Contents

0.1 Outline

- Phase transitions in thermodynamics
- Symmetry-breaking transitions
- If we do not know much: Landau approach
  - free energy and and symmetry considerations
  - second- and first-order transitions
  - coupled order parameters
  - spatial variantion: Ginzburg-Landau
- If we have a microscopic model: e.g. mean-field theory
- gunnea pig: Ising model
- How valid is maen-field theory?
- excitations

0.2 Literature

- *Your favorite texbook on statistical mechanics*
- *For Solid-State Physics:* Daniel I. Khomskii , Basic Aspects of the Quantum Theory of Solids: Order and Elementary Excitations, Cambridge University Press, 2010
- *Quantum-Field-Theory focussed, mostly more advanced:*, A. Altland and Ben Simons, Condensed Matter Field Theory, Cambridge University Press, 2006

1 Symmetry breaking: Magnetism and Superconductivity

1.1 Magnetism

Inspired by Khomskii's book.

When describing the interaction of electrons with electro-magnetic field, 'minimal coupling' is usually employed. Momentum operator $\hat{p}$ is here replaced by $\hat{p} + c\vec{A}(\hat{x})$, where vector potential $\vec{A}$ is an operator via its dependence on $\hat{x}$ and the prefactor c depends on gauge choice. In Coulomb gauge $\vec{\nabla}\vec{A} = 0$ so that $\hat{p}$ only acts on the wave function and commutes with $\vec{A}$. The kinetic energy in the Hamiltonian then contains terms

- $\propto \hat{p}^2$: Usual kinetic energy, together with the potential from the ions, which is still there, it gives the usual band energies $\epsilon_{\vec{k}}$.
- $\propto \hat{p}\vec{A}$: Closer analysis (see Pauli equation) shows that this term can be written as $\vec{B}\cdot(\vec{L}+2\vec{S})$. Such an interaction between magnetic field and existing magnetic moments is *paramagnetism*.
- $\propto \vec{A}^2$: These terms $\propto \vec{B}^2$ do not involve existing magnetic moments, but can be interpreted as the interaction between $\vec{B}$ and induced moments $\propto B$. The induced moments are opposite to the magnetic field, this is *diamagnetism*.

We will here not discuss diamagnetism further and will also not have time to devote to the quantum Hall effect, but will instead focus on paramagnetism and magnetic order.

1.1.1 Paramagnetism: Existing moments without interactions

First, we will discuss how the electron spin reacts to a magnetic field, in two limits: a metal with itinerant and non-interacting electrons and an insulator with localized moments. Orbital angular momentum is here left out, because for itinerant electrons, it is often 'quenched', i.e., forced to be 0, which is certainly the case in s bands. The localized-moment discussion carries over to general angular momenta without much change.

1.1.1.1 Magnetic susceptibility of non-interacting electrons

Non-interacting electrons were discussed in Sec. **??**: quantities can be obtained from the density of states by filling it using the Fermi function. The magnetic susceptibility expresses how strong a magnetization can be induced by a given magnetic field, i.e., by the difference in particle numbers for up and down electrons. For non-interacting electrons, it is determined by the balance between the magnetic energy that can be gained by turning some 'down' electrons into'up' and the energy that is lost by then having to occupy higher-energy states due to the Pauli principle.

A magnetic field $\vec{B}$ along z shifts the one-particle energies $\epsilon(\vec{k}) \to \epsilon(\vec{k}) \pm \frac{g\mu_B}{2}|B|$. Both the up and the down states are then filled up to some chemical potential that is the same for both spins in equilibrium μ. (Otherwise, electrons would flip spin until it evens out.) However, when filling electrons into states, it is technically equivalent to not shift the bands and instead make the chemical potential spin dependent $\mu \to \mu \mp \frac{g\mu_B}{2}|B|$.

Using the Sommerfeld approach, we had seen in Eq. (**??**) that the number of electrons can be obtained by (i) filling the states up to the chemical potential $\mu(T)$ and (ii) setting $\mu(T) = E_F$. This approximate approach instead of a more accurate calculation neglects corrections quadratic in temperature T. Making the particle number spin dependent and noting that the density of states per spin is half the total $\rho_\sigma(\omega) = \frac{1}{2}\rho(\omega)$, we find

$$\frac{N_\sigma}{N} = \int_0^{E_F \pm \frac{g\mu_B}{2}\sigma|B|} \mathrm{d}\omega\, \rho_\sigma(\omega) + \mathcal{O}(T^2) = \frac{1}{2} \int_0^{E_F \pm \frac{g\mu_B}{2}\sigma|B|} \mathrm{d}\omega\, \rho(\omega) + \mathcal{O}(T^2)\,, \tag{1.1}$$

with $\sigma = \pm 1$ for up and down. In the magnetization, most of the integrals cancel and only the density of states around E_F actually contributes, giving the approximation

$$M = \frac{N_\uparrow - N_\downarrow}{N} = \frac{1}{2} \int_{E_F - \frac{g\mu_B}{2}|B|}^{E_F + \frac{g\mu_B}{2}|B|} \mathrm{d}\omega\, \rho(\omega) + \mathcal{O}(T^2) = \frac{g\mu_B}{2}|B|\rho(E_F) + \mathcal{O}(T^2) \quad \text{and}$$

$$\chi = \frac{\partial M}{\partial B} = \underbrace{\frac{g}{2}}_{\approx 1} \mu_B \rho(E_F) + \mathcal{O}(T^2)\,. \tag{1.2}$$

Temperature thus only enters in second order.

Up to second order in temperature, the susceptibility is a constant determined by the density of states at the Fermi level. A large susceptibility can be obtained in systems with many states near the Fermi level: small 'distortions' of the bands can then have a large impact, a mechanism similar to that discussed in Sec. **??** as favoring ordering. The scenario discussed here applies to noninteracting bands, more specifically to metals, because the $\rho(E_F) = 0$ in insulators and their susceptibility would thus vanish. Due to the approximation made, it can only be considered valid if the magnetic energy scale $\mu_B B$ is much smaller than the Fermi energy, which is usually fulfilled.

1.1.1.2 Magnetic susceptibility of non-interacting spins

The susceptibility of localized spin – as opposed to the previously discussed itinerant electrons – turns out to be quite different. Even if insulating, such a localized-spin material has a paramagnetic susceptibility with a clear temperature dependence.

The Hamiltonian of a single spin in a magnetic field is of course $H = -\mu \vec{S}\vec{B}$. Selecting the spin-quantization axis z to be parallel to $\vec{B}$, this becomes for a system of non-interacting localized electrons

$$H = -\frac{g\mu_B}{2}\vec{B}\sum_i \vec{\sigma}_i = -\frac{g\mu_B}{2} B \sum_i \sigma_i^z \tag{1.3}$$

where σ are the Pauli matrices and index i runs over lattice sites. The magnetization and consequently the susceptibility are likewise simply sums over the lattice sites

$$M = \frac{\sum_i \langle M_i \rangle}{N} = \frac{g\mu_B}{2N} B \sum_i \langle \sigma_i^z \rangle \quad \text{and} \tag{1.4}$$

$$\chi = \frac{g\mu_B}{2N} \sum_i \langle \sigma_i^z \rangle \, . \tag{1.5}$$

The susceptibility has thus only to be obtained for one single spin in a magnetic field. As it can only have two states, the involved sums over all states j are easily evaluated:

$$\langle \sigma_i^z \rangle = \frac{\sum_j \langle \sigma_i^z \rangle_j e^{-\beta E_j}}{\sum_j e^{-\beta E_j}} = \frac{e^{-\beta(-\frac{g\mu_B}{2}B)} - e^{-\beta(\frac{g\mu_B}{2}B)}}{e^{-\beta(-\frac{g\mu_B}{2}B)} + e^{-\beta(\frac{g\mu_B}{2}B)}} = \tanh \frac{g\mu_B B}{2k_b T} \tag{1.6}$$

and

$$M = \frac{g\mu_B}{2N} \sum_i \tanh \frac{g\mu_B B}{2k_b T} = \frac{g\mu_B}{2} \tanh \frac{g\mu_B B}{2k_b T} \approx \mu_B \tanh \frac{g\mu_B B}{2k_b T} \, . \tag{1.7}$$

This result has a qualitatively different temperature and magnetic-field dependence than that for itinerant electrons.

The previously discussed itinerant-electron susceptibility was obtained for $\mu_B B \ll E_F$. Here, the energies to compare are $\mu_B B$ vs. $k_B T$. For weak magnetic fields resp. high temperature, (1.7) becomes

$$\chi \approx \frac{\partial}{\partial B} g^2 \mu_B^2 \frac{B}{4k_b T} = \frac{g^2 \mu_B^2}{4k_b T} = \frac{C}{T} \, . \tag{1.8}$$

This is the so-called Curie-law applied to non-interacting spins. Experimentally, $\chi \propto 1/T$ at high T is a sign for the presence of localized moments. At low T, interactions between spins tend to become important.

1.1.2 Interacting moments and ordered states

After noninteracting electrons/spins, we now consider interacting systems, which can order magnetically. For itinerant electrons at weak interactions, where a perturbation theory is valid, we had discussed ferromagnetism in the 'Stoner' picture in Sec. **??** and more general types[1] of magnetic/charge order in Sec. **??**.

We had there used a mean-field decoupling $n_i n_j \to \langle n_i \rangle n_j + n_i \langle n_j \rangle - \langle n_i \rangle \langle n_j \rangle$ and we will use an analogous approximation for localized spins. Let us, however, recall, what approximation is made in a typical mean-field treatment:

$$\begin{aligned} n_i n_j &= \left(\langle n_i \rangle + (n_i - \langle n_i \rangle)\right)\left(\langle n_j \rangle + (n_j - \langle n_j \rangle)\right) = \\ &= \langle n_i \rangle \langle n_j \rangle + (n_i - \langle n_i \rangle)\langle n_j \rangle + \langle n_i \rangle (n_j - \langle n_j \rangle) + \underbrace{(n_i - \langle n_i \rangle)(n_j - \langle n_j \rangle)}_{\approx 0} \approx \\ &\approx n_i \langle n_j \rangle + n_j \langle n_i \rangle - \langle n_i \rangle \langle n_j \rangle \end{aligned} \tag{1.9}$$

[1] These are more realistic than Stoner ferromagnetism in the case of the single-band Hubbard model.

The neglected term is thus the correlation between deviations from the average: A mean-field approach considers these fluctuations to be either very small (so that their square is negligible) or independent of each other.

At strong electron-electron interactions, doubly occupied sites cost high energy so that electrons can hardly hop at half filling. If interactions are strong enough to make such a half-filled system insulating, this so-called 'Mott'-insulator has localized spins.

For localized spins, plausible interactions are of the form

$$H = \sum_{i,j} J_{i,j} \vec{S}_i \vec{S}_j \ . \tag{1.10}$$

These 'Heisenberg' interaction are 'simple' in the sense that they preserve the full rotational symmetry in spin space and do thus not assume further specifics of the system. In practice, some anisotropies are often present due to spin-orbit/lattice coupling. In reading papers and books, one has to take into account that different conventions exists concerning the notation of Heisenberg Hamiltonians:

- Sign of $J_{i,j}$: Sometimes, the Hamiltonian is written with an overall minus sign, so that ferromagnetic couplings become positive.
- Counting of pairs (i, j): Either each bond is only counted once or the sums over i and j go both over *all* sites, so that each bond is counted twice with $J_{i,j} = J_{j,i}$. Gives a factor of two in $J_{i,j}$.

Here, we count each bond twice and $J_{i,j} > 0$ denotes antiferromagnetic coupling favoring opposite spins.

1.1.2.1 Mean-field treatment

In a mean-field decoupling, we use $\vec{S}_i \vec{S}_j \to \langle \vec{S}_i \rangle \vec{S}_j + \vec{S}_i \langle \vec{S}_j \rangle - \langle \vec{S}_i \rangle \langle \vec{S}_j \rangle$ and the Hamiltonian on one specific bond becomes

$$H_{a,b} = -\mu \vec{B} \left(\vec{S}_a + \vec{S}_b \right) + J_{a,b} \langle \vec{S}_b \rangle \vec{S}_a + J_{a,b} \vec{S}_b \langle \vec{S}_a \rangle - J_{a,b} \langle \vec{S}_b \rangle \langle \vec{S}_a \rangle = \sum_{i=a,b} \left(-\mu \vec{B} + J_{i,\bar{i}} \langle \vec{S}_{\bar{i}} \rangle \right) \vec{S}_i - \text{const.}$$

where $\bar{i} = a$ (b) for $i = b$ (a). $\vec{S}_i$ is here still an operator, while $\langle \vec{S}_i \rangle$ is a number and the term containing only these numbers will not be discussed any more at present. (It is important if one wants to calculated the overall energy, though.) The effect of "the other" spin $\vec{S}_{\bar{i}}$ onto $\vec{S}_i$ has here been rewritten into an additional magnetic field.

Extending this treatment to all bonds and discarding the constant, the spin $\vec{S}_i$ sees a total effective field containing both the external field $\vec{B}$ and an internal field coming from all spins it is connected to via non-zero $J_{i,j}$:

$$H_i = -\mu \vec{B}_{i,\text{eff}} \vec{S}_i = \left(-\mu \vec{B} + \sum_j J_{i,j} \langle \vec{S}_j \rangle \right) \vec{S}_i \ . \tag{1.11}$$

Of course, $\vec{S}_i$ itself also has an effect on the other spins, and this "back-feeding" is severely affected by the mean-field approach. Assuming we want to discuss ferromagnetic order, we

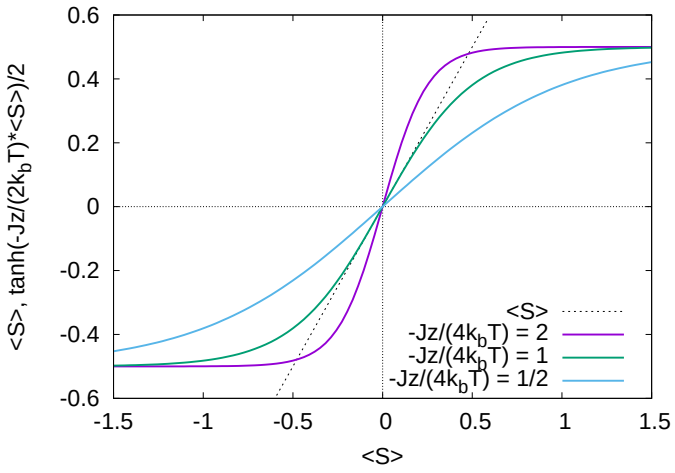


Figure 1.1: Graphical solution of Eq. (1.15) for $z = 4$ (two-dimensional square lattice), ferromagnetic $J = -1$ (setting the relevant unit of energy/temperature) and three temperatures $k_B T = \frac{z(-J)}{4} \cdot x = T_c \cdot x = 2, 1, \frac{1}{2}$ with the critical temperature T_c.

further assume $\langle \vec{S}_i \rangle \to \langle \vec{S} \rangle = \langle S \rangle \vec{e}^z$, i.e. uniform order with spins along z, the direction of $\vec{B}$. The total effective field acting on each spin is then

$$-\mu B_{i,\text{eff}} = -\mu B + \langle S \rangle \sum_j J_{i,j} \tag{1.12}$$

and for uniform nearest-neighbor coupling

$$\mu B_{\text{eff}} = \mu B - zJ\langle S \rangle \,, \tag{1.13}$$

where z is the coordination number, the number of neighbors.

The effective magnetic field acts on each spin separately (in the mean-field approximation!) and one can use (1.6). For spin 1/2, one then gets

$$g\mu_B B_{\text{eff}} = g\mu_B B - zJ\langle S \rangle \quad \text{resp.} \tag{1.14}$$

$$\langle S \rangle = \frac{1}{2}\langle \sigma^z \rangle = \frac{1}{2} \tanh \frac{g\mu_B B_{\text{eff}}}{2k_b T} = \frac{1}{2} \tanh \frac{g\mu_B B - zJ\langle S \rangle}{2k_b T} \,. \tag{1.15}$$

The average spin $\langle S \rangle$ is here the important variable (all other letters denote constants) that expresses the system's reaction.

The difference to non-interacting spins becomes clearest for $B \to 0$ [2]: In this case, Eq. (1.7) for noninteracting spins clearly implies $\langle S \rangle = \tanh 0 = 0$. For interacting spins, $\langle S \rangle = 0$ is also always a solution of Eq. (1.15), but possibly not the only one. To see whether there are other solutions, it is helpful to plot both the left side of (1.15), i.e. $\langle S \rangle$ and the right side $\frac{1}{2} \tanh \frac{-zJ\langle S \rangle}{2k_b T}$ as a function of $\langle S \rangle$ each and to check whether they cross at any points except $\langle S \rangle = 0$. The slope of $\langle S \rangle$ is one and as the hyperbolic tangent has is maximal slope

[2] The role of $\vec{B}$ has here only been to select a spin-quantization axis. This is not necessary, but a helpful trick.

at $\langle S\rangle = 0$, the curves can only have more than one crossing, if that maximal slope is larger than one, see Fig. 1.1. This criterion yields $\frac{-zJ}{4k_bT} = 1$ as the point separating the parameter space with one solution ($\langle S\rangle = 0$) from that with three. For temperatures above the 'critical' one, we have one solution, for temperatures below, three.

It turns out that the free energy of the additional solutions with $\langle S\rangle \neq 0$ is lower than that of $\langle S\rangle = 0$ for $T < T_C$, i.e., it is energetically favorable to have a finite magnetization. This breaks the original symmetry of the Hamiltonian, where positive and negative S^z are equivalent – in fact, for $\vec{B} = 0$, all directions of $\vec{S}$ are equivalent. This mechanism is called 'spontaneous symmetry breaking'.

Let us now investigate the magnetic susceptibility of interacting spins in a magnetic field. It is determined by the derivative of the magnetization per site $M = g\mu_B\langle S\rangle$ with respect to the magnetic field B, i.e.,

$$\chi = \frac{\partial M}{\partial B} = g\mu_B \frac{\partial\langle S\rangle}{\partial B} = \frac{g\mu_B}{2}\frac{\partial}{\partial B}\tanh\frac{g\mu_B B - zJ\langle S\rangle}{2k_bT} \,. \tag{1.16}$$

As in Eq. (1.7) for non-interacting spins, we approximate the hyperbolic tangent by a straight line, an approximation valid for small arguments, i.e., small external and internal magnetic fields compared to temperature. One then finds

$$\chi \approx \frac{g\mu_B}{2}\frac{\partial}{\partial B}\frac{g\mu_B B - zJ\langle S\rangle}{2k_bT} = \frac{g^2\mu_B^2}{4k_bT} + \underbrace{\frac{z(-J)}{4k_b}}_{=T_C}\frac{1}{T}\underbrace{g\mu_B\frac{\partial\langle S\rangle}{\partial B}}_{=\chi}$$
$$\chi\left(1 - \frac{T_C}{T}\right) = \frac{C}{T} \tag{1.17}$$

where the constant $C = \frac{g^2\mu_B^2}{4k_b}$ is the same as in Eq. (1.7). This expression can be rewritten to give

$$\chi = \frac{C}{T - T_C} \,, \tag{1.18}$$

Curie's law for ferromagnets. It starts to diverge at higher temperatures $T \to T_C > 0$, because even a small magnetic field can result in a large magnetization once the system is close to ordering spontaneously. (Nevertheless, the approximation of the tanh implies that the treatment is no longer valid at T_C.)

This consideration can even be extended to antiferromagnetic coupling. In that case, one finds the opposite sign in the denominator of (1.18), i.e., a "negative ordering temperature" indicates antiferromagnetic order: At low T, a magnetic field finds it harder to magnetize the sample than would be expected for independent spins, because they want to align in an alternating pattern.

1.1.2.2 Excitations and Validity of the Mean-field Treatment

The mean-field approximation gives a finite critical temperature T_C regardless of rotation symmetry of the spin, as an Ising Hamiltonian, where spins can only be $\pm 1/2$ would give the same result. The critical temperature is also $T_C > 0$ in all dimensions, even if the presence

of z makes it smaller in lower dimensions. After a look at low-energy excitations, we will however see that the picture is somewhat more complicated than mean-field theory suggests.

In the exercise, one problem showed that the elementary excitations in the 1D Heisenberg ferromagnet have vanishing energy for $\vec{k} \to 0$. With some hand waving, this can be generalized for Hamiltonians with continuous symmetry (e.g., Heisenberg spins, but also "x-y models", where spins live within a plane). Low-temperature states with their spontaneous magnetization aligned along equivalent directions must have the same energy, i.e., one has a continuous order parameter. If continuous path through equivalent directions can be chosen, then one can introduce one single long-wavelength twist through the system. Locally, this twists each spin a little bit from the purely ferromagnetic direction and thus costs a little bit of energy. However, for an infinitely large system, one can make the twist arbitrarily small, so that the state is practically ferromagnetic everywhere and the energy cost becomes arbitrarily small as well. [3] One can thus argue that ordered states of systems with continuous symmetry should have excitations with energies going to 0, so-called 'Goldstone' modes.

Exercise 3 on sheet 4 dealt with the impact of low-energy acoustic phonons on the average ion positions. The conclusion had been that the quantum fluctuations caused by the phonons actually cause the average deviation from the equilibrium position to diverge in one dimension (even at $T = 0$) and two dimensions (at $T > 0$). This argument can now be generalized: The low-energy excitations present due to a continuous order parameter make deviations from that order parameter diverge and melt the order. This happens in 1D regardless of the temperature and in 2D at finite temperature. These no-go statements about long-range order in low dimensions are referred to as "Mermin-Wagner theorem".

If the order parameter is discrete rather than continuous, e.g., in the case of the Ising model, one can not make use of the infinitely slow distortion to argue in favor of zero-energy excitations. In fact, the excitation spectrum is in these cases gapped. Long-range order can then exist at finite T in two dimensions and at $T = 0$ in one dimension. Any finite (even small) T in one dimension nevertheless destroys truly long-rang order: Even if a "mistake" in the spin pattern costs energy Δ and is suppressed by a probability $\mathrm{e}^{-\beta\Delta}$, this unlikely event will still arise *somewhere* in an infinitely long chain. While order can "go around" such isolated defects in higher demensions, defects are fata to order in one dimension.

1.2 (Ginzburg-)Landau Theory

Also inspired by Khomskii's book.

After having studied magnetic ordering, we are now going to discuss very generally approach to ordering transitions. It is also a mean-field approach, but in contrast to the mean-field treatment of the Heisenberg model discussed above, it does not start from a microscopic model, but rather from global symmetry considerations. It can thus not be used to derive macroscopic properties from microscopic ingredients, but rather to find the a theoretical description consistent with observed macroscopic facts. Its strength is that it needs very few ingredients.

The goal is to describe a system that potentially has some ordering transition.

[3] As long as interactions are either finite in range or at least decay sufficiently fast with distance.

- The order parameter η can be a scalar (e.g. magnetization of the Ising model), a vector (e.g. magnetization of ferromagnetic and many antiferromagnetic Heisenberg models) a tensor (e.g. order parameters of some antiferromagnetic patters), it can be real or complex. We start with a real scalar for simplicity.
- The order parameter can also depend on space, the corresponding approach is then called "Ginzburg-Landau" theory, while "Landau theory" might often refers to the homogeneous case. We sart with a uniform order parameter.
- At high temperature, the order parameter is assumed to be $\eta = 0$, i.e., the system is disordered.
- At temperatures just below the ordering temperature, the order parameter is assumed to be rather small (ideally, but not necessarily, going to $\eta \to 0$ at $T \to T_C$), so that expansions in its powers make sense. (Note that this is fulfilled in the case discussed above, because a slope infinitesimally below 1 yields finite-$\langle S \rangle$ solutions infinitesimally away from 0.)

The approach is then to write the free energy as an expansion in powers of the order parameter η:

$$\begin{aligned}\Phi(\eta, T, p, \dots) = \Phi_0(T, p, \dots) + \alpha(T, p, \dots)\eta + A(T, p, \dots)\eta^2 + C(T, p, \dots)\eta^3 + \\ + B(T, p, \dots)\eta^4 + F(T, p, \dots)\eta^5 + D(T, p, \dots)\eta^6 + \dots \end{aligned} \tag{1.19}$$

The parameters depend continuously on temperature T and can additionally depend on other quantities, e.g., pressure. We focus here on temperature dependence and furthermore assume that the parameters depend 'nicely' on T, i.e., are continuously differentiable. Any jumps in thermodynamic quantities are supposed to come from phase transitions captured by η. Physically, this means that we exclude here the presence of other order parameters that might spontaneously break symmetry.

Equilibrium values for η are saddle points of Φ, i.e.,

$$\begin{aligned}\left.\frac{\partial \Phi(\eta, T, p, \dots)}{\partial \eta}\right|_{T,p,\cdots=\text{const.}} = \alpha(T, p, \dots) + 2A(T, p, \dots)\eta + \\ + 3C(T, p, \dots)\eta^2 + 4B(T, p, \dots)\eta^3 + \cdots = 0 \end{aligned} \tag{1.20}$$

In order for $\eta = 0$ to be the only solution for all $T > T_C$, the linear coefficient has to vanish $\alpha(T, p, \dots) = 0$. In many cases, a system's (expected) inversion symmetry will moreover require positive and negative η to be equivalent, so that all odd powers drop out. The simplest case is then found for $B(T \approx T_C, p, \dots) > 0$, where we only have to keep two terms, the quadratic and the fourth power:

$$\Phi(\eta, T) = \Phi_0 + A(T)\eta^2 + B(T)\eta^4 \quad \text{resp.} \tag{1.21}$$

$$\eta \cdot \left(A(T) + 2B(T)\eta^2\right) = 0\,. \tag{1.22}$$

Clearly, $\eta = 0$ is always a solution. However, two more solutions

$$\eta = \pm\sqrt{-\frac{A(T)}{2B(T)}} = \pm\sqrt{\frac{|A(T)|}{2B(T)}} \tag{1.23}$$

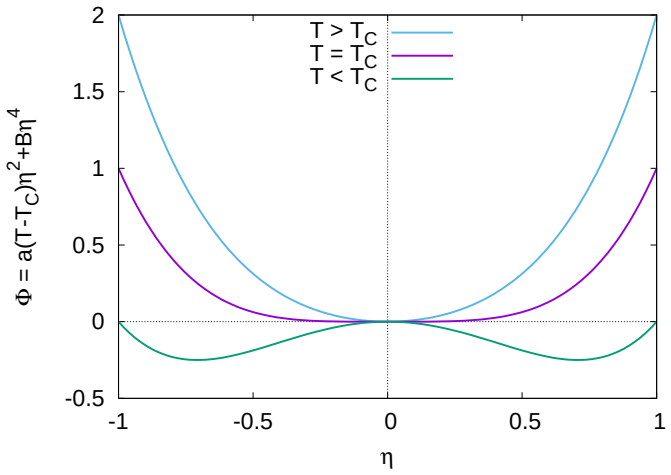


Figure 1.2: Simplest free energy (1.21) describing a second-order phase transition for three temperatures $T > T_C$, $T = T_C$ and $T < T_C$. At $T > T_C$, both the quadratic and the fourth-order term are positive, so that the only minimum is at $\eta = 0$. At $T = T_C$, the coefficient of η^2 goes through 0, so that it is negative at $T < T_C$. For large η, the fourth-order term dominates, but at small η, the quadratic term induces free-energy minima a finite η.

are found if and only if $A(T) < 0$. (Note that we required $B(T \approx T_C) > 0$.) At T_C, coefficient A must consequently change sign. One can easily check that for $A < 0$, these additional solutions give minima, while $\eta = 0$ becomes a local maximum. Focussing on the leading order in temperature, we can set $A(T \approx T_C) = a \cdot (T - T_C) + \mathcal{O}((T - T_C)^2)$ with $a > 0$.[4] For an illustration, see Fig. 1.2. The order parameter just below T_C is then given by

$$\eta = \pm\sqrt{\frac{a \cdot (T_C - T)}{2B}}\,, \tag{1.24}$$

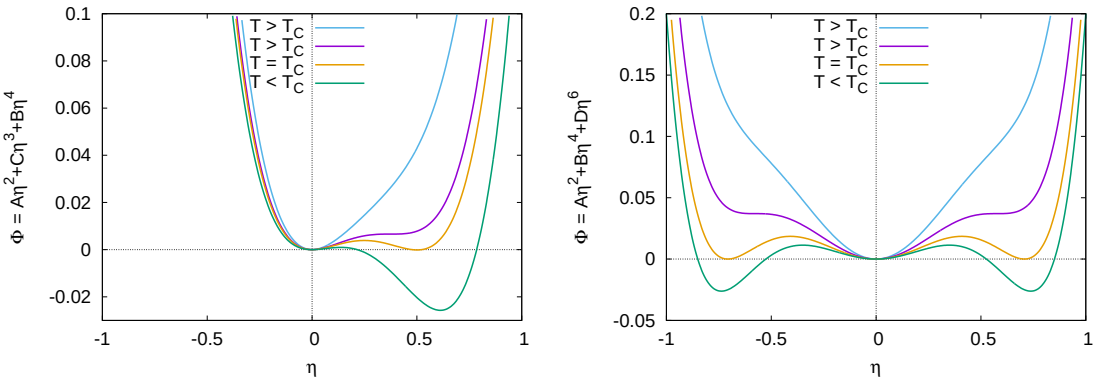
i.e., a square-root dependence. When more exact approaches are used, the 'critical exponent' turns out not to be 1/2, but pure symmetry breaking transitions in fact have an order parameter growing continuously from 0, i.e., the are second order as mean-field theory suggests and as we are going to discuss next.

1.2.1 Second-order transition

An order parameter continuous at the transition is a strong hint in favor of a higher-order phase transition, but let us now in detail discuss the analytic behavior of the optimal $\Phi_{\min}(T \approx T_C)$ and its derivatives. The equilibrium order parameter $\eta_{\min}$ around T_C is close to zero and lowest order gives:

$$\Phi_{\min} = \Phi_0 + A\eta_{\min}^2 + \dots = \Phi_0 \begin{cases} +\,0 & \text{for } \eta_{\min}^2 = 0 \text{ at } T > T_C, \\ -\frac{a^2}{2B}(T - T_C)^2 + \dots & \text{for } \eta_{\min}^2 = \frac{a\cdot(T-T_C)}{2B} \text{ at } T \le T_C, \end{cases} . \tag{1.25}$$

[4] $a = 0$ and $A \propto (T - T_C)^3$ is theoretically possible, but extremely unlikely, as no symmetry enforces this and it would thus need fine tuning.



(a) (b)

Figure 1.3: Two cases of free energies (1.19) describing a first-order phase transition. In (a), the system lacks inversion symmetry, so that a third-order term in η exists. In (b), the system is inversion symmetric, i.e., no odd powers in η are present. The fourth-order coefficient $B(T \approx T_C)$ is here negative. The sixth order becomes then important and the transitions becomes first order.

This functions is continuous at T_C.

The first derivative with respect to T, keeping all other parameters like p or V constant, is the entropy

$$S = -\frac{\partial \Phi}{\partial T} = \underbrace{-\frac{\partial \Phi_0}{\partial T}}_{=S_0} + \begin{cases} 0 & \text{for } T > T_C, \\ \frac{a^2}{B}(T - T_C) + \dots & \text{for at } T \le T_C \end{cases}, \tag{1.26}$$

which is also continuous at T_C, i.e., there is no *first*-order phase transition.

The next derivative gives the specific heat, i.e.,

$$c_p = T\frac{\partial S}{\partial T} = T\frac{\partial S_0}{\partial T} + \begin{cases} 0 & \text{for } T > T_C, \\ T \cdot \frac{a^2}{B} + \dots & \text{for at } T \le T_C \end{cases}, \tag{1.27}$$

which finally has a jump at T_C, i.e., there is a *second-order* phase transition.

1.2.2 Weakly first-order transitions

If the order parameter jumps to a large value at T_C, an expansion like Eq. (1.19) is not valid. But if the jump is small, the "weakly first-order" transition may still tractable.

For a first-order transition, the order parameter η should jump from 0 to a finite value at T_C. There are two basic ways to achieve this. One arises in systems without inversion symmetry. In that case, odd powers of η are allowed in the free energy, only the linear term has to vanish to allow a disordered high-T phase. For small η, adding the third power to the scenario of Fig. 1.2 leads to an asymmetry at high T that develops into a local minimum, which in turn finally dives below 0 – it is then separated from the $\eta = 0$ minimum by a maximum, see Fig. 1.3(a), and the optimal η jumps.

A second possibility can also arise in systems with inversion symmetry, i.e., with only even powers of η. If the coefficient B of the fourth-order term is negative around T_C, higher-orders become important. As $B < 0$ implies that $\Phi(\eta)$ becomes more and more negative with growing $|\eta|$, the equilibrium $|\eta|$ is large enough for the sixth order to matter. If its coefficient D positive, it will eventually penalize large η and it is enough to consider the quadratic, fourth- and sixth-order terms. For $B < 0$ and $D > 0$, varying A then leads to the scenario depicted in Fig. 1.3(b): local minima at finite *eta* arise and finally sink below the $\eta = 0$ minimum. Again, the order parameter jumps at T_C.

1.2.3 Inhomogeneous States and gradients in the order parameter

The Landau approach can be extended to include variations in the order parameter, this is called Ginzburg-Landau theory. An example might be a magnetic spiral state that is almost ferromagnetic, but with a slowly varying preferred direction. Another example is a superconductor with a boundary or with magnetic flux quanta going through it: Superconductivity has to vanish at the boundary or in the magnetic flux, and then continuously reaches its bulk value. The order parameter then becomes a function of position $\eta = \eta(\vec{r})$ and the free energy an integral over the free-energy density.

As in the uniform case, symmetry considerations suggest the terms that should be included in the free-energy integral. If variable order parameters are to be allowed, the free energy might in lowest order depend on their gradients. The free energy of inversion symmetric systems must not distinguish a direction $\vec{r}$ from $-\vec{r}$, so that terms linear in the gradient are forbidden. Squares of the gradient are then the lowest order in which variability can enter. A plausible form of the first few orders in the free energy is then

$$\Phi = \int \mathrm{d}^2 r \, \left(A\eta^2(\vec{r}) + B\eta^4(\vec{r}) + G(\nabla\eta(\vec{r}))^2 + E(\nabla^2\eta(\vec{r}))^2 \right) . \tag{1.28}$$

A positive $G > 0$ suppresses gradients in η, a homogeneous solution is then preferable and one would go back to using the simpler approach discussed above. At negative G, gradients can become favorable, but positive E would prevent them from becoming too large.

One application of the Ginzburg-Landau approach are magnetic spirals, in that case, one assumes that the local magnetization $\vec{M}$ slowly rotates with constant absolute value, e.g., $M_x = \cos \vec{q}\vec{r}$ and $M_y = \sin \vec{q}\vec{r}$. It is the favorable to go into momentum space, where the gradients become multiplications with $\vec{q}$. One then finds $-|G|\vec{q}^2|M|^2 + E\vec{q}^4|N|^2 = 0$, which yields the optimal $\vec{q}$ characterizing the spiral. (A and B determine the optimal $|M|^2 = \eta^2$.)

Before we discuss Ginzburg-Landau theory for superconductors, we are going to look at the microscopic ingredients of superconductivity. Historically, the phenomenological Ginzburg-Landau approach came actually came first.

1.3 Superconductivity

Inspired by Prof. Muramatsu's and Prof. Timm's lecture notes.

Presumably the most interesting feature of superconductivity is the vanishing resistance. (Even though magnetic levitation is not too bad either.) One thus needs some electronic state that is prevented from scattering. An important experimental clue was the isotope effect, where the superconducting T_C depends on the number of neutrons in the ions. The

only way for ionic mass to influence electron dynamic would be via its impact on phonon frequency. This dependence of electronic properties on phonons indicates that this is a case where the adiabatic approximation breaks down and coupling of electrons to lattice *dynamics*, i.e. phonons, has to be taken into account.

In Sec. **??**, we had discussed the adiabatic approximation that decouples electron and lattice dynamics. On p. **??**, we had found that the energy scale of the neglected electron-phonon interaction is smaller than the other scales, but not very much smaller. We had also found that the most 'dangerous' interaction terms involved different electronic eigenstates, so that the adiabatic approximation is best justified if the electronic states are well separated. This is clearly not fulfilled in a metal, where the existence of a Fermi surface implies excited states at arbitrarily low energies. Nevertheless, it is a (perhaps surprisingly) good starting point, to which we now add again as a perturbation the interaction between electrons and lattice degrees of freedom, i.e., phonons.

1.3.1 Phonon-mediated Electron-Electron Interaction

We first need a second-quantization description of electron-phonon interaction. The coupling between ions and electrons is via Coulomb attraction, i.e., a density-density interaction. The ionic density is given by a sum over ions j that can be displaced by $\vec{u}_j$ from their equilibrium positions $\vec{R}_j$. For simplicity, we assume all ions to be equivalent and a one-site unit cell, i.e., a purely elemental solid with a Bravais lattice. The potential V between an electron at $\vec{r}$ and an ion at $\vec{R}_j + \vec{u}_j$ depends on the distance $\vec{r} - \vec{R}_J - \vec{u}_j$, leading to

$$H_{\text{el.-ion}} = \int \mathrm{d}^3 r \; n_{\text{el.}}(\vec{r}) \cdot \sum_j V(\vec{r} - \vec{R}_J - \vec{u}_j) \tag{1.29}$$

where $n_{\text{el.}}(\vec{r})$ is the electron density. Expressing it in terms of Bloch electrons via field operators (**??**) gives

$$n_{\text{el.}}(\vec{r}) = \frac{1}{V} \sum_{\vec{k},\vec{k}',\sigma} \mathrm{e}^{i(\vec{k}-\vec{k}')\vec{r}} c^\dagger_{\vec{k}',\sigma} c_{\vec{k}',\sigma} \; . \tag{1.30}$$

As $\vec{u}_j$ will be small compared to $\vec{R}_j$, a Taylor-series expansion of V is justified:

$$V(\vec{r} - \vec{R}_j - \vec{u}_j) \approx V(\vec{r} - \vec{R}_j) - \vec{\nabla} V(\vec{r} - \vec{R}_j) \cdot \vec{u}_j \; . \tag{1.31}$$

From the first term (zeroth order in $\vec{u}$), we get the ionic potential determining the electronic bands. The second term gives a coupling between electrons and ion displacement $\vec{u}_j$, whose Fourier transform is in turn given by phonon operators, see Eq. (**??**). The vector character of $\vec{u}$ is taken into account by the three phonon branches ν. Combining all the Fourier components gives

$$H_{\text{el.-phonon}} = \sum_{\vec{G}} \sum_{\vec{k},\sigma} \sum_{\vec{q},\nu} g_{\vec{G},\vec{q},\nu} c^\dagger_{\vec{k}+\vec{q}+\vec{G},\sigma} c_{\vec{k},\sigma} \left(b_{\vec{q},\nu} + b^\dagger_{-\vec{q},\nu} \right) \; . \tag{1.32}$$

The sum over reciprocal-lattice vectors $\vec{G}$ is usually restricted to $\vec{G} = 0$, i.e., "Umklapp" processes are neglected. Moreover, the solid will now be assumed to be isotropic, so that all three branches give the same. This finally yields

$$H_{\text{el.-phonon}} = \sum_{\vec{k},\sigma} \sum_{\vec{q},\nu} g_{\vec{q}} c^\dagger_{\vec{k}+\vec{q},\sigma} c_{\vec{k},\sigma} \left(b_{\vec{q},\nu} + b^\dagger_{-\vec{q},\nu} \right) \; , \tag{1.33}$$

where momentum is exchanged between the electron and phonon sectors, i.e., their dynamics are coupled.

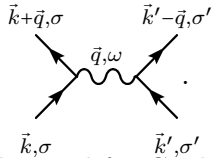
Even though the phonons are important in superconductivity, we are not really interested in them, but only in their effect on the electrons. The next step is consequently to "integrate the phonons out" and obtain the effective phonon-mediated electron-electron interaction. In this approach, we assume that mostly phonons with some finit frequency $\omega \approx \omega_D$ are important. This sounds perhaps weird, as the phonons in elemental Bravais lattices have $\omega \propto |\vec{q}| \to 0$ in the zone center, but the density of states is much larger at the zone boundary, where their dispersion is flat.

Without electron-phonon interaction, the ground state is given by a tensor product of the electronic ground state, the filled Fermi sea, and the vacuum for phonons, i.e., no phonons exist. We now perform perturbation theory with the electron-phonon interaction, where we are only interested in terms that remain in the low-energy Hilbert space of no phonons. The electrons, in contrast, are not required to remain in their ground state. In fact, we derive the effective interaction at first without any reference to the Fermi sea and consider its result as an additional term in the electronic Hamiltonian. A cleaner and more complete derivation can be carried out using diagrammatic techniques, but here, only a simplified version will be presented.

In a simple approach, we assume that the changes in electron energy caused by (1.33) are small compared to ω_D, i.e., $\epsilon_{\vec{k}+\vec{q}} - \epsilon_{\vec{k}} \ll \omega_D$. Applying the electron-phonon coupling then leads to a "high-energy" state and perturbation theory can be used. In first order, a phonon is annihilated (not possible, because none exist) or created (possible, but goes out of the interesting Hilbert space), so that it does not contribute. In second order, a contribution can arise if the second application of (1.33) annihilates the phonon created by the first; spins have to be conserved. Neglecting $\epsilon_{\vec{k}+\vec{q}} - \epsilon_{\vec{k}}$ w.r.t ω_D and using $g_{-\vec{q}} = g_{\vec{q}}^*$, one finds

$$\begin{aligned} H^{(2)}_{\text{el.-el.}} &= -\sum_{\vec{k},\sigma}\sum_{\vec{q},\nu}\sum_{\vec{k}',\sigma'}\sum_{\vec{q}',\nu'} \frac{g_{\vec{q}} g_{\vec{q}'}}{\hbar\omega_{\vec{q}'}} c^\dagger_{\vec{k}+\vec{q},\sigma} c_{\vec{k},\sigma} b_{\vec{q},\nu} c^\dagger_{\vec{k}'+\vec{q}',\sigma'} c_{\vec{k}',\sigma'} b^\dagger_{-\vec{q}',\nu'} \delta_{\nu,\nu'}\delta_{\vec{q},-\vec{q}'} \\ &= \sum_{\vec{k},\sigma}\sum_{\vec{q},\nu}\sum_{\vec{k}',\sigma'} \frac{-|g_{\vec{q}}|^2}{\hbar\omega_{\vec{q}}} c^\dagger_{\vec{k}+\vec{q},\sigma} c_{\vec{k},\sigma} c^\dagger_{\vec{k}'-\vec{q},\sigma'} c_{\vec{k}',\sigma'} \underbrace{b_{\vec{q},\nu} b^\dagger_{\vec{q},\nu}}_{=1+n_{\vec{q},\nu}=1} \end{aligned} \tag{1.34}$$

The system can thus gain energy if two electrons are moved from $\vec{k}$ and $\vec{k}'$ to $\vec{k}+\vec{q}$ and $\vec{k}'-\vec{q}$, but this result is only valid if the electronic excitation energies are smaller than ω_D, implying that all involved electronic states have to be close the the Fermi level. The phonons thus mediate an effective attraction between low-energy electrons. In a diagrammatic approach, the corresponding picture is



(1.35)

i.e., extremely similar to that (**??**) obtained for Coulomb interaction.

A more careful analysis can moreover give:

- The dependence of the interaction on the energy transfer between the electrons. In contrast to Coulomb repulsion, the phonon-mediated interaction is frequency depen-

dent [see the ω over the wiggly line in (1.35)] with

$$\propto \frac{\omega_D}{\omega^2 - \omega_D^2} \,, \tag{1.36}$$

if we replace the momentum dependence of $\omega_{\vec{q}}$ by a constant.

This can be motivated if we use (without showing it) that the total diagram (1.35) should conserve energy. One electron should in the end have gained the energy ω lost by the other, i.e., $\epsilon_{\vec{k}+\vec{q}} - \epsilon_{\vec{k}} = -(\epsilon_{\vec{k}'-\vec{q}} - \epsilon_{\vec{k}'}) = \omega$. We can then include the electronic energy in (1.34) combining terms with momentum transfers $\vec{q}$ and $-\vec{q}$ turns out to yield $\frac{\omega_{\vec{q}}}{\omega^2-\omega_{\vec{q}}^2}$. It has to be admitted, though, that this argument is (while popular) not fully watertight: the difference between $\frac{1}{\omega_D}$ and $\frac{\omega_D}{\omega^2-\omega_D^2}$ is of third order, i.e., we would have to include the next order in perturbation theory to be consistent. That said, a more advanced treatment based on Green's functions supports the result.

The dependence on the electronic energy transfer ω in addition to phonon energy $\omega_{\vec{q}}$ indicates that the phonon-mediated interaction is 'dynamic'. This in turn means that the interaction is "retarded", i.e., the second electron feels the effect of the first only a little time later. This can be understood by noting the physical mechanism: The first electron slightly distorts the lattice (i.e., creates phonons), flies away as it is much faster than the ions, and the second electron still meets the phonons when it comes by. Due to this retardation, the electrons are not actually close to each other and do not feel too strong a Coulomb repulsion.

- The competition with Coulomb repulsion, at least for jellium. This is important, because the two effects are of similar size and obviously compete. It turns out that some attraction survives for small energy transfer.
- Renormalization due to the fact that the electron system is an interacting soup with Coulomb repulsion and phonons, e.g., this also renormalizes the "one-particle" energies. This effect smoothes the energy dependence of the interaction, so that is is no longer singular at $\omega = \omega_D$ (where the use of second-order perturbation theory is after all not justified), but again leaves low-energy attraction intact.

We thus end up with an effective electron-electron attraction that can act on electrons that are very close to the Fermi level.

1.3.2 BCS Theory of Superconductivity

Nicely discussed in Altland and Simons

BCS theory presents a variational ansatz for a somewhat further simplified Hamiltonian that can also be extended to electron-electron interactions mediated by processes different from isotropic phonons. The Hamiltonian is supposed to treat the balance between the non-interacting band energy ("kinetic energy"), whose ground state would be the Fermi sea and which consequently suppresses deviation from this state, and an effective electron-electron attraction between electrons that are close to each other in energy. As the final state cannot be expected to be "very far" from the Fermi sea, one can in fact formally extend the attraction onto all single-electron states, because the kinetic energy easily dominates for states far from the Fermi level. The dynamics of the phonon-mediated electron-electron

interaction is thus neglected here, even though it is very important in helping the attraction survive (for low-energy electrons) competition with Coulomb repulsion.

Second, we simplify the momentum dependence of Eq. (1.34), resp. diagram (1.35) by focussing on precesses where the two annihilators (creators) have opposite momentum. The rationale being this step is that the number of such pairs with total momentum 0 is largest, so that these processes tend to dominate anyway. We are going to see that this gives us Cooper pairs with vanishing total momentum. They can have finite momentum in so-called Fulde-Ferrell-Larkin-Ovchinnikov states, which arise in magnetic fields.

With these simplifications, we arrive at

$$H = \sum_{\vec{k},\sigma} (\epsilon_{\vec{k}} - \mu) c^\dagger_{\vec{k},\sigma} c_{\vec{k},\sigma} + \frac{1}{2} \sum_{\substack{\vec{k},\vec{k}' \\ \sigma,\sigma'}} V_{\vec{k},\vec{k}'} c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} \,. \tag{1.37}$$

If the interaction is not mediated by phonons but via some other mechanism, $V_{\vec{k},\vec{k}'}$ can also depend on σ and σ'. Even though phonons have been integrated out, this is still an interacting Hamiltonian and thus hard to solve. Treating this in perturbation theory, e.g. via Feynman diagrams, leads to one of the instances where the series do not converge to the ground state. The ground state has symmetries fundamentally different from a Fermi liquid.

Instead of perturbation theory, we try a mean-field approach, which can incorporate symmetry breaking. Decouplings of the type we have used before would be, e.g., $\langle c^\dagger_{\vec{k},\sigma} c_{\vec{k}',\sigma} \rangle c^\dagger_{-\vec{k},\sigma'} c_{-\vec{k}',\sigma'}$. This would lead to a (spin-)density wave and in a realistic application with a more general Hamiltonian, one should certainly checke whether this type of order is stable. Here, however, a spin-density wave is not the ground state, Instead, we focus on a different mean-field decoupling, namely

$$c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} \to \langle c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} \rangle c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} + c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} \langle c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} \rangle - \langle c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} \rangle \langle c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} \rangle \,. \tag{1.38}$$

Again, we neglect terms quadratic in fluctuations, as in Eq. (1.9).

The mean-field Hamiltonian then becomes

$$\begin{aligned} H_{\mathrm{MF}} = &\sum_{\vec{k},\sigma} (\epsilon_{\vec{k}} - \mu) c^\dagger_{\vec{k},\sigma} c_{\vec{k},\sigma} + \frac{1}{2} \sum_{\sigma,\sigma'} \sum_{\vec{k}'} \underbrace{\sum_{\vec{k}} V_{\vec{k},\vec{k}'} \langle c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} \rangle}_{= \Delta^*_{\sigma,\sigma'}(\vec{k}')} c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} \\ &+ \frac{1}{2} \sum_{\sigma,\sigma'} \sum_{\vec{k}} c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} \underbrace{\sum_{\vec{k}'} V_{\vec{k},\vec{k}'} \langle c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} \rangle}_{= \Delta_{\sigma,\sigma'}(\vec{k})} - \frac{1}{2} \sum_{\sigma,\sigma'} \sum_{\vec{k},\vec{k}'} V_{\vec{k},\vec{k}'} \langle c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} \rangle \langle c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} \rangle \,, \end{aligned} \tag{1.39}$$

where the last term is a constant. The spin symmetry of $\Delta_{\sigma,\sigma'}$ encodes whether the two electrons forming a pair are in a singlet or triplet state. For phonon-mediated interaction, the singlet channel is relevant.

Looking at the simplest case of $V(\vec{k},\vec{k}') = V_0$ and *singlet* pairing Eq. (1.39) leads to

$$H_{\mathrm{MF}} = \sum_{\vec{k},\sigma} (\epsilon_{\vec{k}} - \mu) c^\dagger_{\vec{k},\sigma} c_{\vec{k},\sigma} + \frac{1}{2} \sum_{\vec{k}} \Delta^* c_{-\vec{k},\uparrow} c_{\vec{k},\downarrow} - \frac{1}{2} \sum_{\vec{k}} \Delta^* c_{-\vec{k},\downarrow} c_{\vec{k},\uparrow} \tag{1.40}$$

$$-\frac{1}{2}\sum_{\vec{k}}\Delta c^\dagger_{\vec{k},\uparrow}c^\dagger_{-\vec{k},\downarrow}+\frac{1}{2}\sum_{\vec{k}}\Delta c^\dagger_{\vec{k},\downarrow}c^\dagger_{-\vec{k},\uparrow}-\frac{1}{2V_0}\sum_{\sigma\neq\sigma'}|\Delta|^2= \tag{1.41}$$

$$=\sum_{\vec{k},\sigma}(\epsilon_{\vec{k}}-\mu)c^\dagger_{\vec{k},\sigma}c_{\vec{k},\sigma}-\Delta^*\sum_{\vec{k}}c_{-\vec{k},\downarrow}c_{\vec{k},\uparrow}-\Delta\sum_{\vec{k}}c^\dagger_{\vec{k},\uparrow}c^\dagger_{-\vec{k},\downarrow}-\frac{|\Delta|^2}{V_0}\,. \tag{1.42}$$

This mean-field Hamiltonian is biquadratic, i.e. each term contains exactly two fermion operators, but the terms with two creators or two annihilators mean that we need an extra step before solving it.

A useful trick is a particle-hole transformation applied to the down-electron operators, [5] i.e.

$$h^\dagger_{\vec{k},\downarrow}=c_{-\vec{k},\downarrow},\quad h_{\vec{k},\downarrow}=c^\dagger_{-\vec{k},\downarrow}\,. \tag{1.43}$$

The mean-field Hamiltonian can then be written as

$$H_{\mathrm{MF}}=\sum_{\vec{k}}(c^\dagger_{\vec{k},\uparrow},h^\dagger_{\vec{k},\downarrow})\begin{pmatrix}\epsilon_{\vec{k}}-\mu & -\Delta^*\\ -\Delta & -(\epsilon_{-\vec{k}}-\mu)\end{pmatrix}\begin{pmatrix}c_{\vec{k},\uparrow}\\ h_{\vec{k},\downarrow}\end{pmatrix}-\frac{|\Delta|^2}{V_0} \tag{1.44}$$

with the so-called Nambu spinor

$$\begin{pmatrix}c_{\vec{k},\uparrow}\\ h_{\vec{k},\downarrow}\end{pmatrix}=\begin{pmatrix}c_{\vec{k},\uparrow}\\ c^\dagger_{-\vec{k},\downarrow}\end{pmatrix}\,. \tag{1.45}$$

This is then solved by diagonalizing the $\vec{k}$-dependent 2×2 matrices. Eigenenergies are clearly

$$E_{\vec{k}}=\pm\sqrt{(\epsilon_{\vec{k}}-\mu)(\epsilon_{-\vec{k}}-\mu)+|\Delta|^2}=E_{-\vec{k}}=\pm\sqrt{(\epsilon_{\vec{k}}-\mu)^2+|\Delta|^2}\,, \tag{1.46}$$

where the last equation applies to the very common inversion-symmetric case with $\epsilon_{-\vec{k}}=\epsilon_{\vec{k}}$.

Eigenstates are linear superpositions of up-electron creator and down-electron annihilator. It is then a matter of convention whether one sees this combined object as a creation or as an annihilation operator, a convenient way is the following designation:

$$\begin{pmatrix}\gamma_{\vec{k},\uparrow}\\ \gamma^\dagger_{-\vec{k},\downarrow}\end{pmatrix}=U\begin{pmatrix}c_{\vec{k},\uparrow}\\ c^\dagger_{-\vec{k},\downarrow}\end{pmatrix}=\begin{pmatrix}u^*_{\vec{k}} & v_{\vec{k}}\\ -v^*_{\vec{k}} & u_{\vec{k}}\end{pmatrix}\begin{pmatrix}c_{\vec{k},\uparrow}\\ c^\dagger_{-\vec{k},\downarrow}\end{pmatrix}\quad\text{with}\quad\begin{pmatrix}E_{\vec{k}} & 0\\ 0 & -E_{\vec{k}}\end{pmatrix}=U^\dagger H_{\vec{k}}U \tag{1.47}$$

For real Δ, $u_{\vec{k}}=\cos\theta_{\vec{k}}$ and $u_{\vec{k}}=\sin\theta_{\vec{k}}$ is a good parameterization, with $\tan 2\theta_{\vec{k}}=\frac{-\Delta}{\epsilon_{\vec{k}}-\mu}$. For complex $\Delta_{\vec{k}}=|\Delta_{\vec{k}}|e^{i\phi_{\vec{k}}}$, the phase can be gauged away when doing the particle-hole transformation (1.43). Expressed in γ operators, [6] the Hamiltonian is then

$$H=\sum_{\vec{k}}\left(E_{\vec{k}}\gamma^\dagger_{\vec{k},\uparrow}\gamma_{\vec{k},\uparrow}-E_{\vec{k}}\gamma_{-\vec{k},\downarrow}\gamma^\dagger_{-\vec{k},\downarrow}\right)-\frac{|\Delta|^2}{V_0}=\sum_{\vec{k}}E_{\vec{k}}\left(\gamma^\dagger_{\vec{k},\uparrow}\gamma_{\vec{k},\uparrow}+\gamma^\dagger_{\vec{k},\downarrow}\gamma_{\vec{k},\downarrow}\right)-\sum_{\vec{k}}E_{\vec{k}}+\frac{|\Delta|^2}{|V_0|}$$

[5] One can check that the full collection of operators building the Hamiltonian continue to fulfil fermion anti-commutation rules.

[6] One should check that the number of γ operators is correct, i.e., that there is no double counting of degrees if freedom, and that they fulfill fermionic anti-commutation relations. It is and they do.

$$= \sum_{\vec{k},\sigma} E_{\vec{k}} n_{\vec{k},\sigma} + E_0 \ , \tag{1.48}$$

where the constant E_0 gives the ground-state energy. Since $E_{\vec{k}} \geq 0$, any $\gamma^\dagger$ operator creates an excitation, which is called a Bogoliubov quasi-particle, and raises the energy. The BCS ground state is accordingly the vacuum state of the Bogoliubov operators, i.e. $\gamma_{\vec{k},\sigma}|\psi_{\mathrm{BCS}}\rangle = 0$ for all $\gamma_{\vec{k},\sigma}$.

Using

$$\begin{aligned}\gamma_{-\vec{k},\downarrow}\gamma_{\vec{k},\uparrow}|0\rangle &= (u^*_{\vec{k}} c_{-\vec{k},\downarrow} - v_{\vec{k}} c^\dagger_{\vec{k},\uparrow})(u^*_{\vec{k}} c_{\vec{k},\uparrow} + v_{\vec{k}} c^\dagger_{-\vec{k},\downarrow})|0\rangle \\ &= (u^*_{\vec{k}} c_{-\vec{k},\downarrow} - v_{\vec{k}} c^\dagger_{\vec{k},\uparrow}) u^*_{\vec{k}} \underbrace{c_{\vec{k},\uparrow}|0\rangle}_{=0} + v_{\vec{k}} u^*_{\vec{k}} \underbrace{c_{-\vec{k},\downarrow} c^\dagger_{-\vec{k},\downarrow}|0\rangle}_{=1\cdot|0\rangle} - v^2_{\vec{k}} c^\dagger_{\vec{k},\uparrow} c^\dagger_{-\vec{k},\downarrow}|0\rangle = v_{\vec{k}} \left(u^*_{\vec{k}} + v_{\vec{k}} c^\dagger_{-\vec{k},\downarrow} c^\dagger_{\vec{k},\uparrow}\right)|0\rangle\end{aligned}$$

such a state is given by

$$|\psi_{\mathrm{BCS}}\rangle \propto \prod_{\vec{k},\sigma} \gamma_{\vec{k},\sigma}|0\rangle \propto \prod_{\vec{k}} \left(u^*_{\vec{k}} + v_{\vec{k}} c^\dagger_{-\vec{k},\downarrow} c^\dagger_{\vec{k},\uparrow}\right)|0\rangle \ . \tag{1.49}$$

The last expression is also normalized. This is a superposition of states with different numbers of Cooper pairs, even having a component with no Cooper pair.

This is rather similar to the coherent states of harmonic oscillators and the BCS ground state can indeed be seen as a fermionic counterpart. Except for the normalization, a coherent state for a single boson mode is given by

$$|z\rangle = \mathrm{e}^{z a^\dagger}|0\rangle = \left(1 + z a^\dagger + \frac{1}{2} z^2 (a^\dagger)^2 + \frac{1}{6} z^3 (a^\dagger)^3 + \dots\right)|0\rangle \tag{1.50}$$

which can by generalized to several modes via a product. The same kind of state with a creation operator of a Cooper pair instead of a boson gives a much shorter sum, because the product of two or more fermion creators vanishes:

$$\begin{aligned}u^*_{\vec{k}} \mathrm{e}^{\frac{v_{\vec{k}}}{u^*_{\vec{k}}} c^\dagger_{-\vec{k},\downarrow} c^\dagger_{\vec{k},\uparrow}}|0\rangle &= u^*_{\vec{k}} \Big(1 + \frac{v_{\vec{k}}}{u^*_{\vec{k}}} c^\dagger_{-\vec{k},\downarrow} c^\dagger_{\vec{k},\uparrow} + \frac{1}{2}\Big(\frac{v_{\vec{k}}}{u^*_{\vec{k}}}\Big)^2 \underbrace{c^\dagger_{-\vec{k},\downarrow} c^\dagger_{\vec{k},\uparrow} c^\dagger_{-\vec{k},\downarrow} c^\dagger_{\vec{k},\uparrow}}_{=0} + 0\Big)|0\rangle = \\ &= \left(u^*_{\vec{k}} + v_{\vec{k}} c^\dagger_{-\vec{k},\downarrow} c^\dagger_{\vec{k},\uparrow}\right)|0\rangle \ .\end{aligned} \tag{1.51}$$

For bosonic fields like light, coherent sates do not have a definite particle number (even though the relative variance becomes very small for large average numbers), but they do have a well-defined phase, which is the conjugate variable to photon number. The eigenstates of the Hamiltonian, in contrast, have a precise particle number, but no well-defined phase. It turns out that such an observation carries over to the BCS ground state: It has a well defined phase and "macroscopic phase coherence".

1.3.2.1 BCS Gap equation

From Eq. (1.46), we can conclude that creating a Bogoliubov quasiparticle costs at least energy Δ, i.e., there is a gap in the excitation spectrum. This gap stabilizes on one hand the

symmetry-broken state, but is on the other determined by $\Delta = V_0\langle c^\dagger_{\vec{k},\uparrow} c^\dagger_{-\vec{k},\downarrow}\rangle$, i.e., depends on the ground state. One thus finds (here shown for real Δ) a self-consistency requirement

$$\begin{aligned}\Delta &= V_0 \sum_{\vec{k}} \langle\psi_{\mathrm{BCS}}| c^\dagger_{\vec{k},\downarrow} c^\dagger_{-\vec{k},\uparrow} |\psi_{\mathrm{BCS}}\rangle = V_0 \sum_{\vec{k}} \langle\psi_{\mathrm{BCS}}| (v_{\vec{k}}\gamma_{-\vec{k},\uparrow} + u_{\vec{k}}\gamma^\dagger_{\vec{k},\downarrow})(u_{\vec{k}}\gamma^\dagger_{-\vec{k},\uparrow} - v_{\vec{k}}\gamma_{\vec{k},\downarrow})|\psi_{\mathrm{BCS}}\rangle = \\ &= V_0 \sum_{\vec{k}} (v_{\vec{k}} u_{\vec{k}} \underbrace{\langle\psi_{\mathrm{BCS}}|\gamma_{-\vec{k},\uparrow}\gamma^\dagger_{-\vec{k},\uparrow}|\psi_{\mathrm{BCS}}\rangle}_{=1} - u_{\vec{k}} v_{\vec{k}} \underbrace{\langle\psi_{\mathrm{BCS}}|\gamma^\dagger_{\vec{k},\downarrow}\gamma_{\vec{k},\downarrow}|\psi_{\mathrm{BCS}}\rangle}_{=0}) = \\ &= V_0 \sum_{\vec{k}} v_{\vec{k}} u_{\vec{k}} = V_0 \sum_{\vec{k}} \sin\theta_{\vec{k}} \cos\theta_{\vec{k}} = \frac{V_0}{2} \sum_{\vec{k}} \sin 2\theta_{\vec{k}} = \frac{V_0}{2} \sum_{\vec{k}} \frac{-\Delta}{E_{\vec{k}}} \,, \end{aligned} \tag{1.52}$$

where the last steps require some trig identities. Obviously, $\Delta = 0$ is always a solution. For *negative* $V_0 < 0$ *only*, other solutions $\frac{2}{|V_0|} = \sum_{\vec{k}} \frac{1}{E_{\vec{k}}}$ are possible that have a finite gap.

We have here of course lost the isotope effect, because we extended the attractive interaction from states near the Fermi energy to arbitrary states. If a cut-off at $|\epsilon_{\vec{k}} - \mu| = \hbar\omega_D$ is included, one finds

$$\Delta \propto 2\omega_D \mathrm{e}^{-\frac{1}{V_0 \rho(E_F)}} \tag{1.53}$$

for an approximately constant density of states $\rho(E_F)$ around the Fermi energy.

1.3.2.2 'Unconventional superconductivity' with momentum-dependent gap

One can get superconducting solutions from positive interactions $V_{\vec{k},\vec{k}'} > 0$ as well, but only if the interaction is momentum dependent. Eq. (1.39) is replaced by its variant

$$\begin{aligned} H_{\mathrm{MF}} = &\sum_{\vec{k},\sigma} (\epsilon_{\vec{k}} - \mu) c^\dagger_{\vec{k},\sigma} c_{\vec{k},\sigma} + \frac{1}{2} \sum_{\sigma,\sigma'} \sum_{\vec{k}'} \underbrace{\sum_{\vec{k}} V_{\vec{k},\vec{k}'} \langle c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'}\rangle}_{=\Delta^*_{\sigma,\sigma'}(\vec{k}')} c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma} \\ &+ \frac{1}{2} \sum_{\sigma,\sigma'} \sum_{\vec{k}} c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'} \underbrace{\sum_{\vec{k}'} V_{\vec{k},\vec{k}'} \langle c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma}\rangle}_{=\Delta_{\sigma,\sigma'}(\vec{k})} \underbrace{- \frac{1}{2} \sum_{\sigma,\sigma'} \sum_{\vec{k}} V_{\vec{k},\vec{k}'} \langle c^\dagger_{\vec{k},\sigma} c^\dagger_{-\vec{k},\sigma'}\rangle \sum_{\vec{k}'} \langle c_{-\vec{k}',\sigma'} c_{\vec{k}',\sigma}\rangle}_{=\text{const.}} \,. \end{aligned} \tag{1.54}$$

The off-diagonal element of the resulting equivalent to Eq. (1.40) is now $\vec{k}$-dependent. All formulas carry over, with a $\vec{k}$-dependent gap $\Delta_{\vec{k}} = f(\vec{k})|\tilde{\Delta}|$. The gap equation becomes

$$\Delta_{\vec{k}'} = \sum_{\vec{k}} V_{\vec{k},\vec{k}'} \langle\psi_{\mathrm{BCS}}| c^\dagger_{\vec{k},\downarrow} c^\dagger_{-\vec{k},\uparrow} |\psi_{\mathrm{BCS}}\rangle = \sum_{\vec{k}} V_{\vec{k},\vec{k}'} v_{\vec{k}} u_{\vec{k}} = -\frac{1}{2} \sum_{\vec{k}} \frac{V_{\vec{k},\vec{k}'} \Delta_{\vec{k}}}{E_{\vec{k}}} \,, \tag{1.55}$$

If the momentum dependence of $\Delta(\vec{k})$ includes a sign change between different $\vec{k}$ points, it can 'heal' the wrong sign of a positive interaction $V_{\vec{k},\vec{k}'} > 0$: If the above sum is dominated by a region with positive $\Delta_{\vec{k}^*} > 0$, i.e. $V_{\vec{k}\approx\vec{k}^*,\vec{k}'} \gg$, finite *negative* $\Delta_{\vec{k}'} < 0$ is allowed and can solve the equation.

Moreover, the formalism can be extended to triplet states. There is then one technical aspect that turns out to be related to new physics: When we did the particle-hole transformation Eq. (1.43), we did not introduce any new operators, but just used the $\downarrow$-operators in

a different way. If both creation/annihilation operators in each pairing term have the same spin (e.g. $\uparrow$), this does not help. For most momenta $\vec{k}$, a way out would be to particle-hole transform $c_{\vec{k},\uparrow}$-operators for half the of the Brillouin zone and leave the other half, which contains $c_{-\vec{k},\uparrow}$, untransformed. This scheme breaks down for special momentum points like $\vec{k} = 0, (\pi, 0), \ldots$ where $\vec{k}$ and $-\vec{k}$ are equivalent: Here, the Nambu spinor definitely double counts the physically available fermionic states. If we then undo this in the end, we may in special cases end up with 'half' a fermionic state, which can with some luck be seen as a Majorana fermion.

1.3.3 Ginzburg-Landau Theory of Superconductivity

Even though this was not how it historically arose, the observation of a coherent state with defined phase may somewhat motivate the choice of a complex order parameter in a superconducting Ginzburg -Landau theory. The free energy has to be real and if inhomogeneous states are in principle permitted, a a plausible free energy is

$$\begin{aligned}\Phi &= \int \mathrm{d}^2 r \, \left(A|\psi(\vec{r})|^2 + B|\psi(\vec{r})|^4 + G|\nabla\psi(\vec{r})|^2 + \ldots \right) = \\ &= \int \mathrm{d}^2 r \, \left(a\frac{T - T_C}{T_C}|\psi(\vec{r})|^2 + \frac{b}{2}|\psi(\vec{r})|^4 + \frac{1}{2m}|\nabla\psi(\vec{r})|^2 + \ldots \right) ,\end{aligned} \tag{1.56}$$

where $\psi(\vec{r})$ is the complex and scalar order parameter. In a simple homogeneous system, we expect a homogeneous solution, which is corroborated by experiment, in that case $G > 0$, $\frac{1}{2m}$ is at first just a different way to write this coefficient. For $G > 0$, gradients will be 0 and the analysis can continue like for the magnetic case discussed above and $|\psi| \propto \sqrt{T_C - T}$ turns out to correspond to the superconducting gap Δ of the BCS theory.

The next step is the presence of a magnetic field and we have to find away to include it that is consistent with the known symmetries. In electromagnetism, gauge invariance is the important principle and a gauge invariant way to combine a magnetic field with a complex position-dependent function has already been used when introducing magnetic fields into quantum mechanics: $-i\hbar\nabla \to -i\hbar\nabla - \frac{e^*}{c}\vec{A}$, where charge e^* is some constant and not (yet) known. The gauge invariance means that changing $\vec{A} \to \vec{A} + \vec{\nabla}\Lambda$ does not affect the free energy, because it can be removed by a change in the phase of ψ by $\psi \to \mathrm{e}^{ie^*/(\hbar c)}\psi$ and the phase by itself is not observable.

To analyze Eq. (1.56), we need functional derivatives w.r.t $\vec{A}(\vec{r})$ (the charged order parameter can change the magnetic field), to $\psi(\vec{r})$ and to $\psi^*(\vec{r})$. Some of the steps are worked out in Prof. Muramatsu's lecture notes, let us here just summarize the results:

- The derivative with respect to $\vec{A}$ gives an equation defining a current:

$$\frac{1}{4\pi}\vec{\nabla} \times \vec{B} = \frac{\hbar e^*}{2mc}\left(\psi^* \left(-i\vec{\nabla} - \frac{e^*}{\hbar c}\vec{A} \right)\psi - \psi\left(i\vec{\nabla} - \frac{e^*}{\hbar c}\vec{A} \right)\psi^* \right) \tag{1.57}$$

  As in general $\vec{\nabla} \times \vec{B} = \frac{4\pi}{c}\vec{j}$, we can read off a superconducting current as

$$\vec{j} = \frac{\hbar e^*}{2m}\left(\psi^* \left(-i\vec{\nabla} - \frac{e^*}{\hbar c}\vec{A} \right)\psi - \psi\left(i\vec{\nabla} - \frac{e^*}{\hbar c}\vec{A} \right)\psi^* \right) , \tag{1.58}$$

  where comparison to experiment fixes $e^* = 2e$, i.e., the charge of a Cooper pair.

- Derivatives with respect to ψ and ψ^* give equations for the order parameter:

$$\frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{e^*}{c}\vec{A}\right)^2 \psi(\vec{r}) + a\frac{T - T_C}{T_C}\psi(\vec{r}) + b|\psi(\vec{r})|^2\psi(\vec{r}) = 0 \tag{1.59}$$

These equations can be used to find the mean-field magnetic field and order parameter.