Faculty of Physics

UNIVERSITY OF BELGRADE

Incomplete lecture notes for

QUANTUM MECHANICS

Ivanka Milošević and Milan Damnjanović

Belgrade 2022

PREFACE

Tekst je i ove godine dopunjavan delovima predavanja. Jedan deo zadataka sa vežbi je promenjen, pre svega zbog promene asistenta na predmetu (Marko Milivojević), ali i zbog načina selekcije. I dalje se ovo može smatrati samo pratećim materijalom predavanja, a nikako ne kompletnim i pročišćenim odrazom ispredavanih lekcija. To će, nadamo se, postati sledećih godina. Tekst dobrim delom nije proveravan nakon pisanja, a posebno ne lektorisan. Stoga treba biti skeptičan prema formulama (!), pa čak i nekim formulacijama, jer prilikom elektronskog editovanja svaka nepažnja može da bude vrlo kreativna. Dalje, nekompletan je: mada su sva predavanja obradjena, i u tom smislu tekst pokriva kurs, česte indikacije podnaslova ukazuju na projektovana proširenja neophodna za potpunije sagledavanje pojedinih tema, otkrivajući najvažnije konceptualne celine koje kursom nisu obuhvaćene. Slike i primeri koji treba da ilustruju sadržaje krajnje su redukovani. Zato molim studente da tekst bude pre svega podsetnik za ono sto treba uraditi, a da za konačnu pripremu ispita koriste dopunsku literaturu.

3.10.2015, M.D.

Neplanirano sam odlučio da pročistim ove beleške za predavanja: ispravio sam dosta manjih grešaka (velike se nikad ne uoče) i donekle reorganizovao sadržaj. Ovo je uzrokovano željom da se ovom tekstu omogući da u budućnosti odslika ono u šta sam uvek verovao (a mentori me svakako učvrstili u tome), i što je verovatno najkraća definicija sada već često korišćene odrednice "Beogradska škola": struktura pojmova sama nameće minimalni sadžaj, uz dodatke koji se prave (možda i predaju) samo kada je to zaista potrebno, odnosno kada su to klice (naznake) nekih oblasti fizike. Tako je nova glava o kvantnoj informaciji zasada samo vesnik priče o jednoj od najmodernijih oblasti moderne fizike: nadam se, kao i za ostale delove teksta, da će biti proširena od strane drugih đaka te naše škole, onih koji su značajno doprineli i stalno doprinose da kvantna informacija i kvantno računanje budu danas toliko značajni u razumevanju sveta. Kao i do sada, udžbenik prof. Herbuta je osnovna literatura, a za pojedine lekcije su precizirane moguće specifične dopune iz svetske literature.

10.5.2018, M.D.

Od januara 2020. ovaj tekst zajedno sređujemo i proširujemo (nakon penzionisanja, M.D. je nastavnik na doktorskim studijama), sa istim principima, i sličnim pogledima kao i ranije, nastavljajući skoro poluvekovnu tradiciju Beogradske škole.

1.10.2020., I.M. and M.D.

Contents

Chapter 1

Quantum Kinematics

1.1 Measurement: System, State, Observable

This section introduces fundamental notions of quantum formalism, such as system, state, ensemble, measurement, etc. All these notions are common ingredients of any physical theory, but it turns out that their thorough analysis is not really necessary for classical mechanics. Further, we try to emphasize that these basic notions cannot be introduced independently, i.e. to define one of them the other ones should also be clarified. So, we start with intuitive analysis, being not essentially related to quantum mechanics, in order to single out the main features of the introduced notions. Their rigorous quantum definitions will be given in the next section within three postulates, which are also mutually dependent and, together with definition of measurement, they should be understood as one whole.

1.1.1 Intuitive analysis

System (quantum) is an object $\boldsymbol{S}$ having some previously defined set of properties. This set is exhaustive enough to make distinction between various systems. E.g., each electron has same mass, charge, spin etc., and other particles are differing by some of them. This characterization by well defined properties is time (meaning technology) dependent: a hundred years ago, when mass and charge were the only known characteristics distinguishing elementary particles, many of them would be considered the same. Therefore, recognition of a particular system consists of a sufficient set of *measurements* of its properties.

The set of properties defining a system still leave some non-uniqueness: systems of the same type differ by some additional characteristics. Thus, one can gather systems with certain given values of some (or all) of these additional properties into ensemble ρ. We usually say that various ensembles appearing this way differ by the state (characterizing included systems). In this sense, the notions of ensemble and its state are synonyms: one says "state of the ensemble" and "ensemble in the state" (although the misleading phrases "state of the system", or "system in the state" are usually used instead). An example may be an ensemble of electrons with given momentum, i.e. electrons in a state with given value of momentum. Note that this notion is based on the set of measurements of some properties, as well as the primary notion of the system. Consequently, two ensembles are in the same state if all the measurements performed on them give the same results. Again, the notion of the state is dependent on our knowledge of the possible properties of the system. Important example is that before discovery of the spin, ensembles (i.e. states) differing only in the value of the spin were considered the same.

Intuitively, measurement is any process of determining of some property of an ensemble. In fact, the very idea of a property of the system is that it is observable, i.e. measurable. This means that for each property (observable) A there is at least one measuring device (in this sense, the notions of the physical quantity and of the device measuring it may be identified), apparatus $\boldsymbol{A}$, capable to distinguish between various values of A: each value corresponds to a particular position of the apparatus' pointer. In other words, in the course of the interaction of the apparatus with a particular system one of these values is realized. Two important facts should be emphasized in this context. Firstly, various members of the same ensemble may produce different pointer positions. Therefore, measurements on the ensemble necessarily have statistical nature. Secondly, before the measurement is performed, nothing can be said about the measured property. This may be interpreted such that the measured property does not exist without the apparatus, or even that the property is realized or imposed by measurement.

The statistical nature of the measurement is well known even within the classical framework: a measurement does not give the result with certainty, but many measurements are performed, and the result is obtained by statistical analysis. Therefore, a single measurement is meaningless. Further, since the measurement is an interaction of apparatus and system, it may change, or even destruct the system (this is particularly important for small systems), and it may be impossible to repeat the procedure on the same system. Therefore, in general, measurements are performed on the ensembles of systems. Nevertheless, as it will be stressed out in the analysis of double slit experiment (Subsection 1.1.2), the necessity for the statistical approach and ensembles in quantum mechanics stems from additional, quite substantial reason.

To summarize, measurement means *measurement of some physical observable A on the ensemble (in the state) ρ, giving as the result probability distribution of the possible values of A.* Precisely, let $\sigma(A) = \{a_1, a_2, \dots\}$ be the set of the possible values of A (i.e. of the positions of the pointer of the apparatus, defined independently of ρ); each particular system from the measured ensemble ρ interacts with apparatus successively, and due to this interaction the pointer gets a series of values a_i from $\sigma(A)$. Let in the course of the measurement each value a_i is obtained N_i times. Obviously, $N = \sum_i N_i$ is the number of the systems in the ensemble. Then the result of the measurement is the probability distribution $v(a_i, A, \rho) \overset{\text{def}}{=} N_i/N$. Since the only criterion of validity of any physical theory (and quantum mechanics in particular) is the comparison with the experiment, fundamental task of the theory is to predict the probability distribution $v(a_i, A, \rho)$ for any ρ and for any A, in terms of its formalism. Before proceeding further in this direction, several remarks should be made.

Firstly, we note here that the (quantum) theory can be well-founded only with infinite ensembles, and this will be assumed throughout the text. However, in real experiments N must be finite, but large enough to provide reliable statistics, i.e. the statistics enabling comparison to the theoretical $N = \infty$ limit.

Secondly, since only the measurement results are verifiable, they are the only objective of quantum mechanics and thus quantum mechanical statements always refer to them. Therefore the underlying logic is completely subdued to this fact, with all the statements based on the results of the measurements. Indeed, the logic of quantum mechanics is not classical, Boolean logic (as will be clarified subsequently). Considering the statements as the answers to the corresponding questions, the above formulated task of a physical theory means its ability to give the answers (i.e. predictions) to the questions which can be at least in principle answered in terms of the experimental results. Together with the elementary statements about probabilities of results, also the statements which can be reduced to the elementary ones are legitimate. For example, besides the question about probability of a single value a_i, one can search for the probability $v(\{a_i, a_j, \dots\}, A, \rho)$ to get any of the values from the subset $\{a_i, a_j, \dots\}$ of $\sigma(A)$; obviously, this must satisfy

$$v(\{a_i, a_j, \dots\}, A, \rho) = v(a_i, A, \rho) + v(a_j, A, \rho) + \dots$$

Similarly, mean value of the observable A in the state ρ

$$\langle A \rangle_\rho = \sum_i a_i v(a_i, A, \rho) \tag{1.1}$$

may be looked for.

For further convenience, in the case when $v(a_i, A, \rho) = 1$ for some particular value a_i and consequently $v(a_j, A, \rho) = 0$ in all other cases $i \neq j$, we will say that the state ρ is with the sharp value a_i of the observable A.

Further, it is easy to understand that due to its infinity, each ensemble ρ can be partitioned into (also infinite!) subensembles $\rho_1, \rho_2, \dots$ in many ways. The weight w_I of the subensemble is defined as its relative part in the ensemble. Considering the process of the partition as taking one system after another from ρ and putting it subsequently into one of ρ_I, in the N-th step one finds the numbers N_I of the systems in the subensembles ρ_I, and the corresponding weights $w_I = \lim_{N\to\infty} N_I/N$ can be defined. Obviously, w_I is finite, positive and $\sum_I w_I = 1$. Also, having ensembles $\rho_1, \rho_2, \dots$ (with the same type of the system) one can make their mixture just by joining all of them into the single ensemble ρ. As soon as the mixture is made, it cannot be reconstructed in which way it actually has been formed., i.e. the corresponding partition can in no way be distinguished from the (infinitely many) other partitions.

Being also ensemble, each subensemble is suitable for measurements. Of course, the results of the subensemble measurements are in general mutually different, and also differ from the results of the ensemble measurement. In fact, the results of the subensemble measurements $v(a_i, A, \rho_I)$ are to the ensemble one related by

$$v(a_i, A, \rho) = \sum_I w_I v(a_i, A, \rho_I). \tag{1.2}$$

Still, there are ensembles giving for any observable A the same results $v(a_i, A, \rho_I) = v(a_i, A, \rho_J)$, and therefore $v(a_i, A, \rho) = v(a_i, A, \rho_I)$ for each subensemble ρ_I. These are called homogeneous ensembles or pure states. All other ensembles are called inhomogeneous, or mixed states. Pure states characterize maximally prepared ensembles. In fact,

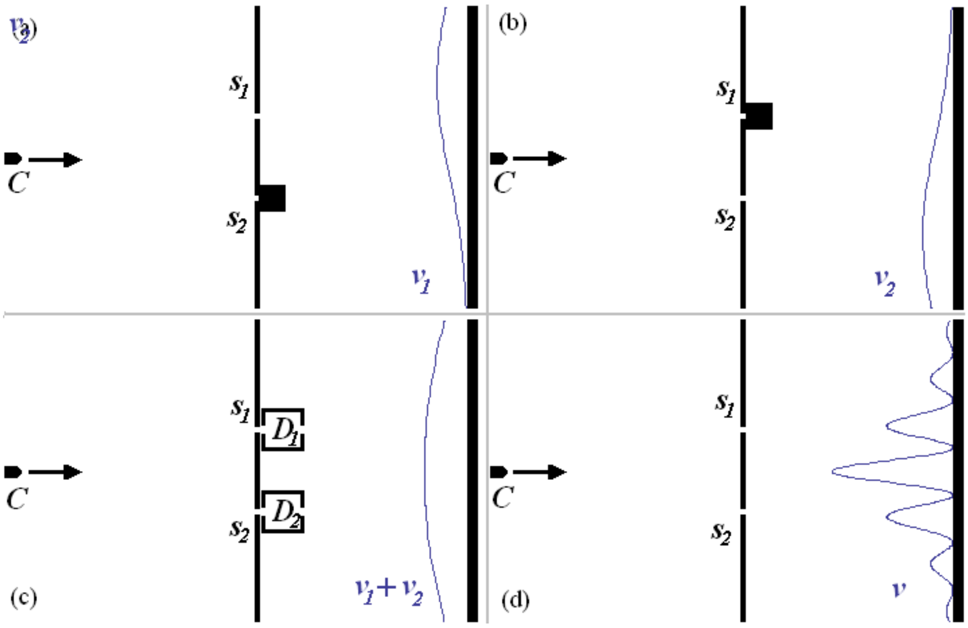

Figure 1.1: Double slit experiment. (a) Device: C is source of collimated photons, s_1 and s_2 are slits on the screen, X is detector of the particles (e.g. photo plate); D_1 and D_2 are detectors of the particles on the slits. (b) Results, being the probability distributions (intensities) $v(x)$ obtained by X: v_i $(i = 1, 2)$ is obtained when only s_i is opened, $v_{1\nLeftrightarrow 2}$ is their normalized sum, while curve v is obtained when both slits are open.

recall that the measurements required to define a particular system leave some freedom for different states which are distinguished between by some additional measurements. If all the *possible additional measurements* are performed, and ensemble is made only of the systems possessing the wanted values of measured observables, the homogeneous ensemble is obviously obtained. Despite we used here the intuitive notion of the possible (additional) measurements, the notion is quite nontrivial, and it will be neatly reconsidered within quantum mechanical formalism, yielding to the one of the most striking differences of classical and quantum physics.

1.1.2 Superposition: Experiments and Phenomenology

Let us neatly consider the well-known double slit experiment. Here we discuss only the simplest scheme. Still, simultaneously this is an experimental scheme capable to analyze all quantum mechanical situations, and with various modifications (e.g. light polarizers) it is used to demonstrate the most subtle questions (there are many texts with detailed interpretation of different setups [7][Ch. I]).

The experimental device is given in Fig. 1.1. The source (C) prepares homogeneous ensemble ψ directed toward the plate. Starting from the source, systems pass through slits s_1 and s_2, and come to the detector. Obviously, detector is measuring device showing position X of the incoming systems, i.e. the result of this measurement is the probability distribution $v(x, X, \psi)$.

If only slit s_1 is open, the result is distribution v_1 (Fig. 1.1, panel (a)), and similarly with only slit s_2 distribution v_2 is obtained (Fig. 1.1, panel (b)). Now, we make two experiments with both slits open. In the first one detectors D_1 and D_2 are switched on (Fig. 1.1, panel (c)), providing thus the information which slit (i.e. path) is chosen. The resulting distribution[1] is $v_{1 \not\Leftrightarrow 2} = v_1 + v_2$. In the second experiment the both detectors are turned off (Fig. 1.1, panel (d)), and now, when the path is unknown, the experimental distribution is v.

In terms of the classical theory, the distribution $v_{1 \not\Leftrightarrow 2}$ corresponds to the ensemble of corpuscules, when the intensity is a sum of the intensities obtained with one of the slits closed. The distribution v is like interference of waves (of course, we did not analyze here the conditions enabling distribution v, like distance between slits, etc.), when the intensities are proportional to the square of fields: $v_1 \sim |\boldsymbol{E}_1|^2$, $v_2 \sim |\boldsymbol{E}_2|^2$, meaning that $v \sim |\boldsymbol{E}_1 + \boldsymbol{E}_2|^2$. However, two experiments show that the same system exhibits both corpuscular and wave-like behavior, depending on the experimental situation. This *wave-corpuscle dualism* essentially means that system is neither the corpuscle nor the wave, but some of its properties, being manifested in appropriate circumstances, can be interpreted in terms of these classical (perhaps intuitive, or, more likely, familiar after courses of classical mechanics) prototypes of behavior. More important is to realize that the interference corresponds to the lack of information on the path of the systems. In fact, we can say that the resulting probability is caused by the interference of possible paths, and not particles (which is excluded by the experiments in which particles come out from the source one after another). Thus, if not measured (by D_1 and D_2 in the above experiment), even the trajectory of the particle does not exist; only a superposition of the all possible paths corresponds to the experimental result. Finally, note that the path is not known without an additional measurement (by D_1 and D_2) because the prepared initial state does not prefer either of the two possibilities allowed by the slits on the screen. Thus, for such a state and such screen there are two equally probable paths. However, quite different situations may be constructed easily: for instance, initial state allows only one path (beam directed toward one slit only), thus paths are not equally probable, or there are more than two slits (possible paths).

1.1.3 Pure States

The afore described experiments, as well as many others performed with pure states can be explained consistently only if the superposition principle is introduced: if two pure states of a system are detected, then any linear combination of these states can be also realized. Obviously this implies that the pure states are vectors in a vector space, which will be called state space. Moreover, the superposition of the states with the same sharp value of a particular property, will be the state with the same value of that property. Otherwise, the notion of the system would be lost, since the superposition of the states of the same system (defined by the fixed values of the system defining properties) could change the value of these properties, and hereby the system.

---

[1] Here the opposite of equivalence $\not\Leftrightarrow$ symbol emphasizes that exactly one of two possible paths has been realized.

POSTULATE 1. – STATES

Pure state of the system S is represented by a unit vector in the state space $\mathcal{S}$ of S, and each vector represents a pure state of S. All the vectors differing only by the phase factor describe the same physical state.

1.1.4 Observables

To complete the description of the measurement process, one must include the property which is measured. Intuitively, it can be assumed that all the directly measured quantities are real, despite the fact that we can combine them to get complex or tensor quantities convenient for some specific purpose. This is to be understood in the sense that outcomes of the measurements are reals.

Further, let us single out some particular value a of the physical property A. Recall that any superposition of the pure states with the sharp value a of A results in the state with the same sharp value of A. This means that in the state space $\mathcal{S}$ there is a subspace $\mathcal{S}_a(A)$ including all such pure states. Consequently, taking into account all the possible measured values of A, the observable A introduces a decomposition of the state space into these subspaces: $\mathcal{S} = \oplus_a \mathcal{S}_a(A)$. These two entities (real values and subspaces) can be interpreted as the spectrum and eigen decomposition of $\mathcal{S}$, i.e. they define uniquely a hermitian operator in $\mathcal{S}$ by its spectral form:

$$A = \sum_a a P_a(A); \tag{1.3}$$

$P_a(A)$ denotes the eigenprojector of A for the eigenvalue a, i.e. the projector onto the eigenspace $\mathcal{S}_a(A)$.

POSTULATE 2. – OBSERVABLES

Each observable biuniquely corresponds to a hermitian operator in the system state space $\mathcal{S}$.

Thus, within the quantum formalism, physical observables are represented by the hermitian operators. On the other hand, this also means that whatever hermitian operator is considered (and in the infinite dimensional Hilbert spaces there are really strange ones), there is at least in principle possibility to construct the measuring device corresponding to it.

It is obvious that the eigenvectors of the various physical observables play an important role in the quantum formalism, as well as in the concrete calculations. It is convenient to denote the eigenvector corresponding to the eigenvalue a of A as $|a, \lambda\rangle$, where the counter λ distinguishes between various (linearly independent) eigenvectors for the same eigenvalue. Thus, an eigenbasis of A is

$$\{|a, \lambda\rangle \mid a \in \sigma(A),\ \lambda = 1, \dots, |\mathcal{S}_a(A)|\}.$$

Having an eigenbasis of A, one can use it to represent vectors and operators in $\mathcal{S}$ by the columns and matrices. Such a representation will be called A-representation.

Particularly, if A is complete or nondegenerate observable, i.e. if its eigenspaces are one-dimensional, $|a\rangle$ uniquely determines the physical state. If A is incomplete, since the commuting observables have common eigenbasis, the notation can be refined using instead of λ the eigenvalues of another commutative observable B. In fact, let us consider the basis $|a,\lambda\rangle$ in $\mathcal{S}_a(A)$, which is also the eigenbasis of B: $B\,|a,\lambda\rangle = b_\lambda\,|a,\lambda\rangle$. The eigenvalues a and b_λ characterizing the common eigenvectors are called compatible. If all the corresponding eigenvalues b_λ are different, then these eigenvalues completely characterize this basis, and dropping the redundant label λ, the vectors are naturally denoted as $|a,b\rangle$. If this is the case for all the eigenvalues of A, the common eigenbasis is unique. In general, the set of the observables $\{A^{(1)},\dots,A^{(N)}\}$ having uniquely determined common eigenbasis $|a^{(1)},\dots,a^{(N)}\rangle$ is called complete set of compatible observables (CSCO). This means that these are the unique vectors which satisfy the system of eigenequations:

$$A^{(i)}\,|a^{(1)},\dots,a^{(N)}\rangle = a^{(i)}\,|a^{(1)},\dots,a^{(N)}\rangle\,,\quad i=1,\dots,N.$$

1.1.5 Probability distribution

To complete the quantum formalism, it remains to interrelate states and observables (already incorporated into formalism) to get the prediction for result of the measurement of a particular observable on the system in some particular state. To this end we use spectral decomposition (1.3) to find the components from the eigenspaces of A of the pure state $|\psi\rangle$: $|\psi\rangle = \sum_a P_a(A)\,|\psi\rangle = \sum_a |\psi_a\rangle$. Since $|\psi\rangle$ is unit vector, the square norms of these components are summed in 1: $\sum_a \|\,|\psi_a\rangle\,\|^2 = 1$, as it is expected from any probability distribution. This hints that the probability to find a is $\|\,|\psi_a\rangle\,\|^2$. A straightforward generalization is that the probability of the results from the interval $[a_1,a_2]$ is $\sum_{a\in[a_1,a_2]}\|\,|\psi_a\rangle\,\|^2$, giving a justification of:

POSTULATE 3. – PROBABILITIES

When the measurement of the observable A is performed on the system in the pure state $|\psi\rangle$, the value from the interval $[a_1,a_2]$ is obtained with the probability

$$v([a_1,a_2],A,|\psi\rangle) = \|P_{[a_1,a_2]}(A)\,|\psi\rangle\,\|^2.$$

Later on this expression will be used also in the equivalent forms:

$$v([a_1,a_2],A,|\psi\rangle) = \langle\psi|\,P_{[a_1,a_2]}(A)\,|\psi\rangle = \mathrm{Tr}\,P_{[a_1,a_2]}(A)\,|\psi\rangle\,\langle\psi|\,. \tag{1.4}$$

For the observables with purely discrete spectrum within considered interval, $P_{[a_1,a_2]}(A)$ is sum $\sum_{a\in[a_1,a_2]} P_a(A)$ of the eigenprojectors related to the enclosed eigenvalues. However, if $[a_1,a_2]$ encloses only continual spectrum of A, there is no eigenvector for a particular continual "eigen-value" a (precisely, the point of the continual spectrum) and therefore $P_a(A)$ vanishes, as well as the related probability, $v(a,A,|\psi\rangle) = 0$. Still, the projector $P_{[a_1,a_2]}(A)$ is well defined with help of the spectral measure $E_a(A)$ associated to A: $P_{[a_1,a_2]}(A) = E_{a_2}(A) - E_{a_1}(A)$; in terms of the (generalized) eigenvectors, for a complete observable A (generalization is obvious), this is $P_{[a_1,a_2]}(A) = \int_{a_1}^{a_2} \mathrm{d}a\,|a\rangle\,\langle a|$. This corresponds to the intuitively expected finite probability to obtain result in some finite interval, and naturally invokes probability density at any point $a \in [a_1, a_2]$. Indeed, using A-representation $|\psi\rangle = \int \mathrm{d}a \psi(a) |a\rangle$ of the state $|\psi\rangle$, the first of (1.4) gives for the interval around a the probability: $v(a - \Delta a/2, a + \Delta a/2], A, |\psi\rangle) = \int_{a-\Delta a/2}^{a+\Delta a/2} |\psi(a')|^2 \,\mathrm{d}a'$. Its differential defines the probability density in a:

$$\tilde{v}(a, A, |\psi\rangle) = \lim_{\Delta a \to 0} \frac{v([a - \Delta a/2, a + \Delta a/2], A, |\psi\rangle)}{\Delta a} = |\psi(a)|^2. \tag{1.5}$$

The postulate clearly introduces the basic elements of quantum logic. Since it involves the state which is measured, and the projector defining the expected event, one can generalize the notion of the postulate: referring to the projector as to the event (thus the subspaces are events, while the elementary events are one dimensional subspaces, i.e. rays) the expression $v(P, |\psi\rangle) = \|P\,|\psi\rangle\,\|^2$ is the probability of that event when the system is prepared in the state $|\psi\rangle$.

Note that each pure state $|\phi\rangle$ defines ray projector $P = |\phi\rangle\langle\phi|$. Then, there exists observable A with the spectral decomposition $A = \alpha P + \sum_a' a P_a(A)$ (summation over the eigenvalues different from α). The probability of the result α in the measurement of A in the initial state $|\psi\rangle$ is easily found by (1.4). Since it can be interpreted as the probability to find state $|\phi\rangle$ when the system is prepared in the state $|\psi\rangle$, it is called transition probability from $|\psi\rangle$ to $|\phi\rangle$:

$$v(1, |\phi\rangle\langle\phi|, |\psi\rangle) = v_{|\psi\rangle \to |\phi\rangle} = |\langle\phi \,|\, \psi\rangle|^2. \tag{1.6}$$

1.1.6 Mixed State

Immediately after introducing the notion of the ensemble, it has been pointed out that besides the homogeneous ones, there are also the inhomogeneous ones. While the former are described by the state vectors, the latter has not yet been incorporated into the formalism. To describe mixed states, we rely once again on the obvious facts concerning the results of measurements.

We consider an inhomogeneous ensemble ρ obtained by mixing, i.e. joining, the homogeneous ensembles: N_i replicas of the homogeneous ensemble described by $|\psi_i\rangle$ is taken together, for $i = 1, \dots, n$. Denoting by $N = \sum_i N_i$ the total number of the joined ensembles, the weight of each homogeneous component is given by $\omega_i = N_i/N$; naturally, weights ω_i satisfy $0 \le \omega_i \le 1$ and $\sum \omega_i = 1$. The probability distribution resulting in the measurement of the observable A is obviously the average of the distributions $v([a_1, a_2], A, |\psi_i\rangle)$ obtained in the measurements on the involved pure states:

$$v([a_1, a_2], A, \rho) = \sum_i \omega_i v([a_1, a_2], A, |\psi_i\rangle) = \mathrm{Tr}\, P_{[a_1,a_2]}(A)\rho, \tag{1.7}$$

where

$$\rho = \sum_i \omega_i |\psi_i\rangle\langle\psi_i|\,. \tag{1.8}$$

Comparing with (1.4), one finds that again $P_{[a_1,a_2]}(A)$ comprises the information on the measuring observable, while the remaining part, ρ, gives all available information about the system, i.e. ensemble. Hence, ρ is the entity describing the ensemble. For example, using A instead of its eigenprojectors in (1.7) one finds the meanvalue:

$$\langle A\rangle = \mathrm{Tr}\,\rho A. \tag{1.9}$$

Being a convex linear combination of the ray projectors $|\psi_i\rangle\langle\psi_i|$, ρ is obviously hermitian, positive and with unit trace. Thus any mixed state is described by a statistical operator. Note that the orthogonality of the states $|\psi_i\rangle$ has not been assumed, and therefore (1.8) is not a spectral form of ρ. Nevertheless, for any statistical operator there is an orthonormal eigenbasis $|r,\lambda\rangle$. However, the spectral decomposition

$$\rho = \sum_{r\lambda} r\,|r,\lambda\rangle\langle r,\lambda| \tag{1.10}$$

has the same form as (1.8), revealing that any statistical operator describes some mixture. In addition, we see that the same mixture is obtained in at least two different ways from pure states: spectral form and (1.8). Consequently, it cannot be reconstructed how the mixture is made, and the question of the pure components of a mixture is in general meaningless. Instead, one can show (see Exercise 1.1) only that the involved pure states are from the range of ρ.

Exercise 1.1: Show that the states $|\psi_i\rangle$ of (1.8) span the range of ρ.

From the construction it is obvious that mixtures are the most general states of physical systems. In addition, Gleason's theorem [8] shows that in any separable Hilbert space (with dimension greater than 2) the probability measures over the subspaces[2] biuniquelly correspond to the statistical operators by (1.7). Particularly, pure states are also mixtures with a single homogeneous ensemble involved, i.e. the statistical operator corresponding to the pure state $|\psi\rangle$ is according to (1.8) the ray projector $|\psi\rangle\langle\psi|$. The ray projectors are exactly the intersection of the sets of the projectors and statistical operators: projectors are positive operators, and their trace is the dimension of their range (thus, only the ray projectors are statistical operators); on the other hand, a statistical operator is idempotent if and only if it has one nondegenerate eigenvalue 1, and the other one 0, with degeneracy $|\mathcal{S}|-1$ (thus, it is a ray projector). For any ray projector P, there is a state $|\psi\rangle$ such that $P = |\psi\rangle\langle\psi|$, and all the vectors $\mathrm{e}^{\mathrm{i}\phi}|\psi\rangle$ differing by an arbitrary phase give the same ray projector. Thus, there is a biunique correspondence between pure physical states are in biunique correspondence to the ray projectors (idempotent mixtures) and, as we have seen, the statistical operator description is free of the phase ambiguity characteristic for the vector notation.

Sharp value a of the observable A in the mixed state ρ is by definition obtained when $\mathrm{Tr}\,P_a(A)\rho = 1$. Using any basis $|a,\lambda\rangle$ in $\mathcal{S}_a(A)$, the eigenprojector becomes $P_a(A) = \sum_\lambda |a,\lambda\rangle\langle a,\lambda|$, and $1 = \mathrm{Tr}\,P_a(A)\rho = \sum_\lambda \langle a,\lambda|\,\rho\,|a,\lambda\rangle$. Since ρ is positive operator, this means that the trace over the eigenspace $\mathcal{S}_a(A)$ is equal to the complete trace, and further that for any vector $|\psi\rangle$ orthogonal to $\mathcal{S}_a(A)$ one has $\langle\psi|\,\rho\,|\psi\rangle = 0$, i.e. that orthocomplement of $\mathcal{S}_a(A)$ is a subspace in the null-space $\mathcal{S}_0(\rho)$ of ρ. In other words, the

---

[2] Probability measure μ over the subspaces is a function μ mapping each subspace of $\mathcal{S}$ into nonnegative real number, such that $\sum_i \mu(\mathcal{S}_i) = 1$ for any orthogonal decomposition $\mathcal{S} = \oplus_i \mathcal{S}_i$.

sharp value a of A reveals that ρ is a mixture of the eigenstates of A for the eigenvalue a. Particularly, if a is nondegenerate, then $\rho = |a\rangle \langle a|$ is the pure state.

For each pure state there is a CSCO for which it is a common eigenvector. Therefore, the measurement of these observables reveal their sharp values. On the contrary, there is not proper mixed state with this property. Therefore, pure states correspond to the maximal possible information about system. This can be also seen in terms of entropy:

$$S = -\mathrm{Tr}\, \rho \ln \rho. \tag{1.11}$$

Calculating trace in the eigenbasis of ρ, with help of the spectral form (1.10), one finds $S = -\sum_i r_i \ln r_i$. Since $1 \geq r_i \geq 0$, the entropy S is always nonnegative. It takes its minimal value $S = 0$ for the states with a single nonvanishing (and automatically nondegenerate) eigenvalue $r_1 = 1$. Obviously, these states are pure states. In accordance with the fact that the pure states are maximally prepared ones, this shows consistency of the above definition of entropy.

In the equilibrium statistical physics of special importance is canonical distribution described by the canonical state

$$\rho_{\mathrm{can}} \overset{\mathrm{def}}{=} Z^{-1} \mathrm{e}^{-H/kT}, \quad Z = \mathrm{Tr}\, \mathrm{e}^{-H/kT}. \tag{1.12}$$

Here, T is temperature of the ensemble, H is hamiltonian (operator of energy) and Z is called statistical sum or partition function.

1.1.7 Measurement

Since all notions of quantum formalism are based on the results of measurements, a rigorous consideration of the measurement process itself is necessary.

Intuitively, measurement process is a series of acts of interactions of the apparatus with the particular members of ensemble; each of the acts involves a single ensemble member and results in a particular position of the part of apparatus called *pointer*. The pointer positions correspond to the eigenvalues of the measured observable, and the distribution of the outcomes coincides with the postulated quantum-mechanical prediction (1.4). Besides these characteristics which must be satisfied by every measurement device, measurements may affect the measured system in many different ways (and even destroy it!), depending on the type of the interaction involved . However, it is convenient to extract only the minimal process which yields the result, and to ignore all other details of the particular measurement process.

According to the instant to which the results refer, measurements may be classified as retrospective and predictive. In the former case, the results are valid immediately before the measurement, giving no information about the future of the systems. In contrast to this class of measurements, predictive measurements are not related to the history of the system, as the obtained results refer to the instants after the measurement. There are also measurements which are both retrospective and predictive. Particularly important type of the predictive measurement are selective measurements, enabling to separate subensembles which correspond to different outcomes. Such measurements are necessary in the preparations of the ensembles with wanted properties. Also, predictive measurement is used in the following idealization, which helps to understand the mentioned minimal change the quantum ensembles undergo in any measurement.

Definition 1.1 *Ideal measurement of the observable A is an interaction of the apparatus with ensemble in arbitrary state $|\psi\rangle$ such that:*

(i) *Each eigenvalue a of A is obtained with the quantum probability:*

$$v(a, A, |\psi\rangle) = \langle\psi| P_a(A) |\psi\rangle ;$$

(ii) *Immediately after the measurement the subensemble selected by the eigenvalue a has the sharp value a of A;*

(iii) *If any observable B which is compatible with A, had sharp value b in $|\psi\rangle$, this sharp value is retained immediately after the measurement.*

In fact, the requirement (i) assures that the considered process is really a quantum measurement, while (ii) formalize the intuitively necessary property of selective measurement: in the repeated measurement of A the selected subensembles show only the value by which they were selected. The statement (iii) enables to follow the change of the state in the course of the ideal measurement.

Theorem 1.1 (Collapse of the state, Projection postulate) *Ideal measurement of the observable $A = \sum_a aP_a(A)$ is performed. Then:*

(i) If the initial state of the system was the pure state $|\psi\rangle$, the state of the subensemble corresponding to the selective measurement of the eigenvalue a is

$$|\psi'_a\rangle = P_a(A) |\psi\rangle / \sqrt{\langle\psi| P_a(A) |\psi\rangle}. \tag{1.13a}$$

(ii) If the system initially was in the mixed state ρ, its state immediately after the (non-selective) measurement is

$$\rho' = \sum_a P_a(A)\rho P_a(A). \tag{1.13b}$$

▬*Proof:* (i) We expand the initial state in the form $|\psi\rangle = \sum_a P_a(A) |\psi\rangle$, and use the normalized components $P_a(A) |\psi\rangle / \sqrt{\langle\psi| P_a(A) |\psi\rangle}$ to construct the projector $Q = \sum_a \frac{P_a(A)|\psi\rangle\langle\psi|P_a(A)}{\langle\psi|P_a(A)|\psi\rangle}$. Note that the denominator is $v(a, A, |\psi\rangle)$. Obviously, $P_a(A)Q = QP_a(A) = \frac{P_a(A)|\psi\rangle\langle\psi|P_a(A)}{v(a,A,|\psi\rangle)}$, i.e. Q commutes with each spectral projector of A, and the product is the ray projector on the corresponding component of $|\psi\rangle$. The first conclusion also means that Q is compatible with A. Further, in the initial state it has sharp value 1: $v(1, Q, |\psi\rangle) = \sum_a \frac{\langle\psi|P_a(A)|\psi\rangle\langle\psi|P_a(A)|\psi\rangle}{v(a,A,|\psi\rangle)} = 1$. According to the third characteristics of the ideal measurement, after the ideal measurement Q retains this sharp value, meaning that in each subensemble its eigenvalue 1 value is detected with certainty. Particularly, in the subensemble selected by the eigenvalue a of A, it has this sharp value, as well as $P_a(A)$. Altogether, this subensemble has sharp value 1 of $QP_a(A)$. As this is the ray projector, the corresponding state is necessarily (1.13a).

(ii) Let $\rho = \sum_i \omega_i |\psi_i\rangle \langle\psi_i|$ be one of the decompositions of the mixed state onto the pure components. This means that the ensemble ρ can be divided into the subensembles $|\psi_i\rangle$ weighted by ω_i. After the selective measurements of different eigenvalues of A on the pure components $|\psi_i\rangle$, one gets the subensembles $\frac{P_a(A)|\psi_i\rangle}{\sqrt{v(a,A,|\psi_i\rangle)}}$ for each a and each i. The number of subsystems in each of these subensembles is $\omega_i v(a, A, |\psi_i\rangle)N$, where N is the total number of systems in the ensemble ρ and $v(a, A, |\psi_i\rangle) = \langle\psi_i| P_a(A) |\psi_i\rangle$. Hence, the resulting state is the mixture of all the subensembles, with weights $\omega_i v(a, A, |\psi_i\rangle)$, i.e. $\rho' = \sum_i \omega_i \sum_a v(a, A, |\psi_i\rangle) \frac{P_a(A)|\psi_i\rangle\langle\psi_i|P_a(A)}{v(a,A,|\psi_i\rangle)} = \sum_a P_a(A)\rho P_a(A)$. ∎

The theorem determines the change of state during the ideal measurement. Any other nonideal measurement may be understood as the minimal one followed by some dynamical process which is essentially irrelevant for the measurement outcome, but cannot be avoided due to the specific interaction of the system with the equipment.

Exercise 1.2: The space of all linear operators in $\mathcal{S}$ is the direct product $\mathcal{S} \otimes \mathcal{S}^*$, with naturally defined scalar product $(A, B) \overset{\text{def}}{=} \mathrm{Tr}\, A^\dagger B$. This space is usually called superspace, while the operators acting in it are called superoperators. Show that for each A the mapping $\Pi_A : \rho \to \rho'$ defined by (1.13b) is a superprojector.

Exercise 1.3: Find the states that are not changed in the ideal measurement of A.

The subensemble resulting from the selective measurement of an eigenvalue of an observable (or a set of compatible observables) is homogeneous for any pure or mixed measured state, if and only if the eigenvalue is nondegenerate. Such a measurements is called complete.

1.1.8 Interference: Superposition and Mixture

Thus far, we have developed the formalism to the level which enables us to fully describe the double-slit experiment. The emitted beam is prepared in the state $|\psi\rangle$ with a sharp value of a non-degenerate observable A. The screen with the slits serves as a measuring device of an observable B, with two emphasized nondegenerate eigenvalues (corresponding to the slits). Thus the slits correspond to the eigenstates $|\psi_1\rangle$ and $|\psi_2\rangle$. The observables A and B are not compatible; precisely, the initial state $|\psi\rangle$ (eigenstate of A) is not the eigenstate of B, and it is assumed that $\langle\psi_1 \,|\, \psi\rangle = \langle\psi_2 \,|\, \psi\rangle$. Then Eq. (1.6) shows that the probabilities to detect any of the two paths behind the screen with the slits are equal. Finally position of the particles is measured, i.e. the third observable X, which is complete.

Thus, when the detectors D_1 and D_2 are switched off , the measurement is nonselective with respect to the singled out eigenvalues of B, but selective in the view of the remaining eigenvalues of B. Effectively, this is a selective measurement of an observable B' with adequately modified spectral form of B. Namely, if $B = b_1 P_1[B] + b_2 P_2[B] + C$ (here $P_i[B] = |\psi_i\rangle \langle\psi_i|$), then $B' = b' P'[B'] + C$, where $P'[B'] = P_1[B] + P_2[B]$ is the eigenprojector. Then, according to (1.13a), the state behind the two slits screen is the superposition

$$|\psi'\rangle = P'[B'] |\psi\rangle / ||P'[B'] |\psi\rangle || = (|\psi_1\rangle + |\psi_2\rangle)/\sqrt{2}.$$

Hence, according to the probability postulate (1.4), the resulting probability is

$$v(x, X, |\psi'\rangle) = \frac{|\langle x \,|\, \psi_1\rangle + \langle x \,|\, \psi_2\rangle|^2}{2} = \frac{\psi_1^*(x)\psi_1(x) + \psi_2^*(x)\psi_2(x) + \psi_1^*(x)\psi_2(x) + \psi_2^*(x)\psi_1(x)}{2}.$$

When the detectors are switched on, the measurement of B is selective in the relevant eigenvalues: each of the taken paths makes $|\psi_1\rangle$ or $|\psi_2\rangle$ subensemble, and, according to (1.13b), the state behind the two-slits screen is the mixture:

$$\rho = P_1[B]\,|\psi\rangle\,\langle\psi|\,P_1[B] + P_2[B]\,|\psi\rangle\,\langle\psi|\,P_2[B] = (|\psi_1\rangle\,\langle\psi_1| + |\psi_2\rangle\,\langle\psi_2|)/2,$$

and in this case, (1.7) gives:

$$v(x, X, \rho) = \frac{1}{2}\psi_1^*(x)\psi_1(x) + \frac{1}{2}\psi_2^*(x)\psi_2(x).$$

Thus, we see that $v_{1\nLeftrightarrow 2}$ is obtained directly in this case, while in the previous one there are additional two terms describing the interference.

These results can be easily generalized. Let $|\psi\rangle = \sum_i \psi_i\,|i\rangle$. Then, the probability to detect the eigenvalue a of the observable A is

$$v(a, A, |\psi\rangle) = \sum_i |\psi_i|^2 v(a, A, |i\rangle) + \sum_{i\neq j} \psi_i\psi_j^*\,\langle a\,|\,i\rangle\,\langle j\,|\,a\rangle\,. \tag{1.14}$$

The first sum is exactly the probability obtained for the mixture $\rho = \sum_i |\psi_i|^2\,|i\rangle\,\langle i|$, while the second sum describes the interference. Also, $v(a, A, |\psi\rangle) = |\sum_i \psi_i\,\langle a\,|\,i\rangle\,|^2$ shows that the probability amplitude $\langle a\,|\,\psi\rangle$ is superposition of the contributing probability amplitudes $\langle a\,|\,i\rangle$. In other words, when the path is not known, the probability amplitudes of the possible paths are superposed, while, if the paths are known, the probabilities simply add.

When a mixed state $\rho = \sum_k \omega_k\,|\psi_k\rangle\,\langle\psi_k|$ is represented in the eigenbasis $\{|a\rangle\}$ of a complete observable A, the matrix elements $\rho_{ab} = \sum_k \omega_k\,\langle a\,|\,\psi_k\rangle\,\langle\psi_k\,|\,b\rangle$ relate the mixed pure states $|\psi_k\rangle$ to the measurement of A. Namely, as the diagonal elements $\rho_{aa} = \sum_k \omega_k v(a, A, |\psi_k\rangle)$ give the probability to detect a system in the state $|a\rangle$, they are called populations (of the states $|a\rangle$). The nondiagonal elements, according to (1.14), are the averaged (over the pure components) interference terms of the states $|a\rangle$ and $|b\rangle$ for the system being in each of the pure components $|\psi_k\rangle$. Thus, their nonvanishing values reveal that although the state is a mixture, there is a kind of remained interference from the coherent subensembles which explains why these elements are called coherences.

1.2 Quantization and State Spaces

1.2.1 Canonical Quantization

Classical description of any physical system is easily established, despite possibly unsatisfactory predictions. It is therefore natural to attempt to build the quantum description using the same ideas, or more precisely, to incorporate somehow the classical notions within quantum formalism, i.e. to find an algorithm to construct quantum formalism starting from the classical theory. To this end, the notions of the classical states and variables should be carefully reconsidered, and compared to their quantum analogues.

Within classical theory a system is described in terms of phase space and the space of variables. The states are points in the phase space, which is a manifold factorized onto the configuration space (generally, this is not vector space) and the momentum space. Usually, the dimensions of the both factors are the same, and proportional to the number of the particles in the system (mathematically precise, the phase space is cotangent bundle of the configuration manifold). Any physical quantity, called classical variable, is a function over the state space, $f(\boldsymbol{q},\boldsymbol{p})$. Thus, all the variables are functions of the coordinates and momenta, which are therefore called fundamental variables. In accordance with the profound continuity principle of physics, these functions are differentiable. Note that in each point of the phase space, all the variables have well defined values, or, using quantum vocabulary, each variable f has sharp value $f(\boldsymbol{q},\boldsymbol{p})$ in any classical state. Therefore, for each pair of classical states there are variables having different values in them, i.e. distinguishing between these states. A particular choice of such variables is the fundamental set $\{\boldsymbol{q},\boldsymbol{p}\}$.

Exercise 1.4: Show that the set of functions over phase space, i.e. operators in the state space, is Lie algebra with respect to the Poison bracket, i.e. commutator: $[\ ,\ ]$ is bilinear, antisymmetric and satisfies Jacobi identity. Also, show that $[X,\ ]$ for fixed (function, i.e. operator)X is a derivative (satisfies Leibniz's rule).

Structurally, the variables form a vector space, since linear combination of differentiable functions is also differentiable. In addition, pointwise multiplication $h(\boldsymbol{q},\boldsymbol{p}) \stackrel{\text{def}}{=} f(\boldsymbol{q},\boldsymbol{p})g(\boldsymbol{q},\boldsymbol{p})$ of variables defines a commutative multiplication $h = fg$, and the space of variables becomes an abelian associative algebra. It is also a Lie algebra, with the Poisson bracket $[a,b] \stackrel{\text{def}}{=} \sum_i(\frac{\partial a}{\partial q_i}\frac{\partial b}{\partial p_i} - \frac{\partial a}{\partial p_i}\frac{\partial b}{\partial q_i})$ as the Lie product.

On the other side, the set of quantum observables is real space of all hermitian operators acting in the state space $\mathcal{S}$. It is associative algebra with respect to the operator multiplication (composition). However, the crucial difference is noncommutativity of this algebra, implying that no state is a common eigen state for each observable, i.e. there is no state with a sharp value for any observable. To understand the difficulty and get a hint how to override it, we start with coordinate. If the corresponding observable $\hat{q}$ has to retain the basic idea of the classical coordinate, it must have continual spectrum $\sigma(q) = \mathbb{R}$. The corresponding eigenvectors $|q\rangle$ (i.e. $\hat{q}\,|q\rangle = q\,|q\rangle$), being the states with the sharp value q of $\hat{q}$, resemble to the classical states with well defined coordinate. However, there is continuum of the classical states with the sharp value q, distinguished by the value of the momentum. In analogy with the phase space, one may try to assume that each eigenvalue q of $\hat{q}$ is degenerate, and that all the degenerate vectors should be labelled by an additional value of the momentum. Accordingly, a complete characterization of these eigenvectors would be $|q,p\rangle$, with the simultaneously satisfied $\hat{q}\,|q,p\rangle = q\,|q,p\rangle$ and $\hat{p}\,|q,p\rangle = p\,|q,p\rangle$, implying that $\hat{q}$ and $\hat{p}$ commute. But this contradicts to the experiments, showing that well collimated beam of particles, although having more or less sharp value of momenta, is far from having sharp position and classical trajectory. Thus, the idea to define states by the values of the fundamental observables in them must be weaken: we have to build the state space with either coordinate or momentum eigenvectors only. This opens the question if these observables are complete.

Further, it seems natural to retain the functional relations between the classical variables in the quantum formalism, since these are well established experimentally and therefore formalism independent. In fact, this cannot be avoided, since the variables, like the angular momentum, are even defined as the functions of the fundamental ones. Thus, the quantum observables that are meaningful in the classical physics (like angular momentum, energy, etc.; in the Section 4.5, spin as a nonclassical degree of freedom, will be discussed) should be the same functions of the observales of coordinates and momenta. However, the product of the hermitian operators is hermitian if and only if they commute, implying that incompatibility of the observables suppress this simple analogy. In addition, commutation relations between the operators have already proved to be essential in the view of the simultaneous measurements. Since the commutator (multiplied by imaginary unit) is a Lie product in the space of observables, it has to be related to the classical Lie product, i.e. to the Poisson bracket.

Thus, the quantization procedure has to resolve the described problems, but also to provide the uniqueness of the quantum formalism.

POSTULATE 4. – QUANTIZATION

The quantum description of a physical system is obtained from the classical one by the following correspondence rules:

1. Linear combination of the variables corresponds to the same linear combination of the observables;
2. The product AB of two variables corresponds to the symmetrized product $(\hat{A}\hat{B}+\hat{B}\hat{A})/2$ of the observables;
3. The quantization is continual: any series $\{\hat{A}_i\}$ of observables corresponding to the convergent series $\{A_i\}$ of variables with limit A, is convergent with the limit $\hat{A}$ being the quantized variable A.
4. The basic set of variables corresponds to the basic set of observables in the way that Poisson bracket $[A,B]_{PZ}$ of any variable A with each basic variable B becomes the commutator $\frac{1}{i\hbar}[A,B]$.

The first rule is obvious, while the second one introduces the minimal change of the product to the symmetrized product, to regain the hermitian observable. However, mathematically, symmetrized product of more than two observables is not unique (e.g. ABA and A^2B+BA^2) unless some convention is invoked; in some other quantization schemes (geometric, Weil's, Schwinger's) this problem does not appear. Next rule, asserting that the limit and the quantization are mutually commuting operations, enables to quantize the variables which are transcendent functions of the fundamental set, using their series over the fundamental set.

The last rule is more subtle. Its role is to enable algebraic construction of the state space. To this end, the fundamental set of observables should be defined first, and the set of observables of coordinates and momenta is a natural choice by the analogy with the classical mechanics. Though this is not sufficient for the theories involving nonclassical degrees of freedom (such as spin or other interior degrees to be introduced much later), as far as quantization of the classical theories is considered, this approach is effective. Namely, in such situations all quantum observables are functions of coordinates and momenta. Accordingly, possible degeneracy of the observable of the coordinate would introduce another observables completing it to a CSCO. However, the last rule requires that classical analog A of an opservable $\hat{A}$ from such a CSCO, thus compatible with $\hat{Q}$, satisfies $[A,Q]_{PZ} = -\frac{\mathrm{d}A}{\mathrm{d}P} = 0$. This means that A is a function of Q only, $A = A(Q)$, and therefore $\hat{A}$ is the same function $\hat{A}(\hat{Q})$ of $\hat{Q}$, which is a contradiction: no function of $\hat{Q}$ can lower the degeneracy of $\hat{Q}$. Imaginary unit in front of the commutator has been anticipated as a way to obtain the hermitian Lie product, while the Planck's constant $\hbar$ with the dimension of action is necessary to compensate the dimension of the product of the coordinate and the momentum in the Poisson bracket. Its value $\hbar = 1.0545710^{-34}$ Js is determined by many very precise experiments. Note that for the coordinates and the momenta this reads:

$$[\hat{Q}_i, \hat{Q}_j] = 0, \quad [\hat{P}_i, \hat{P}_j] = 0, \quad [\hat{Q}_i, \hat{P}_j] = \mathrm{i}\hbar\delta_{ij}. \tag{1.15}$$

In general, canonically conjugate observables $\hat{A}$ and $\hat{B}$ satisfy $[\hat{A}, \hat{B}] = \mathrm{i}\hbar$.

Nevertheless, to enable later inclusion of the nonclassical degrees of freedom, the fundamental set of observables is to be defined in a more general way.

Definition 1.2 *The fundamental set is any set $F = \{\hat{A}_1, \dots, \hat{A}_f\}$ of observables such that:*

1. *the state space $\mathcal{S}$ is irreducible with respect to F, i.e. besides the trivial ones in $\mathcal{S}$ there is no other common invariant subspaces;*
2. *There is at least one subset in F being CSCO, such that the whole common eigen basis of this CSCO can be generated from a single common eigen vector by the functions of the fundamental observables;*
3. *the space $\mathcal{S}$ in which such fundamental set is defined is unique up to isomorphism.*

The first requirement is necessary since if the fundamental set is reducible with $\mathcal{S}_1$ and $\mathcal{S}_2$ being the common invariant subspaces, then the formalism can be completely developed in the two invariant subspaces independently. More striking is the consequence that the dynamics of the system, i.e. the change of the state during the time would be restricted to these subspaces, meaning that there would be in fact two physical descriptions of the same system. This intuitive paradox is prevented by the last requirement.

The second requirement is essential for the construction of the state space. This will be enlighten in the next section, when the state spaces of the elementary systems will be explicitly constructed as the Lebesgue space over the classical configuration space, with help of the family of the operator functions mapping one coordinate eigen vector into the others.

1.3 Construction of the state space

In this section the postulate of quantization will be utilized to construct the state space of the systems with classical analogs. Here come systems with N particles in D dimensions.

1.3.1 Single one-dimensional particle

This simplest system classically is described by the two-dimensional phase space with the coordinates $\{x, p\}$ being the fundamental variables with the commutators $[x, x]_{\mathrm{PZ}} = [p, p]_{\mathrm{PZ}} = 0$ and $[x, p]_{\mathrm{PZ}} = 1$. According to the quantization postulate, in the quantum formalism there is unique state space $\mathcal{H}$ with the opservables of coordinate $\hat{x}$ and momentum $\hat{p}$, such that $[\hat{x}, \hat{x}] = [\hat{p}, \hat{p}] = 0$ and $[\hat{x}, \hat{p}] = \mathrm{i}\hbar$.

***Exercise 1.5**:* Show that differential operator $\hat{p} = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$ is linear and Hermitian in the space of all differentiable wave functions which vanish at the both ends of interval (a, b).

***Exercise 1.6**:* Show that $[\hat{x}, \hat{p}] = \mathrm{i}\hbar$, $[\hat{x}^n, \hat{p}] = \mathrm{i}\hbar n\hat{x}^{n-1}$, $[\hat{A}, \hat{p}] = \mathrm{i}\hbar\frac{d\hat{A}}{dx}$, $[\hat{x}, \hat{p}^n] = \mathrm{i}\hbar n\hat{p}^{n-1}$, $[\hat{x}, \hat{B}] = \mathrm{i}\hbar\frac{d\hat{B}}{dp}$, where $\hat{p} = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$, $A = A(x)$ and $B = B(p)$ are differentiable functions in x and p, respectively.

An immediate consequence of these relations is:

Theorem 1.2 *The observable $\hat{x}$ has pure continual spectrum $C\sigma(\hat{x}) = \mathbb{R}$.*

▬*Proof:* Let us define the one-parameter family of the unitary operators $\hat{U}(q) \overset{\text{def}}{=} \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}q\hat{p}}$, for each real q. Using the exercise 1.6, it is easy to show that $[\hat{x}, \hat{U}(q)] = \sum_{n=0}^{\infty} \frac{(-\frac{\mathrm{i}}{\hbar}q)^n}{n!}[\hat{x}, \hat{p}^n] = \sum_{n=0}^{\infty} \frac{(-\frac{\mathrm{i}}{\hbar}q)^n n\mathrm{i}\hbar}{n!}\hat{p}^{n-1} = q\hat{U}(q)$. Further, if $x \in \mathbb{R}$ is in the continual spectrum of $\hat{x}$, then the (generalized) eigen vector $|x\rangle$ exists such that $\hat{x}\,|x\rangle = x\,|x\rangle$. But then, $\hat{x}(\hat{U}(q)\,|x\rangle) = ([\hat{x}, \hat{U}(q)] + \hat{U}(q)\hat{x})\,|x\rangle = \hat{U}(q)(q + \hat{x})\,|x\rangle = (x+q)(\hat{U}(q)\,|x\rangle)$. This means that $\hat{U}(q)\,|x\rangle$ is also eigen vector for the eigenvalue $x+q$. Therefore, if any real number x is in the spectrum, then all reals are also in the spectrum, since q is arbitrary. ▮

This is in accordance with the intuitive notion of coordinate (and momentum also): the sharp value x of the particle coordinate, i.e. completely localized particle, should be described by the vector $|x\rangle$, corresponding thus to the continual spectrum. Note the appearance of the translational operator family $\hat{U}(q)$, which, as mentioned in the definition of the FSO generates the whole eigen basis of $\hat{x}$ from each of its vectors. In fact, the nonuniqueness of the phase factor is resolved by convention

$$|x\rangle = U(x)\,|0\rangle\,, \forall x \in \mathbb{R}. \tag{1.16}$$

Hence, one can start with the basis $|x\rangle$, make the space over it, and check the other parts of the quantization procedure. At first, recall that $\hat{x}$ is complete observable; if not, it could be completed to CSCO, but then the canonical commutator rules would allow only functions of $\hat{x}$ as the compatible with $\hat{x}$ observables, which cannot complete $\hat{x}$. Further, the uniqueness of the obtained space follows from the fact that the Heissenberg Lie algebra (vector space over $\{\hat{x}, \hat{p}, \hat{I}\}$ with the commutators $[\hat{x}, \hat{p}] = \text{const}$ and $[\hat{x}, \hat{I}] = [\hat{I}, \hat{p}] = 0$) is solvable, meaning that its irreducible representations are either one-dimensional or infinite dimensional one; then, in the latter case, the obtained basis determines coordinate representation with well defined fundamental observables: opertors of coordinate ($\hat{x}$ in its eigenbasis is multiplicative operator), and of momentum (derivative operator $\hat{p} = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}$, proof follows immediately) determine fully, thus uniquely, the complete quantum formalism.

Coordinate representation

Thus, the task of the construction of the space is fulfilled. Each vector can be expanded over the obtained basis $|\psi\rangle = \int_{\mathbb{R}} \psi(x)\,|x\rangle\,\mathrm{d}x$. The corresponding Fourier coefficients $\psi(x) = \langle x\,|\,\psi\rangle$ give the coordinate representation of the state, which is called wave function. Instead of the point in the phase space, the state is therefore described by the functions. The density of probability to find the particle at x is $\psi^*(x)\psi(x)$. The finite norm functions, i.e. the functions over $\mathbb{R}$ satisfying $\int_{\mathbb{R}} |\psi|^2\,\mathrm{d}x < \infty$ form the Lebesgue space $\mathcal{L}(\mathbb{R})$. Note that the vectors $|x\rangle$ are not from this space, since their coordinate representation is Dirac δ-function: $|x'\rangle = \int_{\mathbb{R}} \delta(x-x')\,|x\rangle\,\mathrm{d}x$. In the coordinate representation the observable $\hat{x}$ becomes the multiplicative operator: since $\hat{x}\,|\psi\rangle = \int_{\mathbb{R}} \psi(x)\hat{x}\,|x\rangle\,\mathrm{d}x = \int_{\mathbb{R}}(x\psi(x))\,|x\rangle\,\mathrm{d}x$, the function $\psi(x)$ is mapped by $\hat{x}$ to $x\psi(x)$. Therefore, the domain of the coordinate $\hat{x}$ is the subspace in $\mathcal{L}(\mathbb{R})$ with the functions satisfying that also $\int_{\mathbb{R}} x^2|\psi|^2\,\mathrm{d}x < \infty$. Finally, since $\hat{p} = \mathrm{i}\hbar\left[\frac{\mathrm{d}\hat{U}(q)}{\mathrm{d}q}\right]_{q=0}$, the coordinate representation of $\hat{p}$ is found to be the derivative operator

$$\hat{p} = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}. \tag{1.17}$$

Namely, $\hat{p}\,|\psi\rangle = \mathrm{i}\hbar\left[\frac{\mathrm{d}\hat{U}(q)}{\mathrm{d}q}\right]_{q=0}\int_{\mathbb{R}} \psi(x)\,|x\rangle\,\mathrm{d}x = \mathrm{i}\hbar\left[\frac{\mathrm{d}}{\mathrm{d}q}\int_{\mathbb{R}} \psi(x)\,|x+q\rangle\,\mathrm{d}x\right]_{q=0} =$ $\mathrm{i}\hbar\int_{\mathbb{R}}\left[\frac{\mathrm{d}\psi(x-q)}{\mathrm{d}q}\right]_{q=0}|x\rangle\,\mathrm{d}x = -\mathrm{i}\hbar\int_{\mathbb{R}}\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\,|x\rangle\,\mathrm{d}x$. Its domain is comprised by the functions from $\mathcal{L}(\mathbb{R})$ for which also the image is in $\mathcal{L}(\mathbb{R})$, i.e. such that $\int_{\mathbb{R}} |\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}|^2\,\mathrm{d}x < \infty$.

Momentum representation

The eigenproblem of the momentum operator (1.17) is easily solved. Its spectrum is continual and nondegenerate, $C\sigma(\hat{p}) = \mathbb{R}$, with the corresponding eigenvectors (generalized, distributions) normalized to Dirac's δ-function

$$|p\rangle = \int_{\mathbb{R}} \langle x\,|\,p\rangle\,|x\rangle\,\mathrm{d}x, \quad \langle x\,|\,p\rangle = \psi_p(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}px}, \quad \langle p\,|\,p'\rangle = \delta(p-p') \tag{1.18}$$

known as plane waves.

As the spectrum is nondegenerate, momentum operator is itself a CSCO, enabling to represent vectors and operators in momentum representation. As for the states, they are again represented by functions, $\psi(p) = \langle p\,|\,\psi\rangle$. While the momentum operator in this representation is the multiplicative operator, $\hat{p}\psi(p) = p\psi(p)$, the coordinate operator is the derivative, $\hat{x}\psi(p) = \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}p}\psi(p)$.

Clearly, the form of the plane waves reveals that the transition between coordinate and momentum representation is realized by the Fourier transformation:

$$\psi(p) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}px}\psi(x)\,\mathrm{d}x, \quad \psi(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}px}\psi(p)\,\mathrm{d}p. \tag{1.19}$$

Generally, *Fourier-Plancherel* operator F is defined by $F\,f(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}px}f(x)\,\mathrm{d}x$; in $\mathcal{L}(\mathbb{R})$ it is unitary, but can be analogously defined in some extended spaces (e.g. it interrelates eigenbases of $\hat{p}$ and $\hat{x}$, as can be easily checked).

1.3.2 Three dimensional particle, many particles

According to the Fatu-Fubini theorem, Lebesgue space over $\mathbb{R}^n$ is direct product of the Lebesgue spaces over each $\mathbb{R}$: $\mathcal{L}(\mathbb{R}\oplus\cdots\oplus\mathbb{R})=\mathcal{L}(\mathbb{R})\otimes\cdots\otimes\mathcal{L}(\mathbb{R})$. Thus, each degree of freedom described by the real axis contributes by its Hilbert space $\mathcal{L}(\mathbb{R})$, and the total state space is product of these. For the discrete degrees of freedom, like spin, the corresponding space is finite dimensional.

This gives straightforward algorithm to construct state spaces of arbitrary systems. For a single particle, as its configurational space is three-dimensional (corresponding to three degrees of freedom), the state space is $\mathcal{S}=\mathcal{L}(\mathbb{R}^3)=\mathcal{L}(\mathbb{R})\otimes\mathcal{L}(\mathbb{R})\otimes\mathcal{L}(\mathbb{R})=\mathcal{L}^3(\mathbb{R})$; this space is usually called orbital space. Analogously, for N such particles, the state space becomes $\mathcal{S}=\mathcal{L}(\mathbb{R}^{3N})=\mathcal{L}^{3N}(\mathbb{R})$.

In analogy to the single one-dimensional particle, according to (1.15), there is common eigenbasis $|Q_{11},\dots,Q_{13},\dots,Q_{N3}\rangle$ of all the coordinate operators Q_{ni} ($n=1,\dots,N$, $i=1,2,3$). This basis defines the coordinate representation. Momentum operators of the n-th particle, expressed in the convenient vector form, are represented as:

$$\boldsymbol{p}_n=-\mathrm{i}\hbar\boldsymbol{\nabla}_n=-\mathrm{i}\hbar\Big(\frac{\partial}{\partial Q_{nx}}\boldsymbol{e}_x+\frac{\partial}{\partial Q_{ny}}\boldsymbol{e}_y+\frac{\partial}{\partial Q_{nz}}\boldsymbol{e}_z\Big).$$

Exercise 1.7: The translation operator $U(\boldsymbol{a})$ is defined to be such that $U(\boldsymbol{a})\psi(\boldsymbol{r})=\psi(\boldsymbol{r}-\boldsymbol{a})$. Show that $U(\boldsymbol{a})$ is unitary in $\mathcal{L}(\mathbb{R}^3)$ and express it in terms of $\boldsymbol{p}=-\mathrm{i}\hbar\boldsymbol{\nabla}$.

1.3.3 Subsystems: reduced states and correlation

As discussed in subsection 1.3.2, the system consisted of two subsystems (particles, degrees of freedom like coordinates) is described by the state space $\mathcal{S}=\mathcal{S}_1\otimes\mathcal{S}_2$, where $\mathcal{S}_1$ and $\mathcal{S}_2$ are the subsystems' state spaces. This space is spanned by the states being direct products of the single-particle states, which are called non-correlated or separable states, to distinguish from the correlated states, being their nontrivial linear combinations.

It is intuitively plausible that when such composite system is in the non-correlated state $|\psi\rangle=|\psi^{(1)}\rangle\otimes|\psi^{(2)}\rangle$, the subsystems are in the factor states $|\psi^{(1)}\rangle$ and $|\psi^{(2)}\rangle$, respectively. However, it is not obvious what is the subsystem state in the case of the correlated composite state.

As usual, these questions will be clarified through analysis of measurements. Particularly, as a state of a subsystem is manifested only in subsystem measurements, it should give correct results (1.4) in all the measurements of the first subsystem observables $A^{(1)}$. In the composite space, such observables induce composite observables $A^{(1)}\otimes\mathbf{1}^{(2)}$. Suppose that $A^{(1)}$ is a complete observable in $\mathcal{S}^{(1)}$, with the eigenbasis $|a\rangle$, while $|b\rangle$ is an arbitrary basis in $\mathcal{S}^{(2)}$. Then, any composite state $|\psi\rangle$ may be expanded over non-correlated basis $|a\rangle\otimes|b\rangle$ as

$$|\psi\rangle=\sum_{ab}\psi_{ab}\,|a\rangle\otimes|b\rangle=\sum_a|a\rangle\otimes|\psi_a\rangle\,,\quad |\psi_a\rangle=\langle a\,|\,\psi\rangle=\sum_b\psi_{ab}\,|b\rangle\,. \tag{1.20}$$

Here, $\langle a\,|\,\psi\rangle$ is the partial scalar product of the composite state $|\psi\rangle$ and the first subsystem state $|a\rangle$, giving a state of the second subsystem. In the selective measurement of $A^{(1)}\otimes\mathbf{1}$, the state of the subensemble corresponding to the pointer value a is obtained by the projection with $P_a^{(1)} \otimes \mathbf{1}$, since no requirement on the second subsystem is imposed. The result is $|a\rangle \otimes |\psi_a\rangle$, i.e. the product of the chosen eigenvector with the partial scalar product of the state with that eigenvector.

Of course, the partial scalar product may be defined also with the states of the second system. Using it, one further defines the partial trace of an operator. In fact, if the composite operator has in the Dirac notation the form $A = \sum_{aa'bb'} A^{ab}_{a'b'} |ab\rangle \langle a'b'|$, than its partial traces over two spaces, called the reduced operators, are:

$$A_1 = \mathrm{Tr}_2 A = \sum_b \langle b| A |b\rangle = \sum_{aa'} (\sum_b A^{ab}_{a'b}) |a\rangle \langle a'| \,, \tag{1.21a}$$

$$A_2 = \mathrm{Tr}_1 A = \sum_a \langle a| A |a\rangle = \sum_{bb'} (\sum_a A^{ab}_{ab'}) |b\rangle \langle b'| \,. \tag{1.21b}$$

It is easy to show that like the usual trace, the partial ones are independent on the used bases. Further, if A is hermitian, positive or statistical operator, both its partial traces are of the same types.

Exercise 1.8: Show that partial trace of the hermitian operator is hermitian. Analogously for positive and statistical operators.

After this technical preparation, we come back to the question of subsystem states. On the composite system in the most general (mixed) state ρ, the subsystem observable $A^{(1)} = \sum_a P_a^{(1)}$ is measured. The probability of the results is

$$v(a, A^{(1)} \otimes \mathbf{1}, \rho) = \mathrm{Tr}\,(P_a^{(1)} \otimes \mathbf{1})\rho = \sum_{a'b} (\langle a'| \otimes \langle b|)(P_a^{(1)} \otimes \mathbf{1})\rho(|a'\rangle \otimes |b\rangle) \tag{1.22}$$

$$= \sum_{a'} \langle a'| P_a^{(1)} (\sum_b \langle b| \rho |b\rangle) |a'\rangle = \mathrm{Tr}\, P_a^{(1)} \rho_1 = v(a, A^{(1)}, \rho_1). \tag{1.23}$$

This shows that the reduced density matrix ρ_1 (and analogously defined ρ_2) completely determines the results of the subsystem measurements. In this sense, the reduced statistical operators describe the states of the subsystem.

It is important to note that even in the case of the pure composite state $|\psi\rangle$ (i.e. $\rho = |\psi\rangle \langle\psi|$, the reduced states are not pure. The obtained mixed states in these cases are called the mixtures of the second kind to distinguishes from the usual mixtures obtained by the mixing of the different ensembles. Only for non-correlated composite states $|\psi^{(1)}\rangle \otimes |\psi^{(2)}\rangle$ reduced states are pure states $\rho_i = |\psi^{(i)}\rangle \langle\psi^{(i)}|$.

When initial state is separated, $\rho' = \rho_1 \otimes \rho_2$, the reduced states are just the factors ρ_1 and ρ_2 (this includes also the pure states: for $|\psi\rangle = |\psi^{(1)}\rangle \otimes |\psi^{(2)}\rangle$, the subsystems are in the factor states $|\psi^{(1)}\rangle$ and $|\psi^{(2)}\rangle$) as intuitively guessed at the beginning. This is the reason that such states sometimes are called classical.

Note that the reduced operators ρ_1 and ρ_2 do not completely describe the whole composite system. Indeed, they do not determine the composite state ρ, since different composite states ρ and ρ' may have the same both reduced states. For example, for each ρ giving ρ_1 and ρ_2, also the separable (or noncorrelated) state $\rho' = \rho_1 \otimes \rho_2$ has the same reduced operators. The difference between ρ and ρ' reflects the interrelation of the subsystem states in the composite state, and it is erased by the partial traces. In general, this interrelation is called correlation or entanglement, and the information on it is contained in $\rho - \rho'$, although there are some other derived quantities which can be used to measure it.

It is obvious that entropy of the state cannot be a measure of entanglement: it is sufficient to realize that the entropy of all pure states is 0, independently of their correlation, i.e. the same value is obtained for non-correlated and for correlated pure states. However, some information on the entanglement is obtained from the reduced statistical operators. Indeed, the entropy is generally a measure of the purity of the state (though rough, since there are states with the same entropy and non-similar statistical operators), and can be used to check if the reduced states are mixed, offering some information on the entanglement. *Entanglement entropy* is entropy of the reduced statistical operators. This means that two entropies are found: $S_i = -k\mathrm{Tr}\,\rho_i \ln \rho_i$. Especially, for the composite correlated pure state, the reduced statistical operators have the same spectral form (allowing different null-spaces; see Sectio **??**). Since only eigenvalues and their multiplicity determine the entropy of the statistical operator, this means that $S_1 = S_2$. Thus, this is taken as entanglement entropy of the pure bipartite state, giving insight to entanglement as rough as it is entropy classification of the mixed states.

1.4 Uncertainty relations

Every observable has sharp value in each of its eigenstates, but this is not the case for other states. *Dispersion* of the observable is a statistical measure of this non-sharpness:

$$\Delta_\psi(A) = \sqrt{\langle A^2\rangle_\psi - \langle A\rangle_\psi^2} = \sqrt{\langle (A - \langle A\rangle_\psi)^2\rangle_\psi}. \tag{1.24}$$

It vanishes in the eigenstates of A, and the eigenstates are the only nondispersive states for A. On the other hand, the nonvanishing dispersion in a general state, indicating nontrivial probability distribution, points to uncertainty of the prediction of the particular values of the pointer.

For a pair of observables, A and B, the product of the dispersions $\Delta_\psi(A)\Delta_\psi(B)$ is a measure of the uncertainty of the prediction of the pairs of the two pointers' (for A and B) values in the state ψ. Both observables have sharp values only if ψ is their common eigen state. If the observables are incompatible, such states may not exist at all. This indicates that quite generally the result of the measurement of the two observables contains an uncertainty which is related to the compatibility of the observables.

Theorem 1.3 (Heisenberg uncertainty relation) *For any state (pure or mixed) ψ and any pair of observables A and B the uncertainty of the probability distributions $v(A,\psi)$ and $v(B,\psi)$ has lower bound:*

$$\Delta_\psi(A)\Delta_\psi(B) \geq \frac{1}{2}|\,\langle [A,B]\rangle_\psi\,|. \tag{1.25}$$

▬*Proof:* For the state $|\psi\rangle$ we define auxiliary operators $A' = A - \langle A\rangle_\psi$, $B' = B - \langle B\rangle_\psi$, and vectors $|u\rangle = A'\,|\psi\rangle$, $|v\rangle = B'\,|\psi\rangle$. In the Schwartz inequality $\|\,|u\rangle\,\|\|\,|v\rangle\,\| \geq |\,\langle u\,|\,v\rangle\,|$, on the left is $\Delta_\psi(A)\Delta_\psi(B)$,

while on the right there is the mean $|\langle A'B'\rangle_\psi| = |\langle \frac{A'B'+B'A'}{2}\rangle_\psi + \langle \frac{A'B'-B'A'}{2}\rangle_\psi|$. Recall that the obtained symmetrized and antisymmetrized products are hermitian and scew-hermitian operators, and their means are real and pure imaginary, respectively. Thus, the absolute value of the whole expression is greater than that of its imaginary part. Altogether, $\Delta_\psi(A)\Delta_\psi(B) \geq |\langle A'B'\rangle_\psi| \geq \frac{1}{2}|\langle [A', B']\rangle_\psi| = \frac{1}{2}|\langle [A, B]\rangle_\psi|$.
∎

Exercise 1.9: Prove (1.25) for the mixed states.

In particular, canonically conjugated observables A and B (like coordinate and momentum) satisfy Heisenberg's relation:

$$\Delta_\psi(A)\Delta_\psi(B) \geq \frac{\hbar}{2}. \tag{1.26}$$

The states realizing the minimal uncertainty $|\langle [A, B]\rangle|/2$ are called minimal wave packets. To find them it should be examined when the inequalities used in the proof are equalities. Thus, the first one, the Schwartz inequality, degenerates into equality if and only if the related vectors are collinear, that is when $A'|\psi\rangle = \alpha B'|\psi\rangle$. The second inequality becomes an equality when the mean of the symmetrized product vanishes, i.e. when $0 = \langle\psi|(A'B' + B'A')|\psi\rangle = (\alpha^* + \alpha)\langle B'^2\rangle$. Thus, α must be pure imaginary.

Exercise 1.10: Show that the Gaussian $\psi(x) = \mathrm{e}^{-(\frac{x-a}{2f})^2 + \frac{\mathrm{i}}{\hbar}bx}$ is the coordinate-momentum minimal wave packet. What are a, b and f?

Using the uncertainty relation, one can estimate delocalization of a free particle having mean energy $\langle H\rangle$. Hamiltonian of such a particle has kinetic part only, $H = \frac{\hat{\boldsymbol{p}}^2}{2m}$. Therefore, for the one-dimensional particle: $\langle H\rangle = \frac{\langle \hat{p}^2\rangle}{2m} = \frac{(\Delta\hat{p})^2}{2m} + \frac{\langle\hat{p}\rangle^2}{2m} \geq \frac{(\Delta\hat{p})^2}{2m} \geq \frac{\hbar^2}{8m(\Delta\hat{x})^2}$, while for three dimensional particle this gives $\langle H\rangle \geq \frac{\hbar^2}{8m}(\frac{1}{(\Delta\hat{x})^2} + \frac{1}{(\Delta\hat{y})^2} + \frac{1}{(\Delta\hat{z})^2})$. Obviously, for better localization, more energy is needed.

For commutative observables $\Delta_\psi(A)\Delta_\psi(B) \geq 0$. In fact, $\Delta_\psi(A)\Delta_\psi(B) = 0$ for the common eigenstates, being nondispersive for both observables, in accordance with the introduced notion of compatibility. Still, noncommutative observables may have some common eigenstates (but not a common eigen basis!). In such states $\Delta_\psi(A)\Delta_\psi(B) = 0$.

Chapter 2

Quantum Dynamics

The main task of the physics is to predict events in the future on the basis of the known facts in the past. Unlike to the kinematics, which is hardly related to the system, as concerns only general properties, like number of components and consequent degrees of freedom, the dynamics involves all data characterizing the system: structure, mutual interactions of its components and possible external influences. Still, it must be noted from the very beginning that the exact dynamics is rarely solvable problem, and various approximations are used to simplify it, introducing different types of modeling of the system; this may impact even the kinematical description.

2.1 A classification of the physical systems

The classical mechanics solves the dynamical problem within canonical formalism, with the resulting canonical equations which involves Hamiltonian variable. It arises as the quantity governing the dynamics: through the variational principle, it singles out the geodesic of the motion. Therefore, to enable the direct canonical quantization, one should understand the domain of validity of the classical canonical formalism, and particularly the conditions enabling the hamiltonian approach. To this end an analysis of the relationship of the system and its surrounding is needed; it results in a quite general classification of the physical systems, with singled out types of systems enabling Hamiltonian approach.

In fact, even intuitively, the following classes of the physical systems are dynamically different (in the order of increasing complexity):

(i) System of free particles: there is no interaction between the subsystems (particles), and total Hamiltonian is sum of the single-particle Hamiltonians; the other properties are the same as for the next class.

(ii) Isolated system: although the subsystems interact mutually, they do not interact with the surrounding; thus the hamiltonian involves only system coordinates and momenta, preserving space-time homogeneity.

(iii) Conservative system: it may interact with the surrounding, and Hamiltonian may involve also some parameters characterizing the surrounding (external field breaking space homogeneity); still the system does not exchange the energy, and Hamiltonian

is time independent, i.e. time is homogeneous (e.g. particles in the conservative fields).

(iv) Dynamically separated system: this is a system which interacts with the surrounding in a very general manner, but still its dynamic may be singled out from the dynamics of the surrounding; Hamiltonian is dependent both on the time and on the external field parameters, but the system parameters are not involved in the equation of motion of the surrounding (no feedback), which allows the separation from the surrounding (e.g. electron in the accelerator).

(v) Kinematically separated system: only some degrees of freedom in some large system are used to single out kinematically independent subsystem; the dynamical equations may not be written independently for this part (e.g. atom as a subsystem in a molecule).

Obviously hamiltonian approach may be formulated for the classes (i)-(iv), which are idealized to some extent. While negligible feedback may be very realistic approximation of (iv), isolated systems and free particles models are applicable only when some aspects of dynamics are relevant (note that such ideal systems could not be even detected).

2.2 Postulate of Evolution

States of physical systems are changed during the time (in fact, not only the states, but also the system itself may be changed; such generalizations will not be studied here). Just these changes are the motion or the evolution of the system, and the task of the dynamics is to find the law of the motion. According to the experimental evidence and elementary intuition, there are some basic properties of the evolution of the quantum systems, which are comprised in the following

POSTULATE 5. – EVOLUTION

The evolution of the physical system during the time interval $t \in [t_\mathrm{i}, t_\mathrm{f}]$ is a process in which the initial state $|\psi, t_\mathrm{i}\rangle$ is transformed to the final state $|\psi, t_\mathrm{f}\rangle$ such that:

(i) the change is causal, meaning that to each initial state $|\psi, t_\mathrm{i}\rangle$ uniquely corresponds single final state $|\psi, t_\mathrm{f}\rangle$;

(ii) the superposition of the states is conserved, i.e. if $|\psi, t_\mathrm{i}\rangle = \alpha_1 |\psi_1, t_\mathrm{i}\rangle + \alpha_2 |\psi_2, t_\mathrm{i}\rangle$, then $|\psi, t_\mathrm{f}\rangle = \alpha_1 |\psi_1, t_\mathrm{f}\rangle + \alpha_2 |\psi_2, t_\mathrm{f}\rangle$;

(iii) the number of systems within ensemble is conserved;

(iv) the change is continual with respect to time.

The point out properties of the evolution suffice to describe it formally. Indeed, causality means that the evolution is a map $U : |\psi, t_\mathrm{i}\rangle \to |\psi, t_\mathrm{f}\rangle$, i.e. an operator in the state space $\mathcal{H}$. Then, next property essentially specifies that U is linear operator. To understand the third property, recall that the probability to get the value a of the observable A in the state at t_i is $v(a, A, |\psi, t_\mathrm{i}\rangle) = N_a^{(\mathrm{i})}/N = \langle\psi, t_\mathrm{i}|\, P_a(A)\, |\psi, t_\mathrm{i}\rangle$. Sum of the probabilities for all eigenvalues is 1, manifesting that $|\psi, t_\mathrm{i}\rangle$ is normalized. Assuming that the measurement is performed in t_f, the result $v(a, A, |\psi, t_\mathrm{f}\rangle) = N_a^{(\mathrm{f})}/N$, with unchanged N implies that again $\langle\psi, t_\mathrm{f} \,|\, \psi, t_\mathrm{f}\rangle = 1$. Consequently, the contents of this property is that evolution preserves norm of the state; in this form it is characteristics of non-relativistic quantum mechanics. It significantly restricts the operators suitable to describe the evolution: Wigner's theorem[1] then shows that U is unitary operator (as its linearity has been already established). Precisely, there is the family $U(t_\mathrm{i}, t_\mathrm{f})$ of unitary operators for each two instances $t_\mathrm{i} \leq t_\mathrm{f}$. The last continuity principle, is clearly included when the family is reparameterized in the form $U(t - t_\mathrm{i}, t_\mathrm{i})$, with the first argument explicitly dependent on the elapsed time $t - t_\mathrm{i}$ for $t \in [t_\mathrm{i}, t_\mathrm{f}]$. As it is usual in physics, the continuity is to be understood in the very broad sense, i.e. as the differentiability of some sufficient order.

Summarizing all the consequences of the postulate, we obtain the "integral form" of the law of motion:

$$|\psi, t\rangle = U(t - t_\mathrm{i}, t_\mathrm{i})\, |\psi, t_\mathrm{i}\rangle\,. \tag{2.1}$$

The evolution operator $U(t - t_\mathrm{i}, t_\mathrm{i})$ completely governs the dynamics of the system, and therefore it comprises all the relevant dynamical characteristics. In the classical mechanics, Hamiltonian is the variable defining the dynamics, and therefore it is expected that the evolution operator is related to Hamiltonian observable obtained by the quantization postulate. Indeed, this relation is easily obtained. At first, for $t = t_\mathrm{i}$ evolution is the identity map: $U(0, t_\mathrm{i}) = I$. Further, if an arbitrary instant t_1 in the interval $(t_\mathrm{i}, t_\mathrm{f})$ is singled out, the evolution may be seen as the composition of the evolution from t_i to t_1 and from t_1 to t_f: $U(t_\mathrm{f} - t_\mathrm{i}, t_\mathrm{i}) = U(t_\mathrm{f} - t_1, t_1)U(t_1 - t_\mathrm{i}, t_\mathrm{i})$. Finally, let us fix the initial time t_i and find the infinitesimal increment $\mathrm{d}U(t - t_\mathrm{i}, t_\mathrm{i})$ of U in $t = t_\mathrm{i}$. By definition this is the difference $\mathrm{d}U(0, t_\mathrm{i}) = U(\mathrm{d}t, t_\mathrm{i}) - U(0, t_\mathrm{i})$. On the other hand, as a differential, it is proportional to $\mathrm{d}t$; in this way the observable $H(t_\mathrm{i})$ is defined by $\mathrm{d}U(0, t_\mathrm{i}) = -\frac{\mathrm{i}}{\hbar}H(t_\mathrm{i})\,\mathrm{d}t$. Hence, $-\frac{\mathrm{i}}{\hbar}H(t_\mathrm{i})\,\mathrm{d}t = U(\mathrm{d}t, t_\mathrm{i}) - I$. The differential at some later instant $t > t_\mathrm{i}$ can be found with help of the composition rule of the evolutions from t_i to t and from t to $t + \mathrm{d}t$: $\mathrm{d}U(t - t_\mathrm{i}, t_\mathrm{i}) = U(\mathrm{d}t + t - t_\mathrm{i}, t_\mathrm{i}) - U(t - t_\mathrm{i}, t_\mathrm{i}) = (U(\mathrm{d}t, t) - I)U(t - t_\mathrm{i}, t_\mathrm{i})$. Using the derived expression for the differential of U at $t = t_\mathrm{i}$, one gets that the bracket equals $-\frac{\mathrm{i}}{\hbar}H(t)\,\mathrm{d}t$. This gives the first order differential equation for U and the initial condition:

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}U(t - t_0, t_0) = H(t)U(t - t_0, t_0), \quad U(0, t_0) = I. \tag{2.2}$$

Therefore, this is the equation for the evolution operator $U(t - t_0, t_0)$, in its differential form. Note that this two parameter family is defined through the one parameter family $H(t)$. Besides it has the dimension of energy, the hermitian operator function $H(t)$ determines the dynamics of the system; it should be expected that this is Hamiltonian observable, as it will be shown later on.

The integral form of (2.2) is obtained by formal integration. One gets $U(t - t_0, t_0) = I - \frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} H(t_1)U(t_1 - t_0, t_0)\,\mathrm{d}t_1$, and the iterative substitution on the right of this formal

---

[1] norm preserving operators are either unitary or antiunitary.

solution, gives the Dyson's series:

$$U(t-t_0,t_0)=\sum_{n=0}^{\infty}(-\frac{\mathrm{i}}{\hbar})^n\int_{t_0}^{t}\mathrm{d}t_1\cdots\int_{t_0}^{t_{n-1}}\mathrm{d}t_nH(t_1)\cdots H(t_n). \tag{2.3}$$

Note that the operators from the family $H(t)$ in various instants may be noncommutative, $[H(t_i),H(t_j)]\neq 0$. Therefore, in (2.3) appears the time ordered product, as obviously $t_0\leq t_n\leq\cdots\leq t_1\leq t$. Instead, one can use explicit time ordering T, which rearranges any product of the operators in the product ordered with the increased time:

$$T[A(t_{i_1})\cdots A(t_{i_n})]\stackrel{\text{def}}{=}A(t_1)\cdots A(t_n)\quad(t_1\geq\cdots\geq t_n). \tag{2.4}$$

In (2.3) the time ordering is implicitly given through the upper bound of the integrals over t_i. On the other side, if all the integrations in $t_n,\dots,t_1$ were performed from t_0 to t, the total integration domain would be the n-dimensional cube with the length $t-t_0$. The domain $t_1\leq\cdots\leq t_n$ occupies $\frac{1}{n!}$ of this cube; precisely, to any of $n!$ possible ordering of the instances $t_1,\dots,t_n$, i.e. to any permutation of these instances, corresponds the sub-domain of the same volume. Thus, the integration in (2.3) may be also performed over the cube, but with explicit time ordering, and in all the sub-domains the same value is obtained, altogether $n!$ times:

$$U(t-t_0,t_0)=\sum_{n=0}^{\infty}\frac{(-\frac{\mathrm{i}}{\hbar})^n}{n!}\int_{t_0}^{t}\mathrm{d}t_1\cdots\int_{t_0}^{t}\mathrm{d}t_nT[H(t_1)\cdots H(t_n)]. \tag{2.5}$$

In the case when the operators $H(t)$ and $H(t')$ commute for any instants t and t', the time ordering may be omitted; each of the integrals becomes independent, and the last equation is easily integrated to

$$U(t-t_0,t_0)=\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t}H(t')\,\mathrm{d}t'}. \tag{2.6}$$

For the time independent Hamiltonian $H(t)=H$, being an example of this type, we get:

$$U(t-t_0,t_0)=\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H(t-t_0)}. \tag{2.7}$$

2.3 Schrödinger equation

To get explicitly the law of motion of the states, one combine the equations (2.1) and (2.2) to get the famous time dependent Schrödinger equation:

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\psi,t\right\rangle=H(t)\left|\psi,t\right\rangle,\quad\left|\psi,t\right\rangle_{t=t_0}=\left|\psi,t_0\right\rangle. \tag{2.8}$$

This is the first order in time differential equation, with the initial condition given on the right. Its coordinate representation (assuming position only dependent potential):

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\psi(\boldsymbol{r},t)=H(\boldsymbol{r},\boldsymbol{p},t)\psi(\boldsymbol{r},t)=\left(-\tfrac{\hbar^2}{2m}\triangle+V(\boldsymbol{r})\right)\psi(\boldsymbol{r},t),\quad\psi(\boldsymbol{r},t)_{t=t_0}=\psi(\boldsymbol{r},t_0) \tag{2.9}$$

is known as the time dependent wave equation.

2.3.1 Conservative systems

In a conservative system the Hamiltonian is constant with respect to time, and the evolution is described by (2.7). When it acts on a hamiltonian eigen vector $|E,t_0\rangle$ (for the eigen value E), only the time dependent phase factor appears:

$$|E,t\rangle = U(t-t_0)\,|E,t_0\rangle = \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}\,|E,t_0\rangle\,.$$

This means that the eigenvectors of the Hamiltonian are stationary states in the physical sense, i.e. such a physical state is not changed during the time. On the other side, for any stationary state, i.e. state such that $|\psi,t\rangle = f(t)\,|\psi,t_0\rangle$, the Schrödinger equation reads $(\mathrm{i}\hbar\dot{f}/f)\,|\psi,t_0\rangle = H\,|\psi,t_0\rangle$; since the right side is time independent, the bracket must be a constant, meaning that $|\psi,t_0\rangle$ (and therefore $|\psi,t\rangle$) is an eigen state of the Hamiltonian.

This observation suffices to reduce the evolution problem of the conservative systems to the time independent Hamiltonian eigen problem. Indeed, let $|E,\lambda\rangle$ be an eigen basis of H:

$$H\,|E,\lambda\rangle = E\,|E,\lambda\rangle\,. \tag{2.10}$$

If system in t_0 is described by the vector $|E,\lambda\rangle$, in the later instant t the state vector is $\mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}\,|E,\lambda\rangle$. Then, a quite general state can be expanded over the Hamiltonian eigen basis at any instant t in the form: $|\psi,t\rangle = \sum_{E,\lambda} c_{E\lambda}(t)\,|E,\lambda\rangle$. Note that the time dependence is completely comprised in the expansion coefficients. On the other hand, since the evolution preserves the superposition (by the postulate),

$$|\psi,t\rangle = U(t-t_0,t_0)\,|\psi,t_0\rangle = \sum_{E,\lambda} c_{E\lambda}(t_0)U(t-t_0,t_0)\,|E,\lambda\rangle = \sum_{E,\lambda} c_{E\lambda}(t_0)\mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}\,|E,\lambda\rangle\,,$$

giving that the evolution is defined by

$$c_{E\lambda}(t) = c_{E\lambda}(t_0)\mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}. \tag{2.11}$$

Hence, solving the eigen problem of the Hamiltonian, the evolution is completely determined. In other words, the solutions of (2.8) are completely determined (through (2.11)) by the solutions of (2.10), which is therefore called *time independent Scrödinger equation*.

These conclusions are important for the characterization of the states. In fact, if H is degenerate observable, then using any CSCO $\{H,A,B,\dots\}$ containing H and the observables commuting with H, one gets the stationary basis $|E,a,b,\dots\rangle$, which is called complete classification of the states. Here, the eigenvalues a, b etc. of the observables A, B etc. take the role of the degeneracy counter λ from (2.10). Such states have sharp values of these observables. Consequently, these values are conserved during the evolution, i.e. they are integrals of motion, also called good quantum numbers.

2.3.2 Evolution of the mixed states

The evolution of the mixed states is easily derived from the evolution of the pure ones. In fact, let at t_0 a mixed state is given in its most general form $\rho(t_0) = \sum_i \omega_i\,|\psi_i,t_0\rangle\,\langle\psi_i,t_0|$. Afterwards mixture evolves together with its pure components $|\psi_i,t\rangle$, meaning:

$$\rho(t) = U(t-t_0,t_0)\rho(t_0)U^{\dagger}(t-t_0,t_0). \tag{2.12}$$

The differential form of this equation is obtained as derivative in t of (2.12), with the substitution of (2.2) and its adjoint. This analogue of the Schrödinger equation for mixed states is called Liouville equation:

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\rho(t)=[H(t),\rho(t)],\quad \rho(t)_{t=t_0}=\rho(t_0). \tag{2.13}$$

Exercise 2.1: Prove Hellman-Feynman Theorem: for the hamiltonian parameterized by a real parameter a the normalized eigenstates ($H(a)\left|E(a)\right\rangle = E(a)\left|E(a)\right\rangle$) satisfy

$$\frac{\mathrm{d}}{\mathrm{d}a}E(a)=\left\langle E(a)\right|\frac{\mathrm{d}}{\mathrm{d}a}H(a)\left|E(a)\right\rangle. \tag{2.14}$$

Exercise 2.2: Prove that for any observable A, matrix elements of the commutator $[H,A]$ in the (normalized) stationary basis satisfy

$$\left\langle n,\lambda\right|[H,A]\left|n',\lambda'\right\rangle=(E_n-E_{n'})\left\langle n,\lambda\right|A\left|n',\lambda'\right\rangle. \tag{2.15}$$

In particular, taking the hamiltonian in the form $H=T+U(\boldsymbol{r})$, find the matrix elements of the momentum and force (gradient of the potential) proving virial theorem.

2.3.3 Evolution of the probabilities and means

The evolution of the system causes the change of the observable quantities[2]: as the measurement results (i.e. distribution of probabilities) evolve, this is accompanied with change of the derived quantities, like mean values. Even for the observables being constant in time, the measurement results may vary due to the evolution of the states.

Note that an observable $A(t)$ may be explicit function of time (e.g. if it is related to some variable external field). Then its mean in the state $\rho(t)$ during the time is given as $\langle A\rangle_t=\operatorname{Tr}\rho(t)A(t)$. The derivative in t with help of (2.2) yields $\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_t=\operatorname{Tr}[H(t),\rho(t)]A(t)+\mathrm{i}\hbar\operatorname{Tr}\rho(t)\frac{\mathrm{d}}{\mathrm{d}t}A(t)$. Then, expanding commutator, this becomes:

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_t=\langle[A(t),H(t)]\rangle_t+\mathrm{i}\hbar\left\langle\frac{\partial}{\partial t}A\right\rangle_t \tag{2.16}$$

The derivative of a classical variable in t is based on the notion of its value in the close instants. In the quantum sense, the observable itself may have no particular value, and obtains it in the measurement process. Therefore the derivatives must refer not only to the parametric dependence of A of time, but to the both quantities, observable and state, participating in the measurement process. In this sense, since (2.16) is fulfilled for each state, it can be used to define the time derivative of the observable, and get the classical like equation of motion:

$$\left\langle\frac{\mathrm{d}}{\mathrm{d}t}A\right\rangle_t\overset{\mathrm{def}}{=}\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_t,\quad \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}A=[A(t),H(t)]+\mathrm{i}\hbar\frac{\partial}{\partial t}A. \tag{2.17}$$

---

[2] For this subsection recommended literature is &9 of Ref. [11].

Also, it is straightforward to find how the probability of a particular measurement result depends on time:

$$v_t(a, A, |\psi\rangle) = \| P_a(A) U(t - t_0, t_0) |\psi\rangle \|^2.$$

Analogously, given a state $|\phi\rangle$ fixed as a common eigenvector of a CSCO, the transition probability from arbitrary state $|\psi\rangle$ to $|\phi\rangle$ is

$$v_t(|\psi\rangle \to |\phi\rangle) = | \langle\phi| U(t - t_0, t_0) |\psi\rangle |^2.$$

2.3.4 Uncertainty relation time-energy

Time and energy are related to some extent analogously to coordinate and momenta. Therefore, an uncertainty relation involving these quantities may be expected. Still, since there is no observable of time, such relation must be derived indirectly.

We say that A is appropriate observable for the state $|\psi\rangle$ if it is not explicitly time dependent and $\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_{|\psi\rangle} \neq 0$. For such an observable and state, the minimal time $\tau_A = \Delta_{|\psi\rangle}(A)/|\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_{|\psi\rangle}|$ must pass until the change of mean overwhelms the dispersion $\Delta_{|\psi\rangle}(A)$.

Theorem 2.1 (Time-energy uncertainty relation) *Any non-stationary state $|\psi, t\rangle$ and any appropriate observable A satisfy:*

$$\tau_A \Delta_{|\psi\rangle}(H) \geq \frac{1}{2}\hbar. \tag{2.18}$$

▬*Proof:* According to (1.25), $\Delta_\psi(A)\Delta_\psi(H) \geq \frac{1}{2}|\langle[A, H]\rangle_\psi| = \frac{1}{2}|\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_{|\psi\rangle}| = \frac{\hbar}{2}|\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_{|\psi\rangle}|$, and the relation looked for follows when $\Delta(A)$ is expressed in terms of τ_A. ▮

Since this inequality holds for all the appropriate observables, also the infimum $\tau = \inf\{\tau_A \mid A \text{ appropriate}\}$ satisfies $\tau\Delta_{|\psi\rangle}(H) \geq \frac{1}{2}\hbar$. Hence, τ is the best estimate independent of the particular observable.

In the scattering experiments, the target energy is measured by means of beams of particles. The latter are free particles, with hamiltonian $H = p^2/2m$, well collimated with the dispersion $\Delta p \neq 0$. The beam is in average localized at $\langle x\rangle$ with the dispersion Δx, and the mean value evolves. Hence x is appropriate observable, and the measurement has duration of at least τ_x.

Completely different approach to the time-energy uncertainty relation appears on the context of unstable states. In fact, when the quantum system, like atom, is excited from the ground state, the interaction with the surrounding results usually in the deexcitation. Therefore the excited states have their lifetime, and $\tau_{1/2}$ is the half lifetime of such a state. On the other hand, the excited state energy cannot be exactly found in the experiments, but with some error ΔE. The phenomenological uncertainty relation $\tau_{1/2}\Delta E \geq \hbar$ is experimentally found:

2.3.5 Probability current and continuity equation

Assuming that the one particle Hamiltonian has the form $H = \boldsymbol{p}^2/2m + V(\boldsymbol{r}, t)$, the Schrödinger equation in the coordinate representation and its complex conjugate are:

$$\begin{aligned} \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\psi(\boldsymbol{r}, t) &= -\frac{\hbar^2}{2m}\triangle\psi(\boldsymbol{r}, t) + V(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t), \\ -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\psi^*(\boldsymbol{r}, t) &= -\frac{\hbar^2}{2m}\triangle\psi^*(\boldsymbol{r}, t) + V(\boldsymbol{r}, t)\psi^*(\boldsymbol{r}, t). \end{aligned}$$

Multiplying the first of them by $\psi^*(\boldsymbol{r}, t)$ and the second one by $\psi(\boldsymbol{r}, t)$, and subtracting them, one gets the continuity equation:

$$\frac{\partial\rho}{\partial t} + \mathrm{div}\boldsymbol{j} = 0. \tag{2.19}$$

Here, $\rho(\boldsymbol{r}) = \psi^*(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t) = \tilde{v}(\boldsymbol{r}, \hat{\boldsymbol{r}}, |\psi\rangle)$ is the density of the probability (1.5) of the particle occurrence in r, while

$$\boldsymbol{j} = \frac{\hbar}{2m\mathrm{i}}(\psi^*(\boldsymbol{r}, t)\nabla\psi(\boldsymbol{r}, t) - \psi(\boldsymbol{r}, t)\nabla\psi^*(\boldsymbol{r}, t)) = \frac{1}{m}\mathrm{Re}(\psi^*(\boldsymbol{r}, t)\frac{\hbar}{\mathrm{i}}\nabla\psi(\boldsymbol{r}, t)). \tag{2.20}$$

is the probability current.

Exercise 2.3: ° Show that the probability currents in the stationary states ψ of the free particle and hydrogen atom (Subsection 4.4.2, (4.40) and (4.27)) are:

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = C\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} : \quad \boldsymbol{j}_{\boldsymbol{k}} = \frac{\hbar}{m}\boldsymbol{k}|C|^2; \tag{2.21}$$

$$\psi_{nlm}(\boldsymbol{r}) = R_{nl}(r)Y^l_m(\theta, \varphi) : \quad j_r = j_\theta = 0,\ j_\varphi = \frac{\hbar m}{m_e r\sin\theta}|\psi_{nlm}(\boldsymbol{r})|^2. \tag{2.22}$$

2.4 Simple systems

2.4.1 One-dimensional particle

There is a number of general properties of the solutions of Schrödinger's equation for one-dimensional particles. Here only few of them will be mentioned, mostly in the form of exercises. There are many nice analysis on the Web and literature ([11] is one of the best and condensed as usual).

For finite potential both the function and its derivative are continuous (even in the points where potential is discontinuous).

Unbound states

Exercise 2.4: Show that the stationary states of free particle are plane waves, $\Psi(x, t) = c\mathrm{e}^{-\mathrm{i}\frac{E(k)}{\hbar}t}\mathrm{e}^{\mathrm{i}kx}$, corresponding to eigenenergies $E(k) = \hbar^2k^2/2m$ for each real k, which are also eigenstates for momentum with eigenvalues $p = \hbar k$.

Exercise 2.5: Find the stationary states and eigenenergies for the potential $V(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ U & \text{if } x > 0 \end{cases}$. Show that for finite U there is nonzero probability for the particle to be at position $x > 0$, even for the states with energy less that U.

Bound states

Bound state is a state localized within a finite spatial region, with exponential decrease of the probability density outside it. Bound states correspond to the discrete part of spectrum of Hamiltonian. Arbitrary potential $U(x)$, with $\lim_{x\to\pm\infty} U(x) = 0$, hase at least one bound state if there is a region where it is negative.

Exercise 2.6: Bound states of one-dimensional systems are not degenerate.

Exercise 2.7: (Oscillation theorem) Wave function of the n-the excited bound states has n zeros.

Exercise 2.8: Bound states of one-dimensional system with real potential can be represented by a real wave function. Bound states of one-dimensional system with even potential can be represented by a wave function which is odd or even in inversion.

2.4.2 Harmonic oscillator

The hamiltonian of the one-dimensional harmonic oscillator is

$$H = \frac{1}{2}\frac{p^2}{m} + \frac{1}{2}m\omega^2 q^2. \tag{2.23}$$

Introducing new observables $Q = q\sqrt{m\omega/\hbar}$ and $P = p/\sqrt{m\omega\hbar} = -\mathrm{i}\partial_Q$ (note that P and Q are not canonically conjugated, since $[Q, P] = \mathrm{i}$), one gets

$$H = \hbar\omega H', \quad H' = \frac{1}{2}(P^2 + Q^2). \tag{2.24}$$

Exercise 2.9: (Linear Harmonic Oscillator) Find eigenfunctions and energy spectrum of the Hamiltonian $H = \frac{p_x^2}{2m} + \frac{m\omega^2}{2}x^2$.

Besides direct solution (Exercise 2.9) of the appearing hypergeometric equation, the eigenproblem of the Hamiltonian may be instantly solved noticing that H' is symmetric in P and Q. Therefore it commutes with Fourier-Plancherel operator, and shares its eigenbasis of Hermite functions. Note that the same argument is applicable to other operators essentially symmetric in coordinates and momenta.

Still, conceptually most important method of solving eigenproblem of H is Dirac's one. It introduces the operators

$$a^{\pm} = \frac{1}{\sqrt{2}}(Q \mp \mathrm{i}P) = \frac{1}{\sqrt{2}}(Q \mp \frac{\mathrm{d}}{\mathrm{d}Q}), \tag{2.25}$$

with the following significant, though obvious, properties

$$(a^-)^\dagger = a^+, \quad [a^-, a^+] = 1. \tag{2.26}$$

To this pair of operators the third one, occupancy number operator is accompanied:

$$n = a^+a^-. \tag{2.27}$$

It is manifestly positive, and, since $n\,|x\rangle = 0$ iff $0 = \langle x|\,n\,|x\rangle = \|a^-\,|x\rangle\,\|$, its kernel is $\ker(n) = \ker(a^-)$. The following commutator relation is easily verified:

$$[n, a^{\pm}] = \pm a^{\pm}. \tag{2.28}$$

Exercise 2.10: *Normal (Wick's) order* monomial in creation and annihilation operators is when annihilation operators are after the creation operators. Show that any monomial over a^+ and a^- may be expressed as a polynomial with normal ordered terms.

Exercise 2.11: Show that $[(a^\pm)^s, a^\mp] = \mp s(a^\pm)^{s-1}$, and

$$[(a^-)^m, (a^+)^p] = \sum_{s=1}^{\min\{m,p\}} \frac{p!}{(p-s)!}\frac{m!}{(m-s)!}(a^+)^{p-s}(a^-)^{m-s}.$$

The introduced triple of the operators a^-, a^+ and n is very important in the quantum theory of many particles. At first we note that $a^\pm$ are by (2.25) biuniquely related to the fundamental set of observables $\{x, p\}$, making possible to express all the observables in terms of these two. In addition, any observable may be expanded in the series over a^- and a^+ with the terms being monomials of the type $(a^+)^p(a^-)^m$ (as shown in Exercise 2.10). Particularly important are spectral properties of n.

Theorem 2.2 *The spectrum of occupancy number operator is purely discrete* $\sigma(n) = \{0, 1, 2, \dots\}$ *and nondegenerate, with all the eigenvectors* $|i\rangle$ *determined by the vacuum state* $|0\rangle$ *as*

$$|i\rangle = \frac{1}{\sqrt{i!}}(a^+)^i\,|0\rangle\,. \tag{2.29}$$

▬*Proof:* As a positive operator, n must have an eigenvalue $x > 0$. Let $|x\rangle$ be the corresponding eigenvector: $n\,|x\rangle = x\,|x\rangle$. The commutator relation (2.28) gives $n(a^\pm\,|x\rangle) = (a^\pm n \pm a^\pm)\,|x\rangle = (x \pm 1)(a^\pm\,|x\rangle)$. Obviously, together with x the eigenvalues are also $x-1$ (for $a^-\,|x\rangle$), $x-2$ (for $(a^-)^2\,|x\rangle$), etc. This series would become negative at some point. The only way to avoid such contradiction is that x is integer itself: then one eigenvector is $(a^-)^x\,|x\rangle \sim |0\rangle$. The vacuum vector $|0\rangle$ satisfies $n\,|0\rangle = 0$, meaning that it is from the kernel of n, implying $0 = \langle 0|\,n\,|0\rangle = \langle 0|\,a^+a^-\,|0\rangle = \|a^-\,|0\rangle\,\|^2$. Thus, $a^-\,|0\rangle = 0$, and for $j > x$ the negative part of the spectrum is not generated as $(a^-)^j\,|j\rangle = 0$. The series is then cut here, since all the remaining members vanish, restricting the spectrum of n to nonnegative integers. Assume that this spectrum is not simple; the degeneracy can be removed by a compatible observable. As q and p are functions of a^+ and a, any observable has a form $A = \sum_{pm} a_{pm}(a^+)^p(a^-)^m$. The compatibility with n imposes condition (Exercise 2.12)

$$0 = [A, n] = \sum_{pm} a_{pm}[(a^+)^p(a^-)^m, a^+a^-] = \sum_{pm}(m-p)a_{pm}(a^+)^p(a^-)^m,$$

requiring $a_{pm} = 0$ unless $m = p$. But then $A = \sum_s b_s n^s$, i.e. it is function of n and cannot remove the degeneracy. Hence, n is complete observable. Finally, it is clear that $|i\rangle = c_i(a^+)^i\,|0\rangle$, and only the normalizing coefficient in (2.29) is to be found. Assuming that the basis $\{|i\rangle\}$ is orthonormal, the squared norm of $a^+\,|i\rangle = c\,|i+1\rangle$ is $\langle i|\,a^-a^+\,|i\rangle = i+1$, i.e. $a^+\,|i\rangle = \sqrt{i+1}\,|i+1\rangle$, verifying the proposed result.
▮

Exercise 2.12: Show $[(a^+)^p(a^-)^m, a^+a^-] = (m-p)(a^+)^p(a^-)^m$.

Now it is easy to solve the eigen problem of the harmonic oscillator. Namely, since obviously $H' = n + \frac{1}{2}$, i.e.

$$H = \hbar\omega(n + \frac{1}{2}) \tag{2.30}$$

the eigen vectors of H are exactly the same vectors (2.29) as for n, while the spectrum of H is shifted upward for one half, and then multiplied by $\hbar\omega$. Thus the spectrum consist of equidistant eigenvalues, the difference between successive ones being $\hbar\omega$. The remarkable features of this spectrum allow us to consider that the eigen state $|i\rangle$ describes a system with i free (noninteracting) quasi-particles: each of them contributes by the single quantum $\hbar\omega$. Operator a^+, shifting from $|i\rangle$ to $|i+1\rangle$, acts as the creation operator of these quasi particles, while its adjoint a^- is the annihilation operator. In this context, the eigen value i of n is the number of the quanta in the state $|i\rangle$, and n is called occupation number operator. Usually the annihilation operator a^- is denoted by a, while the creation one a^+ is according to the first property simply its adjoint $a^\dagger$.

To get the eigenfunctions in the Q-representation $\psi_i(Q) = \langle Q\,|\,i\rangle$, we note that the vacuum state is singled out by $a^-\,|0\rangle = 0$, which is by (2.25) the differential equation $\frac{\mathrm{d}\psi}{\mathrm{d}Q} + Q\psi = 0$, with the solution $\psi_0(Q) = \pi^{-\frac{1}{4}}\mathrm{e}^{-\frac{1}{2}Q^2}$. Then applying (2.29) one recognizes Hermite function $\psi_i(Q)$ of the order i:

$$\psi_i(Q) = \frac{1}{\sqrt{\sqrt{\pi}2^i i!}}\mathrm{e}^{-\frac{1}{2}Q^2}H_i(Q). \tag{2.31}$$

Exercise 2.13: Prove (2.31). Hint: show the operator identity: $(Q - \frac{\mathrm{d}}{\mathrm{d}Q}) = (-\mathrm{e}^{\frac{1}{2}Q^2}\frac{\mathrm{d}}{\mathrm{d}Q}\mathrm{e}^{-\frac{1}{2}Q^2})$.

Exercise 2.14: Generalize the above consideration to the system with d harmonic degrees of freedom; its hamiltonian is $H = \sum_{i=1}^{d}(\frac{1}{2}\frac{p_i^2}{m_i} + \frac{1}{2}m_i\omega_i^2 q_i^2)$.

Exercise 2.15: Find the mixed state describing harmonic oscillator in the thermodynamic equilibrium with a heat reservoir at temperature T. Discuss the mean energy in the limits of zero and very high temperature.

2.4.3 Charged particle in a homogeneous magnetic field

For a particle with charge q in electromagnetic field described by the vector and scalar potentials $\boldsymbol{A}$ and V, Hamiltonian is

$$H(t) = \frac{1}{2m}\left[\left(\hat{\boldsymbol{p}} - \frac{q}{c}\vec{A}(\vec{r},t)\right)^2 + qV(\vec{r},t)\right]. \tag{2.32}$$

In the considered case, scalar potential vanishes while the vector one is static, and for conventionally taken magnetic field of intensity B along z-axis, in the Landaus's gauge it is $\boldsymbol{A} = -By\boldsymbol{e}_x$. This leads to the time independent Schrödinger equation (coordinate representation):

$$\frac{1}{2m}\left[(\hat{p}_x + \frac{q}{c}By)^2 + \hat{p}_y^2 + \hat{p}_z^2\right]\psi(x,y,z) = E\psi(x,y,z) = E\mathrm{e}^{\mathrm{i}(xp_x+zp_z)/\hbar}Y(y). \tag{2.33}$$

the last expression manifests that potential is independent of x and z, making total solution factorized to the solution $Y(y)$ along y and plane waves for free motion in the remaining directions. As formulated in the first part of the Lemma on the separation of variables (Section **??**), for each pair of eigenvalues p_x and p_z, of the the corresponding momenta there is their common subspace, where these observables reduce to scalars p_x and p_z. It remains to solve eigenequation of the Hamiltonian H_y of the motion along y parameterized by p_x:

$$H_y = \left[\frac{1}{2m}\hat{p}_y^2 + \frac{1}{2}m\omega^2(y + \frac{cp_x}{qB})^2\right] Y(y) = EY(y). \tag{2.34}$$

Clearly, this describes harmonic oscillator along y, around center at $-\frac{cp_x}{qB}$ (determined by momentum along x) and field dependent frequency (known as *cyclotron frequency*) $\omega = \frac{qB}{mc}$. While the eigenfunction is Hermite function, the harmonic oscillator eigenvalues $E = (n + \frac{1}{2})\hbar\omega$ are in this situation known as *Landau's levels*.

2.5 Pictures

The unitary transformations preserve the complete structure of the quantum mechanics. Mathematically, this refers to the invariance of linear combination and norm of vectors, as well as of the operator product and commutator. These leads to the physically obvious independence of the measurement results on the choice of the basis in the state space. Nevertheless, a particular choice of the basis may be more convenient either for conceptual or practical reasons. Here we will consider three such possibilities, called pictures, related to the dynamics of the system, which proved to be important for various purposes.

In fact, the first one has been introduced in the derivation of the dynamical law. Throughout the section 2.3 it has been tacitly assumed that the states are evolving, while the observables remain unchanged at least for the conservative systems. This picture is used to define the other ones. Indeed, any dynamical picture will be defined by an unitary operator function $T(t)$ over time; in such a picture state and observable are in the instant t given by their Scrödinger values ($|\psi, t; \mathrm{S}\rangle$ and $A_\mathrm{S}(t)$) and the picture operator $T(t)$:

$$|\psi, t; T\rangle = T(t)\,|\psi, t; \mathrm{S}\rangle\,, \quad A_T(t) = T(t)A_\mathrm{S}(t)T^{-1}(t). \tag{2.35}$$

Hence, the Schrödinger picture is obtained with $T(t) = I$, with the evolution law for states given by (2.1) or (2.8), while the observables are considered time independent as far as the system evolution is regarded, although an explicit time dependence $A(t)$ may be imposed by the surrounding (e.g. the apparatus is moving).

2.5.1 Heisenberg's picture

Within this evolution picture the picture operator is inverse to the evolution: $T(t) = U^{-1}(t)$. The equations of motion

$$\begin{aligned} |\psi, t; \mathrm{H}\rangle &= |\psi, 0; \mathrm{S}\rangle = |\psi, 0; \mathrm{H}\rangle\,, & \text{(2.36a)}\\ A_\mathrm{H}(t) &= U^{-1}(t)A_\mathrm{S}(t)U(t), \quad \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}A_\mathrm{H}(t) = [A_\mathrm{H}(t), H(t)] + \mathrm{i}\hbar\frac{\partial}{\partial t}A_\mathrm{H}(t) & \text{(2.36b)}\end{aligned}$$

are easily derived from the Scrödinger's one. Note that the states are time independent, while the observables evolve. The observables are changed for two different reasons. The evolution of the system, incorporated through the picture itself (by the definition on the right of (2.36b)) results in the first (commutator) term in the differential form at the right of (2.36b). Possible explicit (external influence) changes of the measuring apparatus gives the last contribution in (2.36b), where one should be warned that the partial time derivative is defined as $\frac{\partial}{\partial t}A_{\mathrm{H}}(t) \overset{\mathrm{def}}{=} U^{-1}(t)\left[\frac{\partial}{\partial t}A_{\mathrm{S}}(t)\right]U(t)$.

2.5.2 Dirac's picture

This picture, also called the interaction picture, gives suitable point of view for the physical problems with Hamiltonian containing two parts, $H = H_0 + H'$; in fact, usually it is assumed that H' is small with respect to H_0, but for purpose of introducing Dirac picture this is not important. Then the picture is characterized by the operator $T(t) = U_0^{-1}(t)$, where U_0 is defined as the evolution realized by the first part of the Hamiltonian, i.e. satisfying (2.2) for H_0; defining the remaining evolution by $U = U_0U'$, from (2.2) one gets:

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}U_0(t-t_0,t_0) = H_0(t)U_0(t-t_0,t_0), \quad U_0(0,t_0) = I, \tag{2.37}$$

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}U'(t-t_0,t_0) = H'_D(t)U'(t-t_0,t_0), \quad U'(0,t_0) = I. \tag{2.38}$$

It must be understood that although these equations apparently decouple the dynamics related to H_0 and H', the second one, referring to H' implicitly involves H_0 through the Dirac picture of H'. Both the evolutions of the states and observables are nontrivial; from (2.35) follows

$$|\psi,t;\mathrm{D}\rangle = U'(t)\,|\psi,0;\mathrm{D}\rangle\,, \quad A_{\mathrm{D}}(t) = U_0^{-1}(t)A_{\mathrm{S}}(t)U_0(t). \tag{2.39}$$

The differential form of these equations is easily found:

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\psi,t;\mathrm{D}\rangle = H'_{\mathrm{D}}(t)\,|\psi,t;\mathrm{D}\rangle\,, \tag{2.40}$$

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}A_{\mathrm{D}}(t) = [A_{\mathrm{D}}(t),H_{0\mathrm{D}}] + \mathrm{i}\hbar\frac{\partial}{\partial t}A_{\mathrm{D}}(t), \tag{2.41}$$

with $\frac{\partial}{\partial t}A_{\mathrm{D}}(t) \overset{\mathrm{def}}{=} U_0^{-1}(t)\left[\frac{\partial}{\partial t}A_{\mathrm{S}}(t)\right]U_0(t)$.

2.6 Propagator and Green function

Retarded propagator is two instance (t_2, t_1) function derived from evolution operator. It vanishes for $t_2 < t_1$, while otherwise it is matrix element of the evolution operator in the coordinate representation:

$$\begin{aligned} K(\boldsymbol{r}_2,t_2;\boldsymbol{r}_1,t_1) &= \langle\boldsymbol{r}_2,t_2|\,U(t_2-t_1,t_1)\,|\boldsymbol{r}_12,t_1\rangle\,\theta(t_2-t_1) \\ &= \theta(t_2-t_1)\sum_{n,\lambda}\phi^*_{n\lambda}(\boldsymbol{r}_1)\mathrm{e}^{-\mathrm{i}E_n(t_2-t_1)}\phi_{n\lambda}(\boldsymbol{r}_2) \end{aligned} \tag{2.42}$$

Using eigenbasis of the Hamiltonian, $\phi_n(\boldsymbol{r}_2) = \langle \boldsymbol{r}_2 \,|\, E_n \rangle$, satisfying Schrödinger equation $\left(\mathrm{i}\hbar \frac{\partial}{\partial t_2} - H(\boldsymbol{r}_2, \boldsymbol{p}_2) \right) \mathrm{e}^{-\mathrm{i}E_n(t_2 - t_1)} \phi_n(\boldsymbol{r}_2) = 0$, one gets (step function of time is a constant for Hamiltonian operator, its derivative over t_2 is $\delta(t_2 - t_1)$, and $\sum_{n,\lambda} \phi^*_{n\lambda}(\boldsymbol{r}_1)\phi_{n\lambda}(\boldsymbol{r}_2) = \delta(\boldsymbol{r}_2 - \boldsymbol{r}_1)$):

$$\left(\mathrm{i}\hbar \frac{\partial}{\partial t_2} - H(\boldsymbol{r}_2, \boldsymbol{p}_2) \right) K(\boldsymbol{r}_2, t_2; \boldsymbol{r}_1, t_1) = \mathrm{i}\hbar \delta(t_2 - t_1) \delta(\boldsymbol{r}_2 - \boldsymbol{r}_1). \tag{2.43}$$

So, it turns out that propagator is Green function of the Schrödinger equation.

The importance of the (retarded) propagator stems from the fact that knowing wave function at instance t_1, in the later instance t_2 the evolved state $|\psi, t_2\rangle = U(t_2 - t_1, t_1)$ is in the coordinate representation obtained as

$$\psi(\boldsymbol{r}_2, t_2) = \int \mathrm{d}\, r_1^3 \, \langle \boldsymbol{r}_2 | \, U(t_2 - t_1, t_1) \, | \boldsymbol{r}_1 \rangle \, \psi(\boldsymbol{r}_1, t_1) = \int \mathrm{d}\, r_1^3 K(\boldsymbol{r}_2, t_2; \boldsymbol{r}_1, t_1) \psi(\boldsymbol{r}_1, t_1). \tag{2.44}$$

This resembles classical Huygens' principle, where each point of the wavefront becomes a secondary wave source. Note further that the first part of the above is valid even when $t_2 < t_1$; in that case, one can introduce advanced propagator $\langle \boldsymbol{r}_2 | \, U(t_2 - t_1, t_1) \, | \boldsymbol{r}_1 \rangle \, \theta(t_1 - t_2)$.

The propagator techniques are well developed, and essentially invoke another, Feynman's path integral interpretation of quantum mechanics, introduced in the next section.

2.7 Feynman's path integral approach

Feynman showed that quantum dynamics may be understood in a conceptually simple way[3], giving direct insight to the difference between classical and quantum treatment; in addition it offers a shortcut to many important techniques of field theory (both in solid state and particle physics).

The Feynman's approach is based on the expression for the probability amplitude to detect a particle, initially being (in $t_0 = 0$) at x_0, at $x = x_N$ in the later instance $t = t_N$: $\langle q_N, t_N \,|\, q_0, 0 \rangle$. The time interval $[0, t]$ may be divided in N equal parts $\Delta t = t_N / N$, and the amplitude becomes:

$$\langle q_N \,|\, q_0, 0 \rangle = \langle q_N | \underbrace{U(\Delta t) \cdots U(\Delta t)}_{N} | q_0 \rangle \, .$$

Namely, assuming that the system is conservative the evolution operator $U(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar} H t}$ is written as $U(\Delta t)^N$. Further for each $t_i = i \Delta t$ we introduce identity in the form $\int \, \mathrm{d}q_i \, |q_i\rangle \, \langle q_i|$, giving

$$\begin{aligned} \langle q_N \,|\, q_0 \rangle &= \int \mathrm{d}q_1 \cdots \mathrm{d}q_N \, \langle q_N | \, U(\Delta t) \, | q_{N-1} \rangle \cdots \langle q_i + 1 | \, U(\Delta t) \, | q_i \rangle \cdots \langle q_1 | \, U(\Delta t) \, | q_0, 0 \rangle \\ &= \int \mathrm{d}q_1 \cdots \mathrm{d}q_N K(N, N-1) \cdots K(1, 0) \end{aligned} \tag{2.45}$$

---

[3] On the other hand, this view introduced new mathematical techniques, particularly the Wienner measures, which are still developing. This resembles to the Dirac's δ-function, and its influence to functional analysis.

(he $K(i,i-1)$ is short for the propagator $K(q_i,t_i;q_{i-1},t_{i-1})$).

Further we assume that the potential is coordinate dependent only, i.e. $H=\frac{\hat{p}^2}{2m}+V(\hat{q})$, giving $\langle q_i+1|\,U(\Delta t)\,|q_i\rangle = \langle q_i+1|\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(\frac{\hat{p}^2}{2m}+V(\hat{q}))\Delta t}\,|q_i\rangle$. Note that $[\hat{p}\Delta t, V\Delta t]\sim V'(\hat{q})\Delta^2 t$, and in the limit $\Delta t=0$ (i.e. for infinite N) this comutator vanishes, meaning that $\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(\frac{\hat{p}^2}{2m}+V(\hat{q}))\Delta t}\to \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{\hat{p}^2}{2m}}\mathrm{e}^{\frac{\mathrm{i}}{\hbar}V(\hat{q})\Delta t}$. Thus, in the same limit the considered amplitude becomes: $\langle q_i+1|\,U(\Delta t)\,|q_i\rangle = \langle q_i+1|\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\frac{\hat{p}^2}{2m}}\,|q_i\rangle\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}V(\hat{q})\Delta t}$. Introducing the identity in the form of the momentum eigen decomposition, $I=\int \mathrm{d}p\,|p\rangle\,\langle p|$, with plane waves $|p\rangle = \frac{1}{\sqrt{2\pi}}\int \mathrm{d}q\mathrm{e}^{\frac{\mathrm{i}}{\hbar}pq}\,|q\rangle$, one gets $\langle q_i+1|\,U(\Delta t)\,|q_i\rangle = \frac{1}{2\pi}\int \mathrm{d}p\mathrm{e}^{\frac{\mathrm{i}}{\hbar}(p\frac{q_{i+1}-q_i}{\Delta t}-\frac{p^2}{2m})\Delta t}\mathrm{e}^{\frac{\mathrm{i}}{\hbar}V(q_i)\Delta t}$. The last integral may be calculated[4], giving $\langle q_i+1|\,U(\Delta t)\,|q_i\rangle = \sqrt{\frac{m\hbar\mathrm{e}^{-\mathrm{i}\pi/2}}{2\pi\Delta t}}\mathrm{e}^{\frac{\mathrm{i}}{\hbar}[\frac{m}{2}(\frac{q_{i+1}-q_i}{\Delta t})^2-V(q_i)]\Delta t}$. It is easy to recognize action $\int_0^{\Delta t}\mathrm{d}tL(q,\dot{q})$ in the exponent, where the classical lagrangian is $L(q,\dot{q})=\frac{m}{2}\dot{q}^2-V(q)$ (recall that Δt is infinitesimal). Altogether, making the product, the looked for amplitude becomes:

$$\langle q_N\,|\,q_0,0\rangle = \int \mathcal{D}q\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\int_0^t L(q,\dot{q})\,\mathrm{d}t}, \quad \mathcal{D}q=\lim_{N\to\infty}\left(\frac{mN\hbar}{2\pi\mathrm{i}t}\right)^{N/2}\prod_{i=1}^{N-1}\mathrm{d}q_i. \tag{2.46}$$

The integral is the sum of the contributions of all possible trajectories. The infinitesimal volume $\mathcal{D}q$ is up to the phase equal to the mathematically well founded Wienner measure in the space of trajectories. The discrepancy may be formally cured by so called Wick's rotation, making time imaginary (i.e. if $\mathrm{i}t$ is used instead of t). Anyway, as the probabilities are observable quantities, the phase is not relevant. Therefore, (2.46) is the path or functional integral formulation of quantum dynamics. It reveals that the probability amplitude of transition from q_0 to q_N is sum of contributions over all possible paths connecting these two positions. Each path $q_0\to q_N$ is weighted by its amplitude equal to $\mathrm{e}^{\frac{\mathrm{i}}{\hbar}S(q_0\to q_N)}$.

In the classical mechanics only the extremal path is realized. On the other hand, in the quantum mechanics all the paths are realized, but with different weights corresponding to the action along paths. As the action is in the exponent, the classical path corresponds to the stationary, i.e. minimal phase. The weights of the other paths are dumped, as the action is increased, with more emphasized oscillatory character of the integrand. There is also quasi classical limit (of quantum mechanics), where the paths with the action $\int_0^t L(q,\dot{q})\,\mathrm{d}t$ much greater than $\hbar$ are taken into account only (this corresponds to retaining of the leading terms in the expansion of the exponent in (2.46)).

Expansion of the action functional around the classical path (being the minimum of the functional) must be quadratic in the parameters introduced in the space of trajectories. Therefore, the leading terms in such a perturbative approach (the method of stationary phase) are Gaussian integrals. This makes the formalism not only conceptually important, but also efficient in various problems, despite the fact that only Gaussian (besides some trivial ones) path integrals are exactly calculated.

---

[4] It is well known that $\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}(\alpha p+\beta p^2)}\,\mathrm{d}p=\sqrt{\frac{\pi}{\beta}}\mathrm{e}^{\mathrm{i}(\frac{\pi}{4}-\frac{\alpha^2}{4\beta})}$; it is assumed that β is positive, while in our case it is imaginary. This will be discussed after (2.46).

Chapter 3

Galilean Transformations

The space-time symmetries are important in many aspects of system dynamics. Their most significant manifestations are conservation laws (in the isolated systems); namely, these laws are the evidence of symmetries. When the conserved quantities are generators of the Lie-algebra of the symmetry group, they explicitly refer only to the connected component of the group. The other components give rise to conserved parities.

Non-relativistic and relativistic mechanics consider different geometry of the space-time. This is manifested through different set of space-time symmetries, i.e. the transformations leaving space-time unchanged, usually called relativity principle: Galilean one, relevant for non-relativistic mechanics, assumes space-time invariance with respect to Galilean group, in contrast to the Poincare group related to the Einstein relativity.

3.1 Galilean Group

Galilean group $\mathcal{G}$ gathers transformations $g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau)$ acting on the space-time point $(\boldsymbol{r},t)$ by:

$$g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau)(\boldsymbol{r},t)=(R_{\boldsymbol{\phi}}\boldsymbol{r}+\boldsymbol{a}-\boldsymbol{v}t,t+\tau). \tag{3.1}$$

Here, $\boldsymbol{\phi}$ is the 3-dimensional real vector defining the rotation $R_{\boldsymbol{\phi}}$ around $\boldsymbol{\phi}$ for the angle ϕ equal to the length of $\boldsymbol{\phi}$. Also $\boldsymbol{a}$ and $\boldsymbol{v}$ are 3-dimensional real vectors, specifying the translation and boost, while τ is a real scalar, defining the time translation. The action on the momentum $\boldsymbol{p}=m\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}$ is derived directly:

$$g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau)\boldsymbol{p}=R_{\boldsymbol{\phi}}\boldsymbol{p}-m\boldsymbol{v}. \tag{3.2}$$

The structural properties of $\mathcal{G}$ are to be examined first, and then more each of these transformations will be discussed in more detail. The consecutive application of two group elements gives the multiplication law:

$$g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau)g(\boldsymbol{\phi}',\boldsymbol{a}',\boldsymbol{v}',\tau')=g(\boldsymbol{\phi}'',\boldsymbol{a}+R_{\boldsymbol{\phi}}\boldsymbol{a}'-\boldsymbol{v}\tau',R_{\boldsymbol{\phi}}\boldsymbol{v}'+\boldsymbol{v},\tau+\tau'), \tag{3.3}$$

where $\boldsymbol{\phi}''$ is defined by $R_{\boldsymbol{\phi}''}=R_{\boldsymbol{\phi}}R_{\boldsymbol{\phi}'}$. Also, the requirement $g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau)(\boldsymbol{r},t)=(\boldsymbol{r},t)$ gives the identity element, and then, putting it on the right of the (3.3), one solves in the primed parameters to get the inverse of $g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau)$:

$$e = g(0,0,0,0), \qquad g^{-1}(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau) = g(-\boldsymbol{\phi}, -R_{\boldsymbol{\phi}}^{-1}(\boldsymbol{a}+\tau\boldsymbol{v}), -R_{\boldsymbol{\phi}}^{-1}\boldsymbol{v}, -\tau). \tag{3.4}$$

Galilean group is connected ten dimensional Lie group. It is easy to show that when all but one arguments are set to 0, one gets the following subgroups, which are themselves significant in further considerations:

$$\mathrm{SO}(3,\mathbb{R}) = \{g(\boldsymbol{\phi},0,0,0)\}, \quad \text{Rotational group;} \tag{3.5a}$$
$$\boldsymbol{T}_3^{\mathrm{S}} = \{g(0,\boldsymbol{a},0,0)\}, \quad \text{Spatial translational group;} \tag{3.5b}$$
$$\boldsymbol{T}_3^{\mathrm{B}} = \{g(0,0,\boldsymbol{v},0)\}, \quad \text{Boost group;} \tag{3.5c}$$
$$\boldsymbol{T}_1^{\mathrm{T}} = \{g(0,0,0,\tau)\}, \quad \text{Time translational group.} \tag{3.5d}$$

The first two subgroups are built of pure spatial transformations, and together they form the Euclidean group $\mathrm{E}(3) = \boldsymbol{T}_3^{\mathrm{S}} \wedge \mathrm{SO}(3,\mathbb{R})$ (semi-direct product, with translations as the invariant subgroup). The last subgroup consists of the time translations, i.e. the transformations connecting systems with different origin of the time scale, is purely related to time. Both time and spatial coordinates involves only the third subgroup of the boosts, the transformations interrelating the instantly coinciding coordinate systems moving uniformly each with respect to another.

Besides these transformations which are continual functions of the parameters, there are also discrete transformations of the spatial, time and total reversal, $\mathcal{I}_{\mathrm{S}}$, $\mathcal{I}_{\mathrm{T}}$ and $\mathcal{I} = \mathcal{I}_{\mathrm{S}}\mathcal{I}_{\mathrm{T}}$:

$$\mathcal{I}_{\mathrm{S}}(\boldsymbol{r},t) = (-\boldsymbol{r},t), \quad \mathcal{I}_{\mathrm{T}}(\boldsymbol{r},t) = (\boldsymbol{r},-t), \quad \mathcal{I}(\boldsymbol{r},t) = (-\boldsymbol{r},-t). \tag{3.6}$$

Together with the continual transformations, they form extended Galilean group $\bar{\mathcal{G}}$, with $\mathcal{G}$ being its invariant subgroup (as the connected component of the identity) and three additional cosets: $\bar{\mathcal{G}} = \mathcal{G} + \mathcal{I}_{\mathrm{S}}\mathcal{G} + \mathcal{I}_{\mathrm{T}}\mathcal{G} + \mathcal{I}\mathcal{G}$. It is important to note that there is no possibility to verify directly the space-time invariance with respect to these transformations, and therefore they are not a priori included in the relativity principle.

3.2 Galilean transformations of the classical variables

The Galilean transformations are defined by the action in the space-time, and afterward the action on the momenta is derived. This action can be extended to the variables, i.e. functions on the phase space. In the space of variables Galilean group is realized by the canonical transformations, being themselves functions of the variables. This is important in the view of the forthcoming quantization procedure. To begin with, we introduce Poisson bracket action of one variable onto another: to each variable G corresponds the supervariable $\tilde{G}$, which is the linear operator (in the space of variables) acting on the arbitrary variable A as

$$\tilde{G}A \overset{\mathrm{def}}{=} [G,A]_{\mathrm{PZ}} \tag{3.7}$$

Consequently, $\tilde{G}^n A = [G,[G,\ldots[G,A]_{\mathrm{PZ}}\ldots]_{\mathrm{PZ}}]_{\mathrm{PZ}}$. Therefore, to each analytic function $F(G)$ of G the supervariable $\tilde{F}(G)$ may be associated using Taylor expansion of F. In particular, the exponential function

$$\mathrm{e}^{\gamma\tilde{G}} \overset{\text{def}}{=} \sum_{n=0}^{\infty} \frac{\gamma^n}{n!}\tilde{G}^n \tag{3.8}$$

defines an one-parameter group of the transformations associated to G. Therefore, $\tilde{G}$ is the Lie-algebra generator of this group.

The Galilean transformations are in the variable space represented by the following supervariables:

$$\tilde{D}(g(\boldsymbol{\phi},0,0,0)) = \mathrm{e}^{-\boldsymbol{\phi}\cdot\tilde{\boldsymbol{l}}}, \qquad \tilde{D}(g(0,\boldsymbol{a},0,0)) = \mathrm{e}^{-\boldsymbol{a}\cdot\tilde{\boldsymbol{p}}}, \tag{3.9a}$$

$$\tilde{D}(g(0,0,\boldsymbol{v},0)) = \mathrm{e}^{m\boldsymbol{v}\cdot\tilde{\boldsymbol{q}}}, \qquad \tilde{D}(g(0,0,0,\tau)) = \mathrm{e}^{-\tau\tilde{H}}. \tag{3.9b}$$

This is verified by the action on the fundamental variables, i.e. in the phase space:

$$\mathrm{e}^{-\boldsymbol{\phi}\cdot\tilde{\boldsymbol{l}}}\boldsymbol{r} = R_{\boldsymbol{\phi}}\boldsymbol{r}, \qquad \mathrm{e}^{-\boldsymbol{\phi}\cdot\tilde{\boldsymbol{l}}}\boldsymbol{p} = R_{\boldsymbol{\phi}}\boldsymbol{p}; \tag{3.10a}$$

$$\mathrm{e}^{-\boldsymbol{a}\cdot\tilde{\boldsymbol{p}}}\boldsymbol{r} = \boldsymbol{r} + \boldsymbol{a}, \qquad \mathrm{e}^{-\boldsymbol{a}\cdot\tilde{\boldsymbol{p}}}\boldsymbol{p} = \boldsymbol{p}; \tag{3.10b}$$

$$\mathrm{e}^{m\boldsymbol{v}\cdot\tilde{\boldsymbol{q}}}\boldsymbol{r} = \boldsymbol{r}, \qquad \mathrm{e}^{m\boldsymbol{v}\cdot\tilde{\boldsymbol{q}}}\boldsymbol{p} = \boldsymbol{p} - m\boldsymbol{v}; \tag{3.10c}$$

$$\mathrm{e}^{-\tau\tilde{H}}\boldsymbol{r} = \boldsymbol{r} + \frac{\tau}{m}\boldsymbol{p}, \qquad \mathrm{e}^{-\tau\tilde{H}}\boldsymbol{p} = \boldsymbol{p}. \tag{3.10d}$$

In the last equations the hamiltonian of the free particle $H = \boldsymbol{p}^2/2m$ is assumed, since the probe particle in the empty space-time is considered. Therefore, the Lie algebra generators are angular momentum $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$ for rotations, linear momentum $\boldsymbol{p}$ for translations, coordinates $\boldsymbol{q}$ for boosts and Hamiltonian for time translations. Note that all the transformations are canonical. Clearly, (3.10) also defines the action on any other classical variable $A(\boldsymbol{r},\boldsymbol{p})$: on the functions of the fundamental variables action is defined by (3.7), which gives:

$$\tilde{D}(g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau))A(\boldsymbol{r},\boldsymbol{p}) = A(\tilde{D}(g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau))\boldsymbol{r}, \tilde{D}(g(\boldsymbol{\phi},\boldsymbol{a},\boldsymbol{v},\tau))\boldsymbol{p}). \tag{3.11}$$

***Exercise 3.1**: Prove (3.9) in the coordinate form.*

However, although spatial translations and boosts commute, the supervariables corresponding to them do not. Therefore, the obtained set of supervariables is not a representation of $\mathcal{G}$, despite the fact that each subgroup is correctly represented. In fact, there is no faithful representation of the Galilean group. Since the representations faithfully relating transformations and supervariables are necessary, only the Euclidean subgroup and time translations can be simultaneously considered. While the pure geometrical invariance results in the conservation laws of the generators (angular and linear momentum), the time translations are manifested as the conservation of energy. On the other side, the boost invariance leads to the low of uniform motion of the centre of mass of the isolated system (since the kinetic energy is not the same in the boosted systems, the action is also changed, and the Nöther theorem is not applicable [12]).

3.3 Quantization: Wigner theorem

Being classical space-time symmetry, Galilean group is symmetry group of a single (non-relativistic probe) particle, or of a centre of mass of any isolated system. Therefore, observable characteristics of such a system must be same before and after Galilean transformation, i.e. such characteristics are independent on the chosen inertial reference frame.

In the context of quantum mechanics, the observable quantities are various probabilities, which can be reduced to squares of absolute value of scalar products (recall (1.4)). In other words, if Galilean transformation T transform states $|\phi\rangle$ and $|\psi\rangle$ into $|\phi'\rangle$ and $|\psi'\rangle$, then the space-time symmetry requires that

$$|\langle\phi\,|\,\psi\rangle| = |\langle\phi'\,|\,\psi'\rangle| = |\langle\phi|\,T^{\dagger}T\,|\psi\rangle|.$$

Consequently, T is in the state space represented by an operator $D(T)$ preserving absolute value of the scalar products. The famous Wigner theorem shows that such operators are either unitary or antiunitary.

Note that without absolute value in the above relation, only unitary operators would be allowed. In addition, due the absolute value together with $D(T)$ any operator $\mathrm{e}^{\mathrm{i}\chi_T}D(T)$ can be used equally well. This reflects the projective structure of the state space, i.e. that the collinear vectors describe the same state. One can try to use the freedom in the choice of the phases $\mathrm{e}^{\mathrm{i}\chi_T}$ to get the set of the unitary and antiunitary operators forming a representation of the Galilean group. However, this attempt fails, as well as in the classical physics, because Galilean group has no faithful representation. Still, all the important subgroups may be independently quantizied, and Euclidean group can be combined with time translations.

Although Wigner theorem does not specify which operators are represented by unitary or antiunitary operators, it is obvious by the continuity argument: the connected component of the identical element must be represented unitary. On the other hand, product of two antiunitary operators is unitary, meaning that either the whole group or its halving subgroup is unitary represented. Thus, at most two cosets are antiunitary.

Really, the complex conjugated Schrödinger equation (for a particle in the time independent potential) corresponds to the time reversed one with conjugated vectors, meaning that the time reversal is antiunitary in the state space. As the spatial inversion is obviously linear, it follows that purely spatial halving subgroup is linear, while time reversal and total inversion give the cosets having antilinear representations.

3.4 Quantization of the Galilean group

Since the action of the Galilean transformation in the variable space is well defined in terms of Poisson bracket and variables, the postulate of quantization may be directly applied to get the Galilean action in the space of the operators (superspace). Analogously to the supervariables, we define the superoperators with help of the commutator. So, to each operator $\hat{G}$ in the state space $\mathcal{H}$, its superoperator $\hat{\hat{G}}$ is defined by its action on the arbitrary operator $\hat{A}$:

$$\hat{\hat{G}}\hat{A} \overset{\mathrm{def}}{=} [\hat{G}, \hat{A}]. \tag{3.12}$$

Due to the appearance of the term $\mathrm{i}\hbar$ in the canonical quantization, the exponential function $\mathrm{e}^{\gamma\tilde{G}}$ is quantized as $\mathrm{e}^{\frac{\gamma}{\mathrm{i}\hbar}\hat{\hat{G}}}$. Indeed, $\mathrm{e}^{\frac{\gamma}{\mathrm{i}\hbar}\hat{\hat{G}}}\hat{A} = \sum_{n=0}^{\infty}\frac{\gamma^n}{n!}\frac{1}{\mathrm{i}\hbar}[\hat{G},\frac{1}{\mathrm{i}\hbar}[\hat{G},\dots,\frac{1}{\mathrm{i}\hbar}[\hat{G},\hat{A}]\dots]]$, which is obviously the quantized exponential supervariable action.

Exercise 3.2: Show the following operator identities:

$$\mathrm{e}^{B}A\mathrm{e}^{-B} = A + [B,A] + \tfrac{1}{2!}[B,[B,A]] + \tfrac{1}{3!}[B,[B,[B,A]]] + \dots;$$
$$\mathrm{e}^{A}\mathrm{e}^{B} = \mathrm{e}^{A+B}\mathrm{e}^{[A,B]/2} \qquad \text{for } [A,[A,B]] = 0 = [[A,B],B].$$

Therefore, the superoperator representation of the Galilean group is directly obtained from (3.9):

$$\hat{\hat{D}}(g(\boldsymbol{\phi},0,0,0)) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\boldsymbol{\phi}\cdot\hat{\hat{\boldsymbol{l}}}}, \qquad \hat{\hat{D}}(g(0,\boldsymbol{a},0,0)) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\boldsymbol{a}\cdot\hat{\hat{\boldsymbol{p}}}}, \tag{3.13a}$$
$$\hat{\hat{D}}(g(0,0,\boldsymbol{v},0)) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}m\boldsymbol{v}\cdot\hat{\hat{\boldsymbol{q}}}}, \qquad \hat{\hat{D}}(g(0,0,0,\tau)) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\tau\hat{\hat{H}}}. \tag{3.13b}$$

These superoperators are bijective image of the Galilean group, but the same problem as for supervariables appears: representatives of the boosts and translations do not commute as the transformations do. Again, the Euclidean group and the time translations only can be simultaneously quantized.

The superoperator representation is not really appropriate for the usual tasks of quantum mechanics, and the operator representation is looked for to realize the group action in the state space. The Baker-Hausdorff lemma solves this task. It asserts that there is bijection between the exponential superoperators and exponential operators, i.e. to each exponential action $\mathrm{e}^{\hat{\hat{G}}}$ in the operator space uniquely corresponds exponential action $\mathrm{e}^{\hat{G}}$ in the state space, such that superoperator essentially acts as the similarity transformation:

$$\mathrm{e}^{\hat{\hat{G}}}\hat{A} = \mathrm{e}^{\hat{G}}\hat{A}\mathrm{e}^{-\hat{G}}. \tag{3.14}$$

Therefore in the state space, the Galilean transformations are represented by:

$$\hat{D}(g(\boldsymbol{\phi},0,0,0)) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\boldsymbol{\phi}\cdot\hat{\boldsymbol{l}}}, \qquad \hat{D}(g(0,\boldsymbol{a},0,0)) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\boldsymbol{a}\cdot\hat{\boldsymbol{p}}}, \tag{3.15a}$$
$$\hat{D}(g(0,0,\boldsymbol{v},0)) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}m\boldsymbol{v}\cdot\hat{\boldsymbol{q}}}, \qquad \hat{D}(g(0,0,0,\tau)) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\tau\hat{H}}. \tag{3.15b}$$

Note that the result for the translations is known from the chapter 1. Indeed, the operators $\hat{D}(g(0,\boldsymbol{a},0,0))$ representing translations are direct three-dimensional generalization of the translational operator $U(a)$ used in the Theorem 1.2.

Exercise 3.3: Find $\mathrm{e}^{\mathrm{i}\pi\mathcal{I}}\psi(\boldsymbol{r})$, where $\mathcal{I}$ is spatial inversion, i.e. $\mathcal{I}\psi(\boldsymbol{r}) = \psi(-\boldsymbol{r})$.

Exercise 3.4: Find the projectors $P_{\pm}$ on the subspaces spanned by even (P_+) and odd (P_-) functions according to the spatial inversion.

Exercise 3.5: Find the coefficients α and β for which the operator $U_{\alpha\beta} = e^{\alpha x^2+\beta p^2}$ becomes spatial inversion, i.e. rotation for π in the phase space of one-dimensional particle.

3.5 Active and passive interpretations

The Galilean transformations are defined by their action on the space-time points. In this sense, they actively map one point to another, and therefore this interpretation is called active interpretation. On the other hand, there is passive interpretation, describing the same act as the inverse transformation of the referent frame, with no action on the space-time points. Thus, to each active transformation T on the space time-vectors represented in the reference frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_t\}$ by the column $(\boldsymbol{r}, t)$, corresponds the inverse transformation of the frame:

$$\text{Active}: T\begin{pmatrix}\boldsymbol{r}\\ t\end{pmatrix} \to \begin{pmatrix}\boldsymbol{r}'\\ t'\end{pmatrix}; \quad \text{Passive}: T^{-1}\boldsymbol{e}_i \to \boldsymbol{e}'_i, \tag{3.16}$$

such that in the transformed frame the initial point is described as $(\boldsymbol{r}', t')$. Obviously, some of the transformations cannot be really realized actively, e.g. time translations, spatial and time inversion; even more, boosts are defined with help of the reference systems. Due to the inversion of the transformation when one comes from active to the passive interpretation, they are anti-isomorphic. Acting on the variables, being the functions on the phase space, the supervariables (3.9) are essentially passive interpretation, and therefore the sign minus appears in the exponents.

Finally, from the action $\hat{D}(g)\,|\boldsymbol{q}\rangle = |g\boldsymbol{q}\rangle$, it follows that $\langle\boldsymbol{q}|\,\hat{D}(g)\,|\psi\rangle = \langle g^{-1}\boldsymbol{q}\,|\,\psi\rangle$. Thus, coordinate representation of the Galilean transformations is

$$\hat{D}(g)\psi(\boldsymbol{q}) = \psi(g^{-1}\boldsymbol{q}). \tag{3.17}$$

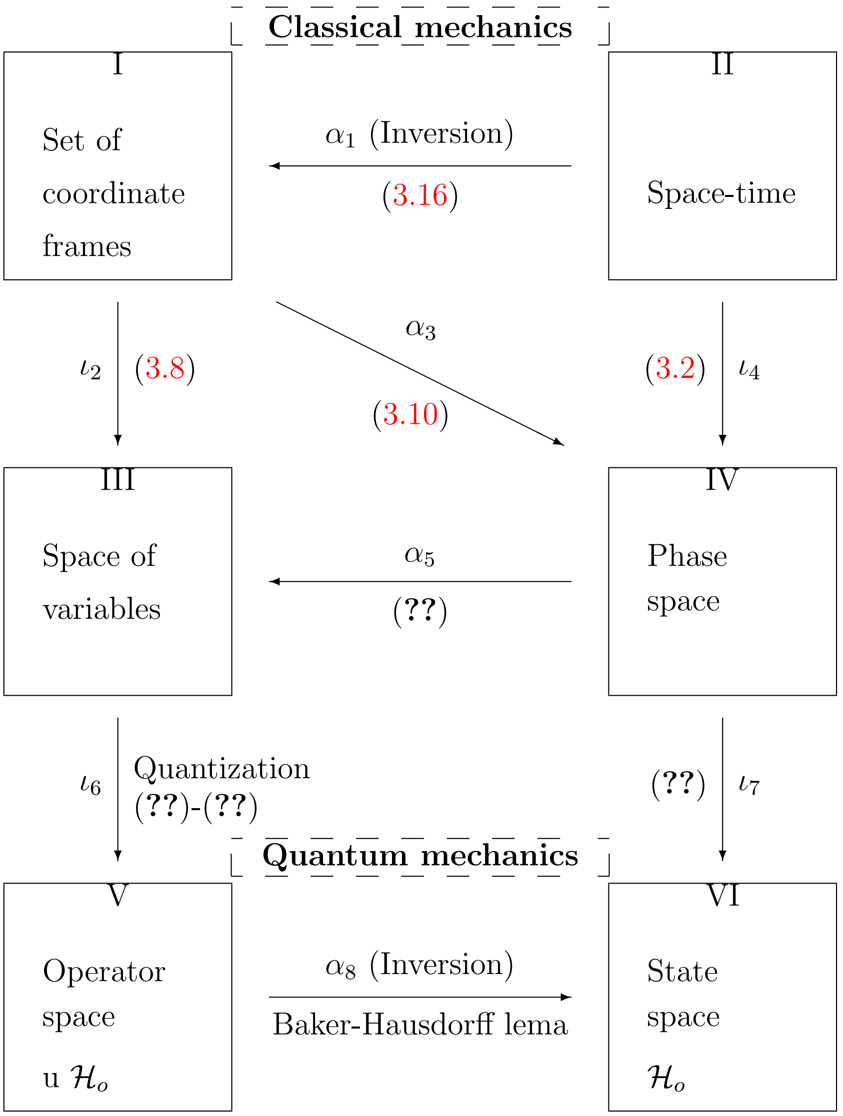

Figure 3.1: Active and passive interpretions of transformation. Left column (kvadrati I, III i V i izomorfizmi ι_2 i ι_6) daje elemente za pasivnu, a desna kolona (kvadrati II, IV i VI i izomorfizmi ι_4 i ι_7) za aktivnu interpretaciju grupa $E(3)$ i $T_3^{(v)}$. Antiizomorfizmi α_1, α_3, α_5 i α_8, kao i izomorfizmi ι_2, ι_4, ι_6 i ι_7 povezuju osnovne podgrupe Galilejeve grupe koje deluju u skupovima označenim sa I do VI. Specijalno: $\iota_2 \overset{\text{def}}{=} \alpha_5 \circ \alpha_3 = \alpha_5 \circ \iota_4 \circ \alpha_1^{-1}$; $\alpha_3 = \iota_4 \circ \iota_1$ je prevoenje vienja izmeu dva opservera; ι_4 je reprezentovanje u fiksiranom koordinatnom sistemu sa implikacijama u impulsnom prostoru.

Chapter 4

Rotations and Angular Momentum

4.1 Elementary Properties of Rotations

As explained in the Chapter 3, rotations are parameterized by the vectors $\boldsymbol{\phi}$, specifying the angle ϕ and the axis (by the unit vector $\hat{\boldsymbol{\phi}} = \boldsymbol{\phi}/\phi$) of each rotation. The defining action of the rotations in the Euclidean space $\mathbb{R}^3$ is given by the special orthogonal matrices, and therefore the rotational group is identified with the group SO(3,$\mathbb{R}$).

The obvious equality $R_{\boldsymbol{\phi}} = R_{(2\pi-\phi)(-\hat{\boldsymbol{\phi}})}$ of the rotations around opposite axes shows that all the rotations are obtained taking all the unit vectors and the angles not greater than π, i.e. that the range of the vectors $\boldsymbol{\phi}$ (manifold of the parameters) is π-ball, a ball of radius π. However, the opposite points of the same diameter of this ball correspond to the same rotation: $R_{\pi\boldsymbol{u}} = R_{-\pi\boldsymbol{u}}$ for arbitrary unit vector $\boldsymbol{u}$. Hence, the rotational group is compact (parameter space is bounded and closed), but double connected, as any diameter of the π-ball is a closed path not homotopic to any closed path in the π-ball's interior.

Further, the first relation in (3.15), $\hat{D}(g(\boldsymbol{\phi},0,0,0) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\boldsymbol{\phi}\cdot\hat{\boldsymbol{l}}}$, shows that rotations are represented in the orbital state space (of a single three-dimensional particle) by the exponents of the operators of the orbital angular momentum. Therefore the components of the angular momentum form the Lie algebra of the rotational group. Using the definition $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$, and the canonical coordinates-momenta commutation relations $[q_i, p_j] = \mathrm{i}\hbar\delta_{ij}$, the commutation relations of the angular momenta are easily derived:

$$[l_i, l_j] = \mathrm{i}\hbar \sum_k \epsilon_{ijk} l_k \tag{4.1}$$

(ϵ is Levi–Civita tensor). This shows that the Lie algebra is so(3,$\mathbb{R}$) (of course, this follows immediately from the general theory), the set of the traceless hermitian operators[1]. Now, several general (and elementary) results of the group theory can be exploited to get additional insight into the basic properties of the rotations and angular momenta.

---

[1] Mathematical definition of the exponential relation between Lie group and its algebra (real!) does not include the imaginary unit, and therefore unitary representations of the group are obtained from the skew hermitian representation of the algebra. This factor enables that the Lie algebra elements in physics are hermitian (physical observables), but non-closeness of the commutator in the space of hermitian operators must be cured by another factor $\mathrm{i}\hbar$ in the commutators, as in (4.1). This may be also understood as a change of basis in the complexified Lie algebra.

First, compactness implies that rotational group is represented by unitary operators, while the angular momenta are hermitian (precisely, equivalent to the unitary and hermitian operators, or alternatively, that there is a scalar product — selected in physics — such that the representative operators are unitary and hermitian, respectively). As a consequence, each representation is decomposable to the irreducible ones, the latter being necessarily finite dimensional.

The fact that the angular momenta form three-dimensional simple Lie algebra so(3,ℝ), enables to construct irreducible representation according to the well elaborated algorithm for the semisimple algebras (Section 4.2.1). Additionally, its rank is one, which reveals that there is a single Cartan element and a single Kazimir operator: conventionally, these are the third component, l_3, and the square, $l^2 = l_1^2 + l_2^2 + l_3^2$, of the angular momentum). Therefore, irreducible representations are classified by a single parameter, different eigenvalues of the Kazimir operator, while a basis within arbitrary irreducible space is counted by the different eigenvalues of the Cartan element as another parameter.

Doubly connected parameter space implies that there is a double (universal) covering, locally isomorphic to SO(3,ℝ), i.e. a simply connected Lie group $\widetilde{\mathrm{SO}(3,\mathbb{R})}$, with two-to-one homomorphism onto SO(3,ℝ), but with the same Lie algebra. As it will be shown later on, this is SU(2) group, with algebra su(2)=so(3,ℝ). Accordingly, the obtained representations of the algebra of the angular momenta do not really produce (as exponentials) the representations of the rotational but of the covering group. Only half of them represent rotations, while the rest are the "double valued" representations, associating a pair of the operators to each rotation; the characteristic feature of the double valued representation, differing them from the ordinary ones, is that rotations for 2π reverse the vectors (of the representative space), thus only rotations for 4π are associated to the identity operator. Ultimate consequence is that the very notion of rotations should be reconsidered: without any physical content specified, the general concept of rotations is defined by the commutation relations (4.1). This essentially introduces the covering group $\widetilde{\mathrm{SO}(3,\mathbb{R})}$ as the one underlying the angular momenta, rather then SO(3,ℝ). The fermionic spin degrees of freedom (Section 4.5), which can be explained only by the double-valued representations, proves that this is not just a mathematical extrapolationartefact, but physical reality.

4.2 Algebra of angular momentum

To stress out that from now on the notion of the angular momentum is as general as possible, free of any specific physical situation, the angular momenta are defined as a triple of the hermitian operators K_i $(i = 1, 2, 3)$ satisfying the commutation relations:

$$[K_i, K_j] = \mathrm{i}\hbar \sum_k \epsilon_{ijk} K_k. \tag{4.2}$$

In this way only the Lie algebra properties are extracted from the initial (prototypic) example, the orbital angular momentum. Indeed, only this is really necessary for the following construction of the representation (neatly following the prescription developed for the semisimple complex Lie algebras).

4.2.1 Irreducible representations of angular momentum

Construction of the irreducible representations starts by choice of a suitable basis (in mathematical language, this corresponds to the standard form of the complexified algebra). Mutually adjoint complex linear combinations

$$K_{\pm} = K_1 \pm \mathrm{i} K_2, \tag{4.3}$$

of K_1 and K_2, together with K_3 form standard components of angular momentum. They span the same space as the three momenta, since (4.3) is solvable in the Cartesian components: $K_1 = \frac{1}{2}(K_+ + K_-)$ and $K_2 = \frac{1}{2\mathrm{i}}(K_+ - K_-)$. Therefore, space is irreducible for the Cartesian components if and only if it is irreducible for the standard ones.

Lemma 4.1 *The standard components satisfy the following commutation relations:*

$$[K_3, K_{\pm}] = \pm\hbar K_{\pm}; \tag{4.4a}$$
$$[K_+, K_-] = 2\hbar K_3; \tag{4.4b}$$
$$[K_3, K_+K_-] = [K_3, K_-K_+] = [K_+K_-, K_-K_+] = 0. \tag{4.4c}$$

▬*Proof:* (a) and (b) are obtained applying (4.3) and then (4.2).
(c) $[K_3, K_\pm K_\mp] = [K_3, K_\pm]K_\mp + K_\pm[K_3, K_\mp] = \pm\hbar K_\pm K_\mp \mp \hbar K_\pm K_\mp = 0$;
$[K_+K_-, K_-K_+] = K_-[K_+K_-, K_+] + [K_+K_-, K_-]K_+ =$
$K_-K_+[K_-, K_+] + K_-[K_+, K_+]K_- + K_+[K_-, K_-]K_+ + [K_+, K_-]K_-K_+ =$
$K_-K_+(-2\hbar K_3) + 2\hbar K_3 K_- K_+ = 0.$ ∎

The operators $K_\pm K_\mp$ are hermitian (an positive), and together with K_3 make a set of compatible observables (as justified by (4.4c)). Thus, they have a common orthonormal eigenbasis. Number m is used to denote eigenvalues $m\hbar$ of K_3 (the units of angular momentum and action are same). In the possibly degenerate eigensubspace $\mathcal{V}_m$ of K_3, a common (for the whole triple) orthonormal eigenbasis is denoted by $|m\lambda\rangle$ ($\lambda = 1, \dots, |\mathcal{V}_m|$):

$$K_3 |m\lambda\rangle = m\hbar |m\lambda\rangle, \quad K_\pm K_\mp |m\lambda\rangle = \alpha_m^\pm(\lambda) |m\lambda\rangle, \tag{4.5}$$

with $\alpha_m^\pm(\lambda)$ being the eigenvalues of $K_\pm K_\mp$.

Lemma 4.2 *For each $|m\lambda\rangle$ the vectors $K_\pm |m\lambda\rangle$ are mutually orthogonal for different λ, and satisfy*

$$K_3(K_\pm |m\lambda\rangle) = (m \pm 1)\hbar(K_\pm |m\lambda\rangle), \tag{4.6a}$$
$$K_\pm K_\mp(K_\pm |m\lambda\rangle) = \alpha_m^\mp(\lambda)(K_\pm |m\lambda\rangle), \tag{4.6b}$$
$$K_\pm K_\mp(K_\mp |m\lambda\rangle) = (\alpha_m^\pm(\lambda) \pm 2\hbar^2(m \mp 1))(K_\mp |m\lambda\rangle). \tag{4.6c}$$

▬*Proof:* All the relations follows from (4.5). To show orthogonality and (4.6b) one regroups the terms:

$$(\langle m\lambda| K_\pm^\dagger)(K_\pm |m\lambda'\rangle) = \langle m\lambda| (K_\mp K_\pm |m\lambda'\rangle) = \alpha_m^\mp(\lambda) \langle m\lambda \,|\, m\lambda'\rangle = \alpha_m^\mp(\lambda)\delta_{\lambda\lambda'};$$

$$(K_\pm K_\mp)(K_\pm |m\lambda\rangle) = K_\pm(K_\mp K_\pm) |m\lambda\rangle = \alpha_m^\mp(\lambda)(K_\pm |m\lambda\rangle).$$

For (4.6a) and (4.6c) it suffices to apply (4.4a) and (4.4c), respectively

$$K_3(K_\pm |m\lambda\rangle) = ([K_3, K_\pm] + K_\pm K_3) |m\lambda\rangle = (\pm 1 + m)\hbar(K_\pm |m\lambda\rangle);$$

$$\begin{aligned} K_\pm K_\mp(K_\mp|m\lambda\rangle) &= ([K_\pm,K_\mp]+K_\mp K_\pm)(K_\mp|m\lambda\rangle) = \\ &= (\pm 2\hbar K_3 + K_\mp K_\pm)(K_\mp|m\lambda\rangle) = (\pm 2\hbar^2(m\mp 1)+\alpha_m^\pm(\lambda))(K_\mp|m\lambda\rangle). \end{aligned}$$

(in the last equation (4.6a) is used). ▮

The lema enlightens that vectors of the basis $|m\lambda\rangle$ of $\mathcal{V}_m$ are transferred by $K_\pm$ into $\mathcal{V}_{m\pm1}$, but remain eigenvectors of $K_\pm K_\mp$ for the fixed λ; still, some of them vanish, since the eigenequations (4.6) are trivially solved by the zero vector. As the rotation group is compact, all its irreducible representations are finite dimensional, and there are maximal k^+ and minimal k^- values of m for which $K_\pm|k^\pm\lambda\rangle = 0$. Thus, the subspace $\mathcal{V}_\lambda^{(k^+,k^-)}$ spanned by the vectors $\{|m\lambda\rangle \,|\, m = k^-,\dots,k^+\}$ is invariant and irreducible for the standard componenets K_3 and $K_\pm$, i.e. for the angular momentum.

Theorem 4.1 *All the irreducible representations of angular momentum are $D^{(k)}(\boldsymbol{K})$ with the maximal weight*

$$k = k^+ = -k^- = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots \tag{4.7}$$

They are $2k+1$-dimensional, and in the standard basis (4.5), denoted as $|km\rangle$ with $m = -k,\dots,k$, of the irreducible space $\mathcal{V}^{(k)}$, the matrices $D^{(k)}(\boldsymbol{K})$ are determined by:

$$K_3|km\rangle = m\hbar|km\rangle\,,\quad K_\pm|km\rangle = c_m^{(k)\pm}\hbar|k,m\pm1\rangle \tag{4.8a}$$

$$c_m^{(k)\pm} = \sqrt{k(k+1)-m(m\pm1)}. \tag{4.8b}$$

▬*Proof:* According to the previous results, in the single irreducible subspace K_3 is nondegenerate, making λ superfluous, and the standard basis is $|m\rangle$, with $m = k^-,\dots,k^+$. A direct consequence of (4.6a) are equations

$$K_3|m\rangle = m\hbar|m\rangle\,,\quad K_\pm|m\rangle = c_m^\pm\hbar|m\pm1\rangle\,,$$

having the form of (4.8a). The coefficients $c_m^\pm$ are to be still determined in accordance with the previously assumed orthonormality. Applying $K_+ = K_-^\dagger$ to act by K_+ on the left and on the right in $\langle m+1|\,K_+\,|m\rangle$, we get $c_{m+1}^{-\,*} = c_m^+$. Further, we already proved that there are minimal, k^- and maximal, k^+, value of m, for which $K_\pm|k^\pm\rangle = 0$, allowing to assume $c_{k^\pm}^\pm = 0$. Then we use (4.4b), to find

$$2\hbar\langle m|\,K_3\,|m\rangle = \langle m|\,[K_+,K_-]\,|m\rangle = \langle m|\,K_+K_- - K_-K_+\,|m\rangle = \|K_-\,|m\rangle\,\|^2 - \|K_+\,|m\rangle\,\|^2.$$

In the view of (4.8a), this reads:

$$2m = |c_m^-|^2 - |c_m^+|^2. \tag{4.9}$$

Now we sum these equations for all values of m, and find (Fig. 4.1):

$$2\sum_{m=k^-}^{k^+} m = \sum_{m=k^-}^{m=k^+}(|c_m^-|^2 - |c_m^+|^2).$$

When we substitute the equalities $|c_{m+1}^-| = |c_m^+|$ and $c_{k^\pm}^\pm = 0$ found previously, the right hand side vanishes (Fig. 4.1), and we get

$$\sum_{m=0}^{k^+-k^-}(k^- + m) = \frac{(k^+ + k^-)(k^+ - k^- + 1)}{2} = 0.$$

As $k^+ \geq k^-$, this implies $k^- = -k^+$. Finally, the difference $k^+ - k^- = 2k^+$ is an integer (since K_+ lifts m from k^- to k^+ in unit steps), meaning that $k = k^+ = 0, \frac{1}{2}, \ldots$, and $|\mathcal{V}^{(k)}| = 2k + 1$. Analogously, summing (4.9) from m to $k = k^+$, we find:

$$|c^-_m|^2 = |c^+_{m-1}|^2 = 2\sum_{s=m}^{k} s = 2(\sum_{s=1}^{k} s - \sum_{s=1}^{m-1} s) = k(k+1) - m(m-1),$$

which is, with a convention on the (common) phase, the value in the theorem. ∎

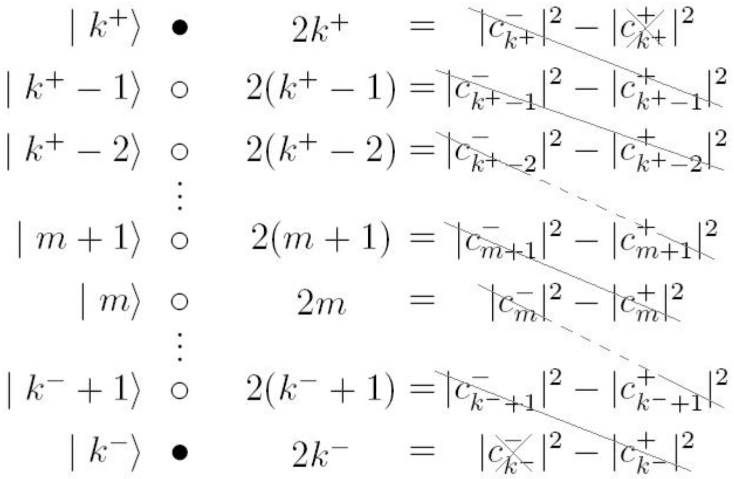

Figure 4.1: Standard basis of an irreducible representation of the angular momenta. For each standard vector $|m\rangle$ (denoted by ○, or ● for maximal and minimal weight) the corresponding equation (4.9) is written, and the terms which are canceled in the course of summation are singled out by the gray lines.

Exercise 4.1: Show that the matrices of the angular momenta in the standard basis are:

$$D^{(k)}(K_3) = \hbar \begin{bmatrix} -k & 0 & 0 & 0 & 0 \\ 0 & -k+1 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & k-1 & 0 \\ 0 & 0 & 0 & 0 & k \end{bmatrix}, D^{(k)}(K_+) = \hbar \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \sqrt{2k} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{4k-2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{4k-2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2k} & 0 \end{bmatrix}, D^{(k)}(K_-) = D^{(k)^\dagger}(K_+). \tag{4.10}$$

The matrices representing rotations can be now easily find from the first equation in (3.15). In particular, a rotation around z-axis is represented as

$$D^{(k)}(R_{\phi \boldsymbol{e}_z}) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\phi D^{(k)}(K_3)} = \mathrm{diag}[\mathrm{e}^{\mathrm{i}k\phi}, \mathrm{e}^{\mathrm{i}(k-1)\phi}, \ldots, \mathrm{e}^{-\mathrm{i}k\phi}].$$

It is important to note that $D^{(k)}(R_{2\pi \boldsymbol{e}_z}) = (\pm)^{2k} I_{2k+1}$, i.e. only for an integer k this is the identity, while this is the case for the rotations for multiples of 4π in the case of half-integral k. As mentioned in Section 4.1, this manifests the fact that the presented construction gives the representation of the Lie algebra of the rotational group, and not of the group itself. Therefore the exponential map of these representations coincides with the representations of the universal covering group SU(2) of the rotational group. Essentially, this is the consequence of the nontrivial fundamental group of the rotational group

SO(3). However, this seemingly artifact of the construction, appears to be a substantial physical property of the particles, as it will be seen in the discussion of the interior degrees of freedom, like spin (Section 4.5). In fact, each geometrical symmetry group G, usually considered as a subgroup of the extended Galilean group, essentially is only a homomorphic image of its double covering $\tilde{G}$ called double group: there is 2 to one homomorphism f from $\tilde{G}$ to G, such that each spin rotation u is mapped to the corresponding spatial rotation R.

4.2.2 Square of the angular momentum

Square of angular momentum is positive observable

$$K^2 = K_1^2 + K_2^2 + K_3^2. \tag{4.11}$$

It is easy to show that K^2 commutes with all the components K_i of angular momentum. Therefore, eigensubspaces of K^2, being invariant under all K_i, contain complete irreducible subspaces. Thus each irreducible subspace $\mathcal{V}^{(k)}$ is a subspace of an eigensubspace of K^2, meaning that in $\mathcal{V}^{(k)}$ operator K^2 acts as a multiplication constant. To find out this constant, we simply act on an arbitrary vector $|km,\lambda\rangle$ by K^2, written in the more convenient form:

$$K^2 = \frac{1}{2}(K_+K_- + K_-K_+) + K_3^2. \tag{4.12}$$

Then, using the results of the Theorem 4.1, we find

$$K^2 \left|km,\lambda\right\rangle = k(k+1)\hbar^2 \left|km,\lambda\right\rangle. \tag{4.13}$$

Exercise 4.2: Show (4.13).

Thus the eigensubspaces of K^2 are multiple irreducible subspaces, i.e. the sum of the irreducible subspaces of the same maximal weight k. This fact enables to find standard basis in any space where the operators of the angular momentum are defined as the common eigenvectors of K_3 and K^2, i.e. of the system of eigenproblems:

$$K^2 \left|km,\lambda\right\rangle = k(k+1)\hbar^2 \left|km,\lambda\right\rangle, \quad K_3 \left|km,\lambda\right\rangle = m\hbar \left|km,\lambda\right\rangle. \tag{4.14}$$

This is the most efficient way to determine standard basis in the state space, which will be systematically used later on.

4.2.3 Discussion

Operators of the angular momentum are related to the rotations by the relations (3.15). Essentially, this means that the angular momentum components span the Lie algebra of the rotational group. Therefore, the method used in the construction of the irreducible representations is the usual prescription in the theory of semisimple Lie algebras. In particular, the first two equation of (4.4) are the so called standard form of such an algebra, with K_3 being the only Cartan vector, while $K_\pm$ are roots. Analogously, (4.6a) manifests the well known relation between the roots and the weights. Finally, as the algebra is of rank 1, it has a single independent Kazimir operator, which is K^2.

It is useful to illustrate some differences in the physical contents of the classical and quantum angular momentum. As usual, the source of them is commutativity of the classical physical variables, in contrast to their quantum analogues.

From the uncertainty relations (4.2) one directly gets:

$$\Delta K_i \Delta K_j \geq \frac{\hbar}{2} \sum_k \epsilon_{ijk} |\langle K_k \rangle|. \tag{4.15}$$

Therefore, unless the average of one component of the angular momentum in the given state vanishes, the dispersions of the other components are nonzero.

Second, if the state $|x\rangle$ is non-dispersive for K_i, i.e. $\Delta K_i = 0$ meaning that it is an eigenstate of K_i, then from (4.15) follows that averages of both other components of angular momentum vanish.

Finally, in the classical mechanics only the component of the angular momentum along the axis of rotation is nonzero; those in the perpendicular directions vanish, as well as their squares. On the contrary, in quantum mechanics the averages $\langle K_i^2 \rangle$ and $\langle K_j^2 \rangle$ vanish only in the state with vanishing total angular momentum, $|k = 0, m = 0, \lambda\rangle$. This shows that the mentioned criterion to determine rotational axis is not applicable in quantum mechanics.

Exercise 4.3: Show that $\langle K_i^2 \rangle_{|x\rangle} = \langle K_j^2 \rangle_{|x\rangle} = 0$ if and only if for all the components $K_p |x\rangle = 0$, i.e. $|x\rangle = |k = 0, m = 0, \lambda\rangle$.

4.3 Orbital Angular Momentum

4.3.1 Coordinate representation

After general considerations we turn back to the original definition of angular momentum, i.e. to the orbital angular momentum $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$. Its Cartesian components $l_i = \sum_{jk} \epsilon_{ijk} r_i p_j$ in the coordinate representation, when momentum components $p_i = -\mathrm{i}\hbar \frac{\partial}{\partial r_i}$ are substituted, become:

$$l_1 = -\mathrm{i}\hbar \left(r_2 \frac{\partial}{\partial r_3} - r_3 \frac{\partial}{\partial r_2} \right), \quad l_2 = -\mathrm{i}\hbar \left(r_3 \frac{\partial}{\partial r_1} - r_1 \frac{\partial}{\partial r_3} \right), \quad l_3 = -\mathrm{i}\hbar \left(r_1 \frac{\partial}{\partial r_2} - r_2 \frac{\partial}{\partial r_1} \right). \tag{4.16}$$

The forthcoming considerations are most efficient performed in spherical coordinates. Also, angular momentum is particularly important (as conserved quantity) for the systems with spherical symmetry, for which spherical coordinates are natural choice. To find the spherical components of the angular momentum we use Cartesian components and apply coordinate transformations: $x = r \sin\theta \cos\varphi$, $y = r \sin\theta \sin\varphi$, $z = r \cos\theta \cos\varphi$. The Jacobian of this transformation relates the coordinate partial derivatives:

$$\begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \\ \frac{\partial}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \\ r\cos\theta\cos\varphi & r\cos\theta\sin\varphi & -r\sin\theta \\ -r\sin\theta\sin\varphi & r\sin\theta\cos\varphi & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}. \tag{4.17}$$

Its inverse is used to express the Cartesian derivatives in terms of spherical coordinates:

$$\begin{aligned}\frac{\partial}{\partial x} &= \sin\theta\cos\varphi\frac{\partial}{\partial r}+\frac{1}{r}\cos\theta\cos\varphi\frac{\partial}{\partial\theta}-\frac{1}{r}\frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial\varphi},\\ \frac{\partial}{\partial y} &= \sin\theta\sin\varphi\frac{\partial}{\partial r}+\frac{1}{r}\cos\theta\sin\varphi\frac{\partial}{\partial\theta}+\frac{1}{r}\frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial\varphi},\\ \frac{\partial}{\partial z} &= \cos\theta\frac{\partial}{\partial r}-\frac{1}{r}\sin\theta\frac{\partial}{\partial\theta}.\end{aligned}\tag{4.18}$$

Knowing this, Cartesian components are easily expressed in terms of spherical coordinates:

$$\begin{aligned}l_x &= \mathrm{i}\hbar(\sin\varphi\frac{\partial}{\partial\theta}+\cot\theta\frac{\partial}{\partial\varphi}),\\ l_y &= -\mathrm{i}\hbar(\cos\varphi\frac{\partial}{\partial\theta}-\cot\theta\frac{\partial}{\partial\varphi}),\\ l_z &= -\mathrm{i}\hbar\frac{\partial}{\partial\varphi}.\end{aligned}\tag{4.19}$$

We also find $l_\pm$ and l^2 as general theory emphasized their importance:

$$l_\pm=\hbar\mathrm{e}^{\pm\mathrm{i}\varphi}\left(\pm\frac{\partial}{\partial\theta}+\mathrm{i}\cot\theta\frac{\partial}{\partial\varphi}\right),\quad l^2=-\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right].\tag{4.20}$$

All the considered operators act in the Lebesgue space $\mathcal{L}(\mathbb{R}^3)$ of the single three dimensional particle, which is the product of the three Lebesgue spaces over Cartesian coordinates: $\mathcal{L}(\mathbb{R}^3)=\mathcal{L}(\mathbb{R}_x)\otimes\mathcal{L}(\mathbb{R}_y)\otimes\mathcal{L}(\mathbb{R}_z)$. On the other side, this space is seemingly product of the Lebesgue spaces over the spherical coordinates. It would be very convenient to retain this product structure in spherical coordinate system, although due to the singularities of this system this is not completely correct (e.g., functions $f(r,\theta,\varphi)$ are acceptable only if for $r=0$ they are independent of θ and φ, i.e. if $f(0,\theta,\varphi)$ is constant; otherwise the value in coordinate origin is undefined). So, these deficiencies are cured by extra conditions on the functions $f(r,\theta,\varphi)$:

$$\mathcal{L}(\mathbb{R}^3)=\left\{f(r,\theta,\varphi)\in\mathcal{L}([0,\infty)_r)\otimes\mathcal{L}([0,\pi]_\theta)\otimes\mathcal{L}([0,2\pi]_\varphi)\,\middle|\,{}^{f(0,\theta,\varphi)=\text{const},}_{\substack{f(r,\theta=0/\pi,\varphi)=f_{0/\pi}(r),\\ f(r,\theta,\varphi)=f(r,\theta,\varphi+2\pi)}}\right\}.\tag{4.21}$$

Note that none of the angular momentum operators (4.19) and (4.20) are r dependent, and therefore they act as identity in the factor space $\mathcal{L}([0,\infty)_r)$. This reflects that the angular momentum generates rotations, which do not affect the radius. A consequence is that the orbital angular momentum is effectively defined in the angular Lebesgue space of the functions over the sphere (Fig. 4.2):

$$\mathcal{L}(S^2)=\left\{f(\theta,\varphi)\in\mathcal{L}([0,\pi]_\theta)\otimes\mathcal{L}([0,2\pi]_\varphi)\,\middle|\,\substack{f(\theta=0/\pi,\varphi)=\text{const}_{0/\pi},\\ f(\theta,\varphi)=f(\theta,\varphi+2\pi)}\right\}.\tag{4.22}$$

The last condition explicates 2π periodicity of all the functions in the Lebesgue space. As a consequence, the orbital angular momentum has only integer irreducible components, as it will be explicated soon.

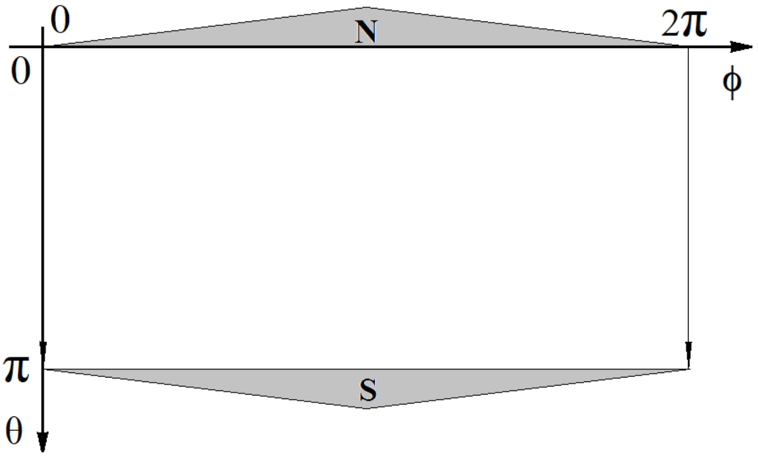

Figure 4.2: Obtaining the sphere S^2 from a rectangle $(\theta,\varphi) \in [0,\pi] \times [0,2\pi]$ by identifications in (4.22). Firstly, since $f(\theta,\varphi) = f(\theta,\varphi+2\pi)$, two vertical arrows are mutually identified to obtain cylinder. Secondly, φ-independence, $f(\theta=0,\varphi) = \mathrm{const}_0$, $f(\theta=\pi,\varphi) = \mathrm{const}_\pi$, is contraction (depicted by shaded triangle) of the upper (bottom) boundary circle of the cylinder into the single north (south) pole point of the resulting sphere.

4.3.2 Standard Basis: Spherical Harmonics

According to the general theory, standard basis is common eigenbasis of l_z and l^2. The ineffectiveness of these operators in the radial factor space allows to apply the theorem on the separation of variables (Appendix **??**), i.e. to consider the eigenproblem in $\mathcal{L}(S^2)$ only: the obtained eigenbasis $Y_l^m(\theta,\varphi) = \langle\theta,\varphi \,|\, lm\rangle$ in this space may be multiplied by any basis $f_n(r)$ of $\mathcal{L}([0,\infty)_r)$, to get basis $f_n(r)Y_l^m(\theta,\varphi) = \langle r,\theta,\varphi \,|\, nlm\rangle$ in the total state space.

In the view of (4.14) functions $Y_l^m(\theta,\varphi)$ are solution of the system of partial differential equations:

$$-\mathrm{i}\tfrac{\partial}{\partial\varphi}Y_l^m(\theta,\varphi) = mY_l^m(\theta,\varphi), \tag{4.23a}$$

$$-\left[\tfrac{1}{\sin^2\theta}\tfrac{\partial^2}{\partial\varphi^2} + \tfrac{1}{\sin\theta}\tfrac{\partial}{\partial\theta}\left(\sin\theta\tfrac{\partial}{\partial\theta}\right)\right] Y_l^m(\theta,\varphi) = l(l+1)Y_l^m(\theta,\varphi). \tag{4.23b}$$

Since l_z effectively acts in the factor space $\mathcal{L}([0,2\pi]_\varphi)$, one again applies theorem on the separation of variables to solve it. Obviously, the first of these equations is solved by the functions $g_l^m(\theta)\mathrm{e}^{\mathrm{i}m\varphi}$, with m restricted to the integer values by the last condition in (4.21). In the second eigenproblem we again use separation of variables: $l^2 = A_\theta \otimes l_z + B_\theta \otimes \mathbf{1}_\varphi$; the factors in $\mathcal{L}([0,2\pi]_\varphi)$ commute, the second one being identity, while the first one is function of l_z. Thus, the eigenfunctions have form $g_l^m(\theta)\mathrm{e}^{\mathrm{i}m\varphi}$; substituting this, the second eigenproblem becomes equation in $g_l^m(\theta)$:

$$\left[\frac{m^2}{\sin^2\theta} - \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - l(l+1)\right] g_l^m(\theta) = 0. \tag{4.24}$$

After transition to new variable $\chi = \cos\theta$ (consequently $\frac{\partial}{\partial\chi} = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}$), this becomes general hypergeometric equation:

$$\left[\frac{\partial^2}{\partial\chi^2} - \frac{2\chi}{1-\chi^2}\frac{\partial}{\partial\chi} + \frac{l(l+1)(1-\chi^2)-m^2}{(1-\chi^2)^2}\right] g_l^m(\chi) = 0. \tag{4.25}$$

The first condition in (4.22) implies that for $m \neq 0$ only functions such that $g_l^m(\chi = \pm 1) = 0$ should be considered. Straightforward application of the general method (Appendix **??**) shows that the for each $l = 0, 1, \dots$ and $m = 0, \pm 1, \dots, \pm l$ there is a single solution of (4.25) in the form of the Legendre polynomial $g_l^m(\chi) = (1-\chi^2)^{\frac{m}{2}} P_{l-|m|}^{(|m|,|m|)}(\chi)$. Finally, including the factors providing normalization

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \,\mathrm{d}\theta Y_l^{m^*}(\theta,\varphi) Y_{l'}^{m'}(\theta,\varphi) = \delta_{ll'}\delta_{mm'}, \tag{4.26}$$

one gets the spherical harmonics as the standard basis:

$$Y_l^m(\theta,\varphi) = \frac{(-)^{l-m}}{2^l l! \sqrt{2\pi}} \sqrt{\frac{2l+1}{2}\frac{(l+m)!}{(l-m)!}} \frac{1}{\sin^{|m|}\theta} \mathrm{e}^{\mathrm{i}m\varphi} \frac{\mathrm{d}^{l-|m|}}{\mathrm{d}(\cos\theta)^{l-|m|}} \sin^{2l}\theta. \tag{4.27}$$

Exercise 4.4: Using the method explained in Appendix **??** prove that (4.27) are solutions of (4.23).

Exercise 4.5: Show transformation rule of the spherical harmonics under the spatial inversion:

$$\mathcal{I}_S Y_l^m(\theta,\varphi) = (-)^l Y_l^m(\theta,\varphi). \tag{4.28}$$

4.4 Central potentials

Quantum numbers l and m are conserved whenever orbital angular momentum commutes with Hamiltonian. In such cases l^2 and l_z may be used together with Hamiltonian (and if necessary some other observables) to form a complete set of compatible observables. In this way Schrödinger equation is simplified, with clear interpretation of the solutions. Kinetical part of Hamiltonian is reduced to the square of momentum, and always commutes with angular momentum components. Hence, compatibility of H and angular momentum is determined by the potential. The condition $[V(\boldsymbol{r}), l_i] = 0$ immediately means that the potential commutes with all rotations: $[V, D(R_{\boldsymbol{\phi}})] = 0$, i.e. $D^{-1}(R_{\boldsymbol{\phi}})VD(R_{\boldsymbol{\phi}}) = V$. Assuming that the potential depends only on the coordinates, the last condition reads that the potential must have spherical symmetry:

$$D^{-1}(R_{\boldsymbol{\phi}})V(\boldsymbol{r})D(R_{\boldsymbol{\phi}}) = V(R_{\boldsymbol{\phi}}\boldsymbol{r}) = V(\boldsymbol{r}). \tag{4.29}$$

This analysis emphasizes that in the problems with spherical symmetry only, the orbital angular momentum has particularly important role, with its components being conserved quantities. Technically this fact is efficiently exploited if l^2 and l_z are included in CSCO together with H.

Obviously, in the spherical coordinates (4.29) shows that the potential depends only on the distance r from the coordinate origin. Indeed, acting on some chosen point $\boldsymbol{r} = (r, \theta, \varphi)$, the set of all rotations generates sphere with the radius r, and along this sphere $V(\boldsymbol{r})$ is constant, meaning that V depends on r only, i.e. (4.29) is satisfied if and only if the potential is central $V(\boldsymbol{r}) = V(r)$.

According to the results of Section 4.3, in such problems the time independent Schrödinger equation is solved by the wave functions of the form $R_{nl}(r)Y_l^m(\theta, \varphi)$. In fact, as in the angular factor space (4.22) of the total space (4.21) l^2 and l_z are CSCO, the spherical harmonics are necessarily angular part of any CSCO in the total space. Therefore, only the radial functions $R_{nl}(r)$ are to be determined for any specific potential. Really, since the Hamiltonian of the particle in the central potential is:

$$H = \frac{p^2}{2m} + V(r) = -\frac{\hbar^2}{2m}\triangle + V(r). \tag{4.30}$$

Further, according to (4.20) the laplacian has the form $\triangle = \frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} - \frac{1}{\hbar^2 r^2}l^2$. Hence, using the theorem on the separation of variables, the time independent Schrödinger equation reduces to the radial part only:

$$\left[\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}r^2\frac{\mathrm{d}}{\mathrm{d}r} + \left(\frac{2m}{\hbar^2}(E_{nl} - V(r)) - \frac{l(l+1)}{r^2}\right)\right]R_{nl}(r) = 0. \tag{4.31}$$

Recall that the radial eigenbasis is to be found solving such an equation for each $l = 0, 1, \dots$. Particularly, as only for $l = 0$ spherical harmonic is constant, for $l > 0$ the first condition in (4.21) implies that $R_{nl}(r = 0) = 0$. The finite norm of the solutions means that $\int_0^\infty r^2\,\mathrm{d}r|R_{nl}(r)|^2 < \infty$.

In some problems it is convenient to use $\tilde{R}_{nl}(r) = rR_{nl}(r)$, simplifying (4.31) to

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{2m}{\hbar^2}(E_{nl} - V(r)) - \frac{l(l+1)}{r^2}\right]\tilde{R}_{nl}(r) = 0. \tag{4.32}$$

Again, for $l > 0$ the first condition in (4.21) implies that for $r \to 0$ the limiting value of the function is $\tilde{R}_{nl}(r)/r \to 0$; the normalization is enabled by $\int_0^\infty \mathrm{d}r|\tilde{R}_{nl}(r)|^2 < \infty$. The form of (4.32) allows to interpret it as one-dimensional (on the half line) Schrödinger's equation, with the potential $V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}$; the latter central barrier repulses particles from the central region.

4.4.1 Free particle

It has been shown that the plane waves are solutions of the time independent Schrödinger equation of the free particle, $V(r) = 0$. These are found taking momentum components as the CSCO compatible with Hamiltonian. Recall that the eigenenergies $E_{\boldsymbol{k}} = E_k = \hbar^2k^2/2m$ are degenerate along the spheres of the radius k.

However, the constant zero potential is spherically symmetric, and therefore CSCO can be formed by H, l^2 and l_z. Consequently, there must be an alternative eigenbasis of H, subdued to this choice of CSCO. As discussed, the latter two give the spherical harmonics in the factor space $\mathcal{L}(S^2)$, and the radial component is completed by the solutions of the radial equation (4.32), which becomes

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2}+\frac{2m}{\hbar^2}E-\frac{l(l+1)}{r^2}\right]\tilde{R}_{nl}(r)=0.$$

Denoting $\frac{2m}{\hbar^2}E$ by k^2 (in accordance with the known solution), it turns out that the variable in equation is $x=kr$; thus

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}+1-\frac{l(l+1)}{x^2}\right]\tilde{R}_{kl}(x)=0. \tag{4.33}$$

Instead of n, the energies are counted by $k\geq 0$. As the equation itself is k independent, the solvability gives no restriction on k; the spectrum is continual (as expected), and for each k the same set of solutions (the functions of x, not of r!) must be obtained. The equation can be solved iteratively. In fact, for $l=0$, it reduces to $\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}+1\right)\tilde{R}_{k0}(x)=0$, solved by $\tilde{R}_{k0}(x)=2\sin x$. On the other hand, the solution can be looked for in the form $\tilde{R}_{kl}(x)=x^{l+1}\chi_{kl}(x)$, when (4.33) becomes equation in χ:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}+\frac{2(l+1)}{x}\frac{\mathrm{d}}{\mathrm{d}x}+1\right]\chi_{kl}(x)=0. \tag{4.34}$$

To solve it, one differentiate the whole equation, getting an equation in $\chi'_{kl}(x)$, which, after substitution $\chi'_{kl}(x)=x\phi_{kl}(x)$ becomes equation in the auxiliary function $\phi_{kl}(x)$:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}+\frac{2(l+2)}{x}\frac{\mathrm{d}}{\mathrm{d}x}+1\right]\phi_{kl}(x)=0.$$

Notably, it is the same as (4.34), but for $l+1$. Therefore, $\chi_{k,l+1}(x)=\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\chi_{kl}(x)$, meaning that $\chi_{k,l}(x)=(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x})^l\chi_{k0}(x)$. This conclusion, together with the found result for $\tilde{R}_{k0}(x)$ straightforwardly gives: $\tilde{R}_{kl}(x)=x^{l+1}(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x})^l 2\frac{\sin x}{x}$. Finally, the solutions of the radial part are:

$$R_{kl}(r)=2(-)^l\frac{r^l}{k^l}(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r})^l 2\frac{\sin kr}{r}=2kj_l(kr). \tag{4.35}$$

Here, $j_l(kr)$ are Bessel's functions of the first kind (integer, or spherical Bessel's functions). Altogether, the generalized eigenbasis for CSCO with angular momentum is

$$\psi_{klm}(\boldsymbol{r})=\langle r\theta\varphi\,|\,klm\rangle=2kj_l(kr)Y_l^m(\theta,\varphi). \tag{4.36}$$

4.4.2 Coulomb potential and Hydrogen like atoms

Hamiltonian of the relative particle of the hydrogen like atom with Z protons in the nucleus is

$$H=\frac{p^2}{2m}-\frac{Ze^2}{r}, \tag{4.37}$$

where $m = m_{\mathrm{e}} m_{\mathrm{N}}/(m_{\mathrm{e}} + m_{\mathrm{N}})$ and $r = |\boldsymbol{R}_{\mathrm{N}} - \boldsymbol{r}_{\mathrm{e}}|$; clearly, $m \approx m_{\mathrm{e}}$ since $m_{\mathrm{e}} \ll m_{\mathrm{N}}$. The radial Schrödinger equation becomes:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{2m}{\hbar^2}\left(E_{nl} + \frac{Ze^2}{r}\right) - \frac{l(l+1)}{r^2}\right]\tilde{R}_{nl}(r) = 0. \tag{4.38}$$

This equation takes more standard form if the involved quantities are reparameterized. Using Bohr radius $a_0 = \hbar^2/m_{\mathrm{e}}e^2 = 0.529\,\mathrm{Å}$, and atomic unit of energy (one Henry equal to two Rydbergs) $e^2/a_0 = 4.36\,10^{-18}\,\mathrm{J} = 27.21\,\mathrm{eV}$, the energy and radius are substituted by the dimensionless $\varepsilon_{nl} = a_0 E_{nl}/e^2$ and $x = r/a_0$, to get the generalized hypergeometric equation:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{2\varepsilon_{nl}x^2 + 2Zx - l(l+1)}{x^2}\right]\tilde{R}_{nl}(x) = 0. \tag{4.39}$$

The standard method of solving of such equation (Appendix **??**) straightforwardly gives that the discrete energy spectrum and the corresponding orthonormal radial functions are for each $n = 1, 2, \dots$ and $l = 0, \dots, n-1$:

$$E_n = -\frac{me^4}{2\hbar^2}\frac{Z^2}{n^2}, \quad R_{nl}(r) = C_{nl}\mathrm{e}^{-\frac{2Z}{n}\frac{me^2}{\hbar^2}r}\left(\frac{me^2}{\hbar^2}r\right)^{l+1} L^{2l+1}_{n-l-1}\left(\frac{2Z}{n}\frac{me^2}{\hbar^2}r\right); \tag{4.40}$$

the normalizing factor is $C_{nl} = \sqrt{\frac{Z(n-l-1)!}{n!(n+l)!}}$, and $L^p_q(x)$ are generalized Laguerre polynomials. Important characteristic of the Coulomb potential is that the energies are degenerate in l, which is due to the hidden symmetry of another group isomorphic to the rotational one (but not related to the spatial rotations). This enlarges the degeneracy of energy levels to $\sum_{l=0}^{n-1}(2l+1) = n^2$.

Finally, note that only bounded (discrete) levels are found above. The continual part of the spectrum is positive part of the real axis. The corresponding states, relevant for various scattering processes, can be found (together with the states of the discrete spectrum) with help of the parabolic coordinates [11, &37,&135].

4.5 Spin

There are several experiments showing that the orbital space $\mathcal{S}_{\mathrm{o}}$ which we are dealing with in the previous sections is not sufficient to describe completely the single particle states.

4.5.1 Zeeman's Effect

Hamiltonian of the particle with the mass m and charge q in the electromagnetic field defined by the scalar and vector potential Φ and $\boldsymbol{A}$ is

$$H = \frac{1}{2m}\left(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right)^2 + q\Phi \tag{4.41}$$

(c is the speed of light). Particularly, the potentials of the homogeneous constant magnetic field $\boldsymbol{B} = B\boldsymbol{e}_z$ can be in the Coulomb gauge ($\mathrm{div}\boldsymbol{A} = 0$) taken in the form $\boldsymbol{A} = \frac{1}{2}\boldsymbol{B} \times \boldsymbol{r}$

and $\Phi = 0$, giving

$$\left(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\right)^2 = p^2 - \frac{q}{2c}[\boldsymbol{p}\cdot(\boldsymbol{B}\times\boldsymbol{r}) + (\boldsymbol{B}\times\boldsymbol{r})\cdot\boldsymbol{p}] + \frac{q^2}{4c^2}(\boldsymbol{B}\times\boldsymbol{r})^2 = p^2 - \frac{q}{c}Bl_z + \frac{q^2}{4c^2}B^2r_\perp^2.$$

Here, $r_\perp$ is the radius vector component perpendicular to the field.

Consequently, when a hydrogen like atom is exposed to such a magnetic field, Hamiltonian of the electron becomes:

$$H = -\frac{\hbar^2}{2m}\triangle - \frac{Ze^2}{r} + \frac{\mu_\mathrm{B}}{\hbar}Bl_z + \frac{e^2B^2}{8mc^2}r_\perp^2, \tag{4.42}$$

where $\mu_\mathrm{B} = \frac{e\hbar}{2mc} = 0.927\ 10^{-23}$ J/T is the Bohr's magneton. The terms involving magnetic field can be roughly estimated by the mean values for the lowest hydrogen like atom states: $\langle r_\perp^2\rangle \sim a_0^2$, $\langle l_z\rangle \sim \hbar$, one finds that $\langle\mu_\mathrm{B}Bl_z\rangle \sim 0.927\ 10^{-23}B$ J/T, $\langle\frac{e^2B^2}{8mc^2}r_\perp^2\rangle \sim 4\ 10^{-26}B^2\mathrm{J/T}^2$. The maximal experimental magnetic fields are of the order 1 T, making therefore the last term by thousands less then the third term. In addition, assuming that the typical energies of the lowest hydrogen like atom states are of the order of 1 eV$= 1.6\ 10^{-19}$ J, the third term is by thousands less than the energy of the isolated atom. Therefore the last term may be neglected, while the rest of Hamiltonian may be treated perturbatively, using the isolated atom Hamiltonian as the nonperturbed one, while the third therm is perturbation. Accordingly, using the expression (6.5) for the first correction of the degenerate level, and noticing that the hydrogen like atom eigenstates $|nlm\rangle$ are also eigenstates for the perturbation, one finds the perturbed energies

$$E_{nm} = E_n + \mu_\mathrm{B}Bm. \tag{4.43}$$

Thus, quite generally, the energy level is split in the $2l+1$ equidistant sublevels. This is called Zeeman's effect. Particularly, for the hydrogen like atoms, with energy levels E_n being degenerate in angular momentum values $l = 0,\dots n-1$, this result implies that each level is in the magnetic field split into $2l_\mathrm{max}+1$ equidistant sublevels E_{nm} for $m = -l,\dots,l$ and $l_\mathrm{max} = n-1$.

However, in the experiments with hydrogen like atoms, it appears that there are even number of sublevels, and this is called anomalous Zeeman's effect. Obviously, it cannot be described within introduced formalism.

4.5.2 Interior degrees of freedom

Further evidence on the incompleteness of the formalism was given by the Stern-Gerlach experiment: passing through the inhomogeneous magnetic field, well collimated monochromatic beam of atoms with a single electron is split into parts. Analysis analogous to that of the Zeeman's effect shows again that the number of obtained beams should be odd, e.g. no splitting can occur for the single s-state electron. But in such cases even number of split beams is found experimentally.

Still, there is a way to explain the described effects within quantum mechanical formalism. The Postulate 4 emphasizes that the total state space is the direct product of the state spaces for each degree of freedom. In the view of this, the results of the previous experiments shows that there is another degree of freedom, being not of the orbital type: the possible values are not parameterized by the continual variable (e.g. coordinate or momentum), but take a finite number of values. Such degrees are called interior degrees of freedom.

POSTULATE 6. – INTERIOR DEGREES

Besides orbital, there are also interior degrees of freedom, contributing to the total state space $\mathcal{S}$ by the finite dimensional interior factor space:

$$\mathcal{S} = \mathcal{S}_{\text{orbital}} \otimes \mathcal{S}_{\text{interior}}.$$

The construction of the total space as the product of the orbital and interior ones is essentially required by the previous Postulate 4. Of course, the same argument shows that if there are several interior degrees of freedom, the interior space is constructed as the direct product of the corresponding interior spaces: $\mathcal{S}_{\text{interior}} = \mathcal{S}^1_{\text{interior}} \otimes \mathcal{S}^2_{\text{interior}} \otimes \dots$. In comparison with the previously introduced spaces based on the particle position (continuously parameterized by coordinates), the only difference is finiteness of the interior spaces.

Finite dimension of the space of function over some set X means that the set itself is finite. Indeed, as any function over finite (meaning also discrete) set is continuous, such a space is spanned by the functions f_x being zero everywhere but in a single x from X, e.g. $f_x(y) \overset{\text{def}}{=} \delta_{xy}$ (Kronecker step function). Obviously, the dimension of the spanned space $\mathcal{L}^2(X)$ is equal to the number of points in X. Consequently, the last postulate means that the interior degrees of freedom are finite sets.

This fact may be considered as an empirical one, although relativistic theory gives a profound proof for it through the analysis of the possible extensions of the Poincare group: Coleman-Mandula theorem asserts that any group having the Poincare group as a subgroup, and describing the space-time symmetries in the physically correct way, must be a direct product of the Poincare group and some compact groups. Therefore, its irreducible representations are products of the representations of the Poincare group with the *finite-dimensional* representations of the compact factor. So, while the Poincare group leads to infinite dimensional orbital space (its unitary irreducible representations are infinite dimensional), the compact factors introduce finite-dimensional interior spaces.

SUDBERRY??? Intrinsic parity

4.5.3 Formalism of the spin $s = \frac{1}{2}$

The simplest nontrivial case of the angular momentum is the representation for $s = 1/2$ (in this special case s is used instead of general k). According to the Theorem 4.1, the dimension of the space $\mathcal{S}^{(\frac{1}{2})}$ of this representation is $|\mathcal{S}^{(\frac{1}{2})}| = 2s + 1 = 2$. A suitable hermitian basis of the corresponding operator space (four dimensional) is formed by $\sigma_0 =$

$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ and three Pauli matrices:

$$\sigma_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \sigma_2 = \begin{pmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{pmatrix}, \quad \sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{4.44}$$

These matrices are used (for more details concerning physically relevant notions and notation related to the space $\mathbb{C}^2$ see Chapter 7) to define the spin matrices by

$$S_i = \frac{1}{2}\hbar\sigma_i \quad (i = 1, 2, 3). \tag{4.45}$$

It is easy to verify that

$$[S_i, S_j] = \mathrm{i}\hbar\sum_k \epsilon_{ijk}S_k, \tag{4.46}$$

showing that these operators represent the angular momentum components in $\mathcal{S}^{(\frac{1}{2})}$. Accordingly, the spin square eigenvalue is $s(s+1)\hbar^2 = \frac{3}{4}\hbar^2$, which also follows from $S^2 = \sum_i S_i^2 = \frac{3}{4}\hbar^2\sigma_0$.

Using the general properties of angular momentum (or directly solving the eigenproblem of S_3), one finds that the standard basis in $\mathcal{S}^{(\frac{1}{2})}$ is $\{|+\rangle = |m = \frac{1}{2}\rangle, |-\rangle = |m = -\frac{1}{2}\rangle\}$, for the eigenvalues $\pm\frac{1}{2}\hbar$ of S_3.

The rotations in the spin space are represented by the operators $D^{(\frac{1}{2})}(R_{\phi\boldsymbol{u}}) = \mathrm{e}^{-\frac{\mathrm{i}}{2}\phi\boldsymbol{u}\cdot\boldsymbol{\sigma}}$. Particularly, in the standard basis to the rotation for ϕ around z axis corresponds the matrix $\begin{pmatrix}\exp(-\frac{\mathrm{i}}{2}\phi) & 0\\ 0 & \exp(\frac{\mathrm{i}}{2}\phi)\end{pmatrix}$. Obviously, for $\phi = 2\pi$ the matrix $-I_2$ is obtained, while the identity transformation is achieved only for $\phi = 4\pi$. This shows that the matrices for spin $\frac{1}{2}$ represent group SU(2).

4.6 Addition of angular momenta

Let us consider state space being product

$$\mathcal{S} = \mathcal{S}_1 \otimes \mathcal{S}_2, \tag{4.47}$$

with angular momenta $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ in the factors spaces. Such situation is very frequent: e.g. orbital and spin momenta of a single electron, or two particles with angular momenta of each of them. It is easily found that the the defining for angular momenta relations $[K_{pi}, K_{pj}] = \mathrm{i}\hbar\sum_k \epsilon_{ijk}K_{pk}$, $(p = 1,\ 2)$ imply that the whole space $\mathcal{S}$ operators

$$K_i = K_{1i} \otimes I_2 + I_1 \otimes K_{2i}, \quad i = 1, 2, 3, \tag{4.48a}$$

also satisfy commutation relations (4.2); in other words they also form angular momentum $\boldsymbol{K}$ being *sum of the momenta* $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$:

$$\boldsymbol{K} = \boldsymbol{K}_1 + \boldsymbol{K}_2. \tag{4.48b}$$

Components of $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are given by $K_{1i} = K_{1i}^{(1)} \otimes I_2$ and $K_{2i} = I_1 \otimes K_{2i}^{(2)}$.

4.6.1 Irreducible subspaces and standard basis

For various purposes it is useful to know irreducible components and standard basis of the composed momentum $\boldsymbol{K}$. It is assumed that irreducible components of the momenta $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are known, i.e. that the decompositions

$$\mathcal{S}_i = \oplus_{k_i,\lambda_i} \mathcal{S}_i^{(k_i\lambda_i)}, \quad i = 1, 2 \tag{4.49}$$

are given. Here, $\mathcal{S}_i^{(k_i\lambda_i)}$ are irreducible subspaces of the momentum $\boldsymbol{K}_i$, where $k_i(k_i+1)\hbar$ is eigenvalue of $\boldsymbol{K}_i^2$ and λ_i counts different irreducible subspaces of the same type when frequency number of k_i is greater than 1.

Irreducible components and Clebsch-Gordan series

The product structure (4.47) of $\mathcal{S}$, together with decomposition (4.49) gives:

$$\mathcal{S} = \oplus_{k_1,k_2} \oplus_{\lambda_1,\lambda_2} \mathcal{S}_1^{(k_1\lambda_1)} \otimes \mathcal{S}_2^{(k_2\lambda_2)}. \tag{4.50}$$

Therefore, to reveal the irreducible components of angular momenta in $\mathcal{S}$ it is sufficient to find out the irreducible components in the product $\mathcal{V}^{(k_1)} \otimes \mathcal{V}^{(k_2)}$ of irreducible subspaces, i.e. Clebsch-Gordan series.

Theorem 4.2 ***(Clebsch-Gordan series of angular momentum)*** *The product of the irreducible spaces of the weight k_1 and k_2 of angular momentum contains exactly once each irreducible subspace with weight $k = |k_1 - k_2|, \dots, k_1 + k_2$:*

$$\mathcal{V}^{(k_1)} \otimes \mathcal{V}^{(k_2)} = \oplus_{k=|k_1-k_2|}^{k_1+k_2} \mathcal{V}^{(k)}. \tag{4.51}$$

▬*Proof:* This is one of the theorems easily derived as a special case of the Weyl's theorem on the semisimple Lie algebras. Here we recall that in each irreducible space $|k_i m_i\rangle$ ($m_i = -k_i, \dots, k_i$) is standard basis. One basis of the direct product space is $|k_1 m_1; k_2 m_2\rangle = |k_1 m_1\rangle \otimes |k_2 m_2\rangle$, and this basis is represented by the points (m_1, m_2) in Fig. 4.3. Vectors of this basis are eigenvectors of K_3:

$$K_3 \, |k_1 m_1; k_2 m_2\rangle = (m_1 + m_2)\hbar \, |k_1 m_1; k_2 m_2\rangle \, .$$

Therefore, vectors with the same sum $m = m_1 + m_2$ span eigensubspaces of K_3. These subspaces are denoted by the inclined lines connecting vectors with the same m; at the end of each line the corresponding $m_1 + m_2$ (generally and in particular example) and the number d_m of vectors on it are written. Obviously, the maximal sum is $k_1 + k_2$ with a single vector $|k_1 k_1; k_2 k_2\rangle$ on it. The sum takes on all the values $k_1 + k_2, k_1 + k_2 - 1, \dots, -(k_1 + k_2)$, and the degeneracy of these values increases by one for $m = k_1 + k_2, k_1 + k_2 - 1, \dots, k_1 - k_2$ (we assumed that $k_1 \geq k_2$), then remains constant for $m = k_1 - k_2, k_1 - k_2 - 1, \dots, k_2 - k_1$, and finally decreases by one for $m = k_2 - k_1, k_2 - k_1 - 1, \dots, -k_1 - k_2$. Thus, in the decomposition appears single irreducible subspace corresponding to the maximal value $k_1 + k_2$; it must contain a single eigenvector of K_3 for all eigenvalues $k_1 + k_2, k_1 + k_2 - 1, \dots, -(k_1 + k_2)$. When this representation is taken into account, all degeneracies should be diminished by 1. Therefore, among the remaining m values, the maximal $m_1 + m_2 - 1$ is now nondegenerate. We repeat the procedure, collecting all lower eigenvalues corresponding to this representation, and subtracting from the diagram. So, all the representations from $k_1 + k_2$ to $k_1 - k_2$ appear once, and after this the diagram is exhausted (as the degeneracies are constant). ▮

In this way the total space reduction (4.50) becomes

$$\mathcal{S} = \oplus_{k_1,k_2} \oplus_{k=|k_1-k_2|}^{k_1+k_2} \oplus_{\lambda_1,\lambda_2} \mathcal{V}_{\lambda_1\lambda_2}^{(k)}. \tag{4.52}$$

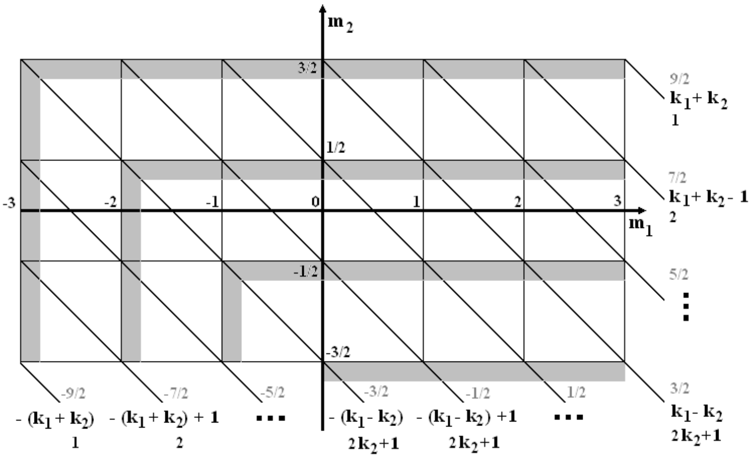

Figure 4.3: Decomposition of the product of irreducible subspaces of angular momentum. Case $k_1 = 3$ and $k_2 = 3/2$ is illustrated. At the end of each inclined line the sum $m_1 + m_2$ for the point (m_1, m_2) on it is depicted (concrete example is given above in gray) and the number of these points (degeneracy of the sum value) is below. Values of $m_1 + m_2$ included in the irreducible subspace with fixed maximal weight (written on the right side of diagram) are the lines crossed by the shaded region corresponding to maximal weight.

Standard basis and Clebsch-Gordan coefficients

After the Clebsch-Gordan series are found, we can proceed to determine the standard basis in the product space. According to the last decomposition, one can use standard basis of the direct product of the irreducible spaces. Namely, if $|k_i, m_i\rangle$ $(i = 1, 2)$ are standard bases in the irreducible spaces $\mathcal{V}^{(k_1)}$ and $\mathcal{V}^{(k_2)}$, then in the product space, since each irreducible component occurs once, the standard basis $|k_1 k_2; km\rangle$ is unique (up to the common phase factor). Therefore the non-correlated $|k_1, m_1; k_2, m_2\rangle = |k_1 m_1\rangle \otimes |k_2 m_2\rangle$ and standard basis are related by the unitary transition matrix, with elements being Clebsch-Gordan coefficients $\langle k_1, m_1; k_2, m_2 \,|\, k_1 k_2; km\rangle$:

$$|k_1 k_2; km\rangle = \sum_{m_1=-k_1}^{k_1} \sum_{m_2=-k_2}^{k_2} \langle k_1, m_1; k_2, m_2 \,|\, k_1 k_2; km\rangle \, |k_1, m_1; k_2, m_2\rangle \,. \tag{4.53}$$

The Clebsch-Gordan coefficients as completely mathematical (i.e. independent of physical context) constants are calculated and tabulated[2] in various forms (including computer programs). Still, it is interesting to exploit the eigensubspaces of K_3 used in the proof of the Theorem 4.2. In the first step we found that the standard vector of the maximal weight is noncorrelated: $|k_1 k_2; k = k_1 + k_2, m = k_1 + k_2\rangle = |k_1, m_1 = k_1; k_2, m_2 = k_2\rangle$. Then, using the operator $K_- = K_{1-} \otimes I_2 + I_1 \otimes K_{2-}$ we find (see Theorem 4.1):

$$K_- \, |k_1 k_2; k_1 + k_2, k_1 + k_2\rangle = c_m^{(k)-} \hbar \, |k_1 k_2; k_1 + k_2, k_1 + k_2 - 1\rangle \,.$$

---

[2] Sometimes, instead of Clebsch-Gordan coefficients, their particular biunique function called $3j$-symbols are used.

This vector belongs to the eigenspace of K_3 for eigenvalue $k_1 + k_2 - 1$. This eigenspace (Fig. 4.3, the second inclined line) is double degenerate. Thus, there is only one (independent) vector orthogonal to $|k_1k_2; k_1 + k_2, k_1 + k_2\rangle$, and it must be $|k_1k_2; k_1 + k_2 - 1, k_1 + k_2 - 1\rangle$. Then, in the triply degenerate subspace of the eigenvalue $m = k_1 + k_2 - 2$ two basis vectors are proportional to $K_-^2 \, |k_1k_2; k_1 + k_2, k_1 + k_2\rangle$ and $K_- \, |k_1k_2; k_1 + k_2 - 1, k_1 + k_2 - 1\rangle$, and the vector orthogonal to these two is (up to a phase) $|k_1k_2; k_1 + k_2 - 2, k_1 + k_2 - 2\rangle$. Repeating procedure all the standard vectors and the corresponding Clebsch-Gordan coefficients are derived.

It is important to stress out that though the decomposition (4.50) involves non-correlated bases differing by indices λ_i, the Clebsch-Gordan coefficient do not depend on them, giving that the total standard basis is:

$$|k_1k_2; km; \lambda_1\lambda_2\rangle = \sum_{\lambda_1,\lambda_2} \sum_{m_1=-k_1}^{k_1} \sum_{m_2=-k_2}^{k_2} \langle k_1, m_1; k_2, m_2 \,|\, k_1k_2; km\rangle \, |k_1, m_1, \lambda_1; k_2, m_2, \lambda_2\rangle \,. \tag{4.54}$$

Note that $\{K_1^2, K_{13}, A_1, K_2^2, K_{23}, A_2\}$ is CSCO with non-correlated basis as common eigenbases, while CSCO for the standard basis (4.54) is $\{K_1^2, A_1, K_2^2, A_2, K^2, K_3\}$.

Wigner-Eckart theorem and selection rules

In the state space with defined angular momentum $\boldsymbol{K}$ there is also a representation $D(SU(2))$. Standard basis $|km\lambda\rangle$, defined by equations (4.14) corresponds to the decomposition of the space to the irreducible subspaces of SU(2): $\mathcal{S} = \oplus_{k\lambda}\mathcal{V}_\lambda^{(k)}$, meaning that (4.14) are manifestation of the general requirement for standard basis $D(U)\,|km\lambda\rangle = \sum_{m'=-k}^{k} D^{(k)}_{m'm}(U)\,|km'\lambda\rangle$.

However, the transformations of SU(2) act also on the operators in $\mathcal{S}$ according to the rule: $\hat{\hat{U}}A = D(U)AD^\dagger(U)$. This generates a representation in the operator space, and the standard basis among the operators can be defined:

$$\hat{\hat{U}}A^{(k)}_{m\lambda} = \sum_{m'=-k}^{k} D^{(k)}_{m'm}(U)A^{(k)}_{m'\lambda}. \tag{4.55}$$

Any set

$$\mathsf{A}^{(k)} = \{A^{(k)}_m \,|\, m = -k, \dots, k\}, \tag{4.56}$$

satisfying (4.55) is called irreducible tensor (operator), and operators $A^{(k)}_m$ are its *components*. Further on subscript λ is omitted, as well as the concrete set of observables (e.g. electrical field, coordinates, etc.) which are dealt with. Again, using Lie algebra, i.e. angular momentum, (4.55) becomes:

$$[K_3, A^{(k)}_m] = m\hbar A^{(k)}_m, \quad [K_\pm, A^{(k)}_m] = \sqrt{(k \mp m)(k \pm m + 1)}\hbar A^{(k)}_{m\pm1}. \tag{4.57}$$

In this way symmetry adapted bases are imposed both in the space of operators and in the state space. It is useful to find the matrix elements of the irreducible tensors between standard basis elements. In this context the whole operator space basis is not looked for;

only a particular irreducible tensor corresponding to considered physical quantities (e.g. operators of momenta, electrical or magnetic field) are of interest, and subscript λ is not used any more (in fact it is contained in the symbol of the quantity, $\boldsymbol{p}$, $\boldsymbol{E}$, $\boldsymbol{H}$, etc.).

Theorem 4.3 ***(Wigner-Eckart theorem)*** *Matrix element of an irreducible tensor component $A^{(k_2)}_{m_2}$ between two standard vectors is factorized to Clebsch-Gordan coefficient (independent on λ_1 and λ) and the reduced matrix element independent on m, m_1 and m_2:*

$$\langle km\lambda|\, A^{(k_2)}_{m_2}\, |k_1m_1\lambda_1\rangle = \langle k_1m_1; k_2m_2\,|\, k_1k_2km\rangle\, \langle k\lambda|\,|A^{(k_2)}|\,|k_1\lambda_1\rangle\,. \qquad (4.58)$$

Note that the first term, Clebsch-Gordan coefficient, is strictly symmetry based and *a priori* known, while physical context is given only by the reduced matrix element, which is constant for the whole multiplets of m, m_1 and m_2. Therefore it can be calculated (or experimentally found) for one particular choice of m's, and then used in all of the $(2k+1)(2k_1+1)(2k_2+1)$ cases with fixed k, k_1 and k_2.

This conclusion is very important, since the matrix elements are basic ingredient in the calculations of the transition probabilities, i.e. in study of any physical process (Section 6.4); they appear even in the perturbative approach to Schrödinger equation (Section 6.1). Due to the Wigner-Eckart theorem, one can *a priori* find selection rules, a list of the pairs of the states between which the transition in the studied process is not possible: whenever the Clebsch-Gordan coefficient vanishes the transition between states $|km\lambda\rangle\, A^{(k_2)}_{m_2}$ and $|k_1m_1\lambda_1\rangle$ cannot be induced by the perturbation $A^{(k_2)}_{m_2}$. Some of these rules are easily seen; for example if k is less than $|k_1-k_2|$ or greater than k_1+k_2, or if $m\neq m_1+m_2$, the transition is forbidden.

Addition of more than two angular momenta

In the case that there are more than two angular momenta the procedure of their addition is the same as for two, only the question of the order of addition appears.

For the three angular momenta two possibilities are usually considered: $\boldsymbol{K} = (\boldsymbol{K}_1 + \boldsymbol{K}_2) + \boldsymbol{K}_3 = \boldsymbol{K}_1 + (\boldsymbol{K}_2 + \boldsymbol{K}_3)$. The difference is not in the irreducible components, but in the corresponding standard bases. In the first case the standard basis is common eigenbasis for CSCO $\{K_1^2, K_2^2, K_3^2, K_{12}^2, K^2, K_3, A_1, A_2, A_3\}$, while in the second case CSCO is $\{K_1^2, K_2^2, K_3^2, K_{23}^2, K^2, K_3, A_1, A_2, A_3\}$; here $\boldsymbol{K}_{ij} = \boldsymbol{K}_i + \boldsymbol{K}_j$.

Considering only irreducible subspaces, i.e. the product $\mathcal{V}^{(k_1)}\otimes\mathcal{V}^{(k_2)}\otimes\mathcal{V}^{(k_3)}$, this results in two standard bases $|(k_1k_2)k_3k_{12}km\rangle$ and $|k_1(k_2k_3)k_{23}km\rangle$. These bases are related by the unitary matrix of $6j$-symbols.

Analogously, for addition of four angular momenta we define $9j$-symbols, etc.

Superselection rule of total momentum

Note that each particle has spin corresponding to an irreducible angular momentum, i.e. being $s = 0, \frac{1}{2}, 1, \dots$, and that orbital momentum is always integer. Therefore, according to the theorem on the addition of angular momenta, total angular momentum of a particle is either half integer or integer. Further, the same rule for addition gives that total momentum of many particles is again either integer or half integer. This is known as the:

Theorem 4.4 *(Superselection rule of momentum) State combining half integer and integer angular momentum cannot be realized in the nature.*

To distinguish between states with integer and half integer angular momentum it is useful to introduce superselection observable $R = R_{2\pi\boldsymbol{u}} = \mathrm{e}^{-\frac{i}{\hbar}2\pi\boldsymbol{u}\cdot\boldsymbol{K}}$. This observable acts as identity on the vectors corresponding to integer k, and reverses the vectors with half integer k. Thus, if there was a state space combing integer and half integer total angular momentum, the corresponding vectors would span the eigensubspaces of R for the eigenvalues 1 and -1. However, the superselection principle asserts that in each state space only one eigenvalue is realized.

In particular, the space of the operators (superspace) is always with integer angular momentum.

4.6.2 Examples and applications

Addition of two spins $s = 1/2$

In this case the dimension of $\mathcal{S}$ is four. While noncorrelated basis is $|s_1 m_{s_1}; s_2 m_{s_2}\rangle$ for $m_{s_i} = \pm 1/2$, or in short $|{+}{+}\rangle = (1,0,0,0)^T$, $|{+}{-}\rangle = (0,1,0,0)^T$, $|{-}{+}\rangle = (0,0,1,0)^T$, $|{-}{-}\rangle = (0,0,0,1)^T$. On the other hand, according to Theorem 4.2 we get $\mathcal{S} = \mathcal{V}^{(0)} + \mathcal{V}^{(1)}$, with standard basis $|\frac{1}{2}\frac{1}{2}00\rangle$, $|\frac{1}{2}\frac{1}{2}1,-1\rangle$, $|\frac{1}{2}\frac{1}{2}10\rangle$, $|\frac{1}{2}\frac{1}{2}11\rangle$. While the maximal weight vectors is $|\frac{1}{2}\frac{1}{2}11\rangle = |{+}{+}\rangle$, all other vectors are found with help of $S_- = \hbar \begin{pmatrix} 0&0&0&0 \\ 1&0&0&0 \\ 1&0&0&0 \\ 0&1&1&0 \end{pmatrix}$:

$$|\tfrac{1}{2}\tfrac{1}{2}10\rangle = (|{+}{-}\rangle + |{-}{+}\rangle)/\sqrt{2}, \quad |\tfrac{1}{2}\tfrac{1}{2}1-1\rangle = |{-}{-}\rangle\,, \quad |\tfrac{1}{2}\tfrac{1}{2}00\rangle = (|{+}{-}\rangle - |{-}{+}\rangle)/\sqrt{2}.$$

Spin harmonics of electron

It is known that orbital angular momentum is integer, and contains all integer irreducible components: $\mathcal{S}_O = \oplus_{l=0}^{\infty}\mathcal{V}^{(l)}$ Taking spin of electron into account, one gets addition in the form $\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{s}$ and the total space $\mathcal{S} = \mathcal{S}_O \otimes \mathcal{S}_s = \oplus_{l=0}^{\infty}\mathcal{V}^{(l)} \otimes \mathcal{V}^{(\frac{1}{2})}$ is decomposed according to (4.51) into irreducible subspaces:

$$\mathcal{S} = \mathcal{V}_{l=0}^{(j=\frac{1}{2})} \oplus \oplus_{l=1}^{\infty}\left(\mathcal{V}_l^{(j=l-\frac{1}{2})} \oplus \mathcal{V}_l^{(j=l+\frac{1}{2})}\right) = \oplus_{j=\frac{1}{2}}^{\infty}\left(\mathcal{V}_{l=j-\frac{1}{2}}^{(j)} \oplus \mathcal{V}_{l=j+\frac{1}{2}}^{(j)}\right). \tag{4.59}$$

Hence, each half integer total angular momentum occurs twice, due to coupling of the spin to the orbital momentum $l = j - \frac{1}{2}$ and to the orbital momentum $l = j + \frac{1}{2}$. In each of the irreducible subspaces $\mathcal{V}_l^{(j)}$ the standard basis $|l\frac{1}{2}; jm_j\rangle$ $(j = l \pm 1/2)$ is with help of the Clebsch-Gordan coefficients expressed as the combination of the product of the basis vectors (Fig. 4.4, right):

$$|l\frac{1}{2}; jm_j\rangle = \sum_{m=-l}^{l}\sum_{m_s=-\frac{1}{2}}^{\frac{1}{2}} \langle lm; \frac{1}{2}m_s \,|\, l\frac{1}{2}; jm_j\rangle \,|lm; \frac{1}{2}m_s\rangle\,.$$

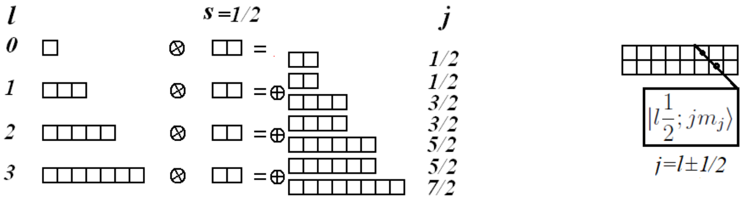

Figure 4.4: Total angular momentum of electron.

As only the Clebsch-Gordan coefficients satisfying $m_j = m + m_s$ may not vanish, taking into account that m_s takes only values $\pm\frac{1}{2}$, the sum (for $l > 0$) reduces to:

$$|l\frac{1}{2};jm_j\rangle = c_+\,|lm_j - \frac{1}{2};\frac{1}{2}\frac{1}{2}\rangle + c_-\,|lm_j + \frac{1}{2};\frac{1}{2},-\frac{1}{2}\rangle\,, \quad c_\pm = \langle lm_j \mp \frac{1}{2};\frac{1}{2},\pm\frac{1}{2}\,|\,l\frac{1}{2};jm_j\rangle\,. \tag{4.60}$$

The coordinate representation (only angular coordinates are considered, and in spin space vectors $|\frac{1}{2},\frac{1}{2}\rangle = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $|\frac{1}{2},-\frac{1}{2}\rangle = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$) becomes:

$$\langle\theta\varphi\,|\,l\frac{1}{2};jm_j\rangle = \begin{pmatrix} c_+Y_l^{m_j-\frac{1}{2}}(\theta,\varphi) \\ c_-Y_l^{m_j+\frac{1}{2}}(\theta,\varphi)\end{pmatrix}. \tag{4.61}$$

This spinor function is called spin (spherical) harmonic.

4.6.3 $L - S$ and $j - j$ coupling

The electronic system in the atom with Z electrons is described by Hamiltonian:

$$H = \sum_{i=1}^{Z}(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i} + v_i^{jj}) + \sum_{i<j}\frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}. \tag{4.62}$$

In the single-particle part besides usual kinetic and Coulomb terms there is spin-orbit interaction $v_i^{jj} = \boldsymbol{l}_i \cdot \boldsymbol{s}_i f(r_i)$, the term commuting with the components of $\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{s}$ (and with $\boldsymbol{l}^2$ and $\boldsymbol{s}^2$). Intuitively it appears as the magnetic dipole of an electron (proportional to its spin) moves in the electric field of the nucleus. Given form can be justified by the perturbation theory within relativistic approach. Depending on the order of the perturbation theory, the form of the function $f(r)$ can be specified (e.g. $f(r) = \frac{m^2c^2}{r}\frac{\mathrm{d}}{\mathrm{d}r}\frac{-Ze^2}{r}$), and for the present purpose it is important only that it is spherically symmetric.

The last term includes many-electron problem, and we assumed that only two-particle interactions exist. Such terms, related to many-body systems, prevent exact solution of the problem, and this is probably the most important difficulty in the contemporary physics. Usual approximate approach is to extract some average single electron potential, i.e. to write the last term in the form

$$\sum_{i<j}\frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} = \sum_i W(r_i) + \sum_{i<j} V_{ij}^{\mathrm{rez}}. \tag{4.63}$$

The residual potential $V^{LS}=\sum_{i<j}V_{ij}^{\mathrm{rez}}$ is the remaining part of the interaction which has not been included by the averaged potential W (more detailed analysis of this approach will be given in the Section 6.2.2). It is assumed to be spin independent.

Adding the averaged potential W to the Coulomb electron-nucleus term, we write Hamiltonian as the sum of the single-particle central Hamiltonian H_0, two-particle term V^{LS}, and spin-orbit (still single particle) term V^{jj}:

$$H = H_0 + V^{LS} + V^{jj}. \tag{4.64}$$

Of course, the total Hamiltonian commutes with the total angular momentum $\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$. However, it is assumed that many-body term depends only on the orbital variables, and commutes both with total orbital and spin angular momenta $\boldsymbol{L}$ and $\boldsymbol{S}$. On the other hand, since coupling electronic orbital and spin degrees of freedom of each electron, V^{jj} commutes with $\boldsymbol{j}$. Therefore, further analysis depends on the relation between spin-orbit and many-body part. It turns out that many-body term is weakly dependent on the atomic number, while spin-orbit term is proportional to it. Therefore, V^{LS} is the dominant for the light atoms, while V^{jj} overwhelms in the case in the second half of periodic system.

Consequently, as the first perturbation to H_0 for light atoms we take V^{LS} term, getting Russel-Saunders or LS-coupling Hamiltonian $H^{LS}=H_0+V^{LS}$; only afterwards the last term V^{jj} is included as the perturbation. The corresponding decomposition of the total space is

$$\mathcal{S}=\mathcal{S}_O\otimes\mathcal{S}_S=(\mathcal{S}_{O1}\otimes\cdots\otimes\mathcal{S}_{OZ})\otimes(\mathcal{S}_{S1}\otimes\cdots\otimes\mathcal{S}_{SZ}),$$

with momenta $\boldsymbol{L}=\boldsymbol{l}_1+\dots\boldsymbol{l}_Z$ and $\boldsymbol{S}=\boldsymbol{s}_1+\dots\boldsymbol{s}_Z$ acting in $\mathcal{S}_O$ and $\mathcal{S}_S$. There is an eigen-basis $|LM;SM_S;\lambda\rangle$ for H^{LS} being standard both for $\boldsymbol{L}$ and $\boldsymbol{S}$, with $(2L+1)(2S+1)$ degenerate eigenspaces. When we add V^{jj}, the perturbation theory predicts that each of these levels $E_{LS\lambda}$ will split: for each $J=|L-S|,\dots,L+S$ an sublevel $E_{LSJ\lambda}$ appears.

In the case of heavy atoms V^{jj} interaction prevails, and the procedure is reversed. Now we start with the jj-coupling, i.e. with $H_{jj}=H_0+V^{jj}$. As this describes non-interacting particles in the external field, for each electron the eigen-basis $|nljm_j\rangle$ can be found, and then the total eigenbasis can be adapted to the set of the compatible observables $J^2, J_3, j_1^2, j_2^2, \dots$, corresponding to decomposition :

$$\mathcal{S}=\mathcal{S}_1\otimes\cdots\otimes\mathcal{S}_Z,\qquad \mathcal{S}_i=\mathcal{S}_{Oi}\otimes\mathcal{S}_{Si}.$$

In each space $\mathcal{S}_i$ acts total electronic angular momentum $\boldsymbol{j}_i=\boldsymbol{l}_i+\boldsymbol{s}_i$. The degeneracy of the eigenenergies depends on Z, and it is high in general. When V^{LS} is included, only J^2 and J_3 remain conserved, and this leads to the splitting of the obtained levels to the $2J+1$ degenerate ones.

Periodic table of elements

Classification of the states of the hydrogen atom and derived rules for addition of the angular momentum give insight to the systematization of the atoms with more electrons according to angular momenta (orbital, spin and total). To this end we anticipate that

Table 4.1: Beginning of the periodic system. Horizontal lines denote fulfilled subshells, while the shells are separated by double lines; both have all angular momenta zero. Quantum numbers n, l and m_s are given only for the electrons out of the last filled (sub)shell. For He there was another possible total spin $S = 1$, which was rejected by Pauli principle. On the other hand, $J = 1/2$ is preferred for B by energy with respect to another possibility $J = 3/2$.

Z	Symbol	n	l	m_s	L	S	J	${}^{2S+1}L_J$	Configuration
1	H	1	0	$\pm\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	${}^2S_{\frac{1}{2}}$	$(1s)$
2	He	1,1	0,0	$\frac{1}{2}, -\frac{1}{2}$	0	0	0	${}^1S_0$	$(1s)^2$
3	Li	2	0	$\pm\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	${}^2S_{\frac{1}{2}}$	(He)$(2s)$
4	Be	2,2	0,0	$\frac{1}{2}, -\frac{1}{2}$	0	0	0	${}^1S_0$	(He)$(2s)^2$
5	B	2	1	$\pm\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	${}^2P_{\frac{1}{2}}$	(He)$(2s)^2(2p)$

- -

state of many-electron system is antisymmetric (i.e. changes sign when any two of electrons are permuted), and consequently Pauli principle (to be derived in the next Chapter) forbidding two electrons to occupy the same state.

At first we make inventory of the involved angular momenta and introduce conventional notation. Each electron has orbital, spin and total angular, denoted as $\boldsymbol{l}$, $\boldsymbol{s}$ and $\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{s}$. For the complete electronic system total orbital, spin and angular momentum, $\boldsymbol{L} = \sum_i \boldsymbol{l}_i$, $\boldsymbol{S} = \sum_i \boldsymbol{s}_i$ and $\boldsymbol{J} = \sum_i \boldsymbol{j}_i$ are defined. For the states with orbital momentum quantum number l equal to 0, 1, 2, 3, etc. we use symbols s, p d, f etc. (and analogously S, P, D, F,…for L).

For each main quantum number $n = 1, 2, \ldots$, possible orbital momenta are $l = 0, \ldots, n-1$; in other words, *shell* n contains *subshells* nl with these values of l. Each of them contains $2(2l+1)$ states with $m = 0, \pm 1, \ldots, \pm l$ and $m_s = \pm\frac{1}{2}$. Starting with known eigenstates and eigenenergies, we begin to fill the states with electrons, starting from the lowest levels.

Thus, for $n = 1$, only $l = m = 0$ is allowed, and the ground state of hydrogen atom is $|n = 1, l = 0, m = 0; s = 1/2, m_s\rangle$, with total momenta $L = l = 0$, $S = s = 1/2$, $J = j = 1/2$. Spectroscopic notation of these angular momenta and electronic configuration is ${}^{2S+1}L_J = {}^2S_{1/2}$ and $(nl) = (1s)$. For the next atoms the higher energy levels are occupied, and angular momenta are given in the table (note that LS-coupling is assumed, in agreement with previous conclusions). Generally, for pairs of electrons the orbital part of state is allowed to be the same, and then the spin part is antisymmetric.

Note that this algorithm is not valid for all elements. For example d-shell is open before $3p$ is filled, etc.

Exercise 4.6: Calculate within the first order perturbation theory the energy of the ground state of the helium atom, neglecting spin-orbit interaction.

Chapter 5

Identical Particles

5.1 Quantum Formalism

Modern physics has no general solution to the many-body problems, despite of their obvious importance. Challenge of finding the method to solve the many-body problem, reinforced by high significance of understanding of complex systems, keeps causing lot of efforts which results in production of extremely sophisticated studies, both conceptually and mathematically, with a number of (more or less successful) approximations. However, even the foundation of the adequate formalism is a valuable result, based on probably the longest sequence of deductions in science (thus, the most impressive and beautiful): from the very idea of identical particles, formalized as the permutational symmetry, powerful enough to discover notions of fermions and bosons, Fock's space, and second quantization with creation and annihilation of particles. Therewith, geometrization of the permutations singles out two-dimensional configuration space, allowing for other types of identical (quasi)-particles, experimental realization of which is one of the main focuses of nowadays physics.

5.1.1 Permutational Indistinguishability and Symmetrization

According to the conclusion of the subsection 1.3.2, the state space of N particles is the product of their single particle states. Thus, in the case of N identical particles, when all the particles are described by the same single particle state space $\mathcal{S}$, the total state space is *a priori* constructed as $\mathcal{S}^N = \mathcal{S} \otimes \cdots \otimes \mathcal{S}$. On the other hand, the intuitive notion of the indistinguishability of the particles, means that permutations among them is not observable: there is no measurement distinguishing between initial and permuted state. As for the pure states this means that the initial and the permuted state vectors are collinear, differing at most by a phase. However, in the space $\mathcal{S}^N$ there are vectors not obeying this requirement. For example, consider non-correlated states, with p-th particle being in the state $|\psi_p\rangle$. The composite state is simply denoted as

$$|\psi_1, \dots, \psi_N\rangle = |\psi_1\rangle_1 \otimes \cdots \otimes |\psi_N\rangle_N \,. \tag{5.1}$$

Note that, by convention, the p-th particle is in the state at the p-th place, and therefore the second index (outside ket) may be omitted. Any permutation π of the particles transform this state into

$$\Delta(\pi)\left|\psi_1,\dots,\psi_N\right\rangle \overset{\text{def}}{=} \cdots\otimes\left|\psi_1\right\rangle_{\pi 1}\otimes\cdots\otimes\left|\psi_N\right\rangle_{\pi N}\otimes\cdots = \left|\psi_{\pi^{-1}1},\dots,\psi_{\pi^{-1}N}\right\rangle. \tag{5.2}$$

If permutation π acts nontrivially acts on the particles being in mutually orthogonal single particle states, the initial and permuted states are orthogonal onto each other,thus non collinear.

This shows that the construction of the state space of identical particles must be further suited to the intuitive requirement of indistinguishability. To this end we emphasize that (5.2) effectively introduces the operators $\Delta(\pi)$ in $\mathcal{S}^N$ associated to the permutation π. Indeed, any single-particle basis $|i\rangle$ ($i=1,\dots,|\mathcal{S}|$) of $\mathcal{S}$ induces the non-correlated basis $|i_1,\dots,i_N\rangle$ ($i_p=1,\dots,|\mathcal{S}|$), and (5.2) is applicable for all the vectors of the basis. This suffices to define linear operators $\Delta(\pi)$. In addition, all the basis vectors are by $\Delta(\pi)$ mapped to the vectors of the same basis, i.e. $\Delta(\pi)$ permutes basis vectors; this means that $\Delta(\pi)$ is a unitary representation (orthonormal single particle basis is assumed) of the symmetric group $\boldsymbol{S}_N$ in $\mathcal{S}^N$.

The basic intuitive requirement of indistinguishability of the initial $|\Psi\rangle$ and the permuted state $\Delta(\pi)\,|\Psi\rangle$ (conveniently, for N particle state capital letters are used) can be now precisely expressed as:

$$\Delta(\pi)\,|\Psi\rangle = \mathrm{e}^{\mathrm{i}\varphi(\pi,\Psi)}\,|\Psi\rangle\,, \tag{5.3}$$

with the phase depending both on permutation and state. In other words, each physically allowable state $|\Psi\rangle$ spans one dimensional subspace of $\mathcal{S}^N$, in which the representation $\Delta(\boldsymbol{S}_N)$ is reduced to the phases $\mathrm{e}^{\mathrm{i}\varphi(\pi,\Psi)}$. Of course, these sub-representations are one dimensional, thus irreducible ones. However, the permutation group $\boldsymbol{S}_N$ for $N>1$ has exactly two one-dimensional irreducible representations (Exercise 5.1), symmetric (identical) $D^{(+)}$ and antisymmetric (alternating) $D^{(-)}$:

$$D^{(\pm)}(\pi) = (\pm)^{\tilde{\pi}}. \tag{5.4}$$

Here, $\tilde{\pi}$ denotes the parity of the permutation π: each permutation is a product of the transpositions, and although this decomposition is not unique, as well as the number of the involved transpositions, this number is uniquely either even, $\tilde{\pi}=0$, or odd, $\tilde{\pi}=1$. Particularly, $D^{(\pm)}(\tau_{pp'})=\pm 1$ for the transposition $\tau_{pp'}$ of any two particles p and p'.

***Exercise 5.1**:* For permutational group $\boldsymbol{S}_N$ with $N>1$ show that:

(a) even permutations form halving subgroup of $\boldsymbol{S}_N$ (alternating subgroup $\boldsymbol{A}_N$);

(b) the only one-dimensional irreducible representations are those defined in (5.4).

***Exercise 5.2**:* Show that $\{\tau_1,\dots,\tau_{N-1}\}$ with $\tau_i \overset{\text{def}}{=} \tau_{i,i+1}$, are generators of $\boldsymbol{S}_N$. Then prove the generator relations:

$$\tau_i^2 = e, \tag{5.5a}$$

$$\tau_i\tau_j = \tau_j\tau_i, \qquad |i-j|\geq 2, \tag{5.5b}$$

$$\tau_i\tau_{i+1}\tau_i = \tau_{i+1}\tau_i\tau_{i+1}, \qquad 1\leq i\leq N-1. \tag{5.5c}$$

The above analysis enlightens that the permutation group $\boldsymbol{S}_N$ is in $\mathcal{S}^N$ represented by $\Delta(\boldsymbol{S}_N)$, which is reducible. Its irreducible components correspond to multiple irreducible subspaces. Physically allowable states form multiple irreducible subspaces $\mathcal{S}^N_\pm$ of the two one-dimensional representations $D^{(\pm)}(\boldsymbol{S}_N)$. These two subspaces, symmetric and antisymmetric $\mathcal{S}^N_\pm$ of $\mathcal{S}^N$ are singled out as the ranges of the corresponding group projectors, symmetrizer and antisymmetrizer:

$$P^{(\pm)} = \frac{1}{N!}\sum_\pi (\pm)^{\tilde{\pi}}\Delta(\pi). \tag{5.6}$$

The only physically allowable states are within these subspaces, and for $|\Psi\rangle \in \mathcal{S}^N_+$ the phase in (5.3) is $\mathrm{e}^{\mathrm{i}\varphi(\pi,\Psi)} = 1$, while for $|\Psi\rangle \in \mathcal{S}^N_-$ one gets $\mathrm{e}^{\mathrm{i}\varphi(\pi,\Psi)} = (-)^{\tilde{\pi}}$. Note that a nontrivial linear combination $\alpha_+|+\rangle + \alpha_-|-\rangle$ of $|\pm\rangle \in \mathcal{S}^N_\pm$, when mapped by a transposition operator, is not collinear to the initial vector: $\Delta(\tau)(\alpha_+|+\rangle + \alpha_-|-\rangle) = \alpha_+|+\rangle - \alpha_-|-\rangle$. Hence, to retain the superposition principle, the system of identical particles must be described within only one of these two subspaces.

These conclusions are summarized in:

POSTULATE 7. – IDENTICAL PARTICLES

The state space of a system of N identical particles with the single particle state space $\mathcal{S}$ is either symmetric $\mathcal{S}^N_+$ or antisymmetric $\mathcal{S}^N_-$ subspace of the direct product space $\mathcal{S}^N$. The choice depends only on the type of the particles: bosons, particles with integer spin, have the symmetrized N particle states, while for the fermions, the particles with half-integer spin, the states are antisymmetrized.

The last part of the Postulate, relating spin to the type of the transformation of the particle states under permutations can be proved within relativistic quantum mechanics, using the locality of interaction [16]. This has deep influence not only on the kinematical, but also on the dynamical properties of the particles: as a consequence of the different permutational symmetry for the systems of bosons and fermions the different statistics, Bose-Einstein or Fermi-Dirac, are applied.

Anyons — Abelian and non-Abelian braiding statistics

In the permutational groups $\boldsymbol{S}_N$ a transposition τ is an involution: $\tau^2 = 1$, i.e. $\tau = \tau^{-1}$. Thus the one-dimensional representation can be either unit or alternating representation, which is the main ingredient in revealing bosons and fermions as the only two types of identical particles. Considering geometric realization of a transposition τ, interchange of two particles is essentially a rotation β of one particle around the other, such that β^2 restores the original mutual positions, i.e. $\beta^2 = 1$. As generally rotations are directed transformations, it is obvious that β and β^{-1} are oppositely directed. Therefore, while β^2 realizes the initial position of the particles by the full rotation (e.g. circle made by the first particle around the second one), $\beta\beta^{-1}$ ends in the same final position passing the half of the circle twice but in the opposite directions (i.e. making half of the circle and returning back along the same way). Clearly, in the latter case, the entire path traveled by a particle can be homotopically squeezed to a point. As for the former case, the path is homotopic to a point in three-dimensional (Fig. 5.1), but not in two-dimensional geometry (Fig. 5.2). Hence, in three-dimensional geometry, all conceivable evolutions of quantum states which are caused by permutations of identical particles are indistinguishable, which is clearly explicated by the Postulate 7 and equation (5.3).

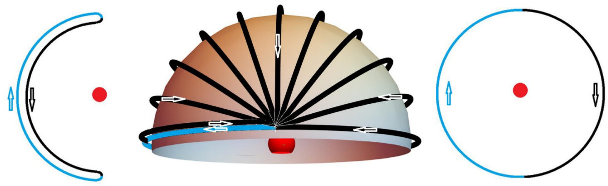

Figure 5.1: System of two identical particles and two oriented paths of one of the particles around the other (left and right figures). In three-dimensional geometry, the homotopy of the two paths can be easily established by a continual series of the interpolating paths (central figure): topological equivalence of the paths.

However, in two-dimensional geometry, the two paths are topologically different, as the other particle is unavoidable obstacle to the homotopical squeezing of the path corresponding to β^2. Consequently, β^2 is not the identity, and moreover, order of the transformation of the particles interchange β is infinite. In general, in the N-particles case, transformations β_i $(i = 1, \ldots, N-1)$ interchanging successive particles (i and $i+1$) generate braid group[2] $\boldsymbol{B}_N$ satisfying the generator relations:

$$\beta_i^2 \neq 1 \tag{5.7a}$$

$$\beta_i\beta_j = \beta_j\beta_i, \qquad |i-j| \geq 2 \tag{5.7b}$$

$$\beta_i\beta_{i+1}\beta_i = \beta_{i+1}\beta_i\beta_{i+1}, \qquad 1 \leq i \leq N-1. \tag{5.7c}$$

Note that the difference with respect to $\boldsymbol{S}_N$ (5.5) makes the first relation only: (5.7a) *vs.* (5.5a). Namely, while the transpositions τ_i (of the permutational group $\boldsymbol{S}_N$) satisfy $\tau_i^2 = 1$, the generators of the braid group $\boldsymbol{B}_N$ do not square to the identity: $\beta_i^2 \neq 1$. Consequently, whilst the order of the permutational group is finite $|\boldsymbol{S}_N| = N!$, the braid group $\boldsymbol{B}_N$ is of infinite order.

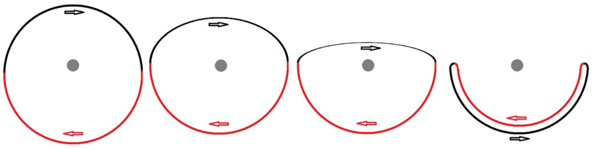

Figure 5.2: System of two identical particles and oriented paths of one of the particles around the other. In two-dimensional configuration the path at the right cannot be homotopically mapped to the paths left to it as it is unavoidable that one of the interpolating paths goes through the other particle: topological inequivalence of the paths.

Braid group can be seen as a set of topological classes of the trajectories of N-particles in the (2+1)-dimensional space-time, Fig.5.3 (note that slice at any moment intersects

$N = 3$ trajectories, i.e. that the conservation of the number of particles is assumed: duplication backwards would correspond to the particle annihilation and creation at an intermediate instant).

For physics remarkable consequence is that the paths corresponding to different braid group transformations are distinguishable, and may correspond to different quantum states. Hence, quantum mechanics does not impose restriction to one-dimensional irreducible representations of braid groups. All this reveals that classification of identical particles in two-dimensional geometry goes beyond the distinction of bosons and fermions, only: there exist also the other types of particles, conveniently named anyons.

Yet another distinction of the two-dimensional geometry is the existence of the nontrivial topological classes of trajectories of distinguishable particles which makes a subgroup of the braid group, made of "pure braids", elements returning each of the particles to its initial position. This is in complete opposition to the topological triviality of the corresponding trajectories in three-dimensional geometry.

The order of the action of the braiding operations onto the quantum states is irrelevant for abelian anyons; they are associated to the one-dimensional (unitary) representations, i.e. the braiding is characterized just by a phase factor, $\Delta(\beta)\,|\Psi\rangle = \mathrm{e}^{\mathrm{i}\varphi}\,|\Psi\rangle$. Although the braiding can, in principle, be described by any value of the exchange phase (hence the name anyon), a stability consideration[10] requires the phase be a rational multiple of 2π.

However, when the braiding act by a multidimensional irreducible representation and the braiding operations do not commute, the particles are called non-Abelian anyons. One of the significant consequences of the non-Abelian anyon statistics is that in systems with a degenerate many-body state, particle exchange leads to the state change and that the final state depends on the order of the exchange operations. This is of particular importance for fault-tolerant topological quantum computation through quasi-particle (Majorana zero modes) braiding[15].

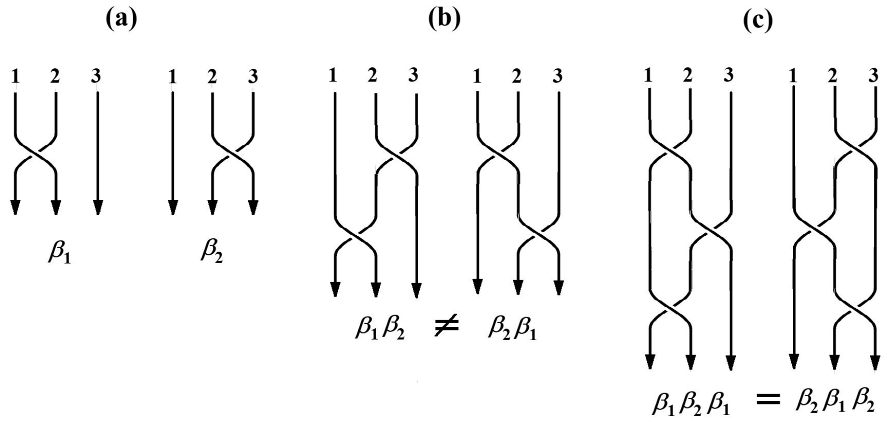

Figure 5.3: Braid group $\boldsymbol{B}_3$: (a) generators; (b& c) generator relations (5.7). By convention, when one trajectory passes over/under the other one, the corresponding exchange is clockwise/counterclockwise. The arrows indicate direction of time. Multiplication of elements is concatenation of the corresponding paths.

Natural question to ask is whether anyons as a novel type of identical particles exist in nature or can be built in laboratory. It turns out that electrons, for instance, when confined to two spatial dimensions still retain Fermi-Dirac statistics but that under specific conditions in systems of (many) electrons *quasi-particle excitations* which obey anyon statistics emerge (e.g., formerly theoretically predicted and recently reported experimental observation of the *fractional quantum-Hall state* anyon braiding statistics[14]).

Finally, note that in one-dimensional configurational space, the quantum statistics is not well defined, since the exchange is not possible without one particle going through the other.

5.1.2 Structure of the state space — occupation numbers

The permutational symmetry of the physical states gives further insight to the structure of the boson and fermion state spaces. At first, it should be noticed that (5.2) shows that the operators $\Delta(\boldsymbol{S}_N)$ in $\mathcal{S}^N$ permute the non-correlated basis vectors $|i_1,\dots,i_N\rangle$. Indeed, their action consists in the rearrangement of the single particle states quantum numbers $i_1,\dots,i_N$. This inspires introduction of the occupation numbers: for the N particle state $|i_1,\dots,i_N\rangle$ one counts the number $n_i(|i_1,\dots,i_N\rangle)$ of the particles occupying the single particle state $|i\rangle$. Thus, each basis state defines $|\mathcal{S}|$ occupation numbers $\boldsymbol{n}(|i_1,\dots,i_N\rangle) = (n_1,\dots,n_{|\mathcal{S}|})$, with the obvious relation $\sum_{i=1}^{|\mathcal{S}|} n_i = N$. More than one basis vector may have the same array $\boldsymbol{n}$. Precisely, all the permutations acting (by the representation $\Delta(\boldsymbol{S}_N)$) on the vector $|i_1,\dots,i_N\rangle$ give the vectors with the same occupation numbers. Hence, the non-correlated basis is by the action of $\boldsymbol{S}_N$ partitioned into the orbits, each of them characterized by the unique $\boldsymbol{n}$. In other words, each $|\mathcal{S}|$-tuple $\boldsymbol{n}$ of the non-negative integers n_s such that $\sum_{s=1}^{|\mathcal{S}|} n_s = N$ defines an orbit of the non-correlated basis.

The permutations interchanging particles being in the same single particle states leave the basis vector invariant (e.g. $\Delta(\tau_{12})\,|i_1, i_2 = i_1, i_3,\dots,i_N\rangle = |i_1, i_2 = i_1, i_3,\dots,i_N\rangle$).‘ Taking from each orbit $\boldsymbol{n}$ a representative basis vector $|i_1,\dots,i_N\rangle$, it is easy to find its stabilizer: it is the group $\boldsymbol{S}_{\boldsymbol{n}} = \boldsymbol{S}_{n_1}\otimes\cdots\otimes\boldsymbol{S}_{n_{|\mathcal{S}|}}$; $\boldsymbol{S}_{n_s}$ contains only the permutations interchanging the particles in the same state $|i_s\rangle$ and leaving invariant all other particles (therefore different factors permute disjoint subsets of the particles and commute mutually). Obviously, the order of the stabilizer is $|\boldsymbol{S}_{\boldsymbol{n}}| = \boldsymbol{n}! \overset{\text{def}}{=} \prod_{s=1}^{|\mathcal{S}|} n_s!$, showing that the number of the basis vectors in the orbit $\boldsymbol{n}$ of $\boldsymbol{S}_N$ is

$$|\boldsymbol{n}| = \frac{|\boldsymbol{S}_N|}{|\boldsymbol{S}_{\boldsymbol{n}}|} = \frac{N!}{n_1!\cdots n_{|\mathcal{S}|}!} = \frac{N!}{\boldsymbol{n}!}. \tag{5.8}$$

Now, recall that each orbit $\boldsymbol{n}$ is a subset of the non-correlated basis. Therefore, it spans an $|\boldsymbol{n}|$-dimensional subspace $\mathcal{S}_{\boldsymbol{n}}$ of $\mathcal{S}^N$, and

$$\mathcal{S}^N = \oplus_{\boldsymbol{n}}\mathcal{S}_{\boldsymbol{n}}. \tag{5.9}$$

As $\mathcal{S}^N$ is spanned by the whole non-correlated basis, one finds $\sum|\boldsymbol{n}| = |\mathcal{S}|^N$. Each of these subspaces $\boldsymbol{S}_{\boldsymbol{n}}$ is invariant under the permutation operators $\Delta(\pi)$, showing that (5.9) is a decomposition of the total space onto the invariant subspaces of the representation $\Delta(\boldsymbol{S}_N)$. Hence, this representation reduces in the occupation number spaces $\mathcal{S}_{\boldsymbol{n}}$ onto the sub-representations $\Delta_{\boldsymbol{n}}(\boldsymbol{S}_N)$ geting the reduced form $\Delta(\boldsymbol{S}_N) = \oplus_{\boldsymbol{n}}\Delta_{\boldsymbol{n}}(\boldsymbol{S}_N)$.

Finally, we perform the (anti)symmetrization, as requested by the Postulate VII. The (anti)symmetric space $\mathcal{S}^N_{\pm}$ is found as the range of the (anti)symmetrizer (5.6). The invariance of $\mathcal{S}^N_{\boldsymbol{n}}$, resulting in the decomposition of $\Delta(\boldsymbol{S}_N)$, implies that the group projectors also reduce in the occupation number subspaces:

$$P^{(\pm)} = \oplus_{\boldsymbol{n}} P^{(\pm)}_{\boldsymbol{n}}, \quad P^{(\pm)}_{\boldsymbol{n}} = \frac{1}{N!}\sum_{\pi}(\pm)^{\tilde{\pi}}\Delta_{\boldsymbol{n}}(\pi). \tag{5.10}$$

Due to the orthogonality of the subspaces $\mathcal{S}^N_{\boldsymbol{n}}$, the operators $P^{(\pm)}_{\boldsymbol{n}}$ are projectors, and their ranges $\mathcal{S}^N_{\boldsymbol{n}\pm}$ are orthogonally summed into the range of $P^{(\pm)}$. The first conclusion emerges: the (anti)symmetrized space $\mathcal{S}^N_{\pm}$ is orthogonal sum of the (anti)symmetrized subspaces $\mathcal{S}^N_{\boldsymbol{n}\pm}$ of the occupation number spaces $\mathcal{S}^N_{\boldsymbol{n}}$.

To find these subspaces for each orbit the orbit representative is chosen so to be the first one when the orbit elements are sorted in the increasing order with respect to the single particle state indices. Simply, it is the non-correlated vector $|\boldsymbol{n}\rangle = |i_1 \leq i_2 \leq \cdots \leq i_n\rangle$ for which from $i_p \leq i_{p'}$ follows that $p \leq p'$, i.e. the single particle state indices do not decrease from left to right (Exercise 5.3). According to the general group theoretical theorem, $\boldsymbol{S}_N$ acts such that exactly $\boldsymbol{n}!$ permutations (a coset of the stabilizer $\boldsymbol{S}_{\boldsymbol{n}}$ with a representative ζ_p, $p = 1, \dots, N!/\boldsymbol{n}!$) map $|\boldsymbol{n}\rangle$ to each of the vectors $|\boldsymbol{n}, p\rangle = \Delta(\zeta_p)|\boldsymbol{n}\rangle$ of the orbit $\boldsymbol{n}$; a vector $|\boldsymbol{n}, p\rangle$ of the orbit is fixed by $\boldsymbol{n}!$ permutations (conjugated subgroup $\zeta_p\boldsymbol{S}_{\boldsymbol{n}}\zeta_p^{-1}$). Therefore, assuming that the single particle basis is orthonormalized (then the used non-correlated vectors of $\mathcal{S}^N$ are also orthonormalized), $\mathrm{Tr}P^{(\pm)}_{\boldsymbol{n}} = \frac{1}{N!}\sum_p \langle\boldsymbol{n}, p|\sum_\pi(\pm)^{\tilde{\pi}}\Delta_{\boldsymbol{n}}(\pi)|\boldsymbol{n}, p\rangle$. As only the elements $\pi = \zeta_p\phi\zeta_p^{-1}$ ($\phi \in \boldsymbol{S}_{\boldsymbol{n}}$) of the stabilizer of $|\boldsymbol{n}, p\rangle$ contribute to the sum over p:

$$\mathrm{Tr}P^{(\pm)}_{\boldsymbol{n}} = \frac{1}{N!}\sum_p \langle\boldsymbol{n}, p|\sum_\phi(\pm)^{\tilde{\phi}}\Delta_{\boldsymbol{n}}(\zeta_p\phi\zeta_p^{-1})|\boldsymbol{n}, p\rangle = \frac{1}{N!}\sum_p \langle\boldsymbol{n}|\sum_\phi(\pm)^{\tilde{\phi}}\Delta_{\boldsymbol{n}}(\phi)|\boldsymbol{n}\rangle\,.$$

Noticing that $N!/\boldsymbol{n}!$ summands in p are p-independent, while the remaining stabilizer elements fix $|\boldsymbol{n}\rangle$, one gets:

$$\mathrm{Tr}P^{(\pm)}_{\boldsymbol{n}} = \frac{1}{\boldsymbol{n}!}\sum_\phi(\pm)^{\tilde{\phi}}. \tag{5.11}$$

Exercise 5.3: For $|\mathcal{S}| = 7$ and $N = 6$ analyze the state $|5, 2, 5, 3, 2, 5\rangle$: find its occupation numbers, stabilizer, standard orbit representative and order of orbit.

Exercise 5.4: Show that $\Delta_{\boldsymbol{n}}(\boldsymbol{S}_N)$ is a ground representation (induced from the unit representation of stabilizer). Using this derive (5.11) as the frequencies $f^{\pm}_{\Delta_{\boldsymbol{n}}}$ of the irreducible representations $D^{(\pm)}(\boldsymbol{S}_N)$ in $\Delta_{\boldsymbol{n}}(\boldsymbol{S}_N)$.

Bosons

In the case of bosons, the summands in (5.11) are parity independent and equal to 1, i.e. $\mathrm{Tr}P^{(+)}_{\boldsymbol{n}} = 1$. Thus, for bosons, each occupation number subspace $\mathcal{S}^N_{\boldsymbol{n}}$ contains exactly

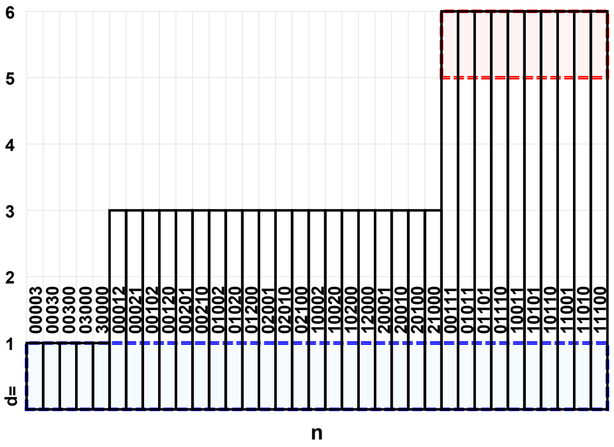

Figure 5.4: Space $\mathcal{S}^N$. Spaces $\mathcal{S}^N_{\boldsymbol{n}}$ ($N = 3$, $|\mathcal{S}| = 5$) are represented by vertical bold rectangles assigned by the occupation numbers; their dimensions are indicated by the light grid labeled on the left. Spaces $\mathcal{S}^N_{\pm}$ are dash-framed, blue and red shaded.

one dimensional symmetric subspace spanned by the normalized symmetric vector $|\boldsymbol{n}^+\rangle$ obtained by the action of the symmetrizer onto $|\boldsymbol{n}\rangle = |i_1 \leq \dots \leq i_N\rangle$:

$$|\boldsymbol{n}^+\rangle = \sqrt{\frac{N!}{\boldsymbol{n}!}} P^{(+)}_{\boldsymbol{n}} |\boldsymbol{n}\rangle = \sqrt{\frac{1}{\boldsymbol{n}!N!}} \sum_{\pi} |i_{\pi^{-1}1}, \dots, i_{\pi^{-1}N}\rangle . \tag{5.12}$$

The normalization factor is easily found, taking into account that the whole orbit with $|\boldsymbol{n}|$ orthonormal vectors appear after the action of the permutations, and each of them exactly $\boldsymbol{n}!$ times.

Thus $\mathcal{S}^N_+$ is orthogonal sum of the one-dimensional symmetrized occupation number subspaces:

$$\mathcal{S}^N_+ = \sum_{\boldsymbol{n}} \mathcal{S}^N_{\boldsymbol{n}+}, \quad \mathcal{S}^N_{\boldsymbol{n}+} = \mathrm{Span}(|\boldsymbol{n}^+\rangle). \tag{5.13}$$

It is obvious that each occupation number $\boldsymbol{n}$ uniquely defines the symmetrized vector $|\boldsymbol{n}^+\rangle$, and in this sense the occupation numbers single out a basis in the bosonic space.

Fermions

Analogously to bosons, the space of fermions is found with help of $P^{(-)}_{\boldsymbol{n}}$. However, for n_s greater than one, the factor $\boldsymbol{S}_{n_s}$ contains halves of even and odd permutations; then, due to Exercise 5.1, one half of the permutations in each factor $\boldsymbol{S}_{n_i}$, and thus in the whole stabilizer $\boldsymbol{S}_{\boldsymbol{n}}$, is even and the rest is odd: they cancel in (5.11), i.e. $\mathrm{Tr}\, P^{(-)}_{\boldsymbol{n}} = 0$, meaning that the subspaces $\mathcal{S}_{\boldsymbol{n}}$ with $\boldsymbol{n}! > 1$ contain no antisymmetric physical states. Only when all $n_s = 0, 1$ the trivial stabilizer (the identity element solely since $\boldsymbol{n}! = 1$) $\boldsymbol{S}_{\boldsymbol{n}}$ is obtained; the sum in (5.11) is reduced to the single identical (even) permutation, and again $P^{(-)}_{\boldsymbol{n}}$ one-dimensional projector ($\mathrm{Tr}\, P^{(-)}_{\boldsymbol{n}} = 1$). This one-dimensional subspace is spanned by the antisymmetric vector

$$|\boldsymbol{n}^-\rangle = \sqrt{N!}P^{(-)}_{\boldsymbol{n}}|\boldsymbol{n}\rangle = \sqrt{\frac{1}{N!}}\sum_{\pi}(-)^{\tilde{\pi}}|i_{\pi^{-1}1} < \cdots < i_{\pi^{-1}N}\rangle\,. \tag{5.14}$$

Likewise the case of bosons, $\mathcal{S}^N_-$ is orthogonal sum of the one-dimensional anti-symmetrized occupation number subspaces, but the sum includes only the terms for which $\boldsymbol{n}! = 1$ (i.e. with the components $n_s = 0, 1$):

$$\mathcal{S}^N_- = \sum_{\boldsymbol{n}\ (n_s<2)} \mathcal{S}^N_{\boldsymbol{n}^-}, \quad \mathcal{S}^N_{\boldsymbol{n}^-} = \mathrm{Span}(|\boldsymbol{n}^-\rangle). \tag{5.15}$$

Again, occupation number $\boldsymbol{n}$ satisfying $n_s < 2$ uniquely defines the antisymmetrized vector $|\boldsymbol{n}^-\rangle$, and such occupation numbers provide a basis in the fermionic space.

5.1.3 States

Pure states

It is shown that the substantial difference between bosons and fermions is the type of symmetrization. This has far reaching consequences on the behavior of the two types of particles.

The most important one is the already emphasized constraint $n_i = 0, 1$ to the fermionic occupation numbers. In fact, this is the famous Pauli exclusion principle: two fermions cannot be in the same state. The most frequent example is that at most two electrons may have the same orbital wave function $\psi(\boldsymbol{r})$: then their states $|\psi, m_s\rangle$ differ in the spin space by the quantum number $m_s = \frac{1}{2}$ and $m_s = -\frac{1}{2}$. This fact is the corner stone of the periodic system of elements.

As there is no analogous requirement for bosons, the different statistics are obeyed by the ensembles of the identical fermions and bosons. This can be easily understood even on the simplest systems. For instance, consider the system of two particles in the states a and b. Classically, there are four states of the pair: a_1a_2, a_1b_2, b_1a_1 and b_1b_2 (the subscript enumerates the particles). So, assuming equal probabilities of these states, one can find both particles in the state a with probability 1/4, the same for the state b, and the probability to find different states of the particles is 1/2. However, in the quantum mechanical treatment, these states are combined into the three symmetric states $|aa\rangle$, $|ab\rangle + |ba\rangle$, $|bb\rangle$ and one antisymmetric state $|ab\rangle - |ba\rangle$. Hence, in the case of bosons, the probability is 1/3 to find both the particles in the state $|a\rangle$, the same for $|b\rangle$, and the same to find the particles in different states. Finally, one finds the fermions in the different states with certainty. While for the fermions the coincident states are completely forbidden, for bosons these are even more probable than for the classical particles. In fact, the total energy of non-interacting bosons is the sum of all the energies of the occupied single particle states; the ground state of such a system is obviously one with all the particles being in the same single particle state of the minimal energy. This effect is called Bose condensation. It is important in understanding of various phenomena, and probably its most famous manifestations are superfluidity and superconductivity.

The same example may illustrate another significant difference between the two types of particles. The antisymmetrized vector (5.14) is correlated for any $\boldsymbol{n}$, while when all bosons are in the same state $|j\rangle$, i.e. when $n_i = N\delta_{ij}$, state $|\boldsymbol{n}^+\rangle$ is non-correlated. In other words, fermions are always correlated: Pauli principle imposes, through the antisymmetrization, at least the exchange or Pauli correlation. Of course, general physical states are nontrivial superposition of the basis vectors $|\boldsymbol{n}^-\rangle$, being in a sense more correlated than $|\boldsymbol{n}^-\rangle$. In the same sense, Pauli correlation is the minimal one; it is purely kinematical, not related to dynamics of the system. On the other hand, depending on the particular dynamics (including external fields and interaction between particles), the non-correlated states may be even preferred for bosons.

Despite these, and other physical differences between bosons and fermions, many related expressions are formally analogous, and to a far extent it is possible to treat them in a unified way. Basically this is due to the common form

$$|\boldsymbol{n}^\pm\rangle = \sqrt{\frac{1}{\boldsymbol{n}!N!}}\sum_\pi (\pm)^{\tilde{\pi}}\Delta(\pi)\,|i_1 \leq \cdots \leq i_N\rangle \tag{5.16}$$

of the occupation number basis vectors; here $\boldsymbol{n}! = 1$ is to be taken for fermions to get (5.14), while otherwise (5.12) is obtained. The normalization factor comes from the fact that among $N!$ summands in (5.16) there are $N!/\boldsymbol{n}!$ mutually orthogonal ones; each of them appears $\boldsymbol{n}!$ times and such a class contributes by $\boldsymbol{n}!$ to the norm $N!\boldsymbol{n}!$.

It is useful to adopt convention that if any occupation number is negative the corresponding vector vanishes: $n_i < 0$ implies $|n^\pm\rangle = 0$. Further, relation $\boldsymbol{n} < \boldsymbol{n}'$ means that $n_s \leq n'_s$ for each s, i.e. that the single particle states involved in $|\boldsymbol{n}^\pm\rangle$ make a subset of those appearing in $|\boldsymbol{n}'^\pm\rangle$; this implies that $|\boldsymbol{n}^\pm\rangle$ and $|\boldsymbol{n}'^\pm\rangle$ describe sets of $N < N'$ particles, such that $|\boldsymbol{n}'^\pm\rangle$ is obtained from $|\boldsymbol{n}^\pm\rangle$ by adding $N' - N$ particles while retaining all the existing states.

As for the fermions, there is a well known convenient way to perform the antisymmetrization (5.16). In fact, this vector is obtained automatically if one calculates Slater determinant $N!^{-1/2}(|i_p\rangle_q)$, using the direct product as the multiplication of the matrix elements, which also means that the factors in the product do not commute. The expansion rules for determinants provide the correct signs of the factors. However, if the signs are always taken to be positive, the definition of permanent is obtained, allowing one to write (5.16) in the form:

$$|\boldsymbol{n}^\pm\rangle = \sqrt{\frac{1}{\boldsymbol{n}!N!}}\begin{vmatrix} |i_1\rangle & |i_1\rangle & \cdots & |i_1\rangle \\ |i_2\rangle & |i_2\rangle & \cdots & |i_2\rangle \\ \cdots & \cdots & \cdots & \cdots \\ |i_N\rangle & |i_N\rangle & \cdots & |i_N\rangle \end{vmatrix}_\pm . \tag{5.17}$$

Here, the sign distinguishes between permanent (+) and determinant (-). Alas, the algebra of permanents is much more cumbersome than that of determinants, making this notation widely used for fermions only. An illustration is easy proof (Exercise 5.5) that the $|n^-\rangle$ is biuniquelly related to the subspace $\mathcal{S}_{\boldsymbol{n}} = \mathrm{Span}(|i_1\rangle, \dots, |i_N\rangle)$ spanned by the occupied single particle states (this subspace of $\mathcal{S}$ should be distinguished from $\mathcal{S}^N_{\boldsymbol{n}}$ of the N-particle space). Further important application is the coordinate representation of the states $|\boldsymbol{n}^\pm\rangle$.

Indeed, with $\psi_i(\boldsymbol{r}_p) = \langle \boldsymbol{r}_p \,|\, i\rangle$, from (5.17) it is obvious that the corresponding wave function $\Psi_{\boldsymbol{n}^\pm}(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N) = \langle \boldsymbol{r}_1,\dots,\boldsymbol{r}_N \,|\, \boldsymbol{n}^\pm\rangle$ is:

$$\Psi_{\boldsymbol{n}^\pm}(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N) = \sqrt{\frac{1}{\boldsymbol{n}!N!}} \begin{vmatrix} \psi_{i_1}(\boldsymbol{r}_1) & \psi_{i_1}(\boldsymbol{r}_2) & \cdots & \psi_{i_1}(\boldsymbol{r}_N) \\ \psi_{i_2}(\boldsymbol{r}_1) & \psi_{i_2}(\boldsymbol{r}_2) & \cdots & \psi_{i_2}(\boldsymbol{r}_N) \\ \cdots & \cdots & \cdots & \cdots \\ \psi_{i_N}(\boldsymbol{r}_1) & \psi_{i_N}(\boldsymbol{r}_2) & \cdots & \psi_{i_N}(\boldsymbol{r}_N) \end{vmatrix}_\pm . \tag{5.18}$$

***Exercise 5.5**:* Show that arbitrary basis $|j_p\rangle$ of $\mathrm{Span}(|i_p\rangle \,|p = 1,\dots,|i_N\rangle)$ gives by (5.16) the same physical state as the basis $|i\rangle$ (i.e. the vector differing from $|\boldsymbol{n}^-\rangle$ at most by a phase).

Being the basis of the state space $\mathcal{S}^N_\pm$, the states $|\boldsymbol{n}^\pm\rangle$ do not exhaust it. The superposition principle inevitably includes their linear combinations. Although some of these linear combinations are again purely kinematically correlated states, made with the different bases of the single particle space, there are (anti)symmetrized states which cannot be written in the form (5.16) for any possible choice of the single particle states. Usually the eigenstates of the real N-particle systems are just such states, and then the additional correlation in comparison to the exchange one is naturally called dynamical correlation.

Mixed States

The methods of statistical physics are quite important in many body problems. They are based on the mixed states, and we consider some of their elementary properties. A statistical operator in $\mathcal{S}^N_\pm$ can be expanded in the form

$$\rho = \sum_{\boldsymbol{n},\boldsymbol{n}'} c_{\boldsymbol{n}\boldsymbol{n}'} \,|\boldsymbol{n}^\pm\rangle\, \langle \boldsymbol{n}'^\pm| \,. \tag{5.19}$$

From this form it is obvious that these operators commute with permutations. Indeed, $\Delta^{-1}(\pi)\rho\Delta(\pi) = \rho$, both for bosons and fermions since the parity of π appears twice. Such operators are called symmetric operators, and only these are relevant (by the same argument as for ρ) for the systems of identical particles.

Subsystems

The reduced states of the clusters of $L \leq N$ particles are suitable to describe measurements on these natural subsystems. Applying previously introduced techniques (Subsection 1.3.3) to the subsystems of L and the remaining $N - L$ particles, the reduced state of the particles $p_1,\dots,p_L$ is the partial trace over other particles:

$$\rho_{p_1\cdots p_L} = \mathrm{Tr}_{\,\hat{p}_1\cdots\hat{p}_L}\rho. \tag{5.20}$$

Here, hat denotes the particles omitted in the trace. That this is a statistical operator in $\mathcal{S}^L$ follows from the general theory. However, it should be checked out wheter it is really a state of L identical particles, i.e. a statistical operator in $\mathcal{S}^L_\pm$ having the form (5.19) with L-particle symmetrized states $|\boldsymbol{n}^\pm_L\rangle$ and $|\boldsymbol{n}'^\pm_L\rangle$ ($\sum_i n_{Li} = \sum_i n'_{Li} = L$). To this end we use the fact that the partial scalar product $\langle i_{L+1},\dots,i_N \,|\, \boldsymbol{n}^\pm\rangle$ is up to the constant just the (anti)symmetrized state of the first L particles (if the set of the quantum numbers in the bra is not a subset of those in the ket, the partial scalar product vanishes). Basically, since $\boldsymbol{n}^\pm$ has terms with the quantum numbers of the occupation number $\boldsymbol{n}$ distributed over particles by all possible permutations, then the partial scalar product with $\langle i_{L+1},\dots,i_N|$ annihilates the terms differing in quantum numbers on the last $N-L$ positions; in the remaining terms, the complementing quantum numbers of $|\boldsymbol{n}^\pm\rangle$ are retained only, but distributed as before over all of the first $N-L$ positions (details are in Exercise 5.7). Finally, as in the partial trace appear only such terms, the reduced state is really a state of L identical particles.

Exercise 5.6: Given bosonic state $|\boldsymbol{n}=(210)\rangle$, find $\langle 1_1\,|\,\boldsymbol{n}\rangle$, $\langle 1_2\,|\,\boldsymbol{n}\rangle$, $\langle 1_1 1_2\,|\,\boldsymbol{n}\rangle$, $\langle 1_1, 2_2\,|\,\boldsymbol{n}\rangle$ and $\langle 2_1, 1_2\,|\,\boldsymbol{n}\rangle$. Compare the corresponding clusters' states. Discuss fermionic analogs.

Exercise 5.7: Prove that the partial scalar product of the N-particle state $|\boldsymbol{n}^\pm\rangle$ with non-correlated last $N-L$ particles vector $|j_{L+1},\dots,j_N\rangle$ with the occupation numbers $\boldsymbol{n}_{N-L}$ is L-particle state:

$$\langle j_{L+1},\dots,j_N\,|\,\boldsymbol{n}^\pm\rangle = \sqrt{\frac{\boldsymbol{n}!L!}{\boldsymbol{n}_L!N!}}(\pm)^{\tilde{\sigma}}\,|\boldsymbol{n}_L^\pm\rangle\,, \quad \boldsymbol{n}_L = \boldsymbol{n}-\boldsymbol{n}_{N-L} \tag{5.21a}$$

(recall that vector vanishes if any of n_{Li} is negative); σ is a permutation (unique in the case of fermions and irrelevant in the case of bosons) such that

$$\Delta(\sigma)\,|i_1\leq\cdots\leq i_N\rangle = |i_{\sigma^{-1}1}\leq\cdots\leq i_{\sigma^{-1}L}, j_{L+1},\dots,j_N\rangle\,.$$

Analogously, for the first L-particle vector $|j_1,\dots,j_L\rangle$ show

$$\langle j_1,\dots,j_L\,|\,\boldsymbol{n}^\pm\rangle = \sqrt{\frac{\boldsymbol{n}!(N-L)!}{\boldsymbol{n}_{N-L}!N!}}(\pm)^{\tilde{\sigma}}\,|\boldsymbol{n}_{N-L}^\pm\rangle\,, \quad \boldsymbol{n}_{N-L} = \boldsymbol{n}-\boldsymbol{n}(j_1,\dots,j_L), \tag{5.21b}$$

with σ defined by $\Delta(\sigma)\,|i_1\leq\cdots\leq i_N\rangle = |j_1,\dots,j_L, i_{\sigma^{-1}(L+1)}\leq\cdots\leq i_{\sigma^{-1}N}\rangle$.
For $L=1$ derive:

$$\langle j\,|\,\boldsymbol{n}^\pm\rangle = \sqrt{\frac{n_j}{N}}(\pm)^{l_j}\,|\boldsymbol{n}_{N-1}^\pm\rangle\,, \quad n_{N-1,i} = n_i-\delta_{ij},\ l_j = \sum_{i<j} n_i. \tag{5.21c}$$

Exercise 5.8: Show the equality:

$$\langle \boldsymbol{n}_{N-L}^\pm\,|\,\boldsymbol{n}^\pm\rangle = \sqrt{\binom{\boldsymbol{n}}{\boldsymbol{n}_L}\Big/\binom{N}{L}}(\pm)^{\tilde{\lambda}}\,|\boldsymbol{n}_L^\pm\rangle\,, \tag{5.22}$$

where $\binom{\boldsymbol{n}}{\boldsymbol{n}_L} \overset{\text{def}}{=} \boldsymbol{n}!/\boldsymbol{n}_L!\boldsymbol{n}_{N-L}!$ with $\boldsymbol{n} = \boldsymbol{n}_L+\boldsymbol{n}_{N-L}$.

It is intuitively clear that due to the indistinguishability of particles, i.e. due to the symmetrization of the state ρ, each of the L-particle subsystem states should be described by essentially the same reduced statistical operator:

$$\rho_{1,\dots,L} = \rho_{p_1,\dots,p_L}, \quad \forall\{p_1,\dots,p_L\}. \tag{5.23}$$

To justify this formally, the matrix elements of the reduced operator $\rho_{1,\dots,L}$ are found:

$$\langle i_1,\dots,i_L|\,\rho_{1,\dots,L}\,|i'_1,\dots,i'_L\rangle = \sum_{j_{L+1},\dots,j_N} \langle i_1,\dots,i_L,j_{L+1},\dots,j_N|\,\rho\,|i'_1,\dots,i'_L,j_{L+1},\dots,j_N\rangle\,.$$

As ρ is symmetric, it can be replaced by $\Delta^{-1}(\pi)\rho\Delta(\pi)$ for any π. Choosing π to map $\{1,\dots,L\}$ into $\{p_1,\dots,p_L\}$, it changes the particles omitted in the bras and kets of partial trace to $p_1,\dots,p_L$, which gives $\rho_{p_1,\dots,p_L}$. On the other hand, commutation of ρ with the permutations of the first L particles, shows that $\rho_{1,\dots,L}$ is symmetric operator.

It has been discussed that the pure states are among the mixed ones singled out as the one dimensional projectors. As the purely kinematically correlated states $|\boldsymbol{n}^{\pm}\rangle$ span the state space, it is useful to find the reduced L-particle states of $\rho = |\boldsymbol{n}^{\pm}\rangle\langle\boldsymbol{n}^{\pm}|$. Using (5.21a), one straightforwardly shows (Exercise 5.9):

$$\rho_{1,\dots,L} = \binom{N}{L}^{-1} \sum_{\boldsymbol{n}_L \leq \boldsymbol{n}} \binom{\boldsymbol{n}}{\boldsymbol{n}_L} |\boldsymbol{n}_L^{\pm}\rangle\langle\boldsymbol{n}_L^{\pm}|, \quad \text{with } \binom{\boldsymbol{n}}{\boldsymbol{n}_L} \overset{\text{def}}{=} \prod_{s=1}^{|\mathcal{S}|} \binom{n_s}{n_{Ls}}. \tag{5.24a}$$

An important consequence is the form of the single particle reduced states ρ_1. The possible states $|\boldsymbol{n}_1\rangle$ are in fact the single particle states: for each s with $n_s > 0$ there is $\boldsymbol{n}_1$ such that $\boldsymbol{n}_{1i} = \delta_{is}$, making $|\boldsymbol{n}_1\rangle = |s\rangle$. Then $\boldsymbol{n}_1! = 1$, and (5.24a) becomes:

$$\rho_1 = N^{-1} \sum_{s=1}^{|\mathcal{S}|} n_s |s\rangle\langle s|. \tag{5.24b}$$

Exercise 5.9: Prove (5.24a)

Exercise 5.10: Find 4-particle reduced state for 6-particle pure state $|\boldsymbol{n}^+ = (2,3,1)^+\rangle$.

Particularly interesting, and for further purposes important, is the structure of these reduced states for fermions. As the occupation number $\boldsymbol{n}$ satisfies $n_i = 0, 1$, (5.24b) becomes $\rho_1 = P_{\boldsymbol{n}}/N$, where $P_{\boldsymbol{n}}$ is projector onto the previously introduced single-particle subspace $\mathcal{S}_{\boldsymbol{n}} = \mathrm{Span}(|i_1\rangle,\dots,|i_N\rangle)$. Also, the general form of the reduced states is simple (Exercise 5.11):

$$\rho_{1,\dots,L} = \frac{1}{\binom{N}{L}} \underbrace{P_{\boldsymbol{n}} \otimes \cdots \otimes P_{\boldsymbol{n}}}_{L} P_L^{(-)}. \tag{5.25a}$$

The most significant special cases of this result are for $L = 1, 2, N$:

$$\rho_1 = \frac{P_{\boldsymbol{n}}}{N}, \quad \rho_{12} = \frac{2}{N(N-1)} (P_{\boldsymbol{n}} \otimes P_{\boldsymbol{n}}) P_2^{(-)}, \quad \rho = \underbrace{P_{\boldsymbol{n}} \otimes \cdots \otimes P_{\boldsymbol{n}}}_{N} P^{(-)}. \tag{5.25b}$$

The last expression explains how the subspace of the single particle occupied state uniquelly determines N-fermion state, as it is anticipated earlier (Exercise 5.5).

Exercise 5.11: Prove (5.25).

Therefore, each N-dimensional single-particle subspace completely determines one antisymmetric physical state; different choices of bases give Slater determinants differing by phase factors. *Vice versa*, each N particle Slater determinant uniquely determines N-dimensional subspace in the single particle state $\mathcal{S}$. In this sense, Slater determinants and single particle subspaces are biuniquelly related.

5.1.4 Operators

Symmetric Operators

Indistinguishability of the particles reflects itself in the operators. Likewise statistical operators, any N-particle observable must be invariant under the particle permutations, i.e. $\Delta^{-1}(\pi)A\Delta(\pi) = A$. Such symmetric operators in $\mathcal{S}^N$ are singled out as the range of the group $\boldsymbol{S}_N$ super-projector

$$\hat{\hat{P}}^{(+)}A = A, \quad \hat{\hat{P}}^{(+)} \stackrel{\text{def}}{=} \frac{1}{N!}\sum_{\pi} \Delta^{-1}(\pi)\ldots\Delta(\pi). \tag{5.26}$$

In other words, the physical observables commute with the permutations, and therefore transform according to the identical representation of $\boldsymbol{S}_N$.

Note that a symmetric operator acts within the physical state spaces (of identical particles), i.e. it maps physical states into the physical states, which is also a necessary physical property. Indeed, this is actually a requirement that $\mathcal{S}^N_\pm$ is an invariant subspace for any physically relevant N-particle operator A, or equivalently, that A commutes with the symmetrizer and the antisymmetrizer,

$$[(\hat{\hat{P}}^{(+)}A), P^{(\pm)}] = 0, \tag{5.27}$$

which is easily verified (Exercise 5.12). Thus, although the symmetric operators are frequently formally written in $\mathcal{S}^N$, their physically relevant part is the reduced operator $AP^{(\pm)} = P^{(\pm)}AP^{(\pm)}$ in $\mathcal{S}^N_\pm$, having the form (5.19). Hence, only symmetric operators will be further considered.

***Exercise 5.12**:* Prove (5.27).

Symmetric N-particle operators are usually constructed from one- or two-particle ones. More generally, an *L-particle operator* $A_{p_1\cdots p_L}$ essentially couples L particles $p_1,\ldots,p_L$: it acts nontrivially in the product of the spaces of these particles, while it is identity in other particle spaces. Using suitable permutations, one straightforwardly obtains equivalent operators for the other L particle clusters: $\Delta^{-1}(\pi)A_{p_1\cdots p_L}\Delta(\pi) = A_{p_{\pi 1}\cdots p_{\pi L}}$. In this way, by the symmetrizer (5.26), the symmetric (in N particle permutations) L-particle operator $A^{[L]} = c\hat{\hat{P}}^{(+)}A_{p_1\cdots p_L}$ is constructed, where the constant c is defined according to the particular physical content of the operator (there is no universal normalization convention as for the states). When $A_{p_1\cdots p_L}$ is symmetric itself, i.e. when it commutes with L particle permutations one gets:

$$A^{[L]} = c\sum_{p_1<\cdots<p_L}^{N} A_{p_1,\ldots,p_L}. \tag{5.28}$$

Otherwise the sum includes various orderings of the same particles; these essentially give the symmetrized (by the super-symmetrizer (5.26)) operator. Thus, any symmetric operator A in $\mathcal{S}^N$ may be expanded into the series $A = \sum_{L=0}^{N} A^{[L]}$.

Typical examples of the symmetric operators are additive quantities like energy (then $c = 1$ in (5.28)): an one-particle operator is just the sum $A^{[1]} = \sum_p A_p$, while $A^{[2]} =$

$\sum_{p_1<p_2} A_{p_1p_2}$ is a two particle operator. Another type are the multiplicative operators, like many symmetry transformations: if D_1 is a single particle operator, then its L-th tensor power $D_1 \otimes \cdots \otimes D_1$ is a symmetric operator in $\mathcal{S}^L$. Note that (anti)symmetrizers $P^\pm$ are symmetric operators, combining all L-particle symmetric operators (for all $L = 1, ..., N$) corresponding to the permutations effectively permuting L-particles; each of these classes of permutations consists of the subclasses with the same cyclic structure (with $N - L$ unit cycles), each defining a symmetric operator.

As the indistinguishability of the particles is now completely incorporated in the quantum formalism, few examples of the symmetric operators discussed below may indicate its profound and far-reaching consequences.

It has been mentioned that the Pauli exclusion principle allows at most two electrons to be in the same orbital state $|\psi\rangle$; then their total single-particle states are $|\psi, m_s = \frac{1}{2}\rangle$ and $|\psi, m_s = -\frac{1}{2}\rangle$, yielding the two particle state (in shorten notation) $\frac{1}{\sqrt{2}}\,|\psi\rangle\,(|{+}{-}\rangle - |{-}{+}\rangle)$. Note that the total spin $S^2 = \boldsymbol{s}_1^2 + \boldsymbol{s}_2^2$ of this state is $S = 0$. In fact, this is the simplest manifestation of the quite general result on the total angular momentum of the fermions in a closed shell (Exercise 5.13): if $2k+1$ fermions occupy the shell of the angular momentum k (this is the set of states $|Q; km\rangle$, $m = -k, \dots, k$, with fixed all other quantum numbers Q, including k in particular), then their total angular momentum vanishes: $K = 0$. This refers to any angular momentum appearing in the concrete problem. Here we slightly generalize the notion of the shell introduced in the physics of atoms in analogy to the hydrogen like model: besides the main quantum number n there appears the orbital angular momentum l, as well as the spin s. Then, a closed shell is formed by the $2(2l+1)$ electrons filling the states with the fixed n and l. It follows that all the total angular momenta, L, S and J of these electrons vanish.

Exercise 5.13: (Closed shell) Prove that the total angular momentum of the $2l + 1$ fermions in the states $\{|lm\rangle \mid m = -l, \dots, l\}$ vanishes.

Symmetry transformations offer examples of symmetric multiplicative operators. In fact, an arbitrary single particle operator $D = D_1$ generates a series of the multiplicative L-particle operators: for each L the (anti)symmetrized tensor power $[D^{\otimes L}]^\pm = \hat{\hat{P}}^{(+)} \otimes_{p=1}^{L} D_p$. Further, recall that when the number of identical fermions is equal to the dimension of the single-particle state space, $N = |\mathcal{S}|$, the total state space $\mathcal{S}_-^N$ is one-dimensional, meaning that $[D^{\otimes|\mathcal{S}|}]^-$ is an one-dimensional operator. Simultaneously, as the operators $D(g)$ form a representation of the group $\boldsymbol{G}$, then the (anti)symmetrized operators $[D^{\otimes|\mathcal{S}|}]^\pm$ also form representations ((anti)symmetrized powers). In particular, in the space of $|\mathcal{S}|$ fermions, this representation is $\det D(g)$. Therefore, the general conclusion is that if in a single particle space there is a representation $D(\boldsymbol{G})$, in the space $\mathcal{S}_-^{|\mathcal{S}|}$ (the fulfilled fermionic case), with $|D| = |\mathcal{S}|$, acts the representation $\det D(\boldsymbol{G})$. The previous example is a special case of this statement. Considering the rotational group (precisely SU(2)), and its irreducible representations $D^{(k)}(SU(2))$, the closed shell corresponds to the representation $\det D^{(k)}(SU(2))$. As one-dimensional, this must be unit representation $D^{(0)}(SU(2))$ (indeed, the determinant is the basis independent product of the eigenvalues, $\prod_{m=-k}^{k} \mathrm{e}^{\mathrm{i}m\phi} = 1$), with zero total angular momentum.

A generalization of this analysis is elaborated in Exercise 5.14. The second equation of the Exercise concerns the reducible representations in the single-fermion space, corresponding to several filled shells. The third part reflects the group-algebra equality $\otimes_{p=1}^{n} D_p = \mathrm{e}^{\sum_{p=1}^{n} A_p}$ relating some n-particle and single-particle symmetric operators. As for the rotational group, this implies that the fulfilled shells may be ignored in the calculation of the total angular momentum.

Exercise 5.14: Show:

1. $[D^{\otimes|D|}]^{-} = \det D$.
2. If $D = D_1 \oplus D_2 \oplus \cdots \oplus D_K$, then $[D^{\otimes|D|}]^{-} = \prod_{i=1}^{K} \det D_i$.
3. If $D = D_1 \otimes D_2 \otimes \cdots \otimes D_K$, then $[D^{\otimes|D|}]^{-} = \prod_{i=1}^{K} (\det D_i)^{|D|-|D_i|}$.

Exercise 5.15: Let $D = A \oplus B$ in $\mathcal{S} = \mathcal{A} \oplus \mathcal{B}$. Show that in the N-fermion space $\mathcal{S}_{-}^{N}$ the states with fulfilled the first $|\mathcal{A}|$ states is a subspace $\mathcal{S}_{\mathcal{A}-}^{N} = \mathcal{A}_{-}^{|\mathcal{A}|} \bigwedge \mathcal{B}_{-}^{N-|\mathcal{A}|} = \mathcal{B}_{-}^{N-|\mathcal{A}|}$, and that in this subspace $D^{\otimes N}$ is reduced to $[D^{\otimes N}]_{\mathcal{A}}^{-} = \det A \; [B^{\otimes(N-|\mathcal{A}|)}]^{-}$. Here, $\bigwedge$ denotes the antisymmetrized product: $\mathcal{A}_{-}^{|\mathcal{A}|} \bigwedge \mathcal{B}_{-}^{N-|\mathcal{A}|} = P^{(-)} \left(\mathcal{A}_{-}^{|\mathcal{A}|} \otimes \mathcal{B}_{-}^{N-|\mathcal{A}|} \right)$.

Mean values and reduced operators

In analogy to the reduced L-particle statistical operators, the L-particle operators in $\mathcal{S}_{p_1} \otimes \cdots \otimes \mathcal{S}_{p_L}$ are from the N-particle ones constructed by partial traces over the remaining particles: $A_{p_1,\dots,p_L} \overset{\text{def}}{=} \mathrm{Tr}_{\hat{p}_1,\dots,\hat{p}_L} A$. As (5.23) is implied by the fact that ρ is symmetric, for all the physically relevant (i.e. symmetric) operators one gets symmetric L-particle operators:

$$A_{1,\dots,L} = A_{p_1,\dots,p_L}, \quad \forall\{p_1,\dots,p_L\}. \tag{5.29}$$

Mean value of the L-particle symmetric operator $A^{[L]}$ in the mixed state ρ is

$$\langle A^{[L]} \rangle_{\rho} = \mathrm{Tr}\, A^{[L]} \rho = \binom{N}{L} \mathrm{Tr}_{1,\dots,L} A_{1,\dots,L} \rho_{1,\dots,L}. \tag{5.30}$$

To see this it suffices to use (5.28): $\mathrm{Tr}\, A^{[L]} \rho = \sum_{p_1<\dots<p_L}^{N} \mathrm{Tr}_{p_1,\dots,p_L} A_{p_1,\dots,p_L} \mathrm{Tr}_{\hat{p}_1,\dots,\hat{p}_L} \rho$. According to (5.29), all the summands are the same, thus giving the binomial factor. Note that $\mathrm{Tr}_{1,\dots,L}$ is the complete trace of the expression after it, and therefore in the applications it will be denoted only by Tr.

Among numerous consequences of (anti)symmetrization, there is appearance of *exchange terms* in mean values for $L > 1$, besides the *direct term*. For example, for distinguishable particles the mean value in a basis state is

$$\langle i_1,\dots,i_N | A_{12} \otimes \mathbb{1}_{3,\dots,N} | i_1,\dots,i_N \rangle = \langle i_1, i_2 | A_{12} | i_1, i_2 \rangle .$$

For identical particles, $\langle \boldsymbol{n}^{\pm} | A^{[2]} | \boldsymbol{n}^{\pm} \rangle$ contains besides this direct, also the exchange contribution $\langle i_1, i_2 | A_{12} | i_2, i_1 \rangle$.

Exercise 5.16: Calculate $\langle \boldsymbol{n}^{\pm} | A^{[2]} | \boldsymbol{n}^{\pm} \rangle$.

Particularly, the energy is an additive quantity, and, assuming only pairwise interactions, the Hamiltonian of N-particle system is $H = H^{[1]} + H^{[2]}$. Single-particle Hamiltonian is $H^{[1]} = \sum_{p=1}^{N} H_p$, where $H_p = T_p + U_p$. Analogously, pairwise interactions V_{pq}, contribute in the sense of (5.28) ($c = 1$), by $H^{[2]} = \frac{1}{2}\sum_{p_1 \neq p_2} V_{p_1 p_2}$. Therefore, the mean value of the Hamiltonian is found by (5.30):

$$E[\rho] = N \mathrm{Tr}\,_1 H_1 \rho_1 + \binom{N}{2} \mathrm{Tr}\,_{12} V_{12} \rho_{12}. \tag{5.31}$$

Suppose that the occupation numbers are defined for the eigenbasis of H_1 in $\mathcal{S}$, i.e. $H_1 = \sum_i \epsilon_i \left|i\right\rangle \left\langle i\right|$. Then for the purely kinematically correlated states $\rho = \left|\boldsymbol{n}^{\pm}\right\rangle \left\langle \boldsymbol{n}^{\pm}\right|$, one gets by (5.24a):

$$E[\rho] = \sum_s n_s \epsilon_s + \sum_{\boldsymbol{n}_2 < \boldsymbol{n}} \binom{\boldsymbol{n}}{\boldsymbol{n}_2} \mathrm{Tr}\, V_{12} \left|\boldsymbol{n}_2^{\pm}\right\rangle \left\langle \boldsymbol{n}_2^{\pm}\right| \tag{5.32}$$

The first term is the energy of the noninteracting particles, and in some simple dynamical approximation only it is used. The second term has not the form of the single-particle mean, which makes the problem too hard to proceed exactly. The main obstacle is correlation. In fact, even for the single particle term, the general expression is found only for the minimally correlated states. Further, to understand that the main complexity is introduced by correlations, we consider very exceptional case of non-correlated two-particle reduced state: $\rho_{12} = \rho_1' \otimes \rho_1'$ (the factors must be equal as the state is symmetrical). Then the second term in (5.31) is easy to find: $\mathrm{Tr}\,_{12} V_{12} \rho_{12} = \mathrm{Tr}\, \rho_1' \mathrm{Tr}\,_2 V_{12} \rho_1'$. We reveal the mean field approximation: the interaction term is the energy of the first particle in the averaged field $\mathrm{Tr}\,_2 V_{12} \rho_1'$ of the others (being in the states ρ_1'), which reduces the problem to the single-particle one. This observation motivates the most of the approximative approaches to the many-body dynamics.

5.2 Second Quantization

Until now the systems with the fixed number of particles are considered. However, it is well known that in various processes particles are created or annihilated. Although nonrelativistic quantum mechanics cannot give real insight and explanation of such processes, its formalism is sufficiently general to treat such situation. Further, in many nonrelativistic problems, usually within some approximations, the dynamics can be well described in terms of quasi-particles: it turns out that excitations of the systems have so much of the properties of the real particles (e.g. momenta, spin, etc.), that it is not only technically convenient, but heuristically very fruitful, to treat them on equal footing. Typically this occurs in the harmonic approximations, retaining only the Hamiltonian terms quadratic in some generalized coordinates. Being itself an archetypical example, the harmonic oscillator problem underlies all such approaches. Finally, even when the number of particles is fixed, the previously developed formalism is deficient in the sense that most of the calculations are performed in the space $\mathcal{S}^N$, with cumbersome appearance of the permutations, although the physical space is only its (anti)symmetrized part. All these motivate the development of the advanced quantum mechanical formalism, called second quantization, which is indispensable tool in the contemporary many-body physics and quantum field theory (including particularly solid state and elementary particles physics).

5.2.1 Fock space

Although nonrelativistic quantum mechanics cannot give completely realistic insight to the processes of the transitions between particles, e.g. decays or creation of particles, the introduced formalism may treat also these situations. To this end one should firstly realize that the state space $\mathcal{S}^N_\pm$, describing the system with N identical particles, is insufficient framework for the systems with the varying number of particles. Obviously, $\mathcal{S}^N_\pm$ ($N = 0, 1, 2\dots$) spaces are to be summed, to get the adequate state space, called Fock space:

$$\mathcal{F}_\pm = \oplus_N \mathcal{S}^N_\pm. \tag{5.33}$$

Recall that if $\{|i\rangle \mid i = 1, \dots, \mathcal{S}\}$ is a basis in the single particle space, then the occupation number vectors $\boldsymbol{n} = (n_1, \dots, n_{|\mathcal{S}|})$ define the states $|\boldsymbol{n}^\pm\rangle$ uniquely. All possible choices of n_i such that $N = \sum_{i=1}^{|\mathcal{S}|} n_i$ give an orthonormal basis in $\mathcal{S}^N_\pm$. Therefore, using the states $|\boldsymbol{n}^\pm\rangle$ for $N = 0, 1, \dots$ one gets the occupation number orthonormal basis of the Fock space $\mathcal{F}_\pm$, and any vector of this space may be expanded over the basis $|\Psi\rangle_\pm = \sum_{\boldsymbol{n}} c^\pm_{\boldsymbol{n}} |\boldsymbol{n}^\pm\rangle$. Particularly, for $N = 0$, one gets the one dimensional vacuum space $\mathcal{S}^0_\pm$, spanned by the vacuum state $|0\rangle$ (of course, $\boldsymbol{n} = 0$ means that $n_1 = \cdots = n_{|\mathcal{S}|} = 0$). Also, for $N = 1$ various occupation numbers are exhausted by one particle being in a particular state $|i\rangle = |0, \dots, 0, 1, 0 \dots, 0\rangle$ (1 in the i^{th} place), while all other states are unoccupied; thus the initial single particle space is spanned by this basis, i.e. $\mathcal{S}^1_\pm = \mathcal{S}$.

Exercise 5.17: Show that the fermionic Fock's space with the finite dimensional single particle space $\mathcal{S}$ is isomorphic to the $|\mathcal{S}|$-qubit space (distinguishable quasi particles with the two-dimensional spaces).

5.2.2 Creation and annihilation operators

Obviously, the change in the number of particle is manifested as the transition from one of the subspaces $\mathcal{S}^N_\pm$ to the other one $\mathcal{S}^{N'}_\pm$. Such a transition may be considered as a series of changes by one particle. Thus, it suffices to study creation or annihilation of one particle in a particular (single-particle) state.

Let N-particle system being in the state $|\boldsymbol{n}^\pm\rangle = |n_1, n_2, \dots\rangle$, the particle in the state $|i\rangle$ is annihilated. The resulting state is $|\boldsymbol{n}'^\pm\rangle$, describing the system of $N-1$ particles, with $n'_j = n_j$ particles in the states $|j(\neq i)\rangle$ and $n'_i = n_i - 1$ particles in the state $|i\rangle$. This observation is used to define conveniently the annihilation operator a_i in the i^{th} state:

$$a_i |n_1, \dots, n_i, \dots\rangle = (\pm)^{l_i(\boldsymbol{n})} \sqrt{n_i}\, |n_1, \dots, n_i - 1, \dots\rangle\,. \tag{5.34}$$

Here, $l_i(\boldsymbol{n}) = \sum_{j=1}^{i-1} n_j$ is the number of particles being in the states j with $j < i$ (for given $|\boldsymbol{n}^\pm\rangle$), i.e. the number of states on the left of the first i. For bosons $(+)^{l_i(\boldsymbol{n})} = 1$ and may be omitted, but we write it to emphasize the unified treatment. The adjoint operator $a^\dagger_i$ can be found considering the matrix elements $\langle n'_1, n'_2, \dots| a^\dagger_i |n_1, n_2, \dots\rangle$. As it acts on the left as the annihilation operator, all the matrix elements vanish except for $\langle n_1, \dots n_i + 1, \dots |$, i.e.

$$\langle \boldsymbol{n}'^{\pm} |\, a_i^\dagger \,| \boldsymbol{n}^{\pm} \rangle = \sqrt{n_i'} (\pm)^{l_i(\boldsymbol{n}')} \delta_{n_1}^{n_1'} \cdots \delta_{n_i}^{n_i'-1} \cdots = \sqrt{1+n_i} (\pm)^{l_i(\boldsymbol{n}')} \delta_{n_1}^{n_1'} \cdots \delta_{1+n_i}^{n_i'} \cdots$$

Recall that for fermions $n_i = 0, 1$, meaning that for $n_i = 1$ the occupation number on the left should be $n_i' = 2$, and bra automatically vanishes, implying that the whole scalar product is always zero; as the only other case is $n_i = 0$, the root may be conveniently, in the case of fermions, changed to $\sqrt{1 - n_i}$. Finally, one gets the creation operator:

$$a_i^\dagger \,| n_1, \dots, n_i, \dots \rangle = (\pm)^{l_i(\boldsymbol{n})} \sqrt{1 \pm n_i}\, | n_1, \dots, n_i + 1, \dots \rangle \,. \tag{5.35}$$

***Exercise 5.18**:* Show that $a_s \,| \boldsymbol{n}^{\pm} \rangle = \sqrt{N} \, \langle s \,| \boldsymbol{n}^{\pm} \rangle$. Derive (5.35) using this. Generalize the first result to get (5.21b) in the form:

$$\langle j_1, \dots, j_L \,| \boldsymbol{n}^{\pm} \rangle = \sqrt{\frac{(N-L)!}{N!}} a_{j_L} \cdots a_{j_1} \,| \boldsymbol{n}^{\pm} \rangle \,. \tag{5.36}$$

Find the form of this expression independent on the number of particles N in the state $| \boldsymbol{n}^{\pm} \rangle$.

Some important properties of the introduced operators follow directly. Action of the annihilation and creation operators, a_i and $a^\dagger{}_i$, for any i, onto the vacuum state:

$$a_i \,| 0 \rangle = 0, \quad a_i^\dagger \,| 0 \rangle = | i \rangle \,. \tag{5.37}$$

The successive action onto arbitrary state for fermions gives $a_i^2 = a_i^{\dagger^2} = 0$, while for bosons: $a_i^k \,| \boldsymbol{n}^{+} \rangle = \sqrt{\frac{n_i!}{(n_i-k)!}} \,| n_1, \dots, n_i - k, \dots \rangle$ and $a_i^{\dagger^k} \,| \boldsymbol{n}^{+} \rangle = \sqrt{\frac{(n_i+k)!}{n_i!}} \,| n_1, \dots, n_i + k, \dots \rangle$. Obviously, any state can be obtained from the vacuum state by creation operators:

$$| \boldsymbol{n}^{\pm} \rangle = \frac{1}{\sqrt{n_1! \cdots n_{|\mathcal{S}|}!}} a_1^{\dagger^{n_1}} \cdots a_{|\mathcal{S}|}^{\dagger^{n_{|\mathcal{S}|}}} \,| 0 \rangle \,. \tag{5.38}$$

Finally, note that the number of particles in the state $| i \rangle$ is easily obtained by application of occupation number operator $\hat{n}_i \overset{\text{def}}{=} a_i^\dagger a_i$, as its eigenvalue:

$$\hat{n}_i \,| \boldsymbol{n}^{\pm} \rangle = n_i \,| \boldsymbol{n}^{\pm} \rangle \tag{5.39}$$

(for bosons, this directly follows from the definitions, while for fermions the cases of $n_i = 0$ and $n_i = 1$ should be separately discussed). Accordingly, since the states with the same number of particles are comprised in the subspaces $\mathcal{S}_{\pm}^{N}$, these are eigenspaces of the number of particles operator $\hat{N} = \sum_i \hat{n}_i = \sum_i a_i^\dagger a_i$:

$$\hat{N} \,| \boldsymbol{n}^{\pm} \rangle = \sum_i n_i \,| \boldsymbol{n}^{\pm} \rangle = N \,| \boldsymbol{n}^{\pm} \rangle \,. \tag{5.40}$$

The creation and annihilation operators correspond to the arbitrary chosen basis in the single particle space $\mathcal{S}$. If the considered basis is continual, then one gets a continual family of the creation/annihilation operators, called field operators. For example, to the coordinate basis $| x \rangle$ correspond the field operators $a^\dagger(x)$ creating a particle at x: $| x \rangle = a^\dagger(x) \,| 0 \rangle$.

5.2.3 Bosonic and fermionic commutation relations

It is obvious that any transition from $|\boldsymbol{n}^{\pm}\rangle$ to $|\boldsymbol{n}'^{\pm}\rangle$ can be obtained by action of the appropriately chosen product of various creation and annihilation operators. Therefore, it is important to manipulate easily with these operators, and the commutation relations of these operators provide such a tool. In addition, the commutation relations are conceptually important, revealing in a specific way the difference between bosons and fermions, with the far reaching consequences.

At first, we look for the commutation relations of two annihilation operators, a_i and a_j. To this end we compare the action of $a_i a_j$ and $a_j a_i$, for $i < j$:

$$\begin{aligned} a_i a_j |\boldsymbol{n}^{\pm}\rangle &= (\pm)^{l_i(\boldsymbol{n})+l_j(\boldsymbol{n})}\sqrt{n_i n_j}\,|\ldots, n_i-1, \ldots, n_j-1, \ldots\rangle\,, \\ a_j a_i |\boldsymbol{n}^{\pm}\rangle &= (\pm)^{l_i(\boldsymbol{n})+l_j(\boldsymbol{n})-1}\sqrt{n_i n_j}\,|\ldots, n_i-1, \ldots, n_j-1, \ldots\rangle\,. \end{aligned}$$

In the bosonic case the both actions give the equal results. However, since in the second equation a_i acts first, it diminishes by one the number of particles in the states with indices less than j, which in the case of fermions is manifested by the change of sign (of the result). Obviously, for $i = j$, the equal results for bosons, and zeros for fermions are obtained. Finally, for $j < i$ one should reverse the equations ordering. Hence, denoting commutator and anti-commutator of the operators as follows $[A, B]_{\pm} \overset{\text{def}}{=} AB \mp BA$, the results obtained can be concisely written in the form $[a_i, a_j]_{\pm} = 0$. Analogously, the same holds for the annihilation operators: $[a_i^{\dagger}, a_j^{\dagger}]_{\pm} = 0$.

Similarly, commutation relations of the annihilation a_i and creation $a_j^{\dagger}$ operators follows from the equalities:

$$\begin{aligned} a_i a_j^{\dagger} |\boldsymbol{n}^{\pm}\rangle &= (\pm)^{l_i(\boldsymbol{n})+l_j(\boldsymbol{n})}\sqrt{n_i(1 \pm n_j)}\,|\ldots, n_i-1, \ldots, n_j+1, \ldots\rangle\,, \\ a_j^{\dagger} a_i |\boldsymbol{n}^{\pm}\rangle &= (\pm)^{l_i(\boldsymbol{n})+l_j(\boldsymbol{n})-1}\sqrt{n_i(1 \pm n_j)}\,|\ldots, n_i-1, \ldots, n_j+1, \ldots\rangle\,, \end{aligned}$$

where $i < j$, while for $i = j$:

$$a_i a_i^{\dagger} |\boldsymbol{n}^{\pm}\rangle = (\pm)^{l_i(\boldsymbol{n})+l_i(\boldsymbol{n})}\sqrt{(1+n_i)(1 \pm n_i)}\,|\ldots, n_i, \ldots\rangle = (1 \pm n_i)\,|\boldsymbol{n}^{\pm}\rangle\,;$$

to show this for fermions, one again notes that the whole expression *a priori* vanishes for $n_i = 1$, so the change of the factor $1 + n_i$ to $1 - n_i$ does not matter, as well as in the case $n_i = 0$. In the opposite order one gets the occupation number operators, i.e. $a_i^{\dagger} a_i |\boldsymbol{n}^{\pm}\rangle = n_i |\boldsymbol{n}^{\pm}\rangle$. So, all the obtained results are comprised in:

$$[a_i, a_j]_{\pm} = [a_i^{\dagger}, a_j^{\dagger}]_{\pm} = 0, \quad [a_i, a_j^{\dagger}]_{\pm} = \delta_{ij}. \tag{5.41}$$

Note that for the field operators, instead of the Kronecker delta, the Dirac's delta function is used.

Exercise 5.19: Reconsider the harmonic oscillator problem in terms of the second quantization: find single particle space and Fock space. What are particles here? Are they fermions or bosons?

5.2.4 Representation of the second quantization

Any orthonormal basis $|i\rangle$ in the single-particle state $\mathcal{S}$ can be used to construct the set of the creation operators $a_i^\dagger$ in $\mathcal{F}_\pm$, and their adjoints annihilation operators a_i. Consequently, instead of the non-correlated bases in $\mathcal{F}$ constructed from the single-particle basis, the corresponding creation and annihilation operators may be used to represent directly the states and the operators in the physical spaces $\mathcal{F}_\pm$.

As for the vectors, their second quantization form is obvious. Indeed, for arbitrary state $|\Psi\rangle = \sum_{\boldsymbol{n}} \Psi_{\boldsymbol{n}} |\boldsymbol{n}^\pm\rangle$ from $\mathcal{F}_\pm$, (5.38) gives $|\Psi\rangle = \sum_{\boldsymbol{n}} \Psi_{\boldsymbol{n}} \frac{1}{\sqrt{\boldsymbol{n}!}} a_1^{\dagger^{n_1}} \cdots a_{|\mathcal{S}|}^{\dagger^{n_{|\mathcal{S}|}}} |0\rangle$. This may be applied also to represent the single particle states. For example, any wave function can be written as $|\psi\rangle = \int \psi(\boldsymbol{r}) a^\dagger(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}\, |0\rangle = \psi^\dagger |0\rangle$, where $\psi^\dagger = \int \psi(x) a^\dagger(x)\, \mathrm{d}x$, and $a^\dagger(\boldsymbol{r})$ is a coordinate field. Particularly, the plane waves are: $|\boldsymbol{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \int \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}} a^\dagger(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}\, |0\rangle$.

The operators also may be expressed in terms of the creation and annihilation operators, as follows from the Dirac notation. Obviously, the operators leaving the number of particles invariant are combined from the product of the same number of creation and annihilation operators. The additive operators among them have very convenient form. For instance, in the cases of one-particle and two-particle operators $T^{[1]} = \sum_p T_p$ and $V^{[2]} = \frac{1}{2} \sum_{p_1 \neq p_2} V_{p_1 p_2}$, with $T_1 = T = \sum_{ss'} T^s_{s'} |s\rangle \langle s'|$ and $V_{12} = V = \sum_{s_1 s_1' s_2 s_2'} V^{s_1 s_2}_{s_1' s_2'} |s_1 s_2\rangle \langle s_1' s_2'|$ (this notation is used to remind of the typical examples of kinetic energy and pairwise interaction, but the result is quite general):

$$T^{[1]} = \sum_{ss'} T^s_{s'} a_s^\dagger a_{s'}, \tag{5.42a}$$

$$V^{[2]} = \frac{1}{2} \sum_{s_1 s_1' s_2 s_2'} V^{s_1 s_2}_{s_1' s_2'} a_{s_1}^\dagger a_{s_2}^\dagger a_{s_2'} a_{s_1'}. \tag{5.42b}$$

Exercise 5.20: Prove that for additive symmetric L-particle operator $A^{[L]} = \sum_{p_1 < \cdots < p_l} A_{p_1,\dots,p_L}$ holds

$$A^{[L]} = \frac{1}{L!} \sum_{\substack{s_1,\dots,s_L \\ s_1',\dots,s_L'}} A^{s_1,\dots,s_L}_{s_1',\dots,s_L'} a_{s_1}^\dagger \cdots a_{s_L}^\dagger a_{s_L} \cdots a_{s_1}, \tag{5.42c}$$

with $A^{s_1,\dots,s_L}_{s_1',\dots,s_L'} = \langle s_1,\dots,s_L | A_{1,\dots,L} | s_1',\dots,s_L'\rangle$.

Exercise 5.21: Express the angular momentum operators in the form (5.42a) and give the alternative proof of the closed shell property (5.13).

In the context of representations, the transitions between them should be considered, too. Taking besides $|a;i\rangle$ the basis $|b;i\rangle$, related to the first one by the unitary transition operator U, $|b;q\rangle = \sum_{iq} U_{iq} |a;i\rangle$, one gets two sets of creation (and annihilation) operators $a_i^\dagger$ and $b_i^\dagger$. Then obviously

$$b_q^\dagger = \sum_{iq} U_{iq} a_i^\dagger, \quad b_q = \sum_{iq} U_{iq}^* a_i, \tag{5.43}$$

meaning that the creation operators transform like the basis, while the annihilation operators transform like the dual basis. The plane waves illustrate how the transition between the bases relates the corresponding creation operators (fields): $a^\dagger(\boldsymbol{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}} a^\dagger(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}$.

Chapter 6

Approximate Methods

Apart from few simple (although important) problems, Schrödinger equation cannot be solved exactly. A number of methods is developed to find the approximate solutions to it. It should be emphasized that these methods are substantially different, and suited for particular purposes.

6.1 Time Independent Perturbations

6.1.1 Perturbed and Unperturbed Hamiltonian

When the Hamiltonian H can be separated into two parts, $H = H_0 + H'$, the first of them being dominant, it is natural to assume that the solutions (eigenvalues and eigenvectors) of the dominant part H_0 are close to those of the total Hamiltonian. This intuitive fact is easily formalized. To this end the parameter $a \in [0, 1]$ is introduced, such that the a-parameterized operator family

$$H(a) = H_0 + aH', \tag{6.1}$$

form a continuous path from $H_0 = H(0)$ to $H = H(1)$. We assume that along the whole path Hamiltonians $H(a)$ have discrete spectrum. If $E_n(0)$ and $|n, \lambda, 0\rangle$ are the eigenvalues and eigenvectors of the unperturbed Hamiltonian, H_0, then the parameter a, which can be as small as necessary, interrelates perturbed and unperturbed eigenvalues and eigenvectors: if $E_n(1)$ is some eigenvalue of H (for $a = 1$), then when a continuously decreases to 0, one gets path $E_n(a)$ of eigenvalue, with $E_n(0)$ being the corresponding eigenvalue of H_0. However, several different values, say $E_n(a)$, $E_{n'}(a)$,..., can have the same initial unperturbed eigenvalue $E_n(0)$: a degenerate level of the unperturbed Hamiltonian may be split by the perturbation into several less degenerate levels[1]. We denote by $E_{ni}(a)$ all the perturbed levels which come together to $E_n^{(0)} = E_{ni}(0)$ in the unperturbed case. Then

---

[1] It is conceivable that several unperturbed levels $E_n(0)$, $E_{n'}(0)$,..., give paths coinciding at $a = 1$ (different differentiable curves may coincide in the countable number of discrete points). This is tacitly excluded in the forthcoming discussion, as intuitively H' is "small" enough to bridge eigenvalues of H_0. Rigorously, some nonperturbed eigenvalues may be close enough; then they are treated together, as it will be clarified soon.

the condition for the corresponding families of the eigenvalues $E_{ni}(a)$ and eigenvectors $|n, i, \lambda; a\rangle$ is:

$$H(a)\,|n, i, \lambda; a\rangle = E_{ni}(a)\,|n, i, \lambda; a\rangle\,. \tag{6.2}$$

In other words, index i counts the splitting of the unperturbed level $E_n(0)$. The reminded degeneracy (if any) of the perturbed levels is counted by the index λ, such that the perturbed eigenvectors are $|ni\lambda; a\rangle$.

6.1.2 Perturbative expansion

These preliminary considerations enable us to expand in a all the terms in the parametric eigen problem (6.2):

$$(H_0 + aH')\sum_p a^p\,|ni\lambda\rangle^{(p)} = \left(\sum_s a^s E_{ni}^{(s)}\right)\left(\sum_p a^p\,|ni\lambda\rangle^{(p)}\right). \tag{6.3}$$

Here one should take into account that $E_{ni}^{(0)} = E_n^{(0)}$ and $|ni\lambda\rangle^{(0)} = |ni\lambda; 0\rangle$. While the zero order energy levels $E_n^{(0)}$ are assumed to be known, the zero order vectors $|ni\lambda; 0\rangle$ are not, since the unperturbed Hamiltonian H_0 cannot single out any natural subbasis within its eigen space. Then, separating different orders in a one gets the series of equations:

$$a^0: \quad H_0\,|ni\lambda; 0\rangle = E_n^{(0)}\,|ni\lambda; 0\rangle\,; \tag{6.4a}$$

$$a^1: \quad H_0\,|ni\lambda\rangle^{(1)} + H'\,|ni\lambda; 0\rangle = E_n^{(0)}\,|ni\lambda\rangle^{(1)} + E_{ni}^{(1)}\,|ni\lambda; 0\rangle\,; \tag{6.4b}$$

$$\cdots \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a^p: \quad H_0\,|ni\lambda\rangle^{(p)} + H'\,|ni\lambda\rangle^{(p-1)} = \sum_{s=0}^{p} E_{ni}^{(s)}\,|ni\lambda\rangle^{(p-s)}\,. \tag{6.4c}$$

The first of these equations is immediate consequence of the assumption that the zero order eigen problem of H is the eigen problem of H_0. In this sense, this trivial equation is a consistency condition for the method. Next we come to the first order equation. If P_n is the projector onto the unperturbed Hamiltonian eigenspace $\mathcal{S}_n$, then $P_n\,|ni\lambda; 0\rangle = |ni\lambda; 0\rangle$ and $P_n H_0 = P_n H_0 P_n = E_n^{(0)} P_n$. When (6.4b) is projected into $\mathcal{S}_n$ multiplying both its both sides by P_n, one gets:

$$E_n^{(0)} P_n\,|ni\lambda\rangle^{(1)} + P_n H' P_n\,|ni\lambda; 0\rangle = E_n^{(0)} P_n\,|ni\lambda\rangle^{(1)} + E_{ni}^{(1)} P_n\,|ni\lambda; 0\rangle\,.$$

The first terms on the both sides cancel, and this becomes the eigen equation for the operator $P_n H' P_n$, which is the perturbation H' cut in the subspace $\mathcal{S}_n$:

$$P_n H' P_n\,|ni\lambda; 0\rangle = E_{ni}^{(1)}\,|ni\lambda; 0\rangle\,. \tag{6.5}$$

Both the first order correction of energy levels and the zero order state vectors are obtained by solving this eigenproblem. Note further that each unperturbed level is up to the first order correction treated independently. In practice this means that one takes some important levels (usually the ground one, and sometimes a few above it). Taking any eigenbasis $|n\mu; 0\rangle$ for each of these levels $E_n^{(0)}$, the secular equation (6.5) is obtained in the matrix form $\langle n\mu;0|\,H'\,|n\nu;0\rangle = E^{(1)}_n\delta_{\mu\nu}$. Its different eigenvalues $E^{(1)}_{ni}$ are the first order corrections in energy, while the corresponding eigenvectors $|ni,\lambda;0\rangle$ are the zero order eigenvectors (for the perturbation H') obtained as the representative columns in any chosen basis $|n\mu;0\rangle$.

This procedure can be further developed for higher order corrections [9, 13]. Here, only the case of non-degenerate unperturbed level is analyzed. Let us mention here that when $E^{(0)}_n$ is a discrete eigenvalue with very close other energies, or a point of the continual spectrum are to be corrected, one should solve equation (6.5), but with P_n being the sum of projectors of the relevant neighboring eigenvalues (e.g. from some interval centered at $E^{(0)}_n$).

It is convenient to have a simple estimate of the error made in the course of the approximation of the energy levels. The mean value in the corrected eigenvectors of the Hamiltonian terms which are excluded from it in the considered level of approximation, gives such an estimate.

Exercise 6.1: *Stark effect in hydrogen atom* Hamiltonian of a hydrogen atom (with Hamiltonian H_0 given by (4.37)) in a homogeneous electrical field E (along z-axis) is $H = H_0 + H'$, $\quad H' = -eEz$. Show that the matrix elements of H' in the eigenbasis of H_0 in the total (spin inclusive) state space are

$$\langle nlmm_s|\,H'\,|n'l'm'm'_s\rangle = \delta_{m_s,m'_s}\delta_{mm'}(\delta_{l,l'+1}V^{+}_{nn'} + \delta_{l,l'-1}V^{-}_{nn'}). \tag{6.6}$$

Show that the first order energy corrections for levels $n = 1,2$ are $E^{(1)}_1 = 0$ and $E^{(1)}_2 = 0, \pm 3eEa_0$.

6.1.3 Higher Corrections for Non-degenerate Level

If $E^{(0)}_n$ is non-degenerate, the indices i and λ are redundant. While obviously the zero order state must be $|n;0\rangle$ (the only one in $\mathcal{S}_n$), (6.5) shows that the first order energy correction is simply the mean of the perturbation in this state:

$$\langle n;0|\,H'\,|n;0\rangle = E^{(1)}_n. \tag{6.7}$$

On the other hand, the projector $P^{\perp}_n = \sum_{m(\neq n),\mu}|m\mu;0\rangle\langle m\mu;0|$ onto the orthocomplement to $\mathcal{S}_n$ is (we assumed non-degeneracy for n^{th} unperturbed level only, while the other ones may be degenerate) is applied to the first order equation, giving the projection onto $\mathcal{S}^{\perp}_n$:

$$\sum_{m(\neq n),\mu}\left(E^{(0)}_m\langle m\mu;0\,|\,n\rangle^{(1)} + \langle m\mu;0|\,H'\,|n;0\rangle\right)|m\mu;0\rangle = E^{(0)}_n\sum_{m(\neq n),\mu}\langle m\mu;0\,|\,n\rangle^{(1)}\,|m\mu;0\rangle\,.$$

Consequently, finding the projections onto the basis vectors $|m\mu;0\rangle$, and assuming that the first order correction $|n\rangle^{(1)}$ of the state $|n;0\rangle$ is orthogonal to $|n;0\rangle$:

$$\langle m\mu;0\,|\,n\rangle^{(1)} = \frac{\langle m\mu;0|\,H'\,|n;0\rangle}{E^{(0)}_n - E^{(0)}_m},\ \text{i.e.}\ \ |n\rangle^{(1)} = \sum_{m(\neq n),\mu}\frac{\langle m\mu;0|\,H'\,|n;0\rangle}{E^{(0)}_n - E^{(0)}_m}\,|m\mu;0\rangle\,. \tag{6.8}$$

Having in mind the expression for the projector $P^{\perp}_n$ onto the orthocomplement of the 1D subspace spanned by $|n;0\rangle$, which can also be rewritten in the form $P^{\perp}_n = I - |n;0\rangle\langle n;0|$, the above expression for the first order correction $|n\rangle^{(1)}$ can be stated in the basis-independent form:

$$|n\rangle^{(1)} = P_n^{\perp} (E_n^{(0)} - H_0)^{-1} P_n^{\perp} H' |n;0\rangle . \tag{6.9}$$

The second order correction of the energy can be found analogously, projecting (6.4c) for $p = 2$ onto $\mathcal{S}_n$:

$$E_n^{(0)} P_n |n\rangle^{(2)} + P_n H' |n\rangle^{(1)} = E_n^{(0)} P_n |n\rangle^{(2)} + E_n^{(1)} P_n |n\rangle^{(1)} + E_n^{(2)} |n;0\rangle .$$

The first terms on the both sides cancel, the second term on the right vanishes due to the orthogonality of $|n\rangle^{(1)}$ and $|n;0\rangle$:

$$|n;0\rangle \langle n;0| H' |n\rangle^{(1)} = E_n^{(2)} |n;0\rangle .$$

Finally, substituting (6.8) one gets

$$E_n^{(2)} = \sum_{m(\neq n),\mu} \frac{|\langle m\mu;0| H' |n;0\rangle|^2}{E_n^{(0)} - E_m^{(0)}} . \tag{6.10}$$

Exercise 6.2: *Stark effect in diatomic molecule* For a diatomic molecule with the fixed inter-atomic distance the effective Hamiltonian in a homogeneous electrical field E (along z-axis) is

$$H = H_0 + H', \quad H_0 = \boldsymbol{l}^2/2mr_0^2, \quad H' = -dE\cos\theta;$$

the angle between molecule axis and z-axis is θ, and $\boldsymbol{d} = q_1\boldsymbol{r}_1 + q_2\boldsymbol{r}_2$ is the dipole moment of the molecule. Show that the first order contribution of the perturbation H' vanishes, while the second order one is

$$E_l^{(2)} = -\sum_{l>0} \frac{\sum_{m=-l}^{l} |\langle l=0, m=0| H' |lm\rangle|^2}{l(l+1)\hbar^2/mr_0^2} = -\frac{|\langle l=0, m=0| H' |l=1, m=0\rangle|^2}{l(l+1)\hbar^2/mr_0^2} .$$

Usually it is interesting to find the corrections to the energy of the ground state, $E_0^{(0)}$, being the minimal among the eigenvalues of H_0. Obviously, the second order correction is always negative, independently of the system. In fact, there is an estimate of this correction:

Exercise 6.3: Show that if there is a finite difference $\triangle E_n^{(0)}$ from the level $E_n^{(0)}$ to the closest to it level, and the dispersion $\triangle H'$ in the state $|n;0\rangle$ is finite, then

$$|E^{(2)}| \leq \frac{\triangle^2 H'}{|\triangle E_n^{(0)}|} . \tag{6.11}$$

6.2 Variational Method

6.2.1 General characteristics

Foundation

The variational method is based on the well known mathematical theorem on the coincidence of the eigenvectors of an observable, with the stationary vectors of the functional defined over the domain of A as the mean $\langle\psi| A |\psi\rangle$ of A. We are going to prove it, in the form suitable for further purposes.

Theorem 6.1 *Let the Hamiltonian H be with purely discrete spectrum, and let the energy functional be defined as:*

$$E[|\psi\rangle] \overset{\text{def}}{=} \frac{\langle\psi| H |\psi\rangle}{\langle\psi \mid \psi\rangle}. \tag{6.12}$$

Then the stationary points of $E[|\psi\rangle]$ coincide with the eigenvectors of H; the corresponding stationary values are the eigenvalues of H.

▬*Proof:* The variation of (6.12) in the form $\langle\psi| (H - E[|\psi\rangle]) |\psi\rangle = 0$, with the stationary condition $\delta E[|\psi\rangle] = 0$ gives equation $\langle\delta\psi| (H - E[|\psi\rangle]) |\psi\rangle + \langle\psi| (H - E[|\psi\rangle]) |\delta\psi\rangle = 0$. Due to the independence of the variation of the vectors, $|\delta\psi\rangle$, and functionals, $\langle\delta\psi|$, both terms in the last equation must vanish, imposing the eigenvalue condition: $(H - E[|\psi\rangle]) |\psi\rangle = 0$. ▮

In addition, the same proof shows that in any invariant subspace in $\mathcal{S}$, the stationary points of E in any invariant subspace are the eigen vectors of H from that subspace (note that besides the vectors, also the variations $|\delta\psi\rangle$ in the proof are from this subspace). Altogether, the variational problem over the entire state space $\mathcal{S}$ is equivalent to the eigenvalue problem of H, and can be used in the context of the dynamics of the system. Still, in this form it actually reduces to eigenproblem, giving no advantage to its solution. Still, the following form of the min-max theorem (also known as variational theorem or Courant–Fischer–Weyl principle) is a fruitful consequence of the statement 6.1:

Theorem 6.2 *The ground state energy (minimal eigenvalue of H) is less than or equal to the mean value of the Hamiltonian in any state:*

$$E_0 \leq E[|\psi\rangle] = \frac{\langle\psi| H |\psi\rangle}{\langle\psi \mid \psi\rangle}. \tag{6.13}$$

▬*Proof:* Clearly $E[|\psi\rangle] = \sum_n E_n v_n \geq E_0 \sum_n v_n = E_0$, where $v_n = v(E_n, H, |\psi\rangle)$ are probabilities. ▮

Now one uses variational formulation of the eigenproblem to introduce the following approximation. Instead of the whole state space $\mathcal{S}$ only its part $\mathcal{T}$, called trial set, is considered. Then the variational method gives the stationary points within $\mathcal{T}$. According to the theorem 6.2, it is expected that the ground state $|0\rangle$ is best approximated within $\mathcal{T}$ by the state $|\psi_0\rangle$, giving the minimal $E[|\psi_0\rangle]$. Of course, the quality of the approximation is determined by the choice of $\mathcal{T}$.

This stress out that the choice of $\mathcal{T}$ is the critical point of the approximation. Therefore, to improve the results all *a priori* known properties of the ground state should be included. For example, frequently the symmetry or number of zeros of the ground state wave function are *a priory* clear, and the trial functions with such characteristics are to be considered. This is also important in approximations of the excited states within this method. If the ground state, either exact $|0\rangle$, or approximate $|\psi_0\rangle$, is known, the trial set is to be chosen among the orthogonal vectors.

Though there is no precise estimate of the error, some results are useful in this context (Exercise 6.4). For example (6.14) shows that any state, including the stationary over $\mathcal{T}$ states, is in the $\triangle H$-neighborhood of an eigenstate; still, it is not specified to which eigenstate, and the criterion cannot serve as an estimate how far from the ground state is the result.

Exercise 6.4: Let $\{|0\rangle, |1\rangle, \dots\}$ and $\{E_0, E_1, \dots\}$ be the exact eigenvectors and the corresponding eigenvalues of the Hamiltonian, in the increasing energy order.
(i) For each physical state $|\psi\rangle$ there is at least one n such that

$$|E_n - E[|\psi\rangle]| \leq \triangle_{|\psi\rangle} H. \tag{6.14}$$

(ii) If the difference between the variational and the exact ground states $|\psi_0\rangle$ and $|0\rangle$ is estimated by $\epsilon \overset{\mathrm{def}}{=} 1 - |\langle \psi_0 \,|\, 0\rangle|$, then for each $|\psi\rangle$ orthogonal to $|\psi_0\rangle$:

$$\epsilon \geq |\langle \psi \,|\, 0\rangle|^2, \quad E[|\psi\rangle] \geq E_1 - \epsilon(E_1 - E_0). \tag{6.15}$$

Standard realizations

To enable variational calculus, i.e. analysis, the trial set quite generally must be a manifold. Therefore, the states are taken in the parameterized form $|\psi(a_1, a_2, \dots)\rangle$. Then the variational derivative is found as the derivative over the parameters, and the stationary state is obtained as a value of the parameters a_1, a_2, etc., satisfying all the equations

$$\frac{\partial}{\partial a_i} \frac{\langle \psi(a_1, a_2, \dots)| H |\psi(a_1, a_2, \dots)\rangle}{\langle \psi(a_1, a_2, \dots) \,|\, \psi(a_1, a_2, \dots)\rangle} = 0. \tag{6.16}$$

Most frequently, this procedure is performed in the coordinate representation.

Particularly, if the trial set is a subspace, then the parameters are naturally chosen as the coefficients a_i in the expansion $|\psi\rangle = \sum_i a_i |\phi_i\rangle$ of the arbitrary vector over some basis $|\phi_i\rangle$ of $\mathcal{T}$. Then, denoting the Hamiltonian matrix elements in this subbasis by $H_{ij} = \langle \phi_i | H |\phi_j\rangle$, the state norm in the equations (6.16) may be restricted to 1, which introduces the mean energy as the Lagrange multiplier:

$$\frac{\partial}{\partial a_i}\left(\sum_{pq} a_p^* a_q H_{pq} - E \sum_{pq} a_p^* a_q\right) = 0, \text{ i.e. } \sum_p a_p^*(H_{pq} - E) = 0 \;\forall q.$$

Since $|\phi_i\rangle$ form the basis in the subspace $\mathcal{T}$, in the last equation only the matrix elements of the restricted operator $H_{\mathcal{T}} = P_{\mathcal{T}} H P_{\mathcal{T}}$ are involved (as usual, $P_{\mathcal{T}}$ is the projector onto $\mathcal{T}$). Hence, this equation is exactly the eigenvalue problem $\langle \psi | (H_{\mathcal{T}} - E) = 0$, its q-th component written in the $|\phi_i\rangle$ representation. Indeed, this is anticipated when immediately after the Theorem 6.1 it is generalized to any invariant subspace of H: the only relevant operator in the variational procedure turns out to be the restriction $\mathcal{H}_{\mathcal{T}}$, for which the subspace $\mathcal{T}$ is invariant. Therefore, the variational problem with linear trial set reduces to the eigenvalue problem of the restriction.

It is also interesting to compare the variational method to the stationary perturbations. To this end let Hamiltonian be $H = H_0 + H'$, while the trial set is chosen as an eigen space of H_0: $\mathcal{T} = \mathcal{S}_n^0$. The stationary states are found as the eigen states of the restricted Hamiltonian $P_n(H_0 + H')P_n = E_n^0 + P_n H' P_n$. Obviously, the states $|ni\lambda; 0\rangle$ are obtained, with encountered correction of energy of the first order: $E[|ni\lambda; 0\rangle] = E_n^0 + E_{ni}^{(1)}$.

6.2.2 Hartree-Fock method

Correlation of many particle states is the main obstacle to the exact dynamical approach. They cannot be avoided, as the (anti)symmetrization and interaction make the noncorrelated states not physically real. As it has been discussed, the minimal correlation is pure kinematical exchange one, introduced by the indistinguishability of the particles. As for the electrons, this Pauli correlation is included by the Hartree-Fock variational method: the trial set comprises all the Slater determinants. It should be immediately stressed out that the dynamical correlation is completely neglected, which restricts the scope of the applications: the typical error in the obtained ground state energy is 0.1 eV is good enough for prediction of the most of the physical (e.g. electro-optical properties), but too large for reliable predictions of the stable molecular configuration. Therefore, Hartree-Fock method has been during the past decade progressively substituted by the density functional methods, to accommodate requirements of high-technology and materials science.

Energy functional

The cornerstone of the Hartree-Fock method is the restriction to the trial set of Slater determinants. This results in the extremely simple form (5.25b) of the corresponding single- and two-particle reduced states, which are in turn enough to find convenient general form (for this trial set!) of the energy functional. Indeed, as at most pairwise interactions are assumed, the Hamiltonian is $H = H^{[1]} + H^{[2]}$, with the single-particle term $H^{[1]} = \sum_{p=1}^{N} H_p$; here $H_p = T_p + U_p$ includes kinetic energy and potential of the possible external field. The pairwise Coulomb interactions V_{pq}, determine $H^{[2]} = \frac{1}{2}\sum_{p_1 \neq p_2} V_{p_1 p_2}$. Therefore, for the Slater determinant composed of the states $|\phi_1\rangle, \dots, |\phi_N\rangle$, making the single particle space projector $P = \sum_{i=1}^{N} |\phi_i\rangle \langle\phi_i| = N\rho_1$, the energy functional, i.e. the mean value of the Hamiltonian, by (5.31) reads:

$$E[P] = \mathrm{Tr}\,_1(H_1 + \frac{1}{2}U_1^{\mathrm{HF}})P = \sum_{p=1}^{N} \langle\phi_p| (H_1 + \frac{1}{2}U_1^{\mathrm{HF}}) |\phi_p\rangle\,, \tag{6.17}$$

with $V = V_{12}$. The crucial point is that the two-particle terms reduce to the formally single particle operator:

$$U_1^{\mathrm{HF}} \overset{\mathrm{def}}{=} 2\mathrm{Tr}\,_2 P_2^{(-)} V P = 2N\mathrm{Tr}\,_2 P_2^{(-)} V \rho_2; \tag{6.18}$$

here, P and ρ_2 are the operators in the second particle space. In other words, the Slater determinants approximation changes the dynamical model to that of the noninteracting particles in the external field $U + \frac{1}{2}U^{\mathrm{HF}}$.

Explicitly, the terms related to the two-particle interactions in (6.17) are:

$$\langle\phi_p| \frac{1}{2}U_1^{\mathrm{HF}} |\phi_p\rangle = \frac{1}{2}\sum_q (\langle\phi_p\phi_q| V |\phi_p\phi_q\rangle - \langle\phi_p\phi_q| V |\phi_q\phi_p\rangle) \tag{6.19a}$$

(note that the term $p = q$ is cancelled). The first term is called direct or Hartree's and the second one exchange or Fock's (the ordinal of the space is given within the vector for clarity):

$$
\begin{aligned}
U_1^{\mathrm{H}} &= \sum_{p=1}^{\infty}\sum_{q=1}^{N} |\phi_p 1\rangle \langle\phi_p 1| \langle\phi_q 2| V |\phi_q 2\rangle = \sum_{q=1}^{N} \langle\phi_q 2| V |\phi_q 2\rangle & \text{(6.19b)}\\
U_1^{\mathrm{F}} &= \sum_{p=1}^{\infty}\sum_{q=1}^{N} |\phi_p 1\rangle \langle\phi_q 1| \langle\phi_p 2| V |\phi_q 2\rangle\,. & \text{(6.19c)}
\end{aligned}
$$

Their coordinate representation (assuming position dependent interaction potential) is

$$
\begin{aligned}
U_1^{\mathrm{H}}(\boldsymbol{r}) &= \int \mathrm{d}\boldsymbol{r}_2 \sum_q \phi_q^*(\boldsymbol{r}_2)\phi_q(\boldsymbol{r}_2)V(\boldsymbol{r},\boldsymbol{r}_2), & \text{(6.19d)}\\
U_1^{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}') &= \sum_q \phi_q^*(\boldsymbol{r}')\phi_q(\boldsymbol{r})V(\boldsymbol{r},\boldsymbol{r}'). & \text{(6.19e)}
\end{aligned}
$$

The Hartree contribution is diagonal in the coordinates, i.e. it is usual multiplicative single-particle potential. In fact, this is simply averaged interaction with other electrons, and due to the exchange it must be corrected by the Fock's term, which is not of this type. The method in which exchange term is neglected (and possibly treated perturbatively later on) is therefore called Hartree's, while the Hartree-Fock method proceeds with both terms.

Equation of extrema

Having thus the energy functional expressed as the single particle operator mean value, we can perform variational calculus to find the extremes. At first, we emphasize again that the Slater determinants are considered only, which results in the approximation to the ground state with the minimal, Pauli, correlation. On the other hand, due to the biunique relation of the N-particle Slater determinants and N-dimensional projectors in the single particle state space, the trial set is formed by the N-dimensional projectors (or subspaces) in $\mathcal{S}$. Such a set is manifold (thus the variational approach is eligible), but the realization of the variational problem is more transparent within the constrained (conditional) variation scheme: we perform the variation over the linear space of the hermitian operators (apparently enlarging the trial set by the physically senseless operators, not corresponding to the states), but compensate this by the constraint $P^2 - P = 0$. This introduces the Lagrange multiplier Λ, and the variation of the energy functional becomes:

$$
\delta E[P] = \delta \mathrm{Tr}_1 \left[\left(H_1 + \mathrm{Tr}_2(P_2^{(-)} V P)\right) P + \Lambda(P^2 - P)\right] = 0.
$$

Here, Λ is matrix of the dimension of the single particle space, in which the N-dimensional projectors P are trial set. The dimension of the projector is automatically provided in the course of the realization of the variational method. Indeed, the initial condition should be appropriately chosen, i.e. such N-dimensional projector is started with, and in the view of the continuity of the procedure, only the manifold of the same dimensional projectors is taken into account. Since $\Delta(\tau)P_2^{(-)}V(\delta P\otimes P)\Delta(\tau)=P_2^{(-)}V(P\otimes\delta P)$, and $\mathrm{Tr}_1\Lambda(P^2-P)=\mathrm{Tr}_1(P\Lambda-\Lambda)P$, the equation for the variational extreme is

$$\mathrm{Tr}\,[(h+\Lambda P+P\Lambda-\Lambda)\delta P]=0,$$

with the single particle effective Hamiltonian $h=H_1+U_1^{\mathrm{HF}}$. The trace is over the single particle space, which is the total space of the equation. The variational equation has the form of the operator scalar product; thus, since the variation is arbitrary hermitian operator, one gets the equivalent algebraic equation for the extremes:

$$h+\Lambda P+P\Lambda-\Lambda=0. \tag{6.20}$$

Multiplying it by P, firstly from left and then from right, and subtracting the obtained equations, we get:

$$[h,P]=0. \tag{6.21}$$

However, we easily verify that whenever (6.21) is satisfied by some h and P, then also the extremal equation (6.20) is solved by $\Lambda=h-hP-Ph$. Therefore, (6.21) is the solution which is looked for, and it is called the Hartree-Fock equation. If in U_1^{HF} the exchange potential is omitted, one gets the Hartree equation of the same form. Note again that the considerably simple algebraic form of the Hartree-Fock equation (6.21) originates in the effectively single particle procedure, enabled by the dynamical correlations neglecting.

Self consistent solution

It is important that U_1^{HF}, and this way h, is itself determined by the state P (precisely, by the Slater determinant corresponding to P), which is looked for. This makes the equation (6.21) nonlinear in P, and the procedure for solving it is necessarily iterative. To understand this approach, consider the final solution: we obtained the Hartree-Fock Hamiltonian h commuting with the N-dimensional projector P. Consequently, within the range of P there are N states (single particle) which are eigenstates of h:

$$h\,|i_p\rangle=\epsilon_p\,|i_p\rangle\,,\quad p=1,\dots,N. \tag{6.22}$$

As the dynamical model is reduced to the independent (noninteracting) electrons, we may assume that the electrons, due to the Pauli principle, occupy these N states. In this sense, the operator h takes the role of the Hamiltonian; nevertheless, the total energy is not the sum of the above eigenvalues, as U^{HF} is multiplied by $\frac{1}{2}$ in (6.17). Further, if there was not mutual interaction of electrons, i.e. if $U=0$, these states would be real single particle states (and h would be real single particle Hamiltonian).

Thus one starts with a Slater determinant constructed by a set of states $|i_1^0\rangle\,,\dots,|i_N^0\rangle$, giving the projector P^0. Usually, they correspond to the noninteracting case of the considered problem; as Hartree-Fock method is usually used to find the atomic or molecular orbitals, one can use the first N hydrogen like orbitals. Then, for this state U^{HF} and h are to be calculated: $U^{\mathrm{HF}}[P^0]=2\mathrm{Tr}_2P_2^{(-)}V^{(12)}P^0$, $h^0=T+U+U^{\mathrm{HF}}[P^0]$. Next, the eigen problem of h^0 is solved, to get the eigenvectors $|i_p^1\rangle$, and the eigenergies ϵ_p^1 $(p=1,\dots,|\mathcal{S}|)$. The lowest N energies select the eigenstates $|i_1^1\rangle\,,\dots,|i_N^1\rangle$ forming new Slater determinant and projector P^1. The procedure is then repeated: in the s-th step we calculate single particle Hamiltonian h^s, with eigenvectors determining the next step:

$$h^s \left| i_p^{s+1} \right\rangle = \epsilon_p^{s+1} \left| i_p^{s+1} \right\rangle , \quad p = 1, \dots, N.$$

The solution is the fixed point of the procedure: after some iterations, new state is same as the previous. If the initial state is considerably well estimated, the procedure converges, providing self consistency of the solution. The obtained solution surely is the extreme of energy (within the trial set of the Slater determinants), but it is not obvious if it is minimum. However, the choice of the lowest energy eigenvectors usually leads to the minimum.

As mentioned above, the total energy of the resulting state is not the sum of the obtained eigenenergies; indeed by (6.17) one finds:

$$E[P] = \sum_i \epsilon_i - \frac{1}{2} \left\langle U^{\mathrm{HF}} \right\rangle .$$

This is because the sum includes each pairwise interaction twice. Therefore, the eigenvalues ϵ_i are not to be interpreted as the energies of the electrons in the corresponding states $|i\rangle$, but the ionization energies necessary to extract a particle which is in that state (Koopmans' theorem).

6.3 Density Functional Theory

The Hartree-Fock method essentially reduces the interacting many-body problem to the noninteracting one, with suitably changed single particle external potential. The later includes a part of the interparticle interactions as the mean field in the assumed state. The cost for the reduction to the single particle problem is neglecting of correlations (except the Pauli exchange), due to the trial functions of the Slater determinant type.

The same concept inspired a bit more precise method called density functional theory. It is based on the following natural considerations. When isolated, a system of particles (electrons) is described by $H_0 = T + V$. Here V comprises mutual pairwise interactions V_{pq}; it is a universal characteristics of the considered type of particles, independent of the external conditions. Therefore, the ground state ρ, and consequently the reduced density matrix $\rho_1 = \mathrm{Tr}_{\hat{1}} \rho$, of this system is completely determined by the external potential U only. However, there is a reverse interrelation between the density and the external potential:

Theorem 6.3 ***[Hohenberg-Kohn, 1964]*** *Reduced density matrix ρ_1 of the ground state uniquely determines the external potential U. Particularly, for the position dependent potential, the density function $\rho_1(\boldsymbol{r}) = \langle \boldsymbol{r} | \rho_1 | \boldsymbol{r} \rangle$ determines U.*

▬*Proof:* Let ρ_1 be the density of the ground state ρ in the external potential U. Let there is another external potential U' yielding the ground state ρ' with the same reduced statistical operator ρ_1. With $H_0 = T + V$, the minimal energy (5.31) of the ground states in the potentials U and U' are related as:

$$E_U[\rho] = \mathrm{Tr}\, H_0 \rho + N \mathrm{Tr}_1 U \rho_1 < E_U[\rho'] = \mathrm{Tr}\, H_0 \rho' + N \mathrm{Tr}_1 U \rho_1 = E_{U'}[\rho'] + N \mathrm{Tr}_1 (U - U') \rho_1 .$$

Analogously, $E_{U'}[\rho'] < E_U[\rho] + N\mathrm{Tr}_1(U'-U)\rho_1$. Then, adding these two equations, one gets contradiction $E_{U'}[\rho'] + E_U[\rho] < E_{U'}[\rho'] + E_U[\rho]$, proving that if in the potential U the reduced ground state density is ρ_1, there is no other potential producing the same ground state density. This is a consequence of the same ρ_1 in the last term; when the external potential is position dependent only, this term reduces to the $\mathrm{Tr}_1(U-U')\rho_1 = \int \mathrm{d}\boldsymbol{r}\rho_1(\boldsymbol{r})(U(\boldsymbol{r})-U'(\boldsymbol{r}))$, involving the density functions, i.e. their equality is sufficient, as stated. ▮

Although in the above proof the non-degenerate ground state is considered, the generalization to degenerate ground state exists in literature. Further, there are densities $\rho_1(\boldsymbol{r})$ being not reduced states of the ground state for any potential. Thus, the theorem means that there is biunique correspondence between potentials U and the densities of the ground states of a many-particle system in some external potential; such densities are called U-representable.

To summarize, a system of many particles of some sort defines itself an inter-particle interaction V, being universal in the sense that it has the same form for these particles independently on the specific circumstances. As the kinetic energy is universal as well, only the external potential characterizes different physical situations. Now, the Hohenberg-Kohn theorem emphasizes the importance of the densities: they biuniquely correspond to these external conditions, which are manifested through the associated ground states. When this is understood, the theorem implies that the ground state is functional of $\rho_1(\boldsymbol{r})$, as well as the ground state mean value of any many body observable A:

$$\langle A\rangle \overset{\text{def}}{=} \mathrm{Tr}\, A\rho = A[\rho_1(\boldsymbol{r})]. \tag{6.23}$$

The last conclusion applied to the energy functional immediately gives:

Theorem 6.4 ***[Hohenberg-Kohn, 1964]*** *The energy functional*

$$E_U[\rho_1] = F[\rho_1] + \mathrm{Tr}_1 U\rho_1 \tag{6.24}$$

reaches its minimum for the ground state reduced density ρ_1 corresponding to U. Here, F is functional universal for the given type of particles.

Instead of ρ_1, to fix the number of particles N, we use $\varrho = N\rho_1$. The function $\varrho(\boldsymbol{r}) = \langle\boldsymbol{r}|\,\varrho\,|\boldsymbol{r}\rangle$ is real N-particle density, due to the trace relation

$$\mathrm{Tr}_1\varrho = \int \mathrm{d}\boldsymbol{r}\varrho(\boldsymbol{r}) = N. \tag{6.25}$$

Density variational principle

It is well known that the N-particle ground state $|\Psi\rangle$ is the minimal state of the energy mean. The variational procedure may be performed in two steps. Firstly, for each fixed $\varrho(\boldsymbol{r})$, one may find $|\Psi_\varrho\rangle$ as the minimum of the energy mean over the states giving same $\varrho(\boldsymbol{r})$. Thus this mean is the functional of $\varrho(\boldsymbol{r})$:

$$E[|\Psi_\varrho\rangle] = \min_{\Phi\to\varrho(\boldsymbol{r})} \langle\Phi|\,H\,|\Phi\rangle = F[\varrho(\boldsymbol{r})] + \int \mathrm{d}\boldsymbol{r}U(\boldsymbol{r})\varrho(\boldsymbol{r}). \tag{6.26}$$

This defines the universal (exterior field independent) functional

$$F[\varrho(\boldsymbol{r})] \overset{\text{def}}{=} \min_{\Phi \to \varrho(\boldsymbol{r})} \langle \Phi | T + V | \Phi \rangle = \langle \Psi_\varrho | T + V | \Psi_\varrho \rangle . \tag{6.27}$$

In the next step, the ground state is found as the minimum of $E[|\Psi_\varrho\rangle]$ over densities, i.e. the states $|\Psi_\varrho\rangle$. The Euler-Lagrange equation for the constrained variational extremes [2] is

$$\frac{\delta F[\varrho(\boldsymbol{r})]}{\delta \varrho(\boldsymbol{r})} + U(\boldsymbol{r}) = \mu, \tag{6.28}$$

where the chemical potential μ is the Lagrange coefficient for the density constraint (6.25).

Determination of the functional $F[\varrho(\boldsymbol{r})]$ would transform (6.28) into equation in ϱ, making N-particle problem exactly solvable. However, the functional is not known. In principle, it is a sum of the kinetic energy ($T[\varrho(\boldsymbol{r})]$) and potential containing the direct (Hartree) term ($D[\varrho(\boldsymbol{r})]$), Fock exchange ($X[\varrho(\boldsymbol{r})]$) and correlation ($C[\varrho]$) functionals[3]. The last two terms are usually joined into the exchange-correlation functional $E_{\text{xc}}[\varrho(\boldsymbol{r})]$. These potentials will be precisely determined later. The kinetic energy and correlation functional are not known exactly.

To handle the kinetic energy, we consider a system of noninteracting particles, i.e. $V = 0$, when essentially the single particle approximation is fulfilled, and instead of F the kinetic energy functional T_0 is obtained. Thus, the Hamiltonian is $H_0 = T + U$, and for fixed $\varrho(\boldsymbol{r})$ one can define $|\Psi_{\varrho 0}\rangle$ and the ground state $|\Psi_0\rangle$:

$$T_0[\varrho(\boldsymbol{r})] \overset{\text{def}}{=} \min_{\Phi \to \varrho(\boldsymbol{r})} \langle \Phi | T | \Phi \rangle = \langle \Psi_{\varrho 0} | T | \Psi_{\varrho 0} \rangle , \tag{6.29}$$

giving the variational minimum by the equation:

$$\frac{\delta T_0[\varrho(\boldsymbol{r})]}{\delta \varrho(\boldsymbol{r})} + U_0(\boldsymbol{r}) = \mu. \tag{6.30}$$

Now, assuming that $\varrho(\boldsymbol{r})$ is both interacting and noninteracting system U-representable, one can think of (6.28) and (6.30) as describing equivalent (via $\varrho(\boldsymbol{r})$) interacting and noninteracting systems. Their equivalence is achieved by the redefined (with respect to $U(\boldsymbol{r})$) external field $U_0(\boldsymbol{r})$ (which is also suited to get the same Lagrange multiplier μ). Thus, the Euler-Lagrange equation for the noninteracting system give Kohn-Sham equations:

Theorem 6.5 ***[Kohn-Sham, 1965]*** *The exact N-particle ground state density $\varrho(\boldsymbol{r})$ is $\varrho = \sum_{p=1}^{N} |p\rangle \langle p|$, where $|p\rangle$ are the N lowest eigenvectors of the Kohn-Sham single-particle Hamiltonian $H_0 = T + U_0$:*

$$H_0 |p\rangle = \epsilon_p |p\rangle , \quad U_0(\boldsymbol{r}) = U(\boldsymbol{r}) + \frac{\delta D[\varrho(\boldsymbol{r})]}{\delta \varrho(\boldsymbol{r})} + \frac{\delta E_{\text{xc}}[\varrho(\boldsymbol{r})]}{\delta \varrho(\boldsymbol{r})}. \tag{6.31}$$

---

[2] Variation of the functional $F[\varrho]$ is defined as $\delta F[\varrho] \overset{\text{def}}{=} \frac{\partial}{\partial \alpha} \left(F[\varrho(\boldsymbol{r}) + \alpha \delta \varrho] \right)_{\alpha=0}$. The functional derivative is the functional of ϱ comprising the part of variation linear in $\delta\varrho$. For the local functional $F[\varrho] = \int \mathrm{d}\boldsymbol{r} f[\boldsymbol{r}, \varrho]$, the variation is $\int \mathrm{d}\boldsymbol{r} f'_\varrho[\boldsymbol{r}, \varrho] \delta\varrho$, making $\frac{\delta F[\varrho]}{\delta \varrho} = f'_\varrho[\boldsymbol{r}, \varrho(\boldsymbol{r})]$.

[3] Exchange term includes correlation due to antisymmetrization only, while the rest is gathered in correlation term; this means that Slater determinant gives $C = 0$, while $C \neq 0$ is provided only by their linear combinations.

This equation can be solved self consistently, in analogy with Hartree-Fock method if all the potentials are known. For the noninteracting Kohn-Sham orbitals (giving $|\Psi_0\rangle$ as their Slater determinant) $T_0[\varrho(\boldsymbol{r})]$ is found exactly, as well as the Hartree direct term, $D[\varrho(\boldsymbol{r})] = \frac{1}{2}\int \mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\varrho(\boldsymbol{r})\varrho(\boldsymbol{r}')v(\boldsymbol{r}-\boldsymbol{r}')$, introducing the potential (6.19d). Since Slater determinant $|\Psi_0\rangle$ involves no correlation, $X[\varrho(\boldsymbol{r})] \overset{\mathrm{def}}{=} \langle\Psi_0|\,V\,|\Psi_0\rangle - D[\varrho(\boldsymbol{r})]$, giving $\langle\Psi_0|\,T + V\,|\Psi_0\rangle = T_0[\varrho(\boldsymbol{r})] + D[\varrho(\boldsymbol{r})] + X[\varrho(\boldsymbol{r})]$. As $F[\varrho(\boldsymbol{r})] = \langle\Psi|\,T+V\,|\Psi\rangle$, and $E_{xc}[\varrho(\boldsymbol{r})] = X[\varrho(\boldsymbol{r})] + C[\varrho(\boldsymbol{r})]$ it remains:

$$C[\varrho] = F[\varrho] - (T_0[\varrho] + D[\varrho] + X[\varrho]) = \langle\Psi|\,(T+V)\,|\Psi\rangle - \langle\Psi_0|\,(T+V)\,|\Psi_0\rangle\,. \tag{6.32}$$

Note that when $V = 0$, single particle case, $X[\varrho(\boldsymbol{r})] = -D[\varrho(\boldsymbol{r})]$, making $C[\varrho(\boldsymbol{r})] = 0$ (since a priori $F[\varrho(\boldsymbol{r})] = T_0[\varrho(\boldsymbol{r})]$). For $V \neq 0$, since $|\Psi\rangle$ is the minimal point of $\langle T+V\rangle$, $C[\varrho(\boldsymbol{r})] \leq 0$. On the contrary, $|\Psi_0\rangle$ minimize $\langle T\rangle$, making $\langle T\rangle_\Psi - \langle T\rangle_{\Psi_0}$ positive. Consequently, the rest of $C[\varrho(\boldsymbol{r})]$, $\langle V\rangle_\Psi - \langle V\rangle_{\Psi_0}$, is negative.

Thus the main problem is determination of correlation functional. This is performed within some approximations. It appears that it is easier to approximate together exchange and correlation. To understand this energy, we consider family of the Hamiltonians, $H_\lambda = T + \lambda V + U$, yielding $|\Psi_\lambda\rangle$ with $\varrho(\boldsymbol{r})$ fixed (adiabatic connection). For $\lambda = 0$ the Kohn-Sham system is obtained, and for $\lambda = 1$ the actual interacting system. Further, we introduce two-particle density operator $\varrho_2^\lambda = N(N-1)\mathrm{Tr}_{\,3,\dots,N}\,|\Psi_\lambda\rangle\langle\Psi_\lambda|$, and two particle density $\varrho_2^\lambda(\boldsymbol{r}',\boldsymbol{r}) = \langle\boldsymbol{r},\boldsymbol{r}'|\,\varrho_2^\lambda\,|\boldsymbol{r},\boldsymbol{r}'\rangle$, giving the joint probability to find particle around $\boldsymbol{r}'$ and another one around $\boldsymbol{r}$. Then

$$\langle V\rangle_{\Psi_\lambda} = \frac{1}{2}\mathrm{Tr}\,V\varrho_2 = \frac{1}{2}\int \mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\varrho_2^\lambda(\boldsymbol{r}',\boldsymbol{r})v(\boldsymbol{r},\boldsymbol{r}'). \tag{6.33}$$

If the conditional probability to find a particle near $\boldsymbol{r}'$ when the other particle is fixed at $\boldsymbol{r}$ is $n_2^\lambda(\boldsymbol{r},\boldsymbol{r}')$, then $\varrho_2^\lambda(\boldsymbol{r}',\boldsymbol{r}) = \varrho(\boldsymbol{r})n_2^\lambda(\boldsymbol{r},\boldsymbol{r}')$. Obviously, $n_2^\lambda(\boldsymbol{r},\boldsymbol{r}')$ is the density of the remaining particles, when one is at $\boldsymbol{r}$, giving $\int \mathrm{d}\boldsymbol{r}'n_2^\lambda(\boldsymbol{r},\boldsymbol{r}') = N-1$. Finally, to extract the exchange-correlation part, one writes $n_2^\lambda(\boldsymbol{r},\boldsymbol{r}') = \varrho(\boldsymbol{r}') + n_{\mathrm{xc}}^\lambda(\boldsymbol{r},\boldsymbol{r}')$. The first term is the direct one, for the noninteracting particles. This can be used in (6.35), written in the form

$$E_{\mathrm{xc}}[\varrho(\boldsymbol{r})] = \int_0^1 \mathrm{d}\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\Psi_\lambda|\,(T{+}\lambda V)\,|\Psi_\lambda\rangle - D[\varrho] = \int_0^1 \mathrm{d}\lambda\,\langle\Psi_\lambda|\,(T{+}V)\,|\Psi_\lambda\rangle - D[\varrho] \tag{6.34}$$

(Hellman-Faynman theorem is used on the right). Altogether, one gets:

$$E_{\mathrm{xc}}[\varrho(\boldsymbol{r})] = \frac{1}{2}\int \mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\varrho(\boldsymbol{r})\bar{n}_{\mathrm{xc}}^\lambda(\boldsymbol{r}',\boldsymbol{r})V(\boldsymbol{r},\boldsymbol{r}'), \quad \bar{n}_{\mathrm{xc}}(\boldsymbol{r}',\boldsymbol{r}) = \int_0^1 \mathrm{d}\lambda n_{\mathrm{xc}}^\lambda(\boldsymbol{r}',\boldsymbol{r}). \tag{6.35}$$

This shows that the exchange-correlation energy is the interaction between a particle and averaged (with respect to the coupling constant λ) exchange-correlation hole around it. This hole appears due to three factors: (i) self-interaction correction, (ii) Pauli principle, (3) potential V. First two of them contribute to the exchange, and the last to correlation.

6.3.1 Uniform electron gas

We consider d-dimensional noninteracting electron gas of the uniform density ϱ. When only translational subgroup of the lattice with all periods equal to L is used, the corresponding eigenorbitals are plane waves $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}/\sqrt{L^d}$, with eigenenergies $\hbar k^2/2m$. Assuming g dimensional interior space ($g = 2$ for spin) the density of orbitals in $\boldsymbol{k}$-space is $gL^d/(2\pi)^d$. On the other hand, $N = \varrho L^d$ electrons from the elementary cell occupy all the states with the momentum $k \leq k_{\mathrm{F}}$ (each with the degeneracy g). This defines the Fermi sphere with radius k_{F}, such that its volume $S_d = \pi^{d/2}k_{\mathrm{F}}^d/\Gamma(d/2+1)$ contains exactly N states. Therefore the density is:

$$\varrho = k_{\mathrm{F}}^d \frac{g}{2^d} \frac{1}{\pi^{d/2}\Gamma(d/2+1)} = \frac{\Gamma(d/2+1)}{\pi^{d/2} r_{\mathrm{S}}^d}, \tag{6.36}$$

where the Seitz radius r_{S} defines the sphere containing single electron.

The average kinetic energy per electron of these orbitals is $t_0 = \frac{g}{N}\frac{\hbar^2}{2m}\sum_{k<k_{\mathrm{F}}} k^2$; using $\boldsymbol{k}$-space orbital density $gL^d/(2\pi)^d$, and $\boldsymbol{k}$-space volume element $\mathrm{d}\boldsymbol{k}^d = d\pi^{d/2}k^{d-1}/\Gamma(d/2+1)\,\mathrm{d}k$, one finds:

$$t_0[\varrho] = \frac{\hbar^2}{2m}\frac{d}{d+2}k_{\mathrm{F}}^2 = \frac{\hbar^2}{2m}\frac{d}{d+2}\frac{4\pi}{g}\Gamma^{2/d}(d/2+1)\varrho^{2/d} = \frac{\hbar^2}{2m}\frac{d}{d+2}\frac{4\Gamma^{2/d}(d/2+1)}{g^{2/d}r_{\mathrm{S}}^2}. \tag{6.37}$$

Particularly, for $g = 2$ and $d = 3$ one gets:

$$T_0[\varrho] = \int \mathrm{d}\boldsymbol{r}\varrho(\boldsymbol{r})t_0[\varrho] = \frac{\hbar^2}{m}\frac{3^{2/3}3\pi^{4/3}}{10}\int \mathrm{d}\boldsymbol{r}\varrho^{5/3}(\boldsymbol{r}). \tag{6.38}$$

6.3.2 Exchange energy

The Fock exchange energy for the Slater determinant Ψ of doubly occupied orbitals ψ_i is functional:

$$E_{\mathrm{x}}[\Psi] = -\frac{1}{2}\sum_{\sigma}\sum_{i,j}\int\int \mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\frac{\phi_{i\sigma}^*(\boldsymbol{r})\phi_{j\sigma}^*(\boldsymbol{r}')\phi_{i\sigma}(\boldsymbol{r}')\phi_{j\sigma}(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{6.39}$$

To this end, all the functionals should be known, which is not the case. However, various approximations are applied. These are based on the assumptions on thee locality of the used functionals. In fact, the functionals having Riesz-Fréchet representations, $F[\rho] \overset{\mathrm{def}}{=} \int \mathrm{d}\boldsymbol{r}\phi(\boldsymbol{r})\rho(\boldsymbol{r})$, are called local. However, there is no direct argument resulting in the locality, and therefore the expansion over derivatives of ρ_1 must be involved to mimic nonlocality. The zero-order term, without derivatives, being thus of the local form, in this expansion gives the local density approximation (LDA). Neglect of the gradient terms corresponds to the uniform gas of particles. Therefore, to get this approximation one looks for the relevant energy terms in the uniform gas, where all of terms are functions of density (being the only parameter of the system). Then, one comes back to the slowly varying density (becoming function of $\boldsymbol{r}$), and construct the average of the energy terms with such density; this average is itself a functional of ρ_1.

6.4 Time dependent perturbations

The goal of the time dependent perturbation theory is to describe the response of a quantum system to an external perturbation. It is assumed that, when isolated, the considered system is described by the Hamiltonian H_0, with the eigensystem

$$H_0 \left| n \right\rangle = E_n \left| n \right\rangle . \tag{6.40}$$

Then, during the time interval $[0, t]$, it is exposed to the external field (in a very broad sense, e.g. arbitrary interaction with some other system). It contributes by an additional interaction term in the Hamiltonian, the perturbation $H'(t)$ which is a (nonvanishing during the interval $[0, t]$) function of time. The most frequent examples are constant or periodic perturbation, which may describe influence of the electromagnetic radiation.

The response of the system is manifested in excitations: $H'(t)$ induces transitions between the various eigenstates of the unperturbed Hamiltonian. Thus, the main task is to find these transition probabilities. Precisely, for the system prepared at $t = 0$ in the eigenstate $\left| n \right\rangle$ (of the isolated Hamiltonian H_0), we look for the probability

$$v(1, \left| n' \right\rangle \left\langle n' \right| , \left| n, t \right\rangle) = v_{n\to n'}(t) = | \left\langle n' \right| U(t) \left| n \right\rangle |^2 \tag{6.41}$$

to find the system at t in some other eigenstate $\left| n' \right\rangle$.

6.4.1 Expansion of the evolution

The task is efficiently solved within Dirac's picture (Sec. 2.5.2). The total evolution, governed by

$$H(t) = H_0 + H'(t), \tag{6.42}$$

is factorized in the form

$$U(t) = U^{(0)}(t)U'(t). \tag{6.43}$$

While $U(t)$ satisfies the equation (2.2), the unperturbed evolution $U^{(0)}$ and the remained factor $U'(t)$ obey (2.37) and (2.38) (shortened notation: $U(t-t_0, t_0) = U(t-0, 0) = U(t)$):

$$\begin{aligned} \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}U^{(0)}(t) &= H_0 U^{(0)}(t), \quad U^{(0)}(0) = I, \\ \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}U'(t) &= H'_D(t)U'(t), \quad U'(0) = I. \end{aligned}$$

The formal integral solution of the last equation

$$U'(t) = I + \frac{1}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t_1 H'_D(t_1)U'(t_1) \tag{6.44}$$

is iterated to get the series

$$U'(t) = \sum_{p=0}^{\infty} U'^{(p)}(t), \quad U'^{(p)}(t) = \frac{1}{\mathrm{i}^p\hbar^p}\int_0^t \cdots \int_0^{t_{p-1}} \mathrm{d}t_1 \cdots \mathrm{d}t_p H'_D(t_1)\cdots H'_D(t_p).$$

Note the time ordering $t \geq t_1 \geq t_2 \geq \cdots \geq 0$. The convergence of the series is assumed. As only the first few terms of the series are considered (neglecting the reminder), this convergence is particularly important, and considerations are restricted to sufficiently small t. When $U'(t)$ is substituted, (6.43) becomes:

$$U(t) = \sum_{p=0}^{\infty} U^{(p)}(t), \quad U^{(p)}(t) = U^{(0)}(t)U'^{(p)}(t). \tag{6.45}$$

Taking into account $H'_D(t) = U^{(0)^\dagger}(t)H'(t)U^{(0)}(t)$ and $U^{(0)}(t_i)U^{(0)^\dagger}(t_j) = U^{(0)}(t_i - t_j)$, the convenient form of the terms is obtained:

$$U^{(p)}(t) = \frac{1}{\mathrm{i}^p\hbar^p}\int_0^t \cdots \int_0^{t_{p-1}} \mathrm{d}t_1 \cdots \mathrm{d}t_p U^{(0)}(t-t_1)H'(t_1)U^{(0)}(t_1-t_2)\cdots H'(t_p)U^{(0)}(t_p). \tag{6.46}$$

6.4.2 Transition amplitudes

As the transition probability (6.41) involves the probability amplitudes $\langle n'|\, U(t)\, |n\rangle$, we have to find the matrix elements $\langle n'|\, U^{(p)}(t)\, |n\rangle$. To simplify this, we insert the identical operators $I = \sum |n_i\rangle \langle n_i|$ after each $H'(t_i)$ in (6.46), which reduces the unperturbed evolution to the phases:

$$\begin{aligned}\langle n'|\, U^{(p)}(t)\, |n\rangle &= \sum_{n_1,\dots,n_{p-1}} \frac{1}{\mathrm{i}^p\hbar^p}\int_0^t \cdots \int_0^{t_{p-1}} \mathrm{d}t_1 \cdots \mathrm{d}t_p \times \\ &\times\ \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}[E_{n'}(t-t_1)+E_{n_1}(t_1-t_2)+\cdots+E_{n_{p-1}}(t_{p-1}-t_p)+E_n t_p]} \times \\ &\times\ \langle n'|\, H'(t_1)\, |n_1\rangle \langle n_1|\, H'(t_2)\, |n_2\rangle \cdots \langle n_{p-1}|\, H'(t_p)\, |n\rangle\,.\end{aligned} \tag{6.47}$$

In practice, the series (6.41) is truncated to a few first terms. We explicitly find them for $p = 0, 1, 2$:

$$\langle n'|\, U^{(0)}(t)\, |n\rangle = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_n t}\delta_{nn'}; \tag{6.48}$$

$$\langle n'|\, U^{(1)}(t)\, |n\rangle = \frac{\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_{n'}t}}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t_1 \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(E_{n'}-E_n)t_1} \langle n'|\, H'(t_1)\, |n\rangle\,; \tag{6.49}$$

$$\begin{aligned}\langle n'|\, U^{(2)}(t)\, |n\rangle &= \sum_{n_1} \frac{\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_{n'}t}}{\mathrm{i}^2\hbar^2}\int_0^t \int_0^{t_1} \mathrm{d}t_1\, \mathrm{d}t_2 \mathrm{e}^{\frac{\mathrm{i}}{\hbar}[(E_{n'}-E_{n_1})t_1+(E_{n_1}-E_n)t_2]} \times \\ &\times\ \langle n'|\, H'(t_1)\, |n_1\rangle \langle n_1|\, H'(t_2)\, |n\rangle\,.\end{aligned} \tag{6.50}$$

The last expressions, as well as the general term (6.47) can be nicely interpreted by diagrams Fig. 6.1. In fact, it turns out that the perturbation in the expression for $U^{(p)}$ acts only in the instances $t_1, \dots, t_p$, and meanwhile the system evolves as isolated. Consequently, the state $|i\rangle$ during the interval (t_i, t_{i+1}) obtains pure phase factor. At the instance t_i, affected by the perturbation $H'(t_i)$, the system abruptly changes state from $|n_{i-1}\rangle$ to $|n_i\rangle$, not conserving its energy (of isolated system). Thus, for $p \geq 2$ there appear intermediate states $|n_1\rangle, \dots, |n_{p-1}\rangle$, which are called virtual states. Transitions to all unperturbed eigenstates at any instance from $[0, t]$ are included by summations over the unperturbed basis, and integration over all intermediate instances.

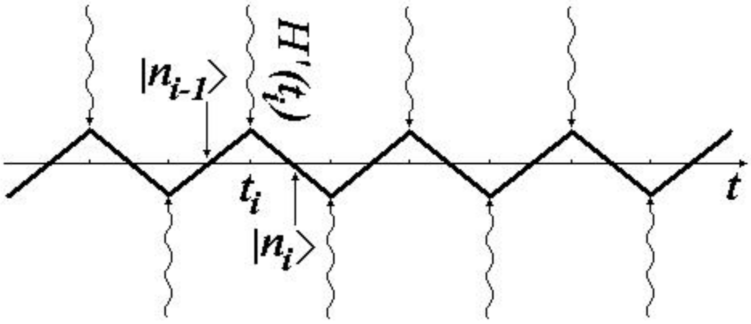

Figure 6.1: Scheme of the terms of the transition amplitudes: wave lines represent perturbation acting in the instances t_i only, and causing system transition from the state $|n_{i-1}\rangle$ to $|n_i\rangle$, while meanwhile the system evolves as isolated.

6.4.3 Transition probabilities

The preceding results enable to expand the transition probabilities. Substituting (6.45) in (6.41) we find

$$\begin{aligned} v_{n\to n'}(t) &= \sum_{p,p'=0}^{\infty} \langle n'|\, U^{(p)}(t)\, |n\rangle \, \langle n'|\, U^{(p')}(t)\, |n\rangle^* \\ &= \sum_{p}^{\infty} \sum_{s=0}^{p} \langle n'|\, U^{(s)}(t)\, |n\rangle \, \langle n'|\, U^{(p-s)}(t)\, |n\rangle^* \, . \end{aligned} \tag{6.51}$$

Using (6.48) we find the first two terms in the transition probabilities:

$$v^{(0)}_{n\to n'}(t) = \delta_{nn'}; \tag{6.52}$$

$$v^{(1)}_{n\to n'}(t) = \delta_{nn'} \frac{2}{\hbar} \mathrm{Im}[\mathrm{e}^{\frac{\mathrm{i}}{\hbar}(E_n - E_{n'})t} \int_0^t \mathrm{d}t_1 \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(E_{n'}-E_n)t_1} \langle n'|\, H'(t_1)\, |n\rangle]; \tag{6.53}$$

$$v^{(2)}_{n\to n'}(t) = v^{(2)}_{n\not\to n'}(t) + \delta_{nn'} v^{(2)}_{n\Rightarrow n'}(t). \tag{6.54}$$

Obviously, in the zeroth and the first order, probabilities forbid transitions (to other states then the initial one $|n\rangle$). Only in the second order, besides the state preserving term $v^{(2)}_{n\Rightarrow n}(t) = 2\mathrm{Re}[\langle n|\, U^{(2)}(t)\, |n\rangle \, \langle n|\, U^{(0)}(t)\, |n\rangle^*]$, there is a probability of the transition to other states. Introducing Bohr's frequencies $\omega_{n'n} = \frac{E_{n'}-E_n}{\hbar}$ one finds the transition probability (in the first effective order):

$$v^{(2)}_{n\not\to n'}(t) = \frac{1}{\hbar^2} \left| \int_0^t \mathrm{d}t_1 \mathrm{e}^{\mathrm{i}\omega_{n'n}t_1} \langle n'|\, H'(t_1)\, |n\rangle \right|^2 . \tag{6.55}$$

Note that the expression is symmetric in $|n\rangle$ and $|n'\rangle$ as $H'(t)$ is hermitian; therefore $v^{(2)}_{n\not\to n'}(t) = v^{(2)}_{n'\not\to n}(t)$, i.e. the opposite transitions occur with same probability. This characteristic of the first order transition probability remains in the higher orders only for the systems with time reversal symmetry.

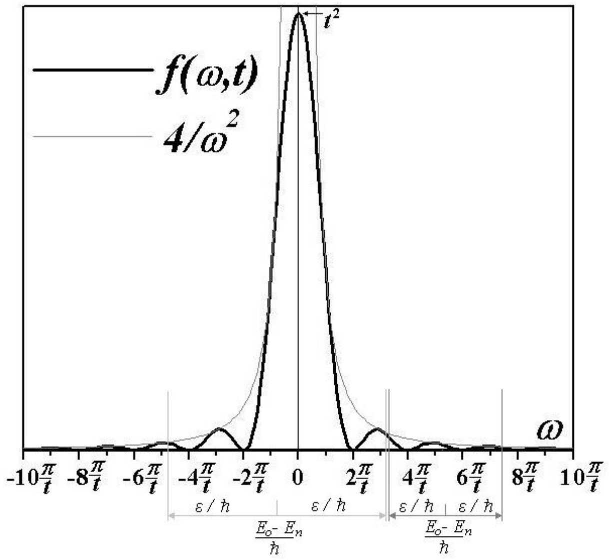

Figure 6.2: Function $f(\omega, t)$: note its periodicity of $2\pi/t$ dumped by the enveloping function $4/\omega^2$. Bellow: intervals of frequency corresponding to Eqns. (6.61) and (6.60).

Constant perturbation and Fermi golden rule

When H' is constant during the considered time interval, the matrix element $\langle n'| H' |n\rangle$ is constant, and the remaining time dependent part is easily integrated giving squared "cardinal sine" function (Fig. 6.2):

$$f(\omega, t) = \left| \int_0^t \mathrm{d}t_1 \mathrm{e}^{\mathrm{i}\omega t_1} \right|^2 = 2\frac{1-\cos\omega t}{\omega^2} = \left(\frac{\sin t\frac{\omega}{2}}{\frac{\omega}{2}} \right)^2 . \tag{6.56a}$$

Two asymptotic properties of this function are important for further analysis:

$$\int_{-\infty}^{\infty} f(\omega, t)\,\mathrm{d}\omega = 2\pi t, \tag{6.56b}$$

$$f(\omega, t \to \infty) = 2\pi t \delta(\omega). \tag{6.56c}$$

In the first (effective) order, the transition probability becomes:

$$v^{(2)}_{n \nrightarrow n'} = \frac{1}{\hbar^2} |\langle n'| H' |n\rangle|^2 f(\omega_{n'n}, t). \tag{6.57}$$

For experimental applications it is useful to consider probability of transition to a group of close states N' (not to a single state). This probability is sum of the probabilities for each of these states:

$$v^{(2)}_{n \nrightarrow N'} = \sum_{n' \in N'} \frac{1}{\hbar^2} |\langle n'| H' |n\rangle|^2 f(\omega_{n'n}, t). \tag{6.58}$$

Particularly frequent is the situation when the energies (eigen-energies of $H^{(0)}$) of the final states are (quasi)continually distributed within the interval $(E_0 - \varepsilon/2, E_0 + \varepsilon/2)$. Then the last sum becomes integral with the contribution of each energy E pondered by its weight, i.e. by the density of states $\rho(E)$:

$$v^{(2)}_{n\nrightarrow N'} = \frac{1}{\hbar^2}\int_{E_0-\varepsilon/2}^{E_0+\varepsilon/2} |\langle n'|\, H'\,|n\rangle|^2 f(\frac{E-E_n}{\hbar}, t)\rho(E)\,\mathrm{d}E. \tag{6.59}$$

For sufficiently small ε the matrix elements and density may be considered almost constant, and taken outside the integral. On the other side, for large t (but still within the applicability of the approximation!), $\varepsilon \gg 2\pi\hbar/t$, and the interval of integration contains a number of periods $2\pi/t$ of $f(\omega, t)$. Two different situations may occur:

(i) $E_n \in (E_0-\varepsilon/2, E_0+\varepsilon/2)$ (Fig. 6.2). Then, since the contribution of the central peak is highly dominant (according to (6.56c)), the integration may be approximately extended to the whole energy axis, and (6.56b) is used to obtain:

$$v^{(2)}_{n\nrightarrow N'} = \frac{2\pi}{\hbar}|\langle n'|\, H'\,|n\rangle|^2\rho(E_n)t. \tag{6.60}$$

This conclusion, that the transition probability to the nearby states linearly increases with time is known as the Fermi's golden rule.

(ii) $E_n \notin (E_0-\varepsilon/2, E_0+\varepsilon/2)$ (Fig. 6.2). Now the central peak is excluded, but the integral may be approximated by its mean over a number of periods. As the mean of the denominator is $(E_0-E_n)^2$, while the mean of $2(1-\cos\omega t)$ over multiple period $[2\pi k/t, 2\pi(k+k')/t]$ (k and k' are integers) is 2, the estimate of the transition amplitude becomes:

$$v^{(2)}_{n\nrightarrow N'} = \frac{2\varepsilon|\langle n'|\, H'\,|n\rangle|^2}{(E_0-E_n)^2}\rho(E_0) \tag{6.61}$$

To conclude, the last equation shows that the transition rate per unit time is constant for the transitions conserving the energy of the isolated system, while it vanishes (t is large) if the energy is changed.

Periodic perturbation and resonances

Another frequent situation is that the perturbation is periodic in time (e.g. optical transitions induced by the electromagnetic waves). Then its general form is

$$H'(t) = A\mathrm{e}^{\mathrm{i}\omega t} + A^\dagger\mathrm{e}^{-\mathrm{i}\omega t}, \tag{6.62}$$

where A is time independent operator. Instead of the complete matrix elements, only time independent part can be put outside the integral in (6.55), resulting in

$$v^{(2)}_{n\nrightarrow n'} = \frac{1}{\hbar^2}\left|\langle n'|\, A\,|n\rangle\int_0^t \mathrm{e}^{\mathrm{i}(\omega_{n'n}+\omega)t_1}\,\mathrm{d}t_1 + \langle n'|\, A^\dagger\,|n\rangle\int_0^t \mathrm{e}^{\mathrm{i}(\omega_{n'n}-\omega)t_1}\,\mathrm{d}t_1\right|^2. \tag{6.63}$$

Again, the case when t is sufficiently large is considered. Then, the first term in the probability amplitude is negligible unless $E_{n'} = E_n - \hbar\omega$, which is the process of emitting of the energy quant $\hbar\omega$; analogously, the second term describes dominantly the absorption of the same quant: $E_{n'} = E_n + \hbar\omega$. Under the same assumption of large t, $t \gg 2\pi/\omega$, these two regions are separated. Therefore, the amplitudes of emission and absorption do not interfere, and the probability is large only for these two processes, which can be analyzed separately. For emission, one gets analogously to (6.57):

$$v^{(2)\mathrm{e}}_{n\nrightarrow n'} = \frac{1}{\hbar^2}|\left\langle n'\right| A \left|n\right\rangle|^2 f(\omega_{n'n}+\omega,t), \quad v^{(2)\mathrm{a}}_{n\nrightarrow n'} = \frac{1}{\hbar^2}|\left\langle n'\right| A^\dagger \left|n\right\rangle|^2 f(\omega_{n'n}-\omega,t). \tag{6.64}$$

This obviously differs with respect to the constant perturbation only in the shift for $\pm\omega$ of the function $f(\omega,t)$. Therefore, all the conclusions are analogous to that case: instead of the conservation of energy (corresponding to the central peak of $f(\omega,t)$) the transition probability per unit time to the group of states around $E_n \mp \omega$ is constant, while vanishes for different final energies. All these results are easily generalized for the more complex periodic perturbations, having the Fourier form

$$H'(t) = \sum_s A_s \mathrm{e}^{\mathrm{i}s\omega t} + A_s^\dagger \mathrm{e}^{-\mathrm{i}s\omega t};$$

each term contributes to the emission/absorption of the quant $s\hbar\omega$.

6.5 Adiabatic approximation

6.5.1 General framework

External intervention on a system is described by time-dependent Hamiltonian. In many situations the dependence is included through some parameters, $\boldsymbol{s} = (s_1,\dots,s_n)$ varying smoothly as $\boldsymbol{s}(t)$ in time within some parameters manifold $\mathcal{M}$ (M-dimensional). Therefore, Hamiltonian gets form $H(t) = H(\boldsymbol{s}(t))$. The evolution becomes a process with time dependent Hamiltonian; the time dependence comes as dependence on the path $\boldsymbol{s}(t)$in $\mathcal{M}$ during the considered time interval (which is mapped to the path $H(\boldsymbol{s}(t))$ in the operators space). If $s(t)$ is "length" parameterizing the path from $s(0) = s_0$, then the time derivative is $\frac{\mathrm{d}}{\mathrm{d}t} = \dot{s}\frac{\mathrm{d}}{\mathrm{d}s}$ where $\dot{s} = \frac{\mathrm{d}s}{\mathrm{d}t}$ is "velocity" along the path[4].

Adiabatic picture

At each instance Hamiltonian eigensystem depends on the position on the path, and can be written as $H(s)\left|n\lambda;s\right\rangle = E_n(s)\left|n\lambda;s\right\rangle$. Eigenvalues $E_n(s)$ (such a function $E_n(s)$ will be called energy band) and eigenprojectors $P_n(s) = \sum_\lambda \left|n\lambda;s\right\rangle\left\langle n\lambda;s\right|$ (on the eigenspaces $\mathcal{S}_n(s)$) are only position (in $\mathcal{M}$) dependent, differentiable along the path, and in this sense these are purely geometrical quantities, independent of time.

Therefore, there exists path-parameterized family of unitary operators $W(s)$ interrelating eigenbases in s_0 and s:

$$W(s)\left|n\lambda;s\right\rangle = \left|n\lambda;s_0\right\rangle, \quad P_n(s) = W^\dagger(s)P_n(0)W. \tag{6.65a}$$

---

[4] Alternatively, $\frac{\mathrm{d}}{\mathrm{d}t} = \dot{\boldsymbol{s}}\cdot\nabla_{\boldsymbol{s}}$ can be used throughout.

This family is not unique[5]; the choice is fixed by initial value $W(0) = \mathbb{1}$ and infinitesimal change defined by Hermitian connection operator $\Omega(s)$:

$$\mathrm{i}\hbar\frac{\mathrm{d}W(s)}{\mathrm{d}s} = -W(s)\Omega(s) \text{ (i.e. } \mathrm{i}\hbar\frac{\mathrm{d}W^\dagger(s)}{\mathrm{d}s} = \Omega(s)W^\dagger(s)), \quad W(s_0) = \mathbb{1}. \tag{6.65b}$$

The operators $\Omega(s)$, which describe evolution of eigenprojectors by $\mathrm{i}\hbar\frac{\mathrm{d}P_n(s)}{\mathrm{d}s} = [\Omega(s), P_n(s)]$, are non-unique, as well: the last equation holds also for $\Omega'(s) = \Omega(s) + \sum_i P_i(s)F_i(s)P_i(s)$, with arbitrary (Hermitian) operator families $F_i(s)$. This freedom enables to choose $\Omega(s)$ such that $P_n(s)\Omega(s)P_n(s) = 0$, meaning that the infinitesimal change of each eigenspace is orthogonal to it. It can be shown (Exercise 6.5) that all these requirements satisfy

$$\Omega(t) \overset{\text{def}}{=} \mathrm{i}\hbar\sum_n P_n(s)\frac{\mathrm{d}P_n(s)}{\mathrm{d}s}. \tag{6.65c}$$

The time is introduced as a passage along the path. Family $W(s(t))$ defines the adiabatic picture, with equations following from the general results (Section 2.5):

$$U_W(s(t)) = W(s(t))U(t), \quad \mathrm{i}\hbar\frac{\mathrm{d}U_W(s(t))}{\mathrm{d}t} = (H_W(s(t)) + \dot{s}\Omega_W(s(t))U_W(s(t)). \tag{6.65d}$$

Observables in this picture get form (2.35): $A_W(s) = W(s)A(s)W^\dagger(s)$; in particular, all eigenprojectors are constant, $[P_n(s(t))]_W = P_n(0)$, yielding $H_W(s) = \sum_n E_n(s)P_n(0)$.

Adiabatic theorem

Next, consider a simple case, when Hamiltonian is equal to $H_W(s(t))$: its path dependence consists in energies varying as $E_n(s)$ (as for $H(s)$), but the eigensubspaces are preserved: $P_n(s) = P_n(0)$. The generated evolution is unitary family $\Phi(t)$ satisfying Schrödinger's equation (2.2):

$$\mathrm{i}\hbar\frac{\mathrm{d}\Phi(t)}{\mathrm{d}t} = H_W(t)\Phi(t), \quad \Phi(0) = \mathbb{1}. \tag{6.66}$$

Manifestly, corresponding connection (analogous to Ω) vanishes (since there is no transition to perpendicular eigensubspaces), while picture operator (analogous to W) is identity as the eigenprojectors are fixed). The equation (6.66) is easily integrated in the form $\Phi(t) = \sum_n \mathrm{e}^{-\mathrm{i}\phi_n(t)}P_n(0)$, showing that each state stationary for $H(0)$ remains stationary, evolving within the same eigenspace; it only obtains the dynamical phase $\phi_n(t) = \int_0^t \mathrm{d}\tau E_n(\tau)/\hbar$.

Otherwise, in general case, evolution involves transitions from one to another eigensubspace, and (6.65d) cannot be directly integrated as in the previous situation. However, the approach for time dependent perturbation may be applied for slow processes, putting $H_0 = H_\mathrm{i}$ and $H'(t) = H(t) - H_0$. If the energy difference $|E_n - E_{n'}|$ between the levels (at same position s) is large in comparison with $|\dot{s}\frac{\mathrm{d}E_n(s)}{\mathrm{d}s}|$, the probability of transition to other bands is suppressed (according to (6.61)), and for an infinitely slow process (with

---

[5] Clearly, this is related to the choice of eigenbasis of $H(s)$, meaning that arbitrary block-diagonal unitary $B = \oplus_n B_n$ (phases for nondegenerate energies) generates another family $W'(s) = W(s)B$.

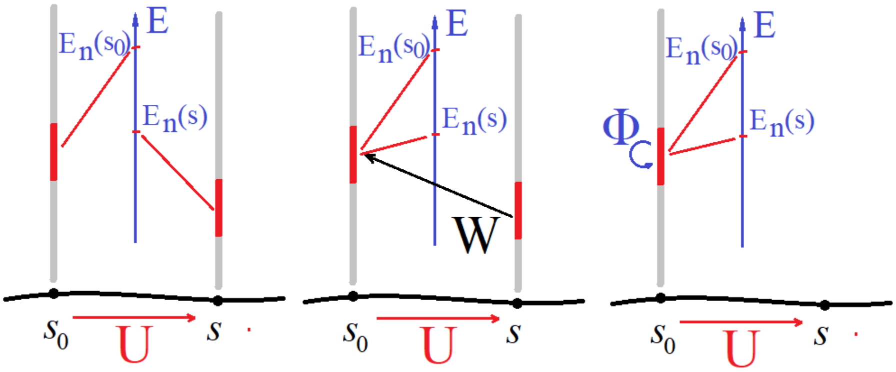

Figure 6.3: Adiabatic picture. Points s_0 and s of the parameter manifold each carry a copy of the system state space (gray) with emphasized eigensubspaces (red) interrelated by the evolution U (Schrödinger's picture, left). Transformation $W(s)$ maps the space associated o s into the space associated to s_0, establishing adiabatic picture. Φ is dynamics with the same evolution of eigenenergies, but the eigensusbspaces fixed (thus depicted in righ pannel, within a single copy of the state space in $\boldsymbol{s}_0$).

$\dot{s} = 0$) it exactly vanishes. In fact, a hint is given by the form of (6.65d): only the path-dependent (thus geometric) change Ω is scaled by $\dot{s}$, in contrast to the term H_W, which has fixed projectors and can be associated to single point in $\mathcal{M}$ (Figure 6.3). Therefore, in the limit of vanishing $\dot{s}$, the connection term in (6.65d) disappears. In the same limit U_W becomes Φ, giving in Schrödinger's picture $\lim_{\dot{s}\to 0} U(t) = W^\dagger(t)\Phi(t)$; this is in accordance with $\mathrm{i}\hbar \frac{\mathrm{d}W^\dagger(s(t))\Phi(t)}{\mathrm{d}t} = (H(t) + \dot{s}\Omega(s(t)))W^\dagger(s(t))\Phi(t)$, which is exact Schrödinger's equation when $\dot{s} = 0$. This is adiabatic behaviour, which can be formulated[13] also as:

Theorem 6.6 ***(Adiabatic theorem)*** *In the course of infinitely slow change of the parameters of Hamiltonian, the system being in the eigenstate $|n = n_i, t_\mathrm{i}\rangle$ of the initial Hamiltonian evolves to the derived by continuity eigenstate $|n = n_f, t_\mathrm{f}\rangle$ of the final Hamiltonian:*

$$\left(\lim_{\dot{s}\to 0} U(t)\right) P_n(0) = P_n(t) \left(\lim_{\dot{s}\to 0} U(t)\right). \tag{6.67}$$

Exercise 6.5: $\{P_n(s) \mid n = 1, 2, ...\}$ is a smooth (in s) family of the orthogonal decompositions of the state space, $\sum_n P_n(s) = \mathbb{1}$. Show:

(a) $P_n(s)\frac{\mathrm{d}P_n(s)}{\mathrm{d}s}P_n(s) = 0$;

(b) $\sum_n P_n(s)\frac{\mathrm{d}P_n(s)}{\mathrm{d}s} = -\sum_n \frac{\mathrm{d}P_n(s)}{\mathrm{d}s}P_n(s)$;

(c) $P_n(s)\Omega(s)P_n(s) = 0$ and $\mathrm{i}\hbar\frac{\mathrm{d}P_n(s)}{\mathrm{d}s} = [\Omega(s), P_n(s)]$ for $\Omega(s) \overset{\mathrm{def}}{=} -\mathrm{i}\hbar\sum_n P_n(s)\frac{\mathrm{d}P_n(s)}{\mathrm{d}s}$.

Adiabatic approximation

Adiabatic theorem refers to the adiabatic limit, when it is exact. Still, for slow (but not infinitely) processes it is used approximately in the form $U(t) \approx (\lim_{\dot{s}\to 0} U(t))$. The approximation describes the system forced to preserve the energy levels linked by continuity, i.e. it forbids the transition between them. It is minimal change of original evolution adapted to this property.

6.5.2 Geometric (Berry's) phase

According to the adiabatic theorem, for a non-degenerate[6] band $E_n(s)$, the eigenvectors $|n;s_0\rangle$ of $H(s_0)$ evolve to the eigenvectors $|n,s(t)\rangle$ of $H(s(t))$. In the special case of the evolution by Φ, it is shown that the obtained phase is purely dynamical $\phi_n(t)$. In general adiabatic case, $U = W^\dagger\Phi$, and (still nondegenerate) $|n;s_0\rangle$ gets additional phase from $W^\dagger$: $|n;s_0,t\rangle = W^\dagger(t)\Phi(t)\,|n;s_0\rangle$ satisfies $\frac{\mathrm{d}|n;s_0,t\rangle}{\mathrm{d}t} = -\frac{\mathrm{i}}{\hbar}(\dot{s}\Omega(s(t)) + E_n(t))\,|n;s_0,t\rangle$; this is solved as $|n;s_0,t\rangle = \mathrm{e}^{\mathrm{i}\gamma_n(t)}\mathrm{e}^{-\mathrm{i}\phi_n(t)}\,|n;s(t)\rangle$, where *geometric* of Berry's phase[4] $\gamma_n(t)$ satisfies $\dot{\gamma}_n(t) = \mathrm{i}\,\langle n,s(t)|\,\frac{\mathrm{d}}{\mathrm{d}s}\,|n,s(t)\rangle\,\dot{s}$.

However, the states at $\boldsymbol{s}$ are taken with arbitrary phase, meaning that a local smooth phase shift $|n,s\rangle \to \mathrm{e}^{\mathrm{i}\alpha(s)}\,|n,s\rangle$ is equally well choice. This clearly affect the geometric phase: $\gamma(s) \to \gamma(s) + \alpha(s_0) - \alpha(s)$. Still, it is not possible to identify all geometrical phases by choice of phase shift. Even on the closed path C it may happen that γ_C cannot be annihilated by phase shift, though this cannot depend on the choice of the phase shifts along the path. In this sense, characterization of loops by set of Berry's phases $\gamma(C) = \{\gamma_n(C)\}$ is essentially a topological characterization of the band structure of Hamiltonian.

6.5.3 Sudden approximation

When Hamiltonian abruptly changes, the time interval necessary for it vanishes. Therefore, Dyson's expansion (2.3) ends with the first term, with unit evolution operator describing this instantaneous process. The state of the system is not changed. Clearly, afterward evolution with changed Hamiltonian continues.

When a process is rapid, still not abrupt, the same approach is applicable only approximately. An estimate of the introduced error is given by the next terms of Dyson's series.

6.5.4 Born-Oppenheimer approximation

Unavoidable and the most famous application of the adiabatic theory appears when complex matter, consisting of ions and electrons is considered[7]. The great difference in masses of the two types of ingredients enables specific approximate separations of degrees of freedom related to the two classes. Let $\boldsymbol{R} = (R_{1x},\dots,R_{Nz})$ denotes the set of all the coordinates of ions, while $\boldsymbol{r} = (r_{1x},\dots,r_{Nz})$ are the coordinates of electrons. Total Hamiltonian is:

$$H = T_{\boldsymbol{R}} + T_{\boldsymbol{r}} + V(\boldsymbol{R},\boldsymbol{r}) = H_{\boldsymbol{r}} + T_{\boldsymbol{R}}, \tag{6.68a}$$

where $T_{\boldsymbol{R}} = \sum_{I=1}^{N} \frac{\hbar^2}{2M_I}\triangle_{\boldsymbol{R}}$, and $T_{\boldsymbol{r}} = \sum_{r=1}^{n} \frac{\hbar^2}{2m}\triangle_{\boldsymbol{r}}$ are ionic and electronic kinetic energies, while the total interaction $V(\boldsymbol{R},\boldsymbol{r}) = V_{\mathrm{ion}}(\boldsymbol{R}) + V_{\mathrm{el}}(\boldsymbol{r}) + V_{\mathrm{ie}}(\boldsymbol{R},\boldsymbol{r})$ is the total interaction consisting of ionic, electronic and electron-ion part. This Hamiltonian gives total Schrödinger equation:

$$H\Psi(\boldsymbol{R},\boldsymbol{r}) = E\Psi(\boldsymbol{R},\boldsymbol{r}). \tag{6.68b}$$

---

[6] In the case of degeneracy, there is a generalization of the geometric phase [18].

[7] In some references adiabatic approximation is identified to Born-Oppenheimer one.

Electronic Hamiltonian

$$H_{\boldsymbol{r}} = T_{\boldsymbol{r}} + V(\boldsymbol{R}, \boldsymbol{r}), \tag{6.69a}$$

does not include ionic kinetic energy, and all other terms commute with ionic coordinate operators[8]. Therefore, it can be understood as a family of the electronic Hamiltonians, parameterized by the ionic configurations $\boldsymbol{R}$, which establishes the framework of the adiabatic model and approximation. Accordingly, the solutions of the electronic Schrödinger equation

$$H_{\boldsymbol{r}}\eta_i(\boldsymbol{R}, \boldsymbol{r}) = \epsilon_i(\boldsymbol{R})\eta_i(\boldsymbol{R}, \boldsymbol{r}), \tag{6.69b}$$

are also parameterized by $\boldsymbol{R}$: for each ionic position $\boldsymbol{R}$ there are electronic states, making basis in the electronic state space, and energies. Accordingly, in this basis, $|\boldsymbol{R}\rangle\,|\boldsymbol{R}; i\rangle$ (the first factor is in ionic space, the second one is basis of the electronic space for ions fixed in $\boldsymbol{R}$), the *adiabatic representation* of the total eigenfunction is expanded as

$$\Psi(\boldsymbol{R}, \boldsymbol{r}) = \sum_i \chi_i(\boldsymbol{R})\eta_i(\boldsymbol{R}, \boldsymbol{r}). \tag{6.70}$$

Note that no approximation is applied yet: this is only *adiabatic picture*[9], where for each $\boldsymbol{R}$ the expansion coefficients are ionic functions $\chi_i(\boldsymbol{R})$. To find them, we substitute (6.70) in (6.68b), apply (6.69b), multiply by η_j^* and finally integrate over $\boldsymbol{r}$, to find out:

$$[T_{\boldsymbol{R}} + \epsilon_j(\boldsymbol{R}) - E]\,\chi_j(\boldsymbol{R}) = \sum_i \Lambda_{ji}\chi_i(\boldsymbol{R}).$$

Matrix Λ has the elements

$$\Lambda_{ji} = \sum_I \frac{\hbar}{2M_I}\int \mathrm{d}\boldsymbol{r}\eta_j^*(\boldsymbol{R}, \boldsymbol{r})\frac{\partial^2\eta_i(\boldsymbol{R}, \boldsymbol{r})}{\partial \boldsymbol{R}_I^2} + \left(\sum_I \frac{\hbar}{M_I}\int \mathrm{d}\boldsymbol{r}\eta_j^*(\boldsymbol{R}, \boldsymbol{r})\frac{\partial\eta_i(\boldsymbol{R}, \boldsymbol{r})}{\partial \boldsymbol{R}_I}\frac{\partial\chi_i(\boldsymbol{R})}{\partial \boldsymbol{R}_I}\right).$$

It comprises all the elements of H non-diagonal in electronic basis functions $\eta_i(\boldsymbol{R}, \boldsymbol{r})$, thus enabling transition between the electronic eigensubspaces during the evolution of the whole system. Obviously, for large ionic masses it is small, which justifies neglecting Λ, which is exactly adiabatic approximation. Then the previous equation becomes ionic Schrödinger equation:

$$[T_{\boldsymbol{R}} + \epsilon_j(\boldsymbol{R})]\,\chi_{mj}(\boldsymbol{R}) = E_{mj}\chi_{mj}(\boldsymbol{R}).$$

This equation contains no electronic degrees of freedom. On the other hand, the electrons appears here through their eigenenergies, which have the role of the external potential defining the ion dynamics.

To summarize, adiabatic approximation suffices to separate electron and ions. Within this model, the electronic equation is to be solved first, and their energies give the potential in which ions move. To study the ground state of the total system, one takes the ground

---

[8] Alternatively, for some purposes also $V_{\mathrm{ion}}(\boldsymbol{R})$ can be omitted from $H_{\boldsymbol{r}}$, with the same conclusions.

[9] In fact, in the Hilbert space $\mathcal{S} = \mathcal{L}(\boldsymbol{R}) \otimes \mathcal{L}(\boldsymbol{r})$ the basis $|\boldsymbol{R}\rangle \otimes |\eta_i, \boldsymbol{R}\rangle$ is used, showing that $\mathcal{S}$ is considered as a fiber bundle with base $\boldsymbol{R}$ and fibre $\mathcal{L}(\boldsymbol{r})$; this, and similar fibrations of the state space are in the background of the topological analyses frequent in modern physics.

state of electrons, with energy ϵ_0, and then solves equation of ions in the external potential ϵ_0.

Despite successful reduction of the problem to the separate equations in electrons and ions, each of these subsystems is usually complex enough, and further approximations are employed. While for the electrons various approximations will be considered later, here we mention that for ions it is usual to adopt Born approximation. It assumes that the ions are in the position $\boldsymbol{R}_0$ corresponding to minimum of the potential ϵ_0. Then, their dynamics is by harmonic approximation reduced to vibrations around this minimum.

6.6 Elementary scattering theory

Collisions of some sort of the particles with a target is one of the most frequent experimental methods to examine the properties of the target and interactions. The scheme of the collision is simple (Fig. **??**): a collimated beam of particles is directed toward the target. After that the scattered particles are detected and studied, and according to this many conclusions about the target are derived. There are various types of the collision processes: spallation and reactions (when some completely new particles are scattered), attachment (the target and the incoming particle together form a new particle) and scattering (the simplest one, when the outgoing particle is the same one as incoming). Particularly, here we are interested in the elastic scattering, in which the kinetic energy of the incoming particle is preserved.

In the experiments with elastic scattering the distribution $\mathrm{d}n(\theta,\varphi)$ of the outgoing particles over the directions is measured. Obviously, this is proportional to the flux Φ_{in} of the incoming (along the z-axis) particles:

$$\mathrm{d}n(\theta,\varphi) = \sigma(\theta,\varphi)\Phi_{\mathrm{in}}\,\mathrm{d}\Omega, \tag{6.71}$$

where the differential cross section $\sigma(\theta,\varphi)$ is the proportionality factor, which comprises all the details on the types of the beam and target, interactions, etc. Note that $\sigma(0,\varphi)$ is not measured quantity, since experimentally it is not easy to distinguish between the scattered particles with $\theta = 0$ scattering angle, and the particles which simply passed by target without interaction (thus in the same direction). If the interaction between beam and target depends only on the particle distance (spherically symmetric), then the cross section is φ-independent. Therefore, it can be integrated over φ, giving $\sigma(\theta) = \int_0^{2\pi}\sigma(\theta,\varphi)\,\mathrm{d}\varphi = 2\pi\sigma(\theta,0)$. Total cross section is the integral $\sigma = \int_0^{\pi}\int_0^{2\pi}\sigma(\theta,\varphi)\,\mathrm{d}\Omega$. Cross sections have the dimension of surface, and usually are measured in Barns: $1\mathrm{Barn}=10^{-24}\mathrm{cm}^2$.

Within the quantum theory, it is possible to give some general relations on the scattering cross section, without detailed analysis of the interaction. To this end the stationary regime is to be considered, when the state $\psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}Et/\hbar}$ of beam is stationary for Hamiltonian $H = T + V(r)$ describing a beam particle interacting with target only. The interaction potential is assumed to decrease with r. The time independent function is of the form (it is not normalized)

$$\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}kz} + f(\theta,\varphi)\frac{\mathrm{e}^{\mathrm{i}kr}}{r}, \tag{6.72}$$

being a superposition of the incoming state (plane wave along z-axis) and the scattered state (spherical wave). It asymptotically (for large r) satisfies time independent Schrödinger equation. The function $f(\theta, \varphi)$ (or $f(\theta)$ for the spherically symmetric interaction) is called scattering amplitude. The number of scattered particles is given by (6.71), with the incoming beam particles in the state of plane wave. Thus their flux is proportional to the probability current (2.21): $\Phi_{\mathrm{in}} = c\hbar k/m$. On the other side, $\mathrm{d}n(\theta, \varphi)$ is proportional to the flux of outgoing particles:

$$\mathrm{d}n(\theta, \varphi) = c\boldsymbol{j}^{\mathrm{out}} \cdot \mathrm{d}\boldsymbol{S} = cj_r^{\mathrm{out}} r^2 \, \mathrm{d}\Omega.$$

The outgoing probability current is easily calculated applying (2.20) to the outgoing state $\psi_{\mathrm{sw}}(\boldsymbol{r}) = f(\theta, \varphi)\mathrm{e}^{\mathrm{i}kr}/r$. The leading term in r is r^{-2}, and in the asymptotic region (which is studied here) only r-component of current appears:

$$j_r^{\mathrm{out}} = \frac{1}{m}\mathrm{Re}(\psi^*(\boldsymbol{r}, t)\frac{\hbar}{\mathrm{i}}\partial_r \psi(\boldsymbol{r}, t)) = \frac{\hbar}{m}k\frac{|f(\theta, \varphi)|^2}{r^2}.$$

Using these two expressions for the number of outgoing particles one finally finds the differential cross section:

$$\sigma(\theta, \varphi) = |f(\theta, \varphi)|^2. \tag{6.73}$$

This expression reveals how the experimental results (expressed by $\sigma(\theta, \varphi)$) are related to the basic quantum mechanical object (the state, given by $f(\theta, \varphi)$). Therefore it is the starting point of the quantum scattering theory, which is very elaborated (and a subject of separate courses). Logically, as a theory relating incoming (beam and target) with outgoing (again beam and target after interaction) it is also a way (equivalent to the previously described) to express the evolution.

Chapter 7

Elements of quantum information

7.1 Qubit space and Bloch sphere representation

Some facts about the simplest nontrivial (with non-commuting operators) space, two-dimensional or *qubit space* $\mathbb{C}^2$ are considered here. As mentioned (Section 4.5.3), suitable hermitian basis of the corresponding operator space (four-dimensional) is formed by the identity matrix $\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and three Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7.1}$$

For the Pauli matrices the following properties are easily verified:

$$\sigma_i^2 = \sigma_0, \quad \mathrm{Tr}\,\sigma_i = 0, \quad \mathrm{Det}\,\sigma_i = -1, \quad \sigma_i\sigma_j + \sigma_j\sigma_i = 2I\delta_{ij}, \quad \sigma_1\sigma_2\sigma_3 = \mathrm{i}\sigma_0. \tag{7.2}$$

It follows that the basis σ_i $(i = 0, \dots, 3)$ is orthogonal with respect to the operator scalar product: $\mathrm{Tr}\,\sigma_i\sigma_j = 2\delta_{ij}$. Therefore, each linear operator can be expanded in the form $A = \sum_{i=0}^{3} a_i\sigma_i$, with complex coefficients $a_i = \frac{1}{2}\mathrm{Tr}\,\sigma_i A$. The last three coefficients can be grouped as a vector $\boldsymbol{a}$, to get $A = a_0\sigma_0 + \boldsymbol{a}\cdot\boldsymbol{\sigma}$. The second factor $\boldsymbol{a}\cdot\boldsymbol{\sigma}$, gives the traceless part of A, while the first one gives the scalar part with trace $2a_0$. For the hermitian operators, the coefficients a_i are real.

Exercise 7.1: Show that $A^2 = 2a_0\boldsymbol{a}\cdot\boldsymbol{\sigma} + \sum_{i=0}^{3} a_i^2$ for $A = \sum_{i=0}^{3} \boldsymbol{a}_i\boldsymbol{\sigma}_i$. When A is a projector?

In particular, the one-dimensional projectors are $P_{\boldsymbol{a}} = \frac{1}{2}(\sigma_0 + \boldsymbol{a}\cdot\boldsymbol{\sigma})$, with arbitrary unit vector $\boldsymbol{a}$ (see Exercise 7.1). Thus, each point $\boldsymbol{a}$ on the unit Bloch's sphere represents a physical pure state. Representing $\boldsymbol{a}$ by the spherical coordinates $(1, \vartheta, \varphi)$ on the Bloch's sphere, $\boldsymbol{a}(\vartheta, \varphi)$, the projector $P_{\boldsymbol{a}}$ has unique, up to a phase, eigenvector (for the eigenvalue 1) $|\boldsymbol{a}; +\rangle = (\cos\frac{\vartheta}{2}, \mathrm{e}^{\mathrm{i}\varphi}\sin\frac{\vartheta}{2})$. The equality $\mathrm{Tr}\,P_{\boldsymbol{a}}P_{\boldsymbol{b}} = \frac{1}{2}(1 + \boldsymbol{a}\cdot\boldsymbol{b})$ means also that the projectors are orthogonal for the antipodal points, $\boldsymbol{b} = -\boldsymbol{a}$, on the sphere (with $\vartheta_{\boldsymbol{b}} = \pi - \vartheta_{\boldsymbol{a}}$ and $\varphi_{\boldsymbol{b}} = \pi + \varphi_{\boldsymbol{a}}$). Hence, the vector $|\boldsymbol{a}; -\rangle = |-\boldsymbol{a}; +\rangle$ spans the null-space of $P_{\boldsymbol{a}}$ and together with $|\boldsymbol{a}; +\rangle$ makes an orthonormal basis in $\mathbb{C}^2$. The operators $\rho_{\boldsymbol{a}} = \frac{1}{2}(\sigma_0 + \boldsymbol{a}\cdot\boldsymbol{\sigma})$ with $|\boldsymbol{a}| < 1$ are the non-trivial mixtures.

Hence, in the Bloch representation, qubits are represented as vectors of the Bloch ball. In particular, the pure states are represented by vectors of the Bloch sphere (unit vectors).

Exercise 7.2: Show that the spaces of traceless and of scalar observables are orthocomplements.

Exercise 7.3: a) Show that $\langle \boldsymbol{x}| \, \boldsymbol{a}\cdot\boldsymbol{\sigma} \, |\boldsymbol{x}\rangle = \boldsymbol{a}\cdot\boldsymbol{x}$ for $\boldsymbol{a}$ and $\boldsymbol{x}$ from $\mathbb{R}^3$ and $|\boldsymbol{x}| = 1$.
b) Show that operator $\rho_{\boldsymbol{a}} = \frac{1}{2}(\sigma_0 + \boldsymbol{a}\cdot\boldsymbol{\sigma})$ with $|\boldsymbol{a}| \leq 1$ is positive-definite and statistical, with $|\boldsymbol{a}| = 0$ and $|\boldsymbol{a}| = 1$ corresponding to maximally ($\rho_0 = \frac{1}{2}$) and minimally (pure) mixed states.

7.2 Bipartite systems: Schmidt decomposition

Theorem 7.1 ***(Schmidt decomposition)*** *Pure composite state $|\Psi\rangle$ of a bipartite system has the reduced states with mutually equivalent spectral forms*

$$\rho_i \overset{\mathrm{def}}{=} \mathrm{Tr}_{\hat{i}} \, |\Psi\rangle \langle\Psi| = \sum_s r_s \, |i;s\rangle \langle i;s| \quad (i = 1,2), \tag{7.3a}$$

where convex set $\{r_1, r_2, \dots\}$ and orthonormal vectors $|i;s\rangle$ determine decomposition

$$|\Psi\rangle = \sum_s \sqrt{r_s} \, |1;s\rangle \otimes |2;s\rangle \, . \tag{7.3b}$$

▬*Proof:* This is a particular form of the singular value decomposition (Exercise 7.4). Indeed, the composite state space $\mathcal{S} = \mathcal{S}_1 \otimes \mathcal{S}_2$ is bijective to the space of the linear operators from $\mathcal{S}_2$ into $\mathcal{S}_1$, with an identification: $|\Psi\rangle = \sum_s \sigma_s \, |u_s\rangle \otimes |v_s\rangle$ and $\Psi = \sum_s \sigma_s \, |u_s\rangle \otimes \langle v_s|$. Then $\rho = |\Psi\rangle \langle\Psi|$ gives the partial traces $\rho_1 = \sum_s \sigma_s^2 \, |u_s\rangle \langle u_s|$ and $\rho_2 = \sum_s \sigma_s^2 \, |v_s\rangle \langle v_s|$. ∎

Exercise 7.4: Prove singular value decomposition (SVD): Any $n_1 \times n_2$ matrix can be factorized as $M = U\Sigma V^{+}$, where U and V are $n_1 \times n_1$ and $n_2 \times n_2$ unitary matrices, while Σ is $n_1 \times n_2$ diagonal matrix with only the first r positive elements $\sigma_i > 0$ $(i = 1, \dots, r \leq n_1, n_2)$ called *singular values.*

Schmidt decomposition theorem shows that the reduced states ρ_i are mutually equivalent on their respective ranges (while the null-spaces may differ). Precisely, these ranges $\mathcal{R}(\rho_i)$ are subspaces of the same dimension in the subsystems' state spaces $\mathcal{S}_i$, and there is a bijection of the subsystems' operators supported on these ranges. In particular, such paired observables are called *twins*, and subsystem measurement of one of them is automatically a measurement of the other one, with the corresponding reduction [17].

7.3 Entanglement, hidden variables, Bell's theorem

Two-cubit antisymmetric state $|\Psi-\rangle = \frac{|+-\rangle - |-+\rangle}{\sqrt{2}}$ serves as a simple example of entanglement and its (sometimes unexpected) features. It is important that the state $|\Psi-\rangle$ is unique in the sense that arbitrary (orthonormal) basis of $\mathbb{C}^2$ gives the same state (up to a phase; see Exercise 7.5). In the context of this Section, this state is associated to the pair of photons with opposite polarizations, which propagate in opposite directions (only their polarization space is explicitly taken into account).

Exercise 7.5: Prove uniqueness of the antisymmetric state: for the orthonormal bases $|a;i\rangle$ and $|b;i\rangle$: $|a;+\rangle \, |a;-\rangle - |a;-\rangle \, |a;+\rangle = \mathrm{e}^{\mathrm{i}\phi}(|b;+\rangle \, |b;-\rangle - |b;-\rangle \, |b;+\rangle)$.

Assume that the experimental setup consists of the emitter of the pairs of photons in the joint state $|\Psi-\rangle$ and a pair of the (possibly distant) detectors with polarizations set along the directions $\boldsymbol{a}$ and $\boldsymbol{b}$. Photons coincidentally come to the detectors, while decision about $\boldsymbol{a}$ and $\boldsymbol{b}$ can be made after the particles are prepared (and emitted) in the state $|\Psi-\rangle$. Each detector has possible outcomes ±1, corresponding to polarization paralel/antiparalel to its direction ($\boldsymbol{a}$ and $\boldsymbol{b}$). Thus, the outcome of the described coincidental measurement is a pair (a,b) of these values, and the result of the experiment is the probability distribution $v((a,b),(A,B),|\Psi\rangle)=\frac{n_{ab}}{n}$ (with $n=\sum_{a,b=\pm1}n_{ab}$). These experimental results yield statistical correlation defined as the average of ±1 for coincident/opposite outcomes:

$$C(\boldsymbol{a},\boldsymbol{b})=-\sum_{ab}(-1)^{\frac{a+b}{2}}v((a,b),(A,B),|\Psi\rangle)=\frac{n_{++}-n_{+-}-n_{-+}+n_{--}}{n}. \tag{7.4}$$

Quantum mechanics treats the apparatuses as the observables $A=(\boldsymbol{a}\cdot\boldsymbol{\sigma})\otimes\mathbb{1}$ and $B=\mathbb{1}\otimes(\boldsymbol{b}\cdot\boldsymbol{\sigma})$, with eigenvalues ±1: these are local detectors, which measure particle polarization along them, and the outcomes ±1, evidencing parallel/antiparallel polarization, correspond to the single-particle eigenvectors $|\boldsymbol{a};\pm\rangle$ and $|\boldsymbol{b};\pm\rangle$ (in the Bloch sphere notation for $\mathbb{C}^2$). In the initial state $|\Psi-\rangle$ the quantum mechanical prediction is:

$$C_{\mathrm{qm}}(\boldsymbol{a},\boldsymbol{b})=\langle\Psi-|\,(\boldsymbol{a}\cdot\boldsymbol{\sigma}\otimes\mathbb{1})(\mathbb{1}\otimes\boldsymbol{b}\cdot\boldsymbol{\sigma})\,|\Psi-\rangle=-\boldsymbol{a}\cdot\boldsymbol{b}. \tag{7.5}$$

***Exercise 7.6**:* Prove (7.5).

As discussed above, the state $|\Psi-\rangle$ is invariant with respect to the chosen single-particle basis. Thus, coincidental outcomes with the antiparallel polarizations (and other possibilities) are retained even in the case when $\boldsymbol{a}=-\boldsymbol{b}$ is determined after the photons are prepared (sent to detectors), which means that the result on the first apparatus instantaneously determines the second one (Einstein's "spooky action at a distance"). To remedy this, the existence of hidden variables (there is also a number of related notions in the literature, starting [6] with "element of reality" as a property which is in any experiment certainly observed) is considered. Namely, it is conceivable that there is an underlying quantity $\lambda\in\Lambda$ determining uniquely the outcomes (one of ±1) on both sides, which itself is defined at the instance of separation of particles, thus before the measurements are coincidentally performed and polarizations decided. Then, for each value of λ, and the experimental choice $\boldsymbol{a}$ and $\boldsymbol{b}$, there are functions $A(\boldsymbol{a},\lambda)$ and $B(\boldsymbol{b},\lambda)$ giving the detected values ±1. Therefore, if $p(\lambda)$ is probability to prepare system described by λ, the hidden variable model predicts correlation:

$$C_{\mathrm{hv}}(\boldsymbol{a},\boldsymbol{b})=\int_{\Lambda}A(\boldsymbol{a},\lambda)B(\boldsymbol{b},\lambda)p(\lambda)d\lambda. \tag{7.6}$$

7.3.1 Bell's inequalities

It turns out that the hidden variable theories can be ruled out by the experiment. In fact, Bell [3] noticed that within such theories some quantities satisfy so called Bell's inequalities, which are broken by the quantum mechanical prediction. There is a number of these inequalities, each suited to some experiment; the most famous and illustrated bellow, also refereed as Bell's, is proposed [5] few years later.

It is assumed that on the both sides the apparatuses may be in either of two positions corresponding to measuring polarization $\boldsymbol{a}$ and $\boldsymbol{a}'$ on one, and $\boldsymbol{b}$ or $\boldsymbol{b}'$ on the other side. Consider the quantity $S = C(\boldsymbol{a},\boldsymbol{b}) - C(\boldsymbol{a},\boldsymbol{b}') + C(\boldsymbol{a}',\boldsymbol{b}) + C(\boldsymbol{a}',\boldsymbol{b}')$. Within a hidden variable approach, S_{hv} for the state $|\Psi-\rangle$ is easily calculated. In fact, $A(\boldsymbol{a},\lambda)$ and $B(\boldsymbol{b},\lambda)$, $A(\boldsymbol{a}',\lambda)$ and $B(\boldsymbol{b}',\lambda)$ are ± 1 depending on λ, making that among $B(\boldsymbol{b},\lambda) + B(\boldsymbol{b}',\lambda)$ and $B(\boldsymbol{b},\lambda) - B(\boldsymbol{b}',\lambda)$ one takes value 0, and the other 2 or -2. Hence, shortening notation of $X(\boldsymbol{x},\lambda)$ by $X_{\boldsymbol{x}}$, this gives

$$A_{\boldsymbol{a}}B_{\boldsymbol{b}} - A_{\boldsymbol{a}}B_{\boldsymbol{b}'} + A_{\boldsymbol{a}'}B_{\boldsymbol{b}} + A_{\boldsymbol{a}'}B_{\boldsymbol{b}'} = A_{\boldsymbol{a}}(B_{\boldsymbol{b}} - B_{\boldsymbol{b}'}) + A_{\boldsymbol{a}'}(B_{\boldsymbol{b}} + B_{\boldsymbol{b}'}) \leq 2,$$

and averaging by (7.6) in λ yields:

$$S_{\text{hv}}(\boldsymbol{a},\boldsymbol{a}',\boldsymbol{b},\boldsymbol{b}') = C_{\text{hv}}(\boldsymbol{a},\boldsymbol{b}) - C_{\text{hv}}(\boldsymbol{a},\boldsymbol{b}') + C_{\text{hv}}(\boldsymbol{a}',\boldsymbol{b}) + C_{\text{hv}}(\boldsymbol{a}',\boldsymbol{b}') \leq 2.$$

Quantum approach (7.5), with polarizations $\boldsymbol{a} = \boldsymbol{e}_z$, $\boldsymbol{a}' = \boldsymbol{e}_x$ (or $A = \sigma_z$, $A' = \sigma_x$), $\boldsymbol{b} = (1, \vartheta = \frac{3\pi}{4}, \varphi = \pi)$ and $\boldsymbol{b}' = (1, \frac{\pi}{4}, \pi)$ (with $B = \frac{\sigma_z+\sigma_x}{\sqrt{2}}$ and $B' = \frac{\sigma_z-\sigma_x}{\sqrt{2}}$), gives $C_{\text{qm}}(\boldsymbol{a},\boldsymbol{b}) = -C_{\text{qm}}(\boldsymbol{a},\boldsymbol{b}') = C_{\text{qm}}(\boldsymbol{a}',\boldsymbol{b}) = C_{\text{qm}}(\boldsymbol{a}',\boldsymbol{b}') = \frac{1}{\sqrt{2}}$. Hence, quantum prediction $S_{\text{qm}}(\boldsymbol{a},\boldsymbol{a}',\boldsymbol{b},\boldsymbol{b}') = 2\sqrt{2}$ in this case violates Bell's equation for hidden variables. Experiment supports quantum prediction.

Exercise 7.7: Show the *no-signaling theorem*(*no-communication theorem*): measurement of subsystem observable A does not affect the results of the measurement of arbitrary complementary subsystem observable B.

Exercise 7.8: Show the *no-cloning theorem*: There is no unitary operator U in $\mathcal{S}_1 \otimes \mathcal{S}_1$ such that $U|a\rangle \otimes |b\rangle = \mathrm{e}^{\mathrm{i}f(a,b)}|a\rangle \otimes |a\rangle$ for arbitrary $|a\rangle$ and $|b\rangle$.

Appendix A

Technical support

A.1 Separation of Variables

Here we give simple algebraic spectral theorem on the operators acting in the direct product space $\mathcal{S} = \mathcal{S}_1 \otimes \mathcal{S}_2$. The aim is to facilitate the solution of the eigenproblem, reducing it to the factor spaces for two special, but frequent in quantum mechanical problems forms of the considered operators: $A = B_1 \otimes B_2$ and $A = B_1 \otimes B_2 + C_1 \otimes C_2$. While the former is given for the completeness only, in the view of its self obvious outcome and considerably rare physical situations when it is applicable, the latter one refers to wide range of examples.

Lemma A.1 (Separation of variables) *Let $\mathcal{S} = \mathcal{S}_1 \otimes \mathcal{S}_2$.*

(a) *If $A = B_1 \otimes B_2$ with spectral forms $B_i = \sum_{b_i} b_i P_{b_i}(B_i)$ $(i = 1, 2)$, where $P_{b_i}(B_i)$ are the spectral projectors of the factor operators, then the spectral form of A is:*

$$A = \sum_{b_1, b_2} b_1 b_2 P_{b_1}(B_1) \otimes P_{b_2}(B_2). \tag{A.1}$$

Geometrically, this means that the space $\mathcal{S}$ is decomposed as the direct sum $\mathcal{S} = \oplus_{b_1,b_2} \mathcal{S}_1^{(b_1)} \otimes \mathcal{S}_2^{(b_2)}$ of the eigenspaces $\mathcal{S}_1^{(b_1)} \otimes \mathcal{S}_2^{(b_2)}$ of A for the eigenvalues $b_1 b_2$ (not necessarily different; hence, to achieve the spectral decomposition one should rearrange the sum firstly, to group together the same eigenvalues), each of these subspaces being the product of the eigenspaces of the factor operators.

(b) *If $A = B_1 \otimes B_2 + C_1 \otimes C_2$, and B_2 and C_2 are compatible, and the projectors to their common eigensubspaces $\mathcal{S}_2^{(b_2,c_2)}$ are $P_{(b_2,c_2)}$, then*

$$A = \sum_{(b_2,c_2)} \sum_{a(b_2,c_2)} a(b_2, c_2) P_{a(b_2,c_2)}(b_2 B_1 + c_2 C_1) \otimes P_{b_2 c_2}. \tag{A.2}$$

In other words, $\mathcal{S}$ is decomposed into the direct sum $\mathcal{S} = \sum_{(b_2,c_2)} \sum_{a(b_2,c_2)} \mathcal{S}_1^{a(b_2,c_2)} \otimes \mathcal{S}_2^{(b_2,c_2)}$ of the products of the eigenspaces $\mathcal{S}_1^{a(b_2,c_2)}$ for the eigenvalues $a(b_2, c_2)$ (again, not necessarily different) of the operators $A_1(b_2, c_2) = b_2 B_1 + c_2 C_1$ and common eigenspaces $\mathcal{S}_2^{(b_2,c_2)}$ of B_2 and C_2.

(c) *Let $D = A \otimes B$ commutes with C, and $|b\beta\rangle$ and $|c\gamma\rangle$ the eigenbases of B and C (eigenvalues b and c. Then a common eigenbasis of D and C is $\sum_{b\beta} |a; c\gamma, b\beta\rangle \otimes |b\beta\rangle$, where $\{|a; c\gamma, b\beta\rangle = \mid \gamma = 1, \dots, |\mathcal{S}_c|\}$ is eigenbasis of $A_c^{(b\beta)}$ defined as the reduced A in the invariant subspace $\mathcal{S}_c^{(b\beta)}$ of $\mathcal{S}_1$ spanned by $\langle b\beta \mid c\gamma\rangle$.*

▬*Proof:* ***(a)*** The direct products $|b_1; \lambda_1\rangle \otimes |b_1; \lambda_2\rangle$ of the eigenvectors of B_1 and B_2 are basis in $\mathcal{S}$. Thus they are an eigenbasis of A, since $A\,|b_1; \lambda_1\rangle \otimes |b_1; \lambda_2\rangle = b_1 b_2\,|b_1; \lambda_1\rangle \otimes |b_1; \lambda_2\rangle$. Finally, for fixed b_1 and b_2 they span the subspaces $\mathcal{S}_1^{(b_1)} \otimes \mathcal{S}_2^{(b_2)}$ in the decomposition $\mathcal{S} = \oplus_{b_1,b_2} \mathcal{S}_1^{(b_1)} \otimes \mathcal{S}_2^{(b_2)}$.
(b) Operator A is reduced to $A_1(b_2, c_2) \otimes P_{b_2 c_2}$ in each of the spaces $\mathcal{S}_1 \otimes \mathcal{S}_2^{(b_2,c_2)}$, since $A\,|x\rangle\,|(b_2, c_2), \lambda\rangle = ((b_2 B_1 + c_2 C_1)\,|x\rangle) \otimes |(b_2, c_2), \lambda\rangle$ for each $|x\rangle$ from $\mathcal{S}_1$ and $|(b_2, c_2), \lambda\rangle$ from $\mathcal{S}_2^{(b_2,c_2)}$. Therefore, it remains to find the spectral form of $A_1(b_2, c_2)$ in $\mathcal{S}_1$ for each compatible pair (b_2, c_2).
(c) Clearly, $|c\gamma\rangle = \sum_{b\beta} \langle b\beta \mid c\gamma\rangle \otimes \otimes |b\beta\rangle$. Then, due to commutation, eigenspace $\mathcal{S}_c$ is invariant under D, which reduces to D^c such that $D\,|c\gamma\rangle = \sum_{\gamma'} (D^c)_\gamma^{\gamma'}\,|c\gamma'\rangle = \sum_{\gamma' b\beta} (D^c)_\gamma^{\gamma'}\,\langle b\beta \mid c\gamma'\rangle \otimes |b\beta\rangle$; on the other hand, $D\,|c\gamma\rangle = \sum_{b\beta} bA\,\langle b\beta \mid c\gamma\rangle \otimes |b\beta\rangle$, showing that D^c (b, β independent) is equal to bA in $\mathcal{S}_c^{(b\beta)}$. This implies that the last subspace is invariant under A, and the reduced operator $A_c^{(b\beta)}$ is β-independent. The eigenvectors $|a; c\gamma, b\beta\rangle$ ($\gamma = 1, 2, \dots$), when multiplied by $|b\beta\rangle$ are separable eigenvectors of D from the igenspace of C. ▮

A.2 Hypergeometric equation

The most important equations of quantum dynamics are second order differential one. Among them, the class of the *generalized hypergeometric equation* covers most of the problems considered in this textbook:

$$u''(t) + \frac{\tilde{\tau}(t)}{\sigma(t)} u'(t) + \frac{\tilde{\sigma}(t)}{\sigma(t)^2} u(t) = 0, \tag{A.3}$$

where $\sigma(t)$ and $\tilde{\sigma}(t)$ are polynomials of the order less then 3, and $\tilde{\tau}(t)$ a polynomial of the order less than 2. This class is solvable, with the solution standardized within so called special functions. Here we give the three step algorithm to find out these solutions. The method is based on the theory of differential operators in the underlying Lebesgue space $\mathcal{L}([a, b])$.

Step 1. Find all possible constants k such that

$$\pi(t) = \frac{\sigma' - \tilde{\tau}}{2} \pm \sqrt{\left(\frac{\sigma' - \tilde{\tau}}{2}\right)^2 - \tilde{\sigma} + k\sigma}$$

is polynomial (of the order less than 2). This means that we look for k giving the rooted expression the form $(at + b)^2$. For each (at most 2) such k, we find two polynomials $\pi_{k\pm}$, and the next steps are to be performed for each of them.

Step 2. For each $\pi_{k\pm}$ solve the first order differential equation in $\phi(t)$:

$$\frac{\phi'}{\phi} = \frac{\pi_{k\pm}}{\sigma},$$

and calculate polynomials $\tau_{k\pm} = 2\pi_{k\pm} - \widetilde{\tau}$ and constants $\lambda_{k\pm} = -k - \pi'_{k\pm}$. Substituting the solutions $\phi_{k\pm}$ in (A.3) by the change $u = \phi_{k\pm} y_{k\pm}$ we get *hypergeometric equation* in $y = y_{k\pm}$:

$$\sigma(t)y''(t) + \tau(t)y'(t) = \lambda y(t), \tag{A.4}$$

which obviously has the eigenproblem form.

Step 3. Solve *Pearson's equation*

$$(\sigma\rho)' = \tau_{k\pm}\rho, \text{ i.e. } \frac{\rho'}{\rho} = \frac{\tau_{k\pm} - \sigma'}{\sigma},$$

for each k, and among the solutions $\rho_{k\pm}$ select only those being positive (perhaps vanishing on a countable set) and bounded. Then (A.4) can be rewritten in the eigenequation form:

$$Hy(t) = \lambda y(t), \quad \text{with } H = \frac{1}{\rho(t)}\frac{d}{dt}[\sigma(t)\rho(t)\frac{d}{dt}]. \tag{A.5}$$

Table A.1: Classical orthogonal polynomials. Among Jacobi's polynomials are *Legendre's polynomials* $P_n(t) = P_n^{(0,0)}(t)$ and *associated Legendre's polynomials* $P_n^{(m)}(t) = P_n^{(m,m)}(t)$.

Name	(a,b)	$\sigma(t)$	$\tau(t)$	$\rho(t)$	x_n
Jacobi	$(-1,1)$	$1-t^2$	$\beta-\alpha-$ $-(\alpha+\beta+2)t$	$(1-t)^\alpha\times$ $\times(1+t)^\beta$	$P_n^{(\alpha,\beta)}(t) = \frac{(-1)^n \times}{2^n n![(1-t)^\alpha(1+t)^\beta]}$ $\frac{d^n}{dt^n}[(1-t)^{n+\alpha}(1+t)^{n+\beta}]$
Laguerre	$(0,\infty)$	t	$-t+\alpha+1$	$t^\alpha e^{-t}$	$L_n^\alpha(t) = \frac{e^t t^{-\alpha}}{n!}\frac{d^n}{dt^n}[e^{-t}t^{n+\alpha}]$
Hermite	$(-\infty,\infty)$	1	$-2t$	e^{-t^2}	$H_n(t) = \frac{(-1)^n}{\sqrt{2^n n!\sqrt{\pi}}}e^{t^2}\frac{d^n}{dt^n}[e^{-t^2}]$

The rest of the algorithm is based on the

Theorem A.1 *In the subspace* $\{f(t) \in \mathcal{L}^2((a,b),\rho) \mid \sigma(a)\rho(a)f(a) = \sigma(b)\rho(b)f(b) = 0\}$ *of* $\mathcal{L}^2((a,b),\rho)$*, with* ρ *being bounded positive function on* (a,b) *satisfying* $(\sigma\rho)' = \tau\rho$*, operator* H *has purely discrete spectrum* $\{\lambda_n = n\tau' + \frac{n(n-1)}{2}\sigma'' \mid n = 0,1,\dots\}$ *with the corresponding eigenfunctions being orthogonal polynomial given by Rodrigues equation* $x_n(t) = \frac{B_n}{\rho(t)}\frac{d^n}{dt^n}[\sigma^n(t)\rho(t)]$ *(here* B_n *is arbotrary constant, to be determined by the normalization).*

To resume, only for $\lambda = \lambda_n$ (for some n) (A.3) has solution $u = \phi x_n$ satisfying the conditions of the theorem, which is unique (up to the constant). For other values of λ there are no solutions (at least of this type).

It is easy to verify that the obtained functions $x_n(t)$ are polynomials of the order n. Therefore, they can be obtained by the orthonormalization of the set $\{1, t, t^2, \dots\}$ in the space $\mathcal{L}^2((a,b),\rho)$. These *classical orthogonal polynomials* are well studied and standardized (see Table A.1 and [1] for details) according to σ and τ.

Appendix B

Solutions of Exercises

B.0.1 Quantum Kinematics

Solution ??. The following equalities are valid: $[AB, C] = ABC - CAB = ABC - ACB + ACB - CAB = A(BC - CB) + (AC - CA)B = A[B, C] + [A, C]B$.

Solution ??. Writing expression as series in powers of λ we have: $(A - \lambda B)^{-1} = \sum_{n=0}^{\infty} \lambda^n L_n$, where operators L_n are to be determined. Multiplying by $(A - \lambda B)$ one gets $\mathbb{1} = \sum_{n=0}^{\infty} \lambda^n (A - \lambda B) L_n = AL_0 + \sum_{n=0}^{\infty} \lambda^{n+1}(AL_{n+1} - BL_n)$. It is obvious that expression becomes $(A - \lambda B)^{-1} = \sum_{n=0}^{\infty} \lambda^n (A^{-1}B)^n A^{-1} = A^{-1} \sum_{n=0}^{\infty} \lambda^n (BA^{-1})^n$.

Solution ??. $\langle (A - \langle A \rangle)^2 \rangle = \langle A^2 - 2A \langle A \rangle + \langle A \rangle^2 \rangle = \langle A^2 \rangle - 2 \langle A \rangle \langle A \rangle + \langle A \rangle^2 = \langle A^2 \rangle - \langle A \rangle^2$

Solution ??. For $m = 1$ equality holds. Supposing that equality is valid for some $m = n$, for $m = n+1$ using solution ??: $[A^{n+1}, B] = A[A^n, B] + [A, B]A^n = nA^n[A, B] + [A, B]A^n$, using condition that operator A (and any its positive power) commutes with commutator $[A, B]$, we have $[A^{n+1}, B] = (n + 1)A^n[A, B]$.

Solution ??. Assuming that orthonormal basis $\beta = \{|i\rangle\}$ exists, trace can be expressed as $\mathrm{Tr}(|u\rangle \langle v|) = \sum \langle i \,|\, u\rangle \langle v \,|\, i\rangle = \sum \langle v \,|\, i\rangle \langle i \,|\, u\rangle = \langle v| \left(\sum |i\rangle \langle i|\right) |u\rangle = \langle v| \mathbb{1} |u\rangle = \langle v \,|\, u\rangle$.

Solution ??. Operator H is Hermitian iff $(x, Hy) = (Hx, y)$ holds for any two vectors x and y. To shorten notation we'll write $A^\dagger = A^+$ and $A = A^-$. The following set of equalities for any x and y $(x, A^\pm A^\mp y) = (A^\mp x, A^\mp y) = (A^\pm A^\mp x, y)$ proves that operators $A^\pm A^\mp$ are Hermitian. For any vector x, mean value $(x, A^\pm A^\mp x) = (A^\mp x, A^\mp x) = \|A^\mp x\|^2$ is always positive, and for any orthonormal basis $\beta = \{|i\rangle\}$, $\mathrm{Tr} A^\pm A^\mp = \sum (i, A^\pm A^\mp i) = \sum (A^\mp i, A^\mp i) = \sum \|A^\mp i\|^2$. Finally, $\mathrm{Tr} A^\pm A^\mp = \sum \|A^\mp i\|^2 = 0$ iff $\|A^\mp i\| = 0$ for any $i \Rightarrow A^\mp = 0$.

Solution ??. $\mathrm{d}(AB)/\mathrm{d}\lambda = \lim_{\varepsilon\to 0}(A(\lambda + \varepsilon)B(\lambda + \varepsilon) - A(\lambda)B(\lambda))/\varepsilon = \lim_{\varepsilon\to 0}(A(\lambda + \varepsilon)B(\lambda+\varepsilon) - A(\lambda)B(\lambda+\varepsilon) + A(\lambda)B(\lambda+\varepsilon) - A(\lambda)B(\lambda))/\varepsilon = \lim_{\varepsilon\to 0}(A(\lambda+\varepsilon) - A(\lambda))/\varepsilon B(\lambda + \varepsilon) + \lim_{\varepsilon\to 0} A(\lambda)(B(\lambda + \varepsilon) - B(\lambda))/\varepsilon = (\mathrm{d}(A)/\mathrm{d}\lambda)B + A(\mathrm{d}(B)/\mathrm{d}\lambda)$. Differentiating equality $AA^{-1} = \mathbb{1}$ one gets $(\mathrm{d}(A)/\mathrm{d}\lambda)A^{-1} + A(\mathrm{d}(A^{-1})/\mathrm{d}\lambda) = 0 \Rightarrow \mathrm{d}(A^{-1})/\mathrm{d}\lambda = -A^{-1}(\mathrm{d}(A)/\mathrm{d}\lambda)A^{-1}$.

Solution 1.1. For an arbitrary vector $|\psi\rangle$ its image under ρ is $\rho\,|\psi\rangle = \sum_i \omega_i\,|\psi_i\rangle\,\langle\psi_i\,|\,\psi\rangle$. Thus, any vector from the range of ρ is a linear combination of the involved states $|\psi_i\rangle$, and $\mathcal{R}(\rho) < \mathcal{V} \overset{\text{def}}{=} \operatorname{span}\{|\psi\rangle\}$. If $\mathcal{R}(\rho)$ is a proper subspace in $\mathcal{V}$, there exists a nonzero vector $|x\rangle$ from $\mathcal{V}$ such that $\rho\,|x\rangle = 0$. Then we can find an orthonormal basis $|j\rangle$ in $\mathcal{V}$, such that $|1\rangle = |x\rangle$. In this basis $|\psi_i\rangle = \sum_j c_{ji}\,|j\rangle$ and $\rho = \sum_{jj'}\sum_i c_{ji}c^*_{j'i}\omega_i\,|j\rangle\,\langle j'|$, with $c_{ji} = \langle j\,|\,\psi_i\rangle$. Since $|1\rangle$ is an eigenvector of ρ for the eigenvalue 0, it spans one-dimensional invariant subspace for ρ, meaning that for the matrix elements in the first row and the first column of ρ vanish. Particularly, for the first diagonal matrix element one has $\sum_i c_{1i}c^*_{1i}\omega_i = 0$, and in the view of positivity of ω_i this means that all the coefficients c_{1i} vanish, i.e. $|x\rangle = 0$, which contradicts to the assumption.

Solution 1.2. From (1.13b) one finds superoperator $\Pi_A X \overset{\text{def}}{=} \sum_a P_a(A)XP_a(A)$. This superoperator is hermitian: $(X,\Pi_A Y) = \sum_a \operatorname{Tr} X^\dagger P_a(A)YP_a(A) = \sum_a \operatorname{Tr} P_a(A)X^\dagger P_a(A)Y = \operatorname{Tr}(\sum_a P_a(A)XP_a(A))^\dagger Y = (\Pi_A X, Y)$. It is also idempotent:

$$\Pi_A\Pi_A X = \sum_{a,b} P_a(A)P_b(A)XP_b(A)\Pi_A = \sum_a P_a(A)X\Pi_A = \Pi_A X.$$

Solution 1.5. Using partial integration one verifies: $(\psi, p\varphi) = \int_a^b \psi^* p\varphi \mathrm{d}x = \int_a^b \psi^*(-\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}\varphi)\mathrm{d}x = \int_a^b \psi^*(-\mathrm{i}\hbar\varphi')\mathrm{d}x = -\mathrm{i}\hbar\psi^*\varphi|_a^b + \int_a^b(-\mathrm{i}\hbar\psi')^*\varphi\mathrm{d}x = (p\psi,\varphi)$.

Solution 1.6. Applying obvious relation $[A,BC] = B[A,C] + [A,B]C$ successively in $[\hat{x},\hat{p}^n]$ one finds series $[\hat{x},\hat{p}^n] = \hat{p}^{n-1}[\hat{x},\hat{p}] + [\hat{x},\hat{p}^{n-1}]\hat{p} = \mathrm{i}\hbar\hat{p}^{n-1} + \hat{p}^{n-2}[\hat{x},\hat{p}]\hat{p} + [\hat{x},\hat{p}^{n-2}]\hat{p}^2 = 2\mathrm{i}\hbar\hat{p}^{n-1} + [\hat{x},\hat{p}^{n-2}]\hat{p}^2 = \cdots = n\mathrm{i}\hbar\hat{p}^{n-1}$.

Solution 1.7. Since $(U(\boldsymbol{a})\phi, U(\boldsymbol{a})\psi) = \int \mathrm{d}\boldsymbol{r}\phi(\boldsymbol{r}-\boldsymbol{a})\psi(\boldsymbol{r}-\boldsymbol{a}) = \int \mathrm{d}\boldsymbol{r}\phi(\boldsymbol{r})\psi(\boldsymbol{r}) = (\phi,\psi)$, $U(\boldsymbol{a})$ is unitary. The relation with momenta is obtained as in one-dimensional case, due to commutativity: $U(\boldsymbol{a}) = \prod_{i=1}^3 \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}a_i p_i} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\boldsymbol{a}\cdot\boldsymbol{p}}$ (more in Chapter 3).

Solution 1.8. Since for each $|x\rangle$ one has $\langle x|\,A_1\,|x\rangle = \sum_b(\langle x|\otimes\langle b|)A(|x\rangle\otimes|b\rangle)$, the reality and positivity of the terms on the right implies reality and positivity of the left hand side. Finally, if A is statistical, $\operatorname{Tr} A_1 = \sum_{ab}(\langle a|\otimes\langle b|)A(|a\rangle\otimes|b\rangle) = \operatorname{Tr} A = 1$.

Solution 1.9. Consider the general case, i.e. mixed state ρ. Being positive operator, ρ has unique positive root $\rho^{1/2}$. Let us define the auxiliary operators $A' = A - \langle A\rangle_\rho$, $B' = B - \langle B\rangle_\rho$, and $A” = \rho^{1/2}A'$, $B” = \rho^{1/2}B'$. Using the standard scalar product in the operators space $(X,Y) = \operatorname{Tr} A^\dagger B$, it is easy to recognize the Schwarz inequality in the expression $\operatorname{Tr}(A”^\dagger A”)\operatorname{Tr}(B”^\dagger B”) \geq |\operatorname{Tr} B”^\dagger A”|^2$. Note that the left hand side is equal to $\Delta^2_\rho(A)\Delta^2_\rho(B)$, while the right one substituting $B”$ becomes mean $|\,\langle A'B'\rangle_\rho\,|^2 = |\,\langle\frac{A'B'+B'A'}{2}\rangle_\rho + \langle\frac{A'B'-B'A'}{2}\rangle_\rho\,|^2$. Recall that the obtained symmetrized and antisymmetrized products are hermitian and scew-hermitian operators, and their means are real and pure imaginary, respectively. The absolute value of the whole expression is greater than the square of the imaginary part: $\Delta^2_\rho(A)\Delta^2_\rho(B) \geq |\,\langle A'B'\rangle_\rho\,|^2 = \frac{1}{4}|\,\langle[A',B']\rangle_\rho\,|^2 = \frac{1}{4}|\,\langle[A,B]\rangle_\rho\,|^2$.

Solution 1.10. ???

B.0.2 Quantum dynamics

Solution 2.1. Differentiating the expectation value one finds:

$$\frac{\mathrm{d}}{\mathrm{d}a}E(a) = \langle E(a)| \frac{\mathrm{d}}{\mathrm{d}a}H(a)\,|E(a)\rangle + (\frac{\mathrm{d}}{\mathrm{d}a}\langle E(a)|)H(a)\,|E(a)\rangle + \langle E(a)|\,H(a)(\frac{\mathrm{d}}{\mathrm{d}a}\,|E(a)\rangle).$$

Therefore:

$$\frac{\mathrm{d}}{\mathrm{d}a}E(a) = \langle E(a)| \frac{\mathrm{d}}{\mathrm{d}a}H(a)\,|E(a)\rangle + E(a)(\frac{\mathrm{d}}{\mathrm{d}a}\langle E(a)|)\,|E(a)\rangle + E(a)\,\langle E(a)|\,(\frac{\mathrm{d}}{\mathrm{d}a}\,|E(a)\rangle).$$

The last two terms cancel, as the derivative of the norm of the eigenstate, being 1.

Solution 2.2. $\langle n,\lambda|\,[H,A]\,|n',\lambda'\rangle = \langle n,\lambda|\,(HA-AH)\,|n',\lambda'\rangle = (E_n - E_{n'})\,\langle n,\lambda|\,A\,|n',\lambda'\rangle$.

Solution **??**. Straightforwardly apply (2.20).

Solution **??**. It is easy to see that if energy (E) is greater than any potential in infinity ($V_\pm$) than eigenfunction is plane-wave with momentum $k = \sqrt{2m(E-V_\pm)/\hbar^2}$ without boundary condition in that infinity. This means that particle is free iff energy is greater than minimum of the potentials in infinity. In conclusion, the state is bound iff $V_+ > E$.

Solution **??**. Suppose that we have two physically different states (nonlinear) $\psi_1(x)$ and $\psi_2(x)$ heaving the same energy. Than, for any potential $V(x)$ we have $\psi_1''/\psi_1 = 2m(E - V(x))/\hbar^2 = \psi_2''/\psi_2$ leading to equation $(\psi_1'\psi_2 - \psi_1\psi_2')' = 0$. Integrating the last equation we have $\psi_1'\psi_2 - \psi_1\psi_2' = C$ for any x. Constant C vanishes because eigenfunction and its first derivative for boundary states vanishes in both infinities. Furthermore, integrating $\psi_1'/\psi_1 = \psi_2/\psi_2'$ we have $\psi_1 = A\psi_2$ and because states are normalized constant A is phase factor specifying that states $\psi_1(x)$ and $\psi_2(x)$ are the same physical states which contradicts assumptions.

Solution **??**. Consider two eigenfunctions $\psi_n(x)$ and $\psi_{n+1}(x)$ with eigenvalues $E_n < E_{n+1}$. From the equations $\psi_n'' + 2m(E_n - V)\psi_n/\hbar^2 = 0$ and $\psi_{n+1}'' + 2m(E_{n+1} - V)\psi_{n+1}/\hbar^2 = 0$, after simple calculation we obtain $(\psi_n'\psi_{n+1} - \psi_{n+1}'\psi_n)|_a$...???

Solution **??**. Adding and subtracting normal and spatially inverted Schrödinger equation we get two equalities $(V(x) - E)(\psi(x) \pm \psi(-x)) = 0$, $\forall x$, giving that eigenfunctions are either even or odd ($\psi(-x) = \pm\psi(x)$).

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution **??**. ???

Solution 2.4. Obvious.

Solution 2.5. Obvious.

Solution 2.6. If two bound states with the wave functions $\psi_i(x)$ both satisfy $H\psi_i(x) = E\psi_i(x)$ with $H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)$, then, first multiply each equation by another wave function and then subtract the obtained two equations, to get $\psi_2\psi_1'' - \psi_1\psi_2'' = 0$. Consequently $\psi_2\psi_1' - \psi_1\psi_2' = c$, where c is a constant; however, at infinity both functions must vanish, implying $c = 0$. Finally, $\frac{\psi_1'}{\psi_1} = \frac{\psi_2'}{\psi_2}$, which shows that $\psi_1 = d\psi_2$.

Solution 2.7. Proceeding as in Exercise 2.6, assume that the wave functions $\psi_i(x)$ ($i = 1, 2$) satisfy $H\psi_i(x) = E_i\psi_i(x)$ (assumed $E_2 > E_1$); again, first multiply each equation by another wave function and then subtract the obtained two equations, to get $(\psi_1\psi_2' - \psi_2\psi_1')' = \frac{\hbar^2}{2m}(E_1 - E_2)\psi_1\psi_2$. If $\psi_1(x_0) = \psi_2(x_0) = 0$ (at least at $-\infty$), for each x_1 holds

$$\psi_1(x_1)\psi_2'(x_1) - \psi_2(x_1)\psi_1'(x_1) = -\frac{\hbar^2}{2m}(E_2 - E_1)\int_{x_0}^{x_1} \mathrm{d}x\,\psi_1(x)\psi_2(x),$$

which is $\psi_2(x_1)\psi_1'(x_1) = \frac{\hbar^2}{2m}(E_1 - E_2)\int_{x_0}^{x_1} \mathrm{d}x\,\psi_1(x)\psi_2(x)$ for the next greater than x_0 root x_1 of ψ_1. Clearly, sign$\psi_1(x)$ is constant within (x_0, x_1), and it is opposite to sign$\psi_1'(x_1)$. Then, if also sign$\psi_2(x)$ is constant within (x_0, x_1), the equality does not hold (opposite signs of the two sides). https://arxiv.org/pdf/quant-ph/0702260.pdf

Solution 2.8. Conjugate of the of the wave equation $(-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x))\psi(x) = E\psi(x)$ is $(-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x))\psi^*(x) = E\psi^*(x)$ (real potential is naturally assumed[1]). The sum of the two identities is the same equation for real $\psi + \psi^*$. For the parity proof, note that x-reversal makes the same equation for $\psi(x)$ and $\psi(-x)$, i.e. their sum and difference are with the same eigenvalue, and only one can be nonzero (otherwise the degeneracy occurs).

Solution 2.9. Note that the eigenproblem has form of the hypergeometric equation, and use this (Appendix **??**).

Solution **??**. ???

Solution 2.10. This is a consequence of the commutation relation in (2.26), and well known general rules $[A, BC] = [A, B]C + B[A, C]$ and $[AB, C] = A[B, C] + [A, C]B$. Technically, consider a monomial M with a non-ordered sequence between sub-monomials (ordered or not) M_1 and M_2, i.e. $M = M_1 a^- a^+ M_2$ (with $|M|$ operators). Then this

---

[1] This is related to time-reversal invariance of the system.

disorder is rectified at the cost of additional monomial with $|M|-2$ operators: $M = M_1a^-a^+M_2 = M_1a^+a^-M_2 + M_1M_2$.

Solution 2.11. With well known identity $[AB,C] = A[B,C] + [A,C]B$ and (2.26) written in the form $[a^\pm, a^\mp] = \mp 1$, one easily gets $A = [(a^\pm)^s, a^\mp] = [(a^\pm)^{s-1}a^\pm, a^\mp] = (a^\pm)^{s-1}[a^\pm, a^\mp] + [(a^\pm)^{s-1}, a^\mp]a^\pm = \mp(a^\pm)^{s-1} + [(a^\pm)^{s-1}, a^\mp]a^\pm$. Repeating the same procedure in the second term: $A = \mp 2(a^\pm)^{s-1} + [(a^\pm)^{s-2}, a^\mp](a^\pm)^2 ...$; in the $(s-1)$-step the result is: $A = \mp(s-1)(a^\pm)^{s-1} + [a^\pm, a^\mp](a^\pm)^{(s-1)}$, which immediately gives the wanted $A = \mp s(a^\pm)^{s-1}$.

Solution 2.12. Identity $[AB,CD] = A[B,C]D + [A,C]BD + CA[B,D] + C[A,D]B$ is easily verified by direct check or using obvious identities $[A,BC] = [A,B]C + B[A,C]$ and $[AB,C] = A[B,C] + [A,C]B$. Accordingly, $[(a^+)^p(a^-)^m, a^+a^-] =$
$= (a^+)^p[(a^-)^m, a^+]a^- + [(a^+)^p, a^+](a^-)^m a^- + a^+(a^+)^p[(a^-)^m, a^-] + a^+[(a^+)^p, a^-](a^-)^m$. While the second and the third term vanish (operator commutes with its arbitrary function), to the first and to the last applies the first identity of Exercise 2.11, yielding $m(a^+)^p m(a^-)^{m-1}a^- - pa^+(a^+)^{p-1}(a^-)^m = (m-p)(a^+)^p(a^-)^m$, which is looked for.

Solution 2.13. ???

Solution 2.14. ???

Solution 2.15. ???

B.0.3 Galilean transformations

Solution 3.1. ???

Solution 3.2. Introducing family $g(A,B,\alpha) = e^{\alpha B}Ae^{-\alpha B}$ and expanding it around $\alpha = 0$ one gets $g(A,B,\alpha) = \sum \frac{\alpha^n}{n!}\frac{\mathrm{d}^n g}{\mathrm{d}\alpha^n}$, where $\frac{\mathrm{d}^0 g}{\mathrm{d}\alpha^0} = A$, $\frac{\mathrm{d}g}{\mathrm{d}\alpha} = Bg - gB = [B,A]$, $\frac{\mathrm{d}^2 g}{\mathrm{d}\alpha^2} = [B, \frac{\mathrm{d}g}{\mathrm{d}\alpha}] = [B,[B,A]], \ldots$ Setting $\alpha = 1$ one gets the equality.
The second identity follows from the previous one (and the pointed out conditions), with $T(s) = e^{sA}e^{sB}$: $\frac{\mathrm{d}}{\mathrm{d}s}T(s) = AT(s) + T(s)B = (A + e^{sA}Be^{-sA})T(s) = (A + B + [A,B]s)T(s) \Rightarrow T(s) = \mathrm{e}^{(A+B)s + \frac{1}{2}[A,B]s^2}$.

Solution 3.3. For even n, $\mathcal{I}^n$ is unity while for odd n, $\mathcal{I}^n = \mathcal{I}$. Expanding operator $\mathrm{e}^{\mathrm{i}\pi\mathcal{I}}$ one gets $\mathrm{e}^{\mathrm{i}\pi\mathcal{I}}\psi(\mathbf{r}) = \sum_n \frac{(\mathrm{i}\pi)^n}{n!}\mathcal{I}^n\psi(\mathbf{r}) = \sum_k (-)^k \frac{\pi^{2k}}{(2k)!}\psi(\mathbf{r}) + \mathrm{i}\sum_k (-)^k \frac{\pi^{2k+1}}{(2k+1)!}\mathcal{I}\psi(\mathbf{r}) = \cos(\pi)\psi(\mathbf{r}) + \mathrm{i}\sin(\pi)\psi(-\mathbf{r}) = -\psi(\mathbf{r})$

Solution 3.4. Introducing unity $\mathbb{1}$ and space inversion $\mathcal{I}$ operators, one can write any function as sum of even and odd functions: $\psi(\mathbf{r}) = \frac{\psi(\mathbf{r})+\psi(-\mathbf{r})}{2} + \frac{\psi(\mathbf{r})-\psi(-\mathbf{r})}{2} = \frac{\mathbb{1}+\mathcal{I}}{2}\psi(\mathbf{r}) + \frac{\mathbb{1}-\mathcal{I}}{2}\psi(\mathbf{r}) = P_+\psi(\mathbf{r}) + P_-\psi(\mathbf{r})$. Operator $\mathcal{I}$ is Hermitian as it follows from: $(\psi(\mathbf{r}), \mathcal{I}\varphi(\mathbf{r})) = \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}y\int_{-\infty}^{\infty}\mathrm{d}z\psi(\mathbf{r})\varphi(-\mathbf{r}) = -\int_{\infty}^{-\infty}\mathrm{d}x\int_{\infty}^{-\infty}\mathrm{d}y\int_{\infty}^{-\infty}\mathrm{d}z\psi(-\mathbf{r})\varphi(\mathbf{r}) = \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}y\int_{-\infty}^{\infty}\mathrm{d}z\psi(-\mathbf{r})\varphi(\mathbf{r}) = (\mathcal{I}\psi(\mathbf{r}), \varphi(\mathbf{r}))$. Consequently, operators $P_\pm$ are Hermitian, and idempotent $P_\pm^2 = \frac{(\mathbb{1}\pm\mathcal{I})^2}{4} = \frac{\mathbb{1}\pm 2\mathcal{I}+\mathcal{I}^2}{4} = P_\pm$.

Solution 3.5. Operator $U_{\alpha\beta}$ is unitary if α and β are pure imaginary numbers. Operator $U_{\alpha\beta}$ is space inversion if $S_x = U_{\alpha\beta}xU_{\alpha\beta}^\dagger = -x$ and $S_p = U_{\alpha\beta}pU_{\alpha\beta}^\dagger = -p$. The operator equality Prob. **??** gives: $S_x = x + [\alpha x^2 + \beta p^2, x] + \frac{1}{2!}[\alpha x^2 + \beta p^2, [\alpha x^2 + \beta p^2, x]] + \cdots = \sum_{i=0}^{\infty} \frac{1}{2!} K_i$, where $K_{i+1} = [K_i, x]$. It is easy to calculate commutators: $K_0 = x$, $K_1 = -2\mathrm{i}\hbar\alpha p$, $K_2 = -2\mathrm{i}\hbar\alpha 2$

Solution **??**. ???

B.0.4 Angular momentum

Solution 4.2. Obviously, if $K_p \left|x\right\rangle = 0$ $(p = 1, 2, 3)$, then also $\langle K_p^2 \rangle_{|x\rangle} = \left\langle x\right| K_p K_p \left|x\right\rangle = \| K_p \left|x\right\rangle \| = 0$, and $\langle K^2 \rangle_{|x\rangle} = 0$ (and $k = m = 0$).
In the opposite direction, from $\langle K_i^2 \rangle_{|x\rangle} = \langle K_j^2 \rangle_{|x\rangle} = 0$, the same argument implies that $\| K_i \left|x\right\rangle \| = \| K_j \left|x\right\rangle \| = 0$, i.e. $K_i \left|x\right\rangle = K_j \left|x\right\rangle = 0$. However, then $0 = [K_i, K_j] \left|x\right\rangle = K_k \left|x\right\rangle$, for the remaining component K_k.

Solution 4.1. Straightforward application of (4.8) in the representation formula $K_i \left|km\right\rangle = \sum_{m'=-k}^{k} D_{m'm}^{(k)}(K_i) \left|km'\right\rangle$.

Solution 4.2. Use the form (4.12) for K^2, and then substitute the expressions (4.8) for the action of K_3 and $K_\pm$ on the standard basis.

Solution 4.3. As squares of Hermitean operators, K_i^2 are positive, implying that their average vanishes if and only if $K_i \left|x\right\rangle = 0$. Then, if $\langle K_i^2 \rangle_{|x\rangle} = \langle K_j^2 \rangle_{|x\rangle} = 0$, also $[K_i, K_j] \left|x\right\rangle = 0$, and $\langle K^2 \rangle_{|x\rangle} = 0$. This means that $k = m = 0$.

Solution 4.4. ???

Solution 4.5. Since $\mathcal{I}_S(r, \theta, \varphi) = (r, \pi - \theta, \varphi + \pi)$ is the action of spatial inversion in the spherical coordinates, it follows $\mathcal{I}_S Y_l^m(\theta, \varphi) = Y_l^m(\pi - \theta, \varphi + \pi)$. Substituting this in (4.4) the results is straightforwardly obtained.

Solution 4.6. Besides nucleus helium atom contains two electrons, and without spin-orbit interaction the electronic Hamiltonian consists of noninteracting single particle part H_0 (nonperturbed part) and electron-electron Coulomb repulsion (perturbation):

$$H = H_0 + H', \quad H_0 = h_1 + h_2, \quad h_i = T_i - \frac{2e}{r_i}, \quad H' = \frac{e^2}{r_{12}}. \tag{B.1}$$

Eigenenergies of the hydrogen-like Hamiltonians h_i are $\epsilon_n = -\frac{2me^2}{n^2\hbar^2}$, with orbital eigenfunctions $\langle r, \theta, \varphi \,|\, nlm \rangle = R_{nl}(r) Y_l^m(\theta, \varphi)$. Without spin orbit coupling (neglected), the states $\left|nlm\pm\right\rangle$ are degenerate, and the ground single electron state is $\left|100\pm\right\rangle$. Accordingly, the minimal energy $\epsilon_1 = -54.4\,\mathrm{eV}$ is achieved when both electrons are with $n = 1$ (and $l = m = 0$), which due to total antisymmetrization (necessary for fermions as shown in Section 5.1) gives the total state

$$\left|\Psi\right\rangle = \frac{\left|100+\right\rangle\left|100-\right\rangle - \left|100-\right\rangle\left|100+\right\rangle}{\sqrt{2}} = \left|100\right\rangle \frac{\left|+-\right\rangle - \left|-+\right\rangle}{\sqrt{2}}, \quad \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{8}{\pi a_0^3} \mathrm{e}^{-2\frac{r_1 + r_2}{a_0}} \frac{\left|+-\right\rangle - \left|-+\right\rangle}{\sqrt{2}},$$

with $L = 0$ and $S = 0$; clearly, $J = 0$, too. Obviously, the nonperturbed energy of this (ground) state is $E_0^0 = 2\epsilon_1 = -108.8\,\mathrm{eV}$. Using $\int \frac{\mathrm{e}^{-(ar_1 + br_2)}}{|r_1 - r_2|} \,\mathrm{d}\boldsymbol{r}_1 \,\mathrm{d}\boldsymbol{r}_2 = 16\pi^2 \frac{2a^2 + 2b^2 + 6ab}{a^2 b^2 (a+b)^2}$, the first correction (Section 6.1) is found $E_0 = E_0^0 + E_0^{(1)} = -74.8\,\mathrm{eV}$. The experimental value is $E_0^{\mathrm{exp}} = -78.6\,\mathrm{eV}$, thus lower (negative second order contribution).

B.0.5 Identical particles

Solution 5.1. Here $\tau_{pq} = \tau_{pq}^{-1} = \tau_{qp}$ is transposition of the elements p and q; conveniently, $\tau_{ii} = e$. All the transpositions are in the same conjugation class of $\boldsymbol{S}_N$, since any two of them, $\tau_{pp'}$ and $\tau_{qq'}$, are mutually conjugated: $\tau_{pp'} = \tau_{pq}\tau_{p'q'}\tau_{qq'}\tau_{pq}\tau_{p'q'}$. Therefore, they have the same character in each representation; the same numbers correspond to all transpositions in one dimensional representations. As obviously $\tau_{pp'}^2 = e$ (identity element), homomorphism implies $D^2(\tau_{pp'}) = \mathbb{1}$ in any representation. Consequently, in one dimensional representations for all transpositions is simultaneously either $D(\tau_{pp'}) = 1$ or $D(\tau_{pp'}) = -1$. Taking one of these two choices, the number representing any other permutation is product (of even or odd number) of ± factors.

Solution 5.2. To demonstrate that transpositions τ_i generate $\boldsymbol{S}_N$, arbitrary permutation $\pi = \pi^{(0)} = (\pi_1, \dots, \pi_N)$ will be factorized as their product. To this end we start with auxiliary permutation $p = p^{(0)}$ set to the identical element of $\boldsymbol{S}_N$. Then, the minimal position i_1 such that π_{i_1} is greater than π_{i_1+1} is found, and the permutations $p^{(1)} = \tau_{i_1} p^{(0)}$ and $\pi^{(1)} = \pi^{(0)}\tau_{i_1}$ are formed. The procedure is repeated with $\pi^{(1)}$ and $p^{(1)}$ (instead of $\pi^{(0)}$ and $p^{(0)}$), to get $\pi^{(2)}$ and $\pi^{(2)}$, etc. Obviously, for some $k < N$ steps, this results in the identical permutation $\pi^{(k)} = e$, while $\pi^{(k)} = \tau_{i_k} \cdots \tau_{i_1} = \pi$, which is the proposed factorization. The proof of relations (5.5) is straightforward.

Solution 5.3. Occupation numbers $\boldsymbol{n} = (n_1, \dots, n_7) = (0, 2, 1, 0, 3, 0, 0)$ define orbit with the standard representative $|\boldsymbol{n}\rangle = |2, 2, 3, 5, 5, 5\rangle$. The stabilizer of $|\boldsymbol{n}\rangle$ is $\boldsymbol{S}_0 \otimes \boldsymbol{S}_2 \otimes \boldsymbol{S}_1 \otimes \boldsymbol{S}_0\boldsymbol{S}_3 \otimes \boldsymbol{S}_0\boldsymbol{S}_0 = \boldsymbol{S}_2 \otimes \boldsymbol{S}_3$, with $|\boldsymbol{n}| = 12$ elements. Hence, the orbit consists of $N!/12 = 60$ states.

Solution 5.4. $\Delta_{\boldsymbol{n}}(S_N)$ is defined by the action of the permutations on the orbit representative $|\boldsymbol{n}\rangle$: $\Delta_{\boldsymbol{n}}(\pi)\,|\boldsymbol{n}\rangle = |\boldsymbol{n}, p\rangle$ for $\pi = \zeta_p \phi$, which is equivalent to $\Delta_{\boldsymbol{n}}(\boldsymbol{S}_N) = \mathbb{1}(\boldsymbol{S}_{\boldsymbol{n}}) \uparrow \boldsymbol{S}_N$. Thus, the frequencies are $f^{\pm}_{\Delta_{\boldsymbol{n}}} = \mathrm{Tr}\,\boldsymbol{S}_N(\Delta_{\boldsymbol{n}} \otimes D^{(\pm)}) = \mathrm{Tr}\,\boldsymbol{S}_{\boldsymbol{n}}(\mathbb{1} \otimes D^{(\pm)}) = \frac{1}{\boldsymbol{n}!}\sum_\phi (\pm)^{\tilde{\phi}}$.

Solution 5.5. The two bases are necessarily related as $|j_p\rangle = \sum_l U_{lp}\,|i_l\rangle$, with U being an unitary operator (thus $\det(U)$ is pure phase). The new Slater determinant is $|\boldsymbol{n}^-; j\rangle = \frac{1}{\sqrt{N!}} \det |j_p\rangle_q = \frac{1}{\sqrt{N!}} \det \sum_l U_{lp}\,|i_l\rangle_q$. Noticing that the form of the pq-th element is $(U^T\,|i_l\rangle)_{pq}$, the homomorphism of the determinant functional, $\det A \det B = \det AB$, directly gives $|\boldsymbol{n}^-; j\rangle = \det(U)\,|\boldsymbol{n}^-; i\rangle$.

Solution 5.6. Positive $\boldsymbol{S}_3$ permutations are $\{(123), (231), (312)\}$, while $\{(132), (213), (321)\}$ are negative. Hence, up to the normalizing constant, $|(210)\rangle = |112\rangle + |121\rangle + |211\rangle + |121\rangle + |112\rangle + |211\rangle$ (in the order of positive followed by negative permutations of $|112\rangle$; only bosonic case is considered). This gives:

$$\begin{array}{rccccccl}
\pi : & (123) & (231) & (312) & (132) & (213) & (321) & \\
|(210)\rangle \sim & |112\rangle + & |121\rangle + & |211\rangle + & |121\rangle + & |112\rangle + & |211\rangle & \\
\langle 1_1 \,|\, (210)\rangle \sim & |1_2 2_3\rangle + & |2_2 1_3\rangle + & & |2_2 1_3\rangle + & |1_2 2_3\rangle & & \sim |(110)\rangle_{23} \\
\langle 1_2 \,|\, (210)\rangle \sim & |1_1 2_3\rangle + & & |2_1 1_3\rangle + & & |1_1 2_3\rangle + & |2_1 1_3\rangle & \sim |(110)\rangle_{13} \\
\langle 1_1 1_2 \,|\, (210)\rangle \sim & |2_3\rangle + & & & & |2_3\rangle & & \sim |(010)\rangle_3 \\
\langle 1_1 2_2 \,|\, (210)\rangle \sim & & |1_3\rangle + & & |1_3\rangle & & & \sim |(100)\rangle_3 \\
\langle 2_1 1_2 \,|\, (210)\rangle \sim & & & |1_3\rangle + & & & |1_3\rangle & \sim |(100)\rangle_3
\end{array}$$

???INSERT CONSTANTS??? Obviously, for the corresponding clusters the same states are obtained. In the fermionic case the considered state vanishes, as well as the subsystems' states. Starting with nontrivial fermionic state, the same equivalence is easily obtained.

Solution 5.7. All the terms in the sum (5.16) contain the same quantum numbers, and obviously the wanted partial scalar product vanishes if the set of quantum numbers $\{i_1, \dots, i_N\}$ is not a superset of $\{j_{L+1}, \dots, j_N\}$, i.e. when some of the components of $\boldsymbol{n}_L = \boldsymbol{n} - \boldsymbol{n}_{N-L}$ are negative. Otherwise, there exists a permutation σ defined in the exercise, rearranging the quantum numbers $\{i_1 \leq \cdots \leq i_N\}$ such that the last $N-L$ coincide with $|j_{L+1}, \dots, j_N\rangle$ and the first L are ordered. Then the L particle state $|\Psi_L\rangle = \langle j_{L+1}, \dots, j_N \,|\, \boldsymbol{n}^{\pm}\rangle$ is:

$$|\Psi_L\rangle = \frac{(\pm)^{\tilde{\sigma}}}{\sqrt{\boldsymbol{n}! N!}} \sum_{\pi} (\pm)^{\tilde{\pi}} \langle j_{L+1}, \dots, j_N|\, \Delta(\pi)\, |i_{\sigma^{-1}1} \leq \cdots \leq i_{\sigma^{-1}L}, j_{L+1}, \dots, j_N\rangle .$$

Since bra-vector is fixed, the partial scalar product involving the permutation π does not vanish only when π leaves invariant the states of the particles $L+1, \dots, N$. Each such permutation may be written as a product $\pi = \pi_L \pi_{\boldsymbol{n}}$, where $\pi_L \in \boldsymbol{S}_L$ permutes only the first L particles, and $\pi_{\boldsymbol{n}} \in \boldsymbol{S}_{\boldsymbol{n}}$. However, the intersection of the groups $\boldsymbol{S}_L$ and $\boldsymbol{S}_{\boldsymbol{n}}$ is just the group $\boldsymbol{S}_{\boldsymbol{n}_L}$, the stabilizer of the L particle state $|i_{\sigma^{-1}1} \leq \cdots \leq i_{\sigma^{-1}L}\rangle$. Therefore, the factorization $\pi_L \pi_{\boldsymbol{n}}$ is not unique, and in the set of the products $\pi_L \pi_{\boldsymbol{n}}$ each permutation appears exactly $\boldsymbol{n}_L!$ times. One gets

$$|\Psi_L\rangle = \sum_{\pi_L, \pi_{\boldsymbol{n}}} \frac{(\pm)^{\tilde{\sigma} + \tilde{\pi}_L + \tilde{\pi}_{\boldsymbol{n}}}}{\boldsymbol{n}_L! \sqrt{\boldsymbol{n}! N!}} \langle j_{L+1}, \dots, j_N|\, \Delta(\pi_L \pi_{\boldsymbol{n}})\, |i_{\sigma^{-1}1} \leq \cdots \leq i_{\sigma^{-1}L}, j_{L+1}, \dots, j_N\rangle .$$

Finally, one realizes that

$$\langle j_{L+1}, \dots, j_N|\, \Delta(\pi_L \pi_{\boldsymbol{n}})\, |i_{\sigma^{-1}1} \leq \cdots \leq i_{\sigma^{-1}L}, j_{L+1}, \dots, j_N\rangle = \Delta(\pi_L)\, |i_{\sigma^{-1}1} \leq \cdots \leq i_{\sigma^{-1}L}\rangle ,$$

independently of $\pi_{\boldsymbol{n}}$; hence there are $\boldsymbol{n}!$ such terms (all with the same sign: for fermions $\pi_{\boldsymbol{n}}$ is the identical permutation only, and for bosons all terms are with the same sign), giving

$$|\Psi_L\rangle = \frac{(\pm)^{\tilde{\sigma}} \boldsymbol{n}!}{\boldsymbol{n}_L! \sqrt{\boldsymbol{n}! N!}} \sum_{\pi_L} (\pm)^{\tilde{\pi}_L} \Delta(\pi_L)\, |i_{\sigma^{-1}1} \leq \cdots \leq i_{\sigma^{-1}L}\rangle .$$

Summing over π_L, and including normalization factor, one immediately gets (5.21a). Analogously one gets (5.21b). As for $L = 1$ case, the permutation σ moves the first particle in the state $|j\rangle$, to the first place (where this state is cancelled by $\langle j|$). So, there are

$l_j = \sum_{i<j} n_i$ states before, which are effectively cycled by $\sigma = \tau_{12} \cdots \tau_{l_j-1,l_j} \tau_{l_j,l_j+1}$, showing that $l_j = \tilde{\sigma}$.

Solution 5.8. First, note that the permutation σ from the Exercise 5.9 may be given a canonical form $\sigma = \pi_{N-L}\lambda$, where λ is a "riffle shuffle" permutation, which on the set $\{i_1, ..., i_n\}$ with singled out L (and therefore the remaining $N-L$) elements simply separates them, putting the L elements first and $N-L$ after them, without changing the order within subsets, while π_{N-L} permutes only the second subset. Taking into account (5.16) for $N-L$ particles, one gets $\langle \boldsymbol{n}^{\pm}_{N-L} \,|\, \boldsymbol{n}^{\pm} \rangle = \sqrt{\frac{1}{\boldsymbol{n}_{N-L}!(N-L)!}} \sqrt{\frac{\boldsymbol{n}!L!}{\boldsymbol{n}_L!N!}} (N-L)!(\pm)^{\tilde{\lambda}} \, |\boldsymbol{n}^{\pm}_L\rangle$.

Solution 5.9. From (5.21a) it follows straightforwardly:

$$\rho_{1,\dots,L} = \sum_{j_{L+1},\dots,j_N} \langle j_{L+1},\dots,j_N \,|\, \boldsymbol{n}^{\pm}\rangle \langle \boldsymbol{n}^{\pm} \,|\, j_{L+1},\dots,j_N\rangle = \sum_{\boldsymbol{n}_L \leq \boldsymbol{n}} C_{\boldsymbol{n}_L} \frac{\boldsymbol{n}!L!}{\boldsymbol{n}_L!N!} \, |\boldsymbol{n}^{\pm}_L\rangle \langle \boldsymbol{n}^{\pm}_L| \, ,$$

where $C_{\boldsymbol{n}_L}$ shows how many times occupation number $\boldsymbol{n}_L$ appears in the summation over the permutations of the elements of the set $\boldsymbol{j} = (j_{L+1},\dots,j_N)$. To find the weights $C_{\boldsymbol{n}_L}$, we observe that $\boldsymbol{n}_L = \boldsymbol{n} - \boldsymbol{n}(\boldsymbol{j})$ and that when $\boldsymbol{n}(\boldsymbol{j})$ fixed, $\boldsymbol{n}_L$ is also fixed. Therefore, out of $(N-L)!$ permutations of the elements of the set $\boldsymbol{j}$, altogether $\boldsymbol{n}(\boldsymbol{j})!$ permutations yield the same term in the sum over $\boldsymbol{j} = (j_{L+1},\dots,j_N)$. Thus, there are altogether $C_{\boldsymbol{n}_L} = (N-L)!/\boldsymbol{n}(\boldsymbol{j})! = (N-L)!/(\boldsymbol{n}-\boldsymbol{n}_L)!$ terms with same $\boldsymbol{n}_L$. Hence

$$\rho_{1,\dots,L} = \sum_{\boldsymbol{n}_L \leq \boldsymbol{n}} \frac{(N-L)!}{(\boldsymbol{n}-\boldsymbol{n}_L)!} \frac{\boldsymbol{n}!L!}{\boldsymbol{n}_L!N!} \, |\boldsymbol{n}^{\pm}_L\rangle \langle \boldsymbol{n}^{\pm}_L| = \binom{N}{L}^{-1} \sum_{\boldsymbol{n}_L \leq \boldsymbol{n}} \binom{\boldsymbol{n}}{\boldsymbol{n}_L} |\boldsymbol{n}^{\pm}_L\rangle \langle \boldsymbol{n}^{\pm}_L| \, .$$

Solution 5.10. As $(220), (211), (130), (121), (031) \leq (231)$, binomials give:

$$\begin{aligned} \rho_4 &= 3\,|(220)^+\rangle \langle (220)^+| + 6\,|(211)^+\rangle \langle (211)^+| + 2\,|(130)^+\rangle \langle (130)^+| + \\ &+ 6\,|(121)^+\rangle \langle (121)^+| + 3\,|(031)^+\rangle \langle (031)^+| \, . \end{aligned}$$

Solution 5.11. As now $\boldsymbol{n}! = \boldsymbol{n}_L! = 1$ (5.24a) becomes

$$\rho_{1,\dots,L} = \sum_{\boldsymbol{n}_L \leq \boldsymbol{n}} \frac{|\boldsymbol{n}^{\pm}_L\rangle \langle \boldsymbol{n}^{\pm}_L|}{\binom{N}{L}} = \frac{L!}{\binom{N}{L}} \sideset{}{'}\sum_{i_1<\cdots<i_L=1}^{N} P^{(-)}_L \, |i_1,\dots,i_L\rangle \langle i_1,\dots,i_L| \, P^{(-)}_L .$$

The primed sum means that all the indices run over the N single particle states defined by $n_s = 1$ in $\boldsymbol{n}$ (alternatively, we may assume that the single particle space basis $|i\rangle$ starts with these N occupied states and indices run from 1 to N). Due to the antisymmetrizer, the restrictions $i_1 < \cdots < i_L$ in the summation may be omitted: the repeated indices are neglected by the antisymmetrizers, and only unordered permutations of $(i_1 < \cdots < i_L)$ are effectively added. However, corresponding vectors $P^{(-)} \, |i_1,\dots,i_L\rangle$ are up to the sign same as for the ordered one, and the same sign appears both in bra and ket. Hence there are $L!$ same terms, instead of one which must be compensated by divison:

$$\rho_{1,\dots,L} = \frac{1}{\binom{N}{L}} \sideset{}{'}\sum_{i_1,\dots,i_L=1}^{N} P^{(-)}_L \, |i_1,\dots,i_L\rangle \langle i_1,\dots,i_L| \, P^{(-)}_L .$$

It remains to notice that the summations in $i_1, \dots, i_L$ become independent, each giving the projector $P_{\boldsymbol{n}}$, and that the direct product of L these projectors commutes with the projector $P_L^{(-)}$, which can be omitted from one side.

Solution 5.12. By definition $(\hat{\hat{P}}^{(+)}A)P^{(\pm)} = \frac{1}{N!^2}\sum_{\pi\pi'}(\pm)^{\tilde{\pi}}\Delta^{-1}(\pi')A\Delta(\pi')\Delta(\pi)$. Denoting $\pi'\pi$ by π'', one finds $\Delta(\pi')\Delta(\pi) = \Delta(\pi'')$ and $\Delta^{-1}(\pi') = \Delta^{-1}(\pi)\Delta^{-1}(\pi'')$. Substituting summation in π' and π by summation in π'' and π^{-1}, respectively (rearrangement lema), one obtains $(\hat{\hat{P}}^{(+)}A)P^{(\pm)} = P^{(\pm)}(\hat{\hat{P}}^{(+)}A)$.

Solution 5.13. The operators $K_\pm = \sum_{p=1}^{2k+1} k_{p\pm}$ are symmetric (additive) operators, thus commuting with the antisymmetrizer (recall that otherwise the relevant space $\mathcal{S}_-^{2k+1}$ would not be invariant; formally since $\hat{\hat{P}}K_\pm = K_\pm$, this follows from (5.27)). Thus the $2k+1$-particle state (5.14) is defined by the occupation numbers $n_m = 1$, $m = -k, \dots, k$ (using single particle basis $|k, m\rangle$):

$$|\boldsymbol{n}^-\rangle = \sqrt{N!}P_{\boldsymbol{n}}^{(-)}|\boldsymbol{n}\rangle = \sqrt{\frac{1}{N!}}\sum_\pi (-)^{\tilde{\pi}}\Delta(\pi)|k, -k; \dots; k, k\rangle\,,$$

giving

$$K_\pm|\boldsymbol{n}^-\rangle = \sqrt{\frac{1}{N!}}\sum_{p=1}^{2k+1}\sum_\pi (-)^{\tilde{\pi}}k_{p\pm}\Delta(\pi)|k, -k; \dots; k, k\rangle\,.$$

Now note that each summand $k_{p\pm}\Delta(\pi)|k, -k; \dots; k, k\rangle$ is zero: either at position p after action of $\Delta(\pi)$ appears $k_{p\pm}|k, m = \pm k\rangle = 0$, or $k_{p\pm}|k, \pm m\rangle \sim |k, \pm m \pm 1\rangle$, the state in which is another electron already. However, $K_\pm|\boldsymbol{n}^-\rangle = K_\pm P^{(-)}|\boldsymbol{n}^-\rangle = P^{(-)}K_\pm|\boldsymbol{n}^-\rangle$, and we can apply on each summand another antisymmetrization, resulting in the vanishing vector. Finally, $K_\pm|\boldsymbol{n}^-\rangle = 0$ means that $|\boldsymbol{n}^-\rangle$ is of the angular momentum 0 (only standard vector $|K = 0, M = 0\rangle$ obeys this).

Solution 5.14. Acting on the Slater determinant $|1, \dots, N\rangle$ (as $N = |\mathcal{S}|$) one gets
$(\otimes_{p=1}^N D_p)|1, \dots, N\rangle = \frac{1}{N!}\sum_\pi(-1)^\pi\sum_{j_1,\dots,j_N} D_{j_1,\pi^{-1}1}\cdots D_{j_N,\pi^{-1}N}|j_1\rangle\otimes\cdots\otimes|j_N\rangle$, which is (reordering the matrix element) $\frac{1}{N!}\sum_\pi(-1)^\pi\sum_{j_1,\dots,j_N} D_{j_{\pi 1}1}\cdots D_{j_{\pi N}N}|j_1\rangle\otimes\cdots\otimes|j_N\rangle$. Changing $j_{\pi p}$ to j_p gives $\frac{1}{N!}\sum_\pi(-1)^\pi\sum_{j_1,\dots,j_N} D_{j_1 1}\cdots D_{j_N N}|j_{\pi^{-1}1}\rangle\otimes\cdots\otimes|j_{\pi^{-1}N}\rangle = \sum_{j_1,\dots,j_N} D_{j_1 1}\cdots D_{j_N N}P^-|j_1, \dots, j_N\rangle$. The antisymmetrizer annihilates the terms with repeated indices j_p; as $P^-|j_1, \dots, j_N\rangle = (-1)^{\pi(j_1,\dots,j_N)}|1, \dots, N\rangle$, this becomes
$(\sum_{j_1,\dots,j_N}(-1)^{\pi(j_1,\dots,j_N)}D_{j_1 1}\cdots D_{j_N N}) = (\det D)|1, \dots, N\rangle$. The second part of the Exercise is an obvious generalization of the well known equality $\det(A\otimes B) = (\det A)^{|B|}(\det B)^{|A|}$.

Solution 5.15. In general, in the orthonormalized basis $|\boldsymbol{i}\rangle$ ($\boldsymbol{i} = (i_1 < \cdots < i_N)$ are ordered N-tuples of the ordinals of a basis of $\mathcal{S}$) of the $\mathcal{S}_-^N$ according to (5.14):

$$\begin{aligned}\langle\boldsymbol{i}|D^{\otimes N}|\boldsymbol{j}\rangle &= \tfrac{1}{N!}\sum_{\pi,\rho}(-)^{\pi+\rho}D^{i_{\pi^{-1}1}}_{j_{\rho^{-1}1}}\cdots D^{i_{\pi^{-1}N}}_{j_{\rho^{-1}N}} = (\text{with } \alpha = \pi\rho^{-1})\\ &= \tfrac{1}{N!}\sum_\pi\sum_\alpha(-)^\alpha D^{i_1}_{j_{\alpha 1}}\cdots D^{i_N}_{j_{\alpha N}} = \sum_\alpha(-)^\alpha D^{i_1}_{j_{\alpha 1}}\cdots D^{i_N}_{j_{\alpha N}} = ([D^{\otimes N}]^-)^{\boldsymbol{i}}_{\boldsymbol{j}}.\end{aligned}$$

If $a = |A|$ and $n = N - a$ (clearly $a \leq N \leq |S|$), the proposed states are spanned by $|\boldsymbol{a}, \boldsymbol{k}\rangle$; here $\boldsymbol{a} = (1, \dots, a)$, while $\boldsymbol{k}$ is an ordered n-tuple of the numbers $a+1, \dots, |\mathcal{S}|$. When this is applied to the found matrix elements:

$$\begin{aligned}\langle \boldsymbol{a}, \boldsymbol{k}|\, D^{\otimes N}\, |\boldsymbol{a}, \boldsymbol{l}\rangle &= \sum_{\pi} (-)^{\pi} D^{1}_{\pi 1} \cdots D^{a}_{\pi a} D^{k_1}_{l_{\pi 1}} \cdots D^{k_N}_{l_{\pi N}} = \sum_{\pi} (-)^{\pi} A^{1}_{\pi 1} \cdots A^{a}_{\pi a} B^{k_1}_{l_{\pi' 1}} \cdots B^{k_n}_{l_{\pi' n}} \\ &= \sum_{\alpha\nu} (-)^{\alpha+\nu} A^{1}_{\alpha 1} \cdots A^{a}_{\alpha a} B^{k_1}_{l_{\nu 1}} \cdots B^{k_n}_{l_{\nu n}} = \det A\, ([B^n]^-)^{\boldsymbol{k}}_{\boldsymbol{l}}.\end{aligned}$$

Here, $\alpha \in S_a$ and $\nu \in S_n$ ($S_a \otimes S_n < S_N$) are the permutations of the first a and the last n states only, since only the permutations $\pi = \alpha\nu$ may give nonvanishing matrix elements. Also, π' is permutation π, only with new indexes of $\boldsymbol{k}$ and $\boldsymbol{l}$ used, i.e. if $\pi p = q$ and $p, q > a$, then $\pi'(p - a) = (q - a)$ (if $q \leq a$ then the contribution vanish).

Solution 5.17. For finite dimensional $\mathcal{S}$, the Fock's space has finite form $\mathcal{F}_- = \oplus_{N=0}^{|\mathcal{S}|} \mathcal{S}_-^N$ (for $N > |\mathcal{S}|$ the summands are zero-dimensional). Thus, $|\mathcal{F}_-| = \sum_{N=0}^{|\mathcal{S}|} \binom{|\mathcal{S}|}{N} = 2^{|\mathcal{S}|}$, showing the isomorphism.

Solution 5.18. Compare the definition (5.34) to (5.21c). As for $a_s^\dagger$, we again consider its left action in $|\boldsymbol{n}'^{\pm}\rangle\, a_s^\dagger\, |\boldsymbol{n}^{\pm}\rangle = \sqrt{N}(\langle \boldsymbol{n}'^{\pm} \,|\, s\rangle)\, |\boldsymbol{n}^{\pm}\rangle$. Thus, the expression vanishes unless $|\boldsymbol{n}'^{pm}\rangle$ is the unique one satisfying $\langle s \,|\, \boldsymbol{n}'^{\pm}\rangle \sim |\boldsymbol{n}^{\pm}\rangle$. Apparently, this is just $|\boldsymbol{n}^{\pm}\rangle\, |s\rangle$ (everywhere it is assumed that $|s\rangle$ refers to the last particle). However, although it has the correct occupation number, it is not (anti)symmetrized, thus it is not in the Fock space: the adjoint of the operator in $\mathcal{F}_{\pm}$ must be also an operator in the same space. Thus this vector is to be (anti)symmetrized, but previously quantum number s is to be put (i.e. permuted by σ) to the position p_s, such that the states after it are with greater quantum numbers. Altogether, this exactly gives the same result as (5.35). The generalization (5.36) is obtained by the successive application of the first relation. N-independent form is achieved using the number of particles operator: $\langle j_1, \dots, j_L \,|\, \boldsymbol{n}^{\pm}\rangle = a_{j_L} \frac{1}{\sqrt{\hat{N}}} \cdots a_{j_1} \frac{1}{\sqrt{\hat{N}}}\, |\boldsymbol{n}^{\pm}\rangle$.

Solution 5.19. Consider a particle with one dimensional state space. Due to this peculiarity, all the observables are real numbers, among them the hamiltonian being $\hbar\omega$ ($\omega \geq 0$). Further, any N-particle space is again one dimensional, spanned by the vector $|N\rangle$ (as there is only one single particle basis vector, occupation number vector is simply the total number of particles). The particles do not interact, making the N-particle hamiltonian $N\hbar\omega$. The particles may be created or annihilated, thus they must be considered within Fock's space, which is exactly the state space of the harmonic oscillator. It only remains to shift the energy scale for $\bar{\omega}/2$, and the oscillator dynamics is completely reproduced by the described noninteracting particles, which really correspond to the excitations of the harmonic oscillator. The commutation relation (2.27) shows that such excitations are of Bose type.

Solution 5.20. Consider an arbitrary matrix element $A^{\boldsymbol{n}}_{\boldsymbol{n}'} = \langle \boldsymbol{n}^{\pm}|\, A^{[L]}\, |\boldsymbol{n}'^{\pm}\rangle$ in the occupation number basis:

$$A^{\boldsymbol{n}}_{\boldsymbol{n}'} = \sum_{p_1<\cdots<p_l} \sum_{\substack{s_1,\dots,s_L \\ s'_1,\dots,s'_L}} A^{s_1,\dots,s_L}_{s'_1,\dots,s'_L} \langle \boldsymbol{n}^{\pm} \,|\, s_1^{(p_1)},\dots,s_L^{(p_L)}\rangle \langle s_1'^{(p_1)},\dots,s_L'^{(p_L)} \,|\, \boldsymbol{n}'^{\pm}\rangle\,;$$

superscript is particle counter, i.e. noncorrelated vector $|s_1^{(p_1)},\dots,s_L^{(p_L)}\rangle$ describes a system of particles $p_1 < \cdots < p_L$, with the particle p_j in the state s_j. To apply (5.36), one previously should move particles $p_1 < \cdots < p_L$ to the first L places; however, the sign arising (for fermions only) in such a permutation, appears both for bra and ket. Thus, summation in particles gives $\binom{N}{L}$ same terms:

$$A^{\boldsymbol{n}}_{\boldsymbol{n}'} = \binom{N}{L} \sum_{\substack{s_1,\dots,s_L \\ s'_1,\dots,s'_L}} A^{s_1,\dots,s_L}_{s'_1,\dots,s'_L} \langle \boldsymbol{n}^{\pm} \,|\, s_1,\dots,s_L\rangle \langle s'_1,\dots,s'_L \,|\, \boldsymbol{n}'^{\pm}\rangle\,.$$

Now, (5.36) directly gives the equality of the arbitrary matrix elements of the operators in (5.42c).

Solution 5.21. Defining operators a_{Qkm} corresponding to the basis $|Qkm\rangle$, and noticing that $K_{\pm}$ is additive single particle operator such that $\langle Qkm|\, k_{\pm}\, |Q'k'm'\rangle = \alpha^{\pm}_{m'}\delta_{QQ'}\delta_{kk'}\delta_{m\mp1,m'}$, one finds $K_{\pm} = \sum_m \alpha^{\pm}_m a^{\dagger}_m a_{m\mp1}$. The closed shell corresponds to the state $|\boldsymbol{n}^{-}\rangle$, such that the occupation numbers of all the states are 0 except that $n_{Qkm} = 1$, for $m = -l,\dots,l$. Obviously, $K_{\pm}$ maps this state into the vector with occupation numbers greater than 1, i.e. into the vanishing vector.

B.0.6 Approximate Methods

Solution 6.1. The perturbation quantum numbers are $l = 1$ and $m = 0$, and -1 with respect to spatial inversion; recall that the later is $(-1)^l$ for the states. Therefore, the spin space selection rule is $\Delta m_s = 0$, and consequently $\Delta m = 0$, while the transition is allowed for $\Delta l = \pm1$ (since $\Delta l = 0$ is prevented by the spatial inversion). This leaves the wanted form (6.6), with the reduced matrix elements $V^{\pm}$. Clearly, in the first order, there is no correction to E_1 (as $l = 0$). As for $n = 2$, it is fourfold degenerate (nonperturbed, spinless) with the basis $|11\rangle$, $|10\rangle$, $|1,-1\rangle$ and $|00\rangle$. The perturbation matrix elements vanish for $m \neq m'$, as well as when $l = l'$. This means that the linear effect exists only in the subspace spanned by $m = 0$ vectors $|210\rangle$ and $|200\rangle$: the reduced matrix is $H' = eE\begin{pmatrix} \langle 210|z|210\rangle=0 & \langle 210|z|200\rangle \\ \langle 210|z|200\rangle & \langle 200|z|200\rangle=0 \end{pmatrix}$ has eigenvalues $E^{(1)} = \pm|\langle 210|\, z\, |200\rangle\,||e|E = \pm 3eEa_0$.

Solution 6.2. The nonperturbed term is kinetic rotational energy (rigid rotator) with the eigensolutions $E^0_l = l(l+1)\hbar^2/2mr_0^2$ and $Y^m_l(\theta,\varphi)$, respectively. Perturbation is z-component of the vector operator, and since $Y^0_1(\theta,\varphi) \sim \cos\theta$, the matrix elements are:

$$\langle lm|\, H'\, |l'm'\rangle \sim \int_0^{\pi}\int_0^{2\pi} \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi Y^{m*}_l(\theta,\varphi) Y^0_1(\theta,\varphi) Y^{m'}_{l'}(\theta,\varphi).$$

They vanish unless the selection rules $\Delta m = m - m' = 0$ and $\Delta l = l - l' = 0, \pm1$ are satisfied: the first one is conservation of the z-component of angular momentum, while the second one means that $D^{(l)}$ is a component of $D^{(1)} \otimes D^{(l')} = D^{(l'-1)} \oplus D^{(l')} \oplus D^{(l'+1)}$. Analogous considerations for the spatial inversion exclude the term l', as on the left side the parity is $(-1)^(l'+1)$. This immediately means that the first order correction vanishes. The second order perturbation is then obtained straightforwardly.

Solution 6.3. Taking the absolute value of (6.10) one gets the proposed estimation directly:

$$|E_n^{(2)}| \leq \sum_{m(\neq n),\mu} \frac{|\langle m\mu;0|\, H' \,|n;0\rangle\,|^2}{|E_n^{(0)} - E_m^{(0)}|} \leq \frac{\sum_{m(\neq n),\mu} |\langle m\mu;0|\, H' \,|n;0\rangle\,|^2}{|\triangle E_n^{(0)}|} =$$

$$\frac{\langle n;0|\, H'(I - |n;0\rangle\,\langle n;0|)H' \,|n;0\rangle}{|\triangle E_n^{(0)}|} = \frac{\langle {H'}^2\rangle - \langle H'\rangle^2}{|\triangle E_n^{(0)}|}.$$

Solution 6.4. (i) Let E_n be the eigenenergy closest to the $E[|\psi\rangle]$. Then, using the mean in $|\psi\rangle$ of the spectral form $(H - E[|\psi\rangle])^2 = \sum_i (E_i - E[|\psi\rangle])^2 P_i$, one gets

$$\triangle^2_{|\psi\rangle} H \geq \sum_i \langle\psi|\, (E_n - E[|\psi\rangle])^2 P_i \,|\psi\rangle = (E_n - E[|\psi\rangle])^2.$$

(ii) For any vector $|\psi\rangle$ orthogonal to $|\psi_0\rangle$ one finds

$$1 = \langle 0\,|\,0\rangle = \langle 0\,|\,\psi_0\rangle\,\langle\psi_0\,|\,0\rangle + \langle 0\,|\,\psi\rangle\,\langle\psi\,|\,0\rangle + \cdots \leq 1 - \epsilon + |\,\langle 0\,|\,\psi\rangle\,|^2,$$

which proves the left inequality. The right one follows from:
$E[|\psi\rangle] = \sum_{nn'} \langle\psi\,|\,n\rangle\,\langle n|\, H \,|n'\rangle\,\langle n'\,|\,\psi\rangle = \sum_n E_n |\,\langle\psi\,|\,n\rangle\,|^2 =$
$E_0 |\,\langle\psi\,|\,0\rangle\,|^2 - E_1 |\,\langle\psi\,|\,0\rangle\,|^2 + E_1 |\,\langle\psi\,|\,0\rangle\,|^2 + \sum_{n(>0)} E_n |\,\langle\psi\,|\,n\rangle\,|^2 \geq$
$(E_0 - E_1)|\,\langle\psi\,|\,0\rangle\,|^2 + E_1 \sum_n |\,\langle\psi\,|\,n\rangle\,|^2 \geq -(E_0 - E_1)\epsilon + E_1.$

Solution 6.5. $P_n^2 = P_n$ implies $\frac{\mathrm{d}P_n(s)}{\mathrm{d}s} = \frac{\mathrm{d}P_n(s)}{\mathrm{d}s} P_n(s) + P_n(s) \frac{\mathrm{d}P_n(s)}{\mathrm{d}s}$; multiplying again by a projector one gets $P_n(s) \frac{\mathrm{d}P_n(s)}{\mathrm{d}s} P_n(s) = 0$.
Analogously, $\frac{\mathrm{d}\mathbb{1}}{\mathrm{d}s} = \frac{\mathrm{d}\sum_n P_n(s)}{\mathrm{d}s} = \sum_n (P_n(s) \frac{\mathrm{d}P_n(s)}{\mathrm{d}s} + \frac{\mathrm{d}P_n(s)}{\mathrm{d}s} P_n(s)) = 0$.
Due to (b) Ω is Hermitian. $\Omega P_m - P_m \Omega = -\mathrm{i}\hbar \sum_n (P_n \frac{\mathrm{d}P_n}{\mathrm{d}s} P_m - P_m P_n \frac{\mathrm{d}P_n}{\mathrm{d}s}) =$
$-\,\mathrm{i}\hbar(\sum_n (\frac{\mathrm{d}P_n}{\mathrm{d}s} - \frac{\mathrm{d}P_n}{\mathrm{d}s} P_n) P_m - P_m \frac{\mathrm{d}P_m}{\mathrm{d}s}) = \mathrm{i}\hbar(\frac{\mathrm{d}P_m}{\mathrm{d}s} P_m + P_m \frac{\mathrm{d}P_m}{\mathrm{d}s}) = \mathrm{i}\hbar \frac{\mathrm{d}P_m}{\mathrm{d}s}.$

B.0.7 Elements of Quantum Information

Solution 7.1. $A^2 = \sum_{i,j=0}^3 a_i a_j \sigma_i \sigma_j = \sum_{i=0}^3 a_i^2 + 2\sum_{i=1}^3 a_0 a_i \sigma_i + \sum_{1\leq i<j\leq 3}^3 (\sigma_i\sigma_j + \sigma_j\sigma_i)$; result follows from (7.2). Identity operator is obtained for $a_i = 0$ except $a_0 = 1$; projectors with one-dimensional range have trace 1, thus $a_0 = \frac{1}{2}$, while $A^2 = A$ implies $\sum_{i=1}^3 a_i^2 = \frac{1}{4}$, i.e. $A = \frac{1}{2}(\sigma_0 + \boldsymbol{a}\cdot\boldsymbol{\sigma})$ with $|\boldsymbol{a}| = 1$.

Solution 7.2. If $S = s\mathbb{1}$ is scalar observable, then $(S, X) = \mathrm{Tr}\, S^\dagger X = s(\mathrm{Tr}\, X)^*$. This vanishes if and only if $\mathrm{Tr}\, X = 0$. In the opposite direction, if A is orthogonal to all traceless operators, then for each $i \neq j$ A is orthogonal to all operators $|i\rangle\,\langle j|$ (meaning that $A^i_j = 0$, i.e. A is diagonal), as well as to the operators $|i\rangle\,\langle i| - |j\rangle\,\langle j|$, giving that the diagonal elements of A are mutually equal.

Solution 7.3. a) For $\boldsymbol{a} \in \mathbb{R}^3$ with $|\boldsymbol{a}| \leq 1$ the obtained matrix is $\boldsymbol{a} \cdot \boldsymbol{\sigma} = \left(\begin{smallmatrix} a_3 & a_1-\mathrm{i}a_2 \\ a_1+\mathrm{i}a_2 & -a_3 \end{smallmatrix}\right)$. Then the result directly follows substituting $|\boldsymbol{x}\rangle = \mathrm{e}^{\mathrm{i}\gamma}(\cos\frac{\alpha}{2}, \mathrm{e}^{\mathrm{i}\beta}\sin\frac{\alpha}{2})$ with $\boldsymbol{x} = (1, \alpha, \beta)$. b) From the first part $\langle \boldsymbol{x}| \rho_{\boldsymbol{a}} |\boldsymbol{x}\rangle = \frac{1+\boldsymbol{a}\cdot\boldsymbol{x}}{2}$. Clearly, $0 \leq \langle \boldsymbol{x}| \rho_{\boldsymbol{a}} |x\rangle \leq 1$for $|\boldsymbol{a}| \leq 1$, proving that the operator is positive. Obviously its trace is 1. Obviously, if $|\boldsymbol{a}| = 1$ then $\langle \boldsymbol{a}| \rho_{\boldsymbol{a}} |\boldsymbol{a}\rangle = 1$, meaning that $\rho_{\boldsymbol{a}}$ is projector ($|\boldsymbol{a}\rangle$ is vector of maximal expectation, thus the eigenvector, and its eigenvalue is 1).

Solution 7.4. M^+M is positive definite $n_2 \times n_2$ matrix with eigensystem $M^+Mv_i = \sigma_i^2 v_i$. Gather orthonormal columns v_i into V and $u_i = Mv_i/\sigma_i$ into U. Clearly $(U^+MV)_{pq} = \sum_{ij} u_p^* M v_q$ is $\sigma_p \delta_{pq}$ for $p, q \leq r$ and 0 otherwise, meaning that $M = U\Sigma V^+$.

Solution 7.5. This is exactly the established correspondence of the antisymmetrized states and single-particle subspaces. It can be seen by the obvious fact that if $U|a; i\rangle = |b; i\rangle$, then $|\Psi_a; -\rangle = \det U |\Psi_b; -\rangle$, and $\det U = \mathrm{e}^{\mathrm{i}\phi}$ for unitary U.

Solution 7.6. As $|\Psi-\rangle$ is independent of the basis, the basis $|\boldsymbol{a}; \pm\rangle$ is used. Exercise 7.3 shows that $\langle \boldsymbol{a}; \pm| \boldsymbol{b}\cdot\boldsymbol{\sigma} |\boldsymbol{a}; \pm\rangle = \pm \boldsymbol{a}\cdot\boldsymbol{b}$. Therefore, $\langle \Psi-| (\boldsymbol{a}\cdot\boldsymbol{\sigma} \otimes \mathbb{1})(\mathbb{1} \otimes \boldsymbol{b}\cdot\boldsymbol{\sigma}) |\Psi-\rangle = \sum_{a=\pm} \frac{\langle \boldsymbol{a};a|\boldsymbol{a}\cdot\boldsymbol{\sigma}|\boldsymbol{a};a\rangle\langle \boldsymbol{a};-a|\boldsymbol{b}\cdot\boldsymbol{\sigma}|\boldsymbol{a};-a\rangle - \langle \boldsymbol{a};a|\boldsymbol{a}\cdot\boldsymbol{\sigma}|\boldsymbol{a};-a\rangle\langle \boldsymbol{a};-a|\boldsymbol{b}\cdot\boldsymbol{\sigma}|\boldsymbol{a};a\rangle}{2}$. The first factor is the matrix element of the observable $\boldsymbol{a}\cdot\boldsymbol{\sigma}$ between its eigenvectors, hence in the first summands this is $a = \pm 1$, while vanishes in the second (crossed) summand). The remaining terms are expectation of $\boldsymbol{b}\cdot\boldsymbol{\sigma}$ in the states $|-\boldsymbol{a}\rangle$.

Solution 7.7. Let the initial composite state be ρ, and arbitrary subsystem observables $A = \sum_i a_i P_i^{(1)} \otimes \mathbb{1}_2$ and $B = \mathbb{1}_1 \otimes \sum_j b_i Q_j^{(2)}$. Then the measurement of A results in $\rho^A = \sum_i (P_i^{(1)} \otimes \mathbb{1}_2 \rho P_i^{(1)} \otimes$. The measurement of the subsystem observable B gives the same result as it has been performed on the "subsystem state", the partial trace of the whole system. So, if there was no measurement of A, this would be $\rho_2 = \mathrm{Tr}_1[\rho]$, and $\rho_2^A = \sum_i \mathrm{Tr}_1[P_i^{(1)} \otimes \mathbb{1}_2 \rho P_i^{(1)} \otimes \mathbb{1}_2]$. However, choosing in the first subsystem space the basis adapted to the spectral decomposition of A, it becomes transparent that $\rho_2 = \rho_2^A$, i.e. that the results of the measurement of A do not affect the other measurement; actually, the other subsystem have no information if the firs subsystem measurement is performed, either.

Solution 7.8. Assume that $U|a\rangle \otimes |b\rangle = \mathrm{e}^{\mathrm{i}f(a,b)} |a\rangle \otimes |a\rangle$ and $U|a'\rangle \otimes |b\rangle = \mathrm{e}^{\mathrm{i}f(a',b)} |a'\rangle \otimes |a'\rangle$. Then the scalar products of the initial and final states are $\langle a \,|\, a'\rangle$ and $\langle a \,|\, a'\rangle^2$. As U is unitary, these are equal, which restricts the initial states to either collinear or orthogonal, contradicting to the required arbitrariness.

Literature

[1] M. Abramovitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1972. A.2

[2] E. Artin. Theorie der zopfe. *Abh. Math. Semin. Univ. Hambg.*, 4:47–72, 1925. 5.1.1

[3] J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1:195–200, Nov 1964. 7.3.1

[4] M. V. Berry. Quantal phase factors accompanying adiabatic changes. *Proc. Roy. Soc. Lond. A*, 392:45–57, 1984. 6.5.2

[5] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969. 7.3.1

[6] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935. 7.3

[7] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Vol. 3)*. Addison-Wesley, Reading, Massachusetts, the definitive edition (2nd edition) edition, 1964. 1.1.2

[8] A. M. Gleason. Measures on the closed subspaces of a hilbert space. *J. Math. Mech.*, 6:885–93, 1957. 1.1.6

[9] T. Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Springer, Berlin Heidelberg, 1995. 6.1.2

[10] A. Kitaev. *Annals of Physics*, 321:2–111, 2006. 5.1.1

[11] L. D. Landau and E.M. Lifshitz. *Quantum Mechanics Non-Relativistic Theory (Course of Theoretical Physics, Vol. 3)*. Butterworth-Heinemann, Oxford, third edition, 1981. 2, 2.4.1, 4.4.2

[12] B. V. Medvedev. *Nachala teoreticheskoy fiziki: mehanika, teoriya polya, elementy kvantovoy mehaniki (in Russian)*. Nauka, 1977. 3.2

[13] A. Messiah. *Quantum Mechanics*. North-Holland, Amsterdam, 1972. 6.1.2, 6.5.1

[14] J. Nakamura. *Nat. Phys.*, 16:931–936, 2020. 5.1.1

[15] S. Das Sarma, M. Freedman, and C. Nayak. *npj Quantum Information*, 1:15001, 2015. 5.1.1

[16] Raymond F. Streater and Arthur S. Wightman. *PCT, Spin and Statistics, and All That*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 2000. 5.1.1

[17] Milan Vujičić and Fedor Herbut. A quantum-mechanical theory of distant correlations. *Journal of Mathematical Physics*, 25(7):2253–2259, 1984. 7.2

[18] Steven Weinberg. *Lectures on Quantum Mechanics*. Cambridge: Cambridge University Press, 2013. 6