QUANTUM MECHANICS

Lecture notes of a course of 50 lectures taught in 2008
at the University of Cape Town by Professor R.D. Viollier

Contents

1 Hilbert space

1.1 Introduction

The mathematical foundation of Quantum Mechanics (QM) is Functional Analysis (FA), which is based on the unification of Complex Analysis with Linear Algebra. QM and FA were developed concurrently, between 1925 and 1935, in a fruitful collaboration between physicists and mathematicians. The first consistent formulation of QM as a fundamental theory was published by Werner Heisenberg in 1925.

QM differs from other fundamental theories, like Classical Mechanics (CM), Classical Electrodynamics (CE), Statistical Mechanics (SM), Special Relativity (SR) and General Relativity (GR) in mainly two respects:

(i) Complex numbers play a fundamental role in QM. In fact, the imaginary unit $i = \sqrt{-1}$ enters Heisenberg's and Schrödinger's equations of motion, as well as the Born-Jordan quantization condition, while, e.g. Newton's 2nd law, Maxwell's and Einstein's equations involve only real quantities. Thus, complex numbers are of fundamental importance in QM, rather than merely being a convenient calculational tool, as in CM (e.g. in the case of coupled oscillations), in CE (e.g. in the case of electromagnetic waves or AC-circuits) or in SR (e.g. in the pseudo-Euclidean metric of Minkowski space $\mathcal{M}_4$).

(ii) QM is not deterministic, as the concept of probability enters QM fundamentally in the measurement process. This is in marked contrast to SM, where probability is merely introduced as a convenient tool to tackle the statistics of $\sim 10^{24}$ interacting particles, while the microscopic processes are still considered to be deterministic. This is one of the reasons why Albert Einstein did not accept QM as a complete and fundamental theory until his death in 1955.

1.2 Vectors

The Hilbert space $\mathcal{H}$ is an infinite-dimensional vector space over the field of the complex numbers $\mathcal{C}$. The state of a quantum system is completely determined by a "ket" vector in this fictitious space $\mathcal{H}$ which is characterized by an infinite column of complex numbers,

$$|\psi\rangle \underset{\text{def}}{=} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} .$$

The $c_i \in \mathcal{C} (i = 1,2, \cdots)$ are the components or Fourier coefficients of $|\psi\rangle$. The vectors $|\psi\rangle \in \mathcal{H}$ must have finite norm

$$\| \psi \| \underset{\text{def}}{=} \sqrt{\sum_{i=1}^{\infty} |c_i|^2} < \infty .$$

A necessary, but not sufficient condition for this to happen is

$$\lim_{n \to \infty} c_n = 0 .$$

This suggests that the components of c_n with $n > N$ may be neglected for a sufficiently large N.

1.3 Addition of vectors

The sum of two vectors $|\varphi\rangle$ and $|\psi\rangle \in \mathcal{H}$ is defined as

$$|\varphi\rangle + |\psi\rangle \underset{\text{def}}{=} \varphi + \psi = \begin{pmatrix} b_1 + c_1 \\ b_2 + c_2 \\ \vdots \end{pmatrix}$$

$$\text{with} \qquad |\varphi\rangle \underset{\text{def}}{=} \varphi = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix} .$$

If $\| \varphi \| < \infty$ and $\| \psi \| < \infty$, then also $\| \varphi + \psi \| < \infty$, as we prove in tut 1(b). This means that $|\varphi\rangle + |\psi\rangle \in \mathcal{H}$, as well.

- properties:

  (which follow directly from the definition of the sum of two vectors)

$$\begin{aligned}
|\psi\rangle + |\varphi\rangle &= |\varphi\rangle + |\psi\rangle && \text{commutative law} \\
(|\psi\rangle + |\varphi\rangle) + |\chi\rangle &= |\psi\rangle + (|\varphi\rangle + |\chi\rangle) && \text{associative law} \\
|\psi\rangle + |0\rangle = |0\rangle + |\psi\rangle &= |\psi\rangle && \exists \text{ null vector} \\
|0\rangle &\underset{\text{def}}{=} \begin{pmatrix} 0 \\ 0 \\ \vdots \end{pmatrix} \\
|\psi\rangle + |-\psi\rangle &= |-\psi\rangle + |\psi\rangle = |0\rangle && \exists \text{ negative vector} \\
|-\psi\rangle \underset{\text{def}}{=} -|\psi\rangle = -\psi &= \begin{pmatrix} -c_1 \\ -c_2 \\ \vdots \end{pmatrix} .
\end{aligned}$$

Thus, $\mathcal{H}$ forms a commutative group with respect to the addition of vectors.

1.4 Multiplication of a vector with a scalar

- definition:

The product of a vector with a complex number is defined as

$$a|\psi\rangle \underset{\text{def}}{=} a\psi = \begin{pmatrix} ac_1 \\ ac_2 \\ \vdots \end{pmatrix} \quad a \in \mathcal{C} \quad , \quad |\psi\rangle \in \mathcal{H} \ .$$

As $a|\psi\rangle$ has finite norm, i.e.

$$\| a\psi \|^2 = \sum_{i=1}^{\infty} |a|^2 |c_i|^2 = |a|^2 \underbrace{\sum_{i=1}^{\infty} |c_i|^2}_{<\infty} < \infty \ ,$$

it is contained in $\mathcal{H}$, i.e. $a|\psi\rangle \in \mathcal{H}$.

- properties:

(which follow directly from the definition)

$$\left.\begin{aligned} a(|\psi\rangle + |\varphi\rangle) &= a|\psi\rangle + a|\varphi\rangle \\ (a+b)|\psi\rangle &= a|\psi\rangle + b|\psi\rangle \\ (ab)|\psi\rangle &= a(b|\psi\rangle) \\ 1\,|\psi\rangle &= |\psi\rangle \end{aligned}\right\} \text{linearity laws}$$

- definition:

N vectors $|\psi_1\rangle, |\psi_2\rangle, \cdots |\psi_N\rangle$ are linearly independent, if the linear relation

$$z_1|\psi_1\rangle + z_2|\psi_2\rangle + \cdots + z_N|\psi_N\rangle = |0\rangle$$

only holds for $z_1 = z_2 = \cdots = z_N = 0$. Otherwise the N vectors are linearly dependent. Similarly, in Euclidean space $\mathcal{E}_3$, if $\vec{a}$ and $\vec{b}$ point in different directions, $\lambda\vec{a} + \mu\vec{b} = \vec{0}$ can only be fulfilled for $\lambda = \mu = 0$.

1.5 Scalar product

- definition:

To every ordered pair of vectors $|\varphi\rangle$ and $|\psi\rangle \in \mathcal{H}$, we can associate a complex number through the Dirac $\langle$bra$|$ket$\rangle$, introduced by P.A.M. Dirac in 1927, i.e.
c

$$\langle\varphi|\psi\rangle \underset{\text{def}}{=} \varphi^\dagger\psi = (b_1^*, b_2^*, \cdots) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} = \sum_{i=1}^{\infty} b_i^* c_i \ \in \ \mathcal{C} \ , \quad b_i^* \text{ being the complex conjugate of } b_i.$$

Similarly, we have in $\mathcal{E}_3$: $\vec{a}\cdot\vec{b} = a_x\, b_x + a_y\, b_y + a_z\, b_z$.
Euclidean scalar product

The Dirac "bra" vector is characterized by an infinite sequence of complex numbers

$$\langle\varphi| \underset{\text{def}}{=} \varphi^\dagger = (b_1^*, b_2^*, \cdots) \in \mathcal{H}^\dagger \ ,$$

where $\mathcal{H}^\dagger$ is the dual Hilbert space. Thus $\langle\varphi|$ fulfils the same linearity laws as $|\varphi\rangle$. The scalar product $\langle\varphi|\psi\rangle$ is finite if the norms of both $|\varphi\rangle$ and $|\psi\rangle$, are finite. See tut 1(a).

- properties:

(which follow directly from the definition)

$$\left.\begin{aligned} \langle\varphi|\psi+\chi\rangle &= \langle\varphi|\psi\rangle + \langle\varphi|\chi\rangle \\ \langle\varphi|a\psi\rangle &= a\langle\varphi|\psi\rangle \end{aligned}\right\} \text{linearity}$$

$$\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^* \qquad \text{conjugation}$$

$$\langle\varphi|\varphi\rangle > 0 \qquad \text{positivity}$$

$$\text{unless} \quad |\varphi\rangle = |0\rangle \ .$$

- consequences:

$$\begin{aligned} \langle\varphi+\psi|\chi\rangle &= \langle\varphi|\chi\rangle + \langle\psi|\chi\rangle \\ \langle a\varphi|\psi\rangle &= a^*\langle\varphi|\psi\rangle \ . \end{aligned}$$

The norm of the vector $|\psi\rangle$ can now be written as

$$\| \psi \| = \sqrt{\langle\psi|\psi\rangle} = \sqrt{\sum_{i=1}^{\infty} |c_i|^2} \ .$$

- definition:

  $|\varphi\rangle$ and $|\psi\rangle$ are orthogonal to each other

  iff $\langle\varphi|\psi\rangle = \sum_{i=1}^{\infty} b_i^* c_i = 0$ and both $\| \varphi \| \neq 0$ and $\| \psi \| \neq 0$.

  The following inequalities hold:

- Schwarz's inequality:

  $|\langle\varphi|\psi\rangle| \leqslant \| \varphi \| \ \| \psi \|$ (see tut 1(a))

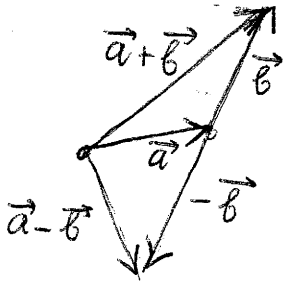
  Similarly, in $\mathcal{E}_3$: $|\vec{a} \cdot \vec{b}| = |\vec{a}||\vec{b}| \cos\gamma| \leqslant |\vec{a}| \, |\vec{b}|$

  γ = the smaller angle between $\vec{a}$ and $\vec{b}$

- Minkowski's inequality:

  $\| \varphi \pm \psi \| \leqslant \| \varphi \| + \| \psi \|$ (see tut 1(b))

  similarly, in $\mathcal{E}_3$:



  $|\vec{a} \pm \vec{b}| \leqslant |\vec{a}| + |\vec{b}|$

- triangular inequality:

  $\| \psi - \varphi \| \leqslant \| \psi - \chi \| + \| \varphi - \chi \|$ (see tut 1(c))

  Similarly, in $\mathcal{E}_3$: $|\vec{a} - \vec{b}| \leqslant |\vec{a}| + |\vec{b}|$ clear!

  replace $\vec{a} \to \vec{a} - \vec{c}$; $\vec{b} \to \vec{b} - \vec{c}$

  $\Rightarrow \underbrace{|\vec{a} - \vec{c} - (\vec{b} - \vec{c})|}_{|\vec{a} - \vec{b}|} \leqslant |\vec{a} - \vec{c}| + |\vec{b} - \vec{c}|$.

- conclusion:

  The scalar product can be used to measure "distances" and "angles" in $\mathcal{H}$. It is, therefore, the metric of $\mathcal{H}$.

1.6 Basis vectors

- definition:

  A convenient set of vectors, spanning up the Hilbert space $\mathcal{H}$, is

$$|\psi_1\rangle = \psi_1 \underset{\text{def}}{=} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \ , \ |\psi_2\rangle = \psi_2 \underset{\text{def}}{=} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \ , \cdots$$

- properties:

  This allows us to write

$$|\psi\rangle = \psi \underset{\text{def}}{=} \sum_{i=1}^{\infty} c_i |\psi_i\rangle = \sum_{i=1}^{\infty} c_i \psi_i = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix}$$

  for all $|\psi\rangle \in \mathcal{H}$, i.e. the set of vectors $|\psi_i\rangle = \psi_i \quad (i = 1, 2, \cdots)$ form a basis of $\mathcal{H}$.

  Similarly, in $\mathcal{E}_3$: every vector $\vec{a}$ can be expanded as

$$\vec{a} = a_x \vec{e}_x + a_y \vec{e}_y + a_z \vec{e}_z = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \ ,$$

  in which $\vec{e}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{e}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

  form a basis of $\mathcal{E}_3$.

  orthonormality:

  The N vectors $|\psi_1\rangle, |\psi_2\rangle \, , \cdots , |\psi_N\rangle$ are obviously linearly independent and orthonormal ($\equiv$ orthogonal and normalized), i.e.

  $\langle\psi_m|\psi_n\rangle = \psi_m^\dagger \psi_n = \delta_{mn} \qquad (m, n = 1, \ \cdots \ , \ N)$.

  Similarly, in $\mathcal{E}_3$:
  $\vec{e}_i \cdot \vec{e}_j = \delta_{ij} \quad (i, j = x, y, z)$ are orthonormal.

  dimension of $\mathcal{H}$:

  As N can be arbitrarily large, the dimension of $\mathcal{H}$ is infinite, i.e.

  $\dim \mathcal{H} = \infty$, while the dimension of $\mathcal{E}_3$ is $\dim \mathcal{E}_3 = 3$.

1.7 Completeness

- Let $|\psi_i'\rangle \quad (i = 1,2, \cdots)$, with $\langle\psi_i'|\psi_k'\rangle = \delta_{ik} \quad (i,k = 1,2, \cdots)$, be a different set of orthonormal vectors in $\mathcal{H}$. This set of vectors is complete in $\mathcal{H}$ (i.e. it is a basis of $\mathcal{H}$), if every vector $|\psi\rangle \in \mathcal{H}$ can be expanded as

$$|\psi\rangle = \sum_{i=1}^{\infty} c_i'|\psi_i'\rangle = \sum_{i=1}^{\infty} c_i|\psi_i\rangle = \text{invariant.}$$

In general $c_i' \neq c_i$ and $|\psi_i'\rangle \neq |\psi_i\rangle \quad (i = 1,2, \cdots)$.

Similarly, we can expand every vector $\vec{a} \in \mathcal{E}_3$ in a different set of orthonormal basis vectors $\vec{e}_i' \cdot \vec{e}_k' = \delta_{ik} \quad (i,k = x,y,z)$, i.e.

$$\vec{a} = a_x' \vec{e}_x' + a_y' \vec{e}_y' + a_z' \vec{e}_z' = a_x\vec{e}_x + a_y\vec{e}_y + a_z\vec{e}_z = \text{invariant.}$$

In general $a_i' \neq a_i$, and $\vec{e}_i' \neq \vec{e}_i, \quad (i = x,y,z)$.

The c_i' are the complex components or Fourier coefficients of $|\psi\rangle$ in the complete and orthonormal set $|\psi_i'\rangle \quad (i = 1,2, \cdots)$ of $\mathcal{H}$, i.e.

$$\begin{aligned} \langle\psi_i'|\psi\rangle &= \langle\psi_i'|\sum_{k=1}^{\infty} c_k'\psi_k'\rangle \\ &= \sum_{k=1}^{\infty} c_k' \underbrace{\langle\psi_i'|\psi_k'\rangle}_{\delta_{ik}} = c_i' \\ \Rightarrow c_i' &= \langle\psi_i'|\psi\rangle = \psi_i'^{\dagger}\psi \in \mathcal{C} \ . \end{aligned}$$

Similarly, in $\mathcal{E}_3$ we have $\quad a_i' = \vec{a} \cdot \vec{e}_i' \in \mathcal{R} \quad (i = x,y,z)$.

- Let $|\psi_i'\rangle \quad (i = 1,2, \cdots)$ be an arbitrary complete and orthonormal set of vectors in $\mathcal{H}$. The Fourier expansions of the vectors $|\varphi\rangle$ and $|\psi\rangle \in \mathcal{H}$ in this basis are

$$|\varphi\rangle = \sum_{i=1}^{\infty} b_i'|\psi_i'\rangle \quad \text{and} \quad |\psi\rangle = \sum_{i=1}^{\infty} c_i'|\psi_i'\rangle \ ,$$

with the Fourier coefficients

$$c_i' = \langle\psi_i'|\psi\rangle \in \mathcal{C} \quad \text{and} \quad b_i' = \langle\psi_i'|\varphi\rangle \in \mathcal{C} \ .$$

The scalar product is thus

$$\begin{aligned} \langle\varphi|\psi\rangle &= \left\langle \sum_{i=1}^{\infty} b_i'\psi_i' \mid \sum_{k=1}^{\infty} c_k'\psi_k' \right\rangle \\ &= \sum_{i,k=1}^{\infty} b_i'^{*} c_k' \underbrace{\langle\psi_i'|\psi_k'\rangle}_{\delta_{ik}} = \sum_{i=1}^{\infty} b_i'^{*} c_i' = \sum_{i=1}^{\infty} b_i^{*} c_i = \text{invariant} \ . \end{aligned}$$

This is valid for every complete and orthonormal set $|\psi_i'\rangle$.
Now inserting the Fourier coefficients, we have

$$\langle\varphi|\psi\rangle = \sum_{i=1}^{\infty} \langle\psi_i'|\varphi\rangle^{*} \langle\psi_i'|\psi\rangle = \sum_{i=1}^{\infty} \langle\varphi|\psi_i'\rangle\langle\psi_i'|\psi\rangle \quad \text{Parseval's equation.}$$

Similarly, in $\mathcal{E}_3$ we have

$$\begin{aligned} \vec{a} \cdot \vec{b} &= (\vec{a} \cdot \vec{e}_x')(\vec{e}_x' \cdot \vec{b}) + (\vec{a} \cdot \vec{e}_y')(\vec{e}_y' \cdot \vec{b}) + (\vec{a} \cdot \vec{e}_z')(\vec{e}_z' \cdot \vec{b}) \\ &= a_x' \, b_x' + a_y' \, b_y' + a_z' \, b_z' = a_x b_x + a_y b_y + a_z b_z = \text{invariant.} \end{aligned}$$

- special case:

For $|\varphi\rangle = |\psi\rangle$, we have

$$\begin{aligned} \| \psi \|^2 = \langle\psi|\psi\rangle &= \sum_{i=1}^{\infty} \langle\psi|\psi_i'\rangle\langle\psi_i'|\psi\rangle = \sum_{i=1}^{\infty} |\langle\psi_i'|\psi\rangle|^2 \quad \text{Bessel's equation} \\ &= \sum_{i=1}^{\infty} |c_i'|^2 = \sum_{i=1}^{\infty} |c_i|^2 = \text{invariant.} \end{aligned}$$

Thus, if the set $|\psi_i'\rangle \ (i = 1,2, \cdots)$ is incomplete in that e.g. $|\psi_k'\rangle$ with $k = 37$ missing, one could not represent vectors $|\psi\rangle$ having a component in the $|\psi_k'\rangle$ direction.

Similarly, in $\mathcal{E}_3$ we have

$$\begin{aligned} \vec{a} \cdot \vec{a} &= (\vec{a} \cdot \vec{e}_x')^2 + (\vec{a} \cdot \vec{e}_y')^2 + (\vec{a} \cdot \vec{e}_z')^2 \\ &= a_x'^2 + a_y'^2 + a_z'^2 = a_x^2 + a_y^2 + a_z^2 = \text{invariant.} \end{aligned}$$

2 Operators

2.1 Definition

- The dynamical variables (i.e. momentum, position vector, angular momentum, energy, etc.) of a quantum system are operators. An operator F maps a general vector $|\psi\rangle \in \mathcal{H}$ onto another vector $|\varphi\rangle \in \mathcal{H}$

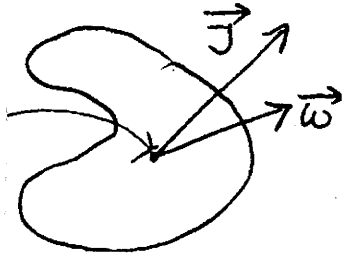
$$|\varphi\rangle \underset{\text{def}}{=} F|\psi\rangle = |F\psi\rangle \in \mathcal{H} \ .$$

As $|F\psi\rangle \in \mathcal{H}$, it must have finite norm, i.e. $\| F\psi \| < \infty$. Similarly, in $\mathcal{E}_3$: The relation between the angular momentum $\vec{J}$, the moment of inertia tensor θ, and the angular velocity $\vec{\omega}$ of an arbitrarily shaped rotating rigid body is given by $\vec{J} = \theta\vec{\omega}$ or

$$\begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix} = \begin{pmatrix} \theta_{xx} & \theta_{xy} & \theta_{xz} \\ \theta_{yx} & \theta_{yy} & \theta_{yz} \\ \theta_{zx} & \theta_{zy} & \theta_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \ ,$$

where $\vec{J}$ and $\vec{\omega}$ point in different directions, in general.

centre-of-mass



2.2 Matrix representation of operators

- Linear operators satisfy the linearity law

$$F(a|\varphi\rangle + b|\psi\rangle) = aF|\varphi\rangle + bF|\psi\rangle \quad a, b \in \mathcal{C}; \quad |\varphi\rangle, |\psi\rangle \in \mathcal{H} \ .$$

We henceforth restrict ourselves to linear operators.

- Let $|\psi_k\rangle$ $(k = 1, 2, \cdots)$ be a complete and orthonormal set of vectors of $\mathcal{H}$.

$$|\varphi\rangle \text{ and } |\psi\rangle \in \mathcal{H} \ \Rightarrow \ \left\{ \begin{array}{lcl} |\varphi\rangle & = & \sum_{k=1}^{\infty} b_k |\psi_k\rangle \\ |\psi\rangle & = & \sum_{k=1}^{\infty} c_k |\psi_k\rangle \end{array} \right. \quad \begin{array}{l} \text{Fourier} \\ \text{expansions} \end{array}$$

with the Fourier coefficients

$$b_k = \langle\psi_k|\varphi\rangle \text{ and } c_k = \langle\psi_k|\psi\rangle \qquad b_k, c_k \in \mathcal{C} \ .$$

Inserting the Fourier expansions for $|\varphi\rangle$ and $|\psi\rangle$ into $|\varphi\rangle = |F\psi\rangle$, we obtain, using the linearity law,

$$\sum_{k=1}^{\infty} b_k |\psi_k\rangle = F \sum_{k=1}^{\infty} c_k |\psi_k\rangle = \sum_{k=1}^{\infty} c_k |F\psi_k\rangle \ .$$

Taking the scalar product from the left with $\langle\psi_i|$

$$\sum_{k=1}^{\infty} b_k \underbrace{\langle\psi_i|\psi_k\rangle}_{\delta_{ik}} = \sum_{k=1}^{\infty} c_k \underbrace{\langle\psi_i|F\psi_k\rangle}_{F_{ik} \in \mathcal{C}} \ ,$$

we arrive at

$$b_i = \sum_{k=1}^{\infty} F_{ik} c_k \ .$$

Here,

$$F_{ik} \underset{\text{def}}{=} \langle\psi_i|F\psi_k\rangle \qquad \in \mathcal{C}$$

are the (finite) matrix elements of the linear operator F in the complete and orthonormal set $|\psi_k\rangle$ of $\mathcal{H}$. In matrix notation the defining equation for an operator is therefore,

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} F_{11}, & F_{12}, & \cdots \\ F_{21}, & F_{22}, & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} \text{ or } \varphi = F\psi \text{ or } b_i = \sum_{k=1}^{\infty} F_{ik} c_k$$

with $b_i \underset{\text{def}}{=} \langle\psi_i|\varphi\rangle = \psi_i^\dagger \varphi$, $c_i \underset{\text{def}}{=} \langle\psi_i|\psi\rangle = \psi_i^\dagger \psi$ and $F_{ik} \underset{\text{def}}{=} \langle\psi_i|F\psi_k\rangle = \psi_i^\dagger F \psi_k$.

Thus linear operators can be represented by infinite square matrices.

2.3 Sum and differences of operators

- <u>definition:</u>

$$\begin{array}{rcl} (G \pm F)|\psi\rangle & \underset{\text{def}}{=} & G|\psi\rangle \pm F|\psi\rangle \\ & = & |G\psi\rangle \pm |F\psi\rangle \\ & = & |(G \pm F)\psi\rangle \end{array}$$

valid for every $|\psi\rangle \in \mathcal{H}$.

Applying this to a complete and orthonormal set of vectors $|\psi_k\rangle$, and taking the scalar product with $\langle\psi_i|$ from the left, we have

$$\begin{array}{rcll} \langle\psi_i|(G \pm F)\psi_k\rangle & = & \langle\psi_i|G\psi_k\rangle \pm \langle\psi_i|F\psi_k\rangle & \\ \Rightarrow \quad (G \pm F)_{ik} & = & G_{ik} \pm F_{ik} & \text{matrix addition and subtraction law.} \end{array}$$

2.4 Product of operators

- definition:

$$(GF)|\psi\rangle \underset{\text{def}}{=} G(F|\psi\rangle) = G|\varphi\rangle$$
$$\text{with } |\varphi\rangle \underset{\text{def}}{=} F|\psi\rangle = |F\psi\rangle \text{ and } |\varphi\rangle, |\psi\rangle \in \mathcal{H} \ .$$

Writing these equations in matrix form, we have

$$\sum_{k=1}^{\infty}(GF)_{ik}\, c_k = \sum_{l=1}^{\infty} G_{il}\, b_l = \sum_{l=1}^{\infty} G_{il} \underbrace{\sum_{k=1}^{\infty} F_{lk} c_k}_{b_l} = \sum_{k,l=1}^{\infty} G_{il}\, F_{lk}\, c_k \ .$$

As c_k is arbitrary in this equation, we may choose e.g. $c_k = \delta_{kn}$

$$\Rightarrow \quad (GF)_{in} = \sum_{l=1}^{\infty} G_{il}\, F_{ln} \ .$$

This matrix multiplication law, already known in Linear Algebra, was rediscovered by W. Heisenberg in 1925, when he was a postdoctoral fellow with M. Born (Göttingen). As in general, $GF \neq FG$ for matrices, this also holds for operators.

- special matrices:

$$aI \underset{\text{def}}{=} \begin{pmatrix} a & 0 & 0 & \cdots \\ 0 & a & 0 & \cdots \\ 0 & 0 & a & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{with } a \in C \ ,$$

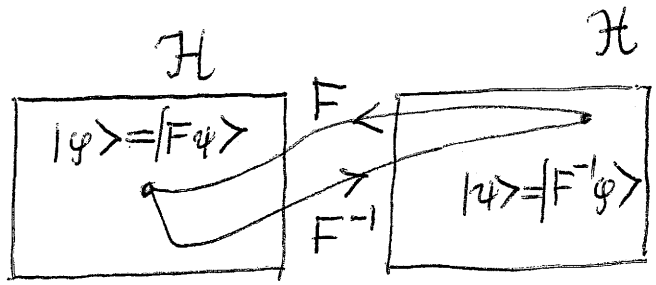
aI commutes with every matrix F. For $a = 1$ this is the unit matrix I.

2.5 Inverse operator

- definition: Let $|\varphi\rangle = F|\psi\rangle = |F\psi\rangle$. We now define the truncated matrix

$$F_N \underset{\text{def}}{=} \begin{pmatrix} F_{11} & \cdots & F_{1N} \\ \vdots & & \vdots \\ F_{N1} & \cdots & F_{NN} \end{pmatrix} \ .$$

If $\lim\limits_{N\to\infty} F_N^{-1} \underset{\text{def}}{=} F^{-1}$ exists, the inverse F^{-1} is uniquely defined and F is invertible. A necessary, but not sufficient condition for this to happen is $\det F_N \neq 0$ for $N > N_0$.



- properties:

(which follow from the definition)

$$(F^{-1}F)\,|\psi\rangle = F^{-1}(F|\psi\rangle) = F^{-1}|\varphi\rangle = |\psi\rangle$$
valid for every $|\psi\rangle \in \mathcal{H}$
thus $F^{-1}F = I$

$$(FF^{-1})\,|\varphi\rangle = F\left(F^{-1}|\varphi\rangle\right) = F|\psi\rangle = |\varphi\rangle$$
valid for every $|\varphi\rangle \in \mathcal{H}$
thus $FF^{-1} = I$

and, therefore,
$FF^{-1} = F^{-1}F = I$
$\Rightarrow F$ and F^{-1} commute

$$(F^{-1})^{-1} = F \qquad \text{(see tut 2)}$$

$$(FG)^{-1} = G^{-1}F^{-1} \text{ (see tut 2).}$$

The quotient F/G has no meaning, unless $[F, G^{-1}] = 0$.

$$(FGH)^{-1} = H^{-1}G^{-1}F^{-1} \quad \text{(see tut 2).}$$

2.6 Powers and functions of operators

- definition: $F^0 \underset{\text{def}}{=} I$

$$F^{n+1} = F\,F^n \quad (n = 0, 1, 2, \cdots) \quad \text{defined recursively.}$$

If F^{-1} defined, then

$$F^{-n-1} = F^{-n}F^{-1} \qquad (n = 1, 2, \cdots)$$

is also defined recursively. Thus all positive and negative powers of F are defined. We can now represent functions of an operator in terms of a Laurent series, i.e.

$$g(F) = \sum_{n=-k}^{\infty} a_n\,F^n, \qquad a_n \in \mathcal{C}\ ,$$

or a Taylor series, i.e.

$$f(F) \underset{\text{def}}{=} \sum_{n=0}^{\infty} \frac{f^n(0)}{n!}\,F^n\ .$$

- example: $\exp(i\omega\sigma_x)$ with $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (see tut 3 and also tut 12).

2.7 Adjoint operator

- definition: $F^\dagger$ is the adjoint operator of F

iff $\langle\varphi|F\psi\rangle \underset{\text{def}}{=} \langle F^\dagger\varphi|\psi\rangle$ for

every $|\psi\rangle\ ,\ |\varphi\rangle \in \mathcal{H}$

- consequence:

Let $|\psi_i\rangle$ $(i = 1, 2, \cdots)$ be a complete and orthonormal set of vectors $\in \mathcal{H}$

$$\langle\psi_n|F\psi_m\rangle = \langle F^\dagger\psi_n|\psi_m\rangle = \langle\psi_m|F^\dagger\psi_n\rangle^*$$

$$\Rightarrow \quad \left(F^\dagger\right)_{mn} = F^*_{nm}$$

$F^\dagger$ is the transposed and complex conjugate matrix F.

- special case: An operator A is Hermitean iff $A = A^\dagger$.

remark: Hermitean operators describe dynamical variables, because their eigenvalues are real, as we will see later.

- properties: (see tut 4)

$$\begin{aligned}
F^{\dagger\dagger} &= F \\
(fF + gG)^\dagger &= f^*\,F^\dagger + g^*\,G^\dagger \quad f, g \in \mathcal{C} \\
(FG)^\dagger &= G^\dagger F^\dagger \\
(FGH)^\dagger &= H^\dagger\,G^\dagger\,F^\dagger \\
F + F^\dagger &\text{ is Hermitean} \\
i(F - F^\dagger) &\text{ is Hermitean} \\
\text{if } A^\dagger &= A,\ B^\dagger = B^\dagger \text{ (Hermitean) and } r \in \mathcal{R}, \text{ then}
\end{aligned}$$

$$\left.\begin{aligned}
(A \pm B)^\dagger &= A \pm B \\
(rA)^\dagger &= rA \\
(AB + BA)^\dagger &= AB + BA \\
[i\,(AB - BA)]^\dagger &= i\,(AB - BA)
\end{aligned}\right\} \text{ also Hermitean.}$$

- definition:

$$\begin{aligned}
[F, G] &\underset{\text{def}}{=} FG - GF \qquad \text{commutator} \\
\{F, G\} &\underset{\text{def}}{=} FG + GF \qquad \text{anticommutator}
\end{aligned}$$

2.8 Basis operators

Let $|\psi_j\rangle$ be a complete and orthonormal set of vectors of $\mathcal{H}$.

- theorem: F can be represented as

$$F = \sum_{j,l=1}^{\infty} F_{jl}|\psi_j\rangle\langle\psi_l|\ ,$$

where $|\psi_j\rangle\langle\psi_l|$ is a basis matrix, having the matrix element 1 in the k-th column and j-th sequence, and all other matrix elements vanishing, i.e.

$$|\psi_j\rangle\langle\psi_l| = \begin{pmatrix} & & l\text{-th column} \\ & \downarrow & \\ 0 & \vdots\ 0 & \\ \cdots & 1\ \cdots & \leftarrow j^{\text{th}} \text{ sequence} \\ 0 & \vdots\ 0 & \end{pmatrix} \qquad (j, l = 1, 2, \cdots).$$

$$= \psi_j \otimes \psi_l^\dagger \qquad \text{direct product}$$

proof: Show that the matrix elements of F come out correctly

(i) apply F to $|\psi_k\rangle$:

$$F|\psi_k\rangle = |F\psi_k\rangle = \sum_{j,l=1}^{\infty} F_{jl}|\psi_j\rangle \underbrace{\langle\psi_l|\psi_k\rangle}_{\delta_{lk}}$$
$$= \sum_{j=1}^{\infty} F_{jk}|\psi_j\rangle$$

(ii) take scalar product with $\langle\psi_i|$:

$$\langle\psi_i|F\psi_k\rangle = \sum_{j=1}^{\infty} F_{jk} \underbrace{\langle\psi_i|\psi_j\rangle}_{\delta_{ij}} = F_{ik} \quad \text{q.e.d.}$$

special case:
The completeness relation can be written as

- theorem: $I = \sum_{n=1}^{\infty} |\psi_n\rangle\langle\psi_n|$

proof: apply I to $|\psi\rangle \in \mathcal{H}$ $\Rightarrow$ $I\,|\psi\rangle = \sum_{n=1}^{\infty} |\psi_n\rangle \underbrace{\langle\psi_n|\psi\rangle}_{c_n} = |\psi\rangle$

valid for every $|\psi\rangle$ q.e.d.

2.9 Expectation values

definition: The expectation value $\overline{F}$ of an operator F in a state $|\psi\rangle$, normalized as

$\||\psi\|^2 = \langle\psi|\psi\rangle = \psi^\dagger\psi = \sum_{i=1}^{\infty} |c_i|^2 = 1$, is given by

$$\overline{F} \underset{\text{def}}{=} \langle\psi|F\psi\rangle = \sum_{i,k=1}^{\infty} c_i^* \langle\psi_i|F\,\psi_k\rangle\, c_k = \sum_{i,k=1}^{\infty} c_i^* F_{ik}\, c_k = \psi^\dagger F\psi \ .$$

$\overline{F}$ depends on the state vector $|\psi\rangle$, of course.

properties: (for proof see tut 5)

$$\overline{fF+gG} = f\,\overline{F} + g\,\overline{G}$$
$$\overline{F^\dagger} = \overline{F}^*$$
$$\overline{F^\dagger F} \geqslant 0 \ .$$

3 Eigenvalues, eigenvectors and axioms of measurement

3.1 Eigenvalue problem

Consider the eigenvalue equation for a Hermitean operator $A = A^\dagger$

$$A\,|\psi\rangle = a\,|\psi\rangle \qquad a \in \mathcal{C}$$
$$|\psi\rangle \in \mathcal{H} \ ,$$

i.e. we are looking for non-trivial ($\||\psi\|| \neq 0$) solutions $|\psi\rangle$, with $A\,|\psi\rangle$ "parallel" to $|\psi\rangle$. The eigenvalue equation can be rewritten as

$$(A - aI)\,|\psi\rangle = |0\rangle \ ,$$

or in matrix form as

$$\begin{pmatrix} A_{11}-a, & A_{12}, & A_{13}, & \cdots \\ A_{21}, & A_{22}-a, & A_{23}, & \cdots \\ A_{31}, & A_{32}, & A_{33}-a, & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \ .$$

A sensible approach towards solving this infinite system of coupled complex linear algebraic equations, is to truncate the Hilbert space after the first N dimensions, calculating the eigenvalues and eigenvectors, and then investigating the limit $N \to \infty$. We may then identify those eigenvalues and eigenvectors which become numerically stable as $N \to \infty$.

The truncated system of N coupled complex linear algebraic equations has non-trivial solutions $|\psi^N\rangle$, with $\||\psi^N\|| \neq 0$, only for

$$P_N(a) \underset{\text{def}}{=} \det(A_N - aI_N) = 0 \quad \text{(see tut 6).}$$

The N eigenvalues a_n^N (n=1,2, $\cdots$ N) are obtained from the zero's of the characteristic polynomial $P_N(a)$ of degree N.

One gets the corresponding N eigenvectors $|\psi_n^N\rangle$ (n = 1,2, $\cdots$) from the solution of the system of complex coupled linear algebraic equations (see tut 6).

$$(A_N - a_n^N I_N)\,|\psi_n^N\rangle = |0\rangle \qquad n = 1, \cdots, N.$$

The normalization of the eigenvectors is arbitrary. Due to the symmetries of A, some of the eigenvalues may be identical, corresponding to a multiple zero of the characteristic polynomial. The maximal number g of linearly independent eigenstates, having the same eigenvalue $a_k = a$ is called the degeneracy of the eigenvalue a_k.

3.2 Orthogonality of the eigenvectors of $A = A^\dagger$

- theorem: for $A = A^\dagger$ Hermitean

  (i) the eigenvalues are real

  (ii) the eigenvectors to different eigenvalues are orthogonal.

proof:

write down
$$\begin{cases} |A\,\psi_m\rangle &= a_m|\psi_m\rangle & \text{(i)} \\ |A\,\psi_n\rangle &= a_n|\psi_n\rangle & \text{(ii)} \end{cases}$$

apply $\langle\psi_n|$ to (i) $\Rightarrow \langle\psi_n|A\,\psi_m\rangle = a_m\langle\psi_n|\psi_m\rangle$ (i)′ eq.(1)

and $\langle\psi_m|$ to (ii) $\Rightarrow \langle\psi_m|A\,\psi_n\rangle = a_n\langle\psi_m|\psi_n\rangle$ (ii)′

take complex conjugate of (ii)′ $\langle\psi_m|A\,\psi_n\rangle^* = a_n^*\langle\psi_m|\psi_n\rangle^*$ (ii)*

interchange arguments of (ii)* $\langle A\,\psi_n|\psi_m\rangle = a_n^*\langle\psi_n|\psi_m\rangle$ (ii)*

use definition of $A^\dagger$ $\langle\psi_n|A^\dagger\,\psi_m\rangle = a_n^*\langle\psi_n|\psi_m\rangle$ (ii)*

use Hermiticity of A $\langle\psi_n|A\,\psi_m\rangle = a_n^*\langle\psi_n|\psi_m\rangle$ (ii)* eq.(2)

subtract eq.(2) from eq.(1)

$$\Rightarrow \quad 0 = (a_m - a_n^*)\,\langle\psi_n|\psi_m\rangle$$

for $m = n$: $(a_n - a_n^*)\,\langle\psi_n|\psi_n\rangle = 0$,
as $\|\psi_n\| \neq 0$ per definition
$\Rightarrow$ $a_n = a_n^*$, i.e. all eigenvalues are real

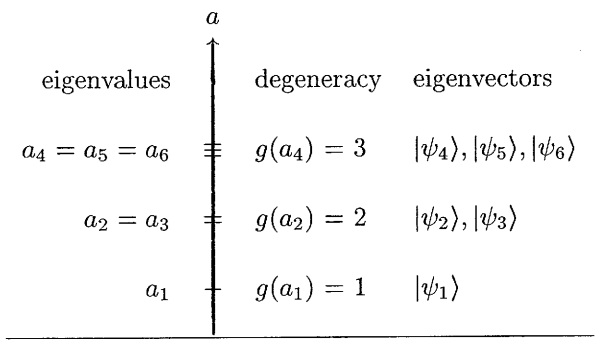
for $a_m \neq a_n$: $\Rightarrow \underbrace{(a_m - a_n)}_{\neq 0}\,\langle\psi_n|\psi_m\rangle = 0$
$\Rightarrow \langle\psi_n|\psi_m\rangle = 0$
i.e. eigenvectors to different eigenvalues are orthogonal (see tut 7)

remark: If $a_m = a_n$ with $m \neq n$, $\langle\psi_n|\psi_m\rangle$ may well be non-zero. Thus the matrix of the scalar products $I'_{nm} \underset{\text{def}}{=} \langle\psi_n|\psi_m\rangle$ has, in general, box-diagonal form.

- example: 6×6 matrix $I'_{n,m} = \langle\psi_n|\psi_m\rangle \neq \delta_{n,m}$
  Here the ψ_n are, in general, not an orthonormal set of vectors.

$$I' = \begin{pmatrix} \times & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \end{pmatrix} \quad \times = \text{nonzero matrix element in general}$$

corresponding to the eigenvalue spectrum:



a_k: eigenvalue

$g(a_k)$: degeneracy of the eigenvalue a_k

$g(a_k)$ is the maximal number of linearly independent eigenvectors that have the same eigenvalue a_k. These eigenvectors are orthogonal to all eigenvectors having different eigenvalues. However, they are not necessarily orthogonal to eigenvectors having the same eigenvalues.

3.3 Orthonormalization of eigenvectors

We are thus faced with the general problem of orthonormalizing g linearly independent eigenvectors $|\varphi_1\rangle, |\varphi_2\rangle, \cdots, |\varphi_g\rangle$, having the same eigenvalue a. This can be done using the Gram-Schmidt algorithm. The orthonormalized vectors will also be eigenvectors of A, with the eigenvalue a, because every linear combination of eigenvectors with the eigenvalue a has the same eigenvalue, as well.

- definitions:

$$
\begin{aligned}
|\chi_1\rangle &\underset{\text{def}}{=} |\varphi_1\rangle \\
|\chi_2\rangle &\underset{\text{def}}{=} c_{21}|\chi_1\rangle + |\varphi_2\rangle \\
|\chi_3\rangle &\underset{\text{def}}{=} c_{31}|\chi_1\rangle + c_{32}|\chi_2\rangle + |\varphi_3\rangle \\
&\vdots \\
|\chi_g\rangle &\underset{\text{def}}{=} c_{g1}|\chi_1\rangle + c_{g2}|\chi_2\rangle + \cdots + c_{g,g-1}|\chi_{g-1}\rangle + |\varphi_g\rangle \ .
\end{aligned}
$$

- observation:

  (i) every $|\chi_k\rangle$ $k = 1, \cdots, g$ is a linear combination of $|\varphi_1\rangle, \cdots, |\varphi_g\rangle$

  (ii) every $|\varphi_k\rangle$ $k = 1, \cdots, g$ is a linear combination of $|\chi_1\rangle, \cdots, |\chi_g\rangle$.

- orthogonalization:

$$
\begin{aligned}
\langle\chi_1|\chi_2\rangle &= c_{21}\||\chi_1\|^2 + \langle\chi_1|\varphi_2\rangle = 0 \\
\Rightarrow \quad c_{21} &= -\frac{\langle\chi_1|\varphi_2\rangle}{\||\chi_1\|^2} \\
\langle\chi_1|\chi_3\rangle &= c_{31}\||\chi_1\|^2 + \langle\chi_1|\varphi_3\rangle = 0 \\
\Rightarrow \quad c_{31} &= -\frac{\langle\chi_1|\varphi_3\rangle}{\||\chi_1\|^2} \\
\langle\chi_2|\chi_3\rangle &= c_{32}\||\chi_2\|^2 + \langle\chi_2|\varphi_3\rangle = 0 \\
\Rightarrow \quad c_{32} &= -\frac{\langle\chi_2|\varphi_3\rangle}{\||\chi_2\|^2} \quad \cdots \text{ etc. .}
\end{aligned}
$$

- normalization:

$$
|\psi_k\rangle = \frac{|\chi_k\rangle}{\||\chi_k\|} \quad k = 1, \cdots, g \quad \text{orthonormal vectors (see tut 8).}
$$

After this Gram-Schmidt orthonormalization procedure, all eigenvectors of $A = A^\dagger$ are mutually orthonormal. If all eigenvectors are included, they thus provide us with a complete $\left(\sum_{i=1}^{\infty} |\psi_i\rangle\langle\psi_i| = I\right)$ and orthonormal ($\langle\psi_i|\psi_j\rangle = \delta_{ij}$) set of vectors of $\mathcal{H}$.

3.4 Expectation values again

Let $|\psi\rangle$ be normalized as $\||\psi\|^2 = \langle\psi|\psi\rangle = \psi^\dagger\psi = 1$. We want to express the expectation value of a dynamical variable A ($\equiv$ Hermitean operator) of a quantum system in the state $|\psi\rangle$ ($\equiv$ vector) in terms of the Fourier coefficients of $|\psi\rangle$ and the matrix elements of A.

- general case:

  If the $|\psi_k\rangle$ ($k = 1, 2, \cdots$) form an arbitrary complete and orthonormal set of vectors of $\mathcal{H}$, the expectation value is

$$
\begin{aligned}
\overline{A} &\underset{\text{def}}{=} \langle\psi|A\psi\rangle = \sum_{i,k=1}^{\infty} c_i^* \, \langle\psi_i|A\psi_k\rangle \, c_k \\
&= \sum_{i,k=1}^{\infty} c_i^* \, A_{ik} \, c_k = \psi^\dagger A\psi \in \mathcal{R} \quad \text{with} \\
&\quad A_{ik} = \langle\psi_i|A\psi_k\rangle \quad \text{matrix element} \\
&\quad \text{and } c_k = \langle\psi_k|\psi\rangle \quad \text{Fourier coefficient.}
\end{aligned}
$$

- special case:

  If the $|\psi'_k\rangle$ ($k = 1, 2, \cdots$) are the complete and orthonormal(ized) set of the eigenvectors of $A = A^\dagger$, with $A|\psi'_k\rangle = a_k|\psi'_k\rangle$, and $a_k \in \mathcal{R}$ as $A = A^\dagger$, the expectation value in the state $|\psi\rangle$ is simply

$$
\begin{aligned}
\overline{A} &= \langle\psi|A\psi\rangle = \sum_{i,k=1}^{\infty} c_i'^* \, \langle\psi_i'|A\psi_k'\rangle \, c_k' \\
&= \sum_{i,k=1}^{\infty} c_i'^* \langle\psi_i'|a_k\psi_k'\rangle c_k' = \sum_{i,k=1}^{\infty} c_i'^* \, a_k \underbrace{\langle\psi_i'|\psi_k'\rangle}_{\delta_{ik}} c_k' \\
&= \sum_{i=1}^{\infty} |c_i'|^2 a_i \ \in \ \mathcal{R} \ .
\end{aligned}
$$

  Here, the $c_i' = \langle\psi_i'|\psi\rangle \in \mathcal{C}$ are the Fourier coefficients in the orthonormal(ized) basis of the eigenvectors $|\psi_i'\rangle$ of $A = A^\dagger$, and a_i are the corresponding (real) eigenvalues.

We can now interpret

$$p_i \underset{\text{def}}{=} |c_i'|^2 > 0\ , \quad \text{with} \quad \sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} |c_i'|^2 = \|\psi\| = 1,$$

as the probability for finding the quantum system in the eigenstate $|\psi_i'\rangle$, with eigenvalue a_i. Thus $\overline{A} \underset{\text{def}}{=} \sum_{i=1}^{\infty} |c_i'|^2 a_i = \sum_{i=1}^{\infty} p_i\ a_i$ is the weighted average of the eigenvalues or expectation value of the dynamical variable in the state $|\psi\rangle$ with $\|\psi\| = 1$. This probabilistic interpretation of the state vector $|\psi\rangle$ was first proposed by M. Born (Göttingen) in 1926, and it has subsequently been adopted in the "Copenhagen interpretation" of the state vector $|\psi\rangle$ of N. Bohr and W. Heisenberg in 1927.

3.5 Axioms of measurement

(i) The measurement of a dynamical variable $A = A^\dagger$ of a quantum system that is initially in the state $|\psi\rangle$, with norm $\|\psi\| = 1$, yields one of its (real) eigenvalues, say $a_k = a$, which may have a degeneracy $g_k = g$.

(ii) The probability for measuring this eigenvalue $a_k = a$ is

$$p(a) = \sum_{k;a_k=a} |c_k'|^2 = \sum_{k;a_k=a} |\langle\psi_k'|\psi\rangle|^2\ .$$

Thus, the measurement process is not deterministic. In fact, QM prevents us from knowing more than the probability of the various possible outcomes of the measurement.

(iii) Measuring the eigenvalue $a_k = a$, the system will "instantly" collapse into a state $|\psi''\rangle$ of the $g(a)$-dimensional subspace of $\mathcal{H}$ with eigenvalue $a_k = a$,

$$|\psi''\rangle = \sum_{k;a_k=a} c_k''|\psi_k'\rangle\ .$$

This state $|\psi''\rangle$ must have unit probability, i.e. $\|\psi''\| = \sum_{k;a_k=a} |c_k''|^2 = 1,$

if the probability is to be conserved during the measurement process.

remarks:

(i) In nature, nothing happens "instantly". However, so far, nobody has been able to measure the time scale of the collapse process of the state vector. In any case, this collapse time must be very short. The time scale must definitely be shorter than the electromagnetic time scale of $\sim 10^{-18}$ sec, but it could be between the nuclear time scale of $\sim 10^{-23}$ sec and the Planck time scale of $\sim 10^{-43}$ sec.

(ii) In order to pinpoint the exact $|\psi''\rangle$ in the subspace of $\mathcal{H}$, one would need to measure further complementary dynamical operators $A', A'', \cdots$. These must commute with each other and A, i.e.

$$[A', A] = [A'', A] = [A', A''] = \cdots = 0$$

so that they would not transcend the subspace of $\mathcal{H}$ created by the previous measurement(s) (see tut 7).

example:

$$\begin{aligned} [L^2, L_z] &= 0 \\ L^2|Y_{lm_l}\rangle &= l(l+1)|Y_{lm_l}\rangle \\ L_z|Y_{lm_l}\rangle &= m_l|Y_{lm_l}\rangle \text{ (see tut 11).} \end{aligned}$$

4 Quantization of a classical theory

Suppose we have a classical theory: how do we find the corresponding quantum theory? There are two possible methods:

(i) canonical quantization (P.A.M. Dirac, 1927)

(ii) path integral quantization (R.P. Feynman, 1948).
Involves new mathematics, i.e. path integrals.

We rely here on canonical quantization.

4.1 Canonical quantization

We proceed in five steps:

Step 1: Take a classical Lagrange function: e.g. one-dimensional motion of a particle with mass m in a time-independent potential energy $V(q)$

$$L = L(q,\dot{q}) = \underbrace{\tfrac{1}{2}\, m\dot{q}^2}_{\substack{\text{kinetic}\\ \text{energy}}} - \underbrace{V(q)}_{\substack{\text{potential}\\ \text{energy}}} .$$

L.V. de Lagrange (Frenchman, 1736-1813, $\sim$ 1755).

L. Euler (Swiss, 1707-1783, $\sim$ 1745).

Determine the momentum p that is canonically conjugate to the generalized coordinate q

$$p \underset{\text{def}}{=} \frac{\partial L}{\partial \dot{q}} = m\dot{q} \ .$$

Set up the Euler-Lagrange equation of motion

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q} \qquad \begin{aligned} \frac{\partial L}{\partial q} &= F \quad \text{generalized force} \\ \frac{\partial L}{\partial \dot{q}} &= p \quad \text{generalized momentum} \end{aligned}$$

$$\Rightarrow \quad \dot{p} = m\ddot{q} = -\frac{\partial V}{\partial q} = F \quad \text{Newton's second law:}$$

one differential equation second order in time for $q(t)$.
I. Newton (Englishman, 1643-1727), work done $\sim$ 1666, but not published until 1687 in the Principia.

Step 2: Evaluate the Hamilton function

$$H(p,q) \underset{\text{def}}{=} \dot{q}p - L(q,\dot{q})$$

theorem:

Hamilton function does not depend explicitly on $\dot{q}$

proof:

$$\frac{\partial}{\partial \dot{q}} H(p,q) = p - \frac{\partial L}{\partial \dot{q}} = 0 \quad \text{(keeping } p \text{ and } q \text{ fixed)} \qquad \text{q.e.d.}$$

consequence:

One has to express the Hamilton function completely in terms of the canonically conjugate variables, p and q, replacing $\dot{q}$ with p/m, i.e.

$$H(p,q) = \frac{p}{m}\, p - \frac{1}{2}\, m\, \left(\frac{p}{m}\right)^2 + V(q,t) = \frac{p^2}{2m} + V(q) \ .$$

example:

harmonic oscillator

$$\begin{aligned} H(p,q) &= \frac{p^2}{2m} + \frac{1}{2} m\omega^2\, q^2 \\ &= \frac{p^2}{2m} + \frac{1}{2}\, f\, q^2 \ . \end{aligned}$$

The Hamilton function

$$H(p,q) \underset{\text{def}}{=} \dot{q}p - L(q,\dot{q})$$

obeys two coupled differential equations of first order in time for $p(t)$ and $q(t)$

$$\left.\begin{aligned} \frac{\partial H}{\partial p} &= \dot{q} \\ \frac{\partial H}{\partial q} &= -\frac{\partial L}{\partial q} = -\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = -\dot{p} \end{aligned}\right\} \begin{array}{l}\text{Hamilton's}\\ \text{equations}\\ \text{of motion}\end{array} \ .$$

W.R. Hamilton (Irishman, 1805-1865, $\sim$ 1834).

example:

For the harmonic oscillator $H = \dfrac{p^2}{2m} + \dfrac{1}{2} m\omega^2 q^2$ we have

$$\frac{\partial H}{\partial p} = \frac{p}{m} = \dot{q} \quad , \quad \frac{\partial H}{\partial q} = m\omega^2 q = -\dot{p}$$

$$\text{or} \quad \left\{ \begin{aligned} m\ddot{q} &= -m\omega^2 q = -fq \quad \text{Newton's second law} \\ p &= m\dot{q} \qquad\qquad\qquad\quad \text{definition of momentum} \end{aligned} \right.$$

Step 3: Quantization

Following M. Born and P. Jordan, we replace the canonically conjugate variables, p and q, by non-commuting linear operators $\hat{p}$ and $\hat{q}$, satisfying the canonical quantization relations (in this section we use hats again to distinguish operators from numbers).

One degree of freedom:

$$[\hat{p}, \hat{q}] = i\hbar\, \hat{I}$$

$\uparrow$

first appearance of
i: imaginary unit
$\hbar$: Planck's constant
in the combination $i\hbar$

Born-Jordan quantization condition: M. Born, P. Jordan (1924)

Generalization to N degrees of freedom:

$$[\hat{p}_l, \hat{q}_k] = -i\hbar\, \delta_{lk}$$

$$[\hat{p}_l, \hat{p}_k] = 0 \qquad l, k = 1, 2, \cdots N$$

$$[\hat{q}_l, \hat{q}_k] = 0 \ .$$

In 1928, P.A.M. Dirac, W. Heisenberg and W. Pauli applied these canonical quantization conditions to Quantum Electrodynamics (QED), the relativistic quantum field theory that describes the electromagnetic interactions of the electrons, positrons and photons.

The quantization conditions can be fulfilled with

(a)	infinite matrices	$\Rightarrow$	matrix mechanics	(W. Heisenberg, 1925)
(b)	differential operators	$\Rightarrow$	wave mechanics	(E. Schrödinger, 1926)
(c)	abstract linear operators	$\Rightarrow$	Dirac's method	(P.A.M. Dirac, 1927).

The equivalence of the first two approaches was shown by W. Pauli in 1926.
P.A.M. Dirac showed, in 1927, that any linear mathematical objects, p and q, which obey the Born-Jordan commutation relations, will do the job.

Step 4: Replace Hamilton function by Hamilton operator

$$H(p, q) \to \hat{H}(\hat{p}, \hat{q}) \ .$$

caution: This recipe may be ambiguous, as the order of the operators, e.g. $\hat{p}\hat{q}$ or $\hat{q}\hat{p}$ matters in QM, in contrast to CM. However, this is not a problem in our example:

$$H(p, q) = \frac{p^2}{2m} + V(q) \qquad \text{Hamilton function}$$

$$\Rightarrow \quad \hat{H}(\hat{p}, \hat{q}) = \frac{\hat{p}^2}{2m} + V(\hat{q}) \qquad \text{Hamilton operator} \ .$$

Step 5: Formulate the equations of motion.

The dynamical equation of motion of QM is most conveniently formulated in the so-called Schrödinger picture, in which the operators $\hat{p}$ and $\hat{q}$ are time-independent, whereas the state vectors $|\psi(t)\rangle$ are time dependent. Between two measurements, $|\psi(t)\rangle$ evolves deterministically according to the time-dependent Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}(\hat{p}, \hat{q})\, |\psi(t)\rangle \ .$$

$\uparrow$

second appearance of $i\hbar$.

remark:

Canonical quantization is not manifestly Lorentz invariant, simply because it treats $\hat{q}$ as an operator, while t is merely a real parameter. However, when this procedure is applied to a Hamilton function describing a classical relativistic theory, it invariably produces a fully Lorentz invariant quantum theory. In fact, QM is not in conflict with Special Relativity (SR), but it is incompatible with General Relativity (GR). This is one of the reasons why an acceptable theory of Quantum Gravity has yet to be found.

5 Pictures

5.1 Unitary transformations

- Unitary transformations of vectors and operators in $\mathcal{H}$ play an important role in the formulation of the dynamical equations of motion of QM. These are similar to orthogonal transformations of vectors and tensors in $\mathcal{E}_3$. Vectors transform under unitary transformations as

$$|\psi'\rangle = U|\psi\rangle \ ,$$

where U is a unitary operator $U^\dagger = U^{-1}$. Unitary transformations in $\mathcal{H}$ are similar to orthogonal transformations in $\mathcal{E}_3$, because they also keep the scalar product invariant, i.e.

$$\begin{aligned} \langle\varphi'|\psi'\rangle &= \langle U\varphi|U\psi\rangle = \langle\varphi|U^\dagger U\psi\rangle \\ &= \langle\varphi|U^{-1}U\psi\rangle = \langle\varphi|I\psi\rangle = \langle\varphi|\psi\rangle \ . \end{aligned}$$

Thus the "lengths" of vectors and the "angles" between vectors, in short: the geometry in $\mathcal{H}$ will not change under unitary transformation.

- The transformation properties of operators can be obtained by requiring that the defining equation for operators,

$$|\varphi\rangle = F|\psi\rangle \leftrightarrow |\varphi'\rangle = F'|\psi'\rangle \ ,$$

be the same in the original and the transformed systems.

In the transformed system we have $|\varphi'\rangle = F'|\psi'\rangle$.

Substituting $|\varphi'\rangle = U|\varphi\rangle$ and $|\psi'\rangle = U|\psi\rangle$, we obtain

$$U|\psi\rangle = F' U|\psi\rangle \ .$$

Applying U^{-1} from the left, we arrive at

$$\begin{aligned} \underbrace{U^{-1}U}_{I}|\varphi\rangle &= U^{-1}F'U|\psi\rangle \\ \text{or} \qquad |\varphi\rangle &= U^{-1}F'U|\psi\rangle \ , \text{ valid for every } |\psi\rangle \in \mathcal{H}. \end{aligned}$$

As in the original system $|\varphi\rangle = F|\psi\rangle$ is also valid for every $|\psi\rangle \in \mathcal{H}$,

we must have $F = U^{-1}F'U$.

The corresponding inverse transformation is easily obtained, sandwiching F between U and U^{-1}

$$UFU^{-1} = \underbrace{UU^{-1}}_{I} F' \underbrace{UU^{-1}}_{I} = F' \ .$$

- summary: The transformation properties for vectors and operators are

	transformation	inverse transformation
vectors:	$\|\psi'\rangle = U\|\psi\rangle$	$\|\psi\rangle = U^{-1}\|\psi'\rangle$
operators:	$F' = U\,F\,U^{-1}$	$F = U^{-1}\,F'\,U$ (see tut 7(f)).

5.2 Schrödinger and Heisenberg pictures

- There is quite some freedom in choosing the time-dependence of the operators, $p(t)$, $q(t)$ and the state vector $|\psi(t)\rangle$. Such a consistent choice is called a picture. This is similar to the choice one has in Euler's classical dynamics of rotating rigid bodies, where one may prefer to formulate the equations of motion in the rotating, body-fixed system, rather than an inertial system, in order to have a constant moment of inertia. Of course, one has to pay a penalty in that the equations of motion in the body-fixed system will be different from those in an inertial system, because of the centrifugal and Coriolis forces.

We now focus again on the Hamilton operator description of one-dimensional motion of a particle of mass m moving in a time-independent potential energy $V(q)$,

$$H(p,q) = \frac{p^2}{2m} + V(q) \ .$$

In CM, $H(p,q)$ is a constant of motion. We will show later that this remains valid in QM. There are two special choices of pictures:

Schrödinger picture: $p_S, q_S, |\psi_S(t)\rangle$ E. Schrödinger (1926)

Heisenberg picture: $p_H(t), q_H(t), |\psi_H\rangle$ W. Heisenberg (1925) .

remark: It is not surprising that the Heisenberg picture was found first, because it is closest to the classical description with time-dependent variables $p(t)$ and $q(t)$.

- Schrödinger picture:

We start with the Schrödinger picture, because we all became acquainted with it first. The time-dependence of the state vector $|\psi_S(t)\rangle$ is given by

$$|\psi_S(t)\rangle = U(t)|\psi_S(0)\rangle \ ,$$

where $U(t) = \exp\left(-\frac{i}{\hbar}Ht\right)$ is the time-evolution operator.
In the Schrödinger picture, the operators p_S and q_S, and every operator function $F(p_S, q_S)$, are time-independent. In particular, the Hamilton operator $H(p_S, q_S) = H = \text{const}$.

theorem: $|\psi_S(t)\rangle$ satisfies the time-dependent Schrödinger equation

proof:

$$\begin{aligned} i\hbar \frac{d}{dt} |\psi_S(t)\rangle &= i\hbar \frac{d}{dt} U(t)|\psi_S(0)\rangle = i\hbar \frac{d}{dt} \exp\left(-\frac{i}{\hbar}Ht\right) |\psi_S(0)\rangle \\ &= i\hbar \left(-\frac{i}{\hbar}H\right) \underbrace{\exp\left(-\frac{i}{\hbar} Ht\right) |\psi_S(0)\rangle}_{|\psi_S(t)\rangle} . \end{aligned}$$

Here we have expanded the exponent in powers of H, differentiated and resummed the powers term by term again.

$i\hbar \frac{d}{dt} |\psi_S(t)\rangle = H|\psi_S(t)\rangle$ is indeed the time-dependent Schrödinger equation. q.e.d.

- Heisenberg picture:

We may choose the Heisenberg picture to coincide with the Schrödinger picture at $t = 0$. In this case, we have

$$|\psi_H\rangle = |\psi_S(0)\rangle \qquad \text{for the state vectors,}$$

$$\left.\begin{aligned} \text{and } p_H(0) &= p_S \\ q_H(0) &= q_S \\ F(p_H(0), q_H(0)) &= F(p_S, q_S) \end{aligned}\right\} \quad \text{for the operators.}$$

We can then write

$$|\psi_H\rangle = \underbrace{U^{-1}(t)\overbrace{U(t)}^{|\psi_S(t)\rangle}|\psi_S(0)\rangle}_{I} = U^{-1}(t)|\psi_S(t)\rangle ,$$

where $U(t) \underset{\text{def}}{=} \exp\left(-\frac{i}{\hbar} Ht\right)$ is the time-evolution operator.

We have shown in section 5.1 that

$$\begin{aligned} \text{if } |\psi\rangle &= U^{-1}|\psi'\rangle \quad \text{holds for state vectors,} \\ \text{then } F &= U^{-1}F'U \quad \text{must hold for operators.} \end{aligned}$$

thus:

$$\begin{aligned} \text{if } |\psi_H\rangle &= U^{-1}(t)|\psi_S(t)\rangle \\ \text{then } p_H(t) &= U^{-1}(t)\, p_S\, U(t) \\ \text{and } q_H(t) &= U^{-1}(t)\, q_S\, U(t) \\ \text{and in general } F(p_H(t), q_H(t)) &= U^{-1}(t)\, F(p_S, q_S)\, U(t) , \end{aligned}$$

with the initial conditions

$$\begin{aligned} p_H(0) &= p_S \\ q_H(0) &= q_S \\ \text{and } F(p_H(0), q_H(0)) &= F(p_S, q_S) . \end{aligned}$$

example:

$$H(p,q) = \frac{p^2}{2m} + V(q) \qquad \text{Hamilton operator}$$

$$U^{-1}(t)\, H(p_S, q_S)\, U(t) = U^{-1}(t) \underbrace{\left(\frac{p_S^2}{2m} + V(q_S)\right)}_{a_0 I + a_1 q_S + a_2 q_S^2 + \cdots} U(t) .$$

Inserting $U(t)\, U^{-1}(t) = I$ wherever appropriate, we obtain

$$\begin{aligned} U^{-1}(t)\, H(p_S, q_S)\, U(t) &= \frac{1}{2m} U^{-1}(t)\, p_S \overbrace{U(t)\, U^{-1}(t)}^{I} p_S\, U(t) \\ &\quad + a_0\, U^{-1}(t)\, I\, U(t) + a_1\, U^{-1}(t)\, q_S\, U(t) + \\ &\quad + a_2\, U^{-1}(t)\, q_S \underbrace{U(t)\, U^{-1}(t)}_{I} q_S\, U(t) + \cdots \\ &= \frac{1}{2m} p_H^2(t) + \underbrace{a_0\, I + a_1\, q_H(t) + a_2\, q_H^2(t) + \cdots}_{V(q_H(t))} \\ &= H(p_H(t), q_H(t)) . \end{aligned}$$

Here we have used $p_H(t) = U^{-1}(t)\, p_S\, U(t)$ and $q_H(t) = U^{-1}(t)\, q_S\, U(t)$.

remark: This also works for negative powers, e.g.

$$U(t)^{-1}\, q_S^{-1}\, U(t) = \left(U(t)^{-1}\, q_S\, U(t)\right)^{-1} = q_H(t)^{-1} .$$

theorem: $H(p_H(t), q_H(t)) = \text{const}$

proof:

$$U(t) \underset{\text{def}}{=} \exp\left(-\frac{i}{\hbar} Ht\right) = \exp\left(-\frac{i}{\hbar} H(p_S, q_S)t\right)$$

$$\Rightarrow \quad [U(t), H(p_S, q_S)] = \left[\exp\left(-\frac{i}{\hbar} H(p_S, q_S)t\right), H(p_S, q_S)\right] = O$$

Thus, as

$$U^{-1}(t)\, H(p_S, q_S)\, U(t) \;=\; \underbrace{U^{-1}(t)\, U(t)}_{I}\, H(p_S, q_S) = H(p_S, q_S) = H$$

and

$$U^{-1}(t)\, H\,(p_S, q_S)\; U(t) \;=\; H\,(p_H(t), q_H(t))\quad ,$$

we conclude that

$$H\,(p_H(t),\, q_H(t)) \;=\; H\,(p_S, q_S) \;=\; H \;=\; \text{const} \quad \text{q.e.d.}$$

- Heisenberg equation of motion

Abbreviating $F(t) \underset{\text{def}}{=} F\,(p(t), q(t))$, i.e. time-dependence of $F(t)$ is implicit through $p(t)$ and $q(t)$ only, we can write

$$F(t) = U^{-1}(t)\, F(0)\, U(t) \quad \text{with} \quad U = \exp\left(-\frac{i}{\hbar}\, Ht\right) \quad .$$

The time-derivative is therefore

$$\begin{aligned}
\frac{d}{dt}\, F(t) \;&=\; \frac{dU^{-1}(t)}{dt}\, F(0)\, U(t) + U^{-1}(t)\, F(0)\, \frac{dU(t)}{dt} \\
&=\; \frac{i}{\hbar}\, H\, \underbrace{U^{-1}(t)\, F(0)\, U(t)}_{F(t)} + \underbrace{U^{-1}(t)\, F(0)\, U(t)}_{F(t)}\, \left(-\frac{i}{\hbar}\, H\right) \\
&=\; \frac{i}{\hbar}\, H\, F(t) - \frac{i}{\hbar}\, F(t)\, H \\
&=\; \frac{i}{\hbar}\, [H, F(t)] = \frac{1}{i\hbar}\, [F(t), H] \quad .
\end{aligned}$$

We thus obtain Heisenberg's equation of motion

$$\frac{d}{dt}\, F\,(p_H(t), q_H(t)) \;=\; \frac{1}{i\hbar}\, [F\,(p_H(t), q_H(t))\,, H]$$

↑
third appearance of $i\hbar$

W. Heisenberg (1925)

- consequence:

The following four statements are causally connected:

(i) $\dfrac{d}{dt}\, F\,(p_H(t), q_H(t)) = 0$

(ii) $[F, H] = 0$

(iii) F and H have common eigenvectors (see tut 7)

(iv) F is a symmetry of H

special case: $\dfrac{d}{dt}\, H = \dfrac{1}{ih}\, [H, H] = 0$

- summary:

Schrödinger picture : $p_S, q_S = \text{const}$

$F(p_S, q_S) = \text{const}$

$i\hbar\, \dfrac{d}{dt}\, |\psi_S(t)\rangle = H|\psi_S(t)\rangle$

Schrödinger equation

Heisenberg picture : $p_H(t), q_H(t),\, |\psi_H\rangle = \text{const}$

$i\hbar\, \dfrac{d}{dt}\, F_H(t) = [F_H(t), H]$

Heisenberg's equation of motion

Later we will introduce the

Dirac or interaction picture : $p_D(t), q_D(t),\, |\psi_D(t)\rangle$

$F\,(p_D(t), q_D(t))$

see tut 29 and chapter 15

5.3 Matrix elements in the Heisenberg and Schrödinger pictures

- In the Schrödinger picture the state vector evolves as

$$|\psi_S(t)\rangle \;=\; U(t)|\psi_S(0)\rangle = \mathrm{e}^{-\frac{i}{\hbar}Ht}\,|\psi_S(0)\rangle$$

$$\text{with} \qquad H \underset{\text{def}}{=} H(p_S,q_S) = \frac{p_S^2}{2m} + V(q_S) = \text{const.}$$

Let $|\psi_k\rangle$ $(k{=}1,2,\cdots)$ be a complete and orthonormal(ized) set of eigenvectors of H, i.e.

$$H|\psi_n\rangle = E_n|\psi_n\rangle \ .$$

The initial state vector $|\psi_S(0)\rangle$ can thus be expanded in terms of the $|\psi_n\rangle$ as

$$|\psi_S(0)\rangle = \sum_{n=1}^{\infty} c_n(0)\,|\psi_n\rangle = |\psi_H\rangle \ ,$$

with Fourier coefficients $c_n(0) = \langle\psi_n|\psi_S(0)\rangle$. For the time-evolution of $|\psi_S(t)\rangle$ we obtain

$$\begin{aligned}
|\psi_S(t)\rangle &= \mathrm{e}^{-\frac{i}{\hbar}Ht}\sum_{n=1}^{\infty} c_n(0)\,|\psi_n\rangle \\
&= \sum_{n=1}^{\infty} c_n(0)\,\mathrm{e}^{-\frac{i}{\hbar}Ht}|\psi_n\rangle = \sum_{n=1}^{\infty} \underbrace{c_n(0)\mathrm{e}^{-\frac{i}{\hbar}E_nt}}_{c_n(t)}\,|\psi_n\rangle \\
&= \sum_{n=1}^{\infty} c_n(t)\,|\psi_n\rangle \ ,
\end{aligned}$$

$$\text{with} \qquad c_n(t) \;=\; c_n(0)\,\mathrm{e}^{-\frac{i}{\hbar}E_nt} \qquad \begin{array}{l}\text{time-dependent}\\ \text{Fourier coefficients.}\end{array}$$

The relation between the Fourier coefficients in the Schrödinger and Heisenberg pictures is thus given by

$$c_n(t) \;=\; \langle\psi_S(t)|\psi_n\rangle \;=\; \mathrm{e}^{-\frac{i}{\hbar}E_nt}\,\langle\psi_H|\psi_n\rangle \;=\; \mathrm{e}^{-\frac{i}{\hbar}E_nt}\,c_n(0) \ .$$

- In the Heisenberg picture, an operator depending on time only through the canonically conjugate variables $p(t)$ and $q(t)$, evolves as

$$F_H(t) = U^{-1}(t)\,F_H(0)\,U(t) = \mathrm{e}^{\frac{i}{\hbar}Ht}\,F_H(0)\,\mathrm{e}^{-\frac{i}{\hbar}Ht} \ .$$

The matrix elements of $F_H(t)$, in terms of the eigenvectors $|\psi_k\rangle$ of H, are therefore

$$\begin{aligned}
\langle\psi_i|F_H(t)\psi_k\rangle &= \langle\psi_i|\mathrm{e}^{\frac{i}{\hbar}Ht}\,F_H(0)\,\mathrm{e}^{-\frac{i}{\hbar}Ht}\,\psi_k\rangle \\
&= \langle\mathrm{e}^{-\frac{i}{\hbar}Ht}\psi_i|F_H(0)\,\mathrm{e}^{-\frac{i}{\hbar}Ht}\,\psi_k\rangle \\
&= \langle\mathrm{e}^{-\frac{i}{\hbar}E_it}\psi_i|F_H(0)\,\mathrm{e}^{-\frac{i}{\hbar}E_kt}\,\psi_k\rangle \\
&= \mathrm{e}^{\frac{i}{\hbar}(E_i-E_k)t}\,\langle\psi_i|F_H(0)\,\psi_k\rangle \ .
\end{aligned}$$

The relation between matrix elements in the Heisenberg and Schrödinger pictures is thus given by

$$(F_H(t))_{ik} \underset{\text{def}}{=} \langle\psi_i|F_H(t)\psi_k\rangle = \mathrm{e}^{\frac{i}{\hbar}(E_i-E_k)t}\,\langle\psi_i|F_S\,\psi_k\rangle \underset{\text{def}}{=} \mathrm{e}^{\frac{i}{\hbar}(E_i-E_k)t}\,(F_S)_{ik} \ ,$$

as $F_H(0) = F_S$ per definition.

6 Heisenberg's uncertainty principle

This principle, discovered by W. Heisenberg in 1927, is arguably the most innovative and central piece of QM. It is due to the non-commutativity of two operators. Here we derive it for two abstract operators *à la* Dirac.

Let A and B be non-commuting Hermitean operators $A = A^\dagger$, $B = B^\dagger$ with $[A,B] \neq 0$.

theorem:

$$\overline{(A-\overline{A}\,I)^2} \cdot \overline{(B-\overline{B}\,I)^2} \geqslant \left(\overline{\frac{i}{2}\,[A,B]}\right)^2 .$$

proof: Let us study the function $f(\lambda)$ of a real variable λ defined as

$$f(\lambda) \underset{\text{def}}{=} \overline{(A+i\lambda B)^\dagger\,(A+i\lambda B)} = \overline{(A^\dagger - i\lambda B^\dagger)\,(A+i\lambda B)}$$

$$= \overline{(A-i\lambda B)\,(A+i\lambda B)} = \overline{A^2} + \lambda^2\,\overline{B^2} + i\lambda\,\overline{[A,B]} .$$

$f(\lambda)$ must be real and non-negative, simply because it can also be written in the form

$$f(\lambda) = \overline{\langle\psi|F^\dagger F\psi\rangle} \quad \text{with} \quad F \underset{\text{def}}{=} A + i\lambda B .$$

Indeed, we have shown in tut 5 that $\langle\psi|F^\dagger F\psi\rangle = \langle F\psi|F\psi\rangle = \|F\psi\|^2 \geqslant 0$. Thus the polynomial of second degree in λ, $f(\lambda) = a\,\lambda^2 + b\,\lambda + c$, must be non-negative for <u>all</u> real values of λ.

The coefficients are

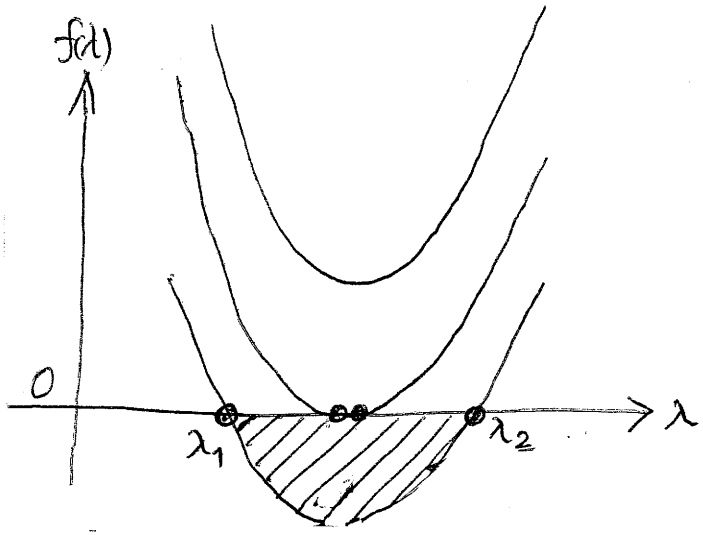
$$c \underset{\text{def}}{=} \overline{A^2} = \langle\psi|A^2\psi\rangle = \langle A^\dagger\psi|A\psi\rangle = \langle A\psi|A\psi\rangle \geqslant 0$$

$$\text{and} \quad a \underset{\text{def}}{=} \overline{B^2} \geqslant 0 .$$

$b = i\overline{[A,B]}$ must also be real, as $f(\lambda) = a\,\lambda^2 + b\,\lambda + c$ is real, but it can be positive, zero or negative. We are thus faced with the possible contradiction that $f(\lambda)$, as a real, non-negative polynomial of a real variable λ, might be negative for some values of λ. The zero's of $f(\lambda)$ are at

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} .$$

There are three possibilities shown here:



The shadowed area is unacceptable, because $f(\lambda) < 0$ for $\lambda_1 < \lambda < \lambda_2$.

The only way to avoid the contradiction $\|F\psi\|^2 < 0$ is to require that the discriminant be $b^2 - 4ac \leqslant 0$, or $4ac \geqslant b^2$, which results in

$$4\,\overline{A^2}\cdot\overline{B^2} \geqslant \overline{i\,[A,B]}^2 .$$

This inequality also holds for operators shifted by a real number,

$$A' \underset{\text{def}}{=} A - \overline{A}\,I \quad \text{and} \quad B' = B - \overline{B}\,I ,$$

because A' and B' are also Hermitean. Interestingly, the commutator does not change if we replace $A \to A'$ and $B \to B'$ as $\left[A - \overline{A}\,I\ ,\ B - \overline{B}\,I\right] = [A,B]$.
We thus conclude

$$\overline{(A-\overline{A}\,I)^2} \cdot \overline{(B-\overline{B}\,I)^2} \geqslant \left(\overline{\tfrac{i}{2}[A,B]}\right)^2$$

$$\text{or} \quad (\Delta A)^2 \cdot (\Delta B)^2 \geqslant \left(\overline{\tfrac{i}{2}[A,B]}\right)^2 .$$

Here we have defined the uncertainties in A and B squared as

$$(\Delta A)^2 \underset{\text{def}}{=} \overline{(A-\overline{A}\,I)^2} \text{ and } (\Delta B)^2 \underset{\text{def}}{=} \overline{(B-\overline{B}\,I)^2} .$$

example:

$$[p,q] = -i\hbar\, I$$

Born-Jordan quantization condition (1924)

$$\Rightarrow \quad (\Delta p)^2\,(\Delta q)^2 = \overline{(p-\overline{p}\,I)^2}\cdot\overline{(q-\overline{q}\,I)^2} \geqslant \left(\overline{\frac{i}{2}(-i\hbar\, I)}\right)^2 = \left(\frac{\hbar}{2}\right)^2 = \frac{\hbar^2}{4}\ .$$

Thus the uncertainty in p and q is

$$\Delta p\,\Delta q \geqslant \frac{\hbar}{2}\ . \qquad \text{W. Heisenberg (1927)}$$

conclusion:

One cannot measure p and q simultaneously with arbitrary precision.

- citations attributed to famous physicists:

  "You can look at the world with the p-eye or with the q-eye.
  But, if you want to look at it with both eyes you get crazy."
  W. Pauli

  "Nobody understands Quantum Mechanics"
  R.P. Feynman

  "God does not throw dice" A. Einstein

7 Linear harmonic oscillator

7.1 Abstract linear operators : Dirac's method

- P.A.M. Dirac (1927): Consider the harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{1}{2}\,fq^2 \qquad \text{Hamilton operator}$$

$$f = m\omega^2 \qquad \text{spring constant}$$

$p^\dagger = p$ and $q^\dagger = q$ are abstract linear Hermitean operators which obey the canonical commutation relation

$$[p,q] = -i\hbar\, I \qquad \text{Born-Jordan quantization condition}$$

- Dirac's brilliant idea: introduce step operators

$$B^\dagger \underset{\text{def}}{=} \frac{1}{\sqrt{2\hbar}}\left[\sqrt{m\omega}\,q - \frac{i}{\sqrt{m\omega}}\,p\right] \quad \text{raising}$$

$$B = \frac{1}{\sqrt{2\hbar}}\left[\sqrt{m\omega}\,q + \frac{i}{\sqrt{m\omega}}\,p\right] \quad \text{lowering}$$

- consequences:

$$BB^\dagger = \frac{1}{2\hbar}\left[m\omega q^2 + \frac{1}{m\omega}\,p^2 - iqp + ipq\right]$$

$$B^\dagger B = \frac{1}{2\hbar}\left[m\omega q^2 + \frac{1}{m\omega}\,p^2 - ipq + iqp\right]$$

difference:

$$[B,B^\dagger] \underset{\text{def}}{=} BB^\dagger - B^\dagger B = \frac{1}{2\hbar}\,2i\,[p,q] = \frac{i}{\hbar}\,(-i\hbar)\,I$$

$$\Rightarrow \quad [B,B^\dagger] = I \qquad \text{commutator}$$

sum:

$$\{B,B^\dagger\} \underset{\text{def}}{=} BB^\dagger + B^\dagger B = \frac{m\omega}{\hbar}\,q^2 + \frac{1}{m\hbar\omega}\,p^2$$

anti-commutator

- Hamilton operator:

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}\, m\omega^2 q^2 = \frac{1}{2}\,\hbar\omega\,\{B, B^\dagger\} \quad \text{anti-commutator} \\ &= \frac{1}{2}\,\hbar\omega\,\left(BB^\dagger + B^\dagger B\right) = \frac{1}{2}\,\hbar\omega\,\left(B^\dagger B + I + B^\dagger B\right) \\ H &= \hbar\omega\,\left(B^\dagger B + \tfrac{1}{2} I\right) \end{aligned}$$

(i) $B^\dagger B$ Hermitean $\Rightarrow$ eigenvalues are real, eigenvectors to different eigenvalues are orthogonal (see section 3.2)

(ii) $\langle\psi|B^\dagger B\psi\rangle = \langle B\psi|B\psi\rangle = \|B\psi\|^2 \geqslant 0$
eigenvalues are non-negative (see tut 5)

- Let $|\psi_\lambda\rangle$ be eigenvector of $B^\dagger B$ with eigenvalue λ

$$B^\dagger B|\psi_\lambda\rangle = \lambda|\psi\rangle$$

assumptions:

(i) λ not degenerate

(ii) $|\psi_\lambda\rangle \in \mathcal{H}$ with $\langle\psi_\lambda|\psi_\lambda\rangle = \|\psi_\lambda\|^2$
non-trivial and normalizable, i.e. $0 < \|\psi_\lambda\| < \infty$.

- theorem: $B^\dagger|\psi_\lambda\rangle$ is eigenvector of $B^\dagger B$ with eigenvalue $\lambda + 1$

proof:

$$\begin{aligned} [B, B^\dagger] &= BB^\dagger - B^\dagger B = I \\ BB^\dagger &= I + B^\dagger B \\ B^\dagger B\, B^\dagger|\psi_\lambda\rangle &= B^\dagger(I + B^\dagger B)|\psi_\lambda\rangle \\ &= B^\dagger(1+\lambda)|\psi_\lambda\rangle \\ &= (\lambda+1)\, B^\dagger|\psi_\lambda\rangle \qquad \text{q.e.d.} \end{aligned}$$

- theorem: $B|\psi_\lambda\rangle$ is eigenvector of $B^\dagger B$ with eigenvalue $\lambda - 1$

proof:

$$\begin{aligned} & BB^\dagger - B^\dagger B = I \\ \Rightarrow \quad & B^\dagger B = BB^\dagger - I \end{aligned}$$

$$\begin{aligned} B^\dagger B\, B|\psi_\lambda\rangle &= (BB^\dagger - I)B|\psi_\lambda\rangle \\ &= (B\, B^\dagger B - BI)|\psi_\lambda\rangle \\ &= B\,(B^\dagger B - I)|\psi_\lambda\rangle = B(\lambda-1)|\psi_\lambda\rangle \\ &= (\lambda-1)\, B|\psi_\lambda\rangle \qquad \text{q.e.d.} \end{aligned}$$

- summary:

$$\begin{aligned} B^\dagger B\ \left(B^\dagger\right)^n|\psi_\lambda\rangle &= (\lambda+n)\left(B^\dagger\right)^n|\psi_\lambda\rangle \\ &\vdots \qquad\qquad \vdots \\ B^\dagger B\ B^\dagger|\psi_\lambda\rangle &= (\lambda+1)\, B^\dagger|\psi_\lambda\rangle \\ B^\dagger B\ |\psi_\lambda\rangle &= \lambda\,|\psi_\lambda\rangle \\ B^\dagger B\ B|\psi_\lambda\rangle &= (\lambda-1)\, B\,|\psi_\lambda\rangle \\ &\vdots \qquad\qquad \vdots \\ B^\dagger B\ B^n|\psi_\lambda\rangle &= (\lambda-n)\, B^n\,|\psi_\lambda\rangle \end{aligned}$$

} discussed

- normalization:

$$\begin{aligned} \langle B^n\psi_\lambda|B^\dagger B\, B^n\,\psi_\lambda\rangle &= (\lambda-n)\,\langle B^n\,\psi_\lambda|B^n\psi_\lambda\rangle \\ &= (\lambda-n)\,\|B^n\psi_\lambda\|^2 \\ = \langle B^{n+1}\psi_\lambda|B^{n+1}\psi_\lambda\rangle &= \|B^{n+1}\psi_\lambda\|^2 \\ &\quad \text{recursion relation} \end{aligned}$$

$$\|B^{n+1}\psi_\lambda\|^2 = (\lambda-n)(\lambda-n+1)\,\cdots\,(\lambda-1)\lambda\,\underbrace{\|\psi_\lambda\|^2}_{1}$$

- problem:

For sufficiently large n, we will end up with $\|B^{n+1}\psi_\lambda\|^2 \leqslant 0$. However, negative probability or "ghost" states are unacceptable in $\mathcal{H}$, as $\|\psi\|^2 \geqslant 0$ for every $|\psi\rangle \in \mathcal{H}$. The only way round this contradiction, is to set $\lambda = n$, yielding $\|B^{n+1}\psi_\lambda\|^2 = 0$. We then have $B^\dagger B|\psi_n\rangle = n|\psi_n\rangle$ and $\|B^{n+1}\psi_n\|^2 = \|B^{n+2}\psi_n\|^2 = \cdots = 0$.

- definition: The state with lowest eigenvalue and unit norm is called the groundstate

$$|\psi_0\rangle \underset{\text{def}}{=} \frac{B^n|\psi_n\rangle}{\|B^n\psi_n\|} \qquad \text{with } \|\psi_0\| = 1$$

- properties:

(i) $$B|\psi_0\rangle = B\,\frac{B^n|\psi_n\rangle}{\|B^n\psi_n\|} = \frac{B^{n+1}|\psi_n\rangle}{\|B^n\psi_n\|} = |0\rangle$$

because $\|B^{n+1}\psi_n\| = 0$

$\Rightarrow$ "B annihilates the ground state $|\psi_0\rangle$"

(ii) $$B^\dagger B|\psi_0\rangle = B^\dagger|0\rangle = |0\rangle = 0|\psi_0\rangle$$

$\Rightarrow$ $|\psi_0\rangle$ is an eigenvector of $B^\dagger B$ to the eigenvalue zero.

(iii) The excited states are given by $|\psi_n\rangle = c_n \left(B^\dagger\right)^n |\psi_0\rangle$.

- We now determine the normalization constant c_n. We want $\|\psi_n\| = 1$ for $n \geqslant 1$. So let us first evaluate

$$\begin{aligned}
\| \left(B^\dagger\right)^n \psi_0\|^2 &\underset{\text{def}}{=} \langle \left(B^\dagger\right)^n \psi_0 | \left(B^\dagger\right)^n \psi_0\rangle \\
&= \langle BB^\dagger \left(B^\dagger\right)^{n-1} \psi_0 | \left(B^\dagger\right)^{n-1} \psi_0\rangle \\
&= \langle \left(I + B^\dagger B\right) \left(B^\dagger\right)^{n-1} \psi_0 | \left(B^\dagger\right)^{n-1} \psi_0\rangle \\
&= (1 + n - 1)\langle \left(B^\dagger\right)^{n-1} \psi_0 | \left(B^\dagger\right)^{n-1} \psi_0\rangle \\
&= n\| \left(B^\dagger\right)^{n-1} \psi_0\|^2 = n(n-1)\, \| \left(B^\dagger\right)^{n-2} \psi_0\|^2 \\
&= n(n-1) \cdots 1 \underbrace{\|\psi_0\|^2}_{=1} = n!
\end{aligned}$$

Thus $$\|\psi_n\|^2 = |c_n|^2 \| \left(B^\dagger\right)^n \psi_0\|^2 = |c_n|^2 n! = 1$$

$$\begin{aligned}
\Rightarrow \quad |c_n| &= \frac{1}{\sqrt{n!}} \\
c_n &= \frac{1}{\sqrt{n!}} \exp\left(i\varphi_n\right) \qquad \varphi_n : \text{ arbitrary phase.}
\end{aligned}$$

Here we have the freedom of fixing this phase φ_n for every excited state differently. However, the simplest phase convention is $\varphi_n = 0 \;\Rightarrow\; c_n = \frac{1}{\sqrt{n!}}$

- summary:

(i) The normalized excited states are given by

$$|\psi_n\rangle = \frac{1}{\sqrt{n!}} \left(B^\dagger\right)^n |\psi_0\rangle \qquad n \in \mathcal{N} \ .$$

(ii) The groundstate is defined as $B|\psi_0\rangle = |0\rangle$ with $\|\psi_0\| = 1$.

(iii) The eigenvalues and eigenvectors of H are given by

$$\begin{aligned}
H\,|\psi_n\rangle &= \hbar\omega \left(B^\dagger B + \tfrac{1}{2} I\right) |\psi_n\rangle \\
&= E_n|\psi_n\rangle = \hbar\omega(n + \tfrac{1}{2})|\psi_n\rangle \ , \qquad n \in \mathcal{N}_0 \ .
\end{aligned}$$

(iv) The zero-point energy $E_0 = \frac{1}{2}\hbar\omega$ is a consequence of the uncertainty principle, as the uncertainties Δp and Δq must obey $\Delta p \cdot \Delta q \geqslant \hbar/2$, even in the ground state of the harmonic oscillator (see tut 15).

(v) theorem:

$$\left.\begin{aligned}
B|\psi_n\rangle &= \sqrt{n}|\psi_{n-1}\rangle \quad \text{lowering} \\
B^\dagger|\psi_n\rangle &= \sqrt{n+1}|\psi_{n+1}\rangle \quad \text{raising}
\end{aligned}\right\} \text{ ladder or step operators}$$

proof:

$$\begin{aligned}
B^\dagger|\psi_n\rangle &= B^\dagger \frac{1}{\sqrt{n!}} \left(B^\dagger\right)^n |\psi_0\rangle \\
&= \frac{\sqrt{n+1}}{\sqrt{(n+1)!}} \left(B^\dagger\right)^{n+1} |\psi_0\rangle
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad B^\dagger|\psi_n\rangle &= \sqrt{n+1}|\psi_{n+1}\rangle \\
BB^\dagger|\psi_n\rangle &= \sqrt{n+1}\, B|\psi_{n+1}\rangle \\
(B^\dagger B + I)|\psi_n\rangle &= \sqrt{n+1}\, B|\psi_{n+1}\rangle \\
(n+1)|\psi_n\rangle &= \sqrt{n+1}\, B|\psi_{n+1}\rangle \\
B\,|\psi_{n+1}\rangle &= \sqrt{n+1}|\psi_n\rangle
\end{aligned}$$

replace n by $n-1$

$\Rightarrow$ $B\,|\psi_n\rangle = \sqrt{n}|\psi_{n-1}\rangle$ q.e.d.

7.2 Matrices in the Schrödinger and Heisenberg pictures

- Schrödinger picture:

We can now get a matrix representation for B and $B^\dagger$ and, therefore, also for p and q, in terms of the complete and orthonormal set of the eigenvectors of $B^\dagger B$.

$$\begin{aligned}\left(B^\dagger B\right)_{kl} &= \langle\psi_k|B^\dagger B\psi_l\rangle \\ &= l\langle\psi_k|\psi_l\rangle = k\delta_{kl} \qquad (k,l=0,1,2,3,\,\cdots)\end{aligned}$$

note: The matrix and vector indices start here with zero for convenience.

$$B^\dagger\,B = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 2 & 0 & \\ 0 & 0 & 0 & 3 & \cdots \\ \vdots & & \vdots & & \ddots \end{pmatrix}$$

The eigenvectors of $B^\dagger B$ are

$$|\psi_0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad , \quad |\psi_1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \quad , \cdots$$

- We now would like to find matrix representations for $B^\dagger$ and B

$$\begin{aligned}\left(B^\dagger\right)_{kl} &\underset{\text{def}}{=} \langle\psi_k\,\underbrace{|B^\dagger\psi_l\rangle}_{\sqrt{l+1}|\psi_{l+1}\rangle} = \underbrace{\langle\psi_k|\psi_{l+1}\rangle}_{\delta_{k,l+1}}\sqrt{l+1} \\ &= \delta_{k,l+1}\,\sqrt{k} \qquad (k,l=0,1,2,\,\cdots)\end{aligned}$$

examples:

$\underline{k=1,\ l=0}$ $\quad\Rightarrow\ \left(B^\dagger\right)_{kl} = \left(B^\dagger\right)_{10} = \sqrt{1}$

$\underline{k=2,\ l=1}$ $\quad\Rightarrow\ \left(B^\dagger\right)_{kl} = \left(B^\dagger\right)_{21} = \sqrt{2}$

$$B^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \sqrt{1} & 0 & 0 & 0 & \\ 0 & \sqrt{2} & 0 & 0 & \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & & \vdots & & \ddots \end{pmatrix}$$

$$B = B^{\dagger\dagger} = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \\ 0 & 0 & 0 & \sqrt{3} & \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & & \vdots & & \ddots \end{pmatrix}$$

- The matrices for p and q can be obtained from

$$\begin{cases} B &\underset{\text{def}}{=} \dfrac{1}{\sqrt{2\hbar}}\left[\sqrt{m\omega}\,q + \dfrac{i}{\sqrt{m\omega}}\,p\right] \\ B^\dagger &= \dfrac{1}{\sqrt{2\hbar}}\left[\sqrt{m\omega}\,q - \dfrac{i}{\sqrt{m\omega}}\,p\right] \end{cases}$$

yielding

$$\begin{cases} q &= \sqrt{\dfrac{\hbar}{2m\omega}}\,\left(B^\dagger + B\right) \\ p &= i\sqrt{\dfrac{m\hbar\omega}{2}}\,\left(B^\dagger - B\right) \end{cases}$$

remarks:

(i) p and q are Hermitean.

(ii) p and q satisfy the Born-Jordan quantization condition.

- Heisenberg picture:
  As $B, B^\dagger, p$ and q are time-independent, these matrices are given in the Schrödinger picture. The time-dependent matrices, expressed in the Heisenberg picture, can be obtained through the transformations

$$\begin{aligned} B(t) &= U^{-1}(t)\, B\, U(t) \quad \text{with } U(t) = \exp\left(-\tfrac{i}{\hbar} H t\right) \\ B^\dagger(t) &= U^{-1}(t)\, B^\dagger\, U(t) \\ p(t) &= U^{-1}(t)\, p\, U(t) \\ q(t) &= U^{-1}(t)\, q\, U(t) \end{aligned}$$

$$\begin{aligned} \langle \psi_m | B(t) \psi_n \rangle &= \langle \psi_m | \mathrm{e}^{\frac{i}{\hbar} H t}\, B\, \mathrm{e}^{-\frac{i}{\hbar} H t}\, \psi_n \rangle \\ &= \langle \mathrm{e}^{-\frac{i}{\hbar} H t}\, \psi_m | B\, \mathrm{e}^{-\frac{i}{\hbar} H t}\, \psi_n \rangle \\ &= \mathrm{e}^{\frac{i}{\hbar}(E_m - E_n)\, t}\, \langle \psi_m | B \psi_n \rangle \\ &= \mathrm{e}^{\frac{i}{\hbar}(E_m - E_n)\, t}\, \sqrt{n}\, \delta_{m,n-1} = \mathrm{e}^{\frac{i}{\hbar}(-\hbar\omega)\, t}\, \sqrt{n}\, \delta_{m,n-1} \\ &= \mathrm{e}^{-i\omega t}\, \sqrt{n}\, \delta_{m,n-1} \end{aligned}$$

$$\begin{aligned} m = 0, \quad n = 1 &\Rightarrow \sqrt{1}\, \mathrm{e}^{-i\omega t} \\ m = 1, \quad n = 2 &\Rightarrow \sqrt{2}\, \mathrm{e}^{-i\omega t} \end{aligned}$$

$$B(t) = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \\ & & & \ddots \end{pmatrix} \mathrm{e}^{-i\omega t}$$

$$B^\dagger(t) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \sqrt{1} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \\ & & & \ddots \end{pmatrix} \mathrm{e}^{i\omega t}$$

$$\Rightarrow \quad p(t) = i \sqrt{\frac{m\hbar\omega}{2}}\, \left(B^\dagger(t) - B(t)\right) = i \sqrt{\frac{m\hbar\omega}{2}}\, \left(B^\dagger\, \mathrm{e}^{i\omega t} - B\, \mathrm{e}^{-i\omega t}\right)$$

$$= i \sqrt{\frac{m\hbar\omega}{2}} \begin{pmatrix} 0 & -\sqrt{1}\mathrm{e}^{-i\omega t} & 0 & 0 \\ \sqrt{1}\mathrm{e}^{i\omega t} & 0 & -\sqrt{2}\mathrm{e}^{-i\omega t} & 0 \\ 0 & \sqrt{2}\mathrm{e}^{i\omega t} & 0 & -\sqrt{3}\mathrm{e}^{-i\omega t} \\ 0 & 0 & \sqrt{3}\mathrm{e}^{i\omega t} & 0 \\ & & & \ddots \end{pmatrix}$$

$$q(t) = \sqrt{\frac{\hbar}{2m\omega}}\, \left(B^\dagger(t) + B(t)\right) = \sqrt{\frac{\hbar}{2m\omega}}\, \left(B^\dagger\, \mathrm{e}^{i\omega t} + B\, \mathrm{e}^{-i\omega t}\right)$$

$$= \sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & \sqrt{1}\mathrm{e}^{-i\omega t} & 0 & 0 \\ \sqrt{1}\mathrm{e}^{i\omega t} & 0 & \sqrt{2}\mathrm{e}^{-i\omega t} & 0 \\ 0 & \sqrt{2}\mathrm{e}^{i\omega t} & 0 & \sqrt{3}\mathrm{e}^{-i\omega t} \\ 0 & 0 & \sqrt{3}\mathrm{e}^{i\omega t} & 0 \\ & & & \ddots \end{pmatrix} .$$

remark: $p_S = p(0)$, $p_H(t) = p(t)$, $q_S = q(0)$, $q_H(t) = q(t)$, compare with section 5.3.

7.3 Wave functions in coordinate and momentum space

- If we decide to represent $\hat{p}$ and $\hat{q}$ by differential operators, we still have a further choice (in this section we again use hats to distinguish operators from numbers) for the representation of the operators in

coordinate space:	or	momentum space: (see tut 9)
$\hat{p} \to -i\hbar \dfrac{\partial}{\partial q}$		$\hat{p} \to p$
$\hat{q} \to q$		$\hat{q} \to i\hbar \dfrac{\partial}{\partial p}$

- The differential operators are supposed to act on complex wave functions

coordinate space:	momentum space:
$\psi(q,t)$ with $\displaystyle\int_{-\infty}^{\infty} \lvert\psi(q,t)\rvert^2\, dq = 1$	$a(p,t)$ with $\displaystyle\int_{-\infty}^{\infty} \lvert a(p,t)\rvert^2\, dp = 1$ (see tut 9)

interpretation:

$\lvert\psi(q,t)\rvert^2\, dq$	$\lvert a(p,t)\rvert^2\, dp$
probability for finding particle in the coordinate interval dq around q	probability for finding particle in the momentum interval dp around p (see tut 9)

- The Born-Jordan quantization conditions are satisfied in both representations

coordinate space:

$$[\hat{p},\hat{p}]\,\psi(q,t) = (-i\hbar)^2\left[\frac{\partial}{\partial q},\frac{\partial}{\partial q}\right]\psi(q,t) = (-i\hbar)^2\left(\frac{\partial^2}{\partial q^2}-\frac{\partial^2}{\partial q^2}\right)\psi(q,t) = 0\cdot\psi(q,t)$$

$$[\hat{q},\hat{q}]\,\psi(q,t) = [q,q]\,\psi(q,t) = \left(q^2-q^2\right)\psi(q,t) = 0\cdot\psi(q,t)$$

$$[\hat{p},\hat{q}]\,\psi(q,t) = -i\hbar\left[\frac{\partial}{\partial q},q\right]\psi(q,t) = -i\hbar\left[\frac{\partial}{\partial q}\left(q\psi(q,t)\right)-q\frac{\partial}{\partial q}\psi(q,t)\right]$$
$$= -i\hbar\,\psi(q,t) \quad \text{q.e.d.}$$

momentum space (see tut 9):

$$[\hat{p},\hat{p}]\;a(p,t) = [p,p]\;a(p,t) = \left(p^2-p^2\right)\;a(p,t) = 0\cdot a(p,t)$$

$$[\hat{q},\hat{q}]\;a(p,t) = (i\hbar)^2\left[\frac{\partial}{\partial p},\frac{\partial}{\partial p}\right]a(p,t) = -\hbar^2\left(\frac{\partial^2}{\partial p^2}-\frac{\partial^2}{\partial p^2}\right)a(p,t) = 0\cdot a(p,t)$$

$$[\hat{p},\hat{q}]\;a(p,t) = i\hbar\left[p,\frac{\partial}{\partial p}\right]a(p,t) = i\hbar\left[p\,\frac{\partial}{\partial p}\,a(p,t)-\frac{\partial}{\partial p}\,\left(p\,a(p,t)\right)\right]$$
$$= -i\hbar\,a(p,t) \quad \text{q.e.d.}$$

- For the harmonic oscillator problem, we can easily get the wave function, e.g. in coordinate space representation, introducing Dirac's step operators

$$\hat{q} \;\to\; q\,, \quad \hat{p}\to i\hbar\,\frac{\partial}{\partial q}$$

$$B \underset{\text{def}}{=} \frac{1}{\sqrt{2\hbar}}\left[\sqrt{m\omega}\,q+\frac{i}{\sqrt{m\omega}}\left(-i\hbar\,\frac{\partial}{\partial q}\right)\right]$$

$$B^\dagger \underset{\text{def}}{=} \frac{1}{\sqrt{2\hbar}}\left[\sqrt{mw}\,q-\frac{i}{\sqrt{m\omega}}\left(-i\hbar\,\frac{\partial}{\partial q}\right)\right]$$

- The groundstate in the coordinate representation is defined as

$$B|\psi_0\rangle \underset{\text{def}}{=} \frac{1}{\sqrt{2\hbar}}\left[\sqrt{m\omega}\,q+\frac{\hbar}{\sqrt{m\omega}}\,\frac{\partial}{\partial q}\right]\psi_0(q) = 0$$

$$\left[\frac{m\omega}{\hbar}\,q+\frac{\partial}{\partial q}\right]\psi_0(q) = 0$$

$$\frac{\partial}{\partial q}\,\psi_0(q) = -\frac{m\omega}{\hbar}\,q\,\psi_0(q)$$

solution: $\psi_0(q) = A\exp\left(-\frac{m\omega}{2\hbar}\,q^2\right)$

normalization:

$$\int_{-\infty}^{\infty}|\psi_0(q)|^2dq = 1 \;\Rightarrow\; A = \sqrt[4]{\frac{m\omega}{\hbar\pi}} \quad \text{(see tut 15)}$$

- The normalized wave functions of the excited states are therefore

$$\psi_n(q) = \frac{1}{\sqrt{n}}\,\left(B^\dagger\right)^n\,\psi_0(q)$$
$$= \frac{1}{\sqrt{n}}\left[\frac{1}{\sqrt{2\hbar}}\left(\sqrt{m\omega}\,q-\frac{\hbar}{\sqrt{m\omega}}\,\frac{\partial}{\partial q}\right)\right]^n\,\psi_0(q)\;,$$

which generate the wave functions $\psi_n(q)$ (see tut 15). One can get the wave functions of the groundstate and excited states in momentum space $a_n(p)$ in a similar fashion.

- summary:

There are many ways of writing the time-dependent Schrödinger equation and its most general solution. We have shown in the case of abstract linear operators and state vectors that

$$H(\hat{p},\hat{q})|\psi(t)\rangle = \left[\frac{\hat{p}^2}{2m}+V(\hat{q})\right]|\psi(t)\rangle = i\hbar\,\frac{d}{dt}\,|\psi(t)\rangle$$

$$|\psi(t)\rangle = \sum_{n=1}^{\infty}c_n|\psi_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_nt}$$

with $\quad H(\hat{p},\hat{q})|\psi_n\rangle = E_n|\psi_n\rangle.$

We can translate this result to infinite square and column matrices

$$H(p,q)\psi(t) = \left[\frac{p^2}{2m} + V(q)\right]\psi(t) = i\hbar\,\frac{d}{dt}\,\psi(t)$$

$$\psi(t) = \sum_{n=1}^{\infty} c_n\,\psi_n\,\mathrm{e}^{-\frac{i}{\hbar}E_n t}$$

with $H(p,q)\,\psi_n = E_n\,\psi_n$, and to differential operators and wave functions in coordinate space

$$H\left(-i\hbar\frac{\partial}{\partial q}, q\right)\,\psi(q,t) = \left[-\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial q^2} + V(q)\right]\,\psi(q,t) = i\hbar\,\frac{\partial}{\partial t}\,\psi(q,t)$$

$$\text{with the solution } \psi(q,t) = \sum_{n=1}^{\infty} c_n\,\psi_n\,(q)\,\mathrm{e}^{-\frac{i}{\hbar}E_n t}$$

and the time-independent Schrödinger equation

$$H\left(-i\hbar\,\frac{\partial}{\partial q}, q\right)\psi_n(q) = E_n\psi_n(q) \ .$$

Finally, in terms of differential operators and wave functions in momentum space, the time-dependent Schrödinger equation is given by

$$H\left(p, i\hbar\,\frac{\partial}{\partial p}\right)\,a(p,t) = \left[\frac{p^2}{2m} + V\left(i\hbar\,\frac{\partial}{\partial p}\right)\right]\,a(p,t) = i\hbar\,\frac{\partial}{\partial t}\,a(p,t) \ ,$$

$$\text{with } \ V\left(i\hbar\frac{\partial}{\partial p}\right) = a_0 + a_1\,i\hbar\,\frac{\partial}{\partial p} + a_2\,\left(i\hbar\,\frac{\partial}{\partial p}\right)^2 + \cdots \ .$$

The most general solution of the time-dependent Schrödinger equation is

$$a(p,t) = \sum_{n=1}^{\infty} c_n\,a_n\,(p)\,\mathrm{e}^{-\frac{i}{\hbar}E_n t} \ ,$$

where $a_n(p)$ fulfils the time-independent Schrödinger equation

$$H\left(p, i\hbar\,\frac{\partial}{\partial p}\right)\,a_n\,(p) = E_n\,a_n\,(p) \ .$$

remark: One usually prefers the coordinate space representation over the momentum space representation, simply in order to avoid differential equations of infinite order in $\frac{\partial}{dq}$ for a general potential. However, in the case of the harmonic oscillator, both representations are equally acceptable.

7.4 Dictionary of the mathematical tools of Quantum Mechanics

As a variety of mathematical tools describe the same quantum physics, there must be a one-to-one correspondence linking these basic tools:

- vectors

$$|\psi\rangle \ \leftrightarrow \ \psi(q) \ \leftrightarrow \ \psi = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} \ .$$

- normalization of vectors

$$\|\psi\|^2 = \int_{-\infty}^{\infty} |\psi(q)|^2 dq = \sum_{n=1}^{\infty} |c_n|^2 = \psi^\dagger\psi < \infty \ .$$

  remark: The square integrable complex wave function $\psi(q)$ is the analogue of the square summable complex Fourier coefficient c_n, the only difference being that the index $q \in \mathcal{R}$ is continuous, while the index $n \in \mathcal{N}$ is discrete.

- sum of two vectors

$$|\varphi\rangle + |\psi\rangle \ \leftrightarrow \ \varphi(q) + \psi(q) \ \leftrightarrow \ \varphi + \psi$$

$$\|\varphi + \psi\| < \infty \quad \text{if} \quad \|\varphi\| < \infty \quad \text{and} \quad \|\psi\| < \infty \ .$$

- multiplication with a complex number

$$a\,|\psi\rangle \ \leftrightarrow \ a\,\psi(q) \ \leftrightarrow \ a\psi$$

$$\|a\psi\| < \infty \ \text{ if } \ \|\psi\| < \infty \ .$$

- linear independence: If the equations

$$\begin{aligned} z_1\,|\psi_1\rangle \ &+ \ \cdots + z_N|\psi_N\rangle = |0\rangle \\ z_1\,\psi_1(q) \ &+ \ \cdots + z_N\,\psi_N(q) = 0 \quad (\text{valid for all } q \in \mathcal{R}) \\ z_1\,\psi_1 \ &+ \ \cdots + z_N\,\psi_N = 0 \ . \end{aligned}$$

  only hold for $z_1 = \cdots = z_N = 0$, the set of N state vectors, wave functions and column matrices are linearly independent.

- scalar product

$$\langle\varphi|\psi\rangle = \int_{-\infty}^{\infty} \varphi(q)^*\,\psi(q)\,dq = \sum_{n=1}^{\infty} b_n^*\,c_n = \varphi^\dagger\psi \ .$$

  remark:

  Here again, we observe the analogy between the product of the wave functions integrated over $q\in\mathcal{R}$, with the product of Fourier coefficients summed over $n\in\mathcal{N}$.

- orthogonality

$$\langle\varphi|\psi\rangle = \int_{-\infty}^{\infty} \varphi(q)^*\ \psi(q)\ dq = \sum_{n=1}^{\infty} b_n^*\ c_n = \varphi^\dagger\psi = 0$$

with $\|\varphi\| \neq 0$ and $\|\psi\| \neq 0$.

- orthonormality of a set of vectors

$$|\psi_1\rangle, |\psi_2\rangle, \cdots \leftrightarrow \psi_1(q), \psi_2(q), \cdots \leftrightarrow \psi_1, \psi_2, \cdots$$

orthonormal iff

$$\langle\psi_i|\psi_k\rangle = \int_{-\infty}^{\infty} \psi_i^*(q)\ \psi_k(q)\ dq = \psi_i^\dagger\ \psi_k = \delta_{ik}\ .$$

- Fourier expansion

$$|\psi\rangle \ = \ \sum_{n=1}^{\infty} c_n\ |\psi_n\rangle \leftrightarrow\ \psi(q) = \sum_{n=1}^{\infty} c_n\ \psi_n(q)\ \leftrightarrow\ \psi = \sum_{n=1}^{\infty} c_n\ \psi_n\ .$$

- Fourier coefficients

$$c_n = \langle\psi_n|\psi\rangle = \int_{-\infty}^{\infty} \psi_n^*(q)\ \psi(q)\ dq = \psi_n^\dagger\ \psi\ .$$

consequence:

$$\begin{aligned}\psi(q) &= \sum_{n=1}^{\infty} c_n\ \psi_n(q) \ = \ \sum_{n=1}^{\infty}\int_{-\infty}^{\infty} \psi_n^*(q')\psi(q')dq'\psi_n(q) \\ &= \int_{-\infty}^{\infty} \underbrace{\sum_{n=1}^{\infty}\psi_n(q)\psi_n^*(q')}_{\delta(q-q')}\ \psi(q')\ dq' = \psi(q)\ .\end{aligned}$$

Here we have introduced Dirac's delta function,

$$\int_{-\infty}^{\infty} \delta(q-q')\ \psi(q')\ dq' \underset{\text{def}}{=} \psi(q)\ ,$$

which is the continuous analogue of the discrete Kronecker delta $\sum_{n=1}^{\infty} \delta_{mn}\ a_n \underset{\text{def}}{=} a_m$

.

- completeness

$$\sum_{n=1}^{\infty} |\psi_n\rangle\langle\psi_n| = I \quad\leftrightarrow\quad \sum_{n=1}^{\infty} \psi_n(q)\ \psi_n^*(q') = \delta(q-q') \text{ Dirac's delta function}$$

$$\leftrightarrow \sum_{n=1}^{\infty} \psi_n\ \otimes\ \psi_n^\dagger = I. \quad \text{Here} \quad \psi_n \otimes \psi_n^\dagger \quad \text{is the direct product matrix.}$$

- linear operators

$$F(\hat{p},\hat{q})\ \leftrightarrow\ F\left(-i\hbar\,\frac{\partial}{\partial q},\ q\right)\ \leftrightarrow\ F(p,q)$$

- matrix elements

$$\begin{aligned}F_{ik} = \langle\psi_i|F(\hat{p},\hat{q})\psi_k\rangle \ &= \ \int_{-\infty}^{\infty} \psi_i^*(q)\ F\left(-i\hbar\,\frac{\partial}{\partial q},\ q\right)\ \psi_k(q)\ dq \\ &= \ \psi_i^\dagger\ F(p,q)\ \psi_k\end{aligned}$$

- Hermiticity of an operator

$$\langle\varphi|F(\hat{p},\hat{q})\psi\rangle \ = \ \langle F(\hat{p},\hat{q})\ \varphi|\psi\rangle$$

$$\begin{aligned}\int_{-\infty}^{\infty} \varphi^*(q)\ F\left(-i\hbar\,\frac{\partial}{\partial q},\ q\right)\ \psi(q)\ dq \ &= \ \int_{-\infty}^{\infty} \left[F\left(-i\hbar\,\frac{\partial}{\partial q},\ q\right)\ \varphi(q)\right]^*\ \psi(q)\ dq \\ \varphi^\dagger\ F(p,q)\psi \ &= \ (F(p,q)\varphi)^\dagger\ \psi\end{aligned}$$

theorem: $-i\hbar\,\dfrac{\partial}{\partial q}$ is Hermitean

proof: through partial integration $\displaystyle\int_{-\infty}^{\infty} \varphi^*(q)\ \left(-i\hbar\,\frac{\partial}{\partial q}\right)\ \psi(q)\ dq$

$$\begin{aligned}&= \int_{-\infty}^{\infty} \left(-i\hbar\,\frac{\partial}{\partial q}\right)\ \varphi^*(q)\psi(q)\ dq - \int_{-\infty}^{\infty} \psi(q)\left(-i\hbar\frac{\partial}{\partial q}\right)\varphi^*(q)dq \\ &= \underbrace{-i\hbar\,\varphi^*(q)\ \varphi(q)\Big|_{-\infty}^{\infty}}_{0} + \int_{-\infty}^{\infty} \left[-i\hbar\frac{\partial}{\partial q}\varphi(q)\right]^*\ \psi(q)\ dq \quad \text{q.e.d.}\end{aligned}$$

Thus, proving the Hermiticity of differential operators always involves partial integration and the boundary conditions of wave functions at infinity.

- eigenvalues of a Hermitean operator $A(\hat{p},\hat{q})$

$$\begin{aligned}\left(A(\hat{p},\hat{q}) - a_n\,\hat{I}\right)\,|\psi_n\rangle &= 0\\ \left(A\left(-i\hbar\frac{\partial}{\partial q}, q\right) - a_n\right)\,|\psi_n(q)\rangle &= 0\\ (A(p,q) - a_n\,I)\;\psi_n &= 0\;.\end{aligned}$$

Finding the eigenvalues of differential operators also requires boundary conditions of the wave functions at infinity which in turn guarantees the square integrability of the wave function.

7.5 Link between Quantum Mechanics and Classical Mechanics

This link is most conveniently established in the Heisenberg picture.

- Quantum Mechanics

  The time-derivative of an operator $F(\hat{p}(t),\hat{q}(t))$ in the Heisenberg picture is given by

$$\frac{d}{dt}F(\hat{p}(t),\hat{q}(t)) = \frac{1}{i\hbar}\left[F(\hat{p}(t),\hat{q}(t)), H(\hat{p}(t),\hat{q}(t))\right]$$

- Classical Mechanics

  theorem:

$$\frac{d}{dt}F(p(t),q(t)) = \{F(p(t),q(t)), H(p(t),q(t))\}_{\text{Poisson}}$$

  definition:

$$\{F,H\}_{\text{Poisson}} \underset{\text{def}}{=} \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial H}{\partial q} \quad \text{Poisson bracket}$$

  proof:

$$\frac{dF}{dt} = \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial p}\dot{p}$$

  but

$$\left.\begin{aligned}\dot{q} &= \frac{\partial H}{\partial p}\\ \dot{p} &= -\frac{\partial H}{\partial q}\end{aligned}\right\}\ \text{Hamilton's equations of motion as derived in section 4.1}$$

  therefore

$$\frac{dF}{dt} = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial H}{\partial q} \underset{\text{def}}{=} \{F,H\}_{\text{Poisson}} \quad \text{q.e.d.}$$

Applying the definition of the Poisson's bracket to p and q, we have, substituting $F = p$ and $H = q$,

$$\{p,q\}_{\text{Poisson}} = \frac{\partial p}{\partial q}\frac{\partial q}{\partial p} - \frac{\partial p}{\partial p}\frac{\partial q}{\partial q} = -1\;.$$

Thus the transition from CM to QM means the substitution of the Poisson bracket with the commutator of the corresponding variables, i.e. divided by $i\hbar$

$$\begin{aligned}\{F,H\}_{\text{Poisson}} &\longleftrightarrow \left[\hat{F},\hat{H}\right]/i\hbar\\ \{p,q\}_{\text{Poisson}} &\longleftrightarrow [\hat{p},\hat{q}]/i\hbar\;.\end{aligned}$$

remarks: This substitution rule only holds in the Heisenberg picture.

W. Heisenberg has been guided by this analogy in his quest for the formulation of the equation of motion for the operators in the Heisenberg picture in 1925.

Classical Mechanics	Quantum Mechanics
W.R. Hamilton (1834)	W. Heisenberg (1925)
$\{p,q\}_{\text{Poisson}} = -1$	$\frac{1}{i\hbar}[\hat{p},\hat{q}] = -\hat{I}$ M. Born, P. Jordan (1924)
$\frac{dq}{dt} = \{q,H\}_{\text{Poisson}} = \frac{\partial H}{\partial p}$	$\frac{d\hat{q}}{dt} = \frac{1}{i\hbar}\left[\hat{q},\hat{H}\right] = \frac{\partial \hat{H}}{\partial \hat{p}}$
$\frac{dp}{dt} = \{p,H\}_{\text{Poisson}} = -\frac{\partial H}{\partial q}$	$\frac{d\hat{p}}{dt} = \frac{1}{i\hbar}\left[\hat{p},\hat{H}\right] = -\frac{\partial \hat{H}}{\partial \hat{q}}$
Hamilton's equations of motion	Heisenberg's equations of motion

We have not yet shown the last equality in the Heisenberg's equations of motion.

- theorem: $\left[\hat{q},\hat{H}\right] = i\hbar\frac{\partial \hat{H}}{\partial \hat{p}}$

  proof: choose momentum space representation

$$\hat{p} \to p, \qquad \hat{q} \to i\hbar\frac{\partial}{\partial p}$$

$$\hat{H}(\hat{p},\hat{q}) \to \hat{H}\left(p, i\hbar\frac{\partial}{\partial p}\right)$$

$$\left[\hat{q},\hat{H}\right] a(p,t) = \left[i\hbar\frac{\partial}{\partial p},\hat{H}\right] a(p,t)$$
$$= i\hbar\frac{\partial}{\partial p}\left(\hat{H}a(p,t)\right) - \hat{H}i\hbar\frac{\partial}{\partial p}a(p,t)$$
$$= i\hbar\frac{\partial\hat{H}}{\partial p}a(p,t) + i\hbar\,\hat{H}\,\frac{\partial a(p,t)}{\partial p} - i\hbar\,\hat{H}\,\frac{\partial a(p,t)}{\partial p}$$
valid for any $a(p,t)$

$$\Rightarrow \quad \left[\hat{q},\hat{H}\right] = i\hbar\frac{\partial\hat{H}}{\partial\hat{p}} \quad \text{q.e.d.}$$

- theorem: $\left[\hat{p},\hat{H}\right] = -i\hbar\frac{\partial\hat{H}}{\partial\hat{q}}$

  proof: choose coordinate space representation

$$\hat{p} \;\to\; -i\hbar\frac{\partial}{\partial q}, \qquad \hat{q} \;\to\; q$$
$$\hat{H}\left(\hat{p},\hat{q}\right) \;\to\; \hat{H}\left(-i\hbar\frac{\partial}{\partial q},q\right)$$

$$\left[\hat{p},\hat{H}\right]\psi(q,t) = \left[-i\hbar\frac{\partial}{\partial q},\hat{H}\right]\psi(q,t)$$
$$= -i\hbar\frac{\partial}{\partial q}\left(\hat{H}\psi(q,t)\right) - \hat{H}\left(-i\hbar\frac{\partial}{\partial q}\right)\psi(q,t)$$
$$= -i\hbar\frac{\partial\hat{H}}{\partial q}\psi(q,t) - i\hbar\,\hat{H}\,\frac{\partial\psi(q,t)}{\partial q} + i\hbar\,\hat{H}\,\frac{\partial\psi(q,t)}{\partial q}$$
valid for any $\psi(q,t)$

$$\Rightarrow \quad \left[\hat{p},\hat{H}\right] = -i\hbar\frac{\partial\hat{H}}{\partial\hat{q}} \quad \text{q.e.d.}$$

- summary:

  The operators in the Heisenberg picture obey the classical equations of motion in operator form. However, in tut 18 we show that the expectation values satisfy the classical equations of motion only under certain conditions. Thus, in general, there will be quantum corrections to the classical theory.

8 Rotation group

Physics happens in 3-dimensional Euclidean space $\mathcal{E}_3$. We thus need some insight into the properties of $\mathcal{E}_3$.

8.1 Vectors and scalars in $\mathcal{E}_3$

- $\mathcal{E}_3$ is a vector space over the field of the real numbers $\mathcal{R}$. Thus we have

  vectors:
$$\vec{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \in \mathcal{E}_3$$
  $a_x, a_y, a_z \in \mathcal{R}$ being the Cartesian components of $\vec{a}$,

  and scalars: $\lambda \in \mathcal{R}$.

- $\dim\mathcal{E}_3 = 3$: Every vector $\vec{a} \in \mathcal{E}_3$ is represented by
$$\vec{a} = a_x\vec{e}_x + a_y\vec{e}_y + a_z\vec{e}_z = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}$$
  with the basis vectors
$$\vec{e}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

8.2 Scalar product of two vectors

For every pair of vectors $\vec{a}, \vec{b} \in \mathcal{E}_3$, the scalar product is defined as
$$\vec{a}\cdot\vec{b} \underset{\text{def}}{=} \vec{a}^{\,T}\vec{b} = a_xb_x + a_yb_y + a_zb_z ,$$
with $\vec{a}^{\,T} \underset{\text{def}}{=} (a_x,a_y,a_z) \in \mathcal{E}_3^T$, being the transposed vector of the dual space $\mathcal{E}_3^T$.

- properties:

$\vec{a}\cdot\left(\lambda\vec{b}+\mu\vec{c}\right)$	$=$	$\lambda\vec{a}\cdot\vec{b}+\mu\vec{a}\cdot\vec{c}$	linearity
$\vec{a}\cdot\vec{b}$	$=$	$\vec{b}\cdot\vec{a}$	symmetry
$\vec{a}\cdot\vec{a}$	$>$	0	positive definiteness

  except for $\vec{a}^{\,T} = (0,0,0)$

- definition:

$$|\vec{a}| \underset{\text{def}}{=} \sqrt{\vec{a}\cdot\vec{a}} = \sqrt{a_x^2 + a_y^2 + a_z^2} \qquad \text{absolute value of } \vec{a}$$

- consequences:

$$\vec{e}_i \cdot \vec{e}_j \underset{\text{def}}{=} \vec{e}_i{}^T \vec{e}_j = \delta_{ij} \qquad (i,j = x,y,z) \ ,$$

i.e. the basis vectors are orthonormal.

remarks:

(i) Schwarz's, Minkowski's and triangular inequalities must hold.

(ii) The scalar product can be used to measure lengths of vectors and angles between vectors, i.e. the geometry of $\mathcal{E}_3$ is governed by an Euclidean metric.

8.3 Orthogonal transformations

- Consider a homogeneous linear transformation R of a vector $\vec{a} \in \mathcal{E}_3$

$$\vec{a}\,' = \begin{pmatrix} a'_x \\ a'_y \\ a'_z \end{pmatrix} = \begin{pmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{pmatrix} \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} = R\vec{a}$$

and similarly for $\vec{b}\,' = R\vec{b}$, with $\vec{b} \in \mathcal{E}_3$. We require R to leave the Euclidean scalar product invariant, i.e.

$$\vec{a}\,' \cdot \vec{b}\,' = (\vec{a}\,')^T \vec{b}\,' = \vec{a}^T R^T R\, \vec{b} = \vec{a}^T \vec{b} = \vec{a}\cdot\vec{b} \ ,$$

valid for every pair $\vec{a}$ and $\vec{b} \in \mathcal{E}_3$

$\Rightarrow$ R is restricted to $R^T R = I_3 \Rightarrow \det R^T \det R = (\det R)^2 = 1$.

- consequence:

$\det R$ is either $+1$ or -1. As $\det R$ is non-zero, it is invertible, i.e.

$$R^{-1}R = RR^{-1} = I_3 \Rightarrow R^{-1} = R^T.$$

definition: A matrix satisfying $R^{-1} = R^T$ is an orthogonal matrix.

- properties:

The infinite set of orthogonal matrices forms the Lie group $O(3,\mathcal{R})$ with respect to matrix multiplication (Sophus Lie, $\sim$ 1865). The group elements are described by continuous real parameters. O stands for orthogonal, 3 for 3-dimensional, and $\mathcal{R}$ for real matrices.

The group properties are the same as those of finite groups:

(i) The composition law of two elements R and S is defined by the matrix product RS. Indeed, if $R^T = R^{-1}$ and $S^T = S^{-1}$ (or $R, S \in O\ (3,\mathcal{R})$), then

$$\left.\begin{aligned} (RS)^T &= S^T R^T \\ (RS)^{-1} &= S^{-1}R^{-1} = S^T R^T \end{aligned}\right\} \Longrightarrow \begin{aligned} &(RS)^T = (RS)^{-1} \\ &\text{or } RS \in O(3,\mathcal{R}) \end{aligned}$$

(ii) $R(ST) = (RS)T$ is due to the associativity of the matrix product.

(iii) There exists a unit element I_3 with $R\,I_3 = I_3\,R = R$.

$I_3 \in O(3,\mathcal{R})$, because $I_3^T = I_3^{-1}$ is also orthogonal.

(iv) To every R there exists an inverse $R^{-1} \in O(3,\mathcal{R})$, as

$$\left(R^{-1}\right)^T = \left(R^T\right)^T = R = \left(R^{-1}\right)^{-1} \quad \text{orthogonal.}$$

remarks:

(i) In general $RS \neq SR$, i.e. the group is non-commutative or non-Abelian.

(ii) There are two disjoint classes of orthogonal transformations, i.e. classes of R, which cannot be joined by varying the real matrix elements R_{ik}, with $R^T = R^{-1}$, continuously: $(\det R)^2 = 1 \Rightarrow$

$\det R = 1$: proper orthogonal transformations $\equiv$ rotations.

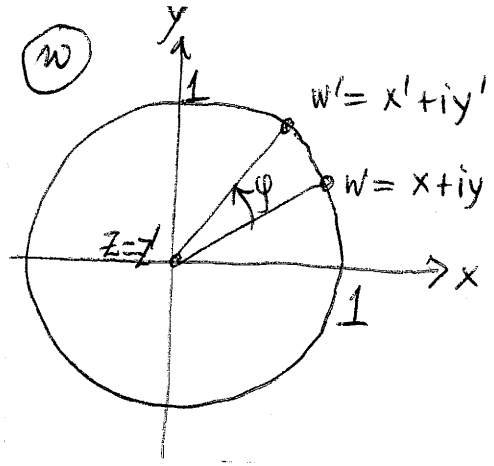
$\det R = -1$: improper orthogonal transformations $\equiv$ reflections.

(iii) While the rotations are exact symmetries, the reflections are only approximate symmetries of the laws of Nature. On a macroscopic scale, we can easily rotate our bodies without a problem, but we cannot transform ourselves into our own reflection images.

examples:

(a) rotation about z-axis:

For a rotation in the complex w-plane we have



$$w' = x' + iy' = e^{i\varphi} w = (\cos\varphi + i\sin\varphi)(x + iy)$$

$$\text{or } x' + iy' = (x\cos\varphi - y\sin\varphi) + i(y\cos\varphi + x\sin\varphi) \ .$$

Equating real and imaginary parts, we arrive at

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{R_z(\varphi)} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

remarks:

(i) Similarly, we can describe rotations around the x- and y-axes.

(ii) The subset of all rotations forms a (Lie) subgroup $SO(3,\mathcal{R}) \in O(3,\mathcal{R})$ because, if $\det R = 1$ and $\det S = 1$, then $\det RS = \det R \det S = 1$.

S for special or unimodular, i.e. $\det R = 1$
O for orthogonal
3 for 3-dimensional
$\mathcal{R}$ for real matrices

(b) total space reflection:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \overbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}^{P} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \underset{\text{def}}{=} P \begin{pmatrix} x \\ y \\ z \end{pmatrix} = - \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

with $\det P = -1$.

remark:

Under P transformation vectors, axial vectors, scalars and pseudoscalars transform as follows:

vectors:	$\vec{r}, \vec{p} \to -\vec{r}, -\vec{p}$
axial vectors:	$\vec{L} = \vec{r} \times \vec{p} \to (-\vec{r}) \times (-\vec{p}) = \vec{r} \times \vec{p} = \vec{L}$
scalars:	$\vec{r} \cdot \vec{p} \to (-\vec{r}) \cdot (-\vec{p}) = \vec{r} \cdot \vec{p}$
pseudoscalars:	$\vec{r} \cdot \vec{L} \to -\vec{r} \cdot \vec{L}$

(c) some simple questions:

* Is P an orthogonal matrix?

$$\left.\begin{array}{rcl} P^2 = I_3 \Rightarrow P & = & P^{-1} \\ P & = & P^T \end{array}\right\} \quad \begin{array}{l} \Rightarrow \quad P^{-1} = P^T \\ \quad \text{orthogonal} \\ \Rightarrow \quad \text{yes} \end{array}$$

* Do the reflections form a subgroup of $O(3,\mathcal{R})$?

(i) $\det I_3 = 1 \Rightarrow$ unit element I_3 is not a reflection $\Rightarrow$ no

(ii) $PP = P^2 = I_3$ product of two parity transformations is not a reflection $\Rightarrow$ no

* Is X an orthogonal matrix?

$$X = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Rightarrow X^T = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ X^{-1} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ as } X^{-1}X = I_3$$

$$\Rightarrow \quad X^T = X^{-1} \quad \text{orthogonal} \Rightarrow \text{yes}$$

* Is X a reflection or a rotation?

$$\det X = 1 \implies \text{rotation}$$

indeed, $X = R_z(\pi)$ with

$$R_z(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

* Is Y a reflection or a rotation?

$$Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \Rightarrow \det Y = -1 \Rightarrow \text{reflection}$$

* How is Y related to the total reflection?

take a look at

$$XP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = Y$$

$$\Rightarrow \; R_z(\pi)\, P = P R_z(\pi) = Y$$

remark: all reflections can be represented by the product of a rotation and a total reflection.

8.4 Infinitesimal rotations

- Sophus Lie's brilliant idea ($\sim$ 1865, Norwegian).
  All you need to specify the rotation group are the infinitesimal rotations in the vicinity of the unit transformation.

example: expanding the rotation matrix around z-axis in powers of φ

$$R_z(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\varphi & 0 \\ \varphi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + O(\varphi^2)$$

$$= I_3 + \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{r_z} \varphi + O(\varphi^2) \; .$$

r_z is called the generator of the rotation (counter-clockwise) around z-axis.
in general:

$$R_k(\varphi) = I + r_k\, \varphi + O(\varphi^2)$$

$$r_k \underset{\text{def}}{=} \left.\frac{dR_k(\varphi)}{d\varphi}\right|_{\varphi=0} \qquad k = x, y, z$$

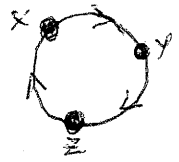
generator for rotations (counter-clockwise) around k-axis

$$r_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} , \; r_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} , \; r_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

Here we have assumed a right-handed coordinate system, and we have generated r_x and r_y through cyclic permutations of the matrix elements $(r_z)_{xy}$ and $(r_z)_{yx}$, i.e.

$$(r_x)_{yz} = -1 \qquad (r_y)_{zx} = -1 \qquad (r_z)_{xy} = -1$$
$$(r_x)_{zy} = 1 \qquad (r_y)_{xz} = 1 \qquad (r_z)_{yx} = 1$$

note: $r_k^T = -r_k$ antisymmetric $\qquad k = x, y, z$



cyclic permutation

- commutation relations: $[r_x, r_y] = ?$

$$\overbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}}^{r_x} \quad \overbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}}^{r_y} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}}_{r_y} \quad \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}}_{r_x} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

subtract

$$[r_x, r_y] = r_x\, r_y - r_y\, r_x = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = r_z \ .$$

By cyclic permutation we arrive at

- Lie algebra of $SO(3, \mathcal{R})$ (in Mathematics):

$$\begin{aligned} [r_x, r_y] &= r_z \\ [r_y, r_z] &= r_x \\ [r_z, r_x] &= r_y \end{aligned}$$

This looks almost like $[L_x, L_y] = i\, L_z$ + cyclic permutations, i.e. the angular momentum commutation relations. However, the L_k are Hermitean, i.e. $L_k^\dagger = L_k \quad (k = x, y, z)$, while the $r_k^T = -r_k$ are real and antisymmetric. To get Hermitean generators, we multiply by i

$$i\, r_k \underset{\text{def}}{=} J_k \ \Rightarrow\ J_k^\dagger = J_k \text{ Hermitean and purely imaginary}$$

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad J_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad J_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ .$$

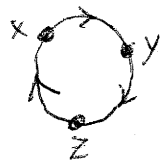
The commutation relations transform into

$$\begin{aligned} [r_x, r_y] = r_z \ \Rightarrow\ [ir_x, ir_y] = -r_z &= i(i\, r_z) \\ [J_x, J_y] &= i\, J_z \ . \end{aligned}$$

We thus arrive at the

- Lie algebra of $SO(3, \mathcal{R})$ (in Physics):

$$\begin{aligned} [J_x, J_y] &= i\, J_z \\ [J_y, J_z] &= i\, J_x \\ [J_z, J_x] &= i\, J_y \end{aligned}$$



cyclic permutation

important observation:

These are the same commutation relations which the Cartesian components of the angular momentum $\vec{L} \underset{\text{def}}{=} \vec{r} \times \vec{p}/\hbar$, i.e. L_x, L_y, L_z fulfil, and which have been derived using the generalized Born-Jordan quantization conditions (see tut 11).

- eigenvalues of J_x, J_y, J_z and J^2

We have shown, in tut 6, that the 3×3 matrix $J_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ has the eigenvalues 1, 0, -1. This is also true for J_x and J_y. We can easily evaluate

$$J_x^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad J_y^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad J_z^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ .$$

Thus

$$J^2 = J_x^2 + J_y^2 + J_z^2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 2I_3 \ .$$

The eigenvalues of J^2 are given by the zeros of the characteristic polynomial. Thus

$$P_3(a) = \det \begin{pmatrix} 2-a & 0 & 0 \\ 0 & 2-a & 0 \\ 0 & 0 & 2-a \end{pmatrix} = (2-a)^3 = 0 \ .$$

$\Rightarrow$ J^2 has the threefold degenerate eigenvalue 2.

$$\left[J^2, J_z\right] = 0 \ , \ \text{because} \ J^2 \propto I_3 \ .$$

- eigenvalues of S_x, S_y, S_z and S^2

The 2×2 Pauli matrices

$$S_x = \frac{\sigma_x}{2} \ , \ S_y = \frac{\sigma_y}{2} \ , \ S_z = \frac{\sigma_z}{2}$$

$$\text{with} \ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Pauli matrices: W. Pauli (1925)
quaternions: W.R. Hamilton (1843)

are Hermitean and fulfil $[S_x, S_y] = i\, S_z$ (+ cyclic permutations), as well (see tut 13). However, S_x, S_y and S_z each have only two eigenvalues, namely $\frac{1}{2}, -\frac{1}{2}$. The 2×2 matrix

$$S^2 = S_x^2 + S_y^2 + S_z^2 = \frac{\sigma_x^2}{4} + \frac{\sigma_y^2}{4} + \frac{\sigma_z^2}{4} = \frac{3}{4}\, I_2$$

has the twofold degenerate eigenvalue $\frac{3}{4}$. As $S^2 \propto I_2$, we have $\left[S^2, S_z\right] = 0$. Later we will see that J^2, J_z describe a spin 1, while S^2, S_z describe a spin $\frac{1}{2}$ particle.

8.5 Finite rotations again

- The functional equation of $R_k(\varphi)$ is

$$R_k(\varphi + \lambda) = R_k(\varphi)\ R_k(\lambda) = R_k(\lambda)\ R_k(\varphi) \ .$$

These rotations commute, because they have the same rotation axis.

strategy: We would like to derive a differential equation for $R_k(\varphi)$

$$\begin{aligned} \frac{dR_k(\varphi)}{d\varphi} &\underset{\text{def}}{=} \lim_{\lambda \to 0} \frac{R_k(\varphi + \lambda) - R_k(\varphi)}{\lambda} \\ &= \lim_{\lambda \to 0} \frac{R_k(\varphi)\ R_k(\lambda) - R_k(\varphi)}{\lambda} \\ &= R_k(\varphi) \lim_{\lambda \to 0} \left[\frac{R_k(\lambda) - I_3}{\lambda}\right] \\ &= \lim_{\lambda \to 0} \left[\frac{R_k(\lambda) - I_3}{\lambda}\right] R_k(\varphi) \ . \end{aligned}$$

We obtained in section 8.4 $\quad R_k(\lambda) = I_3 + \lambda\, r_k + O(\lambda^2)$.

Inserting $R_k(\lambda)$ in the derivatives, we have

$$\Rightarrow \begin{cases} \dfrac{d\, R_k(\varphi)}{d\varphi} &= R_k(\varphi)\ r_k = r_k\ R(\varphi) \\ R_k(0) &= I_3 \end{cases} .$$

This is a first-order differential equation for $R_k(\varphi)$ with the initial condition $R_k(0) = I_3$.

solution:

$$R_k(\varphi) = \exp\left(\varphi r_k\right) = \exp\left(-i\varphi\ J_k\right)$$

$$\text{with} \quad r_k = -iJ_k \ .$$

For J_z we have shown in tut 12

$$R_k(\varphi) = \exp\left(-i\varphi\ J_z\right) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \ .$$

- theorem: $R_z(\varphi)$ is orthogonal

proof:

$$\left(R_z(\varphi)\right)^T = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = R_z(-\varphi)$$

$$\Rightarrow \quad R_z(-\varphi) = (R(\varphi))^T . \qquad \text{eq.(1)}$$

$$\text{But} \quad R_z(\varphi)\, R_z(-\varphi) = R_z(-\varphi)\, R_z(\varphi) =$$

$$= R_z(\varphi - \varphi) = R_z(0) = I_3$$

$$\Rightarrow \quad R_z(-\varphi) = (R_z(\varphi))^{-1} \quad \text{eq.(2)}$$

Comparing eqs.(1) and (2) we conclude

$$(R_z(\varphi))^T = (R_z(\varphi))^{-1} \qquad \text{orthogonal or } R_z(\varphi) \in O(3,\mathcal{R}) \qquad \text{q.e.d.}$$

- theorem: $R_z(\varphi)$ is unimodular

  proof:

$$\det R_z(\varphi) = \begin{vmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= 1 \cdot \begin{vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{vmatrix}$$

$$= \cos^2\varphi + \sin^2\varphi = 1$$

Thus $R_z(\varphi) \in SO(3,\mathcal{R})$ q.e.d.

9 Matrix representation of the Lie algebra of $SO(3,\mathcal{R})$

9.1 Motivation

We have found two linear representations of the Lie algebra of $SO(3,\mathcal{R})$

- 3×3 matrices

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad J_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \qquad J_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

  The eigenvalues of J_x, J_y, J_z are 1, 0, -1 and those of $J^2 = J_x^2 + J_y^2 + J_z^2$ are 2.

- 2×2 matrices

$$S_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad S_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad S_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

  The eigenvalues of S_x, S_y, S_z are $\frac{1}{2}, -\frac{1}{2}$ and those of $S^2 = S_x^2 + S_y^2 + S_z^2$ are $\frac{3}{4}$.

  As the eigenvalues of S_k and J_k differ, these are inequivalent representations of the same Lie algebra.

- question: Are there more such inequivalent complex $n \times n$ matrices with $n = 2,3, \cdots$ satisfying the Lie-algebra of $SO(3,\mathcal{R})$?

  strategy: We first find the eigenvectors and eigenvalues of J^2 and J_z using the Lie-algebra in terms of abstract operators. We then calculate the matrix elements of J_x, J_y, J_z, in the basis of the eigenvectors of J^2 and J_z.

9.2 Eigenvalues and eigenvectors of J^2 and J_z

The Casimir operator of $SO(3,\mathcal{R})$

$$J^2 = J_x^2 + J_y^2 + J_z^2$$

commutes with all the generators of $SO(3,\mathcal{R})$ (see tut 11(b)), i.e.

$$[J^2, J_x] = [J^2, J_y] = [J^2, J_z] = 0 .$$

Thus J^2 and e.g. J_z have common eigenvectors (see tut 7(a))

$$J^2|\psi_{jm}\rangle = \lambda_j|\psi_{jm}\rangle$$

$$J_z|\psi_{jm}\rangle = m|\psi_{jm}\rangle .$$

The index j labels the different eigenvalues of J^2. If $A = A^\dagger$ (Hermitean), then obviously

$$\langle\psi|A^2\psi\rangle = \langle\psi|A^\dagger A\,\psi\rangle = \langle A\psi|A\psi\rangle \geqslant 0 \ .$$

This is valid also for a sum of two square terms with $A^\dagger = A$ and $B^\dagger = B$, i.e.

$$\langle\psi|(A^2+B^2)\psi\rangle = \langle\psi|A^\dagger A\psi\rangle + \langle\psi|B^\dagger B\psi\rangle = \langle A\psi|A\psi\rangle + \langle B\psi|B\psi\rangle \geqslant 0 \ .$$

Choosing $A = J_x$ and $B = J_y$, we have

$$\langle\psi_{jm}|\left(J_x^2+J_y^2\right)\psi_{jm}\rangle = \langle\psi_{jm}|\left(J^2-J_z^2\right)\psi_{jm}\rangle = \left(\lambda_j - m^2\right)\underbrace{\langle\psi_{jm}|\psi_{jm}\rangle}_{1} = \lambda_j - m^2 \geqslant 0 \ .$$

Finally, for a sum of three square terms with $A = A^\dagger$, $B = B^\dagger$ and $C = C^\dagger$, we have

$$\begin{aligned}\langle\psi|\left(A^2+B^2+C^2\right)\psi\rangle &= \langle\psi|A^\dagger A\psi\rangle + \langle\psi|B^\dagger B\psi\rangle + \langle\psi|C^\dagger C\psi\rangle \\ &= \langle A\psi|A\psi\rangle + \langle B\psi|B\psi\rangle + \langle C\psi|C\psi\rangle \geqslant 0 \ .\end{aligned}$$

Choosing $A = J_x$, $B = J_y$ and $C = J_z$ we have

$$\begin{aligned}\langle\psi_{jm}|\left(J_x^2+J_y^2+J_z^2\right)\psi_{jm}\rangle &= \lambda_j\,\underbrace{\langle\psi_{jm}|\psi_{jm}\rangle}_{1} \\ &= \lambda_j \geqslant 0 \ .\end{aligned}$$

We now define step or ladder operators $J_\pm \underset{\text{def}}{=} J_x \pm i\,J_y$ or

$$\begin{aligned}J_x &= \frac{1}{2}\,(J_+ + J_-) \\ J_y &= \frac{1}{2i}\,(J_+ - J_-) \ ,\end{aligned}$$

and

$$J^2 = J_x^2 + J_y^2 + J_z^2 = \frac{1}{2}\,(J_+\,J_- + J_-\,J_+) + J_z^2 \ , \qquad \text{see tut 11(c)}$$

The Lie algebra of $SO(3,\mathcal{R})$ in the Cartan form is thus

$$\begin{aligned}[J_z, J_\pm] &= \pm J_\pm \\ [J_+, J_-] &= 2\,J_z \ ,\end{aligned}$$

with

$$\left[J^2, J_\pm\right] = 0 \quad \text{and} \quad \left[J^2, J_z\right] = 0 \ .$$

- Let

$$\begin{aligned}J^2|\psi_{jm}\rangle &= \lambda_j|\psi_{jm}\rangle \\ J_z|\psi_{jm}\rangle &= m|\psi_{jm}\rangle \ .\end{aligned}$$

<u>theorem:</u> $J_\pm|\psi_{jm}\rangle$ are also eigenvectors of J^2 and J_z with

$$\begin{aligned}J^2\left(J_\pm|\psi_{jm}\rangle\right) &= \lambda_j\left(J_\pm|\psi_{jm}\rangle\right), && \text{same eigenvalue} \\ J_z\left(J_\pm|\psi_{jm}\rangle\right) &= (m\pm 1)J_\pm|\psi_{jm}\rangle, && \text{eigenvalue shifted by } \pm\,1\end{aligned}$$

<u>proof:</u>

$$\begin{aligned}J^2(J_\pm|\psi_{jm}\rangle) &= \left(J_\pm J^2 + \overbrace{\left[J^2, J_\pm\right]}^{=0}\right)|\psi_{jm}\rangle \\ &= \lambda_j\left(J_\pm|\psi_{jm}\rangle\right)\end{aligned}$$

$$\begin{aligned}J_z(J_\pm|\psi_{jm}\rangle) &= \left(J_\pm J_z + \overbrace{\left[J_z, J_\pm\right]}^{\pm J_\pm}\right)|\psi_{jm}\rangle \\ &= J_\pm(J_z \pm I)\,|\psi_{jm}\rangle \\ &= (m\pm 1)\,J_\pm\,|\psi_{jm}\rangle \qquad \text{q.e.d.}\end{aligned}$$

$$\Rightarrow J_\pm\,|\psi_{jm}\rangle = c_\pm(j,m)\,|\psi_{j,m\pm1}\rangle \quad \left.\begin{matrix}\text{raising} \\ \text{lowering}\end{matrix}\right\} \text{ step operators}$$

$$\text{with} \quad c_\pm(j,m) \in \mathcal{C} \ .$$

The values of m for given λ_j are limited by

$$\lambda_j - m^2 \geqslant 0 \ , \ \lambda_j \geqslant 0 \ , \text{ as shown earlier.}$$

$\Rightarrow$ for every value of λ_j there exists a largest acceptable m-value m_{max} and a smallest acceptable m-value m_{min}

$$m_{max} \geqslant m_{min} \ .$$

Acting with $J_\pm$ on $|\psi_{jm}\rangle$ repeatedly, we will increase (decrease) m beyond the bound $\lambda_j - m^2 \geqslant 0$. The only way to avoid this contradiction is for these sequences to terminate at both ends, i.e.

$$J_+ |\psi_{jm_{max}}\rangle = 0$$

$$J_- |\psi_{jm_{min}}\rangle = 0$$

$$\begin{aligned} J^2 &= \tfrac{1}{2}(J_+J_- + J_-J_+) + J_z^2 \\ J_z &= \tfrac{1}{2}(J_+J_- - J_-J_+) \qquad \text{add or subtract} \\ \hline J^2 \pm J_z &= J_\pm J_\mp + J_z^2 \end{aligned}$$

$$\Rightarrow \quad J_\pm J_\mp = J^2 - J_z\,(J_z \mp I) \quad \text{mathematical identity}$$

$$\begin{cases} J_-J_+ |\psi_{jm_{max}}\rangle = 0 \;\Rightarrow \underbrace{[\lambda_j - m_{max}(m_{max}+1)]}_{0} |\psi_{jm_{max}}\rangle = 0 \\ J_+J_- |\psi_{jm_{min}}\rangle = 0 \;\Rightarrow \underbrace{[\lambda_j - m_{min}(m_{min}-1)]}_{0} |\psi_{jm_{min}}\rangle = 0 \end{cases}$$

$$\Rightarrow \quad \lambda_j = m_{max}(m_{max}+1) = m_{min}(m_{min}-1)$$

theorem: $\qquad (m_{min} + m_{max})(m_{max} - m_{min} + 1) = 0$

proof:

$$m_{min}\, m_{max} - m_{min}^2 + m_{min} + m_{max}^2 - m_{max}m_{min} + m_{max} = 0$$

$$\Rightarrow \quad m_{max}^2 + m_{max} = m_{min}^2 - m_{min}$$

$$\Rightarrow \quad m_{max}(m_{max}+1) = m_{min}(m_{min}-1) = \lambda_j \qquad \text{q.e.d.}$$

- conclusion:

One of the factors in the product

$$(m_{min} + m_{max})\,(m_{max} - m_{min} + 1) = 0$$

must be zero. It can only be

$$m_{min} + m_{max} = 0\,,$$

because $m_{max} \geqslant m_{min}$, and thus $m_{max} - m_{min} + 1 > 0$ always.

Thus $\qquad m_{min} = -m_{max}$.

We know that successive values of m differ by 1

$$\Rightarrow \quad m_{max} - m_{min} = n \qquad n = 0, 1, 2, \cdots$$

$$\text{but} \quad m_{max} = -m_{min} \quad \Rightarrow \quad 2\, m_{max} = n$$

$$m_{max} = \frac{n}{2} \qquad m_{min} = -\frac{n}{2}$$

$$m = -\frac{n}{2}, -\frac{n}{2}+1, \cdots, \frac{n}{2}$$

$$2\frac{n}{2} + 1 \qquad \text{values of } m$$

- definition:

$$j \underset{\text{def}}{=} \frac{n}{2} \quad \Rightarrow \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$$

$$\Rightarrow \quad m = -j, -j+1, \cdots, +j \qquad (2j+1) \text{ values of } m$$

$$\Rightarrow \quad \lambda_j = m_{max}\,(m_{max}+1) = m_{min}\,(m_{min}-1) = j(j+1)$$

$$J^2|\psi_{jm}\rangle = j(j+1)|\psi_{jm}\rangle \qquad j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \cdots$$

$$J_z|\psi_{jm}\rangle = m|\psi_{jm}\rangle \qquad \underbrace{m = -j, -j+1, \cdots, +j}_{(2j+1) \text{ values of } m}$$

- summary: What went into this result?

(i) J_x, J_y, J_z Hermitean

(ii) $[J_x, J_y] = i\, J_z$ + cyclic permutations, Lie group $SO(3, \mathcal{R})$.
$\Rightarrow$ Casimir operator $J^2 = J_x^2 + J_y^2 + J_z^2$
commutes with all J_x, J_y and J_z, e.g.
$[J^2, J_z] = 0 \;\Rightarrow$ we can choose simultaneous eigenvectors of J^2, J_z.

(iii) non-negativity of the expectation values
$\langle\psi|(J_x^2 + J_y^2 + J_z^2)\psi\rangle \geqslant 0$; $\langle\psi|(J_x^2 + J_y^2)\psi\rangle \geqslant 0$.

- The surprising consequence of these simple considerations is:

Half-integer eigenvalues for J_z are allowed. This important fact was known to W. Pauli already in early 1925, ahead of the discovery of the electron's spin by S.A. Goudsmit and G.E. Uhlenbeck in mid 1925.

9.3 Ladder operators acting on normalized eigenvectors

- So far we know $J_\pm|\psi_{jm}\rangle = c_\pm(j,m)|\psi_{j,m\pm1}\rangle$, where $c_\pm(j,m) \in C$ may be determined fixing $||\psi_{jm}|| = ||\psi_{j,m\pm1}|| = 1$.

- strategy:

  Calculate $c_\pm(j,m)$ using the mathematical identity
  $J_\mp J_\pm = J^2 - J_z\ (J_z \pm I)$, which we have derived in section 9.2.

- Calculate the norm squared of $J_\pm|\psi_{jm}\rangle$

$$\begin{aligned} ||J_\pm\psi_{jm}||^2 &\underset{\text{def}}{=} \langle J_\pm\psi_{jm}|J_\pm\psi_{jm}\rangle \\ &= \langle c_\pm(j,m)\ \psi_{j,m\pm1}|c_\pm(j,m)\ \psi_{j,m\pm1}\rangle \\ &= |c_\pm(j,m)|^2 \underbrace{\langle\psi_{j,m\pm1}|\psi_{j,m\pm1}\rangle}_{\text{1 per definition}} \end{aligned}$$

- We now use $J_\pm^\dagger = (J_x \pm i\ J_y)^\dagger = \left(J_x^\dagger \mp i\ J_y^\dagger\right) = (J_x \mp i\ J_y) = J_\mp$

  $\Rightarrow$ J_+ and J_- are adjoints of each other

$$\begin{aligned} |c_\pm(j,m)|^2 &= ||J_\pm\psi_{jm}||^2 = \langle J_\pm\ \psi_{jm}|J_\pm\psi_{jm}\rangle = \langle\psi_{jm}|J_\mp J_\pm\psi_{jm}\rangle \\ &= \langle\psi_{jm}|\left(J^2 - J_z(J_z \pm I)\right)\ \psi_{jm}\rangle \quad \text{from section 9.2} \\ &= (j(j+1) - m(m\pm1))\ \underbrace{\langle\psi_{jm}|\psi_{jm}\rangle}_{\text{1 per definition}} \end{aligned}$$

$$\begin{aligned} \Rightarrow |c_\pm(j,m)|^2 &= j(j+1) - m(m\pm1) \\ &= (j\mp m)(j\pm m+1) \\ \Rightarrow c_\pm(j,m) &= \sqrt{(j\mp m)(j\pm m+1)}\ \underbrace{e^{i\varphi_\pm(j,m)}}_{=1}\ , \end{aligned}$$

i.e. we chose the simplest phase convention *à la* Condon and Shortley, i.e. $\varphi_\pm(j,m) = 0$

$$\Rightarrow\ J_\pm|\psi_{jm}\rangle = \sqrt{(j\mp m)(j\pm m+1)}|\psi_{j,m\pm1}\rangle$$

9.4 Generating all the normalized angular momentum eigenstates

strategy:

For a given j, we start with the eigenstate having the smallest magnetic quantum number, i.e. $|\psi_{j,-j}\rangle$, which is defined as

$$J_-|\psi_{j,-j}\rangle = 0 \quad \text{with} \quad ||\psi_{j,-j}|| = 1\ .$$

We may then obtain all further normalized eigenstates $|\psi_{jm}\rangle$, with $m = -j+1, -j+2, \cdots, +j$, by applying J_+ up to $2j$ times on $|\psi_{j,-j}\rangle$.

- theorem:
$$|\psi_{jm}\rangle = \sqrt{\frac{(j-m)!}{(2j)!(j+m)!}}\ J_+^{j+m}\ |\psi_{j,-j}\rangle\ .$$

  proof by induction:

  (i) We show first that the theorem is valid for $m = -j$

$$|\psi_{j,-j}\rangle = \underbrace{\sqrt{\frac{(2j)!}{(2j)!0!}}\ J_+^0}_{I}\ |\psi_{j,-j}\rangle$$

  (ii) Assuming that the theorem is valid for m, we show that it is valid for $m+1$, as well. Using

$$\begin{aligned} J_+|\psi_{jm}\rangle &= \sqrt{(j-m)(j+m+1)}\ |\psi_{j,m+1}\rangle\ , \\ \text{we have} \quad |\psi_{j,m+1}\rangle &= \frac{1}{\sqrt{(j-m)(j+m+1)}}\ J_+\ |\psi_{jm}\rangle\ . \end{aligned}$$

  Inserting the theorem for $|\psi_{jm}\rangle$, we obtain

$$\begin{aligned} |\psi_{j,m+1}\rangle &= \frac{1}{\sqrt{(j-m)(j+m+1)}}\ J_+\ \sqrt{\frac{(j-m)!}{(2j)!(j+m)!}}\ J_+^{j+m}\ |\psi_{j,-j}\rangle \\ &= \sqrt{\frac{(j-m-1)!}{(2j)!(j+m+1)!}}\ J_+^{j+m+1}\ |\psi_{j,-j}\rangle\ . \end{aligned}$$

  This is indeed the original theorem, with m replaced by $m+1$. q.e.d.

9.5 Matrix elements of J_x, J_y and J_z

- We now calculate the matrix elements of $J_\pm$ and J_z in the orthonormal basis of the eigenstates of J^2 and J_z. The only non-zero matrix elements of $J_\pm$ are off-diagonal, i.e.

$$\begin{aligned}(J_\pm)_{m\pm1,m} &= \langle\psi_{j,m\pm1}|J_\pm\psi_{jm}\rangle \\ &= \sqrt{(j\mp m)(j\pm m+1)}\ \underbrace{\langle\psi_{j,m\pm1}|\psi_{j,m\pm1}\rangle}_{1} \ .\end{aligned}$$

The only non-zero matrix elements of J_z are diagonal.i.e.

$$(J_z)_{m,m} = \langle\psi_{jm}|J_z\psi_{jm}\rangle = m\ \underbrace{\langle\psi_{j,m}|\psi_{j,m}\rangle}_{1} \ .$$

- We may also get the matrix elements of J_x and J_y through

$$J_\pm = J_x \pm i\, J_y \quad \Leftrightarrow \quad \begin{cases} J_x &= \dfrac{1}{2}\ (J_+ + J_-) \\ J_y &= \dfrac{1}{2i}\ (J_+ - J_-) \end{cases}$$

$$\begin{aligned}\Rightarrow \quad (J_x)_{m\pm1,m} &= \langle\psi_{j,m\pm1}|J_x\psi_{jm}\rangle = \frac{1}{2}\ \langle\psi_{j,m\pm1}|J_\pm\psi_{jm}\rangle \\ &= \frac{1}{2}\sqrt{(j\mp m)(j\pm m+1)} \\ (J_y)_{m\pm1,m} &= \langle\psi_{j,m\pm1}|J_y\psi_{jm}\rangle = \frac{\pm1}{2i}\ \langle\psi_{j,m\pm1}|J_\pm\psi_{jm}\rangle \\ &= \frac{\pm1}{2i}\ \sqrt{(j\mp m)(j\pm m+1)} \ .\end{aligned}$$

9.6 Examples

- $\underline{j=\frac{1}{2}}$

$$J_k = \begin{pmatrix} (J_k)_{\frac{1}{2},\frac{1}{2}} & (J_k)_{\frac{1}{2},-\frac{1}{2}} \\ (J_k)_{-\frac{1}{2},\frac{1}{2}} & (J_k)_{-\frac{1}{2},-\frac{1}{2}} \end{pmatrix} \quad k=x,y,z$$

$$\begin{aligned} J_x &= \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\sigma_x}{2} \\ J_y &= \begin{pmatrix} 0 & \frac{1}{2i} \\ -\frac{1}{2i} & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\sigma_y}{2} \\ J_z &= \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\sigma_z}{2} \ . \end{aligned}$$

<u>remark:</u>

The Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ were introduced in early 1925. Pauli seemed to be unaware that William Rowan Hamilton (Irishman, 1805-1865) had introduced similar matrices in 1843, in order to generalize the number concept from Leonhard Euler's complex numbers to William Hamilton's non-commuting hypercomplex numbers or quaternions, represented in 4-dimensional space.

The normalized eigenvectors of J^2 and J_z are for $j=\frac{1}{2}$

$$|\psi_{\frac{1}{2},\frac{1}{2}}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |\psi_{\frac{1}{2},-\frac{1}{2}}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ .$$

The eigenvalue equations are

$$\begin{aligned} J^2|\psi_{\frac{1}{2},m}\rangle &= \frac{3}{4}\,|\psi_{\frac{1}{2},m}\rangle \\ J^z|\psi_{\frac{1}{2},m}\rangle &= m\,|\psi_{\frac{1}{2},m}\rangle \quad m=\frac{1}{2},-\frac{1}{2} \ . \end{aligned}$$

- $j = 1$

$$J_k = \begin{pmatrix} (J_k)_{1,1} & (J_k)_{1,0} & (J_k)_{1,-1} \\ (J_k)_{0,1} & (J_k)_{0,0} & (J_k)_{0,-1} \\ (J_k)_{-1,1} & (J_k)_{-1,0} & (J_k)_{-1,-1} \end{pmatrix} \qquad k = x, y, z$$

new representation:

$$J_x = \begin{pmatrix} 0 & \frac{1}{2}\sqrt{2} & 0 \\ \frac{1}{2}\sqrt{2} & 0 & \frac{1}{2}\sqrt{2} \\ 0 & \frac{1}{2}\sqrt{2} & 0 \end{pmatrix}$$

$$J_y = \begin{pmatrix} 0 & -i\,\frac{1}{2}\sqrt{2} & 0 \\ i\,\frac{1}{2}\sqrt{2} & 0 & -i\,\frac{1}{2}\sqrt{2} \\ 0 & i\,\frac{1}{2}\sqrt{2} & 0 \end{pmatrix}$$

$$J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

diagonal

old Cartesian representation:

$$J'_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

$$J'_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}$$

$$J'_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

non-diagonal

Thus the normalized eigenvectors of J^2 and J_z in the new representation are

$$|\psi_{1,1}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\psi_{1,0}\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |\psi_{1,-1}\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

The eigenvalue equations are

$$\begin{aligned} J^2|\psi_{1,m}\rangle &= j(j+1)|\psi_{1,m}\rangle = 2|\psi_{1,m}\rangle \\ J_z|\psi_{1,m}\rangle &= m|\psi_{1,m}\rangle \text{ with } m = 1, 0, -1 \ . \end{aligned}$$

remark:

Similar spin matrices may be obtained for $j = \frac{3}{2}, 2, \cdots$, etc.

summary:

- A general spin state vector describing the spin j of a particle is given by

$$|\psi_j\rangle = \sum_{m=-j}^{j} c_{j,m}|\psi_{j,m}\rangle = \begin{pmatrix} c_{j,j} \\ c_{j,j-1} \\ \vdots \\ c_{j,-j} \end{pmatrix}$$

with $(2j+1)$ components $c_{j,m} \in \mathcal{C}$.

Here, $|c_{j,m}|^2$ is the probability for finding the particle with a spin projection m in the z-direction, if the spin state vector is normalized as

$$\|\psi_j\|^2 = \langle\psi_j|\psi_j\rangle = \sum_{m=-j}^{j} |c_{j,m}|^2 = 1 \ .$$

9.7 Rotation matrices

- The transformation property of spin state vector $|\psi_j\rangle$, under rotation around the k-axis ($k = x, y, z$) counter clockwise with angle φ, is given by

$$|\psi'_j\rangle = D_k^{(j)}(\varphi)\,|\psi_j\rangle = \mathrm{e}^{-i\varphi\, J_k^{(j)}}\,|\psi_j\rangle, \qquad k = x, y, z$$

Here, $J_k^{(j)}$ are the angular momentum matrices describing spin j and $D_k^{(j)}(\varphi)$ the rotation matrix around the k-axis with angle φ counter clockwise.

examples:

- rotation of a $j = \frac{1}{2}$ spin state vector around the z-axis counter clockwise with angle φ

$$|\psi'_{\frac{1}{2}}\rangle = D_z^{(\frac{1}{2})}(\varphi)\,|\psi_{\frac{1}{2}}\rangle = \exp\left(-i\varphi\, J_z^{(\frac{1}{2})}\right)|\psi_{\frac{1}{2}}\rangle = \exp\left(-i\varphi\,\frac{\sigma_z}{2}\right)|\psi_{\frac{1}{2}}\rangle$$

with $D_z^{(\frac{1}{2})}(\varphi) = \exp\left(-i\varphi\frac{\sigma_z}{2}\right) = \begin{pmatrix} \mathrm{e}^{-\frac{i\varphi}{2}} & 0 \\ 0 & \mathrm{e}^{\frac{i\varphi}{2}} \end{pmatrix}$ and $|\psi_{\frac{1}{2}}\rangle = \begin{pmatrix} c_{\frac{1}{2},\frac{1}{2}} \\ c_{\frac{1}{2},-\frac{1}{2}} \end{pmatrix}$.

(see tuts 3 and 23)

- rotation of a $j = 1$ spin state vector around the z-axis counter clockwise with an angle φ

$$|\psi_1'\rangle = D_z^{(1)}(\varphi)|\psi_1\rangle = \mathrm{e}^{-i\varphi J_z^{(1)}}|\psi_1\rangle$$

with $\quad D_z^{(1)}(\varphi) = \mathrm{e}^{-i\varphi J_z^{(1)}} = \begin{pmatrix} \mathrm{e}^{-i\varphi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mathrm{e}^{+i\varphi} \end{pmatrix}$ and $\quad |\psi_1\rangle = \begin{pmatrix} c_{1,1} \\ c_{1,0} \\ c_{1,-1} \end{pmatrix}$.

This is similar to the rotation matrix of a vector around the z-axis counter clockwise with an angle φ (see tuts 22 and 23)

$$\vec{v}\,' \;=\; R_z(\varphi)\vec{v} \;=\; \mathrm{e}^{-iJ_z'\varphi}\;\vec{v}$$

with $\quad R_z(\varphi) = \mathrm{e}^{-i\varphi J_z'} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$.

(see tuts 12 and 23)

remark:

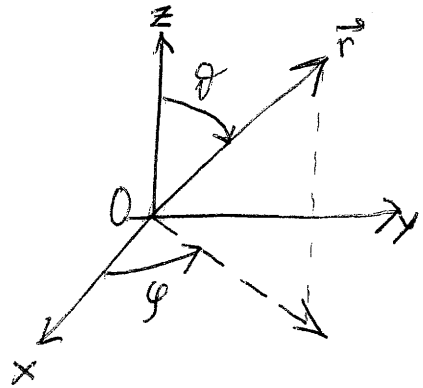
The complex rotation matrix $D_z^{(1)}(\varphi)$ is related to the real rotation matrix $R_z(\varphi)$ via a unitary transformation

$$U\; D_z^{(1)}(\varphi)\; U^{-1} = R_z(\varphi)$$

with $U^\dagger = U^{-1} = \text{const}$ for all φ (see tuts 22 and 23).

10 Spherical harmonics

10.1 Angular momentum operators in spherical coordinates



$$\begin{cases} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{cases}$$

we define the angular momentum

$$\vec{L} \underset{\text{def}}{=} \vec{r}\times\vec{p}/\hbar$$

as a dimensionless quantity.

- theorem: $\qquad L_z = -i\,\dfrac{\partial}{\partial\varphi}$

proof: Apply $\quad -i\,\dfrac{\partial}{\partial\varphi}\quad$ to an arbitrary function $\quad F(x,y,z)$

$$-i\,\frac{\partial}{\partial\varphi}\,F(x,y,z) = -i\,\frac{\partial F}{\partial x}\,\frac{\partial x}{\partial\varphi} - i\,\frac{\partial F}{\partial y}\,\frac{\partial y}{\partial\varphi} - i\,\frac{\partial F}{\partial z}\,\frac{\partial z}{\partial\varphi}$$

$$\begin{cases} \dfrac{\partial x}{\partial\varphi} &= -r\sin\theta\sin\varphi = -y \\ \dfrac{\partial y}{\partial\varphi} &= r\sin\theta\cos\varphi = x \\ \dfrac{\partial z}{\partial\varphi} &= 0 \end{cases}$$

$$\begin{aligned} -i\,\frac{\partial}{\partial\varphi}\,F(x,y,z) &= +i\,\frac{\partial F}{\partial x}\,y \;-\; i\,\frac{\partial F}{\partial y}\,x \\ &= \left(-ix\,\frac{\partial}{\partial y} + iy\,\frac{\partial}{\partial x}\right)F(x,y,z) \end{aligned}$$

$$= \frac{1}{\hbar}\left[x\,\underbrace{\overbrace{\left(-i\hbar\,\frac{\partial}{\partial y}\right)}^{\hbar L_z}}_{p_y} - y\,\underbrace{\left(-i\hbar\,\frac{\partial}{\partial x}\right)}_{p_x}\right]F(x,y,z) \qquad \text{q.e.d.}$$

- theorem: $$(L_x \pm i\, L_y) = \mathrm{e}^{\pm\, i\varphi} \left[\pm\, \frac{\partial}{\partial\theta} + i\, \cot\, \theta\, \frac{\partial}{\partial\varphi} \right]$$

proof: Apply these differential operators to an arbitrary function

$$F(x,y,z) = F\,(r \sin\,\theta\,\cos\,\varphi,\; r\,\sin\,\theta\,\sin\,\varphi,\; r\,\cos\,\theta)$$

$$\begin{aligned}
&\mathrm{e}^{\pm i\varphi} \left[\pm\, \frac{\partial}{\partial\theta} + i\, \cot\,\theta\, \frac{\partial}{\partial\varphi} \right] F(x,y,z) \\
&= \pm\, \mathrm{e}^{\pm i\varphi} \left[\frac{\partial F}{\partial x}\, r\,\cos\,\theta\,\cos\,\varphi + \frac{\partial F}{\partial y}\, r\,\cos\,\theta \sin\varphi - \frac{\partial F}{\partial z}\,\sin\,\theta \right] \\
&\quad + \mathrm{e}^{\pm i\varphi}\, i\, \frac{\cos\,\theta}{\sin\,\theta} \left[-\frac{\partial F}{\partial x}\, r\,\sin\,\theta\,\sin\,\varphi + \frac{\partial F}{\partial y}\, r\,\sin\,\theta\,\cos\,\varphi \right] \\
&= \frac{\partial F}{\partial x}\, \mathrm{e}^{\pm i\varphi}\, r\,\cos\,\theta\,(\pm\cos\,\varphi - i\,\sin\,\varphi) \\
&\quad + \frac{\partial F}{\partial y}\, \mathrm{e}^{\pm i\varphi}\, r\,\cos\,\theta\,(\pm\sin\,\varphi + i\,\cos\,\varphi) \\
&\quad + \frac{\partial F}{\partial z}\,(\mp\, \mathrm{e}^{\mp i\varphi}\, r\,\sin\,\theta) \\
&= \pm\, \frac{\partial F}{\partial x}\, r\,\cos\,\varphi + i\, \frac{\partial F}{\partial y}\, r\,\cos\,\theta + \frac{\partial F}{\partial z}\,[\mp\,(\cos\,\varphi \pm i\,\sin\,\varphi)\; r\,\sin\,\theta] \\
&= \underbrace{\pm\, z\, \frac{\partial F}{\partial x} \mp x\, \frac{dF}{dz}}_{\pm\, i\, L_y\, F} + \underbrace{i\, z\, \frac{\partial F}{\partial y} - iy\, \frac{\partial F}{\partial z}}_{L_x\, F} \qquad \text{q.e.d.}
\end{aligned}$$

10.2 Casimir operator L^2 in spherical coordinates

We now want to calculate L^2 using the mathematical identity

$$L^2 = \frac{1}{2} \underbrace{\Big(L_x + i\, L_y\Big)\Big(L_x - i\, L_y\Big)}_{L_+ L_-} + \frac{1}{2} \underbrace{\Big(L_x - i\, L_y\Big)\Big(L_x + i\, L_y\Big)}_{L_- L_+} + L_z^2 \;.$$

We first calculate

$$\begin{aligned}
L_+ L_- &= \mathrm{e}^{i\varphi} \left[\frac{\partial}{\partial\theta} + i\,\cot\,\theta\, \frac{\partial}{\partial\varphi} \right] \mathrm{e}^{-i\varphi} \left[-\frac{\partial}{\partial\theta} + i\,\cot\,\theta\, \frac{\partial}{\partial\varphi} \right] \\
&= -\mathrm{e}^{i\varphi} \left[-i\, \frac{\partial}{\partial\theta} + \cot\,\theta\, \frac{\partial}{\partial\varphi} \right] \mathrm{e}^{-i\varphi} \left[i\, \frac{\partial}{\partial\theta} + \cot\,\theta\, \frac{\partial}{\partial\varphi} \right] \;.
\end{aligned}$$

Using the operator identity

$$\frac{\partial}{\partial\varphi}\, \mathrm{e}^{-i\varphi} \;=\; \mathrm{e}^{-i\varphi} \left(-i + \frac{\partial}{\partial\varphi} \right) \;,$$

one can shift $\mathrm{e}^{-i\varphi}$ to the left and obtain

$$L_+ L_- = -\mathrm{e}^{i\varphi}\, \mathrm{e}^{-i\varphi} \left(-i\, \frac{\partial}{\partial\theta} + \cot\,\theta \left(-i + \frac{\partial}{\partial\varphi} \right) \right) \left(i\, \frac{\partial}{\partial\theta} + \cot\,\theta\, \frac{\partial}{\partial\varphi} \right) \;,$$

or

$$L_+ L_- = -\left[\frac{\partial^2}{\partial\theta^2} + \cot^2\,\theta \left(-i + \frac{\partial}{\partial\varphi} \right) \frac{\partial}{\partial\varphi} - i\, \frac{\partial}{\partial\theta} \cot\,\theta\, \frac{\partial}{\partial\varphi} + \cot\,\theta \left(-i + \frac{\partial}{\partial\varphi} \right) i\, \frac{\partial}{\partial\theta} \right] \;.$$

Using the operator identity

$$\frac{\partial}{\partial\theta}\, \cot\,\theta \;=\; -\frac{1}{\sin^2\theta} + \cot\,\theta\, \frac{\partial}{\partial\theta} \;,$$

we obtain

$$\begin{aligned}
L_+ L_- = -\Bigg[&\frac{\partial^2}{\partial\theta^2} \;+\; \cot^2\,\theta\, \frac{\partial^2}{\partial\varphi^2} - i\,\cot^2\,\theta\, \frac{\partial}{\partial\varphi} + i\, \frac{1}{\sin^2\theta}\, \frac{\partial}{\partial\varphi} - i\,\cot\,\theta\, \frac{\partial}{\partial\theta}\, \frac{\partial}{\partial\varphi} \\
&+ \cot\,\theta\, \frac{\partial}{\partial\theta} \;+\; i\,\cot\,\theta\, \frac{\partial}{\partial\theta}\, \frac{\partial}{\partial\varphi} \Bigg]
\end{aligned}$$

$$L_+ L_- = -\left[\frac{\partial^2}{\partial\theta^2} + i\, \frac{1 - \cos^2\theta}{\sin^2\theta}\, \frac{\partial}{\partial\varphi} + \cot\,\theta\, \frac{\partial}{\partial\theta} + \cot^2\,\theta\, \frac{\partial^2}{\partial\varphi^2} \right]$$

$$L_+ L_- = -\left[\frac{\partial^2}{\partial\theta^2} + i\, \frac{\partial}{\partial\varphi} + \cot\,\theta\, \frac{\partial}{\partial\theta} + \cot^2\,\theta\, \frac{\partial^2}{\partial\varphi^2} \right] \;.$$

Similarly, we arrive at

$$L_- L_+ = -\left[\frac{\partial^2}{\partial\theta^2} - i\, \frac{\partial}{\partial\varphi} + \cot\,\theta\, \frac{\partial}{\partial\theta} + \cot^2\,\theta\, \frac{\partial^2}{\partial\varphi^2} \right] \;.$$

The final result is thus

$$L^2 \;=\; -\left[\frac{\partial^2}{\partial\theta^2} + \cot\,\theta\, \frac{\partial}{\partial\theta} + \left(1 + \cot^2\,\theta\right)\, \frac{\partial^2}{\partial\varphi^2} \right] \;.$$

or

$$L^2 \;=\; -\left[\frac{1}{\sin\theta}\, \frac{\partial}{\partial\theta} \left(\sin\,\theta\, \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta}\, \frac{\partial^2}{\partial\varphi^2} \right] \;.$$

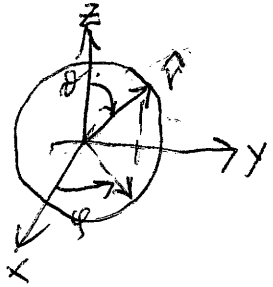
10.3 Orbital angular momentum wave functions

- Based on the canonical commutation relations, we have shown in tut 11 that $\vec{L} = \vec{r} \times \vec{p}/\hbar$ fulfils the Lie algebra of $SO(3,\mathcal{R})$, i.e. $[L_x, L_y] = iL_z$, and cyclic permutations thereof. From chapter 9, we know that the eigenvalues and eigenvectors of $L^2 = L_x^2 + L_y^2 + L_z^2$ and L_z must obey

$$\begin{aligned} L^2|Y_{lm}\rangle &= l(l+1)|Y_{lm}\rangle \\ L_z|Y_{lm}\rangle &= m|Y_{lm}\rangle \end{aligned}$$

$$\begin{aligned} \text{with } m &= -l, -l+1, \cdots, +l \qquad (2l+1) \text{ values} \\ \text{and } l &= 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \cdots \text{ in principle.} \end{aligned}$$

- From sections 10.1 and 10.2, we know that L_x, L_y, L_z can be written in spherical coordinates in terms of θ, φ and their partial derivatives $\frac{\partial}{\partial\theta}$, $\frac{\partial}{\partial\varphi}$ only. Thus one must be able to interpret the eigenfunctions in terms of θ and φ only, defined on the unit sphere.



$\hat{r} = \vec{r}/r = \{\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta\}$

unit vector in $\vec{r}$-direction

$\Rightarrow \quad |Y_{lm}\rangle = Y_{lm}(\theta,\varphi)$

We have shown in chapter 9 and section 10.2

$$L_z Y_{lm}(\theta,\varphi) = -i\frac{\partial}{\partial\varphi} Y_{lm}(\theta,\varphi) = m\, Y_{lm}(\theta,\varphi) \qquad \text{eq.(1)}$$

$$L^2 Y_{lm}(\theta,\varphi) = -\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] Y_{lm}(\theta,\varphi) \qquad \text{eq.(2)}$$

$$= l(l+1)\, Y_{lm}(\theta,\varphi)$$

The most general solution to eq.(1) is $Y_{lm}(\theta,\varphi) = A_{lm}(\theta)\, e^{im\varphi}$. There are two possible periodicity conditions for $Y_{lm}(\theta,\varphi)$ consistent with the Lie algebra of $SO(3,\mathcal{R})$:

"bosonic" periodicity	"fermionic" periodicity
$Y_{lm}(\theta,\varphi+2\pi) = Y_{lm}(\theta,\varphi)$	$Y_{lm}(\theta,\varphi+2\pi) = -Y_{lm}(\theta,\varphi)$
$\Rightarrow \; e^{im(\varphi+2\pi)} = e^{im\varphi}$	$e^{im(\varphi+2\pi)} = -e^{im\varphi}$
$e^{2\pi im} = 1$	$e^{2\pi im} = -1$
$m = 0, \pm1, \pm2, \cdots$	$m = \pm\frac{1}{2}, \pm\frac{3}{2}, \cdots$
$l = 0, 1, 2, \cdots$	$l = \frac{1}{2}, \frac{3}{2}, \cdots$
Single-valued with $Y_{lm}(\theta,\varphi) = Y_{lm}(\theta,\varphi+2\pi)$.	Double-valued with $Y_{lm}(\theta,\varphi) = -Y_{lm}(\theta,\varphi+2\pi)$ and $Y_{lm}(\theta,\varphi+4\pi) = Y_{lm}(\theta,\varphi)$.
Because it reaches the same value after a rotation around z-axis with an angle 2π, it is uniquely defined on the unit sphere.	Because it reaches the same value only after a rotation around z-axis with an angle 4π, it is double valued on the unit sphere.

conclusion:
If we insist that $Y_{lm}(\theta,\varphi)$ should be uniquely defined on the unit sphere for all $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \varphi < 2\pi$, then $l = 0, 1, 2, \cdots$ and $m = 0, \pm1, \pm2, \cdots, \pm l$ are the only possible values for l and m. However, if we admit double-valued functions, then half-integer values of l and m are allowed, as well.

10.4 Construction of the spherical harmonics

- eigenfunction with lowest m-value for given l

We start with the eigenfunction with the smallest m-eigenvalue for a given l, which is defined as

$$L_-\, Y_{l,-l}(\theta,\varphi) = 0 \; .$$

For $m = -l$ we have $Y_{l,-l}(\theta,\varphi) = A_{l,-l}(\theta)\, e^{-il\varphi}$

and L is given by

$$L_- = L_x - i\, L_y = e^{-i\varphi}\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right) .$$

We thus arrive at the differential equation

$$e^{-i\varphi}\left(-\frac{\partial}{\partial\theta}A_{l,-l}(\theta)+l\cot\theta\,A_{l,-l}(\theta)\right)e^{-il\varphi}=0$$

$$\text{or}\quad \frac{d}{d\theta}A_{l,-l}(\theta)-l\cot\theta\,A_{l,-l}(\theta)=0$$

$$\text{or}\quad \frac{1}{A_{l,-l}(\theta)}\frac{d}{d\theta}A_{l,-l}(\theta)=l\cot\theta\ .$$

solution: $\ln A_{l,-l}(\theta)=l\ln\sin\theta+\text{const}$ eq.(3)

proof:

$$\frac{d}{d\theta}\ln A_{l,-l}(\theta)=\frac{A'_{l,-l}(\theta)}{A_{l,-l}(\theta)}=\frac{d}{d\theta}(l\ln\sin\theta+\text{const})$$
$$=l\frac{\cos\theta}{\sin\theta}=l\cot\theta \qquad \text{q.e.d.}$$

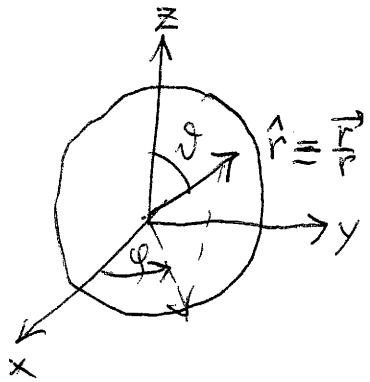
Taking the exponent on both sides of eq.(3) yields

$$A_{l,-l}(\theta)=(\sin\theta)^l\,e^{\text{const}}=a(\sin\theta)^l\ .$$

- normalization of $Y_{l,-l}(\theta,\varphi)$:

$$Y_{l,-l}(\theta,\varphi)=a\,(\sin\theta)^l\,e^{-il\varphi}$$

- calculation of the normalization constant:



The surface of the unit sphere is

$$\int_0^\pi \sin\theta\,d\theta\int_0^{2\pi}d\varphi=2\pi(-\cos\theta)\Big|_0^\pi$$
$$=2\pi(1+1)=4\pi\ .$$

Integrating $|Y_{l,-l}(\theta,\varphi)|^2$ over the unit sphere, we arrive at

$$\int_0^\pi \sin\theta d\theta\int_0^{2\pi}d\varphi\left|Y_{l,-l}(\theta,\varphi)\right|^2=2\pi|a|^2\underbrace{\int_0^\pi(\sin\theta)^{2l+1}\,d\theta}_{\frac{2(l!\,2^l)^2}{(2l+1)!}}=\frac{4\pi|a|^2\,(l!\,2^l)^2}{(2l+1)!}=1$$

$$\Rightarrow\qquad a=\frac{1}{\sqrt{4\pi}}\frac{\sqrt{(2l+1)!}}{2^l\,l!}$$

remark:

Here, we have chosen the arbitrary phase of a to be zero, yielding

$$\Rightarrow\quad Y_{l,-l}(\theta,\varphi)=\overbrace{\frac{1}{\sqrt{4\pi}}\frac{\sqrt{(2l+1)!}}{2^l l!}}^{A_{l,-l}(\theta)}(\sin\theta)^l\,e^{-il\varphi}$$

- building up higher m-quantum numbers for a given l value:

We have obtained in section 9.3

$$Y_{lm}(\theta,\varphi)=\sqrt{\frac{(l-m)!}{(2l)!(l+m)!}}\,L_+^{l+m}\,Y_{l,-l}(\theta,\varphi)\ .\qquad \text{eq.(4)}$$

Let $A(\theta)$ be an arbitrary function of θ

$$L_+\left(A(\theta)\,e^{-il\varphi}\right)=\underbrace{e^{i\varphi}\left(\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\varphi}\right)}_{\text{from section 10.1}}\left(A(\theta)\,e^{-il\varphi}\right)$$
$$=e^{-i(l-1)\varphi}\left[\frac{d}{d\theta}+l\cot\theta\right]A(\theta)$$

theorem:

$$\left[\frac{d}{d\theta}+l\cot\theta\right]A(\theta)=\frac{1}{(\sin\theta)^l}\frac{d}{d\theta}\left[(\sin\theta)^l\,A(\theta)\right]$$

proof:

$$\frac{1}{(\sin\theta)^l}\frac{d}{d\theta}\left[(\sin\theta)^l\,A(\theta)\right]=\frac{1}{(\sin\theta)^l}\left[l\,(\sin\theta)^{l-1}\cos\theta\,A(\theta)+(\sin\theta)^l\frac{dA}{d\theta}\right]$$
$$=\left[l\cot\theta+\frac{d}{d\theta}\right]A(\theta)\qquad \text{q.e.d.}$$

$$\Rightarrow L_+ \left(A(\theta)\, e^{-il\varphi}\right) = e^{-i(l-1)\varphi} \frac{1}{(\sin\theta)^l} \frac{d}{d\theta} \left[(\sin\theta)^l\, A(\theta)\right]$$

$$= -e^{-i(l-1)\varphi} \frac{1}{(\sin\theta)^{l-1}} \frac{d}{d\cos\theta} \left[(\sin\theta)^l\, A(\theta)\right] ,$$

where we have used $d\cos\theta = -\sin\theta\, d\theta$

- Applying this trick n times, we have

$$L_+^n \left(A(\theta)\, e^{-il\varphi}\right) = (-1)^n\, e^{-i(l-n)\varphi} \frac{1}{(\sin\theta)^{l-n}} \left(\frac{d}{d\cos\theta}\right)^n \left[(\sin\theta)^l\, A(\theta)\right] \quad \text{eq.(5)}$$

- Setting $n = l + m$ and using eqs.(4), (5) and

$$A(\theta) = A_{l-l}(\theta) = \frac{1}{\sqrt{4\pi}} \frac{\sqrt{(2l+1)!}}{2^l\, l!} (\sin\theta)^l ,$$

we obtain

$$Y_{l,m}(\theta,\varphi) = \frac{(-1)^{l+m}}{2^l\, l!} \sqrt{\frac{2l+1}{4\pi}} \sqrt{\frac{(l-m)!}{(l+m)!}}$$

$$e^{im\varphi} (\sin\theta)^m \left(\frac{d}{d\cos\theta}\right)^{l+m} (\sin\theta)^{2l} .$$

10.5 Properties of the spherical harmonics

(i) Using Leibniz's rule, one can derive the conjugation property of the spherical harmonics, i.e.

$$Y_{lm}^*(\theta,\varphi) = (-1)^m\, Y_{l-m}(\theta,\varphi) .$$

(ii) Under space reflection

$$\vec{r} \to \vec{r}\,' = P\vec{r} = -\vec{r} \quad \text{with} \quad P = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

or in spherical coordinates

$$(r,\theta,\varphi) \to (r',\theta',\varphi') = (r, \pi = \theta, \varphi + \pi) ,$$

the transformation property of $Y_{lm}(\theta,\varphi)$ is

$$P\, Y_{lm}(\theta,\varphi) = Y_{lm}(\pi - \theta, \varphi + \pi) = (-1)^l\, Y_{lm}(\theta,\varphi) = p\, Y_{lm}(\theta,\varphi) ,$$

with $p = (-1)^l$ the parity quantum number.

(iii) The spherical harmonics are orthonormal:

$$\langle Y_{lm} | Y_{l'm'} \rangle \underset{\text{def}}{=} \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta\, Y_{lm}^*(\theta,\varphi)\, Y_{l'm'}(\theta,\varphi) = \delta_{ll'}\, \delta_{mm'} \quad \text{and}$$

(iv) complete:

Any function $T(\theta,\varphi)$, defined on the unit sphere, e.g. the temperature of the microwave background radiation, can be expanded as

$$T(\theta,\varphi) = \sum_{l,m=0}^{\infty} c_{lm}\, Y_{lm}(\theta,\varphi)$$

with $c_{lm} = \langle Y_{lm} | T \rangle = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta\, d\theta\, Y_{lm}^*(\theta,\varphi)\, T(\theta,\varphi)$ Fourier coefficient.

(v) The eigenvalue equations are

$$L^2 Y_{lm}(\theta,\varphi) = -\left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2}\right] Y_{lm}(\theta,\varphi)$$

$$= l(l+1)\, Y_{lm}(\theta,\varphi) \qquad l = 0, 1, 2, \cdots$$

$$L_z\, Y_{lm}(\theta,\varphi) = -i \frac{\partial}{\partial\varphi} Y_{lm}(\theta,\varphi)$$

$$= m\, Y_{lm}(\theta,\varphi) \qquad m = -l, -l+1, \cdots, l .$$

(vi) The actions of the step or ladder operators on a spherical harmonic is

$$L_\pm\, Y_{lm}(\theta,\varphi) = e^{\pm i\varphi} \left(\pm \frac{\partial}{\partial\theta} + i \cot\theta \frac{\partial}{\partial\varphi}\right) Y_{lm}(\theta,\varphi)$$

$$= \sqrt{(l \mp m)(l \pm m + 1)}\, Y_{lm+1}(\theta,\varphi) .$$

10.6 Vectors, tensors and spinors

We have discussed an extraordinary variety of geometrical objects with peculiar transformation properties under rotations in $\mathcal{E}_3$. We need to categorize their properties.

- spherical tensors and spinors

A column of $(2j+1)$ complex numbers

$$|\psi_j\rangle = \begin{pmatrix} c_{j,j} \\ c_{j,j-1} \\ \vdots \\ c_{j,-j} \end{pmatrix} ,$$

transforming under rotations, around the k-axis with an angle φ counter clockwise, according to $|\psi'_j\rangle = D^{(j)}_k(\varphi)\,|\psi_j\rangle$ with $(k = x, y, z)$ is called a spherical tensor or a spinor as follows:

(a) a spherical tensor if j=0, 1, 2, $\cdots$ non-negative integer

examples:

scalar ($j = 0$) $\equiv$ invariant, spherical tensor of rank 0; 1 component

spherical vector ($j = 1$) $\equiv$ spherical tensor of rank 1; 3 components

spherical tensor ($j = 2$) $\equiv$ spherical tensor of rank 2; 5 components

(b) a spinor if $j = \frac{1}{2}, \frac{3}{2} \cdots$ non-negative half-integer

examples:

spinor ($j = \frac{1}{2}$); 2 components

spinor ($j = \frac{3}{2}$); 4 components

remark: Under rotations, the transformation properties of spinors differ from those of spherical tensors. After a rotation of a spinor around any axis with an angle 2π, we will end up with the negative spinor. Only after a rotation of the spinor by 4π do we get the original spinor back. Spinors are thus representations of $SU(2, \mathcal{C})$ rather than $SO(3, \mathcal{R})$. $j = \frac{1}{2}$ spinors rotate at half the rate of $j = 1$ spherical and Cartesian vectors. However, $SU(2, \mathcal{C})$ and $SO(3, \mathcal{R})$ share the same Lie-algebra.

- Cartesian tensors

A set of 3^r ($r \in \mathcal{N}_0$) real numbers, transforming under rotations as the r-fold products of Cartesian coordinates, is a Cartesian tensor of rank r.

examples:

- scalar $\equiv$ invariant, or Cartesian tensor of rank r=0 (1 component), does not transform;
- Cartesian vector, or Cartesian tensor of rank r=1 (3 components), transforms like x, y, z;
- Cartesian tensor of rank r=2 (9 components), transforms like the products of two coordinates, i.e.

$$\begin{matrix} xx, xy, xz \\ yx, yy, yz \\ zx, zy, zz \end{matrix}$$

e.g. moment of inertia $\theta_{xx}, \theta_{xy}, \cdots$

remark: Cartesian tensors are representations of $SO(3, \mathcal{R})$ rather than $SU(2, \mathcal{C})$.

- relation between spherical and Cartesian tensors: spherical harmonics

It is difficult to compare the transformation properties of vectors, tensors and spinors. But what we can do, is look at the spherical harmonics $Y_{lm}(\theta, \varphi)$, because these transform like spherical tensors, and they can also be written in terms of Cartesian tensor components.

examples:

$$\begin{cases} x &= r \sin\theta \cos\varphi \\ y &= r \sin\theta \sin\varphi \\ z &= r \cos\theta \end{cases}$$

scalar, spherical tensor of rank 0:

$$Y_{0,0}(\theta, \varphi) = \frac{1}{2\sqrt{\pi}} \qquad \text{scalar}$$

Cartesian vector, spherical vector or spherical tensor of rank 1:

$$Y_{1,0}(\theta, \varphi) = \frac{1}{2}\sqrt{\frac{3}{\pi}}\,\cos\theta = \frac{1}{2}\sqrt{\frac{3z}{\pi r}}$$

$$Y_{1,\pm1}(\theta, \varphi) = \mp\frac{1}{2}\sqrt{\frac{3}{2\pi}}\,\sin\theta\,\mathrm{e}^{\pm i\vartheta} = \mp\frac{1}{2}\sqrt{\frac{3}{2\pi}}\,\frac{x \pm iy}{r}$$

Cartesian tensor or spherical tensor of rank 2:

$$Y_{2,0}(\theta, \varphi) = \frac{1}{4}\sqrt{\frac{5}{\pi}}\,(2\cos^2\theta - \sin^2\varphi) = \frac{1}{4}\sqrt{\frac{5}{\pi}}\,\frac{2z^2 - x^2 - y^2}{r^2}$$

$$Y_{2,\pm1}(\theta, \varphi) = \mp\frac{1}{2}\sqrt{\frac{15}{2\pi}}\,\cos\theta\,\sin\theta\,\mathrm{e}^{\pm i\varphi} = \mp\frac{1}{2}\sqrt{\frac{15}{2\pi}}\,\frac{z(x \pm iy)}{r^2}$$

$$Y_{2,\pm2}(\theta, \varphi) = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\,\sin^2\theta\,\mathrm{e}^{\pm 2i\varphi} = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\,\frac{(x \pm iy)^2}{r^2}\ .$$

note: Under rotations r is invariant.

11 Two-body problem

Take a classical Hamiltonian for two distinguishable interacting particles

$$H = \frac{\vec{p}_1{}^2}{2m_1} + \frac{\vec{p}_2{}^2}{2m_2} + V(\vec{r}_1, \vec{r}_2, t),$$

subject to a time-dependent potential energy. The transition to quantum mechanical theory is obtained as

$$\vec{p}_1 \to -i\hbar\vec{\nabla}_1 \ , \qquad \vec{p}_2 \to -i\hbar\vec{\nabla}_2 \ .$$

The Hamilton operator is therefore

$$H = -\frac{\hbar^2}{2m_1}\Delta_1 - \frac{\hbar^2}{2m_2}\Delta_2 + V(\vec{r}_1, \vec{r}_2, t)$$

with the Laplacians

$$\Delta_1 \equiv \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2} \quad mboxand \qquad \Delta_2 \equiv \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2} \ ,$$

and the time-dependent Schrödinger equation for two interacting particles

$$H\,\psi(\vec{r}_1, \vec{r}_2, t) = i\hbar\frac{\partial}{\partial t}\,\psi(\vec{r}_1, \vec{r}_2, t) \ .$$

11.1 Constants of motion

(a) energy: If $V(\vec{r}_1, \vec{r}_2, t) = V(\vec{r}_1, \vec{r}_2)$ is time-independent, then

(i) the Hamiltonian has time-translation symmetry

(ii) time is homogeneous

(iii) energy is conserved

(iv) space and time variables can be separated in the Schrödinger eq., i.e.

ansatz: $\psi(\vec{r}_1, \vec{r}_2, t) = \chi(\vec{r}_1, \vec{r}_2)\,v(t)$

$$\begin{aligned} H\,\psi(\vec{r}_1, \vec{r}_2, t) &= v(t)\,H\,\chi(\vec{r}_1, \vec{r}_2) \\ &= i\hbar\,\chi(\vec{r}_1, \vec{r}_2)\frac{\partial v(t)}{\partial t} \end{aligned} \quad \Big| \text{divide by } \psi$$

$$\underbrace{\frac{1}{\chi(\vec{r}_1, \vec{r}_2)}\,H\,\chi(\vec{r}_1, \vec{r}_2)}_{\substack{F(\vec{r}_1, \vec{r}_2) \\ \text{function of} \\ \text{coordinates only}}} = \underbrace{i\hbar\,\frac{1}{v(t)}\,\frac{\partial v(t)}{\partial t}}_{\substack{f(t) \\ \text{function of} \\ \text{time only}}} = E = \text{const}$$

conservation of energy

$$\Rightarrow \begin{cases} i\hbar\,\dfrac{\partial v(t)}{\partial t} = E\,v(t) & \text{can be solved trivially} \\ & \text{solution} v(t) = v(0)\,\mathrm{e}^{-\frac{iE}{\hbar}\,t} \\ H\,\chi(\vec{r}_1, \vec{r}_2) = E\,\chi(\vec{r}_1, \vec{r}_2) & \text{solution less trivial} \end{cases}$$

time-independent Schrödinger equation for two interacting particles

(b) momentum (or centre-of-mass):

$$H = -\frac{\hbar^2}{2m_1}\Delta_1 - \frac{\hbar^2}{2m_2}\Delta_2 + V(\vec{r}_1, \vec{r}_2)$$

$H\,\chi(\vec{r}_1, \vec{r}_2) = E\,\chi(\vec{r}_1, \vec{r}_2)$ time-independent Schrödinger equation

we introduce

new coordinates:

$$\begin{aligned} X &= \frac{m_1x_1 + m_2x_2}{m_1 + m_2} \ , \quad Y = \frac{m_1y_1 + m_2y_2}{m_1 + m_2} \ , \ \ldots && \text{centre-of-mass coordinates} \\ x &= x_2 - x_1 \ , \quad y = y_2 - y_1 \ , \ \ldots && \text{relative coordinates} \\ M &= m_1 + m_2 \ , \quad m = \frac{m_1m_2}{m_1 + m_2} \end{aligned}$$

total mass reduced mass

theorem: The kinetic energy term can be written as

$$\begin{aligned} -\frac{\hbar^2}{2m_1}\Delta_1 - \frac{\hbar^2}{2m_2}\Delta_2 &= -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}\right) \\ &\quad -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \ . \end{aligned}$$

proof: $x_1, x_2 \to X, x$ coordinate transformation

$$\begin{aligned} \frac{\partial}{\partial x_1} &= \frac{\partial X}{\partial x_1}\frac{\partial}{\partial X} + \frac{\partial x}{\partial x_1}\frac{\partial}{\partial x} \\ \frac{\partial}{\partial x_1} &= \frac{m_1}{m_1 + m_2}\frac{\partial}{\partial X} - \frac{\partial}{\partial x} \\ \frac{\partial^2}{\partial x_1^2} &= \frac{m_1^2}{(m_1 + m_2)^2}\frac{\partial^2}{\partial X^2} - \frac{2m_1}{m_1 + m_2}\frac{\partial}{\partial X}\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2} \\ \frac{\partial}{\partial x_2} &= \frac{\partial X}{\partial x_2}\frac{\partial}{\partial X} + \frac{\partial x}{\partial x_2}\frac{\partial}{\partial x} \\ \frac{\partial}{\partial x_2} &= \frac{m_2}{m_1 + m_2}\frac{\partial}{\partial X} + \frac{\partial}{\partial x} \end{aligned}$$

$$\frac{\partial^2}{\partial x_2^2} = \frac{m_2^2}{(m_1+m_2)^2}\frac{\partial^2}{\partial X^2} + \frac{2m_2}{m_1+m_2}\frac{\partial}{\partial X}\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2}$$

$$\begin{aligned}\frac{1}{m_1}\frac{\partial^2}{\partial x_1^2} + \frac{1}{m_2}\frac{\partial^2}{\partial x_2^2} &= \frac{m_1}{(m_1+m_2)^2}\frac{\partial^2}{\partial X^2} - \frac{2}{m_1+m_2}\frac{\partial}{\partial X}\frac{\partial}{\partial x} + \frac{1}{m_1}\frac{\partial^2}{\partial x^2} \\ &\quad + \frac{m_2}{(m_1+m_2)^2}\frac{\partial^2}{\partial X^2} + \frac{2}{m_1+m_2}\frac{\partial}{\partial X}\frac{\partial}{\partial x} + \frac{1}{m_2}\frac{\partial^2}{\partial x^2} \\ &= \frac{1}{M}\frac{\partial^2}{\partial X^2} + \underbrace{\left(\frac{1}{m_1}+\frac{1}{m_2}\right)}_{\frac{1}{m}}\frac{\partial^2}{\partial x^2}\end{aligned}$$

q.e.d.

If the potential energy is a function of the relative coordinates only, i.e.

$$V(\vec{r}_1,\vec{r}_2) = V(\vec{r}_1-\vec{r}_2) = V(x,y,z) \ ,$$

we have

(i) actio = reactio is valid

(ii) the Hamiltonian has space-translation symmetry

(iii) space is homogeneous

(iv) momentum (or centre-of-mass) conserved

(v) variables $\vec{R}$ and $\vec{r}$ can be separated in the Schrödinger eq.

We thus have the Hamilton operator

$$H = H_{\rm CM} + H_{\rm rel}$$

with

$$H_{\rm CM} = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2}+\frac{\partial^2}{\partial Z^2}\right) \qquad \text{CM kinetic energy term}$$

$$H_{\rm rel} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right) + V(x,y,z) \qquad \begin{array}{l}\text{relative kinetic and}\\ \text{potential energy terms}\end{array}$$

$$H\,\chi(\vec{r}_1,\vec{r}_2) = E\,\chi(\vec{r}_1,\vec{r}_2) \qquad \text{time-independent Schrödinger equation}$$

ansatz: $\chi(\vec{r}_1,\vec{r}_2) = U(\vec{R})\,u(\vec{r})$ separation of CM and relative variables

$$\begin{aligned}H\,\chi(\vec{r}_1,\vec{r}_2) &= u(\vec{r})\,H_{\rm CM}\,U(\vec{R}) + U(\vec{R})\,H_{\rm rel}\,u(\vec{r}) \\ &= E\,U(\vec{R})\,u(\vec{r}) \qquad \left|\begin{array}{l}\text{divide}\\ \text{by } \chi\end{array}\right.\end{aligned}$$

$$\begin{aligned}\underbrace{\frac{1}{U}H_{\rm CM}\,U}_{\begin{array}{l}\text{depends on}\\ \vec{R}\text{ only}\\ \Rightarrow \text{const}\\ E_{\rm CM}\end{array}} + \underbrace{\frac{1}{u}H_{\rm rel}\,u}_{\begin{array}{l}\text{depends on}\\ \vec{r}\text{ only}\\ \Rightarrow \text{const}\\ E_{\rm rel}\end{array}} &= E \\ -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial X^2}+\frac{\partial^2}{\partial Y^2}+\frac{\partial^2}{\partial Z^2}\right)U(X,Y,Z) &= E_{\rm CM}\,U(X,Y,Z) \\ \text{with} \qquad E_{\rm CM} &= \frac{\hbar^2K^2}{2M} \\ \text{solution trivial} \qquad U(\vec{R}) &= U(0)\,e^{\pm i\vec{K}\cdot\vec{R}} \\ \left(-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+V(x,y,z)\right)u(x,y,z) &= E_{\rm rel}\,u(x,y,z) \\ \text{solution less trivial}\end{aligned}$$

conclusion: Similar to the classical two-body problem, the quantum mechanical two-body problem can be reduced to the one-body problem, introducing centre-of-mass and relative coordinates, the total mass M and the reduced mass m. For this to be valid space-time, it has to be homogeneous, i.e. invariant under the transformations

$$\begin{aligned}\vec{r}_i &\to \vec{r}_i\,' = \vec{r}_i + \vec{a} \qquad\qquad i = 1,2 \\ t &\to t' = t+\tau \ .\end{aligned}$$

(c) angular momentum:

If $V(\vec{r}) = V(r)$ does not depend on the orientation in space, then

(i) the Hamiltonian has rotational symmetry

(ii) space is isotropic

(iii) angular momentum is conserved

(iv) angular (ϑ,φ) and radial (r) variables can be separated in the Schrödinger eq.

In spherical coordinates we obtain

$$\left\{\begin{array}{lcl} x &=& r\,\sin\vartheta\,\cos\varphi \\ y &=& r\,\sin\vartheta\,\sin\varphi \\ z &=& r\,\cos\vartheta \end{array}\right. \ .$$

Using the canonical commutation relations, we have proven in tut 22, the identity

$$\hbar^2\,L^2 = (\vec{r}\times\vec{p})\cdot(\vec{r}\times\vec{p}) = r^2p^2 - (\vec{r}\cdot\vec{p})^2 + i\hbar\vec{r}\cdot\vec{p}\,.$$

Inserting

$$\begin{aligned} \vec{p} &= -i\hbar\,\vec{\nabla} \quad \text{and} \\ \vec{r}\cdot\vec{p} &= -i\hbar\,\vec{r}\cdot\vec{\nabla} = -i\hbar\,r\frac{\partial}{\partial r} \quad , \end{aligned}$$

we arrive at

$$\begin{aligned} \Delta &= \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2} \\ &= \frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}-\frac{L^2}{r^2} \end{aligned}$$

with

$$L^2 = -\left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\,\frac{\partial}{\partial\vartheta}\right)+\frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}\right] \quad .$$

The time-independent Schrödinger equation thus becomes

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}-\frac{L^2}{r^2}\right)u+V(r)\,u=E_{\rm rel}\,u \quad .$$

Separation of the radial and angular parts yields

$$u(r,\vartheta,\varphi)=R(r)\,Y(\vartheta,\varphi) \quad ,$$

and dividing by $u(r,\vartheta,\varphi)$ we have

$$-\frac{\hbar^2}{2m}\frac{1}{R}\left(\frac{\partial^2 R}{\partial r^2}+\frac{2}{r}\frac{\partial R}{\partial r}\right)+V(r)+\frac{1}{Y}\frac{\hbar^2 L^2}{2mr^2}Y = E_{\rm rel} \quad .$$

$$\underbrace{\frac{r^2}{R}\left(\frac{\partial^2 R}{\partial r^2}+\frac{2}{r}\frac{\partial R}{\partial r}\right)+\frac{2mr^2}{\hbar^2}\Big(E_{\rm rel}-V(r)\Big)}_{F(r)=\lambda,\ \text{function of } r \text{ only}} = \underbrace{\frac{1}{Y}L^2 Y}_{\substack{G(\vartheta,\varphi)=\lambda \\ \text{function of} \\ \vartheta,\varphi \text{ only}}}$$

$$\Rightarrow \begin{cases} \dfrac{\partial^2 R}{\partial r^2}+\dfrac{2}{r}\dfrac{\partial R}{\partial r}+\dfrac{2m}{\hbar^2}\Big(E_{\rm rel}-V(r)\Big)R-\dfrac{\lambda}{r^2}R=0 \\ \text{radial diff. eq., physics is in here} \\ \\ L^2\,Y=\lambda\,Y \\ \text{angular diff. eq.} \qquad \text{“geometry”} \end{cases} \quad .$$

We have already found the solutions to the angular differential equation

$$L^2\,Y=-\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\,\frac{\partial Y}{\partial\vartheta}\right)-\frac{1}{\sin^2\vartheta}\frac{\partial^2 Y}{\partial\varphi^2}=\lambda\,Y, \qquad \begin{aligned}\lambda &= l\,(l+1) \\ l &= 0,1,2,\cdots\end{aligned} \quad .$$

11.2 Radial differential equation

The radial differential equation to be solved is

$$\frac{d^2R}{dr^2}+\frac{2}{r}\frac{dR}{dr}+\frac{2m}{\hbar^2}\Big(E_{\rm rel}-V(r)\Big)R-\frac{l\,(l+1)}{r^2}R=0 \quad .$$

Introducing the reduced wave function $w=r\,R$,
we get

$$\begin{aligned} \frac{d^2}{dr^2}w &= \frac{d^2}{dr^2}rR=\frac{d}{dr}\Big(R+rR'\Big)=R'+R'+rR'' \\ \Rightarrow \frac{1}{r}w'' &= R''+\frac{2}{r}R' \end{aligned}$$

$$\frac{1}{r}w''+\frac{2m}{\hbar^2}\Big(E_{\rm rel}-V(r)\Big)\frac{w}{r}-\frac{l\,(l+1)}{r^2}\frac{w}{r}=0$$

$$\Rightarrow -\frac{\hbar^2}{2m}\frac{d^2w}{dr^2}+\left(V(r)+\frac{\hbar^2\,l\,(l+1)}{2mr^2}\right)w=E_{\rm rel}\,w \quad .$$

This looks like one-dimensional quantum mechanical motion in an effective potential in the interval $r\in[0,\infty)$

$$V_{\rm eff}(r)=V(r)+\underbrace{\frac{\hbar^2}{2m}\frac{l\,(l+1)}{r^2}}_{\substack{\text{centrifugal} \\ \text{potential}}} \quad .$$

The boundary conditions at infinity determine whether we have a bound or a scattering state:

$$\begin{aligned} w &\to 0 \qquad \text{for } r\to\infty \quad \text{bound states} \\ w &\to \mathrm{e}^{\pm ikr} \quad \text{for } r\to\infty \quad \text{free states} \quad . \end{aligned}$$

12 Special functions

12.1 The confluent hypergeometric function

We need to study the differential equation

$$x\,y''(x) + (\gamma - x)\,y'(x) - \alpha\,y(x) = 0$$

with

$$\begin{aligned} \alpha, \gamma &: \text{complex parameters} \\ x &: \text{complex variable.} \end{aligned}$$

Let us try the ansatz for a solution:

$$\begin{aligned} y(x) &= x^\lambda \sum_{n=0}^{\infty} a_n\, x^n = \sum_{n=0}^{\infty} x^{\lambda+n} a_n \\ y'(x) &= \sum_{n=0}^{\infty} (\lambda+n)\, x^{\lambda+n-1} a_n \\ y''(x) &= \sum_{n=0}^{\infty} (\lambda+n)(\lambda+n-1)\, x^{\lambda+n-2} a_n \end{aligned}$$

$$\sum_{n=0}^{\infty} \left[(\lambda+n)(\lambda+n-1)\, a_n\, x^{\lambda+n-1} + \gamma(\lambda+n)\, a_n\, x^{\lambda+n-1} - (\lambda+n)\, a_n\, x^{\lambda+n} - \alpha\, a_n\, x^{\lambda+n} \right] = 0$$

coefficient with $x^{\lambda-1}, n = 0$:

$$\begin{aligned} \lambda(\lambda-1)\, a_0 + \gamma\lambda\, a_0 &= 0 \\ \lambda(\lambda-1) + \gamma\lambda &= 0 \qquad \text{fundamental equation} \end{aligned}$$

$$\Rightarrow \qquad \lambda = 0 \quad \text{or} \quad \lambda = 1 - \gamma$$

two linearly independent solutions! coefficient with $x^{\lambda+n-1}$, $n = 1, 2, \ldots$:

$$\begin{aligned} (\lambda+n)(\lambda+n-1)\, a_n + \gamma(\lambda+n)\, a_n - (\lambda+n-1)\, a_{n-1} - \alpha\, a_{n-1} &= 0 \\ (\lambda+n)(\lambda+n-1+\gamma)\, a_n - [(\lambda+n-1)+\alpha]\, a_{n-1} &= 0 \end{aligned}$$

$$\Rightarrow \frac{a_n}{a_{n-1}} = \frac{\lambda+n-1+\alpha}{(\lambda+n)(\lambda+n-1+\gamma)}$$

(a) $\lambda = 0$

$$\frac{a_n}{a_{n-1}} = \frac{n-1+\alpha}{n\,(n-1+\gamma)}$$

choose

$$\begin{aligned} a_0 &= 1 \\ a_1 &= \frac{\alpha}{\gamma} \\ a_2 &= \frac{\alpha+1}{2\,(\gamma+1)}\, a_1 = \frac{(\alpha+1)\,\alpha}{2\,(\gamma+1)\,\gamma} \\ a_3 &= \frac{\alpha+2}{3\,(\gamma+2)}\, a_2 = \frac{(\alpha+2)(\alpha+1)\,\alpha}{3\cdot 2\,(\gamma+2)(\gamma+1)\,\gamma} \\ &\vdots \end{aligned}$$

$$\begin{aligned} \Rightarrow y(x) &= \sum_{n=0}^{\infty} \frac{\alpha(\alpha+1)(\alpha+2)\ldots(\alpha+n-1)}{n!\,\gamma(\gamma+1)\ldots(\gamma+n-1)}\, x^n = \sum_{n=0}^{\infty} \frac{\alpha_n\, x^n}{n!\,\gamma_n} \\ a_n &\overset{\text{def}}{=} \alpha(\alpha+1)\ldots(\alpha+n-1) \quad \text{Pochammer symbol} \\ &= {}_1F_1(\alpha;\gamma;x) \qquad \text{well defined if } \gamma \neq 0,-1,-2,\ldots \\ &\quad \text{confluent hypergeometric function} \end{aligned}$$

properties:

(i) convergent for every x, since

$$\lim_{n\to\infty} \frac{a_n}{a_{n-1}} = \lim_{n\to\infty} \frac{\lambda+n-1+\alpha}{(\lambda+n)(\lambda+n-1+\gamma)} = 0$$

"quotient criterion"

(ii) power series terminates if $\alpha = -n$ ($n = 0, 1, 2, 3, \ldots$) polynomial of degree n
$\Rightarrow$ Laguerre polynomials

(iii) $\alpha = \gamma \Rightarrow {}_1F_1(\alpha;\alpha;x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \mathrm{e}^x$

(iv) $\frac{d}{dx}\, {}_1F_1(\alpha;\gamma;x) = \frac{\alpha}{\gamma}\, {}_1F_1(\alpha+1;\gamma+1;x)$

(v) $\gamma\, {}_1F_1(\alpha;\gamma;x) + (\alpha-\gamma)\, {}_1F_1(\alpha;\gamma+1;x) - \alpha\, {}_1F_1(\alpha+1;\gamma+1;x) = 0$
$(2\alpha-\gamma+x)\, {}_1F_1(\alpha;\gamma;x) + (\gamma-\alpha)\, {}_1F_1(\alpha-1;\gamma;x) - \alpha\, {}_1F_1(\alpha+1;\gamma;x) = 0$
there are many other relations involving three ${}_1F_1$ functions

(vi) ${}_1F_1(\alpha;\gamma;x) = \mathrm{e}^x\, {}_1F_1(\gamma-\alpha;\gamma;-x)$
Kummer transformation

(b) $\lambda = 1-\gamma$

$$\frac{a_n}{a_{n-1}} = \frac{\lambda+n-1+\alpha}{(\lambda+n)(\lambda+n-1+\gamma)} \quad \text{with } \lambda = 1-\gamma$$

$$\Rightarrow \frac{a_n}{a_{n-1}} = \frac{1-\gamma+n-1+\alpha}{(1-\gamma+n)(1-\gamma+n-1+\gamma)} = \frac{\alpha-\gamma+n}{(1-\gamma+n)\,n}$$

choose

$$a_0 = 1$$

$$a_1 = \frac{\alpha-\gamma+1}{2-\gamma}$$

$$a_2 = \frac{\alpha-\gamma+2}{(3-\gamma)\,2}\,a_1 = \frac{(\alpha-\gamma+2)(\alpha-\gamma+1)}{(3-\gamma)(2-\gamma)\,2}$$

$$a_3 = \frac{(\alpha-\gamma+3)(\alpha-\gamma+2)(\alpha-\gamma+1)}{(4-\gamma)(3-\gamma)(2-\gamma)\,3\cdot 2}$$

$$\vdots \qquad \vdots$$

$$\Rightarrow y(x) = x^{1-\gamma}\,{}_1F_1(\alpha-\gamma+1;2-\gamma;x)$$

conclusion:

the general solution of the confluent hypergeometric differential equation

$$x\,y''(x) + (\gamma - x)\,y'(x) - \alpha\,y(x) = 0$$

is

$$y(x) = A\,{}_1F_1(\alpha;\gamma;x) + B\,\underbrace{x^{1-\gamma}}_{\text{singular at } x=0}\,{}_1F_1(\alpha-\gamma+1;2-\gamma;x)$$

remarks:

(i) if $\gamma \neq$ integer both functions exist and are linearly independent

(ii) assume $\gamma =$ integer

(a) if $\gamma = 1$	both functions are identical
(b) if $\gamma = 0,-1,-2,\ldots$	only $x^{1-\gamma}\,{}_1F_1(\alpha-\gamma+1;2-\gamma;x)$ exists
(c) if $\gamma = 2,3,\ldots$	only ${}_1F_1(\alpha;\gamma;x)$ exists

Thus for $\gamma = 0,\pm1,\pm2,\ldots$ we need an additional linearly independent solution. This problem is difficult but it can be solved. We will not discuss it further here.

12.2 The gamma function

There is a connection between the confluent hypergeometric and the gamma function.
definition:

$$\Gamma(z) \equiv \int_0^\infty \mathrm{e}^{-t}\,t^{z-1}\,dt \qquad \text{convergent if } \mathrm{Re}\,z > 0$$

$$\Gamma(z+1) = \int_0^\infty \mathrm{e}^{-t}\,t^{z}\,dt$$

$$= -\mathrm{e}^{-t}\,t^z\Big|_0^\infty + z\,\underbrace{\int_0^\infty \mathrm{e}^{-t}\,t^{z-1}\,dt}_{\Gamma(z)}$$

$$\Gamma(z+1) = z\Gamma(z) \qquad \text{functional equation}$$

special cases:

(i) $\Gamma(1) = 1$ $\qquad \int_0^\infty \mathrm{e}^{-t}\,dt = 1$

(ii) $\Gamma(n+1) = n!$

(iii) $\Gamma(1/2) = \int_0^\infty \mathrm{e}^{-t}t^{-1/2}dt = 2\int_0^\infty \mathrm{e}^{-u^2}du = \int_{-\infty}^\infty \mathrm{e}^{-u^2}du = \sqrt{\pi}$

substitution $u = \sqrt{t}$, $\quad du = \dfrac{dt}{2\sqrt{t}}$

(iv)

$$\Gamma(n+1/2) = (n-1/2)(n-3/2)\ldots 1/2\,\Gamma(1/2)$$

$$= \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n}\sqrt{\pi}$$

$$= \frac{(2n-1)!!}{2^n}\sqrt{\pi}$$

$\Gamma(z)$ defined in terms of integral representation only for $\mathrm{Re}\,z > 0$. However, we can define it everywhere using analytic continuation if we assume functional equation valid everywhere except $z = 0, -1, -2,\ldots$

$$\Gamma(z+1) = z\,\Gamma(z)$$

indeed for $|z| < \varepsilon$ assume Taylor expansion exists around $z = 1$

$$\Gamma(z+1) = \underbrace{\Gamma(1)}_{1} + z\,\underbrace{\left(\frac{\Gamma'(1)}{1!} + z\,\frac{\Gamma''(1)}{2!} + z^2\frac{\Gamma'''(1)}{3!} + \ldots\right)}_{f(z)} = 1 + z\,f(z)$$

$f(z)$ holomorphic around $z = 0$, defined ∞-differentiable functional equation:

$$\Gamma(z) = \frac{1}{z}\,\Gamma(z+1) = \frac{1}{z} + f(z)$$

$\Rightarrow z = 0$ simple pole of $\Gamma(z)$ with residue 1.
more generally

$$\Gamma(z) = (z-1)\,\Gamma(z-1) = (z-1)\,(z-2)\,\Gamma(z-2) = (z-1)\ldots(z-n)\,\Gamma(z-n)$$

$$\begin{aligned}\Rightarrow \Gamma(z-n) &= \frac{\Gamma(z)}{(z-1)\ldots(z-n)} = \frac{\frac{1}{z} + f(z)}{(z-1)\ldots(z-n)} \\ &= \frac{1}{z\,(z-1)\ldots(z-n)} + \underbrace{\frac{f(z)}{(z-1)\ldots(z-n)}}_{\text{holomorphic for } |z|<\varepsilon}\end{aligned}$$

for $|z| < \varepsilon$

$$\begin{aligned}\Gamma(z-n) &= \frac{(-1)^n}{n!\,z} + g_n(z) \\ & \quad g_n(z) \text{ holomorphic for } |z| < \varepsilon\end{aligned}$$

$\Rightarrow \Gamma(z)$ has simple poles at $z = 0, -1, -2, \ldots$ with residues $r_n = \dfrac{(-1)^n}{n!}$

further properties without proof:

(i) $\Gamma(z)\,\Gamma(1-z) = \dfrac{\pi}{\sin \pi z} \Rightarrow \Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi}$ as we already know

(ii) $\Gamma(2z) = \dfrac{2^{2z-1}}{\sqrt{\pi}}\,\Gamma(z)\,\Gamma(z+1/2)$

12.3 Euler's beta function

definition:

$$B(x,y) \equiv \int_0^1 t^{x-1}(1-t)^{y-1}\,dt \quad \text{convergent if } \begin{matrix}\text{Re } x > 0 \\ \text{Re } y > 0\end{matrix}$$

properties:

(i) $B(x,y) = \dfrac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}$

(ii) $B(x,y+1) = \dfrac{y}{x}\,B(x+1,y)$

proof of (ii): $\dfrac{\Gamma(x)\,\Gamma(y+1)}{\Gamma(x+y+1)} = \dfrac{y}{x}\,\dfrac{\Gamma(x+1)\,\Gamma(y)}{\Gamma(x+y+1)}$ proof of (i): substitute

$$t = \frac{v}{v+1} \qquad v = \frac{t}{1-t} \qquad dt = \frac{dv}{(1+v)^2}$$

$$t \in [0,1] \quad \Rightarrow \quad v \in [0,\infty)$$

$$B(x,y) = \int_0^\infty v^{x-1}(1+v)^{-x-y}\,dv$$

let us evaluate the integral

$$\begin{aligned}I(x,y) &= \int_0^\infty v^{x-1}\,dv \underbrace{\int_0^\infty \mathrm{e}^{-(1+v)t}\,t^{x+y-1}\,dt}_{\frac{\Gamma(x+y)}{(1+v)^{x+y}}} \\ &= \int_0^\infty \frac{\Gamma(x+y)}{(1+v)^{x+y}}\,v^{x-1}\,dv \\ &= \Gamma(x+y)\,B(x,y)\end{aligned}$$

if we can show that $I(x,y) = \Gamma(x)\,\Gamma(y)$, proof complete

$$\begin{aligned}I(x,y) &= \int_0^\infty\!\!\int_0^\infty v^{x-1}\,\mathrm{e}^{-(1+v)t}\,t^{x+y-1}\,dv\,dt \\ &\quad \text{substitution } (v,t) \to (s,t):\ vt = s,\ t\,dv = ds \\ I(x,y) &= \int_0^\infty\!\!\int_0^\infty \left(\frac{s}{t}\right)^{x-1}\mathrm{e}^{-t-s}\,t^{x+y-1}\,\frac{ds}{t}\,dt \\ &= \int_0^\infty\!\!\int_0^\infty s^{x-1}\,\mathrm{e}^{-s}\,\frac{ds}{t}\,\mathrm{e}^{-t}\,t^{y}\,dt \\ &= \Gamma(x)\,\Gamma(y) \qquad \text{q.e.d.}\end{aligned}$$

12.4 Integral representation of the confluent hypergeometric function

consider the integral

$$\begin{aligned}F(\alpha;\gamma;z) &\equiv \frac{\Gamma(\gamma)}{\Gamma(\alpha)\,\Gamma(\gamma-\alpha)}\int_0^1 \overbrace{\mathrm{e}^{zt}}^{\sum_{n=0}^\infty \frac{z^n t^n}{n!}}\,t^{\alpha-1}(1-t)^{\gamma-\alpha-1}dt \\ &\quad \text{Re } \alpha > 0 \\ &\quad \text{Re } (\gamma-\alpha) > 0 \\ &= \frac{\Gamma(\gamma)}{\Gamma(\alpha)\,\Gamma(\gamma-\alpha)}\sum_{n=0}^\infty \frac{z^n}{n!}\int_0^1 t^{\alpha+n-1}\,(1-t)^{\gamma-\alpha-1}dt\end{aligned}$$

$$\begin{aligned}\text{but} \quad B(\alpha+n,\gamma-\alpha) &= \int_0^1 t^{\alpha+n-1}\,(1-t)^{\gamma-\alpha-1}dt \\ &= \frac{\Gamma(\alpha+n)\,\Gamma(\gamma-\alpha)}{\Gamma(\gamma+n)}\end{aligned}$$

$$\Rightarrow F(\alpha;\gamma;z) = \sum_{n=0}^{\infty} \frac{z^n\,\Gamma(\gamma)\,\Gamma(\alpha+n)\,\Gamma(\gamma-\alpha)}{n!\,\Gamma(\alpha)\,\Gamma(\gamma-\alpha)\,\Gamma(\gamma+n)}$$

$$\text{but}\quad \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} = (\alpha+n-1)\ldots(\alpha+1)\,\alpha\frac{\Gamma(\alpha)}{\Gamma(\alpha)} \equiv (\alpha)_n$$

$$\Rightarrow F(\alpha;\gamma;z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}\,\frac{(\alpha)_n}{(\gamma)_n} \equiv {}_1F_1(\alpha;\gamma;z)$$

asymptotic representation:

$${}_1F_1(\alpha;\gamma;z) = F(\alpha;\gamma;z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\,\Gamma(\gamma-\alpha)}\left[\left(\int_{-\infty}^{1} - \int_{-\infty}^{0}\right) \mathrm{e}^{zt}\,t^{\alpha-1}\,(1-t)^{\gamma-\alpha-1}\,dt\right]$$

two substitutions:

(i) $t = 1 - \dfrac{u}{z}$, $\qquad t \in (-\infty,1] \Rightarrow u \in (\infty,0]$, $\quad dt = -\dfrac{du}{z}$

(ii) $t = -\dfrac{u}{z}$, $\qquad t \in (-\infty,0] \Rightarrow u \in (\infty,0]$, $\quad dt = -\dfrac{du}{z}$

$$\begin{aligned} {}_1F_1(\alpha;\gamma;z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\,\Gamma(\gamma-\alpha)}\Bigg[& \mathrm{e}^z\,z^{\alpha-\gamma}\overbrace{\int_0^{\infty} du\,\mathrm{e}^{-u}\,u^{\gamma-\alpha-1}\left(1-\frac{u}{z}\right)^{\alpha-1}}^{\sim\,\Gamma(\gamma-\alpha)\text{ for }|z|\gg 1} \\ & + (-z)^{-\alpha}\underbrace{\int_0^{\infty} du\,\mathrm{e}^{-u}\,u^{\alpha-1}\left(1+\frac{u}{z}\right)^{\gamma-\alpha-1}}_{\sim\,\Gamma(\alpha)\text{ for }|z|\gg 1}\Bigg] \end{aligned}$$

for $|z| \gg 1$ first term of asymptotic expansion

$${}_1F_1(\alpha;\gamma;z) \sim \frac{\Gamma(\gamma)}{\Gamma(\alpha)}\,\mathrm{e}^z\,z^{\alpha-\gamma} + \frac{\Gamma(\gamma)}{\Gamma(\gamma-\alpha)}\,(-z)^{-\alpha} + \ldots$$

check of the approximation:

(i) ${}_1F_1(\alpha;\alpha;z) = \mathrm{e}^z$, approximation gives the same.

(ii) ${}_1F_1(-2;\gamma;z) = 1 + \dfrac{-2}{\gamma}\,z + \dfrac{(-2)(-2+1)}{\gamma\,(\gamma+1)\,2!}\,z^2 \sim \dfrac{z^2}{\gamma\,(\gamma+1)}$ approximation yields only the leading term

13 The Coulomb problem

13.1 Formulation of the problem

Hydrogen-like atom: nucleus with charge Ze, one electron with charge $-e$, highly ionized

$$V(r) = -\frac{Ze^2}{r} \qquad \text{Coulomb potential}$$

we discuss here only the attractive case: free and bound states
radial differential equation:

$$\frac{d^2w}{dr^2} + \frac{2m}{\hbar^2}\left(E + \frac{Ze^2}{r}\right)w - \frac{l\,(l+1)}{r^2}\,w = 0$$

w : reduced wave function $\quad R = \dfrac{w}{r}$

$$w'' + \left(k^2 + \frac{\lambda}{r}\right)w - \frac{l\,(l+1)}{r^2}\,w = 0$$

where

$$\begin{cases} \lambda &= \dfrac{2mZe^2}{\hbar^2} \\ k^2 &= \dfrac{2mE}{\hbar^2} \end{cases}$$

two possibilities:

(a) $E < 0 \Rightarrow k^2 = -\kappa^2 < 0$: bound states

(b) $E > 0 \Rightarrow k^2 > 0$: free (scattering) states

asymptotic behaviour:

(i) $r \to 0 \quad \Rightarrow \quad w'' - \dfrac{l\,(l+1)}{r^2}\,w = 0 \quad \Rightarrow \quad w \sim r^{l+1}, r^{-l}$

(ii) $r \to \infty$

(a) bound states: $E < 0$
$w'' - \kappa^2\,w = 0, \quad w \sim \mathrm{e}^{-\kappa r}, \mathrm{e}^{+\kappa r}$

(b) free (scattering) states: $E > 0$
$w'' + k^2\,w = 0 \quad w \sim \mathrm{e}^{ikr}, \mathrm{e}^{-ikr}$
we will treat this case later

13.2 Bound states ($E < 0$)

reject r^{-l} since for $r \to 0$ singular: not normalizable (true for $l \geq 1$, also valid for $l = 0$, however one needs more careful argumentation) reject $\mathrm{e}^{\kappa r}$ since for $r \to \infty$ singular: not normalizable.
ansatz:

$$w(r) = \mathrm{e}^{-\kappa r}\,r^{l+1}\,\underbrace{v(r)}_{\text{well behaved function}}$$

$$\Rightarrow \frac{d^2v}{dr^2} + \frac{dv}{dr}\left(\frac{2(l+1)}{r} - 2\kappa\right) + v\,\frac{\lambda - 2(l+1)\kappa}{r} = 0$$

substitute $x = 2\kappa r$

$$\Rightarrow x\,\frac{d^2v}{dx^2} + \frac{dv}{dx}\left[2(l+1) - x\right] - v\left[l + 1 - \frac{\lambda}{2\kappa}\right] = 0$$

compare with

$$x\,v''(x) + (\gamma - x)\,v'(x) - \alpha\,v(x) = 0$$

$$v(x) = {}_1F_1(\alpha;\gamma;x) \qquad \text{regular for } x \to 0$$

exclude the other solution

$$\Rightarrow \begin{cases} \gamma &= 2l+2 \\ \alpha &= l+1-\dfrac{\lambda}{2\kappa} \end{cases}$$

regular solution:

$$R(r) = \text{const} \times r^l\,\mathrm{e}^{-\kappa r}\,{}_1F_1\left(l+1-\frac{\lambda}{2\kappa};2l+2;2\kappa r\right)$$

asymptotic behaviour:

$${}_1F_1(\alpha;\gamma;z) \sim \frac{\Gamma(\gamma)}{\Gamma(\alpha)}\,\mathrm{e}^z\,z^{\alpha-\gamma} + \frac{\Gamma(\gamma)}{\Gamma(\gamma-\alpha)}\,(-z)^{-\alpha}$$

$$\begin{cases} \gamma &= 2l+2 \\ \alpha &= l+1-\dfrac{\lambda}{2\kappa} \end{cases} \qquad z = 2\kappa r$$

$${}_1F_1\left(l+1-\frac{\lambda}{2\kappa};2l+2;2\kappa r\right) \sim \frac{\Gamma(2l+2)}{\Gamma\left(l+1-\frac{\lambda}{2\kappa}\right)}\,\mathrm{e}^{2\kappa r}\,(2\kappa r)^{-l-1-\frac{\lambda}{2\kappa}} + \frac{\Gamma(2l+2)}{\Gamma\left(l+1+\frac{\lambda}{2\kappa}\right)}\,(-2\kappa r)^{-l-1+\frac{\lambda}{2\kappa}}$$

leading term goes like $\mathrm{e}^{2\kappa r}$ for $r \to \infty$ which grows faster than $\mathrm{e}^{-\kappa r}$ decreases $\Rightarrow$ contradiction with $R(r) \to 0$ for $r \to \infty$ (boundary condition). The only way to get rid of this leading term is to have

$$\Gamma\left(l+1-\frac{\lambda}{2\kappa}\right) = \infty \qquad \Gamma\text{-function has poles at}$$

$$l+1-\frac{\lambda}{2\kappa} = -n_r \qquad n_r = 0,1,2,\ldots \quad \text{radial quantum-number}$$

the second term will grow like a power

$$n \equiv l+1+n_r \qquad \text{principal quantum number}$$

condition for eigenvalues:

$$\frac{\lambda}{2\kappa} = n \quad \Rightarrow \quad \kappa^2 = \frac{\lambda^2}{4n^2}$$

$$\begin{cases} \lambda &= \dfrac{2m}{\hbar^2}\,Ze^2 \\ \kappa^2 &= -\dfrac{2mE}{\hbar^2} \end{cases}$$

$$E = -\frac{\hbar^2\kappa^2}{2m} = -\frac{\hbar^2}{2m}\,\frac{\lambda^2}{4n^2}$$

$$= -\frac{\hbar^2}{2m}\left(\frac{2m}{\hbar^2}\right)^2\frac{Z^2e^4}{4n^2}$$

$$E = -\frac{2m}{\hbar^2}\,\frac{Z^2e^4}{4n^2} = -\frac{1}{2}\,mc^2\left(\frac{e^2}{\hbar c}\right)^2\frac{Z^2}{n^2}$$

Balmer formula: J.J. Balmer (Swiss, 1885)

$$E = -\frac{1}{2}\,mc^2\,\alpha^2\,\frac{Z^2}{n^2} = -13.605804\ \text{eV}\,\frac{Z^2}{n^2}$$

$$\begin{cases} n &= 1,2,3,\ldots \\ l &= 0,1,2,\ldots,n-1 \\ (\text{since } l = n-1-n_r) & \\ m &= -l,-l+1,\ldots,l \end{cases}$$

where

$\alpha = \dfrac{1}{137.0359895(61)}$ fine structure constant

$m = \dfrac{m_e\,m_A}{m_e+m_A}$ reduced mass

$m_e = 0.51099906(15)\ \text{MeV}/c^2$ mass of the electron

m_A mass of the nucleus

<u>bound state wavefunction</u>

$$u(r,\vartheta,\varphi) = \text{const} \times r^l \exp(-r/an)\,{}_1F_1\left(-n+l+1;2l+2;\frac{2r}{a}\frac{1}{n}\right)Y_{lm}(\vartheta,\varphi)$$

$$\kappa = \frac{\lambda}{2n}$$

$$a = \frac{\hbar}{mc}\,\frac{1}{\alpha Z} = \frac{\hbar}{mc}\,\frac{\hbar c}{e^2 Z} = \frac{\hbar^2}{me^2Z} \qquad \text{Bohr's radius}$$

definition of Laguerre polynomials: polynomials of degree k

$$L_k(x) \equiv e^x \frac{d^k}{dx^k}\left(x^k e^{-x}\right) \qquad \text{eq.(1)}$$

$$\left(fg\right)^{(k)} = \sum_{n=0}^{k} \binom{k}{n} f^{(k-n)} g^{(n)} \qquad \text{Leibniz' rule}$$

$$\Rightarrow L_k(x) = \sum_{n=0}^{k} \binom{k}{n} (-1)^n \underbrace{k(k-1)\ldots(n+1)}_{k-n \text{ terms}} x^n \qquad \binom{k}{n} = \frac{k!}{n!(k-n)!}$$

$$= \sum_{n=0}^{k} \frac{(-1)^n k!\, k!}{n!\,(k-n)!\,n!} x^n$$

examples:

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= 1 - x \\ L_2(x) &= 2 - 4x + x^2 \\ L_3(x) &= 6 - 18x + 9x^2 - x^3 \end{aligned}$$

$${}_1F_1(-k;1;x) = \sum_{n=0}^{k} \frac{(-k)_n}{n!\,(1)_n} x^n = \sum_{n=0}^{k} \frac{(-k)(-k+1)\ldots(-k+n-1)}{n!\,n!} x^n$$

$$= \sum_{n=0}^{k} \frac{(-1)^n k(k-1)\ldots(k-n+1)}{n!\,n!} x^n = \sum_{n=0}^{k} (-1)^n \frac{k!}{n!\,n!\,(k-n)!} x^n$$

comparison

$$\Rightarrow {}_1F_1(-k;1;x) = \frac{1}{k!} L_k(x)$$

definition of associated Laguerre polynomials: polynomial of degree k

$$L^m_{k+m}(x) \equiv \frac{d^m}{dx^m} L_{k+m}(x) \qquad \text{eq.(2)}$$

$$= (k+m)! \frac{d^m}{dx^m} {}_1F_1(-(k+m);1;x)$$

$$= (k+m)! \frac{(-k-m)_m}{(1)_m} {}_1F_1(-k;m+1;x)$$

$$= (k+m)! \frac{(-1)^m (k+m)(k+m-1)\ldots(k+1)\,k!}{m!\,k!} {}_1F_1(-k;m+1;x)$$

$$= (-1)^m \binom{k+m}{m} (k+m)!\, {}_1F_1(-k;m+1;x) \quad \text{eq.(1) inserted in eq.(2)}$$

$$L^m_{k+m}(x) = \frac{d^m}{dx^m} L_{k+m}(x) = \frac{d^m}{dx^m}\left[e^x \frac{d^{k+m}}{dx^{k+m}}\left(x^{k+m} e^{-x}\right)\right]$$

special cases:

(1)

$$\begin{aligned} L^0_0(x) &= 1 \\ L^0_1(x) &= -x + 1 \\ L^1_1(x) &= -1 \\ L^0_2(x) &= x^2 - 4x + 2 \\ L^1_2(x) &= 2x - 4 \\ L^2_2(x) &= 2 \\ L^0_3(x) &= -x^3 + 9x^2 - 18x + 6 \\ L^1_3(x) &= -3x^2 + 18x - 18 \\ L^2_3(x) &= -6x + 18 \\ L^3_3(x) &= -6 \end{aligned}$$

normalized wavefunction:

$$u_{nlm}(r,\vartheta,\varphi) = -\sqrt{\frac{(n-l-1)!}{2n\,[(n+l)!]^3}\left(\frac{2Z}{an}\right)^{2l+3}} r^l e^{-\frac{Zr}{an}} L^{2l+1}_{n+l}\left(\frac{2Zr}{an}\right) Y_{lm}(\vartheta,\varphi)$$

with

$$\begin{aligned} n &= 1,2,3,\ldots \\ l &= 0,1,2,\ldots,n-1 \\ m &= -l,-l+1,\ldots,l \end{aligned}$$

orthonormal

$$\int u^*_{nlm}(\vec{r})\, u_{n'l'm'}(\vec{r})\, d^3r = \delta_{nn'}\delta_{ll'}\delta_{mm'}$$

but not complete! we also need free states. eigenvalues:

$$E_{nlm} = -\frac{1}{2} mc^2 \left(\frac{e^2}{\hbar c}\right)^2 \frac{Z^2}{n^2}$$

groundstate:

$$u_{100}(r,\vartheta,\varphi) = \frac{1}{\sqrt{\pi\, a^3}} e^{-\frac{r}{a}} \qquad \text{for } Z = 1$$

13.3 Free states ($E > 0$)

eq. for reduced wavefunction

$$w'' + \left(k^2 + \frac{\lambda}{r}\right) w - \frac{l(l+1)}{r^2} w = 0$$

$$\begin{cases} \lambda = \dfrac{2mZe^2}{\hbar^2} \\ k^2 = \dfrac{2mE}{\hbar^2} > 0 \qquad \text{free states} \end{cases}$$

ansatz:

$$w(r) = r^{l+1}\,\mathrm{e}^{-ikr}\,v(r)$$

substitution: $x = 2ikr$

$$\Rightarrow x\,\frac{d^2v}{dx^2} + (2l+2-x)\,\frac{dv}{dx} - \left(l+1+i\,\frac{\lambda}{2k}\right)\,v = 0$$

confluent hypergeometric differential equation

$$x\,v'' + (\gamma - x)\,v' - \alpha\,v = 0$$

$$v(r) = {}_1F_1\left(l+1+i\,\frac{\lambda}{2k}; 2l+2; 2ikr\right)$$

$$\begin{cases} \lambda &= \dfrac{2mZe^2}{\hbar^2} \\ k^2 &= \dfrac{2mE}{\hbar^2} = \dfrac{2m\frac{1}{2}mv_\infty^2}{\hbar^2} \end{cases}$$

$$\begin{cases} \alpha &= l+1+i\,\dfrac{\lambda}{2k} \\ \gamma &= 2l+2 \end{cases}$$

$$\Rightarrow k = \frac{mv_\infty}{\hbar}$$

$$\frac{\lambda}{2k} = \frac{Ze^2}{\hbar v_\infty} \equiv \eta \quad \text{Coulomb parameter}$$

$$R(r) = \text{const}\; r^l\,\mathrm{e}^{-ikr}\,{}_1F_1(l+1+i\eta; 2l+2; 2ikr)$$

theorem: $R(r) = R^*(r)$ provided $\text{const}^* = \text{const}$

proof: ${}_1F_1(\alpha;\gamma;x) = \mathrm{e}^x\,{}_1F_1(\gamma-\alpha;\gamma;-x)$ Kummer transformation

$$\begin{aligned} \text{compl. conj. } R^*(r) &= \text{const}^*\, r^l\,\mathrm{e}^{ikr}\,{}_1F_1(l+1-i\eta; 2l+2; -2ikr) \\ \text{use Kummer: } R(r) &= \text{const}\, r^l\,\mathrm{e}^{-ikr}\,\mathrm{e}^{2ikr}\,{}_1F_1(2l+2-l-1-i\eta; 2l+2; -2ikr) \\ &= \text{const}\, r^l\,\mathrm{e}^{ikr}\,{}_1F_1(l+1-i\eta; 2l+2; -2ikr) \end{aligned}$$

$\Rightarrow R(r)$ real if const real

asymptotic behaviour:

$${}_1F_1(\alpha;\gamma;z) \sim \frac{\Gamma(\gamma)}{\Gamma(\alpha)}\,\mathrm{e}^z\,z^{\alpha-\gamma} + \frac{\Gamma(\gamma)}{\Gamma(\gamma-\alpha)}\,(-z)^{-\alpha}$$

$$\begin{aligned} \Rightarrow R(r) &\sim \text{const}\left[r^l\,\mathrm{e}^{-ikr}\,\frac{\Gamma(2l+2)}{\Gamma(l+1+i\eta)}\,\mathrm{e}^{2ikr}\,(2ikr)^{-l-1+i\eta}\right. \\ &\qquad \left. + r^l\,\mathrm{e}^{-ikr}\,\frac{\Gamma(2l+2)}{\Gamma(l+1-i\eta)}\,(-2ikr)^{-l-1-i\eta}\right] \\ &\sim \text{const}\left[r^l\,\mathrm{e}^{ikr}\,\frac{\Gamma(2l+2)}{\Gamma(l+1+i\eta)}\,\mathrm{e}^{(\ln 2kr + \frac{i\pi}{2})(-l-1+i\eta)}\right. \\ &\qquad \left. + \text{complex conjugate expression}\right] \\ &\sim \text{const}\left[r^l\,\frac{\Gamma(2l+2)}{\Gamma(l+1+i\eta)}\,\mathrm{e}^{ikr}\,\mathrm{e}^{-(l+1)\ln 2kr - \eta\frac{\pi}{2} + i[\eta\,\ln 2kr - (l+1)\frac{\pi}{2}]}\right. \\ &\qquad \left. + \text{complex conjugate expression}\right] \end{aligned}$$

$$\Gamma(l+1+i\eta) \equiv |\Gamma(l+1+i\eta)|\,\mathrm{e}^{i\sigma_l(\eta)}$$

$\sigma_l(\eta)$: Coulomb phase shift

$$\begin{aligned} R(r) &\sim \text{const}\left[r^l\,\frac{\Gamma(2l+2)}{|\Gamma(l+1+i\eta)|}\,(2kr)^{-(l+1)}\,\mathrm{e}^{-\eta\frac{\pi}{2}}\,\mathrm{e}^{i[kr+\eta\,\ln 2kr - (l+1)\frac{\pi}{2} - \sigma_l(\eta)]}\right. \\ &\qquad \left. + \text{complex conjugate expression}\right] \\ &\sim \text{const}\,\frac{1}{r}\,\frac{\Gamma(2l+2)}{|\Gamma(l+1+i\eta)|}\,(2k)^{-(l+1)}\,\mathrm{e}^{-\eta\frac{\pi}{2}}\,2\,\sin\left(kr + \eta\,\ln 2kr - l\frac{\pi}{2} - \sigma_l(\eta)\right) \end{aligned}$$

define normalization constant

$$\text{const} \equiv \frac{|\Gamma(l+1+i\eta)|}{(2l+1)!}\,(2k)^l\,\mathrm{e}^{\eta\frac{\pi}{2}}$$

$$\begin{aligned} \Rightarrow R(r) &= \frac{|\Gamma(l+1+i\eta)|}{(2l+1)!}\,(2kr)^l\,\mathrm{e}^{\eta\frac{\pi}{2}}\,\mathrm{e}^{-ikr}\,{}_1F_1(l+1+i\eta; 2l+2; 2ikr) \\ R(r) &\sim \frac{1}{kr}\,\sin\left(kr + \eta\,\ln 2kr - l\frac{\pi}{2} - \sigma_l(\eta)\right) \\ j_l(kr) &\sim \frac{1}{kr}\,\sin\left(kr - l\frac{\pi}{2}\right) \end{aligned}$$

introduction of the concept of phase shift

if the potential decreases faster than Coulomb, i.e. $\propto r^{-(1+\varepsilon)}$, with $\varepsilon > 0$, then there is no $\eta \ln 2kr$ term

calculation of Coulomb phase shift:

$$\begin{aligned} \sigma_l(\eta) &\equiv \arg\Gamma(l+1+i\eta) \\ \text{or } \Gamma(l+1+i\eta) &= |\Gamma(l+1+i\eta)|\, e^{i\sigma_l(\eta)} \end{aligned}$$

$$\begin{aligned} |\Gamma(l+1+i\eta)|^2 &= (l+i\eta)\,(l-i\eta)\,|\Gamma(l+i\eta)|^2 \\ &= (l^2+\eta^2)\,\left((l-1)^2+\eta^2\right)\ldots\left(1+\eta^2\right)|\Gamma(1+i\eta)|^2 \\ |\Gamma(1+i\eta)|^2 &= \Gamma(1+i\eta)\,\Gamma(1-i\eta) = i\eta\,\Gamma(i\eta)\,\Gamma(1-i\eta) \end{aligned}$$

$$\Gamma(z)\,\Gamma(1-z) = \frac{\pi}{\sin\pi z}$$

$$|\Gamma(1+i\eta)|^2 = \left|i\eta\,\frac{\pi}{\sinh\pi\eta}\right| = \left(\frac{\pi\eta}{\sinh\pi\eta}\right)$$

$$|\Gamma(l+1+i\eta)|^2 = (l^2+\eta^2)\,\left((l-1)^2+\eta^2\right)\ldots\left(1+\eta^2\right)\left(\frac{\pi\eta}{\sinh\pi\eta}\right)$$

$$\Gamma(l+1+i\eta) = (l+i\eta)(l-1+i\eta)\ldots(1+i\eta)\Gamma(1+i\eta)$$

$$\begin{aligned} \sigma_l(\eta) &= \sum_{\lambda=1}^{l} \arctan\frac{\eta}{\lambda} + \sigma_0(\eta) \\ \sigma_0(\eta) &= \arg\Gamma(1+i\eta) \\ &\quad \text{not elementary} \end{aligned}$$

14 Charged particle in an electromagnetic field: Abelian gauge theory

14.1 Classical theory

- definition: Lagrange function

$$L(\vec{r},\dot{\vec{r}},t) \underset{\text{def}}{=} \frac{1}{2}\,m\,\dot{\vec{r}}^2 - e\,\phi + \frac{e}{c}\,\dot{\vec{r}}\cdot\vec{A}$$

e: charge, m: mass } of the particle

ϕ: scalar, $\vec{A}$: vector } potentials

$$\vec{B} \underset{\text{def}}{=} \vec{\nabla}\times\vec{A} \qquad \text{magnetic}$$

$$\vec{E} \underset{\text{def}}{=} -\vec{\nabla}\phi - \frac{1}{c}\,\frac{\partial\vec{A}}{\partial t} \qquad \text{electric} \quad \Big\} \text{ fields}$$

- theorem: Euler-Lagrange equation derived from $L(\vec{r},\dot{\vec{r}},t)$ yields Newton's equation of motion with the Lorentz force.

proof:

momentum canonically conjugate to $\vec{r}$

$$\vec{p} \underset{\text{def}}{=} \frac{\partial L}{\partial\dot{\vec{r}}} = m\,\dot{\vec{r}} + \frac{e}{c}\,\vec{A} \qquad \begin{array}{l}\text{momentum acquires} \\ \text{an extra term } \dfrac{e}{c}\,\vec{A}\end{array}$$

rate of change of momentum

$$\begin{aligned} \frac{d}{dt}\,\vec{p} &= \frac{d}{dt}\,\frac{\partial L}{\partial\dot{\vec{r}}} = m\,\ddot{\vec{r}} + \frac{e}{c}\,\frac{d}{dt}\,\vec{A} \\ &= m\,\ddot{\vec{r}} + \frac{e}{c}\left(\dot{\vec{r}}\cdot\vec{\nabla}\right)\vec{A} + \frac{e}{c}\,\frac{\partial}{\partial t}\,\vec{A}\ . \end{aligned}$$

Here we have used

$$\frac{d}{dt}\,A_k(x,y,z,t) = \dot{x}\,\frac{\partial A_k}{\partial x} + \dot{y}\,\frac{\partial A_k}{\partial y} + \dot{z}\,\frac{\partial A_k}{\partial z} + \frac{\partial A_k}{\partial t} \qquad (k = x,y,z)$$

or in vector notation

$$\frac{d}{dt}\,\vec{A} = \left(\dot{\vec{r}}\cdot\vec{\nabla}\right)\vec{A} + \frac{\partial\vec{A}}{\partial t}\ .$$

generalized force

$$\begin{aligned}\frac{\partial L}{\partial \vec{r}} &= -e\,\vec{\nabla}\,\phi + \frac{e}{c}\,\vec{\nabla}\,\left(\dot{\vec{r}}\cdot\vec{A}\right) \\ &= -e\,\vec{\nabla}\,\phi + \frac{e}{c}\,\left(\dot{\vec{r}}\cdot\vec{\nabla}\right)\,\vec{A} + \frac{e}{c}\,\dot{\vec{r}}\times\left(\vec{\nabla}\times\vec{A}\right)\ .\end{aligned}$$

Here we have used the vector operator identity

$$\dot{\vec{r}}\times\left(\vec{\nabla}\times\vec{A}\right) = \vec{\nabla}\,\left(\dot{\vec{r}}\cdot\vec{A}\right) - \left(\dot{\vec{r}}\cdot\vec{\nabla}\right)\,\vec{A}$$

which follows from the vector identity

$$\vec{a}\times\left(\vec{b}\times\vec{c}\right) = \left(\vec{a}\cdot\vec{c}\right)\,\vec{b} - \left(\vec{a}\cdot\vec{b}\right)\,\vec{c}\ ,$$

substituting $\vec{a}=\dot{\vec{r}},\ \vec{b}=\vec{\nabla},\ \vec{c}=\vec{A}$ and
taking into account that $\vec{\nabla}$ is acting on $\vec{A}$ only.

Euler-Lagrange equation

$$\frac{d}{dt}\,\frac{\partial L}{\partial \dot{\vec{r}}} = \frac{\partial L}{\partial \vec{r}}$$

$$m\,\ddot{\vec{r}} + \frac{e}{c}\,\left(\dot{\vec{r}}\cdot\vec{\nabla}\right)\,\vec{A} + \frac{e}{c}\,\frac{\partial \vec{A}}{\partial t} = -e\,\vec{\nabla}\,\phi + \frac{e}{c}\,\left(\dot{\vec{r}}\cdot\vec{\nabla}\right)\,\vec{A} + \frac{e}{c}\dot{\vec{r}}\times\left(\vec{\nabla}\times\vec{A}\right)$$

$$\Rightarrow\ m\,\ddot{\vec{r}} = \underbrace{-e\,\vec{\nabla}\,\phi - \frac{e}{c}\,\frac{\partial \vec{A}}{\partial t}}_{e\vec{E}} + \underbrace{\frac{e}{c}\,\dot{\vec{r}}\times\left(\vec{\nabla}\times\vec{A}\right)}_{\frac{e}{c}\,\dot{\vec{r}}\times\vec{B}}\ .$$

Lorentz force

We thus conclude that the chosen Lagrangean indeed describes a particle of charge e and mass m moving in electric and magnetic fields, $\vec{E}$ and $\vec{B}$, q.e.d.

- definition: Hamilton function

$$\begin{aligned}H &\underset{\text{def}}{=} \dot{\vec{r}}\cdot\vec{p} - L \quad \text{with} \quad \vec{p} = m\,\dot{\vec{r}} + \frac{e}{c}\,\vec{A} \\ H &= m\,\dot{\vec{r}}^{\,2} + \frac{e}{c}\,\dot{\vec{r}}\cdot\vec{A} - \frac{1}{2}\,m\,\dot{r}^2 + e\,\phi - \frac{e}{c}\,\dot{\vec{r}}\cdot\vec{A} \\ H &= \frac{1}{2}\,m\,\dot{\vec{r}}^{\,2} + e\,\phi\ .\end{aligned}$$

We recall that H is a function of the canonically conjugate variables, $\vec{r}$ and $\vec{p}$ only. Using $\vec{p} = m\dot{\vec{r}} + \frac{e}{c}\,\vec{A}$ to get rid of $\dot{\vec{r}}$, we arrive at

$$H = \frac{\left(\vec{p} - \frac{e}{c}\,\vec{A}\right)^2}{2m} + e\,\phi\ .$$

- principle of minimal substitution

In order to obtain the Hamilton function, coupling a charged massive particle to an electromagnetic field, one needs a Hamilton function describing the non-interacting particle.

example: We start with $H = \frac{\vec{p}^{\,2}}{2m}$, the nonrelativistic energy-momentum-mass relationship.

Introduce the minimal substitution

$$\begin{cases}\vec{p} \to \vec{p} - \frac{e}{c}\,\vec{A} \\ H \to H - e\,\phi\ .\end{cases}$$

For our example, this yields

$$H = \frac{\left(\vec{p} - \frac{e}{c}\,\vec{A}\right)^2}{2m} + e\,\phi\ .$$

remark:

This recipe also works for a Hamilton function describing a non-interacting relativistic particle or a matter field. In fact, all the fundamental non-gravitational interactions of the Standard Model of Particle Physics are obtained via a generalized minimal substitution principle.

14.2 Quantum mechanical theory

- The quantum mechanical theory is obtained by imposing the generalized Born-Jordan quantization conditions

$$\begin{aligned}[p_k, q_l] &= -i\hbar\,\delta_{kl} \qquad\qquad k,l = 1,2,3 \\ [p_k, p_l] &= [q_k, q_l] = 0\ .\end{aligned}$$

These can be satisfied e.g. with differential operators in the coordinate space representation, i.e.

$$\begin{aligned}p_k &\to -i\hbar\,\frac{\partial}{\partial q_k} \qquad q_k \to q_k \qquad k = 1,2,3 \\ \text{or}\ \ \vec{p} &\to -i\hbar\,\vec{\nabla} \qquad \vec{r} \to \vec{r}\ .\end{aligned}$$

The time-dependent Schrödinger equation, describing a massive charged particle interacting with an electromagnetic field, thus becomes

$$H\left(-i\hbar\,\vec{\nabla},\vec{r}\right)\;\psi = i\hbar\,\frac{\partial\psi}{\partial t}$$

$$\text{or}\quad \left[\frac{1}{2m}\,\left(-i\hbar\vec{\nabla}-\frac{e}{c}\,\vec{A}\right)^2+e\,\phi\right]\;\psi = i\hbar\,\frac{\partial\psi}{\partial t}$$

$$\text{or}\quad \left[-\frac{\hbar^2}{2m}\,\Delta+\frac{i\hbar e}{2mc}\left(\vec{\nabla}\cdot\vec{A}+\vec{A}\cdot\vec{\nabla}\right)+\frac{e^2}{2mc^2}\,\vec{A}^{\,2}+e\,\phi\right]\;\psi = i\hbar\,\frac{\partial\psi}{\partial t}$$

- Gauge invariance (discovered by H. Weyl in 1927):

  Directly related to the principle of minimal substitution is the invariance of the interacting Schrödinger equation under local $U(1)$ gauge transformations:

$$\begin{cases} \psi \;\to\; \psi' \;=\; \exp\left(-\dfrac{ie}{\hbar c}\,\chi\right)\,\psi & \text{wave function} \\ \vec{A} \;\to\; \vec{A}\,' \;=\; \vec{A}-\vec{\nabla}\,\chi & \text{vector potential} \\ \phi \;\to\; \phi' \;=\; \phi+\dfrac{1}{c}\,\dfrac{\partial\chi}{\partial t} & \text{scalar potential,} \end{cases}$$

  where $\chi=\chi\,(\vec{r},t)\;\in\;\mathcal{R}$ is an arbitrary function of space-time.

  remarks:

  (i) theorem: The gauge transformations

$$\vec{A}\,' \;=\; \vec{A}-\vec{\nabla}\,\chi \quad\text{and}\quad \phi'=\phi+\frac{1}{c}\,\frac{\partial\chi}{\partial t}$$

$$\text{leave}\qquad \vec{E}\underset{\text{def}}{=}-\vec{\nabla}\phi-\frac{1}{c}\,\frac{\partial\vec{A}}{\partial t}\,,\qquad \vec{B}\underset{\text{def}}{=}\vec{\nabla}\times\vec{A}$$

  and thus also Maxwell's equations in vacuo invariant.

  proof: $\vec{E}\,'=-\vec{\nabla}\phi'-\frac{1}{c}\,\frac{\partial\vec{A}\,'}{\partial t}=-\vec{\nabla}\phi-\frac{1}{c}\,\vec{\nabla}\,\frac{\partial}{\partial t}\chi-\frac{1}{c}\,\frac{\partial\vec{A}}{\partial t}+\frac{1}{c}\,\frac{\partial}{\partial t}\,\vec{\nabla}\chi$

$$=\underbrace{-\vec{\nabla}\phi-\frac{1}{c}\,\frac{\partial\vec{A}}{\partial t}}_{\vec{E}}\;\underbrace{-\frac{1}{c}\vec{\nabla}\,\frac{\partial}{\partial t}\chi+\frac{1}{c}\,\frac{\partial}{\partial t}\,\vec{\nabla}\chi}_{\vec{0}}\;=\;\vec{E}$$

$$\vec{B}\,'=\vec{\nabla}\times\vec{A}\,'\;=\;\underbrace{\vec{\nabla}\times\vec{A}}_{\vec{B}}\;-\;\underbrace{\vec{\nabla}\times\vec{\nabla}\chi}_{\vec{0}}\;=\;\vec{B}\qquad\text{q.e.d.}$$

  (ii) If, at the same time, we transform the wavefunction ψ with an Abelian unitary $U(1)$ transformation, $U=\exp\left(-\frac{ie}{\hbar c}\chi\right)$, where the real phase $\chi=\chi(\vec{r},t)$ can be chosen locally, i.e. differently at every space-time point, the interacting Schrödinger equation is invariant, i.e.

$$\left[\frac{1}{2m}\,\left(-i\hbar\,\vec{\nabla}-\frac{e}{c}\,\vec{A}\right)^2+e\phi\right]\psi=0\;\leftrightarrow\;\left[\frac{1}{2m}\,\left(-i\hbar\,\vec{\nabla}-\frac{e}{c}\,\vec{A}\,'\right)^2+e\phi'\right]\psi'=0$$

  (see tut 30).

  (iii) The relativistic quantum field theory, describing the electroweak and strong interactions of the leptons and quarks, is also invariant under local gauge transformations. In this case the gauge group is non-Abelian and semi-simple, i.e. a direct product of simple gauge groups which do not have invariant subgroups.

$$\underbrace{U(1)_Y\otimes SU(2)_T}_{\substack{\text{electroweak}\\ \text{interaction}}}\;\otimes\;\underbrace{SU(3)_C}_{\substack{\text{strong}\\ \text{interaction}}}$$

  Y: weak hypercharge
  T: weak isospin (lefthanded)
  C: colour.

- Coulomb gauge:

  Because of gauge invariance, we are free to choose, e.g. the Coulomb gauge $\vec{\nabla}\cdot\vec{A}=0$. As this condition is in the form of a scalar product, it is invariant under rotations, but not under Lorentz transformations. Under a gauge transformation, the Coulomb gauge condition transforms as

$$\vec{\nabla}\cdot\vec{A}\,'=\vec{\nabla}\cdot\left(\vec{A}-\vec{\nabla}\chi\right)=\vec{\nabla}\cdot\vec{A}-\vec{\nabla}\cdot\vec{\nabla}\chi\;.$$

  Thus, if we want this gauge condition to be valid in any other gauge-transformed system, i.e.

$$\vec{\nabla}\cdot\vec{A}=\vec{\nabla}\cdot\vec{A}\,'=0\;,$$

  the arbitrary phase function $\chi(\vec{r},t)$ must be restricted to $\vec{\nabla}\cdot\vec{\nabla}\chi=\Delta\chi=0$.

  However, the space of all real functions $\chi(\vec{r},t)$, satisfying the Laplace equation, is still very large.

  In the Coulomb gauge, the time-dependent Schrödinger equation simplifies to

$$\left[-\frac{\hbar^2}{2m}\,\Delta+\frac{i\hbar e}{mc}\,\vec{A}\cdot\vec{\nabla}+\frac{e^2}{2mc^2}\,\vec{A}^{\,2}+e\,\phi\right]\;\psi=i\hbar\,\frac{\partial\psi}{\partial t}\;,$$

$$\text{because}\qquad \vec{\nabla}\cdot\left(\vec{A}\psi\right)=\psi\,\underbrace{\vec{\nabla}\cdot\vec{A}}_{=0}+\vec{A}\cdot\vec{\nabla}\,\psi\;.$$

14.3 Homogeneous magnetic field

- theorem: $\vec{A} = -\frac{1}{2}\,\vec{r}\times\vec{B}$ describes a homogeneous magnetic field $\vec{B} = \text{const.}$

  proof: Using $\vec{a}\times(\vec{b}\times\vec{c}) = (\vec{a}\cdot\vec{c})\,\vec{b} - (\vec{a}\cdot\vec{b})\,\vec{c}$ with $\vec{a} = \vec{\nabla}, \vec{b} = \vec{r}, \vec{c} = \vec{B}$ and taking into account that $\vec{\nabla}$ acts on $\vec{r}$ only, we obtain

$$\begin{aligned}\vec{\nabla}\times\vec{A} &= -\tfrac{1}{2}\,\vec{\nabla}\times\left(\vec{r}\times\vec{B}\right)\\ &= -\tfrac{1}{2}\underbrace{\left(\vec{B}\cdot\vec{\nabla}\right)\vec{r}}_{\vec{B}} + \tfrac{1}{2}\underbrace{\left(\vec{\nabla}\cdot\vec{r}\right)}_{3}\,\vec{B}\ .\end{aligned}$$

  Here, we have used

$$\left(\vec{B}\cdot\vec{\nabla}\right)\vec{r} = \left(B_x\frac{\partial}{\partial x} + B_y\frac{\partial}{\partial y} + B_z\frac{\partial}{\partial z}\right)\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}B_x\\B_y\\B_z\end{pmatrix} = \vec{B}$$

$$\Rightarrow\quad \vec{\nabla}\times\vec{A} = -\tfrac{1}{2}\,\vec{B} + \tfrac{3}{2}\,\vec{B} = \vec{B} \qquad \text{q.e.d.}$$

- theorem: $\vec{A} = -\frac{1}{2}\,\vec{r}\times\vec{B}$ fulfils the Coulomb gauge condition $\nabla\cdot\vec{A} = 0$

  proof:

$$\begin{aligned}\vec{\nabla}\cdot\vec{A} &= -\tfrac{1}{2}\,\vec{\nabla}\cdot\left(\vec{r}\times\vec{B}\right)\\ &= -\tfrac{1}{2}\left(\vec{B}\times\vec{\nabla}\right)\cdot\vec{r}\\ &= -\tfrac{1}{2}\underbrace{\left(B_y\frac{\partial}{\partial z} - B_z\frac{\partial}{\partial y}\right)x}_{=0} + \text{cyclic permutations}\end{aligned}$$

$$\Rightarrow\quad \vec{\nabla}\cdot\vec{A} = 0 \qquad \text{q.e.d.}$$

- The Hamilton operator for $\vec{B} = \text{const}$, with $\vec{A}$ obeying the Coulomb gauge condition $\vec{\nabla}\cdot\vec{A} = 0$, is now

$$\begin{aligned}H &= -\frac{\hbar^2}{2m}\Delta + \frac{i\hbar e}{mc}\,\vec{A}\cdot\vec{\nabla} + \frac{e^2}{2mc^2}\,\vec{A}^2 + e\,\phi\\ &= -\frac{\hbar^2}{2m}\Delta - \frac{i\hbar e}{mc}\,\frac{1}{2}\left(\vec{r}\times\vec{B}\right)\cdot\vec{\nabla} + \frac{e^2}{8mc^2}\left(\vec{r}\times\vec{B}\right)^2 + e\,\phi\\ &= -\frac{\hbar^2}{2m}\Delta + \frac{i\hbar e}{2mc}\,\vec{B}\cdot\left(\vec{r}\times\vec{\nabla}\right) + \frac{e^2}{8mc^2}\left(\vec{r}\times\vec{B}\right)^2 + e\,\phi\\ &= -\frac{\hbar^2}{2m}\Delta - \underbrace{\frac{e\hbar}{2mc}\,(-i\vec{r}\times\vec{\nabla})\cdot\vec{B}}_{\vec{M}_L = \frac{e\hbar}{2mc}\,\vec{L}} + \frac{e^2}{8mc^2}\left(\vec{r}\times\vec{B}\right)^2 + e\,\phi\end{aligned}$$

kinetic energy term	paramagnetic term (linear in B)	diamagnetic term (quadratic in B)	Coulomb term

- remarks:

  (i) $\vec{M}_L = \frac{e\hbar}{2mc}\,\vec{L}$ is the magnetic moment operator associated with the orbital motion of the massive charged particle, since classically as well as quantum mechanically, the potential energy V of the magnetic moment $\vec{M}_L$ in a magnetic field $\vec{B}$ is $V = -\vec{M}_L\cdot\vec{B}$.

  (ii) The absolute value of the ratio of the electron's orbital magnetic moment $\vec{M}_L$ and its angular momentum $\vec{L}$ is Bohr's magneton

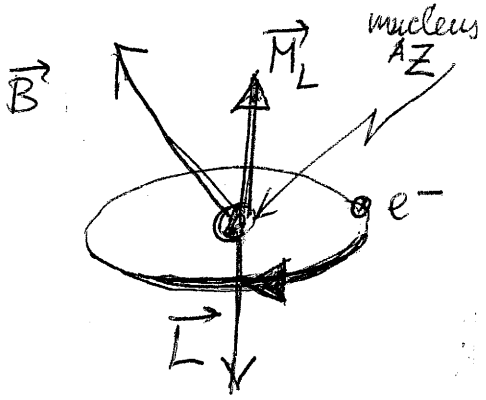
$$\begin{aligned}\mu_B \underset{\text{def}}{=} \left|\frac{\vec{M}_L}{\vec{L}}\right| = \frac{|e|\hbar}{2m_e c} &= 0.578838263\times10^{-8}\ \text{eV/G}\\ &= 0.578838263\times10^{-10}\ \text{MeV/T}.\end{aligned}$$

  (iii) For a proton, this ratio is called the nuclear magneton. It is substantially smaller than the Bohr magneton, i.e.

$$\mu_N \underset{\text{def}}{=} \left|\frac{\vec{M}_L}{\vec{L}}\right|_{m=1} = \frac{|e|\hbar}{2m_p c} = \frac{m_e}{m_p}\,\mu_B = 0.545\times10^{-3}\mu_B\ .$$

14.4 Magnetic moment due to the electron's spin

In mid 1925, S.A. Goudsmit and G.E. Uhlenbeck observed the splitting of a beam of alkali atoms entering a strong inhomogeneous magnetic field, perpendicular to the beam axis and decreasing in the direction of magnetic field in a so-called Stern-Gerlach experiment.



In an external inhomogeneous magnetic field $\vec{B}(\vec{r})$, an atom will, in general, feel both a torque

$$\vec{\tau} \;=\; \vec{M}_L \times \vec{B}(\vec{r}) \;,$$

and a force

$$\vec{F} = -\vec{\nabla}\, V(\vec{r}) \qquad = \; \vec{\nabla}\, \left(\vec{M}_L \cdot \vec{B}(\vec{r})\right) \;.$$

For $\vec{B}(\vec{r}) = B_z(z)\vec{e}_z$ and $\vec{M}_L = M_x\,\vec{e}_x + M_y\,\vec{e}_y$, the torque $\vec{\tau}$ causes the magnetic moment vector $\vec{M}_L$ to precess around the z-axis, while the force $\vec{F}$ is zero.

However, for $\vec{B}(\vec{r}) = B_z(z)\;\vec{e}_z$ and $\vec{M}_L = M_z\,\vec{e}_z$, the torque is zero, while the force is

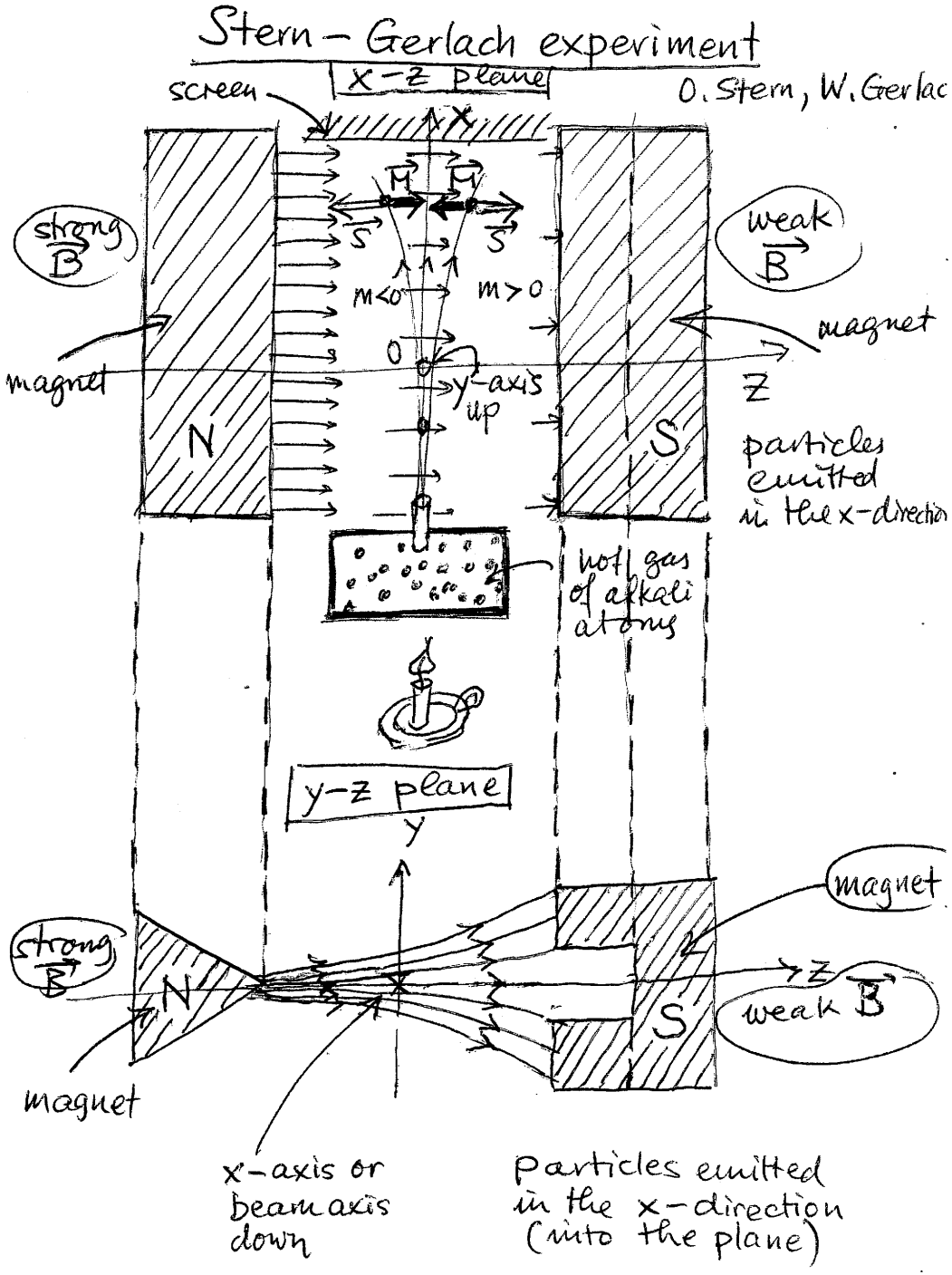
$$\vec{F} = F\,\vec{e}_z = M_z\,\frac{dB_z}{dz}\,\vec{e}_z \;.$$

We thus conclude:

(i) if $M_z > 0$ and $\dfrac{dB_z}{dz} < 0 \;\Rightarrow\; F_z < 0$

(ii) if $M_z < 0$ and $\dfrac{dB_z}{dz} < 0 \;\Rightarrow\; F_z > 0$

(iii) if $M_z = 0$ and $\dfrac{dB_z}{dz} < 0 \;\Rightarrow\; F_z = 0$



- What are our expectations for the outcome of the Stern-Gerlach experiment?

  (i) Alkali atoms, in the ground state, have one electron outside a closed shell, as follows from the exclusion principle postulated by W. Pauli in January 1925. S.A. Goudsmit and G.E. Uhlenbeck argued that the odd electron must be responsible for the total magnetic moment of the atom, because

  (a) the magnetic moments associated with the orbital motion of the electrons in the closed shell will cancel, and

  (b) the nuclear magnetic moments can be neglected as they are smaller by a factor of $\sim 10^{-3}$ than those of the electrons, as we have seen in the case of the orbital magnetic moment of the proton.

  (ii) The odd electron is presumably in a $l = 0$ state, but $l > 0$ cannot be excluded, even for a spherically symmetric potential. The beam would thus split according to the magnetic quantum numbers $m_l = -l, -l+1, \cdots, +l$, as all magnetic substates should be present in a beam emanating from a hot gas due to the equipartition theorem. Thus for $l = 0$ and $m = 0$ we expect no split at all, and for $l = 1$ and $m = 1, 0, -1$, we expect a symmetric split into three beams. However, Goudsmit and Uhlenbeck observed a symmetric split into two beams. It thus seemed as if $l = \frac{1}{2}$ with $m_l = \frac{1}{2}, -\frac{1}{2}$, although l must be integer. Goudsmit and Uhlenbeck concluded that this puzzle may be solved if the odd electron is indeed in a $l = 0$ orbital state as expected, but it has an intrinsic angular momentum called spin with quantum numbers $s = \frac{1}{2}$ and $m_s = \pm\frac{1}{2}$. Indeed, Pauli had already foreseen this possibility of half-integer angular momentum state vectors in his article in the Handbook of Physics, published in early 1925. Associated with the spin should be a magnetic moment. For this brilliant idea, which was initially heavily contested, but eventually turned out to be the correct solution, Goudsmit and Uhlenbeck should have been awarded the Nobel prize.

  (iii) The magnetic moment associated with the electron's orbital motion is $\vec{M}_L = -g_L \frac{|e|\hbar}{2mc} \vec{L}$, with $g_L = 1$, as we have seen. Thus, they suggested $\vec{M}_S = -g_S \frac{|e|\hbar}{2m_e c} \vec{S}$, with $\vec{S} = \frac{\vec{\sigma}}{2}$, σ_k being Pauli's matrices. While we know from the classical theory that $g_L = 1$, there is no classical analogue for the connection of the magnetic moment with spin. Thus the Landé factor g_S must be determined experimentally. Later, we will see that g_S depends on the type of particle $e^-, p, n, d, \mu^-, \cdots$, etc. in question in contrast to g_L.

14.5 Determining g_S in experiment and theory

- experiment: The Landé factors g_S of the charged leptons e^-, e^+, μ^-, and μ^+ have been very accurately determined through experiments on spin precession in a constant magnetic field (e.g. at CERN, Geneva). The most recent g_S-value for the electron is $g_S = 2 \times (1 + 0.001159652209\ \cdots)$. How was this measured?

  (i) spin precession frequency

  The equation of motion for the precession of the magnetic moment $\vec{M}_S$ in a constant magnetic field $\vec{B}$ is

$$\vec{\tau} = \vec{M}_S \times \vec{B} = \frac{d}{dt}\hbar\vec{S} \ .$$

  Inserting $\vec{M}_S = g_S \frac{e\hbar}{2mc} \vec{S}$,

  we have $\frac{d}{dt}\hbar\vec{S} = g_S \frac{e\hbar}{2mc} \vec{S} \times \vec{B}$. Choosing $\vec{B} = B\vec{e}_z$, we obtain

$$\vec{S} \times \vec{B} = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ S_x & S_y & S_z \\ 0 & 0 & B \end{vmatrix} = \{S_y, -S_x, 0\}\ B \ .$$

  Thus, in components, we arrive at

$$\begin{cases} \dot{S}_x &= g_S \frac{eB}{2mc} S_y = \omega_{prec}\, S_y \\ \dot{S}_y &= -g_S \frac{eB}{2mc} S_x = -\omega_{prec}\, S_x \\ \dot{S}_z &= 0 \ , \end{cases}$$

  where $\omega_{prec} = g_S \frac{eB}{2mc}$ is the angular velocity of the spin precession.

  theorem:

  The most general solution to these coupled differential equations is

$$\begin{cases} S_x &= -a\cos(\omega_{prec}\, t + \varphi) \\ S_y &= a\sin(\omega_{prec}\, t + \varphi) \\ S_z &= b \end{cases} \qquad a, b, \omega_{prec} = \text{constants}$$

precessing components / unchanged component

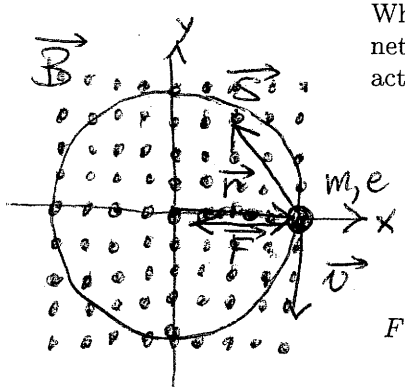
$$\text{Here, } \vec{S} = \underbrace{S_x\, \vec{e}_x + S_y\, \vec{e}_y}_{\vec{S}_\perp} + \underbrace{S_z\, e_z}_{\vec{S}_\parallel}$$

$$a \underset{\text{def}}{=} \sqrt{S_x^2 + S_y^2} = S_\perp^2$$

  proof:

$$\begin{cases} \dot{S}_x &= a\,\omega_{prec}\,\sin(\omega_{prec}\, t + \varphi) = \omega_{prec}\, S_y \\ \dot{S}_y &= a\,\omega_{prec}\,\cos(\omega_{prec}\, t + \varphi) = -\omega_{prec}\, S_x \\ \dot{S}_z &= 0 \end{cases}$$

(ii) cyclotron frequency



While there is no force acting on the magnetic moment, as $\vec{B} = \text{const}$, there is a force acting on the charge of the particle.

Lorentz force = centripetal force

$$\vec{F} = \frac{e}{c}\,\vec{v} \times \vec{B} = -\frac{mv^2}{r}\,\frac{\vec{r}}{r}$$

$$F = \frac{mv^2}{r} = \frac{e}{c}\,vB \quad \Rightarrow \quad \omega_{cycl} = \frac{v}{r} = \frac{eB}{mc}$$

cyclotron angular frequency or velocity

(iii) special case

For $g_S = 2$, the two angular frequencies are equal, i.e.

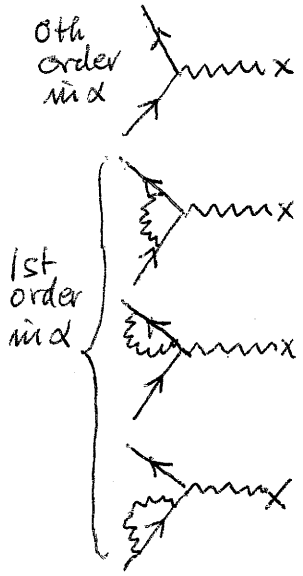
$$\omega_{cycl} = \omega_{prec} = \frac{eB}{mc} \ .$$

Thus, after one full revolution, $\vec{S}$ will again point exactly in the same direction if $g_S = 2$. However if, after a full revolution, $\vec{S}$ points in a different direction, we know $g_S - 2 \neq 0$ with incredible precision, because we can wait for a very large number of revolutions to be completed, limited only by the lifetime of the particle.

- theory

(i) $g_S = 2$ emerges quite naturally from the principle of minimal substitution applied to the relativistic Dirac equation (P.A.M. Dirac, 1928). This is also true for the non-relativistic Pauli equation (W. Pauli, 1926), as we shall see later. Small deviations from $g_S - 2$ can be calculated with astonishing accuracy, in the framework of perturbative quantum electrodynamics (QED). This relativistic quantum field theory was first formulated by W. Pauli, W. Heisenberg and P.A.M. Dirac in 1928, with contributions from E. Fermi in 1932, and it was correctly applied to processes of first order in e, i.e. the emission and absorption of photons by an electron. In 1949, R.P. Feynman, J. Schwinger, S. Tomonaga and F. Dyson found an ingenious trick to hide the singularities, which plagued the higher-order perturbation theory, "under the rug" through the renormalization of the mass and charge of the electron. For this achievement, R.P.Feynman, J. Schwinger and S. Tomonaga were honoured with the Nobel prize in 1965, while F. Dyson was left out.

(ii) J. Schwinger was the first to calculate, in 1949, g_S to first order in the fine structure constant $\alpha = \dfrac{e^2}{\hbar c} = \dfrac{1}{137.03\cdots}$ (i.e. second order in e^2), taking into account Feynman graph contributions to the magnetic moment up to first order in α.



After renormalization Schwinger found

$$g_S = 2\left(1 + \frac{\alpha}{2\pi}\right) + O(\alpha^2) \qquad \alpha = \frac{e}{\hbar c} = \frac{1}{137.03\cdots}$$

fine structure constant

$$= 2(1 + 0.001161409) + O(\alpha^2) \ ,$$

accurate to almost 4 digits. As, in 1949, g_S was experimentally known to more than 3 digits, this was the first significant test of QED.

(iii) Today, QED is, by far, the most accurate theory ever found in Nature, as theory and experiment agree to at least 9 non-trivial digits! The theoretical values are obtained, in perturbation theory up to fourth order in α, in a semi-convergent power series in the fine structure constant

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137.03\cdots} \ .$$

The fact that α includes the fundamental constants of Nature governing Classical Electrodynamics (e), Special Relativity (c) and Quantum Mechanics ($\hbar$), suggests that g_S simultaneously tests these three fundamental theories, which are the pillars of QED, to 9 significant digits. We may therefore, conclude that all these ingredients of QED, must be separately accurate to at least 9 digits, unless there is some strange numerical conspiracy going on between these very different fundamental theories.

At 9 digits the (calculable) weak interactions make an important contribution, and at 10 digits the strong interactions set in. The latter are, unfortunately, not calculable, as the confinement problem has not been solved yet.

14.6 The spin-modified Schrödinger equation

- We now incorporate the spin $\vec{S} = \frac{1}{2}\vec{\sigma}$ of a spin $\frac{1}{2}$ particle in the Schrödinger equation. The Pauli matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are supposed to act on a two-component spinor. Thus ψ must be a two-component spinor wave function

$$\psi(\vec{r},t) = \begin{pmatrix} c_{\frac{1}{2},+\frac{1}{2}}\,(\vec{r},t) \\ c_{\frac{1}{2},-\frac{1}{2}}\,(\vec{r},t) \end{pmatrix} = c_{\frac{1}{2},\frac{1}{2}}\,(\vec{r},t)\,|\psi_{\frac{1}{2},\frac{1}{2}}\rangle + c_{\frac{1}{2},-\frac{1}{2}}\,(\vec{r},t)\,|\psi_{\frac{1}{2},-\frac{1}{2}}\rangle \ ,$$

with $|\psi_{\frac{1}{2},\frac{1}{2}}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|\psi_{\frac{1}{2},-\frac{1}{2}}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Here, the spinors $|\psi_{\frac{1}{2},m_s}\rangle$ are the orthonormal eigenvectors of S^2 and S_z, i.e.

$$\begin{aligned} S^2|\psi_{\frac{1}{2},m_s}\rangle &= \tfrac{3}{4}\,|\psi_{\frac{1}{2},m_s}\rangle \\ S_z|\psi_{\frac{1}{2}\,m_s}\rangle &= m_s|\psi_{\frac{1}{2},m_s}\rangle \end{aligned} \qquad m_s = \tfrac{1}{2}\,,-\tfrac{1}{2} \ .$$

The complex coefficients $c_{\frac{1}{2},\frac{1}{2}}\,(\vec{r},t)$ and $c_{\frac{1}{2},-\frac{1}{2}}\,(\vec{r},t)$ describe the spin up and down states, respectively, z being the quantization axis.

- The time-dependent spin-modified Schrödinger equation for a particle with mass m, charge e and Landé factor g_S moving in an electromagnetic field, is given by

$$H'\,\psi = \left[\frac{1}{2m}\left(-i\hbar\,\vec{\nabla} - \frac{e}{c}\,\vec{A}\right)^2 + e\,\phi\right] I_2\,\psi - g_S\,\frac{e\hbar}{2mc}\,\vec{B}\cdot\vec{S}\,\psi = i\hbar\,I_2\,\frac{\partial\psi}{\partial t}$$

where $g_S = 2\,(1 + 0.001159622209\cdots)$ for the electron and positron.

remarks:

(i) Here the Landé factor g_S is a free parameter that cannot be obtained from the principle of minimal substitution, since $\vec{B}$ is itself gauge invariant and $\vec{S}$ does not involve the coordinates.

(ii) In order to obtain a consistent 2×2 matrix equation, we have multiplied the spin-independent parts with the unit matrix $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, wherever necessary.

- theorem:

$$\vec{\sigma}\cdot\vec{A}\,\vec{\sigma}\cdot\vec{B} = \vec{A}\cdot\vec{B}\,I_2 + i\,\vec{\sigma}\cdot(\vec{A}\times\vec{B})$$

with $[A_k,\sigma_l] = [B_k,\sigma_l] = 0 \qquad (k,l = x,y,z).$

proof:

$$\begin{aligned} \vec{\sigma}\cdot\vec{A}\,\vec{\sigma}\cdot\vec{B} &= (\sigma_x\,A_x + \sigma_y\,A_y + \sigma_z\,A_z)\,(\sigma_x\,B_x + \sigma_y\,B_y + \sigma_z\,B_z) \\ &= \underbrace{\sigma_x^2}_{I_2}\,A_x\,B_x + \underbrace{\sigma_y^2}_{I_2}\,A_y\,B_y + \underbrace{\sigma_z^2}_{I_2}\,A_z\,B_z \\ &\quad + \underbrace{\sigma_x\,\sigma_y}_{i\sigma_z}\,A_x\,B_y + \underbrace{\sigma_y\,\sigma_x}_{-i\sigma_z}\,A_y\,B_x \\ &\quad + \underbrace{\sigma_y\,\sigma_z}_{i\sigma_x}\,A_y\,B_z + \underbrace{\sigma_z\,\sigma_y}_{-i\sigma_x}\,A_z\,B_y \\ &\quad + \underbrace{\sigma_z\,\sigma_x}_{i\sigma_y}\,A_z\,B_x + \underbrace{\sigma_x\,\sigma_z}_{-i\sigma_y}\,A_x\,B_z \\ &= (A_x\,B_x + A_y\,B_y + A_z\,B_z)\,I_2 \\ &\quad + i\sigma_z\,(A_x\,B_y - A_y\,B_x) \\ &\quad + i\sigma_x\,(A_y\,B_z - A_z\,B_y) \\ &\quad + i\sigma_y\,(A_z\,B_x - A_x\,B_z) \\ &= \vec{A}\cdot\vec{B}\,I_2 + i\sigma_x\left(\vec{A}\times\vec{B}\right)_x + i\sigma_y\left(\vec{A}\times\vec{B}\right)_y + i\sigma_z\left(\vec{A}\times\vec{B}\right)_z \\ &= \vec{A}\cdot\vec{B}\,I_2 + i\vec{\sigma}\cdot\left(\vec{A}\times\vec{B}\right) \qquad \text{q.e.d.} \end{aligned}$$

- Pauli's brilliant idea: W. Pauli (1926)

Consider the Hamilton operator H_{Pauli} acting upon the spinor wave function ψ

$$H_{\text{Pauli}}\,\psi = \left\{\frac{1}{2m}\left[\vec{\sigma}\cdot\left(\vec{p} - \frac{e}{c}\,\vec{A}\right)\right]^2 + e\,\phi\,I_2\right\}\,\psi$$

with $\vec{p} = -i\hbar\vec{\nabla}$. Using the mathematical identity $\vec{\sigma}\cdot\vec{A}\,\vec{\sigma}\cdot\vec{B} = \vec{A}\cdot\vec{B}\,I_2 + i\vec{\sigma}\cdot(\vec{A}\times\vec{B})$,

we obtain

$$H_{\text{Pauli}}\psi = \underbrace{\left[\frac{1}{2m}\left(\vec{p}-\frac{e}{c}\vec{A}\right)^2+e\,\phi\right]}_{\text{“doubled” Schrödinger part}} I_2\,\psi + \underbrace{\frac{i\vec{\sigma}}{2m}\cdot\left[\left(\vec{p}-\frac{e}{c}\vec{A}\right)\times\left(\vec{p}-\frac{e}{c}\vec{A}\right)\right]}_{\text{Pauli spin-part}}\psi$$

$$= \left[\frac{1}{2m}\left(\vec{p}-\frac{e}{c}\vec{A}\right)^2+e\,\phi\right] I_2\,\psi + \frac{i\vec{\sigma}}{2m}\cdot\left[\underbrace{\vec{p}\times\vec{p}\,\psi}_{0} - \underbrace{\frac{e}{c}\,\vec{p}\times(\vec{A}\psi) - \frac{e}{c}\,\vec{A}\times\vec{p}\,\psi}_{X=?} + \underbrace{\frac{e^2}{c^2}\,\vec{A}\times\vec{A}\,\psi}_{0}\right] .$$

Here we have to keep in mind that $\vec{p} = -i\hbar\,\vec{\nabla}$ acts on both $\vec{A}$ and ψ in the first term of X, yielding

$$X = -\frac{e}{c}\,\psi\,\vec{p}\times\vec{A} + \frac{e}{c}\,\vec{A}\times\vec{p}\,\psi - \frac{e}{c}\,\vec{A}\times\vec{p}\,\psi$$

$$= +\frac{e}{c}\,\psi\,i\hbar\,\underbrace{\vec{\nabla}\times\vec{A}}_{\vec{B}} = \frac{e}{c}\,\psi\,i\hbar\,\vec{B} = i\hbar\,\frac{e}{c}\,\vec{B}\,\psi \neq 0 .$$

We therefore obtain

$$H_{\text{Pauli}}\,\psi = \left[\frac{1}{2m}\left(\vec{p}-\frac{e}{c}\vec{A}\right)^2+e\,\phi\right] I_2\,\psi + \frac{1}{2m}\,i\,\vec{\sigma}\cdot\frac{e}{c}\,i\hbar\,\vec{B}\,\psi$$

or

$$H_{\text{Pauli}}\,\psi = \underbrace{\left[\frac{1}{2m}\left(\vec{p}-\frac{e}{c}\vec{A}\right)^2+e\,\phi\right]}_{\text{“doubled” Schrödinger part}} I_2\,\psi - \underbrace{\frac{e\hbar}{2mc}\,2\,\frac{\vec{\sigma}}{2}\cdot\vec{B}\,\psi}_{\text{Pauli spin part}} .$$

Comparing $H_{\text{Pauli}}\,\psi$ with

$$H'\,\psi = \left[\frac{1}{2m}\left(\vec{p}-\frac{e}{c}\vec{A}\right)^2+e\phi\right] I_2\,\psi - \frac{e\hbar}{2mc}\,g_S\,\vec{B}\cdot\vec{S}\,\psi ,$$

we conclude that

$$H_{\text{Pauli}} = H' \text{ for } g_S = 2 .$$

Thus, the Landé factor $g_S = 2$ emerges quite naturally from the Pauli equation. This also happens for the relativistic Dirac equation.

- <u>minimal substitution and gauge symmetry</u>

  One can now write the time-dependent Pauli equation in a form exhibiting its $U(1)$ gauge invariance, i.e.

$$\left(i\hbar\,\frac{\partial}{\partial t} - e\,\phi\right) I_2\,\psi = \frac{1}{2m}\left[\vec{\sigma}\cdot\left(\vec{p}-\frac{e}{c}\vec{A}\right)\right]^2\psi$$

$$= \frac{1}{2m}\left[\vec{\sigma}\cdot\left(-i\hbar\,\vec{\nabla}-\frac{e}{c}\vec{A}\right)\right]^2\psi$$

  or

$$i\hbar\,D_0\,I_2\,\psi = \frac{1}{2m}\left[\vec{\sigma}\cdot\left(-i\hbar\,\vec{D}\right)\right]^2\psi ,$$

  with the covariant or gauge invariant derivatives

$$\begin{cases} D_0 \underset{\text{def}}{=} \dfrac{\partial}{\partial t} + i\,\dfrac{e}{\hbar}\,\phi \\ \vec{D} \underset{\text{def}}{=} \vec{\nabla} - i\,\dfrac{e}{\hbar c}\,\vec{A}, \end{cases}$$

  introduced in tut 30.

- <u>remarks</u>

  (i) The Landé factor $g_S = 2$, predicted by the Pauli equation, describes the experimentally measured magnetic moment of the electron very well. The missing 0.1% are consistent with higher-order QED corrections.

  (ii) Pauli's Hamilton operator can be obtained from the non-interacting Hamilton operator $H_{\text{Pauli}} = \dfrac{1}{2m}\,[\vec{\sigma}\cdot\vec{p}]^2$ through a modified principle of minimal substitution

$$\begin{cases} \vec{p} \longrightarrow \vec{p}-\dfrac{e}{c}\,\vec{A} \\ H_{\text{Pauli}} \longrightarrow H_{\text{Pauli}} - e\,\phi\,I_2 \end{cases} \text{ yielding}$$

$$H_{\text{Pauli}} = \frac{1}{2m}\left[\vec{\sigma}\cdot\left(\vec{p}-\frac{e}{c}\vec{A}\right)\right]^2 + e\,\phi\,I_2$$

(iii) It is shown in tut 32 that, for a constant magnetic field $\vec{B}$, H_{Pauli} is also supersymmetric, making it so far the only case in which supersymmetry can be observed in Nature.

(iv) The Pauli equation is obviously non-relativistic, as time and space are treated differently, in fact it is a differential equation that is first order in time and second order in space. A relativistically acceptable differential equation should have derivatives of the same order in space and time.

14.7 Abelian gauge theory: $U(1)$

- covariant derivatives and Special Relativity

The covariant derivatives introduced in tut 30 can be put into a Minkowski vector form

$$D_\mu \underset{\text{def}}{=} \frac{\partial}{\partial x_\mu} - i\,\frac{e}{\hbar c}\,A_\mu \ .$$

In pseudo-Euclidean Minkowski space $\mathcal{M}_4$, we have

$$\frac{\partial}{\partial x_\mu} \underset{\text{def}}{=} \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \\ \frac{\partial}{\partial x_4} \end{pmatrix} \underset{\text{def}}{=} \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \\ \frac{\partial}{\partial ict} \end{pmatrix} \quad \text{space-time gradients}$$

$$A_\mu \underset{\text{def}}{=} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix} \underset{\text{def}}{=} \begin{pmatrix} A_x \\ A_y \\ A_z \\ i\,\phi \end{pmatrix} \quad \text{vector-scalar potentials}$$

$$F_{\mu\nu} \underset{\text{def}}{=} \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} = \begin{pmatrix} 0 & B_z & -B_y & -iE_x \\ -B_z & 0 & B_x & -iE_y \\ B_y & -B_x & 0 & -iE_z \\ iE_x & iE_y & iE_z & 0 \end{pmatrix} \quad \text{electromagnetic field tensor}$$

Here, we have used the definitions

$$\vec{E} \underset{\text{def}}{=} -\vec{\nabla}\,\phi - \frac{1}{c}\,\frac{\partial \vec{A}}{\partial t} \quad \text{for the electric field and}$$

$$\vec{B} \underset{\text{def}}{=} \vec{\nabla} \times \vec{A} \quad \text{for the magnetic field.}$$

The commutators of the covariant space-time derivatives generate the electromagnetic field via

$$\begin{aligned}
[D_\mu, D_\nu]\,\psi &= \left[\frac{\partial}{\partial x_\mu} - i\,\frac{e}{\hbar c}\,A_\mu\ ,\ \frac{\partial}{\partial x_\nu} - i\,\frac{e}{\hbar c}\,A_\nu\right]\psi \\
&= \underbrace{\left[\frac{\partial}{\partial x_\mu}, \frac{\partial}{\partial x_\nu}\right]}_{=0}\psi - i\,\frac{e}{\hbar c}\left[\frac{\partial}{\partial x_\mu}, A_\nu\right]\psi \\
&\quad -i\,\frac{e}{\hbar c}\left[A_\mu, \frac{\partial}{\partial x_\nu}\right]\psi - \frac{e^2}{\hbar^2 c^2}\underbrace{[A_\mu, A_\nu\,]}_{=0}\psi \\
&= -i\,\frac{e}{\hbar c}\,\frac{\partial}{\partial x_\mu}\,(A_\nu\psi) + i\,\frac{e}{\hbar c}\,A_\nu\,\frac{\partial \psi}{\partial x_\mu} \\
&\quad -i\,\frac{e}{\hbar c}\,A_\mu\,\frac{\partial \psi}{\partial x_\nu} + i\,\frac{e}{\hbar c}\,\frac{\partial}{\partial x_\nu}(A_\mu\,\psi) \\
[D_\mu, D_\nu]\,\psi &= -i\,\frac{e}{\hbar c}\left(\frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu}\right)\psi = -i\frac{e}{\hbar c}\,F_{\mu\nu}\,\psi \ .
\end{aligned}$$

known as Bianchi identities. The Maxwell's equations in vacuo are

$$\frac{\partial}{\partial x_\nu}\,F_{\mu\nu} \ = \ 0 \ .$$

special cases:

(i) $\mu = 1, \nu = 2 \ \Rightarrow \ [D_x, D_y] = i\,\frac{e}{\hbar c}\,B_z$ (see tut 30)

(ii) $\mu = 4, \nu = 1 \ \Rightarrow \ [D_4, D_x] = -i\,\frac{e}{\hbar c}\,i\,E_x$

$$= \frac{-i}{c}\,[D_0, D_x]$$

$\Rightarrow \ [D_0, D_x] = i\,\frac{e}{\hbar}\,E_x$ (see tut 30)

remark:

The commutator of the covariant derivatives is proportional to the gauge field tensor also in non-Abelian gauge theories, although it has a somewhat different form.

15 Non-Abelian gauge theory: $SU(2)$

Following C.N. Yang and R.L. Mills (*Phys. Rev.***96**(1954)191), we now want to generalize the gauge principle to include a non-Abelian $SU(2)$ symmetry group, which has led to substantial advances in our understanding of the non-gravitational interactions.

We start here with the Schrödinger equation, describing two different, non-interacting, spinless particles, having the same mass m, i.e.

$$-\frac{\hbar^2}{2m}\,\Delta\,\psi_{\frac{1}{2}}\,(\vec{r},t)\;=\;i\hbar\,\frac{\partial}{\partial t}\,\psi_{\frac{1}{2}}\,(\vec{r},t)\;\;.$$

Here,

$$\psi_{\frac{1}{2}}\,(\vec{r},t)\;\underset{\text{def}}{=}\;\begin{pmatrix} f_{\frac{1}{2},\frac{1}{2}}\,(\vec{r},t) \\ f_{\frac{1}{2},-\frac{1}{2}}\,(\vec{r},t)\end{pmatrix}$$

is a space-time dependent isospinor describing the up and down flavours, $T_3=\frac{1}{2}$ and $T_3=-\frac{1}{2}$, of isospin $T=\frac{1}{2}$. The choice of the isospin projection axis is arbitrary, as the Schrödinger equation is invariant under <u>global</u>, i.e. space-time independent, non-Abelian $SU(2)$ isospin transformations

$$\psi'_{\frac{1}{2}}\,(\vec{r},t)\;=\;\exp\left(-i\,\frac{\vec{\tau}}{2}\cdot\vec{\varphi}\right)\,\psi_{\frac{1}{2}}\,(\vec{r},t)\;\;,$$

where $\vec{\tau}\cdot\vec{\varphi}\underset{\text{def}}{=}\tau^1\varphi^1+\tau^2\varphi^2+\tau^3\varphi^3$ and φ^k $(k=1,2,3)$ are some constant real phases. The components of the isovector operator $\vec{\tau}\underset{\text{def}}{=}\left(\tau^1,\tau^2,\tau^3\right)$ are the Pauli matrices, acting on the two-component isospinor $\psi_{\frac{1}{2}}\,(\vec{r},t)$. The generators $T^k=\frac{1}{2}\,\tau^k$ $(k=1,2,3)$ of the fundamental representation of $SU(2)$ fulfil the Lie algebra

$$\left[\frac{\tau^k}{2},\frac{\tau^l}{2}\right]\;=\;i\,\sum_{m=1}^{3}\,\varepsilon_{klm}\,\frac{\tau^m}{2}\;\;,$$

with

$$\varepsilon_{klm}\;=\;\begin{cases}\;\;1 & \text{for even permutations of 1,2,3}\\ -1 & \text{for odd permutations of 1,2,3}\\ \;\;0 & \text{otherwise}\end{cases}$$

being the structure constants of this Lie group. Similar to the spin rotations, which can be interpreted as rotations in ordinary 3-dimensional Euclidean space $\mathcal{E}_3$, the isospin rotations can be interpreted in a fictitious 3-dimensional Euclidean isospace $\mathcal{T}_3$. The components of the (real) isovector $\vec{\varphi}=(\varphi^1,\varphi^2,\varphi^3)\in\mathcal{T}_3$ define the rotation axis, as well as the rotation angle around this axis in isospace $\mathcal{T}_3$.

We would now like to <u>gauge this $SU(2)$ isospin symmetry</u> by choosing the components (or phases) φ^k <u>locally</u>, i.e. at every space-time point <u>differently</u>: $\varphi^k=\varphi^k(\vec{r},t)$, $k=1,2,3$. This is exactly what we did in the case of $U(1)$ gauge symmetry. In order to incorporate the generalized principle of minimal substitution, which is linked to gauge invariance, we further require that suitably defined generalized space-time derivatives transform under $SU(2)$ gauge transformations like $\psi_{\frac{1}{2}}$, i.e.

$$\mathcal{D}_\mu\,\psi_{\frac{1}{2}}\to\mathcal{D}'_\mu\psi'_{\frac{1}{2}}\;=\;\exp\left(-i\,\frac{\vec{\tau}}{2}\cdot\vec{\varphi}\right)\,\mathcal{D}_\mu\,\psi_{\frac{1}{2}}\;\;.$$

Similar to $U(1)$ gauge theory, this can be achieved by introducing <u>covariant derivatives</u>,

$$\mathcal{D}_\mu\;\underset{\text{def}}{=}\;\frac{\partial}{\partial x_\mu}-i\,\frac{g}{\hbar c}\,\frac{\vec{\tau}}{2}\cdot\vec{\mathcal{A}}_\mu\,(\vec{r},t)\;\;,$$

acting on the isospinor $\psi_{\frac{1}{2}}\,(\vec{r},t)$. Here, the three $SU(2)$ gauge fields $\mathcal{A}^k_\mu\;(\vec{r},t)$ $(k=1,2,3)$ have been combined with the three generators of the $SU(2)$ isospin group $\frac{1}{2}\,\tau^k$ to form an isoscalar product $\frac{1}{2}\,\vec{\tau}\cdot\vec{\mathcal{A}}_\mu(\vec{r},t)$ that is invariant under rotations in isospace $\mathcal{T}_3$. This guarantees that the covariant derivatives do not carry the isovector quantum numbers. The coupling constant g is the analogue of the charge e in Abelian gauge theory. $\mathcal{A}^k_\mu(\vec{r},t)$ $(\mu=1,2,3,4;\;k=1,2,3)$ are vectors, both in 3-dimensional Euclidean isospace $\mathcal{T}_3$, as well as in 4-dimensional pseudo-Euclidean Minkowski space $\mathcal{M}_4$. They are space-time dependent and describe three massless isophotons, similar to the photon in $U(1)$ gauge theory which is described by $A_\mu(\vec{r},t)$.

The commutators of the covariant derivatives fulfil the Bianchi identities

$$[\mathcal{D}_\mu,\mathcal{D}_\nu]\;=\;-\,i\,\frac{g}{\hbar c}\,\vec{\mathcal{F}}_{\mu\nu}\cdot\frac{\vec{\tau}}{2}\;\;,$$

as shown below. The non-Abelian isovector field tensor

$$\vec{\mathcal{F}}_{\mu\nu}\;\underset{\text{def}}{=}\;\frac{\partial\vec{\mathcal{A}}_\nu}{\partial x_\mu}-\frac{\partial\vec{\mathcal{A}}_\mu}{\partial x_\nu}+\frac{g}{\hbar c}\,\vec{\mathcal{A}}_\mu\times\vec{\mathcal{A}}_\nu$$

is the analogue of the Abelian electromagnetic field tensor $F_{\mu\nu}\;\underset{\text{def}}{=}\;\dfrac{\partial A_\nu}{\partial x^\mu}-\dfrac{\partial A_\mu}{\partial x_\nu}$, which we have derived from the Bianchi identity in section 4.7 and tut 30, i.e.

$$[D_\mu,D_\nu]\;=\;-\,i\,\frac{e}{\hbar c}\,F_{\mu\nu}\;\;.$$

However, in contrast to $F_{\mu\nu}$, $\vec{\mathcal{F}}_{\mu\nu}$ is a nonlinear function of the isovector-Minkowski vector potential $\vec{\mathcal{A}}_\mu$. Thus the field equations for $\vec{\mathcal{F}}_{\mu\nu}$ in vacuo, i.e. $\displaystyle\sum_{\nu=1}^{4}\,\partial_\nu\,\vec{\mathcal{F}}_{\mu\nu}=0$, exhibit some nonlinear terms in the gauge field $\vec{\mathcal{A}}_\mu$, describing the self-interactions of the isophotons. In fact, this is the fundamental difference between non-Abelian and Abelian gauge theory, where the electromagnetic field tensor $F_{\mu\nu}$ is given linearly in terms of the photon field A_μ. Accordingly, the electromagnetic field equations in vacuo are $\dfrac{\partial}{\partial x_\nu}\,F_{\mu\nu}=0$ and there are no self-interaction terms of the photon field A_μ in Abelian gauge theory.

We now derive the relation between the $\mathcal{D}_\mu$'s and $\vec{\mathcal{F}}_{\mu\nu}$.

theorem:

$$[\mathcal{D}_\mu, \mathcal{D}_\nu] = -i \, \frac{g}{\hbar c} \, \vec{\mathcal{F}}_{\mu\nu} \cdot \frac{\vec{\tau}}{2}$$

$$\text{with} \quad \vec{\mathcal{F}}_{\mu\nu} \;=\; \frac{\partial \vec{\mathcal{A}}_\nu}{\partial x_\mu} - \frac{\partial \vec{\mathcal{A}}_\mu}{\partial x_\nu} + \frac{g}{\hbar c} \, \vec{\mathcal{A}}_\mu \times \vec{\mathcal{A}}_\nu \;\; .$$

proof:

$$\begin{aligned}
[\mathcal{D}_\mu, \mathcal{D}_\nu] \, \psi_{\frac{1}{2}} \;&=\; \left[\frac{\partial}{\partial x_\mu} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu, \frac{\partial}{\partial x_\nu} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\nu \right] \psi_{\frac{1}{2}} \\
&=\; \left[\frac{\partial}{\partial x_\mu}, \frac{\partial}{\partial x_\nu} \right] \psi_{\frac{1}{2}} - i \, \frac{g}{\hbar c} \left[\frac{\partial}{\partial x_\mu}, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\nu \right] \psi_{\frac{1}{2}} \\
&\quad - i \, \frac{g}{\hbar c} \left[\frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu, \frac{\partial}{\partial x_\nu} \right] \psi_{\frac{1}{2}} - \frac{g^2}{\hbar^2 c^2} \left[\frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\nu \right] \psi_{\frac{1}{2}} \\
&=\; - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \frac{\partial \vec{\mathcal{A}}_\nu}{\partial x_\mu} \, \psi_{\frac{1}{2}} + i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \frac{\partial \vec{\mathcal{A}}_\mu}{\partial x_\nu} \psi_{\frac{1}{2}} \\
&\quad \underbrace{- \frac{g^2}{\hbar^2 c^2} \left[\sum_{k=1}^{3} \frac{\tau^k}{2} \, \mathcal{A}_\mu^k, \sum_{l=1}^{3} \frac{\tau^l}{2} \, \mathcal{A}_\nu^l \right] \psi_{\frac{1}{2}}}_{- \frac{g^2}{\hbar^2 c^2} \sum_{k,l=1}^{3} \mathcal{A}_\mu^k \, \mathcal{A}_\nu^l \underbrace{\left[\frac{\tau^k}{2}, \frac{\tau^l}{2} \right]}_{i \sum_{m=1}^{3} \varepsilon_{klm} \frac{\tau^m}{2}} \psi_{\frac{1}{2}}} \\
&=\; - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \left(\frac{\partial \vec{\mathcal{A}}_\nu}{\partial x_\mu} - \frac{\partial \vec{\mathcal{A}}_\mu}{\partial x_\nu} \right) \psi_{\frac{1}{2}} - i \, \frac{g^2}{\hbar^2 c^2} \sum_{m=1}^{3} \left(\vec{\mathcal{A}}_\mu \times \vec{\mathcal{A}}_\nu \right)^m \frac{\tau^m}{2} \psi_{\frac{1}{2}} \\
&\quad \text{with} \; \left(\vec{\mathcal{A}}_\mu \times \vec{\mathcal{A}}_\nu \right)^m \underset{\text{def}}{=} \sum_{k,l=1}^{3} \varepsilon_{klm} \, \mathcal{A}_\mu^k \, \mathcal{A}_\nu^l \\
&=\; -i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \left(\frac{\partial \vec{\mathcal{A}}_\nu}{\partial x_\mu} - \frac{\partial \vec{\mathcal{A}}_\mu}{\partial x_\nu} \right) \psi_{\frac{1}{2}} - i \, \frac{g^2}{\hbar^2 c^2} \left(\vec{\mathcal{A}}_\mu \times \vec{\mathcal{A}}_\nu \right) \cdot \frac{\vec{\tau}}{2} \psi_{\frac{1}{2}} \\
&=\; -i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \left(\frac{\partial \vec{\mathcal{A}}_\nu}{\partial x_\mu} - \frac{\partial \vec{\mathcal{A}}_\mu}{\partial x_\nu} \right) \psi_{\frac{1}{2}} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \left(\frac{g}{\hbar c} \, \vec{\mathcal{A}}_\mu \times \vec{\mathcal{A}}_\nu \right) \psi_{\frac{1}{2}} \\
&\quad \Rightarrow \;\; [\mathcal{D}_\mu, \mathcal{D}_\nu] \;=\; - \, i \, \frac{g}{\hbar c} \, \vec{\mathcal{F}}_{\mu\nu} \cdot \frac{\vec{\tau}}{2} \\
&\quad \text{with} \;\; \vec{\mathcal{F}}_{\mu\nu} \underset{\text{def}}{=} \frac{\partial \vec{\mathcal{A}}_\nu}{\partial x_\mu} - \frac{\partial \vec{\mathcal{A}}_\mu}{\partial x_\nu} + \frac{g}{\hbar c} \, \vec{\mathcal{A}}_\mu \times \vec{\mathcal{A}}_\nu \;\; \text{q.e.d.}
\end{aligned}$$

Last, but not least, we derive the transformation properties of $\vec{\mathcal{A}}_\mu$ under $SU(2)$ gauge transformations. These are most easily obtained for infinitesimal $SU(2)$ gauge transformations

$$\begin{aligned}
\psi' \;&=\; \left(I_2 - i \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \right) \psi \qquad \text{eq.(1)} \\
\mathcal{D}'_\mu \psi'_\mu \;&=\; \left(I_2 - i \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \right) \mathcal{D}_\mu \psi \;\; , \qquad \text{eq.(2)}
\end{aligned}$$

with $\vec{\alpha} \underset{\text{def}}{=} \frac{g}{\hbar c} \, \vec{\varphi}$ being an infinitesimal space-time dependent vector.
Inserting the inverse of eq.(1)

$$\psi \;=\; \left(I_2 + i \, \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \right) \psi'$$

into eq.(2), we obtain

$$\mathcal{D}'_\mu \, \psi' \;=\; \left(I_2 - i \, \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \right) \mathcal{D}_\mu \left(I_2 + i \, \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \right) \psi' \;\; .$$

Introducing the covariant derivatives yields

$$\left(I_2 \, \frac{\partial}{\partial x_\mu} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}'_\mu \right) \psi' = \left(I_2 - i \, \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \right) \left(I_2 \, \frac{\partial}{\partial x_\mu} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu \right) \left(I_2 + i \, \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \right) \psi'$$

Up to order $\vec{\alpha}$, we thus arrive at

$$\begin{aligned}
\left(I_2 \, \frac{\partial}{\partial x_\mu} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}'_\mu \right) \psi' \;&=\; \left(I_2 \, \frac{\partial}{\partial x_\mu} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu \right) \psi' \\
&\quad - i \, \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \left(I_2 \frac{\partial}{\partial x_\mu} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu \right) \psi' \\
&\quad + \left(I_2 \, \frac{\partial}{\partial x_\mu} - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu \right) i \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \, \psi' \\
- i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}'_\mu \, \psi' \;&=\; - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}'_\mu \, \psi' \\
&\quad \left[- i \, \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \, , \, I_2 \, \frac{\partial}{\partial x_\mu} \right] \psi' \\
&\quad \left[- i \, \vec{\alpha} \cdot \frac{\vec{\tau}}{2} \, , - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu \right] \psi' \\
- i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}'_\mu \, \psi' \;&=\; - i \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}'_\mu \, \psi' \\
&\quad + i \, \frac{\partial \vec{\alpha}}{\partial x_\mu} \cdot \frac{\vec{\tau}}{2} \, \psi' \\
&\quad - \left[\vec{\alpha} \cdot \frac{\vec{\tau}}{2} \, , \, \frac{g}{\hbar c} \, \frac{\vec{\tau}}{2} \cdot \vec{\mathcal{A}}_\mu \right] \psi'
\end{aligned}$$

$$
\begin{aligned}
-\left[\vec{\alpha}\cdot\frac{\vec{\tau}}{2}\,,\,\frac{g}{\hbar c}\,\frac{\vec{\tau}}{2}\cdot\vec{\mathcal{A}}_\mu\right]\psi' &= -\left[\sum_{k=1}^{3}\alpha^k\,\frac{\tau^k}{2}\,,\,\frac{g}{\hbar c}\sum_{l=1}^{3}\frac{\tau^l}{2}\cdot\mathcal{A}_\mu^l\right]\psi' \\
&= -\sum_{k,l=1}^{3}\alpha^k\,\frac{g}{\hbar c}\,A_\mu^l\,\underbrace{\left[\frac{\tau^k}{2}\,,\,\frac{\tau^l}{2}\right]}_{\sum_{m=1}^{3} i\,\varepsilon_{klm}\,\frac{\tau^m}{2}}\psi' \\
&= -i\,\frac{g}{\hbar c}\sum_{k,l,m=1}^{3}\varepsilon_{klm}\,\alpha^k\,A_\mu^l\,\frac{\tau^m}{2}\,\psi' \\
= -i\,\frac{g}{\hbar c}\sum_{m=1}^{3}\left(\vec{\alpha}\times\vec{\mathcal{A}}_\mu\right)^m\,\frac{\tau^m}{2}\,\psi' &= -i\,\frac{g}{\hbar c}\left(\vec{\alpha}\times\vec{\mathcal{A}}_\mu\right)\cdot\frac{\vec{\tau}}{2}\,\psi' \\
\Rightarrow\quad -i\,\frac{g}{\hbar c}\,\frac{\vec{\tau}}{2}\cdot\vec{\mathcal{A}}'_\mu\,\psi' &= -i\,\frac{g}{\hbar c}\,\frac{\vec{\tau}}{2}\cdot\vec{\mathcal{A}}_\mu\,\psi' \\
&\quad + i\,\frac{\partial\vec{\alpha}}{\partial x_\mu}\cdot\frac{\vec{\tau}}{2}\,\psi' \\
&\quad - i\,\frac{g}{\hbar c}\left(\vec{\alpha}\times\vec{\mathcal{A}}_\mu\right)\cdot\frac{\vec{\tau}}{2}\,\psi \\
\Rightarrow\quad I_2\,\vec{\mathcal{A}}'_\mu &= I_2\vec{\mathcal{A}}_\mu + \vec{\alpha}\times\vec{\mathcal{A}}_\mu - \frac{\hbar c}{g}\,\frac{\partial\vec{\alpha}}{\partial x_\mu} \\
\text{with}\qquad \vec{\alpha} &\underset{\text{def}}{=} \frac{g}{\hbar c}\,\vec{\varphi}\ .
\end{aligned}
$$

16 Time-dependent perturbation theory

16.1 General formalism

- Consider the Hamilton operator $H(t) = H_0 + H_1(t)$,

  with H_0 time-independent and $H_1(t)$ time-dependent. $|u_n\rangle$ and E_n are the eigenvectors and eigenvalues of H_0, respectively, i.e.

  $$H_0|u_n\rangle = E_n|u_n\rangle \qquad\qquad n = 1, 2, \cdots\ .$$

  The $|u_n\rangle$'s form a complete and orthonormal set of eigenvectors of H_0. At $t = 0$ the statevector is given by

  $$|\psi(0)\rangle = \sum_{n=1}^{\infty} c_n|u_n\rangle, \ \text{ with } \ c_n = \langle u_n|\psi(0)\rangle \qquad \text{Fourier coefficients.}$$

- The time evolution of this initial state $|\psi(0)\rangle$, due to H_0 alone, is governed by the Schrödinger equation

  $$i\hbar\frac{d}{dt}\,|\psi(t)\rangle = H_0\,|\psi(t)\rangle\ .$$

  Its most general solution that incorporates the initial condition is

  $$|\psi(t)\rangle = \sum_{n=1}^{\infty} c_n\,|u_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_n t}\ .$$

- The time-evolution of the state vector $|\widetilde{\psi}(t)\rangle$, due to $H(t) = H_0 + H_1(t)$, is governed by the time-dependent Schrödinger equation

  $$i\hbar\frac{d}{dt}\,|\widetilde{\psi}(t)\rangle = H\,|\widetilde{\psi}(t)\rangle = \big(H_0 + H_1(t)\big)\,|\widetilde{\psi}(t)\rangle\ . \qquad \text{eq.(1)}$$

  Here we assume that the initial state vector is the same as before, i.e.

  $$|\widetilde{\psi}(0)\rangle = |\psi(0)\rangle = \sum_{n=1}^{\infty} c_n\,|u_n\rangle\ .$$

  A solution of eq.(1) can be attempted in terms of

  $$|\widetilde{\psi}(t)\rangle = \sum_{n=1}^{\infty} c_n(t)\,|u_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_n t}\ ,$$

  with time-dependent Fourier coefficients $c_n(t)$.

strategy: Find a differential equation for the coefficients $c_n(t)$. In mathematics this procedure is called the method of the variation of the constants.

- Inserting
$$|\tilde{\psi}(t)\rangle = \sum_{n=1}^{\infty} c_n(t)\,|u_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_n t}$$
into the time-dependent Schrödinger equation (1), we obtain
$$i\hbar \sum_{n=1}^{\infty} c_n(t)\,\frac{(-i)}{\hbar}\,E_n\,|u_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_n t} + i\hbar \sum_{n=1}^{\infty} \dot{c}_n(t)\,|u_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_n t}$$
$$= \sum_{n=1}^{\infty} c_n(t)\,E_n|u_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_n t} + \sum_{n=1}^{\infty} H_1(t)\,c_n(t)\,|u_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_n t} \quad .$$
The first terms on both sides of the equation cancel.

Applying $\langle u_n|$ from the left, we arrive at
$$i\hbar\,\dot{c}_k(t)\,\mathrm{e}^{-\frac{i}{\hbar}E_k t} = \sum_{n=1}^{\infty} \langle u_k\,|H_1(t)\,u_n\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_n t}\,c_n(t) \quad ,$$
or
$$i\hbar\,\dot{c}_k(t) = \sum_{n=1}^{\infty} [H_1(t)]_{kn}\,\mathrm{e}^{i\omega_{kn}t}\,c_n(t) \quad ,$$
with
$$[H_1(t)]_{kn} \underset{\text{def}}{=} \langle u_k\,|H_1(t)\,u_n\rangle$$
and
$$\omega_{kn} \underset{\text{def}}{=} \frac{E_k - E_n}{\hbar} \quad .$$

theorem:
$$[H_1(t)]_{kn}\,\mathrm{e}^{i\omega_{kn}t} = \Big[\underbrace{\mathrm{e}^{\frac{iH_0 t}{\hbar}}\,H_1(t)\,\mathrm{e}^{-\frac{iH_0 t}{\hbar}}}_{\hat{H}_1(t)}\Big]_{k,n}$$

$\hat{H}_1(t)$: interaction operator in the Dirac picture

$|\hat{\psi}(t)\rangle$: state vector in the Dirac picture

proof:
$$\langle u_k\,|\mathrm{e}^{\frac{iH_0 t}{\hbar}}\,H_1(t)\,\mathrm{e}^{-\frac{iH_0 t}{\hbar}}\,u_n\rangle = \langle \mathrm{e}^{-\frac{iH_0 t}{\hbar}}\,u_k\,|H_1(t)\,\mathrm{e}^{-\frac{iH_0 t}{\hbar}}\,u_n\rangle$$
$$= \langle \mathrm{e}^{-\frac{iE_k t}{\hbar}}\,u_k\,|H_1(t)\,\mathrm{e}^{-\frac{iE_n t}{\hbar}}\,u_n\rangle$$
$$= \langle u_k\,|H_1(t)\,u_n\rangle\,\underbrace{\mathrm{e}^{i\frac{E_k - E_n}{k}t}}_{\mathrm{e}^{i\omega_{kn}t}} \qquad \text{q.e.d.}$$

- Our problem can be formulated in three different ways:

(i) in terms of Fourier coefficients:
$$i\hbar\frac{d}{dt}\,c_k(t) = \sum_{n=1}^{\infty} \Big[\hat{H}(t)\Big]_{kn}\,c_n(t) \quad \text{with} \quad c_n(0) = c_n \quad ,$$

(ii) in terms of a Schrödinger equation:
$$i\hbar\frac{d}{dt}\,|\hat{\psi}(t)\rangle = \hat{H}(t)\,|\hat{\psi}(t)\rangle$$
$$\text{with} \quad |\hat{\psi}(0)\rangle = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} \quad ,$$

(iii) or in terms of a time-evolution operator equation:
$$\frac{d}{dt}\,U(t) = -\frac{i}{\hbar}\,\hat{H}_1(t)\,U(t) \quad \text{with} \quad U(0) = I \quad . \qquad \text{eq.(2)}$$
Here we have introduced the time-evolution operator in the interaction picture
$$|\hat{\psi}(t)\rangle \underset{\text{def}}{=} U(t)\,|\hat{\psi}(0)\rangle \quad .$$

remark:

As, in general, $\Big[\hat{H}_1(t), \hat{H}_1(t')\Big] \neq 0$ for $t \neq t'$, this differential equation cannot be solved in closed form. However, we can get implicit solutions that eventually yield a perturbation expansion in powers of $\hat{H}_1(t)$.

- Integrating eq.(2) from zero to t, we obtain
$$\int_0^t \frac{d}{dt_1}\,U(t_1)\,dt_1 = U(t_1)\Big|_{t_1=0}^{t_1=t} = U(t) - \underbrace{U(0)}_{I} = -\frac{i}{\hbar}\int_0^t dt_1\,\hat{H}_1(t_1)\,U(t_1) \quad ,$$
or the implicit solution
$$U(t) = I + \left(-\frac{i}{\hbar}\right)\int_0^t dt_1\,\hat{H}_1(t_1)\,U(t_1) \quad . \qquad \text{eq.(3)}$$

Inserting eq.(3) repeatedly on the right-hand side of eq.(3), one can easily generate a whole sequence of implicit solutions, like

$$U(t) = I - \frac{i}{\hbar}\int_0^t dt_1\,\hat{H}_1(t_1)\left(I - \frac{i}{\hbar}\int_0^{t_1} dt_2\,\hat{H}_1(t_2)\,U(t_2)\right) \qquad \text{eq.(4)}$$

or

$$U(t) = I + \left(-\frac{i}{\hbar}\right)\int_0^t dt_1\,\hat{H}_1(t_1) + \left(\frac{-i}{\hbar}\right)^2\int_0^t dt_1\,\hat{H}_1(t_1)\int_0^{t_1} dt_2\,\hat{H}_1(t_2)\,U(t_2) \quad \text{eq.(5)}$$

etc.
Approximating $U(t_n)$ by $U(t_n) = I$ on the right-hand sides of eqs.(3), (4), etc., one obtains the following successive approximations of $U(t)$:

to first order:

$$U_1(t) = I + \left(\frac{-i}{\hbar}\right)\int_0^t dt_1\,\hat{H}_1(t_1) \qquad \text{from eq.(3)}$$

to second order:

$$U_2(t) = I + \left(\frac{-i}{\hbar}\right)\int_0^t dt_1\,\hat{H}_1(t_1) + \left(-\frac{i}{\hbar}\right)^2\int_0^t dt_1\,\hat{H}_1(t_1)\int_0^{t_1} dt_2\,\hat{H}_1(t_2)$$

from eq.(4)

$\vdots$

to infinite order:

Dyson's expansion

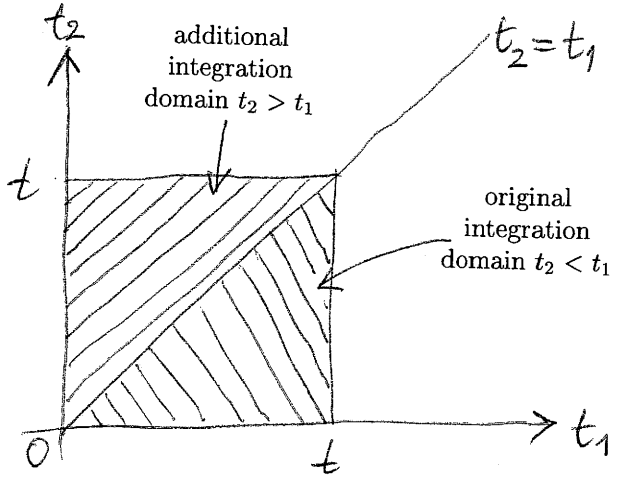
$$U_\infty(t) = I + \sum_{n=1}^{\infty}\left(-\frac{i}{\hbar}\right)^n\int_0^t dt_1\int_0^{t_1} dt_2\cdots\int_0^{t_{n-1}} dt_n\,\hat{H}_1(t_2)\cdots\hat{H}_1(t_n)\ .$$

remarks:

(i) In this approximation scheme we may identify the term $U_0(t) = I$ as the zeroth order approximation.

(ii) In 1949, Freeman Dyson replaced these virtually intractable nestled integrals by multiple integrals, all extending from 0 to t.

To see whether Dyson's idea works, let us look at the second order term in the perturbation expansion:



Dyson's idea is that we should add to the original integral I_2, a further integral

$$I_2' = \int_0^t dt_2\int_0^{t_2} dt_1\,\hat{H}_1(t_2)\,\hat{H}_1(t_1)\ .$$

$I_2 = I_2'$, because I_2' is obtained from I_2 by interchanging the two dummy integration variables $t_1 \leftrightarrow t_2$. However, we note that we cannot combine these two integrals into one single integral, divided by two, because the intergrands are different, i.e. $\hat{H}_1(t_1)\hat{H}_1(t_2) \neq \hat{H}_1(t_2)\,\hat{H}_1(t_1)$. In order to have a single intergrand defined in both integration domains, Dyson introduced the time-ordering operator T with the following property:

$$T\left[\hat{H}_1(t_1)\,\hat{H}_1(t_2)\right] \underset{\text{def}}{=} \begin{cases} \hat{H}_1(t_1)\,\hat{H}_1(t_2) & \text{if } t_2 < t_1 \\ \quad \text{original ordering} & \\ \hat{H}_1(t_2)\,\hat{H}_1(t_1) & \text{if } t_2 > t_1 \\ \quad \text{ordering inverted} & \end{cases}$$

Thus T orders the operators such that their arguments increase from the right to the left. We can therefore write in second order

$$I_2 = \int_0^t dt_1\int_0^{t_1} dt_2\hat{H}_1(t_1)\,\hat{H}(t_2) = \frac{1}{2!}\int_0^t\int_0^t T\left[\hat{H}_1(t_1)\hat{H}(t_2)\right]dt_1 dt_2\ .$$

This idea can be generalized to include multiple integral terms to all orders,

$$\begin{aligned} U_\infty(t) &= I + \sum_{n=1}^{\infty} \left(\frac{-i}{\hbar}\right)^n \int_0^t dt_1 \int_0^t dt_2 \cdots \int_0^{t_{n-1}} dt_n \, \hat{H}_1(t_1)\, \hat{H}_1(t_2) \cdots \hat{H}_1(t_n) \\ &= I + \sum_{n=1}^{\infty} \left(\frac{-i}{\hbar}\right)^n \frac{1}{n!} \int_0^t \cdots \int_0^t T\left[\hat{H}_1(t_1) \cdots \hat{H}_1(t_n)\right] dt_1 \cdots dt_n \end{aligned}$$

remarks:

(i) Here again the time-ordering operator T orders the operators such that their time arguments increase from right to left.

(ii) The factor $n!$ takes care of the number of possible permutations of time variables $t_1, t_2 \cdots, t_n$, and thus identical integrals generated in this way. The simplification of these terms is very useful, because the time-ordered operator products can be easily expanded into normal-ordered products. These can then be interpreted in terms of Feynman graphs (see Relativistic Quantum Field Theory).

16.2 Example: constant perturbation in first order perturbation theory

- We now specialize these equations to first order perturbation theory yielding either

$$U(t) = I - \frac{i}{\hbar} \int_0^t dt' \, \hat{H}_1(t')$$

or

$$|\psi(t)\rangle = |\psi(0)\rangle - \frac{i}{\hbar} \int_0^t dt' \, \hat{H}_1(t') \, |\psi(0)\rangle$$

or

$$c_n(t) = c_n(0) - \frac{i}{\hbar} \int_0^t dt' \sum_{k=1}^{\infty} \left[\hat{H}_1(t')\right]_{nk} c_k(0) \ .$$

Using the latter formula and assuming that only the state $|u_m\rangle$ is populated at $t = 0$, with unit amplitude i.e. $c_k(0) = \delta_{km}$ $(k = 1, 2, \cdots)$, the Fourier coefficients $c_n(t)$ simplify

$$\boxed{\text{for} \quad n \neq m} \quad \text{to}$$

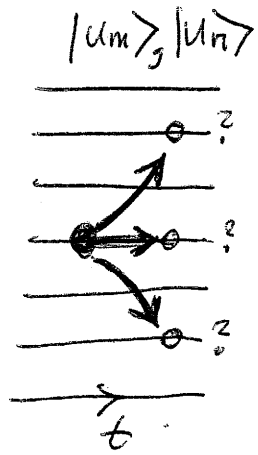
$$\begin{aligned} c_n(t) &= -\frac{i}{\hbar} \int_0^t \left[\hat{H}_1(t')\right]_{nm} dt' \\ &= -\frac{i}{\hbar} \int_0^t [H_1(t')]_{nm} \, e^{i\omega_{nm}t'} dt' \end{aligned}$$

and $\boxed{\text{for} \quad n = m}$ to

$$c_m(t) = 1 - \frac{i}{\hbar} \int_0^t [H_1(t')]_{m,m} \, dt' \ .$$

For a constant perturbation, these integrals yield

$\boxed{\text{for} \quad n \neq m}$



$$\begin{aligned} c_n(t) &= -\frac{i}{\hbar} [H_1]_{nm} \left. \frac{e^{i\omega_{nm}t}}{i\omega_{nm}} \right|_0^t \\ c_n(t) &= -[H_1]_{nm} \frac{e^{i\omega_{nm}t} - 1}{\hbar\omega_{nm}} \end{aligned}$$

and $\boxed{\text{for} \quad n = m}$

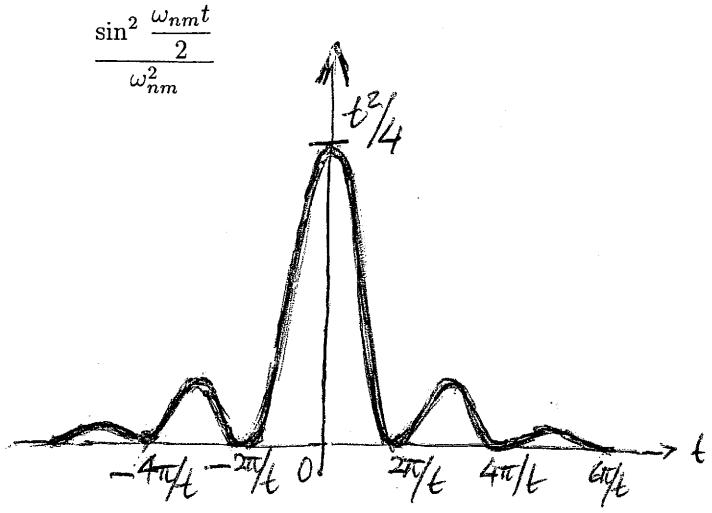
$$c_m(t) = 1 - \frac{i}{\hbar} [H_1]_{mm} \, t \ .$$

- The probability for a transition from a state $|u_m\rangle$ to a state $|u_n\rangle$ $(n \neq m)$ is, in first order perturbation theory, given by

$$P(m \to n) = |c_n(t)|^2 = \frac{|[H_1]_{nm}|^2}{\hbar^2} \frac{e^{i\omega_{nm}t} - 1}{\hbar\omega_{nm}^2} = \frac{|[H_1]_{nm}|^2}{\hbar^2} \frac{4\sin^2 \frac{\omega_{nm}t}{2}}{\omega_{nm}^2} \ ,$$

where we have used the identity

$$\begin{aligned} \left(e^{i\omega t} - 1\right)\left(e^{-i\omega t} - 1\right) &= 2 - \left(e^{i\omega t} + e^{-i\omega t}\right) = 2\left(1 - \cos\omega t\right) \\ &= 2\left(1 - \cos^2\frac{\omega t}{2} + \sin^2\frac{\omega t}{2}\right) = 4\sin^2\frac{\omega t}{2} \ . \end{aligned}$$



For $t \to \infty$, this looks almost like Dirac's delta function.

definition: $$\delta_t(\omega) \underset{\text{def}}{=} \frac{2}{\pi t} \frac{\sin^2 \frac{\omega t}{2}}{\omega^2}$$

- theorem: $$\lim_{t\to\infty} \delta_t(\omega) = \delta(\omega)$$

proof:

(i) $\lim_{t\to\infty} \delta_t(\omega)$ is normalized to unity

$$\int_{-\infty}^{\infty} \delta_t(\omega)\, d\omega = \frac{2}{\pi t} \int_{-\infty}^{\infty} \frac{\sin^2 \frac{\omega t}{2}}{\omega^2}\, d\omega$$

$$= \frac{2}{\pi t} \int_{-\infty}^{\infty} \frac{\sin^2 x}{\left(\frac{2x}{t}\right)^2} \frac{2}{t}\, dx = \frac{1}{\pi} \underbrace{\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2}\, dx}_{\pi} = 1$$

$$\Rightarrow \lim_{t\to\infty} \int_{-\infty}^{\infty} \delta_t(\omega)\, d\omega = 1$$

(ii) $\lim_{t\to\infty} \delta_t(\omega) = 0$ except at $\omega = 0$

For $t > 0$ and $\omega \neq 0$, we have

$$\begin{array}{ccccccc} & 0 & \leqslant & \delta_t(\omega) = \frac{2}{\pi t} & \frac{\sin^2 \frac{\omega t}{2}}{\omega^2} & \leqslant & \frac{2}{\pi t} \frac{1}{\omega^2} \\ \lim\limits_{t\to\infty} & \downarrow & & & \downarrow & & \downarrow \\ & 0 & \leqslant & & \lim\limits_{t\to\infty} \delta_t(\omega) & \leqslant & 0 \end{array}$$

$$\Rightarrow \lim_{t\to\infty} \delta_t(\omega) = 0 \quad \text{for} \quad \omega \neq 0$$

(iii) $\lim_{t\to\infty} \delta_t(\omega) = \infty$ at $\omega = 0$

$$\delta_t(0) = \lim_{\omega\to 0} \delta_t(\omega) = \lim_{\omega\to 0} \frac{2}{\pi t} \frac{\left(\frac{\omega t}{2}\right)^2}{\omega^2} = \frac{t}{2\pi}$$

$$\Rightarrow \lim_{t\to\infty} \delta_t(0) = \lim_{t\to\infty} \frac{t}{2\pi} = \infty$$

In summary, we have shown that $\lim_{t\to\infty} \delta_t(\omega)$ fulfils the three requirements of $\delta(\omega)$, i.e.

(i) $$\int_{-\infty}^{\infty} \delta(\omega)\, d\omega = 1$$

(ii) $$\delta(\omega) = 0 \text{ for } \omega \neq 0$$

(iii) $$\delta(\omega) = \infty \text{ for } \omega = 0\ .$$

remark:

These are the three attributes defining the delta function, as it was originally introduced by P.A.M. Dirac around 1925, also the founder of the mathematical theory of distributions ≡ generalized functions.

We thus conclude that

$$\delta(\omega) = \lim_{t\to\infty} \delta_t(\omega) = \lim_{t\to\infty} \frac{2}{\pi t} \frac{\sin^2 \frac{\omega t}{2}}{\omega^2}$$

is, indeed, a representation of the δ function q.e.d.

remark: For completeness, we mention again that Dirac's function $\delta(x-x')$ is a generalization of Kronecker's $\delta_{nn'}$, where the integer variables n and n' are replaced by continuous real variables x and x'

$$\begin{array}{ccccccc} f(x) & = & \int_{-\infty}^{\infty} f(x')\ \delta(x'-x)\,dx', & & \int_{-\infty}^{\infty} \delta(x'-x)\,dx' & = & 1 \\ \Updownarrow & & \Updownarrow & & \Updownarrow & & \Updownarrow \\ a_n & = & \sum_{n'=1}^{\infty} a_{n'}\,\delta_{n'n}\ , & & \sum_{n'=1}^{\infty} \delta_{n'n} & = & 1 \end{array}$$

- The probability for a transition $|u_m\rangle \to |u_n\rangle$ is

$$P_t(m\to n) = |c_n(t)|^2 = 2\pi\,\frac{\left|[H_1]_{nm}\right|^2}{\hbar^2}\,t\,\frac{2}{\pi t}\,\frac{\sin^2\frac{\omega_{nm}t}{2}}{\omega_{nm}^2}\ ,$$

and the transition probability per unit time is, therefore,

$$p_t(m\to n) = \frac{2\pi}{\hbar^2}\left|[H_1]_{nm}\right|^2\frac{2}{\pi t}\,\frac{\sin^2\omega_{nm}t}{\omega_{nm}^2}\ .$$

In the limit $t\to\infty$, i.e. asymptotically, the transition probability per unit time becomes

$$p_\infty(m\to n) = \frac{2\pi}{\hbar^2}\left|[H_1]_{nm}\right|^2\delta(\omega_{nm})\ ,$$

i.e. independent of time. One may replace the argument of the delta function ω_{nm} by $E_n - E_m$ using the identity

$$\delta(\omega) = \hbar\,\delta(\hbar\omega)\ ,$$

which fulfils all requirements of the delta function, in particular the normalization

$$\int_{-\infty}^{\infty}\hbar\delta(\hbar\omega)\,d\omega = \int_{-\infty}^{\infty}\delta(\omega)\,d\omega = 1\ .$$

We thus obtain Fermi's golden rule (E. Fermi $\sim$ 1930, Nobel Prize 1938)

$$\boxed{p_\infty(m\to n) = \frac{2\pi}{\hbar}\left|[H_1]_{nm}\right|^2\delta(E_n-E_m)}$$

remarks:

(i) Dirac's delta function governs energy conservation. As $\delta(E_n-E_m)=0$ for $E_n\neq E_m$, a constant perturbation is not able to induce a transition, because asymptotically it does not carry energy, as it is not time-dependent.

(ii) Fermi's golden rule can easily be generalized to include perturbation theory to infinite order. The probability to infinite order for a transition from a state $|\hat\psi_m\rangle$ to $|\hat\psi_n\rangle$ is governed by

$$P_t(m\to n) = |<\hat\psi_n|U_\infty(t)\hat\psi_m>\ ,$$

which will also generate an energy conserving delta function $\delta(E_n-E_m)$. In first order perturbation theory, we obtain the terms derived earlier.

16.3 Example: periodic interaction in first order perturbation theory

Let us now discuss a Hermitean perturbation oscillating with a circular frequency ω

$$H_1(t) = F\,e^{-i\omega t} + F^\dagger\,e^{i\omega t}\qquad \omega>0\ . \qquad \text{eq.(1)}$$

Here F is a time independent operator, which carries an energy $E=\pm\hbar\omega$, according to Planck's law. If at $t=0$ the Fourier coefficients are $c_m(0)=\delta_{mk}\quad(k=1,2,\cdots)$, $c_m(t)$ is in first order perturbation theory given by

$$c_n(t) = -\frac{i}{\hbar}\int_0^t [H_1(t')]_{nm}\,e^{i\omega_{nm}t'}\,dt'\ , \qquad \text{eq.(2)}$$

with $n\neq m$ and $\omega_{nm} \underset{\text{def}}{=} \dfrac{E_n-E_m}{\hbar}$.

Inserting eq.(1) into eq.(2), we obtain

$$\begin{aligned} c_n(t) &= -\frac{i}{\hbar}\left\{F_{nm}\int_0^t e^{i(\omega_{nm}-\omega)t'}\,dt'\right. \\ &\quad \left. + \underbrace{(F^\dagger)_{nm}}_{F^*_{mn}}\int_0^t e^{i(\omega_{nm}+\omega)t'}\,dt'\right\} \end{aligned}$$

or

$$c_n(t) = -F_{nm}\,\frac{e^{i(\omega_{nm}-\omega)t}-1}{\hbar(\omega_{nm}-\omega)} - F^*_{mn}\,\frac{e^{i(\omega_{nm}+\omega)t}-1}{\hbar(\omega_{nm}+\omega)}\qquad \text{with}\quad \omega>0\ .$$

remark:

This equation differs from that for a constant perturbation only by the frequency shifts $\pm\omega$, which lead to frequency shifts of $\pm\omega$ in the arguments of the δ-function. The transition probability per unit time has thus two terms

$$p_\infty\,(m \to n) = \frac{2\pi}{\hbar^2}\left|F_{nm}\right|^2 \delta\,(\omega_{nm} - \omega) \;+\; \frac{2\pi}{\hbar^2}\left|F^*_{mn}\right|^2 \delta\,(\omega_{nm} + \omega) \qquad\qquad \omega > 0\ .$$

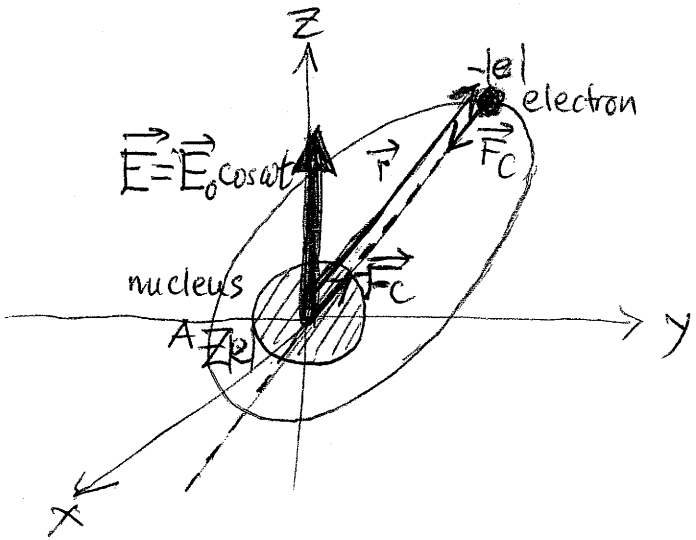
The transition can now take place if $\omega = \pm\omega_{nm}$ or $E_n = E_m \pm \hbar\omega$.

- We now apply these equations to an electron of mass m and charge $-|e|$, bound in a hydrogen-like atom, and subject to an oscillating electric field $\vec{E}(t) = (0, 0, E_z(\omega))\cos\omega t$. The period of the perturbation of the Hamilton operator H_0 causes the system to emit or absorb an energy $\hbar\omega$ via

$$H_1(t) = |e|E_z(\omega)z\ \cos\omega t = |e|E_z(\omega)z\ \frac{\left(\mathrm{e}^{i\omega t} + \mathrm{e}^{-i\omega t}\right)}{2}\ .$$

Here $|e|E_z(\omega)z$ is the potential energy of the electron in the field $\vec{E}$.

Classically, we have



$\vec{r}$ = relative vector between nucleus and electron

The probability per unit time for a transition $|u_m\rangle \to |u_n\rangle$ (with $m \neq n$) in first order is

$$p_\infty\,(m \to n) \;=\; \frac{2\pi}{\hbar^2}\,\frac{e^2\,(E_z(\omega))^2}{4}\left|z_{nm}\right|^2\Big[\delta\,(\omega_{nm} - \omega) + \delta(\omega_{nm} + \omega)\Big]\ .$$
$$\omega > 0$$

- In thermal equilibrium, the energy density of the electric radiation field, at a circular frequency ω in a cavity, consists of three components in mutually orthogonal directions, e.g.

$$u(\omega) \;=\; \frac{1}{8\pi}\,\left(E_x^2(\omega) + E_y^2(\omega) + E_z^2(\omega)\right)$$

(see Classical Electrodynamics).

Only the electric field component in direction of the matrix element of the relative vector

$$\vec{r}_{nm} \underset{\text{def}}{=} \langle u_n|\vec{r}\,u_m\rangle = x_{nm}\,\vec{e}_x + y_{nm}\,\vec{e}_y + z_{nm}\,\vec{e}_z\ ,$$

contributes to the transition. As this comprises only a third of the total energy density, we have

$$p_\infty\,(m \to n) \;=\; \frac{2\pi}{\hbar^2}\,\frac{e^2}{4}\,\frac{1}{3}\,8\pi\,u(\omega)\left|\vec{r}_{nm}\right|^2\Big[\delta\,(\omega_{nm} - \omega) + \delta(\omega_{nm} + \omega)\Big]\ .$$
$$\omega > 0$$

- Integrating over the circular frequency ω yields

$$\int_0^\infty p_\infty(m \to n)\,d\omega = \frac{4\pi^2 e^2}{3\hbar^2}\,\left|\vec{r}_{nm}\right|^2\Big[u(\omega_{nm}) + u(-\omega_{nm})\Big]\ .$$

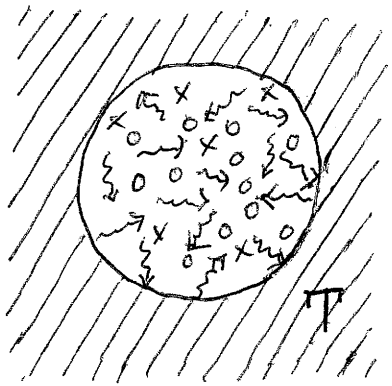
The first term describes the induced emission of a photon $m \to n$, while the second term stands for the induced absorption of a photon, as $-\omega_{nm} = \omega_{mn}$.

17 Radiative Decays

17.1 Radiation in thermal equilibrium with matter

- We would now like to establish relations between the various transition amplitudes, in a gas at temperature T, consisting of atoms with only two eigenstates, i.e. the groundstate n and one excited state m.

A. Einstein (1917), Nobel Prize 1922

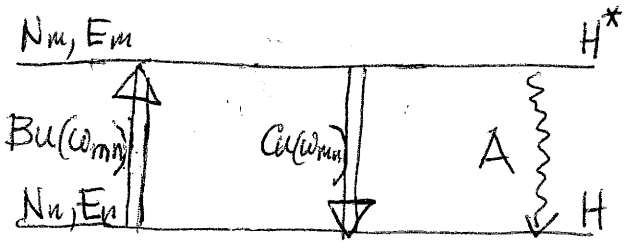


T: temperature of the heat bath

o: atom in the ground state n

x: atom in the excited state m

$\rightsquigarrow$: photon with energy $\hbar\omega_{mn}$



simplest case:

We assume that the degeneracy of the levels is $g_m = g_n = 1$.

induced absorption	induced emission	spontaneous emission
$\gamma + H \to H^*$	$\gamma + H^* \to \underbrace{\gamma + \gamma}_{\text{these photons have the same direction and energy}} + H$	$H^* \to \gamma + H$

N_k: number of atoms in the state $|u_k\rangle$ $\qquad k = m, n$.

E_k: energy of state $|u_k\rangle$.

- In thermal equilibrium at the temperature T, the number of atoms in the state $|u_m\rangle$ is

$$N_m = N \exp\left(-\frac{E_m}{kT}\right)$$ Boltzmann factor (see Statistical Mechanics).

L. Boltzmann ($\sim$ 1871).

The constant N is determined by the total number of atoms in the gas and it drops out in the ratio of the numbers of atoms in the two states, i.e.

$$\frac{N_m}{N_n} = \exp\left(-\frac{E_m - E_n}{kT}\right) = \exp\left(-\frac{\hbar\omega_{mn}}{kT}\right) .$$

The reaction probabilities and numbers of atoms in the excited and groundstates must respect the conservation of the total number of atoms.

$$\underbrace{B\,u(\omega_{mn})}_{\text{probability for induced absorption}} N_n = \Big[\underbrace{A}_{\text{probability for spontaneous emission}} + \underbrace{C u(\omega_{mn})}_{\text{probability for induced emission}} \Big] N_m \ .$$

Dividing both sides by $N_m\, B u(\omega_{mn})$ and using $N_m/N_n = \exp(\hbar\omega_{mn}/kT)$, we get

$$\frac{N_n}{N_m} = \frac{A}{Bu(\omega_{mn})} + \frac{C}{B} = \exp\frac{\hbar\omega_{mn}}{kT}$$

$$\text{or} \quad \frac{A}{Bu(\omega_{mn})} = \exp\frac{\hbar\omega_{mn}}{kT} - \frac{C}{B} \ .$$

Taking the inverse on both sides, we end up with

$$\frac{B}{A}\,u(\omega_{mn}) = \frac{1}{\exp\frac{\hbar\omega_{mn}}{kT} - \frac{C}{B}}$$

$$\text{or} \quad u(\omega_{mn}) = \frac{A/B}{\exp\frac{\hbar\omega_{mn}}{kT} - \frac{C}{B}} \ .$$

Comparing this result with Planck's law

$$u(\omega)\, d\omega = \frac{\hbar\omega^3 / (\pi^2 c^3)}{\exp\frac{\hbar\omega^3}{kT} - 1}\, d\omega \ ,$$

M. Planck (1900)
Nobel Prize 1901

we conclude that, for $\omega = \omega_{mn}$, we have

$$\begin{cases} A/B = \hbar\omega_{nm}^3 / (\pi^2 c^3) \\ C/B = 1 \ . \end{cases}$$

A. Einstein (1917)
Nobel Prize 1922

- The probability per unit time for the spontaneous emission of a photon is, in first order perturbation theory, given by

$$A = \frac{h\omega_{nm}^3}{\pi c^3} \; B = \frac{4e^2\omega_{nm}^3}{3\hbar c^3}\left|\vec{r}_{nm}\right|^2 \quad .$$

Here, ω_{nm} is the circular frequency of the emitted radiation, i.e.

$$\hbar\omega_{mn} = E_m - E_n$$

and

$$r_{nm} = \langle u_n|\vec{r}u_m\rangle \underset{\text{def}}{=} \int u_n^*(\vec{r})\,\vec{r}\,u_m(\vec{r})\,d^3r \quad ,$$

where $e\vec{r}$ is the electric dipole ($E1$) operator.
remarks:

(i) A fully consistent calculation of the decay probability A could only be obtained once Quantum Electrodynamics was formulated and evaluated to first order.

(ii) The process of induced or stimulated emission, predicted by A. Einstein in 1917 on the basis of Planck's formula, lead to the invention of the Maser by Ch. H. Townes in 1954 (Nobel Prize 1964) and the Laser by Ch. H. Townes and A.L. Schawlow in 1956 (Nobel Prize 1981).

- radioactive decay law:

Without radiation field, i.e. $u(\omega_{nm}) = 0$, the rate of change of $N_m(t)$ is given by

$$dN_m(t) \;=\; -A\,N_m(t)\,dt \quad ,$$

where $N_n(t) + N_m(t) = \text{const}$, as well. Here

$N_m(t)$ is the number of atoms in the excited state E_m, and

$N_n(t)$ is the number of atoms in the ground state E_n.

The solution of this differential equation is

$$N_m(t) = N_m(0)\,\mathrm{e}^{-At} = N_m(0)\,\mathrm{e}^{-t/\tau}$$

with $\tau = 1/A$ being the lifetime of the excited state $|u_m\rangle$.

17.2 Example: lifetime of an excited state

initial state: $u_{210}(\vec{r})$ $n=2$, $l=1$, $m=0$

final state: $u_{100}(\vec{r})$ $n=1$, $l=0$, $m=0$

$$u_{n\,l\,m}(r,\theta,\varphi) = -\sqrt{\frac{(n-l-1)!}{2n\,[(n+l)!]^3}\left(\frac{2Z}{an}\right)^{2l+3}}\; r^l \exp\left(-\frac{Zr}{an}\right) L_{n+l}^{2l+1}\left(\frac{2Zr}{an}\right) Y_{lm}(\theta,\varphi)$$

- initial state: $n=2,\ l=1,\ m=0\ ,\ Z=1$

$$u_{2,1,0}(r,\theta,\varphi) = -\sqrt{\frac{1}{4\cdot(3!)^3}\left(\frac{2}{2a}\right)^5}\; r\exp\left(-\frac{r}{2a}\right) \underbrace{L_3^3\left(\frac{2r}{an}\right)}_{-6} \underbrace{Y_{1,0}(\theta,\varphi)}_{\frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta}$$

$$u_{2,1,0}(r,\theta,\varphi) \;=\; -\sqrt{\frac{1}{4\cdot 27\cdot 8}\left(\frac{1}{a}\right)^5}\; r\exp\left(-\frac{r}{2a}\right)(-6)\sqrt{\frac{3}{4\pi}}\cos\theta$$

$$\;=\; -\frac{1}{4\cdot 3\sqrt{3}\sqrt{2}}\,a^{-5/2}\,r\exp\left(-\frac{r}{2a}\right)(-6)\sqrt{\frac{3}{4\pi}}\cos\theta$$

$$u_{2,1,0}(r,\theta,\varphi) \;=\; \frac{1}{4}\,(2\pi)^{-\frac{1}{2}}\,a^{-\frac{5}{2}}\,r\exp\left(-\frac{r}{2a}\right)\cos\theta$$

- final state: $n=1,\ l=0,\ m=0\ ,\ Z=1$

$$u_{1,0,0}(r,\theta,\varphi) = -\sqrt{\frac{1}{2}\left(\frac{2}{a}\right)^3}\exp\left(-\frac{r}{a}\right)\underbrace{L_1^1\left(\frac{2r}{a}\right)}_{-1}\underbrace{Y_{00}(\theta,\varphi)}_{\frac{1}{\sqrt{4\pi}}}$$

$$u_{1,0,0}(r,\theta,\varphi) \;=\; \frac{1}{\sqrt{\pi a^3}}\exp\left(-\frac{r}{a}\right)$$

- matrix element:

$$\vec{r}_{fi} = \int u_{1,0,0}(\vec{r})\, \vec{r}\, u_{2,1,0}(\vec{r})\, d^3r$$

$$\begin{cases} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{cases} \qquad d^3r = r^2\, dr\, \sin\theta\, d\theta\, d\varphi$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_{fi} = \int_0^\infty r^2\, dr \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\varphi\, \frac{e^{-(r/a)}}{a\sqrt{\pi a}} \begin{pmatrix} r\sin\theta\cos\varphi \\ r\sin\theta\sin\varphi \\ r\cos\theta \end{pmatrix} \frac{a^{-5/2}}{4\sqrt{2\pi}}\, r\, e^{-\frac{r}{2a}} \cos\theta$$

but $$\int_0^{2\pi} \sin\varphi\, d\varphi = \int_0^{2\pi} \cos\varphi\, d\varphi = 0$$

$\Rightarrow \quad x_{fi} = y_{fi} = 0 \quad \Rightarrow$ no radiation emitted in the x, y plane

$$z_{fi} = \int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\varphi\, \frac{e^{-\frac{3r}{2a}}}{\pi\, a^4\, 4\sqrt{2}}\, r^4 \sin\theta\cos^2\theta$$

$$= \frac{2\pi}{4\pi\, a^2\sqrt{2}} \int_0^\infty dr\, e^{-\frac{3r}{2a}}\, r^4 \int_0^\pi d\theta\, \sin\theta\cos^2\theta$$

$$= \frac{1}{2\sqrt{2}\, a^4} \left(\frac{2a}{3}\right)^5 \underbrace{\int_0^\infty du\, e^{-u}\, u^4}_{\Gamma(5) = 4!} \underbrace{\int_{-1}^1 \cos^2\theta\, d(\cos\theta)}_{\frac{2}{3}}$$

$$A = \frac{1}{\tau} = \frac{4e^2\omega^3}{3\hbar c^3} |z_{fi}|^2$$

$$= \frac{4e^2\omega^3}{3\hbar c^3} \left[\frac{2^5\, a^5\, 4!\, 2}{2a^4\sqrt{2}\, 3^5 \cdot 3}\right]^2 = \frac{4\, e^2\, \omega^3}{3\hbar\, c^3} \left(\frac{2^8 a}{\sqrt{2}\, 3^5}\right)^2$$

$$= \frac{2^{17}\, a^2\, e^2\, \omega^3}{3^{11}\, \hbar\, c^3}$$

$$E_n = -\frac{\mu\, Z^2\, e^4}{2\hbar^2} \frac{1}{n^2} \quad \text{with} \quad \mu = \frac{m_e\, m_A}{m_e + m_A} \quad \text{reduced mass}$$

$$\boxed{\omega = \frac{E_2 - E_1}{\hbar} = \frac{\mu\, Z^2\, e^4}{2\hbar^3} \left(1 - \frac{1}{4}\right) = \frac{3\,\mu\, Z^2\, e^4}{8\hbar^3}}$$

$$a = \frac{\hbar^2}{\mu\, e^2\, Z} = \frac{\hbar}{\mu c} \frac{1}{Z} \frac{\hbar c}{e^2}$$

radiative decay time

$$\boxed{\tau(2p \to 1s) = \left(\frac{3}{2}\right)^8 \left(\frac{\hbar c}{e^2}\right)^5 \frac{1}{Z^4} \frac{\hbar}{\mu c^2}}$$

$$\frac{e^2}{\hbar c} = \frac{1}{137.035999(50)} \qquad \mu = \frac{m_e\, m_A}{m_e + m_A}$$

examples:

$Z = 1$ $\quad e^- + \mathrm{H}$ $\quad \boxed{\tau = 1.60 \times 10^{-9}\mathrm{sec}}$

$Z = 80$ $\quad e^- + {}^{200}\mathrm{Hg}$ $\quad \boxed{\tau = 3.90 \times 10^{-17}\mathrm{sec}}$

$Z = 2$ $\quad \mu^- + {}^4\mathrm{He}$ $\quad \boxed{\tau = 0.494 \times 10^{-12}\mathrm{sec}}$

where we have used

$$m_{Hg}\, c^2 = \overbrace{(200}^{\text{nucleon number}} \times \overbrace{931.478}^{\text{amu}} - \overbrace{29.5)}^{\text{mass excess}} \text{ MeV} \quad \text{mercury}$$

$m_\mu\, c^2 = 105.658389$ MeV muon

$m_\alpha\, c^2 = 3728.3368$ MeV helium

$\mu = \dfrac{m_\mu\, m_\alpha}{m_\mu + m_\alpha} = 102.7476$ MeV 3% difference!

reduced mass

18 Ionization of Atomic Hydrogen

- A hydrogen atom is subject to a spatially homogeneous electric field in the z-direction that is periodic in time and given by the Hamilton operator interaction

$$H_1(t) = e E_z z \cos\omega t = \frac{1}{2} e E_z z \left(e^{i\omega t} + e^{-i\omega t}\right) .$$

de-excitation ↑ ↑ excitation

For an atom in its groundstate, only the second term leads to a non-zero contribution. The probability per unit time for the emission of the electron is given, in first order perturbation theory, by Fermi's golden rule, i.e.

$$d^2p(i \to f) = \frac{2\pi}{\hbar} \left| \langle u_f | \frac{1}{2} e_z E z u_i \rangle \right|^2 \delta(\hbar\omega - \hbar\omega_{fi})\, d^2n \qquad \text{eq.(1)}$$

with $\hbar\omega_{fi} \underset{\text{def}}{=} E_f - E_i$.

Here we neglect the recoil energy of the proton and the spin of the electron, which are both good approximations.

The initial electron wave function and energy are given by

$$u_f(\vec{r}) = u_{1,0,0}(r,\theta,\varphi) = \frac{1}{\sqrt{\pi a^3}} \exp\left(-\frac{r}{a}\right)$$

$$E_i = E_1 = -\frac{1}{2}\frac{me^4}{\hbar^2} ,$$

respectively, while the final electron wave function and energy are

$$u_f(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k}\cdot\vec{r}}$$

$$E_f = \frac{\hbar^2 k^2}{2m} .$$

We neglect the Coulomb interaction in the final state which is also a good approximation at high energies. The final state wavefunction is normalized using periodic boundary conditions in a large volume V. The differential

$$d^2n = \frac{V}{(2\pi)^3} k^2\, dk\, d\Omega$$

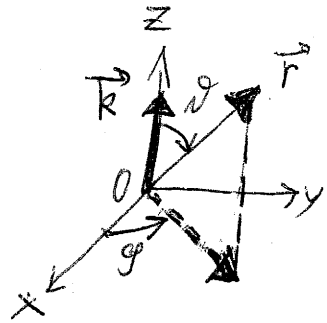
in eq.(1) describes the number of final states available to the electron with wave vectors of absolute values between k and $k + dk$ pointing into the solid angle $d\Omega$.

We now calculate the matrix element

$$\langle u_f \left| \frac{1}{2} eE_z z\, u_i \right\rangle \underset{\text{def}}{=} \frac{eE_z}{2\sqrt{\pi a^3 V}} \underbrace{\int e^{-i\vec{k}\cdot\vec{r}} z e^{-\frac{r}{a}} d^3r}_{J(\alpha,\vec{k})} \qquad \text{eq.(2)}$$

corollary: $$I(\alpha,\vec{k}) \underset{\text{def}}{=} \int e^{i\vec{k}\cdot\vec{r}} \frac{e^{-\alpha r}}{r} d^3r = \frac{4\pi}{k^2+\alpha^2}$$

proof: Choose $\vec{k}$ and the angles θ and φ as shown in this figure.



$$I(\alpha,\vec{k}) = \int_0^\infty r\, e^{-\alpha r} dr \int_0^\pi \sin\vartheta\, d\vartheta\, e^{ikr\cos\vartheta} \int_0^{2\pi} d\varphi$$

$$= -2\pi \int_0^\infty r\, e^{-\alpha r} dr \int_1^{-1} d\cos\vartheta\, e^{ikr\cos\vartheta}$$

$$= 2\pi \int_0^\infty r\, e^{-\alpha r} dr \int_{-1}^{1} dx\, e^{ikrx}$$

$$= \frac{2\pi}{ik} \int_0^\infty \frac{r}{r} e^{-\alpha r} \left(e^{ikr} - e^{-ikr}\right) dr$$

$$= \frac{2\pi}{ik} \int_0^\infty \left[e^{(ik-\alpha)r} - e^{(-ik-\alpha)r}\right] dr$$

$$= \frac{2\pi}{ik} \left[\frac{-1}{ik-\alpha} + \frac{1}{-ik-\alpha}\right]$$

$$= -\frac{2\pi}{ik} \frac{(ik+\alpha)+(ik-\alpha)}{(ik-\alpha)(ik+\alpha)} = \frac{4\pi}{k^2+\alpha^2} \qquad \text{q.e.d.}$$

theorem: $$J(\alpha,\vec{k}) \underset{\text{def}}{=} \int z e^{-i\vec{k}\cdot\vec{r}}\, e^{-\alpha r}\, d^3r = \frac{32\pi i k_z\, \alpha}{(k^2+\alpha^2)^3}$$

proof: $$I(\alpha,\vec{k}) = \frac{4\pi}{k^2+\alpha^2} = \int e^{i\vec{k}\vec{r}}\, \frac{e^{-\alpha r}}{r}\, d^3r$$

$$\frac{\partial}{\partial k_z} I(\alpha,\vec{k}) = \frac{\partial}{\partial k_z}\frac{4\pi}{k^2+\alpha^2} = -\frac{4\pi}{(k^2+\alpha^2)^2}\, 2k_z$$

$$= \frac{\partial}{\partial k_z}\int e^{i\vec{k}\cdot\vec{r}}\, \frac{e^{-\alpha r}}{r}\, d^3r = \int iz\, e^{i\vec{k}\cdot\vec{r}}\, \frac{e^{-\alpha r}}{r}\, d^3r \ .$$

Thus we arrive at

$$\frac{8\pi i k_z}{(k^2+\alpha^2)^2} = \int z\, e^{i\vec{k}\cdot\vec{r}}\, \frac{e^{-\alpha r}}{r}\, d^3r \ .$$

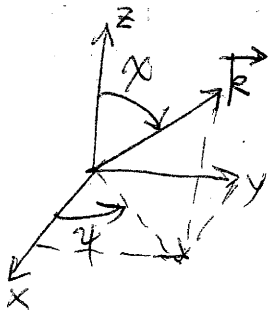
Taking the partial derivative with respect to α, we obtain

$$\frac{\partial}{\partial\alpha}\frac{8\pi i k_z}{(k^2+\alpha^2)^2} = \int z\, e^{i\vec{k}\cdot\vec{r}} \left(-\frac{r}{r}\right) e^{-\alpha r}\, d^3r$$

$$= -\int z\, e^{i\vec{k}\cdot\vec{r}}\, e^{-\alpha r}\, d^3r$$

$$= \frac{-2\cdot 8\pi i k_z}{(k^2+\alpha^2)^3}\, 2z$$

$$= -\frac{32\pi i k_z \alpha}{(k^2+\alpha^2)^3} \ . \qquad \text{q.e.d.}$$



Defining the electron emission angle χ through $k_z = k\cos\chi$, and introducing Bohr's radius of the hydrogen atom, $a = \frac{1}{\alpha}$, we arrive at

$$J(\alpha,\vec{k}) = \int z\, e^{-i\vec{k}\cdot\vec{r}}\, e^{-\frac{r}{a}}\, d^3r = \frac{32\pi i k\cos\chi}{a\left(k^2+\frac{1}{a^3}\right)^3} = \frac{32\pi i a^5 k\cos\chi}{(a^2k^2+1)^3} \ .$$

The ionization probability per unit time is in first order perturbation theory given by

$$d^2p(i\to f) = \frac{2\pi}{\hbar}\left|\langle u_f\left|\frac{1}{2}\, e E_z\, u_i\right\rangle\right|^2 \delta(\hbar\omega-\hbar\omega_{fi})\ \frac{V}{(2\pi)^3\, k^2\, dk\, d\Omega} \ .$$

Inserting the matrix element eq.(2), we obtain

$$d^2p(i\to f) = \frac{2\pi}{\hbar}\,\frac{e^2E_z^2}{\pi a^3V^4}\cos^2\chi\,\frac{2^{10}\pi^2a^{10}k^2}{(1+a^2k^2)^6}\,\delta(\hbar\omega-\hbar\omega_{fi})\frac{V}{(2\pi)^3}\,k^2\,dk\,d\Omega$$

$$\text{or}\quad d^2p(i\to f) = \frac{64\,e^2\,E_z^2\,a^7\,k^2}{\hbar\pi(1+a^2k^2)^6}\cos^2\chi\ \ k^2\,dk\,d\Omega\ \ \delta(\hbar\omega-\hbar\omega_{fi}) \ .$$

With
$$\hbar\omega_{fi} = E_f - E_i \ , \quad E_f = \frac{\hbar^2k^2}{2m} \ , \quad dE_f = \frac{\hbar^2}{2m}\,2k\,dk$$

and
$$k\,dk = \frac{m}{\hbar^2}\,d\,E_f$$

we arrive at

$$dp(i\to f) = \int \frac{64\,e^2\,E_z^2\,a^7\,k^2}{\hbar\pi(1+a^2k^2)^6}\,\delta(\hbar\omega-E_f+E_i)\,\cos^2\chi\ \ \frac{mk}{\hbar^2}\,dE_f\,d\Omega$$

$$\text{or}\qquad dp(i\to f) = \frac{64\,e^2\,E_z^2\,a^7\,k^3\,m}{\hbar^3\pi(1+a^2k^2)^6}\,\cos^2\chi\ \ d\Omega$$

with $\quad E_f - E_i = \hbar\omega \quad$ describing the conservation of energy.

The differential probability for ionization per unit time is

$$\frac{dp(i\to f)}{d\Omega} = \frac{64\,e^2\,E_z^2\,a^7\,k^3\,m}{\hbar^3\pi(1+a^2k^2)^6}\,\cos^2\chi$$

with
$$k = \frac{\sqrt{2m(E_1+\hbar\omega)}}{\hbar^2} = \sqrt{\frac{2m}{\hbar}(\omega-\omega_0)}$$

and
$$E_1 = -\frac{me^4}{2\hbar^2} = -\hbar\omega_0 \ , \quad a = \frac{\hbar^2}{me^2} \ .$$

The circular frequency at threshold is $\omega_0 = \frac{me^4}{2\hbar^3}$. After a little algebra, i.e.

$$1+k^2\,a^2 = 1+\frac{2m}{\hbar}\,(\omega-\omega_0)\,\frac{\hbar^4}{m^2\,e^4} = 1+\frac{\omega-\omega_0}{\omega_0}$$

$$= 1+\frac{\omega}{\omega_0}-1 = \frac{\omega}{\omega_0}$$

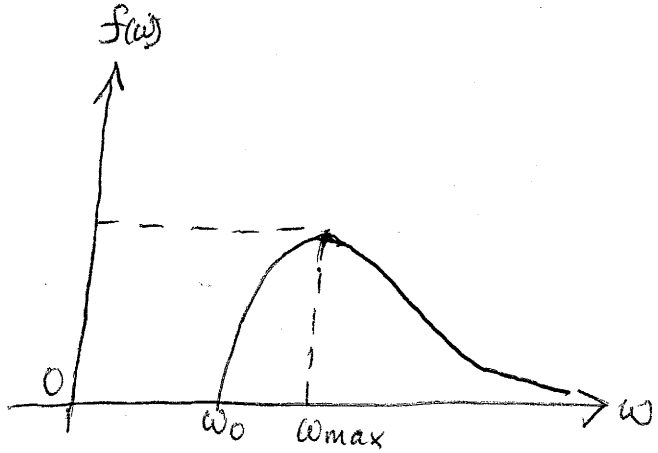
$$\frac{\omega}{\omega_0} - 1 = k^2\, a^2 \quad \Rightarrow \quad k^2 = a^{-2}\left(\frac{\omega}{\omega_0} - 1\right)$$

$$k^3 = a^{-3}\left(\frac{\omega}{\omega_0} - 1\right)^{3/2} ,$$

the differential probability for ionization per unit time becomes

$$\frac{dp(i \to f)}{d\Omega} = \frac{64\, e^2\, E_z^2\, a^4\, m}{\hbar^3 \pi} \underbrace{\left(\frac{\omega}{\omega_0} - 1\right)^{3/2} \left(\frac{\omega_0}{\omega}\right)^6}_{f(\omega)} \cos^2 \chi ,$$

where $f(\omega)$ is a dimensionless spectral ionization probability.



We would like to find the maximum of the function $f(\omega)$, i.e.

$$f(\omega) = \left(\frac{\omega_0}{\omega}\right)^6 \left(\frac{\omega}{\omega_0} - 1\right)^{3/2}$$

$$f'(\omega) = 6\left(\frac{\omega_0}{\omega}\right)^5 \left(-\frac{\omega_0}{\omega^2}\right) \left(\frac{\omega}{\omega_0} - 1\right)^{3/2} + \left(\frac{\omega_0}{\omega}\right)^6 \frac{3}{2} \left(\frac{\omega}{\omega_0} - 1\right)^{\frac{1}{2}} \left(\frac{1}{\omega_0}\right)$$

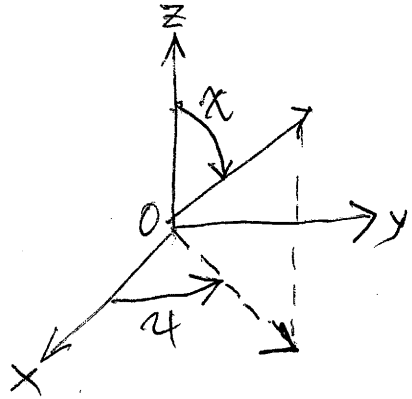
$$f'(\omega) = \left(\frac{\omega_0}{\omega}\right)^5 \left(\frac{\omega}{\omega_0} - 1\right)^{\frac{1}{2}} \left[-\frac{9}{2\omega} + 6\,\frac{\omega_0}{\omega^2}\right] = 0$$

$$\Rightarrow \qquad 6\frac{\omega_0}{\omega_{max}^2} = \frac{g}{2\omega_{max}}$$

$$\omega_{max} = \frac{4}{3}\,\omega_0 .$$

Thus the ionization process is most effective for $\omega_{max} = \frac{4}{3}\,\omega_0$ corresponding to $E_{max} = \frac{4}{3}\,13.6\,\text{eV}$. The electron is preferentially emitted in the direction of the electric field, i.e.

$$\frac{dp(i \to f)}{d\Omega} = \frac{64\, e^2\, E_z^2\, a^4}{\hbar^3 \pi} \left(\frac{\omega}{\omega_0} - 1\right)^{3/2} \left(\frac{\omega_0}{\omega}\right)^6 \cos^2 \chi .$$



$$d\Omega = \sin\chi\, d\chi\, d\psi$$

$$\int_0^{2\pi} d\psi \int_0^{\pi} \sin\chi\, \cos^2\chi\, d\chi = \frac{4\pi}{3}$$

Integrating over the electron emission angles, the spectral ionization probability per unit time is

$$p(i \to f) = \frac{256\, a^4\, e^2\, E_z^2}{3\hbar^3} \left(\frac{\omega_0}{\omega}\right)^6 \left(\frac{\omega}{\omega_0} - 1\right)^{3/2} ,$$

while the total ionization probability per unit time is

$$P(i \to f) = \int_{\omega_0}^{\infty} p(i \to f)\, d\omega = \frac{256\, a^4 e^2 E_z^2}{3\hbar^3} \int_{\omega_0}^{\infty} \left(\frac{\omega_0}{\omega}\right)^6 \left(\frac{\omega}{\omega_0} - 1\right)^{3/2} .$$

<u>remark:</u>

Some 400 kyr after the Big Bang, the electrons and protons combined for the first time in this universe to form hydrogen atoms. However, 600 Myr later, perhaps due to the early quasars, the hydrogen atoms were reionized through strong ionizing radiation. This allowed the mixture of hydrogen atoms, protons and electrons, to form stable hydrogen molecule ions H_2^+, which eventually became neutral H_2 molecules through the capture of additional electrons. Because of the low-lying rotational bands of the H_2 molecules at infrared wavelengths (see tut 31), the massive H_2 clouds could radiate off their gravitational collapse energy efficiently, eventually forming the early stars of the galaxies.

19 Relativistic wave equations

19.1 Recipe for generating wave equations

(i) Take a Hamiltonian describing a free massive particle

$$H = f(\vec{p}, m) \qquad \text{energy-momentum-mass relationship.}$$

(ii) Introduce the electromagnetic interaction via minimal substitution

$$\begin{cases} \vec{p} \to \vec{p} - \dfrac{e}{c}\vec{A} \\ H \to H - e\phi \end{cases} \Rightarrow \quad H - e\phi = f\left(\vec{p} - \frac{e}{c}\vec{A}, m\right) \ .$$

(iii) Make the transition from classical to quantum mechanics, e.g. in coordinate space, by replacing

$$\begin{cases} \vec{p} \to -i\hbar\vec{\nabla} \\ H \to i\hbar\dfrac{\partial}{\partial t} \end{cases} \Rightarrow \quad \left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\psi = f\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}, m\right)\psi$$

and apply this operator equation to a space-time dependent wave function ψ.

19.2 Examples

- Schrödinger's equation:

  E. Schrödinger (1926), Nobel Prize 1933.

$$H = \frac{p^2}{2m}$$

$$\Rightarrow \quad \left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\psi = \frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2\psi$$

  describes a charged, nonrelativistic and spinless particle moving in an electromagnetic field.

- Pauli equation:

  W. Pauli (1926), Nobel Prize 1945.

$$H = \frac{1}{2m}\,\vec{\sigma}\cdot\vec{p}\,\vec{\sigma}\cdot\vec{p}$$

$$\Rightarrow \quad \left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\psi = \frac{1}{2m}\left[\vec{\sigma}\cdot\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)\right]\left[\vec{\sigma}\cdot\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)\right]\psi$$

  describes a charged, nonrelativistic particle with spin $\frac{1}{2}$ and $g_S = 2$ moving in an electromagnetic field.

- Relativistic Schrödinger equation:

  E. Schrödinger, *Ann. Physik* **81** (1926) 109

$$H = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}$$

$$\Rightarrow \quad \left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\psi = \sqrt{\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2 c^2 + m^2c^4}\,\psi$$

$$= mc^2\sqrt{1 + \frac{\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2}{m^2c^2}}\,\psi$$

  describes a charged relativistic massive spinless particle moving in an electromagnetic field. Expanding the square root yields

$$\left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\psi = mc^2\left(1 + \tfrac{1}{2}\frac{\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2}{m^2c^2} - \frac{1}{8}\frac{\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^4}{m^4c^4} + \cdots\right)\psi \ ,$$

  where we have used

$$(1+x)^{\frac{1}{2}} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \cdots x \ll 1 \ .$$

  The rest energy term mc^2 can be "gauged away", introducing the wavefunction $\widetilde{\psi}$

$$\psi \underset{\text{def}}{=} \exp\left(-i\frac{mc^2}{\hbar}t\right)\widetilde{\psi} \ .$$

$$\frac{\partial\psi}{\partial t} = \exp\left(-\frac{imc^2}{\hbar}t\right)\frac{\partial\widetilde{\psi}}{\partial t} - \exp\left(-\frac{mc^2}{\hbar}t\right)\left(i\frac{mc^2}{\hbar}\right)\widetilde{\psi}$$

$$\text{l.h.s.} = i\hbar\frac{\partial}{\partial t}\psi - e\phi\psi$$

$$= \exp\left(-\frac{imc^2}{\hbar}t\right)\left(i\hbar\frac{\partial}{\partial t} + mc^2 - e\phi\right)\widetilde{\psi}$$

$$\text{r.h.s.} = \exp\left(-\frac{imc^2}{\hbar}t\right)\left[mc^2 + \frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2 - \frac{1}{8m^3c^2}\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^4 + \cdots\right]\widetilde{\psi}$$

$$\text{l.h.s.} = \text{r.h.s.}$$

$$\Rightarrow \quad \left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\widetilde{\psi} = \underbrace{\frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2\widetilde{\psi}}_{\text{nonrelativistic term}} \underbrace{-\frac{1}{8m^3c^2}\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^4\widetilde{\psi} + \cdots}_{\text{first relativistic correction term}}$$

relativistic Schrödinger equation

remarks:

(i) We have expanded the square root in powers of the operator $\left(-i\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2$. This series describes the relativistic corrections to the nonrelativistic Schrödinger equation. Of course, this differential equation is not particularly pretty, as time and space derivatives are treated very differently.

(ii) The matrix elements of the higher powers of the operator $\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2$ become increasingly singular and difficult to evaluate in perturbation theory.

- Klein-Gordon equation

E. Schrödinger, *Ann. Physik* **81** (1926) 109
W. Gordon, *Z. Physik* **40** (1926) 117
O. Klein, *Z. Physik* **41** (1927) 407.

A way of bypassing the "relativistic square-root problem" is to apply our procedure for generating wave equations to H^2 instead of H. Indeed,

$$H^2 = \vec{p}^{\,2}c^2 + m^2c^4$$

simply yields

$$\left(i\hbar\frac{\partial}{\partial t} - e\phi\right)^2\varphi = \underbrace{\left[\left(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2 c^2 + m^2c^4\right]\varphi}_{\text{Klein-Gordon equation}}.$$

The Klein-Gordon field describes a massive, charged and spinless, relativistic particle field.

remarks:

(i) Here space and time derivatives are treated equally.

(ii) However, $\varphi(\vec{r},t)$ cannot be interpreted as a quantum mechanical wavefunction or state vector Indeed, one would have to specify $\varphi(\vec{r},t)\big|_{t=0}$ as well as $\frac{\partial}{\partial t}\varphi(\vec{r},t)\big|_{t=0}$, in order to obtain the time-evolution of $\varphi(\vec{r},t)$. This is contrary to the well-established principles of QM, which are based on the link with CM.

(iii) However, $\varphi(\vec{r},t)$ describes a massive charged, relativistic spin-0 field, the would-be wavefunction $\varphi(\vec{r},t)$ fulfils the Klein-Gordon equation $\left(\Box - \left(\frac{mc}{\hbar}\right)^2\right)\varphi(\vec{r},t) = 0$ in the noninteracting case. This is similar to the scalar-vector potential $A_\mu(\vec{r},t)$ that describes the massless spin-1 photon field, and obeys $\Box A_\mu = 0$ in the Lorentz or Feynman gauge $\sum_\mu \frac{\partial}{\partial x_\mu}A_\mu(\vec{r},t) = 0$.

- Klein-Gordon current conservation

We would like to obtain an equation describing the current conservation of the Klein-Gordon field, for $\vec{A} = \phi = 0$ for simplicity.

$$H^2\varphi = \left(\vec{p}^{\,2}c^2 + m^2c^4\right)\varphi$$

$$-\hbar^2\frac{\partial^2}{\partial t^2}\varphi = \left(-\hbar^2\vec{\nabla}^{\,2}c^2 + m^2c^4\right)\varphi \;\Big|\; \text{divide by } \hbar^2c^2$$

$$-\frac{1}{c^2}\frac{\partial}{\partial t^2}\varphi = \left(-\vec{\nabla}^{\,2} + \frac{m^2c^2}{\hbar^2}\right)\varphi$$

$$\left[\left(\vec{\nabla}^{\,2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\varphi - \left(\frac{mc}{\hbar}\right)^2\right]\varphi = 0$$

$$\Box \underset{\text{def}}{=} \Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} = \vec{\nabla}\cdot\vec{\nabla} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$$

d'Alembert operator

multiply by

$\varphi^*\Big|$ $\left[\Box - \left(\frac{mc}{\hbar}\right)^2\right]\varphi = 0$ original Klein-Gordon equation

$\varphi\Big|$ $\left[\Box - \left(\frac{mc}{\hbar}\right)^2\right]\varphi^* = 0$ complex conjugated Klein-Gordon equation

subtract

$$\Rightarrow \quad \varphi^*\,\Box\,\varphi - \varphi\,\Box\,\varphi^* = 0$$

$$\varphi^* \left(\vec{\nabla} \cdot \vec{\nabla} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \varphi - \varphi \left(\vec{\nabla} \cdot \vec{\nabla} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \varphi^* = 0$$

$$-\frac{1}{c^2} \left(\varphi^* \frac{\partial^2}{\partial t^2} \varphi - \varphi \frac{\partial^2}{\partial t^2} \varphi^* \right) + \varphi^* \vec{\nabla} \cdot \vec{\nabla} \varphi - \varphi \vec{\nabla} \cdot \vec{\nabla} \varphi^* = 0$$

$$-\frac{i\hbar}{2m} \left| -\frac{1}{c^2} \frac{\partial}{\partial t} \left(\varphi^* \frac{\partial}{\partial t} \varphi - \varphi \frac{\partial}{\partial t} \varphi^* \right) + \vec{\nabla} \cdot \left(\varphi^* \vec{\nabla} \varphi - \varphi \vec{\nabla} \varphi^* \right) = 0 \right.$$

$$\frac{\partial}{\partial t} \underbrace{\left[\frac{i\hbar}{2mc^2} \left(\varphi^* \frac{\partial}{\partial t} \varphi - \varphi \frac{\partial}{\partial t} \varphi^* \right) \right]}_{\rho} + \vec{\nabla} \cdot \underbrace{\frac{\hbar}{2im} \left[\varphi^* \vec{\nabla} \varphi - \varphi \vec{\nabla} \varphi^* \right]}_{\vec{J}} = 0$$

ρ: probability density differs from Schrödinger's expression $\rho = \varphi^* \varphi$

$\vec{J}$: probability current density exactly the same expression as for the the Schrödinger equation

$$\Rightarrow \quad \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0$$

continuity equation

remarks:

(i) Both, ρ and $\vec{J}$, are real, as in the case of the Schrödinger equation. However, ρ is not positive definite, and thus ρ cannot be interpreted as a probability density! The root of this problem is again that the Klein-Gordon equation contains a second derivative in time.

(ii) The Klein-Gordon equation allows for both positive and negative energy solutions

$$\varphi = A e^{\frac{i}{\hbar}(\vec{p} \cdot \vec{r} - Et)}$$

$$\text{with} \quad E^2 = \vec{p}^{\,2} c^2 + m^2 c^4$$

$$\Rightarrow \quad E = \pm \sqrt{\vec{p}^{\,2} c^2 + m^2 c^4} \; .$$

Thus there is no way that we can restrict ourselves to positive energy solutions only, as these do not form a complete set, and interactions will invariably cause coupling to these negative energy solutions. However, this problem appears in all relativistic equations.

(iii) In order to have a non-zero current, φ must be complex, as in the case of the Schrödinger equation.

20 Dirac equation

20.1 Dirac's brilliant idea

P.A.M. Dirac, *Proc. Roy. Soc.* **A117** (1928) 610; **A118** (1928) 351 Nobel Prize 1933

Dirac's starting point in 1927 was Schrödinger's noninteracting relativistic wave equation

$$i\hbar \frac{\partial}{\partial t} \psi = \sqrt{\left(-i\hbar \vec{\nabla} \right)^2 c^2 + m^2 c^4} \, \psi \; .$$

Initially, Dirac believed, erroneously, that by choosing the positive root, he could eliminate the unwanted negative energy solutions. The challenge he faced in 1927 was that he wanted a differential equation of first order in time and space that describes a spin $\frac{1}{2}$ particle. Thus, he linearized the square root by brute force, using 4 "coefficients" α_x, α_y, α_z and β, i.e.

$$\sqrt{\left(-i\hbar \vec{\nabla} \right)^2 c^2 + m^2 c^4} \, I \psi = \left[\vec{\alpha} \cdot \left(-i\hbar \vec{\nabla} \right) c + \beta mc^2 \right] \psi \; .$$

Of course, Dirac knew about the successful incorporation of spin $S = \frac{1}{2}$ with $g_S = 2$ in the Pauli equation

$$i\hbar I_2 \frac{\partial \psi}{\partial t} = \frac{1}{2m} \left[\vec{\sigma} \cdot \left(-i\hbar \vec{\nabla} \right) \right]^2 \psi \; .$$

Thus it was probably not too far-fetched to postulate that the coefficients $\vec{\alpha} = \{\alpha_x, \alpha_y, \alpha_z\}$ and β could be matrices.

20.2 Properties of the Dirac matrices

The matrices $\vec{\alpha}$ and β are defined through the "linearization condition" of the free Dirac Hamilton operator, i.e.

$$H = \sqrt{\vec{p}^{\,2} c^2 + m^2 c^4} \, I = c \vec{\alpha} \cdot \vec{p} + \beta mc^2 \; ,$$

with $\vec{p} = -i\hbar \vec{\nabla}$.

we note:

(i) α_x, α_y, α_z and β must be Hermitean matrices, as well, if we want H to be Hermitean.

(ii) we must have

$$H^2 = \left(\vec{p}^{\,2} c^2 + m^2 c^4 \right) I = \left(c \vec{\alpha} \cdot \vec{p} + \beta mc^2 \right)^2$$

or in components

$$\begin{aligned} H^2 &= p_x^2\,c^2\,I + p_y^2\,c^2\,I + p_z^2\,c^2\,I + m^2\,c^4\,I \\ &= \left(c\,\alpha_x\,p_x + c\,\alpha_y\,p_y + c\,\alpha_z\,p_z + \beta\,mc^2\right)^2 \\ &= c^2\,\alpha_x^2\,p_x^2 + c^2\,\alpha_y^2\,p_y^2 + c^2\,\alpha_z^2\,p_z^2 + \beta^2\,m^2\,c^4 \quad \text{square terms} \end{aligned}$$

$$\left.\begin{aligned} &+\; c^2\,(\alpha_x\,\alpha_y + \alpha_y\,\alpha_x)\;p_x\,p_y + c^2\;(\alpha_x\,\alpha_z + \alpha_z\alpha_x)\;p_x\,p_z \\ &+\; c^2\,(\alpha_y\,\alpha_z + \alpha_z\,\alpha_y)\;p_y\,p_z + mc^3\,(\alpha_z\,\beta + \beta\,\alpha_z)\;p_x \\ &+\; mc^3\;(\alpha_y\,\beta + \beta\,\alpha_y)\;p_y + mc^3\;(\alpha_z\,\beta + \beta\,\alpha_z)\;p_z \quad . \end{aligned}\right\} \text{mixed terms}$$

- Thus, the Dirac matrices must fulfil a Clifford algebra, i.e.

$$\begin{aligned} \alpha_x^2 &= \alpha_y^2 = \alpha_z^2 = \beta^2 = I \\ \{\alpha_x, \alpha_y\} &= \{\alpha_y, \alpha_z\} = \{\alpha_z, \alpha_x\} = O \\ \{\beta, \alpha_x\} &= \{\beta, \alpha_y\} = \{\beta, \alpha_z\} = O \;. \end{aligned}$$

The matrices chosen by P.A.M. Dirac,

$$\beta = \begin{pmatrix} I_2 & O_2 \\ O_2 & -I_2 \end{pmatrix} \qquad \vec{\alpha} = \begin{pmatrix} O_2 & \vec{\sigma} \\ \vec{\sigma} & O_2 \end{pmatrix} ,$$

indeed, fulfil this algebra. Introducing $\gamma_k = i\beta\alpha_k$ $(k = 1,2,3)$ and $\gamma_4 = \beta$, one can see that these γ_μ are Hermitean and fulfil the Clifford algebra, as well, i.e.

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} \qquad (\mu, \nu = 1,2,3,4) ,$$

This enables us to write the noninteracting Dirac equation in covariant form as

$$\left[\sum_{\mu=1}^{4} \gamma_\mu \frac{\partial}{\partial x_\mu} + \frac{mc}{\hbar}\right] \psi = 0 \;.$$

<u>remarks:</u>

(i) In $H\psi$, the 4×4 matrices $\alpha_x, \alpha_y, \alpha_z, \beta$ act on a complex 4-component Dirac spinor $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$, which describes two spin $\frac{1}{2}$ particles: a <u>particle</u> and an <u>antiparticle</u>.

(ii) The Dirac spinor index has nothing to do with the Lorentz vector index. It is a mere accident that both have four components in $d = 3 + 1$ dimensions.

(iii) By construction, the Dirac spinor ψ is a solution of the free Dirac equation

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi = \left(c\,\vec{\alpha} \cdot \left(-i\hbar\,\vec{\nabla}\right) + \beta\,mc^2\right) \psi \;.$$

But each spinor component also obeys the free Klein-Gordon equation

$$\left(i\hbar \frac{\partial}{\partial t}\right)^2 \psi = H^2\,\psi = \left(c\,\vec{\alpha} \cdot \left(-i\hbar\,\vec{\nabla}\right) + \beta\,mc^2\right)^2 \psi$$

or

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = H^2\,\psi = \left(-\hbar^2\,c^2\,\vec{\nabla}^{\,2} + m^2\,c^4\right) \psi$$

or

$$\left(\Box - \frac{m^2c^2}{\hbar^2}\right) \psi = \left(\sum_{\mu=1}^{4} \left(\frac{\partial}{\partial x_\mu}\right)^2 - \frac{m^2c^2}{\hbar^2}\right) \psi = 0 \;.$$

20.3 Dirac current density conservation

The Dirac equation implies current density conservation:

take the Dirac equation

$$i\hbar \frac{\partial \psi}{\partial t} = \left[c\,\vec{\alpha} \cdot \left(-i\hbar\,\vec{\nabla}\right) + \beta\,mc^2\right] \psi$$

multiply it with $\psi^\dagger$ from the left

$$i\hbar\,\psi^\dagger \frac{\partial \psi}{\partial t} = -i\hbar\,c\,\psi^\dagger\,\vec{\alpha} \cdot \vec{\nabla}\,\psi + mc^2\,\psi^\dagger\,\beta\,\psi \qquad \text{eq.(1)}$$

take the adjoint of Dirac equation

$$-i\hbar \frac{\partial \psi^\dagger}{\partial t} = i\hbar\,c\left(\vec{\nabla}\,\psi^\dagger\right) \cdot \vec{\alpha} + mc^2\,\psi^\dagger\,\beta$$

multiply it with ψ from the right

$$-i\hbar \frac{\partial \psi^\dagger}{\partial t}\,\psi = i\hbar\,c\left(\vec{\nabla}\,\psi^\dagger\right) \cdot \vec{\alpha}\,\psi + mc^2\,\psi^\dagger\,\beta\,\psi \qquad \text{eq.(2)}$$

$\uparrow$ acts only on $\psi^\dagger$ and <u>not</u> on ψ

subtract eq.(2) from eq.(1)

$$i\hbar\left(\psi^\dagger\,\frac{\partial\psi}{\partial t}+\frac{\partial\psi^\dagger}{\partial t}\,\psi\right) = -i\hbar c\,\psi^\dagger\,\vec{\alpha}\cdot\vec{\nabla}\,\psi - i\hbar c\left(\vec{\nabla}\,\psi^\dagger\right)\cdot\vec{\alpha}\,\psi$$

$\uparrow$ acts only on $\psi^\dagger$ and not on ψ

$$= -i\hbar c\,\vec{\nabla}\cdot\left(\psi^\dagger\,\vec{\alpha}\,\psi\right)$$

$$\Rightarrow \quad \frac{\partial}{\partial t}\,\psi^\dagger\psi = -c\,\vec{\nabla}\cdot\left(\psi^\dagger\,\vec{\alpha}\,\psi\right)$$

$$\frac{\partial}{\partial t}\,\rho + \vec{\nabla}\cdot\vec{J} = 0$$

continuity equation or conservation of probability density and current density

with

$$\begin{cases} \rho = \psi^\dagger\psi = \psi_1^*\,\psi_1 + \cdots\,\psi_4^*\,\psi_4 \geqslant 0 & \text{real and positive scalar} \\ \vec{J} = c\,\psi^\dagger\,\vec{\alpha}\,\psi & \text{real vector} \end{cases}$$

both under rotations

indeed:

$$J_k^* = J_k^\dagger = c\,\psi^\dagger\,\alpha_k^\dagger\,\psi^{\dagger\dagger} = c\,\psi^\dagger\,\alpha_k\,\psi = J_k \ . \qquad \text{real}$$

The continuity equation can be put into covariant form as well, i.e.

$$\sum_{\mu=1}^{4}\frac{\partial}{\partial x_\mu}\,J_\mu = \vec{\nabla}\cdot\vec{J} + \frac{\partial\rho}{\partial t} = 0$$

with

$$J_\mu = ic\overline{\psi}\,\gamma_\mu\,\psi \underset{\text{def}}{=} (J_x, J_y, J_z, ic\rho) \ ,$$

where $\overline{\psi}$ is the adjoint Dirac spinor

$$\overline{\psi} \underset{\text{def}}{=} \psi^\dagger\,\gamma_4 \ .$$

Indeed, we have

$$J_4 = ic\overline{\psi}\,\gamma_4\,\psi = ic\psi^\dagger\,\gamma_4\,\gamma_4\,\psi = ic\psi^\dagger\,\beta\beta\,\psi = ic\psi^\dagger\psi = ic\rho$$

$$J_k = ic\overline{\psi}\,\gamma_k\,\psi = ic\psi^\dagger\,\gamma_4\,\gamma_k\,\psi = ic\psi^\dagger\,\beta\,i\beta\alpha_k\,\psi = -c\psi^\dagger\,\alpha_k\,\psi \ . \qquad (k=1,2,3)$$

conclusions:

(i) ρ can be interpreted as a probability density. Thus the free Dirac equation seems to have all the required attributes:
(a) positive definiteness of the probability density
(b) it is first order in time and space derivatives.

(ii) We can now use minimal substitution, keeping $\vec{\alpha}$ and β as chosen by Dirac, to obtain the interacting Dirac equation

$$\left(i\hbar\,\frac{\partial}{\partial t} - e\,\phi\right)\,\psi = \left[\vec{\alpha}\cdot\left(-i\hbar\,\vec{\nabla} - \frac{e}{c}\,\vec{A}\right)\,c + \beta\,mc^2\right]\,\psi \ .$$

or in covariant form

$$\left(\sum_{\mu=1}^{4}\gamma_\mu\,D_\mu + \frac{mc}{\hbar}\right)\,\psi = 0$$

with the covariant derivatives $D_\mu \underset{\text{def}}{=} \dfrac{\partial}{\partial x_\mu} - i\,\dfrac{e}{\hbar c}\,A_\mu \quad (\mu = 1,2,3,4)$.

This interacting Dirac equation is invariant under local gauge transformations, because it is formulated in terms of covariant derivatives. It describes an electron with charge $(-|e|)$ or positron with charge $(+|e|)$, both with mass m and $g_S = 2$, interacting with an electromagnetic field given by $A_\mu = (A_x, A_y, A_z, i\phi)$.

20.4 Interacting Dirac and Maxwell fields

A Dirac field $\psi(\vec{r},t)$, interacting with a Maxwell field given by the vector and scalar potentials $A_\mu(\vec{r},t)$, fulfils the Dirac equation

$$\left[\sum_{\mu=1}^{4}\gamma_\mu\,D_\mu - \frac{mc}{\hbar}\right]\psi = 0 \ .$$

Here the covariant derivatives are defined as

$$D_\mu = \frac{\partial}{\partial x_\mu} - i\frac{e}{\hbar c}\,A_\mu$$

with

$$\frac{\partial}{\partial x_\mu} \underset{\text{def}}{=} \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial ict}\right)$$

and

$$A_\mu \underset{\text{def}}{=} \left(A_x,\, A_y,\, A_z,\, i\phi\right) \ .$$

The 4×4 matrices γ_μ, defined as

$$\gamma_k \underset{\text{def}}{=} \begin{pmatrix} O_2 & -i\sigma_k \\ i\sigma_k & O_2 \end{pmatrix} \qquad (k = 1,2,3)$$

and

$$\gamma_4 \underset{\text{def}}{=} \begin{pmatrix} O_2 & -I_2 \\ I_2 & O_2 \end{pmatrix} ,$$

obey the Clifford algebra

$$[\gamma_\mu, \gamma_\nu] \;=\; 2\delta_{\mu\nu} \qquad (\mu,\nu = 1,2,3,4) \;.$$

The electromagnetic field tensor $F_{\mu\nu}(\vec{r},t)$ is given in terms of the Maxwell field $A_\nu(\vec{r},t)$ as

$$F_{\mu\nu} \underset{\text{def}}{=} \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} \;, \qquad \text{eq.(1)}$$

and it obeys Maxwell's equations

$$\begin{cases} \displaystyle\sum_{\nu=1}^{4} \frac{\partial F_{\mu\nu}}{\partial x_\nu} \;=\; \frac{4\pi}{c} J_\mu & \text{eq.(2)} \\ \dfrac{\partial F_{\mu\nu}}{\partial x_\lambda} + \dfrac{\partial F_{\lambda\mu}}{\partial x_\nu} + \dfrac{\partial F_{\nu\lambda}}{\partial x_\mu} \;=\; 0 \qquad (\lambda,\mu,\nu = 1,2,3,4) \;. & \text{eq.(3)} \end{cases}$$

Eq.(3) is fulfilled identically by eq.(1). Here the current density

$$J_\mu(\vec{r},t) \;=\; J_\mu \underset{\text{def}}{=} \Big(J_x,\, J_y,\, J_z,\, ic\rho\Big) \;,$$

fulfils the continuity equation

$$\sum_{\mu=1}^{4} \frac{\partial J_\mu}{\partial x_\mu} \;=\; \vec{\nabla}\cdot\vec{J} + \frac{\partial \rho}{\partial t} \;=\; 0 \;.$$

In the Lorentz or Feynman gauge,

$$\sum_{\mu=1}^{4} \frac{\partial A_\mu}{\partial x_\mu} \;=\; \vec{\nabla}\cdot\vec{A} + \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \;=\; 0 \;,$$

Maxwell's equations (eq.(2)) assume their simplest form, i.e.

$$\Box\, A_\mu \;=\; -\frac{4\pi}{c} J_\mu \;.$$

We may identify the conserved electric current density $J_\mu(\vec{r},t)$ with that of the Dirac equation, multiplied with e, i.e.

$$J_\mu \;=\; iec\,\overline{\psi}\,\gamma_\mu\,\psi \;,$$

where the adjoint Dirac field is defined as

$$\overline{\psi} \underset{\text{def}}{=} \psi^\dagger\,\gamma_4 \;.$$

The space-time components of the current density are thus

$$\begin{cases} J_k \;=\; -ec\,\psi^\dagger\,\alpha_k\,\psi & (k = 1,2,3) \\ J_4 \;=\; iec\,\psi^\dagger\psi \;. \end{cases}$$

Inserting this conserved current density into Maxwell's equations (eq.(4), we arrive at a system of coupled nonlinear partial differential equations describing the mutual interactions between the Dirac and Maxwell fields, $\psi(\vec{r},t)$ and $A_\mu(\vec{r},t)$,

$$\begin{cases} \left[\displaystyle\sum_{\mu=1}^{4} \gamma_\mu \left(\frac{\partial}{\partial x_\mu} - i\frac{e}{\hbar c} A_\mu\right) - \frac{mc}{\hbar}\right]\psi \;=\; 0 \\ \Box\, A_\mu \;=\; -4\pi\, ei\overline{\psi}\,\gamma_\mu\,\psi \\ \dfrac{\partial A_\mu}{\partial x_\mu} \;=\; 0 \end{cases} .$$

These field equations, which can be solved in a perturbation expansion in powers of the electromagnetic coupling e, are the classical analogue of Quantum Electrodynamics, the theory of the interacctions of the electrons, positrons and photons.