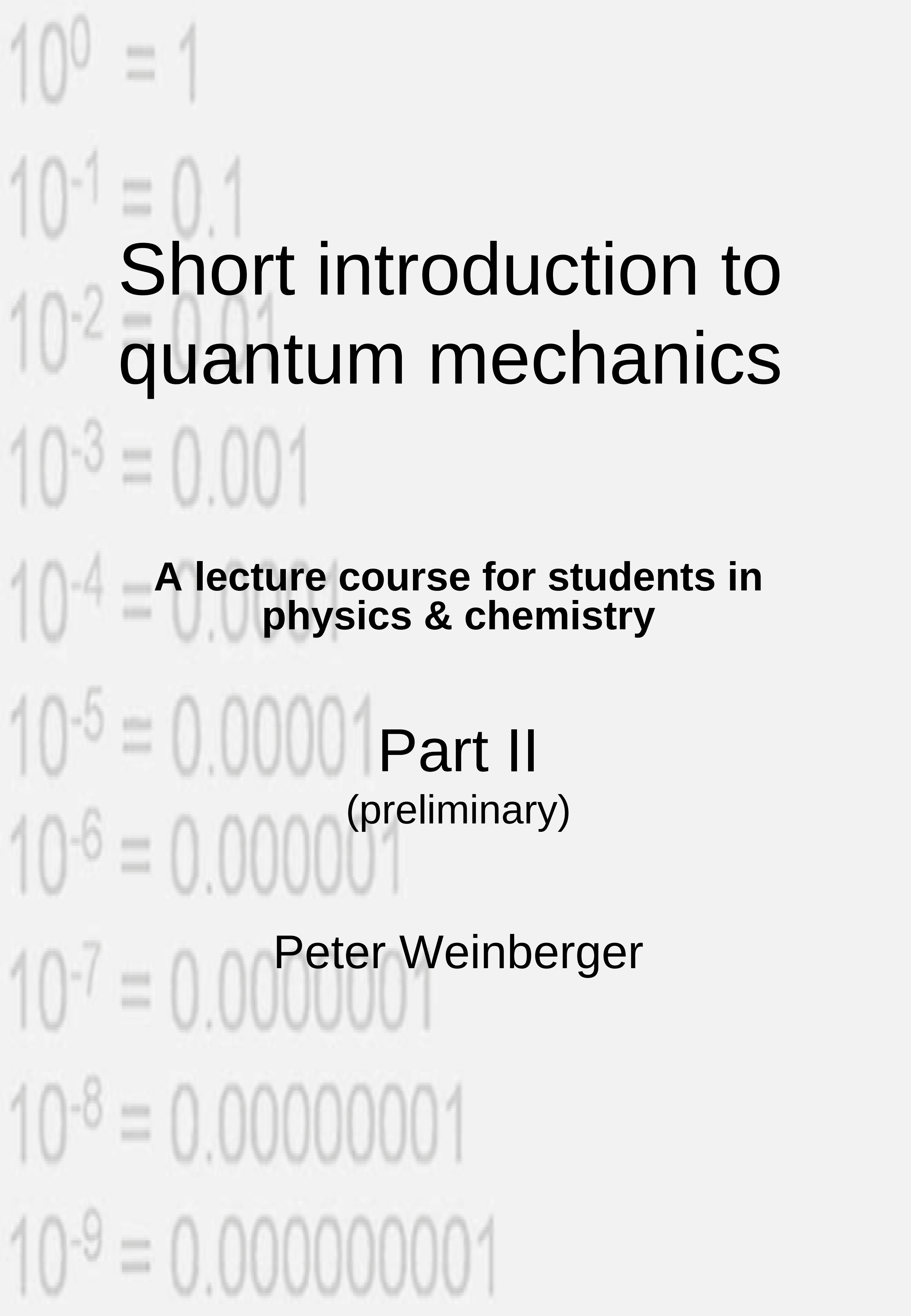


Short introduction to quantum mechanics

A lecture course for students in physics & chemistry

Part II

(preliminary)

Peter Weinberger

Contents

List of Figures

11 The harmonic oscillator

Suppose the Hamilton operator is given by

$$\widehat{H}(x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) \quad , \tag{1}$$

where the potential energy is proportional to the square of the distance from the origin, i.e., corresponds to the force of a spring

$$V(x) = \frac{1}{2}kx^2 \quad , \tag{2}$$

with k being the so-called force constant. In terms of the frequency ω , $\omega = \sqrt{k/m}$, the potential energy therefore can also be written as

$$V(x) = \frac{m\omega^2}{2}x^2 \quad . \tag{3}$$

Using now the following coordinate transformation

$$q = \sqrt{\alpha}x \quad , \alpha = \frac{\omega m}{\hbar} \quad \rightarrow \quad \frac{d^2}{dx^2} = \alpha\frac{d^2}{dq^2} \quad , \tag{4}$$

the Hamilton operator reduces to

$$\widehat{H}(q) = \frac{\hbar\omega}{2}\left\{-\frac{d^2}{dq^2} + q^2\right\} \quad . \tag{5}$$

In the corresponding Schrödinger equation,

$$\left\{-\frac{d^2}{dq^2} + q^2\right\}\psi(q) = \widehat{h}(q)\psi(q) = \frac{2E}{\hbar\omega}\psi(q) = \epsilon\psi(q) \quad , \tag{6}$$

the parameter ϵ is a dimensionless quantity.

11.1 Creation- and annihilation operators

Rather than using "traditional ways" to solve this differential equation in terms of an asymptotic solution and a polynomial, the properties of the momentum operator $\widehat{p}$,

$$\widehat{p} = -i\frac{\partial}{\partial q} \quad , \tag{7}$$

shall be exploited, since according to the uncertainty relation

$$[\widehat{p}, q]_- = \frac{1}{\hbar}[\widehat{p}_x, x]_- = -i \quad \rightarrow \; i[\widehat{p}, q]_- = 1 \quad , \tag{8}$$

and therefore

$$(q + i\widehat{p})(q - i\widehat{p}) = q^2 + \widehat{p}^2 + i[\widehat{p}, q]_- \quad , \tag{9}$$

$$(q-i\widehat{p})(q+i\widehat{p}) = q^2+\widehat{p}^2 - i\,[\widehat{p},q]_- \quad . \tag{10}$$

Using the following abbreviations

$$\widehat{a} = \frac{1}{\sqrt{2}}(q+i\widehat{p}) = \frac{1}{\sqrt{2}}(q+\frac{\partial}{\partial q}) \quad , \tag{11}$$

$$\widehat{a}^\dagger = \frac{1}{\sqrt{2}}(q-i\widehat{p}) = \frac{1}{\sqrt{2}}(q-\frac{\partial}{\partial q}) \quad , \tag{12}$$

the operator $\widehat{h}(q)$ in (6) is given by

$$\widehat{h}(q) = \widehat{p}^2 + q^2 = \widehat{a}\widehat{a}^\dagger + \widehat{a}^\dagger\widehat{a} \tag{13}$$

and the Schrödinger equation therefore by

$$\left\{\widehat{a}\widehat{a}^\dagger + \widehat{a}^\dagger\widehat{a}\right\}\psi(q) = \epsilon\psi(q) \quad . \tag{14}$$

It is easy to show that the commutator of $\widehat{a}$ and $\widehat{a}^\dagger$ is unity,

$$\left[\widehat{a},\widehat{a}^\dagger\right]_- = \widehat{a}\widehat{a}^\dagger - \widehat{a}^\dagger\widehat{a} = 1 \quad , \tag{15}$$

since only (10) has to be subtracted from (9) and the result compared with (8). Multiplying now the commutator from the right with $\psi(q)$,

$$\left\{\widehat{a}\widehat{a}^\dagger - \widehat{a}^\dagger\widehat{a}\right\}\psi(q) = \psi(q) \quad . \tag{16}$$

one can see that the obtain equation is of similar form as the Schrödinger equation (14). Adding and subtracting these two equations (14 - 16) yields the following two equations,

$$\widehat{a}\widehat{a}^\dagger\psi(q) = (\frac{\epsilon}{2}+\frac{1}{2})\psi(q) \quad , \tag{17}$$

$$\widehat{a}^\dagger\widehat{a}\psi(q) = (\frac{\epsilon}{2}-\frac{1}{2})\psi(q) \quad , \tag{18}$$

from which one immediately can read off that $\widehat{a}^\dagger\widehat{a}$ and $\widehat{a}^\dagger\widehat{a}$ have the same eigenfunctions as $\widehat{h}(q)$, however according to the eigenvalues $(\frac{\epsilon}{2}-\frac{1}{2})$ and $(\frac{\epsilon}{2}+\frac{1}{2})$, respectively.

Multiplying (17) from the left with $\widehat{a}^\dagger$,

$$\underbrace{\widehat{a}^\dagger\widehat{a}}_{=(\widehat{a}\widehat{a}^\dagger-1)}\;\widehat{a}^\dagger\psi(q) = \widehat{a}\widehat{a}^\dagger\widehat{a}^\dagger\psi(q) - \widehat{a}^\dagger\psi(q) = (\frac{\epsilon}{2}+\frac{1}{2})\widehat{a}^\dagger\psi(q) \quad , \tag{19}$$

one easily can see that also $\widehat{a}^\dagger\psi(q)$ is an eigenfunction of $\widehat{h}(q)$, but corresponding to the eigenvalue $(\frac{\epsilon}{2}+\frac{3}{2})$:

$$\widehat{a}\widehat{a}^\dagger\left(\widehat{a}^\dagger\psi(q)\right) = (\frac{\epsilon}{2}+\frac{3}{2})\left(\widehat{a}^\dagger\psi(q)\right) \quad . \tag{20}$$

Similarly, by multiplying (18) from the left with $\widehat{a}$,

$$\underbrace{\widehat{a}\widehat{a}^{\dagger}}_{=(\widehat{a}^{\dagger}\widehat{a}+1)}\ \widehat{a}\psi(q) = (\frac{\epsilon}{2} - \frac{1}{2})\widehat{a}\psi(q) \quad , \tag{21}$$

one obtains

$$\widehat{a}^{\dagger}\widehat{a}\,(\widehat{a}\psi(q)) = (\frac{\epsilon}{2} - \frac{3}{2})\,(\widehat{a}\psi(q)) \quad , \tag{22}$$

namely that $\widehat{a}\psi(q)$ is again an eigenfunction of $\widehat{h}(q)$. Repeating these multiplications with $\widehat{a}^{\dagger}$and $\widehat{a}$ from the left n times, one obtains the following remarkable equations:

$$\widehat{a}\widehat{a}^{\dagger}\left((\widehat{a}^{\dagger})^{n}\psi(q)\right) = (\frac{\epsilon}{2} + \frac{1}{2} + n)\left((\widehat{a}^{\dagger})^{n}\psi(q)\right) \quad , \tag{23}$$

$$\widehat{a}^{\dagger}\widehat{a}\,((\widehat{a})^{n}\psi(q)) = (\frac{\epsilon}{2} - \frac{1}{2} - n)\,((\widehat{a})^{n}\psi(q)) \quad . \tag{24}$$

Because of the last two equations $\widehat{a}^{\dagger}$and $\widehat{a}$ are called **step-up** and **step-down operators**, respectively, or more generally **creation** and **annihilation operators**.

Since the expectation value of $\widehat{h}(q)$, $<\widehat{h}(q)>$ has to be positive definite, i.e., cannot be negative,

$$<\widehat{h}(q)> = \int\limits_{-\infty}^{\infty} \psi^{*}(q)\widehat{h}(q)\psi(q)dq = \int\limits_{-\infty}^{\infty} \psi^{*}(q)(\widehat{p}^{2} + q^{2})\psi(q)dq =$$

$$= \underbrace{-\int\limits_{-\infty}^{\infty} \psi^{*}(q)\frac{d^{2}}{dq^{2}}\psi(q)dq}_{partial\ integration} + \int\limits_{-\infty}^{\infty} \psi^{*}(q)q^{2}\psi(q)dq =$$

$$= \underbrace{-\psi^{*}(q)\frac{d\psi(q)}{dq}\mid_{-\infty}^{\infty}}_{=0} + \int\limits_{-\infty}^{\infty}\left(\frac{d\psi^{*}(q)}{dq}\right)\left(\frac{d\psi(q)}{dq}\right)dq + \int\limits_{-\infty}^{\infty} \psi^{*}(q)\psi(q)q^{2}dq \quad ,$$

$$<\widehat{h}(q)> = \int\limits_{-\infty}^{\infty}\left\{\underbrace{\left|\frac{d\psi(q)}{dq}\right|^{2} + |\psi(q)|^{2}\,q^{2}}_{\geq\,0}\right\}dq \geq 0 \quad , \tag{25}$$

the annihilation operator $\widehat{a}$ can only be applied k times, because otherwise the eigenvalue

$$\frac{\epsilon}{2} - \frac{2k+1}{2} = \frac{\epsilon}{2} - \frac{1}{2} - k \tag{26}$$

becomes negative and therefore would correspond to a negative expectation value of $\widehat{h}(q)$. If one denotes the wavefunction that belongs to the smallest possible eigenvalue by $\psi_0(q)$, then obviously the following relation has to be valid

$$\widehat{a}\psi_0(q) = 0 \quad . \tag{27}$$

Multiplying now from the left with the creation operator $\widehat{a}^\dagger$, one easily can see that

$$\widehat{a}^\dagger(\widehat{a}\psi_0(q)) = (\widehat{a}^\dagger\widehat{a})\psi_0(q) = \widehat{a}^\dagger(0) = 0\psi_0(q) = 0 \quad . \tag{28}$$

According to (18), however

$$(\widehat{a}^\dagger\widehat{a})\psi_0(q) = (\frac{\epsilon_0}{2} - \frac{1}{2})\psi_0(q) \quad ,$$

which implies that (28) can only be the case if

$$\epsilon_0 = 1 \quad , \tag{29}$$

because $\psi_0(q) = 0$, $\forall q$, is not an acceptable solution, since the norm is identically zero! The lowest energy eigenvalue of $\widehat{H}(q)$, E_0, (see (6)) is therefore given by

$$E_0 = \frac{\hbar\omega}{2}\epsilon_0 = \frac{\hbar\omega}{2} \quad . \tag{30}$$

11.2 Eigenfunctions

The eigenfunction $\psi_0(q)$ can be obtained by considering the explicit form of $\widehat{a}$ (11), which leads to the following differential equation

$$\widehat{a}\psi_0(q) = \frac{1}{\sqrt{2}}\left(\frac{d\psi_0(q)}{dq} + q\psi_0(q)\right) = 0 \quad , \tag{31}$$

whose solution can be guessed immediately:

$$\psi_0(q) = c_0 \exp(-q^2/2) \quad , \tag{32}$$

since

$$\frac{d\psi_0(q)}{dq} = -c_0 q \exp(-q^2/2) \quad . \tag{33}$$

The normalization constant c_0 follows from the normalization integral

$$\int\limits_{-\infty}^{\infty} \psi_0(q)^*\psi_0(q)dq = c_0^2 \int\limits_{-\infty}^{\infty} \exp(-q^2)dq \quad , \quad \rightarrow c_0 = \pi^{-1/4} \quad . \tag{34}$$

All other eigenfunctions, namely those belonging to higher eigenvalues of $\widehat{H}(q)$ can then be obtained by successive application of the creation operator $\widehat{a}^\dagger$. For $\psi_1(q)$, for example one gets

$$\psi_1(q) = \frac{c_1}{c_0}\widehat{a}^\dagger\psi_0(q) = c_1\widehat{a}^\dagger \exp(-q^2/2) =$$

$$
\begin{aligned}
&= c_1 \left[\frac{1}{\sqrt{2}} (q - \frac{d}{dq}) \right] \exp(-q^2/2) = \\
&= \frac{c_1}{\sqrt{2}} \left[q \exp(-q^2/2) + q \exp(-q^2/2) \right] = \\
&= \sqrt{2} c_1 q \exp(-q^2/2) \quad .
\end{aligned}
\tag{35}
$$

In general therefore the n-th eigenfunction $\psi_n(q)$ is given by

$$
\begin{aligned}
\psi_n(q) &= \frac{c_n}{c_{n-1}} \widehat{a}^\dagger \psi_{n-1}(q) = \frac{c_n}{c_{n-2}} \left(\widehat{a}^\dagger \right)^2 \psi_{n-2}(q) = \quad ... \quad = \\
&= \frac{c_n}{c_0} \left(\widehat{a}^\dagger \right)^n \psi_0(q) \quad ,
\end{aligned}
\tag{36}
$$

or using the explicit form of $\widehat{a}^\dagger$ by

$$
\psi_n(q) = \frac{c_n}{\sqrt{2^n}} \left(q - \frac{d}{dq} \right)^n \exp(-q^2/2) \quad . \tag{37}
$$

This last equation can be rewritten by collecting the constant factors in N_n and multiplying with unity from the left hand side in the following sense:

$$
\begin{aligned}
\psi_n(q) &= N_n \underbrace{\exp(-q^2/2) \exp(q^2/2)}_{=1} \left(q - \frac{d}{dq} \right)^n \exp(-q^2/2) = \\
&= N_n \exp(-q^2/2) \underbrace{\exp(q^2/2) \left(q - \frac{d}{dq} \right)^n \exp(-q^2/2)}_{=H_n(q)} = \\
&= N_n \exp(-q^2/2) H_n(q) \quad .
\end{aligned}
\tag{38}
$$

It can be shown that the explicit evaluation of the normalization constants N_n yields

$$
N_n = \frac{c_n}{\sqrt{2^n}} = \left(2^n n! \sqrt{\pi} \right)^{-1/2} \quad . \tag{39}
$$

The functions $H_n(q)$,

$$
H_n(q) = \exp(q^2/2) \left(q - \frac{d}{dq} \right)^n \exp(-q^2/2) \quad , \tag{40}
$$

are the famous **Hermite polynomials**, which obey the following **recursion relations**

$$
H_{n+1}(q) = 2q H_n(q) - H'_n(q) \quad , \tag{41}
$$

where

$$
H'_n(q) = \frac{d}{dq} H_n(q) = 2n H_{n-1}(q) \quad . \tag{42}
$$

As can be seen from the table below for the first few values of n the Hermite polynomials are very simple polynomials in q:

n	$H_n(q)$
0	1
1	$2q$
2	$4q^2 - 2$
3	$8q^3 - 12q$

Suppose P is the inversion operator, $Pq = -q$, then from the above table, but also from the general form of the Hermite polynomials, the symmetry can be read off quite easily, namely

$$\widehat{P}\psi_n(q) = c_n \exp(-q^2/2) H_n(-q) = (-1)^n c_n \exp(-q^2/2) H_n(q) =$$

$$= (-1)^n \psi_n(q) \quad . \tag{43}$$

For the eigenfunctions $\psi_n(q)$ themselves compact relations can be given in terms of the creation- and annihilation operators

$$\boxed{\widehat{a}^\dagger \psi_n(q) = \sqrt{n+1}\psi_{n+1}(q)} \tag{44}$$

$$\boxed{\widehat{a}\psi_n(q) = \sqrt{n}\psi_{n-1}(q)} \tag{45}$$

as can be exemplified easily from the first two functions

$$\frac{1}{c_1}\psi_1(q) = \frac{1}{c_0}\widehat{a}^\dagger \psi_0(q) = \left[\frac{1}{\sqrt{2}}(q - \frac{d}{dq})\right] \exp(-q^2/2) =$$

$$= \underbrace{\sqrt{2}}_{=\sqrt{n+1}} q \exp(-q^2/2) \quad . \tag{46}$$

It should be noted that the eigenfunctions are products of the exponential function of the argument $-q^2/2$ and a polynomial of order n (Hermite polynomial). They have therefore n nodes (zero locations). Quite clearly eigenfunctions belonging to different eigenvalues $\widehat{h}(q)$, $\widehat{a}\widehat{a}^\dagger$ or $\widehat{a}^\dagger\widehat{a}$ are orthogonal (orthonormal):

$$\int_{-\infty}^{\infty} \psi_n(q)\psi_m(q) dq = \delta_{nm} \quad . \tag{47}$$

For the first four eigenvalues the corresponding eigenfunctions are displayed in *Figure 17*. From this figure one can in particular read off very easily the inversion symmetry for the eigenfunctions (43). Note that the parabola in this figure indicates the potential $V(x) = \frac{1}{2}kx^2$.

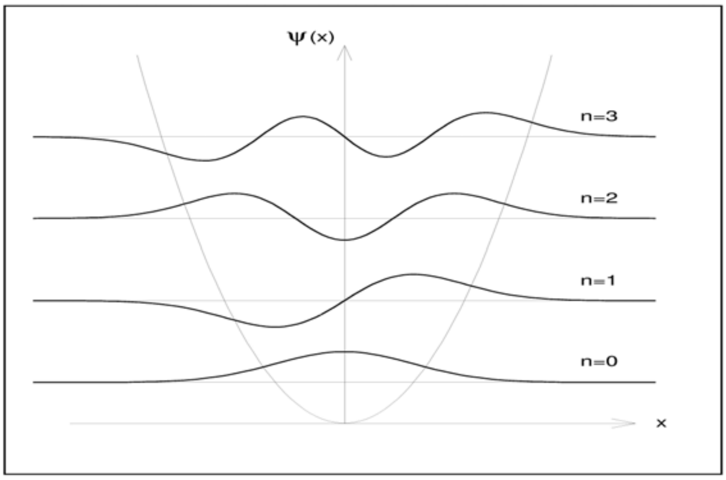


Figure 17: Eigenfunctions of the harmonic oscillator

11.3 Eigenvalues

Finally, the eigenvalues E_n of $\widehat{H}(q)$ can be obtained from the condition that $\epsilon_0 = 1$ (see (30)) and from the eigenvalue equation of $\widehat{a}\widehat{a}^\dagger$, see (17) and (20),

$$\widehat{a}\widehat{a}^\dagger \psi_n(q) = (\frac{\epsilon_n}{2} + \frac{1}{2})\psi_n(q) \quad , \tag{48}$$

$$\widehat{a}\widehat{a}^\dagger (\widehat{a}^\dagger \psi_{n-1}(q)) = (\frac{\epsilon_{n-1}}{2} + \frac{1}{2} + 1)(\widehat{a}^\dagger \psi_{n-1}(q)) \quad , \tag{49}$$

from which directly follows that

$$(\frac{\epsilon_n}{2} + \frac{1}{2}) = (\frac{\epsilon_{n-1}}{2} + \frac{1}{2} + 1) \quad , \tag{50}$$

or

$$\epsilon_n = \epsilon_{n-1} + 2 = \epsilon_0 + 2n = 1 + 2n \quad , \tag{51}$$

i.e.,

$$\boxed{E_n = \tfrac{\hbar\omega}{2}\epsilon_n = \hbar\omega(n + \tfrac{1}{2})} \tag{52}$$

11.4 Selection rules

The harmonic oscillator quite frequently serves in infrared spectroscopy (IR-Spectroscopy) as the most simplest model of interpretation. For a molecule such as

$$ {}^{R}_{R}\rangle \; C = O \quad ,$$

where R denotes a rest such as CH_3- or C_6H_5-, the oscillation of the oxygen constituent in direction of the C-O bond can be considered to be a harmonic oscillation. According to chapter 10, the transition probability for an (absorptive) excitation is proportional to

$$P_{nm} \sim |d_{nm}|^2 \, \delta(E_n - E_m - \hbar\omega) \quad , \tag{53}$$

where $E_n - E_m$ is the energy difference between two eigenvalues of the harmonic oscillator and d_{nm} is the expectation value of the electric dipole moment $d = eq$ (e , elementary electric charge),

$$d_{nm} \equiv \langle d_{nm} \rangle = \int\limits_{-\infty}^{\infty} \psi_n^*(q) e q \psi_m(q) dq \quad . \tag{54}$$

Omitting the normalizations for the eigenfunctions $\psi_m(q)$ and $\psi_n(q)$,

$$\psi_i(q) \sim \exp(-q^2/2) H_i(q) \ ; \quad i = n, m \quad ,$$

the electric dipole moment can easily be calculated using the recursion formulae for the Hermite polynomials in (41) and (42)

$$\boxed{q H_m(q) = \tfrac{1}{2} H_{m+1}(q) + m H_{m-1}(q)} \tag{55}$$

$$\begin{aligned} \langle d_{nm} \rangle \ &\sim \int\limits_{-\infty}^{\infty} \exp(-q^2) H_n(q) \frac{1}{2} H_{m+1}(q) dq + \\ &+ \int\limits_{-\infty}^{\infty} \exp(-q^2) H_n(q) m H_{m-1}(q) dq = \\ = \frac{1}{2} \int\limits_{-\infty}^{\infty} &\psi_n^*(q) \psi_{m+1}(q) dq + m \int\limits_{-\infty}^{\infty} \psi_n^*(q) \psi_{m-1}(q) dq \quad . \end{aligned} \tag{56}$$

Since eigenfunctions to different eigenvalues of $\widehat{H}(q)$ are orthogonal, this implies that

$$\langle d_{nm} \rangle \ = \{ \begin{array}{ll} = 0 \ ; & |n - m| \neq 1 \\ \neq 0 \ ; & otherwise \end{array} \quad . \tag{57}$$

The selection rules for the harmonic oscillator are therefore simply given by

$$\boxed{\Delta n = n - m = \pm 1} \tag{58}$$

and hence the "allowed" energy difference is given by

$$\boxed{\Delta E = E_{n+1} - E_n = \{(n + 1 + 1/2) - (n + 1/2)\} \hbar\omega = \hbar\omega} \tag{59}$$

References

[1] E.Schrödinger, Ann.Physik 79, 489-527 (1926)

[2] L.D.Landau and E.M.Lifschitz, *"Lehrbuch der Theoretischen Physik"*, Volume III, *"Quantenmechanik"*, Akademie Verlag, 1974

12 Rotational motions - the rigid and the non-rigid rotator

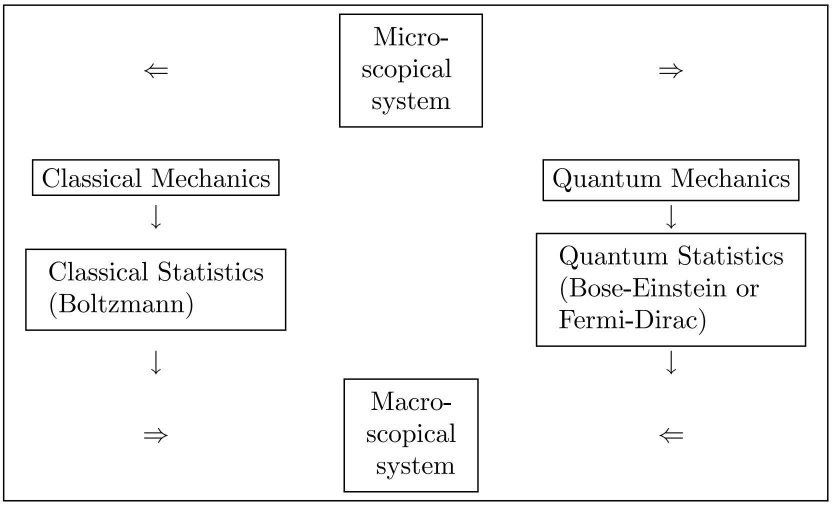
References

[1] E.Schrödinger, Ann.Physik 79, 489-527 (1926)

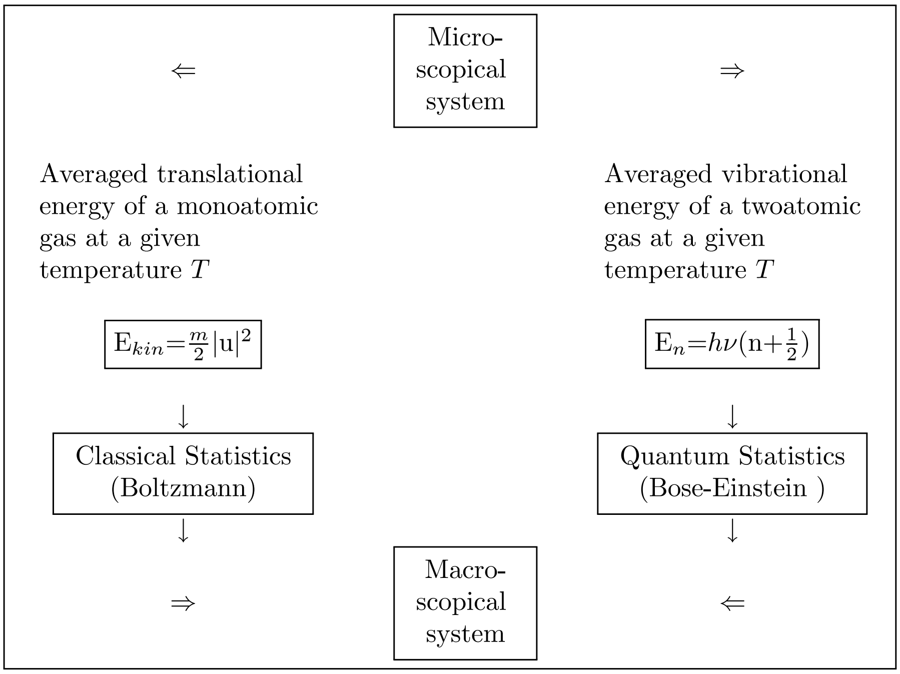
13 Quantum statistics and statistical thermodynamics

13.1 Introduction

Phenomenological thermodynamics provides descriptions of a physical system (gas, liquid etc.) based on **empirical laws**. **Statistical thermodynamics** describes properties of a macroscopical systems in terms of the properties or the interactions of its microscopical parts (particles). This implies that (1) the microscopical properties have to be determined and (2) formal descriptions have to be developed in order to map microscopical variables onto their macroscopical counterparts. Since the number N of particles of a macroscopical system is very large, namely of the order of 10^{23}, it is impossible, but also unnecessary, to give a one-to-one mapping between the variables in the microcosmos and the macrocosmos. These mappings have to based on to averages or probability relations. Averages over a macroscopical system, however, can only be performed if the determination of the microscopical properties is adequate. In principle one could think of the following scheme of descriptions



An example is given in the following table, where m is the mass and u the velocity of the particles and E_n the n-th energy eigenvalue of a harmonic oscillator:


Micro-
scopical
system
⇐
⇒
Averaged translational energy of a monoatomic gas at a given temperature T
Averaged vibrational energy of a twoatomic gas at a given temperature T
$E_{kin}=\frac{m}{2}|u|^2$
$E_n=h\nu(n+\frac{1}{2})$
Classical Statistics
(Boltzmann)
Quantum Statistics
(Bose-Einstein)
⇒
⇐
Macro-
scopical
system

13.2 A short reminder of the theory of probabilities

13.2.1 Relative frequencies

Let $S_1, S_2, \dots, S_n$ be the results of n individual measurements,

$$M_n = \{S_1, S_2, \dots, S_n\} \quad , \tag{1}$$

and $E_1, E_2, \dots, E_m$ incidences in M_n. These incidences are subsets of M_n, $E_i \subset M_n$. If $|E_i|$ and $|M_n|$ denote the number of elements in these sets, then

$$h_n(E_i) = \frac{|E_i|}{|M_n|} \tag{2}$$

denotes the **relative frequency** (occurrence) of the incident E_i in M_n. Quite clearly such a relative frequency is determined by $|M_n|$. The **probability**

$$p(E_i)$$

for the incidence E_i is then defined by the following limiting procedure

$$p(E_i) = \lim_{n\to\infty} h_n(E_i) \quad . \tag{3}$$

The two possible trivial cases, namely $E_i = \{0\}$ (empty set) and $E_i = M_n$ imply immediately that

$$0 \le p(E_i) \le 1 \quad . \tag{4}$$

If $p(E_i) > p(E_j)$ then the incident E_i is more probable than the incident E_j, if $p(E_i) = p(E_j)$ then these two incidents have the same probability.

Consider for example that throwing a dice 20 times yields the following results

$$M_{20} = \{2, 6, 2, 2, 3, 1, 5, 4, 2, 2, 3, 5, 4, 6, 1, 6, 6, 4, 3, 5\} \quad .$$

The incidence to throw a one, a two etc. was then

$$\begin{array}{ll} E_1 = \{1, 1\} & E_4 = \{4, 4, 4\} \\ E_2 = \{2, 2, 2, 2, 2\} & E_5 = \{5, 5, 5\} \\ E_3 = \{3, 3, 3\} & E_6 = \{6, 6, 6, 6\} \end{array}$$

or to have only even or odd numbers by

$$\begin{array}{l} A_1 = \{3, 1, 5, 3, 1, 3, 5\} \\ A_2 = \{2, 6, 2, 2, 4, 2, 2, 4, 6, 6, 6, 4\} \end{array} \quad .$$

The relative frequency of a one was $|E_1| \, / \, |M_{20}| = 1/10$ etc.

Per definition the direct sum of two incidents $E_i \cup E_j$ is an incident that either the incident E_i occurs or the incident E_j. From the above example one immediately can see that

$$A_1 = E_1 \cup E_3 \cup E_5 \quad , \qquad A_2 = E_2 \cup E_4 \cup E_6 \quad .$$

The probability for the sum of two independent incidents E_1 and E_2, $p(E_1+E_2)$ is therefore given by

$$p(E_1+E_2) = \lim_{n\to\infty} \frac{|E_1|+|E_2|}{|M_n|} = \lim_{n\to\infty} \frac{|E_1|}{|M_n|} + \lim_{n\to\infty} \frac{|E_2|}{|M_n|} = \\ = p(E_1) + p(E_2) \quad , \tag{5}$$

i.e., for two independent incidents E_i and E_j the individual probabilities are additive

$$p(E_i+E_j) = p(E_i) + p(E_j) \quad . \tag{6}$$

Finally if the direct sum over all independent incidents E_i is denotes by

$$\bigcup_{i=1}^{m} E_i = M_n \quad , \tag{7}$$

then from (4) follows immediately that

$$p(E_1+E_2+\cdots+E_m) = p(M_n) = \sum_{i=1}^{m} p(E_i) = 1 \quad . \tag{8}$$

13.2.2 Probability functions...

...of discrete variables Let $E_1, E_2, \cdots E_n$ be characteristic "states" of one and the same object that can be mapped onto the numbers $x_1, x_2\ , \cdots x_n$ and let p_1, $p_2, \cdots p_n$ denote the probabilities for the occurrence of these numbers. If there exists a function defined as follows

$$p_k = W(x_k) \quad , \tag{9}$$

then x_k is called a **stochastic variable** and W the **probability function**.

...of continuous variables Consider now for a moment a fortune wheel whose hand can stop anywhere between 0^0 and 360^0. In terms of relative frequencies the probability for one particular position of the hand is identically zero. If, however, one asks for the probability of a position between 180^0 and $180^0+\Delta x$, then this probability is given by

$$p = \frac{\Delta x}{360} = W(180)\Delta x \quad . \tag{10}$$

If therefore $dW(x)$ denotes the probability for the occurrence of a variable of coincidence in the interval x and $x+dx$ then $dW(x)$ is given by

$$dW(x) = w(x)dx \quad , \tag{11}$$

where $w(x)$ is the so-called **probability density**. In analogy to the definition of relative frequencies in (9) the probability density is normalized to unity by

$$\int_{\{x\}} w(x)dx = 1 \quad , \tag{12}$$

where $\{x\}$ denotes the range of x.

13.2.3 Products of probabilities

The product AB of two incidences A and B is an incidence AB where A as well as B occur. If the incidences A and B are "statistically" independent, i.e., the probability of A is independent of the probability of B, then

$$p(AB) = p(A)p(B) \quad . \tag{13}$$

Similar definitions apply for probability densities.

13.2.4 Averages ...

...of variables For discrete variables averages are defined by

$$\overline{x} = \langle x \rangle = \sum_{k=1}^{N} x_k p_k \quad , \tag{14}$$

whereas for continuous variables they are given by

$$\overline{x} = \langle x \rangle = \int\limits_{\{x\}} w(x) x dx \quad . \tag{15}$$

The so-called standard deviation Δx is then a "traditional" measure for the deviation of a single value of x from the corresponding average $\overline{x}$,

$$\Delta x = \sqrt{\overline{(\overline{x} - x)^2}} \quad , \tag{16}$$

whereas

$$\Delta x = \sqrt{\overline{x}^2 - x^2} \quad , \tag{17}$$

is called a "fluctuation".

...of functions Let $F(x_1, x_2, \ldots, x_n)$ be an arbitrary function of the continuous stochastic variables $x_1, x_2, \ldots, x_n$ and let $w(x_1, x_2, \ldots, x_n)$ be the probability density for the product of the corresponding single incidences $w(x_1), w(x_2), \ldots w(x_n)$. The average of F, $\overline{F}$ is then given by

$$\overline{F} = \int\limits_{\{x_1, x_2, \ldots, x_n\}} \cdots \int F(x_1, x_2, \ldots, x_n) w(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n \quad . \tag{18}$$

This equation can be written also in the following form

$$\overline{F} = \int\limits_{F} F W(F) dF \quad , \tag{19}$$

where $W(F)$ is the probability density for the case that $F(x_1, x_2, \ldots, x_n)$ assumes exactly the value F,

$$W(F) = \int\limits_{\{x_1, x_2, \ldots, x_n\}} \cdots \int \delta\left(F - F(x_1, x_2, \ldots, x_n)\right) \times$$

$$\times\; w(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n \quad , \tag{20}$$

and $\delta(x-y)$ is the Dirac distribution- (δ-) function:

$$f(y) = \int\limits_{-\infty}^{\infty} f(x)\delta(x-y)dx \quad , \quad \int\limits_{-\infty}^{\infty} \delta(x)dx = 1 \quad . \tag{21}$$

13.3 Kinetic gas theory

In many aspects the ("classical") kinetic theory of a gas of hard spheres was the precursor to the atomistic picture that in the end lead to the famous Boltzmann distribution. It is therefore worthwhile as a start to recall the main features of this theory.

13.3.1 The "molecular" distribution function

Let

$$f(\mathbf{r},\mathbf{u},t)\underbrace{dxdydz}_{d\mathbf{r}}\underbrace{du_x du_y du_z}_{d\mathbf{u}} \tag{22}$$

be the number of molecules (atoms, hard spheres), which are at the time t at $\mathbf{r}$ in the volume element $d\mathbf{r}$ and whose component velocities are in the intervals $[u_x, u_x + du_x]$, $[u_y, u_y + du_y]$ and $[u_z, u_z + du_z]$. The six-dimensional space of the coordinates x, y, z, u_x, u_y, u_z is called phase space. Integrating over all velocities,

$$\iiint\limits_{\{u_x,u_y,u_z\}} f(\mathbf{r},\mathbf{u},t)d\mathbf{r}d\mathbf{u} = n(\mathbf{r},t)d\mathbf{r} \ , \tag{23}$$

yields the so-called **particle density**, namely the number of (velocity independent) particles which at the time t are located within the volume element $d\mathbf{r}$. Integration over the volume element $d\mathbf{r}$,

$$\iiint n(\mathbf{r},t)d\mathbf{r} = N(t) \ , \tag{24}$$

determines the total number of particles at a given time t. Quite clearly the molecular distribution function $f(\mathbf{r},\mathbf{u},t)$ is not normalized to unity, but to the number of particles. In the following only equilibrium situations shall be considered. Obviously the individual gas molecules can move around quite a bit, the averaged number of particles per volume, however, shall be assumed to be not time dependent ("**steady state approximation**", "**stationary state**"), i.e., to be a constant.

13.3.2 Atomistic picture of the pressure of an ideal gas

Suppose one wants to calculate the pressure p exerted by a monoatomic gas on the walls of the container. Phenomenologically the averaged force $\overline{F}_\perp$ perpendicular to a surface element ΔA is given by

$$\overline{F}_\perp = p\Delta A \ . \tag{25}$$

This force can also be viewed as a time average in the interval $[t, t + \Delta t]$ of the forces exerted by the individual molecules hitting the wall:

$$p\Delta A = (\overline{F}_{\perp})_{gas\ on\ wall} = \frac{1}{\Delta t}\int\limits_{\Delta t} (F_{\perp})_{gas\ on\ wall}\, dt \quad . \tag{26}$$

If $p_{\perp}$ denotes the momentum of a gas particle perpendicular to ΔA, then the force the wall exerts on to this particle is given by

$$(F_{\perp})_{wall\ on\ gas} = \frac{dp_{\perp}}{dt}\ , \tag{27}$$

and oppositely the force exerted by the gas on the wall by

$$(F_{\perp})_{gas\ on\ wall} = -\frac{dp_{\perp}}{dt}\ , \tag{28}$$

The equilibrium obviously is given by

$$(\overline{F}_{\perp})_{wall\ on\ gas} = -\ (\overline{F}_{\perp})_{gas\ on\ wall}\ . \tag{29}$$

Combining therefore (26) and (28) one gets

$$p\Delta A\Delta t = \int\limits_{\Delta t}\left(-\frac{dp_{\perp}}{dt}\right) dt = (p_{\perp})_1 - (p_{\perp})_2 = -\Delta p_{\perp} \quad , \tag{30}$$

where $(p_{\perp})_1$ is the perpendicular component of the momentum of the particle before the collision with the wall and $(p_{\perp})_2$ after the collision. From *Figure 18* one can see that for one particular velocity $\mathbf{u}$, whose direction forms an angle θ, $0 \leq \theta \leq \pi/2$, with the surface normal, all those particles with velocity $\mathbf{u}$ that are within the length $|\mathbf{u}|\,\Delta t$ from the wall will hit the wall within the time interval Δt, i.e., one can see that all those molecules that are within the volume element $dV = \Delta A\,|\mathbf{u}|\,\Delta t \cos\theta$ will arrive at the wall within Δt. If one assumes now a distribution function of velocities, $f(\mathbf{u})$, the number of particles actually hitting the wall is given by

$$f(\mathbf{u})\underbrace{\Delta A\,|\mathbf{u}|\,\Delta t\cos\theta d}_{dV}\mathbf{u}\ . \tag{31}$$

Let m denote the mass of the particles then the momentum perpendicular to the wall for one particular particle is given by $m\,|\mathbf{u}|\cos\theta$ and the total momentum therefore by this momentum times the total number of particles, namely

$$m\,|\mathbf{u}|\cos\theta\ \times\ f(\mathbf{u})\Delta A\,|\mathbf{u}|\,\Delta t\cos\theta d\mathbf{u}\ = m\Delta t\Delta A\left(|\mathbf{u}|^2\cos^2\theta\right) f(\mathbf{u})d\mathbf{u} \quad .$$

The total momentum before the collision ($(p_{\perp})_1$, (30)) is therefore given by

$$(p_{\perp})_1 =\ m\Delta t\Delta A\iiint |\mathbf{u}|^2\cos^2\theta f(\mathbf{u})d\mathbf{u} \quad ; 0 \leq \theta \leq \pi/2\ , \tag{32}$$

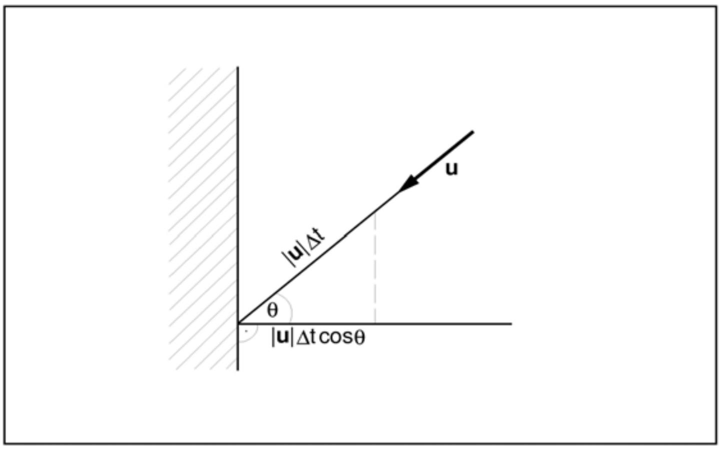


Figure 18: Atomistic picture of the pressure of an ideal gas

whereas after the collision with the wall the individual momentum is given by $-\ m\,|\mathbf{u}|\cos\theta$ and the total momentum by

$$(p_\perp)_2 = -\ m\Delta t\Delta A \iiint |\mathbf{u}|^2 \cos^2\theta f(\mathbf{u}) d\mathbf{u} \quad ; \pi/2 \le \theta \le \pi \ . \tag{33}$$

Going now back to (30), one can see that by combining the last two equations,

$$p\Delta A\Delta t = (p_\perp)_1 - (p_\perp)_2 =$$

$$= m\Delta A\Delta t \iiint |\mathbf{u}|^2 \cos^2\theta f(\mathbf{u}) d\mathbf{u} \quad ; 0 \le \theta \le \pi \quad , \tag{34}$$

or that the pressure p can be written as

$$p = mn \iiint |\mathbf{u}|^2 \cos^2\theta \underbrace{f(\mathbf{u})n^{-1}} d\mathbf{u} \quad ; 0 \le \theta \le \pi \quad , \tag{35}$$

where n is the particle density as given by the total number of particles N divided by the volume V and N is according to the "steady state approximation" a constant. By including the particle density n to the integral, the quantity $f(\mathbf{u})n^{-1}$ is a probability density, since it is normalized to the "total" particle density, i.e., since

$$0 \le f(\mathbf{u})n^{-1} \le 1 \quad .$$

Denoting now by $\overline{u}_\perp^2$ the average of the square of the velocity in z-direction (direction of the surface normal), then the pressure is given by

$$p = mn\overline{u}_\perp^2 \tag{36}$$

If one finally assumes that the velocity distribution is **isotropical**, i.e., all directions of the velocity, u_x, u_y and u_z, are equally probable,

$$\overline{u}_{\perp}^2 = \overline{u}_x^2 = \overline{u}_y^2 = \overline{u}_z^2 = \frac{1}{3}\overline{(u_x^2 + u_y^2 + u_z^2)} = \frac{1}{3}\overline{u}^2 \quad , \tag{37}$$

one gets

$$p = \frac{1}{3} nm\overline{u}^2 \quad . \tag{38}$$

By denoting the averaged kinetic energy of a gas molecule (particle) by $\overline{\epsilon}_t$,

$$\overline{\epsilon}_t = \frac{m}{2}\overline{u}^2 \quad , \tag{39}$$

for an isotropic velocity distribution the pressure is then defined by

$$p = \frac{2}{3} n\overline{\epsilon}_t = \frac{2}{3}\frac{N}{V}\overline{\epsilon}_t \quad . \tag{40}$$

Since the total number of particles N can be written as a multiple of the **Avogardo** (**Loschmidt**) number L , $N = \nu L$, the pressure can be written also in the following form

$$p = \frac{2}{3}\frac{\nu L}{V}\overline{\epsilon}_t \quad , \tag{41}$$

or alternatively one gets

$$pV = \frac{2}{3}\nu L\overline{\epsilon}_t \quad . \tag{42}$$

If one compares this equation with the famous equation for an ideal gas,

$$pV = vRT \quad , \tag{43}$$

where T is the temperature, one immediately can deduce that

$$\overline{\epsilon}_t = \frac{3}{2}\left(\frac{R}{L}\right)T = \frac{3}{2}kT \quad , \tag{44}$$

or alternatively find an "**atomistic-like**" definition for the temperature

$$T = \frac{2}{3}k^{-1}\overline{\epsilon}_t \quad . \tag{45}$$

In the last two equations k is the famous **Boltzmann constant**, k =1.38.10^{-23} J/K. Using (39) and (44),

$$\overline{u}^2 = \frac{3kT}{m} = \frac{3RT}{mL} = \frac{3RT}{\mu} \quad , \tag{46}$$

where $\mu = mL$ is the molecular weight [g] one easily can calculate the averaged velocity $\sqrt{\overline{u}^2}$ at a given temperature T. For He at T=273 K one gets for example $\sqrt{\overline{u}^2}$ ~1300msec^{-1}.

13.3.3 Maxwell-Boltzmann distribution of velocities

Suppose the velocity distribution $f(\mathbf{r},\mathbf{u},t)$ is rewritten as a product of the particle density $n(\mathbf{r},t)$ and a distribution $F(\mathbf{r},\mathbf{u},t)$ normalized to unity,

$$f(\mathbf{r},\mathbf{u},t) = n(\mathbf{r},t)\underbrace{n(\mathbf{r},t)^{-1}f(\mathbf{r},\mathbf{u},t)} = n(\mathbf{r},t)F(\mathbf{r},\mathbf{u},t) \quad . \tag{47}$$

In equilibrium this new distribution function has to be time- and space independent, i.e., $F(\mathbf{r},\mathbf{u},t) \equiv F(\mathbf{u})$. By imposing the following two conditions,

$$F(\mathbf{u}) = F(|\mathbf{u}|) \equiv F(u) \quad , \tag{48}$$

$$F(u) = F_1(u_x)F_2(u_y)F_3(u_z) \quad , \tag{49}$$

namely that $F(\mathbf{u})$ is isotropical and that the components of the velocity are mutually independent, $F(\mathbf{u})$ can be determined.

From (49) follows immediately that

$$\ln F(u) = \ln F_1(u_x) + \ln F_2(u_y) + \ln F_3(u_z) \quad , \tag{50}$$

which partially differentiated with respect to u_x gives

$$\frac{\partial \ln F(u)}{\partial u_x} = \frac{d\ln F(u)}{du}\frac{\partial u}{\partial u_x} = \frac{u_x}{u}\frac{d\ln F(u)}{du} = \frac{d\ln F_1(u_x)}{du_x} \quad , \tag{51}$$

since $u = \sqrt{u_x^2 + u_y^2 + u_z^2}$. From the similar derivatives with respect to u_y and u_z follows

$$\frac{1}{u}\frac{d\ln F(u)}{du} = \frac{1}{u_x}\frac{d\ln F_1(u_x)}{du_x} = \frac{1}{u_y}\frac{d\ln F_2(u_y)}{du_y} = \frac{1}{u_z}\frac{d\ln F_3(u_z)}{du_z} \quad . \tag{52}$$

These expressions can only be equal to each other if they equal the same constant, say -2γ. For example,

$$\frac{1}{u_x}\frac{d\ln F_1(u_x)}{du_x} = -2\gamma \quad . \tag{53}$$

The solution of this differential equation is simply given by

$$F_1(u_x) = a\exp(-\gamma u_x^2) \quad , \tag{54}$$

where a and γ are constants yet to be determined. The distribution function $F(u)$ is therefore given by

$$F(u) = a^3\exp(-\gamma(u_x^2 + u_y^2 + u_z^2)) = a^3\exp(-\gamma u^2) \quad . \tag{55}$$

Since $F(u)$ is normalized to unity, i.e.,

$$\int F(u)d\mathbf{u} = 1 \quad , \quad (d\mathbf{u} = u^2 du \sin\theta d\theta d\phi) \quad , \tag{56}$$

a^3 has to be $(\gamma/\pi)^{3/2}$.

Finally the constant γ can be evaluated[1] by determining the averaged square of the velocity $\overline{u}^2$,

$$\overline{u}^2 = \int u^2 F(u) d\mathbf{u} = (\gamma/\pi)^{3/2} \int\limits_0^\infty \int\limits_0^\pi \int\limits_0^{2\pi} u^4 \exp(-\gamma u^2) du \sin\theta d\theta d\phi =$$

$$= 4\pi(\gamma/\pi)^{3/2} \int\limits_0^\infty u^4 \exp(-\gamma u^2) du = \frac{3}{2\gamma} \quad . \tag{57}$$

If one associates now this result with that for $\overline{u}^2$ in (46), one gets

$$\gamma = \frac{m}{2kT} \quad , \tag{58}$$

i.e., $F(u)$ is given by

$$F(u) = \Big(\frac{m}{2\pi kT}\Big)^{3/2} \exp(-\frac{mu^2}{2kT}) = \Big(\frac{m}{2\pi kT}\Big)^{3/2} \exp(-\frac{\epsilon_t}{kT}) \quad , \tag{59}$$

where $\epsilon_t = mu^2/2$ is the translational energy. Equation (59) is nothing but the famous **Maxwell-Boltzmann (velocity) distribution**.

13.4 Statistical mechanics

13.4.1 Phase space

For one particle the phase space, that is the collection of space ($\mathbf{r}$)- and momentum ($\mathbf{p}$) coordinates of this particle, is 6 dimensional. For N particles the dimension of the phase space is $6N$. If one considers the phase space to be partioned into **cells** of the volume $d\phi$,

$$d\phi = dxdydzdp_xdp_ydp_z \quad , \tag{60}$$

then each particle (atoms, molecules) has to belong to one cell. If furthermore i numbers these cells, then in each cell the number of particles N_i can be assumed to be much larger than one, $N_i \gg 1$, since $d\phi$ only has to be small as compared to the macroscopical dimensions of the system. The fundamental problem of statistical mechanics is to find out how these numbers N_i can be related to the phase space coordinates.

13.4.2 Microstates and macrostates

As is well-known from quantum mechanics, because of the uncertainty relation[2], the space- and momentum coordinates of a microscopical particle cannot be determined simultaneously exact:

$$\Delta x \Delta p_x = \Delta y \Delta p_y = \Delta z \Delta p_z \ \simeq \ h \quad . \tag{61}$$

---

[1] the integral on the rhs of the following equation leads to a Gamma function

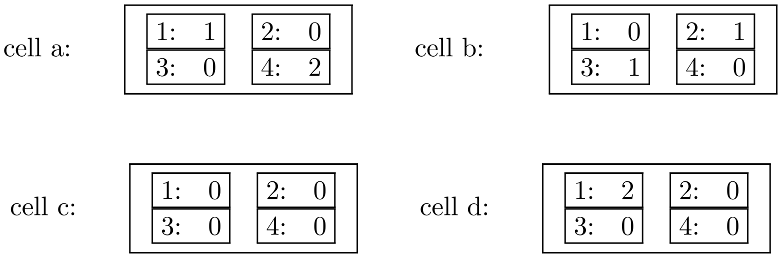
[2] see also the discussion in chapter 1

This implies that the space- and momentum coordinates of such a particle can only be pinned down to the volume h^3 in phase space. Such a volume shall be called a **subcell**. It should be noted that while the volume of a cell, $d\phi$, is arbitrary, the volume of a subcell is fixed to h^3. In general the volume of a cell can be assumed to be large enough so as to contain a large number of subcells. Quite clearly the number of subcells in a particular cell is given by

$$g = \frac{d\phi}{h^3} \quad . \tag{62}$$

The complete **specification of all 6 coordinates** of the **subcell** in which a particular particle is located is called a **microstate**. The **specification** of the **number of particles** in each cell of the phase space is called a **macrostate**.

In order to exemplify these definitions in the following 4 cells, labelled by a,b,c and d are considered below, where each of these cells contains 4 subcells, labelled by 1,2,3 and 4. The number of particles in each subcell is denoted after the semi-colon.



Microstates	Macrostates
one particle: cell a, subcell 1	number of particles in cell a: 3
one particle: cell a, subcell 4	
one particle: cell a, subcell 4	
one particle: cell b, subcell 2	number of particles in cell b: 2
one particle: cell b, subcell 3	
one particle: cell d, subcell 1	*etc*
etc	

Up to now implicitly only monoatomic gases were considered for which it is sufficient to specify the subcell into which a particular atom is located. For molecules for example, one has to deal also with **internal degrees of freedom**, corresponding to vibrational[3] , rotational [4] and electronic energies in addition to the translational energies.

---

[3] see chapter 11

[4] see chapter 12

Thermodynamical probabilities In the following the question will be addressed of how many microstates belong to one given macrostate. Suppose there are only two cells, labelled by i and j, each of them partitioned into 4 subcells. Suppose further that a 4-particle system shall be described. If N_i and N_j denote the number of phasepoints (particles) in these two cells then there are altogether 5 possible macrostates labelled by roman numbers:

$$\underbrace{\overset{i}{\boxed{4}}\quad\overset{j}{\boxed{0}}}_{I}\qquad\underbrace{\overset{i}{\boxed{3}}\quad\overset{j}{\boxed{1}}}_{II}\qquad\underbrace{\overset{i}{\boxed{2}}\quad\overset{j}{\boxed{2}}}_{III}\qquad\underbrace{\overset{i}{\boxed{1}}\quad\overset{j}{\boxed{3}}}_{IV}\qquad\underbrace{\overset{i}{\boxed{0}}\quad\overset{j}{\boxed{4}}}_{V}$$

	I	II	III	IV	V
N_i	4	3	2	1	0
N_j	0	1	2	3	4

The number of microstates that belong to each macrostate is called thermodynamical probability. However, before this number can be determined, one has to distinguish between

bosons: arbitrary number of phase points per cell and

fermions: only two phase points per cell.

The statistics that treats the case of two phase points per cell is called **Fermi-Dirac statistics**, while the other case is the **Bose-Einstein** statistics.

13.4.3 Bose-Einstein statistics

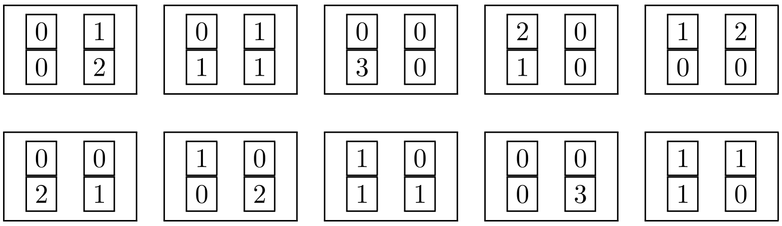
Suppose one starts from the previous example of 4 particles in 2 cells, considers macrostate II

$$\overset{i}{\boxed{3}}\quad\overset{j}{\boxed{1}}$$

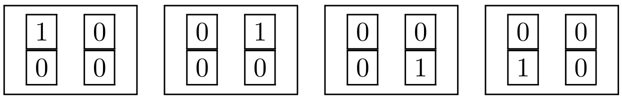
and partitions each cell into 4 subcells. In cell i the three particles can now be distributed into the 4 subcells as follows:

$$\boxed{\begin{array}{cc}3&0\\0&0\end{array}}\quad\boxed{\begin{array}{cc}2&1\\0&0\end{array}}\quad\boxed{\begin{array}{cc}0&2\\0&1\end{array}}\quad\boxed{\begin{array}{cc}1&0\\2&0\end{array}}\quad\boxed{\begin{array}{cc}0&0\\1&2\end{array}}$$

$$\boxed{\begin{array}{cc}1&1\\0&1\end{array}}\quad\boxed{\begin{array}{cc}0&3\\0&0\end{array}}\quad\boxed{\begin{array}{cc}2&0\\0&1\end{array}}\quad\boxed{\begin{array}{cc}0&2\\1&0\end{array}}\quad\boxed{\begin{array}{cc}0&1\\2&0\end{array}}$$



As one can see there are 20 different arrangements (possibilities) to distribute these 3 particles. The thermodynamical probability W_i for cell i is therefore 20. In cell j only one particle has to be distributed in 4 subcells. As one can see from the arrangements below $W_j = 4$.



The product of W_i and W_j, $W = W_i W_j = 80$ is the number of possibilities to arrange macrostate II. In general for an arbitrary number M of cells i with individual probabilities W_i the total probability W is given by

$$W = \prod_{i=1}^{M} W_i \quad . \tag{63}$$

Suppose that the subcells in one particular cell i can be labelled by $1, 2, 3, \ldots, g_i$, whereas the phase points (particles) in this cell are labelled by $a, b, c, \ldots, N_i$. For example

$$\boxed{\text{1ab}} \qquad \boxed{\text{2cde}} \qquad \boxed{\text{3}} \qquad \boxed{\text{4fgh}}$$

implies that in subcell 1 there phase points a and b, in subcell 2 phase points c, d and e, etc. By removing the boxes this gives a **sequence** of numbers and letters that reflects exactly one particular occupation of these 4 subcells. For the present example this sequence is $1ab2cde34fgh$. In general such a sequence has to start with a number between 1 and g_i. The following $(N_i + g_i - 1)$ symbols can then be arranged in an arbitrary manner. The number of permutations is then given by $(N_i + g_i - 1)!$ The number of sequences that start with a number is therefore $g_i(N_i + g_i - 1)!$ All permutations, however, that result from a pairwise interchange of phase points, like in the above examples $1ab$ and $1ba$, represent the same microstate and therefore have to be excluded. There are altogether $N_i!$ such pairwise permutations of letters. Furthermore the numbering of cells has to be in natural order, otherwise $g_i!$ double counting occurs. The total number of microstates in cell i is therefore given by the following individual probability W_i,

$$W_i = \frac{g_i(N_i + g_i - 1)!}{g_i! N_i!} = \frac{(N_i + g_i - 1)!}{(g_i - 1)! N_i!} \quad , \tag{64}$$

and consequently the thermodynamical probability is defined by

$$W = \prod_{i=1}^{M} \frac{(N_i + g_i - 1)!}{(g_i - 1)! N_i!} \quad . \tag{65}$$

By taking the logarithm of W,

$$\ln W = \sum_{i=1}^{M} \left\{ \ln(N_i + \underbrace{g_i - 1}_{\simeq\, g_i})! - \ln(\underbrace{g_i - 1}_{\simeq\, g_i})! - \ln(N_i!) \right\} \quad , \tag{66}$$

and using the Stirling formula,

$$\ln(n!) = n \ln n - n \quad , \quad n \gg 1 \quad , \tag{67}$$

one obtains the following expression,

$$\ln W \simeq \sum_{i=1}^{M} \{ (N_i + g_i) \ln(N_i + g_i) - g_i \ln g_i - N_i \ln N_i \} \quad , \tag{68}$$

where $(g_i - 1)$ was replaced by g_i, since $g_i \gg 1$.

Since particles in motion change their positions in phase space with respect to time, the number of phase points N_i in the various cells also changes with time. Quite clearly, if W has a **maximum** W^0, then also $\ln W$ is at its maximal value. Considering therefore small finite changes δN_i (finite, because the N_i are integer numbers!) and remembering that the number of subcells $g_i = d\phi_i / h^3$ is fixed by the uncertainty relation, such changes results in the following change $\delta \ln W$ for $\ln W$:

$$\begin{aligned} \delta \ln W = \sum_{i=1}^{M} \{ & \ln(N_i + g_i)\delta N_i \; + \underbrace{(N_i + g_i)\delta \ln(N_i + g_i)} \\ & - \ln N_i \delta N_i - \underbrace{N_i \delta \ln N_i} \} \quad . \end{aligned} \tag{69}$$

From the relations

$$\delta \ln N_i = \frac{\delta N_i}{N_i} \quad , \quad \delta \ln(N_i + g_i) = \frac{\delta N_i}{N_i + g_i} \quad , \tag{70}$$

one easily can see that

$$\sum_{i=1}^{M} N_i \delta \ln N_i = \sum_{i=1}^{M} (N_i + g_i) \delta \ln(N_i + g_i) = \sum_{i=1}^{M} \delta N_i \quad , \tag{71}$$

which reduces (69) to

$$\delta \ln W = \sum_{i=1}^{M} \{ \ln(N_i + g_i)\delta N_i \; - \ln N_i \delta N_i \} =$$

$$= \sum_{i=1}^{M} \ln(\frac{N_i + g_i}{N_i}) \delta N_i \quad . \tag{72}$$

The set of N_i for which $\delta \ln W^0 = 0$, shall be denoted by $\{N_i^0\}$,

$$\delta \ln W^0 = \sum_{i=1}^{M} \ln(\frac{N_i^0 + g_i}{N_i^0}) \delta N_i = 0 \quad . \tag{73}$$

and obviously corresponds to the thermodynamical equilibrium. The variation of the N_i, however, is restricted by the following two conditions,

$$\sum_{i=1}^{M} N_i = N = const \quad \rightarrow \quad \sum_{i=1}^{M} \delta N_i = 0 \quad , \tag{74}$$

$$\sum_{i=1}^{M} N_i \epsilon_i = U = const \quad \rightarrow \quad \sum_{i=1}^{M} \delta N_i \epsilon_i = 0 \quad , \tag{75}$$

namely that total number particles and the total energy of the system has to be constant, ϵ_i being the energy of an individual particle. This implies that Lagrange parameters have to be used of the form

$$\lambda_1 \sum_{i=1}^{M} \delta N_i = 0 \quad , \quad \lambda_2 \sum_{i=1}^{M} \delta N_i \epsilon_i = 0 \quad , \tag{76}$$

such that

$$\sum_{i=1}^{M} \left\{ \ln(\frac{N_i^0 + g_i}{N_i^0}) + \lambda_1 + \lambda_2 \epsilon_i \right\} \delta N_i = 0 \quad . \tag{77}$$

Since in (77) each factor of δN_i has to vanish, by choosing for matters of convenience

$$\lambda_1 = -\ln B \quad , \quad \lambda_2 = -\beta \quad , \tag{78}$$

one easily can see that

$$\ln(\frac{N_i^0 + g_i}{N_i^0}) - \ln B_1 - \beta \epsilon_i = 0 \quad , \tag{79}$$

from which immediately follows that

$$\ln(\frac{N_i^0 + g_i}{B N_i^0}) = \beta \epsilon_i \quad \rightarrow \quad \frac{g_i}{N_i^0} = B \exp(\beta \epsilon_i) - 1 \quad . \tag{80}$$

The ratio of the number of phase points N_i per cell and the number of subcells (g_i) is therefore given by

$$\frac{N_i^0}{g_i} = \frac{1}{B \exp(\beta \epsilon_i) - 1} \quad . \tag{81}$$

This ratio is called the **Bose-Einstein distribution function**.

13.4.4 Maxwell-Boltzmann statistics

If the number of particles per cell is much smaller than the number of subcells, i.e.,

$$\begin{aligned} N_i^0 \ll g_i \ \rightarrow \ \frac{N_i^0}{g_i} \ll 1 \rightarrow \ B\exp(\beta\epsilon_i) - 1 \gg 1 \\ \rightarrow \ B\exp(\beta\epsilon_i) - 1 \simeq B\exp(\beta\epsilon_i) \quad , \end{aligned} \tag{82}$$

then

$$\frac{N_i^0}{g_i} = \frac{1}{B}\exp(-\beta\epsilon_i) \quad , \ B \gg 1 \quad . \tag{83}$$

This last equation is exactly of the form of the **Maxwell-Boltzmann distribution** (see also (59)). The inequalities in (82) imply that within a gedanken experiment where the subcells become smaller and smaller, i.e. position and momentum of a particle becomes sharper and sharper, the Bose-Einstein distribution reduces to the Maxwell-Boltzmann distribution.

13.4.5 Partition function

The constant B appearing in (81) and (83) can be obtained from the condition that the number of particles has to be constant. Assuming now always maximal probability, the superscript specifying the "optimal" set of numbers N_i can be dropped. From (74) follows immediately that

$$\sum_{i=1}^{M} N_i = N = \frac{1}{B}\sum_{i=1}^{M} g_i \exp(-\beta\epsilon_i) \quad . \tag{84}$$

The sum appearing on the rhs of this equation is called **partition function** and traditionally is abbreviated by Z or Q

$$Z = \sum_{i=1}^{M} g_i \exp(-\beta\epsilon_i) \quad . \tag{85}$$

With this notation the constant B is given by

$$B = \frac{Z}{N} \tag{86}$$

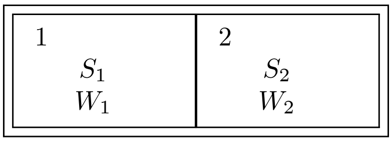
and the Maxwell-Boltzmann distribution can be formulated as

$$N_i = \frac{N}{Z} g_i \exp(-\beta\epsilon_i) \quad . \tag{87}$$

The evaluation of the constant β will be given in the following section.

13.4.6 Entropy and thermodynamical probability

Consider an isolated thermodynamical system that is partitioned into two isolated subsystems 1 and 2 with thermodynamical probabilities W_1 and W_2.



Let S_1 and S_2 be two functions such that in thermodynamical equilibrium

$$S = S_1 + S_2 \quad . \tag{88}$$

S is called **entropy**. If there is a connection between this phenomenological quantity S and the thermodynamical probability $W = W_1 W_2$ (in thermodynamical equilibrium) then S has to be a function of W such that

$$S = S_1 + S_2 = f(W) = f(W_1 W_2) = f(W_1) + f(W_2) \quad . \tag{89}$$

Quite clearly S can only be proportional to the logarithm of the thermodynamical probability W, if

$$S = k \ln W \quad . \tag{90}$$

Going now back to (68)

$$\ln W^0 \equiv \ln W = \sum_{i=1}^{M} \underbrace{(N_i + g_i)\ln(N_i + g_i)}_{} - g_i \ln g_i - N_i \ln N_i \quad , \tag{91}$$

the first term on the rhs of this equation can be rewritten as follows

$$(N_i + g_i)\ln(N_i + g_i) = g_i(1 + \frac{N_i}{g_i}) \ln\left(g_i(1 + \frac{N_i}{g_i})\right) =$$

$$= g_i(1 + \frac{N_i}{g_i}) \left(\ln g_i + \underbrace{\ln(1 + \frac{N_i}{g_i})}_{} \right) \quad . \tag{92}$$

Using the following expansion of $\ln(1+x) = x - x^2/2 + x^3/3 \pm \cdots$; $(-1 \leq x \leq 1)$, it is obvious that for $g_i \gg N_i$ (Maxwell-Boltzmann distribution)

$$\ln(1 + \frac{N_i}{g_i}) \simeq \frac{N_i}{g_i} \quad ,$$

and therefore

$$(N_i + g_i)\ln(N_i + g_i) = g_i \ln g_i + N_i \ln g_i + N_i + \frac{N_i^2}{g_i} \simeq$$

$$\simeq g_i \ln g_i + N_i \ln g_i + N_i \tag{93}$$

In the case of the Maxwell-Boltzmann distribution (91) reduces therefore to

$$\ln W = \sum_{i=1}^{M} N_i \ln g_i + N_i - N_i \ln N_i = \sum_{i=1}^{M} N_i (\ln \frac{g_i}{N_i} + 1) \quad . \tag{94}$$

By using now in this equation the definition for the partition function Z (87), which was derived from the condition that the number of particles has to stay constant,

$$\frac{N_i}{g_i} = \frac{N}{Z} \exp(-\beta\epsilon_i) \rightarrow \ \ln \frac{g_i}{N_i} = \ln \frac{Z}{N} + \beta\epsilon_i \quad , \quad Z = \sum_{i=1}^{M} g_i \exp(-\beta\epsilon_i) \quad , \tag{95}$$

one gets for the entropy S:

$$S = k \ln W = k \sum_{i=1}^{M} N_i (\ln \frac{Z}{N} + \beta\epsilon_i + 1) =$$

$$= k \ln \frac{Z}{N} \sum_{i=1}^{M} N_i + k\beta \underbrace{\sum_{i=1}^{M} \epsilon_i N_i}_{= U} + \ k \sum_{i=1}^{M} N_i =$$

$$= kN \ln \frac{Z}{N} + k\beta U + kN \quad . \tag{96}$$

From classical (phenomenological) thermodynamics, however, it is known that in the case of only one type of particles the partial derivative of the entropy with respect to the energy U at constant volume v is given by the inverse of the temperature T,

$$\left(\frac{\partial S}{\partial U} \right)_v = \frac{1}{T} \quad . \tag{97}$$

In evaluating this derivative,

$$\left(\frac{\partial S}{\partial U} \right)_v = \frac{kN}{Z} \frac{dZ}{d\beta} \left(\frac{\partial \beta}{\partial U} \right)_v + k\beta + kU \left(\frac{\partial \beta}{\partial U} \right)_v \quad , \tag{98}$$

one can make use of the fact that

$$\frac{dZ}{d\beta} = - \sum_{i=1}^{M} \epsilon_i g_i \exp(-\beta\epsilon_i) = - \frac{UZ}{N} \quad , \tag{99}$$

since from the definition of the partition function

$$N_i = \frac{N}{Z} g_i \exp(-\beta\epsilon_i)$$

it follows that

$$\sum_{i=1}^{M} \epsilon_i N_i = U = \frac{N}{Z} \sum_{i=1}^{M} g_i \epsilon_i \exp(-\beta\epsilon_i) \quad . \tag{100}$$

Using now the expression for $dZ/d\beta$ (99) in equation (98) one gets

$$\left(\frac{\partial S}{\partial U}\right)_v = -\frac{kN}{Z}\frac{UZ}{N}\left(\frac{\partial \beta}{\partial U}\right)_v + k\beta + kU\left(\frac{\partial \beta}{\partial U}\right)_v = k\beta \quad . \tag{101}$$

Comparing finally this result with the classical expression in (97) one can see that obviously the constant β has to be of the form

$$\beta = \frac{1}{kT} \quad . \tag{102}$$

Now all Lagrange parameters are evaluated and the Boltzmann distribution in (87) - and for that matter also the Bose-Einstein distribution - can be completed:

$$N_i = \frac{Ng_i}{Z} \exp(-\frac{\epsilon_i}{kT}) \quad . \tag{103}$$

The partition function Z is then given by

$$Z = \sum_{i=1}^{M} g_i \exp(-\frac{\epsilon_i}{kT}) \quad , \tag{104}$$

and the total energy U by

$$U = \frac{N}{Z}\sum_{i=1}^{M} g_i \epsilon_i \exp(-\frac{\epsilon_i}{kT}) = \frac{NkT^2}{Z}\frac{dZ}{dT} = NkT^2 \frac{d\ln Z}{dT} \quad . \tag{105}$$

13.4.7 Fermi-Dirac statistics

In contrast to the Bose-Einstein case, in the Fermi-Dirac case only two phase points can occupy one particular subcell. If one imagines each subcell to be split into two half-subcells, then the number of such half-subcells is given by

$$g_i = \frac{2d\phi}{h^3} \quad , \tag{106}$$

where as should be remembered $d\phi$ is the volume of a cell. Within this construction in each half-subcell there can be only one phase point or none, i.e., these half-subcells can be either occupied or unoccupied. If N_i is the number of particles in cell i then there N_i occupied and $(N_i - g_i)$ unoccupied half-subcells. The number of microstates for each macrostate is then given by

$$W_i = \binom{g_i}{N_i} \equiv \binom{g_i}{g_i - N_i} = \frac{g_i!}{N_i!(g_i - N_i)!} \quad , \tag{107}$$

and the thermodynamical probability by

$$W = \prod_{i=1}^{M} \binom{g_i}{N_i} \quad . \tag{108}$$

By assuming as before in the case of the Bose-Einstein statistics that $g_i, N_i \gg 1$, i.e., that the Stirling formula $\ln(n!) = n \ln n - n$ can be used, the logarithm of the thermodynamical probability is given by

$$\ln W = \sum_{i=1}^{M} g_i \ln g_i - N_i \ln N_i + (N_i - g_i) \ln(g_i - N_i) \quad . \tag{109}$$

This probability has to be a maximum (see also (74))

$$\begin{aligned}\delta \ln W = \sum_{i=1}^{M} \; -N_i \delta \ln N_i - \ln N_i \delta N_i + (N_i - g_i) \delta \ln(g_i - N_i) + \\ + \ln(g_i - N_i) \delta N_i = \\ = \sum_{i=1}^{M} - \ln N_i \delta N_i + \ln(g_i - N_i) \delta N_i = \sum_{i=1}^{M} \ln(\frac{g_i - N_i}{N_i}) \delta N_i = 0\end{aligned} \tag{110}$$

under the condition that the number of particles and the total energy stays constant ($\delta N = \delta U = 0$), which implies that

$$\sum_{i=1}^{M} \left\{ \ln(\frac{g_i - N_i}{N_i}) - \ln B - \beta \epsilon_i \right\} \delta N_i = 0 \quad , \tag{111}$$

where the same convention for the Lagrange parameters is used as in (82). For the ratio of the "optimal" occupation N_i ($\equiv N_i^0$) and g_i one obtains therefore

$$\frac{N_i}{g_i} = \frac{1}{B \exp(\beta \epsilon_i) + 1} \quad . \tag{112}$$

The constant β can be obtained in a similar way as before (105), namely by comparing the partial derivative of the entropy S with respect to the energy U at constant volume v with its classical analogon:

$$\beta = \frac{1}{kT} \quad . \tag{113}$$

In order to determine the constant B in (112) g_i and N_i are thought to be replaced by the following infinitesimal quantities

$$g_i \rightarrow \frac{2}{h^3} dx dy dz dp_x dp_y dp_z = \frac{2}{h^3} d\phi \quad , \tag{114}$$

$$N_i \rightarrow dN_x dN_y dN_z \underbrace{dW_x dW_y dW_z}_{d^3W} \quad , \tag{115}$$

where dW_x, dW_y and dW_z are probability coordinates that correspond to an integration over p_x, p_y and p_z. With this replacement (112) can be written as

$$dN_x dN_y dN_z d^3W = \frac{2}{h^3} \frac{1}{B\exp(\beta\epsilon)+1} dx dy dz dp_x dp_y dp_z \quad . \tag{116}$$

Integrating now over x, y and z yields

$$N d^3W \equiv d^3N = \frac{2}{h^3} v \frac{1}{B\exp(\beta\epsilon)+1} dp_x dp_y dp_z \quad , \tag{117}$$

where v is the volume. For $B\exp(\beta\epsilon_i) \gg 1$, this would lead again to the Maxwell-Boltzmann statistics. For the Fermi-Dirac case, as given for example by a system of electrons, this (unfortunately) is not the case. According to Sommerfeld for fermions the constant B is of the form

$$B = \exp(-\epsilon_m/kT) \quad , \tag{118}$$

where ϵ_m is a temperature-dependent reference energy,

$$\epsilon_m = \epsilon_{m0}(1 + (\frac{\pi kT}{\epsilon_{m0}})^2 + \cdots) \quad , \tag{119}$$

with ϵ_{m0} being a constant. With this ansatz (117) can be rewritten as follows

$$\rho \equiv \frac{d^3N}{dp_x dp_y dp_z} = \frac{2}{h^3} v \frac{1}{\exp(\frac{\epsilon-\epsilon_m}{kT})+1} \quad . \tag{120}$$

Expanding now the exponential in this equation in the usual power series,

$$\exp(\frac{\epsilon-\epsilon_m}{kT}) = 1 + \frac{1}{1!}(\frac{\epsilon-\epsilon_m}{kT}) + \frac{1}{2!}(\frac{\epsilon-\epsilon_m}{kT})^2 + \cdots \quad ,$$

and truncating the series after the first term, one gets

$$\rho = \frac{2}{h^3} v \frac{kT}{2kT+\epsilon-\epsilon_m} \quad . \tag{121}$$

From the limit of $T \to 0$, $\epsilon_m \to \epsilon_{m0}$,

$$\begin{aligned} \lim_{T\to 0} \rho \equiv \rho_0 = \frac{2}{h^3} v \lim_{T\to 0} \frac{kT}{2kT+\epsilon-\epsilon_m} = \\ = \frac{2}{h^3} v \lim_{T\to 0} \frac{d(kT)/dT}{d(2kT+\epsilon-\epsilon_m)/dT} = \frac{v}{h^3} \quad , \end{aligned} \tag{122}$$

one can see that

$$\rho_0 = \{ \begin{array}{ll} v/h^3 & ;\ \epsilon < \epsilon_{m0} \\ 0 & ;\ \epsilon > \epsilon_{m0} \end{array} \tag{123}$$

The constant ϵ_{m0} is called **Fermi energy** and is traditionally denoted by ϵ_F or ϵ_f.

13.5 Partition functions

In the Maxwell-Boltzmann statistics, which shall be applied in the following, the partition function Z (85) is given by a sum over the exponential functions only

$$Z = \sum_{i=1}^{M} \exp(-\epsilon_i/kT) \quad . \tag{124}$$

The energy ϵ_i in each cell i can be thought to be the sum of the translational energy ϵ_{trans} of the center of gravity and an internal energy ϵ_{int}, which in turn can be considered to be decomposed into a sum of the rotational energy ϵ_{rot}, vibrational energy ϵ_{vib}, electronic energy ϵ_{el}, nuclear energy ϵ_{nuc} and chemical energy ϵ_{chem}. In terms of such a decomposition the partition function is given by

$$\begin{aligned} Z &= \sum_{i=1}^{M} \exp(-(\epsilon_{trans} + \epsilon_{rot} + \epsilon_{vib} + \epsilon_{el} + \epsilon_{nuc} + \epsilon_{chem})/kT) = \\ &= \sum_{i=1}^{M} \exp(-\epsilon_{trans}/kT) \exp(-\epsilon_{rot}/kT) \exp(-\epsilon_{vib}/kT) \times \\ &\quad \times \exp(-\epsilon_{el}/kT) \exp(-\epsilon_{nuc}/kT) \exp(-\epsilon_{chem}/kT) = \\ &= Z_{trans} Z_{rot} Z_{vib} Z_{el} Z_{nuc} Z_{chem} \quad . \end{aligned} \tag{125}$$

If therefore the various contributions to ϵ_i correspond to independent motions or can be considered to be independent from each other then the partition function Z is a product of partition functions corresponding to the individual energy contributions. In the following these individual partition functions are discussed for some characteristic contributions to Z.

13.5.1 Translational partition function

For a particle (atoms, molecules etc.) moving in a three-dimensional box of dimensions a, b and c the energy eigenvalues are given[5] by

$$E_{n_x,n_y,n_z} = \frac{\pi^2\hbar^2}{8m} \left\{ \frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right\} \quad . \tag{126}$$

The translational (Boltzmann) partition function is therefore defined by

$$\begin{aligned} Z_{trans} = &\left(\sum_{n_x=1}^{\infty} \exp(-\frac{\hbar^2 n_x^2}{8mkTa^2}) \right) \left(\sum_{n_y=1}^{\infty} \exp(-\frac{\hbar^2 n_y^2}{8mkTb^2}) \right) \times \\ &\times \left(\sum_{n_z=1}^{\infty} \exp(-\frac{\hbar^2 n_z^2}{8mkTc^2}) \right) . \end{aligned} \tag{127}$$

---

[5] see chapter 3, equ.(3.67)

Since the energy differences between two following eigenvalues,

$$\frac{\hbar^2}{8mkTa^2}\left((n_x+1)^2-n_x^2\right)\sim\frac{\hbar^2}{4mkTa^2}n_x \quad , \tag{128}$$

are rather very small for macroscopical dimensions of a, b and c, i.e., since the quantum numbers n_x, n_y and n_z are **quasi-continuous**, the sums in (126) can be replaced by integrals:

$$\sum_{n_x=1}^{\infty}\exp(-\frac{\hbar^2 n_x^2}{8mkTa^2})=\int\limits_0^{\infty}\exp(-\frac{\hbar^2 n_x^2}{8mkTa^2})dn_x=\frac{a(2\pi mkT)^{1/2}}{\hbar} \quad . \tag{129}$$

Identifying finally abc as the volume v, the translational partition function is given by an expression, well-known from classical thermodynamics

$$Z_{trans}=\frac{v}{\hbar^3}(2\pi mkT)^{3/2} \quad . \tag{130}$$

13.5.2 Rotational partition function

For a heteroatomic diatomic molecule like the CO molecule (covering symmetry $C_{\infty v}$), the energy eigenvalues of a **rigid rotator**[6] motion are given by

$$\epsilon_{rot}=j(j+1)\frac{\hbar^2}{2I} \quad , \quad I=\frac{m_1 m_2}{m_1+m_2}R \quad , \tag{131}$$

where m_1 and m_2 are the masses of the two atoms and R the (rigid) distance between them. These eigenvalues are $(2j+1)$-fold degenerated. The corresponding partition function is then defined by

$$Z_{rot}=\sum_{j=0}^{\infty}(2j+1)\exp(-j(j+1)\frac{\Theta_{rot}}{T}) \quad , \quad \Theta_{rot}=\frac{\hbar^2}{2kI} \quad . \tag{132}$$

Since Θ_{rot} has the dimension of a temperature, it is sometimes also called the characteristic temperature of the corresponding (rigid) rotation. For $T\gg\Theta_{rot}$ the sum in (132) can again be replaced by an integral

$$Z_{rot}=\int\limits_0^{\infty}(2j+1)\exp(-j(j+1)\frac{\Theta_{rot}}{T})dj \quad . \tag{133}$$

By replacing $j(j+1)$ by ξ ($(2j+1)dj=d\xi$), one immediately can see that the partition function for a rigid rotation of a two-atomic molecule is given by

$$Z_{rot}=\frac{2IkT}{\hbar^2}=\frac{T}{\Theta_{rot}} \quad . \tag{134}$$

---

[6] see chapter 12

Since for a homoatomic diatomic molecule like N_2 the covering symmetry is $D_{\infty h}$, i.e., a rotation over π perpendicular to the internuclear distance R leaves the molecule invariant, a correction factor of 1/2 has to be augmented. For an arbitrary diatomic molecule the rotational partition function is then given by

$$Z_{rot} = \frac{2IkT}{\kappa \hbar^2} \quad , \quad \kappa = \{ \begin{array}{ll} 2 & D_{\infty h} \\ \\ 1 & C_{\infty v} \end{array} \quad , \tag{135}$$

where κ is sometimes also called symmetry correction factor. In general for a molecule with three principal moments of inertia I_1, I_2 and I_3 and a symmetry correction factor $\kappa > 1$, the corresponding rotational partition function is given by

$$Z_{rot} = \frac{(8\pi I_1 I_2\ I_3)^{1/2}}{\kappa h^3} (kT)^{3/2} \quad . \tag{136}$$

13.5.3 Vibrational partition function

Consider again a diatomic molecule such as CO or N_2. In this case there is only one vibrational degree of freedom. The vibrational energy is then of the form [7]

$$E_{vib} = \hbar\omega(n + \frac{1}{2}) \quad , \tag{137}$$

and the corresponding partition function is given by

$$Z_{vib} = \sum_{n=0}^{\infty} \exp(-\hbar\omega(n + \frac{1}{2})/kT) \quad . \tag{138}$$

In general, however, the energy difference between two subsequent eigenvalues ($\hbar\omega$) is no longer small as compared to kT, and therefore the sum cannot be converted directly into an integral. By factorizing the zero level contribution $\hbar\omega/2$, and making use of the geometrical series,

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots \quad ,$$

one can see that the vibrational partition function

$$Z_{vib} = \exp(-\frac{\hbar\omega}{2kT}) \sum_{n=0}^{\infty} \exp(-\frac{n\hbar\omega}{kT}) =$$

$$= \exp(-\frac{\hbar\omega}{2kT}) \left(1 + \underbrace{\exp(-\frac{\hbar\omega}{kT})}_{x} + \underbrace{\exp(-\frac{2\hbar\omega}{kT})}_{x^2} + \cdots \right) \quad , \tag{139}$$

---

[7] see chapter 11

can (at least) be summed up in a closed form:

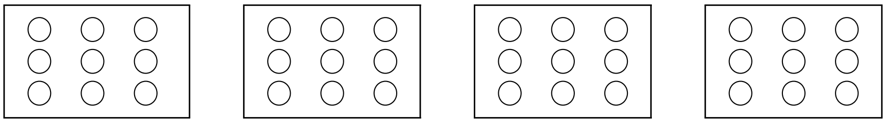
$$Z_{vib} = \frac{\exp(-\frac{\hbar\omega}{2kT})}{\left(1 - \exp(-\frac{\hbar\omega}{kT})\right)} \quad . \tag{140}$$

A polyatomic molecule with s atomic constituents has in general $3s - 6$ vibrational degrees of freedom. Each degree of freedom corresponds then to a so-called normal coordinate and a normal vibrational frequency ω_i.

13.6 Thermodynamical functions

Up to now it was assumed that like in a gas the macroscopic system consists of a large number of particles (single molecules) that do not interact. The discussed thermodynamical properties like the energy U or the entropy S correspond therefore to the partition functions for those particles. For liquid or solid systems, however, one has to take into account the interaction between these particles.

Consider a collection of particles (molecules), which can be termed a "super-molecule", and consider further that the macroscopical system can be viewed as a collection of such "super-molecules", each of them having exactly the same number of particles and the same volume:



If this set of L super-molecules is in thermodynamical equilibrium, then it is called a canonical ensemble of super-molecules. As compared to the large number of internal degrees of freedoms, the three translational degrees of freedom can be neglected. The internal energies E_l of a super-molecule are considered to be quantized, even so they are quite dense. Assuming for the equilibrium a Boltzmann distribution of the kind

$$W_l = \frac{\exp(-E_l/kT)}{\sum\limits_{l=1}^{L} \exp(-E_l/kT)} \quad , \tag{141}$$

the corresponding partition function

$$\widetilde{Z} = \sum_{l=1}^{L} \exp(-E_l/kT) \tag{142}$$

is called **canonical partition ("macroscopical") function**.

13.6.1 Weakly interacting particles

If the super-molecule can be considered as a gas with very small interactions between the individual molecules, then the internal energy E_l is a sum of the internal energies of the individual molecules numbered by α

$$E_l \sim \sum_{\alpha=1}^{N_\alpha} E_{\alpha_m,l} \quad , \tag{143}$$

where the subscript α_m, l denotes the $m-th$ energy eigenvalue of the $\alpha - th$ individual molecule that belongs to the $l - th$ energy eigenvalue of the super-molecule. In this case the canonical partition function reduces to

$$\widetilde{Z} = \sum_{l=1}^{L} \exp\left(-(\sum_{\alpha=1}^{N_\alpha} E_{\alpha_m,l})/kT \right) = \sum_{l=1}^{L} \prod_{\alpha=1}^{N_\alpha} \exp(-E_{\alpha_m,l})/kT =$$

$$= \prod_{\alpha=1}^{N_\alpha} \sum_{l=1}^{L} \exp(-E_{\alpha_m,l})/kT \quad . \tag{144}$$

For the model of an ideal gas, i.e., for the case that there are no interactions within the super-molecule, the energy eigenvalues $E_{\alpha_m,l}$ are independent of l, i.e., $E_{\alpha_m,l} \equiv E_{\alpha_m} \equiv E_m$. The corresponding canonical partition function is then given by

$$\widetilde{Z} = \prod_{\alpha=1}^{N_\alpha} \sum_{m=1}^{L} \exp(-E_{\alpha_m})/kT =$$

$$= \frac{1}{N_\alpha!} \left(\sum_{m=1}^{L} \exp(-E_m)/kT \right)^{N_\alpha} = \frac{1}{N_\alpha!} Z^{N\alpha} \quad , \tag{145}$$

where the factor $1/N_\alpha!$ arises from the fact that there are $N_\alpha!$ possibilities to add up the energy eigenvalues for the super-molecule from N_α individual molecules. It should be noted that in (145) Z is the partition function for an individual molecule.

Internal energy The internal energy of a system of N weakly interacting particles (molecules) can be formulated as

$$U = \sum_{i=1}^{N} N_i E_i = N\overline{E} \quad , \tag{146}$$

where $\overline{E}$ is the averaged energy per particle,

$$\overline{E} = \frac{\sum\limits_{i=1}^{N} N_i E_i}{\sum\limits_{i=1}^{N} N_i} = \frac{\sum\limits_{i=1}^{N} E_i \exp(-E_i/kT)}{\sum\limits_{i=1}^{N} \exp(-E_i/kT)} = \frac{\sum\limits_{i=1}^{N} E_i \exp(-E_i/kT)}{Z} \quad . \tag{147}$$

Since according to (99)

$$\left(\frac{\partial Z}{\partial \beta}\right)_v = -\sum_{i=1}^{N} E_i \exp(-E_i/kT) \quad , \quad \beta = 1/kT \quad , \tag{148}$$

the internal energy can be written as

$$U = -N\left(\frac{\partial \ln Z}{\partial \beta}\right)_v = -\left(\frac{\partial \ln \widetilde{Z}}{\partial \beta}\right)_v \quad , \tag{149}$$

or alternatively as

$$U = -nR\left(\frac{\partial \ln Z}{\partial (1/T)}\right)_v \quad , \quad nR = kN \quad , \tag{150}$$

where R is the ideal gas constant. For the specific heat at constant volume one gets therefore

$$c_v = \frac{1}{n}\left(\frac{\partial U}{\partial T}\right)_v = \frac{R}{T^2}\left(\frac{\partial^2 \ln Z}{\partial (1/T)^2}\right)_v \quad . \tag{151}$$

Entropy The entropy of a system of weakly interacting particles is given by (see also (96))

$$S = kN\ln(\frac{Z}{N}) + k\beta U + kN = kN(\ln(\frac{Z}{N}) + 1) + \frac{U}{T} \quad , \tag{152}$$

or alternatively using the ideal gas constant and (150) by

$$S = nR\left(\ln(\frac{Z}{N}) - \frac{1}{T}\left(\frac{\partial \ln Z}{\partial (1/T)}\right)_v + 1\right) \tag{153}$$

In terms of the macroscopical partition function $\widetilde{Z}$ the entropy can be written as

$$S = k\left(\ln \widetilde{Z} \; - \frac{1}{T}\left(\frac{\partial \ln \widetilde{Z}}{\partial (1/T)}\right)_v\right) \quad , \tag{154}$$

where because of the Sterling formula the one in (153) can be omitted.

Helmholtz free energy Using the definition for the (Helmholtz) free energy F,

$$F = U - TS \quad , \tag{155}$$

it is now easy to see that

$$F = -nRT\left(\ln(\frac{Z}{N}) + 1\right) \quad \text{or} \quad F = -kT\ln \widetilde{Z} \quad . \tag{156}$$

Pressure, Gibbs free energy and enthalpy By recalling the following famous thermodynamical definitions,

$$G = U - TS + pv = F + pv \quad , \tag{157}$$

$$p = -\left(\frac{\partial F}{\partial v}\right)_T \quad , \tag{158}$$

it is again easy to show that the pressure p is given by

$$p = nRT\left(\frac{\partial \ln Z}{\partial v}\right)_T \quad , \tag{159}$$

and the Gibbs free enthalpy G therefore by

$$G = -nRT\left(\ln(\frac{Z}{N}) + 1 - v\left(\frac{\partial \ln Z}{\partial v}\right)_T\right) \quad . \tag{160}$$

For the enthalpy H,

$$H = U + pv \quad , \tag{161}$$

one finally gets

$$H = nRT\left(\left(\frac{\partial \ln Z}{\partial \ln T}\right)_v + \left(\frac{\partial \ln Z}{\partial \ln v}\right)_T\right) \quad . \tag{162}$$

13.7 Summary of thermodynamical functions for weakly interacting particles

In summarizing all the above expressions for the various thermodynamical quantities or functions, one can see at one glance the overall importance of the partition function in statistical thermodynamics:

$$U = -nR\left(\frac{\partial \ln Z}{d(1/T)}\right)_v$$

$$c_v = \frac{R}{T^2}\left(\frac{\partial^2 \ln Z}{\partial(1/T)^2}\right)_v$$

$$S = nR\left(\ln(\frac{Z}{N}) - \frac{1}{T}\left(\frac{\partial \ln Z}{\partial(1/T)}\right)_v + 1\right)$$

$$F = -nRT\left(\ln(\frac{Z}{N}) + 1\right)$$

$$p = nRT\left(\frac{\partial \ln Z}{\partial v}\right)_T$$

$$G = -nRT\left(\ln(\frac{Z}{N}) + 1 - v\left(\frac{\partial \ln Z}{\partial v}\right)_T\right)$$

$$H = nRT\left(\left(\frac{\partial \ln Z}{\partial \ln T}\right)_v + \left(\frac{\partial \ln Z}{\partial \ln v}\right)_T\right)$$

References

[1] E.M.Landau and L.P.Lifschitz, *"Lehrbuch der Theoretischen Physik - Statistische Physik"*, Parts 1 and 2, Akademie Verlag, Berlin, 1979

[2] C.L.Tien and J.H.Lienhard, *"Statistical Thermodynamics"*, Holt-Rinehard-Winston, New York, 1971

[3] R.P.Feynman, *"Statistical Mechanics"*, Addison-Wesley Publishing Company, 1988

[4] J.F.Lee, F.W.Sears and D.L.Turcotte, *"Statistical Thermodynamics"*, Addison-Wesley Publishing Company, 1973

[5] L.K.Nash, *"Elements of Statistical Thermodynamics"*, Addison-Wesley Publishing Company, 1968

[6] F.C.Andrew, *"Equilibrium Statistical Mechanics"*, John Wiley, New York, 1963

[7] O.K.Rice, *"Statistical Mechanics, Thermodynamics and Kinetics"*, Freeman and Co., San Francisco-London, 1967

14 Mathematical foundations of quantum mechanics

References

[1] R.Courant and D.Hilbert, *"Methods of Mathematical Physics"*, John Wiley and Sons, 1989

[2] G.Fano, *"Mathematical Methods of Quantum Mechanics"*, McGraw Hill Comp., 1971

15 Second Quantization

The concept of operators, functionals, and the isomorphism between the abstract Hilbert space and the L_2 function space, i.e. the space of all quadratic integrable functions, became quite familiar in quantum mechanics. In fact most of this book is confined to this concept, which is summarized in chapter 14. Since the so-called second quantization, which is **not** a **new or another form** of **quantization**, but a very elegant and efficient way to describe n-particle systems, is perhaps less known than the "first" quantization, this chapter deals with its general concepts in quite some detail.

15.1 The n-particle space

Suppose ψ_n is an element of the abstract Hilbert space H_n that corresponds to a quantum mechanical system of n identical particles. The configurational space representation of this element is continuous,

$$\psi(x_1, x_2, x_3, ..., x_n; t) = \langle x_1, x_2, x_3, ..., x_n \mid \psi_n(t)\rangle \quad , \tag{1}$$

where $x_i = (\mathbf{r}_i, \sigma_i)$ comprises the space- and spin coordinates of the i-th particle. In configurational space the resolution of the identity and the orthogonality relation is given by

$$\iiint ... \int \mid x_n, ..., x_3, x_2, x_1\rangle\langle x_1, x_2, x_3, ..., x_n \mid dx_1 dx_2 dx_3 ... dx_n \; = 1 \quad , \tag{2}$$

$$\langle x'_n, ..., x'_3, x'_2, x'_1 \mid x_1, x_2, x_3, ..., x_n\rangle = \delta(x_1 - x'_1)...\delta(x_n - x'_n) \quad . \tag{3}$$

Rather than using **continuous representations** in most cases **discrete representations** are needed, as given for example in terms of a basis formed by the eigenfunctions of an (Hermitian) operator with a discrete eigenvalue spectrum.

Let $\{\phi_k(x)\}$ be a complete and orthonormal set of one-particle wave functions,

$$\phi_k(x) = \langle x \mid \phi_k\rangle = \langle x \mid k\rangle \quad , \tag{4}$$

$$\langle \phi_i \mid \phi_j\rangle = \langle i \mid k\rangle = \delta_{ij} \quad , \tag{5}$$

$$\sum_{k=1}^{n} \mid \phi_k\rangle\langle\phi_k \mid = \sum_{k=1}^{n} \mid k\rangle\langle k \mid = 1 \quad , \tag{6}$$

then the transformation to a discrete representation of ψ_n can be expressed by the following generalized resolution of the identity, where k_i stands for all indices k in (6) for the i-th particle,

$$\sum_{k_1, k_2, ..., k_n} \mid k_n, k_{n-1}, ..., k_2, k_1\rangle\langle k_1, k_2, ..., k_{n-1}, k_n \mid =$$

$$= \sum_{\{k_i\}} \mid k_n, k_{n-1}, ..., k_2, k_1\rangle\langle k_1, k_2, ..., k_{n-1}, k_n \mid = 1 \quad , \tag{7}$$

such that

$$\psi_n(x_1, x_2, ..., x_n; t) =$$

$$= \sum_{\{k_i\}} \langle x_1, x_2, ..., x_n \mid k_n, .., k_2, k_1 \rangle \underbrace{\langle k_1, k_2, ...k_n \mid \psi_n(t) \rangle} =$$

$$= \sum_{\{k_i\}} \phi_{k_1}(x_1)\phi_{k_2}(x_2)...\phi_{k_n}(x_n) c_n(k_1, k_2, ...k_n; t) =$$

$$= \sum_{\{k_i\}} \phi_{\{k_i\}}(x_1, x_2, ..., x_n) c_n(k_1, k_2, ...k_n; t) \quad . \tag{8}$$

The functions $c_n(k_1, k_2, ...k_n; t)$ are now the transformed n-particle wave functions and the set of discrete variables $\{k_i\}$ is called a **configuration**. As is well-known a system of n identical particles can be either a system of **bosons** or of **fermions**. This additional symmetry (permutational symmetry) helps tremendously to simplify (8).

15.1.1 Configuration interaction

For bosons the **occupation numbers** n_i of a particular (one-particle) state ϕ_i can be 0, 1, 2, .. , whereas for fermions they are restricted to 0 and 1. These occupation numbers n_i specify how often a particular argument x_i (one-particle wave function $\phi_i(x_i)$) occurs in a given function $\psi(x_1, x_{2,}, .., x_i, ..., x_n; t)$.

Bosons Because of multiple occupations, for a particular configuration $\{k_i\}$ there are more contributions of the same kind to the sum in (8), namely exactly

$$\frac{n!}{n_1!n_2!...n_i!..} \quad . \tag{9}$$

Boson wave functions, $\psi_n^B(x_1, x_2, ..., x_n; t)$, have to be **symmetric** with respect to permutations of the coordinates $x_1, x_2, ..., x_n$, i.e. have to be invariant with respect to the **symmetrization operator** $O_s^{(n)}$,

$$O_s^{(n)} = \frac{1}{n!} \sum_{P \in S_n} P_{(x)} \quad , \tag{10}$$

where S_n is the permutational group,

$$\psi_n^B(x_1, x_2, ..., x_n; t) = O_s^{(n)} \psi_n^B(x_1, x_2, ..., x_n; t) =$$

$$= \sum_{\{k_i\}} O_s^{(n)} \phi_{k_1}(x_1)\phi_{k_2}(x_2)...\phi_{k_n}(x_n) c_n^B(k_1, k_2, ...k_n; t) =$$

$$= \sum_{\{k_i\}'} \frac{n!}{n_1!n_2!...n_i!..} O_s^{(n)} \phi_{k_1}(x_1)\phi_{k_2}(x_2)...\phi_{k_n}(x_n) c_n^B(k_1, k_2, ...k_n; t) \quad . \tag{11}$$

It should be noted that the permutational group is of order $n!$ There are only two one-dimensional irreducible representations, namely the trivial representation and the so-called alternating representation, where each permutation is represented by ϵ_P , see also the below case of fermions. Permutating e.g. (x_1, x_2, x_3) in the alternating representation, ϵ_P is in turn 1, -1 and 1. The symmetrization operator projects[1] onto the subspace of totally symmetric states, the antisymmetrization operator onto the subspace of totally antisymmetric states. The index (x) indicates that the coordinates are permutated.

In (11) $\{k_i\}'$ now denotes the set of all different configurations, the summation is therefore restricted to only those configurations! Denoting $\{k_i\}'$ simply by K, equation (11) can also be written as

$$\psi_n^B(x_1, x_2, ..., x_n; t) = \sum_K \phi_K^B(x_1, x_2, ..., x_n)\overline{c}_K^B(t) \quad , \tag{12}$$

where

$$\phi_K^B(x_1, x_2, ..., x_n) = \sqrt{\frac{n!}{n_1!n_2!...n_i!..}} O_s^{(n)} \phi_{k_1}(x_1)\phi_{k_2}(x_2)...\phi_{k_n}(x_n) \tag{13}$$

is a **normalized** and **symmetric product of single particle wave functions** (**Hartree product**) and

$$\overline{c}_K^B(t) = \sqrt{\frac{n!}{n_1!n_2!...n_i!..}} c_n^B(k_1, k_2, ...k_n; t) =$$

$$= \sqrt{\frac{n!}{n_1!n_2!...n_i!..}} \langle k_1, k_2, ...k_n \mid \psi_n^B(t)\rangle \quad . \tag{14}$$

Because of the normalization of the n-boson wave functions,

$$\langle \psi_n^B(x_1, x_2, ..., x_n; t) \mid \psi_n^B(x_1, x_2, ..., x_n; t)\rangle = 1 \quad , \tag{15}$$

the sum over the squares of $\overline{c}_K^B(t)$ has to be unity,

$$\sum_K \left|\overline{c}_K^B(t)\right|^2 = 1 \quad . \tag{16}$$

Fermions The n-fermion wave function $\psi_n^F(x_1, x_2, ..., x_n; t)$ has to be **antisymmetric** with respect to permutations of the coordinates $x_1, x_2, ..., x_n$, i.e. has to be invariant with respect to the **antisymmetrization operator** $O_{as}^{(n)}$,

$$O_{as}^{(n)} = \frac{1}{n!} \sum_{P \in S_n} \epsilon_P P_{(x)} \quad , \tag{17}$$

$$O_{as}^{(n)} \psi_n^F(x_1, x_2, ..., x_n; t) = \psi_n^F(x_1, x_2, ..., x_n; t) \quad . \tag{18}$$

---

[1] see for example the discussion in the book by Jansen and Boon, p 233 ff

For fermions the occupation numbers n_i are restricted to the set $\{0,1\}$. Using as before a discrete representation,

$$\psi_n^F(x_1,x_2,...,x_n;t) =$$

$$= \sum_{\{k_i\}} O_{as}^{(n)} \phi_{k_1}(x_1)\phi_{k_2}(x_2)...\phi_{k_n}(x_n) c_n^F(k_1,k_2,...k_n;t) \quad , \tag{19}$$

it turns out that the transformed functions $c_n^F(k_1,k_2,...k_n;t)$ have the same symmetry as the original wave function,

$$O_{as}^{(n)} c_n^F(k_1,k_2,...k_n;t) = c_n^F(k_1,k_2,...k_n;t) \quad . \tag{20}$$

Since according to (17) $O_{as}^{(n)}$ acts on the x coordinates, the functions $c_n^F(k_1,k_2,...k_n;t)$ are only determined up to the sign, the set $\{k_1,k_2,...k_n\}$ that characterizes a configuration has to satisfy an additional condition in order to specify configurations uniquely, namely $k_1 < k_2 < ... < k_n$, i.e. this set has to be naturally ordered. Denoting such a naturally ordered sequence of arguments k_i by K

$$K = \{k_1 < k_2 < ... < k_n\} \quad , \tag{21}$$

K is also called an **ordered configuration**. Altogether there are $n!$ such ordered configurations. The n-fermion wave function in (19) can therefore be written as

$$\psi_n^F(x_1,x_2,...,x_n;t) =$$

$$= \sum_{\{k_1<k_2<...<k_n\}} n! O_{as}^{(n)} \phi_{k_1}(x_1)\phi_{k_2}(x_2)...\phi_{k_n}(x_n) c_n^F(k_1,k_2,...k_n;t) =$$

$$= \sum_K \phi_K^F(x_1,x_2,...,x_n)\overline{c}_K^F(t) \quad , \tag{22}$$

where

$$\phi_K^F(x_1,x_2,...,x_n) = \sqrt{n!} O_{as}^{(n)} \phi_{k_1}(x_1)\phi_{k_2}(x_2)...\phi_{k_n}(x_n) \quad , \tag{23}$$

is a **normalized** and **antisymmetric product of one-particle wave functions** (**Slater determinant**) that belongs to a particular ordered configuration K. The transformed functions $\overline{c}_K^F(t)$ comprise the remainder of the right hand side of (22),

$$\overline{c}_K^F(t) = \sqrt{n!} c_n^F(k_1,k_2,...k_n;t) \quad . \tag{24}$$

Because of the normalization for the n-fermion wave function,

$$\langle \psi_n^F(x_1,x_2,...,x_n;t) \mid \psi_n^F(x_1,x_2,...,x_n;t)\rangle = 1 \quad , \tag{25}$$

$$\sum_K \left|\overline{c}_K^F(t)\right|^2 = 1 \quad . \tag{26}$$

Equations (12) and (22), rewritten below in a more compact form

$$\psi_n(x_1,x_2,...,x_n;t) = \{ \begin{array}{ll} \sum\limits_K \phi_K^B(x_1,x_2,...,x_n)\overline{c}_K^B(t) & ;bosons \\ \sum\limits_K \phi_K^F(x_1,x_2,...,x_n)\overline{c}_K^F(t) & ;fermions \end{array} \tag{27}$$

represent nothing but the so-called **configuration interaction** for an n-particle system of either bosons or fermions. It should be noted that in both cases this interaction includes **all possible (allowed) configurations**.

15.2 The occupation number representation

Each configuration corresponds to a set N of occupation numbers,

$$N = \{n_1, n_2, ..., n_n\} \quad , \tag{28}$$

for which the following sum rule applies

$$\sum_{i=1}^{\infty} n_i = n \quad , \tag{29}$$

where n is the number of particles. Because of this unique relation a set $K = \{k_1, k_2, ..., k_n\}$ can be mapped uniquely onto a set $N = \{n_1, n_2, ..., n_n\}$.

15.2.1 Bosons

Using these sets of occupation numbers N for Bosons equation (12) can also be formulated as

$$\psi_n^B(x_1,x_2,...,x_n;t) = \sum_N \phi_N^B(x_1,x_2,...,x_n) f_N^B(t) \quad , \tag{30}$$

where

$$\begin{aligned} \phi_N^B(x_1,x_2,...,x_n) &= \langle x_1,x_2,...,x_n \mid n_n,...,n_2,n_1\rangle^B = \\ &= \sqrt{\frac{n!}{n_1!n_2!...}} O_s^{(n)} \langle x_1,x_2,...,x_n \mid k_n,...,k_2,k_1\rangle^B \quad , \end{aligned} \tag{31}$$

and

$$f_N^B(t) = \langle n_1,n_2,...,n_n \mid \psi_n^B(t)\rangle = \sqrt{\frac{n!}{n_1!n_2!...}} c_n^B(k_1,k_2,...,k_n;t) \quad . \tag{32}$$

The function $f_N^B(t)$ is called **occupation number representation** of an **n-boson wave function** for a given set of occupation numbers N.

15.2.2 Fermions

For fermions one similarly obtains from equation (22)

$$\psi_n^F(x_1, x_2, ..., x_n; t) = \sum_N \phi_N^F(x_1, x_2, ..., x_n) f_N^F(t) \quad , \tag{33}$$

where

$$\phi_N^F(x_1, x_2, ..., x_n) = \langle x_1, x_2, ..., x_n \mid n_n, ..., n_2, n_1 \rangle^F = \sqrt{n!} O_{as}^{(n)} \langle x_1, x_2, ..., x_n \mid k_n, ..., k_2, k_1 \rangle^F \quad , \tag{34}$$

and

$$f_N^F(t) = \langle n_1, n_2, ..., n_n \mid \psi_n^F(t) \rangle = \sqrt{n!} c_n^F(k_1, k_2, ..., k_n; t) \quad . \tag{35}$$

In (34) $\phi_N^F(x_1, x_2, ..., x_n)$ is a **Slater determinant** corresponding to a set of occupations numbers N and $f_N^F(t)$ is called **occupation number representation** of an **n-fermion wave function**.

15.3 The Fock space and the occupation number space

15.3.1 The abstract Fock space

The abstract **Fock space** or **"graded" Hilbert space** $\widehat{H}$ is the direct sum of all n-particle Hilbert spaces,

$$\widehat{H} = H_0 \oplus H_1 \oplus H_2 \oplus ... H_n \oplus ... \quad , \tag{36}$$

where the "zero-particle" space H_0 is the (one-dimensional) space (field) of all complex numbers. In (36) H_1 is the abstract **one-particle Hilbert space**, $H_2 = H_1 \otimes H_1$ is the abstract **two-particle Hilbert space**, and in general $H_n = H_{n-1} \otimes H_1$ is the abstract **n-particle Hilbert space**, where $\otimes$ denotes the tensorial product. For bosons the **symmetric tensorial product** has to be taken, for fermions the **antisymmetric**. The n-boson space (H_n^B) and the n-fermion space (H_n^F) are formally subspaces of the abstract n-particle Hilbert space,

$$H_n^B = O_s^{(n)} H_n \quad , \tag{37}$$

$$H_n^F = O_{as}^{(n)} H_n \quad . \tag{38}$$

If $\widehat{I}_{\widehat{H}}$ denotes the identity operator in the Fock space[2], then this operator can be written as the sum over all projections P_{H_n} into the subspaces H_n,

$$\widehat{I}_{\widehat{H}} = \sum_{n=0}^{\infty} P_{H_n} \quad , \tag{39}$$

where the projection operator $P_{H_n} = \sum\limits_{n=0}^{\infty} \mid n \rangle \langle n \mid$ in turn is the identity operator in the n-particle Hilbert space H_n.

---

[2] In the following all operators in the Fock space are denoted by hats, all representations (matrices) of such operators are underlined and carry a hat! The same applies to a basis (vector) in the Fock space, it is underlined and denoted by a hat.

15.3.2 Realizations of the abstract Fock space

The configurational space The abstract Fock space can be represented by the space $\widehat{H}^{(c)}$ of all quadratic integrable functions of the variables $x_1, x_2, ..., x_n, ...$. In the configurational space $\widehat{H}^{(c)}$ the identity operator (39) is given by

$$\widehat{I}_{\widehat{H}^{(c)}} = \sum_{n=0}^{\infty} P_{H_n^{(c)}} = \sum_{n=0}^{\infty} \int d^n x \mid x_n, ..., x_2, x_1\rangle\langle x_1, x_2, ..., x_n \mid \quad . \tag{40}$$

The occupation number space In order to use the concept of occupation numbers for a realization of the abstract Fock space, two different kinds of representations can be constructed. By selecting in each (n-particle) subspace H_n a basis of vectors

$$\{\mid n_i\rangle\} = \{\mid n_1, n_2, ..n_i, ...; n\rangle\} \quad , \tag{41}$$

such that the following sum rule is fulfilled

$$\sum_{i=1}^{\infty} n_i = n \quad , \tag{42}$$

the identity operator (39) can be represented by

$$\widehat{I}_{\widehat{H}^{(d)}} = \sum_{n=0}^{\infty} P_{H_n^{(d)}} = \sum_{n=0}^{\infty} \left\{ \sum_{\{\mid n_i\rangle\}} \mid n; ..., n_2, n_1\rangle\langle n_1, n_2, ...; n \mid \right\} \quad , \tag{43}$$

where the superscript (d) stands for discrete. $\widehat{H}^{(d)}$ is sometimes called the **discrete occupation number space**.

If one omits the sum rule in (42) as condition and considers the following orthonormal set

$$\{\mid n\rangle = \mid ..., n_2, n_1\rangle \ ; \langle m \mid n\rangle = \delta_{m_1 n_1}\delta_{m_2 n_2}..\delta_{m_i n_i}.... \ \} \quad , \tag{44}$$

then the representation of the identity in the Fock space is given by

$$\widehat{I}_{\widehat{H}^{(o)}} = \sum_{n_1, n_2, ...} \mid ..., n_2, n_1\rangle\langle n_1, n_2, ... \mid \quad , \tag{45}$$

i.e. the set $\{\mid n\rangle\}$ is a realization of a basis in the Fock space and contains the basis sets for the subspaces H_n as subsets. $\widehat{H}^{(o)}$ is called **occupation number space**. If one considers now the following subspaces R_i, $i = 1, 2, ...$, namely the Hilbert spaces corresponding to the occupation numbers $n_i, i = 1, 2....$, where i numbers the one-particle states $(\phi_i(x_i))$,

$$R_i = \{\mid n_i\rangle\} \quad , \tag{46}$$

then for fermions these subspaces $(R_i^F = \{\mid 0\rangle, \mid 1\rangle\})$ are two-dimensional, whereas for bosons the dimensionality of these subspaces $(R_i^B = \{\mid 0\rangle, \mid 1\rangle, \mid$

$2\rangle, ...\}$) is infinite. According to the resolution of the identity in (45), the occupation number space $\widehat{H}^{(o)}$ can be written as the following tensorial product of subspaces,

$$\widehat{H}^{(o)} = R_1 \otimes R_2 \otimes ...R_i \otimes ... \quad , \tag{47}$$

where of course for fermions the antisymmetric and for bosons the symmetric tensorial product has to be taken.

15.4 Representation of states and operators in Fock space

15.4.1 Abstract Fock space

States Let $\underline{\widehat{n}}$ be an orthonormal basis in the abstract Fock space,

$$\underline{\widehat{n}} = \{| 0\rangle, | 1\rangle, | 2\rangle, ..., | n\rangle, ...\} \quad , \tag{48}$$

$$\underline{\widehat{n}}^{\dagger}\underline{\widehat{n}} = \underline{\widehat{1}}_{\widehat{H}} \quad , \quad \widehat{I}_{\widehat{H}} = \sum_{n=0}^{\infty} | n\rangle\langle n | \quad , \tag{49}$$

where $\underline{\widehat{1}}_{\widehat{H}}$ (unit matrix) is the representation of the identity operator $\widehat{I}_{\widehat{H}}$ in $\widehat{H}$. Any arbitrary element $\widehat{\psi} \in \widehat{H}$ can then be represented in this basis,

$$\widehat{\psi} = \sum_{n=0}^{\infty} | n\rangle\langle n | \widehat{\psi} = \sum_{n=0}^{\infty} | n\rangle\psi_n = \underline{\widehat{n}}\underline{\widehat{\psi}} \quad , \tag{50}$$

where $\psi_n = \langle n | \widehat{\psi}$ is its component in the (abstract) Hilbert space H_n. The column vector $\underline{\widehat{\psi}}$ is of the following form

$$\underline{\widehat{\psi}} = \begin{pmatrix} \psi_o \\ \psi_1 \\ \vdots \\ \psi_n \\ \vdots \end{pmatrix} \quad . \tag{51}$$

The norm of $\widehat{\psi} \in \widehat{H}$ is defined by

$$\| \widehat{\psi} \|^2 = \left\{\widehat{\psi} \mid \widehat{\psi}\right\} = \left\{\underline{\widehat{n}}\underline{\widehat{\psi}} \mid \underline{\widehat{n}}\underline{\widehat{\psi}}\right\} = \underline{\widehat{\psi}}^{\dagger} \{\underline{\widehat{n}} \mid \underline{\widehat{n}}\} \underline{\widehat{\psi}} = \sum_{n=0}^{\infty}\langle\psi_n \mid \psi_n\rangle \quad , \tag{52}$$

where $\langle\psi_n \mid \psi_n\rangle$ is the "traditional" scalar product[3] in the (abstract) n-particle Hilbert space H_n. If $\widehat{\psi}$ is a pure n-particle state, say $\widehat{\phi}_N$, then this state is represented by

$$\widehat{\phi}_N = | n\rangle\langle n | \widehat{\phi}_N = \underline{\widehat{n}}\underline{\widehat{\phi}}_N \quad , \tag{53}$$

---

[3] see also chapter 14

where

$$\underline{\widehat{\phi}}_N = \begin{pmatrix} 0 \\ \vdots \\ (\phi_N)_n \\ 0 \\ \vdots \end{pmatrix} \quad . \tag{54}$$

A very particular case is the so-called **vacuum state** $\widehat{\psi}_0$, which formally is identical to the zero basis vector $\mid 0\rangle$

$$\widehat{\psi}_0 = \mid 0\rangle\langle 0 \mid \widehat{\psi}_0 = \mid 0\rangle \underbrace{\langle 0 \mid 0\rangle}_{=1} = \mid 0\rangle = \underline{\widehat{n}}\,\underline{\widehat{\psi}}_0 \quad , \tag{55}$$

where

$$\underline{\widehat{\psi}}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \tag{56}$$

clearly is normalized to unity.

Operators An operator $\widehat{A}$, defined in Fock space, can be represented in the following way

$$\widehat{A} = \sum_{i=0}^{\infty}\sum_{k=0}^{\infty} \mid i\rangle\langle i \mid \widehat{A} \mid k\rangle\langle k \mid = \sum_{i=0}^{\infty}\sum_{k=0}^{\infty} \mid i\rangle A_{ik} \langle k \mid = \underline{\widehat{n}\widehat{A}\widehat{n}}^{\dagger} \quad , \tag{57}$$

where A_{ik} is the ik-th element of the matrix $\underline{\widehat{A}}$,

$$\underline{\widehat{A}} = \begin{pmatrix} A_{00} & A_{01} & \ldots & A_{0n} & \cdots \\ A_{10} & A_{11} & \cdots & \cdots & \cdots \\ \vdots & \vdots & \cdots & \cdots & \cdots \\ A_{n0} & \vdots & \ldots & A_{nn} & \cdots \\ \vdots & \vdots & \ldots & \cdots & \ldots \end{pmatrix} \quad . \tag{58}$$

The operator $\widehat{A}$ maps an element $\widehat{\psi} \in \widehat{H}$ onto an element $\widehat{\phi} \in \widehat{H}$,

$$\widehat{A}\widehat{\psi} = \widehat{\phi} \quad . \tag{59}$$

Represented in the basis $\underline{\widehat{n}}$ this mapping yields

$$\left[\underline{\widehat{n}\widehat{A}\widehat{n}}^{\dagger}\right] \underline{\widehat{n}}\,\underline{\widehat{\psi}} = \underline{\widehat{n}}\,\underline{\widehat{\phi}} \rightarrow \; \underline{\widehat{A}}\,\underline{\widehat{\psi}} = \underline{\widehat{\phi}} \quad or: \sum_{k=0}^{\infty} A_{ik}\psi_k = \phi_i \quad . \tag{60}$$

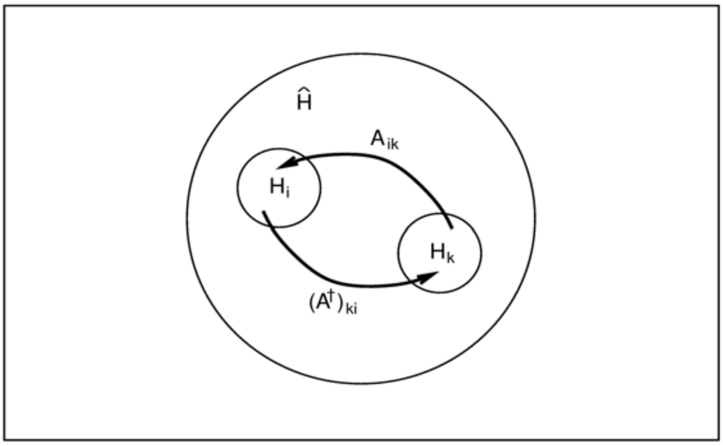


Figure 18: Mappings between Hilbert spaces

This looks pretty much like "ordinary" matrix algebra with a slightly more fancy notation. **It is not**! Maybe this is partially the source of a rather common misunderstanding of the concepts of the second quantization, because the matrix elements A_{ik} are in general **mappings from the k-particle Hilbert space onto the i-particle Hilbert space** (see also *Figure 18*). Quite clearly in order to define the **adjoint operators** $(A_{ik})^\dagger = (A^\dagger)_{ki}$, one carefully has to keep track of properly defined scalar products.

Consider first the following elements in the Hilbert space H_i,

$$\psi_i,\ A_{ik}\psi_k \in H_i \quad ; \quad \psi_k \in H_k \quad , \tag{61}$$

where the scalar product

$$\langle \psi_i \mid A_{ik}\psi_k \rangle \equiv \langle \psi_i \mid A_{ik}\psi_k \rangle_{H_i} \quad , \tag{62}$$

is of course well-defined, which is indicated by the additional subscript H_i. The domain of the mapping A_{ik}, $D_{A_{ik}}$, however, is the space H_k,

$$D_{A_{ik}} = H_k \quad . \tag{63}$$

Since $\psi_k \in H_k$ can always be expressed as a linear combination of elements $\chi_k \in H_k$, the scalar product in (62) is also a linear functional $\mathcal{L}(\psi_k)_{H_k}$ defined in H_k,

$$\mathcal{L}(\psi_k)_{H_k} = \langle \psi_i \mid A_{ik}\psi_k \rangle_{H_i} \quad . \tag{64}$$

Consequently, according to the Fischer-Riesz theorem[4] there exists one and only one element $\phi_k \in H_k$, such that

$$\mathcal{L}(\psi_k)_{H_k} = \langle \phi_k \mid \psi_k \rangle_{H_k} \quad , \tag{65}$$

---

[4] see also chapter 14

where the scalar product $\langle \phi_k \mid \psi_k \rangle_{H_k}$ is defined in H_k. If one defines now the adjoint operator $A^\dagger_{ki}$ such that

$$\phi_k = (A^\dagger)_{ki}\psi_i = \langle k \mid A^\dagger \mid i \rangle \psi_i \quad , \tag{66}$$

then obviously the following crucial equation is valid

$$\boxed{\langle \psi_i \mid \ A_{ik}\psi_k \rangle_{H_i} = \langle (A^\dagger)_{ki}\psi_i \mid \psi_k \rangle_{H_k}} \tag{67}$$

It should be noted that on the lhs of the last equation there is a scalar product in the (abstract) Hilbert space H_i, whereas on the right hand side it is a scalar product in H_k!

Obviously according to (67) two kinds of operators in the abstract Fock space can be distinguished, namely

$$\begin{array}{lll} A_{ik} = 0 & i \neq k & \text{first kind (order) or diagnoal operators} \\ A_{ik} \neq 0 & i \neq k & \text{second kind (order) operators} \end{array} \quad . \tag{68}$$

15.4.2 The configurational space $\widehat{H}^{(c)}$

In the configurational space $\widehat{H}^{(c)}$ the basis $\underline{\widehat{n}}$ is given by

$$\underline{\widehat{n}} = \{\mid 0\rangle, \{x_1\}, \{x_1, x_2\}, ..., \{x_1, x_2, ..., x_n\}, ...\} \quad , \tag{69}$$

from which elements and operators in $\widehat{H}^{(c)}$ can be constructed.

Elements Representations of elements in $\widehat{H}^{(c)}$ are obtained in the following way

$$\widehat{\psi} = \sum_{n=0}^{\infty} \int d^n x \mid x_n, ..., x_2, x_1 \rangle \langle x_1, x_2, ..., x_n \mid \widehat{\psi} = \underline{\widehat{n}}\underline{\widehat{\psi}} \quad , \tag{70}$$

$$\langle x_1, x_2, ..., x_n \mid \widehat{\psi} = \psi_n(x_1, x_2, ..., x_n) \quad ,$$

where $\underline{\widehat{\psi}}$ and the vacuum state $\underline{\widehat{\psi}}_0$ are of the form

$$\underline{\widehat{\psi}} = \begin{pmatrix} \psi_0 \\ \psi_1(x_1) \\ \psi_2(x_1, x_2) \\ \vdots \\ \vdots \end{pmatrix} \quad , \quad \underline{\widehat{\psi}}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \end{pmatrix} \quad . \tag{71}$$

The norm of such an element in $\widehat{H}^{(c)}$ is given by

$$\parallel \widehat{\psi} \parallel^2 = \sum_{n=0}^{\infty} \int d^n x \psi_n^*(x_1, x_2, ..., x_n)\psi_n(x_1, x_2, ..., x_n) \quad . \tag{72}$$

Operators In a similar way representations of operators in $\widehat{H}^{(c)}$ are obtained,

$$\widehat{A} = \sum_{i=0}^{\infty} \sum_{n=0}^{\infty} \int \int d^i x d^k y \mid x_i, ..., x_2, x_1 \rangle A_{ik} \langle y_1, y_2, ..., y_k \mid \quad , \tag{73}$$

where the matrix elements are mappings from one subspace of $\widehat{H}^{(c)}$ to another one. These mappings are general functions of the following kind

$$\begin{aligned} A_{ik} &= \langle x_1, x_2, ..., x_i \mid A_{ik} \mid y_1, y_2, ..., y_k \rangle = \\ &= A_{ik}(x_1, x_2, ..., x_i; y_1, y_2, ..., y_k) \quad . \end{aligned} \tag{74}$$

The adjoint operators are then defined according to (67) as

$$\begin{aligned} &\int d^i x \psi_i^*(x_1, x_2, ..., x_i) \times \\ &\times \left[\int d^k y A_{ik}(x_1, x_2, ..., x_i; y_1, y_2, ..., y_k) \psi_k(y_1, y_2, ..., y_k) \right] = \\ &\int d^k x \psi_k(x_1, x_2, ..., x_k) \quad \times \\ &\times \left[\int d^i y A_{ki}^{\dagger}(x_1, x_2, ..., x_k; y_1, y_2, ..., y_i) \psi_i(y_1, y_2, ..., y_i) \right]^* \quad . \end{aligned} \tag{75}$$

15.4.3 Occupation number space representations

If one wants to construct representations in the **discrete occupation number space** (see (41)-(43)) the corresponding basis is defined by

$$\underline{\widehat{n}}^{(d)} = \{\mid 0 \rangle, \mid \underline{1} \rangle, ..., \mid \underline{n} \rangle ...\} \quad , \tag{76}$$

where the basis $\mid \underline{n} \rangle \equiv \mid n; n_1, n_2, .. \rangle$, $\sum\limits_{i=1}^{\infty} n_i = n$, refers to the sum rule restricted sets of occupation numbers. In terms of this basis a general element in the Fock space is then described by

$$\widehat{\psi} = \sum_{n=0}^{\infty} \mid \underline{n} \rangle \underline{f}_n = \underline{\widehat{n}}^{(d)} \underline{\widehat{\psi}}^{(d)} \quad ; \quad \underline{f}_n = \langle \underline{n} \mid \widehat{\psi} \quad , \tag{77}$$

and has the following norm

$$\parallel \underline{\widehat{\psi}}^{(d)} \parallel^2 = \{ \underline{\widehat{\psi}}^{(d)} \mid \underline{\widehat{\psi}}^{(d)} \} = \sum_{n=0}^{\infty} \langle \underline{f}_n \mid \underline{f}_n \rangle \quad . \tag{78}$$

Representations of operators are formulated in a similar way as before, namely,

$$\widehat{A} = \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \mid \underline{i} \rangle A_{ik}^{(d)} \langle \underline{k} \mid = \underline{\widehat{n}}^{(d)} \underline{\widehat{A}}^{(d)} \left(\underline{\widehat{n}}^{(d)} \right)^{\dagger} \quad ; \quad \underline{A}_{ik}^{(d)} = \langle \underline{i} \mid \widehat{A} \mid \underline{k} \rangle \quad , \tag{79}$$

where the adjoint operators are given by

$$\left(\underline{A}_{ik}^{(d)}\right)^{\dagger} = \left((\widehat{\underline{A}}^{(d)})^{\dagger}\right)_{ki} \quad . \tag{80}$$

Consider now for a given set of occupation numbers $\{n_i\}$ the corresponding projection operator

$$\mid ..., n_i, ...n_2, n_1, \rangle\langle n_1, n_2, ..., n_i, ... \mid \quad , \sum_{i=1}^{\infty} n_i \neq n \quad , \tag{81}$$

then

$$\psi_{\{n_i\}}^{(o)} = \langle n_1, n_2, ..., n_i, ... \mid \widehat{\psi} \tag{82}$$

is the component of $\widehat{\psi}$ in this (one-dimensional) subspace. In principle therefore the matrix element of the Fock space operator $\widehat{A}$ corresponding to the sets $\{n_i\}$ and $\{n_i'\}$,

$$A_{\{n_i\},\{n_i'\}} = \langle n_1, n_2, ..., n_i, ... \mid \widehat{A} \ \mid ..., n_i', ...n_2', n_1', \rangle \quad . \tag{83}$$

maps the component $\psi_{\{n_i\}}^{(o)}$ into a component $\psi_{\{n_i'\}}^{(o)}$. The adjoint representation can formally be written as

$$A_{\{n_i'\},\{n_i\}}^{\dagger} = \langle n_1', n_2', ..., n_i', ... \mid \widehat{A} \ \mid ..., n_i, ...n_2, n_1, \rangle \quad , \tag{84}$$

15.5 Creation and annihilation operators in the abstract Fock space

15.5.1 Annihilation operators

An annihilation operator $\widehat{b}(\phi)$, where ϕ is an arbitrary element from the one-particle Hilbert space, $\phi \in H_1$, is an operator of second kind with the following representation in the Fock space

$$\underline{\widehat{b}}(\phi) = \begin{pmatrix} 0 & b_{01}(\phi) & 0 & 0 & \cdots \\ 0 & 0 & b_{12}(\phi) & 0 & \cdots \\ 0 & 0 & 0 & b_{23}(\phi) & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad , \tag{85}$$

where the matrix elements have the following meaning:

$$b_{n-1,n}(\phi) = \langle n-1 \mid \widehat{b}(\phi) \mid n\rangle = \sqrt{n}\langle\phi \mid \tag{86}$$

The operator $b_{n-1,n}(\phi)$ is defined in H_n and maps an element $\psi_n \in H_n$ into an element in H_{n-1},

$$b_{n-1,n}(\phi)\psi_n = \sqrt{n}\langle\phi \mid \psi_n\rangle \quad ; \quad \langle\phi \mid \psi_n\rangle \in H_{n-1} \quad . \tag{87}$$

If therefore $\widehat{b}(\phi)$ acts onto an element $\widehat{\psi} \in \widehat{H}$, then in each subspace H_n one particle is removed by means of the one-particle state ϕ . The annihilation operator $\widehat{b}(\phi)$ is only unique if $\widehat{\psi}$ describes a system of non-distinguishable particles such as bosons or fermions,

$$\langle\phi \mid \psi_n^B\rangle = O_s^{(n-1)}\langle\phi \mid \psi_n^B\rangle \quad , \tag{88}$$

$$\langle\phi \mid \psi_n^F\rangle = O_{as}^{(n-1)}\langle\phi \mid \psi_n^F\rangle \quad . \tag{89}$$

15.5.2 Creation operators

The adjoint operator of the annihilation operator $\widehat{b}(\phi)$ is the creation operator and is defined in terms of its matrix elements

$$b^\dagger_{n,n-1}(\chi) = \langle n \mid \widehat{b}(\chi) \mid n-1\rangle = \sqrt{n}\chi \quad ; \quad \chi \in H_n \quad . \tag{90}$$

If $\psi_{n-1} \in H_{n-1}$, then the operator $b^\dagger_{n,n-1}(\chi)$ maps ψ_{n-1} into an element $\in H_n$,

$$b^\dagger_{n,n-1}(\chi)\psi_{n-1} = \sqrt{n}\chi\psi_{n-1} \quad . \tag{91}$$

It should be noted that $\chi\psi_{n-1}$ is the direct product of χ and ψ_{n-1}. The creation operator is also an operator of second kind in the Fock space with the following matrix structure

$$\underline{\widehat{b}}^\dagger(\chi) = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ b^\dagger_{10}(\chi) & 0 & 0 & 0 & \cdots \\ 0 & b^\dagger_{21}(\chi) & 0 & 0 & \cdots \\ 0 & 0 & b^\dagger_{32}(\chi) & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad . \tag{92}$$

Since the mapping is from H_{n-1} onto H_n (anti-)symmetrization with respect to the permutational group has to be performed in each subspace. If therefore $\mid n^B\rangle$ and $\mid n^F\rangle$ are properly symmetrized basis sets in H_n,

$$\mid n^B\rangle = O_s^{(n)} \mid n\rangle \quad , \tag{93}$$

$$\mid n^F\rangle = O_{as}^{(n)} \mid n\rangle \quad , \tag{94}$$

then the following symmetry conserving property for creation operators is valid

$$\left[b^\dagger_{n,n-1}(\chi)\right]^{B/F} \psi_{n-1}^{B/F} = \sqrt{n}O_{s/as}^{(n)}\chi\psi_{n-1}^{B/F} \quad , \tag{95}$$

where

$$\left[b^\dagger_{n,n-1}(\chi)\right]^{B} = \langle n \mid O_s^{(n)}\widehat{b}^\dagger(\chi)O_s^{(n-1)} \mid n-1\rangle \quad , \tag{96}$$

$$\left[b^\dagger_{n,n-1}(\chi)\right]^{F} = \langle n \mid O_{as}^{(n)}\widehat{b}^\dagger(\chi)O_{as}^{(n-1)} \mid n-1\rangle \quad . \tag{97}$$

15.5.3 Commutator and anticommutator relations

From the representations of $\widehat{b}(\phi)$ and $\widehat{b}^\dagger(\chi)$, $\phi, \chi \in H_1$ in (85) and (92), one can see that the product operators $\widehat{b}^\dagger(\chi)\widehat{b}(\phi)$ and $\widehat{b}(\phi)\ \widehat{b}^\dagger(\chi)$ are operators of first kind in the Fock space, i.e., are diagonal operators,

$$\widehat{b}^\dagger(\chi)\widehat{b}(\phi) = \begin{pmatrix} b^\dagger_{10}(\chi)b_{01}(\phi) & 0 & 0 & \cdots \\ 0 & b^\dagger_{21}(\chi)b_{12}(\phi) & 0 & \cdots \\ 0 & 0 & b^\dagger_{32}(\chi)b_{23}(\phi) & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix} \quad , \tag{98}$$

$$\widehat{b}(\phi)\widehat{b}^\dagger(\chi) = \begin{pmatrix} b_{01}(\phi)b^\dagger_{10}(\chi) & 0 & 0 & \cdots \\ 0 & b_{12}(\phi)b^\dagger_{21}(\chi) & 0 & \cdots \\ 0 & 0 & b_{23}(\phi)b^\dagger_{32}(\chi) & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix} \quad . \tag{99}$$

For an arbitrary element $\widehat{\psi} \in \widehat{H}$ the action of these product operators is therefore confined to a particular Hilbert space H_n

$$\begin{aligned} \langle n \mid \widehat{b}^\dagger(\chi)\widehat{b}(\phi) \mid n\rangle\psi_n &= \langle n \mid \widehat{b}^\dagger(\chi) \mid n-1\rangle\langle n-1 \mid \widehat{b}(\phi) \mid n\rangle\psi_n = \\ &= b^\dagger_{n,n-1}(\chi)b_{n-1,n}(\phi)\psi_n \quad , \end{aligned} \tag{100}$$

$$\begin{aligned} \langle n \mid \widehat{b}(\phi)\widehat{b}^\dagger(\chi) \mid n\rangle\psi_n &= \langle n \mid \widehat{b}(\phi) \mid n+1\rangle\langle n+1 \mid \widehat{b}^\dagger(\chi) \mid n\rangle\psi_n = \\ &= b_{n,n+1}(\phi)b^\dagger_{n+1,n}(\chi)\psi_n \quad . \end{aligned} \tag{101}$$

Bosons For bosons it can be shown that the last two equations yield

$$b^\dagger_{n,n-1}(\chi)b_{n-1,n}(\phi)\psi^B_n = n\chi\langle\phi \mid \psi^B_n\rangle \quad , \tag{102}$$

$$b_{n,n+1}(\phi)b^\dagger_{n+1,n}(\chi)\psi^B_n = \langle\phi \mid \chi\rangle\psi^B_n + n\chi\langle\phi \mid \psi^B_n\rangle \quad , \tag{103}$$

which subtracted from each other define the commutator of the creation and annihilation operator in Fock space,

$$\boxed{\left[\widehat{b}(\phi), \widehat{b}^\dagger(\chi)\right]_- = \langle\phi \mid \chi\rangle\widehat{I}_B} \tag{104}$$

where $\widehat{I}_B$ is the identity operator in the boson-part of the Fock space. By simply reading off the respective matrix structures, one can see that

$$\boxed{\left[\widehat{b}(\phi), \widehat{b}(\chi)\right]_- = \left[\widehat{b}^\dagger(\phi), \widehat{b}^\dagger(\chi)\right]_- = 0} \tag{105}$$

Fermions Explicitly evaluated for fermions equations (100) and (101) give

$$b^{\dagger}_{n,n-1}(\chi)b_{n-1,n}(\phi)\psi^F_n = n\chi\langle\phi \mid \psi^F_n\rangle \quad , \tag{106}$$

$$b_{n,n+1}(\phi)b^{\dagger}_{n+1,n}(\chi)\psi^F_n = \langle\phi \mid \chi\rangle\psi^F_n - n\chi\langle\phi \mid \psi^F_n\rangle \quad , \tag{107}$$

from which one can see that in this case the anticommutator of the creation and annihilation operator in Fock space is of the following form

$$\boxed{\left[\widehat{b}(\phi), \widehat{b}^{\dagger}(\chi)\right]_+ = \langle\phi \mid \chi\rangle\widehat{I}_F} \tag{108}$$

where $\widehat{I}_F$ now is the identity operator in the fermion-part of the Fock space. In a similar way one obtains the equivalent forms of (105), namely

$$\boxed{\left[\widehat{b}(\phi), \widehat{b}(\chi)\right]_+ = \left[\widehat{b}^{\dagger}(\phi), \widehat{b}^{\dagger}(\chi)\right]_+ = 0} \tag{109}$$

Quite clearly, for bosons as well as for fermions, a special case arises if the one particle functions $\phi = \phi_i$ and $\chi = \phi_j$ are orthogonal, $\langle\phi_i \mid \phi_j\rangle = \delta_{ij}$.

15.6 Creation and annihilation operators in the occupation number space

Acting on the vacuum state $\widehat{\psi}_0 =\mid 0\rangle$, $\widehat{\psi}_0 \in \widehat{H}$, the creation operator yields the following result

$$\widehat{b}^{\dagger}(\phi)\widehat{\psi}_0 = \underline{\widehat{n}\widehat{b}}^{\dagger}(\phi)\underline{\widehat{n}}^{\dagger}\underline{\widehat{n}\widehat{\psi}}_0 =\mid 1\rangle\langle 1 \mid \widehat{b}^{\dagger}(\phi) \mid 0\rangle\underbrace{\langle 0 \mid 0\rangle}_{=1} =$$

$$= \mid 1\rangle b^{\dagger}_{10}(\phi) =\mid 1\rangle\phi \equiv \widehat{\phi}_1 \quad , \tag{110}$$

namely a pure one-particle state in $\widehat{H}$ (see (54)),

$$\widehat{\phi}_1 = \underline{\widehat{n}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \quad . \tag{111}$$

By repeated application of the creation operator $\widehat{b}^{\dagger}(\phi^{\alpha_i})$ on the vacuum state,

$$\widehat{b}^{\dagger}(\phi^{\alpha_n})\widehat{b}^{\dagger}(\phi^{\alpha_{n-1}})....\widehat{b}^{\dagger}(\phi^{\alpha_1}) \mid 0\rangle = \sqrt{n!}O^{(n)}_{s/as}\widehat{\phi}_1^{\alpha_n}\widehat{\phi}_1^{\alpha_{n-1}}...\widehat{\phi}_1^{\alpha_1} \quad , \tag{112}$$

a pure n-particle state in the Fock space can be generated. If one chooses now an **orthonormal set of one-particle functions**, then by means of the last equation and by making use of the transformation described in (4)-(8) a suitable basis for representations in the occupation number space $\widehat{H}^{(o)}$ can be obtained.

15.6.1 Bosons

Let $\mid n_1, n_2, ... : n\rangle^B$ be a pure n-boson state,

$$\mid n_1, n_2, ... : n\rangle^B \in \widehat{H}^{(o)} \quad , \tag{113}$$

$$\mid n_1, n_2, ... : n\rangle^B = \underline{\widehat{n}}^B \underline{\widehat{\phi}}_N^B = (\mid 0\rangle, \mid 1\rangle, ..)^B \begin{pmatrix} 0 \\ \vdots \\ (\phi_N^B)_n \\ 0 \\ \vdots \end{pmatrix} \quad , \tag{114}$$

then by using (32) this state is given by

$$\mid n_1, n_2, ... : n\rangle^B = \mid n\rangle (\phi_N^B)_n = \mid n\rangle \left\{ \sqrt{\frac{n!}{n_1! n_2! ...}} O_s^{(n)} \prod_{i=1}^{n} \phi_{k_i} \right\} \quad . \tag{115}$$

This state, however, as was said before can be only obtained by means of the creation operators in (112). Abbreviating for a moment $\widehat{b}^\dagger(\phi_i)$ by $\widehat{b}_i^\dagger$ and $\widehat{b}(\phi_i)$ by $\widehat{b}_i$, the pure n-boson state has to be derived from repeated application of the creation operator on the vacuum state

$$\mid n_1, n_2, ... : n\rangle^B = \frac{1}{\sqrt{n_1!}} \left[\widehat{b}_1^\dagger\right]^{n_1} \frac{1}{\sqrt{n_2!}} \left[\widehat{b}_2^\dagger\right]^{n2} ... \mid 0\rangle \quad . \tag{116}$$

By acting with $\widehat{b}_i^\dagger$ or $\widehat{b}_i$ on the pure n-boson state one can show that

$$\widehat{b}_i^\dagger \mid n_1, n_2, ... : n\rangle^B = \sqrt{n_i + 1} \mid n_1, n_2, ..., n_i + 1, \ldots : n + 1\rangle^B \quad , \tag{117}$$

$$\widehat{b}_i \mid n_1, n_2, ... : n\rangle^B = \sqrt{n_i} \mid n_1, n_2, ..., n_i - 1, \ldots : n - 1\rangle^B \quad . \tag{118}$$

According to the general commutator relations for the creation and annihilation operator in (104)-(105), these relations are now given by

$$\boxed{\left[\widehat{b}_i, \widehat{b}_j^\dagger\right]_- = \delta_{ij} \widehat{I}_{\widehat{H}_B} \quad , \quad \left[\widehat{b}_i, \widehat{b}_j\right]_- = \left[\widehat{b}_i^\dagger, \widehat{b}_j^\dagger\right]_- = 0} \tag{119}$$

where the product operators $\widehat{b}_i^\dagger \widehat{b}_i$ and $\widehat{b}_i \widehat{b}_i^\dagger$ have the following properties (see also (102) and (103))

$$\widehat{b}_i^\dagger \widehat{b}_i \mid n_1, n_2, ..., n_i, ... : n\rangle^B = n_i \mid n_1, n_2, ..., n_i, ... : n\rangle^B \quad , \tag{120}$$

$$\widehat{b}_i \widehat{b}_i^\dagger \mid n_1, n_2, ..., n_i, ... : n\rangle^B = (n_i + 1) \mid n_1, n_2, ..., n_i, ... : n\rangle^B \quad . \tag{121}$$

Because of equation (120) the product operator $\widehat{b}_i^\dagger \widehat{b}_i$ is also called number operator N_i, which, however, is only defined in certain subspace of $\widehat{H}^{(o)}$, namely R_i , see (46). The sum over these number operators N_i,

$$\boxed{\widehat{N} = \sum_{i=1}^{\infty} N_i} \tag{122}$$

is again an operator in the Fock space and is generally called the **number operator** $\widehat{N}$. The properties of $\widehat{N}$ follow from (120),

$$\widehat{N} \mid n_1, n_2, ..., n_i, ... : n\rangle^B \equiv \widehat{N} \mid n\rangle^B = n \mid n\rangle^B \quad . \tag{123}$$

According to this equation , the number operator $\widehat{N}$ has the following representation in $\widehat{H}^{(o)}$,

$$\mid n\rangle\langle n \mid \widehat{N} \mid m\rangle\langle m \mid = m \underbrace{\mid n\rangle\langle n \mid}_{=1} \underbrace{m\rangle\langle m \mid}_{=1} = m \quad , \tag{124}$$

$$\underline{\widehat{N}} = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 2 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad . \tag{125}$$

15.6.2 Fermions

For fermions the procedure is pretty much the same as for bosons with the exception that only **ordered configurations** (22) can be used. Let $\mid n_1, n_2, ..., n_i, ... : n\rangle^F$ be a pure n-fermion state and $K = \{k_1 < k_2 ... < k_n\}$ such an ordered configuration, then

$$\begin{aligned} \mid n_1, n_2, ..., n_i, ... : n\rangle^F &= \widehat{b}^\dagger(\phi_{k_1})\widehat{b}^\dagger(\phi_{k_2})...\widehat{b}^\dagger(\phi_{k_n}) \mid 0\rangle = \\ &= \mid n\rangle(\phi_K^F)_n = \mid n\rangle \left\{ \sqrt{n!} O_{as}^{(n)} \phi_{k_1}\phi_{k_2}...\phi_{k_n} \right\} \quad . \end{aligned} \tag{126}$$

By acting now with $\widehat{b}^\dagger(\phi_i) = \widehat{b}_i^\dagger$ or $\widehat{b}(\phi_i) = \widehat{b}_i$ on this pure n-fermion state, one gets

$$\begin{aligned} \widehat{b}_i^\dagger \mid n_1, n_2, ..., n_i, ... : n\rangle^F &= \\ = (-1)^{s_i}(1 - n_i) \mid n_1, n_2, ..., n_i + 1, ... : n + 1\rangle^F \end{aligned} \tag{127}$$

and

$$\begin{aligned} \widehat{b}_i \mid n_1, n_2, ..., n_i, ... : n\rangle^F &= \\ = (-1)^{s_i} n_i \mid n_1, n_2, ..., n_i - 1, ... : n - 1\rangle^F \end{aligned} \tag{128}$$

where

$$s_i = \sum_{k=0}^{i-1} n_k \quad . \tag{129}$$

The anticommutator relations follow from the general relations in (108) and (109),

$$\boxed{\left[\widehat{b}_i, \widehat{b}_j^\dagger\right]_+ = \delta_{ij} I_{\widehat{H}_F} \quad , \quad \left[\widehat{b}_i, \widehat{b}_j\right]_+ = \left[\widehat{b}_i^\dagger, \widehat{b}_j^\dagger\right]_+ = 0} \tag{130}$$

The definition of the number operator is the same as for bosons

$$\boxed{\widehat{N} = \sum_{i=1}^{\infty} \widehat{b}_i^\dagger \widehat{b}_i = \sum_{i=1}^{\infty} N_i} \tag{131}$$

15.7 Field operators

In $H_1^{(c)}$, $H_1^{(c)} \subset \widehat{H}^{(c)}$, the projection operator for one particle states is given by

$$P_{H_1^{(c)}} = \int \mid x\rangle\langle x \mid dx \quad , \tag{132}$$

where as is well-known $\mid x\rangle$ strictly speaking is not normalizable in H_1. Formally, however, a pure one-particle state in $\widehat{H}^{(c)}$ can be written as

$$\mid \widehat{x}\rangle = \begin{pmatrix} 0 \\ \mid x\rangle \\ 0 \\ 0 \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} o \\ x \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad , \tag{133}$$

$$\begin{aligned} \mid \widehat{x}\rangle = \sum_n \int d^n x \mid x_n, ..., x_2, x_1\rangle\langle x_1, x_2, ..., x_n \mid \widehat{x}\rangle = \\ = \int \mid x_1\rangle\langle x_1 \mid \widehat{x}\rangle dx_1 = \\ = \underline{\widehat{n}}^{(c)} \begin{pmatrix} 0 \\ \delta(x - x_1) \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad , \end{aligned} \tag{134}$$

where $\delta(x - x_1)$ is the **Dirac Deltafunction** (δ_x).

If one denotes now the creation and annihilation operators in the "traditional" notation by

$$\widehat{\Psi}^\dagger(x) \equiv \widehat{b}^\dagger(x) \quad , \quad \widehat{\Psi}(x) \equiv \widehat{b}(x) \quad , \tag{135}$$

then according to the general form of such operators they have the following matrix elements

$$b^\dagger_{n,n-1}(x) = \sqrt{n} O^{(n)}_{s/as} x \quad , \tag{136}$$

$$b_{n-1,n}(x) = \sqrt{n}\langle x \mid \quad . \tag{137}$$

The operators $\widehat{\Psi}^\dagger(x)$ and $\widehat{\Psi}(x)$ are usually called **field operators**. According to (134) the representations of the matrix elements of the field operators (in configurational space) are given by

$$b^\dagger_{n,n-1}(\delta_x) = \sqrt{n} O^{(n)}_{s/as} \delta(x - x_n) \quad , \tag{138}$$

$$b_{n-1,n}(\delta_x) = \sqrt{n} \int dx_n \delta(x - x_n) \quad . \tag{139}$$

The matrix element $b^{\dagger}_{n,n-1}(\delta_x)$ maps a pure $(n-1)$ particle function $\in H^{(c)}_{n-1}$ onto an n particle function $\in H^{(c)}_{n}$,

$$b^{\dagger}_{n,n-1}(\delta_x)\psi_{n-1}(x_1,x_2,...,x_{n-1}) =$$

$$= \sqrt{n}O^{(n)}_{s/as}\delta(x-x_n)\psi_{n-1}(x_1,x_2,...,x_{n-1}) \quad , \tag{140}$$

whereas by $b_{n-1,n}(\delta_x)$ an n particle function is mapped into a $(n-1)$ particle function

$$b_{n-1,n}(\delta_x)\psi_n(x_1,x_2,...,x_n) = \sqrt{n}\int dx_n\delta(x-x_n)\psi_n(x_1,x_2,...,x_n) =$$

$$= \sqrt{n}\psi_n(x_1,x_2,...,x_{n-1},x) \quad , \tag{141}$$

where in the last equation x now has to be viewed as a (dummy) parameter.

Suppose now a pure n-particle state is denoted by $\mid x_n,...,x_2,x_1\rangle$ with the following formal representations in the abstract Fock space $\widehat{H}$ and in the configurational space $\widehat{H}^{(c)}$, respectively,

$$\begin{pmatrix} 0 \\ 0 \\ O^{(n)}_{s/as}x_1x_2...x_n \\ 0 \\ \vdots \end{pmatrix} \quad , \quad \begin{pmatrix} 0 \\ 0 \\ O^{(n)}_{s/as}\prod_{m=1}^{n}\delta(x_m-x_m) \\ 0 \\ \vdots \end{pmatrix} \quad , \tag{142}$$

then as before with all other representations this state has to be generated from the vacuum state by means of repeated application of the creation operator $\widehat{b}^{\dagger}(x)$ (135),

$$\widehat{b}^{\dagger}(x_n)\mid 0\rangle = \mid x_n\rangle \quad , \quad \widehat{b}^{\dagger}(x_{n-1})\widehat{b}^{\dagger}(x_n)\mid 0\rangle = \sqrt{2}O^{(2)}_{s/as}\mid x_{n-1},x_n\rangle \quad , \tag{143}$$

i.e. is given by

$$\mid x_1,,x_2,...,x_n\rangle^{B/F} = \frac{1}{\sqrt{n!}}\widehat{b}^{\dagger}(x_1)\widehat{b}^{\dagger}(x_2)...\widehat{b}^{\dagger}(x_n)\mid 0\rangle \quad . \tag{144}$$

If one compares this result with the corresponding equations for the occupation number space representations (112), (115) and (127), one can see that the operators $\widehat{b}^{\dagger}(\phi_i)$ and $\widehat{b}^{\dagger}(x)$ are obviously related to each other by the same kind of general transformation (7) as discussed for n particles states right at the beginning of this chapter,

$$\mid x_1,,x_2,...,x_n\rangle = \widehat{b}^{\dagger}(x_1)...\widehat{b}^{\dagger}(x_n)\mid 0\rangle =$$

$$= \sum_{\{k_i\}}\mid k_1,k_2,...,k_n\rangle\langle k_n,k_{n-1},...k_1\mid x_1,,x_2,...,x_n\rangle =$$

$$= \sum_{\{k_i\}} \frac{1}{\sqrt{n!}} \widehat{b}^\dagger(\phi_{k_1}) ... \widehat{b}^\dagger(\phi_{k_n}) \mid 0\rangle \langle k_n, k_{n-1}, ...k_1 \mid x_1, , x_2, ..., x_n\rangle \quad . \tag{145}$$

The relation between the field operators and the creation and annihilation operators is therefore given by

$$\boxed{\widehat{b}^\dagger(x) = \sum_i \widehat{b}^\dagger(\phi_i)\langle \phi_i \mid x\rangle \equiv \sum_i \widehat{b}_i^\dagger \langle i \mid x\rangle} \tag{146}$$

$$\boxed{\widehat{b}(x) = \sum_i \langle x \mid \phi_i\rangle \widehat{b}(\phi_i) \equiv \sum_i \langle x \mid i\rangle \widehat{b}_i} \tag{147}$$

or (assuming orthogonal one particle functions) oppositely by

$$\boxed{\widehat{b}^\dagger(\phi_i) = \int dx \phi_i(x) \widehat{b}^\dagger(x)} \tag{148}$$

$$\boxed{\widehat{b}(\phi_i) = \int dx \phi_i^*(x) \widehat{b}(x)} \tag{149}$$

According to these relations, the commutator (bosons) or anticommutator relations (fermions) for field operators can be defined

$$\boxed{\left[\widehat{b}(x), \widehat{b}^\dagger(x')\right]_- = \delta(x-x')\widehat{I}_{\widehat{H}_B} \quad ; \left[\widehat{b}(x), \widehat{b}(x')\right]_- = \left[\widehat{b}^\dagger(x), \widehat{b}^\dagger(x')\right]_- = 0} \tag{150}$$

$$\boxed{\left[\widehat{b}(x), \widehat{b}^\dagger(x')\right]_+ = \delta(x-x')\widehat{I}_{\widehat{H}_F} \quad ; \left[\widehat{b}(x), \widehat{b}(x')\right]_+ = \left[\widehat{b}^\dagger(x), \widehat{b}^\dagger(x')\right]_+ = 0} \tag{151}$$

from which one clearly can see that in the configurational space the number operator is given by

$$\widehat{N} = \int \widehat{b}^\dagger(x) \widehat{b}(x) dx \quad . \tag{152}$$

15.8 A second quantization formulation for diagonal operators in Fock space

Suppose an operator $\widehat{A}$ in Fock space is of first kind,

$$\underline{\widehat{A}} = \begin{pmatrix} A_{11} & 0 & 0 & \cdots \\ 0 & A_{22} & 0 & \cdots \\ 0 & 0 & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad , \tag{153}$$

and the representations of its matrix elements in configurational space,

$$A_{nn} = \langle n \mid \iint d^n x dy^n \mid x_1, ..., x_n\rangle A_{nn}(x_1, .., x_n; y_1, ..., y_n)\langle y_1, ..., y_n \mid n\rangle \ , \tag{154}$$

are for example of the following form

$$A_{nn}(x_1, .., x_n; y_1, ..., y_n) = A(y_1, ..., y_n) \prod_{m=1}^{n} \delta(y_m - x_m) \quad . \tag{155}$$

Furthermore, it shall be assumed that the function $A(y_1, ..., y_n)$ can be written as a sum of the form

$$A(y_1, ..., y_n) = \sum_{\{l_1, \cdots, l_k\}}^{M} A(y_{l_1}, y_{l_2}, ..., y_{l_k}) \quad , \quad k < n \quad , M = \binom{n}{k} \quad . \tag{156}$$

The matrix elements in (155) can now be reformulated in terms of creation and annihilation operators. Consider first only one term in (156), $A(y_1, ..., y_{k)}$, to be used in (155),

$$\langle n \mid \iint d^n x d^n y \mid x_1, ..., x_n \rangle A(y_1, ..., y_k) \prod_{m=1}^{k} \delta(y_m - x_m) \langle y_1, ..., y_n \mid n \rangle =$$

$$\frac{1}{n!} \langle n \mid \int d^n x \widehat{b}^\dagger(x_1) ... \underbrace{\widehat{b}^\dagger(x_n) \mid 0 \rangle A(x_1, ..., x_k) \langle 0 \mid \widehat{b}(x_n)}_{step-wise} ... \widehat{b}(x_1) \mid n \rangle \quad . \tag{157}$$

By developing the integral step-wise, one gets

$$\frac{(n-k)!}{n!} \langle n \mid \int d^k x \widehat{b}^\dagger(x_1) ... \widehat{b}^\dagger(x_k) A(x_1, ..., x_k) \widehat{b}(x_k) ... \widehat{b}(x_1) \mid n \rangle \quad . \tag{158}$$

If one finally sums according to (156), the matrix elements of the operator $\widehat{A}$ are given by

$$\langle n \mid \widehat{A} \mid n \rangle = \langle n \mid \frac{1}{k!} \int d^k x \widehat{b}^\dagger(x_1) ... \widehat{b}^\dagger(x_k) A(x_1, ..., x_k) \widehat{b}(x_k) ... \widehat{b}(x_1) \mid n \rangle \quad , \tag{159}$$

and therefore the operator by

$$\widehat{A} = \frac{1}{k!} \int d^k x \widehat{b}^\dagger(x_1) ... \widehat{b}^\dagger(x_k) A(x_1, ..., x_k) \widehat{b}(x_k) ... \widehat{b}(x_1) \quad . \tag{160}$$

15.8.1 The Fock formulation of the Hartree-Fock equations

Suppose as an example for the above general formulation that the Hamilton operator H_{nn} for a n-particle system is restricted to two-particle interactions only,

$$H_{nn} = \sum_{i=1}^{n} h_i + \frac{1}{2} \sum_{i \neq j} g_{ij} \quad , \tag{161}$$

then the corresponding operator in Fock space, $\widehat{\mathcal{H}}$ can be formulated immediately using (160),

$$\widehat{\mathcal{H}} = \int dx \widehat{b}^{\dagger}(x) h(x) \widehat{b}(x) + \frac{1}{2} \iint dx_1 dx_2 \widehat{b}^{\dagger}(x_1) \widehat{b}^{\dagger}(x_2) g(x_1, x_2) \widehat{b}(x_2) \widehat{b}(x_1) \quad , \tag{162}$$

or

$$\widehat{\mathcal{H}} = \sum_{i,j} \widehat{b}_i^{\dagger} \langle i \mid h_1 \mid j \rangle \widehat{b}_j + \frac{1}{2} \sum_{i,j,k,l} \widehat{b}_i^{\dagger} \widehat{b}_j^{\dagger} \langle ij \mid g_{12} \mid kl \rangle \widehat{b}_k \widehat{b}_l \quad . \tag{163}$$

In order to evaluate the expectation value of $\widehat{\mathcal{H}}$ in Fock space,

$$\left\{ \widehat{\psi} \mid \widehat{\mathcal{H}} \widehat{\psi} \right\} = \sum_{n=1}^{\infty} \langle \psi_n \mid H_{nn} \psi_n \rangle \quad , \tag{164}$$

each term in this sum has to be evaluated,

$$\langle \psi_n \mid H_{nn} \psi_n \rangle = \langle \psi_n \mid \left\{ \langle n \mid \int dx \widehat{b}^{\dagger}(x) h(x) \widehat{b}(x) \mid n \rangle \right\} \mid \psi_n \rangle \quad +$$

$$\langle \psi_n \mid \left\{ \langle n \mid \frac{1}{2} \iint dx_1 dx_2 \widehat{b}^{\dagger}(x_1) \widehat{b}^{\dagger}(x_2) g(x_1, x_2) \widehat{b}(x_2) \widehat{b}(x_1) \mid n \rangle \right\} \mid \psi_n \rangle \quad . \tag{165}$$

For $x = x'$ the first term on the rhs of the last equation can be rewritten as

$$\langle \psi_n \mid \left\{ \langle n \mid \int dx \widehat{b}^{\dagger}(x) h(x) \widehat{b}(x) \mid n \rangle \right\} \mid \psi_n \rangle \; =$$

$$= \int dx \left\{ h(x) \underbrace{\langle \psi_n \mid \langle n \mid \widehat{b}^{\dagger}(x) \widehat{b}(x') \mid n \rangle \mid \psi_n \rangle}_{=\Gamma^{(1)}(x;x')} \right\}_{x=x'} \quad , \tag{166}$$

where $\Gamma^{(1)}(x, x')$ is the so-called **first order density matrix** :

$$\Gamma^{(1)}(x; x') = \langle \psi_n \mid \langle n \mid \widehat{b}^{\dagger}(x) \widehat{b}(x') \mid n \rangle \mid \psi_n \rangle =$$

$$= \int d^n x \psi_n^*(x_1, x_2, ..., x_n) \sum_{l=1}^{n} \delta(x - x_l) \psi_n(x_1, x_2, ..., x_l, x', x_{l+1}, ..., x_n) =$$

$$= n \int \psi_n^*(x, x_2, ..., x_n) \psi_n(x', ..., x_n) dx_2 dx_3 ... dx_n \quad . \tag{167}$$

In a similar way the second term in (165) can be rewritten,

$$\langle \psi_n \mid \langle n \mid \iint dx_1 dx_2 \widehat{b}^{\dagger}(x_1) \widehat{b}^{\dagger}(x_2) g(x_1, x_2) \widehat{b}(x_2) \widehat{b}(x_1) \mid n \rangle \mid \psi_n \rangle \; =$$

$$= \iint g(x_1, x_2) \Gamma^{(2)}(x_1, x_2; x_1, x_2) dx_1 dx_2 \quad , \tag{168}$$

where $\Gamma^{(2)}(x_1, x_2; x_1, x_2)$ is the so-called **second order density matrix**,

$$\Gamma^{(2)}(x_1, x_2; x_1, x_2) =$$

$$= \frac{1}{2} \int d^n x' \psi_n^*(x_1', x_2', x_3', ..., x_n') \times$$

$$\times \sum_{k \neq l = 1}^{n} \delta(x_1 - x_k')\delta(x_2 - x_l')\psi_n(x_1', x_2', x_3', ..., x_n') =$$

$$= \binom{n}{2} \int \psi_n^*(x_1, x_2, x_3', ..., x_n')\psi_n(x_1, x_2, x_3', ..., x_n')dx_3'...dx_n' \quad . \qquad (169)$$

15.9 Quantization of matter fields - the relation between first and second quantization

The time-dependent Schrödinger equation for a single particle moving in its own field is given by

$$(h(x) + V(x))\,\psi(x,t) = i\hbar\frac{\partial}{\partial t}\psi(x,t) \quad , \qquad (170)$$

where $x = (\mathbf{r}, \sigma)$ comprises the space $(\mathbf{r})$ and spin (σ) coordinates. The one-particle operator contains the kinetic energy and some potential term $U(\mathbf{r})$,

$$h(x) = -\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{r}) \quad , \qquad (171)$$

and $V(x)$ is the so-called self-interaction contribution,

$$V(x) = e^2 \int \frac{\psi^*(x')\psi(x')}{|\mathbf{r} - \mathbf{r}'|} dx' \quad . \qquad (172)$$

The so-called time evolution operator $\widehat{U} = \widehat{U}(t)$,

$$\widehat{\psi}_S(t) = \widehat{U}\widehat{\psi}_S(0) \;\equiv\; \widehat{U}\widehat{\psi}_H \quad , \qquad (173)$$

which transforms any operator $\widehat{A}_S$ in the **Schrödinger picture** into an operator $\widehat{A}_H(t)$ in the **Heisenberg picture**[5],

$$\widehat{A}_H(t) = \widehat{U}\widehat{A}_S\widehat{U}^\dagger \quad , \qquad (174)$$

has the property that

$$\widehat{U}^\dagger\widehat{U} = \widehat{U}\widehat{U}^\dagger = 1 \quad . \qquad (175)$$

Transforming therefore the operator on the lhs of (170) into the Heisenberg picture one gets

$$\left(h_H(x,t) + \widehat{V}_H(x,t)\right) = \widehat{U}\left(h(x) + \widehat{V}(x)\right)\widehat{U}^\dagger \quad . \qquad (176)$$

---

[5] In the following the Schrödinger and the Heisenberg picture are specified by indices S and H, respectively.

If one **attempts to replace** in (170) $\psi(x,t)$ by the field operator $\widehat{b}_H(x,t)$, where the index H indicates that this is an operator in the Heisenberg picture, this equation now reads as

$$\Big(h_H(x,t)+\widehat{V}_H(x,t)\Big)\widehat{b}_H(x,t)=i\hbar\frac{\partial}{\partial t}\widehat{b}_H(x,t)\quad , \tag{177}$$

which in turn applied to an arbitrary element $\widehat{\psi}_H\in\widehat{H}$ yields

$$\Big(h_H(x,t)+\widehat{V}_H(x,t)\Big)\widehat{b}_H(x,t)\widehat{\psi}_H=i\hbar\frac{\partial}{\partial t}\Big(\widehat{b}_H(x,t)\widehat{\psi}_H\Big)\quad . \tag{178}$$

By multiplying this equation from the left with the evolution operator $\widehat{U}$ one gets

$$\Big(h(x)+\widehat{V}(x)\Big)\widehat{b}(x)\widehat{\psi}_S(t)=i\hbar\left(\widehat{U}\frac{\partial}{\partial t}\widehat{b}_H(x,t)\widehat{U}^\dagger\right)\widehat{\psi}_S(t)\quad , \tag{179}$$

since

$$\widehat{U}\widehat{A}_H\widehat{\psi}_H=\widehat{U}\widehat{U}^\dagger\widehat{A}_S\widehat{U}\widehat{\psi}_H=\widehat{A}_S\widehat{\psi}_S(t)\quad . \tag{180}$$

The rhs of (178) can now be rewritten using again the evolution operator

$$i\hbar\frac{\partial}{\partial t}\Big(\widehat{b}_H(x,t)\widehat{\psi}_H\Big)=i\hbar\frac{\partial}{\partial t}\Big(\widehat{b}_H(x,t)\widehat{U}^\dagger\widehat{U}\widehat{\psi}_H\Big)=$$

$$=i\hbar\frac{\partial}{\partial t}\Big(\widehat{b}_H(x,t)\widehat{U}^\dagger\Big)\widehat{\psi}_S(t)\quad , \tag{181}$$

$$i\hbar\frac{\partial}{\partial t}\Big(\widehat{b}_H(x,t)\widehat{\psi}_H\Big)=i\hbar\frac{\partial}{\partial t}\Big(\widehat{U}^\dagger\widehat{b}(x)\widehat{U}\widehat{\psi}_H\Big)=i\hbar\frac{\partial}{\partial t}\Big(\widehat{U}^\dagger\widehat{b}(x)\widehat{\psi}_S(t)\Big)\quad . \tag{182}$$

from the last two equations one obtains therefore the relation

$$i\hbar\frac{\partial}{\partial t}\Big(\widehat{b}_H(x,t)\widehat{U}^\dagger\Big)\widehat{\psi}_S(t)=i\hbar\frac{\partial}{\partial t}\Big(\widehat{U}^\dagger\widehat{b}(x)\Big)\widehat{\psi}_S(t)\quad , \tag{183}$$

which multiplied from the left with $\widehat{U}$ leads to the following equation

$$i\hbar\widehat{U}\frac{\partial}{\partial t}\widehat{b}_H(x,t)\widehat{U}^\dagger\widehat{\psi}_S(t)=$$

$$=\underbrace{i\hbar\widehat{U}\left(\frac{\partial\widehat{U}^\dagger}{\partial t}\right)}_{=-\widehat{\mathcal{H}}}\widehat{b}(x)\widehat{\psi}_S(t)+i\hbar\frac{\partial}{\partial t}\Big(\widehat{b}(x)\widehat{\psi}_S(t)\Big)\quad . \tag{184}$$

In here $\widehat{\mathcal{H}}$,

$$\widehat{\mathcal{H}}=-i\hbar\widehat{U}\frac{\partial\widehat{U}^\dagger}{\partial t}\quad , \tag{185}$$

is a **Hermitian** operator in the Fock space $\widehat{H}$ such that

$$\widehat{\mathcal{H}}\widehat{\psi}=i\hbar\frac{\partial}{\partial t}\widehat{\psi}\quad ,\quad \widehat{\psi}\in\widehat{H}\quad . \tag{186}$$

This property follows directly from the transformation $\widehat{\psi}_S(t) = \widehat{U}\widehat{\psi}_H$,

$$\left(\frac{\partial}{\partial t}\right)\widehat{\psi}_S(t) = \left(\frac{\partial \widehat{U}}{\partial t}\right)\widehat{\psi}_H = \left(\frac{\partial \widehat{U}}{\partial t}\right)\widehat{U}^\dagger\widehat{\psi}_S(t) = \left(\frac{\partial \widehat{U}}{\partial t}\widehat{U}^\dagger\right)\widehat{\psi}_S(t) \quad . \tag{187}$$

and from the unitary properties of the evolution operator, namely $\widehat{U}\widehat{U}^\dagger = 1$,

$$\frac{\partial}{\partial t}\left(\widehat{U}^\dagger\widehat{U}\right) = \frac{\partial \widehat{U}^\dagger}{\partial t}\widehat{U} + \widehat{U}^\dagger\frac{\partial \widehat{U}}{\partial t} = 0 \quad . \tag{188}$$

It is therefore easy to see that

$$\widehat{\mathcal{H}} = i\hbar\widehat{U}^\dagger\frac{\partial \widehat{U}}{\partial t} = -i\hbar\widehat{U}\frac{\partial \widehat{U}^\dagger}{\partial t} = \widehat{\mathcal{H}}^\dagger \tag{189}$$

If therefore $\widehat{U}$ is a diagonal operator in Fock space, which implies that the number of particles is not changed with respect to the time variable t, then also $\widehat{\mathcal{H}}$ is a diagonal operator in $\widehat{H}$.

Using now (184) , (185) and (189) in equation (179) one finally gets

$$\left(\widehat{h}(x) + \widehat{V}(x) + \widehat{\mathcal{H}}\right)\left(\widehat{b}(x)\widehat{\psi}_S(t)\right) = i\hbar\frac{\partial}{\partial t}\left(\widehat{b}(x)\widehat{\psi}_S(t)\right) \quad . \tag{190}$$

Equation (190) can also be formulated in the following way using only operators and the definition of $\widehat{\mathcal{H}}$ in (186),

$$\left(\widehat{h}(x) + \widehat{V}(x)\right)\widehat{b}(x) = \left(-\widehat{\mathcal{H}}\widehat{b}(x) + i\hbar\frac{\partial}{\partial t}\widehat{b}(x)\right) = \left[\widehat{b}(x), \widehat{\mathcal{H}}\right]_- \quad . \tag{191}$$

By transforming this last equation into the Heisenberg picture one gets

$$\widehat{U}^\dagger\left[\widehat{b}(x), \widehat{\mathcal{H}}\right]_-\widehat{U} = \left[\widehat{b}_H(x,t), \widehat{\mathcal{H}}_H\right]_- =$$

$$= \widehat{U}^\dagger\left(\widehat{h}(x) + \widehat{V}(x)\right)\widehat{U}\widehat{U}^\dagger\widehat{b}(x)\widehat{U} = \left(\widehat{h}_H(x,t) + \widehat{V}_H(x,t)\right)\widehat{b}_H(x,t) \quad . \tag{192}$$

This, however, is nothing but the lhs of (177),

$$\left(\widehat{h}_H(x,t) + \widehat{V}_H(x,t)\right)\widehat{b}_H(x,t) = i\hbar\frac{\partial}{\partial t}\widehat{b}_H(x,t) \quad , \tag{193}$$

i.e., one finally can comprise the **essence of the second quantization** by means of an **equation of motion for the field operator in the Heisenberg picture**

$$i\hbar\frac{\partial}{\partial t}\widehat{b}_H(x,t) = \left[\widehat{b}_H(x,t), \widehat{\mathcal{H}}_H\right]_- \quad . \tag{194}$$

15.9.1 A configurational space formulation of the second quantization

Using a configurational representation of (190),

$$\left(\underline{\widehat{h}}(x)+\underline{\widehat{V}}(x)+\underline{\widehat{\mathcal{H}}}\right)\underline{\widehat{b}}(x)\underline{\widehat{\psi}}_S(t)=i\hbar\frac{\partial}{\partial t}\left(\underline{\widehat{b}}(x)\underline{\widehat{\psi}}_S(t)\right)\quad , \tag{195}$$

$\underline{\widehat{h}}(x)$ and $\underline{\widehat{V}}(x)$ are of the following form in $\widehat{H}^{(c)}$

$$\widehat{V}(x)=\underline{\widehat{n}}^{(c)}\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & V_{11}(x_1,x) & 0 & 0\\ 0 & 0 & V_{22}(x_1,x_2,x) & 0\\ 0 & 0 & 0 & \ddots\end{pmatrix}\quad , \tag{196}$$

$$\widehat{h}(x)=\underline{\widehat{n}}^{(c)}\begin{pmatrix}h_{00}(x) & 0 & 0 & 0\\ 0 & h_{11}(x) & 0 & 0\\ 0 & 0 & h_{22}(x) & 0\\ 0 & 0 & 0 & \ddots\end{pmatrix}\quad . \tag{197}$$

If one takes therefore the projections into the subspaces

$$H_0^{(c)},H_1^{(c)},H_2^{(c)},..,H_n^{(c)},..\subset\widehat{H}^{(c)}\quad ,$$

the various components of (195) are found. For $n=0$ one gets

$$\begin{aligned}\langle 0\mid\widehat{h}(x)+\widehat{V}(x)+\widehat{\mathcal{H}}\mid 0\rangle\langle 0\mid\widehat{b}(x)\mid 1\rangle\langle 1\mid\psi_S(t)&=\\ =\langle 0\mid\widehat{h}(x)+\widehat{V}(x)+\widehat{\mathcal{H}}\mid 0\rangle b_{01}(x)\psi_1(t)&=\\ =\underbrace{(h_{00}(x)+V_{00}(x)+\mathcal{H}_{00})}_{=\,h(x)}\underbrace{b_{01}(\delta_x)\psi_1(x_1,t)}_{=\,\psi_1(x,t)}&=i\hbar\frac{\partial}{\partial t}\psi_1(x,t)\ .\end{aligned} \tag{198}$$

By comparing this equation with the definition for $\widehat{\mathcal{H}}$ in (186), one easily can see that in $H_1^{(c)}$

$$\mathcal{H}(x)=h(x)\quad . \tag{199}$$

For $n=1$ one similarly obtains

$$\begin{aligned}\langle 1\mid\widehat{h}(x)+\widehat{V}(x)+\widehat{\mathcal{H}}\mid 1\rangle\langle 1\mid\widehat{b}(x)\mid 2\rangle\langle 2\mid\psi_S(t)&=\\ =\langle 1\mid\widehat{h}(x)+\widehat{V}(x)+\widehat{\mathcal{H}}\mid 1\rangle b_{12}(x)\psi_2(t)&=\\ =\left(h_{11}(x)+V_{11}(x_1,x)+\mathcal{H}_{11}(x_1)\right)b_{12}(\delta_x)\psi_2(x_1,x_2,t)&=\\ =\left(h(x)+\frac{e^2}{|\mathbf{r}-\mathbf{r}_1|}+\mathcal{H}(x_1)\right)\psi_2(x_1,x,t)&=i\hbar\frac{\partial}{\partial t}\psi_2(x_1,x,t)\quad ,\end{aligned} \tag{200}$$

or

$$\mathcal{H}(x_1,x)=h(x)+\mathcal{H}(x_1)+\frac{e^2}{|\mathbf{r}-\mathbf{r}_1|}\quad . \tag{201}$$

In general, for an arbitrary n one obtains a recursion formula for the operator $\widehat{\mathcal{H}}$,

$$\mathcal{H}(x_1, x_2, ..., x_n, x) = h(x) + \mathcal{H}(x_1, x_2, ..., x_n) + \sum_{k=1}^{n} \frac{e^2}{|\mathbf{r} - \mathbf{r}_k|} \quad . \tag{202}$$

It should be noted that $\widehat{\mathcal{H}}$ therefore is nothing but the **Hamilton operator for an n-particle system** in the ordinary (**first**) **quantization**.

References

[1] V.Fock, *"Konfigurationsraum und zweite Quantisierung"*, Z.f.Physik 75, 622-648 (1932)

[2] P.Jordan, *"Zur Methode der zweiten Quantisierung"*, Z.f.Physik 75, 649-653 (1932)

[3] M.Hehenberger, *"Second Quantization"*, Lecture Notes, Uppsala 1971

[4] J.W.Negele and H.Orland, *"Quantum Many-Particle Systems"*, Addison-Wesley Publishing Company, 1988

[5] L.Jansen and M.Boon, *"Theory of Finite Groups. Applications to Physics"*, North Holland Publishing Company, 1967

[6] F.A.Berezin, *"The Method of Second Quantization"*, NAUKA Publishing House, Moscow, 1986

[7] B.R.Judd, *"Second Quantization and Atomic Spectroscopy"*, John Hopkins Press, Baltimore, 1967

16 Resolvents and Green functions

In this chapter basic properties of resolvents and of Green functions are briefly reviewed. These properties are the formal basis of scattering theory to be discussed in one of the following chapters.

16.1 Resolvents

Consider the eigenvalue equation

$$H \mid \psi\rangle = \epsilon \mid \psi\rangle \quad , \tag{1}$$

where H (Hamiltonian) is a linear Hermitian time-independent (differential) operator. The resolvent of H, $G(z)$, is defined by

$$G(z) = (z - H)^{-1} \quad , \tag{2}$$

$$(z - H)(z - H)^{-1} = 1 \quad , \tag{3}$$

where z is a complex variable. $G(z)$ is an analytic function of z, whose singularities constitute the eigenvalue spectrum of H. If $\Delta(z)$ is the square of the distance from z to the closest eigenvalues of H, then the norm of $G(z)$, $\parallel G(z) \parallel$, is given by

$$\parallel G(z) \parallel = 1/\Delta(z) \quad , \tag{4}$$

which implies that $G(z)$ is bounded in the whole complex plane except the eigenvalues of H.

$G(z)$ and H have the same eigenfunctions, since by subtracting both sides of (1) from $z \mid \psi\rangle$ and multiplying the left-hand side with $G(z)$, one gets

$$\begin{aligned} (z - H) \mid \psi\rangle &= (z - \varepsilon) \mid \psi\rangle \quad , \\ G(z)(z - H) \mid \psi\rangle &= (z - \varepsilon)G(z) \mid \psi\rangle \quad , \\ \mid \psi\rangle &= (z - \varepsilon)G(z) \mid \psi\rangle \quad , \\ G(z) \mid \psi\rangle &= (z - \varepsilon)^{-1} \mid \psi\rangle \quad . \end{aligned} \tag{5}$$

16.2 Resolvents and symmetry

Suppose for simplicity that the eigenvalue spectrum of H is only discrete

$$\epsilon_1 < \epsilon_2 < \epsilon_3 < \cdots < \epsilon_n \quad .$$

If P_i denotes the projection operator onto the eigenspace of ϵ_i,

$$P_i \mid \psi\rangle = \mid \psi_i\rangle \quad , \quad P_i P_j = \delta_{ij} \tag{6}$$

$$\sum_i P_i = 1 \quad , \tag{7}$$

then by a similar manipulation as for eqns (5), one can show that

$$(z-\varepsilon_i)^{-1}P_i \mid \psi\rangle = G(z)P_i \mid \psi\rangle \quad . \tag{8}$$

Quite clearly, summing over all i yields

$$\sum_i G(z)P_i \mid \psi\rangle = G(z) \mid \psi\rangle = \sum_i \frac{1}{(z-\varepsilon_i)}P_i \mid \psi\rangle \quad , \tag{9}$$

which can be written as an operator equation as

$$G(z) = \sum_i \frac{1}{(z-\varepsilon_i)}P_i \quad , \tag{10}$$

or in terms of Dirac's 'bra' and 'ket' notation

$$G(z) = \sum_i \frac{\mid \psi_i\rangle\langle\psi_i \mid}{(z-\varepsilon_i)} \quad . \tag{11}$$

If C_i (see also Figure 16.1) denotes a contour in the complex plane encircling the i-th eigenvalue of H, by means of the Cauchy integral formula the projection operator P_i can be expressed as

$$P_i = \frac{1}{2\pi i}\int_{C_i} G(z)dz \quad . \tag{12}$$

Furthermore, if C is a closed contour in the complex plane circumscribing a subset of eigenvalues ϵ_i, $\epsilon_{i+1}, \cdots$, ϵ_{i+n} with corresponding projection operators $P_i, P_{i+1}, \cdots, P_{i+n}$, one gets

$$HP_C = \frac{1}{2\pi i}\int_{C} G(z)zdz \quad , \tag{13}$$

where P_C is the projection operator onto the direct sum of the corresponding eigenspaces.

$G(z)$ has simple poles precisely at the positions of the discrete eigenvalues of H. If H has (also) a continuous spectrum, side limits of the resolvent have to be formed:

$$z = \epsilon + i\delta \quad , \quad \delta > 0$$

$$G^{\pm}(\epsilon) = \lim_{\delta\to 0}(\epsilon - H \pm i\delta)^{-1} \quad , \tag{14}$$

i.e. the resolvent is expressed as a limit of operators in Hilbert space. The meaning of (14) is as follows: if z tends to ϵ from above the real axis, $Im(z) > 0$, $(z-H)^{-1}$ tends to G^+, whereas if z tends to ϵ from below the real axis, $Im(z) < 0$, $(z-H)^{-1}$ tends to G^-. For each eigenvalue of H in the continuum, $G(z)$ has a branch point. For $G(z)$ the continuum therefore corresponds to a dense line of branch points, i.e. to a branch cut.

16.3 Green functions

Any representation (representative) of a resolvent is called a Green function. Usually such representations are denoted by a small or capital G with the spanning vectors of the representation appearing either as arguments or subscripts. For example $G(\mathbf{r},\mathbf{r}',z)$,

$$G(\mathbf{r},\mathbf{r}',z)=\langle \mathbf{r}\mid (z-H)^{-1}\mid \mathbf{r}'\rangle \quad , \tag{15}$$

is called the configuration space representation (representative) of $(z-H)^{-1}$. It is extremely useful to adopt from the outset the Dirac notation

$$\Phi_i(\mathbf{r})=\langle \mathbf{r}\mid \Phi_i\rangle \quad , \quad \Phi_i^*(\mathbf{r})=\langle \Phi_i\mid \mathbf{r}\rangle \quad , \tag{16}$$

$$\sum_i \mid \Phi_i\rangle\langle\Phi_i\mid \quad , \quad \langle\Phi_i\mid\Phi_j\rangle=\delta_{ij} \quad , \tag{17}$$

$$\int d\mathbf{r}\mid \mathbf{r}\rangle\langle\mathbf{r}\mid = 1 \quad , \tag{18}$$

$$\langle \mathbf{r}\mid\mathbf{r}'\rangle=\delta(\mathbf{r}-\mathbf{r}') \quad . \tag{19}$$

Switching representation is then simply a matter of using the resolution of the identity (7) and (18). If, for example, H has only a discrete eigenvalue spectrum of eigenvalues ϵ_i with corresponding orthonormal eigenstates $\mid \Phi_i\rangle$, then the configuration space representation in (15) can easily be switched into what is called the spectral representation of the resolvent $G(z)$:

$$\langle \mathbf{r}\mid (z-H)^{-1}\mid \mathbf{r}'\rangle=\langle \mathbf{r}\mid \sum_i \mid\Phi_i\rangle\langle\Phi_i\mid (z-H)^{-1}\sum_j\mid\Phi_j\rangle\langle\Phi_j\mid\mathbf{r}'\rangle \quad ,$$

$$(z-H)^{-1}\mid\Phi_j\rangle=\frac{1}{(z-\epsilon_j)}\mid\Phi_j\rangle \quad ,$$

$$\langle \mathbf{r}\mid (z-H)^{-1}\mid \mathbf{r}'\rangle=\sum_{i,j}\langle\mathbf{r}\mid\Phi_i\rangle\frac{\langle\Phi_i\mid\Phi_j\rangle}{(z-\epsilon_j)}\langle\Phi_j\mid\mathbf{r}'\rangle=$$

$$=\sum_i\frac{\langle\mathbf{r}\mid\Phi_i\rangle\langle\Phi_i\mid\mathbf{r}'\rangle}{(z-\epsilon_i)}=\sum_i\frac{\Phi_i(\mathbf{r})\Phi_i^*(\mathbf{r}')}{(z-\epsilon_i)} \quad . \tag{20}$$

Of course, eqn (20) can be obtained directly by taking the configurational representation of (11).

16.4 The configuration space representation of $G_0(z)$

If the Hamiltonian H has (also) a continuous spectrum the problem of finding representations of $G(z)$ is much more involved. Suppose H is the Hamiltonian for a free particle, H_0,

$$H_0\mid\mathbf{p}\rangle=p^2\mid\mathbf{p}\rangle \quad , \tag{21}$$

$$(z - H_0)^{-1} \mid \mathbf{p}\rangle = G_0(z) \mid \mathbf{p}\rangle = \frac{1}{z - p^2} \mid \mathbf{p}\rangle \quad , \tag{22}$$

$$z = \varepsilon + i\delta \quad . \tag{23}$$

$$\int d\mathbf{p} \mid \mathbf{p}\rangle\langle\mathbf{p} \mid = 1 \quad , \quad \langle\mathbf{p} \mid \mathbf{p}'\rangle = \delta(\mathbf{p} - \mathbf{p}') \quad , \tag{24}$$

$$\langle\mathbf{r} \mid \mathbf{p}\rangle = (2\pi)^{-3/2} \exp(i\mathbf{p}\cdot\mathbf{r}) \quad , \tag{25}$$

then from (22) the configuration space representation of $G_0(z)$ can be obtained by again using the resolution of the identity:

$$\langle\mathbf{r} \mid G_0(z) \mid \mathbf{r}'\rangle = \int \langle\mathbf{r} \mid \mathbf{p}\rangle\langle\mathbf{p} \mid G_0(z) \mid \mathbf{p}'\rangle\langle\mathbf{p}' \mid \mathbf{r}'\rangle d\mathbf{p}d\mathbf{p}' =$$

$$= \frac{1}{8\pi^3} \int \exp(i\mathbf{p}\cdot\mathbf{r}) \frac{1}{z - p^2} \exp(-i\mathbf{p}\cdot\mathbf{r}') d\mathbf{p} \quad . \tag{26}$$

By setting $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ and changing z to κ^2, where $Re(\kappa) > 0$, one gets

$$\langle\mathbf{r} \mid G_0(z) \mid \mathbf{r}'\rangle = G_0(\mathbf{R}, z) =$$

$$= \frac{1}{8\pi^3} \int_0^{2\pi} d\Phi \int_0^{\pi} \sin\Theta d\Theta \int_0^{\infty} \frac{\exp(ipR\cos\Theta)}{\kappa^2 - p^2} p^2 dp =$$

$$= \frac{1}{4\pi^2} \int_0^{\infty} \frac{1}{ipR} \frac{\exp(ipR) - \exp(-ipR)}{\kappa^2 - p^2} p^2 dp =$$

$$\frac{1}{2\pi R} \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{p\exp(ipR)}{\kappa^2 - p^2} dp \quad , \tag{27}$$

where Θ and Φ are the angles between $\mathbf{p}$ and $\mathbf{R}$. The value of the last integral is not changed by using a semicircular contour in the upper half of the complex plane. Rewriting the integrand in (27) as $v(p)/w(p)$, where $v(p) = p\exp(ipR)$ and $w(p) = \kappa^2 - p^2$, one can see immediately that $v(p)$ is a regular function, whereas $w(p)$ has a simple zero location at $\kappa^2 = p^2$. The residue is therefore given by the value of $v(p)/w'(p)$ at $p^2 = \kappa^2$. According to the two roots of p^2, the contour integration yields the following situation:

Im(z)	*Pole*	*Residue*
$\delta > 0$	$p = \kappa$	$-\exp(i\kappa R)/2$
$\delta < 0$	$p = -\kappa$	$-\exp(-i\kappa R)/2$

The configurational representation of $G_0(z)$ is therefore given by

$$\langle \mathbf{r} \mid G_0(z) \mid \mathbf{r}' \rangle = G_0(\mathbf{r},\mathbf{r}',z) = -\frac{1}{4\pi R} \left\{ \begin{array}{lcl} \exp(i\kappa R) & ; & \delta > 0 \\ \exp(-i\kappa R) & ; & \delta < 0 \end{array} \right. \quad . \tag{28}$$

Since this representation depends on the sign of the imaginary part of z, its is not well defined. From (28) one can see that

$$G_0(\mathbf{r},\mathbf{r}',z)^* = G_0(\mathbf{r},\mathbf{r}',z^*) \quad . \tag{29}$$

The configurational representations of the side limits (14), however, are well defined:

$$\langle \mathbf{r} \mid G_0^{\pm}(\epsilon) \mid \mathbf{r}' \rangle = G_0^{\pm}(\mathbf{r},\mathbf{r}',\epsilon) = -\frac{1}{4\pi R} \left\{ \begin{array}{ll} \exp(\pm i\epsilon^{1/2} R) & \epsilon > 0 \\ 1 & \epsilon = 0 \\ \exp(-\left|\epsilon\right|^{1/2} R) & \epsilon < 0 \end{array} \right. \quad . \tag{30}$$

It should be noted that, because of (3), similar 'formal' operator equations pertain for $G_0^{\pm}(\epsilon)$

$$(\epsilon - H_0) G_0^{\pm}(\epsilon) = G_0^{\pm}(\epsilon)(\epsilon - H_0) = 1 \quad , \tag{31}$$

which, however, do not imply that $(\epsilon - H_0)$ has an inverse!

16.5 Side limits and Dirac delta functions

In general the Hamiltonian has a discrete spectrum $\{\varepsilon_i\}$ and a continuous spectrum $\{\epsilon_j\}$. Using the following notation for the resolution of the identity

$$\sum_i \mid \Phi_i \rangle \langle \mathbf{\Phi}_i \mid + \int dj \mid \Phi_j \rangle \langle \mathbf{\Phi}_j \mid = \sum_k \mid \Phi_k \rangle \langle \mathbf{\Phi}_k \mid \quad , \tag{32}$$

then from the spectral representation of $G(z)$, $z = \epsilon + i\delta$, one can see immediately in (20) that

$$G(\mathbf{r},\mathbf{r}',z)^* = G(\mathbf{r},\mathbf{r}',z^*) \quad , \tag{33}$$

which in turn implies that, for the side limits $G^{\pm}(\mathbf{r},\mathbf{r}',\epsilon)$,

$$G^{+}(\mathbf{r},\mathbf{r}',\epsilon)^* = G^{-}(\mathbf{r},\mathbf{r}',\epsilon) \quad . \tag{34}$$

Taking therefore the difference of $G^{+}(\mathbf{r},\mathbf{r}',\epsilon)$ and $G^{-}(\mathbf{r},\mathbf{r}',\epsilon)$:

$$\overline{G}(\mathbf{r},\mathbf{r}',\epsilon) = G^{+}(\mathbf{r},\mathbf{r}',\epsilon) - G^{-}(\mathbf{r},\mathbf{r}',\epsilon) = 2iIm[G^{+}(\mathbf{r},\mathbf{r}',\epsilon)] \quad , \tag{35}$$

it can be understood that $\overline{G}(\mathbf{r},\mathbf{r}',\epsilon)$ is the discontinuity of the spectral representation of $G(z)$ across the branch cut of $G(z)$.

Since formally the Dirac delta function $\delta(\varepsilon - \epsilon_0)$ can be expressed by the following limiting property (see, in particular Messiah 1969)

$$\lim_{\delta \to 0^+} \left[\frac{1}{\epsilon \pm i\delta - \epsilon_0} \right] = PP \left[\frac{1}{\epsilon - \epsilon_0} \right] \mp i\pi\delta(\varepsilon - \epsilon_0) \quad , \tag{36}$$

where PP stands for principal part, $\overline{G}(\mathbf{r}, \mathbf{r}', \epsilon)$ or $Im[G^{\pm}(\mathbf{r}, \mathbf{r}', \epsilon)]$ can be written as

$$\overline{G}(\mathbf{r}, \mathbf{r}', \epsilon) = -2\pi i \sum_k \langle \mathbf{r} \mid \Phi_k \rangle \langle \mathbf{\Phi}_k \mid \mathbf{r}' \rangle \delta(\varepsilon - \epsilon_k) \quad ,$$

$$Im[G^{\pm}(\mathbf{r}, \mathbf{r}', \epsilon)] = \mp\pi \sum_k \langle \mathbf{r} \mid \Phi_k \rangle \langle \mathbf{\Phi}_k \mid \mathbf{r}' \rangle \delta(\varepsilon - \epsilon_k) \quad . \tag{37}$$

Integration over $\mathbf{r}$ of the diagonal elements ($\mathbf{r} = \mathbf{r}'$) of a configurational space representation defines the trace (Tr) of this representation. Since, however,

$$\int \langle \mathbf{r} \mid \Phi_k \rangle \langle \mathbf{\Phi}_k \mid \mathbf{r} \rangle d\mathbf{r} = \int \langle \Phi_k \mid \mathbf{r} \rangle \langle \mathbf{r} \mid \mathbf{\Phi}_k \rangle d\mathbf{r} =$$

$$= \langle \Phi_k \mid \int \mid \mathbf{r} \rangle \langle \mathbf{r} \mid d\mathbf{r} \mid \mathbf{\Phi}_k \rangle = \langle \Phi_k \mid \mathbf{\Phi}_k \rangle = 1 \quad ,$$

one gets directly the density of states per unit volume, $n(\varepsilon)$,

$$Tr\left[Im[G^{\pm}(\mathbf{r}, \mathbf{r}', \epsilon)] \right] = \mp\pi \sum_k \delta(\varepsilon - \epsilon_k) \quad ,$$

$$n(\varepsilon) = \sum_k \delta(\varepsilon - \epsilon_k) = \mp\pi^{-1} Tr\left[Im[G^{\pm}(\mathbf{r}, \mathbf{r}', \epsilon)] \right] \quad . \tag{38}$$

16.6 Resolvents and perturbation theory

Suppose the Hamiltonian H is given in terms of an unperturbed Hamiltonian H_0 and a perturbation operator λV, where λ is a suitable (real) parameter,

$$H = H_0 + \lambda V \quad . \tag{39}$$

Let $G(z)$ and $G_0(z)$ be the resolvents of H and H_0, respectively,

$$G(z) = (z - H)^{-1} \quad , \quad G_o(z) = (z - H_o)^{-1} \quad . \tag{40}$$

Then $G(z)$ can be expressed in terms of $G_0(z)$ as

$$G(z) = \frac{1}{z - H} = \frac{1}{z - H_0 - \lambda V} = \frac{z - H_o}{z - H_o} \left[\frac{1}{z - H_0 - \lambda V} \right] =$$

$$= \frac{1}{z - H_o} \left[\frac{z - H_o + \lambda V - \lambda V}{z - H_0 - \lambda V} \right] = G_o(z) \left[1 + \frac{\lambda V}{z - H_0 - \lambda V} \right] =$$

$$= G_o(z) \left[1 + \lambda V G(z) \right] = G_o(z) \left[1 + G(z) \lambda V \right] = \left[1 + G(z) \lambda V \right] G_o(z) \quad . \tag{41}$$

On the right-hand side of (41) $G(z)$ can be substituted by the value of $G(z)$ on the left-hand side:

$$G(z) = G_o(z)\,[1 + \lambda V G_o(z)\,[1 + \lambda V G(z)]] \quad .$$

Repeated application of this procedure yields the following series for $G(z)$:

$$G(z) = \sum_{n=0}^{\infty} \lambda^n G_o(z)\,[V G_o(z)]^n \quad . \tag{42}$$

The convergence radius of this series is defined by the norm of the operator $\lambda V G_o(z)$, i.e. $\|\lambda V G_o(z)\|$. The series is absolutely convergent if

$$\|\lambda V\| < \Delta_0(z) \quad , \tag{43}$$

where $\Delta_0(z)$ is the square of the distance of z to the closest eigenvalue of H_o.

Choosing in particular λ to be unity, (41) can be reduced to

$$G(z) = G_o(z)\,[1 + V G(z)] \quad , \tag{44}$$

$$G(z) = [1 + G(z) V]\, G_o(z) \quad . \tag{45}$$

Formally, these two equations can also be obtained by using the operator identity

$$A^{-1} = B^{-1} + B^{-1}(A - B)A^{-1}$$

with $A = z - H$ and $B = z - H_0$ for (44) and A and B interchanged for (45). Equation (44) (or for that matter (45)) is called the **Lippmann-Schwinger equation for the resolvent** $G(z)$.

Defining $\Omega(z)$ as

$$\Omega(z) = [1 + V G(z)] \quad , \tag{46}$$

the Lippmann-Schwinger equation can also be written as

$$\boxed{G(z) = G_o(z)\Omega(z) = \Omega(z) G_o(z)} \tag{47}$$

On the left-hand side of (47) the definition of $G(z)$ can again be used to give

$$\Omega(z) G_o(z) = G_o(z)\,[1 + V G(z)] = G_o(z) + G_o(z) V G(z) =$$

$$= G_o(z) + G_o(z) V \Omega(z) G_o(z) \quad ,$$

$$\Omega(z) = 1 + G_o(z) V \Omega(z) \quad , \tag{48}$$

which is yet another way to formulate a Lippmann-Schwinger equation.

16.7 The T-operator

For many purposes it is convenient to define the following operator $T(z)$ as

$$T(z) = V + VG(z)V \quad . \tag{49}$$

$T(z)$ is an analytic function of z and has the same properties as $G(z)$, namely a pole for each discrete eigenvalue of H, $H = H_0 + V$, and a branch cut for the continuum. Multiplying from the left with $G_o(z)$ one gets

$$G_o(z)T(z) = \left[G_o(z) + G_o(z)VG(z)\right]V = G_o(z)\left[1 + VG(z)\right]V = \quad ,$$

$$G_o(z)T(z) = VG(z) = G(z)V \quad , \tag{50}$$

which can be written in terms of (44) as

$$G_o(z)T(z) = VG_o(z)\left[1 + VG(z)\right] \quad ,$$

$$T(z) = V + VG_0(z)T(z) \quad . \tag{51}$$

Equation (51) is usually called the Lippmann-Schwinger equation for the T operator. because of the definition in (49), side limits can be taken as in the case of $G(z)$:

$$\lim_{\delta \to 0^+} T(\epsilon \pm i\delta) = T^{\pm}(\epsilon) = V + VG^{\pm}(\epsilon)V \quad . \tag{52}$$

Similar to $G(z)$ the adjoint operator of $T(z)$, $T(z)^{\dagger}$, is given by

$$T(z)^{\dagger} = T(z^*) \quad , \tag{53}$$

and $T(z)$ can also be expressed in terms of the operator $\Omega(z)$ as

$$\boxed{T(z) = V\left[1 + VG(z)\right] = V\Omega(z) = \Omega(z)V} \tag{54}$$

16.8 The Lippmann-Schwinger equation

The usefulness of the T operator can be seen by considering the following homogeneous and inhomogeneous problems

$$(\epsilon - H_0) \mid \Phi(\epsilon)\rangle = 0 \quad , \tag{55}$$

$$(\epsilon - H_0) \mid \Psi(\epsilon)\rangle = V \mid \Psi(\epsilon)\rangle \quad , \tag{56}$$

where $\mid \Phi(\epsilon)\rangle$ is an eigenstate of H_0 to the same eigenvalue ϵ as (56). Because of eqn (31), the general solution of eqn (56) is given by

$$\mid \Psi^{\pm}(\epsilon)\rangle = \mid \Phi(\epsilon)\rangle + G_0^{\pm}(\epsilon)V \mid \Psi^{\pm}(\epsilon)\rangle \quad . \tag{57}$$

Using eqn (57) iteratively,

$$\mid \Psi^{\pm}(\epsilon)\rangle = \mid \Phi(\epsilon)\rangle + G_0^{\pm}(\epsilon)V\left[\mid \Phi(\epsilon)\rangle + G_0^{\pm}(\epsilon)V \mid \Psi^{\pm}(\epsilon)\rangle\right] \quad , \textit{etc.}$$

one gets

$$| \Psi^{\pm}(\epsilon)\rangle = | \Phi(\epsilon)\rangle + G_0^{\pm}(\epsilon)T^{\pm}(\epsilon) | \Phi(\epsilon)\rangle = \left[1 + G_0^{\pm}(\epsilon)T^{\pm}(\epsilon)\right] | \Phi(\epsilon)\rangle \quad , \tag{58}$$

or, using eqn (50)

$$| \Psi^{\pm}(\epsilon)\rangle = \left[1 + G^{\pm}(\epsilon)V\right] | \Phi(\epsilon)\rangle \quad , \tag{59}$$

or, using eqn (46)

$$\boxed{| \Psi^{\pm}(\epsilon)\rangle = \Omega^{\pm}(\epsilon) | \Phi(\epsilon)\rangle} \tag{60}$$

Equations (58) - (60) are usually called the **Lippmann-Schwinger equation**. In each of these three equations the solution of $H = H_0 + V$, $| \Psi^{\pm}(\epsilon)\rangle$, is given in terms of the (known) unperturbed eigenstate $| \Phi(\epsilon)\rangle$.

16.9 Green functions and perturbation theory

In principle all that is necessary now is to form representations of the various Lippmann-Schwinger equations discussed previously by making use of the 'resolution of the identity'.

The configuration space representation of the Lippmann-Schwinger equation for the resolvent $G(z)$,

$$G(z) = G_0(z)\left[1 + VG(z)\right] \quad ,$$

is given by

$$\begin{aligned} &\langle \mathbf{r} | G(z) | \mathbf{r}'\rangle = \langle \mathbf{r} | G_0(z) | \mathbf{r}'\rangle + \\ &+ \textstyle\iint \langle \mathbf{r} | G_0(z) | \mathbf{r}_1\rangle\langle \mathbf{r}_1 | V | \mathbf{r}_2\rangle\langle \mathbf{r}_2 | G(z) | \mathbf{r}'\rangle d\mathbf{r}_1 d\mathbf{r}_2 \end{aligned} \quad . \tag{61}$$

However, since in reality V will be a potential given within the framework of the local density functional, the representation of V in configuration space can be assumed to be diagonal

$$\langle \mathbf{r} | V | \mathbf{r}'\rangle = \delta(\mathbf{r} - \mathbf{r}')\langle \mathbf{r} | V | \mathbf{r}\rangle = \delta(\mathbf{r} - \mathbf{r}')V(\mathbf{r}) \quad .$$

This reduces (61) to

$$\begin{aligned} &\langle \mathbf{r} | G(z) | \mathbf{r}'\rangle = \langle \mathbf{r} | G_0(z) | \mathbf{r}'\rangle + \\ &+ \textstyle\int \langle \mathbf{r} | G_0(z) | \mathbf{r}"\rangle\langle \mathbf{r}" | V | \mathbf{r}"\rangle\langle \mathbf{r}" | G(z) | \mathbf{r}'\rangle d\mathbf{r}" \end{aligned} \tag{62}$$

or

$$G(\mathbf{r}, \mathbf{r}', z) = G_0(\mathbf{r}, \mathbf{r}', z) + \int G_0(\mathbf{r}, \mathbf{r}", z)V(\mathbf{r}")G(\mathbf{r}", \mathbf{r}', z)d\mathbf{r}" \tag{63}$$

In particular from the more 'traditional' form (eqn(63)) one can see that the Green function $G(\mathbf{r}", \mathbf{r}', z)$ satisfies an inhomogeneous integral equation, where the integral kernel is given by $G_0(\mathbf{r}, \mathbf{r}", z)V(\mathbf{r}")$.

In momentum space the corresponding representation of eqn (44) can be similarly found:

$$\begin{aligned}&\langle \mathbf{p} \mid G(z) \mid \mathbf{p}'\rangle = \langle \mathbf{p} \mid G_0(z) \mid \mathbf{p}'\rangle + \\ &+ \iint \langle \mathbf{p} \mid G_0(z) \mid \mathbf{p}_1\rangle\langle \mathbf{p}_1 \mid V \mid \mathbf{p}_2\rangle\langle \mathbf{p}_2 \mid G(z) \mid \mathbf{p}'\rangle d\mathbf{p}_1 d\mathbf{p}_2\end{aligned} \quad . \tag{64}$$

Using the relation

$$\langle \mathbf{r} \mid \mathbf{p}\rangle = \frac{1}{(2\pi)^{d/2}} \exp(i\mathbf{p}\cdot\mathbf{r}) \quad , \tag{65}$$

where $d = 1, 2, 3$ is the dimensionality of the problem, one easily can show that eqn (64) is the Fourier transform of (61).

Turning now to eqn (58), which in configuration space representation is given by

$$\begin{aligned}\langle \mathbf{r} \mid \Psi^{\pm}(\epsilon)\rangle \quad &= \langle \mathbf{r} \mid \Phi(\epsilon)\rangle + \langle \mathbf{r} \mid G^{\pm}(\epsilon)V \mid \Phi(\epsilon)\rangle \\ &= \langle \mathbf{r} \mid \Phi(\epsilon)\rangle + \langle \mathbf{r} \mid G_0^{\pm}(\epsilon)T^{\pm}(\epsilon) \mid \Phi(\epsilon)\rangle\end{aligned} \quad , \tag{66}$$

one can see immediately that by using $\mid \mathbf{p}\rangle$ as eigenstates of H_0 (eqn (21)) one gets the equation

$$H_0 \mid \mathbf{p}\rangle = p^2 \mid \mathbf{p}\rangle \quad , \quad p = \sqrt{\epsilon} \quad ,$$

$$\langle \mathbf{r} \mid \Psi^{\pm}(p)\rangle = \langle \mathbf{r} \mid \mathbf{p}\rangle + \langle \mathbf{r} \mid G_0^{\pm}(p)T^{\pm}(p) \mid \mathbf{p}\rangle \quad . \tag{67}$$

16.10 Green functions and scattering theory

Suppose that the representation of V in eqn (67) is diagonal in configuration space and that $V(r)$ decays rapidly as $\mathbf{r}$ approaches infinity. Equation (67) can now be interpreted in terms of a scattering formalism: an incident particle of energy $\epsilon = p^2$ in state $\mid \mathbf{p}\rangle$ is perturbed by $V(r)$ resulting in a modified state $\mid \Psi^{\pm}(p)\rangle$, which will be analyzed as r approaches infinity. Using explicitly the representation of $\mid \mathbf{p}\rangle$ in configuration space (eqn (65)) and the resolution of the identity, eqn (67) is simply given by

$$\begin{aligned}\langle \mathbf{r} \mid \Psi^{\pm}(p)\rangle \quad &= (2\pi)^{-3/2}\left\{\exp(i\mathbf{p}\cdot\mathbf{r}) + \right. \\ &\left. \iint \langle \mathbf{r} \mid G_0^{\pm}(p) \mid \mathbf{r}'\rangle\langle \mathbf{r}' \mid T^{\pm}(p) \mid \mathbf{r}"\rangle \exp(i\mathbf{p}\cdot\mathbf{r}")d\mathbf{r}'d\mathbf{r}"\right\}\end{aligned} \quad . \tag{68}$$

Since

$$|\mathbf{r} - \mathbf{r}'| \underset{\mathbf{r}>>\mathbf{r}'}{\longrightarrow} \mathbf{r} - (\mathbf{r}\cdot\mathbf{r}')/r + O(r^{-2}) \quad , \tag{69}$$

where $O(x)$ stands for 'orders of x' the free-particle Green function $\langle \mathbf{r} \mid G_0^{\pm}(p) \mid \mathbf{r}'\rangle$ (eqn(30)) can be written as

$$\langle \mathbf{r} \mid G_0^{\pm}(p) \mid \mathbf{r}'\rangle \underset{\mathbf{r}>>\mathbf{r}'}{\longrightarrow} -\frac{\exp(\pm ipr)}{4\pi r} \exp\left(\mp ip(\mathbf{r}\cdot\mathbf{r}')/r\right)\left[1 + O(r^{-1})\right] \quad . \tag{70}$$

As r goes to infinity, $\langle \mathbf{r} \mid \Psi^{\pm}(p)\rangle$ has the asymptotic form

$$\begin{aligned}\langle \mathbf{r} \mid \Psi^{\pm}(p)\rangle \quad &\underset{r\to\infty}{\longrightarrow} (2\pi)^{-3/2}\,\{\exp(i\mathbf{p}\cdot\mathbf{r}) - (\exp(\pm ipr)/4\pi r) \times \\ &\times \textstyle\iint \exp\left(\mp ip(\mathbf{r}\cdot\mathbf{r}')/r\right) \langle \mathbf{r}' \mid T^{\pm}(p) \mid \mathbf{r}"\rangle \exp(i\mathbf{p}\cdot\mathbf{r}")d\mathbf{r}'d\mathbf{r}"\big\}\end{aligned} \tag{71}$$

Because of the factor $\exp(\pm ipr)/r$ only $\langle \mathbf{r} \mid \Psi^{+}(p)\rangle$ is physically acceptable, i.e. only the side limits $G_0^{+}(p)$ and $T^{+}(p)$ have to be considered. Inspecting now the argument of the first exponential in the integrand, $ip(\mathbf{r}\cdot\mathbf{r}')/r$, one can see that $p\mathbf{r}/r$ is a vector of length p but with direction $\mathbf{r}/r$. Denoting this vector by $\mathbf{p}_f$, the double integral in eqn (71) is nothing more than a change of representations

$$\iint \exp\left(-i\mathbf{p}_f\cdot\mathbf{r}\right) \langle \mathbf{r} \mid T^{\pm}(p) \mid \mathbf{r}'\rangle \exp(i\mathbf{p}\cdot\mathbf{r}')d\mathbf{r}d\mathbf{r}' = 8\pi^3\langle \mathbf{p}_f \mid T^{\pm}(p) \mid \mathbf{p}\rangle \quad . \tag{72}$$

Equation (71) can therefore be written in th form

$$\begin{aligned}\langle \mathbf{r} \mid \Psi^{\pm}(p)\rangle \quad &\underset{r\to\infty}{\longrightarrow} (2\pi)^{-3/2} \\ &\times \left\{\exp(i\mathbf{p}\cdot\mathbf{r}) + (\exp(ipr)/r)\left[-2\pi^2\langle \mathbf{p}_f \mid T^{\pm}(p) \mid \mathbf{p}\rangle\right]\right\}\end{aligned} \quad , \tag{73}$$

i.e. $\langle \mathbf{r} \mid \Psi^{\pm}(p)\rangle$ can be identified as the familiar scattering wavefunction in elementary collision theory, usually written as

$$\langle \mathbf{r} \mid \Psi^{\pm}(p)\rangle \underset{r\to\infty}{\longrightarrow} (2\pi)^{-3/2}\left[\exp(i\mathbf{p}\cdot\mathbf{r}) + f(\mathbf{p}_f,\mathbf{p})\frac{\exp(ipr)}{r}\right] \quad . \tag{74}$$

The quantity $f(\mathbf{p}_f,\mathbf{p})$ is the so-called scattering amplitude. The scattering amplitude is proportional to a momentum space representation of $T^{\pm}(p)$, $\langle \mathbf{p} \mid T^{\pm}(p) \mid \mathbf{p}'\rangle$, in which $\mathbf{p}$ and $\mathbf{p}'$ have the same length: $p^2 = (p')^2 = \epsilon$. For this very reason this particular representation of $T^{\pm}(p)$ is called the '**on the energy shell T-matrix**'.

References

[1] F.Constantinescu and E.Magyari, *Problems in Quantum Mechanis*, Pergamon Press, Oxford, 1971

[2] E.N.Ecomomou, *Green's functions in Quantum Physics*, Springer-Verlag, 1979

[3] A.Messiah, *Quantum Mechanics*, North Holland, Amsterdam, 1969

[4] R.G.Newton, *Scattering Theory of Waves and Particles*, McGraw-Hill, New York, 1966

[5] G.Rickayzen, *Green's functions and Condensed Matter*, Academic Press, London, 1980

[6] L.S.Rodberg and R.M.Thaler, *Introduction to the Quantum Theory of Scattering*, Academic Press, new York, 1967

[7] J.R.Taylor, *Scattering Theory: the Quantum Theory of non-relativistic Collisions*, Wiley, New York, 1972

[8] P.Weinberger, *Electron Scattering Theory for Ordered and Disordered Matter*, Oxford Science Publications, Oxford, 1990

[9] J.M.Ziman, *Elements of Advanced Quantum Theory*, Cambridge University Press, 1969

17 Relativistic quantum mechanics

17.1 Minkowski-space

Suppose the set of space-time vectors is given by

$$M = \{x^\mu\} \quad , \tag{1}$$

$$x^\mu \equiv (x^0, x^1, x^2, x^3) = (x^0, x^k) = (ct, \mathbf{r}) \quad , \tag{2}$$

$$\mu = 0, 1, 2, 3 \quad , \quad k = 1, 2, 3 \quad ,$$

where $x^0 = ct$ is the time component and $\mathbf{r} = (x^1, x^2, x^3)$ the space component of an arbitrary space-time vector x^μ. For any arbitrary pair of elements $x, y \in M$ the scalar product in M is defined as follows

$$(x, y) \equiv \sum_{\mu=0}^{3} x_\mu y^\mu = x_0 y^0 - \sum_{k=1}^{3} x_k y^k \quad , \tag{3}$$

and in particular therefore the norm as

$$\| \, x \, \| = (x, x) = x_0 x^0 - (\mathbf{r}, \mathbf{r}) \quad . \tag{4}$$

The metric in M is said to be **pseudo-euclidean**, since the **metric tensor** $g_{\mu\nu}$ is of the following form

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -\underline{1}_3 \end{pmatrix} = g^{\mu\nu} \quad . \tag{5}$$

The set M is sometimes also called **Minkowski space**.

In M a vector a is called **contravariant** (usually denoted by e.g. a^μ) if it "transforms like a space-time vector" x^μ and **covariant** (usually denoted by e.g. a_μ) if transforms like $\partial/\partial x^\mu$. The transformation of a contravariant vector by means of the metric tensor $g_{\mu\nu}$ yields a covariant vector:

$$a_\mu = \sum_{v=0}^{3} g_{\mu\nu} a^\nu \equiv g_{\mu\nu} a^\nu \quad , \tag{6}$$

while by the opposite procedure a contravariant vector is obtained:

$$a^\mu = \sum_{v=0}^{3} g^{\mu\nu} a_\nu \equiv g_{\mu\nu} a^\nu \quad . \tag{7}$$

It should be noted that in either case $a_0 = a^0$. The implicit summation over repeated indices as indicated in the last two equations is usually called the

Einstein sum convention. Quite clearly this implies e.g. for the product of the metric tensor with itself that

$$\sum_{\rho=0}^{3} g_{\mu\rho}g^{\rho\nu} \equiv g_{\mu\rho}g^{\rho\nu} = \delta_{\mu}{}^{\nu} \quad , \quad \delta_{\mu}{}^{\nu} = \{ \begin{array}{ll} 1 \, , & \nu = \mu \\ 0 \, , & \nu \neq \mu \end{array} \quad , \tag{8}$$

a unit matrix is the result.

A vector a^{μ} is called a **space-like vector** if its norm $a_{\mu}a^{\mu} < 0$ and oppositely a **time-like vector** if the norm is positive.

Defined in M the **gradient** can be written as a covariant vector ∂_{μ},

$$\partial_{\mu} \equiv \frac{\partial}{\partial x^{\mu}} = (\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}) = (\frac{\partial}{\partial ct}, \boldsymbol{\nabla}) \tag{9}$$

or as contravariant vector ∂^{μ}

$$\partial^{\mu} = (\frac{\partial}{\partial ct}, -\boldsymbol{\nabla}) \quad , \tag{10}$$

whereby

$$\partial_{\mu}\partial^{\mu} \equiv \Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta \tag{11}$$

usually is called **D'Alembert operator**.

If $\mathbf{A} = \mathbf{A}(\mathbf{r}, t)$ denotes the vector potential and $\phi = \phi(\mathbf{r}, t)$ the scalar potential then the electromagnetic field can be written as the following contravariant vector A^{μ}

$$A^{\mu} = (\phi, \mathbf{A}) \tag{12}$$

such that the electric and magnetic field, **E** and **H**, respectively, are given by

$$\mathbf{E} = (E_x, E_y, E_z) = -\boldsymbol{\nabla}\phi - \frac{\partial A}{\partial x^0} \quad , \tag{13}$$

$$\mathbf{H} = (H_x, H_y, H_z) = rot\mathbf{A} \quad . \tag{14}$$

The so-called electro-magnetic field tensor $F_{\mu\nu}$, formally written as

$$F_{\mu\nu} = \frac{\partial A_{\nu}}{\partial x^{\mu}} - \frac{\partial A_{\mu}}{\partial x^{\nu}} \quad , \tag{15}$$

is an antisymmetric tensor in M, whose elements are given by the components of **E** and **H**,

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -H_z & H_y \\ -E_y & H_z & 0 & -H_x \\ -E_z & -H_y & H_x & 0 \end{pmatrix} \quad . \tag{16}$$

The gradient and the electromagnetic field vectors finally can be combined to yield the following four component vector D_{μ}

$$D_{\mu} = \delta_{\mu} + ieA_{\mu} = (\frac{\partial}{\partial x^0} + ie\phi, \boldsymbol{\nabla} - ie\mathbf{A}) \tag{17}$$

17.2 Poincaré and Lorentz transformations

Poincaré transformations are inhomogeneous linear transformations that preserve the quadratic form $x_\mu x^\mu$, i.e. the norm in M. Such a transformation is defined by obtained

$$(x^\mu)' = \Omega^\mu{}_\upsilon x^\upsilon + a^\mu \tag{18}$$

where $(x^\mu)'$ is the transformed vector, $\Omega^\mu{}_\upsilon$ a space-time point operation Ω, which keeps the origin invariant, and a^μ a translation. If denotes $(\Omega \mid a)$ the operator that maps x^μ on $(x^\mu)'$,

$$(\Omega \mid a)x^\mu = \Omega^\mu{}_\upsilon x^\upsilon + a^\mu = (x^\mu)' \quad , \tag{19}$$

then the matrix $\Omega^\mu{}_\upsilon$ is the representations of the corresponding space-time point operation, whereby matrices like $\Omega_{\mu\upsilon}$ and $\Omega^{\mu\nu}$ can be obtained by using the metric tensor $g_{\mu\nu}$ such as in the following transformation

$$\Omega^{\mu\nu} = g_\nu{}^\rho \Omega^\mu{}_\rho \quad .$$

From the condition that the norm is left invariant and that the transformations are real, the properties of the matrices $\Omega_\mu{}^\nu$ can be deduced, namely

$$\Omega^*_{\mu\upsilon} = \Omega_{\mu\upsilon} \quad , \tag{20}$$

$$\Omega_{\mu\upsilon}\Omega^{\mu\lambda} = \Omega_{\nu\mu}\Omega^{\lambda\mu} = \delta_\nu{}^\lambda \quad , \tag{21}$$

$$\det \mid \Omega^\mu_\upsilon \mid = \pm 1 \quad . \tag{22}$$

The set of operators (Ω/a) forms a group, the so-called **Poincaré group**,

$$P = \{(\Omega \mid a) \ / \ (\Omega \mid a)(\Omega' \mid a') = (\Omega\Omega' \mid \Omega a' + a)\} \quad , \tag{23}$$

in which the identity element $(\epsilon \mid 0)$ has the following representation for the pure space-time point operation ϵ

$$D(\epsilon) = \begin{pmatrix} 1 & 0 \\ 0 & \underline{1}_3 \end{pmatrix} \quad . \tag{24}$$

Similarly the pure time-inversion operator $(T \mid 0)$ and pure space inversion operator $(J \mid 0)$ are defined by the representations of their corresponding space-time point operations

$$D(T) = \begin{pmatrix} -1 & 0 \\ 0 & \underline{1}_3 \end{pmatrix} \quad , \quad D(J) = \begin{pmatrix} 1 & 0 \\ 0 & -\underline{1}_3 \end{pmatrix} \quad . \tag{25}$$

The set of operator $(\Omega \mid a)$ for which $\Omega^{00} > 0$, i.e., which preserve the direction of time, form a subgroup $\overline{P} \subset P$ of index two:

$$\overline{P} = \{(\overline{\Omega} \mid a) \ / \ \overline{\Omega}^{00} \geq 0\} \quad , \tag{26}$$

since the complement $(P - \overline{P})$ is defined by

$$(P - \overline{P}) \ = \{(\Omega \mid a) \ / \ (\Omega \mid a) = (\overline{\Omega} \mid a)(T \mid 0)\} \quad . \tag{27}$$

$\overline{P}$ is called **orthochronous Poincaré group**, which in turn has a subgroup of index two, namely the so-called **proper orthochronous Poincaré group** $\overline{P}_{+}$, which is the set of time conserving transformations for which det $\mid \Omega^{\mu}{}_{v} \mid = 1$:

$$\overline{P}_{+} = \{(\overline{\Omega} \mid a) \ / \ \overline{\Omega}^{00} \geq 0, \det \mid \Omega^{\mu}{}_{v} \mid = 1\} \quad . \tag{28}$$

In terms of left cosets, the Poincaré group $P \supset \overline{P} \supset \overline{P}_{+}$ can therefore be written as

$$P = \{\overline{P}, (T \mid 0)\overline{P}\} \quad , \tag{29}$$

$$\overline{P} = \{\overline{P}_{+}, (J \mid 0)\overline{P}_{+}\} \quad . \tag{30}$$

These three Poincaré groups contains as corresponding subgroups all those operations for which the translation $a = 0$:

$$L = \{(\Omega \mid a)\} = \{\overline{L}, (T \mid 0)\overline{L}\} \tag{31}$$

$$\overline{L} = \{\overline{L}_{+}, (J \mid 0)\overline{L}_{+}\} \quad . \tag{32}$$

L is called **Lorentz group**, $\overline{L}$ **orthochronous Lorentz group** and $\overline{L}_{+}$ **proper orthochronous Lorentz group**.

It should be noted that the subset of operators of the Poincaré group that corresponds to pure space translations only also forms a subgroup, the so-called Euclidean group:

$$P \supset E = \{(\Omega \mid a) \ / \ \forall a^{0} = 0\} \quad . \tag{33}$$

The corresponding subgroup of the Lorentz group is the familiar Rotation-Inversion group in the three-dimensional vector space R_3.

17.3 Postulates of relativistic quantum mechanics

Missing

17.4 Dirac equation

For a single particle of charge e and mass m the relativistic Hamilton function is given by[1]

$$H = e\phi + \sqrt{(\mathbf{p} - e\mathbf{A})^2 + m^2} \quad , \tag{34}$$

where ϕ is the scalar potential, $\mathbf{A}$ the vector potential and $\mathbf{p}$ the momentum. Assuming now as discussed in the previous section that the probability density $\rho = \psi^*\psi$ is positive definite then it follows immediately that the corresponding Hamilton operator $\widehat{H}$ has to be Hermitian, since:

$$\int \frac{\partial \rho}{\partial t} d^3x = -\frac{i}{\hbar} \int \left(\psi^* H \psi - (H\psi)^* \psi \right) d^3x = 0 \quad . \tag{35}$$

For the sake of simplicity in the following discussion only the Hamilton function in the absence of a field shall be considered:

$$H = \sqrt{\mathbf{p}^2 + m^2} \quad . \tag{36}$$

Since the lhs of (35) is linear in $\partial/\partial t$ ($\equiv \partial/\partial x^0$ in (9)) this implies that also $\widehat{H}$ on the rhs of (35) has to be linear with respect to $\partial/\partial x^k, k = 1, 3$, i.e., with respect to components of the momentum operator $\mathbf{p}$. This condition is usually called the **condition of relativistic covariance**. If one replaces according to the correspondence principle $E \rightarrow i\partial/\partial t$ and $\mathbf{p} \rightarrow -i\boldsymbol{\nabla}$, one immediately can see that the condition of linearity cannot be fulfilled in a straightforward manner, since the square root is not a linear operator. The Dirac problem, but also the problem of Pauli's spin theory, however, can be associated with a special polynomial algebra.

17.4.1 Polynomial algebras

Let $P_2(x)$ be a second order polynomial of the following form

$$P_2(x) = a_{21} \sum_{i \neq j} x_i x_j + a_{22} \sum_j x_j^2 \quad , \quad i, j = 1, 2, \ldots, m \quad , \tag{37}$$

where the a_{ij} are elements of a symmetric matrix. Consider further that the linear form

$$L(x) = \sum_{j=1}^{m} \alpha_j x_j \tag{38}$$

satisfies the condition

$$P_2(x) + L^2(x) = 0 \quad , \tag{39}$$

then the set of coefficients $\{\alpha_j\}$ has to satisfy the following properties[2]:

$$i = j : \quad [\alpha_i, \alpha_j]_+ = -2a_{22} I \quad , \tag{40}$$

---

[1] for a discussion of classical relativistic dynamics see e.g. the book by Messiah

[2] see in particular the paper by Raghavacharyulu and Menon, 1970

$$i \neq j : \quad [\alpha_i, \alpha_j]_+ = -a_{21} I \quad , \tag{41}$$

where I denotes the identity element in $\{\alpha_j\}$ and $[,]_+$ anticommutators. The set of coefficients $\{\alpha_j\}$ is called an associative algebra. Two special cases carry famous names, namely

$$a_{21} = a_{22} = 0 \quad \rightarrow \quad [\alpha_i, \alpha_j]_+ = 0 \quad , \tag{42}$$

the so-called **Grassmann algebra**[3] and

$$a_{21} = 0 \; , \quad a_{22} = -1 \quad \rightarrow \quad [\alpha_i, \alpha_j]_+ = 2\delta_{ij} \quad , \tag{43}$$

the so-called **Clifford algebra**. Comparing now with (36) one can see that exactly the case of the Clifford algebra is needed in tackling the problem of the linearization of the square root:

$$\underbrace{\sqrt{(\sum_{j=1}^{m} p_j^2)}}_{P_2(p)} = \underbrace{\sum_{j=1}^{m} \alpha_j p_j}_{L(p)} \quad . \tag{44}$$

In the following first the case for $m = 2$ and 3 (Pauli spin theory) is discussed by considering the smallest groups with Clifford algebraic structure and only then in a similar way the Dirac problem ($m = 4$) is addressed.

17.4.2 The Pauli groups

For $m = 2$ the smallest set of elements α_i that shows group closure[4] is given by

$$G_P^{(m=2)} = \{\pm I, \pm\alpha_1, \pm\alpha_2, \pm\alpha_1\alpha_2\} \quad . \tag{45}$$

This group is of order 8 and has 5 classes (C_i), namely $C_1 = \{I\}$, $C_2 = \{-I\}$, $C_3 = \{\pm\alpha_1\}$, $C_4 = \{\pm\alpha_2\}$, $C_5 = \{\pm\alpha_1\alpha_2\}$. There are therefore 5 irreducible representations ($\Gamma_i^{(m=2)}, i = 1, 5$) of dimensions n_i such that

$$\sum_{i=1}^{5} n_i^2 = 8 \quad . \tag{46}$$

This implies that 4 irreducible representations ($\Gamma_i^{(m=2)}, i = 1, \ldots, 4$) have to be one-dimensional and one two-dimensional. Since one-dimensional representations are commutative, i.e., do not satisfy the conditions of a Clifford algebra, only the two-dimensional representation ($\Gamma_5^{(m=2)}$) is of help. The matrices for this irreducible representation are listed below:

$$\Gamma_5^{(m=2)}(\pm I) = \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \Gamma_5^{(m=2)}(\pm\alpha_1) = \pm \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad ,$$

---

[3] this is exactly the algebra of creation and annihilation operators for fermions, see also chapter 15

[4] for the discussion in the following sections see also chapter 18

$$\Gamma_5^{(m=2)}(\pm\alpha_2) = \pm \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \Gamma_5^{(m=2)}(\pm\alpha_1\alpha_2) = \pm \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad . \tag{47}$$

Using this set of matrices it is easy to show that it indeed forms a representation of $G_P^{(m=2)}$ and that these matrices are Clifford algebraic. For the case of $m = 2$ the problem of the linearization of the square root is therefore solved:

$$\sqrt{p_1^2 + p_2^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = p_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + p_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad . \tag{48}$$

For $m = 3$ the smallest set of elements σ_i forming a group is given by

$$G_P^{(m=3)} = \{\pm I, \pm\alpha_1, \pm\alpha_2, \pm\alpha_3, \pm\alpha_1\alpha_2, \pm\alpha_1\alpha_3, \pm\alpha_2\alpha_3, \pm\alpha_1\alpha_2\alpha_3\} \quad . \tag{49}$$

The order of this group is 16. It has 10 classes, namely

$$\begin{gathered} C_1 = \{I\}, \quad C_2 = \{-I\}, \\ C_3 = \{\pm\alpha_1\}, \quad C_4 = \{\pm\alpha_2\}, \quad C_5 = \{\pm\alpha_3\} \\ C_6 = \{\pm\alpha_1\alpha_2\}, \quad C_7 = \{\pm\alpha_1\alpha_3\}, \quad C_8 = \{\pm\alpha_2\alpha_3\} \\ C_9 = \{\alpha_1\alpha_2\alpha_3\}, \quad C_{10} = \{-\alpha_1\alpha_2\alpha_3\}, \end{gathered} \tag{50}$$

and therefore 10 irreducible representations,

$$\sum_{i=1}^{10} n_i^2 = 16 \quad , \tag{51}$$

of which 8 ($\Gamma_i^{(m=3)}, i = 1, 8$) are one-dimensional and two ($\Gamma_i^{(m=3)}, i = 9, 10$) are two-dimensional. Again only the two-dimensional irreducible representations are Clifford algebraic.

For α_1 and α_2 one can use the same matrix representatives as in the $m = 2$ case,

$$\Gamma_9^{(m=3)}(\alpha_1) = \Gamma_5^{(m=2)}(\alpha_1) \quad , \quad \Gamma_9^{(m=3)}(\alpha_2) = \Gamma_5^{(m=2)}(\alpha_2) \quad , \tag{52}$$

provided that the corresponding matrix for α_3 is defined by

$$\Gamma_9^{(m=3)}(\alpha_3) = -i\Gamma_9^{(m=3)}(\alpha_1)\Gamma_9^{(m=3)}(\alpha_2) \quad . \tag{53}$$

The second two-dimensional irreducible representation ($\Gamma_{10}^{(m=3)}$) is by the way the complex conjugate representation of $\Gamma_9^{(m=3)}$. It is rather easy to proof that these two irreducible representations are indeed non equivalent.

For the $m = 3$ case the problem of the linearization of the square root reduces therefore to the following matrix equation:

$$\sqrt{p_1^2 + p_2^2 + p_3^2}\Gamma_9^{(m=3)}(I) = \tag{54}$$

$$= p_1\Gamma_9^{(m=3)}(\alpha_1) + p_2\Gamma_9^{(m=3)}(\alpha_2) + p_3\Gamma_9^{(m=3)}(\alpha_3) \quad . \tag{55}$$

The matrices

$$\Gamma_9^{(m=3)}(\alpha_1) \equiv \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \Gamma_9^{(m=3)}(\alpha_2) \equiv \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad ,$$

$$\Gamma_9^{(m=3)}(\alpha_3) \equiv \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad , \tag{56}$$

carry a famous name. They are the so-called **Pauli spin matrices**, usually - as indicated in the last equ. - denoted simply by σ_1, σ_2 and σ_3. For $m = 2, 3$ the corresponding groups are called **Pauli group** (as indicated by the index P).

17.4.3 The Dirac group

For $m = 4$ the following subset of the Clifford algebra forms the smallest group

$$G_D^{(m=4)} = \{\pm I, \pm\alpha_i \ (i \leq 4), \pm\alpha_i\alpha_j \ (i < j), \pm\alpha_i\alpha_j\alpha_k \ (i < j < k),$$
$$\pm\alpha_1\alpha_2\alpha_3\alpha_4 \equiv \pm\alpha_5\} \quad , \tag{57}$$

where "traditionally" the elements α_i are usually also denoted by γ_μ. The order of this group is 32. It has 17 classes,

$$C_1 = \{I\}, \quad C_2 = \{-I\},$$
$$C_{3-6} = \{\pm\alpha_i \mid i \leq 4\}, \quad C_{7-12} = \{\pm\alpha_i\alpha_j \ \mid \ i < j \leq 4\},$$
$$C_{13-16} = \{\pm\alpha_i\alpha_j\alpha_k \mid \ i < j < k \leq 4\}, \quad C_{17} = \{\pm\alpha_1\alpha_2\alpha_3\alpha_4\}, \tag{58}$$

and therefore 17 irreducible representations. As can be checked in analogy to (46) 16 of these irreducible representations ($\Gamma_i^{(m=4)}, i = 1, \ldots, 16$) are one-dimensional and one is four-dimensional ($\Gamma_{17}^{(m=4)}$). Again only the matrices of the four-dimensional irreducible representation satisfy the conditions of the Clifford algebra. The following matrices

$$\Gamma_{17}^{(m=4)}(\alpha_i) \equiv \alpha_i \equiv \gamma_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \quad , \quad i = 1, 2, 3 \quad , \tag{59}$$

$$\Gamma_{17}^{(m=4)}(\alpha_4) \equiv \beta \equiv \gamma_4 = \begin{pmatrix} \underline{1}_2 & 0 \\ 0 & -\underline{1}_2 \end{pmatrix} \quad , \tag{60}$$

where the σ_i are the Pauli spin matrices and $\underline{1}_2$ is a two-dimensional unit matrix, are irreducible representatives of the elements $\alpha_i \in G_D^{(m=4)}$. These particular representatives, usually denoted simply by α_i and β, are called **Dirac matrices**, $G_D^{(m=4)}$ **Dirac group**.

17.4.4 Relations between the Dirac group and the Pauli groups

The subgroup structure The Dirac group contains the Pauli groups as subgroups,

$$G_P^{(m=2)} \subset G_P^{(m=3)} \subset G_D^{(m=4)} \quad , \tag{61}$$

whereby $G_P^{(m=2)}$ is a normal subgroup in $G_P^{(m=3)}$ and $G_D^{(m=4)}$. This implies that in a coset decomposition of $G_D^{(m=4)}$ in terms of $G_P^{(m=2)}$,

$$G_D^{(m=4)} = \{IG_P^{(m=2)}, \alpha_3 G_P^{(m=2)}, \alpha_4 G_P^{(m=2)}, \alpha_3\alpha_4 G_P^{(m=2)}\} \quad , \tag{62}$$

left and right cosets are identical,

$$\begin{aligned} \alpha_3 G_P^{(m=2)} &= \{\pm\alpha_3, \pm\alpha_3\alpha_1, \pm\alpha_3\alpha_2, \pm\alpha_3\alpha_1\alpha_2\} = \\ &= \{\pm\alpha_3, \pm\alpha_1\alpha_3, \pm\alpha_2\alpha_3, \pm\alpha_1\alpha_2\alpha_3\} = G_P^{(m=2)}\alpha_3 \end{aligned} \tag{63}$$

and that $G_P^{(m=2)}$ consists of complete classes of $G_D^{(m=4)}$ (58) denoted for a moment as $C_i(G_D^{(m=4)})$,

$$\begin{aligned} G_P^{(m=2)} = \; & \{C_1(G_D^{(m=4)}), C_2(G_D^{(m=4)}) \quad , \\ & C_3(G_D^{(m=4)}), C_4(G_D^{(m=4)}), C_5(G_D^{(m=4)})\} \quad . \end{aligned} \tag{64}$$

It should be noted that $G_P^{(m=3)}$ is not a normal subgroup in $G_D^{(m=4)}$, since

$$C_{17}(G_D^{(m=4)}) = C_9(G_D^{(m=3)}) \cup C_{10}(G_D^{(m=3)}) \tag{65}$$

Subduced representations The set of matrices

$$\Gamma_{17}^{(m=4)}(G_P^{(m=2)}) \equiv \{\Gamma_{17}^{(m=4)}(\alpha), \forall\alpha \in G_P^{(m=2)}\}$$

and

$$\Gamma_{17}^{(m=4)}(G_P^{(m=3)}) \equiv \{\Gamma_{17}^{(m=4)}(\alpha), \forall\alpha \in G_P^{(m=3)}\}$$

of course also forms a representation for $G_P^{(m=2)}$ and $G_P^{(m=3)}$, respectively, which, however, is reducible. Such representations are called **subduced representations**. Reducing these two representations (for example by means of the orthogonality relation for characters), one finds the following decompositions into irreducible representations:

$$\Gamma_{17}^{(m=4)}(G_P^{(m=2)}) = 2\Gamma_5^{(m=2)}(G_P^{(m=2)}) \tag{66}$$

and

$$\begin{aligned} \Gamma_{17}^{(m=4)}(G_P^{(m=3)}) &= \Gamma_9^{(m=3)}(G_P^{(m=3)}) + \Gamma_{10}^{(m=3)}(G_P^{(m=3)}) \\ &= \Gamma_9^{(m=3)}(G_P^{(m=3)}) + \left(\Gamma_9^{(m=3)}(G_P^{(m=3)})\right)^* \end{aligned} \tag{67}$$

Since the irreducible representation $\Gamma_{17}^{(m=4)}$ of $G_D^{(m=4)}$ always subduces only the group of the Pauli spin matrices (and their complex

conjugates), there is no way to deal with the problem posed by (36) **in terms of** 2×2 **matrices only, i.e., to linearize properly the square root** $\sqrt{\mathbf{p}^2+m^2}$ **for a three-component vector** p**! In other words: there is no other "truely" relativistic description but the one using the Dirac matrices.**

One can summarize the properties of these three groups very compactly in the below short table:

m	Group-order	# of classes	# of one-dimensional irreps	# of two-dimensional irreps	# of four-dimensional irreps
2	8	5	4	1	0
3	16	10	8	2	0
4	32	17	16	0	1
	2^{m+1}	m^2+1	2^m		

(68)

The so-called **fundamental theorem of Dirac matrices**, namely that a necessary and sufficient condition for a set of 4 matrices γ_i' to be Dirac matrices, i.e., to be irreducible and Clifford algebraic, is that they have to be obtained via a similarity transformation W from the matrices in (59,60):

$$\gamma_i' = W^{-1}\gamma_i W \quad , \quad i=1,4 \quad , \tag{69}$$

is in the context of the Dirac group nothing but the Schur lemma for irreducible representations.

17.5 The Dirac Hamiltonian

In terms of the postulates of relativistic quantum mechanics a Hamilton operator corresponding to the Hamilton function in (36) can now be formulated

$$i\hbar\frac{\partial\psi}{\partial t} = H_D\psi \quad , \tag{70}$$

where H_D is the so-called Dirac Hamiltonian:

$$H_D = c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2 \quad , \tag{71}$$

or using $\hbar = c = 1$,

$$H_D = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta m \quad , \tag{72}$$

$$\boldsymbol{\alpha}\cdot\mathbf{p} = \sum_{k=1}^{3}\alpha_k p_k \quad . \tag{73}$$

Equation (70) is not symmetrical with respect to the space and time derivatives (see also (9) and (10)):

$$\left(i\frac{\partial}{\partial t} - \boldsymbol{\alpha}\cdot\mathbf{p} - \beta m\right)\psi = \left(i\underline{1}_4\frac{\partial}{\partial t} + i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} - \beta m\right)\psi = 0 \quad . \tag{74}$$

Multiplying however from the left with β and denoting the resulting set of Dirac matrices by γ^μ,

$$\gamma^\mu = (\gamma^0, \gamma^1, \gamma^2, \gamma^3) = (\gamma^0, \boldsymbol{\gamma}) = \beta(\underline{1}_4, \boldsymbol{\alpha}) \quad , \tag{75}$$

a covariant form of the Dirac equation is obtained

$$(\gamma^\mu (i\partial_\mu) - m)\, \psi = 0 \quad , \tag{76}$$

whose Hermitian conjugated form is given by

$$(-i\partial_\mu)\psi^\dagger \gamma^{\mu\dagger} - m\psi^\dagger = 0 \quad . \tag{77}$$

If one defines now the following quantity $\overline{\psi}$,

$$\overline{\psi} = \psi^\dagger \gamma^0 \rightarrow \psi^\dagger = \overline{\psi}\gamma^0 \quad , \tag{78}$$

then (77) can be rewritten as

$$(-i\partial_\mu)\overline{\psi}\gamma^\mu - m\overline{\psi}\gamma^0 = 0 \quad . \tag{79}$$

Multiplying now (76) from the left with $\overline{\psi}$ and (79) from the right with ψ,

$$\gamma^\mu \overline{\psi}(i\partial_\mu)\psi - m\overline{\psi}\psi = 0$$

$$(-i\partial_\mu)\overline{\psi}\gamma^\mu\psi - m\overline{\psi}\gamma^0\psi = 0$$

References

[1] A.Messiah, *"Quantum Mechanics"*, North Holland Publishing Company, 1969

[2] E.M.Rose, *"Relativistic Electron Theory"*, John Wiley and Sons, 1961

[3] J.D.Bjorken and S.D.Drell, *"Relativistic Quantum Mechanics"*, McGraw Hill Inc., 1964

[4] L.D.Landau and E.M.Lfschitz, *"Lehrbuch der Theoretischen Physik"*, Band IV, *"Relativistische Quantentheorie"*, Akademie Verlag, 1980

[5] L.Jansen and M.Boon, *"Theory of Finite Groups. Applications to Physics"*, North Holland Publishing Company, 1967

[6] I.V.V.Raghavaryulua and N.B.Menon, J.Math.Phys. 11, 3055 (1970)

18 Group theory and quantum mechanics

19 Time-dependence

20 Single-site and multiple scattering

20.1 The muffin-tin approximation

Suppose that the potential $V_{eff}(\mathbf{r})$ in a Schrödinger- or Dirac-type Kohn-Sham Hamiltonian can be written as follows:

$$V_{eff}(\mathbf{r}) = V(\mathbf{r}) = \sum_i V_{eff}^i(\mathbf{r}_i) = \sum_i V_i(\mathbf{r}_i) \quad , \tag{1}$$

$$\mathbf{r}_i = \mathbf{r} - \mathbf{R}_i \quad ,$$

where the 'individual' potentials $V_i(\mathbf{r}_i)$ are loctaed at positions $\mathbf{r}_i$ such that the domains of any arbitrary pair of potentials V_i and V_j, Δ_i and Δ_j, are disjoint in the tensorial space of spin and configuration

$$\Delta_i \cap \Delta_j = 0 \quad , \qquad \forall i, j \quad .$$

The positions $\mathbf{R}_i$ can be positions of Coulomb singularities and can also refer to 'empty' sites. Equation (1) can be viewed from a very pragmatic standpoint, namely as a partioning of the configuration space into regions of different physical interest. For a semi-inifinite crystal, for example, one can thus talk about 'bulk' regions, 'surface' regions, and a 'vacuum' region. Within an infinite crystal, the positions $\mathbf{R}_i$ are quite 'naturally' provided by the real space lattice. The application of eqn (1) can be greatly simplified by using the approximation

$$V_i(\mathbf{r}_i) = \begin{cases} V_i(r_i) & , \ |\mathbf{r}_i| \leq R_s^i \\ \\ \text{constant} & , \ \text{otherwise} \end{cases} \quad , \tag{2}$$

i.e. by using spherical symmetric potentials $V_i(r_i)$. This approximation is traditionally called the **muffin-tin approximation**, and R_s^i is correspondingly refered to as the muffin-tin radius for the i-th sphere. In addition, for all further applications it will be assumed that all individual potentials $V_i(r_i)$ are 'regular potentials':

$$\lim_{r_i \to 0} r_i^2 V_i(r_i) = 0 \quad , \qquad \forall i \quad . \tag{3}$$

Clearly, the great advantage of the muffin-tin approximation is that in each domain Δ_i the corresponding Hamiltonian H_i has very useful constants of motion, such as L^2 and L_z, for example, in the case of a non-relativistic approach. It seems most appropriate therefore to discuss first the scattering from a single finite range spherical potential well and then adress the multi-centre problem by joining the various single-site solutions together. In the following a single site finite range potential $V(r)$,

$$V(\mathbf{r}) = \begin{cases} V(r) & , \ |\mathbf{r}| \leq R_s \\ \\ \text{constant} & , \ \text{otherwise} \end{cases} \quad , \quad \lim_{r \to 0} r^2 V(r) = 0 \quad , \tag{4}$$

will be considered from various standpoints. First of all, however, the case of zero potential has to be considered.

20.2 The zero potential case

20.2.1 The non-relativistic zero potential case

In the absence of a potential $V(r)$ the radial Schrödinger equation is simply given by

$$\left(\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2} + p^2\right) P_l(r) = 0 \quad , \tag{5}$$

$$P_l(r) = rR_l(r) \quad , \quad p^2 = 2\epsilon \quad , \quad \epsilon \geq 0 \quad . \tag{6}$$

Equation (5) is an ordinary second order linear differential equation and therefore has two linear independent solutions. The regular solution must behave like r^{l+1} (see for example Messiah 1969). The regular solution of (5) is given as $zj_l(z)$, where $z = pr$ and $j_l(z)$ is a spherical Bessel function (Abramowitz and Stegun, 1972):

$$j_l(z) = (\pi/2z)^{1/2} J_{l+1/2}(z) \quad , \tag{7}$$

$$j_l(z) \underset{z\to 0}{\longrightarrow} \frac{z^l}{(2l+1)!!} \quad . \tag{8}$$

The other solution, which behaves like r^{-l} (Messiah 1969), is conveniently chosen to be z times a spherical Neumann function, $n_l(z)$ (Abramowitz and Stegun, 1972):

$$n_l(z) = (\pi/2z)^{1/2} N_{l+1/2}(z) \quad , \tag{9}$$

$$n_l(z) \underset{z\to 0}{\longrightarrow} -z^{-l-1}(2l-1)!! \tag{10}$$

As r approaches infinity the spherical bessel and neumann functions show the following limiting behaviour:

$$j_l(z) \underset{|z|\to\infty}{\longrightarrow} \sin(z - l\pi/2)/z \quad , \tag{11}$$

$$n_l(z) \underset{|z|\to\infty}{\longrightarrow} -\cos(z - l\pi/2)/z \quad . \tag{12}$$

However, as r approaches infinity the centrifugal term in (5) vanishes and the solutions should behave like $\exp(\pm ipr)$, which in fact implies that they behave like $\pm\left[n_l(z) \pm ij_l(z)\right]$. The Hankel functions $h^{\pm}(z)$ (Abramowitz and Stegun, 1972) as used traditionally in scattering theory (see, for example, Lloyd and Smith 1972),

$$h^{\pm}(z) = (\pi/2)\, H^{\pm}_{l+1/2}(z) \ , \quad h^{\pm}(z) = j_l(z) \pm in_l(z) \quad , \tag{13}$$

show the required asymptotic behaviour (multiplied by a phase factor $\exp(\pm i\pi/2)$. They refer also to the irregular solutions: $h^+(z)$ and $h^-(z)$ behave asymptotically as 'incoming' and 'outgoing' spherical waves, respectively. Like the Bessel and Neumann functions they are functions in general of a complex argument z.

The spherical Bessel and Neumann functions are summarized for the first few values of l in the following table.

Spherical Bessel and Neumann functions

l	$j_l(z)$	$n_l(z)$
0	$\sin(z)/z$	$-\cos(z)/z$
1	$\sin(z)/z^2 - \cos(z)/z$	$-\cos(z)/z^2 - \sin(z)/z$
2	$\sin(z)(3z^{-3} - z^{-1}) - 3z^{-2}\cos(z)$	$-\cos(z)(3z^{-3} - z^{-1}) - 3z^{-2}\sin(z)$

For $\epsilon < 0$, p in eqn (6) is purely imaginary. In this case the regular and the irregular solutions are the so-called modified Bessel functions $i_l(pr)$ and the modified Hankel functions $k_l^+(pr)$, respectively:

$$i_l(pr) = (-i)^l j_l(ipr) \quad , \tag{14}$$

$$k_l^+(pr) = (-i)^{-l} h^+(ipr) \quad . \tag{15}$$

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, new York, 1972

[2] A. Messiah, *Quantum Mechanics*, North-Holland, Amsterdam, 1969

[3] P. Weinberger, *Electron Scattering Theory for Ordered and Disordered Matter*, Oxford Science Publications, Oxford, 1990