e}{2}\mathbf{r} \times \mathbf{v} = -\frac{e\mathbf{L}}{2m_o}$$

pointing towards us in this diagram (by the right hand rule)



Magnetic moments in magnetic fields

If we apply a magnetic field $\mathbf{B}$

- classically the energy of an object with magnetic moment μ_d changes by $E_\mu = -\boldsymbol{\mu}_d \cdot \mathbf{B}$

Applying $\mathbf{B}$ along the z-direction to a hydrogen atom

- will make the angular momentum quantized around the z-direction with eigenvalues $m\hbar$
  - or in vector form $m\hbar\hat{\mathbf{z}}$
    - where m goes in integer steps from $-l$ to $+l$

Magnetic moments in magnetic fields

Taking a semiclassical model for the moment

with vector angular momentum $m\hbar\hat{\mathbf{z}}$

and our classical formula $\boldsymbol{\mu}_e = -e\mathbf{L} / 2m_o$

we expect magnetic moments for these electron orbits of

$$\boldsymbol{\mu}_e = -\frac{em\hbar\hat{\mathbf{z}}}{2m_o} \equiv -m\mu_B\hat{\mathbf{z}}$$

where μ_B is called the Bohr magneton $\mu_B = \dfrac{e\hbar}{2m_o}$

Applying a magnetic field $\mathbf{B}$

we therefore expect energy changes for these states

$$E_m = m\mu_B B$$

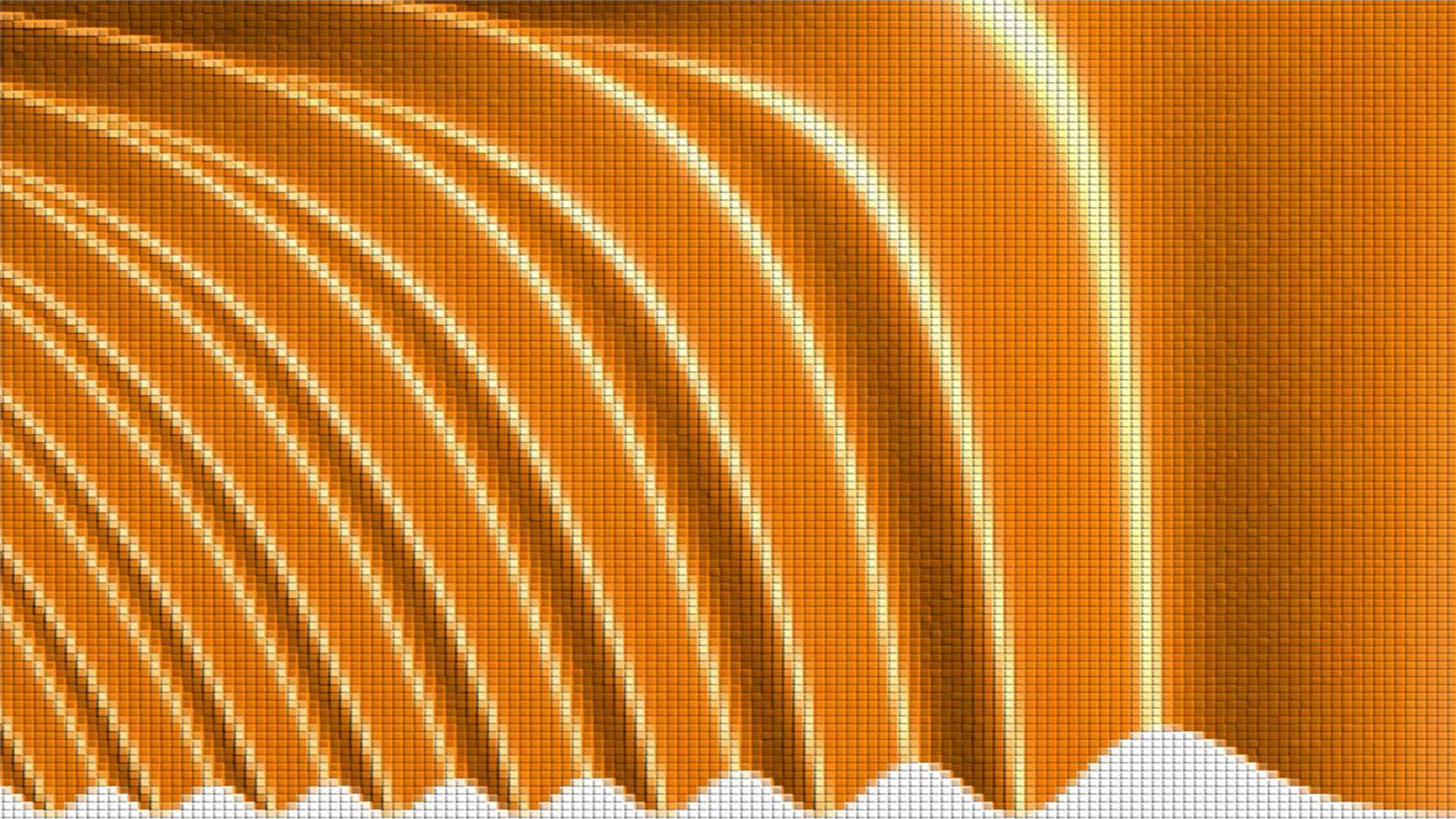
Zeeman effect in hydrogen

So, in a hydrogen atom

- we might expect an applied magnetic field to
  - split the $2l+1$ degenerate energy levels
    - into $2l+1$ different energies – the Zeeman effect
      - e.g., a p state ($l=1$) would split into 3 levels

We should do this calculation quantum mechanically

- e.g., using degenerate perturbation theory
  - with perturbing Hamiltonian operator $\hat{H}_p = (e/2m_o)B\hat{L}_z$
    - but the result (neglecting spin) would be essentially the same

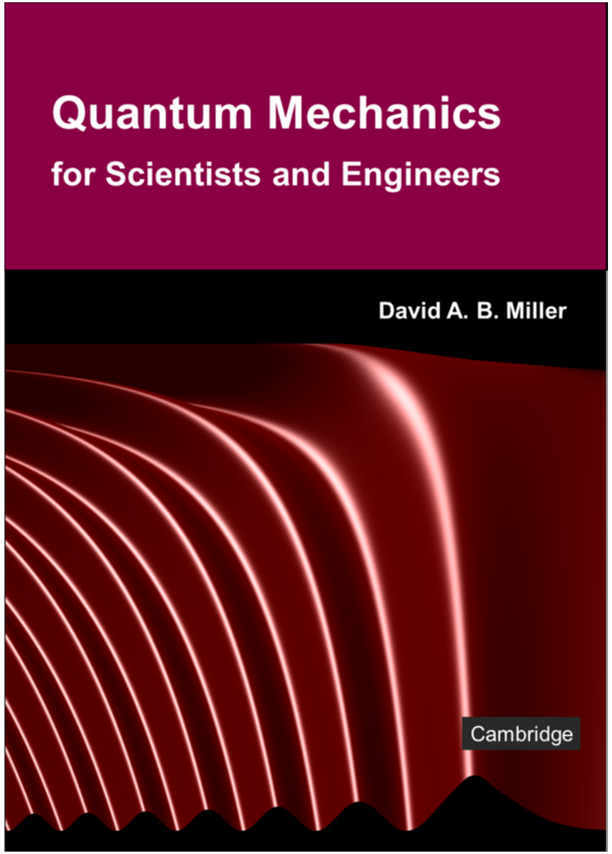


13.1 Electron spin

Slides: Video 13.1.4 Spin angular momentum

Text reference: Quantum Mechanics for Scientists and Engineers

Section 12.2



Electron spin

Spin angular momentum

Quantum mechanics for scientists and engineers

David Miller

Spin angular momentum

To distinguish spin angular momentum from orbital angular momentum

- we use the quantum numbers
  - s rather than l, and
  - σ rather than m

To reconcile this with the quantum mechanics of angular momentum

- to get $2s+1=2$
  - we need $s=1/2$

Spin angular momentum

Hence we assign total spin angular momentum

$\hbar s = \hbar / 2$ to the electron

We say that σ can take values

in integer steps from -s to $+s$

so $\sigma = -1/2$ or $\sigma = +1/2$, and

the corresponding z angular momentum

component in the z direction is $\sigma\hbar$

Spin angular momentum

Incidentally, and somewhat confusingly

- the spin magnetic moment of the electron is not $\sigma\mu_B$
  - but is instead $\mu_e = g\sigma\mu_B$
    - where the so-called gyromagnetic factor $g \simeq 2.0023$

There is no radius of classical orbit of an electron that will give it both

- an angular momentum of $\hbar / 2$ and
- a magnetic moment of $\pm g\mu_B / 2$
  - further confirming that spin cannot be considered as corresponding to a classical orbit of any kind

State vectors for spin angular momentum

Suppose for the moment that we are only interested in the spin properties of the electron

Let us go back and consider

how we would have described an angular momentum state

in the orbital angular momentum case

without describing it explicitly as a function of angle in space

State vectors for orbital angular momentum

Suppose, for example that we considered only

- states with a specific value of l
  - which we can write as $|l\rangle$

In general

- such a state would be some linear combination of the basis states $|l,m\rangle$
  - corresponding to any of the specific allowed values of m, i.e.,

$$|l\rangle = \sum_{m=-l}^{l} a_m |l,m\rangle$$

State vectors for orbital angular momentum

In the case of these states

$$|l\rangle = \sum_{m=-l}^{l} a_m |l,m\rangle$$

each of the states $|l,m\rangle$ can also be written

as one of the spherical harmonic functions in space

and the resulting linear combination $|l\rangle$

can also therefore be written as

a function of angle in space

State vectors for orbital angular momentum

We could also, if we wish, write

$$|l\rangle = \sum_{m=-l}^{l} a_m |l,m\rangle$$

explicitly as a vector

$$|l\rangle \equiv \begin{bmatrix} a_l \\ a_{l-1} \\ \vdots \\ a_{-l+1} \\ a_{-l} \end{bmatrix}$$

Note that the set of functions corresponding to all the possible values of m for a given l

is a complete set for describing any possible function with that value of l

including even the eigenfunctions of $\hat{L}_x$ and $\hat{L}_y$ oriented around the other axes

State vectors for spin angular momentum

In the case of the electron spin

we cannot write the basis functions as functions of angle in space

but we do expect that we can write them using

the same kind of state and vector formalism

as we use for other angular momenta

For electron spin that formalism becomes very simple

Instead of l, we have s

which we know is ½

and instead of m we have σ

State vectors for spin angular momentum

There are however, now only two basis states

$|1/2,1/2\rangle$ and $|1/2,-1/2\rangle$

corresponding to $\sigma = 1/2$ and $\sigma = -1/2$ respectively

Hence, writing our general spin state as $|s\rangle$ we have

$$|s\rangle = a_{1/2}|1/2,1/2\rangle + a_{-1/2}|1/2,-1/2\rangle \equiv a_{1/2}|\uparrow\rangle + a_{-1/2}|\downarrow\rangle \equiv \begin{bmatrix} a_{1/2} \\ a_{-1/2} \end{bmatrix}$$

where we show another common notation

with $|\uparrow\rangle$ being the “spin-up” state $|1/2,1/2\rangle$, and

$|\downarrow\rangle$ being the “spin-down” state $|1/2,-1/2\rangle$

The “up” and “down” conventionally refer to the z direction

State vectors for spin angular momentum

Any possible spin state of the electron can presumably be described this way

Rather obviously

- a state with its magnetic moment in the $+z$ direction
  - the "spin-up" state – will be the state $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$
- a state with its magnetic moment in the $-z$ direction
  - the "spin-down" state – will be the state $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

State vectors for spin angular momentum

The choice of unit amplitudes for these states $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

also assures they are normalized

Normalization here means assuring that

the sum of the modulus squared of the two vector elements is equal to one

i.e., $\left|a_{1/2}\right|^2 + \left|a_{-1/2}\right|^2 = 1$

We could also multiply these states by any unit complex number

and they would still be spin-up and spin-down states respectively

State vectors for spin angular momentum

We might think that these vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

can represent only spin-up and spin-down states

oriented along the z axis

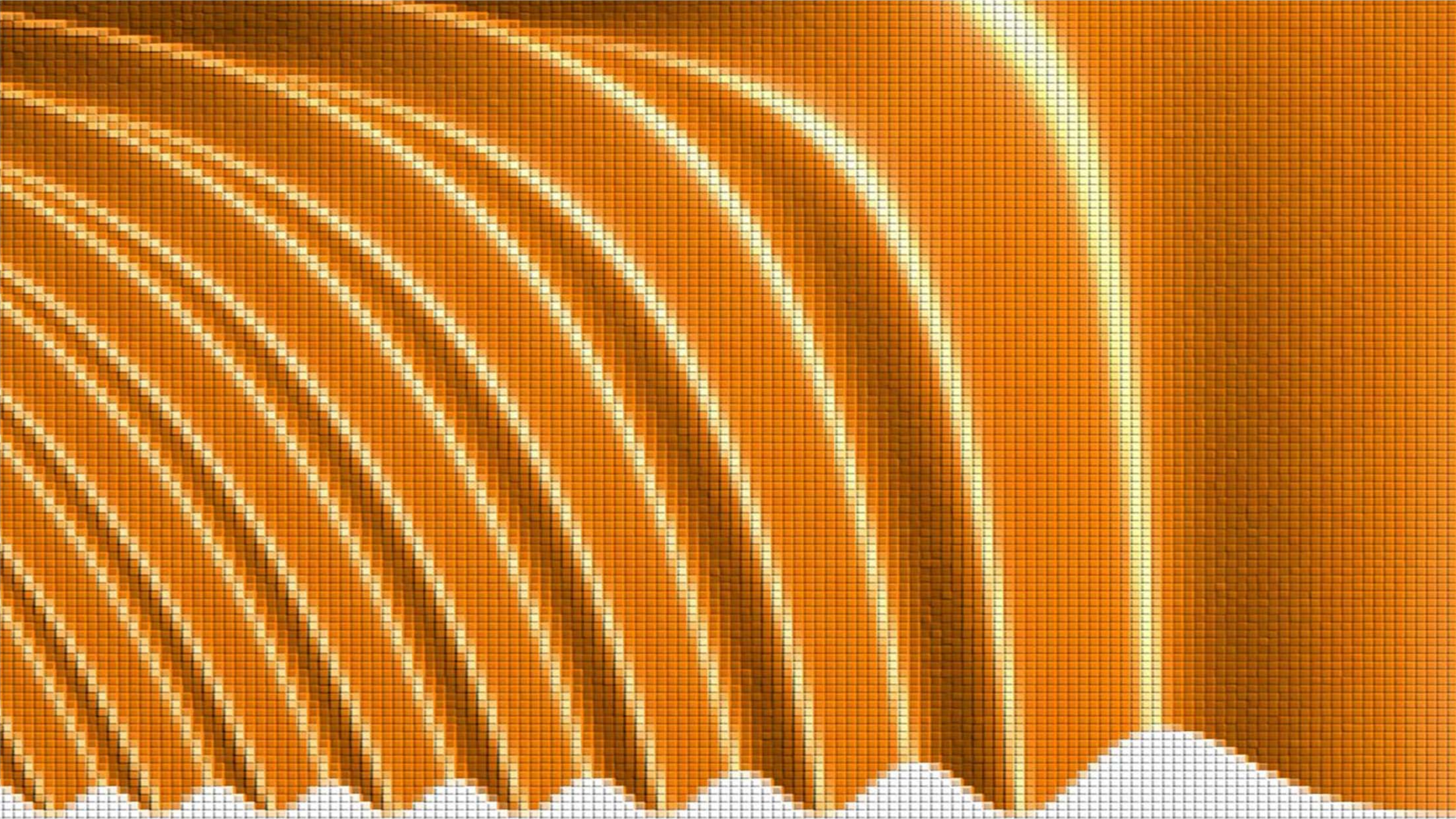
In fact, these two basis vectors can represent any possible spin state of the electron

including spin states with the magnetic moment

oriented along the x direction or

oriented along the y direction

We can show this once we define the spin operators

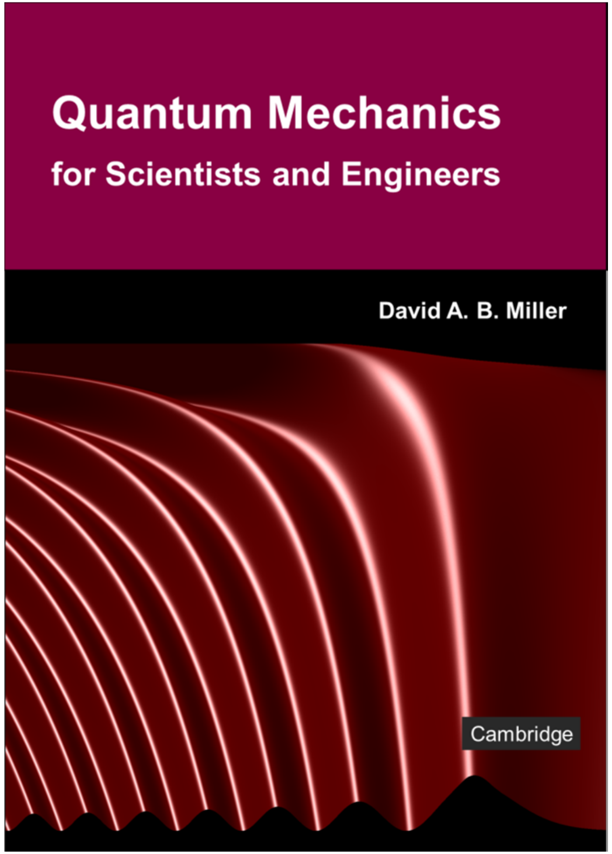


13.1 Electron spin

Slides: Video 13.1.6 Operators for spin angular momentum

Text reference: Quantum Mechanics for Scientists and Engineers

Section 12.3



Electron spin

Operators for spin angular momentum

Quantum mechanics for scientists and engineers

David Miller

Commutation relations for spin operators

So now we ask the spin angular momentum operators

which we write as $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$

by analogy with the orbital angular momentum operators $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$

to obey a set of commutation relations

$$\left[\hat{S}_x, \hat{S}_y\right] = i\hbar\hat{S}_z$$

$$\left[\hat{S}_y, \hat{S}_z\right] = i\hbar\hat{S}_x$$

$$\left[\hat{S}_z, \hat{S}_x\right] = i\hbar\hat{S}_y$$

Commutation relations for spin operators

Commonly we work with the "dimensionless" operators

$\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$

from which the spin angular momentum magnitude $\hbar/2$ has been removed, i.e.,

$\hat{\sigma}_x = 2\hat{S}_x/\hbar$, $\hat{\sigma}_y = 2\hat{S}_y/\hbar$, $\hat{\sigma}_z = 2\hat{S}_z/\hbar$

giving the set of commutation relations

$$\left[\hat{\sigma}_x, \hat{\sigma}_y\right] = 2i\hat{\sigma}_z$$

$$\left[\hat{\sigma}_y, \hat{\sigma}_z\right] = 2i\hat{\sigma}_x$$

$$\left[\hat{\sigma}_z, \hat{\sigma}_x\right] = 2i\hat{\sigma}_y$$

Pauli spin matrices

If we choose to represent the spin function in the vector format

then the operators become represented by matrices

One set of matrix representations of these operators is

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Such matrix representations are known as

Pauli spin matrices

Paulie spin matrices

There is more than one way we could have chosen these

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \ \hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

In fact there is an infinite number of ways

depending on what axis we choose for the spin

This set, which we can call the z representation

is such that the spin-up and spin-down vectors defined previously

are eigenvectors of the $\hat{\sigma}_z$ operator

These operators do obey the commutation relations

Spin operators

We can write the three Pauli spin matrices as one entity, $\hat{\boldsymbol{\sigma}}$

which has components associated with each of the coordinate directions x, y, and z

$$\hat{\boldsymbol{\sigma}} = \mathbf{i}\hat{\sigma}_x + \mathbf{j}\hat{\sigma}_\mathbf{y} + \mathbf{k}\hat{\sigma}_\mathbf{z} \equiv \mathbf{i}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \mathbf{j}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \mathbf{k}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

For completeness, by analogy with the $\hat{L}^2$ operator

we can also define an $\hat{S}^2$ operator

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$$

or a $\hat{\sigma}^2$ operator

$$\hat{\sigma}^2 \equiv \hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\sigma}} = \hat{\sigma}_x^2 + \hat{\sigma}_y^2 + \hat{\sigma}_z^2$$

Spin operators

From the definitions for the Pauli matrices

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

we see that

$$\hat{\sigma}^2 = \hat{\sigma}_x^2 + \hat{\sigma}_y^2 + \hat{\sigma}_z^2 \equiv 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and hence that

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 \equiv \frac{3}{4}\hbar^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = s(s+1)\hbar^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

so any spin ½ vector is an eigenvector of $\hat{S}^2$

with eigenvalue $s(s+1)\hbar^2 = (3/4)\hbar^2$

Spin and orbital angular momentum operators

Just as any spin ½ vector is an eigenvector of $\hat{S}^2$

with eigenvalue $s(s+1)\hbar^2 = (3/4)\hbar^2$

for orbital angular momentum

any linear combination of spherical harmonics

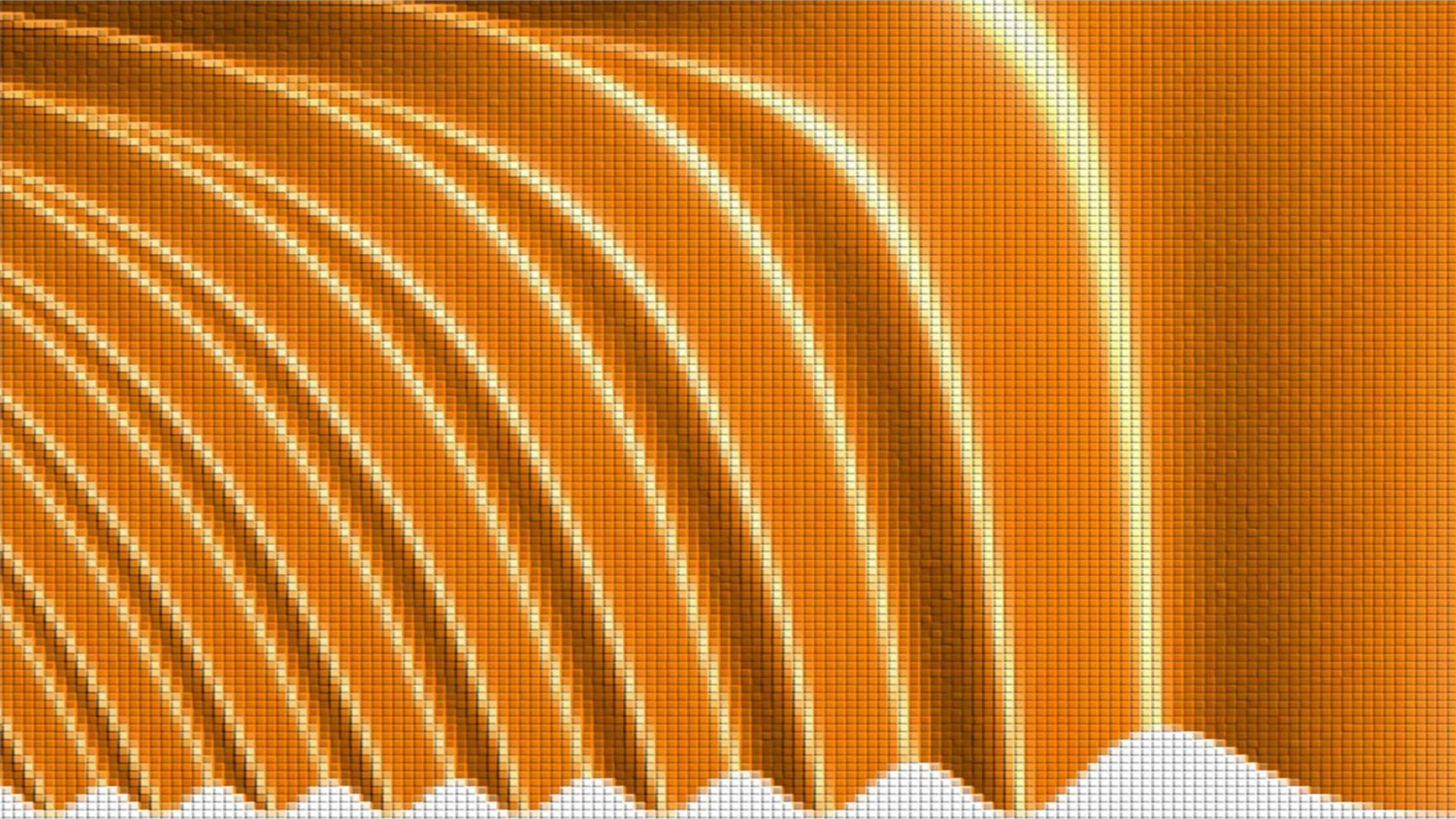
corresponding to a given l value

is an eigenfunction of the $\hat{L}^2$ operator

with eigenvalue $l(l+1)\hbar^2$

so the spin operator behaviors are still analogous to

the behavior of orbital angular momentum operators

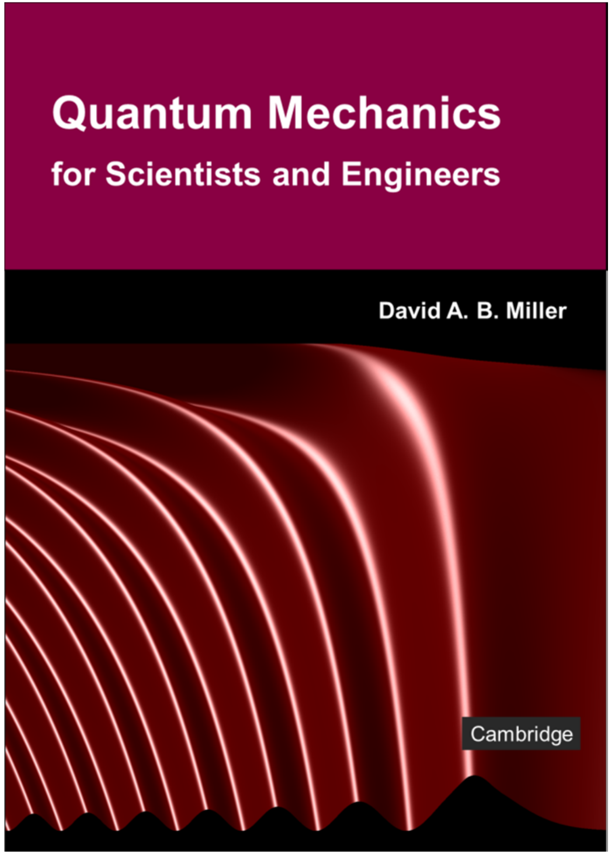


13.2 Spin states

Slides: Video 13.2.1 Visualizing spin states

Text reference: Quantum Mechanics for Scientists and Engineers

Section 12.4



Spin states

Visualizing spin states

Quantum mechanics for scientists and engineers

David Miller

Visualizing spin states

We can therefore write the general spin state as

$$|s\rangle = \cos(\theta/2)|\uparrow\rangle + \exp(i\phi)\sin(\theta/2)|\downarrow\rangle \equiv \begin{bmatrix} \cos(\theta/2) \\ \exp(i\phi)\sin(\theta/2) \end{bmatrix}$$

at least within an overall phase factor

Since $\cos^2(\theta/2) + \sin^2(\theta/2) = 1$

the magnitude of this vector is correctly guaranteed to be unity

and the $\exp(i\phi)$ factor allows for

any relative quantum-mechanical phase between the two components

The Bloch sphere

We now ask for the expectation value $\langle s|\hat{\boldsymbol{\sigma}}|s\rangle$

of the Pauli spin operator $\hat{\boldsymbol{\sigma}}$ with such a state $|s\rangle$

obtaining as the result

after some algebra

what we will call

a “spin polarization” vector $\mathbf{P}_s$

$$\mathbf{P}_s = \langle s|\hat{\boldsymbol{\sigma}}|s\rangle = \mathbf{i}\langle s|\hat{\sigma}_x|s\rangle + \mathbf{j}\langle s|\hat{\sigma}_y|s\rangle + \mathbf{k}\langle s|\hat{\sigma}_z|s\rangle$$

$$= \mathbf{i}\sin\theta\cos\phi + \mathbf{j}\sin\theta\sin\phi + \mathbf{k}\cos\theta$$

The Bloch sphere

E.g., one term in this evaluation of $\mathbf{P}_s$ is, explicitly

$$\langle s|\hat{\sigma}_x|s\rangle = \begin{bmatrix}\cos(\theta/2) & \exp(-i\phi)\sin(\theta/2)\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\cos(\theta/2)\\ \exp(i\phi)\sin(\theta/2)\end{bmatrix}$$

$$= \begin{bmatrix}\cos(\theta/2) & \exp(-i\phi)\sin(\theta/2)\end{bmatrix}\begin{bmatrix}\exp(i\phi)\sin(\theta/2)\\ \cos(\theta/2)\end{bmatrix}$$

$$= \cos(\theta/2)\sin(\theta/2)\left[\exp(-i\phi)+\exp(i\phi)\right]$$

$$= 2\cos(\theta/2)\sin(\theta/2)\cos\phi$$

$$= \sin\theta\cos\phi$$

The Bloch sphere

Now

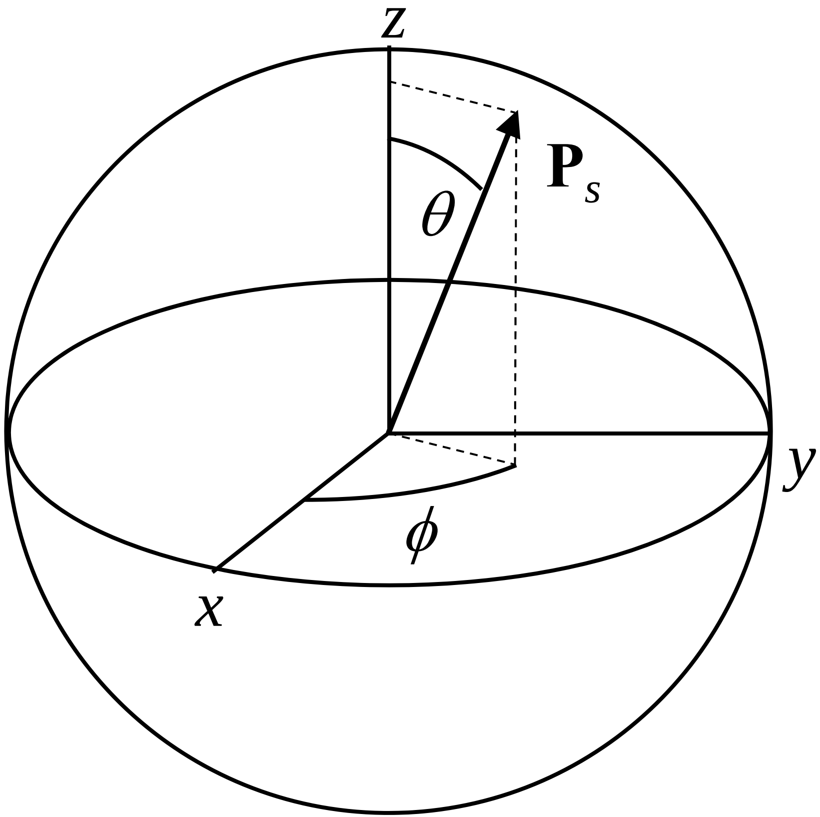
$$\mathbf{P}_s = \mathbf{i}\sin\theta\cos\phi + \mathbf{j}\sin\theta\sin\phi + \mathbf{k}\cos\theta$$

is a vector from the origin

out to a point on a sphere of unit radius

with angle relative to the North pole of θ

and azimuthal angle ϕ



The Bloch sphere

The general spin state $|s\rangle$

can be visualized

in terms of its spin

polarization vector

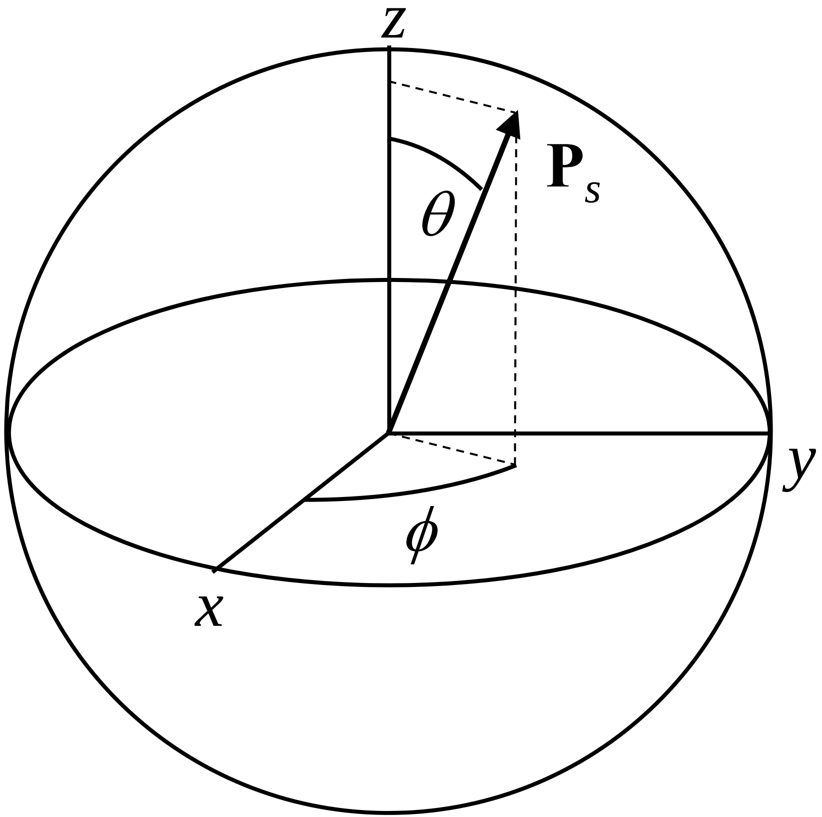
$$\mathbf{P}_s = \langle s|\hat{\boldsymbol{\sigma}}|s\rangle$$

as a vector on a unit

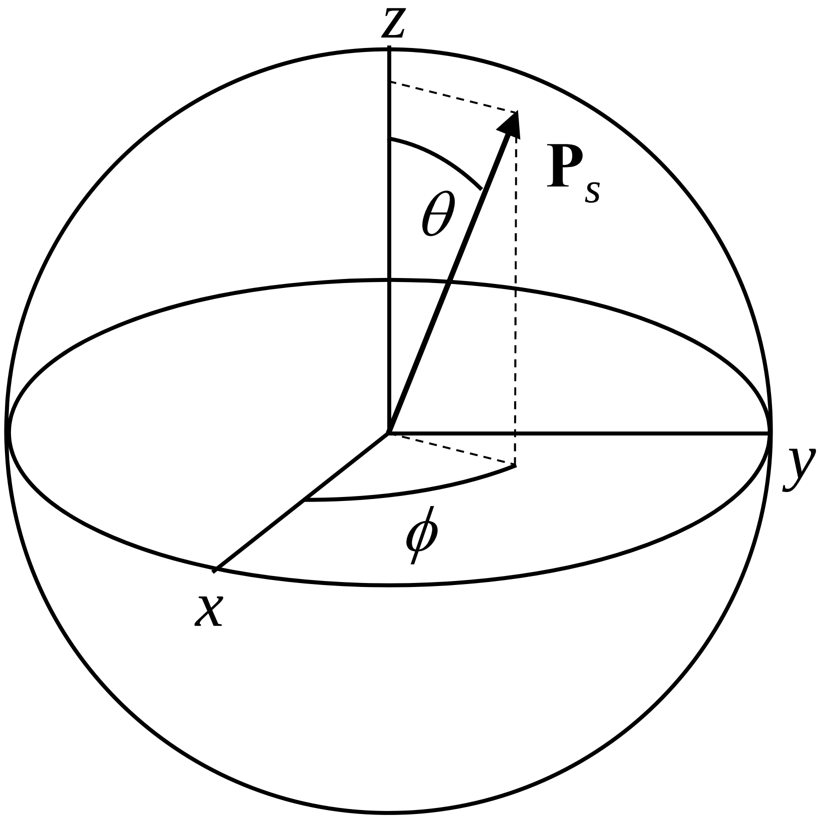
sphere

The North pole corresponds to the state $|\uparrow\rangle$

and the South pole to state $|\downarrow\rangle$



The Bloch sphere

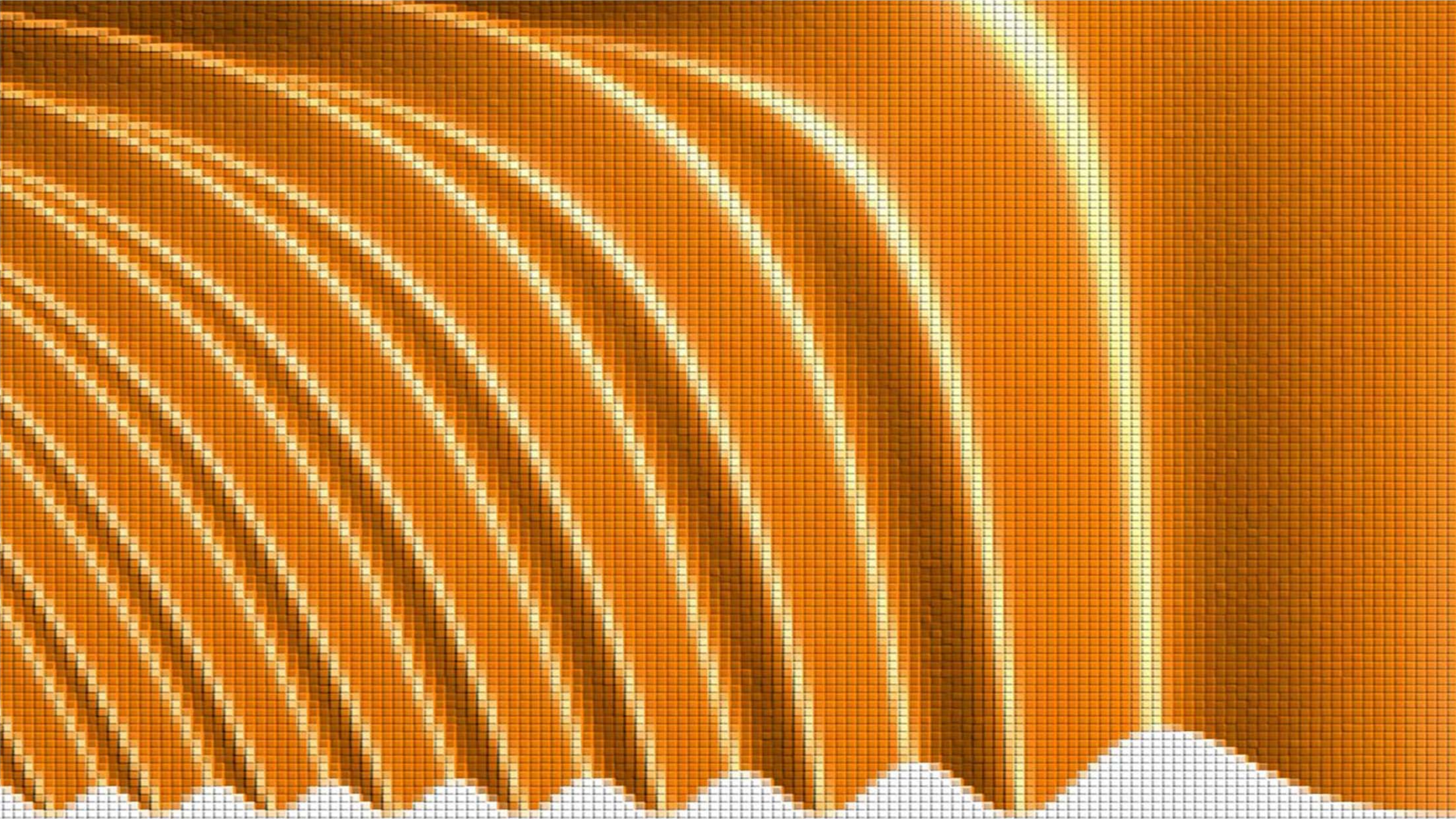


This is called the Bloch sphere

- with the angles θ and ϕ on this sphere

characterizing the spin state

- and the geometrical x, y, and z directions corresponding
  - to the directions of the eigenvectors of the corresponding spin operators

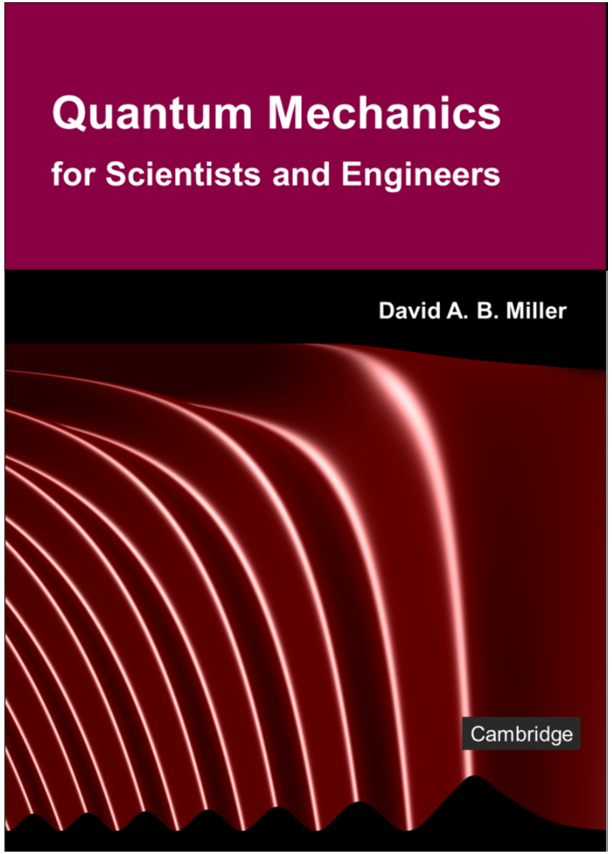


13.2 Spin states

Slides: Video 13.2.3 Wavefunctions, spin and Hilbert space

Text reference: Quantum Mechanics for Scientists and Engineers

Section 12.5



Spin states

Wavefunctions, spin and Hilbert space

Quantum mechanics for scientists and engineers

David Miller

Wavefunctions with spin

Thus if Ψ is to be the most complete representation of the electron state

including spin effects

we might write

$$|\Psi\rangle \equiv \begin{bmatrix} \psi_{\uparrow}(\mathbf{r},t) \\ \psi_{\downarrow}(\mathbf{r},t) \end{bmatrix} \equiv \psi_{\uparrow}(\mathbf{r},t)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \psi_{\downarrow}(\mathbf{r},t)\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

A function of the form $\begin{bmatrix} \psi_{\uparrow}(\mathbf{r},t) \\ \psi_{\downarrow}(\mathbf{r},t) \end{bmatrix}$ is called a "spinor"

Basis functions in combined Hilbert spaces

The basis functions in our new Hilbert space

are all the products of the basis functions in the original separate spaces

For example, if the basis functions for the spatial and temporal function were $\psi_1(\mathbf{r},t)$, $\psi_2(\mathbf{r},t)$, …, $\psi_j(\mathbf{r},t)$, …

then the basis functions when we add spin are

$$\psi_1(\mathbf{r},t)\begin{bmatrix}1\\0\end{bmatrix}, \psi_2(\mathbf{r},t)\begin{bmatrix}1\\0\end{bmatrix}, \ldots, \psi_j(\mathbf{r},t)\begin{bmatrix}1\\0\end{bmatrix}, \ldots,$$

$$\psi_1(\mathbf{r},t)\begin{bmatrix}0\\1\end{bmatrix}, \psi_2(\mathbf{r},t)\begin{bmatrix}0\\1\end{bmatrix}, \ldots, \psi_j(\mathbf{r},t)\begin{bmatrix}0\\1\end{bmatrix}, \ldots$$

Combining functions in Hilbert spaces for spatial problems

This concept of new basis functions being the products of the elements of two basis function sets

is not exclusively a quantum mechanical one

E.g., we might represent a classical spatial function

in one dimensional box of size L_x

as a Fourier series of the form

$$f(x) = \sum_n a_n \exp\left(i2n\pi x / L_x\right)$$

so we have basis functions of the form

$$\exp\left(i2n\pi x / L_x\right)$$

Combining functions in Hilbert spaces for spatial problems

A function in a two-dimensional rectangular box

of sizes L_x and L_y in the respective coordinate directions

can be represented as a Fourier series

$$g(x,y) = \sum_{n,p} a_{n,p} \exp(i2\pi nx / L_x) \exp(i2\pi py / L_y)$$

Here the new basis functions $\exp(i2\pi nx / L_x) \exp(i2\pi py / L_y)$

are the products of the basis functions

of the two Hilbert spaces

associated with the two separate problems of

functions in x and functions in y

Direct product spaces

A Hilbert space formed by

combining two other spaces

and making the new basis functions

the products of the basis functions in the different spaces

is called a direct product space

The spinors exist in a direct product space

formed by the multiplication of

the spatial and temporal basis functions and

the spin basis functions

Dirac notation and direct product spaces

In the electron spin case

we could write the basis functions in Dirac notation as

$$|\psi_1\rangle|\uparrow\rangle, |\psi_2\rangle|\uparrow\rangle, \ldots|\psi_j\rangle|\uparrow\rangle, \ldots, |\psi_1\rangle|\downarrow\rangle, |\psi_2\rangle|\downarrow\rangle, \ldots, |\psi_j\rangle|\downarrow\rangle, \ldots$$

Here, we understand that

the $|\psi_j\rangle$ kets are vectors in one Hilbert space

representing arbitrary spatial and temporal functions

and the $|\uparrow\rangle$ and $|\downarrow\rangle$ kets are vectors in another Hilbert space

representing only spin functions

Dirac notation and direct product spaces

The products $|\psi_j\rangle|\uparrow\rangle$ and $|\psi_j\rangle|\downarrow\rangle$

are vectors in the direct product Hilbert space

Direct products

are “products” of vectors in different Hilbert spaces

to give a new vector in the “direct product” space

and are sometimes written explicitly as $|a\rangle \otimes |b\rangle$

though we will mostly not use this notation

We could also write these products

using any of the notations $|\psi_j \uparrow\rangle \equiv |\psi_j\rangle|\uparrow\rangle \equiv |\psi_j\rangle \otimes |\uparrow\rangle$

Dirac notation and direct product spaces

For example, we could write the basis functions of our direct product time, position and spin Hilbert space as

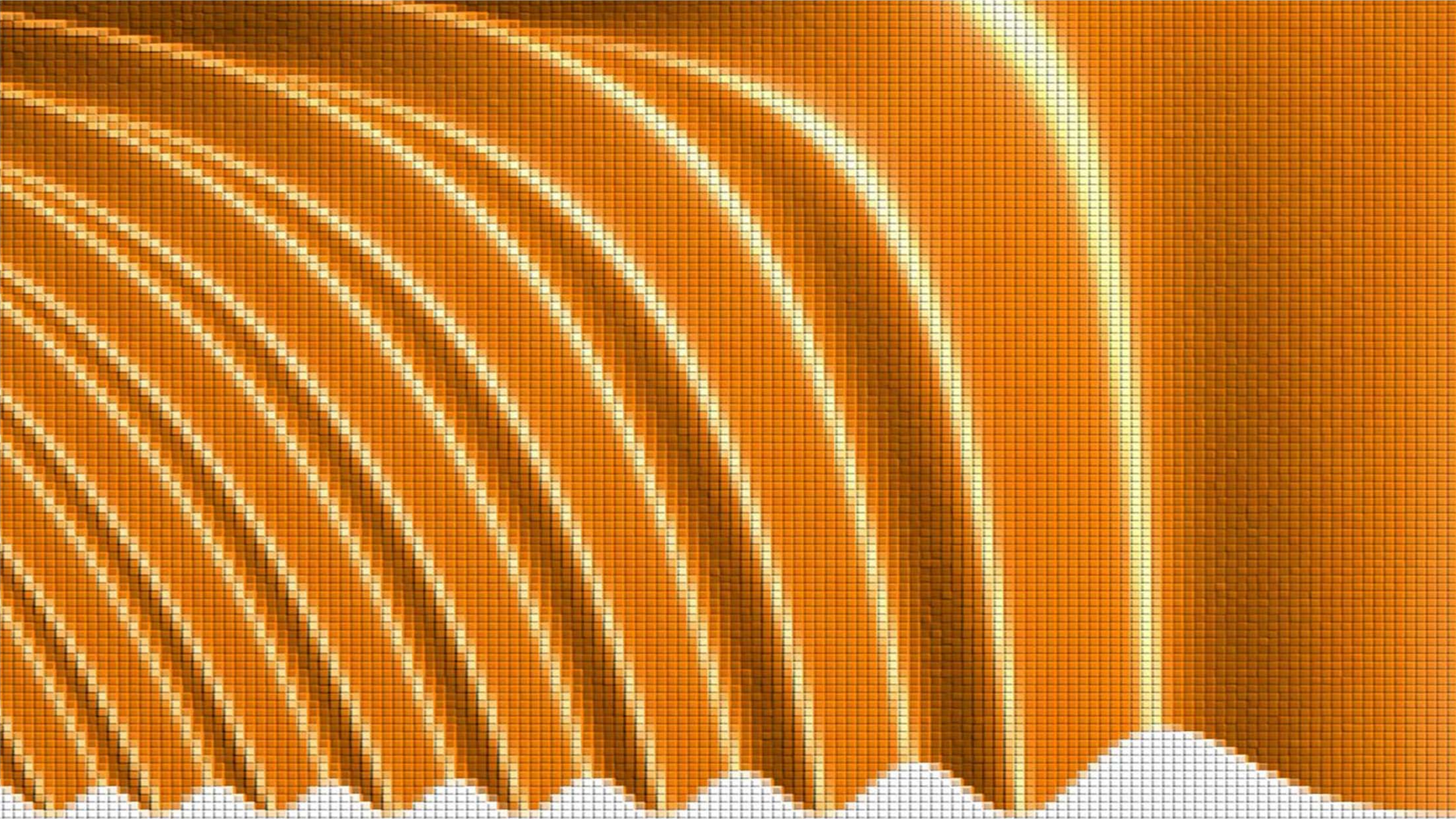
$$|\psi_1 \uparrow\rangle, |\psi_2 \uparrow\rangle, \ldots, |\psi_j \uparrow\rangle, \ldots, |\psi_1 \downarrow\rangle, |\psi_2 \downarrow\rangle, \ldots, |\psi_j \downarrow\rangle, \ldots$$

With our different notations

we could also write

$$|\Psi\rangle \equiv \begin{bmatrix} \psi_\uparrow(\mathbf{r},t) \\ \psi_\downarrow(\mathbf{r},t) \end{bmatrix} \equiv \psi_\uparrow(\mathbf{r},t)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \psi_\downarrow(\mathbf{r},t)\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

as $|\Psi\rangle = |\psi_\uparrow\rangle|\uparrow\rangle + |\psi_\downarrow\rangle|\downarrow\rangle = |\psi_\uparrow \uparrow\rangle + |\psi_\downarrow \downarrow\rangle$

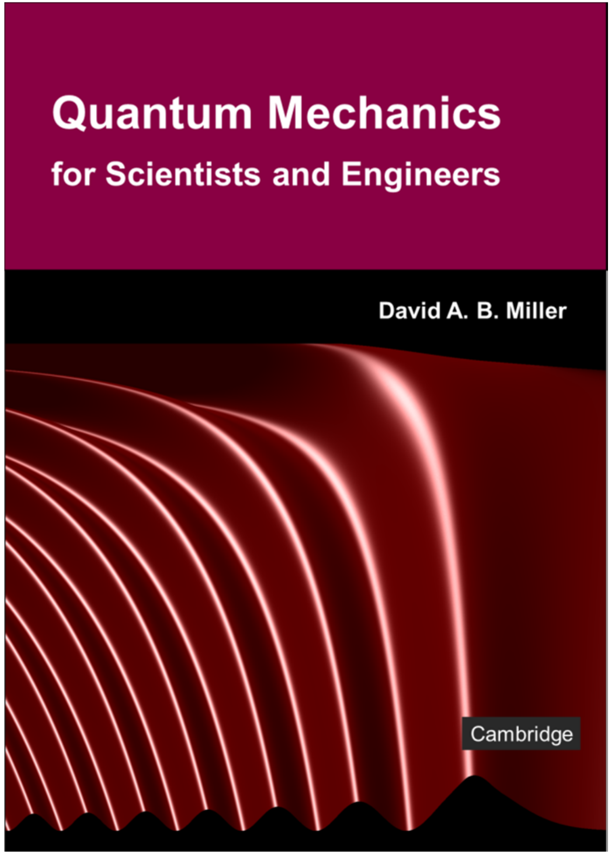


13.2 Spin states

Slides: Video 13.2.5 The Pauli equation

Text reference: Quantum Mechanics for Scientists and Engineers

Section 12.6



Spin states

The Pauli equation

Quantum mechanics for scientists and engineers

David Miller

Energy of a spin in a magnetic field

Classically, because angular momentum is a vector

then electron spin would also be a vector

$$\boldsymbol{\sigma}$$

Then we would expect a magnetic moment

$$\boldsymbol{\mu}_e = g\mu_B\boldsymbol{\sigma}$$

in a vector generalization of

$$\mu_e = g\sigma\mu_B$$

Energy of a spin in a magnetic field

In this classical analogy

with magnetic moment $\boldsymbol{\mu}_e = g\mu_B \boldsymbol{\sigma}$

the energy associated with that magnetic moment

in a magnetic field $\mathbf{B} = \mathbf{i}B_x + \mathbf{j}B_y + \mathbf{k}B_z$

would be $E_S = \boldsymbol{\mu}_e \cdot \mathbf{B} = g\mu_B \boldsymbol{\sigma} \cdot \mathbf{B}$

In the quantum mechanical case, as usual we postulate

an operator instead of the classical quantity

so instead of E_S, we have, with Pauli spin operator $\hat{\boldsymbol{\sigma}}$

$$\hat{H}_S = \frac{g\mu_B}{2} \hat{\boldsymbol{\sigma}} \cdot \mathbf{B} \equiv \frac{g\mu_B}{2} B_x \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \frac{g\mu_B}{2} B_y \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \frac{g\mu_B}{2} B_z \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Energy of a spin in a magnetic field

In $\hat{H}_S = \frac{g\mu_B}{2}\hat{\boldsymbol{\sigma}} \cdot \mathbf{B}$

compared to the classical $E_S = \boldsymbol{\mu}_e \cdot \mathbf{B} = g\mu_B \boldsymbol{\sigma} \cdot \mathbf{B}$

the factor of ½ in the quantum expression

is only because we like to work with Pauli matrices

with eigenvalues of unit magnitude

rather than the half integer magnitude

associated with the spin itself

It does not express any other difference in the physics

The Pauli equation includes this energy term $\hat{H}_S = \frac{g\mu_B}{2}\hat{\boldsymbol{\sigma}} \cdot \mathbf{B}$

The Pauli equation

The Pauli equation also

- treats electromagnetic effects on the electron as a charged particle semiclassically
  - i.e., with classical electric and magnetic fields

So by extension from classical electromagnetism

- it uses $\hat{\mathbf{p}} - e\mathbf{A}$
  - instead of just the momentum operator $\hat{\mathbf{p}} = -i\hbar\nabla$
    - in constructing the rest of the energy terms in the equation

The Pauli equation

Hence, instead of the Schrödinger equation, we have the Pauli equation

$$\left[\frac{1}{2m_o}\left(\hat{\mathbf{p}}-e\mathbf{A}\right)^2+V+\frac{g\mu_B}{2}\hat{\boldsymbol{\sigma}}\cdot\mathbf{B}\right]\Psi=i\hbar\frac{\partial\Psi}{\partial t}$$

Note here that $\Psi\equiv\begin{bmatrix}\psi_{\uparrow}\left(\mathbf{r},t\right)\\ \psi_{\downarrow}\left(\mathbf{r},t\right)\end{bmatrix}$ is a spinor

The Pauli equation is therefore not one differential equation

but is in general two coupled ones

The Pauli equation

More rigorously, since we think of $\hat{\boldsymbol{\sigma}}$ as a 2x2 matrix

we should actually introduce

the corresponding 2x2 identity matrix $\hat{I}_S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

and we should write

$$\left\{ \frac{1}{2m_o} \left(\hat{\mathbf{p}} - e\mathbf{A} \right)^2 \hat{I}_S + V\hat{I}_S + \frac{g\mu_B}{2} \hat{\boldsymbol{\sigma}} \cdot \mathbf{B} \right\} \Psi = i\hbar \frac{\partial \Psi}{\partial t}$$

Then everything in the curly brackets {…} has the

2x2 matrix character needed for the spinor $\Psi \equiv \begin{bmatrix} \psi_\uparrow(\mathbf{r},t) \\ \psi_\downarrow(\mathbf{r},t) \end{bmatrix}$

though in practice $\hat{I}_S$ is just assumed

rather than written explicitly

The Pauli equation

The Pauli equation

$$\left[\frac{1}{2m_o}\left(\hat{\mathbf{p}}-e\mathbf{A}\right)^2+V+\frac{g\mu_B}{2}\hat{\boldsymbol{\sigma}}\cdot\mathbf{B}\right]\Psi=i\hbar\frac{\partial\Psi}{\partial t}$$

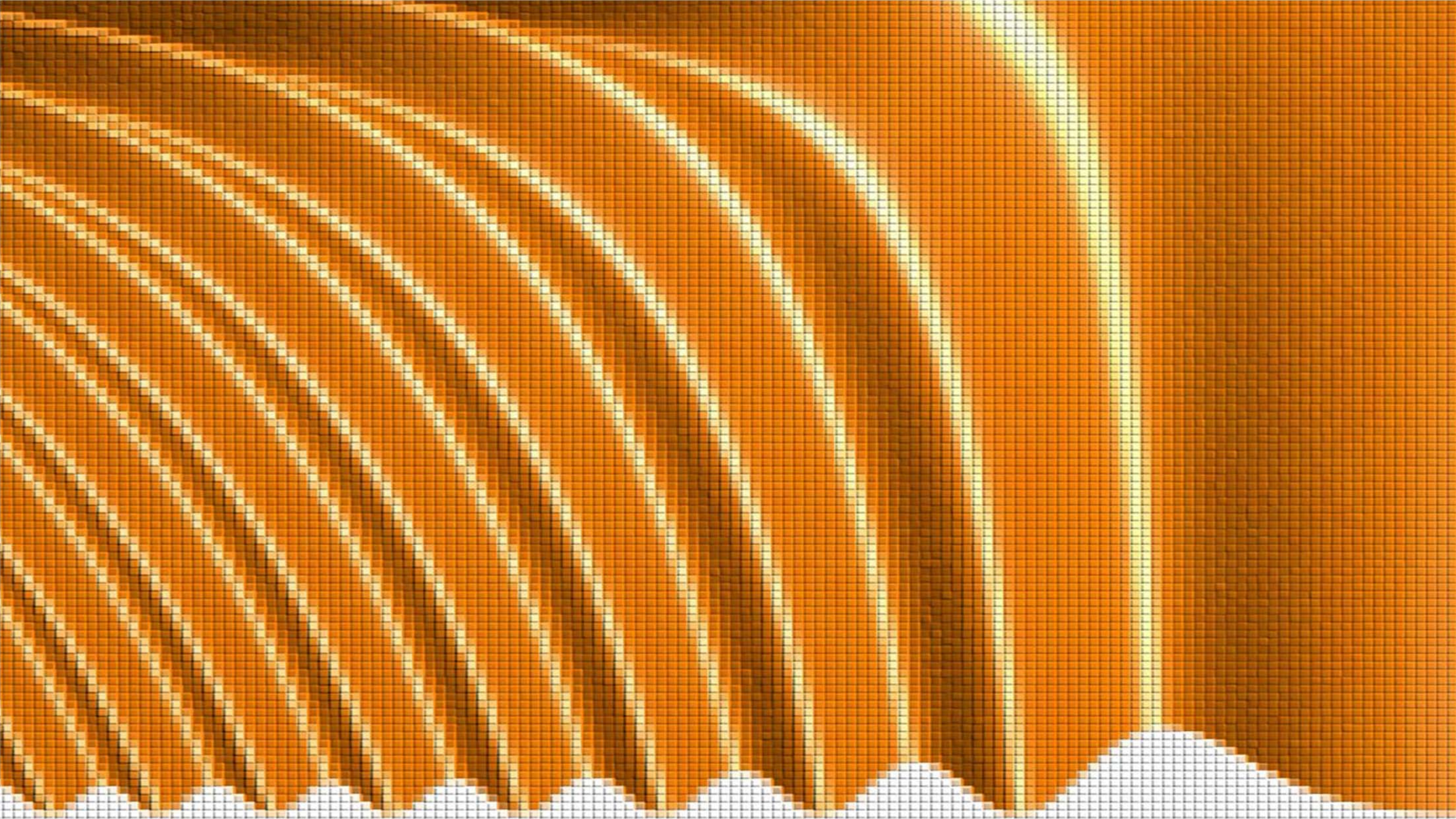
is the starting point for investigating

the effects of magnetic fields on electrons

It can be used, for example

to derive the Zeeman effect rigorously

including the effects of spin

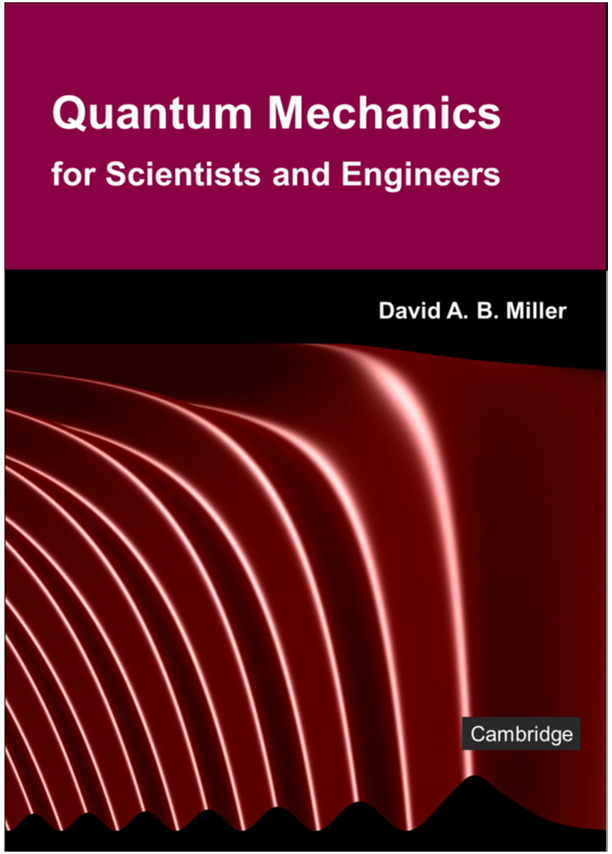


13.2 Spin states

Slides: Video 13.2.7 Where does spin come from?

Text reference: Quantum Mechanics for Scientists and Engineers

Section 12.7

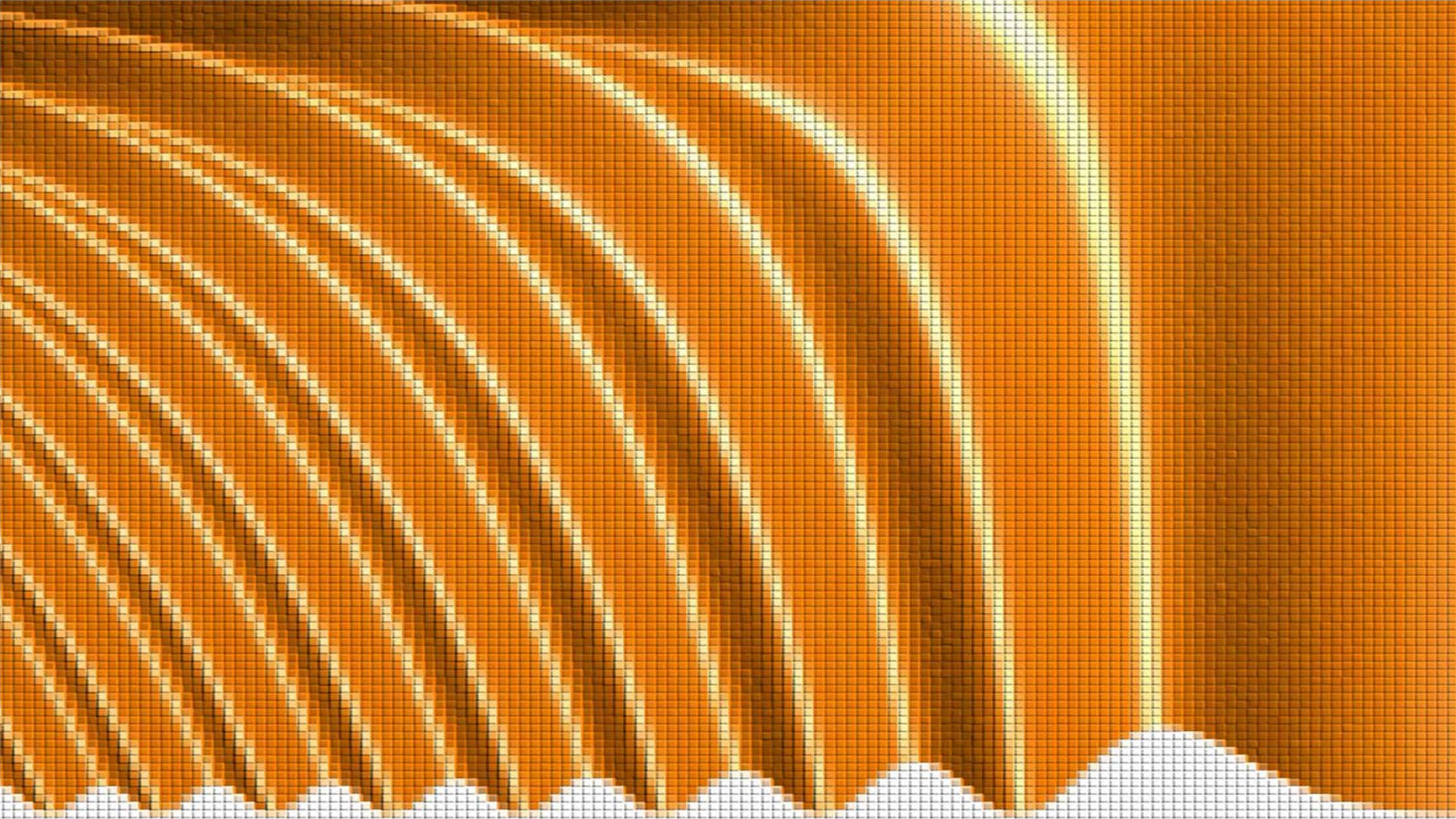


Spin states

Where does spin come from?

Quantum mechanics for scientists and engineers

David Miller

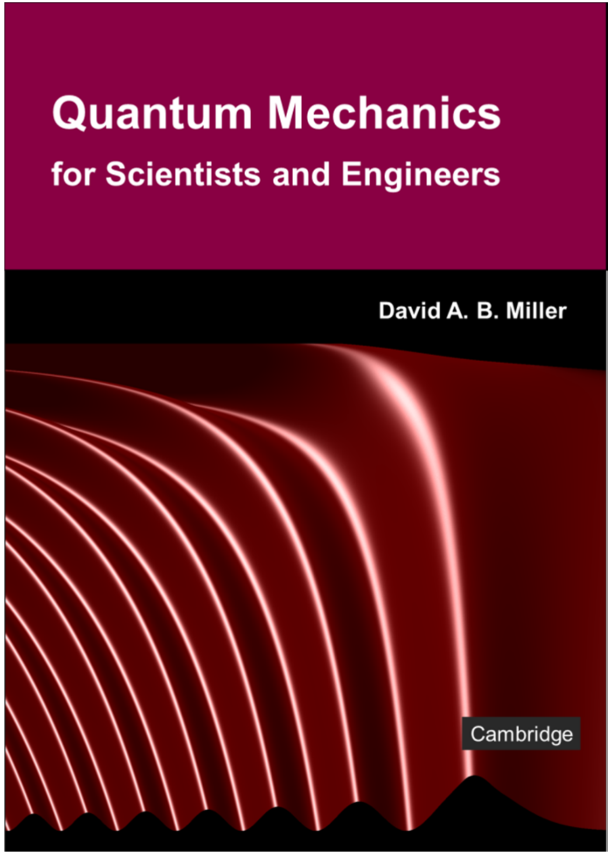


13.3 Identical particles and exchange

Slides: Video 13.3.1 Scattering identical particles

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.1 up to Eq. 13.17



Identical particles and exchange

Scattering identical particles

Quantum mechanics for scientists and engineers

David Miller

Scattering of identical particles

- We imagine that the two electrons are traveling through space
  - each in some kind of wavepacket
- The wavepackets might each be quite localized in space at any given time
  - but they each extend arbitrarily far
    - even though the amplitudes will become small
      - so the wavefunctions always overlap to some degree

path a

path b

Scattering of identical particles

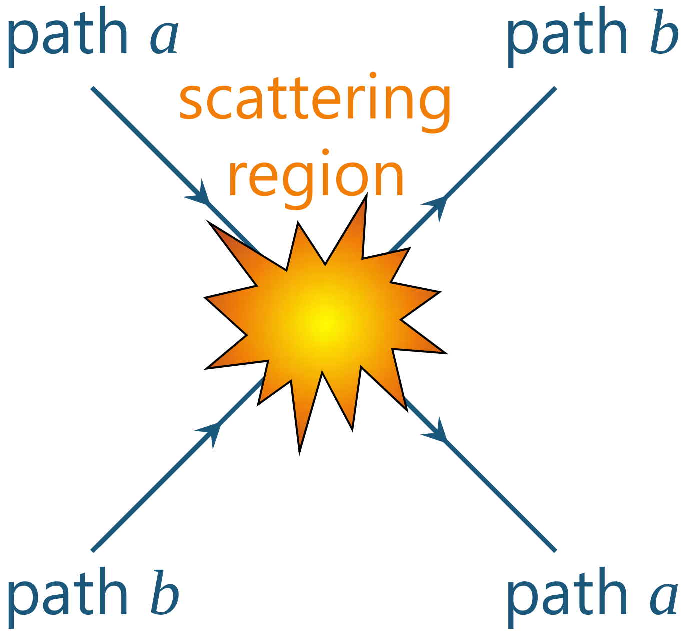
We may find the following argument more convincing

if we imagine that the wavepackets are initially directed towards one another

and that these wavepackets substantially overlap

for some period of time

as they "bounce" off one another



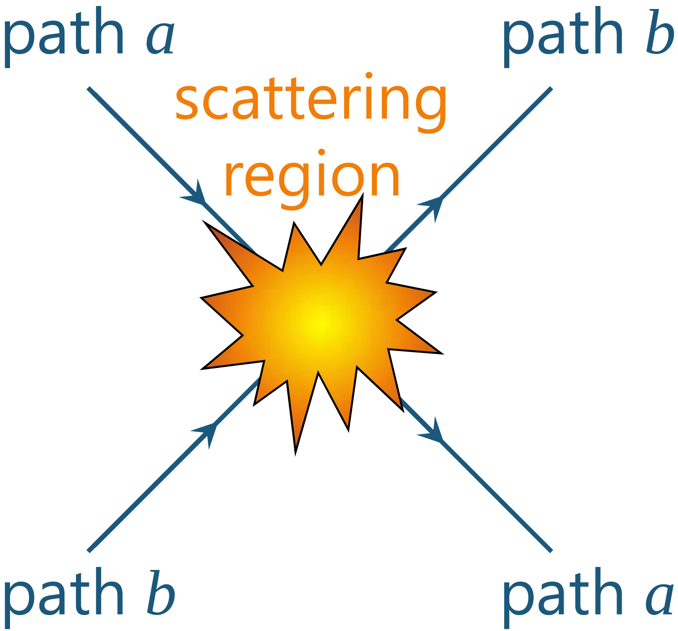
Scattering of identical particles

Now, certainly on the right of the scattering region

- when we measure the electrons
  - possibly finding one near path a
  - and another near path b

because two electrons are absolutely identical

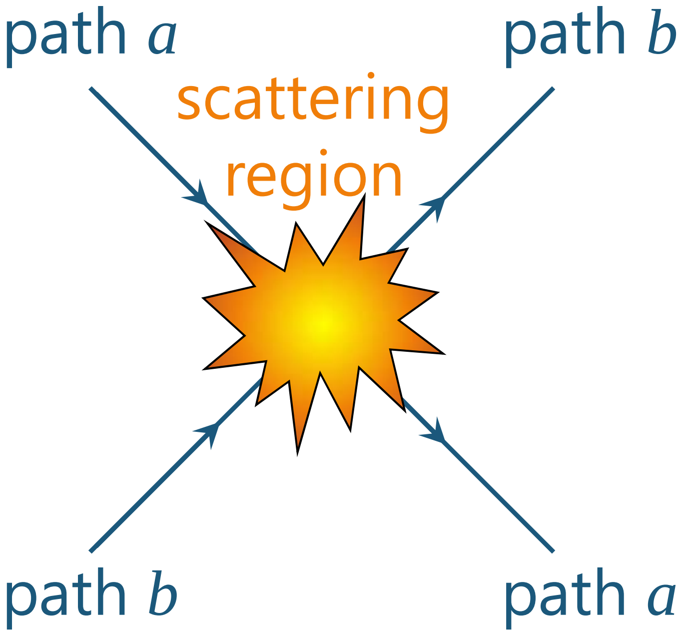
- we have absolutely no way of knowing whether it is
  - electron 1 or electron 2 that we find near any particular path



Scattering of identical particles

We might have good reason to believe

- because of our understanding of the scattering process
  - that if electron 1 started out on path a on the left
    - it is unlikely electron 1 emerged onto path b on the right
      - but we have to accept that it is possible



Scattering of identical particles

Let us write the wavefunction $\psi_a(\mathbf{r})$

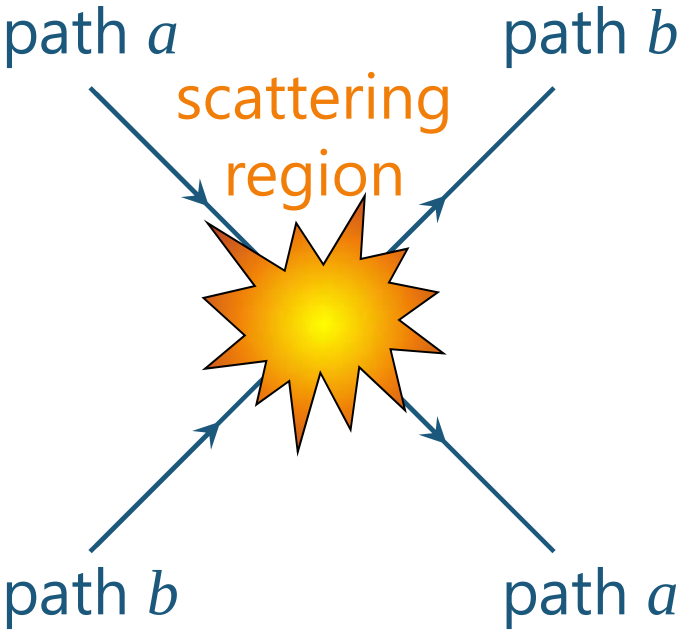
associated with path a

at least on the right of the scattering region

and at some particular time

and similarly write $\psi_b(\mathbf{r})$

for the corresponding wavefunction on path b



Scattering of identical particles

Hence, we might expect that the

two particle wavefunction $\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2)$

on the right can be written as

some linear combination of the

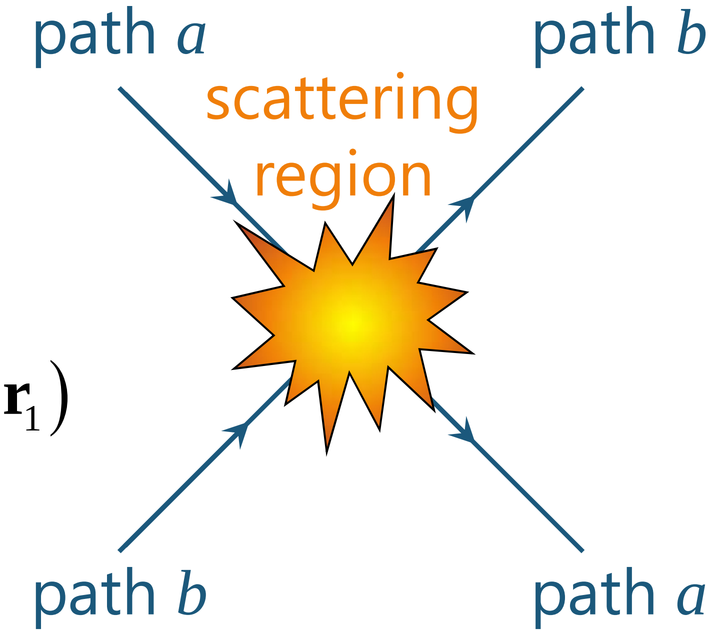
two possible outcomes

$$\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2) = c_{12}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + c_{21}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)$$

where c_{12} is the amplitude for

the outcome that electron 1 is on path a and electron 2 is on path b

and oppositely for amplitude c_{21}



Scattering of identical particles

But we believe electrons to be absolutely identical

so it can make no difference to any measurable outcome if we swap the electrons

We cannot measure the wavefunction itself

but we do expect to be able to measure $|\psi_{tp}|^2$

Swapping the electrons changes

$\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)$ into $\psi_{tp}(\mathbf{r}_2,\mathbf{r}_1)$

and so we conclude that

$$\left|\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)\right|^2 = \left|\psi_{tp}(\mathbf{r}_2,\mathbf{r}_1)\right|^2$$

Scattering of identical particles

Now $\left|\psi_{tp}\left(\mathbf{r}_1,\mathbf{r}_2\right)\right|^2 = \left|\psi_{tp}\left(\mathbf{r}_2,\mathbf{r}_1\right)\right|^2$

means that

$$\psi_{tp}\left(\mathbf{r}_2,\mathbf{r}_1\right) = \gamma\psi_{tp}\left(\mathbf{r}_1,\mathbf{r}_2\right)$$

where γ is some complex number of unit magnitude

We could of course swap the particles again

Since the particles are absolutely identical

we expect this swapping process produces exactly the same result, and so

$$\psi_{tp}\left(\mathbf{r}_1,\mathbf{r}_2\right) = \gamma\psi_{tp}\left(\mathbf{r}_2,\mathbf{r}_1\right)$$

Scattering of identical particles

With both $\psi_{tp}(\mathbf{r}_2,\mathbf{r}_1)=\gamma\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)$ and $\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=\gamma\psi_{tp}(\mathbf{r}_2,\mathbf{r}_1)$

we conclude $\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=\gamma^2\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)$

so, presuming (or postulating) the wavefunction should be restored on this double swap

$$\gamma^2=1$$

So we have only two possibilities for γ

$\gamma=1$ or $\gamma=-1$

hence

$$\boxed{\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=\pm\psi_{tp}(\mathbf{r}_2,\mathbf{r}_1)}$$

Scattering of identical particles

Now we can substitute our general linear combination

$$\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=c_{12}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)+c_{21}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)$$

in $\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=\pm\psi_{tp}(\mathbf{r}_2,\mathbf{r}_1)$ to get

$$\begin{aligned}c_{12}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)+c_{21}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\\ =\pm\left(c_{21}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)+c_{12}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right)\end{aligned}$$

Rearranging, we have

$$\begin{aligned}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)\left[c_{12}\mp c_{21}\right]&=\left[\pm c_{12}-c_{21}\right]\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\\ &=\pm\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\left[c_{12}\mp c_{21}\right]\end{aligned}$$

Scattering of identical particles

But $\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)[c_{12} \mp c_{21}] = \pm\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)[c_{12} \mp c_{21}]$

must hold for all $\mathbf{r}_1$

and in general $\psi_a(\mathbf{r}_1) \neq \psi_b(\mathbf{r}_1)$

since they represent different and largely separate wavepackets

and so we must have

$$c_{12} \mp c_{21} = 0$$

So in $\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2) = c_{12}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + c_{21}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)$

we must have $\boxed{c_{12} = \pm c_{21}}$

Scattering of identical particles

So given that the electrons emerge on paths a and b

we have shown that there are only two possibilities

for the nature of the wavefunction on the right of the scattering volume

Either

$$\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=c\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)+\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right]$$

or

$$\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=c\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)-\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right]$$

where c is in general some complex constant

Scattering of identical particles

We have therefore proved that, on the right

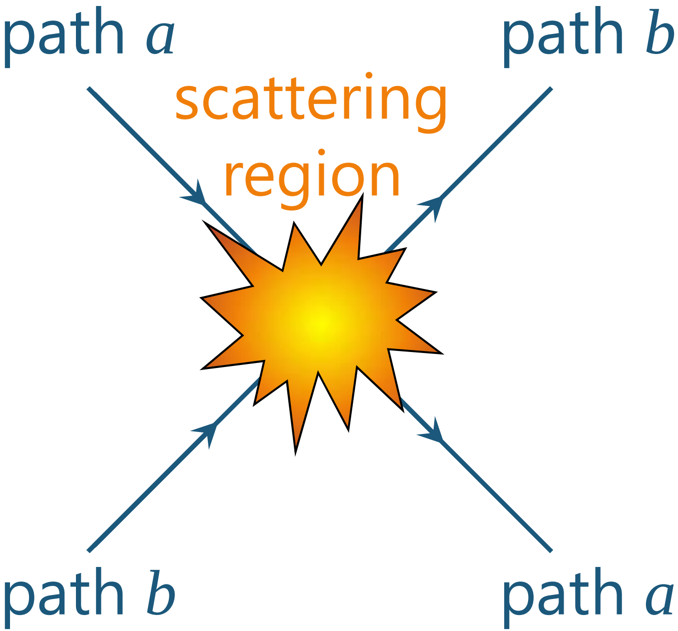
the amplitudes of

the function $\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)$

and the function $\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)$

are equal in magnitude

though possibly opposite in sign



Scattering of identical particles

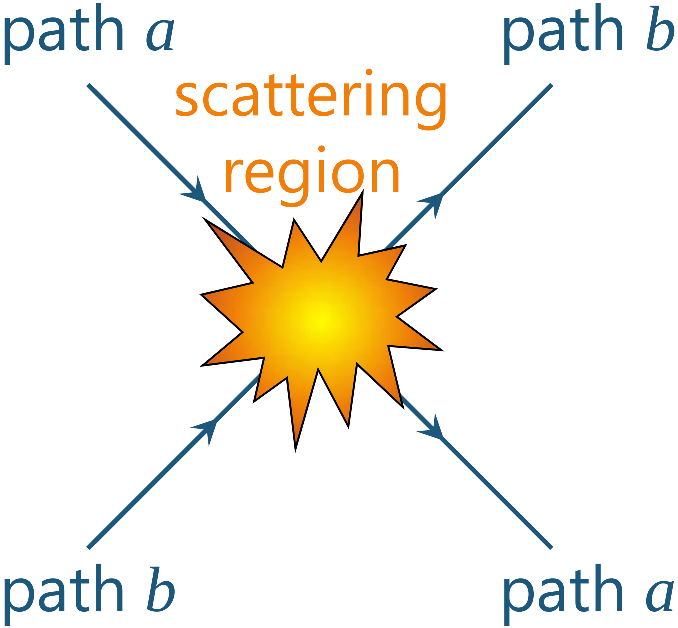
But, we might say

for the electron on path a on the left

the scattering probability into path a on the right

is in general different from

the scattering probability into path b on the right



Scattering of identical particles

How therefore can we have

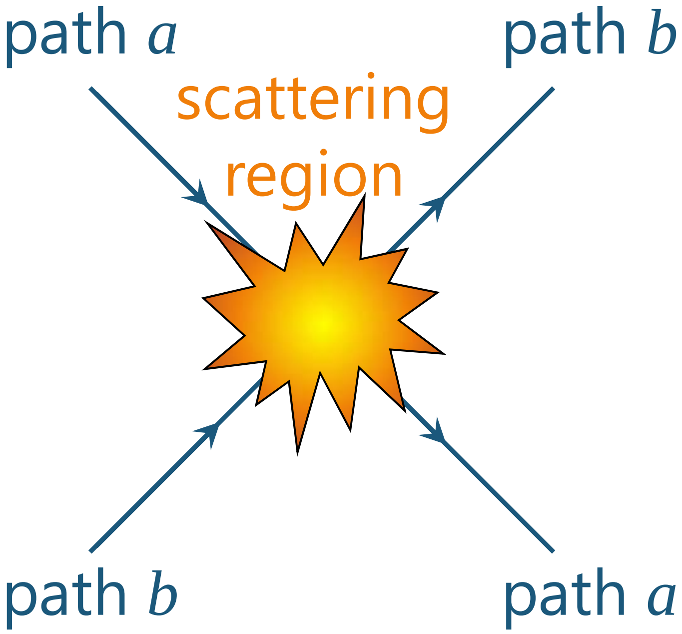
the amplitudes of the two possibilities on the right

$$\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)$$

and

$$\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)$$

being equal in magnitude?



Scattering of identical particles

The resolution of this apparent problem is that

even on the left of the scattering volume

at some time before the scattering

the wavefunction

$$\psi_{tpbefore}\left(\mathbf{r}_1,\mathbf{r}_2\right)$$

must also have had the two possibilities

$\psi_{abefore}\left(\mathbf{r}_1\right)\psi_{bbefore}\left(\mathbf{r}_2\right)$ and $\psi_{abefore}\left(\mathbf{r}_2\right)\psi_{bbefore}\left(\mathbf{r}_1\right)$

being equal in magnitude

Scattering of identical particles

Specifically, then

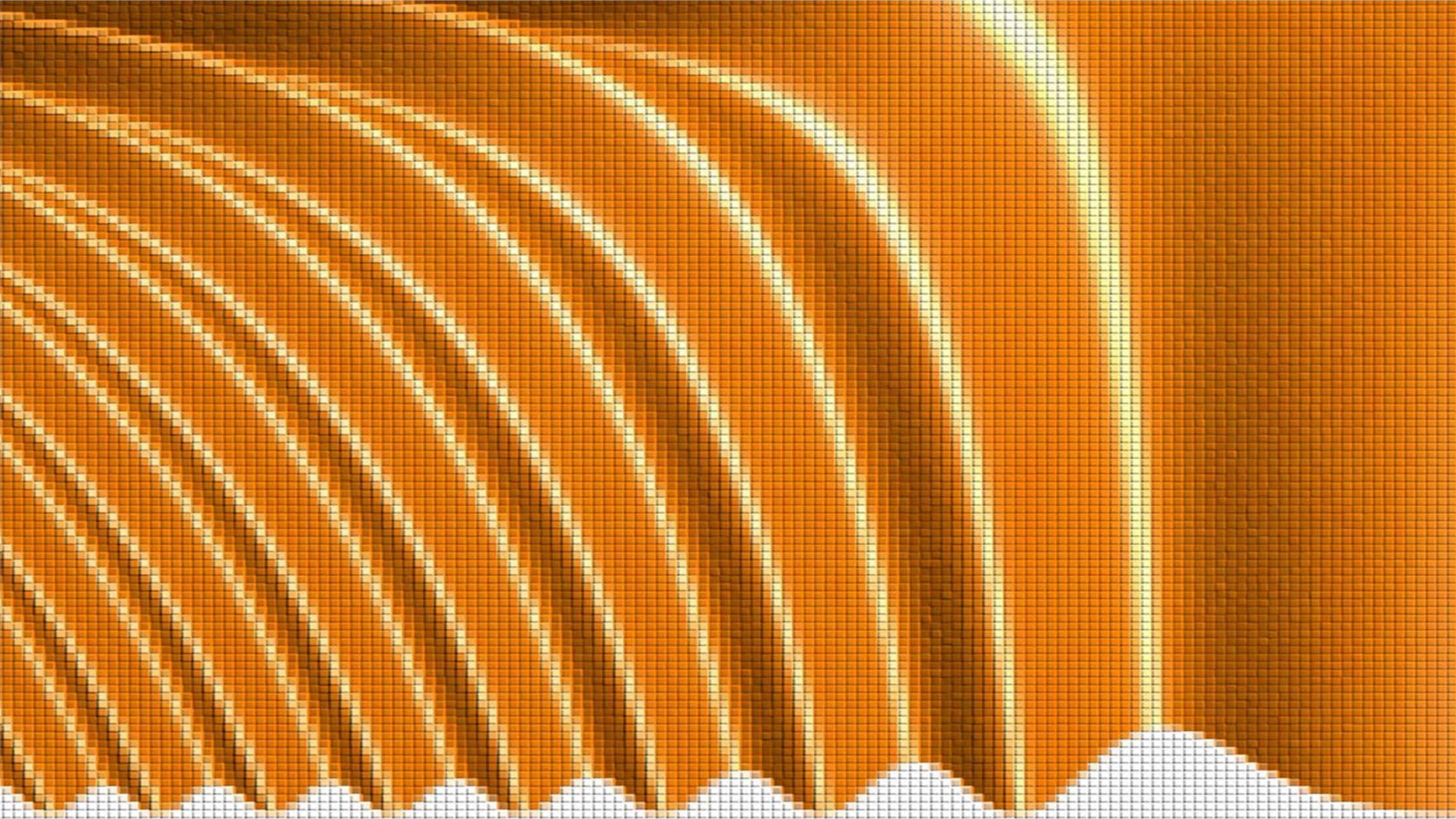
even before the interaction

the wavefunction must have been either

$$\psi_{tpbefore}\left(\mathbf{r}_1,\mathbf{r}_2\right)=c_{before}\left[\psi_{abefore}\left(\mathbf{r}_1\right)\psi_{bbefore}\left(\mathbf{r}_2\right)+\psi_{abefore}\left(\mathbf{r}_2\right)\psi_{bbefore}\left(\mathbf{r}_1\right)\right]$$

or

$$\psi_{tpbefore}\left(\mathbf{r}_1,\mathbf{r}_2\right)=c_{before}\left[\psi_{abefore}\left(\mathbf{r}_1\right)\psi_{bbefore}\left(\mathbf{r}_2\right)-\psi_{abefore}\left(\mathbf{r}_2\right)\psi_{bbefore}\left(\mathbf{r}_1\right)\right]$$

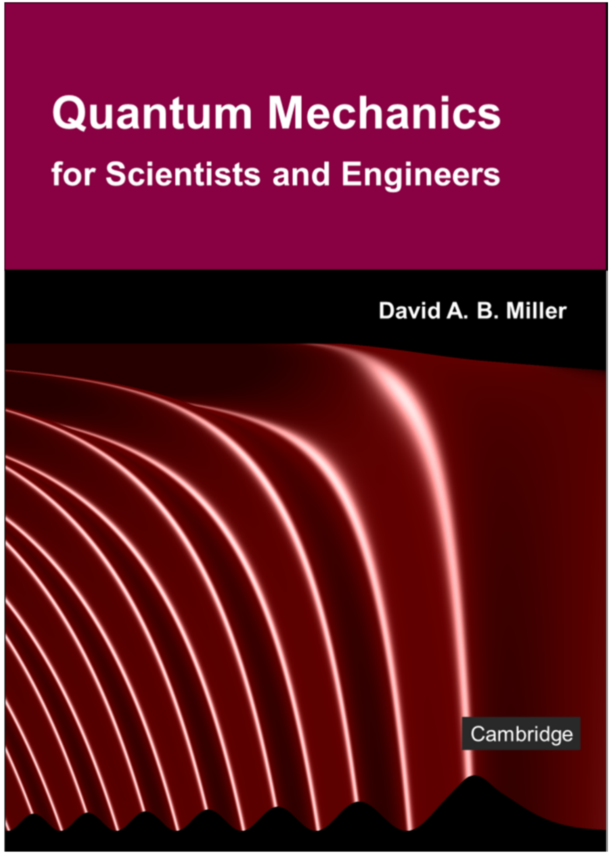


13.3 Identical particles and exchange

Slides: Video 13.3.3 Fermions and bosons

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.1 last 4 paragraphs, and 13.2



Identical particles and exchange

Fermions and bosons

Quantum mechanics for scientists and engineers

David Miller

Bosons

We find that a given kind of particle always corresponds

to only one of the possible choices of γ

All particles corresponding to

$$\gamma = +1$$

i.e., with a wavefunction for a pair of particles in the form

$$\psi_{tp}(\mathbf{r}_1, \mathbf{r}_2) = c\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right]$$

are called bosons

Photons and all particles with integer spin

including also, for example, ^{4}He nuclei

are bosons

Fermions

All particles corresponding to

$$\gamma = -1$$

i.e., a wavefunction for a pair of particles in the form

$$\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2) = c\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) - \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right]$$

are called fermions

Electrons, protons, neutrons

and all particles with half integer spin

are fermions

Pauli exclusion principle

For two fermions, we know the wavefunction is built from the form

$$\psi_{tp}\left(\mathbf{r}_1,\mathbf{r}_2\right)=c\left[\psi_a\left(\mathbf{r}_1\right)\psi_b\left(\mathbf{r}_2\right)-\psi_a\left(\mathbf{r}_2\right)\psi_b\left(\mathbf{r}_1\right)\right]$$

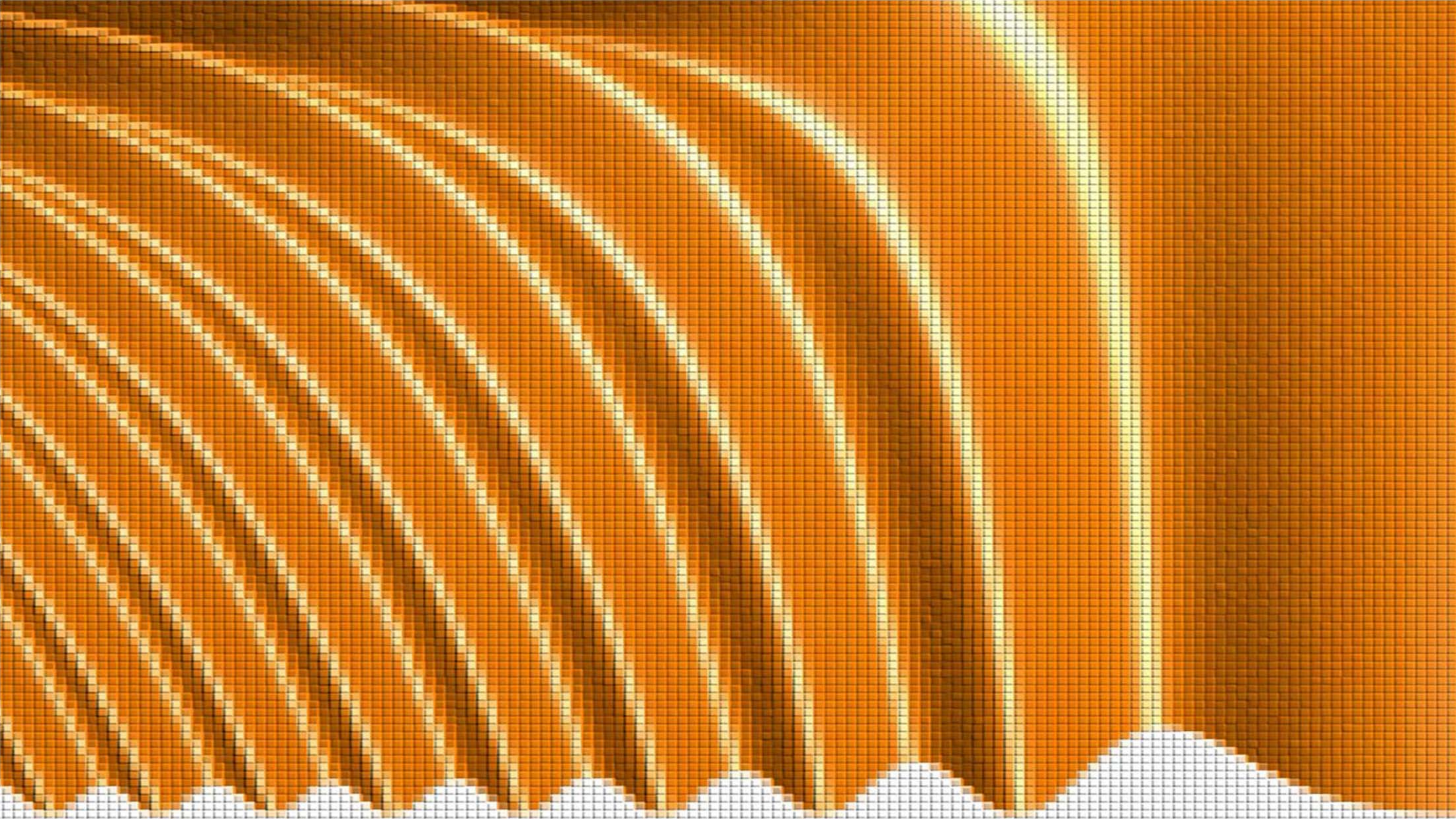
Suppose now that we postulate that the two fermions are in the same single-particle state

say, state a

Then the wavefunction becomes

$$\psi_{tp}\left(\mathbf{r}_1,\mathbf{r}_2\right)=c\left[\psi_a\left(\mathbf{r}_1\right)\psi_a\left(\mathbf{r}_2\right)-\psi_a\left(\mathbf{r}_2\right)\psi_a\left(\mathbf{r}_1\right)\right]=0$$

Note that this wavefunction is zero everywhere

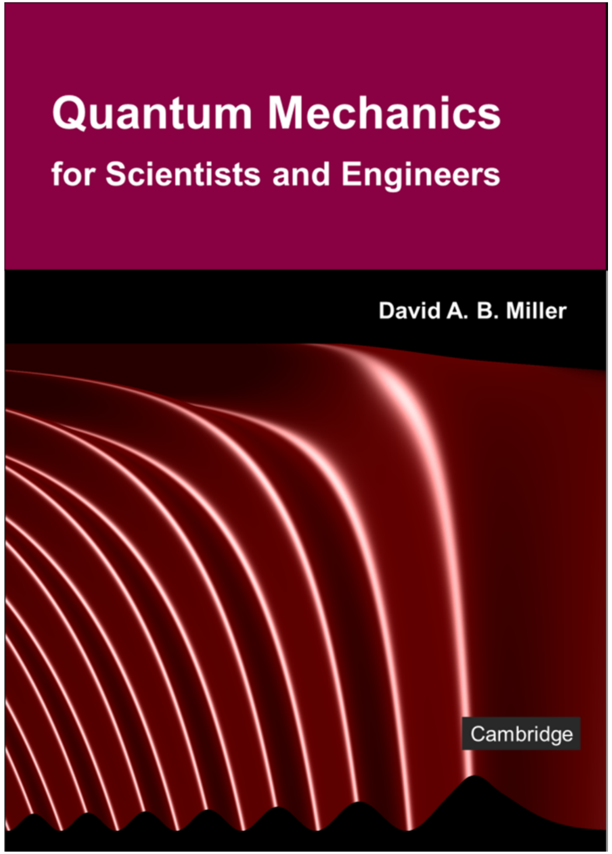


13.3 Identical particles and exchange

Slides: Video 13.3.5 States, single-particle states, and modes

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.3



Identical particles and exchange

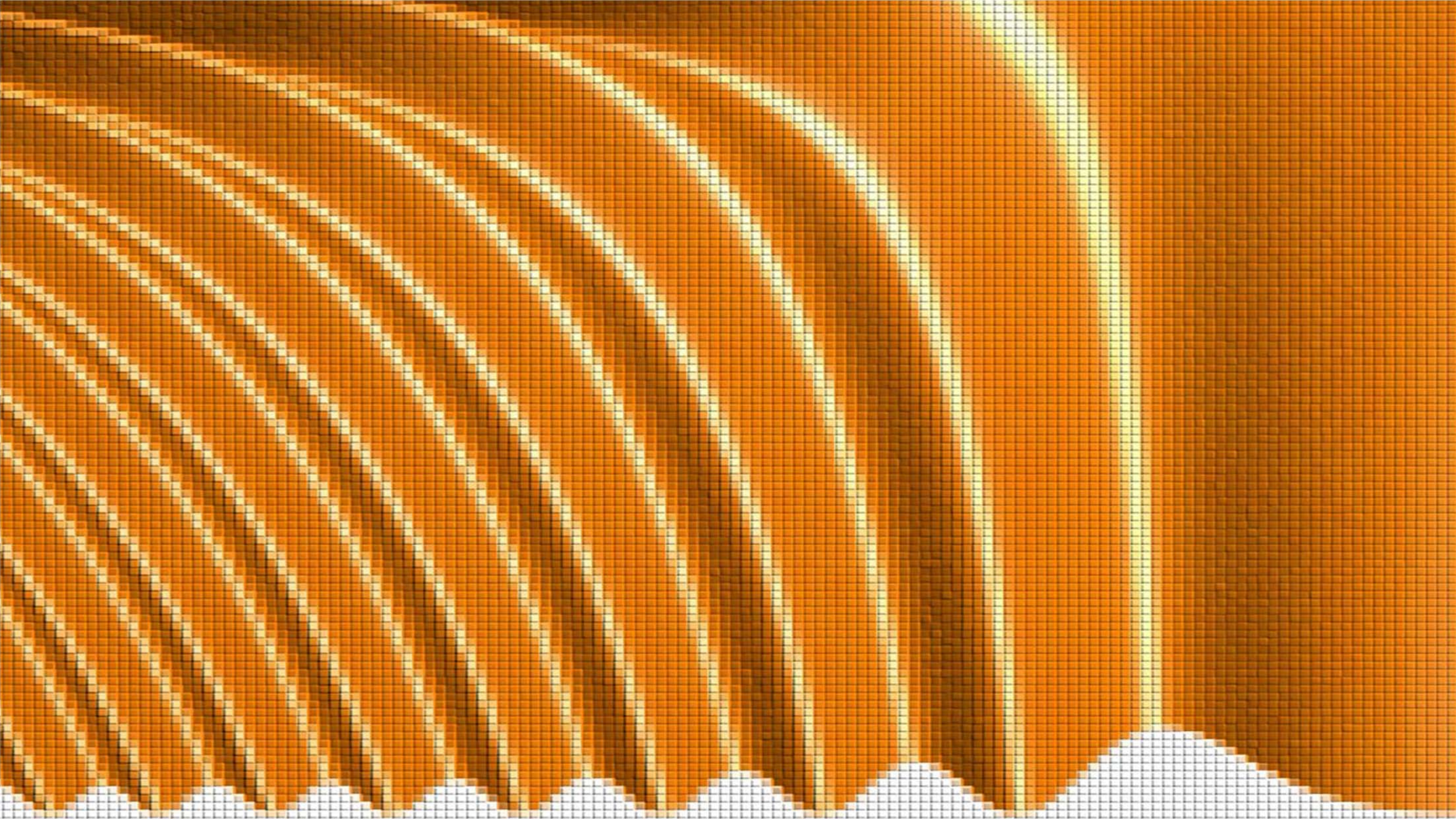
States, single-particle states, and modes

Quantum mechanics for scientists and engineers

David Miller

Fermion single particle states and the state of the system

Boson modes and the state of the system

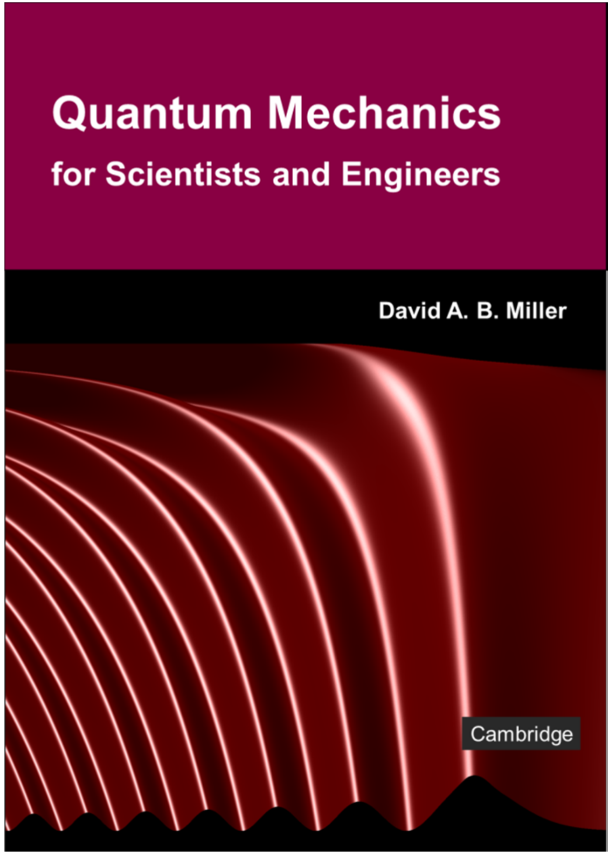


13.3 Identical particles and exchange

Slides: Video 13.3.7 Exchange energy

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.4



Identical particles and exchange

Exchange energy

Quantum mechanics for scientists and engineers

David Miller

Exchange energy

Suppose we have two electrons of identical spin

- They will certainly have a Coulomb repulsion
  - and so we could write the Hamiltonian similarly to the hydrogen atom
    - except here the two particles are identical
      - and the Coulomb potential is repulsive rather than attractive

The Hamiltonian is therefore

$$\hat{H} = -\frac{\hbar^2}{2m_o}\left(\nabla_{\mathbf{r}_1}^2 + \nabla_{\mathbf{r}_2}^2\right) + \frac{e^2}{4\pi\varepsilon_o\left|\mathbf{r}_1 - \mathbf{r}_2\right|}$$

Exchange energy

Because they are fermions

the state of the two particles is built on the form

$$\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=\left(1/\sqrt{2}\right)\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)-\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right]$$

where the individual wavefunctions $\psi_a(\mathbf{r})$ and $\psi_b(\mathbf{r})$ are normalized

and the factor $1/\sqrt{2}$ gives overall normalization

We can also write this in bra-ket notation as

$$|\psi_{tp}\rangle=\left(1/\sqrt{2}\right)\left(|1,a\rangle|2,b\rangle-|2,a\rangle|1,b\rangle\right)$$

where $|1,a\rangle\equiv\psi_a(\mathbf{r}_1)$ and so on

Order in wavefunction and bra or ket products

Note that the order of the products of the wavefunctions does not matter in expressions such as

$$\psi_{tp}(\mathbf{r}_1,\mathbf{r}_2)=\left(1/\sqrt{2}\right)\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)-\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right]$$

and $\left|\psi_{tp}\right\rangle=\left(1/\sqrt{2}\right)\left(\left|1,a\right\rangle\left|2,b\right\rangle-\left|2,a\right\rangle\left|1,b\right\rangle\right)$

Obviously

$$\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)=\psi_b(\mathbf{r}_2)\psi_a(\mathbf{r}_1)$$

since $\psi_a(\mathbf{r}_1)$ and $\psi_b(\mathbf{r}_2)$ are each simply a number for any given value of $\mathbf{r}_1$ or $\mathbf{r}_2$

Order in wavefunction and bra or ket products

For the case of the bra-ket notation

we can similarly state

$$|1,a\rangle|2,b\rangle = |2,b\rangle|1,a\rangle$$

Quite generally

the order of the statement of the vectors

corresponding to different degrees of freedom or dynamical variables

does not matter in direct product spaces

Exchange energy

Now for this two-electron state

$$|\psi_{tp}\rangle = \left(1/\sqrt{2}\right)\left(|1,a\rangle|2,b\rangle - |2,a\rangle|1,b\rangle\right)$$

we evaluate the energy expectation value

$$\langle E\rangle = \langle\psi_{tp}|\hat{H}|\psi_{tp}\rangle$$

I.e.,

$$\langle E\rangle = \frac{1}{2}\begin{bmatrix} \langle 1,a|\langle 2,b|\hat{H}|1,a\rangle|2,b\rangle + \langle 2,a|\langle 1,b|\hat{H}|2,a\rangle|1,b\rangle \\ -\langle 1,a|\langle 2,b|\hat{H}|2,a\rangle|1,b\rangle - \langle 2,a|\langle 1,b|\hat{H}|1,a\rangle|2,b\rangle \end{bmatrix}$$

The first two terms (which are actually equal)

have a straightforward meaning

Exchange energy

Formally evaluating, we have for the first term

$$\langle 1,a|\langle 2,b|\hat{H}|1,a\rangle|2,b\rangle$$

$$=\langle 1,a|\langle 2,b|\left(-\frac{\hbar^2}{2m_o}\left(\nabla^2_{\mathbf{r}_1}+\nabla^2_{\mathbf{r}_2}\right)+\frac{e^2}{4\pi\varepsilon_o|\mathbf{r}_1-\mathbf{r}_2|}\right)|1,a\rangle|2,b\rangle$$

$$=\langle 1,a|\langle 2,b|-\frac{\hbar^2}{2m_o}\nabla^2_{\mathbf{r}_1}|1,a\rangle|2,b\rangle+\langle 1,a|\langle 2,b|-\frac{\hbar^2}{2m_o}\nabla^2_{\mathbf{r}_2}|1,a\rangle|2,b\rangle$$

$$+\langle 1,a|\langle 2,b|\frac{e^2}{4\pi\varepsilon_o|\mathbf{r}_1-\mathbf{r}_2|}|1,a\rangle|2,b\rangle=E_{KEa}+E_{KEb}+E_{PEab}$$

Exchange energy

Here, E_{KEa} is the kinetic energy of an electron in single-particle state a

$$E_{KEa} = \langle 1,a| \langle 2,b| -\frac{\hbar^2}{2m_o}\nabla_{\mathbf{r}_1}^2 |1,a\rangle |2,b\rangle = \langle 1,a| -\frac{\hbar^2}{2m_o}\nabla_{\mathbf{r}_1}^2 |1,a\rangle \langle 2,b|2,b\rangle$$

$$= -\frac{\hbar^2}{2m_o}\int \psi_a^*(\mathbf{r})\nabla^2\psi_a(\mathbf{r})\,d^3\mathbf{r}$$

Note that $\langle 2,b|2,b\rangle = 1$ because the single particle wavefunctions are normalized

Similarly $E_{KEb} = -\frac{\hbar^2}{2m_o}\int \psi_b^*(\mathbf{r})\nabla^2\psi_b(\mathbf{r})\,d^3\mathbf{r}$

Exchange energy

The final contribution, E_{PEab}, is the Coulomb potential energy from the interaction of the

charge density from one electron in single-particle state a and the other in single-particle state b

$$E_{PEab} = \langle 1,a| \langle 2,b| \frac{e^2}{4\pi\varepsilon_o |\mathbf{r}_1 - \mathbf{r}_2|} |1,a\rangle |2,b\rangle = e^2 \int \frac{|\psi_a(\mathbf{r})|^2 |\psi_b(\mathbf{r}')|^2}{4\pi\varepsilon_o |\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}'$$

So all parts of $\langle 1,a| \langle 2,b| \hat{H} |1,a\rangle |2,b\rangle = E_{KEa} + E_{KEb} + E_{PEab}$

are simple kinetic and potential energies

and similarly for $\langle 2,a| \langle 1,b| \hat{H} |2,a\rangle |1,b\rangle$

Exchange energy

So the two terms

$$\frac{1}{2}\left[\langle 1,a|\langle 2,b|\hat{H}|1,a\rangle|2,b\rangle+\langle 2,a|\langle 1,b|\hat{H}|2,a\rangle|1,b\rangle\right]=E_{KEa}+E_{KEb}+E_{PEab}$$

give us the energy we expect semiclassically

the kinetic energies of the two particles and the potential energy from their interaction

But there are two more terms, on the bottom line of

$$\langle E\rangle=\frac{1}{2}\begin{bmatrix}\langle 1,a|\langle 2,b|\hat{H}|1,a\rangle|2,b\rangle+\langle 2,a|\langle 1,b|\hat{H}|2,a\rangle|1,b\rangle \\ -\langle 1,a|\langle 2,b|\hat{H}|2,a\rangle|1,b\rangle-\langle 2,a|\langle 1,b|\hat{H}|1,a\rangle|2,b\rangle\end{bmatrix}$$

These give what is called the exchange energy

an energy term with no classical analog

Exchange energy

We note that, by the Hermiticity of the Hamiltonian

$$\langle 2,a|\langle 1,b|\hat{H}|1,a\rangle|2,b\rangle = \left[\langle 1,a|\langle 2,b|\hat{H}|2,a\rangle|1,b\rangle\right]^*$$

and so the exchange energy can be written

$$E_{EXab} = -\frac{1}{2}\left(\langle 1,a|\langle 2,b|\hat{H}|2,a\rangle|1,b\rangle + \left[\langle 1,a|\langle 2,b|\hat{H}|2,a\rangle|1,b\rangle\right]^*\right)$$

$$= -\mathrm{Re}\left[\int \psi_a^*(\mathbf{r}_1)\psi_b^*(\mathbf{r}_2)\hat{H}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)d^3\mathbf{r}_1 d^3\mathbf{r}_2\right]$$

and finally for the total expectation value of the energy of these two electrons in this state

$$\langle E\rangle = E_{KEa} + E_{KEb} + E_{PEab} + E_{EXab}$$

Validity of single particle calculations

If the function $\psi_a(\mathbf{r})$ is only substantial in a region near to some point $\mathbf{r}_a$

then so also is the function $\nabla^2\psi_a(\mathbf{r})$

Similarly, if the function $\psi_b(\mathbf{r})$ is only significant near to some point $\mathbf{r}_b$

then so also is the function $\nabla^2\psi_b(\mathbf{r})$

So, if the points $\mathbf{r}_a$ and $\mathbf{r}_b$ are far enough apart that there is negligible overlap of the functions $\psi_a(\mathbf{r})$ and $\psi_b(\mathbf{r})$

$$\int \psi_a^*(\mathbf{r}_1)\nabla_{\mathbf{r}_1}^2\psi_b(\mathbf{r}_1)\,d^3\mathbf{r}_1 \simeq 0 \text{ and } \int \psi_b^*(\mathbf{r}_2)\nabla_{\mathbf{r}_2}^2\psi_a(\mathbf{r}_2)\,d^3\mathbf{r}_2 \simeq 0$$

Validity of single particle calculations

Similarly, for such negligible overlap

regardless of the form of the potential energy $V(\mathbf{r}_1,\mathbf{r}_2)$

$$\int \psi_a^*(\mathbf{r}_1)\psi_b^*(\mathbf{r}_2)V(\mathbf{r}_1,\mathbf{r}_2)\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)d^3\mathbf{r}_1 d^3\mathbf{r}_2 \simeq 0$$

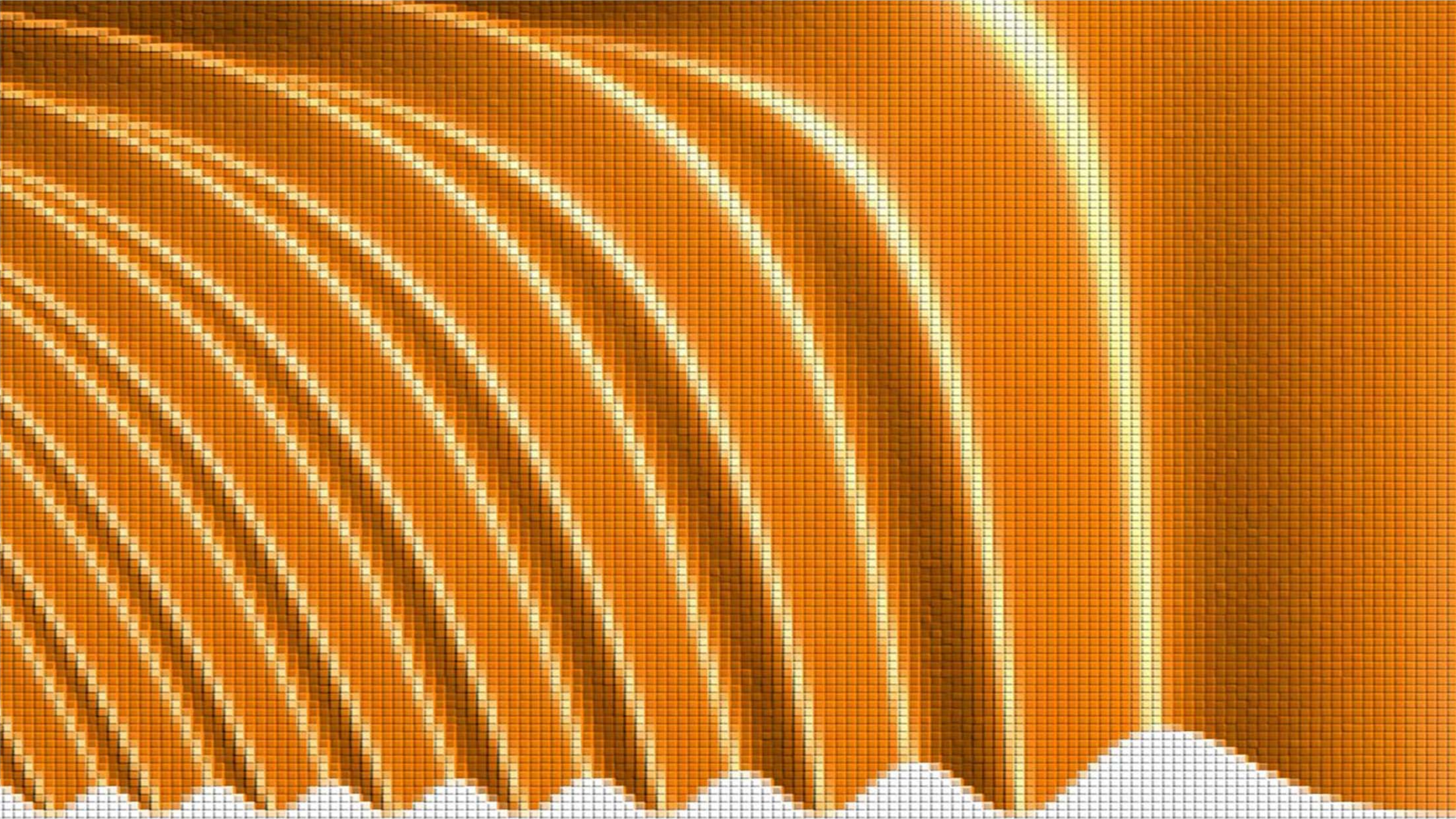
simply because the functions $\psi_a(\mathbf{r})$ and $\psi_b(\mathbf{r})$

do not overlap

Hence

there is only a contribution to the exchange energy

if the individual particle wavefunctions overlap

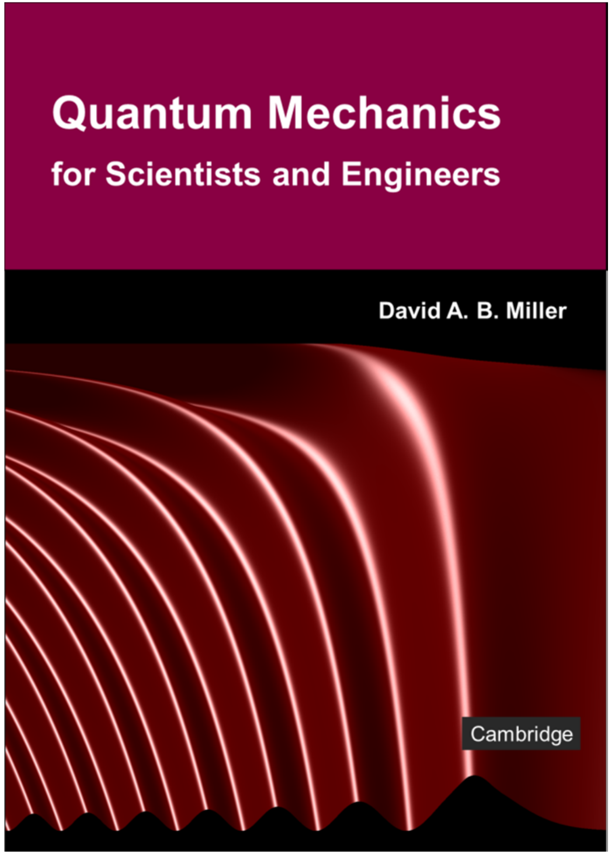


14.1 States of identical particles

Slides: Video 14.1.1 Multiple particle states

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.5



States of identical particles

Multiple particle states

Quantum mechanics for scientists and engineers

David Miller

Extension to more than two non-identical particles

If we had N different (i.e., not identical) particles

that were approximately not interacting

at least in some region of space and time

(e.g., substantially before or after the scattering)

then we could construct the state $|\Psi_{different}\rangle$ for those by

simply multiplying the single-particle states or modes

$$|\psi_{different}\rangle = |1,a\rangle|2,b\rangle|3,c\rangle\ldots|N,n\rangle$$

where the numbers and the letter N refer to the particles

and the small letters refer to the single-particle state

the individual particles are in

More than two bosons

We can write the state as

$$\left|\psi_{\text{identical bosons}}\right\rangle \propto \sum_{\hat{P}} \hat{P} \left| |1,a\rangle |2,b\rangle |3,c\rangle \ldots |N,n\rangle \right\rangle$$

Here $\hat{P}$ is one of the permutation operators

- This is an operator that changes one function in the Hilbert space into another
  - in this case by permuting the particles among the modes

The meaning of the sum is that it is taken over all of those possible distinct permutation operators

More than two bosons

The notation here $\left|\psi_{\textit{identical bosons}}\right\rangle \propto \sum_{\hat{P}} \hat{P}\left|\left|1,a\right\rangle\left|2,b\right\rangle\left|3,c\right\rangle\ldots\left|N,n\right\rangle\right\rangle$

is just a mathematical way of saying we are summing

over all permutations of the N particles

among the chosen set of modes

Incidentally, for this boson case

it is quite allowable for two or more of the modes to be the same mode

e.g., for mode b to be the same mode as mode a

an important and general property of bosons

More than two bosons

Note that, for any given set of modes a, b, c, …n

with given numbers of these bosons in each mode

there is only one possible such boson state of identical particles

The state $\left|\psi_{\text{identical bosons}}\right\rangle \propto \sum_{\hat{P}} \hat{P}\left| |1,a\rangle |2,b\rangle |3,c\rangle \ldots |N,n\rangle \right\rangle$

satisfies the symmetry requirement that

swapping any two particles does not change the sign or amplitude of the state

Swapping particles just corresponds to changing the order of the terms, leaving the sum itself unchanged

More than two fermions

We can write the state for   identical fermions as

$$\left|\psi_{identical\ \ fermions}\right\rangle = \frac{1}{\sqrt{N!}}\sum_{\hat{P}=1}^{N!} \pm\hat{P}\left|\left|1,a\right\rangle\left|2,b\right\rangle\left|3,c\right\rangle\ldots\left|N,n\right\rangle\right\rangle$$

where now by $\pm\hat{P}$ we mean that

we use the + sign

when the permutation corresponds to an even number of pair-wise swaps of the individual particles

and the – sign

when the permutation corresponds to an odd number of pair-wise swaps of the individual particles

More than two fermions

Note that for this state

$$\left|\psi_{identical\ \ fermions}\right\rangle = \frac{1}{\sqrt{N!}}\sum_{\hat{P}=1}^{N!} \pm\hat{P}\left|\left|1,a\right\rangle\left|2,b\right\rangle\left|3,c\right\rangle\ldots\left|N,n\right\rangle\right\rangle$$

- if two of the single-particle states are identical
  - e.g., if $b = a$
    - then the fermion state is exactly zero because
      - for each permutation there is an identical one with opposite sign that exactly cancels it

This is the extension of the Pauli exclusion principle to N particles

Slater determinant

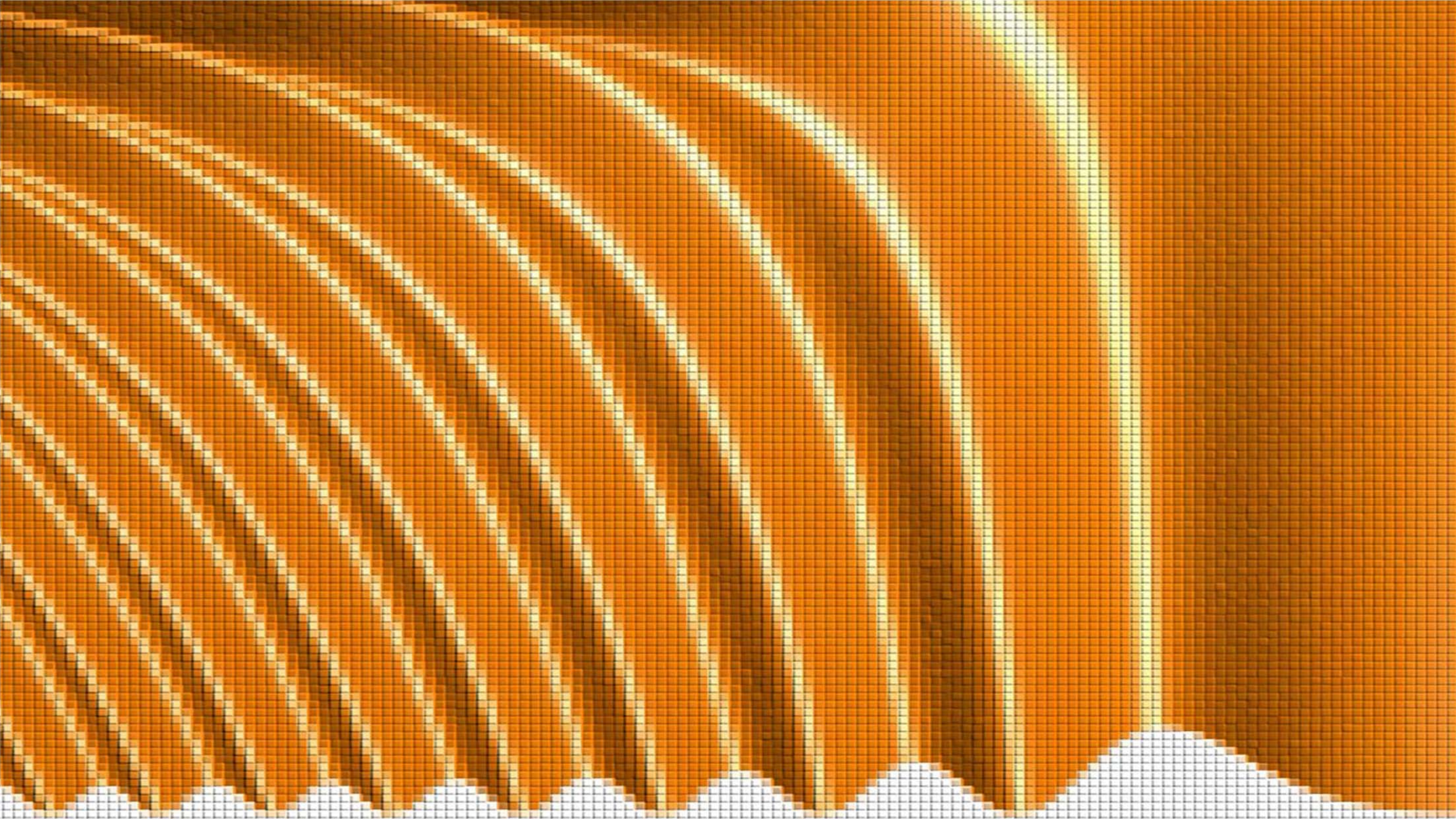
There is a particularly convenient way to write the N particle fermion state

which is called the Slater determinant

$$\left|\psi_{identical\ \ fermions}\right\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} |1,a\rangle & |2,a\rangle & \cdots & |N,a\rangle \\ |1,b\rangle & |2,b\rangle & \cdots & |N,b\rangle \\ \vdots & \vdots & \ddots & \vdots \\ |1,n\rangle & |2,n\rangle & \cdots & |N,n\rangle \end{vmatrix}$$

This is just another way of writing

$$\left|\psi_{identical\ \ fermions}\right\rangle = \frac{1}{\sqrt{N!}}\sum_{\hat{P}=1}^{N!} \pm\hat{P}\left|\,|1,a\rangle|2,b\rangle|3,c\rangle\ldots|N,n\rangle\right\rangle$$

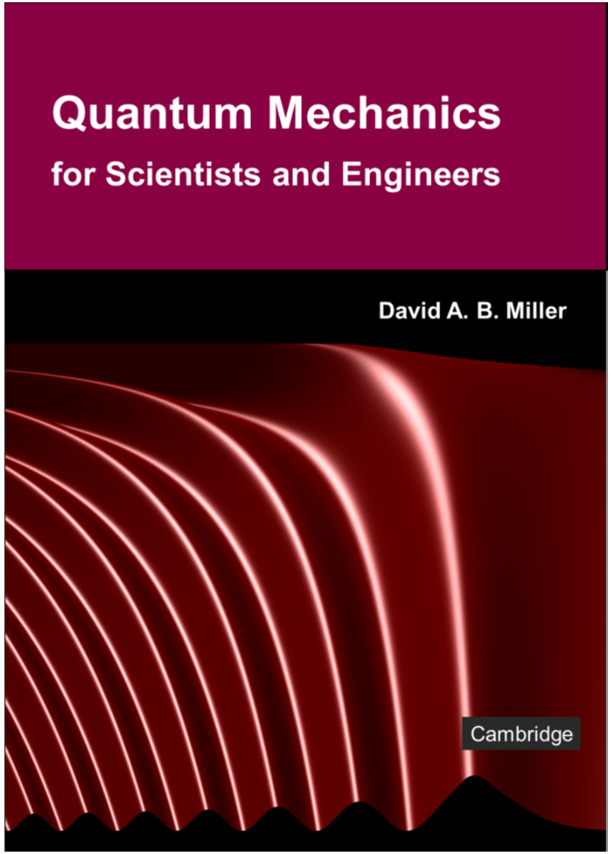


14.1 States of identical particles

Slides: Video 14.1.3 Multiple particle basis functions

Text reference: Quantum Mechanics for Scientists and Engineers

Sections 13.6 up to Eq. 13.50



States of identical particles

Multiple particle basis functions

Quantum mechanics for scientists and engineers

David Miller

Multiple particle basis functions

So we can find some complete basis set

to represent one of the particles

$$\psi_i\left(\mathbf{r}_j\right) \equiv \left| j, i \right\rangle$$

and then formally construct a new basis set

$$\Psi_{ab\cdots n}\left(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N\right) \equiv \left| \Psi_{ab\cdots n} \right\rangle$$

for the N particle system

from products of single particle functions

appropriately symmetrized with respect to exchange

Multiple particle basis functions

Depending on the symmetry with respect to exchange

there are different forms for this basis function

(i) non-identical particles

$$\psi_{ab\cdots n}\left(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_N\right)=\psi_a\left(\mathbf{r}_1\right)\psi_b\left(\mathbf{r}_2\right)\cdots\psi_n\left(\mathbf{r}_N\right)$$

or equivalently $\left|\Psi_{ab\cdots n}\right\rangle=\left|1,a\right\rangle\left|2,b\right\rangle\cdots\left|N,n\right\rangle$

where each of the $\psi_a\left(\mathbf{r}\right)$ may be chosen to be any of the single particle basis functions $\psi_i\left(\mathbf{r}\right)$

(ii) identical bosons $\left|\Psi_{ab\cdots n}\right\rangle\propto\sum_{\hat{P}}\hat{P}\left|1,a\right\rangle\left|2,b\right\rangle\cdots\left|N,n\right\rangle$

(iii) identical fermions $\left|\Psi_{ab\cdots n}\right\rangle=\dfrac{1}{\sqrt{N!}}\sum_{\hat{P}=1}^{N!}\pm\hat{P}\left|1,a\right\rangle\left|2,b\right\rangle\cdots\left|N,n\right\rangle$

Number of basis functions – non-identical particles

For non-identical particles

- there is one basis function for every choice of combination of single particle basis functions

If we imagined there were M possible single particle basis functions

- and there are N particles
  - then there are in general M^N such basis functions for the N particle system

Number of orthogonal states – non-identical particles

So, for N non-identical particles

specifying a state of that N particle system

involves specifying a linear combination

of the M^N different orthogonal N-particle basis functions

Because there are only M^N different orthogonal N-particle basis functions

there can only be M^N different orthogonal N-particle states

even if we now allow them to interact

Number of orthogonal states – non-identical particles

So, for N non-identical particles

even allowing them to interact

there are only M^N possible orthogonal N-particle states

each of which will be a different combination of the M^N different N-particle basis functions

Number of orthogonal states of N non-identical particles,
with M available single-particle states or modes,

$$= M^N$$

Number of distinct basis functions - bosons

In the case of identical bosons

the N-particle basis states

corresponding to different permutations

of the same set of choices of basis modes

are not distinct

and so there are fewer basis states than for non-identical particles

Number of distinct basis functions - bosons

For example, we see that in

$$\left|\Psi_{ab\cdots n}\right\rangle \propto \sum_{\hat{P}} \hat{P}\left|1,a\right\rangle\left|2,b\right\rangle\cdots\left|N,n\right\rangle$$

the state $\left|\Psi_{ab\cdots n}\right\rangle$ is not distinct from $\left|\Psi_{ba\cdots n}\right\rangle$

Since all permutations of the products of basis modes

are already in the sum

these two states are the same sum of products

performed in a different order

Number of distinct basis functions - bosons

The counting of these boson states is complicated, but

- it corresponds to a standard result in permutations and combinations, which is the problem of
  - counting the number of combinations of M things
    - here the single particle states or modes
  - taken N at a time
    - since we always have N particles
  - with repetitions allowed
    - i.e., we can have more than one particle in a mode
  - with the standard result $(M+N-1)!/\left[N!(M-1)!\right]$

Number of distinct basis functions - bosons

For example

- think of M boxes
  - each containing as many blocks as we like of just one color
    - with each box containing a different color of blocks
- We are picking N blocks altogether from these boxes
  - The number of possible different combinations of blocks we can end up with is

$$(M+N-1)!/\left[N!(M-1)!\right]$$

Number of distinct basis functions - bosons

For example

the set of combinations of

2 particles among

3 modes, a, b, and c

allowing repetitions is

ab, ac, bc, aa, bb, cc

giving six in all

which corresponds to

$$(3+2-1)!/\left[2!(3-1)!\right]=6$$

Number of orthogonal states - bosons

Just as for the non-identical particle case

this number of basis states is also

the number of different orthogonal states we can have

for the set of identical boson particles

even if we allow interactions

Number of orthogonal states of N identical bosons,

$$\text{with } M \text{ available modes} = \frac{(M+N-1)!}{N!(M-1)!}$$

Number of distinct basis functions - fermions

Specifically, if there are M choices for the first basis single-particle state a in $\left|\Psi_{ab\cdots n}\right\rangle$

then there are $M - 1$ choices for the second single particle basis state b, and so on down to

$M - N + 1$ choices for the last single particle basis state n

Hence, instead of M^N initial choices, we have only

$$M\left(M-1\right)\cdots\left(M-N+1\right) = M!/\left(M-N\right)!$$

Since the order of the choice of states does not matter

we divide by the number of different orders, $N!$

Number of orthogonal states - fermions

Hence in the identical fermion case

there are $M!/\left[(M-N)!N!\right]$ possible basis states

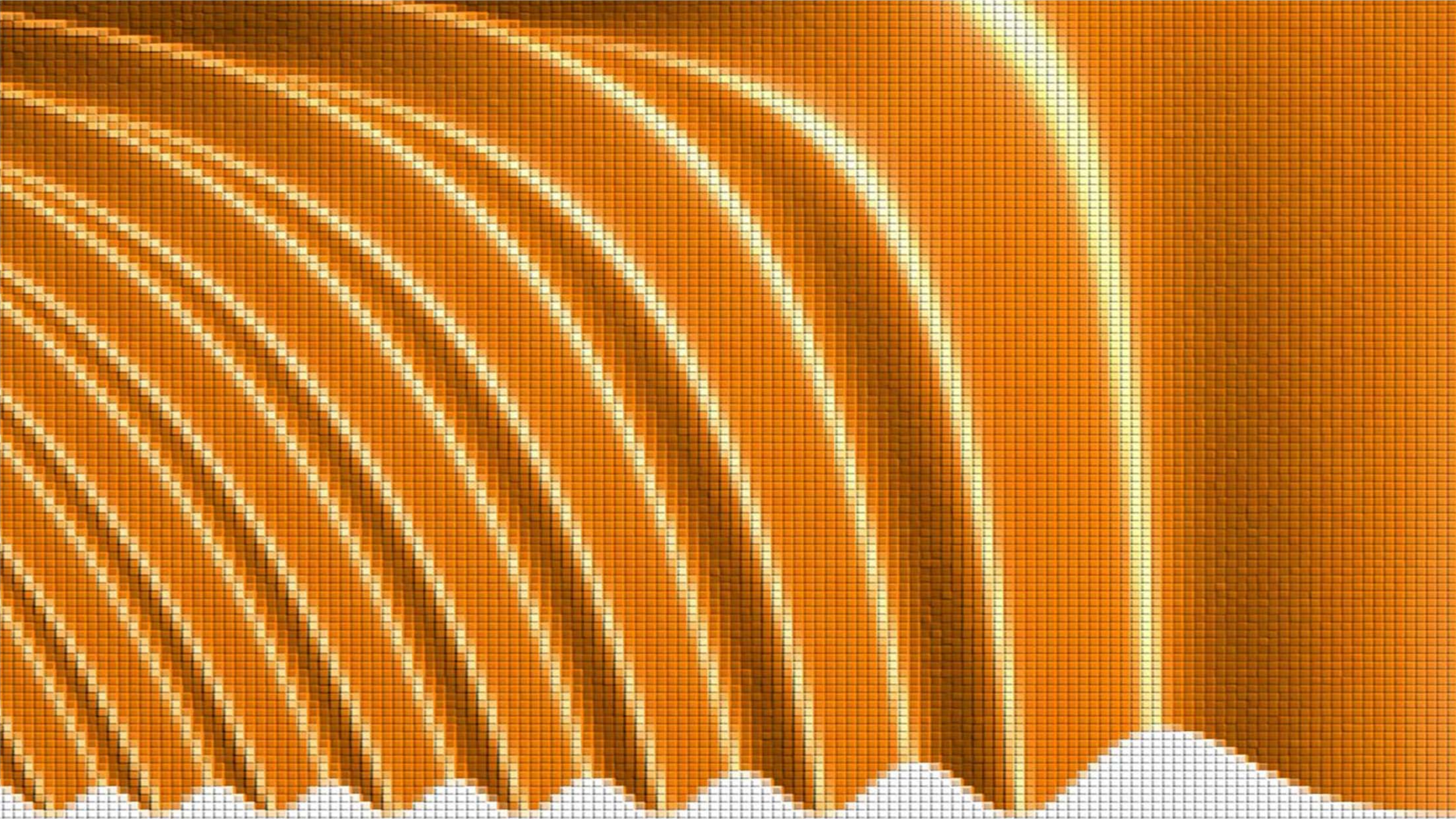
and hence the same number of possible orthogonal states altogether

even if we allow interactions between particles

I.e.,

Number of orthogonal states of N identical fermions, with M available single-particle states

$$=\frac{M!}{(M-N)!N!}$$

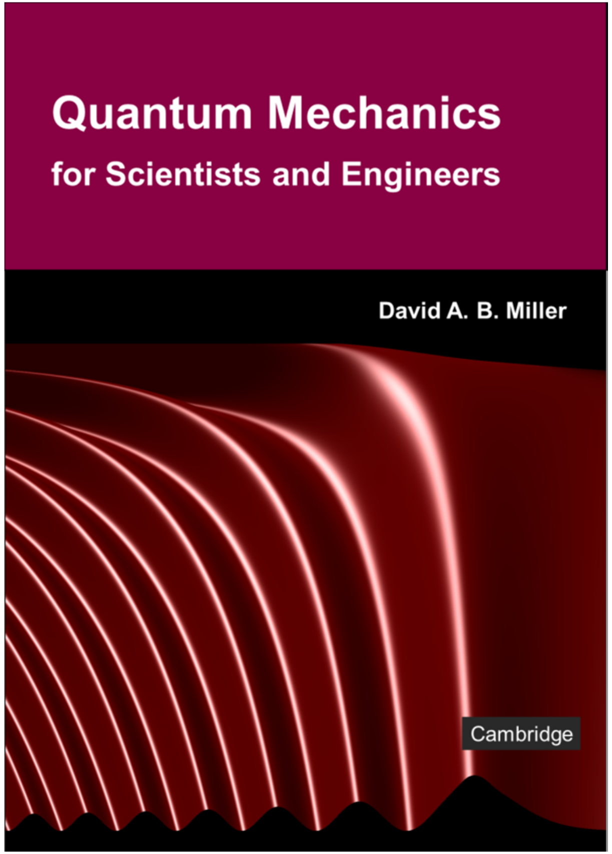


14.1 States of identical particles

Slides: Video 14.1.5 Numbers of states

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.6 subsection “Example numbers of states”



States of identical particles

Numbers of states

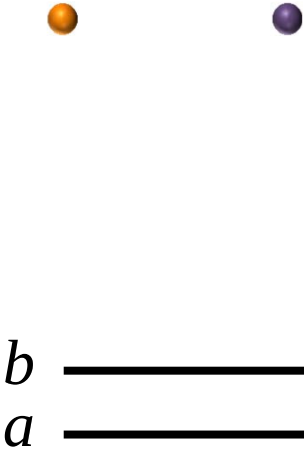
Quantum mechanics for scientists and engineers

David Miller

Example numbers of states

For example, suppose we have two particles

- each of which can be in one of two different single-particle states or modes, a and b



Example numbers of states

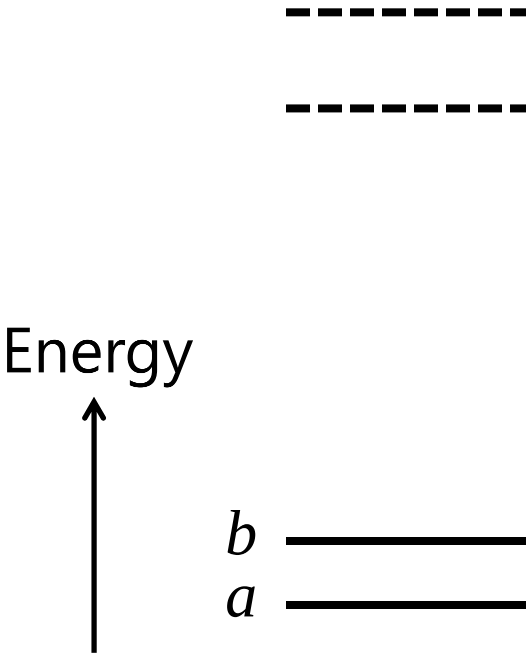
Suppose these particles are in some potential such that

there are two single-particle states or modes quite close in energy

and all other possible states are sufficiently far away in energy

that we can approximately neglect those other states

in our counting



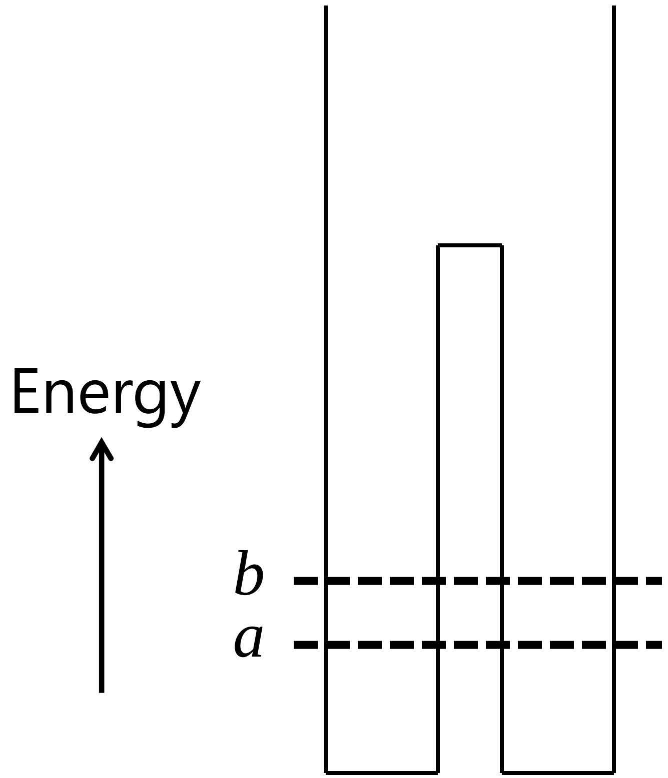
Example numbers of states

We might be considering, for example

two particles in a weakly coupled pair of similar quantum boxes

or a one-dimensional problem

such as coupled potential wells



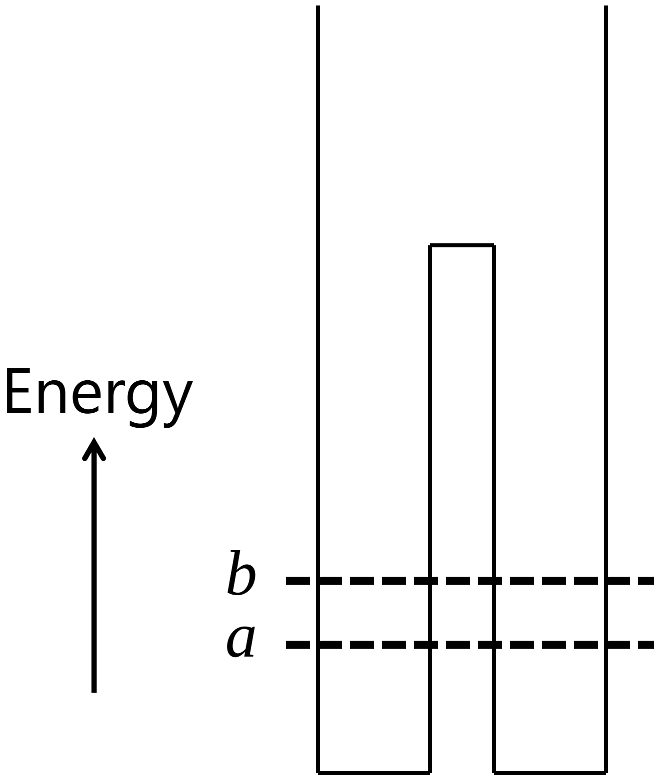
Example numbers of states

Because we know for some other reason that the particles cannot have much energy

- for example, the temperature may be low

  - we presume the particles can only be in one or other

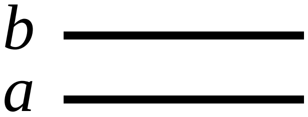
    - of the two lowest coupled single-particle states or modes of these two wells or boxes



Example numbers of states

For each situation we consider

- non-identical particles
- identical bosons, and
- identical fermions
  - these single-particle states or modes
    - might be different
      - e.g., because of exchange energy
  - but that will not affect our argument here
    - which is just counting states



Example numbers of states

We can now write out the possible states in each case

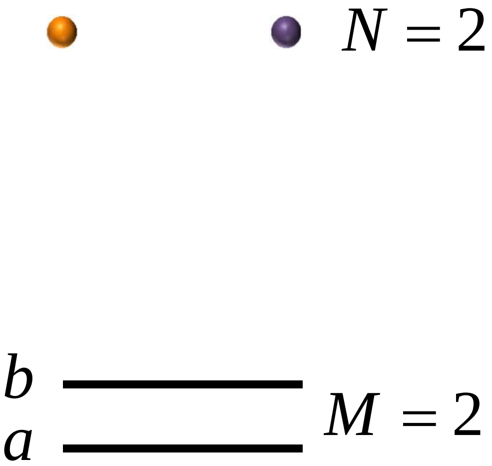
For all of these cases

the number of possible single-particle states or modes of a particle is

$M = 2$

and the number of particles is

$N = 2$

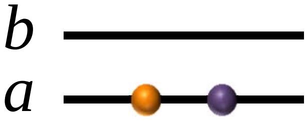


Non-identical particles

For non-identical particles

- such as
  - a proton and a neutron
    - the possible distinct states of this pair of particles are

$$|1,a\rangle|2,a\rangle$$



Non-identical particles

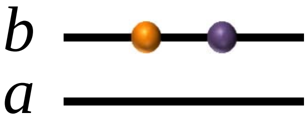
For non-identical particles

such as

a proton and a neutron

the possible distinct states of this pair of particles are

$$|1,a\rangle|2,a\rangle \quad |1,b\rangle|2,b\rangle$$

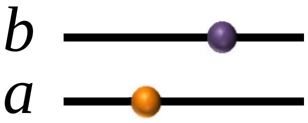


Non-identical particles

For non-identical particles

- such as
  - a proton and a neutron
    - the possible distinct states of this pair of particles are

$|1,a\rangle|2,a\rangle \quad |1,b\rangle|2,b\rangle \quad |1,a\rangle|2,b\rangle$

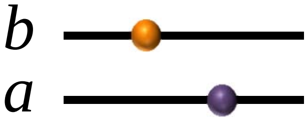


Non-identical particles

For non-identical particles

- such as
  - a proton and a neutron
    - the possible distinct states of this pair of particles are

$$|1,a\rangle|2,a\rangle \quad |1,b\rangle|2,b\rangle \quad |1,a\rangle|2,b\rangle \quad |1,b\rangle|2,a\rangle$$



As we expected from the expression M^N

- there are $2^2 = 4$ states of the pair of particles

Bosons

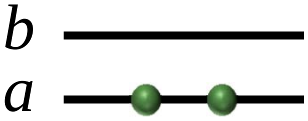
We could consider identical bosons

- such as two ^{4}He (helium-four) atoms (which are bosons)
  - because they are made from 6 particles each with spin ½
    - two protons, two neutrons and two electrons, which therefore have an integer total spin

Bosons

The possible distinct states of this pair of identical bosons are

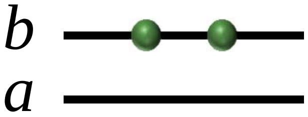
$$|1,a\rangle|2,a\rangle$$



Bosons

The possible distinct states of this pair of identical bosons are

$$|1,a\rangle|2,a\rangle \quad |1,b\rangle|2,b\rangle$$

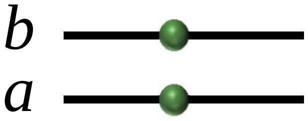


Bosons

The possible distinct states of this pair of identical bosons are

$$|1,a\rangle|2,a\rangle \quad |1,b\rangle|2,b\rangle \quad \frac{1}{\sqrt{2}}\left(|1,a\rangle|2,b\rangle+|2,a\rangle|1,b\rangle\right)$$

Note there is only one way of having the two identical particles in different states



Bosons

In this list of basis states

$$|1,a\rangle|2,a\rangle,\ |1,b\rangle|2,b\rangle,\ \frac{1}{\sqrt{2}}\left(|1,a\rangle|2,b\rangle+|2,a\rangle|1,b\rangle\right)$$

we do not have to write the explicit symmetrized form $|1,a\rangle|2,a\rangle+|2,a\rangle|1,a\rangle$

since it is describing the same state as $|1,a\rangle|2,a\rangle$

and similarly for the state with both particles in the b mode

(The $1/\sqrt{2}$ normalizes the explicitly symmetric combination state)

Bosons

In this list of basis states

$$|1,a\rangle|2,a\rangle,\ |1,b\rangle|2,b\rangle,\ \frac{1}{\sqrt{2}}\left(|1,a\rangle|2,b\rangle+|2,a\rangle|1,b\rangle\right)$$

we therefore have 3 states

which agrees with $(M+N-1)!/\left[N!(M-1)!\right]$

i.e., $(2+2-1)!/\,2!(2-1)!=3$

Note that this is *not* the same as

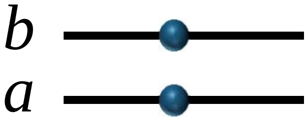
the case of non-identical particles

where we had 4 states

Fermions

For identical fermions, there is only

- one possible state of the pair of particles
  - since the two particles have to be in different single-particle states
    - and here there are only two single-particle states to choose from for each particle



Fermions

So that one (normalized) state is

$$\left(1/\sqrt{2}\right)\left(|1,a\rangle|2,b\rangle-|2,a\rangle|1,b\rangle\right)$$

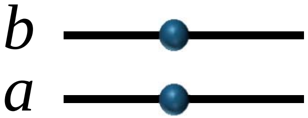
which agrees with the formula

$$M!/\left[\left(M-N\right)!N!\right]$$

which gives $2!/\left(2!0!\right)=1$ state

where we remember

that $0! = 1$



Thermal occupation of states

The differences in the number of available states in the three cases of

- non-identical particles
- identical bosons, and
- identical fermions

leads to very different behavior once we consider the thermal occupation of states

Thermal occupation of states

For example, if we presume that we are at some relatively high temperature

such that the thermal energy, $k_B T$

is much larger than the energy separation of the two single-particle states or modes a and b

but still much less than the energy to the next states

then the thermal occupation probabilities

of all the different allowed combinations of single-particle states or modes

will all tend to be similar

Non-identical particles

For the case of the non-identical particles

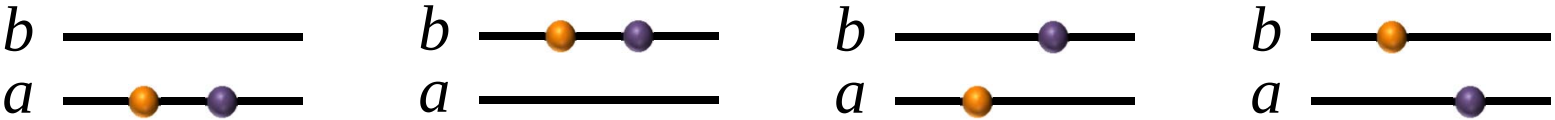
which behave like classical particles as far as the counting of states is concerned

with the 4 states $|1,a\rangle|2,a\rangle$ $|1,b\rangle|2,b\rangle$ $|1,a\rangle|2,b\rangle$ $|1,b\rangle|2,a\rangle$

we therefore expect a probability of ~ ¼

of occupation of each of the states

Therefore, the probability that the two particles are in the same state is ~ ½



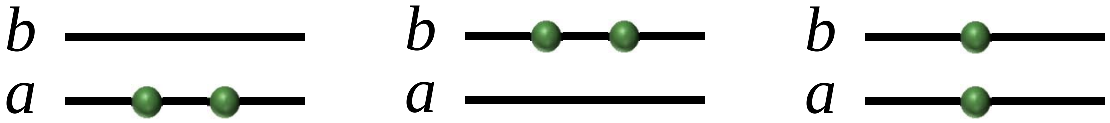
Identical bosons

For the case of the identical bosons

there are only three possible states

so the probability of occupation of any one state

is ~ 1/3



Identical bosons

Two of the two-particle states have

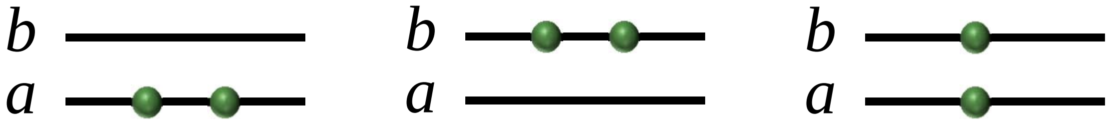
the particles in identical modes $|1,a\rangle|2,a\rangle \quad |1,b\rangle|2,b\rangle$

and only one two-particle state $\left(1/\sqrt{2}\right)\left(|1,a\rangle|2,b\rangle+|2,a\rangle|1,b\rangle\right)$

has the particles in different single particle states

So the probability of finding the two identical bosons in the same single-particle state (mode) is now 2/3

larger than the ½ for the non-identical particle case



Identical fermions

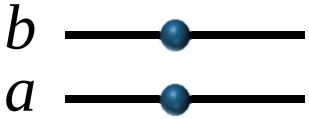
For the case of identical fermions

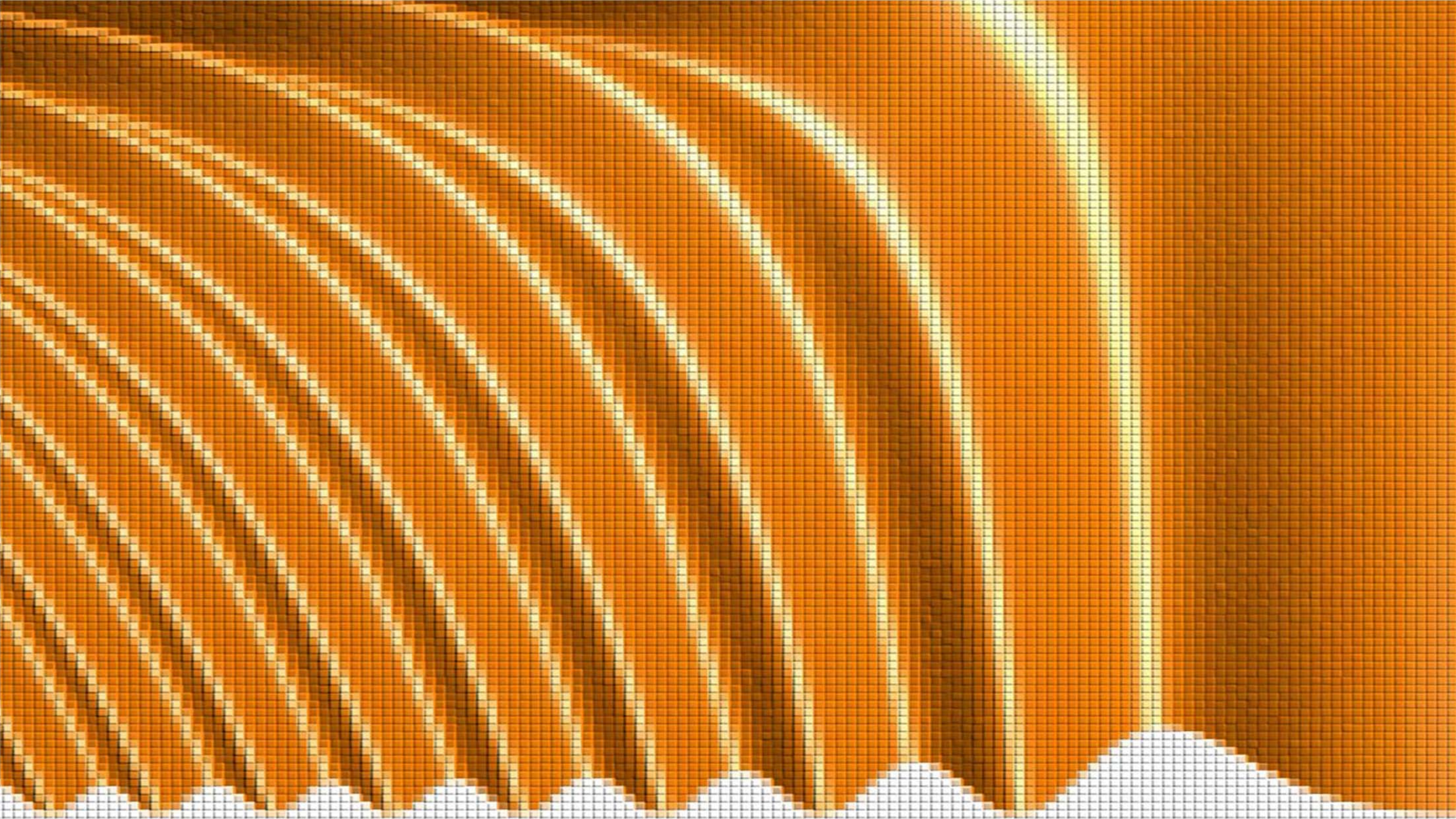
there is only one possible state

$$\left(1/\sqrt{2}\right)\left(|1,a\rangle|2,b\rangle-|2,a\rangle|1,b\rangle\right)$$

which therefore has probability ~1

and it necessarily corresponds to the two particles being in different states



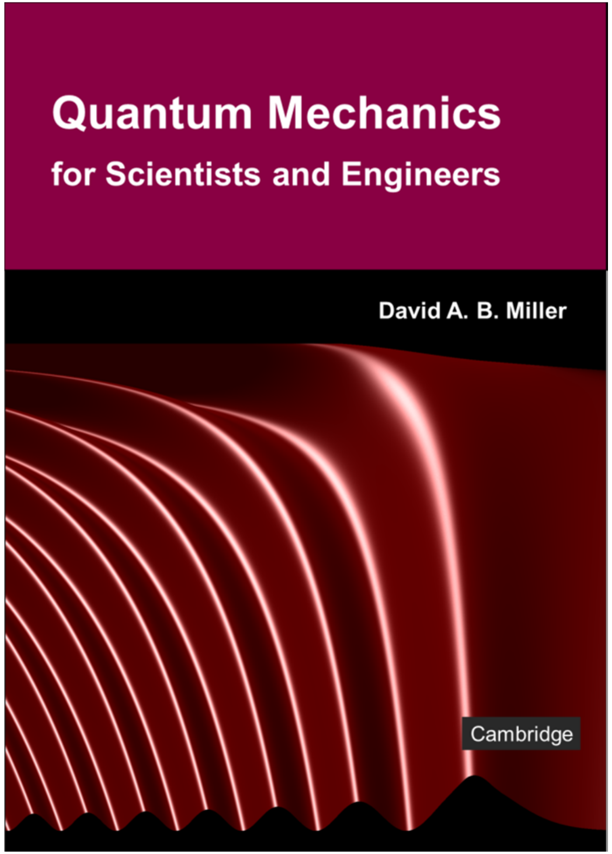


14.1 States of identical particles

Slides: Video 14.1.7 Analogy for counting states

Text reference: Quantum Mechanics for Scientists and Engineers

Sections 13.6 subsection "Bank account analogy for counting states"



States of identical particles

Analogy for counting states

Quantum mechanics for scientists and engineers

David Miller

Bank account analogy

Suppose you have

- an antique jar (*a*) in the kitchen for your spending money
- and a box (*b*) under the bed for your savings money

You put your dollar bills

- each labeled with a unique number
  - into either the antique jar (*a*) or the box (*b*)

a

b

Bank account analogy

This is like the quantum mechanical situation of

non-identical particles (the dollar bills) and

different single-particle states or modes (a or b)

into which they can be put

– the jar or the box

a

b

Bank account analogy

- If I have two dollar bills
  - then there are four possible situations
    - i.e., states of the entire system
      - of two dollar bills in the antique jar and/or the box

a

b

Bank account analogy

bill 1 in the box and bill 2 in the box

a

b

Bank account analogy

bill 1 in the box and bill 2 in the box

bill 1 in the box and bill 2 in the antique jar

a

b

Bank account analogy

- bill 1 in the box and bill 2 in the box
- bill 1 in the box and bill 2 in the antique jar
- bill 1 in the antique jar and bill 2 in the box

a

b

Bank account analogy

- bill 1 in the box and bill 2 in the box
- bill 1 in the box and bill 2 in the antique jar
- bill 1 in the antique jar and bill 2 in the box
- bill 1 in the antique jar and bill 2 in the antique jar

a

b

Bank account analogy

bill 1 in the box and bill 2 in the box

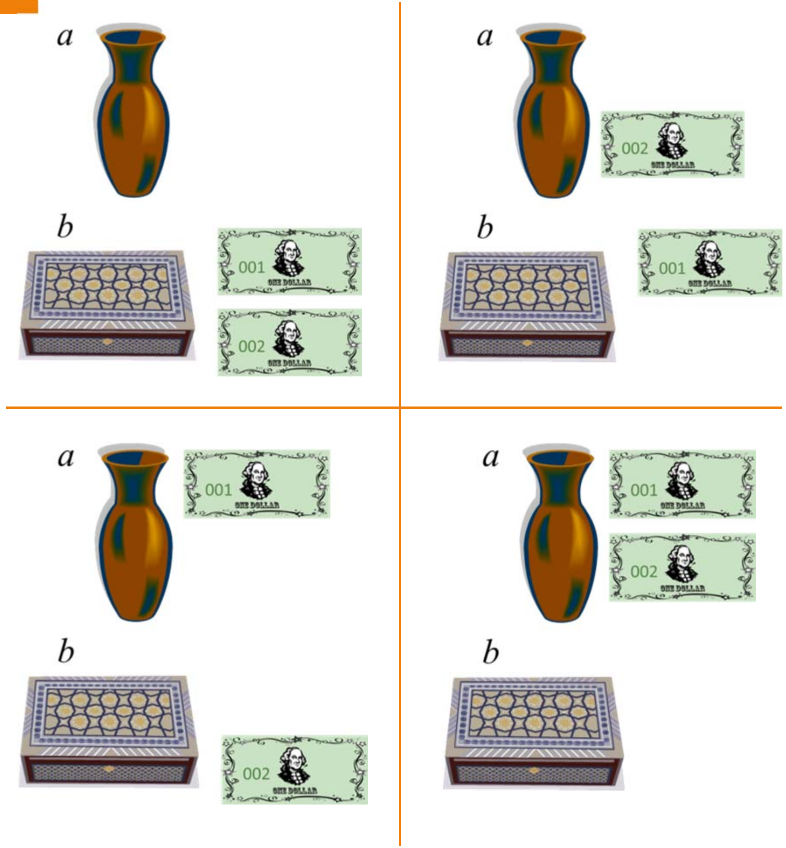
bill 1 in the box and bill 2 in the antique jar

bill 1 in the antique jar and bill 2 in the box

bill 1 in the antique jar and bill 2 in the antique jar

making four states altogether

This reproduces the counting for non-identical particles



Bank account analogy

Consider next that you have two bank accounts

- a checking account (a), and a savings account (b)

You may still have the same amount of money

- $2

You may know how much money you have in each account

- but the dollars are themselves identical in the accounts

So now there are only three possible states

- Two dollars in savings
- One dollar in savings and one in checking
- Two dollars in checking

Bank account analogy

Note that, in these three possible states

- Two dollars in savings
- One dollar in savings and one in checking
- Two dollars in checking

there are

- 2 states with both dollars in the same account
  - but only one in which they are in different accounts

This bank account argument above gives the counting for boson states

Bank account analogy

Consider now that you have two bank accounts

- a checking account (a) and a savings account (b)

but you are living in the Protectorate of Pauliana

- where you may have no more than one dollar in each bank account

Then for your two dollars

- there is only one possible state
  - one dollar in savings
  - one dollar in checking

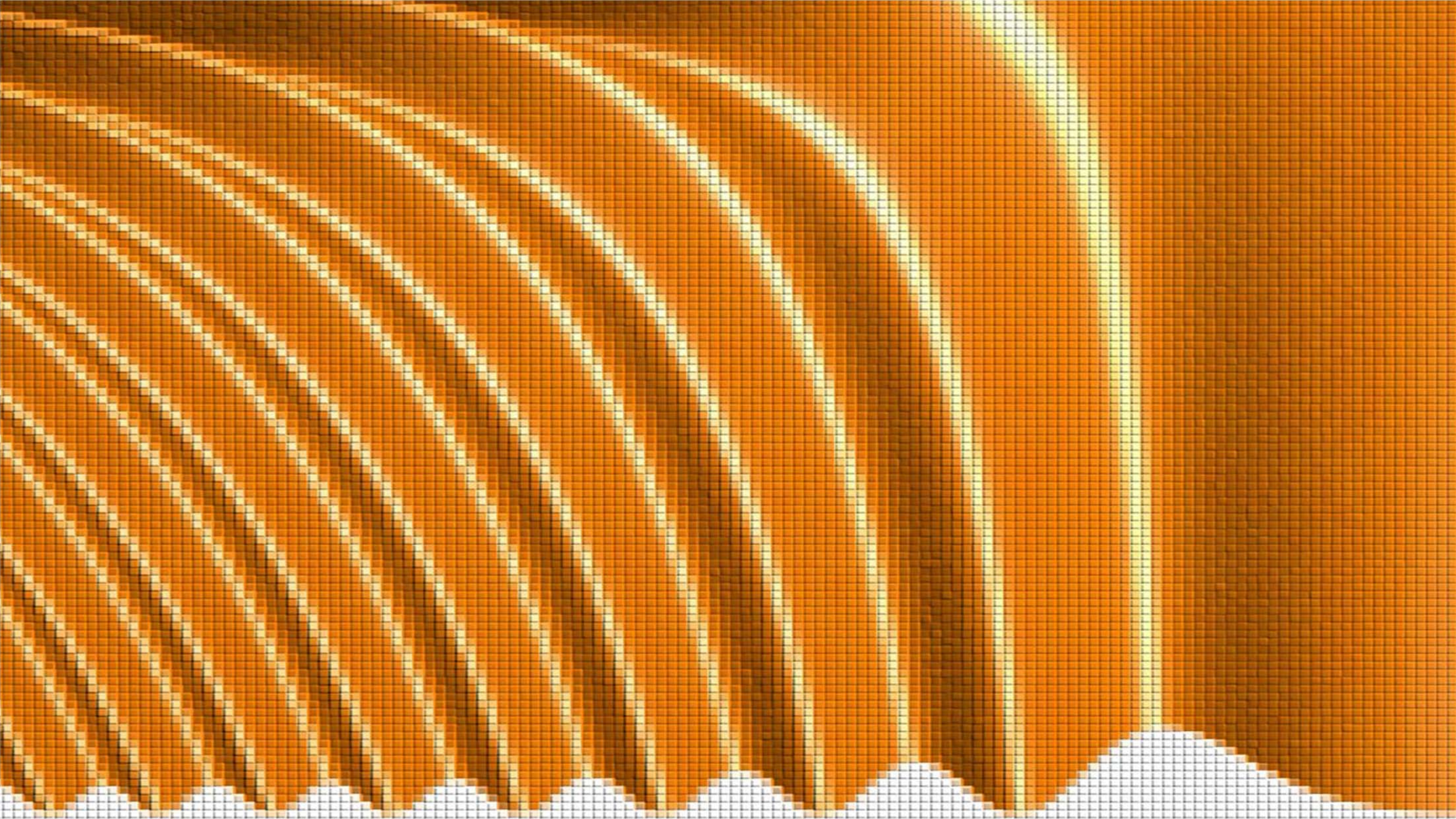
This gives the counting for fermion states

Counting states with two “bank accounts”

For the case of identical fermions
- there is only one possible state for our two dollars
  - with each dollar being in a different bank account

For identical bosons
- there are three possible states for our two dollars
  - in two of which both are in the same bank account
  - and in one of which they are in different bank accounts

For non-identical (classical) particles
- there are four possible states for our dollar bills
  - in two of which both are in the same bank account
  - and in two of which they are in different bank accounts

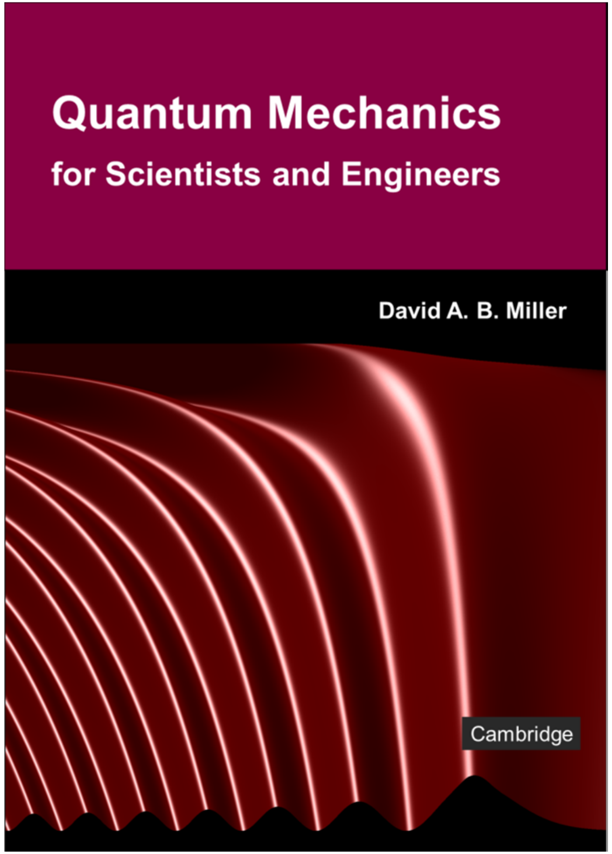


14.2 Multiple particle distributions

Slides: Video 14.2.1 Thermal distributions

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.7



Multiple particle distributions

Thermal distributions

Quantum mechanics for scientists and engineers

David Miller

Thermal distributions

Non-identical particles

Maxwell-Boltzmann

$$P_{MB}(E) = \exp\left[-(E-\mu)/k_B T\right]$$

Identical bosons

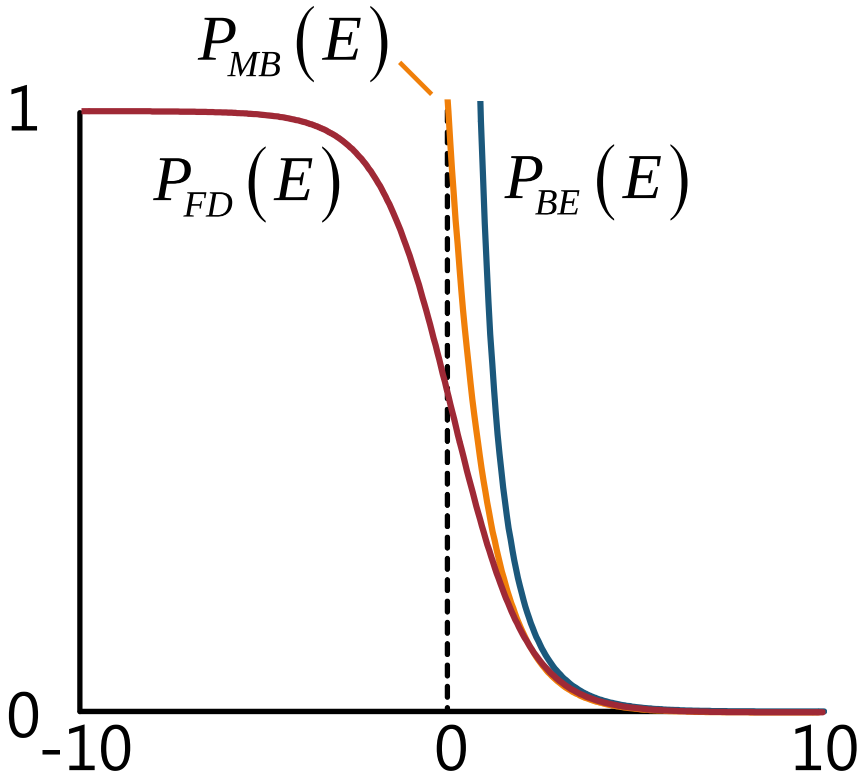
Bose-Einstein

$$P_{BE}(E) = \frac{1}{\exp\left[(E-\mu)/k_B T\right]-1}$$

Identical fermions

Fermi-Dirac

$$P_{FD}(E) = \frac{1}{1+\exp\left[(E-\mu)/k_B T\right]}$$

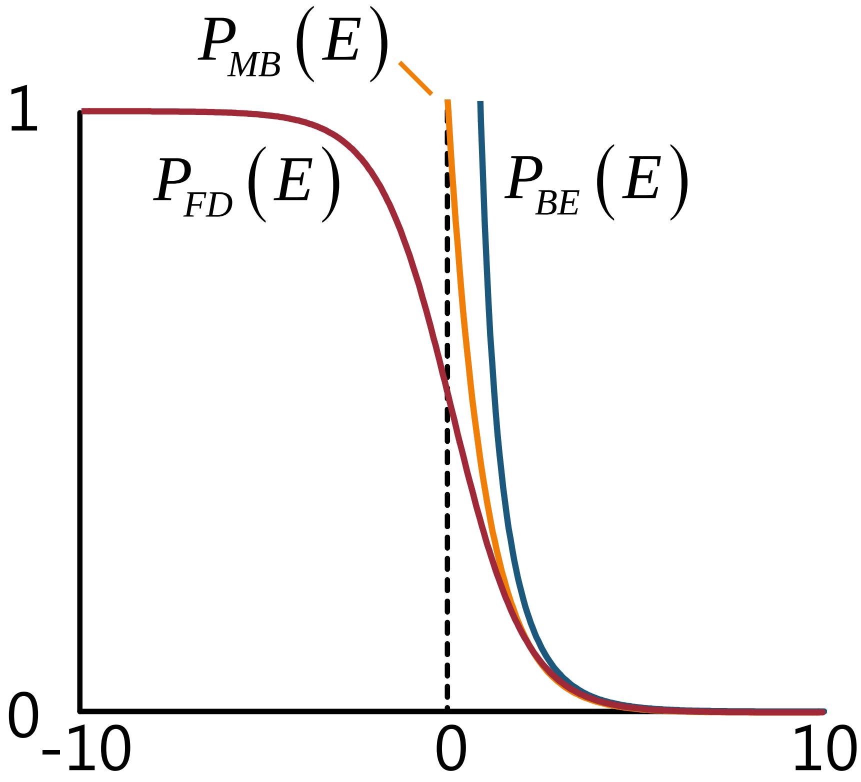


Energy E - μ (in units of $k_B T$)

(μ – chemical potential or Fermi energy)

Thermal distributions

- Identical bosons
  - are more likely to be in the same mode
    - than are classical or non-identical particles
- The Bose-Einstein distribution
  - lies above the Maxwell-Boltzmann distribution



Energy E - μ (in units of $k_B T$)

(μ – chemical potential or Fermi energy)

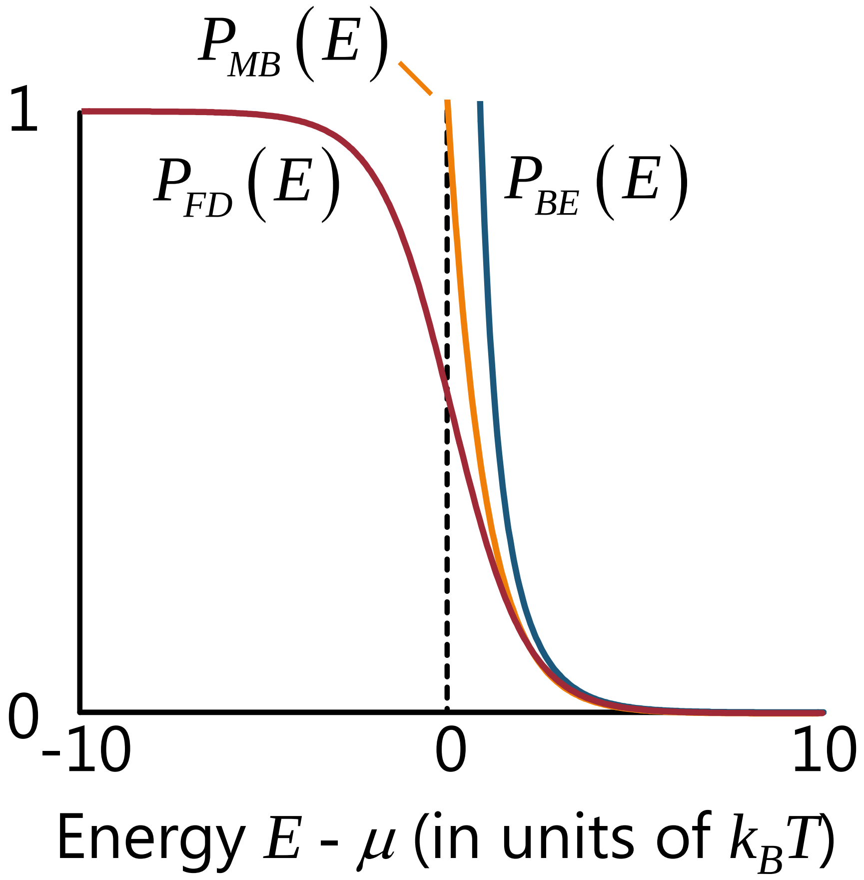
Thermal distributions

Identical fermions

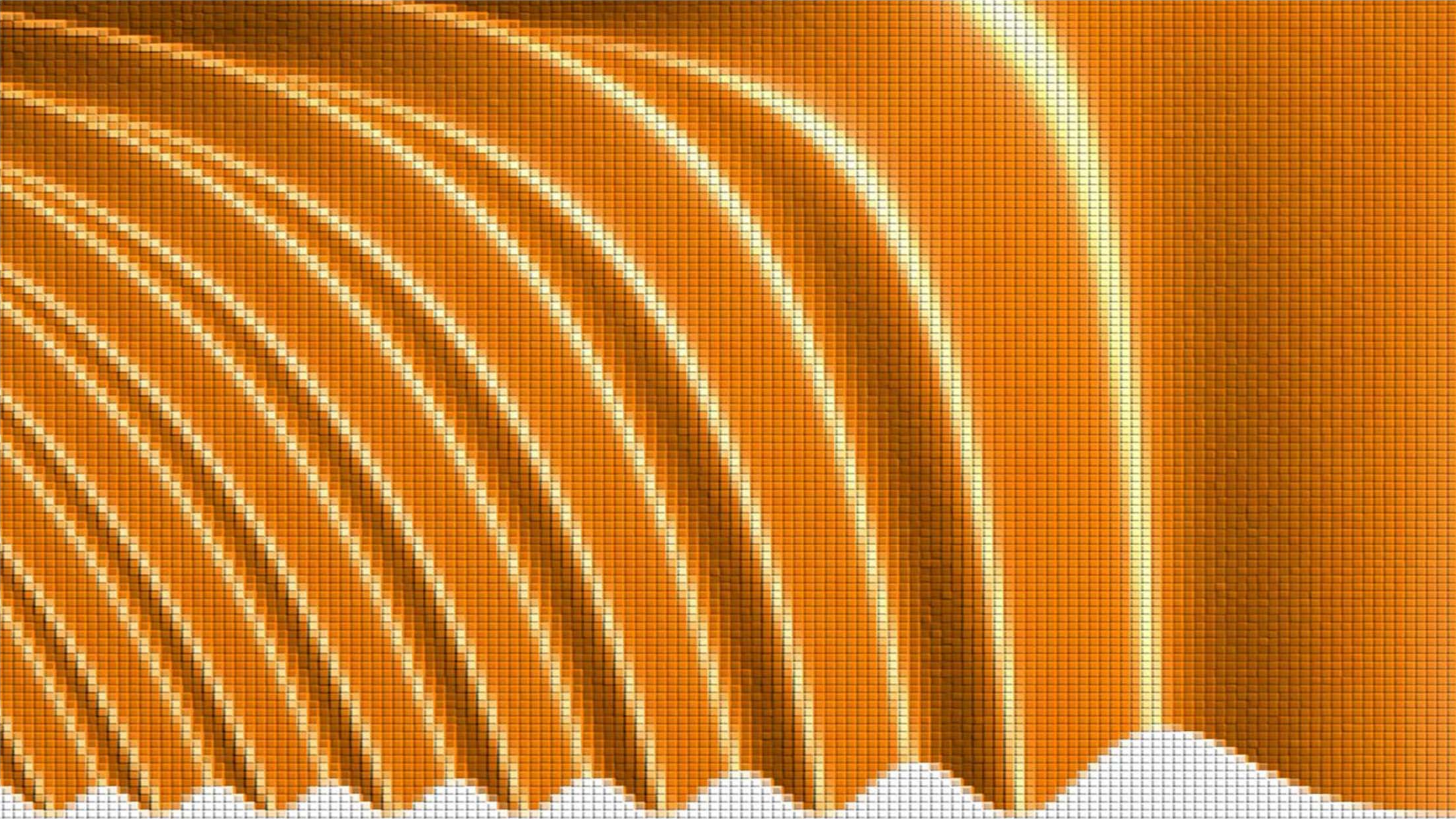
- are less likely to be in the same single-particle state
  - than are classical or non-identical particles
    - In fact, they never are in the same single-particle state

The Fermi-Dirac distribution

- lies below the Maxwell-Boltzmann distribution



(μ – chemical potential or Fermi energy)

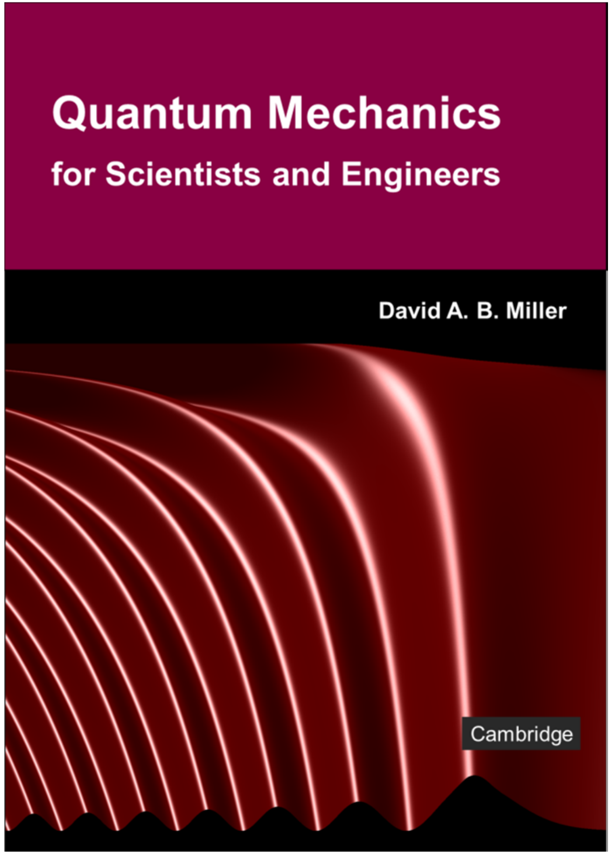


14.2 Multiple particle distributions

Slides: Video 14.2.3 Examples of states of multiple identical particles

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.8



Multiple particle distributions

Examples of states of multiple identical particles

Quantum mechanics for scientists and engineers

David Miller

A filled semiconductor band

Consider a filled valence band in a semiconductor

- in the “single particle” approximation
  - where one electron is assumed to move in an average periodic potential
    - and therefore is in a Bloch state of a particular $\mathbf{k}$ value

The possible Bloch states of a single electron (of a given spin)

- correspond to all the different possible $\mathbf{k}$ values in the band
  - of which there are N_c if there are N_c unit cells in the crystal

A filled semiconductor band

A full band therefore corresponds to

- N_c electrons of each spin in $2N_c$ different single-particle states
  - where the factor of 2 comes from the two spin states associated with each $\mathbf{k}$ value

There is only one such state

- made from these single particle states
  - that obeys the antisymmetry with respect to exchange
    - which is the Slater determinant of all of the single-particle states in the band

N photons in a mode

Photons in a given mode are certainly identical

Photons are very simple identical bosons

Photons do not have excited states of any kind

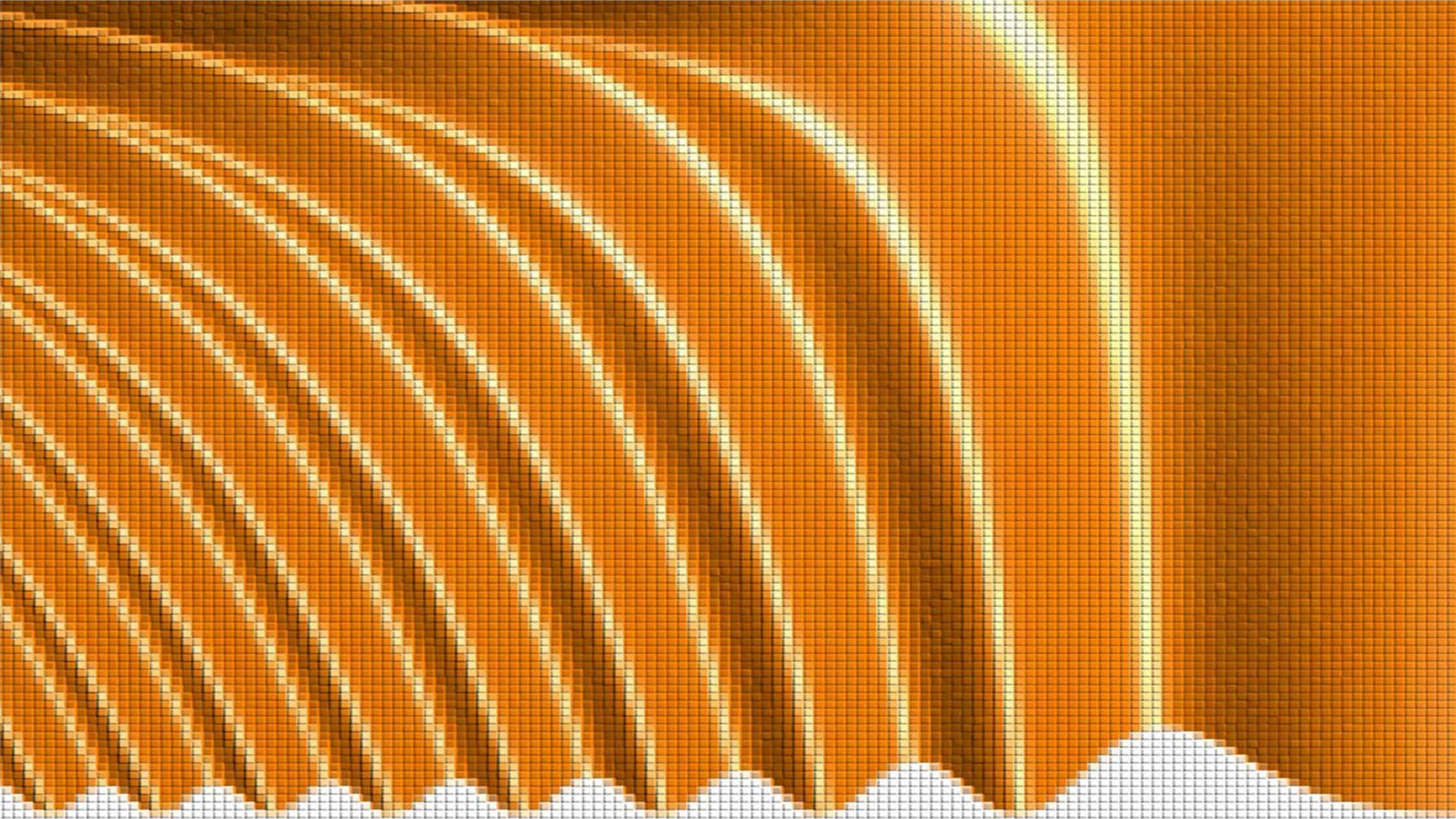
There is therefore no meaning to the idea of identical photons in a given mode

having more than one state to choose from

So, $M = 1$ for photons in a mode

and the number of possible states of the N photons in the mode is simply

$$(1+N-1)!/\left[N!(1-1)!\right]=1$$

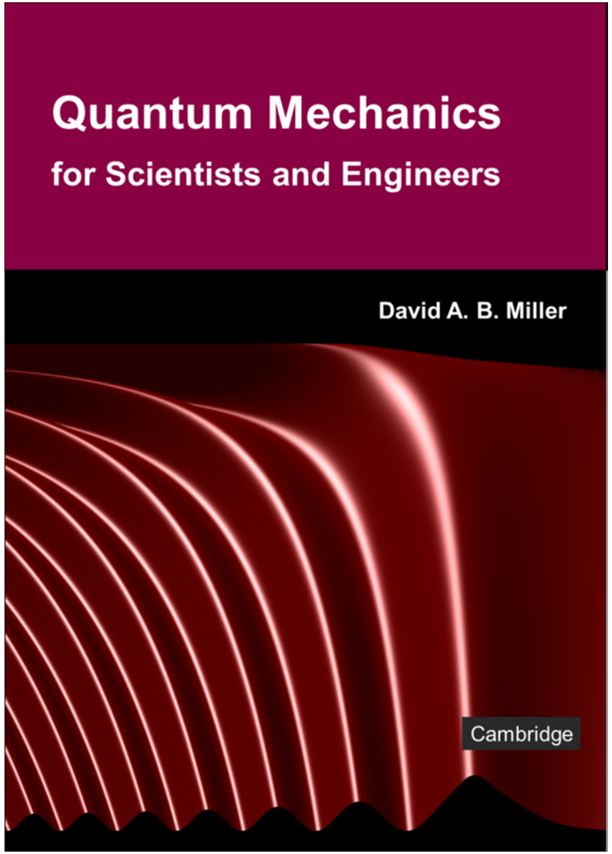


14.2 Multiple particle distributions

Slides: Video 14.2.5 Quantum mechanical particles reconsidered

Text reference: Quantum Mechanics for Scientists and Engineers

Section 13.9 – 13.10

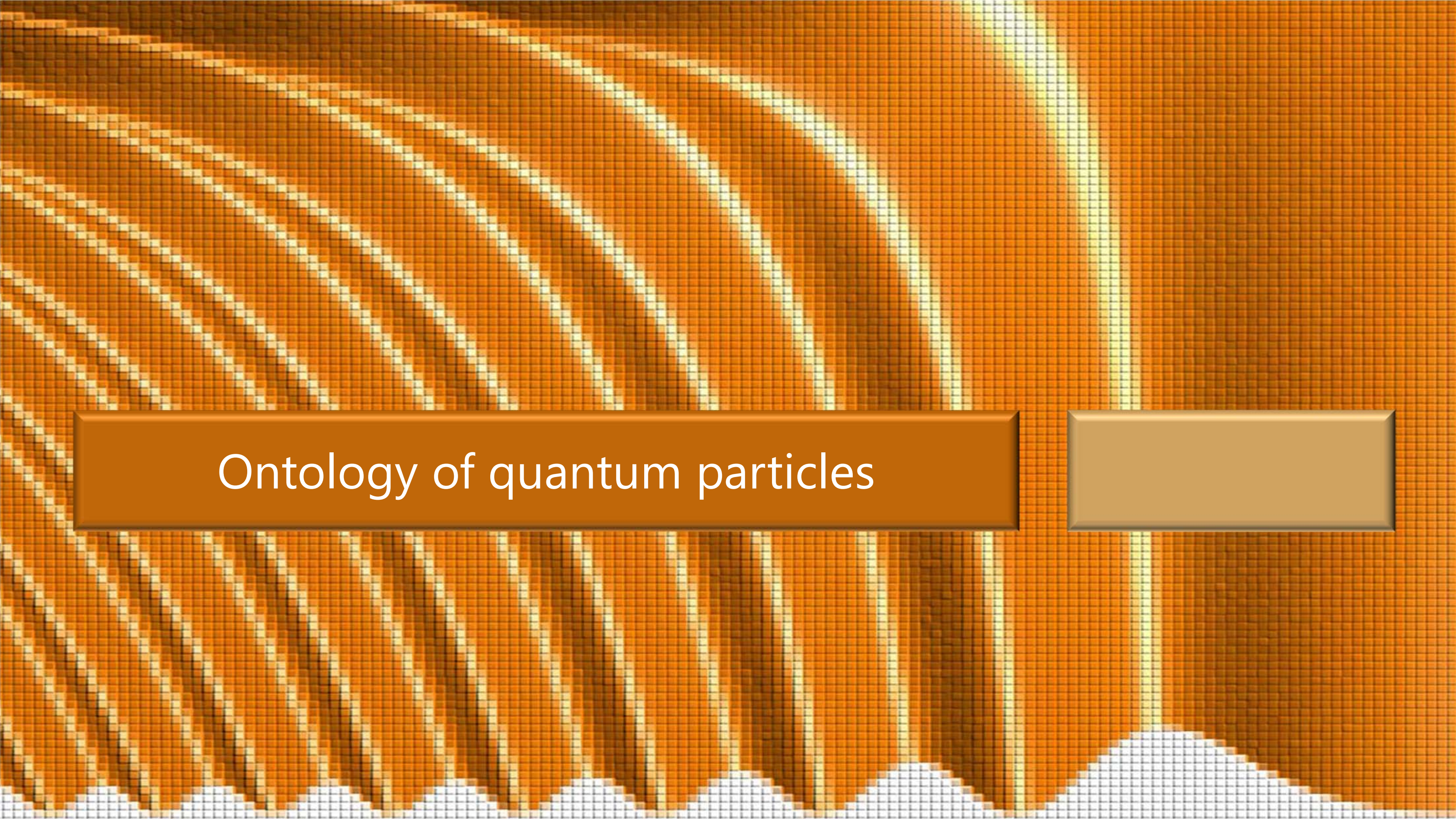


Multiple particle distributions

Quantum mechanical particles reconsidered

Quantum mechanics for scientists and engineers

David Miller



Ontology of quantum particles

Quantum mechanical particles reconsidered

There are fewer ontological problems if we consider levels of excitation of modes

- Instead of “3 photons in mode a and 2 in mode b”
  - we say that mode a is in its third level of excitation, and mode b is in its second level of excitation

Counting is then simple, as in the bank account analogy

It does not matter if we never introduce “particles”

- as long as we have the rules constructed by quantum mechanics for manipulating states
  - it does not matter what words we use

Distinguishable and indistinguishable particles

Distinguishable and indistinguishable particles

So, even if two particles are identical

if there is no reasonable physical process

by which they could be swapped

such “distinguishable” particles can be treated

as if they were non-identical

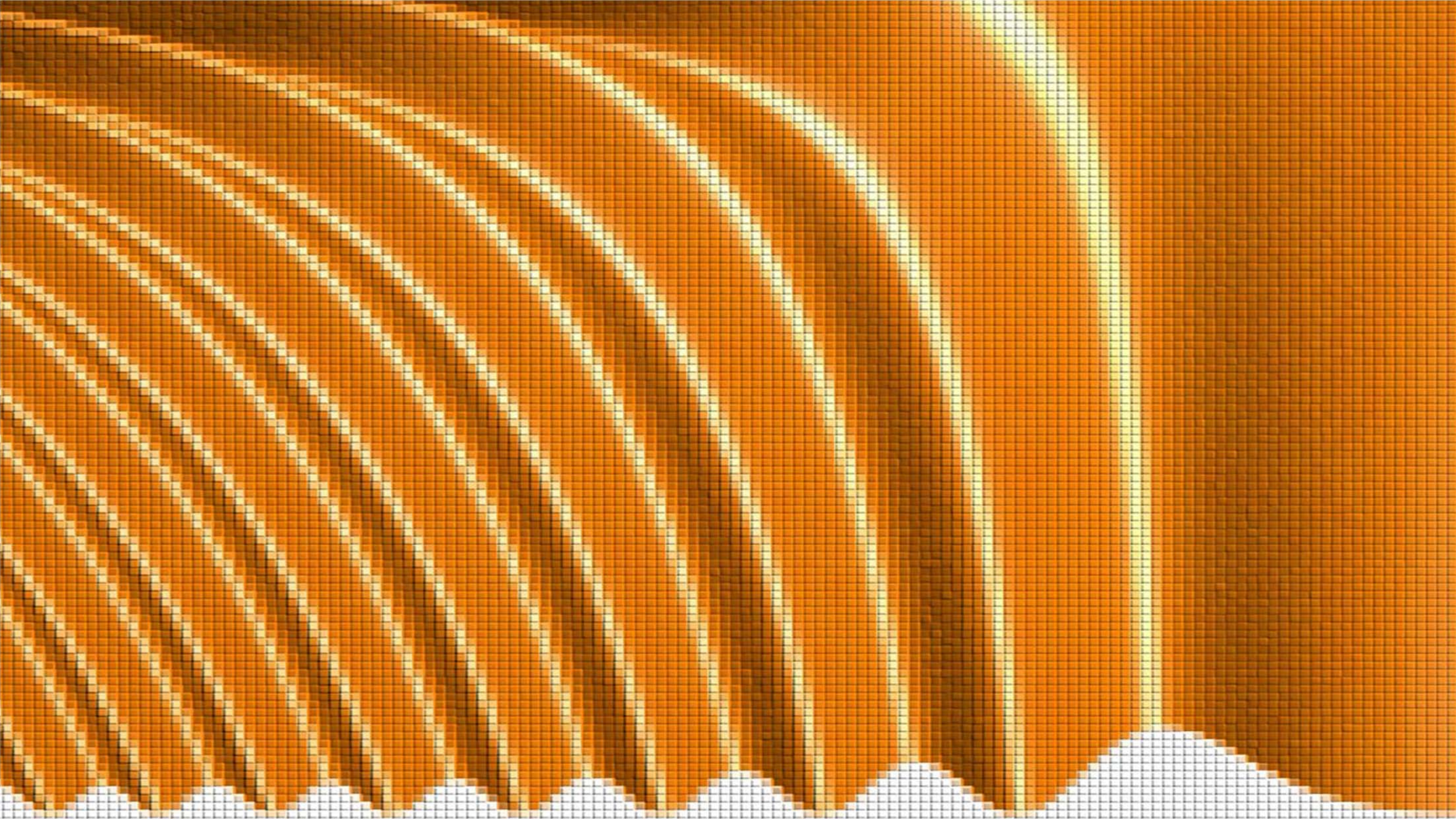
Distinguishable and indistinguishable particles

So, we can say as an approximation that

- two identical particles are distinguishable
  - if the exchange interaction between them is negligibly small
    - Then the "distinguishability" lets us treat them as non-identical particles for practical purposes

Conversely, if we say that two particles are indistinguishable

- because of the possibility of exchange of them
  - then we are saying that we have to symmetrize the state properly with respect to exchange

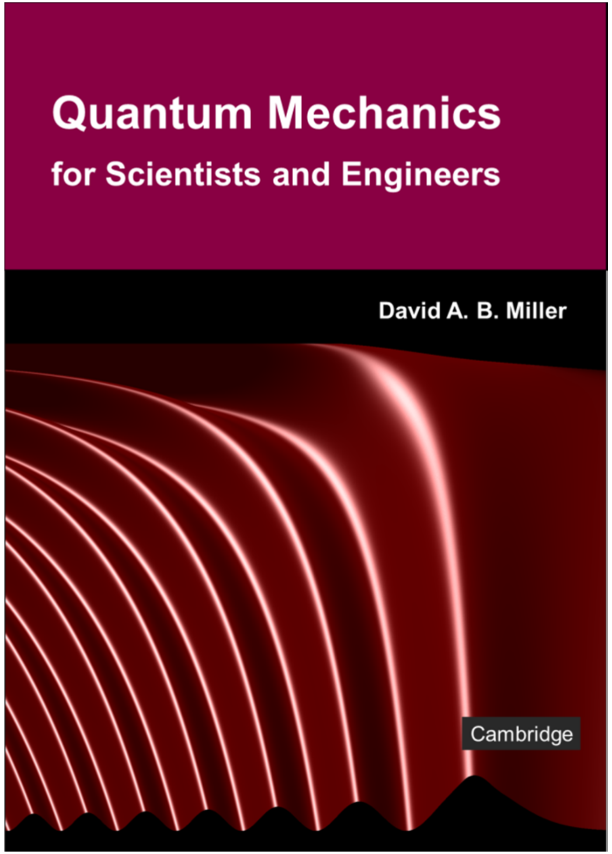


14.3 Operators for harmonic oscillators

Slides: Video 14.3.1 Raising and lowering operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.1 up to Eq. 15.16



Operators for harmonic oscillators

Quantum mechanics for scientists and engineers

David Miller

Operators for harmonic oscillators

Raising and lowering operators

Quantum mechanics for scientists and engineers

David Miller

Harmonic oscillator raising and lowering operators

The harmonic oscillator Schrödinger equation was

$$\hat{H}\psi = \left[-\frac{\hbar^2}{2m}\frac{d^2}{dz^2} + \frac{1}{2}m\omega^2 z^2 \right]\psi = E\psi$$

where ω is the angular frequency of oscillation

Now we introduce a dimensionless distance $\xi = \sqrt{\frac{m\omega}{\hbar}}z$

so we can rewrite the Schrödinger equation as

$$\frac{1}{2}\left[-\frac{d^2}{d\xi^2} + \xi^2 \right]\psi = \frac{E}{\hbar\omega}\psi$$

Harmonic oscillator raising and lowering operators

The term $-d^2/d\xi^2+\xi^2$ reminds us of the difference of two squares of numbers

$$-a^2+b^2=b^2-a^2=(-a+b)(a+b)$$

though here we have the difference that

$d^2/d\xi^2$ is an operator

If we examine a product of this form for our present case, we have

$$\frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right)\times\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)=\frac{1}{2}\left(-\frac{d^2}{d\xi^2}+\xi^2\right)-\frac{1}{2}\left(\frac{d}{d\xi}\xi-\xi\frac{d}{d\xi}\right)$$

Harmonic oscillator raising and lowering operators

$$\frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right)\times\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)=\frac{1}{2}\left(-\frac{d^2}{d\xi^2}+\xi^2\right)-\frac{1}{2}\left(\frac{d}{d\xi}\xi-\xi\frac{d}{d\xi}\right)$$

in addition to the $(1/2)\left(-d^2/d\xi^2+\xi^2\right)$ we want to rewrite

has a second term that we cannot simply discard

Here we note that, for any function $f(\xi)$

$$\left(\frac{d}{d\xi}\xi-\xi\frac{d}{d\xi}\right)f(\xi)=\frac{d}{d\xi}\left(\xi f(\xi)\right)-\xi\frac{d}{d\xi}f(\xi)$$

$$=f(\xi)\frac{d\xi}{d\xi}+\xi\frac{d}{d\xi}f(\xi)-\xi\frac{d}{d\xi}f(\xi)=f(\xi)$$

Harmonic oscillator raising and lowering operators

Since this relation $\left(\frac{d}{d\xi}\xi-\xi\frac{d}{d\xi}\right)f(\xi)=f(\xi)$

is true for any function $f(\xi)$

we can write the commutation relation $\left(\frac{d}{d\xi}\xi-\xi\frac{d}{d\xi}\right)=1$

So from

$$\frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right)\times\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)=\frac{1}{2}\left(-\frac{d^2}{d\xi^2}+\xi^2\right)-\frac{1}{2}\left(\frac{d}{d\xi}\xi-\xi\frac{d}{d\xi}\right)$$

we have $\frac{1}{2}\left(-\frac{d^2}{d\xi^2}+\xi^2\right)=\frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right)\times\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)+\frac{1}{2}$

Harmonic oscillator raising and lowering operators

We can choose to write the "raising" or "creation" operator

$$\boxed{\hat{a}^\dagger \equiv \frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right)}$$

(pronounced "a dagger")

and the "lowering" or "annihilation" operator

$$\boxed{\hat{a} \equiv \frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)}$$

Harmonic oscillator raising and lowering operators

Note these operators $\hat{a}^\dagger \equiv \frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right)$ and $\hat{a} \equiv \frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)$

are Hermitian adjoints of one another

The operator $d/d\xi$ is anti-Hermitian, as shown for d/dz

i.e., $\langle\phi|d/d\xi|\psi\rangle = -\left[\langle\psi|d/d\xi|\phi\rangle\right]^*$ for arbitrary $|\phi\rangle$ and $|\psi\rangle$

though ξ is Hermitian (being the position operator)

Therefore, we see $\hat{a}$ and $\hat{a}^\dagger$ are Hermitian adjoints, i.e.,

$$\langle\phi|\hat{a}|\psi\rangle = \langle\phi|\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)|\psi\rangle = \left[\langle\psi|\frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right)|\phi\rangle\right]^* = \left[\langle\psi|\hat{a}^\dagger|\phi\rangle\right]^*$$

Note neither of these operators is Hermitian, i.e., $\hat{a} \neq \hat{a}^\dagger$

Hamiltonian with raising and lowering operators

From $\frac{1}{2}\left(-\frac{d^2}{d\xi^2}+\xi^2\right)=\frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right)\times\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)+\frac{1}{2}$

and using the definitions for these operators

$$\hat{a}^\dagger \equiv \frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi}+\xi\right) \qquad \hat{a} \equiv \frac{1}{\sqrt{2}}\left(\frac{d}{d\xi}+\xi\right)$$

the Schrödinger equation

$\frac{1}{2}\left[-\frac{d^2}{d\xi^2}+\xi^2\right]\psi=\frac{E}{\hbar\omega}\psi$ becomes $\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)\psi=\frac{E}{\hbar\omega}\psi$

So we can rewrite the Hamiltonian as $\hat{H}\equiv\hbar\omega\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)$

Number operator

We know from the previous harmonic oscillator solution

that the eigenenergy for eigenstate $|\psi_n\rangle$ is

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)$$

so given $\hat{H} \equiv \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)$

we know that

$$\hat{a}^\dagger\hat{a}|\psi_n\rangle = n|\psi_n\rangle$$

Number operator

Since $\hat{a}^\dagger \hat{a} |\psi_n\rangle = n |\psi_n\rangle$

this operator $\hat{a}^\dagger \hat{a}$ obviously has the harmonic oscillator states as its eigenstates

and the number n of the state as its eigenvalue

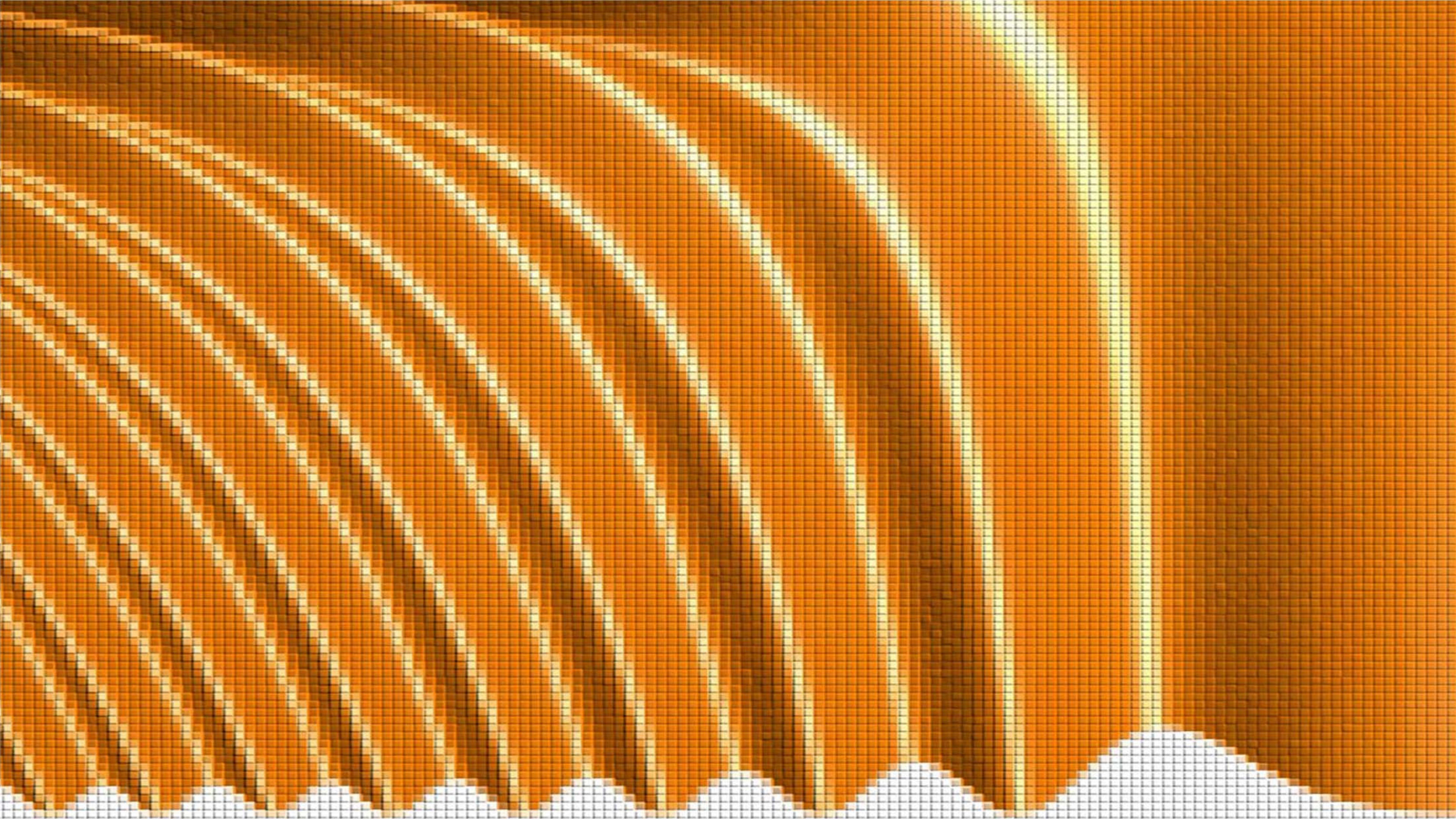
so it is sometimes called

the number operator

$$\hat{N} \equiv \hat{a}^\dagger \hat{a}$$

with the eigenequation

$$\hat{N} |\psi_n\rangle = n |\psi_n\rangle$$

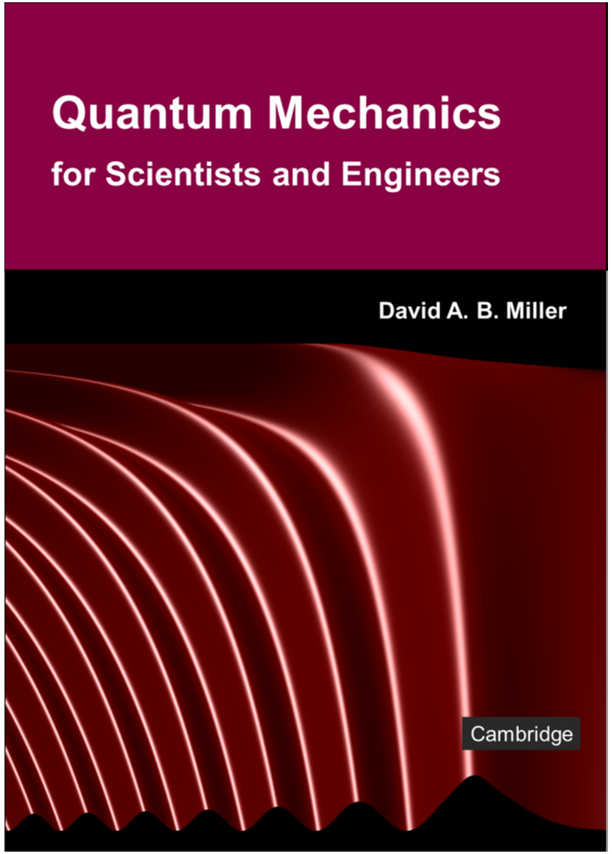


14.3 Operators for harmonic oscillators

Slides: Video 14.3.3 Properties of raising and lower operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.1 from "Properties of raising …"



Operators for harmonic oscillators

Properties of raising and lower operators

Quantum mechanics for scientists and engineers

David Miller

Properties of raising and lowering operators

The operators $\hat{a}$ and $\hat{a}^\dagger$ have a very important property

which is their commutator

Specifically, we find

$$\boxed{\left[\hat{a},\hat{a}^\dagger\right]=\hat{a}\hat{a}^\dagger-\hat{a}^\dagger\hat{a}=1}$$

(and it is a good exercise to prove this)

This relation is so useful that it is worth memorizing

If we use the word “less” in its meaning as “minus”

we have the relatively rhythmic phrase

> “a a dagger less a dagger a equals one”

Properties of raising and lowering operators

We can use this property

$$\left[\hat{a},\hat{a}^{\dagger}\right]=\hat{a}\hat{a}^{\dagger}-\hat{a}^{\dagger}\hat{a}=1$$

together with the property

$$\hat{a}^{\dagger}\hat{a}\left|\psi_{n}\right\rangle=n\left|\psi_{n}\right\rangle$$

to show the reason why these operators are called

raising and lowering operators

or creation and annihilation operators

Properties of raising and lowering operators

Operating on both sides of

$$\hat{a}^{\dagger}\hat{a}|\psi_n\rangle = n|\psi_n\rangle$$ with $\hat{a}$

we have $\hat{a}\left(\hat{a}^{\dagger}\hat{a}\right)|\psi_n\rangle = n\hat{a}|\psi_n\rangle$

Regrouping on the left gives $\left(\hat{a}\hat{a}^{\dagger}\right)\left(\hat{a}|\psi_n\rangle\right) = n\left(\hat{a}|\psi_n\rangle\right)$

and substituting from $\hat{a}\hat{a}^{\dagger} - \hat{a}^{\dagger}\hat{a} = 1$ we have

$$\left(\hat{a}\hat{a}^{\dagger}\right)\left(\hat{a}|\psi_n\rangle\right) = \left(1+\hat{a}^{\dagger}\hat{a}\right)\left(\hat{a}|\psi_n\rangle\right) = n\left(\hat{a}|\psi_n\rangle\right)$$

i.e. $\hat{a}^{\dagger}\hat{a}\left(\hat{a}|\psi_n\rangle\right) = (n-1)\left(\hat{a}|\psi_n\rangle\right)$

Properties of raising and lowering operators

But this expression $\hat{a}^\dagger \hat{a}\left(\hat{a}|\psi_n\rangle\right) = (n-1)\left(\hat{a}|\psi_n\rangle\right)$ means

from $\hat{a}^\dagger \hat{a}|\psi_n\rangle = n|\psi_n\rangle$

that $\hat{a}|\psi_n\rangle$ is simply $|\psi_{n-1}\rangle$

at least within some normalizing constant A_n

Hence

$$\hat{a}|\psi_n\rangle = A_n|\psi_{n-1}\rangle$$

and we see why

the operator $\hat{a}$ is called the lowering operator

because it changes the state $|\psi_n\rangle$ into the state $|\psi_{n-1}\rangle$

Properties of raising and lowering operators

We can perform a similar analysis

operating on both sides of $\hat{a}^\dagger \hat{a} |\psi_n\rangle = n|\psi_n\rangle$ with $\hat{a}^\dagger$

The details of this are left as an exercise

The result is $\hat{a}^\dagger \hat{a}\left(\hat{a}^\dagger |\psi_n\rangle\right) = (n+1)\left(\hat{a}^\dagger |\psi_n\rangle\right)$

Again, we conclude from $\hat{a}^\dagger \hat{a} |\psi_n\rangle = n|\psi_n\rangle$ that

$\hat{a}^\dagger |\psi_n\rangle$ is simply $|\psi_{n+1}\rangle$

at least within some normalizing constant B_{n+1}

Hence $\hat{a}^\dagger |\psi_n\rangle = B_{n+1}|\psi_{n+1}\rangle$ and we see why

the operator $\hat{a}^\dagger$ is called the raising operator

because it changes the state $|\psi_n\rangle$ into the state $|\psi_{n+1}\rangle$

Raising and lowering operators

Incidentally, one way to remember which operator

$$\hat{a}|\psi_n\rangle = A_n|\psi_{n-1}\rangle$$

$$\hat{a}^\dagger|\psi_n\rangle = B_{n+1}|\psi_{n+1}\rangle$$

is which

- is to think of the superscript dagger "$\dagger$" as a "$+$" sign
  - corresponding to raising the state
    - Indeed, it is quite a common notation to use a superscript "$+$" sign

Deducing the normalization coefficients A_n and B_n

Premultiplying $\hat{a}\left|\psi_n\right\rangle = A_n\left|\psi_{n-1}\right\rangle$ by $\left\langle\psi_{n-1}\right|$ gives

$$\left\langle\psi_{n-1}\middle|\hat{a}\middle|\psi_n\right\rangle = A_n$$

Now

$$\left\langle\psi_{n-1}\right|\hat{a} = \left[\hat{a}^\dagger\left|\psi_{n-1}\right\rangle\right]^\dagger$$

which can be rewritten as

$$\left\langle\psi_{n-1}\right|\hat{a} = \left[\hat{a}^\dagger\left|\psi_{n-1}\right\rangle\right]^\dagger = \left[B_n\left|\psi_n\right\rangle\right]^\dagger$$

since from

$$\hat{a}^\dagger\left|\psi_n\right\rangle = B_{n+1}\left|\psi_{n+1}\right\rangle$$

we know

$$\hat{a}^\dagger\left|\psi_{n-1}\right\rangle = B_n\left|\psi_n\right\rangle$$

Deducing the normalization coefficients A_n and B_n

Hence $\langle\psi_{n-1}|\hat{a} = \left[B_n|\psi_n\rangle\right]^\dagger = B_n^*\langle\psi_n|$

so $\langle\psi_{n-1}|\hat{a}|\psi_n\rangle = A_n = B_n^*\langle\psi_n|\psi_n\rangle = B_n^*$

Hence $\hat{a}^\dagger\hat{a}|\psi_n\rangle = A_n\hat{a}^\dagger|\psi_{n-1}\rangle = A_nB_n|\psi_n\rangle = |A_n|^2|\psi_n\rangle = n|\psi_n\rangle$

so $A_n = \sqrt{n}$ within a unit complex constant

which we choose to be +1

so $B_n = \sqrt{n}$ also

So instead of $\hat{a}|\psi_n\rangle = A_n|\psi_{n-1}\rangle$ we have $\boxed{\hat{a}|\psi_n\rangle = \sqrt{n}|\psi_{n-1}\rangle}$

Instead of $\hat{a}^\dagger|\psi_n\rangle = B_{n+1}|\psi_{n+1}\rangle$ we have $\boxed{\hat{a}^\dagger|\psi_n\rangle = \sqrt{n+1}|\psi_{n+1}\rangle}$

Harmonic oscillator eigenfunctions

We know that the harmonic oscillator has a lowest state

which corresponds to $n = 0$

Hence, from $\hat{a}|\psi_n\rangle = \sqrt{n}|\psi_{n-1}\rangle$ we must have $\hat{a}|\psi_0\rangle = 0$

This gives an alternative way of deducing $|\psi_0\rangle \equiv \psi_0(\xi)$

Using the differential operator definition of $\hat{a}$, we have

$$\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi} + \xi\right)\psi_0(\xi) = 0$$

which confirms the solution is (after normalization)

$$\psi_0(\xi) = \left[1/(\pi)^{1/4}\right]\exp\left(-\xi^2/2\right)$$

Harmonic oscillator eigenfunctions

Now, we can construct all the solutions for different n

Successive application of $\hat{a}^\dagger$ to $|\psi_0\rangle$ gives

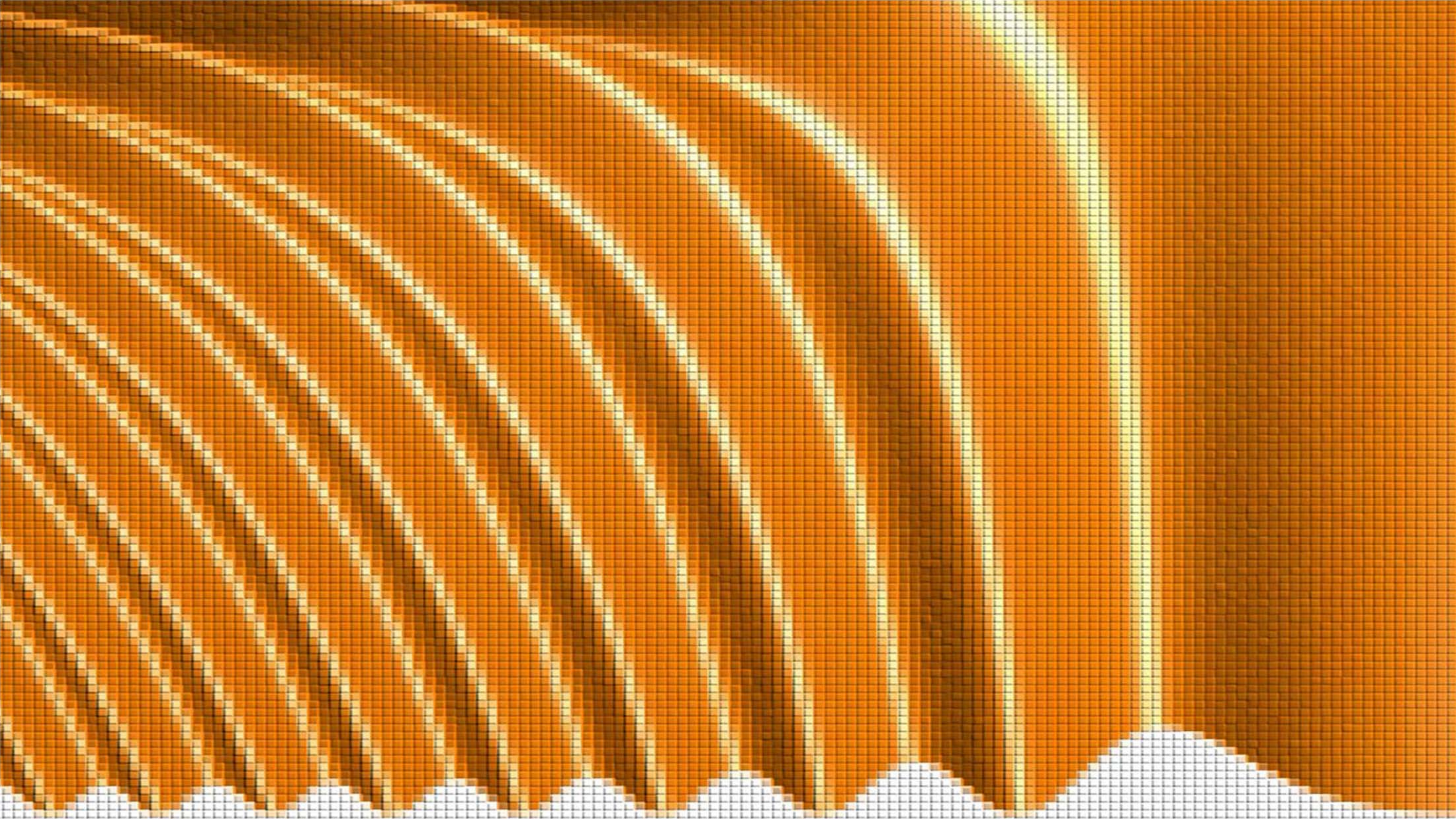
$$\left(\hat{a}^\dagger\right)^n |\psi_0\rangle = \sqrt{n!}|\psi_n\rangle$$

and so the normalized eigenstates can be written as

$$\boxed{|\psi_n\rangle = \frac{1}{\sqrt{n!}}\left(\hat{a}^\dagger\right)^n |\psi_0\rangle}$$

By this approach each eigenfunction can be progressively deduced from preceding ones

We can also use this as a substitution to allow convenient manipulations of the states by operators

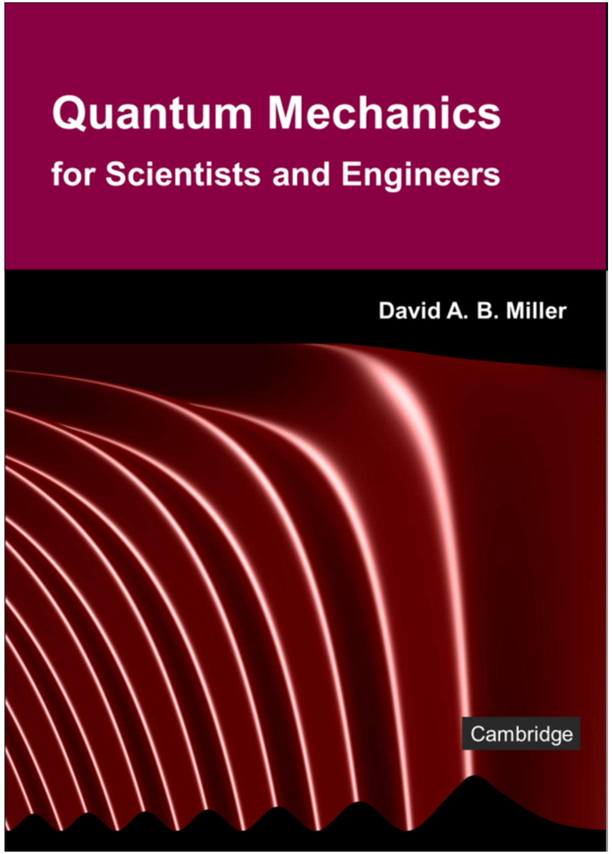


14.3 Operators for harmonic oscillators

Slides: Video 14.3.5 Hamilton's equations

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.2



Operators for harmonic oscillators

Hamilton's equations

Quantum mechanics for scientists and engineers

David Miller

Hamilton’s equations

In classical mechanics

- the Hamiltonian, H, represents the total energy
  - and in the case of one particle in one dimension
    - it is a function of
      - the momentum, p
        - and the position, q

p and q are considered to be independent variables

Hence, in classical mechanics $H = \frac{p^2}{2m} + V(q)$

- where $V(q)$ is the potential energy

Hamilton's equations

The force on the particle

is the negative of the gradient of the potential

(a particle accelerates when going down hill)

i.e., with $H = \frac{p^2}{2m} + V(q)$ the force is $F = -\frac{dV}{dq} \equiv -\frac{\partial H}{\partial q}$

As usual, from Newton's second law

force = rate of change of momentum

so we know that $\boxed{\frac{dp}{dt} = -\frac{\partial H}{\partial q}}$

Hamilton’s equations

With $H = \dfrac{p^2}{2m} + V(q)$

we know also that $\dfrac{\partial H}{\partial p} = \dfrac{p}{m}$

Since $p = mv$ where v is the particle velocity

and, by definition, $v \equiv dq/dt$

we therefore have

$$\boxed{\frac{dq}{dt} = \frac{\partial H}{\partial p}}$$

Hamilton’s equations

The two equations

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \qquad \frac{dq}{dt} = \frac{\partial H}{\partial p}$$

are known as Hamilton’s equations

If the Hamiltonian depends on two quantities p and q

and these quantities and the Hamiltonian obey Hamilton’s equations

then we have found the quantities analogous to momentum and position

Hamilton’s equations

It has been very successful in quantum mechanics

to start with a classical version of the problem

with a Hamiltonian H and quantities p and q

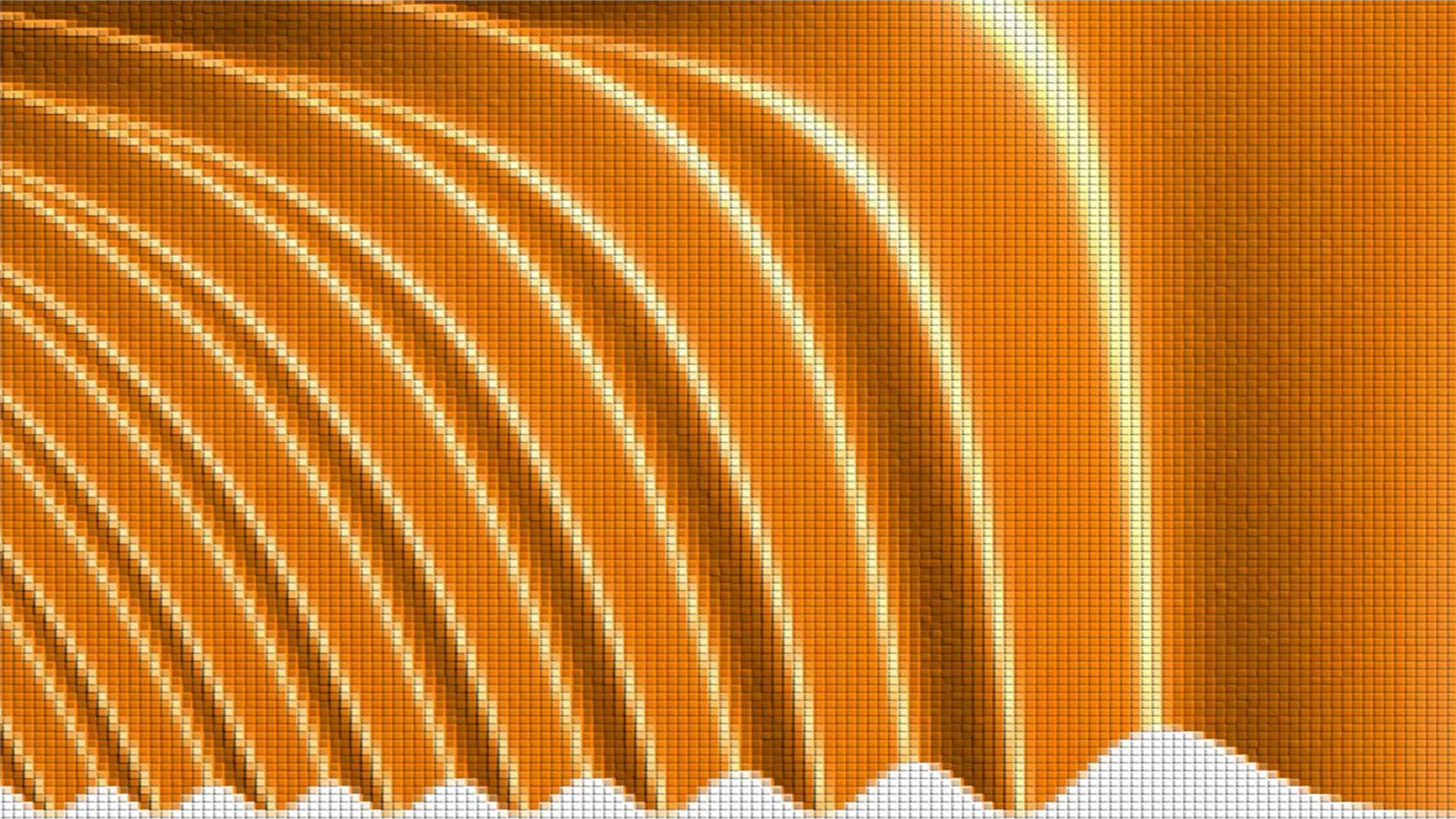
all obeying Hamilton’s equations

and then to propose a quantum version

by substituting a differential operator

$-i\hbar\, d\,/\,dq$ for p

in the corresponding Hamiltonian

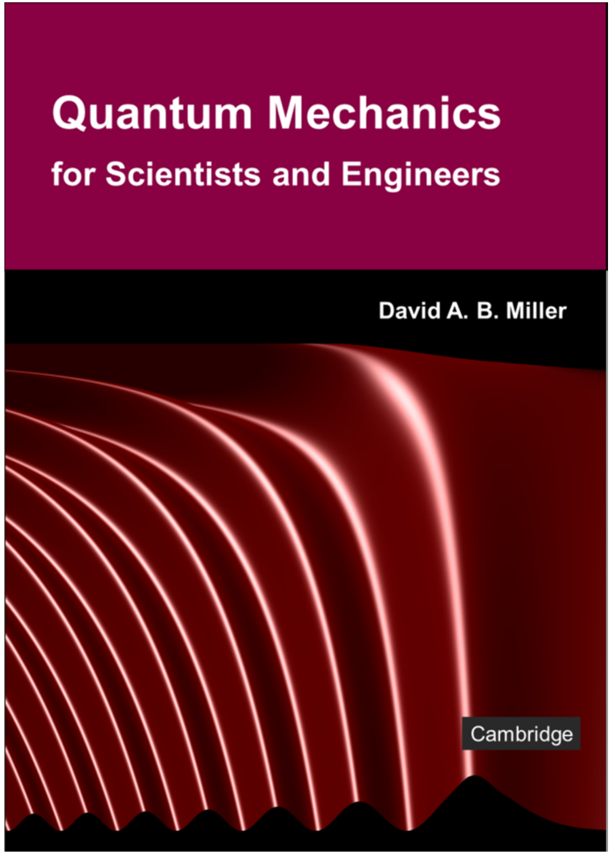


15.1 Quantizing the electromagnetic field

Slides: Video 15.1.1 Field modes as oscillators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.3 subsection "Description of a mode …"



Quantum Mechanics
for Scientists and Engineers

David A. B. Miller

Cambridge

Quantizing the electromagnetic field

Field modes as oscillators

Quantum mechanics for scientists and engineers

David Miller

An electromagnetic field mode

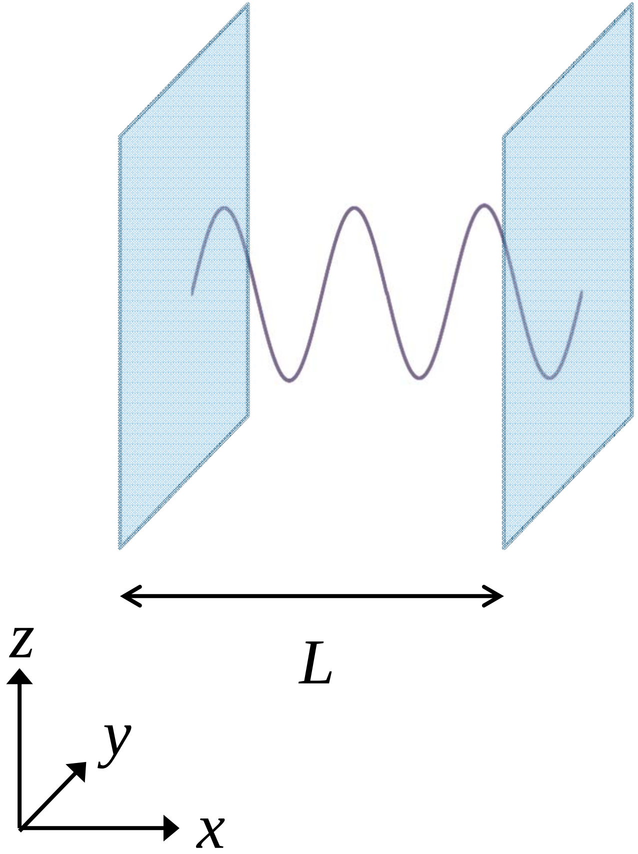
We imagine a box of length L in the x direction

We presume it is arbitrarily large in the other dimensions

and consequently the mode can be described as

a standing plane wave in the x direction

of some wavevector magnitude k



An electromagnetic field mode

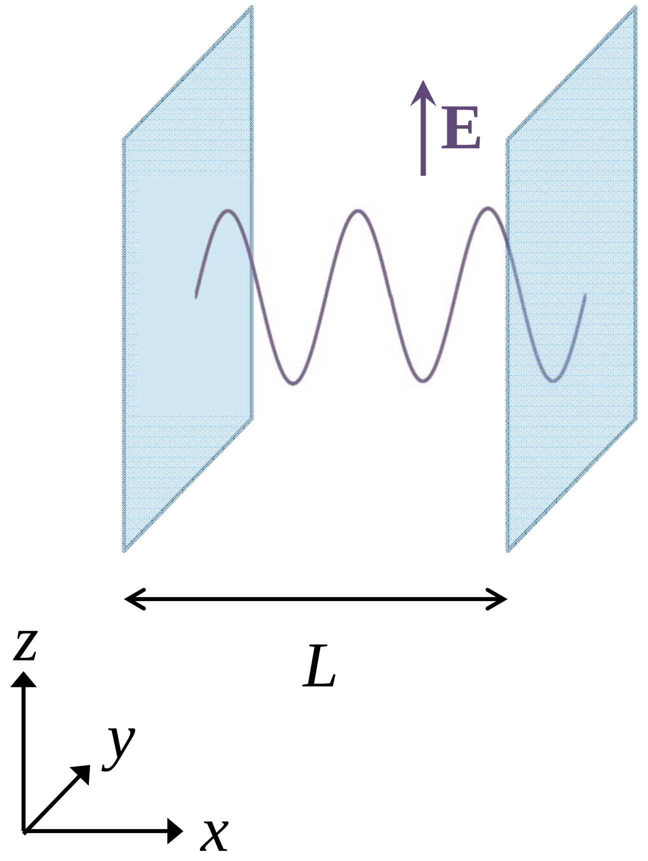
We expect that the electric field $\mathbf{E}$

is perpendicular to the x direction

as both the $\mathbf{E}$ field and the magnetic field $\mathbf{B}$

are transverse to the direction of propagation

for a simple plane electromagnetic wave



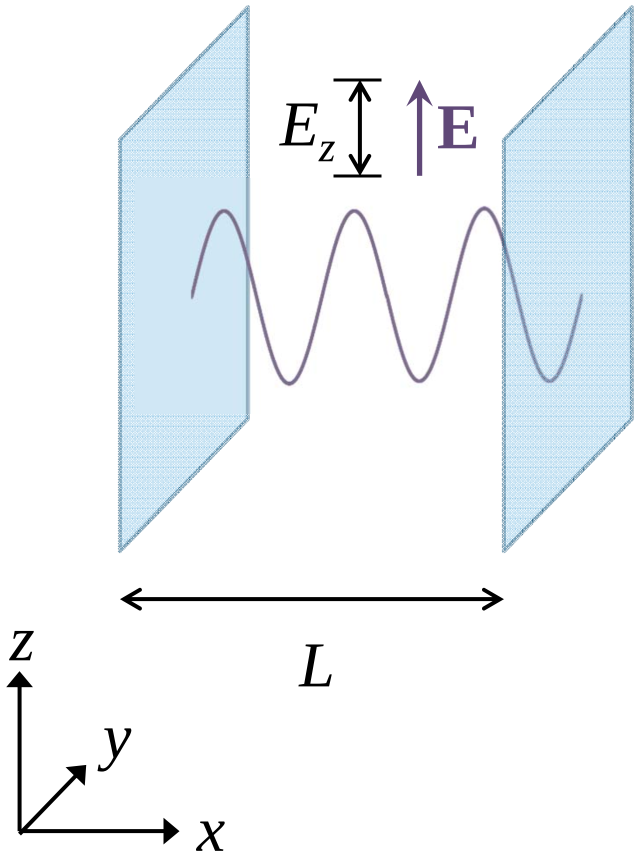
An electromagnetic field mode

We will choose the mode to be polarized in the z direction

with an appropriate amplitude E_z

The $\mathbf{E}$ field in the other two directions

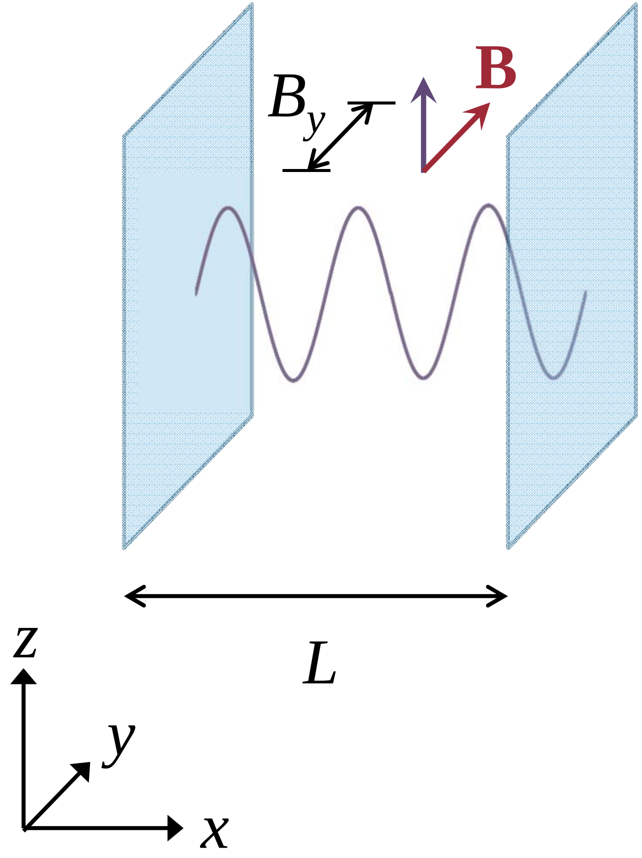
is taken to be zero



An electromagnetic field mode

We also expect that the magnetic field $\mathbf{B}$ is perpendicular to the $\mathbf{E}$ field

- so we choose it polarized in the y direction
  - with amplitude B_y
    - with zero $\mathbf{B}$ field in the other two directions



An electromagnetic field mode

Hence we postulate that

$$E_z = p(t) D \sin kx$$

and

$$B_y = q(t) \frac{D}{c} \cos kx$$

where c is the velocity of light

introduced here for subsequent convenience

D is a constant still to be determined, and

$p(t)$ and $q(t)$ are at the moment simply functions of time yet to be determined

An electromagnetic field mode

We now check that these fields satisfy the appropriate Maxwell's equations

- which will justify all our postulations about these classical fields
  - and will tell us some other required relations between our postulated quantities

We now presume that we are in a vacuum

- so no charge density and no magnetic materials
  - and the permittivity and permeability are their vacuum values of ε_o and μ_o respectively

An electromagnetic field mode

Using the Maxwell equation

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

with E_x and E_y both zero by choice

and noting that $\partial E_z / \partial y = 0$

because we have an infinite plane wave with no variation in the y direction

then we are left with

$$\frac{\partial E_z}{\partial x} = \frac{\partial B_y}{\partial t}$$

An electromagnetic field mode

So with our choices $E_z = p(t) D \sin kx$ and $B_y = q(t) \frac{D}{c} \cos kx$

then our result from Maxwell's equations $\frac{\partial E_z}{\partial x} = \frac{\partial B_y}{\partial t}$

gives the requirement $kpD \cos kx = \frac{D}{c} \frac{\partial q}{\partial t} \cos kx$

so $\frac{dq}{dt} = \omega p$

where $\omega = kc$

An electromagnetic field mode

Similarly, using the Maxwell equation

$$\nabla \times \mathbf{B} = \varepsilon_o \mu_o \frac{\partial \mathbf{E}}{\partial t}$$

with B_x and B_z both zero by choice

and noting that $\partial B_y / \partial z = 0$

because we have an infinite plane wave with no variation in the z direction

then we are left with

$$\frac{\partial B_y}{\partial x} = \varepsilon_o \mu_o \frac{\partial E_z}{\partial t}$$

An electromagnetic field mode

So with our choices $E_z = p(t) D \sin kx$ and $B_y = q(t) \frac{D}{c} \cos kx$

then our result from Maxwell's equations $\frac{\partial B_y}{\partial x} = \varepsilon_o \mu_o \frac{\partial E_z}{\partial t}$

gives the requirement $-kq \frac{D}{c} \sin kx = \varepsilon_o \mu_o \frac{dp}{dt} D \sin kx$

i.e., using the relation $\varepsilon_o \mu_o = \frac{1}{c^2}$

we have $\frac{dp}{dt} = -\omega q$

An electromagnetic field mode

So we have found that our postulated form for the mode of the radiation field

$$E_z = p(t)D\sin kx \qquad B_y = q(t)\frac{D}{c}\cos kx$$

does indeed satisfy the two Maxwell equations

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \nabla \times \mathbf{B} = \varepsilon_o \mu_o \frac{\partial \mathbf{E}}{\partial t}$$

provided we have the relations

$$\frac{dq}{dt} = \omega p \qquad \frac{dp}{dt} = -\omega q$$

between our time-varying amplitudes p and q

An electromagnetic field mode

Differentiating $\frac{dq}{dt} = \omega p$ with respect to time t

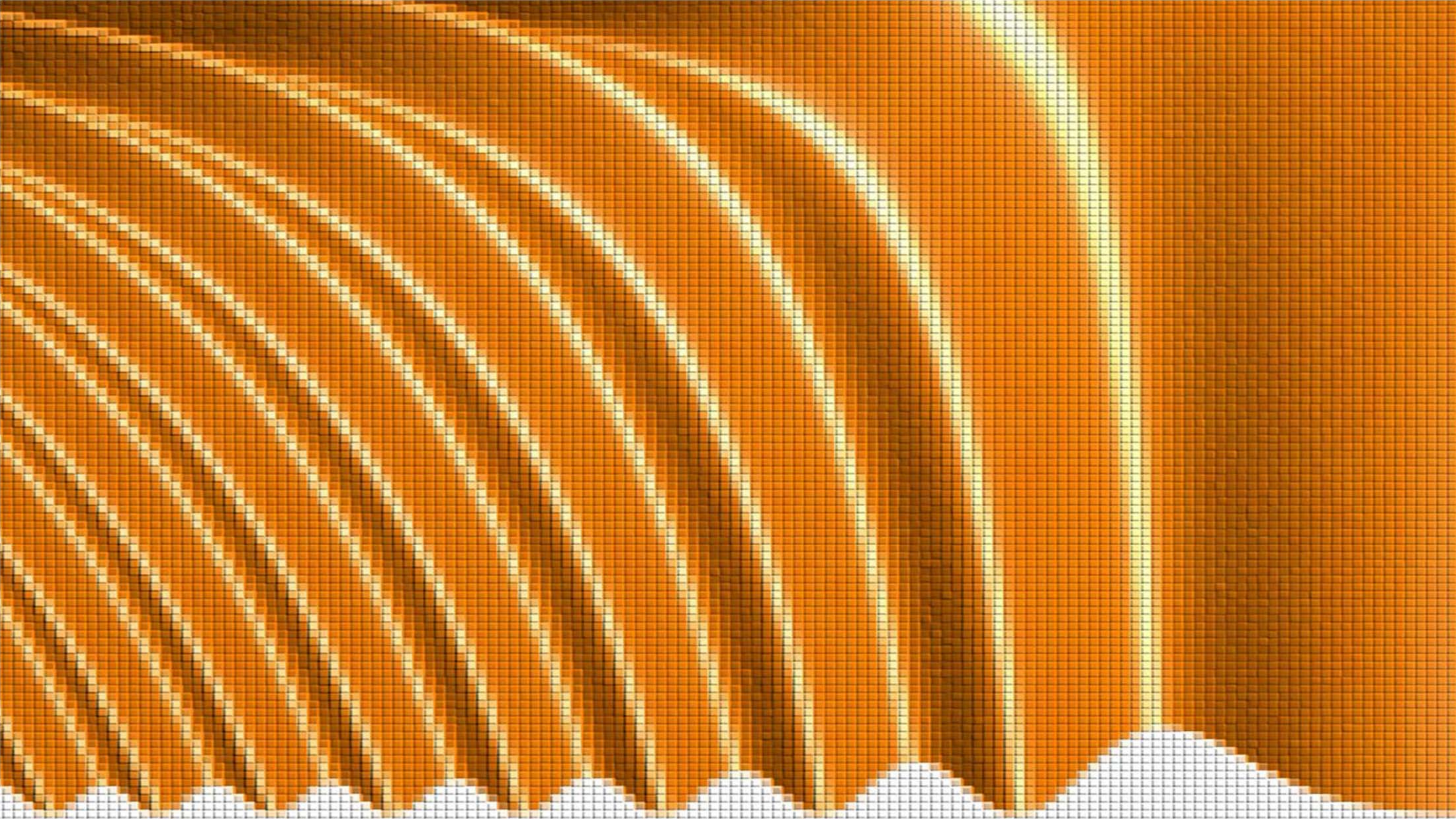
and substituting from $\frac{dp}{dt} = -\omega q$

we find $\frac{d^2 q}{dt^2} = -\omega^2 q$

which means that the electromagnetic mode

does indeed behave exactly like a harmonic oscillator

with oscillation (angular) frequency ω

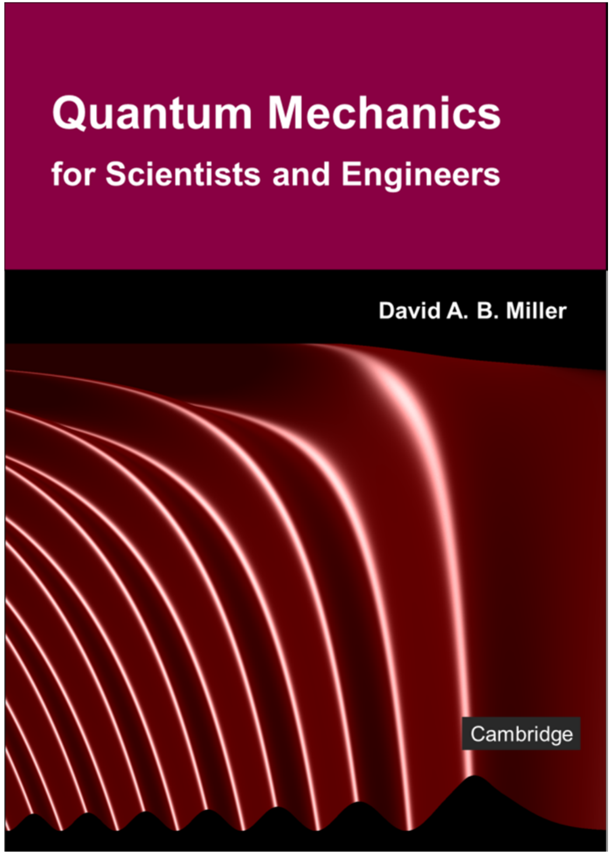


15.1 Quantizing the electromagnetic field

Slides: Video 15.1.3 Electromagnetic mode Hamiltonian

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.3 from subsection “Hamiltonian for …” to the paragraph after Eq. 15.64



Quantizing the electromagnetic field

Electromagnetic mode Hamiltonian

Quantum mechanics for scientists and engineers

David Miller

Hamiltonian for an electromagnetic mode

Formally in an electromagnetic field in a vacuum

the energy density is $W = \frac{1}{2}\left(\varepsilon_o \mathbf{E}^2 + \frac{1}{\mu_o}\mathbf{B}^2\right)$

In a box of length L, then, per unit cross-sectional area, the total energy is the Hamiltonian

$$H = \int_0^L W dx = \frac{D^2}{2}\int_0^L \left(\varepsilon_o p^2 \sin^2 kx + \frac{1}{\mu_o c^2} q^2 \cos^2 kx\right) dx$$

$$= \frac{D^2 \varepsilon_o}{2}\int_0^L \left(p^2 \sin^2 kx + q^2 \cos^2 kx\right) dx = \frac{D^2 L \varepsilon_o}{4}\left[p^2 + q^2\right]$$

Hamiltonian for an electromagnetic mode

With our Hamiltonian $H = \frac{D^2 L \varepsilon_o}{4}\left[p^2 + q^2 \right]$

we now try to choose D so as to get p and q

to correspond to the analogs of momentum and position with this Hamiltonian

by having H, p, and q obey Hamilton's equations

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \qquad \frac{dq}{dt} = \frac{\partial H}{\partial p}$$

Hamiltonian for an electromagnetic mode

If we choose $D = \sqrt{2\omega / L\varepsilon_o}$

then $H = \frac{D^2 L\varepsilon_o}{4}\left[p^2 + q^2\right] = \frac{\omega}{2}\left(p^2 + q^2\right)$

so $\frac{\partial H}{\partial p} = \omega p$ and $\frac{\partial H}{\partial q} = \omega q$

But we already deduced that $\frac{dq}{dt} = \omega p$ and $\frac{dp}{dt} = -\omega q$

so our chosen H, p, and q obey Hamilton's equations

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \qquad \frac{dq}{dt} = \frac{\partial H}{\partial p}$$

Quantizing the Hamiltonian

Having derived a classical Hamiltonian for an electromagnetic mode

we now proceed to quantize it

We postulate that we can substitute the operator $\hat{p} = -i\hbar \frac{d}{dq}$

for the scalar quantity p

of the classical Hamiltonian $H = \frac{\omega}{2}\left(p^2 + q^2\right)$

obtaining $\hat{H} = \frac{\omega}{2}\left[-\hbar^2 \frac{d^2}{dq^2} + q^2\right]$

Quantizing the Hamiltonian

For convenience we define the dimensionless unit $\xi = q / \sqrt{\hbar}$

For future use, we also can define a

dimensionless momentum operator $\hat{\pi} = \hat{p} / \sqrt{\hbar} \equiv -i \frac{d}{d\xi}$

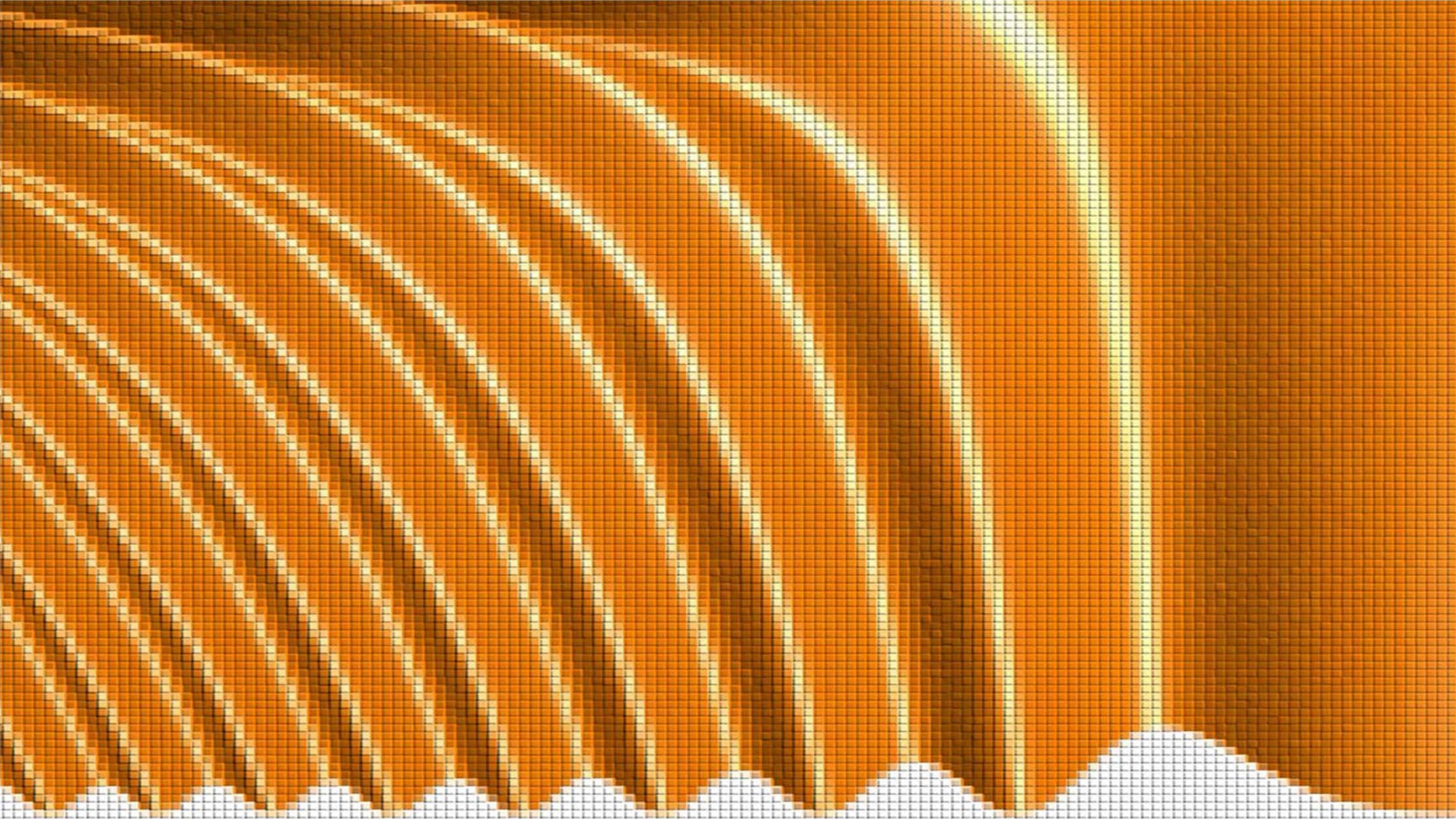
In these dimensionless units

the Hamiltonian can be written in the form

$$\hat{H} = \frac{\hbar\omega}{2}\left[-\frac{d^2}{d\xi^2} + \xi^2\right]$$

identical to the harmonic oscillator Hamiltonian

or in the form $\hat{H} = \frac{\hbar\omega}{2}\left(\hat{\pi}^2 + \xi^2\right)$

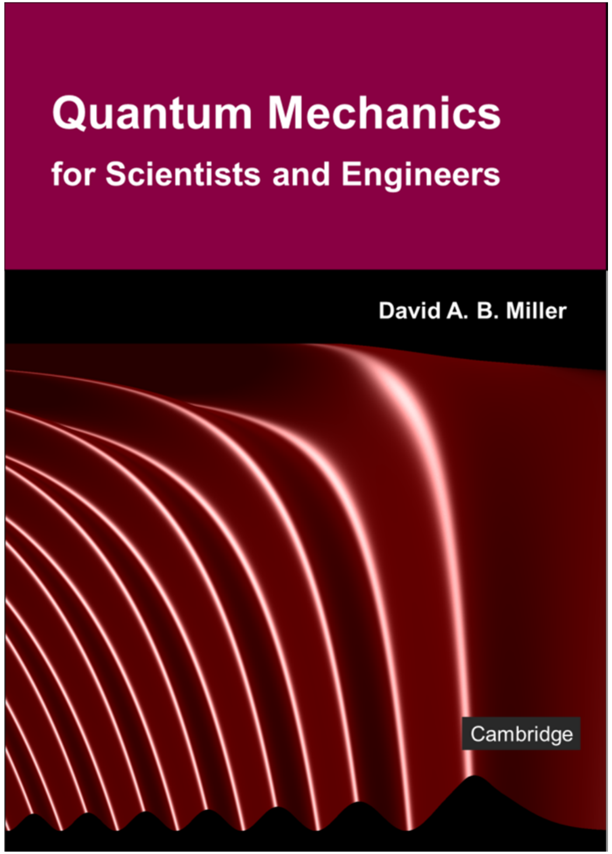


15.1 Quantizing the electromagnetic field

Slides: Video 15.1.5 Quantum states of an electromagnetic mode

Text reference: Quantum Mechanics for Scientists and Engineers

Sections 15.3 from two paragraphs above Eq. 15.65 to end of 15.4



Quantizing the electromagnetic field

Quantum states of an electromagnetic mode

Quantum mechanics for scientists and engineers

David Miller

Mode notation

To show which mode the operators refer to

- we commonly use λ to index the different modes
  - Note λ here does not necessarily mean wavelength
    - It is just an integer index

With that new notation

- we have for a given mode, for example,
  - an angular frequency ω_λ
    - instead of just ω

Mode notation

With that notation, for a given mode we will also have

- a Hamiltonian $\hat{H}_\lambda$
- creation and annihilation operators $\hat{a}_\lambda^\dagger$ and $\hat{a}_\lambda$
- and a number operator $\hat{N}_\lambda$

We can also label the eigenstates similarly as $|\psi_{\lambda n}\rangle$

- being the nth eigenstate associated with the mode

We should also change to using the coordinate ξ_λ

- since each different mode will have its own corresponding coordinate

Mode notation

With this notation, we use some harmonic oscillator results

to give key relations for the electromagnetic mode λ

The Hamiltonian becomes $\hat{H}_\lambda \equiv \hbar\omega_\lambda \left(\hat{a}_\lambda^\dagger \hat{a}_\lambda + \frac{1}{2} \right)$

The number operator becomes $\hat{N}_\lambda \equiv \hat{a}_\lambda^\dagger \hat{a}_\lambda$

and so we have $\hat{N}_\lambda \left|\psi_{\lambda n}\right\rangle = \hat{a}_\lambda^\dagger \hat{a}_\lambda \left|\psi_{\lambda n}\right\rangle = n_\lambda \left|\psi_{\lambda n}\right\rangle$

and the corresponding eigen energy is

$$E_{\lambda n} = \hbar\omega \left(n_\lambda + \frac{1}{2} \right)$$

which grows $\propto n_\lambda$, i.e., as $n_\lambda \hbar\omega$

Mode notation

We also have

the commutation relation $\left[\hat{a}_\lambda, \hat{a}_\lambda^\dagger\right] = \hat{a}_\lambda \hat{a}_\lambda^\dagger - \hat{a}_\lambda^\dagger \hat{a}_\lambda = 1$

the lowering relation $\hat{a}_\lambda \left|\psi_{\lambda n}\right\rangle = \sqrt{n_\lambda}\left|\psi_{\lambda n-1}\right\rangle$

which takes the state with n_λ photons in mode λ

and changes it into the state with $n_\lambda - 1$ photons

so we now call hence we call $\hat{a}_\lambda$

the annihilation operator for mode λ

the raising relation $\hat{a}_\lambda^\dagger \left|\psi_{\lambda n}\right\rangle = \sqrt{n_\lambda + 1}\left|\psi_{\lambda n+1}\right\rangle$

so we similarly call $\hat{a}_\lambda^\dagger$

the creation operator for mode λ

Mode notation

We will also have that trying to take a photon out of an "empty" mode

will lead to zero result $\hat{a}_\lambda |\psi_{\lambda 0}\rangle = 0$

just as the same mathematics did for our simple harmonic oscillator

and we will also be able to write, as before

$$|\psi_{\lambda n}\rangle = \frac{1}{\sqrt{n_\lambda !}} \left(\hat{a}_\lambda^\dagger\right)^{n_\lambda} |\psi_{\lambda 0}\rangle$$

though now we think of it as

the state with n_λ photons in mode λ

Nature of quantum states of modes

We have quantized the electromagnetic mode by analogy

- leading to an abstract set of results
  - Do we have to be this abstract?
    - Can't we use the wavefunction as before?

The wavefunction does have some meaning

- though it is quite different, for example
  - from that of the electron spatial wavefunction

Nature of quantum states of modes

Just as before

we have for the state with no photons in mode λ

$$\psi_{\lambda 0}(\xi_\lambda) = \frac{1}{(\pi)^{1/4}} \exp\left(-\frac{\xi_\lambda^2}{2}\right)$$

but if we now work backwards

to find the physical interpretation of the coordinate ξ_λ

we find

$$B_y = \xi_\lambda \sqrt{\frac{2\mu_o \hbar \omega_\lambda}{L}} \cos kx$$

Nature of quantum states of modes

In other words, ξ_λ is, in a dimensionless form

- the amplitude of the mode of the magnetic field
  - It is not a spatial coordinate

For example, we can interpret $\left|\psi_{\lambda 0}\left(\xi_\lambda\right)\right|^2$

- as the probability that
  - in this state of electromagnetic field mode λ
    - the mode has (dimensionless) amplitude ξ_λ

That probability is therefore the Gaussian, $\left(1/\sqrt{\pi}\right)\exp\left(-\xi_\lambda{}^2\right)$

- We would find related results for the states of the mode with more photons

Wavefunctions in quantized field modes

Though we may sometimes be interested in these

- distributions of magnetic or electric field amplitude
  - which we can deduce from this "wavefunction"

we are generally much less interested in these

- than we were in the
  - probabilities of finding particles at points in space

As a result, in the quantized electromagnetic field

- we make relatively little use of wavefunctions

Wavefunctions in quantized field modes

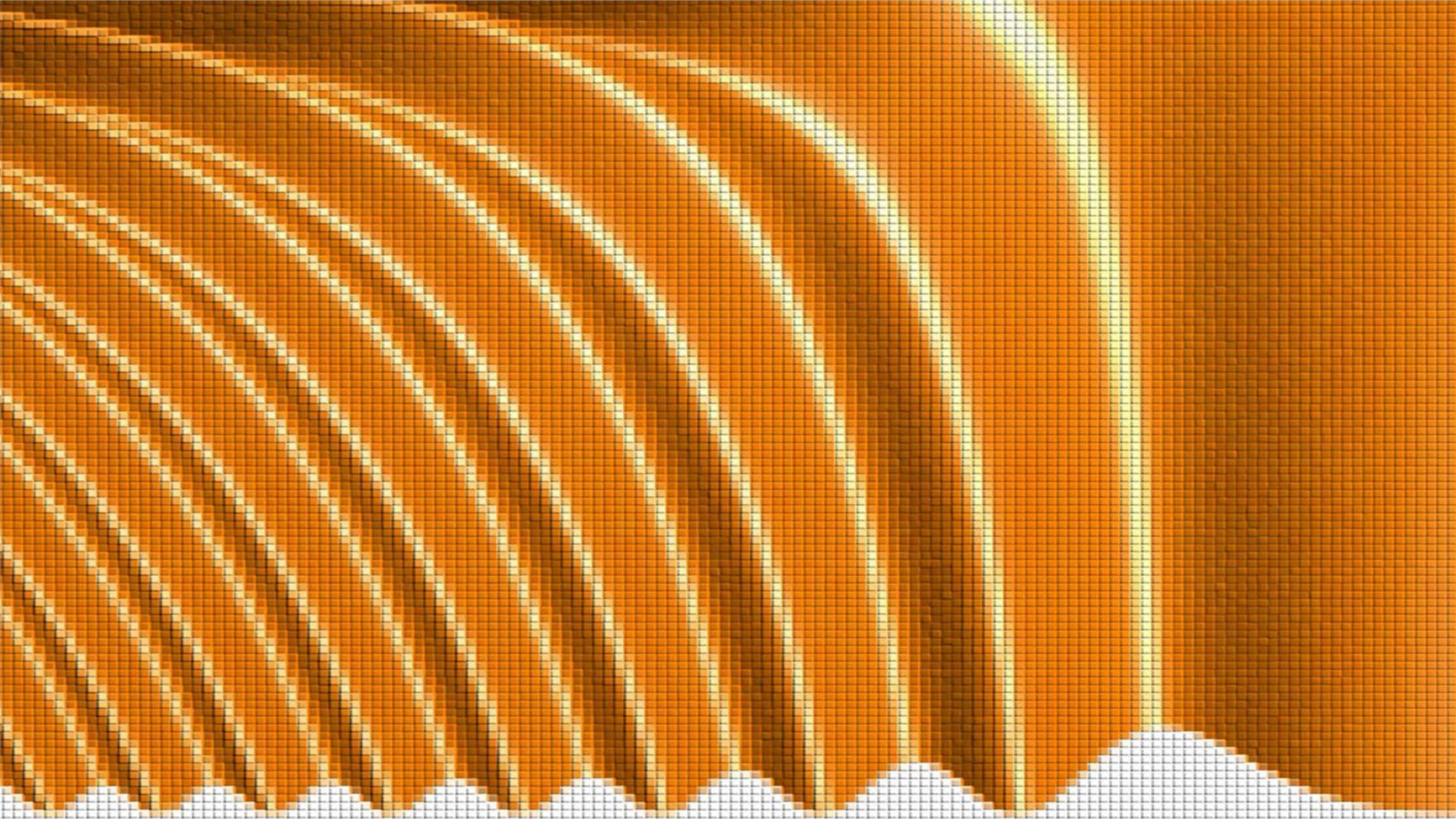
Most of the results we are interested in

- such as processes adding or subtracting photons
  - can more conveniently be described through the use of operators and state vectors

Typically, the basis set and the resulting state we will use

- will not be written as functions, $\psi(\xi_\lambda)$ of the amplitudes, ξ_λ, of the fields in the modes
  - but as basis vectors corresponding to specific numbers of photons in a mode

$$|\psi_{\lambda n}\rangle \equiv |n_\lambda\rangle$$

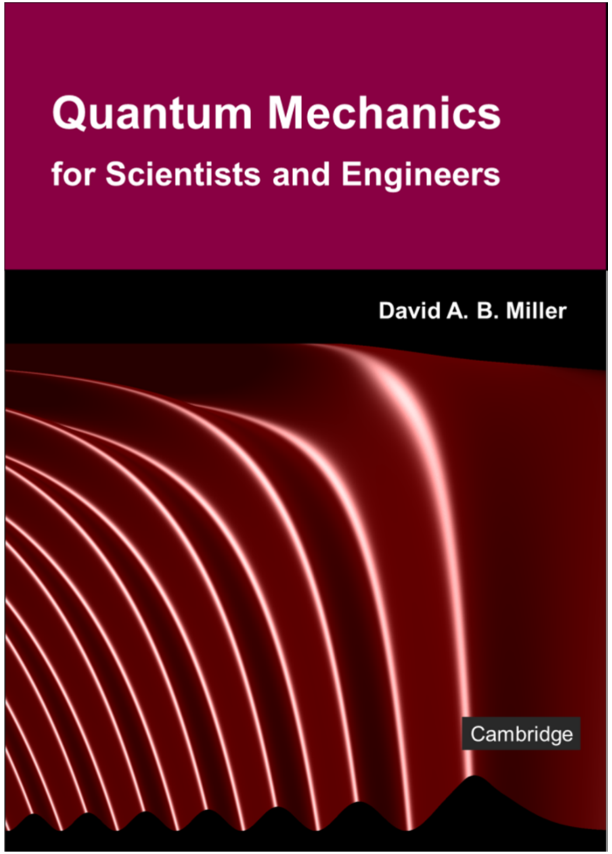


15.1 Quantizing the electromagnetic field

Slides: Video 15.1.7 Field operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.5



Quantizing the electromagnetic field

Field operators

Quantum mechanics for scientists and engineers

David Miller

ξ_λ as an operator

From the original definitions of the creation and annihilation operators in terms of ξ and $d/d\xi$, we have

$$\hat{a}_\lambda^\dagger \equiv \frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi_\lambda}+\xi_\lambda\right) \quad \hat{a}_\lambda \equiv \frac{1}{\sqrt{2}}\left(\frac{d}{d\xi_\lambda}+\xi_\lambda\right)$$

Now we note that we can write

$$\hat{\xi}_\lambda \equiv \frac{1}{\sqrt{2}}\left(\hat{a}_\lambda + \hat{a}_\lambda^\dagger\right)$$

Note that ξ_λ (which we can now write as $\hat{\xi}_\lambda$)

is really an operator, not just a coordinate

just as position $\mathbf{r}$ was really an operator $\hat{\mathbf{r}}$

π_λ as an operator

Similarly

we can also write the dimensionless form of

the generalized momentum operator

defined as in

$$\hat{\pi}_\lambda = -i d / d\xi_\lambda \equiv \hat{p}_\lambda / \sqrt{\hbar}$$

in the form

$$\hat{\pi}_\lambda = \frac{i}{\sqrt{2}}\left(\hat{a}_\lambda^\dagger - \hat{a}_\lambda\right)$$

Field operators

With these definitions of $\hat{\xi}_\lambda$ and $\hat{\pi}_\lambda$

we now substitute back into the relations

$$E_z = p(t) D \sin kx \qquad B_y = q(t)[D/c]\cos kx$$

that defined the mode’s electric and magnetic fields

Instead of scalar quantities for the electric and magnetic fields for this mode

we now have operators

$$\hat{E}_{\lambda z} = i\left(\hat{a}_\lambda^\dagger - \hat{a}_\lambda\right)\sqrt{\frac{\hbar\omega_\lambda}{\varepsilon_o L}}\sin kx \qquad \hat{B}_{\lambda y} = \left(\hat{a}_\lambda^\dagger + \hat{a}_\lambda\right)\sqrt{\frac{\mu_o \hbar\omega_\lambda}{L}}\cos kx$$

Meaning of field operators

Just as before

- if we want to know the average value of a measurable quantity
  - we take the expected value of its operator
    - and the same is true here

For a state $\left|\phi\right\rangle$ of this mode

- we would have

$$\left\langle E_{\lambda z}\right\rangle = \left\langle \phi \middle| \hat{E}_{\lambda z} \middle| \phi \right\rangle$$

$$\left\langle B_{\lambda y}\right\rangle = \left\langle \phi \middle| \hat{B}_{\lambda y} \middle| \phi \right\rangle$$

Commutator for electric and magnetic fields

With these field operators

we can now work out the commutator

for the electric and magnetic fields

Explicitly, from

$$\hat{E}_{\lambda z} = i\left(\hat{a}_\lambda^\dagger - \hat{a}_\lambda\right)\sqrt{\frac{\hbar\omega_\lambda}{\varepsilon_o L}}\sin kx \qquad \hat{B}_{\lambda y} = \left(\hat{a}_\lambda^\dagger + \hat{a}_\lambda\right)\sqrt{\frac{\mu_o \hbar\omega_\lambda}{L}}\cos kx$$

we have

$$\left[\hat{E}_{\lambda z}, \hat{B}_{\lambda y}\right] = i\frac{\hbar\omega_\lambda}{L}\sqrt{\frac{\mu_o}{\varepsilon_o}}\sin kx \cos kx\left[\hat{a}_\lambda^\dagger - \hat{a}_\lambda, \hat{a}_\lambda + \hat{a}_\lambda^\dagger\right]$$

Commutator for electric and magnetic fields

Multiplying out this algebra

$$\left[\hat{E}_{\lambda z},\hat{B}_{\lambda y}\right]=i\frac{\hbar\omega_\lambda}{L}\sqrt{\frac{\mu_o}{\varepsilon_o}}\sin kx\cos kx\left[\hat{a}_\lambda^\dagger-\hat{a}_\lambda,\hat{a}_\lambda+\hat{a}_\lambda^\dagger\right]$$

$$=i\frac{\hbar\omega_\lambda}{L}\sqrt{\frac{\mu_o}{\varepsilon_o}}\sin kx\cos kx$$

$$\times\left(\hat{a}_\lambda^\dagger\hat{a}_\lambda+\hat{a}_\lambda^\dagger\hat{a}_\lambda^\dagger-\hat{a}_\lambda\hat{a}_\lambda-\hat{a}_\lambda\hat{a}_\lambda^\dagger-\hat{a}_\lambda\hat{a}_\lambda^\dagger+\hat{a}_\lambda\hat{a}_\lambda-\hat{a}_\lambda^\dagger\hat{a}_\lambda^\dagger+\hat{a}_\lambda^\dagger\hat{a}_\lambda\right)$$

$$=2i\frac{\hbar\omega_\lambda}{L}\sqrt{\frac{\mu_o}{\varepsilon_o}}\sin kx\cos kx\left[\hat{a}_\lambda^\dagger\hat{a}_\lambda-\hat{a}_\lambda\hat{a}_\lambda^\dagger\right]$$

Commutation relation for fields

But we know that

$$\hat{a}_\lambda^\dagger \hat{a}_\lambda - \hat{a}_\lambda \hat{a}_\lambda^\dagger = -\left[\hat{a}_\lambda \hat{a}_\lambda^\dagger - \hat{a}_\lambda^\dagger \hat{a}_\lambda\right] = -1$$

so from our previous result

$$\left[\hat{E}_{\lambda z}, \hat{B}_{\lambda y}\right] = 2i \frac{\hbar \omega_\lambda}{L} \sqrt{\frac{\mu_o}{\varepsilon_o}} \sin kx \cos kx \left[\hat{a}_\lambda^\dagger \hat{a}_\lambda - \hat{a}_\lambda \hat{a}_\lambda^\dagger\right]$$

we have the commutation relation for the electric and magnetic field operators in this mode

$$\left[\hat{E}_{\lambda z}, \hat{B}_{\lambda y}\right] = -2i \frac{\hbar \omega_\lambda}{L} \sqrt{\frac{\mu_o}{\varepsilon_o}} \sin kx \cos kx$$

Uncertainty principle for electric and magnetic fields

The general form of the commutation relation $\left[\hat{A},\hat{B}\right]=i\hat{C}$

gives the uncertainty principle $\Delta A \Delta B \geq \left|C\right|/2$

so the standard deviations of the expected values

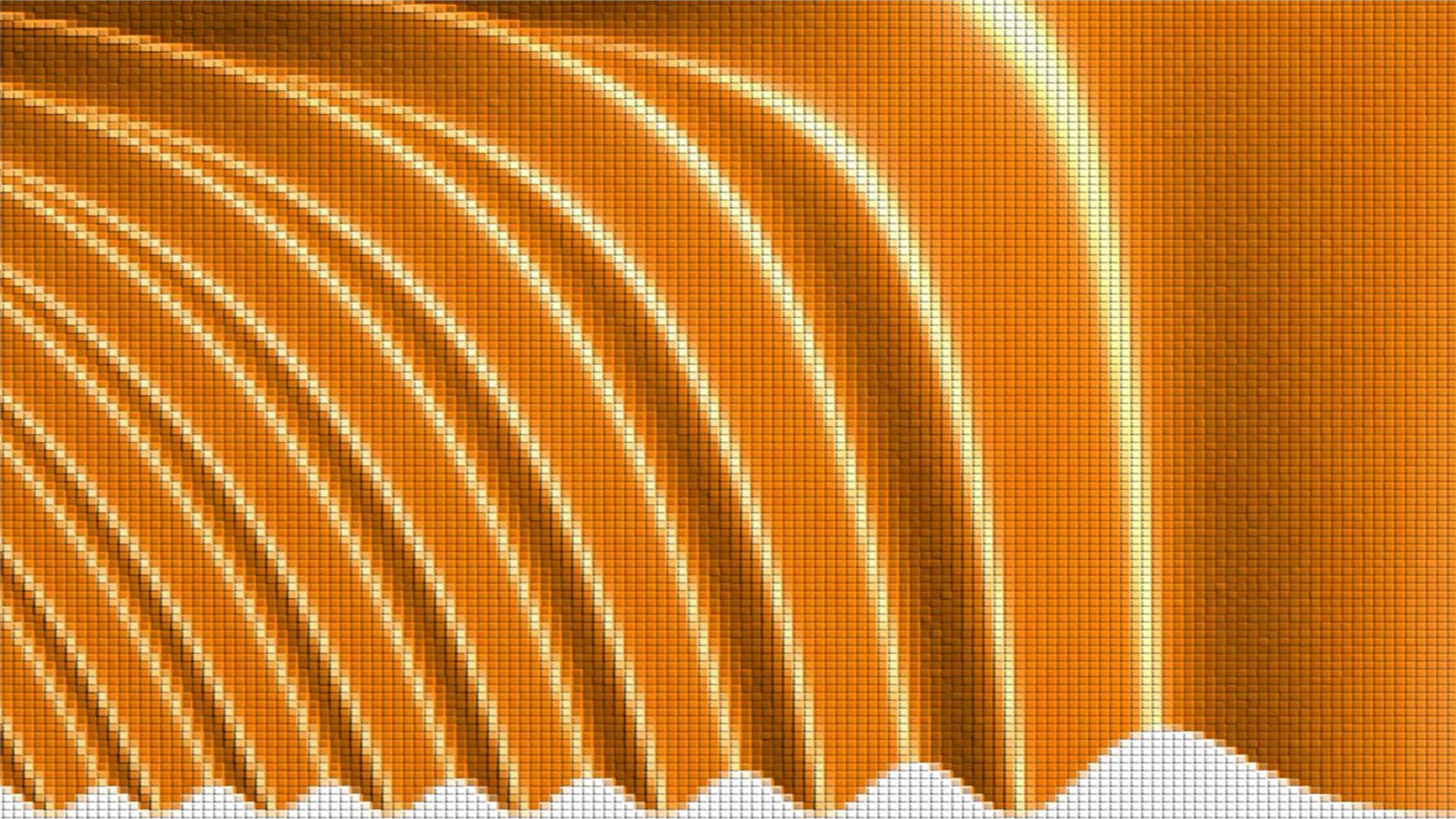
of the electric and magnetic field amplitudes in this mode obey

$$\Delta E_{\lambda z} \Delta B_{\lambda y} \geq \frac{\hbar \omega_\lambda}{L}\sqrt{\frac{\mu_o}{\varepsilon_o}}\sin kx \cos kx$$

These field operators do not commute in general

We cannot in general simultaneously know

both the electric and magnetic field exactly!

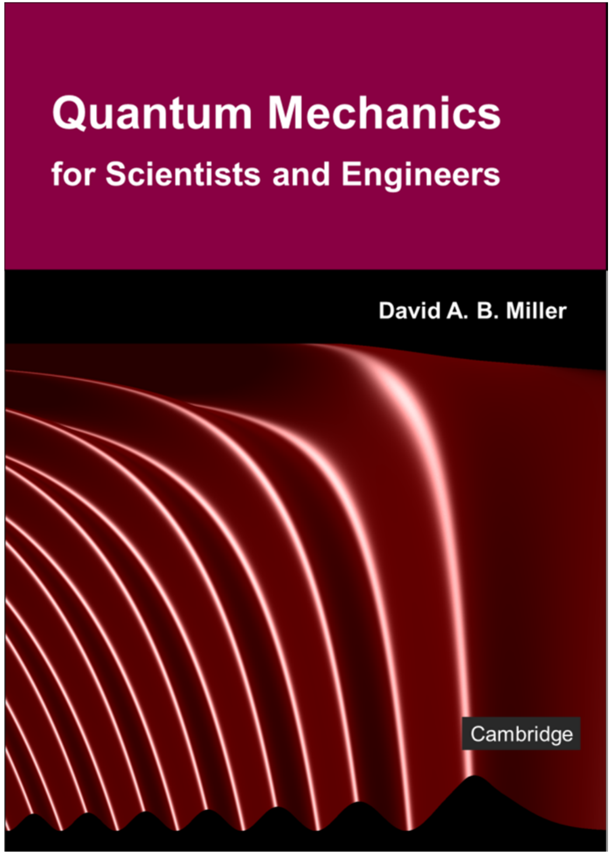


15.2 Quantum states of the electromagnetic field

Slides: Video 15.2.1 Number states

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.6 to end of subsection "Representation of time dependence …"



Quantum states of the electromagnetic field

Quantum mechanics for scientists and engineers

David Miller

Quantum states of the electromagnetic field

Number states

Quantum mechanics for scientists and engineers

David Miller

Number states

The Hamiltonian and number operator eigenstates $|n_\lambda\rangle$

correspond to n_λ photons in the mode

and are known as the number states or Fock states

In these states

the probability of measuring any particular amplitude

$B_{\lambda y}$ in the mode

is distributed according to

the square of the Hermite-Gaussian harmonic oscillator solutions with quantum number n_λ

The $E_{\lambda z}$ amplitudes are similarly distributed

Number states

The expectation values of the electric and magnetic field amplitudes are both zero for any number state, e.g.,

$$\langle n_\lambda | \hat{E}_{\lambda z} | n_\lambda \rangle = i\sqrt{\frac{\hbar\omega_\lambda}{L\varepsilon_o}} \sin kx \langle n_\lambda | \hat{a}_\lambda^\dagger - \hat{a}_\lambda | n_\lambda \rangle$$

$$= i\sqrt{\frac{\hbar\omega_\lambda}{L\varepsilon_o}} \sin kx \left(\sqrt{n_\lambda + 1} \langle n_\lambda | n_\lambda + 1 \rangle - \sqrt{n_\lambda} \langle n_\lambda | n_\lambda - 1 \rangle \right) = 0$$

because the states $|n_\lambda\rangle$, $|n_\lambda - 1\rangle$, and $|n_\lambda + 1\rangle$

are all eigenstates of the same Hamiltonian

and so are orthogonal

and similarly for the magnetic field mode amplitude

Schrödinger and Heisenberg representations

So far, we have used solutions to the time-independent Schrödinger equation for the electromagnetic mode

Here we use the term "Schrödinger equation" in the generalized sense where we mean that

$$\hat{H}\left|\phi\right\rangle = E\left|\phi\right\rangle$$

is a Schrödinger equation for a system

in an eigenstate $\left|\phi\right\rangle$ with eigenenergy E

Explicitly, for the eigenstates of our electromagnetic mode, we have

$$\hat{H}\left|n_\lambda\right\rangle = \left(n_\lambda + 1/2\right)\hbar\omega_\lambda\left|n_\lambda\right\rangle$$

Schrödinger and Heisenberg representations

Generalizing our earlier postulations

- we also postulate here that
  - the time-dependent generalized Schrödinger equation is valid, i.e.,

$$\hat{H}|\phi\rangle = i\hbar\frac{\partial}{\partial t}|\phi\rangle$$

- even if our Hamiltonian is not the one in our original Schrödinger equation for an electron
  - This postulation does appear to work

Time-dependence and number states

With this approach to describing time-dependence

as before, to get the time variation of a given state

we multiply the time-independent energy eigenstates

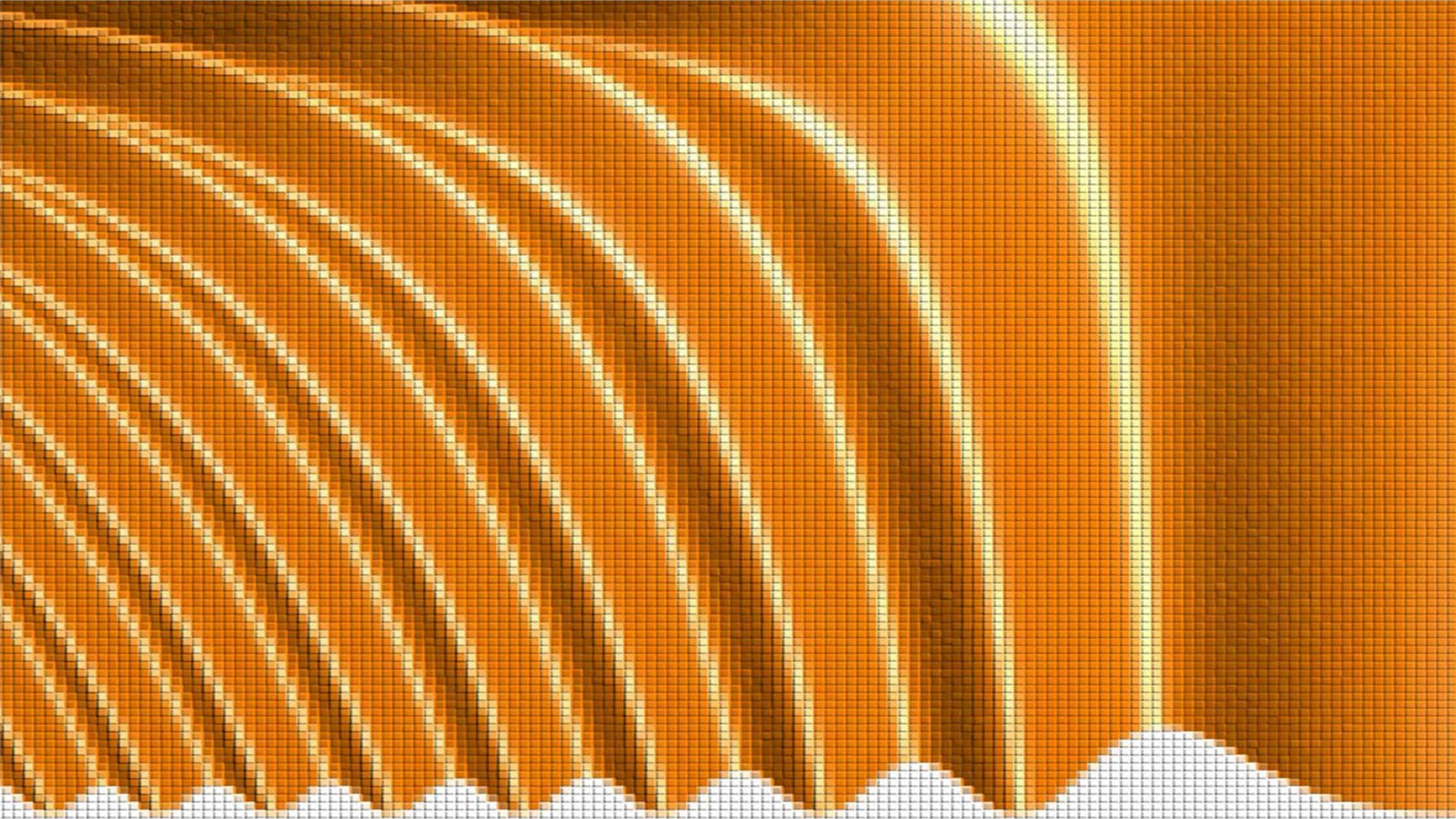
by $\exp\left[-i\left(n_\lambda + 1/2\right)\hbar\omega_\lambda t/\hbar\right] = \exp\left[-i\left(n_\lambda + 1/2\right)\omega_\lambda t\right]$

to make $\hat{H}\left|n_\lambda\right\rangle = \left(n_\lambda + 1/2\right)\hbar\omega_\lambda\left|n_\lambda\right\rangle$

consistent with $\hat{H}\left|\phi\right\rangle = i\hbar\frac{\partial}{\partial t}\left|\phi\right\rangle$

so including time-dependence the number states become

$$\exp\left[-i\left(n_\lambda + 1/2\right)\omega_\lambda t\right]\left|n_\lambda\right\rangle$$

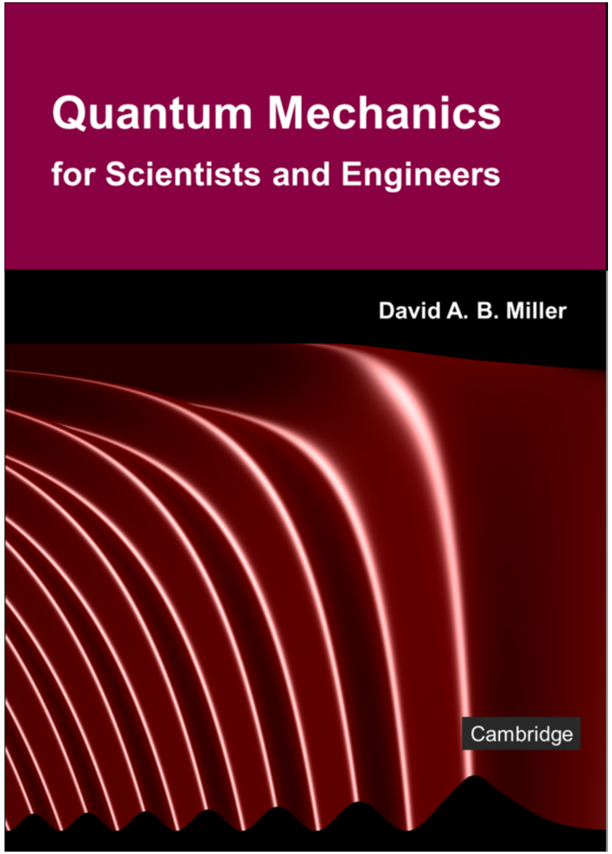


15.2 Quantum states of the electromagnetic field

Slides: Video 15.2.3 The coherent state

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.6 subsection "Coherent state"



Quantum states of the electromagnetic field

The coherent state

Quantum mechanics for scientists and engineers

David Miller

The coherent state

The state that corresponds most closely to the classical field in an electromagnetic mode

is the coherent state

introduced previously as an example with the harmonic oscillator

Using our current notation we can rewrite this as

$$\left|\Psi_{\lambda\bar{n}}\right\rangle = \sum_{n_\lambda=0}^{\infty} c_{\lambda\bar{n}n} \exp\left[-i\left(n_\lambda + \frac{1}{2}\right)\omega_\lambda t\right]\left|n_\lambda\right\rangle$$

where $c_{\lambda\bar{n}n} = \sqrt{\left[\bar{n}^{n_\lambda} \exp\left(-\bar{n}\right)\right] / n_\lambda!}$

The coherent state

In these expansion coefficients $c_{\lambda\bar{n}n} = \sqrt{\dfrac{\bar{n}^{n_\lambda} \exp(-\bar{n})}{n_\lambda !}}$

the quantity $\bar{n}$ will turn out to be the expected value of the number of photons in the mode

As before, note that

$$\left|c_{\lambda\bar{n}n}\right|^2 = \frac{\bar{n}^{n_\lambda} \exp(-\bar{n})}{n_\lambda !}$$

is the Poisson distribution

with mean $\bar{n}$ and standard deviation $\sqrt{\bar{n}}$

The coherent state

Note that, in the coherent state

the number of photons in the mode is not determined

The coefficients $\left|c_{\lambda\bar{n}n}\right|^2$

tell us the probability that

we will find n_λ photons in the mode

if we make a measurement

This number is now found to be distributed

according to a Poisson distribution

The coherent state

It is in fact the case that

- the statistics of the number of photons
    - in an oscillating “classical” electromagnetic field
        - are Poissonian

For example

- if we put a photodetector in a laser beam
    - we will measure a Poissonian distribution
        - of the arrival rates of the photons
            - an effect known as shot noise

Coherent state

Coherent state oscillations with

$$\left|\Psi_{\lambda\bar{n}}\right\rangle \equiv \Psi_{\lambda\bar{n}}\left(\xi_\lambda, t\right)$$

$$= \sum_{n_\lambda=0}^{\infty} c_{\lambda\bar{n}n} \exp\left[-i\left(n_\lambda + \frac{1}{2}\right)\omega_\lambda t\right]\left|n_\lambda\right\rangle$$

where

$$c_{\lambda\bar{n}n} = \sqrt{\frac{\bar{n}^{n_\lambda} \exp\left(-\bar{n}\right)}{n_\lambda !}}$$

and ξ_λ is the magnetic field amplitude

Coherent state

Coherent state oscillations with

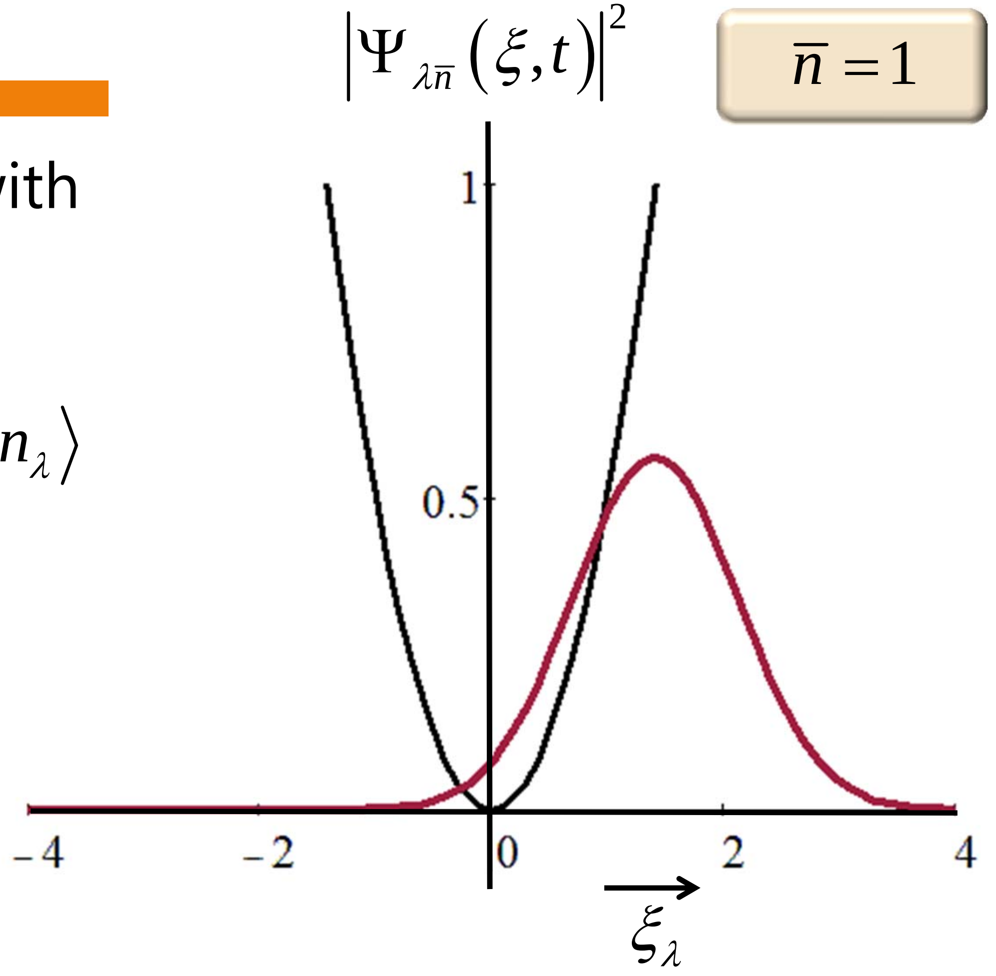
$$\left|\Psi_{\lambda\bar{n}}\right\rangle \equiv \Psi_{\lambda\bar{n}}\left(\xi_\lambda, t\right)$$

$$= \sum_{n_\lambda=0}^{\infty} c_{\lambda\bar{n}n} \exp\left[-i\left(n_\lambda + \frac{1}{2}\right)\omega_\lambda t\right]\left|n_\lambda\right\rangle$$

where

$$c_{\lambda\bar{n}n} = \sqrt{\frac{\bar{n}^{n_\lambda}\exp\left(-\bar{n}\right)}{n_\lambda!}}$$

and ξ_λ is the magnetic field amplitude



Coherent state

Coherent state oscillations with

$$\left|\Psi_{\lambda\bar{n}}\right\rangle \equiv \Psi_{\lambda\bar{n}}\left(\xi_\lambda, t\right)$$

$$= \sum_{n_\lambda=0}^{\infty} c_{\lambda\bar{n}n} \exp\left[-i\left(n_\lambda + \frac{1}{2}\right)\omega_\lambda t\right]\left|n_\lambda\right\rangle$$

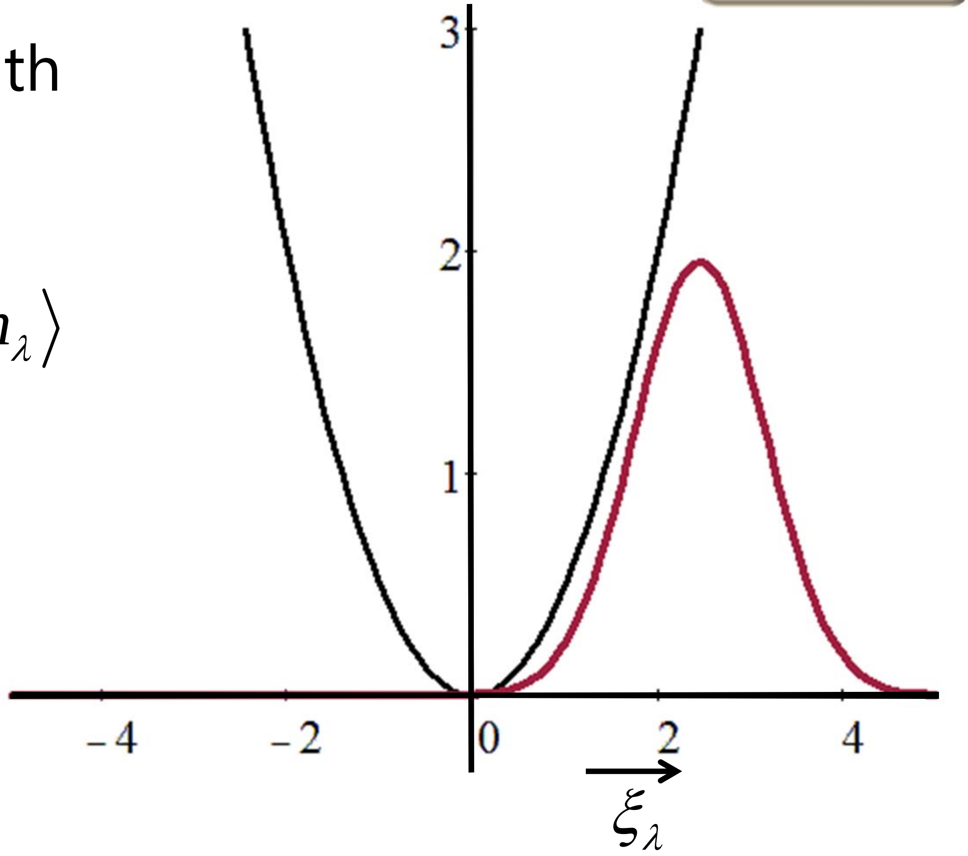
where

$$c_{\lambda\bar{n}n} = \sqrt{\frac{\bar{n}^{n_\lambda}\exp\left(-\bar{n}\right)}{n_\lambda!}}$$

and ξ_λ is the magnetic field amplitude

$$\left|\Psi_{\lambda\bar{n}}\left(\xi,t\right)\right|^2$$

$\bar{n} = 3$



Coherent state

Coherent state oscillations with

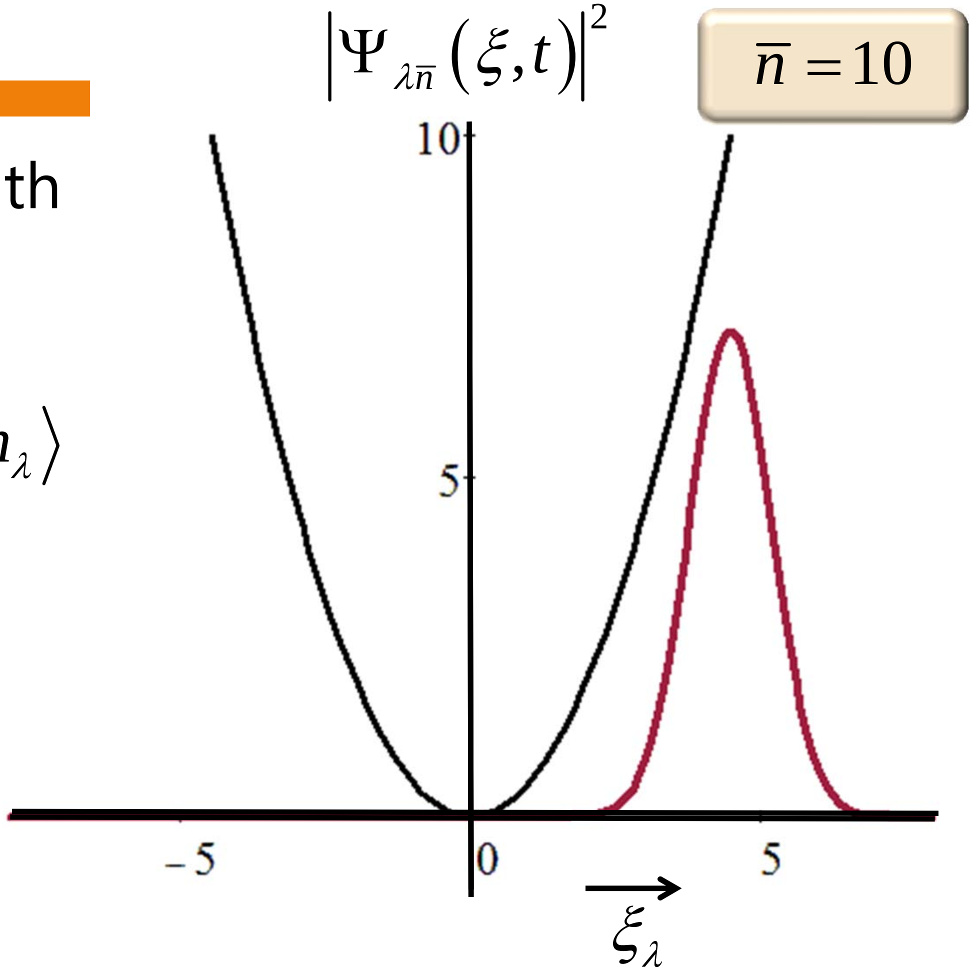
$$\left|\Psi_{\lambda\bar{n}}\right\rangle \equiv \Psi_{\lambda\bar{n}}\left(\xi_\lambda, t\right)$$

$$= \sum_{n_\lambda=0}^{\infty} c_{\lambda\bar{n}n} \exp\left[-i\left(n_\lambda + \frac{1}{2}\right)\omega_\lambda t\right]\left|n_\lambda\right\rangle$$

where

$$c_{\lambda\bar{n}n} = \sqrt{\frac{\bar{n}^{n_\lambda} \exp(-\bar{n})}{n_\lambda !}}$$

and ξ_λ is the magnetic field amplitude



Coherent state

Coherent state oscillations with

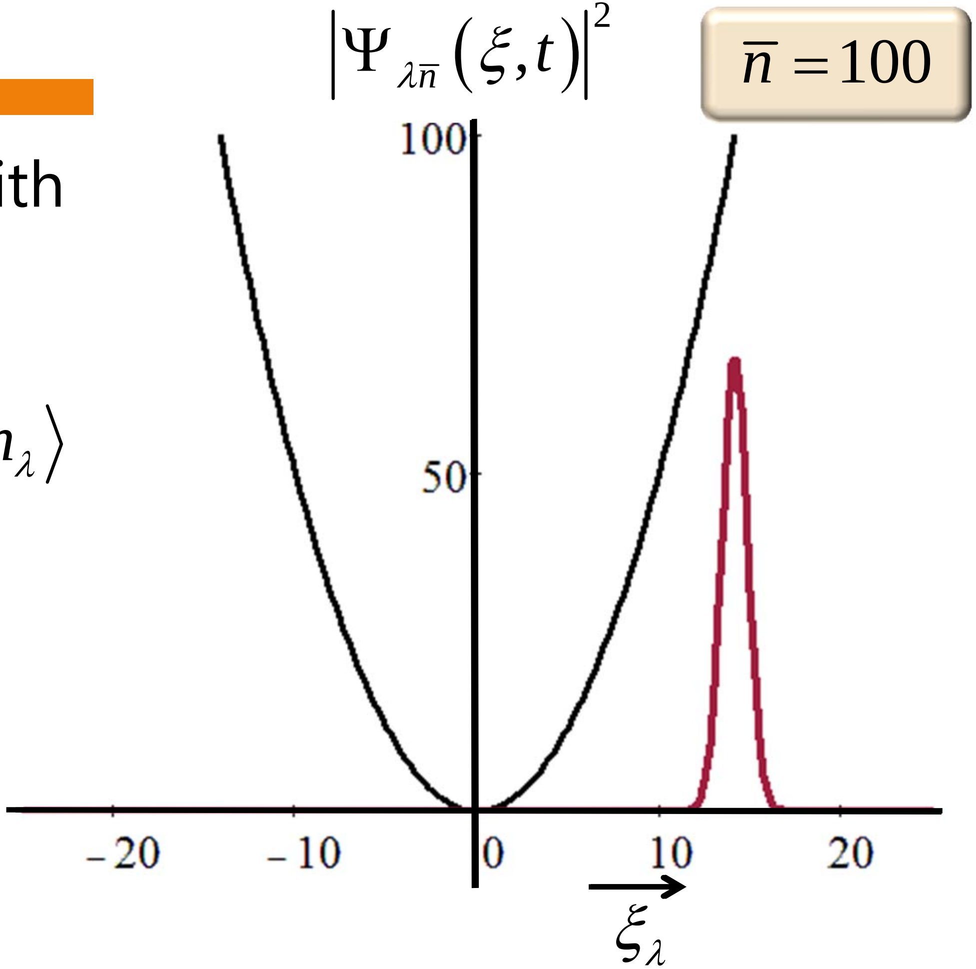
$$\left|\Psi_{\lambda\bar{n}}\right\rangle \equiv \Psi_{\lambda\bar{n}}\left(\xi_\lambda, t\right)$$

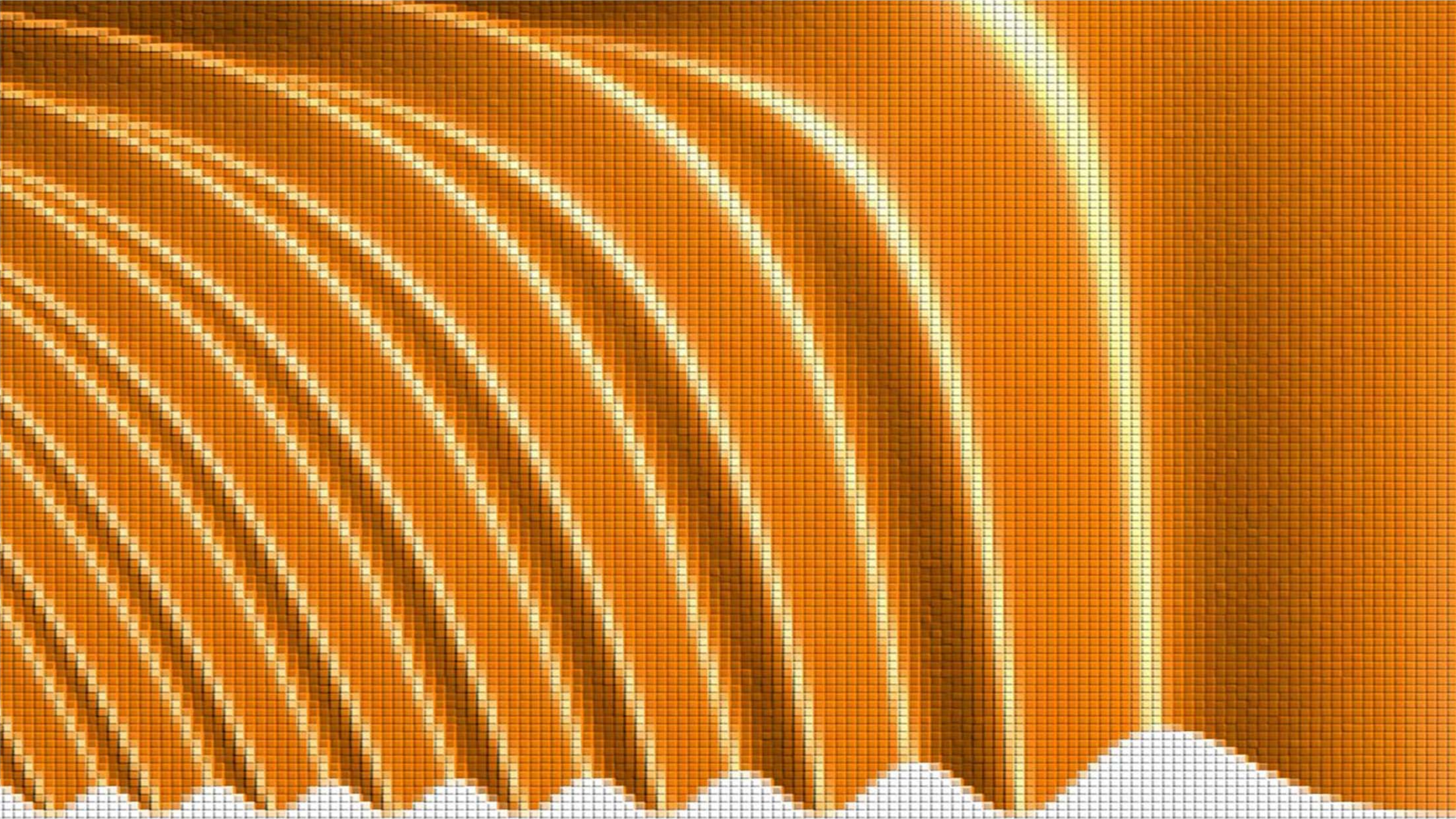
$$= \sum_{n_\lambda=0}^{\infty} c_{\lambda\bar{n}n} \exp\left[-i\left(n_\lambda + \frac{1}{2}\right)\omega_\lambda t\right]\left|n_\lambda\right\rangle$$

where

$$c_{\lambda\bar{n}n} = \sqrt{\frac{\bar{n}^{n_\lambda} \exp\left(-\bar{n}\right)}{n_\lambda !}}$$

and ξ_λ is the magnetic field amplitude



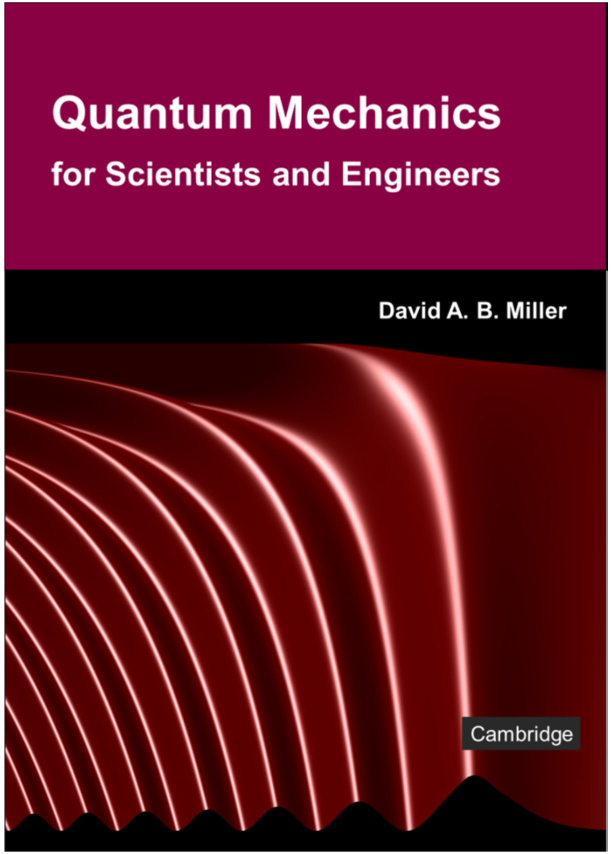


15.2 Quantum states of the electromagnetic field

Slides: Video 15.2.5 Sets of modes

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.7 to start of “Multimode photon states”



Quantum states of the electromagnetic field

Sets of modes

Quantum mechanics for scientists and engineers

David Miller

Sets of classical modes

We postulate a set of classical modes

each of which has the following form

$$\mathbf{E}_\lambda\left(\mathbf{r},t\right)=-p_\lambda(t)D_\lambda\mathbf{u}_\lambda\left(\mathbf{r}\right) \qquad \mathbf{B}_\lambda\left(\mathbf{r},t\right)=q_\lambda(t)\frac{D_\lambda}{c}\mathbf{v}_\lambda\left(\mathbf{r}\right)$$

Here $\mathbf{E}_\lambda$, $\mathbf{B}_\lambda$, $\mathbf{u}_\lambda$, and $\mathbf{v}_\lambda$ are all in general vectors

and D_λ is a constant

The forms we used for our plane wave example

$$E_z = p\left(t\right)D\sin kx \qquad B_y = q\left(t\right)\frac{D}{c}\cos kx$$

correspond to these with

$$\mathbf{u}_\lambda\left(\mathbf{r}\right)=-\hat{\mathbf{z}}\sin\left(kx\right) \qquad \mathbf{v}_\lambda\left(\mathbf{r}\right)=\hat{\mathbf{y}}\cos\left(kx\right)$$

Sets of classical modes

$$\mathbf{E}_\lambda(\mathbf{r},t) = -p_\lambda(t) D_\lambda \mathbf{u}_\lambda(\mathbf{r}) \quad \text{and} \quad \mathbf{B}_\lambda(\mathbf{r},t) = q_\lambda(t)\frac{D_\lambda}{c}\mathbf{v}_\lambda(\mathbf{r})$$

satisfy Maxwell's equations

and the wave equation in free space if we require

$$\nabla \times \mathbf{u}_\lambda(\mathbf{r}) = \frac{\omega_\lambda}{c}\mathbf{v}_\lambda(\mathbf{r}) \qquad \nabla \times \mathbf{v}_\lambda(\mathbf{r}) = \frac{\omega_\lambda}{c}\mathbf{u}_\lambda(\mathbf{r})$$

$$\frac{dq_\lambda}{dt} = \omega_\lambda p_\lambda \qquad \frac{dp_\lambda}{dt} = -\omega_\lambda q_\lambda$$

We presume the classical electromagnetic problem

with its boundary conditions

has been solved to give these modes

Sets of classical modes

We will also presume that

the spatial functions $\mathbf{u}_\lambda(\mathbf{r})$ and $\mathbf{v}_\lambda(\mathbf{r})$

are normalized over the entire volume

and they are all orthogonal

So, for two (possibly different) modes

with indices $\lambda 1$ and $\lambda 2$ respectively

$$\int \mathbf{u}_{\lambda 1}(\mathbf{r}) \cdot \mathbf{u}_{\lambda 2}(\mathbf{r}) d^3\mathbf{r} = \delta_{\lambda 1, \lambda 2}$$

and

$$\int \mathbf{v}_{\lambda 1}(\mathbf{r}) \cdot \mathbf{v}_{\lambda 2}(\mathbf{r}) d^3\mathbf{r} = \delta_{\lambda 1, \lambda 2}$$

Classical superpositions and energy

Consider a classical superposition of such modes

$$\mathbf{E}(\mathbf{r},t)=\sum_{\lambda}-p_{\lambda}(t)D_{\lambda}\mathbf{u}_{\lambda}(\mathbf{r}) \qquad \mathbf{B}(\mathbf{r},t)=\sum_{\lambda}q_{\lambda}(t)\frac{D_{\lambda}}{c}\mathbf{v}_{\lambda}(\mathbf{r})$$

The total energy of such a field is

$$H=\int\frac{1}{2}\left(\varepsilon_{o}\mathbf{E}^{2}+\frac{1}{\mu_{o}}\mathbf{B}^{2}\right)d^{3}\mathbf{r}$$

$$=\frac{1}{2}\varepsilon_{o}\sum_{\lambda1,\lambda2}D_{\lambda1}D_{\lambda2}\int\left[p_{\lambda1}p_{\lambda2}\mathbf{u}_{\lambda1}(\mathbf{r})\cdot\mathbf{u}_{\lambda2}(\mathbf{r})+q_{\lambda1}q_{\lambda2}\mathbf{v}_{\lambda1}(\mathbf{r})\cdot\mathbf{v}_{\lambda2}(\mathbf{r})\right]d^{3}\mathbf{r}$$

where we have also used $1/c^{2}=\varepsilon_{o}\mu_{o}$

Classical superpositions and energy

Using the orthonormality of the $\mathbf{u}_\lambda$ and of the $\mathbf{v}_\lambda$ in

$$H = \frac{1}{2}\varepsilon_o \sum_{\lambda 1,\lambda 2} D_{\lambda 1} D_{\lambda 2} \int \left[p_{\lambda 1} p_{\lambda 2} \mathbf{u}_{\lambda 1}(\mathbf{r}) \cdot \mathbf{u}_{\lambda 2}(\mathbf{r}) + q_{\lambda 1} q_{\lambda 2} \mathbf{v}_{\lambda 1}(\mathbf{r}) \cdot \mathbf{v}_{\lambda 2}(\mathbf{r}) \right] d^3\mathbf{r}$$

eliminates the “cross terms” with different indices, so

$$H = \frac{1}{2}\varepsilon_o \sum_{\lambda} D_\lambda^2 \left(p_\lambda^2 + q_\lambda^2 \right)$$

so we can write a sum of separate Hamiltonians

$$H = \sum_\lambda H_\lambda$$

where $H_\lambda = \frac{1}{2}\varepsilon_o D_\lambda^2 \left(p_\lambda^2 + q_\lambda^2 \right)$

Classical mode Hamiltonians

In each $H_\lambda = \frac{1}{2}\varepsilon_o D_\lambda^2\left(p_\lambda^2 + q_\lambda^2\right)$

if we now choose $D_\lambda = \sqrt{\frac{\omega_\lambda}{\varepsilon_o}}$

then we have $H_\lambda = \frac{\omega_\lambda}{2}\left(p_\lambda^2 + q_\lambda^2\right)$

and the H_λ, p_λ, and q_λ now obey

Hamilton's equations $\frac{dp_\lambda}{dt} = -\frac{\partial H_\lambda}{\partial q_\lambda}$ $\frac{dq_\lambda}{dt} = \frac{\partial H_\lambda}{\partial p_\lambda}$

as we could check

Quantization of individual modes

We can proceed for each mode, postulating

a "momentum" operator for each mode

$$\hat{p}_\lambda = -i\hbar \frac{d}{dq_\lambda}$$

hence changing the classical Hamiltonian

$$H_\lambda = \frac{\omega_\lambda}{2}\left(p_\lambda^2 + q_\lambda^2\right)$$

to the proposed the quantum mechanical one

$$\hat{H}_\lambda = \frac{\omega_\lambda}{2}\left[-\hbar^2 \frac{d^2}{dq_\lambda{}^2} + q_\lambda^2\right]$$

Quantization of individual modes

We next rewrite this Hamiltonian as

$$\hat{H}_\lambda = \frac{\hbar\omega_\lambda}{2}\left[-\frac{d^2}{d\xi_\lambda{}^2} + \xi_\lambda{}^2\right] = \hbar\omega_\lambda\left(\hat{a}_\lambda^\dagger\hat{a}_\lambda + \frac{1}{2}\right)$$

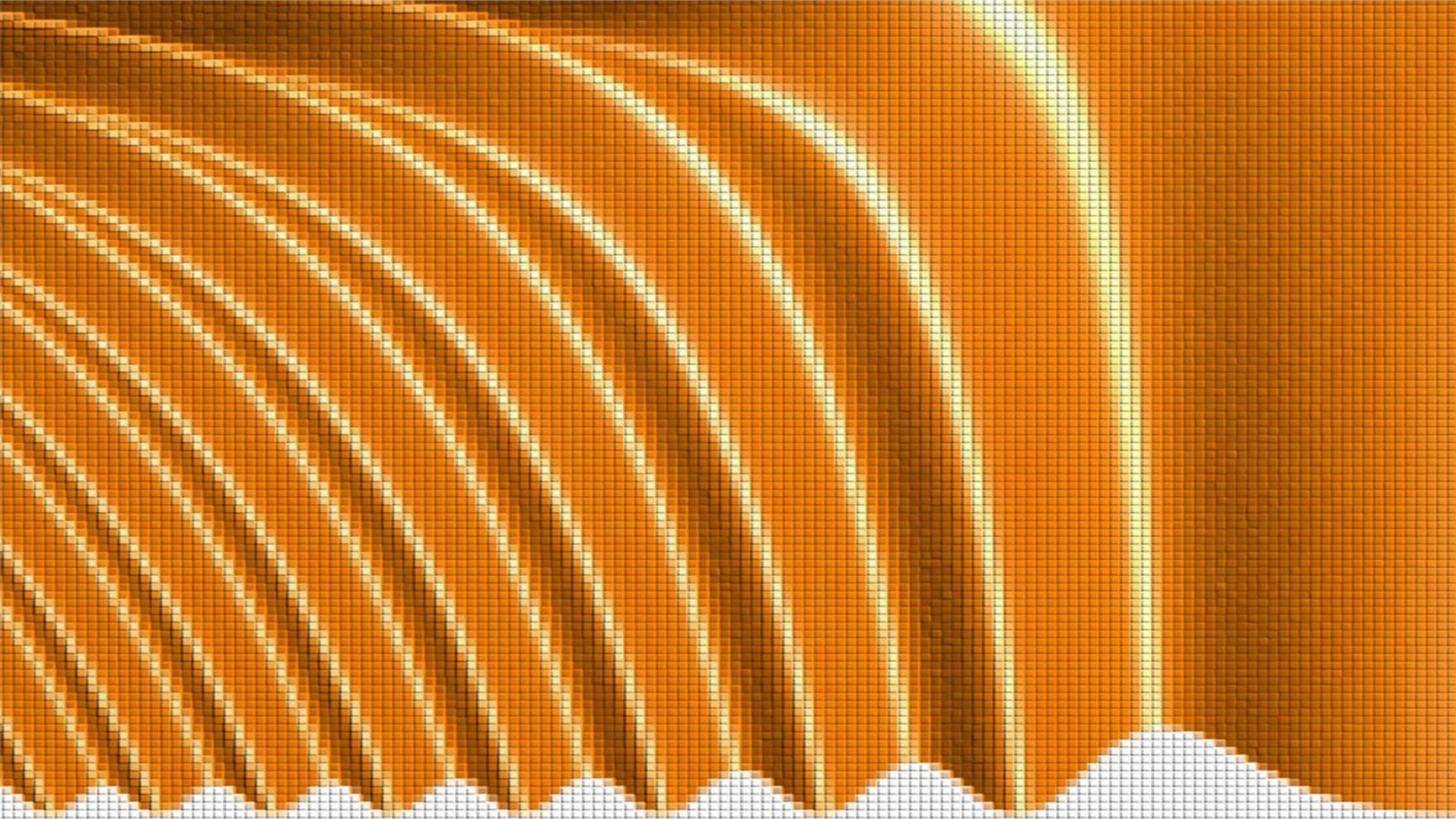
defining dimensionless units $\xi_\lambda = q_\lambda / \sqrt{\hbar}$

and creation and annihilation operators

$$\hat{a}_\lambda^\dagger \equiv \frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi_\lambda} + \xi_\lambda\right) \qquad \hat{a}_\lambda \equiv \frac{1}{\sqrt{2}}\left(\frac{d}{d\xi_\lambda} + \xi_\lambda\right)$$

so the total Hamiltonian for the set of modes is

$$\boxed{\hat{H} = \sum_\lambda \hbar\omega_\lambda\left(\hat{a}_\lambda^\dagger\hat{a}_\lambda + \frac{1}{2}\right)}$$

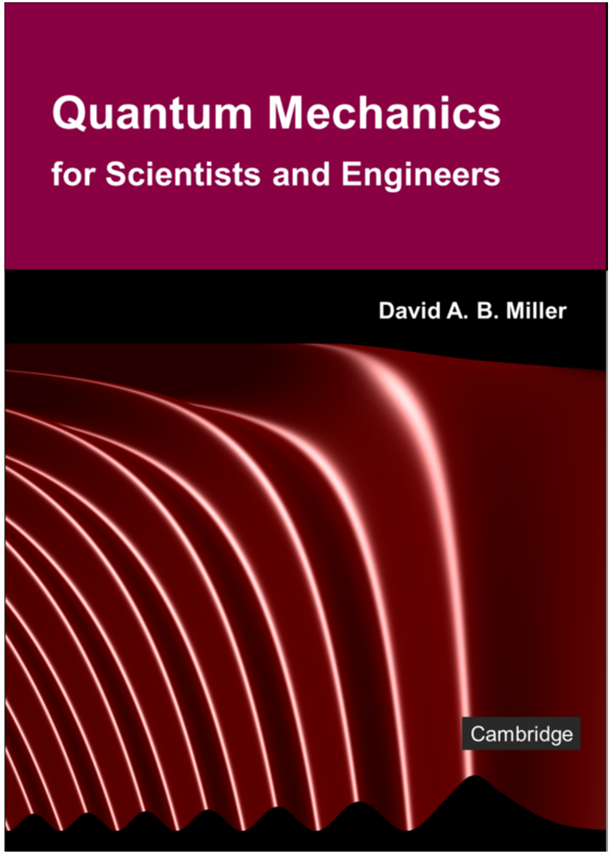


15.2 Quantum states of the electromagnetic field

Slides: Video 15.2.7 Multimode photon states

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.7 subsection "Multimode photon states"



Quantum states of the electromagnetic field

Multimode photon states

Quantum mechanics for scientists and engineers

David Miller

Occupation number representation

For example, the state with

one photon in mode k, three in mode m and none in any other mode

which would be an example of a "multimode" state

could be written as

$$\left|0_a, \ldots, 0_j, 1_k, 0_l, 3_m, 0_n, \ldots\right\rangle$$

where we have labeled the modes progressively with the lower case letters

Single mode operators with multimode states

Just as before

the annihilation operators will have the property

now specific to given mode

$$\hat{a}_\lambda \left| \ldots, n_\lambda, \ldots \right\rangle = \sqrt{n_\lambda} \left| \ldots, (n_\lambda - 1)_\lambda, \ldots \right\rangle$$

with $\hat{a}_\lambda \left| \ldots, 0_\lambda, \ldots \right\rangle = 0$

Similarly $\hat{a}_\lambda^\dagger \left| \ldots, n_\lambda, \ldots \right\rangle = \sqrt{n_\lambda + 1} \left| \ldots, (n_\lambda + 1)_\lambda, \ldots \right\rangle$

and the number operator for a given mode

will still be $\hat{N}_\lambda \equiv \hat{a}_\lambda^\dagger \hat{a}_\lambda$

so $\hat{N} \left| \ldots, n_\lambda, \ldots \right\rangle = n_\lambda \left| \ldots, n_\lambda, \ldots \right\rangle$

Writing multimode states using operators

We can create a multimode state by progressively operating with the appropriate creation operators

starting with the "zero" state or "empty" state

often written simply as $|0\rangle$

For our example state, we could write

$$\left|0_a,\ldots,0_j,1_k,0_l,3_m,0_n,\ldots\right\rangle = \frac{1}{\sqrt{1!3!}}\hat{a}_k^\dagger \hat{a}_m^\dagger \hat{a}_m^\dagger \hat{a}_m^\dagger \left|0\right\rangle$$

where the factor $1/\sqrt{1!3!}$ compensates for the factors introduced by the creation operators in

$$\hat{a}_\lambda^\dagger \left|\ldots,n_\lambda,\ldots\right\rangle = \sqrt{n_\lambda+1}\left|\ldots,\left(n_\lambda+1\right)_\lambda,\ldots\right\rangle$$

keeping the state normalized

Writing multimode states using operators

In general, we can write a state with

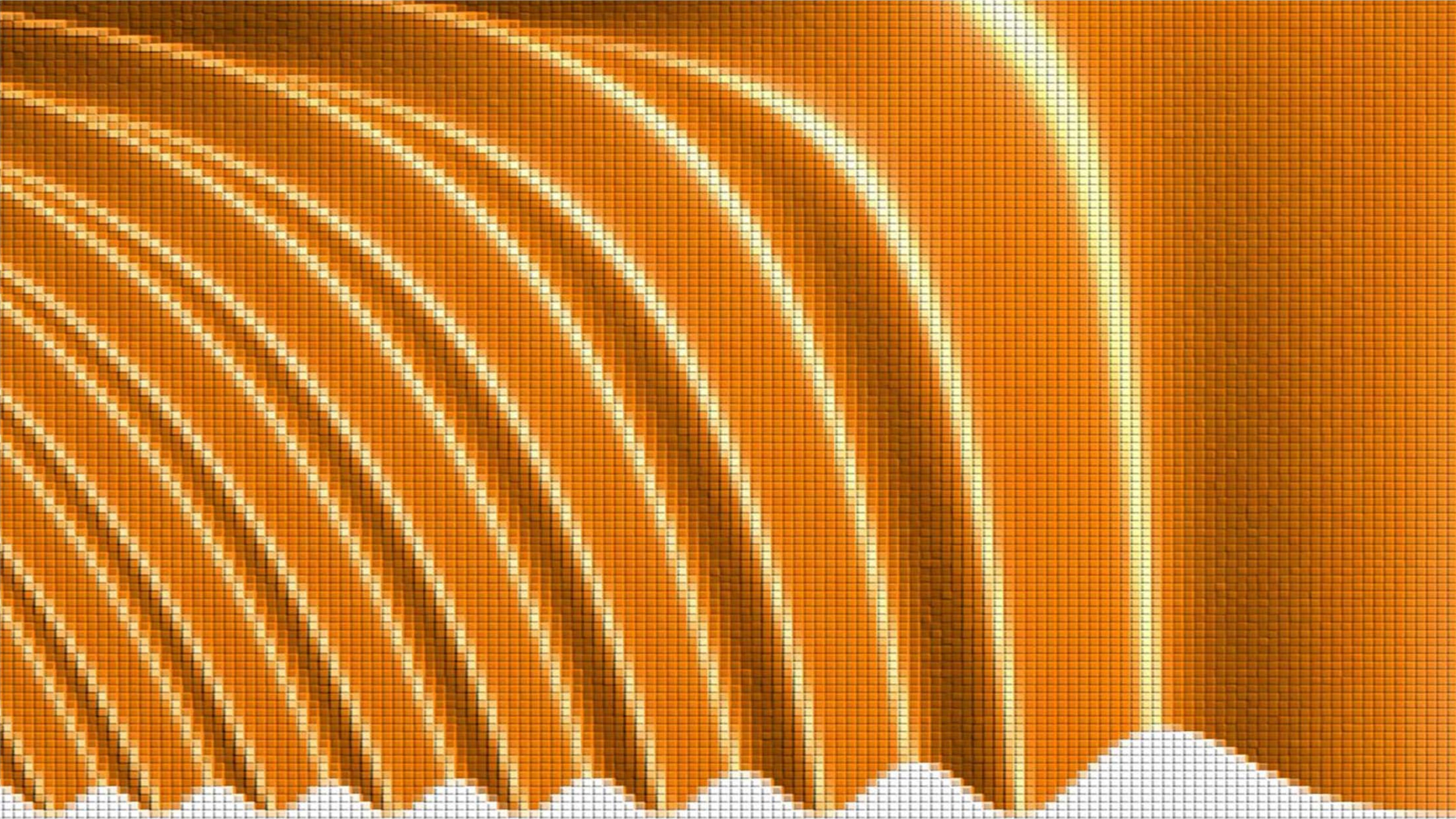
n_1 particles in mode 1

n_2 particles in mode 2

and so on

as

$$\left|n_1, n_2, \ldots, n_\lambda, \ldots\right\rangle = \frac{1}{\sqrt{n_1!n_2!\ldots n_\lambda!\ldots}}\left(\hat{a}_1^\dagger\right)^{n_1}\left(\hat{a}_2^\dagger\right)^{n_2}\ldots\left(\hat{a}_\lambda^\dagger\right)^{n_\lambda}\ldots\left|0\right\rangle$$

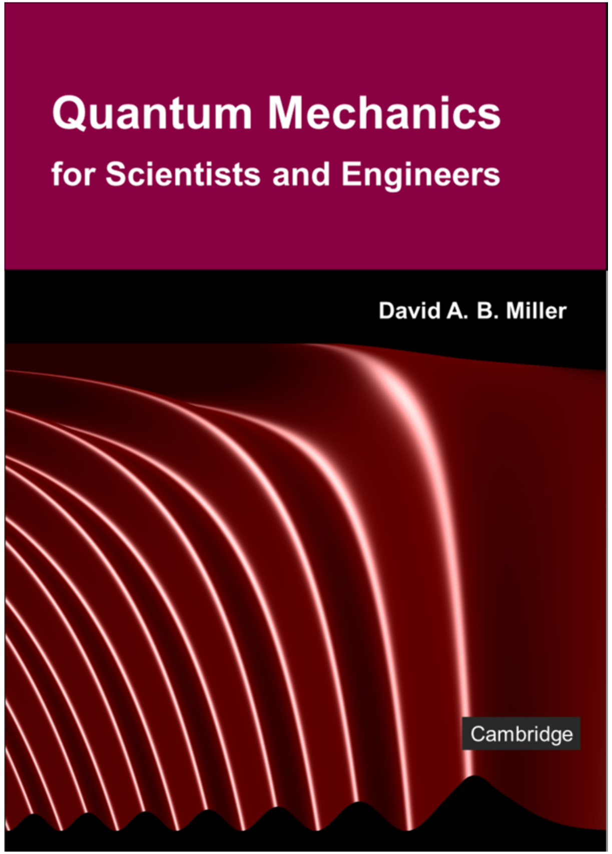


15.2 Quantum states of the electromagnetic field

Slides: Video 15.2.9 Multimode operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 15.7 subsection “Commutation relations …” to end



Quantum states of the electromagnetic field

Multimode operators

Quantum mechanics for scientists and engineers

David Miller

Commutation relations for boson operators

Formally, then, for creation operators

operating on any state

we must have

$$\hat{a}_j^\dagger \hat{a}_k^\dagger = \hat{a}_k^\dagger \hat{a}_j^\dagger$$

or, in the form of a commutation relation

$$\hat{a}_j^\dagger \hat{a}_k^\dagger - \hat{a}_k^\dagger \hat{a}_j^\dagger = 0$$

Similarly, for annihilation operators

it does not matter in what order we destroy particles

and so we similarly have

$$\hat{a}_j \hat{a}_k - \hat{a}_k \hat{a}_j = 0$$

Commutation relations for boson operators

For mixtures of annihilation and creation operators

if we annihilate a boson in one mode

and create one in another

it does not matter in what order we do that either

Only if we are creating and annihilating in the same mode does it matter what order we do this

with a commutation relation we have previously deduced (i.e., $\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = 1$)

Hence in general we can write

$$\hat{a}_j\hat{a}_k^\dagger - \hat{a}_k^\dagger\hat{a}_j = \delta_{jk}$$

Multimode field operators

It is now straightforward to construct the full multimode electric and magnetic field operators

We start from the classical definition of the multimode electric field

$$\mathbf{E}(\mathbf{r},t)=\sum_{\lambda}-p_{\lambda}(t)D_{\lambda}\mathbf{u}_{\lambda}(\mathbf{r})$$

as an expansion in classical field modes

We use the relation $D_{\lambda}=\sqrt{\omega_{\lambda}/\varepsilon_{o}}$ we deduced to get Hamilton's equations, and

we substitute the operator $\hat{p}_{\lambda}$

for the quantity p in each mode

Multimode field operators

We therefore use our previously deduced operator

$$\hat{p}_\lambda = i\sqrt{\frac{\hbar}{2}}\left(\hat{a}_\lambda^\dagger - \hat{a}_\lambda\right)$$

in $\mathbf{E}(\mathbf{r},t) = \sum_\lambda -p_\lambda(t) D_\lambda \mathbf{u}_\lambda(\mathbf{r})$

to obtain the (multimode) electric field operator

$$\hat{\mathbf{E}}(\mathbf{r},t) = i\sum_\lambda \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right)\sqrt{\frac{\hbar\omega_\lambda}{2\varepsilon_o}}\mathbf{u}_\lambda(\mathbf{r})$$

Multimode field operators

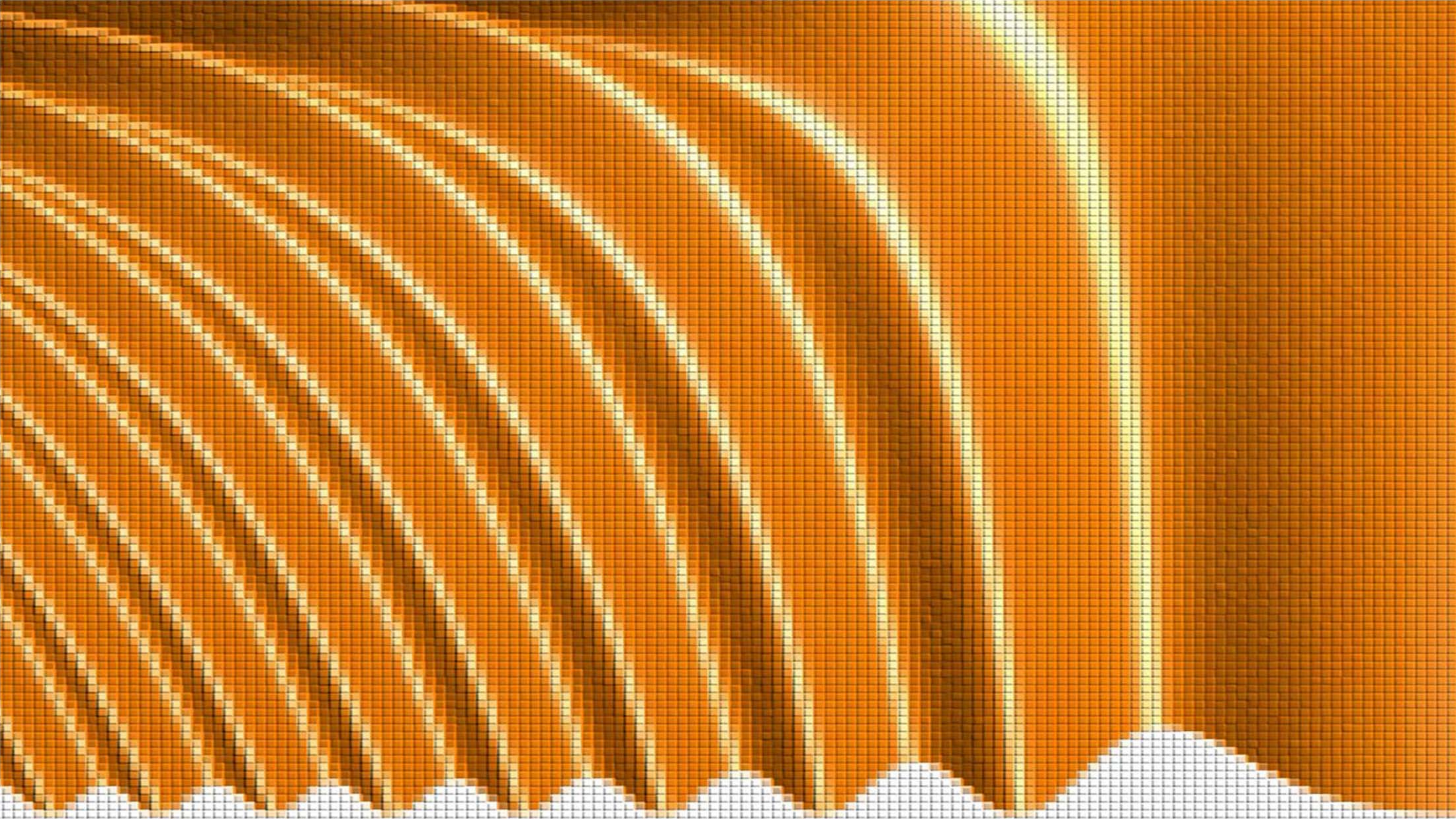
By a similar argument, starting from the classical expression for a multimode magnetic field

$$\mathbf{B}(\mathbf{r},t)=\sum_{\lambda} q_{\lambda}(t)\frac{D_{\lambda}}{c}\mathbf{v}_{\lambda}(\mathbf{r})$$

substituting the operator $\hat{q}_{\lambda} \equiv \sqrt{\frac{\hbar}{2}}\left(\hat{a}_{\lambda}+\hat{a}_{\lambda}^{\dagger}\right)$ for q_{λ}

we obtain

$$\boxed{\hat{\mathbf{B}}(\mathbf{r},t)=\sum_{\lambda}\left(\hat{a}_{\lambda}+\hat{a}_{\lambda}^{\dagger}\right)\sqrt{\frac{\hbar\omega_{\lambda}\mu_{o}}{2}}\mathbf{v}_{\lambda}(\mathbf{r})}$$

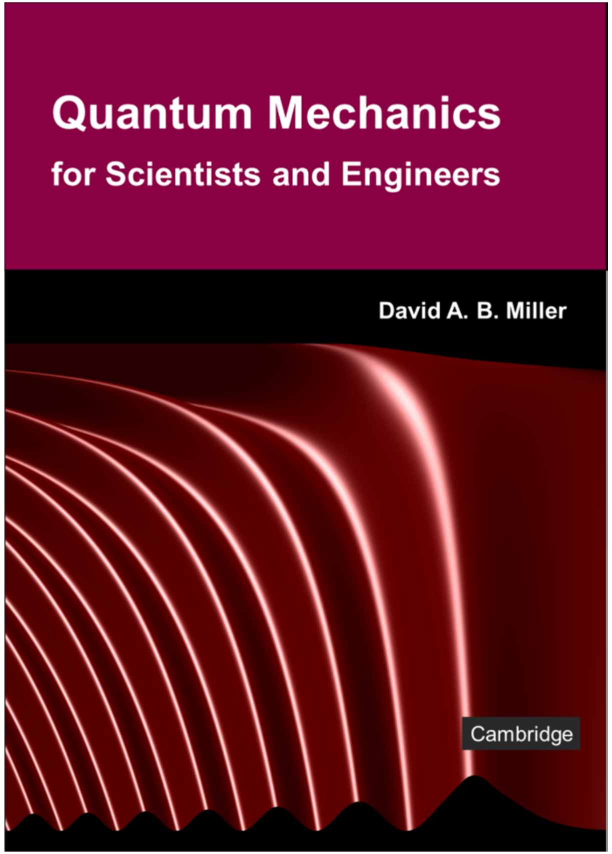


15.3 Fermion annihilation and creation operators

Slides: Video 15.3.1 Describing fermion states

Text reference: Quantum Mechanics for Scientists and Engineers

Section Chapter 16 introduction and Section 16.1 up to beginning of “Fermion creation operators”



Fermion annihilation and creation operators

Quantum mechanics for scientists and engineers

David Miller

Fermion annihilation and creation operators

Describing fermion states

Quantum mechanics for scientists and engineers

David Miller

Description and ordering of multiple fermion states

We can write a basis state for multiple identical fermions as

$$\left|\psi_{N;a,b,\ldots n}\right\rangle = \frac{1}{\sqrt{N!}} \sum_{\hat{P}=1}^{N!} \pm \hat{P} \left| |1,a\rangle |2,b\rangle |3,c\rangle \ldots |N,n\rangle \right\rangle$$

$$\equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} |1,a\rangle & |2,a\rangle & \cdots & |N,a\rangle \\ |1,b\rangle & |2,b\rangle & \cdots & |N,b\rangle \\ \vdots & \vdots & \ddots & \vdots \\ |1,n\rangle & |2,n\rangle & \cdots & |N,n\rangle \end{vmatrix}$$

Here, there are N identical fermions, and they occupy single-particle basis states a, b, … n

Single-particle fermion states

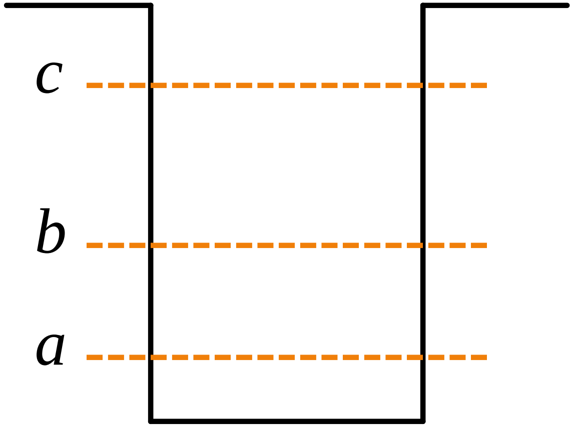
Single-particle basis states are individual states a fermion can occupy

and here each has a lower case letter associated with it

For example

each possible electron state in a potential well or atom

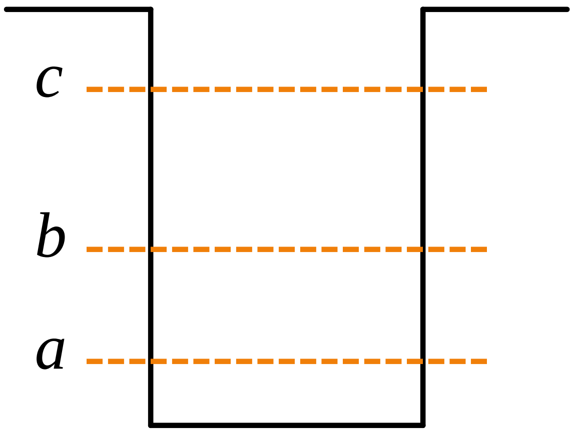
corresponds to a different single-particle basis state here



Multiple fermion basis states

Though $\left|\psi_{N;a,b,\ldots n}\right\rangle =$

$$\frac{1}{\sqrt{N!}}\sum_{\hat{P}=1}^{N!}\pm\hat{P}\left|\left|1,a\right\rangle\left|2,b\right\rangle\left|3,c\right\rangle\ldots\left|N,n\right\rangle\right\rangle$$



might seem to imply that each of the possible states is occupied

that is not in general the case

Very few of the possible single-particle states are likely occupied

in any given multiple fermion basis state

Multiple fermion basis states

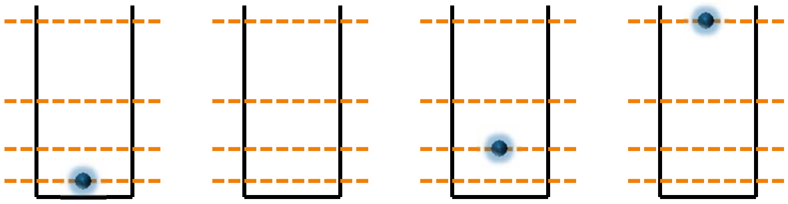
We might have three electrons in with four potential wells

and be considering a (multiple particle) basis state in which there is

1 electron in the ground state of well 1

1 in the 2nd state of well 3

1 in the 4th state of well 4



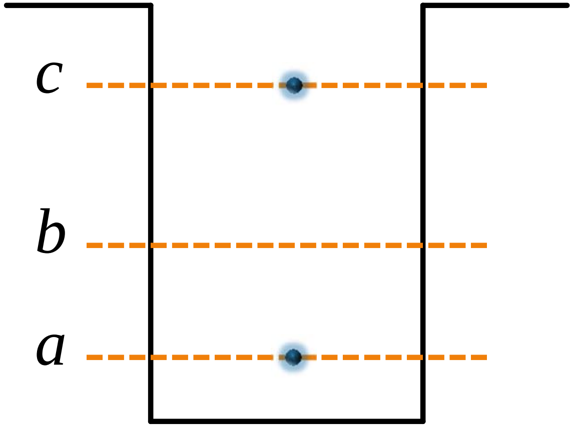
Multiple fermion basis states

A basis state that might have two electrons in one well

- e.g., one on the lowest state and
  - one on the third state
    - though this is not necessarily an eigenstate of the Hamiltonian

The first electron would repel the second electron

- so the second electron would not see a simple square potential



Multiple fermion basis states

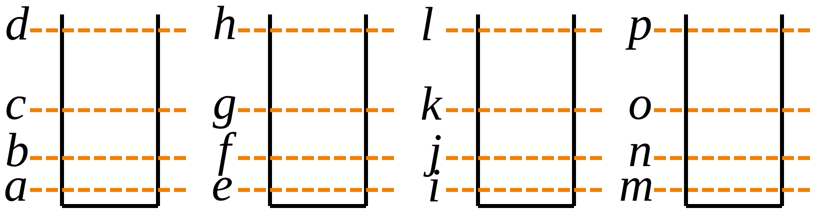
For manipulations in the fermion case, we must

- define one standard order of labeling of the single-particle basis states
    - in the determinants for the multiple fermion basis functions

Multiple fermion basis states

For example, if we had a system with four potential wells

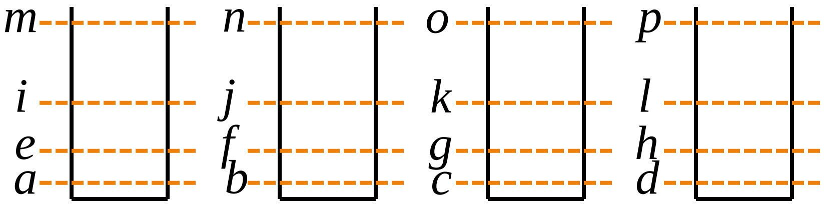
- we might label sequentially all of the states in well 1
  - then next all of the states in well 2
    - and so on



Multiple fermion basis states

We could choose some other labeling sequence

- labeling all of the first states in wells 1 through 4
  - then all of the second states in wells 1 through 4
    - and so on

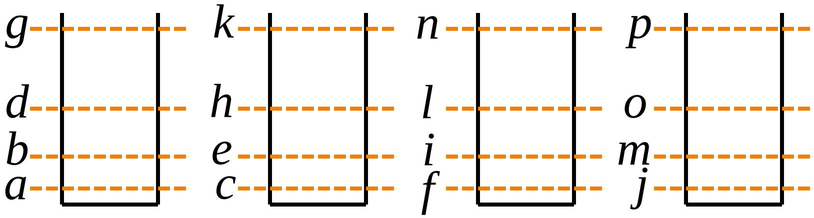


Multiple fermion basis states

We could choose some other labeling sequence

- labeling all of the first states in wells 1 through 4
  - then all of the second states in wells 1 through 4
    - and so on

or we could even choose some more complicated labeling sequence



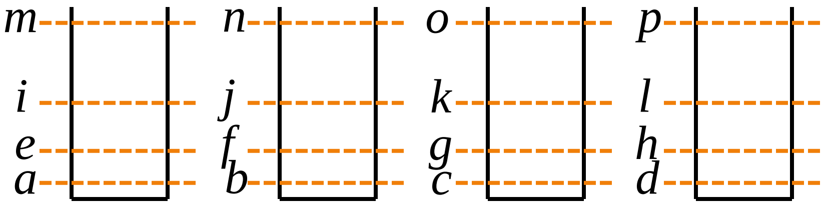
Multiple fermion basis states

It does not matter what sequence we choose

but we have to have one standard labeling sequence

Here, we label the single-particle basis states using the lower case letters

and those will be in alphabetical sequence in our standard order



Occupation number representation

We might, for example, have a basis state corresponding to three identical fermions

one in state b, one in state k, and one in state m

In standard order, we would write that state as

$$\left|\psi_{3;b,k,m}\right\rangle = \frac{1}{\sqrt{3!}}\begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} \equiv \left|0_a,1_b,0_c,\ldots,1_k,0_l,\ldots,1_m,0_n\ldots\right\rangle$$

Here we have also introduced another notation

the occupation number notation

similar to the boson occupation number notation

Occupation number representation

In this occupation number representation, as in the state

$$\left|\psi_{3;\,b,k,m}\right\rangle = \frac{1}{\sqrt{3!}}\begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} \equiv \left|0_a,1_b,0_c,\ldots,1_k,0_l,\ldots,1_m,0_n\ldots\right\rangle$$

0_a in the ket means that the single-particle fermion state (or fermion mode) a is empty, and

1_b means state b is occupied

Because this is a fermion state

the determinant combination of the different fermions to the occupied states is understood

States not in standard order

We could also write a state that was not in standard order for the rows

e.g., $$\left|\psi_{3;\,k,b,m}\right\rangle = \frac{1}{\sqrt{3!}}\begin{vmatrix} |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

To get that state into standard order for the rows

we would have to swap the first and second rows

If we swap two adjacent rows in a determinant

we have to multiply the determinant by -1

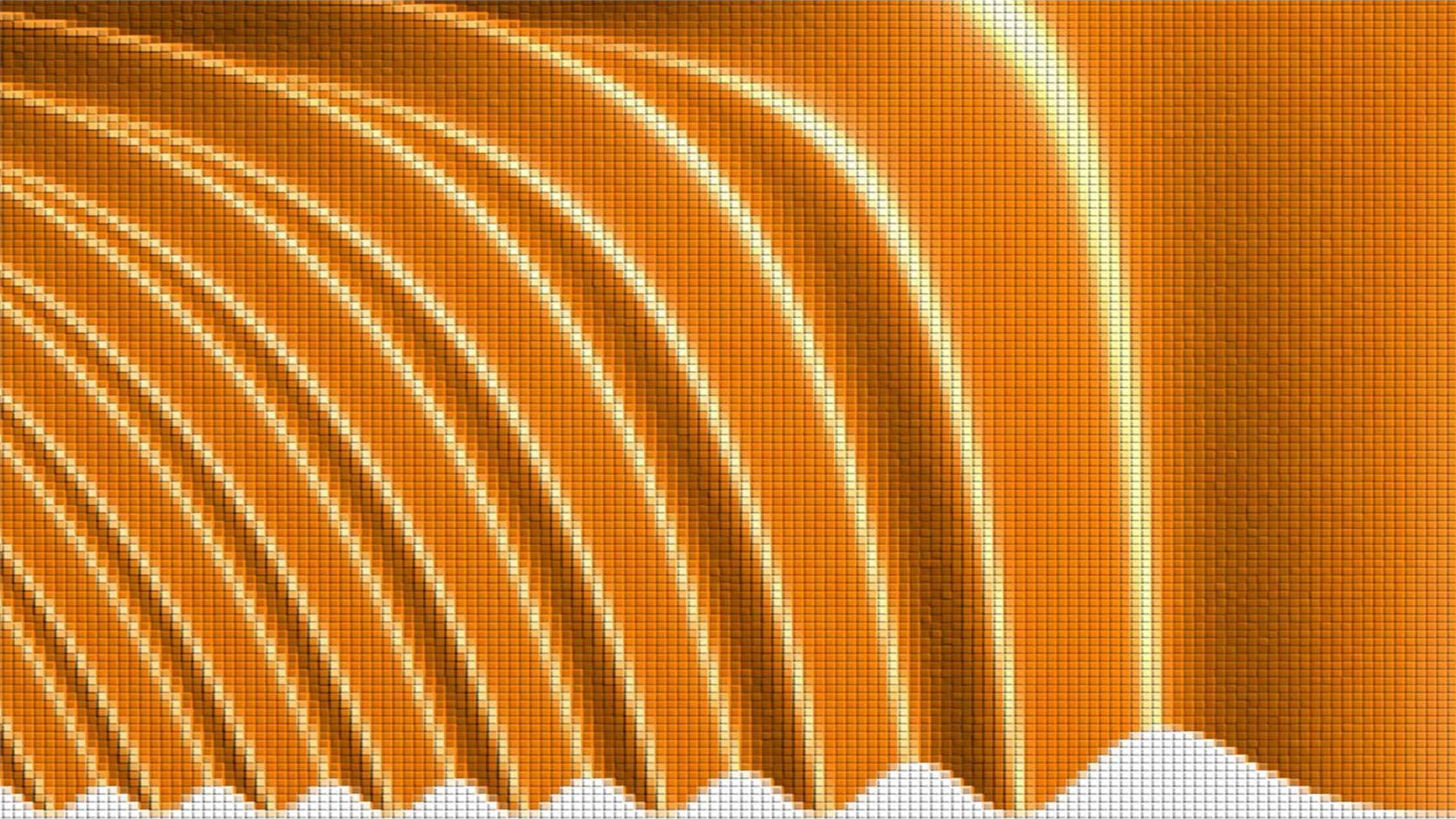
States not in standard order

So, swapping the top two rows, we have

$$\left|\psi_{3;\,k,b,m}\right\rangle = -\frac{1}{\sqrt{3!}}\begin{vmatrix} \left|1,k\right\rangle & \left|2,k\right\rangle & \left|3,k\right\rangle \\ \left|1,b\right\rangle & \left|2,b\right\rangle & \left|3,b\right\rangle \\ \left|1,m\right\rangle & \left|2,m\right\rangle & \left|3,m\right\rangle \end{vmatrix}$$

$$= -\left|0_a, 1_b, 0_c, \ldots, 1_k, 0_{l,} \ldots, 1_m, 0_n \ldots\right\rangle$$

$$= -\left|\psi_{3;\,b,k,m}\right\rangle$$

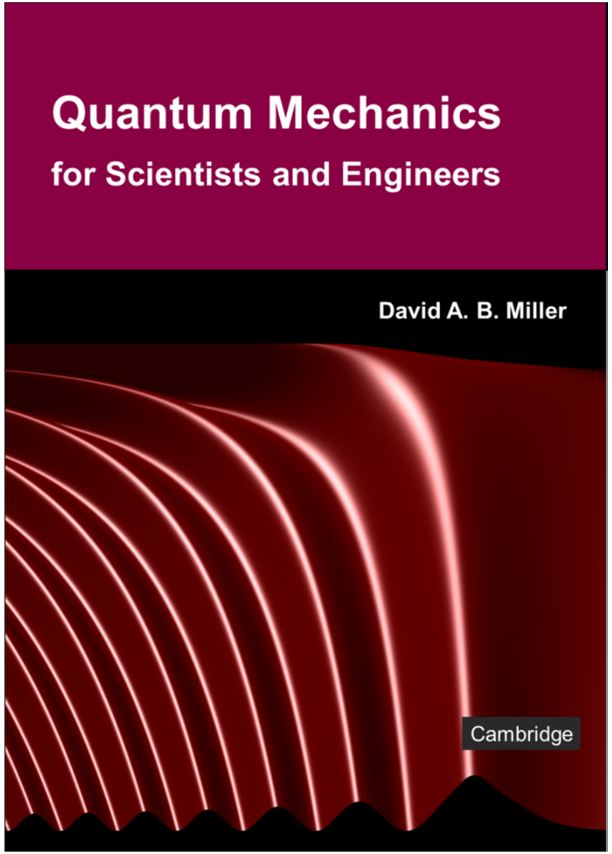


15.3 Fermion annihilation and creation operators

Slides: Video 15.3.3 Creation operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 16.1 subsection “Fermion creation operators”



Fermion annihilation and creation operators

Creation operators

Quantum mechanics for scientists and engineers

David Miller

Fermion creation operators

Now we postulate

a fermion creation operator

for fermion “mode” or single-particle basis state k

and write it as $\hat{b}_k^\dagger$

It must take any state in which single-particle basis state k is empty

and turn it into one

in which this state k is occupied

Constructing the creation operator

Suppose we start with the state where

single-particle states b and m are occupied

but state k and all other states are not

In the permutation notation, we therefore propose that

$\hat{b}_k^\dagger$ has the following effect on that state

$$\hat{b}_k^\dagger \frac{1}{\sqrt{2!}} \sum_{\hat{P}=1}^{2!} \pm \hat{P} \big| |1,b\rangle |2,m\rangle \big\rangle = \frac{1}{\sqrt{3!}} \sum_{\hat{P}=1}^{3!} \pm \hat{P} \big| |1,b\rangle |2,m\rangle |3,k\rangle \big\rangle$$

So, $\hat{b}_k^\dagger$ adds a third particle into the system

and we propose that it adds it to the end of the list

Constructing the creation operator

Adding to the end of the list is equivalent to

adding a row to the bottom of the determinant (and a column to the right)

i.e., now dropping the normalization factors for convenience but keeping the sign behavior

$$\hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \end{vmatrix}$$

Note the sequence in the permutation notation is the same as the sequence down this leading diagonal

Constructing the creation operator

For this case

$$\hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \end{vmatrix} = -\begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

the determinant is not written in standard order

To get this particular determinant into standard order

we need to swap the bottom two rows

and in performing this one swap

we must therefore multiply the determinant by -1

Constructing the creation operator

Suppose now that we add another particle

this time in state j

using the operator $\hat{b}_j^\dagger$

Then we have

$$\hat{b}_j^\dagger \hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto -\hat{b}_j^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

Constructing the creation operator

Suppose now that we add another particle

- this time in state j
  - using the operator $\hat{b}_j^\dagger$

Then we have

$$\hat{b}_j^\dagger \hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto - \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle & |4,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle & |4,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle & |4,m\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle & |4,j\rangle \end{vmatrix}$$

Constructing the creation operator

To get to standard order

- we have to swap the bottom j row with the adjacent m row
  - multiplying by -1

$$\hat{b}_j^\dagger \hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto -(-1) \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle & |4,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle & |4,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle & |4,m\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle & |4,j\rangle \end{vmatrix}$$

Constructing the creation operator

Then we swap the j row

now second from the bottom

with the adjacent k row,

multiplying again by -1

$$\hat{b}_j^\dagger \hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto -(-1)^2 \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle & |4,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle & |4,k\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle & |4,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle & |4,m\rangle \end{vmatrix}$$

Constructing the creation operator

So finally

$$\hat{b}_j^{\dagger}\hat{b}_k^{\dagger}\begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto -(-1)^2 \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle & |4,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle & |4,j\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle & |4,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle & |4,m\rangle \end{vmatrix}$$

Constructing the creation operator

Now suppose we do this two-particle creation operation in the opposite order

- First, similarly to before, but first with $\hat{b}_j^\dagger$
  - and performing the necessary swap of the bottom two rows

$$\hat{b}_j^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto -\begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

Constructing the creation operator

Next, if we operate with $\hat{b}_k^\dagger$

$$\hat{b}_k^\dagger \hat{b}_j^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto -\hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

Constructing the creation operator

Next, if we operate with $\hat{b}_k^\dagger$

adding a k row to the bottom (and a column to the right)

we obtain

$$\hat{b}_k^\dagger \hat{b}_j^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto - \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle & |4,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle & |4,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle & |4,m\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle & |4,k\rangle \end{vmatrix}$$

Constructing the creation operator

Now, however

we only have to swap adjacent rows once

not twice

to get the determinant into standard order

$$\hat{b}_k^\dagger \hat{b}_j^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} \propto \overset{\Downarrow}{+} \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle & |4,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle & |4,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle & |4,m\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle & |4,k\rangle \end{vmatrix}$$

This result is –1 times the result from that of the operators in the order $\hat{b}_j^\dagger \hat{b}_k^\dagger$

Sign behavior for creation operator pairs

For example, we would get the same difference in sign

- if we had considered the pairs of operators
  - $\hat{b}_a^\dagger \hat{b}_k^\dagger$ and $\hat{b}_k^\dagger \hat{b}_a^\dagger$
    - or the pairs $\hat{b}_j^\dagger \hat{b}_n^\dagger$ and $\hat{b}_n^\dagger \hat{b}_j^\dagger$

Note one of the pairs of operators always results in one more swap of adjacent rows than the other

- because it encounters one more row to be swapped

Sign behavior for creation operator pairs

Hence we have the result

valid for any state with single-particle states j and k initially empty

$$\boxed{\hat{b}_j^\dagger \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j^\dagger = 0}$$

In fact

this relation is universally true for any state

as we can now show

Sign behavior for creation operator pairs

For any state in which state k is initially occupied

- the fermion creation operator for that state
  - must have the property that $\hat{b}_k^\dagger \left| \ldots, 1_k, \ldots \right\rangle = 0$
    - because we cannot create two fermions in one single-particle state

Hence when the single-particle state k is occupied

- trivially we have $\hat{b}_j^\dagger \hat{b}_k^\dagger \left| \ldots, 1_k, \ldots \right\rangle = 0$ and $\hat{b}_k^\dagger \hat{b}_j^\dagger \left| \ldots, 1_k, \ldots \right\rangle = 0$

Hence $\hat{b}_j^\dagger \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j^\dagger = 0$ still works here

- because each individual term is zero
  - and similarly when state j is initially occupied

Sign behavior for creation operator pairs

We also trivially get the same result $\hat{b}_j^\dagger \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j^\dagger = 0$

for any initial state with $j = k$

because we are trying to create at least two fermions in the single-particle state

three if it is already occupied

and so we also get zero for both terms

Hence we conclude that

$$\hat{b}_j^\dagger \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j^\dagger = 0$$

is valid for any starting state

Anticommutation relation

A relation of the form $\hat{b}_j^\dagger \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j^\dagger = 0$

is called an anticommutation relation

It is like a commutation relation between operators

but with a plus sign in the middle, rather than

the minus sign of a commutation relation

A notation sometimes used for an anticommutator of two operators is

$$\hat{b}_j^\dagger \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j^\dagger \equiv \left\{ \hat{b}_j^\dagger, \hat{b}_k^\dagger \right\}$$

Now we will progressively develop a family of anticommutation relations for the fermion operators

Formalization of creation operator sign behavior

We see, with our choice that we add the particle in state k initially to the end of the list

or, equivalently, to the bottom of the determinant

and then swap it into place

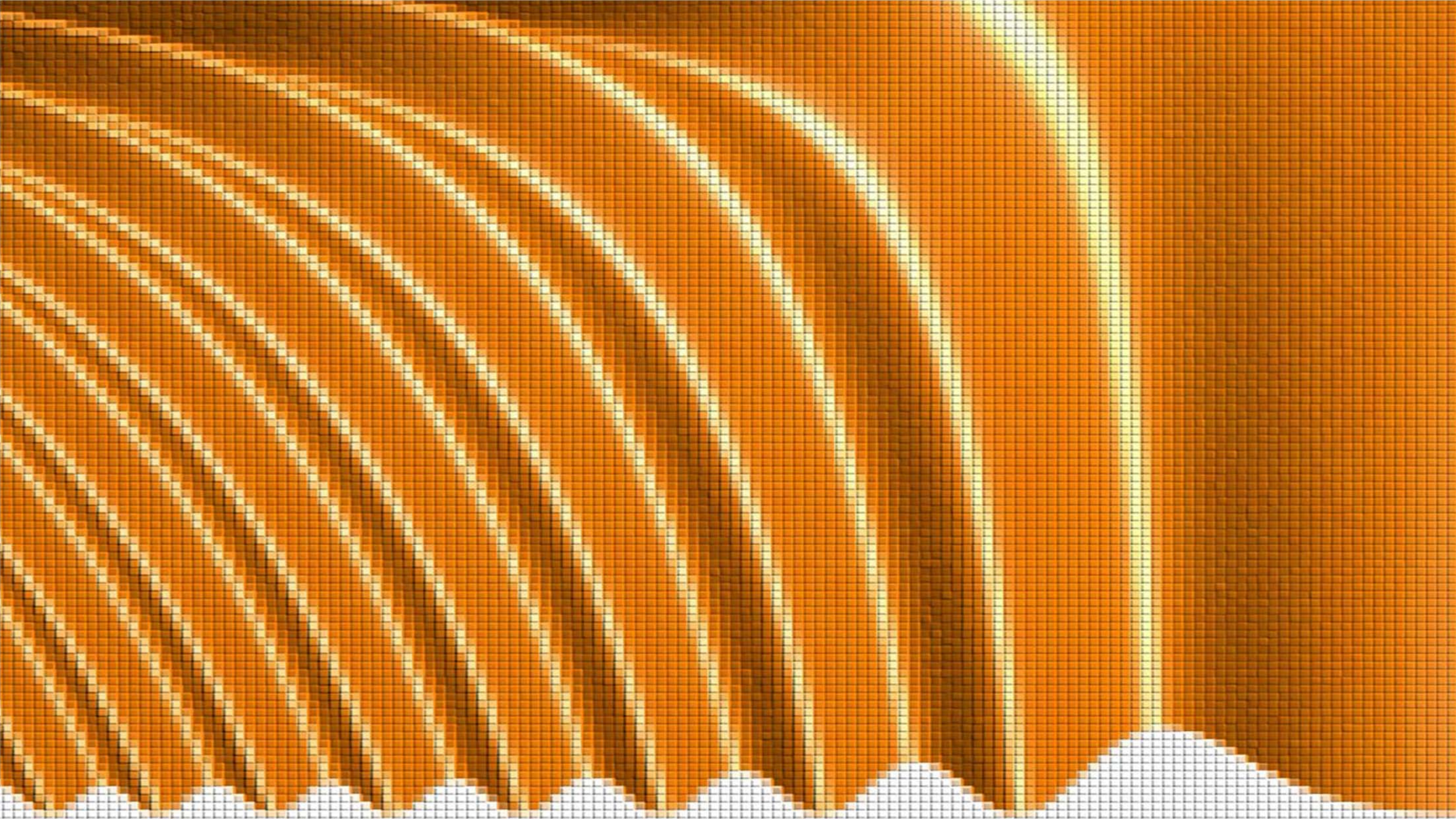
that the number of swaps we have to perform is

the number, S_k, of occupied states that are

after the state k in the standard order

With this definition

we have formally $\hat{b}_k^\dagger |\ldots, 0_k, \ldots\rangle = (-1)^{S_k} |\ldots, 1_k, \ldots\rangle$

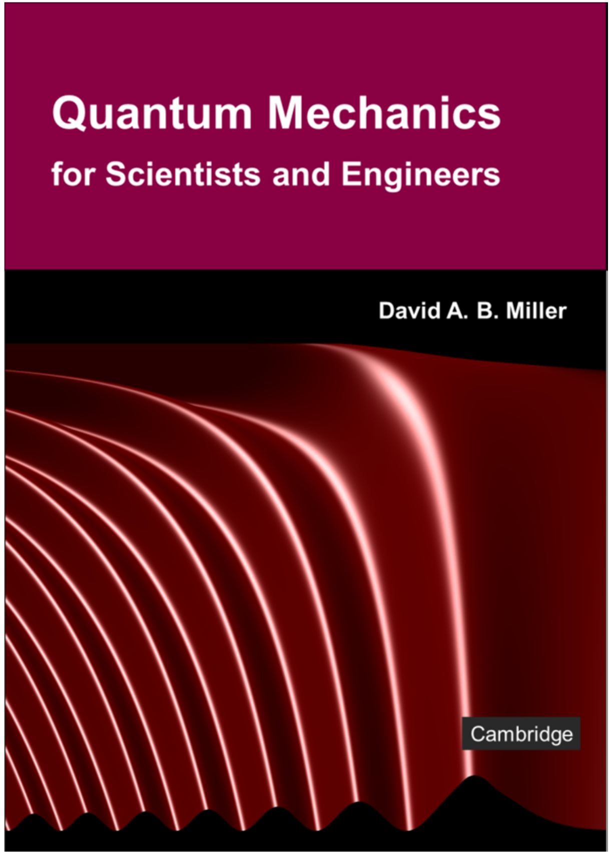


15.3 Fermion annihilation and creation operators

Slides: Video 15.3.5 Annihilation operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 16.1 subsection “Fermion annihilation operators”



Fermion annihilation and creation operators

Annihilation operators

Quantum mechanics for scientists and engineers

David Miller

Fermion annihilation operators

Now we can proceed to define annihilation operators

From $\hat{b}_k^\dagger \left|\ldots,0_k,\ldots\right\rangle = (-1)^{S_k} \left|\ldots,1_k,\ldots\right\rangle$ we can see that

$$\left\langle\ldots,1_k,\ldots\right|\hat{b}_k^\dagger\left|\ldots,0_k,\ldots\right\rangle = (-1)^{S_k}\left\langle\ldots,1_k,\ldots\middle|\ldots,1_k,\ldots\right\rangle = (-1)^{S_k}$$

Taking the complex conjugate

or, actually, the Hermitian adjoint, of both sides

$$\left\langle\ldots,1_k,\ldots\right|\hat{b}_k^\dagger\left|\ldots,0_k,\ldots\right\rangle^\dagger = \left(\hat{b}_k^\dagger\left|\ldots,0_k,\ldots\right\rangle\right)^\dagger\left(\left\langle\ldots,1_k,\ldots\right|\right)^\dagger$$

$$= \left\langle\ldots,0_k,\ldots\right|\hat{b}_k\left|\ldots,1_k,\ldots\right\rangle = (-1)^{S_k}$$

So, we deduce $\hat{b}_k\left|\ldots,1_k,\ldots\right\rangle = (-1)^{S_k}\left|\ldots,0_k,\ldots\right\rangle$

Fermion annihilation operators

Hence

whereas $\hat{b}_k^\dagger$ creates a fermion in single-particle state k

provided that state was empty

we see from $\hat{b}_k \left|\ldots,1_k,\ldots\right\rangle = (-1)^{S_k} \left|\ldots,0_k,\ldots\right\rangle$

$\hat{b}_k$ annihilates a fermion in single-particle state k

provided that state was full

and is called the fermion annihilation operator for state k

Annihilation operator acting on a state

The annihilation operator acting on the Slater determinant progressively

- swaps the row corresponding to state k in the determinant with the one below it
    - until that row gets to the bottom of the determinant
        - in which case we remove it (and the last column) of the determinant
            - in an inverse fashion to the process with the creation operator

Anticommutator for annihilation operators

Using the relation (analogous to $\hat{b}_k^\dagger |\ldots, 1_k, \ldots\rangle = 0$)

$$\hat{b}_k |\ldots, 0_k, \ldots\rangle = 0$$

which merely states that, if the single-particle state k is empty to start with

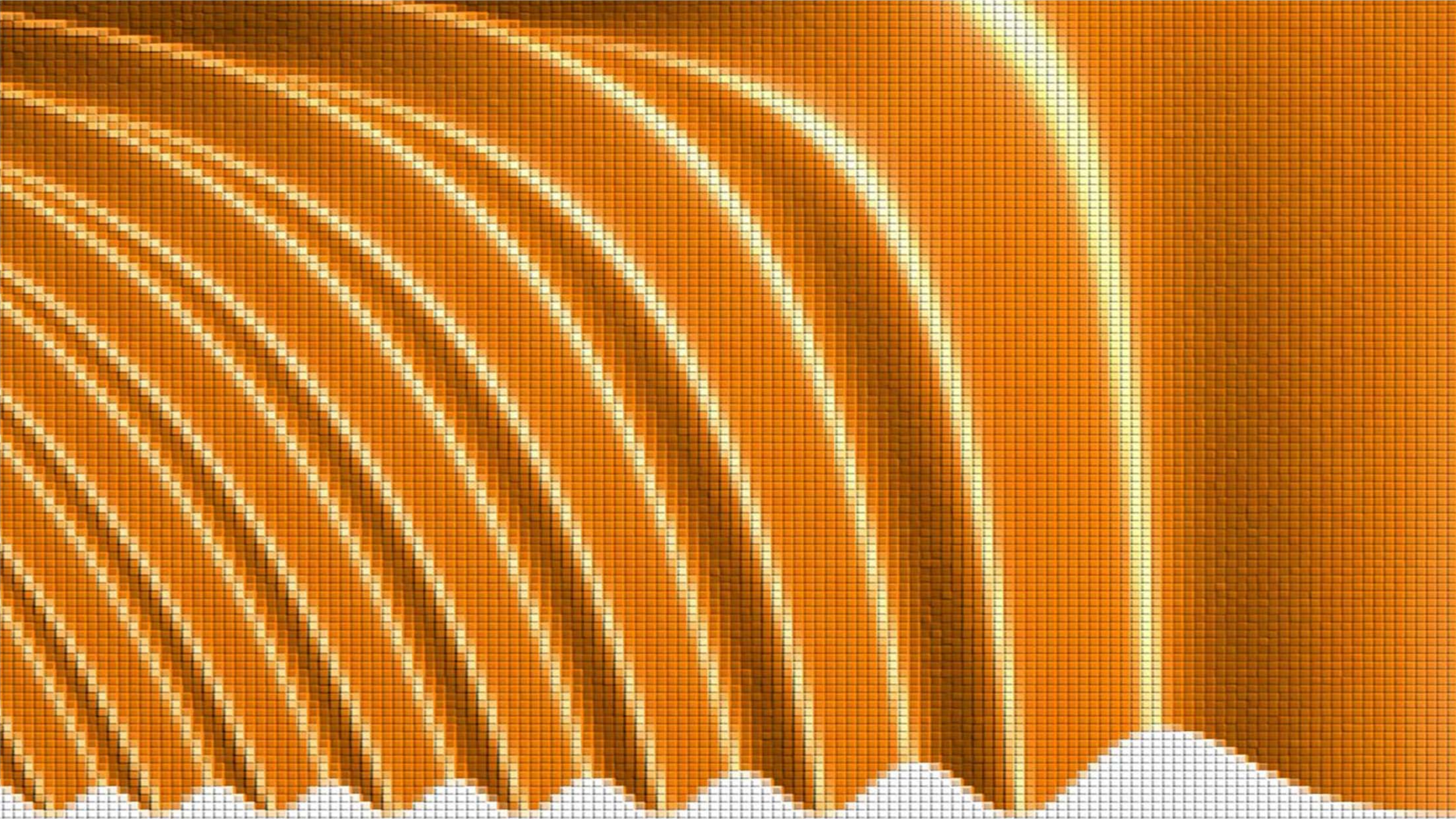
we cannot annihilate another particle from that state

we can argue similarly that

$$\boxed{\hat{b}_j \hat{b}_k + \hat{b}_k \hat{b}_j = 0}$$

which is the annihilation operator anticommutation relation

valid for all states and for $j = k$

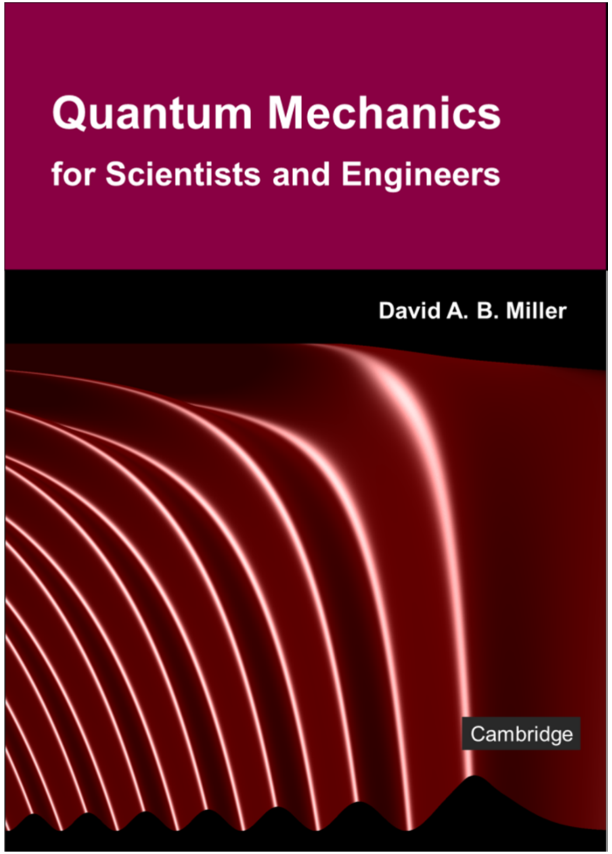


15.3 Fermion annihilation and creation operators

Slides: Video 15.3.7 Mixtures of creation and annihilation operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 16.1 subsection "Mixtures of creation and annihilation operators"



Fermion annihilation and creation operators

Mixtures of creation and annihilation operators

Quantum mechanics for scientists and engineers

David Miller

Mixtures of creation and annihilation operators

Suppose single-particle states b, j, and m are initially occupied

and we operate on this state first with the annihilation operator $\hat{b}_j$

Then we have

$$\hat{b}_j \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} = -\begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix}$$

where we had to swap the j and m rows to get the j row to the bottom

Mixtures of creation and annihilation operators

Now we operate with $\hat{b}_k^\dagger$, obtaining

$$\hat{b}_k^\dagger \hat{b}_j \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} = -\hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |2,m\rangle \end{vmatrix} = \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

where the minus sign is cancelled because we had to swap the k row from the bottom with the m row

Mixtures of creation and annihilation operators

Next let us consider applying these operators in the opposite order, starting now with $\hat{b}_k^\dagger$

$$\hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} = -\begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle & |4,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle & |4,j\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle & |4,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle & |4,m\rangle \end{vmatrix}$$

where we had to swap the k row from the bottom with the m row

Mixtures of creation and annihilation operators

Applying the $\hat{b}_j$ operator now gives

$$\hat{b}_j \hat{b}_k^\dagger \begin{Vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{Vmatrix} = -\hat{b}_j \begin{Vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle & |4,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle & |4,j\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle & |4,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle & |4,m\rangle \end{Vmatrix}$$

$$= -\begin{Vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{Vmatrix}$$

In operating with $\hat{b}_j$, two swaps are required because we have to swap past both the m and k rows.

Mixtures of creation and annihilation operators

As before, we find an additional row swap required with one order of operators rather than the other

The result

$$\hat{b}_k^\dagger \hat{b}_j \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} = \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

is minus the result

$$\hat{b}_j \hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,j\rangle & |2,j\rangle & |3,j\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} = -\begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

Mixtures of creation and annihilation operators

Hence, at least when operating on states when single-particle state j is initially full and single-particle state k is initially empty

$$\boxed{\hat{b}_j \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j = 0}$$

Again, if state j is initially empty, then both pairs of operators will lead to a zero result

and similarly if state k is initially full

Hence, as long as states j and k are different states

this relation is universally true

Mixtures of creation and annihilation operators

The only special case we have to consider more carefully here is for $j = k$

Suppose single-particle state k is initially full

Then we have

$$\hat{b}_k \hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} = 0$$

because $\hat{b}_k^\dagger$ operating on this state gives zero

Mixtures of creation and annihilation operators

For the other order of operators, we have

$$\hat{b}_k^\dagger \hat{b}_k \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix} = -\hat{b}_k^\dagger \begin{vmatrix} |1,b\rangle & |2,b\rangle \\ |1,m\rangle & |1,m\rangle \end{vmatrix} = \begin{vmatrix} |1,b\rangle & |2,b\rangle & |3,b\rangle \\ |1,k\rangle & |2,k\rangle & |3,k\rangle \\ |1,m\rangle & |2,m\rangle & |3,m\rangle \end{vmatrix}$$

It is left as an exercise to repeat this derivation for the situation where state k is initially empty

In both cases, the result is the same

One or other of the pairs returns the original state

and the other pair returns zero

Anticommutation relation for mixed operators

Hence we can say that $\hat{b}_k \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_k = 1$

Putting this together with $\hat{b}_j \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j = 0$

we can write the anticommutation relation for mixed annihilation and creation operators

$$\boxed{\hat{b}_j \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j = \delta_{jk}}$$

Fermion number operator

Finally, we note that $\hat{b}_k^\dagger \hat{b}_k$

- is the fermion number operator for the state k
  - i.e., it will tell us the number of fermions occupying state k

If the state is initially empty

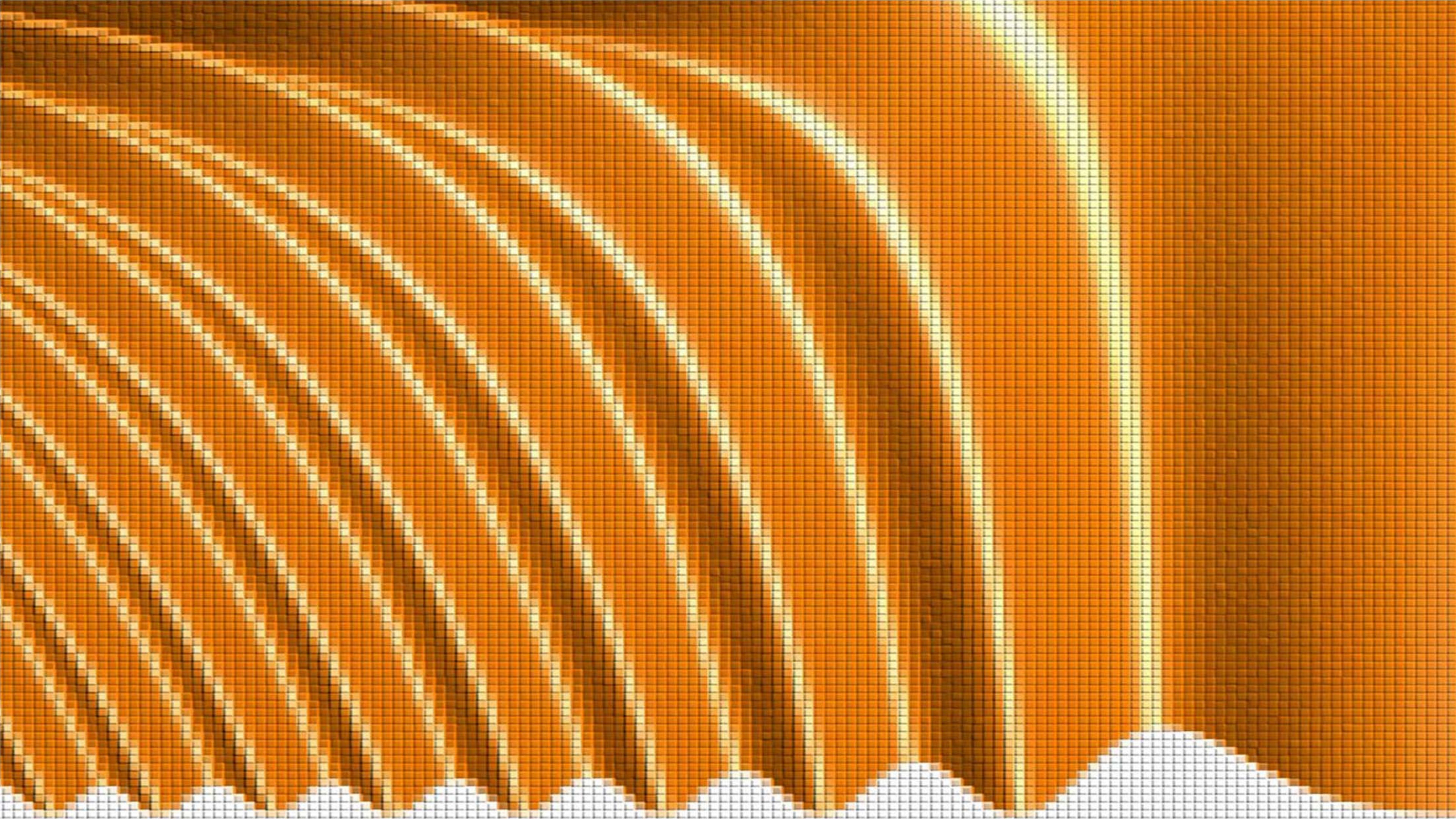
- it will return the value zero

and if the state is initially full

- it will return the value 1

We can write this as

$$\boxed{\hat{N}_k = \hat{b}_k^\dagger \hat{b}_k}$$

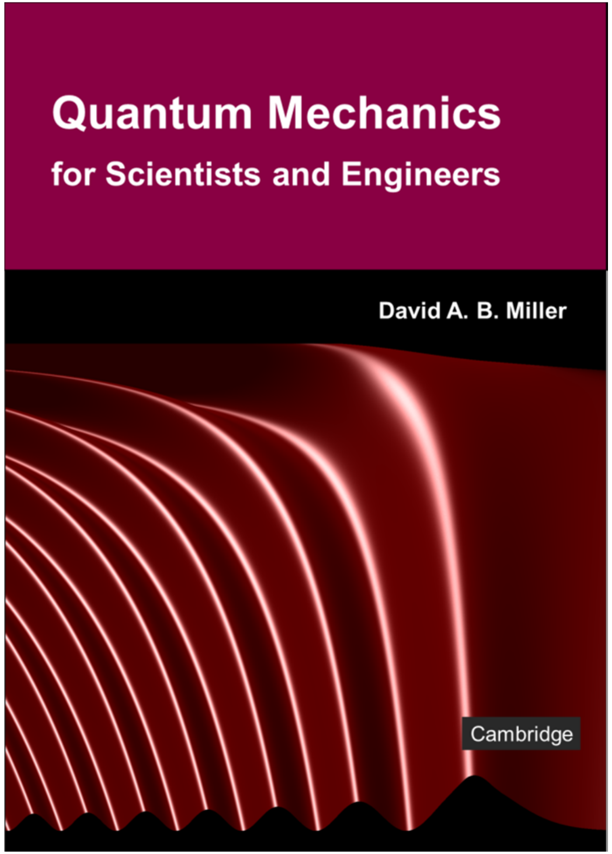


16.1 Fermion wavefunction and Hamiltonian operators

Slides: Video 16.1.1 Wavefunction operator

Text reference: Quantum Mechanics for Scientists and Engineers

Section 16.2



Fermion wavefunction and Hamiltonian operators

Wavefunction operator

Quantum mechanics for scientists and engineers

David Miller

Fermion wavefunction operator

We would like an operator, in occupation number form

representing the quantum mechanical wavefunction

to give a way to transform operators, such as Hamiltonians, from a spatial form

into the occupation number representation

We propose first a wavefunction operator when we have a single particle

$$\boxed{\hat{\psi}(\mathbf{r}) = \sum_j \hat{b}_j \phi_j(\mathbf{r})}$$

where the $\phi_k(\mathbf{r})$ are some complete set for describing functions of space

Fermion wavefunction operator

Suppose a single particle was in state m

i.e., with wavefunction $\phi_m(\mathbf{r})$

We can also write that state as

$$|\ldots 0_l, 1_m, 0_n, \ldots\rangle \equiv \hat{b}_m^\dagger |0\rangle$$

where $|0\rangle$ is the state with

no fermions present in any single-particle state

Then we find that

$$\hat{\psi}(\mathbf{r})|\ldots 0_l, 1_m, 0_n, \ldots\rangle = \hat{\psi}(\mathbf{r})\hat{b}_m^\dagger |0\rangle = \sum_j \phi_j(\mathbf{r})\hat{b}_j \hat{b}_m^\dagger |0\rangle$$

Fermion wavefunction operator

Now in this relation $\hat{\psi}(\mathbf{r})|\ldots 0_l, 1_m, 0_n, \ldots\rangle = \sum_j \phi_j(\mathbf{r}) \hat{b}_j \hat{b}_m^\dagger |0\rangle$

we use the anticommutation relation $\hat{b}_j \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j = \delta_{jk}$

obtaining $\hat{\psi}(\mathbf{r})|\ldots 0_l, 1_m, 0_n, \ldots\rangle = \sum_j \phi_j(\mathbf{r}) \left(\delta_{jm} - \hat{b}_m^\dagger \hat{b}_j \right) |0\rangle$

But $\hat{b}_j |0\rangle = 0$ because an attempt to annihilate a particle that is not there results in a null result

(The right hand side here is actually a vector of zero length)

Hence we have $\hat{\psi}(\mathbf{r})|\ldots 0_l, 1_m, 0_n, \ldots\rangle = \phi_m(\mathbf{r})|0\rangle$

Fermion wavefunction operator

We can see then from $\hat{\psi}(\mathbf{r})|\ldots 0_l, 1_m, 0_n, \ldots\rangle = \phi_m(\mathbf{r})|0\rangle$

- that this operator has successfully extracted the amplitude $\phi_m(\mathbf{r})$

We have also acquired the ket $|0\rangle$ in the result

- which might seem odd
  - but we should have a state vector here
    - because the result of operating on a state vector
      - should be a state vector

Fermion wavefunction operator

By a simple extension of the above algebra

if the particle is in a linear superposition, i.e.,

$$|\psi_S\rangle = \sum_k c_k |\ldots, 1_k, \ldots\rangle$$

where by $|\ldots, 1_k, \ldots\rangle$ we mean the state with

one particle in state k and no other single-particle states occupied

then $\hat{\psi}(\mathbf{r})|\psi_S\rangle = \left(\sum_k c_k \phi_k(\mathbf{r})\right)|0\rangle$

which has extracted the desired linear superposition of wavefunctions

Two-fermion wavefunction operator

Next we propose a wavefunction operator for a two-fermion state

$$\hat{\psi}\left(\mathbf{r}_1,\mathbf{r}_2\right)=\frac{1}{\sqrt{2}}\sum_{j,n}\hat{b}_n\hat{b}_j\phi_j\left(\mathbf{r}_1\right)\phi_n\left(\mathbf{r}_2\right)$$

The $1/\sqrt{2}$ is for normalization of the final result

It is left as an exercise to show that such an operator

operating on a state with two different single-particle states occupied

leads to a linear combination of products of wavefunctions that is correctly antisymmetric with respect to exchange of these two particles

Two-fermion wavefunction operator

That is, if this operator $\hat{\psi}(\mathbf{r}_1,\mathbf{r}_2)=\frac{1}{\sqrt{2}}\sum_{j,n}\hat{b}_n\hat{b}_j\phi_j(\mathbf{r}_1)\phi_n(\mathbf{r}_2)$

acts on a state s that has

one fermion in single-particle state k

and an identical fermion in single-particle state m

i.e., the state $\left|\ldots,1_k,\ldots,1_m,\ldots\right\rangle\equiv\hat{b}_k^\dagger\hat{b}_m^\dagger\left|0\right\rangle$

then

$$\hat{\psi}(\mathbf{r}_1,\mathbf{r}_2)\left|\ldots,1_k,\ldots,1_m,\ldots\right\rangle=\frac{1}{\sqrt{2}}\left[\phi_k(\mathbf{r}_1)\phi_m(\mathbf{r}_2)-\phi_k(\mathbf{r}_2)\phi_m(\mathbf{r}_1)\right]\left|0\right\rangle$$

Multiple fermion wavefunction operator

We can propose to extend such wavefunction operators

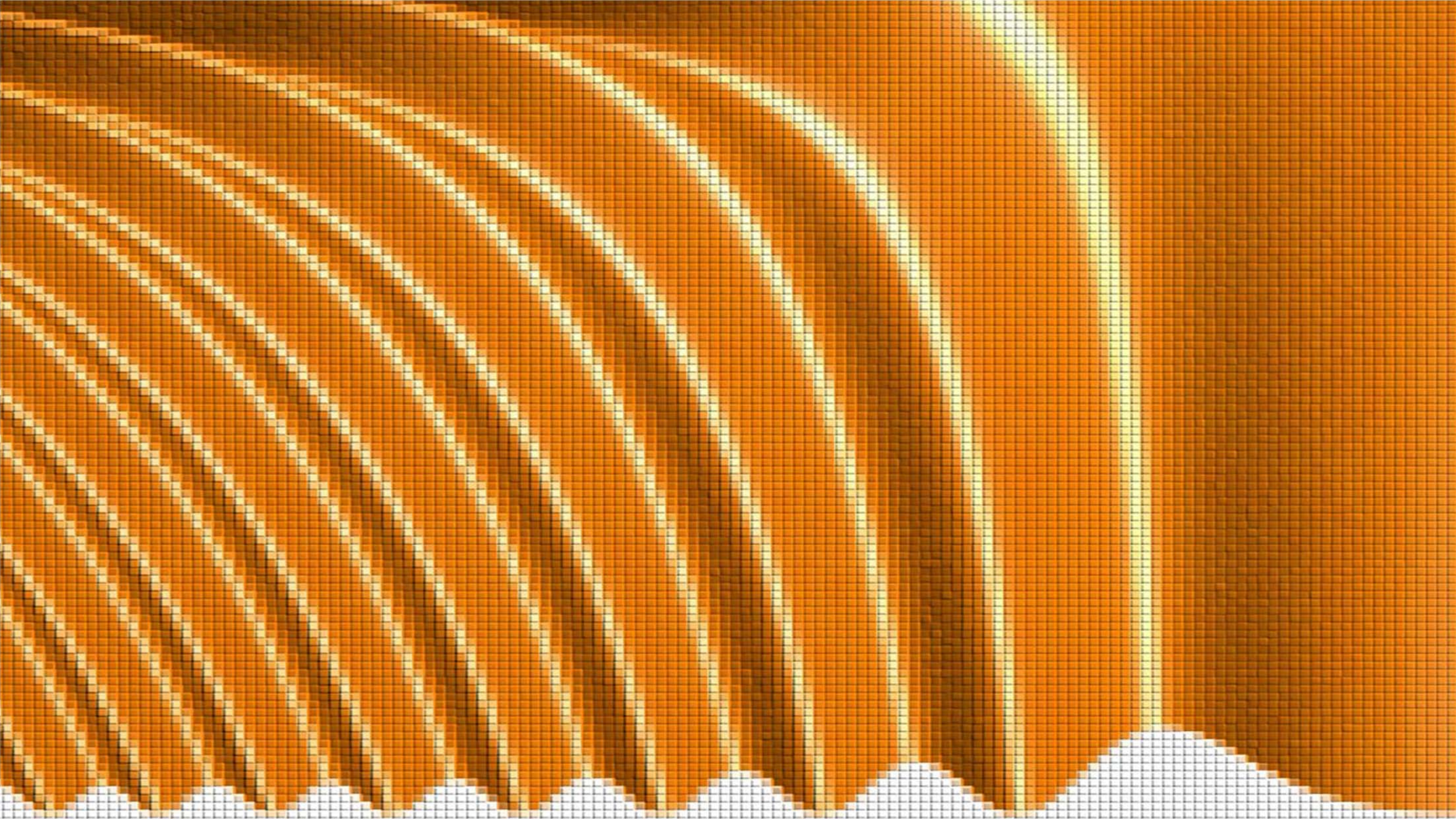
to larger numbers of particles

postulating

$$\hat{\psi}\left(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_N\right)=\frac{1}{\sqrt{N}}\sum_{a,b,\ldots,n}\hat{b}_n\ldots\hat{b}_b\hat{b}_a\phi_a\left(\mathbf{r}_1\right)\phi_b\left(\mathbf{r}_2\right)\ldots\phi_n\left(\mathbf{r}_N\right)$$

with the expectation that these operators will also extract the correct sum of permutations

to give wavefunctions antisymmetric with respect to exchange

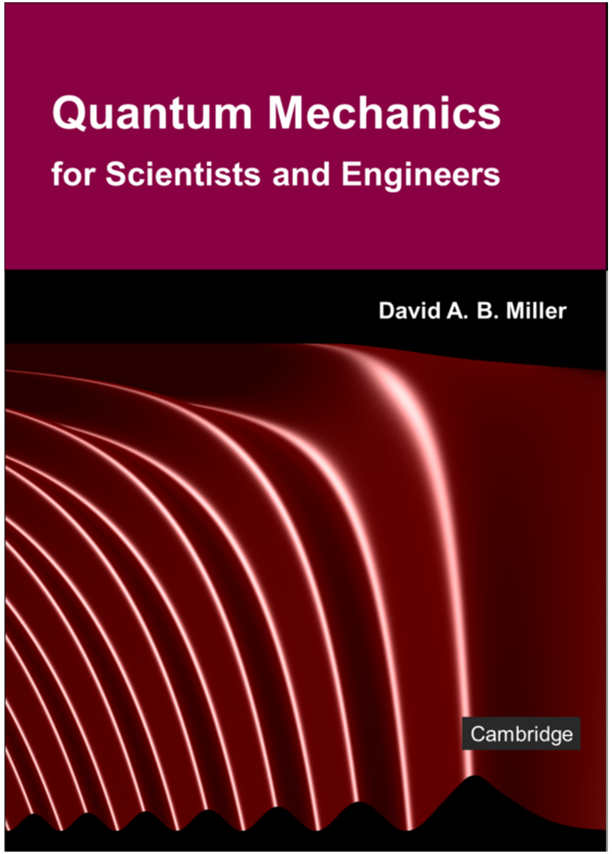


16.1 Fermion wavefunction and Hamiltonian operators

Slides: Video 16.1.3 Representing fermion Hamiltonians

Text reference: Quantum Mechanics for Scientists and Engineers

Section 16.3 up to paragraph after Eq. 16.56



Fermion wavefunction and Hamiltonian operators

Representing fermion Hamiltonians

Quantum mechanics for scientists and engineers

David Miller

Single-particle fermion Hamiltonian

Consider a Hamiltonian for a single fermion

Previously, we had a simple Hamiltonian such as the simplest Schrödinger equation for a single particle

$$\hat{H}_{\mathbf{r}} = -\frac{\hbar^2}{2m}\nabla_{\mathbf{r}}^2 + V(\mathbf{r})$$

For any given state $|\psi\rangle$ of spatial wavefunction $\psi(\mathbf{r})$

the expected value for energy was

$$\langle E\rangle = \langle\psi|\hat{H}_{\mathbf{r}}|\psi\rangle = \int \psi^*(\mathbf{r})\left[-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}}^2 + V(\mathbf{r})\right]\psi(\mathbf{r})\,d^3\mathbf{r}$$

Postulating a Hamiltonian with fermion character

To give the Hamiltonian fermion particle character

we substitute for the wavefunction in

$$\langle E\rangle = \langle\psi|\hat{H}_{\mathbf{r}}|\psi\rangle = \int \psi^*(\mathbf{r})\left[-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}}^2 + V(\mathbf{r})\right]\psi(\mathbf{r})\,d^3\mathbf{r}$$

with the wavefunction operator

generating our desired new fermion operator

instead of the expectation value

Hence we obtain a single particle Hamiltonian operator

$$\hat{H} \equiv \int \hat{\psi}^\dagger(\mathbf{r})\left[-\frac{\hbar^2}{2m}\nabla_{\mathbf{r}}^2 + V(\mathbf{r})\right]\hat{\psi}(\mathbf{r})\,d^3\mathbf{r}$$

Postulating a Hamiltonian with fermion character

Presume that the single-particle basis states with spatial wavefunctions $\phi_m(\mathbf{r})$

are the eigenstates of this single particle Hamiltonian

with corresponding eigenenergies E_m

Now using the wavefunction operator $\hat{\psi}(\mathbf{r}) = \sum_j \hat{b}_j \phi_j(\mathbf{r})$

we obtain

$$\hat{H} = \int \sum_{j,k} \hat{b}_j^\dagger \hat{b}_k \phi_j^*(\mathbf{r}) \left[-\frac{\hbar^2}{2m} \nabla_{\mathbf{r}}^2 + V(\mathbf{r}) \right] \phi_k(\mathbf{r}) d^3\mathbf{r}$$

$$= \sum_{j,k} \hat{b}_j^\dagger \hat{b}_k E_k \int \phi_j^*(\mathbf{r}) \phi_k(\mathbf{r}) d^3\mathbf{r} = \sum_{j,k} \hat{b}_j^\dagger \hat{b}_k E_k \delta_{jk} = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j \equiv \sum_j E_j \hat{N}_j$$

Postulating a Hamiltonian with fermion character

Now we can use this elegant form $\hat{H} = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j$

Consider a state $|\psi\rangle$ as

a linear superposition of the basis states

In the $\mathbf{r}$ representation $|\psi\rangle = \sum_m c_m \phi_m(\mathbf{r})$

or equivalently in the number state notation

$$|\psi\rangle = \sum_m c_m \hat{b}_m^\dagger |0\rangle$$

where $\hat{b}_m^\dagger |0\rangle$ is a convenient way of writing the basis state in which

only single particle state m occupied

Postulating a Hamiltonian with fermion character

The Hermitian conjugate of $|\psi\rangle = \sum_m c_m \hat{b}_m^\dagger |0\rangle$

- is $\langle\psi| = \sum_m c_m^* \langle 0|\hat{b}_m$

So now we can evaluate the energy expectation value

- using our new forms of the state and the Hamiltonian

$$\langle E\rangle = \langle\psi|\hat{H}|\psi\rangle = \sum_{m,n,j} c_m^* c_n E_j \langle 0|\hat{b}_m \hat{b}_j^\dagger \hat{b}_j \hat{b}_n^\dagger |0\rangle$$

Now we simplify $\hat{b}_m \hat{b}_j^\dagger \hat{b}_j \hat{b}_n^\dagger |0\rangle$

- using anticommutation relations

Simplifying with anticommutation relations

A simple algebraic approach to simplify an expression like $\hat{b}_m \hat{b}_j^\dagger \hat{b}_j \hat{b}_n^\dagger |0\rangle$ is to

use the anticommutation relations to push annihilation operators to the right

That will lead to disappearance of terms because

an annihilation operator acting on the empty state gives a zero result

We will therefore keep making substitutions of the form

$$\hat{b}_m \hat{b}_j^\dagger = \delta_{mj} - \hat{b}_j^\dagger \hat{b}_m$$

which is just the anticommutation relation for $\hat{b}_m$ and $\hat{b}_j^\dagger$

Simplifying with anticommutation relations

Hence we have $\hat{b}_m \hat{b}_j^\dagger \hat{b}_j \hat{b}_n^\dagger |0\rangle = \left(\delta_{mj} - \hat{b}_j^\dagger \hat{b}_m\right)\left(\delta_{nj} - \hat{b}_n^\dagger \hat{b}_j\right)|0\rangle$

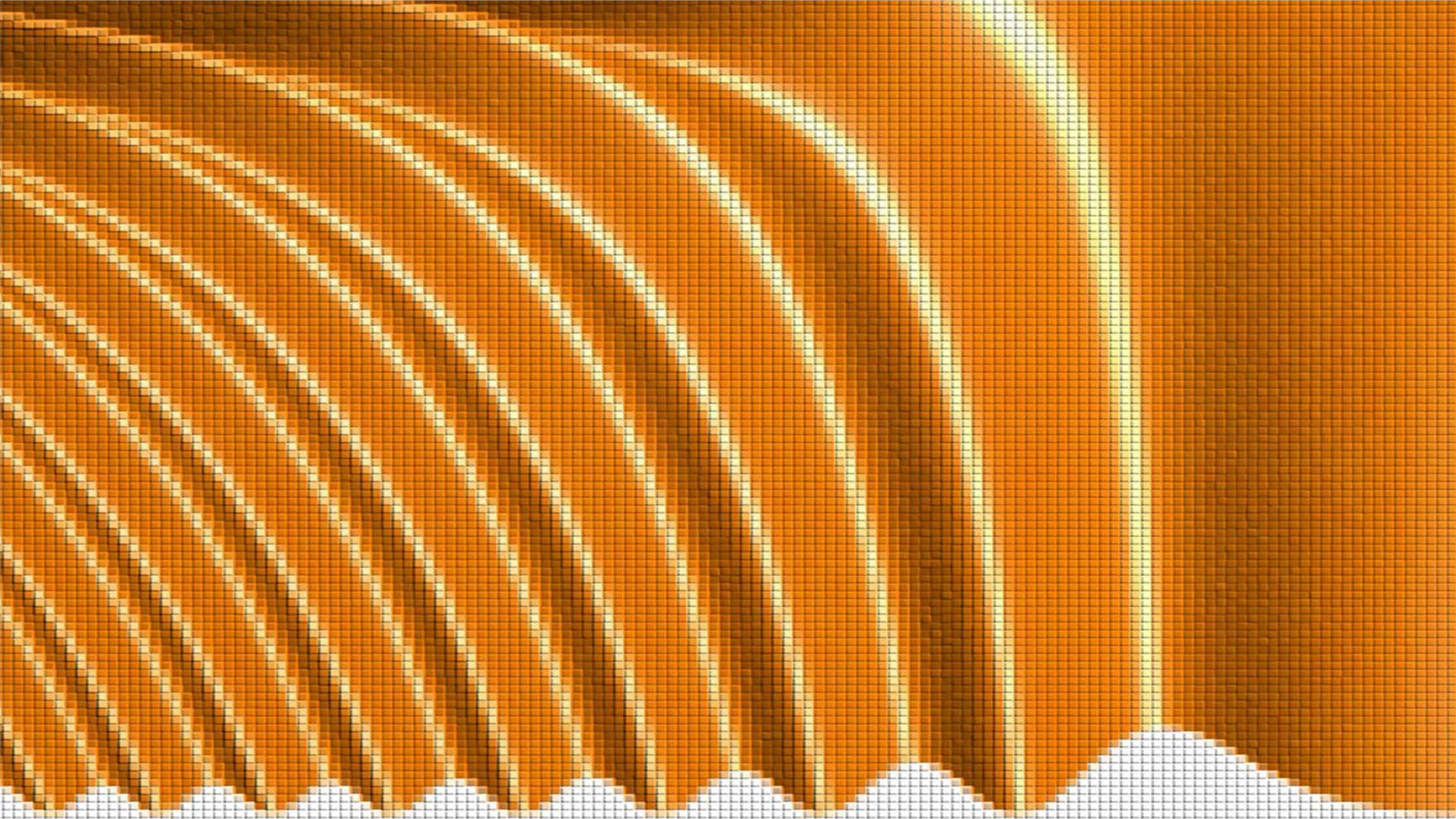
$$= \left(\delta_{mj} - \hat{b}_j^\dagger \hat{b}_m\right)\delta_{nj}|0\rangle = \delta_{mj}\delta_{nj}|0\rangle$$

So substituting back into

$$\langle E\rangle = \langle\psi|\hat{H}|\psi\rangle = \sum_{m,n,j} c_m^* c_n E_j \langle 0|\hat{b}_m \hat{b}_j^\dagger \hat{b}_j \hat{b}_n^\dagger|0\rangle$$

we have $\langle E\rangle = \sum_{m,n,j} c_m^* c_n E_j \delta_{mj}\delta_{nj} \langle 0|0\rangle = \sum_j |c_j|^2 E_j$

which is exactly the result we would have expected based on our previous approach

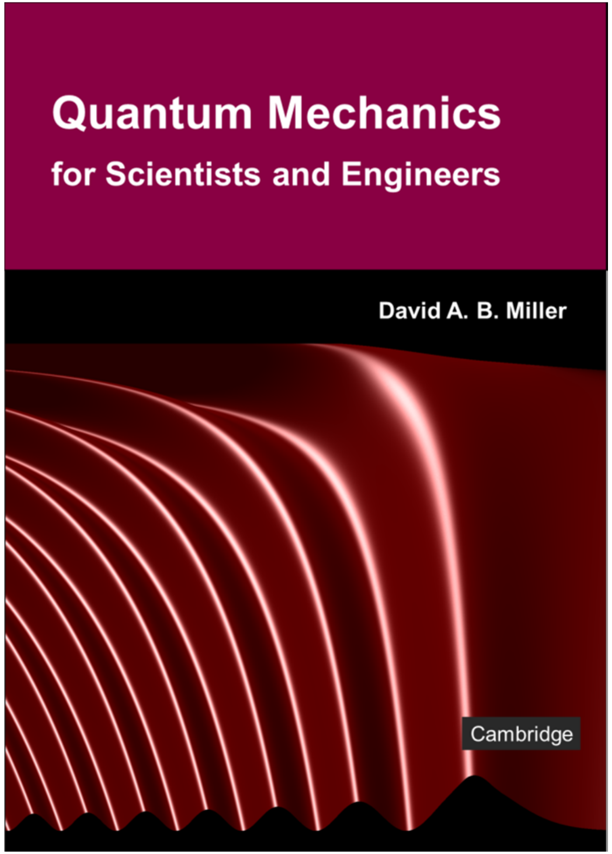


16.1 Fermion wavefunction and Hamiltonian operators

Slides: Video 16.1.5 Fermion Hamiltonians with multiple particle states

Text reference: Quantum Mechanics for Scientists and Engineers

Section 16.3 subsection “Single particle fermion Hamiltonians with multiple particle states”



Fermion wavefunction and Hamiltonian operators

Fermion Hamiltonians with multiple particle states

Quantum mechanics for scientists and engineers

David Miller

Single-particle Hamiltonians with multiple particle states

Suppose then that we have N identical fermions

Fermion i is presumed to have a single-particle Hamiltonian in the original $\mathbf{r}$ form such as

$$\hat{H}_{\mathbf{r}i} = -\frac{\hbar^2}{2m}\nabla^2_{\mathbf{r}_i} + V\left(\mathbf{r}_i\right)$$

Therefore, in the original $\mathbf{r}$ form

the total Hamiltonian for the set of N fermions is

$$\hat{H}_{\mathbf{r}} = \sum_{i=1}^{N} \hat{H}_{\mathbf{r}i}$$

Single-particle Hamiltonians with multiple particle states

We now show that, even for the multiple fermion case

we can still write the total Hamiltonian operator

exactly as in $\hat{H} = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j$

So, regardless of how many particles are in the system

we do not have to change the Hamiltonian for non-interacting fermions

i.e., we do not have to write a sum over all the particles

like $\hat{H}_{\mathbf{r}} = \sum_{i=1}^{N} \hat{H}_{\mathbf{r}i}$

Single-particle Hamiltonians with two particle states

So now we consider a single-particle fermion operator and a multi-fermion system

We now illustrate that

$$\hat{H} = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j$$

is still the Hamiltonian we would deduce when

we create the Hamiltonian using multiple fermion wavefunction operators

We show this Hamiltonian

works also for two fermions ($N = 2$)

Single-particle Hamiltonians with two particle states

Suppose we have a specific two-fermion state with

one fermion in single-particle state k

and one in single-particle state m

We can write that state as

$$\left|\psi_{TP}\right\rangle = \left|\ldots,1_k,\ldots,1_m,\ldots\right\rangle = \hat{b}_k^\dagger \hat{b}_m^\dagger \left|0\right\rangle$$

Evaluating with our new Hamiltonian form we have

$$\left\langle E\right\rangle = \left\langle\psi_{TP}\right|\hat{H}\left|\psi_{TP}\right\rangle = \sum_j \left(\hat{b}_k^\dagger \hat{b}_m^\dagger \left|0\right\rangle\right)^\dagger E_j \hat{b}_j^\dagger \hat{b}_j \hat{b}_k^\dagger \hat{b}_m^\dagger \left|0\right\rangle$$

$$= \sum_j E_j \left\langle 0\right| \hat{b}_m \hat{b}_k \hat{b}_j^\dagger \hat{b}_j \hat{b}_k^\dagger \hat{b}_m^\dagger \left|0\right\rangle$$

Single-particle Hamiltonians with two particle states

Now we simplify $\langle E\rangle = \sum_j E_j \langle 0|\hat{b}_m \hat{b}_k \hat{b}_j^\dagger \hat{b}_j \hat{b}_k^\dagger \hat{b}_m^\dagger |0\rangle$

using the anticommutation relation $\hat{b}_m \hat{b}_j^\dagger = \delta_{mj} - \hat{b}_j^\dagger \hat{b}_m$

to push the annihilation operators to the right

$$\hat{b}_m \hat{b}_k \hat{b}_j^\dagger \hat{b}_j \hat{b}_k^\dagger \hat{b}_m^\dagger |0\rangle = \hat{b}_m \left(\delta_{jk} - \hat{b}_j^\dagger \hat{b}_k\right)\left(\delta_{jk} - \hat{b}_k^\dagger \hat{b}_j\right)\hat{b}_m^\dagger |0\rangle$$

$$= \left(\delta_{jk}\hat{b}_m \hat{b}_m^\dagger - \delta_{jk}\hat{b}_m \hat{b}_k^\dagger \hat{b}_j \hat{b}_m^\dagger - \delta_{jk}\hat{b}_m \hat{b}_j^\dagger \hat{b}_k \hat{b}_m^\dagger + \hat{b}_m \hat{b}_j^\dagger \hat{b}_k \hat{b}_k^\dagger \hat{b}_j \hat{b}_m^\dagger\right)|0\rangle$$

$$= \Big[\delta_{jk}\left(1 - \hat{b}_m^\dagger \hat{b}_m\right) - \delta_{jk}\left(\delta_{mk} - \hat{b}_k^\dagger \hat{b}_m\right)\left(\delta_{mj} - \hat{b}_m^\dagger \hat{b}_j\right)$$

$$- \delta_{jk}\left(\delta_{mj} - \hat{b}_j^\dagger \hat{b}_m\right)\left(\delta_{mk} - \hat{b}_m^\dagger \hat{b}_k\right) + \left(\delta_{mj} - \hat{b}_j^\dagger \hat{b}_m\right)\left(1 - \hat{b}_k^\dagger \hat{b}_k\right)\left(\delta_{mj} - \hat{b}_m^\dagger \hat{b}_j\right)\Big]|0\rangle$$

Single-particle Hamiltonians with two particle states

Now in

$$\hat{b}_m \hat{b}_k \hat{b}_j^\dagger \hat{b}_j \hat{b}_k^\dagger \hat{b}_m^\dagger |0\rangle$$

$$= \Big[\delta_{jk} \left(1 - \hat{b}_m^\dagger \hat{b}_m\right) - \delta_{jk} \left(\delta_{mk} - \hat{b}_k^\dagger \hat{b}_m\right)\left(\delta_{mj} - \hat{b}_m^\dagger \hat{b}_j\right)$$

$$- \delta_{jk} \left(\delta_{mj} - \hat{b}_j^\dagger \hat{b}_m\right)\left(\delta_{mk} - \hat{b}_m^\dagger \hat{b}_k\right) + \left(\delta_{mj} - \hat{b}_j^\dagger \hat{b}_m\right)\left(1 - \hat{b}_k^\dagger \hat{b}_k\right)\left(\delta_{mj} - \hat{b}_m^\dagger \hat{b}_j\right)\Big] |0\rangle$$

we have annihilation operators on the far right

in every expression involving creation and annihilation operators

so all of those terms disappear ($\hat{b}_i |0\rangle = 0$ for any i)

Single-particle Hamiltonians with two particle states

Hence we have

$$\hat{b}_m \hat{b}_k \hat{b}_j^\dagger \hat{b}_j \hat{b}_k^\dagger \hat{b}_m^\dagger |0\rangle = \left(\delta_{jk} - \delta_{jk}\delta_{mk}\delta_{mj} - \delta_{jk}\delta_{mj}\delta_{mk} + \delta_{mj}\right)|0\rangle$$

But, by choice, m and k are different states so

δ_{mk} never has any value other than zero

Hence we have $\hat{b}_m \hat{b}_k \hat{b}_j^\dagger \hat{b}_j \hat{b}_k^\dagger \hat{b}_m^\dagger |0\rangle = \left(\delta_{jk} + \delta_{mj}\right)|0\rangle$

Substituting back into $\langle E \rangle = \sum_j E_j \langle 0 | \hat{b}_m \hat{b}_k \hat{b}_j^\dagger \hat{b}_j \hat{b}_k^\dagger \hat{b}_m^\dagger |0\rangle$

we have $\langle E \rangle = \sum_j E_j \left(\delta_{jk} + \delta_{mj}\right)\langle 0|0\rangle = E_k + E_m$

exactly as expected for two non-interacting fermions

Single-particle Hamiltonians with two particle states

Hence this illustration shows how the Hamiltonian

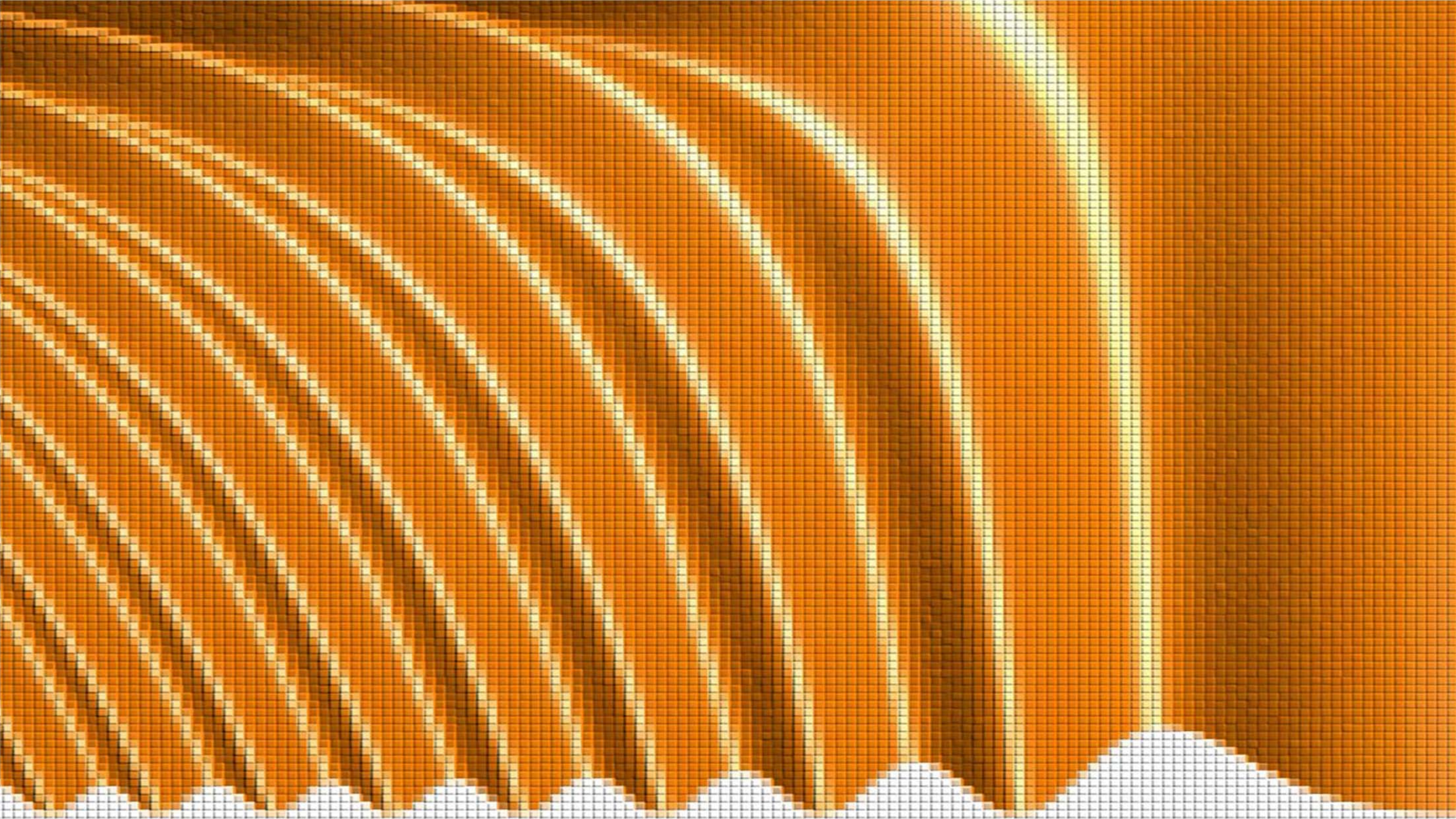
$$\boxed{\hat{H} = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j}$$

also works for multiple particle states.

Unlike the $\mathbf{r}$ representation of the Hamiltonian

we do not have to add separate Hamiltonians for each identical fermion

and hence we have an elegant form of Hamiltonian for multiple fermion systems

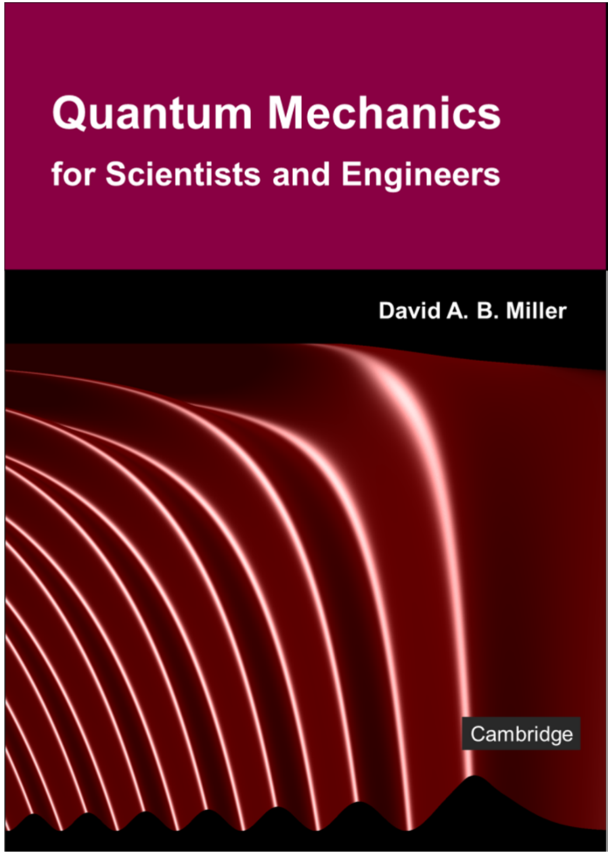


16.2 Fermion operators and multiple particles

Slides: Video 16.2.1 Single-particle fermion operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 16.3 subsection "Representation of general single-particle fermion operators"



Fermion operators and multiple particles

Single-particle fermion operators

Quantum mechanics for scientists and engineers

David Miller

General single-particle fermion operators

Here we consider a system with N fermions

In the $\mathbf{r}$ representation of an operator $\hat{G}_{\mathbf{r}}$

e.g., such as the momentum operator

for a multiple fermion system

we would add all of the operators

corresponding to the coordinates of each particle, i.e.,

$$\hat{G}_{\mathbf{r}} = \sum_{i=1}^{N} \hat{G}_{\mathbf{r}i}$$

where $\hat{G}_{\mathbf{r}i}$ is the operator for a specific particle

e.g., it might be the momentum operator

General single-particle fermion operators

In the annihilation and creation operator formalism

we postulate instead that

$$\hat{G} = \int \hat{\psi}^{\dagger} \hat{G}_{\mathbf{r}} \hat{\psi}\, d^3\mathbf{r}_1 d^3\mathbf{r}_2 \ldots d^3\mathbf{r}_N$$

where

$$\hat{\psi}\left(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N\right) = \frac{1}{\sqrt{N}} \sum_{a,b,\ldots,n} \hat{b}_n \ldots \hat{b}_b \hat{b}_a \phi_a\left(\mathbf{r}_1\right) \phi_b\left(\mathbf{r}_2\right) \ldots \phi_n\left(\mathbf{r}_N\right)$$

is the N-particle fermion wavefunction operator, so

$$\hat{G} = \frac{1}{N} \sum_{i=1}^{N} \sum_{\substack{a,b,\ldots n \\ a',b',\ldots n'}} \hat{b}_{a'}^{\dagger} \hat{b}_{b'}^{\dagger} \ldots \hat{b}_{n'}^{\dagger} \hat{b}_n \ldots \hat{b}_b \hat{b}_a$$

$$\times \int \phi_{a'}^{*}\left(\mathbf{r}_1\right) \phi_{b'}^{*}\left(\mathbf{r}_2\right) \ldots \phi_{n'}^{*}\left(\mathbf{r}_N\right) \hat{G}_{\mathbf{r}i} \phi_a\left(\mathbf{r}_1\right) \phi_b\left(\mathbf{r}_2\right) \ldots \phi_n\left(\mathbf{r}_N\right) d^3\mathbf{r}_1 d^3\mathbf{r}_2 \ldots d^3\mathbf{r}_N$$

General single-particle fermion operators

In

$$\hat{G} = \frac{1}{N}\sum_{i=1}^{N}\sum_{\substack{a,b,\dots n\\ a',b',\dots n'}} \hat{b}_{a'}^{\dagger}\hat{b}_{b'}^{\dagger}\dots\hat{b}_{n'}^{\dagger}\hat{b}_{n}\dots\hat{b}_{b}\hat{b}_{a}$$

$$\times\int \phi_{a'}^{*}\left(\mathbf{r}_1\right)\phi_{b'}^{*}\left(\mathbf{r}_2\right)\dots\phi_{n'}^{*}\left(\mathbf{r}_N\right)\hat{G}_{\mathbf{r}i}\phi_{a}\left(\mathbf{r}_1\right)\phi_{b}\left(\mathbf{r}_2\right)\dots\phi_{n}\left(\mathbf{r}_N\right)d^3\mathbf{r}_1 d^3\mathbf{r}_2\dots d^3\mathbf{r}_N$$

each of the a, b, …, n and each of the a', b', …, n'

ranges over all single-particle fermion states

Now, all the spatial integrals, except the one over r_i

lead to Kronecker deltas of the form $\delta_{k'k}$

forcing $a' = a$, $b' = b$, etc., except for particle i

General single-particle fermion operators

Hence $\hat{G} = \frac{1}{N}\sum_{i=1}^{N} \sum_{a,b,\dots,i1,i2,\dots n} G_{i1i2} \hat{b}_a^\dagger \hat{b}_b^\dagger \dots \hat{b}_{i1}^\dagger \dots \hat{b}_n^\dagger \hat{b}_n \dots \hat{b}_{i2} \dots \hat{b}_b \hat{b}_a$

where $G_{i1i2} = \int \phi_{i1}^*(\mathbf{r}_i) \hat{G}_{\mathbf{r}i} \phi_{i2}(\mathbf{r}_i) d^3\mathbf{r}_i$

We can use the anticommutation relation $\hat{b}_j \hat{b}_k + \hat{b}_k \hat{b}_j = 0$

to progressively swap the operator $\hat{b}_{i2}$

from the right to the center

and the anticommutation relation $\hat{b}_j^\dagger \hat{b}_k^\dagger + \hat{b}_k^\dagger \hat{b}_j^\dagger = 0$

to progressively swap the operator $\hat{b}_{i1}^\dagger$

from the left to the center

General single-particle fermion operators

Each such application of an anticommutation relation results in a sign change

but there are equal number of swaps from the left and from the right

so there is no net sign change in this operation

Hence we have

$$\hat{G} = \frac{1}{N}\sum_{i=1}^{N}\sum_{a,b,\ldots,i1,i2,\ldots n} G_{i1i2}\underbrace{\hat{b}_a^\dagger\hat{b}_b^\dagger\ldots\hat{b}_n^\dagger}_{\text{omitting } \hat{b}_{i1}^\dagger}\hat{b}_{i1}^\dagger\hat{b}_{i2}\underbrace{\hat{b}_n\ldots\hat{b}_b\hat{b}_a}_{\text{omitting } \hat{b}_{i2}}$$

General single-particle fermion operators

In practice with any operator

- in the end we are working out its matrix elements
  - Any two operators with identical matrix elements are equivalent operators

We consider two, possibly different, *N*-fermion basis states

- $|\psi_{1N}\rangle$ and $|\psi_{2N}\rangle$
- and consider matrix elements of the operator $\hat{G}$ in

$$\hat{G} = \frac{1}{N}\sum_{i=1}^{N} \sum_{a,b,\ldots,i1,i2,\ldots n} G_{i1i2} \underbrace{\hat{b}_a^\dagger \hat{b}_b^\dagger \ldots \hat{b}_n^\dagger}_{omitting\ \hat{b}_{i1}^\dagger} \hat{b}_{i1}^\dagger \hat{b}_{i2} \underbrace{\hat{b}_n \ldots \hat{b}_b \hat{b}_a}_{omitting\ \hat{b}_{i2}}$$

  - between such states

General single-particle fermion operators

Because of Pauli exclusion

the only strings of operators that can survive in matrix elements for legal fermion states

are those in which the operators $\hat{b}_a, \hat{b}_b, \ldots \hat{b}_n$

are all different from each other

i.e., correspond to annihilation operators for different single particle states

and are each different from both $\hat{b}_{i1}$ and $\hat{b}_{i2}$

since otherwise we would be trying either to

annihilate two fermions from the same state or

create two fermions in the same state

General single-particle fermion operators

Hence, for these states

since no two states in the string of creation operators or in the string of annihilation operators can be identical

not only do the pairs of annihilation operators anticommute and

the pairs of creation operators anticommute as usual

so also do all the pairs of creation and annihilation operators with different subscripts

other than possibly the pair $\hat{b}_{i1}^{\dagger}\hat{b}_{i2}$

General single-particle fermion operators

Hence in $$\hat{G} = \frac{1}{N} \sum_{i=1}^{N} \sum_{a,b,\ldots,i1,i2,\ldots n} G_{i1i2} \underbrace{\hat{b}_a^\dagger \hat{b}_b^\dagger \ldots \hat{b}_n^\dagger}_{omitting\ \hat{b}_{i1}^\dagger} \hat{b}_{i1}^\dagger \hat{b}_{i2} \underbrace{\hat{b}_n \ldots \hat{b}_b \hat{b}_a}_{omitting\ \hat{b}_{i2}}$$

we can swap the creation operator $\hat{b}_a^\dagger$

all the way from the left

until we get to the left of the corresponding annihilation operator $\hat{b}_a$

only acquiring minus signs as we do so

General single-particle fermion operators

Actually, we acquire an even number of minus signs

- because the number of swaps taken to get to the middle
- is equal to
  - the number to get from the middle to its final position
    - so there is no change in sign in all these swaps

We can repeat this procedure for each creation operator

- other than $\hat{b}_{i1}^{\dagger}$
  - which we do not need to move anyway

General single-particle fermion operators

Hence, with all these swaps, we can rewrite

$$\hat{G} = \frac{1}{N}\sum_{i=1}^{N} \sum_{a,b,\dots,i1,i2,\dots n} G_{i1i2}\, \underbrace{\hat{b}_a^\dagger \hat{b}_b^\dagger \dots \hat{b}_n^\dagger}_{\text{omitting } \hat{b}_{i1}^\dagger} \hat{b}_{i1}^\dagger \hat{b}_{i2}\, \underbrace{\hat{b}_n \dots \hat{b}_b \hat{b}_a}_{\text{omitting } \hat{b}_{i2}}$$

as

$$\hat{G} = \frac{1}{N}\sum_{i=1}^{N} \sum_{a,b,\dots,i1,i2,\dots n} G_{i1i2} \hat{b}_{i1}^\dagger \hat{b}_{i2}\, \underbrace{\hat{b}_n^\dagger \hat{b}_n \dots\, \hat{b}_b^\dagger \hat{b}_b \hat{b}_a^\dagger \hat{b}_a}_{\text{omitting } \hat{b}_{i1}^\dagger \hat{b}_{i2}}$$

or more simply

$$\hat{G} = \frac{1}{N}\sum_{i=1}^{N} \sum_{a,b,\dots,i1,i2,\dots n} G_{i1i2} \hat{b}_{i1}^\dagger \hat{b}_{i2}\, \underbrace{\hat{N}_n \dots \hat{N}_b \hat{N}_a}_{\text{omitting } \hat{b}_{i1}^\dagger \hat{b}_{i2}}$$

General single-particle fermion operators

When this operator $\hat{G} = \frac{1}{N}\sum_{i=1}^{N} \sum_{a,b,\ldots,i1,i2,\ldots n} G_{i1i2}\hat{b}_{i1}^{\dagger}\hat{b}_{i2} \underbrace{\hat{N}_n \ldots \hat{N}_b \hat{N}_a}_{omitting\ \hat{b}_{i1}^{\dagger}\hat{b}_{i2}}$

operates on a specific N-fermion basis state $\left|\psi_{1N}\right\rangle$

the only terms in the summation that can survive

are those for which the list of states $a,b,\ldots n$

corresponds to occupied states in $\left|\psi_{1N}\right\rangle$

and so the sum over a, b, …, n (omitting $i1$ and $i2$)

and the number operators

can be dropped without changing any matrix element

General single-particle fermion operators

Hence we can write $\hat{G} = \frac{1}{N}\sum_{i=1}^{N}\sum_{i1,i2} G_{i1i2}\hat{b}_{i1}^{\dagger}\hat{b}_{i2}$

It makes no difference which fermion we are considering

G_{i1i2} is the same for every fermion

so the sum over i is trivial, and so

$$\hat{G} = \sum_{j,k} G_{jk}\hat{b}_{j}^{\dagger}\hat{b}_{k}$$

where we also further simplified notation by substituting j for $i1$ and k for $i2$

General single-particle fermion operators

This is the general form for a single-particle fermion operator

$$\boxed{\hat{G} = \sum_{j,k} G_{jk} \hat{b}_j^\dagger \hat{b}_k}$$

The Hamiltonian

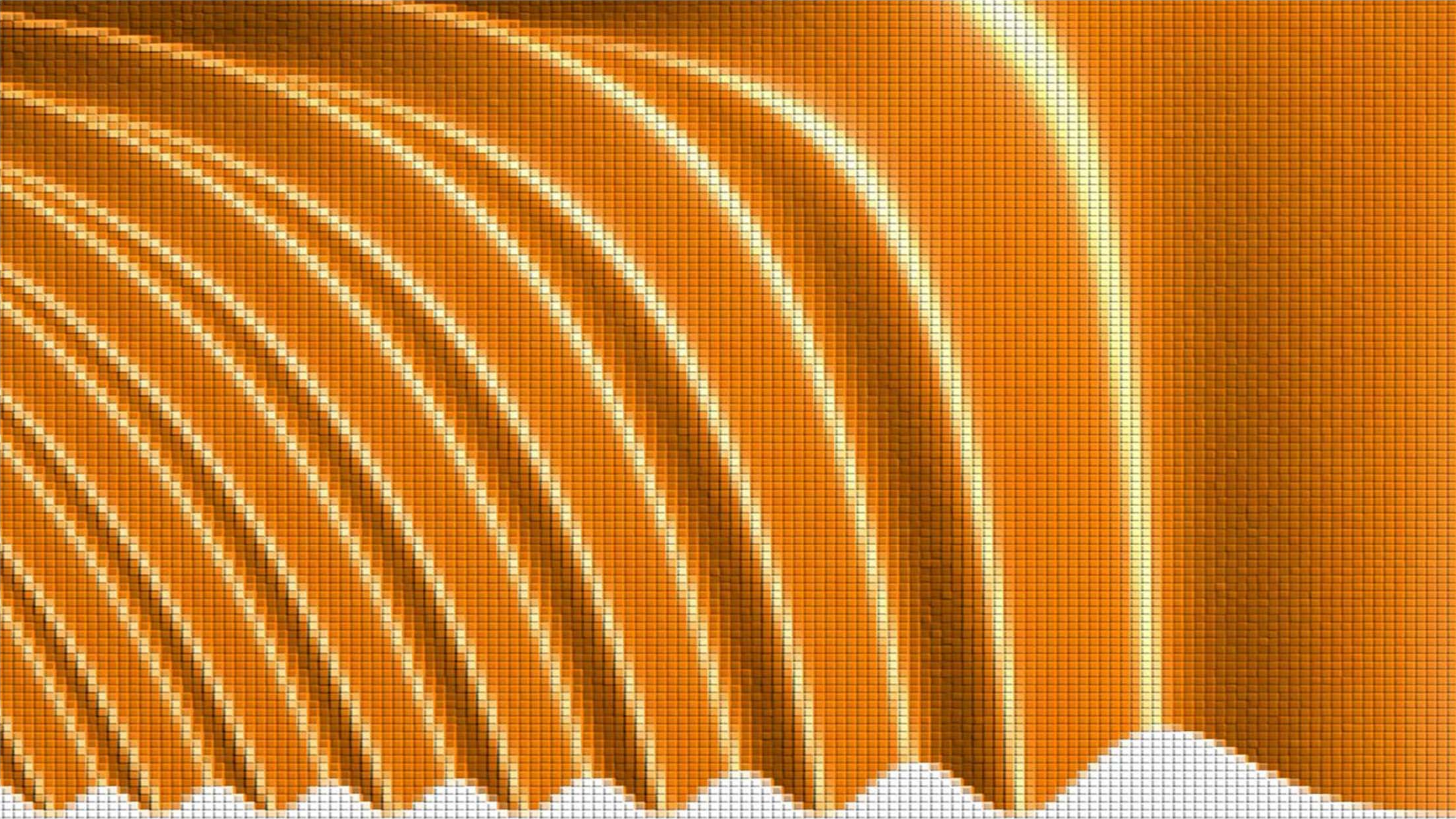
$$\hat{H} = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j$$

is just a special case for a diagonal operator

Hence we have found a very simple form

for the single-particle fermion operator

valid for any number of fermions

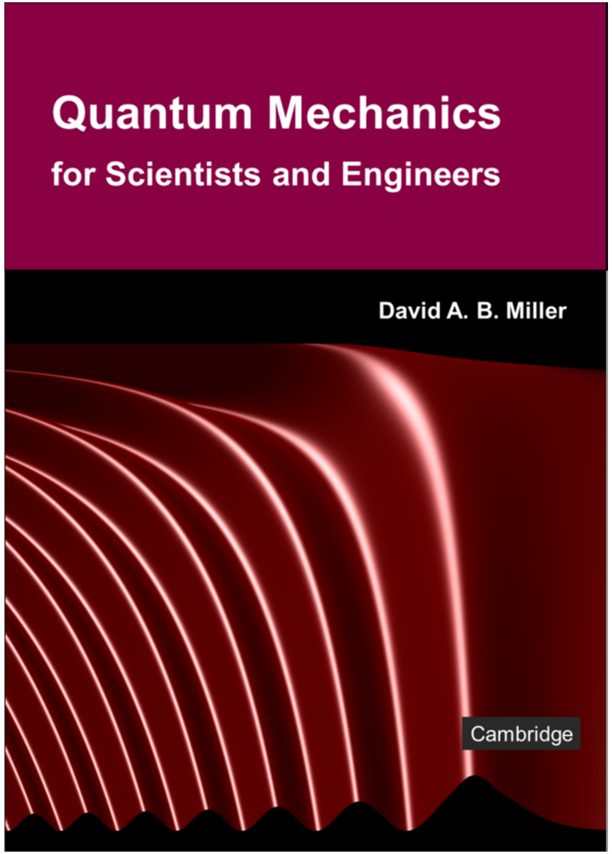


16.2 Fermion operators and multiple particles

Slides: Video 16.2.3 Two-particle fermion operators

Text reference: Quantum Mechanics for Scientists and Engineers

Section 16.3 subsection "Two-particle fermion operators"



Fermion operators and multiple particles

Two-particle fermion operators

Quantum mechanics for scientists and engineers

David Miller

Two-particle fermion operators

Fermions such as electrons interact

e.g., through their Coulomb repulsion

For such cases, we need two-particle operators

In the $\mathbf{r}$ form, we might have an operator $\hat{D}_{\mathbf{r}}\left(\mathbf{r}_1,\mathbf{r}_2\right)$

that depends on the coordinates of both particles

Then we postulate we can write

$$\hat{D}=\int \hat{\psi}^{\dagger}\left(\mathbf{r}_1,\mathbf{r}_2\right)\hat{D}_{\mathbf{r}}\left(\mathbf{r}_1,\mathbf{r}_2\right)\hat{\psi}\left(\mathbf{r}_1,\mathbf{r}_2\right)d^3\mathbf{r}_1 d^3\mathbf{r}_2$$

using the two-fermion wavefunction operator

$$\hat{\psi}\left(\mathbf{r}_1,\mathbf{r}_2\right)=\frac{1}{\sqrt{2}}\sum_{j,k}\hat{b}_k\hat{b}_j\phi_j\left(\mathbf{r}_1\right)\phi_k\left(\mathbf{r}_2\right)$$

Two-particle fermion operators

Substituting this two-particle wavefunction operator

into $\hat{D} = \int \hat{\psi}^\dagger(\mathbf{r}_1, \mathbf{r}_2) \hat{D}_\mathbf{r}(\mathbf{r}_1, \mathbf{r}_2) \hat{\psi}(\mathbf{r}_1, \mathbf{r}_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2$

we have

$$\hat{D} = \frac{1}{2} \sum_{a,b,c,d} \hat{b}_a^\dagger \hat{b}_b^\dagger \hat{b}_d \hat{b}_c \int \phi_a^*(\mathbf{r}_1) \phi_b^*(\mathbf{r}_2) \hat{D}_\mathbf{r}(\mathbf{r}_1, \mathbf{r}_2) \phi_c(\mathbf{r}_1) \phi_d(\mathbf{r}_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2$$

or equivalently

$$\boxed{\hat{D} = \frac{1}{2} \sum_{a,b,c,d} D_{abcd} \hat{b}_a^\dagger \hat{b}_b^\dagger \hat{b}_d \hat{b}_c}$$

where

$$D_{abcd} = \int \phi_a^*(\mathbf{r}_1) \phi_b^*(\mathbf{r}_2) \hat{D}_\mathbf{r}(\mathbf{r}_1, \mathbf{r}_2) \phi_c(\mathbf{r}_1) \phi_d(\mathbf{r}_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2$$

Order of suffixes in two-particle fermion operators

Note in

$$\hat{D} = \frac{1}{2}\sum_{a,b,c,d} D_{abcd}\hat{b}_a^\dagger \hat{b}_b^\dagger \hat{b}_d \hat{b}_c$$

the order of the suffixes on the chain of operators

$\hat{b}_a^\dagger \hat{b}_b^\dagger \hat{b}_d \hat{b}_c$ is not a, b, c, d

The ordering is in the opposite sense for the annihilation operators

This different ordering emerges

from the wavefunction operators

and the properties of Hermitian conjugation

Two-particle operators with multiple particles

We presume that the two-particle fermion operator

$$\hat{D} = \frac{1}{2} \sum_{a,b,c,d} D_{abcd} \hat{b}_a^\dagger \hat{b}_b^\dagger \hat{b}_d \hat{b}_c$$

would remain unchanged as we changed the system

to have more than two fermions in it

The arguments would be similar to those for the single-particle fermion operator $\hat{G} = \sum_{j,k} G_{jk} \hat{b}_j^\dagger \hat{b}_k$

So we presume this is a general statement

for a two-particle fermion operator

in this annihilation and creation operator approach

Electrons interacting through the Coulomb potential

For two electrons (of the same spin) with Coulomb repulsion

the Hamiltonian in the $\mathbf{r}$ form is

$$\hat{H}_{\mathbf{r}}\left(\mathbf{r}_1,\mathbf{r}_2\right)=-\frac{\hbar^2}{2m_o}\left(\nabla_{\mathbf{r}_1}^2+\nabla_{\mathbf{r}_2}^2\right)+\frac{e^2}{4\pi\varepsilon_o\left|\mathbf{r}_1-\mathbf{r}_2\right|}$$

Hence our two particle operator formalism gives us

$$\hat{H}=\frac{1}{2}\sum_{a,b,c,d}H_{abcd}\hat{b}_a^\dagger\hat{b}_b^\dagger\hat{b}_d\hat{b}_c$$

where H_{abcd} is defined analogously to

$$D_{abcd}=\int\phi_a^*\left(\mathbf{r}_1\right)\phi_b^*\left(\mathbf{r}_2\right)\hat{D}_{\mathbf{r}}\left(\mathbf{r}_1,\mathbf{r}_2\right)\phi_c\left(\mathbf{r}_1\right)\phi_d\left(\mathbf{r}_2\right)d^3\mathbf{r}_1d^3\mathbf{r}_2$$

Electrons interacting through the Coulomb potential

Suppose specifically we have the two-fermion state

where one electron is in the basis state $\phi_k(\mathbf{r})$

and the other is in the basis state $\phi_m(\mathbf{r})$

i.e., the two-particle state can be written

$$\left|\psi_{TP}\right\rangle = \hat{b}_k^\dagger \hat{b}_m^\dagger \left|0\right\rangle$$

We evaluate the expectation value of the energy

using the Hamiltonian $\hat{H} = \frac{1}{2}\sum_{a,b,c,d} H_{abcd}\hat{b}_a^\dagger \hat{b}_b^\dagger \hat{b}_d \hat{b}_c$

for this state, i.e.

$$\left\langle\psi_{TP}\right|\hat{H}\left|\psi_{TP}\right\rangle = \frac{1}{2}\left\langle 0\right| \sum_{a,b,c,d} H_{abcd}\hat{b}_m \hat{b}_k \hat{b}_a^\dagger \hat{b}_b^\dagger \hat{b}_d \hat{b}_c \hat{b}_k^\dagger \hat{b}_m^\dagger \left|0\right\rangle$$

Electrons interacting through the Coulomb potential

Now

$$\langle 0|\hat{b}_m\hat{b}_k\hat{b}_a^\dagger\hat{b}_b^\dagger\hat{b}_d\hat{b}_c\hat{b}_k^\dagger\hat{b}_m^\dagger|0\rangle =$$

$$\delta_{ak}\delta_{bm}\delta_{ck}\delta_{dm}+\delta_{am}\delta_{bk}\delta_{cm}\delta_{dk}-\delta_{am}\delta_{bk}\delta_{ck}\delta_{dm}-\delta_{ak}\delta_{bm}\delta_{cm}\delta_{dk}$$

the proof of which is left as an exercise

Hence we have for the energy expectation value

$$\langle\psi_{TP}|\hat{H}|\psi_{TP}\rangle=\frac{1}{2}\langle 0|\sum_{a,b,c,d}H_{abcd}\hat{b}_m\hat{b}_k\hat{b}_a^\dagger\hat{b}_b^\dagger\hat{b}_d\hat{b}_c\hat{b}_k^\dagger\hat{b}_m^\dagger|0\rangle$$

$$=\frac{1}{2}\left(H_{kmkm}+H_{mkmk}-H_{mkkm}-H_{kmmk}\right)$$

Electrons interacting through the Coulomb potential

In $\langle \psi_{TP} | \hat{H} | \psi_{TP} \rangle = \frac{1}{2}\left(H_{kmkm} + H_{mkmk} - H_{mkkm} - H_{kmmk} \right)$

explicitly, we have

$$H_{kmkm} = H_{mkmk} = \int \phi_k^*(\mathbf{r}_1)\phi_m^*(\mathbf{r}_2)\hat{H}_{\mathbf{r}}\phi_k(\mathbf{r}_1)\phi_m(\mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2$$

and

$$H_{kmmk} = H_{mkkm}^* = \int \phi_k^*(\mathbf{r}_1)\phi_m^*(\mathbf{r}_2)\hat{H}_{\mathbf{r}}\phi_m(\mathbf{r}_1)\phi_k(\mathbf{r}_2)d^3\mathbf{r}_1 d^3\mathbf{r}_2$$

These are exactly the same terms as previously calculated using the $\mathbf{r}$ formalism

Electrons interacting through the Coulomb potential

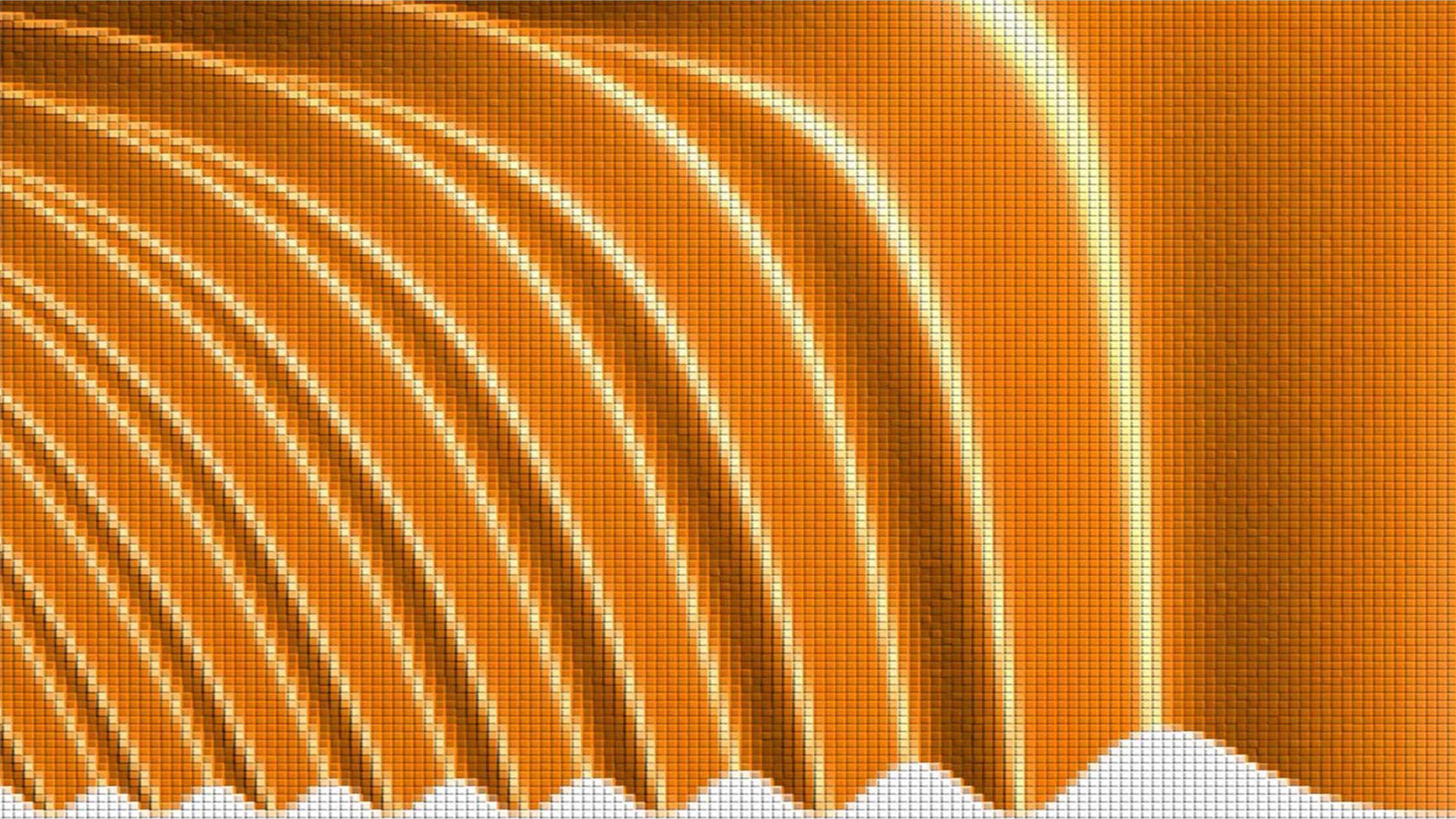
Remember in $\langle \psi_{TP} | \hat{H} | \psi_{TP} \rangle = \frac{1}{2}\left(H_{kmkm} + H_{mkmk} - H_{mkkm} - H_{kmmk} \right)$

H_{kmkm} or equivalently $(1/2)(H_{kmkm} + H_{mkmk})$ is the sum of the kinetic energies for the two particles and the Coulomb potential energy for two electrons

so it is the energy we would calculate if the particles were not identical

$-(1/2)(H_{mkkm} + H_{kmmk})$ is the exchange energy

Hence this approach does reproduce the results of our previous $\mathbf{r}$ formalism

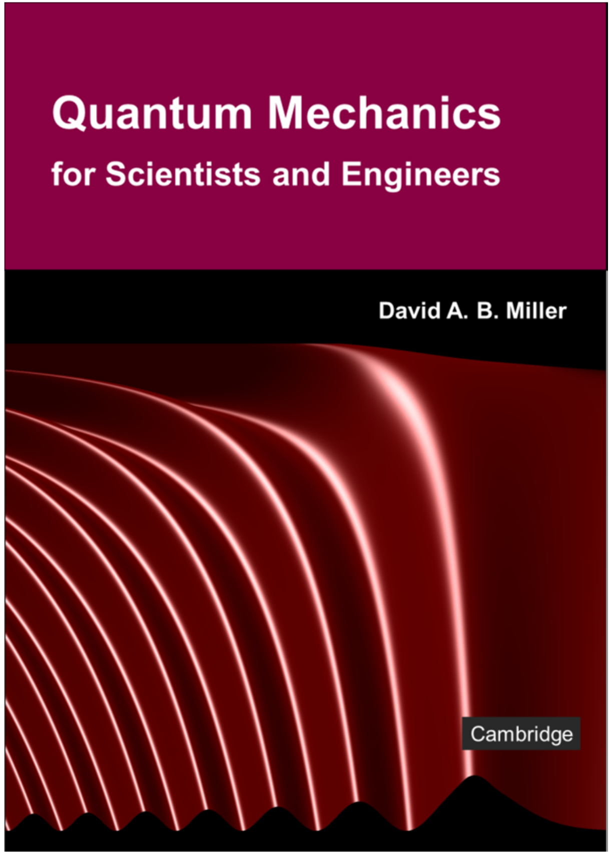


16.3 Interaction of different particles

Slides: Video 16.3.1 States with different kinds of particles

Text reference: Quantum Mechanics for Scientists and Engineers

Section 17.1



Interaction of different particles

States with different kinds of particles

Quantum mechanics for scientists and engineers

David Miller

States with different kinds of particles

For a state with

- one fermion in fermion state k, and one in state q
- one photon in photon mode λ_d and three in photon mode λ_s

we could write the state in a list form

- or alternatively using creation operators acting on the empty state as

$$\left|\ldots,0_j,1_k,0_l,\ldots 0_p,1_q,0_r,\ldots;\ \ldots,0_{\lambda c},1_{\lambda d},0_{\lambda e},\ldots,0_{\lambda r},3_{\lambda s},0_{\lambda t},\ldots\right\rangle$$

$$\equiv\left|N_{fm};N_{bn}\right\rangle\equiv\frac{1}{\sqrt{3!}}\hat{b}_k^\dagger\hat{b}_q^\dagger\hat{a}_{\lambda d}^\dagger\left(\hat{a}_{\lambda s}^\dagger\right)^3\left|0\right\rangle$$

States with different kinds of particles

In $\left|\ldots,0_j,1_k,0_l,\ldots 0_p,1_q,0_r,\ldots;\ \ldots,0_{\lambda c},1_{\lambda d},0_{\lambda e},\ldots,0_{\lambda r},3_{\lambda s},0_{\lambda t},\ldots\right\rangle$

$$\equiv\left|N_{fm};N_{bn}\right\rangle\equiv\frac{1}{\sqrt{3!}}\hat{b}_k^{\dagger}\hat{b}_q^{\dagger}\hat{a}_{\lambda d}^{\dagger}\left(\hat{a}_{\lambda s}^{\dagger}\right)^3\left|0\right\rangle$$

N_{fm} is the mth possible list of occupied fermion states

here the list $\ldots,0_j,1_k,0_l,\ldots 0_p,1_q,0_r,\ldots$

and similarly N_{bn} is the nth possible list of occupied boson states

here the list $\ldots,0_{\lambda c},1_{\lambda d},0_{\lambda e},\ldots,0_{\lambda r},3_{\lambda s},0_{\lambda t},\ldots$

Note now that the empty state $\left|0\right\rangle$ is one that is empty both of this kind of fermion and this kind of boson

Commutation relations for different particles

We postulate that creation and annihilation operators for different particles commute under all conditions

Specifically for the boson and fermion operators we would have

$$\hat{b}_j^\dagger \hat{a}_\lambda^\dagger - \hat{a}_\lambda^\dagger \hat{b}_j^\dagger = 0 \qquad \hat{b}_j \hat{a}_\lambda - \hat{a}_\lambda \hat{b}_j = 0$$

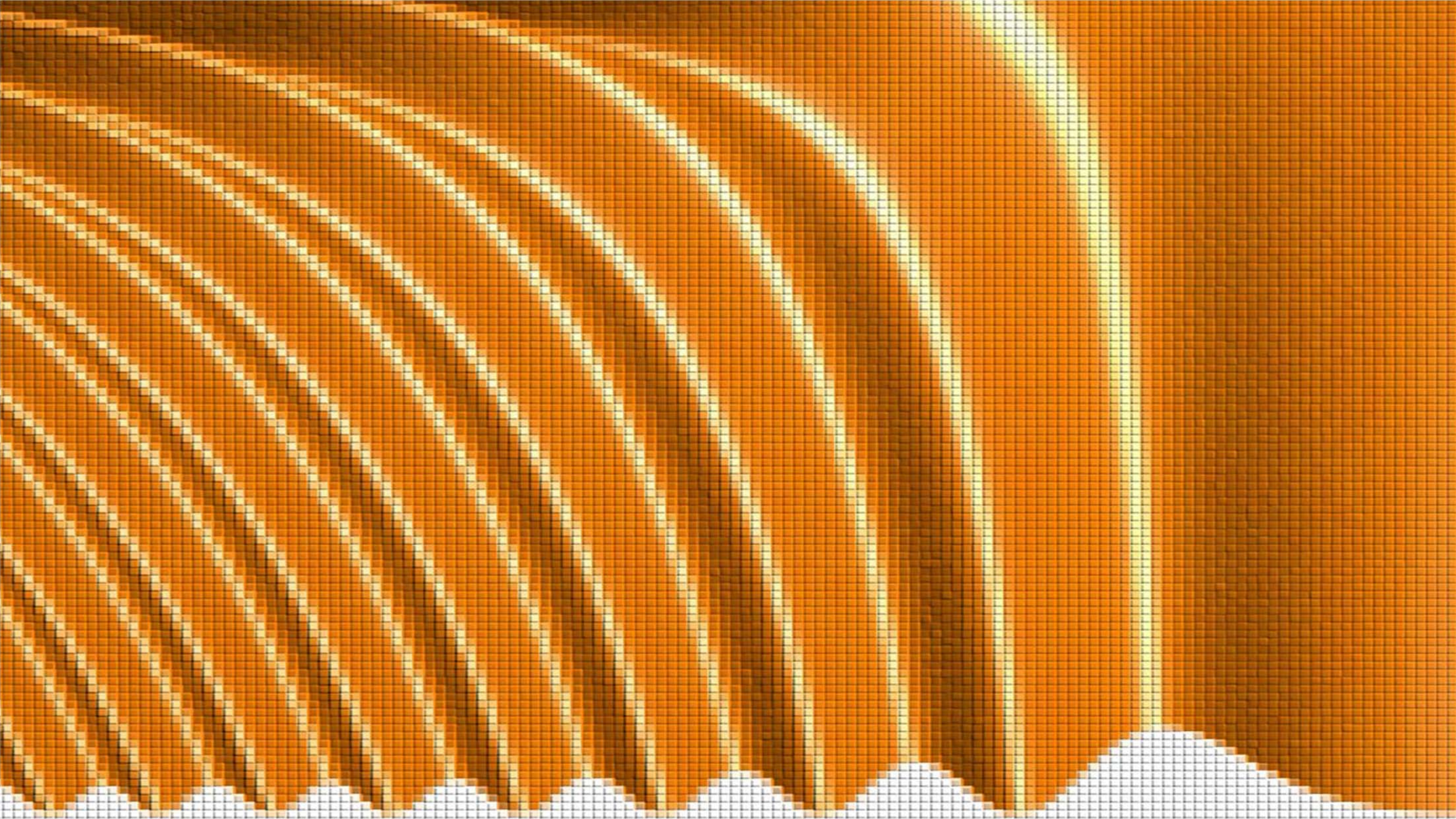
$$\hat{b}_j^\dagger \hat{a}_\lambda - \hat{a}_\lambda \hat{b}_j^\dagger = 0 \qquad \hat{b}_j \hat{a}_\lambda^\dagger - \hat{a}_\lambda^\dagger \hat{b}_j = 0$$

Note similar relations also would hold

for annihilation and creation operators

corresponding to two different kinds of fermions

such as electrons and protons

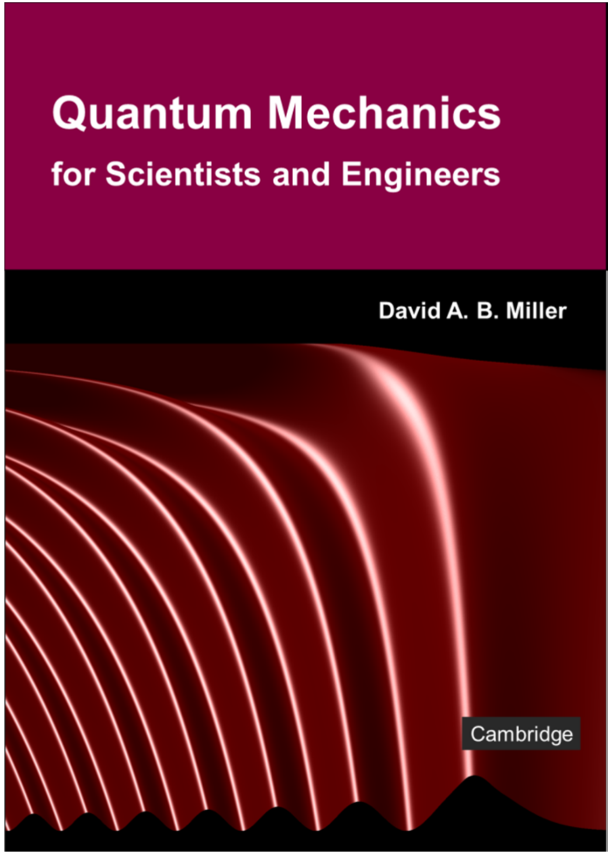


16.3 Interaction of different particles

Slides: Video 16.3.3 Electron – photon interaction

Text reference: Quantum Mechanics for Scientists and Engineers

Section 17.2



Interaction of different particles

Electron – photon interaction

Quantum mechanics for scientists and engineers

David Miller

Electron-photon interaction

Suppose first we had mathematically "turned off"

any interaction between electrons and photons

Because there is no interaction for the moment

the resulting Hamiltonian is the sum of the separate fermion (electron) and boson (photon) Hamiltonians

$$\hat{H}_o = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j + \sum_\lambda \hbar\omega_\lambda \hat{a}_\lambda^\dagger \hat{a}_\lambda$$

The sum over j is over all possible single-particle fermion states (modes)

The sum over λ is over all possible photon modes

Electric dipole interaction – boson operators

Previously, for the electric dipole interaction

we had, from a semiclassical view of the energy of an electron at position $\mathbf{r}_i$ in an electric field $\mathbf{E}$

$$\hat{H}_{scedr} = e\,\mathbf{E}\cdot\mathbf{r}$$

Substituting the multimode electric field operator $\hat{\mathbf{E}}$

for the classical field $\mathbf{E}$ gives

for any specific electron i

$$\hat{H}_{ed\mathbf{r}i} = \sqrt{-1}\,e\sum_{\lambda}\left(\hat{a}_{\lambda} - \hat{a}_{\lambda}^{\dagger}\right)\sqrt{\frac{\hbar\omega_{\lambda}}{2\varepsilon_o}}\mathbf{u}_{\lambda}\left(\mathbf{r}_i\right)\cdot\mathbf{r}_i$$

Electric dipole interaction – fermion operators

For N electrons

- if we stayed in our previous $\mathbf{r}$ form
  - we would have to add all these Hamiltonians

$$\hat{H}_{ed\mathbf{r}} = \sum_{i=1}^{N} \sqrt{-1}\, e \sum_{\lambda} \left(\hat{a}_{\lambda} - \hat{a}_{\lambda}^{\dagger} \right) \sqrt{\frac{\hbar \omega_{\lambda}}{2\varepsilon_{o}}} \mathbf{u}_{\lambda}\left(\mathbf{r}_{i}\right) \cdot \mathbf{r}_{i}$$

Now we want to transform this Hamiltonian in $\mathbf{r}$ form

- into the fermion occupation number form also
  - To do so, we formally use the N-fermion wavefunction operators

Electric dipole interaction – fermion operators

Because the fermion and boson operators commute with one another

the boson operators also commute with the (fermion) wavefunction operators, and so we can write

$$\hat{H}_{ed} = \int \hat{\psi}^{\dagger} \hat{H}_{ed\mathbf{r}} \hat{\psi}\, d^3\mathbf{r}_1 d^3\mathbf{r}_2 \ldots d^3\mathbf{r}_N$$

$$= \int \hat{\psi}^{\dagger} \left[\sum_{i=1}^{N} \sqrt{-1}\, e \sum_{\lambda} \left(\hat{a}_{\lambda} - \hat{a}_{\lambda}^{\dagger} \right) \sqrt{\frac{\hbar\omega_{\lambda}}{2\varepsilon_o}} \mathbf{u}_{\lambda}\left(\mathbf{r}_i\right) \cdot \mathbf{r}_i \right] \hat{\psi}\; d^3\mathbf{r}_1 d^3\mathbf{r}_2 \ldots d^3\mathbf{r}_N$$

$$= \sum_{\lambda} \left(\hat{a}_{\lambda} - \hat{a}_{\lambda}^{\dagger} \right) \int \hat{\psi}^{\dagger} \left[\sum_{i=1}^{N} \sqrt{-1}\, e \sqrt{\frac{\hbar\omega_{\lambda}}{2\varepsilon_o}} \mathbf{u}_{\lambda}\left(\mathbf{r}_i\right) \cdot \mathbf{r}_i \right] \hat{\psi}\; d^3\mathbf{r}_1 d^3\mathbf{r}_2 \ldots d^3\mathbf{r}_N$$

Rewriting the single-particle fermion operator

Note first that this operator

$$\hat{H}_{ed} = \sum_{\lambda} \left(\hat{a}_{\lambda} - \hat{a}_{\lambda}^{\dagger} \right) \int \hat{\psi}^{\dagger} \left[\sum_{i=1}^{N} \sqrt{-1}\, e \sqrt{\frac{\hbar \omega_{\lambda}}{2\varepsilon_{o}}} \mathbf{u}_{\lambda}\left(\mathbf{r}_{i} \right) \cdot \mathbf{r}_{i} \right] \hat{\psi}\; d^{3}\mathbf{r}_{1} d^{3}\mathbf{r}_{2} \ldots d^{3}\mathbf{r}_{N}$$

is a “single-particle” operator

- It is a sum of terms
  - each of which only depends on the coordinates of one particle
    - There are no parts to it that depend on relative separations of two particles, for example

Rewriting the single-particle fermion operator

Now we remember that, when we had such an operator

in the $\mathbf{r}$ form as a sum of single-particle operators for each of N fermions

$$\hat{G}_{\mathbf{r}} = \sum_{i=1}^{N} \hat{G}_{\mathbf{r}i}$$

then the result of rewriting it in creation and annihilation operator form was

$$\hat{G} = \int \hat{\psi}^{\dagger} \hat{G}_{\mathbf{r}} \hat{\psi} \, d^3\mathbf{r}_1 d^3\mathbf{r}_2 \ldots d^3\mathbf{r}_N \;\; = \sum_{j,k} G_{jk} \hat{b}_j^{\dagger} \hat{b}_k$$

with $G_{jk} = \int \phi_j^*(\mathbf{r}) \hat{G}_{\mathbf{r}} \phi_k(\mathbf{r}) d^3\mathbf{r}$

Rewriting the single-particle fermion operator

So our operator

$$\hat{H}_{ed} = \sum_{\lambda} \left(\hat{a}_{\lambda} - \hat{a}_{\lambda}^{\dagger} \right) \int \hat{\psi}^{\dagger} \left[\sum_{i=1}^{N} \sqrt{-1}\, e \sqrt{\frac{\hbar \omega_{\lambda}}{2\varepsilon_o}} \mathbf{u}_{\lambda}\left(\mathbf{r}_i\right) \cdot \mathbf{r}_i \right] \hat{\psi}\, d^3\mathbf{r}_1 d^3\mathbf{r}_2 \ldots d^3\mathbf{r}_N$$

similarly becomes

$$\boxed{\hat{H}_{ed} = \sum_{j,k,\lambda} H_{ed\lambda jk} \hat{b}_j^{\dagger} \hat{b}_k \left(\hat{a}_{\lambda} - \hat{a}_{\lambda}^{\dagger} \right)}$$

with $H_{ed\lambda jk} = \int \phi_j^{*}\left(\mathbf{r}_i\right) \hat{H}_{ed\lambda \mathbf{r}i} \phi_k\left(\mathbf{r}_i\right) d^3\mathbf{r}_i$

$$= \sqrt{-1}\, e \sqrt{\frac{\hbar \omega_{\lambda}}{2\varepsilon_o}} \int \phi_j^{*}\left(\mathbf{r}_i\right) \left[\mathbf{u}_{\lambda}\left(\mathbf{r}_i\right) \cdot \mathbf{r}_i \right] \phi_k\left(\mathbf{r}_i\right) d^3\mathbf{r}_i$$

Rewriting the single-particle fermion operator

In $\hat{H}_{ed} = \sum_{j,k,\lambda} H_{ed\lambda jk} \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right)$

with $H_{ed\lambda jk} = \sqrt{-1}\, e \sqrt{\frac{\hbar\omega_\lambda}{2\varepsilon_o}} \int \phi_j^*\left(\mathbf{r}_i\right)\left[\mathbf{u}_\lambda\left(\mathbf{r}_i\right)\cdot\mathbf{r}_i\right]\phi_k\left(\mathbf{r}_i\right) d^3\mathbf{r}_i$

all the details of the specific form of the single-particle fermion states and of the electromagnetic modes

are contained within the constants $H_{ed\lambda jk}$

Operators and processes

The annihilation and creation operators identify processes that could occur given appropriate starting states

We can open up the operator expression

$$\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) = \hat{b}_j^\dagger \hat{b}_k \hat{a}_\lambda - \hat{b}_j^\dagger \hat{b}_k \hat{a}_\lambda^\dagger$$

If (fermion) state k was occupied and state j was empty

and we had at least one photon in mode λ

then we could have a process corresponding to $\hat{b}_j^\dagger \hat{b}_k \hat{a}_\lambda$

that involves annihilating a photon in mode λ

and changing an electron from state k to state j

i.e., absorption of a photon to change the electron state

Operators and processes

Similarly, in

$$\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) = \hat{b}_j^\dagger \hat{b}_k \hat{a}_\lambda - \hat{b}_j^\dagger \hat{b}_k \hat{a}_\lambda^\dagger$$

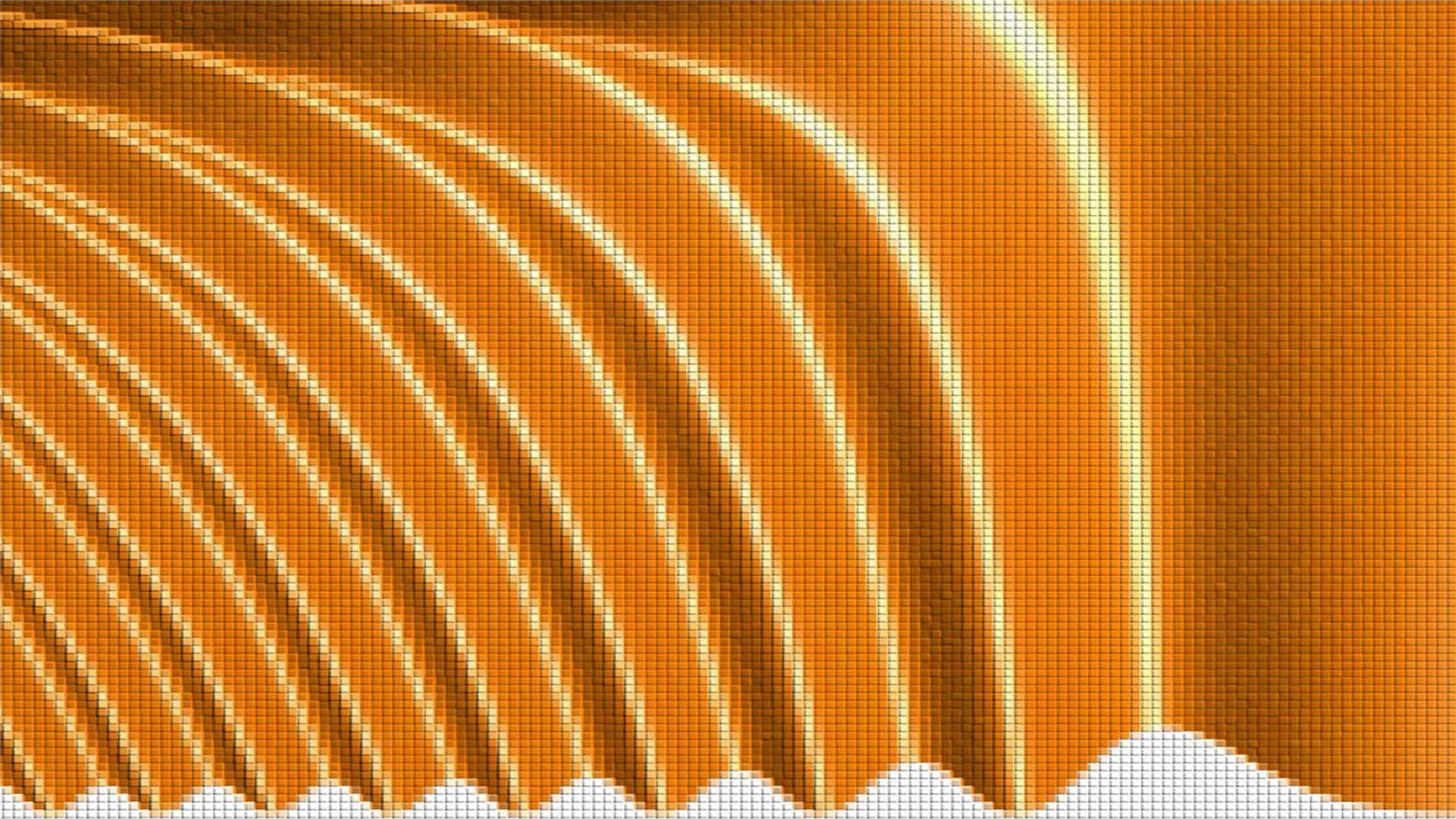
the process corresponding to the operators $\hat{b}_j^\dagger \hat{b}_k \hat{a}_\lambda^\dagger$

is one of emission of a photon

as an electron goes from state k to state j

We will evaluate transition rates for such processes

once we have discussed time-dependent perturbation theory for this formalism

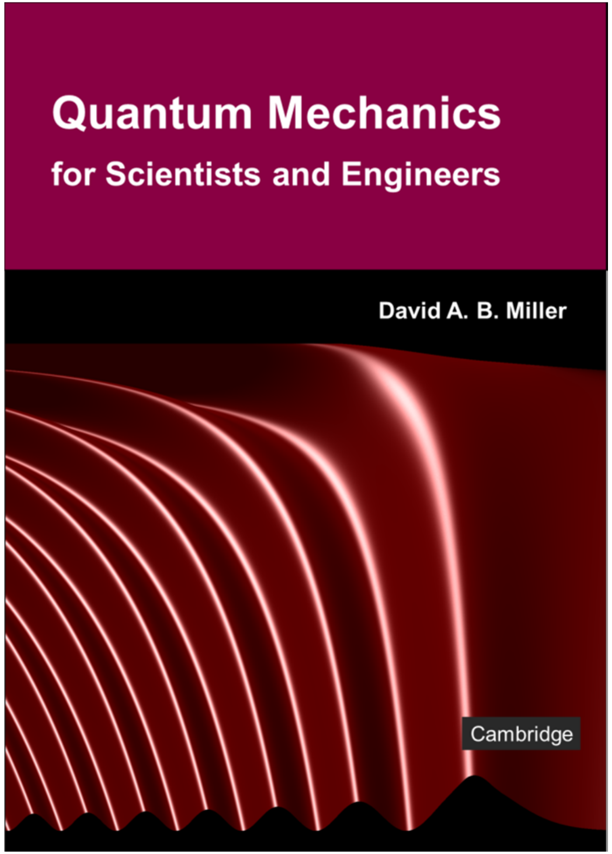


16.3 Interaction of different particles

Slides: Video 16.3.5 Rewriting perturbation theory

Text reference: Quantum Mechanics for Scientists and Engineers

Sections 17.3



Interaction of different particles

Rewriting perturbation theory

Quantum mechanics for scientists and engineers

David Miller

Time-dependent perturbation theory

We will have an unperturbed Hamiltonian, $\hat{H}_o$

- such as the one for non-interacting fermions and bosons

$$\hat{H}_o = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j + \sum_\lambda \hbar\omega_\lambda \hat{a}_\lambda^\dagger \hat{a}_\lambda$$

Then we will consider the interactions between particles

- such as the electric dipole interaction discussed before for electrons and photons
  - as a perturbation

$$\hat{H}_{ed} = \sum_{j,k,\lambda} H_{ed\lambda jk} \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right)$$

Perturbation theory with occupation number states

Now our basis states must describe both

the occupation of each single-particle electron state

and the occupation of each boson mode

Hence we write our basis states in the form used in a previous example

$$\left|\ldots,0_j,1_k,0_l,\ldots0_p,1_q,0_r,\ldots;\ \ldots,0_{\lambda c},1_{\lambda d},0_{\lambda e},\ldots,0_{\lambda r},3_{\lambda s},0_{\lambda t},\ldots\right\rangle$$

$$\equiv\left|N_{fm};N_{bn}\right\rangle\equiv\frac{1}{\sqrt{3!}}\hat{b}_k^\dagger\hat{b}_q^\dagger\hat{a}_{\lambda d}^\dagger\left(\hat{a}_{\lambda s}^\dagger\right)^3|0\rangle$$

Perturbation theory with occupation number states

Specifically, the mth state of this entire (non-interacting) fermion-boson system

- or the mth basis state of an interacting system
  - can be written as $\left|N_{fm};N_{bm}\right\rangle$
    - where N_{fm} is the list of all the occupation numbers
      - of each possible single-particle fermion state
    - and N_{bm} is similarly the list of all the occupation numbers
      - of each possible boson mode

Perturbation theory with occupation number states

These states will be the eigenstates of the unperturbed Hamiltonian

which we take as $\hat{H}_o = \sum_j E_j \hat{b}_j^\dagger \hat{b}_j + \sum_\lambda \hbar\omega_\lambda \hat{a}_\lambda^\dagger \hat{a}_\lambda$

Analogous to the zero-order perturbation theory equation $\hat{H}_o \left|\psi_n\right\rangle = E_n \left|\psi_n\right\rangle$

we now write this in the form $\hat{H}_o \left|N_{fm}; N_{bm}\right\rangle = E_m \left|N_{fm}; N_{bm}\right\rangle$

where E_m would be the energy of this fermion-boson system in state m

in the absence of any interaction between the fermions and bosons

Perturbation theory with occupation number states

The actual system state is some linear superposition $|\psi\rangle$

where we expand this state in our multiple fermion and multiple boson occupation number basis

i.e., analogous to the previous expansion

$$|\Psi\rangle = \sum_n a_n(t)\exp(-iE_n t/\hbar)|\psi_n\rangle$$

we have $|\psi\rangle = \sum_m c_m \exp(-iE_m t/\hbar)|N_{fm};N_{bm}\rangle$

explicitly including the time factors $\exp(-iE_m t/\hbar)$

so we can leave them out of the states $|N_{fm};N_{bm}\rangle$

Perturbation theory with occupation number states

Note again that, in

$$|\psi\rangle = \sum_m c_m \exp\left(-iE_m t / \hbar\right)\left|N_{fm}; N_{bm}\right\rangle$$

in contrast to previous approaches

which treated perturbations as external phenomena

E_m is the energy of the complete (unperturbed) fermion-boson system in this state

not merely the energy of the fermion

Perturbation theory with occupation number states

We now proceed exactly as before to set up time-dependent perturbation theory

with the only differences being that

the basis states $\left|N_{fm};N_{bm}\right\rangle$

are states of the entire system

and the (unperturbed) eigenvalues E_m

are those of the entire (unperturbed) system

here of fermions and bosons

Perturbation theory with occupation number states

So, as before, with $\left|\psi\right\rangle = \sum_m c_m \exp\left(-iE_m t/\hbar\right)\left|N_{fm}; N_{bm}\right\rangle$

in $i\hbar \frac{\partial}{\partial t}\left|\psi\right\rangle = \left(\hat{H}_o + \hat{H}_p\right)\left|\psi\right\rangle$

eliminating terms on both sides using

$$\hat{H}_o \left|N_{fm}; N_{bm}\right\rangle = E_m \left|N_{fm}; N_{bm}\right\rangle$$

and premultiplying by $\left\langle N_{fq}; N_{bq}\right|$ for state q of the fermion-boson system

we obtain, analogously to previous results

$$i\hbar \dot{c}_q \exp\left(-iE_q t/\hbar\right) = \sum_m c_m \exp\left(-iE_m t/\hbar\right)\left\langle N_{fq}; N_{bq}\right|\hat{H}_p\left|N_{fm}; N_{bm}\right\rangle$$

Perturbation theory with occupation number states

Taking the usual perturbation approach

of basing the first-order change in wavefunctions

on the zeroth-order state

i.e., on the unperturbed wavefunctions

we have

analogously to $\dot{a}_i^{(1)}(t) = \frac{1}{i\hbar}\sum_n a_n^{(0)} \exp(i\omega_{in} t)\langle \psi_i | \hat{H}_p(t) | \psi_n \rangle$

$$\dot{c}_q^{(1)} \simeq \frac{1}{i\hbar}\sum_m c_m^{(0)} \exp\left[-i\left(E_m - E_q\right)t/\hbar\right]\left\langle N_{fq}; N_{bq} \middle| \hat{H}_p \middle| N_{fm}; N_{bm} \right\rangle$$

Perturbation theory with occupation number states

In $\dot{c}_q^{(1)} \simeq \frac{1}{i\hbar} \sum_m c_m^{(0)} \exp\left[-i\left(E_m - E_q\right)t/\hbar\right] \left\langle N_{fq}; N_{bq} \middle| \hat{H}_p \middle| N_{fm}; N_{bm} \right\rangle$

as before, we typically presume the system

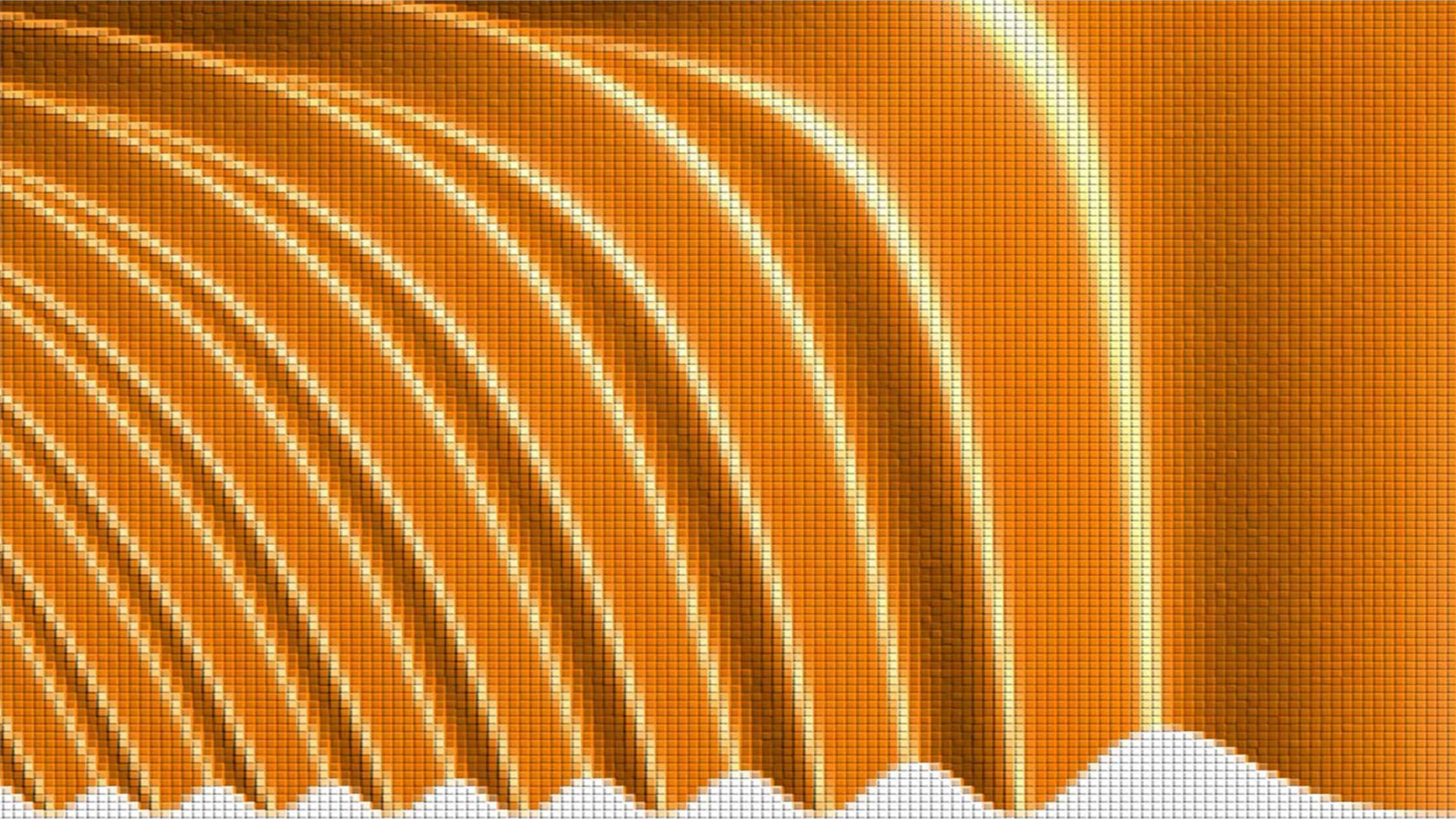
here an electron-photon one

starts in some specific basis state s of the unperturbed problem so that $c_s^{(0)} = 1$

and all other such coefficients are zero

so we have

$$\dot{c}_q^{(1)} \simeq \frac{1}{i\hbar} \exp\left[i\left(E_q - E_s\right)t/\hbar\right] \left\langle N_{fq}; N_{bq} \middle| \hat{H}_p \middle| N_{fs}; N_{bs} \right\rangle$$

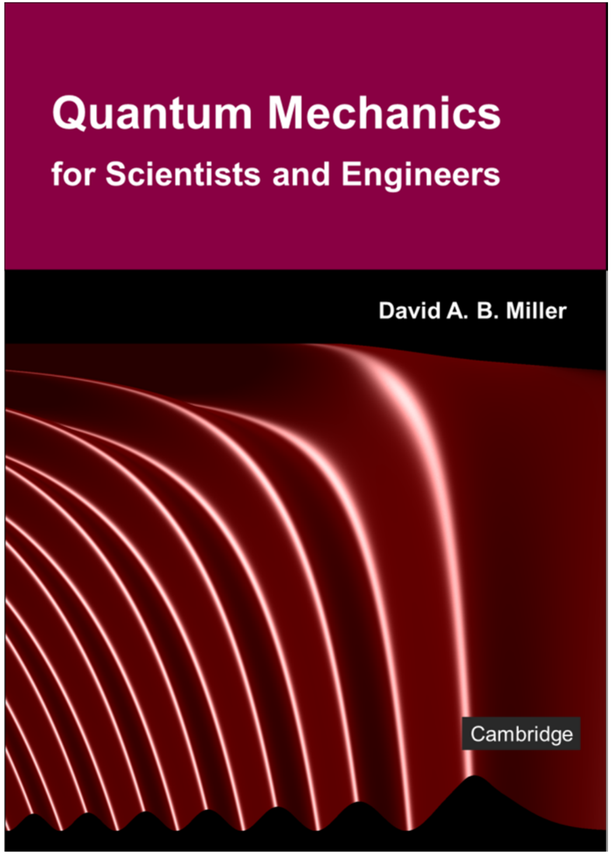


16.3 Interaction of different particles

Slides: Video 16.3.7 Photon absorption

Text reference: Quantum Mechanics for Scientists and Engineers

Section 17.4 through end of subsection "Absorption"



Interaction of different particles

Photon absorption

Quantum mechanics for scientists and engineers

David Miller

Photon emission and absorption

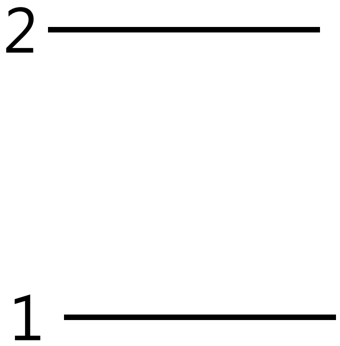
We take the electric dipole perturbing Hamiltonian

in annihilation and creation operator form

$$\hat{H}_p = \hat{H}_{ed} = \sum_{j,k,\lambda} H_{ed\lambda jk}\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right)$$

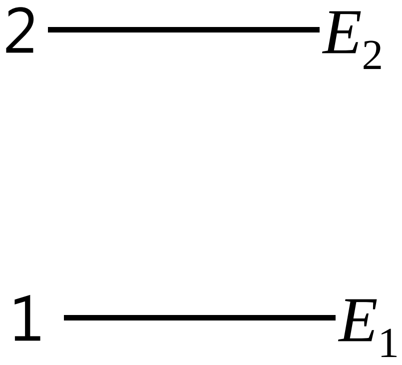
For simplicity, we presume we have only one electron

and that it has only two states of interest



Photon emission and absorption

State 1 – the lowest electron state
  with energy E_1
State 2 – the upper electron state
  with energy E_2
We consider the three possible processes of
  photon absorption
  spontaneous emission
  and stimulated emission


2 E_2

1 E_1

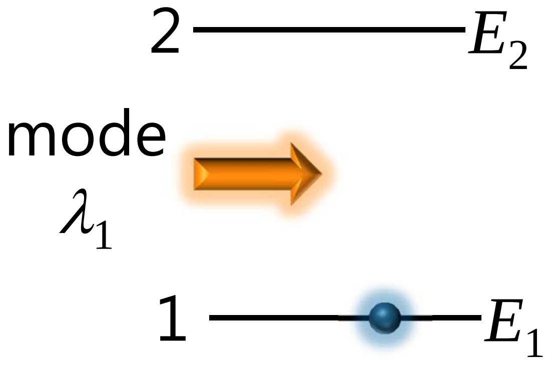
Photon absorption

Suppose

- the electron is initially in state 1
  - the lower state
- there is one photon in mode λ_1
  - and there are no photons in any other modes

Then we can write the initial state as

$$\left| N_{fs}; N_{bs} \right\rangle = \hat{b}_1^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle$$



Photon absorption

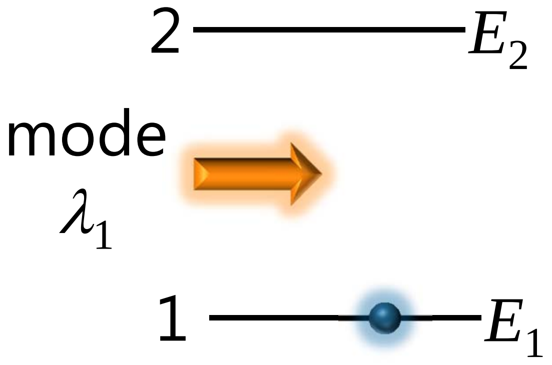
This state will have an energy

$$E_s = E_1 + \hbar\omega_{\lambda_1}$$

From now on

- for simplicity
  - we omit the zero point energy of the harmonic oscillator
    - i.e., the additional $\frac{1}{2}\hbar\omega_{\lambda_1}$ terms

This change merely corresponds to a choice of energy origin



Photon absorption

In our pertubation theory result

$$\dot{c}_q^{(1)} \simeq \frac{1}{i\hbar}\exp\left[i\left(E_q - E_s\right)t/\hbar\right]\left\langle N_{fq}; N_{bq}\left|\hat{H}_p\right|N_{fs}; N_{bs}\right\rangle$$

- as time evolves
  - to have any non-zero result for the amplitude c_q of state q
    - as required for any "transition" of the system to state q
      - we must have the matrix element $\left\langle N_{fq}; N_{bq}\left|\hat{H}_p\right|N_{fs}; N_{bs}\right\rangle$ be non-zero

Photon absorption

Now we have $\hat{H}_p \left| N_{fs}; N_{bs} \right\rangle = \sum_{j,k,\lambda} H_{ed\lambda jk} \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) \hat{b}_1^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle$

Examining the sequence of operators, we have

$$\begin{aligned}
\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) \hat{b}_1^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle &= \hat{b}_j^\dagger \hat{b}_k \hat{b}_1^\dagger \left(\hat{a}_\lambda \hat{a}_{\lambda_1}^\dagger - \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \right) \left| 0 \right\rangle \\
&= \hat{b}_j^\dagger \left(\delta_{k1} - \hat{b}_1^\dagger \hat{b}_k \right) \left(\delta_{\lambda\lambda_1} + \hat{a}_{\lambda_1}^\dagger \hat{a}_\lambda - \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \right) \left| 0 \right\rangle \\
&= \hat{b}_j^\dagger \delta_{k1} \left(\delta_{\lambda\lambda_1} - \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \right) \left| 0 \right\rangle \\
&= \delta_{k1} \delta_{\lambda\lambda_1} \hat{b}_j^\dagger \left| 0 \right\rangle - \delta_{k1} \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle
\end{aligned}$$

Photon absorption

With this result

$$\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \hat{b}_1^\dagger \hat{a}_{\lambda_1}^\dagger \left|0\right\rangle = \delta_{k1}\delta_{\lambda\lambda_1}\hat{b}_j^\dagger \left|0\right\rangle - \delta_{k1}\hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \left|0\right\rangle$$

in $\hat{H}_p \left|N_{fs}; N_{bs}\right\rangle = \sum_{j,k,\lambda} H_{ed\lambda jk} \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \hat{b}_1^\dagger \hat{a}_{\lambda_1}^\dagger \left|0\right\rangle$

only two possible choices for state q, i.e., state $\left|N_{fq}; N_{bq}\right\rangle$

give non-zero results

when we form $\left\langle N_{fq}; N_{bq}\right| \hat{H}_p \left|N_{fs}; N_{bs}\right\rangle$

Either, the first possibility, $\left|N_{fq}; N_{bq}\right\rangle = \hat{b}_j^\dagger \left|0\right\rangle$

or, the second possibility, $\left|N_{fq}; N_{bq}\right\rangle = \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \left|0\right\rangle$

Photon absorption

Consider the first possible "final" state $\left| N_{fq}; N_{bq} \right\rangle = \hat{b}_j^\dagger \left| 0 \right\rangle$

which is the state with

one electron in state j

and no photons in any modes

This state will have energy $E_q = E_j$

which leads to

$$\dot{c}_q^{(1)} \simeq \frac{1}{i\hbar} \exp\left[i\left(E_j - E_1 - \hbar\omega_{\lambda_1} \right) t / \hbar \right] \sum_{k,\lambda} H_{ed\lambda jk} \delta_{k1} \delta_{\lambda\lambda_1} \left\langle 0 \middle| \hat{b}_j \hat{b}_j^\dagger \middle| 0 \right\rangle$$

$$= \frac{1}{i\hbar} \exp\left[i\left(E_j - E_1 - \hbar\omega_{\lambda_1} \right) t / \hbar \right] H_{ed\lambda_1 j1}$$

Fermi's Golden Rule revisited

We can repeat the Fermi Golden Rule derivation

in the current notation

Now in $\dot{c}_q^{(1)} \simeq \frac{1}{i\hbar} \exp\left[i\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)t/\hbar\right] H_{ed\lambda_1 j1}$

we integrate over time

By definition, we choose $c_q^{(1)}\left(t=0\right)=0$

since we regard the system as starting in the specified initial state at $t = 0$

Hence integrating from $t = 0$ to t_o, we have

$$c_q^{(1)}\left(t_o\right) = -\frac{H_{ed\lambda_1 j1}}{E_j - E_1 - \hbar\omega_{\lambda_1}}\left\{\exp\left[i\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)t_o/\hbar\right]-1\right\}$$

Fermi's Golden Rule revisited

I.e., $c_q^{(1)}(t_o) = -\dfrac{H_{ed\lambda_1 j1}}{E_j - E_1 - \hbar\omega_{\lambda_1}}\left\{\exp\left[i\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)t_o/\hbar\right] - 1\right\}$

$$= -2iH_{ed\lambda_1 j1}\exp\left[i\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)t_o/2\hbar\right]\frac{\sin\left[\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)t_o/2\hbar\right]}{E_j - E_1 - \hbar\omega_{\lambda_1}}$$

So

$$\left|c_q^{(1)}(t_o)\right|^2 = 4\left|H_{ed\lambda_1 j1}\right|^2\frac{\sin^2\left[\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)t_o/2\hbar\right]}{\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)^2}$$

$$= \frac{2\pi}{\hbar}t_o\left|H_{ed\lambda_1 j1}\right|^2\left\{\frac{1}{t_o}\frac{2\hbar}{\pi}\frac{\sin^2\left[\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)t_o/2\hbar\right]}{\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)^2}\right\}$$

Fermi’s Golden Rule revisited

Now the function $$\left\{ \frac{1}{t_o} \frac{2\hbar}{\pi} \frac{\sin^2\left[\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)t_o / 2\hbar\right]}{\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)^2} \right\}$$

is a sharply peaked function near $E_j - E_1 - \hbar\omega_{\lambda_1} = 0$

and it has unit area when integrated over this energy argument

(note that $\int_{-\infty}^{\infty}\left[\left(\sin^2 x\right)/x^2\right]dx = \pi$)

Hence in the limit of large t_o

it can be replaced by a delta function

Fermi’s Golden Rule revisited

So $\left|c_q^{(1)}\left(t_o\right)\right|^2 = \frac{2\pi}{\hbar} t_o \left|H_{ed\lambda_1 j1}\right|^2 \left\{ \frac{1}{t_o} \frac{2\hbar}{\pi} \frac{\sin^2\left[\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right) t_o / 2\hbar\right]}{\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)^2} \right\}$

becomes $\left|c_q^{(1)}\left(t_o\right)\right|^2 = \frac{2\pi}{\hbar} t_o \left|H_{ed\lambda_1 j1}\right|^2 \delta\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)$

which gives a steadily rising occupation probability for this state q

Hence the transition rate is

$$w_q \simeq \frac{2\pi}{\hbar} \left|H_{ed\lambda_1 j1}\right|^2 \delta\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)$$

Fermi’s Golden rule to select final states

Now, for $j = 1$, the δ-function

$$\delta\left(E_j - E_1 - \hbar\omega_{\lambda_1}\right)$$

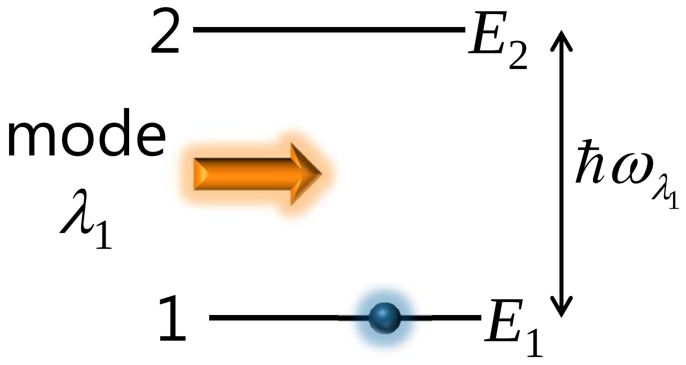
vanishes for any finite $\hbar\omega_{\lambda_1}$

So the only final state q that will give a transition rate

is the state $j = 2$

with the corresponding restriction that

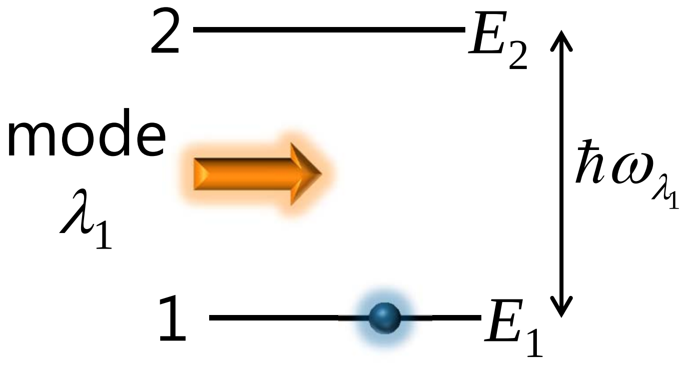
$$E_2 - E_1 = \hbar\omega_{\lambda_1}$$



Absorption process

Hence our process is as follows

- We start with
  - one photon in mode λ_1 and the electron in state 1



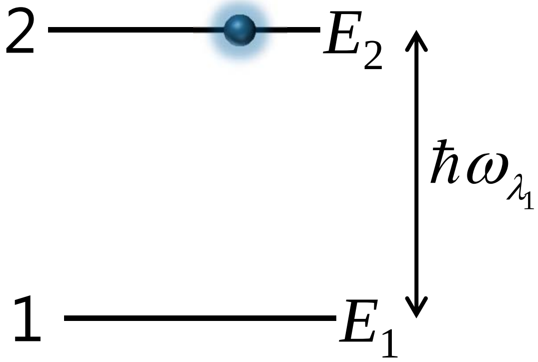
Absorption process

Hence our process is as follows

- We start with
  - one photon in mode λ_1 and the electron in state 1
- We finish with
  - no photons and the electron in state 2

This is a normal absorption process

- now requiring the destruction of the photon in the process



Fermi's Golden rule to select final states

From our operators, the other final possibility would be

$$\left|N_{fq};N_{bq}\right\rangle=\hat{b}_j^\dagger\hat{a}_\lambda^\dagger\hat{a}_{\lambda_1}^\dagger\left|0\right\rangle$$

with a corresponding energy $E_q=E_j+\hbar\omega_\lambda+\hbar\omega_{\lambda_1}$

This term would actually correspond to photon emission

We now have a photon in mode λ

With starting energy $E_s=E_1+\hbar\omega_{\lambda_1}$ as before

the δ-function argument in Fermi's Golden Rule is

$$E_q-E_s=E_j-E_1+\hbar\omega_\lambda$$

But this cannot be close to zero because $E_j-E_1\geq 0$

and $\hbar\omega_\lambda$ is also positive

Fermi's Golden rule to select final states

Hence on integrating over time as above

- this term associated with final state $\left| N_{fq}; N_{bq} \right\rangle = \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle$
  - will not give rise to any steady transition rate

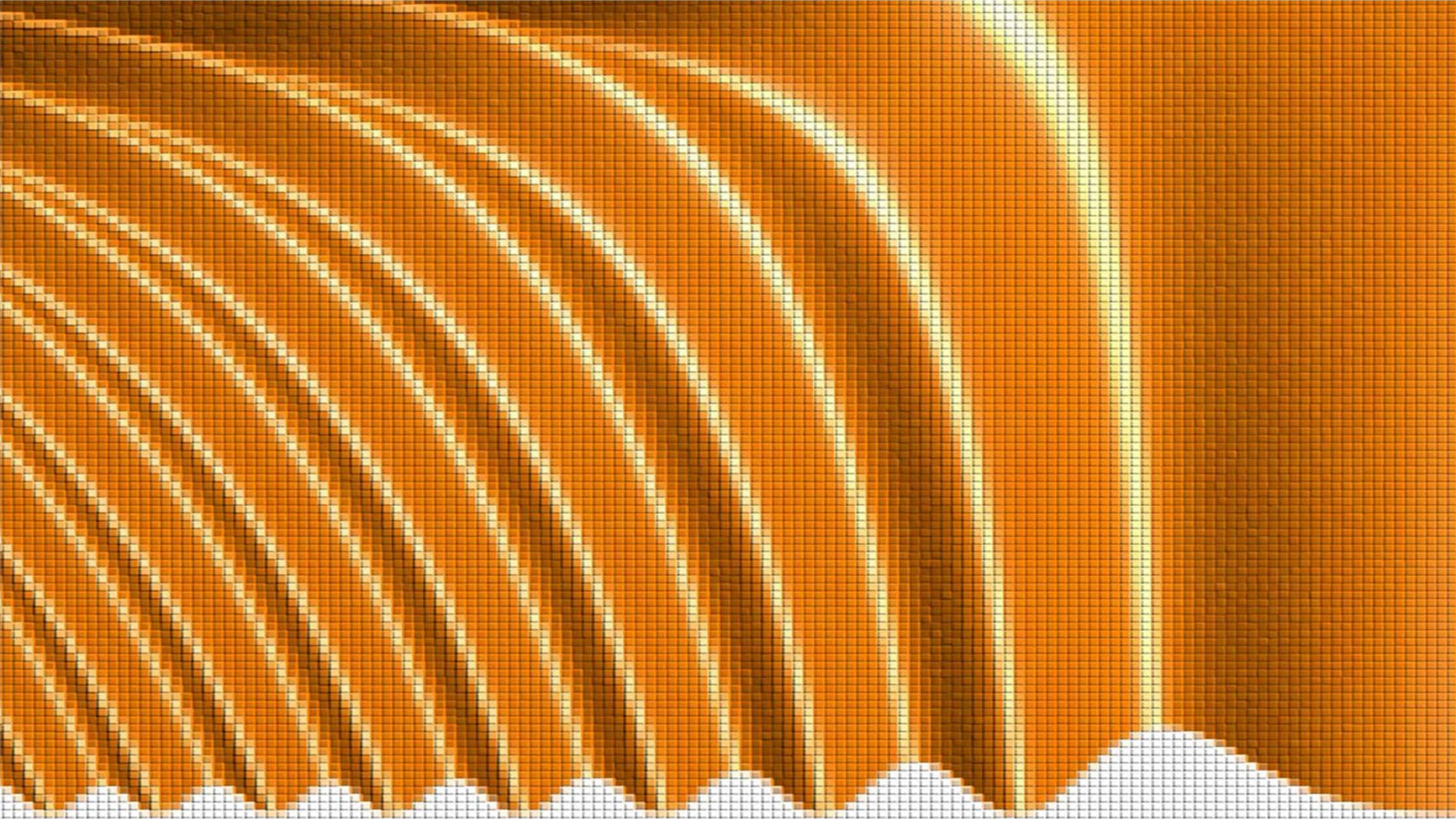
Hence this possibility can be discarded here

Simply put, we cannot emit a photon here

- because there is no lower energy for the electron to go to

Though this might seem obvious

- we have formally derived this conclusion here
  - not merely presumed it is obvious

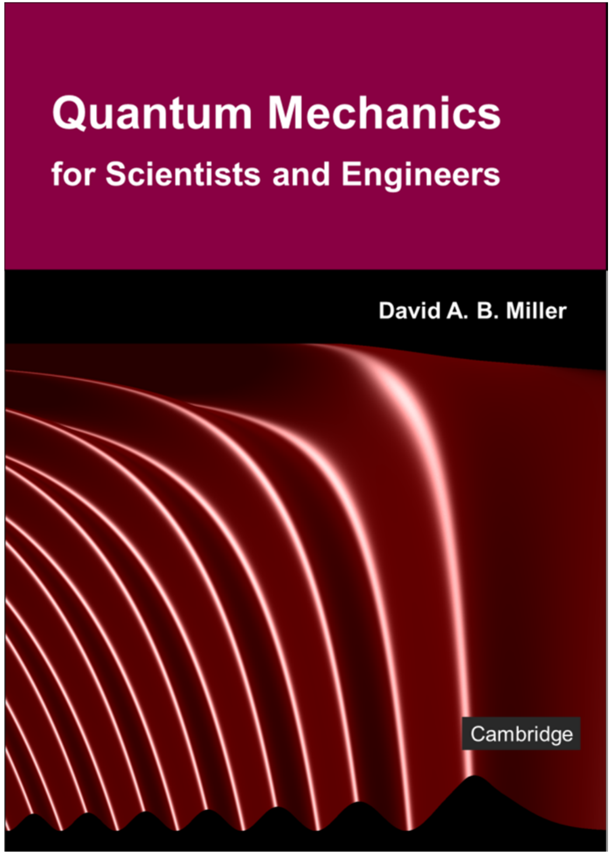


17.1 Spontaneous and stimulated emission

Slides: Video 17.1.1 Spontaneous emission

Text reference: Quantum Mechanics for Scientists and Engineers

Section 17.4 subsection "Spontaneous emission"



Spontaneous and stimulated emission

Spontaneous emission

Quantum mechanics for scientists and engineers

David Miller

Spontaneous emission

Suppose

- the electron is initially in state 2
  - the upper state
- there are no photons in any mode

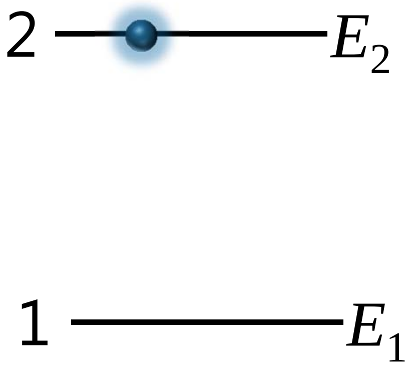
Now we can write the initial state as

$$\left| N_{fs}; N_{bs} \right\rangle = \hat{b}_2^\dagger \left| 0 \right\rangle$$

- with energy $E_s = E_2$

Note semiclassically with no electromagnetic field

- there would be no transitions



Spontaneous emission

Forming $\hat{H}_p \left| N_{fs}; N_{bs} \right\rangle$ with $\hat{H}_p = \hat{H}_{ed} = \sum_{j,k,\lambda} H_{ed\lambda jk} \hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right)$

with our starting state $\left| N_{fs}; N_{bs} \right\rangle = \hat{b}_2^\dagger \left| 0 \right\rangle$

we encounter the string of operators

$$\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) \hat{b}_2^\dagger \left| 0 \right\rangle = \hat{b}_j^\dagger \hat{b}_k \hat{b}_2^\dagger \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) \left| 0 \right\rangle$$

$$= \hat{b}_j^\dagger \left(\delta_{2k} - \hat{b}_2^\dagger \hat{b}_k \right) \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger \right) \left| 0 \right\rangle$$

$$= -\delta_{2k} \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \left| 0 \right\rangle$$

Spontaneous emission

With this string of operators $\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \hat{b}_2^\dagger \left|0\right\rangle = -\delta_{2k} \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \left|0\right\rangle$

from $\hat{H}_p \left|N_{fs}; N_{bs}\right\rangle$

to get a non-zero result for $\left\langle N_{fq}; N_{bq}\right| \hat{H}_p \left|N_{fs}; N_{bs}\right\rangle$

we must therefore choose for state q

$$\left|N_{fq}; N_{bq}\right\rangle = \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \left|0\right\rangle$$

which is the state with

the electron now in state j

and a photon in mode λ

This state q has energy $E_q = E_j + \hbar\omega_\lambda$

Spontaneous emission

We put this state q

which has some specific choice of mode λ

and electron state j

into our general perturbation theory expression

$$\dot{c}_q^{(1)} \simeq \frac{1}{i\hbar}\sum_m c_m^{(0)} \exp\left[-i\left(E_m - E_q\right)t/\hbar\right]\left\langle N_{fq};N_{bq}\left|\hat{H}_p\right|N_{fm};N_{bm}\right\rangle$$

So $\dot{c}_q^{(1)} \simeq \frac{1}{i\hbar}\exp\left[i\left(E_j - E_2 + \hbar\omega_\lambda\right)t/\hbar\right]\sum_k H_{ed\lambda jk}\delta_{k2}\left\langle 0\left|\hat{a}_\lambda \hat{b}_j \hat{a}_j^\dagger \hat{b}_\lambda^\dagger\right|0\right\rangle$

$$= \frac{1}{i\hbar}\exp\left[i\left(E_j - E_2 + \hbar\omega_\lambda\right)t/\hbar\right]H_{ed\lambda j2}$$

Spontaneous emission

Integrating $\dot{c}_q^{(1)} = \frac{1}{i\hbar} \exp\left[i\left(E_j - E_2 + \hbar\omega_\lambda \right) t / \hbar \right] H_{ed\lambda j2}$

and taking $\left| c_q^{(1)} \right|^2$ to get the transition rate, gives

$$w_q = \left(2\pi / \hbar \right) \left| H_{ed\lambda j2} \right|^2 \delta\left(E_j - E_2 + \hbar\omega_\lambda \right)$$

As before, for any finite $\hbar\omega_\lambda$

the only possible choice for the final state is $j = 1$

if there is to be any transition rate

with the requirement $E_2 - E_1 = \hbar\omega_\lambda$

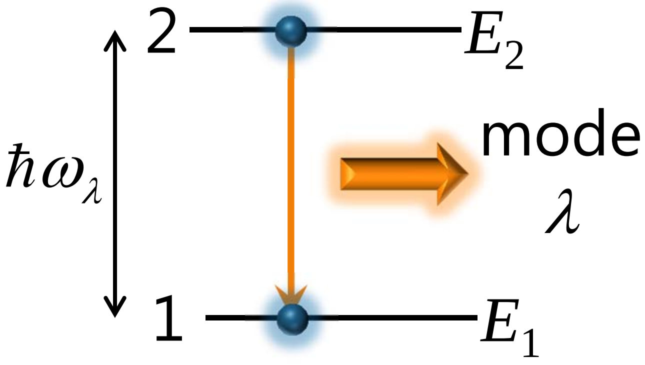
i.e., we have $w_q = \frac{2\pi}{\hbar} \left| H_{ed\lambda 12} \right|^2 \delta\left(E_1 - E_2 + \hbar\omega_\lambda \right)$

Spontaneous emission

This transition process is
  spontaneous emission

The electron starts in its higher state 2
  with no photons present

and ends in its lower state 1
  with one photon present



Spontaneous emission

This photon can be in any mode λ

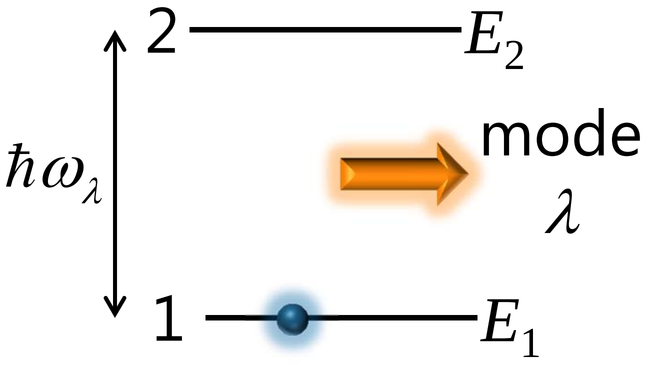
with the correct photon energy to match the energy separation

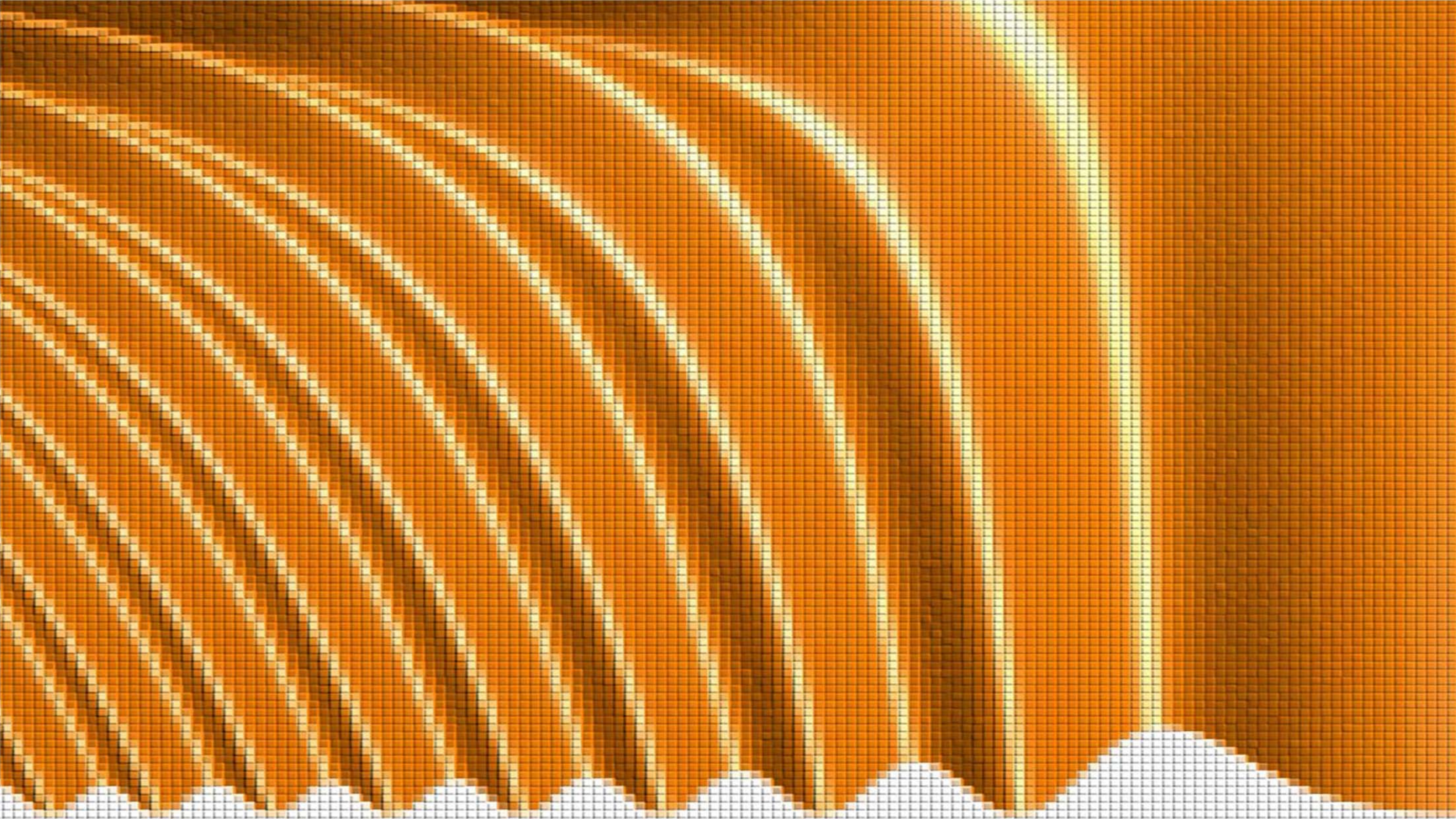
and for which $H_{ed\lambda 12}$ is not zero for some other reason

This process has emerged naturally

as a consequence of quantizing the electromagnetic field

requiring essentially no additional physics except quantization



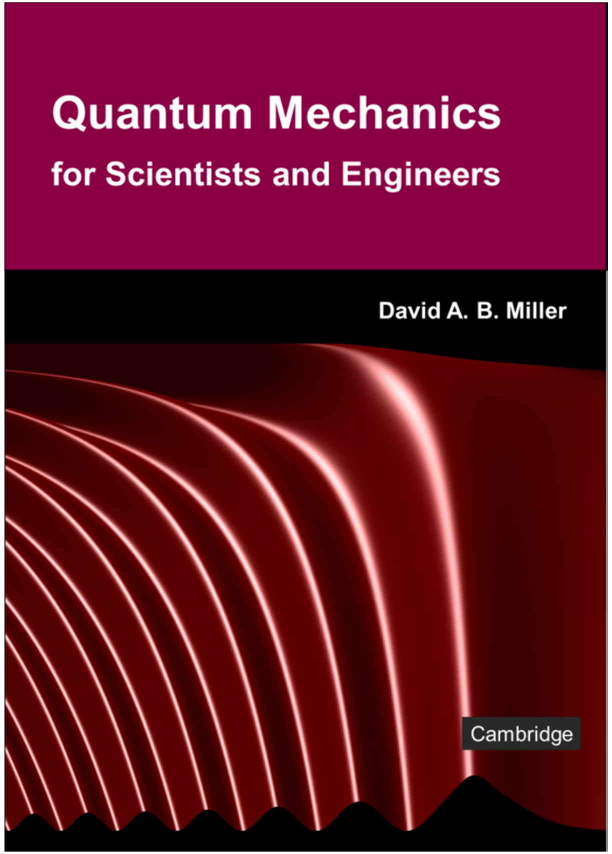


17.1 Spontaneous and stimulated emission

Slides: Video 17.1.3 Stimulated emission

Text reference: Quantum Mechanics for Scientists and Engineers

Section 17.4 subsections “Stimulated emission” and “Multiple-photon case”



Spontaneous and stimulated emission

Stimulated emission

Quantum mechanics for scientists and engineers

David Miller

Stimulated emission

Suppose now we have

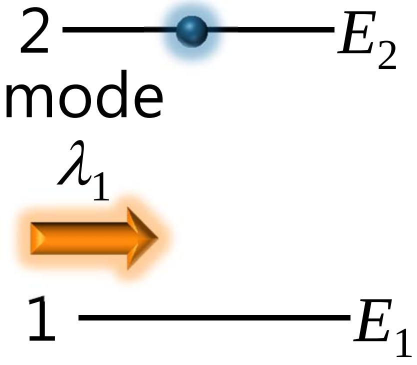
- a photon in mode λ_1 and
- an electron in its upper state 2

The initial state is therefore

$$\left| N_{fs}; N_{bs} \right\rangle = \hat{b}_2^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle$$

with an energy

$$E_s = E_2 + \hbar \omega_{\lambda_1}$$



Stimulated emission

Then, with algebra similar to that used before

$$\hat{b}_j^\dagger \hat{b}_k \left(\hat{a}_\lambda - \hat{a}_\lambda^\dagger\right) \hat{b}_2^\dagger \hat{a}_{\lambda_1}^\dagger \left|0\right\rangle = \delta_{k2}\delta_{\lambda\lambda_1}\hat{b}_j^\dagger \left|0\right\rangle - \delta_{k2}\hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \left|0\right\rangle$$

The first term $\delta_{k2}\delta_{\lambda\lambda_1}\hat{b}_j^\dagger \left|0\right\rangle$ is simply the absorption term

but this will vanish in Fermi’s Golden Rule because

there is no electron state into which we can absorb

given that we are starting in the upper state

The second term has two possibilities in the summation

$\lambda \neq \lambda_1$ both final photons in different modes

$\lambda = \lambda_1$ both final photons in the same mode

Final two photons in different modes

Suppose $\lambda \neq \lambda_1$ so $\hat{H}_p \left| N_{fs}; N_{bs} \right\rangle$ leads to a term $\propto \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle$

Then for some specific λ

to get a non-zero result for $\left\langle N_{fq}; N_{bq} \right| \hat{H}_p \left| N_{fs}; N_{bs} \right\rangle$

the final state will have to be

$$\left| N_{fq}; N_{bq} \right\rangle = \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle$$

with energy $E_q = E_j + \hbar\omega_\lambda + \hbar\omega_{\lambda_1}$

corresponding to a state with

the electron in level j and

a photon in each of the different modes λ and λ_1

Final two photons in different modes

We will have, for some specific λ and j

$$\dot{c}_q^{(1)} \simeq \frac{-1}{i\hbar} \exp\left[i\left(E_j - E_2 + \hbar\omega_\lambda \right) t / \hbar \right] H_{ed\lambda j2} \left\langle 0 \right| \hat{a}_{\lambda_1} \hat{a}_\lambda \hat{b}_j \hat{b}_j^\dagger \hat{a}_\lambda^\dagger \hat{a}_{\lambda_1}^\dagger \left| 0 \right\rangle$$

$$= \frac{-1}{i\hbar} \exp\left[i\left(E_j - E_2 + \hbar\omega_\lambda \right) t / \hbar \right] H_{ed\lambda j2}$$

giving a transition rate $w_q = \frac{2\pi}{\hbar} \left| H_{ed\lambda j2} \right|^2 \delta\left(E_j - E_2 + \hbar\omega_\lambda \right)$

for which the only possibility here for non-zero transition rate is $j = 1$ and $E_2 - E_1 = \hbar\omega_\lambda$

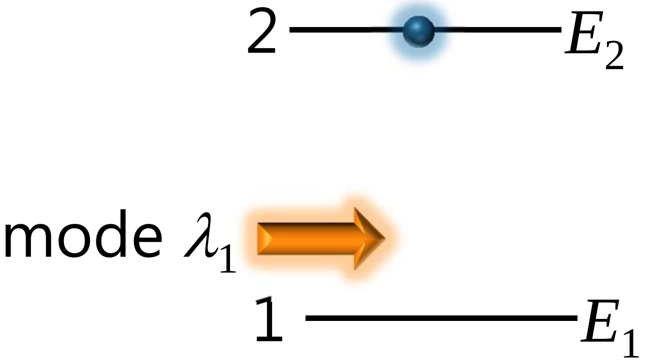
This process is just spontaneous emission into mode λ

with the same transition rate as before

Spontaneous emission

So this process starts with

- the electron in state 2
- a photon in mode λ_1



Spontaneous emission

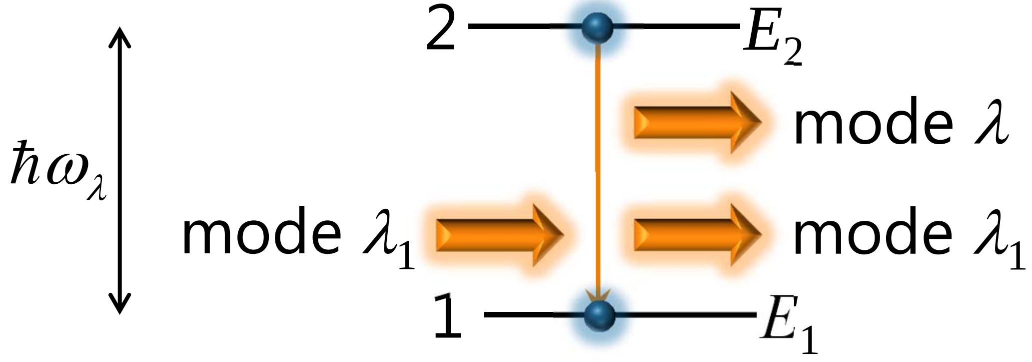
So this process starts with

- the electron in state 2
- a photon in mode λ_1

and ends with

- the electron in state 1
- a photon in mode λ_1
- and a photon spontaneously emitted into mode λ

The spontaneous emission in mode λ is unaffected by the photon in mode λ_1



Final two photons in the same mode

Suppose $\lambda = \lambda_1$ so $\hat{H}_p \left| N_{fs}; N_{bs} \right\rangle$ leads to a term $\propto \hat{b}_j^\dagger \left(\hat{a}_{\lambda_1}^\dagger \right)^2 \left| 0 \right\rangle$

So to get a non-zero result from $\left\langle N_{fq}; N_{bq} \right| \hat{H}_p \left| N_{fs}; N_{bs} \right\rangle$

the final (normalized) state has to be

$\left| N_{fq}; N_{bq} \right\rangle = \frac{1}{\sqrt{2!}} \hat{b}_j^\dagger \left(\hat{a}_{\lambda_1}^\dagger \right)^2 \left| 0 \right\rangle$ with an energy $E_q = E_j + 2\hbar\omega_{\lambda_1}$

Note that, to have a normalized state here

we have had to introduce the factor $1/\sqrt{2!}$

as in the general normalized multiple boson state

$$\left| n_1, n_2, \ldots, n_\lambda, \ldots \right\rangle = \frac{1}{\sqrt{n_1! n_2! \ldots n_\lambda! \ldots}} \left(\hat{a}_1^\dagger \right)^{n_1} \left(\hat{a}_2^\dagger \right)^{n_2} \ldots \left(\hat{a}_\lambda^\dagger \right)^{n_\lambda} \ldots \left| 0 \right\rangle$$

Final two photons in the same mode

Hence from $\left\langle N_{fq}; N_{bq} \middle| \hat{H}_p \middle| N_{fs}; N_{bs} \right\rangle$

we are obtaining a term

$$H_{ed\lambda_1 j2} \langle 0 | \frac{1}{\sqrt{2!}} \left(\hat{a}_{\lambda_1} \right)^2 \hat{b}_j \hat{b}_j^\dagger \left(\hat{a}_{\lambda_1}^\dagger \right)^2 | 0 \rangle$$

$$= \sqrt{2!}\, H_{ed\lambda_1 j2} \langle 0 | \frac{1}{\sqrt{2!}} \left(\hat{a}_{\lambda_1} \right)^2 \hat{b}_j \frac{1}{\sqrt{2!}} \hat{b}_j^\dagger \left(\hat{a}_{\lambda_1}^\dagger \right)^2 | 0 \rangle = \sqrt{2} H_{ed\lambda_1 j2}$$

The $\sqrt{2}$ is very important

It comes from having two photons in the mode

and leads to a larger amplitude for this process

than for spontaneous emission

Final two photons in the same mode

Hence for this process we have

$$\dot{c}_q^{(1)} \simeq \frac{-1}{i\hbar} \exp\left[i\left(E_j - E_2 + \hbar\omega_{\lambda_1} \right) t / \hbar \right] \sqrt{2}\, H_{ed\lambda_1 j2}$$

leading to a transition rate into this final state of

$$w_q = \frac{2\pi}{\hbar} 2 \left| H_{ed\lambda_1 j2} \right|^2 \delta\left(E_j - E_2 + \hbar\omega_{\lambda_1} \right)$$

for which the only possibility for finite transition rate is with $j = 1$ and $E_2 - E_1 = \hbar\omega_{\lambda_1}$

with a corresponding transition rate, finally, of

$$w_q = \frac{2\pi}{\hbar} 2 \left| H_{ed\lambda_1 12} \right|^2 \delta\left(E_1 - E_2 + \hbar\omega_{\lambda_1} \right)$$

Stimulated emission

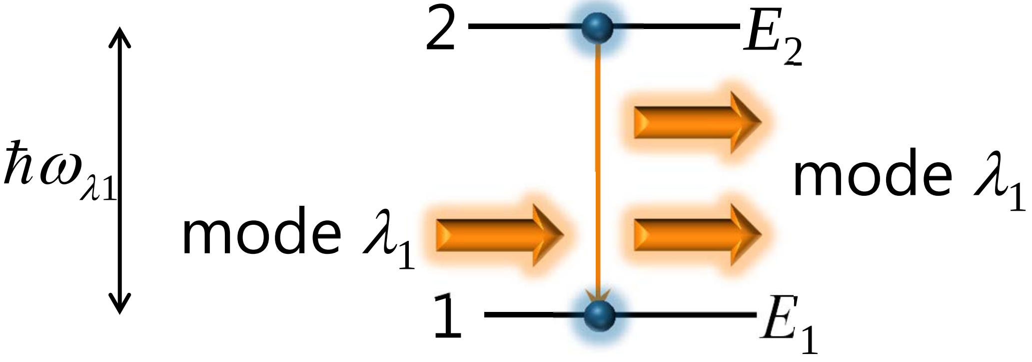
The process starts with

- the electron in state 2
- a photon in mode λ_1

and ends with

- the electron in state 1
- two photons in mode λ_1

This process is stimulated emission



Stimulated emission

Note the additional factor of 2 in the transition rate

$$w_q = \frac{2\pi}{\hbar} 2 \left| H_{ed\lambda_1 12} \right|^2 \delta\left(E_1 - E_2 + \hbar\omega_{\lambda_1} \right)$$

Note that, other things being equal

e.g., matrix elements and energies

the transition rate into the mode already occupied with a photon

is twice as high as the spontaneous emission into an unoccupied mode

Bosons want to go into modes that are already occupied!

Multiphoton case

It is left as an exercise to analyze the case of n_{λ_1} photons initially in mode λ_1

The result for stimulated emission is

$$w_q = \frac{2\pi}{\hbar}\left(n_{\lambda_1}+1\right)\left|H_{ed\lambda_1 12}\right|^2 \delta\left(E_1 - E_2 + \hbar\omega_{\lambda_1}\right)$$

with the transition rate into the mode λ_1 being

$n_{\lambda_1}+1$ times larger than the spontaneous rate into an otherwise similar mode

Multiphoton case

Spontaneous emission in any other mode λ

- is unaffected by the presence of n_{λ_1} photons in mode λ_1
  - as can be shown directly by considering the multiphoton case

The result for absorption with n_{λ_1} photons initially in mode λ_1

- can similarly be shown to be a transition rate

$$w_q = \frac{2\pi}{\hbar} n_{\lambda_1} \left| H_{ed\,\lambda_1 12} \right|^2 \delta\left(E_2 - E_1 - \hbar\omega_{\lambda_1} \right)$$

- where the absorption rate is proportional to the number of photons in the mode
  - as we would expect

Multiphoton case

Note specifically in the absorption transition rate

$$w_q = \frac{2\pi}{\hbar} n_{\lambda_1} \left| H_{ed\lambda_1 12} \right|^2 \delta\left(E_2 - E_1 - \hbar\omega_{\lambda_1} \right)$$

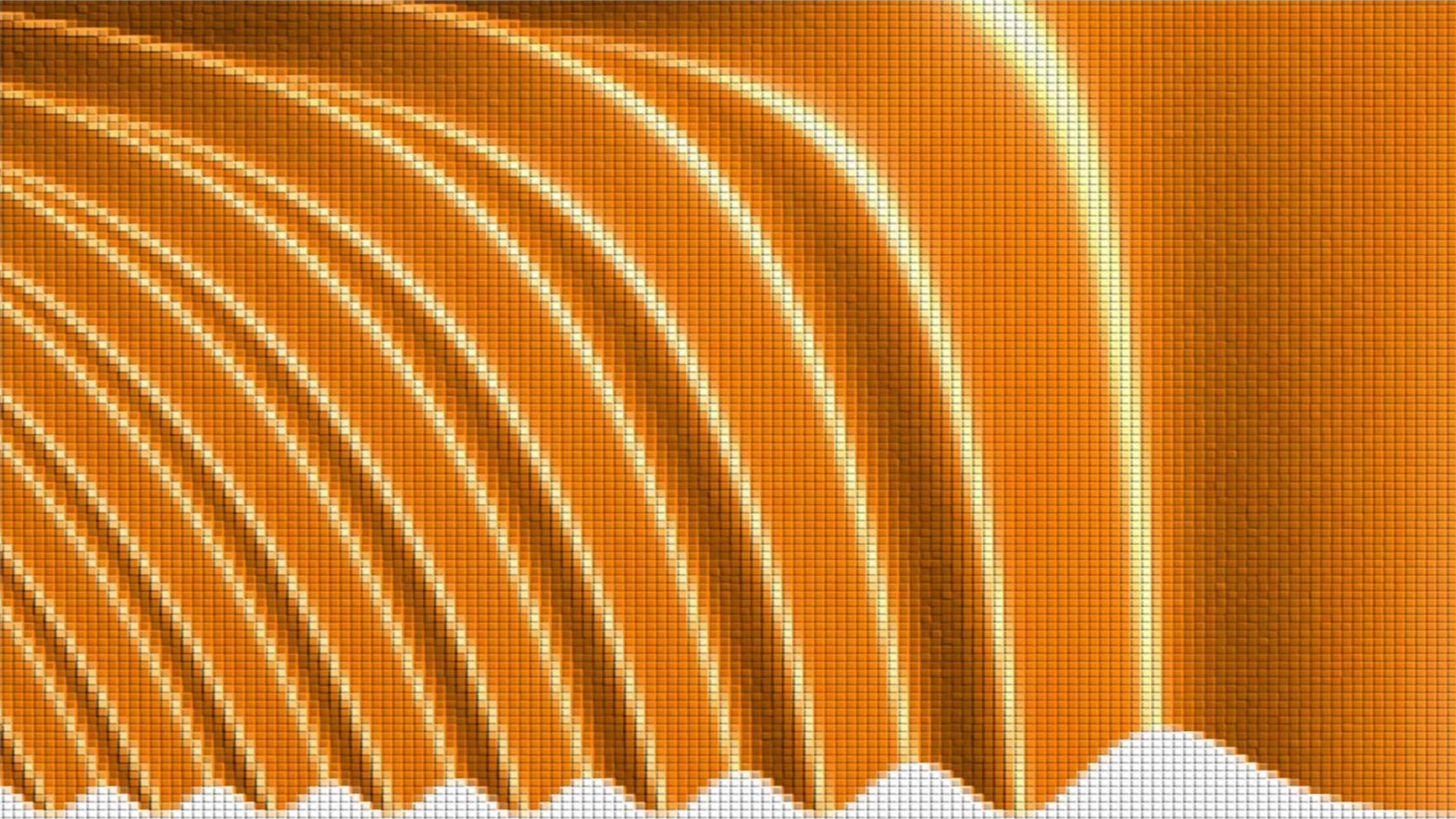
that we wrote the matrix element $H_{ed\lambda_1 12}$

not the matrix element $H_{ed\lambda_1 21}$

Given the definition of $H_{ed\lambda_1 jk}$ above

$$H_{ed\lambda jk} = \sqrt{-1}\, e \sqrt{\frac{\hbar\omega_\lambda}{2\varepsilon_o}} \int \phi_j^*\left(\mathbf{r}_i\right)\left[\mathbf{u}_\lambda\left(\mathbf{r}_i\right).\mathbf{r}_i\right]\phi_k\left(\mathbf{r}_i\right) d^3\mathbf{r}_i$$

we see that $H_{ed\lambda_1 12} = H^*_{ed\lambda_1 21}$ so $\left| H_{ed\lambda_1 12} \right|^2 = \left| H_{ed\lambda_1 21} \right|^2$

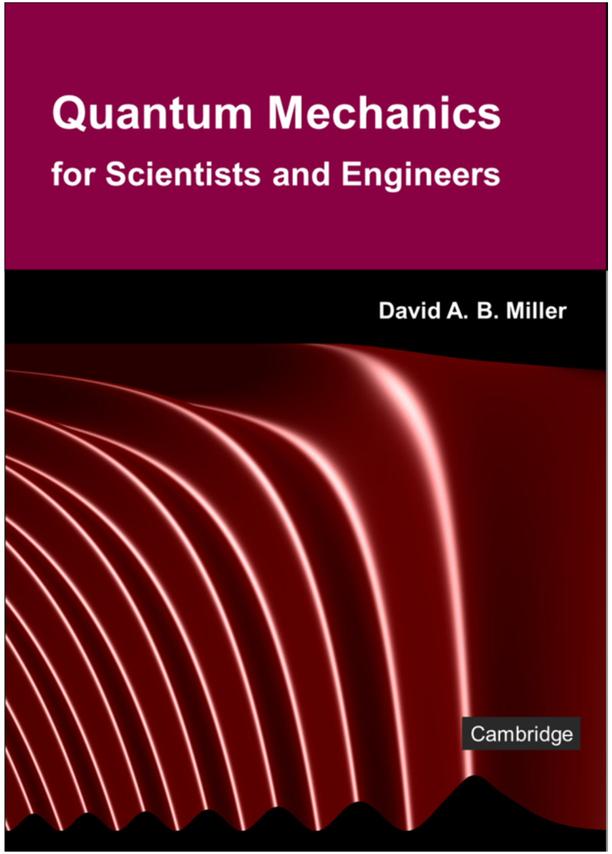


17.1 Spontaneous and stimulated emission

Slides: Video 17.1.5 Total spontaneous emission rate

Text reference: Quantum Mechanics for Scientists and Engineers

Section 17.4 subsection "Total spontaneous emission rate"



Spontaneous and stimulated emission

Total spontaneous emission rate

Quantum mechanics for scientists and engineers

David Miller

Total spontaneous emission rate

We presume we start off with

the electron in an excited state (here state 2), and

no photons in any modes

The total spontaneous transition rate will be

the sum of the transition rates into all possible final states q through spontaneous emission

$$W_{spon} = \sum_q w_q$$

where w_q is the spontaneous emission rate into some specific mode λ

Total spontaneous emission rate

The specific transition rate into mode λ is

$$w_q = \frac{2\pi}{\hbar} \left| H_{ed\lambda 12} \right|^2 \delta\left(E_1 - E_2 + \hbar\omega_\lambda \right)$$

Since here we are presuming the electron starts in state 2 and ends in state 1

the sum $W_{spon} = \sum_q w_q$ over all possible final states

reduces to a sum over all possible photon modes λ

$$W_{spon} = \frac{2\pi}{\hbar} \sum_\lambda \left| H_{ed\lambda 12} \right|^2 \delta\left(E_1 - E_2 + \hbar\omega_\lambda \right)$$

Matrix element evaluation

To calculate the transition rate

- we start with the matrix element $H_{ed\lambda 12}$

We presume the field is approximately uniform

- over the size scale of the quantum system
  - because the wavelength of the light is long
    - compared to the size of an atom

So if the quantum system is centered at some position $\mathbf{r}_o$

- we can just use the field at that point
  - simplifying the calculation of the matrix element

Matrix element evaluation

So in the matrix element

$$H_{ed\lambda jk} = i\, e\sqrt{\frac{\hbar\omega_\lambda}{2\varepsilon_o}}\int \phi_j^*(\mathbf{r})\left[\mathbf{u}_\lambda(\mathbf{r})\cdot\mathbf{r}\right]\phi_k(\mathbf{r})\,d^3\mathbf{r}$$

we simplify by replacing $\mathbf{u}_\lambda(\mathbf{r})$ by $\mathbf{u}_\lambda(\mathbf{r}_o)$

obtaining $H_{ed\lambda jk} \simeq i\, e\sqrt{\frac{\hbar\omega_\lambda}{2\varepsilon_o}}\mathbf{u}_\lambda(\mathbf{r}_o)\cdot\mathbf{r}_{jk}$

where $\mathbf{r}_{jk} = \int \phi_j^*(\mathbf{r})\mathbf{r}\phi_k(\mathbf{r})\,d^3\mathbf{r}$

Plane wave modes

We need a normalizable form for plane wave modes

We imagine we have a cubic box of volume V_b

It is common for mathematical convenience to use running waves with periodic boundary conditions

though one could use standing waves and get the same result for a large box

The resulting modes have the form

$$\mathbf{u}_{\lambda}(\mathbf{r}) = \mathfrak{e}\frac{1}{\sqrt{V_b}}\exp\left(i\mathbf{k}_{\lambda}\cdot\mathbf{r}\right)$$

where $\mathfrak{e}$ is a unit vector in the polarization direction of the electric field

Plane wave modes

These modes are normalized over the box of volume V_b

The allowed values of k_x are spaced by $2\pi / L_x$

- where L_x is the length of the box in the x direction
  - and similarly for the y and z directions
  - leading to a density of modes in $\mathbf{k}$-space of $V_b / (2\pi)^3$

For such propagating waves

- we will also have two distinct polarization directions
  - though we will handle polarization properties directly

Plane wave modes

We approximate the sum over the modes λ

by an integral over $\mathbf{k}$ with this density of states $V_b/(2\pi)^3$

and also formally a sum over the two possible polarizations, i.e.,

$$\sum_{\lambda}\ldots \to \sum_{polarizations}\int \frac{V_b}{(2\pi)^3}d^3\mathbf{k}_{\lambda}\ldots$$

In considering the polarizations

we choose polarization directions at right angles to one another

and at right angles to $\mathbf{k}_{\lambda}$

Plane wave modes

Specifically we choose polarization directions

relative to (vector) matrix element $\mathbf{r}_{12}$

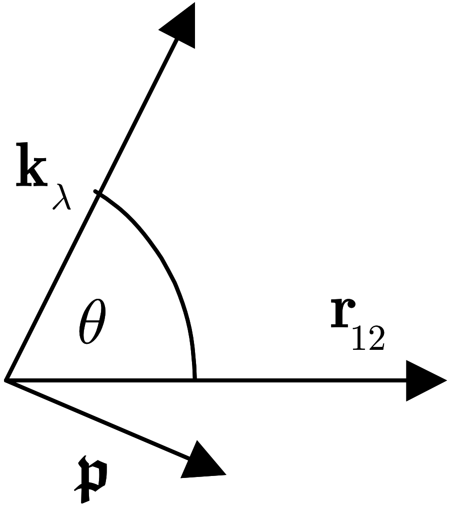
Here we choose one polarization $\mathfrak{p}$

to be in the plane of the vectors $\mathbf{k}_\lambda$ and $\mathbf{r}_{12}$

With this choice

the other polarization direction is perpendicular to $\mathbf{r}_{12}$ and so

$\mathbf{u}_\lambda(\mathbf{r}_o)\cdot\mathbf{r}_{12}$ vanishes for this polarization



Plane wave modes

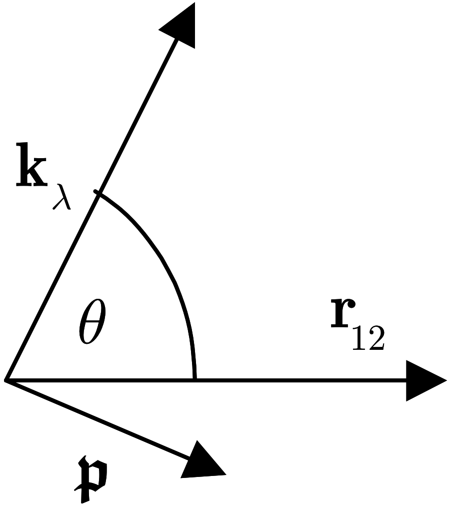
Hence in our summation

we need only retain the first polarization

For this choice, we therefore find that

$$\mathbf{u}_\lambda\left(\mathbf{r}_o\right)\cdot\mathbf{r}_{12} = u_\lambda\left(\mathbf{r}_o\right)r_{12}\sin\theta$$

(the non-bold quantities refer to the vector magnitudes)



Total spontaneous transition rate

Now we can use all these results to rewrite W_{spon}, i.e.,

$$W_{spon} = \frac{2\pi}{\hbar}\sum_{\lambda}\left|H_{ed\lambda 12}\right|^2 \delta\left(E_1 - E_2 + \hbar\omega_\lambda\right)$$

$$= \frac{2\pi}{\hbar}\int \frac{V_b}{\left(2\pi\right)^3}\left| ie\sqrt{\frac{\hbar\omega_\lambda}{2\varepsilon_o}}\frac{1}{\sqrt{V_b}}\exp\left(i\mathbf{k}_\lambda \cdot \mathbf{r}_o\right) r_{12}\sin\theta\right|^2 \delta\left(E_1 - E_2 + \hbar\omega_\lambda\right) d^3\mathbf{k}_\lambda$$

$$= \frac{e^2\left|r_{12}\right|^2}{8\pi^2\varepsilon_o}\int \omega_\lambda \sin^2\theta\, \delta\left(E_1 - E_2 + \hbar\omega_\lambda\right) d^3\mathbf{k}_\lambda$$

$$= \frac{e^2\left|r_{12}\right|^2}{8\pi^2\varepsilon_o}\int_{k_\lambda=0}^{\infty}\int_{\theta=0}^{\pi} \omega_\lambda\, \delta\left(E_1 - E_2 + \hbar\omega_\lambda\right)\sin^2\theta\; 2\pi\sin\theta k_\lambda^2 d\theta dk_\lambda$$

Total spontaneous transition rate

In $W_{spon}=\frac{e^2\left|r_{12}\right|^2}{8\pi^2\varepsilon_o}\int_{k_\lambda=0}^{\infty}\int_{\theta=0}^{\pi}\omega_\lambda\,\delta\left(E_1-E_2+\hbar\omega_\lambda\right)\sin^2\theta\,2\pi\sin\theta k_\lambda^2 d\theta dk_\lambda$

with $ck_\lambda=\omega_\lambda$ and changing variables to $\hbar ck_\lambda\equiv\hbar\omega_\lambda$,

$$W_{spon}=\frac{e^2\left|r_{12}\right|^2}{4\pi\varepsilon_o\hbar^4c^3}\int_{\hbar\omega_\lambda=0}^{\infty}\hbar\omega_\lambda\,\delta\left(E_1-E_2+\hbar\omega_\lambda\right)\left(\hbar\omega_\lambda\right)^2 d\hbar\omega_\lambda\int_0^{\pi}\sin^3\theta d\theta$$

Given that $\int_0^{\pi}\sin^3\theta d\theta=4/3$

we finally have the total spontaneous emission rate

$$W_{spon}=\frac{e^2\left|r_{12}\right|^2\omega_{12}^3}{3\pi\varepsilon_o\hbar c^3}\quad \text{where}\quad \omega_{12}=\left(E_2-E_1\right)/\hbar$$

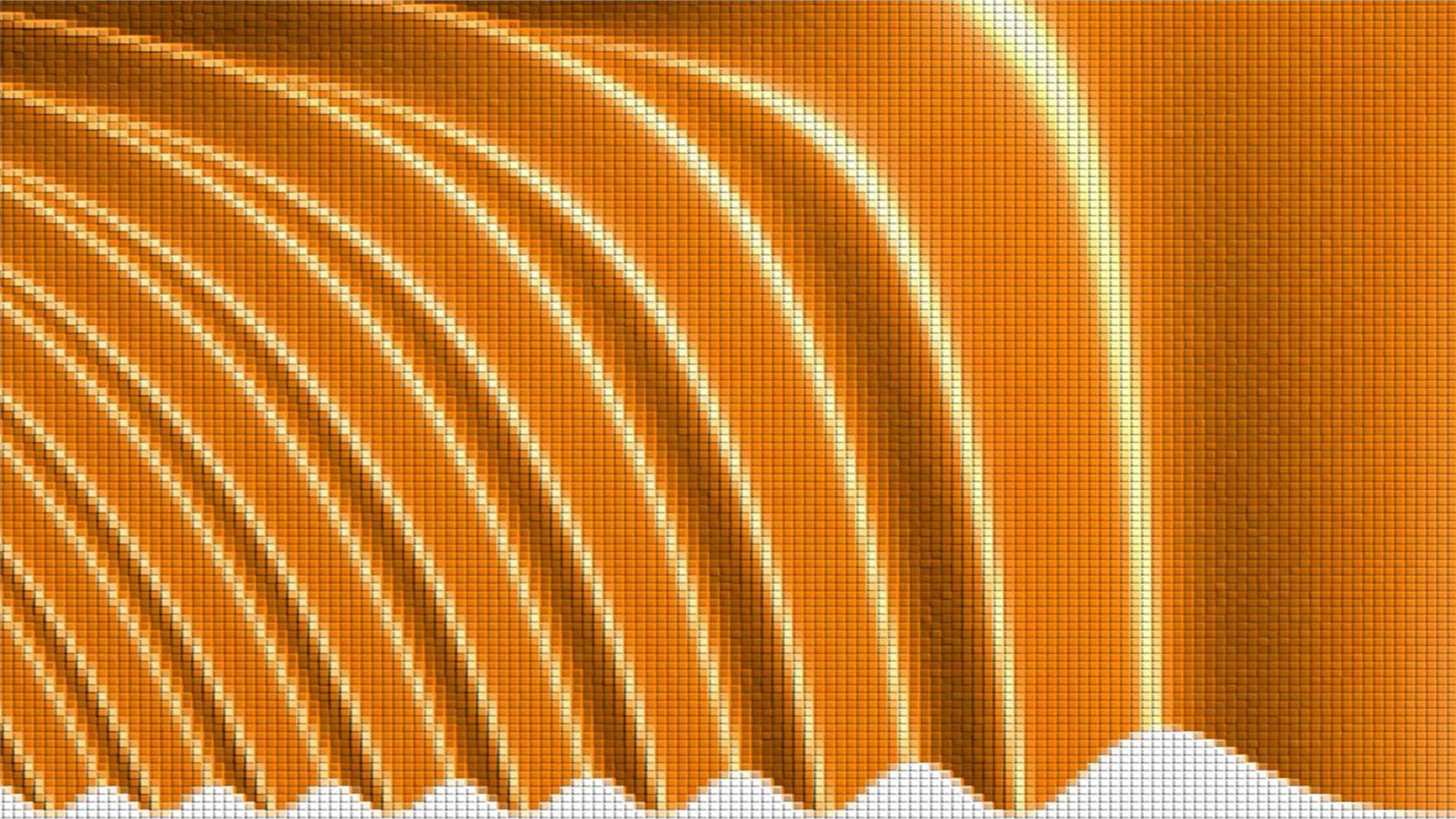
Natural lifetime

Such a rate gives a natural lifetime, τ_{nat}, for a state

$$\tau_{nat} = 1/W_{spon}$$

A quantum mechanical system sitting in empty space in an excited state

- will decay on average over this timescale to its lower state, emitting a photon
  - The direction of the mode into which the photon is emitted is random
    - though weighted somewhat by the polarization effects

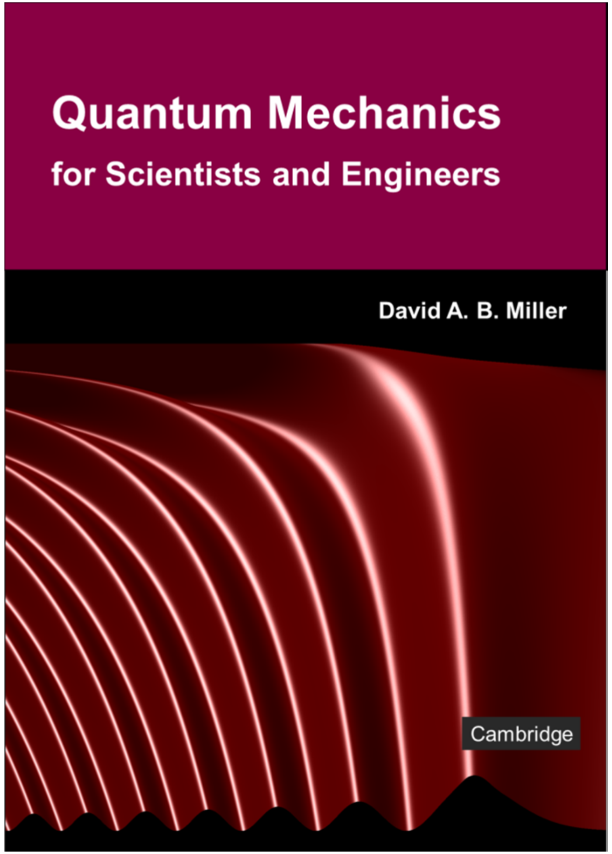


17.2 Mixed states and the density matrix

Slides: Video 17.2.1 Introduction to mixed states and the density matrix

Text reference: Quantum Mechanics for Scientists and Engineers

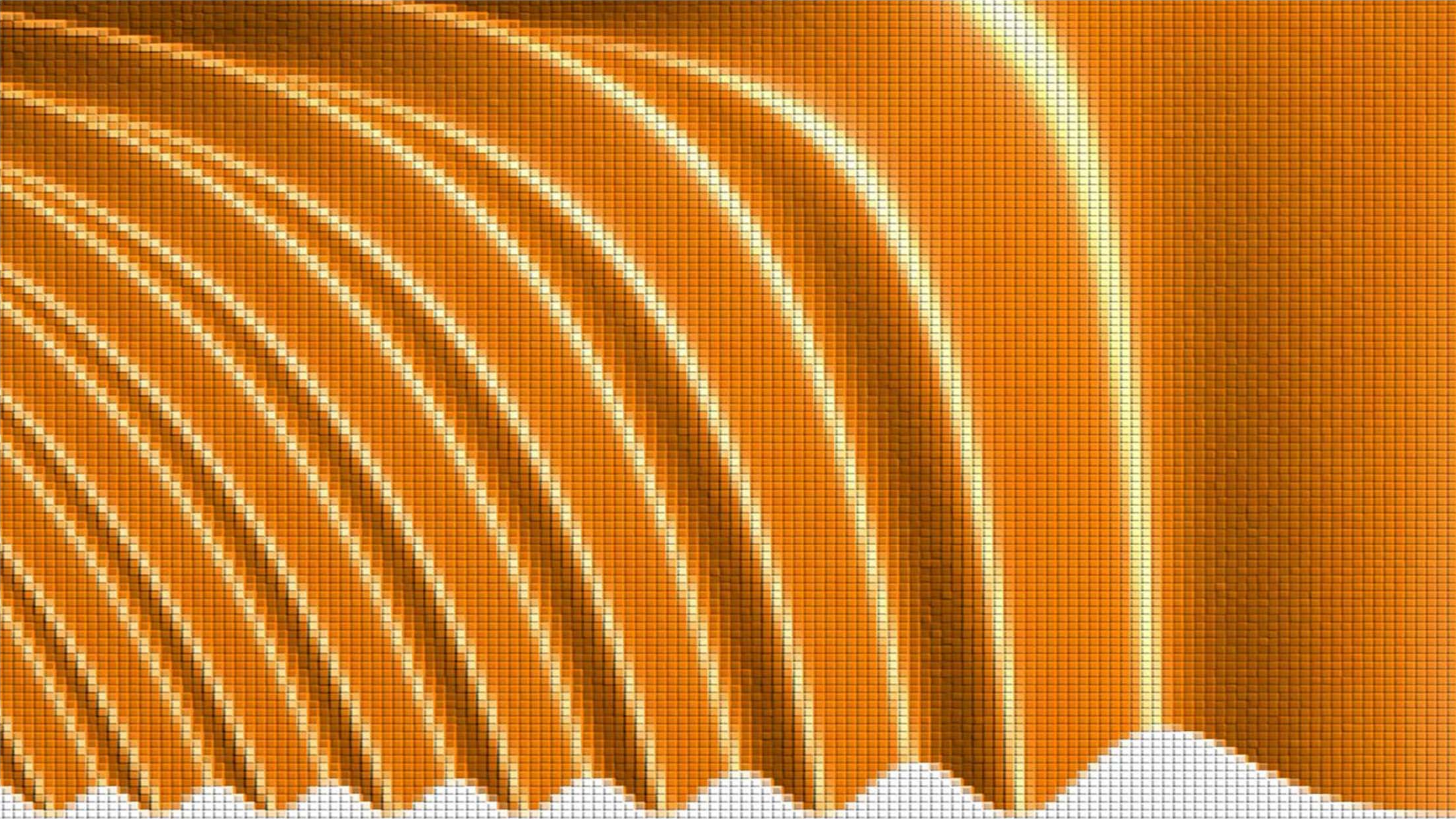
Chapter 14 introduction



Mixed states and the density matrix

Quantum mechanics for scientists and engineers

David Miller

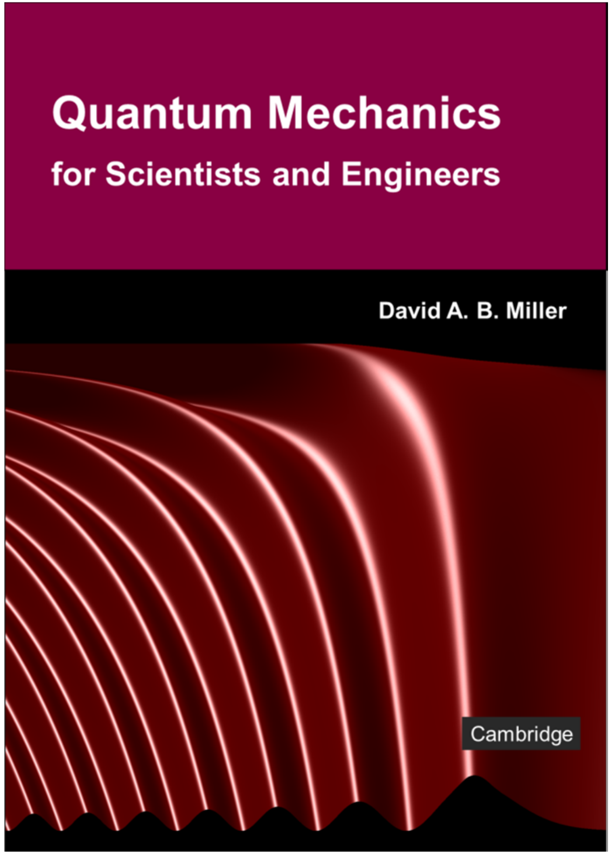


17.2 Mixed states and the density matrix

Slides: Video 17.2.2 Pure and mixed states

Text reference: Quantum Mechanics for Scientists and Engineers

Section 14.1 (first part)



Mixed states and the density matrix

Pure and mixed states

Quantum mechanics for scientists and engineers

David Miller

Pure and mixed states

So far the only randomness we have considered

- is from quantum-mechanical measurement

Consider, e.g., the state of polarization of a photon

- So far, we could write a general state of polarization as

$$|\psi\rangle = a_H |H\rangle + a_V |V\rangle$$

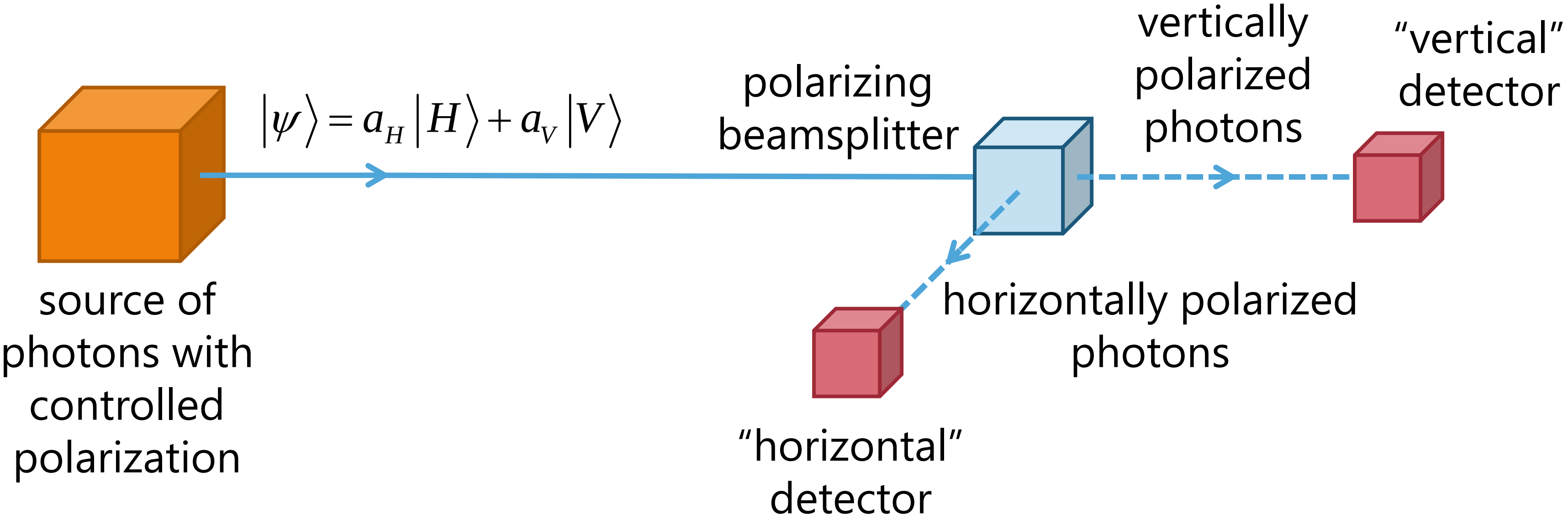
  - where $|H\rangle$ is a horizontally polarized photon state
    - and $|V\rangle$ is a vertically polarized one

Pure and mixed states

Suppose we measure such a state $|\psi\rangle = a_H |H\rangle + a_V |V\rangle$

using e.g., a polarizing beamsplitter that separates horizontal and vertical polarizations

to different outputs with different detectors

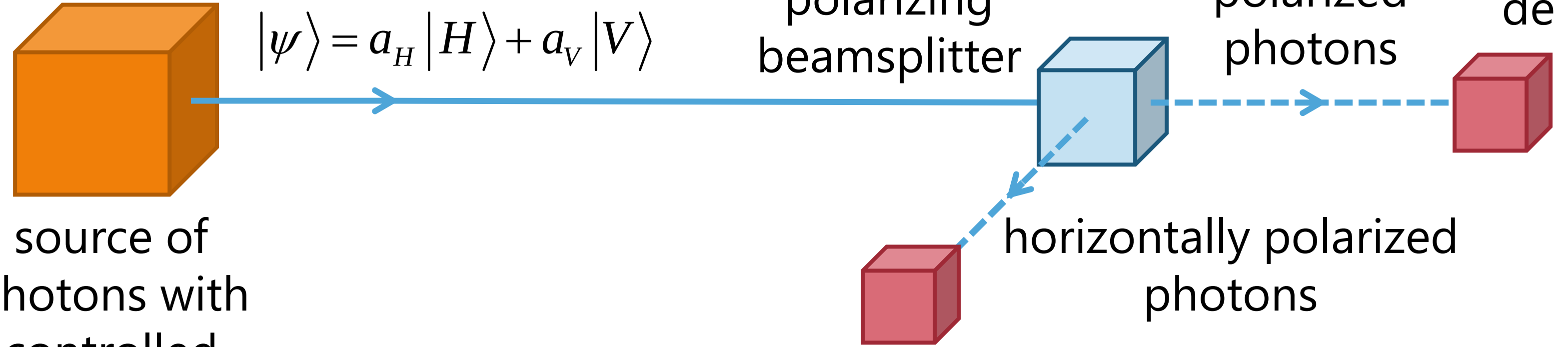


Pure and mixed states

For this state, $|\psi\rangle = a_H |H\rangle + a_V |V\rangle$ we expect probabilities

$|a_H|^2$ of measuring horizontal polarization

$|a_V|^2$ of measuring vertical polarization



More complete description of optical polarization

Since we must have $|a_H|^2 + |a_V|^2 = 1$ by normalization

we could also choose to write

$$a_H = \cos\theta \quad a_V = \exp(i\delta)\sin\theta$$

$\delta = 0$ corresponds to linear polarization

θ is then the angle of the optical electric vector relative to the horizontal axis

When $\delta \neq 0$, the field is in general "elliptically polarized"

the most general state of polarization

$\delta = \pm\pi / 2$ with $\theta = 45°$ give

right and left circular polarization

Polarization filter

We can always build a polarizing filter or compensator

- to allow passage of a photon of any specific polarization
  - 100% of the time
    - even when starting with elliptical polarization

We could arrange to delay only the horizontal polarization by a compensating amount $-\delta$

- to make the photon linearly polarized
  - then rotate the polarization or apparatus by angle θ so that the photon was always passed through
    - to, e.g., the vertical detector

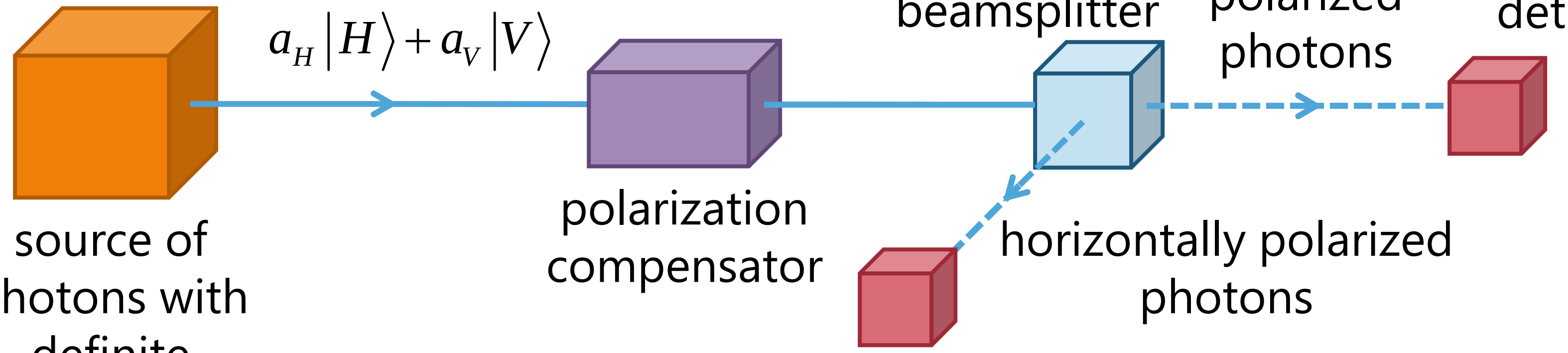
Pure states

When we can make a polarization filter or compensator

so we get 100% of the photons to one detector

we say that the photons are in a "pure" state

here $\left|\psi\right\rangle = a_H\left|H\right\rangle + a_V\left|V\right\rangle$



$a_H\left|H\right\rangle + a_V\left|V\right\rangle$

source of photons with definite polarization

polarization compensator

polarizing beamsplitter

vertically polarized photons

"vertical" detector

horizontally polarized photons

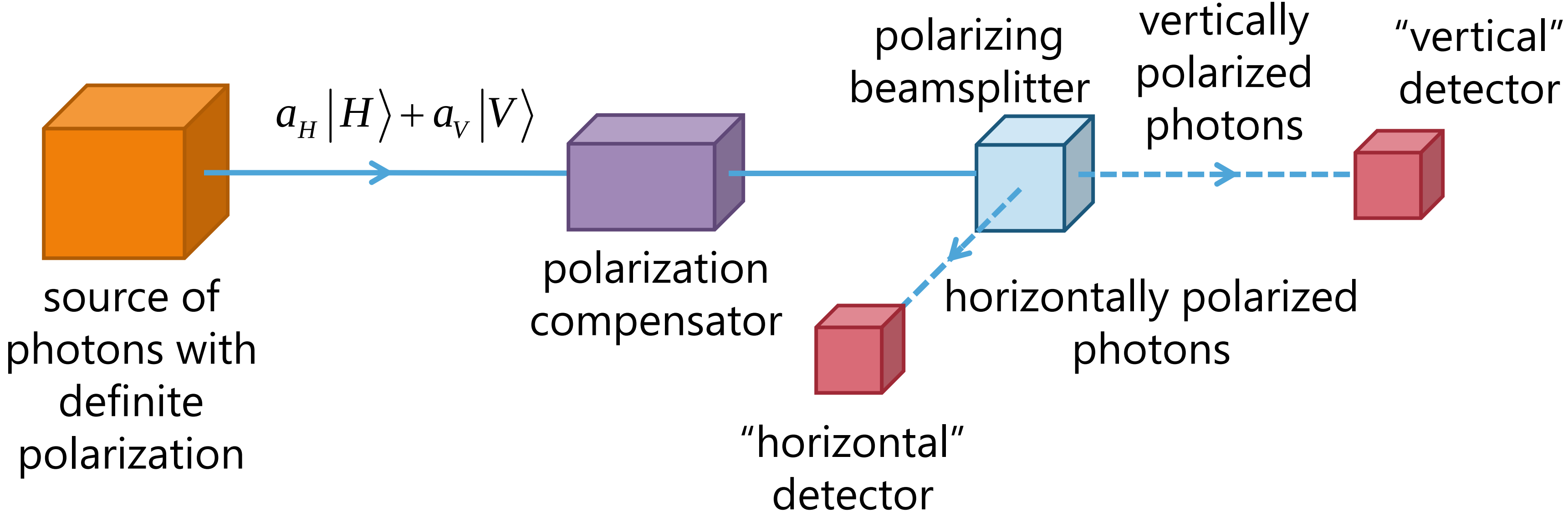
"horizontal" detector

Pure states

All states considered so far have been pure states

A "compensator" could be made to pass any particles in any one specific quantum mechanical state

with 100% efficiency to one detector



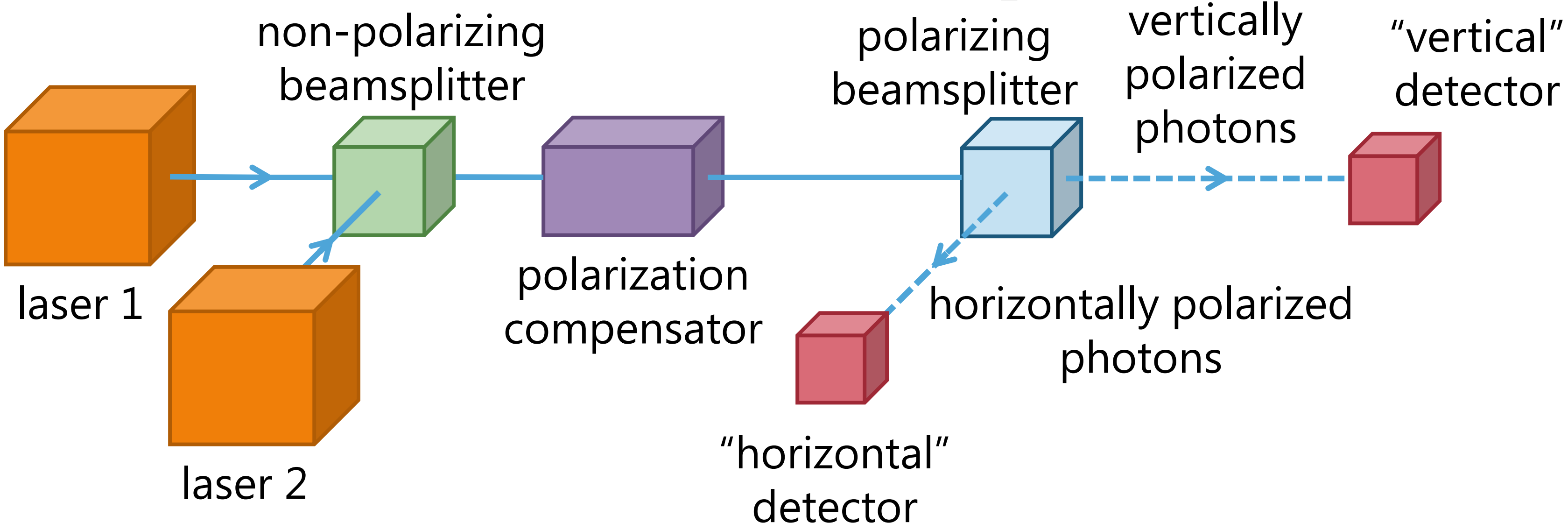
Mixed states

Suppose we have a beam that is a mixture

from two different independent lasers, “1” and “2”

Presume laser 1 contributes a fraction P_1 of the photons

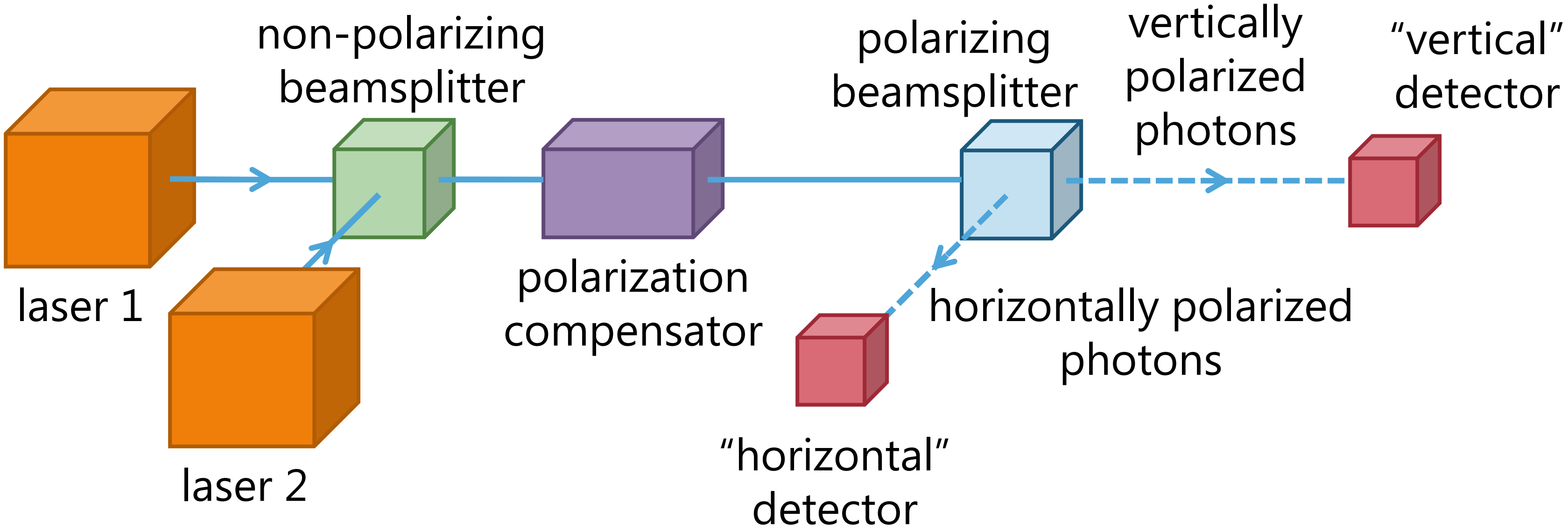
and laser 2 contributes a fraction P_2



Mixed states

Then the probability a given photon is from laser 1 is P_1

and similarly there is probability P_2 it is from laser 2



non-polarizing beamsplitter

laser 1

laser 2

polarization compensator

polarizing beamsplitter

vertically polarized photons

"vertical" detector

horizontally polarized photons

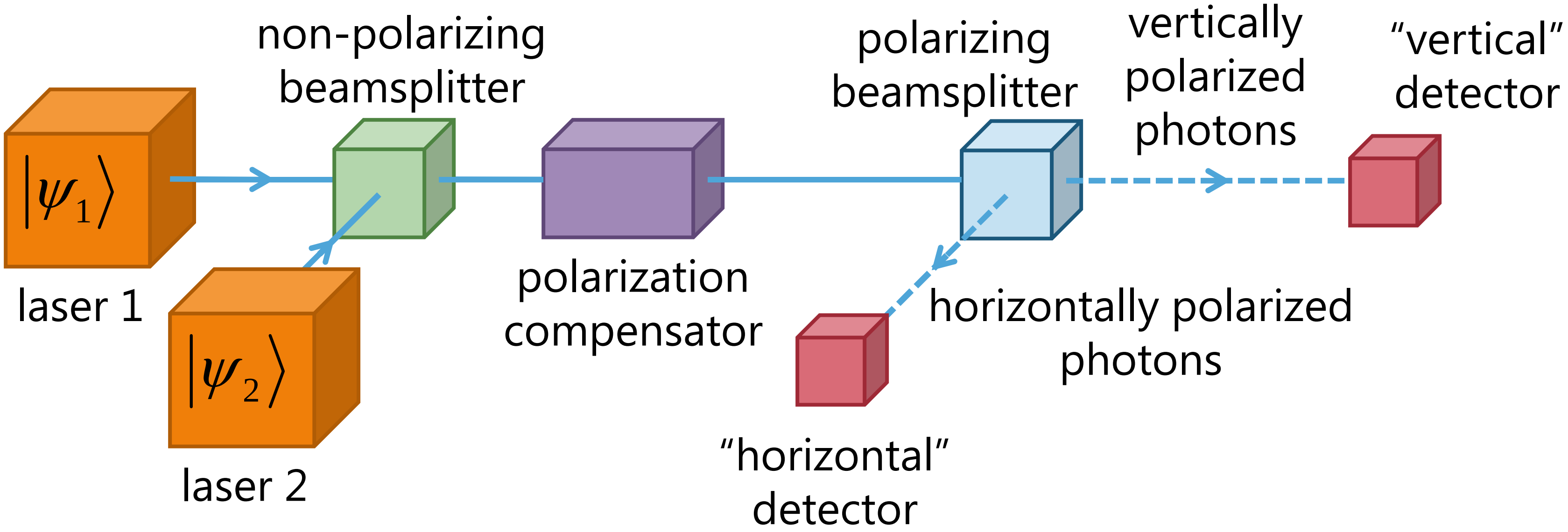
"horizontal" detector

Mixed states

Presume also that these two lasers give uncorrelated photons

of two possibly different polarization states

$|\psi_1\rangle$ and $|\psi_2\rangle$ respectively

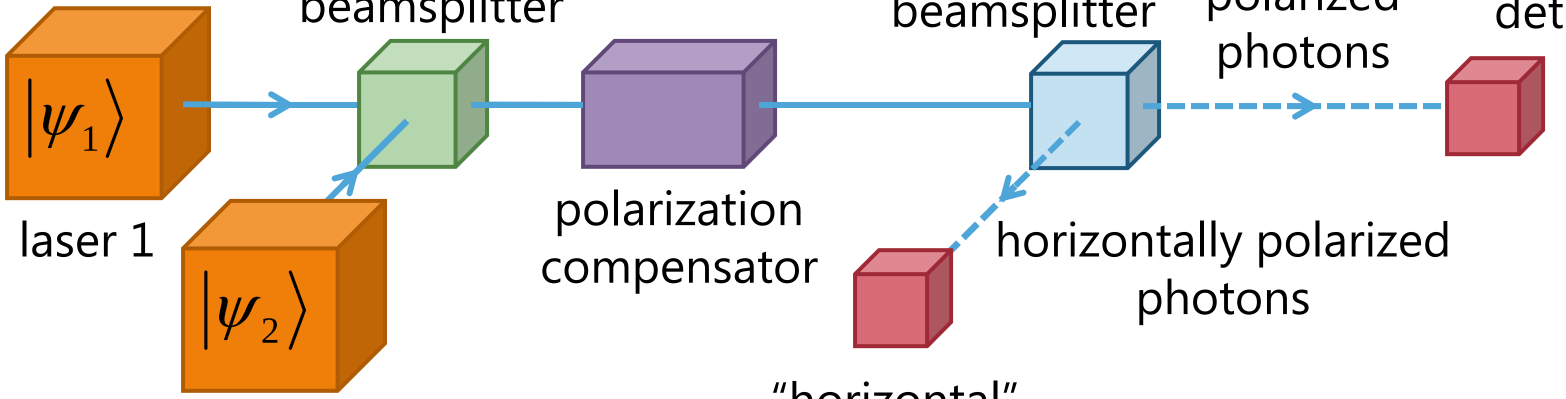


Mixed states

There is now no one setting of the compensator

that in general will pass all the photons from both lasers

to the vertical detector

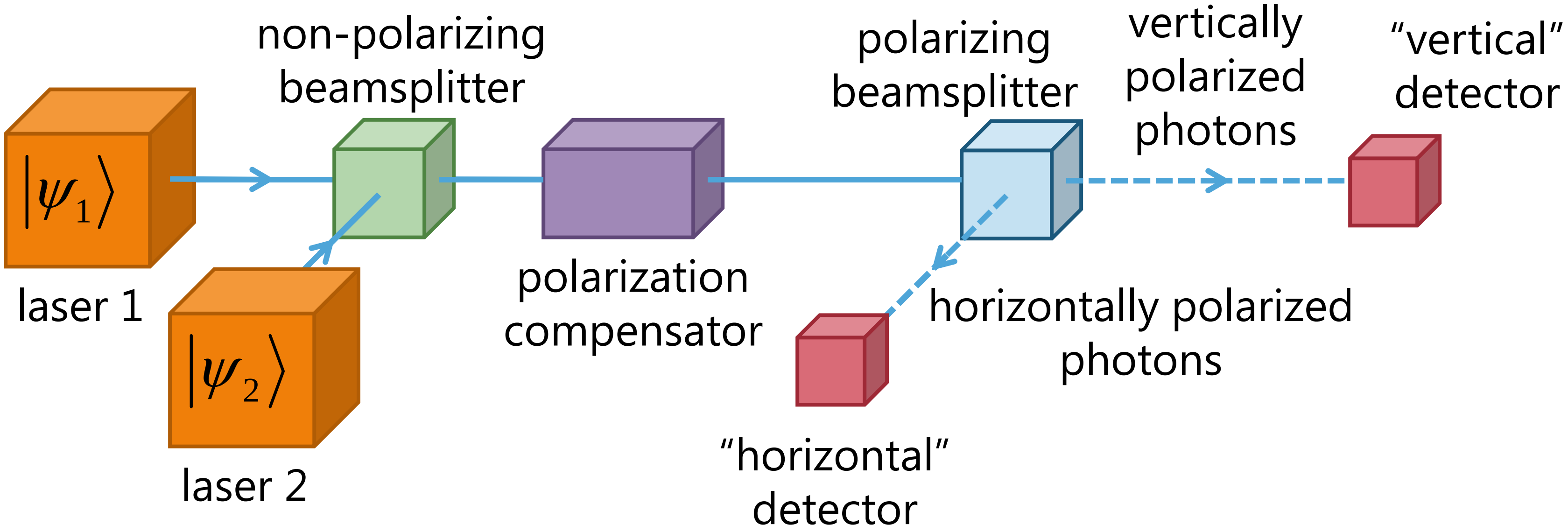


Mixed states

Hence, we cannot simply write this state

as some linear combination such as $b_1 |\psi_1\rangle + b_2 |\psi_2\rangle$

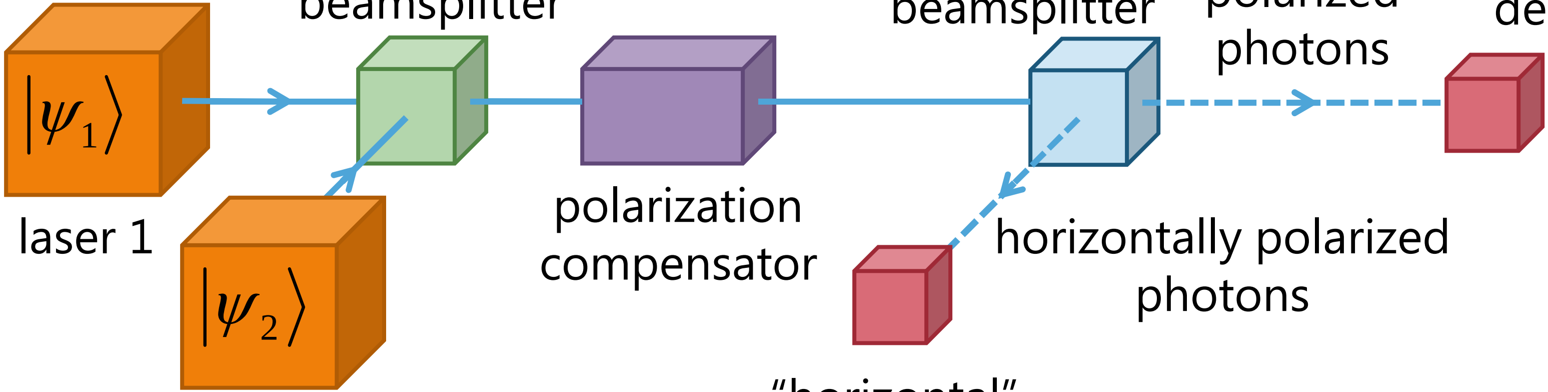
of the two different polarization states



Mixed states

If we were able to do that, we could construct a polarizing filter

to pass 100% of the photons



non-polarizing beamsplitter

polarizing beamsplitter

vertically polarized photons

“vertical” detector

laser 1

laser 2

polarization compensator

horizontally polarized photons

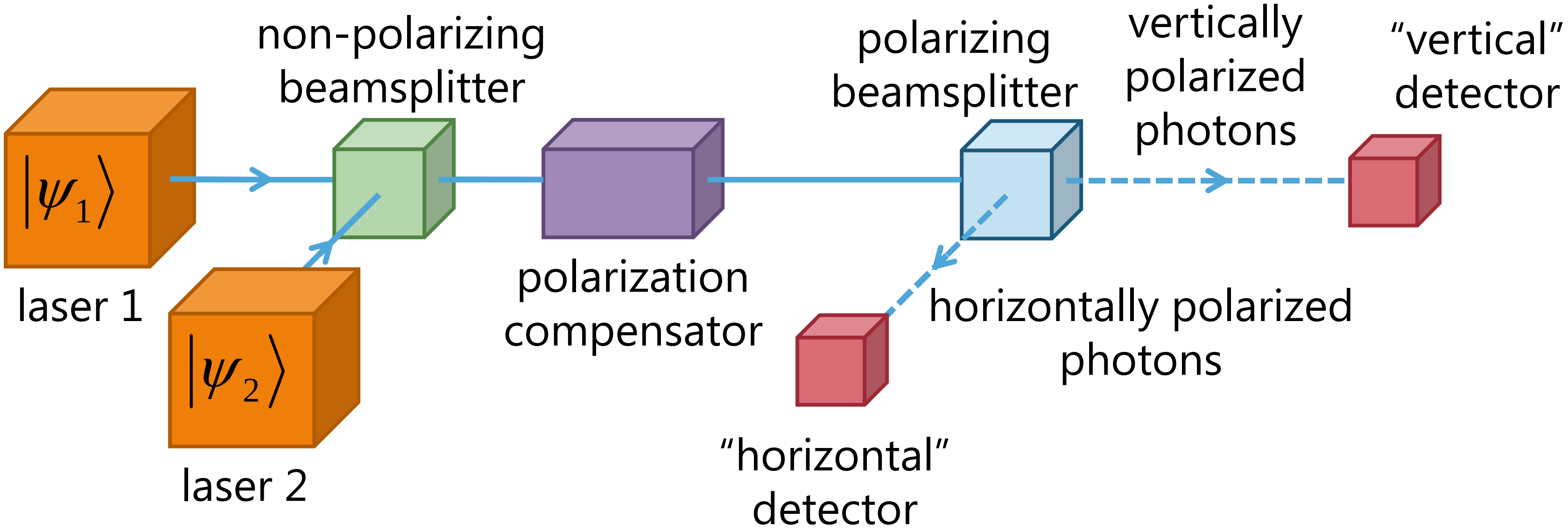
“horizontal” detector

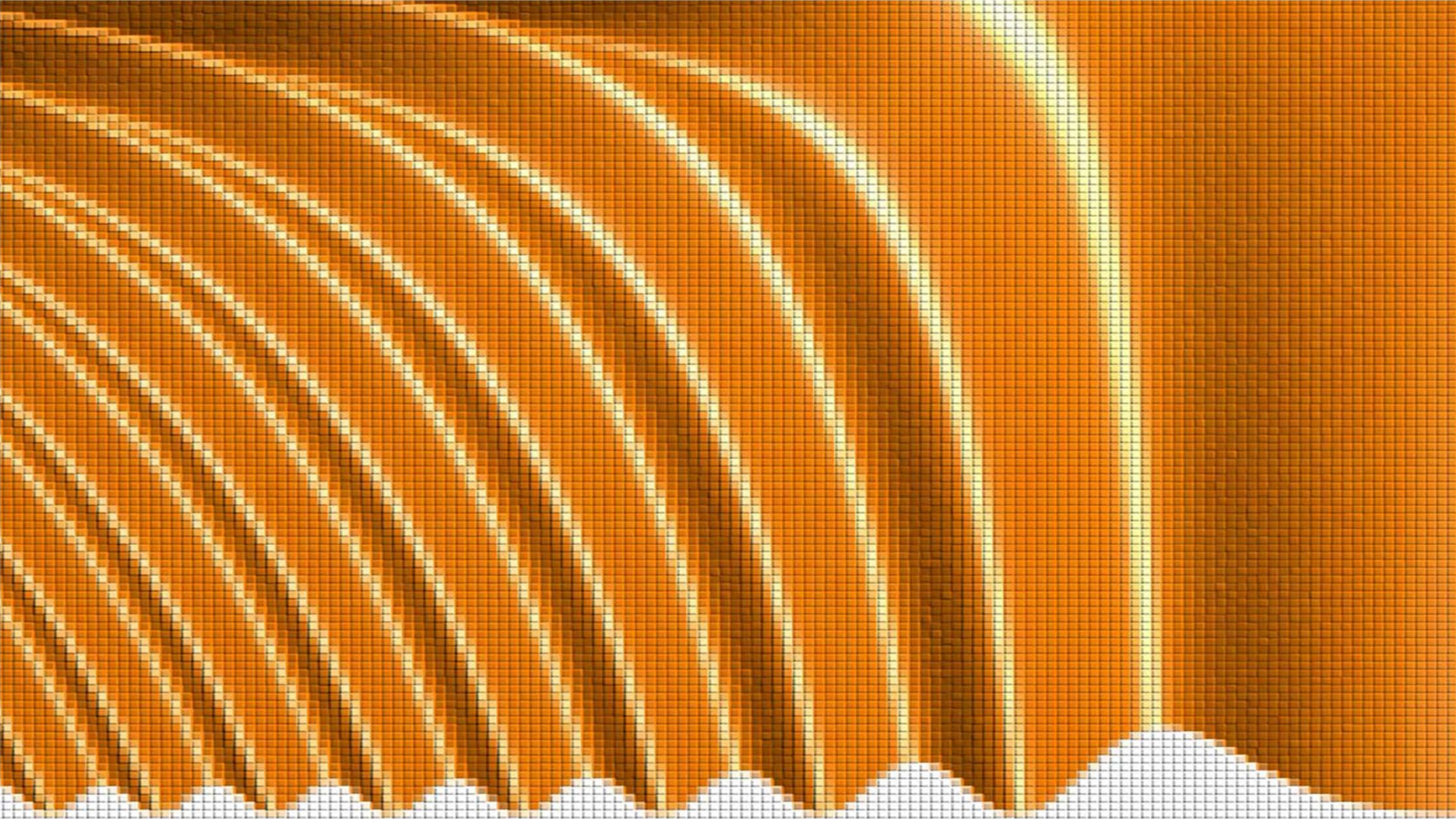
Mixed states

So the state of these photons is described differently

as a “mixed state”

in contrast to a “pure state” like $|\psi\rangle = a_H |H\rangle + a_V |V\rangle$



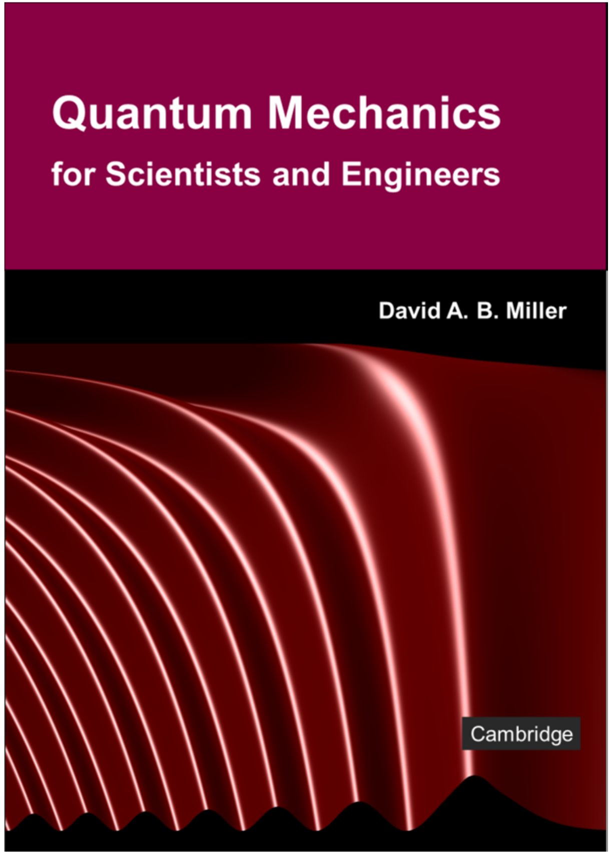


17.2 Mixed states and the density matrix

Slides: Video 17.2.4 Mixed states with potential wells

Text reference: Quantum Mechanics for Scientists and Engineers

Section 14.1 (second part)



Mixed states and the density matrix

Mixed states with potential wells

Quantum mechanics for scientists and engineers

David Miller

Pure states for potential wells

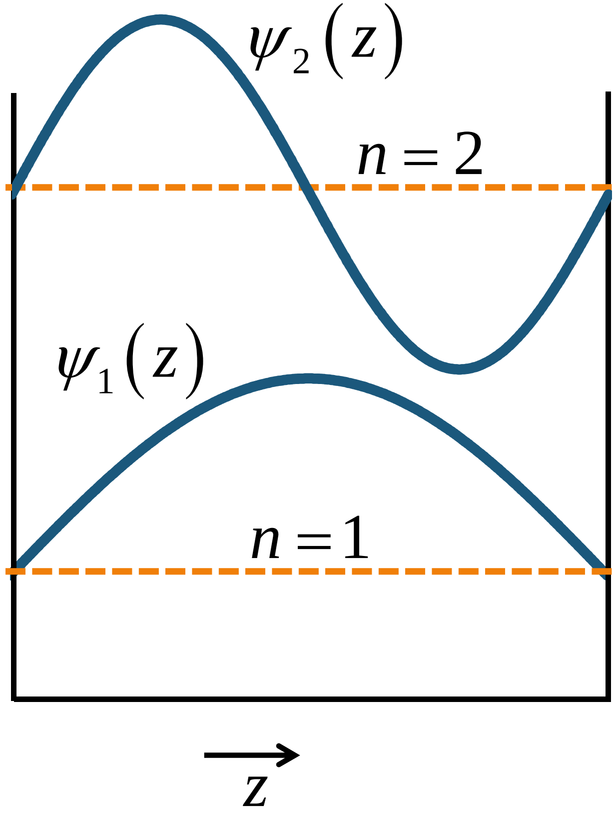
Suppose, for some particle with mass

- we have a potential well
  - such as the "infinite" one-dimensional potential well

We put it in a pure state

- that is an equal linear superposition of the lowest two states of this well

$$|\psi\rangle = \left(1/\sqrt{2}\right)\left(|\psi_1\rangle + |\psi_2\rangle\right)$$



Pure states for potential wells

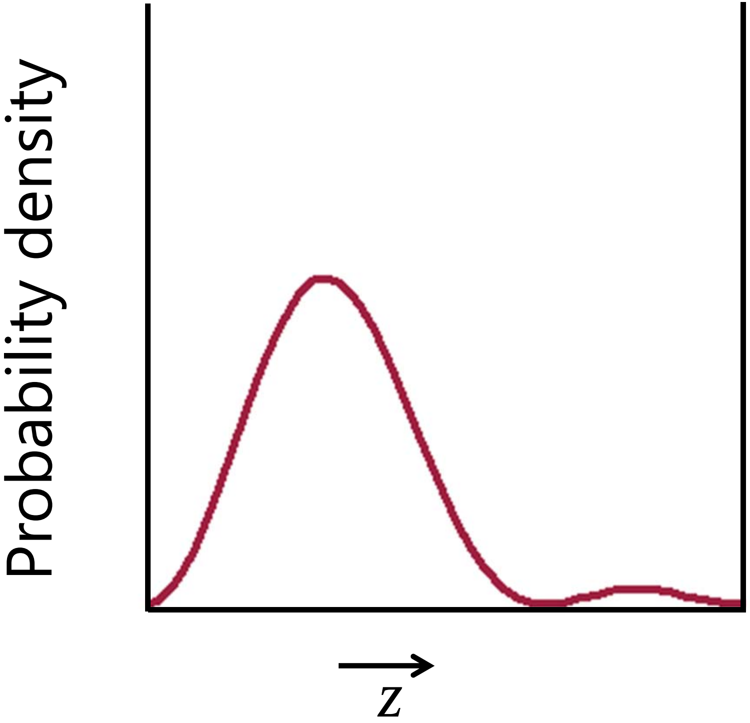
In such a superposition

$$|\psi\rangle = \left(1/\sqrt{2}\right)\left(|\psi_1\rangle + |\psi_2\rangle\right)$$

the position of this particle

will oscillate back and forwards

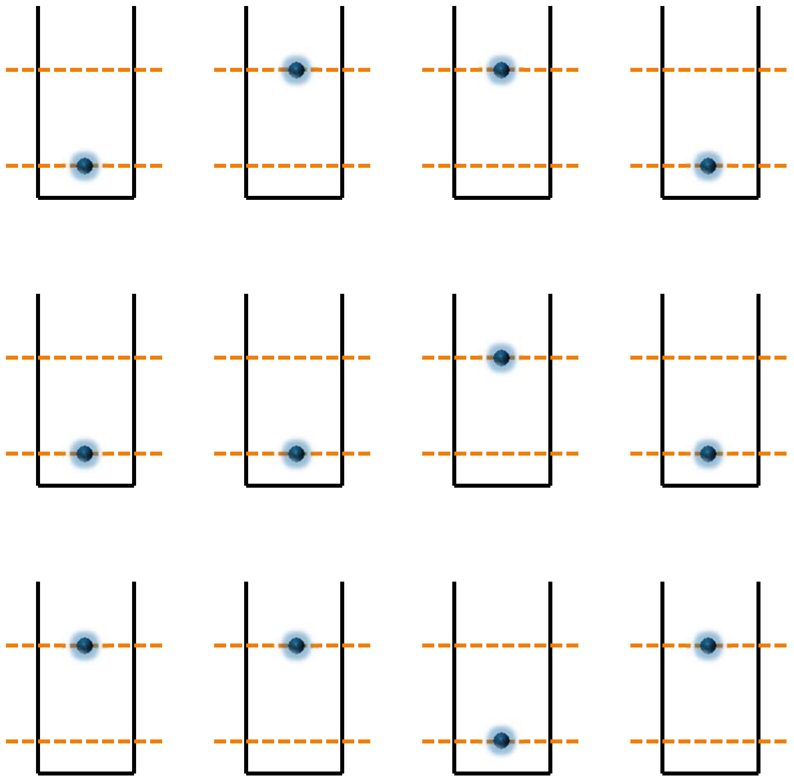
because of the different time-evolution factors for the two energy eigenstates



Mixed states for potential wells

Suppose instead we take an ensemble of identical potential wells

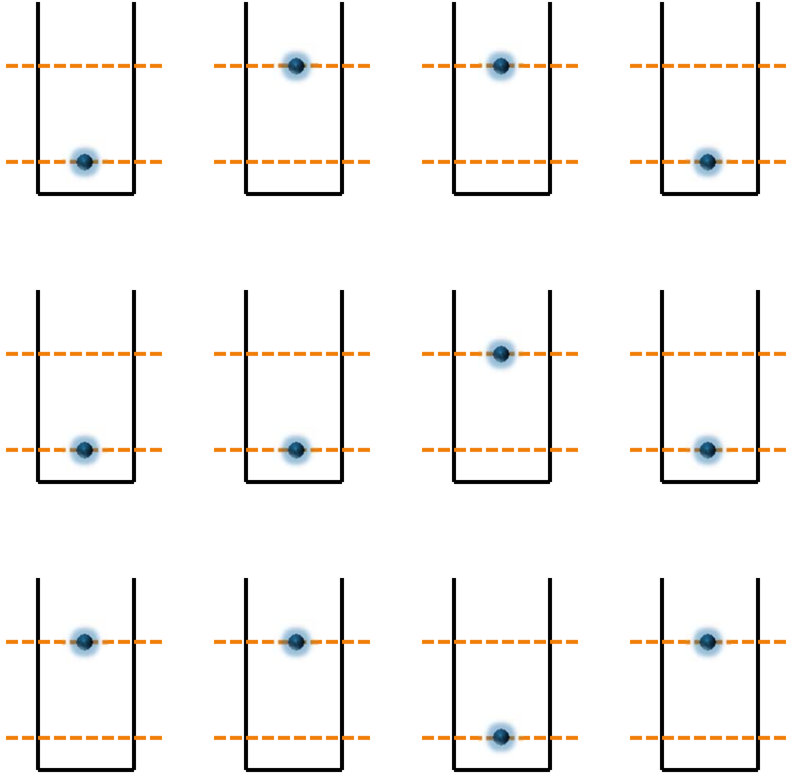
- and randomly prepare
  - half of them with the particle in the lowest state
  - and half of them with the particle in the second state



Mixed states for potential wells

Statistically, since we do not know which wells are which

- at least before performing any measurements
  - each of these wells is in a mixed state
    - with 50% probability of being in either the first or second state



Mixed states for potential wells

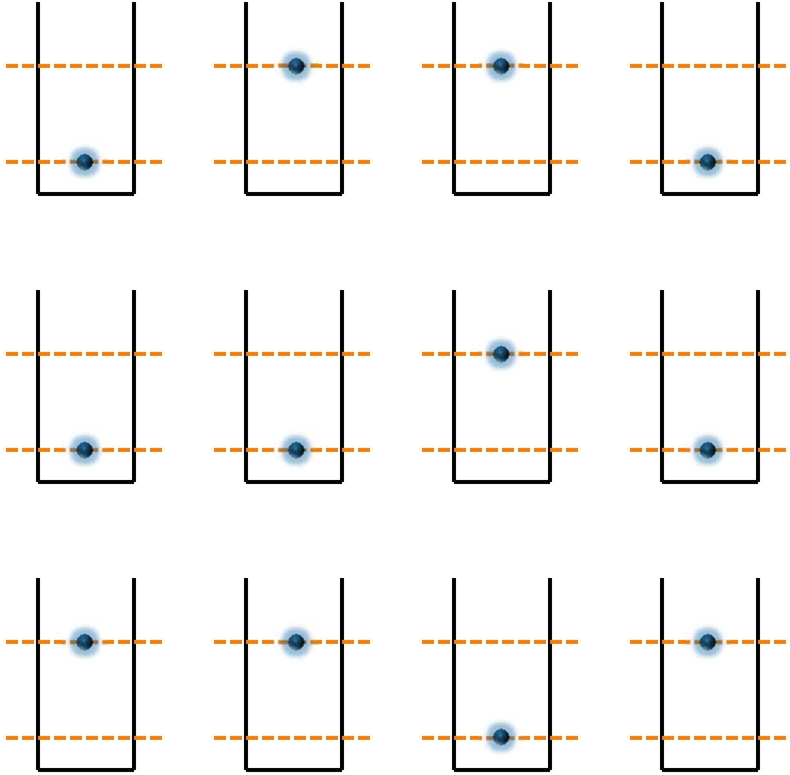
Now we evaluate the expectation value $\langle z \rangle$

of the $\hat{z}$ position operator for each potential well

In each well

$\langle z \rangle$ evaluates to the position of the center of the well

since each of these wavefunctions is equally balanced about the center



Mixed states for potential wells

The "ensemble average" $\overline{\langle z \rangle}$

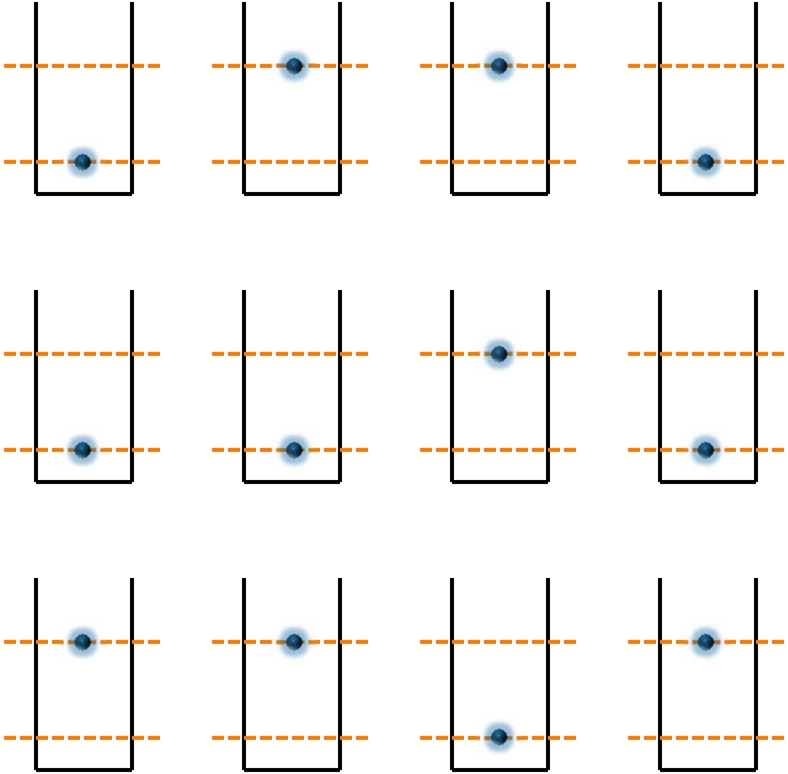
of expectation values from the different wells

is also zero

and there is no oscillation in time

Again it would not be correct simply to write the mixed state

as a linear combination of the form $b_1|\psi_1\rangle + b_2|\psi_2\rangle$



Pure state in skewed wells

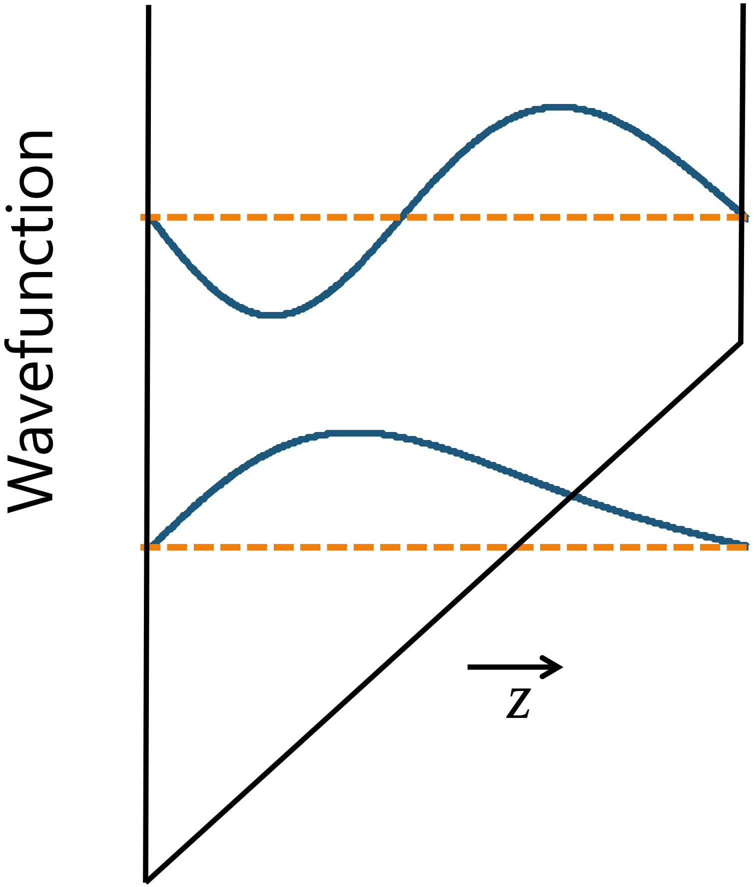
Suppose each well is skewed

- by applying electric field in the z direction
  - for a particle like an electron in the well

Then $\langle z \rangle$ is different for the first and second states of the well

- with $\langle z \rangle = z_1$ for the first state
- and $\langle z \rangle = z_2$ for the second state

For the pure state, we still expect oscillation



Mixed states for skewed wells

For a mixed state

with probabilities P_1 and P_2 respectively that

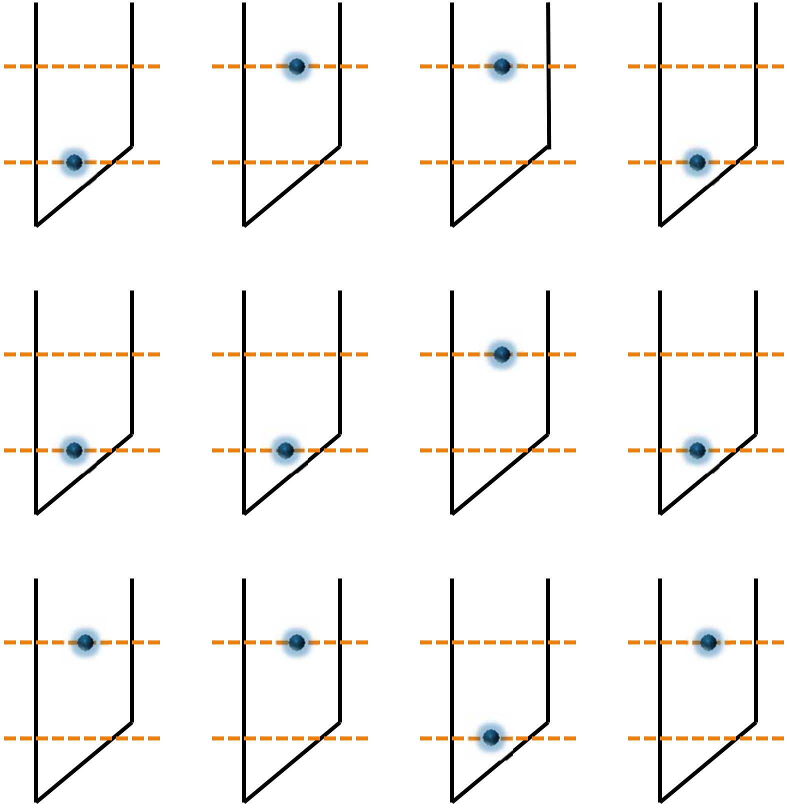
we had prepared a given well in the first or second state

we would still have no oscillation

The ensemble average value of the measured position

would now be

$$\overline{\langle z \rangle} = P_1 z_1 + P_2 z_2 \equiv \sum_{j=1}^{2} P_j \langle \psi_j | \hat{z} | \psi_j \rangle$$



Mixed state ensemble average expectation value

Generalizing from $\overline{\langle z \rangle} = P_1 z_1 + P_2 z_2 \equiv \sum_{j=1}^{2} P_j \langle \psi_j | \hat{z} | \psi_j \rangle$

we expect the ensemble average expectation value

for some operator $\hat{A}$ corresponding to an observable quantity

can be written $\overline{\langle A \rangle} = \sum_j P_j \langle \psi_j | \hat{A} | \psi_j \rangle$

for some set of different quantum mechanical state preparations $|\psi_j\rangle$

made with respective probabilities P_j

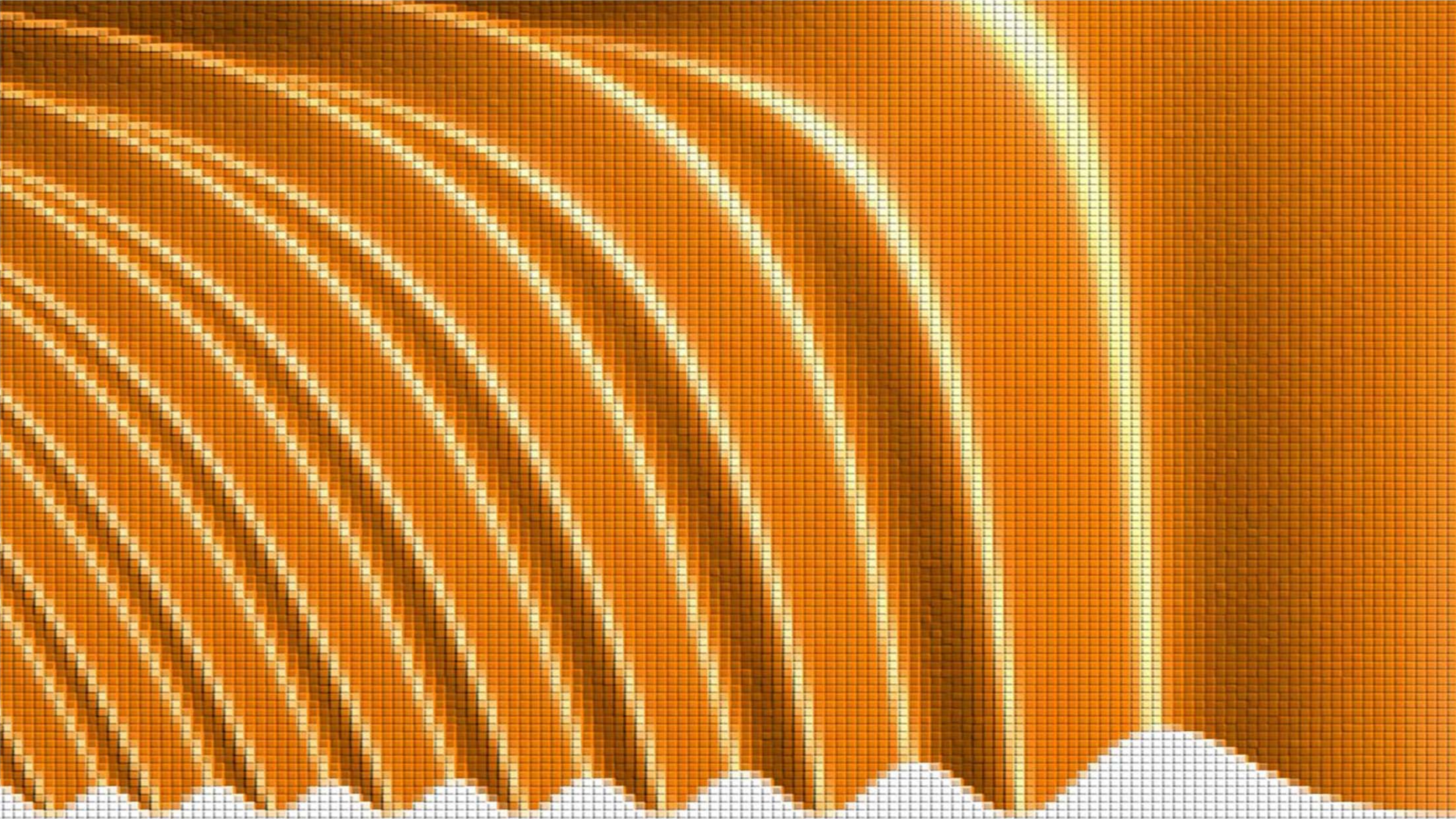
Mixed states

In considering mixed states

- where we are saying that the quantum mechanical state could be any of a set of different possibilities $\left|\psi_j\right\rangle$
  - with respective probabilities P_j
    - the different $\left|\psi_j\right\rangle$ need not be orthogonal

For example, we could be considering several different polarization states close to one another in angle

- perhaps from some fluctuation in time in the precise output polarization of some laser
  - giving a mixed state of many different possible similar but not identical polarizations

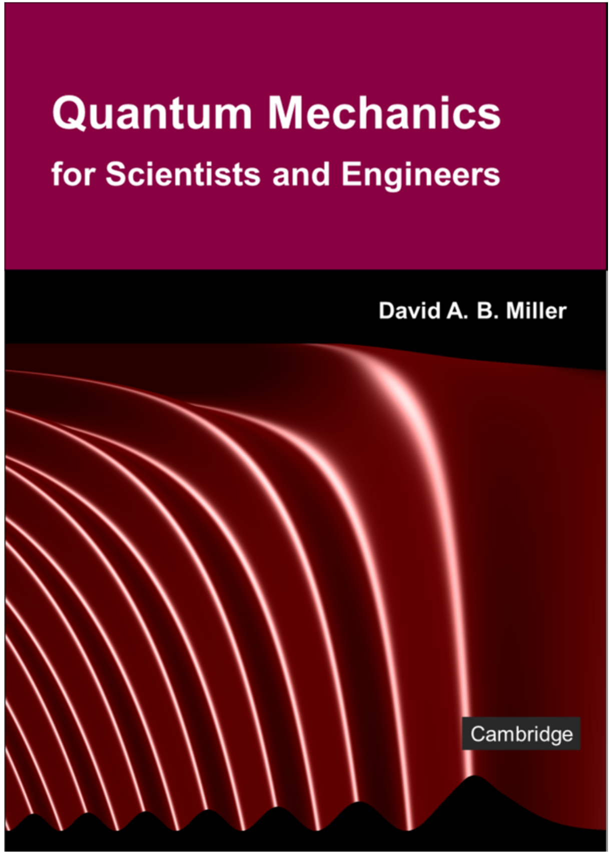


17.2 Mixed states and the density matrix

Slides: Video 17.2.6 Representing mixed states

Text reference: Quantum Mechanics for Scientists and Engineers

Sections 14.2 and 14.3 through Eq. 14.8



Mixed states and the density matrix

Representing mixed states

Quantum mechanics for scientists and engineers

David Miller

Density operator

However we are going to represent the mixed state

it must obviously contain the probabilities P_j

and the pure states $\left|\psi_j\right\rangle$

but it must not simply be a linear combination of the states

The structure we propose instead is the density operator

$$\boxed{\rho = \sum_j P_j \left|\psi_j\right\rangle\left\langle\psi_j\right|}$$

This is an operator because

it contains the outer products $\left|\psi_j\right\rangle\left\langle\psi_j\right|$ of state vectors

Density operator

In $\rho = \sum_j P_j \left|\psi_j\right\rangle\left\langle\psi_j\right|$

- we deliberately leave the "hat" off this operator
  - to emphasize that its physical meaning and use
    - are quite different from other operators we have considered

ρ is not an operator representing some physical observable

- Rather, ρ is representing the state
  - in general, a mixed state
    - of the system

Density operator

If ρ is a useful way of representing the mixed state

- it must allow us to calculate quantities like
  - the ensemble average measured value $\overline{\langle A \rangle}$
    - for any physical observable with operator $\hat{A}$

If we can evaluate $\overline{\langle A \rangle}$ for such any physical observable

- then ρ will be the most complete way we can have of describing this mixed quantum mechanical state
  - because it will tell us the value we will get of any measurable quantity
    - to within our underlying statistical uncertainties

Density matrix and ensemble average values

To understand the properties of the density operator

we write it in terms of a complete orthonormal basis $|\phi_m\rangle$

First we expand each of the pure states $|\psi_j\rangle$ in this set

obtaining $|\psi_j\rangle = \sum_u c_u^{(j)} |\phi_u\rangle$

Then we use this expansion and its adjoint

in our definition $\rho = \sum_j P_j |\psi_j\rangle\langle\psi_j|$ to obtain

$$\rho = \sum_j P_j \left(\sum_u c_u^{(j)} |\phi_u\rangle \right) \left(\sum_v \left(c_v^{(j)} \right)^* \langle\phi_v| \right) = \sum_{u,v} \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right) |\phi_u\rangle\langle\phi_v|$$

Density matrix and ensemble average values

Written $\rho = \sum_{u,v}\left(\sum_j P_j c_u^{(j)}\left(c_v^{(j)}\right)^*\right)|\phi_u\rangle\langle\phi_v| \equiv \sum_{u,v}\rho_{uv}|\phi_u\rangle\langle\phi_v|$

the matrix representation of ρ is now clear

We have for a matrix element in this basis

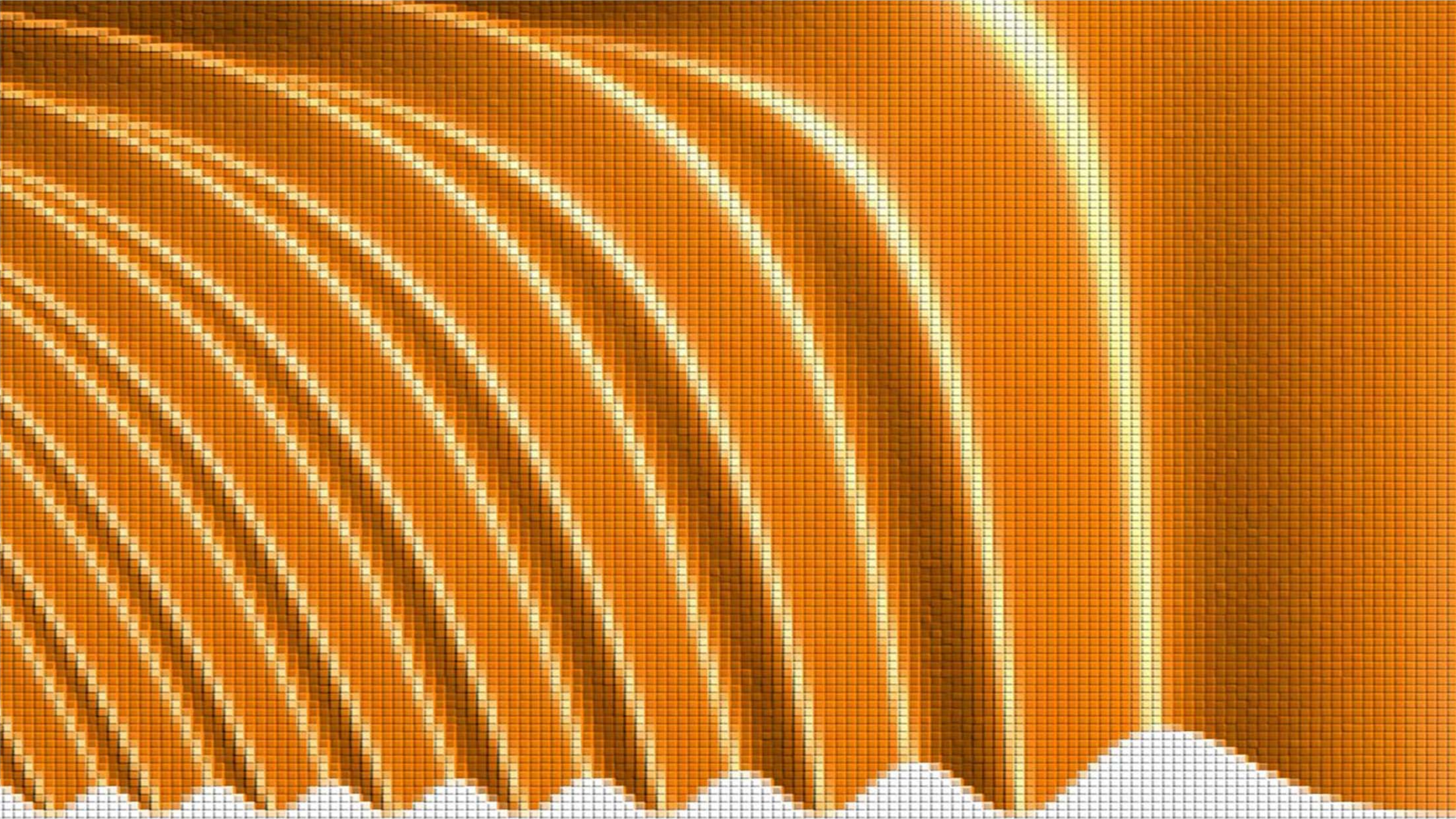
$$\rho_{uv} \equiv \langle\phi_u|\rho|\phi_v\rangle = \sum_j P_j c_u^{(j)}\left(c_v^{(j)}\right)^* \equiv \overline{c_u c_v^*}$$

Here we have also introduced and defined the idea of

the ensemble average of the coefficient product $\overline{c_u c_v^*}$

We now usually talk of ρ as the density matrix

with matrix elements ρ_{uv}

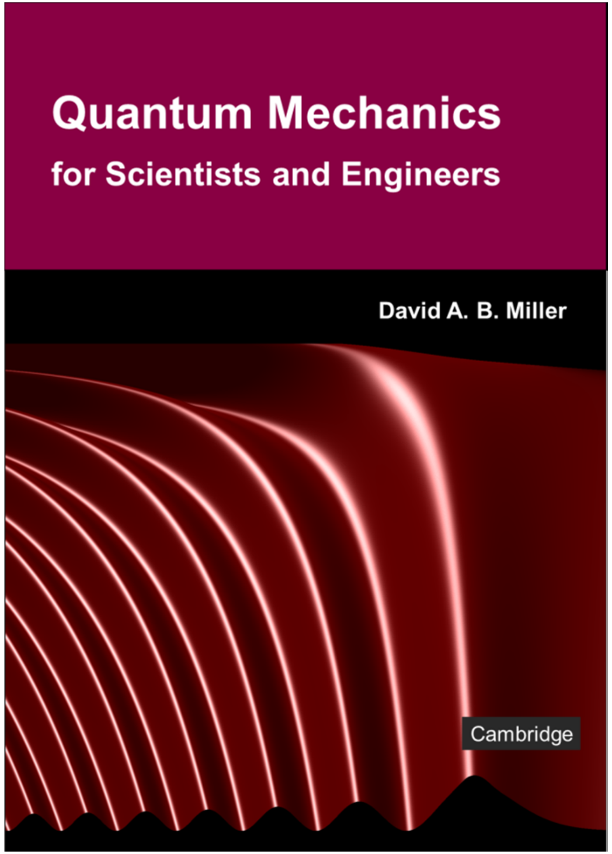


17.2 Mixed states and the density matrix

Slides: Video 17.2.8 Properties of the density matrix

Text reference: Quantum Mechanics for Scientists and Engineers

Section 14.3 (remainder)



Mixed states and the density matrix

Properties of the density matrix

Quantum mechanics for scientists and engineers

David Miller

Properties of the density matrix

We can deduce several properties from our definition of the density matrix

First, the density matrix is Hermitian

Explicitly $\rho_{vu} \equiv \sum_j P_j c_v^{(j)} \left(c_u^{(j)}\right)^* = \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)}\right)^*\right)^* = \rho_{uv}^*$

Because the density matrix is Hermitian

so also is the density operator

since the density matrix is just a representation of the density operator

Properties of the density matrix

Second, the diagonal elements ρ_{mm}

give us the probabilities of finding the system in a specific one of the states $|\phi_m\rangle$

$c_m^{(j)}(c_m^{(j)})^* \equiv |c_m^{(j)}|^2$ is the probability for a pure state j

that we will find the system in state m

Hence adding these up with probabilities P_j

gives the overall probability of finding the system in state m in the mixed state

(The off-diagonal elements are a measure of the "coherence" between different states and we will return to discuss this later)

Properties of the density matrix

Third, the sum of the diagonal elements of the density matrix is unity

i.e., remembering that we can formally write the sum of the diagonal elements of some matrix or operator

as the trace (Tr) of the matrix or operator

$$Tr\left(\rho\right)=\sum_m \rho_{mm} = \sum_m \sum_j P_j \left|c_m^{(j)}\right|^2 = \sum_j P_j \sum_m \left|c_m^{(j)}\right|^2 = \sum_j P_j = 1$$

because the state $\left|\psi_j\right\rangle$ is normalized (so $\sum_m |c_m^{(j)}|^2 = 1$)

and the sum of all the probabilities P_j of the various states $\left|\psi_j\right\rangle$ in the mixed state must be 1

Ensemble average value from the density matrix

Consider an operator $\hat{A}$ for some physical observable

and specifically consider the product $\rho\hat{A}$

$$\rho\hat{A} = \sum_{u.v} \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right) |\phi_u\rangle\langle\phi_v| \hat{A}$$

We can therefore write some diagonal element of the resulting matrix as

$$\left\langle \phi_q \middle| \rho\hat{A} \middle| \phi_q \right\rangle = \sum_{u.v} \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right) \langle \phi_q | \phi_u \rangle \left\langle \phi_v \middle| \hat{A} \middle| \phi_q \right\rangle$$

$$= \sum_{u.v} \left(\sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \right) \delta_{qu} \left\langle \phi_v \middle| \hat{A} \middle| \phi_q \right\rangle = \sum_v \sum_j P_j c_q^{(j)} \left(c_v^{(j)} \right)^* \left\langle \phi_v \middle| \hat{A} \middle| \phi_q \right\rangle$$

Ensemble average value from the density matrix

Then the sum of the all of these diagonal elements

$$\left\langle \phi_q \middle| \rho\hat{A} \middle| \phi_q \right\rangle = \sum_v \sum_j P_j c_q^{(j)} \left(c_v^{(j)} \right)^* \left\langle \phi_v \middle| \hat{A} \middle| \phi_q \right\rangle$$

is $\sum_q \left\langle \phi_q \middle| \rho\hat{A} \middle| \phi_q \right\rangle = \sum_j P_j \left(\sum_v \left(c_v^{(j)} \right)^* \left\langle \phi_v \right| \right) \hat{A} \left(\sum_q c_q^{(j)} \left| \phi_q \right\rangle \right)$

$$= \sum_j P_j \left\langle \psi_j \middle| \hat{A} \middle| \psi_j \right\rangle$$

Note this is the same as the ensemble average value $\overline{\langle A \rangle}$

of the expectation value of the operator $\hat{A}$

for this mixed state, as written before

Ensemble average value from the density matrix

Hence we have a key result of density matrix theory

$$\boxed{\overline{\langle A \rangle} = Tr\left(\rho \hat{A}\right)}$$

The density matrix can be used to deduce any measurable ensemble average property of a mixed state

Hence the density matrix gives a full description of a mixed state

Note that this result

is independent of the basis used to calculate the trace

The basis $|\phi_m\rangle$ could be any set

that is complete for the problem of interest

Ensemble average value from the density matrix

Note also that

if we have the system in a pure state $|\psi\rangle$

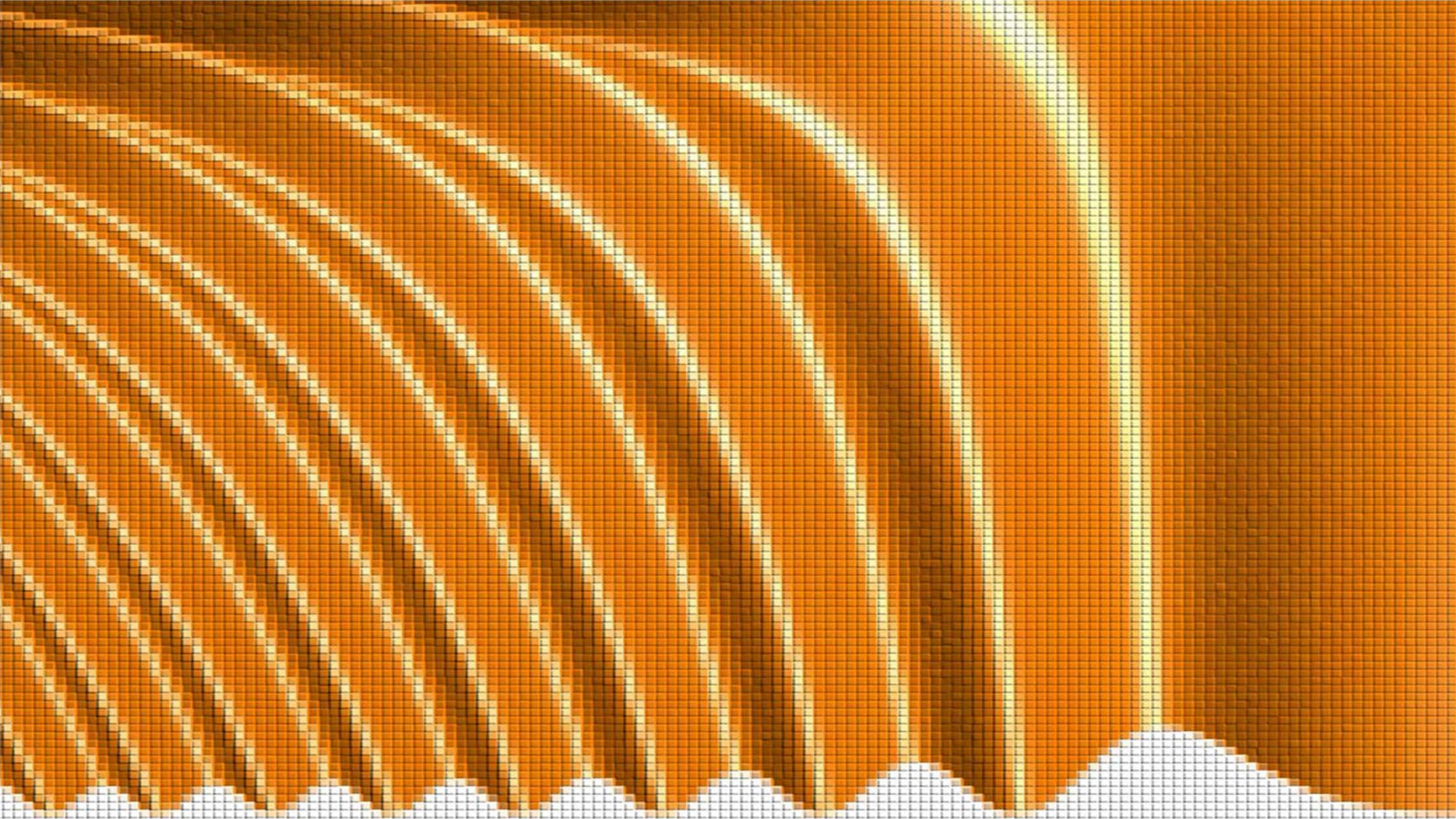
in which case $P = 1$ for that state

and is zero for any other pure state

then we recover the usual result for the expectation value

i.e., $Tr\left(\rho\hat{A}\right) = \langle\psi|\hat{A}|\psi\rangle = \langle A\rangle$

so the density matrix description gives the correct answers for pure or mixed states

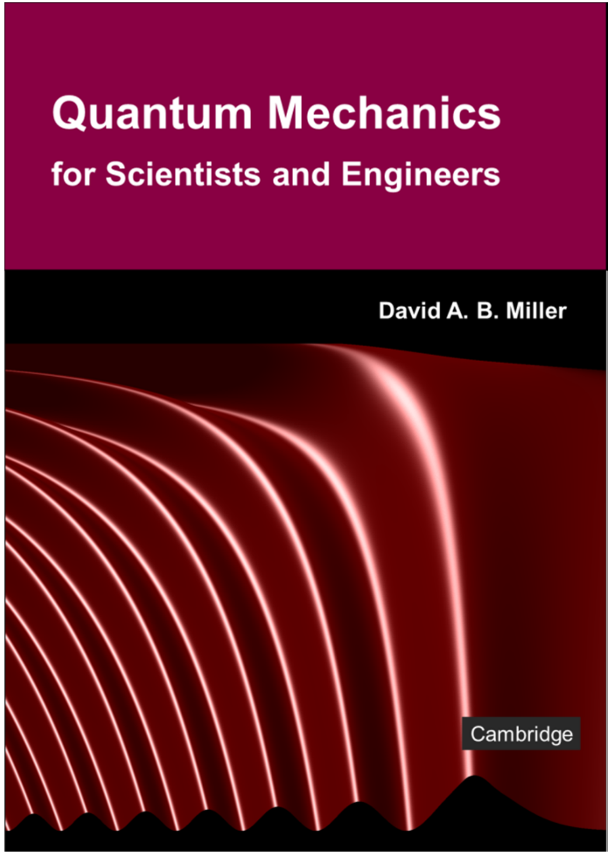


17.2 Mixed states and the density matrix

Slides: Video 17.2.10 Time-evolution of the density matrix

Text reference: Quantum Mechanics for Scientists and Engineers

Section 14.4



Mixed states and the density matrix

Time-evolution of the density matrix

Quantum mechanics for scientists and engineers

David Miller

Time-evolution of the density matrix

We can directly calculate the time-evolution of the density matrix

We start with the Schrödinger equation for any particular one of the pure states in our mixed state

i.e., $\hat{H}\left|\psi_j\right\rangle = i\hbar\frac{\partial}{\partial t}\left|\psi_j\right\rangle$

and we substitute using $\left|\psi_j\right\rangle = \sum_u c_u^{(j)}\left|\phi_u\right\rangle$

to obtain $i\hbar\sum_n \frac{\partial c_n^{(j)}(t)}{\partial t}\left|\phi_n\right\rangle = \sum_n c_n^{(j)}(t)\hat{H}\left|\phi_n\right\rangle$

Here all the time dependence of the state is in the $c_n^{(j)}(t)$

Time-evolution of the density matrix

Now operating from the left with $\langle \phi_m |$

we have $$\langle \phi_m | i\hbar \sum_n \frac{\partial c_n^{(j)}(t)}{\partial t} | \phi_n \rangle = \langle \phi_m | \sum_n c_n^{(j)}(t) \hat{H} | \phi_n \rangle$$

i.e., $$i\hbar \frac{\partial c_m^{(j)}(t)}{\partial t} = \sum_n c_n^{(j)}(t) H_{mn}$$

where $H_{mn} = \langle \phi_m | \hat{H} | \phi_n \rangle$

is a matrix element of the Hamiltonian

Time-evolution of the density matrix

We can take the complex conjugate of both sides of

$$i\hbar \frac{\partial c_m^{(j)}(t)}{\partial t} = \sum_n c_n^{(j)}(t) H_{mn}$$

Noting that $\hat{H}$ is Hermitian, i.e., $H_{mn}^* = H_{nm}$ we have

$$-i\hbar \frac{\partial \left(c_m^{(j)}(t)\right)^*}{\partial t} = \sum_n \left(c_n^{(j)}(t)\right)^* H_{nm}$$

or trivially changing indices

$$-i\hbar \frac{\partial \left(c_n^{(j)}(t)\right)^*}{\partial t} = \sum_s \left(c_s^{(j)}(t)\right)^* H_{sn}$$

Time-evolution of the density matrix

But from our previous result for

the elements of the density matrix $\rho_{uv} = \sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^*$

we can take the time derivative to obtain

$$\frac{\partial \rho_{mn}}{\partial t} = \sum_j P_j \left(c_m^{(j)} \frac{\partial \left(c_n^{(j)} \right)^*}{\partial t} + \left(c_n^{(j)} \right)^* \frac{\partial c_m^{(j)}}{\partial t} \right)$$

Now we can substitute using our previous results

$$i\hbar \frac{\partial c_m^{(j)}(t)}{\partial t} = \sum_q c_q^{(j)}(t) H_{mq} \qquad -i\hbar \frac{\partial \left(c_n^{(j)}(t) \right)^*}{\partial t} = \sum_s \left(c_s^{(j)}(t) \right)^* H_{sn}$$

(We changed the summation index to q in the first result)

Time-evolution of the density matrix

Hence $\dfrac{\partial \rho_{mn}}{\partial t} = \sum_j P_j \left(c_m^{(j)} \dfrac{\partial \left(c_n^{(j)} \right)^*}{\partial t} + \left(c_n^{(j)} \right)^* \dfrac{\partial c_m^{(j)}}{\partial t} \right)$ becomes

$$\frac{\partial \rho_{mn}}{\partial t} = \sum_j P_j \left(\frac{i}{\hbar} c_m^{(j)} \sum_q \left(c_q^{(j)} \right)^* H_{qn} - \frac{i}{\hbar} \left(c_n^{(j)} \right)^* \sum_s c_s^{(j)} H_{ms} \right)$$

$$= \frac{i}{\hbar} \left\{ \sum_q \left[\sum_j P_j c_m^{(j)} \left(c_q^{(j)} \right)^* \right] H_{qn} - \sum_s H_{ms} \left[\sum_j P_j c_s^{(j)} \left(c_n^{(j)} \right)^* \right] \right\}$$

Note, though, that by definition $\rho_{uv} \equiv \sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^*$

so we can substitute for the terms in […]

Time-evolution of the density matrix

Hence we have

$$\frac{\partial \rho_{mn}}{\partial t} = \frac{i}{\hbar}\left(\sum_q \rho_{mq} H_{qn} - \sum_s H_{ms} \rho_{sn}\right) = \frac{i}{\hbar}\left(\left(\rho \hat{H}\right)_{mn} - \left(\hat{H}\rho\right)_{mn}\right) = \frac{i}{\hbar}\left[\rho, \hat{H}\right]_{mn}$$

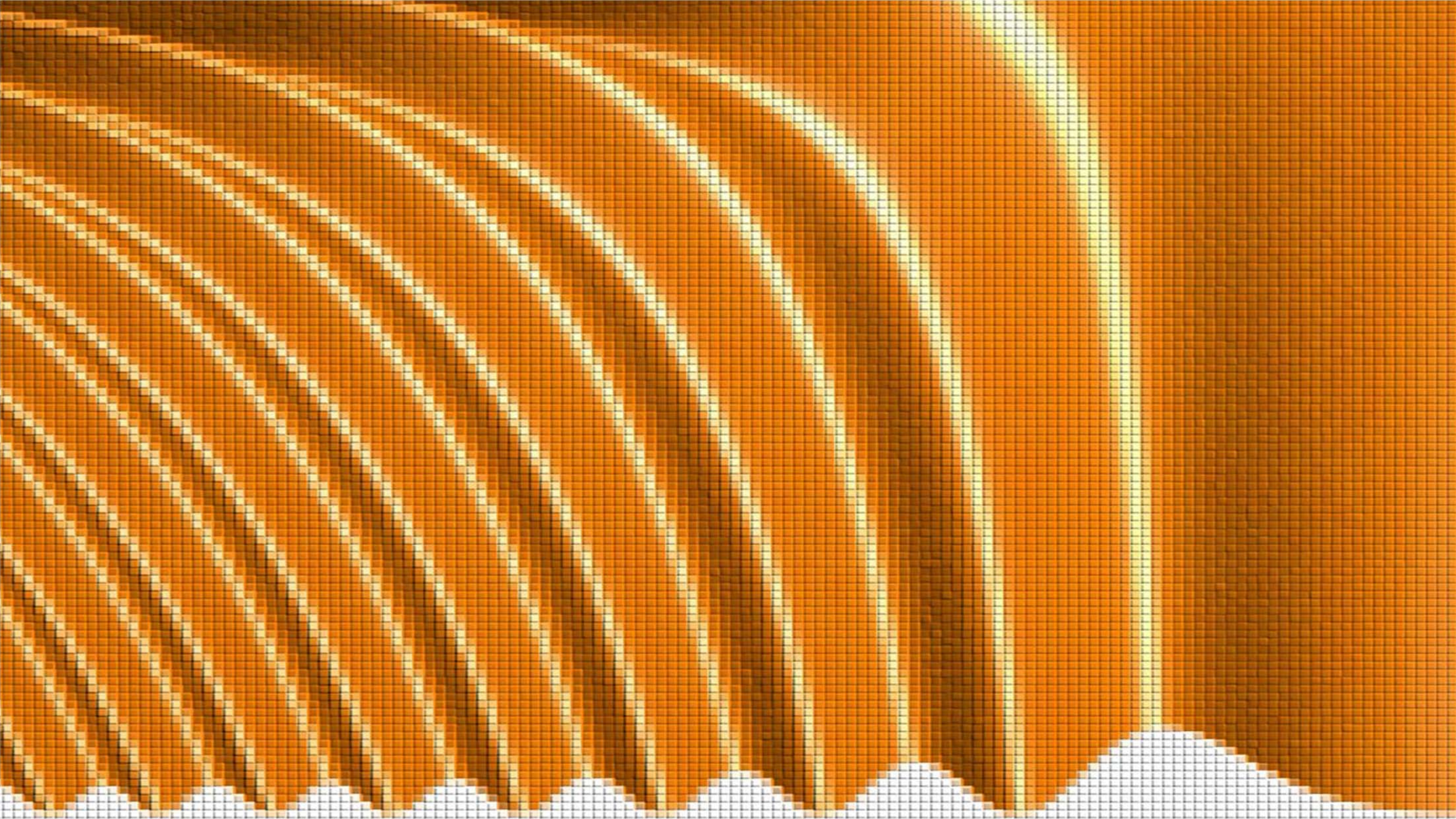
Since this is true for every matrix element *mn*

we have

$$\boxed{\frac{\partial \rho}{\partial t} = \frac{i}{\hbar}\left[\rho, \hat{H}\right]}$$

which tells us how the mixed state evolves in time

and is a key result for the density matrix

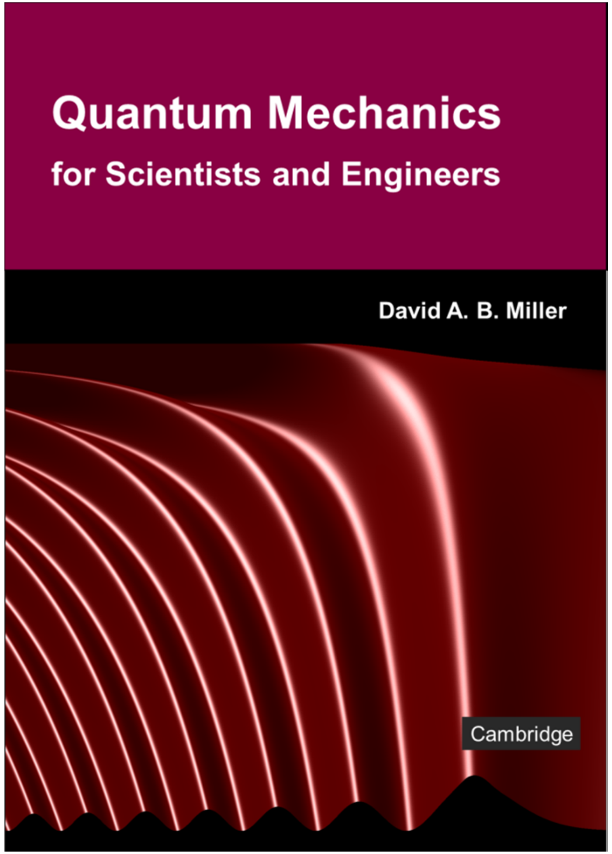


17.3 The density matrix and optical absorption

Slides: Video 17.3.1 Induced dipole and a two-level system

Text reference: Quantum Mechanics for Scientists and Engineers

Section 14.5 through Eq. 14.35



The density matrix and optical absorption

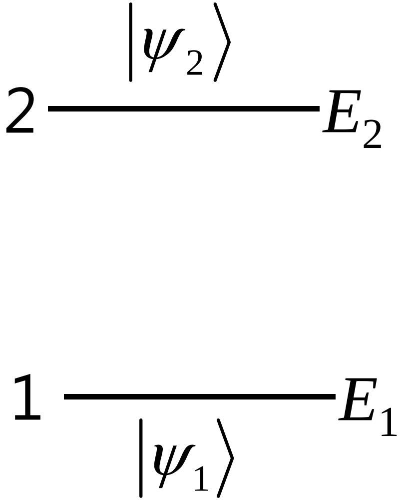
Induced dipole and a two-level system

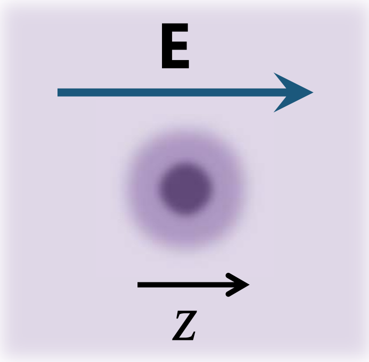
Quantum mechanics for scientists and engineers

David Miller

Two-level system

- Take a two-level system
  - with energies E_1 and E_2
    - and eigenfunctions $|\psi_1\rangle$ and $|\psi_2\rangle$
- Presume the system is much smaller than an optical wavelength
  - so an incident optical field **E** will be uniform across the system
    - and take **E** to be polarized in the z direction
      - with magnitude E



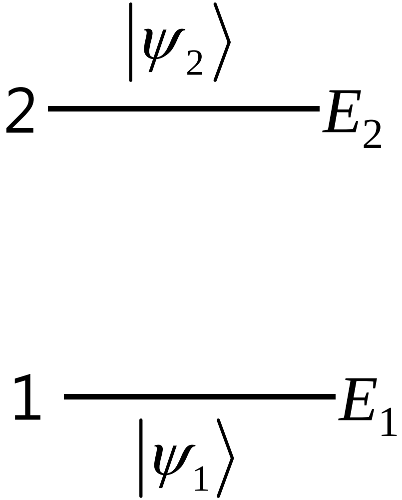


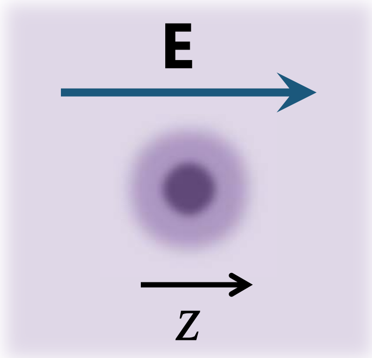
Two-level system

Here we will just treat the interaction with the electric field semiclassically

We take an "electric dipole" interaction

- between the light
  - and the electron in the system
    - so that the energy change
      - on displacing by an amount z
        - is $e\mathrm{E}z$





Interaction of light with a two-level system

Hence we can take the (semiclassical) perturbing Hamiltonian as

$$\hat{H}_p = e\mathrm{E}z \equiv -\mathrm{E}\hat{\mu}$$

where $\hat{\mu}$ is what we will call

the electric dipole operator

with matrix elements

$$\mu_{mn} = -e\left\langle \psi_m \middle| z \middle| \psi_n \right\rangle$$

so that the matrix elements of the perturbing Hamiltonian become

$$\left(\hat{H}_p\right)_{mn} \equiv H_{pmn} = -\mathrm{E}\mu_{mn}$$

Interaction of light with a two-level system

We choose the states $|\psi_1\rangle$ and $|\psi_2\rangle$

to have definite parity in the z direction

so with our definition $\mu_{mn} = -e\langle\psi_m|z|\psi_n\rangle$

$$\mu_{11} = \mu_{22} = 0$$

and hence with our definition $H_{pmn} = -\mathrm{E}\mu_{mn}$

$$H_{p11} = H_{p22} = 0$$

We are free to choose the relative phase of the two wavefunctions such that μ_{12} is real

so that we have

$$\mu_{12} = \mu_{21} \equiv \mu_d$$

Interaction of light with a two-level system

Hence the dipole operator can be written $\hat{\mu} = \begin{bmatrix} 0 & \mu_d \\ \mu_d & 0 \end{bmatrix}$

and the perturbing Hamiltonian is $\hat{H}_p = \begin{bmatrix} 0 & -\mathrm{E}\mu_d \\ -\mathrm{E}\mu_d & 0 \end{bmatrix}$

The unperturbed Hamiltonian $\hat{H}_o$

is just a 2 x 2 diagonal matrix on this basis

with E_1 and E_2 as the diagonal elements

So the total Hamiltonian is $\hat{H} = \hat{H}_o + \hat{H}_p = \begin{bmatrix} E_1 & -\mathrm{E}\mu_d \\ -\mathrm{E}\mu_d & E_2 \end{bmatrix}$

Interaction of light with a two-level system

The density matrix is also a 2 x 2 matrix

because there are only two basis states under consideration here

and in general we can write it as

$$\rho = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$$

for this two-level system

The dipole of the system

We have not yet defined the system's state

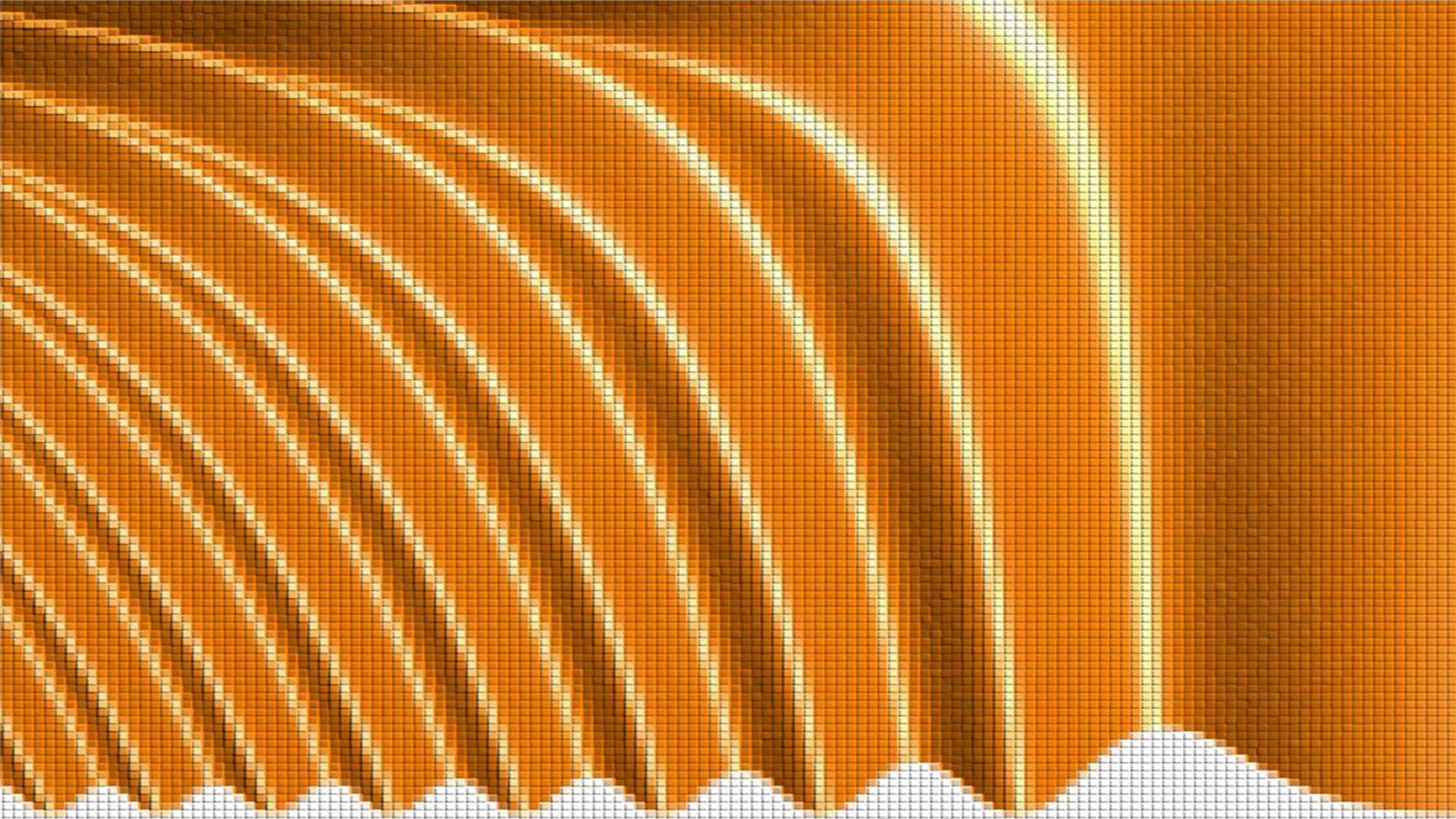
but we can use $\overline{\langle A \rangle} = Tr\left(\rho \hat{A}\right)$ to write

$$\overline{\langle \mu \rangle} = Tr\left(\rho \hat{\mu}\right)$$

Using $\hat{\mu} = \begin{bmatrix} 0 & \mu_d \\ \mu_d & 0 \end{bmatrix}$ and $\rho = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}$ we have

$$\rho \hat{\mu} = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \begin{bmatrix} 0 & \mu_d \\ \mu_d & 0 \end{bmatrix} = \begin{bmatrix} \rho_{12}\mu_d & \rho_{11}\mu_d \\ \rho_{22}\mu_d & \rho_{21}\mu_d \end{bmatrix}$$

Hence $\overline{\langle \mu \rangle} = \mu_d \left(\rho_{12} + \rho_{21}\right)$

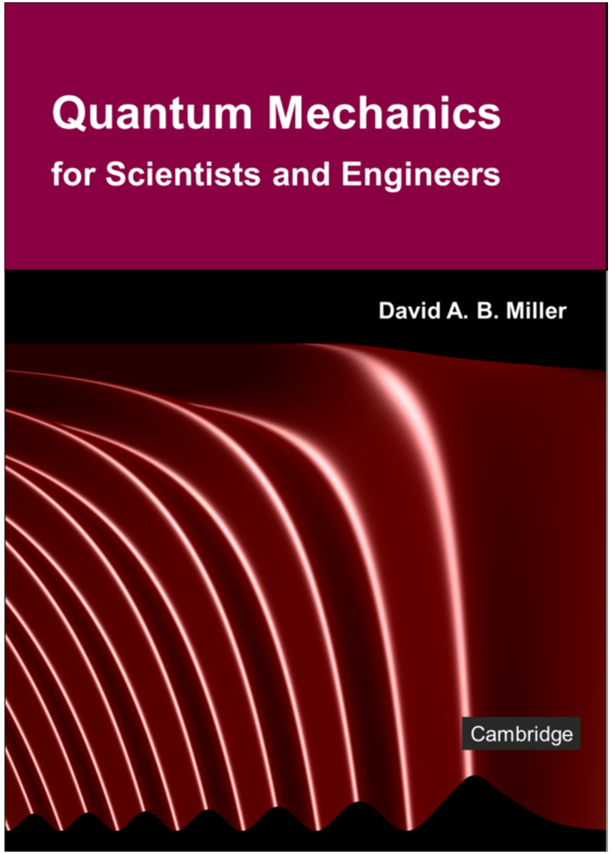


17.3 The density matrix and optical absorption

Slides: Video 17.3.3 Behavior of the density matrix in time

Text reference: Quantum Mechanics for Scientists and Engineers

Section 14.5 from Eq. 14.35 through the paragraph after Eq. 14.40



The density matrix and optical absorption

Behavior of the density matrix in time

Quantum mechanics for scientists and engineers

David Miller

Behavior of the density matrix in time

We have, from $\partial\rho/\partial t = (i/\hbar)\left[\rho,\hat{H}\right]$ with the definitions

$$\rho = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \text{ and } \hat{H} = \hat{H}_o + \hat{H}_p = \begin{bmatrix} E_1 & -\mathrm{E}\mu_d \\ -\mathrm{E}\mu_d & E_2 \end{bmatrix}$$

$$\frac{d\rho}{dt} = \frac{i}{\hbar}\left(\rho\hat{H} - \hat{H}\rho\right)$$

$$= \frac{i}{\hbar}\left(\begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}\begin{bmatrix} E_1 & -\mathrm{E}\mu_d \\ -\mathrm{E}\mu_d & E_2 \end{bmatrix} - \begin{bmatrix} E_1 & -\mathrm{E}\mu_d \\ -\mathrm{E}\mu_d & E_2 \end{bmatrix}\begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix}\right)$$

$$= \frac{i}{\hbar}\begin{bmatrix} -\mathrm{E}\mu_d\left(\rho_{12} - \rho_{21}\right) & -\mathrm{E}\mu_d\left(\rho_{11} - \rho_{22}\right) + \left(E_2 - E_1\right)\rho_{12} \\ -\mathrm{E}\mu_d\left(\rho_{22} - \rho_{11}\right) + \left(E_1 - E_2\right)\rho_{21} & -\mathrm{E}\mu_d\left(\rho_{21} - \rho_{12}\right) \end{bmatrix}$$

Behavior of the density matrix in time

Taking the “2 – 1” element of both sides in

$$\frac{d\rho}{dt} = \frac{i}{\hbar}\begin{bmatrix} -\mathrm{E}\mu_d\left(\rho_{12}-\rho_{21}\right) & -\mathrm{E}\mu_d\left(\rho_{11}-\rho_{22}\right)+\left(E_2-E_1\right)\rho_{12} \\ -\mathrm{E}\mu_d\left(\rho_{22}-\rho_{11}\right)+\left(E_1-E_2\right)\rho_{21} & -\mathrm{E}\mu_d\left(\rho_{21}-\rho_{12}\right) \end{bmatrix}$$

with $\hbar\omega_{21} = E_2 - E_1$ gives

$$\frac{d\rho_{21}}{dt} = \frac{i}{\hbar}\left[\left(\rho_{11}-\rho_{22}\right)\mathrm{E}\mu_d - \left(E_2-E_1\right)\rho_{21}\right] = -i\omega_{21}\rho_{21} + i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{11}-\rho_{22}\right)$$

Behavior of the density matrix in time

From the diagonal elements in

$$\frac{d\rho}{dt} = \frac{i}{\hbar}\begin{bmatrix} -\mathrm{E}\mu_d\left(\rho_{12}-\rho_{21}\right) & -\mathrm{E}\mu_d\left(\rho_{11}-\rho_{22}\right)+\left(E_2-E_1\right)\rho_{12} \\ -\mathrm{E}\mu_d\left(\rho_{22}-\rho_{11}\right)+\left(E_1-E_2\right)\rho_{21} & -\mathrm{E}\mu_d\left(\rho_{21}-\rho_{12}\right) \end{bmatrix}$$

we can examine the population difference $\rho_{11}-\rho_{22}$ between the lower and upper states

Using the Hermiticity of ρ

which tells us that $\rho_{12}=\rho_{21}^*$

we have $\frac{d}{dt}\left(\rho_{11}-\rho_{22}\right)=2i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{21}-\rho_{21}^*\right)$

Behavior of the density matrix in time

Solving $$\frac{d\rho_{21}}{dt} = -i\omega_{21}\rho_{21} + i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{11} - \rho_{22}\right)$$

and $$\frac{d}{dt}\left(\rho_{11} - \rho_{22}\right) = 2i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{21} - \rho_{21}^{*}\right)$$

covers any possible behavior of this idealized system

Note: this is not a perturbation theory analysis

Density matrix and relaxation times

Consider a fractional population difference $\rho_{11} - \rho_{22}$

between the "lower" and "upper" states

Suppose that, in equilibrium, with no applied fields

this difference would have a value $(\rho_{11} - \rho_{22})_o$

Then experience might tell us that

because of mechanisms such as

collisions with the walls of a box or with other atoms

or by spontaneous emission

such systems often settle back down again to $(\rho_{11} - \rho_{22})_o$

with an exponential decay with some time constant T_1

Density matrix and relaxation times

Then we could hypothesize that we could add a term to

$$\frac{d}{dt}\left(\rho_{11}-\rho_{22}\right)=2i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{21}-\rho_{21}^*\right)$$

to give

$$\frac{d}{dt}\left(\rho_{11}-\rho_{22}\right)=2i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{21}-\rho_{21}^*\right)-\frac{\left(\rho_{11}-\rho_{22}\right)-\left(\rho_{11}-\rho_{22}\right)_o}{T_1}$$

For E = 0, this expression would give exponential decay

back to $\left(\rho_{11}-\rho_{22}\right)=\left(\rho_{11}-\rho_{22}\right)_o$

with time constant T_1

Density matrix and relaxation times

We have to consider a similar process also

for the off-diagonal elements of the density matrix

as in $\dfrac{d\rho_{21}}{dt} = -i\omega_{21}\rho_{21} + i\dfrac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{11} - \rho_{22}\right)$

To understand this, we need to understand

the meaning of the off-diagonal elements ρ_{21} and ρ_{12}

which we remember are defined with a relation

$$\rho_{uv} \equiv \left\langle \phi_u \middle| \rho \middle| \phi_v \right\rangle = \sum_j P_j c_u^{(j)} \left(c_v^{(j)} \right)^* \equiv \overline{c_u c_v^*}$$

Dephasing

Within any given pure state j

the product $c_u^{(j)}\left(c_v^{(j)}\right)^*$ is in general oscillating

If we have expanded in energy eigenstates $|\phi_u\rangle$ and $|\phi_v\rangle$

of the unperturbed system

there is a time-dependence $\exp\left(-iE_u t/\hbar\right)$ built into $c_u^{(j)}$

and a time-dependence $\exp\left(iE_v t/\hbar\right)$ built into $\left(c_v^{(j)}\right)^*$

so the product $c_u^{(j)}\left(c_v^{(j)}\right)^*$

has an underlying oscillation of the form

$$\exp(-i(E_u - E_v)t/\hbar)$$

Dephasing

As time evolves, the system can get scattered

- from pure state j into another pure state k
  - with some probability
    - possibly even a state in which ρ_{11} and ρ_{22} are unchanged
      - but in which the phases of the coefficients $c_1^{(k)}$ and $c_2^{(k)}$ are different

At any given time, therefore, we may have an ensemble of

- different possibilities for the quantum mechanical state
  - all possibly with different phases of oscillation

Dephasing

In our mixed state

if we have sufficiently many such random phases that are sufficiently different

then the ensemble average of a product $c_u c_v^*$

for different u and v, i.e., $\overline{c_u c_v^*}$

will average out to zero

But this ensemble average is simply

the off-diagonal density matrix element $\rho_{uv} \equiv \overline{c_u c_v^*}$

Hence, off-diagonal elements contain information about the coherence of the populations in different states

Dephasing

The processes that scatter into states with different phases

can be called "dephasing" processes

The simplest model is that

dephasing processes cause

an exponential settling

of any off-diagonal element

to zero

with some time constant T_2

Dephasing

Hence we postulate adding a term $-\rho_{21}/T_2$ to

$$\frac{d\rho_{21}}{dt} = -i\omega_{21}\rho_{21} + i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{11}-\rho_{22}\right)$$

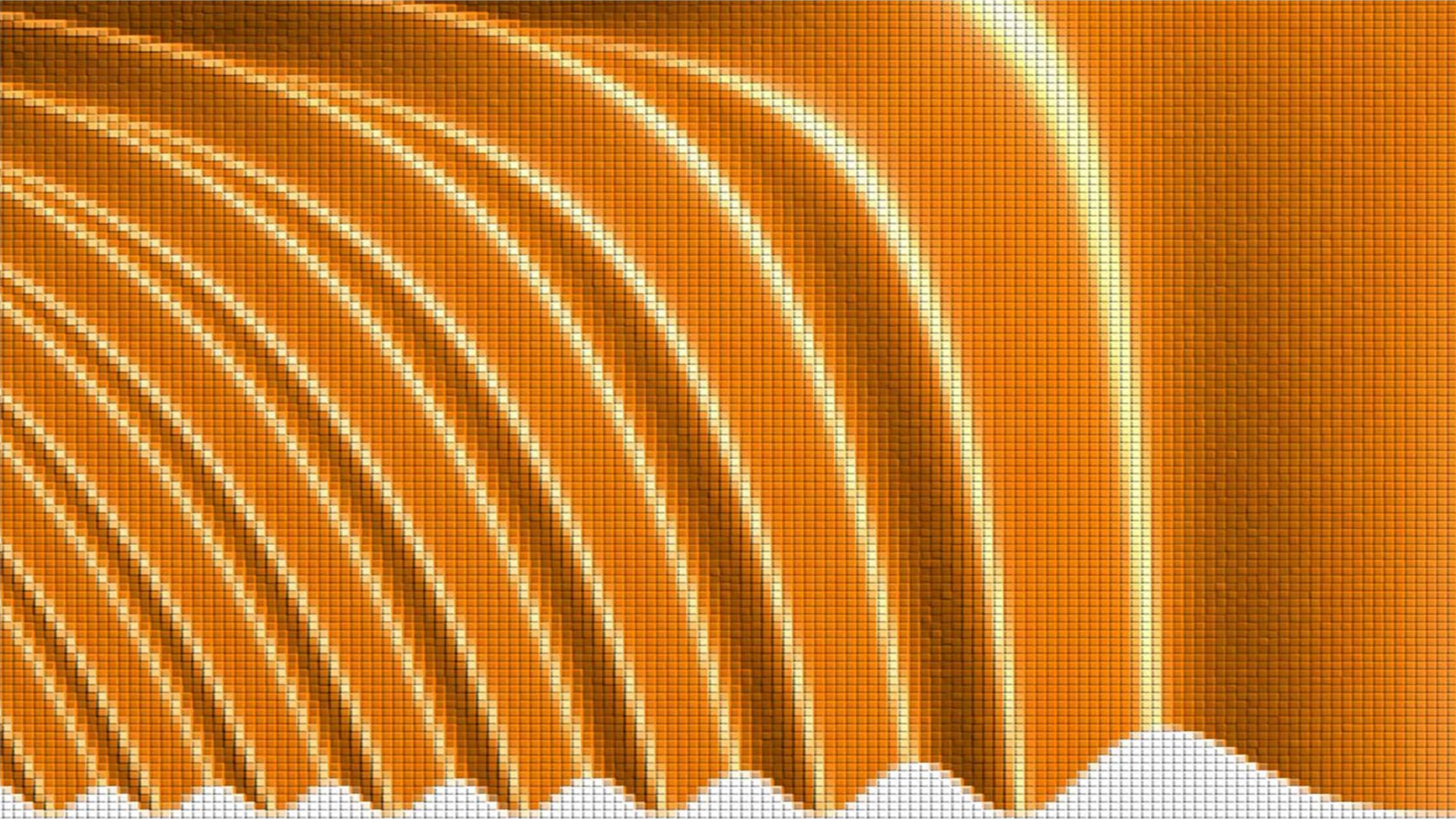
to obtain

$$\frac{d\rho_{21}}{dt} = -i\omega_{21}\rho_{21} + i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{11}-\rho_{22}\right) - \frac{\rho_{21}}{T_2}$$

In the absence of an optical field E

ρ_{21} would execute an oscillation at approximately frequency ω_{21}

decaying to zero approximately exponentially with a dephasing time constant T_2

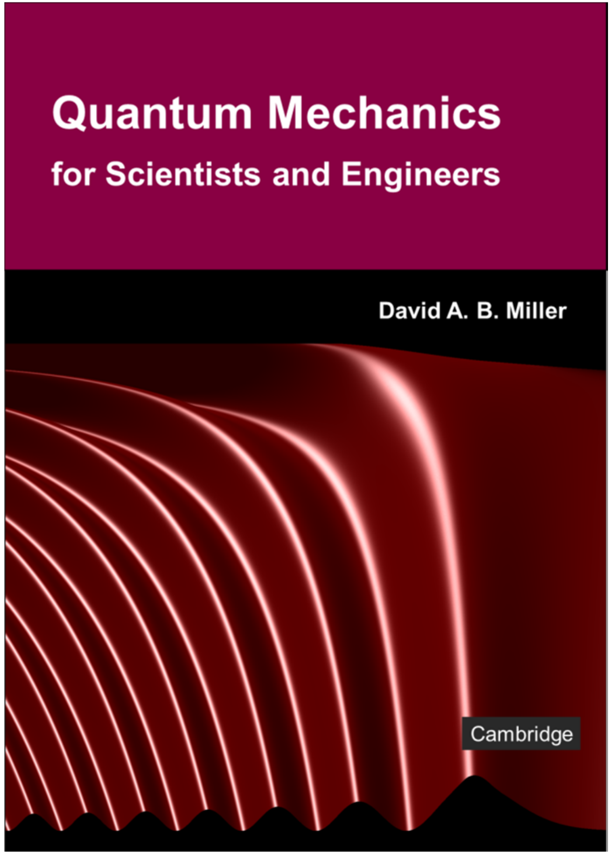


17.3 The density matrix and optical absorption

Slides: Video 17.3.5 Behavior with oscillating field

Text reference: Quantum Mechanics for Scientists and Engineers

Section 14.5 from paragraph above Eq. 14.41 through Fig. 14.1



The density matrix and optical absorption

Behavior with oscillating field

Quantum mechanics for scientists and engineers

David Miller

Behavior with oscillating field

We want see what happens when

we apply an oscillating electric field

$$\mathrm{E}(t) = \mathrm{E}_o \cos \omega t = \frac{\mathrm{E}_o}{2}\left(\exp(i\omega t) + \exp(-i\omega t)\right)$$

to our two-level system

We can simplify our algebra and results

if we define new “slowly varying” quantities

$$\beta_{21}(t) = \rho_{21}(t)\exp(i\omega t) \qquad \beta_{12}(t) = \rho_{12}(t)\exp(-i\omega t)$$

Using these quantity takes out the underlying oscillation at frequency ω from our algebra

Behavior with oscillating field

We can rewrite $\frac{d\rho_{21}}{dt} = -i\omega_{21}\rho_{21} + i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{11} - \rho_{22}\right) - \frac{\rho_{21}}{T_2}$

and $\frac{d}{dt}\left(\rho_{11} - \rho_{22}\right) = 2i\frac{\mu_d}{\hbar}\mathrm{E}\left(\rho_{21} - \rho_{21}^{*}\right) - \frac{\left(\rho_{11} - \rho_{22}\right) - \left(\rho_{11} - \rho_{22}\right)_o}{T_1}$

using $\beta_{21}(t) = \rho_{21}(t)\exp(i\omega t)$

and dropping all terms $\propto \exp(\pm 2i\omega t)$

on the presumption that such terms will average out to zero over timescales of cycles and hence

they will make relatively little contribution to the resulting values of $\rho_{11} - \rho_{22}$ and β_{12}

Bloch equations

Hence we obtain, approximately

$$\frac{d}{dt}\left(\rho_{11}-\rho_{22}\right)=i\frac{\mu_d}{\hbar}\mathrm{E}_o\left(\beta_{21}-\beta_{21}^{*}\right)-\frac{\left(\rho_{11}-\rho_{22}\right)-\left(\rho_{11}-\rho_{22}\right)_o}{T_1}$$

$$\frac{d\beta_{21}}{dt}=i\left(\omega-\omega_{21}\right)\beta_{21}+i\frac{\mu_d}{2\hbar}\mathrm{E}_\mathrm{o}\left(\rho_{11}-\rho_{22}\right)-\frac{\beta_{21}}{T_2}$$

These equations are often known as the Bloch equations

They were first derived in the field of magnetic resonance

Dipole average

We defined $\beta_{21}(t) = \rho_{21}(t)\exp(i\omega t)$ and $\beta_{12}(t) = \rho_{12}(t)\exp(-i\omega t)$

We know the density matrix is Hermitian, so $\rho_{12} = \rho_{21}^*$

so $\beta_{12} \equiv \rho_{12}\exp(-i\omega t) = \rho_{21}^*\exp(-i\omega t) = \beta_{21}^*$

We now evaluate the ensemble average

of the dipole moment of the system

which we previously deduced was $\overline{\langle\mu\rangle} = \mu_d(\rho_{12} + \rho_{21})$

We have $\overline{\langle\mu\rangle} = \mu_d\left(\beta_{12}\exp(i\omega t) + \beta_{21}\exp(-i\omega t)\right)$

$$= 2\mu_d\left[\operatorname{Re}(\beta_{21})\cos\omega t + \operatorname{Im}(\beta_{21})\sin\omega t\right]$$

where we used our result $\beta_{12} = \beta_{21}^*$ from above

Solving in the steady state

Now let us solve in the "steady state"

with a steady monochromatic field and

when the system has settled down

In steady state $\rho_{11} - \rho_{22}$

the population difference between the states

will have settled to some value

so $d(\rho_{11} - \rho_{22}) / dt = 0$

Similarly, any coherent responses will have settled down

to follow the appropriate driving field terms

so we expect $d\beta_{21} / dt = 0$ also

Solving in the steady state

So, setting the left-hand sides of both

$$\frac{d}{dt}\left(\rho_{11}-\rho_{22}\right)=i\frac{\mu_d}{\hbar}\mathrm{E}_o\left(\beta_{21}-\beta_{21}^*\right)-\frac{\left(\rho_{11}-\rho_{22}\right)-\left(\rho_{11}-\rho_{22}\right)_o}{T_1}$$

and $$\frac{d\beta_{21}}{dt}=i\left(\omega-\omega_{21}\right)\beta_{21}+i\frac{\mu_d}{2\hbar}\mathrm{E}_o\left(\rho_{11}-\rho_{22}\right)-\frac{\beta_{21}}{T_2}$$

to zero

we can solve the resulting simultaneous linear equations in the two variables β_{21} and $(\rho_{11}-\rho_{22})$

the details of which are left as an exercise

Solutions in the steady state

With $\Omega = \mu_d \mathrm{E}_o / 2\hbar$, the results are

$$\rho_{11} - \rho_{22} = \left(\rho_{11} - \rho_{22}\right)_o \frac{1 + \left(\omega - \omega_{21}\right)^2 T_2^2}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1}$$

$$\mathrm{Im}\left(\beta_{21}\right) = \frac{\Omega T_2 \left(\rho_{11} - \rho_{22}\right)_o}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1}$$

$$\mathrm{Re}\left(\beta_{21}\right) = \frac{\left(\omega_{21} - \omega\right)\Omega T_2^2 \left(\rho_{11} - \rho_{22}\right)_o}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1}$$

Behavior with oscillating field

Presume that we have some large number N

of such systems ("atoms") per unit volume

The population difference (per unit volume) between

the number in the lower state and

the number in the higher state

is therefore

$$\Delta N = N\left(\rho_{11} - \rho_{22}\right)$$

and in the absence of the optical field

the population difference is

$$\Delta N_o = N\left(\rho_{11} - \rho_{22}\right)_o$$

Population difference with oscillating field

Using $\Delta N = N\left(\rho_{11} - \rho_{22}\right)$ and $\Delta N_o = N\left(\rho_{11} - \rho_{22}\right)_o$

we can rewrite

$$\rho_{11} - \rho_{22} = \left(\rho_{11} - \rho_{22}\right)_o \frac{1 + \left(\omega - \omega_{21}\right)^2 T_2^2}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1}$$

as

$$\boxed{\Delta N = \Delta N_o \frac{1 + \left(\omega - \omega_{21}\right)^2 T_2^2}{1 + \left(\omega - \omega_{21}\right)^2 T_2^2 + 4\Omega^2 T_2 T_1}}$$

This result tells us how the population difference varies

as a function of optical intensity ($\propto \Omega^2$)

and frequency ω

Polarization with oscillating field

In general in electromagnetism

- the (static) polarization P is defined as $P = \varepsilon_o \chi E$
  - where χ is the susceptibility

When we have an oscillating field

- the response of the medium
  - and hence the polarization
    - can be out of phase with the electric field
      - and then it is convenient
        - to generalize the idea of susceptibility

Susceptibility with oscillating field

We can formally think of the susceptibility

as a complex quantity with real and imaginary parts

χ' and χ'' respectively

or equivalently we can explicitly write

the response to a real field $\mathrm{E}_o \cos \omega t$

as $\mathrm{P} = \varepsilon_o \mathrm{E}_o \left(\chi' \cos \omega t + \chi'' \sin \omega t \right)$

It is also generally true in electromagnetism that

the polarization is the dipole moment per unit volume

Hence here we can also write

$$\mathrm{P} = N \overline{\langle \mu \rangle}$$

Susceptibility with oscillating field

Hence using $\overline{\langle \mu \rangle} = 2\mu_d \left[\mathrm{Re}(\beta_{21}) \cos \omega t + \mathrm{Im}(\beta_{21}) \sin \omega t \right]$

$$\mathrm{Im}(\beta_{21}) = \frac{\Omega T_2 (\rho_{11} - \rho_{22})_o}{1 + (\omega - \omega_{21})^2 T_2^2 + 4\Omega^2 T_2 T_1}$$

$$\mathrm{Re}(\beta_{21}) = \frac{(\omega_{21} - \omega) \Omega T_2^2 (\rho_{11} - \rho_{22})_o}{1 + (\omega - \omega_{21})^2 T_2^2 + 4\Omega^2 T_2 T_1}$$

$$\mathrm{P} = \varepsilon_o \mathrm{E}_o \left(\chi' \cos \omega t + \chi'' \sin \omega t \right)$$

$$\mathrm{P} = N \overline{\langle \mu \rangle}$$

we can write explicit formulas for χ' and χ''

Susceptibility with oscillating field

We obtain

$$\chi'(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{(\omega_{21} - \omega) T_2}{1 + (\omega - \omega_{21})^2 T_2^2 + 4\Omega^2 T_2 T_1}$$

$$\chi''(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{1}{1 + (\omega - \omega_{21})^2 T_2^2 + 4\Omega^2 T_2 T_1}$$

In electromagnetism, $\mathrm{Re}(\chi) \equiv \chi'$, the in-phase response

is responsible for refractive index

and the quadrature (i.e., 90° shifted) response, $\mathrm{Im}(\chi) \equiv \chi''$

is responsible for optical absorption

Small field susceptibility

For a “two-level” system for small electric field amplitude

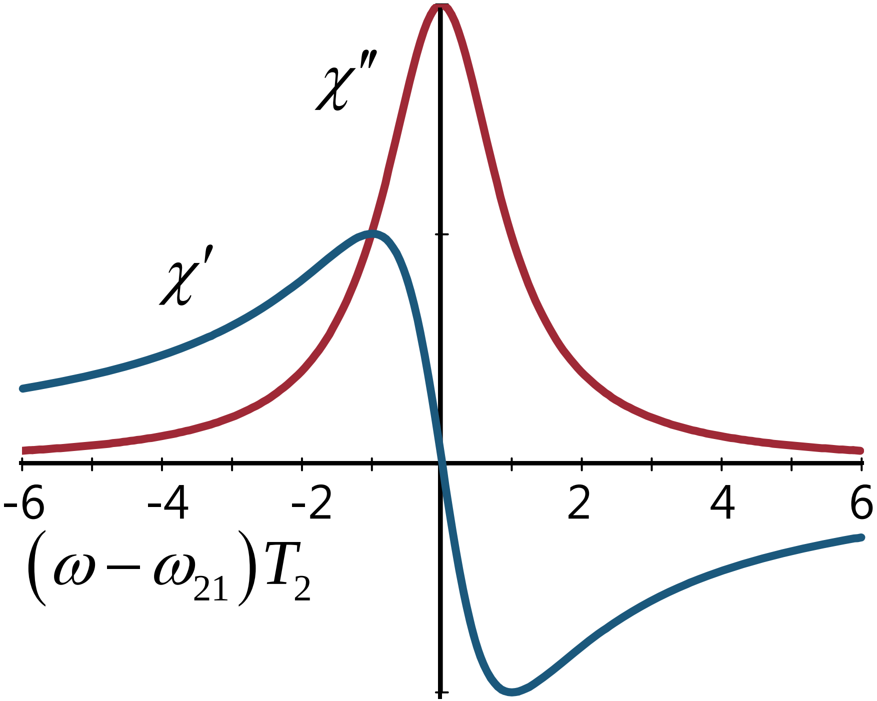
then $\Omega = \mu_d \mathrm{E}_o / 2\hbar \simeq 0$

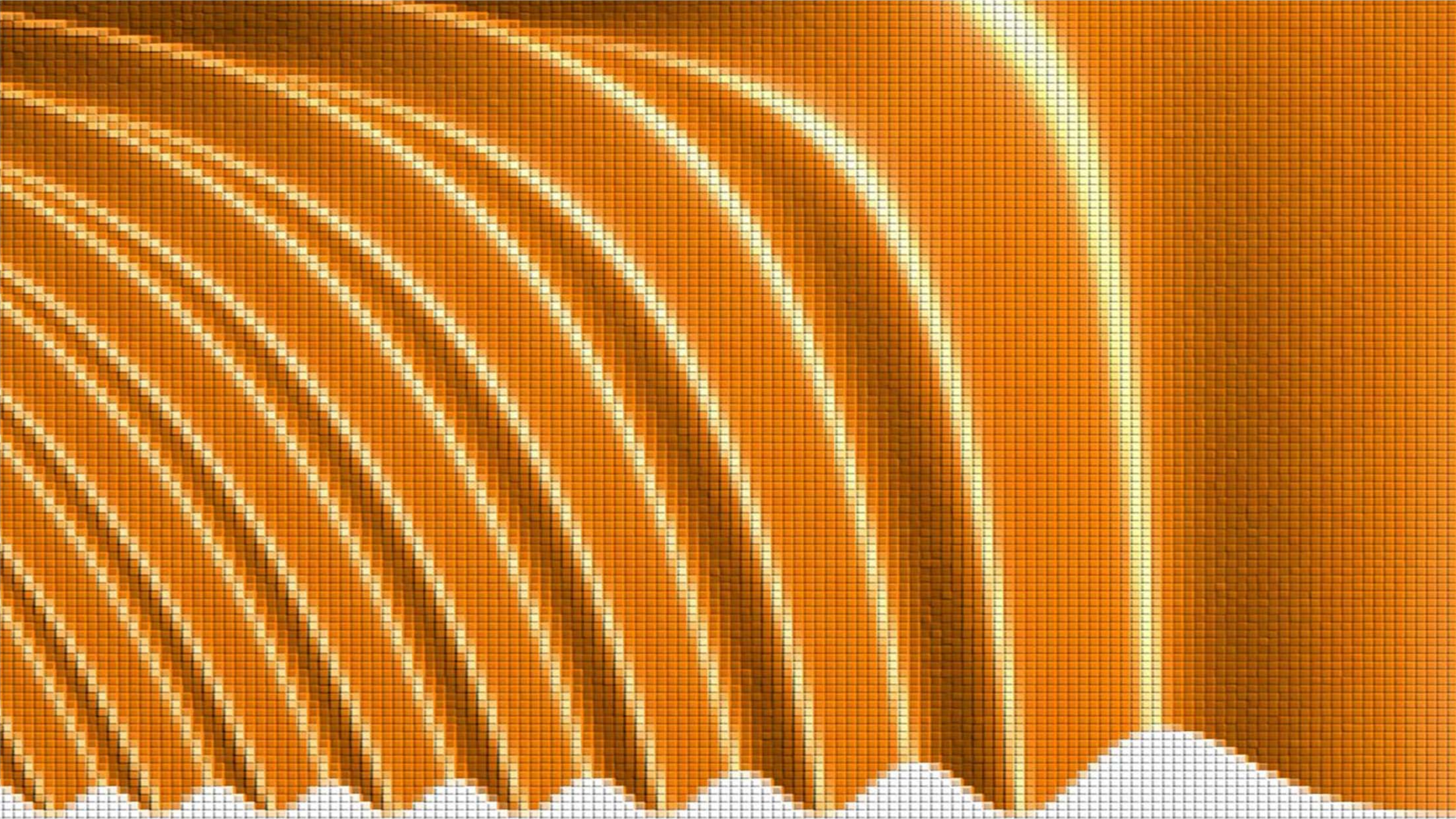
and we have the normal “linear” refraction variation

$$\chi'(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{(\omega_{21} - \omega) T_2}{1 + (\omega - \omega_{21})^2 T_2^2}$$

and Lorentzian absorption

$$\chi''(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{1}{1 + (\omega - \omega_{21})^2 T_2^2}$$



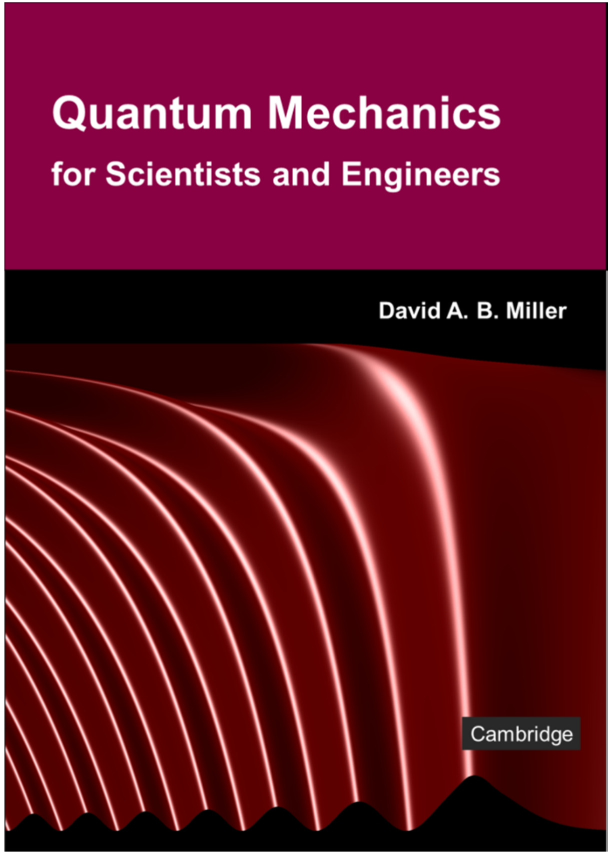


17.3 The density matrix and optical absorption

Slides: Video 17.3.7 Density matrix and nonlinear optics

Text reference: Quantum Mechanics for Scientists and Engineers

Section 14.5 after Fig. 14.1, and 14.6



The density matrix and optical absorption

Density matrix and nonlinear optics

Quantum mechanics for scientists and engineers

David Miller

Absorption saturation

In $\chi''(\omega) = \frac{\mu_d^2 T_2 \Delta N_o}{\varepsilon_o \hbar} \frac{1}{1 + (\omega - \omega_{21})^2 T_2^2 + 4\Omega^2 T_2 T_1}$

Ω^2 is proportional to the electric field squared

which is proportional to the intensity I of the light

Hence we can write $4\Omega^2 T_2 T_1 \equiv I / I_S$

where I_S is called the saturation intensity

Hence, for example, on resonance ($\omega_{21} = \omega$), we have

$$\chi''(\omega) \propto 1/(1 + I / I_S)$$

This equation describes “absorption saturation”

often seen with the high intensities from lasers

Density matrix and perturbation theory

Now we would start with $\frac{\partial \rho_{mn}}{\partial t} = \frac{i}{\hbar}\left[\rho, \hat{H}\right]_{mn}$

for the time evolution of the density matrix

instead of Schrödinger's equation

We could generalize the relaxation time approximation

now writing a proposed set of relations

$$\frac{\partial \rho_{mn}}{\partial t} = \frac{i}{\hbar}\left[\rho, \hat{H}\right]_{mn} - \gamma_{mn}\left(\rho_{mn} - \rho_{mno}\right)$$

ρ_{mno} is the equilibrium value for ρ_{mn}

and γ_{mn} is its "relaxation rate"

Density matrix and perturbation theory

One then starts with equations like

$$\frac{\partial \rho_{mn}}{\partial t} = \frac{i}{\hbar}\left[\rho, \hat{H}\right]_{mn} - \gamma_{mn}\left(\rho_{mn} - \rho_{mno}\right)$$

instead of the time-dependent Schrödinger equation

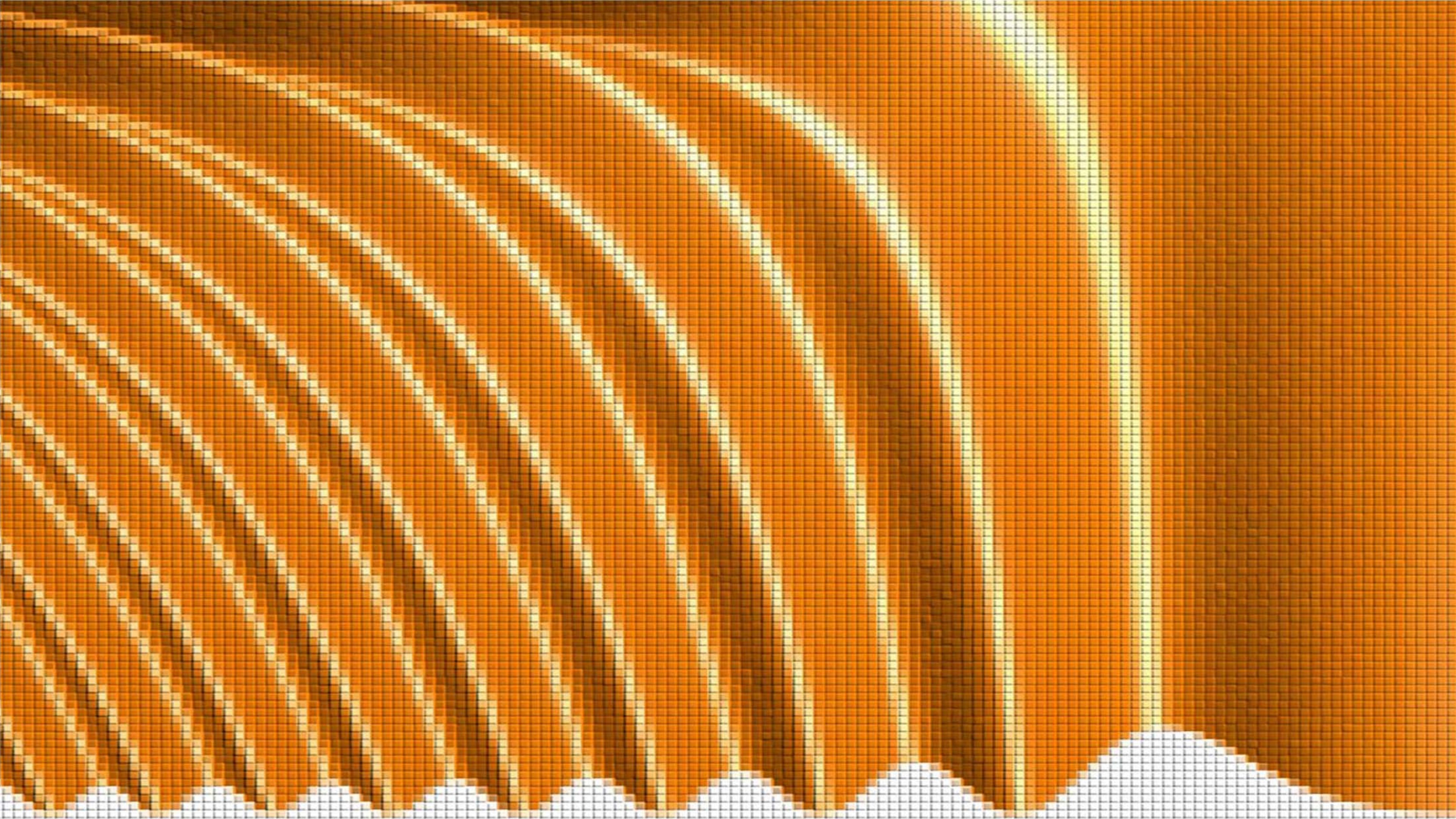
and constructs a perturbation theory just as before

This density matrix version is the one commonly used

for calculating non-linear optical coefficients

eliminating the singularities

when the transition energy and the photon energy coincide

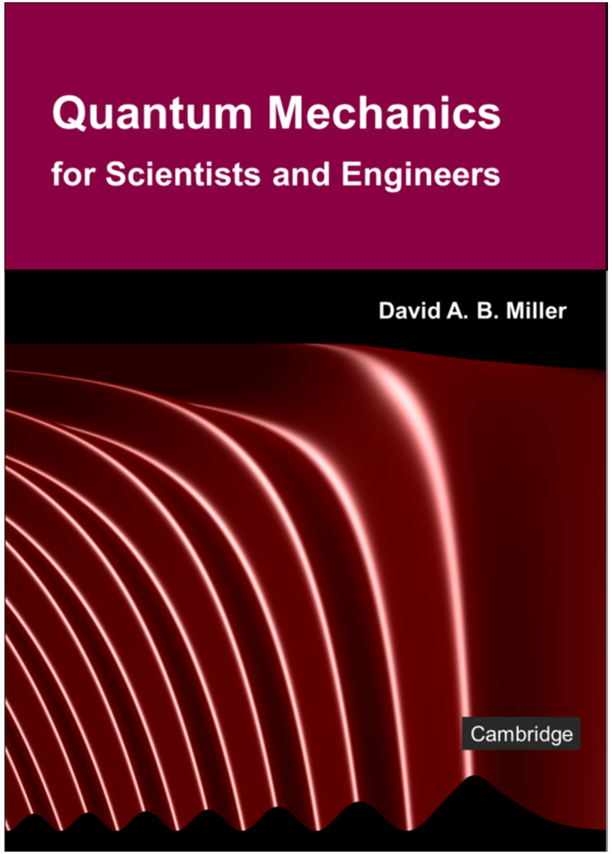


18.1 Quantum measurements and encryption

Slides: Video 18.1.1 Representing information

Text reference: Quantum Mechanics for Scientists and Engineers

Chapter 18 introduction

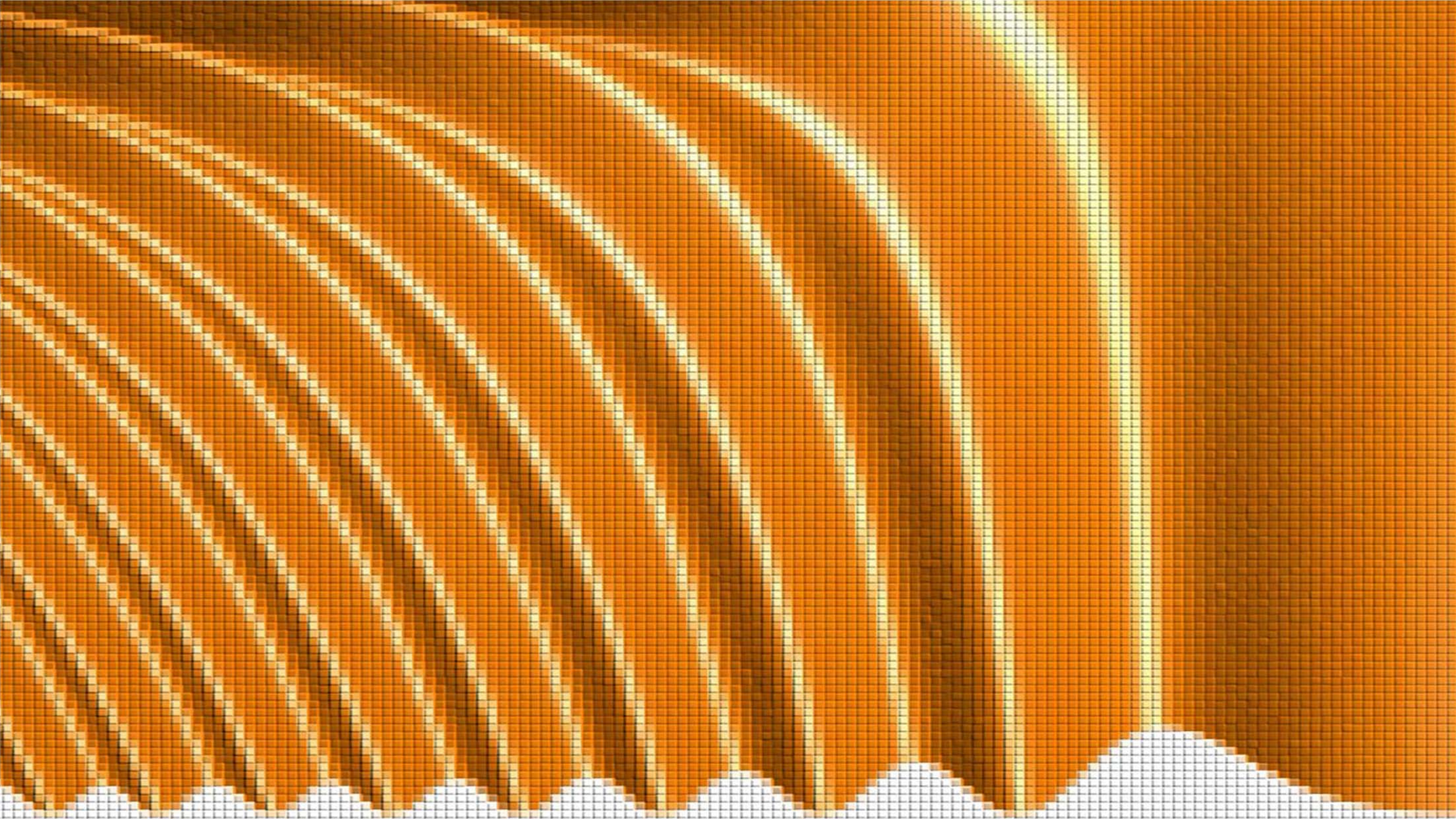


Quantum measurements and encryption

Representing information

Quantum mechanics for scientists and engineers

David Miller

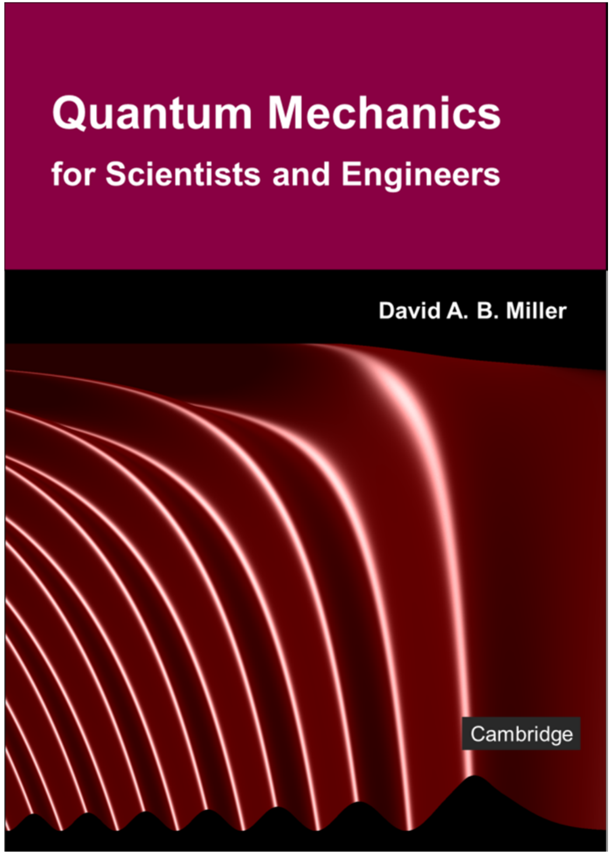


18.1 Quantum measurements and encryption

Slides: Video 18.1.2 Collapse and the no-cloning theorem

Text reference: Quantum Mechanics for Scientists and Engineers

Sections 18.1 and 18.2 through "No cloning theorem" subsection



Quantum Mechanics
for Scientists and Engineers

David A. B. Miller

Cambridge

Quantum measurements and encryption

Collapse and the no-cloning theorem

Quantum mechanics for scientists and engineers

David Miller

Quantum mechanical measurement

To interpret a quantum mechanical calculation

so that we can predict the result of a measurement

we say that

if the quantum mechanical calculation says the state of the system is $|\psi\rangle$

then the average value we will measure for some quantity A is given by

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle$$

where $\hat{A}$ is the operator associated with the quantity A

Quantum mechanical measurement

The measurement is a statistical process

- We must repeat the experiment many times
  - from the start
    - including the process that puts the system into the state
      - and take the average answer

We also find that

- every measurement we make returns a value
  - corresponding to one of the eigenvalues A_n of $\hat{A}$

Quantum mechanical measurement

Not every measurement returns the same value

- If we decompose the state into a linear combination
  - of the normalized eigenstates $|\psi_n\rangle$ of the operator $\hat{A}$
    - i.e., $|\psi\rangle = \sum_n a_n |\psi_n\rangle$
  - then we find that
    - the probability of measuring a particular eigenvalue
      - is given by $|a_n|^2$

This probabilistic result is known as the "Born Rule"

Quantum mechanical measurement

If no external influence is applied in the meantime

- any subsequent measurements on this system
  - will always give the same answer A_n
    - on measuring the quantity A

This behavior is called the

- "collapse of the wavefunction"

Measuring a quantity A appears to force it into one of its eigenstates

- As far as we know, this collapse is totally random
  - with probabilities from the Born Rule

Proof of the no-cloning theorem

We can show that

starting from

the first system in an arbitrary state $|\psi_{1a}\rangle$

and the second system in some prescribed starting state $|\psi_{2s}\rangle$

we cannot in general

create the second system in the state $|\psi_{2a}\rangle$

leaving the first system in state $|\psi_{1a}\rangle$

This is the no-cloning theorem

Proof of the no-cloning theorem

In this proof, our initial state of the two systems is therefore

the (direct product) state $|\psi_{1a}\rangle|\psi_{2s}\rangle$

We then imagine that we have some operation

that, over time, turns this state into the state $|\psi_{1a}\rangle|\psi_{2a}\rangle$

This operation is just some time-evolution operation

that we can describe by a (unitary) linear operator $\hat{T}$

such as the one $\hat{T} = \exp\left[-i\hat{H}(t-t_o)/\hbar\right]$

we devised before

where t is the time we finish

and t_o is the time when we started

Cloning system properties

We presume we have engineered our cloning system

- to give $\hat{T}$ the required properties

Specifically we need at least two properties for $\hat{T}$

- First we want $\hat{T}$ to perform the operation

$$|\psi_{1a}\rangle|\psi_{2a}\rangle = \hat{T}|\psi_{1a}\rangle|\psi_{2s}\rangle$$

  - cloning the state a of system 1 into system 2
- Second, for some orthogonal state $|\psi_{1b}\rangle$ of system 1
  - we want $|\psi_{1b}\rangle|\psi_{2b}\rangle = \hat{T}|\psi_{1b}\rangle|\psi_{2s}\rangle$
    - cloning the state b of system 1 also into system 2

There is no problem in general with constructing such a $\hat{T}$

Cloning a linear superposition state

The problem comes when we want to clone a linear superposition state

Suppose the initial state of system 1

is the linear superposition

$$|\psi_{1Sup}\rangle = \frac{1}{\sqrt{2}}\left(|\psi_{1a}\rangle + |\psi_{1b}\rangle\right)$$

Hence the initial state of the pair of systems is

$$\frac{1}{\sqrt{2}}\left(|\psi_{1a}\rangle + |\psi_{1b}\rangle\right)|\psi_{2s}\rangle = \frac{1}{\sqrt{2}}\left(|\psi_{1a}\rangle|\psi_{2s}\rangle + |\psi_{1b}\rangle|\psi_{2s}\rangle\right)$$

Cloning a linear superposition state

By postulation in quantum mechanics

the operators are linear

Operating on a linear superposition

must give the linear superposition of the operations

$$\hat{T}\frac{1}{\sqrt{2}}\left(|\psi_{1a}\rangle|\psi_{2s}\rangle+|\psi_{1b}\rangle|\psi_{2s}\rangle\right)=\frac{1}{\sqrt{2}}\left(\hat{T}|\psi_{1a}\rangle|\psi_{2s}\rangle+\hat{T}|\psi_{1b}\rangle|\psi_{2s}\rangle\right)$$

$$=\frac{1}{\sqrt{2}}\left(|\psi_{1a}\rangle|\psi_{2a}\rangle+|\psi_{1b}\rangle|\psi_{2b}\rangle\right)$$

This is not the result we wanted for our cloning operation

Cloning a linear superposition state

We wanted our cloning operation to take

the system 1 superposition state $\left(1/\sqrt{2}\right)\left(|\psi_{1a}\rangle+|\psi_{1b}\rangle\right)$

and change system 2 from its initial state $|\psi_{2s}\rangle$

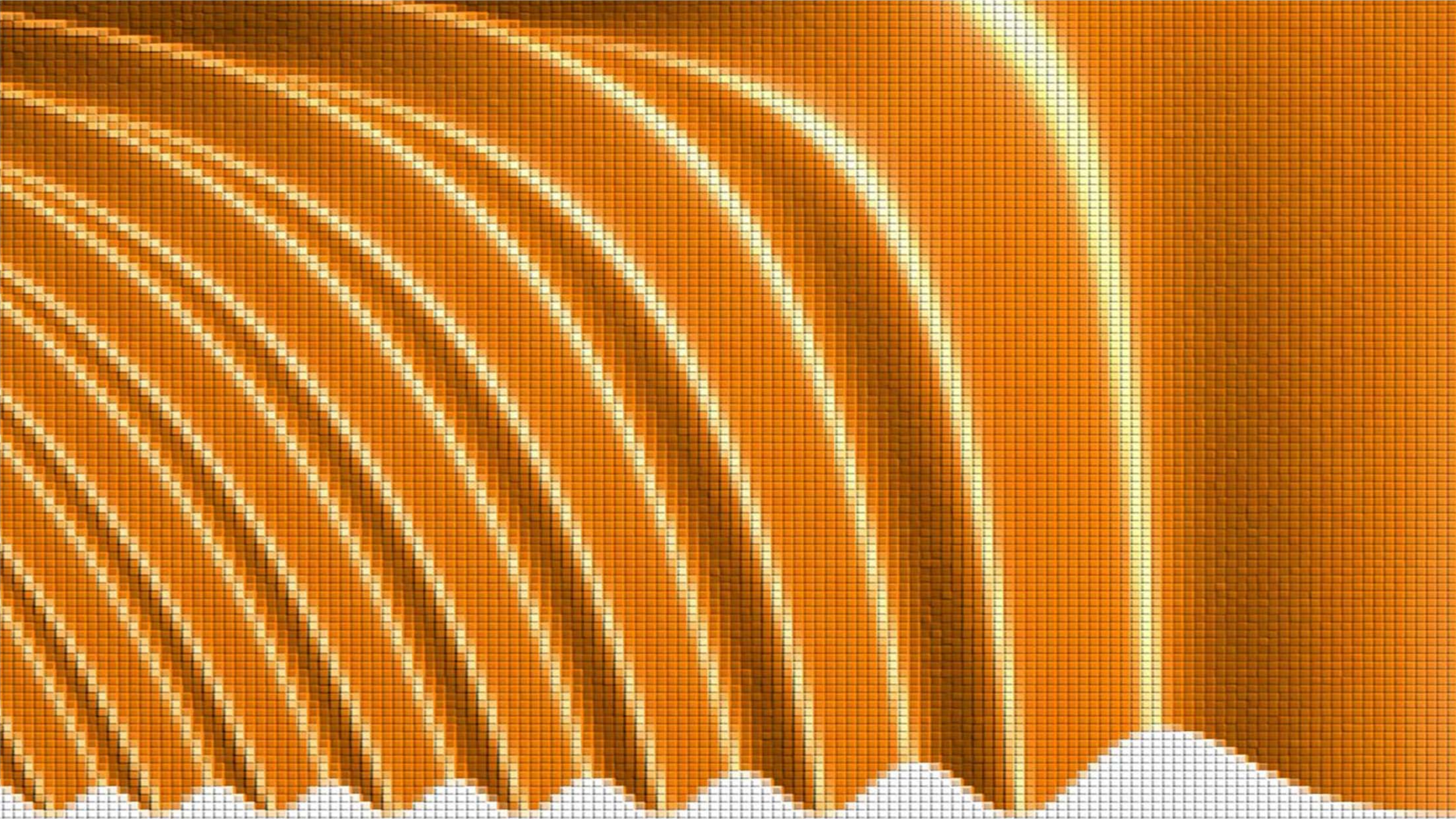
into the superposition $\left(1/\sqrt{2}\right)\left(|\psi_{2a}\rangle+|\psi_{2b}\rangle\right)$

while leaving system 1 in $\left(1/\sqrt{2}\right)\left(|\psi_{1a}\rangle+|\psi_{1b}\rangle\right)$

So we wanted the overall initial state $\left(1/\sqrt{2}\right)\left(|\psi_{1a}\rangle+|\psi_{1b}\rangle\right)|\psi_{2s}\rangle$

to lead to the final state $(1/2)\left(|\psi_{1a}\rangle+|\psi_{1b}\rangle\right)\left(|\psi_{2a}\rangle+|\psi_{2b}\rangle\right)$

but instead we got $\left(1/\sqrt{2}\right)\left(|\psi_{1a}\rangle|\psi_{2a}\rangle+|\psi_{1b}\rangle|\psi_{2b}\rangle\right)$

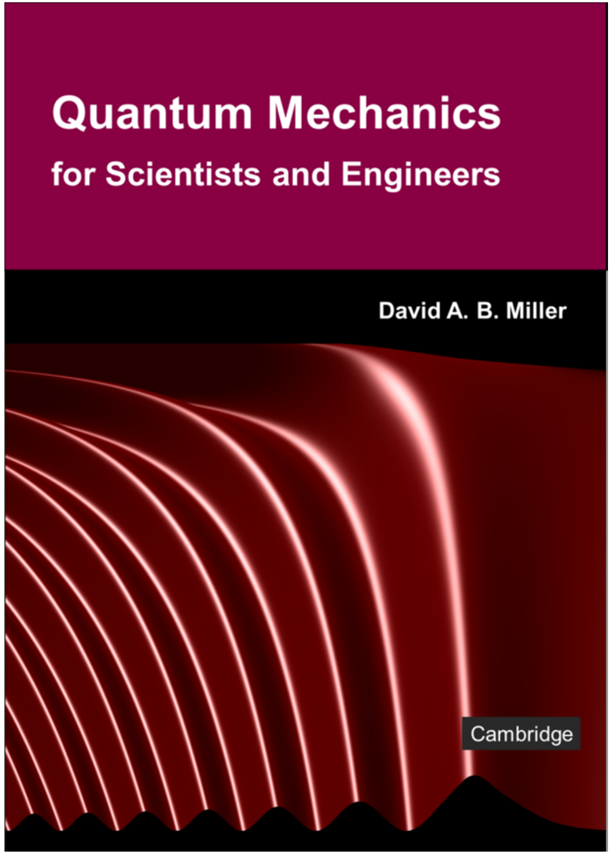


18.1 Quantum measurements and encryption

Slides: Video 18.1.4 Quantum cryptography

Text reference: Quantum Mechanics for Scientists and Engineers

Section 18.2 subsection “A simple quantum encryption scheme”



Quantum measurements and encryption

Quantum cryptography

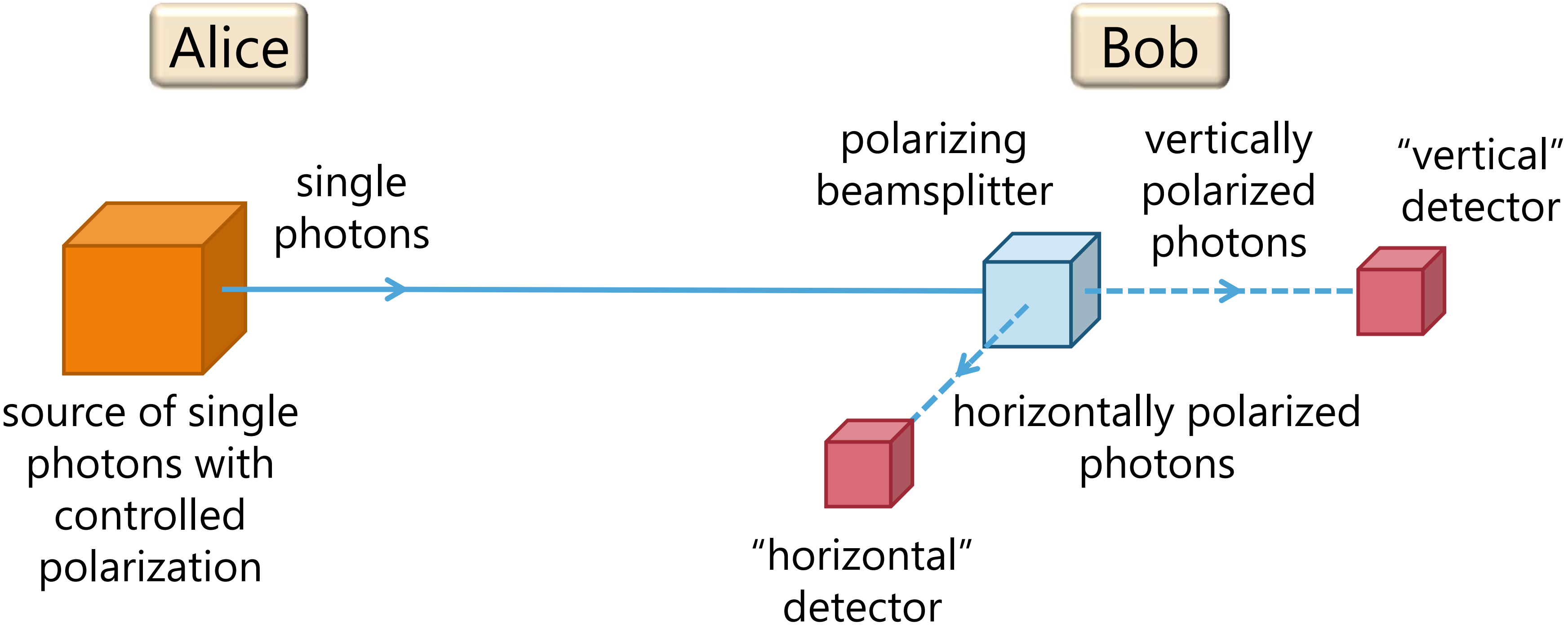
Quantum mechanics for scientists and engineers

David Miller

Simple quantum encryption

Alice wants to communicate securely with Bob

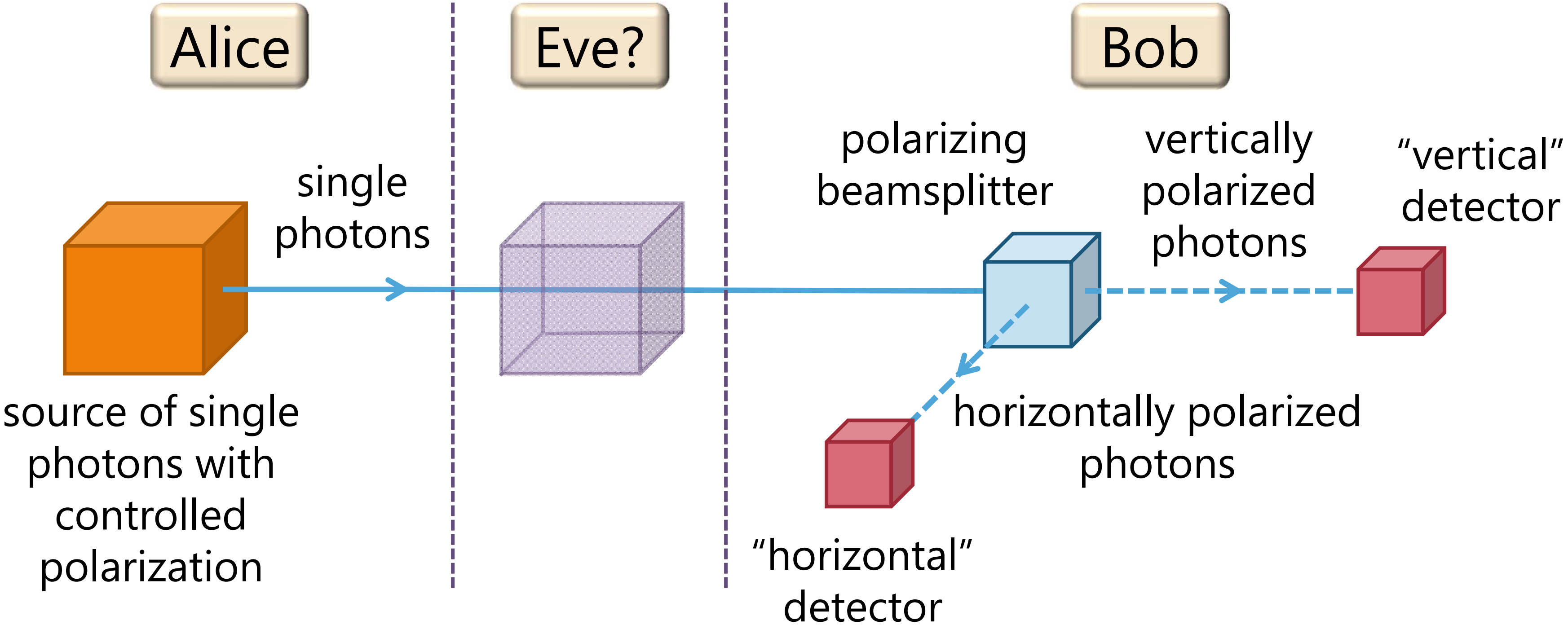
by sending single photons of specific polarizations



Simple quantum encryption

But there may be an “eavesdropper”, Eve

trying to secretly read their messages



Simple quantum encryption

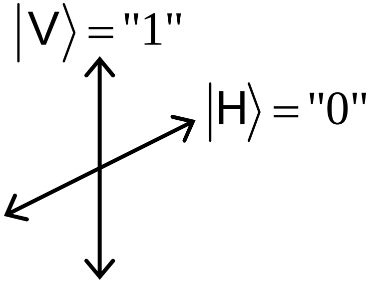
Suppose, first, for a "1"

- Alice sends a vertically polarized photon
  - a photon in the state $|V\rangle$

For a "0"

- she sends a horizontally polarized photon
  - a photon in the state $|H\rangle$

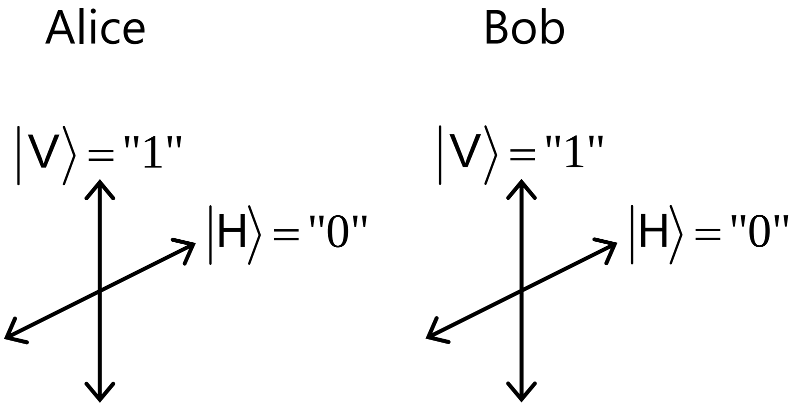
Alice



Simple quantum encryption

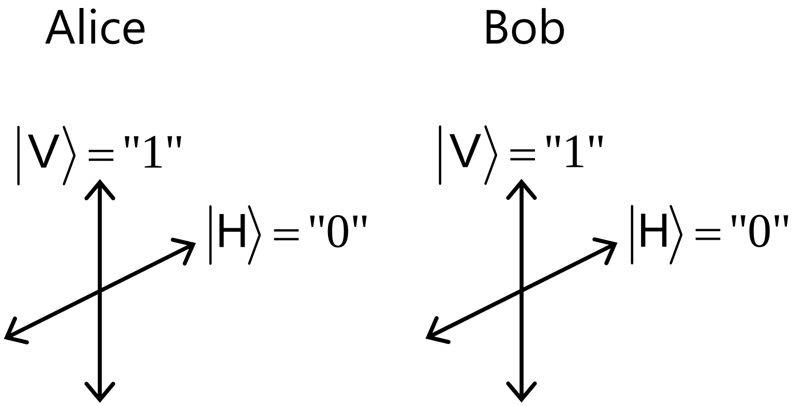
Bob’s apparatus

- separates the two polarizations
  - to different single-photon detectors
    - vertical for “1”
    - horizontal for “0”



Simple quantum encryption

- This scheme is not secure
- Eve could insert a detection system like Bob's into the path
  - receive the photon from Alice
    - write down the answer
- then, using a transmission system like Alice's
  - retransmit the photon on to Bob
    - with Alice and Bob being unaware of her interception


Alice

Bob

$|V\rangle = "1"$

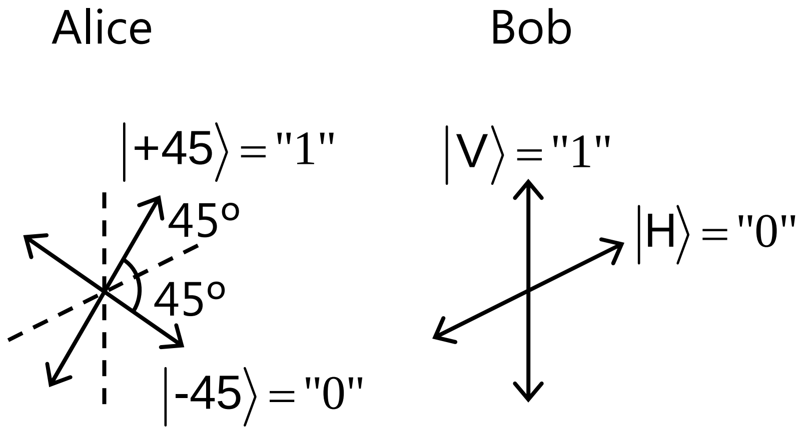
$|H\rangle = "0"$

$|V\rangle = "1"$

$|H\rangle = "0"$

Simple quantum encryption

- Now rotate by 45° the apparatus Alice uses for transmission
- Now Alice transmits
  - a "1" using the state $\left|+45\right\rangle$ and
  - a "0" using the state $\left|-45\right\rangle$
- If Bob leaves his apparatus unchanged
  - he will receive no information
- Note $\left|+45\right\rangle = \left(1/\sqrt{2}\right)\left(\left|H\right\rangle + \left|V\right\rangle\right)$
  - and $\left|-45\right\rangle = \left(1/\sqrt{2}\right)\left(\left|H\right\rangle - \left|V\right\rangle\right)$



Simple quantum encryption

Since $|+45\rangle = \left(1/\sqrt{2}\right)\left(|H\rangle + |V\rangle\right)$

and $|-45\rangle = \left(1/\sqrt{2}\right)\left(|H\rangle - |V\rangle\right)$

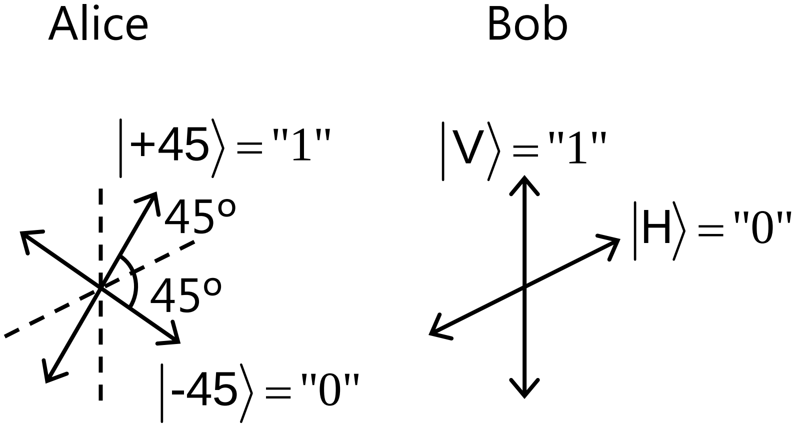
no matter which state Alice sends

Bob's apparatus will give the answer

$|H\rangle$ half the time

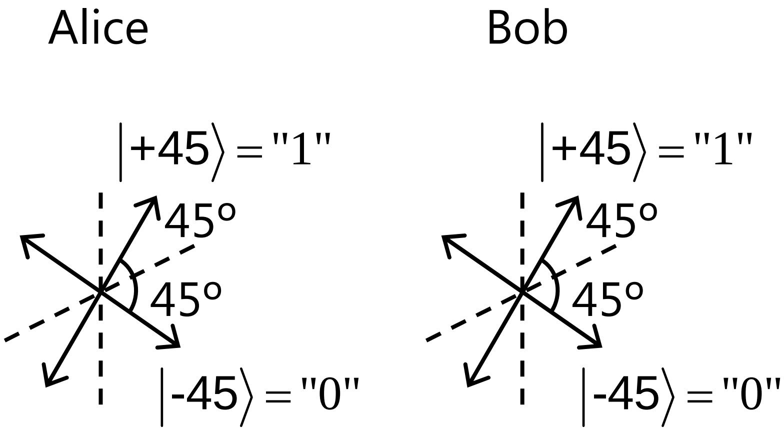
and the answer

$|V\rangle$ half the time



Simple quantum encryption

- Now Alice and Bob each rotate their apparatus by 45°
  - Then they can send information just as before
- If Eve interposes her apparatus
  - oriented horizontally and vertically
- she will receive no information, and
  - Bob and Alice will deduce their message is being intercepted



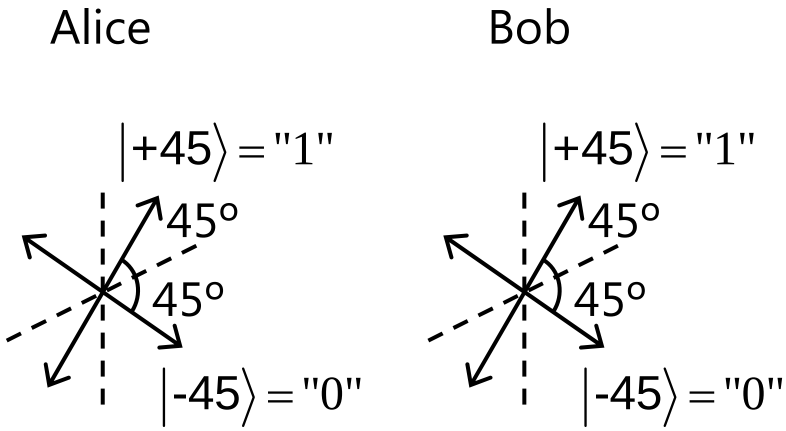
Simple quantum encryption

Bob and Alice can monitor errors

- talking on the telephone and checking
  - quite openly and publicly

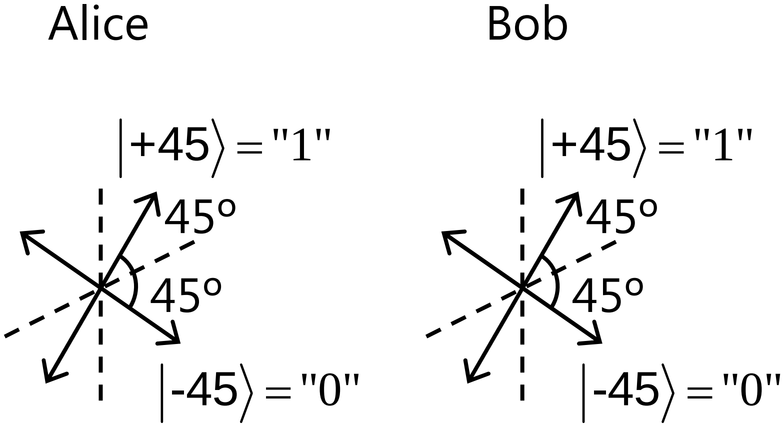
to see that they are sending and receiving

- the same bits
  - on some test cases



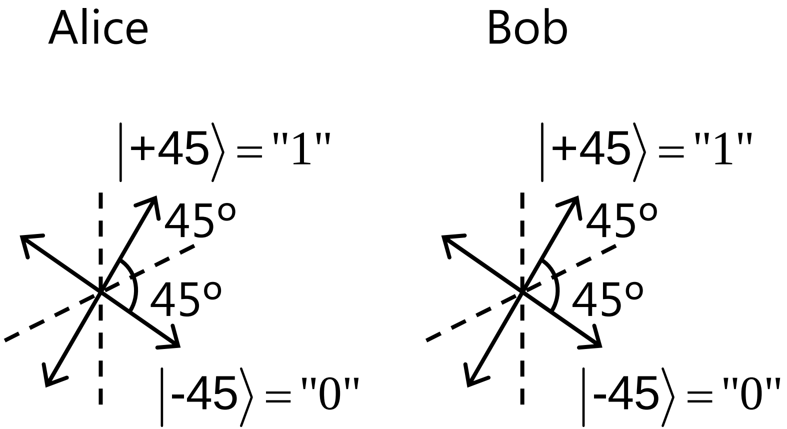
Simple quantum encryption

If Eve has interposed herself
in this horizontal and vertical way
half the bits apparently received by Bob
will turn out to be wrong
and Alice and Bob will know to
discard all of the bits and to
send out a search party
to find Eve and her apparatus



Simple quantum encryption

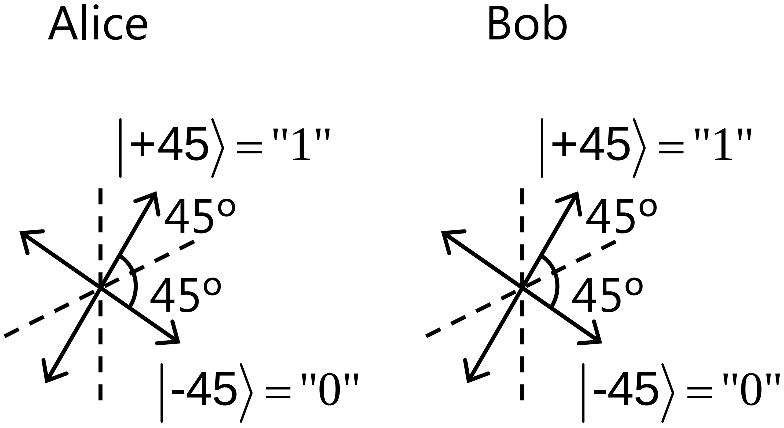
- Eve
  - seeing the approaching search party
    - realizes her apparatus is set incorrectly
      - and retreats to come back another day
        - then setting her apparatus in the 45° fashion



Simple quantum encryption

Alice and Bob might by that time

- have changed back
  - but there is a 50% chance that
    - Eve could set her apparatus correctly
      - and a 50% chance of interception
        - is likely too high for Alice and Bob



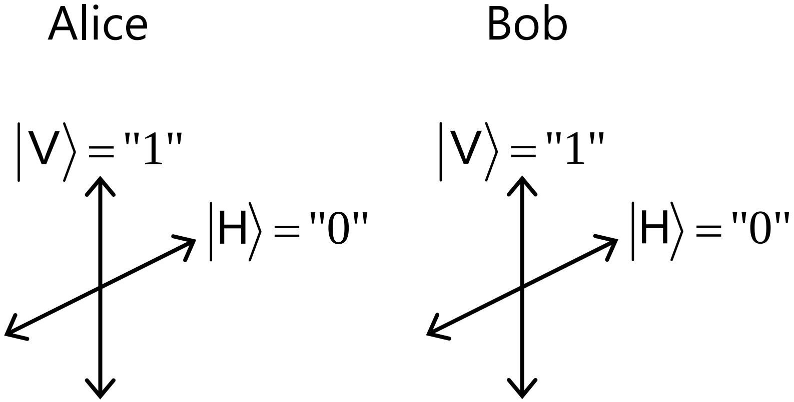
Simple quantum encryption

The trick to thwarting Eve is that
Alice and Bob
for each time they want to try
to communicate a bit
each separately and randomly
choose between
the horizontal-vertical
setting of their apparatus
and the 45° one

Simple quantum encryption

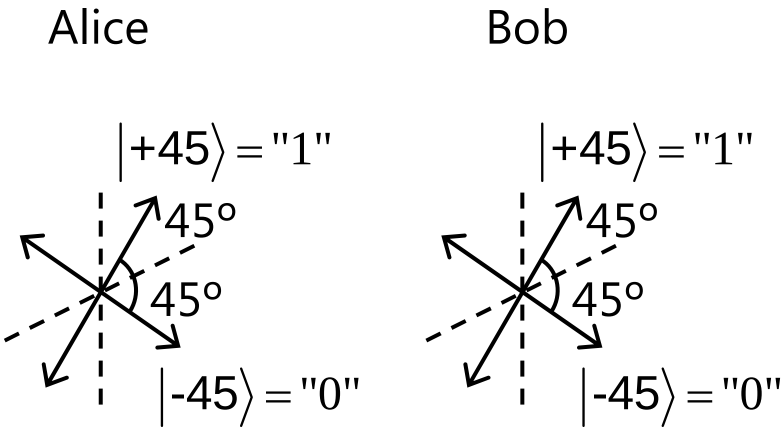
This leads to four possibilities

- In two of these
  - their transmission is meaningful



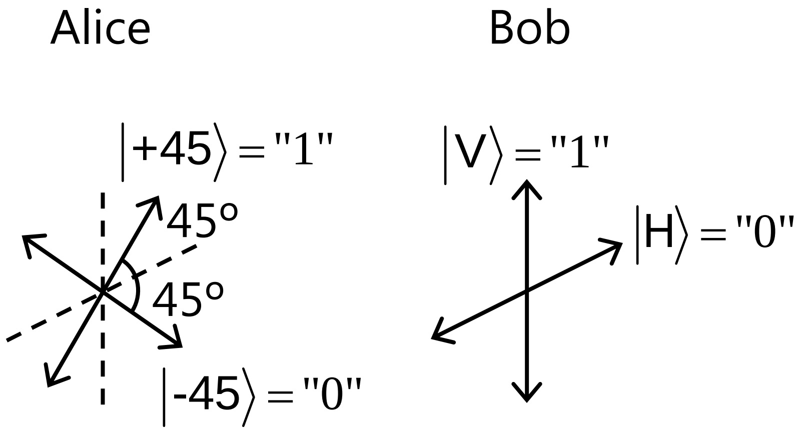
Simple quantum encryption

- This leads to four possibilities
  - In two of these
    - their transmission is meaningful



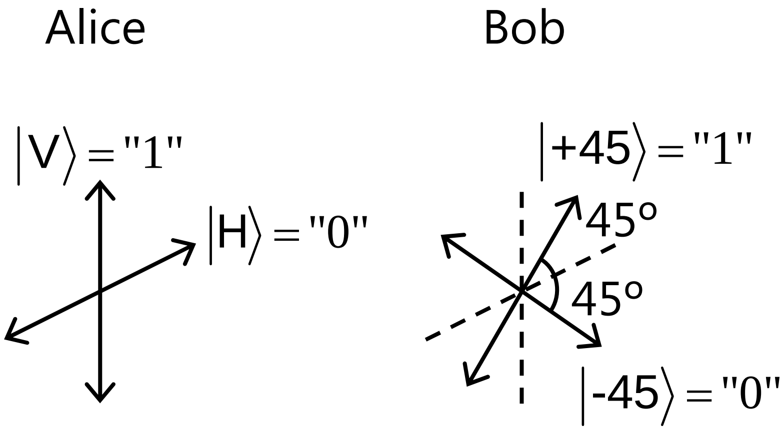
Simple quantum encryption

- This leads to four possibilities
  - In two of these
    - their transmission is meaningful
  - In the other two
    - no information is exchanged



Simple quantum encryption

This leads to four possibilities
- In two of these
  - their transmission is meaningful
- In the other two
  - no information is exchanged



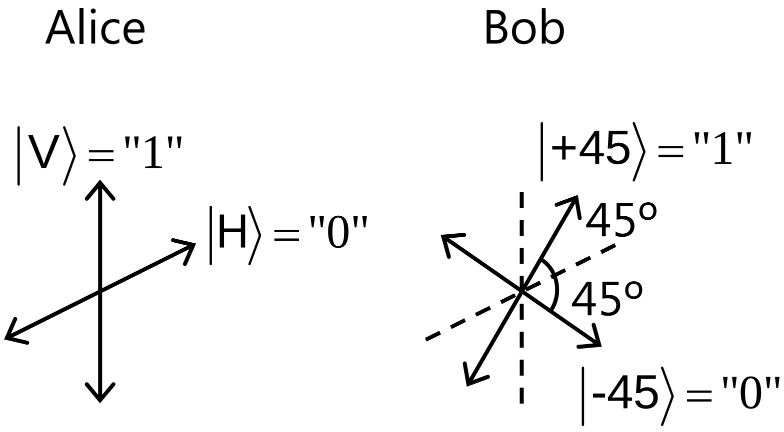
Simple quantum encryption

All that is necessary now

- for successful secure information exchange by Alice and Bob
  - is for them again to

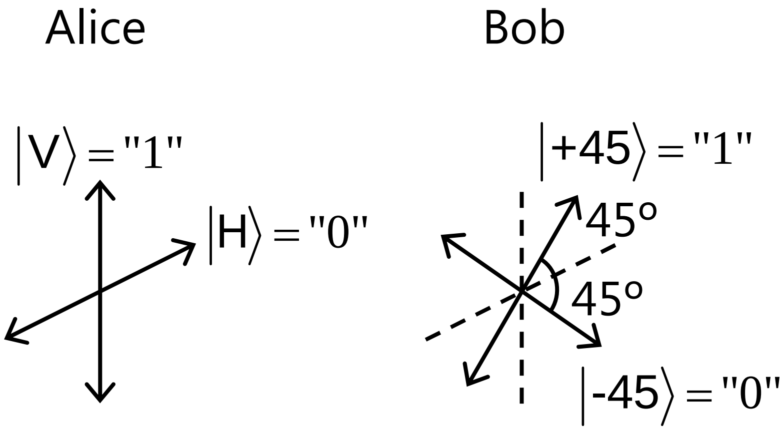
call one another up openly on the telephone and

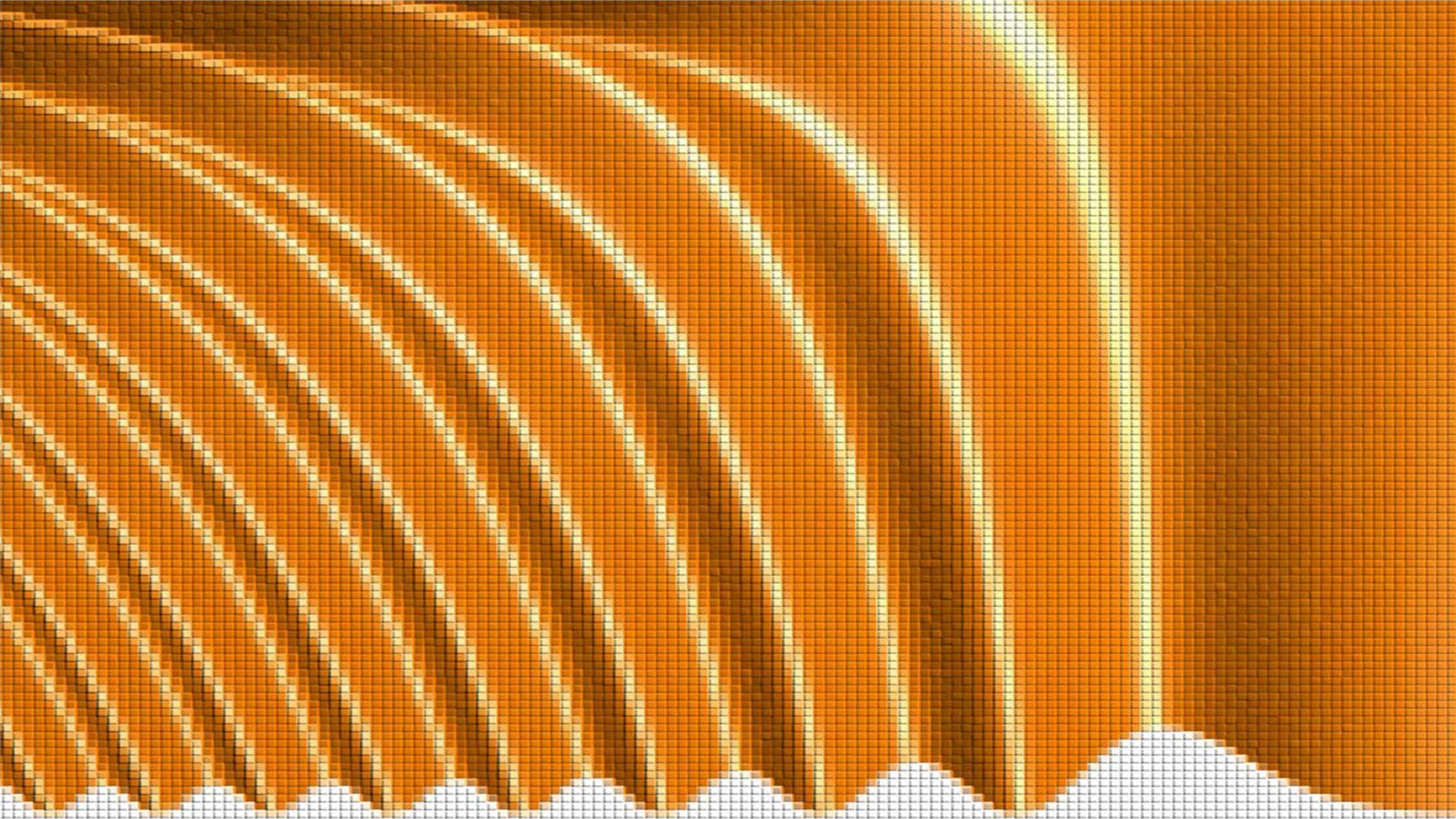
- agree when their polarizers were set the same
  - which they can do without ever revealing what information was exchanged in each case



Simple quantum encryption

- For each case where their polarizers were set the same
  - Alice notes what bit she sent
  - Bob notes what bit he received
- They openly compare some of these bits
  - to see if there are errors introduced by Eve
- If not, they now share secret information
  - in the rest of the bits



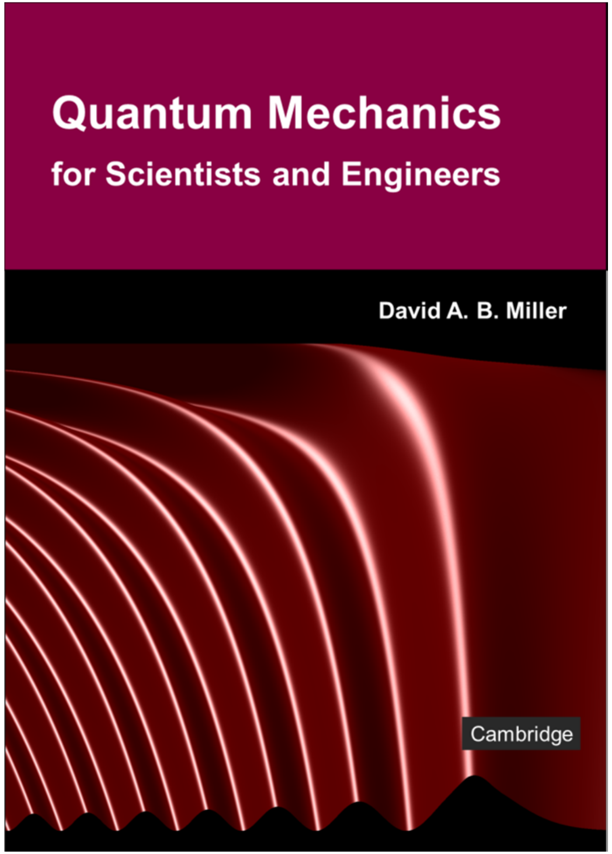


18.2 Quantum computing, teleportation and entanglement

Slides: Video 18.2.1 Entanglement

Text reference: Quantum Mechanics for Scientists and Engineers

Section 18.3



Quantum computing, teleportation and entanglement

Entanglement

Quantum mechanics for scientists and engineers

David Miller

Entanglement

We need to re-examine the states
  of more than one particle

Suppose we have two particles
  e.g., two photons

Entanglement

Photon 1 is in one of a set of possible states $|\psi_m\rangle_1$

e.g., going to the left

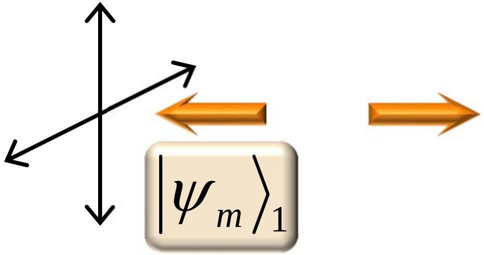
in a particular spatial mode (beam shape)

with a specific frequency

with the different possible states being

vertical

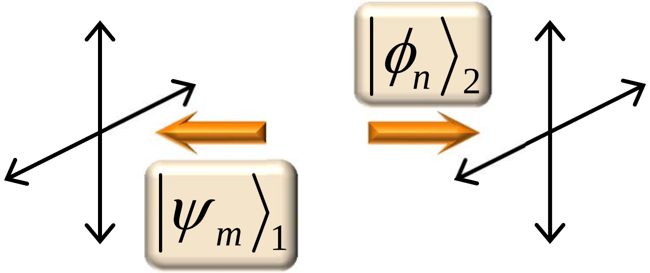
or horizontal polarization



Entanglement

Photon 2 is similarly in one of a set of possible states $|\phi_n\rangle_2$

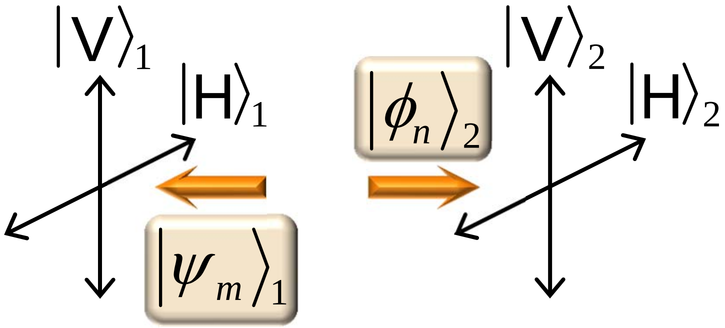
- e.g., going to the right
  - in a particular spatial mode (beam shape)
    - with a specific frequency
      - with the different possible states being
        - vertical
        - or horizontal polarization



Entanglement

Then, appropriate basis states
for the left-going photon 1
would be $|\mathrm{H}\rangle_1$ and $|\mathrm{V}\rangle_1$
where H and V refer to
horizontal and vertical polarization

Similarly, appropriate basis states
for the right-going photon 2
would be $|\mathrm{H}\rangle_2$ and $|\mathrm{V}\rangle_2$



Entanglement

A possible state of these two photons is

$$|H\rangle_1 |V\rangle_2$$

which is the left-going photon

horizontally polarized

and the right-going photon

vertically polarized

Other examples with obvious meanings include

$$|H\rangle_1 |H\rangle_2 \quad |V\rangle_1 |V\rangle_2 \quad |V\rangle_1 |H\rangle_2$$

Entanglement

We can express other polarizations of a given photon

as linear combinations of horizontal and vertical

For example

the state $\left(1/\sqrt{2}\right)\left(|\mathrm{H}\rangle_1 + |\mathrm{V}\rangle_1\right)$ describes

a left-going photon

polarized at an angle of 45°

Hence, a state like $\left(1/\sqrt{2}\right)\left(|\mathrm{H}\rangle_1 + |\mathrm{V}\rangle_1\right)|\mathrm{H}\rangle_2$

describes a left-going photon polarized at 45°

and a right-going photon horizontally polarized

Entanglement

So far, we have assigned each photon a definite polarization

- just as we could classically
  - But, these states are not the only ones allowed by quantum mechanics

For example, consider the following state of the two photons

$$\left|\Phi^{+}\right\rangle_{12}=\frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2}+\left|\mathrm{V}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}\right)$$

- i.e., a linear superposition of
  - the state where both photons are horizontally polarized
  - and the state where both photons are vertically polarized

Entanglement

A state like this $\left|\Phi^{+}\right\rangle_{12}=\frac{1}{\sqrt{2}}\left(|\mathrm{H}\rangle_{1}|\mathrm{H}\rangle_{2}+|\mathrm{V}\rangle_{1}|\mathrm{V}\rangle_{2}\right)$

is a linear superposition

of two of the states we considered already

Quantum mechanically, it is a valid state of the system

It is a vector in the 4-dimensional Hilbert space that

describes the polarization state of two photons

a direct product space in which

$|\mathrm{H}\rangle_{1}|\mathrm{H}\rangle_{2}$, $|\mathrm{V}\rangle_{1}|\mathrm{V}\rangle_{2}$, $|\mathrm{H}\rangle_{1}|\mathrm{V}\rangle_{2}$, and $|\mathrm{V}\rangle_{1}|\mathrm{H}\rangle_{2}$

are appropriate orthonormal basis vectors

Entangled states

The state $\left|\Phi^{+}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_1\left|\mathrm{H}\right\rangle_2 + \left|\mathrm{V}\right\rangle_1\left|\mathrm{V}\right\rangle_2\right)$

is very nonclassical

It cannot be factorized into a product of

a state of particle 1 and

a state of particle 2

States that cannot be factorized into a product

of the states of individual systems on their own

are said to be entangled

Entangled states

In such an entangled state $\left|\Phi^{+}\right\rangle_{12}=\frac{1}{\sqrt{2}}\left(|\mathrm{H}\rangle_1|\mathrm{H}\rangle_2+|\mathrm{V}\rangle_1|\mathrm{V}\rangle_2\right)$

particle 1 does not have a definite state of its own

independent of the state of particle 2

Imagine we measure the polarization

of the left-going photon (photon 1)

and find it is horizontal

Then we have collapsed the overall state

into one that now only has terms in $|\mathrm{H}\rangle_1$

So the state of the whole system now is $|\mathrm{H}\rangle_1|\mathrm{H}\rangle_2$

Bell states

There are three other states like the one already considered

that together constitute the four Bell states

$$\left|\Phi^{+}\right\rangle_{12}=\frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2}+\left|\mathrm{V}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}\right) \quad \left|\Phi^{-}\right\rangle_{12}=\frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2}-\left|\mathrm{V}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}\right)$$

$$\left|\Psi^{+}\right\rangle_{12}=\frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}+\left|\mathrm{V}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2}\right) \quad \left|\Psi^{-}\right\rangle_{12}=\frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}-\left|\mathrm{V}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2}\right)$$

These four Bell states are orthogonal

and are a complete basis

for describing any such two-particle system

with two basis states per particle (here, $\left|\mathrm{H}\right\rangle$ and $\left|\mathrm{V}\right\rangle$)

Entangled states

For the two particles considered here

each with two basis states

the required Hilbert space is four-dimensional

so the most general quantum mechanical state

of these two photons is

$$|\psi\rangle = c_{\mathrm{HH}}|\mathrm{H}\rangle_1|\mathrm{H}\rangle_2 + c_{\mathrm{HV}}|\mathrm{H}\rangle_1|\mathrm{V}\rangle_2 + c_{\mathrm{VH}}|\mathrm{V}\rangle_1|\mathrm{H}\rangle_2 + c_{\mathrm{VV}}|\mathrm{V}\rangle_1|\mathrm{V}\rangle_2$$

where now we need four (generally complex) coefficients

the four different c's

to specify the state of just two photons

Entangled states

For three particles

we need eight coefficients

For four particles

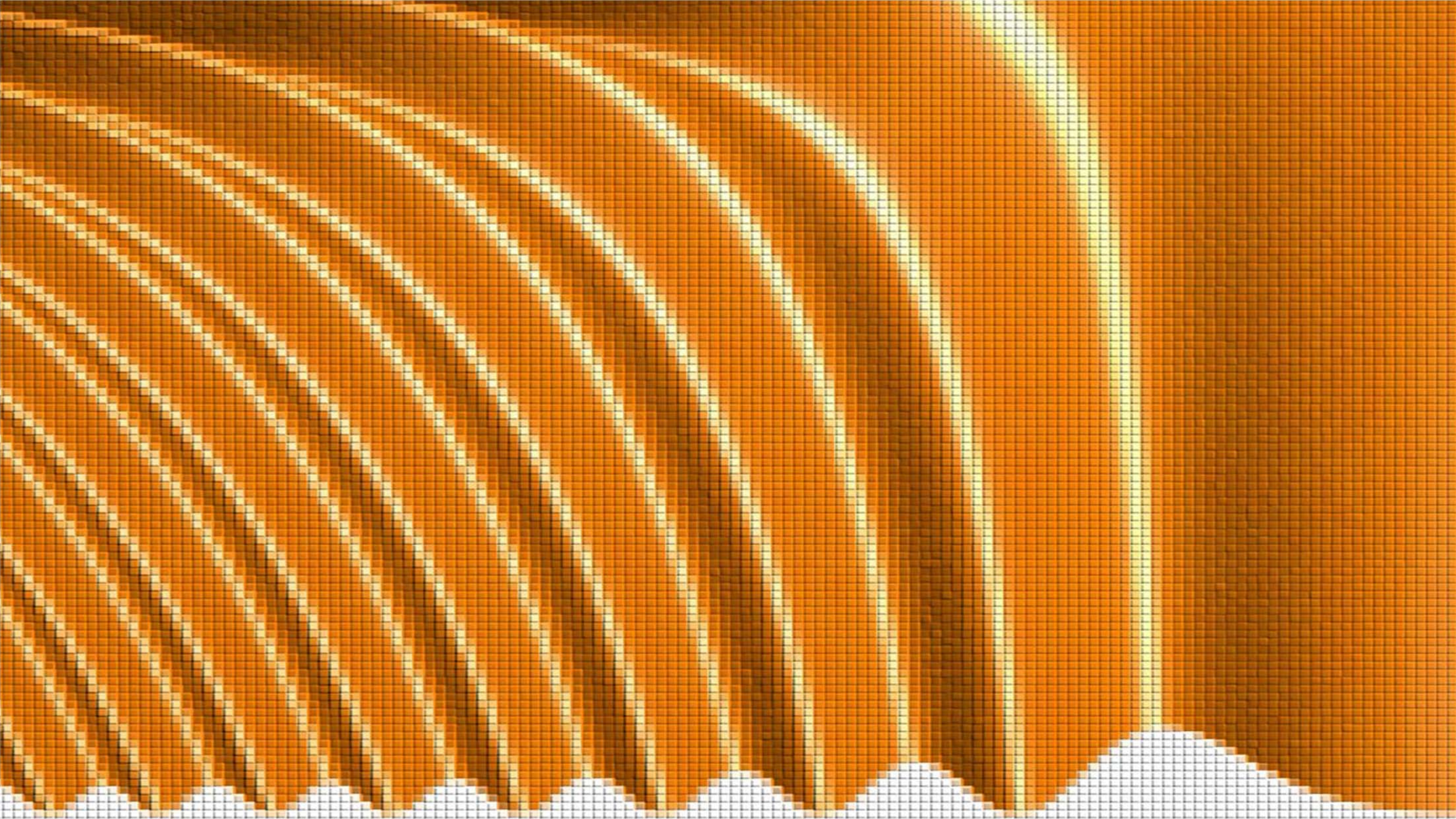
sixteen coefficients

and so on

leading to 2^N coefficients for N particles

300 particles would therefore require 2^{300} coefficients

a number that may be larger than the number of atoms in the observable universe!

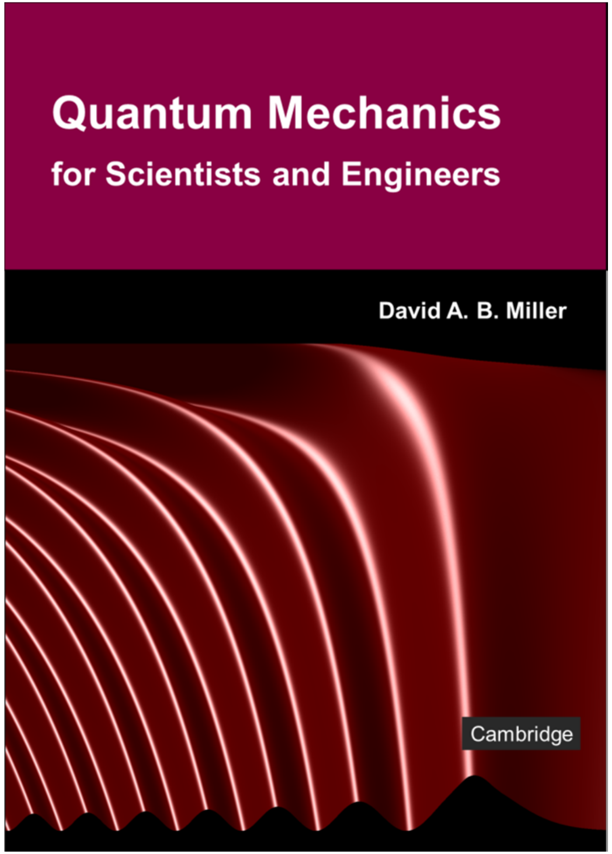


18.2 Quantum computing, teleportation and entanglement

Slides: Video 18.2.3 Quantum computing and qubits

Text reference: Quantum Mechanics for Scientists and Engineers

Section 18.4



Quantum computing, teleportation and entanglement

Quantum computing and qubits

Quantum mechanics for scientists and engineers

David Miller

Representing quantum information – qubits

The qubit state can be written $|\psi\rangle = c_0|0\rangle + c_1|1\rangle \equiv \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$

where $|0\rangle$ is the quantum mechanical state that represents "0"

for example

a horizontal polarization state $|\mathrm{H}\rangle$ of a photon

a spin-down state $|\downarrow\rangle$ of an electron, or

a ground state $|g\rangle$ of an atom

Representing quantum information – qubits

Similarly, in $|\psi\rangle = c_0|0\rangle + c_1|1\rangle \equiv \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$

$|1\rangle$ is the quantum mechanical state representing "1"

and could be physically represented by, for example

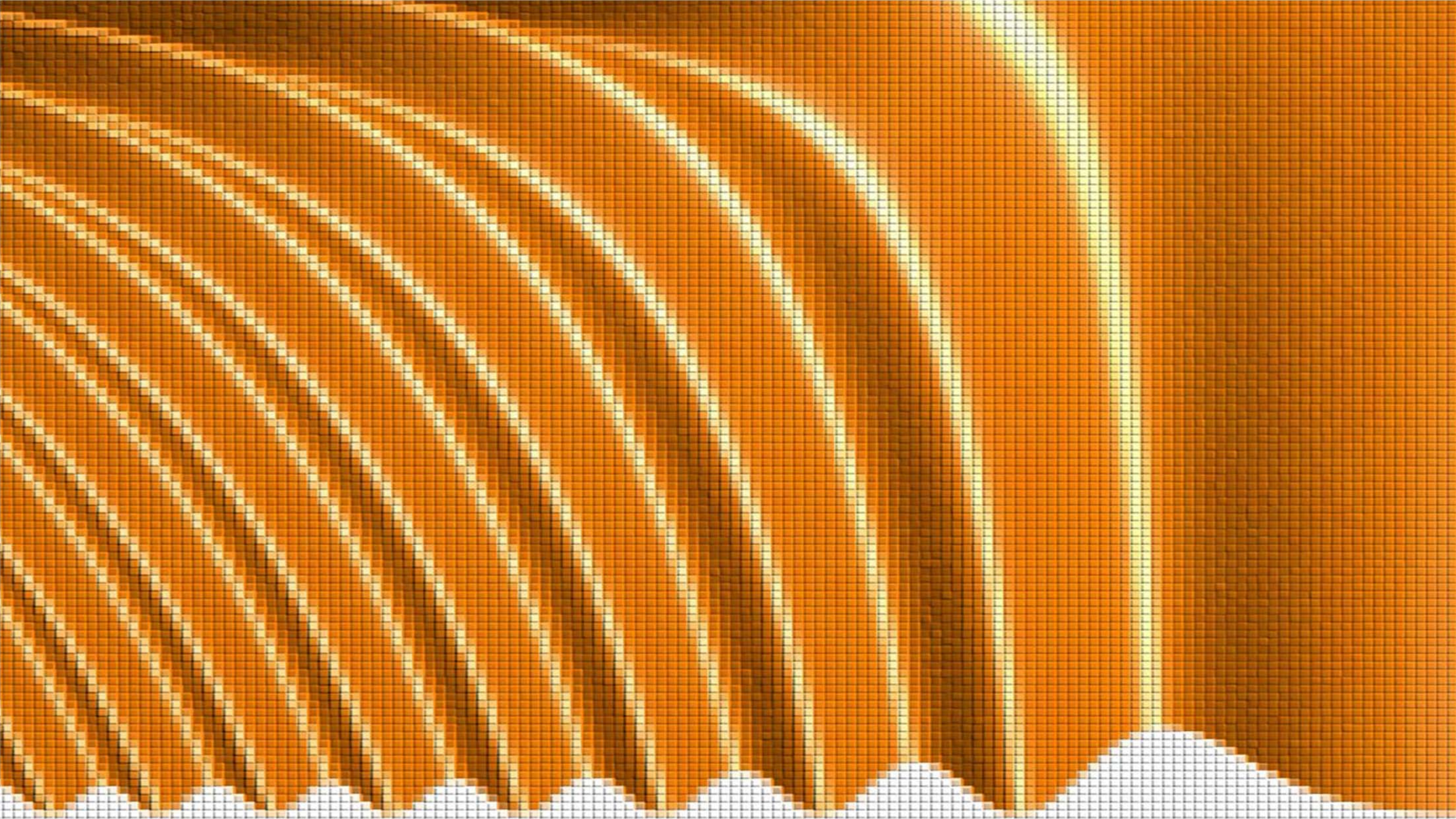
vertical polarization $|\mathrm{V}\rangle$

spin-up $|\uparrow\rangle$, or

an excited atomic state $|e\rangle$

Because of normalization

$$|c_0|^2 + |c_1|^2 = 1$$

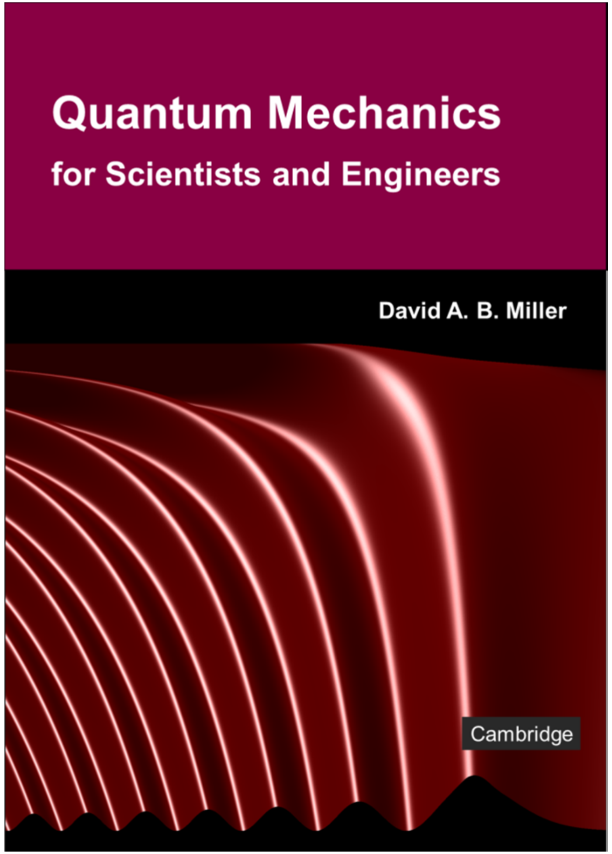


18.2 Quantum computing, teleportation and entanglement

Slides: Video 18.2.5 Running a quantum computer

Text reference: Quantum Mechanics for Scientists and Engineers

Section 18.4



Quantum computing, teleportation and entanglement

Running a quantum computer

Quantum mechanics for scientists and engineers

David Miller

One-qubit gates and operations

A qubit itself can be written as a two element vector

- such as we could use to represent a spin state
  - such as $\begin{bmatrix} a \\ b \end{bmatrix}$

The necessary basic operations for a quantum computer

- can be written as four different operations

Three of these are operations on a single qubit

- We can write these operations as 2 x 2 matrices
  - representing the corresponding unitary operators

One-qubit gates and operations

One possible set of single qubit operations is

$$\hat{U}_H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \hat{U}_Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \hat{U}_{NOTX} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

These unitary operators are known as

Hadamard, $\hat{U}_H$

Z, $\hat{U}_Z$, and

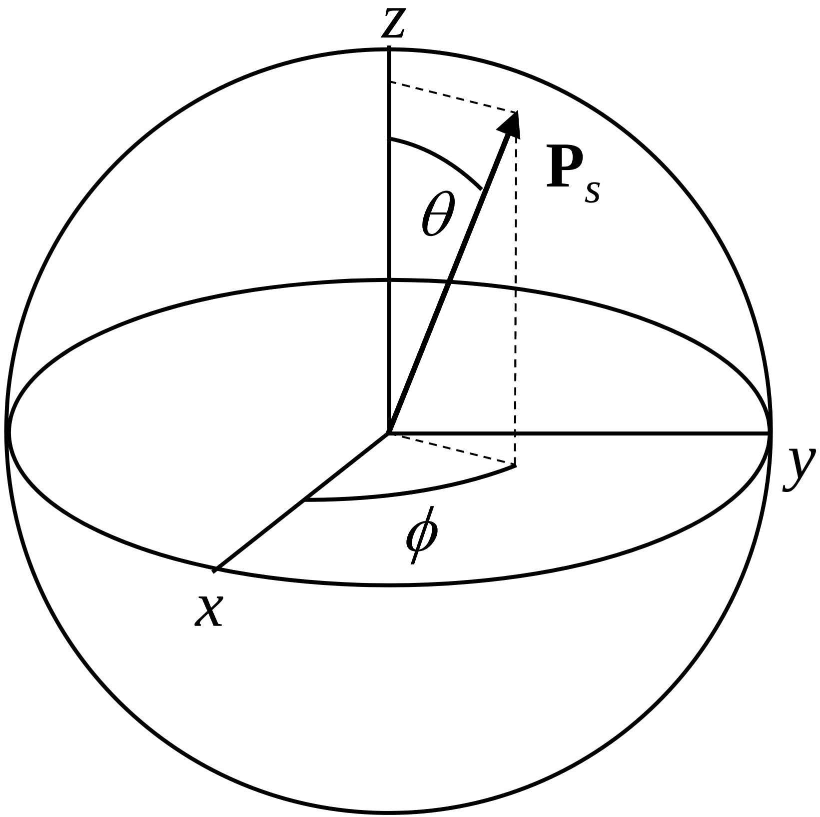
NOT X, $\hat{U}_{NOTX}$

operators

One-qubit operations

Using the Bloch sphere
- we can represent a qubit
  - as a vector, such as
    - the spin polarization vector
  - pointing from
    - the center of a sphere
      - to its surface

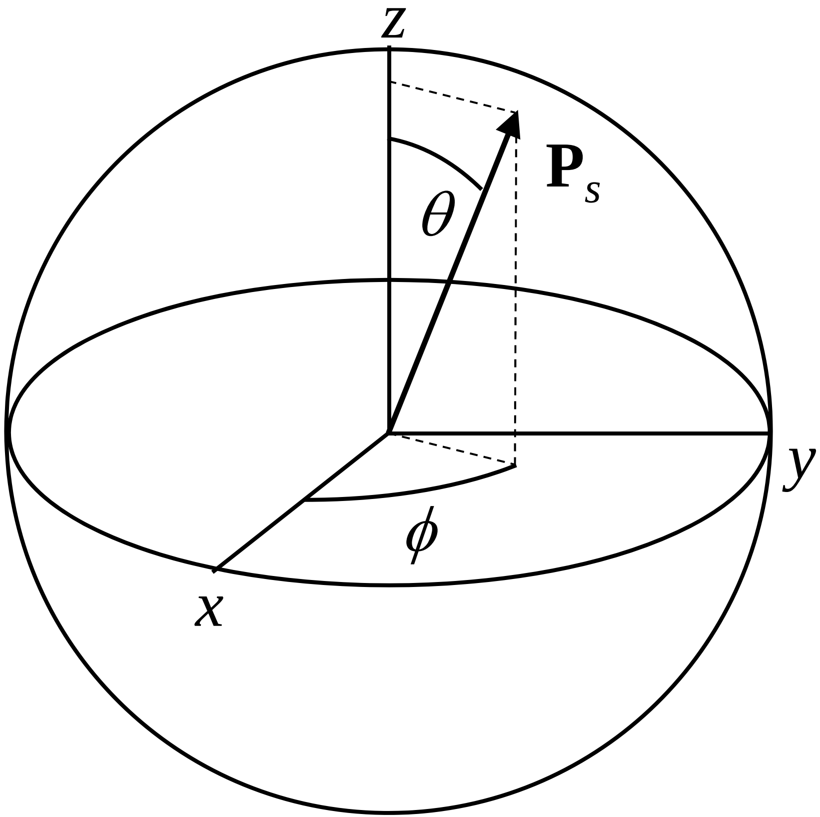
Single qubit operations
- correspond to rotations
  - of the vector on the sphere



One-qubit operations

Single qubit operations can be achieved

- for spins
  - by appropriate pulses of magnetic fields
- for two-level “atomic” systems
  - by pulses of electromagnetic fields
- for photons
  - by changing the polarization



Two-qubit gates

The fourth required operation

- interacts two qubits in a "Controlled-NOT" (C-NOT)
  - One qubit is called the control
  - The other is called the target

If the control is $|0\rangle$

- the target qubit is passed through unchanged
  - but if the control is $|1\rangle$, the target qubit is inverted
    - a target qubit of state $|0\rangle$ is changed to state $|1\rangle$
    - and a target qubit of $|1\rangle$ is changed to state $|0\rangle$
      - hence the name Controlled-NOT

Two-qubit gates

A two-qubit state is a vector in a four-dimensional Hilbert space

that is, like a state of two photons on different paths

$$|\psi\rangle = c_{\mathrm{HH}}|\mathrm{H}\rangle_1|\mathrm{H}\rangle_2 + c_{\mathrm{HV}}|\mathrm{H}\rangle_1|\mathrm{V}\rangle_2 + c_{\mathrm{VH}}|\mathrm{V}\rangle_1|\mathrm{H}\rangle_2 + c_{\mathrm{VV}}|\mathrm{V}\rangle_1|\mathrm{V}\rangle_2$$

which we could rewrite in the notation

$$|\psi\rangle = \begin{array}{l} c_{00}|0\rangle_{control}|0\rangle_{target} + c_{01}|0\rangle_{control}|1\rangle_{target} \\ + c_{10}|1\rangle_{control}|0\rangle_{target} + c_{11}|1\rangle_{control}|1\rangle_{target} \end{array} \equiv \begin{bmatrix} c_{00} \\ c_{01} \\ c_{10} \\ c_{11} \end{bmatrix}$$

Two-qubit gates

The corresponding operator

in this four-dimensional Hilbert space

for this C-NOT gate

can be written

$$\hat{U}_{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

as we will now demonstrate

Two-qubit gates

For example, the input state with

the control as a logic 0 and the target as a logic 1

when written in our form

$$|\psi\rangle = \begin{matrix} c_{00}|0\rangle_{control}|0\rangle_{target} + c_{01}|0\rangle_{control}|1\rangle_{target} \\ +c_{10}|1\rangle_{control}|0\rangle_{target} + c_{11}|1\rangle_{control}|1\rangle_{target} \end{matrix} \equiv \begin{bmatrix} c_{00} \\ c_{01} \\ c_{10} \\ c_{11} \end{bmatrix}$$

has $c_{00} = 0$, $c_{01} = 1$, $c_{10} = 0$, and $c_{11} = 0$

Writing that state as a column vector gives $|\psi\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

Two-qubit gates

Starting with this state

and operating with $\hat{U}_{CNOT}$ gives

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

which is just the state we started with

As intended, the target qubit passes through unchanged

if the control qubit is logic 0

Two-qubit gates

Alternatively, we could choose an input state where

the control qubit is a logic 1 and the target is a logic 1

that is, $c_{00} = 0$, $c_{01} = 0$, $c_{10} = 0$, and $c_{11} = 1$

Then our input state can be written

$$\left|\psi\right\rangle = \begin{array}{l} c_{00}\left|0\right\rangle_{control}\left|0\right\rangle_{target} + c_{01}\left|0\right\rangle_{control}\left|1\right\rangle_{target} \\ +c_{10}\left|1\right\rangle_{control}\left|0\right\rangle_{target} + c_{11}\left|1\right\rangle_{control}\left|1\right\rangle_{target} \end{array} \equiv \begin{bmatrix} c_{00} \\ c_{01} \\ c_{10} \\ c_{11} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Two-qubit gates

Acting on this input state with $\hat{U}_{CNOT}$ gives

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

This output state is $c_{00} = 0$, $c_{01} = 0$, $c_{10} = 1$, and $c_{11} = 0$

- The target qubit is now a logic 0
  - it has been "flipped"
- and the control bit remains at logic 1

Hypothetical two-qubit gate

We could imagine two two-level systems

a "control" system

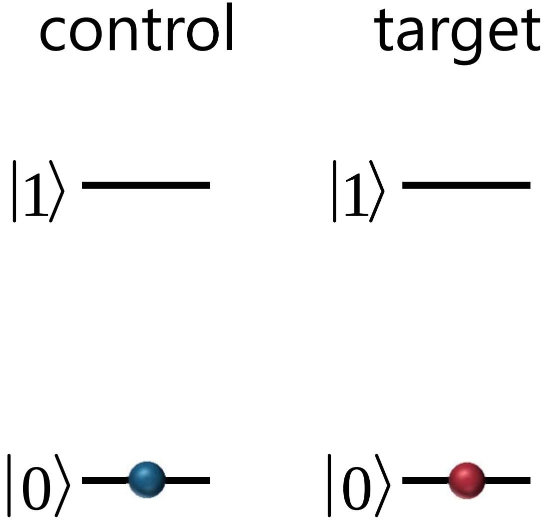
and a "target" system

We shine a "clocking" light pulse at the target system

If the control qubit is in its $|0\rangle$ state

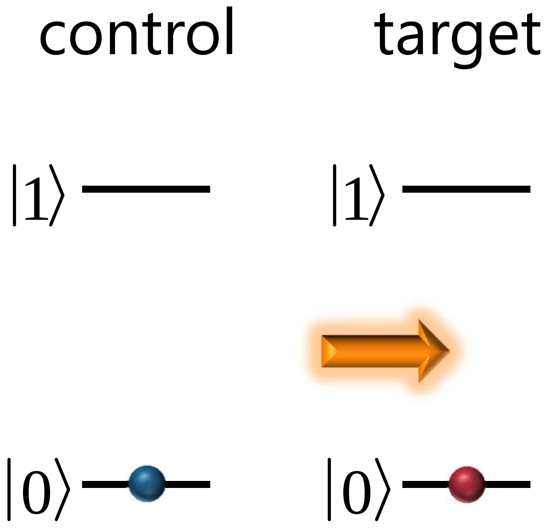
then this clock pulse does nothing to the target qubit system

leaving it in its lower state



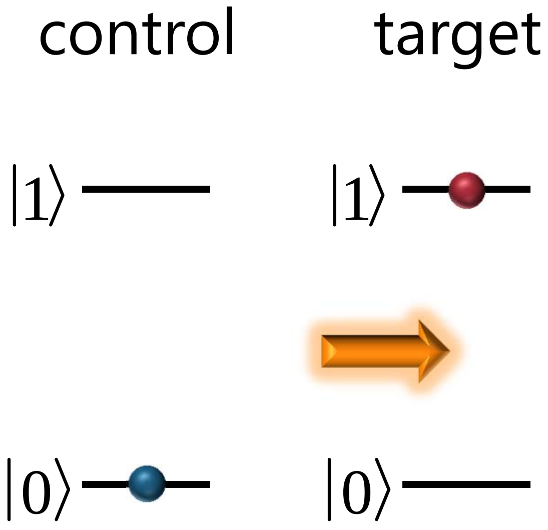
Hypothetical two-qubit gate

- We could imagine two two-level systems
  - a “control” system
  - and a “target” system
- We shine a “clocking” light pulse at the target system
- If the control qubit is in its $|0\rangle$ state
  - then this clock pulse does nothing to the target qubit system
    - leaving it in its lower state



Hypothetical two-qubit gate

- We could imagine two two-level systems
  - a “control” system
  - and a “target” system
- We shine a “clocking” light pulse at the target system
- If the control qubit is in its $|0\rangle$ state
  - then this clock pulse does nothing to the target qubit system
    - leaving it in its lower state
      - or its upper state



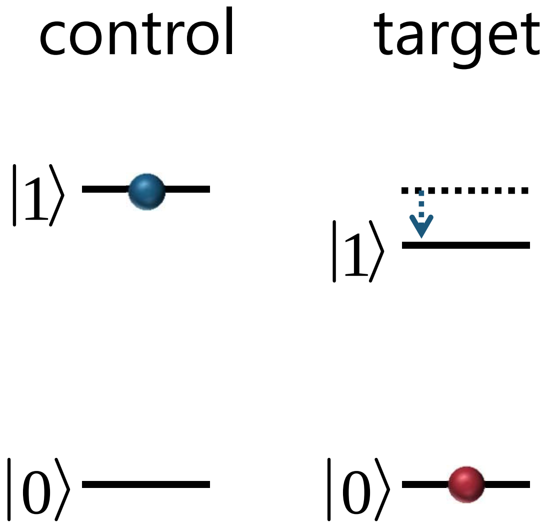
Hypothetical two-qubit gate

If the control qubit is in its $|1\rangle$ state

- perhaps it changes a transition frequency in the target qubit system
  - through some interaction between the control and target qubit systems

With this change in transition energy

- the target qubit system could then be sensitive to the “clock” pulse
  - which then flips target qubit state
    - implementing the C-NOT function



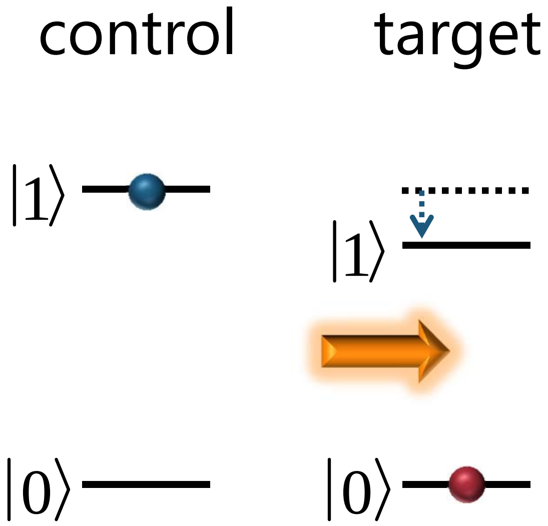
Hypothetical two-qubit gate

If the control qubit is in its $|1\rangle$ state

- perhaps it changes a transition frequency in the target qubit system
  - through some interaction between the control and target qubit systems

With this change in transition energy

- the target qubit system could then be sensitive to the "clock" pulse
  - which then flips target qubit state
    - implementing the C-NOT function



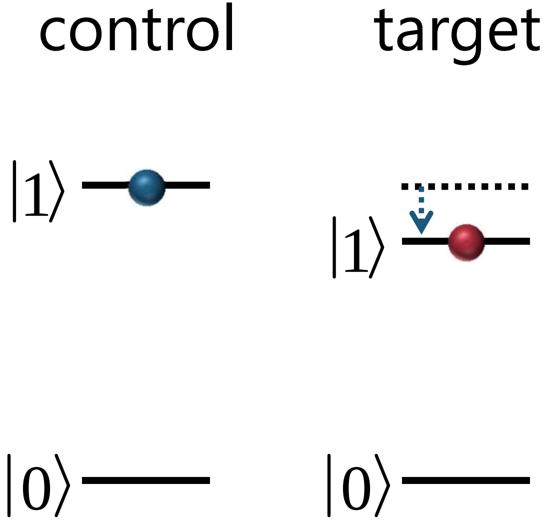
Hypothetical two-qubit gate

If the control qubit is in its $|1\rangle$ state

- perhaps it changes a transition frequency in the target qubit system
  - through some interaction between the control and target qubit systems

With this change in transition energy

- the target qubit system could then be sensitive to the “clock” pulse
  - which then flips target qubit state
    - implementing the C-NOT function



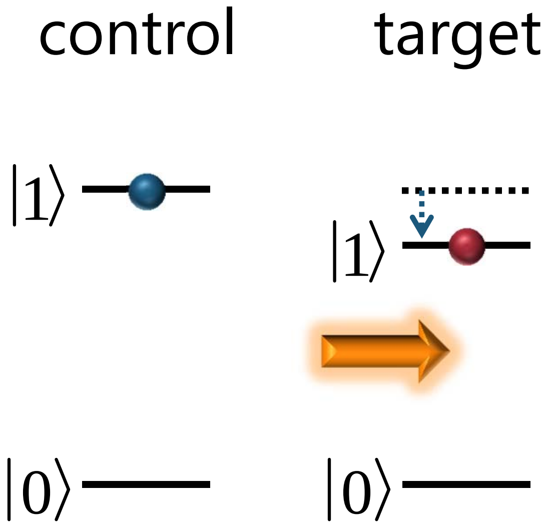
Hypothetical two-qubit gate

If the control qubit is in its $|1\rangle$ state

- perhaps it changes a transition frequency in the target qubit system
  - through some interaction between the control and target qubit systems

With this change in transition energy

- the target qubit system could then be sensitive to the "clock" pulse
  - which then flips target qubit state
    - implementing the C-NOT function



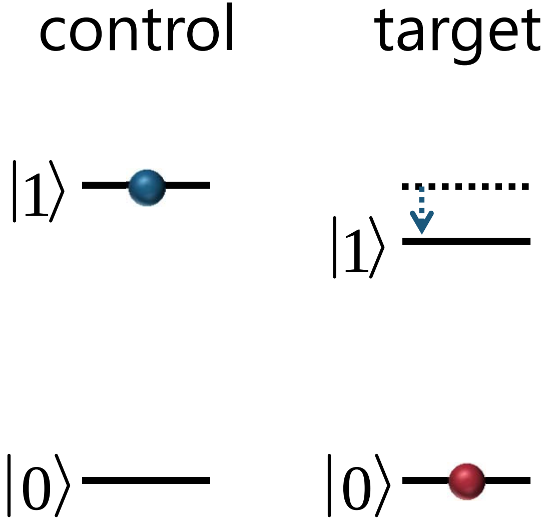
Hypothetical two-qubit gate

If the control qubit is in its $|1\rangle$ state

- perhaps it changes a transition frequency in the target qubit system
  - through some interaction between the control and target qubit systems

With this change in transition energy

- the target qubit system could then be sensitive to the "clock" pulse
  - which then flips target qubit state
    - implementing the C-NOT function



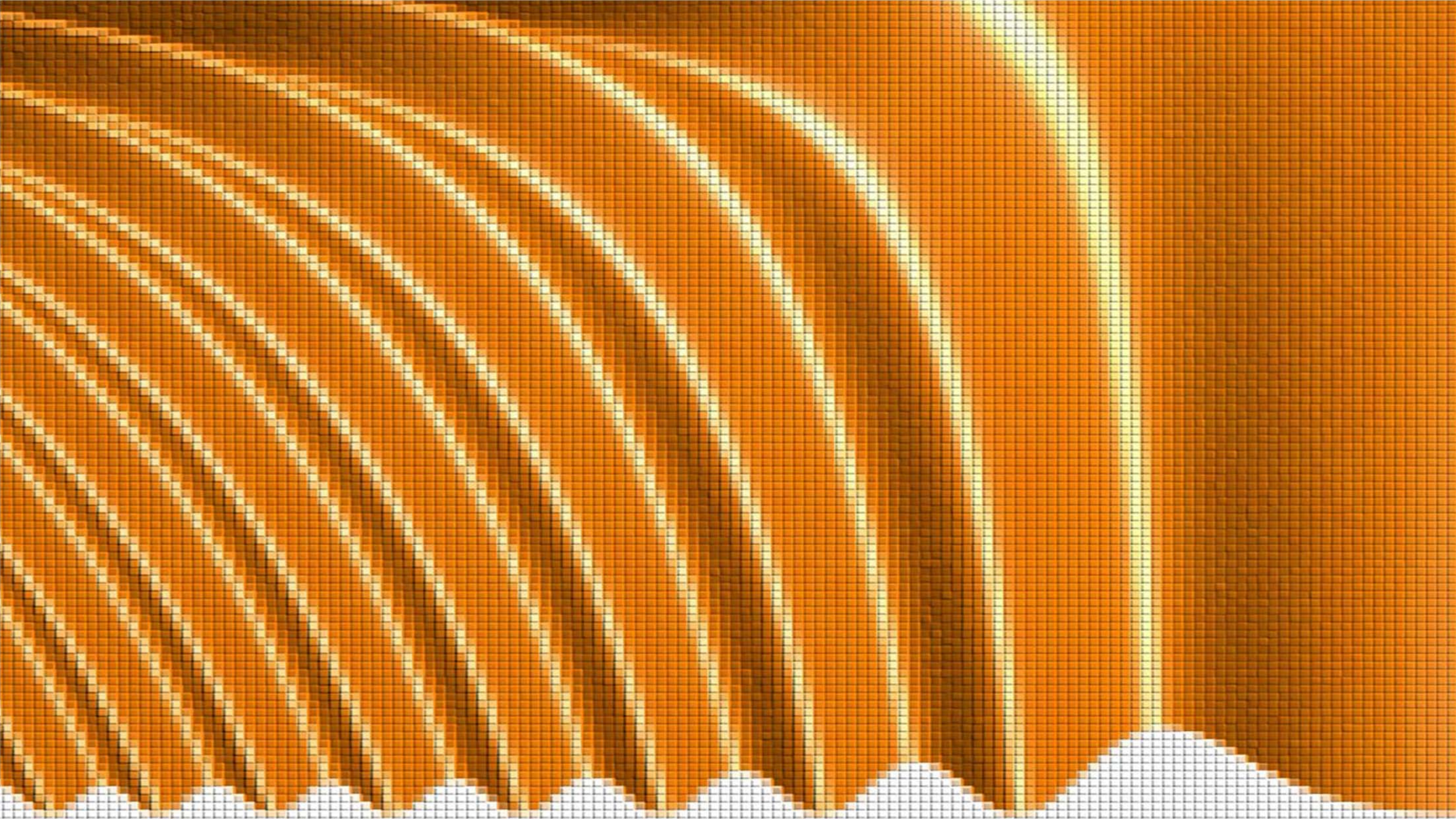
Example approaches to two-qubit gates

The above example was hypothetical

- and simplified

Example systems for real two-qubit gates include

- ions in ion traps
- superconducting flux and charge qubits
- quantum dots, and
- spins in semiconductor impurities

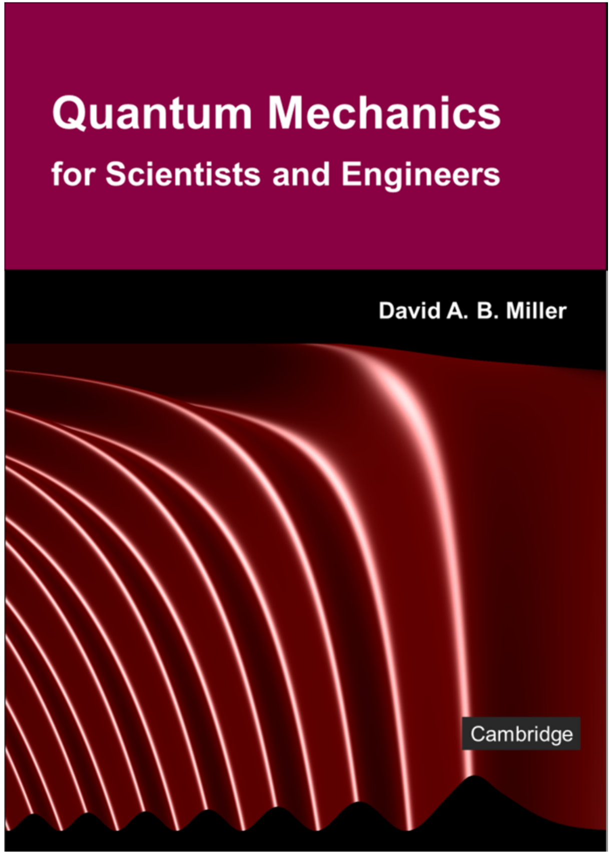


18.2 Quantum computing, teleportation and entanglement

Slides: Video 18.2.7 Quantum teleportation

Text reference: Quantum Mechanics for Scientists and Engineers

Section 18.5



Quantum computing, teleportation and entanglement

Quantum teleportation

Quantum mechanics for scientists and engineers

David Miller

Quantum teleportation

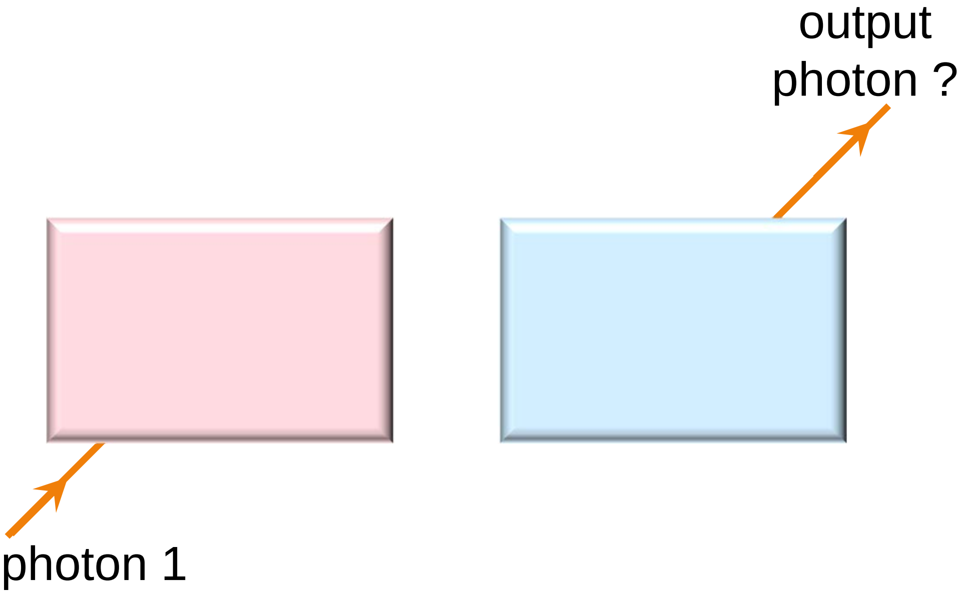
The idea of quantum teleportation is to transfer a quantum state from one place to another without transferring the specific carrier of that state

Quantum teleportation

Suppose photon 1 is in an
- unknown superposition of horizontal and vertical polarization

We want the output photon in the same state
- but without sending photon 1 there

We may even destroy (absorb) photon 1



Quantum teleportation

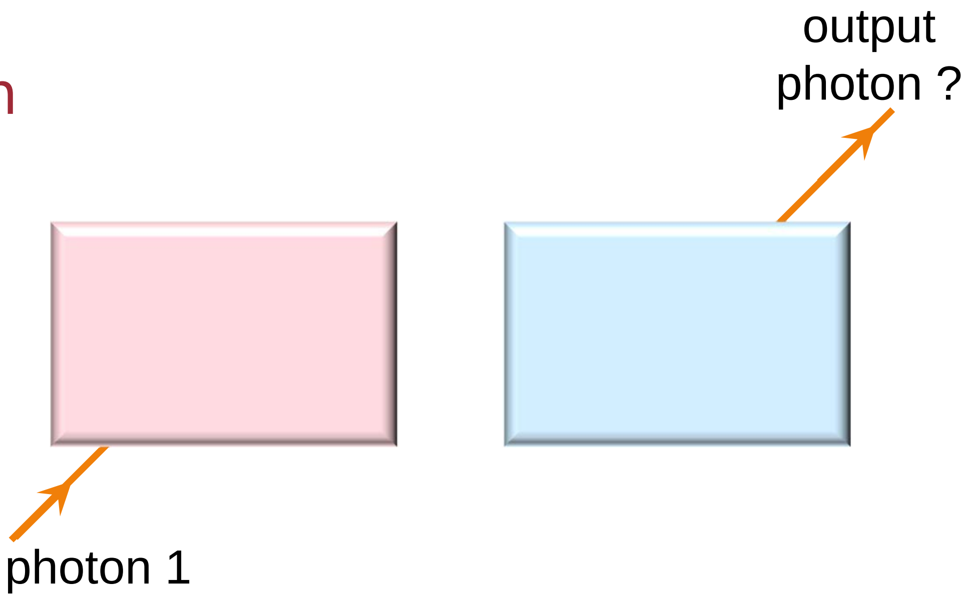
But we know

from the no-cloning theorem that

we cannot clone photon 1

to produce another (output) photon

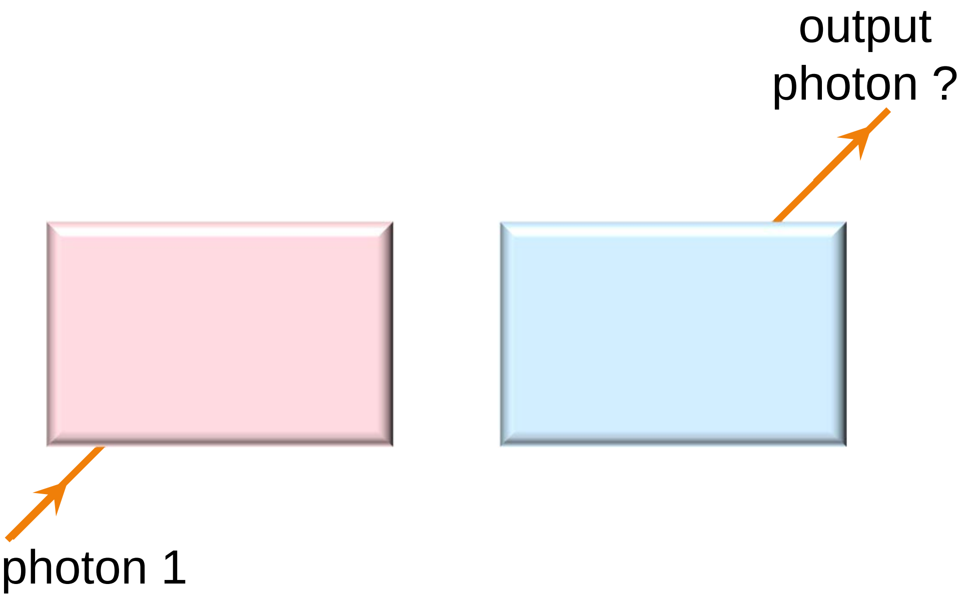
in the same arbitrary superposition



Quantum teleportation

We also know that

- simply measuring photon 1
  - e.g., with a polarizing beamsplitter
    - together with photodetectors
- will not reliably tell us
  - the full quantum state of photon 1



Quantum teleportation

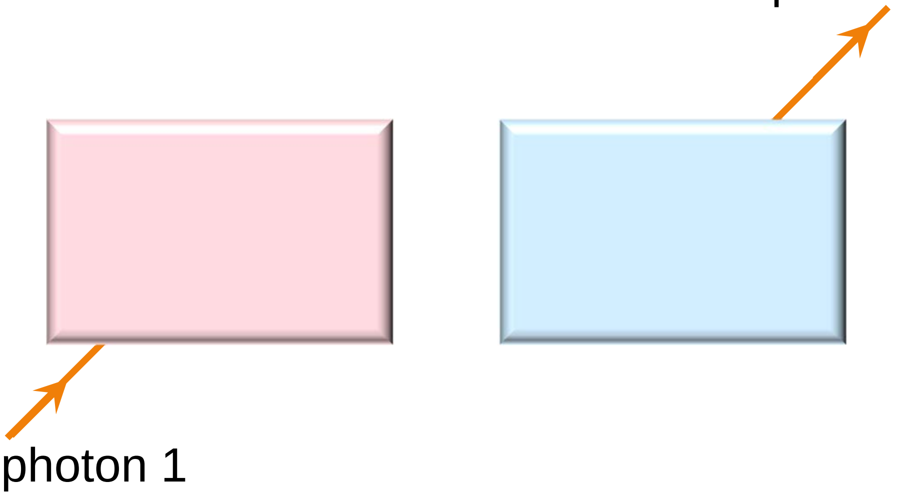
We end up

statistically “collapsing” the state

and throwing away information

about the original quantum state

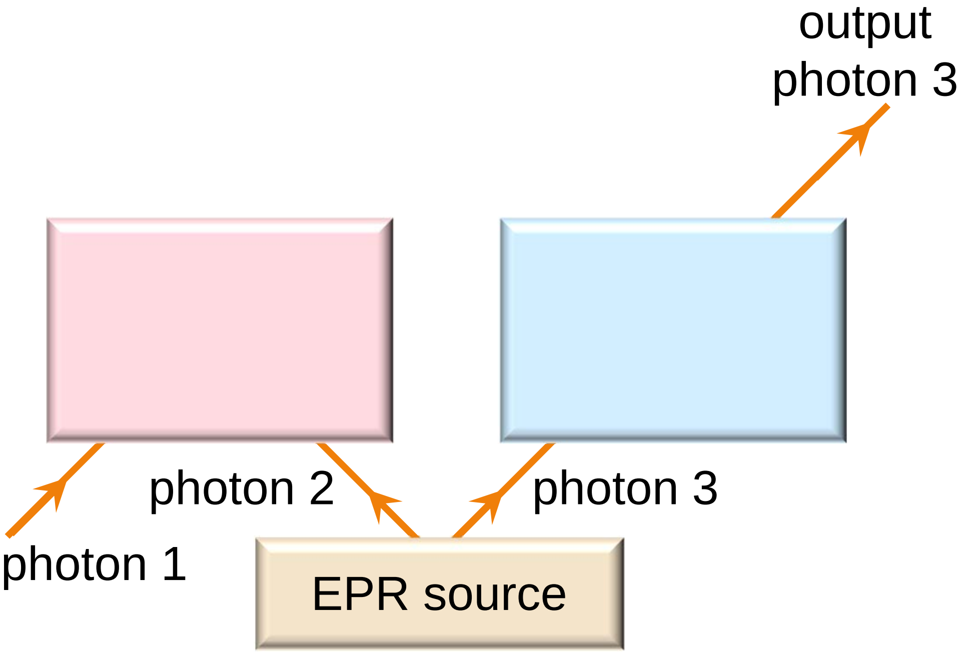
of the photon



Quantum teleportation

The key to quantum teleportation

- is to "share entanglement"
  - by sharing an "EPR" pair of photons
    - that are in a Bell (entangled) state



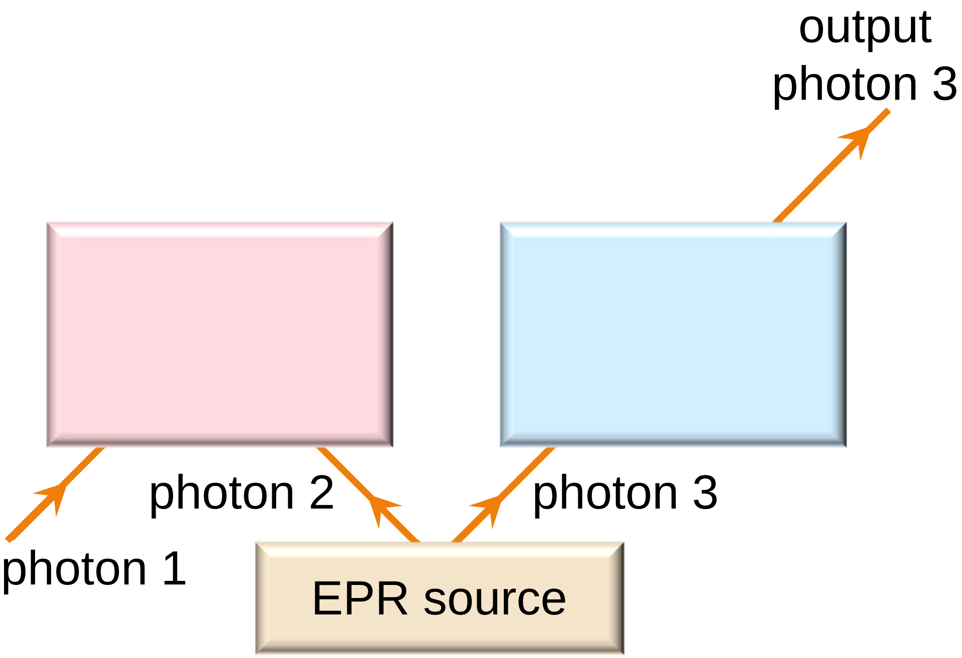
Quantum teleportation

The EPR photon pair is presumed to be in the Bell state

$$|\Psi^-\rangle_{23} = \frac{1}{\sqrt{2}}\left(|\mathrm{H}\rangle_2 |\mathrm{V}\rangle_3 - |\mathrm{V}\rangle_2 |\mathrm{H}\rangle_3\right)$$

The input photon is in an (unknown) superposition of horizontal and vertical polarizations

$$|\psi\rangle_1 = c_\mathrm{H} |\mathrm{H}\rangle_1 + c_\mathrm{V} |\mathrm{V}\rangle_1$$



Quantum teleportation

The state of all three photons, therefore, can be written

$$\left|\Psi\right\rangle_{123}=\frac{1}{\sqrt{2}}\left(c_{\mathrm{H}}\left|\mathrm{H}\right\rangle_{1}+c_{\mathrm{V}}\left|\mathrm{V}\right\rangle_{1}\right)\left(\left|\mathrm{H}\right\rangle_{2}\left|\mathrm{V}\right\rangle_{3}-\left|\mathrm{V}\right\rangle_{2}\left|\mathrm{H}\right\rangle_{3}\right)$$

A core trick in the teleportation is to note that

this state can be rewritten as

$$\begin{aligned}\left|\Psi\right\rangle_{123}=\frac{1}{2}\Big[&\left|\Phi^{+}\right\rangle_{12}\left(c_{\mathrm{H}}\left|\mathrm{V}\right\rangle_{3}-c_{\mathrm{V}}\left|\mathrm{H}\right\rangle_{3}\right)+\left|\Phi^{-}\right\rangle_{12}\left(c_{\mathrm{H}}\left|\mathrm{V}\right\rangle_{3}+c_{\mathrm{V}}\left|\mathrm{H}\right\rangle_{3}\right)\\&+\left|\Psi^{+}\right\rangle_{12}\left(-c_{\mathrm{H}}\left|\mathrm{H}\right\rangle_{3}+c_{\mathrm{V}}\left|\mathrm{V}\right\rangle_{3}\right)-\left|\Psi^{-}\right\rangle_{12}\left(c_{\mathrm{H}}\left|\mathrm{H}\right\rangle_{3}+c_{\mathrm{V}}\left|\mathrm{V}\right\rangle_{3}\right)\Big]\end{aligned}$$

Quantum teleportation

Note that we have managed to write this state

$$|\Psi\rangle_{123} = \frac{1}{2}\Big[|\Phi^+\rangle_{12}\left(c_{\mathrm{H}}|\mathrm{V}\rangle_3 - c_{\mathrm{V}}|\mathrm{H}\rangle_3\right) + |\Phi^-\rangle_{12}\left(c_{\mathrm{H}}|\mathrm{V}\rangle_3 + c_{\mathrm{V}}|\mathrm{H}\rangle_3\right)$$

$$+|\Psi^+\rangle_{12}\left(-c_{\mathrm{H}}|\mathrm{H}\rangle_3 + c_{\mathrm{V}}|\mathrm{V}\rangle_3\right) - |\Psi^-\rangle_{12}\left(c_{\mathrm{H}}|\mathrm{H}\rangle_3 + c_{\mathrm{V}}|\mathrm{V}\rangle_3\right)\Big]$$

in terms of the Bell states of photons 1 and 2

$$|\Phi^+\rangle_{12} = \frac{1}{\sqrt{2}}\left(|\mathrm{H}\rangle_1|\mathrm{H}\rangle_2 + |\mathrm{V}\rangle_1|\mathrm{V}\rangle_2\right) \qquad |\Phi^-\rangle_{12} = \frac{1}{\sqrt{2}}\left(|\mathrm{H}\rangle_1|\mathrm{H}\rangle_2 - |\mathrm{V}\rangle_1|\mathrm{V}\rangle_2\right)$$

$$|\Psi^+\rangle_{12} = \frac{1}{\sqrt{2}}\left(|\mathrm{H}\rangle_1|\mathrm{V}\rangle_2 + |\mathrm{V}\rangle_1|\mathrm{H}\rangle_2\right) \qquad |\Psi^-\rangle_{12} = \frac{1}{\sqrt{2}}\left(|\mathrm{H}\rangle_1|\mathrm{V}\rangle_2 - |\mathrm{V}\rangle_1|\mathrm{H}\rangle_2\right)$$

Quantum teleportation

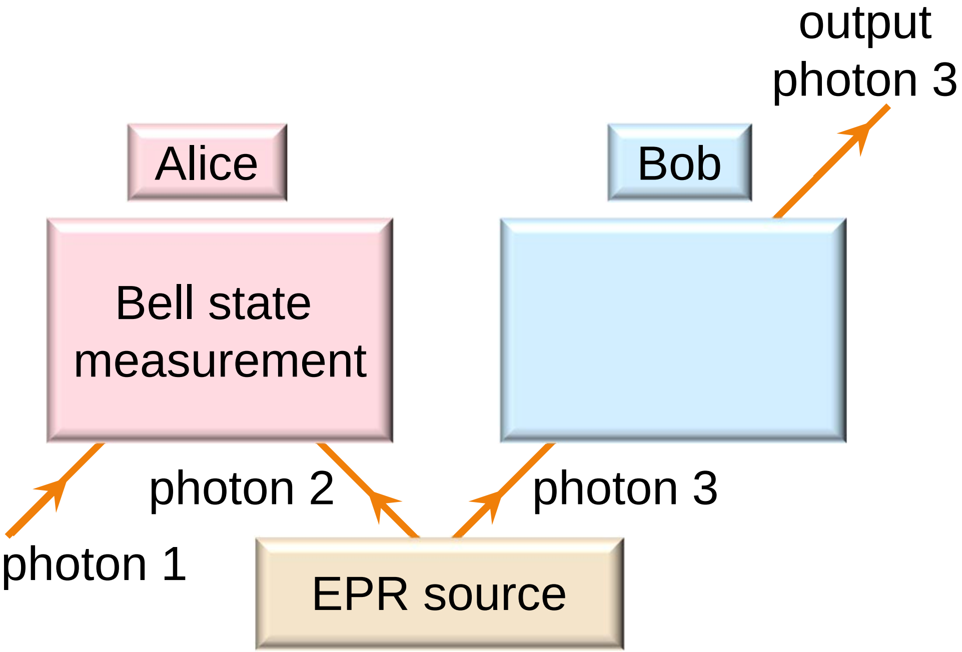
If we now make a measurement

in Alice’s Bell state measurement box

of the Bell state of photons 1 and 2

we collapse the state into

just one of the four Bell state terms



Quantum teleportation

For example, suppose Alice measures $|\Phi^-\rangle_{12}$

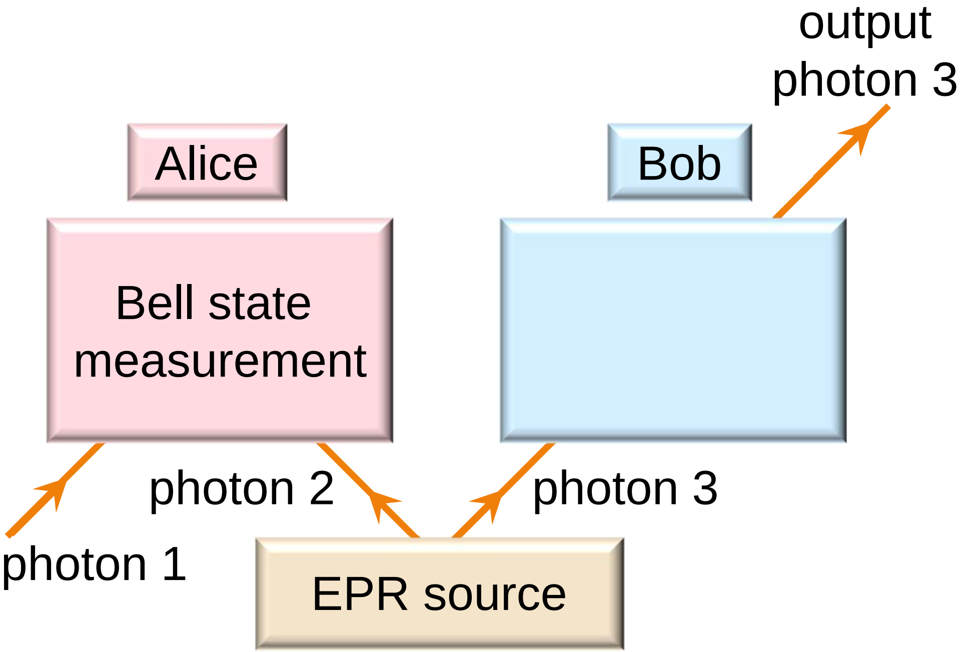
an answer we can know classically

because it is the result of a measurement

then the overall system of three photons

would now be in the state

$$|\Psi\rangle_{123} = \frac{1}{2}|\Phi^-\rangle_{12}\left(c_{\mathrm{H}}|\mathrm{V}\rangle_3 + c_{\mathrm{V}}|\mathrm{H}\rangle_3\right)$$



Quantum teleportation

Because Alice can tell Bob the result of her measurement

by communication over an ordinary classical channel

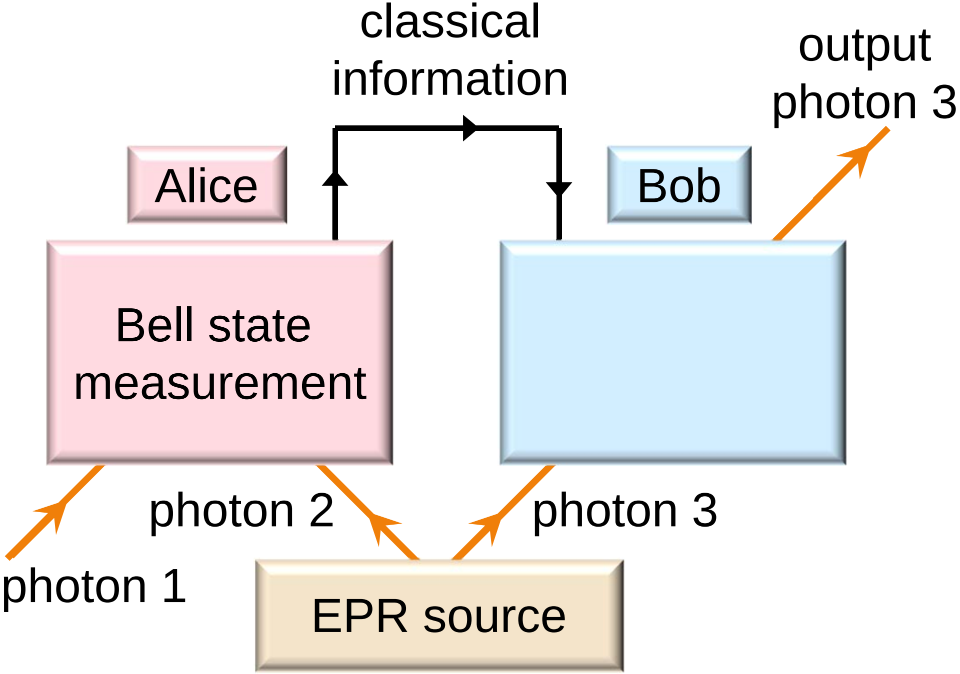
e.g., a telephone line

Bob now knows that

photon 3 is in the state

$$c_H |V\rangle_3 + c_V |H\rangle_3$$

(though he does not know c_H and c_V)



Quantum teleportation

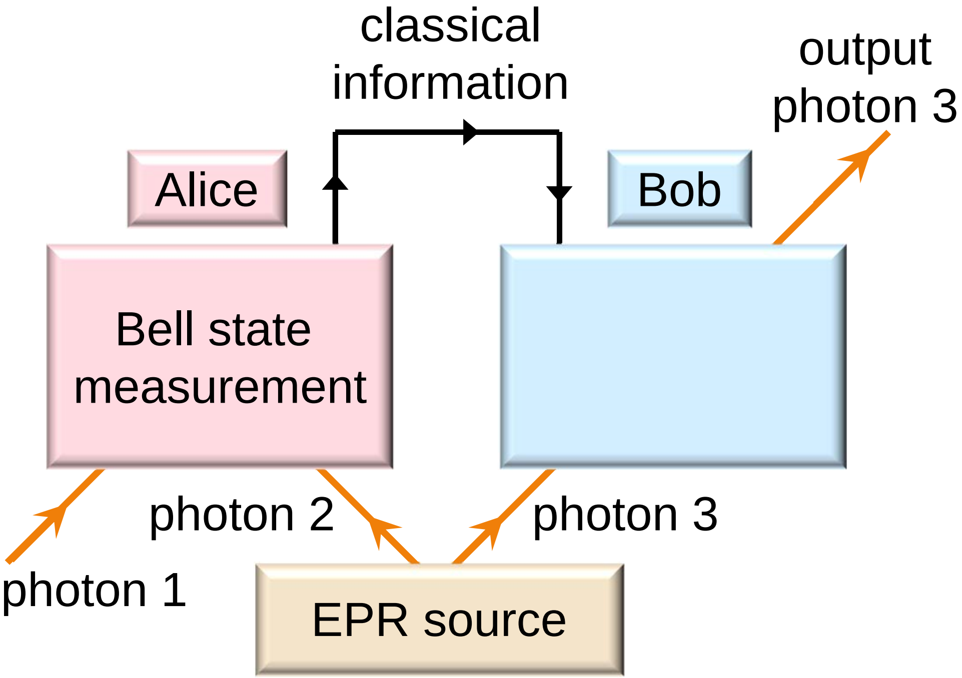
This state $c_{\mathrm{H}}|\mathrm{V}\rangle_3 + c_{\mathrm{V}}|\mathrm{H}\rangle_3$

is not the same as the original state of photon 1

which was by definition

$$c_{\mathrm{H}}|\mathrm{H}\rangle_1 + c_{\mathrm{V}}|\mathrm{V}\rangle_1$$

but that is easily fixed



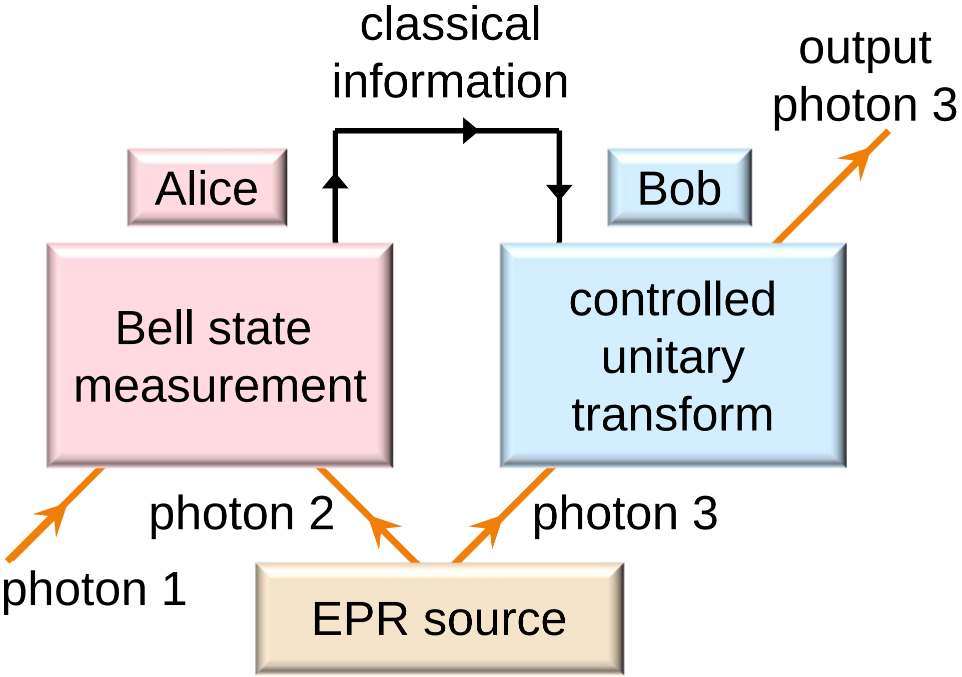
Quantum teleportation

Bob could rotate the polarization 90° clockwise

- turning vertical polarization into horizontal and horizontal into –vertical
- i.e., $|V\rangle \rightarrow |H\rangle$ and $|H\rangle \rightarrow -|V\rangle$

and insert a half wave plate

- to delay the vertical polarization by 180°
  - turning c_V to $-c_V$



Quantum teleportation

By this controlled unitary transformation, Bob has

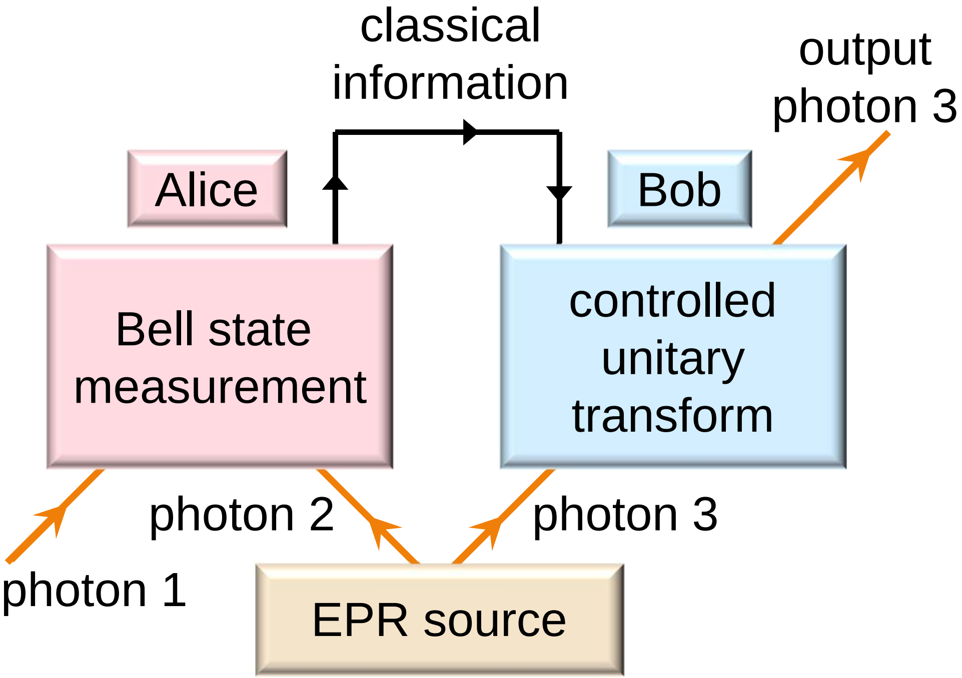
changed $c_{\mathrm{H}}|\mathrm{V}\rangle_3 + c_{\mathrm{V}}|\mathrm{H}\rangle_3$

into $c_{\mathrm{H}}|\mathrm{H}\rangle_3 + c_{\mathrm{V}}|\mathrm{V}\rangle_3$

Photon 3 is now in exactly the same state as photon 1 was

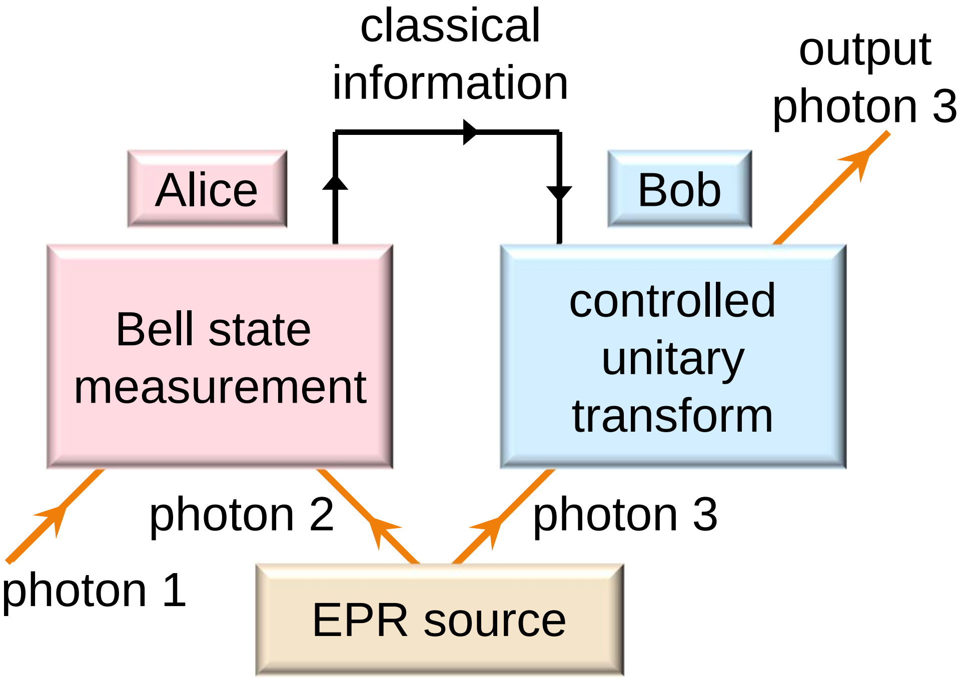
without either Alice or Bob ever knowing what that state was

i.e., without knowing c_{H} and c_{V}



Quantum teleportation

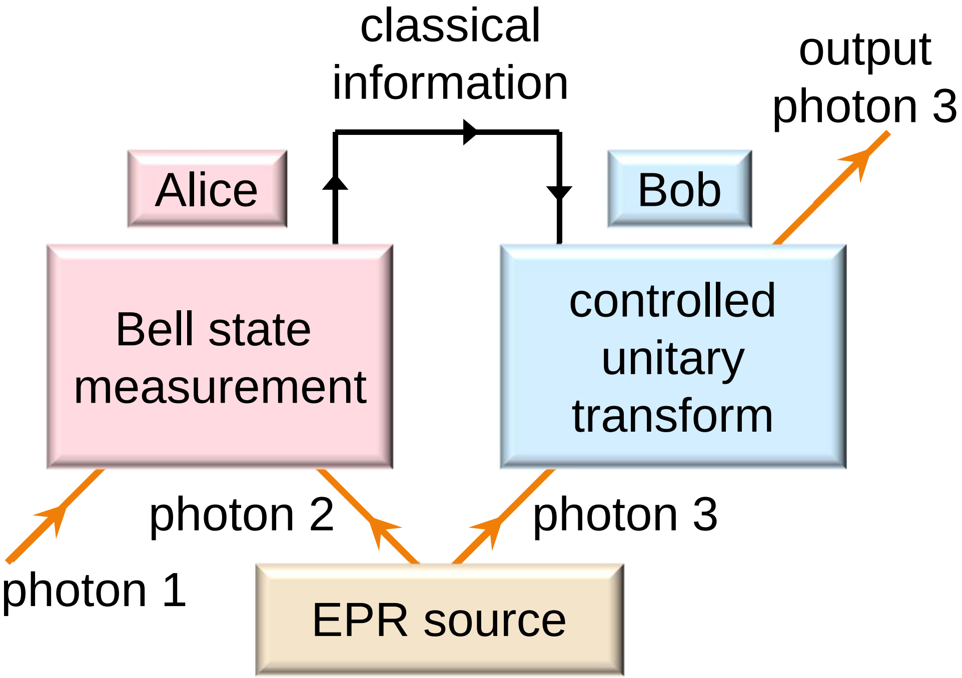
- For other results from Alice’s Bell state measurement
  - Bob implements other polarization manipulations
    - but those present no fundamental problem
      - he could, for example
        - use electrically-controlled phase shifters



Quantum teleportation

In general

- Bob implements a specific unitary transformation on photon 3
  - a combination here of phase delays and polarization rotations
- that depends on the outcome of Alice's Bell state measurement



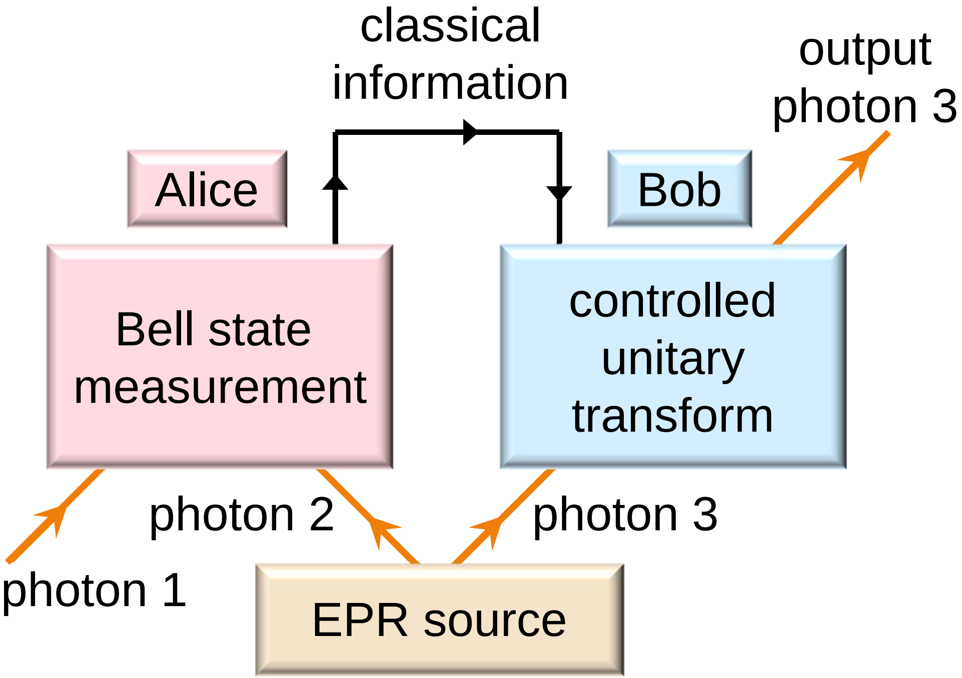
Quantum teleportation

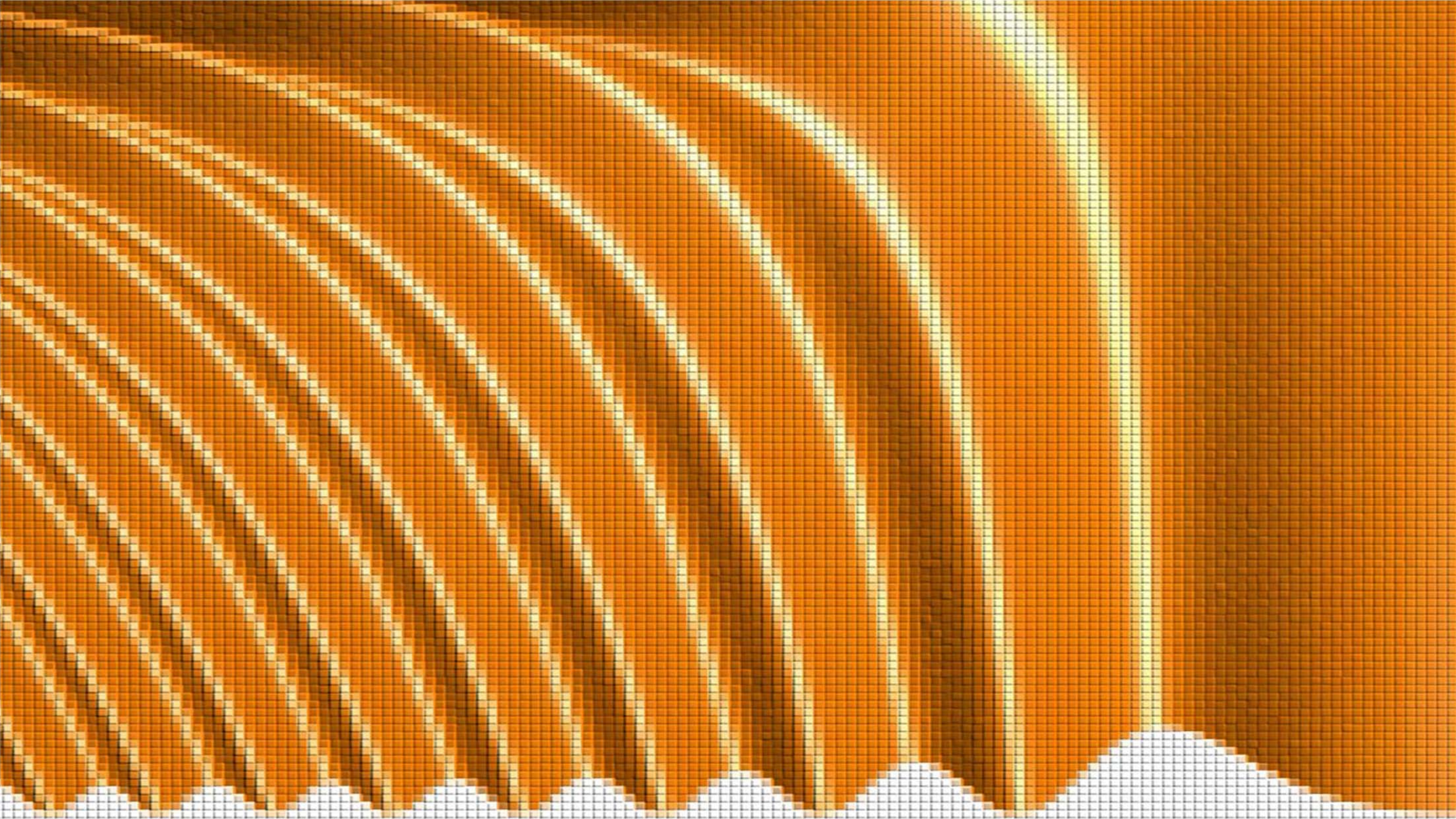
Hence

- for any result from Alice
  - Bob can put photon 3
    - into exactly the same state
    - as photon 1 originally had

thus completing

- the teleportation of the
  - quantum mechanical state



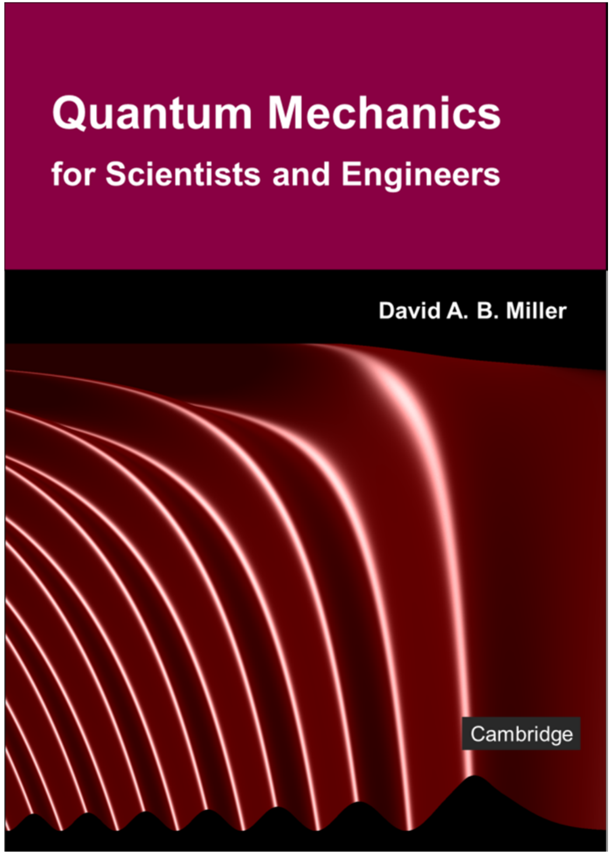


18.3 Hidden variables and Bell’s inequalities

Slides: Video 18.3.1 EPR paradox and hidden variables

Text reference: Quantum Mechanics for Scientists and Engineers

Section 19.1



Hidden variables and Bell's inequalities

EPR paradox and hidden variables

Quantum mechanics for scientists and engineers

David Miller

EPR pairs

We can create two distinguishable particles

an EPR (Einstein-Podolsky-Rosen) pair

in a quantum mechanical superposition state

of the form of one of the Bell states

For example, for two photons 1 and 2

going in different directions

a state like $\left|\Phi^{+}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2} + \left|\mathrm{V}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}\right)$

EPR pairs

Such a state $\left|\Phi^{+}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2} + \left|\mathrm{V}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}\right)$

is a linear superposition of

the state where the two photons

are both horizontally polarized

and the state where the two photons

are both vertically polarized

EPR pairs

In this state $\left|\Phi^{+}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_1\left|\mathrm{H}\right\rangle_2 + \left|\mathrm{V}\right\rangle_1\left|\mathrm{V}\right\rangle_2\right)$

if one measures one of the photons in a state $\left|H\right\rangle$

according to quantum mechanics

the state of both particles is forced to collapse

into the one element $\left|H\right\rangle_1\left|H\right\rangle_2$

in the linear superposition

and a measurement on the other photon

is now bound to give the result $\left|H\right\rangle$ also

EPR pairs

In this state $\left|\Phi^{+}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2} + \left|\mathrm{V}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}\right)$

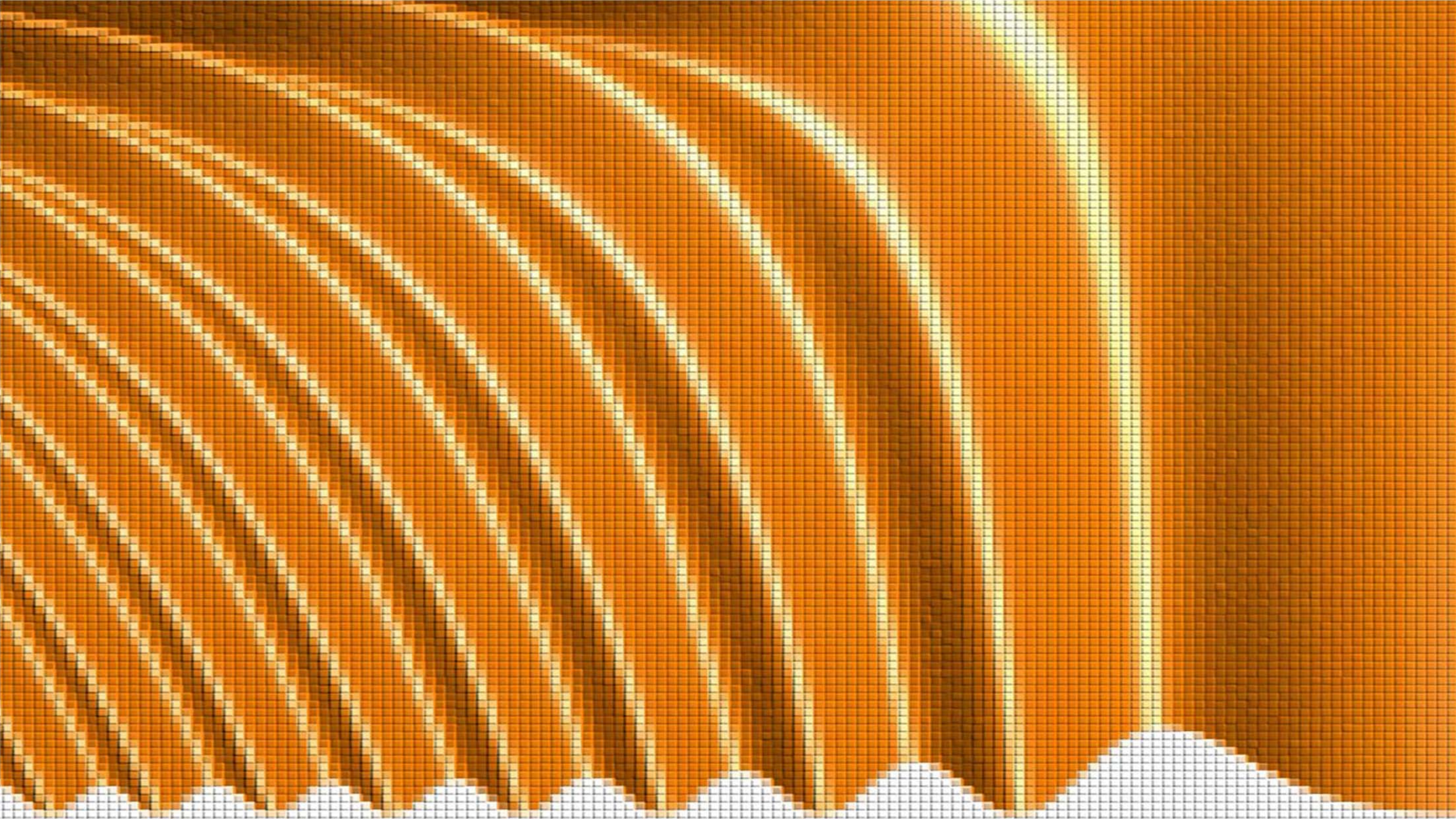
similarly, measuring the result $|V\rangle$ for one photon

will lead

according to quantum mechanics

to the inescapable conclusion that

the other photon will also be in the state $|V\rangle$

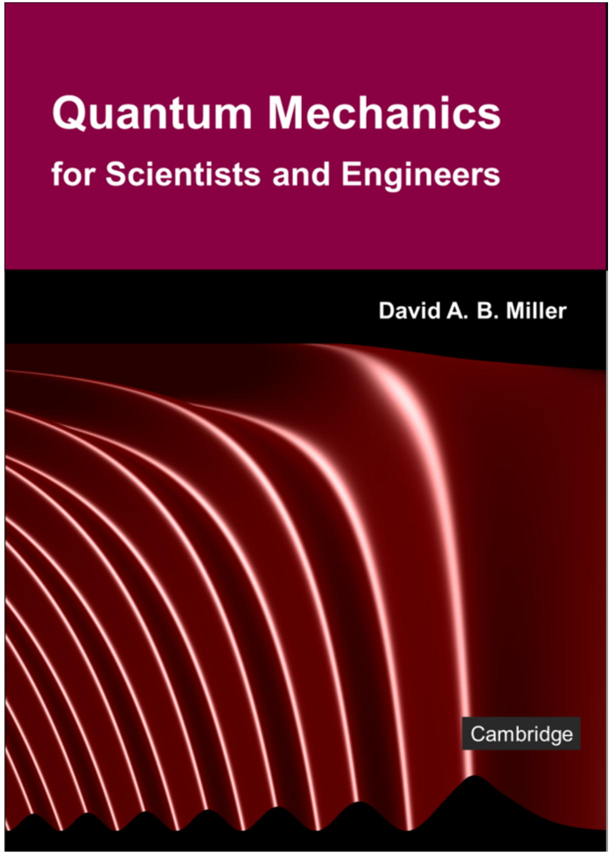


18.3 Hidden variables and Bell's inequalities

Slides: Video 18.3.3 Bell's inequalities

Text reference: Quantum Mechanics for Scientists and Engineers

Section 19.1



Hidden variables and Bell's inequalities

Bell’s inequalities

Quantum mechanics for scientists and engineers

David Miller

EPR experiment

$$\left|\Phi^{+}\right\rangle_{12}=\frac{1}{\sqrt{2}}\left(\left|\mathrm{H}\right\rangle_{1}\left|\mathrm{H}\right\rangle_{2}+\left|\mathrm{V}\right\rangle_{1}\left|\mathrm{V}\right\rangle_{2}\right)$$

An "EPR" pair of photons

in a $\left|\Phi^{+}\right\rangle_{12}$ Bell state

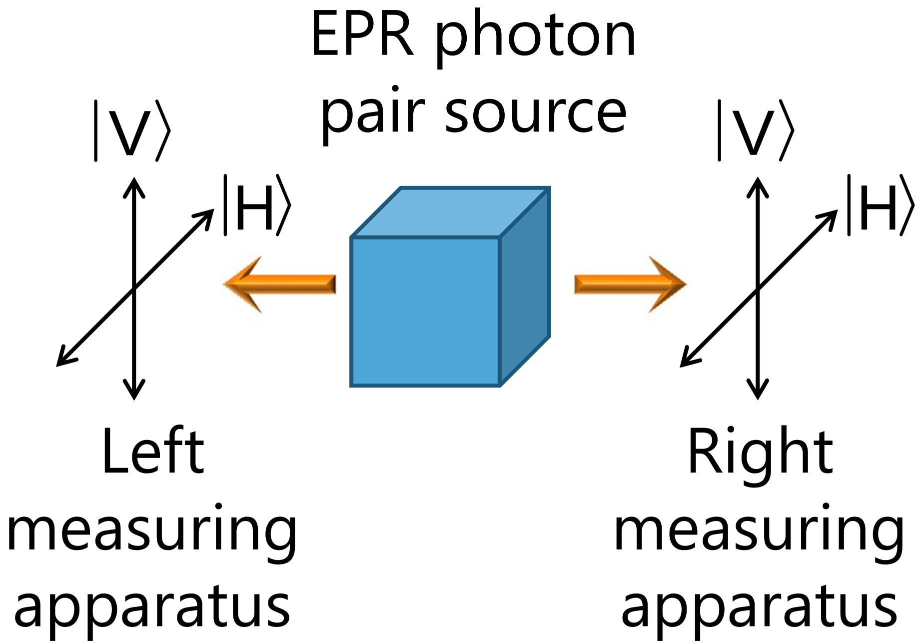
travel to two different measuring apparatuses

with their axes aligned

Quantum mechanics predicts

if we measure one photon to be horizontal

then we will find the other photon is also horizontal



EPR experiment

$$|\Phi^+\rangle_{12} = \frac{1}{\sqrt{2}}\left(|\mathrm{H}\rangle_1 |\mathrm{H}\rangle_2 + |\mathrm{V}\rangle_1 |\mathrm{V}\rangle_2\right)$$

Similarly if we measure one photon to be vertical

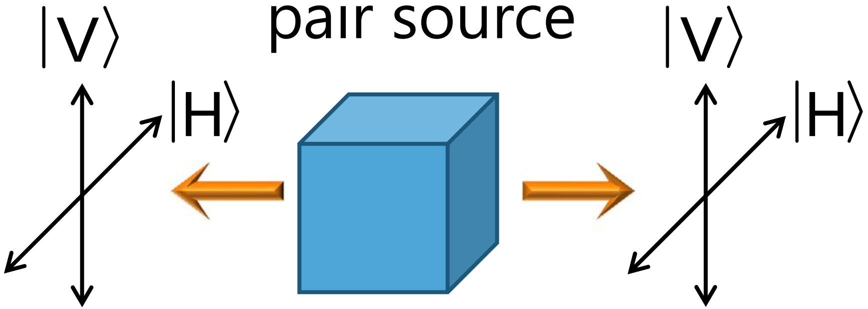
- the other photon will also be measured to be vertical

This is the behavior we find also in experiments

With aligned axes

- both apparatuses measure the same polarization
  - for these Bell-state photons

EPR photon pair source



Left measuring apparatus

Right measuring apparatus

Simple example of a Bell's inequality

Before considering our example

consider the following statement

The number of young women $\leq$

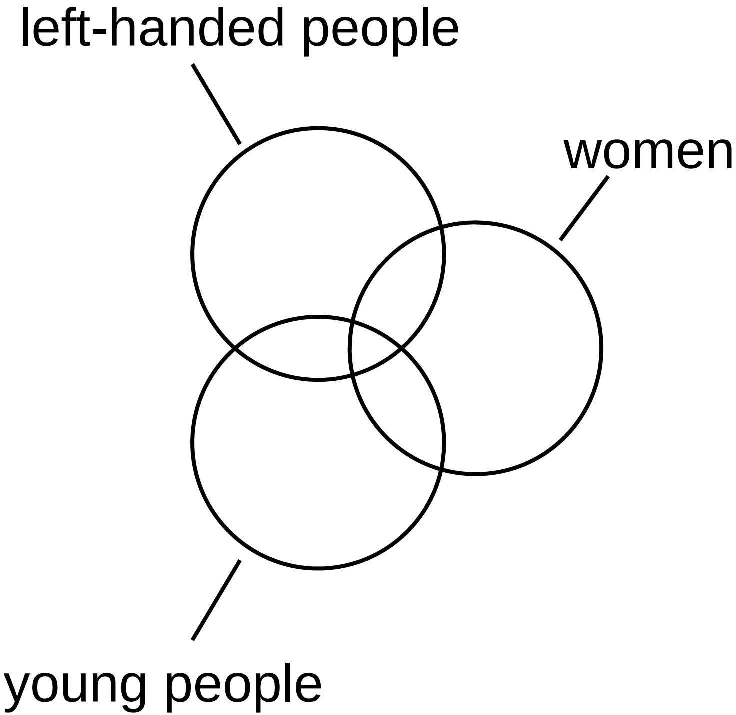
the number of right-handed young people

+ the number of left-handed women

This statement is correct

though it may be easier to see with a Venn diagram

Venn diagram proof

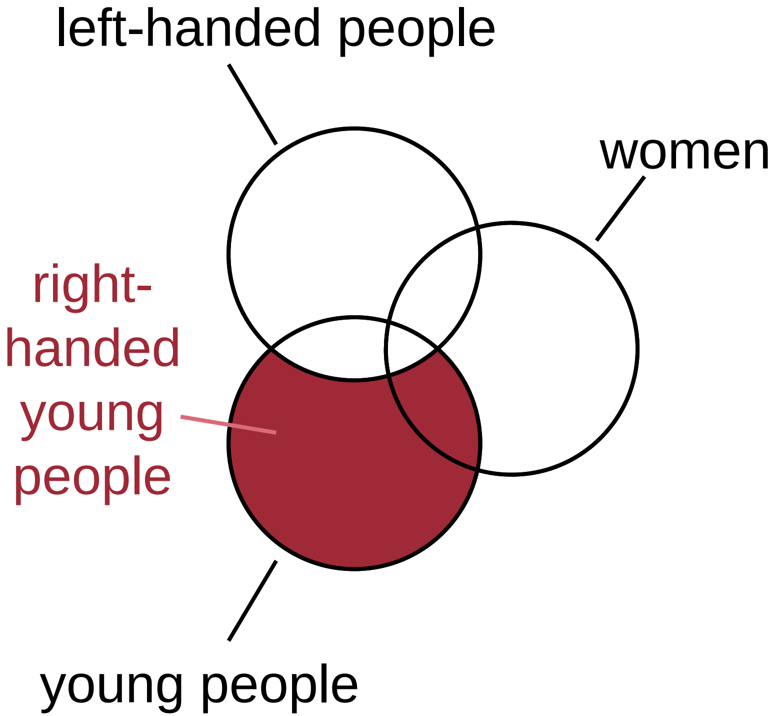


Consider a Venn diagram

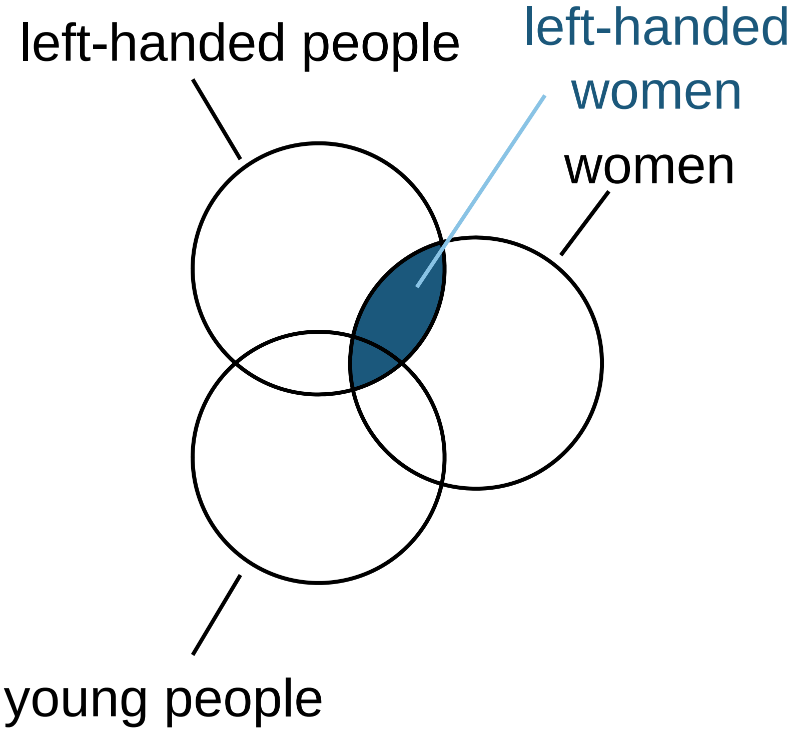
- showing
  - the set of left-handed people
  - the set of women
  - the set of young people

Venn diagram proof

We can show the sets of
right-handed young people



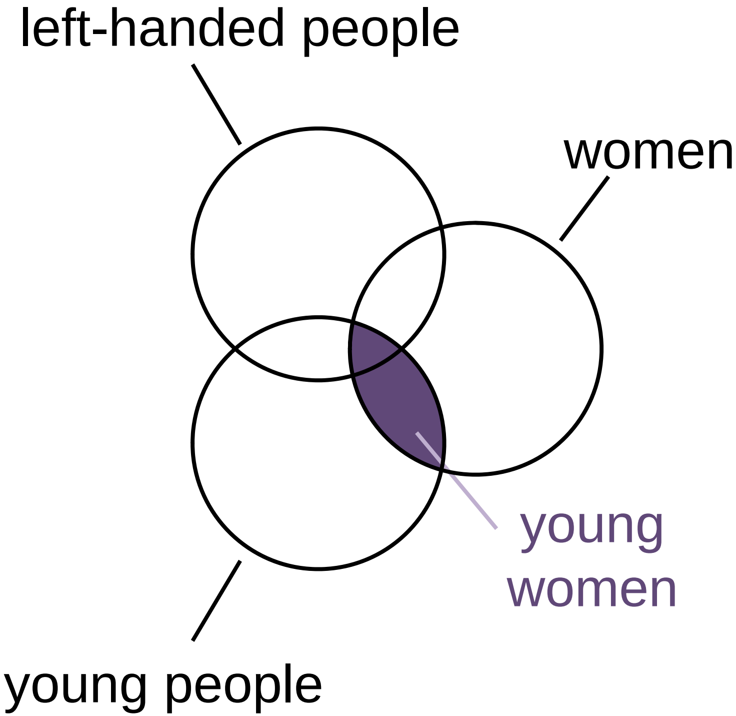
Venn diagram proof



We can show the sets of

- right-handed young people
- left-handed women

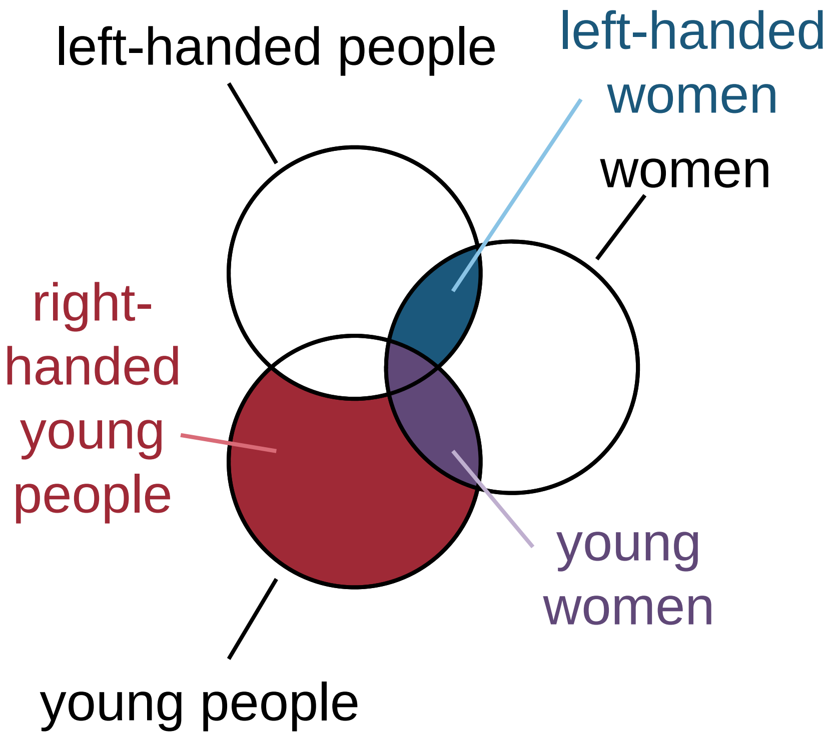
Venn diagram proof



We can show the sets of

- right-handed young people
- left-handed women
- young women

Venn diagram proof



The set of

right-handed young people

plus the set of

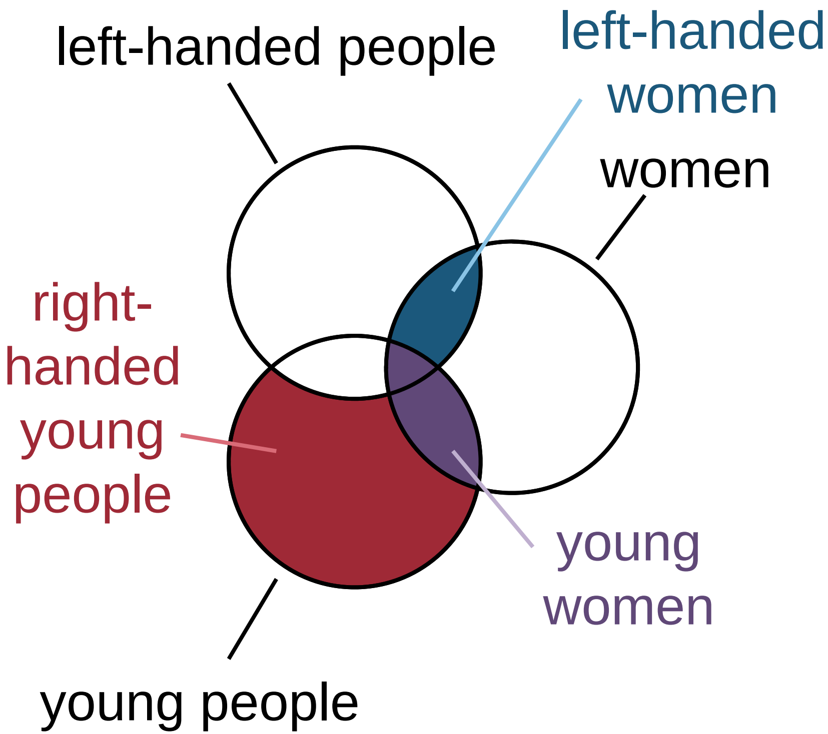
left-handed women

includes all

young women

plus possibly other people too

Venn diagram proof



The set of

right-handed young people

plus the set of

left-handed women

includes all

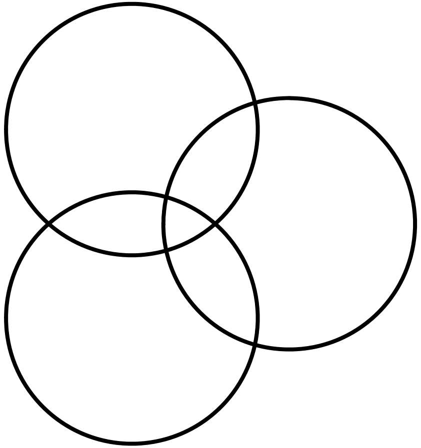
young women

plus possibly other people too

So, # young women ≤

right-handed young people

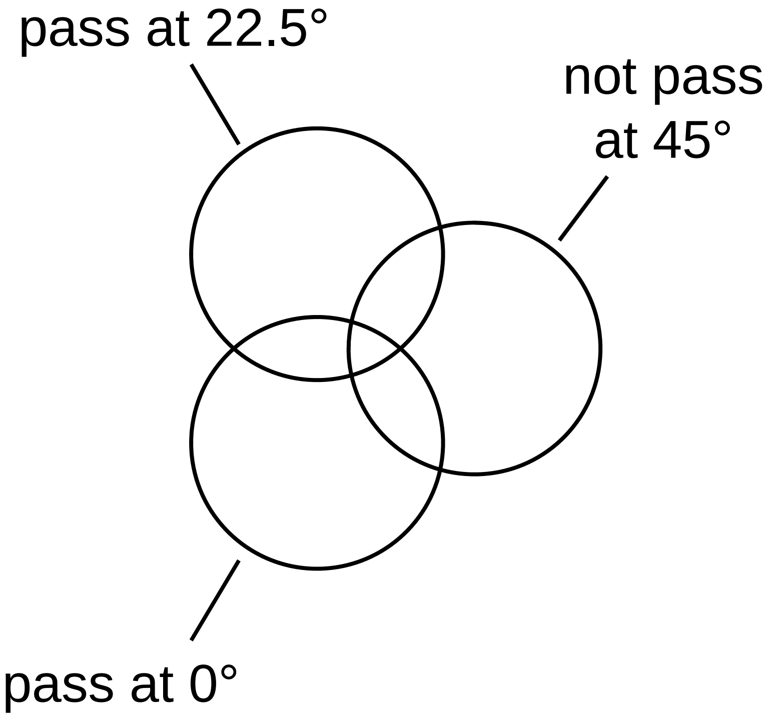
+ # left-handed women



Local variable theory

- So we can draw a Venn diagram
- Each possible value of the local hidden attribute or variable
  - corresponding to a particular measurable set of behaviors
    - with polarizers at any angle
      - is represented by a point on this Venn diagram

Local variable theory



We will be interested in three possible angles for a polarizer

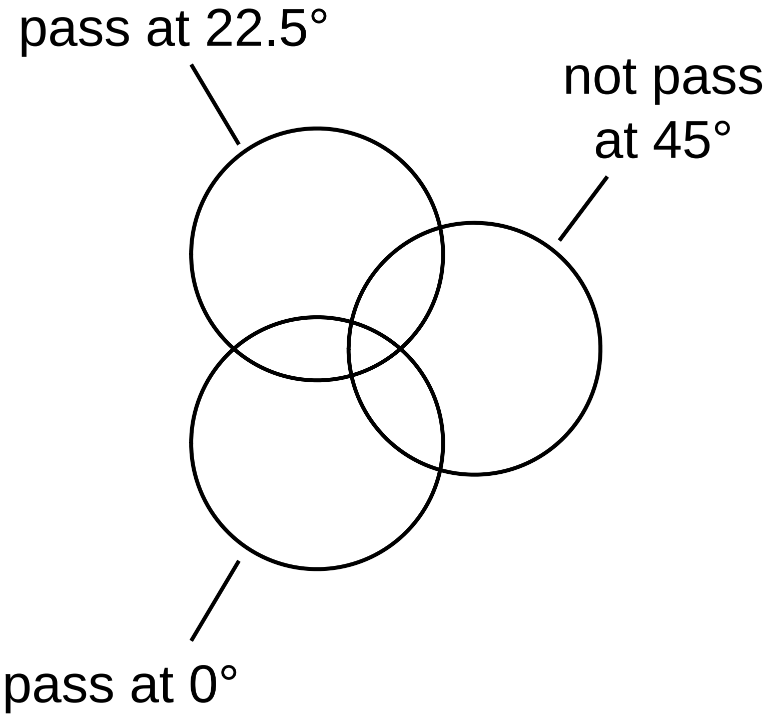
- 0°, 22.5°, or 45°
  - and three specific possible results for any one experiment
    - pass at 22.5°
    - pass at 0°
    - not pass at 45°

Local variable theory

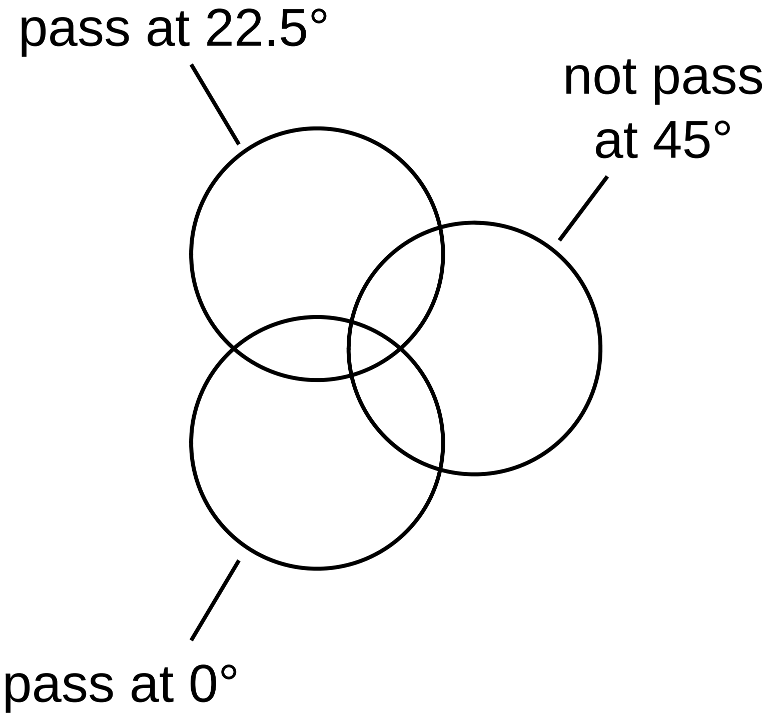
All of these three regions can overlap

and still be in agreement with our observations

on what happens with photons and polarizers

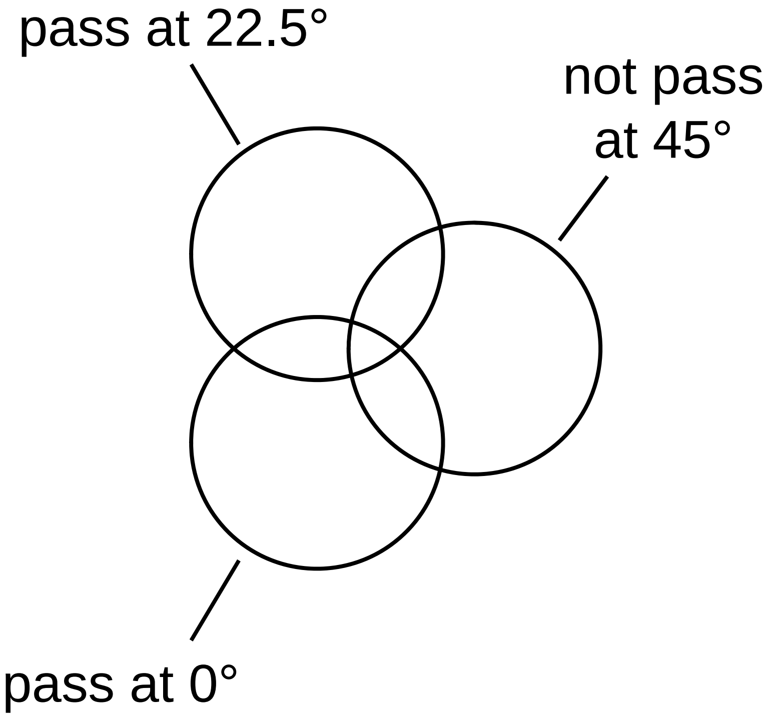


Local variable theory



- We can only perform one test on a given photon
  - with a polarizer set at 0°, 22.5°, or 45°
    - because that test may change the state of the photon in some way

Local variable theory



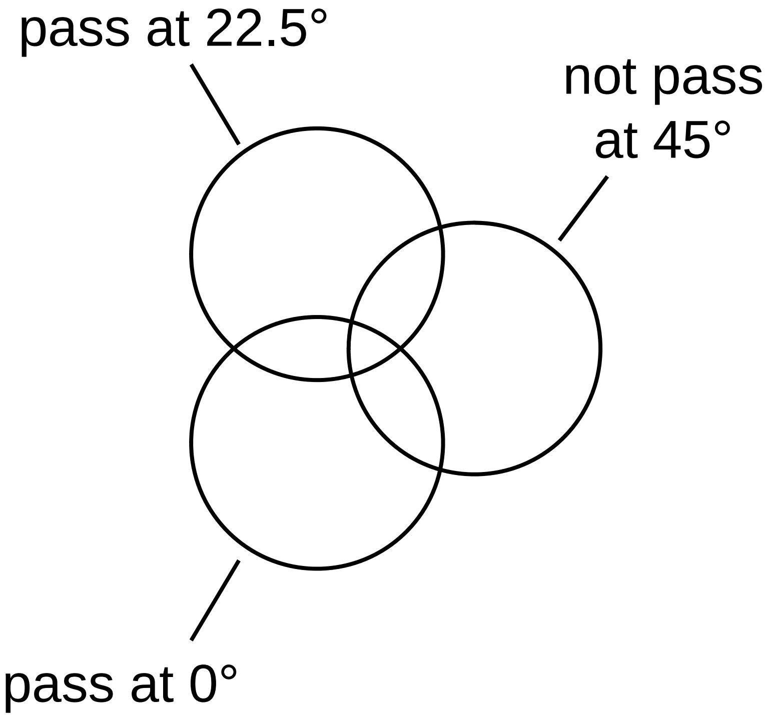
With our EPR photon pair source, we have

- two photons to use in two different experiments
  - one on the left, and one on the right

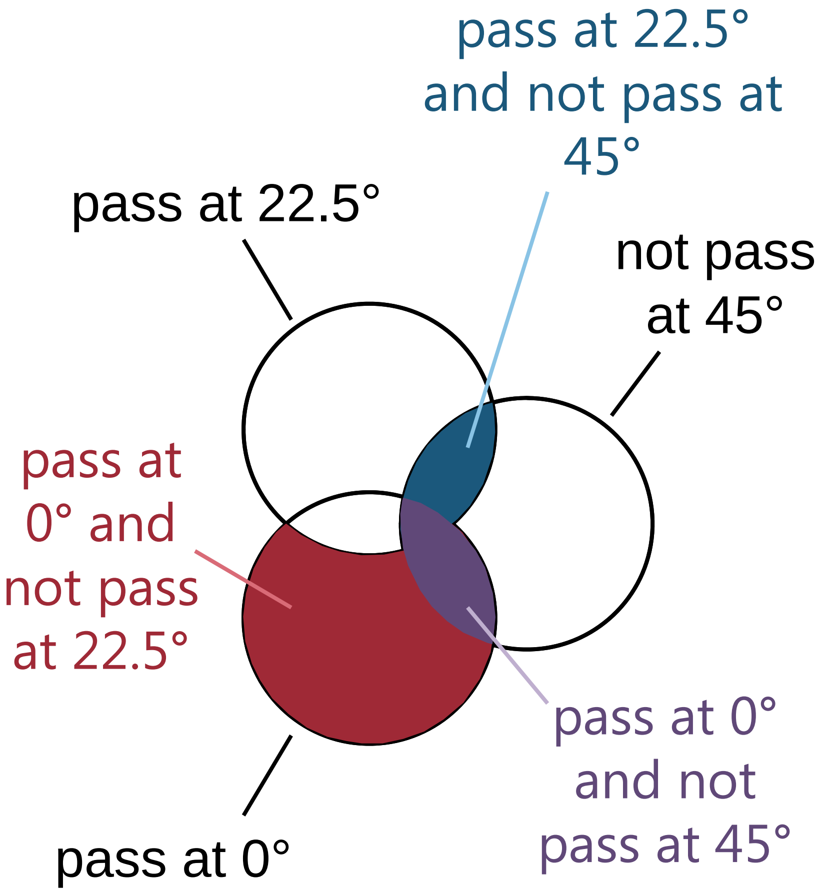
and we already know that photons prepared this way

- always behave identically
  - for identically set polarizers

Local variable theory



- So, both photons must correspond to the same point on the Venn diagram
  - So, we can use one of these photons for one test
    - and the other photon for another test
      - and probe overlap regions



A Bell's inequality

The probability that one photon will pass at 0° while the other will not pass at 45°

≤ the probability that one photon will pass at 0° and the other will not pass at 22.5°

\+ the probability that one photon will pass at 22.5° and the other will not pass at 45°

Bell's inequalities and experiment

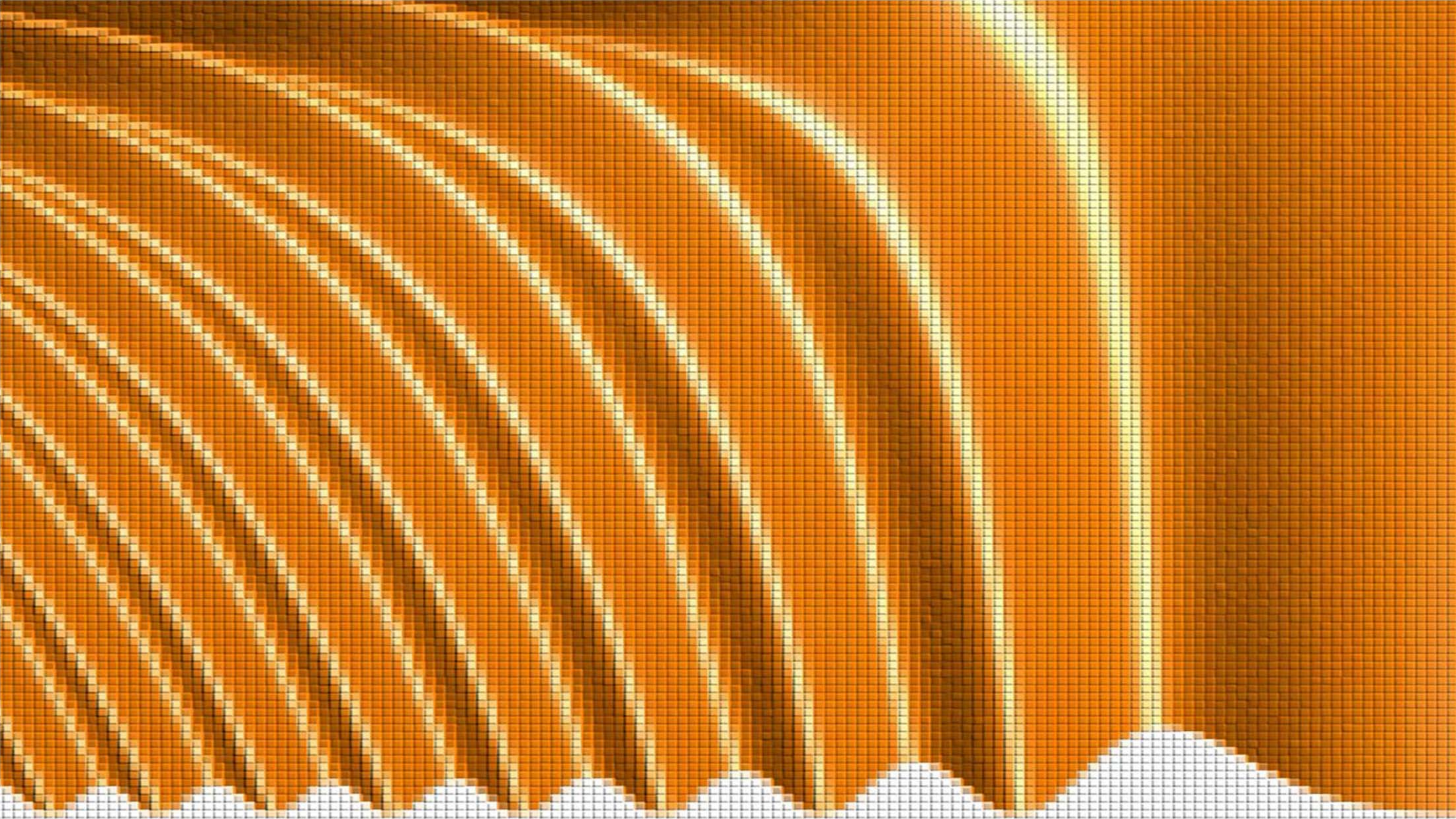
If we find an experiment that violates this inequality

- then we have to throw out deterministic local hidden variable theories
  - e.g., the idea that the photon has a variable that it carries with it that determines the result of the polarization measurement

Experiments do violate this inequality

- Therefore deterministic local hidden variable theories
  - cannot explain reality

This conclusion is independent of the correctness of quantum mechanics

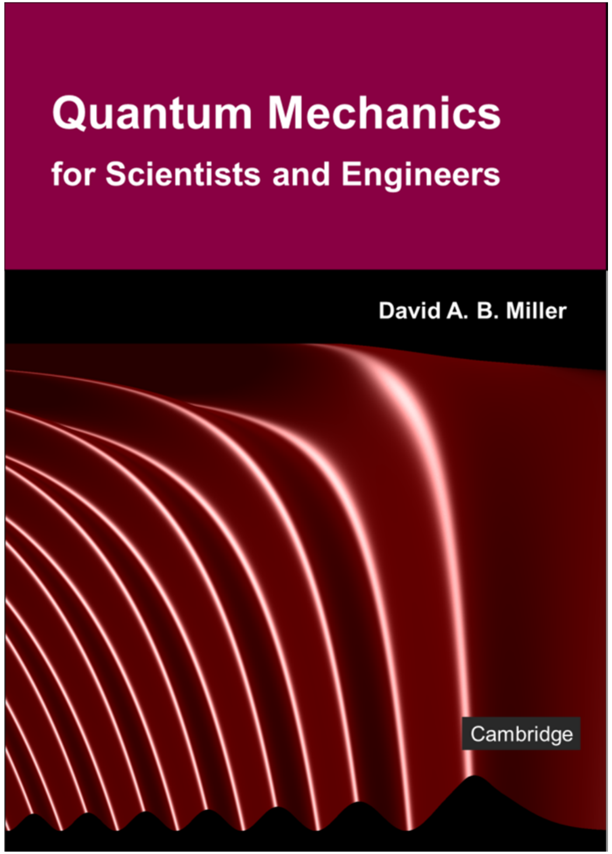


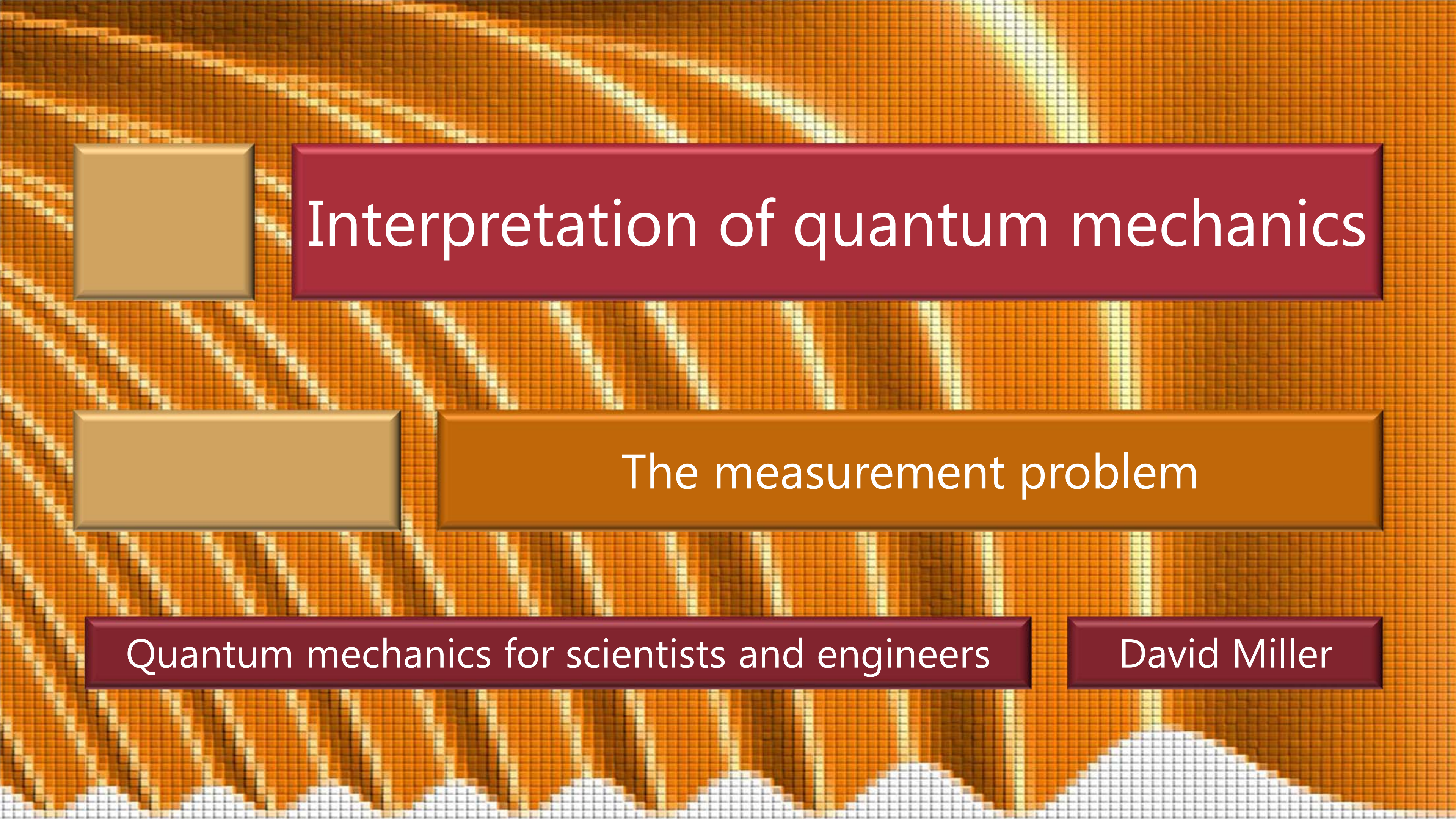
19.1 Interpretation of quantum mechanics

Slides: Video 19.1.1 The measurement problem

Text reference: Quantum Mechanics for Scientists and Engineers

Section 19.2





Interpretation of quantum mechanics

The measurement problem

Quantum mechanics for scientists and engineers

David Miller

Proof of the measurement problem

Suppose that the system starts out

in one of the eigenstates of the quantity

such as electron spin

being measured by the apparatus

Hence, for the initial eigenstate $|\uparrow\rangle$

the result of the measurement process

will be the state

$$\hat{M}|\uparrow\rangle = |\uparrow\rangle$$

When measured in an eigenstate

the system stays in that eigenstate

Proof of the measurement problem

Similarly for the other possible initial eigenstate

$$\hat{M}\left|\downarrow\right\rangle = \left|\downarrow\right\rangle$$

So far, this agrees with our observation

But suppose instead that the system starts

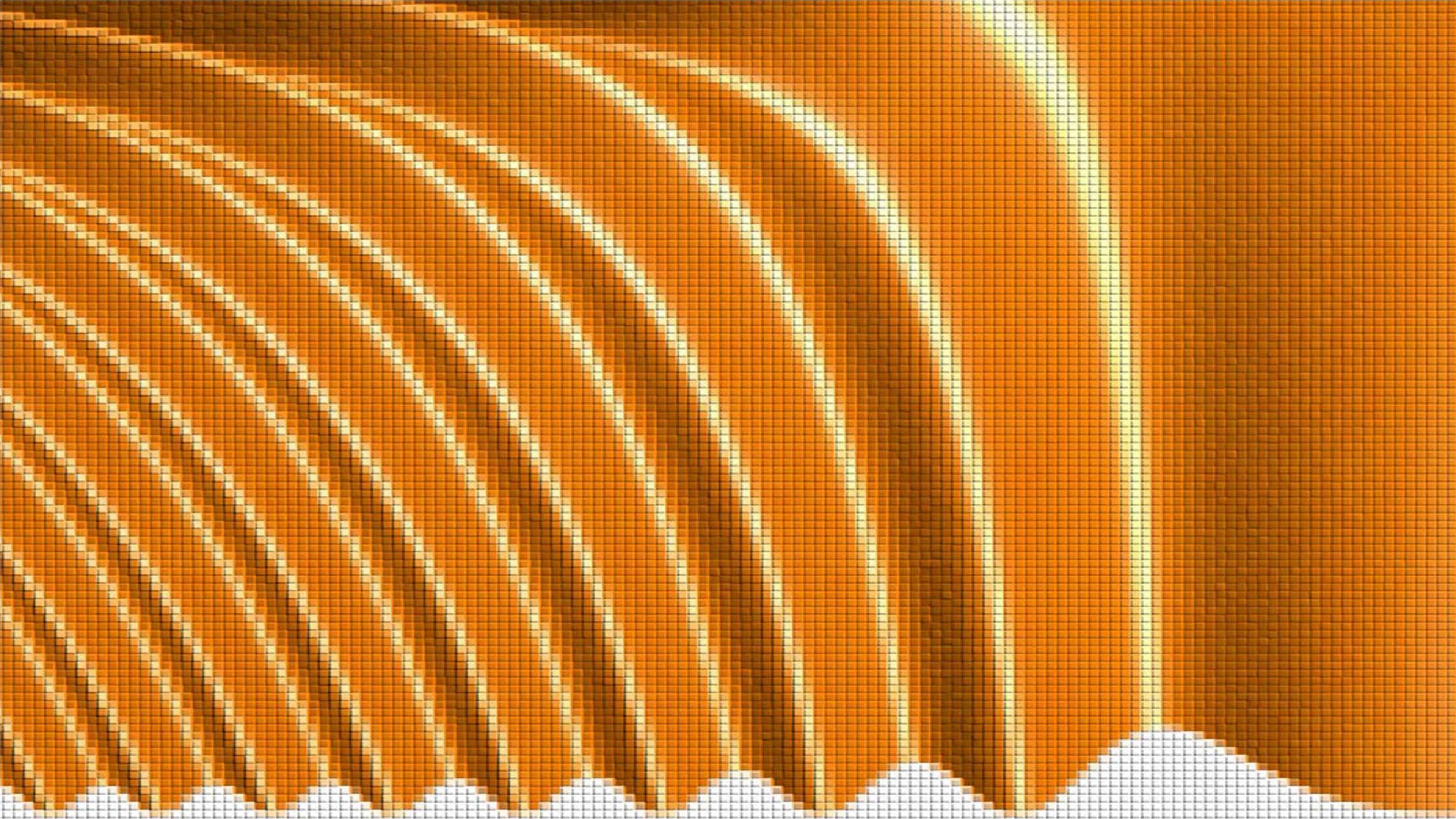
in a linear superposition state

Then on operating on that state

because of the linearity of $\hat{M}$, we have

$$\hat{M}\left(a_{\uparrow}\left|\uparrow\right\rangle + a_{\downarrow}\left|\downarrow\right\rangle\right) = a_{\uparrow}\hat{M}\left|\uparrow\right\rangle + a_{\downarrow}\hat{M}\left|\downarrow\right\rangle = a_{\uparrow}\left|\uparrow\right\rangle + a_{\downarrow}\left|\downarrow\right\rangle$$

Note the resulting state is a linear superposition also

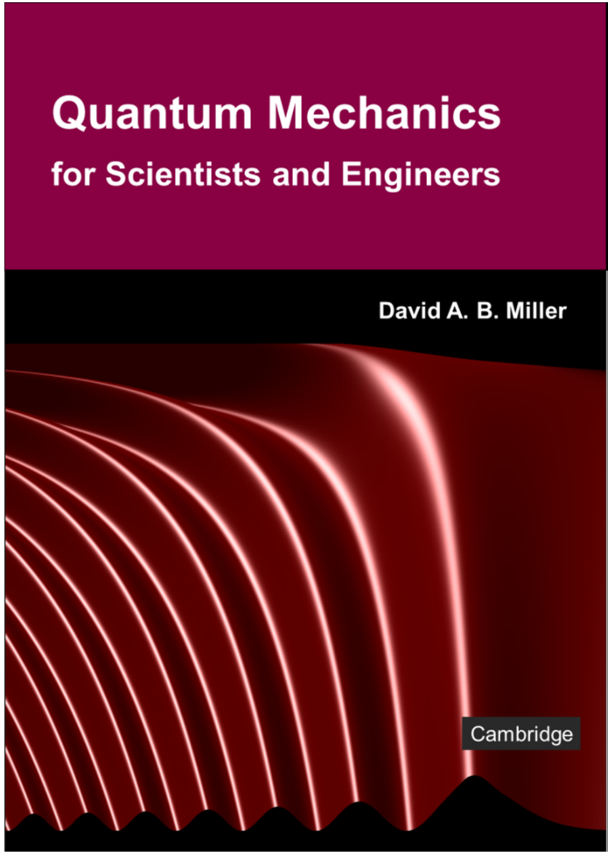


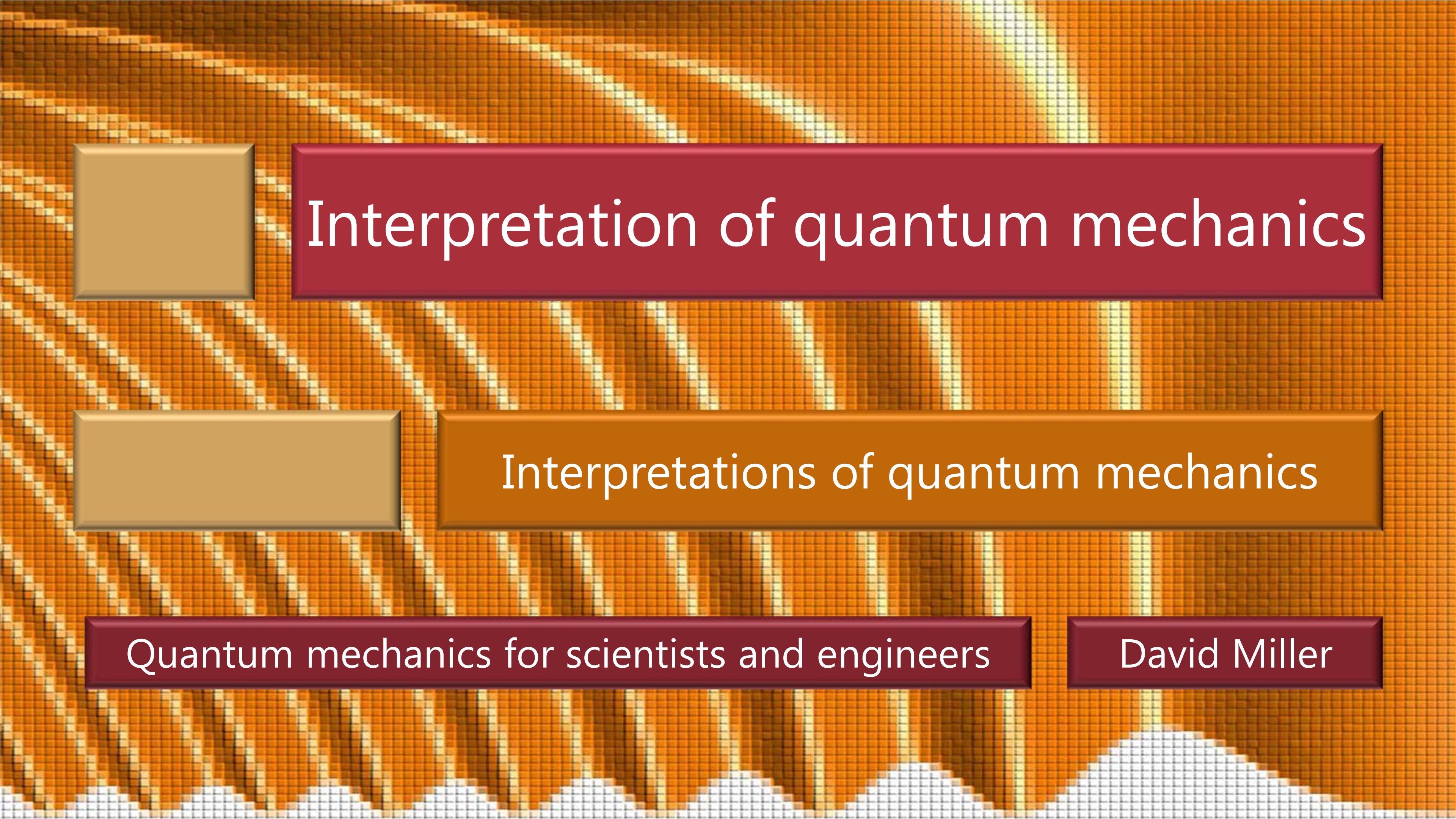
19.1 Interpretation of quantum mechanics

Slides: Video 19.1.3 Interpretations of quantum mechanics

Text reference: Quantum Mechanics for Scientists and Engineers

Section 19.3





Interpretation of quantum mechanics

Interpretations of quantum mechanics

Quantum mechanics for scientists and engineers

David Miller

The Standard Interpretation

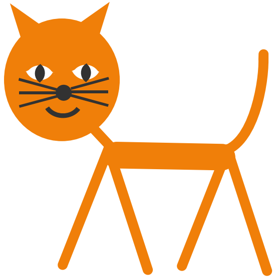
Schrödinger’s cat

The classic illustration of the absurdity of the Standard Interpretation

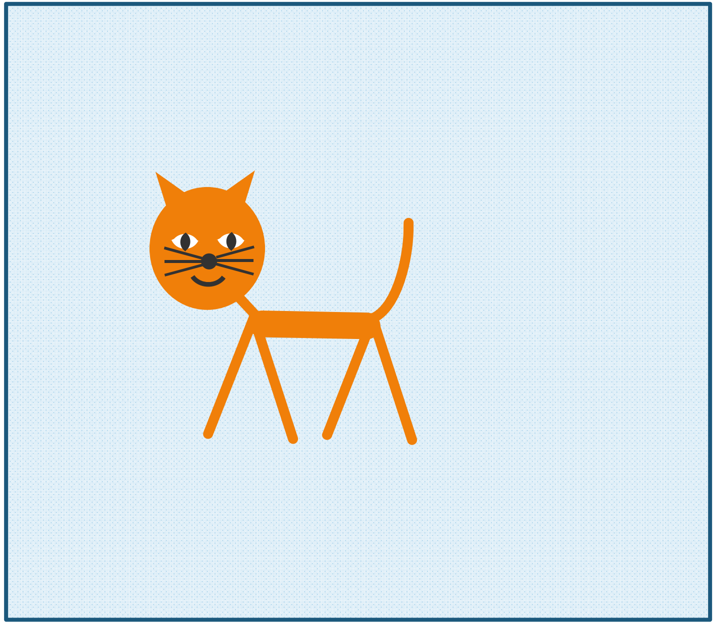
Schrödinger's cat

The classic illustration of the absurdity of the Standard Interpretation

is Schrödinger's cat



Schrödinger's cat



Schrödinger imagines that

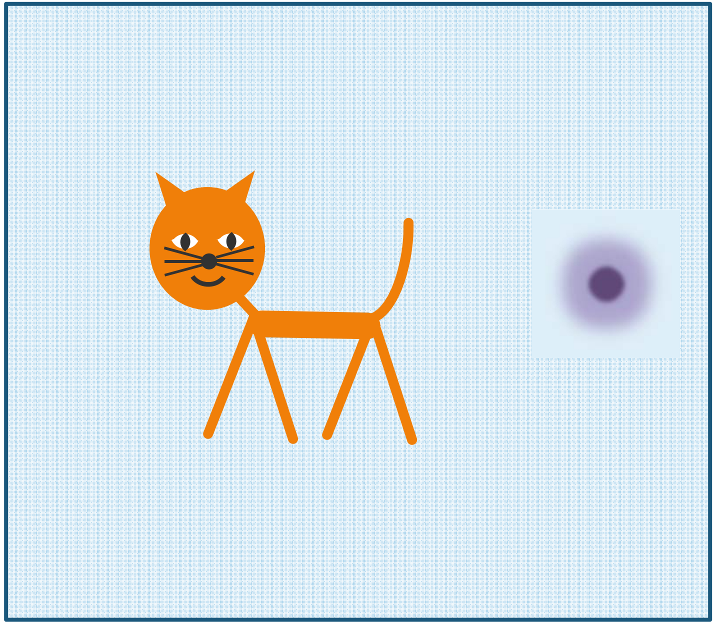
he puts his cat in a box

Schrödinger's cat

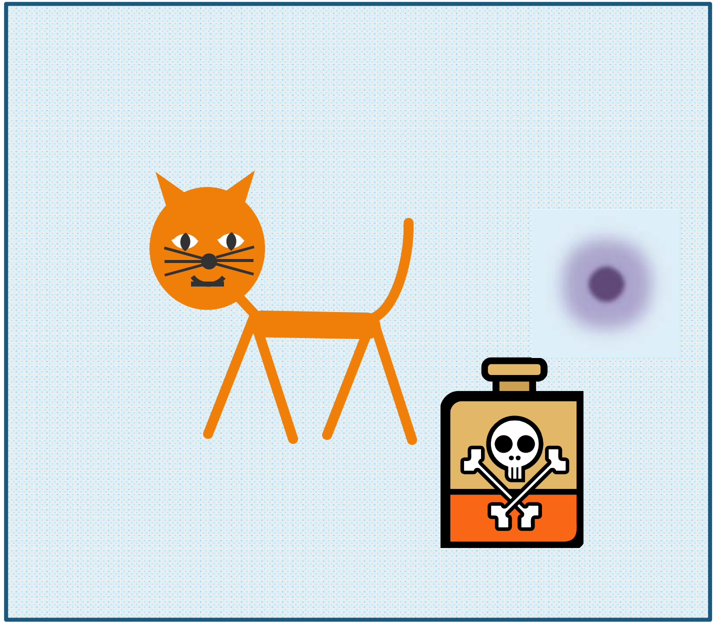
Schrödinger imagines that

he puts his cat in a box

He also puts an atom in the box



Schrödinger’s cat



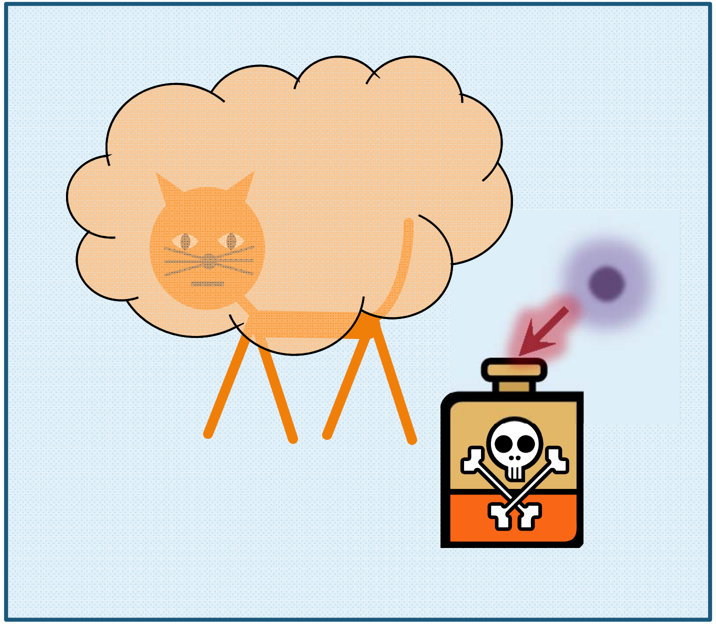
Schrödinger imagines that

he puts his cat in a box

He also puts an atom in the box

and a bottle of poison

Schrödinger’s cat



The atom can decay radioactively

a “random” quantum mechanical process

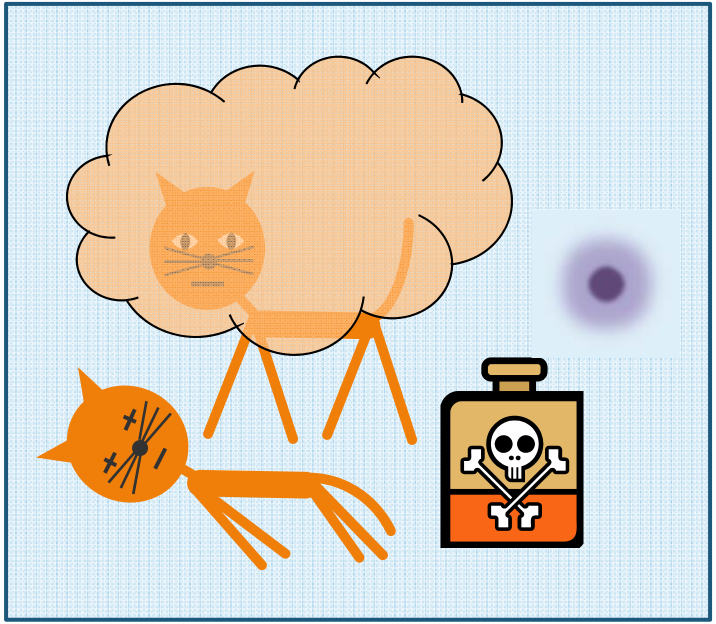
which may or may not occur

If it occurs

it triggers the release

of poison gas from the bottle

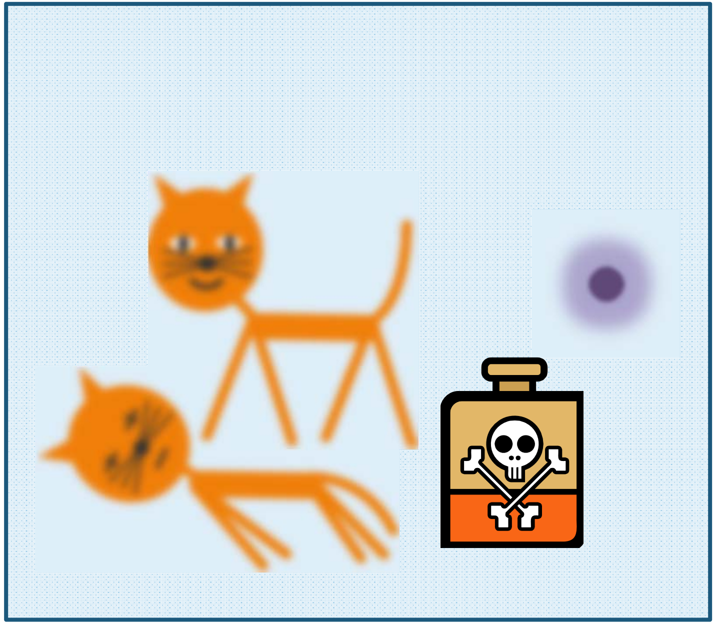
Schrödinger’s cat



The atom can decay radioactively

- a “random” quantum mechanical process
  - which may or may not occur
- If it occurs
  - it triggers the release
    - of poison gas from the bottle
      - killing the cat

Schrödinger’s cat



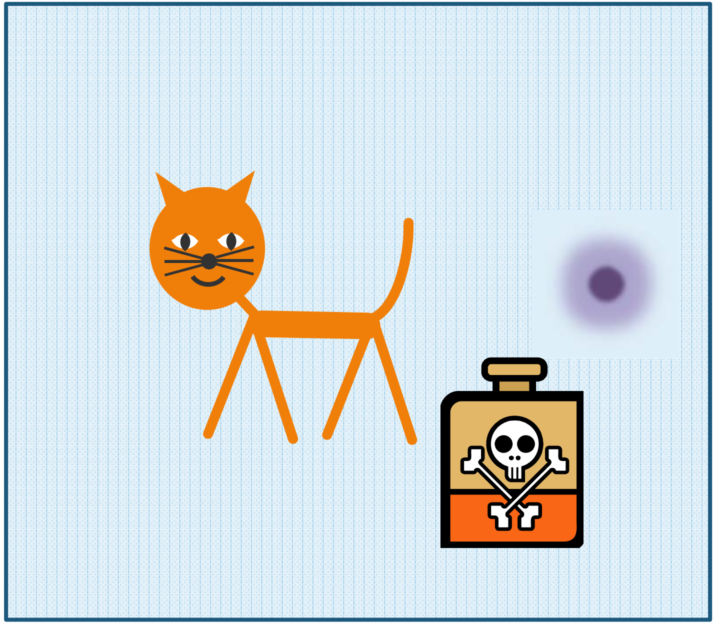
But if we cannot see inside the box

- what is the state of the cat?
  - Is it alive or dead?

If a quantum mechanical state only collapses when it is observed

- “observer created reality”
  - then the cat should be in a linear superposition of alive and dead

Schrödinger's cat



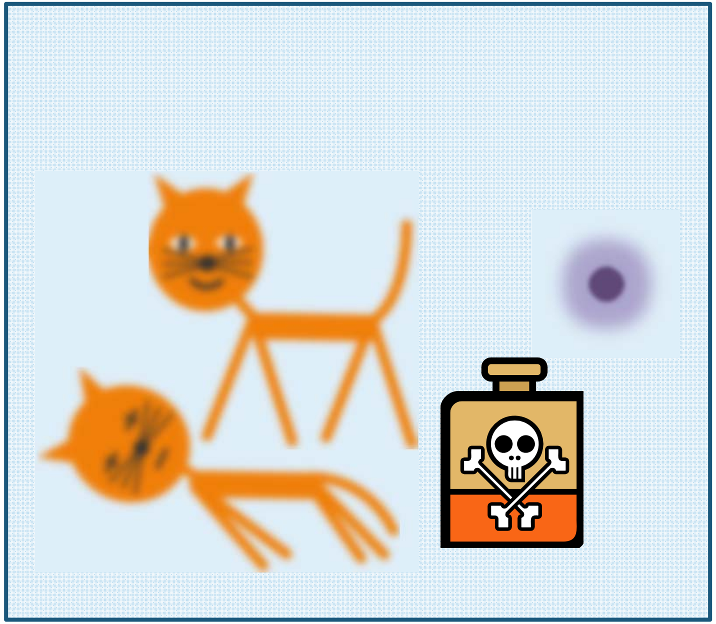
But if we cannot see inside the box

- what is the state of the cat?
  - Is it alive or dead?

If a quantum mechanical state only collapses when it is observed

- "observer created reality"
  - then the cat should be in a linear superposition of alive and dead
  - until we open the box and observe one state or the other

Schrödinger's cat



Schrödinger argues

 superposition is absurd for a cat

Whether it is absurd or not

 and whether absurdity

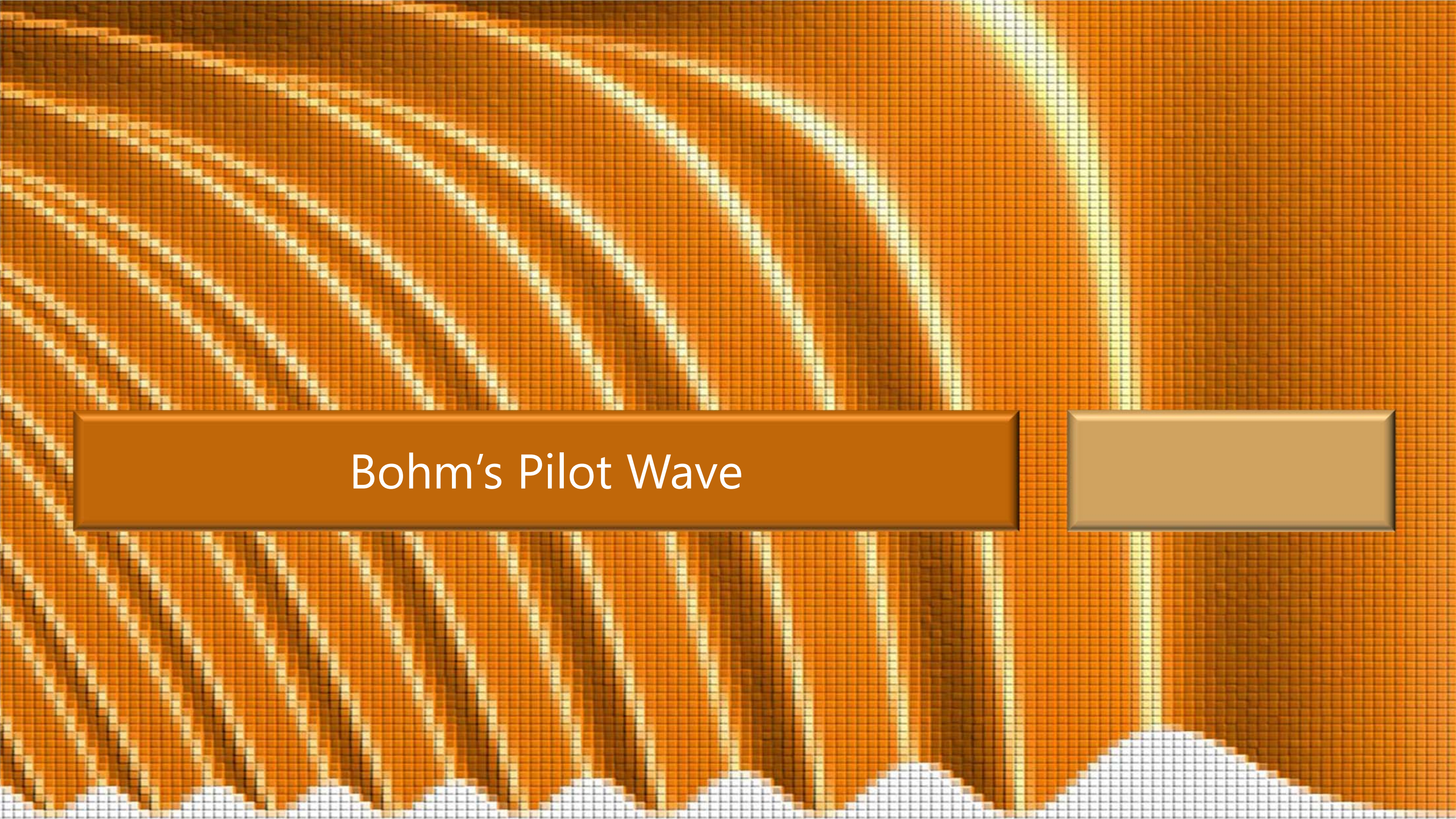
  should be given any weight

   in deciding the validity of a quantum theory

    provided the theory agrees with experiment

     are both matters of opinion

The Copenhagen Interpretation



Bohm’s Pilot Wave

Bohm's Pilot Wave

We start with the time-dependent Schrödinger equation

$$\frac{-\hbar^2}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial\psi}{\partial t}$$

and then we make a mathematical choice to write

$$\psi(\mathbf{r},t) = R(\mathbf{r},t)\exp\left(iS(\mathbf{r},t)\right)$$

where R and S are real quantities

Any complex function can be represented in this way

Bohm's Pilot Wave

If we substitute $\psi(\mathbf{r},t) = R(\mathbf{r},t)\exp\left(iS(\mathbf{r},t)\right)$

into $\frac{-\hbar^2}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial\psi}{\partial t}$

then, after some algebra

we can deduce the equation

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V + Q = 0$$

where $Q = -\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}$

Hamilton-Jacobi equation

The equation $\frac{\partial S}{\partial t}+\frac{\left(\nabla S\right)^2}{2m}+V+Q=0$

but without the "quantum potential" $Q=-\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}$, i.e.,

$$\frac{\partial S}{\partial t}+\frac{\left(\nabla S\right)^2}{2m}+V=0$$

is known as the Hamilton-Jacobi equation

of classical mechanics

It reproduces the usual classical behavior of a particle

S is then the "action" or Hamilton's principal function

and the momentum is $\mathbf{p}=\nabla S$

Hamilton-Jacobi equation

This equation $\frac{\partial S}{\partial t}+\frac{(\nabla S)^2}{2m}+V=0$

is a completely deterministic equation

in which position and momentum are both simultaneously well defined

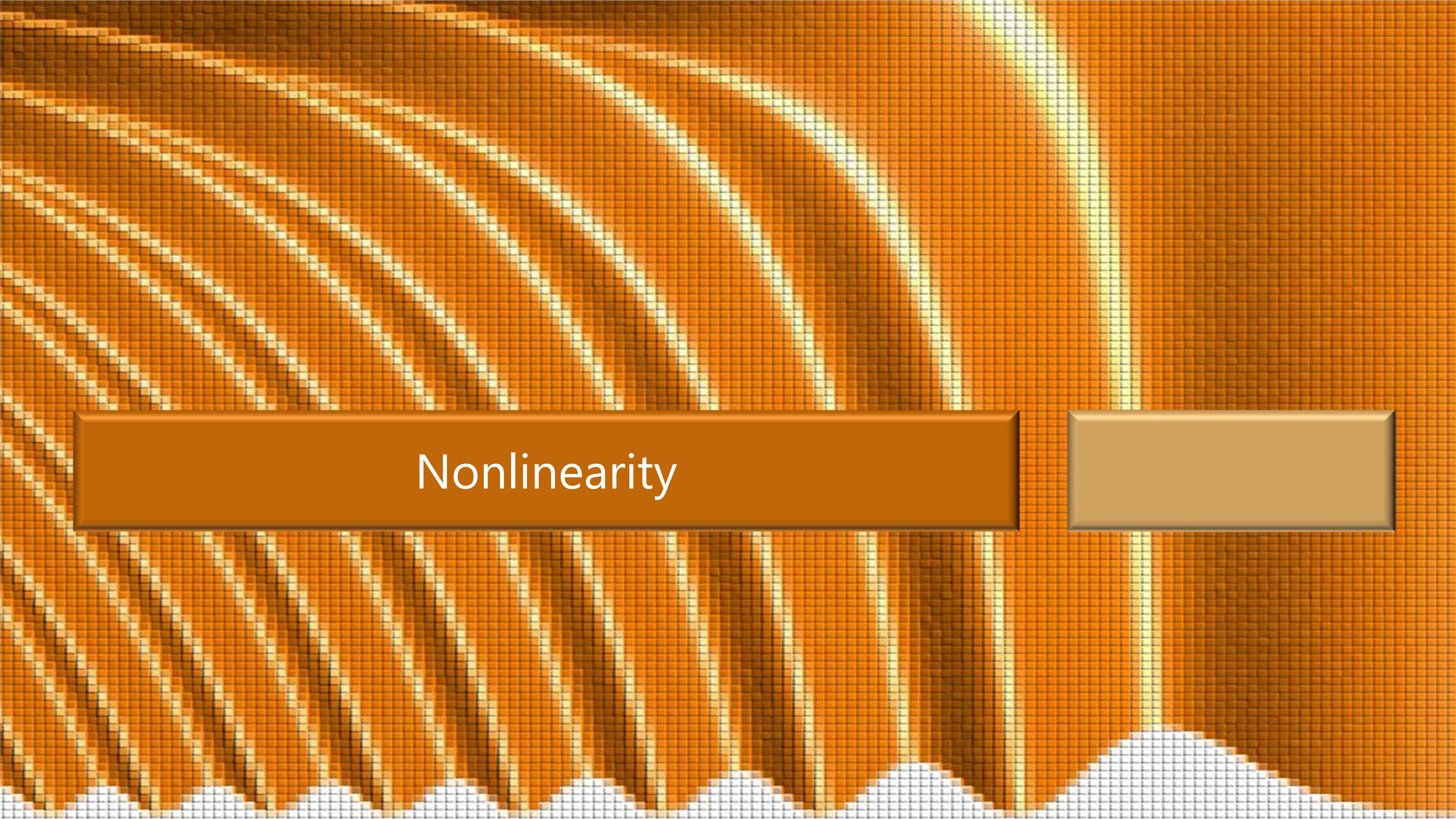
For a large wavepacket and a large mass

then the quantum potential Q is a very small correction

and hence, even using the full form with $Q=-\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}$

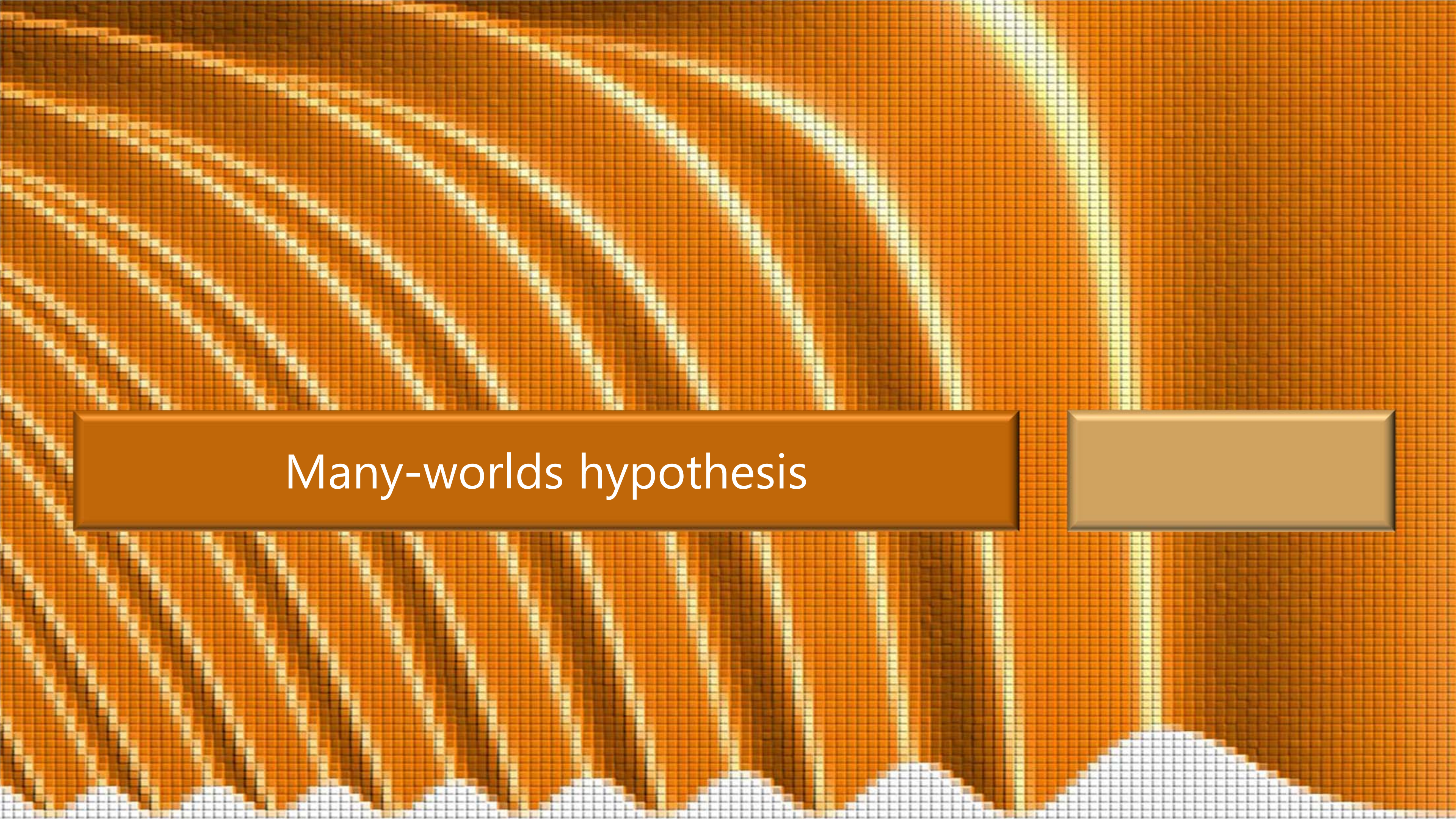
for such effectively classical situations

we obtain the familiar classical behavior



Nonlinearity

Distinction between matter and mind



Many-worlds hypothesis

Epilogue