Quantum Field Theory I

by Prof. Achim Rosch Winter semester 2024/2025

Contents

| 1 | Clas | ssical H | Field Theory | 2 | | | | | | | |
|---|------|------------------------------------|---|----------|--|--|--|--|--|--|--|
| | 1.1 | Lagrange Formalism: Harmonic Chain | | | | | | | | | |
| | | 1.1.1 | A simple example: harmonic chain | 2 | | | | | | | |
| | | 1.1.2 | Continuum limit | 3 | | | | | | | |
| | | 1.1.3 | Euler-Lagrange equations for field theories | 4 | | | | | | | |
| | | 1.1.4 | Solution of the Euler-Lagrange equation | 6 | | | | | | | |
| | 1.2 | onal analysis | 7 | | | | | | | | |
| | 1.3 | Hamil | ton's formalism | 9 | | | | | | | |
| 2 | Seco | Second Quantization 12 | | | | | | | | | |
| | 2.1 | Canon | ical Quantization | 12 | | | | | | | |
| | | 2.1.1 | Commutator | 13 | | | | | | | |
| | | 2.1.2 | Phonons | 14 | | | | | | | |
| | 2.2 | Identie | cal particles and Second Quantization | 17 | | | | | | | |
| | | 2.2.1 | Hilbert Space | 17 | | | | | | | |
| | | 2.2.2 | Fock Space | 20 | | | | | | | |
| | | 2.2.3 | Creation and Annihilation Operators | 21 | | | | | | | |
| | 2.3 | Opera | tors in Fock space | 25 | | | | | | | |
| | | 2.3.1 | Basis change | 25 | | | | | | | |
| | | 2.3.2 | Field operators in Real and Momentum space | 26 | | | | | | | |
| | | 2.3.3 | Single particle operators | 28 | | | | | | | |
| | | 2.3.4 | Two-particle operators | 31 | | | | | | | |
| | | 2.3.5 | Hamiltonian with spin | 32 | | | | | | | |
| | | 2.3.6 | Dynamics of quantum fields | 34 | | | | | | | |
| | 2.4 | Electro | ons in solids | 35 | | | | | | | |
| | | 2.4.1 | Defining solids | 35 | | | | | | | |
| | | 2.4.2 | Bloch theorem | 38 | | | | | | | |
| | | 2.4.3 | Fermi sea and Fermi surface | 39 | | | | | | | |
| | | 2.4.4 | Lattice models | 40 | | | | | | | |
| | 2.5 | Hubba | ard and Heisenberg models | 43 | | | | | | | |
| | | 2.5.1 | Hubbard model | 43 | | | | | | | |
| | | 2.5.2 | Heisenberg model | 44 | | | | | | | |

| | 2.6 | Magne | tism and Spin waves | |
|---|-----|---|---|--|
| | | 2.6.1 | Magnon excitations in ferromagnets | |
| | | 2.6.2 | Magnon excitations in antiferromagnets | |
| 3 | Pat | h Integ | rals 57 | |
| | 3.1 | Constr | ucting path integrals | |
| | 3.2 | Applica | $ations \dots \dots$ | |
| | 3.3 | Statisti | ical physics and imaginary time | |
| | | 3.3.1 | Imaginary time | |
| | | 3.3.2 | Infinities in path integrals | |
| 4 | Fun | ctional | Integrals 70 | |
| | 4.1 | Functio | onal Integrals and Second Quantization | |
| | | 4.1.1 | Coherent states: Bosons | |
| | | 4.1.2 | Coherent states: Fermions | |
| | | 4.1.3 | Functional integrals for \mathcal{Z} | |
| | 4.2 | Gaussia | an integrals and Wick's theorem | |
| | 4.3 | Green's | s functions | |
| | | 4.3.1 | Definitions | |
| | | 4.3.2 | Free particle | |
| | | 4.3.3 | Matsubara frequencies | |
| | 4.4 | Summa | ations and examples | |
| 5 | Dia | gramat | ic Perturbation Theory 92 | |
| | 5.1 | Pertur | bation theory for Green's functions | |
| | 5.2 | Self en | ergy and Hartree-Fock approximation | |
| | 5.3 | Retarded Green's functions and Lehmann representation 105 | | |
| | 5.4 | Quasi j | particles $\ldots \ldots 110$ | |
| | 5.5 | Convergence of Perturbation Theory | | |
| | 5.6 | Screeni | ing of long-ranged Coulomb interactions | |
| | | 5.6.1 | Effective potential, polarization, and RPA approximation 121 | |
| | | 5.6.2 | Thomas Fermi approximation | |
| | | 5.6.3 | Plasma oscillations | |
| | | 5.6.4 | Absorption and particle-hole pairs | |
| | | 5.6.5 | Friedel oscillations | |

| | 5.7 | .7 Controlled approximations: large $N \ldots \ldots \ldots \ldots \ldots \ldots$ | | | |
|---|--|---|--|------------------------------------|-----------------------|
| | 5.8 | 8 Ladder approximation and Bethe-Salpeter equation | | | 0 |
| | | 5.8.1 Betl | he-Salpeter equation | 13 | 1 |
| | | 5.8.2 Lad | der approximation and bound states | 13 | 3 |
| | | 5.8.3 Out | look: Vertex corrections and conservation laws | 13 | 4 |
| | | | | | |
| | | | | | |
| 6 | Pro | bing Matte | er: Linear Response | 13' | 7 |
| 6 | Pro | b ing Matte Linear Res _l | er: Linear Response ponse Theory and Kubo formula | 13 ′ 13′ | 7 7 |
| 6 | Pro 6.1 6.2 | bing Matte Linear Resp Thermodyn | er: Linear Response ponse Theory and Kubo formula | 13 ' 13' 13' | 7 7 |
| 6 | Pro 6.1 6.2 6.3 | bing Matte Linear Resp Thermodyn Optical Cor | er: Linear Response bonse Theory and Kubo formula | 13' 13' 13' 14 | 7 7 9 3 |
| 6 | Pro 6.1 6.2 6.3 6.4 | bing Matte Linear Resp Thermodyn Optical Cor Scattering l | er: Linear Response bonse Theory and Kubo formula | 13' 13 13 14 14 | 7 7 9 3 7 |

Introduction

This course, Quantum Field Theory I, gives an introduction to quantum field theory (QFT), the language in which most of modern physics is formulated. Whether one wants to describe the basic properties of a piece of copper or the scattering of elementary particles in high energy colliders, the method of choice is QFT. QFT is a very active research field driven by experiment and theory. Nature challenges us to find new types of quantum field theory and new approximation methods.

The course focuses on the formulation of QFT based on functional integrals and discusses problems motivated by solid state physics. The methods are, however, equally relevant for high-energy physics. A central goal of QFT I is to learn diagrammatic perturbation theory, a powerful tool for analytic and also numerical theoretical methods. The lecture is complemented by the course QFT II, which covers topics like spontaneous symmetry breaking, phase transitions and renormalization group methods.

The lecture follows roughly (but not too closely) the book of Altland and Simon "Condensed Matter Field Theory", which is an excellent reference but there are also a number of other textbooks connecting QFT and solid-state theory, e.g., of Shankar, Fradkin, Nagaosa and others.

The skrip is partially based on the work of Lionel Jeevan Dmello who has typed in LaTeX formulas and created Feynman diagrams and most of the figures. In the winter semester 2024/2025, I have filled in the text. Unfortunately, the present version of the script has very many typos but I hope that it will, nevertheless, be useful to the reader.

Achim Rosch

1 Classical Field Theory

1.1 Lagrange Formalism: Harmonic Chain

Motivation: We will start the lecture with a simple example: a chain of harmonic oscillators. This will be our first example of a quantum theory with an infinite number of degrees of freedom and we will use it to clarify what a field is. We will use it to define two versions of a field theory: a discreet one and, importantly, also a continuum version where fields depend on space and time. The chapter also serves as a reminder of some of the concepts of analytical mechanics like Lagrange functions, Euler-Lagrange equations and actions. It turns out that these concepts from classical mechanics will become extremely important when formulating QFT. If you are not familiar (any more) with Lagrange functions and related concepts, it is now the best time to consult your favorite book on analytical mechanics.

1.1.1 A simple example: harmonic chain



Figure 1: Harmonic oscillator chain with rest length a

Consider a chain of point particles with masses 'm' coupled by harmonic springs (fig1). If the rest length of the springs is 'a', then, when the chain is in equilibrium, the position of the i^{th} mass is given by:

$$R = R_i^0 = a \cdot i \tag{1}$$

Using this, we define the position of the i^{th} mass with respect to its equilibrium position as:

$$\phi_i = R_i - R_i^0 \tag{2}$$

Therefore, we can write down the Lagrange function of the harmonic chain in terms of ϕ . As usually, the Lagrange function is given as the difference of kinetic and potential energy.

$$\mathcal{L} = T - V = \sum_{i} \frac{1}{2} m \dot{\phi}_{i}^{2} - \frac{\kappa}{2} (\phi_{i} - \phi_{i+1})^{2}$$
(3)

The corresponding action is the time-integral over the Lagrange function

$$S[\phi] = \int_{t_1}^{t_2} \mathcal{L}(\phi_i, \dot{\phi}_i) dt$$
(4)

The corner stone of analytical mechanics is Hamilton's principle. It states that the equation of motion can be obtained by asking the following question: for which functions $\phi(t)$ is S extremal provided that we fix the fields at the initial and final time ($\phi_i(t_1) = \phi_i^{\text{initial}}$ and $\phi_i(t_2) = \phi_i^{\text{final}}$). In short, we required that S does not vary to linear order, $\delta S = 0$, when we consider small variations of the fields, $\phi_i(t) + \delta\phi(t)$.

Outlook: Why is Hamilton's principle valid? While one can derive it from Newton's equation of motion, classical physics cannot give an answer to this 'why' question. Surprisingly, we will find that we will be able to explain this after we have understood how actions enter the quantum theory. But for this, we will have to wait for chapter 3.

1.1.2 Continuum limit

For the moment, we use our knowledge from analytical mechanics, where we learned that Hamilton's principle can be used to derive the Euler - Lagrange equation

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}_i} \right) = \frac{\partial \mathcal{L}}{\partial \phi_i} \tag{5}$$

Below, we will show how such a derivation works.

Our next goal is to develop a theory for weak excitations of our chain which is valid for wavelength λ large compared to the lattice spacing. We assume $\lambda \gg a$ and $\phi_{i+1} - \phi_i \ll a$.

In this limit, the individual particles labeled by i are not important any more and we can instead use a continuous field $\phi(x, t)$, with

$$x \equiv a \cdot n \tag{6}$$

$$\sqrt{a}\,\phi(a\cdot n,t) = \phi_n(t) \tag{7}$$

The prefactor \sqrt{a} in the second line is a convention. Note that 'x' in the above equation acts as a label, similar to *i* before, and should not be confused with the operator x in single-particle quantum mechanics.

As we are interested in the limit of smooth fields, $a \ll \lambda$, we can use Taylor expansions in a to analyze the theory

$$\phi_{n+1} - \phi_n = \sqrt{a}(\phi(x+a) - \phi(x)) \approx a^{\frac{3}{2}} \frac{\partial \phi}{\partial x}$$

and, $\sum_n \to \int_0^L \frac{1}{a} dx$ (8)

For example, we can approximate

$$\sum_{n} (\phi_n - \phi_{n+1})^2 \approx \int_0^L \frac{1}{a} dx \, \left(a^{\frac{3}{2}} \frac{\partial \phi}{\partial x}\right)^2 = a^2 \int_0^L \left(\frac{\partial \phi}{\partial x}\right)^2$$

We can use a similar manipulation for the first term in the action (4) to arrive at

$$S = \int_{t_1}^{t_2} dt \int_0^L dx \, \mathscr{L}(\phi, \partial_x \phi, \partial_t \phi) \tag{9}$$

$$\mathscr{L}(\phi, \partial_x \phi, \partial_t \phi) = \frac{m}{2} \dot{\phi}^2 - \frac{\kappa a^2}{2} (\partial_x \phi)^2 \tag{10}$$

Here \mathcal{L} is called **Lagrange density** which depends in our example on the field and its spatial and temporal derivatives. It is related to the Lagrange function by

$$L[\phi] = \int_0^L \mathcal{L}(\phi, \partial_x \phi, \partial_t \phi) dx$$

1.1.3 Euler-Lagrange equations for field theories

To obtain the equation of motion, we apply Hamiliton's principle. Thus we consider small variations ($\epsilon \ll 1$) of the fields, while keeping them fixed at the initial and finial time.

$$\phi(x,t) \to \phi(x,t) + \varepsilon \eta(x,t)$$

with $\varepsilon \ll 1$, and $\eta(x,t_1) = 0 = \eta(x,t_1) \,\forall x$ (11)

Our goal is to search for solutions

$$\delta S[\phi] \stackrel{!}{=} 0 \tag{12}$$

where we compute the change of action linear in ϵ from

$$\delta S = S[\phi - \varepsilon \eta] - S[\phi] = \varepsilon \int dt dx \, \frac{\partial \mathscr{L}}{\partial \phi} \eta + \frac{\partial \mathscr{L}}{\partial (\partial_t \phi)} \dot{\eta} + \frac{\partial \mathscr{L}}{\partial (\partial_x \phi)} \partial_x \eta + \mathcal{O}(\varepsilon^2)$$

The next step, is to perform a partial integration of the second and third term using

$$\frac{\partial \mathscr{L}}{\partial (\partial_t \phi)} \dot{\eta} = \frac{d}{dt} \left(\eta \frac{\partial \mathscr{L}}{\partial (\partial_t \phi)} \right) - \eta \frac{d}{dt} \frac{\partial \mathscr{L}}{\partial (\partial_t \phi)}$$
$$\frac{\partial \mathscr{L}}{\partial (\partial_x \phi)} \partial_x \eta = \frac{d}{dx} \left(\eta \frac{\partial \mathscr{L}}{\partial (\partial_x \phi)} \right) - \eta \frac{d}{dx} \frac{\partial \mathscr{L}}{\partial (\partial_x \phi)}$$

to arrive at

$$\int dt dx \,\varepsilon \eta \left(\frac{\partial \mathscr{L}}{\partial \phi} - \frac{d}{dt} \frac{\partial \mathscr{L}}{\partial (\partial_t \phi)} - \frac{d}{dx} \frac{\partial \mathscr{L}}{\partial (\partial_x \phi)} \right) \\ + \varepsilon \int dx \,\eta \, \frac{\partial \mathscr{L}}{\partial (\partial_t \phi)} \Big|_{t=t_1}^{t=t_2} + \varepsilon \int dt \,\eta \, \frac{\partial \mathscr{L}}{\partial (\partial_x \phi)} \Big|_{x=0}^{x=L} \stackrel{!}{=} 0 \quad (13)$$

where the second term vanishes due to our boundary conditions (11), $\int dx \, \eta \, \frac{\partial \mathscr{L}}{\partial(\partial_t \phi)} \Big|_{t=t_1}^{t=t_2} = 0$. The third term will be discussed below but for now we focus on the first term.

Equation 13 has to hold for all possible $\eta(x, t)$. This is only possible, if the bracket vanishes

$$\frac{d}{dt}\frac{\partial\mathscr{L}}{\partial(\partial_t\phi)} + \frac{d}{dx}\frac{\partial\mathscr{L}}{\partial(\partial_x\phi)} = \frac{\partial\mathscr{L}}{\partial\phi}$$
(14)

Equation 14 is the Euler-Lagrange equation for Lagrange density \mathscr{L} .

In the example above, we were considering a single real field $\phi(x,t)$ in one space dimension. We can easily generalize this to vector field $\vec{\phi}(\vec{x})$ in *d* dimensions with $\vec{x} = (x_0, x_1, \dots, x_d), x_0 = t$. In this case, the Euler-Lagrange equation is written as

$$\sum_{\mu=0}^{d} \frac{d}{dx_{\mu}} \frac{\partial \mathscr{L}}{\partial(\partial_{\mu}\phi_{i})} = \frac{\partial \mathscr{L}}{\partial\phi_{i}}$$
(15)

where the sum includes both time and space derivatives.

We have still to discuss the last term in Eq. (13). It describes the effects of spatial boundary conditions at x = 0 and x = L. Here, we usually need some input from physics. What situation do we want to describe? We consider the following cases:

- Fixed bounday conditions: Figure 2a. Here $\phi(0) = 0 = \phi(L) \Rightarrow \eta(0) = 0 = \eta(L)$. This is, e.g., relevant for the string of a violin which is fixed at its ends.
- **Periodic boundary conditions:** Figure 2b. Here $\phi(0) = \phi(L)$. This is the most 'convenient' boundary condition for theoretical calculations, because the resulting system is translationally invariant, but it can only be used in cases where one is not interested in boundary effects.
- **Open boundaries:** Figure 2c. Here $\eta(0), \eta(L)$ are arbitrary. This implies, from the boundary condition in equation 13, that $\frac{\partial \mathscr{L}}{\partial(\partial_x \phi)}\Big|_{x=0}^{x=L} = 0$. Here, $\partial_x \phi \Big|_{x=0}^{x=L} = 0$, which means there is no tension. This is relevant, e.g., for the boundary of a solid.



(c) Open boundaries

Figure 2: Boundary conditions for harmonic chain

1.1.4 Solution of the Euler-Lagrange equation

Now let us quickly solve the harmonic chain in the continuum limit. The Lagrange density was given by

$$\mathscr{L}(\phi,\partial_t\phi,\partial_x\phi) = \frac{1}{2}m\dot{\phi}^2 - \frac{\kappa a^2}{2}(\partial_x\phi)^2 \tag{16}$$

Plugging this into the Euler-Lagrange equation (equation 14) we get the following equation of motion

$$m\ddot{\phi} - \kappa a^2 \partial_x^2 \phi = 0 \tag{17}$$

which is the familiar "wave equation":

$$\frac{\partial^2 \phi}{\partial t^2} - c^2 \frac{\partial^2 \phi}{\partial x^2} = 0 \tag{18}$$

where we can identify the constant $c = \sqrt{\frac{\kappa a^2}{m}}$ with the velocity of sound.

For an infinitely large system, a general solution of the equation can be written as

$$\phi(x,t) = \phi_R(x-ct) + \phi_L(x+ct),$$

with ϕ_R, ϕ_L being arbitrary functions.

For fixed boundaries i.e., $\phi(0) = 0 = \phi(L)$, it is convenient to expand the solutions into standing waves

$$\phi(x,t) = \sum a_n \sin(\kappa_n x) \cos(\omega_n t + \varphi_n),$$

with $\kappa_n = \frac{n\pi}{L}$ and $\omega_n^2 = c^2 \kappa_n^2$.

Outlook: Above, we considered the wave equation as the solution of a toy model. It turns out that such simple wave equations (or, more precisely, their quantum

version) can be used to describe *exactly* the sound excitations of solids in the long wavelength and small amplitude regime, even for complex solids. In this case, one can view the Lagrange density (16) (and its three-dimensional generalizations) as the leading term in a Taylor expansion of \mathcal{L} in terms of field and their derivatives. One can then show, that higher-order terms have a vanishingly small contribution whenever the wavelength is much larger than the lattice constant.

1.2 Functional analysis

Motivation: In this section, we learn to know a few basic terms from functional analysis which are useful. We are not aiming at any rigor and we will not discuss any complicated situations (like fields defined on curved manifolds).

The central object is a field theory is – obviously – the field.

Field:

A field ϕ is a map from a manifold M to a target manifold T.

$$\phi: M \longmapsto T$$
$$z \in M \longmapsto \phi(z) \in T$$

Example: Manifold $M = [t_1, t_2] \times [0, L]$, z = (t, x). Target Manifold $T = \mathbb{R}$ or $T = \{\text{Operators on some Hilbert space}\}$

Central to our program are objects like the action. An action is a number computed for a given field configuration in space and time. Such objects are called functionals.

Functional:

A Functional F maps fields to real (or complex) numbers.

$$F:\phi\longmapsto F[\phi]\in\mathbb{R}$$



Figure 3: Visualisation of a mapping from fields to real numbers

Example: Action $S[\phi] = \iint dx dt \,\mathscr{L}(\phi, \partial_t \phi, \partial_x \phi)$

As a next step, we generalize the notion of derivative from functions to functionals. A standard partial derivative is defined as

$$\frac{\partial f(\vec{x})}{\partial x_i} = \left. \frac{f(\vec{x} + \varepsilon \hat{e}_i) - f(\vec{x})}{\varepsilon} \right|_{\varepsilon \to 0} \tag{19}$$

where \hat{e}_i is a unit vector in direction i = x, y, z. The analog of the partial derivative is the functional derivative.

Functional derivative:

$$\frac{\delta F[\phi]}{\delta \phi(z_0)} = \lim_{\varepsilon \to 0} \frac{F[\phi(z) + \varepsilon \delta(z - z_0)] - F[\phi(x)]}{\varepsilon}$$
(20)

where $\delta(z - z_0)$ is the dirac delta function centered at $z = z_0$. Here we change the argument of the field only at one point z_0 and track how the functional changes. If the field is defined as function of space and time, we use instead $\delta^d(\vec{r} - \vec{r_0})\delta(t - t_0)$ when calculating the variation with respect to the field $\phi(\vec{r_0}, t_0)$. And in case of a mulicomponent field, we multiply the δ function with \hat{e}_i

$$\frac{\delta F[\phi]}{\delta \phi_i(\vec{r_0}, t_0)} = \lim_{\varepsilon \to 0} \frac{F[\phi(z) + \varepsilon \hat{e}_i \delta(\vec{r} - \vec{r_0}) \delta(t - t_0)] - F[\phi(x)]}{\varepsilon}$$
(21)

We also generalise the standard Taylor expansion

$$f(\vec{x} + \delta \vec{x}) = f(\vec{x}) + \sum_{i} \frac{\partial f}{\partial x_i} \delta x_i + \mathcal{O}(\delta x_i^2)$$
(22)

from functions to functionals and obtain

$$F[\phi_0 + \Delta\phi] = F[\phi_0] + \int dz' \,\frac{\delta F}{\delta\phi(z')} \Delta\phi(z') + \mathcal{O}(\Delta\phi^2) \tag{23}$$

(24)

Note that the summation has been replace by an integral over z.

To check that we did everything correctly, let us consider $\Delta \phi = \varepsilon \delta(z - z_0)$. We obtain

$$F[\phi_0 + \varepsilon \delta(z - z_0)] = F[\phi_0] + \int dz' \frac{\delta F}{\delta \phi(z')} \varepsilon \delta(z' - z_0) + \mathcal{O}(\varepsilon^2)$$

= $F[\phi_0] + \varepsilon \frac{\delta F}{\delta \phi(z_0)} + \mathcal{O}(\varepsilon^2)$ (25)

which is, indeed, consistent with our definition (20).

As an example, we recall, how we derived the Euler-Lagrange equation by considering variations of the action $S[\phi] = \iint dx dt \mathscr{L}(\phi_i, \partial_\mu \phi_i)$. We obtained

$$\delta S = S[\phi + \varepsilon \eta] - S[\phi] = \iint dx dt \,\varepsilon \eta_i(x, t) \left(\frac{\partial \mathscr{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial (\partial_\mu \phi_i)}\right) + \text{boundary terms}$$

Comparing this to Eq. (23), we can identify

$$\frac{\delta S}{\delta \phi_i(x_0, t_0)} = \frac{\partial \mathscr{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial (\partial_\mu \phi_i)}$$

Thus, we have a new way to write the Euler-Lagrange equations

$$\delta S = 0 \iff \frac{\delta S}{\delta \phi} = 0 \iff \text{Euler-Lagrange.}$$
 (26)

Finally, we would like to point out that one can use standard product and chain rules for functionals, for example

- Product rule: $\frac{\delta(FG)}{\delta\phi} = \frac{\delta F}{\delta\phi} \cdot G + F \cdot \frac{\delta G}{\delta\phi}$
- Chain rule: $\frac{\delta g(F)}{\delta \phi} = g'(F) \cdot \frac{\delta F}{\delta \phi}$

where we denote functionals by capital letters, F, G, and functions g by small letters.

1.3 Hamilton's formalism

Motivation: Besides the Lagrange formalism, the Hamilton formalism is the second cornerstone of analytical mechanics. It is especially useful to take the step from classical mechanics to quantum mechanics (by replacing Poisson brackets by commutators). A similar step (called canonical quantization) can also be taken in field theory – but we will later also find an approach of formulating quantum theories in terms of the Lagrange formalism only. For now, we will focus on the question of how to generalize concepts of momentum and Poisson brackets to fields defined in the continuum.

We start with a reminder of how one introduces momenta and the Hamilton function in classical mechanics. Starting from the Lagrange formalism the momentum p_i conjugate to position q_i is obtain from

$$p_i = \frac{\partial \mathcal{L}}{\partial q_i} \tag{27}$$

Thus, we can easily also define a **momentum field** using our new knowledge of functional derivatives

$$\Pi(x) = \frac{\delta \mathcal{L}}{\delta \dot{\phi}(x)} = \frac{\partial \mathscr{L}(\phi, \partial_x \phi, \partial_t \phi)}{\partial (\partial_t \phi)}$$
(28)

Similarly, the Hamiltonian function in classical physics is defined by

$$H(q_i, p_i) = \sum_i \vec{p}_i \partial_t q_i - L(q_i, \partial_t q_i)|_{\partial_t q_i = \partial_t q_i(q_j, p_j)}$$
(29)

where, importantly, one has to replace all time derivatives of fields by momenta using Eq. (27).

We can generalize this directly to our field theory by defining

Hamilton density:
$$\mathscr{H}(\phi, \partial_x \phi, \Pi) = \Pi \dot{\phi} - \mathscr{L}(\phi, \partial_x \phi, \dot{\phi})\Big|_{\dot{\phi} = \dot{\phi}(\phi, \partial_x \phi, \Pi)}$$

Hamilton function: $\mathcal{H}[\phi, \Pi] = \int d^d \vec{x} \,\mathscr{H}$

As above, one has to eliminate time derivatives of fields using Eq. (28).

Let us apply this to our prime example of a field theory, the harmonic chain in the continuum limit. Using the Lagrange density from Eq. (16) to calculate the momentum field we obtain

$$\Pi(x) = \frac{\partial \mathscr{L}(\phi, \partial_x \phi, \partial_t \phi)}{\partial (\partial_t \phi)} = m\dot{\phi}(x).$$

From this we can calculate the Hamilton density and hence the Hamilton function:

$$\mathcal{H}(\phi, \partial_x \phi, \Pi) = \Pi \frac{\Pi}{m} - \frac{\Pi^2}{2m} + \frac{\kappa a^2}{2} (\partial_x \phi)^2$$
$$\implies \mathcal{H} = \frac{\Pi^2}{2m} + \frac{\kappa a^2}{2} (\partial_x \phi)^2$$
$$\mathcal{H}[\phi, \Pi] = \int dx \left(\frac{\Pi^2}{2m} + \frac{\kappa a^2}{2} (\partial_x \phi)^2\right).$$
(30)

We can easily recognize the kinetic and potential energy expressed in terms of the two fields $\Pi(x,t)$ and $\phi(x,t)$.

Finally, we want to see how **Poisson brackets** are modified in a continuum theory. Poisson brackets are extremely useful: one can use them to write the Hamiltonian equations of motions in a simplified way, they are convenient to discuss the things like canonical transformations or the mathematical structure of time-evolution, and, finally, they are useful for the step from classical physics to quantum physics, where commutators take over the role of Poisson brackets.

To introduce Poisson brackets, we take a step back from the continuum theory to the discretized theory where we originally started. For discretized field ϕ_n and the associated momenta $\Pi_n = \frac{\partial \mathcal{L}(\phi_n, \dot{\phi}_n)}{\partial \dot{\phi}_n}$, the Poisson brackets are defined as

$$\{A,B\} = \sum_{n} \frac{\partial A}{\partial \Pi_{n}} \frac{\partial B}{\partial \phi_{n}} - \frac{\partial A}{\partial \phi_{n}} \frac{\partial B}{\partial \Pi_{n}}$$

With this definition, the fundamental Poisson brackets of positions and momenta are given by

$$\{\Pi_n, \phi_m\} = \delta_{nm} \tag{31}$$

Now we have to take the continuum limit of those equations. As a first step, we recall that we definied $\phi_n = \sqrt{a}\phi(x = na)$. To be consistent with our definition of $\Pi(x)$, Eq. (28), we have to identify $\Pi_n = \Pi(an) \cdot \frac{1}{\sqrt{a}}$. Using these prefactors and the usual rules for approximating sums by integrals, Eq. (8) we obtain

$$\{A,B\} = \sum \frac{\partial A}{\partial \Pi_n} \frac{\partial B}{\partial \phi_n} - \frac{\partial A}{\partial \phi_n} \frac{\partial B}{\partial \Pi_n} \quad \Rightarrow \quad \{A,B\} = \int dx \frac{\delta A}{\delta \Pi_n} \frac{\delta B}{\delta \phi_n} - \frac{\delta A}{\delta \phi_n} \frac{\delta B}{\delta \Pi_n} \\ \left\{ \frac{\Pi_n}{\sqrt{a}}, \frac{\phi_m}{\sqrt{a}} \right\} = \frac{\delta_{nm}}{a} \quad \Rightarrow \quad \{\Pi(x), \phi(x')\} = \delta(x - x') \tag{32}$$

Finally, we can use the definitions collected above to write **Hamilton's equation** of motion in a compact form

$$\dot{\Pi}(x,t) = \{\mathcal{H}, \Pi(x,t)\} = -\frac{\delta \mathcal{H}(t)}{\delta \phi(x,t)}$$

$$\dot{\phi}(x,t) = \{\mathcal{H}, \phi(x,t)\} = -\frac{\delta \mathcal{H}(t)}{\delta \Pi(x,t)}$$
(33)

In this first section, we have – with very large speed and without providing detailed derivations – sketched how we can obtain classical field equations. We were able to use the formalism typically introduced in analytical mechanics classes to analyze field theories. Our prime example of a field theory has been the harmonic chain but we can use very similar approaches to discuss, e.g., electrodynamics and the associated Maxwell equations.

Our next major goal will be to find a way of how to quantize the classical field.

2 Second Quantization

The figure 4 gives an overview over different routes which can be taken to obtain a quantum field theory. One option is to begin from classical field theory and to derive quantum field theory via a procedure known as 'canonical quantization', which will be subject of the following chapter. Later, we will learn about a direction connection of classical actions and quantum field theory using directly the action. Here, one has to note that one cannot 'derive' a quantum theory from a classical theory – while there is no problem in proceeding in the opposite direction. A more reliable approach is, however, to start from our knowledge of few-particle quantum mechanics, which we will be able to generalize to quantum field theory, using a procedure called second quantization, which is the focus of Sec. 2.2.



Figure 4: Different ways to arrive at QFT

2.1 Canonical Quantization

Motivation: In this chapter, we learn a relatively straightforward way to quantize a field theory, called canonical quantization, which is based on simply replacing Poisson brackets by commutators. This works perfectly for the example of the harmonic chain and allows to introduce phonons, the quantum excitations of a vibrating solids. As a side effect, this section will serve as a reminder on how one solves harmonic oscillators problems algebraically. The corresponding algebra will be important when learning the math of second quantization in the next chapter.

2.1.1 Commutator

The transition from classical physics to quantum physics is not unique. This is obvious from the fact, that we have to replace commuting fields by non-commuting operators. If you have an expression like AB (e.g., A = x and B = p) in classical physics, it is unclear whether you should replace it by $\hat{A}\hat{B}$, $\hat{B}\hat{A}$ or something like $\sqrt{\hat{A}\hat{B}}\sqrt{\hat{A}}$ on the quantum side, if \hat{A} and \hat{B} are non-commuting operators.

But a successful rule of thumb for the transition to quantum mechanics, is to replace the Poisson brackets from classical physics by commutators following the recipe:

$$\begin{aligned}
\mathcal{H}(p_i, q_j) &\longrightarrow \mathcal{H}(\hat{p}_i, \hat{q}_j) \\
\{p_i, q_j\} &= \delta_{ij} \longrightarrow [\hat{p}_i, \hat{q}_j] = -i\hbar\delta_{ij}
\end{aligned} \tag{34}$$

What is the justification for that? The reason is that we want to keep the equation of motion compatible

$$\partial_t A = \{\mathcal{H}, A\} \longrightarrow -i\hbar \partial_t \hat{A} = [\hat{\mathcal{H}}, \hat{A}]$$
(35)

Thus, we can hope that the quantum theory will reproduce in the classical limit the correct equations.

Therefore, **canonical quantization** means that we follow precisely those rules also in a field theory by postulating that we simply replace number-valued fields by operator-valued fields

$$\{\Pi_{\alpha}(\vec{x}),\phi_{\beta}(\vec{x}')\} = \delta_{\alpha\beta}\delta^{d}(\vec{x}-\vec{x}') \longrightarrow \left[\hat{\Pi}_{\alpha}(\vec{x}),\hat{\phi}_{\beta}(\vec{x}')\right] = -i\hbar\delta_{\alpha\beta}\delta^{d}(\vec{x}-\vec{x}') \quad (36)$$

where $\hat{\pi}$ and $\hat{\phi}$ are now operators. The following remarks apply

- ϕ , Π are real-number valued fields in the equation above, and, thus, $\hat{\phi}$, $\hat{\Pi}$ are Hermitian operators (which have real eigenvalues).
- To define a quantum theory, we should better fix the Hilbert space we will do this later.
- For fermions there are similar rules, where one uses anti-commutators instead of commutators, but we will not use this in the lecture but go for "second quantization" instead.
- The does not work directly if $\Pi_i = \frac{\partial \mathscr{L}}{\partial(\partial_t \phi_i)}$ has no unique solution which happens, e.g., in Gauge theories, with $\frac{\partial \mathscr{L}}{\partial(\partial_t A^0)} = 0$. For such situations, there exist alternative methods for quantization (quantization with constraints) invented by Dirac, but we will not do this in this lecture.

2.1.2 Phonons

We want to apply this program directly to our harmonic chain. The Hamilton operator is obtained by replacing in the Hamilton function, Eq. (30),

$$\hat{\mathcal{H}} = \int_0^L dx \, \left(\frac{\hat{\Pi}^2(x)}{2m} + \frac{\kappa a^2}{2} (\partial_x \hat{\phi})^2 \right) \tag{37}$$

To obtain something which looks more familiar and to simplify the theory, we write this in the momentum basis by Fourier transforming the fields as follows: **Fourier transformed fields:**

$$\hat{\phi}(x) = \frac{1}{\sqrt{L}} \sum_{n} \hat{\phi}_{k_n} e^{-ik_n x}$$
$$\hat{\Pi}(x) = \frac{1}{\sqrt{L}} \sum_{n} \hat{\Pi}_{k_n} e^{ik_n x}$$
(38)

with, $k_n = \frac{2\pi}{L}n$ for periodic boundary conditions

Inverse Fourier Transforms:

$$\hat{\phi}_{k_n} = \frac{1}{\sqrt{L}} \int_0^L dx \, e^{ik_n x} \hat{\phi}(x)$$

$$\hat{\Pi}_{k_n} = \frac{1}{\sqrt{L}} \int_0^L dx \, e^{-ik_n x} \hat{\Pi}(x)$$
and, $\hat{\phi}_{k_n}^{\dagger} = \hat{\phi}_{-k_n}, \hat{\Pi}_{k_n}^{\dagger} = \hat{\Pi}_{-k_n}$ with $\hat{\phi}^{\dagger}(x) = \hat{\phi}(x)$ as $\hat{\Pi}^{\dagger}(x) = \hat{\Pi}(x)$

$$(39)$$

Using this we can compute the commutation relations of the Fourier-transformed fields

$$\begin{bmatrix} \hat{\Pi}_{k_n}, \hat{\phi}_{k_m} \end{bmatrix} = \frac{1}{L} \int_0^L dx \int_0^L dx' \, e^{-ik_n x} \left[\hat{\Pi}(x), \hat{\phi}(x') \right] e^{ik_m x'} = -i\hbar\delta_{nm}$$
where we have used $\begin{bmatrix} \hat{\Pi}(x), \hat{\phi}(x') \end{bmatrix} = -i\hbar\delta(x - x')$
and $\frac{1}{L} \int_0^L dx \, e^{-i(k_n - k_m)x} = \delta_{k_n, k_m}$
(40)

We now rewrite the Hamiltonian in terms of the new fields and obtain after some straightforward algebra

$$\hat{\mathcal{H}} = \sum_{n} \frac{1}{2m} \hat{\Pi}_{k_n} \hat{\Pi}_{-k_n} + \frac{\kappa a^2}{2} k_n^2 \hat{\phi}_{k_n} \hat{\phi}_{-k_n} \tag{41}$$

where we used again that $\frac{1}{L} \int_0^L dx \, e^{\pm i(k_n+k_m)x} = \delta_{k_n,-k_m}$. What have we gained? In the new Hamiltonian, the momenta are decoupled, k_n only couples to $-k_n$ and we 'almost' obtain for each k_n something like a Harmonic oscillator with a springconstant $\kappa(k_n a)^2$ and a corresponding oscillator frequency ω_{k_n} with

$$\omega_k = \sqrt{\frac{\kappa a^2}{m}} |k| = c|k|$$

where c is the speed of sound defined above.

It is now the time to recall how one solves the harmonic oscillator problem in quantum mechanics using simple algebraic models. Here we use it as a 'trick' to simply our problem, later we will see that the math is useful way beyond the problem of a harmonic oscillator. Below, I assume that the reader is familiar with this solution – if not it is time to look this up.

The solution of the Harmonic oscillator starts by introducing new operators a and a^{\dagger} defined as

$$a^{\dagger} = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} - \frac{i}{m\omega} \hat{p} \right), \ a = (a^{\dagger})^{\dagger}$$
 (42)

The prefactors in the formula above have been chosen for two reason: first, to obtain simple commutation relations

$$[a, a^{\dagger}] = \frac{m\omega}{2\hbar} \left[\hat{x} + \frac{i}{m\omega} \hat{p}, \hat{x} - \frac{i}{m\omega} \hat{p} \right] = 1 \text{ using, } [\hat{x}, \hat{p}] = i\hbar$$
(43)

The second reason for choosing the prefactors was to simplify the Hamiltonian

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 \hat{x}^2 = \hbar\omega \left(a^{\dagger}a + \frac{1}{2}\right)$$

The operator $\hat{n} = a^{\dagger}a$ can then shown to have integer eigenvalues and corresponding eigenstates $|n\rangle$, with $\hat{n}|n\rangle = n|n\rangle$, $n \in \mathcal{N}_0$. The groundstate of the Harmonic oscillator can be obtained from the equation $a|0\rangle = 0$.

It turns out, that we can apply exactly the same trick to our field theory ("hats" on operators are dropped for convenience). We define raising and lowering operators for each momentum k separately

$$a_{k} = \sqrt{\frac{m\omega_{k}}{2\hbar}} \left(\phi_{k} + \frac{i}{m\omega_{k}}\Pi_{-k}\right); \quad a_{k}^{\dagger} = \sqrt{\frac{m\omega_{k}}{2\hbar}} \left(\phi_{-k} - \frac{i}{m\omega_{k}}\Pi_{k}\right)$$
(44)

This gives rise to simple commutation relations

$$[a_k, a_{k'}^{\dagger}] = \delta_{kk'}$$

$$[a_k, a_{k'}] = [a_k^{\dagger}, a_{k'}^{\dagger}] = 0$$
(45)

and we can reexpress the original fields in the new operators

$$\phi_k = \frac{1}{2} \sqrt{\frac{2\hbar}{m\omega_k}} (a_k + a_{-k}^{\dagger}); \quad \Pi_k = \frac{i}{2} \sqrt{2\hbar m\omega_k} (a_k^{\dagger} - a_{-k})$$
(46)

Substituting equations (46) into the Hamiltonian, Eq. (41), we get: (using $\omega_k = c \cdot |k| = \omega_{-k}$)

$$\begin{aligned} \mathcal{H} &= \sum_{k} \frac{1}{2m} \Pi_{k} \Pi_{-k} + \frac{1}{2} m \omega_{k}^{2} \phi_{k} \phi_{-k} \\ &= \sum_{k} \frac{1}{2m} (-1) \frac{1}{4} 2m \hbar \omega_{k} (a_{k}^{\dagger} - a_{-k}) (a_{-k}^{\dagger} - a_{k}) + \frac{1}{2} m \omega_{k}^{2} \frac{1}{4} \frac{2\hbar}{m \omega_{k}} (a_{k} + a_{-k}^{\dagger}) (a_{-k} + a_{k}^{\dagger}) \\ &= \frac{\hbar \omega_{k}}{4} \sum_{k} (-a_{k}^{\dagger} a_{-k}^{\dagger} + a_{k}^{\dagger} a_{k} + a_{-k} a_{-k}^{\dagger} - a_{-k} a_{k} + a_{k} a_{-k} + a_{k} a_{k}^{\dagger} + a_{-k}^{\dagger} a_{-k} + a_{-k}^{\dagger} a_{k}^{\dagger}) \end{aligned}$$

Using $[a_k, a_{k'}] = [a_k^{\dagger}, a_{k'}^{\dagger}] = 0$, Eq. (45) we get

$$\mathcal{H} = \frac{\hbar\omega_k}{4} \sum_k (a_k^{\dagger}a_k + a_{-k}a_{-k}^{\dagger} + a_ka_k^{\dagger} + a_{-k}^{\dagger}a_{-k})$$

Finally, we simplify this using $a_k a_k^{\dagger} = 1 + a_k^{\dagger} a_k$ and $a_{-k} a_{-k}^{\dagger} = 1 + a_{-k}^{\dagger} a_{-k}$ using Eq. (45). Thus we obtain a very simple result

$$\mathcal{H} = \sum_{k} \hbar \omega_k \left(a_k^{\dagger} a_k + \frac{1}{2} \right) \tag{47}$$

We can also rewrite our field operator in terms of a_k^{\dagger} and a_k using our definitions,

$$\phi(\vec{x}) = \frac{1}{\sqrt{L}} \sum_{k} \phi_k e^{-i\vec{k}\cdot\vec{x}} = \frac{1}{\sqrt{L}} \sum_{k} \sqrt{\frac{\hbar}{2m\omega_k}} (a_k + a_{-k}^{\dagger}) e^{-i\vec{k}\cdot\vec{x}}$$

to arrive at

$$\phi(\vec{x}) = \frac{1}{\sqrt{L}} \sum_{k} \sqrt{\frac{\hbar}{2m\omega_k}} \left(a_k^{\dagger} e^{i\vec{k}\cdot\vec{x}} + a_k e^{-i\vec{k}\cdot\vec{x}} \right)$$
(48)

Now we come to the most important and most interesting part of this exercise, the interpretation of the result.

We start with a perhaps, unexpected, question. What defines a (quantum) particle? First, a particle is something you can count, one can have 0, 1, 2... particles. Second, we expect that it carries momentum $\hbar k$ and an energy $E(p) = E(\hbar k)$ and we should be able to count how many particles we have with a given momentum k_n . (With a classical particle, one can also associate both position and momentum with, but in quantum physics this is not done directly because of Heisenberg uncertainty principle. To approximate a classical particle one has to build wave packets using, e.g., superposition of different k modes).

Now we look at our Hamiltonian. The operator $a_k^{\mathsf{T}} a_k$, has spectrum \mathbb{N}_0 . We now take the bold step of interpreting this as the number of **particles** with a given momentum $\hbar k$. Those quantum particles get the name **phonons**. Their energy is $\hbar \omega_k = c |\hbar k|$. Thus, the total energy of the system is obtained by multiplying the energy of a particle with the number of particles then simply $\hbar \omega_k$ times the number of particles.

Here, the operator a_k^{\dagger} 'creates' a phonon with momentum k. Why is this so? It increases $\hat{n}_k = a_k^{\dagger} a_k$ by one. Similarly, a^{\dagger} can destroy a phonon. We will come back to these formulas in more detail in the next chapter.

One cannot overemphasize the importance of this insight which follows from a straightforward reinterpretation of the math. We found that the excitations of the harmonic chain are a new type of particle, the phonon, whose property has almost nothing to do with the particles from which the chain is made. Phonons describe quantized collective excitations of the harmonic chain and like other particles, they are characterized by momentum and energy.

Note that precisely the same physics and same arguments can be used to quantize the electromagnetic field by introducing photons – this is a topic covered in the advanced quantum mechanics course.

2.2 Identical particles and Second Quantization

Motivation: We will now take a completely different route to quantization of manyparticle systems. We will start from the quantum theory of single particles and build from there the Hilbert space of a many-particle system. We will see that the main insight will come by thinking about the question that two elementary particles can – as a matter of principle – not be distinguished from each other. We will see that this determines the mathematical structure of the theory and will also guides us to find the ideal type of operators to be used in such settings.

2.2.1 Hilbert Space

Our starting point is single-particle quantum mechanics, where we pick some arbitrary (but complete) set of basis states of the single particle Hilbert space. **Basis:** 1-particle Hilbert space \mathscr{H}_1 with basis $|\alpha_i\rangle$ or corresponding wave functions $\psi_i(\vec{r})$

Now we want to consider n quantum particles instead. For the moment, we assume that they are **distinguishable** (we have, e.g., one neutron, one proton and one electron. Thus, the Hilbert space \mathscr{H}_n^d of n distinguishable particles is just the (tensor-) product of n single particle Hilbert spaces. For an N-dimensional single-particle Hilbert space this means that there are now N^n basis vectors , the dimension This means, that the N^n basis vectors $|\alpha_{j_1}\rangle|\alpha_{j_2}\rangle\cdots |\alpha_{j_n}\rangle$

$$\mathcal{H}_{n}^{d} = \underbrace{\mathcal{H}_{1} \otimes \cdots \otimes \mathcal{H}_{1}}_{n \text{ times}}$$
$$|\psi\rangle = \sum_{\substack{c \\ j_{1}j_{2}\cdots j_{n} \\ \text{amplitude}}} \underbrace{|\alpha_{j_{1}}\rangle|\alpha_{j_{2}}\rangle \cdots |\alpha_{j_{n}}\rangle}_{\text{particle } k \text{ in state } j_{k}}$$

Alternatively, we can use the position basis and the wavefunction has the form

$$\psi = \psi(r_1, \cdots, r_n)$$

In the latter formulation, we just specify where each of the particle sits and $\psi(r_1, \dots, r_n)$ is the corresponding quantum amplitude.

But what should we do when two particles are **indistinguishable**? Take two electrons (with the same spin). Consider $\psi(a, b)$ and $\psi(b, a)$. The $\psi(a, b)$ describes that particle 1 sits at position a while particle 2 is located at b. In contrast, $\psi(b, a)$ is the quantum amplitude for particle 1 being at b, while 2 is at a. But this makes no sense, when 1 and 2 are indistinguishable! In this case,

 $\psi(r_1, r_2)$ and $\psi(r_2, r_1)$ have to describe same state.

Next, we have to find out, how we can adjust the mathematical formulation to take into account this fundamental fact. Let us try to formalize this idea. For this, we introduce an operator \mathbb{P}_{ij} with the job to exchange particle *i* and particle *j* (this is sometime called transposition). It is defined by

$$\mathbb{P}_{ij}\psi(r_1,\cdots,r_i,\cdots,r_j,\cdots)=\psi(r_1,\cdots,r_j,\cdots,r_i,\cdots)$$

or, equivalently,

$$\mathbb{P}_{ij}|\alpha_{m_1}\rangle\cdots|\alpha_{m_i}\rangle\cdots|\alpha_{m_j}\rangle\cdots|\alpha_{m_n}\rangle = |\alpha_{m_1}\rangle\cdots|\alpha_{m_j}\rangle\cdots|\alpha_{m_i}\rangle\cdots|\alpha_{m_n}\rangle \qquad (49)$$

Now, we demand that for **indistinguishable particles** $|\Psi\rangle$ and $\mathbb{P}_{ij}|\Psi\rangle$ have to describe the same state. From this, we conclude that

$$\mathbb{P}_{ij}|\Psi\rangle = \begin{array}{c} e^{i\phi} \\ \uparrow \\ \text{observable} \end{array} |\Psi\rangle \tag{50}$$

where the factor $e^{i\phi}$ has been introduced because we know that two states differing only by a global phase cannot be distinguished in any experiment.

The operator, which we defined above, has another important property: when we apply it twice, we obviously get back the same state

$$\mathbb{P}_{ij}^2 = \mathbb{1} \tag{51}$$

This harmless equation appears to be a trivial consequence of how we defined permutations. From a more fundamental point of view, we should, however, have asked a different question. What are mathematically consistent ways to define the exchange of two particles and how would we do this physically? This is a subtle and deep question (which we will not address here). It turns out that the answer to this question depends on the dimension of the system and the topology of space. In three dimension Eq. 51 has to hold, while in two dimensions also $\mathbb{P}_{ij}^2 \neq 1$ is allowed (this is related to the fact that a loop around a point can be contracted in 3D but not in 2D). We will ignore this subtlety and just assume that Eq. (51) is valid.

$$\mathbb{P}_{ij}^2 = \mathbb{1} \implies \mathbb{P}_{ij}^2 \psi = e^{2i\phi} \psi = \psi \implies e^{i\phi} = \pm 1$$

Thus, our mathematical analysis shows that there are two possibilities allowed mathematically, $e^{i\phi} = \pm 1$. We now have to go back to experiment and ask which of these options is realized by nature. The answer is that for some particles it is the first option while it is the second option for other particles. Thus, there are two types of particles

Bosons:
$$e^{i\phi} = 1 \quad \Leftrightarrow \quad \psi(r_1, r_2) = \psi(r_2, r_1) \quad \Leftrightarrow \quad \text{symmetric wave functions}$$

and

Fermions: $e^{i\phi} = -1 \quad \Leftrightarrow \quad \psi(r_1, r_2) = -\psi(r_2, r_1) \quad \Leftrightarrow \quad \text{anti-symmetric wave functions}$

As mentioned above, in two dimensions, further options are possible, which are called anyons (realized as excitations in certain quantum Hall systems). For Lorentz

invariant theories, there is a further constraint, called **spin statistic theorem**, which states that bosons always have integer spin $(0, 1, \dots)$, while fermions carry an half integer spin $(\frac{1}{2}, \frac{3}{2}, \dots)$. For excitations in a solid, there is no such constrait. Our next goal is to formally define the Hilbert space of n indistinguishable bosons or fermions. For this purpose, let us introduce S_n , the **set of all permutations** \mathbb{P} of n particles. A general permutation operator \mathbb{P} is obtained by the product of several two-particle exchange operators

$$\mathbb{P} = \mathbb{P}_{i_1 j_1} \mathbb{P}_{i_2 j_2} \cdots \mathbb{P}_{i_m j_m}$$
$$(-1)^{\mathbb{P}} \equiv (-1)^m = \begin{cases} (-1) & \text{odd number of exchanges} \\ 1 & \text{even number of exchanges} \end{cases}$$

We can use this, to define the Hilbert space of n bosons. We only allow for those wave functions, which are **symmetric** under the exchange of particles.

$$\mathscr{H}_{n}^{S} = \left\{ |\psi\rangle \in \mathscr{H}_{n}^{d} \, \big| \, \mathbb{P} |\psi\rangle = |\psi\rangle \, \forall \, \mathbb{P} \in S_{n} \right\}$$
(52)

Note that this Hilbert space is much smaller than the Hilbert space of distinguishable particles \mathscr{H}_n^d as we declared all non-symmetric wave functions to be not physical and thus not part of the Hilbert \mathscr{H}_n^S space.

Similarly, we define the Hilbert space of fermions, which is only allowed to contain anti-symmetrized wave functions

$$\mathscr{H}_{n}^{A} = \left\{ |\psi\rangle \in \mathscr{H}_{n}^{d} \, \big| \, \mathbb{P} |\psi\rangle = -|\psi\rangle \, \forall \, \mathbb{P} \in S_{n} \right\}$$
(53)

Two examples of allowed wave functions are

$$\psi(r_1, r_2) + \psi(r_2, r_1) \in \mathscr{H}_n^S$$

$$\psi(r_1, r_2) - \psi(r_2, r_1) \in \mathscr{H}_n^A$$

For fermions one can use the fact that determinants are completely anti symmetric to define so-called **Slater determinants**

$$\psi_{\alpha_1,\dots,\alpha_n}(r_1,\dots,r_n) = \frac{1}{\sqrt{n!}} \det \begin{bmatrix} \psi_{\alpha_1}(r_1) & \psi_{\alpha_1}(r_2) & \cdots & \psi_{\alpha_1}(r_n) \\ \psi_{\alpha_2}(r_1) & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \psi_{\alpha_n}(r_1) & \cdots & \cdots & \psi_{\alpha_n}(r_n) \end{bmatrix} \in \mathscr{H}_n^A$$
(54)

They can serve as basis states of \mathscr{H}_n^A . An example for n=2 is

$$\frac{1}{\sqrt{2}} \det \begin{bmatrix} \psi_{\alpha_1}(r_1) & \psi_{\alpha_1}(r_2) \\ \psi_{\alpha_2}(r_1) & \psi_{\alpha_2}(r_2) \end{bmatrix} = \frac{1}{\sqrt{2}} (\psi_{\alpha_1}(r_1)\psi_{\alpha_2}(r_2) - \psi_{\alpha_1}(r_2)\psi_{\alpha_2}(r_1)).$$

An important consequence of the anti-symmetrization is the **Pauli principle:** Two fermions cannot occupy the same state. This follows from

$$\mathbb{P}_{12}|\alpha\rangle|\alpha\rangle = |\alpha\rangle|\alpha\rangle \neq -|\alpha\rangle|\alpha\rangle \quad \Rightarrow \quad |\alpha\rangle|\alpha\rangle \notin \mathscr{H}^A.$$

2.2.2 Fock Space

Motivation: It is important to realize that we did something 'stupid' in the previous chapter. We obtained the physical Hilbert space of indistinguishable particles \mathscr{H}_n^A or \mathscr{H}_n^S by embedding it in an unphysical Hilbert space \mathscr{H}_n^d . We made the mistake to use objects like $\psi(r_1, r_2, \ldots r_n)$ which labels the particles with indices $1, 2, \ldots$ and then, we had to repair this mistake by symmetrization or anti-symmetrization. And we had to pay a high prize for that. Consider, for example the Slater determinant, Eq. (54). When one would try to write it out, it would contain n! terms. Already for 100 fermions, these are more terms than atoms in the universe!

Can we avoid working with $\psi(r_1, r_2, \ldots r_n)$ and thus avoid working with an unphysical Hilbert space \mathscr{H}_n^d ? Here the key insight is that there is one thing one can do with indistinguishable particles: one can count their number. Thus, our strategy will be

Indistinguishable particles: **just count** how frequently a given state is occupied

Before introducing formal definitions, let us make a simple example. Starting point is always the choice of a single-particle basis. For example, let us use the threedimensional single-particle Hilbert space \mathscr{H}_1 with basis vectors $\{|\alpha_1\rangle, |\alpha_2\rangle, |\alpha_3\rangle\}$. The associated single-particle wavefunctions are $\psi_{\alpha_1}(r), \psi_{\alpha_2}(r), \psi_{\alpha_3}(r)$.

Now, assume 2 bosons in state $|\alpha_1\rangle$, 2 in $|\alpha_3\rangle$. This information uniquely defines a wave function which we write as $|2, 0, 2\rangle$ (with an hopefully obvious notation). More precisely, this wavefunction is defined by

$$|2,0,2\rangle \propto \sum_{\mathbb{P}\in S_4} \mathbb{P}|\alpha_1\rangle |\alpha_1\rangle |\alpha_3\rangle |\alpha_3\rangle \propto \sum_{\mathbb{P}\in S_4} \mathbb{P}\psi_{\alpha_1}(r_1)\psi_{\alpha_1}(r_2)\psi_{\alpha_3}(r_3)\psi_{\alpha_3}(r_4)$$

(we are omitting the normalization factors here for simplicity).

We can do the same thing for fermions. Consider, for example, two fermions, with one fermion in the first state, and one in the third state. We write the wave function as $|1, 0, 1\rangle$ which is defined by

$$|1,0,1\rangle = \frac{1}{\sqrt{2}}(|\alpha_1\rangle|\alpha_3\rangle - |\alpha_3\rangle|\alpha_1\rangle) = \frac{1}{\sqrt{2}}(\psi_{\alpha_1}(r_1)\psi_{\alpha_3}(r_2) - \psi_{\alpha_3}(r_1)\psi_{\alpha_1}(r_2))$$

Let us generalize this properly. We start by defining a new Hilbert space, the **Fock Space**, which is the sum of the Hilbert spaces with $0, 1, \ldots$ particles

$$\mathscr{F} = \mathscr{F}_0 \oplus \mathscr{F}_1 \oplus \mathscr{F}_2 \oplus \cdots$$
 with $\mathscr{F}_n = \begin{cases} \mathscr{H}_n^S & \text{for Bosons} \\ \mathscr{H}_n^A & \text{for Fermions} \end{cases}$

Here \mathscr{F}_0 is the 1-dimensional Hilbert space with basis $|0\rangle = |0, 0, 0, \cdots \rangle$ which describes that there is no particle present.

To define a basis in the Fock space, we first need a **single-particle basis** $|\alpha_n\rangle$. Then, we simply count how many particles are in a single-particle state. Using the notation introduced above, our Fock-space basis vectors are given by

$$|n_1, n_2, \cdots \rangle \in \mathscr{F}$$
 with, $\begin{cases} n_i \in \mathbb{N}_0 & \text{for Bosons} \\ n_i \in \{0, 1\} & \text{for Fermions} \end{cases}$ (55)

and a general wave function can thus be written as

$$|\psi\rangle = \sum_{n_1, n_2, \dots} c_{n_1, n_2, \dots} |n_1, n_2, \dots \rangle$$

The advantage of the Fock space and its basis is, that we work from the beginning only with physical states and avoid the overhead of having to symmetrize or antisymmetrize wave functions. This is a major gain. There is also a useful side-effect: we can now describe **superpositions of state which have different numbers of particles**, which was not possible before. This is necessary for all processes where the particle number changes, like the emission or absorption of photons or phonons, or the annihilation of electrons and positrons. It will also be essential for a theory of superconductivity.

Note that **bound states** of an even number of fermions behave as bosons, while bound states of an odd number of fermions behave effectively as fermions. The reason is simple: when one exchanges n fermions, one obtains a factor $(-1)^n$. For example, protons and neutrons are bound states of three quarks and therefore fermions. A hydrogen atom, one proton and one electron, acts like a boson in situations where it can be approximated as a point-like particle. In contrast, its isotop, a heavy hydrogen atom with one proton, one neutron and one electron acts like a fermion.

2.2.3 Creation and Annihilation Operators

Motivation: After we have defined our Hilbert space, the next step is to consider operators acting on the Hilbert space. We started our introduction of the Fock space from the observation that the one thing which we can easily do with identical particles is to count them. Therefore, we will have to find out which operators can be used for counting. Furthermore, we will have to learn of how to change particle numbers, i.e., how to do +1 and -1. Thus, the goal of this section will be to find operators which just do that.

To count bosonic particles, we need an operator with spectrum \mathbb{N}_0 . We had precisely such an operator in the algebraic solution of the harmonic oscillator. Thus, we want to borrow the math *without* borrowing the physics. We have already discussed in Eq. (42) how creation and annihilation operators show up in this context. But in the following, we just need their algebra, and, importantly, the spectrum of the counting operator $\hat{n} = a^{\dagger}a$. The essential properties of the Harmonic oscillator operators were

$$[a, a^{\dagger}] = 1, \quad \hat{n} = a^{\dagger}a, \quad \hat{n}|n\rangle = n|n\rangle, n \in \mathbb{N}_{0}$$
$$a^{\dagger}|n\rangle = \sqrt{n+1}|n\rangle, \quad a|n\rangle = \sqrt{n}|n-1\rangle, \quad a|0\rangle = 0$$

This motivates us to **define** a new operator which acts on the Fock space. We can fully define an operator, by specifying how it acts on the basis vectors of a Hilbert space, which were given by $|n_1, n_2, ... \rangle$ with $n_i \in \mathbb{N}_0$ for bosons and $n_i \in \{0, 1\}$ for fermions.

Motivated by the harmonic oscillator example, we **define** creation operators a_i^{\dagger} and annihilation operators a_i by

$$a_{i}^{\dagger}|n_{1},\cdots,n_{i},\cdots\rangle = \sqrt{n_{i}+1}|n_{1},\cdots,n_{i}+1,\cdots\rangle$$
$$a_{i}|n_{1},\cdots,n_{i},\cdots\rangle = \sqrt{n_{i}}|n_{1},\cdots,n_{i}-1,\cdots\rangle$$
(56)

Note that there is no harmonic oscillator anywhere in the problem and what we are doing is completely independent of whatever the Hamiltonian will be. The operators a_i^{\dagger} and a_i remove bosonic particles in given single-particle states.

Can we do something similar for Fermions? In this case, we have to make sure that the operators know about the Pauli principle. We **define** creation and annihilation operators in this case by

$$a_{i}^{\dagger}|n_{1}, \cdots, 1, \cdots \rangle = 0 \text{ Pauli principle}$$

$$a_{i}^{\dagger}|n_{1}, \cdots, 0, \cdots \rangle = (-1)^{\sum_{j < i} n_{j}} |n_{1}, \cdots, 1, \cdots \rangle$$

$$a_{i}|n_{1}, \cdots, 0, \cdots \rangle = 0$$

$$a_{i}|n_{1}, \cdots, 1, \cdots \rangle = (-1)^{\sum_{j < i} n_{j}} |n_{1}, \cdots, 0, \cdots \rangle$$
(57)

Note that we included a strange-looking factor $(-1)^{\sum_{j < i} n_j}$ in the definition. It depends on how we ordered the single-particle states and appears to be non-local and complicated. The opposite is true. This factor in the definition will help us to encode the antisymmetrization of fermionic wave functions in an elegant way. This can, for example, be seen by showing that

$$a_j^{\dagger}a_i^{\dagger}|0\rangle = -a_i^{\dagger}a_j^{\dagger}|0\rangle.$$

The reader is encouraged to check how this relation follows from the definition and that it is valid independently of whether i < j, j < i or i = j.

Both for fermions and bosons our definitions were designed such that

$$a_i = (a_i^{\dagger})^{\dagger}$$

Also, in both cases, it follows from the definitions that

$$a_i|0\rangle = 0$$
, where $|0\rangle$ is the vacuum: $|0\rangle = |0, 0, 0, \cdots \rangle$
 $|n_1, n_2, \cdots \rangle = \prod_i \frac{1}{\sqrt{n_i!}} \left(a_i^{\dagger}\right)^{n_i} |0\rangle$

where in the case of fermions the ordering in the product plays a role. We order the terms such that the creation operator a_i^{\dagger} with i < j is always to the left of a_j^{\dagger} . Thus, we avoid complications with the factor $(-1)^{\sum_{j < i} n_j}$ in the definition.

Our next step is to work out the **algebra** of the operators starting from their definition. For the **bosonic case** we find directly

$$i \neq j: \left\{ \begin{array}{l} a_i^{\dagger} a_j^{\dagger} = a_j^{\dagger} a_i^{\dagger} \\ a_i^{\dagger} a_j = a_j a_i^{\dagger} \end{array} \right\} \equiv \text{ symmetric wavefunction}$$
$$i = j: \quad a_i a_i^{\dagger} - a_i^{\dagger} a_i | n_1, \cdots, n_i, \cdots \rangle$$
$$= \left(\left(\sqrt{n_i + 1} \right)^2 - \left(\sqrt{n_i} \right)^2 \right) | n_1, \cdots, n_i, \cdots \rangle$$
$$= | n_1, \cdots, n_i, \cdots \rangle$$

We can collect this insights into the **bosonic commutation relations**

$$\left[a_i, a_j^{\dagger}\right] = \delta_{ij}, \quad \left[a_i^{\dagger}, a_j^{\dagger}\right] = \left[a_i, a_j\right] = 0$$

Here [A, B] = AB - BA is the **commutator** between A and B. Now we proceed in the same way for **fermionic operators** and find

$$i \neq j : \left\{ \begin{array}{l} a_i^{\dagger} a_j^{\dagger} = -a_j^{\dagger} a_i^{\dagger} \\ a_i^{\dagger} a_j = -a_j a_i^{\dagger} \end{array} \right\} \quad \text{due to } (-1)^{\sum_{j < i} n_j} \\ \equiv \text{ antisymmetric wavefunction} \\ i \neq j : \left\{ \begin{array}{l} a_i a_i^{\dagger} | n_1, \cdots, n_i, \cdots \rangle \\ a_i^{\dagger} a_i | n_1, \cdots, n_i, \cdots \rangle \end{array} = (1 - n_i) | n_1, \cdots, n_i, \cdots \rangle \\ a_i^{\dagger} a_i | n_1, \cdots, n_i, \cdots \rangle \end{array} = n_i | n_1, \cdots, n_i, \cdots \rangle \right\}$$

Again, we can collect the result in a compact form, the **fermionic anticommuta**tion relations

$$\left\{a_i, a_i^{\dagger}\right\} = \delta_{ij}, \quad \left\{a_i^{\dagger}, a_j^{\dagger}\right\} = \left\{a_i, a_j\right\} = 0$$
(58)

where, $\{A, B\} = AB + BA$ is the **anticommutator** between A and B. Importantly, the 'ugly' term $(-1)^{\sum_{j < i} n_j}$ nowhere shows up in the algebra.

The anticommutation relations ensure automatically, that the wave function is completely antisymmetric, for example,

$$a_1^{\dagger}a_2^{\dagger}a_3^{\dagger}|0\rangle = -a_2^{\dagger}a_1^{\dagger}a_3^{\dagger}|0\rangle = a_2^{\dagger}a_3^{\dagger}a_1^{\dagger}|0\rangle$$

We have achieved this property without having to work, e.g., with Slater determinants (54)!

Above, we started from a definition of the Fock space, followed by a definition of creation and annihilation operators, which lead to the corresponding operator algebra. An alternative route is also possible: one can start from the operator algebra, postulate the existence of a unique vacuum state $|0\rangle$ and then used creation operators to build the state of the Hilbert space. The Stone-von Neumann theorem guarantees that (up to unitary transformations) this program gives a unique construction of our Hilbert space. Let us work out a few consequences of the operator algebra. From $\left\{a_i^{\dagger}, a_j^{\dagger}\right\} = 0$ follows for i = j immediately that $\left(a_i^{\dagger}\right)^2 = 0$. This is the Pauli principle, there can be only one fermion per state.

Both for fermions and bosons, it follows from our definitions (please check) that the operator

$$\hat{n}_i = a_i^{\dagger} a_i$$

counts the number of particles in state i,

$$\hat{n}_i | n_1, \cdots, n_i, \cdots \rangle = n_i | n_1, \cdots, n_i, \cdots \rangle.$$

Thus we find

$$\hat{n}_i a_i^{\dagger} | \cdots \rangle = a_i^{\dagger} (\hat{n}_i + 1) | \cdots \rangle$$

Both for bosons and fermions, we thus obtain the 'counting algebra'

$$[\hat{n}_i, a_i^{\dagger}] = a_i^{\dagger}, \qquad [\hat{n}_i, a_i] = -a_i$$

which encodes the physics that a_i^{\dagger} and a_i raise and lower the value of \hat{n}_i by one, respectively.

The reader is encouraged to check this relation explicitly for, e.g., the fermionic case using only the fundamental commutation relations, Eq. (58), and the definition of \hat{n}_i .

$$[a_i^{\dagger}a_i, a_i^{\dagger}] = a_i^{\dagger}a_ia_i^{\dagger} - a_i^{\dagger}a_i^{\dagger}a_i = a_i^{\dagger} - 2a_i^{\dagger}a_i^{\dagger}a_i = a_i^{\dagger}.$$

An important strategy to solve such problems is called **normal ordering**, which describes the procedure to use commutation and anticommutation relations to move all creation operators as far to the left as possible and all destruction operators as much as possible to the right. In the case shown above, one can then also use that $(a_i^{\dagger})^2 = 0$ for fermions as discussed above.

Above, we considered a system with either one species of bosons or one species of fermions. But we can easily extend this to a system with several types of bosons with creation operator $b_{i,\alpha}^{\dagger}$ and several types of fermions with creation operator $f_{i,\beta}^{\dagger}$. Here α and β enumerate different types of particles, e.g., for fermions this could be electron, proton, neutron for $\beta = 1, 2, 3$. In this case, the commutation relations are given by

$$\{f_{j,\beta'}, f_{i,\beta}^{\dagger}\} = \delta_{i,j}\delta_{\beta,\beta'}, \quad \{f_{j,\beta'}, f_{i,\beta}\} = \{f_{j,\beta'}^{\dagger}, f_{i,\beta}^{\dagger}\} = 0$$

$$[b_{j,\alpha'}, b_{i,\alpha}^{\dagger}] = \delta_{i,j}\delta_{\alpha,\alpha'}, \quad [b_{j,\alpha'}, b_{i,\alpha}] = [b_{j,\alpha'}^{\dagger}, b_{i,\alpha}^{\dagger}] = 0$$

$$[b_{j,\alpha}, f_{i,\beta}^{\dagger}] = [f_{j,\beta}, b_{i,\alpha}^{\dagger}] = [f_{j,\beta}, b_{i,\alpha}] = [f_{j,\beta}^{\dagger}, b_{i,\alpha}^{\dagger}] = 0.$$

In this section, we have defined the stage on which quantum field theory takes place: the Fock space. Importantly, we avoided the introduction of unphysical Hilbert spaces and reduced the definition of basis states to the counting of particles. We also learned about operators $\hat{n}_i = a_i^{\dagger} a_i$ which can be used for counting and learned how to do +1 and -1 with the help of creation and annihiliation operators.

2.3 Operators in Fock space

Motivation: Our next goal, will to express all observables and all other operators – especially, including the Hamilton operator – in terms of the creation and annihilation operators. We will find that this is not only mathematically elegant, but also simplifies the formulation of many-particle systems considerably. Moreover, it often allows for a straightforward and simplified interpretation of physical processes. Thus, a central goal of the section is not only to understand the math but also to build physical intuition.

2.3.1 Basis change

When defining creation operators a_{λ}^{\dagger} , we did this relative to a previously fixed singleparticle basis $|\lambda\rangle$ with

$$|\lambda\rangle = a_{\lambda}^{\dagger}|0\rangle$$

Thus, we first explore what happens when we change the basis.

Let us thus introduce a new basis $|\hat{\lambda}\rangle$ with

$$|\tilde{\lambda}\rangle = \sum_{\lambda} |\lambda\rangle \langle\lambda|\tilde{\lambda}\rangle$$

with the unitary matrix $U_{\tilde{\lambda}\lambda} = \langle \lambda | \tilde{\lambda} \rangle$

We now want to introduce new operators $\tilde{a}_{\tilde{\lambda}}^{\dagger}$ in such a way, that they create particles in the state $|\tilde{\lambda}\rangle$. Thus, we have to demand that

$$|\tilde{\lambda}
angle = \tilde{a}^{\dagger}_{\tilde{\lambda}}|0
angle = \sum_{\lambda} U_{\tilde{\lambda}\lambda} a^{\dagger}_{\lambda}|0
angle$$

Therefore, we chose the new operators as

$$\tilde{a}_{\tilde{\lambda}}^{\dagger} = \sum_{\lambda} U_{\tilde{\lambda}\lambda} a_{\lambda}^{\dagger} = \sum_{\lambda} \langle \lambda | \tilde{\lambda} \rangle a_{\lambda}^{\dagger}$$

$$\tilde{a}_{\tilde{\lambda}} = \left(\tilde{a}_{\tilde{\lambda}}^{\dagger} \right)^{\dagger} = \sum_{\lambda} \langle \tilde{\lambda} | \lambda \rangle a_{\lambda}$$
(59)

We should now check, that the new operators obey the same (anti-) commutation relations as the old operators. Let us do this check for fermions by computing

$$\left\{\tilde{a}_{\tilde{\lambda}}^{\dagger},\tilde{a}_{\tilde{\lambda}'}\right\} = \sum_{\lambda\lambda'} \langle\lambda|\tilde{\lambda}\rangle\langle\tilde{\lambda}'|\lambda'\rangle \left\{a_{\lambda}^{\dagger},a_{\lambda'}\right\} = \sum_{\lambda} \langle\tilde{\lambda}'|\lambda\rangle\langle\lambda|\tilde{\lambda}\rangle = \langle\tilde{\lambda}'|\tilde{\lambda}\rangle = \delta_{\tilde{\lambda}'\tilde{\lambda}}$$
(60)

Indeed, using that both $|\lambda\rangle$ and $|\tilde{\lambda}\rangle$ form a complete normalized basis of our Hilbert space, we find that our new operators obey the correct algebra.

We will use the basis change frequently in the following sections.

2.3.2 Field operators in Real and Momentum space

A very useful and frequently used basis are position eigenstates which allows to find directly a description in (real) space. Position eigenstates are δ functions, $\delta(\vec{r} - \vec{x})$. We use the following notations and relations

$$|\vec{x}\rangle \equiv \delta(\vec{r} - \vec{x})$$

$$\langle \vec{x} | \phi \rangle = \int d^3 r \, \delta^3(\vec{r} - \vec{x}) \phi(\vec{r}) = \phi(\vec{x})$$

$$\langle \vec{x} | \vec{x}' \rangle = \int d^3 r \, \delta^3(\vec{r} - \vec{x}) \delta^3(\vec{r} - \vec{x}') = \delta^3(\vec{x} - \vec{x}')$$
(61)

Note, that the eigenstates are **not normalized** to one, $\langle \vec{x} | \vec{x} \rangle = \delta(0) = \infty$, thus we will have to adapt slightly the formulas of the previous section.

Our central goal will be to define a **field operator** $\psi^{\dagger}(\vec{x})$ which creates a particle in a position-eigenstate localized at position \vec{x}

$$|\psi^{\dagger}(\vec{x})|0\rangle = |\vec{x}\rangle$$

Using Eq. (59) and a set of creation operators a_{λ}^{\dagger} which creates particles in the single-particle wavefunction $\Phi_{\lambda}(\vec{x}) = \langle \vec{x} | \lambda \rangle$, we define

$$\psi^{\dagger}(\vec{x}) = \sum_{\lambda} \langle \lambda | \vec{x} \rangle \, a_{\lambda}^{\dagger}, \text{ where } a_{\lambda}^{\dagger} = \int d^3x \, \phi(\vec{x}) \psi^{\dagger}(\vec{x})$$

This leads to

$$\int d^3x \, \phi_{\lambda}(\vec{x}) \psi^{\dagger}(\vec{x}) = \sum_{\lambda'} \int d^3x \langle \lambda' | \vec{x} \rangle \langle \vec{x} | \lambda \rangle a^{\dagger}_{\lambda'} = \sum_{\lambda'} \int d^3x \langle \lambda' | \lambda \rangle a^{\dagger}_{\lambda'} = a^{\dagger}_{\lambda}$$

Let us try to interpret this formal result

$$a_{\lambda}^{\dagger} = \int d^3x \,\phi_{\lambda}(\vec{x})\psi^{\dagger}(\vec{x}) \tag{62}$$

How to we create a particle with the wavefunction $\Phi_{\lambda}(\vec{x})$? We use the operator $\Psi^{\dagger}(\vec{x})$ which creates a particle at position \vec{x} and build a superposition weighted by $\Phi_{\lambda}(\vec{x})$.

Here, it is important to emphasize that $\Psi(x)$ has **nothing to do with a wavefunction**. The use of the letter Ψ is, perhaps, unfortunate (but common). $\Psi(\vec{x})$ is instead an operator which destroys a particle at position \vec{x} . Also \vec{x} is (other than in single-particle quantum mechanics) **not an operator**, it is just a label which tells at which point in space the particle is removed. A similar statement applies to the field operator $\Phi(\vec{x})$ which we introduced to describe the harmonic chain, see Eq. (48). Also this is an operator, acting on position \vec{x} is just a certain operator acting on the system at position \vec{x} . The same is true for the electric and magnetic fields, $\vec{E}(\vec{x})$ and $\vec{B}(\vec{x})$ which are promoted to field operators in quantum field theory. Operators like $\Psi(\vec{x})$, $\Phi(\vec{x})$, $\vec{E}(\vec{x})$, or $\vec{B}(\vec{x})$ are called **quantum fields** acting in real space. By repeating the calculation of Eq. (60), we find that the **(anti-) commutators** of our field operators are given by

Fermions:
$$\{\psi(\vec{x}), \psi^{\dagger}(\vec{x}')\} = \langle \vec{x} | \vec{x}' \rangle = \delta^3(\vec{x} - \vec{x}')$$

Boson: $[\psi(\vec{x}), \psi^{\dagger}(\vec{x}')] = \delta^3(\vec{x} - \vec{x}')$

Let us do a few examples with the goal to practice the use of our quantum field. As discussed above, a single particle wavefunction $\phi(\vec{x})$ is obtained from

$$|\phi\rangle = \int d^3x \,\phi(\vec{x})\psi^{\dagger}(\vec{x})|0\rangle$$

Now let us compute the amplitude that the particle is located at position \vec{x}_0 by calculating the overlap with $|\vec{x}_0\rangle = \psi^{\dagger}(\vec{x}_0)|0\rangle$. We assume that we consider fermions here but the result has to be the same for bosons (as we consider a single particle only)

$$\begin{aligned} \langle \vec{x}_0 | \phi \rangle &= \langle 0 | \psi(\vec{x}_0) \int d^3 x \, \phi(\vec{x}) \psi^{\dagger}(\vec{x}) | 0 \rangle \\ &= \langle 0 | \int d^3 x \, \phi(\vec{x}) \delta^3(\vec{x} - \vec{x}_0) | 0 \rangle - \langle 0 | \int d^3 x \, \phi(\vec{x}) \Psi^{\dagger}(\vec{x}) \Psi(\vec{x}_0) | 0 \rangle \\ &= \phi(\vec{x}_0) \, \langle 0 | 0 \rangle = \phi(\vec{x}_0) \end{aligned}$$

To obtain the second line, we used $\{\Psi(\vec{x}_0), \Psi^{\dagger}(\vec{x})\} = \delta^3(\vec{x} - \vec{x}_0)$ or, equivalently, $\Psi(\vec{x}_0)\Psi^{\dagger}(\vec{x}) = \delta^3(\vec{x} - \vec{x}_0) - \Psi^{\dagger}(\vec{x})\Psi(\vec{x}_0)$ for **normal ordering**, which means for putting the destruction operators to the right and creation operators to the left with the help of (anti-) commutation relations. To get the third line, we can the simply use the definition of the vacuum, $\Psi(\vec{x}_0)|0\rangle = 0$.

Our next goal is to discuss another very important single particle basis: **momentum space**. We need operators, which create particles in momentum eigenstate. Single particle eigenstates of the momentum operator have the form $e^{i\vec{k}\cdot\vec{x}}$ but we need a convention for their normalization. This is done differently by different textbooks and different publication. We will use the following convention: we consider a box of length L in d dimensions with periodic boundary conditions, which enforces that momenta \vec{k} (written here for d = 3) take discrete values

$$\vec{k}_n = \begin{pmatrix} k_{n_1} \\ k_{n_2} \\ k_{n_3} \end{pmatrix}, \quad k_{n_i} = n \frac{2\pi}{L} = n_i \Delta k, \ n_i \in \mathbb{Z}, i = 1, \dots, d.$$

In this finite box, we simply normalize the eigenstates of the momentum operator to 1

$$\langle \vec{k}_n | \vec{k}_m \rangle = \delta_{\vec{k}_m, \vec{k}_n}, \qquad \langle \vec{x} | \vec{k} \rangle = \frac{1}{\sqrt{V}} e^{i \vec{k} \cdot \vec{x}}$$

Thus, using the results of the previous section, we the operater creating a particle in this eigenstate is given by

$$c_{\vec{k}}^{\dagger} = \int \frac{d^3x}{\sqrt{V}} e^{i\vec{k}\cdot\vec{x}}\psi^{\dagger}(\vec{x}), \qquad \psi^{\dagger}(\vec{x}) = \sum_{k} \frac{1}{\sqrt{V}} e^{-i\vec{k}\cdot\vec{x}}c_{\vec{k}}^{\dagger} \tag{63}$$

and the resulting (anti-) commutation relations are

for Fermions:
$$\begin{cases} c_{\vec{k}}, c_{\vec{k}'}^{\dagger} \\ \\ \text{for Bosons:} \end{cases} = \langle \vec{k} | \vec{k}' \rangle = \delta_{\vec{k}, \vec{k}'} \\ c_{\vec{k}}, c_{\vec{k}'}^{\dagger} \end{bmatrix} = \delta_{\vec{k}, \vec{k}'}$$
(64)

In practice, we will most often use the thermodynamics limit $V \to \infty$ resulting in $\Delta k = \frac{2\pi}{L} \to 0$. In this limit, sums over momenta are written as integrals.

$$\sum_{k_x} \dots = \frac{1}{\Delta k_x} \sum_{k_x} \Delta k_x \dots \xrightarrow{V \to \infty} \frac{L}{2\pi} \int dk_x \dots$$
(65)

or in d = 3, we get for $V \to \infty$

$$\sum_{\vec{k}} \cdots \longrightarrow V = \frac{d^3k}{(2\pi)^3} \dots$$

For example, the total number of particles is written for $V \to \infty$

$$N = \sum_{\vec{k}} c_{\vec{k}}^{\dagger} c_{\vec{k}} = V \int \frac{d^3k}{(2\pi)^3} c_{\vec{k}}^{\dagger} c_{\vec{k}} = \int d^3r \,\phi(\vec{r})^{\dagger} \phi(\vec{r}) \tag{66}$$

In a finite volume, $\frac{1}{V} \int d^3 \vec{x} e^{i(\vec{k}_n - \vec{k})_m} = \delta_{\vec{k}_n, \vec{k}_m}$, while for $V \to \infty$

$$\int d^3\vec{x} \ e^{i\vec{k}\vec{x}} = (2\pi)^3 \delta^3(\vec{k})$$

which follows from the rules for Fourier transformations. Thus, one can write (within our definitions) for, e.g., the anti commutator of Fermions

$$\left\{c_{\vec{k}}, c^{\dagger}_{\vec{k}'}\right\} = \frac{(2\pi)^3}{V} \delta^3(\vec{k} - \vec{k}') \text{ for } V \to \infty$$

$$\tag{67}$$

This is equivalent to Eq. (64) but usually it is more convenient to use Eq. (64) only. The formula above is mainly used to match to definitions used by other textbooks or publications.

As a result of this section, we now know how to create quantum particles either in position or momentum space.

2.3.3 Single particle operators

Our next goal is to write all types of operators using creation and annihilation operators. Consider, for example, the Hamilton operator describing N interacting

electrons, which we can split into two parts

$$\mathcal{H} = \underbrace{\sum_{i=1}^{N} \frac{P_i^2}{2m} + V(r_i)}_{\text{single particle operators}} + \underbrace{\frac{1}{2} \sum_{r_i \neq r_j} V(r_i - r_j)}_{\text{two-particle operators}}$$

We will first consider single particle operators, which are written as a sum over particle index i. Within the N-particle Hilbert space, we write such an operator (for either bosons or fermions) as

$$A^{(N)} = \sum_{i=1}^{N} A_i$$

First, we want to promote this to an operator acting on Fock space

$$A = A^{(0)} \oplus A^{(1)} \oplus A^{(2)} \oplus \cdots$$

To make progress, we introduce the eigenbasis $\{|\lambda_i\rangle\}$ of $A^{(1)}$. In this basis, the single-particle operator $A^{(1)}$ is diagonal

$$A^{(1)} = \sum_{i} \lambda_{i} |\lambda_{i}\rangle \langle\lambda_{i}|$$

For this specific basis, we introduce Fock-space eigenstates $|n_{\lambda_1}, n_{\lambda_2}, \cdots \rangle$ where the $|n_{\lambda_i}\rangle$ specify how many fermions or bosons occupy the state with quantum number λ_i . For this basis, we know exactly how A acts

$$A|n_{\lambda_1}, n_{\lambda_2}, \cdots \rangle = \sum_i \lambda_i n_{\lambda_i} |n_{\lambda_1}, n_{\lambda_2}, \cdots \rangle = \sum_i \lambda_i a^{\dagger}_{\lambda_i} a_{\lambda_i} |n_{\lambda_1}, n_{\lambda_2}, \cdots \rangle$$

Therefore, in this basis we find

$$A = \sum_{i} \lambda_{i} a_{\lambda_{i}}^{\dagger} a_{\lambda_{i}} = \sum_{ij} \underbrace{\langle \lambda_{i} | A^{(1)} | \lambda_{j} \rangle}_{\delta_{\lambda_{i}} \lambda_{j}} a_{\lambda_{i}}^{\dagger} a_{\lambda_{j}}, \tag{68}$$

where we will soon see, why we use in the second equation a more complicated formula. For this specific basis, we now know how A looks in second quantization. Thus, our final task is to use the results from the previous section to choose any other basis. Let us assume, that we want to introduce creation operators $a_{\alpha_i}^{\dagger}$ which create particle in the single-particle states $|\alpha_i\rangle$. Using the results of the previous section, the previous operators can be expressed by the new operators

$$a_{\lambda_i}^{\dagger} = \sum_k \langle \alpha_k | \lambda_i \rangle a_{\alpha_k}^{\dagger}, \qquad a_{\lambda_j} = \sum_{k'} \langle \lambda_j | \alpha_{k'} \rangle a_{\alpha_{k'}}.$$

Using this in Eq. (68), we obtain

$$\implies A = \sum_{i,j,k,k'} \langle \alpha_k | \lambda_i \rangle \langle \lambda_i | A^{(1)} | \lambda_j \rangle \langle \lambda_j | \alpha_{k'} \rangle a^{\dagger}_{\alpha_k} a_{\alpha_{k'}}$$

and we arrive at our final result

$$A = \sum_{k,k'} \langle \alpha_k | A^{(1)} | \alpha_{k'} \rangle a^{\dagger}_{\alpha_k} a_{\alpha_{k'}}$$
(69)

It is useful to practice how to interpret such terms. To this end, let us write the single-particle operator in this basis

$$A^{(1)} = \sum_{k,k'} |\alpha_k\rangle \langle \alpha_k | A^{(1)} | \alpha_{k'}\rangle \langle \alpha_{k'} |$$
(70)

The physical interpretation of the operator is that it transfers a quantum particle from state $|\alpha_{k'}\rangle$ to the state $|\alpha_k\rangle$ with the amplitude $\langle \alpha_k | A^{(1)} | \alpha_{k'} \rangle$. We can literally use the same interpretation for our many-particle operator A in Eq. (69). $a_{\alpha_{k'}}$ removes a particle in state $|\alpha_{k'}\rangle$, while $a^{\dagger}_{\alpha_k}$ creates one in $|\alpha_k\rangle$ and this happens with the transition matrix element $\langle \alpha_k | A^{(1)} | \alpha_{k'} \rangle$.

Let us consider a few examples. First, the total momentum of an N particle is given (in first quantization) by $\vec{P} = \sum_{i=1}^{N} -i\hbar \frac{\partial}{\partial \vec{r_i}}$. In momentum and real-space basis, the relevant single-particle matrix elements are given by

$$\langle \vec{k}' | -i\hbar \vec{\nabla} | \vec{k} \rangle = \hbar \vec{k} \delta_{\vec{k},\vec{k}'}$$
$$\langle \vec{x}_0 | -i\hbar \vec{\nabla} | \vec{x}'_0 \rangle = \int d\vec{r} \,\delta(\vec{r} - \vec{x}_0) \left(-i\hbar \frac{\partial}{\partial \vec{r}} \right) \delta(\vec{r} - \vec{x}'_0) = -i\hbar \vec{\nabla} \delta(\vec{x}_0 - \vec{x}'_0)$$

thus, we obtain in second quantization

$$\vec{P} = \sum_{k} \hbar \vec{k} c_{\vec{k}}^{\dagger} c_{\vec{k}}$$
$$= \int dx_0 dx'_0 (-i\hbar) \delta'(x_0 - x'_0) \psi^{\dagger}(x_0) \psi(x'_0) = \int dx \, \psi^{\dagger}(x) \left(-i\hbar \frac{\partial}{\partial x}\right) \psi(x)$$

Similarly, the kinetic energy $T = \sum_{i=1}^{N} \frac{-\hbar^2}{2m} \vec{\nabla}_i^2$ and the potential energy $U = \sum_{i=1}^{N} U(r_i)$ are written as

$$T = \sum_{\vec{k}} \frac{\hbar^2 k^2}{2m} c_{\vec{k}}^{\dagger} c_{\vec{k}} = \int d^3 x \, \psi^{\dagger}(x) \left(\frac{-\hbar^2}{2m} \vec{\nabla}^2\right) \psi(x) = \int d^3 x \, \frac{\hbar^2}{2m} (\vec{\nabla} \psi^{\dagger}(x)) (\vec{\nabla} \psi(x))$$
$$U = \int d^3 x \, U(\vec{x}) \psi^{\dagger}(\vec{x}) \psi(\vec{x}).$$

We would like to emphasize that the description in terms of second quantization is much more elegant and efficient than in first quantization. To describe, e.g., the 10^{23} of a solid, one has to introduce in 1st quantization 10^{23} coordinates $\vec{r_i}$, by labeling each electron with a label *i* which is unphysical as they are indistinguishable.

One way to see that these are really unphysical coordinates consider the operator \vec{r}_1^x , the x-coordinate of the first particle. As electrons are indistinguishable, this

operator has no physical meaning. Mathematically, r_1^x is also not a legal operator as it does not map elements of the Hilbertspace \mathcal{H}_N^A of totally antisymmetric wave functions onto itsself, $r_1^x |\Psi\rangle \notin \mathcal{H}_N^A$.

In contrast, in second quantization, one never introduces unphysical operators. Quantities like the kinetic and potential energy are written as simple integrals or sums over creation and annihiliation operators.

2.3.4 Two-particle operators

As a last step, we have to translate two-particle operators into the language of second quantization. Here the prime example is the potential energy, which in first quantization is written as

$$V = \sum_{i < j} V(r_i, r_j) = \frac{1}{2} \sum_{i \neq j} V(r_i, r_j)$$

For a given basis $|\lambda_i\rangle$, one first calculates matrix elements in the 2-particle Hilbert space

$$\langle \lambda_1 | \langle \lambda_2 | V^{(2)} | \lambda_{2'} \rangle | \lambda_{1'} \rangle = \int d^3 r_1 d^3 r_2 \, \psi^*_{\lambda_1}(r_1) \psi^*_{\lambda_2}(r_2) V(r_1, r_2) \psi_{\lambda_{2'}}(r_2) \psi_{\lambda_{1'}}(r_1).$$

Using these matrix elements one can write (we do this here without a detailed derivation) the operator in second quantization in analogy to Eq. 69

$$V = \frac{1}{2} \sum_{\lambda_1 \lambda_2 \lambda_{1'} \lambda_{2'}} \langle \lambda_1, \lambda_2 | V^{(2)} | \lambda_{1'}, \lambda_{2'} \rangle a^{\dagger}_{\lambda_1} a^{\dagger}_{\lambda_2} a_{\lambda_{2'}} a_{\lambda_{1'}}$$
(71)

both for bosons and fermions. Note that the order of operators is here important as $a_{\lambda_{2'}}a_{\lambda_{1'}} = -a_{\lambda_{1'}}a_{\lambda_{2'}}$ for fermions.

The electron-electron interaction $V^{\text{int}} = \frac{1}{2} \sum_{i \neq j} V^{\text{int}}(\vec{r}_i - \vec{r}_j)$ with $V^{\text{int}}(\vec{r}) = \frac{e^2}{4\pi\epsilon_0 |\vec{r}|}$ is written in second quantization in the real-space basis as

$$V^{\text{int}} = \frac{1}{2} \int d^3x \, d^3x' \, V^{\text{int}}(\vec{x} - \vec{x}') \psi^{\dagger}(\vec{x}) \psi^{\dagger}(\vec{x}') \psi(\vec{x}') \psi(\vec{x}')$$

Let us derive more carefully how the same operator is written in momentum basis, as we will need that later

$$\begin{aligned} \langle \vec{k}_1 \vec{k}_2 | V^{\text{int}}(1-2) | \vec{k}_1' \vec{k}_2' \rangle &= \frac{1}{V^2} \int d^3 r_1 \, d^3 r_2 \, e^{i(\vec{k}_1' - \vec{k}_1) \vec{r}_1} e^{i(\vec{k}_2' - \vec{k}_2) \vec{r}_2} V^{\text{int}}(r_1 - r_2) \\ \text{using } \vec{r}_1 &= \vec{R} + \frac{\vec{r}}{2}, \ \vec{r}_2 &= \vec{R} - \frac{\vec{r}}{2} \text{ and } \begin{vmatrix} 1 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{vmatrix} = 1 \\ &= \frac{1}{\mathcal{V}^2} \int d^3 r \, d^3 R \, e^{i(\vec{k}_1' + \vec{k}_2' - (\vec{k}_1 + \vec{k}_2)) \vec{R}} V^{\text{int}}(\vec{r}) e^{i(\vec{k}_1' - \vec{k}_1 - (\vec{k}_2' - \vec{k}_2)) \vec{r}} \\ &= \frac{1}{\mathcal{V}} \delta_{\vec{k}_1 + \vec{k}_2, \vec{k}_1' + \vec{k}_2'} V^{\text{int}}(\vec{k}_1' - \vec{k}_1) \end{aligned}$$

where we used the Fourier transformation of the potential, $V^{\text{int}}(\vec{k}) = \int d^3x V^{\text{int}}(\vec{x}) e^{i\vec{k}\vec{x}}$. Thus we obtain

$$V^{\text{int}} = \frac{1}{2\mathcal{V}} \sum_{\vec{k}_1, \vec{k}_2, \vec{k}_1', \vec{k}_2'} V^{\text{int}} (\vec{k}_1' - \vec{k}_1) \delta_{\vec{k}_1 + \vec{k}_2, \vec{k}_1' + \vec{k}_2'} c_{\vec{k}_1}^{\dagger} c_{\vec{k}_2}^{\dagger} c_{\vec{k}_2'} c_{\vec{k}_1'}$$
$$= \frac{1}{2\mathcal{V}} \sum_{\vec{k}_1, \vec{k}_2, \vec{q}} V^{\text{int}} (\vec{q}) c_{\vec{k}_1 + \vec{q}}^{\dagger} c_{\vec{k}_2 - \vec{q}}^{\dagger} c_{\vec{k}_2} c_{\vec{k}_1}$$

We will later heavily use simple pictures to display such formulas. As a first example, we can use the following picture for the formula



Here incoming lines denote destruction operators, outgoing lines creation operators. The wiggly line stands for the interaction potential. Each line is labeled by its momentum. Note that at each vertex the sum of incoming momenta is equal to the sum of outgoing momenta.

We have now reached a major milestone: we know how to write all relevant physical operators in terms of second quantization.

2.3.5 Hamiltonian with spin

Up to now, we have omitted one important detail: electrons carry an extra quantum number, the spin. But also for other quantum numbers (e.g., the color index of a quark or the band index of an electron in a solid, discussed later), one may ask how they can be added to the formalism of second quantization. That turns out to be very easy: one just has to add an extra index to all creation operators. For example, in the case of spin we write $\sigma = \uparrow / \downarrow = \pm \frac{1}{2}$ and introduce corresponding field operators

 $\psi^{\dagger}_{\sigma}(\vec{r}) \quad {\rm or} \quad c^{\dagger}_{\vec{k},\sigma} \quad {\rm creates \ particle \ with \ spin \ } \sigma = \uparrow / \downarrow$

Let us practice using that operator. In first quantization, the wave function of an electron with spin is written as a two-component spinor $\vec{\varphi}(\vec{r}) = \begin{pmatrix} \varphi_{\uparrow}(\vec{r}) \\ \varphi_{\downarrow}(\vec{r}) \end{pmatrix}$. Here $\varphi_{\sigma}(\vec{r}) \in \mathbb{C}$ is the quantum mechanical amplitude that the particle is at position \vec{r} and has the spin σ . In second quantization the same single-particle wave function
is written as

$$|\varphi\rangle = \int d^3x \,\varphi_{\uparrow}(\vec{x})\psi^{\dagger}_{\uparrow}(\vec{x}) + \varphi_{\downarrow}(\vec{x})\psi^{\dagger}_{\downarrow}(\vec{x})|0\rangle$$

Our next goal is to write the Hamiltonian of interacting electrons in presence of magnetic field, $\vec{s}_i =$ spin of particle *i*, now taking into account the presence of a magnetic field

$$\mathcal{H}^{(N)} = \sum_{i=1}^{N} \frac{\left(\vec{p_i} - e\vec{A}(\vec{r_i})\right)^2}{2m} + U(\vec{r_i}) + \frac{1}{2} \sum_{\vec{r_i} \neq \vec{r_j}} V(\vec{r_i} - \vec{r_j}) - g\mu_B \vec{B}(\vec{r_i})\vec{s_i}$$

Here the magnetic field enters twice, both via the vector potential, $\nabla \times \vec{A}(\vec{r}) = \vec{B}(\vec{r})$ and also by the Zeeman coupling term of magnetic field with the spin.

We can directly write this Hamiltonian in it's second-quantized form using the field operators $\psi^{\dagger}(\vec{r})$ and $\psi(\vec{r})$

$$\mathcal{H} = \sum_{\sigma=\uparrow/\downarrow} \int d^3x \, \psi^{\dagger}_{\sigma}(\vec{x}) \frac{\left(-i\hbar\vec{\nabla} - e\vec{A}(\vec{x})\right)^2}{2m} \psi_{\sigma}(\vec{x}) + U(\vec{x})\psi^{\dagger}_{\sigma}(\vec{x})\psi_{\sigma}(\vec{x})$$
$$+ \frac{1}{2} \sum_{\sigma,\sigma'} \int d^3x \, d^3x' \, V(\vec{x} - \vec{x}')\psi^{\dagger}_{\sigma}(\vec{x})\psi^{\dagger}_{\sigma'}(\vec{x}')\psi_{\sigma'}(\vec{x}')\psi_{\sigma}(\vec{x})$$
$$- g\mu_b \int d^3x \, \vec{B}(\vec{x}) \sum_{\alpha,\beta=\uparrow/\downarrow} \psi^{\dagger}_{\alpha}(\vec{x}) \frac{\vec{\sigma}_{\alpha\beta}}{2} \psi_{\beta}(\vec{x})$$

Describes almost all properties in solids! Compared to the previously derived formulas, we simply added the spin index to the field. The local charge density is, for example, now given by the operator

$$\rho(\vec{x}) = \sum_{\sigma=\uparrow/\downarrow} \psi^{\dagger}_{\sigma}(\vec{x}) \psi_{\sigma}(\vec{x}), \tag{72}$$

which enters the potential term $U(\vec{x})$ and also the Coulomb interactions. To obtain the last term, one can simply use our master-formula Eq. (69). Here one can use that the single-particle matrix elements for the single-particle wave function $\vec{\varphi}(\vec{r}) = \begin{pmatrix} \varphi_{\uparrow}(\vec{r}) \\ \varphi_{\downarrow}(\vec{r}) \end{pmatrix}$ are given by $\int \vec{\varphi}(\vec{r})^{\dagger} (\vec{B}(\vec{r}) \cdot \frac{\vec{\sigma}}{2}) \vec{\varphi}(\vec{r})$ and as basis states one uses $\delta(\vec{r} - \vec{r}_0)(1, 0)^T$ and $\delta(\vec{r} - \vec{r}_0)(0, 1)^T$ for the two spin components.

For later use, we also give the formula for the Hamiltonian in momentum space for a special case, namely $U(\vec{r}) = 0, \vec{A} = 0, \vec{B} = 0$. In this case, one obtains

$$\mathcal{H} = \sum_{\sigma=\uparrow/\downarrow} \sum_{\vec{k}} \varepsilon(\vec{k}) c^{\dagger}_{\vec{k},\sigma} c_{\vec{k},\sigma} + \frac{1}{2V} \sum_{\sigma\sigma'\vec{k}\vec{k'}\vec{q}} V(\vec{q}) c^{\dagger}_{\vec{k}+\vec{q},\sigma} c^{\dagger}_{\vec{k'}-\vec{q},\sigma'} c_{\vec{k'},\sigma'} c_{\vec{k}\sigma}$$

Graphically, the interaction term takes the following form



Note that the lines entering the left and right vertex carry the spin index σ and σ' , respectively. In the Hamiltonian one sums over those indices. The physics reason for that is that the Coulomb interaction couples to the total charge density $\rho(\vec{x})$.

2.3.6 Dynamics of quantum fields

Finally, it is an instructive exercise to write the Hamiltonian equations of motion for field operators.

Let us start with a reminder that one can use in quantum mechanics equivalently use the Schrödinger picture and the Heisenberg picture. In the Schrödinger picture the time-evolution arises from Schrödinger's equation $i\hbar\partial_t |\psi\rangle = \mathcal{H}|\psi\rangle$.

Here it is useful to introduce the time-evolution operator U_t defined by $|\psi(t)\rangle = U_t |\psi(0)\rangle$ with

$$i\hbar\partial_t U_t = \mathcal{H}U_t$$

solved by $U_t = e^{-i\mathcal{H}t/\hbar}$ for a static Hamiltonian. The expectation value of an operator is then written as

$$\langle A \rangle = \langle \psi(t) | A | \psi(t) \rangle = \langle \psi(0) | U_t^{\dagger} A U_t | \psi(0) \rangle$$

When using the Heisenberg picture, one re-interprets the last equation: while one uses a static wavefunction $|\psi(0)\rangle$, the operators obtain a time dependence given by $A^{\mathcal{H}}(t) = U_t^{\dagger} A U_t$. From this definition, one obtains the well-known Heisenberg equations of motion (we assume that A is time-independent in the Schrödinger picture)

$$i\hbar\partial_t A^{\mathcal{H}}(t) = \left[A^{\mathcal{H}}(t), \mathcal{H}\right]$$

The goal of this section is to obtain the Heisenberg equations of motion for the field operator. As an example, we use the Hamiltonian

$$\mathcal{H} = \int d^3x \, \left(-\frac{\hbar^2}{2m} \right) \psi^{\dagger}(\vec{x}) \nabla^2 \psi(\vec{x}) + \mathcal{U}(\vec{x}) \psi^{\dagger}(\vec{x}) \psi(\vec{x}) + \frac{1}{2} \int d^3x \, d^3x' \, V(\vec{x} - \vec{x}') \psi^{\dagger}(\vec{x}) \psi^{\dagger}(\vec{x}') \psi(\vec{x}') \psi(\vec{x})$$

Using the commutation or anticommutation relations for field operators, we find after some algebra both for fermions and bosons the Heisenberg equations of motion

$$\begin{split} i\hbar\partial_t\psi(\vec{r},t) &= [\psi(\vec{r},t),\mathcal{H}] \\ &= -\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r},t) + \mathcal{U}(\vec{r})\psi(\vec{r},t) + \int d^3x'\,\psi^{\dagger}(\vec{x}',t)\psi(\vec{x}',t)V(\vec{x}'-\vec{r})\psi(\vec{r},t) \end{split}$$

Outlook: These equations have a remarkable similarity with the Schrödinger equation but it is important to realized that this is an equation for a field operator and not for a single field. Thus, due to the product of three operators in the last term, this is an equation which is in practice impossible to be solved directly. We will later learn, that for Bose-Einstein condensates one can approximately replace the operator $\psi(t)$ by its expectation value $\Phi(t) = \langle \psi(t) \rangle$. The resulting non-linear Schrödinger equation (or Gross-Pitaevskii equation) can be solved numerically and is useful to describe experiments with ultracold bosonic atoms.

2.4 Electrons in solids

Motivation: In this course we will use mainly examples from solid state theory when applying methods and concepts of quantum field theory. Therefore, we we will need throughout the course basic notions of solid state theory. In the following three sections, we will develop some central concepts. At the same time, we use these sections also to practice the use of creation and annihilation operators. A major goal will be to obtain an intuitive understanding for them.

2.4.1 Defining solids

We start with a number of useful definitions. A solid is build from an array of periodically arranged atoms. To describe such periodic structures, we use a **Bravais** lattice, see Fig. 5a with

$$\vec{R}_n = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 , n_i \in \mathbb{Z}.$$

It is very useful to classify such lattices by their symmetries (rotations, mirror, etc.). There are 14 symmetry different Bravais lattices. The one with the highest symmetry is the cubic lattice where

$$a_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}, a_2 = \begin{pmatrix} 0 \\ a \\ 0 \end{pmatrix}, a_3 = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix}$$

The Bravais lattice with the lowest symmetry is, e.g., called 'triclinic' where all $|\vec{a}_i|$ are different and point in arbitrary directions.

It is useful to think about the 3D lattice as a set of **unit cells** connected by vectors of the Bravais lattice. One possible choice of such a unit cell is the Wigner Seitz



(c) Wigner Seitz cell (area inside dashed lines)

Figure 5: Sketches of a 2-dimensional Bravais lattice, the atomic configuration of NaCl and for the construction of a Wigner-Seitz cell.

cell, see Fig. 5c, defined by

$$C = \left\{ \vec{x} \in \mathbb{R}^3 \, \middle| \, |\vec{x}| \le |\vec{x} - \vec{R}_n| \, \forall \, \vec{n} \ne \{0, 0, 0\} \right\}$$

This choice of a unit cell has the advantage that it has – by construction – the same symmetries as the Bravais lattice.

Unit cells are a useful concept as they allow for a unique decomposition of a position vector in a Bravais-lattice vector and a rest

$$\vec{r} = \vec{R}_n + \vec{x}, \ \vec{x} \in C.$$

A key quantity for our analysis of electrons in a solid will be the periodic potential created in such systems

$$U(\vec{r} + \vec{R}_n) = U(\vec{r}).$$

230 different so-called space groups describe the possible symmetries of U.

We next study the Fourier transform of $U(\vec{r})$

$$U(\vec{r}) = \sum_{i} U_{\vec{G}_i} e^{i\vec{G}_i\vec{r}}.$$

The periodicity of U puts strong constraints on the allowed Fourier components,

$$U(\vec{r} + \vec{R}_n) = U(\vec{r}) \implies e^{i\vec{G}_i\vec{R}_n} = 1 \,\forall \,\vec{G}_i, \vec{R}_n.$$

Only Fourier components can contribute which belong the **reciprocal lattice** with reciprocal-lattice vectors

$$\vec{G}_n = n_1 \vec{G}_1 + n_2 \vec{G}_2 + n_3 \vec{G}_3, n_i \in \mathbb{Z}$$

with $\vec{G}_i \vec{a}_j = 2\pi \delta_{ij}, i, j = 1, 2, 3$

where the \vec{a}_i are the basis of the Bravais lattice. The reader can check that $\vec{G}_i \vec{R}_n$ is always an integer multiple of 2π with these definitions, and therefore $e^{i\vec{G}_i\vec{R}_n} = 1$ for all i and n. For the above case \vec{G}_1 is given by:

$$\vec{G}_1 = 2\pi \, \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}$$

and $\vec{G}_{2,3}$ are obtained by cyclically changing indices in the formula, $1 \to 2 \to 3 \to 1$. An important concept, frequently used by us, is the 1st **Brillouin zone**, which we abbreviate by **1. BZ**. Is is defined as the Wigner-Seitz cell in reciprocal space,

$$1.BZ = \left\{ \vec{k} \in \mathbb{R}^3 \, \middle| \, |\vec{k}| \le |\vec{k} - \vec{G}_n| \, \forall \, \vec{n} \ne \{0, 0, 0\} \right\}$$

We will use it to obtain a unique decomposition of vectors \vec{K} in reciprocal space

$$\vec{K} = \vec{G}_n + \vec{\kappa}$$
, where $\vec{\kappa} \in 1.$ BZ.

Let us consider as an example the case of NaCl. In this case, one obtains a so-called face-centered cubic lattice (see Fig. 5b, the blue balls are located on the edges and in the middle of the faces of a cube). As basis vectors of the Bravais lattice one can choose

$$\vec{a}_1 = \frac{a}{2} \begin{pmatrix} 0\\1\\1 \end{pmatrix}, \vec{a}_2 = \frac{a}{2} \begin{pmatrix} 1\\0\\1 \end{pmatrix}, \vec{a}_3 = \frac{a}{2} \begin{pmatrix} 1\\1\\0 \end{pmatrix}$$

The Na and Cl atoms can then be found at positions

Na at
$$\vec{R}_n + \begin{pmatrix} 0\\0\\0 \end{pmatrix}$$
, Cl at $\vec{R}_n + \frac{a}{2} \begin{pmatrix} 0\\0\\1 \end{pmatrix}$.

where \vec{R}_n are the vectors of the Bravais lattice.

From this, we compute the basis vectors of the reciprocal lattice

$$\vec{G}_1 = \frac{2\pi}{a} \begin{pmatrix} -1\\ 1\\ 1 \end{pmatrix}, \vec{G}_2 = \frac{2\pi}{a} \begin{pmatrix} 1\\ -1\\ 1 \end{pmatrix}, \vec{G}_3 = \frac{2\pi}{a} \begin{pmatrix} 1\\ 1\\ -1 \end{pmatrix}$$

The corresponding Bravais lattice forms a space-centered cubic lattice (corners of cubes plus a point in the middle of the cube).

2.4.2 Bloch theorem

A central question concerns the form of single electron eigenfunction in the presence of a periodic potential. The Bloch theorem puts strong restrictions on their shape: they are 'essentially' plane waves but modified by a periodic function. Furthermore, it shows how the band-structure is defined with the 1. BZ.

Theorem: The eigenfunctions of a single-particle Hamiltonian with a periodic potential,

$$\mathcal{H} = -\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})$$
, with $U(\vec{r} + \vec{R}_n) = U(\vec{r}) \; \forall \vec{R}_n$ in a Bravais lattice

can be chosen as

$$\phi_{m,\vec{\kappa}}(\vec{r}) = e^{i\vec{\kappa}\vec{r}} u_{m,\vec{\kappa}}(\vec{r}), \text{ with } u_{m,\vec{\kappa}}(\vec{r}+\vec{R}_n) = u_{m,\vec{\kappa}}(\vec{r}) \,\forall \,\vec{R}_n$$

Here, m is the **band index**, $\vec{\kappa} \in 1.BZ$

The eigenenergies $\varepsilon_m(\vec{\kappa})$ are periodic, $\varepsilon_m(\vec{\kappa} + \vec{G}_n) = \varepsilon_m(\vec{\kappa})$ for $\vec{G}_n \in$ reciprocal lattice. They define the **band-structure** of the system.

While the Block theorem is a theorem on a single-particle Hamiltonian, it is still instructive to use the language of second quantization for its proof.

Proof: We start by writing the Hamiltonian in momentum space.

$$\mathcal{H} = \sum_{\vec{k}} \frac{(\hbar \vec{k})^2}{2m} c_{\vec{k}}^{\dagger} c_{\vec{k}} + \sum_{\vec{k}} \sum_{\vec{G}_n} U_{\vec{G}_n} c_{\vec{k}+\vec{G}_n}^{\dagger} c_{\vec{k}} = \sum_{\vec{\kappa}\in 1\text{BZ}} \mathcal{H}_{\vec{\kappa}}$$
$$\mathcal{H}_{\vec{\kappa}} = \sum_{\vec{G}_n, \vec{G}_m} c_{\vec{\kappa}+\vec{G}_n}^{\dagger} A_{n,m}^{\vec{\kappa}} c_{\vec{\kappa}+\vec{G}_m}^{\dagger} \text{ with } A_{n,m}^{\vec{\kappa}} = \frac{\hbar^2 (\vec{\kappa}+\vec{G}_n)^2}{2m} \delta_{n,m} + U_{\vec{G}_n-\vec{G}_m}$$

Importantly, by translational invariance, the momentum \vec{k} only couples with momenta $\vec{k} + \vec{G}_n$. This allows to split the Hamiltonian in sectors labeled by a vector $\vec{\kappa}$ in the first Brillouin zone. For fixed $\vec{\kappa}$, $A^{\vec{\kappa}}$ is a hermitian matrix. By diagonalization one obtains eigenvalues $\varepsilon_{i,\vec{\kappa}}$, $i \in \mathbb{N}$ and eigenvectors eigenvectors $u_{i,\vec{\kappa},\vec{G}_n}$. We can use those, to bring our 2nd quantized Hamiltonian into diagonal form

$$\mathcal{H}_0 = \sum_n \sum_{\vec{\kappa} \in 1\text{BZ}} \varepsilon_{n,\vec{\kappa}} d^{\dagger}_{n,\vec{\kappa}} d_{n,\vec{\kappa}}$$
(73)

by defining

$$d_{n,\vec{\kappa}}^{\dagger} = \sum_{\vec{G}_m} c_{\vec{\kappa}+\vec{G}_m}^{\dagger} u_{n,\vec{\kappa},\vec{G}_m} = \frac{1}{\sqrt{\mathcal{V}}} \int d^d r \, \phi_{n,\vec{\kappa}}(\vec{r}) \psi^{\dagger}(\vec{r})$$

Here, we define the single-particle wave function by

$$\phi_{n,\vec{\kappa}}(\vec{r}) = \sum_{\vec{G}_m} e^{i(\vec{\kappa} + \vec{G}_m)\vec{r}} u_{n,\vec{\kappa},\vec{G}_m}$$

Note that the operator $c_{n,\vec{\kappa}}^{\dagger}$ creates electrons in band *n* with energy $\varepsilon_{n,\vec{\kappa}}$.

Setting $\phi_{n,\vec{\kappa}}(\vec{r}) \stackrel{!}{=} e^{i\vec{\kappa}\vec{r}} u_{n,\vec{\kappa}}(\vec{r})$ we get:

$$u_{n,\vec{\kappa}}(\vec{r}) = \sum_{\vec{G}_m} e^{i\vec{G}_m\vec{r}} u_{n,\vec{\kappa},\vec{G}_m} \implies u_{n,\vec{\kappa}}(\vec{r}+\vec{R}_n) = u_{n,\vec{\kappa}}(\vec{r})$$

By construction the matrix $A_{\vec{\kappa}}$ has the property $A_{\vec{\kappa}+\vec{G}_n} = A_{\vec{\kappa}}$ and therefore the eigenenergies have to be periodic, $\varepsilon_{n,\vec{\kappa}+\vec{G}_m} = \varepsilon_{j,\vec{\kappa}}$ which completes the proof of the Bloch theorem. Note that the proof is constructive: we can use the description above also as a numerical algorithm to compute the bandstructure and the Boch wavefunctions.

2.4.3 Fermi sea and Fermi surface

Let us for now ignore the effects of electron-electron interactions and focus on the properties of non-interacting particles in a periodic potential. Remarkably, it turns out that the picture developed here remains qualitatively correct even when one properly considers interactions – why that is the case, will become clear only much later in the lecture.

For now, let us start with the question on how to obtain the groundstate of the Hamiltonian $\mathcal{H}_0 = \sum_n \sum_{\vec{\kappa} \in 1\text{BZ}} \varepsilon_{n,\vec{\kappa}} d^{\dagger}_{n,\vec{\kappa}} d_{n,\vec{\kappa}}$, Eq. (73), when we add a macroscopic number of electrons, N, to the system, such that the electron density n = N/V is finite. Using Pauli principle, the state with lowest energy is obtained, if we fill up all states up to a maximal energy, the **Fermi energy**. The ground state wavefunction is therefore the so-called **Fermi sea** with

$$|FS\rangle = \prod_{\varepsilon_{m,\vec{\kappa}} < \varepsilon_F} d^{\dagger}_{m,\vec{\kappa}} |0\rangle$$

where $N = \langle \hat{N} \rangle = \langle FS | \sum_{m,\vec{\kappa}} d^{\dagger}_{n,\vec{\kappa}} d_{n,\vec{\kappa}} | FS \rangle = \sum_{\varepsilon_{m,\vec{\kappa}} < \varepsilon_F} 1$ is the number of occupied states (the spin-degree of freedom can be absorbed in the band index m, if bands are spin-degenerate, each band contributes twice to the sum).

As we will see later, all physical properties of metals are dominated by electrons with $\varepsilon_{m,\vec{\kappa}} \approx \varepsilon_F$ (as $k_B T \ll \epsilon_F$ even at room temperature for most metals). We call the area in the 1. BZ with $\varepsilon_{m,\vec{\kappa}} \approx \varepsilon_F$ the **Fermi surface**.

We will later mainly study grand-canonical ensembles with the density matrix $\rho \sim e^{-\beta(H_0-\mu\hat{N})}$, where $\hat{N} = \sum_{m,\vec{\kappa}} d^{\dagger}_{n,\vec{\kappa}} d_{n,\vec{\kappa}}$ is, as above, the operator counting the number of electrons. Thus, we will often consider

$$\mathcal{H}_0 - \mu \hat{N} = \sum_{n,\vec{\kappa}} (\varepsilon_{m,\vec{\kappa}} - \mu) d^{\dagger}_{m,\vec{\kappa}} d_{m,\vec{\kappa}}$$

Sometimes, we will simply include μ in the definition of energy, $\varepsilon_{m,\vec{\kappa}} \to \varepsilon_{m,\vec{\kappa}} - \mu$. In the limit $T \to 0$, Fermi energy and chemical potential become equivalent

$$\lim_{T \to 0} \mu = \varepsilon_F$$

Sometimes it is useful to interpret $|FS\rangle$ as the new vacuum of the system. We can do that by defining

$$f_{m,\vec{\kappa}}^{\dagger} = \begin{cases} d_{m,\vec{\kappa}}^{\dagger} & \text{for } \varepsilon_{m,\vec{\kappa}} > \mu \\ d_{m,\vec{\kappa}} & \text{for } \varepsilon_{m,\vec{\kappa}} < \mu \end{cases}$$

With this definition it follows that

$$f_{m,\vec{\kappa}}|FS\rangle = 0\,\forall\,m,\vec{\kappa}$$

which justifies calling $|FS\rangle$ the 'vacuum'. The Hamiltonian in these variables is given by

$$\mathcal{H}_0 - \mu \hat{N} = \sum_{m,\vec{\kappa}} |\varepsilon_{m,\vec{\kappa}} - \mu| f^{\dagger}_{m,\vec{\kappa}} f_{m,\vec{\kappa}} + \text{constant}$$

and one can see directly, that any excitation relative to the ground-state costs a positive energy.

2.4.4 Lattice models

Motivation: An important step for the understanding of physical phenomena is the use of simplified model systems. Often one asks the question what the simplest model could be which describes a given effect. Furthermore, even if one aims at a full quantitative of, e.g., a material, one does not need to know all the details at energy scales of several electron volt (1 eV corresponds to a temperature scale of 10000 K) but one is only interested at low energies. Thus, one can focus on a few bands which contribute to the Fermi surface. Furthermore, our goal is to obtain an intuitive but (semi-) quantitative picture, how electrons move in a sold.

Starting from the observation, that the electronic wavefunctions remain peaked close to atomic orbitals, we want to develop an intuitive picture based on electrons, which tunnels from an atomic orbital n at site i to another atomic orbital m at site j with rate $t_{i,j}^{n,m}$.

We can translate this physical picture in a very direct way into the language of second quantization by introducing a creation operator $a_{n,i}^{\dagger}$ which creates an electron at site i in orbital n (we will discuss later that nominally one should replace the atomic orbitals by orthonormal Wannier functions, but we can ignore this detail for the moment). We can now simply write the Hamiltonian as

$$\mathcal{H}_{0} = \sum_{i,j,n,m} t_{i,j}^{n,m} a_{m,j}^{\dagger} a_{n,i} + \sum_{i,n,m} \Delta_{n,m} a_{m,i}^{\dagger} a_{n,i},$$

$$\left\{ a_{m,j}^{\dagger}, a_{n,i} \right\} = \delta_{i,j} \,\delta_{n,m}$$

$$(74)$$

where $\Delta_{n,m}$ encodes the rate of hopping and local energies between different orbitals at the same site. Often one chooses a basis, where this term is diagonal, $\Delta_{n,m} = \epsilon_n \delta_{n,m}$. Models like \mathcal{H}_0 are called "tight binding" models. They are both intuitive and can give a precise quantitative description of the electronic band structure close to the Fermi energy by either using the $t_{i,j}^{n,m}$ as fitting parameters or by determining them from more refined calculations, see below. Consider for example, a cubic lattice, where electrons reside in s orbitals. Typically the hopping rate will be largest for nearest neighbors on the cubic lattice, and thus one obtains (suppressing the spin index)

$$\mathcal{H}_0 = -t \sum_{\langle i,j \rangle} a_i^{\dagger} a_j \tag{75}$$

with hopping rate t and the symbol $\langle i, j \rangle$ = denotes that we sum over nearest neighbors on the given lattice. This problem can be diagonalized by using creation operators in momentum space

$$c_{\vec{\kappa}}^{\dagger} = \frac{1}{\sqrt{N}} \sum_{j} e^{i\vec{\kappa}\vec{R}_j} a_j^{\dagger}, \text{ and } a_j^{\dagger} = \frac{1}{\sqrt{N}} \sum_{\vec{\kappa}} e^{-i\vec{\kappa}\vec{R}_j} c_{\vec{\kappa}}^{\dagger}$$

Neighboring sites in the cubic lattice are connected by 6 vectors $\vec{\Delta}_n$, $n = 1, \ldots, 6$

$$\vec{R}_i = \vec{R}_j + \vec{\Delta}_n$$
, where $\vec{\Delta}_n = (\pm 1, 0, 0)^T, (0, \pm 1, 0)^T, (0, 0, \pm 1)^T$

where a is the lattice constant. Such an extremely simple approach can already gives semi-quantitatively good descriptions of many systems, which can easily be refined by including, e.g., next-nearest neighbor hopping rates.

Using these definitions in Eq. (75) we obtain

$$\mathcal{H}_{0} = -\frac{t}{N} \sum_{\vec{\kappa},\vec{\kappa}',n,j} e^{-i\vec{\kappa}(\vec{R}_{j}+\vec{\Delta}_{n})} e^{i\vec{\kappa}'\vec{R}_{j}} c_{\vec{\kappa}}^{\dagger} c_{\vec{\kappa}'} = -t \sum_{\vec{\kappa},\vec{\kappa}',n} \underbrace{\left(\frac{1}{N} \sum_{j} e^{i(\vec{\kappa}'-\vec{\kappa})\vec{R}_{j}}\right)}_{i\vec{\kappa}''} e^{i\vec{\kappa}\vec{\Delta}_{n}} c_{\vec{\kappa}}^{\dagger} c_{\vec{\kappa}'}$$
$$= \sum_{\vec{\kappa}} \left(-t \sum_{n} e^{i\vec{\kappa}\vec{\Delta}_{n}}\right) c_{\vec{\kappa}}^{\dagger} c_{\vec{\kappa}} = \sum_{\vec{\kappa}} \varepsilon_{\vec{\kappa}} c_{\vec{\kappa}}^{\dagger} c_{\vec{\kappa}}$$

Thus,

$$\mathcal{H}_0 = \sum_{\vec{\kappa}} \varepsilon_{\vec{\kappa}} c^{\dagger}_{\vec{\kappa}} c_{\vec{\kappa}}, \qquad \varepsilon_{\vec{\kappa}} = -t \sum_n e^{i\vec{\kappa}\vec{\Delta}_n} = -2t(\cos(\kappa_x a) + \cos(\kappa_y a) + \cos(\kappa_z a))$$

Above, we described how one can make a reasonable guess of the band structure of a materials with a small number of fitting parameters. One can also proceed the opposite way. Let us assume that one already has a full solution of the problem, i.e., one knows the eigen energies and the Bloch wave function. How one can obtain those in a good approximation even for electrons with Coulomb interactions is not part of this lecture (it is covered in the solid-state theory course) but refined numerical codes exist which provide precisely that. Then, it is still useful to obtain a proper tight-binding model. The process of obtaining those is called **down-folding**. Here one proceeds the following way. We start by selecting a subset of M electronic bands n_1, n_2, \dots, n_M which should include at least all of those, which contribute to the Fermi surface. We denote the corresponding Bloch states by $\varphi_{n_i,\vec{\kappa}} = e^{i\vec{\kappa}\vec{r}}u_{n_i,\vec{\kappa}}(\vec{r})$, $\vec{\kappa} \in 1.BZ$. As a next step, we use those to construct a so-called **Wannier-function**, which is localized close to the site \vec{R}_i of an atom.

$$\phi_{m,\vec{R}_i}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\substack{\vec{\kappa} \in 1.\text{BZ} \\ j = 1, 2, \cdots, M}} \mathcal{U}_{m,n_j}^{\vec{\kappa}} \varphi_{n_j,\vec{\kappa}}(\vec{r}) e^{-i\vec{\kappa}\vec{R}_i}$$

There is a substantial arbitrariness in choosing the unitary matrix $\mathcal{U}_{m,n_j}^{\vec{k}}$. Here, the strategy is to make a choice so that the resulting wavefunction is as much as possible localized close to \vec{R}_i (the resulting Wannier functions are called 'maximally localized'). If the Wannier function is not sufficiently localized one has to add further bands.

Outlook: If one chooses bands with a net topological winding number, which is a kind of knot in their wavefunctions, then it is necessary to add further bands to achive a net zero winding number. Only then, the Wannier functions are decaying exponentially at long distances and only then one obtains a reasonable tight-binding model based on a localized Wannier function. For the description of various topological materials (defined here as materials which hosts such bands) one therefore needs to include a minimal set of bands even for the simplest possible tight-binding model. While Wannier functions have been introduced in 1937 and have been heavily used since then, this important relation of Wannier functions with topology was fully realized only 70 years later.

By construction, Wannier functions are orthonormal to each other, even when computed at different sites (in contrast to standard atomic orbitals)

$$\int \phi_{n,\vec{R_i}}^* \phi_{n',\vec{R_j}} = \delta_{n,n'} \delta_{i,j}$$

Introducing the operators $c_{n,\vec{\kappa}}^{\dagger}$ which creates an electron in the Bloch state $\varphi_{n,\vec{\kappa}}$, $\vec{\kappa} \in 1.BZ$, we define the creation operator of electrons in a Wannier state by

$$a_{m,i}^{\dagger} = \frac{1}{\sqrt{N}} \sum_{\vec{\kappa}} \mathcal{U}_{m,n_i}^{\vec{\kappa}} c_{n_i,\vec{\kappa}}^{\dagger} e^{-i\vec{\kappa}\vec{R}_i}$$
(76)

Fits to the band-structures can then be used to determine the hopping rates from one orbital to the next and to derive microscopically the tight-binding model shown in Eq. (74).

Up to now, we considered a model of non-interacting electrons. In reality, there will be **interactions**. Such an interacting model takes the form

$$\mathcal{H} = \sum_{i,j} t_{i,j} a_i^{\dagger} a_j + \sum U_{ijkl} a_i^{\dagger} a_j^{\dagger} a_k a_l, \qquad (77)$$

where we used a compact notation where $i = (m_i, \vec{R}_i, m, \sigma_i)$ denotes the position, the band index and the spin-index.

What are the interaction matrix elements? For electrons, their physical origin is the Coulomb interaction, $V(r) = \frac{e^2}{4\pi\epsilon_0\epsilon r}$. Therefore, one might be tempted to compute the interaction from

$$U_{ijkl} \stackrel{?}{=} \frac{1}{2} \int d^3r d^3r' \,\phi^*_{m_i,\vec{R}_i,\sigma_i}(\vec{r}) \phi^*_{m_j,\vec{R}_j,\sigma_j}(\vec{r}') \phi_{m_k,\vec{R}_k,\sigma_k}(\vec{r}') \phi_{m_l,\vec{R}_l,\sigma_l}(\vec{r}) V(\vec{r}-\vec{r}')$$

This is, however, not a good approximation because it ignores that the Coulomb interaction also deforms the wavefunction of all the other bands, not included in the tight-binding description and one has to use more refined methods (e.g., recalculating the bands while fixing the charge on a single atom) to compute the interactions.

In practice, numerical band structure packages based on "density function theory' can be used to obtain relatively reliable values for the band structures, Wannier functions and hopping rates in a wide range (but not all) materials. A reliable calculation of the interaction matrix elements is more difficult but also a lot of progress has been made in this respect in the recent years.

An alternative and very successful approach is to 'guess' appropriate tight-binding models and interactions (often guided by insights gained from band structure calculations) and considering the unknowns in these models as parameters which can be varied to obtain different physical effects. We will follow this line of argument in the next section.

2.5 Hubbard and Heisenberg models

Motiviation: Simple model systems play an extremely important role to guide our understanding of physical systems. The most important model for interacting electrons is the Hubbard model, the most important for magnetism in quantum systems is the Heisenberg model. We will discuss both and show how the Heisenberg model can be derived from the Hubbard model. These models define some of the most studied quantum field theories in the context of solid-state physics.

2.5.1 Hubbard model

Often in metals one can ignore the long-ranged part of the interactions (due to screening) and one can focus instead only in interactions arising when electrons occupy the same orbital. The simplest possible model of interacting electrons is the Hubbard model

$$\mathcal{H} = -t \sum_{\langle i,j \rangle, \sigma=\uparrow,\downarrow} c^{\dagger}_{i,\sigma} c_{j,\sigma} + U \sum_{i} n_{i,\uparrow} n_{i,\downarrow}$$

with $n_{i,\sigma} = c_{i,\sigma}^{\dagger} c_{i,\sigma}$. The first term describes hopping from one atom to the next (one most often considers square lattices in 2d or cubic lattices in 3d). When there are two electrons on the same site $(n_{i,\uparrow} = n_{i,\downarrow} = 1)$, this costs an extra energy U > 0, arising from the Coulomb repulsion of two electrons in the same orbital.

The Hubbard model is extremely simple, but shows a remarkably **rich physics**, which is amazing taking into account that there are (besides the lattice structure) **only three relevant parameters**, the strength of interactions $\frac{U}{t}$, the density of electrons n, i.e., the number of electrons per site, and the temperature T/t. It is also one of the most studied models in physics. A search for 'Hubbard model' reveals more than 17.000 publications.

It can be used to explore and explain, e.g., the properties of weakly interacting fermions, how strong interactions can induce insulators (see below), how ferro- and

antiferromagnetism (see blow) can arise, how repulsive interactions can induce exotic superconductivity and many more things. In one spatial dimensions, one can surprisingly solve it exactly (with massive mathematical effort based on the Bethe ansatz), in d = 2,3 many features are not fully understood. In d > 1, it is also extremely difficult to solve (with the exception of a few limiting cases). Monte-Carlo methods, for example, fail for electronic densities different from 1 electron per site. (due to the infamous minus-sign problem, which you might have discussed in your computational physics class).

For some solids, the Hubbard model may be a decent, but never perfect approximation as electrons have long-ranged interactions. An almost perfect realization can, however, be realized with fermionic atoms captured in the standing wave of a laser (the first experiment was done in the Bloch group in Munich while our group provided the theory, U. Schneider, *et al.*, Science (2008)).

A prime strategy to understand such models is to analyze limiting cases. Especially easy is the non-interacting limit $U \rightarrow 0$, where one obtains a metal with a relatively simple Fermi surface. This can be the starting point, to analyze weak interactions, which we will do later in this course.

For now, we will discuss the opposite limit of $U \gg t$.

2.5.2 Heisenberg model

We will focus on the case where one has exactly one electron per site, n = 1, also called "half filling". To understand the limit $U \gg t$, we first solve the problem for t = 0

$$\mathcal{H} - \mu \mathcal{N} = U \sum_{i} n_{i,\uparrow} n_{i,\downarrow} - \mu \sum_{i} n_{i,\uparrow} + n_{i,\downarrow}$$

As different sites decouple, we can consider just one site with the spectrum

$$n = 0 \quad E_0 = 0 n = 1 \quad E_1 = -\mu n = 2 \quad E_2 = U - 2\mu$$

An occupation of n = 1 is obtained when the n = 1 states have the lowest energy, e.g., for $\mu = \frac{U}{2}$. In this case, we have $E_2 = E_0 = 0$, while $E_1 = -U/2$. Importantly, there are two states with n = 1. Either a spin-up electron or a spin-down electron can occupy a site $|\uparrow\rangle = c_{\uparrow}^{\dagger}|0\rangle$, $|\downarrow\rangle = c_{\downarrow}^{\dagger}|0\rangle$. For a system with N sites, the ground state has therefore a gigantic degeneracy of 2^N , while adding or removing a particle costs the energy of U/2.

Our goal is to understand what happens when we consider a small but finite hopping t. First, assume that an electron hops to a neighboring site. This leaves one site empty and one double occupied. This costs the energy

$$E_2 + E_0 - 2E_1 = U \gg t.$$

This huge energy costs implies that the electrons will remain localized on each site (up to small quantum fluctuations). Such states of matter are called **Mott insulators**.

We want to understand the property of this Mott insulator for $U \gg t$ but temperatures small compared to $U, T \ll U$. In this limit, we can ignore thermal excitations with n = 0 and n = 2 electrons (sometimes called holes and doublons). Instead, we want to obtain an **effective theory** in the 2^N dimensional low energy space, which we identified above. The rest of the Hilbert space (of dimension $4^N - 2^N$) will only contribute by inprinting perturbative corrections on the low-energy space, which we want to determine below.

How can we formulate a theory in the low-energy subspace? We first have to introduce the relevant operators describing the local degree of freedom for each site, given by the spin of the system. For this, we use the spin operators

$$\vec{S}_{i} = \frac{1}{2} \sum_{\alpha,\beta} c_{i,\alpha}^{\dagger} \vec{\sigma}_{\alpha\beta} c_{i,\beta}$$
(78)

When one calculates the operator \vec{S}_i^2 , one finds that it is given by

$$\vec{S}_{i}^{2} = \begin{cases} 0 & \text{for } n_{i} = n_{i,\uparrow} + n_{i,\downarrow} = 0, 2\\ \frac{1}{2} \left(\frac{1}{2} + 1\right) & \text{for } n_{i} = 1 \end{cases}$$

Thus, in our low-energy space, we can set \vec{S}_i^2 and describe the system with the operators \vec{S}_i only, which obey the algebra

$$\left[S_{j}^{\alpha}, S_{k}^{\beta}\right] = i\delta_{jk}\varepsilon^{\alpha\beta\gamma}S_{j}^{\gamma}$$

How can a model of the low-energy sector look like? Suprisingly, one can guess the answer. First, we expect that interactions remain local, only nearest neighbors are expected to interact in the lmit of large U. Furthermore, we expect that the effective model is invariant under rotations of the spin (this is a property of the original Hubbard model, Eq. (78), which is not fully obvious). This fixes completely the form of the effective low-energy Hamiltonian, the famous **Heisenberg model**

$$\mathcal{H} = J \sum_{\langle i,j \rangle} \vec{S}_i \vec{S}_j \tag{79}$$

More precisely, this is the spin-1/2 Heisenberg model as in our case the local degrees of freedom are spin-1/2 objects, $\vec{S}_i^2 = \frac{1}{2} \left(\frac{1}{2} + 1\right)$. There are also other variants using spins of size s or classical spins.

We need a small calculation to estimate the value of J. Compare two different states for two neighboring sites

- $|\uparrow,\uparrow\rangle$: no hopping due to Pauli principle
- $|\uparrow,\downarrow\rangle$: hopping possible \implies energy gain

To calculate the energy gain, we use perturbation theory to calculate the change in energy. The reader may recall from his quantum mechanics course, that for a Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \Delta \mathcal{H}$ the shift of eigenenergies due to a small $\Delta \mathcal{H}$ reads

$$E \approx E_0 + \langle 0|\Delta \mathcal{H}|0\rangle - \sum_n \frac{|\langle n|\Delta \mathcal{H}|0\rangle|^2}{E_n - E_0} + \mathcal{O}(\Delta \mathcal{H}^3)$$

To use this formula for our problem, we identify \mathcal{H}_0 with the Hubbard model at t = 0, set for two neighboring sites

$$|0\rangle = |\uparrow,\downarrow\rangle = c^{\dagger}_{1\uparrow}c^{\dagger}_{2\downarrow}|0,0\rangle, \qquad \Delta \mathcal{H} = -t\sum_{\sigma=\uparrow,\downarrow}c^{\dagger}_{1,\sigma}c_{2,\sigma} + c^{\dagger}_{2,\sigma}c_{1,\sigma}.$$

Thus, we obtain $\Delta \mathcal{H}|0\rangle = -t(|\uparrow\downarrow,0\rangle + |0\uparrow\downarrow\rangle)$. This is an eigenstate of \mathcal{H}_0 , its energy is U times larger than that of the initial state. From this, we conclude that

$$-\sum_{n} \frac{|\langle n|\Delta \mathcal{H}|0\rangle|^2}{E_n - E_0} = -2\frac{t^2}{U}$$

We therefore conclude that the state $|\uparrow,\downarrow\rangle$ gains from virtual hopping processes energy, while there is no such energy gain for $|\uparrow,\uparrow\rangle$ due to Pauli's principle or, technically, because $\Delta \mathcal{H}|\uparrow,\uparrow\rangle = 0$. Comparing these two energies, we find

$$E_{\uparrow\uparrow} - E_{\uparrow\downarrow} = 2\frac{t^2}{U}.$$

Our goal is to determine the parameter J in the Heisenberg model. Therefore we compute the same quantity for the Heisenberg model.

$$E_{\uparrow\uparrow} - E_{\uparrow\downarrow} = \langle\uparrow,\uparrow|J\vec{S}_1\vec{S}_2|\uparrow,\uparrow\rangle - \langle\uparrow,\downarrow|J\vec{S}_1\vec{S}_2|\uparrow,\downarrow\rangle = J\frac{1}{2}\frac{1}{2} - \frac{1}{2}\left(-\frac{1}{2}\right) = \frac{J}{2}$$

Comparing the two formulas for $E_{\uparrow\uparrow} - E_{\uparrow\downarrow}$ fixes the value of J in the Heisenberg model in the limit $U \gg t$,

$$J \approx 4\frac{t^2}{U} + O(t^3/U^2)$$

This coupling of spins is called **exchange coupling**, a slightly different variant of the same physics arises in models where to magnetic atoms are coupled by another non-magnetic site (e.g., an oxygen atom in a Cu-O-Cu bond). In this case, one uses the term **super exchange coupling** instead.

Above, we did not really derive the Heisenberg model but rather made an eductated guess how it should look like, followed by a calculation of the remaining free parameters. While this works in this specific case and to lowest order in t/U, one needs a more systematic approach to obtain, e.g., corrections to higher order. Below, we give a very brief sketch of this powerful method, which goes under the name **Schrieffer-Wolff transformation**. Variants of it can also be applied whenever one tries to derive an effective low-energy model using a Hamiltonian formalism. It can, e.g., be used to derive systematically relativistic corrections to the Schrödinger equation starting from the Dirac equation. In our case, in can be used to derive the Heisenberg model (and corrections to it) from the Hubbard model at $U \gg t$. To derive an **effective low-energy Hamiltonian**, one starts by splitting the Hilbert space into low energy and high energy states $|e\rangle$, $|h\rangle$ and writes the Hamiltonian in the form $\mathcal{H} = \mathcal{H}_0 + \varepsilon \Delta \mathcal{H}_{hl}$. In our example, the Hubbard model with $U \gg t$, \mathcal{H}_0 is obtained by setting t = 0. \mathcal{H}_0 defines what we mean by low- and high energy states (typically we choose them as eigenstates of H_0). Importantly, H_0 does not include matrix elements mixing high energy and low energy states, $\langle e|\mathcal{H}_0|h\rangle = 0$. For notational simplicity, we also assume that $\langle e|\Delta \mathcal{H}|e'\rangle = \langle h|\Delta \mathcal{H}|h'\rangle = 0$ (if not, we add these terms to \mathcal{H}_0 . With all these definitions, our Hamiltonian takes the following form

$$\mathcal{H} = \begin{pmatrix} \text{high} & \text{low} \\ \mathcal{H}_0 & \varepsilon \Delta \mathcal{H} \\ \hline \varepsilon \Delta \mathcal{H} & \mathcal{H}_0 \end{pmatrix} \begin{array}{l} \text{high} \\ \text{low} \\ \end{array}$$

if we choose a basis where we sort the basis vectors with the high-energy states first and low-energy states second. Note that the matrix elements mixing high and low energies only arise due to $\Delta \mathcal{H}$.

To obtain an effective low-energy model up to corrections of order ϵ^n , we have to get rid of all terms (to this order in ϵ) which mix the low-energy and the high-energy space. Thus, we want to perform a unitary transformation

$$\mathcal{H} \longrightarrow \tilde{\mathcal{H}} = e^{-iS\varepsilon} \mathcal{H} e^{iS\varepsilon}, \quad \text{with } S = S^{\dagger},$$

such that all matrix elements $\langle h | \hat{\mathcal{H}} | l \rangle$ vanish to order ε^n . We will focus below on n = 1, the most studied case. By Taylor expansion in ϵ , we find

$$\tilde{\mathcal{H}} = \mathcal{H} - i\varepsilon[S, \mathcal{H}] - \frac{1}{2}\varepsilon^2[S, [S, \mathcal{H}]] + \cdots$$
$$\approx \mathcal{H}_0 + \varepsilon\Delta\mathcal{H} - i\varepsilon[S, \mathcal{H}_0] - i\varepsilon^2[S, \Delta\mathcal{H}] - \frac{1}{2}\varepsilon^2[S, [S, \mathcal{H}_0]] + \mathcal{O}(\varepsilon^3)$$
(80)

For n = 1, we therefore have to find an operator S' with the property that $i[S, \mathcal{H}_0] = \Delta \mathcal{H}$. This will cancel with the $\epsilon \Delta \mathcal{H}$ which we want to eliminate. In our specific example, we have $\mathcal{H}_0 = U \sum_i n_{i,\uparrow} n_{i,\downarrow}$ and $\Delta \mathcal{H} = -t \sum_{\langle i,j \rangle} c^{\dagger}_{i,\sigma} c_{j,\sigma}$ and one finds that

$$S = i \frac{t}{U} \sum_{\langle i,j \rangle} (n_{i,-\sigma} (1 - n_{j,-\sigma}) - n_{j,-\sigma} (1 - n_{i,\sigma})) c_{i,\sigma}^{\dagger} c_{j,\sigma}$$

has the desired property. Thus, we found a unitary transformation which cancels to order ϵ the terms mixing high and low energies. Importantly, the same unitary transformation creates to order ϵ^2 new terms in the low-energy space. Evaluating these terms in Eq. (80) (we skip this somewhat lengthy calculation here) we obtain indeed the Heisenberg model with $J = \frac{4t^2}{U}$. Compared to our first approach, this is a much more complicated calculation but it is systematic and can be used to calculate systematically corrections to higher order in t/U.

Outlook: The method which we sketched above is powerful and can be used to derive in a Taylor expansion in t/U an effective low-energy model. As low-energy models play such an important role in physics (even the standard model of particle physics is most likely an effective low-energy model), it is interesting to ask the question, how well justified the method is both from a mathematical and physical point of view. Like many other expansion methods used in physics, it turns out to be **not** a converging Taylor expansion with a finite radius of convergence. Even for very small, but finite $t \ll U$, it breaks down at some point. This can be seen from the fact, that flipping $n^* \approx U/J \sim (U/t)^2$ spins in the Heisenberg model, we reach the energy $n^*J \approx U$ and there is an overlap of high-energy and low-energy states. Thus, at order ϵ^{n^*} the Schrieffer-Wolff transformation will break down. Experimentally, one can address this by measuring the lifetime τ_h of a high-energy state with energy U. One finds that the lifetime becomes exponentially large and is roughly given by $\tau_h \sim (U/t)^{n^*} \sim e^{(U/t)^2 \ln(U/t)}$. This physics is impossible to access using a Taylor expansion in t/U only.

Let us discuss some basic properties of the Heisenberg model, Eq. (79), e.g., on a three-dimensional cubic lattice. At finite temperature, the physics is determined by the ratio T/J only. Thus, it is useful to consider two limits

- for $T \gg J$ the spins are disordered and point in random directions
- for $T \ll J$, there is instead antiferromagnetic long-range order, schematically something like

We will discuss excitations in this antiferromagnet in more detail below.

The Heisenberg model was derived assuming that there is exactly one electron per site in the Hubbard model. What happens when one adds or removes a few electrons? What is the resulting low-energy Hamiltonian for $U \gg t$? Consider, for example, the case that one removes electrons, so that the density of electrons is now smaller than 1, $n = 1 - \epsilon$. In this case, the low energy states have either occupation 0 or 1, while the doubly occupied states with 2 electrons are still part of the high-energy Hilbert space. One can follow the program sketched above, to derive again a lowenergy Hamiltonian in the limit of small doping and large U. This low-energy model is called the **tJ model** given by

$$\mathcal{H}_{t\mathscr{J}} = -t \sum_{\langle i,j \rangle} \tilde{c}_{i,\sigma}^{\dagger} \tilde{c}_{j,\sigma} + J \sum_{\langle i,j \rangle} \vec{S}_i \vec{S}_j + \mathcal{O}\left(\frac{t^3}{U^3}, \frac{t^2}{U}(1 - \langle n \rangle)\right)$$

with $J = 4t^2/U$ as above, for the spin operators one has to use the definition of Eq. (78) and we use the definition

$$\tilde{c}_{i,\sigma}^{\dagger} = c_{i,\sigma}^{\dagger} (1 - n_{i,-\sigma})$$

to make sure that the hopping terms do not create doubly occupied sites. The tJ model is also a widely studied model in physics, which will, however, play no further role in this lecture.

2.6 Magnetism and Spin waves

Motivation: Understanding the properties of a quantum field theory requires us to understand the low-energy excitations of a system. In this chapter we explore magnons, the low-energy excitations of ferromagnets and antiferromagnets. This will also serve as an exercise in using second quantization and in deriving low-energy theories.

The starting point of our study is the Heisenberg model defined by

$$\mathcal{H} = J \sum_{\langle i,j \rangle} \vec{S}_i \vec{S}_j$$
$$[S_n^{\alpha}, S_m^{\beta}] = i \varepsilon_{\alpha\beta\gamma} \delta_{nm} S_n^{\gamma}$$

but in contrast to our discussion above, which was restricted to half-integer spins, we will allow for an arbitrary value of $\vec{S}_i^2 = s(s+1)$ implying that S_z can take the eigenvalues $-s, -s+1, \cdots, s-1, s$, resulting in a $(2s+1)^N$ dimensional Hilbert space for an N-site system, where possible values for s are $1/2, 1, 3/2, \ldots$

Experimentally, larger spins arise from having several electrons per atom with their spin aligned parallel. The reason for that is called Hund's rule coupling (due to Pauli principle electrons with the same spin do not come close to each other, reducing the amount of repulsive Coulomb energy).

Both sighs of J can occur. For J > 0, one obtains **antiferromagnetic order** for cubic lattices considered below (more complex ordering patterns may arise on, e.g., a triangular lattice).

J < 0 implies **ferromagnetic order**, where all spins tend to align parallel. Most ferromagnets in nature are metals (and therefore not described by the Heisenberg model, which does not include charge degrees of freedom), because for most insulators a variant of the antiferromagnetic (super-) exchange derived in the previous section dominates. However, there are also insulating ferromagnets, including magnetite, Fe₃O₄ (used first in China to build, e.g., a compass). To obtain ferromagnetism in insulators one needs either a more complicated internal magnetic structure (magnetite is actually a ferrimagnet, hosting internally also an antiferromagnetic order) or a more complex orbital structure. We will, however, focus only on the qualitative properties predicted by the simplified Heisenberg model.

2.6.1 Magnon excitations in ferromagnets

We first consider the ferromagnetic case, J < 0. The ferromagnet has the unique property, that we know analytically the ground state: all spins align parallel. One such state is

$$|0\rangle = |\uparrow\uparrow\uparrow\cdots\rangle,$$

$$\mathcal{H}|0\rangle = J\sum_{\langle i,j\rangle} \vec{S}_i \vec{S}_j |0\rangle = E_0 |0\rangle, \qquad E_0 = -\frac{|J|}{4} \frac{zNs^2}{2}$$

Here z is the number of nearest neighbors of a given site, thus z = 2d on a ddimensional cubic lattice, s is the size each spin, and we replace J = -|J| to emphasize that the energy is negative. The ground state of a system with N sites has the total spin S = Ns. For such a spin, the spin in z-direction, $S_{\text{total}}^z = \sum_i S_i^z$, can take the values $-S, -S + 1, \ldots, S$ and we obtain the

ground state degeneracy:
$$2Ns + 1$$
.

This is a high-degeneracy in the thermodynamic limit but the entropy per site, $S/N = \frac{1}{N} \ln(2Ns + 1)$ still vanishes in the thermodynamical limit, $N \to \infty$ (consistent with the third-law of thermodynamics).

To simplify notations, let us now consider s = 1/2 (but almost identical formulas apply also for arbitrary s). Remarkably, one can also give an analytical formula for so-called **one-magnon excitations** of the ground state excitation:

$$|\vec{K}\rangle = \sum_{i} e^{i\vec{K}\vec{R}_{i}} | \uparrow \uparrow \uparrow \downarrow_{\uparrow} \uparrow \uparrow \uparrow \rangle = \sum_{i} e^{i\vec{K}\vec{R}_{i}}S_{i}^{-}|0\rangle$$

with

$$S_{i}^{-} = S_{i}^{x} - iS_{i}^{y} \stackrel{\uparrow}{=} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_{i}, \qquad S_{i}^{+} = S_{i}^{x} + iS_{i}^{y} \stackrel{\uparrow}{=} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}_{i}.$$

To show that this is indeed an eigenstate of the Hamiltonian, we rewrite it in a somewhat more convenient form

$$\mathcal{H} = -|J| \sum_{\langle i,j \rangle} S_i^z S_j^z + \frac{1}{2} \left(S_i^+ S_j^- + S_i^- S_j^+ \right).$$
(81)

Next, we want to calculate $\mathcal{H}|K\rangle$. For this, we have to compute what $S_i^+S_j^-$ for neighboring sites

$$S_i^+ S_j^- |\uparrow \dots \uparrow \underset{i \, j}{\downarrow} \uparrow \dots \rangle = |\uparrow \dots \uparrow \underset{i \, j}{\uparrow} \dots \rangle$$

Here the spin-down moves from site i to site j. Now we use this to compute

$$(\mathcal{H} - E_0)|\vec{K}\rangle = -|J| \left(\frac{-z_1^2}{2} + \frac{1}{2} \sum_i e^{i\vec{K}\cdot\vec{\Delta}_i} + e^{-i\vec{K}\cdot\vec{\Delta}_i} \right) |\vec{K}\rangle = E_{\vec{K}} |\vec{K}\rangle$$

where the vectors $\vec{\Delta}_i$, $i = 1, \ldots, z$ connect a given spin to its z nearest neighbors. The energy of the one-magnon excitation is therefore given by

$$E_{\vec{K}} = \frac{|J|}{2} \sum_{i=1}^{z} (1 - \cos(\vec{\Delta}_i \vec{K})) \approx \frac{|J|}{2} \vec{K}^2 a^2 \text{ (cubic lattice)}$$

Note that the energy of this one-magnon mode vanishes for $\vec{K} \to 0$, E(K = 0) = 0. This is an example for the **Goldstone theorem**: in systems with short-ranged interaction where a continuous symmetry is spontaneously broken, there exist an excitation with zero energy, a "massless Goldstone mode". While we do not prove the theorem here, it is very intuitive: obviously, it does not cost any energy to rotate the spins, but the groundstate is breaks spin-rotation symmetry. Thus there should be an excitation of zero energy.

Above, we performed an exact calculation which was possible because the exact wavefunction of both the groundstate and the one-magnon state are known in this case. In most systems this is not possible and one needs an approximate method. We mainly want to use this method for the antiferromagnet in the next section, but we practice it here in a somewhat simpler setting. Our plan is to make a **semiclassical approximation**, i.e., an approximation which expands around the classical limit of the problem. In our case, the problem gets more and more classical, when we increase the size s of each spin. But let us see how the problem simplifies in the limit $s \gg 1$.

For this, we first look at the commutation relations of the spin.

$$\left[S_m^z, S_n^{\pm}\right] = \pm S_i^{\pm} \delta_{n,m} \tag{82}$$

$$\left[S_m^+, S_n^-\right] = 2\delta_{n,m} S_m^z \tag{83}$$

If we are now in a large s ferromagnet, we can approximate the operator S_z on the right site by its expectation value $S_m^z \approx \langle S_m^z \rangle \approx s$ and we find that the commutation relation of S^+ and S^- look (up to a trivial factor) very similar to that of bosons

$$\left[\frac{S_m^+}{\sqrt{2s}}, \frac{S_n^-}{\sqrt{2s}}\right] \approx \delta_{m,m}$$

But we need a mathematically more rigorous approach. The Hilbert space of a boson with occupation $n = 0, 1, ..., \infty$ is infinite dimensional while the Hilbert space of a spin of size s has just 2s + 1 states. But we can embedd the spin-Hilbert space into the larger bosonic Hilbert space. For example, for s = 1 there are three states with $S_z = 1, 0, -1$ and we can map those three states to n = 0, 1, 2 bosonic particles,

$$S_z = -1 \iff n = 2$$
$$S_z = 0 \iff n = 1$$
$$S_z = 1 \iff n = 0$$

With this identification, we should be able to rewrite all spin operators in terms of bosonic creation and annhibition operators. The corresponding formula is called **Holstein-Primakoff transformation**

$$S_m^- = a_m^{\dagger} \left(2s - a_m^{\dagger} a_m\right)^{\frac{1}{2}},$$

$$S_m^+ = \left(2s - a_m^{\dagger} a_m\right)^{\frac{1}{2}} a_m,$$

$$S_m^z = S - a_m^{\dagger} a_m$$
(84)

One can check that the operators on the right side obey exactly the same commutation relation than the ones on the left side. The above formulas are exact but they include square-root terms, which are unpleasant to deal with. However, the equation strongly simplify in a **semiclassical approximation** which is also called **large-s approximation**, when we assume $s \gg 1$ and perform a Taylor expansion in 1/s. We obtain, for example,

$$(2s - a_m^{\dagger} a_m)^{\frac{1}{2}} = \sqrt{2s} \left(1 - \frac{a_m^{\dagger} a_m}{2s}\right)^{\frac{1}{2}} \approx \sqrt{2s} - \frac{1}{\sqrt{2s}} \frac{1}{2} a_m^{\dagger} a_m + \cdots$$

and the Hamiltonian is approximated as

$$\mathcal{H} \approx -\frac{|J|Nzs^2}{2} - |J|s \sum_{\langle m,n \rangle} -\underbrace{(a_m^{\dagger}a_m + a_n^{\dagger}a_n)}_{\text{from } S_m^z S_n^z} + a_m^{\dagger}a_n + a_n^{\dagger}a_m.$$

The Hamiltonian is easily diagonalized by introducing the operators in momentum space, $a_k^{\dagger} = \sum_i e^{i\vec{k}\vec{R}_i}a_i^{\dagger}$ and we get

$$\mathcal{H} \approx E_0 + \sum_k \varepsilon_k a_k^{\dagger} a_k + \mathcal{O}(S^0)$$
$$\varepsilon_k = s|J| \sum_{i=1}^z 1 - \cos(\vec{\Delta}_i \vec{k})$$

which (accidentially) turns out to be the exact result even for $s = \frac{1}{2}$. By using Taylor expansions to higher order, one can study how the spin waves interact with each other. This problem is much more easy to treat for bosonic variables than for spins. Therefore this is – even for a ferromagnet – the prefered method to study magnons.

2.6.2 Magnon excitations in antiferromagnets

Now we want to apply the same analysis to the antiferromagnetic Heisenberg model.

$$\mathcal{H} = J \sum_{\langle i,j \rangle} \vec{S}_i \vec{S}_j, \ J > 0 \tag{85}$$

The ground state of this model is (at least for d > 1) not exactly known for any finite s but we can use an expansion in $\frac{1}{s}$. We consider a "bipartite lattice" (e.g., cubic lattice) which can be split up in "A" and "B" sublattices, where all the nearest neighbors of lattice A points belong to sublattice B.

In this setting, the ground state for $s = \infty$ is obtained by, e.g., orienting all spins on the A sublattice is \hat{z} direction, and all spins on the B-sublattice antiparallel to it

We have to adjust the Holstein-Primakoff transformation, Eq. (84), a bit to the new situation such that

$$S^- \sim \begin{cases} a^{\dagger} & \text{on A sublattice} \\ a & \text{on B sublattice} \end{cases}$$

Therefore, we use different formulas for the A and B sublattice

$$S_{m}^{-} = \begin{cases} a_{m}^{\dagger} (2s - a_{m}^{\dagger} a_{m})^{\frac{1}{2}} & m \in A \\ (2s - a_{m}^{\dagger} a_{m})^{\frac{1}{2}} a_{m} & m \in B \end{cases}, \quad S_{m}^{z} = \begin{cases} s - a_{m}^{\dagger} a_{m} & m \in A \\ -(s - a_{m}^{\dagger} a_{m}) & m \in B \end{cases}$$
(87)

We plug this into our Hamiltonian

$$\mathcal{H} = J \sum_{\langle i,j \rangle} S_i^z S_j^z + \frac{1}{2} \left(S_i^+ S_j^- + S_i^- S_j^+ \right)$$

using that the sum is over nearest neighbors only. Thus, when $i \in A \implies j \in B$ and $i \in B \implies j \in A$. After Taylor expansion in 1/s we obtain

$$\mathcal{H} = -NJzs^2 + Js \sum_{\langle i,j \rangle} (a_i^{\dagger}a_i + a_j^{\dagger}a_j + a_i^{\dagger}a_j^{\dagger} + a_ia_j) + \mathcal{O}(s^0)$$
(88)

Compare to the ferromagnet, a new aspect is that we get terms of the form $a_i^{\dagger} a_j^{\dagger}$. This implies that the bosonic vacuum $|0\rangle$ is **not** an eigenstate of the Hamiltonian. To make progress, we go to Fourier space as in the case of the ferromagnet. For the new terms one gets, for example,

$$\sum_{\langle m,n\rangle} a_m^{\dagger} a_n^{\dagger} = \frac{1}{2} \frac{1}{N} \sum_{i=1}^z \sum_n \sum_{k,k'} e^{-i\vec{k}(\vec{R}_n + \vec{\Delta}_i)} a_{\vec{k}}^{\dagger} e^{-i\vec{k'}\vec{R}_n} a_{\vec{k'}}^{\dagger}$$
$$= \sum_k a_{\vec{k}}^{\dagger} a_{-\vec{k}}^{\dagger} \frac{z}{2} \gamma_{\vec{k}}$$
$$\gamma_{\vec{k}} = \frac{2}{z} \frac{1}{2} \sum_i e^{-i\vec{k}\vec{\Delta}_i}$$

For a cubic lattice, $\Delta_i = (\pm 1, 0, 0)^T, (0, \pm 1, 0)^T, (0, 0, \pm 1)^T, i = 1, ..., 6$, and thus γ_k is given by

$$\gamma_k = \frac{1}{3}(\cos(k_x a) + \cos(k_y a) + \cos(k_z a))$$

After Fourier transformation, our Hamiltonian takes the form

$$\mathcal{H} = E_0 + Jsz \sum_{\substack{\vec{k} \in 1.BZ \\ k_x > 0}} \begin{pmatrix} a_k^{\dagger} & a_{-k} \end{pmatrix} \underbrace{\begin{pmatrix} 1 & \gamma_k \\ \gamma_k & 1 \end{pmatrix}}_{C_k} \begin{pmatrix} a_k \\ a_{-k}^{\dagger} \end{pmatrix}$$

where we sum only over half of the 1st BZ, $k_x > 0$, so that each term occurs only once.

Our remaining task is to diagonalize C_k by a canonical transformation, i.e., by a transformation which does not change the commutation relations. Such transformations are called **bosonic Bogoliubov transformations**. We introduce new variables

$$\tilde{a}_k = u_k a_k + v_k a_{-k}^{\dagger}, \qquad \tilde{a}_{-k}^{\dagger} = u_{-k}^* a_{-k}^{\dagger} + v_{-k}^* a_k$$
(89)

or, in a more compact notation,

$$\psi_k = \begin{pmatrix} a_k \\ a_{-k}^{\dagger} \end{pmatrix}, \qquad \tilde{\psi}_k = T_k \psi_k, \qquad T_k = \begin{pmatrix} u_k & v_k \\ v_{-k}^* & u_{-k}^* \end{pmatrix}$$

As stated above, for a canonical transformation the commutation relations should not be modified

$$[\tilde{a}, \tilde{a}^{\dagger}] = [a, a^{\dagger}], \qquad [\tilde{a}, \tilde{a}] = [a, a] = 0$$

This condition, we can also write as

$$\left[\psi_{ki},\psi_{kj}^{\dagger}\right] = \sigma_{ij}^{z} = \left[\tilde{\psi}_{ki},\tilde{\psi}_{kj}^{\dagger}\right] = \left[(T_{k}\psi_{k})_{i},(T_{k}\psi_{k})_{j}^{\dagger}\right] = T_{k}\sigma^{z}T_{k}^{\dagger}$$

Thus, the 2×2 matrix T_k has to have the property that

 $\sigma^z = T_k \sigma^z T_k^{\dagger}$, or, equivalently, $T_k^{-1} = \sigma^z T_k^{\dagger} \sigma^z$

Such matrices are, perhaps, familiar from the physics of Lorentz boosts and they take the form

$$T_k = \begin{pmatrix} \cosh(\eta_k) & \sinh(\eta_k) \\ \sinh(\eta_k) & \cosh(\eta_k) \end{pmatrix}$$

While it is not difficult to obtain T_k directly, we will avoid that in the following. To make progress, we rewrite the Hamiltonian in the new variables.

$$\mathcal{H} = E_0 + \sum_{\substack{\vec{k} \in 1.\text{BZ} \\ k_x > 0}} \psi_k^{\dagger} C_k \psi_k = E_0 + \sum_{\vec{\psi}_k} \tilde{\psi}_k^{\dagger} \tilde{T}_k^{\dagger} C_k \tilde{T}_k \tilde{\psi}_k, \text{ with } \tilde{T}_k = T_k^{-1}$$

We now want to choose \tilde{T}_k such that $\tilde{C}_k = \tilde{T}_k^{\dagger} C_k \tilde{T}_k$ is diagonal. For this, we use the following trick. Consider the matrix $\sigma_z \tilde{C}_k$

$$\sigma^{z}\tilde{C}_{k} = \sigma^{z}\tilde{T}_{k}^{\dagger}\underbrace{\sigma^{z}\sigma^{z}}_{\mathbb{1}}C_{k}\tilde{T}_{k} = \underbrace{\sigma^{z}\tilde{T}_{k}^{\dagger}\sigma^{z}}_{\tilde{T}_{k}^{-1}}\sigma^{z}C_{k}\tilde{T}_{k} = \tilde{T}_{k}^{-1}\sigma^{z}C_{k}\tilde{T}_{k}$$

From this equation one finds that \tilde{T}_k actually diagonalizes the matrix

$$\sigma^z C_k = Jsz \begin{pmatrix} 1 & \gamma_k \\ -\gamma_k & 1 \end{pmatrix}$$

The eigenvalues of this 2×2 matrix are simply $\pm Jsz\sqrt{1-\gamma_k^2}$. After diagonalization, our Hamiltonian takes the form

$$\mathcal{H} = E_0 + \sum_{\vec{k} \in 1.BZ} E_{\vec{k}} \tilde{a}_{\vec{k}}^{\dagger} \tilde{a}_{\vec{k}}, \text{ with } E_{\vec{k}} = Jsz\sqrt{1 - \gamma_{\vec{k}}^2}$$
(90)



Figure 6: Old (red) and new (blue) unit cell and 1.BZ

We observe that $E_{\vec{k}}$ vanishes not only for $\vec{k} = 0$ but also for $\vec{k} = \vec{Q} = \frac{\pi}{a} (\pm 1, \pm 1, \pm 1)^T$. Let us Taylor expand the energy around those points for a three-dimensional cubic lattice (z = 6). In this case we find

$$E_{\vec{k}} \approx 2Js\sqrt{3} \begin{cases} a|\vec{k}| & \text{for } |\vec{k}| \to 0\\ a|\vec{k} - \vec{Q}| & \text{for } |\vec{k} - \vec{Q}| \to 0 \end{cases}$$

For antiferromagnetic spin waves the energy is linear in the momentum k.

Let us add a bit of interpretation on the role of momentum here. In the antiferromagnetic state, the spins alternate from atom to atom, see Fig. 6**a** where a two-dimensional example is shown in the figure. Therefore, one should use a new unit cell (e.g., the tilted blue square in Fig. 6**a**). In 3d, the new basis vectors of the new Bravais lattice (connecting sites with the same spin) can be chosen as $a(1,1,0)^T$, $a(1,0,1)^T$, and $a(0,1,1)^T$ forming an fcc lattice. A corresponding basis in reciprocal space is $\vec{G}_1 = \frac{\pi}{a}(1,1,-1)^T$, $\vec{G}_2 = \frac{\pi}{a}(1,-1,1)^T$, $\vec{G}_3 = \frac{\pi}{a}(-1,1,1)^T$. Thus, the vectors $\vec{Q} = \frac{\pi}{a}(\pm 1, \pm 1, \pm 1)^T$ are actually all part of the reciprocal lattice and correspond to the vector $\vec{k} = 0$ when projected into the correct 1. BZ of the symmetry broken state. The result in Eq. (90) uses the old 1.BZ. It is more physical to project everything the a new 1.BZ based on the new reciprocal lattice vectors \vec{G}_i . If we do this, we obtain two bands which we can, e.g., define by the creation operators

$$a_{1,\vec{k}}^{\dagger} = a_{\vec{k}}^{\dagger}, \qquad a_{2,\vec{k}}^{\dagger} = a_{\vec{k}+\vec{G}_1}^{\dagger}$$

Using that $E_{\vec{k}+\vec{G}_i} = E_{\vec{k}}$, we can write the Hamiltonian Eq. (90) also as

$$\mathcal{H} = E_0 + \sum_{\substack{\vec{k} \in \text{new}\\1.\text{BZ}}} E_{\vec{k}} (\tilde{a}_{1,\vec{k}}^{\dagger} \tilde{a}_{1,\vec{k}} + \tilde{a}_{2,\vec{k}}^{\dagger} \tilde{a}_{2,\vec{k}})$$

In the new 1st BZ, $E_{\vec{k}}$ vanishes only for $\vec{k} = 0$ and we get two $\vec{k} = 0$ Goldstone modes, corresponding to uniform rotations of the magnetization either around the x or y axis.

We close this section by discussing a phenomenon called **quantum fluctuations**. The ground state of the antiferromagnet is a complicated wave function where, e.g., the expectation value $\langle S_i^z \rangle$ is not simply given by $\pm s$. Strictly speaking, a sketch of the groundstate like in Eq. (86) is not correct. Mathematically, we see that from the presence of $a^{\dagger}a^{\dagger}$ terms in the Hamiltonian Eq. (86). This implies, that the bosonic vacuum is not the groundstate, both the number of bosons and the amplitude of the magnetization shows quantum fluctuations.

Let us sketch how one can compute the quantum fluctuations. Consider, for example, the magnetization at site 0.

$$\langle \mathrm{GS}|S_0^z|\mathrm{GS}\rangle = s - \langle \mathrm{GS}|a_0^{\dagger}a_0|\mathrm{GS}\rangle\rangle$$

where $|\text{GS}\rangle$ is the ground state wave function of our antiferromagnet. We can rewrite the last term using the eigenmodes $\tilde{a}_{\vec{k}}^{\dagger}$ of the Hamiltonian and the transformation Eq. (89)

$$\langle \mathrm{GS}|a_0^{\dagger}a_0|\mathrm{GS}\rangle\rangle = \langle \mathrm{GS}|\frac{1}{N}\sum_{k,k'}a_k^{\dagger}a_{k'}|\mathrm{GS}\rangle\rangle = \langle \mathrm{GS}|\frac{1}{N}\sum_{k,k'}(u_k\tilde{a}_k^{\dagger} + v_k\tilde{a}_{-k})(u_{k'}\tilde{a}_{k'} + v_{k'}\tilde{a}_{-k'}^{\dagger})|\mathrm{GS}\rangle$$
$$= \frac{1}{N}\sum_k v_k^2 = \int_{1.\mathrm{BZ}}\frac{d^3k}{(2\pi)^3}\sinh^2(\eta_k) \neq 0 \tag{91}$$

Within our approximation, the groundstate $|GS\rangle$ is the state where no \tilde{a} bosons are present

$$\tilde{a}_{\vec{k}} |\mathrm{GS}\rangle = 0,$$

while $a_{\vec{k}}|\text{GS}\rangle \neq 0$. This relation (and normal-ordering) was used to obtain the last line in Eq. (91). Above, we have not worked out the last integral (as the result is boring in d = 3) but we have shown that our approximation is able to capture quantum fluctuations, at least to leading order in 1/s.

Outlook: The expansion in 1/s is a powerful method to compute analytically the properties of magnets. But does a method designed for large s give reasonable results for s = 1/2, which is often the experimentally relevant case? This is not obvious and one can clearly not expect quantitatively correct results. It turns, however, out that in most systems, including antiferromagnets on cubic lattices, the properties obtained in large s approximations are essentially correct. While the prefactors may not be fully accurate, all qualitative features, like the linear energy-momentum relation at small k, also occur for s = 1/2 as seen both from numerical approaches and experiments. The situation is more tricky (and more interesting) for models with **frustration**. Frustration describes situations where there is no clear candidate for a groundstate, because, e.g., some interactions prefer a ferromagnetic alignment, while other favor antiferromagnetism. In such systems highly exotic non-magnetic states can occur, so-called spin liquids, often desribed by Gauge theories – a topic of the QFT II lecture.

3 Path Integrals

Motivation: In this section, we will discuss a way to do quantum mechanics which - surprisingly - works without wavefunctions and without Hamilton operators. In this chapter, we will start by considering a single quantum-mechanical particle, but the concepts developed here, will later be generalized to quantum field theories.

Let us first recall some absolutely basic concepts underlying quantum mechanics. Arguably, the most fundamental idea underlying quantum mechanics (and quantum field theory) is the concept of a **probability amplitude** $A_i \in \mathbb{C}$ which is a complex number associated to a state. Measurable quantities are probabilities obtained from $|A_i|^2$. A second fundamental concept is the **superposition principle**, probability amplitudes are additive. By adding probability amplitudes, we obtain constructive and destructive interference, which is very different from simply adding probabilities, $|A_1 + A_2|^2 \neq |A_1|^2 + |A_2|^2$.

Let us consider the probability amplitude for a quantum system to move in the time interval t from an initial state $|i\rangle$ to a final state $|f\rangle$, $A_{fi}(t) = \langle f|e^{-i\mathcal{H}t/\hbar}|i\rangle$. Now we use one of the most important tricks in math and physics, we insert an identity operator and rewrite this as a superposition of amplitudes

$$A_{fi}(t) = \langle f|e^{-i\mathcal{H}t/\hbar}|i\rangle = \sum_{c} \langle f|e^{-i\mathcal{H}t/2\hbar}|c\rangle\langle c|e^{-i\mathcal{H}t/2\hbar}|i\rangle = \sum_{c} A_{fc}\left(\frac{t}{2}\right)A_{ci}\left(\frac{t}{2}\right)$$

We have splitted the process $|i\rangle \rightarrow |f\rangle$ into $|i\rangle \rightarrow |c\rangle \rightarrow |f\rangle$, where each subprocess takes half the time and we have to sum over all intermediate states $|c\rangle$.

Instead of splitting the process in two parts, we can also split it into N parts,

$$|i\rangle \rightarrow |c_1\rangle \rightarrow |c_2\rangle \rightarrow \cdots \rightarrow |c_{N-1}\rangle \rightarrow |f\rangle$$

each taking the time $\epsilon = t/N$. We obtain

$$A_{fi}(t) = \sum_{c_1} \sum_{c_2} \cdots \sum_{c_{N-1}} A_{fc_N}(\varepsilon) \cdots A_{c_2c_1}(\varepsilon) A_{c_1i}(\varepsilon)$$
(92)

If we can find (or postulate) a formula for the transition amplitude $A_{c'c}(\varepsilon)$ for very small ε , we can build a new type of quantum mechanics (or quantum field theory) which is based on the evaluation of a lot of sums. No notion of wave functions or operators is needed, we only need transition amplitudes $A_{c_2c_1}(\varepsilon)$ for very short times, $\varepsilon \to 0$. We will precisely derive such formulas in the following.

3.1 Constructing path integrals

Motivation: In the following, we will *derive* path integrals starting from the known formulation of quantum mechanics based on wave functions and operators. Alternatively, we could have started with an educated postulate and then used this to derive quantum mechanics in its usual form.

For a static Hamiltonian, the time evolution of the wave function is given by

$$|\psi(t)\rangle = e^{-i\mathcal{H}t/\hbar}|\psi(0)\rangle$$

The standard Schrödinger equation is formulated for a wave function $\psi(\vec{q}, t) = \langle \vec{q} | \psi(t) \rangle$, where $| \vec{q} \rangle$ is a position-operator eigenstate (we are using the letter \vec{q} for position variables here instead of \vec{x} or \vec{r} used in the previous sections). Inserting $\mathbb{1} = \int d^3 q_i | \vec{q_i} \rangle \langle \vec{q_i} |$, we obtain

$$\psi(\vec{q},t) = \langle \vec{q} | \psi(t) \rangle = \int d^3q \, \langle \vec{q} | e^{-i\mathcal{H}t/\hbar} | \vec{q_i} \rangle \langle \vec{q_i} | \psi(0) \rangle$$

Thus, all information on the time evolution is encoded in the **propagator** or **Green function** defined by

$$U_{\vec{q}_f, t_f; \vec{q}_i, t_i} = \Theta(t_f - t_i) \langle \vec{q}_f | e^{-i\mathcal{H}(t_f - t_i)/\hbar} | \vec{q}_i \rangle$$

The propagator describes the quantum mechanical amplitude to go from an initial position $\vec{q_i}$ at time t_i to a final position $\vec{q_f}$ at time t_f . The θ function has been added to ensure that the final time is always after the initial time (we omit this factor in the following, assuming that $t_f > t_i$)

In the spirit of our introductory remarks, we spit the time evolution operator into N pieces. This step is called **Trotter decomposition**.

$$e^{-i\mathcal{H}t/\hbar} = \left(e^{-i\mathcal{H}\varepsilon/\hbar}\right)^N$$
, where $\varepsilon = \frac{t}{N}$.

Thus we can write

$$U = \langle \vec{q_f} | \mathbb{1}e^{-i\mathcal{H}\varepsilon/\hbar} \mathbb{1} \cdots \mathbb{1}e^{-i\mathcal{H}\varepsilon/\hbar} | \vec{q_i} \rangle$$
(93)

There is some freedom in how to write the **idendity operator** "1". We want to choose it in such a way, that $\mathbb{1}e^{-i\mathcal{H}\varepsilon/\hbar}\mathbb{1}$ is easy to compute in the limit $\varepsilon \to 0$. A very convenient choice (as we will see below) is using both position and momentum eigenstates.

$$\mathbb{1} = \mathbb{1} \cdot \mathbb{1} = \int d^d q_n \, |\vec{q_n}\rangle \langle \vec{q_n}| \, \int \frac{d^d p_n}{(2\pi\hbar)^d} |\vec{p_n}\rangle \langle \vec{p_n}| \tag{94}$$

with

$$\langle \vec{q_n} | \vec{p_n} \rangle = e^{i \vec{p_n} \vec{q_n} / \hbar} \tag{95}$$

Note that we use a different normalization convention here compared to the previous paragraph. The index n is used to replace the nth identity operator in Eq. (93) when we count from right to left. When evaluating Eq. (93), we need for $\mathcal{H} = \frac{\hat{p}^2}{2m} + V(\hat{\vec{q}})$

$$\begin{split} \langle \vec{p}_{n+1} | e^{-i\mathcal{H}\varepsilon/\hbar} | \vec{q}_n \rangle &\approx_{\substack{\uparrow\\\varepsilon \to 0}} \langle \vec{p}_{n+1} | 1 - i\left(\frac{\hat{\vec{p}}^2}{2m} + V(\hat{\vec{q}})\right) \frac{\varepsilon}{\hbar} | \vec{q}_n \rangle + \mathcal{O}(\varepsilon^2) \\ &\approx \langle \vec{p}_{n+1} | \vec{q}_n \rangle \, e^{-i\left(\frac{\vec{p}_{n+1}^2}{2m} + V(\vec{q}_n)\right)\frac{\varepsilon}{\hbar}} + \mathcal{O}(\epsilon^2) \end{split}$$

As we used both momentum and position eigenstates, we could to leading order in ϵ evaluate all $\hat{\vec{p}}$ operators to the left and all $\hat{\vec{q}}$ operators to the right. Importantly, now all operators are gone as \vec{p}_{n+1}^2 and \vec{q}_n are just integration variables, not operators.

Using furthermore

$$\langle \vec{q}_{n+1} | \vec{p}_{n+1} \rangle \langle \vec{p}_{n+1} | \vec{q}_n \rangle = e^{i \vec{p}_{n+1} (\vec{q}_{n+1} - \vec{q}_n)/\hbar}$$

we can also eliminate all wave functions and obtain our first example of a path integral

$$\langle \vec{q}_{f} | e^{-i\mathcal{H}t/\hbar} | \vec{q}_{i} \rangle =$$

$$\lim_{N \to \infty} \int \prod_{n=1}^{N-1} \frac{d^{d}p_{n} d^{d}q_{n}}{(2\pi\hbar)^{d}} \int \frac{d^{d}p_{N}}{(2\pi\hbar)^{d}} e^{-i\frac{\Delta t}{\hbar}\sum_{n=0}^{N-1} \left(\frac{\vec{p}_{n+1}^{2}}{2m} + V(\vec{q}_{n}) - \vec{p}_{n+1}\frac{(\vec{q}_{n+1} - \vec{q}_{n})}{\Delta t}\right) }$$

$$\text{with } \vec{q}_{0} = \vec{q}_{i}, \vec{q}_{N} = \vec{q}_{f}.$$

$$(96)$$

where we use $\Delta t = t/N = \epsilon$. We will write this below in a nicer way, but we can start to find an interpretation of the formula. We start by observing that $\int \frac{d^d q d^d p}{(2\pi\hbar)^d}$ integral over "phase space", the combination of position and momentum space. We will interpret $\vec{q_n} = \vec{q}(t_n)$ (and $\vec{p_n} = \vec{p}(t_n)$) with $t_n = n\epsilon = \frac{n}{N}t$ as the position (and momentum) of a particle at the time t_n . Thus, we define a discretized version of a **path in phase space** $\vec{q}(t')$ (and $\vec{p}(t')$, respectively) with $0 \le t' \le t$, see figure.



Now let us look at the type of integral we are doing. When integrating over $\vec{q_n}$ and $\vec{p_n}$ we effectively sum over all possible (discretized) paths in phase space

$$\int \mathscr{D}(\vec{p}, \vec{q}) \dots \equiv \lim_{N \to \infty} \int d^d q_1 d^d q_2 \cdots d^d q_{N-1} \int d^d p_1 \cdots d^d p_N \frac{1}{(2\pi\hbar)^{Nd}} \dots$$
$$= \underset{\text{in phase space with initial point } \vec{q_i} \text{ and final point } \vec{q_f}}{\sup}$$
(97)

Here, we have to take the limit $N \to \infty$ or, equivalenty, $\epsilon = t/N \to 0$. This is called the **continuum limit**, where consecutive time points come closer and closer to each other. We write the sum over all paths in phase space $\int \mathscr{D}(\vec{p}, \vec{q})$. The integral (and therefore the formalism which we develop in this section) is called **path integral**, or, more precisely, path integral in phase space (as we will soon introduce a somewhat simpler version in real-space only). A tricky question is whether the path $\vec{q}(t')$ and $\vec{p}(t')$ become in any sense continuous in this limit. From the way, how we introduced them, this is far from obvious and will be addressed later. At least one can hope that too large fluctuations of fields are suppressed by terms like $e^{i\vec{p_n}(\vec{q_n}-\vec{q_{n-1}})}$ which rapidly average to zero if, e.g., $\vec{q_n} - \vec{q_{n-1}}$ are far from each other.

For now, let us just express the **warning** that there is no reason to believe here in smooth functions. When we do ordinary integrals over sufficiently smooth ordinary functions, we know that how discretizations are done precisely should not matter. We can **not** expect this to be the case here! Therefore, the path integral is only defined for the discretization scheme of Eq. (96). Thus, we typically cannot use a continuum limit for actual calculation but we should only trust the discretized version.

After all these warnings, it is, nevertheless, nicer to write our results in a continuum version. We can do this, if we never forget, that actual calculations should be done in case of doubt with the discretized version. Thus we write for $N \to \infty$, $\Delta t = t/N \to 0$

$$\begin{aligned} \Delta t \sum_{n=0}^{N-1} &\to \int_0^t dt', \\ \frac{\vec{q}_{n+1} - \vec{q}_n}{\Delta t} &\to \dot{\vec{q}}(t = t_n) \\ \Delta t \sum_n \frac{\vec{p}_n}{2m} + V(\vec{q}_{n-1}) &\to \int_0^t dt' H(\vec{p}(t'), \vec{q}(t')) \end{aligned}$$

where, importantly, $H(\vec{p}, \vec{q})$ is not an operator but simply the **Hamilton function** $H = \frac{\vec{p}^2}{2m} + V(\vec{q}).$

Thus, we will write our path integral over phase space as

$$\langle \vec{q_f} | e^{-i\mathcal{H}t/\hbar} | \vec{q_i} \rangle = \int_{\vec{q}(0) = \vec{q_i}} \mathscr{D}(\vec{p}, \vec{q}) \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, \mathscr{D}(\vec{p}, \vec{q}) \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, \mathcal{D}(\vec{p}, \vec{q}) \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, \mathcal{D}(\vec{p}, \vec{q}) \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, \mathcal{D}(\vec{p}, \vec{q}) \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q_f}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q}, \vec{q}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q}, \vec{q}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q}, \vec{q}}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q}, \vec{q}} \, e^{\frac{i}{\hbar} \int_0^t dt' \, (\vec{p} \cdot \vec{q} - H(\vec{p}, \vec{q}))}_{\vec{q}(t) = \vec{q}, \vec{q}} \, e^{\frac{i}{\hbar} \vec{q}}_{\vec{q}(t) = \vec{q}, \vec{q}}} \, e^{\frac{i}{\hbar} \vec{q}}_{\vec{q}(t) = \vec{q}, \vec{q}} \, e^{\frac{i}{\hbar} \vec{q}}_{\vec{q}(t) = \vec{q}, \vec{q}}} \, e^{\hat$$

where, as repeatedly states, the object on the right-hand side is really defined by Eq. (96).

When we look at Eq. (96), we realize that one can easily do the \vec{p}_n integrals, as they are simple Gaussian integrals

$$\int \frac{d^d p_n}{(2\pi\hbar)^d} e^{-i\frac{\Delta t}{\hbar} \left(\frac{\vec{p}_n^2}{2m} - \vec{p}_n \frac{(\vec{q}_n - \vec{q}_{n-1})}{\Delta t}\right)} = \left(\frac{m}{i\,\Delta t\,2\pi\hbar}\right)^{\frac{d}{2}} e^{\frac{i}{\hbar}\frac{m}{2} \left(\frac{\vec{q}_n - \vec{q}_{n-1}}{\Delta t}\right)^2 \Delta t} \tag{98}$$

We will later derive general formulas on how to do efficiently **Gaussian integrals** in all types of situations. As this is our first one, we will do it slowly. The general strategy contains three steps: (i) First, one completes the square, here

$$\frac{\vec{p}_n^2}{2m} - \vec{p}_n \frac{(\vec{q}_n - \vec{q}_{n-1})}{\Delta t} = \frac{\left(\vec{p}_n - m\frac{(\vec{q}_n - \vec{q}_{n-1})}{\Delta t}\right)^2}{2m} - \frac{m\left(\frac{(\vec{q}_n - \vec{q}_{n-1})}{\Delta t}\right)^2}{2}$$

Second, one shifts the integration variable

$$\vec{p_n} \to \vec{p_n} + m \frac{(\vec{q_n} - \vec{q_{n-1}})}{\Delta t}$$

As a final step, one has to do simple Gaussian integrals of the type

$$\int dx e^{-ax^2} = \sqrt{2\pi/a}$$

There is one issue here: in our case $a = i\alpha$ is purely imaginary but convergence is only guaranteed for Rea > 0. This is an issue also for a numerical evaluation of path integrals which typically are not converging. For now, we assume that there is a tiny real part of $a = i\alpha + \epsilon$ which ensures convergence and we take the limit $\epsilon \to 0$ at the end of the calculation. Collecting all prefactors, we then end up with Eq. (98).

Using this formula, we can integrate over all \vec{p}_n and obtain in d dimensions

$$\left(\frac{m}{i\,\Delta t\,2\pi\hbar}\right)^{\frac{Nd}{2}}e^{\frac{i}{\hbar}\sum_{n}\frac{m}{2}\left(\frac{\vec{q}_{n}-\vec{q}_{n-1}}{\Delta t}\right)^{2}\Delta t}$$

and we can identify the term in the exponent

$$\sum_{n} \frac{m}{2} \left(\frac{\vec{q}_n - \vec{q}_{n-1}}{\Delta t}\right)^2 \Delta t = \int_0^t \frac{1}{2} m \dot{\vec{q}}^2(t') dt'$$

Collecting all results, we obtain a remarkably simple formula for **Feynman's path** integral

$$\langle \vec{q}_f | e^{-i\mathcal{H}t/\hbar} | \vec{q}_i \rangle = \int_{\vec{q}_i}^{\vec{q}_f} \mathscr{D}(\vec{q}) e^{i\mathcal{S}[\vec{q}]/\hbar}$$

$$\mathcal{S}[\vec{q}] = \int_0^t dt' L(\vec{q}, \dot{\vec{q}}) = \int_0^t dt' \left(\frac{1}{2}m\dot{\vec{q}}(t') - V(\vec{q}(t'))\right)$$
(99)

Here the sum over all paths starting at $\vec{q}(0) = \vec{q_i}$ and ending at $\vec{q}(t) = \vec{q_f}$ is defined as

$$\int \mathscr{D}(\vec{q}) \cdots = \lim_{N \to \infty} \left(\frac{Nm}{it2\pi\hbar} \right)^{\frac{Na}{2}} \int \prod_{n=1}^{N-1} d^d q_n \dots$$
(100)

We have obtained a remarkable result. The laws of quantum mechanics are obtained by postulating that each particles takes all possibles paths from an initial point to a final point and we have to weight each path by

$$e^{i\mathcal{S}[\vec{q}]/\hbar}$$

where $S[\vec{q}]$ is the classical action computed along the path. This path is in general unrelated to the classical path, we have instead to consider **all** paths. Planck called \hbar the "quantum of action" (Wirkungsquantum), because of its units, without knowing that the combination $e^{iS[\vec{q}]/\hbar}$ can be viewed as the most central quantity in any quantum theory. This extremely elegant formula, mainly developed by Feynman, holds not only for single-particle quantum system but – as we will see – also in quantum field theory. It is remarkable, that the laws of quantum mechanics can simply obtained using a superposition principle, the classical action and Plancks constant to compute the weight of each path.

Thus, we found a new formulation of quantum mechanics, without using concepts like operators, Schrödinger equations etc. We could just started with Eq. (99) as a postulate and derive ordinary quantum mechanics from there.

As discussed above, the convergence of the path integral is mathematically a bit tricky as one has to be very careful in taking the limit $N \to \infty$ correctly.

What happens in the **classical limit**? In this limit the action is large compared to \hbar ,

 $\Delta S \gg \hbar.$

Consider a few different path which are close by

$$e^{i\mathcal{S}[\vec{q}]/\hbar} + e^{i\mathcal{S}[\vec{q}+\delta\vec{q}]/\hbar} + \dots \approx 0$$

If S varies strongly, then these terms will have the tendency to add up to zero due to destructive interference. But there is an exception: if $S[\vec{q}] \approx S[\vec{q} + \delta \vec{q}]$, then neighboring paths add constructively. Thus, in the classical limit only paths with

 $\delta \mathcal{S} = 0$

can contribute. We have thus **derived Hamilton's variational principle** of classical mechanics starting from a quantum theory. We can also see immediately, that both maxima, minimal and saddlepoints are equally valid classical paths.

3.2 Applications

As the simplest example, let us do the path integral "brute force" for a free particle in d = 1 dimensions. Thus, we have to consider the classical action

$$S = \int_0^t dt' \, \frac{1}{2} m \dot{x}(t')^2 \tag{101}$$

for path obeying the boundary condition $x(0) = q_i, x(t) = q_f$. Here it is convenient to use

$$\vec{x}(t') = (\vec{q}_f - \vec{q}_i)\frac{t'}{t} + \delta \vec{x}(t') \text{ with } \delta \vec{x}(0) = \delta \vec{x}(t) = 0$$
 (102)

This implies

$$S = \frac{1}{2}m\frac{(\vec{q}_f - \vec{q}_i)^2}{t} + \int_0^t dt' \,\frac{1}{2}m\left(\delta \dot{\vec{x}}(t')\right)^2$$

Therefore, we obtain

$$\langle \vec{q_f} | e^{-i\mathcal{H}t/\hbar} | \vec{q_i} \rangle = e^{\frac{i}{\hbar} \frac{1}{2}m \frac{(\vec{q_f} - \vec{q_i})^2}{t}} \langle 0 | e^{-i\mathcal{H}t/\hbar} | 0 \rangle$$

Without any work, we have already computed the dependence of the propagator on coordinates, and the only job left is to calculate the t dependence of the prefactor. to do this, we go back to the discretized version of the path integral

$$\langle 0|e^{-i\mathcal{H}t/\hbar}|0\rangle = \lim_{N \to \infty} \left(\frac{m}{i\Delta t 2\pi\hbar}\right)^{\frac{N}{2}} \int \prod_{i=1}^{N-1} dx_i \, e^{\frac{i}{\hbar}\sum_i \frac{1}{2}m\left(\frac{x_{i+1}-x_i}{\Delta t}\right)^2 \Delta t}.$$

We do these integrals step by step, starting with x_1 using the method discussed below Eq. (98). Collecting all terms with x_1 , we compute

$$\int dx_1 e^{\frac{i}{\hbar} \frac{1}{2}m \left(\frac{(x_2 - x_1)^2}{\Delta t} + \frac{(x_1 - x_0)^2}{\Delta t}\right)} = \int dx_1 e^{\frac{i}{\hbar} \frac{1}{2} \frac{m}{\Delta t} (2x_1^2 + x_0^2 + x_2^2 - 2x_1(x_0 + x_2))}$$
(103)

$$= \left(\frac{2\pi\Delta t\hbar}{2im}\right)^{\frac{1}{2}} e^{\frac{i}{\hbar}\frac{m}{2}\frac{(x_0-x_2)}{2\Delta t}^2} \tag{104}$$

To obtain all terms with x_2 , we have to multiply this result with $\exp(\frac{i}{\hbar}\frac{m}{2}\frac{(x_2-x_3)^2}{\Delta t})$. Integrating over x_2 we get:

$$\left(\frac{2\pi\Delta t\hbar}{2im}\right)^{\frac{1}{2}} \left(\frac{2\pi\frac{2}{3}\Delta t\hbar}{\left(1+\frac{1}{2}\right)im}\right)^{\frac{1}{2}} e^{\frac{i}{\hbar}\frac{m}{2}\frac{\left(x_{0}-x_{3}\right)^{2}}{3\Delta t}^{2}}$$

We repeat this for all other integrations. The term $x_0 - x_N$ vanishes as $x_0 = x_N = 0$. In the denominator, we get the product

$$\left(1+\frac{1}{1}\right)\left(1+\frac{1}{2}\right)\cdots\left(1+\frac{1}{N-1}\right) = \frac{2}{1}\cdot\frac{3}{2}\cdot\frac{4}{3}\cdots\frac{N}{N-1} = N.$$

Collecting all prefactors, we get

$$\langle 0|e^{-i\mathcal{H}t/\hbar}|0\rangle = \left(\frac{m}{2\pi i\hbar t}\right)^{\frac{1}{2}}.$$
(105)

Therefore, the propagator (or Green function) of a free particle is given by:

$$\mathcal{G}(q_f, q_i, t) = \langle q_f | e^{-i\mathcal{H}t/\hbar} | q_i \rangle = \left(\frac{m}{2\pi i\hbar t}\right)^{\frac{1}{2}} e^{\frac{i}{\hbar}m\frac{(\vec{q}_f - \vec{q}_i)^2}{t}}.$$
(106)

The reader is encouraged to calculate the same quantity within the standard formulation of quantum mechanics, which can easily be done here by multiplying with the identity operator $\sum_{k} |k\rangle \langle k|$ and by performing a single Gaussian integral.

3.3 Statistical physics and imaginary time

Motivation: Above, we considered the propagator of a single particle starting from an initial state and ending in a final state. For a many-particle quantum system, like a solid, we typically ask different sets of questions. More often than not, the system which we want to describe (or its initial state) be a at finite temperature. Therefore, we have to investigate how to describe thermal states using path integrals and later functional integrals.

3.3.1 Imaginary time

A quantum system in thermal equilibrium is described by the density matrix

$$\rho = \frac{1}{\mathcal{Z}} e^{-\beta \mathcal{H}}, \qquad \beta = \frac{1}{k_B T}$$

Here Z is the partition sum, which is used above as a normalization constant of the density matrix, but can also be used to compute all thermodynamics properties of a quantum (or classical) system.

$$\mathcal{Z} = \operatorname{Tr}(e^{-\beta \mathcal{H}}) = e^{-\beta F(T)}$$

where F is the free energy of the system at temperature T.

Expectation values of operators are obtained from

$$\langle A \rangle = \operatorname{Tr}(\rho A).$$

Our discussion of thermal equilibrium starts from the observation that $e^{-i\mathcal{H}t/\hbar}$ and $e^{-\beta\mathcal{H}}$ look very similar. It seems, that one has to simply do a replacement

$$t \longrightarrow -i\hbar\beta \tag{107}$$

to transform one term into the other. This is called a **Wick rotation**, which is a rotation in the complex plane. Time is a real-valued variable, which by transformation goes on the imaginary axis.

The concept of **imaginary time** is made more precise by deriving a path integral for the partition sum. Here, we follow exactly the same steps as for the (real-time) propagator, starting from the **Trotter decomposition** and the insertion of suitable identity operators

$$\operatorname{Tr}(e^{-\beta\mathcal{H}}) = \int d^d x \, \langle x | e^{-\beta\mathcal{H}} | x \rangle = \int d^d x \, \langle x | \mathbb{1}e^{-\beta\mathcal{H}/N} \mathbb{1} \cdots \mathbb{1}e^{-\beta\mathcal{H}/N} | x \rangle$$

Compared to the previous derivation of Eq. (96), we only have to replace

$$\begin{array}{cccc} x(t'), \ 0 \leq t' \leq t & \longrightarrow & x(\tau), \ 0 \leq \tau \leq \hbar\beta \\ t' & \longrightarrow & \tau = it' \end{array}$$

Furthermore, because of the trace we have to use periodic boundary conditions

$$x(0) = x(\hbar\beta)$$
 or, within the discrete formulation $x_0 = x_n$ (108)

Let us do the replacement $t \longrightarrow -i\hbar\beta$ directly for the action

$$e^{\frac{i}{\hbar}\int_{0}^{t}dt'\frac{m}{2}\dot{x}^{2}-V(x)} \xrightarrow{t \to -i\hbar\beta} e^{\frac{i}{\hbar}\int_{0}^{-i\hbar\beta}dt'\frac{m}{2}\dot{x}^{2}-V(x)}$$
$$= \begin{vmatrix} t' = -i\tau \\ dt' = -id\tau \\ \partial_{t'} = i\partial_{\tau} \end{vmatrix} = e^{-\frac{1}{\hbar}\int_{0}^{\hbar\beta}d\tau\frac{m}{2}\dot{x}^{2}(\tau)+V(x(\tau))}$$
$$e^{\frac{i}{\hbar}\int_{0}^{t}dt'p(t')\dot{x}(t')} \xrightarrow{e^{-\frac{1}{\hbar}\int_{0}^{\hbar\beta}d\tau(-i)p(\tau)\dot{x}(\tau)}}$$

Thus, we obtain for our two formulations of the path integral (with and without momentum integration)

$$\begin{aligned} \mathcal{Z} &= \int_{x(0)=x(\hbar\beta)} \mathscr{D}(x,p) e^{-\frac{\mathcal{S}[x,p]}{\hbar}} = \int_{x(0)=x(\hbar\beta)} \mathscr{D}(x) e^{-\frac{\mathcal{S}[x]}{\hbar}} \\ \mathcal{S}[x] &= \int_{0}^{\hbar\beta} d\tau \, \frac{m}{2} \dot{x}^{2}(\tau) + V(x(\tau)) \\ \mathcal{S}[x,p] &= \int_{0}^{\hbar\beta} d\tau \, (-i) p \dot{x} + \mathcal{H}(p,x) \end{aligned}$$

The action(s) in imaginary time are called **euclidian action**. This name arises because the minus-sign in the relativistic metric $(\Delta s)^2 = (\Delta x)^2 - (c\Delta t)^2$ changes to a plus sign $(\Delta s)^2 = (\Delta x)^2 + (c\Delta \tau)^2$ and one obtains an euclidian metric. Imaginary time acts similar to an **extra dimension**. At T > 0, the system is, however, finite in this direction $0 \le \tau \le \hbar\beta$, but the size of the system in imaginary-time direction becomes larger and larger for $T \to 0$, or, equivalently, $\beta \to \infty$. In this context it can be useful to realize that the kinectic energy $\frac{m}{2}\dot{x}^2(\tau)$ has the same form as, e.g., the elastic energy $\frac{\kappa}{2}(\partial_x \phi)^2$ of a harmonic chain.

In imaginary time, the action is the sum, not the difference of kinetic and potential energy,

$$\mathcal{S} = \mathcal{S}_{kin} + \mathcal{S}_{pot}.$$

and an other important change are the periodic boundary conditions.

How do we compute expectation values

$$\langle A \rangle = \operatorname{Tr}(\rho A) = \frac{1}{\mathcal{Z}} \int dx' \, \langle x' | e^{-\beta \mathcal{H}} \mathbb{1}A | x' \rangle$$
 (109)

of an operator $A = A(\hat{p}, \hat{x})$?

We have to look up the derivation of the path integral, where we have to compute $\langle p_1 | A | x_0 \rangle$. Thus, we should **normal order** the operators \hat{x} and \hat{p} such that \hat{p} is on the left side. For example, we write

$$A = \hat{x}^2 \hat{p} + \hat{p} \hat{x}^2 = 2\hat{p} \hat{x}^2 + [\hat{x}^2, \hat{p}] = 2\hat{p} \hat{x}^2 + 2i\hbar x \equiv A_N(\hat{p}, \hat{x})$$

Now it is easy to compute $\langle p_1 | A_N(\hat{p}, \hat{x}) | x_0 \rangle = A_N(p_1, x_0) \langle p_1 | x_0 \rangle$. Thus, the momentum variable carries the index 1 instead of 0, which means that it is evaluated at a time-slice $\Delta \tau = \frac{\hbar \beta}{N}$ later than x.

$$\langle A \rangle = \frac{\int \mathscr{D}(p,x) A_N(p(\Delta\tau), x(0)) e^{\frac{-S}{\hbar}}}{\int \mathscr{D}(p,x) e^{-\frac{S}{\hbar}}}, \text{ with } \Delta\tau = \frac{\hbar\beta}{N}$$
(110)

These infinitesimal time slices are sometimes very important as they carry the information on the ordering of operators.

Let us check how the **classical limit** is obtained for the imaginary-time path integral

$$\lim_{\hbar \to 0} e^{-\frac{1}{\hbar} \mathcal{S}[p,x]} = \lim_{\hbar \to 0} e^{-\frac{1}{\hbar} \int_0^{\hbar \beta} \dots} = e^{-\beta \mathcal{H}(p,x)}.$$

As the length of the integral shrinks to zero, we just take the values of the fields at $\tau = 0$. Due to the periodic boundary conditions also $\partial_{\tau} x$ vanishes in this limit.

Above, we have been careful to keep factors of \hbar . For practical calculations, we will often set $\hbar = 1$, effectively measuring time in units of 1/energy.

An important aspect is that functional integrals in imaginary time have much **convergence properties**. While summing over $e^{iS[x]/\hbar}$ implies summing over complex numbers with modulus 1 which is a non-converging sum, a summation over $e^{-S[x]/\hbar}$ is – at least if S[x] is real as in our example – a nicely converging sum of positive numbers. Therefore a frequently used strategy is to evaluate path integrals (and functional integrals) in imaginary time only whenever possible. We will learn later strategies of how to obtain (via analytic continuation) information on what happens for real times using imaginary-time results.

Let us consider the **harmonic oscillator** in one dimension as an example. We therefore consider the eucledian action

$$S = \frac{m}{2} \int_0^{\hbar\beta} d\tau' \, (\dot{x}^2 + \omega^2 x^2), \, x(0) = x(\hbar\beta)$$
(111)

Here it is useful to introduce the Fourier transformation of the $x(\tau)$ taking into account the periodic boundary conditions, $x(0) = x(\hbar\beta)$.

$$x(\tau) = \frac{1}{\hbar\beta} \sum_{n=-\infty}^{\infty} e^{-i\omega_n \tau} x_{\omega_n}, \qquad x_{\omega_n} = \int_0^{\hbar\beta} d\tau \, e^{i\omega_n \tau} x(\tau)$$

Due to the periodic boundary conditions, we have

$$\omega_n = \frac{2\pi}{\hbar\beta} n, n \in \mathbb{Z}, \qquad x_{-\omega_n} = x_{\omega_n}^*.$$

The discrete frequencies, which arise due to our periodic boundary conditions are called **Matsubara frequencies**. If we rewrite the action in the new variables, we obtain

$$\int_{0}^{\hbar\beta} d\tau \, \dot{x}^{2} = \frac{1}{\hbar^{2}\beta^{2}} \sum_{n,m} x_{\omega_{n}} x_{\omega_{m}} (-i\omega_{n}) (-i\omega_{m}) \underbrace{\int_{0}^{\hbar\beta} d\tau \, e^{-i(\omega_{n}+\omega_{m})\tau}}_{\hbar\beta\delta_{n,-m}}$$
$$= \frac{1}{\hbar\beta} \sum_{n} \omega_{n}^{2} x_{\omega_{n}} x_{\omega_{-n}}$$

Adding all terms, we find

$$S = \frac{m}{2} \frac{1}{\hbar\beta} \sum_{n} (\omega^2 + \omega_n^2) |x_{\omega_n}|^2.$$
(112)

In the path integral, we have to integrate over all periodic paths or, equivalently, over all x_{ω_n} . More precisely, we should take into account that $x_{-\omega_n} = x^*_{\omega_n}$ (as $x(\tau) \in \mathbb{R}$). We therefore write

$$x_{\omega_n} = a_n + ib_n, \quad x_{-\omega_n} = a_n - ib_n \quad \text{for } n \ge 1, \qquad x_{\omega_n = 0} = a_0$$

with $a_n, b_n \in \mathbb{R}$. Therefore, the action in terms of these variables takes the form

$$\frac{1}{\hbar}\mathcal{S} = \frac{1}{2} \left(\frac{m}{\beta\hbar^2} \left(\omega^2 a_0^2 + 2\sum_{n=1}^{\infty} (\omega^2 + \omega_0^2)(a_n^2 + b_n^2) \right) \right)$$

and the path integral is now simply an integral over all a_n and b_n . Using standard Gaussian integrals,

$$\frac{\int x^2 e^{-\frac{1}{2}\alpha x^2 dx}}{\int e^{-\frac{1}{2}\alpha x^2} dx} = \frac{1}{\alpha}$$

we obtain

$$\langle a_0^2 \rangle = \left(\frac{m}{\beta\hbar^2}\omega^2\right)^{-1}, \qquad \langle a_n^2 \rangle = \langle b_n^2 \rangle = \left(\frac{m}{\beta\hbar^2}2(\omega^2 + \omega_n^2)\right)^{-1}.$$

Our first goal is to compute the expectation value of the potential energy $\left\langle \frac{1}{2}m\omega^2 x^2 \right\rangle$. We therefore have to compute

$$\langle (x(0))^2 \rangle = \frac{1}{\hbar^2 \beta^2} \sum_{n,m} \langle x_{\omega_n} x_{\omega_m} \rangle = \frac{1}{\hbar^2 \beta^2} \sum_n \langle x_{\omega_n} x_{-\omega_n} \rangle$$
$$= \frac{1}{\hbar^2 \beta^2} \left(\langle a_0^2 \rangle + 2 \sum_{n=1}^{\infty} \langle a_n^2 + b_n^2 \rangle \right) = \frac{1}{\beta} \frac{1}{m} \sum_{n=-\infty}^{\infty} \frac{1}{\omega^2 + \omega_n^2}$$
$$= \frac{\hbar}{2m\omega} \coth\left(\frac{\hbar\omega}{2k_BT}\right)$$

where the last equality can be derived using analytical-continuation techniques which we will treat later.

The expectation value of the potential energy is therefore given by

$$E_{pot} = \frac{1}{2}m\omega^2 \langle x^2 \rangle = \frac{\hbar\omega}{4} \coth\left(\frac{\hbar\omega}{2k_BT}\right) \approx \begin{cases} \frac{1}{2}k_BT & \text{for } k_BT \gg \hbar\omega\\ \hbar\omega/4 & \text{for } k_BT \ll \hbar\omega \end{cases}$$

In the classical regime, $k_B T \gg \hbar \omega$, we recover the classical result $\frac{1}{2}k_B T$, while for T = 0 we obtain that half of the ground-state energy $\hbar \omega/2$ is potential energy, the other half therefore has to be the kinetic energy.

3.3.2 Infinities in path integrals

As a next step, we want to calculate $\frac{1}{2}m\langle \dot{x}^2\rangle$. This is simple because we have calculated above already $\langle x^2 \rangle$ and the formulas are identical up to a factor ω_n^2 arising from the time-derivatives,

$$\frac{1}{2}m\langle \dot{x}^2 \rangle = \frac{1}{2}m\frac{1}{\beta^2}\sum_n \omega_n^2 \langle x_{\omega_n} x_{\omega_{-n}} \rangle = \frac{1}{2}\frac{1}{\beta}\sum_n \frac{\omega_n^2}{\omega^2 + \omega_n^2} = \infty.$$
(113)

This is clearly a divergent sum as each term approaches 1 at large ω_n . This has several consequences: First, the quantity which we calculated **cannot** be the kinetic energy. Second, the result tells us, that the paths which we sum over in the functional integral **cannot** be smooth and differentiable, as otherwise the average $\langle \dot{x}^2 \rangle$ would be finite.

To compute the expectation value of the kinetic energy, $\left\langle \frac{p^2}{2m} \right\rangle$, we have to be more careful. Thus, we go back to the definition of the path integral. Obviously, we have to use the version where the integral over momenta has not been done,

$$\left\langle \frac{p^2}{2m} \right\rangle = \frac{\int \mathscr{D}(p,x) \, e^{-\mathcal{S}} \frac{p_1^2}{2m}}{\int \mathscr{D}(p,x) \, e^{-\mathcal{S}}}$$

Here, we use the proper discretized definition of the path integral. The integral over p_1 has the form (with $\varepsilon = \frac{\beta}{N}$)

$$\int \frac{dp_1}{2\pi\hbar} \frac{p_1^2}{2m} e^{-\varepsilon \frac{p_1^2}{2m} + ip_1(x_1 - x_0)}$$

Shifting the integration variable, $p_1 \to p_1 + \frac{im(x_1 - x_0)}{\varepsilon}$, the integral takes the form

$$\int \frac{dp_1}{2\pi\hbar} \frac{\left(p_1 + \frac{im}{\varepsilon}(x_1 - x_0)\right)^2}{2m} e^{-\varepsilon \frac{p_1^2}{2m} - \varepsilon \left(\frac{x_1 - x_0}{\varepsilon}\right)^2} \tag{114}$$

When multiplying out $(p_1 + \frac{im}{\varepsilon}(x_1 - x_0))^2$, we obtain three terms. The mixed term linear in p_1 vanishes after integration. The p_1^2 gives

$$\frac{\int \frac{dp_1}{2\pi\hbar} \frac{p_1^2}{2m} e^{-\varepsilon \frac{p_1^2}{2m}}}{\int \frac{dp_1}{2\pi\hbar} e^{-\varepsilon \frac{p_1^2}{2m}}} = \frac{1}{2\varepsilon}, \ \varepsilon = \frac{\beta}{N}$$

This contribution is clearly divergent for $N \to \infty$, but we have not yet taken into account the term proportional to $\left(\frac{im}{\varepsilon}(x_1 - x_0)\right)^2 = -(m\dot{x})^2$.

Adding both terms, we find

$$\left\langle \frac{p^2}{2m} \right\rangle = \frac{1}{2\varepsilon} - \frac{1}{2}m\langle \dot{x}^2 \rangle = \frac{1}{2} \left(\frac{N}{\beta} - \frac{1}{\beta} \sum_n \frac{\omega_n^2}{\omega^2 + \omega_n^2} \right)$$
$$= \frac{1}{2} \left(\frac{1}{\beta} \sum_n \left(1 - \frac{\omega_n^2}{\omega^2 + \omega_n^2} \right) \right)$$
$$= \frac{1}{2} \left(\frac{1}{\beta} \sum_n \frac{\omega^2}{\omega^2 + \omega_n^2} \right) = \left\langle \frac{1}{2}m\omega^2 x^2 \right\rangle$$

We obtain that the expectation value of kinetic and potential energy of the harmonic oscillator are equal and, importantly, both are finite. The total energy is therefore given by

$$\langle H \rangle = 2 \times \frac{1}{2} m \omega^2 \langle x^2 \rangle = \frac{\hbar \omega}{2} \coth\left(\frac{\hbar \omega}{2k_B T}\right) \approx \begin{cases} k_B T & \text{for } k_B T \gg \hbar \omega\\ \hbar \omega/2 & \text{for } k_B T \ll \hbar \omega \end{cases}$$
and we obtain for $T \to 0$ the ground state energy and for high temperatures the correct classical result.

The main lesson, which we learned from this exercise – worth to remember – is that at least sometimes it is **necessary to use the discrete definition of the path integral** to obtain the correct result. Furthermore, when summing up over all paths, one cannot expect that the paths are all smooth as $\langle \dot{x}^2 \rangle = \infty$.

4 Functional Integrals

4.1 Functional Integrals and Second Quantization

4.1.1 Coherent states: Bosons

Motivation: Our next step is to generalize the concept of a path integral from single-particle physics to second quantization and thus many-particle physics. In the derivation of the path-integral, a central step was the introduction of suitably chosen identity operator. In the case of single-particle quantum mechanics, we used eigenstates of the position and of the momentum operators, $|\vec{x}\rangle$ and $|\vec{p}\rangle$.

Our goal has to find states $|\vec{\alpha}\rangle$ which make it easy to evaluate $\langle \vec{\alpha}' | H | \vec{\alpha} \rangle$, provided that H is written in the language of second-quantization. This motivates us, to search for **eigenstates of the annihilation operators**, a_i or $a_{\vec{k}}$, also called **coherent state**.

For a **single bosonic mode** the coherent state is defined by

$$a|\alpha\rangle = \alpha|\alpha\rangle$$

Note that a is **not** a hermitian operator. Therefore, eigenvalues will in general **not** be real and eigenvectors will **not** be orthogonal, as we will show below.

To find the coherent state, we expand $|\alpha\rangle$ in particle-number eigenstates

$$|\alpha\rangle = \sum_{n} c_n (a^{\dagger})^n |0\rangle$$

To compute $a|\alpha\rangle$, we move the annihilation operator to the right (normal ordering) using repeatedly the commutation relations $[a, a^{\dagger}] = 1$.

Thus, we find

$$a|\alpha\rangle = \sum_{n} c_n \left(n(a^{\dagger})^{n-1} + (a^{\dagger})^n a \right) |0\rangle = \sum_{n} c_n n(a^{\dagger})^{n-1} |0\rangle$$

By demanding $a|\alpha\rangle = \alpha |\alpha\rangle$, we obtain the condition

$$c_{n+1}(n+1) = \alpha c_n \implies c_n = \alpha^n \frac{C_0}{n!}$$

Therefore, we obtain the so-called **bosonic coherent state**

 $|\alpha\rangle = e^{\alpha a^{\dagger}}|0\rangle$

By construction, the coherent state is an eigenstate of the annihiliation operator, $a|\alpha\rangle = \alpha |\alpha\rangle$. In contrast, applying a^{\dagger} is equivalent to taking the derivative with respect to α ,

$$a^{\dagger}|\alpha\rangle = a^{\dagger}e^{\alpha a^{\dagger}}|0\rangle = \frac{\partial}{\partial\alpha}e^{\alpha a^{\dagger}}|0\rangle = \frac{\partial}{\partial\alpha}|\alpha\rangle.$$

As a next step, we have to compute the overlap to two coherent states. Using $(a^{\dagger})^{n}|0\rangle = \sqrt{n!}|n\rangle$, we obtain

$$\langle \alpha' | \alpha \rangle = \sum_{n,n'} \langle n' | \frac{(\alpha'^*)^{n'}}{n'!} \sqrt{n'!} \frac{(\alpha)^n}{n!} \sqrt{n!} | n \rangle = \sum_n \frac{(\alpha'^* \alpha)^n}{n!} = e^{\alpha'^* \alpha}.$$

Thus, coherent states are *not* orthonormal to each other.

Our next major goal will be to find a formula for the identity operator $\mathbb{1} = \sum_{n} |n\rangle \langle n|$. To find this operator, we start from

$$|\alpha\rangle\langle\alpha| = \sum_{n_1,n_2} \frac{(\alpha^*)^{n_1}}{\sqrt{n_1!}} \frac{\alpha^{n_2}}{\sqrt{n_2!}} |n_2\rangle\langle n_1|.$$

Next, we use the integral identity

$$\int \frac{d\alpha^* d\alpha}{2\pi i} e^{-\alpha^* \alpha} (\alpha^*)^{n_1} \alpha^{n_2} = \delta_{n_1, n_2} n_1!$$

We will discuss such complex integrals (and how they are computed) in detail later. For the moment, we just note that one can treat α and α^* as independent variables. Furthermore, when needed one can always write $\alpha = \alpha_1 + i\alpha_2$ with $\alpha_1, \alpha_2 \in \mathbb{R}$ and integrate over the two real variables α_1 and α_2 instead using

$$\int \frac{d\alpha^* d\alpha}{2\pi i} f(\alpha^*, \alpha) \underset{\substack{\alpha = \alpha_1 + i\alpha_2 \\ \alpha^* = \alpha_1 - i\alpha_2}}{\uparrow} \int \frac{d\alpha_1 d\alpha_2}{\pi} f(\alpha_1 - i\alpha_2, \alpha_1 + i\alpha_2).$$

where we used that the Jacobi determinant for the variable transformation from α^* and α to α_1 and α_2 is given by $\begin{vmatrix} 1 & -i \\ 1 & i \end{vmatrix} = 2i$. Combining these formulas, we find

$$\int \frac{d\alpha^* d\alpha}{2\pi i} e^{-\alpha^* \alpha} |\alpha\rangle \langle \alpha| = \sum_{n_1, n_2} \frac{1}{\sqrt{n_1!}} \frac{1}{\sqrt{n_2!}} \delta_{n_1, n_2}(n_1!) |n_2\rangle \langle n_1| = \sum_{n_1} |n_1\rangle \langle n_1| = \mathbb{1}.$$

This formula was our goal: we have found a way to write the identity operator in terms of eigenstates of the annihilation operator!

In a final step, we have to generalize the formulas from one bosonic mode with annihilation operator a to many bosonic modes with annihilation operators a_i . As the a_i commute with each other, we can simply define

$$|\vec{\alpha}\rangle = e^{\sum_{i} \alpha_{i} a_{i}^{\dagger}} |0\rangle \quad \Rightarrow \quad a_{i} |\vec{\alpha}\rangle = \alpha_{i} |\vec{\alpha}\rangle$$

With these bosonic coherent states, the identity operator is written as

$$\mathbb{1} = \int \prod_{j} \frac{d\alpha_{j}^{*} d\alpha_{j}}{2\pi i} e^{-\sum_{j} \alpha_{j}^{*} \alpha_{j}} |\vec{\alpha}\rangle \langle \vec{\alpha} |.$$

We will see later that this formula will be the key to derive the analog of a path integral for bosonic systems!

Furthermore, one can compute the trace of an arbitrary operator A using $\text{Tr}(A) = \text{Tr}(A\mathbb{1}) = \sum_{\vec{n}} \langle \vec{n} | A \mathbb{1} | \vec{n} \rangle$ to obtain

$$\operatorname{Tr}(A) = \int \prod_{j} \frac{d\alpha_{j}^{*} d\alpha_{j}}{2\pi i} e^{-\sum_{j} \alpha_{j}^{*} \alpha_{j}} \langle \vec{\alpha} | A | \vec{\alpha} \rangle.$$

Outlook: We will mainly use in the lecture coherent states as a 'trick' to obtain the analog of the path integral for a many-particle system. Nevertheless, coherent states have also many important physical applications. For example, the light of a laser is very well described by a coherent state. Similarly, a coherent state can also be used as an approximate wavefunction for weakly interacting Bose-Einstein condensate of bosons.

4.1.2 Coherent states: Fermions

After the successful construction of eigenstates of the bosonic annihiliation operator, we want to use the same construction for fermions. We would like to define eigenstates of **fermionic annhilation operators**

$$c_i |\vec{\eta}\rangle = \eta_i |\vec{\eta}\rangle$$

Here, we face a major problem. Fermions anticommute, $c_i c_j = -c_j c_i$. Using this operator identity we find

$$c_i c_j |\vec{\eta}\rangle = \eta_i \eta_j |\vec{\eta}\rangle = -\eta_j \eta_i |\vec{\eta}\rangle.$$

This equation proves that fermionic annihilation operators cannot have complexvalued eigenfunctions.

At this point, one might be tempted to just give up. Instead, we take a more courageous step: we define a new type of 'numbers', which are different from real or complex numbers because they anticommute. They have the following properties

| $\eta_i \eta_j = -\eta_j \eta_i,$ | |
|--|--|
| $\eta_i \alpha = \alpha \eta_i$ | if $\alpha \in \mathbb{C}$ or $\alpha = $ bosonic operator |
| $\eta_i c^\dagger = -c^\dagger \eta_i$ | if c^{\dagger} fermionic |

These objects are called **Graßmann** variables or Graßmann numbers.

Formally one can study the associative Graßmann algebra of polynomials with complex prefactors $c_n, c_{n,m}, c_{n,m,p}, \dots \in \mathbb{C}$ of the type

$$c^{0} + \sum_{i=1}^{n} c_{i}\eta_{i} + \sum_{i,j=1}^{n} c_{ij}\eta_{i}\eta_{j} + \cdots$$

n anticommuting Graßmann variables can be written as suitably chosen anticommuting $2^n \times 2^n$ matrices, but we will not use such a representation. Instead, we will treat them just as a very special type of anticommuting numbers.

From the defining property $\eta_i \eta_j = -\eta_j \eta_i$, it follows directly, that $\eta_i \eta_i = \eta_i^2 = 0$. This property will become important, when we define functions of Graßmann variables. Functions of Graßmann variables are defined by their Taylor expansion

$$f(\eta) = f(0) + \eta f'(0).$$

The Taylor expansion stops after the first term, because $\eta^2 = 0$. A function of two Graßmann variables therefore always takes the form

$$f(\eta_1, \eta_2) = c_{00} + c_{10}\eta_1 + c_{01}\eta_2 + c_{11}\eta_1\eta_2,$$

where the c are the corresponding Taylor expansion coefficients. For k variables the corresponding formula reads

$$f(\eta_1, \cdots, \eta_k) = \sum_{n=0}^k \left(\sum_{i_1, i_2, \cdots, i_n=1}^k \frac{\partial^n f}{\partial \eta_{i_1} \partial \eta_{i_2} \cdots \partial \eta_{i_n}} \bigg|_{\eta=0} \eta_{i_n} \eta_{i_{n-1}} \cdots \eta_{i_1} \right).$$

We can now also **define** in a purely formal way a derivative operator using

$$\begin{aligned} \partial_{\eta_i}(\eta_j) &= \delta_{i,j}, \\ \partial_{\eta_i}\partial_{\eta_j} &= -\partial_{\eta_j}\partial_{\eta_i}, \\ \partial_{\eta_i}(\eta_j \cdot) &= -\eta_j \partial_{\eta_i}(\cdot), \ i \neq j \end{aligned}$$

Using these rules, one finds, for example, that $\partial_{\eta_i}(\eta_j \eta_i) = -\eta_j$. For our purpose, it will be more important to **define an integral over Graßmann numbers**.

How can we define an integral over an object which is not even a number? What could be a useful and reasonable definition of an integral? We require that the value of an integral should be a complex number. Consider, for example, the integral $\int d\eta_i 1$. We want that (i) the value of the integral is a number, and (ii) that the object is anticommuting, because $d\eta_i$ should be an anticommuting object. This leaves only one option: the integral can only be zero, because zero is the only anticommuting number, $0\alpha = -\alpha 0 = 0$. Next, consider $\int d\eta_i \eta_i$. This is the product of two anticommuting objects $d\eta_i$ (whatever that is) and η_i . We take now the freedom to define the value of this integral to be 1. We will have to show later, that this is a useful definition. But for the moment, we **define integration over Graßmann numbers** by

$$\int d\eta_i = 0$$
$$\int d\eta_i \, \eta_i = 1$$

Let us practice this definition by computing a few integrals.

$$\int d\eta \, f(\eta) = \int d\eta \, (f(0) + f'(0)\eta) = f'(0)$$

This was suprisingly easy. We were able to compute the integral over an arbitrary function f by using that the Taylor expansion stops after the first term (as $\eta^2 = 0$). Using that $d\eta_i$ are anticommuting objects, $d\eta_1\eta_2 = -\eta_2 d\eta_1$, we can also compute more complex integrals

$$\int d\eta_1 d\eta_2 \left(c_{00} + c_{10}\eta_1 + c_{01}\eta_2 + c_{11}\eta_1\eta_2 \right) = -c_{11}$$

We will come back to integrations soon, but first we **define the fermionic coher**ent state

$$|\eta\rangle = e^{-\sum_i \eta_i c_i^{\dagger}} |0\rangle$$

where η_i are Graßmann numbers.

Let us check, whether these are, indeed, eigenstates of the fermionic annihilation operator. Consider the case of a single fermionic mode and let us first compute

$$ce^{-\eta c^{\dagger}}|0\rangle = c(1-\eta c^{\dagger})|0\rangle = -c\eta c^{\dagger}|0\rangle = \eta cc^{\dagger}|0\rangle = \eta|0\rangle$$

This, we should compare to

$$\eta e^{-\eta c^{\dagger}} |0\rangle = \eta (1 - \eta c^{\dagger}) |0\rangle = \eta |0\rangle$$

As the two equations give the same result, we have shown that

$$c|\eta\rangle = \eta|\eta\rangle$$

Repeating this exercise for many fermionic modes, we can show that

$$c_i |\eta\rangle = \eta_i |\eta\rangle$$

Let us also define a corresponding 'bra' states

$$\langle \eta | = \langle 0 | e^{-\sum_i c_i \bar{\eta}_i} = \langle 0 | e^{\sum_i \bar{\eta}_i c_i}$$

Here, we will interpret $\bar{\eta}_i$ and η_i as completely **independent** Graßmann numbers. The bar in $\bar{\eta}_i$ does not have the meaning of complex conjugation here

Here are a few useful properties.

$$[\eta_i c_i^{\dagger}, \eta_j c_j^{\dagger}] = 0 \implies |\eta\rangle = \prod_i (1 - \eta_i c_i^{\dagger}) |0\rangle.$$

Multiplication with c_i^{\dagger} can be written as a derivative operator

$$c_i^{\dagger} |\eta\rangle = -\frac{\partial}{\partial \eta_i} |\eta
angle$$

The overlap of two coherent states is given by

$$\langle \eta' | \eta \rangle = e^{\sum_i \bar{\eta}'_i \eta_i}$$

This can be shown using

$$\langle \eta' | \eta \rangle = \langle 0 | \prod_{i,j} (1 + \bar{\eta}'_i c_i) (1 - \eta_j c_j^{\dagger}) | 0 \rangle = \prod_i (1 + \bar{\eta}'_i \eta_i) = e^{\sum_i \bar{\eta}'_i \eta_i},$$

where we used the anticommutation relations and $\langle 0|c_i^{\dagger}=c_i|0\rangle=0$

We will also need a formula for the identity operator. It looks identical to the bosonic case

$$\mathbb{1} = \int \prod_{\alpha} d\bar{\eta}_{\alpha} d\eta_{\alpha} \, e^{-\sum_{i} \bar{\eta}_{i} \eta_{i}} |\eta\rangle \langle \eta|$$

Let us check this formula for a single fermionic mode

$$\int d\bar{\eta} d\eta \left(1 - \bar{\eta}\eta\right) (1 - \eta c^{\dagger}) |0\rangle \langle 0| (1 + \bar{\eta}c) = \int d\bar{\eta} d\eta \left((-\bar{\eta}\eta) |0\rangle \langle 0| - \eta(-\bar{\eta})c^{\dagger} |0\rangle \langle 0|c\right)$$
$$= |0\rangle \langle 0| + |1\rangle \langle 1|$$

Let us also compute the overlap of a Fock state $|\psi_i\rangle$ with a fermionic coherent state

$$\begin{aligned} |\psi_i\rangle &= |i_1, i_2, \cdots, i_k\rangle \equiv c_{i_1}^{\dagger} c_{i_2}^{\dagger} \cdots c_{i_k}^{\dagger} |0\rangle \\ \implies \langle i_1, i_2, \cdots, i_k |\eta\rangle &= \langle 0 | c_{i_k} \cdots c_{i_1} |\eta\rangle = \eta_{i_k} \cdots \eta_{i_2} \eta_{i_1} \end{aligned}$$

Our next goal, is to compute the trace of an operator, which we will need, e.g., to be able to obtain the partition sum. The trace is defined by a sum over all Fock states and we use our identity operator to rewrite it in terms of fermionic coherent states

$$\operatorname{Tr} A = \sum_{i} \langle \psi_{i} | A | \psi_{i} \rangle = \sum_{i} \langle \psi_{i} | \int \prod_{j} d\bar{\eta}_{j} d\eta_{j}, e^{-\sum_{k} \bar{\eta}_{k} \eta_{k}} | \eta \rangle \langle \eta | A | \psi_{i} \rangle$$
$$= \sum_{i} \int \prod_{j} d\bar{\eta}_{j} d\eta_{j}, e^{-\sum_{k} \bar{\eta}_{k} \eta_{k}} \langle -\eta | A | \psi_{i} \rangle \langle \psi_{i} | \eta \rangle$$

Here, we assumed that A is a bosonic operator, containing an **even number** of fermionic creation and annihilation operators. When going from the first to the second line, we had to exchange $\langle \psi_i | \eta \rangle$ and $\langle \eta | A | \psi_i \rangle$. This produced a minus sign for each η_i , which we encoded in $\langle -\eta | A | \psi_i \rangle$.

Finally, using $\sum_i |\psi_i\rangle \langle \psi_i| = 1$ we obtain the formula for the trace of an operator

$$\operatorname{Tr} A = \int \prod_{j} d\bar{\eta}_{j} d\eta_{j}, e^{-\sum_{k} \bar{\eta}_{k} \eta_{k}} \langle -\eta | A | \eta \rangle.$$

Again, this formula looks very similar to the fermionic case. A major difference is, however, the minus sign in $\langle -\eta | A | \psi_i \rangle$ which will have far-reaching physical consequences as we will see later.

4.1.3 Functional integrals for Z

We are now ready to derive the formula for functional integrals, the generalization of path integrals to quantum field theories.

We start by first normal-ordering our Hamiltonian by putting all creation operators a^{\dagger} to the left, and all *a* to right. The normal-ordered *H* takes, for example, the following form

$$\tilde{\mathcal{H}}(a^{\dagger},a) = \mathcal{H} - \mu \mathcal{N} = \sum_{i,j} t_{ij} a_i^{\dagger} a_j - \mu \sum_i a_i^{\dagger} a_i + \sum_{i,j,k,l} V_{ijkl} a_i^{\dagger} a_j^{\dagger} a_k a_l$$
(115)

where we have already added a chemical potential for later convenience.

In the following, we want to derive formulas which are valid both for fermions and for bosons. Therefore we need a few definitions:

$$\xi = \begin{cases} -1 & \text{for fermions} \\ 1 & \text{for bosons} \end{cases}, \tag{116}$$

$$\psi = (\psi_1, \psi_2, \cdots) = \begin{cases} \text{Graßmann valued for fermions} \\ \text{Complex valued for bosons} \end{cases},$$
(117)

$$d(\bar{\psi},\psi) = \begin{cases} \prod_i d\bar{\psi}_i d\psi_i & \text{for fermions } |\psi\rangle \text{ coherent state} \\ \prod_j \frac{d\bar{\psi}_j d\psi_j}{2\pi i} = \prod_i \frac{d\Re(\psi)d\Im(\psi)}{\pi} & \text{for bosons } (\bar{\psi}=\psi^*) \end{cases}, \quad (118)$$

Now, we are ready to derive the function integral for the partition sum (using $\Delta \tau = \frac{\beta}{N}$) by using a Trotter decomposition where we insert the identity operators of the previous two sections

$$\begin{aligned} \mathcal{Z} &= \operatorname{Tr}\left(e^{-\beta\tilde{\mathcal{H}}(a^{\dagger},a)}\right) \\ &= \int d(\bar{\psi}_{0},\psi_{0}) e^{-\sum_{i}\bar{\psi}_{0_{i}}\psi_{0_{i}}} \langle \xi\psi_{0}|e^{-\beta\tilde{\mathcal{H}}/N} \mathbb{1}e^{-\beta\tilde{\mathcal{H}}/N} \mathbb{1}\cdots \mathbb{1}e^{-\beta\tilde{\mathcal{H}}/N}|\psi_{0}\rangle \\ &= \int \prod_{k=0}^{N-1} d(\bar{\psi}_{k},\psi_{k}) e^{-\sum_{i}\bar{\psi}_{k_{i}}\psi_{k_{i}}} \langle \xi\psi_{0}|e^{-\Delta\tau\tilde{\mathcal{H}}}|\psi_{N-1}\rangle \langle \psi_{N-1}|e^{-\Delta\tau\tilde{\mathcal{H}}}\cdots |\psi_{1}\rangle \langle \psi_{1}|e^{-\Delta\tau\tilde{\mathcal{H}}}|\psi_{0}\rangle \end{aligned}$$

Note the factor ξ in $\langle \xi \psi_0 |$ which takes into account the extra minus sign which we found when deriving the trace formula for fermions.

For small $\Delta \tau$, we can now easily evaluate the relevant overlaps using $e^{-\Delta \tau \tilde{\mathcal{H}}(a^{\dagger},a)} \approx 1 - \Delta \tau \tilde{\mathcal{H}}(a^{\dagger},a) + O((\Delta \tau)^2)$. We then put the annihilation operators to the left and

creation operators to the right using that $|\psi_m\rangle$ are eigenstates of the annihilation operator. Thus we obtain

$$\langle \psi_{m+1} | e^{-\Delta \tau \tilde{\mathcal{H}}(a^{\dagger}, a)} | \psi_m \rangle = e^{\sum_i \bar{\psi}_{m+1_i} \psi_{m_i} - \Delta \tau \tilde{\mathcal{H}}(\bar{\psi}_{m+1}, \psi_m)} + \mathcal{O}(\Delta \tau^2)$$
(119)

where the first term arises from the overlap $\langle \psi_{m+1} | \psi_m \rangle$.

D

Combining the formulas, we obtain our central result, the **functional integral in imaginary time** it its discretized version

$$\mathcal{Z} = \lim_{N \to \infty} \int_{\substack{\bar{\psi}_N = \xi \bar{\psi}_0 \\ \psi_N = \xi \psi_0}} \mathcal{D}(\bar{\psi}, \psi) e^{-\mathcal{S}[\bar{\psi}, \psi]}$$
$$\mathcal{S} = \Delta \tau \sum_{n=0}^{N-1} \left(\sum_i \left(\frac{\bar{\psi}_{n,i} - \bar{\psi}_{n+1,i}}{\Delta \tau} \right) \psi_{n,i} + \tilde{\mathcal{H}}(\bar{\psi}_{n+1}, \psi_n) \right)$$
$$\mathcal{I}(\bar{\psi}, \psi) = \prod_{n=0}^{N-1} d(\bar{\psi}_n, \psi_n)$$

The formula is valid both for bosons and fermions but there are important differences. First, one integrates over complex fields in the bosonic case and over Graßmann fields for fermions. Furthermore, there is the harmless-looking but decisive difference in the boundary condition, $\psi_N = \xi \psi_0$ with $\xi = -1$ for fermions and $\xi = 1$ for bosons.

As in the case of the path integral, the continuum limit is ill-defined but is, nevertheless, frequently used to write results in a more compact way. In this case, sums become integrals and we can write

$$\begin{aligned} \mathcal{Z} &= \int_{\psi(\beta)=\xi\psi(0)}^{\beta} \mathscr{D}(\bar{\psi},\psi) \, e^{-\mathcal{S}[\bar{\psi},\psi]} \\ \mathcal{S} &= \int_{0}^{\beta} d\tau \, \left(\sum_{i} \underbrace{\bar{\psi}_{i}(\tau) \partial_{\tau} \psi_{i}(\tau)}_{\text{"Berry phase"}} + \tilde{\mathcal{H}}(\bar{\psi},\psi) \right) \\ &= \int_{0}^{\beta} d\tau \, \sum_{i,j} \bar{\psi}_{i} \left((\partial_{\tau} - \mu) \delta_{ij} + t_{ij} \right) \psi_{j} + \sum_{i,j,k,l} V_{ijkl} \bar{\psi}_{i}(\tau) \bar{\psi}_{j}(\tau) \psi_{k}(0) \psi_{l}(0) \end{aligned}$$

4.2 Gaussian integrals and Wick's theorem

Motivation: In the absence of interaction, the action is quadratic in the field. Thus, we will have to perform a variant of a Gaussian integral. A major goal in this lecture will be how to do perturbation theory in the interaction strength. How this goes is ruled by the math of Gaussian integrals, which we will study in this section.

We begin with the well-known Gaussian integral over a single real variable

$$\int dx \, e^{-\frac{ax^2}{2}} = \sqrt{\frac{2\pi}{a}}, \quad \text{for } \operatorname{Re}(a) > 0,$$

where the last condition is needed for convergence.

Next, we want to do integrals over complex numbers, z = x + iy. Instead of integrating over x and y, one can also integrate over z = x + iy and $\bar{z} = x - iy$ using the Jacobian

$$\det \begin{pmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{pmatrix} = 2i$$
$$\frac{d\bar{z}dz}{2i} = dxdy$$
(120)

and therefore

Therefore, the Gaussian integral over a single complex number is given by

$$\int \frac{d\bar{z}dz}{2\pi i} e^{-a\bar{z}z} = \int \frac{dxdy}{\pi} e^{-a(x^2+y^2)} = \frac{1}{a}, \quad \text{for } \operatorname{Re}(a) > 0,$$
(121)

The corresponding Graßmann integral is computed from the Taylor expansion

$$\int d\bar{\eta} d\eta \, e^{-a\bar{\eta}\eta} = \int d\bar{\eta} d\eta \left(1 - a\bar{\eta}\eta\right) = a$$

Here, the integral is always converging, independent of the sign of a.

Next, we will need the matrix version. To ensure convergence, let us assume that the $n \times n$ matrix A has the property that $\operatorname{Re}(A) = \frac{1}{2}(A + A^{\dagger})$ is positive definite (i.e., all eigenvalues are positive and not zero). Then convergence is guaranteed in the integral below. Here, we also add linear terms, using the trick 'completing the square'

$$\int \prod_{i} \frac{dz_{i}^{*} dz_{i}}{2\pi i} e^{-z^{\dagger} A z + w^{\dagger} z + z^{\dagger} \tilde{w}} = \begin{vmatrix} z \to z + A^{-1} \tilde{w} \\ z^{\dagger} \to z^{\dagger} + w^{\dagger} A^{-1} \end{vmatrix}$$
$$= \int \prod_{i} \frac{dz_{i}^{*} dz_{i}}{2\pi i} e^{-z^{\dagger} A z + w^{\dagger} A^{-1} \tilde{w}}$$
$$= \det(A)^{-1} e^{w^{\dagger} A^{-1} \tilde{w}}$$
(122)

One can most easily show that the integral 1/det(A) if A is diagonalizable: in this case one first diagonalizes A, then the integral splits into product of integrals over a single complex variable, Eq. (121), and the determinant is the product of eigenvalues. The formula is, however, also valid if A cannot be diagonalized.

The corresponding formula for Graßmann variables is very similar. The only difference is that here convergence is not an issue, A can be an arbitrary complex matrix with an inverse. In this case, we obtain

$$\int \prod_{i} d\bar{\eta}_{i} d\eta_{i} e^{-\bar{\eta}^{\mathsf{T}} A \eta + \bar{v}^{\mathsf{T}} \eta + \bar{\eta}^{\mathsf{T}} v} = \det(A) e^{\bar{v}^{\mathsf{T}} A^{-1} v}$$
(123)

where \bar{v}, v are independent vectors of Graßmann variables.

Let us sketch briefly how one can prove that. Consider first the case $\bar{v}, v = 0$. We then have to Taylor-expand the integral up to order n (if A is an $n \times n$ matrix). Checking the rules of integrations for Graßmann variables in Sec. 4.1.2, we realize that the only integral which is finite and can contribute is of the form

$$\int \prod_{i} d\bar{\eta}_{i} d\eta_{i} \prod_{j} \eta_{j} \bar{\eta}_{j} = 1.$$

What is the prefactor of the integral? Each column and each row of A has to show up once, $A_{1p_1}A_{2p_2}\cdots A_{np_n}$ and the prefactor $(-1)^p$ is obtained from ordering of the Graßmann variables. If one writes this out carefully, one obtains the definition of the determinant, $\sum (-1)^p A_{1p_1}A_{2p_2}\cdots A_{np_n} = \det(A)$, if p counts the number of permutations. Finally, to show that one can use the same 'completing the square' trick as in the complex case, one has to show that shifting variables does not change the value of the integral. For a single variable, we can easily check that

$$\int d\eta f(\eta + \eta') \underset{(\eta + \eta')^2 = 0}{\stackrel{\uparrow}{=}} f'(0) = \int d\eta f(\eta).$$

In the following chapter, we will need an important identity **for Gaussian integrals**, called **Wick's theorem**. Our goal will be to calculate expectation values relative to a Gaussian weight of the form

$$\langle \cdots \rangle_0 = \frac{\int \cdots e^{-\bar{\eta}A\eta}}{\int e^{-\bar{\eta}A\eta}}$$

Here the basic trick is to use Eq. (122) and Eq. (123), taking derivatives with respect to the fields w and v. In the fermionic case, for example, we take derivatives of $e^{\bar{\eta}v+\bar{v}\eta}$ with respect to \bar{v}, v

$$\langle \bar{\eta}_i \eta_j \rangle_0 = \left\langle -\frac{\partial}{\partial v_i} \frac{\partial}{\partial \bar{v}_j} e^{\bar{v}^{\mathsf{T}} \eta + \bar{\eta}^{\mathsf{T}} v} \right\rangle \Big|_{v=0} = -\frac{\partial}{\partial v_i} \frac{\partial}{\partial \bar{v}_j} e^{\bar{v}^{\mathsf{T}} A^{-1} v} \Big|_{v=0}$$

$$= -(A^{-1})_{ji} = -\langle \eta_j \bar{\eta}_i \rangle$$

$$(124)$$

The same trick also works for higher orders

$$\langle \eta_1 \eta_2 \bar{\eta}_3 \bar{\eta}_4 \rangle = \left. \frac{\partial}{\partial v_1} \frac{\partial}{\partial v_2} \frac{\partial}{\partial \bar{v}_3} \frac{\partial}{\partial \bar{v}_4} e^{\bar{v}^{\mathsf{T}} A^{-1} v} \right|_{v=0} \tag{125}$$

Thus, we have shown that calculating expectation values in Gaussian integrals is as easy as taking derivatives of $e^{\bar{v}^T A^{-1}v}$. The resulting rules for computing expectation values are called **Wick's theorem**. We will not derive it here (all the relevant math can be found above) but simply state the rules. It turns out that the formulas are the same for real, complex, or Graßmann Gaussian integrals if we set $\xi = -1$ in the Graßmann case, while setting $\xi = 1$ for real or complex Gaussian integrals. The Wick theorem is written as a sum over permutations and takes the form

$$\langle \psi_{j_1}\psi_{j_2}\cdots\psi_{j_n}\bar{\psi}_{i_n}\cdots\bar{\psi}_{i_1}\rangle_0 = \sum_{p\in S_n} \xi^{|p|} A_{j_1i_{p_1}}^{-1} A_{j_2i_{p_2}}^{-1}\cdots A_{j_ni_{p_n}}^{-1}$$
(126)

where $(-1)^p = -1$ if an odd number of permutations have been performed. Wick's theorem is an identity only valid for Gaussian integrals (over complex or Graßmann variables). Only for Gaussian integrals, it is possible to express higher-order correlation functions in terms of correlation functions of the type $\langle \psi_i \bar{\psi}_j \rangle_0 = A_{i,j}^{-1}$, encoding the width of the Gaussian.

For Graßmann fields, one finds, for example,

$$\langle \psi_1 \psi_2 \bar{\psi}_3 \bar{\psi}_4 \rangle = A_{14}^{-1} A_{23}^{-1} - A_{13}^{-1} A_{24}^{-1}$$

while the corresponding formula for bosnic fields is obtained using only + signs.

$$\langle \psi_1 \psi_2 \bar{\psi}_3 \bar{\psi}_4 \rangle = A_{14}^{-1} A_{23}^{-1} + A_{13}^{-1} A_{24}^{-1}$$

A useful notation to remember and evaluate the terms proceeds in the following way. One first finds all possibilities to pair Ψ_i with $\bar{\Psi}_j$ and denotes the pairing by a line. Then, 'disentagle' the pairs as in the example below. In case of Graßmann variables, add a minus sign whenever exchanging two Graßmann fields. This is best shown in an example

$$\langle \psi_1 \psi_2 \bar{\psi}_3 \bar{\psi}_4 \rangle = \psi_1 \psi_2 \bar{\psi}_3 \bar{\psi}_4 + \psi_1 \psi_2 \bar{\psi}_3 \bar{\psi}_4 = \psi_1 \bar{\psi}_4 \psi_2 \bar{\psi}_3 - \psi_1 \bar{\psi}_3 \psi_2 \bar{\psi}_4 = A_{14}^{-1} A_{23}^{-1} - A_{13}^{-1} A_{24}^{-1} + \dots$$

Here we denoted by + and - whether the term will receive an overall minus sign (in the fermionic case) after 'disentangling' everything. For bosonic fields, the result is the same but all prefactors are positive.

Here is an example with six Graßmann fields

$$\begin{split} \langle \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \rangle &= \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 + \psi_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+ u_1 \psi_2 \psi_3 \bar{\psi}_4 \bar{\psi}_5 \bar{\psi}_6 \\ &+$$

4.3 Green's functions

4.3.1 Definitions

Motivation: The quantity which is most easily calculated in a functional integral are Green's functions (or Green functions). Green's functions usually show up in the theory of linear differential equation, where they describe the response to a δ function. Here, we introduce them in a different variant, as expectation values of products of creation and annihiliation operators. In this section, we first focus on a special variant of the Green's function: the time-ordered Green's function in imaginary time. This turns out to be the object which can be most easily calculated in the functional integral. Only in later chapters we will realize how we can connect this imaginary-time quantity to measurable objects (retarded Green's functions, introduced later).

We start by considering time-dependent operators in the Heisenberg picture, $O(t) = e^{iHt/\hbar}Oe^{-iHt/\hbar}$. We want to bring the operators in a special order, called **time ordering**. Time-ordering just describes that we put the operator at late times to the right and at earlier times to the right. Here, we introduce a time-ordering operator \mathcal{T} defined by

$$\mathcal{T}[O_1(t_1)O_2(t_2)\cdots O_n(t_n)] = \xi^p O_{p_1}(t_{p_1})O_{p_2}(t_{p_2})\cdots O_{p_n}(t_{p_n})$$

with $t_{p_1} > t_{p_2} > \cdots > t_{p_n}, \ \xi = \begin{cases} -1 & \text{Fermions} \\ 1 & \text{Bosons} \end{cases}$

The factor ξ^p is -1 if we need to exchange an odd number of fermionic operators to reach the desired ordering. Note that $\mathcal{T}[O_1(t_1)O_2(t_2)] \neq O_1(t_1)O_2(t_2)$ for $t_1 < t_2$ as operators at different times typically do not commute. What do we do if two operators are evaluated on the same time? In this case, we use normal-odering, i.e., we put creation operators left and annhilation operators to the right.

A very useful application of the time-ordering formalism is that it allows to write the time-evolution operator U for a time-dependent Hamiltonian $\mathcal{H} = \mathcal{H}(t)$ in a compact form. The time-evolution operator, defined by $|\Psi(t)\rangle = U_t |\Psi(0)\rangle$ obeys the Schrödinger equation

$$i\hbar \frac{d}{dt}U_t = \mathcal{H}(t)U_t$$

For a static, time-independent Hamiltonian, this is solved by $U_t = e^{-iHt/\hbar}$ but this does not work for a time-dependent $\mathcal{H}t$ (why?). The problem is solved (at least formally) by writing instead

$$U_t = \mathcal{T} e^{-i \int_0^t dt' \,\mathcal{H}(t')}.$$

Consider, e.g., the second term in the Taylor expansion of exp of $\frac{d}{dt}U_t$, where one

has to compute

$$\mathcal{T}\left[\frac{d}{dt}\int_0^t dt_1 \,\mathcal{H}(t_1)\int_0^t dt_2 \,\mathcal{H}(t_2)\right] = \mathcal{T}\left[\mathcal{H}(t)\int_0^t dt_2 \,\mathcal{H}(t_2) + \int_0^t dt_1 \,\mathcal{H}(t_1)\mathcal{H}(t)\right]$$
$$= 2H(t)\int_0^t dt_2 \,\mathcal{H}(t_2)$$

Only because of the time-ordering operator $\mathcal{H}(t)$ is moved to the left of the expression. This leads to

$$\partial_t U_t = -\frac{i}{\hbar} \underbrace{\mathcal{H}(t)}_{\text{largest } t \text{ on left side!}} U(t)$$

and thus U_t solves the Schrödinger equation.

We want to apply these concepts also in imaginary time. As a first step, we have to define Heisenberg operators in imaginary time. For creation and annihilation operators, they take the form

$$a_{\alpha}^{\dagger}(\tau) = e^{\tau(\mathcal{H}-\mu\mathcal{N})} a_{\alpha}^{\dagger} e^{-\tau(\mathcal{H}-\mu\mathcal{N})}$$
$$a_{\alpha}(\tau) = e^{\tau(\mathcal{H}-\mu\mathcal{N})} a_{\alpha} e^{-\tau(\mathcal{H}-\mu\mathcal{N})}$$

Note that with these definitions, $(a_{\alpha}(\tau))^{\dagger} = a_{\alpha}^{\dagger}(-\tau) \neq a_{\alpha}^{\dagger}(\tau)$.

We define next the time-ordered Green's function in imaginary time

$$\mathcal{G}(\alpha_1\tau_1,\alpha_2\tau_2) = \mathcal{G}_{\alpha_1,\alpha_2}(\tau_1-\tau_2) = -\langle \mathcal{T}a_{\alpha_1}(\tau_1)a_{\alpha_2}^{\dagger}(\tau_2)\rangle$$
$$= \begin{cases} -\xi \langle a_{\alpha_2}^{\dagger}(\tau_2)a_{\alpha_1}(\tau_1) \rangle & \tau_2 \ge \tau_1\\ -\langle a_{\alpha_1}(\tau_1)a_{\alpha_2}^{\dagger}(\tau_2) \rangle & \tau_2 < \tau_1 \end{cases}$$

where $\langle \cdots \rangle = \frac{1}{Z} \operatorname{Tr} \left(\cdots e^{-\beta((\mathcal{H}-\mu\mathcal{N}))} \right)$ denotes expectation values in a thermal state. We are using grand-canonical ensembles with a chemical potential μ . In the following formulas, however, we will absorb the chemical potential term in \mathcal{H} using $\mathcal{H}-\mu\mathcal{N} \to \mathcal{H}$ to make the formulas a bit shorter and more easy to read.

Above, we claimed that time-ordered Green's function can easily be computed with functional integrals. To see why, simply insert all of our definitions. We obtain

$$\tau_{2} \geq \tau_{1} : -\xi \langle a_{\alpha_{2}}^{\dagger}(\tau_{2}) a_{\alpha_{1}}(\tau_{1}) \rangle = -\frac{\xi}{\mathcal{Z}} \operatorname{Tr} \left(e^{-(\beta - \tau_{2})\mathcal{H}} a_{\alpha_{2}}^{\dagger} e^{-(\tau_{2} - \tau_{1})\mathcal{H}} a_{\alpha_{1}} e^{-\tau_{1}\mathcal{H}} \right)$$

$$\tau_{2} < \tau_{1} : -\langle a_{\alpha_{1}}(\tau_{1}) a_{\alpha_{2}}^{\dagger}(\tau_{2}) \rangle = -\frac{1}{\mathcal{Z}} \operatorname{Tr} \left(e^{-(\beta - \tau_{1})\mathcal{H}} a_{\alpha_{1}} e^{-(\tau_{1} - \tau_{2})\mathcal{H}} a_{\alpha_{2}}^{\dagger} e^{-\tau_{2}\mathcal{H}} \right)$$

We observe that the time-evolution operator in each case is simply split into three pieces. In the first line, e.g., from 0 to τ_1 from τ_1 to τ_2 and, finally, from τ_2 to β .

We can now follow our standard program: we split the time evolution from 0 to β into N pieces (Trotter), $\Delta \tau = \frac{\beta}{N}$, and introduce N identity operators 1, which are coherent states. The only difference compared to the functional integral for Z are the extra operators a_{α_1} and $a_{\alpha_2}^{\dagger}$ which are easily evaluated with the help of the

coherent states

$$a_{\alpha_1} \underbrace{|\psi_{n_1}\rangle}_{\text{coherent state}} = \psi_{n_1,\alpha_1} |\psi_{n_1}\rangle, \qquad \underbrace{\langle\psi_{n_2+1}|}_{\text{coherent state}} a_{\alpha_2}^{\dagger} = \langle\psi_{n_2+1} |\psi_{n_2+1,\alpha_2}\rangle$$

where $n_1 = N \frac{\tau_1}{\beta}$ and $n_2 = N \frac{\tau_2}{\beta}$ are the time steps corresponding to the time τ_1 and τ_2 , respectively. Note that the time-index of the destruction operator is shifted by 1 as we use the coherent state to the left of the operator. Thus, the corresponding field is evaluated an infinitesimal time step $\epsilon = \beta/N$ later.

Thus, we find both in the bosonic and fermionic case that the imaginary-time timeordered Green function (both for $\tau_1 > \tau_2$ and $\tau_2 \ge \tau_1$) is given by

$$\mathcal{G}(\alpha_1\tau_1,\alpha_2\tau_2) = \mathcal{G}_{\alpha_1,\alpha_2}(\tau_1-\tau_2) = -\frac{\int \mathscr{D}(\bar{\psi},\psi)\,\psi_{\alpha_1}(\tau_1)\bar{\psi}_{\alpha_2}(\tau_2+\varepsilon)e^{-\mathcal{S}[\bar{\psi},\psi]}}{\int \mathscr{D}(\bar{\psi},\psi)\,e^{-\mathcal{S}[\bar{\psi},\psi]}}$$

This holds in general, time-ordered correlation functions of operators can directly be written as correlation functions of the corresponding fields in functional integrals.

Let us evaluate the Green function for a time difference of β

$$\begin{aligned} \mathcal{G}_{\alpha_1,\alpha_2}(\beta) &= -\langle a_{\alpha_1}(\beta) a_{\alpha_2}^{\dagger}(0) \rangle = -\frac{1}{\mathcal{Z}} \operatorname{Tr} \left(e^{-(\beta-\beta)\mathcal{H}} a_{\alpha_1} e^{-\beta\mathcal{H}} a_{\alpha_2}^{\dagger} \right) \\ &= -\frac{1}{\mathcal{Z}} \operatorname{Tr} \left(e^{-\beta\mathcal{H}} a_{\alpha_2}^{\dagger} a_{\alpha_1} \right) = \xi \mathcal{G}_{\alpha_1,\alpha_2}(0) \end{aligned}$$

This boundary condition is directly related to the boundary conditions of fields in the functional integral and we will show later that it has profound consequences

The Green's function
$$\mathcal{G}_{\alpha_1,\alpha_2}(\tau)$$
 is $\begin{cases} \text{periodic} & \text{in } \beta \text{ for } \text{fermions} \\ \text{antiperiodic} & \text{in } \beta \text{ for } \text{fermions} \end{cases}$

4.3.2 Free particle

As a next step, we evaluate the partition sum and the Green's function for noninteracting particles where the Hamiltonian and the action are given by

$$\begin{aligned}
\mathcal{H} &= \sum_{\alpha} (\varepsilon_{\alpha} - \mu) a_{\alpha}^{\dagger} a_{\alpha} \\
\mathcal{S} &= \sum_{\alpha} \int_{0}^{\beta} d\tau \, \bar{\psi}_{\alpha}(\tau) (\partial_{t} + (\varepsilon_{\alpha} - \mu)) \psi_{\alpha}(\tau) \\
&= \sum_{\substack{\alpha \\ \uparrow \\ \text{discrete} \\ \text{definition}}} \sum_{\alpha} \left(\bar{\psi}_{\alpha 0} \ \bar{\psi}_{\alpha 1} \ \cdots \ \bar{\psi}_{\alpha N-1} \right) \mathcal{S}^{(\alpha)} \begin{pmatrix} \bar{\psi}_{\alpha 0} \\ \bar{\psi}_{\alpha 1} \\ \vdots \\ \bar{\psi}_{\alpha N-1} \end{pmatrix}
\end{aligned} \tag{127}$$

with

$$\mathcal{S}^{(\alpha)} = \begin{pmatrix} 1 & & & -\xi a \\ -a & 1 & & \\ & -a & 1 & & \\ & & \ddots & \\ & & & -a & 1 \end{pmatrix} \leftarrow \text{distinguishes bosons/fermions}$$
$$a = 1 - \frac{\beta}{N} (\varepsilon_{\alpha} - \mu)$$

Here, we used the discreet definition of the functional integral, which is the safest way to obtain the correct the result for the partition sum and the Green's function.

According to Eq. (122) and Eq. (123), the Gaussian functional integral is obtained from the determinant of $\mathcal{S}^{(\alpha)}$, which can easily be evaluated

$$\det(S^{(\alpha)}) \underset{\text{row}}{\stackrel{\uparrow}{\underset{\text{row}}{1}}} 1 + (-1)^{N-1} (-\xi a) (-a)^{N-1} = 1 - \xi \left(1 - \frac{\beta}{N} (\varepsilon_{\alpha} - \gamma)\right)^{N} \underset{\substack{\uparrow\\N \to \infty}{=}}{=} 1 - \xi e^{-\beta(\varepsilon_{\alpha} - \mu)}$$

Therefore, the partition sum is obtained as

$$\mathcal{Z}_{0} = \begin{cases} \prod_{\alpha} \frac{1}{\det(S^{(\alpha)})} = \prod_{\alpha} \frac{1}{1 - e^{-\beta(\varepsilon_{\alpha} - \mu)}} & \text{bosons} \\ \prod_{\alpha} \det(S^{(\alpha)}) = \prod_{\alpha} 1 + e^{-\beta(\varepsilon_{\alpha} - \mu)} & \text{fermions} \end{cases}$$

which is the result well known from statistical physics.

To compute the Green's function, we need the inverse of $\mathcal{S}^{(\alpha)}$. Due to the simple structure of $\mathcal{S}^{(\alpha)}$, it can also be evaluated analytically

$$S^{(\alpha)^{-1}} = \frac{1}{1 - \xi a^{N}} \begin{pmatrix} 1 & \xi a^{N-1} & \xi a^{N-2} & \cdots & \cdots & \xi a \\ a & 1 & \xi a^{N-1} & & \xi a^{2} \\ a^{2} & a & 1 & & \xi a^{3} \\ a^{3} & a^{2} & a & 1 & & \vdots \\ \vdots & & & \ddots & \vdots \\ a^{N} - 1 & \cdots & \cdots & \cdots & \cdots & \xi a^{N-1} \end{pmatrix}$$

We can rewrite this in the limit $N \to \infty$

$$\text{for } i \leq j : \left(S^{(\alpha)^{-1}}\right)_{ij} = \frac{a^{i-j}}{1-\xi a^N} \\ = \underbrace{\frac{\left(1 - \frac{\beta}{N}(\varepsilon_{\alpha} - \mu)\right)^{\frac{N}{\beta}(\tau_i - \tau_j)}}{1 - \xi \left(1 - \frac{\beta}{N}(\varepsilon_{\alpha} - \mu)\right)^N} \xrightarrow[N \to \infty]{} \frac{e^{-\varepsilon_{\alpha}(\tau_i - \tau_j)}}{1 - \xi e^{-\beta(\varepsilon_{\alpha} - \mu)}} \\ \text{similarly } i < j : \left(S^{(\alpha)^{-1}}\right)_{ij} \longrightarrow \frac{\xi e^{-\varepsilon_{\alpha}(\beta - (\tau_i - \tau_j))}}{1 - \xi e^{-\beta(\varepsilon_{\alpha} - \mu)}}$$

Thus we obtain for the Green's function

$$\mathcal{G}_{\alpha_1,\alpha_2}(\tau_1-\tau_2) = \langle \mathcal{T}a_{\alpha_1}(\tau_1)a_{\alpha_2}^{\dagger}(\tau_2) \rangle = \lim_{N \to \infty} -\delta_{\alpha\alpha'} \left(S^{(\alpha)^{-1}} \right)_{ij} = \delta_{\alpha\alpha'} g_{\alpha}(\tau_1-\tau_2)$$

Bosons:
$$g_{\alpha}(\tau) = -e^{(\varepsilon_{\alpha}-\mu)\tau} \left[\Theta(\tau-\varepsilon)\left(1+n_{B}(\varepsilon_{\alpha}-\mu)\right)+\Theta(-\tau+\varepsilon)n_{B}(\varepsilon_{\alpha}-\mu)\right]$$

 $n_{B}(\omega) = \frac{1}{e^{\beta\omega}-1}$ Bose function, $\varepsilon = \frac{\beta}{N} \to 0, \Theta(\tau) = \begin{cases} 0 & \tau < 0\\ 1 & \tau > 0 \end{cases}$
Fermions: $g_{\alpha}(\tau) = -e^{(\varepsilon_{\alpha}-\mu)\tau} \left[\Theta(\tau-\varepsilon)\left(1-n_{F}(\varepsilon_{\alpha}-\mu)\right)-\Theta(-\tau+\varepsilon)n_{F}(\varepsilon_{\alpha}-\mu)\right]$
 $n_{F}(\omega) = \frac{1}{e^{\beta\omega}+1}$ Fermi function

Let us check what happens when we set τ to zero.

$$g_{\alpha}(0) = -\xi \langle a_{\alpha}^{\dagger} a_{\alpha} \rangle = \begin{cases} -n_B(\varepsilon_{\alpha} - \mu) \\ n_F(\varepsilon_{\alpha} - \mu) \end{cases}$$

Thus, we recover that Fermi and Bose functions describe the occupation of particles. Note that were only able to obtain this result by doing all computations carefully at finite N, taking the limit $N \to \infty$ only at the end. Luckily, we will have to do this somewhat painful exercise only once and the calculations also get easier when working in Fourier space, which we want to consider next.

4.3.3 Matsubara frequencies

A frequently used standard trick to simplify all types of problems it to do a Fourier transformation. While we argued above, that the only safe way to do a calculation is to use the discreet definition of the functional integral, we will now try to work directly in the continuum limit (but we will have to readjust the resulting formula slightly in the end).

$$S_0 = \sum_{\alpha} \int_0^\beta d\tau \, \bar{\psi}_{\alpha}(\tau) (\partial_t + (\varepsilon_{\alpha} - \mu)) \psi_{\alpha}(\tau)$$

with $\psi_{\alpha}(0) = \xi \psi_{\alpha}(\beta)$

We use standard Fourier transformations

$$\bar{\psi}_{\alpha}(\tau) = \frac{1}{\sqrt{\beta}} \sum_{n} \bar{\psi}_{\alpha,n} e^{i\omega_{n}\tau} \qquad \bar{\psi}_{\alpha,n} = \frac{1}{\sqrt{\beta}} \int_{0}^{\beta} d\tau \, \bar{\psi}_{\alpha}(\tau) e^{-i\omega_{n}\tau}$$
$$\psi_{\alpha}(\tau) = \frac{1}{\sqrt{\beta}} \sum_{n} \psi_{\alpha,n} e^{-i\omega_{n}\tau} \qquad \psi_{\alpha,n} = \frac{1}{\sqrt{\beta}} \int_{0}^{\beta} d\tau \, \psi_{\alpha}(\tau) e^{i\omega_{n}\tau}$$

Here, it is important to recall that the fermionic /bosonic fields have antiperiodic/periodic boundary conditions, respectively. Therefore, we have to adjust the so-called **Matsubara frequencies** accordingly

$$\omega_n = \begin{cases} \frac{2\pi}{\beta}n & \text{bosons}\\ \frac{2\pi}{\beta}\left(n+\frac{1}{2}\right) & \text{fermions} \end{cases}, n \in \mathbb{Z}$$

with

Using $\int_0^\beta e^{i(\omega_n - \omega_m)\tau} = \beta \, \delta_{\omega_n, \omega_m}$, we obtain for the action

$$\mathcal{S}_0 = \sum_{\alpha} \sum_{\omega_n} \bar{\psi}_{\alpha,n} (-i\omega_n + (\varepsilon_\alpha - \mu))\psi_{\alpha,n}$$

which is diagonal in Matsubara indices n.

Also interactions can be expressed in Matsubara modes

$$S_{\text{int}} = \int_{0}^{\beta} d\tau \, V_{\alpha\beta,\gamma\rho} \bar{\psi}_{\alpha}(\tau) \bar{\psi}_{\beta}(\tau) \psi_{\gamma}(\tau) \psi_{\rho}(\tau)$$
$$= \frac{1}{\beta} \sum_{n_{1},n_{2},n_{3},n_{4}} V_{\alpha\beta,\gamma\rho} \bar{\psi}_{\alpha,n_{1}} \bar{\psi}_{\beta,n_{2}} \psi_{\gamma,n_{3}} \psi_{\rho,n_{4}} \, \delta_{n_{1}+n_{2},n_{3}+n_{4}}$$

where the Kronecker delta describes that the sum of frequencies of destruction operators matches that of creation operators, a formula which we will identify later with energy conservation.

The functional integral is then obtained by summing over all Fourier modes

$$\mathcal{Z} = \int \mathscr{D}(\bar{\psi}_n, \psi_n) \, e^{-(\mathcal{S}_0 + \mathcal{S}_{\text{int}})}$$

We can also express Green's functions in terms of Fourier modes.

$$\mathcal{G}_{\alpha\alpha'}(\tau) = -\langle \psi_{\alpha}(\tau)\bar{\psi}_{\alpha'}(\varepsilon)\rangle = \frac{1}{\beta} \sum_{\omega_n} e^{-i\omega_n\tau} \mathcal{G}_{\alpha\alpha'}(i\omega_n) e^{i\omega_n\varepsilon}$$
$$\mathcal{G}_{\alpha\alpha'}(i\omega_n) = \int_0^\beta d\tau \, e^{i\omega_n\tau} \, \mathcal{G}_{\alpha\alpha'}(\tau) = -\langle \psi_{\alpha,n}\bar{\psi}_{\alpha,n}\rangle$$

When deriving the equation above, we used that $\langle \psi_{\alpha,n} \bar{\psi}_{\alpha',m} \rangle = 0$ for $n \neq m$. This is equivalent to the observation that our imaginary-time Green's function depends only on the time difference, $\mathcal{G}_{\alpha\alpha'}(\tau_1, \tau_2) = \mathcal{G}_{\alpha\alpha'}(\tau_1 - \tau_2)$. Note the – very important – factor ϵ in the first equation. This is what we need to recover the correct expressions obtained from the discreet definition of the functional integral. ϵ can be identified with β/N and we have to take the limit $\epsilon \to 0$ at the end of the calculation.

Let us consider, how ϵ shows up when computing a Fourier transformation.

$$\int_{0}^{\beta} d\tau \, e^{i\omega_{n}\tau} \mathcal{G}_{\alpha\alpha'}(\tau) = -\int_{0}^{\beta} d\tau \, e^{i\omega_{n}\tau} \langle \psi_{\alpha}(\tau)\bar{\psi}_{\alpha'}(\varepsilon) \rangle -\int_{0}^{\beta} d\tau \, e^{i\omega_{n}\tau} \frac{1}{\beta} \sum_{n_{1},n_{2}} e^{i\omega_{n_{1}}\tau} e^{i\omega_{n_{2}}\varepsilon} \underbrace{\langle \psi_{\alpha,n_{1}}\bar{\psi}_{\alpha',n_{2}} \rangle}_{=0 \text{ for } n_{1}\neq n_{2}} = -\langle \psi_{\alpha,n}\bar{\psi}_{\alpha',n} \rangle e^{i\omega_{n}\varepsilon}$$

Here, taking the limit $\epsilon \to 0$ has no consequence. This is different when computing, e.g., $\mathcal{G}_{\alpha\alpha'}(\tau = 0)$, which we will consider in the next section.

Finally, we compute the Green's function for Matsubara frequencies in the absence of interaction. As the action is diagonal

$$e^{-\mathcal{S}_0} = e^{-\sum_{\alpha,n} \bar{\psi}_{\alpha,n}(-i\omega_n + (\varepsilon_\alpha - \mu))\psi_{\alpha,n}}$$

we can read off the result without the need of any matrix inversion and obtain

$$\mathcal{G}^0_{\alpha}(i\omega_n) = \frac{1}{i\omega_n - (\varepsilon_{\alpha} - \mu)}$$

4.4 Summations and examples

Motivation: One of the main goals of the next section of the lecture will be to develop efficient techniques to compute observables using perturbation theory. For this, we will use Feynman diagrams, which are a convenient way to visualize formulas.



Importantly, when evaluating these formulas, we will almost always have to perform summations over Matsubara frequencies. In this section, we will learn the necessary mathematical technique.

Our goal is to calculate a sum of the type

$$\frac{1}{\beta} \sum_{\omega_n} f(i\omega_n) \quad \text{with} \quad \omega_n = \frac{2\pi}{\beta} \begin{cases} n & \text{bosons} \\ n + \frac{1}{2} & \text{fermions} \end{cases}$$

Here, our plan is to use the magic of complex analysis, or, more precisely, the residue theorem. Here, we need functions which have single poles at Matsubara frequencies. Perhaps surpisingly, the Bose- and Fermi functions have precisely this property.

Let us start with the **Bose function**

$$n_B(z) = \frac{1}{e^{\beta z} - 1}$$

which we now consider as a function of a complex variable z. This function has poles for $z = i\omega_n = \frac{2\pi i}{\beta}n$. To see that, let us Taylor-expand $n_B(z)$ around $i\omega_n$

$$n_B(i\omega_n + z) = \frac{1}{e^{i2\pi n}e^{\beta z} - 1} \approx \frac{1}{\beta z}$$
 for $z \to 0$

A similar statement holds for the Fermi function

$$n_F(z) = \frac{1}{e^{\beta z} + 1}$$

when we consider fermionic Matsubara frequencies $i\omega_n = \frac{2\pi i}{\beta} \left(n + \frac{1}{2}\right)$,

$$n_F(i\omega_n + z) = \frac{1}{e^{i2\pi(n + \frac{1}{2})}e^{\beta z} + 1} = \frac{1}{1 - e^{\beta z}} \approx -\frac{1}{\beta z} \text{ for } z \to 0.$$



Figure 7: Pole structure of the Bose (red) and Fermi (blue) functions and contour Γ used for the integral in Eq. (128).

We can now use the residuum theory by integrating along a contour enclosing all poles of the Bose- or Fermi function, see Fig.7. Provided that f(z) is an analytic function with no poles within the contour, we find

$$\frac{1}{\beta} \sum_{\omega_n} f(i\omega_n) = \xi \oint_{\Gamma} \frac{dz}{2\pi i} n_{B/F}(z) f(z)$$
(128)

where the factor $\xi = \pm 1$ arises as the Fermi function (where $\xi = -1$) has poles of the type $-1/(\beta z)$ while the Bose function ($\xi = 1$) has poles $1/(\beta z)$.

Having replaced the sum by an integral may not sound like a big advantage, but now we can use the power of complex analysis (Cauchy's integral theorem) to deform the contour in a convenient way.

Therefore, we will follow the following recipe. (i) Write all sums as **contour integrals**, (ii) analyze **poles and possible branch-cuts** of f(z), (iii) **deform the contour** "conveniently", using, for example, that points at infinity may not contribute to the integral, if $n(z)f(z) < \frac{1}{z^{\alpha}}$ for $\alpha > 1$.

Let us calculate, for example, the expectation value of $\langle a_{\alpha}^{\dagger}a_{\alpha}\rangle$ for the Hamiltonian $\mathcal{H} = \sum_{\alpha} (\epsilon_{\alpha} - \mu) a_{\alpha}^{\dagger}a_{\alpha}$, where the Green's function is given by $1/(i\omega_n - (\epsilon_{\alpha} - \mu))$. The equal-time expectation value is written as a sum over all Matsubara frequencies, which we rewrite as a contour integral.

$$\begin{aligned} \langle a_{\alpha}^{\dagger} a_{\alpha} \rangle &= \langle \bar{\psi}_{\alpha}(\varepsilon) \psi_{\alpha}(0) \rangle = -\xi \frac{1}{\beta} \sum_{n} e^{i\omega_{n}\varepsilon} \mathcal{G}_{\alpha}(i\omega_{n}) \\ &= -\xi \frac{1}{\beta} \sum_{n} e^{i\omega_{n}\varepsilon} \frac{1}{i\omega_{n} - (\varepsilon_{\alpha} - \mu)} = -\oint_{\Gamma} \frac{dz}{2\pi i} n_{B/F}(z) \frac{1}{z - (\varepsilon_{\alpha} - \mu)} e^{z\varepsilon} \end{aligned}$$

Note the factor $e^{z\varepsilon}$. In this example, the function $f(z) = \frac{e^{z\varepsilon}}{z - (\varepsilon_{\alpha} - \mu)}$ has a simple pole at $z = (\varepsilon_{\alpha} - \mu)$, see Fig. 8. As the next step, we deform the contour of Fig. 7 to that of Fig. 8.

Next, we have to find out, whether there is a contribution from the circle located at



Figure 8: Contour integral for a function with a simple pole.

 $|z| = \infty$ in the complex plane. Here, we use that

$$\frac{e^{\varepsilon z}}{e^{\beta z} \mp 1} \sim \begin{cases} e^{-\beta \operatorname{Re} z} & \text{for } \operatorname{Re} z \to \infty \\ \mp e^{\epsilon \operatorname{Re} z} & \text{for } \operatorname{Re} z \to -\infty \end{cases} \to 0 \quad \text{for} \quad |\operatorname{Re} z| \to \infty.$$

In combination with the 1/z factor from the Green's function, this ensures that the circle at infinity does not contribute. Note the importance of the $e^{\varepsilon z}$ term to reach that conclusion.

Thus, we conclude that the only contribution to the integral comes from the pole at $\epsilon_{\alpha} - \mu$ and we obtain the well-known result

$$\langle a_{\alpha}^{\dagger}a_{\alpha}\rangle = -\oint_{\Gamma}\frac{dz}{2\pi i}\,n_{B/F}(z)\frac{1}{z-(\varepsilon_{\alpha}-\mu)}e^{z\varepsilon} = n_{B/F}(\varepsilon_{\alpha}-\mu).$$

The Bose and Fermi function describe the occupation of states. Apparently there is a surprising link of the analytical properties of the Bose- and Fermi functions in the complex plane to their physical significance.

Above, we were considering functions with a single pole. A much more common situation are branch cuts. As we will see later, they naturally arise in interacting systems. Here, we consider the free energy of a non-interacting system. In this case, we have shown above that

$$e^{-\beta F} = \mathcal{Z} = C \prod_{\alpha} (i\omega_n - (\varepsilon_\alpha - \mu))^{-\xi}$$

where C is some (uninteresting) normalization constant.

Thus the free energy can be written as a sum over Matsubara frequencies, which we again rewrite as a contour integral.

$$F = \operatorname{const} + \xi \frac{1}{\beta} \sum_{\omega_n, \alpha} \ln \left(i\omega_n - (\varepsilon_\alpha - \mu) \right) e^{i\omega_n \varepsilon}$$
$$= \operatorname{const} + \sum_{\alpha} \oint_{\Gamma} \frac{dz}{2\pi i} n_{B/F}(z) \ln \left(z - (\varepsilon_\alpha - \mu) \right) e^{z\varepsilon}$$



Figure 9: Branch cut of the logarithm (green) and the chosen contour for the integration.

Here, we have to consider the analytical structure of $\ln z$. $\ln z$ is defined by the equation $e^{\ln z} = z$ but there is some ambiguity as one can always add $i2\pi$ to the exponent (and therefore to the logarithm). One can, however, not avoid that $\ln z$ is a discontinuous function somewhere. For convenience we define this branch cut to be on the negative real axis. For $z = -|r| + i\epsilon$, we use

$$\ln(-r \pm i\epsilon) = \ln|r| \pm i\pi$$

consistent with $e^{\ln |r| \pm i\pi} = -|r|$. Thus, the imaginary part of $\ln z$ jumps from $i\pi$ to $-i\pi$ upon crossing the branch cut in Fig. 9.

Next, we deform the initial contour of integration, the black line in Fig. 9, to the blue line in Fig. 9, using as before that contributions at infinity vanish as $\frac{e^{\varepsilon z}}{e^{\beta z} \mp 1} \rightarrow 0$. The two blue contours are at $\omega + \pm i\epsilon$ and contribute with opposite sign due to the different orientation of the two contours. For $\epsilon \rightarrow 0$, the only contribution can arise from the discontinuity of $\ln z <$

$$F = C + \sum_{\alpha} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} \underbrace{n_{B/F}(\omega)}_{\text{use } n_{B/F}(\omega \pm \varepsilon) = n_{B/F}(\omega)}_{\text{up to pole of } n_{B}(\omega)} (\ln(\omega + i\varepsilon - (\epsilon_{\alpha} - \mu))) - \ln(\omega - i\varepsilon - (\epsilon_{\alpha} - \mu)))$$
$$= C + \sum_{\alpha} \int_{-\infty}^{\infty} \frac{d\omega}{\pi} n_{B/F}(\omega) \underbrace{\operatorname{Im}(\ln(\omega + i0 - (\epsilon_{\alpha} - \mu))))}_{\pi\Theta((\epsilon_{\alpha} - \mu) - \omega)}$$
$$= C + \sum_{\alpha} \int_{-\infty}^{(\epsilon_{\alpha} - \mu)} d\omega n_{B/F}(\omega)$$
$$= \left| n_{B/F}(\omega) = \pm \frac{1}{\beta} \frac{\partial}{\partial \omega} \ln(1 \mp e^{-\beta\omega}) \right| = C' \pm T \sum_{\alpha} \ln(1 \mp e^{-\beta(\epsilon_{\alpha} - \mu)})$$

where the upper sign is for bosons and the lower sign for fermions. One can check

that this recovers the well-known result from statistical physics

$$Z = e^{-\beta F} = \begin{cases} \prod_{\alpha} \sum_{n=0}^{\infty} e^{-\beta(\epsilon_{\alpha}-\mu)n} & \text{for bosons} \\ \prod_{\alpha} \left(1 + e^{-\beta(\epsilon_{\alpha}-\mu)}\right) & \text{for fermions} \end{cases}.$$

These examples show how one can use insights from complex analysis to compute the necessary sums over bosonic or fermionic Matsubara sums.

5 Diagramatic Perturbation Theory

Motivation: For practically all quantum field theories, an exact computation of physical quantities is impossible and one has to rely on approximation methods. In the following sections, we will learn about the most important approximation method: diagrammatic perturbation theory. Here 'diagrammatic' refers to an idea introduced by Feynman: instead of writing formulas, we draw diagrams ('Feynman diagrams') to represent those formulas. The backbone of diagrammatic perturbation theory is a simple Taylor expansions in the strength of interaction but we will learn tricks to simplify the latter.

5.1 Perturbation theory for Green's functions

As a first step, we split our action into a part quadratic in the fields, S_0 , and the rest, which we call interactions, S_{int} .

$$\mathcal{S} = \mathcal{S}_0 + \lambda \mathcal{S}_{int}, \ \mathcal{S}_0$$
 quadratic in fields, $\mathcal{S}_0 = \sum_{ij} \bar{\psi}_i \left(g^0\right)_{ij}^{-1} \psi_j$

Our primiary goal in this section is to compute the Green's function in perturbation theory.

$$\mathcal{G} = -\langle \psi_i \bar{\psi}_j \rangle = -\frac{\int \mathscr{D}(\bar{\psi}, \psi) e^{-(\mathcal{S}_0 + \mathcal{S}_{\text{int}})} \psi_i \bar{\psi}_j}{\int \mathscr{D}(\bar{\psi}, \psi) e^{-(\mathcal{S}_0 + \mathcal{S}_{\text{int}})}}$$
(129)

Thus, we will first do a Taylor expansion in λ (or, equivalently, in S_{int}) and we want to use Wick's theorem to evaluate terms like $\int e^{S_0} \psi_i \bar{\psi}_j (S_{\text{int}})^n$.

Here, we face the practical problem that a lots and lots of terms are generated but it turns out that many of them either cancel with each other or give identical contributions. I turns out that it is extremely useful to replace the formulas which we obtain by "diagrams", schematic picture, which represent formulas. We will compute once how often each diagram contributes to obtain their combinatorial prefactors.

Before we do this in full generality, let us do a simple example. We start with the non-interacting problem S_0

$$\mathcal{S}_{0} = \sum_{i,j} \bar{\psi}_{i} \left(g^{0}\right)_{ij}^{-1} \psi_{j} \implies -\langle \psi_{i} \bar{\psi}_{j} \rangle_{0} = \frac{\int \mathscr{D}(\bar{\psi}, \psi) \left(-\psi_{i} \bar{\psi}_{j}\right) e^{-\mathcal{S}_{0}}}{\int \mathscr{D}(\bar{\psi}, \psi) e^{-\mathcal{S}_{0}}} = \left(g^{0}\right)_{ij}$$

The application, we have in mind is that S_0 describes, e.g., Fermions with single band with dispersion $\epsilon_{\vec{k}}$, which are interacting with Coulomb interactions

$$\begin{split} \mathcal{S}_{0} &= \sum_{\substack{n,\sigma=\uparrow/\downarrow\\\vec{k}\in 1\text{B.Z}}} \bar{\psi}_{n,\vec{k},\sigma} \left(-(i\omega_{n}-\varepsilon_{\vec{k}}) \right) \psi_{n,\vec{k},\sigma} \\ \mathcal{S}_{\text{int}} &= \frac{1}{2} \int d^{3}x \, d^{3}x' \, V(\vec{x}-\vec{x}') \int_{0}^{\beta} d\tau \, \sum_{\sigma,\sigma'} \bar{\psi}_{\sigma}(\vec{x},\tau) \bar{\psi}_{\sigma}'(\vec{x}',\tau) \psi_{\sigma}'(\vec{x},\tau) \\ &= \frac{T}{2} \frac{1}{\mathcal{V}} \sum_{\vec{k},\vec{k}',\vec{q}} V_{\vec{q}} \bar{\psi}_{\vec{k}+\vec{q},\sigma} \bar{\psi}_{\vec{k}'-\vec{q},\sigma'} \psi_{\vec{k}',\sigma'} \psi_{\vec{k},\sigma}, \qquad \text{where, } V_{\vec{q}} = \int dr \, V(\vec{r}) e^{i\vec{q}\cdot\vec{r}} \end{split}$$

To obtain some intuition, we first consider 1st order perturbation theory, linear in $S_{\rm int}$.

$$\begin{aligned} \mathcal{G} &= -\langle \psi_2 \bar{\psi}_1 \rangle = -\frac{\int \mathscr{D}(\psi, \psi) e^{-(\mathcal{S}_0 + \mathcal{S}_{\text{int}})} \psi_2 \psi_1}{\int \mathscr{D}(\bar{\psi}, \psi) e^{-(\mathcal{S}_0 + \mathcal{S}_{\text{int}})}} \\ &\approx -\langle \psi_2 \bar{\psi}_1 \rangle_0 + \left\langle \psi_2 \bar{\psi}_1 \int_0^\beta d1' d2' \, \bar{\psi}_{1'} \bar{\psi}_{2'} \psi_{2'} \psi_{1'} \frac{V(2' - 1')}{2} \right\rangle_0 \\ &- \langle \psi_2 \bar{\psi}_1 \rangle_0 \left\langle \iint_0^\beta d1' d2' \, \bar{\psi}_{1'} \bar{\psi}_{2'} \psi_{2'} \psi_{1'} \frac{V(2' - 1')}{2} \right\rangle_0 + \mathcal{O}(V^2) \end{aligned}$$

Here, we used a short-hand notation, where $\int d1' = \int d\tau'_1 d\vec{x}'_1$ is an integral over both time and space, and $\psi_{1'} = \psi(\tau'_1, \vec{x}'_1)$ while $V(2' - 1') = V(\vec{x}'_2 - \vec{x}'_1)$.

Next, we will apply Wick's theorem, Eq- (126))

$$\langle \psi_{j_1}\psi_{j_2}\cdots\psi_{j_n}\bar{\psi}_{i_n}\cdots\bar{\psi}_{i_1}\rangle_0 = \sum_{p\in S_n} \xi^{|p|} A_{j_1i_{p_1}}^{-1} A_{j_2i_{p_2}}^{-1}\cdots A_{j_ni_{p_n}}^{-1}$$

for $S_0 = \bar{\psi}A\psi$. For our example, this gives rise to 3!+2!+1! = 9 terms. To write these terms, we use a graphical representation, based on **Feynman diagrams** (defined more precisely later). We denote

$$-\langle \psi_2 \bar{\psi}_1 \rangle = 2 - 4 - 1$$
$$-V(1-2) = 1 - 2$$

Using this notation and Wick's theorem, we find (up to signs which we will analyze more carefully later) that



These are our first examples of Feynman diagrams. We observe two things: (i) a lot of terms cancel, e.g., term 2 and 3 cancels with the last two terms. (ii) some of the diagrams which have topologically the same shape corresponds (possibly after renaming the integration variables 1' and 2' to the exactly the same term. For example, the following three terms are identical



Our next goal is to identify the general principle behind these two observations.

First, we analyze under what conditions diagrams cancel. Here, one obtains the following important theorem

Linked cluster theorem:

All "disconnected diagrams" cancel in perturbation theory when correlation functions (or free energy) is calculated

Here, 'connected' and 'disconnected' refers to the Feynman diagrams. Are they fully connected (term 1,4,5,6,7 in Eq. (130)) or not (terms 2,3,8,9)?

Proof: We consider the expectation value $\langle X(\psi) \rangle$, and write $\langle \cdots \rangle^{\text{conn}}$ for fully connected diagrams. Using the Taylor expansion in \mathcal{S}_{int} , we obtain

$$\begin{split} \langle X(\psi) \rangle &= \frac{\left\langle X(\psi) \sum \frac{(-1)^n}{n!} (\mathcal{S}_{\rm int})^n \right\rangle_0}{\left\langle \sum \frac{(-1)^n}{n!} (\mathcal{S}_{\rm int})^n \right\rangle_0} = \frac{\sum_{n,k} \frac{(-1)^n}{n!} \binom{n}{k} \langle X(\psi) (\mathcal{S}_{\rm int})^k \rangle_0^{\rm conn} \langle (\mathcal{S}_{\rm int})^{n-k} \rangle}{\langle \cdots \rangle} \\ &= \frac{\sum_{k \le n} \frac{(-1)^k}{k!} \langle X(\psi) (\mathcal{S}_{\rm int})^k \rangle_0^{\rm conn} \frac{(-1)^{n-k}}{(n-k)!} \langle (\mathcal{S}_{\rm int})^{n-k} \rangle}{\langle \cdots \rangle} \\ &= \frac{1}{\langle \cdots \rangle} \sum_k \frac{(-1)^k}{k!} \langle X(\psi) (\mathcal{S}_{\rm int})^k \rangle_0^{\rm conn}}{k!} \sum_{\substack{k'=0 \\ \langle e^{-\mathcal{S}_{\rm int}} \rangle_0}} \frac{(-1)^{k'}}{k!} \langle (\mathcal{S}_{\rm int})^k \rangle_0^{\rm conn}} \\ &= \sum_k \frac{(-1)^k}{k!} \langle X(\psi) (\mathcal{S}_{\rm int})^k \rangle_0^{\rm conn}}{k!} \end{split}$$

This shows, that the disconnected terms in the numerator cancel exactly the terms from the Taylor expansion of the denominator.

As a next step, we have to work out **combinatorial prefactors**. For this, we consider the diagrams to order n, which contain n interaction lines and n integrals over time.



According to Wick's theorem, we have to consider all possible ways to connect the ingoing and outgoing lines of the diagram. After we have found a valid connected diagram, we can just exchange the integration variables and internal labels τ_1, \dots, τ_n . Doing so, produces n! diagrams with the same value. This factor n! cancels the factor 1/n! from the Taylor expansion of $e^{-S_{\text{int}}}$. Furthermore, we can exchange the top and bottom of each diagram. This gives 2^n equivalent diagrams. The factor 2^n is canceled by a factor $\left(\frac{1}{2}\right)^n$ form $S_{\text{int}} = \frac{1}{2} \iint \cdots$.

We conclude that **topologically distinct diagrams** carry a prefactor ± 1 .

To obtain the correct sign, we observe that all fermionic lines either go from the initial to the final point of the diagram or, alternatively, form closed loops. Let us first look at what types of contractions form a closed look. We recall the structure of our interaction

$$\mathcal{S}_{\text{int}} = \frac{1}{2} \int V(1-2)\bar{\psi}_1 \bar{\psi}_2 \psi_2 \psi_1 = \frac{1}{2} \int V(1-2)\bar{\psi}_1 \psi_1 \bar{\psi}_2 \psi_2,$$

where at each vertex (i.e., where an interaction line begins) has an incoming and an outgoing fermionic line. A closed loop has always the shape



and therefore arises from a contraction of the type

$$\bar{\psi}_1\psi_1\bar{\psi}_2\psi_2\bar{\psi}_3\psi_3\cdots\cdots\bar{\psi}_n\psi_n=(-1)\psi_n\bar{\psi}_1\psi_1\bar{\psi}_2\cdots\psi_{n-1}\bar{\psi}_n$$

Therefore each closed loop contributes a minus sign.

In contrast, consider a line which starts at the creation operator ψ_i^{\dagger} and ends at ψ_f correspinding to the diagram



Here the contractions have a different structure

$$\psi_{f}\bar{\psi}_{i}\bar{\psi}_{1}\psi_{1}\bar{\psi}_{2}\psi_{2}\cdots\cdots\bar{\psi}_{n}\psi_{n} = (-1)(-1)\psi_{f}\bar{\psi}_{1}\psi_{1}\bar{\psi}_{2}\cdots\psi_{n-1}\bar{\psi}_{n}\psi_{n}\bar{\psi}_{i}$$

and do not carry a minus sign. We are now ready to collect all of our results in a set of rules, the so-called Feynman rules, for drawing Feynman diagrams and for translating the diagrams into formulas.

Feynman rules for correlation functions

The setting

Feynman rules are used to compute correlation functions and other physical quantities in perturbation theory in the strength of interactions. We use the following setting: $S = S_0 + S_{\text{int}}$ where the action is either written in position/time or momentum/frequency space. α_i are band indices.

$$S_{0} = \sum_{\sigma} \int_{0}^{\beta} d\tau \int d^{d}x \, \bar{\Psi}_{\sigma}(x,\tau) (\partial_{\tau} + \frac{\nabla^{2}}{2m} + \mu + V(x)) \Psi_{\sigma}(x,\tau)$$

$$= \sum_{\sigma,\sigma'} \int_{0}^{\beta} d\tau d\tau' \int d^{d}x d^{d}x' \, \bar{\Psi}_{\sigma}(x,\tau) \left(-g_{0}^{-1}(x,\tau;x',\tau')\right) \Psi_{\sigma'}(x',\tau')$$

$$= \sum_{\kappa \in 1.BZ, \sigma, \alpha, \omega_{n}} \bar{\Psi}_{\kappa,\alpha,\sigma,\omega_{n}} (-(i\omega_{n} - (\epsilon_{\kappa}^{\alpha} - \mu)))) \Psi_{\kappa,\alpha,\sigma,\omega_{n}}$$

$$= \sum_{\kappa \in 1.BZ, \sigma,\alpha, \omega_{n}} \bar{\Psi}_{\kappa,\alpha,\sigma,\omega_{n}} \left(-(g^{0})_{\kappa,\alpha,\sigma}^{-1}(i\omega_{n})\right) \Psi_{\kappa,\alpha,\sigma,\omega_{n}}$$

$$S_{\text{int}} = \frac{1}{2} \sum_{\sigma\sigma'} \int d^{d}x \, d^{d}x' \, V(x - x') \bar{\Psi}_{\sigma}(x) \bar{\Psi}_{\sigma'}(x') \Psi_{\sigma'}(x') \Psi_{\sigma}(x)$$

$$= \frac{1}{2} \frac{T}{V_{\kappa,\kappa',q,\omega_{n},\omega_{n'},\Omega_{m},\sigma\sigma',\alpha_{i}}} V_{\alpha_{1}\alpha_{2}}^{\alpha_{4}\alpha_{3}}(q)$$

$$\times \bar{\Psi}_{\kappa+q,\alpha_{4},\sigma,\omega_{n}+\Omega_{m}} \bar{\Psi}_{\kappa'-q,\alpha_{3},\sigma',\omega_{n'}-\Omega_{m}} \Psi_{\kappa',\alpha_{2},\sigma',\omega_{n'}} \Psi_{\kappa,\alpha_{1},\sigma,\omega_{n}}$$

Above, we diagonalized the non-interacting problem in momentum space, which is only possible if the potential is periodic, $V(\boldsymbol{x}) = V(\boldsymbol{x} + \boldsymbol{R}_n)$ and thus $g_0^{-1}(\boldsymbol{x} + \boldsymbol{R}_n, \tau; \boldsymbol{x}' + \boldsymbol{R}_n, \tau')$.

The following Feynman rules have been derived using (i) a Taylor expansion in S_{int} , (ii) Wick's theorem which allows to express Gaussian integrals in terms of Green functions, (iii) the linked cluster theorem to eliminate disconnected diagrams, (iv) an evaluation of combinatorial prefactors of topologically equivalent diagrams & an analysis of sign occuring when reordering Grassmann variables.

The rules

For perturbation theory to order n, i.e., proportional to V^n follow the following recipe:

- 1) consider all connected, topologically inequivalent diagrams with n interaction lines $\checkmark \checkmark \checkmark$ and directed propagators \longrightarrow . Propagators begin/end at interaction lines and/or the creation/annihiliation operators of the time-ordered imaginary time correlation functions which is calculated. Spin is conserved at each vertex.
- 2a) position/time space:

Associate to each vertex a point in space-time. Replace \longrightarrow by $g_0(2,1) = g_0(x_2\tau_2, x_1\tau_1) = -\langle \Psi_2 \bar{\Psi}_1 \rangle$ and each $\bigvee \longrightarrow$ by $-V(2,1) = -V(x_2 - x_1)\delta(\tau_2 - \tau_1)$ (note: sign due to $e^{-S_{int}}!$)

2b) momentum/frequency space:

Associate with each line the momentum \mathbf{k} and a Matsubara frequency ω_n (fermionic for fermionic lines) and further quantum numbers (spin, band index, ...), such that at each vertex the momenta and frequencies sum to zero (momentum and energy conservation). Use arrows in the interaction lines to denote the direction of momentum flow. Replace

 $\overset{\mathbf{k}, \alpha \ \omega_{n}, \sigma}{\longrightarrow} = g_{\boldsymbol{k}\alpha\sigma}(i\omega_{n}) = \frac{1}{i\omega_{n} - (\epsilon_{\boldsymbol{k}}^{\alpha} - \mu)} \text{ and } \overset{\alpha_{i}}{\overset{\alpha_{i}}{\longrightarrow}} -V_{\alpha_{1}\alpha_{2}}^{\alpha_{i}\alpha_{3}}(\boldsymbol{q})$ where α_{1}, α_{2} and α_{3}, α_{4} are band indices of incoming and outgoing lines, respectively. For fermions $\omega_{n} = \frac{2\pi}{\beta}(n + \frac{1}{2})$ are fermionic Matsubara frequencies. Note that differences and sums of fermionic Matsubara frequencies are bosonic Matsubara frequencies (e.g. Ω_{m} in S_{int}).

- 3) Multiply for each Fermi loop (closed fermionic line without external vertices) by (-1).
- 4) Integrate/sum over all internal variables position/time space: $\int_0^\beta d\tau \int d^d x$ momentum/frequency space: $\frac{1}{\beta} \sum_{\omega_n} \int \frac{d^d k}{(2\pi)^d}$ spin sums: extra faktor 2 per Fermi-loop

For bosons (e.g. atoms with even number of electrons+protons+neutrons) the same rules apply: replace fermionic by bosonic Matsubara frequency and there is no factor (-1) for bosonic loops (for phonons and photons most often a different definition of the free Green's function is used).

The diagrammatic rules for the free energy F can for example be obtained from a coupling constant integration (see lecture). Similar rules apply for other situations. Note that for non-standard actions it is often useful to find your own Feynman rules. As our first example, let us expand the Green's function up to second order in perturbation theory. We obtain two diagrams in linear order, n = 1, and ten topologically different diagrams in second order, n = 2.



where the double line stands for our approximation of the full Green's function which contains the interaction corrections. We can compare that to a blind application of Wick's theorem, which would have given 153 terms instead. In the next section, we will discover that one actually has to compute only 4 terms of the 12 terms.

For now, we want to practice to translate the diagrams into formulas. We start with one example in position space

$$1 \longrightarrow \underbrace{f_{1'}}_{1'} \underbrace{f_{2'}}_{2'} = \int d\vec{x}_1' d\tau_1' d\vec{x}_2' d\tau_2' g_0(\vec{x}_1' \tau_1', \vec{x}_1 \tau_1) \left(-V(\vec{x}_1' - \vec{x}_2')\right) g_0(\vec{x}_2' \tau_2', \vec{x}_1' \tau_1') g_0(\vec{x}_2 \tau_2, \vec{x}_2' \tau_2')$$

Here is an other example in momentum space, where the application of the Feynman rules gives

Note the extra minus sign and the extra factor 2 because the diagram contained a loop.

Above, we have given Feynman rules for the case of electron-electron interactions. But one can formulate and derive Feynman rules for all types of other problems which have the structure $S = S_0 + S_{int}$ provided that S_0 is quadratic in the relevant fields. Below, we sketch a few examples.

• The scattering from a fixed potential is described by

$$\Delta S = \int_0^\beta d\tau d^3x \sum_\sigma U(\vec{x}) \bar{\psi}_\sigma(\vec{x}) \psi_\sigma(\vec{x})$$
$$= \sum_{\omega_n} \sum_\sigma \int \frac{d^3k}{(2\pi)^3} U_{\vec{q}} \psi^{\dagger}_{\vec{k}+\vec{q},\sigma}(i\omega_n) \psi^{\dagger}_{\vec{k},\sigma}(i\omega_n)$$

Diagrammatically, this can, e.g., be drawn as

 $\overset{v_{\vec{q}}}{\underset{\vec{k}}{\otimes}}$. In a macroscopic $\vec{k} \cdot \vec{k} + \vec{q}$

sample with disorder, a considerable simplification can be reached by averaging either over the position of the defects or a Gaussian-distributed random potential, which leads to another set of diagrammatic rules (not discussed here).

• When electron-phonon interactions or the interaction of electrons with light is considered, one has to treat actions of the type (written in momentum space)

$$\mathcal{S}_{\rm int} = \int_0^\beta \sum_{\vec{k}, \vec{q}} g_{\vec{k}, \vec{q}} \bar{\Psi}_{\vec{k} + \vec{q}}(\tau) \Psi_{\vec{k}}(\tau) (a_{\vec{q}}(\tau) + \bar{a}_{-\vec{q}}(\tau)) \, d\tau.$$

where $\bar{a}_{\vec{q}}$ is the complex field representing the bosonic creation operator of a phonon or photon. The function $g_{\vec{k},\vec{q}}$ parametrizes the strength and functional form of the electron-phonon or electron-photon coupling. In this case, vertices

describing the emission and absorption of the bosons have the form \checkmark or

, respectively, where the dashed line represents the Green's function of \checkmark

the boson and the vertex (=the point connecting the lines) gets the weight $g_{\vec{k},\vec{q}}$ depending on the momenta of the incoming fermionic line and the outgoing (or ingoing) bosonic line. Typical diagrams then have, for example, the form

• Similarly, one can derive Feynman diagrams for field theories with either complex or real fields (where the two cases have different combinatorial prefactors). If an action is, for example, given by

$$S = \int J \left(\nabla\phi\right)^2 + \frac{\alpha}{3}\phi^3 + \frac{\beta}{4}\phi^4,$$

a typical diagram may be draws as



where the lines correspond to the propagator of the Φ field, $\frac{1}{Jk^2}$, while triangle and square are proportional to the coupling constants α and β , $\Delta \equiv \alpha$, $\Box \equiv \beta$.

We end this section with a note on how to calculate the free energy $F = -\frac{1}{\beta} \ln(Z)$. Here, apparently only closed loops contribute



but the weight of each diagram is not 1 anymore, as we now have to do the Taylor expansion of the logarithm.

To derive the correct prefactors of each diagram, one can use the following trick. We start from

$$e^{-\beta F(\lambda)} = \int e^{-(S_0 + \lambda S_{\text{int}})},$$

where we have added a factor λ in front of S_{int} . Our goal is to determine $F(\lambda = 1)$. For this, we consider

$$\frac{\partial F}{\partial \lambda} = \frac{\partial}{\partial \lambda} \left(-\frac{1}{\beta} \ln(Z) \right) = \frac{1}{\beta} \frac{1}{Z} \int S_{\text{int}} e^{-(S_0 + \lambda S_{\text{int}})} = \frac{1}{\beta} \langle S_{\text{int}} \rangle$$
$$\implies F(\lambda = 1) = F_0 + \int_0^1 d\lambda \frac{1}{\beta} \langle S_{\text{int}} \rangle$$

Now, we can use our previously derived Feynmal rules to calculate $\langle S_{\text{int}} \rangle$. A term of order λ^n in perturbation theory for S_{int} , which contains n + 1 interaction lines, then simply gets the prefactor $\int_0^1 d\lambda \, \lambda^n = \frac{1}{n+1}$.

5.2 Self energy and Hartree-Fock approximation

Motivation: We have derived the Feynman rules for perturbation theory based on a Taylor expansion in the interactions, S_{int} . We will discuss below why this bruteforce Taylor expansion is almost never a good idea. In most cases, we will have to resum perturbation theory, effectively including an infinite number of diagrams. An important example of this approach will be discussed in the following. Let us start by taking a closer look at the formulas derived in the previous section. For example, consider one of the terms arising in linear-order perturbation theory.



We find that this term in the perturbative expansion is characterized by a double pole at $i\omega_n = \epsilon_{\vec{k}} - \mu$. This double pole is completely unphysical and things get worse when higher order of perturbation theory are considered.

To understand what is going on here, let us look at a simpler example, the term

$$\frac{1}{x-a}.$$

Here, a simply shifts the position of the pole from x = 0 to x = a. If we perform a Taylor-expansion in a to finite order, something bad is happening

$$\frac{1}{x-a} = \underbrace{\frac{1}{x} + \frac{a}{x^2} + \frac{a^2}{x^3}}_{\text{Singular at } x=0} + \cdots$$

We obtain apparent singularities at x = 0 which get worse and worse when we consider higher orders of perturbation theory. The exact expression, however, does not have any singularity at x = 0.

We are facing precisely the same problem for our Green's function. To solve the problem, we need a technique to reorganize the terms in the perturbation theory in a way which allows for the shift of poles. When we look at the perturbative expansion of the Green's function, we realize that there is a simple recurring pattern. We can write



where the symbol $\bigotimes \equiv$ is defined by the diagrams



Let us make this idea more precise. We define an object called **self-energy**.

Definition: Self energy $\Sigma_{\vec{k}}(i\omega_n) = \bigotimes$ = Sum of all "one-particle irreducible diagrams" = diagrams which do **not** decompose into 2 parts if one fermionic line — is cut

With this definition, we can write the exact Green's function as



The second line is a elegant way to rewrite the infinite sum as a finite sum. The reader is encourages to check that the first line can be recovered from the second line by iteratively inserting the definition for the exact Green's function again and again. This is the famous **Dyson equation** which in formulas reads

$$\mathcal{G}_{\vec{k}}(i\omega_n) = g^0_{\vec{k}}(i\omega_n) + g^0_{\vec{k}}(i\omega_n)\Sigma_{\vec{k}}(i\omega_n)\mathcal{G}_{\vec{k}}(i\omega_n)$$
$$\iff \mathcal{G}_{\vec{k}}(i\omega_n) = \left[\mathbb{1} - g^0_{\vec{k}}(i\omega_n)\Sigma_{\vec{k}}(i\omega_n)\right]^{-1}g^0_{\vec{k}}(i\omega_n)$$

where the second equation follows from the first one by solving for the exact Green's function $\mathcal{G}_{\vec{k}}(i\omega_n)$.

In a general multi-band system the Green's function and the self-energy is a matrix (in the band-index-space or in spin-space) and one has to perform a matrix inversion to compute the Green's function. However, if we consider only a single band (or a self-energy which is a diagonal matrix), we can write directly

$$\mathcal{G}_{\vec{k}}(i\omega_n) = \frac{1}{g_{\vec{k}}^0(i\omega_n)^{-1} - \Sigma_{\vec{k}}(i\omega_n)} = \frac{1}{i\omega_n - \varepsilon_{\vec{k}} - \Sigma_{\vec{k}}(i\omega_n)}$$

This is the version which we will use most often.

One can also use diagrams to write the same formula

$$= \frac{}{1 - } = \frac{1}{(- -)^{-1} - ()}$$

Up to now, all formulas including were exact, provided that the exact self-energy was used and provided that all sums properly converge. But the exact self-energy is unknown. The basic idea is now, that instead of using perturbation theory to compute \mathcal{G} , we use instead perturbation theory to compute the self-energy Σ . Plugging the approximate Σ into the Dyson equation the amounts to an infinite resummation of perturbation theory. Importantly, we thereby avoid the divergencies discussed at the beginning of this section and we are able to describe situations where interaction correction shift the location of poles.

We will start, by computing the self-energy to linear order in perturbation theory, n = 1.

$$\Sigma_{\vec{k}}(i\omega_n) = \underbrace{\left(\begin{matrix} k', i\omega'_n \\ q = 0 \end{matrix}\right)}_{\textbf{Hartree}} + \underbrace{\left(\begin{matrix} k', i\omega'_n \\ \vec{k'}, i\omega'_n \end{matrix}\right)}_{\textbf{Fock}} \\ = \int \frac{d^3k'}{(2\pi)^3} \frac{1}{\beta} \sum_{\omega'_n} g_{\vec{k}}(i\omega_n) \left(-V(0)(-1)(\overset{\text{spin}}{(2)} - V(\vec{k} - \vec{k'})) \right) \\ \stackrel{\uparrow}{\text{Fermi}}_{\text{loop}} \\ = \int \frac{d^3k'}{(2\pi)^3} n_F(\varepsilon_{\vec{k}} - \mu) \left(2V(0) - V(\vec{k} - \vec{k'}) \right)$$
(131)

Here we used our previously derived result that $\frac{1}{\beta} \sum_{\omega'_n} g_{\vec{k}}(i\omega_n) = n_F(\epsilon_{\vec{k}} - \mu)$. The two diagrams carry the names Hartree and Fock diagram. Both are independent of frequency ω_n which implies that we can interpret them as a change of the dispersion

$$\epsilon_{\vec{k}} \to E_{\vec{k}} = \epsilon_{\vec{k}} + \Sigma_{\vec{k}}.$$

The Hartree term describes that a single particle sees an effective potential from all the other particles, $V_{\text{eff}}(\vec{r}) = \int d^3 \vec{r}' V(\vec{r} - \vec{r}') n(\vec{r}')$, where $n(\vec{r}')$ is the local density of particles. In the simple example considered above (translationally invariant system with a single band), $n(\vec{r}') = 2 \int \frac{d^3k'}{(2\pi)^3} n_F(\varepsilon_{\vec{k}} - \mu)$ is a constant independent of space. Therefore also the effective potential is a constant in this case, which can be absorbed in a shift of a chemical potential. But in a more complicated setting (many bands or no translational symmetry), the Hartree correction becomes important. The Fock term encodes a quantum correction to the Hartree term, which takes the Pauli principle into account. To see this (and as a consistency check), let us consider fermions without spin (implying that we have to replace the factor 2 in the Hartree term by 1) and purely local interactions $V(\vec{r}) = U\delta(\vec{r})$ (or $V(\vec{r}_i - \vec{r}_j) = g\delta_{i,j}$ on a lattice). According to the Pauli-principle, two spinless Fermions are never at the same position. Therefore, a purely local interaction has exactly no effect. Indeed, for local interaction the Fourier transformation of the interaction potential is just a constant, $V(\vec{k}) = V(0) = \text{ const}$. From this we obtain that $V(0) - V(\vec{k} - \vec{k'}) = 0$. Hartree and Fock terms cancel exactly and the self-energy is zero as required by Pauli's principle for this specific model.

When looking at the Hartree-Fock equations and their interpretation, we realize a problem. The occupation of electrons is in Eq. (131) computed from $n_F(\epsilon_{\vec{k}} - \mu)$, where $\epsilon_{\vec{k}}$ is the dispersion of the non-interacting system. We do have, however, a better approximation for the dispersion in the interaction system, $E_{\vec{k}} = \epsilon_{\vec{k}} + \Sigma_{\vec{k}}^{\text{HF}}$. How can we improve our formulas, taking the 'corrected' dispersion into account? This can be achieved by replacing the non-interacting 'bare' Green's function in the formula for the self-energy by the interaction-corrected 'full' Green's function. If we do that within the Hartree-Fock scheme, we arrive at the **self consistent Hartree-Fock approximation**. Diagrammatically, this is written as

$$= \frac{1}{g_{\vec{k}}^0(i\omega_n)^{-1} - \Sigma_{\vec{k}}(i\omega_n)},$$

$$\Sigma_{\vec{k}}(i\omega_n) = \bigcirc + \swarrow + \bigcirc + \bigcirc + \bigcirc + \bigcirc + \frown + \bigcirc + \cdots$$

$$= \swarrow + \bigcirc + \bigcirc + \bigcirc + \bigcirc + \bigcirc + \bigcirc + \cdots$$

Using the full Green's function to compute the self-energy is equivalent to the resummation of infinitely many diagrams as shown in the last line, which is obtained by recursively using the definition of the full Green's function.

In formulas, the self-consistency equation can be written as

$$\Sigma_{\vec{k}}(i\omega_n) = E_{\vec{k}} - \varepsilon_{\vec{k}} = \int \frac{d^3k'}{(2\pi)^3} n_F(E_{\vec{k}} - \mu) \left(2V(0) - V(\vec{k} - \vec{k'})\right)$$
(132)

Here, $E_{\vec{k}}$ shows up twice on the left and right-hand side of the equation and it has to be calculated **self-consistently**. In practice this means that one starts with a guess of $E_{\vec{k}}$, computes the right-hand side with this guess and then one updates the value of $E_{\vec{k}}$. This is repeated until convergence is reached.

The self-consistent Hartree Fock approximation is an effective single-particle approximation, which means that it just leads to a modified band-structure but is not able to describe effects like the inelastic scattering of electrons (to be discussed
later). Like most 'simple' approximations it can be derived in many different ways. One useful way is to view it as an approximation of the interaction term. Here, one replaces the exact Hamiltonian by the Hartree-Fock Hamiltonian $\mathcal{H}_{\rm HF}$ which is quadratic in the creation and annihilation operators

$$\mathcal{H} = \mathcal{H}_{0} + \frac{1}{2} \sum_{\sigma,\sigma'} \int dx dx' V(x - x') \psi_{\sigma}^{\dagger}(x) \psi_{\sigma}(x) \psi_{\sigma'}^{\dagger}(x') \psi_{\sigma'}(x')$$

$$\approx \mathcal{H}_{\mathrm{HF}} = \mathcal{H}_{0} + \underbrace{\sum_{\sigma,\sigma'} \int dx dx' V(x - x') \psi_{\sigma}^{\dagger}(x) \psi_{\sigma}(x) \langle \psi_{\sigma'}^{\dagger}(x') \psi_{\sigma'}(x') \rangle_{\mathrm{HF}}}_{\mathrm{Hartree}}$$

$$- \underbrace{\sum_{\sigma,\sigma'} \int dx dx' V(x - x') \psi_{\sigma}^{\dagger}(x) \psi_{\sigma'}^{\dagger}(x') \langle \psi_{\sigma'}^{\dagger}(x') \psi_{\sigma}(x) \rangle_{\mathrm{HF}}}_{\mathrm{Hartree}}$$

Fock = exchange interaction

where $\langle \cdots \rangle_{\rm HF}$ is calculated self-consistently using $\mathcal{H}_{\rm HF}$. Thus, we can write

$$\mathcal{H}_{\rm HF} = H_0 + \int dx dx' \,\psi^{\dagger}_{\sigma}(x) \left(\bigodot + \bigodot \right) \psi_{\sigma'}(x')$$

One can show (we are not doing this here) that the ground state of \mathcal{H}_{HF} can also be obtained by a variational approach where one uses ground-states of non-interacting Hamiltonians (slater determinants) as an ansatz for the wave function

$$E_{\rm HF} = \min_{|\psi\rangle \in \rm non-interacting ground states} \langle \psi | \mathcal{H} | \psi \rangle$$

In this sense, the Hartree-Fock approximation is the 'best possible' effective singleparticle approximation as it comes closest to the true groundstate $\min_{\text{all } |\psi\rangle} \langle \psi | \mathcal{H} | \psi \rangle \leq E_{\text{HF}}$.

Outlook: While the Hartee-Fock equations are very simple to be solved in the singleband model used above, they are numerically highly demanding for a continuum system with 1/r Coulomb interactions and an external potential (like a solid, where the external potential of electrons comes from the ions). This is mainly due to the non-local nature of the Fock term, which makes the self-consistency equation to converge only very slowly. Therefore, one uses in such cases often simpler effective single-particle approximations, e.g., the *local density approximation* (LDA), which may be covered in the solid-state theory course, but not in this lecture.

A remarkable feature of the self-consistent Hartree-Fock equation is that – due to the infinite resummation of diagrams – it is able to describe spontaneous symmetry breaking. For example, for sufficiently strong interactions a magnetic solution can be obtained where $g_{\vec{k},\uparrow}(i\omega_n) \neq g_{\vec{k},\downarrow}(i\omega_n)$. We will discuss this in more detail within the QFT II lecture.

5.3 Retarded Green's functions and Lehmann represetation

Motivation: Using Matsubara frequencies and imaginary times is useful for doing calculations, but real experiments happen at real times with real frequencies. While

we postpone a complete discussion on how to compute experimental observables to a later chapter, we will find here a simple (but generally applicable) way to relate the computed quantities to measurable ones. This will also be important to find a physical interpretation of objects like the self-energy.

As a start, we go back to the definition of the time-ordered imaginary-time Green's function. Here, we pretend that we know the exact eigenstates $|n\rangle$ and eigenenergies E_n of the fully interacting system, $\tilde{\mathcal{H}}|n\rangle = E_n|n\rangle$ with $\tilde{\mathcal{H}} = \mathcal{H} - \mu \mathcal{N}$. Using those, we obtain

$$\mathcal{G}_{AB}(\tau) = -\langle \mathcal{T}A(\tau)B(0) \rangle \quad \text{e.g. } A = a_{\vec{k}}, B = a_{\vec{k}}^{\dagger}$$
$$= -\frac{1}{\mathcal{Z}} \operatorname{Tr} \left(e^{-(\beta - \tau)\tilde{\mathcal{H}}} A \mathbf{1} e^{-\tau \tilde{\mathcal{H}}} B \right)$$
$$= -\frac{1}{\mathcal{Z}} \sum_{n,m} e^{-(\beta - \tau)E_n} \langle n|A|m \rangle e^{-\tau E_m} \langle m|B|n \rangle$$

Next, we perform a Fourier transformation using either fermionic or bosonic Matsubara frequencies. Here we use

$$\int_{0}^{\beta} d\tau \, e^{(i\omega_n - \Delta E)\tau} = \frac{1}{i\omega_n - \Delta E} e^{(i\omega_n - \Delta E)\tau} \Big|_{0}^{\beta} = \frac{1}{i\omega_n - \Delta E} \left(\xi e^{-\Delta E\beta} - 1\right),$$

where, as always, $\xi = -1$ in the fermionic case. Thus, we obtain the so-called **Lehmann representation** of the correlation function

$$\mathcal{G}_{AB}(i\omega_n) = \int_0^\beta d\tau \, e^{i\omega_n \tau} \mathcal{G}_{AB}(\tau)$$

= $\frac{1}{\mathcal{Z}} \sum_{n,m} \frac{1}{i\omega_n - (E_m - E_n)} \left(e^{-\beta E_n} - \xi e^{-\beta E_m} \right) \langle n|A|m \rangle \langle m|B|n \rangle$

Now, we have to compare this to an object which is experimentally measurable. We will discuss in Sec. 6 the following problem. Consider a Hamiltonian of the form $H = H_0 + \lambda(t)B$ with some operator B. In this case, we will show that in the limit of small $\lambda(t)$ the change of $\langle A \rangle$ is linear in λ

$$\langle A(t) \rangle - \langle A(t) \rangle_{\lambda=0} = \int G^R_{AB}(t-t')\lambda(t')dt'.$$

Here, $G_{AB}^{R}(t)$ is the **retarded Green's function**. They are defined – for real times – for either bosonic or fermionic operators by

$$\mathcal{G}_{AB}^{R}(t) = -\frac{i}{\hbar}\Theta(t) \begin{cases} \langle [A(t), B(0)] \rangle & \text{bosonic} \\ \langle \{A(t), B(0)\} \rangle & \text{fermionic} \end{cases}$$

with $A(t) = e^{it\tilde{\mathcal{H}}/\hbar}Ae^{-it\tilde{\mathcal{H}}/\hbar}$, $\tilde{\mathcal{H}} = \mathcal{H} - \mu \mathcal{N}$ using the standard real-time Heisenberg picture (whether or not to include $-\mu \mathcal{N}$ in the time-evolution operator is a question

of definition here). We call the correlation function **retarded** because of the thetafunction $\theta(t)$, which guarantees that the operator A is evaluated always after the operator B. Here, we call an operator fermionic (bosonic) if it includes an odd (even) number of fermionic operators. **Warning:** Note that we often set $\hbar = 1$ in this script and thus factors of \hbar are omitted in the following.

An example of a bosonic operator in an electronic system is the electric current and we will see in Sec. 6 that the electric conductivity can be obtained from the retarded correlation function. Similarly, a neutron-scattering experiment measures, for example, retarded spin correlation functions.

Now, we use the same trick as above to write the retarded Green's function in terms of the exact many-body eigenstates of the Hamiltonian

Next, we apply a Fourier transformation (using standard, real-valued frequencies) and obtain the Lehmann representation of a retarded correlation function

$$\begin{aligned} \mathcal{G}_{AB}^{R}(\omega) &= \int_{-\infty}^{\infty} dt \, e^{i(\omega+i\epsilon)t} \mathcal{G}_{AB}^{R}(t) \\ &= \frac{1}{\mathcal{Z}} \sum_{n,m} \frac{1}{\omega - (E_m - E_n) + i\epsilon} \langle n|A|m \rangle \langle m|B|n \rangle \left(e^{-\beta E_n} - \xi e^{-\beta E_m} \right) \end{aligned}$$

where, importantly, the factor $i\epsilon$ is needed to obtain the correct $\theta(t)$ function on the time axis. This can be seen most easily by considering

$$\int_{-\infty}^{\infty} \frac{1}{\omega - E + i\epsilon} e^{-i\omega t} \frac{d\omega}{2\pi} = -i\theta(t)e^{-iEt}$$

This is evaluated by contour integration. For t > 0 (t < 0) one can close the contour on the lower (upper) half-plane, where $e^{-izt} \propto e^{-i(i\text{Im}z)t} = e^{-\text{Im}z)t} \to 0$ for Im $z \to \mp \infty$. Then, one uses the residue theorem. As for $\epsilon > 0$ the pole is in the lower half plane at $z = E_0 - i\epsilon$, only t > 0 contributes.

Comparing the two expressions for the retarded correlation function $\mathcal{G}_{AB}^{R}(\omega)$ and the time-ordered imaginary-time correlation function $\mathcal{G}_{AB}(i\omega_{n})$, we see that the formulas are identical upon replacing $i\omega_{n} \longrightarrow \omega + i\epsilon$.

$$\lim_{\substack{\epsilon \to 0\\\epsilon > 0}} \mathcal{G}_{AB}(i\omega_n \longrightarrow \omega + i\epsilon) = \mathcal{G}^R_{AB}(\omega)$$
(133)

Thus, retarded Green's function are obtained by a simple **analytic continuation** $i\omega_n \longrightarrow \omega + i\epsilon$. This program works for both fermionic and bosonic correlation functions.

A very useful concept is to define $\mathcal{G}_{AB}(z)$, $z \in \mathbb{C}$ in the full *complex plane* by replacing $i\omega_n \to z$. Then, we obtain a function which is analytic both in the upper and lower complex plane where poles or a branch cut (if the E_m are dense) on the real axis only.

Above, we discussed general properties of correlation functions. Now, we will focus on Green's functions, choosing, for example, $B = a_{\vec{k}}^{\dagger}$ and $A = a_{\vec{k}}$

$$g_{\vec{k}}(\tau) = -\langle \mathcal{T}a_{\vec{k}}(\tau)a_{\vec{k}}^{\dagger}(0) \rangle$$

The **spectral function** is defined as the imaginary part of the retarded correlation function times a factor -2. In the case of fermions, the Lehmann representation is given by

$$A_{\vec{k}}(\omega) = -2\mathrm{Im}(\mathcal{G}_{\vec{k}}^{R}(\omega))$$

= $2\pi \frac{1}{\mathcal{Z}} \sum_{n,m} |\langle m|a_{\vec{k}}^{\dagger}|n\rangle|^{2} \left(e^{-\beta E_{m}} + e^{-\beta E_{n}}\right) \delta(\omega - (E_{m} - E_{n}))$
 $\mathrm{Im}\left(\frac{1}{\omega+i0}\right) = -i\pi\delta(\omega)$

Outlook: We will show later that **angular-resolved photoemission experiment** measure the product of Fermi function and spectral function, $n_F(\omega)A_{\vec{k}}(\omega)$. Furthermore, one can use scanning tunneling microscopy to measure $\sum_{\vec{k}} A_{\vec{k}}(\omega)$.

By definition, the spectral function is always positive, $A_{\vec{k}}(\omega) \ge 0$ for all \vec{k}, ω . Furthermore, the integral over frequencies is normalized to 1

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} A_{\vec{k}}(\omega) = g_{\vec{k}}^R(t=0^+) = \langle \{a_{\vec{k}}, a_{\vec{k}}^\dagger\} \rangle = 1$$

Thus, we can interpret $A_{\vec{k}}(\omega)$ as a probability. It is roughly the probability that a single-particle excitation with momentum \vec{k} has the energy ω . This can be seen most easily for T = 0 In this case $e^{-\beta E_n}$ projects onto the ground-state $|0\rangle$ of the interacting system with energy $E_0 < E_n$. We obtain in this case

$$A_{\vec{k}}(\omega)|_{T=0} = \begin{cases} \sum_{n} |\langle n|a_{\vec{k}}^{\dagger}|0\rangle|^2 \, 2\pi\delta(\omega - (E_n - E_0)) & \omega > 0 \quad \text{create particles} \\ \sum_{n} |\langle n|a_{\vec{k}}|0\rangle|^2 \, 2\pi\delta(\omega + (E_n - E_0)) & \omega < 0 \quad \text{create holes} \end{cases}$$

Here for $\omega > 0$, $|\langle n | a_{\vec{k}}^{\dagger} | 0 \rangle|^2$ is the probability that the state $a_{\vec{k}}^{\dagger} | 0 \rangle$ is an eigenstate with an energy $\omega = E_n - E_0$ above the ground-state.

In the absence of interactions, the (time-ordered imaginary-time) Green's function is $\frac{1}{i\omega_n - (\varepsilon_{\vec{k}} - \mu)}$ and thus the spectral function is simply given by

$$A_{\vec{k}}(\omega) = 2\pi\delta(\omega - (\varepsilon_{\vec{k}} - \mu))$$

This describes that an excitations with momentum \vec{k} has a fixed energy $\varepsilon_{\vec{k}}$. Later, we will explore how this changes when interactions are taken into account.

We close this section by discussing some useful properties of spectral functions in the complex plane. As can be seen directly from the Lehmann representation, we can use the spectral function to compute the Green's function in the full complex plane

$$g_{\vec{k}}(z) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \, \frac{A_{\vec{k}}(\omega)}{z - \omega} \tag{134}$$

where, as discussed above, the retarded Green's function is given by $g_{\vec{k}}^R(\omega) = g_{\vec{k}}(\omega + i\epsilon)$. Sometimes, it is also useful to define the 'advanced' Green's function $g_{\vec{k}}^A(\omega) = g_{\vec{k}}(\omega - i\epsilon)$.

The Green's function is (as mentioned above) analytic both in the upper and lower complex plane but has a branch cut on the real axis, as sketched in the figure.



As $A_{\vec{k}}(\omega)$ is real, we have

$$g_{\vec{k}}(\omega - i\epsilon) = g_{\vec{k}}(\omega + i\epsilon)^*$$

and thus the imaginary part of $g_{\vec{k}}(z)$ jumps when one crosses the branch cut, Im $g_{\vec{k}}^R(\omega + i\epsilon) = -\text{Im} g_{\vec{k}}(\omega - i\epsilon)$ and $g_{\vec{k}}(\omega + i\epsilon) - g_{\vec{k}}(\omega - i\epsilon) = -iA_{\vec{k}}(\omega)$. A useful relation is

$$\frac{1}{\omega + i\epsilon} = \frac{\omega}{\omega^2 + \epsilon^2} - i\frac{\epsilon}{\omega^2 + \epsilon^2} = PV\left(\frac{1}{\omega}\right) - i\pi\delta(\omega)$$

where PV stands for 'principle value' and serves as a reminder that for the purpose of integration one should replace $1/\omega$ by $\frac{\omega}{\omega^2 + \epsilon^2}$, taking the limit $\epsilon \to 0$ after integration. With this notation, one can compute the real part of the Green's function from its imaginary part and vice versa

$$\operatorname{Re} g_{\vec{k}}^{R}(\omega) = \int_{\mathrm{PV}} \frac{d\omega'}{\pi} \frac{\operatorname{Im} g_{\vec{k}}^{R}(\omega')}{\omega' - \omega}$$
$$\operatorname{Im} g_{\vec{k}}^{R}(\omega) = -\int_{\mathrm{PV}} \frac{d\omega'}{\pi} \frac{\operatorname{Re} g_{\vec{k}}^{R}(\omega')}{\omega' - \omega}$$

This is an example of a **Kramers-Kronig relation**, which also applies to other retarded correlation functions like dielectric functions, magnetic susceptibilities, optical conductivities discussed in Sec. 6. Such relations are frequently used, e.g., in experiments to construct complex functions when only real- or imaginary parts have been measured.

Outlook: In this section, we have seen that we can use analytic continuation $i\omega_n \rightarrow \omega + i\epsilon$ to obtain experimentally measurable frequency-dependent quantities from our imaginary-time formalism. In the chapter on linear-response theory, chapter 6, we will develop a more general theory of this effect.

While it is simple to do the analytic continuation in analytic calculations, it is a major problem is settings where $\mathcal{G}(i\omega_n)$ is computed numerically. In this case, one knows this function only for a finite number of frequencies ω_n , often with a numerical error. Even if the numerical error is small and the numbers of frequencies is large, there is not a unique analytical continuation. One possible strategy of how to address the problem is called 'maximum entropy' and is based on the idea that one searches for the smoothest spectral function on the real axis which is consistent with the numerically obtained result of the Green's function at imaginary Matsubara, $\mathcal{G}(i\omega_n)$. The two quantities are related by Eq. (134).

5.4 Quasi particles

Motivation: In the previous section, we have shown that the spectral function $A_{\vec{k}}(\omega)$ gives the probability that a fermionic excitation with momentum \vec{k} has the frequency ω . Our next goal is to understand how interactions affect this quantity. We will start with some general considerations on how the self-energy affects the spectral function but ultimately we want to discuss one of the most powerful concepts in physics: the concept of a quasi-particle, an effective excitation arising in an interacting system.

For a single-band model the Green's function in the presence of interaction is given by

$$\mathcal{G}_{\vec{k}}^{R}(\omega) = \frac{1}{\omega + i\epsilon - \left(\varepsilon_{\vec{k}} - \mu + \Sigma_{\vec{k}}^{R}(\omega)\right)} \quad \text{with } \Sigma_{\vec{k}}^{R}(\omega) = \Sigma_{\vec{k}}(\omega + i\epsilon)$$

and therefore the spectral function is given by

$$A_{\vec{k}}(\omega) = -2\operatorname{Im} \mathcal{G}_{\vec{k}}^{R}(\omega) = -2\frac{\operatorname{Im} \Sigma_{\vec{k}}^{R}(\omega)}{\left(\omega - \left(\varepsilon_{\vec{k}} - \mu + \operatorname{Re} \Sigma_{\vec{k}}^{R}(\omega)\right)\right)^{2} + \left(\operatorname{Im} \Sigma_{\vec{k}}^{R}(\omega)\right)^{2}}$$

An extra $i\epsilon$ is not needed as the self-energy will generically have its own imaginary part.



Our goal is to find and discuss possible sharp peaks $A_{\vec{k}}(\omega)$ in the spectral function as function of ω . Such a peak indicates that an excitation with momentum \vec{k} has a more-or-less well defined energy at the position $E_{\vec{k}}$ of the peak. If a well-defined peak exist, we call $E_{\vec{k}}$ the **quasi-particle energy**. To obtain a peak, the denominator of the spectral function has to be small. This happens under two conditions, first $\operatorname{Im} \Sigma_{\vec{k}}^{R}(\omega)$ has to be small (this will be discussed below) and the first term in the denominator has to vanish.

$$E_{\vec{k}} - \left(\varepsilon_{\vec{k}} - \mu + \operatorname{Re}\Sigma_{\vec{k}}(E_{\vec{k}})\right) = 0 \tag{135}$$

To get more information on the peak, we Taylor-expand the Green's function around $\omega \approx E_{\vec{k}}$

$$\mathcal{G}_{\vec{k}}^{R}(\omega) \approx \frac{1}{\alpha(\omega - E_{\vec{k}}) + i\beta} = \frac{Z_{\vec{k}}}{\omega - E_{\vec{k}} + i\Gamma_{\vec{k}}}$$
$$A_{\vec{k}}^{R}(\omega) = -2\mathrm{Im}\mathcal{G}_{\vec{k}}^{R}(\omega) \approx Z_{\vec{k}}\frac{2\Gamma_{\vec{k}}}{(\omega - E_{\vec{k}})^{2} + \Gamma_{\vec{k}}^{2}}$$
(136)

with the quasi-particle weight

$$Z_{\vec{k}} = \frac{1}{\alpha} = \left(1 - \left. \frac{\partial}{\partial \omega} \operatorname{Re} \Sigma_{\vec{k}}(\omega) \right|_{\omega = E_{\vec{k}}} \right)^{-}$$

and the quasi-particle scattering rate

$$\Gamma_{\vec{k}} = -Z_{\vec{k}} \operatorname{Im} \Sigma_{\vec{k}} (E_{\vec{k}})$$

The three quantities, $E_{\vec{k}}, Z_{\vec{k}}, \Gamma_{\vec{k}}$, describe the quasi-particle peak in the spectral function. $E_{\vec{k}}$ is the **position** of the peak, $\Gamma_{\vec{k}}$ is its **width**, and $Z_{\vec{k}}$ it its **weight**, defined by taking the integral close to the peak only, $Z_{\vec{k}} \approx \int_{\text{peak}} \frac{d\omega}{2\pi} A_{\vec{k}}(\omega)$. The height of the peak is thereby $2Z_{\vec{k}}/\Gamma_{\vec{k}}$. The scattering rate determines how fast the Green function decays in time, $G_{\vec{k}}(t) \approx Z_{\vec{k}}e^{-iE_{\vec{k}}t}e^{-\Gamma_{\vec{k}}t}$.

Below, we will show that $\Gamma_{\vec{k}}$ becomes very small at low temperatures and close to the Fermi energy. Thus the peak becomes very sharp because there exist an excitation

with a long lifetime. These excitations, the **quasi particles**, are best suited describe the interacting many-particle systems. they are the relevant low-energy excitation arising from the interplay of the band-structure **and** the interactions.

The quasi-particle velocity, $\tilde{v}_{\vec{k}} = \frac{\partial E_{\vec{k}}}{\partial \vec{k}}$, is computed by taking the derivative of Eq. (135).

$$\frac{\partial E_{\vec{k}}}{\partial \vec{k}} - \left(\frac{\partial \varepsilon}{\partial \vec{k}} + \frac{\partial}{\partial \vec{k}} \operatorname{Re} \Sigma_{\vec{k}}(E_{\vec{k}}) + \frac{\partial}{\partial \omega} \operatorname{Re} \Sigma_{\vec{k}}(\omega) \Big|_{\omega = E_{\vec{k}}} \cdot \frac{\partial E_{\vec{k}}}{\partial \vec{k}} \right) = 0$$

which is solved to give

$$\tilde{v}_{\vec{k}} = \frac{\partial E_{\vec{k}}}{\partial \vec{k}} = Z_{\vec{k}} \left(\frac{\partial \varepsilon}{\partial \vec{k}} + \frac{\partial}{\partial \vec{k}} \operatorname{Re} \Sigma_{\vec{k}}(E_{\vec{k}}) \right)$$

It is customary to discuss the velocity in terms of the **effective mass** m^* of quasi particles, obtained from the bare electron mass m and the ratio of the quasi-particle velocity and the bare velocity $\vec{v}_{\vec{k}} = \frac{\partial \varepsilon}{\partial \vec{k}}$

$$m^* = m \frac{|\vec{v}_{\vec{k}}|}{|\tilde{v}_{\vec{k}}|} = m \frac{1}{Z_{\vec{k}}} \frac{|\vec{v}_{\vec{k}}|}{|\vec{v}_{\vec{k}} + \frac{\partial}{\partial \vec{k}} \operatorname{Re} \Sigma_{\vec{k}}(E_{\vec{k}})|}$$

where \vec{k} is evaluated at the Fermi surface.

Outlook: The effective mass of quasi-particle can be vastly different from the effective mass of bare electrons. The most extreme case are called heavy-fermion systems like, for example, CeCu₆, where the quasi-particle weight is tiny, $Z_{\vec{k}} \sim 10^{-3}$, due to strong interaction effects and thus $\frac{m^*}{m} \sim 1000$ which implies that quasi particles are $1000 \times$ slower and that quantities like the specific heat are 1000 fold enhanced (at low T) compared to the specific heat of ordinary metals.

A useful way to think about the quasi-particle weight is to view it as measuring the square of a quasi-particle wave function and the wave function of an electron in the absence of interactions. Equivalently, one can expand the operators formally in the following way,

$$\begin{array}{ccc} a^{\dagger}_{\vec{k},\sigma} \approx \sqrt{Z_{\vec{k}}} & c^{\dagger}_{\vec{k},\sigma} & + \sum A^{\vec{k}}_{\vec{k}_{1}\sigma,\vec{k}_{2}\sigma',\vec{k}_{3}\sigma'} c^{\dagger}_{\vec{k}_{1}\sigma} c^{\dagger}_{\vec{k}_{2}\sigma'} \delta_{\vec{k}_{1}+\vec{k}_{2}-\vec{k}_{3}-\vec{k}} + \cdots \\ \uparrow & \uparrow & \uparrow & \uparrow \\ \text{creates} & \text{creates} & \text{creates} & \text{quasi} \\ \text{electron} & \text{quasi particle} & \text{hole} \end{array}$$

One way to measure the quasi-particle weight is to consider the occupation function of electrons in a metal at T = 0

$$n_{\vec{k}} = \langle a_{\vec{k}}^{\dagger} a_{\vec{k}} \rangle = \frac{1}{\beta} \sum_{\omega_n} g_{\vec{k}}(i\omega_n) = -\oint \frac{dz}{2\pi i} n_F(z) g_{\vec{k}}(z)$$
$$= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} n_F(z) A_{\vec{k}}(\omega) \stackrel{=}{\underset{T=0}{\uparrow}} \int_{-\infty}^{0} \frac{d\omega}{2\pi} A_{\vec{k}}(\omega)$$

As we will show below, the quasi-particle scattering rate vanishes, $\Gamma_{\vec{k}} = 0$, at the Fermi momentum $\vec{k} = \vec{k}_F$, where $E_{\vec{k}_F} = 0$. This implies that the spectral function

obtains an infinite simally sharp peak, $A_{\vec{k}}(\omega) \approx Z_{\vec{k}} 2\pi \delta(\omega - E_{\vec{k}})$ which changes from negative to positive energies when the Fermi surface is crossed. This leads to a jump in $n_{\vec{k}}$ with the amplitude $Z_{\vec{k}_F}$ as shown in the figure.



$$Z_{\vec{k}} = \text{jump of } n_{\vec{k}} \text{ at } \vec{k} = \vec{k}_F$$

Outlook: The theory of interacting quasi-particles goes under the name **Fermi liq-uid theory** (or Landau's Fermi liquid theory, after Lev Landau) but is not covered in this course. Its starting point is to first ignore the scattering of quasi particles, $\Gamma_{\vec{k}} \approx 0$. In this limit, the density of quasi-particle excitations at fixed momentum \vec{k} is conserved, $[\mathcal{H}_{\rm QP}, \tilde{n}_{\vec{k},\sigma}] = 0$ and one can write down an effective quasi-particle Hamiltonian

$$\mathcal{H}_{\rm QP} \approx \sum_{\vec{k},\sigma=\uparrow,\downarrow} E_{\vec{k}} \tilde{n}_{\vec{k},\sigma} + \sum_{\substack{\vec{k},\vec{k'}\\\sigma,\sigma'}} f_{\vec{k},\vec{k'}}^{\sigma,\sigma'} \tilde{n}_{\vec{k},\sigma} \tilde{n}_{\vec{k'},\sigma'} + \cdots$$

where the first term is given by the quasi-particle energy identified above, while the other terms encode interaction effects. Fermi liquid theory (which in its full form also contains an equation describing the scattering of quasi particles) allows to calculate exactly some of the low-temperature properties of metals based on a small number of fitting parameters.

All of the discussion given above relies on the assumption that the scattering rate of quasi particles is small. Therefore, we will now calculate this central quantity. As the calculation is so important, we will do it twice: first, we will use a set of qualitative arguments and then (in a more sketchy second part) show how this can be seen by calculating the relevant Feynman diagrams.

For the qualitative argument, we consider an initial state at where a single quasi particle with $E_1 > 0$ (zero corresponds to the Fermi energy in our notations) starts to interact with the quasi-particles with energies E < 0 forming the Fermi sea of occupied states in the ground state at T = 0. Next, we consider a single scattering process, see figure, where the quasi-particle loses energy, $E_1 \rightarrow E'_1$ and another particle gains energy, $E_2 \rightarrow E'_2$.



From our initial conditions, we obtain $E_2 < 0$ and the Pauli principle enforces that $0 < E'_1 < E_1$ and $0 < E'_2$. Combining this with energy conservation, $E_2 = E'_1 + E'_2 - E_1$, we obtain $0 > E_2 > E'_1 - E_1$. Now, we count how many energies are available for scattering

$$\int_0^{E_1} dE_1' \int_{-(E_1 - E_1')}^0 dE_2 = \frac{1}{2} E_1^2.$$

The phase-space for scattering is reduced proportional to E_1^2 , when E_1 gets closer to the Fermi energy 0. This strongly suggest that the scattering rate is proportional to E_1^2

$$\Gamma_{\vec{k}} \propto \begin{cases} E_{\vec{k}}^2 & \text{for } T \ll E_{\vec{k}} \ll \varepsilon_F \\ T^2 & \text{for } E_{\vec{k}} \ll T \ll \varepsilon_F \end{cases}$$
(137)

In the second line, we guessed the result for finite temperatures. T has the effect to smear out the sharp Fermi surface, creating extra excitations with a typical energy of $E_{\vec{k}} \sim T$ (we use units where $k_B = \hbar = 1$). Thus, one can expect that one can replace the relevant energies by T.

Two effects have not been taken into account above. First, while energy conservation was essential for our argument, we ignored momentum conservation. In d > 1momentum conservation is easily fulfilled and - as it turns out - does not lead to any qualitative changes of the result. Furthermore, we only considered scattering of two particles. What happens if n > 2 particles are involved? In this case, the phase space is proportional to E_1^{2n-2} and thus one can neglect this effect for sufficiently small E_1 or temperatures. In conclusion, our qualitative argument strongly suggests that the scattering rate becomes very small close to the Fermi surface and for low T.

Next, we want to check this result using diagrams. We have to compute the imaginary part of the self-energy



As Hartree and Fock diagrams do not have any imaginary parts, we focus on two diagrams arising in second-order perturbation theory



Using the Feynman rules, these pictures translate into the following formula

$$\delta\Sigma_{\vec{k}}(i\Omega) = \frac{1}{\beta} \sum_{i\omega_1} \frac{1}{\beta} \sum_{i\omega_2} \int \frac{d^3\vec{k}_1}{(2\pi)^3} \int \frac{d^3\vec{k}_2}{(2\pi)^3} g_{\vec{k}_1}(i\omega_1) g_{\vec{k}_2}(i\omega_2) g_{\vec{k}+\vec{k}_2-\vec{k}_1}(i\Omega+i\omega_2-i\omega_1) \\ \cdot \left(V(\vec{k}-\vec{k}_1)V(\vec{k}_1-\vec{k}_2) - 2\left(V(\vec{k}-\vec{k}_1)\right)^2 \right)$$
(138)

We first focus on the frequency summations and observe that the sums have the following structure

$$A(i\Omega) = \frac{1}{\beta} \sum_{i\omega_1} g_1(i\omega_1)\chi_{23}(i\Omega - i\omega_1)$$
$$\chi_{23}(i\omega_B) = \frac{1}{\beta} \sum_{i\omega_2} g_2(i\omega_2)g_3(i\omega_B + i\omega_2)$$

where $i\omega_B$ is the difference of two fermionic Matsubara frequencies and therefore a bosonic Matsubara frequency.



Figure 10: Pole structure for above equations. Red: branch cuts, thin blue lines: final contour which runs directly above and below the two branch cuts.

Let us first focus on $A(i\Omega)$, where we use the our standard trick to write the sum as a contour integral. Now, we have to observe that $g_1(z)$ has poles or a branch cut on the real axis (lower red line in the figure), Im z = 0. In contrast, $\chi_{23}(i\Omega - z)$ has branch cuts or poles at $\text{Im } z = i\Omega$ (upper red line). But as shown in the figure, we can deform the contour in such a way, that it runs directly above and below the two branch cuts. One pair of paths can be parametrized by $z = \pm \omega \pm i\epsilon$, the other one by $z = i\Omega \pm \omega \pm i\epsilon$, where ω runs from $-\infty$ to ∞ . Using that $g_1(w + i\epsilon) - g_1(\omega - i\epsilon) = 2i\text{Im } g_1^R(\omega + i\epsilon)$ and the anlog equation for χ_{23} , we obtain

$$A(i\Omega) = -\oint \frac{dz}{2\pi i} n_F(z) g_1(z) \chi_{23}(i\Omega - z)$$

= $-\int \frac{d\omega'}{\pi} n_F(\omega') \operatorname{Im} [g_1(\omega' + i\epsilon)] \chi_{23}(i\Omega - \omega')$
 $+ n_F(\omega' + i\Omega) g_1(\omega' + i\Omega) \operatorname{Im} [\chi_{23}(-(\omega' + i0))]$

Next, we observe that $n_F(\omega' + i\Omega) = \frac{1}{e^{\beta(\omega'+i\Omega)}+1} = \frac{1}{-e^{\beta\omega'}+1} = -n_B(\omega')$ as $\Omega = \frac{2\pi}{\beta} \left(n + \frac{1}{2}\right)$ is a fermionic Matsubara frequency. Furthermore, $\operatorname{Im} \chi(-\omega' - i\varepsilon) = \frac{2\pi}{\beta} \left(n + \frac{1}{2}\right)$

 $-\text{Im}\,\chi(-\omega'+i\varepsilon)$. The two minus signs cancel and thus we find that

$$A(i\Omega) = -\int \frac{d\omega'}{\pi} n_F(\omega') \operatorname{Im} \left[g_1(\omega' + i\epsilon) \right] \chi_{23}(i\Omega - \omega') + n_B(\omega') g_1(\omega' + i\Omega) \operatorname{Im} \left[\chi_{23} \left(-\omega' + i0 \right) \right]$$

Next, we perform the analytic continuation $i\Omega \to \Omega + i\epsilon$. It is important, that we do this step only **after** we are done with shifting contours.

$$A(\Omega + i\epsilon) = -\int \frac{d\omega'}{\pi} n_F(\omega') \operatorname{Im} \left[g_1(\omega' + i\epsilon)\right] \chi_{23}(\Omega - \omega' + i\epsilon) + n_B(\omega') g_1(\omega' + \Omega + i\epsilon) \operatorname{Im} \left[\chi_{23}\left(-\omega' + i0\right)\right]$$

Finally, we shift the integration variable in the second term $\omega' \to \omega' - \Omega$ and take the imaginary part of the full expression to arrive at

$$\operatorname{Im} A(\Omega + i0) = -\int \frac{d\omega}{\pi} \left(n_F(\omega) + n_B(\omega - \Omega) \right) \operatorname{Im} g_1(\omega + i\epsilon) \operatorname{Im} \chi_{23}(\Omega - \omega' + i\epsilon)$$
(139)

From our previous argument, we expect that the imaginary part of the self-energy vanishes for T = 0 and $\Omega = 0$. This, we can already see at this stage of the calculation, as for $T \to 0$, we find that $n_F(\omega) + n_B(\omega) \approx 0$ as $n_B(\omega < 0) \to -1$ for $T \to 0$. Using exactly the same strategy, we can also show that

$$\operatorname{Im} \chi_{23}(\omega_B + i\varepsilon) = -\int \frac{d\omega'}{\pi} \left(n_F(\omega') - n_F(\omega' + \omega_B) \right) \operatorname{Im} g_2(\omega' + i\varepsilon) \operatorname{Im} g_3(\omega' + \omega_B + i\varepsilon)$$
(140)

This quantity vanishes for $\omega_B \to 0$.

Finally, we use that $\operatorname{Im} g_{\alpha}(\omega + i\epsilon) = -\pi\delta(\omega - \varepsilon_{\alpha})$ (we use a notation, where the chemical potential is included in the definition of the energy). Combining Eqs. (139) and (140), we see that $\operatorname{Im} A(\Omega + i0)$ is therefore proportional to a product of three δ -functions. With two of them, we do the integrations over ω and ω' . Now, we combine everything to compute the imaginary part of the self-energy, $\delta \Sigma_{\vec{k}}(\omega + i\epsilon)$ using Eq. (138). Evaluating the self energy at $\omega = \epsilon_{\vec{k}}$, which is the energy of the particle, which we want to study, we obtain

$$\operatorname{Im} \delta\Sigma_{\vec{k}}(\varepsilon_{\vec{k}}) = -\int \frac{d^{3}k_{1}}{(2\pi)^{3}} \frac{d^{3}k_{2}}{(2\pi)^{3}} \frac{d^{3}k_{3}}{(2\pi)^{3}} \delta\left(\vec{k} + \vec{k}_{2} - (\vec{k}_{1} + \vec{k}_{3})\right) \delta(\varepsilon_{\vec{k}} + \varepsilon_{\vec{k}_{2}} - (\varepsilon_{\vec{k}_{1}} + \varepsilon_{\vec{k}_{3}})) \\ \cdot \left(2\left(V(\vec{k} - \vec{k}_{1})\right)^{2} - V(\vec{k} - \vec{k}_{1})V(\vec{k}_{1} - \vec{k}_{3})\right) \cdot \left(\underbrace{n_{F}(\varepsilon_{\vec{k}_{2}})\left(1 - n_{F}(\varepsilon_{\vec{k}_{1}})\right)\left(1 - n_{F}(\varepsilon_{\vec{k}_{3}})\right)}_{\text{first term}} + \underbrace{n_{F}(\varepsilon_{\vec{k}_{1}})n_{F}(\varepsilon_{\vec{k}_{3}})\left(1 - n_{F}(\varepsilon_{\vec{k}_{2}})\right)}_{\text{second term}}\right) \right)$$
(141)

To write the equation above in a form which is easy to interpret, the following identity has been used

$$n_B(\varepsilon_1 - \varepsilon_2) \left(n_F(\varepsilon_2) - n_F(\varepsilon_1) \right) = n_F(\varepsilon_1) \left(1 - n_F(\varepsilon_2) \right),$$

which simply follows from the formulas for the Bose- and Fermi functions.

We can now try to find an interpretation of the equation given above. First, consider the two δ functions only, which ensure that momentum and energy are conserved. Initially, our particle has momentum \vec{k} and energy $\epsilon_{\vec{k}}$ and it scatters from a particle with momentum \vec{k}_2 and energy $\epsilon_{\vec{k}_2}$. They scatter into states with momenta \vec{k}_1 and \vec{k}_3 in such a way that the sum of initial energies and momenta equals the sum of final-state energies and momenta.

The first term in the last line of Eq. (141) simply describes that for the scattering process to happen, it is necessary that the initial state \vec{k}_2 is occupied and – due to Pauli principle – the two final states have to be empty, see sketch below. The second term can be interpreted as the scattering of hole-excitations. A hole is a missing electron and thus has occupation $1 - n_F(\epsilon)$ rather than $n_F(\epsilon)$. Similarly, the term describing the Pauli-blocking of holes in the final state is $n_F(\epsilon') = 1 - (1 - n_F(\epsilon'))$.



Finally, one has to perform the momentum integrations in Eq. (141), either numerically or analytically, which is non-trivial and requires some extra work. The results depends in the spatial dimension of the system. For example, for $\epsilon_{\vec{k}} \ll T$ one obtains

$$d = 3: \ \Gamma_{\vec{k}} \approx -\operatorname{Im} \Sigma_{\vec{k}}(\varepsilon_{\vec{k}}) \propto T^2 \quad \varepsilon_{\vec{k}} \ll T$$
$$d = 2: \ \Gamma_{\vec{k}} \approx -\operatorname{Im} \Sigma_{\vec{k}}(\varepsilon_{\vec{k}}) \propto T^2 \ln\left(\frac{\varepsilon_F}{T}\right) \quad \varepsilon_{\vec{k}} \ll T$$
$$d = 1: \ \Gamma_{\vec{k}} \approx -\operatorname{Im} \Sigma_{\vec{k}}(\varepsilon_{\vec{k}}) \propto T$$

In two and three dimensions, the scattering rate $\Gamma_{\vec{k}}$ at low temperatures is much smaller than the typical energy of thermal excitations, given by T, $\Gamma_{\vec{k}} \ll T$. Thus, one can approximate the system as weakly interacting quasi particles. Thus, one can use Fermi liquid theory. This is, however, not valid in d = 1, where the scattering is not weak!

Outlook: Both our qualitative argument and the direct perturbative calculations show that the scattering of electrons (or, more precisely, of electronic quasi particles) is strongly suppressed at low temperatures and low frequencies. Here temperature and frequency have to be compared to the Fermi energy, which is most metals is several 10.000 Kelvin. This suppression of scattering is the main reason why Fermi liquid theory is valid and why we can understand the properties of metals using weakly interacting quasi-particles. The situation is different in d = 1, where the picture of weakly interacting Fermions breaks down. A calculation of the real part of the self energy in d = 1 reveals logarithmic divergencies at T = 0 which indicate that perturbation theory is not valid. Luckily, in this case it turns out that chargeand spin-density wave excitations are weakly interacting, which is the basis of the Luttinger-liquid theory describing 1D metals (unfortunately, not covered in this course).

5.5 Convergence of Perturbation Theory

Perturbation theory is ultimately based on a Taylor expansion. In our case, we consider actions of the form

$$\mathcal{S} = \mathcal{S}_0 + \lambda \mathcal{S}_{\text{int}}, \text{ where } \mathcal{S}_0 = -\bar{\psi}g_0^{-1}\psi$$

and calculate some observable in a Taylor expansion in the strength of interaction λ . Thus, the obvious question arises whether the Taylor series is converging or not and what that implies for our results.

Let us recall some results on the convergence of a Taylor series $\sum_n c_n z^n$. This series converges (absolutely) for |z| < r, where r is the so-called convergence radius. For |z| > r, in contrast, the Taylor series does not converge. For example, the series

$$\sum_{0}^{\infty} a^n z^n = \frac{1}{1 - az}$$

converges for |z| > 1/a but diverges for |z| > 1/a.

Thus, we can ask the question about the convergence radius of a Taylor expansion in the strength of interaction. We will show below that for most physically relevant problems, like interacting many-particle systems in the thermodynamic limit, **the radius of convergence is zero**.

As an example, consider the integral

$$I(g) = e^{-f(g)} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx \, e^{-x^2 - gx^4}.$$
(142)

The integral is obviously divergent for all g < 0. This implies immediately that the convergence radius of the Taylor series has to be zero.

We can use this type of argument to analyze other problems. Consider a problem with Coulomb interactions of the form $\frac{e^2}{|r-r'|}$. Assume that this problem has a finite convergence radius, when we do perturbation theory in e. Now consider the replacement $e \rightarrow ie$, which is an allowed coupling constant within the convergence radius. Here, we are replacing electron-electron repulsion by electron-electron attraction! When we do something as drastic, it is very plausible that the solid is not stable anymore (within quantum electrodynamics, QED, even the vacuum would be unstable because of the proliferation of electron-positron pairs). Thus, something non-analytic is happening, in contradiction to our assumption of a finite convergence radius. Similar arguments can be made also for local interactions in the Hubbard model where one obtains superconductivity for negative U. Thus, we conclude that naive **perturbation theory converges almost never** in the models of interest to us.



Figure 11: Absolute value of the difference of the exact value of I(g) and its perturbative value at g = 0.04 computed in *n*th order perturbation theory. As the perturbation theory is not converging, the error grows for large orders of *n*.

What does this imply? First, one has to note that the asymptotic expansion arising from non-converging sums can sometimes be very useful. Let us consider the perturbation theory for f(g) which can be computed to high order using programs like Mathematica

$$f(g) \approx \frac{3g}{4} - 3g^2 + \frac{99g^3}{4} - 306g^4 + \frac{50139g^5}{10} - 102096g^6 + \frac{69533397g^7}{28} + O(g^7)$$

The coefficients of the perturbation theory grow exponentially, the prefactor of the g^{20} term is approximately $3 \cdot 10^{28}$. Nevertheless, if g is very small, one can still use low orders of perturbation theory. This is show in Fig. 11. For g = 0.04 one can compute I(q) with an error of 10^{-4} when one uses 6th order perturbation theory. Typically for a non-converging, asymptotic series, the error does, however, increase when one considers higher orders of perturbation theory! The situation which we encounter here arises, for example, when in high-energy physics the cross-section of electron-electron scattering in is computed within quantum electrodynamics (QED). In this case the effective coupling constant, the fine-structure constant, $\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c} \approx$ $\frac{1}{137,036}$ is very small and thus low-order perturbation theory can be useful in this case. Calculation of quantities like the g-factor of the electron are consistent with experiment on a level of 10^8 and corrections are routinely calculated up to order α^4 . Higher order terms are difficult to compute both due to a large number of terms and badly converging integrations. If one were able to evaluate perturbation theory to very high order, the results would, however, get worse as perturbation theory is not converging. In most solid-state problems, one faces a more difficult situation: there is not small parameter and kinetic energies and interaction terms are of similar size and bare perturbation theory is often not only quantitatively but also qualitatively wrong.

One can, however, improve convergence massively using **resummation methods**. For example, **Borel summation** is a methods from mathematics which allows to make sense of non-convering perturbative expansions of the sort $h(g) = \sum_k a_k g^k$. Here, one considers instead $H(t) = \sum_k \frac{a_k}{k!} t^k$ which may have much better convergence properties. If H(t) is converging, using $\int_0^\infty (gt)^n e^{-t} dt = g^n n!$, one can then compute h(g) from $h(g) = \int_0^\infty H(gt) e^{-t} dt$.



Figure 12: Comparison of resummed perturbation theory for f(g) with the exact result (dashed line) as function of g. Already for n = 1, equivalent to self-consistent Hartree-Fock, one obtains a very good approximation for all values of g. The result for n = 2 has an error of less than 0.2%. In contrast, bare perturbation theory cannot be used for $g \gtrsim 0.1$.

In physics, typically different resummation methods are used. One powerful method is to readjust the expansion point. In the example (142), we can rewrite the exponent of I(g) as

$$x^{2} + gx^{4} = ax^{2} + \underbrace{(1-a)x^{2} + gx^{4}}_{\text{perturbation}}$$

In linear order perturbation theory, one obtains, for example,

$$f(g) \approx \frac{\log(a)}{2} + \frac{(1-a)}{2a} + \frac{3g}{4a^2}$$

Which value of a should we use? One possible choice is to demand that the final result, calculated in n order perturbation theroy, is approximately independent of a, $\partial f/\partial a = 0$. For n = 1 this method is equivalent to **self-consistent Hartree-Fock** (as discussed before, this is the 'best' Gaussian approximation). We obtain for for n = 1, that $a = \frac{1}{2} \left(\sqrt{12g + 1} + 1 \right)$ and thus we get

$$f(g) \approx \frac{\log\left[\frac{1}{2}\left(\sqrt{12g+1}+1\right)\right]}{2} + \frac{1 - \frac{1}{2}\left(\sqrt{12g+1}+1\right)}{\sqrt{12g+1}+1} + \frac{3g}{\left(\sqrt{12g+1}+1\right)^2}$$

In Fig. 12, we show that this formula gives evan for n = 1 accurate results not only for small g but, surprisingly, also for very large g. If one uses instead n = 3, the relative error becomes smaller than 0.2%.

Outlook: That our resummation method works so well in the case described above is, perhaps, accidential. But the example shows, that with field-theoretical resummation techniques one can obtain good results even in situations when naive perturbation theory is divergent and, importantly, even when coupling constants are not small. This is important, as for practically all applications in solid-state physics, there are no obviously small parameters.

5.6 Screening of long-ranged Coulomb interactions

Motivation: In almost all situations, we have to use resummation methods to obtain meaningful results from perturbation theory. One such resummation technique discussed in previous chapters was the introduction of the self energy. Another example is self-consistent Hartree-Fock theory. In this and the following sections, we will consider further resummation techniques to be able to describe important physical phenomena like screening or the formation of bound states.

When doing perturbation theory for the Coulomb interaction, $V(r) = \frac{e^2}{4\pi\epsilon_0 r}$, one realizes that one obtains many divergent contributions arising because the Coulomb potential decays very slowly. Consider, for example, a uniformly charged block of length $L \times L \times L$ with charge density ρ . In this case, the Coulomb energy per volume $V = L^3$ diverges for large L

$$\frac{E_C}{V} = \frac{1}{2V} \int d^3r d^3r' V(|\vec{r} - \vec{r'}|) \rho^2 \propto \rho^2 L^2.$$

As a consequence, a 3D solid is always charge neutral. The charge of the electrons is canceled by the charge of the ions.

Taking this into account in a simplified way, we use in the following the 'jellium model'. We replace the ions by a uniformly charges background, $\rho_{ion}(\vec{r}) = \text{const.}$, which is exactly compensated by the uniform average charge of the electrons. This is equivalent to using the following Coulomb potential for electron-electron interactions in Fourier space

$$V(\vec{q}) = \begin{cases} \frac{4\pi e^2}{4\pi\epsilon_0 q^2} & \vec{q} \neq 0\\ 0 & \vec{q} = 0 \end{cases} \text{ (compensated by background)}$$

As a consequence, the Hartree term, which is proportional to $V(\vec{q}=0)$, vanishes.

5.6.1 Effective potential, polarization, and RPA approximation

To avoid unphysical divergencies, we have to resum perturbation theory. Here we follow the same arguments which lead us to the concept of the self-energy. Let us consider the so-called effective potential, defined as the sum of all connected diagrams which begin and end with an interaction line.



Here, $i\Omega$ is a bosonic Matsubara frequency. Consider, for example, the last term. As all three interaction lines have the momentum \vec{q} , this term is proportional to $1/q^6$

and thus highly divergent in the small q limit. This unphysical divergence implies that we have to resum perturbation theory. We do this by introducing the analog of the self energy, which is called **polarization**, $\Pi_{\vec{q}}(i\Omega)$.

$$= \cdots + \cdots = \frac{\cdots}{1 - \cdots}$$

In formulas, this equation reads

$$V_{\text{eff}}(\vec{q}, i\Omega) = \frac{V(\vec{q})}{1 + \Pi_{\vec{q}}(i\Omega)V(\vec{q})} = \frac{V(\vec{q})}{\epsilon(\vec{q}, i\Omega)} = \frac{1}{V_{\vec{q}}^{-1} + \Pi_{\vec{q}}(i\Omega)}$$
(143)

Here, the **polarization** $\Pi_{\vec{q}}(i\Omega)$ is defined as



Here the dots denote points where an interaction line with momentum \vec{q} and frequency $i\Omega$ has to start or end.

Note that in Eq. (143), we have introduced a frequency and momentum-dependent dielectric function: $\epsilon(\vec{q}, i\Omega) = 1 - \prod_{\vec{q}}(i\Omega)V(\vec{q})$, which allows to write

$$V_{\text{eff}}(\vec{q}, i\Omega) = \frac{4\pi e^2}{4\pi\epsilon_0\epsilon(\vec{q}, i\Omega)q^2}.$$

This dielectric function is used in electrodynamics, in formulas like $\vec{D} = \epsilon_0 \epsilon_r \vec{E}$. In insulators, one can often approximate ϵ by a constant, but this is not possible in a metal.

The formulas given above are (within the jellium model) exact. Now we consider the simplest possible approximation for the polarization, the so called **Random Phase Approximation (RPA)**. Here, we compute $\Pi_{\vec{q}}(i\Omega)$ to lowest order in perturbation theory

$$\Pi_{\vec{q}}^{\text{RPA}}(i\Omega) \approx \bigodot = -\frac{2}{\beta} \sum_{i\omega_n} \int \frac{d^3k}{(2\pi)^3} g_{\vec{k}}^0(i\omega_n) g_{\vec{k}+\vec{q}}^0(i\omega_n+i\Omega)$$
$$= -2 \int \frac{d^3k}{(2\pi)^3} \frac{n_F(\varepsilon_{\vec{k}}-\mu) - n_F(\varepsilon_{\vec{k}+\vec{q}}-\mu)}{i\Omega + \varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}}}$$

5.6.2 Thomas Fermi approximation

RPA is a powerful approximation, which we will analyze piece by piece in different limits. Within the **Thomas Fermi approximation**, we consider the limit of small \vec{q} and Ω and replace $\Pi_{\vec{q}}(i\Omega)$ by $\Pi_{\vec{q}=0}(\Omega=0)$

$$\Pi_{\vec{q}=0}(\Omega=0) \approx -2\int \frac{d^3k}{(2\pi)^3} \frac{(-1)n'_F(\varepsilon_{\vec{k}}-\mu) \cdot \frac{\partial \varepsilon_{\vec{k}}}{\partial \vec{k}} \cdot \vec{q}}{(-1)\frac{\partial \varepsilon_{\vec{k}}}{\partial \vec{k}} \cdot \vec{q}} = -2\int \frac{d^3k}{(2\pi)^3} n'_F(\varepsilon_{\vec{k}}-\mu)$$
$$\approx 2\int \frac{d^3k}{(2\pi)^3} \,\delta(\varepsilon_{\vec{k}}-\mu) = N_F^{\text{tot}} \quad \text{density of states at Fermi surface}$$

In the last line, we consider a metal where the temperature is much smaller than the Fermi energy, which allows to replace the derivative of the Fermi function with a negative δ function.

Thus, for $\vec{q}, \Omega \to 0$, we obtain

$$V_{\text{eff}}(\vec{q}) \approx \frac{4\pi e^2}{4\pi\epsilon_0 \left(q^2 + \left(\frac{1}{\lambda}\right)^2\right)}, \qquad V_{\text{eff}}(\vec{r}) = \frac{e^2}{4\pi\epsilon_0 |\vec{r}|} e^{-\frac{|\vec{r}|}{\lambda}}$$
(144)
with $\lambda = \left(e^2 N_F^{\text{tot}}/\epsilon_0\right)^{-\frac{1}{2}}$ Thomas-Fermi screening length

Let us check the formula for the effective potential in real space

$$\int \frac{d^3q}{(2\pi)^3} V_{\text{eff}}(\vec{q}) e^{i\vec{q}\cdot\vec{r}} = \frac{2\pi}{(2\pi)^3} \int_{-1}^{1} du \int q^2 dq \, V_{\text{eff}}(q) e^{iqru} = \frac{1}{(2\pi)^2} \int_{0}^{\infty} dq \, \frac{e^{iqr} - e^{-iqr}}{iqr} q^2 \frac{4\pi e^2}{4\pi\epsilon_0 \left(q + \frac{i}{\lambda}\right) \left(q - \frac{i}{\lambda}\right)} = \frac{e^2}{4\pi\epsilon_0 r} \int_{-\infty}^{\infty} \frac{dq}{2\pi i} \frac{e^{iqr}}{\left(q + \frac{i}{\lambda}\right) \left(q - \frac{i}{\lambda}\right)} 2q = \frac{e^2}{4\pi\epsilon_0 |r|} e^{-\frac{|r|}{\lambda}}$$

Thus, in a metal, the effective potential decays in a metal exponentially at large distances (up to Friedel-oscillations discussed below). The physical mechanism behind that is called **screening**. Around a positive test-charge a cloud of negatively charge electrons accumulates in such a way that the net charge is zero. Remarkably, our simple resummation strategy captures this effect. The size of the screening cloud is set by the Thomas-Fermi screening length λ .

5.6.3 Plasma oscillations

Next, we consider the limit of large frequencies but small momenta, $q \to 0$. We also do the analytic continuation, $i\Omega \to \Omega + i0$ (we are writing here *i*0 instead of *i* ϵ for an infinitesimal imaginary part to avoid confusion with the dielectric function)

$$\begin{split} \lim_{\vec{q}\to 0} \Pi_{\vec{q}}(\Omega+i0) &= -2\int \frac{d^3k}{(2\pi)^3} \frac{n_F(\varepsilon_{\vec{k}}-\mu) - n_F(\varepsilon_{\vec{k}+\vec{q}}-\mu)}{\Omega+i0 + \varepsilon_{\vec{k}} - \varepsilon_{\vec{k}+\vec{q}}} \\ &\approx -2\int \frac{d^3k}{(2\pi)^3} \underbrace{\frac{n_F(\varepsilon_{\vec{k}}-\mu) - n_F(\varepsilon_{\vec{k}+\vec{q}}-\mu)}{\Omega+i0}}_{=0} - \frac{\left(\frac{\partial\varepsilon_{\vec{k}}}{\partial\vec{k}}\vec{q}\right)\left(\vec{q}\frac{\partial}{\partial\vec{k}}n_F(\varepsilon_{\vec{k}}-\mu)\right)}{\Omega^2} \\ &\approx \sum_{i,j} q_i q_j \int \frac{d^3k}{(2\pi)^3} n_F(\varepsilon_{\vec{k}}-\mu) \frac{\partial^2\varepsilon_{\vec{k}}}{\partial k_i \partial k_j} \frac{1}{\Omega^2}(-2) \end{split}$$

In the second line, we perform a Taylor expansion of the denominator. The first term in the second line vanishes exactly as one can shift the \vec{k} integration by $-\vec{q}$, thus only the $1/\Omega^2$ term survives, which is evaluated in the third line using a partial integration. Thus, we find for the dielectric function in this limit

$$\epsilon(\vec{q} \to 0, \Omega) = 1 + V_0(\vec{q}) \Pi_{\vec{q}}(i\Omega) \approx 1 - \frac{\omega_p^2}{\Omega^2}$$
$$\omega_p^2 = 2 \int \frac{d^3k}{(2\pi)^3} n_F(\varepsilon_{\vec{k}} - \mu) \frac{\partial^2 \varepsilon_{\vec{k}}}{\partial k^2} e^2 / \epsilon_0$$

where ω_p is called the **plasma frequency**. Note that the *q* dependence has canceled. For a quadratic dispersion $\varepsilon_{\vec{k}} = \frac{k^2}{2m}$, the plasma frequency is given by

$$\omega_p = \sqrt{\frac{ne^2}{m\epsilon_0}},$$

where n is the particle density and m the electron mass.

At $\Omega = \omega_p$ the dielectric function vanishes and one obtains a pole in the effective potential, $V_{\text{eff}} = \frac{V(\vec{q})}{\epsilon(q,\Omega)} = \frac{V(\vec{q})}{1 - \frac{\omega_p^2}{\Omega^2}}$, see figure. Such poles describe collective oscillations of a system, the **plasma oscillations**. Such oscillations of the positive relative to the negative charges do occur in any system of charged particles and therefore also in a plasma, i.e., in a ionized gas, which explains the name.



Figure 13: Effective potential diverging at $\Omega = \omega_p^2$

Let us re-derive the physics of plasma oscillations in a more basic approach. This will help to convince ourselves that the result derived above is exact and not an artifact of an approximation. We consider a solid with volume V and investigate the uniform oscillations of all electrons relative to the background given by the ions as shown in the following sketch.



We assume that all electrons move by Δx relative to background charge. We obtain a kind of plate capacitor where the charge density on the surface (with units charge per area) is given by $en\Delta x$. Using Maxwell's equation $\nabla E = \rho/\epsilon_0$, we obtain an electric field inside the solid of $E = en\Delta x/\epsilon_0$. The energy stored in this electric field is given by $\frac{1}{2}\epsilon_0 \int E^2 = \frac{1}{2}\epsilon_0(en/\epsilon_0)^2(\Delta x)^2 V$. We interpret this as the energy of an harmonic oscillator $\frac{1}{2}D(\Delta x)^2$. As the total mass of the electrons is given by M = mnV, the frequency of the harmonic oscillator is obtained as

$$\omega_p = \sqrt{\frac{D}{M}} = \sqrt{\frac{ne^2}{m\epsilon_0}}$$

which is exactly the result which we obtained above from RPA. Note that there is no screening in this large-frequency limit and plasma oscillations arise necessarily in a system with long-ranged interaction.

Outlook: Collective plasma oscillations of electrons play an important role in technology, e.g., for building antennas. They have also important consequences in superconductors and in the theory of the Higgs effect in high-energy physics. In both cases, instead of gapless Goldstone modes and gapless photons, one obtains oscillations at the corresponding plasma frequency (in high-energy physics, this is related to the mass of the W and Z bosons, the analogs of photons responsible for the electroweak interactions).

5.6.4 Absorption and particle-hole pairs

The absorption of electromagnetic waves is controlled by the imaginary part of the dielectric function and therefore of $\Pi_{\vec{q}}(\Omega + i0)$. Within RPA it is given by

$$\operatorname{Im} \Pi_{\vec{q}}(\Omega + i0) = 2 \int \frac{d^3k}{(2\pi)^3} \left(n_F(\varepsilon_{\vec{k}} - \mu) - n_F(\varepsilon_{\vec{k} + \vec{q}} - \mu) \right) \pi \delta(\Omega - (\varepsilon_{\vec{k} + \vec{q}} - \varepsilon_{\vec{k}}))$$

It has a simple physical interpretation: it counts the number of electron-hole pairs with energy Ω and momentum \vec{q} . This can be seen from the fact that for T = 0 and $\Omega > 0$, there is only a contribution when $\varepsilon_{\vec{k}} < \mu$ (hole excitation) while simultaneously $\varepsilon_{\vec{k}+\vec{q}} > \mu$) (electronic excitation).

For small Ω one finds

$$\operatorname{Im} \Pi_{\vec{q}}(\Omega) \approx -N_F^{\text{tot}} \frac{\Omega}{v_F |\vec{q}|} \quad \text{for} \ \ \Omega < v_F |q|, |\vec{q}| \ll 2k_F$$

where v_F is the Fermi velocity and N_F^{tot} the total density of states at the Fermi energy.

Particle-hole pairs have the maximum possible momentum of $2k_F$ and thus

$$\operatorname{Im} \Pi_{\vec{q}}(\Omega \to 0) = 0 \quad \text{for} \quad |\vec{q}| > 2k_F.$$

The following sketch shows in which frequency and momentum regime the imaginary parts are finite, due to the excitation of particle-hole pairs (or plasma oscillations)



5.6.5 Friedel oscillations

Finally, we consider the limit $\Omega = 0, T = 0$ but we now focus on large momenta, close to $2k_F$. This limit is most easily analyzed for a quadratice dispersion, $\varepsilon_{\vec{k}} = \frac{k^2}{2m}$, where one can do the momentum integrals analyticially. In d = 3 one finds (we do not show the derivation)

$$\Pi_{\vec{q}}(\Omega=0) = -e^2 \frac{mk_F}{\pi^2} \left(\frac{1}{2} + \frac{1-x^2}{4x} \ln \left| \frac{1+x}{1-x} \right| \right), \qquad x = \frac{q}{2k_F}$$

For x = 1 or $q \approx 2k_F$ this function has a weak non-analyticity $\Pi_{\vec{q}} \approx C_1 + C_2 |q - 2k_F| \ln(|q - 2k_F|)$. This non-analyticity therefore also shows up in the q-dependence of the effective potential

$$V_{\text{eff}}(q \approx 2k_F) \approx \tilde{C} + \tilde{C}' |q - 2k_F| \ln(|q - 2k_F|).$$

If one does a Fourier transformation of this function back to real space, one obtains

$$V_{\rm eff}(\vec{r}) \approx \frac{e^2}{4\pi\epsilon_0 r} e^{-\frac{r}{\lambda}} + \tilde{\tilde{C}} \underbrace{\frac{\cos(2k_F r + \varphi)}{r^d}}_{\text{Friedel}}$$

in d dimensions. Thus, at large distances the effective potential does **not** decay exponentially as suggested by the Thomas Fermi approximation but at T = 0 instead with a powerlaw $1/r^d$ in d dimensions. This is an effect arising from the sharp Fermi surface of a metal. The prefactor oscillates with $\cos(2k_F r)$ and the effect has the name **Friedel oscillations**.



Outlook: Friedel oscillations have an important technological application when one builds magnetic sensors used, e.g., in magnetic hard disks. Albert Fert and the Peter Grünberg received the Nobel prize for the Giant Magneto Resistance (GMR) effect used here (Peter Grünberg did this research in Jülich, made his habilitation in Cologne and was teaching from 1992 until his retirement in 2004 as a professor in Cologne). GMR devices are made from two ferromagnetic layers separated by a metallic layer.



The magnetic layers trigger Friedel oscillations in the the metal decaying with 1/r which leads to an effective interaction of the metallic layers which can be fine-tuned with the thickness of the layer. The sensor then uses the resistivity difference of parallel and antiparallel magnetic configurations.

5.7 Controlled approximations: large N

Motivation: We have argued that to obtain physically meaningful results, one typically has to resum an infinite subset of an infinite number of diagrams. But how does one choose the right subset of diagrams? What can be a guiding principle? There is not a simple answer to this question, a lot depends on what physical effect one wants to describe.

Hear are a few useful strategies:

- One strategy is to analyze and resum the **most diverging** terms in perturbation theory. This lead us, for example, to the idea to introduce the self energy together with the Dyson equation to resum an infinite set of diagrams. Later, the same strategy was used by us to compute the effective potential from polarization diagrams.
- Often, it is also a good idea to go for the **simplest** possible approximation which captures a given physical effect. For example, RPA was both simple and surprisingly rich in physics.
- A strategy, which is not covered in this course, is to be guided by principles like conservation laws or gauge invariance. So-called **conserving approximations** are build in such a way that they obey all relevant conservation laws.
- In the follow-up lecture, QFT II, two further powerful methods will be introduced, **effective field theories** and the **renormalization group**.

• In this section, we will discuss an example of a **controlled approximation**, where the selection is diagrams is controlled by the smallness of some parameter α .

For a **controlled approximation** one needs – or has to invent – a small dimensionless parameter $\alpha \ll 1$. There are many possible choices for α . Here are a few examples:

- In QED, for the scattering of high-energy electrons, one uses that the dimensionless interaction strength $\frac{e^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137}$ is small.
- The parameter r_s is defined as the ratio of the average distance electron and the Bohr radius a_B , $r_s = \frac{1}{a_B} \left(\frac{1}{\text{electron density}}\right)^{\frac{1}{3}}$. Using either $\alpha = r_s$ or $\alpha = 1/r_s$, one obtains completely different approximations.
- One can also use $\alpha = \frac{1}{\text{dimension}}$, $\alpha = \frac{\text{wavelength of an electron}}{\text{scattering length}}$, $\alpha = \frac{1}{\text{spin size}}$, the latter expansion we used to derive the spin-wave theory for magnets in a systematic way.
- A powerful set of these approaches are called **large-N approximation**. Here, one identifies $\alpha = \frac{1}{N}$, where N is some extra large parameter related to the number of components of the fields.

Let us make a concrete example. Electrons carry a spin 1/2, where spin rotations are described by the group SU(2). We now generalize this to rotations described by the group SU(N) instead (more precisely, we use the fundamental representation of SU(N)). We achieve this simply by generalization the creation operator of an electron to

$$\psi^{\dagger}_{\alpha}(\vec{r}), \qquad \alpha = 1, \cdots, N.$$

The 'spin index' now runs over N instead of just two components. An SU(N) symmetric version of the interaction term now takes the form

$$\mathcal{H}_{\rm int} = \frac{1}{2} \frac{1}{N} \int d^3 r d^3 r' \sum_{\alpha, \alpha'=1, \cdots, N} \psi^{\dagger}_{\alpha}(\vec{r}) \psi_{\alpha}(\vec{r}) V(\vec{r}-\vec{r'}) \psi^{\dagger}_{\alpha'}(\vec{r'}) \psi_{\alpha'}(\vec{r'}).$$

For N = 2 this is equivalent to the usual Coulomb interaction. Very important is the prefactor $\frac{1}{N}$ in front of the Coulomb interaction, which is needed to make the limit $N \to \infty$ well-defined. How does N enter in the Feynman rules? First, each interaction line obtains a factor 1/N but one obtains an extra factor of N for each fermionic loop as one has not to sum over N fermionic species.

Let us look at a few self-energy diagrams and analyze how they depend on N



It is obvious, that we can drop some of the diagrams for large N but have to keep others. More precisely, for a given order in the interaction strength, we need the maximal number of Fermi loops. These are precisely the diagrams which we summed up using RPA.

$$\underset{\mathcal{O}(\frac{1}{N})}{\overset{\mathcal{O}(\frac{1}{N^2}N)}} + \underset{\mathcal{O}(\frac{1}{N^3}N^2)}{\overset{\mathcal{O}(\frac{1}{N^3}N^2)}} + \cdots = -\frac{1}{N} \frac{V_{\vec{q}}}{1 + \prod_{\vec{q}}^{\text{RPA}}V_{\vec{q}}}$$

Here, when one adds an interaction line, one also adds a Fermi loop. Thus all diagrams are proportional to 1/N. All corrections to RPA like



are suppressed by extra powers of 1/N and can be neglected for $N \to \infty$. Thus, **RPA becomes exact for** $N \to \infty$.

$$\lim_{N \to \infty} V_{\text{eff}}(\vec{q}) = \frac{1}{N} \frac{V_{\vec{q}}}{1 + \prod_{\vec{q}}^{\text{RPA}} V_{\vec{q}}}$$

where $\Pi_{\vec{q}}^{\text{RPA}} = \bullet$ is the particle-hole bubble of RPA (here defined without the factor N).

The leading self-energy correction is then

$$\lim_{N \to \infty} \Sigma_{\vec{q}}(\Omega) = \underbrace{}_{}$$

The same argument can also be used for more complicated quantities like correlation functions of the type $\langle \sum_{\alpha} \Psi^{\dagger}_{\alpha}(\vec{r}) \Psi_{\alpha}(\vec{r}) \sum_{\alpha'} \Psi^{\dagger}_{\alpha'}(\vec{r'}) \Psi_{\alpha'}(\vec{r'}) \rangle$ (needed to compute charge susceptibilities), where to leading order one has to take into account three diagrams



We have shown that RPA becomes exact for large N but why does this help when we are interested in the case N = 2? Why does it help to have a **controlled approximation**? We now know that a model exist, where RPA is quantitatively valid. This implies that the approximation scheme will not violate any basic physical principles like conservation laws. The example above shows, that it also guides the computation of more complex quantities. Furthermore, it is the starting point for a systematic improvement of the approximation.



Figure 14: T-matrix in d = 1 for $V_0 = -0.5$, $m^* = 1$. The presence of a bound state gives rise to a δ -function in Im $T(\omega + i\epsilon)$ at $\omega = E_B$, where $E_B = -V_0^2/2m^* = -0.125$ is the binding energy.

5.8 Ladder approximation and Bethe-Salpeter equation

Motivation: Different resummation methods are needed depending on the physical problem. In this section, we investigate which diagrams have to be resummed to capture exactly the limit of a low density of particles. The resulting vertex equations (or Bethe-Salpeter equations) are, however, have, however, a much broader range of applicability.

Our goal is to find an approximation which becomes exact in the low-density limit. It should therefore capture also the physics of just two particles. Our problem is that the two-particle interaction $V(\vec{x}_1 - \vec{x}_2)$ often is not small, and quantities like bound states are not accessible in perturbation theory. As a concrete example, which can be solved analytically, we can consider the scattering from a δ function

$$V(\vec{x}_1 - \vec{x}_2) = V_0 \delta^3(\vec{x}_1 - \vec{x}_2)$$

but we are also interested in more general interaction potential.

Before we attack the problem with many-particles in the low-density limit, let us focus on just two particles. By going to relative coordinates, $\vec{r} = \vec{x}_1 - \vec{x}_2$, such a two-particle problem can be rewritten as a potential scattering of a particle with an effective mass $\frac{1}{m^*} = \frac{1}{m_1} + \frac{1}{m_2}$ (check, e.g., the solution of the hydrogen problem in any QM textbook).

How would we treat such a single-particle scattering problem diagramatically? We would have to sum up the following diagrams

where the dashed line with a cross stands, in momentum space, for the Fourier transform of the potential $V_{\vec{q}}$. $T_{\vec{k}\vec{k}'}$ is called the T-matrix of the problem with

$$T_{\vec{k}\vec{k}'} = V_{\vec{k}-\vec{k}'} + \sum_{\vec{k}''} V_{\vec{k}-\vec{k}''} g^0_{\vec{k}''} V_{\vec{k}''-\vec{k}'} + \sum_{\vec{k}'',\vec{k}'''} V_{\vec{k}-\vec{k}''} g^0_{\vec{k}''} V_{\vec{k}''-\vec{k}''} g^0_{\vec{k}'''} V_{\vec{k}''-\vec{k}'} + \cdots$$

For $V_{\vec{q}} = V_0 = const.$, i.e. for a δ -function interaction in real space, the equation simplifies to a geometric series.

$$T^{R}(\omega) = T(\omega + i\epsilon) = \frac{V_{0}}{1 - V_{0} g^{R}_{00}(\omega)}, \qquad g^{R}_{00}(\omega) = \int \frac{d^{d}k}{(2\pi)^{d}} g^{R}_{\vec{k}}(\omega)$$

As a concrete example, let us consider the case $\varepsilon_{\vec{k}} = \frac{k^2}{2m^*}$ in d = 1, where the local Green's function is given by

$$g_{00}^{R}(\omega) = \int \frac{dk}{2\pi} \frac{1}{\omega - \frac{k^{2}}{2m^{*}} + i\epsilon} = -\frac{1}{\sqrt{2m^{*}}} \frac{1}{(-(\omega + i\epsilon))^{\frac{1}{2}}}$$
(145)

A bound state shows up as a pole in the *T*-matrix, i.e., for $1 - V_0 g_{00}^R(\omega) = 0$ at $\omega = E_B$. The position of the pole is the binding energy, see Fig. 14. In the example shown above, a bound state occurs always exist for $V_0 < 0$ in d = 1 with $E_B = -\frac{V_0^2}{2m^*}$.

Above, we sketched very briefly how one can solve the single-particle potentialscattering problem. But we are interested in a different problem: we want to consider a system with a finite density of particles. What diagrams do we have to resum in this case in such a way that we are correctly treating the two-particle problem also (including bound states)? The discussion given above, suggests that we should keep the following diagrams

This set of diagrams is called (for obvious reasons) "ladder approximation" in the particle-particle channel.

For other problems (e.g., for bound states of an electron and a hole in a semiconductor) one instead considers the ladder diagrams in the particle-hole channel by reverting the direction of one of the Green's functions.



5.8.1 Bethe-Salpeter equation

Before we try to evaluate the ladder-approximation, we take a step back and consider a more general problem. Let us assume, that we want to evaluate (in a functional integral) the expectation value of four fields, where two particles are created and two are destroyed (this is therefore a two-particle correlation function)

$$\langle \psi_{2'}\psi_{1'}\bar{\psi}_1\bar{\psi}_2\rangle = 1 = 1' - 1' + 2 = 2' - 2' + 2'$$

Here, the indices stand, e.g., for momentum and frequencies, $1 = (\vec{q}_1, \vec{\Omega}_1)$. The double line is the full Green's function (including self-energy corrections) $\rightarrow = \mathcal{G}_{\vec{q}}(i\Omega)$, while the box, which depends on four frequencies and momenta is called a **vertex**. It is the sum of all possible connected diagrams beginning and ending with two fermionic lines.

As a next step, we want to generalize the concept of the Dyson equation and the self-energy (introduced to compute a Green's function) to such a vertex. This can be done with a so-called **Bethe-Salpeter equation**. Many different versions of this equations exists. For the case we are interested in, they take the following form



where the blue dashed object is called **irreducible vertex**. It is defined as the sum of all connected diagrams, which do not fall into two pieces when cutting **two** fermionic lines running in parallel.



Let us write the Bethe-Salpeter equation in a formula. As a short-hand notation, we use $\vec{K} = (i\Omega, \vec{K})$ to denote both frequencies and momenta. With this notation, our vertex depends on three frequency/momenta indices only and we use the following convention



Due to our diagrammatic rules, the sum of the incoming momenta (and frequencies) has to be equal to the sum of outgoing frequencies and momenta, thus $2\vec{K}$ is the total momentum of the two particles, which is conserved during the interaction. If we introduce a similar notation for the **irreducible vertex** $I^1_{\vec{K}}(\vec{k},\vec{k'})$, the Bethe-Salpeter equation takes the form

$$V_{\vec{K}}(\vec{k},\vec{k}') = I_{\vec{K}}^{1}(\vec{k},\vec{k}') + \frac{1}{2} \sum_{\vec{k}''} I_{\vec{K}}^{1}(\vec{k},\vec{k}'') g_{\vec{K}+\frac{\vec{k}''}{2}} g_{\vec{K}-\frac{\vec{k}''}{2}} V_{\vec{K}}(\vec{k}'',\vec{k}')$$
(146)

Thus, for a given irreducible vertex $I^1_{\vec{K}}(\vec{k}, \vec{k}')$, the Bethe-Salpeter equation is a **linear integrale equation** with a single internal integration over frequency and momentum.

The prefactor " $\frac{1}{2}$ " in front of the second terms is needed to avoid double counting, arising from the fact that there are two ways to connect outgoing to ingoing lines,





Figure 15: Contour used for the evaluation of $A(\vec{K}, i\Omega)$.

5.8.2 Ladder approximation and bound states

We can now come back to the **ladder approximation**. The ladder approximation is the lowest order (linear in the interaction) approximation for the irreducible vertex.



The graphical representation given above is a bit misleading as the Green's function are **not** part of the irreducible vertex. In formulas, this is $I^1_{\vec{K}}(\vec{k},\vec{k}') = -V^{\text{int}}_{\vec{k}/2+\vec{k}'/2} - V^{\text{int}}_{\vec{k}/2-\vec{k}'/2}$, where V^{int} is the interaction potential (not to be confused with the vertex). Furthermore, we consider for simplicity local interactions, $V^{\text{int}}_{\vec{q}} = U = const.$ as in the Hubbard model.

As a warm-up, we consider one of the internal summation within the ladder diagram

$$\vec{K} + \frac{\vec{k}}{2} \xrightarrow{\vec{K} - \vec{q}} \vec{K} + \frac{\vec{k}'}{2} = \sum_{\substack{\mathbf{q} = (i\omega_n, \vec{q}) \\ \vec{K} - \frac{\vec{k}}{2} \\ \vec{K} - \frac{\vec{k}}{2} \\ \vec{K} - \frac{\vec{k}}{2} \\ \vec{K} - \frac{\vec{k}}{2} \\ \vec{K} - \frac{\vec{k}'}{2} \\ \vec{K} - \frac{\vec{k}'$$

Importantly, we observe that the object is independent of \vec{k} and $\vec{k'}$ if we use a local interaction, $V_{\mathbf{q}}^{\text{int}} = U$. Therefore, within the ladder approximation and for local bare interactions, also the full vertex depends only on \vec{K} , but not on \vec{k} and $\vec{k'}$, $V_{\vec{K}}(\vec{k},\vec{k'}) = V_{\vec{K}}$ and the Bethe-Salpeter equation takes the simple form

$$V_{\vec{K}} = -2U - U \underbrace{\sum_{\mathbf{q}} g_{\vec{K}-\mathbf{q}} \, g_{\vec{K}+\mathbf{q}}}_{A_{\vec{K}}} V_{\vec{K}}$$

solved by

$$V_{\vec{K}} = -U - \frac{U}{1 + UA_{\vec{K}}}.$$
(147)

where

$$\begin{split} A_{\vec{K}} &= A(\vec{K}, i\Omega) = \frac{1}{\beta} \sum_{\omega_n} \int \frac{d^3q}{(2\pi)^3} g_{\vec{K}-\vec{q}}(i\Omega - i\omega_n) g_{\vec{K}+\vec{q}}(i\Omega + i\omega_n) \\ &= -\int \frac{d^3q}{(2\pi)^3} \oint \frac{dz}{2\pi i} n_F(z) \frac{1}{i\Omega - z - \varepsilon_{\vec{K}-\mathbf{q}}} \frac{1}{i\Omega + z - \varepsilon_{\vec{K}+\mathbf{q}}} \\ &= \int \frac{d^3q}{(2\pi)^3} \left[\frac{n_F(-\varepsilon_{\vec{K}-\mathbf{q}})(-1)^3}{2i\Omega - (\varepsilon_{\vec{K}-\mathbf{q}} + \varepsilon_{\vec{K}+\mathbf{q}})} + \frac{n_F(\varepsilon_{\vec{K}+\mathbf{q}})(-1)^2}{2i\Omega - (\varepsilon_{\vec{K}-\mathbf{q}} + \varepsilon_{\vec{K}+\mathbf{q}})} \right] \\ &= -\int \frac{d^3q}{(2\pi)^3} \frac{1 - n_F(\varepsilon_{\vec{K}+\mathbf{q}}) - n_F(\varepsilon_{\vec{K}-\mathbf{q}})}{2i\Omega - (\varepsilon_{\vec{K}-\mathbf{q}} + \varepsilon_{\vec{K}+\mathbf{q}})} \end{split}$$

Fig. 15 shows the contour in the complex plane used in the evaluation of the frequency sum

Finally, let us consider the limit when the density of fermions becomes very small. In this case, we can approximate $n_F(\varepsilon_{\vec{K}\mp\vec{q}})\approx 0$. Furthermore, we consider the case $\varepsilon_{\vec{k}}=\frac{k^2}{2m}$. We obtain

$$A(\vec{K}, i\Omega) = -\int \frac{d^3q}{(2\pi)^3} \frac{1}{2i\Omega - \left(2\frac{K^2}{2m} + 2\frac{q^2}{2m}\right)} \stackrel{=}{\uparrow} -g_{00}\left(2i\Omega - \frac{K^2}{m}\right)$$

$$\stackrel{\text{compare with scattering}}{\underset{\text{problem for } m^* = \frac{m}{2}}$$

We can compare this to the scattering problem, where the T matrix was computed using the local Green's function, see Eq. (145), of a problem with the effective mass $m^* = m/2$ and we obtain the same equation and the same condition $1 - Ug_0 0(\tilde{\omega}) = 0$ for the bound state, if we identify $\tilde{\omega}$ with $2\Omega - K^2/m$. This shows, that our vertex equation correctly describes the formation of bound states in the low-density limit. Furthermore, we can use it, e.g., to explore how a finite density of particles affects the bound-state formation. The physics, which we discussed here, is, for example, important to understand excitons in doped semiconductors.

5.8.3 Outlook: Vertex corrections and conservation laws

Outlook: Above, we have shown how one can use Bethe-Salpeter equation to obtain non-perturbative results for vertices, focussing on the physics of the formation of bound states. Vertices and vertex equations play an important role in the computation of all types of correlation functions.

In many applications, one encounters a slightly different variant of the vertex equa-

tion, based on vertices with three instead of four legs. An example is given below:



Here, I^2 is the irreducible vertex defined as the sum of all diagrams with four legs, which do not fall into two parts when two antiparallel moving Green's function lines are cut. $\Gamma_{\vec{q},\vec{k}}$ is the sum of all connected diagrams which start with $\Gamma^0_{\vec{q},\vec{k}}$ on the left and end with two Green's function on the right.

The form of $\Gamma^0_{\vec{q},\vec{k}}$ depends on application. Below, we will calculate correlation function of charge

$$\rho_{\vec{q}} = \sum_{\vec{k}} c^{\dagger}_{\vec{k}+\vec{q}} c_{\vec{k}} = \sum_{\vec{k}} c^{\dagger}_{\vec{k}+\frac{\vec{q}}{2}} c_{\vec{k}-\frac{\vec{q}}{2}}$$

In this case, one needs $\Gamma^0_{\vec{q},\vec{k}} = 1$. Instead, if one wants to study correlations of the current density (needed for the computation of the conductivity, see next chapter)

$$\vec{j}_{\vec{q}} = \frac{1}{2} \sum_{\vec{k}} \left(\vec{v}_{\vec{k}+\frac{\vec{q}}{2}} + \vec{v}_{\vec{k}-\frac{\vec{q}}{2}} \right) c^{\dagger}_{\vec{k}+\frac{\vec{q}}{2}} c_{\vec{k}-\frac{\vec{q}}{2}}$$

one uses $\Gamma^0_{\vec{q},\vec{k}} = \frac{1}{2} \left(\vec{v}_{\vec{k}+\frac{\vec{q}}{2}} + \vec{v}_{\vec{k}-\frac{\vec{q}}{2}} \right)$, where $\vec{v}_k = \partial \epsilon_{\vec{k}} / \partial \vec{k}$ is the velocity.

To show the importance of vertex correction, let us calculate a simple quantity, the correlation function of the total charge

$$Q = \sum_{\vec{k}} c^{\dagger}_{\vec{k}} c_{\vec{k}}$$

The retarded correlation function is defined by

$$\Pi^{R}_{QQ}(t) = -i\Theta(t)\langle [Q(t), Q(0)] \rangle$$

As $[Q, \mathcal{H}] = 0$ (charge conservation), we obtain $Q(t) = e^{i\mathcal{H}t}Qe^{-i\mathcal{H}t} = Q$ and therefore $\Pi^R_{QQ}(t) = 0$ for all t. Similarly, also the Fourier transformation of this quantity $\Pi^R_{QQ}(\Omega)$ vanishes. This is, however, not obvious when one calculates diagrams for this quantity. Here, we consider the time-ordered correlation function in imaginary

time. This quantity can be calculated diagrammatically



where the dot in the equations above stands for $\vec{k} = \delta_{\vec{k},\vec{k}'}$ as we are calculating the correlation function of $Q = \sum_{\vec{k}} c_{\vec{k}}^{\dagger} c_{\vec{k}}$. In the last line, we used the (three-leg) vertex introduced above with an ingoing momentum of $\vec{q} = 0$.

Up to now, everything was exact. Now, we use a – frequently used, but very dangerous – approximations: we *ignore* vertex correction and consider only the following diagram

$$\Pi_{QQ}(i\Omega) \approx \bigoplus_{\vec{k}} = \sum_{\vec{k}} \frac{1}{\beta} \sum_{i\omega_n} \int \frac{d^3k}{(2\pi)^3} g_{\vec{k}}(i\omega_n) g_{\vec{k}}(i\omega_n + i\Omega)$$
$$\implies \operatorname{Im} \Pi_{QQ}(\Omega + i\epsilon) = \sum_{\vec{k}} \int d\omega \ (n_F(\omega) - n_F(\omega + \Omega)) \operatorname{Im} g_k(\omega) \operatorname{Im} g_k(\omega + \Omega) \neq 0$$

When ignoring vertex correction, we obtain a finite result. In contrast, we have argued above, that the exact result for this quantity is 0. Thus, ignoring vertex correction turned out to be a disaster: we obtained completely wrong result, violating basic conservation laws. This shows, that vertex corrections can become very important. For conserved quantities, there is exact cancellation of vertex and self energy corrections!

How can one avoid these problems? One solution was discussed in the section on controlled calculations. If one does a controlled calculation, considering, e.g., all diagrams to a given order in $\frac{1}{N}$, then, automatically, the results will be correct.

The alternative is to construct approximations to the self-energy and to vertices in a consistent way. This can be accomplished using so-called Luttinger-Ward functionals (or Ward-identities). This is, however, a topic beyond the scope of this lecture.

6 Probing Matter: Linear Response

Motivation: What is measured in a typical experiment? In quantum mechanics courses, we teach that an (idealized) measurement determines the eigenvalues of some Hermitian operator. Very few experiments are, however, of that type with a few notable exceptions like the Stern-Gerlach experiment on electronic spins, a number of experiments in quantum optics, or, most importantly, measurements used in quantum computers. In the accelerators used for high-energy physics, in contrast, one typically measures the outcome of collission of particles (differential cross sections). In experiments on solid state systems, one either measures thermodynamic quantities, or probes matter using electromagnetic waves, electrons or neutrons. One measures, for example, the (optical) conductivity, uses (angular-resolved) photoemission, Raman scattering, Neutron scattering, or scanning tunneling microscopy, to name a few examples.

The common logic behind most of these experiments is that one finds a way to perturb the system slightly and then measures how the system changes in response to this perturbation. A large class of such experiments can be described by **linear** response theory.

6.1 Linear Response Theory and Kubo formula

We consider the following setup. A system is perturbed weakly and one measure the change of some observables $\langle A(t) \rangle$.

$$\mathcal{H} = \mathcal{H}_{0} - \lambda(t) \hat{B} \qquad \text{measure } \langle \hat{A}(t) \rangle \qquad (148)$$

$$\uparrow \qquad \text{some operator} \qquad \text{interacting many} \qquad \text{particle system}$$

The sign is a bit of a convention here, we have chosen it such that $\langle \hat{B} \rangle$ grows for $\lambda > 0$ to minimize the energy.

To solve this time-dependent problem for weak $\lambda(t)$, we consider the time evolution operator \mathcal{U} with

$$i\hbar\partial_t\mathcal{U}=\mathcal{H}\mathcal{U}.$$

which allows to compute the time evolution of an arbitrary wave function, $|\psi(t)\rangle = \mathcal{U}|\psi(0)\rangle$.

For $\lambda = 0$, the time-evolution operator is given by $\mathcal{U}(\lambda = 0) = e^{\frac{-i}{\hbar}\mathcal{H}_0 t}$, for $\lambda \neq 0$ we make the ansatz

$$\mathcal{U} = e^{-\frac{i}{\hbar}\mathcal{H}_0 t} \mathcal{S}(t, t_0)$$

with $S(t_0, t_0) = 1$ assuming that $\lambda = 0$ for $t < t_0$, i.e., before the experiment starts. With this ansatz we obtain

$$i\hbar\partial_t \mathcal{U} = \mathcal{H}_0 \mathcal{U} + e^{-i\mathcal{H}_0 t/\hbar} i\hbar\partial_t \mathcal{S} \stackrel{!}{=} \left(\mathcal{H}_0 - \lambda(t)\hat{B}\right) \mathcal{U}$$
$$\implies i\hbar\partial_t \mathcal{S} = -\lambda(t)e^{i\mathcal{H}_0 t/\hbar}\hat{B}e^{-i\mathcal{H}_0 t/\hbar} \mathcal{S} = -\lambda(t)B_{\mathcal{H}_0}(t)\mathcal{S}(t)$$

with

$$B_{\mathcal{H}_0}(t) = e^{i\mathcal{H}_0 t/\hbar} \hat{B} e^{-i\mathcal{H}_0 t/\hbar}.$$

is the operator in the Heisenberg picture relative computed for the unperturbed system, \mathcal{H}_0

Integrating up the last equation, one obtain S in integral form

$$\begin{split} \mathcal{S}(t,t_0) &= 1 + \frac{i}{\hbar} \int_{t_0}^t dt' \ \lambda(t') B_{\mathcal{H}_0}(t') \mathcal{S}(t',t_0) \\ &\approx 1 + \frac{i}{\hbar} \int_{t_0}^t dt' \ \lambda(t') B_{\mathcal{H}_0}(t') \end{split}$$

In the last line, we took the limit $\lambda \ll 1$, keeping only perturbations to linear order in λ .

Now, we can use the result to compute the change of an observable A. We assume that before the experiment, the system was in a thermal equilibrium state described by the density matrix $\rho_0 = e^{-\beta \mathcal{H}}/Z$. We obtain to linear order in λ

$$\begin{aligned} \langle A \rangle &= \operatorname{Tr} \left(\rho_0 \ \mathcal{U}^{\dagger}(t) A \mathcal{U}(t) \right) \\ &= \operatorname{Tr} \left(\rho_0 \ e^{i\mathcal{H}_0 t/\hbar} A e^{-i\mathcal{H}_0 t/\hbar} \right) \\ &+ \frac{i}{\hbar} \int_{t_0}^t dt' \ \operatorname{Tr} \left(\rho_0 \left(-\lambda(t') B_{\mathcal{H}_0}(t') e^{i\mathcal{H}_0 t/\hbar} A e^{-i\mathcal{H}_0 t/\hbar} + e^{i\mathcal{H}_0 t/\hbar} A e^{-i\mathcal{H}_0 t/\hbar} \lambda(t') B_{\mathcal{H}_0}(t') \right) \right) \\ &= \langle A \rangle_0 + \int_{t_0}^t dt' \ \frac{i}{\hbar} \langle [A_{\mathcal{H}_0}(t), B_{\mathcal{H}_0}(t')] \rangle \lambda(t') \end{aligned}$$

Thus, the change of A, $\delta \langle A \rangle = \langle A(t) \rangle - \langle A \rangle_0$, is given by the so-called **Kubo** formula

$$\delta \langle A(t) \rangle = \int_{-\infty}^{\infty} dt' \ \chi^R_{AB}(t-t')\lambda(t')$$
$$\chi^R_{AB}(t-t') = \frac{i}{\hbar} \Theta(t-t') \langle [A_{\mathcal{H}_0}(t), B_{\mathcal{H}_0}(t')] \rangle_{\mathcal{H}_0}$$

 $\chi^{R}_{AB}(t-t')$ is called a (generalized) **susceptibility** or a **retarded** AB **correlation function**. The susceptibility is a property of the equilibrium system which can be used to describe to linear order the response to a weak perturbation.

Most often, we are interested in the response to periodic perturbation $\lambda(t) \propto \cos(\omega t)$ and therefore need the Fourier transformation of the Kubo formula

$$\delta \langle A(\omega) \rangle = \chi^R_{AB}(\omega) \lambda(\omega)$$

In Sec. 5.3, we have already shown (using the Lehmann representation) that the retarded correlation function showing up in the Kubo formula can be obtained by analytic continuation from the imaginary-time time-ordered correlation function

$$\chi^R_{AB}(\omega) = \chi_{AB}(i\omega_n \mapsto \omega + i\epsilon).$$

An important feature of the Kubo formula is **causality**, encoded in the property that

$$\chi_{\rm AB}(t - t') = 0 \qquad \text{for } t > t'$$

This implies, that when $\lambda(t)$ changes, the response happens only later. We have already discussed in Sec. 5.3, how this property is encoded in the analytic properties of $\chi_{AB}(z)$ in the complex plane, which has to be analytic in the upper half of the complex plane.

6.2 Thermodynamic Stability and Fluctuation-Dissipation Theorem

Motivation: The laws of thermodynamics and the stability of matter imply strong constraints on the properties of correlation functions.

The 2nd law of thermodynamics states, that entropy always increases. We will need a variant of this law, which can be formulated in the following way: we cannot build a machine (perpetuum mobile of the second kind) which extracts work from a system in thermal equilibrium without heating some other reservoir.

To explore this physics, let us calculate how a time-dependent $\lambda(t)$ affects the energy of our system described by the Hamiltonian \mathcal{H}_0 , where $\mathcal{H} = \mathcal{H}_0 - \lambda(t)A$. For small λ we find

$$\frac{d}{dt}\langle \mathcal{H}_0 \rangle = -\frac{i}{\hbar} \langle [\mathcal{H}_0, \mathcal{H}] \rangle = -\frac{i}{\hbar} \lambda(t) \langle [A, \mathcal{H}_0] \rangle$$
$$= \lambda(t) \int dt' \; \frac{\partial}{\partial t} \chi^R_{AA}(t - t') \lambda(t')$$

where we used the Kubo formula for small λ and

$$\frac{\partial}{\partial t}\chi^{R}_{AA}(t) = \frac{\partial}{\partial t} \left(\frac{i}{\hbar} \Theta(t) \langle [A_{\mathcal{H}_{0}}(t), A_{\mathcal{H}_{0}}(0)] \rangle_{\mathcal{H}_{0}} \right)$$
$$= \frac{i}{\hbar} \Theta(t) \langle [-\frac{i}{\hbar} [A_{\mathcal{H}_{0}}(t), \mathcal{H}_{0}], \mathcal{H}_{0}], A_{\mathcal{H}_{0}}(0)] \rangle_{\mathcal{H}_{0}} = -\frac{i}{\hbar} \chi^{R}_{[A, \mathcal{H}_{0}], A}(t)$$

Next, we consider $\lambda(t) = \lambda_0 \cos(\Omega t) = \frac{\lambda_0}{2} (e^{i\Omega t} + e^{-i\Omega t})$ and use that

$$\int dt' \ \chi^R_{AA}(t-t')e^{\pm i\Omega t'} = \int dt' \ d\omega \ e^{-i\omega(t-t')}\chi^R_{AA}(\omega)e^{\pm i\Omega t'} = \chi^R_{AA}(\pm\Omega)e^{\pm i\Omega t}$$

to obtain

$$\frac{d}{dt}\langle \mathcal{H}_0 \rangle = \frac{\lambda_0^2}{2} \cos(\Omega t) \left((-i\Omega) e^{-i\Omega t} \chi_{AA}^R(\Omega) + (i\Omega) e^{i\Omega t} \chi_{AA}^R(-\Omega) \right)$$

When we average this quantity over time, all oscillatory terms drop and we find

$$\overline{\frac{d}{dt}\langle \mathcal{H}_0 \rangle} = \frac{\lambda_0^2}{2} \frac{1}{2} \left((-i\Omega) \chi_{AA}^R(\Omega) + (i\Omega) \chi_{AA}^R(-\Omega) \right)$$
$$= \frac{\lambda_0^2}{2} \Omega \operatorname{Im} \chi_{AA}^R(\Omega)$$

as $\chi^R_{AA}(-\omega) = \chi^R_{AA}(\omega)^*$.

Now, we will not be able to extract energy from the system, therefore the energy \mathcal{H}_0 will always increase. We conclude that

thermodynamics stability
$$\implies \overline{\frac{d}{dt}\langle \mathcal{H}_0 \rangle} \ge 0$$
, energy is dissipated
 $\implies \Omega \operatorname{Im} \chi^R_{AA}(\Omega) \ge 0$ for all Ω

Alternatively, one can also use the Lehmann representation to show that $\Omega \operatorname{Im} \chi^R_{AA}(\Omega) \geq 0$. This shows that thermodynamics is a consistent theory but when one does an approximate calculation this property is not always guaranteed.

Another conclusion is that Im $\chi_{AA}(\Omega)$, is proportional to the absorption of energy and thus describes **dissipation**. The real part, in contrast, gives the reactive response.

We can use a Kramers-Kronig relation , see Sec. 5.3, and $\omega \operatorname{Im} \chi \omega \ge 0$ to compute the sign of the susceptibility evaluated at zero frequency

$$\chi_{AA}^{R}(0) = \operatorname{Re} \chi_{AA}^{R}(0) = \int \frac{d\omega}{\pi} \frac{\operatorname{Im} \chi_{AA}^{R}(\omega)}{\omega} \ge 0$$

This inequality can also be derived either from thermodynamics stability (the system responds by lowering the energy of $\langle H - \lambda A \rangle$) or from the Lehmann representation directly. For finite ω , the real part can be either positive or negative.

Outlook: While susceptibilities as defined above always have to be positive, this is not true for the magnetic susceptibility, $\chi = -\frac{d^2F}{dB^2}$. There is an extra **diamagnetic** contribution, technically arising from the fact that the Hamiltonian is not of the form $\mathcal{H}_0 - \lambda A$ with $\lambda = B$ and A the magnetization (as assumed above), but there is an extra diamagnetic contribution quadratic in the vector potential, discussed in Sec. 6.3.

Next, we will discuss an important exact identity for correlation function, which goes under the name Fluctuation - Dissipation theorem

Motivation: Consider a classical example, the Brownian motion of a particle with coordinate $\vec{R}(t)$ in a liquid. It's equation of motion is described by

$$m\frac{d^2}{dt^2}\vec{R} = \underbrace{-\gamma\frac{d}{dt}\vec{R}}_{\text{friction}} + \underbrace{\vec{F}(t)}_{\substack{\text{random}\\\text{force}}}$$

where γ is the friction constant and F is a random force with time average $\langle F \rangle = 0$ and $\langle F_i(t)F_j(t') \rangle = \delta_{ij}\delta(t-t')C$ and we would like to know the value of the
constant C. Physically, both the friction constant and the random force arise from the scattering of the particle with the bath molecules. Due to the random force, the particle always moves and we can compute the average kinetic energy in ddimensions, $\left\langle \frac{1}{2}m(d\vec{R}/dt)^2 \right\rangle = \frac{dC}{4\gamma}$. From thermodynamics, we know that this average is given by $\frac{d}{2}k_BT$. Thus, we conclude that

$$C = 2\gamma k_B T, \qquad \langle F_i(t)F_j(t')\rangle = \delta_{ij} 2\gamma k_B T \delta(t-t').$$

The strength of the fluctuations of the random force is determined by temperature and by the friction constant γ .

We want to search for a similar relation for quantum system. The amount of fluctuations of an operator A (to simplify notations, we assume $\langle A \rangle = 0$) at a frequency ω is defined by the correlation function

$$C^{>}_{AA}(\omega) = \int_{-\infty}^{\infty} dt \ e^{i\omega t} \langle A(t)A(0) \rangle$$

For this quantity, we derive a Lehmann representation using the exact eigenstates of the Hamiltonian

$$C_{AA}^{>}(\omega) = \sum_{n,m} \int dt \ e^{i\omega t} \frac{1}{\mathcal{Z}} \langle n | e^{-\beta \mathcal{H}} e^{i\mathcal{H}t/\hbar} A e^{-i\mathcal{H}t/\hbar} | m \rangle \langle m | A | n \rangle$$
(149)

$$= \frac{1}{\mathcal{Z}} \sum_{n,m} 2\pi \delta(\omega + (E_n - E_m)/\hbar) e^{-\beta E_n} \langle n|A|m \rangle \langle m|A|n \rangle$$
(150)

We compare this to the Lehmann representation of the imaginary part of the susceptibility derived in Sec. 5.3

$$\operatorname{Im} \chi_{AA}(\omega) = \frac{1}{\hbar \mathcal{Z}} \sum_{n} \pi \delta(\omega - (E_m - E_n)) \underbrace{\left(e^{-\beta E_n} - \xi e^{-\beta E_m}\right)}_{e^{-\beta E_n} (1 - \xi e^{-\beta \hbar \omega})} |\langle n|A|m \rangle|^2$$

and obtain

$$C_{AA}^{>}(\omega) = 2 \frac{\operatorname{Im} \chi_{AA}^{R}(\omega)}{1 - \xi e^{-\beta\omega}}.$$

This is the **fluctuation-dissipation theorem**. For bosonic operators, $\xi = 1$, one can use the Bose function $n_B(E) = 1/(e^{\beta E} - 1)$ to write the fluctuation-dissipation theorem in the form

$$C_{AA}^{>}(\omega) = 2\hbar \left(1 + n_B(\hbar\omega)\right) \operatorname{Im} \chi_{AA}^{R}(\omega)$$

While the left-hand side describes the fluctuation of the system, $\text{Im} \chi^R_{AA}(\omega)$ gives information on dissipation, as discussed above.

It is useful, to explore the fluctuation-dissipation theorem also in the **classical limit**

$$\hbar\omega \ll k_B T.$$

For small ω , the imaginary part of the susceptibility is linear in ω and therefore we define the friction constant

$$\gamma_A = \lim_{\omega \to 0} \frac{\operatorname{Im} \chi^R_{AA}(\omega)}{\omega}$$

with $\frac{\operatorname{Im}\chi_{AA}(\omega)}{\omega} \approx \gamma_A = const.$ Using that $\lim_{\omega \to 0} (1 + n_B(\hbar \omega) = \frac{k_B T}{\hbar \omega}$, we obtain the fluctuation-dissipation theorem in the classical limit

$$C^{>}_{AA}(\omega) \approx 2k_B T \gamma_A \quad \iff \quad \langle A(t)A(0) \rangle \approx 2k_B T \gamma_A \,\delta(t-t')$$

This is precisely the form of the fluctuation-dissipation theorem which we discussed above for the Brownian motion of a particle.

Finally, we want to discuss special problems which arise when calculating the **sus-ceptibility of conserved quantities**. The number of electrons, N, is, for example, conserved. The energy is another example of a conserved quantity. Or in models without spin-orbit coupling the magnetization is conserved.

If $[A, \mathcal{H}_0] = 0$, then A(t) = A and

$$\chi^R_{AA}(t) = -i\langle [A(t), A(0)] \rangle = -i\langle [A, A] \rangle = 0$$

This simply reflects that a conserved quantity does not change.

The static, thermodynamic response, however, is finite. Consider the equilibrium density matrix

$$\rho(\lambda) = rac{e^{-eta(\mathcal{H}_0 - \lambda A)}}{\mathcal{Z}(\lambda)}.$$

We can compute the change of A in response to a small change of λ .

$$\langle A \rangle_{\lambda} - \langle A \rangle_{\lambda=0} = \operatorname{Tr}\left((\rho(\lambda) - \rho(0))A\right) \stackrel{[A,\mathcal{H}_0]=0}{=} \underbrace{\beta\left(\langle A^2 \rangle - \langle A \rangle^2\right)}_{\chi_{AA}^{\mathrm{th}}} \lambda = \chi_{AA}^{\mathrm{th}} \lambda$$

This defines the **thermodynamic susceptibility** χ^{th}_{AA} .

How can we compute the thermodynamic susceptibility using retarded correlation functions? The trick is to consider the response at finite momentum. Instead of the conserved quantity A, we investigate

$$A_{\vec{q}} = \int d\vec{x} \ e^{i\vec{q}\vec{x}}a(\vec{x})$$

with $A_{\vec{q}=0} = A$ and $[\mathcal{H}_0, A_q] = 0$ only for $\vec{q} = 0$. Here, $a(\vec{x})$ is the density of the conserved quantity, e.g., the charge density for A = Q. Now, we can define and compute the susceptibility at finite \vec{q}

$$\chi_q^R(\omega) = -i \int dt \; \Theta(t) e^{i\omega t} \langle [A_q(t), A_q(0)] \rangle$$

While $\lim_{\omega\to 0} \lim_{q\to 0} \chi_q(\omega) = 0$, as discussed above, we obtain the thermodynamic susceptibility taking the limits the other way around

$$\chi_{AA}^{\rm th} = \lim_{q \to 0} \lim_{\omega \to 0} \chi_q(\omega).$$

The situation is exactly opposite, when one calculates the dc conductivity of, e.g., a metal (see next chapter). In this case, one studies the response to an electric field $\vec{E}(\vec{x},t) = \vec{E}_{\vec{q}}(\omega) \operatorname{Re} e^{i(\vec{q}\vec{r}-\omega t)}$ with

$$\vec{j}_{\vec{q}}(\omega) = \sigma_{\vec{q}}(\omega)\vec{E}_{\vec{q}}(\omega).$$

In this case, there is no current flowing for a static potential. Thus, one finds that $\lim_{q\to 0} \lim_{\omega\to 0} \sigma_q(\omega) = 0$. To compute the **dc conductivity** for a uniform electric field, one takes the opposite limit

$$\sigma = \lim_{\omega \to 0} \lim_{\vec{q} \to 0} \sigma_{\vec{q}}(\omega).$$

The **transport limit**, first $\vec{q} \to 0$ and then $\omega \to 0$ is opposite to the **susceptibility limit**, first $\omega \to 0$ and then $\vec{q} \to 0$, discussed above.

6.3 Optical Conductivity

Arguably, the most important quantity of a metal is its **conductivity**. It is defined by the linear relation of current and electric field (valid for weak electric fields)

$$\langle j_i(t) \rangle = \int dt' \ \sigma_{ij}(t-t') E(t')$$
$$\langle j_i(\omega) \rangle = \sigma_{ij}(\omega) E_j(\omega),$$

where \vec{j} is the current density. The frequency-dependent $\sigma(\omega)$ is called the optical conductivity.

How does one measure $\sigma(\omega)$? An apparent problem is that the electric field inside a metal will typically not be the same as the field outside of a metal. In practice, one therefore measures the resistivity, $\rho(\omega) = \sigma(\omega)^{-1}$, defined by $E_i = \rho_{ij}j$. Here, the current density is imprinted on the system externally and one measures the drop of the voltage.



Four contacts are needed for this experiment to inject and extract currents and to measure the voltage drop, thus it is called a 4-point probe measurements. It has the advantage that there is no contribution from the resistivity of the contacts.

For large frequencies, in contrast, one measures instead the transmission and/or reflection of light.



From this one can extract the complex-valued dielectric constant $\epsilon(\omega)$. Here, one starts from the Maxwell equation $\frac{1}{\mu_0} \vec{\nabla} \times \vec{B} = \vec{j} + \epsilon_0 \partial_t \vec{E}$. After a Fourer transformation, the right-hand side is written as

$$\vec{j}(\omega) - i\omega\epsilon_0 \vec{E}(\omega) = (\sigma(\omega) - i\omega\epsilon_0)\vec{E}(\omega) \equiv (-i\omega)\epsilon(\omega)\epsilon_0 \vec{E}(\omega)$$

which allows to identify

$$\epsilon(\omega) = 1 - \frac{\sigma(\omega)}{i\omega\epsilon_0}.$$

Equivalently, one can start from $\vec{D} = \epsilon_0 \epsilon \vec{E} = \epsilon_0 \vec{E} + \vec{P}$, identifying $\partial_t \vec{P} = \vec{j}$.

How does a time-dependent electric field enter the Hamiltonian? We can start from

$$\vec{E} = -\vec{\nabla}\phi - \dot{\vec{A}}$$

and either introduce a space-dependent potential $\phi(\vec{x}) = -\vec{x}\vec{E}$ or a constant vector potential,

$$\vec{A}(\omega) = \frac{\vec{E}(\omega)}{i\omega}.$$

while setting $\Phi = 0$ (by choosing the appropriate Gauge). We choose the latter option to avoid a space-dependent perturbation.

Gauge-invariance largely fixes how the vector potential enters the Hamiltonian. For a single-band model with dispersion $\mathcal{E}(\vec{p})$ one simply has to replace \vec{p} by $\vec{p} - e\vec{A}$ (minimal substitution). Thus, the Hamiltonian in the presence of a vector potential is given by

$$\begin{aligned} \mathcal{H} &= \sum_{\vec{p},\sigma} c^{\dagger}_{\vec{p}\sigma} \, \mathcal{E} \! \left(\vec{p} - e\vec{A} \right) \, c_{\vec{p}\sigma} + \mathcal{H}_{\text{int}} \\ &\approx \mathcal{H}_0 + \mathcal{H}_{\text{int}} - \vec{\mathcal{J}} \cdot \vec{A} + \frac{1}{2} d_{ij} A_i A_j + O(A^3) \end{aligned}$$

where we used a Taylor expansion in \vec{A} in the second line with

$$\mathcal{J}_{i} = e \sum_{\vec{p},\sigma} c^{\dagger}_{\vec{p}\sigma} \frac{\partial \mathcal{E}(\vec{p})}{\partial p_{i}} c_{\vec{p}\sigma}, \qquad d_{ij} = e^{2} \sum_{\vec{p},\sigma} c^{\dagger}_{\vec{p}\sigma} \frac{\partial^{2} \mathcal{E}(\vec{p})}{\partial p_{i} \partial p_{j}} c_{\vec{p}\sigma}.$$

The (gauge invariant) physical current operator is given by

$$\mathcal{J}_i^{\mathrm{phy}} = -\frac{\partial \mathcal{H}}{\partial A_i} = \mathcal{J}_i - d_{ij}A_j.$$

How do we know that? We can either derive this formula (we are not doing this here) using the continuity equation, $\partial_t \rho + \vec{\nabla} \vec{j} = 0$, or derive the Maxwell equation

as an Euler-Lagrange equation starting form $S = S_{\text{matter}} + S_{\text{Maxwell}}$. This allows to identify the term above as the current.

Next, we can use the Kubo formula to compute the current induced to linear order in \vec{A}

$$\langle \mathcal{J}_{i}^{\mathrm{phy}(t)} \rangle = \int_{\substack{\uparrow \\ \text{use gauge} \\ \text{with } \phi = 0}} dt' \left(\chi_{\mathcal{J}_{i}\mathcal{J}_{j}}^{R}(t-t') - \underbrace{\langle d_{ij} \rangle \delta(t-t')}_{\text{diamagnetic}} \right) A_{j}(t')$$

Using $j_i = \sigma_{ij} E_j$, $j_i = \langle \mathcal{J}_i^{\text{phy}(t)} \rangle / V$ (V is the volume), $\vec{A}(\omega) = \frac{E(\omega)}{i\omega}$, we obtain

$$\sigma_{ij}(\omega) = \frac{1}{V} \frac{\Pi_{ij}(\omega) - \langle d_{ij} \rangle}{i\omega}, \qquad \Pi_{ij}(\omega) = \chi^R_{\mathcal{J}_i \mathcal{J}_j}(\omega)$$

The conductivity is thus computed from a current-current correlation function. In metals (but not in superconductors), the relation $\Pi_{ij}(\omega = 0) = \langle d_{ij} \rangle$ has to hold to obtain a finite conductivity at $\omega = 0$. This can be derived using the continuity equation for charge. As sometimes approximations do not fulfill this relation, one often only computes the real-part of the conductivity (obtaining the imaginary part from a Kramers-Kronig relation if necessary)

$$\operatorname{Re} \sigma(\omega) = \frac{1}{V} \frac{\operatorname{Im} \Pi(\omega)}{\omega}.$$

As an example, let us calculate $\operatorname{Re} \sigma(\omega = 0)$ without vertex corrections. This is, in general, an unjustified approximation as we will discuss in more detail below. Nevertheless, it allows to obtain quickly a useful result. The current operator for a single band with dispersion $\epsilon_{\vec{k}}$ is given by $J_i = e \sum_{\vec{k},\sigma} v_{\vec{k}}^i c_{\vec{k},\sigma}^\dagger c_{\vec{k},\sigma}$ with the velocity $v_{\vec{k}}^i = \partial \epsilon_{\vec{k}} / \partial k_i$. Neglecting vertex corrections, we obtain

$$\Pi_{ij}(i\Omega) \approx \bigcirc = \frac{1}{V} \sum_{\vec{k},\sigma} \frac{1}{\beta} \sum_{i\omega_n} \int \frac{d^3k}{(2\pi)^3} ev_{\vec{k}}^j g_{\vec{k}}(i\omega_n) ev_{\vec{k}}^i g_{\vec{k}}(i\omega_n + i\Omega)$$

After performing the sum and analytic continuation, we find

$$\operatorname{Im}\Pi_{ij}(\Omega+i\epsilon) = 2\sum_{\vec{k},\sigma} \int \frac{d\omega}{\pi} e^2 v_{\vec{k}}^i v_{\vec{k}}^j \left(n_F(\omega) - n_F(\omega+\Omega)\right) \operatorname{Im} g_k(\omega) \operatorname{Im} g_k(\omega+\Omega)$$

Without an external magnetic field and for a system with, e.g., cubic symmetry, the off-diagonal terms vanish. Taking the limit $\Omega \to 0$, we find

$$\sigma_{ii} = \lim_{\Omega \to 0} \frac{\operatorname{Im} \Pi_{ii}(\Omega)}{\Omega} = 2 \int \frac{d^3k}{(2\pi)^3} \int \frac{d\omega}{\pi} \ e^2 (v_{\vec{k}}^i)^2 \left(-n'_F(\omega)\right) (\operatorname{Im} g_k(\omega))^2.$$

For T much smaller than the Fermi energy, we can approximate $-n'_F(\omega) = \delta(\omega)$ to perform the ω integral. All contributions to the integral come close to the Fermi energy and thus we evaluate both the self-energy and the velocities at the Fermi energy and at $\omega = 0$. Furthermore, we multiply with $1 = \int d\epsilon \delta(\epsilon - \epsilon_{\vec{k}})$ and replace $\epsilon_{\vec{k}} = \epsilon$ in the integral

$$\int d\epsilon \, (\operatorname{Im} g_k(\omega=0))^2 = \int d\epsilon \left(\frac{\Gamma_{k_F}}{(0-\epsilon - \operatorname{Re} \Sigma_{k_F}(0))^2 + \Gamma_{k_F}^2} \right)^2 = \frac{\pi}{2} \frac{1}{\Gamma_{k_F}}$$

with the scattering rate evaluated at the Fermi surface

$$\Gamma_{k_F} = \operatorname{Im} \Sigma_{k_F}(\omega = 0).$$

The same integral ensures that ϵ is always small, and we can identify $2 \int \frac{d^3k}{(2\pi)^3} \delta(\epsilon - \epsilon_{\vec{k}}) = N_F$ with the total density of states at the Fermi energy, N_F . Combining everything, we obtain

$$\sigma \approx e^2 N_F \left\langle \frac{(v_{\vec{k}}^i)^2}{2\Gamma_{k_F}} \right\rangle_{\rm FS}$$

where $\langle a_{\vec{k}} \rangle_{\rm FS} = \frac{\int d^3k \, \delta(\epsilon_{\vec{k}}) a_{\vec{k}}}{\int d^3k \delta(\epsilon_{\vec{k}})}$ denotes an average over the Fermi surface. For a spherical Fermi surface, $\epsilon_k = \hbar^2 k^2 / (2m)$, the density of states at $\mu = \hbar^2 k_F^2 / (2m)$ is given by $2 \int \frac{d^3k}{(2\pi)^3} \delta(\frac{\hbar^2 k^2}{2m} - \mu) = 2 \frac{4\pi}{(2\pi)^3} \frac{k_F^2}{\hbar^2 k_F / m}$ and $\left\langle (v_{\vec{k}}^i)^2 \right\rangle_{\rm FS} = \frac{1}{3} v_F^2 = \frac{1}{3} \left(\frac{\hbar k_F}{m}\right)^2$. Therefore, we obtain $4\pi k^3 = 1$

$$\sigma \approx e^2 2 \frac{4\pi}{3} \frac{k_F^3}{m} \frac{1}{2\Gamma_{k_F}}$$

Using the density of electrons, $n = 2\frac{4}{3}\pi k_F^3 \frac{1}{(2\pi)^3}$, this result can be written as

$$\sigma \approx \frac{e^2 n}{m} \tau, \qquad \frac{1}{\tau} = 2\Gamma_{k_I}$$

This is the famous Drude formula, where we identify the scattering time by $\tau = 1/(2\Gamma_{k_F})$. The phenomenological Drude formula was derived phenomenologically in 1900 a long time before it was know that electrons are fermions. Nevertheless, it captures surprisingly main features of the conduction of electrons. Importantly, the conductivity is inversely proportional to the scattering rate.

Outlook: The result given above have been derived without taking vertex corrections into account (which often require to solve integrale equations, which is much more work). How good or how bad the approximation is, depends on the mechanism responsible for electron-electron scattering. For short-ranged disorder potentials (described by δ -functions), for example, it turns out that vertex correction vanish exactly. An other extreme case is a system with perfect momentum conservation, where the conductivity is infinite, while $\Gamma_{k_F} \propto T^2$ is finite. In this case, the result derived above is completely wrong, because vertex correction cancel exactly the selfenergy corrections considered above. In most realistic situations (e.g., scattering from impurity potentials with a finite-range or electron-electron scattering for large Fermi surfaces), vertex corrections change the prefactor of the conductivity by a factor of order 1. Thus, the calculation given above is at least qualitatively correct.

6.4 Scattering Experiments

Motivation: There is an important class of experiments, which cannot directly be described by the Kubo formula. These are scattering experiments, where one shoots particles or waves onto a target and observes what comes out, as shown in the sketch



Examples for such experiments are X-ray scattering, neutron scattering, or (angular-resolved) photo emission.

In these experiments, one effectively measures that scattering probability or transition rate

$$\Gamma_{\vec{q}_{\mathrm{out}},\vec{q}_{\mathrm{ir}}}$$

that an incoming particle with momentum \vec{q}_{in} produces an outgoing particle with momentum \vec{q}_{out} .

During such a scattering experiment, the momentum and energy transfer

$$\Delta \vec{q} = \vec{q}_{\rm in} - \vec{q}_{\rm out}, \qquad \Delta E = \mathcal{E}(\vec{q}_{\rm in}) - \mathcal{E}(\vec{q}_{\rm out})$$

describe how much momentum and energy is deposited to the target. Here $\mathcal{E}(\vec{q})$ is the energy of the test-particle, e.g., of a neutron.

If we assume that the interaction of the test-particle with the solid is weak (e.g., for neutrons), we can use perturbation theory to calculate the scattering rate. More precisely, we use **Fermi's golden rule** formula (which the reader has probably derived in their quantum mechanics course)

$$\Gamma_{\vec{q}_{\text{out}},\vec{q}_{\text{in}}} = \frac{2\pi}{\hbar} \sum_{i,f} |\langle \vec{q}_{\text{out}}, f | \Delta \mathcal{H} | \vec{q}_{\text{in}}, i \rangle|^2 \, \delta \left(\mathcal{E}(\vec{q}_{\text{in}}) + E_i - \left(\mathcal{E}(\vec{q}_{\text{out}}) + E_f \right) \right) \underbrace{\frac{e^{-\beta E_i}}{\mathcal{Z}}}_{\substack{\text{occupation of initial states}}}$$

Here, $|\vec{q}_{\rm in}, i\rangle$ ($|\vec{q}_{\rm out}, f\rangle$ denotes the initial (final) state which is the product of an incoming plane wave of the test particle (e.g., a neutron) and the eigenstates $|i\rangle$, $|f\rangle$ eigenstates of \mathcal{H}_0 , the Hamiltonian of the system which is probed.

In the following, we use the of first quantization describing the neutron position with \vec{R} and its spin with neutron spin \vec{S} . The neutron interacts, for example, locally the nuclei in the solid, $\psi_n^{\dagger}(\vec{r})\psi_n(\vec{r})$, as described by the Hamiltonian

$$\Delta \mathcal{H}_c = \int d^3 r \ V(\vec{R} - \vec{r}) \rho_n(\vec{r}).$$

Another important coupling arises from dipole-dipole interactions of the neutron spin and the electron spins

$$\Delta \mathcal{H}_s = \sum_{ij} \int d^3 r \underbrace{S_i U^{ij}(\vec{R} - \vec{r})}_{\text{dipole-dipole}} \underbrace{\psi^{\dagger}_{\alpha}(\vec{r}) \sigma^{j}_{\alpha\beta} \psi_{\beta}(\vec{r})}_{\text{electron spin}} \underbrace{\psi^{\dagger}_{\alpha}(\vec{r}) \sigma^{j}_{\alpha\beta} \psi_{\beta}(\vec{r})}_{\text{density}}$$

To evaluate the matrix element in the Golden-rule formula for, e.g. $\Delta \mathcal{H}_c$, we first compute,

$$\begin{aligned} \langle \vec{q}_{\text{out}} | \Delta \mathcal{H}_c | \vec{q}_{\text{in}} \rangle &= \int d^3 R d^3 r \ e^{i(\vec{q}_{\text{in}} - \vec{q}_{\text{out}})\vec{R}} V(\vec{R} - \vec{r}) \hat{\rho}_n(\vec{r}) \\ &= V_{-\Delta \vec{q}} \ \hat{\rho}_{-\Delta \vec{q}} \end{aligned}$$

For $\Delta \mathcal{H}_s$, we have to take the neutron spin into account (which we suppressed in the formulas given above). Denoting the neutron wave function by $|\vec{q}, \alpha\rangle$, $\alpha = \uparrow, \downarrow$, we identify

$$\langle \vec{q}_{\text{out}} \alpha_{\text{out}} | \Delta \mathcal{H}_s | \vec{q}_{\text{in}} \alpha_{\text{in}} \rangle = \sum_{i,j=1,2,3} \frac{\sigma_{\alpha_{\text{out}}\alpha_{\text{in}}}^i}{2} \ U_{-\Delta \vec{q}}^{ij} \ \hat{s}_{-\Delta \vec{q}}^j$$

where the operator $s^{j}_{-\Delta \vec{q}}$ is the Fourier transform of the spin density $\sum_{\alpha\beta} \psi^{\dagger}_{\alpha}(\vec{r}) \sigma^{j}_{\alpha\beta} \psi_{\beta}(\vec{r})$ of the electrons.

Let us first consider the case, where \mathcal{H}_c dominates. In this case, we find

$$\Gamma = \frac{2\pi}{\hbar} \sum_{i,f} |\langle f | V_{-\Delta \vec{q}} \, \hat{\rho}_{-\Delta \vec{q}} | i \rangle|^2 \frac{e^{-\beta E_i}}{\mathcal{Z}} \delta(E_f - E_i - \Delta E)$$

which is the same type of formula, which we have seen for the Lehmann representation of the correlation function $C^{>}(\omega)$ in Eq. (150). Thus, we can rewrite it as

$$\Gamma = \frac{2\pi}{\hbar} |V_{\Delta \vec{q}}|^2 \int_{-\infty}^{\infty} \frac{dt}{2\pi} \operatorname{Tr} \left(e^{\beta \mathcal{H}} e^{i\mathcal{H}t} \hat{\rho}_{\Delta \vec{q}} e^{-i\mathcal{H}t} \hat{\rho}_{-\Delta \vec{q}} \right) e^{i\Delta Et}$$
$$= \frac{1}{\hbar} |V_{\Delta \vec{q}}|^2 \underbrace{\int dt \ \langle \hat{\rho}_{\Delta \vec{q}}(t) \hat{\rho}^*_{\Delta \vec{q}}(0) \rangle e^{i\Delta Et}}_{C^>(\Delta E)}$$

Next, we use the **fluctuation-dissipation theorem** to express it in terms of the imaginary part of a charge-charge susceptibility

$$\Gamma_c(\Delta \vec{q}, \Delta E) = \frac{2\pi}{\hbar} |V(\Delta \vec{q})|^2 \cdot 2 \left(1 + n_B(\Delta E)\right) \operatorname{Im} \chi^R_{\rho^*_{\Delta \vec{q}} \rho_{\Delta \vec{q}}}(\Delta E),$$

A similar formula applies for the contribution arising from the spin-spin interaction. In this case, the result depends on the polarization of the incoming and outgoing spin which can be worked out using the formulas given above (not shown here for simplicity)

$$\Gamma_s(\Delta \vec{q}, \Delta E) \propto (1 + n_B(\Delta E)) \operatorname{Im} \chi^R_{s^*_{\Lambda \vec{\sigma}} * s^*_{\Lambda \vec{\sigma}}}(\Delta E)_s$$

In most of neutron scattering experiments, one effectively averages over the spin orientations. If one uses **polarized neutron scattering** instead, one obtains information on the spin-orientation, which can, e.g., be used to distinguish the signal arising from the two scattering mechanisms.

Thus, we found that **neutron scattering measures density-density correlations of nuclei and spin-spin correlations of electrons**. Taking into account, that $\operatorname{Im} \chi^{R}_{\rho^{*}_{\Delta \vec{q}} \rho_{\Delta \vec{q}}}(\Delta E)$ is an odd function of ΔE , we can compute the ratio of the probability that the neutron either gains or loses the energy ΔE .

$$\frac{\Gamma(\Delta \vec{q}, \Delta E)}{\Gamma(\Delta \vec{q}, -\Delta E)} = \frac{1 + n_B(\Delta E)}{-(1 + n_B(-\Delta E))} = \frac{\frac{e^{\beta \Delta E}}{e^{\beta \Delta E} - 1}}{-\frac{1}{1 - e^{\beta \Delta E}}} = e^{\beta \Delta E}$$

This is an example of a **detailed balance** relation. Such relations follow from thermal equilibrium and ensure that two reservoirs interacting via the exchange of neutrons or other type of radiation equilibrate with each other.

Above, we discussed neutron scattering but very similar formulas apply for other scattering experiment. One important example is **angular resolved neutron scattering**, where one excites the system by absorbing photons (e.g., from a laser or a synchrotron) and measures the direction and energy of an outgoing electron. The interaction of the electromagnetic wave with matter is described by

$$\Delta H = -\int dx \ \vec{j}(x)\vec{A}(x) + \mathcal{O}(A(x)^2) \sim \sum \Gamma_{\vec{k},\vec{q}} c^{\dagger}_{\vec{k}+\vec{q}} c_{\vec{k}} a_{\vec{q}},$$

where we used that the vector potential is linear in the photon creation and annihilation operators, $a_{\vec{q}}^{\dagger}, a_{\vec{q}}$, and we omitted all spin- and polarization quantum numbers. The precise form of the matrix element $\Gamma_{\vec{k},\vec{q}}$ is not important here. If the energy of the outgoing electron is sufficiently large, one can approximately ignore the interacting of the outgoing electron and approximate the final state in the Golden-rule formula as $c_{\vec{q}_{out}}^{\dagger}|f\rangle$, while the initial state is given by $a_{\vec{q}_{in}}^{\dagger}|i\rangle$, where $|f\rangle$, $|i\rangle$ are eigenstates of the measured system. In this case, from ΔH only one term contributes

$$\Delta H \propto c_{\vec{q}_{\text{out}}}^{\dagger} c_{\vec{q}_{\text{out}} - \vec{q}_{\text{in}}} a_{\vec{q}_{\text{in}}}.$$

Thus, angular resolved photoemission measures the correlation function of $c_{\vec{q}_{out}-\vec{q}_{in}}$. Repeating the derivation given above, one obtains that

$$\Gamma(\Delta \vec{q}, \Delta E) \propto n_F(-\Delta E) A_{-\Delta \vec{q}}(-\Delta E), \quad A_{\Delta \vec{q}}(\omega) = \operatorname{Im} g_{\Delta \vec{q}}(\omega)$$

Thus, angular resolve neutron scattering directly measures the fermionic **spectral function** multiplied with a Fermi function.

There are many other experiments, which can be understood from the calculation of correlation function. In a scanning tunneling microscopy (STM) experiment, the term ΔH describes the tunneling from the tip to the sample and one measures the current *I* through the probe as a function of the applied voltage. Using again perturbation theory in ΔH , one can show that

$$\frac{\partial I}{\partial V} \propto \operatorname{Im} \mathcal{G}^{R}_{\vec{r},\vec{r}}(\omega)$$

where \vec{r} is the position of the tip and $\mathcal{G}^{R}_{\vec{r},\vec{r}}(\omega)$ is the real-space Green's function where an electron is created and destroyed at position \vec{r} .

Another class of experiments are Nuclear Magnetic Resonance (NMR), Electron Spin Resonance (ESR), or Muon Spin Resonance (μ SR), where one measures how oscillating magnetic fields are absorbed by local (nuclear, electronic or muon) spins inside of the material in the presence of an external field. Effectively, one measures the coupling of the spins to their environment, governed by local spin-spin correlation functions. One important application of NMR is Magnetic Resonance Imaging (MRI), used in medicine.

There are, however, situations where this paradigm cannot be applied. What can one do, if

6.5 Outlook: Beyond Linear Response

Many experiments build on the concept of weakly perturbing the system. We have shown that they can be described by equilibrium correlation functions. We have focused on the response linear in the external perturbation, but one can generalize the concept also to non-linear effects, by calculating the response to 2nd or 3rd order in the applied perturbation. Even then, one can start from an imaginary time path integral and use analytic continuation techniques to obtain the real-time response. This strategy fails, however, when the system is far from thermal equilibrium. Consider, for example, the problem of a **quench**, defined by the sudden change of a Hamiltonian. Starting from an initial state with a density matrix $\rho_0 = e^{-\beta H_0}/Z_0$, one investigates the time-evolution with a different Hamiltonian $H \neq H_0$,

$$\rho(t) = e^{-iHt/\hbar} \rho_0 e^{iHt/\hbar}.$$

Expectation values, for example, are then computed from $\langle A(t) \rangle = tr[A\rho(t)]$.

Here, one can apply our standard approach to derive a function integral directly for $\rho(t)$ by applying a Trotter decomposition for both time evolution operators $e^{\pm iHt/\hbar}$ to the left and right of ρ_0 (and, possibly also a Trotter decomposition of ρ_0 .

The resulting theory is called **Keldysh formalism**. Due to the two time-evolution operators, it includes two time variables (called contours) running from 0 to t in the example discussed above. Thus, Green's function are promoted to 2×2 matrices $G_{\alpha_1\alpha_2}(t_1, t_2)$, where α_1 and α_2 encode on which of the two time contours t_1 and t_2 are located. One obtains the same type of Feynman diagrams with slightly modified Feynman rules, which take into account the matrix structure of the Keldysh Green's functions.

Thus, one can copy many resummation techniques from equilibrium problems. For example, one arrives at a quantum version of the Boltzmann equation, when one combines the Dyson equation with a self-consistent evaluation of the self-energy. The Keldysh formalism – in combination with further diagrammatic approximations – is therefore often the starting point to describe quantum systems far from thermal equilibrium.