Quantum Mechanics II

G. M. Pastor
Institut für Theoretische Physik
Fachbereich Mathematik und Naturwissenschaften
Universität Kassel

January 12, 2023

Contents

1 Fundamental concepts

The purpose of this chapter is to provide a quick revision of the mathematical formalism in quantum mechanics, which is usually covered in an introductory course.

1.1 State vectors and Hilbert space

In quantum mechanics a physical state is represented by a *state vector*, also called simply vector or ket, which is an element of a complex Hilbert space, i.e., a vector space with an inner product. The state vector, denoted for example by $|\,\alpha\,\rangle$, defines the state of a quantum mechanical system and has therefore the following two main properties:

i) It contains the answer to all the questions we are allowed to ask about a physical system, i.e., to measure. In other words the knowledge of $|\,\alpha\,\rangle$ allows one to derive the results of all measurements.

ii) The knowledge of the state $|\,\alpha, t\,\rangle$ of a system at a time t suffices to predict the state $|\,\alpha, t'\,\rangle$ at all times $t' > t$.

These two fundamental properties are of course common to the notion of a state in any mechanical theory, e.g., in classical mechanics and statistical mechanics, even though the state itself may have a very different nature (positions and momenta, or density matrix, for example).

The foundations of quantum mechanics can be formulated in the form of four main postulates.

Postulate I: The principle of superposition

For all $|\,\alpha\,\rangle$ and $|\,\beta\,\rangle$ physical states of a system, the linear combination $|\,\gamma\,\rangle = c_1\,|\,\alpha\,\rangle + c_2\,|\,\beta\,\rangle$ with c_1 and $c_2 \in \mathbb{C}$ is a possible physical state.

This closure condition, together with the following usual algebraic properties

i) associate and commutative laws for the sum of states,

ii) distributive laws with respect to a multiplication of sum of states by a scalar $c \in \mathbb{C}$,

iii) existence of neutral element $\vec{0}$ such that $|\,\alpha\,\rangle + \vec{0} = |\,\alpha\,\rangle$ for all $|\,\alpha\,\rangle$ and

iv) existence of an inverse $-|\,\alpha\,\rangle$ such that $-|\,\alpha\,\rangle + |\,\alpha\,\rangle = \vec{0}$ for all $|\,\alpha\,\rangle$

implies that the ensemble of all the physical states of a system forms a linear vector space $\mathbb{V}$.

Given a vector $|\,\alpha\,\rangle$ and any scalar $c \neq 1$, $c \in \mathbb{C}$, the vector $c\,|\,\alpha\,\rangle$ is mathematically different from $|\,\alpha\,\rangle$. However, it represents the same physical state, i.e., it has the same properties. The ensemble of all vectors $\{c\,|\,\alpha\,\rangle$ with $c \in \mathbb{C}\}$ is called a *ray*, by analogy to a straight line in $\mathbb{R}^n$. A physical state corresponds actually to a ray, not to a vector.

Postulate II: The inner product

An *inner product* $\langle \alpha \,|\, \beta \rangle$ is defined for any pair of vectors in $\mathbb{V}$ with the usual properties:

i) Hermiticity:

$$\langle \beta \,|\, \alpha \rangle = \langle \alpha \,|\, \beta \rangle^* \quad \forall \,|\, \alpha \rangle \text{ and } |\, \beta \rangle \in \mathbb{V}.$$

ii) Linearity at right:

$$\langle \gamma \,|\, c_1 \, \alpha + c_2 \, \beta \rangle = c_1 \, \langle \gamma \,|\, \alpha \rangle + c_2 \, \langle \gamma \,|\, \beta \rangle \text{ for all } c_1 \text{ and } c_2 \in \mathbb{C}.$$

Formally, one may wish to write

$$|\, \delta \rangle = c_1 \,|\, \alpha \rangle + c_2 \,|\, \beta \rangle$$

and request

$$\langle \gamma \,|\, \delta \rangle = c_1 \, \langle \gamma \,|\, \alpha \rangle + c_2 \, \langle \gamma \,|\, \beta \rangle.$$

One says that the inner product is linear at right.

Combining i) and ii), we have

$$\begin{aligned} \langle \, c_1 \, \alpha + c_2 \, \beta \,|\, \gamma \rangle &= \langle \gamma \,|\, c_1 \, \alpha + c_2 \, \beta \rangle^* \\ &= (c_1 \, \langle \gamma \,|\, \alpha \rangle + c_2 \, \langle \gamma \,|\, \beta \rangle)^* \\ &= c_1^* \, \langle \alpha \,|\, \gamma \rangle + c_2^* \, \langle \beta \,|\, \gamma \rangle. \end{aligned}$$

This property is known as the *antilinearity* of the inner product at left.

iii) Positive definiteness:

$$\langle \alpha \,|\, \alpha \rangle \geq 0 \quad \forall \,|\, \alpha \rangle \in \mathbb{V}$$

and

$$\langle \alpha \,|\, \alpha \rangle = 0 \;\Rightarrow\; |\, \alpha \rangle = 0.$$

This allows one to normalize any non-zero state $|\, \alpha \rangle$ as

$$|\, \tilde{\alpha} \rangle = \frac{|\, \alpha \rangle}{\sqrt{\langle \alpha \,|\, \alpha \rangle}}$$

and to introduce a metric (i.e., the notion of distance) in $\mathbb{V}$.

Two vector states are said to be *orthogonal* if

$$\langle \alpha \,|\, \beta \rangle = \langle \beta \,|\, \alpha \rangle = 0.$$

A vector space with an inner product is called a *Hilbert space*.

1.2 Dual vector space: Bras

The dual vector space $\mathbb{V}^*$ of a complex vector space $\mathbb{V}$ is the *vector space* of all *linear* functions $f : \mathbb{V} \to \mathbb{C}$. $\mathbb{V}$ and $\mathbb{V}^*$ have the same dimension. A simple way to realize the isomorphism $\mathbb{V} \leftrightarrow \mathbb{V}^*$ is to define the linear function

$$f_v(w) = \langle\, v \,|\, w \,\rangle \quad \forall\ w \in \mathbb{V}$$

as the element of $\mathbb{V}^*$ associated to $v \in \mathbb{V}$ ($f_v : \mathbb{V} \to \mathbb{C}$).

For each state vector or ket $|\, \alpha \,\rangle$ there is an associated element of the dual space $\mathbb{V}^*$ denoted by $\langle\, \alpha \,|$ and known as *bra*, which satisfies the following dual correspondence (DC) or isomorphism:

$$|\, \alpha \,\rangle \overset{\text{DC}}{\longleftrightarrow} \langle\, \alpha \,|.$$

Since $\langle\, c_1\, \alpha + c_2\, \beta \,|\, \gamma \,\rangle = \Big(c_1^* \langle\, \alpha \,| + c_2^* \langle\, \beta \,|\Big) |\, \gamma \,\rangle$, we have

$$c_1 \,|\, \alpha \,\rangle + c_2 \,|\, \beta \,\rangle \overset{\text{DC}}{\longleftrightarrow} c_1^* \langle\, \alpha \,| + c_2^* \langle\, \beta \,|$$

and in particular

$$c \,|\, \alpha \,\rangle \overset{\text{DC}}{\longleftrightarrow} c^* \langle\, \alpha \,|.$$

The dual space of bras $\langle\, \alpha \,|$ is obviously *isomorph* to the vector space $\mathbb{V}$. Moreover, $(\mathbb{V}^*)^* \equiv \mathbb{V}$. We may therefore regard the multiplication by a bra from the left

$$\langle\, \alpha \,| : \mathbb{V} \to \mathbb{C}$$

as a linear function from $\mathbb{V}$ to $\mathbb{C}$, and the multiplication by a ket from the right

$$|\, \alpha \,\rangle : \mathbb{V}^* \to \mathbb{C}.$$

as a linear function from $\mathbb{V}^*$ to $\mathbb{C}$.

1.3 Physical observables and the orthonormal basis derived from them

We consider a physical quantity or *observable* A and denote by

$$\{a', a'', \ldots\}$$

the ensemble of all possible *real values* that the measurement of this observable can yield. In general, a single observable or real number a' is not enough for defining the underlying physical state $|\, a' \,\rangle$ univocally. In order to avoid this ambiguity we request that A actually stands for a set of observables, known as *complete set of compatible observables*, which define the state $|\, a \,\rangle$ univocally. Consequently, each value a' stands for a set of quantum

numbers. For example, $a' \equiv$ position x', y', z' and spin projection σ', or $a' \equiv n'l'm'\sigma'$ for an electron in a spherically symmetric potential. In this context "complete" means that the measurement of A characterizes the physical state completely (besides a multiplicative factor, of course). No other observable can be found that is compatible with those in A, i.e., that is simultaneously measurable with arbitrary precision.

Postulate III: The relation to experiment

Let $|\, a' \,\rangle$ be one of the normalized state vectors corresponding to the physical state for which a measurement of A gives with certainty the value a'. We postulate that for any normalized state $|\, \alpha \,\rangle \in \mathbb{V}$ ($\langle\, \alpha \,|\, \alpha \,\rangle$=1), the probability $P_\alpha(a')$ of measuring the value a' of the observable A when the system is in the state $|\, \alpha \,\rangle$ is given by

$$P_\alpha(a') = |\langle\, \alpha \,|\, a' \,\rangle|^2.$$

The normalization of $|\, \alpha \,\rangle$ is not essential at this point. Taking into account that the probabilities satisfy $\sum_{a'} P(a') = 1$, the proportionality relation $P(a') \propto |\langle\, \alpha \,|\, a' \,\rangle|^2$ is sufficient to define $P(a')$ even if $|\, \alpha \,\rangle$ is not normalized.

An immediate consequence of the previous postulate is that

$$\langle\, a'' \,|\, a' \,\rangle = 0 \quad \text{for all } a' \neq a'',$$

since otherwise an A-measurement on $|\, a' \,\rangle$ could have a'' as a result. Let us recall that a' and a'' stand for sets of complete compatible observables. Thus,

$$\langle\, a' \,|\, a'' \,\rangle = \delta_{a'a''} \quad \forall\ a', a''.$$

The states $\{|\, a' \,\rangle, |\, a'' \,\rangle, \ldots\}$ form an orthonormal basis set in the Hilbert space of the system under consideration.

Postulate IV: Basis completeness

By hypothesis, actually by construction, the basis set $\{|\, \alpha' \,\rangle, |\, \alpha'' \,\rangle, \ldots\}$ is *complete*, i.e., it spans the Hilbert space $\mathbb{V}$ of physical states, since the list of values $\{a', a'', \ldots\}$ contains the ensemble of all possible results of an A-measurement.

Any state $|\, \beta \,\rangle$ can then be written as

$$|\, \beta \,\rangle = \sum_{a'} c_{a'} \,|\, a' \,\rangle.$$

The coefficients are unique since

$$\langle\, a'' \,|\, \beta \,\rangle = \sum_{a'} c_{a'} \,\langle\, a'' \,|\, a' \,\rangle = c_{a''}.$$

Thus, we can always write

$$|\, \beta \,\rangle = \sum_{a'} |\, a' \,\rangle\, \langle\, a' \,|\, \beta \,\rangle.$$

Taking into account that this holds for all $|\,\beta\,\rangle \in \mathbb{V}$ we have

$$\sum_{a'} |\,a'\,\rangle\,\langle\,a'\,| = \mathbb{1},$$

where $\mathbb{1}$ is the identity operator in $\mathbb{V}$. This is known as the *completeness relation*. Clearly, the converse is also true, i.e., whenever

$$\sum_{a'} |\,a'\,\rangle\,\langle\,a'\,| = \mathbb{1} \quad \Rightarrow \quad \{|\,a'\,\rangle, |\,a''\,\rangle, \ldots\} \text{ spans the vector space } \mathbb{V}.$$

This is an extremely useful way of writing the identity operator $\mathbb{1}$. It is probably the most important arithmetic identity derived from Dirac's bra-ket notation.

Notice that there are many different physically relevant complete sets of compatible observables A, B, C, etc. with spectra $\{a', a'', \ldots\}$, $\{b', b'', \ldots\}$, etc. It is the goal of transformation theory to work out the algebra for changing the representation of states. In this context one speaks of the A-representation, the B-representation, etc. For example, for a single spinless particle one may consider the coordinate representation $|\,\vec{x}\,\rangle$ or the momentum representation $|\,\vec{p}\,\rangle$.

1.4 Inner product in component form

The completeness relation allows us to derive an explicit expression for the inner product in terms of the components of the kets in a given basis. Let

$$|\,\alpha\,\rangle = \sum_{a'} c_{a'}\,|\,a'\,\rangle = \sum_{a'} |\,a'\,\rangle\,\langle\,a'\,|\,\alpha\,\rangle$$

and

$$|\,\beta\,\rangle = \sum_{a'} d_{a'}\,|\,a'\,\rangle = \sum_{a'} |\,a'\,\rangle\,\langle\,a'\,|\,\beta\,\rangle.$$

The inner product between these states is given by

$$\begin{aligned}\langle\,\alpha\,|\,\beta\,\rangle &= \langle\,\alpha\,|\,\mathbb{1}\,|\,\beta\,\rangle = \langle\,\alpha\,|\,\textstyle\left(\sum_{a'}|\,a'\,\rangle\,\langle\,a'\,|\right)|\,\beta\,\rangle \\ &= \sum_{a'} \langle\,\alpha\,|\,a'\,\rangle\,\langle\,a'\,|\,\beta\,\rangle = \sum_{a'} \langle\,a'\,|\,\alpha\,\rangle^*\,\langle\,a'\,|\,\beta\,\rangle.\end{aligned}$$

Consequently,

$$\langle\,\alpha\,|\,\beta\,\rangle = \sum_{a'} c^*_{a'}\,d_{a'}.$$

Notice that the coefficients or components of the left-hand-side vector state $|\,\alpha\,\rangle$ are conjugated.

1.5 Linear operators and their matrix elements

Consider an arbitrary linear operator $\hat{X} : \mathbb{V} \to \mathbb{V}$ and let $\hat{X}$ be defined by its action on the A-basis:[1]

$$\hat{X} \,|\, a' \,\rangle = \sum_{a''} c_{a''} \,|\, a'' \,\rangle.$$

Multiplying with $\langle\, a''' \,|$ from the left we have

$$\langle\, a''' \,|\, \hat{X} \,|\, a' \,\rangle = c_{a'''}$$

or

$$\hat{X} \,|\, a' \,\rangle = \sum_{a''} |\, a'' \,\rangle \,\langle\, a'' \,|\, \hat{X} \,|\, a' \,\rangle.$$

This could have been obtained more simply as

$$\hat{X} \,|\, a' \,\rangle = \mathbb{1}\, \hat{X} \,|\, a' \,\rangle = \sum_{a''} |\, a'' \,\rangle \,\langle\, a'' \,|\, \hat{X} \,|\, a' \,\rangle.$$

The *matrix elements* $\langle\, a'' \,|\, \hat{X} \,|\, a' \,\rangle$ are of course complex in general. The knowledge of $\langle\, a'' \,|\, \hat{X} \,|\, a' \,\rangle$ for all $|\, a' \,\rangle$ and $|\, a'' \,\rangle$ defines the operator completely, since

$$\hat{X} = \mathbb{1}\, \hat{X}\, \mathbb{1} = \sum_{a' a''} |\, a'' \,\rangle\langle\, a'' \,|\, \hat{X} \,|\, a' \,\rangle \,\langle\, a' \,|.$$

The array of matrix elements $\langle\, a'' \,|\, \hat{X} \,|\, a' \,\rangle$ is called the *matrix of the operator* $\hat{X}$ in the A-representation.

The representation of the product of two operators $\hat{X}$ and $\hat{Y}$ is easily obtained:

$$\begin{aligned} \hat{X}\, \hat{Y} &= \left(\sum_{a'} |\, a' \,\rangle \,\langle\, a' \,| \right) \hat{X} \left(\sum_{a''} |\, a'' \,\rangle \,\langle\, a'' \,| \right) \hat{Y} \left(\sum_{a'''} |\, a''' \,\rangle \,\langle\, a''' \,| \right) \\ &= \sum_{a'\, a'''} |\, a' \,\rangle \underbrace{\sum_{a''} \langle\, a' \,|\, \hat{X} \,|\, a'' \,\rangle \,\langle\, a'' \,|\, \hat{Y} \,|\, a''' \,\rangle}_{\langle\, a' \,|\, \hat{X}\, \hat{Y} \,|\, a''' \,\rangle} \langle\, a''' \,|. \end{aligned}$$

It follows that

$$\langle\, a' \,|\, \hat{X}\, \hat{Y} \,|\, a''' \,\rangle = \sum_{a''} \langle\, a' \,|\, \hat{X} \,|\, a'' \,\rangle \,\langle\, a'' \,|\, \hat{Y} \,|\, a''' \,\rangle,$$

as in the usual multiplication of matrices.

[1] A hat $\hat{}$ is used throughout these notes to distinguish operators from numbers.

The *trace of an operator* is defined by

$$\mathrm{Tr}\,\{\hat{X}\} = \sum_{a'} \langle\, a' \,|\, \hat{X} \,|\, a' \,\rangle,$$

and the *determinant* as

$$\det\{\hat{X}\} = \det\{\langle\, a' \,|\, \hat{X} \,|\, a'' \,\rangle\}.$$

As in the case of matrices the trace and determinant are independent of the representation, i.e., of the basis.

1.6 Projection operator $|\, a' \,\rangle \langle\, a' \,|$

Consider a state $|\, a' \,\rangle$ that is normalized to 1 ($\langle\, a' \,|\, a' \,\rangle = 1$). The operator

$$\hat{P}_{a'} = |\, a' \,\rangle \langle\, a' \,|$$

is a projection operator since

$$\begin{aligned} \hat{P}_{a'}^2 &= \hat{P}_{a'}\, \hat{P}_{a'} = |\, a' \,\rangle \langle\, a' \,|\, a' \,\rangle \langle\, a' \,| \\ &= \hat{P}_{a'}. \end{aligned}$$

For any vector $|\, \alpha \,\rangle = \sum_{a''} c_{a''} \,|\, a'' \,\rangle$ we have

$$\hat{P}_{a'} \,|\, \alpha \,\rangle = |\, a' \,\rangle \langle\, a' \,|\, \alpha \,\rangle = c_{a'} \,|\, a' \,\rangle.$$

Consequently, $\hat{P}_{a'} = |\, a' \,\rangle \langle\, a' \,|$ projects any state in the direction of $|\, a' \,\rangle$.

1.7 Outer product $|\, \beta \,\rangle \langle\, \alpha \,|$

A more general form of operator is the *outer product*

$$\hat{P}_{\beta\alpha} = |\, \beta \,\rangle \langle\, \alpha \,|.$$

Applying $\hat{P}_{\beta\alpha} = |\, \beta \,\rangle \langle\, \alpha \,|$ to a given state $|\, \gamma \,\rangle$, one obtains $|\, \beta \,\rangle \langle\, \alpha \,|\, \gamma \,\rangle$, which points in the direction of $|\, \beta \,\rangle$ with a component equal to the projection of $|\, \gamma \,\rangle$ along the direction of $|\, \alpha \,\rangle$.

1.8 Transpose, conjugate and adjoint of operators

The following transformations of operators are particularly useful. Given an operator $\hat{X}$ we define, as in the case of matrices, the transpose, complex conjugate and adjoint of $\hat{X}$

as follows:

$$\begin{array}{lll}
\text{transpose} & \hat{X}^t = \sum_{a'a''} |\, a' \,\rangle \langle\, a'' \,|\, \hat{X} \,|\, a' \,\rangle \langle\, a'' \,| & \Rightarrow \quad \left(\hat{X}^t\right)^t = \hat{X} \\
\text{complex conjugate} & \hat{X}^* = \sum_{a'a''} |\, a' \,\rangle \langle\, a' \,|\, \hat{X} \,|\, a'' \,\rangle^* \langle\, a'' \,| & \Rightarrow \quad \left(\hat{X}^*\right)^* = \hat{X} \\
\text{adjoint} & \hat{X}^\dagger = \sum_{a'a''} |\, a' \,\rangle \langle\, a'' \,|\, \hat{X} \,|\, a' \,\rangle^* \langle\, a'' \,| & \Rightarrow \quad \left(\hat{X}^\dagger\right)^\dagger = \hat{X} \\
& = \sum_{a'a''} |\, a' \,\rangle \underbrace{\langle\, \hat{X}\, a' \,|\, a'' \,\rangle}_{\langle\, a' \,|\, \hat{X}^\dagger \,|\, a'' \,\rangle} \langle\, a'' \,|. &
\end{array}$$

A particularly useful and defining property of the adjoint operator $\hat{X}^\dagger$ is

$$\langle\, \alpha \,|\, \hat{X}\beta \,\rangle = \langle\, \hat{X}^\dagger \alpha \,|\, \beta \,\rangle.$$

If $\hat{X}^t = \hat{X}$ we say that $\hat{X}$ is *symmetric*, if $\hat{X}^* = \hat{X}$ we say that $\hat{X}$ is *real*, and if $\hat{X}^\dagger = \hat{X}$ we say that $\hat{X}$ is *hermitian*.

One can easily show that

$$(\hat{X}\,\hat{Y})^t = Y^t\,\hat{X}^t,$$

$$(\hat{X}\,\hat{Y})^\dagger = Y^\dagger\,\hat{X}^\dagger,$$

$$(c\,\hat{X})^* = c^*\,\hat{X}^*,$$

and

$$(c\,\hat{X})^\dagger = c^*\,\hat{X}^\dagger \quad \forall\, c \in \mathbb{C}.$$

Examples:

i) Consider the outer product $|\, \alpha \,\rangle \langle\, \beta \,| = \hat{P}_{\alpha\beta}$. It holds that

$$\hat{P}_{\alpha\beta}^\dagger = \left(\sum_{a'a''} |\, a' \,\rangle \langle\, a' \,|\, \alpha \,\rangle \langle\, \beta \,|\, a'' \,\rangle \langle\, a'' \,| \right)^\dagger = \sum_{a'a''} |\, a' \,\rangle \langle\, a'' \,|\, \alpha \,\rangle^* \langle\, \beta \,|\, a' \,\rangle^* \langle\, a'' \,|$$
$$= \sum_{a'a''} |\, a' \,\rangle \langle\, a' \,|\, \beta \,\rangle \langle\, \alpha \,|\, a'' \,\rangle \langle\, a'' \,| = |\, \beta \,\rangle \langle\, \alpha \,| = \hat{P}_{\beta\alpha}.$$

ii) For a projection operator we have

$$\hat{P}_\alpha^\dagger = (|\, \alpha \,\rangle \langle\, \alpha \,|)^\dagger = \hat{P}_\alpha \quad \Rightarrow \quad \hat{P}_\alpha \text{ is hermitian.}$$

iii) $\mathrm{Tr}\,\{\hat{X}\,\hat{Y}\} = \mathrm{Tr}\,\{\hat{Y}\,\hat{X}\}$.

iv) $\mathrm{Tr}\,\{|\,\alpha\,\rangle\,\langle\,\beta\,|\} = \langle\,\beta\,|\,\alpha\,\rangle$. This implies that for states belonging to an orthonormal basis

$$\mathrm{Tr}\,\{|\,a'\,\rangle\,\langle\,a''\,|\} = \delta_{a'a''}.$$

1.9 Operators and the ket $\leftrightarrow$ bra correspondence

Let $\hat{X}$ be a linear operator $\hat{X} : \mathbb{V} \to \mathbb{V}$ and consider an arbitrary state vector $|\,\alpha\,\rangle$:

$$|\,\gamma\,\rangle = \hat{X}\,|\,\alpha\,\rangle.$$

We would like to obtain the bra $\langle\,\gamma\,|$ associated to $|\,\gamma\,\rangle$:

$$|\,\gamma\,\rangle = \hat{X}\,|\,\alpha\,\rangle = \sum_{a'a''} |\,a'\,\rangle\,\langle\,a'\,|\,\hat{X}\,|\,a''\,\rangle\,\langle\,a''\,|\,\alpha\,\rangle.$$

The corresponding bra is given by

$$\begin{aligned}\langle\,\gamma\,| = \langle\,\hat{X}\,\alpha\,| &= \sum_{a'a''} \langle\,a'\,|\,\hat{X}\,|\,a''\,\rangle^*\,\langle\,a''\,|\,\alpha\,\rangle^*\,\langle\,a'\,| \\ &= \sum_{a'a''} \langle\,\alpha\,|\,a''\,\rangle\,\langle\,a''\,|\,\hat{X}^\dagger\,|\,a'\,\rangle\,\langle\,a'\,| \\ &= \langle\,\alpha\,|\,\hat{X}^\dagger.\end{aligned}$$

If we denote $(|\,\alpha\,\rangle)^\dagger = \langle\,\alpha\,|$ we have $(\hat{X}\,|\,\alpha\,\rangle)^\dagger = \langle\,\alpha\,|\,\hat{X}^\dagger \;\Rightarrow\; \langle\,\hat{X}\,\alpha\,| = \langle\,\alpha\,|\,\hat{X}^\dagger$. Only for *hermitian operators* we have

$$(\hat{X}\,|\,\alpha\,\rangle)^\dagger = \langle\,\alpha\,|\,\hat{X} \qquad (\hat{X}^\dagger = \hat{X}).$$

Examples of ket-bra manipulations:

For the sake of clarity we use a dot "$\cdot$" to distinguish the inner product between states from the kets or bras obtained by applying linear operators to them.

i) $\langle\,\beta\,|\,\hat{X}\,|\,\alpha\,\rangle = (\langle\,\beta\,|) \cdot (\hat{X}\,|\,\alpha\,\rangle) = \Big[(\langle\,\alpha\,|\,\hat{X}^\dagger) \cdot (|\,\beta\,\rangle)\Big]^* = \langle\,\alpha\,|\,\hat{X}^\dagger\,|\,\beta\,\rangle^*$.

ii) $\langle\,\beta\,|\,\hat{X}\,|\,\alpha\,\rangle = \langle\,\beta\,| \cdot |\,\hat{X}\,\alpha\,\rangle = \langle\,\hat{X}\,\alpha\,| \cdot |\,\beta\,\rangle^* = \Big[(\langle\,\alpha\,|) \cdot (\hat{X}^\dagger\,|\,\beta\,\rangle)\Big]^* = \langle\,\alpha\,|\,\hat{X}^\dagger\,|\,\beta\,\rangle^*$.

iii) $\langle\,\beta\,|\,\hat{X}\,|\,\alpha\,\rangle = \langle\,X^\dagger\,\beta\,|\,\alpha\,\rangle = \langle\,\alpha\,|\,\hat{X}^\dagger\,\beta\,\rangle^* = \langle\,\alpha\,|\,\hat{X}^\dagger\,|\,\beta\,\rangle^*$.

1.10 Changes between representations: Unitary operators

Let $\{|\,a'\,\rangle, \ldots |\,a^k\,\rangle \ldots\}$ be the basis of the A-representation (i.e., derived from observable A) and $\{|\,b'\,\rangle, \ldots |\,b^k\,\rangle \ldots\}$ be the basis of the B-representation or B-basis. For example, A may correspond to the projection of the angular momentum along the z-axis and B along the x-axis. Since the two bases are orthonormal and complete, the descriptions of the physical states are equivalent. They both span the same physical space $\mathbb{V}$. We are interested in finding how these equivalent descriptions are related, i.e., how to perform the corresponding change of basis or change of representation.

The operator $\hat{U}_{ab}$ that connects the two representations transforms the B-basis into the A-basis as

$$\hat{U}_{ab}\,|\,b^k\,\rangle = |\,a^k\,\rangle \quad \forall\ k = 1, 2, \ldots$$

This operator is assumed to be linear in order to respect the superposition of states:

$$\hat{U}_{ab}\,(c_1\,|\,\alpha\,\rangle + c_2\,|\,\beta\,\rangle) = c_1\,\hat{U}_{ab}\,|\,\alpha\,\rangle + c_2\,\hat{U}_{ab}\,|\,\beta\,\rangle \qquad \text{with } c_1 \text{ and } c_2 \in \mathbb{C}.$$

Geometrically, one may picture $\hat{U}_{ab}$ as a "rotation" mapping the $|\,b^k\,\rangle$ states (regarded as unit vectors or axes) into the $|\,a^k\,\rangle$ states. It is easy to see that the operator $\hat{U}_{ab}$ is given by

$$\hat{U}_{ab} = \sum_k |\,a^k\,\rangle\,\langle\,b^k\,|.$$

This simple form relies on the orthonormality of the basis A and B. Indeed, we have

$$\left(\sum_k |\,a^k\,\rangle\,\langle\,b^k\,|\right)|\,b^l\,\rangle = \sum_k |\,a^k\,\rangle\,\delta_{kl} = |\,a^l\,\rangle.$$

The matrix form of $\hat{U}_{ab}$ in the old representation (or B-basis) and in the new representation (or A-basis) are given by

$$\begin{aligned}\hat{U}_{ab} &= \sum_k |\,a^k\,\rangle\,\langle\,b^k\,| \\ &= \sum_{lk} |\,b^k\,\rangle\,\langle\,b^k\,|\,a^l\,\rangle\,\langle\,b^l\,| \\ &= \sum_{lk} |\,a^k\,\rangle\,\langle\,b^k\,|\,a^l\,\rangle\,\langle\,a^l\,|.\end{aligned}$$

The matrix elements $\langle\,a^k\,|\,\hat{U}_{ab}\,|\,a^l\,\rangle = \langle\,b^k\,|\,a^l\,\rangle$ are called *transformation functions* $[\langle\,\text{old basis } (k)\,|\,\text{new basis } (l)\,\rangle]$. The reason for this name is the role of these overlaps in the transformation of the components of vector states:

$$|\,\alpha\,\rangle = \sum_l |\,a^l\,\rangle\,\langle\,a^l\,|\,\alpha\,\rangle$$

$$\langle\,b^k\,|\,\alpha\,\rangle = \sum_l \langle\,b^k\,|\,a^l\,\rangle\,\langle\,a^l\,|\,\alpha\,\rangle = \sum_l |\,b^l\,\rangle\,\langle\,b^l\,|\,\alpha\,\rangle.$$

The inverse operator $(\hat{U}_{ab})^{-1} = \hat{U}_{ba}$ is given by

$$(\hat{U}_{ab})^{-1} = \hat{U}_{ba} = \sum_k |\, b^k \,\rangle\langle\, a^k \,| = \hat{U}^\dagger_{ab}.$$

It is straightforward to see that $\sum_k |\, b^k \,\rangle\langle\, a^k \,|\, \hat{U}_{ab} = \mathbb{1}$. Operators satisfying $\hat{U}^\dagger = \hat{U}^{-1}$ are called *unitary operators*. Linear operators transforming an orthonormal basis in another orthonormal basis are always unitary and vice versa. Finally, one may easily show that the product of two unitary transformations $a \to b \to c$ is the unitary transformation $a \to c$.

$$\hat{U}_{ab}\,\hat{U}_{bc} = \sum_{kl} |\, a^k \,\rangle\langle\, b^k \,|\, b^l \,\rangle\langle\, c^l \,| = \sum_l |\, a^l \,\rangle\langle\, c^l \,|.$$

1.11 Invariance of the inner product under unitary transformations

Theorem: A linear operator $\hat{U}$ is unitary if and only if it preserves the scalar product, i.e., the scalar product is invariant under the transformation $\hat{U}$.

Let us assume $\hat{U} = \hat{U}^\dagger$ and denote the transformed states

$$|\, \alpha' \,\rangle = \hat{U}\,|\, \alpha \,\rangle \quad \Leftrightarrow \quad \langle\, \alpha' \,| = \langle\, \alpha \,|\, \hat{U}^\dagger$$

and

$$|\, \beta' \,\rangle = \hat{U}\,|\, \beta \,\rangle$$

for any $|\, \alpha \,\rangle$ and $|\, \beta \,\rangle \in \mathbb{V}$. It is clear that the inner product between the transformed states

$$\langle\, \alpha' \,|\, \beta' \,\rangle = \langle\, \alpha \,|\, \hat{U}^\dagger\,\hat{U}\,|\, \beta \,\rangle = \langle\, \alpha \,|\, \beta \,\rangle$$

is invariant under $\hat{U}$.

Conversely, if the inner product is preserved, i.e., $\langle\, \hat{U}\,\alpha \,|\, \hat{U}\,\beta \,\rangle = \langle\, \alpha \,|\, \beta \,\rangle \;\; \forall\; |\, \alpha \,\rangle$ and $|\, \beta \,\rangle \in \mathbb{V}$, this holds in particular for any basis of $\mathbb{V}$. Denoting

$$|\, b^k \,\rangle = \hat{U}\,|\, a^k \,\rangle \qquad k = 1,\, 2,\, \ldots$$

we have

$$\langle\, b^k \,|\, b^l \,\rangle = \langle\, \hat{U}\,a^k \,|\, \hat{U}\,a^l \,\rangle = \langle\, a^k \,|\, a^l \,\rangle = \delta_{kl}.$$

Thus, $|\, b^k \,\rangle$ with $k = 1,\, 2,\, \ldots$ is an orthonormal basis, which implies that $\hat{U}$ is unitary.

1.12 Group of unitary transformations

We have seen that the product of two unitary transformations is unitary. In fact we have

$$(\hat{U}_1\,\hat{U}_2)^\dagger = \hat{U}_2^\dagger\,\hat{U}_1^\dagger \quad \text{and}$$

$$(\hat{U}_1\,\hat{U}_2)^\dagger\,(\hat{U}_1\,\hat{U}_2) = \hat{U}_2^\dagger\,\hat{U}_1^\dagger\,\hat{U}_1\,\hat{U}_2 = \hat{U}_2^\dagger\,\hat{U}_2 = \mathbb{1}.$$

Since $\mathbb{1}$ is obviously unitary, the unitary transformations of a system form a group with the product or composition as the group operation. Symmetry subgroups such as translations, rotations, and time displacements are extremely important in physics.

1.13 Transformation of operators: Equivalent operators

Consider a linear operator $\hat{X} : \mathbb{V} \to \mathbb{V}$ that maps

$$|\,\alpha\,\rangle \xrightarrow{\hat{X}} |\,\beta\,\rangle = \hat{X}\,|\,\alpha\,\rangle.$$

Let $\hat{U}$ be an arbitrary invertible linear operator (not necessarily unitary) and let $\hat{U}^{-1}$ denote its inverse. We look for the operator $\hat{X}'$ that maps the transformed state $|\,\alpha'\,\rangle = \hat{U}\,|\,\alpha\,\rangle$ in the state $|\,\beta'\,\rangle = \hat{U}\,|\,\beta\,\rangle$:

$$|\,\alpha'\,\rangle = \hat{U}\,|\,\alpha\,\rangle \xrightarrow{\hat{X}'} |\,\beta'\,\rangle = \hat{U}\,|\,\beta\,\rangle = \hat{U}\,\hat{X}\,|\,\alpha\,\rangle.$$

Taking into account that

$$\hat{U}\,|\,\alpha\,\rangle \xrightarrow{\hat{X}'} \hat{U}\,\hat{X}\,|\,\alpha\,\rangle$$

we have

$$\hat{X}'\,\hat{U}\,|\,\alpha\,\rangle = \hat{U}\,\hat{X}\,|\,\alpha\,\rangle \quad \forall\,|\,\alpha\,\rangle.$$

Choosing $|\,\alpha\,\rangle = \hat{U}^{-1}\,|\,\beta\,\rangle$ we obtain

$$\hat{X}'\,|\,\beta\,\rangle = \hat{U}\,\hat{X}\,\hat{U}^{-1}\,|\,\beta\,\rangle \quad \forall\,|\,\beta\,\rangle$$

or equivalently,

$$\hat{X}' = \hat{U}\,\hat{X}\,\hat{U}^{-1}.$$

In the particular case of unitary transformations, we have $\hat{U}^\dagger = \hat{U}^{-1}$ and therefore

$$\hat{X}' = \hat{U}\,\hat{X}\,\hat{U}^\dagger.$$

This kind of transformation $\hat{X} \to \hat{X}'$ is known as a *unitary transformation of operators*. The operators X and X' are said to be *unitary equivalent*.

In particular for $\hat{U} = \hat{U}_{ba} = \sum_k |\,b^k\,\rangle\,\langle\,a^k\,|$ we have

$$\hat{X}' = \sum_{k,\,l} |\,b^k\,\rangle\,\langle\,a^k\,|\,\hat{X}\,|\,a^l\,\rangle\,\langle\,b^l\,|.$$

The following relations between unitary equivalent operators can be easily demonstrated:

i) $\mathrm{Tr}\,\{\hat{X}'\} = \mathrm{Tr}\,\{\hat{U}\,\hat{X}\,\hat{U}^\dagger\} = \mathrm{Tr}\,\{\hat{U}^\dagger\,\hat{U}\,\hat{X}\} = \mathrm{Tr}\,\{\hat{X}\}$.

ii) $\det(\hat{U}\,\hat{X}\,\hat{U}^\dagger) = \det\hat{X}' = \det\hat{X}$, since $\det(\hat{X}\,\hat{Y}) = \det(\hat{Y}\,\hat{X})\ \forall\ \hat{X},\hat{Y}$.

iii) $\det\hat{X}^\dagger = [\det\ \hat{X}]^*$, since $\det\hat{X} = \det\hat{X}^t\ \forall\ \hat{X}$.

iv) $|\det\hat{U}| = 1$ for all unitary $\hat{U}$, since $\det(\hat{X}\,\hat{Y}) = \det(\hat{X})\,\det(\hat{Y})\ \forall\ \hat{X}$ and $\hat{Y}$, and $\det(\hat{U}\,\hat{U}^\dagger) = 1$ [$\hat{U}^\dagger = \hat{U}^{-1}$ and $\det(\hat{U}^\dagger) = \det(\hat{U})^*$].

v) $(\hat{U}\,\hat{X}\,\hat{U}^\dagger)^\dagger = \hat{U}\,\hat{X}^\dagger\,\hat{U}^\dagger \quad\Rightarrow\quad (X')^\dagger = (X^\dagger)'$.

1.14 Observables

We can now turn to the question of constructing operators for the observables themselves. The informations at our disposal are the following:

i) The possible values of A are $\{a', a'', \ldots\}$,

ii) the states $\{|\,a'\,\rangle, |\,a''\,\rangle, \ldots\}$ have definite values of A, i.e., a measurement of the observable A on $|\,a'\,\rangle$ yields with certainty a', and

iii) for any normalized state vector $|\,\alpha\,\rangle$ ($\langle\,\alpha\,|\,\alpha\,\rangle = 1$) the probability of measuring the value a' of A is given by $|\langle\,\alpha\,|\,a'\,\rangle|^2$.

The *operator* $\hat{A}$ associated to the *observable* A is defined by requiring a most simple and physically transparent condition, namely, that for all state vectors $|\,\alpha\,\rangle$ the average $\langle\,\alpha\,|\,\hat{A}\,|\,\alpha\,\rangle$ of the linear operator $\hat{A}$ coincides with the *mean value*, also known as *expectation value*, $\langle\,A\,\rangle_\alpha$ of a series of measurements of the observable A in the state $|\,\alpha\,\rangle$.[2] The expectation value is given by

$$\langle\,A\,\rangle_\alpha = \sum_{a'} a'\,|\langle\,\alpha\,|\,a'\,\rangle|^2,$$

where $P_\alpha(a') = |\langle\,\alpha\,|\,a'\,\rangle|^2$ is the probability for measuring the value a' of the observable A in $|\,\alpha\,\rangle$ ($\langle\,\alpha\,|\,\alpha\,\rangle = 1$). We must thus have

$$\langle\,\alpha\,|\,\hat{A}\,|\,\alpha\,\rangle = \sum_{a'} \langle\,\alpha\,|\,a'\,\rangle\,a'\,\langle\,a'\,|\,\alpha\,\rangle \quad \forall\ |\,\alpha\,\rangle.$$

[2] At this point it is useful to recall that two linear operators $\hat{A}$ and $\hat{B}$ having the same mean values on all states $|\,\alpha\,\rangle$ of a vector space are necessarily identical ($\langle\,\alpha\,|\,\hat{A}\,|\,\alpha\,\rangle = \langle\,\alpha\,|\,\hat{A}\,|\,\alpha\,\rangle\ \forall|\,\alpha\,\rangle \Rightarrow \hat{A} = \hat{B}$). This means that a linear operator is unambiguously defined by its mean value on all states. The proof is left as an exercise.

The operators $\hat{A}$ and $\sum_{a'} |\, a' \,\rangle\, a' \,\langle\, a' \,|$ have the same average value for all states and are therefore identical (see footnote 2). Consequently,

$$\hat{A} = \sum_{a'} |\, a' \,\rangle\, a' \,\langle\, a' \,|.$$

It follows that the ket $|\, a' \,\rangle$ is an eigenvector of $\hat{A}$ with the eigenvalue a':

$$\hat{A} \,|\, a' \,\rangle = a' \,|\, a' \,\rangle.$$

Applying the operator $\hat{A}$ on a state $|\, a' \,\rangle$, which has a defined value of A, gives the vector state $a' \,|\, a' \,\rangle$, which is proportional to $|\, a' \,\rangle$. Moreover, $\hat{A}$ is hermitian, since all possible values a' of the observable A are real:

$$\hat{A} = \sum_{a'} |\, a' \,\rangle\, a' \,\langle\, a' \,| \quad \Rightarrow \quad \hat{A}^\dagger = \sum_{a'} |\, a' \,\rangle\, a'^* \,\langle\, a' \,| = \hat{A} \quad (a' \in \mathbb{R}).$$

We arrive at the fundamental conclusion that *the operators of all physical observables are hermitian.*

An important general property of hermitian operators is that the eigenvectors $|\, a' \,\rangle$ and $|\, a'' \,\rangle$ corresponding to different eigenvalues $a' \neq a''$ are always orthogonal. Indeed, we can compute $\langle\, a'' \,|\, \hat{A} \,|\, a' \,\rangle$ in two alternative ways. First, acting with $\hat{A}$ on the ket $|\, a' \,\rangle$ we obtain

$$\langle\, a'' \,|\, \hat{A} \,|\, a' \,\rangle = a' \,\langle\, a'' \,|\, a' \,\rangle. \tag{1.1}$$

Second, acting with $\hat{A}^\dagger = \hat{A}$ on the bra $\langle\, a'' \,|$ and using that $a'' \in \mathbb{R}$ we obtain

$$\langle\, a'' \,|\, \hat{A} \,|\, a' \,\rangle = a''^* \,\langle\, a'' \,|\, a' \,\rangle = a'' \,\langle\, a'' \,|\, a' \,\rangle. \tag{1.2}$$

Subtracting Eqs. (1.1) and (1.2) we conclude that $\langle\, a'' \,|\, a' \,\rangle = 0$ whenever $a' \neq a''$.

In the presence of degeneracies, i.e., $a' = a''$ for linearly independent eigenstates $|\, a' \,\rangle$ and $|\, a'' \,\rangle$, it is possible to orthogonalize the eigenstates within each degenerate subspace by performing appropriate linear combinations among them. One concludes that the eigenstates of hermitian operators form (or can be chosen to form) a complete orthonormal basis of $\mathbb{V}$.

If the Hilbert space $\mathbb{V}$ has a finite dimension, the eigenvalues of $\hat{A}$ can be determined by solving the equation

$$\det\,(A - \lambda \mathbb{1}) = 0.$$

In fact, $\hat{A} \,|\, a' \,\rangle = a' \,|\, a' \,\rangle \Leftrightarrow (A - \lambda \mathbb{1}) \,|\, a' \,\rangle = 0 \Leftrightarrow \hat{A} - \lambda \hat{\mathbb{1}}$ is singular $\Leftrightarrow \det\,(A - \lambda \mathbb{1}) = 0$. For infinite dimensions the eigenvalue equation usually takes the form of a differential equation like the Schrödinger equation or the spherical harmonic differential equation.

1.15 Unitary equivalent observables

Consider now two equivalent representations A and B of our Hilbert space $\mathbb{V}$ associated to two different sets of f compatible observables $\hat{A}_1, \hat{A}_2, \ldots \hat{A}_f$.

$$\hat{A}_i \,|\, a_1^k \, a_2^k \, \ldots \, a_f^k \,\rangle = a_i^k \,|\, a_1^k \, a_2^k \, \ldots \, a_f^k \,\rangle,$$

where $a_1^k \, a_2^k \, \ldots \, a_f^k$ are the f quantum numbers characterizing the element k of the complete A-basis. Similarly for the B-basis

$$\hat{B}_i \,|\, b_1^k \, b_2^k \, \ldots \, b_f^k \,\rangle = b_i^k \,|\, b_1^k \, b_2^k \, \ldots \, b_f^k \,\rangle.$$

The unitary transformation $\hat{U}_{ba}$ connecting the two representations is given by

$$\hat{U}_{ba} = \sum_k |\, b_1^k \, \ldots \, b_f^k \,\rangle \,\langle\, a_1^k \, \ldots \, a_f^k \,|,$$

from which it follows as usual

$$\hat{U}_{ba} \,|\, a_1^k \, \ldots \, a_f^k \,\rangle = |\, b_1^k \, \ldots \, b_f^k \,\rangle \quad \text{and} \quad |\, a_1^k \, \ldots \, a_f^k \,\rangle = \hat{U}_{ba}^\dagger \,|\, b_1^k \, \ldots \, b_f^k \,\rangle.$$

Applying $\hat{U}_{ba}$ to $\hat{A}_i \,|\, a_1^k \, \ldots \, a_f^k \,\rangle = a_i^k \,|\, a_1^k \, \ldots \, a_f^k \,\rangle$ we obtain

$$\hat{U}_{ba} \, \hat{A}_i \, \hat{U}_{ba}^\dagger \,|\, b_1^k \, \ldots \, b_f^k \,\rangle = a_i^k \,|\, b_1^k \, \ldots \, b_f^k \,\rangle.$$

This shows that each element of the B-basis $|\, b_1^k \, \ldots \, b_f^k \,\rangle$ is a simultaneous eigenket of the f transformed observables $\hat{U}_{ba} \, \hat{A}_i \, \hat{U}_{ba}^\dagger$. One concludes that *unitary equivalent observables have the same spectra.* For example, $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ have the same spectrum, also $\hat{x}$, $\hat{y}$ and $\hat{z}$, or $\hat{p}_x$, $\hat{p}_y$ and $\hat{p}_z$.

In addition, we know that all $|\, b_1^k \, \ldots \, b_f^k \,\rangle$ are also eigenkets of the complete set of observables $\hat{B}_1 \, \ldots \, \hat{B}_f$. Therefore,

$$\left[\hat{B}_i, \; \hat{U}_{ba} \, \hat{A}_j \, \hat{U}_{ba}^\dagger\right] = 0 \quad \forall \; i, \, j = 1, \, \ldots \, f.$$

Since the $\{\hat{B}_i\}$ are complete, any operator $\hat{A}_j$ that commutes with all of them is a function of $\{\hat{B}_i\}$. All the eigenvalues and eigenvectors of the $\hat{A}_j$'s are functions of $\{\hat{B}_i\}$.

In general, we have simply $\hat{B}_i = \hat{U}_{ba} \, \hat{A}_i \, \hat{U}_{ba}^\dagger \;\; \forall \, i$, or this one-to-one identity can be achieved by a trivial reordering. In fact, one usually derives the $\hat{B}_i$ by transforming another complete set $\{A_i\}$. For instance, starting from $A = \{L^2, L_z\}$ one obtains $B = \{L^2, L_x\}$ by a rotation of the coordinate system. In this case $b_i^k = a_i^k$ for all $i = 1, \ldots \, f$ and all k, i.e., *the spectra are the same in both representations.* Therefore, is is usually safe to assume that the observables that define all the different representations are unitary equivalent, so that the spectra coincide in all representations.

Finally, it is important to observe that all the algebraic relations or physical laws between observables, whether compatible or not, are invariant under unitary transformations. This was physically expected since unitary transformations connect *equivalent observables*. Consider, for example, the commutation relations between angular momentum operators

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

where x, y, z refer to some coordinate system, and let $\hat{U}$ be the unitary transformation corresponding to some rotation of the coordinate system:

$$\begin{cases} x \to x' \\ y \to y' \\ z \to z'. \end{cases}$$

The transformed operators are then given by

$$\begin{cases} \hat{U}\,\hat{L}_x\,\hat{U}^\dagger = \hat{L}_{x'} \\ \hat{U}\,\hat{L}_y\,\hat{U}^\dagger = \hat{L}_{y'} \\ \hat{U}\,\hat{L}_z\,\hat{U}^\dagger = \hat{L}_{z'}. \end{cases}$$

It follows that

$$\hat{L}_x\,\hat{L}_y - \hat{L}_y\,\hat{L}_x = i\,\hbar\,\hat{L}_z \;\Leftrightarrow\; \hat{U}\,\hat{L}_x\,\hat{U}^\dagger\,\hat{U}\,\hat{L}_y\,\hat{U}^\dagger - \hat{U}\,\hat{L}_y\,\hat{U}^\dagger\,\hat{U}\,\hat{L}_x\,\hat{U}^\dagger = i\,\hbar\,\hat{U}\,\hat{L}_z\,\hat{U}^\dagger$$

and consequently

$$\left[\hat{L}_{x'}, \hat{L}_{y'}\right] = i\,\hbar\,\hat{L}_{z'}.$$

The same holds for $[\hat{x}_i, \hat{p}_i] = i\,\hbar\,\delta_{ij} \;\Leftrightarrow\; [\hat{x}'_i, \hat{p}'_i] = i\,\hbar\,\delta_{ij}$. One can actually easily prove that

$$\hat{U}\,f(\hat{A})\,\hat{U}^\dagger = f(\hat{U}\,\hat{A}\,\hat{U}^\dagger) = f(\hat{A}')$$

for any function f allowing a Taylor or Laurent expansion.

1.16 Continuous spectra

The formalism discussed so far applies, from a mathematical perspective, to Hilbert spaces of finite dimension N, i.e., to discrete bounded spectra. Essentially all the results also apply for $N \to \infty$, i.e., discrete unbounded spectra where the dimension is countable infinite [e.g., $\hat{L}^2 = (l+1)\,l$ with $l \in \mathbb{N}$]. Continuous spectra are actually much more subtle. Nevertheless, the formalism that we discussed so far for discrete spectra can be generalized to observables like the position $\vec{r} = (x, y, z)$ or the momentum $\vec{p} = (p_x, p_y, p_z)$, for which the outcome of a measurement is a continuous variable.

Let us denote by $\hat{\xi}$ the operator of an observable with a continuous spectrum. The eigenvalues are denoted by ξ' and the eigenkets by $|\,\xi'\,\rangle$. Thus we have

$$\hat{\xi}\,|\,\xi'\,\rangle = \xi'\,|\,\xi'\,\rangle$$

and

$$\langle\,\xi'\,|\,\hat{\xi} = \xi'\,\langle\,\xi'\,|,$$

since the eigenvalues are real. Therefore $\hat{\xi}^\dagger = \hat{\xi}$. The orthonormalization of state vectors takes the form

$$\langle\,\xi\,|\,\xi'\,\rangle = \delta(\xi - \xi').$$

If ξ refers to a set of observables $\xi_1, \ldots \xi_f$ with continuous spectra, $\delta(\xi - \xi')$ should be understood as

$$\delta(\xi - \xi') = \delta(\xi_1 - \xi_1')\;\delta(\xi_2 - \xi_2')\;\ldots\;\delta(\xi_f - \xi_f').$$

For example, for the position eigenstates we have

$$\hat{\vec{r}}\,|\,\vec{r}\,'\,\rangle = \vec{r}\,'\,|\,\vec{r}\,'\,\rangle$$

and

$$\langle\,\vec{r}\,|\,\vec{r}\,'\,\rangle = \delta^{(3)}(\vec{r} - \vec{r}\,') = \delta(x - x')\;\delta(y - y')\;\delta(z - z').$$

The basic measurement or "filter" operator takes the form

$$|\,\xi\,\rangle\,\langle\,\xi\,|.$$

In the case of a discrete spectrum (e.g., the z-axis projection of the angular momentum J_z of an atom) the operator

$$\hat{P}_m = |\,m\,\rangle\,\langle\,m\,|$$

corresponds to an experimentally realizable filtration. However, in the case of a continuum spectrum, a single isolated value of ξ cannot be measured. For example, in a linear momentum filter made with a magnetic field and slits, a finite spread Δp is always present due to the finite width of the slit. The same holds for a detector of the position of an electron which has necessarily a finite volume or area. The operator $|\,\xi\,\rangle\,\langle\,\xi\,|$ is only meaningful in the context of an integral over an arbitrary small though finite interval $\Delta\xi$. In fact, $|\,\xi\,\rangle\,\langle\,\xi\,|$ is *not* a projector operator, since $(|\,\xi\,\rangle\,\langle\,\xi\,|)^2 = |\,\xi\,\rangle\,\langle\,\xi\,|\,\xi\,\rangle\,\langle\,\xi\,|$ cannot be defined [$\langle\,\xi\,|\,\xi\,\rangle = \delta(0)$ would be equal to the delta function at the origin].

The projector operator onto a finite interval $\xi - \Delta\xi \le \xi' \le \xi + \Delta\xi$ of values ξ' of the observable $\hat{\xi}$ around ξ is given by

$$\hat{P}_{\xi,\Delta\xi} = \int_{\xi-\Delta\xi}^{\xi+\Delta\xi} |\,\xi'\,\rangle\,\langle\,\xi'\,|\;d\xi'.$$

It is easy to see that $\hat{P}^2_{\xi,\Delta\xi} = \hat{P}_{\xi,\Delta\xi}$ is a proper projector operator.

The probability of measuring a value of the observable $\hat{\xi}$ in the interval $[\xi - \Delta\xi, \xi + \Delta\xi]$ when the system is in the state $|\,\alpha\,\rangle$ is given in analogy to the discrete case by

$$\int_{\xi-\Delta\xi}^{\xi+\Delta\xi} |\langle\,\xi'\,|\,\alpha\,\rangle|^2\, d\xi' \;\xrightarrow{\Delta\xi\to 0}\; 2\,|\langle\,\xi\,|\,\alpha\,\rangle|^2\,\Delta\xi.$$

Our understanding of a continuous spectrum is based on the observation that we continue to measure values of ξ in the interval $[\xi-\Delta\xi, \xi+\Delta\xi]$, even as $\Delta\xi$ becomes arbitrarily small and, moreover, that for very small $\Delta\xi$ the associated probability becomes proportional to $\Delta\xi$.

We may now summarize the basic relationships for continuous spectra. For the sake of comparison we sometimes also give the corresponding relations for discrete spectra, using the notation $\hat{\xi}\,|\,\xi_n\,\rangle = \xi_n\,|\,\xi_n\,\rangle$ for eigenstates and eigenvalues of the discrete spectrum.

The completeness relation is given by

$$\int d\xi\;|\,\xi\,\rangle\,\langle\,\xi\,| = \mathbb{1} \tag{1.3}$$

or $\sum_n |\,\xi_n\,\rangle\,\langle\,\xi_n\,| = \mathbb{1}$ in the discrete case. Thus, an arbitrary state $|\,\alpha\,\rangle$ is expressed in the ξ representation as

$$|\,\alpha\,\rangle = \int d\xi\;|\,\xi\,\rangle\,\langle\,\xi\,|\,\alpha\,\rangle. \tag{1.4}$$

The normalization condition for a ket $|\,\alpha\,\rangle$ reads

$$1 = \langle\,\alpha\,|\,\alpha\,\rangle = \int d\xi\;\langle\,\alpha\,|\,\xi\,\rangle\,\langle\,\xi\,|\,\alpha\,\rangle = \int d\xi\;|\langle\,\xi\,|\,\alpha\,\rangle|^2 \tag{1.5}$$

and the inner product is given by

$$\langle\,\alpha\,|\,\beta\,\rangle = \int d\xi\;\langle\,\alpha\,|\,\xi\,\rangle\,\langle\,\xi\,|\,\beta\,\rangle. \tag{1.6}$$

Notice that Eqs. (1.4)–(1.6) follow immediately from the completeness relation (1.3).

The operators of observables can be written as

$$\hat{\xi} = \int d\xi'\;|\,\xi'\,\rangle\,\xi'\,\langle\,\xi'\,|,$$

which is analogous to the expression $\hat{\xi} = \sum_n |\,\xi_n\,\rangle\,\xi_n\,\langle\,\xi_n\,|$ known for discrete eigenvalues. The matrix elements between eigenstates are given by

$$\langle\,\xi\,|\,\hat{\xi}\,|\,\xi'\,\rangle = \xi\;\delta(\xi - \xi'),$$

while for discrete spectra we have $\langle\, \xi_n \,|\, \hat{\xi} \,|\, \xi_l \,\rangle = \xi_n \, \delta_{nl}$.

The transformation functions between the representations $\{|\, \xi \,\rangle\}$, $\{|\, \zeta \,\rangle\}$ and $\{|\, \eta \,\rangle\}$ satisfy the following relations that can be easily derived using the completeness and orthonormalization conditions:

$$\langle\, \xi \,|\, \zeta \,\rangle = \int d\eta \, \langle\, \xi \,|\, \eta \,\rangle \, \langle\, \eta \,|\, \zeta \,\rangle$$

and

$$\delta\,(\xi - \xi') = \langle\, \xi \,|\, \xi' \,\rangle = \int d\eta \, \langle\, \xi \,|\, \eta \,\rangle \, \langle\, \eta \,|\, \xi' \,\rangle.$$

The most familiar observable with a continuous spectrum is the position. We know that if the electron is in the vector state $|\, \alpha \,\rangle$ the probability of finding it in the interval $[x, x+dx]$ (in one dimension) is $|\langle\, x \,|\, \alpha \,\rangle|^2 \, dx$. Thus, the transformation function $\langle\, x \,|\, \alpha \,\rangle$ is nothing but the usual wave function $\Psi_\alpha(x) = \langle\, x \,|\, \alpha \,\rangle$. Position and momentum representations are particularly important examples involving continuous spectra. They are discussed in the following.

1.17 Position eigenkets and position measurement

The position eigenkets $|\, x' \,\rangle$ satisfy

$$\hat{x} \,|\, x' \,\rangle = x' \,|\, x' \,\rangle.$$

They form a complete set in one dimension (1D). Thus $\int dx \,|\, x \,\rangle \, \langle\, x \,| = \mathbb{1}$ and

$$|\, \alpha \,\rangle = \int_{-\infty}^{+\infty} dx \,|\, x \,\rangle \, \langle\, x \,|\, \alpha \,\rangle \qquad \forall \,|\, \alpha \,\rangle.$$

The probability of measuring a position in the interval $[x, x + \Delta x]$ when the electron is in the state $|\, \alpha \,\rangle$ is given by

$$P(x, x + \Delta x) = \int_x^{x+\Delta x} |\langle\, x' \,|\, \alpha \,\rangle|^2 \, dx' \;\; \xrightarrow{\Delta x \to 0} \;\; |\langle\, x \,|\, \alpha \,\rangle|^2 \, \Delta x.$$

For a normalized state we have

$$\langle\, \alpha \,|\, \alpha \,\rangle = \int_{-\infty}^{+\infty} \langle\, \alpha \,|\, x \,\rangle \, \langle\, x \,|\, \alpha \,\rangle \, dx = \int_{-\infty}^{+\infty} |\langle\, x \,|\, \alpha \,\rangle|^2 \, dx = 1.$$

Therefore, the probability $P(x_1, x_2)$ is normalized to $1 = P(-\infty, +\infty)$.

The previous relations can be immediately generalized to 3 dimensions where $\vec{r} = (x, y, z) = (x_1, x_2, x_3)$:

$$\begin{cases} \hat{x} \,|\, \vec{r}\,' \,\rangle = x' \,|\, \vec{r}\,' \,\rangle \\ \hat{y} \,|\, \vec{r}\,' \,\rangle = y' \,|\, \vec{r}\,' \,\rangle \\ \hat{z} \,|\, \vec{r}\,' \,\rangle = z' \,|\, \vec{r}\,' \,\rangle, \end{cases}$$

$$\int d^3r \, |\, \vec{r} \,\rangle \, \langle \, \vec{r} \,| = \mathbb{1},$$

$$|\, \alpha \,\rangle = \int d^3r \, |\, \vec{r} \,\rangle \, \langle \, \vec{r} \,|\, \alpha \,\rangle,$$

$$P(\vec{r}, d^3r) = |\langle \, \vec{r} \,|\, \alpha \,\rangle|^2 \, d^3r,$$

and

$$[\hat{x}_i, \hat{x}_j] = 0 \qquad \forall \; i, \, j = 1, 2, \text{ and } 3.$$

The orthonormality condition for the position eigenstates reads

$$\langle \, x \,|\, x' \,\rangle = \delta \, (x - x') = \langle \, x' \,|\, x \,\rangle$$

in 1D, or

$$\langle \, \vec{r} \,|\, \vec{r}\,' \,\rangle = \delta^{(3)}(\vec{r} - \vec{r}\,') = \delta \, (x - x') \, \delta \, (y - y') \, \delta \, (z - z')$$

in 3D. The expansion coefficient $\langle \, x \,|\, \alpha \,\rangle$ entering the linear combination or superposition

$$|\, \alpha \,\rangle = \int dx \, |\, x \,\rangle \, \langle \, x \,|\, \alpha \,\rangle$$

is a complex function of x that defines the state $|\, \alpha \,\rangle$ completely. Moreover, it holds that

$$|\langle \, x \,|\, \alpha \,\rangle|^2 \, dx$$

is the probability of finding the particle in an interval of size dx around x (i.e., $x' \in [x, x + dx]$). Therefore,

$$\Psi_\alpha(x) = \langle \, x \,|\, \alpha \,\rangle$$

is the *wave function* associated to the state $|\, \alpha \,\rangle$, as introduced in the traditional (historical) formulation of quantum wave mechanics. In other words, the wave function is the coefficient $\langle \, x \,|\, \alpha \,\rangle$ defining the projection of the state $|\, \alpha \,\rangle$ on the position eigenket $|\, x \,\rangle$. This is *not* an additional assumption. It derives simply from the general postulate that the inner product

$$|\langle \, a' \,|\, \alpha \,\rangle|^2$$

represents the probability of measuring the value a' of any observable A when the particle is in the state $|\, \alpha \,\rangle$ ($\langle \, \alpha \,|\, \alpha \,\rangle = 1$).

The postulates of wave-mechanics apply to one-particle systems and can be easily generalized to any number of particles. However, Dirac's formulation in terms of kets and

bras is more general, since it is valid for any kind of state or system $|\,\alpha\,\rangle$, even if the total number of particles is not well defined as in the relativistic case.

All usual expressions involving the wave function $\Psi_\alpha(x)$ can be *derived* straightforwardly. A few examples follow:

Inner product:

$$\langle\,\beta\,|\,\alpha\,\rangle = \int dx\,\langle\,\beta\,|\,x\,\rangle\,\langle\,x\,|\,\alpha\,\rangle = \int dx\;\Psi^*_\beta(x)\;\Psi_\alpha(x).$$

Remember that we have shown that $\langle\,\beta\,|\,\alpha\,\rangle$ is *independent of the representation.*

Linear expansion of the wave function:

The coefficients $c_a = \langle\,a\,|\,\alpha\,\rangle$ of the expansion

$$|\,\alpha\,\rangle = \sum_{a'} |\,a\,\rangle\,\langle\,a\,|\,\alpha\,\rangle = \sum_{a'} c_a\,|\,a\,\rangle$$

satisfy

$$\langle\,x\,|\,\alpha\,\rangle = \sum_a \langle\,x\,|\,a\,\rangle\,\langle\,a\,|\,\alpha\,\rangle = \sum_a c_a\,\langle\,x\,|\,a\,\rangle.$$

Denoting by

$$u_a(x) = \langle\,x\,|\,a\,\rangle$$

the eigenfunctions of the operator $\hat{A}$ (i.e., the wave function associated to the eigenstate $|\,a\,\rangle$), we obtain

$$\Psi_\alpha(x) = \sum c_a\,u_a(x).$$

This is the usual expansion of $\Psi_\alpha(x)$ in terms of the complete set of eigenfunctions $u_a(x)$ (superposition principle).

Matrix elements of operators:

Let us write $\langle\,\alpha\,|\,\hat{A}\,|\,\beta\,\rangle$ in terms of $\Psi_\alpha(x)$ and $\Psi_\beta(x)$:

$$\begin{aligned}\langle\,\alpha\,|\,\hat{A}\,|\,\beta\,\rangle &= \int dx \int dx'\,\langle\,\alpha\,|\,x'\,\rangle\,\langle\,x'\,|\,\hat{A}\,|\,x\,\rangle\,\langle\,x\,|\,\beta\,\rangle \\ &= \int dx \int dx'\;\Psi^*_\alpha(x')\,\langle\,x'\,|\,\hat{A}\,|\,x\,\rangle\;\Psi_\beta(x).\end{aligned}$$

Thus, in order to calculate $\langle\,\alpha\,|\,\hat{A}\,|\,\beta\,\rangle$, we need to know the matrix elements $\langle\,x'\,|\,\hat{A}\,|\,x\,\rangle$ of the operator $\hat{A}$ in position-space or "coordinate" representation.

$\hat{A}$ can be a local operator $\langle\, x \,|\, \hat{A} \,|\, x' \,\rangle = \delta\,(x - x')\,\langle\, x \,|\, \hat{A} \,|\, x \,\rangle$ or in more rare cases a non-local one. In the former case the expression for $\langle\, \alpha \,|\, \hat{A} \,|\, \beta \,\rangle$ simplifies considerably. For example, for the interaction $\hat{V}$ of the electron with an external potential $v(x)$ we have

$$\langle\, x \,|\, \hat{V} \,|\, x' \,\rangle = \delta\,(x - x')\, v(x),$$

and we recover the well-known expression

$$\langle\, \alpha \,|\, \hat{V} \,|\, \beta \,\rangle = \int dx\; \Psi^*_\alpha(x)\; v(x)\; \Psi_\beta(x).$$

To compute the matrix elements of the momentum operator $\hat{p}$ between arbitrary states $|\, \alpha \,\rangle$ and $|\, \beta \,\rangle$ we may borrow for the moment the expression of the momentum operator known from wave mechanics:

$$\hat{p}\, \Psi_\alpha(x) = -i\,\hbar\, \frac{\partial}{\partial x}\; \Psi_\alpha(x). \tag{1.7}$$

Later on we shall derive this relation by defining $\hat{p}$ as the generator of infinitesimal translations. The left-hand side is the wave function of the state $\hat{p}\,|\, \alpha \,\rangle$ and in the right-hand side we find $\Psi_\alpha(x) = \langle\, x \,|\, \alpha \,\rangle$, the wave function of the state $|\, \alpha \,\rangle$. Thus, in ket-bra notation we can write Eq. (1.7) as

$$\langle\, x \,|\, \hat{p} \,|\, \alpha \,\rangle = -i\,\hbar\, \frac{\partial}{\partial x}\; \langle\, x \,|\, \alpha \,\rangle. \tag{1.8}$$

Note that the derivative $\dfrac{\partial}{\partial x}$ acts on x, which is the position defining the bra $\langle\, x \,|$. Since this holds for all $|\, \alpha \,\rangle$ we have

$$\langle\, x \,|\, \hat{p} = -i\,\hbar\, \frac{\partial}{\partial x}\; \langle\, x \,|.$$

Hermitic conjugation yields

$$\hat{p}^\dagger \,|\, x \,\rangle = \hat{p} \,|\, x \,\rangle = i\,\hbar\, \frac{\partial}{\partial x}\; |\, x \,\rangle,$$

which gives the result of applying $\hat{p}$ to the ket of defined position $|\, x \,\rangle$. Taking $|\, \alpha \,\rangle = |\, x' \,\rangle$ in Eq. (1.8) and using that $\langle\, x \,|\, x' \,\rangle = \delta\,(x - x')$ we have

$$\langle\, x \,|\, \hat{p} \,|\, x' \,\rangle = -i\,\hbar\, \frac{\partial}{\partial x}\; \delta\,(x - x').$$

The matrix element of $\hat{p}$ between arbitrary states is given by

$$\begin{aligned} \langle\, \beta \,|\, \hat{p} \,|\, \alpha \,\rangle &= \int dx\; \langle\, \beta \,|\, x \,\rangle\, \langle\, x \,|\, \hat{p} \,|\, \alpha \,\rangle \\ &= -i\,\hbar \int dx\; \Psi^*_\beta(x)\; \frac{\partial\, \Psi_\alpha(x)}{\partial x}, \end{aligned}$$

where we have used the completeness relation (1.3) and Eq. (1.8).

1.18 Momentum representation

Besides the coordinates of the particle, another very important complete set of observables is given by the momentum. We consider a basis of states with defined momentum $\hat{p}$:

$$\hat{p}\,|\,p'\,\rangle = p'\,|\,p'\,\rangle$$

with

$$\langle\,p\,|\,p'\,\rangle = \delta\,(p-p').$$

For simplicity we restrict ourselves here to 1D. The completeness relation reads $\int dp\,|\,p\,\rangle\,\langle\,p\,|$ and thus an arbitrary state can be written as

$$|\,\alpha\,\rangle = \int dp\,|\,p\,\rangle\,\langle\,p\,|\,\alpha\,\rangle. \tag{1.9}$$

As usual $|\langle\,p\,|\,\alpha\,\rangle|^2\,dp$ is proportional to the probability of measuring a value of the momentum in the interval $[p, p+dp]$. The function

$$\phi_\alpha(p) = \langle\,p\,|\,\alpha\,\rangle$$

is known as the momentum or *momentum-space wave function* of the state $|\,\alpha\,\rangle$. Assuming that $|\,\alpha\,\rangle$ is normalized we have

$$\langle\,\alpha\,|\,\alpha\,\rangle = \int dp\,\langle\,\alpha\,|\,p\,\rangle\,\langle\,p\,|\,\alpha\,\rangle = \int |\phi_\alpha(p)|^2\,dp = 1.$$

In order to connect the two equivalent representations $\Psi_\alpha(x)$ and $\phi_\alpha(p)$ we need the transformation function $\langle\,x\,|\,p\,\rangle$ from the x-representation to the p-representation. Eq. (1.9) implies

$$\Psi_\alpha(x) = \langle\,x\,|\,\alpha\,\rangle = \int dp\,\langle\,x\,|\,p\,\rangle\,\phi_\alpha(p). \tag{1.10}$$

The transformation function $\langle\,x\,|\,p\,\rangle$ depends on both x and p. If we regard it as a function of x for fixed p, $\langle\,x\,|\,p\,\rangle$ is nothing but the coordinate wave function of the eigenstate $|\,p\,\rangle$ of the momentum operator $\hat{p}$. Of course, $\langle\,p\,|\,x\,\rangle = \langle\,x\,|\,p\,\rangle^*$ now regarded as a function of p for fixed x is the momentum wave function of the eigenstate $|\,x\,\rangle$ of the position operator $\hat{x}$.

Let us obtain $\langle\,x\,|\,p\,\rangle$ by using the properties of $\hat{p}$. From Eq. (1.8) we know that

$$\langle\,x\,|\,\hat{p}\,|\,\alpha\,\rangle = -i\,\hbar\,\frac{\partial}{\partial x}\,\langle\,x\,|\,\alpha\,\rangle.$$

Taking $|\,\alpha\,\rangle = |\,p\,\rangle$ we have

$$\langle\,x\,|\,\hat{p}\,|\,p\,\rangle = -i\,\hbar\,\frac{\partial}{\partial x}\,\langle\,x\,|\,p\,\rangle$$

and

$$\frac{i}{\hbar}\, p\, \langle\, x\,|\, p\,\rangle = \frac{\partial}{\partial x}\, \langle\, x\,|\, p\,\rangle.$$

Straightforward integration yields

$$\langle\, x\,|\, p\,\rangle = N\, e^{\frac{i}{\hbar}\, p\, x},$$

where N is a constant to be determined from the normalization condition

$$\langle\, p\,|\, p'\,\rangle = \delta\,(p - p'). \tag{1.11}$$

On the one side we have

$$\langle\, p\,|\, p'\,\rangle = \int_{-\infty}^{+\infty} dx\, \langle\, p\,|\, x\,\rangle\, \langle\, x\,|\, p'\,\rangle = N^2 \int dx\; e^{-i\, \frac{p\, x}{\hbar}}\; e^{i\, \frac{p'\, x}{\hbar}} = N^2 \int_{-\infty}^{+\infty} dx\; e^{i\, \frac{(p - p')\, x}{\hbar}}. \tag{1.12}$$

On the other side the usual Fourier-integral representation of the δ-function reads

$$\delta(p - p') = \frac{1}{2\,\pi} \int_{-\infty}^{+\infty} d\lambda\; e^{i(p - p')\,\lambda} = \frac{1}{2\,\pi\,\hbar} \int_{-\infty}^{+\infty} dx\; e^{\frac{i(p - p')\, x}{\hbar}}, \tag{1.13}$$

where we have replaced $\lambda = x/\hbar$. Replacing Eqs. (1.12) and (1.13) in Eq. (1.11) one obtains $N^2 = 1/(2\pi\hbar)$ and

$$\langle\, x\,|\, p\,\rangle = \frac{1}{\sqrt{2\pi\hbar}}\, e^{\frac{i}{\hbar}\, p\, x}.$$

From Eq. (1.10) we finally have

$$\Psi_\alpha(x) = \frac{1}{\sqrt{2\,\pi\,\hbar}} \int dp\; e^{\frac{i}{\hbar}\, p\, x}\; \phi_\alpha(p)$$

and

$$\phi_\alpha(p) = \frac{1}{\sqrt{2\,\pi\,\hbar}} \int dx\; e^{-\frac{i}{\hbar}\, p\, x}\; \Psi_\alpha(x).$$

The coordinate and momentum wave functions are the Fourier transform of each other.

The formalism can be immediately generalized to 3 dimensions. The position and momentum eigenstates are defined by

$$\hat{\vec{r}}\,|\,\vec{r}\,'\,\rangle = \vec{r}\,'\,|\,\vec{r}\,'\,\rangle \quad \text{with} \quad \langle\,\vec{r}\,|\,\vec{r}\,'\,\rangle = \delta(\vec{r} - \vec{r}\,')$$

and

$$\hat{\vec{p}}\,|\,\vec{p}\,'\,\rangle = \vec{p}\,'\,|\,\vec{p}\,'\,\rangle \quad \text{with} \quad \langle\,\vec{p}\,|\,\vec{p}\,'\,\rangle = \delta(\vec{p} - \vec{p}\,').$$

The completeness relation reads

$$\mathbb{1} = \int d^3r \, |\,\vec{r}\,\rangle \, \langle\,\vec{r}\,| = \int d^3p \, |\,\vec{p}\,\rangle \, \langle\,\vec{p}\,|,$$

which implies

$$|\,\alpha\,\rangle = \int d^3r \, |\,\vec{r}\,\rangle \, \langle\,\vec{r}\,|\,\alpha\,\rangle = \int d^3p \, |\,\vec{p}\,\rangle \, \langle\,\vec{p}\,|\,\alpha\,\rangle.$$

The matrix elements of $\hat{\vec{p}}$ are given by

$$\langle\,\beta\,|\,\hat{\vec{p}}\,|\,\alpha\,\rangle = -i\,\hbar \int d^3r \; \Psi^*_\beta(\vec{r}) \; \vec{\nabla} \; \Psi_\alpha(\vec{r})$$

and the momentum eigenfunctions read

$$\langle\,\vec{r}\,|\,\vec{p}\,\rangle = \frac{1}{(2\,\pi\,\hbar)^{3/2}} \; e^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{r}}.$$

2 Measurements in quantum mechanics

The measurement processes in classical and quantum physics show strong differences as well as some analogies:

1) In both quantum and classical mechanics the result of a measurement is a macroscopic quantity that a human can read.

2) In classical mechanics a measurement can be made in such a soft way that the state of the system is *not* necessarily altered by the measurement process. Consider, for example, the negligible effect of photons on a train or even dust particles, whose reflection allows us to determine their position.

3) In contrast, in quantum mechanics, a measurement is the outcome of an interaction between the considered microscopic system and a macroscopic classical apparatus which results, in most cases, in a change of state of the quantum system. This is something intrinsic to the microscopic world and cannot be overcome by refinements of the measurement process.

P. M. Dirac summarizes the effect of a measurement on a quantum system as follows: "A measurement always causes a system to jump into an eigenstate of the dynamical variable that is being measured". After some thought, this sounds perfectly reasonable. Let us recall that an eigenstate is a state for which a measurement yields with certainty always the same result. Now consider a second measurement immediately after the first one. Since the state of a system cannot change at an arbitrary large speed, it is clear that the second measurement at a time $t' \to t$ should give with certainty the same result as the first one. After all, classical systems also behave this way: The positions of a train at two arbitrarily close times are the same. We conclude that an arbitrarily precise (also known as sharp) measurement of a given dynamical variable projects the state of the system onto an eigenstate of this dynamical variable with eigenvalue equal to the result of the measurement. However, what does this "jumping into an eigenstate" mean? How does it actually happen?

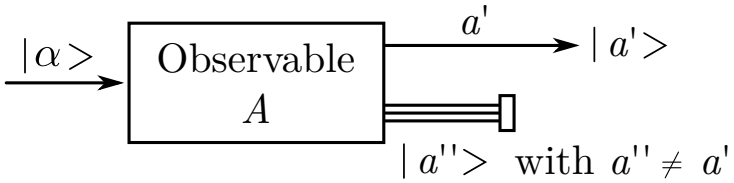


Figure 1: Arbitrarily sharp or selective measurement after which the the quantum state is an eigenstate of the observable A with the just measured eigenvalue a'.

2.1 The Stern-Gerlach measurement

In order to illustrate the kind of physical processes behind the interaction between a quantum system and a classical measuring apparatus, we consider a Stern-Gerlach measurement of the projection of the magnetic moment or angular momentum of an atom.

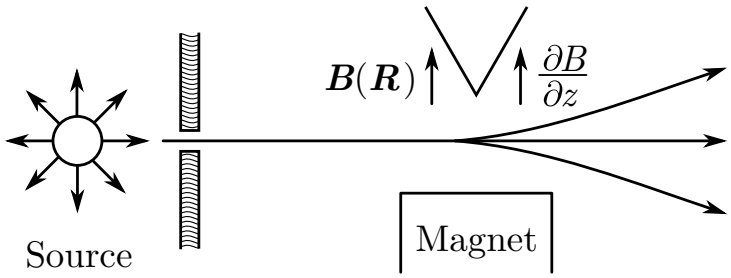


Figure 2: Schematic illustration of the Stern-Gerlach experiment.

The energy of the atom when the center of mass is in the position $\vec{R}$ and the electronic state is $|\, nlm \,\rangle$ is given by

$$E_{nlm}(\vec{R}) = E^0_{nl} - \mu_0\, m\, B(\vec{R}) \quad \Rightarrow \quad -\frac{\partial E_m}{\partial \vec{R}} = \mu_0\, m\, \frac{\partial B}{\partial \vec{R}}.$$

The wave function of the atom can be written in the form

$$\Psi = \sum_m c_m\, u_m(\vec{R}, t)\; \Phi_m(\vec{r}, \vec{R}),$$

which is a linear combination of products of a wave packet $u_m(\vec{R}, t)$ of the center-of-mass coordinate $\vec{R}$ and a wave function $\Phi_n(\vec{r}, \vec{R})$ of the internal degrees of freedom $\vec{r}$ of the atom, i.e., describing the motion of the electron relative to the nucleus. In the adiabatic approximation the wave packet follows the classical trajectory given by the Ehrenfest theorem:

$$u_m(\vec{R}, t) = u_m\Big(\vec{R} - \vec{R}_0(t)\Big)\,, \tag{2.1}$$

where $M\ddot{\vec{R}}_0 = -\dfrac{\partial E_m}{\partial \vec{R}}$. Notice that E_m depends on the internal state m of the atom:

$$E_m(\vec{R}) = \langle\, u_m(\vec{R})\, \Phi_m(\vec{R}, \vec{r}\,) \,|\, \hat{H} \,|\, u_m(\vec{R})\, \Phi_m(\vec{r}, \vec{R}) \,\rangle = E^0_{nl} - \mu_0\, m\, B(\vec{R}). \tag{2.2}$$

For the Stern-Gerlach experiment we prepare the system in a collimated wave packet $u^0(R)$ and in some unknown state $\sum_m c_m\, \Phi_m(r)$ of the electrons in the atom.

The following three different situations need to be considered:
I. Before the magnet:

$$\Psi_0(t) = \underbrace{u^0(\vec{R}, t)}_{\substack{\text{The same} \\ \text{for all } m}} \underbrace{\sum_m c_m\, \Phi_m(\vec{r}\,)}_{\substack{\text{Independent of } \vec{R} \\ \text{since no field}}} \tag{2.3}$$

For simplicity, the time dependence of $\Phi_m(\vec{r})$ corresponding to the obvious phase factor $e^{-\frac{i}{\hbar}E_m\,t}$ is not explicitly indicated here. This can be regarded as a time dependence of the coefficients c_m.

II. Within the magnet:

Each part of the wave function, corresponding to the different states $\phi_m(\vec{r})$ of the electrons in the atom, evolves in time as

$$u^0(\vec{R},t)\,\Phi_m(\vec{r}) \to u_m(\vec{R},t)\,\Phi_m(\vec{R},\vec{r},t) \tag{2.4}$$

according, either to the Newton equation for $\vec{R}_0$ and Eq. (2.1), or to the Schrödinger equation with the corresponding eigenenergy given by Eq. (2.2). Since the equation of time evolution is *linear*, the time dependence of the complete wave function

$$\Psi_0(t) \to \Psi(t) = \sum_m c_m\,u_m(\vec{R},t)\;\Phi_m(\vec{R},\vec{r})$$

is the superposition of the time dependences given by Eq. (2.4).

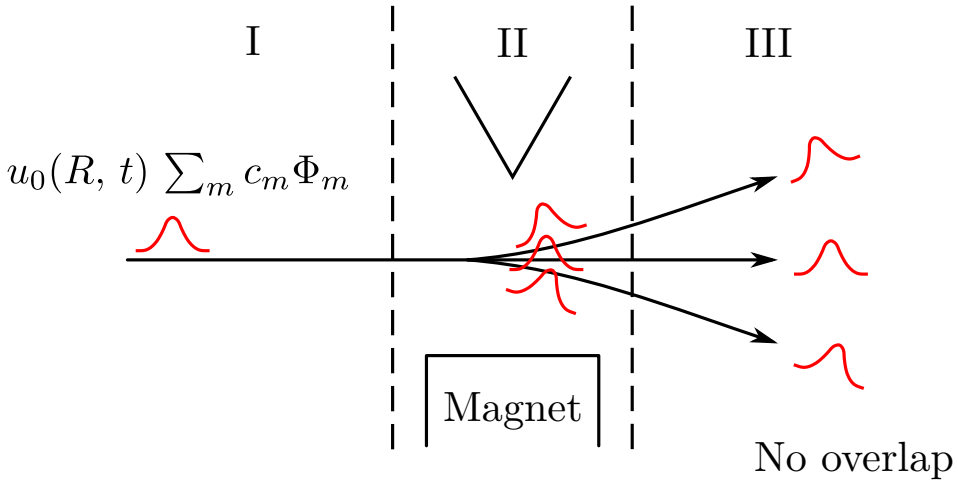


Figure 3: Evolution of the wave function of the center-of-mass coordinate before, within and after the magnet.

III. After the magnet:

The electronic wave functions $\phi_m(\vec{r})$ no longer depend on the center of mass coordinate $\vec{R}$, since there is no field:

$$\Psi_\infty(\vec{R},\vec{r},t) = \sum_m c_m\,u_m(\vec{R},t)\;\Phi_m(\vec{r})$$

with $\sum_m c_m^2 = 1$. Notice that after the measurement $\Psi_\infty(t)$ correlates the state of the atom m with the center of mass wave function $u_m(\vec{R},t)$.

A few remarks are due:

1) The Stern-Gerlach *measurement* modifies the wave function from $\Psi_0(t)$ to $\Psi_\infty(t)$. This is a consequence of the interaction of the atom with the magnetic field, i.e., with the measuring apparatus.

2) Ψ_∞ remains a *pure state* without altering the relative phases of the various coefficients c_m. This is a simple superposition of the states $u_m(\vec{R},t)\,\Phi_m(\vec{r})$.

3) We can now compute the probability P_m of finding the atom, i.e., its center of mass coordinate $\vec{R}$, in the volume V_m after the passage through the magnet:

$$\begin{aligned} P_m &= \int_{V_m} d^3R \int d^3r\ |\Psi_\infty(t)|^2 \\ &= \int_{V_m} d^3R \int d^3r \sum_{m'm''} c^*_{m'}\, u^*_{m'}(\vec{R},t)\,\Phi^*_{m'}(\vec{r})\ c_{m''}\, u_{m''}(\vec{R},t)\,\Phi_{m''}(\vec{r}). \end{aligned}$$

Taking into account that the electronic wave functions for different m are orthogonal to each other, i.e.,

$$\int d^3r\ \Phi^*_{m'}(\vec{r})\,\Phi_{m''}(\vec{r}) = \delta_{m'm''},$$

we obtain

$$P_m = \sum_{m'} |c_{m'}|^2 \underbrace{\int_{V_m} d^3R\ |u_{m'}(\vec{R})|^2}_{\delta_{m'm}} = |c_m|^2,$$

where we have used that $u_{m'}(\vec{R})$ is zero inside V_m for $m' \neq m$, or equivalently, that $u_m(\vec{R})$ vanishes outside V_m. In other words, we take advantage of the fact that, as a result of the Ehrenfest force on the center of mass and of the subsequent uniform rectilinear motion, the wave packets $u_{m'}(\vec{R})$ of the center of mass do not overlap. Therefore, finding the atom in the volume V_m implies that its magnetic state is m.

In conclusion, we measure the internal magnetic state m of the atom by measuring the position of the center of mass $\vec{R}$ of the atom. In order to derive this result the following fundamental properties have been used:

1) The superposition principle, i.e., the linearity of the Schrödinger equation.

2) The fundamental postulate that $\int d^3r\ |\Psi_m(\vec{r},\vec{R})|^2$ gives the probability distribution of the center of mass coordinate $\vec{R}$.

3) A measurement must produce a situation where the occurrence of a macroscopic event unambiguously determines the microscopic physical property of the quantum system. In our case, this one-to-one relation reads "the atom is in volume $V_m \Leftrightarrow$ the magnetic moment of the atom is $\mu_0\, m$".

If one now analyses the state of the atom $\Phi(\vec{r})$ after having detected its position inside V_m, one finds that the atom is in the state $\Phi_m(r)$. In fact, the probability distribution for the relative coordinate $\vec{r}$ of the electron after the passage through the magnet is given by

$$P_\infty(\vec{r}) = \int d^3R\, |\Psi_\infty(\vec{r},\vec{R})|^2 = \int d^3R\, |\textstyle\sum_m c_m\, u_m(\vec{R})\, \Phi_m(\vec{r})|^2$$
$$= \sum_m |c_m|^2\, |\Phi_m(\vec{r})|^2, \tag{2.5}$$

where we have used that $\int u_m^*(\vec{R})\, u_{m'}(\vec{R})\, d^3R = \delta_{mm'}$, since there is no overlap between the different $u_m(\vec{R})$ or between the different V_m. Notice that after the measurement there is a separation of the different internal states m, and that the relative phases between the different c_m do not enter in $P_\infty(\vec{r})$.

The probability distribution $P_\infty(\vec{r})$ [Eq. (2.5), after passing through the apparatus] should be contrasted with the probability distribution for $\vec{r}$ before entering the apparatus. From Eq. (2.3) we have

$$\Psi_0(\vec{r},\vec{R},t) = u_0(\vec{R},t) \sum_m c_m\, \Phi_m(\vec{r})$$
$$\Rightarrow \quad P_0(\vec{r}) = \int d^3R\, |\Psi_0(\vec{r},\vec{R},t)|^2 = |\textstyle\sum_m c_m\, \Phi_m(\vec{r})|^2.$$

The measurement alters the state from Ψ_0 to Ψ_∞. In the original state Ψ_0 the relative phases of the coefficients c_m matter. This is not the case after the measurement, since the apparatus has separated the different m components of the wave function in disjoint regions in space. It should be however noted that if one would not measure the position of the center of mass, one could in principle remix the different $c_m\, u_m(\vec{R})\, \Phi_m(\vec{r},\vec{R})$ and in this case the relative phases *would indeed matter* again. However, as long as the u_m do not overlap, it is impossible to see the difference. Only $|c_m|^2$ appears on the result of any measurement.

In quantum mechanics measuring means preparing the system in a specific state or in an ensemble of specific states if the measurement is not extremely precise (e.g., in a continuum spectrum). A measurement can be regarded as a filtering process. If one has a state with defined momentum p_x, this means that one has "measured p_x". Moreover, if one measures the position of an electron at x_0, its state changes from $\Psi(x)$ to $\delta(x - x_0)$. The position of the electron is thereby defined.

The statistical character of the measurement process is intrinsic to the quantum mechanical nature of the microscopic world, and cannot be removed by a *more detailed theory* involving further variables or more precise initial conditions. This behavior is conceptually very different from what we encounter in classical physics. In the macroscopic limit a physical system has well defined values of its dynamical variables (position, linear or angular momentum, energy, etc.) irrespectively of the fact that a measurement is performed or not. Moreover, these values need not be altered by the measurement process.

The probabilistic random outcome of a measurement in quantum mechanics should not be interpreted as a lack of determinism. Quantum dynamics is indeed 100% deterministic, since the state of the system is not defined by the classical dynamical variables, but by the wave function or vector state. A knowledge of the latter at any given time allows one to predict univocally the state of the system and the results of all measurements at any future time.

2.2 Compatible versus incompatible observables

(Bild)

Consider two compatible observable A and B ($[\hat{A}, \hat{B}] = 0$). In this case a complete basis set of the form $|\, a_j, b^i \,\rangle$ exists in which both A and B have definite values. Assuming that the spectrum of A is non degenerate, the measurement of A followed by a measurement of B can be illustrated as follows:

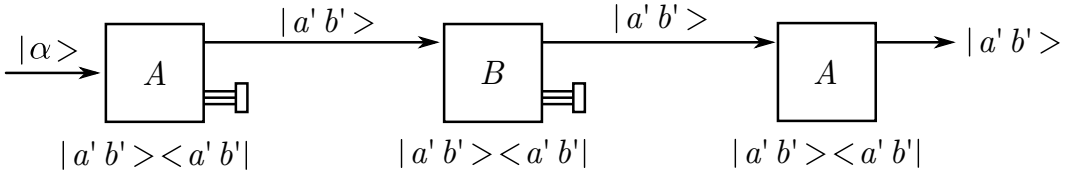


Figure 4: Schematic illustration of the subsequent measurements of two compatible non-degenerate observables A and B.

In general, however, we will find more than one mutually orthogonal states $|\, a', b^i \,\rangle$ having the well-defined value a' of A. In this case the eigenvalue a' is degenerate. If there are degeneracies the first measurement corresponds to the projector

$$\sum_i |\, a', b^i \,\rangle \langle\, a', b^i \,|.$$

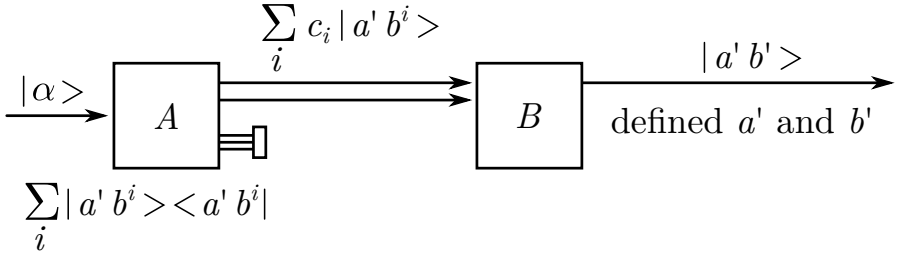


Figure 5: Schematic illustration of the subsequent measurements of two compatible observables A and B with degenerate eigenvalues.

In contrast, when the observable A and B are incompatible (i.e., $[\hat{A}, \hat{B}] \neq 0$) there is no common complete basis of eigenstates of $\hat{A}$ and $\hat{B}$. As an example of incompati-

ble observables let is consider the spin-1/2 operator $\hat{\vec{s}} = (\hat{s}_x, \hat{s}_y, \hat{s}_z)$ for which we have $[\hat{s}_x, \hat{s}_y] = i\hbar\hat{s}_z$ and the cyclic permutations). In this case the state with defined spin projection along the axes x, y and z (e.g., $|\, x, + \,\rangle$ for an up-spin state along x) are the superposition of both up and down states along the other directions:

$$|\, x, \pm \,\rangle = \frac{1}{\sqrt{2}} \left(|\, z, + \,\rangle \pm |\, z, - \,\rangle \right)$$
$$|\, y, \pm \,\rangle = \frac{1}{\sqrt{2}} \left(|\, z, + \,\rangle \pm i\, |\, z, - \,\rangle \right)$$
$$|\, z, \pm \,\rangle = \frac{1}{\sqrt{2}} \left(|\, x, + \,\rangle \pm |\, x, - \,\rangle \right).$$

Let us consider the case of succesive measurements of the observables A, B and C which can be illustrated as follows:

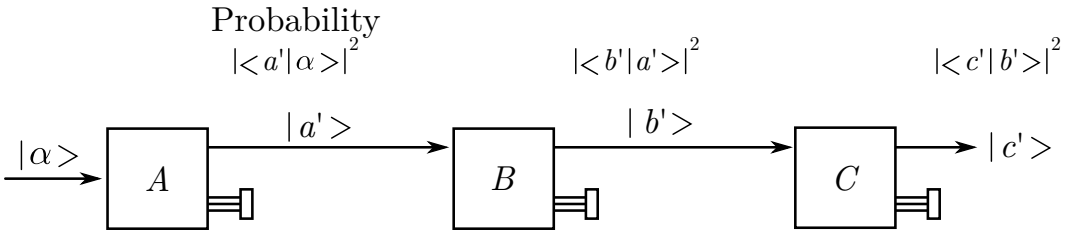


Figure 6: Schematic illustration of the subsequent measurements of the incompatible observables A, B and C.

The first measurement of A has the role of preparing the system in a well-defined state $|\, a' \,\rangle$ with eigenvalue a' for further investigation. The idea is to anylyze how the result of a measurement of the observable C will depend or not on the fact that an intermediate measure of B is performed between the preparation (measurement of A to yield $|\, a' \,\rangle$) and the final mesurement of C.

$$\begin{aligned} \text{Prob}\,(a', b', c') &= |\langle\, b' \,|\, a' \,\rangle|^2 \, |\langle\, c' \,|\, b' \,\rangle|^2 \quad \leftarrow \quad \text{Product of probabilities} \\ &= p(a', b') \quad p(b', c'). \end{aligned}$$

Now suppose we are interested in the probability of obtaining the result c', starting from a', irrespectively of the result of the measurement B. To this aim we perform the following steps:

1) Filter $|\, a' \,\rangle$
2) Measure b'
3) Measure c'
4) Sum over all values of b'.

This is given by

$$P(a', B, c') = \sum_{b'} |\langle\, b' \,|\, a' \,\rangle|^2 \, |\langle\, c' \,|\, b' \,\rangle|^2$$

$$= \sum_{b'} \langle\, a' \,|\, b' \,\rangle \, \langle\, b' \,|\, c' \,\rangle \, \langle\, c' \,|\, b' \,\rangle \, \langle\, b' \,|\, a' \,\rangle. \tag{2.6}$$

Let us now assume that we *do not measure* B at all. For example, we could switch off the B-field. What do we expect for $P(a', c')$?

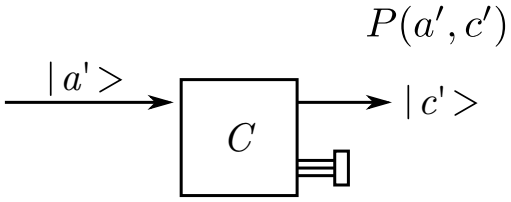


Figure 7: Schematic illustration of the measurements of observable C on a state with well-defined A.

$$P(a', c') = |\langle\, c' \,|\, a' \,\rangle|^2 = \langle\, a' \,|\, c' \,\rangle \, \langle\, c' \,|\, a' \,\rangle$$

$$= \sum_{b'b''} \langle\, a' \,|\, b' \,\rangle \, \langle\, b' \,|\, c' \,\rangle \, \langle\, c' \,|\, b'' \,\rangle \, \langle\, b'' \,|\, a' \,\rangle.$$

Is this the same as $P(a', B, c')$? The answer is clearly no. Something has been altered by measuring b' even if we disregard the result of this measurement. The fact that we can always regard $|\, a' \,\rangle$ as a superposition

$$|\, a' \,\rangle = \sum_{b'} |\, b' \,\rangle \, \langle\, b' \,|\, a' \,\rangle$$

of $|\, b' \,\rangle$ states makes no difference. In the first case we define, "filter" or ascertain the value of B, even if do not care about the outcome of the measurement of B, while in the second case we do not alter the state in any way, we do not perturb $|\, a' \,\rangle$, we just imagine it as a superposition of $|\, b' \,\rangle$ states. In the second case, recording the probability or "measuring" changes the outcome of the second measurement since measuring means *projecting* (Dirac).

One may ask oneself, when does the intermediate measuement of B not alter the result of measuring C, i.e., when do we have $P(a', B, c') = P(a', c')$. For this to hold we need two conditions. First A and B must be of course compatible (i.e., $[\hat{A}, \hat{B}] = 0$) and in addition the spectrum of A must be nondegenerate. Only in this case the measurement of B does not alter the incoming state $|\, a' \,\rangle$ except for an irrelevant phase factor.[3] Another

[3] As a simple example showing that $[\hat{A}, \hat{B}] = 0$ alone does not suffice consider $\hat{A} = \hat{s}^2$, $\hat{B} = \hat{s}_z$ and $\hat{C} = \hat{s}_x$.

possibility is that all three observables are compatible (i.e., $[\hat{A}, \hat{B}] = [\hat{A}, \hat{C}] = [\hat{B}, \hat{C}] = 0$). In any case it is clear that this behavior, namely, that a measurement is an interaction and that in general it alters the outcome of further measurements, constitutes an inherently quantum mechanical phenomenon which is characteristic of incompatible observables.

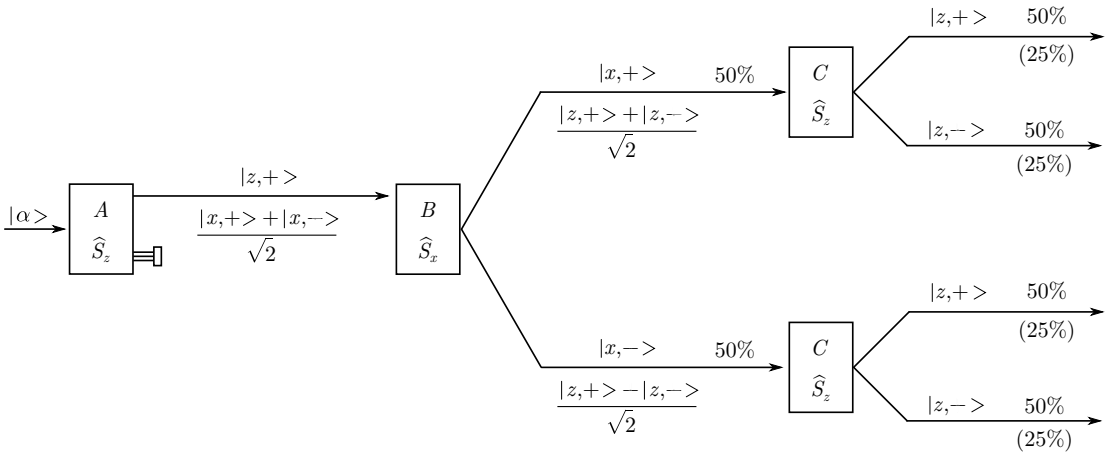
Example:
Let us consider the case where the first preparation mesurement of our experiment is $\hat{A} = \hat{s}_z$ and the initial value of the observable is $a' = 1/2$, or equivalently, for the sake of clarity, $a' = +$. We therefore have $|\, a' \,\rangle = |\, z, + \,\rangle$. Let the intermediate measuement be $\hat{B} = \hat{s}_x$ and the third and final mesurement be again $\hat{C} = \hat{s}_z$. Knowing that

$$|\, a' \,\rangle = |\, z, + \,\rangle = \frac{1}{\sqrt{2}} \left(|\, x, + \,\rangle + |\, x, - \,\rangle \right)$$

and that

$$|\, x, \pm \,\rangle = \frac{1}{\sqrt{2}} \left(|\, z, + \,\rangle \pm |\, x, - \,\rangle \right)$$

it is easy to see that measuring B and summing over all possible outcomes $+$ and $-$ one obtains $P(a, B, c) = P(z+, B, z+) = 1/2$, whereas without measuring B we obviously have $P(a, c) = P(z+, z+) = 1$. This result is illustrated in the following picture.



3 Symmetry in Quantum Mechanics

3.1 Introduction

In this chapter we exploit systematically the various symmetries of quantum mechanical (QM) systems in order to derive the observables associated to the various transformations and their properties. In isolated systems these observables are conserved quantities. Consequently, the stationary states can be classified according to the values of the corresponding covariant property. In these cases, symmetry arguments are certainly the most powerful from a practical point of view. However, also systems that are not isolated, or which lack of symmetry, must have these symmetry related observables, despite the fact that they are not good quantum numbers. One of the goals of this chapter is to derive the physical properties associated to symmetry transformations.

The general validity of symmetry arguments certainly is a great advantage, but it also implies some inconveniences. We will define, for example, the Hamilton operator and the energy of a system from the transformation properties of QM states under time shift, without being at all concerned about the number of particles, the nature of the interactions, or any other system specific details. Furthermore, we shall derive the concepts of linear and angular momentum without needing to request that the system possess translational or rotational symmetry. The underlying formalism, concepts and conclusions are therefore applicable to a wide variety of situations, ranging from atoms, molecules, solids or even the radiation field. However, this generality precludes us from deriving explicit forms for the various operators in terms of the dynamical variables. The situation is pretty much the same as in classical physics, where energy conservation in an isolated system does not give us any clue on how the Hamiltonian should look like in terms of the generalized coordinates and momenta. In the present QM II course we will rely on the background of QM I and on the correspondence principle in order to provide the link to the various applications and to the physical properties.

Besides the practical usefulness of symmetry arguments (transformation laws, conserved quantities, degeneracies, selection rules, etc.) the purpose of this chapter is to reveal the internal structure of the theory of quantum mechanics, or at least to provide with the basic background. However, the symmetry of the equations or the symmetry of eigenstates should not make us forget that in most experimental situations the QM states are not symmetric. It is the task of theory to extract the symmetry principles from asymmetric phenomena, and to derive the consequences of these principles, which can be tested by further observation. In fact, there are many cases where a well-accepted symmetry "principle" needed to be revised in the light of newer experimental observations. Parity conservation and locality principles are examples to be discussed in the following. Therefore, symmetry principles are not obvious *a priori*.

3.2 Symmetries and conservation laws in classical mechanics

Consider a classical system with Hamiltonian

$$H = H(q_1 \dots q_n, p_1 \dots p_n)$$

and suppose that the system's energy is unchanged upon an infinitesimal translation $q_i \to q_i + \delta q$ of one of its generalized coordinates (e.g., $q_i = x \to x + \delta x$). The invariance of H implies

$$\delta H = \frac{\partial H}{\partial q_i} \delta q = 0 \,,$$

or equivalently

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = 0 \,.$$

Therefore, the conjugated momentum is conserved:

$$\boxed{\text{Invariance upon translation}} \iff \boxed{\text{Momentum conservation}}$$

What happens if the system is invariant upon rotations around a given axis $\hat{n}$? Let us consider for simplicity a one-particle system with $\vec{r} = (x, y, z)$ and $\vec{p} = (p_x, p_y, p_z)$. The changes in position and momentum after an infinitesimal rotation $\delta\phi$ are given by

$$\delta\vec{r} = \hat{n} \times \vec{r}\, \delta\phi$$

and

$$\delta\vec{p} = \hat{n} \times \vec{p}\, \delta\phi \,.$$

If H is invariant under an infinitesimal rotation around the axis $\hat{n}$ we have

$$\begin{aligned}
0 = \delta H &= \sum_{i=1}^{3} \frac{\partial H}{\partial r_i} \delta r_i + \frac{\partial H}{\partial p_i} \delta p_i \\
&= -\dot{\vec{p}} \cdot (\hat{n} \times \vec{r})\, \delta\phi + \dot{\vec{r}} \cdot (\hat{n} \times \vec{p})\, \delta\phi \\
&= \left[-\hat{n} \cdot (\vec{r} \times \dot{\vec{p}}) - \hat{n} \cdot (\dot{\vec{r}} \times \vec{p}) \right] \delta\phi \\
&= -\delta\phi \frac{d}{dt} \left(\hat{n} \cdot (\vec{r} \times \vec{p}) \right) \\
&= -\delta\phi \frac{d}{dt} (\hat{n} \cdot \vec{L}).
\end{aligned}$$

Consequently, the projection of the angular momentum $\vec{L}$ along the rotation axis is conserved:

$$\boxed{\text{Invariance upon rotation around } \hat{n}} \Longleftrightarrow \boxed{\vec{L}\cdot\hat{n} \text{ is conserved}}$$

In classical mechanics the derivation of conservation laws is straightforward because we know how the dynamical variables, i.e., the classical state of the system, which is given by $\vec{r}_i$ and $\vec{p}_i$, transforms under the given symmetry operation (translation, rotation, etc.). In quantum mechanics the state is defined by a ket $|\,\Psi\,\rangle$, which is an element of a linear vector space (Hilbert space). In order to understand and exploit the symmetries in quantum mechanics we must first of all derive how the quantum mechanical states transform upon the various symmetry operations.

In the following we consider two physically equivalent descriptions of the same quantum state. The relation between the kets $|\,\Psi\,\rangle$ and $|\,\Psi'\,\rangle$ corresponding to these descriptions can be expressed in terms of an operator $\hat{D}$ whose form depends on the type of transformation:

$$|\,\Psi\,\rangle \longrightarrow |\,\Psi'\,\rangle = \hat{D}\,|\,\Psi\,\rangle\,.$$

Our first goal is to characterize the properties of $\hat{D}$. We shall see that if the descriptions of a physical system in terms of $|\,\Psi\,\rangle$ and $|\,\Psi'\,\rangle$ are equivalent, $\hat{D}$ needs to be either a *linear unitary operator* (the most usual case) or an *antilinear antiunitary operator* (in the case of time inversion).

3.3 Equivalent descriptions: Wigner's theorem

Consider two experimenters O and O' using two different reference systems. For example, the orientation of the coordinate frames of O and O' can be rotated or translated with respect to each other, or the chosen origins of time can be different. On the one side, O uses the kets $|\,\alpha\,\rangle$ for describing the physical states of the system, which are expressed in the orthonormal basis $\{|\,a_m\,\rangle\}$. For example, we can have $|\,a_m\,\rangle = |\,jm\,\rangle$, where $|\,jm\,\rangle$ are the eigenstates of J^2 and of the projection J_z along the z axis of O. On the other side, O' uses the kets $|\,\alpha'\,\rangle$ and a basis set $\{|\,a'_m\,\rangle\}$. The latter can be the eigenstates of J^2 and $J_{z'}$ where z' is the quantization axis chosen by O'. We denote these by $|\,jm\,\rangle'$ in order to stress that we are considering *the same physical state* as O but according to the reference frame of O'.

Imagine that O prepares a state $|\,p\,\rangle$ of momentum p traveling along the y axis and that he/she measure the probability of finding this state in an eigenstate $|\,jm\,\rangle$ of J^2 and J_z. The result of his/her measurement would then be $|\langle\, jm\,|\,p\,\rangle|^2$. Now O' prepares the same state (following the same procedure) and obtains a state $|\,p\,\rangle'$ traveling with the same momentum p and the same energy along the y' axis. If O' measures the probability of finding $|\,p\,\rangle'$ in an eigenstate of J^2 and $J_{z'}$ he/she finds $|'\langle\, jm\,|\,p\,\rangle'|^2$. Since the descriptions of O and O' are physically equivalent we must have

$$|\langle\, jm\,|\,p\,\rangle|^2 = |'\langle\, jm\,|\,p\,\rangle'|^2.$$

In fact, the previous condition must hold for any pair of equivalent states.

In order to formulate the previous statements in general terms we should clarify the notation. O uses the states

$$|\,\alpha\,\rangle = \sum_n c_n\,|\,a_n\,\rangle \tag{3.1}$$

and the basis $|\,a_n\,\rangle$, while O' uses the states

$$|\,\alpha'\,\rangle = \sum_n c'_n\,|\,a'_n\,\rangle \tag{3.2}$$

and the basis $|\,a'_n\,\rangle$. The properties of $|\,\alpha\,\rangle$, $|\,\beta\,\rangle$, etc. as described by O are the same as the properties of $|\,\alpha'\,\rangle$, $|\,\beta'\,\rangle$, etc. as described by O'. Therefore, we have

$$|\langle\,\alpha\,|\,\beta\,\rangle| = |\langle\,\alpha'\,|\,\beta'\,\rangle|.$$

Note that only the absolute values are the same, since only probabilities can be measured.

We would like to characterize the properties of the operator $\hat{D}$ connecting two such equivalent descriptions:

$$|\,\alpha'\,\rangle = \hat{D}\,|\,\alpha\,\rangle\,,$$

with the inverse one-to-one mapping

$$|\,\alpha\,\rangle = \hat{D}^{-1}\,|\,\alpha'\,\rangle.$$

Both basis sets are assumed to be complete

$$\sum_n |\,a_n\,\rangle\langle\,a_n\,| = \sum_n |\,a'_n\,\rangle\langle\,a'_n\,| = \mathbb{1}\,.$$

Without loss of generality we may choose the phase of the states $|\,a'_n\,\rangle$ so that $\hat{D}\,|\,a_n\,\rangle = |\,a'_n\,\rangle$ for all basis states. The most general states $|\,\alpha\,\rangle$ and $|\,\alpha'\,\rangle$ can be expressed in the corresponding basis as

$$|\,\alpha\,\rangle = \sum_n c_n\,|\,a_n\,\rangle \qquad \text{and} \qquad |\,\alpha'\,\rangle = \sum_n c'_n\,|\,a'_n\,\rangle\,.$$

The question is the following: What is the relation between c_n and c'_n when $\hat{D}$ connects the two equivalent descriptions? What are the possible relations that preserve the probability

$$|\langle\,\alpha\,|\,\beta\,\rangle|^2 = |\langle\,\alpha'\,|\,\beta'\,\rangle|^2 \quad \forall\,|\,\alpha\,\rangle \text{ and } |\,\beta\,\rangle\,?$$

Since the superposition principle must be respected, there are two possibilities, namely, $\hat{D}$ can be either a *linear operator* or an *antilinear operator* [1].

(1) Let us first assume that $\hat{D}$ is *linear*. Starting from Eq. (3.1),

$$|\,\alpha\,\rangle = \sum_n c_n\,|\,a_n\,\rangle$$

we obtain

$$|\,\alpha'\,\rangle = \hat{D}\,|\,\alpha\,\rangle = \sum_n c_n\,\hat{D}\,|\,a_n\,\rangle = \sum_n c_n\,|\,a'_n\,\rangle.$$

Comparing with Eq. (3.2) we conclude that, if $\hat{D}$ is linear, then $c'_n = c_n\ \forall\, n$. Moreover, for any

$$|\,\alpha\,\rangle = \sum_n c_n\,|\,a_n\,\rangle \quad \text{and} \quad |\,\beta\,\rangle = \sum_n d_n\,|\,a_n\,\rangle$$

The inner product is given by

$$\langle\,\alpha\,|\,\beta\,\rangle = \sum_{\substack{n\\m}} c^*_m\,d_n \langle\,a_m\,|\,a_n\,\rangle = \sum_n c^*_n\,d_n = \sum_n (c'_n)^*\,d'_n = \langle\,\alpha'\,|\,\beta'\,\rangle.$$

Consequently, if $\hat{D}$ is linear, then $\hat{D}$ preserves the inner product and is therefore *unitary*.

(2) The other possibility is that $\hat{D}$ is an *antilinear* operator. Using the definition of antilinear operators, namely, $\hat{\Lambda}\,(c\,|\,\alpha\,\rangle + |\,\beta\,\rangle) = c^*\,\hat{\Lambda}|\,\alpha\,\rangle + \hat{\Lambda}|\,\beta\,\rangle$, we have

$$\begin{aligned} |\,\alpha'\,\rangle &= \hat{D}\,|\,\alpha\,\rangle = \hat{D}\left(\sum_n c_n\,|\,a_n\,\rangle\right) \\ &= \sum_n c^*_n\,\hat{D}\,|\,a_n\,\rangle = \sum_n c^*_n\,|\,a'_n\,\rangle. \end{aligned}$$

Comparing with Eq. (3.2) we conclude that, if $\hat{D}$ is antilinear, $c'_n = c^*_n$. Moreover, for any

$$|\,\alpha'\,\rangle = \sum_n c^*_n\,|\,a'_n\,\rangle \ \text{ and } \ |\,\beta'\,\rangle = \sum_n d^*_n\,|\,a'_n\,\rangle$$

we have

$$\langle\,\alpha'\,|\,\beta'\,\rangle = \sum_n c_n\,d^*_n = \langle\,\alpha\,|\,\beta\,\rangle^* = \langle\,\beta\,|\,\alpha\,\rangle,$$

which of course satisfies

$$|\langle\,\alpha'\,|\,\beta'\,\rangle| = |\langle\,\alpha\,|\,\beta\,\rangle|.$$

In this case $\hat{D}$ is said to be *anti-unitary*, since it is antilinear and the inner products between the transformed states are the complex conjugate of the original ones.

Notice that antiunitary operators are <u>not</u> linear. In quantum mechanics the phase cannot be measured. Therefore the phase of the scalar product need not be preserved in a symmetry transformation. One often says, that quantum mechanics is a theory on a *ray space*, since the physical meaning of states and the equivalence between them holds for rays of vectors. The ray associated to a ket $|\,\Psi\,\rangle$ is defined by the set of states $\left\{e^{i\gamma}\,|\,\Psi\,\rangle \,\text{with}\, \gamma \in \mathbb{R}\right\}$. Dropping the often artificial normalization condition, the ray associated with $|\,\Psi\,\rangle$ is the one-dimensional subspace generated by $|\,\Psi\,\rangle$ (i.e., $\{\lambda|\,\Psi\,\rangle \,\text{with}\, \lambda \in \mathbb{C}\}$).

Wigner's theorem states that unitary and antiunitary transformations are the only two possibilities of connecting equivalent representations, once the phase convention $\hat{D}\,|\,a_n\,\rangle =$

$|\, a'_n \,\rangle$ between the basis states has been met. The proof is actually simple and may be found in Gottfried's book [1].

3.4 Unitary versus antiunitary transformations

The purpose of this section is to clarify the differences between unitary and antiunitary operators and to assess the relevance of the latter. Let $\hat{\Lambda}$ be an antiunitary transformation satisfying $\Lambda\,|\, a_n \,\rangle = |\, a'_n \,\rangle$ for some complete basis $\{|\, a_n \,\rangle\}$. We then have, for any $|\, \alpha \,\rangle = \sum_n |\, a_n \,\rangle\langle\, a_n \,|\, \alpha \,\rangle$, that

$$|\, \alpha' \,\rangle = \hat{\Lambda}\,|\, \alpha \,\rangle = \sum_n \langle\, a_n \,|\, \alpha \,\rangle^*\, \hat{\Lambda}\,|\, a_n \,\rangle = \sum_n |\, a'_n \,\rangle\langle\, a_n \,|\, \alpha \,\rangle^*.$$

Note that $\hat{\Lambda}$ is defined in relation to a given specific basis $\{|\, a_n \,\rangle\}$, since $\hat{\Lambda}\,|\, a_n \,\rangle = |\, a'_n \,\rangle$. The expansion coefficients $\langle\, a_n \,|\, \alpha \,\rangle$ of an arbitrary state $|\, \alpha \,\rangle = \sum_n \langle\, a_n \,|\, \alpha \,\rangle |\, a_n \,\rangle$ *in this specific basis* are conjugated under the action of $\hat{\Lambda}$.

It is easy to show that if $\hat{\Lambda}_1$ and $\hat{\Lambda}_2$ are antiunitary transformations then $\hat{U} = \hat{\Lambda}_1\,\hat{\Lambda}_2$ is unitary. First, $\hat{U}$ is linear:

$$\begin{aligned} \hat{U}(c|\, \alpha \,\rangle + |\, \beta \,\rangle) &= \hat{\Lambda}_1\hat{\Lambda}_2(c|\, \alpha \,\rangle + |\, \beta \,\rangle) = \hat{\Lambda}_1(c^*\hat{\Lambda}_2|\, \alpha \,\rangle + \hat{\Lambda}_2|\, \beta \,\rangle) = c\hat{\Lambda}_1\hat{\Lambda}_2|\, \alpha \,\rangle + \hat{\Lambda}_1\hat{\Lambda}_2|\, \beta \,\rangle \\ &= c\hat{U}|\, \alpha \,\rangle + \hat{U}|\, \beta \,\rangle)\,. \end{aligned}$$

And second, $\hat{U}$ preserves the scalar product:

$$\langle\, \hat{U}\alpha \,|\, \hat{U}\beta \,\rangle = \langle\, \hat{\Lambda}_1(\hat{\Lambda}_2\alpha) \,|\, \hat{\Lambda}_1(\hat{\Lambda}_2\beta) \,\rangle = \langle\, \hat{\Lambda}_2\beta \,|\, \hat{\Lambda}_2\alpha \,\rangle = \langle\, \alpha \,|\, \beta \,\rangle\,,$$

where we have used that $\langle\, \hat{\Lambda}\beta \,|\, \hat{\Lambda}\alpha \,\rangle = \langle\, \alpha \,|\, \beta \,\rangle$ for antiunitary $\hat{\Lambda}$. In particular one concludes that $\hat{\Lambda}^2$ is always unitary.

Antiunitary operators are often written as $\hat{\Lambda} = \hat{U}\hat{K}$ where $\hat{U}$ is the linear unitary transformation that maps the corresponding basis set

$$\hat{U}\,|\, a_n \,\rangle = |\, a'_n \,\rangle,$$

and $\hat{K}$ is the operator of complex conjugation of the expansion coefficients *leaving the basis states unchanged:*

$$\hat{K}\,|\, a_n \,\rangle = |\, a_n \,\rangle$$

and

$$\hat{K}c_n\,|\, a_n \,\rangle = c_n^*\,\hat{K}\,|\, a_n \,\rangle = c_n^*\,|\, a_n \,\rangle.$$

If the basis is changed, for instance $|\, b_n \,\rangle = i\,|\, a_n \,\rangle$, the roles of $\hat{K}$ and $\hat{U}$ need to be redefined in order to keep the antilinear operator $\hat{\Lambda}$ unchanged. Since $\hat{K}\,|\, a_n \,\rangle = |\, a_n \,\rangle$ and $\hat{K}\,|\, b_n \,\rangle = |\, b_n \,\rangle$ it is clear that $\hat{K}$ is basis dependent.

How important are antiunitary transformations in quantum mechanics?

Suppose that the two equivalent descriptions can be connected by a transformation which depends on a continuous parameter (e.g., a translation, rotation, or time displacement). Let R_1 and R_2 be two of such transformations connecting equivalent descriptions. It is clear that the composition $R_3 = R_2\, R_1$ is also a transformation connecting equivalent descriptions. Moreover, the corresponding representation $\hat{D}$ of these transformations in the Hilbert space of quantum states must obviously satisfy the composition rule

$$\hat{D}(R_3) = \hat{D}(R_2)\, \hat{D}(R_1).$$

If the transformation R is continuous, we can always construct a "half-way" transformation $R_{1/2}$ such that

$$R_{1/2}\, R_{1/2} = R\,,$$

which implies

$$\hat{D}(R) = \hat{D}(R_{1/2})^2.$$

Consequently, $\hat{D}(R)$ is unitary, since the product of two unitary or antiunitary operators is necessarily unitary.

One concludes that antiunitary transformations can only be relevant for discrete symmetries, such as reflexions, space inversion or time reversal. In fact, they are only relevant for the latter, since the representations of the other two correspond, as we shall see, to unitary transformations.

3.5 Groups

By definition, a group is a set of elements $G = \{r, s, t, \ldots\}$ between which a composition or multiplication law denoted by "$\circ$" is defined, which satisfies the following four properties:

1) Closure: $\boxed{r \circ s \in G}\ \forall\ r, s \in G$,

2) Associative law: $\boxed{r \circ (s \circ t) = (r \circ s) \circ t}\ \forall\ r, s, t \in G$,

3) Existence of neutral element, usually denoted by 1 or e (from the German word *Einzelelement*) which satisfies $\boxed{r \circ e = e \circ r = r}\ \forall\ r \in G$, and

4) Existence of an inverse element: $\forall\ r \in G, \quad \exists\ r^{-1} \in G$ such that $\boxed{r \circ r^{-1} = r^{-1} \circ r = e}$.

In addition, one says that a group is *commutative* or *Abelian* if $r \circ s = s \circ r$ for all $r, s \in G$. Groups can be discrete (finite or infinite) or continuous, as illustrated the following examples.

1) Discrete groups

i) $G = \{e\}$ is the trivial group.

ii) The set of two integers $G = \{1, -1\}$ and number multiplication:

$\circ$	1	-1
1	1	-1
-1	-1	1

(Multiplication table)

iii) $\mathbb{Z}$ with the arithmetic sum $\circ \equiv +$ as operation. In this case $e = 0$ and $r^{-1} = -r$. This group is infinite but discrete.

iv) The nth complex roots of unity $\mathbb{Z}_n = \{z \in \mathbb{C} \text{ such that } z^n = 1\}$ form a finite discrete group with $\circ$ being the usual complex number multiplication. This group is isomorphic to the group C_n of the n rotations around a given axis with angles that are multiples of $2\pi/n$.

v) The permutations of N elements, i.e., the set of all bijective mappings P in the natural interval $[1, N]$ ($P : [1, N] \overset{\text{bijective}}{\longrightarrow} [1, N]$). For example, for $N = 3$ we denote them as

$$P \equiv \begin{pmatrix} 1 & 2 & 3 \\ P(1) & P(2) & P(3) \end{pmatrix}.$$

This is an important group for the theory of indistinguishable particles.

2) Continuous groups

i) The set of all rotations around a given axis $\hat{n}$ in $\mathbb{R}^3$ forms a group with the composition of rotations as operation. This group is Abelian and isomorphic to the real numbers with the sum as operation ($\{\mathbb{R}, +\}$).

ii) The translations in $\mathbb{R}^3$. This group is also Abelian and isomorphic to $\mathbb{R}^3$ itself with the vector sum as operation.

iii) The following non-Abelian matrix groups with the multiplication as composition law are particularly important in physics:

$U(n)$: Unitary matrices in $\mathbb{C}^{n\times n}$.
$SU(n)$: Special $n \times n$ unitary complex matrices having $\det(U) = 1$.
For $n = 2$, this group describes the rotations of spin $1/2$ spinors.
$O(n)$: Orthogonal matrices in $\mathbb{R}^{n\times n}$, which satisfy $O^t = O^{-1}$.
$SO(n)$ Special orthogonal real $n \times n$ matrices having $\det(O) = 1$.
For $n = 3$ these are the rotation matrices in $\mathbb{R}^3$.

All the above mentioned continuous groups are *Lie Groups* or smooth groups, i.e., groups where the elements depend smoothly on the defining parameters (e.g., rotation axis and

angle) so that the notion of differentiability exist. Groups that are smooth manifolds are Lie Groups. To each Lie group there is an associated *Lie algebra*, which corresponds to the vector space tangent to the group around the identity. For example, if we consider the group $U(n)$ of unitary matrices, the corresponding Lie algebra is given by the matrices of the form

$$U = \mathbb{1} + i\varepsilon G + O(\varepsilon^2)\,,$$

where the real number $\varepsilon \to 0$ is infinitesimally small ($\varepsilon^2 \simeq 0$) and G is an hermitian matrix.

The Lie algebra describes *infinitesimal transformations*. The matrix or operator $\hat{G}$, being hermitian, usually has a clear physical or geometrical meaning. As we shall see, many important problems related to the representation of Lie groups are solved by considering the corresponding Lie algebra, where the group product is linearized ($\varepsilon^2 \simeq 0$). The group properties for arbitrary finite transformations can be then derived by integration.

Why is the concept of group so important in quantum mechanics and in physics in general? Probably the main reason is that the ensemble of all transformations which leave a quantum-mechanical system invariant or, more generally, which connect equivalent descriptions, forms quite naturally a group (see Sec. 3.3) . The reader may easily convince himself that all 4 group axioms (closure, associative law, existence of neutral and inverse elements) are satisfied. Still, our actual interest as physicist is not so much in the properties of the group itself [for instance, in the properties of SO(3) matrices describing rotations in $\mathbb{R}^3$] but rather in the consequences of applying symmetry transformations to the kets and bras defining the quantum-mechanical states of physical systems. As we know, the quantum states $|\,\Psi\,\rangle$ are elements of a linear vector space, thanks to the superposition principle. Therefore, we need to understand the connection between groups and vector spaces (Hilbert spaces) which is crystallized by the central concept of *representation of a group*. It is in fact the theory of group representations what constitutes the foundation of the theory of symmetry in quantum mechanics.

3.6 Representation of a group

Given a group G (with elements $s \in G$) one speaks of representation of the group G when for each element s of G there is a linear operator $\hat{D}(s)$ acting in a vector space $\mathcal{H}$ (the so-called representation space) such that the composition of the operators $\hat{D}(s)$ acting in $\mathcal{H}$ corresponds to the multiplication of elements in G. In physical or geometrical terms, $\hat{D}(s)$ represents the action of the group-element s in the vector space $\mathcal{H}$.

In more formal terms, the set $\{\hat{D}(s)\colon \mathcal{H} \to \mathcal{H}$ with $s \in G$ and $\hat{D}(s)$ linear$\}$ is a *representation of G in the vector space $\mathcal{H}$* if
1. $\hat{D}(e) = \mathbb{1}$ = Identity operator in $\mathcal{H}$, and
2. $\hat{D}(s\,r) = \hat{D}(s)\,\hat{D}(r) \quad \forall s, r \in G$.
These two fundamental requirements, together with the group properties of G, imply that all the operators $\hat{D}(s)$ of a representation are invertible and that $\hat{D}(s^{-1}) = [\hat{D}(s)]^{-1}$, since $\mathbb{1} = \hat{D}(e) = \hat{D}(s\,s^{-1}) = \hat{D}(s)\,\hat{D}(s^{-1})$ as well as $\mathbb{1} = \hat{D}(s^{-1}\,s) = \hat{D}(s^{-1})\,\hat{D}(s)$.

Examples:
1) Consider the group of rotation matrices $R \in SO(3)$ and a classical particle whose state is defined by its position and momentum $(x, y, z, p_x, p_y, p_z) \in \mathbb{R}^6$. The 6×6 matrices

$$\hat{D}(R) = \left(\begin{array}{c|c} R & 0 \\ \hline 0 & R \end{array}\right)$$

describing the change of the classical state upon rotation

$$\begin{pmatrix} x' \\ y' \\ z' \\ p'_x \\ p'_y \\ p'_z \end{pmatrix} = \hat{D}(R) \begin{pmatrix} x \\ y \\ z \\ p_x \\ p_y \\ p_z \end{pmatrix}$$

is a representation of SO(3) in $\mathbb{R}^6$. Representations that fall into blocks for all group elements R are called reducible.

2) Consider the $(2j+1)$-dimensional space spanned by the $|\, jm \,\rangle$ eigenstates of J^2 and J_z as the representation space and the SO(3) rotation group. We know that upon rotation the kets $|\, jm \,\rangle$ transform into a linear combination of states having the same j but involving in general all the other $|\, jm' \,\rangle$ having different J_z projections m':

$$\hat{D}(R)\,|\, jm \,\rangle = \sum_{m'=-j}^{j} |\, jm' \,\rangle D^{(j)}_{m'm}\,,$$

where the complex unitary matrix

$$D^{(j)}_{m'm} = \langle\, jm' \,|\, e^{-(i/\hbar)\,\hat{\vec{J}}\cdot\hat{n}\,\phi} \,|\, jm \,\rangle\,,$$

sometimes called Wigner functions, is a representation of the rotation group having the dimension $2j+1$. The unit vector $\hat{n}$ defines the axis of the rotation R and ϕ its angle. $\hat{\vec{J}}$ stands for the angular momentum operator. This is known as the $(2j+1)$-irreducible representation of SO(3).

3.7 Reducible versus irreducible representations

A representation $\{\hat{D}(r) \,\text{with}\, r \in G\}$ of the group G in the vector space $\mathcal{H}$ is said to be *reducible* if a nontrivial subspace $\mathcal{S} \subset \mathcal{H}$ exists ($\mathcal{S} \neq \vec{0}$ and $\mathcal{S} \neq \mathcal{H}$) such that $\forall r \in G$ and $\forall\, |\alpha\,\rangle \in \mathcal{S}$

$$\hat{D}(r)\,|\,\alpha\,\rangle \in \mathcal{S}\,.$$

In this case $\mathcal{S}$ is said to be an *invariant subspace* under $\hat{D}$. Otherwise, i.e., if the only invariant subspaces are the trivial ones $\{\vec{0}\}$ and $\mathcal{H}$, the representation is said to be *irreducible*.

It is easy to show that a reducible representation $\hat{D}(r)$ falls into separate blocks for all r by adapting the basis of $\mathcal{H}$ to $\mathcal{S}$. Given a basis $\{\,|\,1\,\rangle \ldots |\,n\,\rangle\}$ of $\mathcal{S}$, where n is the dimension of $\mathcal{S}$, one may construct a basis of $\mathcal{H}$ by adding mutually orthogonal states $|\,k\,\rangle$ $(k > n)$ such that $\langle\, k\,|\,i\,\rangle = 0$ for $k > n$ and $i \leq n$. The invariance of $\mathcal{S}$ for all r implies $\langle\, k\,|\,\hat{D}(r)\,|\,i\,\rangle = 0$ as well as $\langle\, k\,|\,\hat{D}(r^{-1})\,|\,i\,\rangle = 0$ for $i \leq n$ and $k > n$. Consequently,

$$0 = \langle\, k\,|\,\hat{D}(s^{-1})\,|\,i\,\rangle = \langle\, k\,|\,\hat{D}^{\dagger}(s)\,|\,i\,\rangle = \langle\, i\,|\,\hat{D}(s)\,|\,k\,\rangle,$$

which, together with $\langle\, k\,|\,\hat{D}(r)\,|\,i\,\rangle = 0$, shows that $\hat{D}(r)$ can be split in blocks for all r.

Notice that if F is a subgroup of G and $\hat{D}(r)$ a representation of G, the assignment $\hat{D}(r)$ for $r \in F$ is a representation of the subgroup F. However, a representation which is irreducible when considered as a representation of the larger group G, need not be irreducible when it is regarded as a the representation of the subgroup. For example, the $2j+1$-irreducible representation of SO(3) becomes diagonal when restricted to the subgroup of rotations around the z axis:

$$\langle\, jm'\,|\,e^{-(i/\hbar)\,\hat{J}_z\,\phi}\,|\,jm\,\rangle = \delta_{m\,m'}\,e^{-(i/\hbar)\,m\,\phi}\,.$$

Consequently, the $2j+1$-dimensional representation $\{|\,j,m\,\rangle$ with $m = -j,\ldots,j\}$ is reduces to $2j+1$ distinct one-dimensional representations when the subgroup of rotations around the z axis is considered.

3.8 Continuous symmetry transformations in quantum mechanics

We consider a general transformation

$$|\,\Psi\,\rangle \xrightarrow{\hat{S}} |\,\Psi'\,\rangle$$

acting on an arbitrary quantum state $|\,\Psi\,\rangle$ and yielding an equivalent state $|\,\Psi'\,\rangle$. One may regard $\hat{S}$ as a passive transformation connecting two equivalent representations of the same quantum state (e.g., corresponding to different coordinate systems). Or one can regard $\hat{S}$ as an active transformation of the state $|\,\Psi\,\rangle$ yielding a new equivalent state, for example, the result of rotating or translating in 3D space. In these lectures the latter active perspective is adopted in most cases.

In the following we focus on continuous transformations, where $\hat{S}$ depends on some real parameters s (e.g., $s \equiv$ time shift, displacement in a given direction, rotation angle, etc.). We have shown that $\hat{S}$ is a linear unitary operator:

$$|\,\Psi'\,\rangle = \hat{S}(s)\,|\,\Psi\,\rangle$$

with $\hat{S}^{\dagger}\,\hat{S} = \mathbb{1}$. Since $s = 0$ means no transformation at all we have $\hat{S}(0) = \mathbb{1}$. Moreover, the transformation corresponding to $s_1 \circ s_2$ is the composition of the transformations associated to s_1 and s_2, i.e., $\hat{S}(s_1 \circ s_2) = \hat{S}(s_1)\,\hat{S}(s_2)$. Consequently, $\hat{S}$ is a representation of the group of all the symmetry transformations parameterized by s in the Hilbert space of our QM system.

We consider now an *infinitesimal transformation*

$$\hat{S}(ds) = 1 + \delta\hat{S} + O(\delta\hat{S}^2)$$

with $\| \delta\hat{S} \| \ll 1$. Since $\hat{S}$ is unitary, we have

$$\begin{aligned} 1 &= [1 + \delta\hat{S} + O(\delta\hat{S}^2)]^\dagger \, [1 + \delta\hat{S} + O(\delta\hat{S}^2)] \\ &= 1 + (\delta\hat{S} + \delta\hat{S}^\dagger) + O(\delta\hat{S}^2) \, . \end{aligned}$$

This implies that $\delta\hat{S} + \delta\hat{S}^\dagger = 0$ or, in other words, that $\delta\hat{S}$ is antihermitian. It is therefore convenient to write

$$\delta\hat{S} = \frac{-i}{\hbar} \, \hat{G} \, ds \, ,$$

where $ds \to 0$ and $\hat{G}^\dagger = \hat{G}$ is hermitian.

The infinitesimal transformations are then given by

$$\hat{S}(ds) = 1 - \frac{i}{\hbar} \, \hat{G} \, ds + O(ds^2) \, .$$

The operator $\hat{G}$ is known as the *generator* of the group of symmetry transformations $\hat{S}$ or, more precisely, the generator of the infinitesimal transformations $\hat{S}(ds)$. Notice that $\hat{G}$ depends crucially on the type of transformation that one is considering, since the changes in the physical states resulting from, for example, time displacements, rotations around the x-axis, rotations around the z-axis, etc. are profoundly different.

Example:

Consider an infinitesimal translation $\delta\vec{r}$ of a single-particle wave function $\Psi(\vec{r})$. A simple geometrical argument shows that

$$\Psi'(\vec{r}) \;\; = \;\; \Psi(\vec{r} - \delta\vec{r}) \, .$$

It follows that

$$\begin{aligned} \Psi'(\vec{r}) &= \Psi(\vec{r}) - \delta\vec{r} \cdot \vec{\nabla}\Psi + O(\delta r^2) \\ &= (1 - \delta\vec{r} \cdot \vec{\nabla})\Psi(\vec{r}) + O(\delta r^2) \\ &= \left[1 - \frac{i\delta\vec{r}}{\hbar} \cdot (-i\hbar\vec{\nabla})\right] \Psi(\vec{r}) + O(\delta r^2) \, . \end{aligned}$$

Consequently, the generator of infinitesimal translations is the projection $\hat{G} = -i\hbar \, \hat{n} \cdot \vec{\nabla}$ of the momentum operator $\hat{\vec{p}}$ along the direction $\hat{n} = \delta\vec{r}/\delta r$ of the translation.

In the case of time displacements and rotations the generators are the Hamiltonian and the angular momentum operator, respectively. One may actually say that in quantum mechanics these operators are defined by the property of being the generators of the corresponding infinitesimal transformations.

3.9 Finite Transformations

In many cases of interest the generator of infinitesimal transformations $\hat{G}$ is independent of the parameter s which characterizes a finite continuous transformation. For example, since space is homogeneous, the generator of translations along any given direction $\hat{n}$ (i.e., the corresponding component $\hat{n} \cdot \hat{\vec{p}} = -i\hbar\, \hat{n} \cdot \vec{\nabla}$ of the momentum operator) is independent of the position of the system or the extent $s = \hat{n} \cdot \Delta\vec{r}$ of any previous translation. Similarly, since time is homogeneous, the generator of time displacements in an isolated system (no time-dependent external fields) is independent of the time t at which the infinitesimal time propagation is performed. In these cases the operator $\hat{G}$ generating an infinitesimal transformation

$$\hat{S}(ds) = 1 - \frac{i}{\hbar}\,\hat{G}\,ds + O(\delta s^2) \tag{3.3}$$

around the identity cannot depend on s. In this case the operator $\hat{S}(ds)$ is always the same [Eq. (3.3)] irrespectively of any other transformations $\hat{S}(s)$ with finite s which might have already been performed. However, this is not the ony physical situation that one may face. For example, if the system is not isolated, it can be subject to external time-dependent fields and therefore the nature of an infinitesimal time propagation, given by the Hamiltonian $\hat{H}(t)$, can depend on the time t at which the infinitesimal displacement takes place.

We consider the most general situation in which $\hat{G}$ depends on s and search for the representation of finite transformations

$$|\,\Psi'\,\rangle = \hat{S}(s)\,|\,\Psi\,\rangle.$$

Imagine that, starting from some state $|\Psi\,\rangle$, the system is subject to the finite-s transformation $\hat{S}(s)$ and that it is subsequently transformed according to the infinitesimal $\hat{S}(ds)$. The result is then $|\Psi'\,\rangle = \hat{S}(ds)\,\hat{S}(s)\,|\Psi\,\rangle$. Clearly, this must be the same as performing just one transformation of the state $|\Psi\,\rangle$ with the paramter $s + ds$, namely, $|\Psi'\,\rangle = \hat{S}(s + ds)\,|\Psi\,\rangle$. This implies

$$\hat{S}(s + ds) = \hat{S}(ds)\,\hat{S}(s)\,,$$

which is equivalent to the representation property of $\hat{S}$. It follows that

$$\hat{S}(s + ds) - \hat{S}(s) = [\hat{S}(ds) - 1]\hat{S}(s)$$

$$\hat{S}(s + ds) - \hat{S}(s) = -\frac{i}{\hbar}\,\hat{G}(s)\,ds\,\hat{S}(s)$$

and finally

$$i\hbar\frac{d\,\hat{S}(s)}{ds} = \hat{G}(s)\,\hat{S}(s). \tag{3.4}$$

This important differential equation follows directly from the group-representation property of symmetry transformations and the definition of their generator $\hat{G}$ around the identity (Lie Algebra).

In the case of translations, rotations or time displacements of isolated systems, where $\hat{G}$ is independent of s, Eq. (3.4) can be integrated straightforwardly to yield

$$\hat{S}(s) = e^{\frac{-i}{\hbar}\hat{G}\,s}, \tag{3.5}$$

where we have used the condition $\hat{S}(0) = 1$. This is the general form of the *representation of a continuous transformation* connecting equivalent descriptions. In the following sections we discuss the most relevant examples of transformations between equivalent descriptions: time evolution, translations and rotations.

Notice that the differential equation (3.4) for $\hat{S}(s)$ holds for continuous transformations in general, even if $\hat{G}$ depends on s (e.g., when the Hamiltonian depends on time). In this case $\hat{G}(s)$ is the generator of an infinitesimal transformation ds occurring after the transformation $S(s)$, namely,

$$\hat{S}(s + ds) = [1 - \frac{i}{\hbar}\hat{G}(s)\,ds]\,\hat{S}(s), \tag{3.6}$$

which leads straightforwardly to Eq. (3.4). For translations in time, for example, $\hat{G}(s)$ is given by the Hamiltonian $\hat{H}(t)$ at time t. Furthermore note that $\hat{G}(s)$ and $\hat{S}(s)$ need not commute with each other (see Sec. 3.10). This happens when $[\hat{G}(s), \hat{G}(s')] \neq 0$ for $s \neq s'$. Therefore, the order of the operators in Eq. (3.4) matters in general.

Alternative derivation of the representation of finite transformations:

It is instructive to derive Eq. (3.5) by using the group-representation properties of S and the generator of infinitesimal transformations $\hat{G}$. Splitting a finite-s transformation into the product of N transformations of magnitude $ds = s/N$, we may write

$$\hat{S}(s) = \hat{S}(N\frac{s}{N}) = \hat{S}(\frac{s}{N})^N = (1 + \Delta\hat{S})^N\,,$$

where we have assumed that $\hat{G}$ and thus $\Delta\hat{S}(s/N)$ are independent of s. In the limit $N \to \infty$ we may replace $\Delta\hat{S}$ by

$$\Delta\hat{S} = -\frac{i}{\hbar}\hat{G}\,\frac{s}{N}\,,$$

and therefore,

$$\hat{S}(s) = \lim_{N\to\infty}\left(1 - \frac{i}{\hbar}\frac{\hat{G}s}{N}\right)^N.$$

Recalling that $\lim_{N\to\infty}\left(1+\frac{\Delta}{N}\right)^N = e^{\Delta}$, we have[4]

$$\hat{S}(s) = e^{\frac{-i}{\hbar}\hat{G}s}.$$

As in the previous derivation we have used the representation property of $\hat{S}$, the fact that $\hat{G}$ is the generator of infinitesimal transformations and, in the present case, that $\hat{G}$ is independent of s.

3.10 Displacement in Time

We consider a system which is in *not necessarily isolated*, for example, an atom subject to some time-dependent electromagnetic field. At time t_0 the system is some initial state $|\,\alpha\,\rangle$. For $t > t_0$ it evolves to the state $|\,\alpha\, t_0, t\,\rangle$, where we have explicitly recorded that this state is the result of the time evolution of $|\,\alpha\,\rangle = |\,\alpha\, t_0, t_0\,\rangle$ prepared at $t = t_0$. This notation is not superfluous, since in the most general situation $|\,\alpha\, t_0, t\,\rangle$ may depend not only on the initial ket $|\,\alpha\,\rangle$, but also on the time t_0 at which the time evolution started.

The ket $|\,\alpha\,\rangle$ is characterized by a complete set of quantum numbers a and the sum of the probabilities of finding the system in any of the states $|\,a\,\rangle$ is

$$\sum_a |\,c_a(t_0)\,|^2 = \langle\,\alpha\,|\,\alpha\,\rangle = 1,$$

where we have used that

$$|\,\alpha\,\rangle = \sum_a c_a(t_0)\,|\,a\,\rangle$$

and that the basis states are properly normalized ($\langle a|a'\rangle = \delta_{aa'}$). The same holds for the state $|\,\alpha\, t_0, t\,\rangle$ at any later time t, i.e.,

$$\langle\,\alpha\, t_0, t\,|\,\alpha\, t_0, t\,\rangle = 1 \quad \forall\, t.$$

Thus, the norm is conserved at all times for any initial state $|\,\alpha\,\rangle$. Consequently, the time evolution operator $\hat{U}(t, t_0)$, which is defined by

$$|\,\alpha\, t_0, t\,\rangle = \hat{U}(t, t_0)\,|\,\alpha\, t_0, t_0\,\rangle\,,$$

is unitary. Let us recall that linear operators that preserve the norm of any state also preserve the inner product between any two states and are thus unitary. In addition, $\hat{U}(t, t_0)$ must satisfy the group composition or causality property. Indeed, if the time propagation from $t_0 \to t$ yields

$$|\,\alpha\, t_0, t\,\rangle = \hat{U}(t, t_0)\,|\,\alpha\, t_0, t_0\,\rangle\,, \tag{3.7}$$

[4] $\left(1+\frac{\Delta}{N}\right)^N = \sum_{k=0}^{N}\binom{N}{k}\left(\frac{\Delta}{N}\right)^k = \sum_{k=0}^{N}\frac{N!}{k!(N-k!)}\frac{\Delta^k}{N^k} = \sum_{k=0}^{N}\underbrace{\frac{N(N-1)\ldots(N-k+1)}{N^k}}_{\longrightarrow 1 \text{ for } N\to\infty \text{ and all finite } k.}\frac{\Delta^k}{k!}$

a further propagation from $t \to t'$ yields

$$|\,\alpha\, t_0, t'\,\rangle = \hat{U}(t', t)\,|\,\alpha\, t_0, t\,\rangle\,.$$

This must be the same as the direct propagation

$$|\,\alpha\, t_0, t'\,\rangle = \hat{U}(t', t_0)\,|\,\alpha\, t_0, t_0\,\rangle$$

from $t_0 \to t'$. Therefore

$$\hat{U}(t', t_0) = \hat{U}(t', t)\,\hat{U}(t, t_0). \tag{3.8}$$

It is also easy to see that $\hat{U}$ satisfies all the remaining representation properties, namely, $\hat{U}(t, t) = 1 \;\; \forall\; t$ and $[\hat{U}(t_1, t_2)]^{-1} = [\hat{U}(t_1, t_2)]^{\dagger} = \hat{U}(t_2,\, t_1)$. In sum, changing the origin of time from t_0 to t

$$|\,\Psi\,\rangle = |\,\alpha\, t_0\,\rangle \longrightarrow |\,\Psi'\,\rangle = |\,\alpha\, t_0, t\,\rangle$$

constitutes a change of representation between two *equivalent descriptions*. Time evolution connects equivalent descriptions through a unitary transformation.

Let us first of all consider an infinitesimal time displacement $\delta t \to 0$, in which case we have

$$\hat{U}(t + \delta t, t) = 1 + \delta\hat{U}(t) + O(\delta\hat{U}^2)\,.$$

It follows that

$$1 = \hat{U}\,\hat{U}^{\dagger} = 1 + \delta\hat{U}(t)^{\dagger} + \delta\hat{U}(t) + O(\delta\hat{U}^2)\,.$$

It is then convenient to express $\delta\hat{U}$ in terms of the generator $\hat{H}(t)$ of the Lie algebra as

$$\delta\hat{U}(t) = \frac{-i}{\hbar}\hat{H}(t)\,\delta t.$$

The generator of infinitesimal time displacements is the time-dependent Hamiltonian $\hat{H}(t)$. Of course, $\delta\hat{U} + \delta\hat{U}^{\dagger} = 0$ implies that $\hat{H}(t) = \hat{H}(t)^{\dagger}$ and vice versa.

As in any continuous transformation [see Eq. (**??**)] we have

$$i\hbar\,\frac{\partial\hat{U}(t, t_0)}{\partial t} = \hat{H}(t)\,\hat{U}(t, t_0) \tag{3.9}$$

with the initial condition

$$\hat{U}(t_0, t_0) = 1.$$

Taking into account that $|\,\alpha\, t_0, t\,\rangle = \hat{U}(t, t_0)\,|\,\alpha\, t_0, t_0\,\rangle$, one concludes that

$$\begin{aligned} i\hbar\,\frac{\partial\,|\,\alpha\, t_0, t\,\rangle}{\partial t} &= i\hbar\,\frac{\partial\hat{U}(t, t_0)}{\partial t}\,|\,\alpha\, t_0, t_0\,\rangle \\ &= \hat{H}(t)\,\hat{U}(t, t_0)\,|\,\alpha\, t_0, t_0\,\rangle \\ &= \hat{H}(t)\,|\,\alpha\, t_0, t\,\rangle, \end{aligned} \tag{3.10}$$

which is indeed the well-known Schrödinger equation. It is clear that one could start from Eq. (3.10) as in introductory courses, define $\hat{U}(t,\, t_0)$ through Eq. (3.7), and obtain the differential Eq. (3.9). The present approach has the merit of revealing that time evolution is a particular case of the more general continuous transformation between equivalent descriptions which, as such, satisfies a number of properties: representation or composition property, unitarity of $\hat{U}$, hermiticity of the generator $\hat{H}$ of infinitesimal transformations, etc. As we shall see, these fundamental properties are shared by all other continuous symmetry transformations.

We seek now for the solution of $\hat{U}(t, t_0)$ in terms of $\hat{H}(t)$ or, in the language of group theory, we look for the representation of time displacements for finite time intervals. Since we cannot integrate Eq. (3.9) straightforwardly as in the case of a time-independent generator $\hat{G}$, we proceed by iteration. The differential equation

$$i\hbar \frac{\partial}{\partial t} \hat{U}(t, t_0) = \hat{H}(t)\, \hat{U}(t, t_0)$$

can be formally integrated to yield

$$\hat{U}(t, t_0) = 1 - \frac{i}{\hbar} \int_{t_0}^{t} \hat{H}(t')\, \hat{U}(t', t_0)\, dt' \,, \tag{3.11}$$

where we have used that $\hat{U}(t_0, t_0) = 1$. This type of equation is sometimes referred to as Dyson equation. It relates the propagator $\hat{U}(t, t_0)$ from t_0 to t with the propagators $\hat{U}(t', t_0)$ up to all previous times $t' < t$, thus respecting causality. We may now set $t = t'$ and $t' = t''$ in Eq. (3.11) and substitute the thus obtained $\hat{U}(t', t_0)$ on the right-hand side of the same Eq. (3.11) to yield

$$\hat{U}(t, t_0) = 1 - \frac{i}{\hbar} \int_{t_0}^{t} \hat{H}(t')\, dt' + \left(\frac{-i}{\hbar}\right) \int_{t_0}^{t} dt'\, \hat{H}(t') \left(\frac{-i}{\hbar}\right) \int_{t_0}^{t'} dt''\, \hat{H}(t'')\, \hat{U}(t'', t_0) \,.$$

By repeating the substitution recursively a new term involving increasing powers of $\hat{H}$ is added at each iteration. Assuming that the process converges, one obtains the Dyson series

$$\hat{U}(t, t_0) = 1 + \sum_{n=1}^{\infty} \left(\frac{-i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_{n-1}} dt_n\, \hat{H}(t_1)\, \hat{H}(t_2) \ldots \hat{H}(t_n) \,. \tag{3.12}$$

Notice that the integrand is *time ordered*, i.e., $t_1 > t_2 > \ldots > t_n$. According to the usual operator-algebra convention, the operators $\hat{H}$ having the lowest times act first since they appear to the right. This is a consequence of causality, as rooted in the composition or representation property of $\hat{U}(t, t_0)$, namely,

$$\hat{U}(t_1, t_3) = \hat{U}(t_1, t_2)\, \hat{U}(t_2, t_3)$$

for $t_1 > t_2 > t_3$.

Alternatively, we can derive $\hat{U}(t,t_0)$ by using successively the group-representation or composition property of $\hat{U}$ and the fact that $\hat{H}(t)$ is the generator of an infinitesimal time displacement at time t. We can in fact write $\hat{U}(t,t_0)$ as the succession of n short-time propagations:

$$\hat{U}(t,t_0) = \lim_{n\to\infty} \hat{U}(t_n,t_{n-1})\,\hat{U}(t_{n-1},t_{n-2})\,\ldots\hat{U}(t_2,t_1)\,\hat{U}(t_1,t_0) \tag{3.13}$$

with $t = t_n \geq t_{n-1} \geq \ldots t_2 \geq t_1 \geq t_0$. Using that $\hat{U}(t_j,t_{j-1}) = 1 - \frac{i}{\hbar}\hat{H}(t_j)\,\delta t_j$, where $\delta t_j = t_j - t_{j-1} = (t-t_0)/n$ tends to 0 in the limit of $n\to\infty$, and substituting it in Eq. (3.13) on obtains

$$\begin{aligned}
\hat{U}(t,t_0) &= \lim_{n\to\infty}\left\{\left(1-\frac{i}{\hbar}\,\hat{H}(t_n)\,\delta t_n\right)\left(1-\frac{i}{\hbar}\,\hat{H}(t_{n-1})\,\delta t_{n-1}\right)\ldots\right. \\
&\qquad\left.\ldots\left(1-\frac{i}{\hbar}\,\hat{H}(t_2)\,\delta t_2\right)\left(1-\frac{i}{\hbar}\,\hat{H}(t_1)\,\delta t_1\right)\right\} \\
&= 1+\sum_j\left(\frac{-i}{\hbar}\right)\hat{H}(t_j)\,\delta t_j+\sum_{j>i}\left(\frac{-i}{\hbar}\right)^2\hat{H}(t_j)\,\hat{H}(t_i)\,\delta t_i\,\delta t_j+ \qquad (3.14)\\
&\quad+\sum_{j>i>k}\left(\frac{-i}{\hbar}\right)^3\hat{H}(t_j)\,\hat{H}(t_i)\,\hat{H}(t_k)\,\delta t_i\,\delta t_j\,\delta t_k+\ldots \\
&= 1-\left(\frac{i}{\hbar}\right)\int_{t_0}^{t}\hat{H}(t')dt'+\left(\frac{-i}{\hbar}\right)^2\int_{t_0}^{t}\hat{H}(t')dt'\int_{t_0}^{t'}dt''\hat{H}(t'')+ \\
&\quad+\left(\frac{-i}{\hbar}\right)^3\int_{t_0}^{t}\hat{H}(t')dt'\int_{t_0}^{t'}dt''\hat{H}(t'')\int_{t_0}^{t''}dt'''\hat{H}(t''')+\ldots, \qquad (3.15)
\end{aligned}$$

which coincides with the expression (3.12) derived by iteration of the integral equation (3.11) for $\hat{U}(t,t_0)$.

The complex time-ordered Dyson equation (3.12) or (3.15) can be significantly simplified if the Hamiltonian $\hat{H}(t)$ commutes with itself at different times, i.e., if

$$[\hat{H}(t),\hat{H}(t')] = 0 \quad \forall\, t,t', \tag{3.16}$$

since in this case the time ordering is irrelevant.[5] Taking into account that there are $n!$ different and mutually exclusive ways of ordering the time variables $t_1,\ldots t_n$, we obtain

$$\begin{aligned}
&\int_{t_0}^{t}dt_1\int_{t_0}^{t_1}dt_2\ldots\int_{t_0}^{t_n}dt_n\,\hat{H}(t_1)\,\hat{H}(t_2)\ldots\hat{H}(t_n) = \\
&\qquad = \frac{1}{n!}\int_{t_0}^{t}dt_1\ldots\int_{t_0}^{t}dt_n\,\hat{H}(t_1)\ldots\hat{H}(t_n) = \frac{1}{n!}\left[\int_{t_0}^{t}\hat{H}(t')\,dt'\right]^n.
\end{aligned}$$

[5] Find nontrivial physical examples of time-dependent perturbations satisfying the condition (3.16) as well as a few which do not.

Consequently, if $[\hat{H}(t), \hat{H}(t')] = 0$, we may write

$$\hat{U}(t, t_0) = \exp\left\{ \frac{-i}{\hbar} \int_{t_0}^{t} \hat{H}(t')\, dt' \right\}$$

which, in the particular case of a time-independnet $\hat{H}$, yields the well-known expression

$$\hat{U}(t, t_0) = e^{-\frac{i}{\hbar}\hat{H}(t-t_0)}.$$

What does it physically mean that $[\hat{H}(t), \hat{H}(t')] = 0 \ \ \forall\, t, t'$? Consider an eigenstate of $\hat{H}(t_0)$: $\hat{H}(t_0)\,|\, n\, t_0, t_0 \,\rangle = E_n(t_0)\,|\, n\, t_0, t_0 \,\rangle$. Then $[\hat{H}(t), \hat{H}(t_0)] = 0$ implies that $\hat{H}(t)$ and $\hat{H}(t_0)$ can be diagonalized simultaneously. Therefore, $|\, n\, t_0, t_0 \,\rangle$ remains an eigenstate of $\hat{H}(t)$ for all $t \geq t_0$. In fact, $\hat{H}(t_0)\,\hat{H}^k(t)\,|\, n\, t_0, t_0 \,\rangle = E_n(t_0)\,\hat{H}^k(t)\,|\, n\, t_0, t_0 \,\rangle$ for all $k \in \mathbb{N}$. Therefore, $\hat{H}^k(t)\,|\, n\, t_0, t_0 \,\rangle$ and $\hat{U}(t, t_0)\,|\, n\, t_0, t_0 \,\rangle = |\, n\, t, t_0 \,\rangle$ are eigenstates of $\hat{H}(t_0)$ with the energy $E_n(t_0)$. Assuming for simplicity that the states are non-degenerate, we conclude that $\hat{H}(t)\,|\, n\, t_0, t_0 \,\rangle = \alpha\,|\, n\, t_0, t_0 \,\rangle$ with $\alpha \in \mathbb{R}$. Consequently, $|\, n\, t_0, t_0 \,\rangle$ is also an eigenstate of $\hat{H}(t)$ for all $t \geq t_0$. Moreover, since $[\hat{H}(t_0), \hat{U}(t, t_0)] = 0 \ \ \forall\ t \geq t_0 \ \Rightarrow\ |\, n\, t_0, t \,\rangle = \hat{U}(t, t_0)\,|\, n\, t_0, t_0 \,\rangle$ is an eigenstate of $\hat{H}(t_0)$ for all $t \geq t_0$. This explains the simple form of $\hat{U}(t, t_0) = e^{-i/\hbar \int_{t_0}^{t} \hat{H}(t')dt'}$ and the fact that the eigenstates of $\hat{H}(t_0)$ are not modified by the time evolution, except for their phases and eigenenergies. They remain stationary states at all times t.

3.11 Feynman's ordering technique

In the most general case where $[\hat{H}(t), \hat{H}(t')] \neq 0$, for example, a spin-1/2 particle in a magnetic field which changes direction as a function of t, we must use Dyson's equation in its full complexity. In order to simplify the operator calculus, Feynman introduced the *time ordering operator* which is defined by

$$\hat{\mathcal{T}}\{\hat{H}(t_1)\,\hat{H}(t_2)\} = \begin{cases} \hat{H}(t_1)\,\hat{H}(t_2) \text{ if } t_1 > t_2, \text{ and} \\ \hat{H}(t_2)\,\hat{H}(t_1) \text{ if } t_2 > t_1. \end{cases}$$

Note that $\hat{\mathcal{T}}\{\hat{H}(t_1)\,\hat{H}(t_2)\} = \hat{\mathcal{T}}\{\hat{H}(t_2)\,\hat{H}(t_1)\}$, which means that the order of the operators inside the time-ordering operator $\hat{\mathcal{T}}$ is immaterial. Consequently, as long as the operators remain within $\hat{\mathcal{T}}$, one may make all usual manipulations of commuting algebra and in particular take advantage of all know algebraic relations, such as $e^A e^B = e^{A+B}$, as if $\hat{H}(t)$ would commute at different times. Only at the end of the manipulations, when the time-ordering operator needs to be removed in order to recover the conventional notation, the time order must be disentangled.

Using $\hat{\mathcal{T}}$ we can write $\hat{U}(t,t_0)$ in a more compact form:

$$\begin{aligned}\hat{U}(t,t_0) &= 1+\sum_{n=1}^{\infty}\left(\frac{-i}{\hbar}\right)^n\frac{1}{n!}\,\hat{\mathcal{T}}\left\{\int_{t_0}^{t}dt_1\ldots\int_{t_0}^{t}dt_n\,\hat{H}(t_1)\ldots\hat{H}(t_n)\right\}\\ &= 1+\hat{\mathcal{T}}\left\{\sum_{n=1}^{\infty}\left(\frac{-i}{\hbar}\right)^n\frac{1}{n!}\left[\int_{t_0}^{t}dt'\,\hat{H}(t')\right]^n\right\}\\ &= \hat{\mathcal{T}}\left\{\exp\left[\frac{-i}{\hbar}\int_{t_0}^{t}dt'\,\hat{H}(t')\right]\right\}.\end{aligned}$$

The definition of $\hat{\mathcal{T}}$ allows all algebraic manipulations of $\hat{H}(t)$ inside the time ordering operator, since the order in which the operators act is not given by the order in which they are written but by the time t in $\hat{H}(t)$, i.e., the time at which they act [$\hat{\mathcal{T}}\{\hat{H}(t_1)\,\hat{H}(t_2)\} = \hat{\mathcal{T}}\{\hat{H}(t_2)\,\hat{H}(t_1)\}$]. Note that at the end of the calculation the expressions need to be disentangled by applying the definition of $\hat{\mathcal{T}}$ in order to recover the conventional notation in which the operators written at right act first. The interested reader should not hesitate to read and enjoy Feynman's original paper, which is remarkably well written and easy to follow, at least for the first part [see R. P. Feynman, Phys. Rev. **84**, 108 (1951)].

3.12 Transformation of Observables

So far we have mainly discussed the effect of symmetry transformations on the physical states. Let us now focus on how symmetry transformations modify the operators associated to observables. For this purpose we consider a transformation $\hat{S}$ connecting two equivalent descriptions: $|\,\alpha'\,\rangle = \hat{S}\,|\,\alpha\,\rangle$ and $|\,\beta'\,\rangle = \hat{S}\,|\,\beta\,\rangle$ for any kets $|\,\alpha\,\rangle$ and $|\,\beta\,\rangle$. Furthermore, we assume in the following that $\hat{S}$ is unitary, not antiunitary, although not necessarily continuous. Given an observable $\hat{A}$, the matrix elements between the transformed states (e.g., time displaced, translated, rotated, reflected, etc.) are

$$\langle\,\alpha'\,|\,\hat{A}\,|\,\beta'\,\rangle = \langle\,\alpha\,|\,\hat{S}^\dagger\hat{A}\,\hat{S}\,|\,\beta\,\rangle.$$

Instead of transforming the states $|\,\alpha\,\rangle \to |\,\alpha'\,\rangle$ we can transform the observable $\hat{A} \to \hat{A}_T$ according to

$$\hat{A}_T = \hat{S}^\dagger\hat{A}\,\hat{S}, \tag{3.17}$$

which satisfies the defining relation

$$\langle\,\alpha\,|\,\hat{A}_T\,|\,\beta\,\rangle = \langle\,\alpha'\,|\,\hat{A}\,|\,\beta'\,\rangle. \tag{3.18}$$

This means that an observer O' regarding some property A of the transformed system can use either the operator $\hat{A}$ with his/her own transformed kets $|\,\alpha'\,\rangle$ and $|\,\beta'\,\rangle$ or, equivalently, the original kets $|\,\alpha\,\rangle$ and $|\,\beta\,\rangle$ and the transformed operator $\hat{A}_T$ according to the unitary transformation (3.17).

Example:
Consider a state $|\,\Psi\,\rangle$ of a single particle in $\mathbb{R}^3$ which is rotated around the z axis by an angle φ. As usual, positive φ corresponds to a counterclockwise rotation when one looks from the positive z axis towards the origin. The state $|\,\Psi\,\rangle$ is arbitrary. If useful, it can be pictured by its wave function $\Psi(\vec{r}) = \langle\,\vec{r}\,|\,\Psi\,\rangle$, which can be an orbital pointing along the positive x axis, for example. The rotation leads to the state $|\,\Psi'\,\rangle = \hat{S}|\,\Psi\,\rangle$, where $\hat{S} = e^{-\frac{i}{\hbar}\hat{L}_z\,\varphi}$, as we shall see later in this chapter. In coordinate representation $\Psi'(\vec{r}) = \Psi(R^{-1}\vec{r})$ where R is the corresponding SO(3) rotation matrix. We would like to compute the average linear momentum

$$\vec{p}\,' = \langle\,\Psi'\,|\,\hat{\vec{p}}\,|\,\Psi'\,\rangle$$

or the dipole moment

$$\vec{r}\,' = \langle\,\Psi'\,|\,\hat{\vec{r}}\,|\,\Psi'\,\rangle$$

in the rotated state $|\,\Psi'\,\rangle$ without calculating $|\,\Psi'\,\rangle$ explicitly, i.e., in terms of averages of some appropriate operator in the original sate $|\,\Psi\,\rangle$. How can we do that? Well, according to the above discussion [Eqs. (3.17) and (3.18)] we know that the operators we are looking for are

$$\hat{\vec{p}}_T = \hat{S}^\dagger\,\hat{\vec{p}}\,\hat{S}. \tag{3.19}$$

and

$$\hat{\vec{r}}_T = \hat{S}^\dagger\,\hat{\vec{r}}\,\hat{S}. \tag{3.20}$$

These equations are to be understood as vector identities (i.e., $\hat{p}_{\alpha\,T} = \hat{S}^\dagger\,\hat{p}_\alpha\,\hat{S}$ and $\hat{r}_{\alpha\,T} = \hat{S}^\dagger\,\hat{r}_\alpha\,\hat{S}$ for $\alpha = x$, y and z).

Calculating $\hat{\vec{p}}_T$ according to Eq. (3.20) involves some algebra and, admittedly, one does not quite understand what one is actually doing with this operator transformation. In fact, it should be possible to guess the precise form of $\hat{\vec{p}}_T$ in terms of $\hat{\vec{p}} = -i\hbar\vec{\nabla}$ by means of simple geometrical arguments, without any operator algebra. One could make a drawing of how the vectors $\hat{\vec{p}}$ and $\hat{\vec{r}}$, and the averages thereof, should change upon rotation. With that in hand one could calculate $\hat{\vec{p}}_T$ according to Eq. (3.20) explicitly by using $\hat{S} = e^{-\frac{i}{\hbar}\hat{L}_z\,\varphi}$. An elegant way to do this is to consider a finite rotation φ followed by an infinitesimal one $\delta\varphi$ in order to obtain two coupled differential equations for $\hat{p}_{xT}$ and $\hat{p}_{yT}$.

Exercise:
How do the explicit calculation of $\hat{\vec{p}}_T$ and the geometrical guess compare? Is it possible to explain in simple words what the transformed operator $\hat{\vec{p}}_T$ represents? The qualifier "transformed" is certainly general, since it apply to any type of transformation $\hat{S}$. How could one call $\hat{\vec{p}}_T$ more precisely in the present case? Would a similar reasoning apply to other vector properties, for example, in the case of the dipole moment $\vec{r}\,' = \langle\,\Psi'\,|\,\hat{\vec{r}}\,|\,\Psi'\,\rangle$ of the rotated state $|\,\Psi'\,\rangle$? These questions are discussed in some detail in Sec. 3.30.

3.13 Invariant observables

An observable $\hat{A}$ is said to be *invariant under the transformation* $\hat{S}$ if and only if

$$\hat{A}_T = \hat{S}^\dagger \hat{A}\, \hat{S} = \hat{A}.$$

This is equivalent to the commutation relation

$$\hat{A}\, \hat{S} - \hat{S}\, \hat{A} = [\hat{A}, \hat{S}] = 0,$$

which applies to both continuous and discrete unitary transformations $\hat{S}$. Physically, the invariance of an observable A upon $\hat{S}$ means that all the matrix elements of the operator $\hat{A}$ between the transformed states $|\,\alpha'\,\rangle$ and $|\,\beta'\,\rangle$ are the same as between the original states $|\,\alpha\,\rangle$ and $|\,\beta\,\rangle$.

In the case of continuous transformations it is interesting to consider the changes in the observables associated with an infinitesimal parameter change δs. Expressing

$$\hat{S}(\delta s) = 1 - \frac{i}{\hbar}\, \hat{G}\, \delta s$$

in terms of the generator $\hat{G}$, and using Eq. (3.17), we obtain

$$\hat{A}_T = \hat{S}^\dagger \hat{A}\, \hat{S} = \left(1 + \frac{i}{\hbar}\, \hat{G}\, \delta s\right)\, \hat{A}\, \left(1 - \frac{i}{\hbar}\, \hat{G}\, \delta s\right),$$

which to leading order in δs reads

$$\hat{A}_T = \hat{A} + \frac{i}{\hbar}[\hat{G}, \hat{A}]\, \delta s.$$

Consequently, an observable $\hat{A}$ is invariant under the transformations generated by $\hat{G}$ if and only if $[\hat{G}, \hat{A}] = 0$. For example, an observable commuting with $\hat{\vec{P}}$, $\hat{\vec{J}}$ or $\hat{H}$ is invariant under translations, rotations or time evolution, respectively.

Exercise:
Find nontrivial examples of observables $\hat{A}$ and specific physical situations which are invariant under either translations, rotations or time evolution. Verify that $[\hat{A}, \hat{G}] = 0$ for the corresponding generators $\hat{G} = \hat{\vec{P}}$, $\hat{\vec{J}}$ and $\hat{H}$.

3.14 Heisenberg and Schrödinger pictures

The most famous example of the above discussed mathematical freedom of transforming kets or operators is probably the equivalence between Heisenberg's and Schrödinger's description of time evolution. In the Schrödinger picture the time evolution of the states is given by

$$|\,\alpha\, t_0, t\,\rangle_S = \hat{U}(t, t_0)\,|\,\alpha\,\rangle$$

where $|\,\alpha\,\rangle \equiv |\,\alpha\, t_0, t_0\,\rangle$. The matrix elements of $\hat{A}$ at time t are

$$\langle\, \alpha\, t_0, t\,|\,\hat{A}\,|\,\beta\, t_0, t\rangle_S = \langle\,\alpha\,|\,\hat{U}^\dagger(t,t_0)\,\hat{A}\,\hat{U}(t,t_0)\,|\,\beta\,\rangle$$

and therefore the transformed operator $\hat{A}_H$ is given by

$$\hat{A}_H(t,t_0) = \hat{U}^\dagger(t,t_0)\,\hat{A}\,\hat{U}(t,t_0). \tag{3.21}$$

The average values can be calculated using any of the two representations

$$\langle\,\alpha(t)\,|\,\hat{A}_S\,|\,\alpha(t)\rangle_S = \langle\,\alpha\,|\,\hat{A}_H(t)\,|\,\alpha\rangle_H,$$

where $\hat{A}_S = \hat{A}$ is the operator in Schrödinger's picture, which can show an explicit time dependence, $|\,\alpha\rangle_S = |\,\alpha\, t_0, t\rangle_S$ is the time-dependent Schrödinger ket, $\hat{A}_H$ is the operator $\hat{A}$ in the Heisenberg representation, and $|\,\alpha\rangle_H = |\,\alpha\, t_0, t_0\rangle_S$ is the time-independent Heisenberg ket.

The Heisenberg equation of motion for the operator $\hat{A}_H$ is

$$i\hbar\frac{d\hat{A}_H}{dt}(t,t_0) = i\hbar\frac{\partial\hat{A}_H}{\partial t}(t,t_0) + [\hat{A}_H(t,t_0), \hat{H}_H(t,t_0)]\,,$$

where

$$\frac{\partial\hat{A}_H}{\partial t}(t,t_0) = \hat{U}^\dagger(t,t_0)\,\frac{\partial\hat{A}_S(t)}{\partial t}\,\hat{U}(t,t_0)$$

is the Heisenberg operator associated to the partial derivative (explicit time dependence) of the Schrödinger operator $\hat{A}$. The Hamilton operator in the Heisenberg representation is given, as for any other Heisenberg operator, by

$$\hat{H}_H(t,t_0) = \hat{U}(t,t_0)^\dagger\,\hat{H}_S\,\hat{U}(t,t_0).$$

If $\hat{H}_S$ commutes with $\hat{U}(t,\, t_0)$, for example, when $\hat{H}$ commutes with itself at different times, we have simply $\hat{H}_H = \hat{H}_S = \hat{H}$.

In order to derive the Heisenberg equation of motion we have taken the time derivative of Eq. (3.21) and used Eq. (3.9) for $\hat{U}$ and $\hat{U}^\dagger$, which gives

$$i\,\hbar\,\frac{\partial\hat{U}}{\partial t} = \hat{H}(t)\,\hat{U}(t,t_0)$$

and

$$-i\,\hbar\,\frac{\partial\hat{U}^\dagger}{\partial t} = \hat{U}^\dagger(t,t_0)\,\hat{H}(t).$$

This implies

$$\begin{aligned} i\,\hbar\,\frac{d\hat{A}_H}{dt} &= i\,\hbar\,\hat{U}^\dagger\,\frac{\partial\hat{A}_S}{\partial t}\,\hat{U} - \hat{U}^\dagger\hat{H}\,\hat{U}\hat{U}^\dagger\,\hat{A}_S\hat{U} + \hat{U}^\dagger\hat{A}_S\,\hat{U}\hat{U}^\dagger\,\hat{H}\,\hat{U} \\ &= i\,\hbar\frac{\partial\hat{A}_H}{\partial t} + (\hat{A}_H\,\hat{H}_H - \hat{H}_H\,\hat{A}_H). \end{aligned}$$

3.15 Mixed states and density matrices

In order to avoid possible confusions, it is useful to discuss at this point mixed ensembles and their time dependence, a subject familiar from statistical mechanics. So far we have considered pure states $|\,\alpha\,\rangle$, $|\,\beta\,\rangle$, etc. of a QM system and investigated how these states transform under different symmetry operations or time evolution. However, this is certainly not the only physical situation one may encounter. Suppose we consider an *incoherent mixture* of states $|\,\alpha_i\,\rangle$ each having a probability w_i but bearing no phase correlation whatsoever. This is called a *mixture* or a *mixed ensemble*. The fractional population of each of the pure states $|\,\alpha_i\,\rangle$ satisfies the normalization condition $\sum_i w_i = 1$. We also require $\langle\,\alpha_i\,|\,\alpha_i\,\rangle = 1 \quad \forall\, i$. However, note that the $|\,\alpha_i\,\rangle$ need not be orthogonal to each other. For example, we have complete freedom to combine in an electron beam

i) 30 % of electrons polarized with spin up in the x direction: $|\,\alpha_1\,\rangle = \dfrac{|+\rangle + |-\rangle}{\sqrt{2}}$

ii) 30 % spin down along z: $|\,\alpha_2\,\rangle = |-\rangle$

iii) 20 % spin up along z: $|\,\alpha_3\,\rangle = |+\rangle$, and

iv) 20 % spin up along y: $|\,\alpha_4\,\rangle = \dfrac{|+\rangle + i\,|-\rangle}{\sqrt{2}}$.

The average of observable $\hat{A}$ when a large number of measurements is done on such a mixed ensemble is given by

$$\langle\langle\,\hat{A}\,\rangle\rangle = \sum_i w_i\,\langle\,\alpha_i\,|\,\hat{A}\,|\,\alpha_i\,\rangle,$$

where $\langle\,\alpha_i\,|\,\hat{A}\,|\,\alpha_i\,\rangle$ is the usual QM average of observable $\hat{A}$ in the pure state $|\,\alpha_i\,\rangle$ and $\langle\langle\ \ldots\ \rangle\rangle$ is the statistical average of $\langle\,\alpha_i\,|\,\hat{A}\,|\,\alpha_i\,\rangle$ over the ensemble with fractional populations w_i.

In order to express $\langle\langle\,\hat{A}\,\rangle\rangle$ in an invariant form we introduce in the previous equation $\sum_a |\,a\,\rangle\langle\,a\,| = \sum_{a'} |\,a'\,\rangle\langle\,a'\,| = 1$ and rewrite it as

$$\begin{aligned}
\langle\langle\,\hat{A}\,\rangle\rangle &= \sum_{\substack{i\\ a,a'}} w_i\,\langle\,\alpha_i\,|\,a\,\rangle\langle\,a\,|\,\hat{A}\,|\,a'\,\rangle\langle\,a'\,|\,\alpha_i\,\rangle \\
&= \sum_{a,a'}\langle\,a'\,|\,\left(\sum_i w_i\,|\,\alpha_i\,\rangle\langle\,\alpha_i\,|\right)|\,a\,\rangle\langle\,a\,|\,\hat{A}\,|\,a'\,\rangle \\
&= \sum_a \langle\,a\,|\,\hat{A}\,\underbrace{\left(\sum_i w_i\,|\,\alpha_i\,\rangle\langle\,\alpha_i\,|\right)}_{\substack{\text{Depends only on}\\ \text{the ensemble}}}|\,a\,\rangle.
\end{aligned}$$

In a more compact form we have

$$\langle\langle\,\hat{A}\,\rangle\rangle = \mathrm{Tr}\{\hat{\varrho}\,\hat{\mathrm{A}}\}$$

where

$$\hat{\varrho} = \sum_i w_i \,|\, \alpha_i \,\rangle\langle\, \alpha_i \,|$$

is the *density matrix operator*, which depends only on the ensemble, i.e., on the involved pure states $|\, \alpha_i \,\rangle$ and the corresponding weights $w_i \geq 0$ $\big(\sum_i w_i = 1$ and $\langle\, \alpha_i \,|\, \alpha_i \,\rangle = 1\big)$.

The density operator $\hat{\varrho}$ satisfies the following important properties:

1) $\hat{\varrho}^\dagger = \hat{\varrho}$, since $w_i \geq 0$.

2) $\mathrm{Tr}\,\hat{\varrho} = 1$, since $\mathrm{Tr}\,\hat{\varrho} = \sum_{\mathrm{i}} \mathrm{w_i} \sum_{\mathrm{a}} \langle\, \mathrm{a} \,|\, \alpha_{\mathrm{i}} \,\rangle\langle\, \alpha_{\mathrm{i}} \,|\, \mathrm{a} \,\rangle \;=\; \sum_{\mathrm{i}} \mathrm{w_i} \underbrace{\langle\, \alpha_i \,|\, \alpha_i \,\rangle}_{=1} = \sum_{\mathrm{i}} \mathrm{w_i} = 1.$

3) The eigenvalues λ_k of $\hat{\varrho}$ satisfy $0 \leq \lambda_k \leq 1$. This is not so obvious since the $|\, \alpha_i \,\rangle$ need not be orthogonal to each other. Note that the w_i are <u>not</u> the eigenvalues of $\hat{\varrho}$. However, it can easily be proven by noting that

$$\hat{\varrho} \,|\, \Psi_k \,\rangle = \lambda_k \,|\, \Psi_k \,\rangle \quad \Leftrightarrow \quad \sum_i w_i \,|\, \alpha_i \,\rangle\, \langle\, \alpha_i \,|\, \Psi_k \,\rangle = \lambda_k \,|\, \Psi_k \,\rangle$$

$$\Rightarrow \lambda_k = \sum_i w_i \,|\, \langle\, \Psi_k \,|\, \alpha_i \,\rangle \,|^{\,2} \geq 0.$$

In other words, $\hat{\varrho}$ is positive definite, since $\langle\, \Psi \,|\, \hat{\varrho} \,|\, \Psi \,\rangle \geq 0 \quad \forall \,|\, \Psi \,\rangle$. Moreover, taking into account that $\sum_k \lambda_k = \mathrm{Tr}\,\hat{\varrho} = 1 \quad \Rightarrow \quad 0 \leq \lambda_{\mathrm{k}} \leq 1.$

4) $\mathrm{Tr}\{\hat{\varrho}^2\} \leq 1$.

This can be proved by recalling that $\hat{\varrho}^2 = \sum_k \lambda_k^2 \,|\, \Psi_k \,\rangle\langle\, \Psi_k \,|$ and

$$\left(\sum_k \lambda_k\right)^2 = \sum_k \lambda_k^2 + \sum_{k \neq l} \lambda_k \, \lambda_l = 1.$$

Knowing that $\sum_{k \neq l} \lambda_k \, \lambda_l \geq 0$ we have

$$\mathrm{Tr}\,\hat{\varrho}^2 = \sum_{\mathrm{k}} \lambda_{\mathrm{k}}^2 \leq 1.$$

5) $\mathrm{Tr}\{\hat{\varrho}^2\} = 1 \Leftrightarrow \hat{\varrho} = |\, \alpha_{\mathrm{k}} \,\rangle\langle\, \alpha_{\mathrm{k}} \,|$.

In other words, $\mathrm{Tr}\{\hat{\varrho}^2\} = 1$ implies that $w_i = 0 \;\forall\; i \neq k$, and $w_k = 1$ for only one state k, i.e., $\hat{\varrho} = |\, \alpha_k \,\rangle\langle\, \alpha_k \,|$. Only in this case we have $\hat{\varrho}^2 = \hat{\varrho}$, which means that we are dealing with a pure state. This can easily be seen by noting that

$$\sum_k \lambda_k^2 = 1 \quad \Rightarrow \quad \sum_{k \neq l} \lambda_k \lambda_l = 0.$$

Thus, there is not a single pair of $\lambda_k\lambda_l \neq 0$ ($\lambda_k \geq 0\ \forall k$). Since all products $\lambda_k\lambda_l = 0$, we can only have one $\lambda_k \neq 0 \ \Rightarrow\ \lambda_k = 1$. Consequently, $\hat{\varrho} = |\,\alpha_k\,\rangle\langle\,\alpha_k\,|$, and therefore $\hat{\varrho}^2 = |\,\alpha_k\,\rangle\langle\,\alpha_k\,|$.

6) Consequently, $\mathrm{Tr}\{\hat{\varrho}^2\}$ allows us to distinguish pure states from mixtures. $\mathrm{Tr}\{\hat{\varrho}^2\}$ is not only independent of the representation, but also independent of time. Therefore, pure states cannot evolve into mixtures (and vice versa) as long as the dynamics follows a unitary time-evolution operator $\hat{U}(t,t_0)$. The proof is straightforward by noting that

$$\begin{aligned}
\hat{\varrho}^2 &= \sum_{ij} |\,\alpha_i\,\rangle\, w_i\, \langle\,\alpha_i\,|\,\alpha_j\,\rangle w_j \langle\,\alpha_j\,| \\
\mathrm{Tr}\{\hat{\varrho}^2\} &= \sum_a \sum_{ij} \langle\,a\,|\,\alpha_i\,\rangle\, w_i\, \langle\,\alpha_i\,|\,\alpha_j\,\rangle w_i \langle\,\alpha_j\,|\,a\,\rangle \\
&= \sum_{ij} \langle\,\alpha_j\,|\,\alpha_i\,\rangle\langle\,\alpha_i\,|\,\alpha_j\,\rangle w_i\, w_j.
\end{aligned}$$

Since $\hat{U}(t,t_0)$ is unitary it does not modify the scalar products $\langle\,\alpha_i\,|\,\alpha_j\,\rangle$ as a function of time.

7) Finally, one may say that the entropy of a system described by $\hat{\rho}$, which is given by $S = -k_B\langle\ln(\hat{\rho})\rangle = -k_B\mathrm{Tr}\{\hat{\rho}\,\ln(\hat{\rho})\}$ is also independent of time, as long as the time evolution is unitary. Moreover, $S = 0$ if and only if $\hat{\rho}$ describes a pure state, i.e., $\hat{\rho} = |\Psi_k\rangle\langle\Psi_k|$ for some k.

3.16 Time dependence of the density-matrix operator

We consider $\hat{\varrho}(t)$ in the Schrödinger picture:

$$\hat{\varrho}_S(t) = \sum_i w_i\, |\,\alpha_i\, t_0, t\,\rangle\langle\,\alpha_i\, t_0, t\,|,$$

where the Schrödinger kets depend on time as $|\,\alpha_i\, t_0, t\rangle_S = |\,\alpha_i\, t_0, t\,\rangle = \hat{U}(t,t_0)\,|\,\alpha_i\, t_0, t_0\,\rangle$. Knowing that

$$i\hbar\,\frac{\partial}{\partial t}\,|\,\alpha_i\, t_0, t\,\rangle = \hat{H}(t)\,|\,\alpha_i\, t_0, t\,\rangle$$

and

$$-i\hbar\,\frac{\partial}{\partial t}\langle\,\alpha_i\, t_0, t\,| = \langle\,\alpha_i\, t_0, t\,|\,\hat{H}(t),$$

we have

$$i\hbar\,\frac{\partial\hat{\varrho}_S}{\partial t} = \sum_i w_i \left(\hat{H}\,|\,\alpha_i\, t_0, t\,\rangle\,\langle\,\alpha_i\, t_0, t\,| \;-\; |\,\alpha_i\, t_0, t\,\rangle\,\langle\,\alpha_i\, t_0, t\,|\,\hat{H}\right)$$

$$i\hbar \frac{\partial \hat{\varrho}_S}{\partial t} = [\hat{H}, \hat{\varrho}_S(t)].$$

This is known as the *Liouville equation*. Note this is the equation for the time dependence of the operator $\hat{\varrho}_S$ in the *Schrödinger picture*. $\hat{\varrho}_S$ depends on t due to the time dependence of the Schrödinger kets $|\, \alpha_i\, t_0, t\,\rangle$.

As for any operator, in the Heisenberg picture we have

$$\hat{\varrho}_H(t) = \hat{U}^\dagger(t, t_0)\, \hat{\varrho}_S(t)\, \hat{U}(t, t_0).$$

Recalling that $\hat{A}_H = \hat{U}^\dagger(t, t_0)\, \hat{A}_S\, \hat{U}(t, t_0)$ we can calculate $\langle\langle\, \hat{A}\,\rangle\rangle$ in the Heisenberg picture as

$$\begin{aligned} \langle\langle\, \hat{A}\,\rangle\rangle &= \mathrm{Tr}\{\hat{\varrho}_{\mathrm{H}}\, \hat{\mathrm{A}}_{\mathrm{H}}\} = \mathrm{Tr}\{\hat{\mathrm{U}}^\dagger \hat{\varrho}_{\mathrm{S}}\, \hat{\mathrm{U}}\hat{\mathrm{U}}^\dagger \hat{\mathrm{A}}_{\mathrm{S}}\, \hat{\mathrm{U}}\} \\ &= \mathrm{Tr}\{\hat{\varrho}_{\mathrm{S}}\, \hat{\mathrm{A}}_{\mathrm{S}}\}, \end{aligned}$$

which coincides, as expected, with the Schrödinger result. Notice that in the Heisenberg picture

$$\begin{aligned} \hat{\varrho}_H &= \sum_i w_i \underbrace{\hat{U}^\dagger(t, t_0)|\, \alpha_i\, t_0, t\,\rangle}_{|\, \alpha_i\, t_0, t_0\,\rangle} \underbrace{\langle\, \alpha_i\, t_0, t\,|\, \hat{U}(t, t_0)}_{\langle\, \alpha_i\, t_0, t_0\,|} \\ &= \sum_i w_i\, |\, \alpha_i\, t_0, t_0\,\rangle\langle\, \alpha_i\, t_0, t_0\,| = \hat{\varrho}_S(t_0) \end{aligned}$$

is independent of t. In the case of statistical ensembles the density matrix takes the role of the wave function. It defines the mechanical state of the ensemble and all physical properties can be derived from it. As the wave function, $\hat{\rho}$ depends on t in the Schrödinger picture and is independent of t in the Heisenberg picture.

3.17 Symmetry, constants of motion and degeneracies

Consider a symmetry operator $\hat{S}$ whose infinitesimal transformations are generated by $\hat{G}$, i.e.,

$$\hat{S}(\delta s) = 1 - \frac{i}{\hbar}\, \hat{G}\, \delta s + O(\delta s^2)\,,$$

and suppose that the Hamiltonian is invariant under $\hat{S}$. This means that

$$\hat{S}^\dagger \hat{H}\, \hat{S} = \hat{H} \quad \Leftrightarrow \quad [\hat{H}, \hat{S}] = 0 \quad \Leftrightarrow \quad [\hat{H}, \hat{G}] = 0\,.$$

Consequently $[\hat{G}, \hat{U}(t, t_0)] = 0$. Assuming that $\hat{G}$ does not depend explicitly on time, we obtain in the Heisenberg picture that

$$\frac{d\hat{G}_H}{dt} = 0 \quad \text{or equivalently} \quad \hat{G}_H = \hat{G}_S.$$

One concludes that $\hat{G}$ is a constant of motion if $\hat{H}$ is invariant under the symmetry transformation $\hat{S}$ generated by $\hat{G}$.

Examples: i) $\hat{H}$ is invariant under translations $\Leftrightarrow \hat{\vec{P}}$ is a constant of motion.
ii) $\hat{H}$ is invariant under rotations $\Leftrightarrow \hat{\vec{J}}$ is a constant of motion.

One may easily verify that the eigenstates of $\hat{G}$ at $t = t_0$ remain eigenstates with the same eigenvalue for all times, when $\hat{H}$ is invariant under $\hat{G}$. Knowing that $[\hat{H}, \hat{G}] = 0 \ \ \forall\, t$, we also have $[\hat{U}(t, t_0), \hat{G}] = 0 \ \ \forall\, t$. Let $\hat{G}\,|\,g\,t_0, t_0\,\rangle = g\,|\,g\,t_0, t_0\,\rangle$. For any $t > t_0$ we have

$$|\,g\,t_0, t\,\rangle \;=\; \hat{U}(t, t_0)\,|\,g\,t_0, t_0\,\rangle$$

and therefore

$$\begin{aligned}
\hat{G}\,|\,g\,t_0, t\,\rangle &= \hat{G}\,\hat{U}(t, t_0)\,|\,g\,t_0, t_0\,\rangle \\
&= \hat{U}(t, t_0)\,\hat{G}\,|\,g\,t_0, t_0\,\rangle \\
&= g\,\hat{U}(t, t_0)\,|\,g\,t_0, t_0\,\rangle \\
&= g\,|\,g\,t_0, t\,\rangle.
\end{aligned}$$

Conversely, if $\hat{G}$ and $\hat{H}$ commute, the eigenstates of $\hat{H}$ remain eigenstates of $\hat{H}$ with the same energy when they are transformed according to $\hat{S}$. This leads to *symmetry induced degeneracies* in the energy spectrum. In order to prove this statement, let us consider an energy eigenstate $|\,n\,\rangle$ such that $\hat{H}\,|\,n\,\rangle = E_n\,|\,n\,\rangle$. From $[\hat{H}, \hat{G}] = 0$ it follows that $[\hat{H}, \hat{S}] = 0$ as well as

$$\hat{H}(\hat{S}\,|\,n\,\rangle) = \hat{S}\,\hat{H}\,|\,n\,\rangle = E_n\,\hat{S}\,|\,n\,\rangle.$$

Therefore $\hat{S}|\,n\,\rangle$ is also an eigenstate of $\hat{H}$ with the same energy E_n. If $\hat{S}\,|\,n\,\rangle = e^{i\gamma}|\,n\,\rangle$ with $\gamma \in \mathbb{R}$, nothing new is found, since $|\,n\,\rangle$ is itself invariant under $\hat{S}$, i.e., it is an eigenstate of $\hat{S}$. But if $\hat{S}\,|\,n\,\rangle$ and $|\,n\,\rangle$ are linearly independent, a degeneracy is revealed.

Very often $\hat{S}$ is a function of some parameter λ which characterizes or labels the different elements of the symmetry group which can be discrete or continuous. For example, λ may represent the extent and direction of a spatial displacement or the angle and axis of a rotation. It can also label the $N!$ permutations of the coordinates of N identical particles. Clearly, if $\hat{H}\,|\,n\,\rangle = E_n\,|\,n\,\rangle$, all $\hat{S}(\lambda)\,|\,n\,\rangle$ are stationary states with the same energy. The subspace $\mathcal{S}$ spanned by the kets $\hat{S}(\lambda)\,|\,n\,\rangle$ for all $\hat{S}(\lambda)$ in the symmetry group is obviously invariant and thus constitutes an appropriate vector space for a representation of the group. Notice that ensemble of kets $\hat{S}(\lambda)\,|\,n\,\rangle$ for different $\hat{S}(\lambda)$ in the group gives very often a linearly dependent set of vector, most certainly so when the order of the group is larger that the dimension of the underlying Hilbert space as, for example, in the case of continuous groups. However, even when linear dependences are removed and an orthonormal basis of states spanning the generated invariant subspace is obtain, it is not always obvious *a priori* to know whether the representation obtained by applying all $\hat{S}(\lambda)$ in the group to a given ket $|\,n\,\rangle$ is irreducible or not. This depends on the considered ket $|\,n\,\rangle$ and group, as illustrated by the examples below.

Examples:

Consider a Hamiltonian which is rotationally invariant, i.e., $[\hat{D}(R), \hat{H}] = 0$ for all rotations $R \in \mathrm{SO}(3)$. Since

$$\hat{D}(R) = e^{\frac{-i}{\hbar}\hat{\vec{J}}\cdot\hat{n}\,\phi}\,,$$

this implies $[\hat{H}, \hat{\vec{J}}] = 0$ and therefore $[\hat{H}, \hat{J}_\pm] = [\hat{H}, \hat{J}^2] = 0$. The fact that $[\hat{H}, \hat{J}^2] = 0$, $[\hat{H}, \hat{J}_z] = 0$ and $[\hat{J}_z, \hat{J}^2] = 0$ allows us to form simultaneous eigenstates of $\hat{H}$, $\hat{J}^2$ and $\hat{J}_z$ which we denote by $|\, n\, j\, m\,\rangle$. The previous symmetry arguments tell us that

$$\hat{D}(R)\,|\, n\, j\, m\,\rangle$$

have all the same energy for any R. As usual, infinitesimal rotations are enough to clarify the situation, since $\hat{J}_x\,|\, n\, j\, m\,\rangle$, $\hat{J}_y\,|\, n\, j\, m\,\rangle$ and $\hat{J}_z\,|\, n\, j\, m\,\rangle$ must have the same energy. Using that $[\hat{J}_z, \hat{H}] = 0$ does not help us any further, since applying $\hat{J}_z$ to any ket $|\, n\, j\, m\,\rangle$ gives us $\hat{J}_z\,|\, n\, j\, m\,\rangle = m\,|\, n\, j\, m\,\rangle$, which is proportional to $|\, n\, j\, m\,\rangle$. In contrast, the operators $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$ are extremely useful since

$$\hat{J}_\pm\,|\, n\, j\, m\,\rangle = \lambda\,|\, n\, j\, m \pm 1\,\rangle \quad \text{with} \quad \lambda = \sqrt{(j \pm m + 1)(j \mp m)}\,,$$

where we have used, as assumed, that the initial state $|\, n\, j\, m\,\rangle$, and thus all the states obtained by rotation of it, have a well-defined total angular momentum $\hat{J}^2$ ($\hat{J}^2|\, n\, j\, m\,\rangle = j(j+1)|\, n\, j\, m\,\rangle$). Therefore, all $2j+1$ states having $m = -j, \ldots j$ are degenerate. Clearly, this $(2j+1)$-dimensional representation of the SO(3) group is irreducible. Indeed, starting from *any single vector* in this subspace (i.e., any linear combination of states $|\, n\, j\, m\,\rangle$) all other $|\, n\, j\, m'\,\rangle$ with $m' \neq m$ are obtained by performing arbitrary rotations (even infinitesimal rotations) in SO(3) and linear combinations thereof. In other words, there is no smaller, strictly included subspace which remains invariant once all SO(3) rotations are applied.

It is worth noting that the situation is quite different when the symmetry is lower, i.e., when only a smaller subgroup of rotations leaves the Hamiltonian unchanged. For example, suppose that the system is invariant only upon the rotations R_z around the z axis and consider a starting state $|n\, j\rangle$ that is not an eigenstate of $\hat{J}_x$, $\hat{J}_y$ or $\hat{J}_z$ but still has a well defined $\hat{J}^2$. To be more explicit, let us assume that $|n\, j\rangle$ is an eigenstate of $\hat{\vec{J}}\cdot\hat{n}$ with maximal projection $m = j$ for some tilted direction $\hat{n}$, for example, a direction within the xz plane. It is then geometrically clear that the rotations around the z axis yield more than one linearly independent state, in some cases maybe even enough to obtain all $2j + 1$ states $|\, n\, j\, m\,\rangle$ having $m = -j, \ldots j$ (e.g., in the case of p orbitals having $j = 1$). All these states will have the same energy as the initial one. However, it is also clear that the thus obtained representation is reducible, since the group of rotations around the z axis is Abelian and therefore all its irreducible representations have dimension 1. In fact, one may construct linear combinations giving an eigenstates $|\, n\, j\, m_z\,\rangle$ of $\hat{J}_z$ with eigenvalue m_z which is invariant upon any R_z and thus corresponds to a non-trivial invariant subspace.

Another interesting example is the case of permutations of the particle coordinates in wavefunctions, which is most relevant in the context of identical particles. Consider a simple-product wavefunction $\Psi(x_1, x_2, \ldots, x_N) = \varphi_1(x_1)\varphi_1(x_2)\ldots\varphi_1(x_N)$ and the group of all permutations P. If one applies all permutations of the coordinates to Ψ (i.e., $\hat{P}\Psi(x_1, x_2, \ldots, x_N) = \varphi_1(x_{P(1)})\varphi_1(x_{P(2)})\ldots\varphi_1(x_{P(N)})$ one obtains functions which have the same energy, since the Hamiltonian of identical particles is independent of the particle numbering. Clearly, the vector space spanned by the resulting $N!$ functions is invariant with respect to all permutations. Therefore, it constitutes an appropriate representation space of the permutation group. Still, it is also easy to see that this representation is reducible, since the fully symmetrized and fully antisymmetrized products are well-known invariant subspaces, yielding each of them a one-dimensional representation.

3.18 Spatial translations

We consider the group of all spatial translations which are parameterized by the displacement vector $\vec{a} \in \mathbb{R}^3$. The representation of the translation group is given by the operator $\hat{U}(a)$ which relates any function $\Psi(\vec{r})$ with the translated one

$$\Psi'(\vec{r}) = \hat{U}(\vec{a})\,\Psi(\vec{r}) = \Psi(\vec{r} - \vec{a}).$$

Clearly $\hat{U}(\vec{a})$ is linear, preserves the norm of Ψ and is therefore unitary. It satisfies the usual group-representation properties:

$$\hat{U}(\vec{0}) = \mathbb{1},$$

and

$$\hat{U}(\vec{a}_1 + \vec{a}_2) = \hat{U}(\vec{a}_1)\,\hat{U}(\vec{a}_2),$$

which implies

$$\hat{U}(-\vec{a}) = \hat{U}(\vec{a})^{-1} = \hat{U}(a)^{\dagger}.$$

The translation group is an Abelian group, since the sum of vectors in $\mathbb{R}^3$ is commutative. Accordingly, $\hat{U}(\vec{a}_1 + \vec{a}_2) = \hat{U}(\vec{a}_2 + \vec{a}_1)$ and all the operators $\hat{U}$ in the representation commute with each other.

The generator of infinitesimal translations is the momentum operator $\hat{\vec{p}}$. For small translations $\delta\vec{a}$ we have

$$\Psi'(\vec{r}) = \Psi(\vec{r} - \delta\vec{a}) = \Psi(\vec{r}) - \vec{\nabla}\Psi(\vec{r}) \cdot \delta\vec{a} + \mathcal{O}(\delta a^2),$$

which implies

$$\hat{U}(\delta\vec{a}) = 1 - \delta\vec{a} \cdot \vec{\nabla} = 1 - \frac{i}{\hbar}\,\delta\vec{a} \cdot \hat{\vec{p}}.$$

Since translations along different directions commute, all momentum components commute:

$$[\hat{p}_i, \hat{p}_j] = 0 \quad \forall\ i, j = x, y \,\text{and}\, z\,.$$

This well-known property, usually inferred from the commutation of partial derivatives, appears here as consequence of the fact that $\hat{\vec{p}}$ is the generator of translations, which is independent of its coordinate representation.

Taking into account that $\hat{\vec{p}} = -i\hbar\vec{\nabla}$ is independent of $\vec{a}$, and using Eq. (3.5) we easily obtain the representation of a finite translation $\vec{a}$ as

$$\hat{U}(\vec{a}) = e^{-\frac{i}{\hbar}\hat{\vec{p}}\cdot\vec{a}}. \tag{3.22}$$

The translated kets $|\,\alpha, \vec{a}\,\rangle$ are related to the original $|\,\alpha\,\rangle$ by

$$|\,\alpha, \vec{a}\,\rangle = \hat{U}(\vec{a})\,|\,\alpha\,\rangle = e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}}\,|\,\alpha\,\rangle\,.$$

This generalizes to arbitrary kets the relation between translated wavefunctions $\Psi'_\alpha(\vec{r}) = \Psi_\alpha(\vec{r} - \vec{a}) = \langle\,\vec{r}\,|\,\alpha\,\vec{a}\,\rangle$ and the original $\Psi_\alpha(\vec{r}) = \langle\,\vec{r}\,|\,\alpha\,\rangle$. In particular the eigenstates of the momentum operator $\hat{\vec{p}}$ only change their phase upon translation. In other words, the eigenstates of $\hat{\vec{p}}$ are invariant under translations. Since the eigenstates of $\hat{\vec{p}}$ form a complete basis, we conclude that all the irreducible representations of the translation group are one-dimensional, as expected for an Abelian group.

According to the discussion in Sec. 3.17 the Hamiltonian is translationally invariant if and only if $[\hat{H}, \hat{\vec{p}}] = 0$. This also holds for any observable $\hat{A}$ commuting with $\vec{p}$. In general, the translated operator is given by

$$\hat{A}_T = e^{\frac{i}{\hbar}\hat{\vec{p}}\cdot\vec{a}}\,\hat{A}\,e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}},$$

where $\vec{a}$ indicates the translation vector. In particular for an infinitesimal translation $\delta\vec{a}$ we have

$$\hat{A}_T = \hat{A} + \frac{i}{\hbar}\,\delta\vec{a}\cdot[\hat{\vec{p}}, \hat{A}] + O(\delta a^2)\,. \tag{3.23}$$

Let us now apply this formalism to the coordinate operators $\hat{x}_k = x$, y and z. Our purpose is to compare the transformed operator $\hat{x}_{k_T}$, as given by Eq. (3.23), with the corresponding translated operator, as obtained from simple geometrical considerations, in order to derive the commutation rules between position and momentum operators.

The operator $\hat{x}_k$ and any averages of it are unchanged by a translation along an orthogonal direction $j \neq k$. This can be verified, for example, for the operator $\hat{x}$, by noting that any matrix element $\langle\,\alpha\,|\,\hat{x}\,|\,\beta\,\rangle = \int \Psi^*_\alpha(x,y,z)\;\hat{x}\;\Psi_\beta(x,y,z)\,dx\,dy\,dz$ is independent of the origin of the y and z axes. Therefore, we have $\hat{x}_{k_T} = \hat{x}_k$ for $j \neq k$. Combining this information with Eq. (3.23) we obtain

$$\hat{x}_k = \hat{x}_{k_T} = \hat{x}_k + \frac{i}{\hbar}\,\delta a\,[\hat{p}_j, \hat{x}_k],$$

where δa is an infinitesimal translation along the direction $j \neq k$. It follows that

$$[\hat{x}_k, \hat{p}_j] = 0 \quad \text{for } j \neq k.$$

In the case of a translation of $\hat{x}_k$ along the direction k, we have $\hat{x}_k \to \hat{x}_{k_T} = \hat{x}_k + \delta a$. Indeed, any average $\langle \hat{x}_k \rangle$ is shifted by the extent of the translation δa. Replacing in Eq. (3.23) we have,

$$\hat{x}_k + \delta a = \hat{x}_{k_T} = \hat{x}_k + \frac{i}{\hbar}\,\delta a\,[\hat{p}_k, \hat{x}_k]$$

and therefore $1 = \dfrac{i}{\hbar}\,[\hat{p}_k, \hat{x}_k]$. Thus, we recover the fundamental commutation relations

$$[\hat{x}_k, \hat{p}_j] = i\hbar\,\delta_{kj},$$

which are at the origin of Heisenberg's uncertainty relations between momentum and position. Notice that from the point of view of symmetry this is a direct consequence of the definition of $\hat{\vec{p}}$ as the generator of infinitesimal translations.

From $[\hat{x}_k, \hat{p}_k] = i\hbar$ one can easily derive two useful identities:

$$[\hat{x}_k,\, G(\hat{\vec{p}})] = i\hbar\,\frac{\partial G}{\partial \hat{p}_k}$$

and

$$[\hat{p}_k, F(\hat{\vec{x}})] = -i\hbar\,\frac{\partial F}{\partial \hat{x}_k},$$

where $F, G : \mathbb{R}^3 \to \mathbb{C}$ are differentiable functions.[6]

Finally, it is interesting to verify explicitly that

$$|\,\vec{x}\,' + \vec{a}\,\rangle = \hat{U}(\vec{a})\,|\,\vec{x}\,'\,\rangle = e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}}\,|\,\vec{x}\,'\,\rangle$$

is indeed an eigenstate of $\hat{x}$ with eigenvalue $x + a$. Taking into account the commutation relation

$$[\hat{\vec{x}}, e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}}] = i\hbar\left(\frac{-i}{\hbar}\,\vec{a}\right) e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}} = \vec{a}\,e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}},$$

it follows that

$$\begin{aligned}
\hat{\vec{x}}\,|\,\vec{x}\,' + \vec{a}\,\rangle &= \hat{\vec{x}}\,e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}}\,|\,\vec{x}\,'\,\rangle \\
&= \left(\vec{a}\,e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}} + e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}}\,\hat{\vec{x}}\right)\,|\,\vec{x}\,'\,\rangle \\
&= (\vec{a} + \vec{x}\,')\,e^{\frac{-i}{\hbar}\hat{\vec{p}}\cdot\vec{a}}\,|\,\vec{x}\,'\,\rangle \\
&= (\vec{a} + \vec{x}\,')\,|\,\vec{x}\,' + \vec{a}\,\rangle.
\end{aligned}$$

Let us note in passing that this shows that the the spectrum of the coordinate operators $\hat{x}$, $\hat{y}$ and $\hat{z}$ are continuous and unbound ($x_i \in \mathbb{R}$) since $\vec{a} \in \mathbb{R}^3$ is arbitrary.

[6] It is sufficient to show that $[\hat{x}, \hat{p}^n] = i\,\hbar\,n\,\hat{p}^{n-1}\ \forall\ n \geq 0$. For $n = 0$ and $n = 1$ the relation is valid. For $n > 1$ we proceed by induction:

$$[\hat{x}, \hat{p}^n] = [\hat{x}, \hat{p}]\,\hat{p}^{n-1} + \hat{p}\,[\hat{x}, \hat{p}^{n-1}] = i\,\hbar\,\hat{p}^{n-1} + \hat{p}\,[\hat{x}, \hat{p}^{n-1}]. \tag{3.24}$$

Assuming that the relation is valid for $m < n$ we have

$$[\hat{x}, \hat{p}^n] = i\,\hbar\,\hat{p}^{n-1} + \hat{p}\,(n-1)\,i\,\hbar\,\hat{p}^{n-2} = i\,\hbar\,n\,\hat{p}^{n-1}\,, \tag{3.25}$$

which proves the statement.

3.19 Translations in momentum space

A translation in momentum space relates the descriptions of two observers whose reference systems move with a constant velocity $\vec{v}$ with respect to each other. According to the special principle of relativity all the descriptions corresponding to different inertial systems are equivalent. In these lectures, however, we usually prefer to take the perspective that the operators act on the quantum states $|\,\Psi\,\rangle$, in the present case by transferring a given velocity $\vec{v}$ or momentum $\vec{q}$ to the kets. From this point of view, translating a wave packet in momentum space propels $\Psi(\vec{r})$ in a particular direction:

$$\Psi(\vec{r}) = \int \langle \vec{r}|\vec{p}\rangle\, \langle \vec{p}|\Psi\rangle\, dp^3 \rightarrow \Psi'(\vec{r}) = \int \langle \vec{r}|\vec{p}+\vec{q}\rangle\, \langle \vec{p}|\Psi\rangle\, dp^3\,.$$

We are therefore looking for the operator $\hat{U}(\vec{q})$ which, acting on a state with well-defined momentum $|\,\vec{p}\rangle$, yields the state $|\,\vec{p}+\vec{q}\rangle$. It is easy to see that for any constant vector $\vec{q}$ in $\mathbb{R}^3$ ($\vec{q}$ in units of momentum) the operator

$$\hat{U}(\vec{q}) = e^{\frac{i}{\hbar}\vec{q}\cdot\hat{\vec{x}}}, \tag{3.26}$$

where $\hat{\vec{x}}$ is the coordinate operator, performs a translation from $\vec{p}\,'$ to $\vec{p}\,' + \vec{q}$. $\hat{U}(q)$ is of course unitary since $\vec{q} \in \mathbb{R}^3$ and $\hat{\vec{x}}$ is hermitic. Using the commutation relation

$$[\hat{\vec{p}}, e^{\frac{i}{\hbar}\vec{q}\cdot\hat{\vec{x}}}] = -i\hbar \left(\frac{i}{\hbar}\right) \vec{q}\, e^{\frac{i}{\hbar}\vec{q}\cdot\hat{\vec{x}}} = \vec{q}\, e^{\frac{i}{\hbar}\vec{q}\cdot\hat{\vec{x}}},$$

one obtains

$$\begin{aligned}
\hat{\vec{p}}\, \hat{U}(\vec{q})\, |\,\vec{p}\,'\,\rangle &= \hat{\vec{p}}\, e^{i\vec{q}\cdot\hat{\vec{x}}}\, |\,\vec{p}\,'\,\rangle \\
&= \left(\hat{\vec{q}}\, e^{i\vec{q}\cdot\hat{\vec{x}}} + e^{i\vec{q}\cdot\hat{\vec{x}}}\, \hat{\vec{p}}\right) |\,\vec{p}\,'\,\rangle \\
&= \left(\vec{q} + \vec{p}\,'\right) \hat{U}(\vec{q})\, |\,\vec{p}\,'\,\rangle.
\end{aligned}$$

Thus, $\hat{U}(q)|\,\vec{p}\,'\,\rangle$ is an eigenstate of $\hat{\vec{p}}$ with eigenvalue $\vec{p}\,' + \vec{q}$.

$$\hat{U}(\vec{q})|\,\vec{p}\,'\,\rangle = e^{\frac{i}{\hbar}\vec{q}\cdot\hat{\vec{x}}}\, |\,\vec{p}\,'\,\rangle = |\,\vec{p}\,' + \vec{q}\rangle$$

One concludes that minus the coordinate operator $\hat{G} = -\hat{x}$ is the generator of the translations in momentum space. Notice the change of sign of the prefactor, from $(-i)$ in front of $\hat{\vec{p}}$ for the translations in real space [Eq. (3.22)] to i in front of $\hat{\vec{x}}$ for the translations in momentum space [Eq. (3.26)].

3.20 Galileo transformations

The knowledge of the translation operators in both coordinate and momentum spaces allows us to investigate the problem of Galileo invariance in non-relativistic quantum mechanics. We introduce the Galileo transformations at time t_0, which are given by

$$\hat{\vec{x}}_\alpha\,' = \hat{\vec{x}}_\alpha + \vec{v}\, t_0$$

and

$$\hat{\vec{p}}_\alpha{}' = \hat{\vec{p}}_\alpha + m_\alpha \vec{v},$$

where $m_\alpha > 0$ is the mass of particle α, which represents the proportionality constant relating velocity and momentum. As usual $\hat{\vec{x}}_\alpha$ ($\hat{\vec{x}}'_\alpha$) and $\hat{\vec{p}}_\alpha$ ($\hat{\vec{p}}'_\alpha$) refer to the position and momentum operators of the particle α for the reference (transformed) system. A Galileo transformation between two inertial systems involves a translation in space given by $\vec{v}\,t_0$ and a translation in momentum space given by $m_\alpha \vec{v}$. The continuous parameter characterizing the extent of the transformation is the relative velocity $\vec{v} \in \mathbb{R}^3$. The operator for a Galileo transformation at time t with velocity $\vec{v}$ is denoted by $\hat{\Gamma}(t, \vec{v})$. This operator is unitary since the two descriptions are equivalent and the transformation depends on the continuous parameter $\vec{v}$. For an infinitesimal velocity difference $\delta\vec{v}$ the coordinate and momentum translations, and thus $\hat{\Gamma}(t, \vec{v})$, are all infinitesimal.

In the case of a many particle system the position operators $\hat{\vec{x}}_\alpha$ for different particles α commute, i.e., $[\hat{\vec{x}}_\alpha, \hat{\vec{x}}_\beta] = 0 \quad \forall\ \alpha, \beta$. The translation in momentum space $\vec{q}_\alpha$ for an ensemble of particles is obtained by performing the transformation for each individual particle:

$$\hat{U}(\{\vec{q}_\alpha\}) = \prod_\alpha e^{\frac{i}{\hbar}\,\vec{q}_\alpha \cdot \hat{\vec{x}}_\alpha} = e^{\frac{i}{\hbar}\,\sum_\alpha \vec{q}_\alpha \cdot \hat{\vec{x}}_\alpha},$$

where we have used that $e^{\hat{A}}\, e^{\hat{B}} = e^{\hat{A}+\hat{B}}$ when $[\hat{A},\, \hat{B}] = 0$. Similarly, for a real space translation $\vec{a}$ of an ensemble of particles we have

$$\hat{U}(\vec{a}) = \prod_\alpha e^{-\frac{i}{\hbar}\,\hat{\vec{p}}_\alpha \cdot \vec{a}} = e^{-\frac{i}{\hbar}\,\sum_\alpha \hat{\vec{p}}_\alpha \cdot \vec{a}} = e^{-\frac{i}{\hbar}\,\hat{\vec{P}} \cdot \vec{a}},$$

where $\hat{\vec{P}} = \sum_\alpha \hat{\vec{p}}_\alpha$ is the total momentum operator. As in the case of the coordinates, the momentum operators of different particles commute.

It is clear that coordinate and momentum translations do not commute in general. However, for infinitesimal $\delta\vec{a}$ and $\delta\vec{q}_\alpha$, only the linear terms need to taken into account. In the absence of products their order is immaterial. Indeed, for infinitesimal $\delta\vec{v}$ we have $\delta\vec{a} = t\,\delta\vec{v}$ and $\delta\vec{q}_\alpha = m_\alpha \delta\vec{v}$ and therefore

$$\hat{\Gamma}(t, \delta\vec{v}) \;=\; 1 + \frac{i}{\hbar}\left(\sum_\alpha \underbrace{m_\alpha\, \delta\vec{v}}_{\delta\vec{q}_\alpha} \cdot \hat{\vec{x}}_\alpha - \sum_\alpha \underbrace{t\,\delta\vec{v}}_{\delta\vec{a}} \cdot \hat{\vec{p}}_\alpha\right). \tag{3.27}$$

Translation in momentum space — Translation in real space

In this context it is useful to introduce the operator $\hat{\vec{R}} = \dfrac{\sum_\alpha m_\alpha\, \hat{\vec{x}}_\alpha}{M}$ of the center-of-mass coordinate, where $M = \sum_\alpha m_\alpha$ is the total mass. Replacing $\hat{\vec{R}}$ and the total momentum

operator $\hat{\vec{P}} = \sum_\alpha \hat{\vec{p}}_\alpha$ in Eq. (3.27) we obtain

$$\hat{\Gamma}(t, \delta\vec{v}) = 1 + \frac{i}{\hbar}\, \delta\vec{v} \cdot (M\hat{\vec{R}} - t\,\hat{\vec{P}}) + O(\delta v^2)\,.$$

The generator of Galileo transformations is therefore $\hat{G} = (t\hat{\vec{P}} - M\hat{\vec{R}}) \cdot \hat{n}$, where $\hat{n} = \delta\vec{v}/\delta v$ is the direction of the relative motion. The representation of a finite-velocity transformation is then given by

$$\hat{\Gamma}(t, \vec{v}) = e^{\frac{i}{\hbar}\,\vec{v}\cdot(M\hat{\vec{R}} - t\,\hat{\vec{P}})}$$

since $\hat{G}$ is independent of $\vec{v}$.

3.21 Galileo invariance

The special principle of relativity requires the laws of quantum mechanics to be the same on all different inertial systems. Galileo invariance is the mathematical formulation of this fundamental principle in the limit of small relative velocities $v \ll c$. The fact that the equation of motion in non-relativistic quantum mechanics must satisfy Galileo invariance sets strong constraints on the form of the time-evolution operator and thus on the Hamiltonian as a function of $\hat{\vec{x}}_\alpha$ and $\hat{\vec{p}}_\alpha$. The purpose of this section is to derive and discuss the consequences this invariance.

Suppose the observer O prepares a state $|\,\alpha\, t_0, t_0\,\rangle$ at time $t = t_0$, and then undertakes the following two steps:

1. First he lets this state evolve to time t_1 thereby obtaining

$$|\,\alpha\, t_0, t_1\,\rangle = \hat{U}(t_1, t_0)\,|\,\alpha\, t_0, t_0\,\rangle.$$

2. And then, at time $t = t_1$, he transforms this state into the moving reference system with velocity $\vec{v}$ to obtain

$$\begin{aligned} |\,\alpha\, t_0, \vec{v}, t_1\,\rangle &= \hat{\Gamma}(t_1, \vec{v})\,|\,\alpha\, t_0, t_1\,\rangle \\ &= \hat{\Gamma}(t_1, \vec{v})\,\hat{U}(t_1, t_0)\,|\,\alpha\, t_0, t_0\,\rangle. \end{aligned} \tag{3.28}$$

The ket $|\,\alpha\, t_0, \vec{v}, t_1\,\rangle$ in the moving frame can now be tested by all kinds of measurements. For instance, for any state $|\,\beta'\,\rangle$ in the moving reference frame we can compute $\langle\,\beta'\,|\,\alpha\, t_0, \vec{v}, t_1\,\rangle$ and measure the probability $|\langle\,\beta'\,|\,\alpha\, t_0, \vec{v}, t_1\,\rangle|^2$ of finding the system in the state $|\,\beta'\,\rangle$.

Now imagine that the observer O' proceeds in the other way:

1. Starting from the same state

$$|\,\alpha\, t_0, t_0\,\rangle,$$

we transform it to the moving frame at time $t = t_0$ and obtain

$$|\,\alpha\, t_0, \vec{v}, t_0\,\rangle = \hat{\Gamma}(t_0, \vec{v})\,|\,\alpha\, t_0, t_0\,\rangle.$$

2. This ket now evolves in time according to the dynamics in the moving frame which must be given by *the same* time-evolution operator $\hat{U}(t, t_0)$, since the laws of physics are the same in all inertial systems. Note that the dynamics is <u>not</u> the same, since the momenta of the kets $|\,\alpha\, t_0, t_0\,\rangle$ and $|\,\alpha\, t_0, \vec{v}, t_0\,\rangle$ are different. What the principle of relativity requires is that the operator of time evolution $\hat{U}(t, t_0)$ is written in the same way as a function of $\hat{\vec{x}}_\alpha$ and $\hat{\vec{p}}_\alpha$ in both reference frames. Propagating in the moving frame from t_0 to t_1 we obtain

$$\begin{aligned} |\,\alpha\, t_0, \vec{v}, t_1\,\rangle &= \hat{U}(t_1, t_0)\,|\,\alpha\, t_0, \vec{v}, t_0\,\rangle \\ &= \hat{U}(t_1, t_0)\,\hat{\Gamma}(t_0, \vec{v})\,|\,\alpha\, t_0, t_0\,\rangle\,. \end{aligned} \tag{3.29}$$

The two procedures are physically equivalent since the laws of physics are the same on the two frames. Moreover, since $\hat{\Gamma}(t, \vec{v})$ is a continuous transformation between equivalent descriptions, it is unitary. Therefore the inner products are preserved. Comparing Eqs. (3.28) and (3.29) we have

$$\hat{\Gamma}(t_1, \vec{v})\,\hat{U}(t_1, t_0) = \hat{U}(t_1, t_0)\,\hat{\Gamma}(t_0, \vec{v})\,. \tag{3.30}$$

This relation expresses the invariance of $\hat{U}(t_1,\, t_0)$ upon Galileo transformations, i.e., $\hat{U}_T(t_1,\, t_0) = \hat{U}(t_1,\, t_0)$, since according to Eq. (3.17)

$$\hat{U}_T(t_1,\, t_0) = \hat{\Gamma}^\dagger(t_1,\, \vec{v}\,)\,\hat{U}(t_1,\, t_0)\,\hat{\Gamma}(t_0,\, \vec{v}\,). \tag{3.31}$$

See also Sec. 3.13.

Since the invariance relation (3.30) holds for finite transformations, it also holds for infinitesimal ones. We would like to draw the consequences in this simpler limit. Knowing that

$$\hat{\Gamma}(t, \delta\vec{v}) = 1 + \frac{i}{\hbar}\,\delta\vec{v}\cdot(M\hat{\vec{R}} - t\hat{\vec{P}})$$

for infinitesimal $\delta\vec{v}$, and replacing in Eq. (3.30), we obtain

$$(M\hat{\vec{R}} - t_1\hat{\vec{P}})\,\hat{U}(t_1, t_0) = \hat{U}(t_1, t_0)\,(M\hat{\vec{R}} - t_0\hat{\vec{P}})\,,$$

which leads to

$$\Big[M\hat{\vec{R}}\,\hat{U}(t_1, t_0) - \hat{U}(t_1, t_0)M\hat{\vec{R}}\Big] = t_1\hat{\vec{P}}\,\hat{U}(t_1, t_0) - t_0\,\hat{U}(t_1, t_0)\,\hat{\vec{P}}\,.$$

We are primarily interested in translationally invariant systems for which $[\hat{\vec{P}}, \hat{U}(t,t_0)] = 0$ for all t. Multiplying from the left with $\hat{U}^\dagger(t_1,t_0)$ we obtain

$$M\,\hat{U}^\dagger(t_1,t_0)\;\hat{\vec{R}}\,\hat{U}(t_1,t_0) - M\hat{\vec{R}} = (t_1-t_0)\,\hat{\vec{P}}, \tag{3.32}$$

or equivalently,

$$\hat{\vec{R}}_H(t_1) = \hat{\vec{R}}(t_0) + (t_1-t_0)\,\frac{\hat{\vec{P}}}{M}.$$

This is the time dependence of the operator of the center of mass $\hat{\vec{R}}_H$ of an isolated system in the Heisenberg picture! Notice that $\hat{\vec{P}}_H = \hat{\vec{P}}_S$ since $[\hat{U}, \hat{\vec{P}}_S] = 0$.

The Eq. (3.32) imposes serious constraints on $\hat{U}$ and on the Hamiltonian $\hat{H}$ as a function of $\hat{\vec{p}}_\alpha$ and $\hat{\vec{x}}_\alpha$, which are most interesting to reveal. For this purpose we develop $\hat{U}(t_1,t_0)$ in Eq. (3.32) for an infinitesimal time interval $\delta t = t_1 - t_0$ and obtain

$$\frac{i}{\hbar}\,\hat{H}\,M\,\hat{\vec{R}}\,\delta t - \frac{i}{\hbar}\,M\,\hat{\vec{R}}\,\hat{H}\,\delta t = \overbrace{(t_1-t_0)}^{\delta t}\;\hat{\vec{P}} + O(\delta t^2),$$

which is equivalent to

$$\left[M\hat{\vec{R}}, \hat{H}\right] = i\hbar\,\hat{\vec{P}}.$$

Replacing the definition of $\hat{\vec{R}}$ and $\hat{\vec{P}}$ we have

$$\sum_\alpha m_\alpha \underbrace{[\hat{\vec{x}}_\alpha, \hat{H}]}_{i\hbar\,\partial\hat{H}/\partial\vec{p}_\alpha} = i\hbar \sum_\alpha \hat{\vec{p}}_\alpha$$

which can be written as

$$\sum_\alpha m_\alpha \frac{\partial \hat{H}}{\partial \vec{p}_\alpha} = \sum_\alpha \hat{\vec{p}}_\alpha. \tag{3.33}$$

Since all the variables $\vec{p}_\alpha$ are independent, and since Eq. (3.33) holds for any choice of m_α or of the number of particles, we must have

$$m_\alpha \frac{\partial \hat{H}}{\partial \vec{p}_\alpha} = \hat{\vec{p}}_\alpha\,.$$

Straightforward integration reveals the *explicit form* of the Hamiltonian in non-relativistic quantum mechanics:

$$\hat{H} = \sum_\alpha \frac{\hat{p}_\alpha^2}{2m_\alpha} + W(\hat{\vec{x}}_1, \ldots, \hat{\vec{x}}_n)\,,$$

where $W(\vec{x}_1, \ldots, \vec{x}_n)$ is an arbitrary translational invariant function of the particle coordinates. Typically W takes the form of a two particle interaction

$$W(x_1, \ldots, x_n) = \sum_{\alpha<\beta} w(\,|\,\vec{x}_\alpha - \vec{x}_\beta\,|\,)\,.$$

For instance, in the case of the Coulomb interaction we have

$$W = \sum_{\alpha<\beta} \frac{e^2}{|\vec{x}_\alpha - \vec{x}_\beta|} \,.$$

In conclusion, the well-known form of the Hamiltonian of an ensemble of particles, and in particular the form of the kinetic energy operator in non-relativistic quantum mechanics, follow entirely from Galileo invariance! Any reference to the classical mechanics or to the inelegant correspondence principle is absolutely unnecessary.

3.22 Rotations

We turn now our attention to the most interesting and complex group among the continuous groups that we are going to consider in this course. Our first main objectives are i) to characterize the rotations in $\mathbb{R}^3$, i.e., the elements of the group SO(3), ii) to find the representation of this group in quantum mechanical Hilbert spaces, and iii) to derive the generator of infinitesimal rotations.

A few basic properties of rotations and rotation matrices in $\mathbb{R}^3$ should be recalled:

1. A rotation in $\mathbb{R}^3$ is characterized by its axis of rotation $\hat{n}$ and the rotation angle ϕ.
2. Rotation matrices are real and preserve scalar product among vectors.
3. They are therefore orthogonal: $O^t O = O\, O^t = 1$.
4. Consequently, $\det O^t = \det O \;\Rightarrow\; \det O^2 = 1 \;\Rightarrow\; \det O = \pm 1$ with $O \in \mathbb{R}^{3\times 3}$.
5. The rotation matrices R have $\det R = 1$ since they do not change right-handed systems into left-handed systems. This forms the group of *special orthogonal matrices* SO(3).
6. Any orthogonal matrix O can be written either as an SO(3) matrix, if $\det O = 1$, or as πO, where $\pi = -\mathbb{1}$ is the inversion matrix and O has $\det O = 1$ (space inversion implies $x \to -x,\ y \to -y,\ z \to -z$). For example, the reflexion with respect to the xy plane

   $$\sigma_z = \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix} \text{ can be written as } \sigma_z = -R_z(\pi) \text{ with } R_z(\pi) = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix}$$

   being a rotation of an angle $\phi = \pi$ around the z axis.

The multiplication properties of the representation operator $\hat{D}(R)$ in the Hilbert space $\mathcal{H}$ of the quantum mechanical system are the same as the properties of the rotations in SO(3). Let us recall that

$$\hat{D}(R_1)\, \hat{D}(R_2) = \hat{D}(R_1 R_2)$$

and

$$\hat{D}(\mathbb{1}) = \mathbb{1}$$

as in any representation. We can therefore derive the commutation rules for $\hat{D}(R)$ and for the generator of rotations from the commutation properties of the matrices R in SO(3). The rotation matrix around the $\hat{z}$ axis is given by

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

where ϕ is the rotation angle. For an infinitesimal rotation $\delta\phi = \varepsilon$ we have

$$R_z(\varepsilon) = \begin{pmatrix} 1 - \frac{\varepsilon^2}{2} & -\varepsilon & 0 \\ \varepsilon & 1 - \frac{\varepsilon^2}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} + O(\varepsilon^3).$$

For the rotations around the axes $\hat{x}$ and $\hat{y}$ we have

$$R_x(\varepsilon) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \frac{\varepsilon^2}{2} & -\varepsilon \\ 0 & \varepsilon & 1 - \frac{\varepsilon^2}{2} \end{pmatrix} \quad \text{and} \quad R_y(\varepsilon) = \begin{pmatrix} 1 - \frac{\varepsilon^2}{2} & 0 & \varepsilon \\ 0 & 1 & 0 \\ -\varepsilon & 0 & 1 - \frac{\varepsilon^2}{2} \end{pmatrix},$$

which can be obtained from $R_z(\varepsilon)$ by the cyclic permutation $(x \to y,\ y \to z,\ z \to x)$. Direct matrix multiplication yields the commutator

$$\begin{aligned} [R_x(\varepsilon), R_y(\varepsilon)] &= R_x(\varepsilon)\, R_y(\varepsilon) - R_y(\varepsilon)\, R_x(\varepsilon) \\ &= \begin{pmatrix} 0 & -\varepsilon^2 & 0 \\ \varepsilon^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + O(\varepsilon^3) \\ &= R_z(\varepsilon^2) - 1 + O(\varepsilon^3). \end{aligned} \tag{3.34}$$

Notice that the commutator on the left-hand side is proportional to ε^2, which is the angle of the rotation around the z axis on the right-hand side. Cyclic permutations of x, y and z yield the other commutation relations since the matrices $R_x(\varepsilon)$, $R_y(\varepsilon)$ and $R_z(\varepsilon)$ satisfy the cyclic permutation rules. In the following we will use the commutation relation (3.34) between the matrices R_x, R_y and R_z of the SO(3) group in order to derive the commutation relations between the generators $\hat{J}_x$, $\hat{J}_y$ and $\hat{J}_z$ of the group representation in an arbitrary Hilbert space. In fact, the commutation relation $[R_x(\varepsilon), R_y(\varepsilon)] = R_z(\varepsilon^2) - 1 + O(\varepsilon^3)$ among infinitesimal rotations in $\mathbb{R}^3$ implies that in any Hilbert space the same commutation relation $[\hat{D}(\hat{x}, \varepsilon), \hat{D}(\hat{y}, \varepsilon)] = \hat{D}(\hat{z}, \varepsilon^2)] - 1 + O(\varepsilon^3)$ must hold among the representation operators $\hat{D}(\hat{n}, \varepsilon)$ of infinitesimal rotations around the corresponding axes $\hat{n}$.

We consider the representation of infinitesimal rotations $\delta\phi = \varepsilon$ around the axes $\hat{x}$, $\hat{y}$ and $\hat{z}$, which we write, as usual, in the form

$$\hat{D}(\hat{x},\,\varepsilon) = 1 - \frac{i}{\hbar}\,\hat{J}_x\,\varepsilon\,,$$
$$\hat{D}(\hat{y},\,\varepsilon) = 1 - \frac{i}{\hbar}\,\hat{J}_y\,\varepsilon\,,$$

and

$$\hat{D}(\hat{z},\,\varepsilon) = 1 - \frac{i}{\hbar}\,\hat{J}_z\,\varepsilon\,,$$

where $\hat{J}_x$, $\hat{J}_y$ and $\hat{J}_z$ are the generators of infinitesimal rotations around the Cartesian axes. This can be written in a more compact form as

$$\hat{D}(\hat{n},\delta\phi) = 1 - \frac{i}{\hbar}\,\hat{n}\cdot\hat{\vec{J}}\,\delta\phi, \tag{3.35}$$

where $\hat{n}$ denotes the axis of rotation and $\hat{\vec{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$ is by definition the angular momentum operator in the Hilbert space of our quantum mechanical system. Expressing the commutator among rotations around different axes (3.34) in terms of the representation $\hat{D}(R)$ one obtains

$$\begin{aligned}
[\hat{D}(\hat{x},\varepsilon),\hat{D}(\hat{y},\varepsilon)] &= \hat{D}(\hat{x},\varepsilon)\;\hat{D}(\hat{y},\varepsilon) - \hat{D}(\hat{y},\varepsilon)\;\hat{D}(\hat{x},\varepsilon)\\
&= \left(1 - \frac{i}{\hbar}\,\hat{J}_x\varepsilon + O_x(\varepsilon^2)\right)\,\left(1 - \frac{i}{\hbar}\,\hat{J}_y\varepsilon + O_y(\varepsilon^2)\right)\\
&\quad - \left(1 - \frac{i}{\hbar}\,\hat{J}_y\varepsilon + O_y(\varepsilon^2)\right)\,\left(1 - \frac{i}{\hbar}\,\hat{J}_x\varepsilon + O_x(\varepsilon^2)\right)\\
&= \frac{i^2}{\hbar^2}\,\hat{J}_x\,\hat{J}_y\,\varepsilon^2 - \frac{i^2}{\hbar^2}\,\hat{J}_y\,\hat{J}_x\,\varepsilon^2 + O(\varepsilon^3)\\
&= \frac{i^2\varepsilon^2}{\hbar^2}\,(\hat{J}_x\,\hat{J}_y - \hat{J}_y\,\hat{J}_x).
\end{aligned}$$

Taking in account the relation (3.34) among the rotation matrices R and the fundamental property of the represenatation of a group we deduce that the commutator $[\hat{D}(\hat{x},\varepsilon),\hat{D}(\hat{y},\varepsilon)]$ must be equal to

$$\hat{D}(\hat{z},\varepsilon^2) - 1 = \frac{-i}{\hbar}\,\hat{J}_z\,\varepsilon^2,$$

which implies

$$[\hat{J}_x,\hat{J}_y] = i\hbar\,\hat{J}_z.$$

This means that the famous commutation rule between the components of angular momentum operator $\hat{\vec{J}}$ is a consequence of i) the representation property of $\hat{D}(R)$, ii) the

fact that $\hat{\vec{J}}$ is the generator of infinitesimal rotations, and iii) the commutation rules of the rotations R_x, R_y and R_z in SO(3).

By permuting the indices x, y and z cyclically we obtain

$$[\hat{J}_i, \hat{J}_j] = i\hbar\, \varepsilon_{ijk}\, \hat{J}_k. \tag{3.36}$$

Notice that this fundamental commutation rule holds irrespectively of the nature of the Hilbert space in which $\hat{\vec{J}}$ operates (single particle, integer or half-integer spin, many body, bosonic, fermionic, etc.).

Using the definition (3.35) of the angular momentum $\hat{\vec{J}}$ as generator of infinitesimal rotations and the general expression (3.4) of the representation of finite transformations in continuous groups we obtain the representation

$$\hat{D}(\hat{n},\, \phi) = e^{-\frac{i}{\hbar}\, \hat{\vec{J}}\cdot\hat{n}\, \phi} \tag{3.37}$$

of a rotation with finite angle ϕ around the axis $\hat{n}$ (see Sec. 3.9).

3.23 Rotation of spin 1/2 operators

As a first application we consider a single spin-1/2 particle. Let us recall the spin-1/2 operators

$$\hat{S}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\hbar}{2} \left(|-\rangle\langle +| + |+\rangle\langle -| \right),$$

$$\hat{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{i\hbar}{2} \left(|-\rangle\langle +| - |+\rangle\langle -| \right)$$

and

$$\hat{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2} \left(|+\rangle\langle +| - |-\rangle\langle -| \right).$$

Given the ket

$$|\,\alpha\,\rangle = \begin{pmatrix} \langle +|\,\alpha\,\rangle \\ \langle -|\,\alpha\,\rangle \end{pmatrix} = \begin{pmatrix} c_+ \\ c_- \end{pmatrix},$$

the rotated state $|\,\alpha\,\rangle \xrightarrow{R} |\,\alpha\,\rangle_R$ is obtained by applying a rotation operator of an angle ϕ around the axis $\hat{n}$:

$$|\,\alpha_R\,\rangle = e^{-\frac{i}{\hbar}\hat{\vec{S}}\cdot\hat{n}\, \phi}\, |\,\alpha\,\rangle.$$

From Sec. 3.12 we know that for any operator of an observable $\hat{A}$, the rotated operator $\hat{A}^R$, which is defined by the relation $\langle\,\beta\,|\,\hat{A}^R\,|\,\alpha\,\rangle = \langle\,\beta_R\,|\,\hat{A}\,|\,\alpha_R\,\rangle$, is given by

$$\begin{aligned} \hat{A}^R &= \hat{D}^\dagger(R)\, \hat{A}\, \hat{D}(R) \\ &= e^{\frac{i}{\hbar}\hat{\vec{S}}\cdot\hat{n}\, \phi}\ \hat{A}\ e^{\frac{-i}{\hbar}\hat{\vec{S}}\cdot\hat{n}\, \phi}. \end{aligned}$$

Let us consider the effect of a rotation with an angle ϕ around the z axis on the spin operator $\hat{S}_x$ corresponding to the $\hat{x}$ axis:

$$\begin{aligned}\hat{S}_x^R &= \frac{\hbar}{2}\, e^{i\,\frac{\hat{S}_z}{\hbar}\,\phi}\,\{\,|+\rangle\langle -| \,+\, |-\rangle\langle +|\,\}\, e^{-i\,\frac{\hat{S}_z\,\phi}{\hbar}} \\ &= \frac{\hbar}{2}\left\{e^{i\,\phi/2}\,|+\rangle\langle -|\, e^{i\,\phi/2} + e^{-i\,\phi/2}\,|-\rangle\langle +|\, e^{-i\,\phi/2}\right\} \\ &= \frac{\hbar}{2}\left(e^{i\,\phi}\,|+\rangle\langle -| \,+ e^{-i\,\phi}\,|+\rangle\langle -|\,\right) \\ &= \cos\phi\,\hat{S}_x - \sin\phi\,\hat{S}_y.\end{aligned}$$

This means that $\hat{\vec{S}}$ behaves like a vector under rotations. Indeed, for the x component we have shown that

$$\langle\,\hat{S}_x\,\rangle \xrightarrow{R} \langle\,\hat{S}_x\,\rangle_R = \cos\phi\,\langle\,\hat{S}_x\,\rangle - \sin\phi\,\langle\,\hat{S}_y\,\rangle.$$

Moreover, one easily obtains

$$\hat{S}_y^R = \hat{S}_x\,\sin\phi + \hat{S}_y\,\cos\phi$$

and

$$\hat{S}_z^R = \hat{S}_z.$$

Consequently, this can be written in a more compact and transparent form as

$$\hat{\vec{S}}_R = \begin{pmatrix}\hat{S}_x^R\\ \hat{S}_y^R\\ \hat{S}_z^R\end{pmatrix} = \begin{pmatrix}\cos\phi & -\sin\phi & 0\\ \sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\hat{S}_x\\ \hat{S}_y\\ \hat{S}_z\end{pmatrix} = R\,\hat{\vec{S}},$$

where R is the SO(3) rotation matrix! This is a particular case of the general *vector operator transformation*

$$\hat{V}_k^R = \sum_{l=1}^{3} R_{kl}\,\hat{V}_l\,,$$

which holds for any operator $\hat{\vec{V}} = (\hat{V}_1, \hat{V}_2, \hat{V}_3)$ satisfying the commutation rules

$$[\hat{V}_i, \hat{J}_j] = i\hbar\,\varepsilon_{ijk}\,\hat{V}_k\,,$$

where $\hat{\vec{J}} = (\hat{J}_1, \hat{J}_2, \hat{J}_3)$ is the generator of rotations. In the present case we have $[\hat{S}_x, \hat{S}_y] = i\hbar\,\hat{S}_z$. The same commutation rule $[\hat{S}_x, \hat{J}_y] = i\hbar\,\hat{S}_z$ holds for the total angular momentum $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$ since $\hat{\vec{S}}$ and $\hat{\vec{L}}$ commute. Other examples of vector operators are $\hat{\vec{r}}$ and $\hat{\vec{p}}$ (e.g., $[\hat{x}, \hat{L}_y] = i\hbar\,\hat{z}$).

3.24 Spin precession

The Hamiltonian describing the coupling of the spin of an electron with an external magnetic field $\vec{B}$ is given by

$$\hat{H} = \frac{e}{mc}\,\hat{\vec{S}}\cdot\vec{B},$$

where $\hat{\vec{S}}$ is the spin operator and $e>0$ is the electronic charge. For $\vec{B}=B\,\hat{z}$ this reads

$$\hat{H} = \omega\,\hat{S}_z$$

with $\omega = \dfrac{e\,B}{mc}$. The time evolution operator is given by

$$\hat{U}(t,0) = e^{-\frac{i}{\hbar}\hat{H}\,t} = e^{-\frac{i}{\hbar}\hat{S}_z\,\omega\,t}.$$

This is precisely the rotation operator for an angle $\phi=\omega\,t$. According to our previous calculation of rotated operators $\hat{\vec{S}}_R$, the Heisenberg operators are given by

$$\hat{S}_x^H(t) = \hat{S}_x(0)\,\cos\omega\,t - \hat{S}_y(0)\,\sin\omega\,t\,,$$

$$\hat{S}_y^H(t) = \hat{S}_x(0)\,\sin\omega\,t + \hat{S}_y(0)\,\cos\omega\,t\,,$$

and

$$\hat{S}_z^H(t) = \hat{S}_z(0),$$

from which the corresponding averages $\langle\,\hat{S}_x\,\rangle$, $\langle\,\hat{S}_y\,\rangle$ and $\langle\,\hat{S}_z\,\rangle$ are obtained. Clearly, after a time $T=2\pi/\omega$, $\hat{\vec{S}}^H$ and $\langle\,\hat{\vec{S}}\,\rangle$ return to the original direction. However, for the Schrödinger kets $|\,\alpha\,\rangle$ the time requiered to recover exactly the same state is $4\pi/\omega$, since

$$\begin{aligned}|\,\alpha, t_0=0, t\,\rangle &= e^{-\frac{i}{\hbar}\hat{S}_z\,\omega\,t}\,(\,|+\rangle\langle+|\,\alpha\,\rangle + |-\rangle\langle-|\,\alpha\,\rangle\,)\\ &= e^{-i\frac{\omega\,t}{2}}\,|+\rangle\langle+|\,\alpha\,\rangle + e^{i\frac{\omega\,t}{2}}\,|-\rangle\langle-|\,\alpha\,\rangle.\end{aligned}$$

In other words, the period of precession of the measurable average spin and the Heisenberg operator $\hat{\vec{S}}^H(t)$ is $T=2\pi/\omega$, while the period of oscillation of the quantum states is twice as long.

3.25 2π rotation of spinor states

From the previous expression for $\hat{\vec{S}}_R$ one observes that $\hat{\vec{S}}$ is invariant after a 2π rotation, which means that $\hat{\vec{S}}$ is single valued, as expected for any observable. However, the spinor kets show a 4π periodicity. In fact, the rotation of $|\alpha\rangle$ around the z axis gives

$$|\,\alpha_{R_z}\,\rangle = e^{-i\,\frac{\hat{S}_z\,\phi}{\hbar}}\,|\,\alpha\,\rangle = \begin{pmatrix} e^{-i\,\phi/2} & 0\\ 0 & e^{i\,\phi/2}\end{pmatrix}\begin{pmatrix}\chi_+\\ \chi_-\end{pmatrix},$$

where $\chi_+ = \langle +|\alpha\rangle$ and $\chi_- = \langle +|\alpha\rangle$, which implies

$$|\,\alpha_{R_z=2\pi}\,\rangle = -\,|\,\alpha\,\rangle.$$

Of course, this does not affect the average value $\langle\,\hat{A}\,\rangle = \langle\,\hat{A}\,\rangle_{R=2\pi}$ of any observable A. However, this remarkable phase change, actually the fact that a periodicity of 4π is predicted in the phase, could be observed in an *interference experiment.*

This has actually been reported by S. A. Werner *et al.* in "Observation of the Phase Shift of a Neutron due to Precession in a Magnetic Field," Phys. Rev. Lett. **35**, 1053 (1975) as illustrated in Fig. 8.

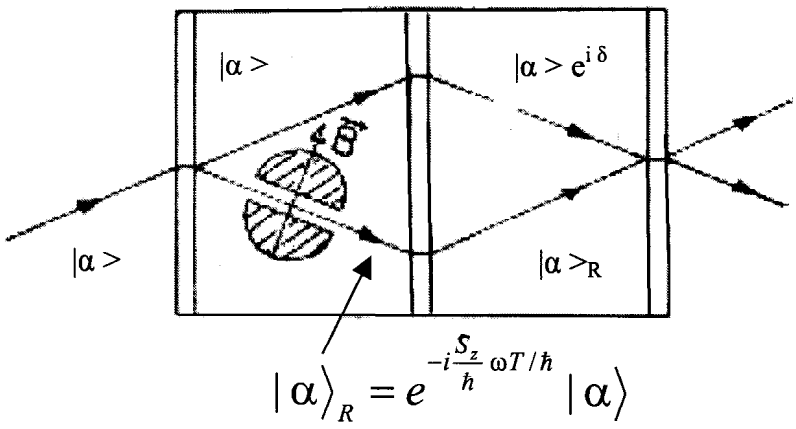


Figure 8: Observation of the Phase Shift of a Neutron due to Precession in a Magnetic Field after S. A. Werner *et al.*, PRL **35**, 1053 (1975).

T is the passage time which depends on the apparatus, on the energy and momentum of the neutrons, and on the magnet length, but which is independent of the magnetic field strength B. Let us denote by δ the phase difference between the two paths for $B = 0$. The spin rotation frequency is given by

$$\omega = \frac{g_n e\hbar\, B}{2mp\, c} = g_n\, \mu_N\, B,$$

where g_n is the neutron gyromagnetic factor and μ_N the nuclear magneton. Notice that $\omega \propto B$ and thus $\omega T \propto B$. Therefore, the expected periodicity of the signal as a function of B corresponds to $\omega T = 4\pi$.

The interference takes place between the states

$$|\,\alpha\,\rangle = e^{i\delta}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \qquad \text{and} \qquad |\,\alpha\,\rangle_R = e^{-i\,\frac{\hat{S}_z}{\hbar}\omega T}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} c_1 & e^{-i\,\omega T/2} \\ c_2 & e^{i\,\omega T/2} \end{pmatrix}.$$

In the interference region we have

$$|\,\alpha\,\rangle + |\,\alpha_R\,\rangle = \begin{pmatrix} c_1\ e^{i\delta} & [1+e^{-i(\omega T/2+\delta)}] \\ c_2\ e^{i\delta} & [1+e^{i(\omega T/2-\delta)}] \end{pmatrix}.$$

The probability of finding a neutron (beam intensity) is

$$\begin{aligned} I &= |\,c_1\,|^2 \left[1+\cos^2\left(\frac{\omega T}{2}+\delta\right)+2\,\cos\left(\frac{\omega T}{2}+\delta\right)+\sin^2\left(\frac{\omega T}{2}+\delta\right)\right] \\ &\quad + |\,c_2\,|^2 \left[1+\cos^2\left(\frac{\omega T}{2}-\delta\right)+2\,\cos\left(\frac{\omega T}{2}-\delta\right)+\sin^2\left(\frac{\omega T}{2}-\delta\right)\right] \\ &= 2\,|\,c_1\,|^2\left(1+\cos\delta\ \cos\frac{\omega T}{2}-\sin\delta\ \sin\frac{\omega T}{2}\right) \\ &\quad + 2\,|\,c_2\,|^2\left(1+\cos\delta\ \cos\frac{\omega T}{2}+\sin\delta\ \sin\frac{\omega T}{2}\right) \\ &= 2\left(1+\cos\delta\ \cos\frac{\omega T}{2}+\underbrace{(|\,c_2\,|^2-|\,c_1\,|^2)}_{\langle\ \rangle=0\ \text{in an unpolarized beam}}\sin\delta\ \sin\frac{\omega T}{2}\right). \end{aligned}$$

In conclusion,

$$I \propto 1+\cos\delta\ \cos\frac{\omega T}{2}\ \leftarrow\ \text{periodicity } 4\pi!$$

Basically, this is what the experimentalists have measured, once δ has been tuned in order that $\cos\delta \cong 1$. In fact, one would not be able to measure the effect if $\cos\delta \simeq 0$ (i.e., $\delta = \pi/2,\ 3\pi/2$, etc).

3.26 Representation of rotations by means of Euler angles

Let us recall the representation of rotations in terms of Euler angles, which is known from classical mechanics. An arbitrary rotation R, which brings the reference axes x, y and z to new orientations given by the axes x', y' and z' can be performed as the succession of the three following rotations:

1. Rotation around axis z with an angle α such that the rotated y axis, which is denoted by $\tilde{y}$, is perpendicular to the plane spanned by $\hat{z}$ and $\hat{z}'$. Only then a rotation around $\tilde{y}$ can bring $z \to z'$.

2. Rotation around $\tilde{y}$ with an angle β equal to the angle between z and z', such that z points along its final direction z'

3. A final rotation around z' with an angle γ equal to the angle between $\tilde{y}$ and y', which now lie on the plane perpendicular to z', which brings $\tilde{y}$ along its final orientation along y'. The x axis goes thereby by itself along x'.

In this way we have

$$R = R_{z'}(\gamma)\, R_{\tilde{y}}(\beta)\, R_z(\alpha). \tag{3.38}$$

Although geometrically clear, this representation is very inconvenient since the axes of rotation $\tilde{y}$ and z' depend on the actual rotation. We would therefore like to express all the rotations in terms of the angles α, β and γ using the fixed reference axes x, y and z. For this purpose we note the following equivalence relations between rotations around different axes:

$$R_{\tilde{y}}(\beta) = R_z(\alpha)\, R_y(\beta)\, R_z(-\alpha) \tag{3.39}$$

$R_z(-\alpha)$: brings $\tilde{y}$ to y.
$R_y(\beta)$: rotates around y with an angle β.
$R_z(\alpha)$: rotates y back to $\tilde{y}$.

and

$$R_{z'}(\gamma) = R_{\tilde{y}}(\beta)\, R_z(\gamma)\, R_{\tilde{y}}(-\beta). \tag{3.40}$$

$R_{\tilde{y}}(-\beta)$: brings z' to z.
$R_z(\gamma)$: rotates around z with an angle γ.
$R_{\tilde{y}}(\beta)$: brings z back to z'.

Thus, using Eq. (3.39) we can write Eq. (3.40) as

$$R_{z'}(\gamma) \;=\; \underbrace{R_z(\alpha)\, R_y(\beta)\, R_z(-\alpha)}_{R_{\tilde{y}}(\beta)}\; R_z(\gamma)\; \underbrace{R_z(\alpha)\, R_y(-\beta)\, R_z(-\alpha)}_{R_{\tilde{y}}(\beta)}.$$

Notice that $R_z(\gamma)$ and $R_z(\alpha)$ commute since they are rotations around the same axis. Consequently,

$$R_{z'}(\gamma) = R_z(\alpha)\, R_y(\beta)\, R_z(\gamma)\, R_y(-\beta)\, R_z(-\alpha) \tag{3.41}$$

can also be expressed in terms of rotations around the reference axes. Finally, replacing Eqs. (3.39) and (3.41) in Eq. (3.38) we can express the complete rotation as

$$\begin{aligned} R &= R_{z'}(\gamma)\, R_{\tilde{y}}(\beta)\, R_z(\alpha) \\ &= \underbrace{R_z(\alpha)\, R_y(\beta)\, R_z(\gamma)\, R_y(-\beta)\, R_z(-\alpha)}_{R_{z'}(\gamma)}\; \underbrace{R_z(\alpha)\, R_y(\beta)\, R_z(-\alpha)}_{R_{\tilde{y}}(\beta)}\; R_z(\alpha) \\ &= R_z(\alpha)\, R_y(\beta)\, R_z(\gamma). \end{aligned}$$

The operator $\hat{D}(R)$ of the representation of the rotation having Euler angles α, β and γ is therefore

$$\begin{aligned}\hat{D}(\alpha,\beta,\gamma) &= \hat{D}_z(\alpha) \quad \hat{D}_y(\beta) \quad \hat{D}_z(\gamma) \\ &= e^{-i\frac{J_z}{\hbar}\alpha} \quad e^{-i\frac{J_y}{\hbar}\beta} \quad e^{-i\frac{J_z}{\hbar}\gamma}\,, \end{aligned} \tag{3.42}$$

where we have used Eq. (3.37). This expression cannot be simplified straightforwardly since $[\hat{J}_y, \hat{J}_z] \neq 0$.

Example: Spin 1/2 systems

We would like to obtain $\hat{D}(\alpha,\beta,\gamma)$ explicitly for the case $J = 1/2$ by using Eq. (3.42) and expressing it in matrix form on the basis of eigenstates $|\, j,m \,\rangle$ of $\hat{J}^2$ and $\hat{J}_z$. Since the basis states are eigenstates of $\hat{J}_z$, the only complicated step is the application of $e^{-i\frac{\hat{J}_y}{\hbar}\beta}$. We consider the irreducible representation or Wigner coefficients

$$d^{(j)}_{m'm}(\beta) = \langle\, jm' \,|\, \exp\left\{\frac{-i\,\hat{J}_y\beta}{\hbar}\right\} |\, jm \,\rangle$$

corresponding to a rotation around the y axis. For $j = 1/2$ we have

$$e^{-i\frac{\hat{J}_y}{\hbar}\beta} = e^{-i\frac{\sigma_y\beta}{2}} = \sum_{k=0}^{+\infty} \frac{1}{k!}\left(\frac{-i\beta}{2}\right)^k (\sigma_y)^k\,, \tag{3.43}$$

where $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is a Pauli matrix. Knowing that $(\sigma_y)^2 = \mathbb{1}$ and splitting even-k and odd-k terms in the Taylor expansion (3.43) we obtain

$$\begin{aligned}\hat{D}^{(1/2)}(\beta) &= \sum_{k=0}^{+\infty} \frac{1}{(2k)!}(-1)^k \left(\frac{\beta}{2}\right)^{2k} \mathbb{1} + \sum_{k=0}^{+\infty} \frac{-i}{(2k+1)!}(-1)^k \left(\frac{\beta}{2}\right)^{2k+1} \sigma_y \\ &= \cos\left(\frac{\beta}{2}\right)\mathbb{1} + \sin\left(\frac{\beta}{2}\right)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{bmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{bmatrix}. \end{aligned}$$

Including the contributions of the rotations around the z axis, the rotation operator of spin-1/2 kets (spinors) for arbitrary rotation angles α, β and γ is given by

$$\begin{aligned}\hat{D}^{(1/2)}(\alpha,\beta,\gamma) &= \begin{pmatrix} e^{-i\frac{\alpha}{2}} & 0 \\ 0 & e^{i\frac{\alpha}{2}} \end{pmatrix}\begin{pmatrix} \cos\beta/2 & -\sin\beta/2 \\ \sin\beta/2 & \cos\beta/2 \end{pmatrix}\begin{pmatrix} e^{-i\frac{\gamma}{2}} & 0 \\ 0 & e^{i\frac{\gamma}{2}} \end{pmatrix} \\ &= \begin{bmatrix} e^{-i(\alpha+\gamma)/2}\cos(\beta/2) & -e^{-i(\alpha-\gamma)/2}\sin(\beta/2) \\ e^{i(\alpha-\gamma)/2}\sin(\beta/2) & e^{i(\alpha+\gamma)/2}\cos(\beta/2) \end{bmatrix}. \end{aligned}$$

This is the $j = 1/2$ irreducible representation of the rotation operator $\hat{D}(\alpha, \beta, \gamma)$. Notice that $\hat{D}^{(1/2)}(0, 2\pi, 0) = -\mathbb{1}$ and

$$\hat{D}(0, \pi, 0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.44}$$

This reflects the change of sign of a spin-1/2 spinor upon a 2π rotation around any axis (here the y axis). The Eq. (3.44) shows how the (-1) phase shift comes up by combining two π rotations, each one of the form $|+\rangle \to |-\rangle$ and $|-\rangle \to -|+\rangle$, where $|\pm\rangle \equiv |1/2, \pm 1/2\rangle$.

3.27 Physical significance of $\hat{D}^{(j)}(R)$:

Consider a state $|jm\rangle$ with defined angular momentum j and $\hat{J}_z = m$ and rotate it according to some element R of SO(3). The result is

$$|jm, R\rangle = \hat{D}(R)\,|jm\rangle,$$

where $\hat{D}(R)$ is the representation of the rotation operator. Using the completeness of the $\{|jm\rangle\}$ basis one has

$$|jm, R\rangle = \sum_{j'm'} |j'm'\rangle\langle j'm'|\,\hat{D}(R)\,|jm\rangle.$$

This can be significantly simplified by noting that, $\langle j'm'|\,\hat{D}(R)\,|jm\rangle = \delta_{jj'}\langle jm'|\,\hat{D}(R)\,|jm\rangle$, since there are no matrix elements of $\hat{D}(R)$ between different j's. In other words, rotations do not change the absolute value of the angular momentum. In fact,

$$\hat{J}^2\hat{D}(R)\,|jm\rangle = \hat{D}(R)\,\hat{J}^2\,|jm\rangle = \hbar^2 j(j+1)\,\hat{D}(R)\,|jm\rangle$$

since $[\hat{\vec{J}}, \hat{J}^2] = \vec{0}$. This means that $\hat{D}(R)\,|jm\rangle$ is an eigenstate of $\hat{J}^2$ with eigenvalue $j(j+1)\hbar^2$. Consequently,

$$|jm, R\rangle = \sum_{m'} |jm'\rangle \underbrace{\langle jm'|\,\hat{D}(R)\,|jm\rangle}_{d^{(j)}_{m'm}},$$

where $d^{(j)}_{m'm} = \langle jm'|\,\hat{D}(R)\,|jm\rangle$ is the matrix form of the $2j+1$ irreducible representation of SO(3). Therefore, the matrix element $d^{(j)}_{m'm}(R)$ as a function of m' for m and j fixed represents the probability amplitude for a R-rotated ket $|jm\rangle$ to be found in the state $|jm'\rangle$.

3.28 General form of $\hat{D}^{(j)}(\alpha, \beta, \gamma)$

The general form of the irreducible representations of SO(3) can be obtained in a similar way as the above discussed spin-1/2 case. Given the Euler angles α, β and γ we can write

$$\begin{aligned} d^{(j)}_{mm'}(\alpha,\beta,\gamma) &= \langle\, jm \,|\, e^{-i\frac{\hat{J}_z\alpha}{\hbar}}\; e^{-i\frac{\hat{J}_y\beta}{\hbar}}\; e^{-i\frac{\hat{J}_z\gamma}{\hbar}} \,|\, jm' \,\rangle \\ &= e^{-i(m\alpha+m'\gamma)} \langle\, jm \,|\, e^{-i\frac{\hat{J}_y\beta}{\hbar}} \,|\, jm' \,\rangle. \end{aligned}$$

The non-trivial part is always the representation

$$d^{(j)}_{mm'}(\beta) = \langle\, jm \,|\, e^{-i\frac{\hat{J}_y\beta}{\hbar}} \,|\, jm' \,\rangle$$

of the rotation around the y axis, for which a close expression in terms of sums of powers of $\cos(\beta/2)$ and $\sin(\beta/2)$ can be obtained [1]:

$$\begin{aligned} d^{(j)}_{mm'}(\beta) &= \sum_k (-1)^{k+m-m'} \frac{\sqrt{(j+m)!\,(j-m)!\,(j+m')!\,(j-m')!}}{(j+m'-k)!\,k!\,(j-k-m)!\,(k+m-m')!} \times \\ &\quad \times \left(\cos\frac{\beta}{2}\right)^{2j-2k-m+m'} \left(\sin\frac{\beta}{2}\right)^{2k+m-m'}, \end{aligned}$$

where the sum over k runs over all integer values of k for which the factorials in the denominator are non-negative. In a more compact form we can write

$$\begin{aligned} d^{(j)}_{mm'}(\beta) =& (-1)^{m-m'} \sqrt{\frac{(j+m)!\,(j-m)!}{(j+m')!\,(j-m')!}} \left(\cos\frac{\beta}{2}\right)^{2j} \times \\ &\times \sum_k (-1)^k \binom{j+m'}{k} \binom{j-m'}{j-k-m} \left(\tan\frac{\beta}{2}\right)^{m-m'+2k}. \end{aligned}$$

These matrix elements satisfy the following useful relations

$$\begin{aligned} d^{(j)}_{mm'}(\beta) &= (-1)^{m-m'}\, d^{(j)}_{m'm}(\beta)\,, \\ d^{(j)}_{mm'}(\beta) &= d^{(j)}_{-m',-m}(\beta) \end{aligned}$$

and

$$d^{(j)}_{mm'}(\beta) = d^{(j)}_{m'm}(-\beta)\,.$$

The verification is left as an exercise.

3.29 Transformation of scalar and spinor fields under rotation

We consider a spinless particle in a state $|\,\Psi\,\rangle$ with wave function $\Psi(\vec{x}) = \langle\,\vec{x}\,|\,\Psi\,\rangle$ and we would like to determine how $\Psi(\vec{x})$ changes upon rotation. We know that

$$|\,\Psi\,\rangle \xrightarrow{R} |\,\Psi'\,\rangle = \hat{D}(R)\,|\,\Psi\,\rangle.$$

For a scalar field we therefore have

$$\Psi(\vec{x}) \to \Psi'(\vec{x}) = \Psi(R^{-1}\vec{x})\,.$$

This can be explicitly demonstrated by noting that

$$\hat{D}^{\dagger}(R)\,|\,\vec{x}\,\rangle = \hat{D}(R^{-1})\,|\,\vec{x}\,\rangle = |\,R^{-1}\vec{x}\,\rangle\,,$$

since $\vec{x}$ transforms like a vector or, in other words, the rotation of state located at $\vec{x}$ is a state located at $R\,\vec{x}$. Therefore,

$$\langle\,R^{-1}\vec{x}\,| = \langle\,\vec{x}\,|\,\hat{D}(R)\,,$$

which implies

$$\Psi'(\vec{x}) = \langle\,\vec{x}\,|\,D(R)\,|\,\Psi\,\rangle = \langle\,R^{-1}\vec{x}\,|\,\Psi\,\rangle = \Psi(R^{-1}\vec{x})\,.$$

We focus now on *spinor fields* corresponding to $S = 1/2$. Consider an electron in a spin-up state along the z axis, which we denote by $|\,1/2\,\rangle$, $|\,+\,\rangle$, or $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ in Pauli's two component form. What is the probability amplitude of finding the electron with its spin pointing in the direction $\hat{n}$, more precisely, with spin-up or spin-down along the axis $\hat{n}$?

If $\hat{n}$ has polar and azimuthal angles (θ, φ), the rotation that brings $\hat{z} \to \hat{n}$ has Euler angles $\alpha = \varphi,\ \beta = \theta$ and $\gamma = 0$. Using the corresponding rotation operator for spin $1/2$, i.e.,

$$\hat{D}^{(1/2)}(\alpha = \varphi, \beta = \theta, \gamma = 0) = \begin{bmatrix} e^{-i\,\varphi/2}\cos(\theta/2) & -e^{-i\,\varphi/2}\sin(\theta/2) \\ e^{i\,\varphi/2}\sin(\theta/2) & e^{i\,\varphi/2}\cos(\theta/2) \end{bmatrix}, \tag{3.45}$$

we can express $|\,+,\hat{z}\,\rangle$ in terms of $|\,+,\hat{n}\,\rangle$ and $|\,-,\hat{n}\,\rangle$ as

$$|\,+,\hat{z}\,\rangle = |\,+,\hat{n}\,\rangle\,\langle\,+,\hat{n}\,|\,+,z\,\rangle \quad + \quad |\,-,\hat{n}\,\rangle\,\langle\,-,\hat{n}\,|\,+,z\,\rangle\,,$$

where $|\,+,\hat{n}\,\rangle$ ($|\,-,\hat{n}\,\rangle$) refers to a spin-up (spin-down) state along the axis $\hat{n}$. Alternatively, we can obtain $|\,+,\hat{n}\,\rangle$ by rotating $|\,+,\hat{z}\,\rangle$ in the direction of $\hat{n}$, which corresponds to multiplying $\hat{D}^{(1/2)}(R)$ by $\begin{pmatrix}1\\0\end{pmatrix} = |\,+,\hat{z}\,\rangle$.

Recalling that $\begin{pmatrix}1\\0\end{pmatrix} = |\,+,\hat{z}\,\rangle$ and $\begin{pmatrix}0\\1\end{pmatrix} = |\,-,\hat{z}\,\rangle$, and using Eq. (3.45) for $\hat{D}^{(1/2)}(R)$ we obtain

$$|\,+,\hat{n}\,\rangle = e^{-i\,\varphi/2}\cos(\theta/2)\,|\,+,z\,\rangle \;+\; e^{i\,\varphi/2}\sin(\theta/2)\,|\,-,z\,\rangle$$

and

$$\langle +, z \,|\, +, \hat{n} \,\rangle = e^{-i\,\varphi/2} \cos(\theta/2).$$

Finally, it is interesting to consider the rotation of an arbitrary spin-1/2 state $|\,\Psi\,\rangle$ which we write as

$$|\,\Psi\,\rangle = \sum_{\mu=-1/2}^{1/2} |\,\Psi_\mu\,\rangle\,|\,\mu, z\,\rangle = |\,\Psi_{1/2}\,\rangle \begin{pmatrix} 1 \\ 0 \end{pmatrix} + |\,\Psi_{-1/2}\,\rangle \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

where $|\,\Psi_\mu\,\rangle$ with $\mu = 1/2$ or $-1/2$ is a ket in the usual Hilbert space of the classical variables and $|\,\mu, z\,\rangle$ is a vector in the two dimensional spin space.

In order to determine the coordinate representation of $|\,\Psi\,\rangle$ we take the scalar product with $\langle\,\vec{x}\,|$ and obtain

$$\Psi(x) = \langle\,\vec{x}\,|\,\Psi\,\rangle = \sum_\mu \Psi_\mu(\vec{x})\,|\,\mu\,\rangle = \begin{bmatrix} \Psi_{1/2}(\vec{x}) \\ \Psi_{-1/2}(\vec{x}) \end{bmatrix} = \begin{bmatrix} \Psi_+(\vec{x}) \\ \Psi_-(\vec{x}) \end{bmatrix}.$$

The rotated ket is then given by

$$|\,\Psi'\,\rangle = \hat{D}_{\rm orb}(R)\;\hat{D}^{(1/2)}(R)\,|\,\Psi\,\rangle,$$

where $\hat{D}_{\rm orb}(R)$ acts on the spatial degrees of freedom exactly as in the spinless case and $\hat{D}^{(1/2)}(R)$ acts on the spin degree of freedom. The explicit form of the opertors is well known:

$$\hat{D}_{\rm orb}(R) = e^{-\frac{i}{\hbar}\,\hat{\vec{L}}\cdot\hat{n}\,\phi}$$

and

$$\hat{D}_{\rm spin}(R) = \hat{D}^{(1/2)}(R) = e^{-\frac{i}{\hbar}\,\hat{\vec{S}}\cdot\hat{n}\,\phi}.$$

Notice that $\hat{\vec{L}}$ and $\hat{\vec{S}}$ act on different variables. Therefore we have $[\hat{\vec{L}}, \hat{\vec{S}}] = 0$. It follows that

$$\hat{D}(R) = \hat{D}_{\rm orb}(R)\;\hat{D}_{\rm spin}(R) = e^{-\frac{i}{\hbar}\,\hat{\vec{J}}\cdot\hat{n}\,\phi}$$

with $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$. Thus,

$$\begin{aligned} \Psi'(\vec{x}) = \langle\,\vec{x}\,|\,\Psi'\,\rangle &= \sum_\mu \langle\,\vec{x}\,|\,\hat{D}_{\rm orb}(R)\,|\,\Psi_\mu\,\rangle\;\hat{D}^{(1/2)}(R)\,|\,\mu\,\rangle \\ &= \sum_{\mu,\,\mu'} \underbrace{\langle\,R^{-1}\vec{x}\,|\,\Psi_\mu\,\rangle}_{\Psi_\mu(R^{-1}\vec{x})}\;|\,\mu'\,\rangle\;\underbrace{\langle\,\mu'\,|\,\hat{D}^{(1/2)}(R)\,|\,\mu\,\rangle}_{\hat{D}^{(1/2)}_{\mu'\mu}(R)}. \end{aligned}$$

The transformed state can be written as

$$\Psi'(\vec{x}) = \sum_{\mu'} \Psi'_{\mu'}(x)\,|\,\mu'\,\rangle,$$

where

$$\Psi'_{\mu'}(x) = \sum_{\mu} \hat{D}^{(1/2)}_{\mu'\mu}(R)\ \Psi_{\mu}(R^{-1}\vec{x}),$$

or in matrix form as

$$\begin{bmatrix} \Psi'_{+}(\vec{x}) \\ \Psi'_{-}(\vec{x}) \end{bmatrix} = \hat{D}^{(1/2)}(R) \begin{bmatrix} \Psi_{+}(R^{-1}\vec{x}) \\ \Psi_{-}(R^{-1}\vec{x}) \end{bmatrix}.$$

A pair of functions $\Psi_+(x)$ and $\Psi_-(x)$ that transforms in this way under rotations is known as spin-1/2 field, $S = \frac{1}{2}$ spinor field or, simply spinor field. The above discussion of the rotation of spin $1/2$ quantum states can be generalized to particles having an arbitrary intrinsic spin S.

3.30 Vector operators

We consider the Hilbert space $\mathbb{V}$ of kets $|\,\alpha\,\rangle$ and the representation $\hat{D}(R)$ of the rotation group in $\mathbb{V}$. Under the rotation R the ket $|\,\alpha\,\rangle$ is changed into $|\,\alpha'\,\rangle = \hat{D}(R)\,|\,\alpha\,\rangle$. The corresponding transformation of the operator $\hat{A}$ is

$$\hat{A}_T = \hat{D}^{\dagger}(R)\,\hat{A}\,\hat{D}(R).$$

We are interested in discussing the properties of particular sets of 3 operators $\hat{V}_i$ with $i = 1,\ 2$ and 3 (or equivalently $\hat{V}_x,\ \hat{V}_y$ and $\hat{V}_z$) which transform under rotations in the same way as conventional vectors in $\mathbb{R}^3$. One therefore defines that $\hat{\vec{V}} = (\hat{V}_1,\ \hat{V}_2,\ \hat{V}_3)$ is a *Cartesian vector operator* if and only if for all rotations R the transformed operators

$$\hat{V}_i^T = \hat{D}^{\dagger}(R)\,\hat{V}_i\,\hat{D}_R$$

satisfy

$$\hat{D}^{\dagger}(R)\ \hat{V}_i\ \hat{D}(R) = \sum_{j=1}^{3} R_{ij}\,\hat{V}_j, \tag{3.46}$$

for $i = 1-3$, where R_{ij} is the SO(3) *matrix* of the rotation R. This vector transformation relation can be written in the more compact and clear matrix-vector form

$$\hat{D}^{\dagger}(R)\ \hat{\vec{V}}\ \hat{D}(R) = R\,\hat{\vec{V}}\,. \tag{3.47}$$

Notice that the vector components $\hat{V}_i$ are operators acting on $\mathbb{V}$, while the $R_{ij} \in \mathbb{R}$ are the real elements of an SO(3) matrix. Therefore, in terms of the matrix elements between quantum states $|\,\alpha\,\rangle$, $|\,\beta\,\rangle$ and their transformed $|\,\alpha'\,\rangle = \hat{D}|\,\alpha\,\rangle$ and $|\,\beta'\,\rangle = \hat{D}|\,\beta\,\rangle$, the condition (3.46) is equivalent to

$$\langle\,\alpha'\,|\,\hat{V}_i\,|\,\beta'\,\rangle = \langle\,\alpha\,|\,\hat{D}^{\dagger}(R)\ \hat{V}_i\ \hat{D}(R)\,|\,\beta\,\rangle = \sum_{j} R_{ij}\,\langle\,\alpha\,|\,\hat{V}_j\,|\,\beta\,\rangle.$$

for all $|\,\alpha\,\rangle$ and $|\,\beta\,\rangle$.

Vector operators satisfy specific commutation rules with the angular momentum operator $\hat{\vec{J}}$. In order to find them we consider an infinitesimal rotation

$$\hat{D}(R) = 1 - \frac{i}{\hbar}\hat{\vec{J}}\cdot\hat{n}\,\delta\phi + O\left(\delta\phi^2\right) \tag{3.48}$$

around an arbitrary axis $\hat{n}$. Replacing Eq. (3.48) for $\hat{D}(R)$ in the left-hand side of the definition (3.46) one obtains

$$\hat{V}_i + \frac{i}{\hbar}\,\delta\phi\,[\hat{\vec{J}}\cdot\hat{n}, \hat{V}_i] = \sum_j R_{ij}(\hat{n},\delta\phi)\,\hat{V}_j\,, \tag{3.49}$$

where $R_{ij}(\hat{n},\delta\phi)$ is the matrix of an infinitesimal rotation around $\hat{n}$. For $\hat{n}$ along the $\hat{z}$ axis we have

$$R_{ij}(\hat{z},\delta\phi) = \begin{pmatrix} 1 & -\delta\phi & 0 \\ \delta\phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and thus

$$\hat{\vec{V}} + \frac{i}{\hbar}\,\delta\phi\,[\hat{J}_z, \hat{\vec{V}}] = \begin{pmatrix} 1 & -\delta\phi & 0 \\ \delta\phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\,\hat{\vec{V}}\,. \tag{3.50}$$

Similar relations are obtained for rotations around the axis x and y by cyclic permutation of the indices.

Splitting the vector identity (3.50) in components we obtain

$$\hat{V}_x + \frac{i}{\hbar}\,\delta\phi\,[\hat{J}_z, \hat{V}_x] = \hat{V}_x - \delta\phi\,\hat{V}_y + O(\delta\phi^2)\,,$$

$$\hat{V}_y + \frac{i}{\hbar}\,\delta\phi\,[\hat{J}_z, \hat{V}_y] = \delta\phi\,\hat{V}_x + \hat{V}_y + O(\delta\phi^2)\,,$$

and

$$\hat{V}_z + \frac{i}{\hbar}\,\delta\phi\,[\hat{J}_z, \hat{V}_z] = \hat{V}_z \qquad\qquad + O(\delta\phi^2)\,,$$

which implies

$$[\hat{V}_x, \hat{J}_z] = -i\hbar\,\hat{V}_y\,,$$
$$[\hat{V}_y, \hat{J}_z] = i\hbar\,\hat{V}_x\,,$$

and

$$[\hat{V}_z, \hat{J}_z] = 0.$$

Finally, taking into account that the indices can be cyclically permuted, we conclude that

$$[\hat{V}_i, \hat{J}_j] = i\hbar\, \varepsilon_{ijk}\, \hat{V}_k \tag{3.51}$$

or equivalently

$$[\hat{J}_i, \hat{V}_j] = i\hbar\, \varepsilon_{ijk}\, \hat{V}_k.$$

It is easy to see, reproducing the algebraic step in reversed order, that the commutation rules (3.51) between $\hat{\vec{V}}$ and $\hat{\vec{J}}$ lead to Eq. (3.46) provided that the rotation R is infinitesimal. Consequently, Eq. (3.51) is a necessary and sufficient condition characterizing vector operators for infinitesimal rotations. Moreover, one can show, by using the group properties of $\hat{D}(R)$, that Eq. (3.51) is also a sufficient condition in order that $\hat{\vec{V}}$ satisfies the vector-operator transformation law (3.47) for finite rotations. Let us assume that the vector transformation property

$$\hat{D}^\dagger(R_i)\, \hat{\vec{V}}\, \hat{D}(R_i) = R_i\, \hat{\vec{V}}$$

holds for infinitesimal rotations R_i. One expresses a finite rotation as the succession of n rotations around the same axis $\hat{n}$, i.e., the matrix product $R = R_n\, R_{n-1}\, \ldots\, R_2\, R_1$. Knowing that $\hat{D}$ is a representation of SO(3), we have

$$\begin{aligned}
\hat{D}^\dagger(R)\, \hat{\vec{V}}\, \hat{D}(R) &= \hat{D}^\dagger(R_n)\, \ldots\, \underbrace{\hat{D}^\dagger(R_1)\, \hat{\vec{V}}\, \hat{D}(R_1)}_{R_1\, \hat{\vec{V}}}\, \ldots\, \hat{D}(R_n) \\
&= R_1\, \hat{D}^\dagger(R_n)\, \ldots\, \hat{D}^\dagger(R_2)\, \hat{\vec{V}}\, \hat{D}(R_2)\, \ldots\, \hat{D}(R_n) \\
&= R_1\, R_2\, \ldots\, R_n\, \hat{\vec{V}} = R\, \hat{\vec{V}}.
\end{aligned}$$

Here we have used that R_i are simple SO(3) matrices (not quantum mechanical operators) which therefore commute with all $\hat{D}^\dagger(R)$. The order of the matrices R_i in the product is irrelevant since they commute because the rotation axis is the same. One concludes that Eq. (3.51) is equivalent to Eq. (3.46) for all R and can thus be used to identify vector operators.

Notice the remarkable simplicity of the transformation law of vector operators and the intuitive geometric association between "rotating a ket" $|\,\alpha\,\rangle$ as $\hat{D}(R)\,|\,\alpha\,\rangle$ in the Hilbert space $\mathbb{V}$, "rotating an operator" as $\hat{D}^\dagger(R)\, \hat{\vec{V}}\, \hat{D}(R)$, and rotating a vector operator as $R\, \hat{\vec{V}}$ in $\mathbb{R}^3$. Like in the case of scalars (i.e., invariants upon rotation), basically all vector operators have a clear physical meaning: position $\vec{r} = (x, y, z)$, momentum $\vec{p} = (p_x, p_y, p_z)$, their vector product $\vec{L} = \vec{r} \times \vec{p}$ (a pseudo vector actually), the intrinsic spin $\vec{S}$, the total angular momentum $\vec{J} = \sum_i \vec{j}_i$ itself, as well as the individual components $\vec{j}_i$ in a many-particle system. The reader is encouraged to verify that Eq. (3.51) is satisfied in all these cases.

3.31 Tensor operators

The notion of vector $V_i \to V_i' = \sum_j R_{ij}\, V_j$ can be extended to more complex objects called Cartesian tensors of rank k, which are defined by requiring the following transformation law under rotation:

$$T'_{i_1,i_2\ldots i_k} = \sum_{j_1,j_2\ldots j_k} R_{i_1 j_1}\, R_{i_2 j_2} \ldots R_{i_k j_k}\, T_{j_1,j_2\ldots j_k}\,.$$

A Cartesian tensor of rank 1 is nothing but a vector. An example of a simple tensor of rank 2 is the dyadic

$$T_{ij} = U_i\, V_j$$

where $\vec{U}$ and $\vec{V}$ are vectors. It follows that $T_{ij}{}' = U_i'\, V_j' = \sum_{kl}\, R_{il}\, R_{jk}\, U_l\, V_k$.

The problem with Cartesian tensors is that they are reducible. Thus, they do not transform in the simplest possible form. For instance, any tensor of second rank of the form $U_i V_i$ can be express as the sum of three distinct contribution:

$$U_i V_j \;=\; \underbrace{\frac{\vec{U}\cdot\vec{V}}{3}\,\delta_{ij}}_{\text{scalar}} + \underbrace{\frac{(U_i V_j - U_j V_i)}{2}}_{\substack{\text{antisymmetric tensor}\\ \varepsilon_{ijk}(\vec{U}\times\vec{V})_k}} + \underbrace{\left(\frac{(U_i V_j + U_j V_i)}{2} - \frac{\vec{U}\cdot\vec{V}}{3}\,\delta_{ij}\right)}_{\text{traceless symmetric}}.$$

The first term is a scalar, the second one an antisymmetric tensor of rank 2 and the third one a symmetric tensor of rank 2. Clearly, upon rotations, these three terms either remain invariant (scalar product) or transform into other terms having the same symmetric or antisymmetric character. Thus, the dyadic is reducible.

In quantum mechanics one therefore considers the *irreducible tensor operators of rank* k. They are given by a set of $(2k+1)$ components $q = -k, -k+1, \ldots, k-1, k$ which by definition transform under rotation as

$$\hat{T}_q'^{(k)} = \hat{D}^\dagger(R)\, \hat{T}_q^{(k)}\, \hat{D}(R) = \sum_{q'=-k}^{k} d^{*(k)}_{qq'}(R)\, \hat{T}_{q'}^{(k)}\,. \tag{3.52}$$

where $d^{*(k)}_{qq'}(R)$ are the complex conjugate of the matrix elements of the irreducible k-dimensional representation of the $SO(3)$ group whose physical significance and general form where presented in Secs. 3.27 and 3.28. Replacing R by R^{-1} the transformation law can be written as

$$\hat{D}(R)\, \hat{T}_q^{(k)}\, \hat{D}^\dagger(R) = \sum_{q'=-k}^{k} d^{(k)}_{q'q}(R)\, \hat{T}_{q'}^{(k)}.$$

Examples of irreducible tensor operators are *scalars*, like

$$\hat{T}_0^{(0)} = \hat{\vec{U}}\cdot\hat{\vec{V}}\,,$$

having rank $k = 0$, and spherical tensors of rank 1, which are obtained from Cartesian vector operators as

$$\hat{T}^{(1)}_{+1} = -\frac{1}{\sqrt{2}}\,(\hat{V}_x + i\,\hat{V}_y) = -\frac{1}{\sqrt{2}}\,\hat{V}_+,$$
$$\hat{T}^{(1)}_{-1} = \frac{1}{\sqrt{2}}\,(\hat{V}_x - i\,\hat{V}_y) = \frac{1}{\sqrt{2}}\,\hat{V}_-$$

and

$$\hat{T}^{(1)}_{0} = \hat{V}_z.$$

The previous examples can be generalized to a whole family of tensor operators which derive from spherical harmonics. In order to motivate the definition of spherical tensors of rank k it is useful to review the transformation law for spherical harmonics $Y_{lm}(\theta, \varphi)$. Consider a spinless particle in a potential having spherical symmetry. We know that the eigenfunctions can be written as

$$\Psi_{lm}(\vec{r}) = \langle\, \vec{r}\,|\, l\, m\,\rangle = R_l(r)\; Y_{lm}(\theta, \varphi),$$

where $\vec{r}$ is the position vector whose spherical coordinates are r, θ and φ. As already discussed, this is a consequence of the rotational symmetry of the Hamiltonian, which implies that the $2l + 1$ states with different m are degenerate. Since the angular dependence is common to all problems with spherical symmetry it is useful to isolate it by defining a *direction eigenket* $|\,\hat{n}\,\rangle$ given by

$$Y_{lm}(\theta, \varphi) = \langle\, \hat{n}\,|\, l\, m\,\rangle = Y_{lm}(\hat{n}).$$

$Y_{lm}(\theta, \varphi)$ is the amplitude for a state $|\, l\, m\,\rangle$ to be found in the direction $\hat{n}$ specified by θ and φ. If we rotate the direction eigenket $|\,\hat{n}\,\rangle$ we obtain $\hat{D}(R)\,|\,\hat{n}\,\rangle = |\,\hat{n}'\,\rangle$ where $\hat{n}' = R\,\hat{n}$ points along the rotated direction $\hat{n}'$.

Let us see how the $Y_{lm}(\theta, \varphi)$ transforms under rotations. We know that

$$|\,\hat{n}\,\rangle \xrightarrow{R} \hat{D}(R)\,|\,\hat{n}\,\rangle = |\,n'\,\rangle$$

and we want to express the probability amplitude

$$Y_{lm}(\theta', \varphi') = Y_{lm}(\hat{n}') = \langle\, \hat{n}'\,|\, l\, m\,\rangle$$

of finding the state $|\, lm\,\rangle$ along the direction $\hat{n}'$ in terms of the probability amplitudes

$$Y_{lm'}(\theta, \varphi) = Y_{lm'}(\hat{n}) = \langle\, \hat{n}\,|\, l\, m'\,\rangle.$$

of finding the states $|\, lm'\,\rangle$ along the direction $\hat{n}$. We start from

$$Y_{lm}(\hat{n}') = \langle\, n'\,|\, l\, m\,\rangle = \langle\, \hat{D}(R)\,\hat{n}\,|\, l\, m\,\rangle = \langle\, \hat{n}\,|\, \hat{D}(R^{-1})\,|\, l\, m\,\rangle\,. \tag{3.53}$$

Knowing that $\hat{D}(R)$ has no matrix elements between states having different l, we can write

$$\begin{aligned}\hat{D}(R^{-1})\,|\,l\,m\,\rangle &= \sum_{m'} |\,l\,m'\,\rangle\langle\,l\,m'\,|\,\hat{D}(R^{-1})\,|\,l\,m\,\rangle \\ &= \sum_{m'} |\,l\,m'\,\rangle\, d^{(l)}_{m'm}(R^{-1}),\end{aligned} \tag{3.54}$$

where $d^{(l)}_{m'm}(R^{-1}) = \langle\, lm'\,|\,\hat{D}(R^{-1})\,|\,lm\,\rangle$ stands for the $(2l+1)$-dimensional irreducible representation.

Multiplying Eq. (3.54) by $\langle\, n\,|$ we have

$$\langle\,\hat{n}\,|\,\hat{D}(R^{-1})\,|\,l\,m\,\rangle = \sum_{m'} \langle\,\hat{n}\,|\,l\,m'\,\rangle\, d^{(l)}_{m'm}(R^{-1}). \tag{3.55}$$

Replacing Eq. (3.55) in Eq. (3.53) we obtain

$$\langle\,n'\,|\,l\,m\,\rangle = \sum_{m'}\langle\,\hat{n}\,|\,l\,m'\,\rangle\, d^{(l)}_{m'm}(R^{-1})$$

or equivalently

$$Y_{lm}(\hat{n}') = \sum_{m'}\, d^{(l)}_{m'm}(R^{-1})\, Y_{lm'}(\hat{n}).$$

Using that $\hat{n}' = R\,\hat{n}$ this can be rewritten as

$$Y_{lm}(R\,\hat{n}) = \sum_{m'}\, d^{(l)}_{m'm}(R^{-1})\, Y_{lm'}(\hat{n})$$

or

$$Y_{lm}(R\,\hat{n}) = \sum_{m'}\, d^{(l)}_{mm'}(R)^*\, Y_{lm'}(\hat{n})\,. \tag{3.56}$$

One concludes that $d^{(l)}_{m'm}(R^{-1}) = d^{(l)}_{mm'}(R)^*$ represents the probability amplitude for finding the state $Y_{lm}(\theta',\varphi')$ with quantization axis $\hat{n}'$ in the state $Y_{lm'}(\theta,\varphi)$ with quantization axis $\hat{n}$. Comparing Eq. (3.56) with the definition of irreducible tensors (3.52) we conclude that the spherical harmonics Y_{lm} transform as the m component of an irreducible spherical tensor of rank l.

Let us now consider a vector operator $\hat{\vec{V}} = (\hat{V}_x,\ \hat{V}_y,\ \hat{V}_z)$ and construct operators of the form $Y_{lm}(\hat{\vec{V}})$ by replacing

$$\hat{n}_x = \frac{x}{r}\ \text{ by }\ \hat{V}_x\,, \qquad \hat{n}_y = \frac{y}{r}\ \text{ by }\ \hat{V}_y\,, \qquad \text{and} \quad \hat{n}_z = \frac{z}{r}\ \text{ by }\ \hat{V}_z$$

in $Y_{lm}(\hat{n})$. In this way we obtain, for example for $l = 1$,

$$Y_{10} = \sqrt{\frac{3}{4\pi}}\,\cos\theta = \sqrt{\frac{3}{4\pi}}\,\frac{z}{r} \longrightarrow \hat{T}^{(1)}_0 = \sqrt{\frac{3}{4\pi}}\,\hat{V}_z$$

and

$$Y_{1,\pm1} = \mp\sqrt{\frac{3}{4\pi}}\,\frac{x \pm i\,y}{\sqrt{2}\,r} \longrightarrow \hat{T}^{(1)}_{\pm1} = \sqrt{\frac{3}{4\pi}}\,\left(\mp\frac{\hat{V}_x \pm i\,\hat{V}_y}{\sqrt{2}}\right).$$

Similarly for $l = 2$ we have, for example,

$$Y_{2,\pm2} = \mp\sqrt{\frac{15}{32\pi}}\,\frac{(x \pm i\,y)^2}{r^2} \longrightarrow \hat{T}^{(2)}_{\pm2} = \sqrt{\frac{15}{32\pi}}\,\left(\hat{V}_x \pm i\,\hat{V}_y\right)^2. \tag{3.57}$$

It is easy to see that, if $\hat{\vec{V}}$ is a Cartesian vector operator, the operator $Y_{lm}(\hat{\vec{V}})$ is the m component of an irreducible spherical tensor of rank l. One should note first of all that

$$\hat{D}^\dagger(R)\; Y_{lm}(\hat{\vec{V}})\; \hat{D}(R) = Y_{lm}[\hat{D}^\dagger(R)\; \hat{\vec{V}}\; \hat{D}(R)],$$

since one can express $Y_{lm}(\hat{n})$ as a polynomial in $\hat{n}_x,\ \hat{n}_y$ and $\hat{n}_z$, and $\hat{D}^\dagger\,\hat{A}^k\,\hat{D} = (\hat{D}^\dagger\,\hat{A}\,\hat{D})^k$ for any operator $\hat{A}$ and any unitary operator $\hat{D}$. Moreover, taking into account that $\hat{\vec{V}}$ is a Cartesian vector we have

$$\hat{D}^\dagger(R)\; \hat{\vec{V}}\; \hat{D}(R) = R\,\hat{\vec{V}}\,.$$

Therefore, using Eq. (3.56), we obtain

$$\hat{D}^\dagger(R)\; Y_{lm}(\hat{\vec{V}})\; \hat{D}(R) = Y_{lm}(R\,\hat{\vec{V}}) = \sum_{m'}\; d^{(l)}_{mm'}(R)^*\; Y_{lm'}(\hat{\vec{V}})\,.$$

Comparing with the definition of spherical tensor (3.52) one concludes that

$$Y_{lm}(\vec{V}) = T^{(k)}_q \quad \text{with } k = l \text{ and } q = m.$$

Notice, however, that the definition of spherical tensor is more general than $Y_{lm}(\hat{\vec{V}})$, since it includes operators that cannot be written as $Y_{lm}(\hat{\vec{V}})$. An example is $(\hat{U}_x + i\,\hat{U}_y)\;(\hat{V}_x + i\,\hat{V}_y)$ with Cartesian vectors $\hat{\vec{U}}$ and $\hat{\vec{V}}$, which transforms as $\hat{T}^{(2)}_2$ ($k = 2$ and $q = 2$) like $(\hat{V}_x + i\,\hat{V}_y)^2$. See the example given in Eq. (3.57).

3.32 Commutation relations for irreducible tensors

Following the same reasoning as in the case of Cartesian vector operators we can relate the condition defining spherical tensor operators

$$\hat{D}^\dagger(R)\; \hat{T}^{(k)}_q\; \hat{D}(R) = \sum_{q'=-k}^{k}\; d^{(k)}_{qq'}(R)^*\; \hat{T}^{(k)}_{q'} \tag{3.58}$$

with the following commutation relations

$$\left[\hat{J}_z, \hat{T}_q^{(k)}\right] = \hbar\, q\, \hat{T}_q^{(k)} \tag{3.59}$$

and

$$\left[\hat{J}_\pm, \hat{T}_q^{(k)}\right] = \hbar\, \sqrt{(k \mp q)\,(k \pm q + 1)}\, \hat{T}_{q\pm1}^{(k)} \tag{3.60}$$

for $q = -k, \ldots, k$. These commutation relations are equivalent to the previous definition (3.58) of tensors of rank k in terms of $\hat{D}(R)$ and the irreducible $(2k+1)$-dimensional representation $d_{q\,q'}^{(k)}(R)$.

For the proof one considers infinitesimal rotations generated by $\hat{\vec{J}} \cdot \hat{n}$, where $\hat{n}$ is an arbitrary unit vector. We shall then choose $\hat{n} = \hat{x}$, $\hat{y}$, and $\hat{z}$ along the coordinate axes in order to construct the commutation relations for $\hat{J}_z$ from $\hat{z}$ and for $\hat{J}_\pm$ from $\hat{x} \pm i\,\hat{y}$. We start from the representation

$$\hat{D}(R) = 1 - \frac{i}{\hbar}\, \hat{\vec{J}} \cdot \hat{n}\, \varepsilon$$

of an infinitesimal rotation R and the definition

$$\hat{D}^\dagger(R)\; \hat{T}_q^{(k)}\; \hat{D}(R) = \sum_{q'}\; d_{qq'}^{(k)}(R)^*\; \hat{T}_{q'}^{(k)} = \sum_{q'}\; d_{q'q}^{(k)}(R^{-1})\; \hat{T}_{q'}^{(k)}$$

of spherical tensors. It follows that

$$\left(1 + \frac{i}{\hbar}\, \hat{\vec{J}} \cdot \hat{n}\, \varepsilon\right)\; \hat{T}_q^{(k)}\; \left(1 - \frac{i}{\hbar}\, \hat{\vec{J}} \cdot \hat{n}\, \varepsilon\right) = \sum_{q'}\; \hat{T}_{q'}^{(k)}\; \langle\, k\, q' \,|\, 1 + \frac{i}{\hbar}\, \hat{\vec{J}} \cdot \hat{n}\, \varepsilon \,|\, k\, q \,\rangle$$

and

$$[\hat{\vec{J}} \cdot \hat{n},\; \hat{T}_q^{(k)}] = \sum_{q'}\; \hat{T}_{q'}^{(k)}\; \langle\, k\, q' \,|\, \hat{\vec{J}} \cdot \hat{n} \,|\, k\, q \,\rangle\,.$$

We can now set $\hat{n} = \hat{z}$ to obtain

$$[\hat{J}_z,\; \hat{T}_q^{(k)}] = \hbar\, q\; \hat{T}_q^{(k)}\,,$$

and set $\hat{n} = \hat{x} \pm i\,\hat{y}$ to obtain

$$[\hat{J}_\pm, \hat{T}_q^{(k)}] = \sum_{q'}\; \hat{T}_{q'}^{(k)}\; \langle\, k\, q' \,|\, \hat{J}_\pm \,|\, k\, q \,\rangle,$$

which implies

$$[\hat{J}_\pm, \hat{T}_q^{(k)}] = \hbar\, \sqrt{(k \mp q)\,(k \pm q + 1)}\; \hat{T}_{q\pm1}^{(k)}\,.$$

This proofs that spherical tensor operators satisfy the commutation relations (3.59) and (3.60). It is easy to see that the converse holds for infinitesimal rotations R, simply by repeating the reasoning in reverse order. Finite rotations can be split into many infinitesimal rotations in the same way as discussed at the end of Sec. 3.30. Thus, the commutation relations (3.59) and (3.60) are a necessary and sufficient condition for spherical tensor operators of rank k.

3.33 Matrix elements of tensor operators: The Wigner Eckart Theorem

We can now turn our attention to the calculation of the matrix elements

$$\langle\, \alpha',\, j'\, m' \,|\, \hat{T}^{(k)}_q \,|\, \alpha,\, j\, m \,\rangle$$

of irreducible tensor operators between eigenstates of the total angular momentum J^2 and J_z. These matrix elements take a particularly simple form which exploits to a maximum the symmetry of the full rotational group. The final conclusion, known as Wigner-Eckart theorem, and the resulting selection rules are extremely important, not only in atomic physics, spectroscopy and the theory of radiation, but also in a number of phenomena in condensed-matter physics (e.g., magnetic phenomena) where a description in terms of localized atomic-like wave functions is a physically meaningful starting point.

First of all we can easily prove the *m-selection rule*. Let $|\, \alpha, jm \,\rangle$ and $|\, \alpha', j'm' \,\rangle$ be two kets with well-defined J^2 and J_z. Since $[\hat{J}_z, \hat{T}^{(k)}_q] = \hbar\, q\, \hat{T}^{(k)}_q$ we have[7]

$$0 = \langle\, \alpha', j'm' \,|\, \hat{J}_z\, \hat{T}^{(k)}_q - \hat{T}^{(k)}_q \hat{J}_z - q\, \hat{T}^{(k)}_q \,|\, \alpha, jm \,\rangle = (m' - m - q)\, \langle\, \alpha', j'm' \,|\, \hat{T}^{(k)}_q \,|\, \alpha, jm \,\rangle.$$

Therefore

$$\langle\, \alpha', j'm' \,|\, \hat{T}^{(k)}_q \,|\, \alpha, jm \,\rangle = 0 \quad \text{unless} \quad m' = m + q\,. \tag{3.61}$$

We can picture this fundamental selection rule by saying that the q component of an spherical irreducible tensor $\hat{T}^{(k)}_q$ carries an orbital angular momentum q around the z axis!

The *Wigner-Eckart theorem* is actually much more powerful. It can be stated as follows:

$$\langle\, \alpha', j'm' \,|\, \hat{T}^{(k)}_q \,|\, \alpha, jm \,\rangle = \langle\, jk; mq \,|\, jk; j'm' \,\rangle\, \frac{\langle\, \alpha' j' \,\|\, \hat{T}^{(k)} \,\|\, \alpha j \,\rangle}{\sqrt{2j+1}}\,,$$

where $\langle\, \alpha' j' \,\|\, \hat{T}^{(k)} \,\|\, \alpha j \,\rangle$ is the so-called *reduced matrix element* which is independent of m, m' and q and where the scalar products $\langle\, jk; mq \,|\, jk; j'm' \,\rangle$ are the *Clebsch-Gordan or Wigner coefficients*

$$\langle\, j_1 j_2;\, m_1 m_2 \,|\, j_1 j_2;\, j'm' \,\rangle$$

for adding the angular momenta j and k to j' ($j_1 = j, j_2 = k \to j'$ and m') or in more compact form $\vec{j} \oplus \vec{k} = \vec{j}'$. We recover of course the m-selection rule $(m_1 + m_2 = m') \to m + q = m'$.

The problem of computing matrix elements between wave functions having spherical symmetry is separated in two perfectly distinct parts:

[7] For simplicity we set $\hbar = 1$ in the following.

(1) On the one side we have the reduced matrix element

$$\langle\, \alpha' j' \,\|\, \hat{T}^{(k)} \,\|\, \alpha j \,\rangle$$

which is obtained by computing $\langle\, \alpha', j'm' \,|\, \hat{T}^{(k)}_q \,|\, \alpha, jm \,\rangle$ *once*, for one value of m, m' and q, and by dividing by the Clebsch-Gordan coefficient $\langle\, jk;\, mq \,|\, jk;\, j'm' \,\rangle$ corresponding to the chosen m, m' and q. For this one needs of course to have $\langle\, jk; mq \,|\, jk; j'm' \,\rangle \neq 0$. As suggested by the notation, $\langle\, \alpha' j' \,\|\, \hat{T}^{(k)} \,\|\, \alpha j \,\rangle$ is independent of m, m' and q. The factor $\dfrac{1}{\sqrt{2j+1}}$ is just a convention.

(2) On the other side we have all the dependence on m, m' and q —the so-called angular dependence— which is taken into account by universal (i.e., problem independent) geometrical factors given by the Clebsch-Gordan (CG) coefficients. Let us recall the definition of the CG coefficients:

$$|\, j_1 j_2;\, jm \,\rangle = \sum_{m_1,\, m_2} |\, j_1 j_2;\, m_1 m_2 \,\rangle\, \langle\, j_1 j_2;\, m_1 m_2 \,|\, j_1 j_2;\, jm \,\rangle,$$

where the kets on the right-hand side are eigenstates of $\hat{J}_{1z}$ and $\hat{J}_{2z}$, i.e.,

$$\hat{J}_{1z} \,|\, j_1 j_2;\, m_1 m_2 \,\rangle = m_1 \,|\, j_1 j_2;\, m_1 m_2 \,\rangle\,,$$
$$\hat{J}_{2z} \,|\, j_1 j_2;\, m_1 m_2 \,\rangle = m_2 \,|\, j_1 j_2;\, m_1 m_2 \,\rangle\,,$$

and the kets on the left-hand side are eigenstates of $\hat{J}^2$ and $\hat{J}_z$ with $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2$, i.e.,

$$(\hat{\vec{J}}_1 + \hat{\vec{J}}_2)^2 \,|\, j_1 j_2; jm \,\rangle = j(j+1) \,|\, j_1 j_2; jm \,\rangle\,,$$

$$(\hat{J}_{1z} + \hat{J}_{2z}) \,|\, j_1 j_2; jm \,\rangle = m \,|\, j_1 j_2; jm \,\rangle\,.$$

The CG coefficients can be obtained from the recursion relations to be discussed later on. They are also tabulated in most quantum mechanics and group-theory books [2].

Before closing this section it is worth to mention that the practical importance of the Wigner-Eckart theorem goes well beyond spherically symmetric problems, since any operator can be written as the sum of spherical tensor operators and any quantum state can be written as the superposition of eigenstates of the angular momentum.

3.34 The recursion relation for the Clebsch-Gordon coefficients

We compute $\langle\, j_1 j_2; m_1 m_2 \,|\, J_\pm \,|\, j_1 j_2; jm \,\rangle$ in two equivalent ways. First we operate with $(J_\pm)^\dagger = J_\mp$ on the bra or left-hand side (LHS) and second we operate with $J_\pm$ on the ket or right-hand side (RHS):

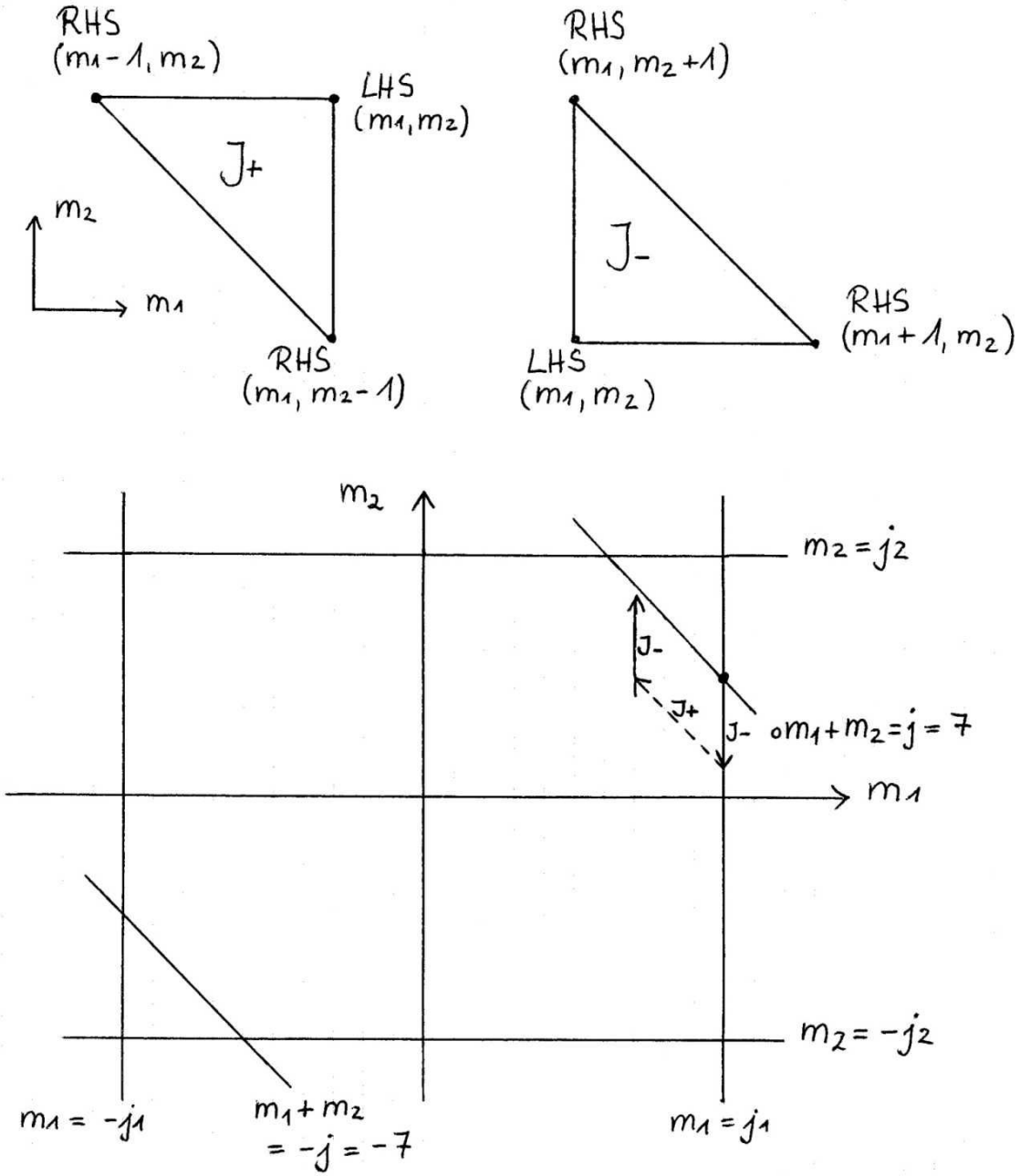
$$\langle\, j_1 j_2; m_1 m_2 \,|\, J_\pm \,|\, j_1 j_2; jm \,\rangle =$$

$$\text{LHS} \quad \longrightarrow \quad \sqrt{(j \mp m)\,(j \pm m + 1)}\, \langle\, j_1 j_2; m_1 m_2 \,|\, j_1 j_2; j\ m{\pm}1 \,\rangle =$$

$$\text{RHS} \quad \longrightarrow \quad = \sqrt{(j_1 \pm m_1)\,(j_1 \mp m_1 + 1)}\, \langle\, j_1 j_2; m_1 {\mp} 1\ m_2 \,|\, j_1 j_2; jm \,\rangle +$$

$$+ \sqrt{(j_2 \pm m_2)\,(j_2 \mp m_2 + 1)}\, \langle\, j_1 j_2; m_1\ m_2 {\mp} 1 \,|\, j_1 j_2; jm \,\rangle. \qquad (3.62)$$

m is defined entirely by m_1 and m_2 by the condition $m_1 + m_2 = m \pm 1$. The recursion relations (3.62) are triangular relations that can be illustrated by the following figure:



Given one coefficient, for example, for $m_1 = j_1$ and $m_2 + m_1 = j$ we can determine all others without any ambiguity. The value of the first coefficient is given by the normalization

condition

$$\sum_{\substack{m_1\\m_2}} |\langle\, j_1 j_2;\, m_1 m_2 \,|\, j_1 j_2; jm \,\rangle|^2 = 1$$

and by the sign convention and phase convention. One can have complex CG coefficients or, to be more precise, the solutions of the recursion relations can be complex. However, in this case the complex conjugate is also always a solution. The CG coefficients are therefore usually chosen to be real.

3.35 Selection rules for matrix elements

The triangular relation for the the CG coefficients implies that

$$\langle\, \alpha j' m' \,|\, T_q^{(k)} \,|\, \alpha\, jm \,\rangle = 0$$

unless

$$|\, j - k \,| \le j' \le j + k. \tag{3.63}$$

This can be interpreted as if $T_q^{(k)}$ would transfer an angular momentum $J = k$ with projection $J_z = q$ to the ket $|\, \alpha\, jm \,\rangle$! In the following, examples of applications of this relation are given.

Spin-orbit coupling $\hat{\vec{L}} \cdot \hat{\vec{S}}$

A tensor $T_0^{(0)}$ of rank 0 is a scalar. For example, the scalar product of two vector operators like $\hat{\vec{L}}$ and $\hat{\vec{S}}$ transforms like a scalar. In this case

$$\langle\, \alpha', j'm' \,|\, \vec{L} \cdot \vec{S} \,|\, \alpha\, jm \,\rangle = \delta_{jj'}\, \delta_{mm'}\, \frac{\langle\, \alpha' j \,\|\, \vec{L} \cdot \vec{S} \,\|\, \alpha j \,\rangle}{\sqrt{2j+1}}.$$

The matrix elements are diagonal in j and m, and independent of m. They can only depend on α and α', i.e., on other quantum numbers like n, L and S. For example, if $|\, \alpha\, j\, m \,\rangle = |\, L\, S\, J\, m \,\rangle$, we have

$$\begin{aligned}\langle\, L\, S\, J\, m \,|\, \vec{L} \cdot \vec{S} \,|\, L\, S\, J\, m \,\rangle &= \langle\, L\, S\, J\, m \,|\, \frac{J^2 - L^2 - S^2}{2} \,|\, L\, S\, J\, m \,\rangle \\ &= \frac{1}{2}\big[J(J+1) - L(L+1) - S(S+1)\big],\end{aligned}$$

which depends on J and $\alpha \equiv LS$. For a given LS multiplet (L and S fixed) the spin-orbit energy satisfies

$$E \propto \frac{1}{2}\, A\, J(J+1)$$

with A independent of J and therefore

$$\Delta E_{J,J-1} = \frac{1}{2}\, A\, \big[J(J+1) - (J-1)J\big] = A\, J.$$

This is known as the Landé interval rule (1923).

Matrix elements of vector operators:

Let us consider a vector operator $\vec{V}$ like the spin moment $\hat{\vec{S}}$, the orbital magnetic moment $\hat{\vec{L}}$, or the dipole matrix element of $\vec{x}$ in the theory of radiation. Given $\vec{V}$ as Cartesian vector, it can be written in terms of $V_{\pm1,0} = T^{(1)}_q$. The selection rules (3.63) imply that the matrix elements of vector operators vanish unless

$$\boxed{\begin{array}{l}\Delta m = m' - m = \pm 1, 0 \\ \Delta j = j' - j = \pm 1, 0\end{array}}$$

but

$$\boxed{\text{no } j = 0 \to j' = 0 \text{ transition!}}$$

The latter is a consequence of the triangular relation $|\, j - k \,| \le j' \le |\, j + k \,|$ which cannot be fulfilled for $k = 1$ and $j = j' = 0$. It would require $1 \le 0 \le 1$. In the case of vector operators, the Wigner-Eckart theorem takes a particularly intuitive and powerful form that is discussed in detail in the following section.

3.36 The projection theorem for vector operators

We consider $\langle\, \alpha' jm' \,|\, V_q \,|\, \alpha jm \,\rangle$, where V_q is an irreducible tensor of rank 1 and the matrix elements are taken between states having the same $j' = j$. Since $\vec{J} \cdot \vec{V} = J_0 V_0 - J_{+1} V_{-1} - J_{-1} V_{+1}$ and recalling that $J_{\pm1} = \mp J_\pm / \sqrt{2}$ and $J_0 = J_z$, we have, for the elements diagonal in m,

$$\begin{aligned}\langle\, \alpha' jm \,|\, \vec{J} \cdot \vec{V} \,|\, \alpha jm \,\rangle \;&=\; m\hbar \,\langle\, \alpha' jm \,|\, V_0 \,|\, \alpha jm \,\rangle + \\ &\quad + \frac{\hbar}{\sqrt{2}} \sqrt{(j+m)\,(j-m+1)}\; \langle\, \alpha' jm-1 \,|\, V_{-1} \,|\, \alpha jm \,\rangle \\ &\quad - \frac{\hbar}{\sqrt{2}} \sqrt{(j-m)\,(j+m+1)}\; \langle\, \alpha' jm+1 \,|\, V_{+1} \,|\, \alpha jm \,\rangle\,.\end{aligned}$$

By the Wigner–Eckart theorem for $j = j'$ we have

$$\langle\, \alpha' jm' \,|\, V_q \,|\, \alpha jm \,\rangle = \langle\, \alpha' j \,\|\, V \,\|\, \alpha j \,\rangle\, c(m, q, j),$$

where $c(m, q, j) = \langle\, j1; m\, q \,|\, j1; j\, m+q \,\rangle / \sqrt{2j+1}$ is independent of α, α' and V. We may then write

$$\langle\, \alpha' jm \,|\, \vec{J} \cdot \vec{V} \,|\, \alpha jm \,\rangle \;=\; c_{jm} \,\langle\, \alpha' j \,\|\, V \,\|\, \alpha j \,\rangle = c_j \,\langle\, \alpha' j \,\|\, V \,\|\, \alpha j \,\rangle\,.$$

The last step results from the fact that $\vec{J} \cdot \vec{V}$ is a scalar, and therefore c_{jm} cannot depend on m. Since c_j is independent of $\vec{V}$, we can apply the equality to $\vec{V} = \vec{J}$ and $\alpha = \alpha'$ to

get

$$\langle\, \alpha jm \,|\, \vec{J}^{\,2} \,|\, \alpha jm \,\rangle = \hbar^2 j\,(j+1) = c_j \,\langle\, \alpha j \,\|\, \vec{J} \,\|\, \alpha j \,\rangle\,.$$

We can now get rid of c_j. Thus,

$$\frac{\langle\, \alpha' jm \,|\, \vec{J}\cdot\vec{V} \,|\, \alpha jm \,\rangle}{\hbar^2 j\,(j+1)} = \frac{\langle\, \alpha' j \,\|\, \vec{V} \,\|\, \alpha j \,\rangle}{\langle\, \alpha j \,\|\, \vec{J} \,\|\, \alpha j \,\rangle}\,.$$

In order to compute $\langle\, \alpha' jm' \,|\, V_q \,|\, \alpha jm \,\rangle$ we return to the Wigner–Eckart theorem and observe that the ratio between matrix elements of operators of the same rank (and same component q) are independent of m, m', and q since this angular dependence is always given by the Clebsch–Gordon (Wigner) coefficients:

$$\underbrace{\frac{\langle\, \alpha' jm' \,|\, V_q \,|\, \alpha jm \,\rangle}{\langle\, \alpha jm' \,|\, J_q \,|\, \alpha jm \,\rangle}}_{\text{independent of } \alpha} \;=\; \frac{\langle\, \alpha' j \,\|\, \vec{V} \,\|\, \alpha j \,\rangle}{\langle\, \alpha j \,\|\, \vec{J} \,\|\, \alpha j \,\rangle} = \frac{\langle\, \alpha' jm \,|\, \vec{J}\cdot\vec{V} \,|\, \alpha jm \,\rangle}{\hbar^2 j\,(j+1)}\,.$$

Therefore,

$$\langle\, \alpha' jm' \,|\, V_q \,|\, \alpha jm \,\rangle = \overbrace{\frac{\langle\, \alpha' jm \,|\, \vec{J}\cdot\vec{V} \,|\, \alpha jm \,\rangle}{\hbar^2 j\,(j+1)}}^{\text{independent of } m \text{ and } q!} \langle\, jm' \,|\, J_q \,|\, jm \,\rangle\,,$$

where we have replaced $\langle\, \alpha jm' \,|\, J_q \,|\, \alpha jm \,\rangle = \langle\, jm' \,|\, J_q \,|\, jm \,\rangle$ since it is independent of α. Taking into account that the matrix elements $\langle\, \vec{J}\cdot\vec{V} \,\rangle$ are independent of q, we can return, if we wish, to the more familiar Cartesian coordinates and write in vector form

$$\boxed{\langle\, \alpha' jm' \,|\, \vec{V} \,|\, \alpha jm \,\rangle = \frac{\langle\, \alpha' jm \,|\, \vec{J}\cdot\vec{V} \,|\, \alpha jm \,\rangle}{\hbar^2\, j(j+1)} \langle\, jm' \,|\, \vec{J} \,|\, jm \,\rangle\,.}$$

This is the projection theorem which has many applications in radiation theory, magnetism, etc. It can be interpreted by saying that the averages or matrix elements of any vector operator $\langle\, \vec{V} \,\rangle$ within the same multiplet $|\, jm \,\rangle$ must be parallel to $\langle\, \vec{J} \rangle$. Notice that the relation $\langle\, \vec{V} \,\rangle \propto \langle\, \vec{J} \,\rangle$ holds only within the same multiplet jm.

Application to magnetic impurities in insulators

We intend to calculate the splitting of a multiplet $|\, J\,L\,S\,m \,\rangle$ of a magnetic impurity in an insulator by neglecting the effects of the crystal field, which is a pretty good approximation for rare earths. The magnetic field $\vec{B} = \vec{\nabla} \times \vec{A}$ enters in the Hamiltonian as in classical mechanics by modifying the kinetic energy

$$T = \frac{p^2}{2m} \longrightarrow \left(\vec{p} + \frac{e}{c}\,\vec{A}\right)^2 \quad \text{with} \quad e > 0$$

$$T = \frac{1}{2m}\,p^2 + \frac{e}{2mc}\,(\vec{p}\cdot\vec{A} + \vec{A}\cdot\vec{p}) + \frac{e^2}{2mc^2}\,A^2.$$

Since $\vec{p}\cdot\vec{A}-\vec{A}\cdot\vec{p}=-i\,\hbar\vec{\nabla}\cdot\vec{A}$ it is convenient to use the Coulomb gauge $\vec{\nabla}\cdot\vec{A}=0$ and rewrite T as

$$T=\frac{p^2}{2m}+\frac{e}{mc}\,(\vec{A}\cdot\vec{p})+\frac{e^2}{2mc^2}\,A^2.$$

For a uniform magnetic field we have $\vec{A}=\dfrac{1}{2}\,\vec{B}\times\vec{r}$ and thus

$$T=\frac{p^2}{2m}+\frac{e}{2mc}\,(\vec{B}\times\vec{r})\cdot\vec{p}+\frac{e^2}{2mc^2}\,\frac{(\vec{B}\times\vec{r})^2}{4}.$$

Using that $(\vec{a}\times\vec{b})\cdot\vec{c}=\vec{a}\cdot(\vec{b}\times\vec{c})$ we have

$$T=\frac{p^2}{2m}+\frac{e\,\hbar}{2mc}\,\vec{B}\cdot\vec{l}+\frac{e^2}{2mc^2}\,\frac{(\vec{B}\times\vec{r})^2}{4}$$

where $\hbar\vec{l}=\vec{r}\times\vec{p}$ is the angular momentum operator. If $\langle\,\vec{l}\,\rangle\neq 0$, the linear paramagnetic term dominates over the quadratic diamagnetic one. In addition we have the interaction of the electron's intrinsic spin with the magnetic field which is given by

$$\Delta H_{\rm spin}=g_0\,\mu_B\,\vec{S}\cdot\vec{B},$$

where $\mu_B=\dfrac{e\,\hbar}{2mc}=0.579\times10^{-8}\,\dfrac{eV}{G}$ is the Bohr magneton, and $g_0=2.0023\simeq 2$ is the electronic g factor. Finally, after summing over all the electrons of the magnetic impurity, the interaction with the magnetic field reads

$$\begin{aligned}\Delta H&=\mu_B\,\vec{B}\cdot\sum_i\,(\vec{l}_i+g_0\,\vec{s}_i)\\&=\mu_B\,\vec{B}\cdot(\vec{L}+g_0\,\vec{S}),\end{aligned}$$

where $\vec{L}$ and $\vec{S}$ are the total orbital and spin momentum operators.

In order to determine the Zeeman splitting of the energy levels at first-order perturbation theory we need to calculate

$$\langle\,J\,L\,S\,m\,|\,\vec{L}+g_0\,\vec{S}\,|\,J\,L\,S\,m\,\rangle.$$

Applying the projection theorem for the vector operator $\vec{L}+g_0\,\vec{S}$ we have

$$\begin{aligned}\langle\,J\,L\,S\,m\,|\,\vec{L}\;\;&+\;\;g_0\,\vec{S}\,|\,J\,L\,S\,m\,\rangle=\\&=\;\;\frac{\langle\,J\,L\,S\,m\,|\,(\vec{L}+g_0\,\vec{S})\cdot\vec{J}\,|\,J\,L\,S\,m\,\rangle}{\hbar^2\,(J+1)J}\,\langle\,J\,L\,S\,m\,|\,\vec{J}\,|\,J\,L\,S\,m\,\rangle\\&=\;\;g(J\,L\,S)\,\langle\,J\,L\,S\,m\,|\,\vec{J}\,|\,J\,L\,S\,m\,\rangle,\end{aligned}$$

where the gyromagnetic Landé factor $g(J\,L\,S)$ is independent of m since it is a scalar. The energy levels are then split in the magnetic field $\vec{H}=H\,J_z$ as

$$E(J\,L\,S\,m)=\mu_B\,H\,g(J\,L\,S)\,m.$$

This leads to the well-known Curie law for magnetic impurities in insulators. We can compute $g(J\,L\,S)$ as follows

$$L^2 = (\vec{J} - \vec{S})^2 = J^2 + S^2 - 2\vec{J}\cdot\vec{S}$$

$$S^2 = (\vec{J} - \vec{L})^2 = J^2 + L^2 - 2\vec{J}\cdot\vec{L}.$$

$$g(J\,L\,S) = \frac{\langle\, J\,L\,S\,m \,|\, (J^2 + L^2 - S^2) + g_0\,(J^2 + S^2 - L^2) \,|\, J\,L\,S\,m\,\rangle}{2\hbar^2\, J(j+1)}$$

$$= \frac{g_0 + 1}{2} + \frac{(g_0 - 1)}{2}\,\frac{[S(S+1) - L(L+1)]}{J(J+1)}.$$

Approximating g_0 by $g_0 = 2$ this reduces to

$$g(J\,L\,S) = \frac{3}{2} + \frac{1}{2}\,\frac{[S(S+1) - L(L+1)]}{J(J+1)}.$$

3.37 Proof of the Wigner–Eckart theorem

The proof of the theorem is instructive and rather simple. Let us recall the recursion relation for the Clebsch–Gordan coefficients:

$$\begin{aligned}\langle\, j_1 j_2; jm \,|\, J_\pm \,|\, j_1 j_2; m_1 m_2\,\rangle &= \sqrt{(j \pm m)\,(j \mp m + 1)}\;\langle\, j_1 j_2; j\; m\mp 1 \,|\, j_1 j_2; m_1 m_2\,\rangle \\ &= \sqrt{(j_1 \mp m_1)\,(j_1 \pm m_1 + 1)}\;\langle\, j_1 j_2; jm \,|\, j_1 j_2; m_1 \pm 1\; m_2\,\rangle + \\ &\quad + \sqrt{(j_2 \mp m_2)\,(j_2 \pm m_2 + 1)}\;\langle\, j_1 j_2; jm \,|\, j_1 j_2; m_1\; m_2 \pm 1\,\rangle\,.\end{aligned}$$

This relation can easily be proven by using $J_\pm |\, jm\,\rangle = \sqrt{(j \mp m)(j \pm m + 1)}\;|\, j\; m\pm 1\,\rangle$.

In view of the discussion of the Wigner–Eckart theorem ($j + k \to j'$) on should consider the following change of variables:

$$\begin{aligned} j_1 &\to j \qquad & m_1 &\to m\,, \\ j_2 &\to k \qquad & m_2 &\to q\,, \\ j &\to j' \qquad & m &\to m'\,. \end{aligned}$$

One then obtains

$$\begin{aligned}\sqrt{(j' \pm m')\,(j' \mp m' + 1)}\;\langle\, jk; j'm' \mp 1 \,|\, jk; mq\,\rangle &= \\ &= \sqrt{(j \mp m)\,(j \pm m + 1)}\;\langle\, jk; j'm' \,|\, jk; m \pm 1\; q\,\rangle + \\ &\quad \sqrt{(k \mp q)\,(k \pm q + 1)}\;\langle\, jk; j'm' \,|\, jk; m\; q \pm 1\,\rangle\,.\end{aligned}$$

This is the recursion relation for adding j and k to obtain j'.

Now we can use $[J_\pm, T^{(k)}_q] = \sqrt{(k \mp q)\,(k \pm q + 1)}\; T^{(k)}_{q\pm1}$, not without observing the analogy with $J_\pm \,|\, jm \,\rangle = \sqrt{(j \mp m)\,(j \pm m + 1)}\; |\, jm \pm 1 \,\rangle$. We then have

$$J_\pm \, T^{(k)}_q = \sqrt{(k \mp q)\,(k \pm q + 1)}\; T^{(k)}_{q\pm1} + T^{(k)}_q \, J_\pm \,.$$

This is as if one would raise the component q of the tensor $T^{(k)}_q$. Everything is pretty much like $J_\pm = J_{1\pm} + J_{2\pm}$ in the Clebsch–Gordan recursion. Taking matrix elements, we have

$$\begin{aligned}
\langle\, \alpha' j' m' \,|\, J_\pm \, T^{(k)}_q \,|\, \alpha j m \,\rangle \;&=\; \sqrt{(j' \pm m')\,(j' \mp m' + 1)}\; \langle\, \alpha' j' m' \mp 1 \,|\, T^{(k)}_q \,|\, \alpha j m \,\rangle \\
&=\; \sqrt{(k \mp q)\,(k \pm q + 1)}\; \langle\, \alpha' j' m' \,|\, T^{(k)}_{q\pm1} \,|\, \alpha j m \,\rangle \,+ \\
&\quad + \sqrt{(j \mp m)\,(j \pm m + 1)}\; \langle\, \alpha' j' m' \,|\, T^{(k)}_q \,|\, \alpha j m \pm 1 \,\rangle \,,
\end{aligned}$$

which are precisely the same recursion relations as for the Clebsch–Gordan coefficients. Since the solution to the recursion is *unique up to a constant*, we have

$$\langle\, \alpha' j' m' \,|\, T^{(k)}_q \,|\, \alpha j m \,\rangle = \langle\, \alpha' j' \,\|\, T^{(k)} \,\|\, \alpha j \,\rangle \, \langle\, j k; j' m' \,|\, j k; m q \,\rangle \,.$$

The "constant" is independent of the variables m, q, and m' corresponding to m_1, m_2, and m in the recursion relations (3.62) for the CG coefficients, but depends on j, k, and j', which correspond to j_1, j_2, and j. This constant is by definition $\langle\, \alpha' j' \,\|\, T^{(k)} \,\|\, \alpha j \,\rangle$ up to some normalization factor like $(1/\sqrt{2j+1})$.

3.38 Discrete Symmetries: Parity or space inversion

The discrete transformations of main interest are parity or space inversion, lattice translations as found in periodic solids, and time inversion symmetry. In the following we discuss parity transformations.

We consider a space inverted state $|\, \alpha' \,\rangle$ that is obtained from the original state $|\, \alpha \,\rangle$ by applying the *parity operator* $\hat{\Pi}$:[8]

$$|\, \alpha' \,\rangle = \hat{\Pi} \,|\, \alpha \,\rangle.$$

The O(3) matrix associated to space inversion is

$$R^{(\text{parity})} = -\mathbb{1} = \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix},$$

which commutes with all rotations of SO(3). In accordance with our usual notation $\hat{D}(R)$ for the representation of SO(3), now extended to O(3), we can write $\hat{D}(-\mathbb{1}) = \hat{\Pi}$.

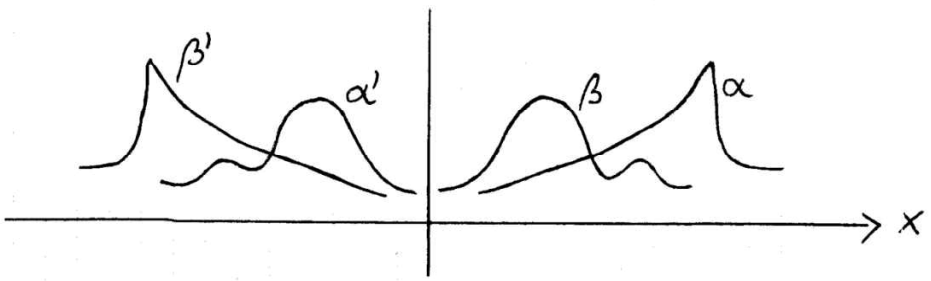
[8] For the sake of clarity a hat "$\hat{}$" is used to distinguish operators from numbers.

In order to define the properties of $\hat{\Pi}$ we require, first of all, that

$$\langle\, \alpha' \,|\, \hat{x} \,|\, \beta' \,\rangle = \langle\, \alpha \,|\, \hat{\Pi}^\dagger \, \hat{x} \, \hat{\Pi} \,|\, \beta \,\rangle = -\langle\, \alpha \,|\, \hat{x} \,|\, \beta \,\rangle$$

for all $|\, \alpha \,\rangle$ and $|\, \beta \,\rangle$ (see Figure) and in particular

$$\langle\, \alpha \,|\, \hat{\Pi}^\dagger \, \hat{x} \, \hat{\Pi} \,|\, \alpha \,\rangle = -\langle\, \alpha \,|\, \hat{x} \,|\, \alpha \,\rangle.$$



This implies

$$\hat{\Pi}^\dagger \, \hat{x} \, \hat{\Pi} = -\hat{x}$$

and using that $\hat{\Pi}$ is unitary,[9] i.e., $\hat{\Pi}^\dagger = \hat{\Pi}^{-1}$, we have

$$\hat{\Pi}^{-1} \, \hat{x} \, \hat{\Pi} = -\hat{x}$$

or

$$\boxed{\hat{x} \, \hat{\Pi} = -\hat{\Pi} \, \hat{x}.}$$

As expected, the position and parity operators $\hat{x}$ and $\hat{\Pi}$ anticommute.

It is easy to see that $\Pi \,|\, x' \,\rangle = e^{i\,\delta} \,|\, -x' \,\rangle$, where $|\, x' \,\rangle$ is the eigenket of the position operator with eigenvalue x' and $\delta \in \mathbb{R}$. In fact,

$$\hat{x} \, \hat{\Pi} \,|\, x' \,\rangle = -\hat{\Pi} \, \hat{x} \,|\, x' \,\rangle = -x' \, \hat{\Pi} \,|\, x' \,\rangle.$$

Since the eigenstates $|\, x \,\rangle$ are nondegenerate, we must have

$$\hat{\Pi} \,|\, x' \,\rangle = e^{i\,\delta} \,|\, -x' \,\rangle.$$

It is convenient to choose the convention $e^{i\,\delta} = 1$, in which case

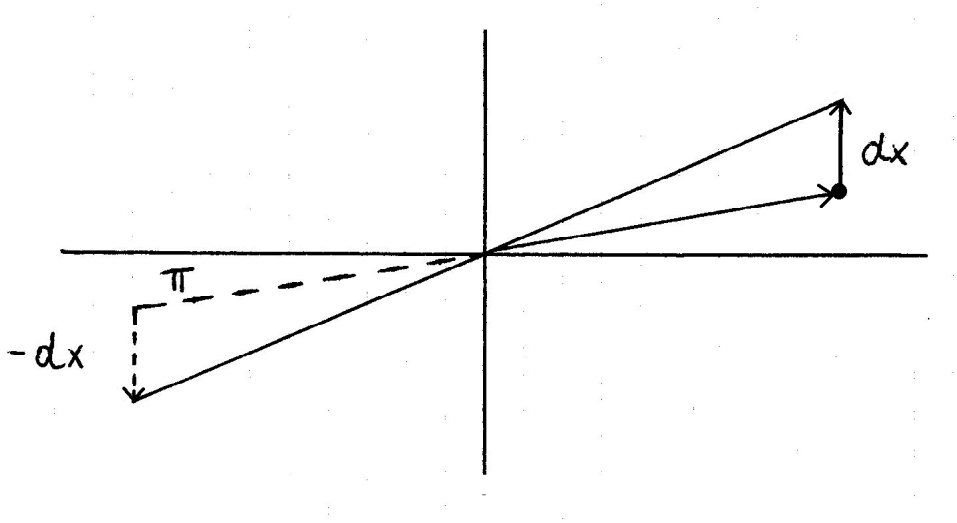
$$\hat{\Pi}^2 \,|\, x' \,\rangle = \hat{\Pi} \,|\, -x' \,\rangle = |\, x' \,\rangle. \tag{3.64}$$

Thus, $\hat{\Pi}^2 = 1$, which implies $\hat{\Pi}^\dagger = \hat{\Pi}^{-1} = \hat{\Pi}$. One concludes that $\hat{\Pi}$ is both hermitic and unitary.[5]

[9] According to Wigner's theorem (Secs. 3.3 and 3.4) a discrete symmetry operation like $\hat{\Pi}$ can be unitary or antiunitary. The reason for $\hat{\Pi}$ being unitary relies on its behavior with respect to spatial translations as discussed in Sec. 3.39.

3.39 Parity and momentum operator

In order to infer the commutation relation between $\hat{\Pi}$ and the momentum operator $\hat{\vec{p}}$, we must consider the relation between $\hat{\Pi}$ and the translation operator $\hat{\mathcal{T}}$. From the picture it is clear that the translation $d\vec{x}$ followed by space inversion is equivalent to space inversion followed by the translation $-d\vec{x}$.



This implies

$$\hat{\Pi}\,\hat{\mathcal{T}}(d\vec{x}) = \hat{\mathcal{T}}(-d\vec{x})\,\hat{\Pi}.$$

Recalling that $\hat{\mathcal{T}}(d\vec{x}) = 1 - \dfrac{i}{\hbar}\,\hat{\vec{p}}\cdot d\vec{x}$ we have

$$\hat{\Pi}\left(1 - \frac{i}{\hbar}\,\hat{\vec{p}}\cdot d\vec{x}\right) = \left(1 + \frac{i}{\hbar}\,\hat{\vec{p}}\cdot d\vec{x}\right)\hat{\Pi}$$

or equivalently

$$\boxed{\hat{\Pi}\,\hat{\vec{p}} = -\hat{\vec{p}}\,\hat{\Pi}.}$$

One concludes that $\hat{\Pi}$ and $\hat{\vec{p}}$ anticommute. Notice that if $\hat{\Pi}$ would be antiunitary, one would have $[\hat{\vec{p}}, \hat{\Pi}] = 0$ and $\hat{\Pi}\,|\,\vec{p}\,\rangle = e^{i\,\delta}\,|\,\vec{p}\,\rangle$ in contrast to the classical notion of inversion of velocity and momentum after space inversion. One concludes that $\hat{\Pi}$ is unitary as assumed in the previous section.

3.40 Parity and angular momentum

The O(3) matrix of the parity transformation $R^{(\text{parity})} = -\mathbb{1}$ commutes with all SO(3) rotation matrices R. Thus the representation or group property of O(3) implies

$$\hat{D}(-\mathbb{1}\,R) = \hat{D}(-\mathbb{1})\,\hat{D}(R) = \hat{\Pi}\,\hat{D}(R)$$

and

$$\hat{D}[R\,(-\mathbb{1})] = \hat{D}(R)\,\hat{D}(-\mathbb{1}) = \hat{D}(R)\,\hat{\Pi}.$$

Therefore

$$\hat{D}(R)\,\hat{\Pi} = \hat{\Pi}\,\hat{D}(R)$$

for all rotations R. Since $\hat{D}(R) = 1 - \dfrac{i}{\hbar}\,\hat{\vec{J}}\cdot\hat{n}\,\delta\phi$ for an infinitesimal rotation $\delta\phi$ around the axis $\hat{n}$ we have

$$[\hat{\Pi}, \hat{\vec{J}}] = 0.$$

This holds for $\hat{\vec{J}} = \hat{\vec{L}},\ \hat{\vec{J}} = \hat{\vec{S}}$ and $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$. In the case of the orbital momentum $\hat{\vec{L}}$,

$$[\hat{\Pi}, \hat{\vec{L}}] = [\hat{\Pi}, \hat{\vec{x}} \times \hat{\vec{p}}] = 0$$

also follows from the anticommutations $\{\hat{\Pi}, \hat{\vec{x}}\} = \{\hat{\Pi}, \hat{\vec{p}}\} = 0$.

In quantum mechanics, as in classical mechanics, $\vec{x}$, $\vec{p}$ and $\vec{J}$ behave like vectors under rotations. One distinguishes between *polar vectors* like $\hat{\vec{x}}$ and $\vec{p}$, which are *odd under parity*, and *axial vectors or pseudovectors* like $\vec{J}$, which are *even under parity*.

3.41 The wave function and parity transformations

The transformation property of the wave function $\Psi_\alpha(\vec{x}) = \langle\,\vec{x}\,|\,\alpha\,\rangle$ can be readily obtained from $\{\hat{\Pi}, \hat{\vec{x}}\} = 0$. Denoting as usual the transformed state by $|\,\alpha'\,\rangle = \hat{\Pi}\,|\,\alpha\,\rangle$ we have

$$\Psi_{\alpha'}(\vec{x}) = \langle\,\vec{x}\,|\,\alpha'\,\rangle = \langle\,\vec{x}\,|\,\hat{\Pi}\,\alpha\,\rangle = \langle\,\hat{\Pi}\,\vec{x}\,|\,\alpha\,\rangle = \langle\,-\vec{x}\,|\,\alpha\,\rangle = \Psi_\alpha(-\vec{x}),$$

where we have used that $\hat{\Pi}|\,x\,\rangle = |\,-x\,\rangle$ from Eq. (3.64). One concludes that

$$\Psi_{\alpha'}(\vec{x}) = \Psi_\alpha(-\vec{x}).$$

The eigenstates of the parity operator

$$\hat{\Pi}\,|\,\alpha\,\rangle = \pm|\,\alpha\,\rangle$$

are even $(+)$ or odd $(-)$ if and only if the corresponding wave function $\Psi_\alpha(\vec{x})$ satisfies

$$\Psi_\alpha(-\vec{x}) = \pm\Psi_\alpha(\vec{x}).$$

The eigenstates of the orbital angular momentum $\hat{\vec{L}}$ can certainly be chosen to be parity eigenstates, i.e., to have defined parity, since $[\hat{\vec{L}}, \hat{\Pi}] = 0$. An eigenstate of $\hat{L}^2$ and $\hat{L}_z$ has a wave function of the form

$$\Psi_{\alpha lm}(\vec{x}) = \langle\,\vec{x}\,|\,\alpha\,l\,m\,\rangle = R_\alpha(r)\,Y_{lm}(\theta,\varphi).$$

The transformation $\vec{x} \to -\vec{x}$ implies

$$\begin{aligned} & r \to r, \\ & \theta \to \pi - \theta \quad \text{and} \quad \cos\theta \to -\cos\theta, \\ & \varphi \to \varphi + \pi \quad \text{and} \quad e^{i\,m\,\varphi} \to (-1)^m\, e^{i\,m\,\varphi}. \end{aligned}$$

To see how $|\,\alpha\,l\,m\,\rangle$ behaves under parity transformation we consider first the case $m = 0$ and recall that

$$Y_{l0}(\theta, \varphi) = P_l(\cos\theta),$$

where $P_l(x)$ is the Legendre polynomial of order l, that involves only powers of $\cos\theta$ with the same parity as l. Consequently,

$$Y_{l0}(\pi - \theta, \varphi + \pi) = P_l(-\cos\theta) = (-1)^l\, P_l(\cos\theta) = (-1)^l\, Y_{l0}(\theta, \varphi).$$

Thus, $\hat{\Pi}\,|\,\alpha\,l\,0\,\rangle = (-1)^l\,|\,\alpha\,l\,0\,\rangle$. Moreover, $[\hat{\vec{L}}, \hat{\Pi}] = 0 \;\Rightarrow\; [\hat{L}_\pm, \hat{\Pi}] = 0$ and therefore

$$\hat{\Pi}\,|\,\alpha\,l\,m\,\rangle = (-1)^l\,|\,\alpha\,l\,m\,\rangle \quad \forall\, m.$$

3.42 Parity invariant Hamiltonians

Consider a parity invariant Hamiltonian

$$\hat{H} = \hat{\Pi}^\dagger \hat{H}\, \hat{\Pi}$$

or equivalently

$$[\hat{H}, \hat{\Pi}] = 0.$$

It is easy to see that the eigenstates of $\hat{H}$ can be chosen to have defined parity. In fact the projector operators

$$\hat{P}_\pm = \frac{1}{2}\,(1 \pm \hat{\Pi})$$

satisfy $\hat{P}_\pm^2 = \hat{P}_\pm$ and $1 = \hat{P}_+ + \hat{P}_-$. Consequently, for any state $|\,\alpha\,\rangle$ we have

$$|\,\alpha\,\rangle = \hat{P}_+\,|\,\alpha\,\rangle + \hat{P}_-\,|\,\alpha\,\rangle.$$

Moreover,

$$\hat{\Pi}\,\hat{P}_\pm = \frac{1}{2}\,(\hat{\Pi} \pm \hat{\Pi}^2) = \frac{1}{2}\,(\hat{\Pi} \pm 1) = \pm\hat{P}_\pm$$

and therefore $\hat{\Pi}\,\hat{P}_\pm\,|\,\alpha\,\rangle = \pm\hat{P}_\pm\,|\,\alpha\,\rangle$. The projected state $\hat{P}_\pm\,|\,\alpha\,\rangle$ is either zero or proportional to an eigenstate of $\hat{\Pi}$ with eigenvalue ± 1. Taking into account that

$$[\hat{\Pi}, \hat{H}] = 0 \;\Rightarrow\; [\hat{P}_\pm, \hat{H}] = 0,$$

we have $\hat{H}\,\hat{P}_\pm\,|\,\alpha\,\rangle = \hat{P}_\pm\,\hat{H}\,|\,\alpha\,\rangle = E_\alpha\,\hat{P}_\pm\,|\,\alpha\,\rangle$. In other words, if $|\,\alpha\,\rangle$ is an eigenstate of $\hat{H}$ with energy E_α, then the projected state $\hat{P}_\pm\,|\,\alpha\,\rangle$ is either zero or an eigenstate of $\hat{H}$ with the same energy.

In conclusion, the eigenstates of $\hat{H}$ can be chosen to have defined parity if $[\hat{H},\hat{\Pi}] = 0$. If there is no degeneracy the eigenstates of $\hat{H}$ must have defined parity, otherwise not necessarily as in the case of plane waves

$$e^{i\,\vec{p}\cdot\vec{x}/\hbar} = \underbrace{\cos\left(\frac{\vec{p}\cdot\vec{x}}{\hbar}\right)}_{\text{even}} + \underbrace{i\,\sin\left(\frac{\vec{p}\cdot\vec{x}}{\hbar}\right)}_{\text{odd}}.$$

An example of energy eigenstates with defined parity is the harmonic oscillator

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\,\hat{x}^2.$$

The annihilation operator

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\,\left(\hat{x} + \frac{i\,\hat{p}}{m\omega}\right)$$

is linear in $\hat{x}$ and $\hat{p}$ and therefore

$$\{\hat{\Pi},\hat{a}\} = \{\hat{\Pi},\hat{a}^\dagger\} = 0.$$

The ground state $|\,0\,\rangle$ is non-degenerate and has even parity since its coordinate representation is a Gaussian. Taking into account that

$$\hat{\Pi}(\hat{a}^\dagger)^n = (-1)^n\,(\hat{a}^\dagger)^n\,\hat{\Pi},$$

the nth excited state

$$|\,n\,\rangle = (\hat{a}^\dagger)^n\,|\,0\,\rangle$$

has parity $(-1)^n$. This is a particular example of a symmetric potential in one dimension.

3.43 Parity selection rule

Consider two states $|\,\alpha\,\rangle$ and $|\,\beta\,\rangle$ with defined parities $\varepsilon_\alpha = \pm 1$ and $\varepsilon_\beta = \pm 1$:

$$\hat{\Pi}\,|\,\alpha\,\rangle = \varepsilon_\alpha\,|\,\alpha\,\rangle$$

$$\hat{\Pi}\,|\,\beta\,\rangle = \varepsilon_\beta\,|\,\beta\,\rangle.$$

Consider an operator $\hat{A}$ that is even under parity (e.g., $\hat{\vec{L}}$), i.e.,

$$\hat{\Pi}^\dagger\,\hat{A}\,\hat{\Pi} = \hat{A} \;\Leftrightarrow\; [\hat{\Pi},\hat{A}] = 0.$$

Then the following selection rule holds

$$\langle\, \alpha \,|\, \hat{A} \,|\, \beta \,\rangle = 0 \;\; \text{unless} \;\; \varepsilon_\alpha = \varepsilon_\beta .$$

Operators that are even under parity can have non-vanishing matrix elements only between states of the same parity. In fact,

$$\langle\, \alpha \,|\, \hat{A} \,|\, \beta \,\rangle = \langle\, \alpha \,|\, \hat{\Pi}^\dagger \, \hat{\Pi} \, \hat{A} \, \hat{\Pi}^\dagger \, \hat{\Pi} \,|\, \beta \,\rangle = \varepsilon_\alpha \, \varepsilon_\beta \, \langle\, \alpha \,|\, \hat{A} \,|\, \beta \,\rangle .$$

Therefore, $\varepsilon_\alpha \neq \varepsilon_\beta \;\Rightarrow\; \langle\, \alpha \,|\, \hat{A} \,|\, \beta \,\rangle = 0$ ($\hat{A}$ even).

In a similar way one can show that if $\hat{A}$ has odd parity, e.g., $\hat{\vec{x}}$, $\hat{\vec{p}}$, it has non-vanishing matrix elements only between states with different parity. Indeed, odd parity means $\hat{\Pi} \, \hat{A} = -\hat{A} \, \hat{\Pi}$, which implies

$$\langle\, \alpha \,|\, A \,|\, \beta \,\rangle = \langle\, \alpha \,|\, \hat{\Pi}^\dagger \, \hat{\Pi} \, \hat{A} \, \hat{\Pi}^\dagger \, \hat{\Pi} \,|\, \beta \,\rangle = \varepsilon_\alpha \, \varepsilon_\beta \, \langle\, \alpha \,|\, \Pi \; A \; \Pi^+ \,|\, \beta \,\rangle = -\varepsilon_\alpha \, \varepsilon_\beta \, \langle\, \alpha \,|\, A \,|\, \beta \,\rangle .$$

In particular, optical transitions connect states with different parity since the relevant matrix elements are $\langle\, n \,|\, \hat{\vec{p}} \,|\, i \,\rangle$ or $\langle\, n \,|\, \hat{\vec{x}} \,|\, i \,\rangle$. This is known as the *Laporte rule*, that was phenomenologically known from atomic optical spectra even before quantum mechanics was formulated.

Another interesting consequence of the parity selection rule is that the dipole moment $\langle\, \alpha \,|\, \hat{\vec{x}} \,|\, \alpha \,\rangle$ and average momentum $\langle\, \alpha \,|\, \hat{\vec{p}} \,|\, \alpha \,\rangle$ of parity eigenstates always vanish. Let $\hat{\Pi} \,|\, \alpha \,\rangle = \varepsilon_\alpha \,|\, \alpha \,\rangle$ with $\varepsilon_\alpha = \pm 1$. Since $\hat{\vec{x}} \, \hat{\Pi} + \hat{\Pi} \, \hat{\vec{x}} = 0$ and $\hat{\vec{p}} \, \hat{\Pi} + \hat{\Pi} \, \hat{\vec{p}} = 0$, we have

$$0 = \langle\, \alpha \,|\, \hat{\vec{x}} \, \hat{\Pi} + \hat{\Pi} \, \hat{\vec{x}} \,|\, \alpha \,\rangle = 2 \, \varepsilon_\alpha \langle\, \alpha \,|\, \hat{\vec{x}} \,|\, \alpha \,\rangle \;\Rightarrow\; \langle\, \alpha \,|\, \hat{\vec{x}} \,|\, \alpha \,\rangle = 0$$

and similarly

$$\langle\, \alpha \,|\, \hat{\vec{p}} \,|\, \alpha \,\rangle = 0.$$

In particular non-degenerate eigenstates of parity invariant Hamiltonians always have vanishing dipole moment $\langle\, \alpha \,|\, \hat{\vec{x}} \,|\, \alpha \,\rangle = 0$ and average momentum $\langle\, \alpha \,|\, \hat{\vec{p}} \,|\, \alpha \,\rangle = 0$.

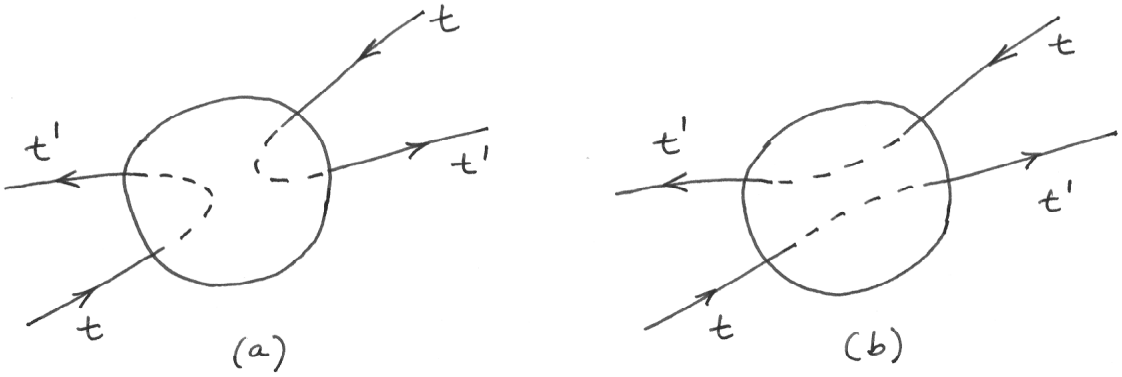
4 Identity of particles

The physical world and in particular atoms, molecules and solids are made of many interacting particles. In principle the N body wave function $\Psi(x_1 \dots x_N)$ and the Schrödinger equation $i\hbar\frac{\partial\Psi}{\partial t} = H\Psi$ contain all the information for describing the physical states of a given system of N identical particles and their time evolution. However, neither the Schrödinger equation is easy to solve in general nor the wave function $\Psi(x_1 \dots x_N)$ appears to be the most practical representation of many-particle states. The representation in terms of *occupation numbers* usually known as *second quantization* provides a much more appropriate framework. Most of the theoretical developments on the many-body properties of matter, in particular concerning magnetism and nanostructures, rely on this formalism. An alternative, very important approach, which circumvents the explicit use of the wave function, is density functional theory which shall be discussed later on.

Second quantization is a formalism in which the occupation numbers of an arbitrary complete set of single-particle states play the role of independent variables, instead of the coordinates x_i of the individual particles as in the usual wave function $\Psi(x_1 \dots x_N)$. The transitions between different many-particle states can be visualized as changes in the occupations of simpler single-particle orbitals. This is particularly useful in order to formulate, visualize and understand the physics of models of the electronic and magnetic properties of matter. For instance, the relevant orbitals responsible for magnetism (the d orbitals in the case of transition metals) can be focused by an appropriate choice of the single-particle basis. The most important interactions and energy scales can then be readily identified. In the following we shall recall the most important steps in the formalism of second quantization for fermions and bosons. These notes are based on the books by Landau & Lifshitz and Fetter & Wallecka to which the reader should refer for further details [3, 4].

4.1 The principle of indistinguishability of identical particles

In classical physics the particles preserve their individuality despite having the same physical properties. A "numbering" of the particles is possible with which one can follow the trajectory of each particle individually along its path. This applies, whatever the number of particles is, and in particular for particles that are exactly identical in all respects. In quantum mechanics the situation is entirely different since the notion of deterministic path ceases to have any meaning as a consequence of Heisenberg's uncertainty principle. Even if a numbering of the particles were possible at some time, for example if we measure the position of each particle in the system at time t, there is no possibility of tracking the positions of the particles at any future (or past) time $t' > t$ ($t' < t$), since the coordinates have no definite values even at times arbitrarily close to t. If we then localize (or measure) an electron at given instant $t' > t$, it is impossible to say which electron (among the N previously localized ones) has arrived at this point. The lack of a single deterministic path can be illustrated in the following scattering picture [5]:



In quantum mechanics identical particles entirely lose their individuality and become completely indistinguishable. *No experimental measurement can ever remove this indistinguishability.* This is the principle of indistinguishability of identical particles, which has many far-reaching consequences.

Consider two observers O and O' who prepare two physically identical quantum mechanical states but adopt two different conventions for labeling the electronic coordinates. For example, O and O' measure the position of the particles at x_1 and x_2 or they prepare the scattering of two identical wave packets. Let $|\,\Psi\,\rangle$ be the state considered by O and $\Psi(x_1, x_2) = \langle\, x_1, x_2\,|\,\Psi\,\rangle$ the coordinate wave function. And let $|\,\Psi'\,\rangle$ be the state considered by O' with the coordinate wave function $\Psi'(x_1, x_2) = \langle\, x_1, x_2\,|\,\Psi'\,\rangle$ with $\Psi'(x_1, x_2) = \Psi(x_2, x_1)$. The principle of indistinguishability of identical particles states that $|\,\Psi\,\rangle$ and $|\,\Psi'\,\rangle$ are equivalent representations of the same physical states with completely equivalent physical properties. Therefore, for any state $|\,\beta\,\rangle$, the probability of finding $|\,\Psi\,\rangle$ or $|\,\Psi'\,\rangle$ in $|\,\beta\,\rangle$ must be the same. This means that

$$|\langle\,\beta\,|\,\Psi\,\rangle|^2 = |\langle\,\beta\,|\,\Psi'\,\rangle|^2 \quad \forall\,|\,\beta\,\rangle.$$

In particular for $|\,\beta\,\rangle = |\,\Psi\,\rangle$ we have

$$|\langle\,\Psi\,|\,\Psi\,\rangle|^2 = 1 = |\langle\,\Psi\,|\,\Psi'\,\rangle|^2.$$

Consequently, taking into account that $\langle\,\Psi'\,|\,\Psi'\,\rangle = 1$, we must have

$$|\,\Psi'\,\rangle = e^{i\alpha}|\,\Psi\,\rangle.$$

Then

$$\langle\, x_1, x_2\,|\,\Psi'\,\rangle = e^{i\alpha}\langle\, x_1, x_2\,|\,\Psi\,\rangle$$

or

$$\Psi'(x_1, x_2) = e^{i\alpha}\,\Psi(x_1, x_2),$$

and thus

$$\Psi(x_2, x_1) = e^{i\alpha}\, \Psi(x_1, x_2).$$

By repeating the interchange we have $\Psi(x_1, x_2) = e^{i\alpha}\, \Psi(x_2, x_1) = e^{2i\alpha}\, \Psi(x_1, x_2)$ which implies that $e^{2i\alpha} = 1 \quad \Rightarrow \quad e^{i\alpha} = \pm 1$.

The previous arguments can immediately be generalized to any two particles x_i, x_j in an N-particle system. Therefore,

$$\Psi(x_1, \ldots x_i \ldots x_j \ldots x_N) = \pm \Psi(x_1, \ldots x_j \ldots x_i \ldots x_N)$$

for all i and j. Since the particles are indistinguishable, the same sign holds for any two particles in the system. The wave function $\Psi(x_1, \ldots x_N)$ is either fully symmetrical with respect to interchange of variables or fully antisymmetrical. The superposition of states with different symmetry is not possible since the resulting wave function would neither be symmetrical nor antisymmetrical.

The particles in nature are thus divided in two disjoint groups. The particles having symmetrical wave functions are called *bosons* and are said to obey Bose-Einstein statistics. The particles with antisymmetrical wave functions are called *fermions* and are said to obey Fermi-Dirac statistics. The property of being a boson or a fermion is of course a fundamental property that depends on the nature of the particle. Experiment shows that there is a one-to-one correspondence between the fermionic or bosonic character and the intrinsic spin of the particles: bosons are particles with integer spin, while fermions are particles with half-integer spin. Most elementary particles are fermions (e^-, e^+, p, n). However, photons and a number of elementary excitations in condensed matter (phonons, magnons, etc.) are bosons.

In the case of complex particles (e.g., an α particle) the interchange of two particles can be regarded as the simultaneous interchange of its constituents. Therefore, the statistics of complex particles is fermionic if the number of elementary fermions is odd, or bosonic if the number of fermions is even. Thus, an α particle composed by two protons and two neutrons is a boson. ^{3}He atoms are fermions while ^{4}He atoms are bosons. This has crucial consequences on the low-temperature properties of these two isotopes. Notice that the integer/half-integer rule holds also for complex particles since an even (odd) number of half-integer elementary particles corresponds to a total spin which is integer (half-integer).

To conclude this section, let us consider the effect of an arbitrary permutation on the coordinates. Let $P : [1, N] \to [1, N]$ be a permutation in the natural interval $[1, N]$ which we denote by

$$P = \begin{pmatrix} 1 & 2 & \ldots & N \\ P(1) & P(2) & \ldots & P(N) \end{pmatrix}$$

or simply $P = [P(1), P(2), \ldots P(N)]$. The *order of the permutation* is the smallest integer m such that P^m is equal to the identity. The *sign of the permutation* P is given by

$\mathrm{sgn}(P) = (-1)^p$ where p is the number of transpositions needed to realize P, or equivalently, the number of transpositions required to bring the sequence $[P(1), P(2), \ldots P(N)]$ into the normal ordering $[1, 2, \ldots N]$. For example,

$$P = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \text{ has } p = 1, \quad \text{while} \quad P = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \text{ has } p = 2\ .$$

Although the decomposition of a permutation P in transpositions and the number of the latter are not unique, the parity of p is always the same. Thus, $\mathrm{sgn}(P)$ is well defined.

It is then easy to see that for bosons

$$\Psi\left(x_{P(1)}, x_{P(2)}, \ldots x_{P(N)}\right) = \Psi(x_1, \ldots x_N),$$

while for fermions

$$\Psi\left(x_{P(1)}, x_{P(2)}, \ldots x_{P(N)}\right) = (-1)^p\ \Psi(x_1, \ldots x_N).$$

A note on permutations:

A permutation P is a bijective mapping of the natural interval $[1, N]$ onto itself, i.e., $P\colon [1, N] \to [1, N]$ bijective. We may denote it by

$$P = \begin{pmatrix} 1 & 2 & ... & N \\ P(1) & P(2) & ... & P(N) \end{pmatrix},$$

where the order of the columns is irrelevant and the inverse P^{-1} is obtained by transposing the 2 rows. For example,

$$P = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}$$

and

$$P^{-1} = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

The set of all permutations in $[1, N]$ forms a finite group of order $N!$, which is known as the *symmetric group*. It is easy to verify that all group-defining properties are satisfied:

(1) Closure: The composition PQ of any two permutations

$$P = \begin{pmatrix} 1 & 2 & ... & N \\ p_1 & p_2 & ... & p_N \end{pmatrix}$$

and

$$Q = \begin{pmatrix} q_1 & q_2 & ... & q_N \\ 1 & 2 & ... & N \end{pmatrix}$$

given by

$$PQ = \begin{pmatrix} q_1 & q_2 & ... & q_N \\ p_1 & p_2 & ... & p_N \end{pmatrix}$$

is also a permutation.

(2) The identity or neutral element e is given by

$$e = \begin{pmatrix} 1 & 2 & ... & N \\ 1 & 2 & ... & N \end{pmatrix}.$$

(3) The composition of bijective functions is associative, i.e.,

$$P(QR) = (PQ)R.$$

(4) Each permutation P has a unique inverse P^{-1}, which is obtained by transposing the two rows defining P:

$$P^{-1} = \begin{pmatrix} p_1 & p_2 & ... & p_N \\ 1 & 2 & ... & N \end{pmatrix}.$$

This group is not only very important in physics but also quite complex, since it is noncommutative or non-Abelian, i.e., in general $PQ \neq QP$.

Cycles:
A more practical notation for permutations consists in expressing them as the product of cycles. For instance, in

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 7 & 5 & 2 & 6 & 3 \end{pmatrix}$$

we have a cycle consisting of

$$1 \to 4, \quad 4 \to 5, \quad 5 \to 2, \quad 2 \to 1,$$

another cycle given by

$$3 \to 7, \quad 7 \to 3$$

and finally the one-element cycle

$$6 \to 6\,.$$

One usually denotes cycles as a list of consecutive integers $(c_1, c_2, \ldots, c_m)$ such that $P(c_i) = c_{i+1}$ for $1 \leq i \leq m-1$ and $P(c_m) = c_1$. For example,

$$(1452), \quad (37) \quad \text{and} \quad (6)\,,$$

where it is implied that the last element of the cycle goes to the first one. Notice that the order of the integers denoting the cycle is irrelevant. For example,

$$(1452) \equiv (5214) \equiv (2145)\,.$$

Moreover, if two cycles C_1 and C_2 involve different integers, the order in which they are performed is immaterial, i.e., products between cycles commute provided that they have no element in common ($C_1 \cap C_2 \Rightarrow C_1 C_2 = C_2 C_1$).

Since the symmetric group is finite, for all permutations P there is an integer m such that

$$P^m = e\,.$$

The smallest of these integers is known as the *order of* P. In the case of cycles, the smallest number m for which

$$C^m = 1$$

is equal to the cycle length. For example, $(3,7)^2 = e$ and $(1452)^4 = e$. It is easy to see that the order of an arbitrary permutation P is equal to the minimum common multiple of the length of all its cycles.

Transposition decomposition:
Any cycle and thus any permutation can be written as the product of two-element cycles or transpositions. An arbitrary cycle $C = (c_1, c_2, ..., c_m)$ can be decomposed in the following $m-1$ transpositions:

$$(c_1, c_2, \ldots c_m) = (c_1, c_2)(c_2, c_3)\ldots(c_{m-1}, c_m)\,,$$

where we multiply, as usual, from right to left. This can also be written as

$$\begin{array}{ccccccc|c}
c_1 & c_2 & \ldots & & c_{m-1} & c_m & | & (c_{m-1}, c_m) \\
c_1 & c_2 & \ldots & & c_m & c_{m-1} & | & (c_{m-2}, c_{m-1}) \\
c_1 & c_2 & \ldots & c_m & c_{m-2} & c_{m-1} & & \\
 & & & & \ldots & & & \\
c_1 & c_2 & c_m & \ldots & c_{m-2} & c_{m-1} & | & (c_2, c_3) \\
c_1 & c_m & c_2 & \ldots & c_{m-2} & c_{m-1} & | & (c_1, c_2) \\
c_m & c_1 & c_2 & \ldots & c_{m-2} & c_{m-1} & &
\end{array}$$

where the pairs in brackets indicate the transpostion which has been applied to the cycle on the left. The number of transpositions in a product PQ of permutations is clearly the sum of the number of transpositions with decompositions in P and Q. Moreover, the number of transpositions in the inverse P^{-1} is the same as the number of transpositions in P. Permutations which are decomposed in an even (odd) number of transpositions are said to be even (odd). The identity is even and all transpositions are odd. A cycle is even if its length p is odd and vice versa.

The decomposition of a permutation P into a product of transpositions is clearly not unique. However, the parity of the number of transpositions involved in any of these decompositions is well defined. In order to show that consider two decompositions of the permutation P:

$$P = T_1\,T_2\,...\,T_k \quad \Leftrightarrow \quad P^{-1} = T_k\,T_{k-1}\,...\,T_2\,T_1$$

and

$$P = T'_1\,T'_2\,\ldots\,T'_m\,.$$

It follows that

$$e = P^{-1}P = (T_k\,T_{k-1}\,\ldots\,T_2\,T_1)\,(T'_1\,T'_2\,\ldots\,T'_m)$$

is even, which implies $m + k$ is even or that k and m have the same parity. This allows us to define the *sign of* P as

$$\mathrm{sgn}(P) = (-1)^{-p}$$

where p is the number of transpositions in any decomposition of P. Clearly, $\mathrm{sgn}(P)$ is an homomorphism of the symmetric group in $[-1, 1]$ since

$$\mathrm{sgn}(PQ) = \mathrm{sgn}(P)\,\mathrm{sgn}(Q) \quad \text{and} \quad \mathrm{sgn}(e) = 1\,.$$

The set of all even permutations forms a subgroup, since the product of any two even permutations is even. This subgroup is known (for historical reasons) as the alternating group G_N. Its order is $N!/2$. Any permutation in the symmteric group S_N is either an element of G_N or can be written as $P = TG$ where T is a transposition and $G \in G_N$. This can be seen by noting that for any permutation $P \notin G_N$ we have $T^{-1}P \in G_N$, where T is some transposition. Indeed, $TG = TG' \Leftrightarrow G = G'$ since $T^2 = 1$. One concludes that the number of odd and even permutations are equal to $N!/2$. Finally, if at all necessary, one may easily convince oneself that the identity does not allow any decomposition in an odd number of tranpositions. For $N = 1$ and $N = 2$ it is clear that the identity cannot be written as a single cycle. Each of the smaller cycles has to be in the corresponding subgroup of elements. Since the number of transpositions in the identity is even for smaller N, the assertion holds for all N.

Self-invariance of a group

Let $r \in G$ whre G is a group. Then $\{rs \,\text{with}\, s \in G\} = G$. In other words rs with $r \in G$ fixed and $s \in G$ arbitrary runs through all G. This implies

$$\sum_{s\in G} f(s) = \sum_{s\in G} f(rs).$$

Moreover, the set of all $\{s^{-1} \,\text{with}\, s \in G\} = G$. Therefore, we also have

$$\sum_{s\in G} f(s) = \sum_{s\in G} f(s^{-1}).$$

We will often apply these properties to permutations, namely,

$$\sum_{P} f(P) = \sum_{P} f(QP) = \sum_{P} f(P^{-1}).$$

4.2 Many-particle fermion states

4.2.1 Occupation numbers and Slater determinants

In the following we focus on fermions. Let us consider an arbitrary complete set of orthonormal single-particle wave functions (often known as spin orbitals):
$\{\varphi_1(x),\ \varphi_2(x), \ldots \varphi_\infty(x)\} = \{\varphi_\alpha,\ \alpha = 1, 2, \ldots \infty\}$ where

$$\sum_{\alpha} \varphi^*_\alpha(x)\ \varphi_\alpha(x') = \delta(x - x')$$

and

$$\int dx\ \varphi_\alpha^*(x)\ \varphi_\beta(x) = \delta_{\alpha\beta}.$$

Note that throughout this chapter we use a compact notation for the particle's coordinates $x_i \equiv (\vec{r}_i, \sigma_i)$, where $\vec{r}_i$ refers to the position and σ_i to the projection of the particle spin along the z axis. Consequently, $\int dx$ actually stands for

$$\int dx \equiv \sum_\sigma \int d^3r.$$

We search for a general expression of $\Psi(x_1, \ldots x_N)$ for fermions in terms of $\varphi_\alpha(x)$ with $\alpha = 1, \ldots \infty$. Any antisymmetrical wave function $\Psi(x_1, \ldots x_N)$ can be expressed as a superposition (linear combination) of antisymmetrized products of N single-particle states, since a superposition of states with different symmetry would be neither symmetric nor antisymmetric. It is, however, instructive to show this explicitly by constructing the antisymmetric elementary N-particle states.

Without loss of generality we expand

$$\Psi(x_1, \ldots x_N) = \sum_{K_1=1}^{\infty} \sum_{K_2=1}^{\infty} \ldots \sum_{K_N=1}^{\infty} c\,(K_1, K_2, \ldots K_N)\ \ \varphi_{K_1}(x_1)\ \varphi_{K_2}(x_2) \ldots \varphi_{K_N}(x_N).$$

Notice that the indices K_i run over the single-particle basis whereas x_i refers to the coordinates of the particles ($1 \leq K_i \leq \infty$ and $1 \leq i \leq N$). Using the orthonormality of the $\varphi_\alpha(x)$ we have

$$c\,(K_1, \ldots K_N) = \int dx_1, \ldots dx_N\ \ \varphi_{K_1}^*(x_1) \ldots \varphi_{K_N}^*(x_N)\ \ \Psi(x_1, \ldots x_N).$$

Taking into account that $\Psi(x_i, x_j) = -\Psi(x_j, x_i)$ for all i and j, it is easy to see that

$$c\,(K_1, \ldots K_i, \ldots K_j, \ldots K_N) = -c\,(K_1, \ldots K_j, \ldots K_i, \ldots K_N).$$

To show this, simply interchange the indices in the wave function and change the integration variables appropriately. One concludes that if $K_i = K_j$ for any i and j the coefficient $c\,(K_1, \ldots K_i, \ldots K_j, \ldots K_N) = 0$. In other words, two fermions cannot occupy the same state. This is known as *Pauli exclusion principle*. Of course, the index K_i includes the labeling of states having different spin projections. Thus one often says that two electrons having the same spin cannot occupy the same orbital.

For a general permutation P, we have

$$c\left(K_{P(1)} \ldots K_{P(N)}\right) = (-1)^p\ c\,(K_1 \ldots K_N).$$

Therefore, all the coefficients corresponding to a given set of spin orbitals $K_1, \ldots K_N$ in any order can be expressed in terms of the coefficient $c\,(K_1, \ldots K_N)$ having $K_1 < K_2 <$

$\ldots < K_N$. The order of $K_1, \ldots K_N$ only afects the nontrivial phase factor $(-1)^p$. We denote the coefficient having "normal ordering" by

$$c\,(K_1 < K_2 < \ldots < K_N).$$

Since the order of the states is fixed in $c\,(K_1 < K_2 < \ldots < K_N)$ this coefficient depends only on the choice of the states $K_1, \ldots K_N$. In other words, $c\,(K_1 < \ldots < K_N)$ depends only on the *occupation numbers* $n_\alpha = 0$ or 1 of the single-particle states $\alpha = 1, 2, \ldots \infty$. It is then useful to write

$$c\,(K_1 < K_2 < \ldots < K_N) = \frac{1}{\sqrt{N!}}\, c\,(n_1, n_2, \ldots n_\infty)\,,$$

where $n_\alpha = 1$ for $\alpha = K_1, K_2, \ldots$ and K_N and $n_\alpha = 0$ otherwise. Notice that $\sum_\alpha n_\alpha = N$ is the number of particles.

We may then replace the sum

$$\sum_{K_1=1}^{\infty} \sum_{K_2=1}^{\infty} \ldots \sum_{K_N=1}^{\infty} \ldots$$

with $K_i \neq K_j\ \forall i, j$ by a sum

$$\sum_{n_1, n_2, \ldots n_\infty=0}^{1} \sum_{P} \ldots$$

over all possible occupation numbers $n_1, n_2, \ldots n_\infty$ (having $n_\alpha = 0, 1$ and $\sum_\alpha n_\alpha = N$) and over all permutations of the order of $K_1, \ldots K_N$. Formally, we may write

$$\Psi(x_1, \ldots x_N) = \sum_{n_1, n_2, \ldots n_\infty=0}^{1} \sum_{P}\ c\left(K_{P(1)}, \ldots K_{P(N)}\right)\ \varphi_{K_{P(1)}}(x_1) \ldots \varphi_{K_{P(N)}}(x_N)$$

where the sum $\sum_{n_1, n_2 \ldots n_\infty}$ runs over all possible choices of the N occupied states ($\sum_\alpha n_\alpha = N$) and the sum $\sum_P$ takes care of the order of $K_1, \ldots K_N$.

Referring all the $c\left(K_{P(1)}, \ldots K_{P(N)}\right)$ to the coefficient in normal ordering we have

$$\begin{aligned}
\Psi(x_1, \ldots x_N) &= \sum_{n_1, \ldots n_\infty=0}^{1} \sum_{P} (-1)^p\, \frac{1}{\sqrt{N!}}\, c\,(n_1, \ldots n_\infty)\ \varphi_{K_{P(1)}}(x_1) \ldots \varphi_{K_{P(N)}}(x_N) \\
&= \sum_{n_1, \ldots n_\infty=0}^{1} c\,(n_1, \ldots n_\infty)\ \frac{1}{\sqrt{N!}} \sum_{P} (-1)^p\ \varphi_{K_{P(1)}}(x_1) \ldots \varphi_{K_{P(N)}}(x_N) \\
&= \sum_{n_1, \ldots n_\infty=0}^{1} c\,(n_1, \ldots n_\infty)\ \frac{1}{\sqrt{N!}} \sum_{P} (-1)^p\ \varphi_{K_1}(x_{P(1)}) \ldots \varphi_{K_N}(x_{P(N)}) \\
\Psi(x_1, \ldots x_N) &= \sum_{n_1, \ldots n_\infty=0}^{1} c\,(n_1, \ldots n_\infty)\ \Phi_{n_1, \ldots n_\infty}(x_1, x_2, \ldots x_N)
\end{aligned}$$

where

$$\Phi_{n_1,\ldots n_\infty}(x_1,\ldots x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{K_1}(x_1) & \varphi_{K_1}(x_2) & \ldots & \varphi_{K_1}(x_N) \\ \varphi_{K_2}(x_1) & \varphi_{K_2}(x_2) & \ldots & \varphi_{K_2}(x_N) \\ \vdots & \vdots & & \vdots \\ \varphi_{K_N}(x_1) & \varphi_{K_N}(x_2) & \ldots & \varphi_{K_N}(x_N) \end{vmatrix}$$

is the simplest fully antisymmetrical state that can be formed by the superposition of N single-particle states. $\Phi_{n_1,\ldots n_\infty}$ is known as a *Slater determinant*. It depends only on the occupation numbers $n_1,\ldots n_\infty$, since the order of the K_i is such that $K_1 < K_2 < \ldots < K_N$. We may then write

$$\langle\, x_1,\ldots x_N \,|\, n_1,\ldots n_\infty \,\rangle = \Phi_{n_1,\ldots n_\infty}(x_1,\ldots x_N)$$

where $|\, n_1,\ldots n_\infty \,\rangle$ is the ket with occupation numbers $n_1,\ldots n_\infty$. Clearly $\Phi_{n_1,\ldots n_\infty}$ or $|\, n_1,\ldots n_\infty \,\rangle$ form a complete set of N particle states ($\sum_\alpha n_\alpha = N$) that are the antisymmetric superposition of single-particle states φ_α.

The orthonormality of the Slater determinants $\Phi_{n_1,n_2,\ldots n_\infty}$ will be demonstrated in Sec. 4.2.3. At this point, however, it is worth noting that it implies the orthormality of the kets with defined occupation numbers, namely,

$$\langle\, n_1, n_2,\ldots n_\infty \,|\, n'_1, n'_2,\ldots n'_\infty \,\rangle = \delta_{n_1 n'_1}\, \delta_{n_2 n'_2} \ldots \delta_{n_\infty n'_\infty}\,,$$

since

$$\begin{aligned} \langle\, n_1,\ldots n_\infty \,|\, n'_1,\ldots n'_\infty \,\rangle &= \int dx_1,\ldots dx_n\, \overline{\langle\, x_1,\ldots x_N \,|\, n_1,\ldots n_\infty \,\rangle} \langle\, x_1,\ldots x_N \,|\, n'_1,\ldots n'_\infty \,\rangle \\ &= \int dx_1,\ldots dx_n\, \Phi^*_{n_1,\ldots n_\infty}(x_1,\ldots x_N)\, \Phi_{n'_1,\ldots n'_\infty}(x_1,\ldots x_N) \\ &= \delta_{K_1 K'_1}\, \delta_{K_2 K'_2} \ldots \delta_{K_N K'_N} \\ &= \delta_{n_1 n'_1}\, \delta_{n_2 n'_2} \ldots \delta_{n_\infty n'_\infty}, \end{aligned}$$

where $K_1 < K_2 < \ldots < K_N$, $K'_1 < K'_2 < \ldots < K'_N$ and we have used that $\int |\, x_1,\ldots x_N \,\rangle\, \langle\, x_1,\ldots x_N \,|\, dx_1,\ldots dx_N = \mathbb{1}$.

A note on matrix determinants

The *Leibniz Formula* for the determinant of an $n \times n$ matrix A is

$$\det A = \sum_P (-1)^p\, A_{1P(1)}\, A_{2P(2)} \ldots A_{nP(n)}$$

where $p = O(P)$ is the order of the permutation P. This can be written in component form as

$$\det A = \sum_P (-1)^p \prod_{i=1}^{n} A_{iP(i)},$$

where the sum $\sum_P$ runs over all the $n!$ permutations of n integers.

It is easy to show that $\det A = \det A^t$ where $(A^t)_{ij} = A_{ji}$ is the transposed matrix. The permutations are all bijective functions and the order of the factors $A_{iP(i)}$ in the product $\prod_{i=1}^{n}$ is immaterial. Therefore we can change, for each P, the index of the product by $j = P(i)$ or equivalently $i = P^{-1}(j)$. Thus

$$\det A = \sum_P (-1)^p \prod_{j=1}^{n} A_{P^{-1}(j)j}.$$

In any group, and in particular in the permutation group, the inverse element is unique, so we may replace $\sum_P$ by $\sum_{P^{-1}}$. Moreover, given two permutations P_1 and P_2 with $P_3 = P_1 P_2$ we have that the orders of P_i, i.e., the number of transpositions, satisfy

$$p_3 = p_1 + p_2 \quad \text{for} \quad P_3 = P_1 P_2.$$

In particular $\mathbb{1} = PP^{-1} \quad \Rightarrow \quad O(\mathbb{1}) = O(P) + O(P^{-1})$. Thus

$$1 = (-1)^{O(\mathbb{1})} = (-1)^{O(P)}\,(-1)^{O(P^{-1})} \quad \Rightarrow \quad (-1)^{O(P^{-1})} = (-1)^p.$$

Finally, redefining $Q = P^{-1}$ we have

$$\det A = \sum_Q (-1)^q \prod_{j=1}^{n} A_{Q(j)j} = \det(A^t).$$

4.2.2 Creation, annihilation, and number operators

The notation $|\, n_1, \ldots n_\infty \,\rangle$ is very cumbersome since we need to explicitly indicate the occupation number of an infinite number of single particle states, most of which are zero ($\sum_\alpha n_\alpha = N$) and, even more important, most of which do not change in an elementary interaction or scattering process among the particles. To overcome the problem, the idea is to use a reference state, in our case the empty state

$$|\, 0 \,\rangle = |\, n_1 = 0, n_2 = 0, \ldots n_\infty = 0 \,\rangle$$

and to introduce operators that add the electrons one by one. These operators act in the Hilbert space spanned by the $|\, n_1, n_2 \ldots n_\infty \,\rangle$ for all $n_\alpha = 0, 1$. They are called *creation operators* since they increase the occupation numbers by 1.

The simplest state is a one-electron state $|\, 1_\alpha \,\rangle = |\, 0 \ldots 1_\alpha \ldots 0 \,\rangle$ whose coordinate representation is $\langle\, x \,|\, 1_\alpha \,\rangle = \varphi_\alpha(x)$. We require that the creation operator $\hat{a}^\dagger_\alpha$ satisfies[10]

$$\hat{a}^\dagger_\alpha |\, 0 \,\rangle = |\, 1_\alpha \,\rangle$$

[10] A hat $\hat{}$ is used to distinguish operators from numbers.

where $|\,0\,\rangle$ is the empty state. Two-electron states are obtained by adding an electron on top of a single electron state using the same creation operators:

$$\hat{a}^\dagger_\beta |\,1_\alpha\,\rangle = \hat{a}^\dagger_\beta\,\hat{a}^\dagger_\alpha\,|\,0\,\rangle.$$

They must represent the same physical state as the Slater determinant constructed using the orbitals $\varphi_\beta(x)$ and $\varphi_\alpha(x)$, namely, *with the rows respecting the order in which the fermions have been added.* More precisely,

$$\begin{aligned}
\langle\, x_1, x_2\,|\,\hat{a}^\dagger_\beta\,\hat{a}^\dagger_\alpha\,|\,0\,\rangle \quad &= \tfrac{1}{\sqrt{2}} \begin{vmatrix} \varphi_\beta(x_1) & \varphi_\beta(x_2) \\ \varphi_\alpha(x_1) & \varphi_\alpha(x_2) \end{vmatrix} \\
&= -\tfrac{1}{\sqrt{2}} \begin{vmatrix} \varphi_\alpha(x_1) & \varphi_\alpha(x_2) \\ \varphi_\beta(x_1) & \varphi_\beta(x_2) \end{vmatrix} \\
&= -\langle\, x_1, x_2\,|\,\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,|\,0\,\rangle\,.
\end{aligned}$$

Therefore,

$$\hat{a}^\dagger_\beta\,\hat{a}^\dagger_\alpha\,|\,0\,\rangle = -\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,|\,0\,\rangle \quad \forall\,\alpha,\beta.$$

The operators $\hat{a}^\dagger_\beta$ and $\hat{a}^\dagger_\alpha$ anticommute, and in particular $(a^\dagger_\alpha)^2\,|\,0\,\rangle = 0 \quad \forall\,\alpha$.

The present choice of phase for the states $\hat{a}^\dagger_\beta\,\hat{a}^\dagger_\alpha\,|\,0\,\rangle$, that leads to the anticommutation rule $\{\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\} = 0$, is of course not the only possible one. However, a major advantage of the anticommutation rule introduced by Jordan and Wigner in 1928 is that the antisymmetry of the many-body states is taken into account by the operator algebra. The sign $(-1)^p$ associated to the permutation that brings a given Slater determinant into "normal ordering" $(K_1 < K_2 < \ldots < K_N)$ is given by the order of the operators $\hat{a}^\dagger_\alpha$. A further advantage of this choice will become clear when computing the matrix elements of operators between different basis states $|\,n_1, n_2, \ldots n_\infty\,\rangle$.

One can proceed of course recursively by adding the particles one by one. For a general N-particle state we require

$$\langle\, x_1, \ldots x_N\,|\,\hat{a}^\dagger_{K_1}\,\hat{a}^\dagger_{K_2}\ldots\hat{a}^\dagger_{K_N}\,|\,0\,\rangle = \frac{1}{\sqrt{N!}}\ \det \begin{vmatrix} \varphi_{K_1}(x_1) & \ldots & \varphi_{K_1}(x_N) \\ \varphi_{K_2}(x_1) & \ldots & \varphi_{K_2}(x_N) \\ \vdots & & \vdots \\ \varphi_{K_N}(x_1) & \ldots & \varphi_{K_N}(x_N) \end{vmatrix}.$$

Notice that this holds for any order of the operators $\hat{a}^\dagger_{K_i}$, even if some K_i are repeated, provided that the order of the K_i in both sides of the equation is the same. Since the previous identity holds for any $|\,x_1, \ldots x_N\,\rangle$ we must have the operator identity

$$\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta = -\hat{a}^\dagger_\beta\,\hat{a}^\dagger_\alpha \quad \forall\,\alpha,\beta.$$

Thus, $\{\hat{a}^\dagger_\alpha, \hat{a}^\dagger_\beta\} = 0$, and in particular $(\hat{a}^\dagger_\alpha)^2 = 0 \quad \forall\,\alpha$.

The previous identification between a product of creation operators acting on the vacuum state $|\,0\,\rangle$ and a Slater determinant is valid for any order of the $\hat{a}^\dagger_{K_i}$. If appropriate, the sign of the Slater determinant changes accordingly. For the basis states $|\,n_1, n_2, \ldots n_\infty\,\rangle$, however, the order of the K_i is fixed by convection to be $K_1 < K_2 < \ldots < K_N$. Thus,

$$\langle\, x_1, \ldots x_N \,|\, n_1, n_2, \ldots n_\infty \,\rangle = \frac{1}{\sqrt{N!}} \det \begin{vmatrix} \varphi_{K_1}(x_1) & \ldots & \varphi_{K_1}(x_N) \\ \varphi_{K_2}(x_1) & \ldots & \varphi_{K_2}(x_N) \\ \vdots & & \vdots \\ \varphi_{K_N}(x_1) & \ldots & \varphi_{K_N}(x_N) \end{vmatrix}$$

with $K_1 < K_2 < \ldots < K_N$. Consequently, we may write

$$|\, n_1, n_2, \ldots n_\infty \,\rangle = (\hat{a}^\dagger_1)^{n_1}\, (\hat{a}^\dagger_2)^{n_2} \ldots (\hat{a}^\dagger_\infty)^{n_\infty} |\, 0 \,\rangle$$

which ensures that the $a^\dagger_i$ act in normal or conventional order.

Annihilation or destruction operators:

Once the commutation properties of the creation operators $a^\dagger_\alpha$ are known, it is easy to derive the commutation rules and matrix elements of the hermitic conjugate operator

$$(\hat{a}^\dagger_\alpha)^\dagger = \hat{a}_\alpha.$$

Since $\hat{a}^\dagger_\alpha\,|\,0\,\rangle = |\,1_\alpha\,\rangle \;\Rightarrow\; 1 = \langle\, 1_\alpha \,|\, \hat{a}^\dagger_\alpha \,|\, 0\,\rangle \;\Rightarrow\; \langle\, 0 \,|\, \hat{a}_\alpha \,|\, 1_\alpha\,\rangle = 1 \;\Rightarrow\; \hat{a}_\alpha\,|\,1_\alpha\,\rangle = |\,0\,\rangle$. In the last step we have used that $\langle\, 0 \,|\, n_1, \ldots n_\infty \,\rangle = 0$ for all $|\, n_1, \ldots n_\infty \,\rangle$ having $\sum_\alpha n_\alpha > 0$. Thus a_α reduces the occupation of the single-particle state α by 1. They are therefore known as *annihilation or destruction operators.*

The anticommutation rules for $\hat{a}^\dagger_\alpha$, namely $\{\hat{a}^\dagger_\alpha,\, \hat{a}^\dagger_\beta\} = 0 \;\forall\, \alpha, \beta \;\Rightarrow\; \{\hat{a}_\alpha,\, \hat{a}_\beta\} = 0 \;\forall\, \alpha, \beta$. In particular we have

$$(\hat{a}^\dagger_\alpha)^2 = 0 \;\Rightarrow\; (\hat{a}_\alpha)^2 = 0.$$

Since $\hat{a}_\alpha\,|\,1_\alpha\,\rangle = |\,0\,\rangle$ and $\hat{a}^2_\alpha = 0$, we have

$$\hat{a}_\alpha\,|\,0\,\rangle = 0 \;\; \forall\, \alpha.$$

Actually, the vacuum state can be defined by this property.

In order to generalize the previous relations to arbitrary basis states $|\,n_1, \ldots n_\infty\,\rangle$ we consider

$$\begin{aligned} \hat{a}^\dagger_\alpha\,|\,n_1, \ldots n_\alpha, \ldots n_\infty\,\rangle &= \hat{a}^\dagger_\alpha\, (\hat{a}^\dagger_1)^{n_1}, \ldots (\hat{a}^\dagger_\alpha)^{n_\alpha}, \ldots (\hat{a}^\dagger_\infty)^{n_\infty}\,|\,0\,\rangle \\ &= \delta_{n_\alpha 0}\,(-1)^{S_\alpha}\,|\,n_1, \ldots n_\alpha + 1, \ldots n_\infty\,\rangle \\ &= \delta_{n_\alpha 0}\,(-1)^{S_\alpha}\,|\,n_1, \ldots 1_\alpha, \ldots n_\infty\,\rangle \end{aligned}$$

where $S_\alpha = \sum_{\beta<\alpha} n_\beta$. Then we have

$$\langle\, n_1, \ldots n'_\alpha, \ldots n_\infty \,|\, \hat{a}^\dagger_\alpha \,|\, n_1, \ldots n_\alpha, \ldots n_\infty \,\rangle = \delta_{n_\alpha 0}\, (-1)^{S_\alpha}\, \delta_{n'_\alpha 1}.$$

Conjugating one obtains

$$\langle\, n_1, \ldots n_\alpha, \ldots n_\infty \,|\, \hat{a}_\alpha \,|\, n_1, \ldots n'_\alpha, \ldots n_\infty \,\rangle = \delta_{n_\alpha 0}\, \delta_{n'_\alpha 1}\, (-1)^{S_\alpha},$$

which implies that

$$\hat{a}_\alpha \,|\, n_1, \ldots n_\alpha, \ldots n_\infty \,\rangle = \delta_{n_\alpha 1}\, (-1)^{S_\alpha} \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle.$$

To be more specific,

$$\begin{cases} \hat{a}^\dagger_\alpha \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle = (-1)^{S_\alpha} \,|\, n_1, \ldots 1_\alpha, \ldots n_\infty \,\rangle \\ \hat{a}^\dagger_\alpha \,|\, n_1, \ldots 1_\alpha, \ldots n_\infty \,\rangle = 0 \end{cases} \tag{4.1}$$

and

$$\begin{cases} \hat{a}_\alpha \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle = 0 \\ \hat{a}_\alpha \,|\, n_1, \ldots 1_\alpha, \ldots n_\infty \,\rangle = (-1)^{S_\alpha} \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle\ . \end{cases} \tag{4.2}$$

We may now demonstrate a very important anticommutation rule, namely, $\{a_\alpha, a^\dagger_\alpha\} = 1$. To this end we consider

$$\begin{aligned} (\hat{a}^\dagger_\alpha\, \hat{a}_\alpha + \hat{a}_\alpha\, \hat{a}^\dagger_\alpha) \,|\, n_1, \ldots n_\alpha, \ldots n_\infty \,\rangle \;&=\; \hat{a}^\dagger_\alpha\, \hat{a}_\alpha \,|\, n_1, \ldots 1_\alpha, \ldots n_\infty \,\rangle\, \delta_{n_\alpha 1} + \\ &+\; \hat{a}_\alpha\, \hat{a}^\dagger_\alpha \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle\, \delta_{n_\alpha 0}\ , \end{aligned}$$

where we have split the ket in the two possible occupations of the spinorbital α ($n_\alpha = 0$ or 1). Applying Eq. (4.2) on the first term ($n_\alpha = 1$) and Eq. (4.1) on the second term ($n_\alpha = 0$) we obtain

$$\begin{aligned} (\hat{a}^\dagger_\alpha\, \hat{a}_\alpha + \hat{a}_\alpha\, \hat{a}^\dagger_\alpha) \,|\, n_1, \ldots n_\alpha, \ldots n_\infty \,\rangle \;&=\; \delta_{n_\alpha 1}\, (-1)^{S_\alpha}\, \hat{a}^\dagger_\alpha \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle + \\ &+\; \delta_{n_\alpha 0}\, (-1)^{S_\alpha}\, \hat{a}_\alpha \,|\, n_1, \ldots 1_\alpha, \ldots n_\infty \,\rangle\ . \end{aligned}$$

Applying again Eqs. (4.1) and (4.2) as appropriate, we finally obtain

$$\begin{aligned} (\hat{a}^\dagger_\alpha\, \hat{a}_\alpha + \hat{a}_\alpha\, \hat{a}^\dagger_\alpha) \,|\, n_1, \ldots n_\alpha, \ldots n_\infty \,\rangle \;&=\; \delta_{n_\alpha 1} \,|\, n_1, \ldots 1_\alpha, \ldots n_\infty \,\rangle + \delta_{n_\alpha 0} \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle \\ &=\; |\, n_1, \ldots n_\alpha, \ldots n_\infty \,\rangle\,, \end{aligned}$$

which proves the statement. A similar calculation yields the anticommutation relation $\{\hat{a}_\alpha, \hat{a}^\dagger_\beta\} = 0$ for $\alpha \neq \beta$. On the one hand we have

$$\begin{aligned} \hat{a}_\alpha\, \hat{a}^\dagger_\beta \,|\, n_1, \ldots n_\alpha, \ldots n_\infty \,\rangle \;&=\; \hat{a}_\alpha\, \hat{a}^\dagger_\beta\, (\hat{a}^\dagger_1)^{n_1} \ldots (\hat{a}^\dagger_\alpha)^{n_\alpha} \ldots (\hat{a}^\dagger_\infty)^{n_\infty} \,|\, 0 \,\rangle \\ &=\; \hat{a}_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}^\dagger_\alpha\, (\hat{a}^\dagger_1)^{n_1} \ldots (\hat{a}^\dagger_\infty)^{n_\infty} \,|\, 0 \,\rangle\, \delta_{n_\alpha 1}\, (-1)^{S_\alpha} \\ &=\; \hat{a}_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}^\dagger_\alpha \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle\, \delta_{n_\alpha 1}\, (-1)^{S_\alpha}\,, \end{aligned}$$

where we have used that only the case $n_\alpha = 1$ gives a nonvanishing result, since otherwise applying $\hat{a}_\alpha$ would yield zero ($\alpha \neq \beta$) and we have moved the operator $a^\dagger_\alpha$ to the left by taking into account the phase factor $(-1)^{S_\alpha}$ due to any possible anticommutations. Using known commutation relations we obtain

$$\begin{aligned} \hat{a}_\alpha\, \hat{a}^\dagger_\beta \,|\, n_1, \ldots n_\alpha, \ldots n_\infty \,\rangle &= (-1)^{S_\alpha+1}\, \delta_{n_\alpha 1}\, \underbrace{\hat{a}_\alpha\, \hat{a}^\dagger_\alpha}_{1-\hat{a}^\dagger_\alpha\, \hat{a}_\alpha}\, \hat{a}^\dagger_\beta \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle \\ &= (-1)^{S_\alpha+1}\, \delta_{n_\alpha 1}\, \hat{a}^\dagger_\beta \,|\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle. \end{aligned} \tag{4.3}$$

On the other hand, concerning the operations in commuted order, we have

$$\hat{a}^\dagger_\beta\, \hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle = \hat{a}^\dagger_\beta\, \delta_{n_\alpha 1}\, (-1)^{S_\alpha} |\, n_1, \ldots 0_\alpha, \ldots n_\infty \,\rangle\,, \tag{4.4}$$

where have simply applied Eq. (4.2). Comparing Eqs. (4.3) and (4.4) we conclude that these operators anticommute:

$$\hat{a}^\dagger_\beta\, \hat{a}_\alpha = -\hat{a}_\alpha\, \hat{a}^\dagger_\beta.$$

We may finally summarize the fermion creation-annihilation commutation rules

$$\begin{aligned} \{\hat{a}^\dagger_\alpha,\, \hat{a}^\dagger_\beta\} &= 0 \\ \{\hat{a}_\alpha,\, \hat{a}_\beta\} &= 0 \\ \{\hat{a}_\alpha,\, \hat{a}^\dagger_\beta\} &= \delta_{\alpha\beta}\,, \end{aligned}$$

which are central to the second quantization algebra.

Number operators

The number operator is defined by

$$\hat{n}_\alpha = \hat{a}^\dagger_\alpha\, \hat{a}_\alpha$$

and is obviously hermitic ($\hat{n}^\dagger_\alpha = \hat{n}_\alpha$). Its name is justified by the fact that it counts the number of particles in the orbital α:

$$\begin{aligned} \hat{a}^\dagger_\alpha\, \hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle &= \hat{a}^\dagger_\alpha\, \delta_{n_\alpha 1}\, (-1)^{S_\alpha} |\, n_1, \ldots 0_\alpha \ldots n_\infty \,\rangle \\ &= \delta_{n_\alpha 1}\, |\, n_1 \ldots n_\infty \,\rangle \\ &= n_\alpha \,|\, n_1, \ldots n_\infty \,\rangle \end{aligned}$$

where $n_\alpha = 0$ or 1 is the occupation number of orbital α. Consequently,

$$\hat{n}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle = n_\alpha \,|\, n_1, \ldots n_\infty \,\rangle.$$

Another important property is

$$[\hat{n}_\alpha, \hat{n}_\beta] = 0.$$

This can be shown by noting that for $\alpha \neq \beta$ $\hat{a}^\dagger_\alpha\, \hat{a}_\alpha\, \hat{a}^\dagger_\beta = -\hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\alpha = \hat{a}^\dagger_\beta\, \hat{a}^\dagger_\alpha\, \hat{a}_\alpha$, or equivalently

$$[\hat{a}^\dagger_\beta, \hat{n}_\alpha] = 0 \;\; \forall\, \alpha \neq \beta.$$

Hermitic conjugation using that $\hat{n}_\alpha^\dagger = \hat{n}_\alpha$ yields

$$[\hat{a}_\beta, \hat{n}_\alpha] = 0 \quad \forall\, \alpha \neq \beta.$$

In conclusion, the $\hat{n}_\alpha$ constitute a complete set of commuting operators in the Hilbert space spanned by $|\, n_1, \ldots n_\infty \,\rangle$. In fact, they are the complete set of compatible observables that define the occupation-number representation.

The eigenvalues of $\hat{n}_\alpha$ are 0 or 1 since $\hat{n}_\alpha^2 = \hat{a}_\alpha^\dagger\, \hat{a}_\alpha\, \hat{a}_\alpha^\dagger\, \hat{a}_\alpha = \hat{a}_\alpha^\dagger\, (1 - \hat{a}_\alpha^\dagger\, \hat{a}_\alpha)\, \hat{a}_\alpha = \hat{a}_\alpha^\dagger\, \hat{a}_\alpha = \hat{n}_\alpha$. This implies that $\hat{n}_\alpha$ is a projection operator.

Single-particle and two-particle processes:

A *single-particle process* is a transition between many-body states, for example, kets $|\, n_1, \ldots n_\infty \,\rangle$ with well-defined occupation numbers or any superposition thereof, in which only one particle changes its state (see Fig. 4.2.2). Mathematically, this corresponds to

$$\begin{aligned} \hat{a}_\alpha^\dagger\, \hat{a}_\beta\, |\, n_1, \ldots n_\infty \,\rangle &= n_\beta\, (-1)^{S_\beta}\, \hat{a}_\alpha^\dagger\, |\, n_1, \ldots 0_\beta, \ldots n_\infty \,\rangle \\ &= n_\beta\, (-1)^{S_\beta}\, (1 - n_\alpha)\, (-1)^{S_\alpha}\, |\, n_1, \ldots 1_\alpha, \ldots 0_\beta, \ldots n_\infty \,\rangle. \end{aligned}$$

In the case of fermions, for the result not to be zero, we need that the initial spinorbital is occupied and final one is empty, i.e., $n_\beta = 1$ and $n_\alpha = 0$. Thus, we may write

$$\hat{a}_\alpha^\dagger\, \hat{a}_\beta\, |\, n_1, \ldots 0_\alpha, \ldots 1_\beta, \ldots n_\infty \,\rangle = n_\beta (1 - n_\alpha)\, (-1)^{\Sigma(\alpha+1, \beta-1)}\, |\, n_1, \ldots 1_\alpha, \ldots 0_\beta, \ldots n_\infty \,\rangle.$$

Here we assumed $\alpha < \beta$ and introduced the notation $\Sigma(\alpha, \beta) = \sum_{\gamma=\alpha}^{\beta} n_\gamma$, so that the number of tranpositions to bring a single creation operator from the position corresponding to α to the position corresponding to β is given by

$$\Sigma(\alpha + 1, \beta - 1) = \sum_{\alpha < \gamma < \beta} n_\gamma.$$

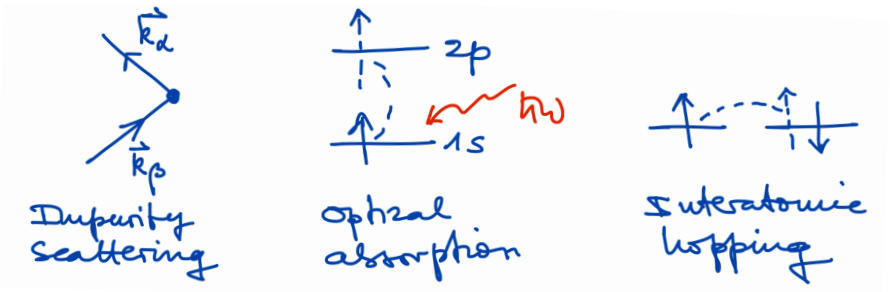


Figure 9: Illustration of single-particle processes

Notice that

$$S_\beta + S_\alpha = \sum_{\gamma<\beta} n_\gamma + \sum_{\gamma<\alpha} n_\gamma = 2\sum_{\gamma<\alpha} n_\gamma + \sum_{\gamma=\alpha}^{\beta-1} n_\gamma = 2\sum_{\gamma<\alpha} n_\gamma + \sum_{\gamma=\alpha+1}^{\beta-1} n_\gamma$$
$$= 2\sum_{\gamma<\alpha} n_\gamma + \Sigma\left(\alpha+1, \beta-1\right),$$

where we have used that $n_\alpha = 0$. These single-particle transitions can represent, for example, a scattering process $\hat{a}^\dagger_{k'}\hat{a}_k$ from a plane wave or Bloch state with wave vector k to a state with wave vector k', as they take place at impurities and interfaces, or the hopping $\hat{a}^\dagger_{j\beta\sigma}\hat{a}_{i\alpha\sigma}$ of an electron with spin σ from the orbital α of atom i to the orbital β of atom j. These processes are such that the occupation of only one single-particle state changes.

Two-particle processes involve a pair of particles whose state is modified as the result of their interaction (see Fig. 4.2.2). In the occupation-number representation this is described by saying that the occupation numbers of two initial spinorbitals γ and δ are reduced and the occupation number of two empty ones α and β is increased. Let us consider the operator $\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\gamma\,\hat{a}_\delta$ with $\alpha < \beta < \gamma < \delta$, which can be written as $\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\gamma\,\hat{a}_\delta = \hat{a}^\dagger_\alpha\,\hat{a}_\delta\,\hat{a}^\dagger_\beta\,\hat{a}_\gamma$ since all the indices were assumed to be different. Applying this operator to an arbitrary many-body state we obtain

$$\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\gamma\,\hat{a}_\delta\,|\,n_1,\ldots n_\infty\,\rangle = n_\delta\,n_\gamma\,(1-n_\beta)\,(1-n_\alpha)\,(-1)^{\Sigma(\beta+1,\gamma-1)+\Sigma(\alpha+1,\delta-1)}\times$$
$$|\,n_1,\ldots 1_\alpha,\ldots 1_\beta,\ldots 0_\gamma,\ldots 0_\delta,\ldots n_\infty\,\rangle\,,$$

where $\Sigma(\beta+1,\gamma-1)+\Sigma(\alpha+1,\delta-1)$ counts the number of transpositions involved in bringing one operators to its crorresponding normal-order position from δ to α and another one from γ to β. Notice that for $\alpha<\beta<\gamma<\delta$ the transition $\hat{a}^\dagger_\beta\,\hat{a}_\gamma$ does not modify the sum $\Sigma(\alpha+1,\delta-1) = \sum_{\alpha<\gamma<\delta} n_\gamma$. An example of a two-particle process is the

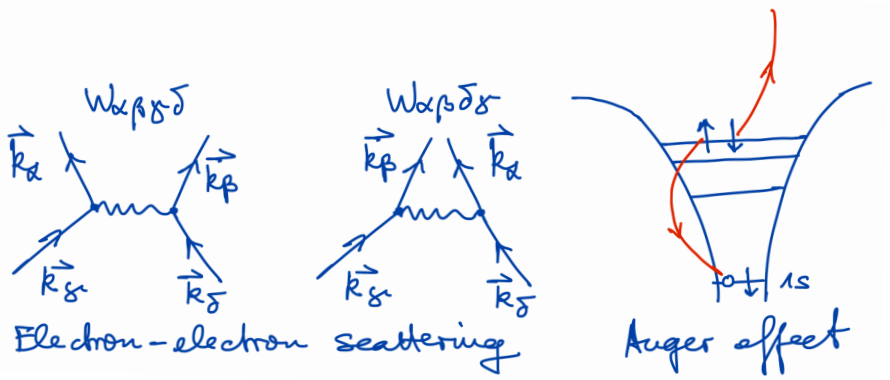


Figure 10: Illustration of two-particle processes

Coulomb scattering between a pair of electrons from the states k_1 and k_2 to the states k_3 and k_4, which is described by $\hat{a}^\dagger_{k_4}\,\hat{a}^\dagger_{k_3}\,\hat{a}_{k_2}\,\hat{a}_{k_1}$.

It is clear that the total number of particles is conserved in both single-particle and two-particle transitions, as confirmed by the fact that the operator of the total number of particles

$$\hat{N} = \sum_\alpha \hat{n}_\alpha$$

commutes with both $\hat{a}^\dagger_\alpha\,\hat{a}_\beta$ and $\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\gamma\,\hat{a}_\delta$.

4.2.3 Matrix elements of single-particle operators

Before we compute the matrix elements of operators between Slater determinants it is instructive to verify their orthonormalization of the wave functions

$$\langle\, x_1, \ldots x_N \,|\, n_1, \ldots n_\infty \,\rangle = \frac{1}{\sqrt{N!}} \sum_P (-1)^p\, \varphi_{K_{P(1)}}(x_1)\, \varphi_{K_{P(2)}}(x_2)\, \ldots\, \varphi_{K_{P(N)}}(x_N),$$

where $K_1 < K_2 < \ldots < K_N$ are the states occupied in $|\, n_1, \ldots n_\infty \,\rangle$, and the sum runs over all permutations P. We would like to calculate

$$\begin{aligned}\langle\, n'_1, \ldots n'_\infty \,|\, n_1, \ldots n_\infty \,\rangle \;\; &= \;\; \frac{1}{N!} \sum_{P,Q} (-1)^{p+q} \int dx_1 \ldots dx_N \\ &\quad \varphi^*_{K'_{P(1)}}(x_1) \ldots \varphi^*_{K'_{P(N)}}(x_N)\; \varphi_{K_{Q(1)}}(x_1) \ldots \varphi_{K_{Q(N)}}(x_N).\end{aligned}$$

The orthogonality of the single-particle orbitals, i.e.,

$$\int dx\, \varphi^*_\alpha(x)\, \varphi_\beta(x) = \delta_{\alpha\beta},$$

implies that $K'_i = K_i \;\; \forall \; i$ if the integral $\langle\, n'_1, \ldots n'_\infty \,|\, n_1, \ldots n_\infty \,\rangle$ is to be non-zero. Moreover, we must also have $Q(i) = P(i)$ for each P, in which case the integral is equal to 1 for all P. One obtains

$$\langle\, n'_1, \ldots n'_\infty \,|\, n_1, \ldots n_\infty \,\rangle = \frac{1}{N!} \sum_P (-1)^{2p}\, \delta_{n_1 n'_1} \ldots \delta_{n_\infty n'_\infty} = \delta_{n_1 n'_1} \ldots \delta_{n_\infty n'_\infty},$$

that is exactly the orthonormalization condition.

We consider now a general single-particle operator

$$F^{(1)} = \sum_{i=1}^N f^{(1)}(x_i)$$

in coordinate representation, for example,

$$F^{(1)} = T = \sum_{i=1}^N \left(\frac{-\hbar^2}{2m} \right) \nabla_i^2$$

or

$$F^{(1)} = V = \sum_{\substack{i=1 \\ \sigma}}^{N} v_\sigma(\vec{r}_i)$$

with $v_+ - v_- = \mu_B B_z(\vec{r})$. Notice that $f^{(1)}(x_i)$ acts only on one variable. The matrix elements of $f^{(1)}$ between single-particle spin-orbitals φ_α and φ_β are given by

$$f^{(1)}_{\alpha\beta} = \int dx\, \varphi^*_\alpha(x)\; f^{(1)}(x)\; \varphi_\beta(x).$$

As usual, α and β label the single particle orbitals while $x \equiv (\vec{r}, \sigma)$ refers to both the electron's position $\vec{r}$ and spin $\sigma =\uparrow$ or $\downarrow$. Accordingly, $\int dx$ stands for $\sum_\sigma \int d^3r$.

We would like to determine the matrix elements of $F^{(1)}$ between many-body N-particle states $|\, n_1, \ldots n_\infty \,\rangle$ with definite occupation numbers n_α:

$$\langle\, n'_1, \ldots n'_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle = \frac{1}{N!} \sum_{P,Q} (-1)^{p+q}$$
$$\times \int dx_1 \ldots dx_N \;\; \varphi^*_{K'_{P(1)}}(x_1) \ldots \varphi^*_{K'_{P(N)}}(x_N) \left[\sum_{i=1}^{N} f^{(1)}(x_i) \right] \varphi_{K_{Q(1)}}(x_1) \ldots \varphi_{K_{Q(N)}}(x_N).$$

Taking into account that $f^{(1)}(x_i)$ acts only on the variable x_i, one concludes that the integral is zero if the list of orbitals $K'_1, \ldots K'_N$ (occupied in $|\, n'_1, \ldots n'_\infty \,\rangle$) and the list of orbitals $K_1, \ldots K_N$ (occupied in $|\, n_1, \ldots n_\infty \,\rangle$) differ by two or more. In other words, the occupation numbers $n'_1, \ldots n'_\infty$ and $n_1, \ldots n_\infty$ must be the same or differ by at most one.

Let us first calculate the diagonal elements. In this case $n_{\alpha'} = n_\alpha \;\; \forall\, \alpha$ and $K_i = K_i{}' \;\; \forall\, i$. For each permutation P we must have

$$P(i) = Q(i) \;\; \forall\, i,$$

since otherwise we would have

$$\int \varphi^*_{K_{P(i)}}(x)\; \varphi_{K_{Q(i)}}(x)\, dx = 0$$

for some i. Taking this into account we obtain

$$\langle\, n_1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle = \frac{1}{N!} \sum_{P} (-1)^{2p}$$
$$\times \int dx_1 \ldots dx_N\; \varphi^*_{K_{P(1)}}(x_1) \ldots \varphi^*_{K_{P(N)}}(x_N) \;\; \left[\sum_{i=1}^{N} f^{(1)}(x_i) \right] \;\; \varphi_{K_{P(1)}}(x_1) \ldots \varphi_{K_{P(N)}}(x_N).$$

The integral is clearly independent of P. Consequently,

$$\begin{aligned}
\langle\, n_1, \quad \ldots \quad n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle = \\
&= \sum_{i=1}^{N} \int dx_1 \ldots dx_N \; \varphi^*_{K_1}(x_1) \ldots \varphi^*_{K_N}(x_N) \; f^{(1)}(x_i) \;\; \varphi_{K_1}(x_1) \ldots \varphi_{K_N}(x_N) \\
&= \sum_{i=1}^{N} \int \varphi^*_{K_i}(x) \; f^{(1)}(x) \; \varphi_{K_i}(x) \, dx \\
&= \sum_{i=1}^{N} f^{(1)}_{K_i K_i} \\
&= \sum_{\alpha=1}^{\infty} n_\alpha \, f^{(1)}_{\alpha\alpha}.
\end{aligned}$$

In conclusion

$$\langle\, n_1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle = \langle\, n_1, \ldots n_\infty \,|\, \sum_\alpha \hat{n}_\alpha \, f^{(1)}_{\alpha\alpha} \,|\, n_1, \ldots n_\infty \,\rangle.$$

In this way the diagonal matrix elements of single-particle operators have been expressed in terms of the occupation numbers and the single-particle integrals $f^{(1)}_{\alpha\alpha}$.

The non-vanishing off-diagonal elements have the form

$$\begin{aligned}
\langle\, n_1, \ldots 1_\alpha, \ldots 0_\beta, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots 0_\alpha, \ldots 1_\beta, \ldots n_\infty \,\rangle = \frac{1}{N!} \sum_{i=1}^{N} \sum_{P,Q} (-1)^{p+q} \\
\times \int dx_1 \ldots dx_N \; \varphi^*_{K'_{P(1)}}(x_1) \ldots \varphi^*_{K'_{P(N)}}(x_N) \; f^{(1)}(x_i) \; \varphi_{K_{Q(1)}}(x_1) \ldots \varphi_{K_{Q(N)}}(x_N),
\end{aligned}$$

where we assumed $\alpha < \beta$. The integrals vanish unless $K_{Q(i)} = \beta$ and $K'_{P(i)} = \alpha$ and $K_{Q(j)} = K'_{P(j)} \;\forall\; j \neq i$. For a given β, there are $(N-1)!$ permutations Q such that $Q(i) = m$ with $K_m = \beta$. Moreover, for each Q, there is only one permutation P satisfying $K'_{P(i)} = \alpha$ and $K'_{P(j)} = K_{Q(j)}$, since all the orbitals K'_i are different $(i = 1, \ldots N)$. Under these conditions the sum over P can be removed and the integrals are then independent of Q and i, provided that $K_{Q(i)} = \beta$. They are equal to

$$\int dx \; \varphi^*_\alpha(x) \; f^{(1)}(x) \; \varphi_\beta(x) = f^{(1)}_{\alpha\beta}.$$

In order to determine the sign $(-1)^{p+q}$ one considers first the identity permutation $Q(i) = i$. The single-particle states are

$$\begin{array}{lccccccccc}
0_\alpha \ldots 1_\beta \to \vec{K} = & K_1, \ldots & K_{l-1}, & K_l, & K_{l+1}, & K_{l+2} \ldots & \boxed{K_m = \beta,} & K_{m+1}, \ldots & K_N \\
1_\alpha \ldots 0_\beta \to \vec{K}' = & \underbrace{K_1}_{=K'_1}, \ldots & K_{l-1}, & \underbrace{\boxed{K'_l = \alpha,}}_{=K'_l} & \underbrace{K_l}_{=K'_{l+1}}, & \underbrace{K_{l+1}}_{=K'_{l+2}}, \ldots & \underbrace{K_{m-1}}_{=K'_m}, & \underbrace{K_{m+1}}_{=K'_{m+1}}, \ldots & \underbrace{K_N}_{=K'_N}.
\end{array}$$

The permutation $P = P_0$ for which the integral is non-vanishing is therefore

$$P_0 = \begin{pmatrix} 1 & \dots & l-1 & l & l+1 & \dots & m-1 & m & m+1 & \dots & N \\ 1 & \dots & l-1 & l+1 & l+2 & \dots & m & l & m+1 & \dots & N \end{pmatrix}.$$

The order $O(P_0)$ of P_0 is $p_0 = \sum_{\alpha<\gamma<\beta} n_\gamma$. Since $Q(i) = i$, the order of Q is $q = 0$ and $p + q = p_0 = \sum_{\alpha<\gamma<\beta} n_\gamma$. One may already note here that the same result is obtained for $\beta < \alpha$.

For an arbitrary perturbation Q such that $Q(i) = m$ (knowing that $K_m = \beta$), we must choose

$$P = P_0 Q$$

since

$$\begin{cases} K'_{P(j)} = K'_{P_0(Q(j))} = K_{Q(j)} \quad \forall\, j \neq i \quad (Q(j) \neq m) \\ K'_{P(i)} = K'_{P_0(Q(i))} = K'_{P_0(m)} = K'_l = \alpha \quad (Q(i) = m). \end{cases}$$

Finally, the order of P is $p = O(P_0Q) = O(P_0) + O(Q) = p_0 + q$ and thus $p+q = p_0 + 2q$. This implies $(-1)^{p+q} = (-1)^{p_0}$ so that

$$(-1)^{p+q} = (-1)^{\sum_{\alpha<\gamma<\beta} n_\gamma}$$

is the same for all Q. We may now write

$$\langle\, n_1, \dots 1_\alpha, \dots 0_\beta, \dots n_\infty \,|\, F^{(1)} \,|\, n_1, \dots 0_\alpha, \dots 1_\beta, \dots n_\infty \,\rangle = \\ = (-1)^{\sum_{\alpha<\gamma<\beta} n_\gamma} f^{(1)}_{\alpha\beta} \, \frac{1}{N!} \sum_{i=1}^{N} \underbrace{\sum_{Q} \delta_{m,Q(i)}}_{(N-1)!}.$$

Taking into account that there are $(N-1)!$ permutations Q yielding $Q(i) = m$ ($K_m = \beta$), N terms in $\sum_{i=1}^{N}$, and only one P yielding a non-vanishing integral for each Q, we finally have

$$\langle\, n_1, \dots 1_\alpha, \dots 0_\beta, \dots n_\infty \,|\, F^{(1)} \,|\, n_1, \dots 0_\alpha, \dots 1_\beta, \dots n_\infty \,\rangle = (-1)^{\Sigma(\alpha+1,\beta-1)} f^{(1)}_{\alpha\beta},$$

where as usual $\Sigma(\alpha+1, \beta-1) = \sum_{\alpha<\gamma<\beta} n_\gamma$. This can be written in terms of creation and destruction operators as

$$\boxed{\langle\, n_1, \dots 1_\alpha, \dots 0_\beta, \dots n_\infty \,|\, f^{(1)}_{\alpha\beta}\, \hat{a}^\dagger_\alpha\, \hat{a}_\beta \,|\, n_1, \dots 0_\alpha, \dots 1_\beta, \dots n_\infty \,\rangle.}$$

The result is the same for $\beta < \alpha$ since

$$\langle\, n_1, \dots 0_\beta, \dots 1_\alpha, \dots n_\infty \,|\, F^{(1)} \,|\, n_1, \dots 1_\beta, \dots 0_\alpha, \dots n_\infty \,\rangle = \\ = \langle\, n_1, \dots 1_\beta, \dots 0_\alpha, \dots n_\infty \,|\, F^{(1)} \,|\, n_1, \dots 0_\beta, \dots 1_\alpha, \dots n_\infty \,\rangle^* \\ = (-1)^{\Sigma(\beta+1,\alpha-1)}\, f^{(1)*}_{\beta\alpha} = (-1)^{\Sigma(\beta+1,\alpha-1)}\, f^{(1)}_{\alpha\beta}.$$

We may then write in the occupation number representation

$$\boxed{\hat{F}^{(1)} = \sum_{\alpha\beta} f^{(1)}_{\alpha\beta}\, \hat{a}^\dagger_\alpha\, \hat{a}_\beta.}$$

Alternative derivation of the sign in one-particle matrix elements

Our starting point is

$$\langle\, n_1, \ldots 1_\alpha, \ldots 0_\beta, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots 0_\alpha, \ldots 1_\beta, \ldots n_\infty \,\rangle = \frac{1}{N!} \sum_{i=1}^{N} \sum_{P,Q} (-1)^{p+q}$$

$$\times \int dx_1 \ldots dx_N\; \varphi^*_{K_{P(1)}}(x_1) \ldots \varphi^*_{K_{P(N)}}(x_N)\; f^{(1)}(x_i)\; \varphi_{K_{Q(1)}}(x_1) \ldots \varphi_{K_{Q(N)}}(x_N).$$

Since the sums run over all permutations we may always set $x_i = x_1$ in $f^{(1)}(x_i)$ and replace the $\sum_{i=1}^{N}$ by a factor N. The matrix element we are looking for is given by

$$\frac{N}{N!} \sum_{P,Q} (-1)^{p+q} \int dx_1 \ldots dx_N\; \varphi^*_{K_{P(1)}}(x_1) \ldots \varphi^*_{K_{P(N)}}(x_N)\; f^{(1)}(x_1)\; \varphi_{K_{Q(1)}}(x_1) \ldots \varphi_{K_{Q(N)}}.$$

It is clear that in order to obtain a non-vanishing result we must have $Q(1) = m$ where $K_m = \beta$, and $P(1) = l$ where $K_l = \alpha$. All other $(N-1)!$ permutations of the other variables must be the same in $Q(j)$ and $P(j)$ since otherwise the integral vanishes. The number of transpositions $q + p$ can be calculated as follows. The permutation Q brings

$$1 \xrightarrow{Q} m \quad \text{and then} \quad m \xrightarrow{Q^{-1}} 1.$$

In both Q and Q^{-1} the number of transpositions is q. Concerning P, we know that it brings

$$1 \xrightarrow{P} l \quad \text{and then} \quad PQ^{-1} \quad \text{brings} \quad m \to l.$$

The number of transpositions in PQ^{-1} is $p + q = \sum_{\alpha<\gamma<\beta} n_\gamma$, and therefore $(-1)^{p+q} = (-1)^{\sum_{\alpha<\gamma<\beta} n_\gamma}$. The contribution of the permutation of the other $N-1$ indices cancels out in PQ^{-1}.

Since $(-1)^{p+q}$ coincides with the sign for bringing $m \to l$ (or $l \to m$) we may conclude at once that it is the same sign as the one involved in

$$\langle \ldots 1_\alpha, \ldots 0_\beta, \ldots \,|\, \hat{a}^\dagger_\alpha\, \hat{a}_\beta \,|\, \ldots 0_\alpha, \ldots 1_\beta, \ldots \rangle.$$

4.2.4 Matrix elements of two-particle operators

Consider a general two-particle operator given by

$$F^{(2)} = \sum_{i<j} f^{(2)}(x_i, x_j) = \frac{1}{2} \sum_{i \neq j} f^{(2)}(x_i, x_j),$$

for example, the Coulomb interaction between electrons

$$W = \sum_{i<j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|}.$$

In analogy with the case of single-particle operators we define the matrix elements between pairs of single-particle states as

$$f^{(2)}_{\alpha\beta\gamma\delta} = \int dx_1\, dx_2 \;\; \varphi^*_\alpha(x_1)\; \varphi^*_\beta(x_2)\; f^{(2)}(x_1, x_2)\; \varphi_\gamma(x_1)\; \varphi_\delta(x_2).$$

Note the precise order of the orbital indices and of the integration variables. It is easy to see that $f^{(2)}_{\alpha\beta\gamma\delta} = f^{(2)}_{\beta\alpha\delta\gamma}$ and $f^{(2)}_{\alpha\beta\gamma\delta} = (f^{(2)}_{\gamma\delta\alpha\beta})^*$, since $f^{(2)}(x_1, x_2) = f^{(2)}(x_2, x_1)$. However, notice that the order of the orbital indices is significant. For example, $f^{(2)}_{\alpha\beta\alpha\beta} = \int d^3r_1\, d^3r_2 \;\; |\varphi_\alpha(\vec{r}_1)|^2\; (\mathrm{e}^2/|\vec{r}_1 - \vec{r}_2|)\; |\varphi_\beta(\vec{r}_2)|^2$ is the so-called direct integral corresponding to the electrostatic Coulomb repulsion between the electronic densities $|\varphi_\alpha(\vec{r}_1)|^2$ and $|\varphi_\beta(\vec{r}_1)|^2$, whereas $f^{(2)}_{\alpha\beta\beta\alpha} = \int d^3r_1\, d^3r_2 \;\; \varphi^*_\alpha(\vec{r}_1)\; \varphi_\beta(\vec{r}_1)\; (\mathrm{e}^2/|\vec{r}_1 - \vec{r}_2|)\; \varphi^*_\beta(\vec{r}_2)\; \varphi_\alpha(\vec{r}_2)$ is the so-called exchange integral which is in genral smaller and would even vanish if the orbitals α and β occupy disjoint regions is space.

We look for the expression of the operator $F^{(2)}$ in occupation number representation, also known as second quantization, in terms of $\hat{a}^\dagger_\alpha,\; \hat{a}_\alpha$ and $f^{(2)}_{\alpha\beta\gamma\delta}$. For this purpose we need to determine the matrix elements between many-body states $|\, n_1, \ldots n_\infty \,\rangle$. One proceeds, as for $F^{(1)}$, starting from

$$\langle\, n'_1, \ldots n'_\infty \,|\, F^{(2)} \,|\, n_1, \ldots n_\infty \,\rangle = \frac{1}{N!} \sum_{P,Q} (-1)^{p+q}$$

$$\times \int dx_1 \ldots dx_N \;\; \varphi^*_{K'_{P(1)}}(x_1) \ldots \varphi^*_{K'_{P(N)}}(x_N) \left[\sum_{i<j} f^{(2)}(x_i, x_j) \right] \varphi_{K_{Q(1)}}(x_1) \ldots \varphi_{K_{Q(N)}}(x_N).$$

These matrix elements vanish unless the occupations $\{n'_\alpha\}$ and $\{n_\alpha\}$ differ in at most 2 states. The situation is similar to the case of a one-particle operator. The integrals involving $f^{(2)}(x_i, x_j)$ are independent of the pair of coordinates (x_i, x_j) since we may always re-order the coordinates on both kets and bras (i.e., on both sets of orbitals K_i and K'_i) to bring x_i and x_j to the first and second places. We may then fix (x_i, x_j) equal to (x_1, x_2) and multiply by the number of pairs $N(N-1)/2$.

Let us consider the most interesting case where the 4 involved single-particle states are different and let us set for simplicity $\alpha < \beta < \gamma < \delta$. We have then to calculate

$$\langle\, n_1, \ldots 1_\alpha, \ldots 1_\beta, \ldots 0_\gamma, \ldots 0_\delta, \ldots n_\infty \,|\, F^{(2)} \,|\, n_1, \ldots 0_\alpha, \ldots 0_\beta, \ldots 1_\gamma, \ldots 1_\delta, \ldots n_\infty \,\rangle =$$
$$= \frac{N(N-1)}{2N!} \sum_{P,Q} (-1)^{p+q} \int dx_1 \ldots dx_N$$
$$\times \varphi^*_{K'_{P(1)}}(x_1) \varphi^*_{K'_{P(2)}}(x_2) \ldots \varphi_{K'_{P(N)}}(x_N)\; f^{(2)}(x_1, x_2) \;\; \varphi_{K_{Q(1)}}(x_1) \varphi_{K_{Q(2)}}(x_2) \ldots \varphi_{K_{Q(N)}}(x_N).$$

In order to get a non-vanishing result we must have one of the following four possibilities:

i) $K_{Q(1)} = \gamma$, $K_{Q(2)} = \delta$ and $K'_{P(1)} = \alpha$, $K'_{P(2)} = \beta$

ii) $K_{Q(1)} = \gamma$, $K_{Q(2)} = \delta$ and $K'_{P(1)} = \beta$, $K'_{P(2)} = \alpha$

iii) $K_{Q(1)} = \delta$, $K_{Q(2)} = \gamma$ and $K'_{P(1)} = \alpha$, $K'_{P(2)} = \beta$

iv) $K_{Q(1)} = \delta$, $K_{Q(2)} = \gamma$ and $K'_{P(1)} = \beta$, $K'_{P(2)} = \alpha$

For $i, j \geq 3$ we must always have $K_{Q(i)} = K'_{P(i)}$. In all cases there are $(N-2)!$ permutations Q that satisfy the conditions $(K_{Q(1)}, K_{Q(2)}) = (\gamma, \delta)$ or (δ, γ). For a given Q the permutation P is entirely defined except for the order of the orbitals α and β, i.e., $(K'_{P(1)}, K'_{P(2)}) = (\alpha, \beta)$ or (β, α).

The situations can be summarized in the following table:

	P(1)	P(2)	Q(1)	Q(2)
I	α	β	γ	δ
II	β	α	γ	δ
III	α	β	δ	γ
IV	β	α	δ	γ

First case: On the one side the permutation Q brings the index of orbital γ to position 1 and the index of δ to position 2:

$$\gamma \xrightarrow{Q} 1$$

$$\delta \longrightarrow 2.$$

The order of Q is the same as the order of Q^{-1} which is given by the number of transpositions involved in

$$1 \xrightarrow{Q^{-1}} \gamma$$

$$2 \longrightarrow \delta.$$

On the other side the permutation P must bring

$$\alpha \xrightarrow{P} 1$$

$$\beta \longrightarrow 2.$$

Thus, $p+q$ is the number of transpositions involved in rearranging the indices of the orbitals such that

$$\alpha \rightarrow \gamma$$
$$\beta \rightarrow \delta.$$

Second case: A similar analysis shows that in this case $p+q$ is the number of transpositions involved in bringing

$$\alpha \rightarrow \delta$$
$$\beta \rightarrow \gamma.$$

This is the same as in case 1 plus one, as can be easily seen by bringing $\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \rightarrow \begin{pmatrix}\gamma\\ \delta\end{pmatrix}$ like in case 1 and then transposing γ and δ.

It is important to note that the details of the permutations among the other indices $i, j \geq 3$ corresponding to the single-particle states that do not change play no role since they appear both in Q and in the permutation P that matches the indices. As already observed there are $(N-2)!$ such permutations. The integral is independent of this reordering among the indices corresponding to identical single-particle states.

We recover then 2 terms given by

$$\boxed{\frac{1}{2}(-1)^{p+q}\left(f^{(2)}_{\alpha\beta\gamma\delta} - f^{(2)}_{\beta\alpha\gamma\delta}\right)}$$

which come from

$$\begin{array}{ll} K_{Q(1)} = \gamma & K'_{P(1)} = \alpha \\ K_{Q(2)} = \delta & K'_{P(2)} = \beta \end{array} \quad \text{and} \quad \begin{array}{ll} K'_{P(1)} = \beta \\ K'_{P(2)} = \alpha \end{array} \quad \text{or} \quad .$$

Here the sign $(-1)^{p+q}$ is the one corresponding to $\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \rightarrow \begin{pmatrix}\gamma\\ \delta\end{pmatrix}$.

Third and fourth cases: In the third case we have

$$\begin{array}{ll} K_{Q(1)} = \delta & K'_{P(1)} = \alpha \\ K_{Q(2)} = \gamma & K'_{P(2)} = \beta. \end{array} \quad \text{and}$$

The sign of this contribution is the one corresponding to $\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \rightarrow \begin{pmatrix}\delta\\ \gamma\end{pmatrix}$. This is given by $(-1)^{p+q+1}$, where $p+q$ refers to the number of transpositions involved in $\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \rightarrow \begin{pmatrix}\gamma\\ \delta\end{pmatrix}$.

In the fourth case we have

$$\begin{array}{ll} K_{Q(1)} = \delta & K'_{P(1)} = \beta \\ K_{Q(2)} = \gamma & K'_{P(2)} = \alpha. \end{array} \quad \text{and}$$

This has the sign corresponding to $\binom{\beta}{\alpha} \to \binom{\delta}{\gamma}$ which is $(-1)^{p+q}$. Putting the contributions of the third and fourth cases together we obtain

$$\boxed{\frac{1}{2}\,(-1)^{p+q}\,\left(f^{(2)}_{\beta\alpha\delta\gamma} - f^{(2)}_{\alpha\beta\delta\gamma}\right).}$$

Finally, summing all four cases and using that $f^{(2)}_{\alpha\beta\gamma\delta} = f^{(2)}_{\beta\alpha\delta\gamma}$ we obtain

$$\boxed{\begin{aligned}\langle\, n_1, \ldots 1_\alpha, \quad \ldots \quad 1_\beta, \ldots 0_\gamma, \ldots 0_\delta, \ldots n_\infty \,|\, F^{(2)} \,|\, n_1, \ldots 0_\alpha, \ldots 0_\beta, \ldots 1_\delta, \ldots 1_\gamma, \ldots n_\infty \,\rangle \\ = \quad (-1)^{p+q}\,(f_{\alpha\beta\gamma\delta} - f_{\alpha\beta\delta\gamma})\,,\end{aligned}}$$

where $(-1)^{p+q}$ is the sign of the permutation that brings the pair $\binom{\alpha}{\beta}$ to $\binom{\gamma}{\delta}$, i.e., $\alpha \to \gamma$ and $\beta \to \delta$.

Calculating the sign $(-1)^{p+q}$ in terms of n_α is possible though boring. However, one may convince oneself rapidly that this is the same sign as the one involved in changing the occupation numbers by applying creation and destruction operators as follows:

$$\begin{aligned}\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\delta\,\hat{a}_\gamma \quad | \quad & n_1, \ldots 0_\alpha, \ldots 0_\beta, \ldots 1_\gamma, \ldots 1_\delta, \ldots n_\infty \,\rangle = \\ = \quad & \hat{a}^\dagger_\alpha\,\hat{a}_\gamma\,\hat{a}^\dagger_\beta\,\hat{a}_\delta \,|\, n_1, \ldots 0_\alpha, \ldots 0_\beta, \ldots 1_\gamma, \ldots 1_\delta, \ldots n_\infty \,\rangle = \\ = \quad & \hat{a}^\dagger_\alpha\,\hat{a}_\gamma\,\hat{a}^\dagger_\beta\,\hat{a}_\delta\,(\hat{a}^\dagger_1)^{n_1}\ \ldots \underset{\alpha}{\sqcup} \ldots \underset{\beta}{\sqcup} \ldots\, \hat{a}^\dagger_\gamma \ldots \hat{a}^\dagger_\delta \ldots (\hat{a}^\dagger_\infty)^{n_\infty}\,|\,0\,\rangle = \\ = \quad & (\hat{a}^\dagger_1)^{n_1} \ldots \underset{\alpha}{\sqcup} \ldots \underset{\beta}{\sqcup} \ldots\, \hat{a}^\dagger_{\underset{\gamma}{\alpha}} \ldots \hat{a}^\dagger_{\underset{\delta}{\beta}} \ldots (\hat{a}^\dagger_\infty)^{n_\infty}\,|\,0\,\rangle = \\ = \quad & (-1)^{p+q}\,|\, n_1, \ldots 1_\alpha, \ldots 1_\beta, \ldots 0_\gamma, \ldots 0_\delta, \ldots n_\infty \,\rangle.\end{aligned}$$

Using this property and taking into account all possible pairs we can write the two particle operator as

$$\begin{aligned}\hat{F}^{(2)} \quad &= \quad \sum_{\alpha<\beta}\sum_{\gamma<\delta}\left(f^{(2)}_{\alpha\beta\gamma\delta} - f^{(2)}_{\alpha\beta\delta\gamma}\right)\,\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\delta\,\hat{a}_\gamma \\ &= \quad \sum_{\alpha<\beta}\left(\sum_{\gamma<\delta} f^{(2)}_{\alpha\beta\gamma\delta}\,\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\delta\,\hat{a}_\gamma + \sum_{\gamma<\delta} f^{(2)}_{\alpha\beta\delta\gamma}\,\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\gamma\,\hat{a}_\delta\right).\end{aligned}$$

The two terms in the previous expression look the same if we interchange the indices γ and δ in one of them, for instance in the second one. After this change of notation one of the sums runs over $\gamma < \delta$ and the other over $\delta < \gamma$. Noting that $\gamma = \delta$ (or $\alpha = \beta$) does not contribute, we finally have

$$\hat{F}^{(2)} = \sum_{\alpha<\beta}\sum_{\gamma,\delta} f^{(2)}_{\alpha\beta\gamma\delta}\,\hat{a}^\dagger_\alpha\,\hat{a}^\dagger_\beta\,\hat{a}_\delta\,\hat{a}_\gamma$$

or equivalently

$$\boxed{\hat{F}^{(2)} = \frac{1}{2} \sum_{\substack{\alpha\,\beta \\ \gamma\delta}} f^{(2)}_{\alpha\beta\gamma\delta}\, \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\delta\, \hat{a}_\gamma.}$$

In the last step we have used that $\hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\delta\, \hat{a}_\gamma = \hat{a}^\dagger_\beta\, \hat{a}^\dagger_\alpha\, \hat{a}_\gamma\, \hat{a}_\delta$ and $f^{(2)}_{\alpha\beta\gamma\delta} = f^{(2)}_{\beta\alpha\delta\gamma}$.

This *second quantization* expression for the two particle operator $F^{(2)}$ is in all points equivalent to the usual first quantization expression in coordinate, momentum or any other representation, since it has the same matrix elements in a complete basis set. Among other advantages, notice that the present second quantization form is valid for any number of fermions.

4.2.5 Fermion fields

We search for an expression for $\hat{F}^{(1)}$ and $\hat{F}^{(2)}$ that is explicitly independent of the choice of the single-particle basis set. Starting from the expression for $F^{(1)}$ we obtain

$$\begin{aligned}
\hat{F}^{(1)} &= \sum_{\alpha\beta} f^{(1)}_{\alpha\beta}\, \hat{a}^\dagger_\alpha\, \hat{a}_\beta \\
&= \sum_{\alpha\beta} \left[\int dx\, \varphi^*_\alpha(x)\, f^{(1)}(x)\, \varphi_\beta(x) \right] \hat{a}^\dagger_\alpha\, \hat{a}_\beta \\
&= \int dx \left(\sum_\alpha \varphi^*_\alpha(x)\, \hat{a}^\dagger_\alpha \right) f^{(1)}(x) \left(\sum_\beta \varphi_\beta(x)\, \hat{a}_\beta \right).
\end{aligned}$$

This can be written in a compact form as

$$\boxed{\hat{F}^{(1)} = \int dx\, \hat{\Psi}^\dagger(x)\, f^{(1)}(x)\, \hat{\Psi}(x),}$$

where $\hat{\Psi}^\dagger(x) = \sum_\alpha\, \varphi^*_\alpha(x)\, \hat{a}^\dagger_\alpha$ is the field creation operator and $\hat{\Psi}(x) = \sum_\alpha\, \varphi_\alpha(x)\, \hat{a}_\alpha$ is the field annihilation operator.

A similar calculation shows that the two-particle operators are given by

$$\boxed{\hat{F}^{(2)} = \frac{1}{2} \int dx\, dx'\, \hat{\Psi}^\dagger(x)\, \hat{\Psi}^\dagger(x')\, f^{(2)}(x, x')\, \hat{\Psi}(x')\, \hat{\Psi}(x).}$$

While x and x' are dummy integration variables and $f^{(2)}(x, x') = f^{(2)}(x', x)$ due to the identity of particles, *the order of* $\hat{\Psi}^\dagger(x)$ *and* $\hat{\Psi}^\dagger(x')$ *with respect to* $\hat{\Psi}(x)$ *and* $\hat{\Psi}(x')$ *matters.* This is important in order to match the order of the operators $\hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\delta\, \hat{a}_\gamma$ and to ensure the hermiticity of $\hat{F}^{(2)}$. The proof is left as an exercise.

It is often useful to replace x by $(\vec{r},\sigma)$ and $\hat{\Psi}(\vec{r},\sigma) = \hat{\Psi}(x)$ by $\Psi_\sigma(\vec{r})$. Assuming for simplicity that $f^{(1)}$ has no matrix elements between states having different spins and that $f^{(2)}$ is independent of σ we obtain

$$\hat{F}^{(1)} = \sum_\sigma \int \hat{\Psi}^\dagger_\sigma(\vec{r})\; f^{(1)}_\sigma(\vec{r})\; \hat{\Psi}_\sigma(\vec{r})\, d^3r$$

and

$$\hat{F}^{(2)} = \frac{1}{2}\sum_{\sigma\sigma'} \int d^3r\, d^3r'\; \hat{\Psi}^\dagger_\sigma(\vec{r})\, \hat{\Psi}^\dagger_{\sigma'}(\vec{r}')\; f^{(2)}(\vec{r},\vec{r}')\; \hat{\Psi}_{\sigma'}(\vec{r}')\, \hat{\Psi}_\sigma(\vec{r}).$$

The second quantization formalism can be extended straightforwardly to three or more particle interactions.

4.2.6 Analogies between 1st and 2nd quantization: Second quantization rules

A number of analogies can be pointed out:

i) In first quantization we consider kets of the form

$$|\,\Psi\,\rangle = \sum_\alpha a_\alpha\, |\,\alpha\,\rangle \quad \text{with} \quad \sum_\alpha |a_\alpha|^2 = 1 \quad \text{and} \quad \langle\,\alpha\,|\,\beta\,\rangle = \delta_{\alpha\beta}.$$

The average of a single particle operator f is given by

$$\langle\, f\,\rangle = \langle\,\Psi\,|\,f\,|\,\Psi\,\rangle = \sum_{\alpha\beta} a^*_\alpha\, a_\beta\, f_{\alpha\beta},$$

where $f_{\alpha\beta} = \langle\,\alpha\,|\,f\,|\,\beta\,\rangle$. This is analogous to the expression for the operator $\hat{F}^{(1)}$ in 2nd quantization:

$$\hat{F}^{(1)} = \sum_{\alpha\beta} f^{(1)}_{\alpha\beta}\, \hat{a}^\dagger_\alpha\, \hat{a}_\beta.$$

Notice that the matrix elements are the same. Just the coefficients have changed from numbers to operators:

$$a^*_\alpha \to \hat{a}^\dagger_\alpha$$
$$a_\alpha \to \hat{a}_\alpha.$$

ii) The expansion of the single-particle wave function $\Psi(x)$ in first quantization in terms of a complete set of orbitals $\varphi_\alpha(x)$ is given by

$$\Psi(x) = \sum_\alpha a_\alpha\, \varphi_\alpha(x),$$

while the second quantization field operator is given by $\hat{\Psi}(x) = \sum_\alpha \varphi_\alpha(x)\, \hat{a}_\alpha$.

iii) In terms of the wave function $\Psi(x)$, the mean value of a single-particle operator in 1st quantization is

$$\langle f^{(1)} \rangle = \int \Psi^*(x)\ f^{(1)}(x)\ \Psi(x)\, dx,$$

while the second quantization form of the operator $F^{(1)} = \sum_i f^{(1)}(x_i)$ is

$$\hat{F}^{(1)} = \int \hat{\Psi}^\dagger(x)\, f^{(1)}(x)\ \hat{\Psi}(x)\, dx.$$

iv) For the two-particle operators the situation is more delicate. The analogy works fine for bosons assuming the two bosons occupy the same state $\Psi(x)$ (one-particle state). For fermions we cannot do that, but there is still a striking analogy between

$$\langle f^{(2)} \rangle = \frac{1}{2} \int dx\, \Psi^*(x) \left(\int dx'\, \Psi^*(x')\ f^{(2)}(x, x')\ \Psi(x') \right) \Psi(x)$$

and the second quantization form of $\hat{F}^{(2)}$.

A second quantization "rule" can then be formulated:

1. Express the mean value of an operator in terms of a single-particle wave function $\Psi(x)$. In the case of two-particle operators do the average of each variable at a time. The single particle state $\Psi(x)$ must be normalized to 1.
2. Then the second quantization operator corresponding to this observable is obtained by replacing the wave function $\Psi(x)$ by the field operator $\hat{\Psi}(x)$ and $\Psi^*(x)$ by $\hat{\Psi}^\dagger(x)$. The hermitian operators $\hat{\Psi}^\dagger(x)$ and the conjugate wave functions $\Psi^*(x)$ are to be written at the left of $\hat{\Psi}(x)$.

These analogies explain the origin of the name "second quantization" ($a_\alpha \to \hat{a}_\alpha$). Nevertheless, it is clear that it is more approriate to regard this formalism as a change of representation, namely, to the occupation-number representation of many-body states.

4.2.7 Properties of $\hat{\Psi}(x)$ and $\hat{\Psi}(x)$

It is clear that $\hat{\Psi}_\sigma(\vec{r})$ decreases by one the number of electrons with spin σ, since it is a linear combination of destruction operators. Similarly, $\hat{\Psi}^\dagger_\sigma(\vec{r})$ increases the number of σ electrons by one.

A particularly important property of the field operators $\hat{\Psi}^\dagger_\sigma(\vec{r})$ and $\hat{\Psi}_\sigma(\vec{r})$ is that they are independent of the particular choice of the single-particle basis. In order to show this we must first of all clarify how the creation and destruction operators transform when the single-particle states are changed.

Let $\hat{a}^\dagger_\alpha$ be the creation operator for a single-particle state $\varphi_\alpha(x)$, i.e.,

$$\langle x \,|\, \hat{a}^\dagger_\alpha \,|\, 0 \rangle = \varphi_\alpha(x),$$

and let

$$\chi_\beta(x) = \sum_\alpha U_{\beta\alpha}\,\varphi_\alpha(x)$$

be the states of another basis set with $U^\dagger U = U\ U^\dagger = \mathbb{1}$. If the creation operator associated to χ_β is denoted by $\hat{c}^\dagger_\beta$, we have

$$\langle\, x\,|\,\hat{c}^\dagger_\beta\,|\,0\,\rangle = \chi_\beta(x) = \sum_\alpha U_{\beta\alpha}\,\varphi_\alpha(x) = \sum_\alpha U_{\beta\alpha}\,\langle\, x\,|\,\hat{a}^\dagger_\alpha\,|\,0\,\rangle = \langle\, x\,|\sum_\alpha U_{\beta\alpha}\,\hat{a}^\dagger_\alpha\,|\,0\,\rangle.$$

Consequently, $\hat{c}^\dagger_\beta = \sum_\alpha U_{\beta\alpha}\,\hat{a}^\dagger_\alpha$ transforms like χ_β with respect to basis changes. If we now construct the field operators in terms of the new basis $\chi_\beta(x)$ and the corresponding creation operators we obtain

$$\begin{aligned}
\hat{\Psi}^\dagger(x) &= \sum_\beta \chi^*_\beta(x)\,\hat{c}^\dagger_\beta = \sum_\beta \left(\sum_\alpha U^*_{\beta\alpha}\,\varphi^*_\alpha(x)\right)\left(\sum_{\alpha'} U_{\beta\alpha'}\,\hat{a}^\dagger_{\alpha'}\right) \\
&= \sum_{\alpha\alpha'} \varphi^*_\alpha(x)\,\hat{a}^\dagger_{\alpha'}\,\underbrace{\sum_\beta U^\dagger_{\alpha\beta}\,U_{\beta\alpha'}}_{\delta_{\alpha\alpha'}} \\
&= \sum_\alpha \varphi^*_\alpha(x)\,\hat{a}^\dagger_\alpha.
\end{aligned}$$

Therefore, $\hat{\Psi}^\dagger(x)$ and $\hat{\Psi}(x)$ are independent of the choice of the single-particle basis.

Anticommutation rules for $\hat{\Psi}(x)$ and $\hat{\Psi}^\dagger(x)$

Let us first consider the annihilation operators

$$\{\hat{\Psi}(x), \hat{\Psi}(x')\} = \sum_{\alpha\beta}\ \varphi_\alpha(x)\ \{\hat{a}_\alpha, \hat{a}_\beta\}\ \varphi_\beta(x').$$

Therefore

$$\{\hat{a}_\alpha, \hat{a}_\beta\} = 0\ \ \forall\,\alpha,\beta \quad\Leftrightarrow\quad \boxed{\{\hat{\Psi}(x), \hat{\Psi}(x')\} = 0\ \ \forall\,x,x'.}$$

To prove the converse multiply $\{\hat{\Psi}(x), \hat{\Psi}(x')\}$ by $\varphi^*_{\beta'}(x')$ and $\varphi^*_{\alpha'}(x)$ and integrate with respect to x and x'. Using the orthonormality of the single-particle orbitals $\varphi_\alpha(x)$ one obtains $\{\hat{a}_{\alpha'}, \hat{a}_{\beta'}\} = 0\ \ \forall\,\alpha',\beta'$.

Obviously, for the creation operators we have

$$\{\hat{\Psi}^\dagger(x), \hat{\Psi}^\dagger(x')\} = 0\ \ \forall\,x,x'.$$

Finally, let us consider the anticommutator

$$\{\hat{\Psi}(x), \hat{\Psi}^\dagger(x')\} = \sum_{\alpha\beta}\ \varphi_\alpha(x)\,\{\hat{a}_\alpha, \hat{a}^\dagger_\beta\}\,\varphi^*_\beta(x'). \tag{4.5}$$

Since $\{\hat{a}_\alpha, \hat{a}_\beta^\dagger\} = \delta_{\alpha\beta}$, we conclude that

$$\{\hat{\Psi}(x), \hat{\Psi}^\dagger(x')\} = \delta(x - x'). \tag{4.6}$$

Recalling that $\delta(x - x') = \delta_{\sigma\sigma'}\ \delta(\vec{r} - \vec{r}')$ we may write the anticommutation relation in the more usual form

$$\{\hat{\Psi}_\sigma(\vec{r}), \hat{\Psi}_{\sigma'}^\dagger(\vec{r}')\} = \delta_{\sigma\sigma'}\ \delta(\vec{r} - \vec{r}').$$

It is interesting to note that the converse is also true, namely,

$$\{\hat{\Psi}(x), \hat{\Psi}^\dagger(x')\} = \delta(x - x')\ \ \forall x, x' \quad \Rightarrow \quad \{\hat{a}_\alpha, \hat{a}_\beta^\dagger\} = \delta_{\alpha\beta}.$$

This can be shown straightforwardly by noting that $\hat{a}_\alpha = \int \varphi_\alpha^*(x)\, \hat{\Psi}(x)\, dx$. Applying the bilinearity of the anticommutator and using Eq. (4.6) we have

$$\{\hat{a}_\alpha, \hat{a}_\beta^\dagger\} = \int dx\, dx'\ \varphi_\alpha^*(x)\, \{\hat{\Psi}(x), \hat{\Psi}^\dagger(x')\}\ \varphi_\beta(x') = \int dx\, dx'\ \varphi_\alpha^*(x)\ \delta(x - x')\ \varphi_\beta(x') = \delta_{\alpha\beta}.$$

In conclusion, the formulation of the anticommutation rules between the creation and destruction operator $\hat{a}_\alpha$ and $\hat{a}_\beta^\dagger$ associated to a single-particle complete basis is equivalent to the anticommutation rules for the field operators at all points in space.

4.2.8 Interpretation of the field operators $\hat{\Psi}^\dagger(x)$ and $\hat{\Psi}(x)$

We already know that $\hat{\Psi}^\dagger(x')$ is a linear combination of creation operators $\hat{\Psi}^\dagger(x') = \sum_\alpha\ \varphi^*(x')\, \hat{a}_\alpha^\dagger$. Therefore, it creates an electron in a particular state in coordinate representation. Moreover, this state is independent of the choice of the basis! Let us find out how it looks like:

$$\begin{aligned}
\langle\, x \,|\, \hat{\Psi}^\dagger(x') \,|\, 0 \,\rangle &= \langle\, x \,|\, \sum_\alpha\ \varphi_\alpha^*(x')\, \hat{a}_\alpha^\dagger \,|\, 0 \,\rangle \\
&= \sum_\alpha\ \varphi_\alpha^*(x')\, \langle\, x \,|\, \hat{a}_\alpha^\dagger \,|\, 0 \,\rangle \\
&= \sum_\alpha\ \varphi_\alpha^*(x')\, \varphi_\alpha(x) = \delta(x - x').
\end{aligned}$$

This implies that $\hat{\Psi}^\dagger(x')$ creates an electron in a state with definite coordinate x' (i.e., $|\, x' \,\rangle$). Equivalently, $\Psi_\sigma^\dagger(\vec{r})$ creates an electron of spin σ at point $\vec{r}$ (i.e., $|\, \vec{r}\,\sigma \,\rangle$. While this is not a common practice, one can actually identify the field operator

$$\Psi_\sigma^\dagger(\vec{r}) \equiv \hat{c}_{\vec{r}\sigma}^\dagger$$

as the usual creation operator of an electron at point $\vec{r}$ with spin σ.

To conclude this section it is interesting to note that the space coordinate $\vec{r}$ is a *parameter* of $\hat{\Psi}_\sigma(\vec{r})$ and $\hat{\Psi}_\sigma^\dagger(\vec{r})$ very much like the time t. In second quantization the position is sort of "degraded" to a parameter like the time. It is in this way that the required equivalent handling of time and space, that is necessary for relativistic quantum mechanics, is achieved.

4.3 Many-particle boson states

Bosons are characterized by symmetric wave functions, i.e.,

$$\Psi(x_{P(1)}, x_{P(2)}, \ldots, x_{P(N)}) = \Psi(x_1, \ldots x_N)$$

for any permutation $\mathbb{P} : [1, N] \to [1, N]$. As in the fermion case we may expand Ψ as

$$\Psi(x_1 \ldots x_N) = \sum_{K_1=1}^{\infty} \ldots \sum_{K_N=1}^{\infty} a(K_1, \ldots K_N)\, \varphi_{K_1}(x_1) \ldots \varphi_{K_N}(x_N)$$

where $\varphi_\alpha(x)$ $(\alpha = 1, 2, \ldots \infty)$ is a complete orthonormal set of single-particle states. Clearly,

$$a(K_1, \ldots K_N) = \int dx_1 \ldots dx_N\, \varphi^*_{K_1}(x_1) \ldots \varphi^*_{K_N}(x_N)\, \Psi(x_1, \ldots x_N)$$

from which we have

$$a(K_{P(1)}, \ldots K_{P(N)}) = a(K_1, \ldots K_N) \quad \forall\ P.$$

We conclude that the order of the orbitals $K_1, K_2, \ldots K_N$ on which the coefficient a depends is irrelevant. Only the number of appearances n_α of each basis orbital α on the list $K_1, K_2, \ldots K_N$ matters. One refers to $n_\alpha \geq 0$ as the occupation number of the orbital α. It is therefore meaningful to make the dependence on the occupation numbers evident by introducing the coefficients

$$\tilde{c}(n_1, n_2, \ldots n_\infty) = a(K_1, \ldots K_N)$$

corresponding to n_α appearencces of the orbital φ_α on the list $K_1, K_2, \ldots K_N$ of the right hand side. Obviously, we have $\sum_\alpha n_\alpha = N$.

The normalization condition on Ψ

$$\int dx_1 \ldots dx_N\, |\Psi(x_1, \ldots x_N)|^2 = 1$$

and $\langle\, \varphi_\alpha \,|\, \varphi_\beta \,\rangle = \delta_{\alpha\beta}$ imply

$$\sum_{K_1=1}^{\infty} \ldots \sum_{K_N=1}^{\infty} |a(K_1, \ldots K_N)|^2 = 1. \tag{4.7}$$

Notice that each group of indices $K_1, \ldots K_N$ with a given order appears only once. We can now make use of the definition of $\tilde{c}\,(n_1, n_2, \ldots n_\infty)$ and express the sum in Eq. (4.7) as

$$\sum_{n_1=0}^{N} \ldots \sum_{n_\infty=0}^{N} |\tilde{c}\,(n_1, n_2, \ldots n_\infty)|^2 \left(\sum_{\substack{K_1, \ldots K_N \\ (n_1, \ldots n_\infty)}} 1 \right) = 1,$$

where the sum over $K_1, \ldots K_n$ in brackets is subject to the constraint that the occupation numbers are $n_1, n_2, \ldots n_\infty$. This sum is equal to the number of ways of choosing n_1 indices $K_i = 1$, n_2 indices $K_i = 2$, n_α indices $K_i = \alpha$, etc.. This is given by $N! \,/\, (n_1!\, n_2! \ldots n_\infty!)$, where $\sum_\alpha n_\alpha = N$, and therefore

$$\sum_{n_1,\ldots n_\infty} |\tilde{c}\,(n_1, \ldots n_\infty)|^2 \, \frac{N!}{n_1! \ldots n_\infty!} = 1.$$

We may now express $\Psi(x_1, \ldots x_N)$ using the relations between $a(K_1, \ldots K_N)$, $a(K_{P(1)}, \ldots K_{P(N)})$ and $\tilde{c}\,(n_1, \ldots n_\infty)$ as

$$\Psi(x_1, \ldots x_N) = \sum_{n_1,\ldots n_\infty} \tilde{c}\,(n_1, \ldots n_\infty) \sum_{P_{\neq K_i}} \varphi_{K_{P(1)}}(x_1) \ldots \varphi_{K_{P(N)}}(x_N),$$

where $\sum_{P_{\neq K_i}}$ is the sum over all possible permutations of the indices $i = 1, \ldots, N$ such that $K_{P(i)} \neq K_i$ for some i. The number of such permutations is $\frac{N!}{n_1! \ldots n_\infty!}$. One may easily convince oneself about the latter by expressing the total number of permutations of N elements $K_1, \ldots K_N$ (that is equal to $N!$) as the number of permutations among different elements times the number of permutations of equal elements among themselves, which is equal to $n_1! \ldots n_\infty!$. In other words, each different ordering of the single-particle states appears $n_1! \ldots n_\infty!$ times if one performs all the $N!$ permutations. We only exclude counting the same product twice, i.e., we exclude the cases where $K_i = K_{P(i)}$ for all i.

Boson wave functions can then be written as

$$\Psi(x_1, \ldots x_N) = \sum_{n_1,\ldots n_\infty} \tilde{c}\,(n_1, \ldots n_\infty) \left(\frac{N!}{n_1! \ldots n_\infty!} \right)^{\frac{1}{2}} \Phi^S_{n_1,\ldots n_\infty}(x_1, \ldots x_N)$$

where

$$\Phi^S_{n_1,\ldots n_\infty}(x_1, \ldots x_N) = \sqrt{\frac{n_1! \ldots n_\infty!}{N!}} \sum_{P_{\neq K_i}} \varphi_{K_{P(1)}}(x_1) \ldots \varphi_{K_{P(N)}}(x_N)$$

is a *completely symmetrized wave function* corresponding to N particles having the occupation numbers $n_1, n_2, \ldots n_\infty$. Notice the *orthonormalization* of $\Phi^S_{n_1,\ldots n_\infty}$ since

$$\int \left[\Phi^S_{n'_1,\ldots n'_\infty}(x_1, \ldots x_N) \right]^* \Phi^S_{n_1,\ldots n_\infty}(x_1, \ldots x_N)\, dx_1 \ldots dx_N = \delta_{n_1 n'_1}\, \delta_{n_2 n'_2} \ldots\, \delta_{n_\infty n'_\infty}.$$

It is now convenient to redefine the expansion coefficients as

$$c\,(n_1, \ldots n_\infty) = \tilde{c}\,(n_1, \ldots n_\infty) \left(\frac{N!}{n_1! \ldots n_\infty!} \right)^{\frac{1}{2}}$$

that satisfy

$$\sum_{n_1,\ldots n_\infty} |c\,(n_1,\ldots n_\infty)|^2 = 1$$

and in terms of which we have

$$\boxed{\Psi(x_1,\ldots x_n) = \sum_{n_1,\ldots n_\infty} c\,(n_1,\ldots n_\infty)\,\Phi^S_{n_1,\ldots n_\infty}(x_1,\ldots x_N).}$$

In analogy to the case of fermions, we have shown that the most general symmetric wave function Ψ can be expressed as a superposition of completely symmetrized N-particle states $\Phi^S_{n_1,\ldots n_\infty}(x_1,\ldots x_N)$ that only depend on the *occupation numbers* of the single-particle states $\varphi_\alpha(x)$.

4.3.1 Matrix elements of single-particle operators

The previous result shows that it is possible to change the representation of states to one in which the occupation numbers $n_1,\ldots n_\infty$ are the dynamical variables. We consider then the ket states $|\, n_1,\ldots n_\infty \,\rangle$ such that

$$\langle\, x_1,\ldots x_N \,|\, n_1,\ldots n_\infty \,\rangle = \Phi^S_{n_1,\ldots n_\infty}(x_1,\ldots x_N)$$

and we seek for the expression of the matrix elements of single-particle operators among the $\Phi^S_{n_1,\ldots n_\infty}$.

Let us consider a single-particle operator of the form

$$F^{(1)} = \sum_{i=1}^{N} f^{(1)}(x_i)$$

and introduce the single-particle matrix elements as in the fermion case

$$f^{(1)}_{\alpha\beta} = \int dx\; \varphi^*_\alpha(x)\; f^{(1)}(x)\; \varphi_\beta(x).$$

The matrix elements between many-body states are given by

$$\langle\, n'_1,\ldots n'_\infty \,|\, F^{(1)} \,|\, n_1,\ldots n_\infty \,\rangle = \frac{\sqrt{n'_1!\ldots n'_\infty!}\,\sqrt{n_1!\ldots n_\infty!}}{N!}\; N$$
$$\times \sum_{P_{\neq K'_i}} \sum_{Q_{\neq K_i}} \int dx_1\ldots dx_N\; \varphi^*_{K'_{P(1)}}(x_1)\ldots\varphi^*_{K'_{P(N)}}(x_N)\; f^{(1)}(x_1)\; \varphi_{K_{Q(1)}}(x_1)\ldots\varphi_{K_{Q(N)}}(x_N)$$

where we have used that Φ^S are completely symmetric and therefore

$$\int \ldots f^{(1)}(x_i)\ldots = \int \ldots f^{(1)}(x_j)\ldots \quad \forall\, i,j.$$

It is clear that in order for the integral to give a non-vanishing result we must have $K'_{P(j)} = K_{Q(j)}$ for all $j \geq 2$. Therefore, the occupation numbers in $|\, n'_1, \ldots n'_\infty \,\rangle$ and $|\, n_1, \ldots n_\infty \,\rangle$ can differ in *at most the occupation of one orbital*. We are thus interested in two cases: the diagonal elements $\langle\, n_1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle$ and the off-diagonal elements of the form $\langle\, n_1, \ldots n_\alpha, \ldots n_\beta - 1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1 \ldots n_\alpha - 1, \ldots n_\beta, \ldots n_\infty \,\rangle$.

Diagonal elements

The diagonal elements are given by

$$\langle\, n_1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle = \frac{n_1! \ldots n_\infty!}{N!}\, N$$
$$\times \sum_{P \neq K_i} \sum_{Q \neq K_i} \int dx_1 \ldots dx_N\, \varphi^*_{K_{P(1)}}(x_1) \ldots \varphi^*_{K_{P(N)}}(x_N)\, f^{(1)}(x_1)\, \varphi_{K_{Q(1)}}(x_1) \ldots \varphi_{K_{Q(N)}}(x_N).$$

Clearly, we must have $P(j) = Q(j),\ \forall\, j \geq 2$ and thus $P \equiv Q$. Note that $\sum_{P \neq K_i}$ involves always different products of $\varphi_{K_i}(x)$, i.e., $K_{P_{(i)}} \neq K_i$ for some i. We can then write

$$\langle\, n_1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle = \frac{n_1! \ldots n_\infty!}{N!}\, N \sum_{Q \neq K_i} \int dx\, \varphi^*_{K_{Q(1)}}(x)\, f^{(1)}(x)\, \varphi_{K_{Q(1)}}(x).$$

Let us consider the permutations Q satisfying $Q(1) = i$ and let $K_i = K_{Q(1)} = \alpha$. In this case the integrals are all equal to $f^{(1)}_{\alpha\alpha}$, so that

$$\langle\, n_1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle = \frac{n_1! \ldots n_\infty!}{(N-1)!} \sum_{\alpha_{\rm occ}} f^{(1)}_{\alpha\alpha} \sum_{\substack{Q \neq K_i \\ K_{Q(1)} = \alpha}} 1. \tag{4.8}$$

The number of permutations Q among different K_i is $\dfrac{N!}{n_1! \ldots n_\infty!}$. Now, the number of such permutations that satisfy $K_{Q(1)} = \alpha$ is

$$\sum_{\substack{Q \neq K_i \\ K_{Q(1)} = \alpha}} 1 = \frac{(N-1)!}{n_1! \ldots (n_\alpha - 1)! \ldots n_\infty!}. \tag{4.9}$$

In fact, one may easily verify that

$$\sum_{\alpha_{\rm occ}} \frac{(N-1)!}{n_1! \ldots (n_\alpha - 1)! \ldots n_\infty!} = \sum_{\alpha_{\rm occ}} \frac{(N-1)!\, n_\alpha}{n_1! \ldots n_\alpha! \ldots n_\infty!} = \frac{(N-1)!}{n_1! \ldots n_\infty!} \sum_\alpha n_\alpha = \frac{N!}{n_1! \ldots n_\infty!}.$$

Replacing Eq. (4.9) in Eq. (4.8) we obtain

$$\begin{aligned} \langle\, n_1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\infty \,\rangle &= \frac{n_1! \ldots n_\infty!}{(N-1)!} \sum_{\alpha_{\rm occ}} f^{(1)}_{\alpha\alpha}\, \frac{n_\alpha\, (N-1)!}{n_1! \ldots n_\alpha! \ldots n_\infty!} \\ &= \sum_\alpha n_\alpha\, f^{(1)}_{\alpha\alpha}\,, \end{aligned}$$

which is thus in closed final form.

Off-diagonal matrix elements

Assuming in order to be explicit that $n_\alpha \geq 1$ and $n_\beta \geq 1$, the off-diagonal elements have the form

$$\begin{aligned}
&\langle\, n_1, \ldots n_\alpha, \ldots n_\beta - 1, \ldots n_\infty \,|\, F^{(1)} \,|\, n_1, \ldots n_\alpha - 1, \ldots n_\beta, \ldots n_\infty \,\rangle = \\
&\quad = \frac{\sqrt{n_1! \ldots n_\alpha! \ldots (n_\beta - 1)! \ldots n_\infty!}\, \sqrt{n_1! \ldots (n_\alpha - 1)! \ldots n_\beta! \ldots n_\infty!}}{N!}\, N \\
&\quad \times \sum_{P_{\neq K'_i}} \sum_{Q_{\neq K_i}} \int \phi^*_{K'_{P(1)}}(x_1) \ldots \phi^*_{K'_{P(N)}}(x_N)\, f^{(1)}(x_1)\, \phi_{K_{Q(1)}}(x_1) \ldots \phi_{K_{Q(N)}}(x_N)\, dx_1 \ldots dx_N \\
&\quad = \frac{n_1! \ldots n_\infty!}{(N-1)!\, \sqrt{n_\alpha}\, \sqrt{n_\beta}}\, f^{(1)}_{\alpha\beta} \sum_{\substack{P_{\neq K'_i} \\ K'_{P(1)} = \alpha}} \sum_{\substack{Q_{\neq K_i} \\ K_{Q(1)} = \beta}} \delta_{K_{Q(1)}\beta}\, \delta_{K'_{P(1)}\alpha} \prod_{j \geq 2} \delta_{K_{Q(j)} K'_{P(j)}}.
\end{aligned}$$

The integrals are zero unless $K_{Q(j)} = K'_{P(j)} \;\; \forall\, j \geq 2$, $K_{Q(1)} = \beta$, and $K'_{P(1)} = \alpha$, since φ_β appears one more time on the right than on the left, and φ_α appears on the left one more time than on the right.

The number of permutations Q among the different K_i that have $K_{Q(1)} = \beta$ is

$$\frac{(N-1)!}{n_1! \ldots (n_\alpha - 1)! \ldots (n_\beta - 1)! \ldots n_\infty!} = \frac{(N-1)!}{n_1! \ldots n_\infty!}\, n_\alpha\, n_\beta$$

and the number of permutations P among different K'_i that have $K'_{P(1)} = \alpha$ is

$$\frac{(N-1)!}{n_1! \ldots (n_\alpha - 1)! \ldots (n_\beta - 1)! \ldots n_\infty!} = \frac{(N-1)!\, n_\alpha\, n_\beta}{n_1! \ldots n_\infty!}.$$

Therefore,

$$\langle \ldots n_\alpha, \ldots n_\beta - 1, \ldots \,|\, F^{(1)} \,|\, \ldots n_\alpha - 1, \ldots n_\beta, \ldots \rangle = \sqrt{n_\alpha}\, \sqrt{n_\beta}\, f^{(1)}_{\alpha\beta},$$

and we can now drop the restriction $n_\alpha, n_\beta \geq 1$.

4.3.2 Matrix elements of two-particle operators

We consider a general two-particle operator

$$F^{(2)} = \frac{1}{2} \sum_{i \neq j} f^{(2)}(x_i, x_j)$$

and the two-particle matrix elements

$$f^{(2)}_{\alpha\beta\gamma\delta} = \int dx\, dx'\, \phi^*_\alpha(x)\, \phi^*_\beta(x')\, f^{(2)}(x, x')\, \phi_\gamma(x)\, \phi_\delta(x').$$

The matrix elements between the many-body states $|\, n_1, \ldots n_\infty \,\rangle$ and $|\, n'_1, \ldots n'_\infty \,\rangle$ vanish unless the occupations in the two states differ in at most two single-particle orbitals. This is, of course, due to the fact that $f^{(2)}(x_i, x_j)$ involves only two coordinates.

There are different cases to consider. Let us look first at ($\alpha \neq \beta \neq \gamma \neq \delta$):

$$\langle \ldots n_\alpha, \ldots n_\beta, \ldots n_\gamma - 1, \ldots n_\delta - 1, \ldots \,|\, F^{(2)} \,|\, \ldots n_\alpha - 1, \ldots n_\beta - 1, \ldots n_\gamma, \ldots n_\delta, \ldots \rangle =$$
$$= \frac{(\ldots n_\alpha! \ldots n_\beta! \ldots (n_\gamma - 1)! \ldots (n_\delta - 1)! \ldots)^{1/2}\ (\ldots (n_\alpha - 1)! \ldots (n_\beta - 1)! \ldots n_\gamma! \ldots n_\delta! \ldots)^{1/2}}{N!}$$
$$\times \frac{N(N-1)}{2} \sum_{P_\neq} \sum_{Q_\neq} \int \phi^*_{K'_{P(1)}}(x_1)\ \ \phi^*_{K'_{P(2)}}(x_2) \ldots \phi^*_{K'_{P(N)}}(x_N)\ f^{(2)}(x_1, x_2)$$
$$\times \phi_{K_{Q(1)}}(x_1)\ \ \phi_{K_{Q(2)}}(x_2) \ldots \phi_{K_{Q(N)}}(x_N)\, dx_1 \ldots dx_N.$$

There are four possible choices for the permutations P and Q that yield non-vanishing contributions:

$K_{P(1)}$	$K_{P(2)}$	$K_{Q(1)}$	$K_{Q(2)}$
α	β	γ	δ
β	α	γ	δ
α	β	δ	γ
β	α	δ	γ

In addition we must require, of course, $K'_{P(j)} = K_{Q(j)}$ for all $j \geq 3$. The number of permutations Q among different K_i having $K_{Q(1)} = \gamma$ (or δ) and $K_{Q(2)} = \delta$ (or γ) is $(N-2)!\,/\,[n_1! \ldots (n_\alpha - 1)! \ldots (n_\beta - 1)! \ldots (n_\gamma - 1)! \ldots (n_\delta - 1)!\,]$. The number of permutations P among different K'_i having $K'_{P(1)} = \alpha$ (or β) and $K'_{P(2)} = \beta$ (or α) is the same. Assuming, for the sake of being explicit, that $\alpha < \beta$ and $\gamma < \delta$, we can then write

$$\begin{aligned} \langle\, n_\alpha,\ n_\beta,\quad n_\gamma\ \ -1,\ n_\delta - 1 \,|\, F^{(2)} \,|\, n_\alpha - 1,\ n_\beta - 1,\ n_\gamma,\ n_\delta \,\rangle = \\ = \sqrt{n_\alpha}\,\sqrt{n_\beta}\,\sqrt{n_\gamma}\,\sqrt{n_\delta}\,\frac{1}{2}\left(f^{(2)}_{\alpha\beta\gamma\delta} + f^{(2)}_{\beta\alpha\gamma\delta} + f^{(2)}_{\alpha\beta\delta\gamma} + f^{(2)}_{\beta\alpha\delta\gamma}\right) \\ = (n_\alpha\, n_\beta\, n_\gamma\, n_\delta)^{1/2}\,\left(f^{(2)}_{\alpha\beta\gamma\delta} + f^{(2)}_{\alpha\beta\delta\gamma}\right). \end{aligned}$$

Let us consider another example of two-particle transition that does not appear in the case of fermions:

$$\langle\, n_\alpha,\ n_\beta,\ n_\gamma - 2 \,|\, F^{(2)} \,|\, n_\alpha - 1,\ n_\beta - 1,\ n_\gamma \,\rangle \quad \text{with} \quad \alpha < \beta \text{ and } \gamma \neq \alpha, \beta.$$

There are two possibilities for P and Q leading to non-vanishing matrix elements:

P(1)	P(2)	Q(1)	Q(2)
α	β	γ	γ
β	α	γ	γ

And there are $\dfrac{(N-2)!}{\dots(n_\alpha-1)!\dots(n_\beta-1)!\dots(n_\gamma-2)!\dots}$ different permutations Q having $K_{Q(1)} = K_{Q(2)} = \gamma$, and of course the same number of permutations P with $K'_{P(1)} = \alpha$ (or β) and $K'_{P(2)} = \beta$ (or α). We then have

$$\begin{aligned}\langle\, n_\alpha,\ n_\beta,\ n_\gamma{-}2 \,|\, F^{(2)} \,|\, n_\alpha{-}1,\ n_\beta{-}1,\ n_\gamma \,\rangle &= \sqrt{n_\alpha}\,\sqrt{n_\beta}\,\sqrt{n_\gamma}\,\sqrt{n_\gamma{-}1}\left(\frac{f_{\alpha\beta\gamma\gamma}+f_{\beta\alpha\gamma\gamma}}{2}\right)\\ &= [n_\alpha\, n_\beta\, n_\gamma\, (n_\gamma{-}1)]^{1/2}\ f_{\alpha\beta\gamma\gamma}.\end{aligned}$$

Finally, we must consider the case in which a pair of bosons occupying the same orbital γ change their state into the same orbital α:

$$\langle\, n_\alpha,\ n_\gamma - 2 \,|\, F^{(2)} \,|\, n_\alpha - 2,\ n_\gamma \,\rangle$$

where α and γ are arbitrary including the case $\alpha = \gamma$ in which no occupations change at all. There is only one possibilities for P and Q that leads to a non-vanishing result, namely, $P(1)$ and $P(2)$ pointing to α and $Q(1)$ and $Q(2)$ pointing to γ. There are $\dfrac{(N-2)!}{\dots(n_\alpha-2)!\dots(n_\gamma-2)!\dots}$ different permutations Q having $K_{Q(1)} = K_{Q(2)} = \gamma$, and of course the same number of permutations P with $K'_{P(1)} = K'_{P(2)} = \alpha$. We then have

$$\langle\, n_\alpha,\ n_\gamma{-}2 \,|\, F^{(2)} \,|\, n_\alpha{-}2,\ n_\gamma \,\rangle = \frac{1}{2}\,[n_\alpha\,(n_\alpha{-}1)\,n_\gamma(n_\gamma{-}1)]^{1/2}\ f_{\alpha\alpha\gamma\gamma}\,.$$

This concludes the calculation of matrix elements among the $\phi^S_{n_1,\dots n_\infty}(x_1,\dots x_N)$. We are now ready to introduce the creation and destruction operators which play the central role in the second quantization method.

4.3.3 Creation and annihilation operators

In order to describe transitions between different states $|\, n_1,\dots n_\infty \,\rangle$ having different occupation numbers we introduce the annihilation operator a_α which is defined by

$$\boxed{\hat{a}_\alpha \,|\, n_1,\dots n_\alpha,\dots n_\infty \,\rangle = \sqrt{n_\alpha}\,|\, n_1,\dots n_\alpha - 1,\dots n_\infty \,\rangle.}$$

In particular,

$$\hat{a}_\alpha \,|\, n_1,\dots 0_\alpha,\dots n_\infty \,\rangle = 0.$$

Notice that, in contrast with the fermionic case, $\hat{a}_\alpha \,|\, n_1,\dots n_\infty \,\rangle$ is not normalized to one, even if $n_\alpha \neq 0$, except for $n_\alpha = 1$. Actually, for fermions we only have $n_\alpha = 0$ or 1. Applying the annihilation operators $\hat{a}_\alpha$ and $\hat{a}_\beta$ successively we obtain, for $\alpha \neq \beta$,

$$\begin{aligned}\hat{a}_\beta\,\hat{a}_\alpha \,|\, n_1,\dots n_\alpha,\dots n_\infty \,\rangle &= \hat{a}_\beta\,\sqrt{n_\alpha}\,|\, n_1,\dots n_\alpha - 1,\dots n_\infty \,\rangle\\ &= \sqrt{n_\beta}\,\sqrt{n_\alpha}\,|\, n_1,\dots n_\alpha - 1,\dots n_\beta - 1,\dots n_\infty \,\rangle\\ &= \hat{a}_\alpha\,\hat{a}_\beta \,|\, n_1,\dots n_\infty \,\rangle.\end{aligned}$$

Therefore,

$$\boxed{[\hat{a}_\alpha,\, \hat{a}_\beta] = 0 \quad \forall\, \alpha, \beta}$$

and thus

$$\boxed{\left[\hat{a}^\dagger_\alpha,\, \hat{a}^\dagger_\beta\right] = 0 \quad \forall\, \alpha, \beta.}$$

This is consistent with the fact that $\langle\, x_1, \ldots x_N \,|\, n_1, \ldots n_\infty \,\rangle$ is completely symmetric, or, in other words, that $\langle\, x_1, \ldots x_N \,|\, \prod_\alpha (\hat{a}^\dagger_\alpha)^{n_\alpha} \,|\, \mathrm{vac} \,\rangle$ is proportional to $\langle\, x_1, \ldots x_N \,|\, n_1, \ldots n_\infty \,\rangle$.

In the case of fermions we exploited the freedom of choosing the *phase* of $\hat{c}^\dagger_k \,|\, n_1, \ldots n_\infty \,\rangle$ to incorporate the sign of the matrix elements in the ordering of the fermion operators $\hat{c}^\dagger_k$. Here, in the case of bosons, we introduce the multiplicative factor $\sqrt{n_\alpha}$ to take into account the fact that the number of single-particle transitions between $\alpha \to \beta$ orbitals depends on the occupation numbers of these orbitals. As we have seen, it is proportional to $\sqrt{n_\alpha}\, \sqrt{n_\beta}$.

The matrix elements of $\hat{a}^\dagger_\alpha$ in the occupation-number representation are easily obtained from

$$\langle\, n'_1, \ldots n'_\infty \,|\, \hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle = \sqrt{n_\alpha}\; \delta_{n_1\, n'_1} \ldots \delta_{n'_\alpha\, n_\alpha - 1} \ldots \delta_{n'_\infty\, n_\infty}.$$

After hermitic conjugation and using that $n'_\alpha = n_\alpha - 1$ we obtain

$$\langle\, n_1, \ldots n_\infty \,|\, \hat{a}^\dagger_\alpha \,|\, n'_1, \ldots n'_\infty \,\rangle = \sqrt{n'_\alpha + 1}\; \delta_{n_1\, n'_1} \ldots \delta_{n_\alpha\, n'_\alpha + 1} \ldots \delta_{n'_\infty\, n_\infty}$$

or equivalently $\boxed{\hat{a}^\dagger_\alpha \,|\, n_1, \ldots n_\infty \,\rangle = \sqrt{n_\alpha + 1}\, |\, n_1, \ldots n_\alpha + 1, \ldots n_\infty \,\rangle.}$

This implies that the *number operator*

$$\hat{n}_\alpha = \hat{a}^\dagger_\alpha\, \hat{a}_\alpha$$

satisfies

$$\begin{aligned} \hat{n}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle = \hat{a}^\dagger_\alpha\, \hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle &= \sqrt{n_\alpha}\; \hat{a}^\dagger_\alpha \,|\, n_1, \ldots n_\alpha - 1, \ldots n_\infty \,\rangle \\ &= n_\alpha \,|\, n_1, \ldots n_\infty \,\rangle. \end{aligned}$$

The commutation rule $[\hat{a}_\alpha, \hat{a}^\dagger_\beta] = 0$ for $\alpha \neq \beta$ follows immediately from

$$\hat{a}_\alpha\, \hat{a}^\dagger_\beta \,|\, n_1, \ldots n_\infty \,\rangle = \sqrt{n_\alpha}\, \sqrt{n_\beta + 1}\, |\, n_1, \ldots n_\alpha - 1, \ldots n_\beta + 1, \ldots n_\infty \,\rangle.$$

Consequently,

$$\boxed{[\hat{n}_\alpha, \hat{n}_\beta] = 0 \quad \forall\, \alpha, \beta.}$$

One concludes that the set of occupation numbers n_α constitutes a *complete set of compatible observables*, the $\hat{n}_\alpha$ being the associated operators. Notice that the multiplicative

factor $\sqrt{n_\alpha}$ in the definition of $\hat{a}_\alpha$ is crucial for this property to hold, i.e., for $\hat{n}_\alpha$ to be given by the same simple form as in the case of fermions.

Finally, we observe that

$$\begin{aligned}\hat{a}_\alpha\,\hat{a}_\alpha^\dagger\,|\,n_1,\ldots n_\infty\,\rangle &= \hat{a}_\alpha\,\sqrt{n_\alpha+1}\,|\,n_1,\ldots n_\alpha+1,\ldots n_\infty\,\rangle \\ &= (n_\alpha+1)\,|\,n_1,\ldots n_\alpha,\ldots n_\infty\,\rangle.\end{aligned}$$

Using the properties of the number operator $\hat{n}_\alpha = \hat{a}_\alpha^\dagger\,\hat{a}_\alpha$ we conclude that

$$\boxed{\hat{a}_\alpha\,\hat{a}_\alpha^\dagger - \hat{a}_\alpha^\dagger\,\hat{a}_\alpha = [\hat{a}_\alpha,\hat{a}_\alpha^\dagger] = 1.}$$

In summary, starting from the definition of $\hat{a}_\alpha$ on the complete set of many-particle states $|\,n_1,\ldots n_\infty\,\rangle$, namely

$$\hat{a}_\alpha\,|\,n_1,\ldots n_\infty\,\rangle = \sqrt{n_\alpha}\,|\,n_1,\ldots n_\alpha-1,\ldots n_\infty\,\rangle,$$

we have derived the *boson-operator commutation rules*

$$[\hat{a}_\alpha,\hat{a}_\beta] = [\hat{a}_\alpha^\dagger,\hat{a}_\beta^\dagger] = 0 \quad \forall\,\alpha,\beta \tag{4.10}$$

and

$$[\hat{a}_\alpha,\hat{a}_\beta^\dagger] = \delta_{\alpha\beta}. \tag{4.11}$$

The converse of the previous statement is also true. One can actually start from the commutation rules and derive the matrix elements of $\hat{a}_\alpha$, $\hat{a}_\alpha^\dagger$ and $\hat{n}_\alpha$ among the basis states $|\,n_1,\ldots n_\infty\,\rangle$.

To prove this we assume that the commutation rules (4.10) and (4.11) for $\hat{a}_\alpha$ and $\hat{a}_\beta^\dagger$ hold. Then one also has

$$[\hat{n}_\alpha,\hat{a}_\alpha] = [\hat{a}_\alpha^\dagger,\hat{a}_\alpha]\,\hat{a}_\alpha = -\hat{a}_\alpha \tag{4.12}$$

and

$$[\hat{n}_\alpha,\hat{a}_\alpha^\dagger] = \hat{a}_\alpha^\dagger\,[\hat{a}_\alpha,\hat{a}_\alpha^\dagger] = \hat{a}_\alpha^\dagger, \tag{4.13}$$

where we have defined as usual $\hat{n}_\alpha = \hat{a}_\alpha^\dagger\,\hat{a}_\alpha$. Eqs. (4.10) and (4.11) also imply that the operators $\hat{n}_\alpha$ commute, i.e., $[\hat{n}_\alpha,\hat{n}_\beta] = 0\ \forall\,\alpha,\beta$. We consider a basis set of simultaneous eigenstates of all $\hat{n}_\alpha$, i.e., $\hat{n}_\alpha\,|\,n_1,\ldots n_\infty\,\rangle = n_\alpha\,|\,n_1,\ldots n_\infty\,\rangle$, which exists, since the $\hat{n}_\alpha^\dagger = \hat{n}_\alpha$ are hermitian and commute.

The eigenvalues n_α are all non-negative, since

$$\begin{aligned}n_\alpha = \langle\,n\,|\,\hat{n}_\alpha\,|\,n\,\rangle = \langle\,n\,|\,\hat{a}_\alpha^\dagger\,\hat{a}_\alpha\,|\,n\,\rangle &= \sum_m \langle\,n\,|\,\hat{a}_\alpha^\dagger\,|\,m\,\rangle\langle\,m\,|\,\hat{a}_\alpha\,|\,n\,\rangle \\ &= \sum_m |\langle\,m\,|\,\hat{a}_\alpha\,|\,n\,\rangle|^2 \geq 0.\end{aligned}$$

Let $|\, n_1, \ldots n_\infty \,\rangle$ be an eigenstate of $\hat{n}_\alpha \;\forall\; \alpha$. Using Eq. (4.12) it follows that

$$\begin{aligned}\hat{n}_\alpha \, \hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle &= (\, [\hat{n}_\alpha, \hat{a}_\alpha] + \hat{a}_\alpha \, \hat{n}_\alpha) \;|\, n_1, \ldots n_\infty \,\rangle \\ &= -\hat{a}_\alpha + n_\alpha \, \hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle \\ &= (n_\alpha - 1) \; \hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle.\end{aligned}$$

This means that $\hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle$ is either zero or it is an eigenstate of $\hat{n}_\alpha$ with eigenvalue $(n_\alpha - 1)$. Consequently, the eigenvalues n_α of $\hat{n}_\alpha$ must be all positive integers or zero, since otherwise we would have negative eigenvalues for $\hat{n}_\alpha$. The latter is, as we have shown, impossible.

Taking this into account we can always write

$$\hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle = A \,|\, n_1, \ldots n_\alpha - 1, \ldots n_\infty \,\rangle$$

where $|\, n_1, \ldots n_\alpha - 1, \ldots n_\infty \,\rangle$ is an element of the basis set of eigenstates of all $\hat{n}_\alpha$ and A is a constant. The normalization of the basis states $|\, n_1, \ldots n_\infty \,\rangle$ implies

$$|A|^2 = \langle\, n_1, \ldots n_\infty \,|\, \hat{a}^\dagger_\alpha \, \hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle = n_\alpha.$$

Therefore, besides an irrelevant phase factor, we have $A = \sqrt{n_\alpha}$ and

$$\boxed{\hat{a}_\alpha \,|\, n_1, \ldots n_\infty \,\rangle = \sqrt{n_\alpha} \,|\, n_1, \ldots n_\alpha - 1, \ldots n_\infty \,\rangle.}$$

This proves the converse.

4.3.4 Second quantization form of one- and two-particle operators

Once we have calculated the matrix elements of $\hat{F}^{(1)}$ and $\hat{F}^{(2)}$ between states with definite occupation numbers $|\, n_1, \ldots n_\infty \,\rangle$, and knowing the definition and matrix elements of $\hat{a}_\alpha$ and $\hat{a}^\dagger_\alpha$, it is quite straightforward to express $\hat{F}^{(1)}$ and $\hat{F}^{(2)}$ in second quantization form. For the one-particle operators we have

$$\langle\, n_\alpha, \; n_\beta - 1 \,|\, \hat{F}^{(1)} \,|\, n_\alpha - 1, \; n_\beta \,\rangle = \sqrt{n_\alpha} \; \sqrt{n_\beta} \; f^{(1)}_{\alpha\beta}$$

for $\alpha \neq \beta$ and

$$\langle \ldots n_\alpha, \ldots \,|\, \hat{F}^{(1)} \,|\, \ldots n_\alpha, \ldots \,\rangle = \sum_\alpha \, n_\alpha \; f^{(1)}_{\alpha\alpha}.$$

for the diagonal elements. Therefore,

$$\boxed{\hat{F}^{(1)} = \sum_{\alpha\beta} \; f^{(1)}_{\alpha\beta} \; \hat{a}^\dagger_\alpha \, \hat{a}_\beta \,.}$$

For the two-particle operators we have

$$\begin{aligned}\langle\, n_\alpha, \; n_\beta, n_\gamma - 1, \; n_\delta - 1 \,|\, \hat{F}^{(2)} \,|\, n_\alpha - 1, \; n_\beta - 1, \; n_\gamma, \; n_\delta \,\rangle = \\ = \sqrt{n_\alpha} \; \sqrt{n_\beta} \; \sqrt{n_\gamma} \; \sqrt{n_\delta} \; \left(f^{(2)}_{\alpha\beta\gamma\delta} + f^{(2)}_{\alpha\beta\delta\gamma} \right)\end{aligned}$$

for $\alpha \neq \beta$ and $\delta \neq \gamma$, and

$$\langle\, n_\alpha,\ n_\beta, n_\gamma - 2 \,|\, \hat{F}^{(2)} \,|\, n_\alpha - 1,\ n_\beta - 1,\ n_\gamma \,\rangle = \\ = \frac{1}{2}\, \sqrt{n_\alpha}\, \sqrt{n_\beta}\, \sqrt{n_\gamma}\, \sqrt{n_\gamma - 1}\, \left(f^{(2)}_{\alpha\beta\gamma\gamma} + f^{(2)}_{\beta\alpha\gamma\gamma} \right)$$

for the transitions of the form $\gamma\gamma \to \alpha\beta$ or $\beta\alpha$. A similar hermitic conjugate expression holds for the transitions of the form $\alpha\beta$ or $\beta\alpha \to \gamma\gamma$. Finally, the matrix elements for two-particle transition of the form $\alpha\alpha \to \gamma\gamma$ are given by

$$\langle\, n_\alpha,\ n_\gamma - 2 \,|\, F^{(2)} \,|\, n_\alpha - 2,\ n_\gamma \,\rangle = \frac{1}{2}\, [n_\alpha\, (n_\alpha - 1)\, n_\gamma (n_\gamma - 1)]^{1/2}\ f_{\alpha\alpha\gamma\gamma}\,.$$

Consequently, we can write

$$\begin{aligned}
\hat{F}^{(2)} &= \sum_{\alpha<\beta} \sum_{\gamma<\delta} \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\gamma\, \hat{a}_\delta\ \left(f^{(2)}_{\alpha\beta\gamma\delta} + f^{(2)}_{\alpha\beta\delta\gamma} \right) \\
&\quad + \frac{1}{2} \sum_{\alpha<\beta} \sum_{\gamma} \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}^2_\gamma\ \left(f^{(2)}_{\alpha\beta\gamma\gamma} + f^{(2)}_{\beta\alpha\gamma\gamma} \right) + \frac{1}{2} \sum_{\alpha} \sum_{\gamma<\delta} \hat{a}^{\dagger\,2}_\alpha\, \hat{a}_\gamma\, \hat{a}_\delta\ \left(f^{(2)}_{\alpha\alpha\gamma\delta} + f^{(2)}_{\alpha\alpha\delta\gamma} \right) \\
&\quad + \frac{1}{2} \sum_{\alpha,\gamma} \hat{a}^{\dagger\,2}_\alpha\, \hat{a}^2_\gamma\, f^{(2)}_{\alpha\alpha\gamma\gamma} \\
&= \sum_{\alpha<\beta} \sum_{\gamma\neq\delta} \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\gamma\, \hat{a}_\delta\, f^{(2)}_{\alpha\beta\gamma\delta} + \sum_{\alpha<\beta} \sum_{\gamma} \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}^2_\gamma\, f^{(2)}_{\alpha\beta\gamma\gamma} \\
&\quad + \frac{1}{2} \sum_{\alpha} \sum_{\gamma\neq\delta} \hat{a}^{\dagger\,2}_\alpha\, \hat{a}_\gamma\, \hat{a}_\delta\, f^{(2)}_{\alpha\alpha\gamma\delta} + \frac{1}{2} \sum_{\alpha,\gamma} \hat{a}^{\dagger\,2}_\alpha\, \hat{a}^2_\gamma\, f^{(2)}_{\alpha\alpha\gamma\gamma} \\
&= \sum_{\alpha<\beta} \sum_{\gamma,\delta} \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\gamma\, \hat{a}_\delta\, f^{(2)}_{\alpha\beta\gamma\delta} + \frac{1}{2} \sum_{\alpha} \sum_{\gamma,\delta} \hat{a}^{\dagger\,2}_\alpha\, \hat{a}_\gamma\, \hat{a}_\delta\, f^{(2)}_{\alpha\alpha\gamma\delta} \\
&= \frac{1}{2} \sum_{\alpha\neq\beta} \sum_{\gamma,\delta} \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\gamma\, \hat{a}_\delta\, f^{(2)}_{\alpha\beta\gamma\delta} + \frac{1}{2} \sum_{\alpha} \sum_{\gamma,\delta} \hat{a}^{\dagger\,2}_\alpha\, \hat{a}_\gamma\, \hat{a}_\delta\, f^{(2)}_{\alpha\alpha\gamma\delta}
\end{aligned}$$

and finally

$$\hat{F}^{(2)} = \frac{1}{2} \sum_{\substack{\alpha,\beta \\ \gamma,\delta}} \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\gamma\, \hat{a}_\delta\, f^{(2)}_{\alpha\beta\gamma\delta}.$$

Knowing that $[\hat{a}_\delta, \hat{a}_\gamma] = 0$ it is convenient to bring $\hat{F}^{(2)}$ in the form

$$\boxed{\hat{F}^{(2)} = \frac{1}{2} \sum_{\substack{\alpha\beta \\ \gamma\delta}} \hat{a}^\dagger_\alpha\, \hat{a}^\dagger_\beta\, \hat{a}_\delta\, \hat{a}_\gamma\, f^{(2)}_{\alpha\beta\gamma\delta}}$$

in order to make the hermiticity evident (since $f^{(2)}_{\alpha\beta\gamma\delta} = f^{(2)}_{\beta\alpha\delta\gamma}$) and in order to have the same formal expression as for fermions. However, notice that for bosons the terms with $\alpha = \beta$ and/or $\gamma = \delta$, and even $\alpha = \beta = \gamma = \delta$ contribute!

The one-particle and two-particle operators have the same interpretation as in the case of fermions. The symmetry of the underlying many-particle states is taken into account by the commutation rules of the creation and annihilation operators.

4.3.5 Boson fields

The second quantization expression of single- and two-particle operators can be brought (as in the case of fermions) in a form which is explicitly independent of the single-particle basis. One defines the *annihilation field-operator*

$$\hat{\Psi}(x) = \sum_{\alpha=1}^{\infty} \varphi_\alpha(x)\, \hat{a}_\alpha$$

and the *creation field-operator*

$$\hat{\Psi}^\dagger(x) = \sum_{\alpha=1}^{\infty} \varphi_\alpha^*(x)\, \hat{a}_\alpha^\dagger,$$

where $\varphi_\alpha(x) = \langle\, x \,|\, a_\alpha^\dagger \,|\, 0 \,\rangle$ are the single-particle basis states in coordinate representation.

The boson field operators have similar properties as the fermion field operators provided that one replaces the anticommutators by commutators whenever appropriate. The following properties can be easily demonstrated:

$$[\hat{\Psi}(x), \hat{\Psi}(x')] = [\hat{\Psi}^\dagger(x), \hat{\Psi}^\dagger(x')] = 0 \quad \forall\, x, x',$$

$$[\hat{\Psi}(x), \hat{\Psi}^\dagger(x')] = \delta(x - x'),$$

$$\hat{F}^{(1)} = \int dx\ \hat{\Psi}^\dagger(x)\ f^{(1)}(x)\ \hat{\Psi}(x),$$

and

$$\hat{F}^{(2)} = \frac{1}{2} \int dx\, dx'\ \hat{\Psi}^\dagger(x)\, \hat{\Psi}^\dagger(x')\ f^{(2)}(x, x')\ \hat{\Psi}(x')\, \hat{\Psi}(x).$$

Although the order of $\hat{\Psi}(x)$ and $\hat{\Psi}(x')$ is immaterial in the case of bosons, one keeps the same expression as for fermions which is explicitly hermitic since $f(x, x') = f(x', x)^*$.

Finally, the following properties should be mentioned:

1) $\hat{\Psi}(x)$ $[\hat{\Psi}^\dagger(x)]$ decreases (increases) the number of particles by one.

2) The analogies between first quantization averages and second quantization operators that were discussed for fermions also hold for bosons.

3) The creation operators $\hat{a}_\alpha^\dagger$ transform like the basis states $\varphi_\alpha(x)$, i.e.,

$$\langle\, x \,|\, \hat{a}_\alpha^\dagger \,|\, 0 \,\rangle = \varphi_\alpha(x).$$

Therefore, if $\chi_\beta(x) = \sum_\alpha U_{\beta\alpha}\, \varphi_\alpha(x)$ then the creation operator $\hat{c}_\beta^\dagger$ associated to χ_β is $\hat{c}_\beta^\dagger = \sum_\alpha U_{\beta\alpha}\, \hat{a}_\alpha^\dagger$.

4) $\hat{\Psi}^\dagger(x)$ and $\hat{\Psi}(x)$ are independent of the choice of the single-particle basis.

5) $\hat{\Psi}^\dagger(x')$ creates a state that is localized at x' (provided the single-particle basis is complete) since

$$\begin{aligned} \langle\, x \,|\, \hat{\Psi}^\dagger(x') \,|\, 0 \,\rangle &= \sum_\alpha \phi_\alpha^*(x')\, \langle\, x \,|\, \hat{a}_\alpha^\dagger \,|\, 0 \,\rangle \\ &= \sum_\alpha \phi_\alpha^*(x')\, \phi_\alpha(x) = \delta(x - x'). \end{aligned}$$

The demonstration of these statements is analogous to the fermion case.

5 Exchange Interaction

In non-relativistic theory and in the absence of a magnetic field the Hamiltonian of an N-particle system is independent of the spin variables of the particles, since the kinetic energy is independent of spin (non-relativistic limit) and the interactions are also independent of the spin orientation of the particles. For instance, in the case of electrons, the interactions are electrical Coulomb repulsions. At first sight one might therefore be tempted to believe that the eigenstates of an N-particle system should be independent of the spin quantum numbers of each particle, and in particular of the total spin S of the system. However, this is not true because the indistinguishability of identical particles imposes the wave function to be symmetrical (bosons) or antisymmetrical (fermions) with respect to the interchange of any pair of particle coordinates $x_i \equiv (\vec{r}_i, \sigma_i)$ and $x_j = (\vec{r}_j, \sigma_j)$. This introduces correlations in the probability-density distribution of finding a pair of particles at points $\vec{r}$ and $\vec{r}\,'$, particularly when $\vec{r}$ and $\vec{r}\,'$ are close and the interaction is strongest. We have actually shown that the symmetry or antisymmetry property of the wave function with respect to interchange holds independently of the considered complete set of observables defining the state of the particles (e.g., position $\vec{r}_i$ and spin σ_i, or momentum $\vec{p}_i$ and spin σ_i). However, all the coordinates have to be exchanged at the same time. Therefore, a subtle dependence of the energy spectrum on S appears, since the symmetry with respect to interchange of the spin part of the wave function depends on the value of the total spin S. This holds despite the fact that the spin variables do not enter explicitly in the many-body Schrödinger equation. The purpose of this section is to discuss these indistinguishability correlations often called *exchange interactions*.

In the absence of magnetic field the non-relativisitic Schrödinger equation does not depend on the spin variables σ_i. Therefore, we may write the N particle wave functions of the stationary states in the form of a product

$$\Psi(x_1, x_2, \ldots x_N) = \phi(\vec{r}_1, \vec{r}_2, \ldots \vec{r}_N)\ \chi(\sigma_1, \sigma_2, \ldots \sigma_N),$$

where ϕ depends only on the Cartesian coordinates, and χ on the spin variables. ϕ is known as the *coordinate wave-function* and χ as the *spin wave-function*. The Hamiltonian acts on the spatial coordinates and determines $\phi(\vec{r}_1 \ldots \vec{r}_N)$. The function χ remains arbitrary *a priori*, as far as finding the eigenstates of $\hat{H}$ is concerned.

To be explicit, let us first consider a system of two identical particles and let $\hat{P}_{12}$ be the operator that permutes the variables of particles 1 and 2. Since the particles are identical, $\hat{H}$ commutes with the operator $\hat{P}_{12}$ that permutes the two spatial coordinates. In fact, for any function $\phi(x_1, x_2)$ we have

$$\hat{P}_{12}\ \hat{H}(1,2)\ \phi(1,2) = \hat{H}(2,1)\ \phi(2,1) = \hat{H}(1,2)\ \phi(2,1) = \hat{H}(1,2)\ \hat{P}_{12}\ \phi(1,2),$$

where we have simply used that H is invariant with respect to interchange. Thus $\hat{P}_{12}\,\hat{H} = \hat{H}\,\hat{P}_{12}$. Moreover, $\hat{P}_{12}$ commutes with the total momentum $\hat{P} = \sum_i \hat{p}_i$ and with the total angular momentum $\hat{\vec{L}} = \sum_i \hat{\vec{l}}_i$. $\hat{P}_{12}$ obviously preserves the norm and

cannot be antiunitary, since it commutes with the translation operator $e^{-i\,\hat{\vec{P}}\cdot\vec{a}/\hbar}$ and the rotation operator $e^{-i\,\hat{\vec{L}}\cdot\hat{n}\,\phi/\hbar}$. Remember that an antiunitary operator involves complex conjugation (see Sec. 3.4). Therefore, $\hat{P}_{12}$ is unitary. Moreover, since $\hat{P}_{12}^2 = 1$, the eigenvalues of $\hat{P}_{12}$ are ± 1. We may then choose the solutions of the coordinate or orbital part of the two-particle Schrödinger equation to be symmetric or antisymmetric with respect to interchange, i.e.,

$$\phi(\vec{r}_1, \vec{r}_2) = \pm\, \phi(\vec{r}_2, \vec{r}_1).$$

Of course, as any two variable function, the two-particle spin wave functions $\chi(\sigma_1, \sigma_2)$ can also be choosen to be symmetric or antisymmetric with respect to interchange.

In the following we discuss the symmetry properties of the complete physical wave functions $\Psi(x_1,\, x_2) = \phi(\vec{r}_1, \vec{r}_2)\chi(\sigma_1, \sigma_2)$, including the spin variables, for different kinds of identical particles.

5.1 Spin zero particles

In this case we are dealing with bosons and the spin factor χ is absent altogether. Thus,

$$\Psi(x_1, x_2) = \phi(r_1, r_2)$$

must be symmetric. Consequently, only the *symmetric solutions* of the Schrödinger equation are acceptable physical states for two spin-zero identical particles. Moreover, the wave function of a two-particle state in the center of mass frame can only depend on $\vec{r} = \vec{r}_1 - \vec{r}_2$. In the absence of an external field the solutions of the Schrödinger equation can be classified according to angular momentum l of the relative motion (i.e., $\phi_l(\vec{r}_1,\, \vec{r}_2) = R(r)\, Y_{lm}(\hat{r})$, where $\vec{r} = \vec{r}_2 - \vec{r}_1$). Therefore, since the interchange of the two particles is equivalent to the inversion of the relative coordinate, we have

$$\phi_l(\vec{r}_1, \vec{r}_2) = (-1)^l\, \phi_l(\vec{r}_2, \vec{r}_1),$$

where l is the orbital momentum of the relative motion. We conclude that two identical spin-zero particles can only have *even values of the total angular momentum.*

Example: The decay of a ^{8}Be nucleus in two α-particles can only lead to final states with even angular momentum.

5.2 Two spin 1/2 particles

We consider a two-electron system such as the He atom or H_2. The complete wave function $\Psi(x_1\, x_2) = \chi(\sigma_1\, \sigma_2)\ \phi(r_1, r_2)$ must be antisymmetric with respect to interchange. Thus, we have two possibilities:

$$\phi(\vec{r}_1, \vec{r}_2) = \phi(\vec{r}_2, \vec{r}_1) \ \text{ and } \ \chi(\sigma_1, \sigma_2) = -\chi(\sigma_2, \sigma_1)$$

or

$$\phi(\vec{r}_1, \vec{r}_2) = -\phi(\vec{r}_2, \vec{r}_1) \quad \text{and} \quad \chi(\sigma_1, \sigma_2) = \chi(\sigma_2, \sigma_1).$$

Since the Hamiltonian is independent of spin, it obviously commutes with the total spin operator $\hat{\vec{S}} = \sum_i \hat{\vec{s}}_i = \hat{\vec{s}}_1 + \hat{\vec{s}}_2$. Therefore χ can be chosen to be an eigenfunction of $\hat{S}^2$ and $\hat{S}_z$. For $S_z = 1$ the only state that can be formed with two spin-1/2 particles is

$$\chi_1^1 = |+,+\rangle = |\, m_1 = 1/2,\ m_2 = 1/2\,\rangle\,.$$

This is a triplet state ($S = 1$), since $\hat{S}^+ |+,+\rangle = (s_1^+ + s_2^+)\,|+,+\rangle = 0$, and is obviously symmetric with respect to the interchange of the spin variables m_1 and m_2.

The other triplet states with $S_z = 0$ and -1 are also symmetric since $\hat{S}^-$ commutes with $\hat{P}_{12}$:

$$\chi_0^1 = \frac{\hat{S}^-}{\sqrt{2}}\,\chi_1^1 = \frac{1}{\sqrt{2}}\,(\,|+,-\rangle + |-,+\rangle\,)$$

and

$$\chi_{-1}^1 = \frac{\hat{S}^-}{\sqrt{2}}\,\chi_0^1 = |-,-\rangle.$$

The singlet state that is obtained by orthogonalization with respect to χ_0^1. It is given by

$$\begin{aligned}\chi_0^0 &= \frac{1}{\sqrt{2}}\,(\,|+,-\rangle - |-,+\rangle\,) \\ &= \frac{1}{\sqrt{2}}(|\, m_1 = 1/2, m_2 = -1/2\,\rangle - |\, m_1 = -1/2, m_2 = 1/2\,\rangle)\end{aligned}$$

and is obviously antisymmetric with respect to interchange of m_1 and m_2. We therefore conclude that the only possible states of a two-electron system are

$$\begin{array}{ccc} \text{Symmetric solution} & & \text{spin singlet} \\ \phi(\vec{r}_1, \vec{r}_2) = \phi(\vec{r}_2, \vec{r}_1) & \text{and} & S = 0 \\ \text{even } l & & \text{(antiparallel spins)} \end{array}$$

or

$$\begin{array}{ccc} \text{Antisymmetric solution} & & \text{spin triplet} \\ \phi(\vec{r}_1, \vec{r}_2) = -\,\phi(\vec{r}_2, \vec{r}_1) & \text{and} & S = 1 \\ \text{odd } l & & \text{(parallel spins).} \end{array}$$

In other words, a symmetric (even l) solution $\phi(\vec{r}_1, \vec{r}_2)$ of the Schrödinger equation is an acceptable physical state only in combination with a singlet spin state, i.e., only if the two electrons form a singlet (antisymmetric) spin state. Conversely, singlet spin states are only possible if the coordinate wave function is symmetric with respect to interchange of $\vec{r}_1$ and $\vec{r}_2$. Analogously, antisymmetric orbital wave functions $\phi(\vec{r}_1, \vec{r}_2) = -\phi(\vec{r}_2, \vec{r}_1)$ are physically acceptable only if the corresponding spin wave-function $\chi(\sigma_1, \sigma_2)$ is a triplet.

This means that the whole coordinate Hilbert space accessible to singlet and triplet spin states are orthogonal to each other. Notice that this fundamental correlation between the symmetries of the coordinate and spin wave functions always holds. Even though states not having a well-defined total S and L are of course possible by superposition, all terms in a linear combination have to be physical states, i.e., they must all be antisymmetric with respect to interchange of the complete set of observables $\vec{r}$ and σ.

As a result the energy of a two-electron system crucially depends on the *total spin* S. This is the direct consequence of the principle of indistinguishability of identical particles. This one-to-one correspondence between *energy level* and the value of the *total spin* is characteristic of a two-electron system (i.e., two spin 1/2-system). In fact the correspondence between energy level and total S also holds for an arbitrary number of electrons. However, it does not hold for particles having higher intrinsic spin $s \geq 1$.

In order to illustrate some examples, let us recall the conventional spectroscopic notations ${}^{2S+1}L_J$ with $L \equiv S, P, D$, etc. for $L = 0, 1, 2$, etc. and $|L - S| \leq J \leq |L + S|$. For two electrons (or protons) the only possible states are

$$ {}^1S_0,\ {}^3P_{0,1,2}\ ,\ {}^1D_2,\ \text{etc.}, $$

while for distinguishable particles (e.g., a proton and a neutron in deuterium) we also find ${}^3S_1$ and ${}^1P_1$ states. This demonstrates, if at all necessary, that the one-to-one correspondence between total spin and orbital symmetry is the manifestation of the fundamental quantum mechanical indistinguishability of identical particles.

5.3 Two particles with arbitrary spin s

We consider the spin states $|\, m_1, m_2 \,\rangle$ for particles 1 and 2, where

$$ \hat{s}_1^z \,|\, m_1,\, m_2 \,\rangle = m_1 \,|\, m_1,\, m_2 \,\rangle\,, $$
$$ \hat{s}_2^z \,|\, m_1,\, m_2 \,\rangle = m_2 \,|\, m_1,\, m_2 \,\rangle\,, $$

and $-s \leq m_1, m_2 \leq s$. We would like to construct the states $|\, S, S_z \,\rangle$ with definite total S and S_z and at the same time find out the behavior of $|\, S, S_z \,\rangle$ with respect to interchange of the spin variables, m_1 and m_2 of the two particles.

The state with maximal $S = 2s$ is straightforwardly obtained, since the only option is to set $m_1 = s$ and $m_2 = s$:

$$ \left.\begin{aligned} S &= 2s \\ S_z &= 2s \end{aligned}\right\} \quad \Longrightarrow \quad \chi_{2s}^{2s} = |\, s, s \,\rangle\,. $$

χ_{2s}^{2s} is symmetric with respect to the interchange of the spin variables of the particles for all values of s ($S = 2s$). The state χ_{2s-1}^{2s}, with the same maximal $S = 2s$ and reduced projection $S_z = S - 1$ along the z-axis, is obtained by applying $\hat{S}^- = \hat{s}_1^- + \hat{s}_2^-$ to χ_{2s}^{2s}:

$$ \chi_{2s-1}^{2s} \propto \hat{S}^- \,|\, s, s \,\rangle = \sqrt{2s}\ (\,|\, s-1, s \,\rangle + |\, s, s-1 \,\rangle\,)\,. $$

The factor $\sqrt{2s}$ is obtained by applying the known equation

$$\hat{J}^-|\, j,m\,\rangle = \sqrt{(j+m)(j-m+1)}\,|\, j,m-1\,\rangle$$

for $m=j=s$. χ^{2s}_{2s-1} is of course symmetric, since $\hat{P}_{12}$ commutes with $\hat{\vec{S}} = \hat{\vec{s}}_1 + \hat{\vec{s}}_2$ and thus $[\hat{P}_{12}, \hat{S}^-] = 0$. The operator $\hat{S}^-$ preserves, therefore, the symmetry properties with respect to interchange.

The state with $S = 2s-1$ and maximal $S_z = S = 2s-1$ is a linear combination of $|\, s,\, s-1\,\rangle$ and $|\, s-1,\, s\,\rangle$, which is orthogonal to χ^{2s}_{2s-1}. Thus, it can be obtained by orthogonalization to χ^{2s}_{2s-1}:

$$\chi^{2s-1}_{2s-1} = \frac{1}{\sqrt{2}}\,(\,|\, s,s-1\,\rangle - |\, s-1,s\,\rangle\,) \quad (S=2s-1,\ S_z=S).$$

This state is clearly antisymmetric with respect to interchange of m_1 and m_2 for all values of s. Applying again $\hat{S}^-$ we obtain two states: χ^{2s}_{2s-2} and χ^{2s-1}_{2s-2}, which are, respectively, symmetric and antisymmetric, since $\hat{S}^-$ preserves the interchange symmetry ($[\hat{S}^-, \hat{P}_{12}] = 0$). In addition the state χ^{2s-2}_{2s-2} is necessarily symmetric, since it needs to be orthogonal to

$$\chi^{2s}_{2s-2} \propto a\,|\, s-2,s\,\rangle + b\,|\, s-1,s-1\,\rangle + a\,|\, s,s-2\,\rangle$$

and

$$\chi^{2s-1}_{2s-2} \propto |\, s-2,s\,\rangle - |\, s,s-2\,\rangle.$$

It is easy to convince oneself that at each step one obtains, alternatively, a symmetric and an antisymmetric state as the parity of S changes, i.e., symmetric states for $S = 2s,\, 2s-2,\, 2s-4, \ldots$ and antisymmetric states for $S = 2s-1,\, 2s-3,\, 2s-5, \ldots$.

In order to prove it, let us show that if χ^S_S is symmetric (antisymmetric) with respect to interchange, then χ^{S-1}_{S-1} is necessarily antisymmetric (symmetric). We know that

$$\chi^S_{S-1} = \hat{s}_1^-\ \chi^S_S + \hat{s}_2^-\ \chi^S_S\,.$$

It follows that

$$\chi^{S-1}_{S-1} \propto \hat{s}_1^-\ \chi^S_S - \hat{s}_2^-\ \chi^S_S\,, \tag{5.1}$$

since the right-hand side is orthogonal to χ^S_{S-1} and to all $\chi^{S'}_m$ with $S' > S$. Note that $\hat{s}_1^-\ \chi^S_S$ and $\hat{s}_2^-\ \chi^S_S$ have the same norm and are orthogonal to each other, since they have different s^z_1 or s^z_2.

Taking into account that there is only one state χ^{S-1}_{S-1}, the previous orthogonality conditions defines it univocally. Finally, it is easy to see that χ^{S-1}_{S-1} is antisymmetric (symmetric) when χ^S_S is symmetric (antisymmetric). From Eq. (5.1) we have

$$\begin{aligned}\hat{P}_{12}\ \chi^{S-1}_{S-1} &\propto \hat{P}_{12}\ (\hat{s}_1^-\ \chi^S_S) - \hat{P}_{12}\ (\hat{s}_2^-\ \chi^S_S)\\ &= \hat{s}_2^-\ \hat{P}_{12}\ \chi^S_S - \hat{s}_1^-\ \hat{P}_{12}\ \chi^S_S\,.\end{aligned}$$

A symmetric χ_S^S, i.e., $\hat{P}_{12}\ \chi_S = \chi_S$, implies therefore

$$\hat{P}_{12}\ \chi_{S-1}^{S-1} = -\chi_{S-1}^{S-1}\,,$$

while an antisymmetric χ_S^S, i.e., $\hat{P}_{12}\ \chi_S^S = -\chi_S^S$, implies

$$\hat{P}_{12}\ \chi_{S-1}^{S-1} = \chi_{S-1}^{S-1}\,.$$

For $S = 2s,\ 2s-2,\ 2s-4,\ \ldots$ the states are even, while for $2s-1,\ 2s-3,\ \ldots$ they are odd. As a result the symmetry with respect to interchange for the spin function χ is $(-1)^{2s-S}$, i.e., for $S = 2s$ always symmetric and then alternating as S decreases. This concludes the characterization of the spin part.

The symmetry of the *complete wave function* $\Psi(x_1\, x_2)$ is even for bosons (integer s) and odd for fermions (half-integer s). This can be expressed in a compact form as

$$\Psi(x_1, x_2) = (-1)^{2s}\ \Psi(x_2, x_1)\,.$$

Consequently, the orbital wave function must transform like $(-1)^{2s}\,/\,(-1)^{2s-S} = (-1)^S$, which depends only on S for both bosons and fermions.

The following combinations between spin and orbital wave functions are possible:

$$\begin{array}{lcl} \boxed{\text{Even } S} & & \phi(\vec{r}_1, \vec{r}_2) = \phi(\vec{r}_2, \vec{r}_1) \\ \chi \sim (-1)^{2s} & & \text{even orbital wave function} \\ & & \boxed{\text{even } l} \end{array}$$

or

$$\begin{array}{lcl} \boxed{\text{Odd } S} & & \phi(\vec{r}_1, \vec{r}_2) = -\phi(\vec{r}_2, \vec{r}_1) \\ \chi \sim (-1)^{2s+1} & & \text{odd orbital wave function} \\ & & \boxed{\text{odd } l} \end{array}$$

These correspondences hold for both fermions ($2s$ = odd) and bosons ($2s$ = even). For even S, $\chi \sim (-1)^{2s}$ is odd for fermions and even for bosons. In both cases, even S goes with an even coordinate wave function (even l). For instance, the orbital part of a singlet ($S = 0$) state is symmetric for two $s = 0$ bosons as well as for two $s = 1/2$ fermions. For odd S, $\chi \sim (-1)^{2s+1}$ is even for fermions and odd for bosons. In both cases, odd S goes with an odd orbital wave function (odd l) and, conversely, even S with an even symmetry of the orbital part. Thus, for two identical particles, the knowledge of S defines the interchange symmetry of the orbital wave function.

In conclusion, for a given energy level of the solution of the Schrödinger equation, only certain (even or odd) values of the total spin S are possible. However, the relation between energy level and total spin is <u>not</u> one-to-one, i.e., a given level with even (odd) l can correspond to any even (odd) S.

5.4 Spin dependence of the two-particle density distribution: Spin-1/2 particles

The fact that the symmetry of the orbital wave function $\phi(\vec{r}_1, \vec{r}_2)$ depends on the total spin S implies that the electronic pair-distribution function also depends on the total spin of the two-particle state. In order to discuss this effect, let us consider a pair of electrons and neglect the electron-electron interaction. The Hamiltonian is given by $\hat{H} = \hat{H}_1 + \hat{H}_2$ with

$$\hat{H}_i = \frac{-\hbar^2}{2m} \nabla_i^2 + v(\vec{r}_i).$$

Since $\hat{H}$ is the sum of two terms acting, respectively, on $\vec{r}_1$ and $\vec{r}_2$ we can express the two-electron eigenstates as a properly symmetrized or antisymmetrized product of single-particle wave functions $\omega_A(\vec{r})$ and $\omega_B(\vec{r})$, which are eigenstates of $\hat{H}_i$:

$$\hat{H}_1 \, \omega_A(\vec{r}_1) = \varepsilon_A \, \omega_A(r_1)$$

and

$$\hat{H}_2 \, \omega_B(\vec{r}_2) = \varepsilon_B \, \omega_B(r_2).$$

One can easily verify that

$$\hat{H} \, \omega_A(r_1) \, \omega_B(r_2) = (\varepsilon_A + \varepsilon_B) \, \omega_A(r_1) \, \omega_B(r_2)$$

and

$$\hat{H} \, \omega_B(r_1) \, \omega_A(r_2) = (\varepsilon_A + \varepsilon_B) \, \omega_B(r_1) \, \omega_A(r_2).$$

The normalized two-particle orbital wave-functions with defined interchange symmetry are given by

$$\phi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2(1 \pm |S_{AB}|^2)}} \, [\omega_A(r_1) \, \omega_B(r_2) \pm \omega_A(r_2) \, \omega_B(r_1)] \tag{5.2}$$

where the $+$ $(-)$ sign corresponds to the singlet (triplet) spin state, and

$$S_{AB} = \int d^3r \, \omega_A^*(\vec{r}) \, \omega_B(\vec{r})$$

is the overlap between the single-particle orbitals. As already discussed, the spin part of the wave function can be chosen to be a singlet or a triplet state, since $[\hat{H}, \, \hat{\vec{S}}] = 0$.

Let us consider for simplicity the case $S_{AB} = 0$. The probability of finding *one electron* at $\vec{r}$ is then given by

$$\begin{aligned}
\frac{1}{2} \, n(\vec{r}) &= \int |\phi(\vec{r}, \vec{r}\,')|^2 \, d^3r' = \\
&= \frac{1}{2} \int \left\{ |\omega_A(\vec{r})|^2 \, |\omega_B(\vec{r}\,')|^2 + |\omega_A(\vec{r}\,')|^2 \, |\omega_B(\vec{r})|^2 \pm \right. \\
&\quad \left. \pm \left[\omega_A^*(\vec{r}) \, \omega_B^*(\vec{r}\,') \, \omega_A(\vec{r}\,') \, \omega_B(\vec{r}) + \omega_A^*(r') \, \omega_B^*(r) \, \omega_A(r) \, \omega_B(r') \right] \right\} d^3r' \\
\frac{1}{2} \, n(\vec{r}) &= \frac{1}{2} \left(|\omega_A(\vec{r})|^2 + |\omega_B(\vec{r})|^2 \right).
\end{aligned}$$

Consequently, the antisymmetry or symmetry of the orbital wave function (5.2) has no effect on the *one-particle density distribution*, which is always equal to the superposition of the single-particle densities ($S_{AB} = 0$).

In order to observe the consequences of interchange symmetry, and thus a dependence on S, we need to consider the two-particle density $|\phi(\vec{r}_1, \vec{r}_2)|^2$, i.e., the probability density of finding one electron at $\vec{r}_1$ *and* the other at $\vec{r}_2$. This is given by

$$\begin{aligned} |\phi(\vec{r}_1, \vec{r}_2)|^2 &= \frac{1}{2} \left\{ |\omega_A(\vec{r}_1)|^2 \, |\omega_B(\vec{r}_2)|^2 + |\omega_A(\vec{r}_2)|^2 \, |\omega_B(\vec{r}_1)|^2 \right\} \pm \\ &\pm \mathrm{Re} \left\{ \omega_A(\vec{r}_1) \, \omega_B(\vec{r}_2) \, \omega_A^*(\vec{r}_2) \, \omega_B^*(\vec{r}_1) \right\} . \end{aligned} \tag{5.3}$$

Now the total spin S of the electron pair plays a role (even for $S_{AB} = 0$) provided that $\omega_A(\vec{r}) \, \omega_B(\vec{r})$ is not identically zero in all space (non-overlapping densities). The first two terms in Eq. (5.3) are positive, independent of S and, of course, its sum is symmetric. The last term in Eq. (5.3) is called the *exchange-pair density*. For $\vec{r}_1 = \vec{r}_2 = \vec{r}$ it is positive and equal to $|\omega_A(r)|^2 \, |\omega_B(\vec{r})|^2$. The exchange-pair density enhances (reduces) the probability of finding the two electrons in close-by positions ($\vec{r}_1 \simeq \vec{r}_2 = \vec{r}$) in the singlet (triplet) states.

Of course the exchange pair density is non-zero only in regions in space where the $|\omega_A(r)|^2 \, |\omega_B(\vec{r})|^2 \neq 0$, i.e., where the single-orbital densities overlap. If the densities do not overlap at all, the antisymmetrization of $\Psi(x_1 \, x_2)$ is irrelevant. In particular the states describing classical particles, which can be localized in non-overlapping positions, need not be symmetrized or antisymmetrized.

As illustrated in the figure, the electron-electron interaction is stronger in the singlet state, since the electrons are closer together as compared with the triplet state due to the antisymmetry of $\phi(\vec{r}_1 \, , \vec{r}_2)$. However, the attractive electron-nucleus interactions are stronger in the singlet state of a molecule, which leads to stronger molecular bonding (Pauli pairing).

5.5 The Heitler-London method for H_2

Heitler and London proposed in 1927 a method for determining the ground state of H_2 based on the combination of ground state wave functions $\omega_A(\vec{r}_1)$ and $\omega_B(\vec{r}_2)$ of the H atom. In atomic units we have

$$\omega_A(r) = \frac{1}{\sqrt{\pi}} \, e^{-r} \qquad \leftarrow \sqrt{\frac{\alpha^3}{\pi}} \, e^{-\alpha r} \;\; \text{optimize } \alpha$$

$$\phi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2(1 \pm S_{AB}^2)}} \left[\omega_A(\vec{r}_1) \, \omega_B(\vec{r}_2) \pm \omega_B(\vec{r}_1) \, \omega_A(\vec{r}_2) \right] ,$$

where the + sign corresponds to a total spin $S = 0$ and the − sign to $S = 1$. The single-particle wave functions are centered at the position of the nuclei $\vec{R}_a$ and $\vec{R}_b$. For example, $\omega_A(\vec{r}) = (\alpha^3/\pi)^{1/2} \exp\{-\alpha\,|\,\vec{r} - \vec{R}_a|\}$. The Hamiltonian can be written as

$$\hat{H} = \underbrace{-\frac{1}{2}\,\nabla_1^2 - \frac{1}{r_{1a}}}_{\hat{H}_1} \underbrace{-\frac{1}{2}\,\nabla_2^2 - \frac{1}{r_{2b}}}_{\hat{H}_2} - \frac{1}{r_{1b}} - \frac{1}{r_{2a}} + \frac{1}{r_{12}} + \frac{1}{R},$$

where $r_{1a} = |\vec{r}_1 - \vec{R}_a|$, $r_{1b} = |\vec{r}_1 - \vec{R}_b|$, $r_{2a} = |\vec{r}_2 - \vec{R}_a|$, $r_{2b} = |\vec{r}_2 - \vec{R}_b|$, $r_{12} = |\vec{r}_1 - \vec{r}_2|$ and $R = |\vec{R}_a - \vec{R}_b|$. The average energy $E_\pm = \langle\, \phi_\pm \,|\, \hat{H} \,|\, \phi_\pm \,\rangle$ for $\alpha = 1$ is given by

$$E_\pm = -1 + \frac{E_0 \pm E_1}{1 \pm S_{AB}^2} + \frac{1}{R},$$

where

$$\begin{aligned} E_0 &= \int |\omega_A(\vec{r}_1)|^2\, |\omega_B(\vec{r}_2)|^2 \left(-\frac{1}{r_{1b}} - \frac{1}{r_{2a}} + \frac{1}{r_{12}}\right) d^3r_1\, d^3r_2 \\ &= -\int \frac{|\omega_A(\vec{r}_1)|^2}{r_{1b}}\, d^3r_1 - \int \frac{|\omega_B(\vec{r}_2)|^2}{r_{2a}}\, d^3r_2 + \int \frac{|\omega_A(\vec{r}_1)|^2\, |\omega_B(\vec{r}_2)|^2}{r_{12}}\, d^3r_1\, d^3r_2 \end{aligned}$$

and

$$E_1 = \int \omega_A^*(\vec{r}_1)\, \omega_B(\vec{r}_1)\, \omega_A(\vec{r}_2)\, \omega_B^*(\vec{r}_2) \left(-\frac{1}{r_{1b}} - \frac{1}{r_{2a}} + \frac{1}{r_{12}}\right) d^3r_1\, d^3r_2.$$

E_0 is a single-particle term, which is weakly attractive at large distances and repulsive at short distances (including the $1/R$ term). It can be interpreted in terms of classical electrostatics, as the interaction of each atomic density with the nucleus of the other atoms, plus the electrostatic repulsion between the two atomic densities $|\omega_A(\vec{r}_1)|^2$ and $|\omega_B(\vec{r}_2)|$.

The so-called "exchange terms" involve hopping between the states A and B due to $1/r_{1b}$ and $1/r_{2a}$, as well as the repulsive term $1/r_{12}$. Notice that the electron-electron repulsion contribution is weaker in the triplet. However, the attractive terms are far more important. Therefore, the ground state is a singlet.

5.6 Symmetry with respect to interchange

In the case of two particles we have shown that the coordinate wave-function must be either symmetrical or antisymmetrical with respect to interchange of the coordinates $\vec{r}_1$ and $\vec{r}_2$:

$$\phi(r_1, r_2) = \pm\, \phi(r_2, r_1).$$

This does not result from the principle of indistinguishability, but is simply a consequence of the fact that the Hamiltonian commutes with permutation P_{12}. The particles are indeed identical and therefore, as in classical mechanics, the Hamiltonian is unchanged

by any permutation of their coordinates. The quantum mechanical indistinguishability of identical particles plays here no role. This symmetry should not be confused with the *physical interchange* of the particles, which requires the exchange of a complete sets of the dynamical variables of the particles, for instance, position $\vec{r}_i$ and spin σ_i.

In the general case of a system with N particles the solution of the Schrödinger equation need not be symmetrical or antisymmetrical with respect to interchange of the coordinates $\vec{r}_i$ of any two particles. The identity of particles implies, however, that the Hamiltonian is invariant with respect to the interchange of coordinates $\vec{r}_i \leftrightarrow \vec{r}_j$ for all i, j. Therefore, if $\phi(\vec{r}_1, \dots \vec{r}_N)$ is a solution of the Schrödinger equation and $\hat{P}$ is the operator performing a permutation of the variables, then $\hat{P}\,\phi(\vec{r}_1, \dots \vec{r}_N)$ is also a solution. This follows from the fact that $\hat{H}$ and $\hat{P}$ commute. In fact, for any function $\phi(1, \dots N)$ it holds

$$\begin{aligned}\hat{P}\,\hat{H}(1, \dots N)\;\phi(1, \dots N) &= \hat{H}(P(1), \dots P(N))\;\phi(P(1), \dots P(N)) = \\ &= \hat{H}(1, \dots N)\;\phi(P(1), \dots P(N)) \\ &= \hat{H}\,\hat{P}\,\phi(1, \dots N).\end{aligned}$$

Since all particles are identical their order does not matter in $\hat{H}$. Therefore, $[\hat{H}, \hat{P}] = 0\;\forall\;\hat{P}$. In addition, for any solution ϕ of the stationary Schrödinger equation

$$\hat{H}\,\phi = E\,\phi$$

we have

$$\hat{P}\,\hat{H}\,\phi = E\,\hat{P}\,\phi\,,$$

and

$$\hat{H}\,\hat{P}\,\phi = E\,\hat{P}\,\phi \quad ([\hat{H}, \hat{P}] = 0)\,.$$

Thus,

$$\hat{P}\,\phi\,(\vec{r}_1, \dots \vec{r}_N) = \phi\,(\vec{r}_{P(1)}, \dots \vec{r}_{P(N)})$$

is also a solution for any permutation of the variables P: $[1, N] \to [1, N]$. Let us recall that there are $N!$ permutations

$$P = \begin{pmatrix} 1 & 2 & \dots & N \\ P(1) & P(2) & \dots & P(N) \end{pmatrix}.$$

While the invariance of $\hat{H}$ under permutations implies that $[\hat{H}, \hat{P}] = 0$ for all $\hat{P}$, different permutations do not commute in general with each other. For this reason, we cannot construct all the solutions of the Schrödinger equations in such a way that they are all symmetrical, antisymmetrical, or even that they transform into themselves ($\hat{P}\,\phi \propto \phi$) for all $\hat{P}$. Mathematically, one would say that the irreducible representations of the permutation group are not all one-dimensional.

The best we can do is to determine the types of symmetries of the functions $\phi\,(\vec{r}_1, \ldots \vec{r}_N)$ under permutations of their variables. We would like to construct sets of functions

$$R = \{\phi_1(\vec{r}_1, \ldots \vec{r}_N),\ \phi_2(\vec{r}_1, \ldots \vec{r}_N),\ \ldots\ \phi_K(\vec{r}_1, \ldots \vec{r}_N)\}$$

such that

$$\hat{P}\,\phi_i(\vec{r}_1, \ldots \vec{r}_N) = \sum_{j=1}^{K} D_{ji}(P)\,\phi_j(\vec{r}_1, \ldots \vec{r}_N).$$

The linear K-dimensional subspace R spanned by these functions is called an *invariant representation subspace*. All the functions in R transform into each other under any permutation of the variables, actually into linear combinations of each other:

$$D_{ji}(P) = \int \phi_j^*(\vec{r}_1, \ldots \vec{r}_N)\ \hat{P}\,\phi_i(\vec{r}_1, \ldots \vec{r}_N)\,d^3r_1 \ldots d^3r_N$$

is the matrix of the operator $\hat{P}$ in the representation basis $\{\phi_1, \ldots \phi_K\}$. The mapping $P \to D(P)$ from the group of all permutations in the linear matrix vector space $\mathbb{C}^{K\times K}$ is a representation of the permutation group in R, since $D(\mathbb{1}) = \mathbb{1}$ and $D(P\,Q) = D(P)\,D(Q)$.

We can assume that $\{\phi_1, \ldots \phi_K\}$ is a set of orthonormal functions since the orthonormalization does not affect the invariance. The set $\{\phi_1, \ldots \phi_N\}$ is the basis of the invariant subspace. For an arbitrary state $\Psi = \sum_i \alpha_i\,\phi_i$ in R we have

$$\hat{P}\,\Psi = \sum_i \alpha_i\,\hat{P}\,\phi_i = \sum_{ij} \alpha_i\,D_{ji}(P)\,\phi_j = \sum_j \left(\sum_i D_{ji}(P)\,\alpha_i\right)\,\phi_j.$$

We are interested in the irreducible invariant subspaces, i.e., those in which the matrix $D_{ij}(P)$ cannot be separated into smaller blocks for all P. We already know two of those irreducible representations: *Symmetric functions*

$$\phi_S(\vec{r}_1, \ldots \vec{r}_N) = \sum_P \phi(\vec{r}_{P(1)}, \ldots \vec{r}_{P(N)}),$$

which satisfy $\hat{P}\,\phi_S = \phi_S\ \forall\ P$, and *antisymmetric functions*

$$\phi_A(\vec{r}_1, \ldots \vec{r}_N) = \sum_P (-1)^p\,\phi(\vec{r}_{P(1)}, \ldots \vec{r}_{P(N)}).$$

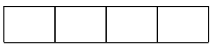
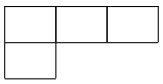
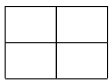
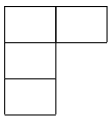
In this case, $\hat{P}\,\phi_A = (-1)^p\,\phi_A\ \forall P$, where $p = O(P)$ is the order of the permutation. These are one-dimensional (thus irreducible) representations, provided that ϕ_S and ϕ_A are not zero.

We would now like to sketch how higher dimensional representations can be constructed. The basic idea is to *symmetrize* with respect to some variables and *antisymmetrize* with

respect to the others. For this purpose we separate the variables in all possible sets $N_1, N_2, \ldots N_K$ such that $N_1 + N_2 + \ldots + N_K = N$. For example, for $N = 4$ we have

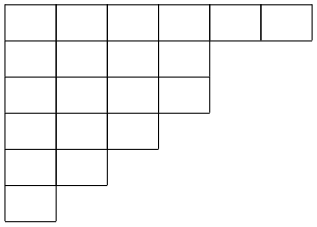
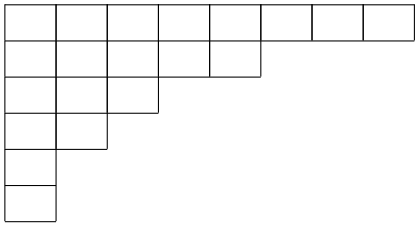
a) 4
b) $3 + 1$
c) $2 + 2$
d) $2 + 1 + 1$
e) $1 + 1 + 1 + 1$.

This is conveniently represented in the form of Young diagrams (Young 1901):

a)  b)  c)  d)  e) 

Note that each diagram has its "dual", which is obtained by exchanging rows and columns, i.e., by the transposition $a \leftrightarrow e$, $b \leftrightarrow d$ and $c \leftrightarrow c$.

To avoid ambiguity the rows are placed in order of decreasing length, so that we have a succession of rows and columns as illustrated in the following examples:

Each diagram corresponds to an irreducible representation, i.e., to a type of symmetry and thus to a given degenerate energy level. The way to proceed is the following:

1) Distribute the variables or indices among the different cells.

2) Symmetrize the (arbitrary) solution of the Schrödinger equation $\phi(\vec{r}_1, \ldots \vec{r}_N)$ with respect to the variables in each row.

3) Choose one variable from each row and antisymmetrize the wave function with respect to these variables. These variables are now "used", i.e., they are eliminated from the diagram.

4) Continue to choose one variable from each row and antisymmetrize with respect to them until only a one-cell column is left.

In this way the resulting function has been *symmetrized* with respect to the variables in each row and *antisymmetrized* with respect to the variables in each column.

Once the procedure is repeated for all possible distributions of the variables among the rows of the diagram, we obtain a set of functions which transform into linear combinations of each other when the variables are permuted in any way ($\forall\ P$). Note that the distribution of the variables among the cells in the same row is irrelevant since they are symmetrized. Each Young diagram corresponds to a given permutation symmetry. By constructing all diagrams we find all possible types of symmetry.

Example:

Let us consider a system with 3 particles. The possible separations of these 3 variables are 3, $2+1$, and $1+1+1$. The first case is the fully symmetric function and the third one is the fully antisymmetric function. We focus therefore in the non trivial case $2+1$. The possible distributions of the variables for this diagram are

1	2
3	

1	3
2	

2	3
1	

Symmetrization and antisymmetrization yields

$$
\begin{aligned}
&\mathrm{Sym}_{1\,2}, \mathrm{Asym}_{1\,3} && \Psi_1 = \phi(1,2,3) + \phi(2,1,3) - \phi(3,2,1) - \phi(2,3,1) \\
&\mathrm{Sym}_{1\,2}, \mathrm{Asym}_{2\,3} && \Psi_2 = \phi(1,2,3) + \phi(2,1,3) - \phi(1,3,2) - \phi(3,1,2) \\
&\mathrm{Sym}_{1\,3}, \mathrm{Asym}_{1\,2} && \Psi_3 = \phi(1,3,2) + \phi(3,1,2) - \phi(2,3,1) - \phi(3,2,1) \\
&\mathrm{Sym}_{1\,3}, \mathrm{Asym}_{2\,3} && \Psi_4 = \phi(1,3,2) + \phi(3,1,2) - \phi(1,2,3) - \phi(2,1,3) = -\Psi_2 \\
&\mathrm{Sym}_{2\,3}, \mathrm{Asym}_{1\,2} && \Psi_5 = \phi(2,3,1) + \phi(3,2,1) - \phi(1,3,2) - \phi(3,1,2) = -\Psi_3 \\
&\mathrm{Sym}_{2\,3}, \mathrm{Asym}_{1\,3} && \Psi_6 = \phi(2,3,1) + \phi(3,2,1) - \phi(2,1,3) - \phi(1,2,3) = -\Psi_1 .
\end{aligned}
$$

We may verify the invariance and irreducibility of the subspace. Denoting the transposition of variables ij by P_{ij}, and noting that there are only 3 different transpositions P_{12}, P_{23} and P_{13} one obtains

$$
\begin{aligned}
&P_{12}\,\Psi_1 = \Psi_2 \\
&P_{23}\,\Psi_1 = \Psi_3 \\
&P_{13}\,\Psi_1 = -\Psi_1 \\
&P_{12}\,\Psi_2 = \Psi_1 \\
&P_{23}\,\Psi_2 = -\Psi_2 \\
&P_{13}\,\Psi_2 = -\Psi_3 \\
&P_{12}\,\Psi_3 = -\Psi_3 \\
&P_{23}\,\Psi_3 = \Psi_1 \\
&P_{13}\,\Psi_3 = -\Psi_2 \\
&\text{and, even if redundant,} \\
&P \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \Psi_3 = -\Psi_3 \,.
\end{aligned}
$$

Notice that not all the functions obtained in this way are linearly independent. The subspace spanned is certainly invariant and irreducible, but the basis may be redundant and needs to be simplified.

All these orbital wave functions correspond to a degenerate level of the Schrödinger equation. The presence of degeneracies is a consequence of the fact that different permutations do not commute with each other. These orbital functions ϕ and orbital symmetries have to be combined with spin functions χ, whose symmetry under permutation must be such that the *complete wave function* $\Psi(x_1, \ldots x_N) = \phi(\vec{r}_1, \ldots \vec{r}_N)\ \chi(\sigma_1, \ldots \sigma_N)$ has the usual symmetry or antisymmetry property with respect to "physical" interchange (fermions or bosons).

Before closing this section let us revisit two simpler examples:

1) $\boxed{1}\boxed{2}\boxed{3}\boxed{4}$ $\quad\rightarrow\quad \Psi = \phi(1\,2\,3\,4) + \phi(2\,1\,3\,4) + \phi(3\,2\,1\,4) + \phi(4\,2\,3\,1) + \ldots$

This diagram gives a symmetric function that transforms into itself under all P.

2) $\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array}$ $\quad\rightarrow\quad \Psi = \phi(1\,2\,3) - \phi(2\,1\,3) - \phi(3\,2\,1) + \phi(2\,3\,1) + \phi(3\,1\,2) - \phi(1\,3\,2)$

This diagram gives a fully antisymmetric function, which transforms into $\pm$ itself according to $(-1)^p$. The symmetrical and antisymmetric cases always correspond to one-dimensional representations.

5.7 Symmetry of the spin functions $\chi(\sigma_1, \ldots \sigma_N)$

For the spin functions the symmetry adapted functions are given by the Young diagrams with $\sigma_1, \ldots \sigma_N$ taking the place of $\vec{r}_1, \ldots, \vec{r}_N$.

5.7.1 Examples of spin Young diagrams for $S = 1/2$

1) $\boxed{+}\boxed{-}$ $\quad\rightarrow\quad |+\rangle_1 |-\rangle_2 + |-\rangle_1 |+\rangle_2 \quad\rightarrow\quad$ triplet $(S = 1)$

2) $\begin{array}{|c|}\hline + \\ \hline - \\ \hline \end{array}$ $\quad\rightarrow\quad |+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2 \quad\rightarrow\quad$ singlet $(S = 0)$

3) $\begin{array}{|c|c|}\hline + & - \\ \hline + \\ \cline{1-1}\end{array}$ $\rightarrow$ $\Big(|+\rangle_1|-\rangle_2+|-\rangle_1|+\rangle_2\Big)|+\rangle_3-|+\rangle_1|+\rangle_2|-\rangle_3-|-\rangle_1|+\rangle_2|+\rangle_3$

$= |+\rangle_1\Big(|-\rangle_2|+\rangle_3-|+\rangle_2|-\rangle_3\Big)$

$\rightarrow$ $= |+\rangle_1|-\rangle_2|+\rangle_3+|-\rangle_1|+\rangle_2|+\rangle_3-|+\rangle_1|-\rangle_2|+\rangle_3-|+\rangle_1|+\rangle_2|-\rangle_3$

$= |-\rangle_1|+\rangle_2|+\rangle_3-|+\rangle_1|+\rangle_2|-\rangle_3.$

This is a doublet ($S = 1/2,\ S_z = 1/2$), as can be easily verified by applying $\hat{S}^+ = \hat{s}_1^+ + \hat{s}_2^+ + \hat{s}_3^+$.

4) The Young diagram for $S = 3/2$ can be simply written as

$\begin{array}{|c|c|c|}\hline \ \ & \ \ & \ \ \\ \hline\end{array}$ $\rightarrow$ S=3/2 ,

where we have omitted specifying all possible choices of the values "+" and "−" of the spin wave functions, since this only affects the value of the S_z. The spin wave function corresponding to this diagram is always symmetric.

Actually, the addition angular momentum can be written with the help of Young diagrams as:

$\underset{1/2}{\begin{array}{|c|}\hline \ \ \\ \hline\end{array}} \otimes \underset{1/2}{\begin{array}{|c|}\hline \ \ \\ \hline\end{array}} = \underset{1}{\begin{array}{|c|c|}\hline \ \ & \ \ \\ \hline\end{array}} \oplus \underset{0}{\begin{array}{|c|}\hline \ \ \\ \hline \ \ \\ \hline\end{array}}$,

$\underset{1}{\begin{array}{|c|c|}\hline \ \ & \ \ \\ \hline\end{array}} \otimes \underset{1/2}{\begin{array}{|c|}\hline \ \ \\ \hline\end{array}} = \underset{3/2}{\begin{array}{|c|c|c|}\hline \ \ & \ \ & \ \ \\ \hline\end{array}} \oplus \underset{1/2}{\begin{array}{|c|c|}\hline \ \ & \ \ \\ \hline \ \ \\ \cline{1-1}\end{array}}$,

and

$\underset{0}{\begin{array}{|c|}\hline \ \ \\ \hline \ \ \\ \hline\end{array}} \otimes \underset{1/2}{\begin{array}{|c|}\hline \ \ \\ \hline\end{array}} = \underset{1/2}{\begin{array}{|c|c|}\hline \ \ & \ \ \\ \hline \ \ \\ \cline{1-1}\end{array}}$.

There is no $\begin{array}{|c|}\hline \ \ \\ \hline \ \ \\ \hline \ \ \\ \hline\end{array}$ since for spin 1/2 two cells are necessarily equal, which leads to a vanishing function after antisymmetrization.

5.7.2 Examples of spin functions for $3e^-$

$$\boxed{+}\boxed{+}\boxed{+} \quad \rightarrow \quad \chi = |+\rangle_1 |+\rangle_2 |+\rangle_3 \qquad S = 3/2,\ S_z = 3/2$$

$$\boxed{+}\boxed{+}\boxed{-} \quad \rightarrow \quad \chi = |+\rangle_1 |+\rangle_2 |-\rangle_3 + |+\rangle_1 |-\rangle_2 |+\rangle_3 + |-\rangle_1 |+\rangle_2 |+\rangle_3 \qquad S = 3/2,\ S_z = 1/2$$

$$\boxed{+}\boxed{-}\boxed{-} \quad \rightarrow \quad \chi = (+--) + (-+-) + (--+) \qquad S = 3/2,\ S_z = 1/2$$

$$\boxed{-}\boxed{-}\boxed{-} \quad \rightarrow \quad \chi = (---) \qquad S = 3/2,\ S_z = -3/2$$

Another possible diagram for $S_z = 1/2$ is

$$\begin{array}{|c|c|}\hline + & + \\ \hline - \\ \cline{1-1}\end{array} \quad \rightarrow \quad \begin{cases} \chi_1 = |+\rangle_1 |+\rangle_2 |-\rangle_3 - |-\rangle_1 |+\rangle_2 |+\rangle_3 \\ \chi_2 = |+\rangle_1 |+\rangle_2 |-\rangle_3 - |+\rangle_1 |-\rangle_2 |+\rangle_3 \end{cases} \qquad S = 1/2,\ S_z = 1/2$$

In one case we have antisymmetrized $1 \leftrightarrow 3$ and in the other $2 \leftrightarrow 3$. These are the two linearly independent solutions for $S_z = \dfrac{1}{2}$ and $S = \dfrac{1}{2}$.

One can also try $\begin{array}{|c|c|}\hline - & + \\ \hline + \\ \cline{1-1}\end{array}$ which gives only one non-vanishing function

$$\begin{aligned} \chi &= |-\rangle_1 |+\rangle_2 |+\rangle_3 + |+\rangle_1 |-\rangle_2 |+\rangle_3 - |+\rangle_1 |+\rangle_2 |-\rangle_3 - |+\rangle_1 |-\rangle_2 |+\rangle_3 \\ &= |-\rangle_1 |+\rangle_2 |+\rangle_3 - |+\rangle_1 |+\rangle_2 |-\rangle_3 \\ &= -\chi_1 \quad \leftarrow \text{ linearly dependent.} \end{aligned}$$

Note that χ_1 and χ_2 form an irreducible representation since $P_{12}\,\chi_1 = \chi_2$, $P_{23}\,\chi_1 = \chi_1 - \chi_2$, $P_{13}\,\chi_1 = -\chi_2$, $P_{12}\,\chi_2 = \chi_1$, $P_{23}\,\chi_2 = -\chi_2$, and $P_{13}\,\chi_2 = -\chi_1$.

5.8 Symmetry of the complete wave function $\Psi(x_1, \ldots x_N)$

The question is what Young diagram of the spin function corresponds to a given diagram of the coordinate function, so that the complete wave function has the right symmetry. The choice depends of course on the boson or fermion character of the particles.

5.8.1 Bosons

$\Psi(x_1, \ldots x_N)$ must be symmetric. For these to be so, one has to use the same Young diagrams for the coordinate and spin parts. The complete wave function is given by definite bilinear combinations of the two:

$$\Psi(x_1, \ldots x_N) = \sum_{ij} A_{ij}\ \phi_i(\vec{r}_1, \ldots \vec{r}_N)\ \chi_j(\sigma_1, \ldots \sigma_N),$$

where ϕ_i and χ_j belong to the same irreducible representation.

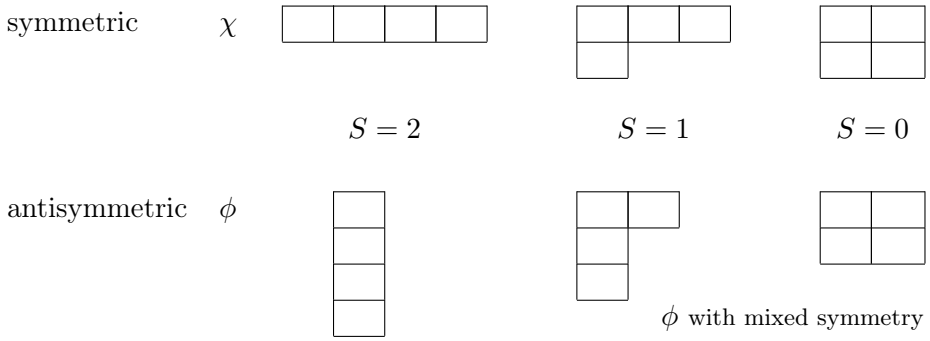
5.8.2 Fermions

$\Psi(x_1, \ldots x_N)$ must be antisymmetrical with respect to all variables. This is obtained by combining coordinate functions and spin functions obtained from Young diagrams that are the "dual" of each other, i.e., the diagram of the spin part is obtained by interchanging rows and columns of the coordinate diagram.

5.8.3 Spin 1/2 particles

The spin variables can take only two values $\sigma_i = \pm 1/2$. Therefore the Young diagrams can have only one or two rows at most, since antisymmetrization of equal variables gives zero. The number of possible permutation symmetries is equal to the number of partitions of $N = N_1 + N_2$ with $N_2 \leq N_1$, i.e., $N_2 = 0, 1, \ldots N/2$ for N even and $N_2 = 0, 1, \ldots (N-1)/2$ for N odd.

For spin $1/2$ each type of symmetry, each Young diagram, corresponds to a definite total spin S of the system. For example, for $N = 4$



A coordinate diagram with n columns corresponds to $S = n/2$. The relation between permutation symmetry and total spin S holds only for spin $1/2$ particles.

5.9 Complete wave function for N identical bosons

The purpose of this section is to provide an explicit form for the complete wave-function of N identical bosons. Our starting point is an irreducible representation of the permutation group with dimension f which is given by the basis $\{\phi_1, \phi_2, \ldots \phi_f\}$. The

$\phi_i = \phi_i(\vec{r}_1, \ldots \vec{r}_N)$ are orthonormal, and transform under permutations like

$$\hat{P}\,\phi_i = \sum_j D_{ji}(P)\,\phi_j.$$

Given an arbitrary coordinate wave function $\Psi = \sum_k \alpha_k\,\phi_k$ we have

$$\hat{P}\,\Psi = \sum_k \alpha'_k\,\phi_k,$$

where

$$\begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_f \end{pmatrix} = D(P) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_f \end{pmatrix}. \tag{5.4}$$

The validity of this matrix equation can be verified by noting that

$$\hat{P}\,\Psi = \sum_k \alpha_k\,\hat{P}\,\phi_k = \sum_k \alpha_k \sum_j D_{jk}(P)\,\phi_j = \sum_j \underbrace{\left[\sum_k D_{jk}(P)\,\alpha_k\right]}_{\alpha'_j} \phi_j.$$

Together with $\{\phi_1, \ldots \phi_f\}$ we have an irreducible representation of the spin functions of dimension $f = 2s + 1$ given by $\{\chi_1, \chi_2, \ldots \chi_f\}$ with $\chi_i = \chi_i(\sigma_1, \ldots \sigma_N)$ that transform under permutation like

$$P\,\chi_i = \sum_j D_{ji}(P)\,\chi_j.$$

For an arbitrary $\chi = \sum_k \beta_k\,\chi_k$ we have

$$P\,\chi = \sum_k \beta'_k\,\chi_k,$$

where

$$\begin{pmatrix} \beta'_1 \\ \vdots \\ \beta'_f \end{pmatrix} = D(P) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_f \end{pmatrix}. \tag{5.5}$$

The complete wave function Ψ of the coordinates and spin must be symmetrical since we are dealing with bosons. This means that Ψ must be invariant with respect to

permutations or, in more geometrical terms, it must behave like a "scalar". Moreover, according to Eqs. (5.4) and (5.5), we know that the components $\alpha_1, \ldots, \alpha_f$ of the coordinate wave-function ϕ and the components $\beta_1, \ldots, \beta_f$ of the spin wave-function χ transform like "vectors", where $D(P)$ is the "rotation" matrix. And how many scalars can be constructed with the products of the components of two vectors in an f-dimensional linear space? In fact, only one scalar can be constructed with two vectors, namely the scalar product. Therefore, we have that the complete boson wave-function is given by

$$\Psi = \sum_{i=1}^{f} \phi_i^*(\vec{r}_1, \ldots \vec{r}_N)\ \chi_i(\sigma_1, \ldots \sigma_N).$$

In order to prove that Ψ is invariant, one should recall that

$$D(P) \text{ is unitary} \quad \Leftrightarrow \quad [D(P)]^{-1} = D(P)^\dagger \tag{5.6}$$

and that

$$D(P)^{-1} = D(P^{-1}). \tag{5.7}$$

Eq. (5.6) is clear, since the permutation of variables does not change the norm of any coordinate or spin wave-function. Eq. (5.7) follows from the representation properties of D [i.e., $D(\mathbb{1}) = \mathbb{1}$ and $D(PQ) = D(P)\ D(Q)$] and the fact that $\hat{P}\,\hat{P}^{-1} = \mathbb{1}$. Then we have

$$\begin{aligned}
\hat{P}\,\Psi &= \sum_{i=1}^{f} \hat{P}_{\text{coord}}\,\phi_i^*\ \hat{P}_{\text{spin}}\,\chi_i = \\
&= \sum_{i=1}^{f} \left[\sum_{j=1}^{f} D_{ji}(P)\,\phi_j\right]^* \left[\sum_{k=1}^{f} D_{ki}(P)\,\chi_k\right] \\
&= \sum_{k,j=1}^{f} \left[\sum_{i=1}^{f} D_{ji}^*(P)\ D_{ki}(P)\right] \phi_j^*\,\chi_k.
\end{aligned}$$

Using that $D(P) = [D(P^{-1})]^\dagger$, or equivalently $D_{ji}^*(P) = D_{ij}(P^{-1})$, one obtains

$$\begin{aligned}
\hat{P}\,\Psi &= \sum_{kj} \left[\sum_{i=1}^{f} D_{ij}(P^{-1})\ D_{ki}(P)\right] \phi_j^*\,\chi_k \\
&= \sum_{kj} \underbrace{\left[D(P)\ D(P^{-1})\right]_{kj}}_{\delta_{kj}}\ \phi_j^*\,\chi_k \\
&= \sum_{k} \phi_k^*\,\chi_k = \Psi.
\end{aligned}$$

One can show (either by using the projection and orthogonality relations of group-representation theory or by using Schur's Lemma) that the scalar product is the only way of constructing a scalar with the direct product the two f-dimensional representations.

The situation is known from the addition of two angular momenta $L = S$. In this case J runs from $0 = |L - S| \leq J \leq L + S$ and there is only one possible $J = 0$ state (scalar under rotations).

$$\begin{aligned} &L + S \qquad (k = 0) \\ &L + S - k,\ L - 1 + S - (k - 1),\ L - 2 + S - (k - 2), \ldots L - k + S. \end{aligned}$$

There are $k + 1$ terms that are linearly independent; k of them belong to the higher J's. We have only one state per value of J and only one of them is an invariant state with $J = 0$.

5.10 Invariant bilinear combination of $\{\phi_i\}$ and $\{\chi_i\}$: The scalar product[11]

Consider two irreducible representations $\{\phi_1, \ldots \phi_{f_\alpha}\}$ and $\{\chi_1, \ldots \chi_{f_\beta}\}$ of a group (e.g., the permutation group) having dimensions f_α and f_β:

$$\hat{P}\,\phi_i = \sum_{j=1}^{f_\alpha} D^{(\alpha)}_{ji}(P)\,\phi_j$$

and

$$\hat{P}\,\chi_k = \sum_{j=1}^{f_\beta} D^{(\beta)}_{jk}(P)\,\chi_j.$$

The direct product $\{\phi_i^*\,\chi_k$ with $1 \leq i \leq f_\alpha$ and $1 \leq k \leq f_\beta\}$ is a representation with representation matrix given by the tensor product of $D^{(\alpha)*}$ and $D^{(\beta)}$:

$$\begin{aligned} P\,\phi_i^*\,\chi_k &= \sum_j D^{(\alpha)*}_{ji}(P)\ \phi_j^* \sum_l D^{(\beta)}_{lk}(P)\,\chi_l \\ &= \sum_{jl} D^{(\alpha)*}_{ji}(P)\ D^{(\beta)}_{lk}(P)\ \phi_j^*\,\chi_l. \end{aligned} \tag{5.8}$$

We know that

$$\left.\begin{aligned} D_{ij}(P)^\dagger \quad &= [D_{ij}(P)]^{-1} = D_{ij}(P^{-1}) \\ &= D^*_{ji}(P) \end{aligned}\right\}\ D^*_{ji}(P) = D_{ij}(P^{-1})$$

[11] Sections 5.10 and 5.11 should be skipped, unless the reader is particularly interested in the group theoretical background.

and therefore

$$P\,\phi_i^*\,\chi_k = \sum_{jl} D_{ij}^{(\alpha)}(P^{-1})\; D_{lk}^{(\beta)}(P)\;\phi_j^*\,\chi_l.$$

Let us consider the case where $\alpha \equiv \beta$ and let us search for an *arbitrary bilinear combination* of ϕ_i^* and χ_k that is *invariant for all* P. Let

$$\Psi = \sum_{ik} \lambda_{ki}\,\phi_i^*\,\chi_k \quad \text{with arbitrary } \lambda_{ki} \in \mathbb{C},$$

it follows that

$$P\,\Psi = \sum_{jl} \left(\sum_{ik} D_{lk}(P)\;\lambda_{ki}\;D_{ij}(P^{-1}) \right) \phi_j^*\,\chi_l.$$

The condition $\hat{P}\,\Psi = \Psi$ implies

$$\sum_{ik} D_{lk}(P)\;\lambda_{ki}\;D_{ij}(P^{-1}) = \lambda_{lj}.$$

Regarding D and λ as matrices ($D(P^{-1}) = D(P)^{-1}$) we have

$$D(P)\,\lambda\,D(P)^{-1} = \lambda \quad \Leftrightarrow \quad D(P)\,\lambda = \lambda\,D(P)$$

for all P. By *Schur's Lemma* we know that a matrix that commutes with *all* the elements of an *irreducible representation* must be proportional to the identity matrix.[12] This implies

$$\lambda_{ki} = e^{i\,\varphi}\,\delta_{ki}$$

and consequently the scalar product is the only invariant bilinear in ϕ_j^* and χ_l.

Alternative proof using group theory:

Given a representation of a group G having dimension f and character $\chi(G)$ we have $f = \sum_\alpha a^{(\alpha)}\,f_\alpha$, where f_α is the dimension of the irreducible representation α and $a^{(\alpha)}$ is the number of times α appears in the decomposition of the representation of dimension f. One can show using the orthogonality relations of representation theory that

$$a^{(\alpha)} = \frac{1}{g} \sum_G \chi(G)\,\chi^{(\alpha)}(G),$$

where g is the order of the group and the sum runs over all the elements of the group (see, for example, Ref. [3]).

[12] See, for instance, H. Boerner, Darstellungen von Gruppen (Springer, Berlin, 1967).

For the permutation group the number of times the symmetrical one-dimensional representation appears in the direct product is given by

$$a^{(0)} = \frac{1}{N!} \sum_P \chi(P)\,, \tag{5.9}$$

since $\chi^{(0)}(P) = 1 \;\; \forall\, P$. For the direct product of $D^{*(\alpha)}_{ji}(P)$ and $D^{(\alpha)}_{ji}(P)$ the character is

$$\chi(P) = \chi^{(\alpha)}(P)^*\, \chi^{(\alpha)}(P) = |\chi^{(\alpha)}(P)|^2. \tag{5.10}$$

This can be shown by noting, from Eq. (5.8), that the matrix elements of the product representation $D_\otimes$ are given by

$$\langle\, j\, l \,|\, D_\otimes(P) \,|\, i\, k \,\rangle = D^{(\alpha)}_{ji}(P)^*\, D^{(\alpha)}_{lk}(P)$$

The character $\chi(P)$ of $D_\otimes(P)$ is then given by

$$\begin{aligned}\chi(P) &= \sum_{ik} D^{(\alpha)}_{ii}(P)^*\, D^{(\alpha)}_{kk}(P) \\ &= \chi^{(\alpha)}(P)^*\, \chi^{(\alpha)}(P) \\ &= |\chi^{(\alpha)}(P)|^2\,.\end{aligned}$$

Replacing Eq. (5.10) in Eq. (5.9) we obtain

$$a^{(0)} = \frac{1}{N!} \sum_P |\chi^{(\alpha)}(P)|^2 = 1$$

for any irreducible representation α. Here we have used the "normalization" relation

$$\sum_G |\chi^{(\alpha)}(G)|^2 = g\,, \tag{5.11}$$

where G is any group of order g. Let us recall that Eq. (5.11) holds if and only if α is irreducible. Consequently, there is only one symmetric state in the direct product of coordinate and spin functions obtained from a given irreducible representation.

5.11 Symmetry projection operators

Group theory allows us to express an arbitrary function as the sum of functions which transform according to the irreducible representations of a given group (see Ref. [3], p. 336):

$$\Psi = \sum_\alpha \sum_i \Psi^{(\alpha)}_i,$$

where

$$\Psi^{(\alpha)}_i = \frac{f_\alpha}{g} \sum_G G^{(\alpha)}_{ii}\, \hat{G}\, \Psi. \tag{5.12}$$

If we are looking for the projection $\Psi_i^{(0)}$ on the symmetric state $\hat{P}\,\Psi_i^{(0)} = \Psi_i^{(0)}$, we have $G_{ii}^{(0)} = 1$ and $f_0 = 1$. Thus,

$$\Psi_i^{(0)} = \frac{1}{g} \sum_G \hat{G}\,\Psi,$$

a result one could have guessed from the start. Let us consider any simple product

$$\Psi_{ik} = \phi_i^{(\alpha)*}\,\chi_k^{(\beta)},$$

to which one applies a group transformation $\hat{G}$ that yields

$$\hat{G}\,\phi_i^{*(\alpha)}\,\chi_k^{(\beta)} = \sum_{lm} G_{li}^{(\alpha)*}\,G_{mk}^{(\beta)}\,\chi_m^{(\beta)}\,\phi_l^{(\alpha)*}\,.$$

Using Eq. (5.12) one obtains

$$\Psi_{ik}^{(0)} = \frac{1}{g} \sum_{lm} \phi_l^{(\alpha)*}\,\chi_m^{(\beta)} \sum_G G_{li}^{(\alpha)*}\,G_{mk}^{(\beta)}\,. \tag{5.13}$$

The fundamental orthogonality relations of representation theory are (Ref. [3], p. 333)

$$\sum_G G_{li}^{(\alpha)*}\,G_{mk}^{(\beta)} = \frac{g}{f_\alpha}\,\delta_{\alpha\beta}\,\delta_{lm}\,\delta_{ik}. \tag{5.14}$$

Replacing Eq. (5.14) in Eq. (5.13) one obtains

$$\Psi_{ik}^{(0)} = \frac{\delta_{\alpha\beta}}{f_\alpha} \sum_{lm} \delta_{ik}\,\delta_{lm}\,\phi_l^{(\alpha)*}\,\chi_m^{(\beta)}$$

and finally

$$\Psi_{ik}^{(0)} = \frac{\delta_{\alpha\beta}\,\delta_{ik}}{f_\alpha} \sum_l \phi_l^{(\alpha)*}\,\chi_l^{(\alpha)}.$$

In conclusion, in order to obtain a non-vanishing result, we must take the *same irreducible representation* $\alpha = \beta$, and the functions ϕ_i and χ_k must transform according to the *same column* of the irreducible representation $i = k$. The only fully symmetrical bilinear combination is the scalar product. Note that this is the same situation as in the addition of angular momentum where $|L - S| \leq J \leq L + S$. In this case one also needs $L = S$ in order to get $J = 0$ (i.e., a rotational invariant state).

6 The Hartree-Fock approximation

We focus here on the many-body electron problem. We have shown that given a single-particle basis $\{\phi_{1\sigma}(\vec{r}), \phi_{2\sigma}(\vec{r}), \ldots\}$ one can construct a complete many-body basis set for the N-electron problem by considering the Slater determinants

$$|\, n_1, n_2, \ldots \rangle = \prod_{\alpha=1}^{\infty} (c^\dagger_{\alpha\uparrow})^{n_{\alpha\uparrow}} \prod_{\beta=1}^{\infty} (c^\dagger_{\beta\downarrow})^{n_{\beta\downarrow}} \,|\, 0 \,\rangle.$$

Therefore, one may seek the solution of Schrödinger equation by considering linear combinations of Slater determinants. One would of course like that the number of Slater determinants with significant contribution remains as limited as possible and, at the same time, one would like to achieve a physical interpretation of these main "configurations".

The idea of the Hartree-Fock method is to focus on *one Slater determinant* denoted by $|\, HF \,\rangle$ and to search for the set of N single-particle orbitals that provide the best possible description of the ground state using just $|\, HF \,\rangle$. Having no other *a priori* information, the notion of best Slater determinant is defined by the requirement that the approximation $E_{HF} = \langle\, HF \,|\, \hat{H} \,|\, HF \,\rangle$ to the ground-state energy be minimal.

In many cases of interest in current research, the single-Slater-determinant state does not provide a satisfactory description of the ground state, even at a qualitative level. These situations are known as strongly correlated or intrinsically multiconfigurational. Nevertheless, the Hartree-Fock approximation has proven to be extremely useful in atomic and molecular physics, in a large variety of chemistry problems and even in condensed-matter physics. It provides a basis for further improvements and it allows to introduce the concepts of self-consistent field and electron correlations.

6.1 The Hartree-Fock energy E_{HF}

Consider the Hamilton operator H in second quantization using the single-particle basis $\{\phi_{i\sigma}\}$

$$H = \sum_{\substack{ij \\ \sigma}} (T^\sigma_{ij} + V^\sigma_{ij})\; c^\dagger_{i\sigma}\, c_{j\sigma} + \frac{1}{2} \sum_{\substack{ijkl \\ \sigma\sigma'}} W^{\sigma\sigma'}_{ijkl}\; c^\dagger_{i\sigma}\, c^\dagger_{j\sigma'}\, c_{l\sigma'}\, c_{k\sigma},$$

where

$$T^\sigma_{ij} = \int \phi^*_{i\sigma}(\vec{r})\; \frac{p^2}{2m}\; \phi_{j\sigma}(\vec{r})\, d^3r,$$

$$V^\sigma_{ij} = \int \phi^*_{i\sigma}(\vec{r})\; v_\sigma(\vec{r})\; \phi_{j\sigma}(\vec{r})\, d^3r,$$

and

$$W^{\sigma\sigma'}_{ijkl} = \int\int \phi^*_{i\sigma}(\vec{r})\; \phi^*_{j\sigma'}(\vec{r}\,')\; \frac{e^2}{|\vec{r} - \vec{r}\,'|}\; \phi_{k\sigma}(\vec{r})\; \phi_{l\sigma'}(\vec{r}\,')\, d^3r\, d^3r'.$$

The external potential is given by $v_\sigma = v(\vec{r}) - \sigma\mu_B\, 2B(\vec{r})$ with $\mu_B = \dfrac{e\hbar}{2mc} = 0,927 \times 10^{-20}$ erg/Gauss and $\sigma = \pm 1/2$. For simplicity we have assumed that the magnetic field is collinear along the spin-quantization direction.

The Hartree-Fock state or Slater determinant is given by

$$|\, HF \,\rangle = \left(\prod_{\alpha_{\rm occ}} c^\dagger_\alpha \right) |\, 0 \,\rangle = \left(\prod_{\alpha=1}^{n_+} c^\dagger_{\alpha\uparrow} \right) \left(\prod_{\beta=1}^{n_-} c^\dagger_{\beta\downarrow} \right) |\, 0 \,\rangle$$

and the average energy by

$$E_{HF} = \langle\, HF \,|\, H \,|\, HF \,\rangle = \langle\, HF \,|\, \hat{T} + \hat{V} \,|\, HF \,\rangle + \langle\, HF \,|\, \hat{W} \,|\, HF \,\rangle,$$

where

$$\langle\, HF \,|\, \hat{T} + \hat{V} \,|\, HF \,\rangle = \sum_\sigma \sum_{i=1}^{n_\sigma} \left(T^\sigma_{ii} + V^\sigma_{ii} \right),$$

$$\langle\, HF \,|\, \hat{W} \,|\, HF \,\rangle = \frac{1}{2} \sum_{\substack{ij \\ \sigma\sigma'}} \left(W^{\sigma\sigma'}_{ijij} - \delta_{\sigma\sigma'}\, W^{\sigma\sigma}_{ijji} \right),$$

$$\langle\, HF \,|\, \hat{T} \,|\, HF \,\rangle = \sum_\sigma \sum_{i=1}^{n_\sigma} \int \phi^*_{i\sigma}(\vec{r}) \left(-\frac{\hbar^2}{2m}\, \nabla^2\, \phi_{i\sigma}(\vec{r}) \right) d^3r,$$

and

$$\langle\, HF \,|\, \hat{V} \,|\, HF \,\rangle = \sum_\sigma \int d^3r\; v_\sigma(\vec{r}) \sum_{i=1}^{n_\sigma} |\phi_{i\sigma}(\vec{r})|^2 = \sum_\sigma \int d^3r\; v_\sigma(\vec{r})\; n_\sigma(\vec{r})\, d^3r.$$

Here we have introduced the spin-resolved electron density

$$n_\sigma(\vec{r}) = \sum_{i=1}^{n_\sigma} |\phi_{i\sigma}(\vec{r})|^2 \qquad [n(\vec{r}) = n_\uparrow(\vec{r}) + n_\downarrow(\vec{r})]\,.$$

There are two types of terms resulting from the electron-electron interaction. The *direct terms* are given by

$$W_{ijkl}\; \underbrace{c^\dagger_{i\sigma}\; \overbrace{c^\dagger_{j\sigma'}\; c_{l\sigma'}}\; c_{k\sigma}}$$

for $i = k$ and $j = l$. Their contribution to the energy is

$$\begin{aligned} E_H &= \frac{1}{2} \sum_{\substack{ij \\ \sigma\sigma'}} \int \phi^*_{i\sigma}(\vec{r})\, \phi^*_{j\sigma'}(\vec{r}\,') \, \frac{e^2}{|\vec{r} - \vec{r}\,'|} \, \phi_{j\sigma'}(\vec{r}\,')\, \phi_{i\sigma}(\vec{r})\, d^3r\, d^3r' \\ &= \frac{e^2}{2} \sum_{\sigma\sigma'} \int \frac{n_\sigma(\vec{r})\, n_{\sigma'}(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} \, d^3r\, d^3r' \\ &= \frac{e^2}{2} \int d^3r\, d^3r' \, \frac{n(\vec{r})\, n(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} \\ &= \frac{1}{2} \sum_{\substack{ij \\ \sigma\sigma'}} U_{ij}^{\sigma\sigma'}. \end{aligned}$$

This is known as *Hartree energy* or classical Coulomb interaction energy.

In addition we have the *exchange terms* that result from the contraction $i\sigma = l\sigma'$ and $j\sigma' = k\sigma$ in

$$W_{ijkl}\, c^\dagger_{i\sigma} \overbrace{c^\dagger_{j\sigma'}\, c_{l\sigma'}}\, c_{k\sigma},$$

which is only possible for $\sigma = \sigma'$. Their contribution to the energy is

$$\begin{aligned} E_X &= -\frac{1}{2} \sum_{\substack{ij \\ \sigma}} \int \phi^*_{i\sigma}(\vec{r})\, \phi^*_{j\sigma}(\vec{r}\,') \, \frac{e^2}{|\vec{r} - \vec{r}\,'|} \, \phi_{i\sigma}(\vec{r}\,')\, \phi_{j\sigma}(\vec{r})\, d^3r\, d^3r' \\ &= -\frac{1}{2} \sum_{\substack{ij \\ \sigma}} J_{ij}^{\sigma}. \end{aligned}$$

As we shall see, this term is always stabilizing ($E_X < 0$), since $J_{ij} \geq 0 \;\forall\; i, j$.

A few important properties deserve to be noted:

i) The *exchange integrals* J_{ij}^σ satisfy

$$\left(J_{ij}^\sigma\right)^* = \int \phi_{i\sigma}(\vec{r})\, \phi_{j\sigma}(\vec{r}\,') \, \frac{e^2}{|\vec{r} - \vec{r}\,'|} \, \phi^*_{i\sigma}(\vec{r}\,')\, \phi^*_{j\sigma}(\vec{r})\, d^3r\, d^3r' = J_{ij}^\sigma$$

and are therefore all real.

ii) $J_{ii}^\sigma = U_{ii}^{\sigma\sigma}$, which implies that the exchange terms cancel exactly the self-interaction terms that are actually present in the classical direct interaction ($i = j, \sigma = \sigma'$).

iii) One can show that all the exchange integrals J_{ij}^{σ} are positive. This means that the exchange energy is always negative. The antisymmetry of the N-electron wave function tends to reduce the Coulomb energy with respect to the classical Hartree term. This effect is so important that it leads to the first Hund's rule in atoms (maximum S for a given shell) and the ferromagnetism of transition metals like Fe, Co and Ni.

In order to prove that $J_{ij}^{\sigma} \geq 0$ it is useful to write the exchange integral as

$$J_{ij}^{\sigma} = e^2 \int \frac{F(\vec{r})\; F^*(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d^3r\, d^3r',$$

where $F(\vec{r}) = \phi_{i\sigma}^*(\vec{r})\; \phi_{j\sigma}(\vec{r})$. For simplicity the indices $ij\sigma$ in $F(\vec{r})$ have been dropped. J_{ij} is then given by

$$J_{ij}^{\sigma} = e^2 \int \Phi(\vec{r})\; F(\vec{r})\, d^3r,$$

where

$$\Phi(\vec{r}) = \int \frac{F^*(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d^3r'.$$

Knowing that $\nabla^2 \left(\dfrac{1}{|\vec{r} - \vec{r}\,'|} \right) = -4\pi\, \delta(\vec{r} - \vec{r}\,')$ we have

$$\nabla^2\, \Phi(\vec{r}) = -4\pi\, F^*(\vec{r})$$

and

$$J_{ij}^{\sigma} = -\frac{e^2}{4\pi} \int \Phi(\vec{r})\; \nabla^2\; \Phi^*(\vec{r})\, d^3r.$$

Using the relation $\vec{\nabla} \cdot (\phi\, \vec{\nabla}\psi) = \vec{\nabla}\phi \cdot \vec{\nabla}\psi + \phi\, \nabla^2\psi$ for $\phi = \psi = \Phi$ we have

$$J_{ij}^{\sigma} = \frac{e^2}{4\pi} \left[\int |\vec{\nabla}\, \Phi(\vec{r})|^2\, d^3r - \underbrace{\int \vec{\nabla} \cdot (\Phi\, \vec{\nabla}\Phi)\; d^3r}_{\substack{=0 \text{ for any} \\ \text{localized basis}}} \right] = \frac{e^2}{4\pi} \int |\vec{\nabla}\, \Phi(\vec{r})|^2\, d^3r.$$

Consequently, $J_{ij}^{\sigma} \geq 0$ for all $ij\sigma$.

6.2 The Hartree-Fock equations

The previous expression for the energy $E_{HF} = \langle\, HF \,|\, H \,|\, HF \,\rangle$ is valid for any Slater determinant state formed by superposition of N orbitals $\phi_{i\sigma}$. Let us now determine the optimal set of occupied orbitals by minimizing E_{HF} with respect to $\phi_{i\sigma}$. For this we have to impose the orthogonality and normalization condition on the $\phi_{i\sigma}$. Therefore we seek for the stationary points of

$$F = E_{HF} - \sum_{\substack{ij \\ \sigma}} \varepsilon_{ij}^{\sigma} \left(\langle\, \phi_{i\sigma} \,|\, \phi_{j\sigma} \,\rangle - \delta_{ij} \right)$$

that are given by

$$\delta E_{HF} - \sum_{\substack{ij \\ \sigma}} \varepsilon_{ij}^{\sigma}\, \delta \langle\, \phi_{i\sigma} \,|\, \phi_{j\sigma} \,\rangle = 0. \tag{6.1}$$

It is easy to show that the matrix of Lagrange multipliers $\varepsilon_{ij}^{\sigma}$ is hermitic. Since E_{HF} is real for all $\phi_{i\sigma}$, δE_{HF} is real. Complex conjugation of Eq. (6.1) yields

$$\delta E_{HF} - \sum_{\substack{ij \\ \sigma}} \varepsilon_{ij}^{\sigma *}\, \delta \langle\, \phi_{i\sigma} \,|\, \phi_{j\sigma} \,\rangle^{*} = 0.$$

Using that $\langle\, \phi_{i\sigma} \,|\, \phi_{j\sigma} \,\rangle^{*} = \langle\, \phi_{j\sigma} \,|\, \phi_{i\sigma} \,\rangle$, and changing the summation indices we can write

$$\delta E_{HF} - \sum_{\substack{ij \\ \sigma}} \varepsilon_{ji}^{\sigma *}\, \delta \langle\, \phi_{i\sigma} \,|\, \phi_{j\sigma} \,\rangle = 0. \tag{6.2}$$

Subtracting Eqs. (6.1) and (6.2) we obtain

$$\sum_{\substack{ij \\ \sigma}} \left[\varepsilon_{ij}^{\sigma} - \varepsilon_{ij}^{\sigma *} \right] \delta \langle\, \phi_{i\sigma} \,|\, \phi_{j\sigma} \,\rangle = 0.$$

This implies $\varepsilon_{ij}^{\sigma} = \varepsilon_{ji}^{\sigma *}\ \forall\, ij\,\sigma$ since $\delta \langle\, \phi_{i\sigma} \,|\, \phi_{j\sigma} \,\rangle$ is arbitrary.

An important property of a Slater determinant is its invariance with respect to any unitary transformation among the occupied orbitals. Physically, this means that in a many-body state formed by the superposition of N single-particle states, the precise form of each individual state is not important, but it is rather the subspace spanned by these N states what matters.

Consider the following unitary transformation

$$\phi'_j = \sum_i S_{ji}\, \phi_i \tag{6.3}$$

with $S^\dagger = S^{-1}$ and $|\det S| = 1$. The Slater determinant of the transformed states is given by

$$|\,HF\,\rangle' = \frac{1}{\sqrt{N!}}\,\det\begin{bmatrix} \phi'_1(\vec{r}_1) & \dots & \phi'_1(\vec{r}_N) \\ \vdots & & \vdots \\ \phi'_N(\vec{r}_1) & \dots & \phi'_N(\vec{r}_N) \end{bmatrix} = \frac{1}{\sqrt{N!}}\,\det(\Phi').$$

It is easy to see that $\Phi' = S\Phi$, i.e.,

$$\begin{bmatrix} \phi'_1(\vec{r}_1) & \dots & \phi'_1(\vec{r}_N) \\ \vdots & & \vdots \\ \phi'_N(\vec{r}_1) & \dots & \phi'_N(\vec{r}_N) \end{bmatrix} = S \begin{bmatrix} \phi_1(\vec{r}_1) & \dots & \phi_1(\vec{r}_N) \\ \vdots & & \vdots \\ \phi_N(\vec{r}_1) & \dots & \phi_N(\vec{r}_N) \end{bmatrix}.$$

Therefore, $|\,HF\,\rangle' = \det(S)\,|\,HF\,\rangle = |\,HF\,\rangle$. Besides an irrelevant phase factor ($|\det S| = 1$) unitary transformations among the occupied orbitals do not modify the Slater determinant.

One can then profit from this degree of freedom in the choice of the individual single-particle orbitals to diagonalize the matrix of Lagrange multipliers $\varepsilon^\sigma_{ij} = \varepsilon^{\sigma *}_{ji}$, which is hermitic. Eq. (6.3) can be written as

$$\begin{pmatrix} \phi'_1 \\ \vdots \\ \phi'_N \end{pmatrix} = S \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_N \end{pmatrix} \quad \Rightarrow \quad S^\dagger \begin{pmatrix} \phi'_1 \\ \vdots \\ \phi'_N \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_N \end{pmatrix}$$

and $(\phi^*_1 \dots \phi^*_N) = (\phi'_1 \dots \phi'_N)\,S$. Consequently,

$$\sum_{ij} \phi^*_i\,\varepsilon_{ij}\,\phi_j = (\phi^*_1 \dots \phi^*_N)\,\varepsilon \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_N \end{pmatrix} = (\phi'^*_1 \dots \phi'_N)\,S\,\varepsilon\,S^\dagger \begin{pmatrix} \phi'_1 \\ \vdots \\ \phi'_N \end{pmatrix}.$$

In the following we can therefore assume, without loss of generality, that

$$\varepsilon^\sigma_{ij} = \delta_{ij}\,\varepsilon_{i\sigma}.$$

The Hartree-Fock equations yielding the optimal $\phi_i(\vec{r})$ are obtained by varying $\delta\phi^*_{i\sigma}$ and

$\delta\phi_{i\sigma}$ as independent variables. The variation with respect to $\delta\,\phi^*_{i\sigma}$ yields

$$-\frac{\hbar^2}{2m}\qquad \nabla^2\,\phi_{i\sigma}(\vec{r}) + v_\sigma(\vec{r})\,\phi_{i\sigma}(\vec{r}) + \sum_{j\sigma'}\int \phi^*_{j\sigma'}(\vec{r}\,')\,\frac{e^2}{|\vec{r}-\vec{r}\,'|}\,\phi_{j\sigma'}(\vec{r}\,')\,\phi_{i\sigma}(\vec{r})\,d^3r' -$$
$$-\sum_j\int d^3r'\,\phi^*_{j\sigma}(\vec{r}\,')\,\frac{e^2}{|\vec{r}-\vec{r}\,'|}\,\phi_{i\sigma}(\vec{r}\,')\,\phi_{j\sigma}(\vec{r}) = \varepsilon_{i\sigma}\,\phi_{i\sigma}(\vec{r}).$$

One can simplify these equations by defining the *Hartree potential*

$$v_H(\vec{r}) = \sum_{j\sigma'}\int d^3r'\,|\phi_{j\sigma'}(\vec{r}\,')|^2\,\frac{e^2}{|\vec{r}-\vec{r}\,'|} = e^2\int d^3r'\,\frac{n(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} \tag{6.4}$$

and the exchange potential

$$v^\sigma_X(\vec{r},\vec{r}\,') = \frac{e^2}{|\vec{r}-\vec{r}\,'|}\,\sum_j\,\phi_{j\sigma}(\vec{r})\,\phi^*_{j\sigma}(\vec{r}\,')\;. \tag{6.5}$$

Notice that $v^\sigma_X(\vec{r},\vec{r}\,')$ is a nonlocal potential. The second factor in Eq. (6.5) is nothing but the spin-σ density matrix $\gamma_{\sigma\sigma}(\vec{r},\,\vec{r}\,') = \langle\,\hat{\Psi}^\dagger_\sigma(\vec{r}\,')\,\Psi_\sigma(\vec{r})\,\rangle$ in the Hartree-Fock (HF) Slater determinant, whose diagonal element $\gamma_{\sigma\sigma}(\vec{r},\,\vec{r}) = \langle\,\hat{n}_\sigma(\vec{r})\,\rangle = n_\sigma(\vec{r})$ represents the density of electrons with spin σ.

The HF equations then read

$$-\frac{\hbar^2}{2m}\,\nabla^2\,\phi_{i\sigma}(\vec{r}) + [v_\sigma(\vec{r}) + v^\sigma_H(\vec{r})]\;\phi_{i\sigma}(\vec{r}) - \int d^3r'\;v^\sigma_X(\vec{r},\vec{r}\,')\;\phi_{i\sigma}(\vec{r}\,') = \varepsilon_{i\sigma}\,\phi_{i\sigma}(\vec{r}). \tag{6.6}$$

This is a set of N self-consistent equations for the single-particle orbitals $\phi_{i\sigma}(\vec{r})$ that is usually solved by iterative methods. Note that the effect of the $N-1$ electrons $j\sigma' \neq i\sigma$ on the electron $i\sigma$ is taken into account by an *effective single-particle Schrödinger equation.*

A number of properties of the HF equations should be pointed out:

1) $v_X(\vec{r},\vec{r}\,') = v_X(\vec{r}\,',\vec{r})^*\;\Rightarrow\;v^\sigma_X$ is hermitic. The solutions of the HF equations can be chosen to be orthonormal as requested at the start.

2) There are 4 physically distinct terms entering the effective single-particle equation: the kinetic energy, the interaction with external potential $v_\sigma(\vec{r})$ (generated, for example, by the nuclei, and applied electric and magnetic fields), the effective potential $v^\sigma_H(\vec{r})$ due to the direct electron-electron interaction with the other $N-1$ electrons, and finally the exchange interaction with the other $N-1$ electrons, that is given by the non-local potential $v^\sigma_X(\vec{r},\vec{r}\,')$.

3) The single-particle form of the HF equations suggests that the eigenvalues ε_i^σ can be interpreted as effective eigenenergies of individual electron states. Indeed, if one assumes that the single-particle states are unchanged upon ionization the $\varepsilon_{i\sigma}$ correspond to the energy required to remove electron $i\sigma$. This is known as Koopmans' theorem.

On the one side the single-particle eigenvalues are related to the matrix elements of kinetic, potential and Coulomb energy. Indeed, multiplying the HF equations by $\phi^*_{i\sigma}(\vec{r})$ and integrating, one obtains

$$\varepsilon_{i\sigma} = T^\sigma_{ii} + V^\sigma_{ii} + \sum_{j\sigma'} \left(W^{\sigma\sigma'}_{ijij} - \delta_{\sigma\sigma'}\, W^{\sigma\sigma}_{ijji} \right). \tag{6.7}$$

On the other side the Hartree-Fock energy can be written as

$$E_{HF}[\{n_{i\sigma}\}] = \sum_{i\sigma} \left(T^\sigma_{ii} + V^\sigma_{ii}\right) n_{i\sigma} + \frac{1}{2} \sum_{\substack{ij \\ \sigma\sigma'}} \left(W^{\sigma\sigma'}_{ijij} - \delta_{\sigma\sigma'}\, W_{ijji} \right) n_{i\sigma}\, n_{j\sigma'},$$

where $n_{i\sigma} = 1$ for the occupied orbitals and zero otherwise. The change in the Hartree-Fock energy after removal of the electron $k\sigma$ can easily be computed *by assuming that the other orbitals remain unchanged*. This is actually a crude approximation particularly for finite systems. Under this assumption one has

$$\begin{aligned} E_{HF}[\{n_{i\mu}\}] = & \sum_{\substack{i\mu \\ \neq k\sigma}} \left(T^\mu_{ii} + V^\mu_{ii}\right) n_{i\mu} + \left(T^\sigma_{kk} + V^\sigma_{kk}\right) n_{k\sigma} + \\ & + \frac{1}{2} \sum_{\substack{ij\neq k \\ \mu\mu'\neq\sigma}} \left(W^{\mu\mu'}_{ijij} - \delta_{\mu\mu'}\, W^{\mu\mu}_{ijji} \right) n_{i\mu}\, n_{j\mu'} + \\ & + \sum_{j\mu'} \left(W^{\sigma\mu'}_{kjkj} - \delta_{\sigma\mu'}\, W^{\sigma\sigma}_{kjjk} \right) n_{j\mu'}\, n_{k\sigma}, \end{aligned}$$

where the indices μ and $\mu' = \pm 1$ refer here to the spin. Consequently, the energy to remove electron $k\sigma$ ignoring the relaxation of the other orbitals is $E\left[n_{k\sigma} = 0\right] - E\left[n_{k\sigma} = 1\right] = -\varepsilon_{k\sigma}$ [see Eq. (6.7)].

However, note that the energy to remove a second electron does not have such a simple expression even if one continues to ignore orbital relaxations. While the HF equations are free of self-interaction, it is also true that the interaction between electrons i and j is counted twice when computing the single-particle eigenvalues: once in $\varepsilon_{i\sigma}$ and once in $\varepsilon_{j\sigma'}$. This is an intrinsic limitation of the self-consistent mean-field approach.

4) As a result of this *double counting* of electron-electron interactions the sum of the eigenvalues is not equal to the Hartree-Fock energy. Indeed, multiplying the HF equations by $\phi^*_{i\sigma}(r)$, integrating over $\vec{r}$, and summing over $i\sigma$, one obtains

$$\sum_{i\sigma} \varepsilon_{i\sigma} = \langle T \rangle + \langle V \rangle + 2 \langle W \rangle = E_{HF} + E_H + E_X.$$

5) One should also note that the exchange potential $v_X(\vec{r}, \vec{r}\,')$ cancels the self-interaction term in $v_H(\vec{r})$ exactly. In fact, for $j = i$ we have

$$v^\sigma_X(r, r') \xrightarrow{j=i} \int d^3r' \, \frac{\phi_{i\sigma}(\vec{r})\, \phi^*_{i\sigma}(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} \, \phi_{i\sigma}(r') = \phi_{i\sigma}(\vec{r}) \int d^3r' \, \frac{|\phi_{i\sigma}(\vec{r}\,')|^2}{|\vec{r} - \vec{r}\,'|}.$$

The Hartree-Fock equations are thus self-interaction free.

6) If one neglects the exchange term, except for the obvious self-interaction part, one obtains the *Hartree equations*. In this case we have an *orbital-dependent* effective Hamiltonian, which poses some difficulties concerning orthogonality of the different $\phi_{i\sigma}$. The orbital-dependent interaction potential reads

$$\tilde{v}^{(i\sigma)}_H(\vec{r}) = \int \sum_{j\sigma' \neq i\sigma} |\phi_{j\sigma'}(\vec{r}\,')|^2 \, \frac{e^2}{|\vec{r} - \vec{r}\,'|} \, d^3r' = e^2 \int d^3r' \, \frac{[n(\vec{r}\,') - |\phi_{i\sigma}(\vec{r}\,')|^2]}{|\vec{r} - \vec{r}\,'|}.$$

It is interesting to point out that these equations can be derived by using as ansatz wave function a simple *non-antisymmetrized* product of single-particle orbitals. This is, of course, not justified, since such states are unphysical.

6.3 From Hartree to Hartree-Fock: The exchange hole $\rho^\sigma_x(\vec{r}, \vec{r}\,')$

The purpose of this section is to provide a local interpretation of the consequences of the antisymmetry of the many-electron wave functions, also know as symmetry correlations or exchange effects, by introducing the concept of exchange hole. Incidentally, this is also useful as an introduction to the concept of exchange-correlation hole, which allows us to obtain a graphical picture of correlations in the context of density-funtional theory.

This section is still in construction ...

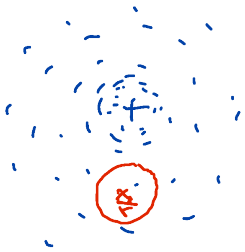
Density and density matrix

$$n_\sigma(\vec{r}) = \langle \Psi \,|\, \hat{\Psi}^\dagger_\sigma(\vec{r}) \hat{\Psi}_\sigma(\vec{r}) \,|\, \Psi \rangle,$$

where
Density matrix

$$\gamma^{(1)}_{\sigma\sigma'}(\vec{r}, \vec{r}\,') = \langle \Psi \,|\, \hat{\Psi}^\dagger_{\sigma'}(\vec{r}\,') \hat{\Psi}_\sigma(\vec{r}) \,|\, \Psi \rangle$$
$$\gamma_{\sigma\sigma}(\vec{r}, \vec{r}) = n_\sigma(\vec{r})$$

$$\Psi_\sigma^\dagger(\vec{r}) = \sum_i \Phi_{i\sigma}^*(\vec{r})\,\hat{c}_{i\sigma}^\dagger$$
$$\Psi_\sigma(\vec{r}) = \sum_i \Phi_{i\sigma}(\vec{r})\,\hat{c}_{i\sigma}$$



are the field operators.

For a Slater determinant $|\,HF\,\rangle$:

$$n_\sigma(\vec{r}) = \sum_{\substack{i \\ occ.}} |\Phi_i(\vec{r})|^2$$

$$\gamma_{\sigma\sigma'}^{(1)}(\vec{r},\vec{r}\,') = \langle\, HF \,|\, \underbrace{\sum_j \Phi_{j\sigma}^*(\vec{r}\,')\,\hat{c}_{j\sigma}^\dagger}_{\Psi_\sigma^\dagger(\vec{r}\,')} \underbrace{\sum_i \Phi_{i\sigma}(\vec{r})\,\hat{c}_{i\sigma}}_{\Psi_\sigma(\vec{r})} \,|\, HF \,\rangle$$

Only $i = j \;\Rightarrow$
$$= \sum_j \Phi_{j\sigma}^*(\vec{r}\,')\Phi_{j\sigma}(\vec{r}) \underbrace{\langle\, HF \,|\, \hat{c}_{j\sigma}^\dagger \hat{c}_{j\sigma} \,|\, HF \,\rangle}_{n_{j\sigma}}$$

For a Slater determinant:

$$\gamma_{\sigma\sigma'}^{(1)}(\vec{r},\vec{r}\,') = \sum_{\substack{j \\ occ.}} \Phi_{j\sigma}^\dagger(\vec{r}\,')\Phi_{j\sigma}(\vec{r})$$

Hartree energy

$$E_H = \frac{e^2}{2} \iint d^3r d^3r' \sum_{\sigma\sigma'} \frac{n_\sigma(\vec{r}) n_{\sigma'}(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}$$

Exchange energy

$$E_x = -\frac{e^2}{2}\iint d^3r d^3r' \sum_{\substack{ij\\ \sigma}} \Phi^*_{i\sigma}(\vec{r})\Phi^*_{j\sigma}(\vec{r}')\frac{1}{|\vec{r}-\vec{r}'|}\Phi_{i\sigma}(\vec{r}')\Phi_{j\sigma}(\vec{r})$$

$$= -\frac{e^2}{2}\sum_\sigma \iint d^3r d^3r' \underbrace{\sum_i \Phi^*_{i\sigma}(\vec{r})\Phi_{i\sigma}(\vec{r}')}_{\gamma_\sigma(\vec{r}',\vec{r})} \underbrace{\sum_j \Phi^*_{j\sigma}(\vec{r}')\Phi_{j\sigma}(\vec{r})}_{\gamma_\sigma(\vec{r},\vec{r}')} \frac{1}{|\vec{r}-\vec{r}'|}$$

Due to $\gamma(\vec{r},\vec{r}') = \gamma(\vec{r}',\vec{r})^* \Rightarrow$

$$E_x = -\frac{e^2}{2}\sum_\sigma \iint d^3r d^3r' |\gamma_\sigma(\vec{r},\vec{r}')|^2 \frac{1}{|\vec{r}-\vec{r}'|}$$

$\Rightarrow E_X \leq 0$

$$E_x = -\frac{e^2}{2}\sum_\sigma \iint d^3r d^3r' \underbrace{\frac{|\gamma_\sigma(\vec{r},\vec{r}')|^2}{n_\sigma(\vec{r}')}}_{\rho_x^\sigma(\vec{r}')} \frac{1}{|\vec{r}-\vec{r}'|} n_\sigma(\vec{r}')$$

$$= \frac{e^2}{2}\sum_\sigma \int d^3r' \underbrace{\left[-\int d^3r \frac{|\gamma_\sigma(\vec{r},\vec{r}')|^2}{n_\sigma(\vec{r})}\frac{1}{|\vec{r}-\vec{r}'|}\right]}_{\text{Attractive potential}} n_\sigma(\vec{r}')$$

Exchange density

$$\rho_x^\sigma(\vec{r},\vec{r}') = \frac{|\gamma_\sigma(\vec{r},\vec{r}')|^2}{n_\sigma(\vec{r})}$$

$$\rho_x^\sigma(\vec{r},\vec{r}) = \frac{|\gamma_\sigma(\vec{r},\vec{r})|^2}{n_\sigma(\vec{r})} = n_\sigma(\vec{r})$$

$$\int \rho_x^\sigma(\vec{r},\vec{r}')d^3r' = 1 \quad \forall r$$

$$\int \rho_x^\sigma(\vec{r},\vec{r}')d^3r' = \frac{1}{n_\sigma(\vec{r})}\int d^3r' \sum_j \Phi^*_{j\sigma}(\vec{r})\Phi_{j\sigma}(\vec{r}')\sum_i \Phi^*_{i\sigma}(\vec{r}')\Phi_{i\sigma}(\vec{r})$$

$$= \frac{1}{n_\sigma(\vec{r})}\sum_j |\Phi_{j\sigma}(\vec{r})|^2 = 1.$$

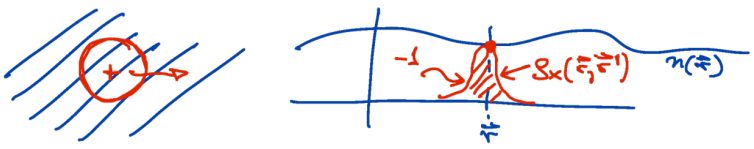


Figure 11: Exchange-hole

$$\Rightarrow E_x = -\frac{e^2}{2}\sum_\sigma \iint d^3r d^3r' \rho_x^\sigma(\vec{r},\vec{r}')\frac{n_\sigma(\vec{r})}{|\vec{r}-\vec{r}'|}$$

A more elegant derivation of the Hartree-Fock (or any) variational equations

Consider an arbitrary $\delta\widetilde{\Phi}_i$ and with that construct two variations $\delta\Phi_i$ of the orbitals $\delta\Phi_i = i\delta\widetilde{\Phi}_i$:

$$E = T + V + E_H + E_x$$

$$T + V = \sum_k \int \Phi_k^*(\vec{r})\left(\frac{-\hbar^2}{2m}\nabla^2 + v(\vec{r})\right)\Phi_k(\vec{r})d^3r.$$

$$\delta\Phi_i = \delta\widetilde{\Phi}_i \Rightarrow \delta T = \int \underbrace{\delta\widetilde{\Phi}_i^*}_{\delta\Phi_i^*}(\vec{r})\left(\frac{-\hbar^2}{2m}\nabla^2 + v(\vec{r})\right)\Phi_i(\vec{r})d^3r + \int \Phi_i(\vec{r})\left(\frac{-\hbar^2}{2m}\nabla^2 + v(\vec{r})\right)\underbrace{\delta\widetilde{\Phi}_i}_{\delta\Phi_i}(\vec{r})d^3r$$

$$\delta\Phi_i = i\delta\widetilde{\Phi}_i \Rightarrow \delta T = -i\int \delta\widetilde{\Phi}_i^*(\vec{r})\left(\frac{-\hbar^2}{2m}\nabla^2 + v(\vec{r})\right)\Phi_i(\vec{r})d^3r + i\int \Phi_i(\vec{r})\left(\frac{-\hbar^2}{2m}\nabla^2 + v(\vec{r})\right)\delta\widetilde{\Phi}_i(\vec{r})d^3r$$

$$\Rightarrow \quad \int \delta\widetilde{\Phi}_i^*(\vec{r})\underbrace{\left(\frac{-\hbar^2}{2m}\nabla^2 + v(\vec{r})\right)\Phi_i(\vec{r})}_{\Rightarrow \quad =0.} d^3r = 0 \quad \forall\delta\widetilde{\Phi}_i$$

Consider one of the spin orbitals $\Phi_k(x)$ of the single-particle basis and construct the ket $|\,\Psi\,\rangle$ consisting in the superposition of states $\hat{\Psi}^\dagger(x)|\,vac\,\rangle$ with defined position x weighted with the orbital $\Phi_k(x)$:

$$|\,\Psi\,\rangle = \int dx\,\Phi_k(x)\hat{\Psi}^\dagger(x)|\,vac\,\rangle. \tag{6.8}$$

It is easy to see that

$$\begin{aligned}
|\,\Psi\,\rangle &= \int dx\,\hat{\Psi}^\dagger(x)|\,vac\,\rangle\Phi_k(x) \\
&= \sum_i \hat{c}_i^\dagger|\,vac\,\rangle \underbrace{\int \Phi_i^*(x)\Phi_k(x)dx}_{\delta_{ik}} \\
&= \hat{c}_k^\dagger|\,vac\,\rangle,
\end{aligned}$$

where we have used the definition of $\hat{\Psi}^\dagger(x) = \sum_i \Phi_i^*(x_i)\hat{c}_i^\dagger$. This shows that our occupation number representation (6.8) of the one-particle state $|\,\Psi\,\rangle$ is physically correct.

In the case of N particle states the ket $|\,\Psi\,\rangle$ corresponding to the wave function $\Psi(x,...,x_N)$ is given by

$$|\,\Psi\,\rangle = \frac{1}{\sqrt{N!}} \int dx_1...dx_N\, \hat{\Psi}^\dagger(x_1)...\hat{\Psi}^\dagger(x_N)|\,vac\rangle \Psi(x_1,...,x_N). \tag{6.9}$$

It is easy to verify that the normalization is correct:

$$\langle\,\Psi\,|\,\Psi\,\rangle = \frac{1}{\sqrt{N!}} \int dx_1...dx_N, dx'_1,...,dx'_N\, \underbrace{\langle\, vac\,|\overbrace{\hat{\Psi}(x_N)...\hat{\Psi}(x'_1)\hat{\Psi}^\dagger(x_1)...\hat{\Psi}^\dagger(x_N)}^{N!\delta(x_1-x'_1)...\delta(x_N-x'_N)}|\,vac\,\rangle}_{\Psi(x_1...x_N)\Psi^*(x'_1...x'_N)}$$

$$= \int dx_1...dx_N |\Psi(x_1,...,x_N)|^2$$

We want to apply the just obtained expression for $|\,\Psi\,\rangle$ in occupation-number representation in terms of the N particle coordinate wave function $\Psi(x_1,...,x_N)$ in order to derive the well-known expression

$$\rho(x,x') = N \int dx_2...dx_N\, \Psi(x_1,x_2,...,x_N)\Psi^*(x',x_2,...,x_N) \tag{6.10}$$

for the single-particle density matrix $\gamma(x,x')$. Our starting point is the definition of $\gamma(x,x')$ in second quantization:

$$\gamma(x,x') = \langle\,\Psi\,|\,\hat{\Psi}^\dagger(x')\hat{\Psi}(x)\,|\,\Psi\,\rangle \tag{6.11}$$

$$\hat{\Psi}(x)|\,\Psi\,\rangle = N\frac{1}{\sqrt{N!}} \int dx_2...dx_N\, \Psi^\dagger(x_2)...\Psi^\dagger(x_N)\Psi(x,x_2,...,x_N)|\,vac\,\rangle$$

$$\langle\,\Psi\,|\hat{\Psi}^\dagger(x') = N\frac{1}{\sqrt{N!}} \int dx_2...dx_N\, \Psi^*(x',x_2,...,x_N)\langle\,vac\,|\Psi(x_N)...\Psi(x_2)$$

$$\langle\,\Psi\,|\,\hat{\Psi}^\dagger(x')\hat{\Psi}(x)\,|\,\Psi\,\rangle = \underbrace{N^2\frac{1}{N!}(N-1)!}_{N} \int dx_2...dx_N\, \Psi(x,x_2,...,x_N)\Psi^*(x',x_2,...,x_N)$$

Towards DFT: What do we need to know in order to calculate the ground-state energy? Do we need the ground-state wave function?

$$\hat{H} = \sum_\sigma \int d^3r \hat{\Psi}^\dagger_\sigma(\vec{r}) \left(\frac{-\hbar^2}{2m}\nabla^2\right) \hat{\Psi}_\sigma(\vec{r}) + \sum_\sigma \int d^3r \hat{\Psi}^\dagger_\sigma(\vec{r})\hat{\Psi}_\sigma(\vec{r})v(\vec{r}) \tag{6.12}$$

$$+\frac{1}{2},$$

$$E_0 = \langle\,\Psi_0\,|\,\hat{H}\,|\,\Psi_0\,\rangle. \tag{6.13}$$

$$\langle\,\Psi_0\,|\sum_\sigma \hat{\Psi}_\sigma(\vec{r})\hat{\Psi}_\sigma(\vec{r})\,|\,\Psi_0\,\rangle = n(\vec{r}), \qquad \leftarrow \text{density of } e^- \tag{6.14}$$

$$n(\vec{r}) : \mathbb{R}^3 \to \mathbb{R} \qquad \int d^3r\, n(\vec{r}) = N = \#\,\text{particles}.$$

One can prove that

$$\langle\, \Psi_0 \,|\, \sum_\sigma \hat{\Psi}_\sigma(\vec{r})\hat{\Psi}_\sigma(\vec{r}) \,|\, \Psi_0 \,\rangle = N \sum_\sigma \int dx_2 ... dx_N |\Psi(\vec{r}, \sigma, x_2, ..., x_N)|^2,$$

where

$$x_i = \vec{r}_i, \sigma_i \quad \text{and} \quad \int dx_i \equiv \sum_{\sigma_i}.$$

The kinetic energy is more tricky since we have the Laplace operator acting on the operator $\hat{\Psi}(\vec{r})$. However, if we could calculate the density matrix

$$\gamma(\vec{r}', \vec{r}) = \langle\, \Psi_0 \,|\, \sum_\sigma \hat{\Psi}^\dagger_\sigma(\vec{r})\hat{\Psi}(\vec{r}') \,|\, \Psi_0 \,\rangle \tag{6.15}$$

we could obtain $\langle\, \Psi_0 \,|\, \hat{T} \,|\, \Psi_0 \,\rangle$ as

$$\begin{aligned}\langle\, T \,\rangle &= \langle\, \Psi_0 \,|\, \int d^3r \hat{\Psi}(\vec{r}) \left[\frac{-\hbar^2}{2m}\nabla^2_{r'}\hat{\Psi}_\sigma(\vec{r}')\right]_{\vec{r}'=\vec{r}} \,|\, \Psi_0 \,\rangle \\ &= \int d^3r\, \hat{\Psi}(\vec{r}) \left[\frac{-\hbar^2}{2m}\nabla^2_{r'}\gamma(\vec{r}, \vec{r}')\right]_{\vec{r}'=\vec{r}}.\end{aligned}$$

Therefore the 1st order density matrix $\gamma(\vec{r}, \vec{r}')$ would be enough to compute the kinetic energy. One can show that in terms of the wave function in coordinate representation:

$$\gamma(\vec{r}, \vec{r}') = N \sum_\sigma \int dx_2 ... dx_N\, \Psi^*(\vec{r}, \sigma, x_2, ..., x_N)\Psi(\vec{r}', \sigma, x_2, ..., x_N) \tag{6.16}$$

of course $\gamma(\vec{r}, \vec{r}) = n(\vec{r})$. In order to calculate the Coulomb interaction we need to have access to the 2 particle reduced density matrix:

$$\gamma^{(2)}(\vec{r}_1, \vec{r}_2, \vec{r}_1{}', \vec{r}_2{}') = \sum_{\sigma\sigma'} \langle\, \Psi_0 \,|\, \hat{\Psi}^\dagger_\sigma(\vec{r}_1{}')\hat{\Psi}^\dagger_{\sigma'}(\vec{r}_2{}')\hat{\Psi}_{\sigma'}(\vec{r}_2)\hat{\Psi}_\sigma(\vec{r}_1) \,|\, \Psi_0 \,\rangle \tag{6.17}$$

which can be shown to be equal to

$$\gamma^{(2)}(\vec{r}_1, \vec{r}_2, \vec{r}_1{}', \vec{r}_2{}') = \frac{N(N-1)}{2} \int dx_3 ... dx_N\, \Psi^*(\vec{r}_1{}''\sigma, \vec{r}_2{}'\sigma', x_3, ..., x_N)\Psi(\vec{r}_1\sigma, \vec{r}_2\sigma', x_3, ..., x_N) \tag{6.18}$$

The reason why the wave function method work is that you don't need a so good approximation of Ψ_0 in order to obtain a pretty good $\gamma^{(2)}$, $\gamma^{(1)}$ and $n(\vec{r}$.
Knowing $\gamma^{(2)}$ you can obtain $\gamma^{(1)}$ by integration on $(\vec{r}_2, \sigma')$ and from $\gamma^{(1)}$ one gets of course $n(\vec{r}$. The most remarkable and a priori far from obvious thing with DFT is that one can get along, i.e., in principle derive the exact ground-state energy only from $n(\vec{r})$ ... and some restrictions ... and some challenges.

The Hohenberg-Kohn Theorem

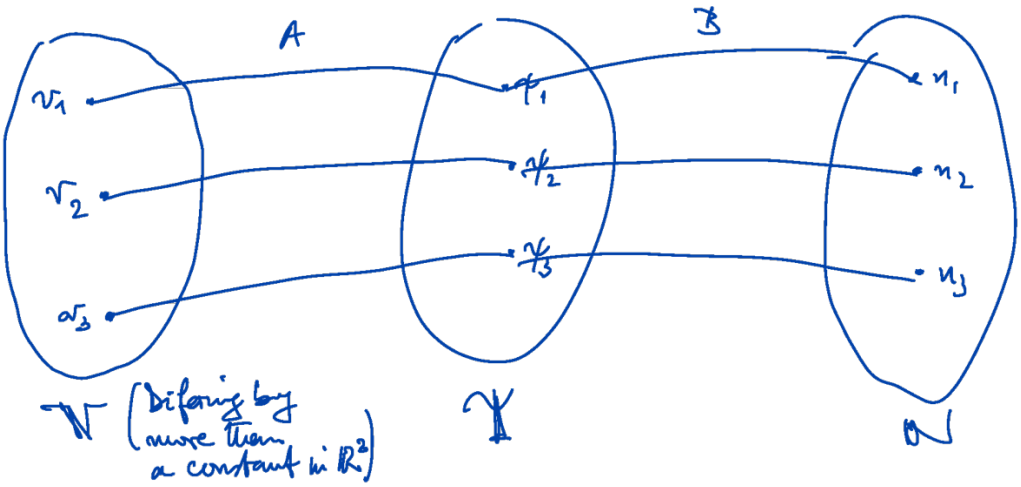


Figure 12: Unique mapping between the potential, a set of degenerated wavefunctions and the densities

i) The ground-state energy of any many-electron system is a functional of the density of electrons $n(\vec{r})$ and

ii) for any given external potential $v(\vec{r})$ the ground-state energy E_0 and density $n_0(\vec{r})$ can be obtained by minimizing a functional of the density.

$$v(\vec{r}) \quad \rightarrow \quad \hat{H} \quad \rightarrow \quad \Psi_0 \quad \rightarrow \quad n(\vec{r}) = \langle\, \Psi_0 \,|\, \sum_\sigma \hat{\Psi}^\dagger_\sigma(\vec{r})\hat{\Psi}(\vec{r}) \,|\, \Psi_0 \,\rangle$$

The mappings A and B are surjective by definition of Ψ and N. Are they injective?

1) Injectivity of A
Assume $v(\vec{r})$ and $v'(\vec{r})$ give the same Ψ_0

$$\hat{H} = \hat{T} + \hat{V} + \hat{W} \quad \Rightarrow \quad \hat{H}|\,\Psi_0\,\rangle = E_0|\,\Psi_0\,\rangle$$
$$\hat{H}' = \hat{T} + \hat{V}' + \hat{W} \quad \Rightarrow \quad \hat{H}'|\,\Psi_0\,\rangle = E_0'|\,\Psi_0\,\rangle$$

$\hat{T}$ and $\hat{W}$ are independent of the problem under study.

$$\Rightarrow \quad (\hat{V} - \hat{V}')|\,\Psi_0\,\rangle = (E_0 - E_0')|\,\Psi_0\,\rangle$$

$$\Rightarrow [v(\vec{r}) - v'(\vec{r})]\Psi_0(\vec{r}, \sigma, x_2, ..., x_N) = cte?\ \Psi_0(\vec{r}, \sigma, x_2, ..., x_N)$$

$$\Rightarrow [v(\vec{r}) - v'(\vec{r})] = cte? \text{ in a region of nonzero measure}$$
$$\Rightarrow [v(\vec{r}) - v'(\vec{r})] = cte? \text{ everywhere.}$$

The potential cannot have infinite barriers. Thus A is bijective $v(\vec{r}) \longleftrightarrow \Psi_0(\vec{r})$.

2) Injectivity of B:
Suppose $|\,\Psi\,\rangle$ and $|\,\Psi\,\rangle'$ yield the same $n(\vec{r})$.

a) $E = \langle\,\Psi\,|\,\hat{H}\,|\,\Psi\,\rangle < \langle\,\Psi'\,|\,\hat{H}'\,|\,\Psi'\,\rangle = \langle\,\Psi'\,|\,\hat{H}' + \hat{V} - \hat{V}'\,|\,\Psi'\,\rangle$
$= E' + \langle\,\Psi'\,|\,\hat{V} - \hat{V}'\,|\,\Psi'\,\rangle = E' + \int d^3r[v(\vec{r}) - v'(\vec{r})]n(\vec{r})$
Repeating the argument starting from $\langle\,\Psi'\,|\,\hat{H}'\,|\,\Psi'\,\rangle$ we get

b) $E' < E + \int d^3r[v'(\vec{r}) - v(\vec{r})]n(\vec{r})$
Adding (a) and (b) we have $E + E' < E' + E$.
$\Rightarrow$ *contradiction*

We conclude that B is injective and therefore bijective:

$$v(\vec{r}) \longleftrightarrow |\,\Psi_0\,\rangle \longleftrightarrow n(\vec{r}).$$

Example:
In the Kohn-Sham method of minmizing the ground-state energy functional the density is exposed as the sum of occupation probabilities $|\varphi_i(\vec{r})|^2$ of single-particle orbitals [the so-called Kohn-Sham orbitals $\varphi_i(\vec{r})$ which are the solutions of the non-interacting Schrödinger equation under the action of an auxiliary external potential $v_S(\vec{r})$:

$$n(\vec{r}) = \sum_i |\varphi_i(\vec{r})|^2 \tag{6.19}$$

with

$$\int d^3r|\varphi_i(\vec{r})|^2 = 1$$

and

$$\frac{-\hbar^2\nabla^2}{2m}\varphi_i(\vec{r}) + v_s(\vec{r})\varphi_i(\vec{r}) = \varepsilon_i\varphi_i(\vec{r}). \tag{6.20}$$

Clearly, $\varphi_i(\vec{r})$ and $v_S(\vec{r})$ are functionals of the density $n(\vec{r})$. They intend to reproduce $n(\vec{r})$ according to 6.19. Uniqueness is guaranteed by the Hohenberg-Kohn theorem for $W = 0$.
Expressing the average non-interacting kinetic energy functional as

$$T_S[n] = \sum_i \int d^3r\varphi_i^*(\vec{r}) - \frac{-\hbar^2\nabla^2}{2m}\varphi_i(\vec{r})$$

show that

$$\delta T_S = -\int d^3r v_S(\vec{r})\delta n(\vec{r})$$

which implies

$$\frac{\delta T_S}{\delta n(\vec{r})} = -v_S(\vec{r}) + \mu,$$

where μ is an arbitrary constant. This is the main property which defines $v_S(\vec{r})$ and leads to the Kohn-Sham equations.

Solutions:

$$\begin{aligned}
T_S[n] &= \sum_i \int \varphi^*(\vec{r}) \left(\frac{-\hbar^2}{2m} \nabla^2 \right) \varphi_i(\vec{r}) d^3r \quad \varphi_i(\vec{r}) = \varphi_i(\vec{r}, n) \\
\delta T_S &= \int d^3r \frac{\delta T_S}{\delta n(\vec{r})} \delta n(\vec{r}) \\
&= \sum_i \int \left[\delta\varphi^*(\vec{r}) \underbrace{\frac{\hat{p}^2}{2m} \varphi_i(\vec{r})}_{[\varepsilon_i - v_S(\vec{r})]\varphi_i(\vec{r})} d^3r + \underbrace{\varphi_i^*(\vec{r}) \frac{\hat{p}^2}{2m}}_{[\varepsilon_i - v_S(\vec{r})]\varphi_i^*(\vec{r})} \delta\varphi_i(\vec{r}) \right] d^3r \\
&= \sum_i \int d^3r \varepsilon_i - v_S(\vec{r})] \left[\delta\varphi^*(\vec{r})\varphi(\vec{r}) + \varphi_i^*(\vec{r})\delta\varphi(\vec{r}) \right] \\
&= \sum_i \varepsilon_i \underbrace{\int d^3r \left(\delta\varphi^*(\vec{r})\varphi(\vec{r}) + \varphi_i^*(\vec{r})\delta\varphi(\vec{r}) \right)}_{\delta\left(\int d^3r |\phi_i|^2\right) = \delta(1) = 0} - \int d^3r v_S(\vec{r}) \underbrace{\sum_i \left(\delta\varphi^*(\vec{r})\varphi(\vec{r}) + \varphi_i^*(\vec{r})\delta\varphi(\vec{r}) \right)}_{\delta n(\vec{r})} \\
&= -\int d^3r v_S(\vec{r}) \delta n(\vec{r})
\end{aligned}$$

7 Density functional theory

The Hartree-Fock approximation is an extremely useful concept in many-particle physics. The self-consistent-field method introduced in this context finds many applications throughout physics and chemistry. Moreover, the so-called post-Hartree-Fock methods derived by using linear combinations of many Slater determinants (either variational or perturbative) can often provide the most accurate results on the ground state and low-lying excited states of atoms and small molecules. However, theoretical methods based on the wave-function $\Psi(x_1, \ldots x_N)$ become rapidly impracticable for studies of nanoparticles, nanostructures and extended systems like wires, films and solids. A different approach to the many-body problem, known as density functional theory, has proven to be most successful in these cases. This theory is based on the fact that, for the study of ground-state properties, the wave function $\Psi(x_1, \ldots x_N)$ can be replaced by the much simpler electronic density $n(\vec{r})$ as the fundamental characteristic of the many-body problem. In the following we present the basic concepts behind this theory, without aiming to pause on mathematical rigor. The interested reader may consult, for example, the books by Dreizler and Gross or by Parr and Yang.[13]

7.1 Levy's constrained search of the ground-state energy

In quantum mechanics the energy of an N-electron state, as any other property, is known to be a functional of the wave function:

$$E\left[|\,\Psi\,\rangle\right] = \langle\,\Psi\,|\,\hat{H}\,|\,\Psi\,\rangle,$$

where $\hat{H}$ is the Hamiltonian, and the wave function $|\,\Psi\,\rangle$ is normalized ($\langle\,\Psi\,|\,\Psi\,\rangle = 1$) and antisymmetric with respect to interchange. The ground-state energy E_0 and ground-state wave-function $|\,\Psi_0\,\rangle$ can be determined by minimization of $E\left[|\,\Psi\,\rangle\right]$ following Rayleigh-Ritz variational principle:

$$E_0 = \min_{|\,\Psi\,\rangle} \left\{\langle\,\Psi\,|\,\hat{T} + \hat{W} + \hat{V}\,|\,\Psi\,\rangle\right\}, \tag{7.1}$$

where $\langle\,\Psi\,|\,\Psi\,\rangle = 1$. $\hat{T}$ and $\hat{W}$ are the kinetic-energy and Coulomb-interaction operators, and

$$\hat{V} = \sum_\sigma \int v(\vec{r})\ \hat{\Psi}^\dagger_\sigma(\vec{r})\ \hat{\Psi}_\sigma(\vec{r})\ d^3r = \int v(\vec{r})\ \hat{n}(\vec{r})\ d^3r$$

is the external-potential operator, which defines the problem under study. It is then insightful, as proposed by Levy, to perform the minimization in two steps:

$$E_0 = \min_{n(\vec{r})} \left\{ \min_{|\,\Psi\,\rangle \to n(\vec{r})} \langle\,\Psi\,|\,\hat{T} + \hat{W} + \hat{V}\,|\,\Psi\,\rangle \right\}, \tag{7.2}$$

[13] R. M. Dreizler and E. K. U. Gross, Density functional theory (Springer, Berlin, 1990); R. G. Parr and W. Yang (Oxford University Press, New York, 1989)

where $\langle\, \Psi \,|\, \hat{n}(\vec{r}) \,|\, \Psi \,\rangle = n(\vec{r})$ represents the electronic density and $\int d^3r\; n(\vec{r}) = N$ gives the number of electrons. The internal minimization or search is constrained in the sense that one does not look for the minimum energy over all antisymmetric N-electron $|\, \Psi \,\rangle$, but only over those $|\, \Psi \,\rangle$ that yield a given density $n(\vec{r})$. The second step then minimizes the resulting functional of $n(\vec{r})$ with respect to all possible density distributions $n(\vec{r})$, so that Eqs. (7.1) and (7.2) coincide. We may thus write

$$E_0 = \min_{n(\vec{r})} E_{LL}[n(\vec{r})]\,, \tag{7.3}$$

where the Levy-Lieb energy E_{LL} is the functional of the density, which is given by

$$\begin{aligned} E_{LL}[n(\vec{r})] &= \min_{|\,\Psi\,\rangle \to n(\vec{r})} \langle\, \Psi \,|\, \hat{T} + \hat{W} + \hat{V} \,|\, \Psi \,\rangle \\ &= \min_{|\,\Psi\,\rangle \to n(\vec{r})} \{\langle\, \Psi \,|\, \hat{T} + \hat{W} \,|\, \Psi \,\rangle\} + \int d^3r\; v(\vec{r})\; n(\vec{r}) \\ &= F_{LL}[n(\vec{r})] + \int d^3r\; v(\vec{r})\; n(\vec{r}). \end{aligned} \tag{7.4}$$

The density functional

$$F_{LL}[n(\vec{r})] = \min_{|\,\Psi\,\rangle \to n(\vec{r})} \langle\, \Psi \,|\, \hat{T} + \hat{W} \,|\, \Psi \,\rangle$$

represents the lowest possible value of the *sum* of the kinetic and Coulomb energies of the system for a given distribution $n(\vec{r})$ of the electrons in $\mathbb{R}^3$. Notice that obtaining $F_{LL}[n(\vec{r})]$ implies a highly non-trivial constrained search over all N-electron states $|\, \Psi \,\rangle$ yielding the density $n(\vec{r})$. $F_{LL}[n(\vec{r})]$ is a *universal functional* of the electronic density $n(\vec{r})$, since it is independent of the specific problem under study, i.e., independent of $v(\vec{r})$. The number of electrons $N = \int d^3r\; n(\vec{r})$, an important constraint along the minimization process yielding $F_{LL}[n(\vec{r})]$, is also given by $n(\vec{r})$.

The Levy-Lieb functional $F_{LL}[n(\vec{r})]$ can always be written as the sum

$$F_{LL}[n(\vec{r})] = T[n(\vec{r})] + W[n(\vec{r})] \tag{7.5}$$

of the kinetic-energy functional $T[n(\vec{r})]$ and the Coulomb-energy functional $W[n(\vec{r})]$, where

$$T[n(\vec{r})] = \langle\, \Psi[n(\vec{r})] \,|\, \hat{T} \,|\, \Psi[n(\vec{r})] \,\rangle$$

and

$$W[n(\vec{r})] = \langle\, \Psi[n(\vec{r})] \,|\, \hat{W} \,|\, \Psi[n(\vec{r})] \,\rangle$$

are the average kinetic and Coulomb energy at the wave function yielding the optimum F_{LL} for the given $n(\vec{r})$. Still, notice that $|\, \Psi[n(\vec{r})] \,\rangle$ and thus $T[n(\vec{r})]$ and $W[n(\vec{r})]$ are

the result of the subtle compromise of minimizing both the kinetic and Coulomb repulsion contributions to the energy of the many-electron system at the same time.

From Eqs. (7.3) and (7.4) we conclude that the ground-state energy of an N-electron system can be obtained by minimizing a functional of $n(\vec{r})$ alone:

$$E_0 = \min_{n(\vec{r})} \{F_{LL}[n(\vec{r})] + \int n(\vec{r})\, v(\vec{r})\, d^3r\}$$

under the constraint $\int d^3r\, n(\vec{r}) = N$. The functional to minimize is given by two main terms. The first one is universal, though not explicitly known and potentially extremely complex. The second one depends on the problem under study and is explicitly known. The fundamental challenge in density-functional theory (DFT) is to find good approximations to $F_{LL}[n(\vec{r})]$.

The electronic density of the ground state must satisfy the Euler-Lagrange equations which are obtained by requiring that

$$\Omega[n(\vec{r})] = E[n(\vec{r})] - \mu \left(\int n(\vec{r})\, d^3r - N \right)$$

is stationary, i.e.,

$$\delta\, \Omega[n(\vec{r})] = \delta F_{LL}[n] + \int \delta n(\vec{r})\, v(\vec{r})\, d^3r - \mu \int \delta n(\vec{r})\, d^3r = 0. \tag{7.6}$$

Using the definition of functional derivative

$$\delta F_{LL} = \int \frac{\delta F_{LL}}{\delta n(\vec{r})}\, \delta n(\vec{r})\, d^3r\,,$$

we can write Eq. (7.6) as

$$\int \left(\frac{\delta F_{LL}}{\delta n(\vec{r})} + v(\vec{r}) - \mu \right) \delta n(\vec{r})\, d^3r = 0\,. \tag{7.7}$$

Since $\delta n(\vec{r})$ is arbitrary we must have

$$\frac{\delta F_{LL}}{\delta n(\vec{r})} + v(\vec{r}) = \mu \tag{7.8}$$

and $\int n(\vec{r})\, d^3r = N$. Given an explicit approximation to $F_{LL}[n(\vec{r})]$, the solution of Eq. (7.8) allows us to obtain the ground-state density $n_0(\vec{r})$ and, from Eq. (7.4), the ground-state energy for any external potential $v(\vec{r})$.

7.2 The Kohn-Sham method

Once we have shown that the ground-state energy can be obtained by minimizing a functional of $n(\vec{r})$ with respect to $n(\vec{r})$, we aim now to implement the variational principle for $E[n(\vec{r})]$ in a differential form. For this purpose one introduces an auxiliary potential $v_S(\vec{r})$ such that the ground state $|\,\Psi_S\,\rangle$ of

$$\hat{H}_S = \hat{T} + \hat{V}_S \quad \text{with} \quad \hat{W} = 0$$

(i.e., without electron-electron interaction) has the same density $n(\vec{r})$ as the interacting electron system. As usual, the single-particle potential operator is given by

$$\hat{V}_S = \sum_\sigma \int \hat{\Psi}^\dagger_\sigma(\vec{r})\, v_S(\vec{r})\, \hat{\Psi}_\sigma(\vec{r})\, \mathrm{d}^3 r,$$

where the auxiliary potential $v_S = v_S[n; \vec{r}\,]$ is an $\vec{r}$-dependent functional of $n(\vec{r})$.

Notice that we are implicitly assuming that $n(\vec{r})$ is non-interacting v-representable, i.e., we assume that there exists a system of N non-interacting electrons ($\hat{W} = 0$) and an external potential $v_S(\vec{r})$ such that the problem

$$v_S(\vec{r}) \;\rightarrow\; |\,\Psi_S\,\rangle \;\rightarrow\; n_S(r) = \langle\,\Psi_S\,|\,\hat{n}(r)\,|\,\Psi_S\,\rangle = n(\vec{r})$$

can be solved. Taking existence for granted, the uniqueness is assured by the Hohenberg-Kohn theorem, which also holds for $\hat{W} = 0$. This theorem states that the mapping between the external potential $v(\vec{r})$ of an interacting many-electron system and the corresponding ground-state density $n(\vec{r})$ is injective. Therefore, $v_S[n;\, \vec{r}\,]$ is a functional of $n(\vec{r})$. The same holds for all the orbitals φ_i, which are the solutions of the single-particle Schrödinger equation

$$\left[\, \frac{-\hbar^2}{2m}\, \nabla^2 + v_S(\vec{r}) \,\right] \varphi_i(\vec{r}) = \varepsilon_i\, \varphi_i(\vec{r})\,.$$

Applying the Hohenberg-Kohn theorem or the Levy-Lieb constraint search to non-interacting systems one concludes that there exists a unique energy functional

$$E_S[n] = T_S[n] + \int v_S(\vec{r})\; n(\vec{r})\, d^3 r$$

for which the condition $\delta E_S[n] = 0$ yields the ground-state density corresponding to $\hat{H}_S$. $T_S[n]$ is the universal kinetic-energy functional for non-interacting particles.

In the absence of degeneracies, the ground state $|\,\Psi_S\,\rangle$ of the non-interacting problem is a Slater determinant (fully antisymmetrized product) which is called Kohn-Sham determinant. The density is then given by

$$n(\vec{r}) = \sum_{i=1}^{N} |\varphi_i(\vec{r})|^2$$

and the non-interacting kinetic energy is given by

$$T_S[n] = \sum_{i=1}^{N} \int \varphi_i^*(\vec{r}) \left(\frac{-\hbar^2}{2m} \nabla^2 \right) \varphi_i(\vec{r})\, d^3r.$$

The main and most remarkable advantage of introducing the set of auxiliary orbitals φ_i is that they allow us to determine easily the kinetic energy $T_S[n]$ of the non-interacting system, as well as its functional derivative.

Applying the extremal condition (7.7) to the non-interacting system we have

$$\delta E_S[n] = \int \left[v_S(\vec{r}) + \frac{\delta T_s}{\delta n(\vec{r})} - \mu \right] \delta n(\vec{r})\, d^3r = 0\,,$$

which implies

$$\frac{\delta T_s}{\delta n(\vec{r})} = -v_S(\vec{r}) + \mu\,. \tag{7.9}$$

This is an important relation which reveals the actual role of auxiliary potential $v_S[n]$ in the theory, as minus the functional derivative of an important uncorrelated contribution $T_S[n]$ to the kinetic energy $T[n]$.

In order to formulate explicit practical approximations to the universal functional $F_{LL}[n]$ it is useful to single out two qualitatively important uncorrelated contributions whose explicit form is known. One therefore writes $F_{LL}[n]$ as

$$F_{LL}[n] = T_S[n] + \frac{e^2}{2} \int \frac{n(\vec{r})\, n(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d^3r\, d^3r' + E_{XC}[n]\,. \tag{7.10}$$

The first term $T_S[n]$ in Eq. (7.10) is the kinetic energy of the auxiliary non-interacting system with the density $n(\vec{r})$. An explicit form for $T_S[n]$ will be given below. The second term is the Hartree energy, which represents the classical electrostatic repulsion of the electronic density $n(\vec{r})$ with itself. Finally, the third yet unknown term is the so-called *exchange and correlation energy-functional*. In fact, as it will be discussed below, this is not a very accurate name.

Comparing Eq. (7.5) and (7.10) one observes that E_{XC} is the sum of two contributions

$$E_{XC}[n] = \underbrace{W[n(r)] - E_H[n]}_{\substack{\text{Exchange and correla-} \\ \text{tion Coulomb energy}}} + \underbrace{T[n] - T_S[n].}_{\text{Correlation kinetic energy}} \tag{7.11}$$

The first one takes into account the effects of the antisymmetry of many-fermion wave functions (exchange) as well as changes in the Coulomb energy beyond the single-determinant level (correlations). The second one incorporates the contribution of electron correlations to the kinetic energy, i.e., the differences between $T[n]$ and the kinetic energy

$T_S[n]$ of the non-interacting (uncorrelated) auxiliary system having the same density $n(\vec{r})$.

For the interacting system we have

$$\delta E[n] = \int \left[v(\vec{r}) + e^2 \int \frac{n(\vec{r}')\, d^3r'}{|\vec{r} - \vec{r}'|} + \frac{\delta E_{XC}}{\delta n(\vec{r})} + \frac{\delta T_S}{\delta n(\vec{r})} - \mu \right] \delta n(\vec{r})\, d^3\vec{r} = 0.$$

Since this holds for all $\delta n(\vec{r})$ the integrand must vanish. Using the relation (7.9) we replace $\dfrac{\delta T_S}{\delta n(\vec{r})}$ and obtain

$$v_S(\vec{r}) = v(\vec{r}) + e^2 \int \frac{n(r')\, d^3r'}{|\vec{r} - \vec{r}'|} + v_{XC}(\vec{r})\,,$$

where

$$v_{XC}[n; \vec{r}] = \frac{\delta E_{XC}}{\delta n(\vec{r})}$$

is the exchange and correlation potential, i.e., the functional derivative of E_{XC} which depends on $n(\vec{r})$ and $\vec{r}$.

The derived functional relation between $v_S[n; \vec{r}]$ and $n(\vec{r})$ allows us to determine the ground-state density and energy by solving the auxiliary non-interacting electron problem. The result is exact provided that $n(\vec{r})$ is non-interacting v-representable. Otherwise we would not be able to express it as the superposition of densities of non-interacting particles. The challenge remains to approximate $E_{XC}[n]$ accurately, since this defines v_{XC}, from which the ground-state density and energy are obtained.

References

[1] K. Gottfried, Quantum Mechanics Volume I: Fundamentals (Benjamin, New York, 1966); K. Gottfried and T.-M. Yan, Quantum Mechanics Volume I: Fundamentals, 2nd edition (Springer, 2004).

[2] M. Tinkham, Group Theory and Quantum Mechanics (McGraw Hill, New York, 1964); Reprinted by Dover, Mineola, 2003.

[3] L.D. Landau and E.M. Lifshitz, Quantum Mechanics Non-relativistic Theory (Pergamon, Oxford, 1958).

[4] A. L. Fetter and J. D. Wallecka, Quantum Theory of Many-particle systems (Mc Graw Hill, San Farncisco, 1971); Reprinted by Dover, Mineola, 2003.

[5] J.J. Sakurai, Modern Quantum Mechanics (Addison Wesley, 1985)

[6] J.J. Sakurai, Advanced Quantum Mechanics (Addison Wesley, Reading, 1967)