Quantum Mechanics Intro

for MSE421: Statistical Mechanics

Kevin Kazuki Huguenin-Dumittan

EPFL, Switzerland

April 3, 2023

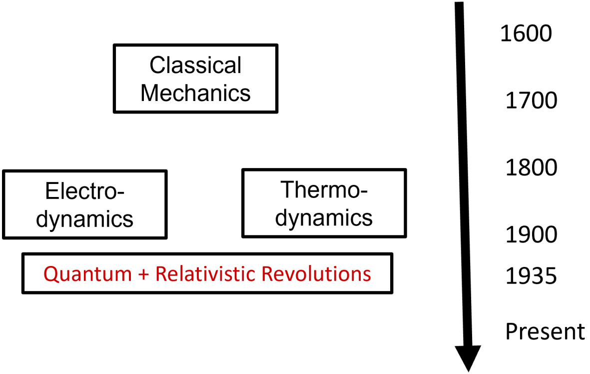
Goal for today: What do these weird expressions mean?

1. Position representation $\psi(x_1, x_2, \ldots, x_n) = \langle x_1, x_2, \ldots, x_n | \psi \rangle$
2. Matrix elements of operator $A_{\nu,\nu'} = \langle \nu | A | \nu' \rangle$
3. Trace of operator $\mathrm{Tr} A = \sum_\nu \langle \nu | A | \nu \rangle$
4. Exponential of operator e^A.

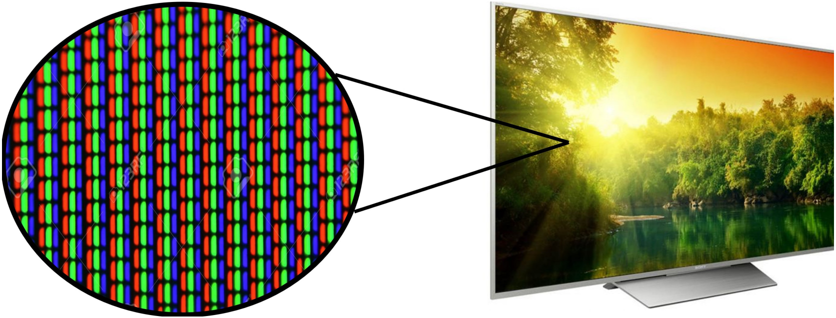
Outline:

1. The flow of ideas: From Planck to Schrödinger
2. The mathematics of quantum mechanics
3. Notation + techniques used in lecture

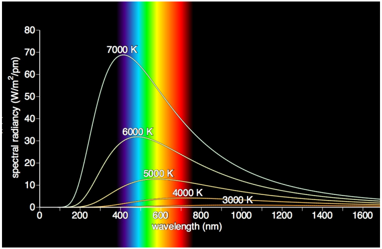
Why was quantum mechanics discovered 100 years ago?

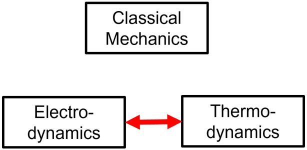


Old quantum theory (1900 – 1925) vs Classical mechanics



Max Planck (1900):
Black body spectrum and Ultraviolet Catastrophe

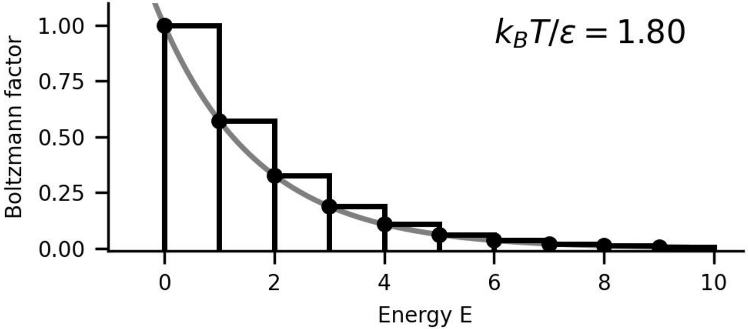




Replacing Sums by Integrals: Basel Problem

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \approx 1.64 \quad (1)$$

$$\approx \int_1^{\infty} \frac{1}{x^2} \mathrm{d}x = 1 \quad (2)$$



A. Einstein's interpretation of Planck's calculation: Continuous version

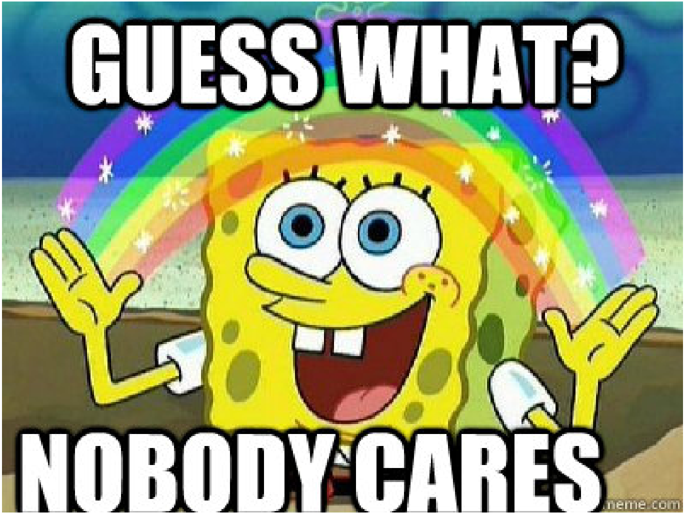
Der Mittelwert der Energie des Massenteilchens ist also:

$$(4) \qquad \bar{E} = \frac{\int E e^{-\frac{N}{RT}E} \, dE}{\int e^{-\frac{N}{RT}E} \, dE} = \frac{RT}{N}.$$

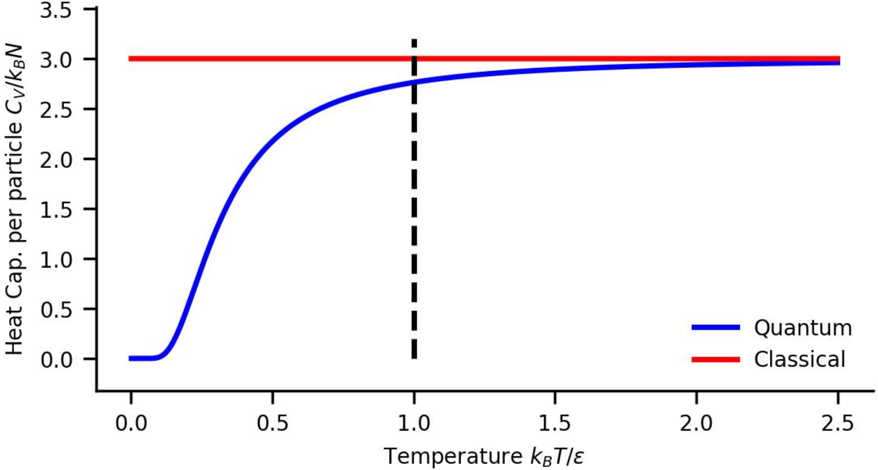
Discrete Sum:

$$\bar{E} = \frac{\int E e^{-\frac{N}{RT}E} \, \omega(E) \, dE}{\int e^{-\frac{N}{RT}E} \, \omega(E) \, dE} = \frac{0 + A \varepsilon e^{-\frac{N}{RT}\varepsilon} + A \, . \, 2\varepsilon e^{-\frac{N}{RT}2\varepsilon} \dots}{A + A \; e^{-\frac{N}{RT}\varepsilon} + A e^{-\frac{N}{RT}2\varepsilon} + \dots}$$

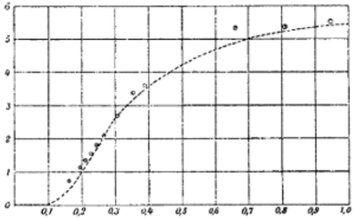
$$= \frac{\varepsilon}{e^{\frac{N}{RT}\varepsilon} - 1}.$$

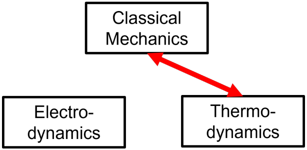

GUESS WHAT?
NOBODY CARES

A. Einstein (1907): Einstein Model of Solids

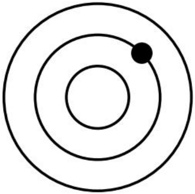


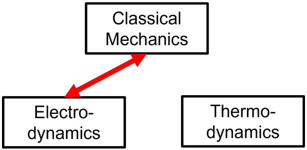
A. Einstein (1907): Einstein Model of Solids



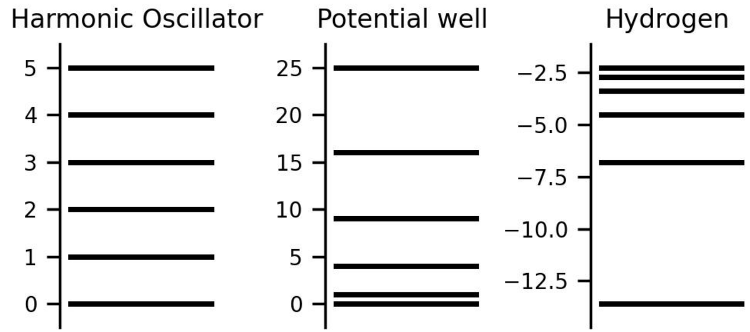


N. Bohr: (1913): Bohr model of atom

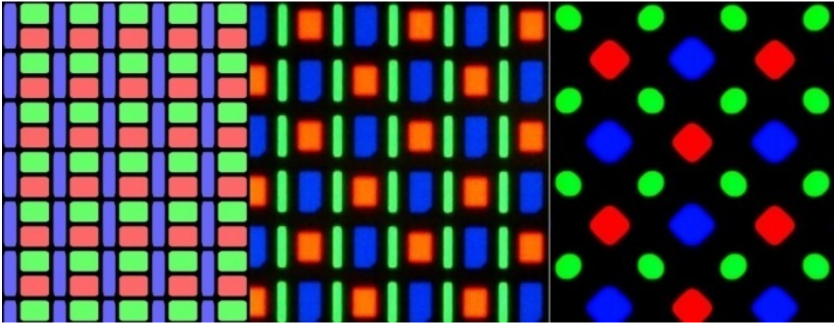




Different systems have different “pixellation” (quantization):



Guess the pixellation rule!



Sommerfeld (1916): Bohr-Sommerfeld Quantization!

$$\oint p \mathrm{d}q = nh \tag{3}$$

. . . with Maslov correction

$$\oint p \mathrm{d}q = \left(n + \frac{1}{2}\right) h \tag{4}$$

But what is the deeper reason behind this quantization?

W. Heisenberg (1925): Matrix mechanics Application: Hamiltonian (diagonalized)

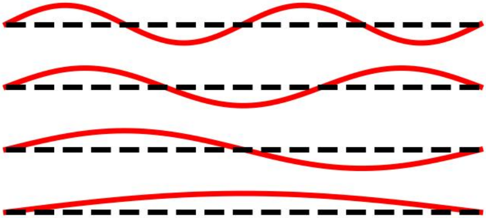
$$\mathcal{H} = \begin{pmatrix} E_1 & & & \\ & E_2 & & \\ & & E_3 & \\ & & & \ddots \end{pmatrix} \tag{5}$$

$$\hat{x} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \\ \vdots & & \ddots & \ddots \end{pmatrix} \quad \hat{p} = -i\hbar \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \\ \vdots & \vdots & & \ddots & \ddots \end{pmatrix} \tag{6}$$

Uncertainty principle: $\hat{x}\hat{p} \neq \hat{p}\hat{x} \Rightarrow \Delta x \Delta p = \frac{\hbar}{2}$.

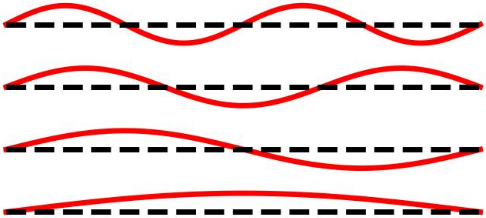
E. Schrödinger (1926): Wave mechanics
Standing wave of strings:

$$\frac{\partial^2 f}{\partial x^2} = -k^2 f \tag{7}$$



E. Schrödinger (1926): Wave mechanics
Standing wave of strings:

$$\frac{\partial^2 f}{\partial x^2} = -k^2 f \tag{8}$$



E. Schrödinger (1926): Wave mechanics
Boundary conditions for string:

$$\sin(kL) = 0 \Leftrightarrow kL = n\pi \Leftrightarrow k = \frac{\pi}{L}n \tag{9}$$

Schrödinger equation:

$$\frac{\partial^2\psi}{\partial x^2} + V\psi = E\psi \tag{10}$$

Mathematics for Quantum Mechanics: Functional Analysis

Functional Analysis $\approx$ infinite dimensional linear algebra

Concept	$\mathbb{R}^n$	Functions
Vector	v	f
Basis	$v = \sum_{j=1}^{n} \alpha_j b_j$	$f = \sum_{j=1}^{\infty} \alpha_j b_j$
Linear Map	Matrix Av	Operator Af
Eigenvalues	$Av = \lambda v$	$Af = \lambda f$
Inner products	$\langle v, w \rangle$	$\langle f, g \rangle$
Norm (length)	$\|v\|$	$\|f\|$
Orthogonality	$v \perp w$	$f \perp g$

What is a vector?

Familiar concept: Vectors are elements of $V = \mathbb{R}^n$ (or $\mathbb{C}^n$).

Proposition

Let $v, w \in V$ and $\alpha \in \mathbb{R}$. Then:

- $v + w \in V$ *(sums of vectors are still in V)*
- $\alpha v \in V$ *(if you scale a vector, it is still in V)*

Let us generalize!

Definition (Vector Space: Simplified)

A set V is a **vector space** over the real numbers, if for any two elements $v, w \in V$ and real number $\alpha \in \mathbb{R}$,

- $v + w \in V$
- $\alpha \cdot v \in V$

In other words, you can add and scale objects, and still stay in the same set.

Counterexample: The set of numbers in the interval $[0, 1]$

Definition (Vector Space)

A **vector space** over the real numbers $\mathbb{R}$ (often just called real vector space) is a set V together with two maps $+$ and $\cdot$, called addition (A) and multiplication (M), such that for any $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$:

- $v + w = w + v$
- $(u + v) + w = u + (v + w)$
- $0 + v = v$
- $\exists v' \in V$ s.t. $v + v' = 0$
- $\alpha(\beta v) = (\alpha\beta)v$
- $1v = v$
- $(\alpha + \beta)v = \alpha v + \beta v$
- $\alpha(v + w) = \alpha v + \alpha w$

More interesting examples:

- The set of continuous functions $C([a,b]) = \{f : [a,b] \to \mathbb{R} : f \text{cont.}\}$
- The set of absolutely integrable functions $L^1([a,b])$ consisting of $f : [a,b] \to \mathbb{R}$ such that

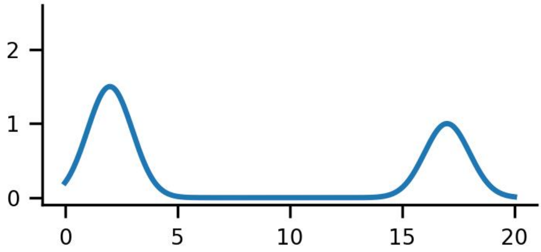
$$\int_a^b |f(x)| dx < \infty. \tag{11}$$

- The set of **square integrable functions** $L^2([a,b])$ consisting of $f : [a,b] \to \mathbb{R}$ such that

$$\int_a^b |f(x)|^2 dx < \infty. \tag{12}$$

Important examples you already knew:

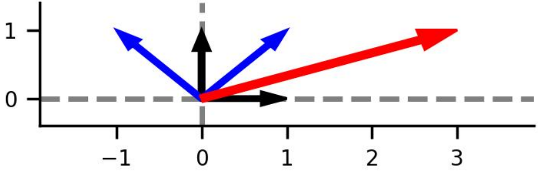
- The set of **wave functions in quantum mechanics** is a vector space (over the complex numbers, rather than real numbers).
- More generally: systems satisfying a **superposition principle**, e.g. acoustic, electromagnetic or elastic deformation waves



Familiar concept: Basis of vectors in $\mathbb{R}^2$.
Example: Standard basis vs second basis $\mathcal{B}' = \{(1,1),(-1,1)\}$

$$v = \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 3b_1 + 1b_2 \tag{13}$$

$$= \begin{pmatrix} 2 \\ -1 \end{pmatrix}_{\mathcal{B}'} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-1)\begin{pmatrix} -1 \\ 1 \end{pmatrix} = 2b_1' + (-1)b_2' \tag{14}$$



Definition (Basis)

A set $(b_1, b_2, \ldots, b_n)$ of vectors in V is called a **basis**, if any vector $v \in V$ can uniquely be written as

$$v = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_n b_n = \sum_{\nu=1}^{n} \alpha_\nu b_\nu. \tag{15}$$

For more general vector spaces: also allow $n \to \infty$:

$$v = \sum_{\nu=1}^{\infty} \alpha_\nu b_\nu \tag{16}$$

Example for functions: Many functions can be completely characterized by the coefficients $\alpha_0, \alpha_1, \alpha_2, \ldots$ of a power series representation

$$f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \cdots = \sum_\nu \alpha_\nu x^\nu \qquad (17)$$

We can interpret

$$f(x) = \sum_\nu \alpha_\nu x^\nu = \sum_\nu \alpha_\nu b_\nu(x), \qquad (18)$$

with basis functions $b_0(x) = 1, b_1(x) = x, b_2(x) = x^2, \ldots$.

Concrete examples:

$$\exp x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots = \begin{pmatrix} 1 \\ 1 \\ \frac{1}{2} \\ \frac{1}{6} \\ \vdots \end{pmatrix} \tag{19}$$

$$\sin x = 0 + x + 0 \cdot x^2 - \frac{1}{6}x^3 + \cdots = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -\frac{1}{6} \\ \vdots \end{pmatrix} \tag{20}$$

Second famous basis for function spaces:

Theorem (Fourier Series)

Let $f : \mathbb{R} \to \mathbb{R}$ *be a continuous and periodic function with period* 2π*. Then there exist coefficients* α_i *such that*

$$f(x) = \alpha_0 + \alpha_1 \sin(x) + \alpha_2 \cos(x) + \alpha_3 \sin(2x) \quad (21)$$
$$+ \alpha_4 \cos(2x) + \alpha_5 \sin(3x) + \alpha_6 \cos(3x) + \dots \quad (22)$$

Trigonometric functions thus form a basis for the space of periodic functions.

Linear maps, Matrices and Operators

Familiar concept: Matrix multiplication

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 7 \end{pmatrix}, \quad v = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \Rightarrow Av = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \tag{23}$$

Proposition

Matrix multiplication satisfies

1. $A(v + w) = Av + Aw$
2. $A(\alpha v) = \alpha Av$

Let's generalize!

Definition

A **linear map** on a vector space V is a map $A : V \to V$ such that

1. $A(v + w) = A(v) + A(w)$
2. $A(\alpha v) = \alpha A(v)$

If V is a function space, we often call A a (linear) **operator**.

Examples of operators with $f = x^3$ and $g = e^x$ (we only check addition here):

1. Derivative and Momentum operators ∂_x and $\hat{p} = -i\hbar\partial_x$:

$$(x^3 + e^x)' = 3x^2 + e^x, \tag{24}$$
$$(f + g)' = f' + g' \tag{25}$$

2. Position operator $\hat{x}$ and functions $V(\hat{x})$:

$$V(x)(x^3 + e^x) = V(x)x^3 + V(x)e^x \tag{26}$$
$$V(f + g) = Vf + Vg \tag{27}$$

3. Energy / Hamilton operator $\hat{\mathcal{H}} = \frac{1}{2m}\hat{p}^2 + V(\hat{x})$

Operators as $\infty \times \infty$-matrices:

$$f = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \end{pmatrix} \tag{28}$$

$$x \cdot f = 0 + \alpha_0 x + \alpha_1 x^2 + \cdots = \begin{pmatrix} 0 \\ \alpha_0 \\ \alpha_1 \\ \vdots \end{pmatrix} \tag{29}$$

Thus:

$$\hat{x} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \\ \vdots & & \ddots & \ddots \end{pmatrix} \tag{30}$$

satisfies:

$$\hat{x}f = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \\ \vdots & & \ddots & \ddots \end{pmatrix} \cdot \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha_0 \\ \alpha_1 \\ \vdots \end{pmatrix} \tag{31}$$

Similarly:

$$f = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots \tag{32}$$

$$f' = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 + \dots \tag{33}$$

yields:

$$\hat{p} = -i\hbar\partial_x = -i\hbar \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \\ \vdots & \vdots & & \ddots & \ddots \end{pmatrix} \tag{34}$$

Eigenvalues, Eigenvectors and Eigenfunctions

Familiar concept for matrices. Generalization is straight forward!

Definition

Let A be a linear map on V. Then, an element v such that

$$Av = \lambda v \tag{35}$$

is called an **eigenvector** of A and λ the corresponding **eigenvalue**. If V is a function space, we also call the vector v an **eigenfunction**.

Equivalence of Schrödinger and Heisenberg's approaches:
The Schrödinger equation (a partial differential equation)

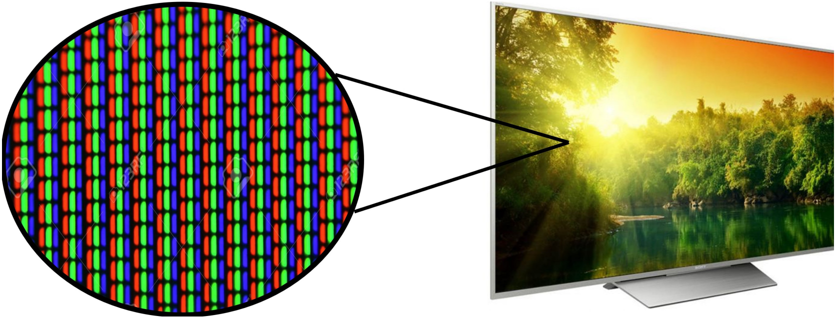
$$-\frac{\hbar^2}{2m}\psi'' + V\psi = E\psi \tag{36}$$

$$\left[-\frac{\hbar^2}{2m}\partial_x^2 + V\right]\psi = E\psi \tag{37}$$

$$\hat{\mathcal{H}}\psi = E\psi \tag{38}$$

is in fact an eigenvalue problem! We need to find the eigenfunctions (eigenvectors) $\psi(x)$ of the "matrix" (operator) $\hat{\mathcal{H}}$ with corresponding eigenvalues E, representing the energy.

Old quantum theory (1900 – 1925) vs Classical mechanics



Inner Products and Geometry

Familiar concept: Inner product of two vectors $v, w \in V = \mathbb{R}^n$:

$$\langle v|w\rangle = v \cdot w = \sum_{\nu=1}^{n} v_\nu w_\nu \tag{39}$$

Many names:

- inner product
- scalar product
- dot product

Many notations:

- $\langle v|w\rangle$
- $\langle v, w\rangle$
- (v, w)
- $v \cdot w$

Proposition (Geometric Properties of Vectors)

- *the* **lenth/size/norm** *of a vector* $\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$ *can be obtained from the inner product using* $\|v\| = \sqrt{\langle v|v\rangle}$.
- *the* **distance** *between the "points"* v, w *can be obtained from* $d(v, w) = \|v - w\|$
- $v, w \neq 0$ *are* **orthogonal** *if:* $\langle v|w\rangle = 0$.

Proposition

The (standard) inner product satisfies the following three properties:

1. **symmetry** $\langle v|w\rangle = \langle w|v\rangle$
2. **bilinearity** = *"distributive law"* = *linearity in both variables*

$$\langle u+v|w\rangle = \langle u|w\rangle + \langle v|w\rangle, \qquad \langle \alpha v|w\rangle = \alpha\langle v|w\rangle \tag{40}$$

$$\langle u|v+w\rangle = \langle u|v\rangle + \langle u|w\rangle, \qquad \langle v|\alpha w\rangle = \alpha\langle v|w\rangle \tag{41}$$

3. **positive definiteness** $\langle v|v\rangle = \|v\|^2 >= 0$, *where equality only occurs for the zero vector* $v = 0 \in \mathbb{R}^n$

Time for generalization!

Definition (Inner Product)

Let V be a vector space. Any "binary operation" $\langle . | . \rangle : V \times V \to \mathbb{R}$ that satisfies conditions (1)-(3) (symmetry, bilinearity and positive definiteness) is called an **inner product**.

Definition (Hilbert Space)

A vector space with an inner product is called an **inner product space**. If it satisfies one more technical condition ("completeness"), it is called a **Hilbert Space**.

Let $(V, \langle.|.\rangle)$ be a Hilbert space. Then:

Definition (Geometric Properties: Generalized)

- The **lenth/size/norm** of a vector is **defined** as $\|v\| = \sqrt{\langle v|v\rangle}$.
- In particular, v is **normalized/a unit vector** iff $\|v\| = 1$, i.e. $\langle v, v\rangle = 1$.
- The **distance** between v, w is **defined** as $d(v, w) = \|v - w\|$
- In particular, $v, w \neq 0$ are called **orthogonal**, iff $\langle v|w\rangle = 0$.

Remark: "iff" means "if and only if"

Theorem

Let V be the function space $L^2([a,b])$ (it also works for $C([a,b])$ etc.) and $f,g \in V$ two functions. Then the integral

$$\langle f|g\rangle = \int_a^b f(x)\cdot g(x)dx \tag{42}$$

defines an inner product $\langle f|g\rangle$ on V.

Note: If we work with complex numbers, we have to take the complex conjugate of the first function, i.e. $\int_a^b \bar{f}(x)\cdot g(x)dx$

Using this inner product and the function space $V = L^2([a, b])$, we get the geometric quantities:

- $\|f\| = \sqrt{\langle f, f, \rangle} = \sqrt{\int_a^b f^2(x)dx}$
- $d(f, g) = \|f - g\| = \sqrt{\int_a^b (f(x) - g(x))^2 dx}$
- Two functions are orthogonal with respect to this inner product, if $\langle f, g \rangle = \int_a^b f(x)g(x)dx = 0$.

Summary of our mathematical journey

Concept	$\mathbb{R}^n$	Functions
Vector	v	f
Basis	$v = \sum_{j=1}^{n} \alpha_j b_j$	$f = \sum_{j=1}^{\infty} \alpha_j b_j$
Linear Map	Matrix Av	Operator Af
Eigenvalues	$Av = \lambda v$	$Af = \lambda f$
Inner products	$\langle v, w \rangle$	$\langle f, g \rangle$
Norm (length)	$\|v\|$	$\|f\|$
Orthogonality	$v \perp w$	$f \perp g$

Fast overview of notation and techniques used in next week's lecture

Notation: State of a system

For functions, we can write $f(x)$ or just f.

Three equivalent notations for the state of a particle:

1. write $\psi(x)$ explicitly
2. just write ψ
3. write $|\psi\rangle$ and call it a "ket"

Why this name? It comes from inner products (brackets $\approx$ "BRA-KET"s)

$$\langle\psi|\phi\rangle \tag{43}$$

$\langle\psi|$ is called the "bra".

Some people write

$$\psi(x) = \langle x|\psi\rangle \tag{44}$$

and call it "position representation".
In higher dimensions:

$$\psi(x, y, z) = \langle xyz|\psi\rangle \tag{45}$$

$$\psi(\boldsymbol{r}) = \langle \boldsymbol{r}|\psi\rangle \tag{46}$$

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \langle \boldsymbol{r}_1, \boldsymbol{r}_2|\psi\rangle \tag{47}$$

Notation: Basis and Matrix Elements

Consider a set of basis vectors $b_1, b_2, \ldots$. We could write them as

$$|b_1\rangle, |b_2\rangle, \ldots, |b_\nu\rangle, \ldots \tag{48}$$

Often, we just write

$$|1\rangle, |2\rangle, \ldots, |\nu\rangle, \ldots \tag{49}$$

Matrix elements in familiar notation:

$A_{\nu\nu'}$ = element in ν-th row, ν'-th column of matrix A.

Alternative notation: $A_{\nu\nu'} = \langle \nu | A | \nu' \rangle$ (shorthand for $\langle b_\nu | A | b_{\nu'} \rangle$)

Example: $\langle 2 | A | 5 \rangle = A_{25}$.

Exponential of Operator

Example: for all $x \in \mathbb{R}$

$$e^x = \exp x = \sum_{\nu=0}^{\infty} \frac{x^\nu}{\nu!} = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots \qquad (50)$$

$$\exp\left(\begin{pmatrix} 2 & 3 \\ 3 & 7 \end{pmatrix}\right) = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 2 & 3 \\ 3 & 7 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 2 & 3 \\ 3 & 7 \end{pmatrix}^2 + \dots \qquad (51)$$

More generally, for any linear map (operator) A:

$$e^A = \exp A = \sum_{\nu=0}^{\infty} \frac{A^\nu}{\nu!} = 1 + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \dots \qquad (52)$$

Special case: Diagonal matrix

$$D = \begin{pmatrix} D_1 & & & \\ & D_2 & & \\ & & D_3 & \\ & & & \ddots \end{pmatrix} \Rightarrow D^\nu = \begin{pmatrix} D_1^\nu & & & \\ & D_2^\nu & & \\ & & D_3^\nu & \\ & & & \ddots \end{pmatrix} \tag{53}$$

Thus:

$$e^D = 1 + D + \frac{1}{2}D^2 + \cdots = \begin{pmatrix} e^{D_1} & & & \\ & e^{D_2} & & \\ & & e^{D_3} & \\ & & & \ddots \end{pmatrix} \tag{54}$$

Application: Hamiltonian (diagonalized)

$$\mathcal{H} = \begin{pmatrix} E_1 & & & \\ & E_2 & & \\ & & E_3 & \\ & & & \ddots \end{pmatrix} \tag{55}$$

Canonical Density matrix ("Boltzmann factor operator")

$$\rho = e^{-\beta\mathcal{H}} = \begin{pmatrix} e^{-\beta E_1} & & & \\ & e^{-\beta E_2} & & \\ & & e^{-\beta E_3} & \\ & & & \ddots \end{pmatrix} \tag{56}$$

From the density matrix ("Boltzmann factor operator")

$$\rho = e^{-\beta\mathcal{H}} = \begin{pmatrix} e^{-\beta E_1} & & & \\ & e^{-\beta E_2} & & \\ & & e^{-\beta E_3} & \\ & & & \ddots \end{pmatrix} \tag{57}$$

Quantum mechanical partition function

$$Q = \mathrm{Tr} e^{-\beta\mathcal{H}} = \sum_{\nu=1}^{\infty} \langle \nu | A | \nu \rangle = \sum_{\nu=1}^{\infty} e^{-\beta E_\nu} \tag{58}$$