

Integrated Course IIb Theoretical Solid State and Statistical Physics

lecture by

PROF. DR. MILENA GRIFONI

during the winter semester 2012/13

revision and layout in LyX by

ANDREAS VÖLKLEIN



Last changed: February 8, 2013

ATTENTION

This script does *not* replace the lecture.

Therefore it is recommended *strongly* to attend the lecture.

Copyright Notice

Copyright © 2012-2013 ANDREAS VÖLKLEIN

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation;

with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled “GNU Free Documentation License”.

Disclaimer of Warranty

UNLESS OTHERWISE MUTUALLY AGREED TO BY THE PARTIES IN WRITING AND TO THE EXTENT NOT PROHIBITED BY APPLICABLE LAW, **THE COPYRIGHT HOLDERS AND ANY OTHER PARTY, WHO MAY DISTRIBUTE THE DOCUMENT AS PERMITTED ABOVE, PROVIDE THE DOCUMENT “AS IS”, WITHOUT WARRANTY OF ANY KIND**, EXPRESSED, IMPLIED, STATUTORY OR OTHERWISE, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NON-INFRINGEMENT, THE ABSENCE OF LATENT OR OTHER DEFECTS, ACCURACY, OR THE ABSENCE OF ERRORS, WHETHER OR NOT DISCOVERABLE.

Limitation of Liability

IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL THE COPYRIGHT HOLDERS, OR ANY OTHER PARTY, WHO MAY DISTRIBUTE THE DOCUMENT AS PERMITTED ABOVE, BE LIABLE TO YOU FOR ANY DAMAGES, INCLUDING, BUT NOT LIMITED TO, ANY GENERAL, SPECIAL, INCIDENTAL, CONSEQUENTIAL, PUNITIVE OR EXEMPLARY DAMAGES, HOWEVER CAUSED, REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF OR RELATED TO THIS LICENSE OR ANY USE OF OR INABILITY TO USE THE DOCUMENT, EVEN IF THEY HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

IN NO EVENT WILL THE COPYRIGHT HOLDERS’/DISTRIBUTOR’S LIABILITY TO YOU, WHETHER IN CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, EXCEED THE AMOUNT YOU PAID THE COPYRIGHT HOLDERS/DISTRIBUTOR FOR THE DOCUMENT UNDER THIS AGREEMENT.

Links

The text of the “GNU Free Documentation License” can also be read on the following site:

<https://www.gnu.org/licenses/fdl-1.3.en.html>

A transparent copy of the recent version of this document can be downloaded from:

<https://github.com/andiv/IK2b>

Literature

Statistical physics:

- TORSTEN FLIESSBACH: *Lehrbuch zur theoretischen Physik IV: Statistische Physik*; Spektrum Akademischer Verlag, 2007; ISBN: 978-3-8274-1684-1

Solid state physics:

- GERD CZYCHOLL: *Theoretische Festkörperphysik*; Springer, 2008; ISBN: 978-3-540-74789-5
- HENRIK BRUUS, KARSTEN FLENSBERG: *Many-body quantum theory in condensed matter physics*; Oxford University Press, 2007; ISBN: 978-0-19-856633-5
- NEIL W. ASHCROFT, N. DAVID MERMIN: *Solid state physics*; Brooks/Cole Cengage Learning, 2011; ISBN: 978-81-315-0052-1

Contents

I	Thermodynamics	1
0	Introduction to Statistical Physics	2
1	Statistical Methods	3
1.1	Probability	3
1.2	Binomial distribution	5
1.3	Normal distribution	7
1.4	Central limit theorem	9
2	Fundamentals of statistical physics	13
2.1	Fundamental postulate	13
2.2	Partition function of the ideal gas	17
2.3	Thermodynamic processes and the first law	20
2.4	Quasi-static processes	21
2.5	Entropy and temperature	23
2.6	Generalized forces	26
2.7	Chemical potential	27
2.8	Thermodynamic potentials	28
2.9	Second and third law of thermodynamics	30
3	Statistical ensembles	32
3.1	Partition functions	32
3.2	Relation to the thermodynamic potentials	34
4	Applications	36
4.1	Ideal gas	36
4.2	Interacting gas	38
4.3	Indistinguishability of particles	41
4.3.1	Occupation number representation	43
4.3.2	Independent particles	44
4.3.3	Fermi-Dirac statistics	45
4.3.4	Bose-Einstein statistics	47
4.3.5	Quantum mechanical corrections	47
4.4	Density of states	49
4.5	Independent Fermi systems	50
4.6	Independent Bose systems	53
4.6.1	Bose-Einstein condensation	53
4.6.2	Black body radiation	59

II Solid State Theory	62
5 Introduction to Solid State theory	64
5.1 The Hamiltonian of a solid	64
5.2 Structural reducibility (adiabatic approximation)	64
5.3 Universality	66
5.4 Symmetries and symmetry breaking	66
6 Independent electrons in periodic potentials	68
6.1 Crystal structure	68
6.2 Bloch theorem	70
6.3 Linear combination of atomic orbitals (LCAO) method	75
6.4 Almost free electrons	81
6.5 Ground state of Bloch electrons	82
6.6 Alkali and noble metals	84
6.7 Semiconductors	85
7 Second quantization	86
7.1 The formalism of second quantization	87
7.2 Representation of one-body and two-body operators in second quantization	89
7.2.1 One-body operators in second quantization	89
7.2.2 Two-body operators	92
8 Interacting electron gas	93
8.1 Jellium model	93
8.2 Energy scales	94
8.3 Electron interaction effects in perturbation theory	95
8.4 Mean field approximation	97
8.5 Dielectric properties and screening	102
8.5.1 Linear response theory	103
8.5.2 Time dependence of operators in second quantization	105
8.5.3 Susceptibility of non-interacting electron gas	106
8.6 Random phase approximation	108
9 Phonons	111
9.1 Acoustic and optical phonons	111
9.2 Electron-phonon interaction	115
Appendix	117
Acknowledgements	117
GNU Free Documentation License	118

Part I.

Thermodynamics

0 Introduction to Statistical Physics

- Within Statistical Physics one considers systems of *many particles*.
(e.g. atoms of a gas or of a liquid, phonons in solids, photons in a plasma, ...)
⇒ AIM: *Macroscopic properties* of systems with *many microscopic* degrees of freedom
- Although each particle obeys the laws of classical or quantum mechanics, due to the large number ($N \approx 10^{23}$ for a mole of gas), the *coupled* equations of motion are not solvable.
⇒ METHOD: The treatment is *statistical*: It is based on assumptions about the *probability* of a set of trajectories or states.
- In Statistical Physics one starts from microscopic quantities and the laws of classical or quantum mechanics. One then introduces statistical hypotheses and on the basis of statistical methods one
 - i) defines macroscopic quantities and
 - ii) deduces relations among them.

TODO: Abb1

1 Statistical Methods

The transition from microscopic to macroscopic degrees of freedom occurs upon *averaging* over the many degrees of freedom. (\Rightarrow reduction to few degrees of freedom)

1.1. Probability

Let us throw a dice:

TODO: Abb2

- i) *Accessible states*: 6 possible outcomes $w_i \in \{1, 2, 3, 4, 5, 6\}$, all other natural numbers in \mathbb{N} are inaccessible.
- ii) *Probability*: The outcome is probabilistic: By throwing the dice N times the result w_i occurs N_i times.

Definition: The probability of outcome w_i is defined as:

$$p(w_i) = p_i := \lim_{N \rightarrow \infty} \frac{N_i}{N} \quad (1.1)$$

Due to $N = \sum_i N_i$ we get:

$$\sum_i p_i = 1 \quad (1.2)$$

- iii) *Equal a priori probability*: For a “fair” dice all the accessible outcomes have equal probability.

$$p_i = p \quad (1.3)$$

(1.2) implies:

$$\sum_{i=1}^6 p = 1 \quad \Rightarrow \quad p = \frac{1}{6}$$

Note: The consequences of this a priori assumptions can be verified a posteriori in experiments.

Note: Besides a dice we could also consider other systems, e.g. a gas of N atoms or of N molecules. Suppose that the energy ε of the atom is measured with some precision $\Delta\varepsilon$. Then we can assign the possible energies $\varepsilon_i = i \cdot \Delta\varepsilon$ with $i \in \mathbb{N}_0$. If we now perform a measurement of the ensemble of N atoms, the probability $p(\varepsilon_i) = p_i$ is again:

$$p(\varepsilon_i) = \frac{N_i}{N}$$

Here N_i is the number of atoms with energy ε_i .

(This quantity will be explicitly calculated later in the course.)

iv) *Addition and multiplication:*

- Let us consider the probability p that by throwing the dice, we get one of the two outcomes w_i of w_j ($i \neq j$). Because the two events are mutually excluding, we have

$$N_{i \text{ or } j} = N_i + N_j$$

and from (1.1) follows:

$$p_{i \text{ or } j} = p_i + p_j \quad (1.4)$$

- If one has two dices, the probability that the first dice is in w_i and the second in w_j is for independent events:

$$p_{ij} = \lim_{N, M \rightarrow \infty} \frac{N_i M_j}{NM} = p_i p_j \quad (1.5)$$

v) *Mean value:*

Experiment 1: Throw a dice N times and obtain N values $w_j = w(t_j)$, $j \in \{1, 2, \dots, N\}$.

$$\Rightarrow \quad \bar{w} = \frac{1}{N} \sum_j w(t_j) \quad \text{time average} \quad (1.6)$$

Experiment 2: Throw N identical dice once and obtain the N values w_j .

$$\bar{\bar{w}} = \frac{1}{N} \sum_j w_j \quad \text{ensemble average} \quad (1.7)$$

Note: For small N is in general $\bar{w} \neq \bar{\bar{w}}$.

Definition: The *thermodynamic limit* is the limit of large systems, ideally $N \rightarrow \infty$.

Assumption: In the thermodynamic limit is $\bar{w} = \bar{\bar{w}}$.

With the definitions (1.1), (1.6) and (1.7) follows:

$$\bar{\bar{w}} = \bar{w} = \sum_{w_i \in \text{accessible states}} w_i p_i \quad (1.8)$$

This relation has predictive power, if the p_i are known. For the case of a dice we have:

$$\bar{w} = \sum_{\text{a.s.}} w_i \cdot \frac{1}{6} = 3,5$$

vi) *Mean square deviation:*

- *mean deviation:*

$$\overline{w - \bar{w}} = \sum_i p_i (w_i - \bar{w}) = \underbrace{\sum_i p_i w_i}_{=\bar{w}} - \bar{w} \underbrace{\sum_i p_i}_{=1} = 0$$

- *mean square deviation:*

$$\overline{(w - \bar{w})^2} = \overline{w^2} - \bar{w}^2 \quad (1.9)$$

- *standard deviation:*

$$\Delta w = \sqrt{\overline{(w - \bar{w})^2}}$$

1.2. Binomial distribution

We consider a system of N “particles”, whose constituents are characterized by only two possible states. This is called a *binary system*.

Examples

- i) *Spin chain*: Consider N spins with $S_z = \pm \frac{1}{2}$.
There are several possible configurations:

$$\uparrow\uparrow\uparrow \dots \uparrow\uparrow \qquad \uparrow\downarrow\downarrow \dots \downarrow\uparrow \qquad \uparrow\uparrow\downarrow\downarrow \dots \uparrow\uparrow$$

- ii) *One dimensional random walk*: Go along the x -axis in steps $\pm\Delta x$ (to the right “+” or to the left “-”) starting from $x = 0$.

$$++++ \dots ++ \qquad +-+- \dots -+ \qquad ++-- \dots ++$$

Let us consider the random walk and define p (respectively q) as the probability of step $+\Delta x$ ($-\Delta x$) with $p+q=1$ and n_+ (n_-) as the number of steps to the right (left) with $n_+ + n_- = N$. Define $m := n_+ - n_-$ to get the position $m\Delta x$ of the walker after N steps. It holds $n_+ = \frac{N+m}{2}$.

We wish to calculate the probability $P_N(m)$ that the walker is in $m\Delta x$ after N steps.

Note that $P_N(m) = W_N(n_+)$ is the probability of having performed n_+ positive steps out of N steps.

Calculation of $P_N(m) = W_N(n_+)$

- i) Consider the sequence $+- - ++$. The probability for this sequence is p^3q^2 .
ii) The sequence has $n_+ = 3$ and $N = 5$. There are other sequences which also have $n_+ = 3$ and $N = 5$. In general the number of configurations $\Omega(N, n_+)$ is called “multiplicity”. For $N = 5$, $n_+ = 3$ it is:

$$\Omega(5, 3) = \frac{5!}{3! \cdot 2!} = \binom{5}{3}$$

- iii) For arbitrary N and n_+ holds similarly:

$$\Omega(N, n_+) = \frac{N!}{n_+!n_-!} = \frac{N!}{n_+!(N-n_+)!} = \binom{N}{n_+} \quad (1.10)$$

So the probability $P_N(m) = W_N(n_+)$ is:

$$W_N(n_+) = \Omega(N, n_+) p^{n_+} q^{N-n_+} \quad (1.11)$$

This is the binomial distribution.

Note: In the following we write n instead of n_+ .

- The Binomial theorem gives:

$$\sum_{n=0}^N W_N(n) = \sum_{n=0}^N \binom{N}{n} p^n q^{N-n} = (p+q)^N = 1$$

– The total number of configurations is:

$$N_{\text{conf}} = \sum_{n=0}^N \Omega(N, n) = \sum_{n=0}^N \binom{N}{n} = \sum_{n=0}^N \binom{N}{n} 1^n 1^{N-n} = (1+1)^N = 2^N$$

– We have:

$$m = n_+ - n_- = n_+ - (N - n_+) = 2n_+ - N$$

Therefore m lies in $\{-N, N\}$ in steps of 2.

Mean values

$$\begin{aligned} \bar{n}_+ = \bar{n} &= \sum_{n=0}^N n W_N(n) = \sum_{n=0}^N n \binom{N}{n} p^n q^{N-n} = \\ &= p \frac{\partial}{\partial p} \underbrace{\left(\sum_{n=0}^N \binom{N}{n} p^n q^{N-n} \right)}_{=(p+q)^N} = pN (p+q)^{N-1} = pN \end{aligned}$$

Analogously follows:

$$\begin{aligned} \bar{n}_- &= qN \\ \bar{m} &= \bar{n} - \bar{n}_- = N(p - q) \end{aligned}$$

Variance

$$\overline{n^2} = \sum_{n=0}^N n^2 \binom{N}{n} p^n q^{N-n} = p \frac{\partial}{\partial p} \left(p \frac{\partial}{\partial p} \left(\sum_{n=0}^N \binom{N}{n} p^n q^{N-n} \right) \right) = Npq + \bar{n}^2$$

Therefore we get the standard deviation:

$$\begin{aligned} \Delta n &= \sqrt{\overline{n^2} - \bar{n}^2} = \sqrt{Npq} \\ \Delta m &= 2\sqrt{Npq} \end{aligned}$$

Law of large numbers

$$\boxed{\frac{\Delta n}{n} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{n}}} \xrightarrow[n \rightarrow \infty]{N \rightarrow \infty} 0 \quad (1.12)$$

I.e. for large numbers N , the distribution $W_N(n)$ approaches a distribution strongly peaked around \bar{n} . Let us e.g. consider $p = q = \frac{1}{2}$, for example for a spin chain in zero magnetic field.

$$W_N(n) = \frac{1}{2^N} \binom{N}{n} = \frac{\Omega(N, n)}{\sum_{m=0}^N \Omega(N, m)} = P_N(m)$$

Therefore it is enough to focus on $\Omega(N, n)$.

Example: $N = 10$

$$\begin{array}{ll} \Omega(N, 0) = 1 & \Omega(N, 3) = 120 \\ \Omega(N, 1) = 10 & \Omega(N, 4) = 210 \\ \Omega(N, 2) = 45 & \Omega(N, 5) = 252 \end{array}$$

The other $\Omega(N, n)$ follow from $\Omega(N, n) = \Omega(N, N - n)$. The total number of possibilities is:

$$\sum_{m=0}^N \Omega(N, m) = 2^{10} = 1024$$

We can plot now $\Omega(N, n)$ or $\tilde{\Omega}(N, m) = \Omega(N, \frac{N+m}{2})$.

TODO: Abb1

1.3. Normal distribution

In the limit of large N (and $Npq \gg 1$) the binomial distribution approaches a Gauss function, also known as *normal distribution*. This is a special case of the *central limit theorem*.

- To this extent it is convenient to investigate the behavior of $\ln(W_N(n))$, which has a smoother variation in n , than $W_N(n)$.
- From (1.11) follows:

$$\ln(W_N(n)) = \ln(N!) - \ln(n!) - \ln((N-n)!) + n \ln(p) + (N-n) \ln(q)$$

Moreover holds:

$$\begin{aligned} \frac{d}{dn} \ln(n!) &\stackrel{n \gg 1}{\approx} \frac{\ln((n+1)!) - \ln(n!)}{n+1-n} = \ln(n+1) \stackrel{n \gg 1}{\approx} \ln(n) \\ \Rightarrow \quad \frac{d}{dn} \ln(W_N(n)) &\approx -\ln(n) + \ln(N-n) + \ln(p) - \ln(q) \stackrel{n=\bar{n}}{=} 0 \end{aligned} \quad (1.13)$$

Whereby $\bar{n} = Np \gg 1$ holds and:

$$\bar{n} = Nq \gg 1 \quad \Leftrightarrow \quad Npq \gg 1 \quad (1.14)$$

After (1.13) the first derivative vanishes at $n = \bar{n}$, and thus follows that $\ln(W_N(n))$ and hence $W_N(n)$ have a stationary point at $n = \bar{n}$.

- Likewise it holds:

$$\frac{d^2}{dn^2} \ln(W_N(n)) = -\frac{1}{n} - \frac{1}{N-n} \stackrel{n=\bar{n}}{=} -\frac{1}{Npq} = -\frac{1}{(\Delta n)^2} < 0$$

Therefore $W_N(n)$ has a maximum at $n = \bar{n}$.

- Furthermore follows:

$$\frac{d^3}{dn^3} \ln(W_N(n)) = \frac{1}{n^2} - \frac{1}{(N-n)^2} \stackrel{n=\bar{n}}{=} \frac{q^2 - p^2}{N^2 p^2 q^2}$$

– Thus the Taylor expansion is:

$$\begin{aligned}\ln W_N(n) &= \ln W_N(\bar{n}) + \frac{d}{dn} W_N(n) \Big|_{\bar{n}} (n - \bar{n}) + \frac{1}{2} \frac{d^2}{dn^2} W_N(n) \Big|_{\bar{n}} (n - \bar{n})^2 + \\ &\quad + \frac{1}{6} \frac{d^3}{dn^3} W_N(n) \Big|_{\bar{n}} (n - \bar{n})^3 + o_0((n - \bar{n})^3) \\ &\approx \ln W_N(\bar{n}) - \frac{1}{2} \frac{(n - \bar{n})^2}{Npq} = \ln W_N(\bar{n}) - \frac{(n - \bar{n})^2}{2(\Delta n)^2}\end{aligned}$$

Gauss distribution

This gives the following probability distribution, called *Gauss distribution*:

$$\boxed{W_N(n) = W_N(\bar{n}) \exp\left(-\frac{(n - \bar{n})^2}{2(\Delta n)^2}\right)} \quad (1.15a)$$

It is a good approximation for:

$$\left| \frac{(\ln W)'''(n - \bar{n})^3}{(\ln W)''(n - \bar{n})^2} \right| \approx \frac{|n - \bar{n}|}{Npq} \ll 1$$

For $n - \bar{n} \approx \Delta n = \sqrt{Npq}$ this means:

$$\frac{1}{\sqrt{Npq}} \ll 1$$

Note: The condition $N \gg 1$ is not sufficient to get (1.15a). For example if $p \ll 1$ it could still be $Np = \bar{n} \ll 1$. One approaches in this case from a binomial distribution the *Poisson distribution*:

$$P(\lambda, n) = \frac{\lambda^n}{n!} e^{-\lambda} \quad \lambda = Np$$

– To determine $W_N(\bar{n})$ we observe:

$$\begin{aligned}1 &= \sum_{n=0}^N W_N(n) \approx \int_0^N dx W_N(x) \approx \int_{-\infty}^{\infty} dx W_N(x) = \\ &= W_N(\bar{n}) \int_{-\infty}^{\infty} dx e^{-\frac{(x - \bar{n})^2}{2(\Delta n)^2}} = W_N(\bar{n}) \sqrt{2\pi} \Delta n\end{aligned}$$

The first approximation is good, because the relative variation of $W_N(n)$ between n and $n + 1$ is small around the average \bar{n} and the second is good, because we make only an exponentially small error. So we get:

$$W_N(n) = \frac{1}{\sqrt{2\pi} \Delta n} e^{-\frac{(n - \bar{n})^2}{2(\Delta n)^2}} \quad (1.15b)$$

- Let us introduce the length l and define $x = nl$, $\bar{x} = \bar{n}l$, $\sigma := \Delta x = \Delta nl = \sqrt{Npq} \cdot l$.
Now the probability density $P(x)$ along the x -axis is defined by setting $P(x) dx$ as the probability to find the random variable between x and $x + dx$.

$$P(x) dx \approx \frac{dx}{l} W_N\left(\frac{x}{l}\right)$$

Now follows from (1.15b) the *normal distribution*:

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \bar{x})^2}{2\sigma^2}\right) \quad (1.16)$$

Note: For $x = (n_+ - n_-)l$ we have $\bar{x} = N(p - q)l$ and $\sigma = 2\sqrt{Npq}l$, but the form (1.16) still holds.

- *Properties of the Gauss distribution:*

$$\int_{-\infty}^{\infty} dx P(x) = 1 \quad \int_{\bar{x}-\nu\sigma}^{\bar{x}+\nu\sigma} dx P(x) = \begin{cases} 0,683 & \nu = 1 \\ 0,954 & \nu = 2 \\ 0,997 & \nu = 3 \end{cases}$$

This means, that an event outside of the 3σ range occurs with probability 0,3%, which is highly improbable!

1.4. Central limit theorem

If the quantity $x = \sum_i s_i$ is the sum of *many* independent random variables, then the probability function $P(x)$ for x is the normal distribution.

Proof

Let us consider a variable x such that

$$x = s_1 + s_2 + \dots + s_N = \sum_{i=1}^N s_i$$

where s_i are *random variables* with:

- $w_i(s_i) ds_i$ is the probability to find the i -th variable with values between s_i and $s_i + ds_i$.

Examples:

- If s_i is the i -th step in a random walk, then x is the particle's position after N steps.
- If s_i is the energy of the i -th atom of a gas, then x is the total energy.

- For all i holds:

$$\begin{aligned} \int_{-\infty}^{\infty} ds w_i(s) &= 1 \\ \bar{s}_i &= \int_{-\infty}^{\infty} ds w_i(s) s < \infty \\ (\Delta s_i)^2 &= \int_{-\infty}^{\infty} ds (s - \bar{s}_i)^2 w_i(s) < \infty \end{aligned}$$

iii) The s_i are independent variables.

Then the average value of x is:

$$\bar{x} = \int_{-\infty}^{\infty} ds_1 W_1(s_1) \dots \int_{-\infty}^{\infty} ds_N W_N(s_N) (s_1 + s_2 + \dots + s_N) = \sum_{i=1}^N \bar{s}_i$$

The mean square deviation is:

$$\begin{aligned} (\Delta x)^2 &= \int_{-\infty}^{\infty} ds_1 W_1(s_1) \dots \int_{-\infty}^{\infty} ds_N W_N(s_N) (x - \bar{x})^2 = \\ &= \int_{-\infty}^{\infty} \dots \left(\sum_{i=1}^N (s_i - \bar{s}_i) \right)^2 = \int_{-\infty}^{\infty} \dots \sum_{i=1}^N (s_i - \bar{s}_i) \sum_{j=1}^N (s_j - \bar{s}_j) = \\ &= \int_{-\infty}^{\infty} \dots \left[\sum_{i=1}^N (s_i - \bar{s}_i)^2 + \underbrace{\sum_{\substack{i,j=1 \\ i \neq j}}^N (s_i - \bar{s}_i)(s_j - \bar{s}_j)}_{\int_{-\infty}^{\infty} \dots = 0} \right] = \\ &= \int_{-\infty}^{\infty} \dots \sum_{i=1}^N (s_i - \bar{s}_i)^2 = \sum_{i=1}^N (\Delta s_i)^2 \end{aligned}$$

Therefore follows:

$$\frac{\Delta x}{\bar{x}} = \frac{\sqrt{\sum_{i=1}^N (\Delta s_i)^2}}{\sum_{i=1}^N \bar{s}_i} = \mathcal{O}_{\infty} \left(\frac{1}{\sqrt{N}} \right)$$

This amounts to the law of large numbers.

E.g. for $w_i(s) = w(s)$ follows $\bar{s}_i = \bar{s}$, $\bar{x} = N\bar{s}$ and $(\Delta x)^2 = N(\Delta s)^2$.

$$\Rightarrow \frac{\Delta x}{\bar{x}} = \frac{\Delta s}{\bar{s}} \cdot \frac{1}{\sqrt{N}} \quad (1.17)$$

Example: Consider a gas of $N = 10^{24}$ atoms. The i -th atom has the energy ε_i with probability $w(\varepsilon_i) = \exp\left(-\frac{\varepsilon_i}{k_B T}\right)$, where k_B Boltzmann constant and T the temperature.

$$\begin{aligned} \bar{\varepsilon} &\approx \Delta \varepsilon \approx k_B T \\ \Rightarrow \frac{\Delta \varepsilon}{\bar{\varepsilon}} &\approx 1 \end{aligned}$$

But $E = \sum_i \varepsilon_i$ is sharply peaked!

$$\frac{\Delta E}{\bar{E}} = \frac{\Delta \varepsilon}{\bar{\varepsilon} \sqrt{N}} \approx 10^{-12}$$

In general the probability density for x is:

$$P(x) = \int_{-\infty}^{\infty} ds_1 w_1(s_1) \dots \int_{-\infty}^{\infty} ds_N w_N(s_N) \delta\left(x - \sum_{i=1}^N s_i\right)$$

The trick is to Fourier transform the delta distribution:

$$\delta(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{iky}$$

With the inverse Fourier transformation of this follows:

$$\begin{aligned} P(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} ds_1 w_1(s_1) \dots \int_{-\infty}^{\infty} ds_N w_N(s_N) e^{ik(s_1+\dots+s_N)} e^{-ikx} = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} \left(\prod_{j=1}^N W_j(k) \right) \end{aligned}$$

$$W_j(k) := \int_{-\infty}^{\infty} ds_j w_j(s_j) e^{iks_j}$$

Let us expand $W_j(k)$ for small k :

$$W_j(k) = 1 + ik\bar{s} - \frac{1}{2}k^2\bar{s}_j^2 + \mathcal{O}_0(k^3)$$

Hence we get:

$$\ln \left(\prod_{j=1}^N W_j(k) \right) = \sum_{j=1}^N \ln \left(1 + ik\bar{s}_j - \frac{1}{2}k^2\bar{s}_j^2 + \mathcal{O}_0(k^3) \right)$$

Expanding the logarithm gives:

$$\ln(1+y) = y - \frac{y^2}{2} + \mathcal{O}_0(y^3)$$

$$\begin{aligned} \Rightarrow \ln \left(\prod_{j=1}^N W_j(k) \right) &= ik \sum_{j=1}^N \bar{s}_j - \frac{1}{2}k^2 \sum_{j=1}^N (\Delta s_j)^2 + \mathcal{O}(k^3) \approx \\ &\approx ik\bar{x} - \frac{k^2}{2} (\Delta x)^2 \end{aligned}$$

Now follows

$$\prod_{j=1}^N W_j(k) \approx e^{ik\bar{x} - \frac{k^2}{2} (\Delta x)^2}$$

and hence:

$$P(x) \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-i(x-\bar{x})k} e^{-(\Delta x)^2 \frac{k^2}{2}}$$

Calculating the integral gives:

$$P(x) = \frac{1}{\sqrt{2\pi}\Delta x} \exp \left(-\frac{(x-\bar{x})^2}{2(\Delta x)^2} \right)$$

$P(x)$ is, as stated before, a Gaussian distribution.

□_{1.4}

Regime of validity

The theorem is based on the truncation of the expansion of the function $W(k)$, which is good for:

$$\left| \frac{\overline{s^{n+1}} k^{n+1}}{\overline{s^n} k^n} \right| \ll 1$$

Our case is $n = 2$. One can estimate:

$$k \approx \frac{1}{\Delta x} \approx \frac{1}{\sqrt{N} \Delta s}$$

$$\Rightarrow \left| \frac{\overline{s^{n+1}}}{\overline{s^n}} k \right| \approx \left| \frac{\overline{s^{n+1}}}{\overline{s^n}} \frac{1}{\Delta s} \right| \frac{1}{\sqrt{N}} \approx \frac{1}{\sqrt{N}} \xrightarrow{N \rightarrow \infty} 0$$

2 Fundamentals of statistical physics

2.1. Fundamental postulate

Microstate

A *microstate* is defined by specifying completely the microscopic state of a system.

- *Example 1:* Let the system be a group of N dices. Hence the microstate r is defined upon assigning the values $n_i \in \{1, 2, \dots, 6\}$ to each of the N dices.

$$r = (n_1, n_2, \dots, n_N) \quad (2.1)$$

The possible number of microstates is:

$$\Omega(N) = 6^N$$

- *Example 2:* Let the system be a physical system with f degrees of freedom. We consider a *closed system*, such that the associated Hamilton function/operator does not depend on time. In this case the Hamilton function/operator yields the *energy* of the microstate.
- *Quantum mechanical case:* As microstates we choose the eigenstates $\{|r\rangle\}$ of the Hamilton operator \hat{H} determined by the time independent Schrödinger equation:

$$\hat{H} |r\rangle = E_r |r\rangle \quad (2.2)$$

Here E_r is the energy of the microstate. For a system with f degrees of freedom the eigenstates depend on f quantum numbers:

$$r = (n_1, n_2, \dots, n_f)$$

Note: We assumed here that the closed system is restricted to a finite Volume V , such that the quantum numbers are discrete.

- *Ideal gas of N particles in a box of volume V :* The interaction between gas particles is neglected (ideal gas). The total Hamilton operator is the sum of the Hamilton operators \hat{h}_ν of the ν -th particle:

$$\hat{H} = \sum_{\nu=1}^N \hat{h}_\nu = \sum_{\nu=1}^N \frac{\hat{p}_\nu^2}{2m} \quad (2.3)$$

Specifically, it holds:

$$\hat{h}_\nu |s\rangle_\nu = \varepsilon_\nu(s) |s\rangle_\nu$$

In the position representation

$$\psi_{s,\nu}(\vec{r}) = \langle \vec{r} | s \rangle_\nu$$

we get:

$$-\frac{\hbar^2}{2m} \nabla^2 \psi_{s,\nu}(\vec{r}) = \varepsilon_\nu(s) \psi_{s,\nu}(\vec{r})$$

For a particle in a box, $\psi_{s,\nu}(\vec{r})$ vanishes at the boundaries. With $V = L^3$ it holds:

$$\begin{aligned} 0 &= \psi_{s,\nu}(0, y, z) = \psi_{s,\nu}(x, 0, z) = \psi_{s,\nu}(x, y, 0) \\ 0 &= \psi_{s,\nu}(L, y, z) = \psi_{s,\nu}(x, L, z) = \psi_{s,\nu}(x, y, L) \end{aligned}$$

This gives for $n_{\nu x}, n_{\nu y}, n_{\nu z} \in \mathbb{N}_{\geq 1}$:

$$\begin{aligned} \psi_{s,\nu}(\vec{r}) &= A \sin\left(\frac{n_{\nu x} \pi x}{L}\right) \sin\left(\frac{n_{\nu y} \pi y}{L}\right) \sin\left(\frac{n_{\nu z} \pi z}{L}\right) \\ \varepsilon_\nu(s) &= \frac{\hbar^2}{2m} \left(\frac{\pi}{L}\right)^2 (n_{\nu x}^2 + n_{\nu y}^2 + n_{\nu z}^2) \end{aligned} \quad (2.4)$$

Therefore the state $|s\rangle_\nu$ and the energy $\varepsilon_\nu(s)$ are fixed by the quantum numbers $n_{\nu x}, n_{\nu y}$ and $n_{\nu z}$. Thus the microstate r of the system of N atoms is defined by assigning the $3N$ quantum numbers $r = (n_1, n_2, \dots, n_{3N})$ with the following order:

$$\begin{aligned} (n_1, n_2, n_3) &= (n_{1x}, n_{1y}, n_{1z}) \\ (n_4, n_5, n_6) &= (n_{2x}, n_{2y}, n_{2z}) \\ &\vdots \end{aligned}$$

Note: There are infinitely many microstates r as each quantum number n_k can take infinitely many values. However, for *fixed energy*, the number becomes *finite*.

Note: Phase-space volume per quantum state

The available states can be represented on a three-dimensional lattice in the space of the $\{p_x, p_y, p_z\}$ with $p_i = \frac{\hbar \pi}{L} n_i$. Due to $n_i > 0$, all *distinct* points are within the octant with $p_i > 0$.

TODO: Abb1

The number of states in a shell with radii between $p = \sqrt{p_x^2 + p_y^2 + p_z^2}$ and $p + dp$ is:

$$\frac{1}{8} \cdot \frac{4\pi p^2 dp}{\left(\frac{\hbar \pi}{L}\right)^3} = \frac{4\pi p^2 dp}{h^3} V$$

In general, the number of states with energy $\varepsilon(p) \leq \varepsilon$ is:

$$\frac{V}{h^3} \cdot 4\pi \int_0^{p_m} p^2 dp = \frac{1}{h^3} \cdot \int_V d^3r \int_{p \leq p_m} d^3p$$

So the minimal volume in phase space is $2\pi\hbar = h$.

Note: Consider the free particle with periodic boundary conditions.

$$\psi(x) = A e^{ikx} \quad \psi(0) = \psi(L) \quad \Rightarrow \quad k = \frac{2\pi n}{L} \quad n \in \mathbb{Z}$$

Because p_x, p_y, p_z can now assume both positive and negative values, one has to consider a full sphere. In this case the infinitesimal number of states is:

$$\frac{4\pi p^2 dp}{\left(\frac{\hbar 2\pi}{L}\right)^3}$$

Again the minimal volume in phase space is $h = 2\pi\hbar$.

- *Classical case*: In the classical case, a system with f degrees of freedom is characterized by assigning the values of f generalized coordinates q_1, q_2, \dots, q_f as well as generalized impulses p_1, p_2, \dots, p_f . Hence there are $2f$ microstates:

$$r = (q_1, \dots, q_f, p_1, \dots, p_f)$$

For an ideal classical gas of N particles there are $2f = 6N$ microstates and a microstate is:

$$r = (\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, \vec{p}_1, \vec{p}_2, \dots, \vec{p}_N) = (x_1, y_1, z_1, x_2, \dots, z_N, p_{x1}, p_{y1}, p_{z1}, \dots, p_{zN})$$

The energy is given by $H(q_1, q_2, \dots, q_f, p_1, \dots, p_f)$.

- *Phase space*: The $2f$ -dimensional space spanned by the variables $\{q_1, p_1, \dots\}$ is called phase-space. So to each classical microstate r is associated a point in phase-space.

TODO: Abb2

For a statistical description we need to be able to *enumerate* the number of microstates r . To this extent we observe that in general it is neither necessary nor possible to give the values (q_i, p_i) exactly. According to the uncertainty principle it holds:

$$\Delta q \Delta p \geq \frac{\hbar}{2} \quad (2.5)$$

We count thus the number of classical states upon observing that for the quantum case to one state corresponds a surface h in phase space. With the $2f$ -dimensional phase space volume V_{ps} the number of states is therefore:

$$N_{\text{states}} = \frac{V_{ps}}{h^f} \quad (2.6)$$

Macrostate

For a many-particle-system with a very large number N of particles, the number of microstates with the same energy is also very large. The system will constantly perform transitions $r \rightarrow r'$ between (accessible) microstates.

In general we are not interested to know details of the evolution of a microstate. Rather we are interested in:

- Which* are the accessible microstates.
- What* is the probability P_r of occurrence of the microstate r .

A *macrostate* is the state of a system defined only through the set of probabilities P_r for every configuration r .

$$\{P_r\} = \{P_1, P_2, \dots\} \quad (2.7)$$

These probabilities are defined according to equation (1.1):

$$P_r = \lim_{M \rightarrow \infty} \frac{M_r}{M} \quad (2.8)$$

Here one has M identical systems and the microstate r occurs upon measurement M_r times.

Fundamental postulate

Let us consider a closed system. From experience we know that independent of the initial preparation, a closed system will evolve after some time to a certain macrostate, which is called *equilibrium state*. The *macroscopically measurable quantities* are those, that attain *constant* values, which are typically averages of a microscopic quantity over all particles. The *fundamental postulate* states:

“A closed system in equilibrium is with equal probability in any of its accessible microstates.” (ergodic hypothesis)

(This is due to the chaotic behavior of practically any system of many particles.)

Note: The fundamental postulate defines the connection between the microscopic structure, i.e. the accessible microstates $\{r\}$, and macroscopic quantities of the equilibrium state, i.e. the probabilities $\{P_r\}$.

– *Example 1:* System of N dices

The number of accessible states is 6^N , so the probability for each is:

$$P_r = \frac{1}{6^N}$$

– *Example 2:* Four electrons in a magnetic field \vec{B}

The relevant degrees of freedom are the spin \hat{s}_z of each electron along the quantization axis defined by $\vec{B} = B\vec{e}_z$.

$$\hat{L}_i = -\hat{\mu}_i \cdot \vec{B} \qquad \hat{\mu}_i = -\mu_B g \frac{\hat{s}_i}{\hbar}$$

So a microstate is characterized by $r = (s_{z,1}, s_{z,2}, s_{z,3}, s_{z,4})$ with $s_{z,j} \in \pm \frac{\hbar}{2}$. Its energy is:

$$E_r = 2\mu_B B \sum_{j=1}^4 s_{z,j}$$

Suppose that the energy in the equilibrium state is $E = 2\mu_B B$.

The accessible states (for distinguishable spins, e.g. sitting on a lattice) are (denoting $s_z = \frac{\hbar}{2}$ by \uparrow and $s_z = -\frac{\hbar}{2}$ by \downarrow):

$$r \in \{(\uparrow\uparrow\downarrow\downarrow), (\uparrow\uparrow\downarrow\uparrow), (\uparrow\downarrow\uparrow\uparrow), (\downarrow\uparrow\uparrow\uparrow)\} =: M$$

For example you can't find $(\uparrow\uparrow\uparrow\uparrow)$, because this state has a different energy. The fundamental postulate yields the probability:

$$P_r = \begin{cases} \frac{1}{4} & \text{for } r \in M \\ 0 & \text{otherwise} \end{cases}$$

Alternative formulation of the fundamental postulate

Let Ω be the number of accessible states. Then the fundamental postulate can be written as:

$$P_r = \begin{cases} \frac{1}{\Omega} & \text{for accessible microstates} \\ 0 & \text{otherwise} \end{cases} \quad (2.9)$$

- For a closed system, Ω depends on the energy E and a family $x = (x_1, \dots, x_n)$ additional external parameters. For this one writes

$$\Omega = \Omega(E, x) \quad (2.10)$$

and calls Ω the *microcanonical partition function*. For a Hamiltonian $H = H(q, p, x)$, the Hamilton operator $\hat{H} = \hat{H}(x)$ depends on x , so in general the energies $E_r = E_r(x)$ also depend on x . For example for a gas of N atoms in a box of volume V we have $x = (V, N)$.

- As the energy E can be assigned only with some uncertainty δE , the partition function is the number of microstates with energies $E_r(x)$ in the range $(E - \delta E, E)$:

$$\Omega(E, x) = \sum_{\substack{r \\ E - \delta E \leq E_r(x) \leq E}} 1 \quad (2.11)$$

This yields:

$$P_r(E, x) = \begin{cases} \frac{1}{\Omega(E, x)} & E - \delta E \leq E_r(x) \leq E \\ 0 & \text{otherwise} \end{cases} \quad (2.12)$$

2.2. Partition function of the ideal gas

The external variables are $x = (V, N)$. These are *extensive* variables, i.e. if you take two of the system and consider them together, these variables double, in contrast to *intensive* variables, which stay the same. We now proceed with the following steps:

$$\hat{H}(V, N) \xrightarrow{1} \varepsilon_r(V, N) \xrightarrow{2} \Omega(E, V, N) \xrightarrow{3} \ln(\Omega) \quad (2.13)$$

Step 1: Eigenvalues of $\hat{H}(V, N)$

- From (2.3) we know:

$$\hat{H}(V, N) = \sum_{\nu=1}^N \hat{h}_\nu$$

So for a microstate $r = (n_1, \dots, n_{3N})$ with $n_k \in \mathbb{N}_{\geq 1}$ we have from (2.4):

$$E_r(V, N) = \sum_{\nu=1}^{3N} \frac{p_\nu^2}{2m} = \sum_{\nu=1}^{3N} \frac{\pi^2 \hbar^2 n_\nu^2}{2mL^2} \quad (2.14)$$

- *Indistinguishability of particles*: When evaluating properties of the ideal gas, we should account for the fact that in quantum mechanics, upon exchanging particles, the wave function changes at most its sign due to the particles indistinguishability, i.e. an exchange of two particles yields the *same state*.

In general for a given microstate $r = (n_1, \dots, n_{3N})$ one can generate $N!$ representations of the *same* macrostate upon exchange of particle indices. Thus when counting states, we have to take care, not to count the same microstate more than once.

Step 2: Evaluation of $\Omega(E, V, N)$

Now we introduce:

$$\phi(E, V, N) := \sum_{\substack{r \\ E_r(V, N) \leq E}} 1 \quad (2.15)$$

Given this, we can calculate the partition function:

$$\Omega(E, V, N) = \phi(E) - \phi(E - \delta E) \quad (2.16)$$

Hence we focus on ϕ .

$$\phi(E) = \sum_{\substack{n_1 \\ E_r \leq E}} \sum_{n_2} \dots \sum_{n_{3N}} \frac{1}{N!}$$

The $\frac{1}{N!}$ accounts for the indistinguishability of the particles. The average momentum $\overline{p_r}$ per particle is:

$$\overline{p_r} = \frac{\pi \hbar}{L} \overline{n_r}$$

For $\overline{n_r} \gg 1$, i.e. $\overline{p_r} \gg \frac{\pi \hbar}{L}$, the steps for increasing a quantum number are small compared to the value of $\overline{p_r}$ and we can replace the sums with integrals:

$$\begin{aligned} \phi(E) &\approx \frac{1}{N!} \underbrace{\int dn_1 \dots \int dn_{3N}}_{E_r \leq E} 1 = \frac{1}{N!} \left(\frac{L}{\pi \hbar} \right)^{3N} \underbrace{\int_0^\infty dp_1 \dots \int_0^\infty dp_{3N}}_{\sum_\nu p_\nu^2 \leq 2mE} 1 = \\ &= \frac{1}{N!} \cdot \frac{1}{2^{3N}} \cdot \frac{1}{(\pi \hbar)^{3N}} \underbrace{\int_{-\infty}^\infty dp_1 \dots \int_{-\infty}^\infty dp_{3N}}_{\sum_\nu p_\nu^2 \leq 2mE} \int_0^L dx_1 \dots \int_0^L dx_n 1 = \\ &= \frac{\text{phase-space volume}}{N! \cdot h^{3N}} \end{aligned} \quad (2.17)$$

The integral

$$\underbrace{\int_{-\infty}^\infty dp_1 \dots \int_{-\infty}^\infty dp_{3N}}_{\sum_\nu p_\nu^2 \leq 2mE} = V_{\sqrt{2mE}}^{3N} \quad (2.18)$$

is the volume of a $3N$ -dimensional sphere of radius $\sqrt{2mE}$. Using the Γ -function with the fundamental property $\Gamma(m+1) = m!$ for $m \in \mathbb{N}$, this volume can be expressed as:

$$V_R^{3N} = c_{3N} R^{3N} \quad c_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)}$$

This gives:

$$\phi(E, V, N) = \frac{1}{N!} \tilde{c}_{3N} V^N E^{\frac{3N}{2}} \quad \tilde{c}_{3N} = c_{3N} \frac{(2m)^{\frac{3N}{2}}}{h^{3N}} \quad (2.19)$$

For example for $\frac{3N}{2} = 10^{24}$ and $\frac{\delta E}{E} = 10^{-5}$ this gives:

$$\frac{\phi(E)}{\phi(E - \delta E)} = \frac{E^{10^{24}}}{(E - \delta E)^{10^{24}}} = \frac{1}{(1 - 10^{-5})^{10^{24}}} \approx \left(1 + \frac{10^{19}}{10^{24}}\right)^{10^{24}} \stackrel{\text{definition of exp}}{\approx} \exp(10^{19})$$

So we can practically neglect $\phi(E - \delta E)$:

$$\boxed{\Omega(E, V) = \phi(E, V) - \phi(E - \delta E, V) \approx \phi(E, V)} \quad (2.20)$$

This is independent of δE .

Step 3: $\ln(\Omega)$

With the approximation $\Gamma(x + 1) \approx \left(\frac{x}{e}\right)^x$ with Euler's number $e \approx 2,718$ and remembering

$$\Gamma(N + 1) = N!$$

for $N \in \mathbb{N}$, we get from equation (2.19):

$$\begin{aligned} \phi(E, V, N) &\approx \left(\frac{e}{N}\right)^N \frac{\pi^{\frac{3N}{2}}}{\left(\frac{3N}{2e}\right)^{\frac{3N}{2}}} \frac{(2m)^{\frac{3N}{2}}}{h^{3N}} V^N E^{\frac{3N}{2}} = \\ &= \left(\frac{V}{N}\right)^N \underbrace{\left(\left(\frac{4}{3}m\pi\frac{e}{h^2}\right)^{\frac{3}{2}}\right)^N}_{=:c} \left(\frac{E}{N}\right)^{\frac{3N}{2}} \end{aligned} \quad (2.21)$$

This gives the microcanonical partition function ideal gas:

$$\ln(\Omega(E, V, N)) \approx \frac{3}{2}N \ln\left(\frac{E}{N}\right) + N \ln \frac{V}{N} + N \ln(c) \quad (2.22a)$$

Note: $f = 3N$ is the number of degrees of freedom.

$$\Omega(E) \sim \left(\frac{E}{N}\right)^{\frac{3N}{2}} \sim \left(\frac{E}{f}\right)^{\gamma f} \quad (2.22b)$$

This dependence holds also for other systems with many degrees of freedom under the condition, that by increasing the energy also the number of accessible states increases. This condition does not hold in general for spin systems.

Example: The energy for N electrons in a magnetic field B is:

$$E_r = 2\mu_B \frac{B}{\hbar} \sum_{\nu=1}^N s_{z,\nu}$$

So only $E \in [E_{\min}, E_{\max}]$ is possible, i. e. the energy is bounded by:

$$E_{\min/\max} = \mp N\mu_B B$$

Since at E_{\min} and E_{\max} only one state is possible, (2.22b) only holds if:

$$\mu_B B \ll E - E_{\min} \ll N\mu_B B$$

2.3. Thermodynamic processes and the first law

- *Mean value of the energy*: From the Hamiltonian $\hat{H}(x)$ we can derive the energies $E_r(x)$. The probabilities P_r then define a macrostate. This allows us to define the mean energy of the macrostate:

$$\bar{E}_r = \sum_{\substack{r \\ E - \delta E \leq E_r(x) \leq E}} P_r E_r(x) \quad (2.23)$$

Because $\delta E \ll E$ holds $\bar{E}_r \approx E$.

- *Thermodynamic process*: Let us consider a situation whereby the system performs a transition from a macrostate a with \bar{E}_a to a macrostate b with \bar{E}_b . We are interested in:

$$\Delta E = E_b - E_a$$

Note: A transition where $E_b \neq E_a$ can only occur, when the system is interacting with the surroundings. The interaction can yield a change of x (and hence of $E_r(x)$) and/or of P_r .

- *Heat and work*: The energy change ΔE can be decomposed into two contributions:
 1. The average energy changes, while the external parameters x are kept constant, so this must be due to a change of the P_r .

$$\Delta E := \Delta Q \quad (2.24a)$$

One says the variation occurs in form of *heat* transferred *to* the system.

$\Delta Q > 0$	heat is added to the system
$\Delta Q < 0$	heat is extracted from the system
$\Delta Q = 0$	the system is thermally isolated (adiabatic process)

2. The parameters x are changed, while the system is thermally isolated.

$$\Delta E := \Delta W \quad (2.24b)$$

One says the variation occurs in form of *work* done *on* the system.

In general holds:

$$\boxed{\Delta E = \Delta Q + \Delta W} \quad (2.24c)$$

This is known as the *first law of thermodynamics*.

Note: If the system is closed holds $\Delta E = 0$ and even $\Delta W = \Delta Q = 0$.

Note: The heat transfer is associated to changes in P_r . This is because if e.g. heat is added, the states with larger energy have larger probability of being occupied.

In general this implies a non-equilibrium probability distribution during the heat transfer. After waiting long enough, a new equilibrium configuration characterized by $P_r(E_b, x_b)$ is reached.

If, however, the heat transfer is *quasi-static*, i.e. very slow, then one can at each instant of time use the initial probability distribution $P_r(E, x)$ with the actual value of the energy.

Note: When the x -parameter is changed, this in general induces a change of $E_r(x)$ and P_r . This is because $E_r = E_r(x)$ and due to the time-dependent perturbation, transitions between microstates can occur and hence P_r can change to a non-equilibrium distribution. But if the process is quasi-static, the change in x does not induce a change in P_r .

– *Example:*

- a) A container with gas is heated up. There is no change in x , so $\Delta E = \Delta Q$.

TODO: Abb3

- b) The volume of gas is reduce/increased. If the system is thermally isolated, we have $\Delta E = \Delta W$.

TODO: Abb4

– *Differentials:* Very often we shall consider infinitesimal changes. Then equation (2.24c) becomes:

$$dE = dW + dQ \quad (2.25)$$

The notation “ d ” means, that ΔQ and ΔW are no *state variables* (“Zustandsgrößen”). In other words

$$\Delta W \neq \Delta W(x) \quad \Delta Q \neq \Delta Q(x)$$

and hence are not set when e.g. the equilibrium state characterized by E and x is reached (in contrast to $E = E(x)$).

To see, why dQ and dW are not associated to state variables, let us look at processes, which yield the *same* final state with $E_b = E_a + \Delta E$ and $V_a = V_b$.

- i) The heat $\Delta Q = \Delta E$ is added.
- ii) The work $\Delta W = \Delta E$ is added to the system by varying in a time dependent way the volume of the gas with $V_a = V_b$ at the end.

So dQ and dW are no exact differentials. However from $E = E(x) = E(x_1, \dots, x_n)$ follows:

$$dE = \sum_{i=1}^n \left(\frac{\partial E}{\partial x_i} \right) dx_i \quad (2.26)$$

So this is an exact differential and hence:

$$\oint dE = 0$$

Here \oint means that we are considering a cyclic process, i.e. $E_a = E_b$.

2.4. Quasi-static processes

– *Definition:* A process is *quasi-static* if the changes on the system are so slow, that it evolves through a sequence of *equilibrium configurations*.

Let τ_{ext} be the timescale, on which changes occur (e.g. heat transfer, change in x) and τ_{rel} the time, the system needs to relax to equilibrium. So a process is quasi-static if and only if:

$$\tau_{\text{ext}} \gg \tau_{\text{rel}}$$

– *Example:*

TODO: Abb5

The gas is expanded by slowly moving the piston and the container is thermally isolated. The external parameter is $x_1 = L(t)$, and hence the parameters characterizing the timescales are:

$$\begin{aligned} v_1 &= \frac{dx_1}{dt} && \text{velocity of the piston} \\ \bar{v} &&& \text{average velocity of gas particles} \end{aligned}$$

$$\Rightarrow \quad \tau_{\text{ext}} \gg \tau_{\text{rel}} \quad \Leftrightarrow \quad v_1 \ll \bar{v}$$

Non quasi-static effects are proportional to $\frac{v_1}{\bar{v}}$. Quantum numbers do not change if this is small.

Generalized forces

Consider a quasi-static process where x_1 changes and x_2, \dots, x_n remain constant. From equation (2.23) follows:

$$\begin{aligned} dE &= d \left(\sum_r P_r E_r(x_1, x_2, \dots, x_n) \right) = \sum_r (dP_r) E_r + \sum_r P_r \frac{\partial E_r}{\partial x_1} dx_1 = \\ &= \sum_r (dP_r) E_r + \frac{\overline{\partial E_r}}{\partial x_1} dx_1 \stackrel{\text{quasi-static}}{=} dQ_{\text{qs}} + \underbrace{\frac{\overline{\partial E_r}}{\partial x_1} dx_1}_{=: dW_{\text{qs}}} \stackrel{\text{adiabatic}}{=} dW_{\text{qs}} \end{aligned} \quad (2.27)$$

– *Definition:* The *generalized force* associated to x_i is:

$$X_i = - \frac{\partial E(x_1, x_2, \dots, x_n)}{\partial x_i} \quad (2.28)$$

Note: The X_i are intensive quantities, associated to the extensive external parameters x_i .

Note: If also other parameters are exchanged, we get:

$$dW_{\text{qs}} = - \sum_{i=1}^n X_i dx_i \quad (2.29)$$

– *Pressure:* Let us consider a gas enclosed in a container with a piston.

TODO: Abb6

$$x_1 = L$$

The gas exerts a force F on the piston, so for an infinitesimal change dL follows:

$$dW_{\text{qs}} = -F dL = -\frac{F}{A} A dL = -P dV \quad (2.30)$$

This is the work performed on the gas. Then equation (2.29) yields:

$$dW_{\text{qs}} = -X_1 dx_1 = \frac{\overline{\partial E_r(V)}}{\partial V} dV$$

And hence we get the microscopic definition of pressure:

$$\boxed{P = - \frac{\overline{\partial E_r(V)}}{\partial V}} \quad (2.31)$$

Note: For a gas the pressure P is always positive.

$$\Rightarrow \quad \begin{cases} dV > 0 & \Rightarrow & dW_{\text{qs}} < 0 & \text{the system performs work} \\ dV < 0 & \Rightarrow & dW_{\text{qs}} > 0 & \text{work is done on the system} \end{cases} \quad (2.32)$$

Pressure of an ideal gas

Consider the quantum mechanical point of view!

As usual let the gas of N particles be contained in a box of volume $V = L_1 L_2 L_3$.

Now we vary $x_1 = L_1$. From equation (2.23) we get:

$$E_r(V, N) = \sum_{\nu=1}^N \sum_{j=1}^3 \frac{\hbar^2 \pi^2}{2m L_j^2} n_{3\nu+j-3}^2$$

$$\mathrm{d}W_{\text{qs}} = \frac{\overline{\partial E_r}}{\partial L_1} \mathrm{d}L_1 = -\frac{2}{L_1} \sum_{\nu=1}^N \frac{\hbar^2 \pi^2}{2m L_1^2} n_{3\nu-2}^2 \mathrm{d}L_1$$

In equilibrium all accessible states are equiprobable, therefore all impulse directions are equiprobable and hence the average kinetic energy is the same for all three directions:

$$\overline{\sum_{\nu=1}^N \frac{\hbar^2 \pi^2}{2m L_1^2} n_{3\nu-2}^2} = \frac{1}{3} \sum_{j=1}^3 \sum_{\nu=1}^N \frac{\hbar^2 \pi^2}{2m L_j^2} n_{3\nu+j-3}^2 = \frac{1}{3} \overline{E_r} = \frac{E}{3}$$

Hence we get:

$$\mathrm{d}W_{\text{qs}} = -\frac{2}{L_1} \frac{E}{3} \mathrm{d}L_1 = -\frac{2}{3} \frac{E}{V} \mathrm{d}V \quad (2.33)$$

This gives the pressure for the ideal gas:

$$\boxed{P = \frac{2}{3} \frac{E}{V}} \quad (2.34)$$

Note: The classical description is discussed by FLIESSBACH in chapter 8, page 60.

Note on reversible and irreversible processes: Corrections to the quasi-static result are of the order of $\frac{v_1}{\bar{v}}$ with $v_1 = \frac{\mathrm{d}x_1}{\mathrm{d}t}$. For a generic process thus follows:

$$\frac{\mathrm{d}E}{E} = \left(\frac{\mathrm{d}E}{E} \right)_{\text{qs}} + \mathcal{O}_0 \left(\frac{v_1}{\bar{v}} \right) = \left(\frac{\mathrm{d}E}{E} \right)_{\text{reversible}} + \left(\frac{\mathrm{d}E}{E} \right)_{\text{irreversible}} \quad (2.35)$$

Here the reversible part would vanish when considering a full expansion-compression cycle. The second term does not vanish and accounts for the positive increase of energy due to the conversion of work into heat. In general holds:

$$\mathrm{d}W \geq \mathrm{d}W_{\text{qs}} \quad (2.36)$$

2.5. Entropy and temperature

Consider two systems A and B in thermal contact with each other and with the external parameters x_A and x_B kept constant.

TODO: Abb7

How is the total energy

$$E = E_B + E_A$$

distributed in equilibrium? Or equivalently, what is the probability $W(E_A)$ for an energy E_A ?

i) Step: In equilibrium all of the accessible states of the *total* system are equiprobable:

$$P_r = \frac{1}{\Omega_0(E)}$$

$\Omega_0(E)$ is the partition function of the composite system.

ii) Step:

$$W(E_A) = \sum_{E_A, E_B=E-E_A}^r P_r = \sum_{E_{r,A}=E_A} \sum_{E_{r,B}=E-E_A} \frac{1}{\Omega_0(E)} = \frac{\Omega_A(E_A) \Omega_B(E-E_A)}{\Omega_0(E)} \quad (2.37)$$

iii) Step: Remember now equation (2.22b):

$$\Omega(E) = c \left(\frac{E}{f} \right)^{\gamma f} \quad \gamma \approx 1$$

Assuming $\gamma_A = \gamma_B = \gamma$ we get:

$$\begin{aligned} \Rightarrow \ln(W(E_A)) &= \ln(\Omega_A(E_A)) + \ln(\Omega_B(E-E_A)) - \ln \Omega_0(E) = \\ &= \gamma f_A \ln E_A + \gamma f_B \ln(E-E_A) + \text{const.} \end{aligned}$$

The function $\ln(W(E_A))$ has a maximum in the interval $[0, E]$ set by:

$$\frac{d}{dE_A} \ln(W(E_A)) = \frac{\gamma f_A}{E_A} - \frac{\gamma f_B}{E-E_A} \stackrel{!}{=} 0$$

So we have a maximum at $E_{A,\max} = \overline{E_A}$ with:

$$\begin{aligned} \frac{\overline{E_A}}{f_A} &= \frac{\overline{E_B}}{f_B} \\ \overline{E_B} &= E - \overline{E_A} \end{aligned} \quad (2.38)$$

So the energy per degree of freedom is the same in the system *A* and *B*!

– Expanding about the maximum yields:

$$\begin{aligned} \ln(W(E_A)) &\approx \ln(W(\overline{E_A})) - \frac{(E - \overline{E_A})^2}{2\Delta E_A^2} \\ \Rightarrow W(E_A) &= W(\overline{E_A}) e^{-\frac{(E - \overline{E_A})^2}{2\Delta E_A^2}} \\ \frac{1}{\Delta E_A^2} &= -\frac{d^2 \ln W}{dE_A^2} = \frac{\gamma f_A}{\overline{E_A}^2} + \frac{\gamma f_B}{\overline{E_B}^2} \end{aligned} \quad (2.39)$$

This equation gives:

$$\Delta E_A = \left(\frac{\gamma f_A}{\overline{E_A}^2} + \frac{\gamma f_B}{\overline{E_B}^2} \right)^{-\frac{1}{2}} < \frac{\overline{E_A}}{\sqrt{\gamma f_A}}$$

So if f_A is very large, the probability $W(E_A)$ sharply peaked around $\overline{E_A}$:

$$W(E_A) \approx \begin{cases} W(\overline{E_A}) & E_A = \overline{E_A} \\ 0 & \text{otherwise} \end{cases} \quad (2.40)$$

This means $E_A = \overline{E_A}$ with $\overline{E_A}$ given by (2.38) sets the condition for equilibrium after heat exchange.

- Entropy S and temperature T are defined as:

$$S = S(E, x) := k_B \cdot \ln(\Omega(E, x)) \quad \text{Entropy} \quad (2.41)$$

$$\frac{1}{T} = \frac{1}{T(E, x)} := \frac{\partial S(E, x)}{\partial E} \quad \text{Temperature} \quad (2.42)$$

Here Ω is the multiplicity of the considered system and k_B is the Boltzmann constant:

$$k_B \approx 1,38 \cdot 10^{-23} \frac{\text{J}}{\text{K}}$$

- The equilibrium condition yields:

$$\ln(\Omega_A(E_A) \Omega_B(E - E_A)) = \max.$$

$$\Leftrightarrow S(E_A) = k_B \ln(\Omega_A \Omega_B) = S_A(E_A, x) + S_B(E - E_A, x) = \max. \quad (2.43)$$

This means:

$$\begin{aligned} 0 &= \frac{\partial S}{\partial E_A} = \frac{\partial S_A}{\partial E_A} - \frac{\partial S_B}{\partial E_B} \\ \Leftrightarrow T_A &= T_B \end{aligned} \quad (2.44)$$

So a temperature difference induces an exchange of heat.

Notes:

- i) Entropy is additive and therefore an extensive variable.
- ii) If there is only one accessible microstate, i.e. $\Omega = 1$, then follows $S = 0$.
- iii) The temperature is a measure of the mean energy per degree of freedom. Indeed we have:

$$\Omega \sim \left(\frac{E}{f}\right)^{\gamma f}$$

$$\frac{1}{k_B T} = \frac{1}{k_B} \frac{\partial S}{\partial E} = \frac{\partial}{\partial E} \ln(\Omega(E, V)) = \frac{\gamma f}{E}$$

Now follows:

$$\boxed{k_B T = \frac{E}{\gamma f}} \quad (2.45)$$

Because E is the energy of the excitations, it is non negative and therefore $T \geq 0$.

Note, that we are going to find for some special systems, where this law does not hold, negative temperatures.

- We consider the case, where the system can also have energy exchange via variation of x . In equilibrium the probability $W(E, x)$ is proportional to the number of configurations $\Omega(E, x)$ with a proportionality constant C :

$$W(E, x) = C \Omega(E, x) = C \exp\left(\frac{S(E, x)}{k_B}\right)$$

Upon expanding $\ln(W(x))$ around the maximum of $S(E, x)$ at \bar{x} and performing the same steps as for $W(E_A)$ we obtain ($x = x_A$):

$$W(x) = \frac{1}{\sqrt{2\pi}\Delta x} \exp\left(-\frac{(x-\bar{x})^2}{2(\Delta x)^2} + \mathcal{O}_0(x-\bar{x})^3\right)$$

$$\Delta x = \sqrt{\frac{-k_B}{\frac{\partial^2 S}{\partial x^2}|_{x=\bar{x}}}}$$

Since x is extensive, it holds $x = \mathcal{O}_\infty(N)$. Additionally we get

$$\Delta x = \mathcal{O}_\infty\left(\left(\frac{N}{N^2}\right)^{-\frac{1}{2}}\right) = \mathcal{O}_\infty(\sqrt{N})$$

and thus:

$$\frac{\Delta x}{x} = \mathcal{O}_\infty\left(\frac{1}{\sqrt{N}}\right) \xrightarrow{N \rightarrow \infty} 0$$

Hence again all microstates are located at the maximum for very large N . And hence the maximum determines also the equilibrium configuration:

$$\boxed{S(E, x) = \max.} \quad (2.46)$$

2.6. Generalized forces

We have introduced the generalized forces in equation (2.28) as:

$$X_i = -\frac{\overline{\partial E_r(x)}}{\partial x_i}$$

Let us now consider:

$$\frac{\partial}{\partial x_i} \ln \Omega(E, x) = \frac{\ln(\Omega(E, x + dx_i)) - \ln(\Omega(E, x))}{dx_i} \stackrel{\star}{=} X_i \frac{\partial \ln \Omega}{\partial E} = \frac{X_i}{k_B T}$$

For \star compare FLIESSBACH, chapter 10, page 78. This yields:

$$\boxed{X_i = k_B T \frac{\partial \ln \Omega(E, x)}{\partial x_i} = T \frac{\partial S(E, x)}{\partial x_i}} \quad (2.47)$$

This equation will provide e. g. the caloric equation of state.

– *Example:* For $x = V$ we have $X = P$ and hence:

$$P = T \frac{\partial S(E, V)}{\partial V}$$

– *Note:* The equilibrium condition (2.46) corresponds to the equality of the associated generalized forces.

TODO: Abb8

$$E = E_A + E_B = \text{const.}$$

$$V = V_A + V_B = \text{const.}$$

In equilibrium holds:

$$S(E_A, V_A) = S_A(E_A, V_A) + S_B(E - E_A, V - V_A) = \max.$$

Then follows:

$$\begin{aligned} \frac{\partial S}{\partial E_A} = 0 & \Rightarrow \frac{1}{T_A} - \frac{1}{T_B} = 0 & \Leftrightarrow & T_A = T_B \\ \frac{\partial S}{\partial V_A} = 0 & \Rightarrow \frac{P_A}{T_A} - \frac{P_B}{T_B} = 0 & \Leftrightarrow & P_A = P_B \end{aligned}$$

– Note on differentials:

$$S(E, x) \Rightarrow dS = \frac{\partial S}{\partial E} dE + \sum_i \frac{\partial S}{\partial x_i} dx_i$$

$$dS = \frac{1}{T} dE + \frac{1}{T} \sum_i X_i dx_i \quad (2.48a)$$

$$dE = T dS - \sum_i X_i dx_i \quad (2.48b)$$

2.7. Chemical potential

So far, we have considered situations with fixed N . However, there can be cases, where N changes. In this case, we define the *chemical potential* μ as the negative generalized force associated to N :

$$\boxed{\mu := -k_B T \frac{\partial \ln \Omega(E, N)}{\partial N}} \quad (2.49)$$

Hence follows:

$$\mu \stackrel{\text{definition of } S}{=} -T \frac{\partial S(E, N, V)}{\partial N} \stackrel{\text{definition of } T}{=} \frac{\partial E(S, N, V)}{\partial N}$$

It is the energy necessary to add an extra particle to the system.

Note: From (2.48b) and (2.49) follows with $E = E(V, N)$ the *Gibbs fundamental form*:

$$\boxed{dE = T dS - P dV + \mu dN} \quad (2.50)$$

Equilibrium condition for exchange of heat, volume and particle number

TODO: Abb9

$$E = E_A + E_B = \text{const.}$$

$$V = V_A + V_B = \text{const.}$$

$$N = N_A + N_B = \text{const.}$$

From $S(E_A, V_A, N_A) = \max.$ we get:

$$T_A = T_B$$

$$P_A = P_B$$

$$\mu_A = \mu_B$$

In general the generalized forces are the same in equilibrium.

2.8. Thermodynamic potentials

Depending on which variables are varied in a thermodynamic process, it might be convenient to introduce a specific thermodynamic function, which is a *state function* with *dimension of energy*. Given a system with energy $E = E(S, V, N)$ one introduces:

– *Internal energy*:

$$E \quad (2.51a)$$

– *Free energy*:

$$F = E - TS \quad (2.51b)$$

– *Enthalpy*:

$$H = E + PV \quad (2.51c)$$

– *Free enthalpy* also called *Gibbs energy*:

$$G = E - TS + PV \quad (2.51d)$$

Then follows:

$$dE = TdS - PdV + \mu dN \quad (2.52a)$$

$$dF = -SdT - PdV + \mu dN \quad (2.52b)$$

$$dH = TdS + VdP + \mu dN \quad (2.52c)$$

$$dG = -SdT + VdP + \mu dN \quad (2.52d)$$

The *natural variables* are therefore:

$$E = E(S, V, N)$$

$$F = F(T, V, N)$$

$$H = H(S, P, N)$$

$$G = G(T, P, N)$$

A fifth thermodynamic potential of interest for later purposes is:

$$J = E - TS - \mu N = F - \mu N \quad (2.53a)$$

$$dJ = -SdT - PdV - Nd\mu \quad (2.53b)$$

In this case the natural variables are:

$$J = J(T, V, \mu)$$

Equilibrium conditions

How does the equilibrium condition “ S maximal” for the closed system $A + B$ translate to equilibrium conditions for the thermodynamic potentials?

$$E_A \text{ is minimal given } S, V, N. \quad (2.54a)$$

$$F_A \text{ is minimal given } T, V, N. \quad (2.54b)$$

$$H_A \text{ is minimal given } T, P, N. \quad (2.54c)$$

$$G_A \text{ is minimal given } S, P, N. \quad (2.54a)$$

Let us proof e.g. (2.54b):

We start again from A and B being in thermal contact with each other. Let us assume, that A is much smaller than B , i.e. B acts as a thermal bath:

$$E_A \ll E = E_A + E_B$$

$$V_A \ll V = V_A + V_B$$

Therefore the contact with B fixes $T := T_B$ and $P := P_B$ for the system A . Suppose now only heat is exchanged.

$$S(E_A) = S_A(E_A, V_A) + S_B(E - E_A, V_B) = \max.$$

Taylor expanding yields:

$$\begin{aligned} S &= S_A + S_B(E, V_B) - \frac{\partial S_B(E_B, V_B)}{\partial E_B} E_A = \\ &= \text{const.} + S_A - \frac{E_A}{T} = \max. \end{aligned}$$

So S being maximal is equivalent to:

$$F_A = E_A - TS_A = \min.$$

Relation between G and μ

From $G = G(T, P, N)$ follows, because G is an extensive quantity and among T, P and N only N is extensive:

$$G = N \cdot g(T, P)$$

From (2.52d) follows:

$$\begin{aligned} \mu &= \frac{\partial G(T, P, N)}{\partial N} = g(T, P) = \frac{G}{N} \\ \Rightarrow G &= N\mu \end{aligned} \quad (2.55)$$

Moreover holds:

$$\begin{aligned} \mu dN + Nd\mu &= dG \stackrel{(2.51d)}{=} -SdT + VdP + \mu dN \\ \Rightarrow d\mu &= -\frac{S}{N}dT + \frac{V}{N}dP := -sdT + vdP \end{aligned}$$

This is the *Duhem-Gibbs relation* and so $\mu = \mu(P, T)$. The equilibrium condition demands μ to be minimal.

Finally, (2.51d) and (2.55) imply:

$$\boxed{J = E - TS - \mu N = -PV} \quad (2.56)$$

For applications see the thermodynamic part of the integrated course IIa.

2.9. Second and third law of thermodynamics

Second law

- *Closed system*: We have demonstrated that a closed system tends to an equilibrium situation, whereby the entropy S acquires its maximal value. So in a process $a \rightarrow b$ the entropy change is:

$$\Delta S = S_b - S_a \geq 0$$

This is the *second law* of thermodynamics for *closed* systems:

$$\boxed{\Delta S \geq 0} \quad (2.57)$$

- *Open system*: Let us consider the differential of the entropy (2.48a):

$$\begin{aligned} dS &= \frac{dE}{T} + \frac{1}{T} \sum_{i=1}^n X_i dx_i \stackrel{(2.29)}{=} \frac{dE}{T} - \frac{dW_{\text{qs}}}{T} = \\ &\stackrel{1.\text{law}}{=} \frac{1}{T} (dQ + dW - dW_{\text{qs}}) \end{aligned}$$

So for a quasi-static process holds:

$$\boxed{dS_{\text{qs}} = \frac{dQ_{\text{qs}}}{T}} \quad (2.58)$$

From $dW \geq dW_{\text{qs}}$ follows directly the *second law* of thermodynamics for *open* systems:

$$\boxed{dS \geq \frac{dQ}{T}} \quad (2.59)$$

Third Law

Now we investigate the limiting behavior of the entropy for $E \rightarrow 0$.

When lowering the energy the system approaches the lowest lying states and in particular its ground state with energy E_0 . Corresponding to E_0 , there exists usually only one possible state of the system, or, if the ground state is degenerate, only a small number of states. In this low temperature regime $\Omega \approx f$ and hence:

$$S \approx k_B \ln(f)$$

At higher energies holds

$$\Omega \sim (E - E_0)^{\gamma f}$$

and hence:

$$S \approx k_B \gamma f \ln(E - E_0) \sim k_B \gamma f$$

Because we have many degrees of freedom, it follows:

$$k_B \ln f \ll k_B \gamma f$$

Therefore the following is an excellent approximation:

$$\boxed{S \xrightarrow{E \rightarrow E_0} 0} \quad (2.60a)$$

Due to $\frac{\partial S}{\partial E} > 0$ follows that with $E \rightarrow E_0$ also converges $T \rightarrow T_0$, whereby T_0 is the smallest achievable temperature. Now the *third law*, also called *Nernst theorem*, states:

$$\boxed{S \xrightarrow{T \rightarrow T_0} 0} \quad (2.60b)$$

Note:

$$\begin{aligned} \frac{1}{k_B T} &= \frac{\partial \ln \Omega}{\partial E} \approx \frac{\gamma f}{E - E_0} \xrightarrow{E \rightarrow E_0} \infty \\ k_B T &\xrightarrow{E \rightarrow E_0} 0 \\ \Rightarrow \quad T_0 &= 0 \end{aligned} \quad (2.60c)$$

3 Statistical ensembles

3.1. Partition functions

Recall: The microcanonical partition function $\Omega(E, x)$ defines the equilibrium state given by the energy E of a closed system via the relation (2.12):

$$P_r(E, x) = \begin{cases} \frac{1}{\Omega(E, x)} & E - \delta E \leq E_r \leq E \\ 0 & \text{otherwise} \end{cases}$$

The ensemble with such a probability is called *microcanonical*.

Canonical ensemble

Similarly, we wish to determine the probability for a subsystem, which can exchange heat with a heat bath of temperature T . The statistical ensemble defined now by $P_r(T, x)$ is called *canonical*.

$$P_r = \begin{cases} P_r(E, x) & \text{microcanonical} \\ P_r(T, x) & \text{canonical (Gibbs ensemble)} \end{cases} \quad (3.1)$$

TODO: Abb10

Now let us determine P_r for the canonical ensemble: The subsystem A can exchange energy (in form of heat) with a thermal bath B .

$$\begin{aligned} A + B &= \text{closed} \\ \Rightarrow E_A + E_B &= \text{const.} \end{aligned}$$

- i) The energies $\{E_r\}$ of the small system should be much smaller than the total energy E :

$$E_r \ll E \quad (3.2)$$

- ii) The probability that A is in a given state r is:

$$P_r = \frac{\overbrace{1}^{A \text{ is in } r} \cdot \Omega_B(E - E_r)}{\Omega(E)} \quad (3.3)$$

- iii) A Taylor expansion yields:

$$\ln(\Omega_B(E - E_r)) \approx \ln(\Omega_B(E)) - \frac{\partial}{\partial E} \ln(\Omega_B(E)) \cdot E_r =$$

$$= \ln(\Omega_B(E)) - \frac{E_r}{k_B T}$$

$$\Rightarrow \Omega_B(E - E_r) \approx \Omega_B(E) e^{-\beta E_r(x)} \quad \beta := \frac{1}{k_B T}$$

This gives the *Boltzmann distribution*:

$$P_r(T, x) = \frac{1}{Z} e^{-\beta E_r(x)} \quad (3.4)$$

From

$$\sum_r P_r = 1$$

follows for the *canonical partition function*:

$$Z = \sum_r e^{-\beta E_r} \quad (3.5)$$

Grandcanonical Ensemble

Analogously one can determine the probability

$$P_r = P_r(T, x, \mu) \quad (3.6)$$

of occurrence of r for the *grandcanonical ensemble*, when particles and heat can be exchanged with a bath characterized by T and μ . The probability is now calculated:

TODO: Abb11

$$E_A + E_B = \text{const.}$$

$$N_A + N_B = \text{const.}$$

$$\text{i) } E_r \ll E, N_r \ll N$$

$$\text{ii) } P_r = \frac{1 \cdot \Omega_B(E - E_r, N - N_r)}{\Omega(E, N)}$$

$$\text{iii) } \ln(\Omega_B(E - E_r, N - N_r)) \approx \ln(\Omega_B(E, N)) - \beta E_r + \beta \mu N_r$$

$$\Omega_B(E - E_r, N - N_r) = \Omega_B(E, N) \cdot \exp(-\beta(E_r(x, N_r) - \mu N_r))$$

And hence follow the *grandcanonical probability*

$$P_r(T, x, \mu) = \frac{1}{Y} e^{-\beta(E_r - \mu N_r)} \quad (3.7)$$

and the *grandcanonical partition function*:

$$Y = \sum_r e^{-\beta(E_r - \mu N_r)} \quad (3.8)$$

Note: In general we have $r = (r', N_r)$, whereby r' is the microstate for a given particle number N_r and hence:

$$Y(T, x, \mu) = \sum_{N_r=0}^N \sum_{r'} e^{-\beta(E_{r'} - \mu N_r)} \approx \sum_{N'=0}^{\infty} Z(T, x, N') e^{\beta \mu N'}$$

3.2. Relation to the thermodynamic potentials

Similar to the case of the microcanonical ensemble, knowledge of Z or Y is enough to determine all relevant thermodynamic quantities and thermodynamic relations (e.g. the caloric equation of state).

$$\hat{H}(V, N) \rightarrow E_r(V, N) \rightarrow \begin{cases} \Omega(E, V, N) \rightarrow S(E, V, N) \\ Z(T, V, N) \rightarrow F(T, V, N) \\ Y(T, V, \mu) \rightarrow J(T, V, \mu) \end{cases} \rightarrow \text{all thermodynamic relations}$$

i) *Relation between Z and F :*

$$Z = Z(T, V, N) \\ \Rightarrow d(\ln Z) \stackrel{\beta = \frac{1}{k_B T}}{=} \frac{\partial \ln(Z)}{\partial \beta} d\beta + \frac{\partial \ln(Z)}{\partial V} dV + \frac{\partial \ln(Z)}{\partial N} dN$$

From (3.5) we know:

$$Z = \sum_r e^{-\beta E_r}$$

$$\frac{\partial \ln(Z)}{\partial \beta} = \frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{1}{Z} \sum_r E_r e^{-\beta E_r} = -\sum_r E_r P_r = -\overline{E_r} = -E \quad (3.9)$$

$$\frac{\partial \ln(Z)}{\partial V} = \frac{1}{Z} \frac{\partial Z}{\partial V} = -\frac{\beta}{Z} \sum_r \frac{\partial E_r}{\partial V} e^{-\beta E_r} = -\beta \overline{\frac{\partial E_r}{\partial V}} = \beta P \quad (3.10)$$

$$\frac{\partial \ln(Z)}{\partial N} = \frac{1}{Z} \frac{\partial Z}{\partial N} = -\frac{\beta}{Z} \sum_r \frac{\partial E_r}{\partial N} e^{-\beta E_r} = -\beta \overline{\frac{\partial E_r}{\partial N}} = -\beta \mu \quad (3.11)$$

So we get

$$d(\ln(Z(\beta, V, N))) = -E d\beta + \beta P dV - \beta \mu dN$$

and thus:

$$d(\ln(Z) + \beta E) = \beta(dE + P dV - \mu dN) \stackrel{(2.52a)}{=} \beta T dS = \frac{dS}{k_B}$$

This gives:

$$S = k_B (\ln(Z) + \beta E) + \text{const.} \quad (3.12)$$

To determine the constant, let us consider a quantum mechanical system with energies $E_0 < E_1 < E_2 < \dots$ and $\Delta := E_1 - E_0$. (Here we assume that there is no degeneracy.)

$$\begin{aligned} Z &= e^{-\beta E_0} + e^{-\beta E_1} + \dots \\ &= e^{-\beta E_0} (1 + e^{-\beta \Delta} + \dots) \xrightarrow[\beta \rightarrow \infty]{T \rightarrow 0 \ (k_B T \ll \Delta)} e^{-\beta E_0} \\ \Rightarrow \ln(Z) &\xrightarrow{T \rightarrow 0} -\beta E_0 \end{aligned}$$

Likewise follows:

$$E = \overline{E_r} = E_0 + E_1 e^{-\beta \Delta} + \dots \xrightarrow{T \rightarrow 0} E_0$$

From the third law $S \xrightarrow{T \rightarrow 0} 0$ follows with (3.12):

$$\begin{aligned} \text{const.} = S - k_B (\ln Z + \beta E) &\xrightarrow{T \rightarrow 0} 0 - k_B (-\beta E_0 + \beta E_0) = 0 \\ \Rightarrow \quad &\boxed{E - TS = F = -k_B T \ln Z} \end{aligned} \quad (3.13)$$

ii) *Relation between Y and J* : One proceeds analogously:

$$Y = \sum_r e^{-\beta(E_r - \mu N_r)}$$

$$\frac{\partial \ln(Y)}{\partial \beta} = -\frac{1}{Y} \sum_r (E_r - \mu N_r) e^{-\beta(E_r - \mu N_r)} = -\overline{(E_r - \mu N_r)} = -E + \mu N \quad (3.14)$$

$$\frac{\partial \ln(Y)}{\partial V} = -\frac{\beta}{Y} \sum_r \frac{\partial E_r}{\partial V} e^{-\beta(E_r - \mu N_r)} = -\beta \overline{\left(\frac{\partial E_r}{\partial V} \right)} = \beta P \quad (3.15)$$

$$\frac{\partial \ln(Y)}{\partial \mu} = -\frac{\beta}{Y} \sum_r (-N_r) e^{-\beta(E_r - \mu N_r)} = \beta \overline{N_r} = \beta N \quad (3.16)$$

$$\begin{aligned} d(\ln(Y(\beta, V, \mu))) &= \frac{\partial \ln Y}{\partial \beta} d\beta + \frac{\partial \ln Y}{\partial V} dV + \frac{\partial \ln Y}{\partial \mu} d\mu = \\ &= (-E + \mu N) d\beta + \beta P dV + \beta N d\mu \\ d(\ln(Y(\beta, V, \mu))) + E d\beta - \mu N d\beta - \beta N d\mu &= \beta P dV \\ d(\ln(Y) + \beta E - \beta \mu N) &= \beta (dE + P dV - \mu dN) = \beta T dS = \frac{dS}{k_B} \end{aligned}$$

$$\Rightarrow \quad S = k_B (\ln(Y) + \beta E - \beta \mu N)$$

$$k_B T \ln(Y) = TS - E + \mu N \stackrel{(2.53a)}{=} -J$$

$$\Rightarrow \quad \boxed{J = -k_B T \ln(Y)} \quad (3.17)$$

4 Applications

4.1. Ideal gas

Calculation of $Z(T, V, N)$

Remember:

$$E_r = \sum_{k=1}^{3N} \frac{p_k^2}{2m} = \sum_{k=1}^{3N} \frac{\hbar^2 k^2}{2m} \quad k = \frac{\pi n_k}{L}$$

Accounting for the indistinguishability of particles yields:

$$\begin{aligned} Z &= \sum_r e^{-\beta E_r} \approx \frac{1}{N!} \int dn_1 \dots \int dn_{3N} e^{-\beta E_r} = \\ &= \frac{V^N}{N! \cdot h^{3N}} \int_{-\infty}^{\infty} dp_1 \dots \int_{-\infty}^{\infty} dp_{3N} e^{-\beta \sum_{k=1}^{3N} \frac{p_k^2}{2m}} = \\ &= \frac{1}{N!} \left(\frac{V}{h^3} \int_{-\infty}^{\infty} d^3p \cdot e^{-\beta \frac{\vec{p}^2}{2m}} \right)^N \end{aligned}$$

In the exponent is $-\beta \frac{\vec{p}^2}{2m}$ and with $p = \|\vec{p}\|$ we get:

$$\begin{aligned} Z &= \frac{1}{N!} \left(\frac{V}{h^3} \int_0^{\infty} 4\pi p^2 dp \cdot e^{-\beta \frac{p^2}{2m}} \right)^N = \\ &\stackrel{x = \frac{p}{\sqrt{2mk_B T}}}{=} \frac{1}{N!} \left(\frac{4\pi V}{h^3} (2mk_B T)^{\frac{3}{2}} \underbrace{\int_0^{\infty} x^2 dx e^{-x^2}}_{=\frac{1}{4}\sqrt{\pi}} \right)^N = \\ &= \frac{1}{N!} \left(\frac{V}{h^3} (2\pi mk_B T)^{\frac{3}{2}} \right)^N \end{aligned}$$

This can be written as

$$\boxed{Z = \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^N} \quad (4.1)$$

with the De Broglie thermal wave length:

$$\lambda = \frac{h}{\sqrt{2\pi mk_B T}} \quad (4.2)$$

This λ can be interpreted as the quantum mechanical wavelength of a particle with an energy of the order $k_B T$:

$$\frac{\hbar^2}{2m} \left(\frac{2\pi}{\lambda} \right)^2 = \pi k_B T$$

Now introduce the single particle partition function

$$z := \sum_{\vec{p}} e^{-\beta \epsilon(\vec{p})} = \frac{V}{h^3} \int d^3 p \cdot e^{-\beta \frac{\vec{p}^2}{2m}} = \frac{V}{\lambda^3} \quad (4.3)$$

to get:

$$\boxed{Z = \frac{z^N}{N!}} \quad (4.4)$$

Note: This equation holds for any system of *independent* particles.

Equation of state of the ideal gas

From (4.4) follows:

$$\ln(Z) = N \ln(z) - \ln(N!) \stackrel{\text{Stirling's formula}}{\underset{\ln(N!) \approx N \ln(N) - N}{\approx}} N \ln\left(\frac{z}{N}\right) + N \quad (4.5)$$

And hence we get:

$$P = \frac{1}{\beta} \frac{\partial \ln(Z)}{\partial V} = N k_B T \frac{\partial \ln(z)}{\partial V} = N k_B T \frac{\partial \ln(V)}{\partial V} = \frac{N k_B T}{V}$$

$$E = -\frac{\partial \ln(Z)}{\partial \beta} = -N \frac{\partial \ln(z)}{\partial \beta} = N \frac{\partial \ln \lambda^3}{\partial \beta} = \frac{3}{2} N \frac{\partial \ln \beta}{\partial \beta} = \frac{3}{2} N k_B T$$

These are the equations of state:

$$PV = N k_B T \quad \text{thermal equation of state} \quad (4.6)$$

$$E = \frac{3}{2} N k_B T \quad \text{caloric equation of state} \quad (4.7)$$

Note: Equation (4.7) is a particular realization of the “equipartition theorem”, which states:

“Each variable occurring *quadratically* in the Hamiltonian yields a contribution of $\frac{1}{2} k_B T$ to the average temperature.”

Calculation of $Y(T, V, \mu)$

From equation (3.8) we get:

$$Y(T, V, N) = \sum_{N=0}^{\infty} Z(T, V, N) e^{\beta \mu N}$$

For the ideal gas follows:

$$Y(T, V, N) = \sum_{N=0}^{\infty} \frac{1}{N!} \frac{V^N}{\lambda^{3N}} e^{\beta \mu N}$$

So with

$$y = \frac{V}{\lambda^3} e^{\beta \mu} \quad (4.8)$$

we get:

$$Y(T, V, N) = \sum_{N=0}^{\infty} \frac{y^N}{N!} = e^y \quad (4.9)$$

We obtain now the various equations of state:

$$N = \frac{1}{\beta} \frac{\partial \ln(Y)}{\partial \mu} = \frac{1}{\beta} \frac{\partial y}{\partial \mu} = \frac{\beta}{\beta} y = \ln(Y) \quad (4.10)$$

$$P = \frac{1}{\beta} \frac{\partial \ln(Y)}{\partial V} = \frac{1}{\beta} \frac{\partial y}{\partial V} = k_B T \frac{e^{\beta \mu}}{\lambda^3} = k_B T \frac{y}{V} = \frac{N k_B T}{V} \quad (4.11)$$

$$E = -\frac{\partial \ln(Y)}{\partial \beta} + \mu N = -\frac{\partial y}{\partial \beta} + \mu N = -\lambda^3 y \frac{\partial \lambda^{-3}}{\partial \beta} = \frac{3}{2} N k_B T \quad (4.12)$$

4.2. Interacting gas

We would like to address the situation, where particles in a dilute system can interact with each other. We start from the Hamiltonian:

$$\hat{H} = \sum_{\nu=1}^N \frac{\hat{p}_{\nu}^2}{2m} + \sum_{\nu=2}^N \sum_{\nu'=1}^{\nu-1} \hat{w}(|\vec{r}_{\nu} - \vec{r}_{\nu'}|) \quad (4.13)$$

In general, this problem can not be solved *exactly*, but it can be solved approximately if the system is *dilute*, i.e.:

$$v = \frac{V}{N} \gg d^3 \quad (4.14)$$

Here d is the average size of a particle.

– The starting point is:

$$\begin{aligned} Y(T, V, \mu) &= \sum_{N=0}^{\infty} Z(N) e^{\beta \mu N} = 1 + \underbrace{Z(1) e^{\beta \mu} + Z(2) e^{2\beta \mu} + \dots}_{=x} \\ \ln(Y) &\stackrel{\ln(1+x) \approx x - \frac{x^2}{2}}{\approx} Z(1) e^{\beta \mu} + \underbrace{\left(Z(2) - \frac{(Z(1))^2}{2} \right)}_{=: Z_2} e^{2\beta \mu} \end{aligned} \quad (4.15)$$

For $\hat{w} = 0$ holds $Z_2 = 0$ due to:

$$Z(N) = \frac{(Z(1))^N}{N!}$$

This also holds true for the next terms, so for $\hat{w} = 0$ we get (cf. (4.9)):

$$\ln(Y) = Z(1) e^{\beta\mu}$$

One can consider Z_2 as a *correction* to the ideal gas result, when only two-body interactions are considered.

– Let us then introduce:

$$Z_1 = Z(1) = \frac{V}{\lambda^3} \quad Z_2 = Z(2) - \frac{(Z(1))^2}{2} \quad (4.16)$$

So we get:

$$\begin{aligned} P &\stackrel{(2.56)}{=} \frac{k_B T \ln(Y)}{V} \stackrel{(3.17)}{=} \frac{k_B T}{V} (Z_1 e^{\beta\mu} + Z_2 e^{2\beta\mu} + \dots) \\ N &= \frac{1}{\beta} \frac{\partial \ln(Y)}{\partial \mu} = Z_1 e^{\beta\mu} + 2Z_2 e^{2\beta\mu} + \dots \end{aligned}$$

Solving the second equation for $e^{\beta\mu}$ gives:

$$e^{\beta\mu} = \frac{N}{Z_1} - 2 \frac{Z_2}{Z_1} e^{2\beta\mu} + \dots$$

Inserting this into the first equation yields:

$$\begin{aligned} \frac{PV}{k_B T} = \ln Y &\approx N - Z_2 e^{2\beta\mu} \stackrel{e^{\beta\mu} \approx \frac{N}{Z_1}}{\approx} N - Z_2 \left(\frac{N}{Z_1} \right)^2 = \\ &= N \left(1 - \frac{Z_2 V}{Z_1^2} \frac{N}{V} \right) =: N \left(1 + \frac{B(T)}{v} \right) \end{aligned} \quad (4.17)$$

The $B(T)$ here is known as *virial coefficient*.

$$\begin{aligned} Z_2 &= \frac{1}{2! \cdot h^6} \int_V d^3 r_1 \int_V d^3 r_2 \int d^3 p_1 \int d^3 p_2 e^{-\beta(\frac{1}{2}(p_1^2 + p_2^2) + w(r_{12}))} = \\ &= \frac{1}{2\lambda^6} \int_V d^3 r_1 \int_V d^3 r_2 e^{-\beta w(r_1 - r_2)} \end{aligned}$$

$$\begin{aligned} B(T) &:= -\frac{Z_2 V}{Z_1^2} = -\frac{V}{\left(\frac{V}{\lambda^3}\right)^2} \left(Z(2) - \frac{(Z(1))^2}{2} \right) = \\ &= -V \left(\frac{\lambda^6}{V^2} \cdot \frac{1}{2\lambda^6} \int_V d^3 r_1 \int_V d^3 r_2 e^{-\beta w(r_{12})} - \frac{1}{2} \right) = \\ &= -\frac{1}{2V} \left(\int_V d^3 r_1 \int_V d^3 r_2 e^{-\beta w(r_{12})} - V^2 \right) = \\ &= -\frac{1}{2V} \left(\int_V d^3 r_1 \int_V d^3 r_2 (e^{-\beta w(r_{12})} - 1) \right) \end{aligned}$$

Now change the variables:

$$\vec{R} := \frac{\vec{r}_1 + \vec{r}_2}{2} \quad \vec{r} := \vec{r}_2 - \vec{r}_1$$

Then follows:

$$\begin{aligned} B(T) &= -\frac{1}{2V} \left(\int_V d^3R \int d^3r \left(e^{-\beta w(r)} - 1 \right) \right) = \\ &= -\frac{1}{2} \left(\int 4\pi r^2 dr \left(e^{-\beta w(r)} - 1 \right) \right) = -2\pi \int dr \cdot r^2 \left(e^{-\beta w(r)} - 1 \right) \end{aligned} \quad (4.18)$$

Note: $B(T)$ is a macroscopically measurable quantity, which yields information about $w(r)$.

- Now make the assumption of an “hard core” potential:

$$w(r) = \begin{cases} \infty & r < d \\ < 0 & r \gtrsim d \\ \approx 0 & r \gg d \end{cases}$$

TODO: Plot $w(r)$

Here d is the order of the particle size. For $r > d$ the potential is attractive due to (instantaneous) dipole interactions. The sketch mimics more realistic potentials (e.g. Lennard-Jones potential).

Due to the attractive potential, the gas exhibits a transition to a liquid phase at low temperatures. The transition occurs, when $k_B T$ is about as large as the attractive potential $|w|$. In the following we only consider:

$$\frac{|w|}{k_B T} \ll 1 \quad \text{for } r \gtrsim d$$

Then holds:

$$e^{-\beta w} - 1 \approx \begin{cases} -1 & r < d \\ -\beta w & r > d \end{cases}$$

$$\Rightarrow \quad B(T) \approx 2\pi \int_0^d dr \cdot r^2 + 2\pi\beta \int_d^\infty dr \cdot r^2 w(r) := b - \frac{a}{k_B T}$$

$$b = \frac{2\pi}{3} d^3 > 0 \quad \quad a = -2\pi \int_d^\infty dr \cdot r^2 w(r) > 0 \quad (4.19)$$

Thus we find:

$$\begin{aligned} P &= \frac{k_B T}{v} \left(1 + \frac{B(T)}{v} \right) = \frac{k_B T}{v} \left(1 + \frac{b - \frac{a}{k_B T}}{v} \right) = \frac{k_B T}{v} \left(1 + \frac{b}{v} \right) - \frac{a}{v^2} \approx \\ &\approx \frac{k_B T}{v} \cdot \frac{1}{1 - \frac{b}{v}} - \frac{a}{v^2} \end{aligned}$$

This yields the equation of the *van-der-Waals gas*:

$$\boxed{P = -\frac{a}{v^2} + \frac{k_B T}{v - b}} \quad (4.20)$$

This gives the corrections to the ideal gas:

- The atoms have finite size, so the accessible volume per atom is $v - b$.
- The attractive interactions for $r > d$ diminish the pressure by $\frac{a}{v^2}$.

Note: The necessary condition for this to work is $B(T) \ll v$, i.e. $\frac{b}{v} \ll 1$ and $\frac{a}{vk_B T} \ll 1$.

Note: At high enough temperatures is $d^3 \gg \lambda^3$ and thus $\frac{d^3}{v} \ll 1$ implies also $\frac{\lambda^3}{v} \ll 1$. But for the ideal gas is $e^{\beta\mu} \approx \frac{\lambda^3}{v}$ and thus follows $\mu < 0$.

Note: $E = -\frac{\partial \ln(Y)}{\partial \beta} + \mu N$ is obtained to be:

$$E = N \left(\frac{3}{2} k_B T - \frac{a}{v} \right) \quad (4.21)$$

4.3. Indistinguishability of particles

So far we have taken into account the indistinguishability of particles by adding the prefactor $\frac{1}{N!}$ in our sum over the microstates. However, indistinguishability has a more profound consequence on the form of a many body wave function. Specifically, let us consider the N -particle Hamiltonian \hat{H}_N

$$\hat{H}_N |\psi\rangle = E_N |\psi\rangle \quad (4.22)$$

and a complete orthonormal single-particle basis $\{|\nu\rangle\}$.

In general any N -particle state lies in the N -th tensor product and thus can be written as a linear combination of products of N single-particle states:

$$|\psi\rangle = \sum_{\nu_1, \dots, \nu_N} A_{\nu_1 \nu_2 \dots \nu_N} |\nu_1\rangle \otimes |\nu_2\rangle \otimes \dots \otimes |\nu_N\rangle \quad (4.23)$$

Example: $N = 2$

$$|\psi\rangle = \sum_{\nu_1, \nu_2} A_{\nu_1 \nu_2} |\nu_1\rangle \otimes |\nu_2\rangle \quad (4.24)$$

Let us introduce the permutation operator \hat{P}_{12} by:

$$\hat{P}_{12} |\psi\rangle = \hat{P}_{12} \sum_{\nu_1, \nu_2} A_{\nu_1 \nu_2} |\nu_1\rangle \otimes |\nu_2\rangle = \sum_{\nu_1, \nu_2} A_{\nu_2 \nu_1} |\nu_1\rangle \otimes |\nu_2\rangle \quad (4.25)$$

Because of the indistinguishability of the particles, $|\psi\rangle$ and $\hat{P}_{12} |\psi\rangle$ must represent the same state and thus follows:

$$\hat{P}_{12} |\psi\rangle = \lambda |\psi\rangle \quad (4.26)$$

Moreover holds:

$$\begin{aligned} \hat{P}_{12}^2 = \mathbb{1} & \quad \Rightarrow \quad \lambda^2 = 1 & \quad \Rightarrow \quad \lambda = \pm 1 \\ & \quad \Rightarrow \quad A_{\nu_1 \nu_2} = \pm A_{\nu_2 \nu_1} \end{aligned} \quad (4.27)$$

Let us for simplicity restrict the sum \sum_{ν_1, ν_2} to two states $|a\rangle$ and $|b\rangle$ only.

– $\lambda = 1$:

$$|\psi\rangle = A_{aa} |a\rangle \otimes |a\rangle + A_{bb} |b\rangle \otimes |b\rangle + A_{ab} (|a\rangle |b\rangle + |b\rangle |a\rangle) \quad (4.28)$$

The state is symmetric upon exchange of the particles 1 and 2.

– $\lambda = -1$: From equation (4.27) follows:

$$A_{\nu_j \nu_j} = 0 \quad (4.29a)$$

$$|\psi\rangle = A_{ab} (|a\rangle |b\rangle - |b\rangle |a\rangle) \quad (4.29b)$$

The state is antisymmetric upon exchange of the particles 1 and 2.

So there are only symmetrized/antisymmetrized basis states. In general this property is closely related to the spin of the considered particles (quasi-particles):

$$\begin{aligned} \text{integer spin} &\leftrightarrow \text{symmetric} \leftrightarrow \text{bosons} \\ \text{half-integer spin} &\leftrightarrow \text{antisymmetric} \leftrightarrow \text{fermions} \end{aligned} \quad (4.30)$$

(cf. W. PAULI: Physical Review 58, 716 (1940))

We have seen in equation (4.23):

$$|\psi\rangle = \sum_{\nu_1, \dots, \nu_N} A_{\nu_1 \nu_2 \dots \nu_N} |\nu_1\rangle \otimes |\nu_2\rangle \otimes \dots \otimes |\nu_N\rangle$$

Here the symmetry requirements are hidden in A_{ν_1, \dots, ν_N} .

So a physically meaningful N -particle basis is the one, which incorporates symmetrized (antisymmetrized) products of single particle states. This is accomplished by symmetrization (antisymmetrization) operators \hat{S}_+ (\hat{S}_-).

Example

$$|\vec{r}_1 \vec{r}_2 \dots \vec{r}_N\rangle |\psi\rangle = \sum_{\nu_1, \dots, \nu_N} A_{\nu_1, \dots, \nu_N} \psi_{\nu_1}(\vec{r}_1) \dots \psi_{\nu_N}(\vec{r}_N) \quad (4.31)$$

$$\hat{S}_{\pm} \prod_{j=1}^N \psi_{\nu_j}(\vec{r}_j) := \frac{\mathcal{N}}{\sqrt{N!}} \begin{vmatrix} \psi_{\nu_1}(\vec{r}_1) & \dots & \psi_{\nu_1}(\vec{r}_N) \\ \vdots & \ddots & \vdots \\ \psi_{\nu_N}(\vec{r}_1) & \dots & \psi_{\nu_N}(\vec{r}_N) \end{vmatrix}_{\pm} \quad (4.32)$$

Here $|\cdot|_+$ is the permanent (without sign change) and $|\cdot|_-$ the normal determinant (slater determinant). Moreover the normalization factor is:

$$\mathcal{N} = \frac{1}{\prod_{\{\nu_i\}} \sqrt{n_{\nu_i}!}} \quad (4.33)$$

Here n_{ν} is the number of times the state $|\nu\rangle$ appears in the set $\{|\nu_1\rangle, \dots, |\nu_N\rangle\}$. It holds:

$$\begin{aligned} n_{\nu} &\in \{0, 1\} && \text{fermions} \\ n_{\nu} &\in \{0, 1, \dots, N\} && \text{bosons} \end{aligned}$$

Example: $N = 3$

$$\hat{S}_{\pm} |\nu_1 \nu_2 \nu_3\rangle = \frac{\mathcal{N}}{\sqrt{3!}} (|\nu_1 \nu_2 \nu_3\rangle \pm |\nu_1 \nu_3 \nu_2\rangle \pm |\nu_3 \nu_2 \nu_1\rangle \pm |\nu_2 \nu_1 \nu_3\rangle + |\nu_2 \nu_3 \nu_1\rangle + |\nu_3 \nu_1 \nu_2\rangle) \quad (4.34)$$

The generic N -particles state is:

$$|\psi\rangle = \sum_{\nu_1, \dots, \nu_N} B_{\nu_1 \dots \nu_N} \hat{S}_{\pm} \underbrace{|\nu_1, \nu_2, \dots, \nu_N\rangle}_{=|\nu_1\rangle \otimes \dots \otimes |\nu_N\rangle} \quad (4.35)$$

4.3.1. Occupation number representation

The symmetrized/antisymmetrized states are fully characterized by the occupation numbers $\{n_{\nu}\}$, where ν runs over *all* single particle states. To such extent one orders the one-particle states in a sequence $\nu_1 < \nu_2 < \nu_3 < \dots$, where energetically lower states come before energetically higher states:

$$\varepsilon_{\nu_1} \leq \varepsilon_{\nu_2} \leq \dots$$

Introduce the notation, called *occupation number representation*:

$$|n_{\nu_1}, n_{\nu_2}, \dots\rangle \quad (4.36)$$

For an N -particle system must hold:

$$\sum_{\nu} n_{\nu} = N \quad (4.37)$$

Note: Due to the Pauli principle, the number of available states for fermions of bosons is very different.

N	fermion basis $ n_{\nu_1}, n_{\nu_2}, \dots\rangle$	boson basis $ n_{\nu_1}, n_{\nu_2}, \dots\rangle$
0	$ 0, 0, \dots\rangle$	$ 0, 0, \dots\rangle$
1	$ 1, 0, \dots\rangle, 0, 1, 0, \dots\rangle, \dots$	$ 1, 0, \dots\rangle, 0, 1, 0, \dots\rangle, \dots$
2	$ 1, 1, 0, \dots\rangle, 1, 0, 1, 0, \dots\rangle, \dots, 0, 1, 1, 0, \dots\rangle, \dots$	$ 2, 0, \dots\rangle, 0, 2, 0, \dots\rangle, \dots, 1, 1, 0, \dots\rangle, \dots$
\vdots	\vdots	\vdots

So the art of particles (bosons/fermions) determines their statistics, i.e. the number of possibilities one has to distribute particles on the occupation number basis states.

In particular, a microstate r with N_r particles is defined upon assigning (n_{ν}) :

$$r = (n_{\nu_1}, n_{\nu_2}, \dots) = (n_{\nu}) \quad (4.38)$$

If the particle number is not fixed, we write:

$$r = (r', N_r) = ((n_{\nu}), N_r) \quad (4.39)$$

A) *Particle number fixed:*

a) One considers only microstates r with $N_r = N$ and evaluates:

$$Z(T, x, N) = \sum_r e^{-\beta E_r(x, N)} \quad (4.40)$$

b) Or one does not constrain the number N_r of the microstate r and evaluates:

$$Y(T, x, \mu) = \sum_r e^{-\beta(E_r(x, N_r) - \mu N_r)} \quad (4.41)$$

Here μ is chosen such that holds:

$$\frac{1}{\beta} \frac{\partial \ln(Y)}{\partial \mu} = \overline{N_r} = N \quad (4.42)$$

B) *Particle number cannot be determined:*

- Particles can be exchanged.
- To this second class belong e.g. systems made of bosonic quasi-particles like photons and phonons. These are excitations and therefore their average number depends on the temperature and cannot be fixed. For these systems, the particle number does not enter the Hamiltonian and therefore the energy E_r of the microstate r is independent on N .

$$\mu = \frac{\partial \overline{E_r}}{\partial N} = 0 \quad (4.43)$$

4.3.2. Independent particles

- For a system of N independent particles holds

$$\hat{H}_N = \sum_{j=1}^N \hat{h}_j \quad \hat{H}_N |\psi\rangle = E_N |\psi\rangle \quad \hat{h}_j |\nu_j\rangle = \varepsilon_j |\nu_j\rangle \quad (4.44)$$

and hence:

$$E_N = \sum_{j=1}^N \varepsilon_j \quad (4.45)$$

In terms of occupation numbers with $N = \sum_{\nu} \bar{n}_{\nu}$:

$$E_N = \sum_{\nu} \bar{n}_{\nu} \varepsilon_{\nu} \quad (4.46a)$$

- For the energy E_r of the microstate $r = (n_{\nu})$ holds:

$$E_r = \sum_{\nu} n_{\nu} \varepsilon_{\nu} \quad (4.46b)$$

Example: Consider a gas of N electrons in a box (with spin included!):

$$(|\nu\rangle) = ((\vec{p}_1, \uparrow), (\vec{p}_1, \downarrow), (\vec{p}_2, \uparrow), (\vec{p}_2, \downarrow), \dots)$$

$$|\vec{p}_1| \leq |\vec{p}_2| \leq |\vec{p}_3| \leq \dots$$

$$\Rightarrow r = (n_{\vec{p}_1}^{\uparrow}, n_{\vec{p}_1}^{\downarrow}, \dots)$$

$$\sum_{\vec{p}, s_z} n_{\vec{p}}^{s_z} = N_r = N$$

Note:

$$E = -\frac{\partial \ln(Y)}{\partial \beta} + \mu N \stackrel{\text{independent}}{\text{particles}} \sum_{\nu} \varepsilon_{\nu} \bar{n}_{\nu} \quad (4.47a)$$

$$N = \frac{1}{\beta} \frac{\partial \ln Y}{\partial \mu} = \sum_{\nu} \bar{n}_{\nu} \quad (4.47b)$$

Here \bar{n}_{ν} is the *average occupation number* of the level ν . So we have:

$$P = -\frac{\partial E_r}{\partial V} = -\sum_{\nu} \frac{\partial \varepsilon_{\nu}}{\partial V} \bar{n}_{\nu} \quad (4.48)$$

Example: Consider an ideal gas with:

$$\varepsilon_{\nu} = \varepsilon_p = \frac{p^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2mL^2}$$

From $L^3 = V$ we get:

$$\frac{\partial \varepsilon_{\nu}}{\partial V} = -\frac{2}{3} \frac{\varepsilon_p}{V}$$

$$\Rightarrow \boxed{P = \frac{2}{3} \frac{E}{V}} \quad (4.49)$$

4.3.3. Fermi-Dirac statistics

Consider a gas of independent fermions and look for \bar{n}_{ν} . To this extent we start from the grandcanonical partition function $Y(T, x, \mu)$.

$$\begin{aligned} Y(T, x, \mu) &= \sum_r e^{-\beta(E_r - \mu N_r)} \stackrel{\text{fermions}}{=} \sum_{n_1 \in \{0,1\}} \sum_{n_2 \in \{0,1\}} \dots e^{-\beta \sum_{\nu} (n_{\nu} \varepsilon_{\nu} - \mu n_{\nu})} = \\ &= \prod_{\nu} \left(\sum_{n_{\nu} \in \{0,1\}} e^{-\beta(\varepsilon_{\nu} - \mu) n_{\nu}} \right) = \prod_{\nu} \left(1 + e^{-\beta(\varepsilon_{\nu} - \mu)} \right) \end{aligned} \quad (4.50)$$

The average occupation number is:

$$\begin{aligned} \bar{n}_{\nu} &= \frac{1}{Y} \sum_{n_1 \in \{0,1\}} \sum_{n_2 \in \{0,1\}} \dots n_{\nu} e^{-\beta \sum_{\nu'} (\varepsilon_{\nu'} - \mu) n_{\nu'}} = \\ &= 1 \cdot e^{-\beta(\varepsilon_{\nu} - \mu)} \frac{1}{Y} \prod_{\nu \neq \nu'} \left(1 + e^{-\beta(\varepsilon_{\nu'} - \mu)} \right) = \\ &= \frac{e^{-\beta(\varepsilon_{\nu} - \mu)}}{1 + e^{-\beta(\varepsilon_{\nu} - \mu)}} \end{aligned}$$

This gives the *Fermi-Dirac* function:

$$\bar{n}_\nu := f(\varepsilon_\nu) := \frac{1}{e^{\beta(\varepsilon_\nu - \mu)} + 1} \quad (4.51)$$

The Fermi-Dirac function f yields the occupation number of a level ν for a system of independent fermions.

Note: For $T = 0$ holds:

$$\lim_{T \rightarrow 0} f(\varepsilon_\nu) = \Theta(\mu - \varepsilon_\nu) = \begin{cases} 1 & \varepsilon_\nu < \mu \\ 0 & \varepsilon_\nu > \mu \end{cases}$$

TODO: Abb12: ground state of $s = \frac{1}{2}$ particles

This ground state configuration is known as the “*Fermi sea*”.

- For fixed N , the chemical potential $\mu = \mu(T, v)$ is a function of T and v . Define the *Fermi energy* as:

$$\varepsilon_F := \lim_{T \rightarrow 0} \mu(T)$$

- For a metal the Fermi energy ε_F is the energy of the *highest occupied level* at zero temperature:

$$\varepsilon_F = \mu(T = 0, v)$$

- For a gapped structure (semi conductor), $\varepsilon_F = \mu(T = 0)$ lies in the middle of the gap.

TODO: Abb13; semi conductor

- The Fermi momentum p_F is defined by:

$$\varepsilon_F = \frac{p_F^2}{2m}$$

- The Fermi wave vector \vec{k}_F is defined by:

$$\vec{p}_F = \hbar \vec{k}_F$$

The Fermi momentum p_F can be determined from:

$$\begin{aligned} N = \sum_\nu n_\nu &= \underbrace{(2s+1)}_{\text{spin degeneracy}} \sum_{\|\vec{p}\| \leq p_F} 1 \approx \frac{V(2s+1)}{h^3} \cdot \frac{4}{3} \pi p_F^3 \stackrel{s=\frac{1}{2}}{=} \frac{V}{h^3} \cdot \frac{8}{3} \pi p_F^3 \\ p_F &= \left(\frac{3}{\pi} \frac{N}{V} \frac{2}{(2s+1)} \right)^{\frac{1}{3}} \frac{h}{2} \stackrel{s=\frac{1}{2}}{=} (3\pi^2)^{\frac{1}{3}} \frac{\hbar}{v^{\frac{1}{3}}} \\ \varepsilon_F &= \left(\frac{3}{\pi} \frac{N}{V} \frac{2}{(2s+1)} \right)^{\frac{2}{3}} \frac{h^2}{4} \stackrel{s=\frac{1}{2}}{=} (3\pi^2)^{\frac{2}{3}} \frac{\hbar^2}{2mv^{\frac{2}{3}}} \end{aligned} \quad (4.52)$$

Note: The Fermi momentum $p_F \sim n^{\frac{1}{3}}$ and the Fermi energy $\varepsilon_F \sim n^{\frac{2}{3}}$. Likewise follows:

$$\begin{aligned} E &= \sum_\nu \varepsilon_\nu \bar{n}_\nu = \underbrace{(2s+1)}_{\text{spin degeneracy}} \sum_{|\vec{p}| \leq p_F} \frac{p^2}{2m} \approx \frac{V}{mh^3} \int_{|\vec{p}| \leq p_F} p^2 d^3p = \\ &= \frac{4\pi V}{mh^3} \int_{p \leq p_F} p^4 dp = \frac{4\pi V}{5mh^3} p_F^5 = \frac{3}{5} \cdot \underbrace{\frac{8\pi V}{3h^3} p_F^3}_{=N} \cdot \frac{1}{2m} p_F^2 = \frac{3}{5} N \varepsilon_F \end{aligned} \quad (4.53)$$

For $T \neq 0$ we have:

TODO: Abb14

4.3.4. Bose-Einstein statistics

Partition function: For bosons any $n_\nu \in \mathbb{N}_0$ is possible. So for a system of independent bosons holds:

$$Y(T, x, \mu) = \sum_r e^{-\beta(E_r - \mu N_r)} = \prod_\nu \left(\sum_{n_\nu} e^{-\beta(\varepsilon_\nu - \mu)n_\nu} \right) \stackrel{\sum_{n=0}^\infty x^n = \frac{1}{1-x}}{=} \prod_\nu \frac{1}{1 - e^{-\beta(\varepsilon_\nu - \mu)}} \quad (4.54)$$

Average occupation number:

$$\begin{aligned} \bar{n}_\nu &= \frac{1}{Y} \sum_r n_\nu e^{-\beta(E_r - \mu N_r)} = \frac{1}{Y} \sum_{n_\nu} n_\nu e^{-\beta(\varepsilon_\nu - \mu)n_\nu} \prod_{\nu \neq \nu'} \frac{1}{1 - e^{-\beta(\varepsilon_{\nu'} - \mu)}} = \\ &= \sum_{n_\nu} \frac{\frac{1}{\beta} \left(\frac{\partial}{\partial \mu} e^{-\beta(\varepsilon_\nu - \mu)n_\nu} \right)}{\left(1 - e^{-\beta(\varepsilon_\nu - \mu)} \right)^{-1}} = \left(1 - e^{-\beta(\varepsilon_\nu - \mu)} \right) \frac{1}{\beta} \frac{\partial}{\partial \mu} \frac{1}{1 - e^{-\beta(\varepsilon_\nu - \mu)}} = \\ &= \frac{e^{-\beta(\varepsilon_\nu - \mu)}}{1 - e^{-\beta(\varepsilon_\nu - \mu)}} \end{aligned}$$

So we get the *Bose-Einstein function*:

$$\boxed{\bar{n}_\nu = \frac{1}{e^{\beta(\varepsilon_\nu - \mu)} - 1}} \quad (4.55)$$

Let us call (4.55) from now on $n_{\text{BE}}(\varepsilon)$.

Note: $n_{\text{BE}}(\varepsilon_\nu) \xrightarrow{\varepsilon_\nu \rightarrow \mu} \infty$

If one also imposes $n_{\text{BE}}(\varepsilon_\nu) \geq 0$, then

$$\varepsilon_\nu - \mu > 0 \quad (4.56)$$

must hold for all ν .

Note: For bosons with $\mu = 0$ (photons, phonons), the Bose-Einstein function reduces to the *Planck function*:

$$n_{\text{BE}}(\varepsilon) \rightarrow \frac{1}{e^{\beta\varepsilon} - 1}$$

4.3.5. Quantum mechanical corrections

We found that for fermions/bosons holds:

$$\bar{n}_\nu = \frac{1}{e^{\beta(\varepsilon_\nu - \mu)} \pm 1}$$

For

$$e^{\beta(\varepsilon_\nu - \mu)} \gg 1$$

we clearly recover the *Boltzmann function*:

$$\bar{n}_\nu = e^{-\beta(\varepsilon_\nu - \mu)}$$

Due to $\varepsilon_1 \leq \varepsilon_2 \leq \dots$ for this

$$e^{\beta(\varepsilon_1 - \mu)} \gg 1 \quad (4.57a)$$

is enough. In this case the semi-classical approximation is valid.

Specific case of the ideal gas

We have $\varepsilon_1 = 0$ and so (4.57a) becomes $e^{\beta\mu} \ll 1$. For the ideal gas holds:

$$e^{\beta\mu} = \frac{\lambda^3}{v}$$

So we get as condition for validity of semi-classical description:

$$\boxed{\lambda^3 \ll v} \quad (4.57b)$$

The higher T , the lower is λ and the better for validity of the semi-classical derivation.

Now we would like to investigate the leading order correction in $e^{\beta\mu}$ to the quantum statistics. Let us then start from $\ln(Y)$:

$$\ln(Y) = \pm \ln \left(\prod_{\nu} \left(1 \pm e^{-\beta(\varepsilon_{\nu}-\mu)} \right) \right) = \pm \sum_{\nu} \ln \left(1 \pm e^{-\beta(\varepsilon_{\nu}-\mu)} \right)$$

We expand in powers of $e^{-\beta(\varepsilon_{\nu}-\mu)}$ and remember $\ln(1 \pm x) = \pm x - \frac{x^2}{2} + \mathcal{O}(x^3)$. We then have:

$$\ln(Y) = \sum_{\nu} e^{-\beta(\varepsilon_{\nu}-\mu)} \mp \frac{1}{2} e^{-2\beta(\varepsilon_{\nu}-\mu)} + \dots$$

This shows that to the leading order no influence of the statistics is seen.

Lowest order

$$\begin{aligned} \ln(Y) &= \sum_{\nu} e^{-\beta(\varepsilon_{\nu}-\mu)} \\ \Rightarrow \quad N &= \frac{1}{\beta} \frac{\partial \ln(Y)}{\partial \mu} = \frac{1}{2s+1} \frac{\lambda^3}{v} e^{\beta\mu} = \ln(Y) \stackrel{J=-k_B T \ln(Y)=-PV}{=} \frac{PV}{k_B T} \\ &\Rightarrow \quad e^{\beta\mu} = \frac{1}{2s+1} \frac{\lambda^3}{V} \end{aligned}$$

Next leading order

$$\begin{aligned} N &= \frac{1}{\beta} \frac{\partial \ln(Y)}{\partial \mu} = \sum_{\nu} \left(e^{-\beta(\varepsilon_{\nu}-\mu)} \mp e^{-2\beta(\varepsilon_{\nu}-\mu)} \right) = \\ &= \ln(Y) \mp \frac{1}{2} \sum_{\nu} e^{-2\beta(\varepsilon_{\nu}-\mu)} = \frac{PV}{k_B T} \mp \frac{1}{2} \sum_{\nu} e^{-2\beta(\varepsilon_{\nu}-\mu)} \end{aligned}$$

So the equation of state is:

$$PV = Nk_B T \mp \frac{1}{2\beta} \sum_{\nu} e^{-2\beta(\varepsilon_{\nu}-\mu)}$$

Ideal gas

$$\begin{aligned} \sum_{\nu} e^{-2\beta(\varepsilon_{\nu}-\mu)} &= \frac{2s+1}{2^{\frac{3}{2}}} \frac{V}{\lambda^3} e^{2\beta\mu} = \frac{2s+1}{2^{\frac{3}{2}}} \frac{V}{\lambda^3} \left(\frac{1}{2s+1} \frac{\lambda^3}{V} \right)^2 = \\ &= \frac{\lambda^3}{2^{\frac{3}{2}} \cdot (2s+1) V} \\ \Rightarrow \quad &\boxed{PV = NK_B T \left(1 + \frac{B_{\text{qm}}(T)}{v} \right)} \end{aligned}$$

$$B_{\text{qm}}(T) = \pm \frac{\lambda^3}{2^{\frac{5}{2}} \cdot (2s+1)}$$

$B_{\text{qm}}(T)$ are the quantum mechanical corrections. For the energy follows:

$$E = \frac{3}{2} N k_B T \left(1 + \frac{B_{\text{qm}}(T)}{v} \right)$$

4.4. Density of states

Let us calculate the average energy and particle number of an ensemble of independent particles.

$$\begin{aligned} E(T, v, \mu) &= \sum_{\nu} \varepsilon_{\nu} \bar{n}_{\nu}(\varepsilon_{\nu}) \\ N(T, v, \mu) &= \sum_{\nu} \bar{n}_{\nu}(\varepsilon_{\nu}) \end{aligned}$$

We introduce the *density of states*:

$$\mathcal{D}(\varepsilon) = \sum_{\nu} \delta(\varepsilon_{\nu} - \varepsilon) \quad (4.58)$$

For the ideal gas holds:

$$\begin{aligned} \mathcal{D}(\varepsilon) &= \sum_{s_z, \vec{p}} \delta(\varepsilon_p - \varepsilon) = (2s+1) \frac{V}{h^3} \int d^3p \delta\left(\frac{p^2}{2m} - \varepsilon\right) = \\ &= (2s+1) \frac{4\pi V}{h^3} \int p^2 dp \delta\left(\frac{p^2}{2m} - \varepsilon\right) = \\ &\stackrel{x=\frac{p^2}{2m}}{=} \stackrel{2pdp=2mdx}{=} (2s+1) \frac{4\pi m V}{h^3} \int \sqrt{2mx} dx \delta(x - \varepsilon) = \\ &= (2s+1) (2\pi V) \frac{(2m)^{\frac{3}{2}}}{h^3} \sqrt{\varepsilon} = (2s+1) \frac{V}{4\pi^2 \hbar^3} (2m)^{\frac{3}{2}} \sqrt{\varepsilon} = \\ &= \frac{3N}{2} \cdot \underbrace{\frac{(2s+1)(2m)^{\frac{3}{2}} v}{6\pi^2 \hbar^3}}_{=\varepsilon_F^{-\frac{3}{2}}} \sqrt{\varepsilon} = \frac{3}{2} \cdot \frac{N}{\varepsilon_F} \cdot \sqrt{\frac{\varepsilon}{\varepsilon_F}} \sim \sqrt{\varepsilon} \end{aligned}$$

The density of states per particle is:

$$d(\varepsilon) := \frac{\mathcal{D}(\varepsilon)}{N} = \frac{3}{2\varepsilon_F} \sqrt{\frac{\varepsilon}{\varepsilon_F}} \quad (4.59)$$

It holds:

$$\begin{aligned} E(T, V, \mu) &= \int_0^\infty d\varepsilon \cdot \varepsilon \mathcal{D}(\varepsilon) n(\varepsilon) \\ N(T, V, \mu) &= \int_0^\infty d\varepsilon \cdot \mathcal{D}(\varepsilon) n(\varepsilon) \end{aligned}$$

4.5. Independent Fermi systems

Sommerfeld expansion

It is an approximation, which enables us to calculate integrals of the form

$$\int d\varepsilon g(\varepsilon) f(\varepsilon)$$

with the Fermi-Dirac function $f(\varepsilon)$ for a system with density of states per particle $d(\varepsilon)$.

1. Step: Introduce:

$$\eta(x) \stackrel{x=\beta(\varepsilon-\mu)}{=} f(\varepsilon) - \Theta(\mu - \varepsilon) = \frac{1}{e^x + 1} - \Theta(-x)$$

TODO: Plot $(e^x + 1)^{-1}$, $\Theta(-x)$, $\eta(x)$

2. Step: Consider:

$$\int_0^\infty d\varepsilon g(\varepsilon) f(\varepsilon) = \int_0^\mu d\varepsilon g(\varepsilon) + \int_0^\infty d\varepsilon g(\varepsilon) \underbrace{(f(\varepsilon) - \Theta(\mu - \varepsilon))}_{=\eta(x)}$$

We expand $g(\varepsilon)$ around μ in the second integral

$$g(\varepsilon) = g(\mu) + g'(\mu)(\varepsilon - \mu) + \frac{g''(\mu)}{2}(\varepsilon - \mu)^2 + \dots$$

to get:

$$\begin{aligned} \int_0^\infty d\varepsilon g(\varepsilon) f(\varepsilon) &\stackrel{x=\beta(\varepsilon-\mu)}{=} \int_0^\mu d\varepsilon g(\varepsilon) + \frac{1}{\beta} \int_{-\beta\mu}^\infty dx \left(g(\mu) + g'(\mu) \frac{x}{\beta} + \dots \right) \eta(x) \approx \\ &\approx \int_0^\mu d\varepsilon g(\varepsilon) + \frac{1}{\beta} \int_{-\infty}^\infty dx \left(g(\mu) + g'(\mu) \frac{x}{\beta} + \dots \right) \eta(x) = \\ &\stackrel{\eta(x) \text{ is odd}}{=} \int_0^\mu d\varepsilon g(\varepsilon) + \frac{g'(\mu)}{\beta^2} \underbrace{\int_{-\infty}^\infty dx \cdot x \eta(x)}_{=2 \int_0^\infty dx \frac{x}{e^x + 1} = \frac{\pi^2}{6}} + \mathcal{O}\left(\left(\frac{T}{\varepsilon_F}\right)^4\right) \approx \\ &\approx \int_0^\mu d\varepsilon g(\varepsilon) + \frac{\pi^2}{6} (k_B T)^2 g'(\mu) \end{aligned} \quad (4.60)$$

3. Step: We still must consider the dependence $\mu = \mu(T)$.

$$\int_0^\mu d\varepsilon g(\varepsilon) = \int_0^{\varepsilon_F} d\varepsilon g(\varepsilon) + \int_{\varepsilon_F}^\mu d\varepsilon g(\varepsilon) \quad (4.61)$$

Recall now $N = \int d\varepsilon \mathcal{D}(\varepsilon) f(\varepsilon)$ to get:

$$\begin{aligned} 1 &= \int d\varepsilon d(\varepsilon) f(\varepsilon) \\ \text{at } T=0 \Rightarrow \quad 1 &= \int_0^{\varepsilon_F} d\varepsilon d(\varepsilon) \end{aligned}$$

Applying (4.60) this gives:

$$\begin{aligned} 1 &= \frac{1}{N} \int_0^\infty d\varepsilon \mathcal{D}(\varepsilon) f(\varepsilon) \approx \int_0^\mu d\varepsilon d(\varepsilon) + \frac{\pi^2}{6} (k_B T)^2 d'(\mu) \\ &\quad \int_{\varepsilon_F}^\mu d\varepsilon d(\varepsilon) \stackrel{\text{mean value theorem}}{\underset{\tilde{\varepsilon} \in [\varepsilon_F, \mu]}{=}} (\mu - \varepsilon_F) d(\tilde{\varepsilon}) \\ 1 &= \underbrace{\int_0^{\varepsilon_F} d\varepsilon d(\varepsilon)}_{=1} + (\mu - \varepsilon_F) d(\tilde{\varepsilon}) + \frac{\pi^2}{6} (k_B T)^2 d'(\mu) \\ \Rightarrow \quad \mu - \varepsilon_F &= -\frac{\pi^2}{6} (k_B T)^2 \frac{d'(\mu)}{d(\tilde{\varepsilon})} \approx -\frac{\pi^2}{6} (k_B T)^2 \frac{d'(\varepsilon_F)}{d(\varepsilon_F)} + o\left(\left(\frac{T}{\varepsilon_F}\right)^2\right) \end{aligned} \quad (4.62)$$

With this we get:

$$\mu = \varepsilon_F + \mathcal{O}(T^2) \quad (4.63)$$

For the ideal gas follows:

$$\mu = \varepsilon_F \left(1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 \right)$$

So the expansion implies on one side $k_B T \ll \varepsilon_F$.

For metals holds $T_F = \frac{\varepsilon_F}{k_B} \approx 10^4$ K and so even at room temperature $\mu(T) \approx \varepsilon_F$ is a very good approximation.

4. Step: Hence we can approximate:

$$\int_0^\infty d\varepsilon g(\varepsilon) f(\varepsilon) = \int_0^{\varepsilon_F} d\varepsilon g(\varepsilon) - \frac{\pi^2}{6} (k_B T)^2 \left(g(\varepsilon_F) \frac{d'(\varepsilon_F)}{d(\varepsilon_F)} - g'(\varepsilon_F) \right) \quad (4.64)$$

Application: Specific heat of metals

The energy per particle is:

$$\frac{E}{N} = \int_0^\infty d\varepsilon \underbrace{\varepsilon d(\varepsilon)}_{=g(\varepsilon)} f(\varepsilon) \approx$$

$$\begin{aligned}
&\approx \int_0^{\varepsilon_F} d\varepsilon \cdot \varepsilon d(\varepsilon) - \frac{\pi^2}{6} (k_B T)^2 (\varepsilon_F d'(\varepsilon_F) - d(\varepsilon_F) - \varepsilon_F d'(\varepsilon_F)) = \\
&= \frac{E_0}{N} + \frac{\pi^2}{6} (k_B T)^2 d(\varepsilon_F)
\end{aligned} \tag{4.65}$$

This gives the specific heat:

$$c_V = \frac{\partial}{\partial T} \frac{E}{N} = \underbrace{\frac{\pi^2}{3} k_B^2 d(\varepsilon_F)}_{=:\gamma} T = \gamma T$$

For the ideal gas is the density of states per particle at the Fermi energy is (see 4.59):

$$d(\varepsilon_F) = \frac{3}{2\varepsilon_F}$$

With this the Sommerfeld formula gives:

$$c_V(T) = \frac{\pi^2 k_B^2}{2\varepsilon_F} T = \gamma T$$

TODO: Plot of the real $c_V(T)$ curve

Note: If one includes also the contribution of lattice vibrations (phonons), one gets (see integrated course IIa):

$$c_V(T) = \gamma T + \alpha T^3 \tag{4.66}$$

The T^3 term is due to the phonons, while the T term comes from the electrons.

Metal	$\frac{\gamma_{\text{measured}}}{\gamma}$
Li	2,17
Na	1,21
K	1,23
Rb	1,22
Cs	1,39
Cu	1,39
Ag	1,00
Au	1,13

For other metals, the agreement between the measured and the theoretically predicted γ is not that good, for example:

$$\text{Fe: } \frac{\gamma_{\text{measured}}}{\gamma} = 8,0 \qquad \text{Bi: } \frac{\gamma_{\text{measured}}}{\gamma} = 0,05$$

Fermi pressure

For nonrelativistic free electrons hold:

$$P(T, V, N) = \frac{2}{3} \frac{E(T, V, N)}{V} = \frac{2}{3} \frac{E_0}{V} + \frac{\pi^2}{6} \cdot \frac{N k_B T}{V} \cdot \frac{k_B T}{\varepsilon_F} \tag{4.67}$$

So at $T = 0$ the pressure in a Fermi gas does not vanish! It remains the *Fermi pressure*. For the ideal gas this is (cf. (4.53)):

$$P_{\text{Fermi}} = P(T = 0, V, N) = \frac{2}{3} \cdot \frac{3}{5} \frac{N}{V} \varepsilon_F = \frac{2}{5} \frac{N}{V} \varepsilon_F$$

So the incompressibility of solids and liquids is (partly) a consequence of the Pauli principle.

Note: The Fermi pressure plays an important role in some models of stars: at equilibrium the matter pressure must compensate the gravitational pressure.

4.6. Independent Bose systems

We discuss some consequences of the Bose-Einstein distribution:

- i) Bose-Einstein condensation (BEC)
- ii) Planck's radiation law

4.6.1. Bose-Einstein condensation

The Bose-Einstein distribution favors a situation in which the greatest part of the population is in the ground orbitat at sufficiently low temperatures. The particles in the ground state are called “condensate”.

- First predicted by Einstein for the ideal gas in 1924.
- Closest experimental demonstration: BEC of an ultracold gas of alkali atoms by E. A. Cornell, W. Ketterle and C. E. Wiemann in 1999.

Ground state occupation

The Bose-Einstein distribution is:

$$n_{\text{BE}}(\varepsilon_\nu) = \frac{1}{e^{\beta(\varepsilon_\nu - \mu)} - 1}$$

For a system of N -particles holds for all states ν :

$$0 \leq n_{\text{BE}}(\varepsilon_\nu) \leq N$$

For the ground state with energy ε_1 we have:

$$n_{\text{BE}}(\varepsilon_1) = \bar{n}_g = \frac{1}{e^{\beta(\varepsilon_1 - \mu)} - 1}$$

In the limit of low temperatures this gives:

$$N = \lim_{T \rightarrow 0} n_{\text{BE}}(\varepsilon_1) = \lim_{T \rightarrow 0} \frac{1}{e^{\beta(\varepsilon_1 - \mu)} - 1} = \lim_{T \rightarrow 0} \frac{k_B T}{\varepsilon_1 - \mu}$$

Since N is a large number, $\mu \lesssim \varepsilon_1$ must hold and moreover:

$$0 \lesssim \frac{\varepsilon_1 - \mu}{k_B T}$$

Note: For the ideal gas holds:

$$\varepsilon_1 = \frac{\hbar^2}{2m} \frac{\pi^2}{L^2} \underbrace{(1 + 1 + 1)}_{n_x = n_y = n_z = 1} = \frac{3}{2} \frac{\hbar^2}{m} \cdot \frac{\pi^2}{L^2} \xrightarrow{V \rightarrow \infty} 0$$

This implies $\mu \leq 0$ and $\lim_{T \rightarrow 0} \mu = 0$.

Critical temperature

It seems that something special happens, when the chemical potential approaches the largest allowed value ε_1 .

At which critical temperature does it happen?

To this extend we calculate N :

$$\begin{aligned} N &= \sum_{\nu} \bar{n}_{\nu} \approx \int d\varepsilon \mathcal{D}(\varepsilon) n_{\text{BE}}(\varepsilon) = \int d\varepsilon \mathcal{D}(\varepsilon) \frac{1}{e^{\beta(\varepsilon-\mu)} - 1} = \\ &= \int d\varepsilon \mathcal{D}(\varepsilon) \sum_{l=1}^{\infty} e^{-\beta(\varepsilon-\mu)l} = \sum_{l=1}^{\infty} e^{\beta\mu l} \int d\varepsilon \mathcal{D}(\varepsilon) e^{-\beta\varepsilon l} \end{aligned}$$

Now we need to specify $\mathcal{D}(\varepsilon)$. We account for a zero-point energy:

$$\mathcal{D}(\varepsilon) \rightarrow \mathcal{D}_m(\varepsilon, \varepsilon_g) = A(\varepsilon - \varepsilon_g)^m \Theta(\varepsilon - \varepsilon_g)$$

We thus get:

$$\begin{aligned} N &= \sum_{l=1}^{\infty} e^{\beta\mu l} \int_{\varepsilon_g}^{\infty} d\varepsilon A(\varepsilon - \varepsilon_g)^m e^{-\beta\varepsilon l} = \\ &\stackrel{x=\varepsilon-\varepsilon_g}{=} \sum_{l=1}^{\infty} \frac{e^{\beta(\mu-\varepsilon_g)l}}{l^{m+1}} \underbrace{\int_0^{\infty} dx \cdot A x^m e^{-\beta x}}_{=Z_1(T)} = \\ &= Z_1(T) \sum_{l=1}^{\infty} \frac{e^{\beta(\mu-\varepsilon_g)l}}{l^{m+1}} \end{aligned}$$

Now we introduced the generalized Riemannian zeta function

$$g_{\nu}(z) = \sum_{l=1}^{\infty} \frac{z^l}{l^{\nu}} \quad (4.68)$$

to get:

$$N = Z_1(T) g_{m+1}\left(e^{\beta(\mu-\varepsilon_g)}\right) \quad (4.69)$$

The critical temperature is defined by

$$\mu(T_c) = \varepsilon_g$$

or:

$$\boxed{\frac{N}{Z_1(T_c)} = g_{m+1}\left(e^{\frac{\mu(T_c)-\varepsilon_g}{k_B T_c}}\right) \stackrel{g_{m+1}(1)=\zeta(m+1)}{=} \zeta(m+1)} \quad (4.70)$$

Here ζ is the Riemannian zeta function.

Where do the remaining particles go? We were calculating N as:

$$N = \sum_{\nu} \bar{n}_{\nu} \approx \int d\varepsilon \mathcal{D}(\varepsilon) n_{\text{BE}}(\varepsilon)$$

Here we missed the occupation of the ground state. However, this is only fine, when $\bar{n}_g \ll N$ holds. What happens for $T > T_c$? In general we must include the particles of the ground state:

$$N = \bar{n}_g + \sum_{\nu \neq g} \bar{n}_\nu \approx \bar{n}_g + \int d\varepsilon \mathcal{D}(\varepsilon) n_{\text{BE}}(\varepsilon) = \bar{n}_g + z_1(T) g_{m+1} \left(e^{\beta(\mu - \varepsilon_g)} \right)$$

For $T \leq T_c$ we get:

$$\begin{aligned} N &= \bar{n}_g + Z_1(T) \zeta(m+1) \stackrel{(4.70)}{=} \bar{n}_g + N \frac{Z_1(T)}{Z_1(T_c)} \\ \Rightarrow \quad \bar{n}_g &= N \left(1 - \frac{Z_1(T)}{Z_1(T_c)} \right) \end{aligned} \quad (4.71)$$

For $T > T_c$ we have $\bar{n}_g \ll N$.

TODO: Abb $\mu(T)$, $\bar{n}_g(T)$

\bar{n}_g is the order parameter in the Bose-Einstein condensation.

Critical density

Likewise, one can keep T constant and vary N (at $V = \text{const.}$). We denote by $n_c = \frac{N_c}{V}$ the *critical density*:

$$\frac{N_c}{Z_1(T)} = \zeta(m+1) \quad (4.72)$$

Application: ideal gas

For the ideal gas holds $\mu_c = \varepsilon_g = 0$, $Z_1 = \frac{V}{\lambda^3}$, $\mathcal{D}(\varepsilon) \sim \varepsilon^{\frac{1}{2}}$, i.e. $m = \frac{1}{2}$, and $\lambda = \frac{h}{\sqrt{2\pi m k_B T}}$. With this we obtain for $T \geq T_c$:

$$N(T, V, \mu) = \frac{V}{\lambda^3} \cdot g_{\frac{3}{2}} \left(e^{\beta\mu} \right) \quad (4.73)$$

The critical temperature is then obtained from:

$$\frac{N(T_c, V, \mu = 0)}{Z_1(T_c)} = \zeta\left(\frac{3}{2}\right)$$

So for the ideal gas we get:

$$\boxed{\frac{\lambda_c^3}{v} = \zeta\left(\frac{3}{2}\right) \approx 2,616} \quad (4.74)$$

The critical temperature is therefore:

$$k_B T_c \approx \frac{2\pi}{\left(\zeta\left(\frac{3}{2}\right)\right)^{\frac{2}{3}}} \frac{h^2}{m v^{\frac{2}{3}}}$$

Example: For the parameters of liquid ^4He with $v = 46 \text{ \AA}^3$ is $T_c = 3,13 \text{ K}$.

- Temperature dependence of the ground state occupation:

$$\frac{Z_1(T)}{Z_1(T_c)} = \left(\frac{T}{T_c}\right)^{\frac{3}{2}}$$

For $T \leq T_c$ follows:

$$\bar{n}_g = N \left(1 - \left(\frac{T}{T_c}\right)^{\frac{3}{2}}\right) \quad (4.75)$$

- The Energy is (cf. FLIESSBACH, chapter 31, page 269):

$$E = \sum_{\vec{p}} \varepsilon_p \bar{n}_p = \frac{3}{2} k_B T \frac{V}{\lambda^3} \cdot g_{\frac{5}{2}}(e^{\beta\mu}) \quad (4.76)$$

This expression is also valid above T_c , because the ground state energy for an ideal gas is zero.

Specific heat of the ideal gas

$$\frac{c_V(T)}{k_B} = \frac{1}{N k_B} \frac{\partial E(T, V, N)}{\partial T}$$

- i) For $T \leq T_c$ we have $\mu = 0$ and:

$$\frac{E}{N} = \frac{3}{2} k_B T \frac{v}{\lambda^3} \zeta\left(\frac{5}{2}\right)$$

Due to $\lambda \sim T^{-\frac{1}{2}}$ and thus

$$\frac{\partial}{\partial T} \left(\frac{T}{\lambda^3}\right) \stackrel{\lambda^{-3} = \alpha T^{\frac{3}{2}}}{=} \alpha \frac{\partial}{\partial T} T^{\frac{5}{2}} = \frac{5}{2} \underbrace{\alpha T^{\frac{3}{2}}}_{=\lambda^{-3}} = \frac{5}{2\lambda^3}$$

follows:

$$\frac{c_V(T)}{k_B} = \frac{15}{4} \frac{v}{\lambda^3} \zeta\left(\frac{5}{2}\right) \stackrel{v = \frac{\lambda_c^3}{\zeta(\frac{5}{2})}}{=} \frac{15}{4} \frac{\lambda_c^3}{\lambda^3} \frac{\zeta(\frac{5}{2})}{\zeta(\frac{3}{2})} = \frac{15}{4} \frac{\zeta(\frac{5}{2})}{\zeta(\frac{3}{2})} \left(\frac{T}{T_c}\right)^{\frac{3}{2}} \quad (4.77)$$

- ii) For $T > T_c$ follows from (4.76):

$$\frac{c_V(T)}{k_B} = \frac{15}{4} \frac{v}{\lambda^3} g_{\frac{5}{2}}(e^{\beta\mu}) + \frac{3}{2} T \frac{v}{\lambda^3} g'_{\frac{5}{2}}(e^{\beta\mu}) \cdot \frac{\partial}{\partial T} (e^{\beta\mu(T,v)})$$

To proceed, let us consider 4.73 for $T \geq T_c$:

$$1 = \frac{v}{\lambda^3} g_{\frac{3}{2}}(e^{\beta\mu})$$

Upon taking $\frac{\partial}{\partial T}$ this yields and the multiplying by λ^3 :

$$0 = \frac{3}{2T} g_{\frac{3}{2}}(e^{\beta\mu}) + g'_{\frac{3}{2}}(e^{\beta\mu}) \frac{\partial}{\partial T} e^{\beta\mu(T,v)}$$

$$\Rightarrow \quad \frac{\partial}{\partial T} e^{\beta\mu(T,v)} = -\frac{3}{2T} \frac{g_{\frac{3}{2}}(e^{\beta\mu})}{g'_{\frac{3}{2}}(e^{\beta\mu})}$$

Thus we get:

$$\begin{aligned} \frac{c_V(T)}{k_B} &= \frac{15}{4} \frac{v}{\lambda^3} g_{\frac{5}{2}}(e^{\beta\mu}) + \frac{3}{2} T \frac{v}{\lambda^3} g'_{\frac{5}{2}}(e^{\beta\mu}) \left(-\frac{3}{2T} \frac{g_{\frac{3}{2}}(e^{\beta\mu})}{g'_{\frac{3}{2}}(e^{\beta\mu})} \right) = \\ &\stackrel{z:=e^{\beta\mu}}{\stackrel{\lambda^3}{v}=g_{\frac{3}{2}}(z)} \frac{15}{4} \cdot \frac{g_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)} - \frac{9}{4} \cdot \frac{g'_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)} \cdot \frac{g_{\frac{3}{2}}(z)}{g'_{\frac{3}{2}}(z)} = \\ &= \frac{15}{4} \cdot \frac{g_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)} - \frac{9}{4} \cdot \frac{g'_{\frac{5}{2}}(z)}{g'_{\frac{3}{2}}(z)} = \frac{15}{4} \cdot \frac{g_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)} - \frac{9}{4} \cdot \frac{g_{\frac{3}{2}}(z)}{g_{\frac{1}{2}}(z)} \end{aligned}$$

In the last step we used:

$$z g'_\nu(z) = g_{\nu-1}(z)$$

This follows directly from the definition (4.68). So for $T > T_c$ the result is:

$$\frac{C_V(T)}{k_B} = \frac{15}{4} \frac{g_{\frac{5}{2}}(e^{\beta\mu})}{g_{\frac{3}{2}}(e^{\beta\mu})} - \frac{9}{4} \frac{g_{\frac{3}{2}}(e^{\beta\mu})}{g_{\frac{1}{2}}(e^{\beta\mu})} \quad (4.78)$$

Note: For $T \gg T_c$ or equivalently $\frac{\lambda^3}{v} \ll 1$ we have $e^{\beta\mu} \ll 1$ and with

$$g_\nu(z) \approx z + \frac{z^2}{2^\nu}$$

for $z \ll 1$ we get:

$$\frac{c_V(T)}{k_B} \approx \frac{3}{2} \left(1 + \frac{e^{\beta\mu}}{2^{\frac{7}{2}}} \right) \quad (4.79)$$

To leading order in $z \approx \frac{\lambda^3}{v} = \left(\frac{T_c}{T}\right)^{\frac{3}{2}} \zeta\left(\frac{3}{2}\right)$ follows.

$$\frac{c_V(T)}{k_B} \approx \frac{3}{2} \left(1 + \frac{\zeta\left(\frac{3}{2}\right)}{2^{\frac{7}{2}}} \left(\frac{T_c}{T}\right)^{\frac{3}{2}} \right)$$

Note:

$$\zeta\left(\frac{1}{2}\right) = \infty$$

$$g_{\frac{1}{2}}(z) \xrightarrow{z \rightarrow 0} \infty$$

This implies for $T = T_c$:

$$\frac{c_V(T_c)}{k_B} = \frac{15}{4} \frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)} \approx 1.925$$

TODO: Abb16; Plot $\frac{c_V}{T}$

So the specific heat is continuous, but not its derivative!

What is the order of the Bose-Einstein condensation phase transition?

Paul Ehrenfest introduced a ordering scheme for phase transitions:

The order of a transition is the lowest derivative of the free energy, that is discontinuous at the critical temperature.

This classification proved not be useful and thus one introduced the modern classification, that is named similarly to the Ehrenfest classification:

A *first order* phase transition involves latent heat, i.e. c_P is singular, while a *second order* phase transition does not involve latent heat, i.e. c_P is finite.

The Bose-Einstein condensation involves latent heat and thus is of first order.

Equation of state for the ideal gas

From

$$P(T, V, N) = \frac{2}{3} \frac{E(T, V, N)}{V}$$

and the above derived expressions for $E(T, V, N)$ and $N(T, V, \mu)$ follows for $T \leq T_c$:

$$P(T, V, N) = \frac{2}{3} \frac{E(T, V, \mu = 0)}{V} = \frac{k_B T}{\lambda^3} \zeta\left(\frac{5}{2}\right) \quad (4.80)$$

This means:

$$\frac{\partial P(T, V, N)}{\partial V} = 0$$

The system does not resist to a change in volume. In other words the compressibility diverges:

$$\kappa_T = -\frac{1}{V} \frac{\partial V(P, T)}{\partial P} \rightarrow \infty \quad (4.81)$$

So c_P is diverging as well. In summary:

TODO: Abb17, abb18

Bose-Einstein condensation of ultracold alkali atoms

The first demonstration was in 1995 by Cornell, Wiemann and Ketterle. (Nobel Prize 2001)

– Why alkali atoms? For the ideal gas holds:

$$T_c \approx \frac{2\pi\hbar^2}{m} \left(\frac{n}{2,612} \right)^{\frac{2}{3}}$$

For a gas in a 3D harmonic trap holds:

$$T_c = \frac{\hbar\omega}{k_B} \left(\frac{N}{\zeta(3)} \right)^{\frac{1}{3}} \approx 0,94 \frac{\hbar\omega}{k_B} N^{\frac{1}{3}}$$

However, the interaction is small for small densities. The condition for the scattering length a is:

$$n|a|^3 \ll 1$$

$$a = \frac{m}{4\pi\hbar} \int d^3r V(r)$$

This is the s -wave scattering length. $V(r)$ is the inter atomic potential for neutral atoms. Low densities are required (good), but temperatures are then very low (bad).

Example: ^{23}Na

$$N \approx 10^6$$

$$n \approx 10^{14} \text{ cm}^{-3}$$

$$\Rightarrow T_c \approx 2 \cdot 10^{-6} \text{ K}$$

How do you reach this temperature?

Use magneto-optical traps and radio-frequency evaporative cooling.

TODO: Abb19

4.6.2. Black body radiation

We consider a metallic cavity at temperature T and want to evaluate the emitted radiation.

- i) *Wave equation:* For the radiation inside the cavity, the Maxwell equations hold. For the electric field $\vec{E}(\vec{r}, t)$ this means:

$$\left(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \vec{E} = 0$$

Using the ansatz

$$\vec{E}(\vec{r}, t) = \vec{E}_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)}$$

yields the dispersion relation:

$$\omega^2 = c^2 \|\vec{k}\|^2$$

$$\Rightarrow \boxed{\omega = c \|\vec{k}\| = ck} \quad (4.82)$$

From

$$\vec{\nabla} \cdot \vec{E} = 0$$

follows:

$$\vec{k} \cdot \vec{E}_0 = 0$$

So the electric field is transverse to the direction of propagation, i.e. two components of \vec{E}_0 are independent. This means there are “two kinds” of photons.

- ii) *Boundary condition:* Consider the cavity as a cube of volume $V = L^3$. (The shape is irrelevant if $L \gg \lambda = \frac{2\pi}{k}$.) The walls are metallic, so $\vec{E}_{\parallel}(\vec{r}, t)$ and $\vec{B}_{\perp}(\vec{r}, t)$ vanish at the boundaries, i.e.:

$$\vec{E}_{\parallel}(\vec{r}, t) = 0$$

$$\vec{B}_{\perp}(\vec{r}, t) = 0$$

This leads to standing waves with:

$$k = \frac{n\pi}{L} \quad \vec{n} = (n_x, n_y, n_z) \in \mathbb{N}^3 \quad n = \sqrt{n_x^2 + n_y^2 + n_z^2} \quad (4.83)$$

The vector \vec{n} specifies the character of the harmonic mode.

- iii) *Quantization*: Planck made the hypothesis that the amplitudes $A_i(t) = E_{0,i}(\vec{r}) \sin(\omega t)$ of each mode are quantized. This in turn implies quantization of the energy of each mode. Assuming a mode polarization index m and a vector \vec{k} it holds:

$$\varepsilon_j = \varepsilon_{\vec{k},m} = \hbar\omega(k) \left(n_{\vec{k},m} + \frac{1}{2} \right) = \hbar\omega(k) \left(n_j + \frac{1}{2} \right) \quad (4.84)$$

Here is $j \in \mathbb{N}_{\geq 1}$ and $\varepsilon_1 \leq \varepsilon_2 \leq \dots$. This means that in the mode with energy $\hbar\omega(k)$ are n_j quanta of light (photons).

- iv) *Energy of microstate r* : $r = (n_1, n_2, \dots) = (n_{\vec{k},m})$ is the microstate with:

$$E_r = \sum_{j=1}^{\infty} \varepsilon_j = \sum_{\vec{k},m} \hbar\omega(k) \left(n_{\vec{k},m} + \frac{1}{2} \right) = E_0 + \sum_{\vec{k},m} \hbar\omega(k) n_{\vec{k},m} \quad (4.85)$$

Due to $n_{\vec{k},m} \in \mathbb{N}$, the associated “particles” are bosons. Moreover the number of particles is not fixed, yielding:

$$\mu = \frac{\overline{\partial E_r}}{\partial N_{\text{ph}}} = 0$$

- v) *Statistics* is obtained from the Bose-Einstein distribution with $\mu = 0$. This gives the *Planck distribution*:

$$\bar{n}_{\vec{k},m} = \bar{n}_k = \frac{1}{e^{\beta\hbar\omega(k)} - 1} = \frac{1}{e^{\beta\hbar ck} - 1} \quad (4.86)$$

- vi) *Average energy*: Due to $E_0 = \infty$, we only consider $E' = E - E_0$. We find:

$$\begin{aligned} E' &= \sum_{\vec{k},m} \varepsilon_k \bar{n}_{\vec{k}} = 2 \sum_{n_x, n_y, n_z} \varepsilon_k \bar{n}_{\vec{k}} = \frac{2V}{(2\pi)^3} \int_{-\infty}^{\infty} d^3k \frac{\hbar ck}{e^{\beta\hbar ck} - 1} = \\ &= \frac{V}{\pi^2} \int_{-\infty}^{\infty} k^2 dk \frac{\hbar ck}{e^{\beta\hbar ck} - 1} \stackrel{x=\beta\hbar ck}{\underset{dk=\frac{1}{\beta\hbar c} dx}} = \frac{V}{\pi^2 \beta^4 \hbar^3 c^3} \int_{-\infty}^{\infty} \frac{x^3 dx}{e^x - 1} = \\ &= \frac{\pi^2 V}{15 \hbar^3 c^3} (k_B T)^4 \end{aligned}$$

With the *Stefan Boltzmann constant*

$$\sigma = \frac{\pi^3 k_B^4}{60 \hbar^3 c^2} = \frac{2\pi^5 k_B^4}{15 \hbar^3 c^2}$$

this gives:

$$E' = \frac{4\sigma}{c} V T^4 \quad (4.87)$$

So the heat capacity is:

$$C_V = \frac{16}{c} \sigma V T^3$$

The dependence on T^3 is typical of bosons with linear dispersion. For example this holds also true for *phonons*.

- vii) *Planck's radiation law*: The integral in the evaluation of E' can also be written as an integral over energy:

$$\frac{E'}{V} = \int_0^\infty d\omega u(\omega)$$

This gives Planck's radiation law, obtained in 1900:

$$u(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\frac{\hbar\omega}{k_B T}} - 1} \quad (4.88)$$

TODO: Plot $u(\omega)$

For $\omega \rightarrow \infty$ holds $u(\omega) \approx e^{-\frac{\hbar\omega}{k_B T}} \omega^3$ (Wien) and for $\omega \rightarrow 0$ holds $u(\omega) \approx \omega^2$ (Rayleigh-Jeans).

The importance of Planck's law is due to the fact that a pure classical treatment would have implied:

$$u_{\text{cl}}(\omega) = \frac{k_B T \omega^2}{\pi^2 c^3} \quad (4.89)$$

This leads to the ultraviolet catastrophe, i.e. the divergence of E' for high ω .

- viii) *Stefan-Boltzmann law*: The emitted power from a surface of area f is (cf. FLIESSBACH, chapter 34, page 303):

$$P_{\text{em}} = \frac{E' c}{4V} f = \sigma f T^4 \quad (4.90)$$

This is the Stefan-Boltzmann law. It enables us e.g. to estimate the temperature of the earth, knowing the temperature of the sun.

- ix) *Application*: Greenhouse effect

Part II.

Solid State Theory

Motivation

We deal with systems in the solid phase. We look in particular at electronic and thermal properties of these systems.

- Because of the large number of constituents, one has to resort to methods of statistical physics. However, the situation is complicated due to:
 - i) presence of a lattice
 - ii) presence of electrostatic interactions
- This new complexity requires novel methods (\rightarrow second quantization formalism, ...) and novel approximations (\rightarrow decoupling of motion of ions and electrons, devise minimal model Hamiltonians which capture the relevant low energy physics).
- We shall primarily focus on “bulk” properties, i.e. properties of large crystalline systems. However, present technologies allow to routinely produce structure of *reduced* dimensionality, e.g. two dimensional electron gas in graphene, quantum wires, nanotubes, quantum dots, single molecules, ..., and also to study hybrids of those, e.g. use carbon nanotubes as a contact for a molecular junction:

TODO: Abb20

5 Introduction to Solid State theory

5.1. The Hamiltonian of a solid

We consider as a starting point a system as being composed of ions (nuclei and closed electron shells) and valence electrons. Moreover we assume that the ions are arranged in a regular lattice at zero temperature.

TODO: Abb21

We are also in the situation that the interaction among the constituents is known (electrostatic interaction). So the Hamiltonian of a solid then acquires a very general form:

$$\hat{H} = \hat{T}_{\text{ion}} + \hat{T}_{\text{el}} + \hat{V}_{\text{ee}} + \hat{V}_{\text{ii}} + \hat{V}_{\text{ei}} \quad (5.1)$$

The \hat{T} are the kinetic energies for the ions and the electrons and the \hat{V} are the potentials for electron-electron, ion-ion and electron-ion interaction. Specifically these look like:

$$\hat{H} = - \sum_{\alpha=1}^M \frac{\hbar^2}{2M_{\alpha}} \nabla_{\alpha}^2 - \sum_{j=1}^N \frac{\hbar^2}{2m} \nabla_j^2 + \sum_{j < k}^N \frac{e^2}{|\vec{r}_j - \vec{r}_k|} + \sum_{\alpha < \beta}^M \frac{Z_{\alpha} Z_{\beta} e^2}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|} - \sum_j \sum_{\alpha} \frac{Z_{\alpha} e^2}{|\vec{r}_j - \vec{R}_{\alpha}|} \quad (5.2)$$

Here Z_{α} is the effective charge number of the ions and due to charge neutrality holds $N = MZ$.

Problem: This Hamiltonian is too complicated to be solved exactly. So some approximations are required.

5.2. Structural reducibility (adiabatic approximation)

Not all components entering entering (5.2) have to be treated simultaneously. The motion of the electrons and the motion of the ions decouple. (Born-Oppenheimer approximation)

The decoupling is based on the observation that, due to their large mass M_{α} , the ions move much slower than the electrons of mass m , i.e. the electrons adapt instantaneously to the actual ion configuration.

$$\frac{m}{M_{\alpha}} \approx 10^{-4}$$

Qualitative proof

Assume that the whole system is in thermal equilibrium. So the electrons and the ions have the same average thermal energy per degree of freedom.

$$\left\langle \frac{mv_i^2}{2} \right\rangle = \left\langle \frac{M_{\alpha}v_{\alpha}^2}{2} \right\rangle = \frac{k_B T}{2}$$

Due to $\frac{m}{M_{\alpha}} \approx 10^{-4}$ the electrons move faster.

Quantitative proof

In order to describe the solid, we should solve the Schrödinger equation:

$$\hat{H}\psi = E\psi$$

Here $\psi = \psi\left((\vec{r}_j), (\vec{R}_\alpha)\right) =: \psi\left(\vec{r}, \vec{R}\right)$ depends on the positions of the electrons and the ions.

- i) Let us consider the Schrödinger equation for the electrons moving in the electrostatic potential described by the *instantaneous* configuration $\{\vec{R}_\alpha\}$ of the ions.

$$\underbrace{\left(\hat{T}_{\text{el}} + \hat{V}_{\text{ee}} + \hat{V}_{\text{ii}} + \hat{V}_{\text{ei}}\right)}_{=\hat{H}_{\text{el}}} \phi_k = \varepsilon_{\text{el},k}(\vec{R}) \phi_k(\vec{r}, \vec{R}) \quad (5.3)$$

$\varepsilon_{\text{el},k}(\vec{R})$ is called *potential energy surface*. The (\vec{R}_α) are only parameters of ϕ_k .

- ii) For a given configuration (\vec{R}_α) , the wave functions ϕ_k form a complete set and can be used to expand the total wave function:

$$\psi(\vec{r}, \vec{R}) = \sum_k \phi_k(\vec{r}, \vec{R}) \chi_k(\vec{R}) \quad (5.4)$$

$$\Rightarrow \hat{H}\psi = \sum_k \left(\hat{H}_{\text{el}} + \hat{T}_{\text{ion}}\right) \phi_k \chi_k = E \sum_k \phi_k \chi_k \quad (5.5)$$

Here we have:

$$\hat{T}_{\text{ion}} = \sum_{\alpha=1}^M \sum_{l \in \{x,y,z\}} \frac{\hat{p}_{\alpha,l}^2}{2M_\alpha} \quad \hat{p}_{\alpha,l} = -i\hbar \frac{\partial}{\partial R_{\alpha,l}}$$

- iii) Now use the completeness of (ϕ_k) to eliminate the electron coordinates by multiplying with ϕ_s^* and integrating over (\vec{r}_j) . We use the product formula

$$\frac{\partial^2}{\partial \vec{R}_\alpha^2} \left(\phi_k(\vec{r}, \vec{R}) \chi_k(\vec{R}) \right) = \phi_k \frac{\partial^2 \chi_k}{\partial \vec{R}_\alpha^2} + 2 \frac{\partial \phi_k}{\partial \vec{R}_\alpha} \cdot \frac{\partial \chi_k}{\partial \vec{R}_\alpha} + \chi_k \frac{\partial^2 \phi_k}{\partial \vec{R}_\alpha^2}$$

to get:

$$\begin{aligned} E \sum_k \underbrace{\int d^{3N}r \cdot \phi_s^* \phi_k}_{=\delta_{s,k}} \chi_k &= \sum_k \underbrace{\int d^{3N}r \cdot \phi_s^* \phi_k}_{=\delta_{s,k}} \left(\varepsilon_k + \hat{T}_{\text{ion}} \right) \chi_k + \\ &+ \sum_k \sum_{\alpha=1}^M \sum_l \int d^{3N}r \cdot \left(\left(\frac{\hat{p}_{\alpha,l}^2}{2M_\alpha} \phi_k \right) \chi_k + 2 \frac{(\hat{p}_{\alpha,l} \phi_k) (\hat{p}_{\alpha,l} \chi_k)}{2M_\alpha} \right) \end{aligned}$$

This gives:

$$\left(\hat{T}_{\text{ion}} + \varepsilon_s(\vec{R}) \right) \chi_s + \sum_k \hat{A}_{sk}(\vec{R}) \chi_k = E \chi_s \quad (5.6)$$

Here the operator \hat{A}_{sk} is:

$$\begin{aligned}\hat{A}_{sk}(\vec{R}) &= \sum_{\alpha=1}^M \sum_{l \in \{x,y,z\}} \frac{\hbar^2}{2M_\alpha} \int d^3N r \cdot \phi_s^* \left(\left(\frac{\partial^2}{\partial R_{\alpha l}^2} \phi_k \right) + 2 \left(\frac{\partial}{\partial R_{\alpha l}} \phi_k \right) \frac{\partial}{\partial R_{\alpha l}} \right) = \\ &=: \hat{A}_{sk,1} + \hat{A}_{sk,2}\end{aligned}$$

Neglecting the contribution from \hat{A}_{sk} , i.e. transitions between different quantum numbers s and k due to the motion of the ions, one gets:

$$\left(\hat{T}_{\text{ion}} + \varepsilon_s(\vec{R}) \right) \chi_s(\vec{R}) = E \chi_s(\vec{R}) \quad (5.7)$$

This is a Schrödinger equation for the ions in the effective potential ε_s .

Note: All effects of the chemical binding among the ions are captured in $\varepsilon_s(\vec{R})$. The equilibrium configuration of the lattice corresponds to minima of $\varepsilon_s(\vec{R})$ at positions $(\vec{R}_{0,\alpha})$, i.e. the crystalline structure.

Note: Neglecting \hat{A}_{sk} amounts to the Born-Oppenheimer approximation. (see CZYCHOLL)

$$\hat{A}_{sk,1} \approx \frac{m}{M_\alpha} \langle \phi_s | \hat{T}_{\text{ee}} | \phi_k \rangle \quad (5.8)$$

$$\hat{A}_{sk,2} \approx \frac{\hbar}{M_\alpha} \langle \hat{p}_{\text{el}} \rangle \langle \hat{p}_{\text{ion}} \rangle \approx \left(\frac{m}{M_\alpha} \right)^{\frac{3}{4}} E_{\text{el}} \quad (5.9)$$

Note: Corrections due to \hat{A}_{sk} can be treated in perturbation theory.

Summary:

$$\left(\hat{T}_{\text{el}} + \hat{V}_{\text{el}} + \hat{V}_{\text{ei}} + \hat{V}_{\text{ii}} \right) \phi_k(\vec{R}, \vec{r}) = \varepsilon_{\text{el},k}(\vec{R}) \phi_k(\vec{R}, \vec{r}) \quad \text{electrons} \quad (5.10a)$$

$$\left(\hat{T}_{\text{ion}} \varepsilon_{\text{el},k}(\vec{R}) \right) \chi_k(\vec{R}) \approx E \chi_k(\vec{R}) \quad \text{ions} \quad (5.10b)$$

5.3. Universality

Focus *not* on the full energy profile of the system, but on its energetically low lying dynamics, for example on the Debye theory for the specific heat of solids.

This is because in the low temperature regime systems tend to behave *universally*, i.e. systems differing in microscopic details (e.g. different ion species) exhibit identical *collective behavior*, e.g. phonons are collective vibration modes of a solid.

As a consequence one can employ simple *model Hamiltonians* without looking at the details of microscopic interaction potentials.

5.4. Symmetries and symmetry breaking

- We know that to symmetries conservation laws are associated. This can help in finding proper eigenfunctions of the underlying problem.
- Interactions can yield a symmetry breaking, i.e. a transition to a new state of matter, where some of the conservation laws present in the initial Hamiltonian are violated for the system's *ground* state.

However, information about symmetries of the original Hamiltonian are contained in the low energy collective excitations, called *Goldstone modes*.

Examples

- A crystal has broken translational symmetry and the collective excitations are phonons.
- A magnet has broken spin rotational symmetries, which leads to magnons as collective excitations.

6 Independent electrons in periodic potentials

In this chapter we shall focus on the electronic problem (5.10a) for the special case in which $\hat{V}_{ee} = 0$. Moreover we assume the ions to sit in their equilibrium position $\vec{R}_{0,\alpha}$ attained at zero temperature and neglect any further effect of \hat{V}_{ii} . We then solve the problem only described by:

$$\begin{aligned}\hat{H} &= \hat{T}_{el} + \hat{V}_{ei} = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m} - \sum_{i=1}^N \sum_{\alpha} \frac{Z_{\alpha} e^2}{|\vec{r}_i - \vec{R}_{0,\alpha}|} = \\ &= \sum_{i=1}^N \underbrace{\left(\frac{\hat{p}_i^2}{2m} + v(\vec{r}_i) \right)}_{=: \hat{h}_i}\end{aligned}\quad (6.1)$$

Here we define:

$$v(\vec{r}_i) := - \sum_{\alpha} \frac{Z_{\alpha} e^2}{|\vec{r}_i - \vec{R}_{0,\alpha}|} \quad (6.2)$$

6.1. Crystal structure

Definitions

1. *Point lattice*: An (infinite) *periodic* lattice is characterized by a point lattice, i.e. the set of *lattice vectors*

$$\vec{R}_{\vec{n}} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3$$

with $\vec{n} \in \mathbb{Z}^3$ and independent vectors \vec{a}_1 , \vec{a}_2 and \vec{a}_3 , known as *primitive lattice vectors*.

TODO: Abb; Example: two-dimensional lattice; $\vec{R} = 2\vec{a}_1 + 3\vec{a}_2$

Clearly, a lattice structure is mapped onto itself under a translation of a lattice vector.

2. *Crystal structure* (cf. ASHCROFT and MERMIN, chapter 4): It is obtained by assigning an atom or group of atoms to each lattice point.

– If an atom is assigned to each lattice point, a *monoatomic Bravais lattice* is obtained.

Examples:

- simple cubic lattice spanned by:

$$\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ a \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix}$$

- bcc lattice spanned by:

$$\frac{a}{2} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \frac{a}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \frac{a}{2} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

- fcc lattice spanned by:

$$\frac{a}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \frac{a}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \frac{a}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

- If a group of atoms is assigned, one speaks of a *lattice with a basis*. The position of the atoms is then specified by $\vec{R} + \vec{\tau}$ where $\vec{\tau}$ is the position of an atom of the basis relative to the lattice point.

Examples:

- A bcc lattice can be seen as a simple cubic lattice with a two point basis:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \frac{a}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

- A fcc lattice can be seen as a simple cubic lattice with a four point basis:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \frac{a}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \frac{a}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \frac{a}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

3. Unit and primitive cells:

- a) A volume of space that, when translated by *all* lattice vectors just fills all space without overlapping itself or leaving voids, is called *primitive (unit) cell*.

A primitive cell must contain precisely one lattice point.

- b) A *unit cell* is a region of space that just fills space without any overlapping when translated through some *subset* of lattice vectors. It is usually chosen to be larger than the primitive cell, but with the symmetry of the lattice.

Example: The primitive cell of the fcc lattice has $\frac{1}{4}$ of the volume of the cubic unit cell.

- c) The Wigner-Seitz cell is the region of space closer to a given lattice point than any other.

TODO: Abb: primitive unit cells (incl. Wigner-Seitz); Non-primitive unit cells

- 4. *Reciprocal space:* Working with periodic structures, it is often convenient to Fourier transform from the direct space to the k -space, also known as the reciprocal space. The reciprocal lattice RL in reciprocal space RS is defined by:

$$\text{RL} = \left\{ \vec{G} \in \text{RS} \mid e^{i\vec{G} \cdot \vec{R}} = 1 \right\}$$

$$\Rightarrow \vec{G} = m_1 \vec{b}_1 + m_2 \vec{b}_2 + m_3 \vec{b}_3$$

The \vec{b}_1 , \vec{b}_2 and \vec{b}_3 are the basis vectors defined (in three dimensions) by:

$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \quad \vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \quad \vec{b}_3 = 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}$$

TODO: Abb: reciprocal lattice

5. *First Brillouin zone*: The first Brillouin zone (1.BZ), or Wigner-Seitz primitive cell, is defined as all \vec{k} in reciprocal space lying closer to $\vec{G} = 0$ than to any other lattice vector $\vec{G} \neq 0$.

$$1.BZ = \left\{ \vec{k} \in \text{RS} \mid \|\vec{k}\| < \|\vec{k} + \vec{G}\| \quad \forall \vec{G} \in \text{RL} \setminus \{0\} \right\}$$

This means, that for all $\vec{q} \in \text{RS}$ exists a $\vec{k} \in 1.BZ$ and a $\vec{G} \in \text{RL}$ such that holds:

$$\vec{q} = \vec{k} + \vec{G}$$

6. *Fourier transform*: The Fourier transform of any in the lattice periodic function

$$\vec{V}(\vec{r}) = V(\vec{r} + \vec{R}_\alpha)$$

is $\tilde{V}(\vec{G})$:

$$V(\vec{r}) = \sum_{\vec{G} \in \text{RL}} \tilde{V}(\vec{G}) e^{i\vec{G} \cdot \vec{r}}$$

See also the integrated course IIa and the ASHCROFT and MERMIN.

6.2. Bloch theorem

We investigate the single particle problem described by:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + v(\vec{r}) \right) \psi(\vec{r}) = \varepsilon \psi(\vec{r}) \quad (6.3)$$

For all lattice vectors \vec{R} holds:

$$v(\vec{r} + \vec{R}) = v(\vec{r})$$

Translation operator

Let us introduce the translation operator $\hat{T}_{\vec{R}}$, defined for all $\phi \in H$ by:

$$\hat{T}_{\vec{R}} \phi(\vec{r}) = \phi(\vec{r} + \vec{R}) \quad (6.4)$$

- Any $\hat{T}_{\vec{R}}$ with a lattice vector \vec{R} commutes with \hat{h} :

$$\begin{aligned} \hat{T}_{\vec{R}} \hat{h} \phi(\vec{r}) &= \hat{T}_{\vec{R}} \left(-\frac{\hbar^2}{2m} \nabla^2 + v(\vec{r}) \right) \phi(\vec{r}) = \\ &= \left(-\frac{\hat{p}^2}{2m} + v(\vec{r} + \vec{R}) \right) \phi(\vec{r} + \vec{R}) = \hat{h} \hat{T}_{\vec{R}} \phi(\vec{r}) \end{aligned}$$

- The translation operators commute with each other:

$$\begin{aligned} \hat{T}_{\vec{R}} \hat{T}_{\vec{R}'} &= \hat{T}_{\vec{R}'} \hat{T}_{\vec{R}} = \hat{T}_{\vec{R} + \vec{R}'} \\ \Rightarrow [\hat{T}_{\vec{R}}, \hat{T}_{\vec{R}'}] &= 0 \end{aligned} \quad (6.5)$$

- We thus can look for simultaneous eigenfunctions of $\hat{T}_{\vec{R}}$ and \hat{h} .

$$\begin{aligned}\hat{h}\psi(\vec{r}) &= \varepsilon\psi(\vec{r}) \\ \hat{T}_{\vec{R}}\psi(\vec{r}) &= C(\vec{R})\psi(\vec{r})\end{aligned}$$

From (6.5) follows:

$$C(\vec{R})C(\vec{R}') = C(\vec{R} + \vec{R}')$$

- The normalization condition

$$1 = \int d^3r |\psi(\vec{r})|^2 \stackrel{\text{variable transformation}}{=} \int d^3r |\psi(\vec{r} + \vec{R})|^2 = \int d^3r |C(\vec{R})|^2 |\psi(\vec{r})|^2$$

gives:

$$\begin{aligned}|C(\vec{R})|^2 &= 1 \\ \Rightarrow C(\vec{R}) &= e^{i\vec{k} \cdot \vec{R}}\end{aligned}$$

First form of the Bloch theorem

The form of the eigenfunctions ψ of $\hat{T}_{\vec{R}}$ is such that:

$$\boxed{\psi(\vec{r} + \vec{R}) = e^{i\vec{k} \cdot \vec{R}}\psi(\vec{r})} \quad (6.6)$$

Bloch functions are functions of this form, which are also eigenfunctions of \hat{h} .

Periodic boundary conditions

Now we write $\vec{R} = N_i \vec{a}_i$ to get:

$$\psi(\vec{r} + N_i \vec{a}_i) = \psi(\vec{r})$$

N_i is chosen such that $L_i = N_i \|\vec{a}_i\|$ is the length of the system in the i -direction. The periodic boundary condition yields:

$$e^{i\vec{k} \cdot N_i \vec{a}_i} = 1 \quad (6.7)$$

With $\vec{b}_i \cdot \vec{a}_j = 2\pi\delta_{ij}$ follows for $n_i \in \mathbb{Z}$:

$$\vec{k} = \sum_i \frac{n_i}{N_i} \vec{b}_i \quad (6.8)$$

The k -space per allowed value of \vec{k} is:

$$V_{\vec{k}} = \frac{\vec{b}_1}{N_1} \cdot \left(\frac{\vec{b}_2}{N_2} \times \frac{\vec{b}_3}{N_3} \right) = \frac{V_{\text{pc}}}{N_1 N_2 N_3} = \frac{V_{\text{pc}}}{N_{\text{cell}}}$$

N_{cell} is the number of allowed values of k per primitive cell with volume V_{pc} . Moreover holds:

$$\begin{aligned}\vec{b}_1 &= 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} & \vec{b}_2 &= 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_2 \cdot (\vec{a}_3 \times \vec{a}_1)} & \vec{b}_3 &= 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_3 \cdot (\vec{a}_1 \times \vec{a}_2)} \\ \Rightarrow V_{\text{pc}} &= \frac{(2\pi)^3}{v} = \frac{(2\pi)^3}{V} N_{\text{cell}} \\ \Rightarrow V_{\vec{k}} &= \frac{(2\pi)^3}{V}\end{aligned}$$

Second form of the Bloch theorem

Define for a Bloch function $\psi_{\vec{k}}(\vec{r})$ the *Bloch factor*:

$$u_{\vec{k}}(\vec{r}) = ae^{-i\vec{k}\cdot\vec{r}}\psi_{\vec{k}}(\vec{r}) \quad (6.9a)$$

With the first form of the Bloch theorem follows:

$$u_{\vec{k}}(\vec{r} + \vec{R}) = u_{\vec{k}}(\vec{r}) \quad (6.9b)$$

The *normalized Bloch function* is:

$$\boxed{\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}}e^{i\vec{k}\cdot\vec{r}}u_{\vec{k}}(\vec{r})} \quad (6.10)$$

Hence, as the plane waves, Bloch waves are delocalized.

The differential equation for $u_{\vec{k}}(\vec{r})$ follows from the Schrödinger equation for ψ :

$$\hat{h}\sqrt{V}\psi_{\vec{k}}(\vec{r}) = \left(-\frac{\hbar^2}{2m}\nabla^2 + v(\vec{r})\right)e^{i\vec{k}\cdot\vec{r}}u_{\vec{k}}(\vec{r}) = \varepsilon e^{i\vec{k}\cdot\vec{r}}u_{\vec{k}}(\vec{r})$$

This gives:

$$\left(\frac{\hbar^2}{2m}\left(\frac{\nabla}{i} + \vec{k}\right)^2 + v(\vec{r})\right)u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k})u_{\vec{k}}(\vec{r}) \quad (6.11)$$

In equation (6.11) the wave vector \vec{k} is only a parameter. Because of the periodicity of $u_{\vec{k}}(\vec{r})$ this problem is equivalent to a boundary problem at the boundary of the unit cell. Therefore, at fixed \vec{k} one expects discrete eigenvalues $\varepsilon_n(\vec{k})$ associated to eigenfunctions $u_{n,\vec{k}}(\vec{r})$.

So as a function of \vec{k} one obtains the n -th band $\varepsilon_n(\vec{k})$.

In general, in Dirac notation, the Schrödinger equation reads:

$$\hat{h}|n, \vec{k}\rangle = \varepsilon_n(\vec{k})|n, \vec{k}\rangle \quad (6.12)$$

Note: The eigenfunctions $u_{n,\vec{k}}$ can be normalized on the unit cell:

$$\frac{1}{V_{\text{cell}}}\int_{V_{\text{cell}}}d^3r u_{n,\vec{k}}^*(\vec{r})u_{n',\vec{k}}(\vec{r}) = \delta_{nn'} \quad (6.13)$$

Then the Bloch functions are also normalized over the full space, i.e. we get:

$$\int_V d^3r \psi_{n,\vec{k}}^*(\vec{r})\psi_{n',\vec{k}'}(\vec{r}) = \delta_{\vec{k},\vec{k}'}\delta_{nn'} \quad (6.14)$$

The completeness relation is:

$$\sum_n u_{n,\vec{k}}^*(\vec{r})u_{n,\vec{k}}(\vec{r}') = V_{\text{cell}}\delta(\vec{r} - \vec{r}') \quad (6.15)$$

Wannier functions

One can build a complete set of orthonormal functions, which, in contrast to the Bloch waves, are localized. These states, known as *Wannier states*, are defined as:

$$w_{\alpha,n}(\vec{r} - \vec{R}_\alpha) := \frac{1}{\sqrt{N_{\text{cell}}}} \sum_{\vec{k} \in \text{1.BZ}} e^{-i\vec{k} \cdot \vec{R}_\alpha} \psi_{n,\vec{k}}(\vec{r}) \quad (6.16)$$

It holds:

$$\begin{aligned} \int d^3r w_{\alpha,n}^*(\vec{r} - \vec{R}_\alpha) w_{\beta,l}(\vec{r} - \vec{R}_\beta) &= \frac{1}{N_{\text{cell}}} \sum_{\vec{k}, \vec{k}' \in \text{1.BZ}} e^{i(\vec{k} \cdot \vec{R}_\alpha - \vec{k}' \cdot \vec{R}_\beta)} \underbrace{\int d^3r \psi_{n,\vec{k}}^*(\vec{r}) \psi_{l,\vec{k}'}(\vec{r})}_{=\delta_{nl}\delta_{\vec{k},\vec{k}'}} \\ &= \frac{1}{N_{\text{cell}}} \sum_{\vec{k} \in \text{BZ}} e^{i\vec{k} \cdot (\vec{R}_\alpha - \vec{R}_\beta)} \delta_{n,l} = \delta_{\alpha,\beta} \delta_{n,l} \end{aligned} \quad (6.17)$$

Vice versa one can express Bloch functions in terms of Wannier functions:

$$\boxed{\psi_{n,\vec{k}}(\vec{r}) = \frac{1}{\sqrt{N_{\text{cell}}}} \sum_{\alpha} e^{i\vec{k} \cdot \vec{R}_\alpha} w_{\alpha,n}(\vec{r} - \vec{R}_\alpha)} \quad (6.18)$$

Energy dispersion

$$\varepsilon_{n,\vec{k}} = \langle n, \vec{k} | \hat{h} | n, \vec{k} \rangle = \int d^3r \psi_{n,\vec{k}}^*(\vec{r}) h(\vec{r}) \psi_{n,\vec{k}}(\vec{r})$$

This means:

$$\begin{aligned} \varepsilon_{n,\vec{k}} &= \frac{1}{N_{\text{cell}}} \sum_{\alpha,\beta} e^{-i\vec{k} \cdot \vec{R}_\alpha} \cdot e^{i\vec{k} \cdot \vec{R}_\beta} \int d^3r w_n^*(\vec{r} - \vec{R}_\alpha) \left(\frac{\hat{p}^2}{2m} + v(\vec{r}) \right) w(\vec{r} - \vec{R}_\beta) = \\ &\stackrel{\substack{\vec{r} - \vec{R}_\beta =: \vec{r}' \\ \vec{r}' \leadsto \vec{r}}}{=} \frac{1}{N_{\text{cell}}} \sum_{\alpha,\beta} e^{-i\vec{k} \cdot (\vec{R}_\alpha - \vec{R}_\beta)} \int d^3r w_n^*(\vec{r} - \vec{R}_\alpha + \vec{R}_\beta) \left(\frac{\hat{p}^2}{2m} + v(\vec{r}) \right) w_n(\vec{r}) = \\ &\stackrel{\substack{\vec{R}_\gamma =: \vec{R}_\alpha - \vec{R}_\beta \\ =}}{=} \frac{1}{N_{\text{cell}}} \sum_{\alpha,\gamma} e^{-i\vec{k} \cdot \vec{R}_\gamma} \int d^3r w_n^*(\vec{r} - \vec{R}_\gamma) \left(\frac{\hat{p}^2}{2m} + v(\vec{r}) \right) w_n(\vec{r}) = \\ &\stackrel{\substack{\sum_{\alpha} = N_{\text{cell}} \\ =}}{=} \sum_{\gamma} e^{-i\vec{k} \cdot \vec{R}_\gamma} \int d^3r w_n^*(\vec{r} - \vec{R}_\gamma) \left(\frac{\hat{p}^2}{2m} + v(\vec{r}) \right) w_n(\vec{r}) = \\ &= \underbrace{\int d^3r w_n^*(\vec{r}) \left(\frac{\hat{p}^2}{2m} + v(\vec{r}) \right) w_n(\vec{r})}_{=: E_n} + \\ &\quad + \sum_{\gamma; \vec{R}_\gamma \neq 0} e^{-i\vec{k} \cdot \vec{R}_\gamma} \underbrace{\int d^3r w_n^*(\vec{r} - \vec{R}_\gamma) \left(\frac{\hat{p}^2}{2m} + v(\vec{r}) \right) w_n(\vec{r})}_{=: \gamma_{nn}(\vec{R}_\gamma) = \gamma_{nn}(-\vec{R}_\gamma)} \end{aligned}$$

$\gamma_{nn}(\vec{R})$ is called the hopping matrix. One can show that the kinetic part $\frac{\vec{p}^2}{2m}$ does not contribute to γ_{nn} and due to $v < 0$ we get $\gamma_{nn} < 0$. So we found:

$$\varepsilon_{n,\vec{k}} = E_n + \sum_{\alpha; \vec{R}_\alpha \neq 0} e^{-i\vec{k} \cdot \vec{R}_\alpha} \gamma_{nn}(\vec{R}_\alpha) \quad (6.19)$$

This is still exact.

Approximation

- Restrict the sum to nearest neighbors only.
- In the calculation of $\gamma_{nn}(\vec{R}_\alpha)$ neglect “three center” integrals:

$$v(\vec{r}) = \sum_{\beta} v_{\text{atom}}(\vec{r} - \vec{R}_\beta) \approx v_{\text{atom}}(\vec{r}) + v_{\text{atom}}(\vec{r} - \vec{R}_\alpha)$$

- *Example*: one-dimensional chain

TODO: Abb22

$$\varepsilon(k) = E + e^{-ika} \gamma(a) + e^{ika} \gamma(-a) = E + 2\gamma(a) \cos(ka)$$

With $-\frac{N_{\text{cell}}}{2} \leq m < \frac{N_{\text{cell}}}{2}$ we have:

$$k_m = \frac{2\pi}{a} \cdot \frac{m}{N_{\text{cell}}}$$

TODO: Abb23

The dispersion is parabolic near $k = 0$.

- *Example*: simple cubic lattice

TODO: Abb24

There are 6 nearest neighbors and define $\gamma := \gamma(a)$ to get:

$$\varepsilon(k) = E + 2\gamma(\cos(k_x a) + \cos(k_y a) + \cos(k_z a))$$

Again the dispersion is parabolic around $\vec{k} = 0$.

- *Note*: Hamilton operator in Wannier basis

$$\hat{h} = \sum_{\alpha, \alpha', n, n'} C_{\alpha, \beta, n, n'} |n, \alpha\rangle \langle n', \alpha'|$$

Here we use:

$$C_{\alpha, \alpha', n, n'} = \langle n, \alpha | \hat{h} | n', \alpha' \rangle = \begin{cases} E_n & \text{for } \alpha = \alpha' \text{ and } n = n' \\ \gamma_{nn'}(\vec{R}_\alpha - \vec{R}_{\alpha'}) & \text{for } \alpha \neq \alpha' \text{ and/or } n \neq n' \end{cases}$$

Now neglect $n \neq n'$ and include only nearest neighbors. So for a given band n we get:

$$\hat{h}_n = \begin{pmatrix} \ddots & & & \\ & E_n & \gamma_{nn} & \\ & \gamma_{nn} & E_n & \gamma_{nn} \\ & & \gamma_{nn} & E_n \\ & & & \ddots \end{pmatrix}$$

Diagonalizing this operator yields again $\varepsilon_{n,\vec{k}}$.

How to evaluate the electronic band structure?

One can envisage, besides numerical “ab initio” methods, two semi-analytical approaches:

method	LCAO	almost free electron
based on	localized states	plane waves
improvement	next nearest atoms	other waves
good for	covalent solids ¹ , transition metals ²	simple metals

6.3. Linear combination of atomic orbitals (LCAO) method

Underlying idea: View the solid as a collection of weakly interacting atoms.

Advantage: The eigenfunctions and eigenvalues of isolated atoms are known.

– *Example:* Na

electronic structure: $1s^2 2s^2 2p^6 3s^1 = [\text{Ne}] 3s^1$

The energy diagram of Na as a function of the inter-atomic distance is:

TODO: Abb25

– *Example:* Si

electronic structure: $[\text{Ne}] 3s^2 3p^2$

TODO: Abb26

LCAO construction

– Step 1: Schrödinger equation

i) *Isolated atom:* The Schrödinger equation yields:

- atomic energies ε_ν
- atomic orbitals φ_ν

$$\hat{h}_{\text{at}} \varphi_\nu (\vec{r} - \vec{R}_\alpha) = \left(-\frac{\hbar^2}{2m} \nabla^2 + v_{\text{at}} (\vec{r} - \vec{R}_\alpha) \right) \varphi_\nu (\vec{r} - \vec{R}_\alpha) = \varepsilon_\nu \varphi_\nu (\vec{r} - \vec{R}_\alpha) \quad (6.20)$$

ii) *Electrons in the crystal:*

$$\hat{h} = \frac{\hat{p}^2}{2m} + v(\vec{r}) = \frac{\hat{p}^2}{2m} + \sum_\alpha v_{\text{at}} (\vec{r} - \vec{R}_\alpha) = \hat{h}_{\text{at}} + \underbrace{\sum_{\beta \neq \alpha} v_{\text{at}} (\vec{r} - \vec{R}_\beta)}_{=\Delta v(\vec{r})} \quad (6.21)$$

Note: The potential Δv yields the effects of the neighboring atoms on the atomic energies and wave functions.

¹insulators (C, N, P) or semiconductors (Si, Ge, As)

²partially filled *d*-shells

- Step 2: Construct extended wave functions satisfying the Bloch theorem using the φ_ν .

$$\psi_{\nu, \vec{k}}(\vec{r}) = \frac{1}{\sqrt{N_{\text{cell}}}} \sum_{\alpha} e^{i\vec{k} \cdot \vec{R}_{\alpha}} \varphi_{\nu}(\vec{r} - \vec{R}_{\alpha}) \quad (6.22)$$

The values of \vec{k} are set by imposing periodic boundary conditions. The Bloch theorem states:

$$\psi_{\nu, \vec{k}}(\vec{r} + \vec{R}_{\alpha}) = e^{i\vec{k} \cdot \vec{R}_{\alpha}} \psi_{\nu, \vec{k}}(\vec{r})$$

- Step 3: Use the functions $\psi_{\nu, \vec{k}}(\vec{r})$ as a complete set for the expansion of the eigenfunctions of \hat{h} :

$$\psi_{n, \vec{k}}(\vec{r}) = \sum_{\nu} c_{n\nu} \psi_{\nu, \vec{k}}(\vec{r}) \quad (6.23)$$

These solve the Schrödinger equation:

$$\hat{h} \psi_{n, \vec{k}}(\vec{r}) = \varepsilon_n(\vec{k}) \psi_{n, \vec{k}}(\vec{r})$$

- Step 4: Obtain the coefficients $c_{n\nu}$.

1. Observe:

$$\begin{aligned} \hat{h} \psi_{n, \vec{k}}(\vec{r}) &= \frac{1}{\sqrt{N_{\text{cell}}}} \sum_{\nu} \sum_{\alpha} c_{n\nu} \left(-\frac{\hbar^2}{2m} \nabla^2 + \sum_{\beta} v_{\text{at}}(\vec{r} - \vec{R}_{\beta}) \right) e^{i\vec{k} \cdot \vec{R}_{\alpha}} \varphi_{\nu}(\vec{r} - \vec{R}_{\alpha}) = \\ &= \frac{1}{\sqrt{N_{\text{cell}}}} \sum_{\nu} c_{n\nu} \sum_{\alpha} \left(\varepsilon_{\nu} + \sum_{\beta \neq \alpha} v_{\text{at}}(\vec{r} - \vec{R}_{\beta}) \right) e^{i\vec{k} \cdot \vec{R}_{\alpha}} \varphi_{\nu}(\vec{r} - \vec{R}_{\alpha}) \end{aligned}$$

2. Multiply from the left with $\psi_{\nu', \vec{k}}^*(\vec{r})$ and integrate in d^3r :

$$\begin{aligned} \int d^3r \psi_{\nu', \vec{k}}^*(\vec{r}) \hat{h} \psi_{n, \vec{k}}(\vec{r}) &= \frac{1}{N_{\text{cell}}} \int d^3r \sum_{\nu} c_{n\nu} \sum_{\gamma} e^{-i\vec{k} \cdot \vec{R}_{\gamma}} \varphi_{\nu'}^*(\vec{r} - \vec{R}_{\gamma}) \cdot \\ &\quad \cdot \sum_{\alpha} \left(\varepsilon_{\nu} + \sum_{\beta \neq \alpha} v_{\text{at}}(\vec{r} - \vec{R}_{\beta}) \right) e^{i\vec{k} \cdot \vec{R}_{\alpha}} \varphi_{\nu}(\vec{r} - \vec{R}_{\alpha}) \end{aligned}$$

3. Use the orthonormality of the atomic wave functions at the *same site*

$$\int d^3r \varphi_{\nu'}^*(\vec{r}) \varphi_{\nu}(\vec{r}) = \delta_{\nu\nu'}$$

to get:

$$\begin{aligned} 0 &= \int d^3r \psi_{\nu', \vec{k}}^*(\vec{r}) \left(\hat{h} - \varepsilon_n(\vec{k}) \right) \psi_{n, \vec{k}}(\vec{r}) = \\ &= \sum_{\nu} c_{n\nu} \left(\varepsilon_{\nu} - \varepsilon_n(\vec{k}) \right) \underbrace{\left(\delta_{\nu\nu'} + \frac{1}{N_{\text{at}}} \sum_{\alpha, \gamma \neq \alpha} e^{-i\vec{k} \cdot \vec{R}_{\gamma}} e^{i\vec{k} \cdot \vec{R}_{\alpha}} \int d^3r \varphi_{\nu'}^*(\vec{r} - \vec{R}_{\gamma}) \varphi_{\nu}(\vec{r} - \vec{R}_{\alpha}) \right)}_{=S_{\nu'\nu}} \end{aligned}$$

$$\begin{aligned}
 & + \sum_{\nu} c_{n\nu} \frac{1}{N_{\text{cell}}} \sum_{\gamma, \alpha} e^{-i\vec{k} \cdot \vec{R}_{\gamma}} e^{i\vec{k} \cdot \vec{R}_{\alpha}} \int d^3r \varphi_{\nu'}^* (\vec{r} - \vec{R}_{\gamma}) \underbrace{\left(\sum_{\beta \neq \alpha} v_{\text{at}} (\vec{r} - \vec{R}_{\beta}) \right)}_{=: K_{\nu'\nu}} \varphi_{\nu} (\vec{r} - \vec{R}_{\alpha}) = \\
 & = \sum_{\nu} c_{n\nu} \left(\left(\varepsilon_{\nu} - \varepsilon_n (\vec{k}) \right) S_{\nu'\nu} + K_{\nu'\nu} \right)
 \end{aligned}$$

So the eigenvalues $\varepsilon_{n, \vec{k}} = \varepsilon_n (\vec{k})$ are obtained from the secular equation:

$$\det \left(\varepsilon_{\nu} S_{\nu'\nu} + K_{\nu'\nu} - \varepsilon_n (\vec{k}) S_{\nu'\nu} \right) = 0 \quad (6.24)$$

The $S_{\nu'\nu}$ and $K_{\nu'\nu}$ are the so called overlap integrals and are defined as:

$$S_{\nu'\nu} := \delta_{\nu'\nu} + \frac{1}{N_{\text{at}}} \sum_{\alpha, \beta \neq \alpha} e^{-i\vec{k} \cdot \vec{R}_{\beta}} e^{i\vec{k} \cdot \vec{R}_{\alpha}} \int d^3r \cdot \varphi_{\nu'}^* (\vec{r} - \vec{R}_{\beta}) \varphi_{\nu} (\vec{r} - \vec{R}_{\alpha}) \quad (6.25)$$

$$K_{\nu'\nu} := \frac{1}{N_{\text{cell}}} \sum_{\gamma, \alpha, \beta \neq \alpha} e^{-i\vec{k} \cdot \vec{R}_{\gamma}} e^{i\vec{k} \cdot \vec{R}_{\alpha}} \int d^3r \varphi_{\nu'}^* (\vec{r} - \vec{R}_{\gamma}) v_{\text{at}} (\vec{r} - \vec{R}_{\beta}) \varphi_{\nu} (\vec{r} - \vec{R}_{\alpha}) \quad (6.26)$$

Note that the atomic orbitals φ_{ν} are in general not orthogonal on different lattice sites.

Tight binding approximation

- $S_{\nu'\nu} \approx \delta_{\nu'\nu}$
- Neglect three center integrals and retain the nearest neighbors (n.n.) in the sum over α in $K_{\nu'\nu}$.

$$\begin{aligned}
 K_{\nu'\nu} \approx & \underbrace{\sum_{\alpha, \vec{R}_{\alpha} \neq 0} \int d^3r \varphi_{\nu'}^* (\vec{r}) v_{\text{at}} (\vec{r} - \vec{R}_{\alpha}) \varphi_{\nu} (\vec{r})}_{=: c_{\nu'\nu} \text{ crystal field}} + \\
 & + \sum_{\alpha \text{ n.n.}} e^{-i\vec{k} \cdot \vec{R}_{\alpha}} \underbrace{\int d^3r \varphi_{\nu'}^* (\vec{r} - \vec{R}_{\alpha}) v_{\text{at}} (\vec{r} - \vec{R}_{\alpha}) \varphi_{\nu} (\vec{r})}_{\gamma_{\nu'\nu}(\vec{R}_{\alpha}) \text{ overlap term}}
 \end{aligned}$$

The secular equation becomes:

$$\det \left(\underbrace{\varepsilon_{\nu} \delta_{\nu\nu'} + c_{\nu\nu'} + \sum_{\alpha \text{ n.n.}} e^{-i\vec{k} \cdot \vec{R}_{\alpha}} \gamma_{\nu\nu'} (\vec{R}_{\alpha})}_{=: \mathcal{H}_{\nu\nu'}(\vec{k})} - \varepsilon_n (\vec{k}) \delta_{\nu\nu'} \right) = 0 \quad (6.27)$$

- *Note:* In some way, we are approximating Wannier functions

$$w_{\alpha, n}^{\text{LCAO}} (\vec{r} - \vec{R}_{\alpha}) = \sum_{\nu} c_{n, \nu} \psi_{\nu} (\vec{r} - \vec{R}_{\alpha})$$

with the additional assumption $S_{\nu'\nu} \approx \delta_{\nu'\nu}$.

- *Note:* The LCAO method is very convenient for problems, where only a few atomic orbitals ν are required.

- *Example:* With one orbital ν only, the secular equation yields (in this case $n \rightarrow \nu$):

$$\varepsilon(\vec{k}) = \varepsilon_\nu + c_{\nu\nu} + \sum_{\alpha \text{ n.n.}} e^{-i\vec{k} \cdot \vec{R}_\alpha} \gamma_{\nu\nu}(\vec{R}_\alpha)$$

Let us e.g. consider an s -band on a cubic lattice (e.g. $3s$ electron of Na on a cubic lattice).

$$\gamma_{ss} = V_{ss\sigma}$$

TODO: Abb overlap between s -orbitals

$$\varepsilon(\vec{k}) = \varepsilon_s + c_{ss} + \sum_{\alpha \text{ n.n.}} e^{-i\vec{k} \cdot \vec{R}} V_{ss\sigma}$$

TODO: Abb direct lattice a ; spanned by $a\vec{e}_x, a\vec{e}_y, a\vec{e}_z$, reciprocal lattice $\frac{2\pi}{a}$, spanned by $\frac{2\pi}{a}\vec{e}_x, \frac{2\pi}{a}\vec{e}_y, \frac{2\pi}{a}\vec{e}_z$

TODO: Abb 1. Brillouin zone with Γ, M, X, R

$$\Gamma = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad X = \frac{\pi}{a} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad M = \frac{\pi}{a} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad R = \frac{\pi}{a} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

In general, accounting for the symmetry of the cubic lattice and with $V_{ss\sigma} =: \gamma$ follows:

$$\varepsilon(\vec{k}) = \underbrace{\varepsilon_s + c_{ss}}_{=: \tilde{\varepsilon}_s} + 2\gamma (\cos(k_x a) + \cos(k_y a) + \cos(k_z a))$$

Hence we get:

$$\varepsilon(\Gamma) = \tilde{\varepsilon}_s + 6\gamma \quad \varepsilon(X) = \tilde{\varepsilon}_s + 2\gamma \quad \varepsilon(M) = \tilde{\varepsilon}_s - 2\gamma \quad \varepsilon(R) = \tilde{\varepsilon}_s - 6\gamma$$

TODO: Abb Plot $\varepsilon(\vec{k})$ path $\Gamma \rightarrow X \rightarrow M \rightarrow R$

- *Note:* The particle dispersion around the Γ -point can be considered as for free electrons with effective electron mass:

$$m^* := -\frac{\hbar^2}{2a^2\gamma}$$

- *Example:* Consider a one-dimensional chain with two orbitals.

TODO: Abb chain with s - and p -orbitals

The secular approximation requires to evaluate the matrix $\mathcal{H}_{\nu'\nu}$. In this case we have:

$$\mathcal{H} = \begin{pmatrix} \tilde{\varepsilon}_s + V_{ss\sigma}(e^{-ika} + e^{ika}) & c_{sp} + V_{sp\sigma}(e^{-ika} - e^{ika}) \\ c_{ps} + V_{sp\sigma}^*(e^{-ika} - e^{ika}) & \tilde{\varepsilon}_p + V_{pp\sigma}(e^{-ika} + e^{ika}) \end{pmatrix}$$

The two orbitals give two bonds. $\tilde{\varepsilon}_p - \tilde{\varepsilon}_s \gg |V_{ss\sigma}|, |V_{sp\sigma}|, |V_{pp\sigma}|$

TODO: Abb s -band, p -band; $V_{ss\sigma} < 0, V_{pp\sigma} > 0$

- *Example:* Consider a two-dimensional square lattice with four orbitals.

TODO: Abb27

For simplicity assume $c_{\nu\nu'} = c_{\nu\nu}\delta_{\nu\nu'}$ and set $\tilde{\varepsilon}_\nu = \varepsilon_\nu + c_{\nu\nu}$:

$$(\mathcal{H}_{\nu\nu'}) = \begin{pmatrix} \langle s | & \langle p_x | & \langle p_y | & \langle p_z | \\ \begin{pmatrix} \tilde{\varepsilon}_s + V_{ss\sigma}g_0 & V_{sp\sigma}g_1 & V_{sp\sigma}g_2^* & 0 \\ V_{sp\sigma}g_1^* & \tilde{\varepsilon}_p + V_{pp\sigma}g_4 + V_{pp\pi}g_3 & 0 & 0 \\ V_{sp\sigma}g_2 & 0 & \tilde{\varepsilon}_p + V_{pp\sigma}g_3 + V_{pp\pi}g_4 & 0 \\ 0 & 0 & 0 & \tilde{\varepsilon}_p + V_{pp\pi}g_0 \end{pmatrix} \end{pmatrix}$$

Here we used the abbreviations:

$$\begin{aligned} g_0 &:= e^{-ik_x a} + e^{ik_x a} + e^{-ik_y a} + e^{ik_y a} & g_1 &:= e^{ik_x a} - e^{ik_x a} \\ g_2 &:= -e^{-ik_y a} + e^{ik_y a} & g_3 &:= e^{-ik_y a} + e^{ik_y a} \\ g_4 &:= e^{ik_x a} + e^{-ik_x a} \end{aligned}$$

TODO: Abb28

Main features:

1. The four orbitals induce four bands.
2. There are degeneracies at the symmetry points.

TODO: Abb29 (neglecting $V_{sp\sigma}$)

– Orbital overlaps for three-dimensional crystals:

$$\text{i)} \quad {}_1 \langle s | \hat{v}_{\text{atom}} | s \rangle_2 = V_{ss\sigma}$$

TODO: Abb30

$$\text{ii)} \quad {}_1 \langle s | \hat{v}_{\text{atom}} | p \rangle_2$$

TODO: Abb31, abb32

$$\text{iii)} \quad {}_1 \langle p | \hat{v}_{\text{atom}} | p \rangle_2 = \left(\hat{d} \cdot \hat{b}_1 \right) \left(\hat{d} \cdot \hat{b}_2 \right) V_{pp\sigma} + \left(\hat{b}_1 - \hat{d} \left(\hat{b}_1 \cdot \hat{d} \right) \right) \left(\hat{b}_2 - \hat{d} \left(\hat{b}_2 \cdot \hat{d} \right) \right) V_{pp\pi}$$

TODO: Abb33

Example

The diamond/zincblend structure consist of two interpenetrating fcc lattices. This can be viewed as an fcc with a two-point basis $\left\{0, \frac{a}{4} (\vec{e}_x + \vec{e}_y + \vec{e}_z)\right\}$.

For the diamond structure both atoms are the same, e.g. for C, Ge or Si, but for the zincblend structure these are different as in $A_{\text{III}}B_{\text{V}}$, e.g. for GaAs or ZnS.

TODO: Abb unit cube of zincblend structure

Consider now the specific example GaAs:

- Ga: $[\text{Ar}] 3d^{10}4s^24p^1$ has two s and one p electrons.
- As: $[\text{Ar}] 3d^{10}4s^24p^3$ has two s and three p electrons.
- Together this are eight electrons.
- The elementary cell consists of one Ga (cation) and one As (anion)
- Basis: $s^c, s^a, p_x^c, p_x^a, p_y^c, p_y^a, p_z^c, p_z^a$
- Nearest neighbors: Remember the secular equation (6.27).
- Overlap:

$$s^c s^a \rightarrow \gamma_{ss}$$

$$s^c p_x^a \rightarrow \gamma_{sp}$$

$$p_x^c p_x^a \rightarrow \gamma_{xx}$$

$$p_x^c p_y^a \rightarrow \gamma_{xy}$$

– Phase factors:

$$\begin{aligned}
 e^{i\vec{q}\cdot\vec{d}_1} + e^{i\vec{q}\cdot\vec{d}_2} + e^{i\vec{q}\cdot\vec{d}_3} + e^{i\vec{q}\cdot\vec{d}_4} &= g_0 & \vec{d}_1 &= \frac{a}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
 e^{i\vec{q}\cdot\vec{d}_1} + e^{i\vec{q}\cdot\vec{d}_2} - e^{i\vec{q}\cdot\vec{d}_3} - e^{i\vec{q}\cdot\vec{d}_4} &= g_1 & \vec{d}_2 &= \frac{a}{4} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \\
 e^{i\vec{q}\cdot\vec{d}_1} - e^{i\vec{q}\cdot\vec{d}_2} + e^{i\vec{q}\cdot\vec{d}_3} - e^{i\vec{q}\cdot\vec{d}_4} &= g_2 & \vec{d}_3 &= \frac{a}{4} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \\
 e^{i\vec{q}\cdot\vec{d}_1} - e^{i\vec{q}\cdot\vec{d}_2} - e^{i\vec{q}\cdot\vec{d}_3} + e^{i\vec{q}\cdot\vec{d}_4} &= g_3 & \vec{d}_4 &= \frac{a}{4} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}
 \end{aligned}$$

With the shorthand notation

$$\begin{aligned}
 {}_1\langle s | \hat{v}_{\text{atom}} | s \rangle_2 &= V_{ss\sigma} =: \gamma_{ss} \\
 {}_1\langle s | \hat{v}_{\text{atom}} | p \rangle_2 &= \frac{V_{sp\sigma}}{\sqrt{3}} =: -\gamma_{sp} \\
 {}_1\langle p_y | \hat{v}_{\text{atom}} | p_y \rangle_2 &= \frac{V_{pp\sigma}}{3} + \frac{2V_{pp\pi}}{3} =: \gamma_{pp} \\
 {}_1\langle p_x | \hat{v}_{\text{atom}} | p_y \rangle_2 &= V_{pp\sigma} \cos^2 \Theta - V_{pp\pi} \cos^2 \Theta = \frac{V_{pp\sigma}}{3} - \frac{V_{pp\pi}}{3} =: \gamma_{xy}
 \end{aligned}$$

we get the LCAO matrix:

$$\mathcal{H} = \begin{matrix} & \begin{matrix} s^c & s^a & p_x^c & p_y^c & p_z^c & p_x^a & p_y^a & p_z^a \end{matrix} \\ \begin{matrix} s^c \\ s^a \\ p_x^c \\ p_x^a \\ p_y^c \\ p_y^a \\ p_z^c \\ p_z^a \end{matrix} & \begin{pmatrix} \varepsilon_s^c & \gamma_{ss}g_0 & 0 & 0 & 0 & \gamma_{sp}g_1 & \gamma_{sp}g_2 & \gamma_{sp}g_3 \\ \gamma_{ss}g_0^* & \varepsilon_s^a & -\gamma_{sp}g_1^* & -\gamma_{sp}g_2^* & -\gamma_{sp}g_3^* & 0 & 0 & 0 \\ 0 & -\gamma_{sp}g_1 & \varepsilon_p^c & -0 & 0 & \gamma_{xx}g_0 & \gamma_{xy}g_3 & \gamma_{xy}g_2 \\ 0 & -\gamma_{sp}g_2 & 0 & \varepsilon_p^c & 0 & \gamma_{xy}g_3 & \gamma_{xx}g_0 & \gamma_{xy}g_1 \\ 0 & -\gamma_{sp}g_3 & 0 & 0 & \varepsilon_p^c & \gamma_{xy}g_2 & \gamma_{xy}g_1 & \gamma_{xx}g_0 \\ \gamma_{sp}g_1^* & 0 & \gamma_{xx}g_0^* & \gamma_{xy}g_3^* & \gamma_{xy}g_2^* & \varepsilon_p^a & 0 & 0 \\ \gamma_{sp}g_2^* & 0 & \gamma_{xy}g_3^* & \gamma_{xx}g_0^* & \gamma_{xy}g_1^* & 0 & \varepsilon_p^a & 0 \\ \gamma_{sp}g_3^* & 0 & \gamma_{xy}g_2^* & \gamma_{xy}g_1^* & \gamma_{xx}g_0^* & 0 & 0 & \varepsilon_p^a \end{pmatrix} \end{matrix}$$

– No general analytic solution is known. Only at some high symmetry points, e.g. the Γ point, an analytic solution can be constructed, because some g_i are zero.

At the Γ point, we have $g_1 = g_2 = g_3 = 0$ and $g_0 = 4$, thus we get:

$$\begin{aligned}
 E &= \frac{\varepsilon_s^c + \varepsilon_s^a}{2} \pm \sqrt{\left(\frac{\varepsilon_s^c - \varepsilon_s^a}{2}\right)^2 + (4\gamma_{ss})^2} && \text{two } s \text{ states} \\
 E &= \frac{\varepsilon_p^c + \varepsilon_p^a}{2} \pm \sqrt{\left(\frac{\varepsilon_p^c - \varepsilon_p^a}{2}\right)^2 + (4\gamma_{xx})^2} && \text{two threefold degenerate } p \text{ states}
 \end{aligned}$$

– For Ge ($[\text{Ar}] 3d^{10}4s^24p^2$) there are eight orbitals, so there are eight energy bands.

– The zincblende bands at the Γ -point are highly degenerate.

TODO: Abb Ge-band structure

6.4. Almost free electrons

See integrated course IIa.

The point of view complimentary to the LCAO one, is to consider the periodic potential as a weak perturbation. In this situation it is convenient to express the Bloch waves as linear combinations of (*delocalized*) *plane waves*, rather than of localized atomic orbitals.

Let us expand the wave function in the plane wave basis $|\vec{k}, \sigma\rangle$. (In the following χ_σ is the spin part, $\chi_\uparrow = (1, 0)$ and $\chi_\downarrow = (0, 1)$, and $\varepsilon_{\vec{k}}^0$, the free energy of the free electron.)

$$\begin{aligned}\psi_\sigma(\vec{r}) &= \sum_{\vec{k}'} c_{\vec{k}'} \langle \vec{r} | \vec{k}' \sigma \rangle = \sum_{\vec{k}'} c_{\vec{k}'} \psi_{\vec{k}', \sigma}^0(\vec{r}) = \frac{1}{\sqrt{V}} \sum_{\vec{k}'} c_{\vec{k}'} e^{i\vec{k}' \cdot \vec{r}} \chi_\sigma \\ \Rightarrow \quad \langle \vec{k}, \sigma | \hat{h} | \psi_\sigma \rangle &= \sum_{\vec{k}'} c_{\vec{k}'} \langle \vec{k}, \sigma | -\frac{\hbar^2}{2m} \nabla^2 + v(\vec{r}) | \vec{k}' \sigma' \rangle = \\ &= \sum_{\vec{k}'} c_{\vec{k}'} \left(\varepsilon_{\vec{k}'}^0 \delta_{\vec{k} \vec{k}'} + \sum_{\vec{G} \in \text{RL}} \tilde{v}(\vec{G}) \delta_{\vec{k}, \vec{k}'+\vec{G}} \right) \delta_{\sigma\sigma'}\end{aligned}$$

Here we used the Fourier transform of the potential:

$$\begin{aligned}v(\vec{r}) &= \sum_{\vec{G} \in \text{RL}} \tilde{v}(\vec{G}) e^{i\vec{G} \cdot \vec{r}} \\ \tilde{v}(\vec{k}) &= \frac{1}{V_{\text{cell}}} \int d^3r e^{-i\vec{k} \cdot \vec{r}} v(\vec{r})\end{aligned}$$

Now we construct the eigenvalue equation:

$$\begin{aligned}c_{\vec{k}} \varepsilon_{\vec{k}}^0 + \sum_{\vec{G} \in \text{RL}} \tilde{v}(\vec{G}) c_{\vec{k}-\vec{G}} &= \varepsilon_n(\vec{k}) c_{\vec{k}} \quad (6.28a) \\ c_{\vec{k}-\vec{G}'} \varepsilon_{\vec{k}-\vec{G}'}^0 + \sum_{\vec{G}''=\vec{G}+\vec{G}' \in \text{RL}} \tilde{v}(\vec{G}-\vec{G}') c_{\vec{k}-\vec{G}''} &= \varepsilon_n(\vec{k}) c_{\vec{k}-\vec{G}'} \\ \sum_{\vec{G} \in \text{RL}} c_{\vec{k}-\vec{G}} \left(\left(\varepsilon_{\vec{k}-\vec{G}}^0 - \varepsilon_n(\vec{k}) \right) \delta_{\vec{G}, \vec{G}'} + \tilde{v}(\vec{G}-\vec{G}') \right) &= 0\end{aligned}$$

This gives the eigenvalue equation:

$$\det \left(\left(\varepsilon_{\vec{k}-\vec{G}}^0 - \varepsilon_n(\vec{k}) \right) \delta_{\vec{G}, \vec{G}'} + \tilde{v}(\vec{G}-\vec{G}') \right) = 0 \quad (6.28b)$$

This is exact. However, the efficiency of this approach crucially depends on the convergence of the plane wave expansion or on the strength of the periodic potential. In other words, the kinetic term should dominate over the non-diagonal potential term.

Let us consider the case in which the free electron picture is only slightly perturbed. Consider for simplicity the one dimensional case:

1. Draw the free electron parabola and identify the Brillouin zones.
TODO: Abb parabola with zone scheme
2. Take into account the periodicity α of the dispersion relation to obtain a reduced zone scheme.
TODO: Abb parabola with reduced zone scheme

3. The degeneracy

$$\tilde{\varepsilon}_{\vec{q}=\frac{\vec{G}}{2}}^0 = \tilde{\varepsilon}_{\vec{q}=\frac{\vec{G}}{2}-\vec{G}}^0$$

at Bragg planes is removed:

$$\varepsilon\left(\vec{q} = \frac{\vec{G}}{2}\right) = \varepsilon_{\frac{\vec{G}}{2}}^0 \pm \left|\tilde{v}(\vec{G})\right|$$

The energy gaps (energy bands) are formed.

4. At the zone boundary holds:

$$\begin{aligned} \frac{\partial}{\partial \vec{q}} \varepsilon(\vec{q}) &= \frac{\hbar^2}{2m} \left(\vec{q} - \frac{\vec{G}}{2} \right) \\ \Rightarrow v(\vec{q}) &= \frac{\partial}{\hbar \partial \vec{q}} \varepsilon(\vec{q}) \Big|_{\vec{q}=\frac{\vec{G}}{2}} = 0 \end{aligned}$$

So the group velocity vanishes in this case and standing waves are formed.

In one dimension holds:

- The most it can occur is twofold degeneracy.
- To leading order in the potential, the free electron description remains correct except at Bragg planes.
- Within an extended-zone scheme one finds:
TODO: Abb extended zone scheme with the first two bands

The main outcome is:

- The periodic potential is treated as a perturbation.
- Main effect: It (often) removes degeneracies at high symmetry points (at Bragg planes).

Example in one dimension: repeated zone scheme

TODO: Abb34

$$\left. \frac{\partial \varepsilon}{\partial k} \right|_{\frac{\vec{G}}{2}} = 0$$

6.5. Ground state of Bloch electrons

The ground state of N_e Bloch electrons is constructed by increasingly filling one-electron levels.

- For each band $\varepsilon_n(\vec{k})$, including spin, $2N_{\text{cell}}$ electrons are allowed.
- The Bravais lattice has n atoms per primitive cell, and thus we have:

$$N_{\text{at}} = nN_{\text{cell}}$$

- Moreover, there are $N_e = N_{\text{at}}Z_e$ valence electrons in total, where Z_e is the number of valence electrons per atom.

Two situations may occur:

1. Some bands are completely filled. All others remain empty.

- The energy difference between the lowest unoccupied level and the highest occupied one is called *band gap*.
- A band gap might occur only if the number of electrons per primitive cell is even.
- Filled bands are inert (cf. ASHCROFT, chapter 12). So at $T = 0$ these are insulators. If the band gap is rather small, one speaks of semi conductors.

2. A number of bands is partially filled.

- The energy of the highest occupied level is called *Fermi energy*. It lies within the energy range of one or more bands.
- The *Fermi surface* separates filled from unfilled levels.

$$\varepsilon_n(\vec{k}) = \varepsilon_F$$

Because $\varepsilon_n(\vec{k})$ has the periodicity of the reciprocal lattice, also ε_F is a k -space surface with such periodicity.

Note: In insulators, the Fermi energy is by definition (as $\varepsilon_F = \lim_{T \rightarrow 0} \mu$) in the middle of the band gap.

- If nZ_e is odd the system is a metal (some bands are partly filled).
- If nZ_e is even, the system is a metal or an insulator depending on the band overlap and on interaction defects. (Example: divalent metals)

TODO: Abb even-odd rule for narrow bands and wide bands

For weak potentials, the Fermi surface can be evaluated starting from the free electron sphere centered at $\vec{k} = 0$, noticing that this will be modified near the Bragg planes.

$$\frac{\partial}{\partial \vec{q}} \varepsilon(\vec{q}) \Big|_{\vec{q} = \frac{\vec{G}}{2}} = 0$$

So the surface of constant energy is orthogonal to a Bragg plane.

TODO: Abb fermi sphere of free electron and modified by Bragg plane

Example: Square lattice

- Bragg planes are the planes intersecting the lines joining the origin and reciprocal lattice points.
- The n -th Brillouin zone is the set of points that can be reached from the first Brillouin zone by crossing $n - 1$ Bragg planes, but no fewer.

TODO: Abb The Nearest through Fifth nearest Neighbors for a point in a square lattice and their Bragg lines

For example consider the first three Brillouin zones and the Fermi surface of this lattice.

TODO: Abb first three BZs; Blue circle: free electron Fermi surface for some electron concentration

The first Brillouin zone is entirely occupied, while the second, third and fourth are only partially occupied.

6.6. Alkali and noble metals

Monovalent metals

Alkali metals (body-centered cubic)	Noble metals (face-centered cubic)
Li: [He] $2s^1$	
Na: [Ne] $3s^1$	
K: [Ar] $4s^1$	Cu: [Ar] $3d^{10}4s^1$
Rb: [Kr] $5s^1$	Ag: [Kr] $4d^{10}5s^1$
Cs: [Xe] $6s^1$	Au: [Xe] $5d^{10}6s^1$

Alkali metals

They have one valence electron and form a bcc lattice in direct space. It holds:

$$\frac{k_F^3}{3\pi^2} = n_e = \frac{N_e}{V} = \frac{2}{a^3}$$

The factor 2 is due to the fact that the bcc lattice has two atoms per conventional cubic cell. This gives:

$$k_F \approx 0,620 \cdot \frac{2\pi}{a}$$

TODO: Abb 1. Brillouin zone: rhombic dodecahedron

The shortest distance from the center Γ of the first Brillouin zone to the zone face is in the direction $[110]$ to the point N :

$$\overline{\Gamma N} = \frac{2\pi}{a} \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + 0} \approx 0,707 \cdot \frac{2\pi}{a} > k_F$$

The prefactor is due to the fact, that $\frac{4\pi}{a}$ is the side of the conventional cubic cell for the fcc lattice.

TODO: Abb Fermi surface; almost no distortion from the sphere

Noble metals

Neglecting the d -electrons³, the noble metals have one valence electron. The direct space lattice is fcc.

$$\frac{k_F^3}{3\pi^2} = n_e = \frac{N_e}{V} = \frac{4}{a^3}$$

The factor 4 is due to the fact that the fcc lattice has four atoms per conventional cubic cell. This gives:

$$k_F \approx 0,782 \cdot \frac{2\pi}{a}$$

³The d -electrons are only relevant for optical properties.

The shortest distance from the center Γ of the first Brillouin zone to the zone face is in the direction $[111]$ to the point L :

$$\overline{\Gamma L} = \frac{2\pi}{a} \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} \approx 0,866 \cdot \frac{2\pi}{a} \approx 1,108 \cdot k_F$$

The prefactor is due to the fact, that $\frac{4\pi}{a}$ is the side of the conventional cubic cell for the bcc lattice.

TODO: Abb Fermi surface for noble metal

Here the Fermi surface lies nearly within the first Brillouin zone, but touches the zone face near L .

TODO: Abb Periodic table of Fermi surfaces; IA: bcc, IIA: Be, Mg: hcp, Ca, Sr: Bcc, Ba: fcc, IB: fcc, IIB: Zn, Cd: hcp, Hg: rhombohedral

Conclusion

For simple metals (alkali and noble metals), one can usually forget the band structure and reduce the problem to that of *free particles*.

6.7. Semiconductors

See integrated course IIa.

Typical situation:

TODO: Abb35

Conclusion

Since all relevant properties of semiconductors are due to excited electrons around the minima of the conduction band or to holes around the maxima of the valence band, one can treat electrons and holes as free particles possessing an effective mass, which can even be negative, defined by:

$$\left(\frac{1}{m^*(n)}\right)_{\alpha\alpha'} := \frac{1}{\hbar^2} \frac{\partial^2 \varepsilon_n(\vec{k})}{\partial k_\alpha \partial k_{\alpha'}}$$

7 Second quantization

Why the need for second quantization?

- The indistinguishability of particles requires that the wave functions must be symmetrized or antisymmetrized. This becomes cumbersome if N is large.
- The first quantization approach is tailored to problems with fixed N .

Refresh the first quantization approach to N -particles systems (cf. section 4.3):

- i) Consider a set of eigenfunctions $|\lambda\rangle = |\psi_\lambda\rangle$ of a single particle Hamiltonian \hat{h}_{sp} :

$$\hat{h}_{\text{sp}} |\psi_\lambda\rangle = \varepsilon_\lambda |\psi_\lambda\rangle$$

Or consider equivalently the set:

$$\{\psi_\lambda(\vec{r}) = \langle \vec{r} | \psi_\lambda \rangle = \langle \vec{r} | \lambda \rangle\}$$

- ii) Any N -particles wave function can be built from the complete orthonormal basis $\{\psi_\lambda(\vec{r})\}$. The N -particles wave function reads:

$$|\lambda_{\nu_1}, \lambda_{\nu_2}, \dots, \lambda_{\nu_N}\rangle = \frac{1}{\sqrt{N!}} \cdot \frac{1}{\sqrt{\prod_{\lambda=0}^{\infty} n_\lambda!}} \sum_{p \in \mathfrak{S}(N)} (-\xi)^{\frac{1-\text{sign}(p)}{2}} |\lambda_{p_1}\rangle \otimes \dots \otimes |\lambda_{p_N}\rangle$$

$$\xi = \begin{cases} +1 & \text{for fermions} \\ -1 & \text{for bosons} \end{cases}$$

- iii) *Occupation number representation*: The basis states for an N -particle system are obtained by listing the occupation number of each basis state.

$$\left\{ |n_{\lambda_1}, n_{\lambda_2}, \dots\rangle \left| \sum_j n_{\lambda_j} = N \right. \right\}$$

Note: The sum over j is determined by the dimension of the single particle Hilbert space. It holds:

$$|\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle = |n_{\lambda_1}, n_{\lambda_2}, \dots\rangle \quad (7.1)$$

For example this means:

$$|\lambda_{\nu_1} = 1, \lambda_{\nu_2} = 3, \lambda_{\nu_3} = 1, \lambda_{\nu_4} = 0, \lambda_{\nu_5} = 5\rangle = |1, 2, 0, 1, 0, 1, 0, 0, \dots\rangle$$

7.1. The formalism of second quantization

- The so-called second quantization formalism is based on the occupation number representation.
- The *basis states* as well as each (many-body) *operator* are uniquely determined by specifying *annihilation/creation* operators, which fulfill *bosonic (fermionic) commutation relations*.

We begin with some abstract definitions:

- Introduce a reference state $|0\rangle := |0, 0, \dots\rangle$ called *the vacuum state*.
- Introduce a set of operators \hat{a}_λ and their adjoints \hat{a}_λ^\dagger such that holds:

$$\hat{a}_\lambda |0\rangle = 0$$

$$\mathcal{N} \prod_{i=1}^N \hat{a}_{\lambda_{\nu_i}}^\dagger |0\rangle = \mathcal{N} \hat{a}_{\lambda_{\nu_1}}^\dagger \cdot \dots \cdot \hat{a}_{\lambda_{\nu_N}}^\dagger |0\rangle := |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle$$

\mathcal{N} is a normalization factor with

$$\mathcal{N}^{-1} = \sqrt{\prod_\lambda n_\lambda}$$

and n_λ is the number of times the state $|\lambda\rangle$ appears in the family $(|\lambda_{\nu_1}\rangle, \dots, |\lambda_{\nu_N}\rangle)$. It holds $n_\lambda \in \mathbb{N}$ for bosons and $n_\lambda \in \{0, 1\}$ for fermions. As we shall proof holds

$$|\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle = |n_{\lambda_1}, n_{\lambda_2}, \dots\rangle \quad (7.2)$$

as in (7.1).

- In order to care for the symmetry of the so constructed states, \hat{a}_λ and \hat{a}_λ^\dagger have to fulfill the following commutation relations:

$$[\hat{a}_\lambda, \hat{a}_\mu^\dagger]_\xi := \hat{a}_\lambda \hat{a}_\mu^\dagger + \xi \hat{a}_\mu^\dagger \hat{a}_\lambda = \delta_{\mu\nu} \quad (7.3)$$

$$[\hat{a}_\lambda, \hat{a}_\mu]_\xi = [\hat{a}_\lambda^\dagger, \hat{a}_\mu^\dagger]_\xi = 0 \quad (7.4)$$

To understand (7.4) observe:

$$|\lambda, \mu\rangle = \mathcal{N} \hat{a}_\lambda^\dagger \hat{a}_\mu^\dagger |0\rangle$$

$$|\mu, \lambda\rangle = \mathcal{N} \hat{a}_\mu^\dagger \hat{a}_\lambda^\dagger |0\rangle \stackrel{(7.4)}{=} -\xi |\lambda, \mu\rangle$$

These are the proper symmetry relations.

On the other hand (7.3) ensures orthonormality:

$$\begin{aligned} \delta_{\lambda\mu} &= \langle 0 | \hat{a}_\lambda \hat{a}_\mu^\dagger | 0 \rangle = \langle 0 | (-\xi) \hat{a}_\mu^\dagger \hat{a}_\lambda + [\hat{a}_\lambda, \hat{a}_\mu^\dagger]_\xi | 0 \rangle = \\ &= \underbrace{\langle 0 | (-\xi) \hat{a}_\mu^\dagger \hat{a}_\lambda | 0 \rangle}_{=0} + \langle 0 | [\hat{a}_\lambda, \hat{a}_\mu^\dagger]_\xi | 0 \rangle = \langle 0 | [\hat{a}_\lambda, \hat{a}_\mu^\dagger]_\xi | 0 \rangle \end{aligned}$$

Under the prescriptions i) - iii), *any* N -body state can be generated by the application of a set of N independent creation operators \hat{a}_λ^\dagger to a vacuum state. This is the completeness of the second quantization approach.

- iv) Define now \mathcal{F}_N as the Hilbert space of states with fixed particle number N , i.e. the linear span of all basis states with $|\lambda_{\nu_1}, \lambda_{\nu_2}, \dots, \lambda_{\nu_N}\rangle$. The space

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathcal{F}_N$$

containing *all* many-body states is called *Fock space*. While the operator algebra of \hat{a}_λ and \hat{a}_λ^\dagger does not close in \mathcal{F}_N , it does in \mathcal{F} .

TODO: Abb36

For fermions holds:

$$\left(\hat{a}_\lambda^\dagger\right)^2 = 0 = \left(\hat{a}_\lambda\right)^2$$

- v) Define the occupation number operator:

$$\hat{n}_\lambda = \hat{a}_\lambda^\dagger \hat{a}_\lambda$$

For it holds:

$$\hat{n}_\lambda \left(\hat{a}_\lambda^\dagger\right)^{n_\lambda} |0\rangle = n_\lambda \left(\hat{a}_\lambda^\dagger\right)^{n_\lambda} |0\rangle$$

This means that $\left(\hat{a}_\lambda^\dagger\right)^{n_\lambda} |0\rangle$ is an eigenstate of \hat{n}_λ with eigenvalue n_λ . Moreover we have:

$$\hat{n}_\lambda |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle = \mathcal{N} \hat{a}_\lambda^\dagger \hat{a}_\lambda \prod_{i=1}^N \hat{a}_{\lambda_{\nu_i}}^\dagger |0\rangle = \sum_{i=1}^N \delta_{\lambda \lambda_{\nu_i}} |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle$$

So $|\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle$ is an eigenstate of \hat{n}_λ with eigenvalue n_λ , which proofs the equality (7.2).

- vi) To change the basis, observe:

$$\begin{aligned} \hat{I} &= \sum_{\lambda} |\lambda\rangle \langle \lambda| \\ \Rightarrow \hat{a}_\lambda^\dagger |0\rangle &= |\tilde{\lambda}\rangle = \sum_{\lambda} |\lambda\rangle \langle \lambda | \tilde{\lambda} \rangle = \left(\sum_{\lambda} \hat{a}_\lambda^\dagger \langle \lambda | \tilde{\lambda} \rangle \right) |0\rangle \end{aligned}$$

This gives:

$$\boxed{\hat{a}_\lambda^\dagger = \sum_{\lambda} \hat{a}_\lambda^\dagger \langle \lambda | \tilde{\lambda} \rangle} \quad \boxed{\hat{a}_\lambda = \sum_{\lambda} \hat{a}_\lambda \langle \tilde{\lambda} | \lambda \rangle} \quad (7.5)$$

Example: Transformation from the coordinate to momentum representation in one dimension for a system of length L goes as follows:

$$\begin{aligned} \hat{a}_{\tilde{\lambda}} &= \hat{a}(x) & \hat{a}_\lambda &= \hat{a}_k \\ \Rightarrow \hat{a}(x) &= \sum_k \hat{a}_k \langle x | k \rangle = \frac{1}{\sqrt{L}} \sum_k \hat{a}_k e^{ikx} \\ \hat{a}_k &= \int_0^L dx \langle k | x \rangle \hat{a}(x) = \frac{1}{\sqrt{L}} \int_0^L dx e^{-ikx} \hat{a}(x) \end{aligned}$$

Note:

$$\left[\hat{a}(\vec{r}_1), \hat{a}^\dagger(\vec{r}_2) \right]_\xi = \delta(\vec{r}_1 - \vec{r}_2)$$

- vii) In second quantization, every operator can be written in terms of the creation and annihilation operators.

7.2. Representation of one-body and two-body operators in second quantization

One-body operators \hat{O}_1 acting in the N -particle Hilbert space \mathcal{F}_N are defined as:

$$\hat{O}_1 = \sum_{i=1}^N \hat{o}_i$$

Here \hat{o}_i is an operator acting on particle i .

Examples

– Kinetic operator:

$$\hat{T} = \sum_{i=1}^N \hat{t}_i = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m} = -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2$$

– One particle potential operator:

$$\hat{V} = \sum_{i=1}^N v(\vec{r}_i) = \sum_{i=1}^N \hat{v}_i$$

7.2.1. One-body operators in second quantization

In general in first quantization holds:

$$\hat{o} = \sum_{\lambda_1 \lambda_2} o_{\lambda_1 \lambda_2} |\lambda_1\rangle \langle \lambda_1| \quad o_{\lambda_1 \lambda_2} = \langle \lambda_1 | \hat{o} | \lambda_2 \rangle$$

Consider a one-body operator that is *diagonal* in the basis ($|\lambda\rangle$), i.e.:

$$\hat{O}_1 = \sum_{i=1}^N \hat{o}_i \quad \hat{o} = \sum_{\lambda} o_{\lambda} |\lambda\rangle \langle \lambda| \quad o_{\lambda} = \langle \lambda | \hat{o} | \lambda \rangle$$

Now we calculate:

$$\begin{aligned} & \left\langle \lambda_{\nu'_1}, \lambda_{\nu'_2}, \dots, \lambda_{\nu'_N} \left| \hat{O}_1 \right| \lambda_{\nu_1}, \lambda_{\nu_2}, \dots, \lambda_{\nu_N} \right\rangle \\ &= \left\langle \lambda_{\nu'_1}, \lambda_{\nu'_2}, \dots, \lambda_{\nu'_N} \left| \sum_{i=1}^N \sum_{\lambda} o_{\lambda} |\lambda\rangle_i \langle \lambda|_i \right| \lambda_{\nu_1}, \lambda_{\nu_2}, \dots, \lambda_{\nu_N} \right\rangle = \\ &= \left\langle \lambda_{\nu'_1}, \lambda_{\nu'_2}, \dots, \lambda_{\nu'_N} \left| \sum_{i=1}^N \sum_{\lambda} o_{\lambda} \delta_{\lambda \lambda_{\nu_i}} \right| \lambda_{\nu_1}, \lambda_{\nu_2}, \dots, \lambda_{\nu_N} \right\rangle \end{aligned}$$

Here we wrote short $|\lambda\rangle_i \langle \lambda|_i$ for the fact that the operator acts on the i -th component of $|\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle$. Using

$$\hat{n}_{\lambda} |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle = \sum_{i=1}^N \delta_{\lambda \lambda_{\nu_i}} |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle$$

we get:

$$\left\langle \lambda_{\nu'_1}, \dots, \lambda_{\nu'_N} \left| \hat{O}_1 \right| \lambda_{\nu_1}, \dots, \lambda_{\nu_N} \right\rangle = \left\langle \lambda_{\nu'_1}, \dots, \lambda_{\nu'_N} \left| \sum_{\lambda} o_{\lambda} \hat{n}_{\lambda} \right| \lambda_{\nu_1}, \dots, \lambda_{\nu_N} \right\rangle$$

This gives:

$$\boxed{\hat{O}_1 = \sum_{\lambda} o_{\lambda} \hat{n}_{\lambda} = \sum_{\lambda} o_{\lambda} \hat{a}_{\lambda}^{\dagger} \hat{a}_{\lambda}} \quad (7.6)$$

For a generic basis $|\tilde{\lambda}\rangle$ we know:

$$\begin{aligned} \hat{a}_{\tilde{\lambda}} &= \sum_{\lambda} \hat{a}_{\lambda} \langle \tilde{\lambda} | \lambda \rangle & \hat{a}_{\tilde{\lambda}}^{\dagger} &= \sum_{\lambda} \hat{a}_{\lambda}^{\dagger} \langle \tilde{\lambda} | \lambda \rangle \\ \hat{a}_{\lambda} &= \sum_{\tilde{\eta}} \hat{a}_{\tilde{\eta}} \langle \lambda | \tilde{\eta} \rangle & \hat{a}_{\lambda}^{\dagger} &= \sum_{\tilde{\nu}} \hat{a}_{\tilde{\nu}}^{\dagger} \langle \tilde{\nu} | \lambda \rangle \end{aligned}$$

With $o_{\tilde{\nu}, \tilde{\eta}} := \langle \tilde{\nu}, \lambda \rangle \langle \lambda, \tilde{\eta} \rangle$ follows:

$$\boxed{\hat{O}_1 = \sum_{\tilde{\eta}, \tilde{\nu}} o_{\tilde{\nu}, \tilde{\eta}} \hat{a}_{\tilde{\nu}}^{\dagger} \hat{a}_{\tilde{\eta}}} \quad (7.7)$$

Example: Kinetic operator

$$\hat{T} = \sum_{i=1}^N \hat{t}_i = \sum_{\mu\nu} t_{\mu\nu} \hat{a}_{\mu}^{\dagger} \hat{a}_{\nu}$$

i) Position representation:

$$\begin{aligned} t_{\mu\nu} &= \langle \mu | \hat{t} | \nu \rangle = \int d^3r \int d^3r' \langle \mu | \vec{r} \rangle \langle \vec{r} | \hat{t} | \vec{r}' \rangle \langle \vec{r}' | \nu \rangle = \\ &= \int d^3r \int d^3r' \psi_{\mu}^*(\vec{r}) \langle \vec{r} | \hat{t} | \vec{r}' \rangle \psi_{\nu}(\vec{r}') = \\ &= \int d^3r \psi_{\mu}^*(\vec{r}) \left(-\frac{\hbar^2}{2m} \nabla_r^2 \right) \psi_{\nu}(\vec{r}) \\ \Rightarrow \quad \hat{T} &= \sum_{\mu, \nu} t_{\mu\nu} \hat{a}_{\mu}^{\dagger} \hat{a}_{\nu} = \int d^3r \underbrace{\left(\sum_{\mu} \psi_{\mu}^*(\vec{r}) \hat{a}_{\mu}^{\dagger} \right)}_{=: \hat{a}^{\dagger}(\vec{r})} \left(-\frac{\hbar^2}{2m} \nabla_r^2 \right) \underbrace{\left(\sum_{\nu} \psi_{\nu}(\vec{r}) \hat{a}_{\nu} \right)}_{=: \hat{a}(\vec{r})} = \\ &= \int d^3r \hat{a}^{\dagger}(\vec{r}) \left(-\frac{\hbar^2}{2m} \nabla_r^2 \right) \hat{a}(\vec{r}) \end{aligned}$$

ii) Momentum representation:

$$\hat{a}^{\dagger}(\vec{r}) = \sum_{\vec{k}} \langle \vec{r} | \vec{k} \rangle^* \hat{a}_{\vec{k}}^{\dagger} = \frac{1}{\sqrt{V}} \sum_{\vec{k}} e^{-i\vec{k} \cdot \vec{r}} \hat{a}_{\vec{k}}^{\dagger}$$

$$\begin{aligned} \Rightarrow \quad \hat{T} &= -\frac{1}{V} \sum_{\vec{k}, \vec{k}'} \int d^3r e^{-i\vec{k} \cdot \vec{r}} \hat{a}_{\vec{k}}^\dagger \frac{\hbar^2}{2m} \nabla_r^2 e^{i\vec{k}' \cdot \vec{r}} \hat{a}_{\vec{k}'} = \\ &= \frac{1}{V} \int d^3r e^{i\vec{k} \cdot \vec{r}} = \delta_{\vec{k}, \vec{0}} \sum_{\vec{k}} \frac{\hbar^2 k^2}{2m} \hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}} \end{aligned}$$

Note: For particles with spin holds:

$$\hat{T} = \sum_{\vec{k}, \sigma} \frac{\hbar^2 k^2}{2m} \hat{a}_{\vec{k}\sigma}^\dagger \hat{a}_{\vec{k}\sigma} = \sum_{\sigma} \int d^3r \hat{a}_{\sigma}^\dagger(\vec{r}) \left(-\frac{\hbar^2}{2m} \nabla_r^2 \right) \hat{a}_{\sigma}(\vec{r})$$

Example: Density operator

$$\hat{\rho}_{\text{tot}} := \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i)$$

In the position representation we get:

$$\rho_{\text{tot}}(\vec{r}) = \hat{a}^\dagger(\vec{r}) \hat{a}(\vec{r})$$

In the momentum representation this gives:

$$\rho_{\text{tot}}(\vec{k}) = \frac{1}{V} \sum_{\vec{k}, \vec{q}} e^{i\vec{q} \cdot \vec{r}} \hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}+\vec{q}}$$

Example: Number operator

Position representation:

$$\hat{N} = \int d^3r \hat{\rho}_{\text{tot}}(\vec{r}) = \int d^3r \hat{a}^\dagger(\vec{r}) \hat{a}(\vec{r})$$

Momentum representation:

$$\hat{N} = \frac{1}{V} \int d^3r \sum_{\vec{q}, \vec{k}} e^{i\vec{q} \cdot \vec{r}} \hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}+\vec{q}} = \sum_{\vec{k}} \hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}} = \sum_{\vec{k}} \hat{n}_{\vec{k}} \quad (7.8)$$

$$\hat{N} |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle = \sum_{\lambda_i} \hat{n}_{\lambda_i} |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle = \underbrace{\sum_{\lambda_i} n_{\lambda_i}}_{=N} |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle = N |\lambda_{\nu_1}, \dots, \lambda_{\nu_N}\rangle$$

Example: One-particle Hamiltonian

$$\hat{H} = \sum_{i=1}^N \hat{h}_i = \sum_{i=1}^N \left(\frac{\hat{p}_i^2}{2m} + v(\vec{r}_i) \right) \quad v(\vec{r}_i + \vec{R}) = v(\vec{r}_i)$$

This is a problem for N electrons in a *periodic* potential. The eigenvector basis vectors are the Bloch states:

$$\begin{aligned} \hat{h} |\vec{k}, n\rangle &= \varepsilon_{\vec{k}, n} |\vec{k}, n\rangle \\ \hat{H} &= \sum_{\vec{k}, n} \varepsilon_{\vec{k}, n} \hat{a}_{\vec{k}, n}^\dagger \hat{a}_{\vec{k}, n} \end{aligned}$$

7.2.2. Two-body operators

A two-body operator in an N -particles system is defined in first quantization as:

$$\hat{O}_2 = \frac{1}{2} \sum_{i,j} \hat{o}_{ij}$$

It follows that these operators are fully characterized by their action on two particles.

Consider the matrix element (primed indices are for particle two):

$$o_{\mu\mu'\lambda\lambda'} = \langle \mu\mu' | \hat{o} | \lambda\lambda' \rangle \stackrel{\text{not symmetrized}}{:=} \langle \mu | \otimes \langle \mu' | \hat{o} | \lambda \rangle \otimes | \lambda' \rangle$$

Computing

$$\langle \mu_1 \mu_2 \dots \mu_N | \hat{O}_2 | \lambda_1 \lambda_2 \dots \lambda_N \rangle$$

with the symmetrized states, we get:

$$\hat{O}_2 = \frac{1}{2} \sum_{\lambda\lambda'\mu\mu'} o_{\mu\mu'\lambda\lambda'} \hat{a}_\mu^\dagger \hat{a}_{\mu'}^\dagger \hat{a}_{\lambda'} \hat{a}_\lambda \quad (7.9)$$

TODO: abb37 Feynman-diagram; interpretation of $o_{\mu\mu'\lambda\lambda'}$

Example: Coulomb interaction

$$\hat{V}_{ee} = \frac{1}{2} \sum_{i,j,i \neq j} \frac{e^2}{\|\vec{r}_i - \vec{r}_j\|} = \frac{1}{2} \sum_{i,j,i \neq j} v_{ee}(\vec{r}_i - \vec{r}_j)$$

In the position representation holds:

$$\begin{aligned} \hat{V}_{ee} &= \frac{1}{2} \sum_{\mu,\mu',\lambda,\lambda'} \left(\int d^3r \int d^3r' \psi_\mu^*(\vec{r}) \psi_{\mu'}^*(\vec{r}') v_{ee}(\vec{r} - \vec{r}') \psi_{\lambda'}(\vec{r}') \psi_\lambda(\vec{r}) \right) \hat{a}_\mu^\dagger \hat{a}_{\mu'}^\dagger \hat{a}_{\lambda'} \hat{a}_\lambda = \\ &= \frac{1}{2} \int d^3r \int d^3r' \hat{a}^\dagger(\vec{r}) \hat{a}^\dagger(\vec{r}') v_{ee}(\vec{r} - \vec{r}') \hat{a}(\vec{r}') \hat{a}(\vec{r}) \end{aligned}$$

Including the spin we get:

$$\hat{V}_{ee} = \frac{1}{2} \sum_{\sigma\sigma'} \int d^3r \int d^3r' \hat{a}_\sigma^\dagger(\vec{r}) \hat{a}_{\sigma'}^\dagger(\vec{r}') v(\vec{r} - \vec{r}') \hat{a}_{\sigma'}(\vec{r}') \hat{a}_\sigma(\vec{r})$$

In the momentum representation, the Coulomb interaction can be seen as a scattering process with momentum transfer \vec{q} .

$$\hat{V}_{ee} = \frac{1}{2V} \sum_{\sigma,\sigma'} \sum_{\vec{q}} \sum_{\vec{k},\vec{k}'} \tilde{v}_{ee}(\vec{q}) \hat{a}_{\vec{k}+\vec{q},\sigma}^\dagger \hat{a}_{\vec{k}'-\vec{q},\sigma'}^\dagger \hat{a}_{\vec{k}',\sigma'} \hat{a}_{\vec{k},\sigma} \quad (7.10)$$

Here \tilde{v}_{ee} is the Fourier transformation:

$$\tilde{v}_{ee}(\vec{q}) := \int d^3r e^{i\vec{q} \cdot \vec{r}} \frac{e^2}{r} = \lim_{\eta \rightarrow 0} \int d^3r e^{i\vec{q} \cdot \vec{r}} e^{-\eta r} \frac{e^2}{r} = \frac{4\pi e^2}{q^2} \quad (7.11)$$

TODO: Abb38

8 Interacting electron gas

Let us remember the Hamiltonian of a solid in the adiabatic approximation:

$$\underbrace{\left(\hat{T}_{ee} + \hat{V}_{ee} + \hat{V}_{ei} + \hat{V}_{ii} \right)}_{=\hat{H}_{\text{el}}} \phi_k = \varepsilon_{\text{el}}(k) \phi_k \quad \text{electrons}$$

$$\left(\hat{T}_{\text{ion}} + \varepsilon_{\text{el}}(k) \right) \chi_k = E \chi_k \quad \text{ions}$$

We now remember

$$\hat{H}_{\text{el}} = \hat{T}_{ee} + \hat{V}_{ee} + \hat{V}_{ei} + \hat{V}_{ii}$$

and in a first approximation *fully* neglect the ion:

$$\hat{H}_{\text{el}} \approx \hat{T}_{ee} + \hat{V}_{ee}$$

$$\hat{V}_{ee} = \frac{1}{2V} \sum_{\vec{k}, \vec{k}', \vec{q}} \sum_{\sigma, \sigma'} \tilde{v}_{ee}(q) \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'}^\dagger \hat{a}_{\vec{k}', \sigma'} \hat{a}_{\vec{k}, \sigma}$$

$$\tilde{v}_{ee}(q) = \frac{4\pi e^2}{q^2}$$

Problem: We have a divergence at $q = 0$.

8.1. Jellium model

To remove the divergence, we observe that the positively charged ions do not only provide a periodic potential for the electrons, but also ensure charge neutrality. To this extend we introduced the local charge density:

$$e\tilde{\rho}(\vec{r}) = e(\rho(\vec{r}) - z\rho_{\text{ion}}(\vec{r}))$$

The jellium approximation is:

$$z\rho_{\text{ion}}(\vec{r}) \approx z\rho_{\text{ion}}^{\text{jellium}} := n_e = \frac{N_e}{V}$$

Within this approximation holds:

$$\hat{V}_{ee} + \hat{V}_{ei} + \hat{V}_{ii} \approx \hat{V}_{ee} + \hat{V}_{ei}^{\text{jellium}} + \hat{V}_{ii}^{\text{jellium}} =: \hat{V}_{ee}^{\text{jellium}}$$

Now the very nice thing is that this turns out to cancel the divergence:

$$\hat{V}_{ee}^{\text{jellium}} = \frac{1}{2V} \sum_{\vec{k}, \vec{k}', \vec{q} \neq 0} \sum_{\sigma, \sigma'} \tilde{v}_{ee}(q) \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'}^\dagger \hat{a}_{\vec{k}', \sigma'} \hat{a}_{\vec{k}, \sigma} \quad (8.1)$$

To see this, let us consider the second quantization expressions for \hat{V}_{ei} and \hat{V}_{ii} :

$$\begin{aligned}\hat{V}_{ei} &= \int d^3r d^3r' \hat{\psi}_{\text{el}}^\dagger(\vec{r}) \hat{\psi}_{\text{ion}}^\dagger(\vec{r}') v_{ei}(\vec{r} - \vec{r}') \hat{\psi}_{\text{ion}}(\vec{r}') \hat{\psi}_{\text{el}}(\vec{r}) \\ \hat{V}_{ii} &= \frac{1}{2} \int d^3r d^3r' \hat{\psi}_{\text{ion}}^\dagger(\vec{r}) \hat{\psi}_{\text{ion}}^\dagger(\vec{r}') v_{ii}(\vec{r} - \vec{r}') \hat{\psi}_{\text{ion}}(\vec{r}') \hat{\psi}_{\text{ion}}(\vec{r})\end{aligned}$$

Within the jellium approximation we get:

$$\begin{aligned}\hat{V}_{ei} &\approx \hat{V}_{ei}^{\text{jellium}} = -\frac{\hat{N}_e}{V} \int d^3r \int d^3r' \hat{\rho}_{\text{el}}(\vec{r}) \frac{e^2}{|\vec{r} - \vec{r}'|} + \frac{\hat{N}_e^2}{V^2} \cdot \frac{1}{2} \int d^3r \int d^3r' \frac{e^2}{|\vec{r} - \vec{r}'|} = \\ &= -\frac{\hat{N}_e^2}{V} \lim_{q \rightarrow 0} \tilde{v}_{ee}(q) + \frac{1}{2} \frac{\hat{N}_e^2}{V} \lim_{q \rightarrow 0} \tilde{v}_{ee}(q) = -\frac{1}{2} \frac{\hat{N}_e^2}{V} \lim_{q \rightarrow 0} \tilde{v}_{ee}(q)\end{aligned}$$

On the other hand, the $q = 0$ contribution from \hat{V}_{ee} is:

$$\begin{aligned}&\frac{1}{2V} \lim_{q \rightarrow 0} \tilde{v}_{ee}(q) \sum_{\vec{k}, \vec{k}'} \sum_{\sigma, \sigma'} \hat{a}_{\vec{k}, \sigma}^\dagger \hat{a}_{\vec{k}, \sigma'}^\dagger \hat{a}_{\vec{k}', \sigma'} \hat{a}_{\vec{k}, \sigma} = \\ &\quad \text{Pauli principle} \quad \frac{1}{2V} \lim_{q \rightarrow 0} \tilde{v}_{ee}(q) \sum_{\vec{k}, \sigma} \sum_{(\vec{k}', \sigma') \neq (\vec{k}, \sigma)} \hat{a}_{\vec{k}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'}^\dagger \hat{a}_{\vec{k}', \sigma'} \hat{a}_{\vec{k}, \sigma} = \\ &= \frac{1}{2V} \lim_{q \rightarrow 0} \tilde{v}_{ee}(q) \underbrace{\sum_{\vec{k}, \sigma} \hat{n}_{\vec{k}, \sigma}}_{\rightarrow N_e} \underbrace{\sum_{(\vec{k}', \sigma') \neq (\vec{k}, \sigma)} \hat{n}_{\vec{k}', \sigma'}}_{\rightarrow N_e - 1} \approx \\ &\quad \stackrel{N_e \gg 1}{\approx} \frac{1}{2} \lim_{q \rightarrow 0} \tilde{v}_{ee}(q) \frac{\hat{N}_e^2}{V}\end{aligned}$$

This compensates $\hat{V}_{ei}^{\text{jellium}} + \hat{V}_{ii}^{\text{jellium}}$.

8.2. Energy scales

We have the interplay between kinetic energy of the electrons and Coulomb repulsion. Which energy scale wins for a normal metal?

– It holds:

$$\frac{1}{n_e} = \frac{V}{N_e} \approx \frac{4}{3} \pi r_0^3$$

So the potential energy per particle is:

$$\varepsilon_{\text{pot}} \approx \frac{e^2}{r_0} \sim n_e^{-\frac{1}{3}}$$

– The kinetic energy per particle is the ground state energy per particle of the noninteracting (ideal) Fermi gas (see equation (4.53) and (4.52)):

$$\varepsilon_{\text{kin}} = \frac{E_{\text{ground state}}}{N_{\text{el}}} = \frac{3}{5} \varepsilon_F = \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} = \frac{3}{5} \frac{\hbar^2}{2m} (3\pi^2 n_e)^{\frac{2}{3}} \sim n_e^{\frac{2}{3}}$$

– It follows:

$$\frac{\varepsilon_{\text{pot}}}{\varepsilon_{\text{kin}}} \sim \frac{n_e^{\frac{1}{3}}}{n_e^{\frac{1}{3}}} = n_e^{-\frac{1}{3}} \xrightarrow{n_e \rightarrow \infty} 0$$

The importance of the Coulomb interaction diminishes with density as a consequence of the *Pauli principle*.

– A parameter to quantify the ratio of the two energy scales is the dimensionless quantity r_s (cf. ASHCROFT and MERMIN):

$$r_s := \frac{r_0}{a_0} = \frac{r_0}{\hbar^2 m^{-1} e^{-2}} = \frac{e^2}{r_0} \cdot \frac{m r_0^2}{\hbar^2} \approx \frac{\varepsilon_{\text{pot}}}{\varepsilon_{\text{kin}}}$$

(With $\Delta p \Delta x \geq \hbar$ follows $\Delta p \approx \frac{\hbar}{r_0}$.)

Metal	r_s
Li	3, 2
Na	3, 9
K	4, 9
Cu	2, 7
Be	1, 9

So simple and noble metals have intermediate values of r_s !

8.3. Electron interaction effects in perturbation theory

We consider

$$\hat{H}_e^{\text{jellium}} = \hat{T}_{ee} + \hat{V}_{ee}^{\text{jellium}} \quad (8.2)$$

and look at $\hat{V}_{ee}^{\text{jellium}}$ as a perturbation. We calculate the ground state energy.

Zeroth order

$$E^{(0)} = \langle \text{FS} | \hat{T}_{ee} | \text{FS} \rangle$$

Here $|\text{FS}\rangle$ (Fermi sea) is the ground state for N_e noninteracting fermions:

$$|\text{FS}\rangle = \prod_{\vec{k}_i, \|\vec{k}_i\| \leq k_F} \prod_{\sigma} \hat{a}_{\vec{k}_i, \sigma}^{\dagger} |0\rangle$$

$$\varepsilon_{k_1} \leq \varepsilon_{k_2} \leq \dots \leq \varepsilon_{k_F}$$

x Hence we get:

$$E^{(0)} = \left\langle \text{FS} \left| \sum_{\vec{k}, \sigma} \frac{\hbar^2}{2m} k^2 \hat{n}_{\vec{k}, \sigma} \right| \text{FS} \right\rangle$$

$$\hat{n}_{\vec{k},\sigma} |\text{FS}\rangle = \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} |\text{FS}\rangle = \begin{cases} 1 & k \leq k_F \\ 0 & k > k_F \end{cases}$$

With this follows:

$$\frac{E^{(0)}}{N} = \varepsilon^{(0)} = \sum_{\vec{k},\sigma} \frac{\hbar^2 k^2}{2m} \langle \text{FS} | \hat{n}_{\vec{k},\sigma} | \text{FS} \rangle = \sum_{\|\vec{k}\| \leq k_F, \sigma} \frac{\hbar^2 k^2}{2m} \stackrel{(4.53)}{=} \frac{3}{5} N \varepsilon_F \approx \frac{2,21}{r_s^2} E_{\text{Ry}}$$

The last approximation is a comparison with the Rydberg energy:

$$E_{\text{Ry}} = \frac{e^2}{2a_0} \approx 13,6 \text{ eV}$$

First order

$$\begin{aligned} E^{(1)} &= \langle \text{FS} | \hat{V}_{ee}^{\text{jellium}} | \text{FS} \rangle = \\ &= \frac{1}{2V} \sum_{\vec{q} \neq 0} \sum_{\vec{k}, \vec{k}', \sigma, \sigma'} \tilde{v}_{ee}(\vec{q}) \cdot \langle \text{FS} | \hat{a}_{\vec{k}+\vec{q},\sigma}^\dagger \hat{a}_{\vec{k}',\sigma'}^\dagger \hat{a}_{\vec{k},\sigma} \hat{a}_{\vec{k}',\sigma'} | \text{FS} \rangle \stackrel{k,k' \leq k_F}{\neq} 0 \end{aligned}$$

We also need $\vec{k}' = \vec{k} + \vec{q}$ and $\sigma = \sigma'$ for this not to vanish.

Now follows:

$$\hat{a}_{\vec{k}+\vec{q},\sigma}^\dagger \hat{a}_{\vec{k}',\sigma'}^\dagger \hat{a}_{\vec{k},\sigma} \hat{a}_{\vec{k}',\sigma'} = \hat{a}_{\vec{k}+\vec{q},\sigma}^\dagger \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k}+\vec{q},\sigma'} \hat{a}_{\vec{k},\sigma'} = -\hat{a}_{\vec{k}+\vec{q},\sigma}^\dagger \hat{a}_{\vec{k}+\vec{q},\sigma'} \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma'}$$

Hence we get:

$$\begin{aligned} E^{(1)} &= -\frac{1}{2V} \sum_{\vec{q} \neq 0} \sum_{\vec{k}, \vec{k}', \sigma, \sigma'} \tilde{v}_{ee}(q) \delta_{\sigma\sigma'} \delta_{\vec{k}+\vec{q}, \vec{k}'} \Theta(k_F - \|\vec{k} + \vec{q}\|) \Theta(k_F - \|\vec{k}\|) = \\ &= -\frac{1}{V} \frac{V^2}{(2\pi)^6} \int d\vec{q} \tilde{v}_{ee}(q) (1 - \delta_{\vec{q},0}) \int d\vec{k} \Theta(k_F - \|\vec{k} + \vec{q}\|) \Theta(k_F - \|\vec{k}\|) \end{aligned}$$

For the integration limits consider:

TODO: Abb39

$$\begin{aligned} 0 &\leq q \leq 2k_F \\ \frac{q}{2k_F} &\leq \cos(\theta_k) \leq 1 \\ \frac{q}{2\cos(\theta_k)} &< k < k_F \end{aligned}$$

(see BRUUS and FLENSBERG)

This yields:

$$\begin{aligned} E^{(1)} &= -\frac{4\pi e^2}{2V} \underbrace{2}_{\text{spin}} \underbrace{(4\pi)}_{q\text{-solid angle}} \frac{V}{(2\pi)^3} \int_{0+}^{2k_F} dq \frac{q^2}{q^2} \underbrace{2}_{\text{symmetry}} \underbrace{(2\pi)}_{\text{integration over } \varphi_k} \cdot \\ &\cdot \int_{\frac{a}{2k_F}}^1 d(\cos \theta_k) \frac{V}{(2\pi)^3} \int_{\frac{a}{2\cos(\theta_k)}}^{k_F} dk \cdot k^2 = -\frac{e^2}{2} V \frac{k_F^4}{2\pi^3} \end{aligned}$$

From this we find:

$$\varepsilon^{(0)} + \varepsilon^{(1)} \approx \left(\frac{2,21}{r_s^2} - \frac{0,916}{r_s} \right) E_{\text{Ry}}$$

TODO: Abb40

$$\begin{aligned} r_s^* &= 4,83 \\ \varepsilon^* &= -0,095 \cdot E_{\text{Ry}} = -1,29 \text{ eV} \end{aligned}$$

For comparison look at:

$$r_s^{\text{Na}} = 3,96 \qquad \varepsilon^{\text{Na}} = -0,08 \text{ Ry} = -1,13 \text{ eV}$$

Second order perturbation theory

One could try to improve the result for the ground state energy by going to second order.

Problem: The contribution diverges!

So summation of all orders in the perturbation theory is needed. (To achieve this, one approximates every summand and only takes the most divergent part. To improve the result one can try to also include the next less divergent part, and so on.)

$$E^{(2)} = \sum_{|\nu\rangle \neq |\text{FS}\rangle} \frac{\langle \text{FS} | \hat{V}_{ee}^{\text{jellium}} | \nu \rangle \langle \nu | \hat{V}_{ee}^{\text{jellium}} | \text{FS} \rangle}{E^{(0)} - E_\nu} = E_{\text{dir}}^{(2)} + E_{\text{exchange}}^{(2)}$$

TODO: Abb41

For non-vanishing contributions must hold:

- i) The intermediate states $|\nu\rangle$ are outside of the Fermi sphere.
- ii) Excited electrons must be put back in the holes left behind.

Then holds e.g. for direct contribution:

$$\begin{aligned} E_{\text{dir}}^{(2)} &= \frac{1}{4V^2} \sum_{|\nu\rangle \neq |\text{FS}\rangle} \sum_{\vec{q} \neq 0} \sum_{\vec{k}, \vec{k}', \sigma, \sigma'} \frac{\tilde{v}_{ee}^2(q)}{E^{(0)} - E_\nu} \\ &\quad \cdot \Theta\left(\left\|\vec{k} + \vec{q}\right\| - k_F\right) \Theta\left(\left\|\vec{k}' - \vec{q}\right\| - k_F\right) \Theta\left(k_F - \left\|\vec{k}\right\|\right) \Theta\left(k_F - \left\|\vec{k}'\right\|\right) \sim \\ &\sim \int_{0+} dq \cdot q^2 \frac{1}{q^4} \cdot \frac{1}{q} \cdot q \cdot q = \int_{0+} dq \frac{1}{q} \approx \lim_{\varepsilon \searrow 0} \ln(q) \end{aligned}$$

So we have logarithmic divergence, which is very bad ...

8.4. Mean field approximation

The Idea is:

TODO: Abb42

The individual particle interacts with an average background, depending usually on the density. So we get an *effective* single particle theory.

Basic concepts of mean field theory

Consider two kinds of particles described by the operators \hat{a}_μ and \hat{b}_ν respectively.

i) Assume the interacting Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{V}_{\text{int}}$$

with:

$$\begin{aligned}\hat{H}_0 &= \sum_{\mu} \varepsilon_{\mu} \hat{a}_{\mu}^{\dagger} \hat{a}_{\mu} + \sum_{\nu} \varepsilon_{\nu} \hat{b}_{\nu}^{\dagger} \hat{b}_{\nu} \\ \hat{V}_{\text{int}} &= \sum_{\nu, \nu', \mu, \mu'} v_{\nu \mu \nu' \mu'} \hat{a}_{\nu}^{\dagger} \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} \hat{a}_{\nu'}\end{aligned}$$

ii) Define now with the average $\langle \cdot \rangle$:

$$\begin{aligned}\hat{d}_{\nu \nu'} &= \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu'} - \langle \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu'} \rangle \\ \hat{e}_{\mu \mu'} &= \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} - \langle \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} \rangle\end{aligned}$$

This gives:

$$\begin{aligned}\hat{V}_{\text{int}} &= \sum_{\nu, \nu', \mu, \mu'} v_{\nu \mu \nu' \mu'} \left(\left(\hat{d}_{\nu \nu'} + \langle \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu'} \rangle \right) \left(\hat{e}_{\mu \mu'} + \langle \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} \rangle \right) \right) = \\ &= \sum_{\nu, \nu', \mu, \mu'} v_{\nu \mu \nu' \mu'} \left(\underbrace{\hat{d}_{\nu \nu'} \hat{e}_{\mu \mu'}}_{\text{interaction part}} + \underbrace{\hat{d}_{\nu \nu'} \langle \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} \rangle + \langle \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu'} \rangle \hat{e}_{\mu \mu'} + \langle \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu'} \rangle \langle \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} \rangle}_{\text{only single-particle operators}} \right)\end{aligned}$$

The mean field approximation amounts to neglecting the interaction part.

$$\hat{V}_{\text{int}} \approx \hat{V}_{\text{MF}} = \sum_{\nu, \nu', \mu, \mu'} v_{\nu \mu \nu' \mu'} \left(\hat{d}_{\nu \nu'} \langle \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} \rangle + \langle \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu'} \rangle \hat{e}_{\mu \mu'} + \langle \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu'} \rangle \langle \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} \rangle \right) \quad (8.3)$$

Looking at \hat{V}_{MF} , we can formulate the mean field procedure in a simple way. If the interaction is the product of two operators \hat{A} and \hat{B} , then follows:

$$\begin{aligned}\hat{H}_{AB} &= \hat{A} \hat{B} \\ \Rightarrow \quad \hat{H}_{\text{MF}, AB} &= \hat{A} \langle \hat{B} \rangle + \langle \hat{A} \rangle \hat{B} - \langle \hat{A} \rangle \langle \hat{B} \rangle\end{aligned}$$

The subtraction ensures the correct average:

$$\langle \hat{H}_{\text{MF}, AB} \rangle = \langle \hat{A} \rangle \langle \hat{B} \rangle$$

Note:

$$\langle \hat{V}_{\text{MF}} \rangle = \sum_{\nu, \nu', \mu, \mu'} v_{\nu \mu \nu' \mu'} \langle \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu'} \rangle \langle \hat{b}_{\mu}^{\dagger} \hat{b}_{\mu'} \rangle$$

This would be the exact form, if the two systems would not be interacting.

- How to find the averages?

In general we should understand $\langle \cdot \rangle$ as being related to the mean field Hamiltonian itself $\hat{H}_{\text{MF}} = \hat{H}_0 + \hat{V}_{\text{MF}}$, i.e. for a system in thermal equilibrium:

$$\langle \hat{A} \rangle = \frac{1}{Z_{\text{MF}}} \text{tr} \left(e^{-\beta \hat{H}_{\text{MF}}} \hat{A} \right)$$

One has to solve a *self-consistent problem*. In other words to evaluate $\bar{n}_{\nu\nu'}^a := \langle \hat{a}_\nu^\dagger \hat{a}_{\nu'} \rangle$ we need a self-consistency loop:

$$\bar{n}_{\nu\nu'}^a \rightarrow \hat{H}_{\text{MF}} \rightarrow \bar{n}_{\nu\nu'}^a \rightarrow \dots$$

- The problem is that the self-consistent procedure might still be intractable for a large number of particles.

In practice one has to *assume* something about the averages $\langle \hat{a}_\nu^\dagger \hat{a}_{\nu'} \rangle$.

Example: Use symmetry arguments or look at zero temperature properties or do something else ...

- *Hartree-Fock approximation* (alike particles):

$$\begin{aligned} \hat{H} &= \hat{H}_0 + \hat{V}_{\text{int}} \\ \hat{H}_0 &= \sum_{\mu} \varepsilon_{\mu} \hat{a}_{\mu}^\dagger \hat{a}_{\mu} \\ \hat{V}_{\text{int}} &= \sum_{\mu, \nu, \mu', \nu'} \hat{a}_{\nu}^\dagger \hat{a}_{\mu}^\dagger \hat{a}_{\mu'} \hat{a}_{\nu'} \end{aligned}$$

- *Mean field approximation:*

$$\begin{aligned} \hat{a}_{\nu}^\dagger \hat{a}_{\mu}^\dagger \hat{a}_{\mu'} \hat{a}_{\nu'} &\approx \hat{a}_{\nu}^\dagger \hat{a}_{\nu'} \langle \hat{a}_{\mu}^\dagger \hat{a}_{\mu'} \rangle + \langle \hat{a}_{\nu}^\dagger \hat{a}_{\nu'} \rangle \hat{a}_{\mu}^\dagger \hat{a}_{\mu'} - \langle \hat{a}_{\nu}^\dagger \hat{a}_{\nu'} \rangle \langle \hat{a}_{\mu}^\dagger \hat{a}_{\mu'} \rangle \pm \\ &\quad \underbrace{\pm}_{\text{bosons/fermions}} \left(\hat{a}_{\nu}^\dagger \hat{a}_{\mu'} \langle \hat{a}_{\mu}^\dagger \hat{a}_{\nu'} \rangle + \langle \hat{a}_{\nu}^\dagger \hat{a}_{\mu'} \rangle \hat{a}_{\mu}^\dagger \hat{a}_{\nu'} - \langle \hat{a}_{\nu}^\dagger \hat{a}_{\mu'} \rangle \langle \hat{a}_{\mu}^\dagger \hat{a}_{\nu'} \rangle \right) = \quad (8.4) \\ &=: \text{Hartree} \pm \underbrace{\text{Fock}}_{\text{exchange part for alike particles}} \end{aligned}$$

- *Hartree-Fock for the interacting electron gas:*

$$\hat{T} + \hat{V}_{ee}^{\text{jellium}} = \sum_{\vec{k}, \sigma} \varepsilon_{\vec{k}} \hat{a}_{\vec{k}, \sigma}^\dagger \hat{a}_{\vec{k}, \sigma} + \frac{1}{2V} \sum_{\vec{q} \neq 0} \sum_{\vec{k}, \vec{k}', \sigma, \sigma'} \tilde{v}_{ee}(\vec{q}) \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}'-\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'} \hat{a}_{\vec{k}, \sigma}$$

In the Hartree-Fock approximation holds:

$$\begin{aligned} \hat{V}_{ee}^{\text{jellium}} &\approx \hat{V}_{ee, \text{MF}}^{\text{jellium}} = \\ &= \frac{1}{2V} \sum_{\vec{q} \neq 0} \sum_{\vec{k}, \vec{k}', \sigma, \sigma'} \tilde{v}_{ee}(\vec{q}) \left(\langle \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}, \sigma} \rangle \hat{a}_{\vec{k}'-\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'} + \right. \\ &\quad - \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}, \sigma} \langle \hat{a}_{\vec{k}'-\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'} \rangle - \langle \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}, \sigma} \rangle \langle \hat{a}_{\vec{k}'-\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'} \rangle - \\ &\quad - \left(\langle \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'} \rangle \hat{a}_{\vec{k}'-\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}, \sigma} + \langle \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'} \rangle \hat{a}_{\vec{k}'-\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}, \sigma} - \right. \\ &\quad \left. \left. - \langle \hat{a}_{\vec{k}+\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}', \sigma'} \rangle \langle \hat{a}_{\vec{k}'-\vec{q}, \sigma}^\dagger \hat{a}_{\vec{k}, \sigma} \rangle \right) \right) \end{aligned}$$

- Translational invariance yields:

$$\begin{aligned}
 \langle \hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}'} \rangle &= \frac{1}{V} \int d^3r \int d^3r' e^{i\vec{k} \cdot \vec{r}} e^{-i\vec{k}' \cdot \vec{r}'} \underbrace{\langle \hat{a}_{\vec{k}}^\dagger(\vec{r}) \hat{a}_{\vec{k}'}(\vec{r}') \rangle}_{=f(\vec{r}-\vec{r}')} \\
 &\stackrel{\vec{R}=\vec{r}+\vec{r}'}{=} \frac{1}{V} \int d^3R \int d^3r' e^{i(\vec{k}-\vec{k}') \cdot \vec{R}} e^{i(\vec{k}'\vec{r}-\vec{k} \cdot \vec{r}')} f(\vec{R}-2\vec{r}') = \\
 &= \frac{1}{V} \int d^3r' \delta_{\vec{k},\vec{k}'} e^{i(\vec{k}'\vec{r}-\vec{k} \cdot \vec{r}')} f(\vec{R}-2\vec{r}') = \\
 &\stackrel{\vec{r}''=\vec{R}-2\vec{r}'}{=} \delta_{\vec{k},\vec{k}'} \frac{1}{V} \int d^3r'' f(\vec{r}'') = \langle \hat{n}_{\vec{k}} \rangle \delta_{\vec{k},\vec{k}'},
 \end{aligned}$$

- Assume spin-independence of the interaction (otherwise ferromagnetic solutions are possible):

$$\langle \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} \rangle = \langle \hat{a}_{\vec{k},-\sigma}^\dagger \hat{a}_{\vec{k},-\sigma} \rangle$$

- The Coulomb interaction only depends on the distance:

$$\begin{aligned}
 v_{ee}(\vec{r}) &= v_{ee}(r) \\
 \Rightarrow \tilde{v}_{ee}(\vec{q}) &= \tilde{v}_{ee}(q) = \tilde{v}_{ee}(-q)
 \end{aligned}$$

Due to the $\delta_{\vec{k},\vec{k}'}$ and $\vec{q} \neq 0$ the Hartree term vanishes and we get:

$$\begin{aligned}
 \hat{V}_{ee,\text{MF}}^{\text{jellium}} &= -\frac{1}{2V} \sum_{\vec{q} \neq 0} \sum_{\vec{k},\sigma} \tilde{v}_{ee}(q) \left(\langle \hat{a}_{\vec{k}+\vec{q},\sigma}^\dagger \hat{a}_{\vec{k}+\vec{q},\sigma} \rangle \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} + \right. \\
 &\quad \left. + \hat{a}_{\vec{k}+\vec{q},\sigma}^\dagger \hat{a}_{\vec{k}+\vec{q},\sigma} \langle \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} \rangle - \langle \hat{a}_{\vec{k}+\vec{q},\sigma}^\dagger \hat{a}_{\vec{k}+\vec{q},\sigma} \rangle \langle \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} \rangle \right)
 \end{aligned}$$

Now substitute $\vec{k} + \vec{q} = \vec{q}'$ and write again \vec{q} for \vec{q}' .

$$\begin{aligned}
 \hat{V}_{ee,\text{MF}}^{\text{jellium}} &= -\frac{1}{2V} \sum_{\vec{q} \neq 0} \sum_{\vec{k},\sigma} \tilde{v}_{ee}(\vec{q} - \vec{k}) \left(\langle \hat{a}_{\vec{q},\sigma}^\dagger \hat{a}_{\vec{q},\sigma} \rangle \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} + \right. \\
 &\quad \left. + \hat{a}_{\vec{q},\sigma}^\dagger \hat{a}_{\vec{q},\sigma} \langle \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} \rangle - \langle \hat{a}_{\vec{q},\sigma}^\dagger \hat{a}_{\vec{q},\sigma} \rangle \langle \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} \rangle \right)
 \end{aligned}$$

From $\tilde{v}_{ee}(\vec{q}) = \tilde{v}_{ee}(-\vec{q})$ we get:

$$\boxed{\hat{V}_{ee,\text{MF}}^{\text{jellium}} = \sum_{\vec{k},\sigma} v_{\text{HF}}(\vec{k}) \hat{a}_{\vec{k},\sigma}^\dagger \hat{a}_{\vec{k},\sigma} - 2 \sum_{\vec{k},\sigma} v_{\text{HF}}(\vec{k}) n(\vec{k}) n_{\vec{k},\sigma}} \quad (8.5a)$$

$$v_{\text{HF}}(\vec{k}) = -\frac{1}{V} \sum_{\vec{q} \neq \vec{k}} \tilde{v}_{ee}(\vec{q} - \vec{k}) n_{\vec{q},\sigma} \quad (8.6)$$

All together gives:

$$\hat{T}_e + \hat{V}_{ee,\text{MF}}^{\text{jellium}} = \sum_{\vec{k},\sigma} \left(\varepsilon_{\vec{k}} + v_{\text{HF}}(\vec{k}) \right) \hat{n}_{\vec{k},\sigma} + \text{const.} \quad (8.5b)$$

Consider the Hartree-Fock contribution at $T = 0$ and assume.

$$n_{\vec{k},\sigma} = \Theta(k_F - k)$$

This is appropriate for fermions if $v_{\text{HF}}(k)$ is a monotonously increasing function of k . Then follows:

$$\begin{aligned} v_{\text{HF}}(\vec{k}) &= -\frac{1}{V} \sum_{\vec{q} \neq \vec{k}} \tilde{v}_{ee}(\vec{q} - \vec{k}) \Theta(k_F - q) = \\ &= -\frac{4\pi e^2}{(2\pi)^3} 2\pi \int_0^{k_F} dq \cdot q^2 \int_{-1}^1 \frac{d(\cos \theta_q)}{k^2 + q^2 - 2qk \cos \theta_q} \end{aligned}$$

With

$$\int_{-1}^1 \frac{d\mu}{1 + a\mu} = \frac{1}{a} \ln \left| \frac{1+a}{1-a} \right|$$

this gives the exchange energy of an electron with \vec{k}, σ :

$$v_{\text{HF}}(\vec{k}) = -e^2 \frac{k_F}{\pi} \left(1 + \frac{k_F^2 - k^2}{2k_F k} \ln \left| \frac{k + k_F}{k - k_F} \right| \right) \quad (8.7)$$

It can be rewritten as:

$$\begin{aligned} v_{\text{HF}}(\vec{k}) &= \frac{e^2 k_F}{\pi} S\left(\frac{k}{k_F}\right) \\ S(y) &= -\left(1 + \frac{1-y^2}{2y} \ln \left| \frac{1+y}{1-y} \right| \right) \end{aligned}$$

TODO: Abb43 ,Plot $S(y)$

$$S(1) = -1 \quad \frac{dS(y)}{dy} \Big|_{y=1} \rightarrow \infty$$

Consequences

1. We can calculate the ground state energy within the Hartree-Fock approximation. The exchange contribution is:

$$E_{\text{G,HF}} = \frac{V}{(2\pi)^3} \sum_{\sigma} \int d\vec{k} v_{\text{HF}}(k) n_{\vec{k},\sigma} = N \cdot \left(-\frac{3}{4} \frac{e^2 k_F}{\pi} \right)$$

This is precisely the same correction as from the first order perturbation theory.

$$\varepsilon_{\text{G,HF}} = \frac{E_{\text{G,HF}}}{N} = -\frac{0,916}{r_s} E_{\text{Ry}}$$

2. The fact that $\frac{dS(y)}{dy}$ diverges for $y = 1$ (i.e. at the Fermi level), implies that the density of states vanishes at ε_F , which is unphysical. In fact holds:

$$\mathcal{D}(\varepsilon) = \sum_{\vec{k},\sigma} \delta\left(\varepsilon - \underbrace{(\varepsilon_{\vec{k}} + v_{\text{HF}}(\vec{k}))}_{=\varepsilon_{\text{tot}}(\vec{k})}\right) = \frac{V}{\pi^2} \int_0^\infty dk \cdot k^2 \delta(\varepsilon - \varepsilon_{\text{tot}}(k)) =$$

$$= \frac{V}{\pi^2} k^2 (\varepsilon) \left(\frac{d\varepsilon_{\text{tot}}(k)}{dk} \right)^{-1}$$

We had already shown for free electron gas:

$$C_V \sim \mathcal{D}_0(\varepsilon_F)$$

This can be measured not to vanish.

The reason for this problem is the divergence of the $\frac{1}{r}$ -Coulomb repulsion of the electrons.

Conclusion

We need theories that go beyond first order perturbation theory or Hartree-Fock. Such theories should at least account for screening effects in the interacting gas. With screening, the Coulomb interaction gets renormalized and no divergence at $q = 0$ anymore.

8.5. Dielectric properties and screening

Let us consider the effect of an external electromagnetic potential φ_{ext} on the interacting electron gas. One consequence is that a charge density is induced in the system:

$$\varrho_{e,\text{ind}}(\vec{r}, t) = e\varrho(\vec{r}, t) = \int d^3r' \int_{t_0}^{\infty} dt' \chi^R(\vec{r}, t; \vec{r}', t') \varphi_{\text{ext}}(\vec{r}', t') \quad (8.8)$$

$\chi^R(\vec{r}, t; \vec{r}', t')$ is the polarizability function also called response function. t_0 is the time, at which the perturbation started.

Once the induced charge is know, the induced potential follows:

$$\begin{aligned} \varphi_{\text{ind}}(\vec{r}, t) &= \int d^3r' u_{ee}(\vec{r} - \vec{r}') \varrho_{e,\text{ind}}(\vec{r}', t) \\ u_{ee}(\vec{r}, t) &= \frac{v_{ee}(\vec{r})}{e^2} \stackrel{\text{SI units}}{=} \frac{1}{4\pi\varepsilon_0 \|\vec{r}\|} \end{aligned}$$

Hence the total potential is:

$$\begin{aligned} \varphi_{\text{tot}}(\vec{r}, t) &= \varphi_{\text{ext}}(\vec{r}, t) + \varphi_{\text{ind}}(\vec{r}, t) = \\ &= \varphi_{\text{ext}}(\vec{r}, t) + \int d^3r' \int d^3r'' \int dt' u_{ee}(\vec{r} - \vec{r}') \chi^R(\vec{r}', t; \vec{r}'', t') \varphi_{\text{ext}}(\vec{r}'', t') = \\ &= \int d^3r'' \int dt' \varepsilon^{-1}(\vec{r}, t; \vec{r}'', t') \varphi_{\text{ext}}(\vec{r}'', t') \end{aligned}$$

Here ε is the dielectric function and it follows:

$$\begin{aligned} \varepsilon^{-1}(\vec{r}, t; \vec{r}', t') &= \delta(\vec{r} - \vec{r}') \delta(t - t') + \int d^3r'' u_{ee}(\vec{r} - \vec{r}'') \chi^R(\vec{r}'', t; \vec{r}', t') \\ \Rightarrow \quad \widetilde{\varepsilon^{-1}}(\vec{q}, \omega) &= 1 + \tilde{u}_{ee}(q) \tilde{\chi}^R(\vec{q}, \omega) \end{aligned} \quad (8.9)$$

$$\tilde{u}_{ee}(q) = \frac{1}{\varepsilon_0 q^2}$$

This holds for system which are homogeneous in space and time.

ε^{-1} is experimentally and $\tilde{\chi}^R$ is theoretically accessible.

In total we get:

$$\tilde{\varphi}_{\text{tot}} = \widetilde{\varepsilon^{-1}} \tilde{\varphi}_{\text{ext}}$$

8.5.1. Linear response theory

The linear response theory is based on the idea that the response of a system to a weak perturbation, proportional to the perturbation itself, can be calculated to the *lowest* order in the perturbation.

Example

The conductivity tensor σ is independent of \vec{E} .

$$\vec{j} = \tilde{\sigma} \vec{E}$$

So we have:

$$\tilde{\varrho}_{e,\text{ind}} = \tilde{\chi}^R \tilde{\varphi}_{\text{ext}}$$

The susceptibility $\tilde{\chi}^R$ is independent of $\tilde{\varphi}_{\text{ext}}$. This allows to evaluate non-equilibrium quantities $(\vec{j}, \tilde{\varrho}_{e,\text{ind}})$ by looking at equilibrium functions $(\tilde{\sigma}, \tilde{\chi}^R, \dots)$.

Formally

Given an external perturbation \hat{H}_{ext} , what is the expectation value $\langle \hat{A} \rangle$ of a given operator \hat{A} to *linear order* in \hat{H}_{ext} ?

Suppose that at time t_0 the external perturbation is applied to the system, driving it out of equilibrium.

$$\hat{H}(t) = \hat{H}_0 + \hat{H}_{\text{ext}}(t) \Theta(t - t_0)$$

TODO: Abb44

For $t < t_0$ we have thermal equilibrium and therefore:

$$\langle \hat{A} \rangle_0 = \text{tr}(\hat{A} \hat{\varrho}_0) = \text{tr} \left(\frac{\hat{A}}{Z} e^{-\beta \hat{H}_0} \right)$$

For $t > t_0$ we have:

$$\langle \hat{A}(t) \rangle = \text{tr}(\hat{A} \hat{\varrho}(t)) = ?$$

We consider \hat{H}_{ext} as a weak perturbation, so that it is convenient to work in deviations from equilibrium. We solve the Liouville-von Neumann equation for $\hat{\varrho}(t) = \hat{\varrho}_0 + \Delta \hat{\varrho}(t)$.

$$\dot{\hat{\varrho}}(t) = -\frac{i}{\hbar} [\hat{H}(t), \hat{\varrho}(t)] = -\frac{i}{\hbar} [\hat{H}_0, \Delta \hat{\varrho}(t)] + \frac{i}{\hbar} [\hat{\varrho}_0, \hat{H}_{\text{ext}}(t)] + o_0(\hat{H})$$

This gives:

$$i \Delta \dot{\hat{\varrho}}(t) - \frac{1}{\hbar} [\hat{H}_0, \Delta \hat{\varrho}(t)] \approx -\frac{1}{\hbar} [\hat{\varrho}_0, \hat{H}_{\text{ext}}(t)]$$

The differential equation is solved by expressing the left hand side as:

$$e^{-\frac{i}{\hbar} \hat{H}_0 t} \left(i \frac{d}{dt} \left(e^{\frac{i}{\hbar} \hat{H}_0 t} \Delta \hat{\varrho}(t) e^{-\frac{i}{\hbar} \hat{H}_0 t} \right) \right) e^{\frac{i}{\hbar} \hat{H}_0 t} = -\frac{1}{\hbar} [\hat{\varrho}_0, \hat{H}_{\text{ext}}(t)]$$

$$\begin{aligned} \mathbf{i} \frac{d}{dt} \left(e^{\frac{\mathbf{i}}{\hbar} \hat{H}_0 t} \Delta \hat{\rho}(t) e^{-\frac{\mathbf{i}}{\hbar} \hat{H}_0 t} \right) &= -e^{\frac{\mathbf{i}}{\hbar} \hat{H}_0 t} \frac{1}{\hbar} \left[\hat{\rho}_0, \hat{H}_{\text{ext}}(t) \right] e^{-\frac{\mathbf{i}}{\hbar} \hat{H}_0 t} = -\frac{1}{\hbar} \left[\hat{\rho}_0, \hat{H}_{\text{ext},I}(t) \right] \\ \Delta \hat{\rho}_I(t) - \Delta \hat{\rho}_I(t_0) &= -\frac{\mathbf{i}}{\hbar} \int_{t_0}^t dt' \left[\hat{H}_{\text{ext},I}(t'), \hat{\rho}_0 \right] \end{aligned}$$

Using the cyclic invariance of the trace we get the *Kubo formula*:

$$\delta \langle \hat{A}(t) \rangle = \langle \hat{A}(t) \rangle - \langle \hat{A} \rangle_0 = \int_{t_0}^{\infty} dt' C_{\hat{A}, \hat{H}_{\text{ext}}}^R(t, t') \quad (8.10)$$

$$C_{\hat{A}, \hat{H}_{\text{ext}}}^R(t, t') = -\frac{\mathbf{i}}{\hbar} \Theta(t - t') \langle [\hat{A}_I(t), \hat{H}_{\text{ext},I}(t')] \rangle_0 \quad (8.11)$$

The correlation function $C_{\hat{A}, \hat{H}_{\text{ext}}}^R(t, t')$ is called *response function*.

$\langle \cdot \rangle_0$ means that we take the zeroth order expectation value. The R means, that it is a *retarded* correlation function, which ensures causality. The $_I$ stands for the interaction representation.

- The inherent non-equilibrium quantity $\delta \langle \hat{A}(t) \rangle$ is expressed as a *retarded* correlation function of the system in equilibrium.
- $\Theta(t - t')$ expresses the causality of the solution, which is why we call it retarded correlation function.

Kubo formula in the frequency domain

Consider the case in which the external Hamiltonian is

$$\hat{H}_{\text{ext}}(t) = \hat{B} \cdot f(t)$$

with a constant operator \hat{B} and a time-dependent element $f(t) \in \mathbb{C}$. Then follows:

$$\delta \langle \hat{A}(t) \rangle = \langle \hat{A}(t) \rangle - \langle \hat{A} \rangle_0 = -\frac{\mathbf{i}}{\hbar} \int_{t_0}^{\infty} dt' \langle [\hat{A}_I(t), \hat{B}_I(t')] \rangle_0 f(t') \Theta(t - t')$$

With

$$C_{\hat{A}, \hat{B}}^R(t, t') = -\frac{\mathbf{i}}{\hbar} \Theta(t - t') \langle [\hat{A}_I(t), \hat{B}_I(t')] \rangle_0 \stackrel{\text{cyclic invariance of the trace}}{=} C_{\hat{A}, \hat{B}}^R(t - t')$$

then follows for $t_0 = -\infty$:

$$\delta \langle \tilde{\hat{A}}(\omega) \rangle = \int_{-\infty}^{\infty} dt e^{\mathbf{i}\omega t} \delta \langle \hat{A}(t) \rangle = \tilde{C}_{\hat{A}, \hat{B}}^R(\omega) \tilde{f}(\omega)$$

We note that the usual definition of the Fourier transform is:

$$\tilde{C}_{\hat{A}, \hat{B}}^R(\omega) = \int_{-\infty}^{\infty} dt e^{\mathbf{i}\omega t} C_{\hat{A}, \hat{B}}^R(t)$$

In order for the Fourier transform to be well defined, the integrand must decay for both plus and minus infinity. For retarded functions as

$$C_{\hat{A}, \hat{B}}^R(t - t') = -\frac{\mathbf{i}}{\hbar} \Theta(t - t') \langle [\hat{A}_I(t), \hat{B}_I(t')] \rangle_0$$

that are zero at negative times, only $+\infty$ can cause a problem. It is then usual to define the Fourier transform as:

$$\tilde{C}_{\hat{A},\hat{B}}^R(\omega) = \lim_{\eta \searrow 0} \int_{-\infty}^{\infty} dt e^{i\omega t} e^{-\eta t} C_{\hat{A},\hat{B}}^R(t)$$

Consider a position dependent external perturbation given by:

$$\hat{H}_{\text{ext}}(t) = \int d^3r \hat{B}(\vec{r}) f(\vec{r}, t)$$

In this case one finds:

$$\delta \langle \tilde{\hat{A}}(\omega) \rangle = \int d^3r \tilde{C}_{\hat{A},\hat{B}}^R \tilde{f}(\vec{r}, \omega)$$

Example (Kubo formula for the dielectric function)

The external perturbation is:

$$\hat{H}_{\text{ext}} = \int d^3r \hat{\varrho}_e(\vec{r}) \varphi_{\text{ext}}(\vec{r}, t)$$

The induced charge is:

$$\langle \hat{\varrho}_{e,\text{ind}} \rangle = \langle \hat{\varrho}_e \rangle - \langle \hat{\varrho}_e \rangle_0$$

This gives:

$$\begin{aligned} \langle \hat{\varrho}_{e,\text{ind}} \rangle &= \varrho_{e,\text{ind}} = \int d^3r' \int_{t_0}^{\infty} dt' C_{\varrho_e(\vec{r}), \varrho_e(\vec{r}')}^R(t, t') \varphi_{\text{ext}}(\vec{r}', t') \\ C_{\varrho_e(\vec{r}), \varrho_e(\vec{r}')}^R(t, t') &= \chi^R(\vec{r}, t, \vec{r}', t') = -\frac{i}{\hbar} \Theta(t - t') \langle [\hat{\varrho}_{e,I}(\vec{r}, t), \hat{\varrho}_{e,I}(\vec{r}', t')] \rangle_0 \end{aligned}$$

χ^R is the *density-density correlation function*.

Note: For the electron gas holds:

$$\begin{aligned} \chi^R(\vec{r}, t, \vec{r}', t') &= \chi^R(\vec{r} - \vec{r}', t - t') \\ \Rightarrow \tilde{\chi}^R(\vec{q}, \omega) &= -\frac{i}{\hbar} \Theta(t - t') \frac{1}{V} \left\langle \left[\tilde{\varrho}_{e,I}(\vec{q}, t), \tilde{\varrho}_{e,I}(-\vec{q}, t') \right] \right\rangle_0 \end{aligned} \quad (8.12)$$

(cf. BRUUS and FLENSBERG, chapter 8.5)

To proceed, we need to know how to evaluate the time evolution of operators in second quantization!

8.5.2. Time dependence of operators in second quantization

In second quantization one deals with operators. Hence also the time evolution is as given in the Heisenberg representation, where the Heisenberg equations hold. Specifically, for a time independent Hamiltonian \hat{H} , a given operator $\hat{A}(t)$ has to obey:

$$\dot{\hat{A}}(t) = \frac{i}{\hbar} [\hat{H}, \hat{A}(t)] + (\partial_t \hat{A})(t) \quad (8.13)$$

Here $\dot{\hat{A}}$ is the time derivative in the Heisenberg picture and $\partial_t \hat{A}$ is the time derivative in the Schrödinger picture.

Example

$$\hat{H} = \sum_{\nu} \varepsilon_{\nu} \hat{a}_{\nu}^{\dagger} \hat{a}_{\nu}$$

The equation of motion is:

$$\begin{aligned} \hat{a}_{\nu}(t) &:= e^{\mathbf{i}\hat{H}\frac{t}{\hbar}} \hat{a}_{\nu} e^{-\hat{H}\frac{t}{\hbar}} \\ \hat{a}_{\nu}^{\dagger}(t) &:= e^{\mathbf{i}\hat{H}\frac{t}{\hbar}} \hat{a}_{\nu}^{\dagger} e^{-\hat{H}\frac{t}{\hbar}} \end{aligned}$$

We simply have to use these equations and the form of the Hamiltonian to get:

$$\begin{aligned} \dot{\hat{a}}_{\nu}(t) &= \frac{\mathbf{i}}{\hbar} [\hat{H}, \hat{a}_{\nu}(t)] = \frac{\mathbf{i}}{\hbar} e^{\mathbf{i}\hat{H}\frac{t}{\hbar}} [\hat{H}, \hat{a}_{\nu}] e^{-\mathbf{i}\hat{H}\frac{t}{\hbar}} = \\ &= \frac{\mathbf{i}}{\hbar} e^{\mathbf{i}\hat{H}\frac{t}{\hbar}} \sum_{\nu'} \varepsilon_{\nu'} \underbrace{[\hat{a}_{\nu'}^{\dagger} \hat{a}_{\nu'}, \hat{a}_{\nu}]}_{=-\delta_{\nu\nu'}} e^{-\mathbf{i}\hat{H}\frac{t}{\hbar}} = -\frac{\mathbf{i}}{\hbar} \varepsilon_{\nu} \hat{a}_{\nu}(t) \end{aligned}$$

This yields:

$$\hat{a}_{\nu}(t) = e^{-\frac{\mathbf{i}}{\hbar} \varepsilon_{\nu} t} \hat{a}_{\nu} \quad (8.14)$$

$$[\hat{a}_{\nu_1}(t_1), \hat{a}_{\nu_2}^{\dagger}(t_2)] = e^{-\frac{\mathbf{i}}{\hbar} \varepsilon_{\nu_1}(t_1-t_2)} \delta_{\nu_1\nu_2} \quad (8.15)$$

8.5.3. Susceptibility of non-interacting electron gas

To start with, we evaluate the susceptibility $\tilde{\chi}^R$ for the non-interacting electron gas. We denote it by $\tilde{\chi}_0^R$.

$$\begin{aligned} \tilde{\chi}^R(\vec{q}, t-t') &= -\frac{\mathbf{i}}{\hbar} \Theta(t-t') \left\langle \left[\tilde{\hat{\rho}}_{e,I}(\vec{q}, t), \tilde{\hat{\rho}}_{e,I}(-\vec{q}, t') \right] \right\rangle_0 \\ \hat{H}_0 &= \sum_{\vec{k}, \sigma} \varepsilon_{\vec{k}} \hat{c}_{\vec{k}, \sigma}^{\dagger} \hat{c}_{\vec{k}, \sigma} \\ \Rightarrow \quad \tilde{\hat{\rho}}_e(\vec{q}, t) &= -e \sum_{\vec{k}, \sigma} \hat{c}_{\vec{k}, \sigma}^{\dagger} \hat{c}_{\vec{k}+\vec{q}, \sigma} e^{-\frac{\mathbf{i}}{\hbar} (\varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}}) t} \end{aligned}$$

Hence follows:

$$\begin{aligned} \tilde{\chi}_0^R(\vec{q}, t-t') &= -\frac{\mathbf{i}}{\hbar} \Theta(t-t') \frac{e^2}{V} \sum_{\vec{k}, \vec{k}', \sigma, \sigma'} \left\langle \left[\hat{c}_{\vec{k}, \sigma}^{\dagger} \hat{c}_{\vec{k}+\vec{q}, \sigma}, \hat{c}_{\vec{k}', \sigma'}^{\dagger} \hat{c}_{\vec{k}'-\vec{q}, \sigma'} \right] \right\rangle_0 \cdot \\ &\quad \cdot e^{\frac{\mathbf{i}}{\hbar} ((\varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}}) t + (\varepsilon_{\vec{k}'} - \varepsilon_{\vec{k}'-\vec{q}}) t')} \end{aligned}$$

With

$$[\hat{c}_{\nu}^{\dagger} \hat{c}_{\mu}, \hat{c}_{\nu'}^{\dagger} \hat{c}_{\mu'}] = \hat{c}_{\nu}^{\dagger} \hat{c}_{\mu'} \delta_{\mu\nu'} - \hat{c}_{\nu'}^{\dagger} \hat{c}_{\mu} \delta_{\nu'\mu}$$

$$\langle \hat{c}_{\vec{k}}^\dagger \hat{c}_{\vec{k}} \rangle_0 = \langle \hat{n}_{\vec{k}} \rangle_0 = f(\varepsilon_{\vec{k}})$$

one gets:

$$\tilde{\chi}_0^R(\vec{q}, t - t') = -\frac{\mathbf{i}}{\hbar} \Theta(t - t') \frac{e^2}{V} \sum_{\vec{k}, \sigma} \left(f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}}) \right) e^{\frac{\mathbf{i}}{\hbar}(\varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}})(t-t')} \quad (8.16)$$

The Fourier transformation is:

$$\begin{aligned} \tilde{\chi}_0^R(\vec{q}, \omega) &= \int_{-\infty}^{\infty} d\tau e^{\mathbf{i}\omega\tau} \tilde{\chi}_0^R(\vec{q}, \tau) e^{-\eta\tau} = \\ &= \frac{e^2}{\hbar V} \sum_{\vec{k}, \sigma} \frac{f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}})}{\frac{1}{\hbar}(\varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}}) + \omega + \mathbf{i}\eta} \end{aligned} \quad (8.17)$$

$\eta \searrow 0$ is a convergence factor. This function is the *Lindhard function* (“the bubble”). We then have:

$$\tilde{\chi}_0^R(\vec{q}, \omega) = \text{Re}(\tilde{\chi}_0^R) + \mathbf{i} \cdot \text{Im}(\tilde{\chi}_0^R)$$

$\text{Re}(\tilde{\chi}_0^R)(\vec{q}, \omega)$ is the principal part of the integral obtained for $\eta = 0$. Using $\text{Im}\left(\frac{1}{a+\mathbf{i}\eta}\right) \rightarrow -\pi\delta(a)$ one gets:

$$\text{Im}(\tilde{\chi}_0^R)(\vec{q}, \omega) = -\frac{\pi e^2}{V\hbar} \sum_{\vec{k}, \sigma} \left(f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}}) \right) \delta\left(\frac{\varepsilon_{\vec{k}}}{\hbar} - \frac{\varepsilon_{\vec{k}+\vec{q}}}{\hbar} + \omega\right) \quad (8.18)$$

TODO: Abb45; electron-hole excitations of free electron gas

Consider $T = 0$, i.e. $f(\varepsilon_k) = \Theta(k_F - k)$, and:

$$f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}}) \neq 0 \quad \begin{cases} \text{if } k > k_F \text{ and } \left| \vec{k} + \vec{q} \right| < k_F & \Rightarrow \omega < 0 \\ \text{if } k > k_F \text{ and } \left| \vec{k} + \vec{q} \right| > k_F & \Rightarrow \omega > 0 \end{cases}$$

Moreover, conservation of energy implies for $\omega > 0$ (this is enough, since the imaginary part of $\tilde{\chi}_0^R$ is odd).

$$0 < \omega = \frac{q^2}{2m} + \frac{\vec{k} \cdot \vec{q}}{m}$$

The maximal and minimal values are ($q \gg k_F$):

$$\begin{aligned} \omega_{\max} &= \frac{q^2}{2m} + v_F q \\ \omega_{\min} &= \frac{q^2}{2m} - v_F q \end{aligned}$$

TODO: Add46

Note: To lowest order in the electron-electron interaction we get:

$$\widetilde{\varepsilon^{-1}}(\vec{q}, \omega) = 1 + \tilde{u}_{ee}(q) \tilde{\chi}_0$$

This quantity diverges for $q \rightarrow 0$. ($\tilde{\chi}_0(\vec{q} = 0, \omega = 0) = -e^2 \mathcal{D}(\varepsilon_F)$) So at this stage, we don't get the screening.

We need to evaluate $\tilde{\chi}^R$ in a non perturbative way. We use the *equation of motion method* within the “Random phase approximation”.

8.6. Random phase approximation

The idea is to find the time evolution of the operator $\hat{\rho}_{e,I}(\vec{q}, t)$ and hence to obtain in turn $\tilde{\chi}^R(q, t - t')$.

Equation of motion method applied to $\tilde{\chi}^R(\vec{q}, t)$ and $\hat{\rho}_I(t)$

(see BRUUS and FLENSBERG, chapter 9.4)

We start from the correlation function:

$$\tilde{\chi}^R(\vec{q}, \vec{k}, \sigma, t - t') = -\frac{i}{\hbar} \Theta(t - t') \left\langle \left[\left(\hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \right) (t), \hat{\rho}_I(-\vec{q}, t') \right] \right\rangle_0$$

We consider $i\hbar \partial_t \tilde{\chi}^R$ and use the Schrödinger equation for the time evolution:

$$\begin{aligned} i\hbar \partial_t \tilde{\chi}^R(\vec{q}, \vec{k}, \sigma, t - t') &= \delta(t - t') \left\langle \left[\left(\hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \right) (t), \hat{\rho}_I(-\vec{q}, t') \right] \right\rangle_0 + \\ &+ \left(\frac{-i}{\hbar} \Theta(t - t') \right) \left(\frac{-\hbar}{i} \right) \left\langle \frac{i}{\hbar} [\hat{H}, \hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma}], \hat{\rho}_I(-\vec{q}, t') \right\rangle_0 \end{aligned}$$

Remember:

$$\hat{H} = \hat{H}_0 + \hat{V}_{ee}^{\text{jellium}}$$

For the first commutator we obtain:

$$\begin{aligned} \left[\hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma}, \hat{\rho}_I(-\vec{q}) \right] &= \sum_{\vec{k}', \sigma'} \left[\hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma}, \hat{c}_{\vec{k}', \sigma'}^\dagger \hat{c}_{\vec{k}' - \vec{q}, \sigma'} \right] = \\ &= \hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k}, \sigma} - \hat{c}_{\vec{k} + \vec{q}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} = \left(\hat{n}_{\vec{k}, \sigma} - \hat{n}_{\vec{k} + \vec{q}, \sigma} \right) \\ \left[\hat{H}_0, \hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \right] &= \left(\varepsilon_{\vec{k}} - \varepsilon_{\vec{k} + \vec{q}} \right) \hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \end{aligned}$$

$$\begin{aligned} \left[\hat{V}_{ee}^{\text{jellium}}, \hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \right] &= -\frac{1}{2V} \sum_{\vec{k}', \vec{q}' \neq 0, \sigma'} \tilde{v}_{ee}(\vec{q}) \left(\hat{c}_{\vec{k} + \vec{q}', \sigma}^\dagger \hat{c}_{\vec{k}' - \vec{q}', \sigma'}^\dagger \hat{c}_{\vec{k}', \sigma'} \hat{c}_{\vec{k} + \vec{q}, \sigma} + \right. \\ &+ \hat{c}_{\vec{k}' + \vec{q}', \sigma'}^\dagger \hat{c}_{\vec{k} - \vec{q}', \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \hat{c}_{\vec{k}', \sigma'} - \hat{c}_{\vec{k}' + \vec{q}', \sigma'}^\dagger \hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q} - \vec{q}', \sigma} \hat{c}_{\vec{k}', \sigma'} - \\ &\left. - \hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k}' - \vec{q}', \sigma'}^\dagger \hat{c}_{\vec{k}', \sigma'} \hat{c}_{\vec{k} + \vec{q} - \vec{q}', \sigma} \right) \end{aligned}$$

This set of equations in general cannot be closed, because the nested commutators generate operators, which contain each time a larger number of \hat{c} and \hat{c}^\dagger .

So truncation schemes are performed in the same spirit of the mean field approximation. The idea is to keep only Hartree contributions, as they are the most diverging ones. Hence we get:

$$\begin{aligned} \left[\hat{V}_{ee}^{\text{jellium}}, \hat{c}_{\vec{k}, \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \right] &\approx -\frac{1}{2V} \sum_{\vec{k}', \vec{q}' \neq 0, \sigma'} \tilde{v}_{ee}(\vec{q}') \left(\hat{c}_{\vec{k} + \vec{q}', \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \underbrace{\left\langle \hat{c}_{\vec{k}' - \vec{q}', \sigma'}^\dagger \hat{c}_{\vec{k}', \sigma'} \right\rangle_0}_{=0} + \right. \\ &+ \underbrace{\left\langle \hat{c}_{\vec{k} + \vec{q}', \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \right\rangle_0}_{=\delta_{\vec{q}, \vec{q}'}} \hat{c}_{\vec{k}' - \vec{q}', \sigma'}^\dagger \hat{c}_{\vec{k}', \sigma'} + \hat{c}_{\vec{k}' + \vec{q}', \sigma'}^\dagger \hat{c}_{\vec{k}', \sigma'} \underbrace{\left\langle \hat{c}_{\vec{k} - \vec{q}', \sigma}^\dagger \hat{c}_{\vec{k} + \vec{q}, \sigma} \right\rangle_0}_{=\delta_{\vec{q}, -\vec{q}'}} + \end{aligned}$$

$$\begin{aligned}
& + \underbrace{\left\langle \hat{c}_{\vec{k}'+\vec{q}',\sigma'}^\dagger \hat{c}_{\vec{k}',\sigma'} \right\rangle_0}_{=0} \hat{c}_{\vec{k}-\vec{q}',\sigma}^\dagger \hat{c}_{\vec{k}+\vec{q},\sigma} - \hat{c}_{\vec{k}'+\vec{q}',\sigma'}^\dagger \hat{c}_{\vec{k}',\sigma'} \underbrace{\left\langle \hat{c}_{\vec{k},\sigma}^\dagger \hat{c}_{\vec{k}+\vec{q}-\vec{q}',\sigma} \right\rangle_0}_{=\delta_{\vec{q},\vec{q}'}} - \\
& - \underbrace{\left\langle \hat{c}_{\vec{k}'+\vec{q}',\sigma'}^\dagger \hat{c}_{\vec{k}',\sigma'} \right\rangle_0}_{=0} \hat{c}_{\vec{k},\sigma}^\dagger \hat{c}_{\vec{k}+\vec{q}-\vec{q}',\sigma} - \hat{c}_{\vec{k},\sigma}^\dagger \hat{c}_{\vec{k}+\vec{q}-\vec{q}',\sigma} \underbrace{\left\langle \hat{c}_{\vec{k}',-\vec{q}',\sigma'}^\dagger \hat{c}_{\vec{k}',\sigma'} \right\rangle_0}_{=0} \\
& - \underbrace{\left\langle \hat{c}_{\vec{k},\sigma}^\dagger \hat{c}_{\vec{k}+\vec{q}-\vec{q}',\sigma} \right\rangle_0}_{=\delta_{\vec{q},\vec{q}'}} \hat{c}_{\vec{k}',-\vec{q}',\sigma'}^\dagger \hat{c}_{\vec{k}',\sigma'} + \text{const.} \Big) = \\
& = -\frac{1}{V} \tilde{v}_{ee}(\vec{q}') \underbrace{\left(\sum_{\vec{k}',\sigma'} \hat{c}_{\vec{k}',-\vec{q}',\sigma'}^\dagger \hat{c}_{\vec{k}',\sigma'} \right)}_{=\hat{\rho}(\vec{q}')} \left(\langle \hat{n}_{\vec{k}+\vec{q},\sigma} \rangle - \langle \hat{n}_{\vec{k},\sigma} \rangle \right)
\end{aligned}$$

This gives:

$$\begin{aligned}
i\hbar \partial_t \tilde{\chi}^R(\vec{q}, \vec{k}, \sigma, t-t') &= \delta(t-t') \left(\langle \hat{n}_{\vec{k},\sigma} \rangle - \langle \hat{n}_{\vec{k}+\vec{q},\sigma} \rangle \right) - \\
& - \left(-\frac{i}{\hbar} \Theta(t-t') \right) \left(\varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}} \right) \left\langle \hat{c}_{\vec{k},\sigma}^\dagger \hat{c}_{\vec{k}+\vec{q}}(t), \hat{\rho}_I(-\vec{q}, t') \right\rangle_0 - \\
& - \left(-\frac{i}{\hbar} \Theta(t-t') \right) \left(\langle \hat{n}_{\vec{k},\sigma} \rangle - \langle \hat{n}_{\vec{k}+\vec{q},\sigma} \rangle \right) \tilde{v}_{ee}(q) \frac{1}{V} \langle [\hat{\rho}_I(\vec{q}, t), \hat{\rho}_I(-\vec{q}, t)] \rangle
\end{aligned}$$

Hence follows:

$$\begin{aligned}
i\hbar \partial_t \tilde{\chi}^R(\vec{q}, \vec{k}, \sigma, t-t') &= \delta(t-t') \left(f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}}) \right) - \\
& - \left(\varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}} \right) \tilde{\chi}^R(\vec{q}, \vec{k}, \sigma, t-t') - \\
& - \left(f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}}) \right) \frac{\tilde{v}_{ee}(q)}{V} \sum_{\vec{k}',\sigma'} \tilde{\chi}(\vec{q}, \vec{k}', \sigma', t-t')
\end{aligned}$$

Let us take the Fourier transformation:

$$\begin{aligned}
\partial_t &\rightarrow -i(\omega + i\eta) \\
\text{FT}(\delta(t)) &= 1
\end{aligned}$$

$$\begin{aligned}
\left(\hbar\omega + i\eta + \varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}} \right) \tilde{\chi}^R(\vec{q}, \vec{k}, \sigma, \omega) &= \frac{f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}})}{1 - \frac{\tilde{v}_{ee}(q)}{V} \sum_{\vec{k}',\sigma'} \tilde{\chi}(\vec{q}, \vec{k}', \sigma', \omega)} \\
\Rightarrow \tilde{\chi}^R(\vec{q}, \vec{k}, \sigma, \omega) &= \frac{f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}})}{\hbar\omega + i\eta + \varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}}} \left(1 - (+) \frac{\tilde{v}_{ee}(q)}{V} \sum_{\vec{k}',\sigma'} \tilde{\chi}(\vec{q}, \vec{k}', \sigma', \omega) \right)
\end{aligned}$$

Remember:

$$\tilde{\chi}^R(\vec{q}, \omega) = \frac{e^2}{V} \sum_{\vec{k},\sigma} \tilde{\chi}^R(\vec{q}, \vec{k}, \sigma, \omega)$$

$$\Rightarrow \quad \tilde{\chi}^R(\vec{q}, \omega) = \tilde{\chi}_0 (1 - (+) \tilde{u}_{ee}(q) \tilde{\chi}^R(\vec{q}, \omega))$$

$$\Rightarrow \quad \tilde{\chi}^R(\vec{q}, \omega) = \frac{\tilde{\chi}_0}{1 + (-) \tilde{u}_{ee}(q) \tilde{\chi}_0} =: \tilde{\chi}_{\text{RPA}}(\vec{q}, \omega)$$

Note: $\tilde{\chi}_{\text{RPA}}$ is obtained from a Dyson series:

$$\begin{aligned} \tilde{\chi}_{\text{RPA}} &= \tilde{\chi}_0 + \tilde{\chi}_0 \tilde{u}_{ee} \tilde{\chi}_{\text{RPA}} = \\ &= \tilde{\chi}_0 + \tilde{\chi}_0 \tilde{u}_{ee} \tilde{\chi}_0 + \tilde{\chi}_0 \tilde{u}_{ee} \tilde{\chi}_0 \tilde{u}_{ee} \tilde{\chi}_{\text{RPA}} = \dots \end{aligned}$$

This tells us that also a diagrammatic solution would be possible.

Note:

$$\widetilde{\varepsilon^{-1}}(\vec{q}, \omega) = 1 + \tilde{u}_{ee}(q) \tilde{\chi}_{\text{RPA}}(\vec{q}, \omega) = 1 + \frac{\tilde{u}_{ee} \tilde{\chi}_0}{1 - \tilde{u}_{ee}(q) \tilde{\chi}_0} = \frac{1}{1 - \tilde{u}_{ee}(q) \tilde{\chi}_0(\vec{q}, \omega)}$$

We hence have screening. (cf. FLENSBERG)

$$\lim_{\vec{q} \rightarrow 0} \tilde{\chi}_0(\vec{q}, \omega) = -e^2 \mathcal{D}(\varepsilon_F)$$

$$\Rightarrow \quad \widetilde{\varepsilon^{-1}} = \frac{1}{1 + \frac{e^2 \mathcal{D}(\varepsilon_F)}{\varepsilon_0 q^2}} = \frac{q^2}{q^2 + k_{\text{TF}}^2}$$

$\frac{2\pi}{k_{\text{TF}}}$ is the Thomas-Fermi wavelength.

$$k_{\text{TF}}^2 = \frac{e^2 \mathcal{D}(\varepsilon_F)}{\varepsilon_0}$$

So we get:

$$\tilde{u}_{ee}^{\text{RPA}} = \widetilde{\varepsilon^{-1}}_{\text{RPA}} \tilde{u}_{ee} = \frac{1}{\varepsilon_0} \frac{1}{q^2 + k_{\text{TF}}^2}$$

In real space this gives:

$$u_{ee}^{\text{RPA}}(\vec{r}) = \frac{e^{-rk_{\text{TF}}}}{r}$$

So the Coulomb interaction is screened on a length k_{TF}^{-1} .

Besides electron-hole interactions, there are also plasmons.

TODO: Abb47

9 Phonons

(see BRUUS and FLENSBERG, chapter 3)

So far we have treated the ions as *fixed* at the positions $\{\vec{R}^0\}$ of the underlying lattice. In this chapter, we study the basic properties of *lattice vibrations*. Under the assumption that the ions perform *small oscillations* about their equilibrium positions, the set of *coupled equations* for the interacting ions can be diagonalized yielding the so-called *normal modes of vibrations*. We shall see that we can represent each normal mode as a harmonic oscillator with a given dispersion relation $\omega_\lambda(\vec{k})$. The quantized vibrations are denoted *phonons*.

Note: Phonons play a fundamental role in our understanding of sound propagation, specific heat, elasticity and electrical resistivity of solids. Moreover, the electron-phonon coupling is at the origin of conventional superconductivity.

9.1. Acoustic and optical phonons

Let us consider the ionic part of the Hamiltonian of a solid:

$$\hat{H}_{\text{ion}} = \hat{T}_{\text{ion}} + \hat{V}_{ii} + \hat{V}_{ei} \quad (9.1)$$

Here we have:

$$\hat{T}_{\text{ion}} = \sum_{j=1}^{N_{\text{ion}}} \frac{\hat{p}_j^2}{2m} \quad \hat{V}_{ii} = \frac{1}{2} \sum_{l \neq j} \hat{v}_{ii}(\vec{R}_l - \vec{R}_j) \quad (9.2)$$

We restrict the discussion here to a monoatomic Bravais lattice. Since the atoms can move a little bit around the equilibrium position $\vec{R}_j^{(0)}$, we write

$$\vec{R}_j = \vec{R}_j^{(0)} + \vec{u}_j(t) \quad (9.3)$$

with the displacement $\vec{u}_j(t)$ from the equilibrium position.

The *classical* equations of motion for the displacements read:

$$M\ddot{\vec{U}}_j(t) = -\frac{\partial V_{ii}}{\partial \vec{u}_j} = -\sum_l \frac{\partial}{\partial \vec{u}_j} v_{ii}(\vec{R}_l - \vec{R}_j)$$

In order to simplify this problem, we proceed in steps:

- i) We expand the ion-ion interaction up to second order in \vec{u} :

$$V_{ii} = \frac{1}{2} \sum_{l \neq j} v_{ii}(\vec{R}_l - \vec{R}_j) \approx V_{ii}^{\text{static}} + \frac{1}{2} \sum_{l \neq j} (\vec{u}_l - \vec{u}_j) \cdot \vec{\nabla} v_{ii}(\vec{R}_l - \vec{R}_j) \Big|_{\vec{R}_l^{(0)}, \vec{R}_j^{(0)}} +$$

$$\begin{aligned}
& + \frac{1}{2} \sum_{l \neq j} \sum_{\mu, \nu \in \{x, y, z\}} (\vec{u}_l - \vec{u}_j)_\mu (\vec{u}_l - \vec{u}_j)_\nu \frac{1}{2} \frac{\partial^2}{(\partial \vec{u}_l)_\mu (\partial \vec{u}_j)_\nu} v_{ii} (\vec{R}_l - \vec{R}_j) \Big|_{\vec{R}_l^{(0)}, \vec{R}_j^{(0)}} = \\
\stackrel{\text{minimum}}{=} & V_{ii}^{\text{static}} + \underbrace{\frac{1}{2} \sum_{l \neq j} \sum_{\mu, \nu} (\vec{u}_l)_\mu D_{\mu\nu} (\vec{R}_l^{(0)} - \vec{R}_j^{(0)}) (\vec{u}_j)_\nu}_{= \frac{1}{2} \sum_{l \neq j} \vec{u}_l \mathbf{D} (\vec{R}_l^{(0)} - \vec{R}_j^{(0)})} \quad (9.4)
\end{aligned}$$

The linear term is zero, since the equilibrium positions are the places, where the sum of the forces on an atom, $\vec{F}_j = \sum_{l \neq j} \vec{\nabla} v_{ii} (\vec{R}_l - \vec{R}_j)$, vanishes.

The D -matrix has some important properties:

- a) Due to the interchangeability of the sequence of the derivative holds:

$$\mathbf{D}^T (\vec{R}^{(0)}) = \mathbf{D} (\vec{R}^{(0)})$$

- b)

$$\sum_{\vec{R}^{(0)}} \mathbf{D} (\vec{R}^{(0)}) = 0$$

This property follows from the fact, that $V_{ii} - V_{ii}^{\text{static}} = 0$ if all the displacements are the same, given e.g. by a vector \vec{d} .

$$0 = \sum_{ij} \vec{d} \mathbf{D} (\vec{R}_i^{(0)} - \vec{R}_j^{(0)}) \vec{d}$$

- c) From the reflection properties of the Bravais lattice follows:

$$\mathbf{D} (\vec{R}^{(0)}) = \mathbf{D} (-\vec{R}^{(0)})$$

- ii) We introduce now the *dynamical matrix*:

$$D_{\mu\nu} (\vec{k}) = \sum_{\vec{R}^{(0)}} D_{\mu\nu} (\vec{R}^{(0)}) e^{-i\vec{k} \cdot \vec{R}^{(0)}}$$

Together with the properties a) - c) we find:

$$\begin{aligned}
D_{\mu\nu} (\vec{k}) &= \frac{1}{2} \left(\sum_{\vec{R}^{(0)}} \mathbf{D} (\vec{R}^{(0)}) e^{-i\vec{k} \cdot \vec{R}^{(0)}} + \sum_{\vec{R}^{(0)}} \mathbf{D} (-\vec{R}^{(0)}) e^{+i\vec{k} \cdot \vec{R}^{(0)}} \right) = \\
&\stackrel{\text{b), c)}}{=} \frac{1}{2} \sum_{\vec{R}^{(0)}} \mathbf{D} (\vec{R}^{(0)}) (e^{-i\vec{k} \cdot \vec{R}^{(0)}} + e^{+i\vec{k} \cdot \vec{R}^{(0)}})
\end{aligned}$$

And hence we get:

$$\mathbf{D} (\vec{k}) = 2 \sum_{\vec{R}^{(0)}} \mathbf{D} (\vec{R}^{(0)}) \sin^2 \left(\frac{\vec{k} \cdot \vec{R}^{(0)}}{2} \right) \quad (9.5)$$

In other words, the matrix $\mathbf{D} (\vec{k})$ is real and symmetric and hence diagonalizable in an orthonormal basis.

iii) Going back to the original coupled problem:

$$M\ddot{\vec{u}}_1 = -\frac{\partial V_{ii}}{\partial \vec{u}_1} \quad \Rightarrow \quad -M\ddot{\vec{u}}_1 = \sum_l \mathbf{D} \left(\vec{R}_l^{(0)} - \vec{R}_1^{(0)} \right) \vec{u}_l \quad (9.6)$$

We look for harmonic solutions:

$$\vec{u}_1(t) = \vec{\epsilon} e^{i(\vec{k} \cdot \vec{R}_1^{(0)} - \omega t)}$$

Insertion in (9.6) yields:

$$M\omega^2 \vec{\epsilon} e^{i(\vec{k} \cdot \vec{R}_1^{(0)} - \omega t)} = \sum_l \mathbf{D} \left(\vec{R}_l^{(0)} - \vec{R}_1^{(0)} \right) \vec{\epsilon} e^{i\vec{k} \cdot (\vec{R}_l^{(0)} - \vec{R}_1^{(0)})} e^{i(\vec{k} \cdot \vec{R}_1^{(0)} - \omega t)}$$

Hence we get:

$$\boxed{M\omega^2 \vec{\epsilon} = \mathbf{D}(\vec{k}) \vec{\epsilon}} \quad (9.7)$$

Because $\mathbf{D}(\vec{k})$ is real and symmetric, there exist for any value of \vec{k} an orthonormal basis set $\{\vec{\epsilon}_{k1}, \vec{\epsilon}_{k2}, \vec{\epsilon}_{k3}\}$ of vectors (polarization vectors), which diagonalizes $\mathbf{D}(\vec{k})$. In that basis holds:

$$\mathbf{D}(\vec{k}) \vec{\epsilon}_{\vec{k},\lambda} = K_{\vec{k},\lambda} \vec{\epsilon}_{\vec{k},\lambda}$$

$K_{\vec{k},\lambda}$ are the eigenvalues of \mathbf{D} and $\vec{\epsilon}_{\vec{k},\lambda} \cdot \vec{\epsilon}_{\vec{k},\lambda'} = \delta_{\lambda\lambda'}$. This in turn yields:

$$M\omega^2 \vec{\epsilon}_{\vec{k},\lambda} = K_{\vec{k},\lambda} \vec{\epsilon}_{\vec{k},\lambda}$$

And so we get the normal mode frequencies:

$$\boxed{\omega_\lambda(\vec{k}) = \sqrt{\frac{K_{\vec{k},\lambda}}{M}}} \quad (9.8)$$

iv) We can also notice that the normal modes

$$\vec{U}_{\vec{k},\lambda}(\vec{R}^{(0)}, t) = \vec{\epsilon}_{\vec{k},\lambda} e^{i(\vec{k} \cdot \vec{R}^{(0)} - \omega_\lambda(\vec{k})t)}$$

satisfy the same equations as the harmonic oscillator. So they can also be quantized.

How to quantize?

We introduce (bosonic) creation and annihilation operators, such that holds:

$$\begin{aligned} \hat{U}_{\vec{k},l} &= l_{\vec{k},\lambda} \frac{1}{\sqrt{2}} \left(\hat{b}_{-\vec{k},\lambda}^\dagger + \hat{b}_{\vec{k},\lambda} \right) \vec{\epsilon}_{\vec{k},\lambda} \\ \hat{P}_{\vec{k},l} &= \frac{\hbar}{l_{\vec{k},\lambda}} \cdot \frac{i}{\sqrt{2}} \left(\hat{b}_{-\vec{k},\lambda} - \hat{b}_{\vec{k},\lambda} \right) \vec{\epsilon}_{\vec{k},\lambda} \end{aligned} \quad (9.9)$$

Here we introduced the oscillator length:

$$l_{\vec{k},\lambda} = \sqrt{\frac{\hbar}{M\omega_\lambda(\vec{k})}}$$

Moreover holds:

$$\left[\hat{b}_{\vec{k}_1, \lambda_1}, \hat{b}_{\vec{k}_2, \lambda_2} \right] = \delta_{\vec{k}_1, \vec{k}_2} \delta_{\lambda_1 \lambda_2}$$

To find the Hamiltonian describing the lattice dynamics, we go back to \hat{H}_{ion} and look at:

$$\hat{H}_{\text{ph}} = \hat{T}_{\text{ion}} + \frac{1}{2} \sum_{i,j} \sum_{\mu,\nu} \left(\hat{\vec{U}}_i \right)_\mu D_{\mu\nu} \left(\hat{\vec{U}}_j \right)_\nu \quad (9.10)$$

Introducing no

$$\begin{aligned} \hat{\vec{P}}_j &= \frac{1}{\sqrt{N_{\text{ion}}}} \sum_{\vec{k} \in 1.\text{BZ}, \lambda} \hat{P}_{\vec{k}, \lambda} e^{i\vec{k} \cdot \vec{R}_j^{(0)}} \\ \hat{\vec{U}}_j &= \frac{1}{\sqrt{N_{\text{ion}}}} \sum_{\vec{k} \in 1.\text{BZ}, \lambda} \hat{U}_{\vec{k}, \lambda} e^{i\vec{k} \cdot \vec{R}_j^{(0)}} \end{aligned}$$

we find:

$$\begin{aligned} \hat{H}_{\text{ph}} &= \frac{1}{2M} \sum_{\vec{k}} \sum_{\lambda, \lambda'} \hat{\vec{P}}_{\vec{k}, \lambda} \cdot \hat{\vec{P}}_{-\vec{k}, \lambda'} + \frac{1}{2} \sum_{\vec{k}, \lambda, \lambda'} \sum_{\mu, \nu} \left(\hat{\vec{U}}_{\vec{k}, \lambda} \right)_\mu \cdot \underbrace{D_{\mu\nu}}_{=M\omega_{\lambda'}^2(-\vec{k})} \cdot \left(\hat{\vec{U}}_{-\vec{k}, \lambda'} \right)_\nu = \\ &\stackrel{(9.9)}{=} \sum_{\vec{k}} \sum_{\lambda} \left(\hat{b}_{\vec{k}, \lambda}^\dagger \hat{b}_{\vec{k}, \lambda} + \frac{1}{2} \right) \hbar \omega_{\lambda}(\vec{k}) \end{aligned} \quad (9.11)$$

It also means:

$$\hat{H}_{\text{ion}} = \hat{T}_{\text{ion}} + \hat{V}_{ii} + \hat{V}_{ei} \underset{\text{approximation}}{\approx}^{\text{harmonic}} \hat{V}_{ii}^{\text{static}} + \hat{H}_{\text{ph}} + \hat{V}_{ei}$$

Then we can rewrite:

$$\begin{aligned} \hat{H} &= \hat{T}_{\text{ion}} + \hat{T}_{ee} + \hat{V}_{ee} + \underbrace{\hat{V}_{ii}}_{\approx \hat{V}_{ii}^{\text{static}} + \hat{V}_{ii}^{\text{harmonic}}} + \underbrace{\hat{V}_{ei}}_{= \hat{V}_{ei}^{\text{static}} + \hat{V}_{e\text{-ph}}} = \\ &= \underbrace{\hat{T}_{\text{el}} + \hat{V}_{ee} + \hat{V}_{ei}^{\text{static}} + \hat{V}_{ii}^{\text{static}}}_{= \hat{H}_{\text{el}}} + \hat{H}_{\text{ph}} + \underbrace{\hat{V}_{ei}^{\text{dynamic}}}_{= \hat{V}_{e\text{-ph}}} + o_0(u^2) \end{aligned}$$

I.e. we are coupling dynamically electrons and ions via the electron-phonon interaction (neglected in the Born-Oppenheimer approximation). Here we have expanded in small displacements \vec{u}_α of the ions around the equilibrium position \vec{R}_α^0 with $\vec{R}_\alpha = \vec{R}_\alpha^0 + \vec{u}_\alpha$.

Note: From

$$D(\vec{k}) = -2 \sum_{\vec{R}^{(0)}} D(\vec{R}^{(0)}) \sin^2 \left(\frac{\vec{k} \cdot \vec{R}^{(0)}}{2} \right)$$

it follows for $\vec{k} \approx 0$:

$$D(\vec{k}) \sim k^2$$

$$\boxed{\omega_{\lambda}(\vec{k}) = c_{\lambda}(\theta_{\vec{k}}, \varphi_{\vec{k}}) k}$$

This is typical of acoustic phonons.

9.2. Electron-phonon interaction

Let us remember

$$\hat{V}_{ei} = - \sum_{j,\alpha} \frac{Ze^2}{\|\vec{R}_\alpha - \vec{r}_j\|}$$

with the electron density:

$$\hat{\rho}_{el}(\vec{r}) = \sum_{j=1}^{N_{el}} \delta(\vec{r} - \vec{r}_j)$$

So we can write:

$$\begin{aligned} \hat{V}_{ei} &= \int d^3r \hat{\rho}_{el}(\vec{r}) \left(- \sum_{\alpha=1}^{N_{ion}} \frac{Ze^2}{\|\vec{R}_\alpha - \vec{r}\|} \right) = \\ &= \int d^3r \hat{\rho}_{el}(\vec{r}) \sum_{\alpha=1}^{N_{ion}} v_{ei}(\vec{r} - \vec{R}_\alpha) \end{aligned}$$

We again expand in \vec{u}_α to first order:

$$v_{ei}(\vec{r} - \vec{R}_\alpha) \approx \underbrace{v_{ei}(\vec{r} - \vec{R}_\alpha^0)}_{\text{static contribution}} - \underbrace{\vec{\nabla}_{\vec{r}} v_{ei}(\vec{r} - \vec{R}_\alpha^0)}_{\text{electron-phonon contribution}} \vec{u}_\alpha$$

This gives

$$\hat{V}_{ei} = \hat{V}_{ei}^{\text{static}} + \hat{V}_{e-ph}$$

with:

$$\hat{V}_{e-ph} = \int d^3r \hat{\rho}_{el}(\vec{r}) \sum_{\alpha=1}^{N_{ion}} \vec{\nabla}_{\vec{r}} v_{ei}(\vec{r} - \vec{R}_\alpha^0) \hat{u}_\alpha$$

Remembering now

$$\hat{u}_\alpha = \frac{1}{\sqrt{N_{\text{cell}}}} \sum_{\vec{k} \in 1.\text{BZ}} \sum_{\lambda} l_{\vec{k},\lambda} \frac{1}{\sqrt{2}} (\hat{b}_{-\vec{k},\lambda}^\dagger + \hat{b}_{\vec{k},\lambda}) \vec{e}_{\vec{k},\lambda} e^{i\vec{k} \cdot \vec{R}_\alpha^0}$$

and (for electrons moving in a periodic potential)

$$\hat{\rho}_{el}(\vec{r}) = \frac{1}{V} \sum_{\vec{k}, \vec{p}, \sigma, n} e^{-i\vec{p} \cdot \vec{r}} \hat{c}_{\vec{k}+\vec{p}, n, \sigma}^\dagger \hat{c}_{\vec{k}, n, \sigma}$$

with the band index n , we get:

$$\hat{V}_{e-ph} = \sum_{\vec{k}, \sigma, n} \sum_{\vec{p}, \lambda} g_\lambda(\vec{p}) \hat{c}_{\vec{k}+\vec{p}, n, \sigma}^\dagger \hat{c}_{\vec{k}, n, \sigma} (\hat{b}_{-\vec{p}, \lambda}^\dagger + \hat{b}_{\vec{p}, \lambda})$$

TODO: Abb48

It can yield effective electron-electron interaction.

Appendix

Acknowledgements

My special thanks goes to Professor Grifoni, who gave this lecture and allowed me to publish this script of the lecture.

I would also like to thank all those, who found errors by careful reading and told me of them.

Andreas Völklein

GNU Free Documentation License

Version 1.3, 3 November 2008

Copyright © 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc.

<<https://fsf.org/>>

Everyone is permitted to copy and distribute verbatim copies of this license document,
but changing it is not allowed

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document “free” in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of “copyleft”, which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The “**Document**”, below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as “**you**”. You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A “**Modified Version**” of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A “**Secondary Section**” is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document’s overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The “**Invariant Sections**” are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The “**Cover Texts**” are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A “**Transparent**” copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not “Transparent” is called “**Opaque**”.

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, L^AT_EX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The “**Title Page**” means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, “Title Page” means the text near the most prominent appearance of the work’s title, preceding the beginning of the body of the text.

The “**publisher**” means any person or entity that distributes copies of the Document to the public.

A section “**Entitled XYZ**” means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as “**Acknowledgements**”, “**Dedications**”, “**Endorsements**”, or “**History**”.) To “**Preserve the Title**” of such a section when you modify the Document means that it remains a section “Entitled XYZ” according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that

these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution

and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled “Endorsements”, provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled “History” in the various original documents, forming one section Entitled “History”; likewise combine any sections Entitled “Acknowledgements”, and any sections Entitled “Dedications”. You must delete all sections Entitled “Endorsements”.

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an “aggregate”

if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the

present version, but may differ in detail to address new problems or concerns. See <https://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

Copyright © YEAR YOUR NAME.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation;

with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with ... Texts." line with this:

with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.