Quantum Mechanics 3:
the quantum mechanics of many-particle systems

Common part of the course (3rd quarter)

W.J.P. Beenakker

Academic year 2018 – 2019

Contents of the common part of the course:

1) Occupation-number representation

2) Quantum statistics (up to § 2.5)

The following books have been used:

F. Schwabl, "Advanced Quantum Mechanics", third edition (Springer, 2005);

David J. Griffiths, "Introduction to Quantum Mechanics", second edition (Prentice Hall, Pearson Education Ltd, 2005);

Eugen Merzbacher, "Quantum Mechanics", third edition (John Wiley & Sons, 2003);

B.H. Bransden and C.J. Joachain, "Quantum Mechanics", second edition (Prentice Hall, Pearson Education Ltd, 2000).

Contents

1 Occupation-number representation

In this chapter the quantum mechanics of identical-particle systems will be worked out in detail. The corresponding space of quantum states will be constructed in the occupation-number representation by employing creation and annihilation operators. This will involve the introduction of the notion of quasi particles and the concept of second quantization.

Similar material can be found in Schwabl (Ch. 1, 2 and 3) and Merzbacher (Ch. 21, 22 and the oscillator part of Ch. 14).

1.1 Summary on identical particles in quantum mechanics

Particles are called identical if they cannot be distinguished by means of specific intrinsic properties (such as spin, charge, mass, $\cdots$).

This indistinguishability has important quantum mechanical implications in situations where the wave functions of the identical particles overlap, causing the particles to be observable simultaneously in the same spatial region. Examples are the interaction region of a scattering experiment or a gas container. If the particles are effectively localized, such as metal ions in a solid piece of metal, then the identity of the particles will not play a role. In those situations the particles are effectively distinguishable by means of their spatial coordinates and their wave functions have a negligible overlap.

For systems consisting of identical particles two additional constraints have to be imposed while setting up quantum mechanics (QM).

- Exchanging the particles of a system of identical particles should have no observable effect, otherwise the particles would actually be distinguishable. This gives rise to the concept of permutation degeneracy, i.e. for such a system the expectation value for an arbitrary many-particle observable should not change upon interchanging the identical particles in the state function. As a consequence, quantum mechanical observables for identical-particle systems should be symmetric functions of the separate 1-particle observables.

- Due to the permutation degeneracy, it seems to be impossible to fix the quantum state of an identical-particle system by means of a complete measurement. Nature has bypassed this quantum mechanical obstruction through the

  symmetrization postulate: identical-particle systems can be described by means of either totally symmetric state functions if the particles are bosons or totally antisymmetric state functions if the particles are fermions. In the non-relativistic QM it is an empirical fact that no mixed symmetry occurs in nature.

In this chapter the symmetrization postulate will be reformulated in an alternative way, giving rise to the conclusion that only totally symmetric/antisymmetric state functions will fit the bill.

Totally symmetric state functions can be represented in the so-called q-representation by $\psi_S(q_1, \cdots, q_N, t)$, where $q_1, \cdots, q_N$ are the "coordinates" of the N separate identical particles. These coordinates are the eigenvalues belonging to a complete set of commuting 1-particle observables $\hat{q}$. Obviously there are many ways to choose these coordinates. A popular choice is for instance $q_j = (\text{spatial coordinate } \vec{r}_j\,, \text{ magnetic spin quantum number } m_{s_j} \equiv \sigma_j\,, \cdots)$, where the dots represent other possible internal (intrinsic) degrees of freedom of particle j. For a symmetric state function we have

$$\boxed{\underset{\hat{P}}{\forall}\ \hat{P}\,\psi_S(q_1, \cdots, q_N, t) \;=\; \psi_S(q_{P(1)}, \cdots, q_{P(N)}, t) \;=\; \psi_S(q_1, \cdots, q_N, t)} \;, \tag{1}$$

with $\hat{P}$ a permutation operator that permutes the sets of coordinates of the identical particles according to

$$q_1 \;\rightarrow\; q_{P(1)} \quad , \quad q_2 \;\rightarrow\; q_{P(2)} \quad , \quad \cdots \quad , \quad q_N \;\rightarrow\; q_{P(N)} \;. \tag{2}$$

Particles whose quantum states are described by totally symmetric state functions are called bosons. They have the following properties:

- bosons have integer spin (see Ch. 4 and 5);
- bosons obey so-called Bose–Einstein statistics (see Ch. 2);
- bosons prefer to be in the same quantum state.

Totally antisymmetric state functions can be represented in the q-representation by $\psi_A(q_1, \cdots, q_N, t)$, with

$$\boxed{\underset{\hat{P}}{\forall}\ \hat{P}\,\psi_A(q_1, \cdots, q_N, t) \;=\; \psi_A(q_{P(1)}, \cdots, q_{P(N)}, t) \;=\; \begin{cases} +\,\psi_A(q_1, \cdots, q_N, t) & \text{even } \hat{P} \\ -\,\psi_A(q_1, \cdots, q_N, t) & \text{odd } \hat{P} \end{cases}} \;. \tag{3}$$

A permutation $\hat{P}$ is called even/odd if it consists of an even/odd number of two-particle interchanges. Particles whose quantum states are described by totally antisymmetric state functions are called fermions. They have the following properties:

- fermions have half integer spin (see Ch. 5);
- fermions obey so-called Fermi–Dirac statistics (see Ch. 2);
- fermions are not allowed to be in the same quantum state.

If the above (anti)symmetrization procedure results in spatial symmetrization, the particles have an increased probability to be in each other's vicinity. Note that this can apply to identical fermions if they happen to be in an antisymmtric spin state. On the other hand, spatial antisymmetrization gives rise to a decreased probability for the particles to be in each other's vicinity. Note that this can apply to identical spin-1 bosons if they happen to be in an antisymmtric spin state.

Isolated non-interacting many-particle systems: many-particle systems with negigible interactions among the particles are called non-interacting many-particle systems. The properties of such systems are determined by the type(s) of particles involved and the possible influence of external potentials on the system (e.g. caused by a magnetic field). We speak of an ideal gas if the number of non-interacting particles is very large and if we place the system in a macroscopic finite enclosure.

> *Isolated non-interacting many-particle systems (such as free-particle systems, ideal gasses, ...) play a central role in this lecture course, both for determining the physical properties of identical-boson/fermion systems and for setting up relativistic QM.*

Since we are dealing here with isolated systems, the total energy is conserved and the state function can be expressed in terms of stationary states

$$\psi_E(q_1,\cdots,q_N,t) \;=\; \exp(-iEt/\hbar)\,\psi_E(q_1,\cdots,q_N)\;,$$
$$\text{with}\quad \hat{H}\,\psi_E(q_1,\cdots,q_N) \;=\; E\,\psi_E(q_1,\cdots,q_N)\;. \tag{4}$$

The Hamilton operator describing a non-interacting many-particle system comprises of pure 1-particle Hamilton operators, i.e. Hamilton operators that depend exclusively on observables belonging to individual single particles. Denoting the 1-particle Hamilton operator of particle j by $\hat{H}_j$, the N-particle Hamilton operator reads

$$\boxed{\hat{H} \;=\; \sum_{j=1}^{N} \hat{H}_j\;, \qquad \text{with}\;\; \left[\hat{H}_j\,,\hat{H}_k\right] \;=\; 0\;\;\text{for all}\;\; j,k=1,\cdots,N}\;. \tag{5}$$

This implies that the Hamilton operators of the individual particles are compatible observables, as expected for non-interacting particles. Suppose now that the 1-particle energy-eigenvalue equation

$$\hat{H}_j\,\psi_{\lambda_j}(q_j) \;=\; E_{\lambda_j}\,\psi_{\lambda_j}(q_j) \tag{6}$$

gives rise to an orthonormal set $\{\psi_{\lambda_j}(q_j)\}$ of energy eigenfunctions belonging to the energy eigenvalues E_{λ_j}, which are labeled by a complete set of quantum numbers. Examples of such complete sets of quantum numbers are for instance $\lambda_j = n_j$ for a linear harmonic

oscillator or $\lambda_j = (n_j\,, \ell_j\,, m_{\ell_j}\,, m_{s_j})$ for a 1-electron atom. For the orthonormal set of energy eigenstates of the complete non-interacting N-particle system we can identify three scenarios.

A) The particles are distinguishable. The orthonormal set of N-particle energy eigenstates comprises of states of the form

$$\psi_E(q_1, \cdots, q_N) \;=\; \psi_{\lambda_1}(q_1)\, \psi_{\lambda_2}(q_2) \cdots \psi_{\lambda_N}(q_N)\;, \qquad \text{with} \quad \boxed{E \;=\; \sum_{j=1}^{N} E_{\lambda_j}}\;. \tag{7}$$

Such so-called product functions describe an uncorrelated system for which the properties of a specific particle can be measured without being influenced by the other particles.

> *The complete set of all product functions spans the space of all possible N-particle states involving distinguishable particles.*

B) The particles are indistinguishable bosons. The orthonormal set of N-particle energy eigenstates comprises of totally symmetric states of the form

$$\psi_S(q_1, \cdots, q_N) \;=\; \frac{1}{\mathcal{N}_{_S}} \sum_{\text{diff.perm.}} \psi_{\lambda_1}(q_{_{P(1)}})\, \psi_{\lambda_2}(q_{_{P(2)}}) \cdots \psi_{\lambda_N}(q_{_{P(N)}})\;,$$

$$\text{with} \quad \mathcal{N}_{_S} \;=\; \sqrt{\text{number of different permutations of } \lambda_1, \cdots, \lambda_N}\;. \tag{8}$$

The possible energy eigenvalues are the same as in equation (7).

> *The space of bosonic N-particle states is spanned by a reduced set of linear combinations of product functions. As expected, this describes a correlated system for which the measurement of the properties of a specific particle is influenced by the other particles.*

C) The particles are indistinguishable fermions. The orthonormal set of N-particle energy eigenstates comprises of totally antisymmetric states of the form

$$\psi_A(q_1, \cdots, q_N) \;=\; \frac{1}{\sqrt{N!}} \sum_{\text{perm.}} (-1)^P\, \psi_{\lambda_1}(q_{_{P(1)}})\, \psi_{\lambda_2}(q_{_{P(2)}}) \cdots \psi_{\lambda_N}(q_{_{P(N)}})$$

$$= \;\frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{\lambda_1}(q_1) & \psi_{\lambda_2}(q_1) & \cdots & \psi_{\lambda_N}(q_1) \\ \psi_{\lambda_1}(q_2) & \psi_{\lambda_2}(q_2) & \cdots & \psi_{\lambda_N}(q_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{\lambda_1}(q_N) & \psi_{\lambda_2}(q_N) & \cdots & \psi_{\lambda_N}(q_N) \end{vmatrix}\;, \tag{9}$$

where the determinant is known as the Slater determinant. The possible energy eigenvalues are again the same as in equation (7), bearing in mind that $\lambda_1, \cdots, \lambda_N$

have to be all different. If two complete sets of quantum numbers λ_j and λ_k coincide, then two of the columns in the Slater determinant are identical and the totally antisymmetric eigenfunction ψ_A vanishes. This is known as the

Pauli exclusion principle for identical fermions: no two identical fermions can be in the same fully specified 1-particle quantum state.

Also the space of fermionic N-particle states is spanned by a reduced set of linear combinations of product functions. This too describes a correlated system for which the measurement of the properties of a specific particle is influenced by the other particles.

1.2 Occupation-number representation

Our aim: we want to construct the space of quantum states (Fock space) for a many-particle system consisting of an arbitrary number of unspecified identical particles. This Fock space makes no statements about the physical scenario that the considered particles are in, such as being subject to interactions, external influences, etc.. It simply is the complex vector space (Hilbert space, to be more precise) that includes all possible many-particle states, on which the quantum mechanical many-particle theory should be formulated. The actual construction of Fock space involves finding a complete set of basis states for the decomposition of arbitrary many-particle state functions. The properties of these basis states will fix the properties of the Fock space.

The general rules for the construction of Fock space are:

- states should not change when interchanging particles;
- in order to guarantee the superposition principle, Fock space should not change when changing the representation of the basis states.

1.2.1 Construction of Fock space

Consider identical particles of an unspecified type and assume $\hat{q}$ to be a corresponding complete set of commuting 1-particle observables. Take the fully specified eigenvalues of these observables to be exclusively discrete: q_j labeled by $j = 1, 2, \cdots$. The corresponding 1-particle basis of normalized eigenstates of $\hat{q}$ is indicated by $\{|q_j\rangle,\ j = 1, 2, \cdots\}$. Subsequently we span the space of many-particle states by means of (special) linear combinations of product functions constructed from these 1-particle basis states, in analogy to the 1-particle energy eigenfunctions that were used in § 1.1 to span the space of non-interacting many-particle states. Since we are dealing with identical particles, a legitimate

many-particle state can make no statement about the identity of a particle in a specific 1-particle eigenstate. Such a many-particle state can at best make statements about the number of particles n_j that reside in a given fully specified 1-particle eigenstate belonging to the eigenvalue q_j. These numbers are called occupation numbers and can take the values $0, 1, \cdots$ (if allowed).

Postulate (replacing the symmetrization postulate): the collective set of hermitian number operators $\hat{n}_1, \hat{n}_2, \cdots$, which count the number of identical particles in each of the 1-particle quantum states $|q_1\rangle, |q_2\rangle, \cdots$, form a complete set of commuting many-particle observables. The employed complete set of 1-particle observables $\hat{q}$ can be chosen freely.

The hidden postulate aspect is that the state space for interacting particles can be constructed from non-interacting building blocks that are based on 1-particle observables. Since the complete set of 1-particle observables can be chosen freely, the superposition principle is automatically incorporated without any restrictions. This will guarantee that no mixed symmetry will occur in Fock space.

Hence, the corresponding set of normalized eigenstates $|n_1, n_2, \cdots\rangle$ will span Fock space completely:

$$\begin{array}{rcl} \text{0-particle state} & : & |\Psi^{(0)}\rangle \equiv |0,0,\cdots\rangle \equiv \text{vacuum state}\,, \\ \text{1-particle states} & : & |\Psi_j^{(1)}\rangle \equiv |0,\cdots,0,n_j{=}1,0,\cdots\rangle \equiv |q_j\rangle\,, \\ & \cdots & \end{array} \tag{10}$$

with

$$\hat{n}_j|\cdots,n_j,\cdots\rangle = n_j|\cdots,n_j,\cdots\rangle\,, \tag{11}$$

where $\cdots$ indicates the other occupation numbers. The representation of states corresponding to this type of basis is called the occupation-number representation.

We will build up Fock space from the vacuum by adding particles to the basis states step by step. To this end we introduce the creation operator $\hat{a}_j^\dagger$, which adds a particle with quantum number q_j to a basis state $|\cdots,n_j,\cdots\rangle$ according to

$$\hat{a}_j^\dagger|\cdots,n_j,\cdots\rangle = \sqrt{n_j{+}1}\,\exp\big(i\alpha_j(\cdots,n_j,\cdots)\big)\,|\cdots,n_j{+}1,\cdots\rangle\,. \tag{12}$$

In principle the phase factor $\exp\big(i\alpha_j(\cdots,n_j,\cdots)\big)$ could depend on the chosen basis state $|\cdots,n_j,\cdots\rangle$ as well as the label j. From the orthonormality of the many-particle basis it

follows automatically that

$$\langle \cdots, n_j, \cdots | \hat{a}_j^\dagger | \cdots, n'_j, \cdots \rangle \quad \stackrel{(12)}{=\!=\!=} \quad \sqrt{n'_j+1}\, \exp\big(i\alpha_j(\cdots, n'_j, \cdots)\big)\, \delta_{n_j, n'_j+1}$$

$$\stackrel{\text{definition}}{=\!=\!=} \quad \langle \cdots, n'_j, \cdots | \hat{a}_j | \cdots, n_j, \cdots \rangle^* \; ,$$

such that

$$\hat{a}_j | \cdots, n_j, \cdots \rangle \;=\; \sqrt{n_j}\, \exp\big(-i\alpha_j(\cdots, n_j{-}1, \cdots)\big)\, | \cdots, n_j{-}1, \cdots \rangle \tag{13}$$

holds for the annihilation operator $\hat{a}_j$. This definition of the creation and annihilation operators is consistent with the notion of the vacuum as a state with no particles, since

$$\boxed{\hat{a}_j | \Psi^{(0)} \rangle \stackrel{(10),(13)}{=\!=\!=} 0 \quad \Rightarrow \quad \langle \Psi^{(0)} | \hat{a}_j^\dagger \;=\; 0} \; . \tag{14}$$

Note that $\hat{a}_j^\dagger \neq \hat{a}_j$, which implies that the creation and annihilation operators themselves are not observables. At this point we can fix the first part of the phase convention of the many-particle basis by specifying the action of $\hat{a}_j^\dagger$ and $\hat{a}_j$ on the basis states $|\Psi^{(0)}\rangle$ respectively $|\Psi^{(1)}\rangle$:

$$\hat{a}_j^\dagger | \Psi^{(0)} \rangle \;\equiv\; | \Psi_j^{(1)} \rangle \quad , \quad \hat{a}_j | \Psi_k^{(1)} \rangle \;\equiv\; \delta_{jk} | \Psi^{(0)} \rangle \quad \Rightarrow \quad \exp\big(i\alpha_j(0, 0, \cdots)\big) \;=\; 1 \; . \tag{15}$$

For arbitrary basis states we then have

$$\big[\hat{n}_j, \hat{a}_{k\neq j}^\dagger\big]\, | \cdots, n_j, \cdots, n_k, \cdots \rangle \stackrel{(11),(12)}{=\!=\!=} \big(n_j - n_j\big)\, \hat{a}_{k\neq j}^\dagger | \cdots, n_j, \cdots, n_k, \cdots \rangle \;=\; 0 \; ,$$

$$\big[\hat{n}_j, \hat{a}_j^\dagger\big]\, | \cdots, n_j, \cdots \rangle \stackrel{(11),(12)}{=\!=\!=} \big([n_j{+}1] - n_j\big)\, \hat{a}_j^\dagger | \cdots, n_j, \cdots \rangle \;=\; \hat{a}_j^\dagger | \cdots, n_j, \cdots \rangle \; ,$$

resulting in the commutator relations

$$\boxed{\big[\hat{n}_j, \hat{a}_k^\dagger\big] \;=\; \delta_{jk} \hat{a}_k^\dagger \qquad \stackrel{\text{herm. conj.}}{=\!=\!=\!\Rightarrow} \qquad \big[\hat{n}_j, \hat{a}_k\big] \;=\; -\,\delta_{jk} \hat{a}_k} \; . \tag{16}$$

The above-defined creation and annihilation operators can be linked directly to the associated number operators. To this end we consider the following matrix elements for arbitrary basis states:

$$\langle \cdots, n_j, \cdots | \hat{a}_j^\dagger \hat{a}_j | \cdots, n'_j, \cdots \rangle \stackrel{(13)}{=\!=\!=} n_j\, \delta_{n_j, n'_j} \; .$$

This matrix is duly diagonal and has the correct occupation numbers as eigenvalues, i.e.

$$\boxed{\hat{n}_j \;=\; \hat{a}_j^\dagger \hat{a}_j} \; . \tag{17}$$

Switching to another discrete 1-particle representation.

Next we switch from the 1-particle basis $\{|q_j\rangle,\ j=1,2,\cdots\}$ belonging to the complete set of observables $\hat{q}$ to a 1-particle basis $\{|p_r\rangle,\ r=1,2,\cdots\}$ belonging to the alternative complete set of observables $\hat{p}$. From the completeness relations it follows that

$$|q_j\rangle \ = \ \sum_r |p_r\rangle\langle p_r|q_j\rangle \qquad \text{and} \qquad |p_r\rangle \ = \ \sum_j |q_j\rangle\langle q_j|p_r\rangle \ , \tag{18}$$

where the transformation matrix $(c_{rj}) \equiv (\langle p_r|q_j\rangle)$ satisfies the unitarity conditions

$$\sum_r \langle q_k|p_r\rangle\langle p_r|q_j\rangle \ = \ \langle q_k|q_j\rangle \ = \ \delta_{jk} \qquad \text{and} \qquad \sum_j \langle p_r|q_j\rangle\langle q_j|p_v\rangle \ = \ \langle p_r|p_v\rangle \ = \ \delta_{rv} \ . \tag{19}$$

Fock space can be set up equally well in terms of these 1-particle eigenstates $|p_r\rangle$ and corresponding creation and annihilation operators $\hat{b}_r^\dagger$ and $\hat{b}_r$. If we take the two vacuum states in both representations to be identical, the following relations should hold for the old 1-particle basis states $|\Psi^{(0)}\rangle$, $|\Psi_j^{(1)}\rangle$ and the new ones $|\Phi^{(0)}\rangle$, $|\Phi_r^{(1)}\rangle$:

$$\begin{aligned}
&\text{vacuum state}: \ |\Phi^{(0)}\rangle \ \equiv \ |0,0,\cdots\rangle \ = \ |\Psi^{(0)}\rangle \ , \\
&\text{1-particle states}: \ \hat{a}_j^\dagger|\Psi^{(0)}\rangle \ = \ |\Psi_j^{(1)}\rangle \ = \ |q_j\rangle \ \overset{(18)}{=\!=\!=} \ \sum_r |p_r\rangle\langle p_r|q_j\rangle \\
&\qquad\qquad = \ \sum_r \langle p_r|q_j\rangle\,|\Phi_r^{(1)}\rangle \ = \ \sum_r \langle p_r|q_j\rangle\,\hat{b}_r^\dagger|\Phi^{(0)}\rangle \ = \ \sum_r \hat{b}_r^\dagger\,\langle p_r|q_j\rangle\,|\Psi^{(0)}\rangle \ .
\end{aligned} \tag{20}$$

Without loss of generality, the (1-particle) relation among the two sets of creation and annihilation operators can be extended to the entire Fock space:

$$\begin{aligned}
\hat{a}_j^\dagger \ = \ \sum_r \hat{b}_r^\dagger\,\langle p_r|q_j\rangle \qquad &\overset{\text{herm. conj.}}{\Longrightarrow} \qquad \hat{a}_j \ = \ \sum_r \hat{b}_r\,\langle q_j|p_r\rangle \ , \\
\hat{b}_r^\dagger \ = \ \sum_j \hat{a}_j^\dagger\,\langle q_j|p_r\rangle \qquad &\overset{\text{herm. conj.}}{\Longrightarrow} \qquad \hat{b}_r \ = \ \sum_j \hat{a}_j\,\langle p_r|q_j\rangle \ .
\end{aligned} \tag{21}$$

Just as in ordinary 1-particle quantum mechanics, the creation of a particle with quantum number q_j is equivalent to a linear superposition of independently created particles with all possible quantum numbers p_r, each with their own amplitude $\langle p_r|q_j\rangle$.

An important cross check of the procedure so far is provided by the total number operator

$$\boxed{\hat{N} \ = \ \sum_j \hat{n}_j} \ . \tag{22}$$

This operator counts the total number of particles of a given identical-particle system, which should be invariant under an arbitrary basis transformation. Indeed we find

$$\hat{N} \ \overset{(17)}{=\!=\!=} \ \sum_j \hat{a}_j^\dagger\hat{a}_j \ \overset{(21)}{=\!=\!=} \ \sum_{j,r,v} \hat{b}_r^\dagger\hat{b}_v\,\langle p_r|q_j\rangle\langle q_j|p_v\rangle \ \overset{(19)}{=\!=\!=} \ \sum_{r,v} \hat{b}_r^\dagger\hat{b}_v\,\delta_{rv} \ = \ \sum_r \hat{b}_r^\dagger\hat{b}_r \ .$$

Construction of Fock space: part one.

For an arbitrary many-particle state $|\Psi\rangle$ it should hold that $\hat{a}_j^\dagger \hat{a}_k^\dagger |\Psi\rangle$ and $\hat{a}_k^\dagger \hat{a}_j^\dagger |\Psi\rangle$ fundamentally describe the same state. This implies a relation of the type

$$\Big(\hat{a}_j^\dagger \hat{a}_k^\dagger \,-\, \exp\big(i\beta(\Psi,j,k)\big)\,\hat{a}_k^\dagger \hat{a}_j^\dagger\Big)\,|\Psi\rangle \;=\; 0\;,$$

where the phase factor $\exp\big(i\beta(\Psi,j,k)\big)$ can depend on the many-particle state $|\Psi\rangle$ as well as the labels j and k.

Next we invoke the superposition principle to prove that $\exp\big(i\beta(\Psi,j,k)\big)$ has to be independent of the many-particle state $|\Psi\rangle$. Consider to this end two different arbitrary many-particle basis states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ in the q-representation. The superposition principle then states that also $|\Psi\rangle = c_1|\Psi_1\rangle + c_2|\Psi_2\rangle$ with $c_{1,2} \in \mathbb{C}$ should describe a legitimate many-particle state. From this it follows that

$$\begin{aligned}
&c_1\,\exp\big(i\beta(\Psi,j,k)\big)\,\hat{a}_k^\dagger \hat{a}_j^\dagger |\Psi_1\rangle \,+\, c_2\,\exp\big(i\beta(\Psi,j,k)\big)\,\hat{a}_k^\dagger \hat{a}_j^\dagger |\Psi_2\rangle \\
&=\; \exp\big(i\beta(\Psi,j,k)\big)\,\hat{a}_k^\dagger \hat{a}_j^\dagger |\Psi\rangle \;=\; \hat{a}_j^\dagger \hat{a}_k^\dagger |\Psi\rangle \;=\; \hat{a}_j^\dagger \hat{a}_k^\dagger \,(c_1|\Psi_1\rangle + c_2|\Psi_2\rangle) \\
&=\; c_1\,\exp\big(i\beta(\Psi_1,j,k)\big)\,\hat{a}_k^\dagger \hat{a}_j^\dagger |\Psi_1\rangle \,+\, c_2\,\exp\big(i\beta(\Psi_2,j,k)\big)\,\hat{a}_k^\dagger \hat{a}_j^\dagger |\Psi_2\rangle.
\end{aligned}$$

Here $\hat{a}_k^\dagger \hat{a}_j^\dagger |\Psi_1\rangle$ and $\hat{a}_k^\dagger \hat{a}_j^\dagger |\Psi_2\rangle$ are again two different basis states. Therefore

$$\underset{\Psi_1,\,\Psi_2}{\forall} \;\exp\big(i\beta(\Psi_1,j,k)\big) \;=\; \exp\big(i\beta(\Psi_2,j,k)\big) \;\equiv\; \exp\big(i\beta(j,k)\big)\;,$$

which implies that

$$\underset{\Psi}{\forall}\;\Big(\hat{a}_j^\dagger \hat{a}_k^\dagger \,-\, \exp\big(i\beta(j,k)\big)\,\hat{a}_k^\dagger \hat{a}_j^\dagger\Big)\,|\Psi\rangle \;=\; 0 \quad\Rightarrow\quad \hat{a}_j^\dagger \hat{a}_k^\dagger - \exp\big(i\beta(j,k)\big)\,\hat{a}_k^\dagger \hat{a}_j^\dagger \;=\; 0\;.$$

Subsequently the representation independence of Fock space can be used by rewriting this operator identity according to equation (21) in terms of an arbitrary alternative 1-particle representation, for which a similar operator identity holds:

$$\sum_{r,v} \langle p_r|q_j\rangle\langle p_v|q_k\rangle\,\big[\,\hat{b}_r^\dagger \hat{b}_v^\dagger - \exp\big(i\beta(j,k)\big)\,\hat{b}_v^\dagger \hat{b}_r^\dagger\,\big] \;=\; 0 \qquad \text{and} \qquad \hat{b}_r^\dagger \hat{b}_v^\dagger - \exp\big(i\beta'(r,v)\big)\,\hat{b}_v^\dagger \hat{b}_r^\dagger \;=\; 0\;.$$

For each unique (r,v)-combination in the p-representation we find

$$\begin{aligned}
&\langle p_r|q_j\rangle\langle p_v|q_k\rangle\,\big[\,\exp\big(i\beta'(r,v)\big) \,-\, \exp\big(i\beta(j,k)\big)\,\big] \\
&+\,\langle p_v|q_j\rangle\langle p_r|q_k\rangle\,\big[\,1 \,-\, \exp\big(i\beta(j,k)\big)\,\exp\big(i\beta'(r,v)\big)\,\big] \;=\; 0\;.
\end{aligned}$$

Since this has to hold true for an arbitrary basis transformation, we eventually obtain

$$\exp\big(i\beta(j,k)\big) \;=\; \exp\big(i\beta'(r,v)\big) \;=\; \pm 1 \;\equiv\; \exp(i\beta)\;.$$

In this way we have found the phase factor $\exp(i\beta)$ to be universal and to give rise to two branches of solutions:

$$\begin{array}{l} \underset{\Psi}{\forall} \quad \left[\hat{a}_j^\dagger, \hat{a}_k^\dagger\right]|\Psi\rangle \;=\; 0 \;\;\Rightarrow\;\; \text{commutation relations}: \quad \left[\hat{a}_j^\dagger, \hat{a}_k^\dagger\right] \;=\; \left[\hat{a}_j, \hat{a}_k\right] \;=\; 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{or} \\ \underset{\Psi}{\forall} \quad \left\{\hat{a}_j^\dagger, \hat{a}_k^\dagger\right\}|\Psi\rangle \;=\; 0 \;\;\Rightarrow\;\; \text{anticommutation relations}: \quad \left\{\hat{a}_j^\dagger, \hat{a}_k^\dagger\right\} \;=\; \left\{\hat{a}_j, \hat{a}_k\right\} \;=\; 0\;, \end{array} \tag{23}$$

introducing the anticommutator

$$\left\{\hat{A}, \hat{B}\right\} \;\equiv\; \hat{A}\hat{B} + \hat{B}\hat{A}\;. \tag{24}$$

Construction of Fock space: part two.

Finally this result can be combined with the general commutation identities (16). To this end we use the fact that a commutator of the type $\left[\hat{A}\hat{B}, \hat{C}\right]$ can be written both in terms of commutators and anticommutators:

$$\boxed{\left[\hat{A}\hat{B}, \hat{C}\right] \;=\; \hat{A}\left[\hat{B}, \hat{C}\right] + \left[\hat{A}, \hat{C}\right]\hat{B} \;=\; \hat{A}\left\{\hat{B}, \hat{C}\right\} - \left\{\hat{A}, \hat{C}\right\}\hat{B}}\;. \tag{25}$$

In this way we find for the commuting set of solutions in equation (23) that

$$\left[\hat{n}_j, \hat{a}_k^\dagger\right] \overset{(17)}{=\!=} \left[\hat{a}_j^\dagger\hat{a}_j, \hat{a}_k^\dagger\right] \;=\; \hat{a}_j^\dagger\left[\hat{a}_j, \hat{a}_k^\dagger\right] + \left[\hat{a}_j^\dagger, \hat{a}_k^\dagger\right]\hat{a}_j \overset{(23)}{=\!=} \hat{a}_j^\dagger\left[\hat{a}_j, \hat{a}_k^\dagger\right] \overset{(16)}{=\!=} \delta_{jk}\hat{a}_j^\dagger\;,$$

and for the anticommuting set of solutions that

$$\left[\hat{n}_j, \hat{a}_k^\dagger\right] \overset{(17)}{=\!=} \left[\hat{a}_j^\dagger\hat{a}_j, \hat{a}_k^\dagger\right] \;=\; \hat{a}_j^\dagger\left\{\hat{a}_j, \hat{a}_k^\dagger\right\} - \left\{\hat{a}_j^\dagger, \hat{a}_k^\dagger\right\}\hat{a}_j \overset{(23)}{=\!=} \hat{a}_j^\dagger\left\{\hat{a}_j, \hat{a}_k^\dagger\right\} \overset{(16)}{=\!=} \delta_{jk}\hat{a}_j^\dagger\;.$$

All in all the creation and annihilation operators should thus satisfy either

$$\boxed{\begin{array}{l} \text{commutation relations}: \; \left[\hat{a}_j^\dagger, \hat{a}_k^\dagger\right] \;=\; \left[\hat{a}_j, \hat{a}_k\right] \;=\; 0 \quad , \quad \left[\hat{a}_j, \hat{a}_k^\dagger\right] \;=\; \delta_{jk}\hat{1}\;, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{the corresponding identical particles are called bosons}, \end{array}} \tag{26}$$

or

$$\boxed{\begin{array}{l} \text{anticommutation relations}: \; \left\{\hat{a}_j^\dagger, \hat{a}_k^\dagger\right\} \;=\; \left\{\hat{a}_j, \hat{a}_k\right\} \;=\; 0 \quad , \quad \left\{\hat{a}_j, \hat{a}_k^\dagger\right\} \;=\; \delta_{jk}\hat{1}\;, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{the corresponding identical particles are called fermions}. \end{array}} \tag{27}$$

In the next chapter it will be shown that these two branches of creation and annihilation operators for identical many-particle systems give rise to utterly different statistical theories. In the exercise class it will be proven that the (anti)commutation relations retain their form under an arbitrary change of representation, which implies that to a given type

of particle belongs a given type of statistics and that mixed statistics is indeed excluded.

In addition it will be shown that the many-particle states

$$\boxed{|n_1, n_2, \cdots\rangle \;\equiv\; \frac{(\hat{a}_1^\dagger)^{n_1}}{\sqrt{n_1!}}\,\frac{(\hat{a}_2^\dagger)^{n_2}}{\sqrt{n_2!}}\,\cdots\,|\Psi^{(0)}\rangle \;\equiv\; \prod_j \frac{(\hat{a}_j^\dagger)^{n_j}}{\sqrt{n_j!}}\,|\Psi^{(0)}\rangle} \tag{28}$$

form the orthonormal basis of Fock space that we set out to construct. This universal basis applies to bosons and fermions alike. Note, though, that the order of the creation operators matters in the fermionic case and does not matter in the bosonic case!

Properties of bosonic many-particle systems (see exercise 1):

- The basis given above corresponds to the explicit phase convention

$$\exp\bigl(i\alpha_j(\cdots, n_j, \cdots)\bigr) \;\equiv\; 1 \tag{29}$$

  in equation (12), such that

$$\boxed{\begin{aligned} \hat{a}_j|\cdots, n_j, \cdots\rangle &= \sqrt{n_j}\,|\cdots, n_j{-}1, \cdots\rangle\;, \\ \hat{a}_j^\dagger|\cdots, n_j, \cdots\rangle &= \sqrt{n_j{+}1}\,|\cdots, n_j{+}1, \cdots\rangle\;. \end{aligned}} \tag{30}$$

- For k-fold annihilation ($k = 0, 1, \cdots$) this becomes

$$(\hat{a}_j)^k|\cdots, n_j, \cdots\rangle \;=\; \sqrt{n_j(n_j{-}1)\cdots(n_j{-}k{+}1)}\,|\cdots, n_j{-}k, \cdots\rangle\;,$$

  which only leads to a sensible result if the series terminates as soon as $n_j - k < 0$. From this we can deduce straightaway that n_j can exclusively take on the values

$$\boxed{n_j \;=\; 0, 1, 2, \cdots}\;, \tag{31}$$

  as expected for the eigenvalues of a number operator.

- The order in which the creation operators occur in the basis states is irrelevant. As a result, the state functions will be totally symmetric in this version of Fock space, as expected for bosonic states. This aspect will be worked out in § 1.3.

Properties of fermionic many-particle systems (see exercise 2):

- As a result of equation (27) we know that $(\hat{a}_j^\dagger)^2 = \frac{1}{2}\{\hat{a}_j^\dagger, \hat{a}_j^\dagger\} = 0$. This we recognize as the anticipated Pauli exclusion principle: no two fermions can occupy the same fully specified 1-particle state.

- The occupation numbers can take on only two values,

$$\boxed{n_j \;=\; 0,1}\;. \tag{32}$$

  This follows from the fact that the number operators are projection operators in this case:

$$\hat{n}_j^2 \;\overset{(17)}{=\!=}\; \hat{a}_j^\dagger \hat{a}_j \hat{a}_j^\dagger \hat{a}_j \;\overset{(27)}{=\!=}\; \hat{a}_j^\dagger (\hat{1} - \hat{a}_j^\dagger \hat{a}_j) \hat{a}_j \;\overset{\hat{a}_j^2 = 0}{=\!=}\; \hat{a}_j^\dagger \hat{a}_j \;=\; \hat{n}_j \;,$$

  concurring with Pauli's exclusion principle.

- The basis given above corresponds to the explicit phase convention

$$\exp\bigl(i\alpha_j(\cdots, n_j, \cdots)\bigr) \;\equiv\; (-1)^{N_{<j}} \quad , \quad N_{<j} \;\equiv\; \sum_{k=1}^{j-1} n_k \tag{33}$$

  in equation (12), such that

$$\boxed{\begin{aligned} \hat{a}_j |\cdots, n_j, \cdots\rangle \;&=\; \delta_{n_j,1}\,(-1)^{N_{<j}}\,|\cdots, n_j{-}1, \cdots\rangle\;, \\ \hat{a}_j^\dagger |\cdots, n_j, \cdots\rangle \;&=\; \delta_{n_j,0}\,(-1)^{N_{<j}}\,|\cdots, n_j{+}1, \cdots\rangle\;. \end{aligned}} \tag{34}$$

- This time the order in which the creation operators occur in the basis states does matter and the state functions will be totally antisymmetric in this version of Fock space, as expected for fermionic states. This latter aspect will be worked out in § 1.3

The state functions of arbitrary identical-particle systems can now be expressed in terms of the basis states (28). Usually these state functions will describe a system with a fixed number of particles. However, in particular in relativistic quantum mechanical situations this may not be true, for instance when particles decay (such as $n \to p\, e^- \bar{\nu}_e$) or when a photon is emitted during de-excitation of an excited atom (see Ch. 4). Maser and laser systems fall into this category as well, as in that case the radiation is built coherently (i.e. in phase). An example of such a coherent state will be worked out in § 1.6. Finally, as artefacts of certain approximation methods situations may arise in which the number of particles is effectively not conserved anymore (see § 1.6 and § 1.7).

1.3 Switching to a continuous 1-particle representation

Since Fock space does not depend on the chosen 1-particle representation, we could also have opted for a complete set of commuting 1-particle observables $\hat{k}$ with a continuous eigenvalue spectrum $\{k\}$. Examples are the position and momentum operators. The corresponding "orthonormal" 1-particle basis is given by $\{|k\rangle\}$. In this case it makes no sense to use an occupation-number representation based on counting, unless the continuous k-space is partitioned in small cells.

In analogy with the previous discussion we will introduce creation and annihilation operators

$$\hat{a}_k^\dagger \;\equiv\; \hat{a}^\dagger(k) \qquad \text{and} \qquad \hat{a}_k \;\equiv\; \hat{a}(k)\;, \tag{35}$$

with the usual properties

$$\boxed{\hat{a}^\dagger(k)|\Psi^{(0)}\rangle \;\equiv\; |k\rangle \qquad \text{and} \qquad \hat{a}(k)|\Psi^{(0)}\rangle \;=\; 0}\;. \tag{36}$$

The properties of these creation and annihilation operators follow directly from the disrete case by performing a 1-particle basis transformation

$$\begin{aligned}
\hat{a}_j^\dagger \;&=\; \int \mathrm{d}k\; \hat{a}^\dagger(k)\langle k|q_j\rangle \qquad \xrightarrow{\text{herm. conj.}} \qquad \hat{a}_j \;=\; \int \mathrm{d}k\; \hat{a}(k)\langle q_j|k\rangle\;, \\
\hat{a}^\dagger(k) \;&=\; \sum_j \hat{a}_j^\dagger \langle q_j|k\rangle \qquad \xrightarrow{\text{herm. conj.}} \qquad \hat{a}(k) \;=\; \sum_j \hat{a}_j \langle k|q_j\rangle\;,
\end{aligned} \tag{37}$$

which can be derived in analogy to equation (21) from

$$\hat{a}_j^\dagger|\Psi^{(0)}\rangle \;=\; |q_j\rangle \;\overset{\text{compl.}}{=\!=}\; \int \mathrm{d}k\; |k\rangle\langle k|q_j\rangle \;\overset{(36)}{=\!=}\; \int \mathrm{d}k\; \langle k|q_j\rangle\, \hat{a}^\dagger(k)|\Psi^{(0)}\rangle\;.$$

For this basis transformation the following unitarity conditions hold:

$$\begin{aligned}
\sum_j \langle k|q_j\rangle\langle q_j|k'\rangle \;&\overset{\text{compl.}}{=\!=}\; \langle k|k'\rangle \;=\; \delta(k-k')\;, \\
\int \mathrm{d}k\; \langle q_j|k\rangle\langle k|q_{j'}\rangle \;&\overset{\text{compl.}}{=\!=}\; \langle q_j|q_{j'}\rangle \;=\; \delta_{jj'}\;.
\end{aligned} \tag{38}$$

Subsequently equations (26) and (27) for a discrete 1-particle representation can be used to obtain

$$\boxed{\left[\hat{a}^\dagger(k),\hat{a}^\dagger(k')\right] \;=\; \left[\hat{a}(k),\hat{a}(k')\right] \;=\; 0 \quad , \quad \left[\hat{a}(k),\hat{a}^\dagger(k')\right] \;=\; \delta(k-k')\hat{1}} \tag{39}$$

for a bosonic system in a continuous 1-particle representation and

$$\boxed{\left\{\hat{a}^\dagger(k),\hat{a}^\dagger(k')\right\} \;=\; \left\{\hat{a}(k),\hat{a}(k')\right\} \;=\; 0 \quad , \quad \left\{\hat{a}(k),\hat{a}^\dagger(k')\right\} \;=\; \delta(k-k')\hat{1}} \tag{40}$$

for a fermionic system in a continuous 1-particle representation.

Proof: the derivation of the vanishing (anti)commutators is trivial. In addition

$$\begin{aligned}
\hat{a}(k)\hat{a}^\dagger(k') \mp \hat{a}^\dagger(k')\hat{a}(k) \;&\overset{(37)}{=\!=}\; \sum_{j,j'} \langle k|q_j\rangle\langle q_{j'}|k'\rangle \left(\hat{a}_j\hat{a}_{j'}^\dagger \mp \hat{a}_{j'}^\dagger\hat{a}_j\right) \;\overset{(26),(27)}{=\!=}\; \sum_j \langle k|q_j\rangle\langle q_j|k'\rangle\hat{1} \\
&\overset{(38)}{=\!=}\; \delta(k-k')\hat{1}\;,
\end{aligned}$$

where the minus (plus) sign refers to bosonic (fermionic) systems.

The representation-independent total number operator $\hat{N}$ can in this case be expressed in terms of the number density operator

$$\boxed{\hat{n}(k) \equiv \hat{a}^\dagger(k)\hat{a}(k)} \tag{41}$$

according to

$$\hat{N} \overset{(22)}{=\!=} \sum_j \hat{n}_j \overset{(17)}{=\!=} \sum_j \hat{a}_j^\dagger \hat{a}_j \overset{(37)}{=\!=} \int \mathrm{d}k \int \mathrm{d}k' \; \hat{a}^\dagger(k)\hat{a}(k') \sum_j \langle k|q_j\rangle\langle q_j|k'\rangle$$

$$\overset{(38)}{=\!=} \int \mathrm{d}k \int \mathrm{d}k' \; \hat{a}^\dagger(k)\hat{a}(k')\delta(k-k') \overset{(41)}{=\!=} \int \mathrm{d}k \; \hat{n}(k) \; . \tag{42}$$

The basis of N-particle state functions takes the following form in the k-representation:

$$\boxed{\left\{ |k_1, \cdots, k_N\rangle \equiv \frac{1}{\sqrt{N!}} \hat{a}^\dagger(k_1) \cdots \hat{a}^\dagger(k_N)|\Psi^{(0)}\rangle : \quad k_1, \cdots, k_N \in \{k\} \right\}} . \tag{43}$$

This is a direct consequence of the fact that

$$\int \mathrm{d}k_1 \cdots \int \mathrm{d}k_N \; |k_1, \cdots, k_N\rangle\langle k_1, \cdots, k_N| = \hat{1}_N \tag{44}$$

is the unit operator in the N-particle subspace, which will be proven in exercise 3. An arbitrary N-particle state function $|\Psi\rangle$ can now be represented in the k-representation by the function

$$\psi(k_1, \cdots, k_N) \equiv \langle k_1, \cdots, k_N|\Psi\rangle \; . \tag{45}$$

As expected, this function is totally symmetric for bosonic systems and totally antisymmetric for fermionic systems. This too will be proven in exercise 3.

1.3.1 Position and momentum representation

The previous discussion can be generalized trivially to cases with mixed 1-particle eigenvalue spectra, which are partially discrete and partially continuous. Two examples of this will be worked out now.

The position representation: as a first example of a mixed representation we consider spin-s particles in the position representation. This representation will be used in chapter 5 when setting up relativistic QM. In the previous expressions we now simply have to substitute

$$\hat{k} \to (\hat{\vec{r}}, \hat{S}_z/\hbar) \quad , \quad \{k\} \to \{\vec{r} \in \mathbb{R}^3 \quad , \quad \sigma = m_s = -s, -s+1, \cdots, s-1, s\} \; ,$$

$$\int \mathrm{d}k \to \sum_\sigma \int \mathrm{d}\vec{r} \quad , \quad \delta(k-k') \to \delta_{\sigma\sigma'}\delta(\vec{r}-\vec{r}\,') \qquad \text{and} \qquad \hat{a}(k) \to \hat{\psi}_\sigma(\vec{r}) \; . \tag{46}$$

Quite suggestively the annihilation oparator $\hat{a}(k)$ is replaced in the position representation by the so-called field operator $\hat{\psi}_\sigma(\vec{r})$. This suggestive notation is introduced in order to facilitate the discussion in § 1.5.

The momentum representation: as a second example of a mixed representation we consider spin-s particles in the momentum representation. This representation is particularly useful for systems that (in first approximation) consist of free particles, as will be used in chapter 5 for analysing relativistic quantum mechanical wave equations. In this case we have to perform the following substitutions:

$$\hat{k} \;\to\; (\hat{\vec{p}}, \hat{S}_z/\hbar) \quad , \quad \{k\} \;\to\; \{\vec{p} \in \mathbb{R}^3 \quad , \quad \sigma = m_s = -s, -s+1, \cdots, s-1, s\} \;,$$

$$\int \mathrm{d}k \;\to\; \sum_\sigma \int \mathrm{d}\vec{p} \quad , \quad \delta(k-k') \;\to\; \delta_{\sigma\sigma'}\,\delta(\vec{p}-\vec{p}\,') \qquad \text{and} \qquad \hat{a}(k) \;\to\; \hat{a}_\sigma(\vec{p}) \;. \tag{47}$$

These two representations are connected through a Fourier transformation:

$$\hat{\psi}_\sigma(\vec{r}) \overset{(37)}{=\!=} \int \mathrm{d}\vec{p}\, \langle \vec{r}\,|\vec{p}\,\rangle\, \hat{a}_\sigma(\vec{p}) \;=\; \int \mathrm{d}\vec{p}\; \frac{\exp(i\vec{p}\cdot\vec{r}/\hbar)}{(2\pi\hbar)^{3/2}}\; \hat{a}_\sigma(\vec{p}) \;,$$

$$\hat{a}_\sigma(\vec{p}) \overset{(37)}{=\!=} \int \mathrm{d}\vec{r}\, \langle \vec{p}\,|\vec{r}\,\rangle\, \hat{\psi}_\sigma(\vec{r}) \;=\; \int \mathrm{d}\vec{r}\; \frac{\exp(-i\vec{p}\cdot\vec{r}/\hbar)}{(2\pi\hbar)^{3/2}}\; \hat{\psi}_\sigma(\vec{r}) \;. \tag{48}$$

In fact this is identical to the connection between the state functions in position and momentum space known from 1-particle QM (see the lecture course Quantum Mechanics 2).

1.4 Additive many-particle quantities and particle conservation

> *Observables: we have succeeded in constructing the state space applicable to systems consisting of an arbitrary number of identical particles and we have investigated what the basis states look like if we employ discrete and/or continuous 1-particle representations. As a next step in the construction of the many-particle toolbox we need to have a closer look at the many-particle observables.*

In Fock space various kinds of many-particle observables are possible, as long as they are duly symmetric under particle interchange. An important class of many-particle observables consists of observables that belong to so-called additive many-particle quantities. These quantities can be categorized by the number of particles that participate as a cluster in the physical quantity. Subsequently these clusters will be summed/integrated over. A characteristic property of these additive many-particle quantities is the existence of a specific representation that makes the corresponding many-particle observable expressible solely in terms of number (density) operators. This automatically implies that different eigenstates of the total number operator $\hat{N}$ are not mixed by such a many-particle observable (which entails particle conservation).

Additive many-particle quantities play a central role in setting up (free-particle) relativistic QM and in the description of systems consisting of weakly interacting identical particles (such as electrons in an atom or conduction electrons in a metal). In lowest-order approximation the particles of a weakly interacting system can be considered non-interacting, which causes the total energy to be expressible as a sum of individual 1-particle energies (see § 1.1). As a first-order correction to this, the interactions between pairs of particles can be added. In this way the particles contribute in sets of two to the energy correction. This procedure can be extended trivially to allow for (progressively smaller) energy corrections involving larger clusters of particles.

1.4.1 Additive 1-particle quantities

Consider a 1-particle observable $\hat{A}$. By employing the completeness relation for the discrete q-representation, this observable can de written as

$$\hat{A} \;=\; \sum_{j,j'} |q_j\rangle \langle q_j|\hat{A}|q_{j'}\rangle \langle q_{j'}| \;.$$

Each individual term occurring in this sum brings the particle from a state $|q_{j'}\rangle$ to a state $|q_j\rangle$, with the matrix element $\langle q_j|\hat{A}|q_{j'}\rangle$ as corresponding weight factor. Using this 1-particle observable a proper many-particle observable can be constructed:

$$\hat{A}^{(1)}_{\rm tot} \;=\; \sum_{\alpha} \hat{A}_\alpha \;,$$

where $\hat{A}_\alpha$ is the 1-particle observable belonging to particle α. The action of $\hat{A}^{(1)}_{\rm tot}$ in Fock space is to simply annihilate a particle in the state $|q_{j'}\rangle$ and subsequently create a particle in the state $|q_j\rangle$ with weight factor $\langle q_j|\hat{A}|q_{j'}\rangle$:

$$\boxed{\hat{A}^{(1)}_{\rm tot} \;=\; \sum_{j,j'} \hat{a}^\dagger_j \langle q_j|\hat{A}|q_{j'}\rangle \hat{a}_{j'} \;=\; \sum_{j,j'} \langle q_j|\hat{A}|q_{j'}\rangle \hat{a}^\dagger_j \hat{a}_{j'}} \;. \tag{49}$$

We speak of an additive 1-particle quantity if the corresponding many-particle observable can be represented in the form (49). This definition hinges on the fact that the given expression has exactly the same form for any discrete 1-particle representation:

$$\hat{A}^{(1)}_{\rm tot} \;\overset{(18),(21)}{=\!=}\; \sum_{r,r',j,j'} \hat{b}^\dagger_r \langle p_r|q_j\rangle \langle q_j|\hat{A}|q_{j'}\rangle \langle q_{j'}|p_{r'}\rangle \hat{b}_{r'} \;\overset{\text{compl.}}{=\!=}\; \sum_{r,r'} \hat{b}^\dagger_r \langle p_r|\hat{A}|p_{r'}\rangle \hat{b}_{r'} \;.$$

Many-particle observables belonging to additive 1-particle quantities can also be translated trivially to a continuous 1-particle representation. By means of the basis transformation (37) and the unitarity conditions (38), equation (49) changes into

$$\boxed{\hat{A}^{(1)}_{\rm tot} \;=\; \int \mathrm{d}k_1 \int \mathrm{d}k_2 \; \hat{a}^\dagger(k_1) \langle k_1|\hat{A}|k_2\rangle \hat{a}(k_2) \;=\; \int \mathrm{d}k_1 \int \mathrm{d}k_2 \; \langle k_1|\hat{A}|k_2\rangle \hat{a}^\dagger(k_1) \hat{a}(k_2)} \;. \tag{50}$$

The general expressions (49) and (50) are valid irrespective of the 1-particle observable $\hat{A}$ having discrete or continuous eigenvalues.

Special cases: let's assume that the 1-particle observable $\hat{A}$ has discrete eigenvalues and is part of the complete set of commuting 1-particle observables $\hat{q}$. In that case

$$\langle q_j|\hat{A}|q_{j'}\rangle \;\equiv\; A_{j'}\,\langle q_j|q_{j'}\rangle \;=\; A_j\,\delta_{jj'} \quad\stackrel{(49)}{\Longrightarrow}\quad \hat{A}^{(1)}_{\rm tot} \;=\; \sum_j A_j\,\hat{n}_j \;.$$

The special feature of this representation is the fact that it merely involves counting particles, in analogy to the total number operator $\hat{N}$. We actually already exploited this concept in § 1.1 while determining the total energy of a non-interacting many-particle system. Apart from additive 1-particle quantities with discrete eigenvalues, also additive 1-particle quantities exist with continuous eigenvalues (see below). In that case a continuous representation can be found such that

$$\langle k_1|\hat{A}|k_2\rangle \;\equiv\; A(k_2)\,\langle k_1|k_2\rangle \;=\; A(k_1)\,\delta(k_1-k_2) \quad\stackrel{(50)}{\Longrightarrow}\quad \hat{A}^{(1)}_{\rm tot} \;=\; \int\!{\rm d}k_1\,A(k_1)\,\hat{n}(k_1) \;.$$

Particle conservation: a characteristic feature of additive many-particle quantities is that the observables only consist of terms with as many creation as annihilation operators. As such these observables commute with the total number operator $\hat{N}$ and therefore do not mix the different eigenstates of $\hat{N}$. In that respect one could speak of "conservation of the number of particles".

Additive 1-particle quantities in position and momentum representation.

In view of future applications, we now take a closer look at a few useful additive 1-particle quantities in the mixed representations belonging to spin-s particles in either position or momentum space.

The total momentum: in the momentum representation the 1-particle momentum operator $\hat{\vec{p}}$ is part of the complete set of observables with respect to which the many-particle Fock space is formulated. As a result, the total momentum operator takes on a particularly simple form in the momentum representation (i.e. merely involving number density operators in momentum space):

$$\hat{\vec{P}}^{(1)}_{\rm tot} \;=\; \sum_\sigma \int\!{\rm d}\vec{p}\,\vec{p}\,\hat{n}_\sigma(\vec{p}) \;=\; \sum_\sigma \int\!{\rm d}\vec{r}\,\hat{\psi}^\dagger_\sigma(\vec{r})\Big(-i\hbar\vec{\nabla}\hat{\psi}_\sigma(\vec{r})\Big) \;. \tag{51}$$

The proof of the last step in this equation involves the Fourier transformation (48) for switching from momentum to position representation (see also App. A for additional details):

$$-i\hbar\int \mathrm{d}\vec{r}\,\hat{\psi}^\dagger_\sigma(\vec{r})\vec{\nabla}\hat{\psi}_\sigma(\vec{r}) \;=\; -i\hbar\int\mathrm{d}\vec{p}\int\mathrm{d}\vec{p}\,'\,\hat{a}^\dagger_\sigma(\vec{p})\,\hat{a}_\sigma(\vec{p}\,')\int\mathrm{d}\vec{r}\;\frac{\exp(-i\vec{p}\cdot\vec{r}/\hbar)}{(2\pi\hbar)^3}\;\vec{\nabla}\exp(i\vec{p}\,'\cdot\vec{r}/\hbar)$$

$$= \int\mathrm{d}\vec{p}\int\mathrm{d}\vec{p}\,'\,\vec{p}\,'\,\hat{a}^\dagger_\sigma(\vec{p})\,\hat{a}_\sigma(\vec{p}\,')\int\mathrm{d}\vec{r}\;\frac{\exp(i\,[\vec{p}\,'-\vec{p}\,]\cdot\vec{r}/\hbar)}{(2\pi\hbar)^3}$$

$$= \int\mathrm{d}\vec{p}\int\mathrm{d}\vec{p}\,'\,\vec{p}\,'\,\hat{a}^\dagger_\sigma(\vec{p})\,\hat{a}_\sigma(\vec{p}\,')\,\delta(\vec{p}-\vec{p}\,') \;=\; \int\mathrm{d}\vec{p}\,\vec{p}\,\hat{n}_\sigma(\vec{p})\;.$$

The total kinetic energy: the corresponding 1-particle observable $\hat{T} = \hat{\vec{p}}^{\,2}/(2m)$ is again diagonal in the momentum representation. In analogy to the previous example we find for the total kinetic-energy operator:

$$\hat{T}^{(1)}_{\mathrm{tot}} \;=\; \sum_\sigma\int\mathrm{d}\vec{p}\;\frac{\vec{p}^{\,2}}{2m}\;\hat{n}_\sigma(\vec{p}) \;=\; \sum_\sigma\int\mathrm{d}\vec{r}\;\hat{\psi}^\dagger_\sigma(\vec{r})\Big(-\frac{\hbar^2}{2m}\;\vec{\nabla}^{\,2}\hat{\psi}_\sigma(\vec{r})\Big)\;. \tag{52}$$

The total spin in the z-direction: in this case the corresponding 1-particle observable $\hat{S}_z$ is part of the complete set of observables of both the position and momentum representation. The total spin operator in the z-direction trivially reads:

$$\hat{\vec{S}}^{(1)}_{\mathrm{tot}}\cdot\vec{e}_z \;=\; \sum_\sigma\int\mathrm{d}\vec{p}\;\hbar\sigma\,\hat{n}_\sigma(\vec{p}) \;=\; \sum_\sigma\int\mathrm{d}\vec{r}\;\hbar\sigma\,\hat{n}_\sigma(\vec{r}) \;\equiv\; \sum_\sigma\hbar\sigma\,\hat{N}_\sigma\;, \tag{53}$$

where $\hat{N}_\sigma$ counts the total number of particles with spin component $\hbar\sigma$ along the z-axis.

The total potential energy in an external field: assume the particles to be subjected to an external potential $V(\vec{r})$. The corresponding 1-particle observable $V(\hat{\vec{r}})$ is obviously diagonal in the position representation, resulting in the following total potential-energy operator:

$$\hat{V}^{(1)}_{\mathrm{tot}} \;\equiv\; \sum_\sigma\int\mathrm{d}\vec{r}\;V(\vec{r})\,\hat{n}_\sigma(\vec{r}) \;=\; \sum_\sigma\int\mathrm{d}\vec{k}\int\mathrm{d}\vec{p}\;\;\mathcal{V}(\vec{k})\,\hat{a}^\dagger_\sigma(\vec{p})\,\hat{a}_\sigma(\vec{p}-\hbar\vec{k})\;, \tag{54}$$

using the definition $V(\vec{r}) \equiv \int\mathrm{d}\vec{k}\;\mathcal{V}(\vec{k})\exp(i\vec{k}\cdot\vec{r})$. The derivation of the right-hand side of this equation follows the line of proof worked out in exercise 4 for a related type of potential.

Second quantization: the additive many-particle observables discussed above resemble quantized versions of the corresponding expectation values of the 1-particle observables in ordinary 1-particle QM, i.e. expectation values with $\psi \to \hat{\psi}$. This formal connection between 1-particle and many-particle quantum theory is indicated somewhat suggestively by the term "second quantization".

1.4.2 Additive 2-particle quantities

In the discrete q-representation the ordered 2-particle basis states are given by

$$|\Psi^{(2)}_{j,k>j}\rangle \;\equiv\; |0,\cdots,0,n_j=1,0,\cdots,0,n_k=1,0,\cdots\rangle \;=\; \hat{a}^\dagger_j\hat{a}^\dagger_k\,|\Psi^{(0)}\rangle\;,$$

$$|\Psi^{(2)}_{jj}\rangle \;\equiv\; |0,\cdots,0,n_j=2,0,\cdots\rangle \;=\; \frac{(\hat{a}^\dagger_j)^2}{\sqrt{2}}\,|\Psi^{(0)}\rangle\;, \tag{55}$$

where the last pair state only has relevance for bosonic systems. With the help of the corresponding completeness relation, a 2-particle observable $\hat{B}$ can be written as

$$\hat{B} \;=\sum_{j,j',k,k'} |\Psi^{(2)}_{jk}\rangle\,\langle\Psi^{(2)}_{jk}|\hat{B}|\Psi^{(2)}_{j'k'}\rangle\,\langle\Psi^{(2)}_{j'k'}|\;.$$

Each individual term occurring in this sum brings the particle pair from a state $|\Psi^{(2)}_{j'k'}\rangle$ to a state $|\Psi^{(2)}_{jk}\rangle$, with the matrix element $\langle\Psi^{(2)}_{jk}|\hat{B}|\Psi^{(2)}_{j'k'}\rangle$ as corresponding weight factor. Using this 2-particle observable a proper many-particle observable can be constructed:

$$\hat{B}^{(2)}_{\rm tot} \;=\; \frac{1}{2}\sum_{\alpha\,,\beta\neq\alpha}\hat{B}_{\alpha\beta}\;,$$

where $\hat{B}_{\alpha\beta}$ is the 2-particle observable belonging to particles α and $\beta\neq\alpha$. The factor of $1/2$ is introduced here to avoid double counting. The corresponding action of this observable in Fock space simply reads

$$\boxed{\hat{B}^{(2)}_{\rm tot} \;=\; \frac{1}{2}\sum_{j,j',k,k'}\hat{a}^\dagger_j\hat{a}^\dagger_k\langle\Psi^{(2)}_{jk}|\hat{B}|\Psi^{(2)}_{j'k'}\rangle\,\hat{a}_{k'}\hat{a}_{j'} \;=\; \frac{1}{2}\sum_{j,j',k,k'}\langle\Psi^{(2)}_{jk}|\hat{B}|\Psi^{(2)}_{j'k'}\rangle\,\hat{a}^\dagger_j\hat{a}^\dagger_k\hat{a}_{k'}\hat{a}_{j'}}\;. \tag{56}$$

This method of writing all creation operators on the left and all annihilation operators on the right in a many-particle operator is usually referred to as normal ordering. We speak of an additive 2-particle quantity if the corresponding many-particle observable can be represented in the form (56). This type of expression has the same form for any discrete 1-particle representation, bearing in mind that just like we have seen in § 1.4.1 each creation/annihilation operator in the expression is linked to a corresponding annihilation/creation operator that is hidden in one of the basis states in the matrix element. Switching to a continuous representation yields accordingly

$$\boxed{\hat{B}^{(2)}_{\rm tot} \;=\; \frac{1}{2}\int{\rm d}k_1\,\cdots\int{\rm d}k_4\,\langle k_1,k_2|\hat{B}|k_3,k_4\rangle\,\hat{a}^\dagger(k_1)\hat{a}^\dagger(k_2)\hat{a}(k_4)\hat{a}(k_3)}\;. \tag{57}$$

Again the generic expressions (56) and (57) are valid irrespective of the 2-particle observabele $\hat{B}$ having discrete or continuous eigenvalues.

Special cases: let's assume that the 2-particle observable $\hat{B}$ has discrete eigenvalues and is diagonal in the q-representation. Then

$$\langle \Psi^{(2)}_{jk} | \hat{B} | \Psi^{(2)}_{j'k'} \rangle \;\equiv\; B_{j'k'} \langle \Psi^{(2)}_{jk} | \Psi^{(2)}_{j'k'} \rangle \;=\; B_{jk}\, \delta_{jj'}\, \delta_{kk'}$$

$$\stackrel{(56)}{\Longrightarrow} \quad \hat{B}^{(2)}_{\rm tot} \;=\; \frac{1}{2} \sum_{j,k} B_{jk}\, \hat{a}^\dagger_j\, \hat{a}^\dagger_k\, \hat{a}_k\, \hat{a}_j \;=\; \frac{1}{2} \sum_{j,k} B_{jk} (\hat{n}_j\, \hat{n}_k - \hat{n}_j\, \delta_{jk}) \;,$$

where we have used that

$$\hat{a}^\dagger_j\, \hat{a}^\dagger_k\, \hat{a}_k\, \hat{a}_j \;\stackrel{(16),(17)}{=\!=}\; \hat{a}^\dagger_j\, \hat{a}_j\, \hat{a}^\dagger_k\, \hat{a}_k - \hat{a}^\dagger_j\, \delta_{jk}\, \hat{a}_j \;\stackrel{(17)}{=\!=}\; \hat{n}_j\, \hat{n}_k - \hat{n}_j\, \delta_{jk} \;=\; \text{pair number operator} \,. \quad (58)$$

Of course, additive 2-particle quantites may also have continuous eigenvalues (see below). In that case a continuous representation exists such that $\hat{B}$ becomes diagonal, allowing to express $\hat{B}^{(2)}_{\rm tot}$ in terms of pair density operators belonging to that representation.

Example: spatial pair interactions in position and momentum representation.

Consider a system comprising of spin-s particles with spatial pair interactions described by the observable $U(\hat{\vec{r}}_1 - \hat{\vec{r}}_2)$. This observable is obviously diagonal in the position representation, resulting in the following total operator for spatial pair interactions:

$$\hat{U}^{(2)}_{\rm tot} \;\equiv\; \frac{1}{2} \sum_{\sigma_1,\sigma_2} \int \mathrm{d}\vec{r}_1 \int \mathrm{d}\vec{r}_2\; U(\vec{r}_1 - \vec{r}_2)\, \hat{\psi}^\dagger_{\sigma_1}(\vec{r}_1)\, \hat{\psi}^\dagger_{\sigma_2}(\vec{r}_2)\, \hat{\psi}_{\sigma_2}(\vec{r}_2)\, \hat{\psi}_{\sigma_1}(\vec{r}_1) \;. \quad (59)$$

In the momentum representation this becomes (see exercise 4)

$$\hat{U}^{(2)}_{\rm tot} \;=\; \frac{1}{2} \sum_{\sigma_1,\sigma_2} \int \mathrm{d}\vec{k} \int \mathrm{d}\vec{p}_1 \int \mathrm{d}\vec{p}_2\; \mathcal{U}(\vec{k})\, \hat{a}^\dagger_{\sigma_1}(\vec{p}_1)\, \hat{a}^\dagger_{\sigma_2}(\vec{p}_2)\, \hat{a}_{\sigma_2}(\vec{p}_2 + \hbar\vec{k})\, \hat{a}_{\sigma_1}(\vec{p}_1 - \hbar\vec{k}) \;, \quad (60)$$

using the definition $U(\vec{r}) \equiv \int \mathrm{d}\vec{k}\, \mathcal{U}(\vec{k}) \exp(i\vec{k}\cdot\vec{r})$.

1.5 Heisenberg picture and second quantization

In preparation for the relativistic treatment of quantum mechanics in chapter 5, we now address the time dependence of the creation and annihilation operators in the Heisenberg picture of quantum mechanics.

Consider to this end an isolated identical-particle system described by a Hamilton operator of the following form:

$$\hat{H}_{\rm tot} \;=\; \hat{H}^{(1)}_{\rm tot} + C\hat{1} \qquad (C \in \mathbb{R}) \;, \quad (61)$$

where the term $\hat{H}^{(1)}_{\rm tot}$ is an additive observable of the type described in § 1.4.1. Next we address the time evolution of this many-particle system in the Heisenberg picture. As explained in the lecture course Quantum Mechanics 2, all state functions are time-independent in the Heisenberg picture and all time dependence of the system resides in

the time evolution of the operators. The differences with the many-particle formulation in the Schrödinger picture, which we have used so far, can be summarized as follows:

$$\begin{aligned}
&\underline{\text{Schrödinger picture}} \;:\; \text{state functions} \;\rightarrow\; |\Psi(t)\rangle \;=\; \exp\bigl(-i\hat{H}_{\text{tot}}\,t/\hbar\bigr)\,|\Psi(0)\rangle\;, \\
&\qquad\qquad\qquad\quad\;\; \text{observables} \;\rightarrow\; \hat{A}(t)\;, \\
&\underline{\text{Heisenberg picture}} \;:\; \text{state functions} \;\rightarrow\; |\Psi_{_H}(t)\rangle \;=\; |\Psi(0)\rangle\;, \\
&\qquad\qquad\qquad\quad\;\; \text{observables} \;\rightarrow\; \hat{A}_{_H}(t) \;=\; \exp\bigl(i\hat{H}_{\text{tot}}\,t/\hbar\bigr)\,\hat{A}(t)\,\exp\bigl(-i\hat{H}_{\text{tot}}\,t/\hbar\bigr)\;,
\end{aligned} \tag{62}$$

where we have assumed that both pictures coincide at the time stamp $t = 0$. For example we have

$$\hat{H}_{\text{tot}_{_H}}(t) \;=\; \hat{H}_{\text{tot}}\;, \tag{63}$$

since we are dealing here with an isolated system. In the Heisenberg picture the operators satisfy the so-called Heisenberg equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{A}_{_H}(t) \;=\; -\,\frac{i}{\hbar}\,\bigl[\hat{A}_{_H}(t),\hat{H}_{\text{tot}}\bigr] \;+\; \Bigl(\,\frac{\partial\hat{A}(t)}{\partial t}\,\Bigr)_{_H}\;. \tag{64}$$

Next we investigate the implications for the creation and annihilation operators.

The discrete case: for a generic discrete 1-particle representation we get

$$\hat{H}^{(1)}_{\text{tot}} \;\overset{(49)}{=\!=}\; \sum_{r,r'}\langle p_r|\hat{H}\,|p_{r'}\rangle\,\hat{b}^{\dagger}_{r}\hat{b}_{r'} \;\overset{(63)}{=\!=}\; \sum_{r,r'}\langle p_r|\hat{H}\,|p_{r'}\rangle\,\hat{b}^{\dagger}_{r_{_H}}(t)\,\hat{b}_{r'_{_H}}(t)\;,$$

with $\hat{H}$ being the 1-particle Hamilton-operator belonging to the additive part of the many-particle operator $\hat{H}_{\text{tot}}$. From this it follows that

$$\begin{aligned}
\bigl[\hat{b}_{v_{_H}}(t),\hat{H}_{\text{tot}}\bigr] \;&=\; \sum_{r,r'}\langle p_r|\hat{H}\,|p_{r'}\rangle\,\bigl[\hat{b}_{v_{_H}}(t),\hat{b}^{\dagger}_{r_{_H}}(t)\,\hat{b}_{r'_{_H}}(t)\bigr] \;\overset{(26),(27)}{=\!=}\; \sum_{r,r'}\langle p_r|\hat{H}\,|p_{r'}\rangle\,\hat{b}_{r'_{_H}}(t)\,\delta_{vr} \\
&=\; \sum_{r'}\langle p_v|\hat{H}\,|p_{r'}\rangle\,\hat{b}_{r'_{_H}}(t)\;.
\end{aligned}$$

The Heisenberg equation of motion then reads

$$\boxed{\,i\hbar\,\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{b}_{v_{_H}}(t) \;=\; \sum_{r}\langle p_v|\hat{H}\,|p_r\rangle\,\hat{b}_{r_{_H}}(t)\,}\;, \tag{65}$$

where we have used that the creation and annihilation operators are time-independent in the Schrödinger picture. This equation of motion strongly resembles the Schrödinger equation for a 1-particle state function written in the p-representation:

$$i\hbar\,\frac{\mathrm{d}}{\mathrm{d}t}\,|\psi(t)\rangle \;=\; \hat{H}\,|\psi(t)\rangle \quad\overset{\langle p_v|*}{\Longrightarrow}\quad i\hbar\,\frac{\mathrm{d}}{\mathrm{d}t}\,\langle p_v|\psi(t)\rangle \;\overset{\text{compl.}}{=\!=}\; \sum_{r}\langle p_v|\hat{H}\,|p_r\rangle\,\langle p_r|\psi(t)\rangle\;.$$

The coupled linear differential equations (65) can be decoupled if we choose a representation for which $\langle p_v|\hat{H}\,|p_r\rangle = E_r\,\langle p_v|p_r\rangle = E_v\,\delta_{rv}$, such that

$$i\hbar\,\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{b}_{v_{_H}}(t) \;=\; E_v\,\hat{b}_{v_{_H}}(t) \quad\Rightarrow\quad \hat{b}_{v_{_H}}(t) \;=\; \exp(-iE_vt/\hbar)\,\hat{b}_{v_{_H}}(0) \;\equiv\; \exp(-iE_vt/\hbar)\,\hat{b}_v\;.$$

Free particles in position space: in the Schrödinger picture the position-space annihilation operator is given by the time-independent field operator $\hat{\psi}_\sigma(\vec{r}\,)$. The corresponding time-dependent operator in the Heisenberg picture is concisely written as

$$\boxed{\big(\hat{\psi}_\sigma(\vec{r}\,)\big)_{_H}(t) \;\equiv\; \hat{\psi}_\sigma(\vec{r},t)}\,, \tag{66}$$

which can be translated to momentum space by means of the Fourier transformation

$$\hat{\psi}_\sigma(\vec{r},t) \;=\int\mathrm{d}\vec{p}\;\frac{\exp(i\vec{p}\cdot\vec{r}/\hbar)}{(2\pi\hbar)^{3/2}}\;\hat{a}_{\sigma_{_H}}(\vec{p},t)\;. \tag{67}$$

For free particles the Hamilton operator is given by the total kinetic-energy operator $\hat{T}^{(1)}_{\rm tot}$:

$$\hat{H}_{\rm tot} \;\overset{(52)}{=\!=}\; \sum_{\sigma'}\int\mathrm{d}\vec{p}\,'\;\frac{\vec{p}\,'^{\,2}}{2m}\;\hat{n}_{\sigma'}(\vec{p}\,') \;\overset{(63)}{=\!=}\; \sum_{\sigma'}\int\mathrm{d}\vec{p}\,'\;\frac{\vec{p}\,'^{\,2}}{2m}\;\hat{n}_{\sigma'_{_H}}(\vec{p}\,',t)\;. \tag{68}$$

In the momentum representation the Heisenberg equation of motion reads

$$i\hbar\,\frac{\partial}{\partial t}\,\hat{a}_{\sigma_{_H}}(\vec{p},t) \;\overset{(64),(68)}{=\!=\!=}\; \sum_{\sigma'}\int\mathrm{d}\vec{p}\,'\;\frac{\vec{p}\,'^{\,2}}{2m}\;\big[\hat{a}_{\sigma_{_H}}(\vec{p},t)\,,\hat{n}_{\sigma'_{_H}}(\vec{p}\,',t)\big] \;\overset{(16)}{=\!=}\; \frac{\vec{p}^{\,2}}{2m}\,\hat{a}_{\sigma_{_H}}(\vec{p},t)$$

$$\Rightarrow\quad \boxed{\hat{a}_{\sigma_{_H}}(\vec{p},t) \;=\; \exp(-iE_{\vec{p}}\,t/\hbar)\,\hat{a}_\sigma(\vec{p}\,) \quad\text{with}\quad E_{\vec{p}}\;\equiv\;\frac{\vec{p}^{\,2}}{2m}}\;. \tag{69}$$

This implies the following for the field operator $\hat{\psi}_\sigma(\vec{r},t)$ in the position representation:

$$i\hbar\,\frac{\partial}{\partial t}\,\hat{\psi}_\sigma(\vec{r},t) \;\overset{(67),(69)}{=\!=\!=}\; \int\mathrm{d}\vec{p}\;\frac{\exp(i\vec{p}\cdot\vec{r}/\hbar)}{(2\pi\hbar)^{3/2}}\;\frac{\vec{p}^{\,2}}{2m}\,\hat{a}_{\sigma_{_H}}(\vec{p},t)$$

$$\overset{(67)}{\Longrightarrow}\quad \boxed{i\hbar\,\frac{\partial}{\partial t}\,\hat{\psi}_\sigma(\vec{r},t) \;=\; -\,\frac{\hbar^2}{2m}\,\vec{\bigtriangledown}^{\,2}\,\hat{\psi}_\sigma(\vec{r},t)}\;. \tag{70}$$

- In the Heisenberg picture the many-particle field operator $\hat{\psi}_\sigma(\vec{r},t)$ satisfies an equation that is identical to the free Schrödinger equation for a 1-particle wave function in position space.[1] For this reason $\hat{\psi}_\sigma(\vec{r},t)$ is also known as the Schrödinger field. Equation (70) can be regarded as a quantized version of the "classical" Schrödinger equation for quantum mechanical wave functions. Also this direct link between a 1-particle wave equation and a many-particle theory formulated in the Heisenberg picture is a manifestation of the concept of second quantization.

[1] This relation extends to situations where potential-energy terms are added to the Hamilton operator.

Second quantization will play a crucial role in the transition from non-relativitic to relativistic QM. We will be forced to abandon the idea of a 1-particle quantum mechanical theory based on a relativistic wave equation and replace it by a corresponding many-particle formulation that constitutes the starting point for quantum field theory.

- As we have seen, the above-given equation of motion (wave equation) for $\hat{\psi}_\sigma(\vec{r},t)$ can be solved (decoupled) straightforwardly by means of a Fourier decomposition (plane-wave expansion):

$$\hat{\psi}_\sigma(\vec{r},t) \overset{(67),(69)}{=\!=\!=} \int \mathrm{d}\vec{p}\, \frac{\exp(i\vec{p}\cdot\vec{r}/\hbar)}{(2\pi\hbar)^{3/2}}\, \exp(-iE_{\vec{p}}t/\hbar)\,\hat{a}_\sigma(\vec{p})\ . \tag{71}$$

  In fact this is a manifestation of particle–wave duality, as each annihilation/creation of a particle is automatically accompanied by a corresponding plane-wave factor. At a later stage we will make use of this aspect to come up with an alternative, many-particle formulation of relativistic QM.

1.6 Examples and applications: bosonic systems

In the preceding discussion we have systematically used the generic term "particles" to talk about the fundamental quantum-mechanical objects of a specific quantum system. It seems obvious what these particles should be: intuitively we expect that the particles belonging to an atomic gas system should be the atoms themselves. However, in the remainder of this chapter we will see that the particle interpretation of quantum systems in terms of the corresponding fundamental energy quanta can change radically if the interaction parameters of the system change. This will teach us to be more flexible with the term "particles" than we are used to. This is actually one of the most profound new aspects of many-particle quantum mechanics that is covered in this lecture course.

1.6.1 The linear harmonic oscillator as identical-particle system

We kick off with one of the simplest and at the same time one of the most widely used quantum systems: the linear harmonic oscillator. Why the oscillator has such an important role to play in quantum physics will become clear at a later stage.

A linear harmonic oscillator is by definition a 1-particle system. However, it also possesses the hallmark features of an identical-particle system, which will play a crucial role in the description of particles in the context of relativistic QM. A completely worked out example along this line will be the photon description of electromagnetic fields in chapter 4. In order

to highlight the analogy between a linear harmonic oscillator and an identical-particle system we first rewrite the Hamilton operator:

$$\hat{H} \;=\; \frac{\hat{p}_x^2}{2m} + \frac{1}{2}\,m\omega^2\,\hat{x}^2 \;\equiv\; \frac{\hbar\omega}{2}\left(\hat{P}^2+\hat{Q}^2\right)\,,$$

$$\text{with}\quad \hat{P} \;\equiv\; \hat{p}_x\sqrt{\frac{1}{m\hbar\omega}} \qquad \text{and} \qquad \hat{Q} \;\equiv\; \hat{x}\sqrt{\frac{m\omega}{\hbar}}\,. \tag{72}$$

Subsequently we introduce the raising, lowering en number operator according to

$$\underline{\text{raising operator}} \;:\; \hat{a}^\dagger \;\equiv\; \frac{1}{\sqrt{2}}\left(\hat{Q}-i\hat{P}\right)\,,$$

$$\underline{\text{lowering operator}} \;:\; \hat{a} \;\equiv\; \frac{1}{\sqrt{2}}\left(\hat{Q}+i\hat{P}\right)\,,$$

$$\underline{\text{number operator}} \;:\; \hat{n} \;\equiv\; \hat{a}^\dagger\hat{a}\,. \tag{73}$$

The dimensionless position and momentum operators $\hat{Q}$ and $\hat{P}$ have the following fundamental properties:

$$\hat{Q}^\dagger \;=\; \hat{Q} \;=\; \frac{\hat{a}+\hat{a}^\dagger}{\sqrt{2}} \quad,\quad \hat{P}^\dagger \;=\; \hat{P} \;=\; \frac{\hat{a}-\hat{a}^\dagger}{i\sqrt{2}} \qquad \text{and} \qquad \left[\hat{Q},\hat{P}\right] \;=\; \frac{1}{\hbar}\left[\hat{x},\hat{p}_x\right] \;=\; i\hat{1}\,. \tag{74}$$

In the last step we have used that the coordinate x and momentum p_x form a conjugate pair, which implies that the corresponding operators satisfy the usual (canonical) quantum mechanical quantization conditions. From this we can deduce the following features:

- $\hat{a}^\dagger \neq \hat{a}$, i.e. raising and lowering operators are not hermitian and therefore do not correspond to an observable quantity.
- $\hat{n}^\dagger = (\hat{a}^\dagger\hat{a})^\dagger = \hat{n}$, i.e. the number operator does correspond to an observable quantity.
- the raising and lowering operators satisfy bosonic commutation relations:

$$\left[\hat{a},\hat{a}^\dagger\right] \;=\; \tfrac{1}{2}\left[\hat{Q}+i\hat{P},\hat{Q}-i\hat{P}\right] \overset{(74)}{=\!=} \hat{1} \qquad \text{and} \qquad \left[\hat{a}^\dagger,\hat{a}^\dagger\right] \;=\; \left[\hat{a},\hat{a}\right] \;=\; 0\,. \tag{75}$$

- from the relation

$$\hat{P}^2+\hat{Q}^2 \overset{(74)}{=\!=} \left(\frac{\hat{a}-\hat{a}^\dagger}{i\sqrt{2}}\right)^2 + \left(\frac{\hat{a}+\hat{a}^\dagger}{\sqrt{2}}\right)^2 \;=\; \hat{a}^\dagger\hat{a}+\hat{a}\hat{a}^\dagger$$

  it follows that the Hamilton operator is linked directly to the number operator:

$$\hat{H} \;=\; \frac{\hbar\omega}{2}\left(\hat{P}^2+\hat{Q}^2\right) \;=\; \frac{\hbar\omega}{2}\left(\hat{a}^\dagger\hat{a}+\hat{a}\hat{a}^\dagger\right) \overset{(75)}{=\!=} \hbar\omega\left(\hat{n}+\tfrac{1}{2}\hat{1}\right)\,. \tag{76}$$

A linear harmonic oscillator effectively behaves like an identical-particle system consisting of non-interacting bosons. These bosons can take on precisely one value for the 1-particle energy, i.e. $E_1 = \hbar\omega$. As such they can be regarded as the energy quanta belonging to the considered vibrational/oscillatory motions. These quanta are created by $\hat{a}^\dagger$, annihilated by $\hat{a}$ and counted by $\hat{n}$. Except for a constant term (zero-point energy) $\frac{1}{2}\hbar\omega$, the Hamilton-operator takes the form of a many-particle observable belonging to an additive 1-particle quantity.

> *The fundamental energy quanta of an oscillator system can be interpreted as non-interacting bosonic particles. Nature displays many different bosonic vibration/oscillation quanta, in particular those that provide a particle interpretation to wave phenomena in classical physics. Well-known examples are photons, related to electromagnetic waves (see Ch. 4), and phonons, related to collective vibrational modes of lattices and other sound waves (see exercise 5).*

In accordance with § 1.2.1 the energy eigenstates of the linear harmonic oscillator can be written as

$$|n\rangle \overset{(28)}{=\!=} \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}\,|0\rangle\,, \qquad \text{with} \quad \hat{H}\,|n\rangle \overset{(11)}{=\!=} \hbar\omega\big(n+\tfrac{1}{2}\big)\,|n\rangle \qquad (n=0,1,2,\cdots)\,. \quad (77)$$

The meaning of the designations "raising" and "lowering" becomes apparent if we consider the action of the creation and annihilation operators:

$$\hat{a}^\dagger|n\rangle \overset{(30)}{=\!=} \sqrt{n+1}\,|n+1\rangle \quad \Rightarrow \quad \hat{H}\,\hat{a}^\dagger|n\rangle \;=\; \hbar\omega\big(n+1+\tfrac{1}{2}\big)\,\hat{a}^\dagger|n\rangle\,,$$

$$\hat{a}|n\rangle \overset{(30)}{=\!=} \sqrt{n}\,|n-1\rangle \quad \Rightarrow \quad \hat{H}\,\hat{a}|n\rangle \;=\; \hbar\omega\big(n-1+\tfrac{1}{2}\big)\,\hat{a}|n\rangle\,. \quad (78)$$

The raising (lowering) operator changes a given energy eigenstate into the energy eigenstate belonging to the next excited (previous de-excited) energy level.

Remark: if we would have considered a fermionic oscillator

$$\hat{H} \;=\; \hbar\omega\big(\hat{n}+C\,\hat{1}\big) \;=\; \hbar\omega\big(\hat{a}^\dagger\hat{a}+C\,\hat{1}\big) \qquad (C\in\mathbb{R})\,,$$

$$\text{with} \quad \big\{\hat{a},\hat{a}^\dagger\big\} \;=\; \hat{1} \qquad \text{and} \qquad \big\{\hat{a}^\dagger,\hat{a}^\dagger\big\} \;=\; \big\{\hat{a},\hat{a}\big\} \;=\; 0\,, \quad (79)$$

Pauli's exclusion principle would have allowed for only two energy eigenvalues, $\hbar\omega C$ and $\hbar\omega\,(C+1)$ corresponding to $n=0$ and $n=1$ respectively. Still, the operators $\hat{a}^\dagger$ and $\hat{a}$ would have been raising and lowering operators, since the commutator identity $\big[\hat{a},\hat{n}\big]=\hat{a}$ holds equally well in the fermionic case.

1.6.2 Forced oscillators: coherent states and quasi particles

Let's now add an interaction to the linear harmonic oscillator of § 1.6.1 according to

$$\hat{H} \;=\; \hbar\omega\big(\hat{n} - \lambda\hat{a}^\dagger - \lambda^*\hat{a} + \tfrac{1}{2}\hat{1}\big) \;=\; \hbar\omega\big(\hat{n} + \tfrac{1}{2}\hat{1}\big) - \hbar\omega\,(\mathrm{Re}\lambda)\big(\hat{a}^\dagger + \hat{a}\big) - \hbar\omega\,(\mathrm{Im}\lambda)\big(i\hat{a}^\dagger - i\hat{a}\big)$$

$$\overset{(72),(74)}{=\!=\!=} \frac{\hat{p}_x^2}{2m} + \frac{1}{2}\,m\omega^2\hat{x}^2 - \hat{x}\,\mathrm{Re}\lambda\sqrt{2m\hbar\omega^3} - \hat{p}_x\,\mathrm{Im}\lambda\sqrt{2\hbar\omega/m} \qquad (\lambda\in\mathbb{C})\;. \tag{80}$$

Systems of this type are called forced linear harmonic oscillators. They are omnipresent in nature. For instance, taking $\lambda = q\mathcal{E}/\sqrt{2m\hbar\omega^3}$ and $\mathcal{E}\in\mathbb{R}$ we obtain an interaction term of the form $\hat{V} = -q\hat{x}\mathcal{E}$, which we recognize as the interaction term of a charged linear harmonic oscillator with charge q subjected to a classical electric field $\mathcal{E}$ along the x-axis. The striking feature of the additional interaction term is that it does not conserve the number of oscillator quanta, since

$$\big[\hat{n},\hat{H}\big] \;=\; -\hbar\omega\big[\hat{n},\lambda\hat{a}^\dagger + \lambda^*\hat{a}\big] \;\overset{(16)}{=\!=\!=}\; -\hbar\omega\,(\lambda\hat{a}^\dagger - \lambda^*\hat{a}) \;\neq\; 0\;.$$

This is a first example of what is referred to as a non-additive interaction, i.e. an interaction that does not conserve the total number of particles.

A new particle interpretation: nature is riddled with systems that involve non-additive interactions. For instance, such systems are ubiquitous in particle physics (e.g. through the production and decay of particles), quantum optics/spectroscopy (e.g. in laser systems and when photons are being absorbed/emitted by atoms) and quantum electronics (e.g. in describing charge flows).

> *In scenarios involving non-additive interactions it usually pays to switch to an alternative particle interpretation, i.e. an alternative description of the interacting quantum system in terms of non-interacting energy quanta that are called quasi particles. Expressed in terms of these new energy quanta the Hamilton operator will take on an additive 1-particle form, which causes the number of quasi particles to be conserved. Furthermore, this implies the existence of a representation that will make the Hamilton operator expressible in terms of number operators for quasi particles.*

To get a feel for what a quasi particle could be, consider a particle that propagates through a medium while influencing other particles and while being influenced by other particles. The combination of particle and ambient influences can collectively behave as a separate free entity. For instance, a conduction electron propagating through a piece of metal can excite vibrations (phonons) in the metal-ion lattice, while in turn the lattice vibrations can affect the motions of the conduction electrons. The degree by which the quasi particles will differ from the original particles depends on the nature and strength of the ambient interactions.

In all examples covered in this lecture course, the original particle interpretation is retrieved in the limit of vanishing ambient influences. If the original particles are bosons (fermions) then the same should be true for the quasi particles.

Therefore, while switching from particles to quasi particles, one set of bosonic (fermionic) creation and annihilation operators will be replaced by another set. As such, it will not come as a surprise that a particle–quasi particle transition actually boils down to performing a unitary transformation in Fock space, since operator identities such as bosonic commutation relations (fermionic anticommutation relations) will be invariant in that case.

Quasi particles and coherent states (Roy J. Glauber, 1963): in case of a forced linear harmonic oscillator we can switch to new bosonic quasi-particle operators $\hat{c}$ and $\hat{c}^\dagger$ by means of the transformation

$$\hat{c} \;=\; \hat{a} - \lambda\hat{1} \qquad , \qquad \hat{c}^\dagger \;=\; \hat{a}^\dagger - \lambda^*\hat{1} \qquad \Rightarrow \qquad \big[\hat{c},\hat{c}^\dagger\big] \;=\; \big[\hat{a},\hat{a}^\dagger\big] \;=\; \hat{1}\;. \tag{81}$$

The number operator for quasi particles $\hat{\tilde{n}}$ now reads

$$\hat{\tilde{n}} \;=\; \hat{c}^\dagger\hat{c} \;\overset{(81)}{=\!=}\; \hat{n} \,-\, \lambda\hat{a}^\dagger \,-\, \lambda^*\hat{a} \,+\, |\lambda|^2\hat{1} \quad \Rightarrow \quad \hat{H} \;=\; \hbar\omega\big(\hat{\tilde{n}} \,+\, \tfrac{1}{2}\hat{1} \,-\, |\lambda|^2\hat{1}\big)\;, \tag{82}$$

allowing us to express the energy eigenstates of the systen in terms of quasi-particle states:

$$|\tilde{n}\rangle_\lambda \;\overset{(28)}{=\!=}\; \frac{(\hat{c}^\dagger)^n}{\sqrt{n!}}\,|\tilde{0}\rangle_\lambda \quad , \quad \hat{H}\,|\tilde{n}\rangle_\lambda \;=\; \hbar\omega\big(\tilde{n}+\tfrac{1}{2}-|\lambda|^2\big)\,|\tilde{n}\rangle_\lambda \qquad (\tilde{n}=0,1,2,\cdots)\;. \tag{83}$$

The ground state of the forced linear harmonic oscillator is given by the state without any quasi particles (quasi-particle vacuum) $|\tilde{0}\rangle_\lambda \equiv |\lambda\rangle$, which satisfies

$$\boxed{\hat{a}\,|\lambda\rangle \;\overset{(81)}{=\!=}\; \big(\hat{c}+\lambda\hat{1}\big)\,|\lambda\rangle \;\overset{\text{vacuum}}{=\!=}\; \lambda|\lambda\rangle}\;. \tag{84}$$

This quasi-particle vacuum is an eigenstate of the original annihilation operator $\hat{a}$ belonging to the eigenvalue λ. Such a state is called a coherent state, where the term coherent refers to the fact that all original many-particle states contribute to this state with related phases, as opposed to incoherent situations where the phases are completely random (see Ch. 2).

Proof: express the state $|\lambda\rangle$ in terms of the original many-particle basis according to

$$|\lambda\rangle \;\equiv\; \sum_{n=0}^{\infty} C_n|n\rangle \;\overset{(84)}{\Longrightarrow}\; 0 \;=\; \sum_{n=0}^{\infty} C_n\big(\hat{a}-\lambda\hat{1}\big)|n\rangle \;\overset{(78)}{=\!=}\; \sum_{n=0}^{\infty} C_n\big(\sqrt{n}\,|n-1\rangle - \lambda|n\rangle\big)$$

$$\overset{\text{basis}}{\Longrightarrow}\; C_{n+1} \;=\; \frac{\lambda C_n}{\sqrt{n+1}} \;=\; \cdots \;=\; \frac{\lambda^{n+1}}{\sqrt{(n+1)!}}\,C_0\;,$$

where the summation parameter n is increased by one in the term containing $|n-1\rangle$. Next we have to normalize the state $|\lambda\rangle$:

$$1 \equiv \langle\lambda|\lambda\rangle = |C_0|^2 \sum_{n,n'=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} \frac{(\lambda^*)^{n'}}{\sqrt{n'!}} \langle n'|n\rangle \overset{\text{orth. basis}}{=\!=\!=} |C_0|^2 \sum_{n=0}^{\infty} \frac{|\lambda|^{2n}}{n!} = |C_0|^2 \exp\big(|\lambda|^2\big) .$$

By choosing $C_0 = \exp(-|\lambda|^2/2)$ we finally obtain

$$\boxed{|\lambda\rangle = \sum_{n=0}^{\infty} \exp(-|\lambda|^2/2) \frac{\lambda^n}{\sqrt{n!}} |n\rangle} , \tag{85}$$

which can be rewritten by employing equation (28):

$$|\lambda\rangle = \sum_{n=0}^{\infty} \exp(-|\lambda|^2/2) \frac{\lambda^n}{\sqrt{n!}} \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}} |0\rangle = \exp(-|\lambda|^2/2) \exp(\lambda \hat{a}^\dagger)|0\rangle . \tag{86}$$

Indeed we observe that all original many-particle states feature in such a coherent state, with the expectation value for the occupation number and associated quantum mechanical uncertainty given by

$$\bar{n} \equiv \langle\lambda|\hat{n}|\lambda\rangle = \langle\lambda|\hat{a}^\dagger\hat{a}|\lambda\rangle \overset{(84)}{=\!=\!=} |\lambda|^2 \langle\lambda|\lambda\rangle = |\lambda|^2 ,$$

$$\Delta n \equiv \sqrt{\langle\lambda|\hat{n}^2|\lambda\rangle - \langle\lambda|\hat{n}|\lambda\rangle^2} = \sqrt{\langle\lambda|(\hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + \hat{a}^\dagger\hat{a})|\lambda\rangle - \bar{n}^2} \overset{(84)}{=\!=\!=} \sqrt{\bar{n}} . \tag{87}$$

Here we used that $\hat{n}^2 = \hat{a}^\dagger\hat{a}\hat{a}^\dagger\hat{a} \overset{(75)}{=\!=\!=} \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}$. The larger the value of $|\lambda| = \sqrt{\bar{n}}$ becomes, the larger the average number will be of original oscillator quanta that are present in the ground state of the forced oscillator and the smaller the relative error $\Delta n/\bar{n} = 1/\sqrt{\bar{n}} = 1/|\lambda|$ on this number will be!

The relation between particle number and phase: writing the complex eigenvalue λ as $\lambda = |\lambda| \exp(i\phi_\lambda)$, the corresponding coherent state has the characteristic feature that the many-particle basis states $|n\rangle$ occur with phases that are related by $\exp(in\phi_\lambda)$:

$$|\lambda\rangle \overset{(86)}{=\!=\!=} \sum_{n=0}^{\infty} \exp(-|\lambda|^2/2) \frac{\big(|\lambda|\,\hat{a}^\dagger\big)^n}{n!} \exp(in\phi_\lambda)|0\rangle = \exp(-|\lambda|^2/2) \exp\big(|\lambda| e^{i\phi_\lambda}\hat{a}^\dagger\big)|0\rangle .$$

$$\Rightarrow \quad -i\frac{\partial}{\partial\phi_\lambda}|\lambda\rangle = \hat{a}^\dagger|\lambda|e^{i\phi_\lambda}|\lambda\rangle = \hat{a}^\dagger\lambda|\lambda\rangle \overset{(84)}{=\!=\!=} \hat{a}^\dagger\hat{a}|\lambda\rangle = \hat{n}|\lambda\rangle .$$

Since the coherent states form a complete set (see remark 1), the following relation between particle number and phase emerges

$$-i\frac{\partial}{\partial\phi_\lambda} = \hat{n}$$

in the coherent-state representation. This resembles the quantum mechanical relation between conjugate pairs of coordinates and momenta, such as the coordinate x and the corresponding momentum p_x for which we have the position-space relation $\hat{p}_x = -i\hbar\partial/\partial x$. As such we expect to find a kind of particle number–phase uncertainty relation of the form

$$\Delta n \, \Delta\phi_\lambda \geq 1/2 \; , \tag{88}$$

although it is not possible to define a unique phase operator in QM in view of the fact that ϕ_λ is defined on the interval $[0, 2\pi)$ only. For a fixed-n state we would therefore expect to have a completely undetermined phase. Actually we have already made use of this in § 1.2 by exploiting the freedom to choose the overall phase of the occupation-number basis states. For an arbitrary coherent state we expect to find a compromise between uncertainty in particle number and uncertainty in phase (see remark 2)

Pseudo-classical properties of the coherent state $|\lambda\rangle$: we know that

$$\langle\lambda|(\hat{a}^\dagger)^k(\hat{a})^l|\lambda\rangle \;=\; (\lambda^*)^k(\lambda)^l \;=\; \langle\lambda|\hat{a}^\dagger|\lambda\rangle^k\langle\lambda|\hat{a}|\lambda\rangle^l$$

for normal-ordered operators, seemingly bypassing quantum mechanical commutation relations by allowing the operators to be effectively replaced by complex parameters. Well, hold your horses $\cdots$ this classical behaviour does not persist once we consider operators that are not normal ordered, such as

$$\langle\lambda|\hat{a}\,\hat{a}^\dagger|\lambda\rangle \;=\; \langle\lambda|(\hat{a}^\dagger\hat{a} + \hat{1})|\lambda\rangle \;=\; |\lambda|^2 + 1 \;=\; \langle\lambda|\hat{a}|\lambda\rangle\langle\lambda|\hat{a}^\dagger|\lambda\rangle + 1 \; .$$

The commutation relations lead to relative corrections of $\mathcal{O}\big(1/|\lambda|^2\big) = \mathcal{O}(1/\bar{n})$ to classical behaviour. In the high occupation-number limit $|\lambda| \gg 1$ this correction is suppressed and a classical situation is approached that allows you to effectively replace the relevant many-particle operators by complex numbers in the coherent-state expectation value. Moreover, it will be shown in exercise 6 that the relative overlap between different coherent states decreases substantially in that case (see also remark 2), allowing the coherent states to be viewed as pseudo classical states consisting of an almost fixed, very large number of oscillator quanta.

Unitary transformation: finally it will be shown in exercise 6 that a coherent state can be obtained from the ground state of the linear harmonic oscillator. As expected this will be achieved by means of a unitary transformation in Fock space:

$$|\lambda\rangle \;=\; \exp(\lambda\hat{a}^\dagger - \lambda^*\hat{a})\,|0\rangle \; . \tag{89}$$

Coherent states are particularly useful in describing electromagnetic radiation that is produced coherently through stimulated emission. For instance, masers and lasers fall into this category as well as electromagnetic fields produced by

classical electric currents (see Ch. 4). Coherent states also feature prominently in the description of systems with macroscopically occupied 1-particle states, such as Bose–Einstein condensates (see Ch. 2) and classical electromagnetic fields (see Ch. 4).

Remark 1: overcomplete set (no exam material).

The collection of all coherent states

$$\big\{\,|\lambda\rangle \;=\; \sum_{n=0}^{\infty} \exp(-|\lambda|^2/2)\,\frac{\lambda^n}{\sqrt{n!}}\,|n\rangle \;:\; \lambda\in\mathbb{C}\,\big\} \tag{90}$$

forms an overcomplete set. To this end we write the complex eigenvalue λ in terms of polar coordinates in the complex plane as $\lambda = |\lambda|\exp(i\phi_\lambda)$. That allows us to derive the following generalized completeness relation for coherent states:

$$\int_{-\infty}^{\infty}\mathrm{d}(\mathrm{Re}\,\lambda)\int_{-\infty}^{\infty}\mathrm{d}(\mathrm{Im}\,\lambda)\;|\lambda\rangle\langle\lambda| \;=\; \int_{-\infty}^{\infty}\mathrm{d}(\mathrm{Re}\,\lambda)\int_{-\infty}^{\infty}\mathrm{d}(\mathrm{Im}\,\lambda)\;\exp(-|\lambda|^2)\sum_{n,n'=0}^{\infty}\frac{\lambda^n(\lambda^*)^{n'}}{\sqrt{n!\,n'!}}\,|n\rangle\langle n'|$$

$$\overset{\text{polar coord.}}{=\!=\!=} \sum_{n,n'=0}^{\infty}\frac{|n\rangle\langle n'|}{\sqrt{n!\,n'!}}\int_0^{\infty}\mathrm{d}|\lambda|\,|\lambda|^{n+n'+1}\exp(-|\lambda|^2)\overbrace{\int_0^{2\pi}\mathrm{d}\phi_\lambda\,\exp\big(i\phi_\lambda[n-n']\big)}^{2\pi\,\delta_{nn'}}$$

$$\overset{z=|\lambda|^2}{=\!=\!=} \pi\sum_{n=0}^{\infty}\frac{|n\rangle\langle n|}{n!}\int_0^{\infty}\mathrm{d}z\;z^n\exp(-z) \;=\; \pi\sum_{n=0}^{\infty}|n\rangle\langle n| \;\overset{\text{compl.}}{=\!=\!=}\; \pi\,\hat{1}\;, \tag{91}$$

where it is used that $\int_0^\infty \mathrm{d}z\;z^n\exp(-z) = n!$ for $n = 0, 1, \cdots$. An arbitrary many-particle state is therefore readily expressible in terms of coherent states!

Remark 2: two more pseudo-classical properties (no exam material).

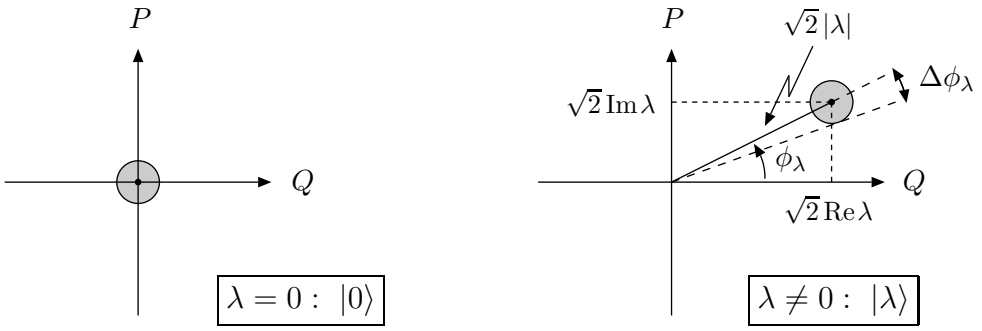
A coherent state is also an example of a quantum state with minimal uncertainty that is divided equally over position and momentum operators. By combining the results on the previous page with equation (74), it can be proven straightforwardly that

$$(\Delta P)^2 \;=\; \langle\lambda|\hat{P}^2|\lambda\rangle - \langle\lambda|\hat{P}|\lambda\rangle^2 = \;-\tfrac{1}{2}\langle\lambda|(\hat{a}-\hat{a}^\dagger)^2|\lambda\rangle + \tfrac{1}{2}\langle\lambda|(\hat{a}-\hat{a}^\dagger)|\lambda\rangle^2 = \;\tfrac{1}{2}\;,$$

$$(\Delta Q)^2 \;=\; \langle\lambda|\hat{Q}^2|\lambda\rangle - \langle\lambda|\hat{Q}|\lambda\rangle^2 = \;+\tfrac{1}{2}\langle\lambda|(\hat{a}+\hat{a}^\dagger)^2|\lambda\rangle - \tfrac{1}{2}\langle\lambda|(\hat{a}+\hat{a}^\dagger)|\lambda\rangle^2 = \;\tfrac{1}{2}\;,$$

which indeed yields $(\Delta x)(\Delta p_x) \overset{(72)}{=\!=\!=} \hbar(\Delta P)(\Delta Q) = \hbar/2$. The quantum mechanical expectation values and uncertainties in the (Q,P)-plane are illustrated schematically in the picture displayed below. From the quantum mechanical uncertainties (i.e. the grey

circles with radius $1/\sqrt{2}$) it can be read off that for increasing $|\lambda|$ both the relative uncertainty in $|\lambda|$ and the absolute uncertainty in the phase ϕ_λ get smaller. This is in marked contrast to the case $\lambda = 0$, for which the phase ϕ_λ is completely undetermined! As will be shown in exercise 6, for $|\lambda| \gg 1$ a classical situation is approached with substantially decreased relative overlap between coherent states. In addition, upon close inspection we indeed observe the anticipated uncertainty relation between particle number and phase: $\sqrt{2}\,|\lambda|\,\Delta\phi_\lambda \geq 1/\sqrt{2} \stackrel{(87)}{\Longrightarrow} \Delta n\,\Delta\phi_\lambda \geq 1/2$.



We finalize by noting that a coherent state is the quantum state that comes closest in approximating the classical motion of a non-interacting linear harmonic oscillator. Equation (65) tells us that the annihilation operator $\hat{a}$ of a linear harmonic oscillator system is given in the Heisenberg picture by

$$\hat{a}_{_H}(t) \,=\, \hat{a}\,\exp(-i\omega t) \quad \Rightarrow \quad \hat{a}_{_H}(t)\,|\lambda\rangle \,=\, \lambda\,\exp(-i\omega t)\,|\lambda\rangle \,=\, |\lambda|\,\exp\big(i[\phi_\lambda - \omega t]\big)\,|\lambda\rangle\ ,$$

implying that

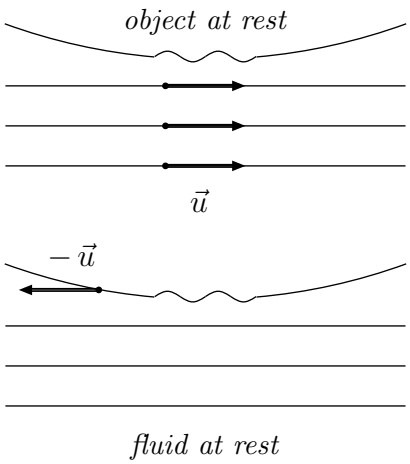
$$\langle\lambda|\hat{Q}_{_H}(t)|\lambda\rangle \,=\, \sqrt{2}\,\mathrm{Re}\big(\lambda\,\exp(-i\omega t)\big) \qquad \text{and} \qquad \langle\lambda|\hat{P}_{_H}(t)|\lambda\rangle \,=\, \sqrt{2}\,\mathrm{Im}\big(\lambda\,\exp(-i\omega t)\big)$$

indeed reproduce the time-dependent motion of a linear harmonic oscillator with a minimum of quantum mechanical uncertainty.

1.6.3 Superfluidity

As a second application of the concept of quasi particles, we consider the quantum-mechanical low-temperature phenomenon of superfluidity. The explicit example that will be investigated will teach us that a low-energy spectrum of excitations in a bosonic many-particle system can change radically upon switching from a non-interacting (ideal-gas) scenario to a scenario in which the particles repel each other weakly. Before we go into the details, first the concept of superfluidity will be introduced.

We speak of a superfluid flow inside a medium if it is not subject to energy leakage through friction. By means of an elegant argument (formulated by L.D. Landau in 1941) it can be established when a superfluid flow is possible. To this end we consider a fluid that flows at uniform velocity $\vec{u}$ past a macroscopic object. Next we switch to a frame of reference in which the fluid is at rest (at absolute zero of temperature) and the object moves at velocity $-\vec{u}$. Friction (viscosity) occurs only then, when the object can convert part of its kinetic momentum and energy into excitations in the fluid. Let's have a closer look at such an excitation with momentum $\vec{p}$ and energy $\epsilon(\vec{p})$. The conserved total momentum and energy of the combined system, consisting of the fluid and the macroscopic object with mass M, is given by



$$\vec{P} \;=\; -M\vec{u} \;=\; \vec{p} + M\vec{u}' \qquad \text{and} \qquad E \;=\; \frac{1}{2}M\vec{u}^{\,2} \;=\; \epsilon(\vec{p}) \;+\; \frac{1}{2}M\vec{u}'^{\,2}\;,$$

with $\vec{u}'$ being the velocity of the object after exciting the fluid. By merging both equations one obtains

$$\epsilon(\vec{p}) + \vec{p}\cdot\vec{u} \;=\; -\frac{\vec{p}^{\,2}}{2M}\;.$$

From this it follows that the excitation is energetically not possible if the flow velocity of the fluid satisfies the condition

$$u \;\equiv\; |\vec{u}| \;<\; \frac{\epsilon(\vec{p})}{|\vec{p}|}\;.$$

If the spectrum of fluid excitations $\epsilon(\vec{p})$ displays a minimum (critical velocity)

$$\boxed{u_c \;=\; \Big(\frac{\epsilon(\vec{p})}{|\vec{p}|}\Big)_{\rm min} \;>\; 0}\;, \tag{92}$$

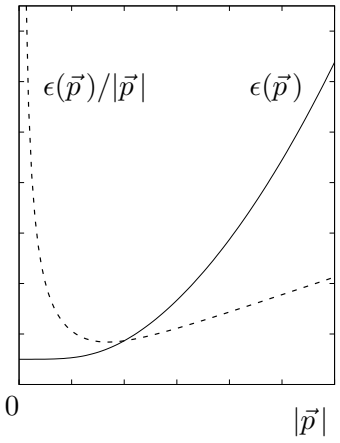
then superfluidity can set in at velocities $u < u_c$ provided that the temperature is sufficiently low. This is the so-called Landau criterion for superfluidity.

We can distinguish a few special shapes (dispersion relations) of the excitation spectrum.

- $\epsilon(\vec{p}) = \vec{p}^{\,2}/(2m)$: this corresponds to the energy spectrum of free non-relativistic particles. Example: a gas consisting of non-interacting massive particles. In that case $\epsilon(\vec{p})/|\vec{p}| = |\vec{p}|/(2m)$ does not possess a positive minimum and according to Landau's criterion no superfluidity can occur.

- $\epsilon(\vec{p}) = c_s|\vec{p}|$: this corresponds to the energy spectrum of free massless quanta. Example: quantized wave phenomena, such as sound waves in a fluid or crystal (see exercise 5). In that case $\epsilon(\vec{p})/|\vec{p}|$ equals the speed of propagation c_s of the waves inside the medium. According to Landau's criterion, superfluidity can occur for flow velocities $u < c_s$. This example will feature prominently in the explicit calculation presented below.

- The excitation spectrum has a finite energy gap $\Delta = \lim_{|\vec{p}|\to 0} \epsilon(\vec{p})$ between the ground state and the first excited state. Example: the energy gap that features in the energy spectrum of superconductors as a result of attractive interactions between special pairs (Cooper pairs) of conduction electrons. In that case $\epsilon(\vec{p})/|\vec{p}|$ clearly displays a minimum and superfluidity can occur if the corresponding flow velocity is low enough.



1.6.4 Superfluidity for weakly repulsive spin-0 bosons (part 1)

As an example we investigate the low-energy excitation spectrum of weakly repulsive bosons. To this end we consider a system consisting of a very large, constant number N of identical spin-0 particles with mass m. The particles are contained inside a large cube with edges L and periodic boundary conditions, giving rise to a discrete momentum spectrum:

$$\{\vec{p} = \hbar\vec{k} \,:\, k_{x,y,z} \;=\; 0, \pm 2\pi/L, \pm 4\pi/L, \cdots\} \;. \tag{93}$$

In addition we assume the temperature to be low enough for the system to be effectively in the ground state, i.e. effectively we are dealing with a $T = 0$ system.[2]

Without interactions among the particles: the total kinetic energy operator of the non-interacting identical-particle system is diagonal in the momentum representation:

$$\hat{T}_{\rm tot} \;=\; \sum_{\vec{k}} \frac{\hbar^2\vec{k}^{\,2}}{2m}\, \hat{a}^\dagger_{\vec{k}}\, \hat{a}_{\vec{k}} \;=\; \frac{1}{2} \sum_{\vec{k}\neq\vec{0}} \frac{\hbar^2\vec{k}^{\,2}}{2m} \left(\hat{a}^\dagger_{\vec{k}}\, \hat{a}_{\vec{k}} + \hat{a}^\dagger_{-\vec{k}}\, \hat{a}_{-\vec{k}}\right) , \tag{94}$$

with kinetic-energy eigenvalues

$$E^{(0)} \;=\; \sum_{\vec{k}} \frac{\hbar^2\vec{k}^{\,2}}{2m}\, n^{(0)}_{\vec{k}} \;. \tag{95}$$

[2]The relevant details regarding temperature dependence and regarding the quantization aspects of the container can be found in chapter 2.

The second expression for $\hat{T}_{\rm tot}$ is given for practical purposes only (see below). It exploits the symmetry of the momentum summation under inversion of the momenta. By adding the total momentum operator

$$\hat{\vec{P}}_{\rm tot} \;=\; \sum_{\vec{k}} \hbar\vec{k}\;\hat{a}^\dagger_{\vec{k}}\,\hat{a}_{\vec{k}} \;=\; \frac{1}{2}\sum_{\vec{k}\neq\vec{0}} \hbar\vec{k}\,\left(\hat{a}^\dagger_{\vec{k}}\,\hat{a}_{\vec{k}} - \hat{a}^\dagger_{-\vec{k}}\,\hat{a}_{-\vec{k}}\right)\,, \tag{96}$$

we can readily read off the particle interpretation belonging to the creation and annihilation operators used. Particles with energy $\hbar^2\vec{k}^{\,2}/(2m)$ and momentum $\hbar\vec{k}$ are created by $\hat{a}^\dagger_{\vec{k}}$, annihilated by $\hat{a}_{\vec{k}}$ and counted by $\hat{n}_{\vec{k}} = \hat{a}^\dagger_{\vec{k}}\,\hat{a}_{\vec{k}}$. The occupation number $n^{(0)}_{\vec{k}}$ indicates how many particles can be found in the given momentum eigenstate in the absence of mutual interactions. The ground state of the non-interacting N-particle system automatically has $n^{(0)}_{\vec{0}} = N$ and $n^{(0)}_{\vec{k}\neq\vec{0}} = 0$.

Including a weak repulsive interaction among the particles: in analogy with exercise 4, the inclusion of weak pair interactions that depend exclusively on the distance between the two particles will lead to an additive many-particle interaction term of the form

$$\hat{V} \;=\; \frac{1}{2}\sum_{\vec{k},\vec{k}',\vec{q}} \mathcal{U}(q)\,\hat{a}^\dagger_{\vec{k}}\,\hat{a}^\dagger_{\vec{k}'}\,\hat{a}_{\vec{k}'+\vec{q}}\,\hat{a}_{\vec{k}-\vec{q}}\,, \tag{97}$$

where

$$\mathcal{U}(q) \;\equiv\; \mathcal{U}(|\vec{q}|) \;=\; \frac{1}{V}\int_V \mathrm{d}\vec{r}\;\mathcal{U}(r)\exp(-i\,\vec{q}\cdot\vec{r}) \tag{98}$$

is the Fourier transform of the spatial pair interaction per unit of volume. As argued in exercise 4, the pair momentum of the interacting particles remains conserved under the interaction (owing to translational symmetry) and $\mathcal{U}(q)$ merely depends on the absolute value of $\vec{q}$ (owing to rotational symmetry).

What do we expect for the energy eigenstates in the ineracting case?

- First of all the state $|n_{\vec{0}} = N, n_{\vec{k}\neq\vec{0}} = 0\rangle$ will no longer be the ground state of the interacting many-particle system. By applying $\hat{V}$ we observe

$$\begin{aligned}\hat{V}\,|n_{\vec{0}} = N, n_{\vec{k}\neq\vec{0}} = 0\rangle \;&=\; \frac{1}{2}\sum_{\vec{q}} \mathcal{U}(q)\,\hat{a}^\dagger_{\vec{q}}\,\hat{a}^\dagger_{-\vec{q}}\,\hat{a}_{\vec{0}}\,\hat{a}_{\vec{0}}\,|n_{\vec{0}} = N, n_{\vec{k}\neq\vec{0}} = 0\rangle \\ &=\; \frac{1}{2}\,N(N-1)\,\mathcal{U}(0)\,|n_{\vec{0}} = N, n_{\vec{k}\neq\vec{0}} = 0\rangle \\ &\quad+\; \frac{1}{2}\,\sqrt{N(N-1)}\sum_{\vec{q}\neq\vec{0}} \mathcal{U}(q)\,|n_{\vec{0}} = N-2, n_{\vec{q}} = n_{-\vec{q}} = 1, \text{ other } n_{\vec{k}} = 0\rangle\,.\end{aligned}$$

  Therefore we expect the true ground state to receive explicit (small) contributions from pairs of particles that occupy excited 1-particle eigenstates with opposite momentum. In that way the net momentum $\vec{P}_{\rm tot} = \vec{0}$ indeed remains unaffected.

- In the non-interacting case the low-energy N-particle excitations involve just a few particles in excited 1-particle states with $\mathcal{O}(h/L)$ momenta. Since all particles then come with $\mathcal{O}(L)$ de Broglie wavelengths, we expect the quantum mechanical influence of the particles to extend across the entire system. Upon including repulsive effects, we expect a single excitation to involve more than just bringing a single particle into motion and hence to require more energy than in the non-interacting case.

  Approximations for weakly repulsive dilute Bose gases (Bogolyubov, 1947):
  if the pair interaction is sufficiently weak and repulsive, we expect for the low-energy N-particle states that still almost all particles occupy the 1-particle ground state, i.e. $N - \langle \hat{n}_{\vec{0}} \rangle \ll N$. This allows us to apply two approximation steps to simplify the many-particle interaction term, which hold as long as the gas is sufficiently dilute to avoid too many particles from ending up in excited 1-particle states.

Approximation 1: for obtaining the low-energy N-particle states the effect from interactions among particles in excited 1-particle states can be neglected. This boils down to only considering interaction terms with two or more creation and annihilation operators that belong to the 1-particle ground state:

$$\hat{V} \;\approx\; \frac{1}{2}\,\mathcal{U}(0)\,\overbrace{\hat{a}^\dagger_{\vec{0}}\,\hat{a}^\dagger_{\vec{0}}\,\hat{a}_{\vec{0}}\,\hat{a}_{\vec{0}}}^{\hat{n}^2_{\vec{0}}-\hat{n}_{\vec{0}}} \;+\; \mathcal{U}(0)\,\overbrace{\hat{a}^\dagger_{\vec{0}}\,\hat{a}_{\vec{0}}}^{\hat{n}_{\vec{0}}}\sum_{\vec{k}\neq\vec{0}}\overbrace{\hat{a}^\dagger_{\vec{k}}\,\hat{a}_{\vec{k}}}^{\hat{n}_{\vec{k}}} \;+\; \frac{1}{2}\,\overbrace{\hat{a}^\dagger_{\vec{0}}\,\hat{a}_{\vec{0}}}^{\hat{n}_{\vec{0}}}\sum_{\vec{q}\neq\vec{0}}\mathcal{U}(q)\Big(\hat{a}^\dagger_{\vec{q}}\,\hat{a}_{\vec{q}}+\hat{a}^\dagger_{-\vec{q}}\,\hat{a}_{-\vec{q}}\Big)$$
$$+\;\frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\mathcal{U}(q)\Big(\hat{a}_{\vec{0}}\,\hat{a}_{\vec{0}}\,\hat{a}^\dagger_{\vec{q}}\,\hat{a}^\dagger_{-\vec{q}}+\hat{a}^\dagger_{\vec{0}}\,\hat{a}^\dagger_{\vec{0}}\,\hat{a}_{\vec{q}}\,\hat{a}_{-\vec{q}}\Big)\;.$$

- The first term corresponds to the configuration $\vec{k}=\vec{k}'=\vec{q}=\vec{0}$, where all creation and annihilation operators refer to the 1-particle ground state.

- The remaining terms cover situations where only two out of three momenta vanish. The second term corresponds to the configurations $\vec{k}=\vec{q}=\vec{0}$ and $\vec{k}'=\vec{q}=\vec{0}$, the third term to $\vec{k}'=\vec{k}-\vec{q}=\vec{0}$ and $\vec{k}=\vec{k}'+\vec{q}=\vec{0}$, and the last term to $\vec{k}-\vec{q}=\vec{k}'+\vec{q}=\vec{0}$ and $\vec{k}=\vec{k}'=\vec{0}$.

Subsequently the number operator $\hat{n}_{\vec{0}}$ can be replaced everywhere by $\hat{N}-\sum\limits_{\vec{q}\neq\vec{0}}\hat{n}_{\vec{q}}$ and the total number of particles N can be taken as fixed and very large:

$$\hat{V} \;\approx\; \frac{N(N-1)}{2}\,\mathcal{U}(0) \;+\; \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\mathcal{U}(q)\Big(N\,\hat{a}^\dagger_{\vec{q}}\,\hat{a}_{\vec{q}}+N\,\hat{a}^\dagger_{-\vec{q}}\,\hat{a}_{-\vec{q}}+\hat{a}_{\vec{0}}\,\hat{a}_{\vec{0}}\,\hat{a}^\dagger_{\vec{q}}\,\hat{a}^\dagger_{-\vec{q}}+\hat{a}^\dagger_{\vec{0}}\,\hat{a}^\dagger_{\vec{0}}\,\hat{a}_{-\vec{q}}\,\hat{a}_{\vec{q}}\Big)\;.$$

Approximation 2, traditional approach: in Bogolyubov's approach it is used that the operators $\hat{a}_{\vec{0}}\,\hat{a}_{\vec{0}}$ and $\hat{a}^{\dagger}_{\vec{0}}\,\hat{a}^{\dagger}_{\vec{0}}$ can be effectively replaced by N when applied to the lowest-energy N-particle states. In principle this could involve extra phase factors $\mathrm{e}^{2i\phi_0}$ and $\mathrm{e}^{-2i\phi_0}$, however, these can be absorbed into a redefinition of the remaining creation and annihilation operators. This approach suggests that we are dealing with an approximately classical situation, where the fact that the operators $\hat{a}_{\vec{0}}$ and $\hat{a}^{\dagger}_{\vec{0}}$ do not commute only affect the considered N-particle states in a negligible way (*as if they were coherent states with* $|\lambda| \gg 1$). This assumption is plausible, bearing in mind that $\sqrt{N-n} \approx \sqrt{N}$ if $N \gg n$. As a result, magnitude-wise the action of $\hat{a}_{\vec{0}}$ and $\hat{a}^{\dagger}_{\vec{0}}$ on states with $N - \langle \hat{n}_{\vec{0}} \rangle \ll N$ will be effectively the same. In this way the following effective approximation is obtained for the total Hamilton operator $\hat{H}_{\mathrm{tot}} = \hat{T}_{\mathrm{tot}} + \hat{V}$ applicable to the lowest lying energy eigenstates:

$$\begin{aligned} \hat{H}_{\mathrm{tot}} \;\approx\; & \frac{N(N-1)}{2}\,\mathcal{U}(0) \,+\, \frac{1}{2}\sum_{\vec{q}\neq\vec{0}} \Big(\frac{\hbar^2 q^2}{2m} + N\mathcal{U}(q)\Big)\Big(\hat{a}^{\dagger}_{\vec{q}}\,\hat{a}_{\vec{q}} + \hat{a}^{\dagger}_{-\vec{q}}\,\hat{a}_{-\vec{q}}\Big) \\ & + \,\frac{1}{2}\sum_{\vec{q}\neq\vec{0}} N\mathcal{U}(q)\Big(\hat{a}^{\dagger}_{\vec{q}}\,\hat{a}^{\dagger}_{-\vec{q}} + \hat{a}_{-\vec{q}}\,\hat{a}_{\vec{q}}\Big)\,. \end{aligned} \tag{99}$$

Owing to this second approximation the total number of particles is no longer conserved under the interactions. The justification for such an approach is purely thermodynamic by nature (see Ch. 2): "the physical properties of a system with a very large number of particles do not change when adding/removing a particle". Since approximations are unavoidable for the description of complex many-particle systems, non-additive quantities of this type are ubiquitous in condensed-matter and low-temperature physics. Note, however, that the total momentum of the many-particle system remains conserved under the interaction, since each term in $\hat{H}$ adds as much momentum as it subtracts.

Approximation 2, but this time conserving particle number (based on the bachelor thesis by Leon Groenewegen): in order to avoid that particle number is not conserved during the second approximation step we can again exploit approximation 1 and write $N \approx \hat{a}_{\vec{0}}\,\hat{a}^{\dagger}_{\vec{0}}$. Without loss of accuracy this allows to rewrite the total Hamilton operator $\hat{H}_{\mathrm{tot}} = \hat{T}_{\mathrm{tot}} + \hat{V}$ as

$$\begin{aligned} \hat{H}_{\mathrm{tot}} \;\approx\; & \frac{N(N-1)}{2}\,\mathcal{U}(0) \,+\, \frac{1}{2}\sum_{\vec{q}\neq\vec{0}} \Big(\frac{\hbar^2 q^2}{2mN} + \mathcal{U}(q)\Big)\Big(\hat{a}_{\vec{0}}\,\hat{a}^{\dagger}_{\vec{0}}\,\hat{a}^{\dagger}_{\vec{q}}\,\hat{a}_{\vec{q}} + \hat{a}_{\vec{0}}\,\hat{a}^{\dagger}_{\vec{0}}\,\hat{a}^{\dagger}_{-\vec{q}}\,\hat{a}_{-\vec{q}}\Big) \\ & + \,\frac{1}{2}\sum_{\vec{q}\neq\vec{0}} \mathcal{U}(q)\Big(\hat{a}_{\vec{0}}\,\hat{a}_{\vec{0}}\,\hat{a}^{\dagger}_{\vec{q}}\,\hat{a}^{\dagger}_{-\vec{q}} + \hat{a}^{\dagger}_{\vec{0}}\,\hat{a}^{\dagger}_{\vec{0}}\,\hat{a}_{-\vec{q}}\,\hat{a}_{\vec{q}}\Big) \end{aligned}$$

$$\equiv \frac{N(N-1)}{2}\,\mathcal{U}(0) + \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\Big(\frac{\hbar^2q^2}{2m} + N\mathcal{U}(q)\Big)\Big(\hat{b}^\dagger_{\vec{q}}\,\hat{b}_{\vec{q}} + \hat{b}^\dagger_{-\vec{q}}\,\hat{b}_{-\vec{q}}\Big)$$

$$+ \frac{1}{2}\sum_{\vec{q}\neq\vec{0}} N\mathcal{U}(q)\Big(\hat{b}^\dagger_{\vec{q}}\,\hat{b}^\dagger_{-\vec{q}} + \hat{b}_{-\vec{q}}\,\hat{b}_{\vec{q}}\Big)\,, \tag{100}$$

where the operators

$$\hat{b}^\dagger_{\vec{q}\neq\vec{0}} = \frac{\hat{a}_{\vec{0}}\,\hat{a}^\dagger_{\vec{q}}}{\sqrt{N}} = \frac{\hat{a}^\dagger_{\vec{q}}\,\hat{a}_{\vec{0}}}{\sqrt{N}} \qquad \text{and} \qquad \hat{b}_{\vec{q}\neq\vec{0}} = \frac{\hat{a}^\dagger_{\vec{0}}\,\hat{a}_{\vec{q}}}{\sqrt{N}} = \frac{\hat{a}_{\vec{q}}\,\hat{a}^\dagger_{\vec{0}}}{\sqrt{N}} \tag{101}$$

have a much clearer physical interpretation. By means of $\hat{b}^\dagger_{\vec{q}}$ a particle is excited from the 1-particle ground state to the 1-particle state with momentum $\hbar\vec{q} \neq \vec{0}$, whereas $\hat{b}_{\vec{q}}$ describes the de-excitation of an excited particle to the 1-particle ground state. The total number of particles is not affected in this way! However, the total Hamilton operator is identical to equation (99), with the non-additive character of $\hat{H}_{\text{tot}}$ simply following from the fact that the interaction can both excite pairs of particles from the 1-particle ground state and make them fall back to the ground state. As a result of the macroscopic occupation of the 1-particle ground state, the operators $\hat{b}^\dagger_{\vec{q}}$ and $\hat{b}_{\vec{q}}$ approximately behave as ordinary creation and annihilation operators:

$$\big[\hat{b}_{\vec{q}}, \hat{b}_{\vec{q}'}\big] = 0 \quad \text{and} \quad \big[\hat{b}_{\vec{q}}, \hat{b}^\dagger_{\vec{q}'}\big] = \frac{1}{N}\big[\hat{a}^\dagger_{\vec{0}}\,\hat{a}_{\vec{q}}, \hat{a}^\dagger_{\vec{q}'}\,\hat{a}_{\vec{0}}\big] = \big[\hat{a}_{\vec{q}}, \hat{a}^\dagger_{\vec{q}'}\big]\frac{\hat{n}_{\vec{0}}}{N} - \frac{\hat{a}^\dagger_{\vec{q}'}\,\hat{a}_{\vec{q}}}{N} \approx \big[\hat{a}_{\vec{q}}, \hat{a}^\dagger_{\vec{q}'}\big]\,.$$

More details can be found in the bachelor thesis by Leon Groenewegen.

> *In analogy with § 1.6.2 it is opportune to switch now to a quasi-particle description that turns the approximated interacting system into a non-interacting quasi-particle system, with a corresponding total Hamilton operator that is both additive and diagonal (i.e. decoupled).*

In this quasi-particle description we expect that we have to combine the operator pairs $\hat{b}^\dagger_{\vec{q}}$ and $\hat{b}_{-\vec{q}}$ as well as $\hat{b}_{\vec{q}}$ and $\hat{b}^\dagger_{-\vec{q}}$, since

- both operators within such a pair describe the same change in momentum and therefore affect the total momentum of the many-particle system in the same way;
- the quasi-particle number operators will then generate the correct terms $\hat{b}^\dagger_{\vec{q}}\,\hat{b}_{\vec{q}}$, $\hat{b}_{-\vec{q}}\,\hat{b}^\dagger_{-\vec{q}}$, $\hat{b}^\dagger_{\vec{q}}\,\hat{b}^\dagger_{-\vec{q}}$ and $\hat{b}_{-\vec{q}}\,\hat{b}_{\vec{q}}$.

How to find the correct quasi-particle description will be addressed in the next intermezzo.

1.6.5 Intermezzo: the Bogolyubov transformation for bosons

Consider a many-particle system consisting of identical bosons that can occupy two fully specified 1-particle quantum states $|q_1\rangle$ and $|q_2\rangle$, such as the momentum states $|\vec{q}\rangle$ and $|-\vec{q}\rangle$ in § 1.6.4. The corresponding creation and annihilation operators are given by $\hat{a}_1^\dagger, \hat{a}_2^\dagger$ and $\hat{a}_1, \hat{a}_2$. The corresponding Fock space is spanned by the basis states $|n_1, n_2\rangle$ as given in equation (28), where $n_{1,2}$ represents the number of identical particles in each of the two 1-particle quantum states. Next we consider a non-additive Hamilton operator of the form

$$\hat{H} \;=\; E(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2) \,+\, \Delta(\hat{a}_1^\dagger \hat{a}_2^\dagger + \hat{a}_2 \hat{a}_1) \qquad\qquad (E > 0 \;\text{ and }\; \Delta \in \mathbb{R}) \;. \tag{102}$$

Evidently, the total number of particles is not conserved by such an operator, as the occupation number of each of the two quantum states is raised by one or lowered by one in the Δ terms. However, by means of a so-called Bogolyubov transformation it can be cast into an additive form up to a constant term.[3] Such a transformation has the generic form

$$\hat{c}_1 \;\equiv\; u_1 \hat{a}_1 + v_1 \hat{a}_2^\dagger \qquad , \qquad \hat{c}_2 \;\equiv\; u_2 \hat{a}_2 + v_2 \hat{a}_1^\dagger \qquad\qquad (u_{1,2} \;\text{ and }\; v_{1,2} \in \mathbb{R}) \;. \tag{103}$$

The real constants u_1, u_2, v_1 and v_2 will be chosen in such a way that the operators $\hat{c}_{1,2}^\dagger$ and $\hat{c}_{1,2}$ satisfy the same bosonic commutator algebra as $\hat{a}_{1,2}^\dagger$ and $\hat{a}_{1,2}$. This will allow us to formulate a new particle interpretation, where $\hat{c}_{1,2}^\dagger$ and $\hat{c}_{1,2}$ describe the creation and annihilation of quasi particles. In order to guarantee that both $\hat{a}_{1,2}^\dagger, \hat{a}_{1,2}$ and $\hat{c}_{1,2}^\dagger, \hat{c}_{1,2}$ satisfy bosonic commutation relations, the following conditions should hold:

$$\boxed{\begin{array}{l} u_1 v_2 - v_1 u_2 \;=\; 0 \qquad \text{and} \qquad u_1^2 - v_1^2 \;=\; u_2^2 - v_2^2 \;=\; 1 \\ \Rightarrow \qquad u_1 = +\,u_2 \quad , \quad v_1 = +\,v_2 \qquad \text{and} \qquad u_1^2 - v_1^2 \;=\; 1 \\ \quad \text{or} \quad u_1 = -\,u_2 \quad , \quad v_1 = -\,v_2 \qquad \text{and} \qquad u_1^2 - v_1^2 \;=\; 1 \;. \end{array}} \tag{104}$$

Proof: the commutation relations $\big[\hat{c}_1, \hat{c}_1\big] = \big[\hat{c}_2, \hat{c}_2\big] = \big[\hat{c}_1, \hat{c}_2^\dagger\big] = 0$ follow directly from the bosonic commutation relations for $\hat{a}_{1,2}^\dagger$ and $\hat{a}_{1,2}$. The indicated conditions for $u_{1,2}$ and $v_{1,2}$ then simply follow from the fact that the other commutators

$$\big[\hat{c}_1, \hat{c}_2\big] \;\overset{(103)}{=\!=}\; \big[u_1 \hat{a}_1 + v_1 \hat{a}_2^\dagger, u_2 \hat{a}_2 + v_2 \hat{a}_1^\dagger\big] \;\overset{(26)}{=\!=}\; (u_1 v_2 - v_1 u_2)\, \hat{1} \;,$$

$$\big[\hat{c}_1, \hat{c}_1^\dagger\big] \;\overset{(103)}{=\!=}\; \big[u_1 \hat{a}_1 + v_1 \hat{a}_2^\dagger, u_1 \hat{a}_1^\dagger + v_1 \hat{a}_2\big] \;\overset{(26)}{=\!=}\; (u_1^2 - v_1^2)\, \hat{1} \;,$$

$$\big[\hat{c}_2, \hat{c}_2^\dagger\big] \;\overset{(103)}{=\!=}\; \big[u_2 \hat{a}_2 + v_2 \hat{a}_1^\dagger, u_2 \hat{a}_2^\dagger + v_2 \hat{a}_1\big] \;\overset{(26)}{=\!=}\; (u_2^2 - v_2^2)\, \hat{1} \;,$$

have to satisfy the usual bosonic commutation relations.

[3] If the states $|q_1\rangle$ and $|q_2\rangle$ carry opposite quantum numbers, such as the momentum in the states $|\vec{q}\rangle$ and $|-\vec{q}\rangle$ in § 1.6.4, then both $\hat{H}$ and the Bogolyubov transformation conserve these quantum numbers.

The sign in this transformation can be chosen freely. Usually one chooses the plus sign, resulting in the following generic form for a bosonic Bogolyubov transformation:

$$\boxed{\hat{c}_1 \;\equiv\; u_1\,\hat{a}_1 + v_1\,\hat{a}_2^\dagger \qquad , \qquad \hat{c}_2 \;\equiv\; u_1\,\hat{a}_2 + v_1\,\hat{a}_1^\dagger} \tag{105}$$

with inverse

$$\boxed{\hat{a}_1 \;=\; u_1\,\hat{c}_1 - v_1\,\hat{c}_2^\dagger \qquad , \qquad \hat{a}_2 \;=\; u_1\,\hat{c}_2 - v_1\,\hat{c}_1^\dagger}\;. \tag{106}$$

In the literature one usually opts for a parametrization of u_1 and v_1 in terms of a real parameter η according to $u_1 = \cosh\eta$ and $v_1 = \sinh\eta$, which automatically satisfies the condition $u_1^2 - v_1^2 = 1$.

Bringing the non-additive Hamilton operator of equation (102) in additive form.
To this end we consider two combinations of quasi-particle number operators. First of all

$$\begin{aligned}
\hat{c}_1^\dagger\,\hat{c}_1 - \hat{c}_2^\dagger\,\hat{c}_2 \;&\overset{(105)}{=\!=}\; (u_1\,\hat{a}_1^\dagger + v_1\,\hat{a}_2)(u_1\,\hat{a}_1 + v_1\,\hat{a}_2^\dagger) \;-\; (u_1\,\hat{a}_2^\dagger + v_1\,\hat{a}_1)(u_1\,\hat{a}_2 + v_1\,\hat{a}_1^\dagger) \\
&\overset{(26)}{=\!=}\; u_1^2(\hat{a}_1^\dagger\,\hat{a}_1 - \hat{a}_2^\dagger\,\hat{a}_2) \;-\; v_1^2(\hat{a}_1\,\hat{a}_1^\dagger - \hat{a}_2\,\hat{a}_2^\dagger) \\
&\overset{(26)}{=\!=}\; (u_1^2 - v_1^2)(\hat{a}_1^\dagger\,\hat{a}_1 - \hat{a}_2^\dagger\,\hat{a}_2) \;\overset{(104)}{=\!=}\; \hat{a}_1^\dagger\,\hat{a}_1 - \hat{a}_2^\dagger\,\hat{a}_2\;.
\end{aligned} \tag{107}$$

This expression tells us that certain quantum numbers are conserved under the transition from particles to quasi particles, provided that these quantum numbers take on opposite values in the states $|q_1\rangle$ and $|q_2\rangle$ (see the footnote on p. 38). In § 1.6.6 we will use this property to conserve the momentum quantum numbers and thereby the total momentum while performing the transformation, as anticipated in § 1.6.4. Secondly we have

$$\hat{c}_1^\dagger\,\hat{c}_1 + \hat{c}_2^\dagger\,\hat{c}_2 + \hat{1} \;=\; (u_1^2 + v_1^2)(\hat{a}_1^\dagger\,\hat{a}_1 + \hat{a}_2^\dagger\,\hat{a}_2 + \hat{1}) \;+\; 2\,u_1 v_1\,(\hat{a}_1^\dagger\,\hat{a}_2^\dagger + \hat{a}_2\,\hat{a}_1)\;. \tag{108}$$

This implies that for $|\Delta| < E$ the non-additive Hamilton operator $\hat{H}$ in equation (102) can be rewritten as

$$\begin{aligned}
\hat{H} \;&=\; E(\hat{a}_1^\dagger\,\hat{a}_1 + \hat{a}_2^\dagger\,\hat{a}_2 + \hat{1}) \;+\; \Delta(\hat{a}_1^\dagger\,\hat{a}_2^\dagger + \hat{a}_2\,\hat{a}_1) \;-\; E\,\hat{1} \\
&=\; \sqrt{E^2 - \Delta^2}\;(\hat{c}_1^\dagger\,\hat{c}_1 + \hat{c}_2^\dagger\,\hat{c}_2 + \hat{1}) \;-\; E\,\hat{1}\;.
\end{aligned} \tag{109}$$

Proof: based on equation (108) we are looking for a factor C such that $C(u_1^2 + v_1^2) = E$ and $2Cu_1v_1 = \Delta$. This implies that E and C should have the same sign and that

$$E^2 - \Delta^2 \;=\; C^2(u_1^2 - v_1^2)^2 \;\overset{(104)}{=\!=}\; C^2 \qquad \overset{E,C>0}{\Longrightarrow} \qquad |\Delta| < E \quad \text{and} \quad C \;=\; \sqrt{E^2 - \Delta^2}\;.$$

As promised at the start of this intermezzo, the Hamilton operator has been cast into a form that consists exclusively of a unit operator and number operators for quasi particles. In exercise 7 it will be shown that the corresponding ground state (i.e. the state without quasi-particle excitations) is comprised of coherently created pairs of particles. In § 1.7.2 the fermionic version of all this will be derived by employing a similar procedure.

1.6.6 Superfluidity for weakly repulsive spin-0 bosons (part 2)

The approximation (100) for the total Hamilton operator at the end of § 1.6.4 is exactly of the form discussed in the previous intermezzo. Therefore, an appropriate set of Bogolyubov transformations can be performed to recast the Hamilton operator in the additive form

$$\begin{aligned}\hat{H} \;&\approx\; \frac{N(N-1)}{2}\,\mathcal{U}(0) \;-\; \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\Big(\frac{\hbar^2q^2}{2m}+N\mathcal{U}(q)-\epsilon_{\vec{q}}\Big) \;+\; \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\epsilon_{\vec{q}}\,(\hat{c}^\dagger_{\vec{q}}\,\hat{c}_{\vec{q}}+\hat{c}^\dagger_{-\vec{q}}\,\hat{c}_{-\vec{q}}) \\ &=\; \frac{N(N-1)}{2}\,\mathcal{U}(0) \;-\; \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\Big(\frac{\hbar^2q^2}{2m}+N\mathcal{U}(q)-\epsilon_{\vec{q}}\Big) \;+\; \sum_{\vec{q}\neq\vec{0}}\epsilon_{\vec{q}}\,\hat{c}^\dagger_{\vec{q}}\,\hat{c}_{\vec{q}}\;, \qquad (110)\end{aligned}$$

with corresponding quasi-particle excitation spectrum

$$\boxed{\epsilon_{\vec{q}} \;=\; \epsilon_{-\vec{q}} \;=\; \frac{\hbar^2q^2}{2m}\sqrt{1+\frac{4mN\mathcal{U}(q)}{\hbar^2q^2}}}\;. \qquad (111)$$

For each pair of momenta $\vec{q}$ and $-\vec{q}$ with $\vec{q}\neq\vec{0}$ a bosonic Bogolyubov transformation of the type

$$\hat{c}_{\vec{q}} \;\equiv\; u_{\vec{q}}\,\hat{b}_{\vec{q}}+v_{\vec{q}}\,\hat{b}^\dagger_{-\vec{q}} \quad , \quad \hat{c}_{-\vec{q}} \;\equiv\; u_{\vec{q}}\,\hat{b}_{-\vec{q}}+v_{\vec{q}}\,\hat{b}^\dagger_{\vec{q}}$$

should be used, based on the following energy parameters in equations (102) and (109):

$$E \;\to\; \frac{1}{2}\Big(\frac{\hbar^2q^2}{2m}+N\mathcal{U}(q)\Big) \quad , \quad \Delta \;\to\; \frac{1}{2}\,N\mathcal{U}(q)\;.$$

For the total momentum operator we find with the help of approximation 2 on p. 37 that

$$\begin{aligned}\hat{\vec{P}}_{\rm tot} \;&\overset{(96)}{=\!=}\; \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\hbar\vec{q}\,\big(\hat{a}^\dagger_{\vec{q}}\,\hat{a}_{\vec{q}}-\hat{a}^\dagger_{-\vec{q}}\,\hat{a}_{-\vec{q}}\big) \;\approx\; \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\hbar\vec{q}\,\frac{\hat{a}_{\vec{0}}\,\hat{a}^\dagger_{\vec{0}}}{N}\,\big(\hat{a}^\dagger_{\vec{q}}\,\hat{a}_{\vec{q}}-\hat{a}^\dagger_{-\vec{q}}\,\hat{a}_{-\vec{q}}\big) \qquad (112) \\ &\overset{(101)}{=\!=}\; \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\hbar\vec{q}\,\big(\hat{b}^\dagger_{\vec{q}}\,\hat{b}_{\vec{q}}-\hat{b}^\dagger_{-\vec{q}}\,\hat{b}_{-\vec{q}}\big) \overset{(107)}{=\!=}\; \frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\hbar\vec{q}\,\big(\hat{c}^\dagger_{\vec{q}}\,\hat{c}_{\vec{q}}-\hat{c}^\dagger_{-\vec{q}}\,\hat{c}_{-\vec{q}}\big) \;=\; \sum_{\vec{q}\neq\vec{0}}\hbar\vec{q}\,\hat{c}^\dagger_{\vec{q}}\,\hat{c}_{\vec{q}}\;.\end{aligned}$$

Now we can read off the new particle interpretation. Quasi particles with energy $\epsilon_{\vec{q}}$ and momentum $\hbar\vec{q}$ are created by $\hat{c}^\dagger_{\vec{q}}$, annihilated by $\hat{c}_{\vec{q}}$ and counted by $\hat{c}^\dagger_{\vec{q}}\hat{c}_{\vec{q}}$. The ground state of the new many-particle system still contains no quanta with momentum $\hbar\vec{q}\neq\vec{0}$. However, both the composition of the ground state in terms of the original particles and the shape (dispersion relation) of the elementary excitation spectrum have changed as a result of the interaction.

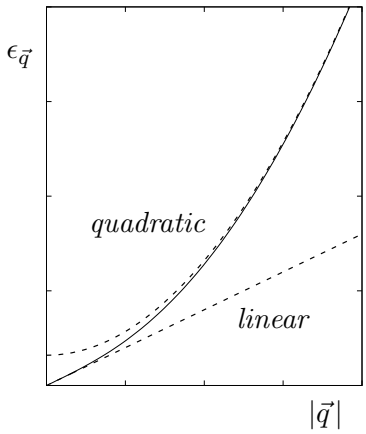
1) Approximated excitation spectrum for weakly repulsive spin-0 bosons:

- For large excitation energies $\hbar^2 q^2 \gg mN|\mathcal{U}(q)|$ the quasi-particle excitation spectrum $\epsilon_{\vec{q}} \approx \hbar^2 q^2/(2m) + N\mathcal{U}(q)$ hardly differs from the non-interacting spectrum. As such, the quasi particles have the same properties as the original bosons.

- For small excitation energies $\hbar^2 q^2 \ll mN\mathcal{U}(q)$ the quasi-particle excitation spectrum $\epsilon_{\vec{q}} \approx \hbar q\sqrt{N\mathcal{U}(q)/m} \approx \hbar q\sqrt{N\mathcal{U}(0)/m}$ changes substantially. The quasi particles describe massless quanta, i.e. quantized sound waves inside the considered medium with speed of propagation

$$c_s \;=\; \lim_{q\to 0}\frac{\epsilon_{\vec{q}}}{\hbar q} \;=\; \sqrt{\frac{N\mathcal{U}(0)}{m}} \qquad \text{for} \qquad \mathcal{U}(0) \;=\; \frac{1}{V}\int_V \mathrm{d}\vec{r}\,\mathcal{U}(r)\;. \tag{113}$$

The low-energy quasi-particle interpretation of the interacting system therefore differs fundamentally from the original particle interpretation of the non-interacting system. Moreover, we have gone from a free-particle system that does not allow for the possibility of superfluidity to an interacting system that can display superfluidity if $u < u_c \approx c_s$.



2) Approximated ground state for weakly repulsive spin-0 bosons: another thing that has changed substantially is the composition of the ground state of the interacting many-particle system in terms of the original particles (see exercise 7).

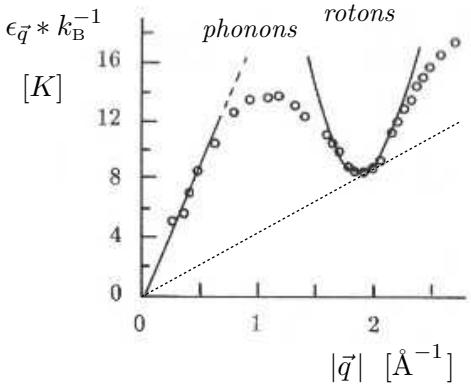
- In the non-interacting case all particles occupy the 1-particle ground state with momentum $\vec{0}$ and energy 0 (i.e. $n^{(0)}_{\vec{0}} = N$ and $n^{(0)}_{\vec{k}\neq\vec{0}} = 0$). This condensate has spatial correlations on all distance scales in view of the corresponding macroscopic de Broglie wavelength.

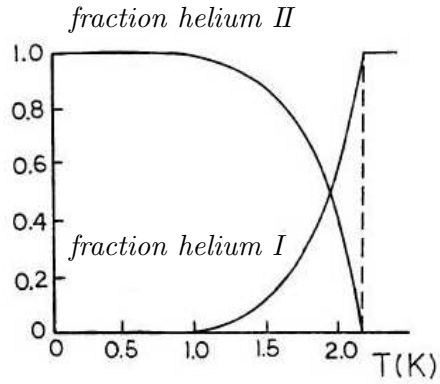
- The latter still holds in the interacting case. However, the particle excitations for $\vec{k} \neq \vec{0}$ are replaced by quasi-particle excitations with the same momentum. In the new ground state none of these quasi-particle excitations are occupied (i.e. $n_{\vec{k}=\vec{0}} \approx N$ and for the quasi particles $\tilde{n}_{\vec{k}\neq\vec{0}} = 0$). This situation without excited quasi particles differs markedly from the situation without excited particles. Written in terms of the original particle interpretation the new condensate actually contains particles that are not in the ground state. To be more precise, it contains coherently excited particle pairs with opposite momenta. The increase in kinetic energy is in that case compensated by the decrease in repulsive interaction energy.

Remark: if the average spatial pair interaction $\mathcal{U}(0)$ would have been attractive, i.e. if $\mathcal{U}(0) < 0$, then the ground state would have been unstable. This can be read off directly from the spectrum (111), which would become complex at very low energies. In that case it would be energetically favourable for the system to have a large number of low-energy particle pairs outside of the 1-particle ground state, which invalidates the assumptions of the previous approach. The fermionic version of such a pairing effect and the ensuing "pair-bonding" instability play a crucial role in the quantum mechanical description of superconductivity. A first exposure to this phenomenon can be found in exercise 9.

1.6.7 The wonderful world of superfluid ^{4}He: the two-fluid model

The many-particle system worked out above is often used to model the low-energy excitations in liquid ^{4}He, which becomes superfluid below $T_\lambda = 2.18\,\mathrm{K}$ (P.L. Kapitsa, J.F. Allen and A.D. Misener, 1937). At this point we have to add a critical note. The pair interactions cannot be truly weak, as required by the approximation method, as we are dealing with a fluid rather than a gas. As a consequence, a second branch of excitations enters the spectrum at higher energies, lowering the critical velocity for superfluidity.





At the absolute zero of temperature ^{4}He is completely superfluid, i.e. no low-energy excitations are excited thermally. At a temperature around 0.9 K a noticable influence of thermally induced excitations sets in. On the temperature interval $0 < T < T_\lambda$ there are effectively two fluids. On the one hand there is the excitation-free condensate. This fluid, referred to as helium II, is superfluid and carries no thermal energy. On the other hand there is the collection of thermally induced quasi-particle excitations. This fluid, referred to as helium I, carries the thermal energy and gives rise to friction. For $T > T_\lambda$ the influence of the helium II component is negligible. By means of this so-called two-fluid model a few surprising phenomena can be understood.

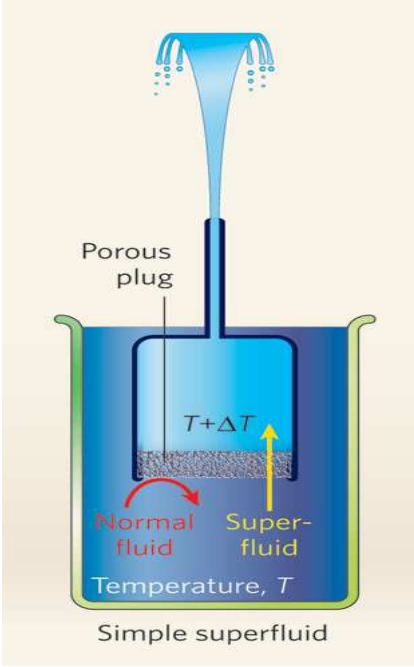
How to recognize T_λ (P.L. Kapitsa, 1937): a remarkable superfluid phenomenon is that the boiling of liquid helium abruptly subsides the moment that the temperature drops below T_λ. Helium I will flow away from any spot where the fluid is locally warmer, thereby carrying away thermal energy, whereas non-thermal helium II will flow towards that spot in

order to compensate for the drop in mass density. This superfluid helium II is characterized by an infinitely efficient heat conductance, making it effectively impossible to have a temperature gradient in the fluid. Already for temperatures just below T_λ the heat transport becomes a millionfold more efficient and gas bubbles will have no time to form.

Friction for $T < T_\lambda$: objects moving through ^{4}He at temperatures below T_λ experience friction exclusively from the helium I component and not from the helium II component. This allows us to experimentally determine the helium I and helium II fluid fractions (see the picture on the previous page). For temperatures below about 0.9 K the fluid behaves almost entirely as a superfluid, giving rise to a persistent current once the fluid is set in motion (e.g. at higher temperature).

Frictionless flow through porous media: as a result of friction the helium I component can not flow through very narrow capillary channels. However, the superfluid helium II component can do this without the need for a pressure difference between the two ends of such a channel. So, ^{4}He can flow through porous surfaces if $T < T_\lambda$. This flow will be completely frictionless. Such a situation where a selective superfluid flow occurs is called a super leak.

Fisher and Pickett, Nature **444**, *2006*



The fontain effect (J.F. Allen and H. Jones, 1938): consider an experimental set-up consisting of two containers with ^{4}He that are connected by means of a porous plug. Both containers are cooled down to the same temperature below T_λ. If the temperature in one of the containers is raised slightly, e.g. by shining a pocket torch on it, then the number of thermally induced excitations will increase in that container. As a result, the helium I component increases at the expense of the helium II component. In order to compensate for the difference in helium II concentration, helium II from the other container will pass through the porous plug. However, a compensatory reverse flow of helium I to the other container, where the helium I concentration is lower, is hampered by friction. Consequently, an increased fluid concentration will accumulate in the heated container (heat pump). By providing the heated container with a capillary safety-relief outlet, a spectacular helium fountain (fontain effect) is produced.

Creeping helium film: ^{4}He has as an additional property that the mutual van der Waals bonding is weaker than the van der Waals bonding to other atoms. Because of this a 30 nm thick 2-dimensional helium film (Rollin film) attaches itself to the entire wall of a closed helium container. If part of this helium film could flow or drip to a lower level within the container (see picture), then a superfluid helium II flow will start that terminates only when the helium level reaches its energetic optimum everywhere in the container. During this process the helium liquid is seemingly defying gravity.

In case you want to see some video evidence for the bold statements that were made about the weird and wacky world of superfluid ^{4}He, then you are advised to have a look at *http://www.youtube.com/watch?v=2Z6UJbwxBZI* or to do your own YouTube search.

1.7 Examples and applications: fermionic systems

1.7.1 Fermi sea and hole theory

Consider a fermionic many-particle system consisting of a very large, constant number N of electrons with mass m. The electrons are contained inside a large cube with edges L and periodic boundary conditions, giving rise to a discrete momentum spectrum:

$$\{\vec{p} = \hbar\vec{k} \,:\, k_{x,y,z} \;=\; 0, \pm 2\pi/L, \pm 4\pi/L, \cdots\} \;. \tag{114}$$

Without mutual interactions among the electrons the total kinetic-energy operator of the non-interacting identical-particle system can be written in diagonal form in the momentum representation:

$$\hat{T}_{\rm tot} \;=\; \sum_{\vec{k}} \sum_{m_s = \pm 1/2} \frac{\hbar^2 \vec{k}^{\,2}}{2m}\, \hat{a}^\dagger_{\vec{k},m_s} \hat{a}_{\vec{k},m_s} \;, \tag{115}$$

with kinetic energy eigenvalues

$$E \;=\; \sum_{\vec{k},m_s} \frac{\hbar^2 \vec{k}^{\,2}}{2m}\, n_{\vec{k},m_s} \;. \tag{116}$$

By adding the total momentum operator and total spin operator along the z-direction

$$\hat{\vec{P}}_{\rm tot} \;=\; \sum_{\vec{k},m_s} \hbar\vec{k}\, \hat{a}^\dagger_{\vec{k},m_s} \hat{a}_{\vec{k},m_s} \qquad \text{and} \qquad \hat{\vec{S}}_{\rm tot} \cdot \vec{e}_z \;=\; \sum_{\vec{k},m_s} m_s \hbar\, \hat{a}^\dagger_{\vec{k},m_s} \hat{a}_{\vec{k},m_s} \tag{117}$$

we can trivially read off the particle interpretation belonging to the creation and annihilation operators used. Particles with kinetic energy $\hbar^2\vec{k}^{\,2}/(2m)$, momentum $\hbar\vec{k}$ and spin component $m_s\hbar$ along the z-axis are created by $\hat{a}^\dagger_{\vec{k},m_s}$, annihilated by $\hat{a}_{\vec{k},m_s}$ and counted by $\hat{n}_{\vec{k},m_s} = \hat{a}^\dagger_{\vec{k},m_s}\hat{a}_{\vec{k},m_s}$. The occupation number $n_{\vec{k},m_s}$ indicates how many particles can be found in the given momentum and spin eigenstate in the absence of mutual interactions.

Ground state: for the ground state of the non-interacting N-electron system this automatically implies that $n_{\vec{k},m_s}=1$ if $|\vec{k}| \le k_F$ and $n_{\vec{k},m_s}=0$ if $|\vec{k}| > k_F$. Such a system is called a completely degenerate electron gas. The fully specified 1-particle energy levels are occupied from the bottom up by one electron up to the Fermi-energie $E_F = \hbar^2 k_F^2/(2m)$. In that case

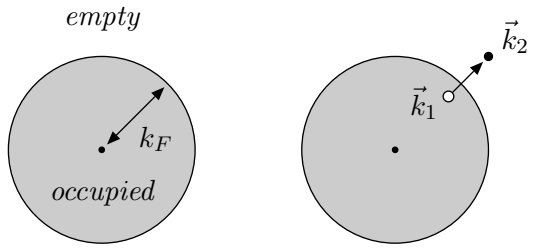
$$N = \sum_{\vec{k},m_s} n_{\vec{k},m_s} = 2\sum_{|\vec{k}|\le k_F} \quad \text{and} \quad E_{\text{ground}} = 2\sum_{|\vec{k}|\le k_F} \frac{\hbar^2\vec{k}^{\,2}}{2m} \ . \tag{118}$$

The corresponding ground state is given by

$$|\Psi_{\text{ground}}\rangle = \prod_{|\vec{k}|\le k_F} \hat{a}^\dagger_{\vec{k},\frac{1}{2}}\hat{a}^\dagger_{\vec{k},-\frac{1}{2}}|\Psi^{(0)}\rangle \quad \Rightarrow \quad \hat{a}_{\vec{k},m_s}|\Psi_{\text{ground}}\rangle = 0 \quad \text{if} \quad |\vec{k}| > k_F \ . \tag{119}$$

This ground state, better known as the Fermi sea, has vanishing total momentum and spin, but a substantial kinetic energy. Even in the ground state the electron gas exerts pressure (see Ch. 2)!

Excited states: the simplest excited states can be obtained by exciting a single electron in the Fermi sea to an energy that places it above the Fermi sea (see picture). As the states inside the Fermi sea are all occupied, excitations are (energetically) easier from energy levels near the edge of the Fermi sea than from lower-lying energy levels. This aspect plays a decisive role in the thermal properties of fermionic many-particle systems at moderate temperatures (see Ch. 2). An excitation of this type can be represented in creation and annihilation language in the following way:



ground state: full Fermi sea — *electron–hole excitation*

$$|\Psi_{\text{ex}}\rangle = \hat{a}^\dagger_{\vec{k}_2,m_{s_2}}\hat{a}_{\vec{k}_1,m_{s_1}}|\Psi_{\text{ground}}\rangle \qquad \text{if} \qquad |\vec{k}_1| \le k_F < |\vec{k}_2| \ . \tag{120}$$

Since the excitation creates a hole in the Fermi sea, it also goes under the name of electron–hole excitation. In the quantum mechanical approach of hole theory this aspect

is highlighted by employing the Bogolyubov transformations

$$\hat{a}_{\vec{k},m_s} \equiv \hat{c}^{\dagger}_{-\vec{k},-m_s} \quad \text{if} \quad |\vec{k}| \leq k_F \qquad \text{and} \qquad \hat{a}^{\dagger}_{\vec{k},m_s} \equiv \hat{c}^{\dagger}_{\vec{k},m_s} \quad \text{if} \quad |\vec{k}| > k_F \tag{121}$$

to switch to a description where a hole in the Fermi sea is upgraded to the status of quasi particle with opposite quantum numbers. In that case the Fermi sea constitutes a state without any holes in the sea or excitations above it, i.e. the Fermi sea is the corresponding quasi-particle vacuum state! An excited state of the above type simply boils down to quasi-particle pair creation. The hole-theory approach is particularly useful for the description of electrical conduction and for Dirac's attempt to set up relativistic QM.

> *This approach is only possible for fermions: we have actually made use of Pauli's exclusion principle, which states that* $n_{\vec{k},m_s} \in [0,1]$*, in order to interchange the role of empty and occupied in the particle interpretation of hole theory. The advantage of doing this is that the Fermi sea (ground state) is the vacuum state within this alternative particle interpretation and that an excited state corresponds to the creation of a quasi-particle pair. In the hole-theory approach conservation of the number of (quasi) particles is sacrificed deliberately for achieving ease of use!*

Weak interactions among the electrons: for weakly repulsive interactions we can in general use perturbation theory, utilizing the above-mentioned non-interacting situation as unperturbed starting point. According to hole theory the Fermi sea can be treated as vacuum state. For weakly attractive interactions the situation changes substantially. In exercise 9 it will be shown how the presence of the Fermi sea of occupied states can give rise to the so-called Cooper instability. This entails that two electrons that occupy states above a Fermi sea of occupied states will always form a "bound" pair with finite binding energy as long as their mutual interaction is attractive. In contrast to a 1-particle bound state in a potential well, this many-particle binding effect occurs irrespective of the strength of the interaction. Moreover, it cannot be described by means of perturbation theory.

1.7.2 The Bogolyubov transformation for fermions

Also a fermionic version of the Bogolyubov transformation of § 1.6.5 exists. To this end we consider exactly the same 2-level system as described on page 38. Only after equation (103) things will start to deviate. So, also here the transformation will be defined by $\hat{c}_1 = u_1 \hat{a}_1 + v_1 \hat{a}^{\dagger}_2$ and $\hat{c}_2 = u_2 \hat{a}_2 + v_2 \hat{a}^{\dagger}_1$.

In contrast to the bosonic case, both $\hat{a}^{\dagger}_{1,2}, \hat{a}_{1,2}$ and $\hat{c}^{\dagger}_{1,2}, \hat{c}_{1,2}$ have to satisfy fermionic anticommutation relations. That implies the following conditions for the real constants

$u_{1,2}$ and $v_{1,2}$:

$$\boxed{\begin{array}{l} u_1 v_2 + v_1 u_2 \;=\; 0 \qquad \text{and} \qquad u_1^2 + v_1^2 \;=\; u_2^2 + v_2^2 \;=\; 1 \\ \Rightarrow \qquad u_1 = +\,u_2 \quad , \quad v_1 = -\,v_2 \qquad \text{and} \qquad u_1^2 + v_1^2 \;=\; 1 \\ \quad \text{or} \quad u_1 = -\,u_2 \quad , \quad v_1 = +\,v_2 \qquad \text{and} \qquad u_1^2 + v_1^2 \;=\; 1\;. \end{array}} \tag{122}$$

Proof: the anticommutation relations $\{\hat{c}_1, \hat{c}_1\} = \{\hat{c}_2, \hat{c}_2\} = \{\hat{c}_1, \hat{c}_2^\dagger\} = 0$ follow directly from the fermionic anticommutation relations for $\hat{a}_{1,2}^\dagger$ and $\hat{a}_{1,2}$. The indicated conditions for $u_{1,2}$ and $v_{1,2}$ then simply follow from the fact that the other anticommutators

$$\{\hat{c}_1, \hat{c}_2\} \overset{(103)}{=\!=} \{u_1\hat{a}_1 + v_1\hat{a}_2^\dagger, u_2\hat{a}_2 + v_2\hat{a}_1^\dagger\} \overset{(27)}{=\!=} (u_1 v_2 + v_1 u_2)\hat{1}\;,$$

$$\{\hat{c}_1, \hat{c}_1^\dagger\} \overset{(103)}{=\!=} \{u_1\hat{a}_1 + v_1\hat{a}_2^\dagger, u_1\hat{a}_1^\dagger + v_1\hat{a}_2\} \overset{(27)}{=\!=} (u_1^2 + v_1^2)\hat{1}\;,$$

$$\{\hat{c}_2, \hat{c}_2^\dagger\} \overset{(103)}{=\!=} \{u_2\hat{a}_2 + v_2\hat{a}_1^\dagger, u_2\hat{a}_2^\dagger + v_2\hat{a}_1\} \overset{(27)}{=\!=} (u_2^2 + v_2^2)\hat{1}\;,$$

have to satisfy the usual fermionic anticommutation relations. The sign in this transformation can be chosen freely. Usually one chooses the upper sign (i.e. $u_1 = u_2$ and $v_1 = -\,v_2$), resulting in the following generic form for a fermionic Bogolyubov transformation:

$$\boxed{\hat{c}_1 \;\equiv\; u_1\hat{a}_1 + v_1\hat{a}_2^\dagger \qquad , \qquad \hat{c}_2 \;\equiv\; u_1\hat{a}_2 - v_1\hat{a}_1^\dagger} \tag{123}$$

with inverse

$$\boxed{\hat{a}_1 \;=\; u_1\hat{c}_1 - v_1\hat{c}_2^\dagger \qquad , \qquad \hat{a}_2 \;=\; u_1\hat{c}_2 + v_1\hat{c}_1^\dagger}\;. \tag{124}$$

In the literature one usually opts for a parametrization of u_1 and v_1 in terms of a real parameter θ according to $u_1 = \cos\theta$ and $v_1 = \sin\theta$, which automatically satisfies the condition $u_1^2 + v_1^2 = 1$.

Special case: if $u_1 = u_2 = 0$, then the signs of both v_1 and v_2 can be chosen independently. Usually one will choose $v_1 = v_2 = 1$, as we did in § 1.7.1 while describing the quasi particles in hole theory.

The fermionic quasi-particle vacuum: removing a particle from a particular state is in the fermionic case equivalent with the creation of a hole in the occupation of that state. The type of quasi particle we are dealing with here is therefore a linear combination of a particle and a hole. This type of quasi particle is used abundantly in condensed-matter physics for describing systems of interacting fermions such as conduction electrons. In exercise 8 the fermionic quasi-particle vacuum $|\tilde{0}\rangle$ will be expressed in terms of the original basis states

$|n_1, n_2\rangle$, resulting in the compact expression

$$\boxed{|\tilde{0}\rangle \overset{\text{up to a phase factor}}{=\!=\!=\!=\!=\!=\!=\!=\!=\!=\!=} u_1|0,0\rangle - v_1|1,1\rangle = \big(u_1 - v_1\hat{a}_1^\dagger\hat{a}_2^\dagger\big)|0,0\rangle} \ . \tag{125}$$

Bringing the non-additive Hamilton operator of equation (102) in additive form.

To this end we consider two combinations of quasi-particle number operators. First of all

$$\begin{aligned}
\hat{c}_1^\dagger\hat{c}_1 - \hat{c}_2^\dagger\hat{c}_2 &\overset{(123)}{=\!=\!=} (u_1\hat{a}_1^\dagger + v_1\hat{a}_2)(u_1\hat{a}_1 + v_1\hat{a}_2^\dagger) \ - \ (u_1\hat{a}_2^\dagger - v_1\hat{a}_1)(u_1\hat{a}_2 - v_1\hat{a}_1^\dagger) \\
&\overset{(27)}{=\!=\!=} u_1^2(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2) \ - \ v_1^2(\hat{a}_1\hat{a}_1^\dagger - \hat{a}_2\hat{a}_2^\dagger) \\
&\overset{(27)}{=\!=\!=} (u_1^2 + v_1^2)(\hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2) \overset{(122)}{=\!=\!=} \hat{a}_1^\dagger\hat{a}_1 - \hat{a}_2^\dagger\hat{a}_2 \ .
\end{aligned} \tag{126}$$

This expression tells us again that certain quantum numbers are conserved under the transition from particles to quasi particles, provided that these quantum numbers take on opposite values in the states $|q_1\rangle$ and $|q_2\rangle$.

Secondly we have

$$\hat{c}_1^\dagger\hat{c}_1 + \hat{c}_2^\dagger\hat{c}_2 - \hat{1} \ = \ (u_1^2 - v_1^2)(\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2 - \hat{1}) \ + \ 2u_1v_1(\hat{a}_1^\dagger\hat{a}_2^\dagger + \hat{a}_2\hat{a}_1) \ . \tag{127}$$

The non-additive Hamilton operator $\hat{H}$ in equation (102) can then be rewritten as

$$\begin{aligned}
\hat{H} &= E(\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2 - \hat{1}) \ + \ \Delta(\hat{a}_1^\dagger\hat{a}_2^\dagger + \hat{a}_2\hat{a}_1) \ + \ E\,\hat{1} \\
&= \pm\sqrt{E^2+\Delta^2}\,(\hat{c}_1^\dagger\hat{c}_1 + \hat{c}_2^\dagger\hat{c}_2 - \hat{1}) \ + \ E\,\hat{1} \ .
\end{aligned} \tag{128}$$

Proof: based on equation (127) we are looking for a factor C such that $C(u_1^2 - v_1^2) = E$ and $2Cu_1v_1 = \Delta$. From this it follows that $E^2 + \Delta^2 = C^2(u_1^2 + v_1^2)^2 \overset{(122)}{=\!=\!=} C^2$, where the sign of C is determined by the sign of $(u_1^2 - v_1^2)/E$.

As desired, the Hamilton operator has been cast into a form that consists exclusively of a unit operator and number operators for quasi particles. For the description of electron holes in § 1.7.1 we have actually used a special version of this where $\hat{c}_1 = \hat{a}_2^\dagger$ en $\hat{c}_2 = \hat{a}_1^\dagger$. In that case $E(\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2) = -E(\hat{a}_1\hat{a}_1^\dagger + \hat{a}_2\hat{a}_2^\dagger - 2) = -E(\hat{c}_2^\dagger\hat{c}_2 + \hat{c}_1^\dagger\hat{c}_1 - 2)$. From the perspective of the Fermi energy, the creation of a hole in the Fermi sea results in a negative energy contribution. As noted in the proof given above, the sign of the energy of the quasi particle is fixed by the sign of $(u_1^2 - v_1^2)/E$, which indeed indicates whether the considered type of quasi particle is more particle than hole (positive sign) or more hole than particle (negative sign).

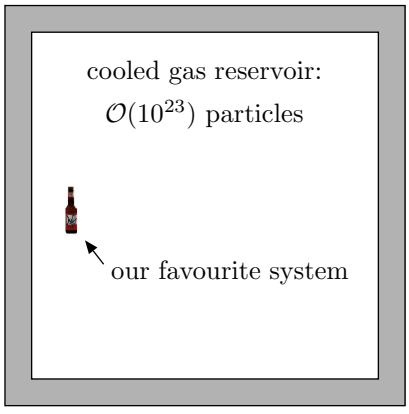
2 Quantum statistics

In this chapter the concept of mixed quantum mechanical ensembles will be introduced. Based on this concept the various types of quantum statistics will be derived, which govern many-particle systems that are in thermal equilibrium.

Similar material can be found in Griffiths (the quantum statistics part in Ch. 5), Merzbacher (Ch. 15, 16, 22) and Bransden & Joachain (Ch. 10, 14).

Up to now we have been working with pure quantum mechanical ensembles belonging to a pure quantum state, i.e. a collection of identical, independent, identically prepared systems that can be described by a single state function $|\psi\rangle$. This state function can be determined up to a phase by performing a complete measurement on a complete set of commuting observables. The systems that make up the ensemble are independent of each other and can consist of individual particles, interacting particle clusters (such as atoms/molecules) or even entire many-particle systems (such as gases). These pure quantum states can be realized experimentally by employing dedicated filters, such as a Stern–Gerlach filter for selecting a pure spin state.

Question: *What happens if not all the degrees of freedom that are relevant for the considered systems are taken into account, thereby leaving out certain quantum information?*



This situation actually reflects the unavoidable reality of the quantum mechanical world that we experience. During a quantum experiment the studied system is for instance prepared by filters and subsequently subjected to the environment of the experimental set-up (gases, electromagnetic fields, etc.). All these environmental influences caused by the contact with the macroscopic outside world in fact determine the properties of the quantum mechanical ensemble that eventually will be measured. Examples of this are quantum systems that have their themodynamical properties fixed by the contact with a macroscopic gas reservoir present within the measurement apparatus, or a particle beam that is produced by a particle source and subsequently passes through a set of filters. It is not feasible to determine the state function of the entire experimental apparatus by means of a complete measurement. The macroscopic outside world simply has way too many degrees of freedom. In such situations the only practical approach is to explicitly integrate out (read: leave out) the degrees of freedom of the outside world, thereby reducing the quantum mechanical description to the space spanned by the degrees of free-

dom of the studied system itself. Note: sometimes it is not really necessary to integrate out certain degrees of freedom, but it simply happens to be more convenient to do so. An example is the polarization of a particle beam, for which it is convenient not to have to deal with the spatial degrees of freedom (see § 2.2).

Integrating out degrees of freedom: to get an idea of what happens when integrating out quantum mechanical degrees of freedom, we consider a spin-1/2 system described by the following normalized pure state function in the position representation:

$$\chi(\vec{r}) \;=\; \begin{pmatrix} \psi_+(\vec{r}) \\ \psi_-(\vec{r}) \end{pmatrix} \;=\; \psi_+(\vec{r})\,\chi_{\frac{1}{2},\frac{1}{2}} \;+\; \psi_-(\vec{r})\,\chi_{\frac{1}{2},-\frac{1}{2}} \;,$$

$$\text{with} \qquad \int \mathrm{d}\vec{r}\,\psi^*_\lambda(\vec{r})\psi_{\lambda'}(\vec{r}) \;=\; W_\lambda\,\delta_{\lambda\lambda'} \qquad (\lambda,\lambda' = \pm) \qquad \text{and} \qquad W_+ + W_- = 1\;.$$

For convenience we have taken $\psi_\pm(\vec{r})$ to be orthogonal. Next we integrate out the spatial degrees of freedom in order to reduce the quantum mechanical description to the two-dimensional spin space. Consider to this end an arbitrary spin observable $\hat{A}$, which acts as a 2×2 matrix A in spin space and which has no action in position space. The corresponding expectation value for the state $\chi(\vec{r})$ then reads

$$\langle\hat{A}\rangle \;=\; \int \mathrm{d}\vec{r}\,\chi^\dagger(\vec{r})\,A\,\chi(\vec{r}) \;=\; \mathrm{Tr}\Big(A\int \mathrm{d}\vec{r}\,\chi(\vec{r})\chi^\dagger(\vec{r})\Big) \;\equiv\; \mathrm{Tr}(A\rho)\;,$$

$$\text{with} \quad \rho \;=\; \overbrace{\int \mathrm{d}\vec{r}\,\langle\vec{r}|\chi\rangle\,\langle\chi|\vec{r}\rangle}^{\vec{r}\text{ integrated out}} \;=\; \begin{pmatrix} W_+ & 0 \\ 0 & W_- \end{pmatrix} \;=\; \underline{\text{density matrix}} \text{ in spin space}$$

$$\text{and} \quad \mathrm{Tr} \;=\; \text{trace in spin space}\;.$$

Here we used that A commutes with $\vec{r}$ and we exploited the fact that for D-dimensional vectors v and $D\times D$ matrices M the following holds:

$$\sum_{i,j=1}^{D} v_i^* M_{ij}\,v_j \;=\; \sum_{i,j=1}^{D} M_{ij}\,(v_j v_i^*) \;\equiv\; \sum_{i,j=1}^{D} M_{ij}W_{ji} \;=\; \mathrm{Tr}(MW)\;.$$

The expectation value for the complete system splits up according to

$$\langle\hat{A}\rangle \;=\; W_+\,\mathrm{Tr}\big(A\,\chi_{\frac{1}{2},\frac{1}{2}}\chi^\dagger_{\frac{1}{2},\frac{1}{2}}\big) \;+\; W_-\,\mathrm{Tr}\big(A\,\chi_{\frac{1}{2},-\frac{1}{2}}\chi^\dagger_{\frac{1}{2},-\frac{1}{2}}\big) \;=\; W_+\langle\hat{A}\rangle_+ \;+\; W_-\langle\hat{A}\rangle_-$$

into two independent expectation values, one for the pure spin state $\chi_{\frac{1}{2},\frac{1}{2}}$ with weight $W_+ \geq 0$ and one for the pure spin state $\chi_{\frac{1}{2},-\frac{1}{2}}$ with weight $W_- = 1 - W_+ \geq 0$. The

quantum mechanical description in spin space is now determined by the density matrix ρ, which is composed of statistical weights and projection operators on pure states:

$$\rho \;=\; W_+ \chi_{\frac{1}{2},\frac{1}{2}}\chi^\dagger_{\frac{1}{2},\frac{1}{2}} \;+\; W_- \chi_{\frac{1}{2},-\frac{1}{2}}\chi^\dagger_{\frac{1}{2},-\frac{1}{2}} \;=\; W_+ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \; W_- \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Mixed ensembles: integrating out quantum mechanical degrees of freedom gives rise to so-called mixed ensembles. These ensembles are incoherent mixtures (i.e. statistical mixtures) of pure subensembles, implying double statistics:

- on the one hand the pure quantum mechanical state functions that describe the pure subensembles have a statistical interpretation;
- on the other hand the complete mixed ensemble comprises of a statistical mixture of such subensembles.

The quantum physics in such scenarios is determined by the density operator, which contains the relevant properties of the mixed ensemble. These relevant properties are the pure quantum mechanical state functions that describe the pure subensembles and the corresponding statistical weights.

> *You might be inclined to think that a mixed ensemble somehow corresponds to a kind of superposition of pure state functions. As will become clear from the example given below, this is not the case.*

Mixed ensembles vs pure ensembles: consider the previous example with equal statistical weights $W_\pm = 0.5$, resulting in a statistical mixture with density matrix $\rho = \frac{1}{2}I$ for which 50% of the particles are in the spin state $\chi_{\frac{1}{2},\frac{1}{2}}$ and 50% in the spin state $\chi_{\frac{1}{2},-\frac{1}{2}}$. Such an incoherent mixture of pure spin states $\chi_{\frac{1}{2},\pm\frac{1}{2}}$ cannot be represented as a linear combination of states such as $(\chi_{\frac{1}{2},\frac{1}{2}} + \chi_{\frac{1}{2},-\frac{1}{2}})/\sqrt{2}$, in spite of the fact that such a pure state indeed corresponds to a 50% chance to measure the particles in each of the two spin states. However, other measurable quantities exist that *will* result in completely different measurement results in both situations. For instance, the state $(\chi_{\frac{1}{2},\frac{1}{2}} + \chi_{\frac{1}{2},-\frac{1}{2}})/\sqrt{2}$ is a pure eigenstate of the spin operator $\hat{S}_x$ belonging to the eigenvalue $\hbar/2$. Upon measuring the spin in the x-direction the outcome of the measurement will be fixed: 100% of the particles will correspond to the measurement value $\hbar/2$. For the mixed ensemble, however, we have

$$\langle \hat{S}_x \rangle \;=\; \frac{\hbar}{2}\,\mathrm{Tr}(\sigma_x \rho) \;\overset{\hat{\rho} = \frac{1}{2}\hat{1}}{=\!=\!=} \; \frac{\hbar}{4}\,\mathrm{Tr}(\sigma_x) \;=\; 0 \;\neq\; \frac{\hbar}{2}\,,$$

since none of the two types of particles on which the measurement is performed actually is in an eigenstate of $\hat{S}_x$. An incoherent mixture in fact implies that the phases of

the separate components in the mixture are unrelated, which excludes a linear combination of pure state functions. The ensemble splits up in *independent subensembles*, each with a specific spin state and statistical weight.

Density operators find applications in many branches of physics. In this chapter we will encounter a few characteristic examples from scattering theory and statistical physics (with applications in solid-state physics, nuclear physics, astrophysics and low-temperature physics). Density operators are also the tools of choice for the description of the state-function collapse of a microscopic system as caused by the influence of a macroscopic piece of measurement equipment.

2.1 The density operator (J. von Neumann, 1927)

Consider an ensemble consisting of a statistical mixture of N independent subensembles, each described by a normalized pure state function

$$|\psi^{(\alpha)}\rangle \;\equiv\; |\alpha\rangle \;, \qquad \text{with} \quad \langle\alpha|\alpha\rangle \;=\; 1 \qquad (\alpha = 1, \cdots, N) \;. \tag{129}$$

Each subensemble is a collection of identical, independent, identically prepared systems described by the corresponding state function $|\alpha\rangle$. The independent systems represented by the various state functions $|\alpha\rangle$ are of identical composition, but they differ by the state they are in. There is no need for these states to be eigenstates of one and the same observable, so that in general $\langle\alpha|\alpha'\rangle \neq 0$ if $\alpha \neq \alpha'$. An example could be an electron beam of which 40% of the particles are polarized along the z-axis, 30% along the y-axis and 30% along the x-axis. This beam is composed of three independent subensembles that are each described by a pure spin function. The three types of systems described by these pure spin functions all consist of one electron, but the spin is quantized in a different direction in each of the three cases.

From expectation value to ensemble average: take $\{|n\rangle\}$ to be an orthonormal set of eigenstates belonging to a complete set of commuting observables of the considered type of system. For discrete values of n we then have

$$\langle n|n'\rangle \;=\; \delta_{nn'} \qquad \text{and} \qquad \sum_n |n\rangle\langle n| \;=\; \hat{1} \;. \tag{130}$$

For n-values belonging to the continuous part of the spectrum, as usual $\delta_{nn'}$ should be replaced by $\delta(n-n')$ and the sum by an integral. Next we can decompose the normalized pure state functions $|\alpha\rangle$ in terms of this basis according to

$$|\alpha\rangle \;\overset{(130)}{=\!=}\; \sum_n |n\rangle\langle n|\alpha\rangle \;\equiv\; \sum_n c_n^{(\alpha)}\,|n\rangle \;, \qquad \text{with} \quad \langle\alpha|\alpha\rangle \;=\; \sum_n |c_n^{(\alpha)}|^2 \;=\; 1 \;. \tag{131}$$

Let $\hat{A}$ be an observable pertaining to the considered type of system, with corresponding (classical) dynamical variable $\mathcal{A}$. For each of the normalized states $|\alpha\rangle$ the expectation value (= quantum mechanical average) of the dynamical variable $\mathcal{A}$ is given by

$$\langle \hat{A} \rangle_\alpha \equiv \langle \alpha | \hat{A} | \alpha \rangle \stackrel{(130)}{=\!=} \sum_{n,n'} \langle \alpha | n' \rangle \langle n' | \hat{A} | n \rangle \langle n | \alpha \rangle \stackrel{(131)}{=\!=} \sum_{n,n'} c_n^{(\alpha)} c_{n'}^{(\alpha)\,*} \langle n' | \hat{A} | n \rangle \ . \quad (132)$$

Subsequently we take into account that we are dealing with a statistical mixture of pure subensembles. To this end we indicate the statistical weight of the pure state $|\alpha\rangle$ in the ensemble by W_α, corresponding to the fraction of systems in the state $|\alpha\rangle$. In short, the factor W_α is the probability for the system to be in the state $|\alpha\rangle$. The ensemble average (= statistical average) $[\hat{A}]$ of the dynamical variable $\mathcal{A}$ can be expressed in a representation independent way as

$$\boxed{[\hat{A}] = \sum_{\alpha=1}^{N} W_\alpha \langle \hat{A} \rangle_\alpha \ , \qquad \text{with} \quad W_\alpha \in [0,1] \quad \text{and} \quad \sum_{\alpha=1}^{N} W_\alpha = 1} \ . \quad (133)$$

Definition of the density operator: by means of equation (132) the expression for the ensemble average can be rewritten as

$$[\hat{A}] = \sum_{\alpha=1}^{N} \sum_{n,n'} W_\alpha \langle n | \alpha \rangle \langle \alpha | n' \rangle \langle n' | \hat{A} | n \rangle \equiv \sum_{n,n'} \langle n | \hat{\rho} | n' \rangle \langle n' | \hat{A} | n \rangle \ , \quad (134)$$

where the density operator $\hat{\rho}$ of the ensemble is composed of projection operators on the pure states:

$$\boxed{\hat{\rho} = \sum_{\alpha=1}^{N} W_\alpha |\alpha\rangle\langle\alpha|} \ . \quad (135)$$

Since $W_\alpha \in \mathbb{R}$, this density operator is hermitian and therefore

$$\underset{\psi_1,\psi_2}{\forall} \langle \psi_2 | \hat{\rho} | \psi_1 \rangle^* \stackrel{(135)}{=\!=} \sum_{\alpha=1}^{N} W_\alpha \langle \psi_2 | \alpha \rangle^* \langle \alpha | \psi_1 \rangle^* = \sum_{\alpha=1}^{N} W_\alpha \langle \psi_1 | \alpha \rangle \langle \alpha | \psi_2 \rangle = \langle \psi_1 | \hat{\rho} | \psi_2 \rangle \ .$$

In the n-representation the density operator is characterized by the so-called density matrix

$$\rho_{nn'} = \langle n | \hat{\rho} | n' \rangle \stackrel{(131),(135)}{=\!=\!=} \sum_{\alpha=1}^{N} W_\alpha c_n^{(\alpha)} c_{n'}^{(\alpha)\,*} \ . \quad (136)$$

In matrix language the ensemble average of the dynamical variable $\mathcal{A}$ simply reads

$$[\hat{A}] \stackrel{(134)}{=\!=} \sum_{n,n'} \langle n | \hat{\rho} | n' \rangle \langle n' | \hat{A} | n \rangle \stackrel{(130)}{=\!=} \sum_{n} \langle n | \hat{\rho} \hat{A} | n \rangle = \boxed{\mathrm{Tr}(\hat{\rho}\hat{A}) = [\hat{A}]} \ . \quad (137)$$

Properties of the density operator:

- Since the density operator $\hat{\rho}$ is hermitian, an orthonormal basis exists of density-operator eigenstates $\{|k\rangle\}$ with real eigenvalues $\{\rho_k\}$. Such an eigenvalue

$$\rho_k \;=\; \langle k|\hat{\rho}|k\rangle \;\stackrel{(136)}{=\!=}\; \sum_{\alpha=1}^{N} W_\alpha\, |c_k^{(\alpha)}|^2 \;\in\; [0,1] \tag{138}$$

  can be interpreted as the probability for the system to be in the pure state $|k\rangle$.

- Conservation of probability implies

$$\boxed{\mathrm{Tr}(\hat{\rho}) \;=\; 1}\,. \tag{139}$$

  Proof: $\mathrm{Tr}(\hat{\rho}) \stackrel{(137)}{=\!=} [\hat{1}] \stackrel{(133)}{=\!=} \sum_{\alpha=1}^{N} W_\alpha \langle\hat{1}\rangle_\alpha \stackrel{\langle\alpha|\alpha\rangle=1}{=\!=\!=} \sum_{\alpha=1}^{N} W_\alpha \stackrel{(133)}{=\!=} 1$.

- For arbitrary ensembles we have

$$\boxed{\mathrm{Tr}(\hat{\rho}^2) \;\leq\; 1}\,. \tag{140}$$

  Proof: in terms of the eigenvalues $\rho_k \in [0,1]$ of $\hat{\rho}$ one readily finds

$$[\hat{\rho}] \stackrel{(137)}{=\!=} \mathrm{Tr}(\hat{\rho}^2) \;=\; \sum_k \rho_k^2 \;\leq\; \sum_k \rho_k \;=\; \mathrm{Tr}(\hat{\rho}) \stackrel{(139)}{=\!=} 1\,.$$

- The following criterion holds for a system in a pure state:

$$\boxed{\mathrm{Tr}(\hat{\rho}^2) \;=\; 1 \quad\Leftrightarrow\quad \hat{\rho}^2 \;=\; \hat{\rho} \quad\Leftrightarrow\quad \text{the system is in a pure state}}\,. \tag{141}$$

  Proof ($\Leftarrow$): assume that the system is in the pure state $|\lambda\rangle$. In that case $W_\alpha = \delta_{\alpha\lambda}$ and the density operator $\hat{\rho} \stackrel{(135)}{=\!=} |\lambda\rangle\langle\lambda|$ is a projection operator on $|\lambda\rangle$. This automatically implies that $\hat{\rho}^2 = \hat{\rho}$ and $\mathrm{Tr}(\hat{\rho}^2) = \mathrm{Tr}(\hat{\rho}) = 1$.

  Proof ($\Rightarrow$): let $\mathrm{Tr}(\hat{\rho}^2) = \mathrm{Tr}(\hat{\rho}) = 1$. In terms of the orthonormal basis $\{|k\rangle\}$ of eigenstates of $\hat{\rho}$ we then have

$$\sum_k \rho_k^2 \;=\; \sum_k \rho_k \;=\; 1 \quad\stackrel{\rho_k^2 \leq \rho_k}{\Longrightarrow}\quad \exists_\lambda\; \rho_k \;=\; \delta_{k\lambda}$$

$$\Rightarrow \quad \text{spectral decomposition}: \;\hat{\rho} \;=\; \sum_k \rho_k |k\rangle\langle k| \;=\; |\lambda\rangle\langle\lambda| \;=\; \hat{\rho}^2\,.$$

  Because the density operator can apparently be expressed as a projection operator $\hat{\rho} = |\lambda\rangle\langle\lambda|$, it follows automatically that we are dealing with a pure ensemble with the system being in the pure state $|\lambda\rangle$.

If the system happens to be in a pure state $|\lambda\rangle$, then

$$\hat{\rho} \;=\; |\lambda\rangle\langle\lambda| \;\equiv\; \hat{\rho}_\lambda \tag{142}$$

and all eigenvalues of $\hat{\rho}$ equal 0 except for one eigenvalue 1 belonging to the eigenstate $|\lambda\rangle$. In this situation the ensemble average of the dynamical variable $\mathcal{A}$ obviously coincides with the expectation value for the state $|\lambda\rangle$:

$$[\hat{A}] \;=\; \mathrm{Tr}(\hat{\rho}_\lambda \hat{A}) \;\stackrel{(133)}{=\!=}\; \langle \hat{A} \rangle_\lambda \;=\; \langle\lambda|\hat{A}|\lambda\rangle \;. \tag{143}$$

2.2 Example: polarization of a spin-1/2 ensemble

As a first application of the density-matrix formalism we consider a beam of spin-1/2 particles. A good understanding of the properties of such a beam is of particular interest for scattering experiments in high-energy physics, since at least one of the beams usually involves spin-1/2 particles (such as electrons, positrons, protons or antiprotons). In scattering experiments the beams are sufficiently dilute to treat the beam particles as independent. In this sense the beam should not be regarded as an identical-particle system anymore, but rather as a (mixed) ensemble of beam particles that can be used to perform a repeated quantum mechanical scattering experiment.

The pure states of the beam particles can for instance be decomposed in terms of the basis $\{|\vec{p}, m_s\rangle : \; \vec{p} \in \mathbb{R}^3 \;\text{ and }\; m_s = \pm 1/2\}$, where $\vec{p}$ is the momentum eigenvalue of the beam particle and $m_s\hbar$ the spin component along the z-axis. Next we leave out the momentum variables from the quantum mechanical description and choose to work with the reduced density matrix in the 2-dimensional spin space. Spin information is actually an important tool to analyse scattering data in detail.

Preparatory step: combining two pure beams.

Consider two beams. Beam a consists of N_a spin-1/2 particles prepared in the pure spin state $|\chi^{(a)}\rangle$ and beam b consists of N_b spin-1/2 particles prepared in the pure spin state $|\chi^{(b)}\rangle$. Subsequently both beams are merged. This results in a statistical mixture of the two pure subensembles with weights $W_a = N_a/(N_a+N_b)$ and $W_b = N_b/(N_a+N_b)$. The density operator of this mixed ensemble is given by

$$\hat{\rho} \;=\; W_a |\chi^{(a)}\rangle\langle\chi^{(a)}| \;+\; W_b |\chi^{(b)}\rangle\langle\chi^{(b)}| \;. \tag{144}$$

With respect to the standard basis $\{|\chi_1\rangle \equiv \chi_{\frac{1}{2},\frac{1}{2}} \;,\; |\chi_2\rangle \equiv \chi_{\frac{1}{2},-\frac{1}{2}}\}$ we have

$$|\chi^{(a)}\rangle \;=\; c_1^{(a)}|\chi_1\rangle + c_2^{(a)}|\chi_2\rangle \qquad \text{and} \qquad |\chi^{(b)}\rangle \;=\; c_1^{(b)}|\chi_1\rangle + c_2^{(b)}|\chi_2\rangle \;, \tag{145}$$

giving rise to the density matrix

$$\rho \overset{(136)}{=\!=} \begin{pmatrix} W_a |c_1^{(a)}|^2 + W_b |c_1^{(b)}|^2 & W_a c_1^{(a)} c_2^{(a)\,*} + W_b c_1^{(b)} c_2^{(b)\,*} \\ W_a c_1^{(a)\,*} c_2^{(a)} + W_b c_1^{(b)\,*} c_2^{(b)} & W_a |c_2^{(a)}|^2 + W_b |c_2^{(b)}|^2 \end{pmatrix} \tag{146}$$

in spin space. If the original pure beam states happen to coincide with the chosen basis states of spin space, i.e. $|\chi^{(a)}\rangle = |\chi_1\rangle$ and $|\chi^{(b)}\rangle = |\chi_2\rangle$, then the density matrix simplifies to a diagonal matrix:

$$\rho \;=\; \begin{pmatrix} W_a & 0 \\ 0 & W_b \end{pmatrix} \;=\; \begin{pmatrix} \dfrac{N_a}{N_a+N_b} & 0 \\ 0 & \dfrac{N_b}{N_a+N_b} \end{pmatrix}. \tag{147}$$

Polarization of an arbitrary spin-1/2 ensemble: consider the 2×2 density matrix ρ in spin space. This matrix can be decomposed as

$$\rho \;=\; A_0\, I_2 \,+\, \vec{A}\cdot\vec{\sigma} \qquad (A_0 \in \mathbb{C} \quad \text{and} \quad A_{x,y,z} \in \mathbb{C})\,, \tag{148}$$

in terms of the 2-dimensional identity matrix I_2 and Pauli spin matrices σ_x, σ_y and σ_z that together form a basis for 2×2 matrices (see App. B). From the trace identities (B.5) for these basis matrices we can deduce that

$$\mathrm{Tr}(\rho) \;=\; A_0\,\mathrm{Tr}(I_2) + \vec{A}\cdot\mathrm{Tr}(\vec{\sigma}) \overset{(B.5)}{=\!=} 2A_0 \overset{(139)}{=\!=} 1\,,$$

$$[\sigma_j] = \mathrm{Tr}(\rho\sigma_j) = A_0\,\mathrm{Tr}(\sigma_j) + \sum_k A_k\,\mathrm{Tr}(\sigma_k\sigma_j) \overset{(B.5)}{=\!=} \sum_k 2A_k\,\delta_{jk} = 2A_j \qquad (j,k = x,y,z)\,.$$

In terms of the spin operator $\hat{\vec{S}} = \hbar\vec{\sigma}/2$ the density matrix is given by

$$\boxed{\rho \;=\; \frac{1}{2}\Big(I_2 + \vec{P}\cdot\vec{\sigma}\Big) \;=\; \frac{1}{2}\begin{pmatrix} 1+P_z & P_x - iP_y \\ P_x + iP_y & 1-P_z \end{pmatrix} \quad , \quad \vec{P} \;=\; [\vec{\sigma}\,] \;=\; \frac{2}{\hbar}\,[\hat{\vec{S}}\,] \in \mathbb{R}^3}\,. \tag{149}$$

The density matrix is determined by three (real) parameters $P_{x,y,z}$ that together form the polarization vector $\vec{P}$ of the ensemble. The fact that ρ has three degrees of freedom is relatively easy to comprehend. A complex 2×2 matrix that is hermitian and has a unit trace is fixed by $2\times(2\times 2)\times\frac{1}{2} - 1 = 3$ independent real parameters.

The physical interpretation of the polarization vector $\vec{P}$: to this end we make use of the fact that the eigenvalues λ of the matrix 2ρ should satisfy

$$(1+P_z-\lambda)(1-P_z-\lambda) - P_x^2 - P_y^2 \;=\; (1-\lambda)^2 - \vec{P}^{\,2} \;=\; 0 \quad \Rightarrow \quad \lambda \;=\; 1 \pm |\vec{P}|\,. \tag{150}$$

Hence, in diagonal form the density matrix looks as follows:

$$\rho^{\text{diag}} \;=\; \frac{1}{2}\begin{pmatrix} 1+|\vec{P}| & 0 \\ 0 & 1-|\vec{P}| \end{pmatrix}, \tag{151}$$

which follows from equation (149) by choosing the spin quantization axis (z-axis) parallel to $\vec{P}$. Next define the spin eigenvectors $\chi_\uparrow$ and $\chi_\downarrow$ for spin quantization along $\vec{P}$, so that $(\vec{P}\cdot\vec{\sigma})\chi_\uparrow = |\vec{P}|\chi_\uparrow$ and $(\vec{P}\cdot\vec{\sigma})\chi_\downarrow = -|\vec{P}|\chi_\downarrow$. Then it follows from equation (147) that

$$\frac{1}{2}(1+|\vec{P}|) \;=\; W_\uparrow \;=\; \frac{N_\uparrow}{N_\uparrow+N_\downarrow} \qquad \text{and} \qquad \frac{1}{2}(1-|\vec{P}|) \;=\; W_\downarrow \;=\; \frac{N_\downarrow}{N_\uparrow+N_\downarrow}$$

$$\Rightarrow \quad |\vec{P}| \;=\; W_\uparrow - W_\downarrow \;=\; \frac{N_\uparrow - N_\downarrow}{N_\uparrow+N_\downarrow} \;\in\; [0,1]\,, \tag{152}$$

where $N_\uparrow$ and $N_\downarrow$ count the number of spin-1/2 particles in the spin states $\chi_\uparrow$ and $\chi_\downarrow$. In that way $|\vec{P}|$ can be readily interpreted as the degree of polarization of the ensemble and $\vec{e}_{_P} = \vec{P}/|\vec{P}|$ as the direction of polarization.

Special cases:

- If the ensemble refers to a pure spin state, then we know because of equation (141) that $\rho = \rho^2$. From this it follows directly that ρ is now determined by two real parameters and a sign (as a result of squaring): by means of equation (149) as well as the relation

$$(\vec{\sigma}\cdot\vec{A})(\vec{\sigma}\cdot\vec{B}) \;\overset{(B.6)}{=\!=}\; (\vec{A}\cdot\vec{B})\,I_2 \,+\, i\,\vec{\sigma}\cdot(\vec{A}\times\vec{B}) \qquad\qquad (\vec{A},\vec{B}\in\mathbb{R}^3) \tag{153}$$

  the condition $\rho=\rho^2$ amounts to

$$\rho \;=\; \frac{1}{2}\Big(I_2+\vec{P}\cdot\vec{\sigma}\Big) \;=\; \rho^2 \;=\; \frac{1}{4}\Big(I_2+\vec{P}\cdot\vec{\sigma}\Big)^2 = \frac{1}{2}\left(\frac{1+\vec{P}^{\,2}}{2}\,I_2+\vec{P}\cdot\vec{\sigma}\right) \;\Rightarrow\; |\vec{P}| \;=\; 1\,.$$

  In diagonal form the density matrix is given by the projection operator

$$\rho^{\text{diag}}_{\text{pure}} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} :\; 100\%\text{ polarization with quantization axis parallel to }\vec{P}. \tag{154}$$

  This corresponds to maximal order (read: maximal quantum information), with all particles polarized in the direction of the polarization vector $\vec{P}$.

- For $0<|\vec{P}|<1$ the ensemble is partially polarized, represented by the inequality $\frac{1}{2} < \text{Tr}(\rho^2) = \frac{1}{2}(1+\vec{P}^{\,2}) < 1$.

- We speak of an unpolarized ensemble if $|\vec{P}| = 0$. In that case

$$\boxed{\rho_{\vec{P}=\vec{0}} \;=\; \tfrac{1}{2}\, I_2} \tag{155}$$

and $\mathrm{Tr}(\rho^2)$ takes on its minimal value $\mathrm{Tr}(\rho^2) = \frac{1}{2}$. In this situation as many particles are in the spin "↑" state as in the spin "↓" state. As such we are dealing with an equal admixture of two totally polarized subensembles, one with spin parallel to the quantization axis and one with spin antiparallel to the quantization axis. Note, though, that the direction of this quantization axis can be chosen freely! An unpolarized ensemble is in fact an example of a completely random ensemble with maximal disorder (see § 2.4).

- An ensemble (beam) with degree of polarization $|\vec{P}|$ can be viewed as being composed of a totally polarized part and an unpolarized part. In diagonal form this reads

$$\rho^{\mathrm{diag}} \;\overset{(151)}{=\!=}\; |\vec{P}| \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1-|\vec{P}|}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Freedom to choose: the density matrix (155) of a completely random ensemble does not depend on the choice of representation for the considered space, in contrast to the density matrix (154) for a pure ensemble. Moreover, a given mixed ensemble can be decomposed in different ways in terms of pure ensembles. For instance, a mixture with 20% of the particles polarized in the positive x-direction, 20% in the negative x-direction, 30% in the positive z-direction and 30% in the negative z-direction results in a net unpolarized ensemble, since $0.2\,(\rho_{\vec{P}=\vec{e}_x} + \rho_{\vec{P}=-\vec{e}_x}) + 0.3\,(\rho_{\vec{P}=\vec{e}_z} + \rho_{\vec{P}=-\vec{e}_z}) = \frac{1}{2}\, I_2$.

2.3 The equation of motion for the density operator

As a next step towards quantum statistics we determine the equation of motion for the density operator in the Schrödinger picture. Consider to this end a statistical mixture of pure states that is characterized at $t = t_0$ by the density operator

$$\hat{\rho}(t_0) \;=\; \sum_{\alpha=1}^{N} W_\alpha |\alpha(t_0)\rangle\langle\alpha(t_0)| \;. \tag{156}$$

Assume the weights W_α of the statistical mixture not to change over time. Then the density operator evolves according to

$$|\alpha(t)\rangle \;=\; \hat{U}(t,t_0)|\alpha(t_0)\rangle \quad\Rightarrow\quad \hat{\rho}(t) \;=\; \sum_{\alpha=1}^{N} W_\alpha |\alpha(t)\rangle\langle\alpha(t)| \;=\; \hat{U}(t,t_0)\,\hat{\rho}(t_0)\,\hat{U}^\dagger(t,t_0) \;. \tag{157}$$

As shown in the lecture course Quantum Mechanics 2, the evolution operator $\hat{U}(t,t_0)$ satisfies the differential equation

$$i\hbar\,\frac{\partial}{\partial t}\,\hat{U}(t,t_0) \;=\; \hat{H}(t)\hat{U}(t,t_0) \;, \tag{158}$$

where $\hat{H}(t)$ is the Hamilton operator belonging to the type of system described by the ensemble. From this we obtain the following equation of motion for the density operator:

$$\boxed{i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) \;=\; \big[\hat{H}(t), \hat{\rho}(t)\big]}\,, \tag{159}$$

which is better known as the Liouville equation. This is the quantum mechanical analogue of the equation of motion for the phase-space probability density in classical statistical mechanics, which can be formulated in terms of Poisson brackets as $\partial\rho_{\mathrm{cl}}/\partial t = -\{\rho_{\mathrm{cl}}, \mathcal{H}_{\mathrm{cl}}\}$. For this reason the name "density operator" was given to $\hat{\rho}$.

Note: $\hat{\rho}(t)$ does not possess the typical time evolution that we would expect for a quantum mechanical operator. Since $\hat{\rho}(t)$ is defined in terms of state functions, it is time independent in the Heisenberg picture and time dependent in the Schrödinger picture. This is precisely the opposite of the behaviour of a normal quantum mechanical operator.

Time evolution of the ensemble average of the dynamical variable $\mathcal{A}$: the ensemble average $[\hat{A}]$ as defined in § 2.1 satisfies the evolution equation

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}[\hat{A}] \;&\overset{(137)}{=\!=}\; \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Tr}(\hat{\rho}\hat{A}) \;=\; \mathrm{Tr}\big(\hat{\rho}\,\frac{\partial \hat{A}}{\partial t}\big) \;+\; \mathrm{Tr}\big(\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t}\,\hat{A}\big) \\
&\overset{(159)}{=\!=}\; \mathrm{Tr}\big(\hat{\rho}\,\frac{\partial \hat{A}}{\partial t}\big) \;-\; \frac{i}{\hbar}\,\mathrm{Tr}\big(\hat{H}\hat{\rho}\hat{A} - \hat{\rho}\hat{H}\hat{A}\big) \;=\; \mathrm{Tr}\big(\hat{\rho}\,\frac{\partial \hat{A}}{\partial t}\big) \;-\; \frac{i}{\hbar}\,\mathrm{Tr}\big(\hat{\rho}\hat{A}\hat{H} - \hat{\rho}\hat{H}\hat{A}\big) \\
&\overset{(137)}{=\!=}\; \boxed{\big[\,\frac{\partial \hat{A}}{\partial t}\,\big] \;-\; \frac{i}{\hbar}\,[\hat{A}\hat{H} - \hat{H}\hat{A}] \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\,[\hat{A}]}\,,
\end{aligned} \tag{160}$$

where $\partial\hat{A}/\partial t$ refers to the explicit time dependence of $\hat{A}$. In the penultimate step we have used that the trace is invariant under cyclic permutations, i.e. $\mathrm{Tr}(\hat{A}\hat{B}\hat{C}) = \mathrm{Tr}(\hat{C}\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{C}\hat{A})$. The evolution equation (160) has the same form as the evolution equation for the expectation value of a dynamical variable in ordinary QM. However, this time we have averaged twice in view of the double statistics!

2.4 Quantum mechanical ensembles in thermal equilibrium

> *As we have seen, there are marked differences between pure ensembles (with maximal order) and completely random ensembles (with maximale disorder). We are now going to investigate the generic differences a bit closer.*

Let $\{|k\rangle\}$ be an orthonormal set of eigenstates of $\hat{\rho}$ corresponding to the eigenvalues $\{\rho_k\}$. As such, these eigenstates together span the (reduced) D-dimensional space on which the pure state functions of the subensembles are defined. This dimensionality D indicates the maximum number of independent quantum states that can be identified in the reduced

space on which we have chosen to consider the quantum mechanical systems. For instance, the density matrix in the spin-1/2 spin space has dimensionality $D = 2$. With respect to this basis we have

$$\rho_{\text{pure}} \;=\; \begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & \varnothing & \\ & & 0 & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & \varnothing & & & & \ddots & \\ & & & & & & 0 \end{pmatrix} \qquad \text{vs} \qquad \rho_{\text{random}} \;=\; \frac{1}{D} \begin{pmatrix} 1 & & & & \\ & 1 & & \varnothing & \\ & & \ddots & & \\ & \varnothing & & 1 & \\ & & & & 1 \end{pmatrix}.$$

Examples of these extreme forms of density matrices have been given in equations (154) and (155) in § 2.2. In case of a pure ensemble the density matrix is simply the projection matrix on the corresponding pure state vector, which depends crucially on the chosen representation of the D-dimensional space. In case of a completely random ensemble, the D orthonormal basis states each receive the same statistical weight $1/D$ to guarantee that $\text{Tr}(\hat{\rho}) = 1$. In that case each state is equally probable and the density matrix is proportional to the D-dimensional identity matrix, which does not depend at all on the chosen representation.

In order to quantify the differences we introduce the quantity

$$\boxed{\sigma \;\equiv\; -\,\text{Tr}(\hat{\rho}\ln\hat{\rho}) \;\overset{\text{compl.}}{=\!=}\; -\sum_{k,k'=1}^{D} \langle k|\hat{\rho}|k'\rangle\langle k'|\ln\hat{\rho}\,|k\rangle \;=\; -\sum_{k=1}^{D} \rho_k \ln(\rho_k) \;\equiv\; \frac{S}{k_{\text{B}}}} \,, \quad (161)$$

where ρ_k represents the probability to find the system in the pure basis state $|k\rangle$. This quantity is minimal for pure states and maximal for a completely random ensemble:

$$\boxed{\sigma_{\text{pure}} \;=\; 0 \qquad \text{vs} \qquad \sigma_{\text{random}} \;=\; -\sum_{k=1}^{D} \frac{1}{D} \ln(1/D) \;=\; \ln(D)} \,. \quad (162)$$

That σ_{random} is maximal will be proven in § 2.4.3. In accordance with classical thermodynamics, it can be deduced from $\sigma = S/k_{\text{B}}$ and $\sigma_{\text{random}} = \ln(D)$ that the quantity S should be interpreted as the quantum mechanical entropy[4] and k_{B} as the well-known Boltzmann constant. You could even say that this definition of the entropy is superior to the classical one. In classical mechanics there is no such thing as counting states. At best one could work with phase-space volumes that have to be made dimensionless

[4]This definition of entropy is also used in information theory in the form of the so-called Shannon entropy $-\sum_k P_k \ln(P_k)$, which is an inverse measure for the amount of information that is encoded in the probability distribution $\{P_1, \cdots, P_N\}$. Each type of probability distribution in fact corresponds to a specific type of ensemble. In the lecture course Statistical Mechanics a relation will be established between the thermodynamical and quantum mechanical definitions of the entropy of a canonical ensemble.

by means of an arbitrary normalization factor. For this reason the classical entropy can only be defined up to an additive constant. In QM, however, there is a natural unit of counting/normalization in the form of $\hbar$.

2.4.1 Thermal equilibrium (thermodynamic postulate)

A quantum mechanical ensemble is regarded as being in thermal equilibrium if it effectively satisfies the equilibrium conditions

$$\boxed{\frac{\partial}{\partial t}\hat{H} \;=\; 0 \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho} \;=\; 0} \quad \xRightarrow{(159)} \quad \big[\hat{H},\hat{\rho}\big] \;=\; 0 \tag{163}$$

and if the corresponding constant entropy is maximal. This last condition on the entropy is the quantum mechanical analogue of the classical second law of thermodynamics, which states that the entropy of macroscopic systems in non-equilibrium configurations always increases until equilibrium is reached.

In thermal equilibrium all averaged system quantities $[\hat{A}]$ are constant, provided that $\partial\hat{A}/\partial t = 0$. For example this holds for the average energy of the system, i.e. $\hat{A} = \hat{H}$.

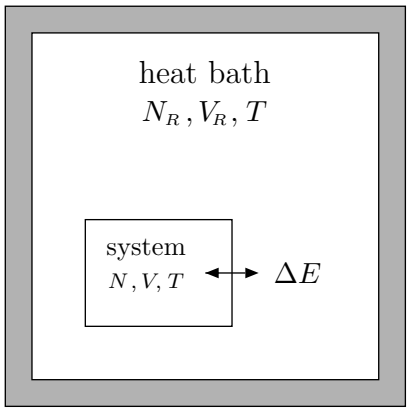
Proof : $\frac{\mathrm{d}}{\mathrm{d}t}[\hat{A}] \overset{(160),\,\partial\hat{A}/\partial t=0}{=\!=\!=} \mathrm{Tr}\big(\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t}\hat{A}\big) \overset{(163)}{=\!=\!=} 0\,.$

In thermal equilibrium the density operator commutes with the Hamilton operator. Therefore an orthonormal set of simultaneous eigenfunctions $\{|k\rangle\}$ of both $\hat{H}$ and $\hat{\rho}$ exists:

$$\hat{H}|k\rangle \;=\; E_k|k\rangle \qquad \text{and} \qquad \hat{\rho}|k\rangle \;=\; \rho_k|k\rangle\;. \tag{164}$$

Using this basis we will maximize the entropy in three different scenarios in § 2.4.2 – 2.4.4.

2.4.2 Canonical ensembles (J.W. Gibbs, 1902)



(N,V,T)-ensemble

closed bottle in a water container

Consider a system with a constant number N of identical particles inside a macroscopic finite enclosure with fixed volume V. Assume the considered system to be in thermal equilibrium with a very large heat bath (outside world), with which it is in weak energy contact. Together the heat bath and the embedded system form a closed system (isolated system).

Weak energy contact: by weak energy contact we mean that the embedded system and heat bath can exchange energy, since only the total energy of the entire isolated system is conserved, but that they do so without having a noticeable influence on each other's energy spectra. Since the mutual interaction is that weak, both the heat bath and the embedded system can be in a definite energy eigenstate at every instance of time (separation of variables). The system and heat bath merely depend on each other through the weak energy contact, which makes the energy of the system time dependent. This weak energy contact does have a role to play, since it provides the mechanism by which the system can reach thermal equilibrium with the heat bath. A weak energy contact can for instance be achieved by separating the system from the heat bath by means of a macroscopic partition wall. This allows energy to be transferred by means of collisions with the atoms in the wall. In that case only a negligible fraction of the enclosed particles will feel the presence of the partition wall if the average de Broglie wavelength of the particles is much smaller than the dimensions of the enclosure (see § 2.5). For instance this applies to systems with sufficiently large (macroscopic) enclosures and/or sufficiently high temperatures.

> *The heat bath determines the mixture: we will see that in thermal equilibrium all energy eigenstates of a closed system have equal probability, i.e. a closed system is described by a completely random ensemble. For a fixed total energy of the entire closed system the number of energy eigenstates of the heat bath is largest if the energy of the heat bath is highest and therefore the energy of the embedded N-particle system is lowest. Subsequently the degrees of freedom of the large heat bath are integrated out, which causes the embedded N-particle systems with the lowest energies to automatically have the highest statistical weights. This gives rise to a so-called canonical ensemble, which is sometimes referred to as (N, V, T)-ensemble. This is a statistical mixture of N-particle systems in specific energy eigenstates. As a result of the weak energy contact with the heat bath, such an N-particle system in thermal equilibrium will not have a fixed energy value. However, the ensemble average $\bar{E}$ will remain constant in thermal equilibrium.*

The canonical density operator: in order to determine the density operator of a canonical ensemble in thermal equilibrium, we have to maximize the entropy under the additional constraints $\mathrm{Tr}(\hat{\rho}) = 1$ and $\mathrm{Tr}(\hat{\rho}\hat{H}) = [\hat{H}] \equiv \bar{E} = \text{constant}$. To this end we employ the Lagrange-multiplier method (see App. C) by varying with respect to the eigenvalues ρ_k as

well as the Lagrange multipliers, i.e.

$$0 \;=\; \delta\Big(\sigma-\beta\big\{\mathrm{Tr}(\hat{\rho}\hat{H})-\bar{E}\big\}-\lambda\big\{\mathrm{Tr}(\hat{\rho})-1\big\}\Big) \;\overset{(164)}{=\!=}\; \delta\Big(\beta\bar{E}+\lambda-\sum_k \rho_k\big\{\ln(\rho_k)+\beta E_k+\lambda\big\}\Big)$$

for all variations $\delta\rho_k$, $\delta\beta$ and $\delta\lambda$. Note: there is no need to vary with respect to the energy eigenvalues E_k, as these energy eigenvalues are independent of the statistical mixture in case of a weak energy contact. From the Lagrange-multiplier method it follows automatically that the constraint conditions have to be satisfied as well as the condition

$$\underset{\delta\rho_k}{\forall}\quad \sum_k \delta\rho_k\big\{\ln(\rho_k)+1+\beta E_k+\lambda\big\} \;=\; 0 \qquad\Rightarrow\qquad \rho_k \;=\; \exp(-1-\lambda)\exp(-\beta E_k)\;.$$

The Lagrange multiplier λ can now be eliminated by means of the constraint condition

$$\mathrm{Tr}(\hat{\rho}) \;=\; \exp(-1-\lambda)\sum_{k'}\exp(-\beta E_{k'}) \;=\; 1\;,$$

which implies for ρ_k that

$$\rho_k \;=\; \frac{\exp(-\beta E_k)}{\sum\limits_{k'}\exp(-\beta E_{k'})}\;. \tag{165}$$

In these sums the summation runs over all fully specified N-particle energy eigenvalues, including the complete degree of degeneracy of these energy levels. Subsequently we use the Lagrange multiplier β to define the temperature of the considered ensemble:

$$\boxed{T \;\equiv\; \frac{1}{k_{\mathrm{B}}\beta}}\;. \tag{166}$$

Out of the complete disorder of the closed system, the contact with the heat bath has created order in the energy distribution of the embedded N-particle system. Note: if the N-particle ground state is not degenerate, then the canonical ensemble will change in the low-temperature limit $T\to 0$ ($\beta\to\infty$) into a pure ensemble where all N-particle systems will be in the ground state.

Finally we introduce the canonical partition function (normalization factor)

$$\boxed{Z_{\scriptscriptstyle N}(T) \;\equiv\; \sum_{k'}\exp(-\beta E_{k'}) \;=\; \sum_{k'}\langle k'|\exp(-\beta\hat{H})|k'\rangle \;=\; \mathrm{Tr}\big(\exp(-\beta\hat{H})\big)}\;. \tag{167}$$

In spectral decomposition the density operator of a canonical ensemble then reads

$$\hat{\rho} \;\overset{(164)}{=\!=}\; \sum_k \rho_k\,|k\rangle\langle k| \;\overset{(165)}{=\!=}\; \sum_k \frac{\exp(-\beta E_k)}{Z_{\scriptscriptstyle N}(T)}\,|k\rangle\langle k| \;=\; \frac{1}{Z_{\scriptscriptstyle N}(T)}\exp(-\beta\hat{H})\sum_k |k\rangle\langle k|$$

$$\overset{\text{compl.}}{\Longrightarrow} \qquad \boxed{\hat{\rho}(\hat{H}) \;=\; \frac{1}{Z_{\scriptscriptstyle N}(T)}\exp(-\beta\hat{H}) \;=\; \frac{\exp(-\beta\hat{H})}{\mathrm{Tr}\big(\exp(-\beta\hat{H})\big)}}\;. \tag{168}$$

The label N in $Z_N(T)$ refers to the number of particles of the considered type of embedded system. An arbitrary averaged physical quantity of the embedded system is then given by the ensemble average

$$\boxed{[\hat{A}] \;=\; \frac{1}{Z_N(T)}\,\mathrm{Tr}\big(\hat{A}\exp(-\beta\hat{H})\big)}\,, \tag{169}$$

allowing us to derive the average system energy from the partition function:

$$\boxed{\bar{E} \;=\; [\hat{H}] \;=\; \frac{\mathrm{Tr}\big(\hat{H}\exp(-\beta\hat{H})\big)}{Z_N(T)} \;=\; \frac{-\,\partial Z_N(T)/\partial\beta}{Z_N(T)} \;=\; -\,\frac{\partial}{\partial\beta}\ln\big(Z_N(T)\big)}\,. \tag{170}$$

Number of particles and the canonical-ensemble concept: the embedded systems of a canonical ensemble can just as well consist of a single particle as of a macroscopic number of particles (such as a gas). A popular way to link up with the classical thermodynamics of ideal gases is to single out one particle in a larger gas system and consider the corresponding (single-particle) canonical ensemble. This approach of treating the rest of the gas as part of the heat bath has its limitations, though. The canonical-ensemble concept implies that the influence exerted by the particles of the heat bath on the particle of the embedded system is exclusively limited to a weak energy contact. This obviously breaks down if the particle density is too high, causing the quantum mechanical overlap between the identical particles in embedded system and heat bath to become relevant. In that case also the effective quantum mechanical interaction induced by the (anti)symmetrization procedure should be taken into account. These quantum mechanical effects will become more pronounced with increasing particle densities (see § 2.5 – 2.7). In the two examples given below we will nevertheless consider two canonical ensembles of systems that consist of a single particle. The reason to do so is to first link up with classical statistical physics before zooming in on the quantum mechanical aspects. In order to study these many-particle aspects, the grand-canonical ensemble approach will turn out to be more handy.

Equipartition of energy: as a standard example we consider a canonical ensemble of systems that each consist of a single free spin-0 particle with mass m that is contained in a macroscopic enclosure (box). Since quantum mechanical many-particle aspects evidently do not play a role in a 1-particle system, we expect the average energy per particle in that case to obey the classical principle of equipartition of energy for an ideal gas. This states that each kinetic and elastic degree of freedom of the considered type of system (i.e. a position/momentum degree of freedom that contributes quadratically to the Hamiltonian) will contribute $k_{\mathrm{B}}T/2$ to the average energy. This is indeed confirmed by the quantum mechanical calculation (see exercise 13), which tells us that the quantum mechanical definitions of entropy, temperature and Boltzmann constant are consistent with their classical thermodynamic counterparts.

Example: magnetization. Consider a canonical ensemble of systems that each consist of a single free electron that is experiencing a constant homogeneous magnetic field in the z-direction, $\vec{\mathcal{B}} = \mathcal{B}\vec{e}_z$. In spin space this corresponds to an interaction

$$\hat{H}_{B}^{\text{spin}} \;=\; -\hat{\vec{\mathcal{M}}}_S \cdot \vec{\mathcal{B}} \;=\; \frac{2\mu_{B}\mathcal{B}}{\hbar}\,\hat{S}_z \;, \tag{171}$$

in terms of the Bohr magneton μ_{B}. The full Hamilton operator $\hat{H} = \hat{H}^{\text{spatial}} + \hat{H}_{B}^{\text{spin}}$ of the electron commutes with $\hat{S}_z$ and therefore the same should hold for $\hat{\rho}(\hat{H})$. Hence, the density matrix is diagonal in spin space if we use a basis of eigenvectors of $\hat{S}_z$ to describe it. For a given spatial energy level the density operator has the following form in spin space:

$$\rho^{\text{spin}} \;=\; \frac{1}{\exp(-\beta\mu_{B}\mathcal{B}) + \exp(\beta\mu_{B}\mathcal{B})} \begin{pmatrix} \exp(-\beta\mu_{B}\mathcal{B}) & 0 \\ 0 & \exp(\beta\mu_{B}\mathcal{B}) \end{pmatrix} . \tag{172}$$

This ensemble corresponds to the following polarization vector:

$$\begin{aligned} \rho^{\text{spin}} \;&=\; \frac{\cosh(\beta\mu_{B}\mathcal{B})\, I_2 - \sinh(\beta\mu_{B}\mathcal{B})\,\sigma_z}{2\cosh(\beta\mu_{B}\mathcal{B})} \;=\; \frac{1}{2}\Big(I_2 - \tanh(\beta\mu_{B}\mathcal{B})\,\sigma_z\Big) \\ \xrightarrow{(149)} \quad \vec{P} \;&=\; -\tanh(\beta\mu_{B}\mathcal{B})\vec{e}_z \;=\; \frac{2}{\hbar}\,[\hat{S}_z]\,\vec{e}_z \;. \end{aligned} \tag{173}$$

From this we can derive the magnetization, i.e. the average magnetic moment per electron in the direction of the magnetic field:

$$\overline{M}_S \;\equiv\; \frac{1}{\mathcal{B}}\,[\hat{\vec{\mathcal{M}}}_S \cdot \vec{\mathcal{B}}] \;=\; -\frac{2\mu_{B}}{\hbar}\,[\hat{S}_z] \;=\; \mu_{B}\tanh(\beta\mu_{B}\mathcal{B}) \;. \tag{174}$$

If $\beta\mu_{B}\mathcal{B} \ll 1$, for instance when the magnetic field is weak or the temperature high, then we obtain Curie's law:

$$\overline{M}_S \;\approx\; \mu_{B}\,(\beta\mu_{B}\mathcal{B}) \;\xrightarrow{(166)}\; \frac{\mu_{B}^2}{k_{\text{B}}T}\,\mathcal{B} \quad\Rightarrow\quad \chi_S(T) \;\equiv\; \Big(\frac{\mathrm{d}\overline{M}_S}{\mathrm{d}\mathcal{B}}\Big)_{\mathcal{B}=0} = \;\frac{\mu_{B}^2}{k_{\text{B}}T} \;\propto\; T^{-1} \;, \tag{175}$$

where $\chi_S(T)$ is called the paramagnetic susceptibility of the spin ensemble.

2.4.3 Microcanonical ensembles

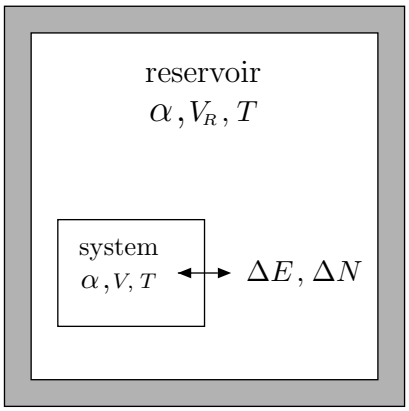
If we do not impose the constraint $\mathrm{Tr}(\hat{\rho}\hat{H}) = \bar{E}$ in the previous derivation, then we will obtain the ensemble with maximum disorder once we maximize the entropy. After all, there is no heat bath to integrate out and create any order in the ensemble. We speak in that case of a microcanonical ensemble or (N, V, E)-ensemble, which determines the thermodynamical properties of a closed system in thermal equilibrium. Because of the absence of energy contact, the system energy is constant and not just the average value of

the system energy. If we take the density matrix ρ to have dimensionality D, then we find by employing the Lagrange-multiplier method that

$$\delta\Big(\sigma - \lambda\big\{\mathrm{Tr}(\hat{\rho}) - 1\big\}\Big) \overset{(164)}{=\!=\!=} \delta\Big(\lambda - \sum_{k=1}^{D} \rho_k\big\{\ln(\rho_k) + \lambda\big\}\Big) = 0$$

for all variations $\delta\rho_k$ and $\delta\lambda$. From this it follows that $\mathrm{Tr}(\hat{\rho}) = 1$ and $\ln(\rho_k) + 1 + \lambda = 0$. By combining the two conditions $\rho_k = \exp(-1-\lambda)$ and $\mathrm{Tr}(\hat{\rho}) = D\exp(-1-\lambda) = 1$, we arrive at the typical diagonal form $\rho_k = 1/D$ as given in § 2.4 for a completely random ensemble. This also proves the conjecture that the density matrix of a completely random ensemble gives rise to the absolute maximum of entropy. Note also that from equation (168) it can be read off that a microcanonical ensemble coincides with the high-temperature limit $T \to \infty$ $(\beta \to 0)$ of a canonical ensemble.

2.4.4 Grand canonical ensembles (J.W. Gibbs, 1902)



(α, V, T) or (μ, V, T)-ensemble

open bottle in a water container

Finally, we consider a quantum mechanical ensemble that satisfies the same criteria as for a canonical ensemble with the exception that the embedded system is allowed to exchange both energy and particles with a reservoir. The (open) many-particle systems described by the ensemble have a variable number of particles. However, the total many-particle Hamilton operator $\hat{H}_{\mathrm{tot}}$ of such a many-particle system satisfies certain particle-number conservation laws.[5] We know that the density operator has to be a constant of motion and that any order in the ensemble is imparted by the contact with the reservoir. As such, we expect the density operator to depend exclusively on observables that belong to conserved quantities that can be exchanged between reservoir and open system. In the considered case that would be the total many-particle Hamilton operator and the conserved

[5] For particles that are not constrained by conservation laws, such as thermal photons, only a canonical-ensemble approach is formally applicable. This will be explained in chapter 4 during the discussion of photon-gas systems.

combinations of total number operators. Let's for convenience assume that the open system consists of one type of particle, i.e. we exclude the possibility of particle mixtures, and that no reactions can occur that affect the number of particles (such as particle decays). In that case $\hat{H}_{\rm tot}$ has to be additive, i.e. $\big[\hat{H}_{\rm tot},\hat{N}\big]=0$, to guarantee that $\hat{N}$ is a constant of motion. In this way the number of particles plays the same role as energy did in the canonical case, simply because the open system has no fixed number of particles but the complete closed system including the reservoir does. In thermal and diffusive equilibrium, the resulting type of ensemble is called a grand canonical ensemble, (α,V,T)-ensemble, or (μ,V,T)-ensemble.

Apart from the average energy we now also have to impose a constraint on the average number of particles of the open system while maximizing the entropy. For this purpose we use the total number operator $\hat{N}$ that counts the total number of particles in the open system, which has eigenvalues $0,1,2,\cdots$. Apart from the Hamilton operator also the density operator is additive, in order to guarantee that the entropy is additive. Hence, the total number operator commutes with both the Hamilton operator and the density operator. Bear in mind, though, that these two operators *will depend on the precise number of particles present in the open system*. This gives rise to the following set of mutually commuting observables $\hat{N}$, $\hat{H}_{\rm tot}(\hat{N})$ and $\hat{\rho}(\hat{N})$, with corresponding orthonormal basis $\{|k,N\rangle\}$:

$$\hat{N}|k,N\rangle \;=\; N|k,N\rangle \quad , \quad \hat{H}_{\rm tot}(\hat{N})|k,N\rangle \;=\; E_k(N)|k,N\rangle$$

$$\text{and} \qquad \hat{\rho}(\hat{N})|k,N\rangle \;=\; \rho_k(N)|k,N\rangle \;. \tag{176}$$

The grand canonical density operator: in order to determine the density operator of a grand canonical ensemble in thermal and diffusive equilibrium we have to maximize the entropy under the additional constraints $\mathrm{Tr}(\hat{\rho})=1$, $[\hat{H}_{\rm tot}(\hat{N})]\equiv\bar{E}_{\rm tot}=$ constant and $[\hat{N}]\equiv\bar{N}=$ constant. By means of the Lagrange-multiplier method we find

$$\begin{aligned} 0 \;&\overset{}{=\!=\!=}\; \delta\Big(\sigma-\alpha\big\{[\hat{N}]-\bar{N}\big\}-\beta\big\{[\hat{H}_{\rm tot}(\hat{N})]-\bar{E}_{\rm tot}\big\}-\lambda\big\{\mathrm{Tr}(\hat{\rho})-1\big\}\Big) \\ &\overset{(176)}{=\!=\!=}\; \delta\Big(\alpha\bar{N}+\beta\bar{E}_{\rm tot}+\lambda-\sum_N\sum_k \rho_k(N)\big\{\ln\big(\rho_k(N)\big)+\alpha N+\beta E_k(N)+\lambda\big\}\Big) \end{aligned} \tag{177}$$

for all variations $\delta\rho_k(N)$, $\delta\alpha$, $\delta\beta$ and $\delta\lambda$. From this it follows automatically that the constraint conditions have to be satisfied as well as the condition

$$\underset{\delta\rho_k(N)}{\forall}\quad \sum_N\sum_k \delta\rho_k(N)\big\{\ln\big(\rho_k(N)\big)+1+\alpha N+\beta E_k(N)+\lambda\big\} \;=\; 0\;.$$

Combined with the constraint $\mathrm{Tr}(\rho) = 1$ this results in

$$\rho_k(N) \;=\; \frac{\exp\bigl(-\beta E_k(N) - \alpha N\bigr)}{\sum\limits_{N'}\sum\limits_{k'} \exp\bigl(-\beta E_{k'}(N') - \alpha N'\bigr)} \;. \tag{178}$$

As in equation (166), the temperature T of the ensemble is defined through $T \equiv 1/k_{\mathrm{B}}\beta$. Finally, we introduce the grand canonical partition function

$$\boxed{\mathcal{Z}(\alpha, T) \;\equiv\; \sum_{N'}\sum_{k'} \exp\bigl(-\beta E_{k'}(N') - \alpha N'\bigr) \;\overset{(167)}{=\!=}\; \sum_{N'} \exp(-\alpha N')\, Z_{N'}(T)} \tag{179}$$

and rewrite the grand-canonical density operator as

$$\boxed{\hat{\rho} \;=\; \frac{1}{\mathcal{Z}(\alpha, T)} \exp\bigl(-\beta \hat{H}_{\mathrm{tot}}(\hat{N}) - \alpha \hat{N}\bigr) \;=\; \hat{\rho}\bigl(\hat{H}_{\mathrm{tot}}(\hat{N}), \hat{N}\bigr)} \;. \tag{180}$$

The quantity $\exp(-\alpha)$ can be regarded as the fugacity of the ensemble, which describes how easy it is to move a particle from the reservoir to the open system:

- if $\exp(-\alpha)$ is small, than contributions from small N values dominate and quantum mechanical many-particle aspects are not important;
- if $\exp(-\alpha)$ is not small, than particle exchange is easier and quantum mechanical many-particle aspects are relevant.

Often the fugacity is written as $\exp(\beta\mu)$, with $\mu = -\alpha/\beta$ known as the chemical potential. The average total energy and the average number of particles of the open system can be extracted directly from the grand canonical partition function:

$$\boxed{\bar{E}_{\mathrm{tot}} \;=\; -\frac{\partial}{\partial\beta} \ln\bigl(\mathcal{Z}(\alpha, T)\bigr) \qquad \text{and} \qquad \bar{N} \;=\; -\frac{\partial}{\partial\alpha} \ln\bigl(\mathcal{Z}(\alpha, T)\bigr)} \;. \tag{181}$$

Remark: in scenarios that do either involve chemical reactions, particle decays or relativistic energies, we have to abandon conservation laws for individual types of particles. However, in those cases certain combinations of particle numbers often will be conserved. While maximizing the entropy, these generalized conservation laws will have to be imposed as constraints on the corresponding ensemble averages. As before this can be implemented by means of the Lagrange-multiplier method, assigning to each conservation law a specific Lagrange multiplier and a related chemical potential. For instance, for relativistic energies particle–antiparticle pair creation can occur, so that only the difference between the number of particles and antiparticles remains conserved (which is equivalent to charge conservation). Another example is provided by a statistical mixture of A-, B- and C-particles for which the reversible reaction $A \rightleftharpoons B + C$ is in chemical equilibrium. In that case the particle-number combination $2N_{_A} + N_{_B} + N_{_C}$ is conserved.

Possible realization: if we are dealing with a vary large number of particles inside a macroscopic volume, then the grand canonical ensemble can be realized straightforwardly by partitioning the macroscopic volume in very many identical cells. The inherent condition is that the cells themselves are again sufficiently macroscopic to guarantee that the energy contact with the reservoir (read: the other cells) can be regarded as weak. Repeated measurements can be performed on such an ensemble by simply measuring locally. The internal structure of a star can, for instance, be determined along these lines by partitioning the star in macroscopic volume elements that are of negligible size compared to the dimensions of the star itself (see exercise 18). The relevant "measurable" quantities are in that case the local particle densities and the local total energy density.

2.4.5 Summary

The differences between the three types of ensembles reside in the increasing degree of order that is imparted on the embedded systems by the contact with the reservoir. In the microcanonical case this contact is extremely weak. It is that weak that we could effectively speak of the absence of any contact, causing E and N to be effectively fixed. Thermal equilibrium is in that case equivalent with the absence of any order. In the canonical case energy transfer occurs between reservoir and embedded system, with the total energy being conserved. Since the reservoir has so many more degrees of freedom than the embedded system, in thermal equilibrium order is generated in the energy mix of the embedded system with a preference for lower energies (i.e. higher reservoir energies). In the grand canonical case we add to this the possibility of particle exchange, with the corresponding particle conservation law translating into extra order in the particle mix of the embedded open system. The three different types of ensembles give rise to identical physical results if the relative fluctuations around the equilibrium values $\bar{E}_{\text{tot}}$ and $\bar{N}$ are very small, causing the energy and number of particles of the embedded system to be effectively fixed. It is then a matter of choice which type of ensemble is more suitable for performing calculations (which in most cases will be the grand canonical ensemble). For instance this applies to systems consisting of a very large number of particles, in view of the extra factors $1/\sqrt{\bar{N}}$ that in general occur in the average relative fluctuations on the measurements.

A Fourier series and Fourier integrals

Fourier decompositions are used quite frequently in these lecture notes. We briefly list the relevant aspects of these decompositions in this appendix.

A.1 Fourier series

Systems with a ring-shaped structure and systems that are enclosed spatially can be described in terms of state functions with a certain degree of periodicity. Consider to this end a 1-dimensional function $f(x)$, for which the function as well as the spatial derivative are stepwise continuous on the interval $x \in \big(-L/2, L/2\big)$ and which additionally satisfies the spatial periodicity condition $f(x) = f(x+L)$. From Fourier analysis it follows (see the lecture course "Trillingen en Golven") that such a function can be decomposed in terms of a basis of periodic plane waves (Fourier modes):

$$f_n(x) \;=\; \frac{1}{\sqrt{L}}\exp(2\pi i n x/L) \;\equiv\; \frac{1}{\sqrt{L}}\exp(ikx) \qquad (n \,=\, 0, \pm 1, \pm 2, \cdots)\;. \tag{A.1}$$

We readily recognize these functions as being momentum eigenstates belonging to the quantized momentum eigenvalues $\hbar k$. This basis satisfies the orthonormality relation

$$\frac{1}{L}\int\limits_{x_0}^{x_0+L} \mathrm{d}x\, \exp\big(ix[k-k']\big) \;=\; \delta_{kk'}\;, \tag{A.2}$$

where x_0 can be chosen arbitrarily. Since x is usually restricted to the interval $\big(-L/2, L/2\big)$, a handy choice is $x_0 = -L/2$. By employing the completeness relation for periodic Fourier modes, the periodic function $f(x)$ can be written as a Fourier series:

$$f(x) \;=\; \sum_{n=-\infty}^{\infty} \frac{A_n}{\sqrt{L}}\exp(2\pi i n x/L) \tag{A.3}$$

$$\;=\; \frac{A_0}{\sqrt{L}} \,+\, \sum_{n=1}^{\infty}\Big[\,\frac{A_n + A_{-n}}{\sqrt{L}}\cos(2\pi n x/L) \,+\, i\,\frac{A_n - A_{-n}}{\sqrt{L}}\sin(2\pi n x/L)\Big]$$

$$\;\equiv\; \frac{B_0}{\sqrt{L}} \,+\, \sum_{n=1}^{\infty}\sqrt{\frac{2}{L}}\,\Big[B_n\cos(2\pi n x/L) + C_n\sin(2\pi n x/L)\Big]\;, \tag{A.4}$$

with

$$A_n \;=\; \frac{1}{\sqrt{L}}\int\limits_{-L/2}^{L/2} \mathrm{d}x\, f(x)\exp(-2\pi i n x/L)\;. \tag{A.5}$$

Periodic systems vs enclosed systems: the close relation between periodic systems and systems that are contained inside an enclosure can be seen as follows. If we assume the impenetrable walls of the enclosure to be positioned at $x = 0$ and $x = L$, the state functions

have to satisfy the continuity condition $\psi(0) = \psi(L) = 0$ (see § 2.5). Inside the enclosure this function coincides with a periodic function $f(x) = f(x + 2L)$ with odd parity, so that $f(x) = -f(-x)$. As a result of this extra condition, only the standing waves $\propto \sin(\pi n x/L)$ for $n = 1, 2, \cdots$ contribute to the Fourier series (A.4), which indeed guarantees that $f(0) = f(L) = 0$. The periodicity length is $2L$ now and at the same time n is allowed to take on positive values only. The number of allowed quantized k values is in this case identical to the number of k values for a system with periodicity length L, with the exception of the absence of the quantum number $k = 0$. The presence of the edge of the box manifests itself in the absence of the constant Fourier mode with quantum number $k = 0$. The influence of the edge of the enclosure is only felt when the quantum states with the lowest absolute values for the momentum (and thus the largest de Broglie wavelengths) contribute significantly to the description of the system. This happens for instance in bosonic gases at low temperatures (see Ch. 2).

A.2 Fourier integrals

The Fourier decomposition of non-periodic quadratically integrable functions can be obtained by taking the continuum limit $L \to \infty$ of the periodic case given above. During this transition the resulting basis of plane waves should be normalized to a δ function. This is based on the fact that $k = 2\pi n/L$, causing the k-spectrum to become very closely spaced for $L \to \infty$ and the summation over n to become an integral over k:

$$\sum_{n=-\infty}^{\infty} \quad \stackrel{L\to\infty}{\longrightarrow} \quad \frac{L}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k \ . \tag{A.6}$$

The relevant basis functions are the momentum eigenfunctions belonging to the momentum eigenvalues $\hbar k$:

$$f_k(x) \ = \ f_n(x) \sqrt{\frac{L}{2\pi}} \ \stackrel{(A.1)}{=\!=\!=} \ \frac{1}{\sqrt{2\pi}} \exp(ikx) \qquad (k \in \mathbb{R}) \ , \tag{A.7}$$

with orthonormality relation

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}x \, \exp\bigl(ix[k-k']\bigr) \ = \ \delta(k-k') \ . \tag{A.8}$$

The Fourier series now becomes a Fourier transformation (Fourier integral):

$$f(x) \ = \ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k \, A(k) \exp(ikx) \qquad \text{and} \qquad A(k) \ = \ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x \, f(x) \exp(-ikx) \ . \tag{A.9}$$

A.2.1 Definition of the δ function

The δ function in equation (A.8) is defined on an arbitrary continuous test function $g(x)$:

$$\underset{g}{\forall} \quad g(x) \;=\; \int\limits_{-\infty}^{\infty} \mathrm{d}x'\, g(x')\delta(x-x') \;. \tag{A.10}$$

Since $g(x)$ is arbitrary, $\delta(x-x')$ should vanish for $x \neq x'$. The peak of $\delta(x-x')$ at $x = x'$ should be such that the integral over $\delta(x-x')$ around $x = x'$ always yields one. The δ function therefore only exists as a limit of a series of functions that progressively become more peaked around $x = x'$. The result of such a limiting procedure is called a distribution. An example of this is given by the integral

$$\delta_\epsilon(x) \;=\; \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k\, \exp(ikx - \epsilon|k|) \;=\; \frac{1}{2\pi}\Big(\frac{1}{ix+\epsilon} - \frac{1}{ix-\epsilon}\Big) \;=\; \frac{1}{\pi}\,\frac{\epsilon}{x^2+\epsilon^2} \qquad (\epsilon > 0) \;.$$

For decreasing ϵ the function $\delta_\epsilon(x)$ becomes smaller for all $x \neq 0$, whereas simultaneously

$$\int\limits_{x_1<0}^{x_2>0} \mathrm{d}x\; \delta_\epsilon(x) \;\overset{v=x/\epsilon}{=\!=}\; \frac{1}{\pi} \int\limits_{x_1/\epsilon}^{x_2/\epsilon} \frac{\mathrm{d}v}{v^2+1} \;=\; \frac{1}{\pi}\Big(\arctan(x_2/\epsilon) - \arctan(x_1/\epsilon)\Big) \;\overset{\epsilon\downarrow 0}{\longrightarrow}\; 1 \;.$$

The following series representations of the δ function will be used in these lecture notes:

$$\delta(x) \;=\; \lim_{\epsilon\downarrow 0} \frac{1}{\pi}\,\frac{\epsilon}{x^2+\epsilon^2} \;=\; \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k\, \exp(ikx) \;, \tag{A.11}$$

$$\delta(x) \;=\; \lim_{\epsilon\downarrow 0} \frac{\theta(x+\epsilon) - \theta(x)}{\epsilon} \;=\; \theta'(x) \;, \tag{A.12}$$

where $\theta(x)$ is the usual Heaviside step function.

The following handy identities hold for the δ function:

$$g(x)\delta(x-c) \;=\; g(c)\delta(x-c) \;, \tag{A.13}$$

as well as

$$\delta\big(h(x)\big) \;=\; \sum_j \frac{1}{|h'(x_j)|}\,\delta(x-x_j) \qquad \text{for} \quad h(x_j) = 0 \quad \text{and} \quad h'(x_j) \neq 0 \;, \tag{A.14}$$

provided that $h(x)$ is sufficiently well-behaved close to the roots of the equation $h(x) = 0$.

In three dimensions, finally, the following compact notation can be used:

$$\delta(\vec{r}) \;\equiv\; \delta(x)\delta(y)\delta(z) \quad \Rightarrow \quad \frac{1}{(2\pi)^3} \int\limits_{-\infty}^{\infty} \mathrm{d}k_x \int\limits_{-\infty}^{\infty} \mathrm{d}k_y \int\limits_{-\infty}^{\infty} \mathrm{d}k_z\, \exp(i\vec{k}\cdot\vec{r}) \;\overset{(A.11)}{=\!=}\; \delta(\vec{r}) \;. \tag{A.15}$$

B Properties of the Pauli spin matrices

Since the world of QM is littered with spin-1/2 particles (such as electrons, nucleons, etc.), we give here a short summary of the properties of the Pauli spin matrices

$$\sigma_x \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_y \;=\; \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma_z \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{B.1}$$

These 2×2 matrices are linked directly to the spin operator $\hat{\vec{S}}$ for spin-1/2 particles:

$$\vec{\sigma} \;\equiv\; \frac{2}{\hbar}\,\hat{\vec{S}} \,. \tag{B.2}$$

From the fundamental commutator algebra for the spin operator and the fact that $\hat{S}_z$ has eigenvalues $\pm\hbar/2$, the following properties of the Pauli spin matrices can be derived:

$$\begin{aligned} &\sigma_k \;=\; \sigma_k^\dagger \quad , \quad \mathrm{Tr}(\sigma_k) \;=\; 0 \quad , \quad \det(\sigma_k) \;=\; -1 \qquad (k = x,y,z) \\ &\text{and} \quad \big[\sigma_x, \sigma_y\big] \;=\; 2i\sigma_z \qquad \text{(cyclic)} \,. \end{aligned} \tag{B.3}$$

Denoting the identity matrix in spin space by I_2, the following additional identities hold for the Pauli spin matrices:

$$\begin{aligned} &\sigma_k^2 \;=\; I_2 \qquad (k = x,y,z) \qquad \text{and} \qquad \sigma_x\sigma_y \;=\; -\,\sigma_y\sigma_x \;=\; i\sigma_z \qquad \text{(cyclic)} \\ &\Rightarrow \quad \big\{\sigma_j, \sigma_k\big\} \;=\; 2\delta_{jk} I_2 \qquad (j\,,k = x,y,z) \,. \end{aligned} \tag{B.4}$$

This allows us to complete the relevant trace identities in spin space to

$$\mathrm{Tr}(I_2) \;=\; 2 \quad , \quad \mathrm{Tr}(\sigma_j) \;\overset{(B.3)}{=\!=}\; 0 \qquad \text{and} \qquad \mathrm{Tr}(\sigma_j\sigma_k) \;\overset{(B.4)}{=\!=}\; 2\delta_{jk} \qquad (j\,,k = x,y,z) \,. \tag{B.5}$$

Moreover, for arbitrary 3-dimensional vectors $\vec{A}$ and $\vec{B}$ the following holds:

$$(\vec{\sigma}\cdot\vec{A})(\vec{\sigma}\cdot\vec{B}\,) \;=\; \frac{1}{2}\sum_{j,k} A_j\,B_k\big(\,\big[\sigma_j,\sigma_k\big] + \big\{\sigma_j,\sigma_k\big\}\,\big) \;\overset{(B.3),(B.4)}{=\!=\!=}\; (\vec{A}\cdot\vec{B}\,)I_2 + i\vec{\sigma}\cdot(\vec{A}\times\vec{B}\,) \,, \tag{B.6}$$

using that

$$\sum_{j,k} \frac{A_j\,B_k}{2}\big[\sigma_j,\sigma_k\big] \;\overset{(B.3)}{=\!=}\; i\sigma_z(A_x B_y - A_y B_x) + i\sigma_x(A_y B_z - A_z B_y) + i\sigma_y(A_z B_x - A_x B_z) \,.$$

Combined with the identity matrix the Pauli spin matrices form a basis of 2×2 matrices. As such, an arbitrary 2×2 matrix A can be decomposed as

$$A \;=\; A_0\,I_2 \;+\; \vec{A}\cdot\vec{\sigma} \qquad (A_0 \in \mathbb{C} \quad \text{and} \quad A_{x,y,z} \in \mathbb{C}) \,. \tag{B.7}$$

C Lagrange-multiplier method

Suppose we want to determine the extrema of a quantity $F(q_1, \cdots, q_N)$ while simultaneously trying to satisfy the boundary condition $f(q_1, \cdots, q_N) = C$. In that case it makes good sense to employ the so-called Lagrange-multiplier method. This consists in looking for solutions to the variational equation

$$\delta\Big(F(q_1, \cdots, q_N) \,-\, \lambda(q_1, \cdots, q_N)\big\{ f(q_1, \cdots, q_N) - C \big\}\Big) \;=\; 0 \;, \tag{C.1}$$

where δ denotes the action of independently varying the coordinates $q_1, \cdots, q_N$ as well as the Lagrange multiplier λ. This results in

$$\sum_{j=1}^{N} \Big(\frac{\partial F}{\partial q_j} \,-\, (f-C)\,\frac{\partial \lambda}{\partial q_j} \,-\, \lambda\,\frac{\partial f}{\partial q_j} \Big)\,\delta q_j \,-\, (f-C)\,\delta\lambda \;=\; 0 \qquad \underset{\delta\lambda,\,\delta q_j}{\forall}$$

$$\Rightarrow \quad f(q_1, \cdots, q_N) \;=\; C \qquad \text{and} \qquad \frac{\partial F}{\partial q_j} \;=\; \lambda\,\frac{\partial f}{\partial q_j} \qquad (j = 1, \cdots, N)\;. \tag{C.2}$$

In this way the boundary condition is indeed satisfied automatically. By substituting the obtained solution(s) for $q_1, \cdots, q_N$ in the boundary condition $f(q_1, \cdots, q_N) = C$, the Lagrange multiplier λ can be related to the constant C. The Lagrange-multiplier method is used in these lecture notes to maximize the entropy of ensembles in thermal equilibrium while simultaneously imposing certain additional constraints.

Index