

MEAN-FIELD APPROACH TO MAGNETISM

Zoltán Néda (Babes-Bolyai University, Cluj, Romania)

1. Basic elements of statistical physics

Statistical physics deals with systems composed of a very large number of interacting particles. Since the dynamics of such large number of particles cannot be described by integrating the equations of motion, statistical physics elaborated special methods for deriving relevant information about the system as whole. The starting point for studying a system composed of a large particle number is to fix the external conditions, applied to the system. These external conditions will limit the state-space in which the system can evolve, and will also fix the probability of the allowed states. The most used ensemble in such sense is the canonical ensemble, where in case of pure mechanical interactions the V (volume), T (temperature) and N (particle number) are fixed. This means that we consider a closed system in contact with a heat-bath at temperature T . Due to the heat-exchange with the heat reservoir the energy can fluctuate. We learn from thermodynamics that in such conditions the relevant thermodynamic potential is the free energy F . The aim of statistical physics is to determine thus the $F(V,N,T)$ functional relation for the studied system, based on an elaborated microscopic model. The knowledge of F is sufficient to obtain all other relevant information by using only simple thermodynamics.

The weight factors (probabilities) of the allowed microstates (i) of the system are given as:

$$p_i = \frac{\exp(-E_i/kT)}{Z} \quad (1)$$

where E_i is the internal energy of the system in the state-space point i and Z is the canonical sum which is also called canonical partition function.

$$Z = \sum_{\{i\}} \exp(-E_i/kT) \quad (2)$$

The relevant thermodynamic potential of the system can be then calculated from the value of Z as:

$$F = -kT \ln(Z) \quad (3)$$

Once the value of Z is known, everything else can be calculated with simple thermodynamics. Solving thus theoretically a microscopic model in statistical physics implies the calculation of a huge sum. A magnetic system is in canonical ensemble when T , N and the external magnetic field H is fixed. In case of magnetic system one aim to determine thus $Z(T,V,H)$, and consequently using (3) the $F(T,V,H)$ free energy. The magnetization of the system is determined as

$$M = -\frac{1}{\mu_0} \frac{\partial F}{\partial H} \quad (4)$$

The most important experimentally measurable material dependent quantities that characterize the magnetic materials is the *susceptibility* (χ). Susceptibility can be determined now as:

$$\chi = \lim_{H \rightarrow 0} \frac{1}{V} \frac{\partial M}{\partial H} = -\frac{1}{V\mu_0} \frac{\partial^2 F}{\partial H^2} \quad (5)$$

For a thermodynamic system in canonical ensemble the energy fluctuates in time. In case of magnetic systems the magnetization also fluctuates in time.

Similarly with classical systems, a quantum mechanical system is in a canonical ensemble when T , V (H for magnetic systems) and N are fixed. For a quantum mechanical system one can define a set of stationary states over which the canonical sum in (2) is calculated. In order to determine thus the partition function, one has to first solve the stationary Schrödinger equation, and determine the

stationary states i and the energy E_i corresponding to these ones. This is however a quite complicated task by itself and very seldom is possible. Calculating the stationary states is equivalent with diagonalizing the \hat{H} Hamiltonian operator. Fortunately, there is however another possibility to calculate Z , without the need of diagonalizing the Hamiltonian operator. The partition function can be computed as

$$Z = \text{Tr}[\exp(-\hat{H}/kT)] \quad (6)$$

where Tr denotes the trace of an operator, and the operator $\exp(-\hat{H}/kT)$ is defined as:

$$\exp(-\hat{H}/kT) = \hat{I} + \frac{1}{1!} \hat{H} + \frac{1}{2!} \hat{H}^2 + \dots + \frac{1}{n!} \hat{H}^n + \dots \quad (7)$$

Computing the trace of the above determined operator is thus another alternative. This trace can be computed in any complete base, and not only in the base that diagonalizes the Hamiltonian. Once the value of Z is known, calculations of the relevant physical quantities is the same as in the classical limit.

2. Interactive versus non-interactive systems

In statistical physics an exact solution of the thermodynamic model means the exact calculation of the partition function Z . This is possible however only for a few and simple systems. The majority of the models that are interesting for practical applications do not allow an exact calculation of the partition function, calculating the sum in (2) being way to difficult. In case of systems composed by non-interacting particles calculating the partition function is simple, due to the fact that one can decompose the partition function in the product of the partition functions calculated for each particle apart. The math is quite simple, first for classical (not quantum mechanical) particles one can naturally write:

$$Z = \sum_i \exp(-\beta E_i) = \sum_{i_1, i_2, \dots, i_N} \exp(-\beta E_{i_1, i_2, \dots, i_N}) \quad (8)$$

where (i_1, i_2, \dots, i_N) denotes the possible microstates of the particles that compose the thermodynamic system. Since the system is not interactive, these one-particle states really exist. For classical non-interactive particles one can also write

$$E_i = \sum_{k=1}^N E_{i_k} \quad (9)$$

and thus:

$$Z = \sum_{i_1, i_2, \dots, i_N} \exp[-\beta(E_{i_1} + E_{i_2} + \dots + E_{i_N})] = \sum_{i_1} \exp(-\beta E_{i_1}) \sum_{i_2} \exp(-\beta E_{i_2}) \dots \sum_{i_N} \exp(-\beta E_{i_N}) \quad (10)$$

Denoting by:

$$Z_k = \sum_{i_k} \exp(-\beta E_{i_k}) \quad (10)$$

the partition function for particle k , the partition function of the whole system writes as:

$$Z = Z_1 \cdot Z_2 \cdot \dots \cdot Z_N \quad (11)$$

For systems made up of identical and non-interacting particles $Z_1 = Z_2 = \dots = Z_N$ and one gets

$Z = (Z_1)^N$, and if one takes into account that the particles are indistinguishable, we get:

$$Z = \frac{1}{N!} (Z_1)^N \quad (12)$$

The evaluation of the Z_1 partition function for one particle is usually quite straightforward, and in such sense non-interactive systems are easy to handle analytically. Since the above method does not hold for systems of interactive particles, handling such systems is quite difficult. If one cannot

calculate exactly the partition function, the natural way to work is to do some approximations. Several approximation methods are known between them the easiest one is the mean-field approximation. Other well-known approximations are the low and high temperature expansion, renormalization, scaling or the use of numerical methods.

3. Phase-transitions and critical phenomena

Phase-transitions are characteristic for thermodynamic systems composed of interacting entities. We speak about phase-transitions whenever the properties of a thermodynamic system change in a qualitative and many times discontinuous manner. These transitions happen usually near some critical points or surfaces in the parameter-space. The behavior of the system in the vicinity of such critical regions is known as critical behavior. Phase transitions are usually a direct consequence of the interactions between the elements composing the thermodynamic system.

The most well-known classification of phase-transitions is due to Ehrenfest, based on the behavior of the F free energy. According to this classification the order of the phase-transition is given by the lowest derivative of the free energy as a function of some thermodynamic variable that becomes discontinuous at the transition point. First-order phase-transitions are those where the first derivative of F (which is related to density or order parameter) becomes discontinuous (Figure 1). In the second-order phase transitions, second order derivatives (related to specific heat, susceptibility, etc...) become discontinuous. Such phase transitions are characteristic for example in ferromagnetic systems in the neighborhood of the Curie temperature. In principle according to the Ehrenfest classifications one can have third, fourth and higher order phase-transitions as well.

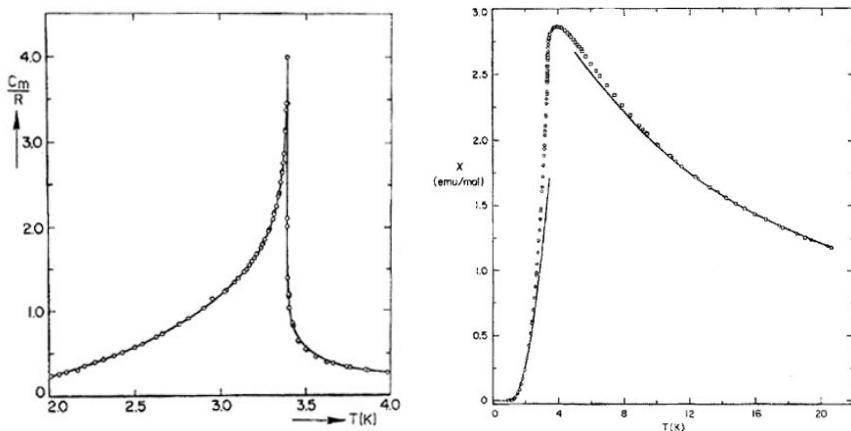


Figure 1. Magnetic specific heat (left) and susceptibility (right) for DyPO_4 as a function of the reduced temperature. The points (o) represent experimental results, the solid line represents the results of a calculation based on high- and low-temperature series expansions with one adjustable constant (from J.C. Wright *et al.*, Phys. Rev. B **3**, 843 (1971)).

Nowadays we prefer to speak about first-order phase transitions whenever latent heat is involved (melting of ice, boiling of water). We refer to second order phase-transitions as continuous phase transitions. The order parameter of the system behaves continuously in the vicinity of the critical point and we observe power-law behavior or divergences for the susceptibility, specific heat and correlation length.

The most important quantity for characterizing a phase-transition is the **order parameter**. The order parameter is usually an adimensional quantity, having zero value in one phase and non-zero value for the other phase, it is meant to characterize the onset of order at the phase-transition point.

For liquid-solid or liquid-gas transitions the order parameter is related to the density of the system. For magnetic phase-transitions the order-parameter is related to the spontaneous magnetization of the system. For example in the para-ferromagnetic phase transition, the order parameter is related to the average spontaneous magnetization (M) per spin. An adimensional order parameter between 0 and 1 can be obtained as:

$$m = \left| \frac{\vec{M}}{\mu_B g N S} \right| = \left| \frac{\langle \vec{S} \rangle}{S} \right| \quad (13)$$

where $\mu_B = \frac{e\hbar}{2m_e}$ is the Bohr-magneton and g is the giromagnetic factor. Note that S is the spin momentum generating the magnetic dipoles and is in adimensional units, the unit \hbar is already included in the value of the Bohr magneton. This order parameter is zero in the paramagnetic phase and non-zero in the ferromagnetic one.

We illustrate in Figure 2 the behavior of this order-parameter in the vicinity of the critical temperature. From the figure one observes that the spontaneous magnetization goes to zero at the critical temperature. The convergence to zero is according to a power-law.

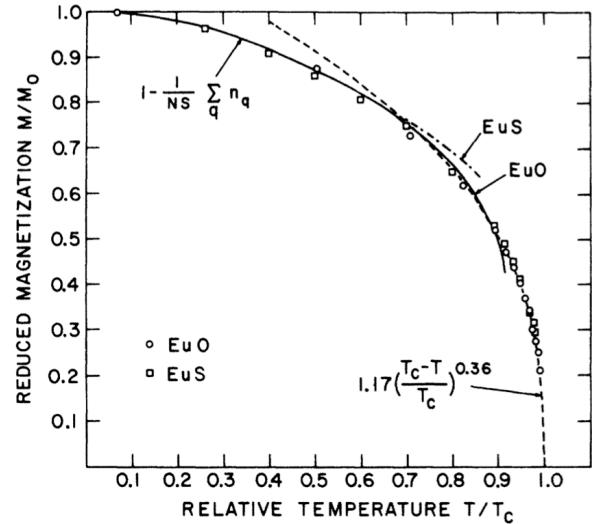
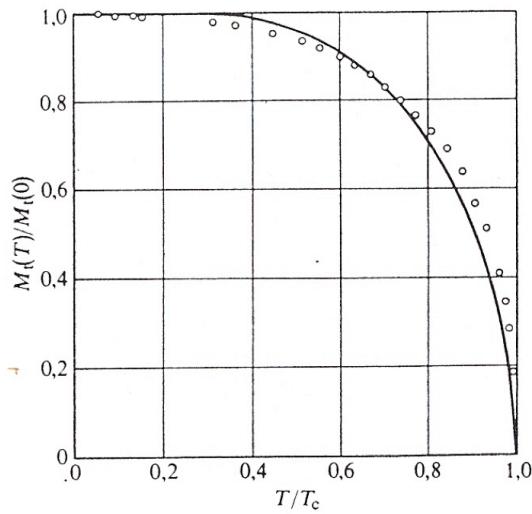


Figure 2. Left: spontaneous magnetization of Nickel as a function of temperature. Experimental results are from Weiss and Forrer (Annn. Phys. Vol. 5, 153 (1926)) and the curve is the mean-field approximation considering $S=1/2$. Right: Temperature dependence of the reduced magnetization of EuO and EuS. The solid and dash-dotted curves were calculated using self-consistent spin-wave-theory.

For second order-phase transitions the second derivatives of the free energy (specific heat, susceptibility) and the correlation length diverges as a power-law at the phase-transition point. This behavior is known as critical behavior and is characterized by critical exponents. Let us consider quantity A which is a second order derivative of the free-energy (specific heat or susceptibility) as a function of a thermodynamic parameter. If the system exhibits a second-order phase-transition at temperature T_c , one can write that:

$$A(T)_{T \rightarrow T_c} \approx C \left| 1 - \frac{T}{T_c} \right|^a \quad (14)$$

The exponent α is known as the critical exponent. On Figure 3 we illustrate the validity of such scaling relations for the susceptibility near the phase-transition point.

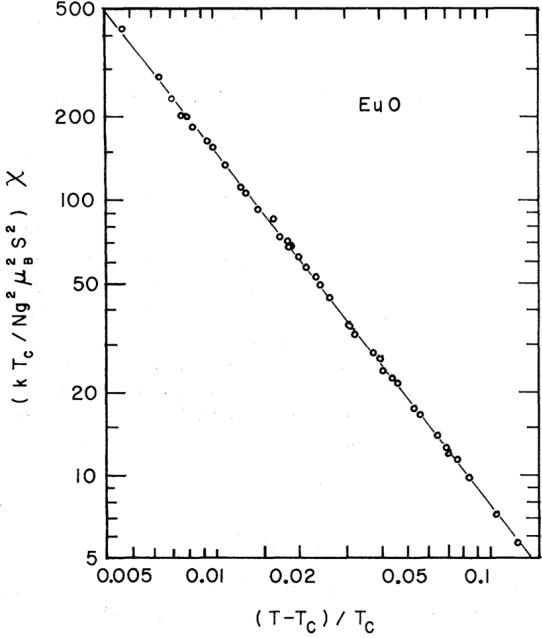


Figure 3. Temperature dependence of the susceptibility of EuO above the critical temperature. The line has a slope $\gamma = 1.3$

For magnetic systems one can define the critical exponents for the order parameter (magnetization) (β), susceptibility (γ), specific heat (α) and correlation length (μ):

$$\begin{aligned} m &= m_0 \left| 1 - \frac{T}{T_c} \right|^\beta \\ \chi &= \chi_0 \left| 1 - \frac{T}{T_c} \right|^{-\gamma} \\ C_v &= C_0 \left| 1 - \frac{T}{T_c} \right|^{-\alpha} \\ \xi &= \xi_0 \left| 1 - \frac{T}{T_c} \right|^{-\mu} \end{aligned} \quad (15)$$

Critical exponents are important quantities for studying phase-transitions, since it presents the feature of universality: it does not depend on the microscopic details of the interaction among the particles, it depends only on the symmetry properties and dimensionality of the thermodynamic system. Successful models elaborated for real thermodynamic systems have to reproduce their measured values, this being a test of the model. An important feature is that the critical exponents are not independent of each other, they are linked by so-called scaling laws. For example in two-dimensional systems it is valid the Rushbrooke identity:

$$\alpha + 2\beta + \gamma = 2 \quad (16)$$

Another generally valid hyper-scaling equation is: $d\mu = 2 - \alpha$, where d is the dimensionality of the system. The reason behind these scaling laws are that ultimately each characteristic quantity is linked analytically to the correlation length. The reason behind the divergence of the specific heat and susceptibility is the divergence of the correlation length. In vicinity of the critical point fluctuations also increases strongly, and eventually diverges. The specific heat and susceptibility is linked to the fluctuations by the well-known equations:

$$\begin{aligned} C_V &= \frac{1}{kT^2 N} (\langle E^2 \rangle - \langle E \rangle^2) \\ \chi &= \frac{1}{kTN} (\langle M^2 \rangle - \langle M \rangle^2) \end{aligned} \quad (17)$$

For time-dependent phenomena, one can also define the critical exponent of the relaxation times. If we are interested in an A measurable quantity, which relaxes in time to its equilibrium value A_0 , the relaxation time is defined as:

$$\tau = \frac{\int_0^\infty t |A(t) - A_0|}{\int_0^\infty |A(t) - A_0|} \quad (18)$$

As the correlation length diverges in the neighborhood of the critical temperature, the relaxation time also diverges. The two quantities are linked by the so-called dynamical critical exponent (ζ):

$$\tau \propto \xi^\zeta \quad (19)$$

One of the aims of the statistical physics is to study models that exhibit phase-transitions and to calculate the critical point and critical exponents of this system. Since exact analytical solutions are rarely possible, the values computed for the critical parameters and critical exponents are also only rough estimates. One approximation that usually is quite straightforward to implement is the **mean-field approximation**. As the dimensionality of the model increases the mean-field approximation gets better and better, and eventually after a critical dimension it leads exact values for the critical exponents.

4. Mean-field approximation for gases

In order to understand the essence of mean-field approximations let us consider first the case of gases. The ideal gas model considers non-interactive point-like particles. For such a mechanical system the Hamiltonian writes as:

$$E = \sum_i E_i = \sum_i \frac{1}{2m} (p_{xi}^2 + p_{yi}^2 + p_{zi}^2) \quad (20)$$

where (p_x, p_y, p_z) denotes the components of the momentum. The partition function for one particle can be easily estimated as:

$$Z_1 = \int_{\{x_1\}} \int_{\{y_1\}} \int_{\{z_1\}} \int_{\{p_{x1}\}} \int_{\{p_{y1}\}} \int_{\{p_{z1}\}} n(x_1, y_1, z_1, p_{x1}, p_{y1}, p_{z1}) \exp[-\beta(p_{x1}^2 + p_{y1}^2 + p_{z1}^2)/2m] dx_1 dy_1 dz_1 dp_{x1} dp_{y1} dp_{z1} \quad (21)$$

where $n(x_1, \dots, p_{z1})$ denotes the local density of the allowed state-space points of one particle. Taking into account that for the ideal gas $n(x_1, \dots, p_{z1}) = \text{const}$, one gets

$$Z = \frac{1}{N!} (Z_1)^N = \frac{1}{N!} (K V T^{3/2})^N, \quad (22) \text{ with } K \text{ a constant.}$$

Correspondingly we get $F = -\frac{3}{2}NkT \ln(KVT) + kT[N \ln(N) - N]$ (23)

which leads us to the equation of states:

$$p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \Rightarrow PV = NkT.$$

Considering interaction between the particles makes the model more complicated. Let us introduce now interactions between the particles that depend only on the distance between the particles. The Hamiltonian of the system writes now as:

$$E = \sum_i \frac{1}{2m}(p_{xi} + p_{yi} + p_{zi}) + \frac{1}{2} \sum_{i \neq j} U(r_{ij}), \quad (24)$$

where r_{ij} denotes the distance between particles i and j , and U is the inter-particle potential energy.

Due to the second interaction term in the Hamiltonian, calculation of the partition function becomes very complicated. The complexity of the problem can be reduced by introducing a pair-correlation function $g(r)$ that denotes the probability to have another particle at distance between r and $r+dr$, around each particle. Assuming this pair-correlation function similar for all particles, the total energy of the system can be written as a separate sum for all particles in the system. This is equivalent with the fact that all particles fill a U_{eff} mean, or effective field that is generated by the others. This approach is thus a typical mean-field approximation for the complex interactive system:

$$\frac{1}{2} \sum_{i \neq j} U(r_{ij}) \approx \frac{1}{2} \sum_{i=1}^N \int_0^\infty g(r)U(r)dr = \sum_i U_{eff} \quad (25)$$

In principle determining the $g(r)$ pair-correlation function and, thus the U_{eff} mean-field is a hard task, and would also require the solution of the initial many-body problem. However, as a first order approximation one can consider the distribution of the molecules homogeneous in space (which is not really the case), and one would obtain than $g(r) \approx 4\pi r^2 \rho$, with $\rho = N/V$ the average concentration of the molecules. The effective potential can be calculated as:

$$U_{eff} \approx \frac{1}{2} \int_0^\infty 4\pi r^2 \rho U(r)dr = \rho C \quad (26)$$

Here C is a negative constant. By using this approximation the system is decoupled, and one can use again the well-established method for calculating the partition function in case of non-interacting systems. The partition function for one particle writes as:

$$Z_1 = \int_{\{x_1\}} \int_{\{y_1\}} \int_{\{z_1\}} \int_{\{p_{x1}\}} \int_{\{p_{y1}\}} \int_{\{p_{z1}\}} n(x_1, y_1, z_1, p_{x1}, p_{y1}, p_{z1}) \exp[-\beta(p_{x1}^2 + p_{y1}^2 + p_{z1}^2)/2m + U_{eff}] dx_1 dy_1 dz_1 dp_{x1} dp_{y1} dp_{z1} \quad (27)$$

Taking into account also the finite volume, b , of each molecule (which reduces the state-space points after the spatial coordinates), one immediately gets:

$$Z = \frac{(Z_1)^N}{N!} = \frac{[T^{3/2} K(V - Nb) e^{-\beta nC}]^N}{N!} \quad (28)$$

After calculating the free energy one gets for the thermal equation of state:

$$(P - n^2 C)(V - Nb) = NkT \quad (29)$$

which is the well-known Van der Waals equation.

5. Mean field approximation for ferromagnets

Let us first study the case when we have localized magnetic moments, generated by angular momentums \vec{J} :

$$\vec{\mu} = \mu_B g \vec{J} \quad (30)$$

(Here $\mu_B = \frac{e\hbar}{2m_e}$ is the Bohr-magneton and g is the giromagnetic factor. Note that J is in adimensional units, the unit \hbar is already included in the value of the Bohr magneton).

The total magnetization of the system writes as:

$$\vec{M} = \sum_i \vec{\mu}_i \quad (31)$$

In case the system is in interaction only with the external field H (dipolar interactions between the magnetic moments are neglected), the total energy of the system writes as

$$E = \sum_i E_i = -\mu_0 \vec{H} \sum_i \vec{\mu}_i = -\mu_0 \mu_B g \vec{H} \sum_i \vec{J}_i = -\mu_0 \mu_B g J H \sum_i \cos(\theta_i) \quad (32)$$

where θ_i denotes the angle between the magnetic field and the orientation of the angular momentum. (The direction of the z axis of the coordinate system is chosen in the direction of the magnetic field H .) Since we have a non-interacting system it is enough to compute the value of Z_1 . Assuming classical angular momentums, with arbitrary (non-quantized) orientations in space, one gets:

$$Z_1 = \int_{\varphi} \int_{\theta} A(\theta, \varphi) \exp(\beta \mu_0 \mu_B g J H \cos(\theta)) d\varphi d\theta, \quad (33)$$

where $A(\varphi, \theta) = A \sin(\theta) d\theta d\varphi$ is the density of states for the angular momentum vector (A is a constant).

A simple calculation yields:

$$Z_1 = 4\pi A \left[\frac{\sinh(t)}{t} \right] \text{ with } t = \beta \mu_0 \mu_B g J H. \quad (34)$$

After calculating the free-energy $F = -kT N \ln(Z_1)$, one gets:

$$M = -\frac{1}{\mu_0} \left(\frac{\partial F}{\partial H} \right)_{T,N} = \mu_B g J N \left[ctgh(t) - \frac{1}{t} \right] = \mu_B g J N L(t) \quad (35)$$

We denote here by $L(t)$ the Langevin function: $L(t) = ctgh(t) - \frac{1}{t}$. (36)

The susceptibility becomes:

$$\chi = \frac{1}{V} \left(\frac{\partial M}{\partial H} \right)_{H \rightarrow 0} = \frac{N \beta \mu_0 \mu_B^2 g^2 J^2}{V} \left[\frac{1}{t^2} - \frac{1}{\sinh^2(t)} \right]_{t \rightarrow 0} = \mu_0 \frac{\mu_B^2 g^2 J^2 N}{3kTV} \propto \frac{1}{T} \quad (37)$$

If one considers the quantum mechanical nature of the angular momentum vector, and takes into account the allowed

$J_z \in \{-J, -J+1, -J+2, \dots, J-1, J\}$ values, and repeats the same calculations, one would get instead:

$$Z_1 = \frac{\sinh \left[t \left(J + \frac{1}{2} \right) \right]}{\sinh \left(\frac{t}{2} \right)} \quad (38)$$

This leads to:

$$M = N\mu_B g \left[\left(J + \frac{1}{2} \right) ctgh \left(t \frac{2J+1}{2J} \right) - \frac{1}{2} ctgh \left(\frac{t}{2J} \right) \right] = N\mu_B g J B_J(t) \quad (39)$$

where

$$B_J(x) = \frac{2J+1}{2J} \coth \left(\frac{2J+1}{2J} x \right) - \frac{1}{2J} \coth \left(\frac{x}{2J} \right) \text{ denotes the Brillouin function.}$$

The susceptibility becomes:

$$\chi = \mu_0 \frac{\mu_B^2 g^2 J(J+1)N}{3Vkt} \propto \frac{1}{T} \text{ leading to the same Curie law.}$$

Let us consider now interacting vectorial magnetic moments, leading to a spontaneous magnetization. Let us consider that the magnetic moment originates from the spin momentum. We consider thus spins fixed on the sites of a lattice that interact via exchange interactions that are of pure quantum-mechanical nature. The exchange interaction decays very quickly with the separation distance, so it is assumed that only first-neighbor spins are interacting. The exchange interaction energy of the system writes as:

$$E = -j \sum_{\langle k,l \rangle} \vec{S}_k \cdot \vec{S}_l \quad (40)$$

where we denote by j the exchange interaction constant, and the brackets under the summation sign suggests that the sum is taken only on nearest neighbor pairs. We assume in the first attempt that the spins are classical ones, which means that we treat the problem in classical statistical physics, and we assume that the spins are classical vectors with all allowed spatial directions.

We mention here that the quantum-mechanical model based on equation (40) would be the Heisenberg model, where the Hamiltonian operator of the system would be:

$$\hat{H} = -j \sum_{\langle k,l \rangle} \hat{\vec{S}}_k \cdot \hat{\vec{S}}_l = -j \sum_{\langle k,l \rangle} (\hat{S}_k^x \hat{S}_l^x + \hat{S}_k^y \hat{S}_l^y + \hat{S}_k^z \hat{S}_l^z), \text{ assuming that} \\ [\hat{S}^i, \hat{S}^k] = i\hbar \epsilon_{ikl} \hat{S}^l$$

Exact quantum-statistical solution of this model is known however only in one dimension (chain of quantum-spins). The Ising model is a simpler special case, where the quantum number of spins is 1/2 and the interactions between the x and y components are neglected.

$$\hat{H} = -j \sum_{\langle k,l \rangle} S_k^z S_l^z \text{ with } S_i^z = \pm 1/2$$

Although the Ising model is the simplest quantum-spin model that exhibits a para-ferromagnetic phase transition, it cannot be exactly solved in three dimensions. In one dimension several hundred different exact solutions are known (exact calculation of the partition function). In two-dimension the exact solution is much harder and it was first given by Lars Onsager in 1944 using the transfer matrix method. Since then, several other exact solutions were given on 2D lattices. Exact solution in the three-dimensional case is not known however. The simplest solution is again the mean-field approximation, which works independently of the dimensionality of the considered lattice.

We will consider in the following the mean-field solution of the semi-classical vector spin model with the Hamiltonian given in (40). For the sake of completeness let us add also the interaction of the magnetic moments generated by the spins with an external magnetic field. The Hamiltonian of the system writes as:

$$E = -j \sum_{\langle k,l \rangle} \vec{S}_k \cdot \vec{S}_l - \bar{H} \mu_0 \mu_B g \sum_k \vec{S}_k \quad (41)$$

We want to calculate the partition function of the system:

$$Z = \sum_{\{\vec{S}_i\}} e^{-\beta E(\vec{S}_i)} \quad (42)$$

Summation refers here on all allowed $\{\vec{S}_i\}$ spin configurations and due to the first-neighbor interaction topology exact calculation of the partition function is very complicated. The simplest approach we can do is again the mean-field approximation. The mean-field approximation considers that every spin is interacting NOT with its real neighbors, but with “mean-neighbors” ($\langle \vec{S} \rangle$), or with a field generated by the mean orientation of the spins. This is schematically illustrated in Figure 4.

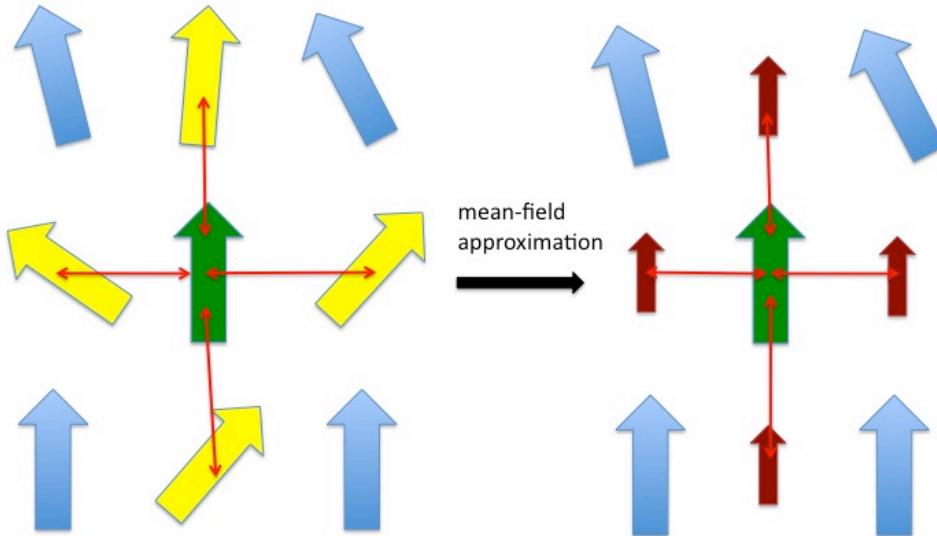


Figure 4. Schematic illustration for the mean-field approximation on a 2D square lattice

The Hamiltonian (energy) of the system becomes in this approximation:

$$E = -j\langle \vec{S} \rangle q \sum_i \vec{S}_i - \vec{H} \mu_0 \mu_B g \sum_i \vec{S}_i = -(jq\langle \vec{S} \rangle + \vec{H} \mu_0 \mu_B g) \sum_i \vec{S}_i = \sum_i E_i \quad (42)$$

Another method to arrive to the same mean-field Hamiltonian is the following:

Let us consider the evident inequality

$$\vec{S}_i = \langle \vec{S}_i \rangle + (\vec{S}_i - \langle \vec{S}_i \rangle) = \langle \vec{S} \rangle + (\vec{S}_i - \langle \vec{S} \rangle) \quad (43)$$

where the second term is the fluctuation term. The Hamiltonian can be written now as:

$$E = -j/2 \sum_{\{k,l\}} [\langle \vec{S} \rangle + (\vec{S}_k - \langle \vec{S} \rangle)] [\langle \vec{S} \rangle + (\vec{S}_l - \langle \vec{S} \rangle)] - H \mu_0 \mu_B g \sum_i \vec{S}_i \quad (44)$$

The factor $1/2$ arises from the fact that now the summation is performed two times on both (k,l) pairs. The first term can be expended as:

$$\begin{aligned}
E_1 &= -j/2 \sum_{\{k,l\}} \left[\langle \vec{S} \rangle + (\vec{S}_k - \langle \vec{S} \rangle) \right] \left[\langle \vec{S} \rangle + (\vec{S}_l - \langle \vec{S} \rangle) \right] = \\
&= -j/2 \left[- \sum_{\{k,l\}} \langle \vec{S} \rangle \langle \vec{S} \rangle + 2 \sum_k \vec{S}_k \left(\sum_l \langle \vec{S} \rangle \right) + \sum_{\{k,l\}} (\vec{S}_k - \langle \vec{S} \rangle)(\vec{S}_l - \langle \vec{S} \rangle) \right]
\end{aligned} \tag{45}$$

Now, in this approach we assume fluctuations small, and neglect thus the last term, which contains the square of these fluctuations. In the second term the summation over l refers to the neighbors on spin k . Assuming that each spin has q neighbors, we can write:

$$E_1 = -j/2 \left[- \sum_{\{k,l\}} \langle \vec{S} \rangle^2 + 2q \langle \vec{S} \rangle \sum_i \vec{S}_i \right] \tag{46}$$

The total energy writes thus:

$$E = j/2 \sum_{\{k,l\}} \langle \vec{S} \rangle^2 - jq \langle \vec{S} \rangle \sum_i \vec{S}_i - H \mu_0 \mu_B g \sum_i \vec{S}_i \tag{47}$$

which differs from the energy written in (42) only in a constant factor (first term). The statistical properties of the Hamiltonians in (47) and (42) are thus the same.

By doing this approximations we have decomposed our interacting particle system in a system of particles (spins) where interactions are only with an external field. This external field is however given both by the real external field (\vec{H}) and the mean-field generated by the neighbors ($jq \langle \vec{S} \rangle$). If the value of $\langle \vec{S} \rangle$ would be known, the problem can be solved in an analogous manner with the paramagnetic (non-interacting systems). We assume again the direction of the z axis in the direction of the external magnetic field. The rotational symmetry of the spins is broken now by the direction of the external magnetic field, so we would get that the direction of $\langle \vec{S} \rangle$ is also in the z axis direction and its modulus is $\langle S \rangle$. Let us assume thus that $\langle \vec{S} \rangle$ is known and proceed in the same manner as in the case of non-interacting magnetic moments. Quite straightforwardly we would immediately get that:

$$M = \mu_B g N S L(f) \tag{48}$$

$$f = \beta S [jq \langle S \rangle + \mu_0 \mu_B g H] \tag{49}$$

We can write however, that: $M = N \mu_B g \langle S \rangle$, leading thus to an implicit equation for the value of $x = \langle S \rangle$:

$$x = S L [\beta S (jqx + \mu_0 \mu_B g H)] \tag{50}$$

By solving equation (50) one immediately gets the value of magnetization and from there the the value of susceptibility can be also calculated. However, due to the $cosh$ function in the Langevin function the equation cannot be analytically solved. Approximations and a graphical solution can be attempted.

Let us study first the case without no external magnetic field ($H = 0$). In such cases we have:

$$x = S L [x \beta S jq] \tag{51}$$

We attempt now a graphical solution and use the form of the Langevin function (Figure 5)

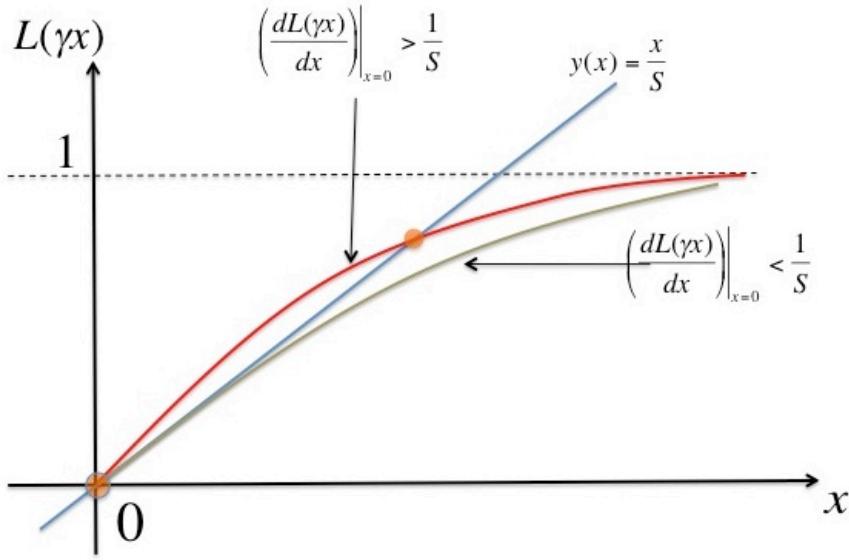


Figure 5. Graphical solution of the $L(yx) = \frac{x}{S}$ equation, depending on the value of γ .

The nature of the solution depends on the slope of the $L(yx)$ function at $x=0$. The slope at $x=0$ is given by:

$$\left. \frac{dL(yx)}{dx} \right|_{x=0} = \frac{1}{3}\gamma \quad (52)$$

This defines thus in equation (52) the critical temperature: T_c :

$$\frac{1}{3}\gamma = \frac{jqS}{3kT_c} = \frac{1}{S} \rightarrow T_c = \frac{jqS^2}{3k} \quad (53)$$

If $T > T_c$ than only one solution is possible $x=0$, which means no spontaneous magnetization ($M=0$). For $T < T_c$ there is however another $x \neq 0$ ($M \neq 0$) solution as well, which becomes stable, and the $x=0$ solution is unstable.

From the used mean-field approximation one concludes thus that there is hint for a phase-transition in the system. There is a critical temperature in the system, below which the system is ferromagnetic and spontaneous magnetization appears by a spontaneous symmetry breaking. Above the critical temperature the system is paramagnetic and no spontaneous magnetization is possible.

Mean-field approximation leads thus for an elegant explanation of the paramagnetic-ferromagnetic phase transition. It predicts the value of the critical temperature, and also predicts values for the main critical exponents of the system. As we will see in the following, these values for low dimensional systems are quite far from the ones that are given (or expected) from an exact solution.

If instead of classical vector spins, we would have worked with semi-classical spins where the quantization in the z axis direction is taken into account we would have get the same equation (50) where instead of the $L(x)$ Langevin function the $B(x)$ Brillouin function would appear:

$$x = SB_S[\beta S(jqx + \mu_0\mu_B gH)]. \quad (54)$$

Using the fact that $\lim_{x \rightarrow 0} B_S(x) = \frac{S+1}{3S}x$, we immediately get that:

$$T_C = \frac{S(S+1)jq}{3k} \quad (55)$$

For the case of the Ising model ($S=1/2$) on a square lattice ($q=4$) one would get: $T_C^{\text{mean-field}} = \frac{j}{k}$, in comparison with the exact result $T_C = \frac{j}{k} \frac{1}{2\ln(1+\sqrt{2})} \approx 0.567 \frac{j}{k}$. For the three dimensional case (cubic lattice) one gets $T_C^{\text{mean-field}} = \frac{3j}{2k}$ in comparison with the believed $T_C \approx 1.11 \frac{j}{k}$ exact value. As the dimensionality of the system increases this mean-field approach gets thus better and better. In one dimension however the results are totally wrong! In a 1D lattice one can quickly solve exactly the problem and gets $T_C = 0$, which means that there is no phase transition for $T > 0$ temperatures and at any finite temperature the system is paramagnetic. Spontaneous order is thus not possible in 1D. In contrast with this exact results the mean field approach would give $T_C^{\text{mean-field}} = \frac{j}{2k}$, proving once again that one has to be very careful with trusting mean-field results in low-dimensional systems.

Tc for the Ising model:

Table 5.2 Exact and approximate Curie temperatures for the Ising model (in units of zJ/k_B).

lattice	d	z	mean-field	Oguchi	exact
linear chain	1	2	1	0.782	0.000
square	2	4	1	0.944	0.567
simple cubic	3	6	1	0.974	0.752
bcc	3	8	1	0.985	0.794
fcc	3	12	1	0.993	0.816

Table 1 Critical temperatures for the Ising model ($S=1/2$) in various approaches.

One can give a very simple explanation for the fact why there is no phase-transition in a 1D classical spin-chain with nearest neighbor interactions. Let us consider a chain of Ising spins all pointing upwards as in Figure 6 (top). Turning downward one of the spins necessitates an energy $4j$. This is shown in the middle line of Figure 6. From here on however all spins can be turned downwards without any added energy, and one can immediately get the configuration in the bottom line of Figure 6. This suggests that an ordered state (all spins up) cannot be stable at any finite temperature, since any tiny energy fluctuation can turn inversely the total magnetization. In other words we cannot have spontaneous magnetic ordering at finite temperatures in 1D systems. An exact solution based for example on the transfer matrix method will convince us immediately about the relevance of this simple reasoning.

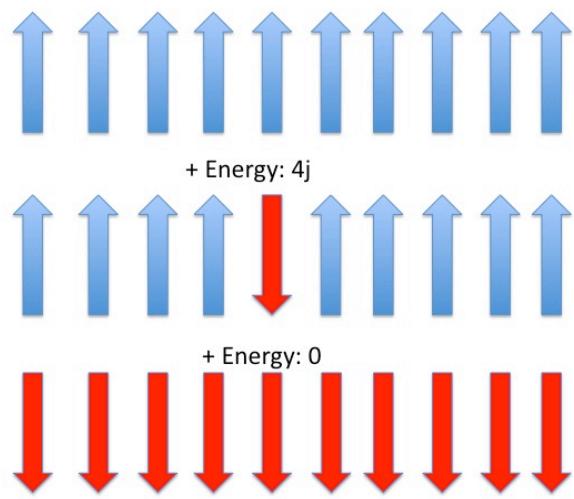


Figure 6. Effect of small fluctuations on an ordered Ising chain.

5.1 Presence of magnetic field.

In the presence of an external magnetic field the graphical solution of equation (50) is given in Figure 7:

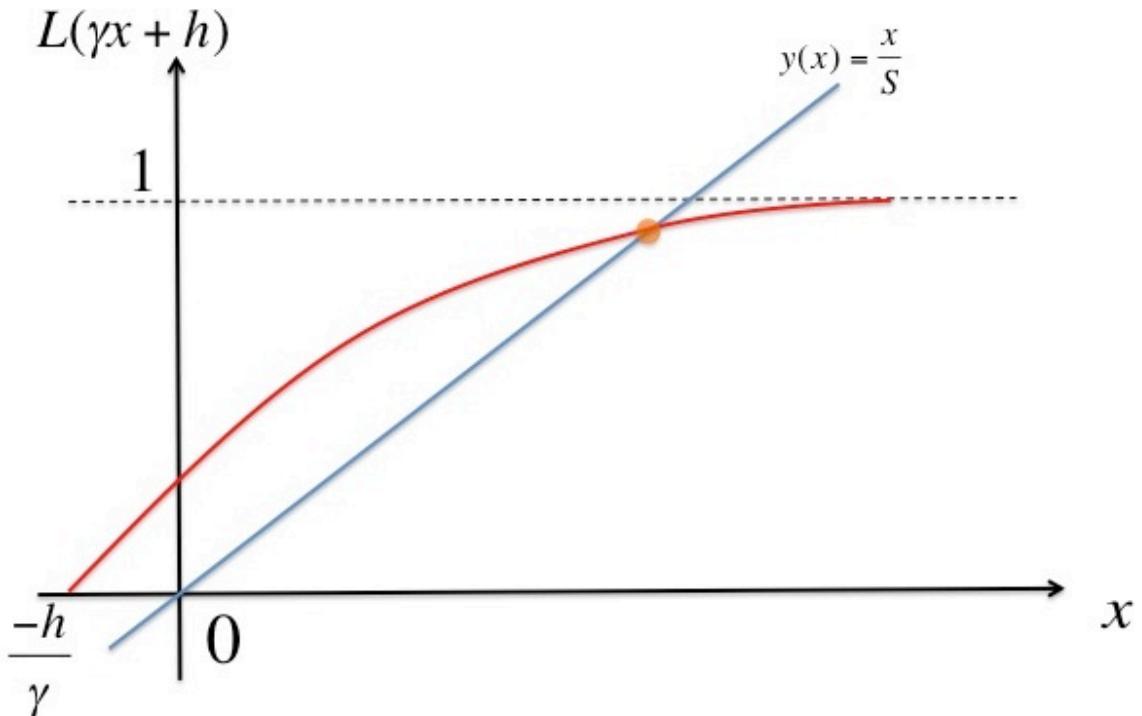


Figure 7. Graphical solution of equation (50) in the presence of an external magnetic field.

The graphic suggests, that for $H \neq 0$ there is always a solution with $M \neq 0$.

Let us determine now the $M(H)$ behavior and the susceptibility in the vicinity of T_c . For the sake of simplicity we work with vector spins. Let us assume $H \rightarrow 0$, and in the limit of $T > T_c$ we have $\langle S \rangle \rightarrow 0$. Consequently in equation (50) we obtain $f \rightarrow 0$. In this limit we get:

$$L(f) \approx \frac{1}{3}f - \frac{1}{45}f^3 + \frac{2}{945}f^5 - \dots \quad (56)$$

We limit ourself to the first term, and we get thus from equation (50):

$$\frac{\langle S \rangle}{S} = \frac{1}{3}(\beta S(jq\langle S \rangle + \mu_0\mu_B gH)) \quad (57)$$

$$\text{Assuming that the total magnetization is: } M = Ng\mu_B\langle S \rangle \quad (58)$$

We immediately get for the $H \rightarrow 0$ limit:

$$M = \frac{NS^2\mu_0\mu_B^2g^2}{3k} \frac{H}{T - T_c} \quad (59)$$

which yields: $\chi = \frac{1}{T - T_c} \frac{NS^2\mu_0\mu_B^2g^2}{3k}$, the well-known Curie-Weiss law for ferromagnetic materials.

If instead of the classical vector spins, we would have worked with quantized z components, we would have got the results:

$$\chi = \frac{1}{T - T_c} \frac{NS(S+1)\mu_0\mu_B^2g^2}{3k} \quad (60)$$

This result shows that the critical exponent of the susceptibility is $\gamma = 1$. Experimental results for high enough temperatures show a trend that is in good agreement with the prediction of the mean-field approach (Figure 8). It is also observable from the figure that in the vicinity of the critical temperature the critical exponent differs from the $\gamma = 1$ mean-field prediction.

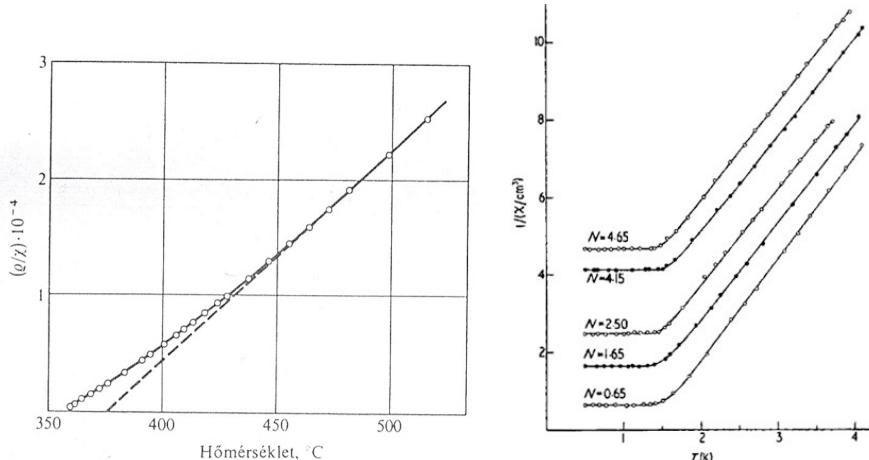


Figure 8. The reciprocal of the susceptibility as a function of the temperature. For high enough temperatures the Curie-Weiss law is valid. In the left panel we present experimental results for Ni (Weiss and Forrer Ann. Phys. Vol. 5, 153 (1926)). On the right panel we plot the reciprocal susceptibility parallel to the crystal c-axis as a function of temperature for LiHoF_4 . The experimental points correspond to measurements for five different sample shapes, characterized by demagnetizing factors, N (A.H. Cooke, D.A. Jones, J.F.A. Silva and M.R. Wells, J. Phys. C **8**, 4083 (1975))

In the vicinity of the critical temperature, however the experimentally obtained critical exponent for the susceptibility of ferromagnetic materials is $\gamma \approx 1.3$. As an example see the results obtained on EuO presented on Figure 3.

5.2 Spontaneous magnetization in the $T \rightarrow T_c(T < T_c)$ limit.

Mean-field theory can also give us the behavior of the $M(T)$ curves in the vicinity of T_c for $H=0$. In order to get this we use the fact that $\langle S \rangle \rightarrow 0$ in this limit. We work again with classical vector-spins and use the expansion of the Langevin function in the $f \rightarrow 0$ limit. In this expansion we keep now the first two terms and set $H=0$. We get:

$$\langle S \rangle \approx S \left(\frac{f}{3} - \frac{f^3}{45} \right) \quad (61)$$

By using the $T_c = \frac{jqS^2}{3k}$ value, with some straightforward math we get:

$$\langle S \rangle^2 = \frac{5}{3} \frac{T^2}{T_c^2} \left(1 - \frac{T}{T_c} \right) \quad (62)$$

which gives for the total spontaneous magnetization in the limit $T \rightarrow T_c(T < T_c)$ of the sample:

$$M \approx N\mu_B g S \sqrt{\frac{5}{3}} \frac{T}{T_c} \left(1 - \frac{T}{T_c} \right)^{\frac{1}{2}} \quad (63)$$

This results predicts the critical exponent $\beta = 1/2$, in comparison with experiments that would yield values around 0.36 (see for example the experimental results presented in Figure 2). This discrepancy is due to the used mean-field approximation. Better approximation methods (low temperature expansion, renormalization or spin-wave theory) would yield results in good agreement with experiments.

6. Mean-field theory of antiferromagnets

The mean-field theory can be developed in the same analogous manner for antiferromagnets as well. For showing this, let us consider a simple antiferromagnetic structure where two topologically equivalent sub-lattices co-exist, and the antiferromagnetic exchange interaction only between the two sub-lattices. In two-dimensions the situation would be the one shown in Figure 9:

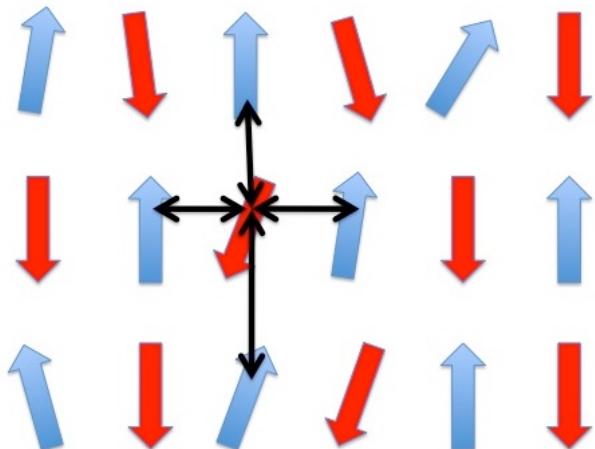


Figure 9. Illustration of the considered antiferromagnetic interaction topology.

In this simple case in the mean-field approximation we consider the two sub-lattice of spins separately (the blue spins are supposed to be on sub-lattice A, and the red spins are supposed to be on sub-lattice B). In such manner the mean-field Hamiltonian of the system will write as:

$$E = -j \sum_{i \in A} \vec{S}_i \left(q \langle \vec{S}_B \rangle \right) - j \sum_{i \in B} \vec{S}_i \left(q \langle \vec{S}_A \rangle \right) - \mu_0 \mu_B g \vec{H} \sum_{i \in A} \vec{S}_i - \mu_0 \mu_B g \vec{H} \sum_{i \in B} \vec{S}_i = \sum_i (E_i^A + E_i^B) \quad (64)$$

with $j < 0$ and :

$$\begin{aligned} E_i^A &= \left[-j \left(q \langle \vec{S}_B \rangle \right) - \mu_0 \mu_B g \vec{H} \right] \vec{S}_i \\ E_i^B &= \left[-j \left(q \langle \vec{S}_A \rangle \right) - \mu_0 \mu_B g \vec{H} \right] \vec{S}_i \end{aligned} \quad (65)$$

For solution we assume now that again the z axis is chosen in the direction of the external magnetic field. Consequently the average spin orientations are also in the direction of the z axis, so we work thus only in modulus and assume $\langle S_A \rangle = -\langle S_B \rangle$.

The self-consistent equations for $\langle S_A \rangle$ and $\langle S_B \rangle$ are similar with the ones obtained for ferromagnets (we write them again for the case of classical vector-spins). We get a system of transcendental equations for $\langle S_A \rangle$ and $\langle S_B \rangle$:

$$\begin{aligned} \frac{\langle S_A \rangle}{S} &= L \left[\beta S \left(jq \langle S_B \rangle + \mu_0 \mu_B g H \right) \right] \\ \frac{\langle S_B \rangle}{S} &= L \left[\beta S \left(jq \langle S_A \rangle + \mu_0 \mu_B g H \right) \right] \end{aligned} \quad (66)$$

Assuming $\langle S_A \rangle = -\langle S_B \rangle$ and $H=0$ we have the same graphical solution for both of them, as in the case of ferromagnets. We obtain a critical temperature, named as Neel temperature:

$$T_N = \frac{|jqS^2|}{3k} \quad (67)$$

In case of quantized z components of the S spins one would get: $T_N = \frac{|jqS(S+1)|}{3k}$. For $T < T_c$ the system orders anti-ferromagnetically and for $T > T_c$ it behaves as a paramagnet.

The behavior of the susceptibility has however a different form from the one obtained in ferromagnetic systems due to the two oppositely oriented sub-lattice. In the limit $T > T_c$ and $H \rightarrow 0$ one can make the same expansion for the Langevin function, and keeping only the first term we get:

$$\begin{aligned} \frac{\langle S_A \rangle}{S} &= \frac{\beta S}{3} \left(jq \langle S_B \rangle + \mu_0 \mu_B g H \right) \\ \frac{\langle S_B \rangle}{S} &= \frac{\beta S}{3} \left(jq \langle S_A \rangle + \mu_0 \mu_B g H \right) \end{aligned} \quad (68)$$

The total magnetization of the system writes now as: $M = \frac{N}{2} g \mu_B (\langle S_A \rangle - \langle S_B \rangle)$ and can be obtained immediately by subtracting equations (68) :

$$M = N \frac{S^2 \mu_0 \mu_B^2 g^2}{3k} \frac{1}{T + T_N} H, \quad (69)$$

leading to the susceptibility: $\chi = N \frac{S^2 \mu_0 \mu_B^2 g^2}{3k} \frac{1}{T + T_N}$ (70)

7. Mean-field theory for ferrimagnets

The simplest view on ferrimagnets is as an uncompensated antiferromagnet. Let us consider again a very simple case when two equivalent sub-lattices are present, and the only difference is that the spins on the two sub-lattice have different values (Figure 10) and the exchange interaction between the components can be also different ($j_{AB} \neq j_{BA}$).

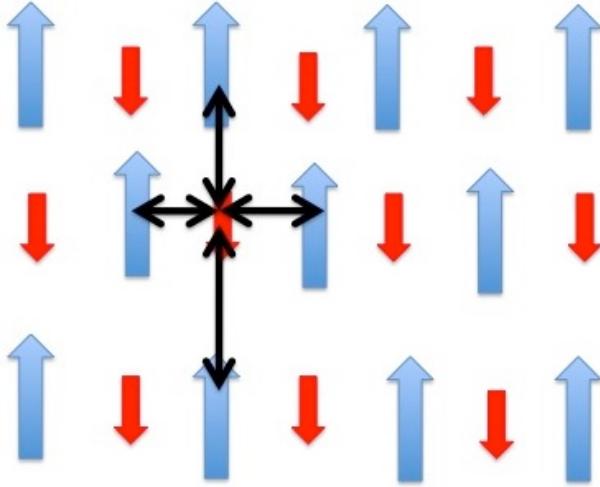


Figure 10. Illustration of the considered ferrimagnetic interaction topology.

In this case we can write up the same equations as in the case of antiferromagnets taking into account that now $\langle S_A \rangle \neq \langle S_B \rangle$

$$\begin{aligned} \frac{\langle S_A \rangle}{S_A} &= L \left[\beta S_A (j_{AB} q \langle S_B \rangle + \mu_0 \mu_B g H) \right] \\ \frac{\langle S_B \rangle}{S_B} &= L \left[\beta S_B (j_{BA} q \langle S_A \rangle + \mu_0 \mu_B g H) \right] \end{aligned} \quad (71)$$

Solving the system of equations (71) for $\langle S_A \rangle$ and $\langle S_B \rangle$, we can compute

$$M(T, H) = \frac{N}{2} (\langle S_A \rangle - \langle S_B \rangle) \mu_B g, \quad (72)$$

and depending on the S_A, S_B, j_{AB}, j_{BA} values very different behavior for the spontaneous magnetization and susceptibility is observed. In Figure 11 We sketch the possible curves one might observe:

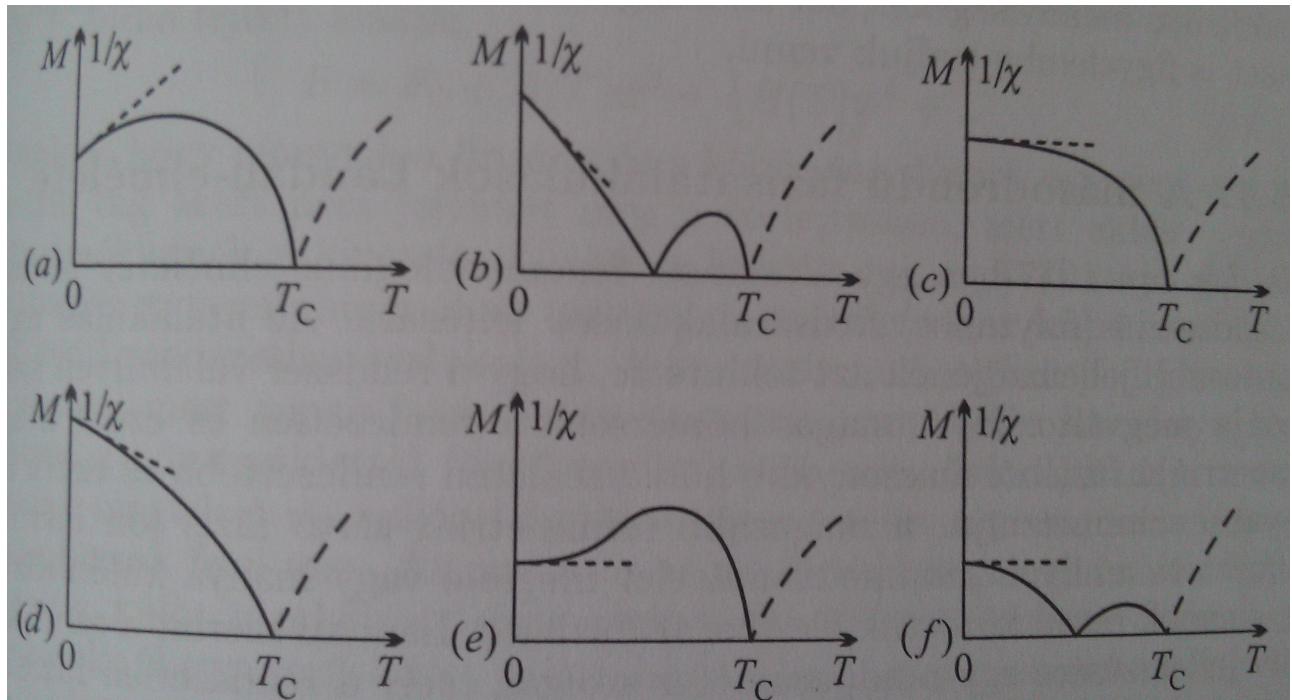


Figure 11: Temperature dependence of the spontaneous magnetization and inverse susceptibility for various choices of the S_A, S_B, j_{AB}, j_{BA} parameters.

8. General case: Interactions with more neighbors and the case of more sub-lattices

Similarly with the antiferromagnetic and ferromagnetic systems discussed above, one could consider system with local interactions, but with more distant neighbors and more sub-lattices with different spin values. The mean-field approach for such systems would be very similar. If the sub-lattices are denoted as $X_w \rightarrow w = 1, 2, \dots, K$ one can write in the mean-field approximation:

$$\frac{\langle S_{X_w} \rangle}{S_{X_w}} = L \left[\beta S_A \left(\sum_r j_{rw} q_{rw} \langle S_{X_r} \rangle + \mu_0 \mu_B g H \right) \right] \quad (73)$$

where $r, w \in \{1, 2, \dots, K\}$. This leads to a system of coupled K equations that can be solved numerically.

9. Mean-field approach to itinerant magnetism

In case of most of the metals that exhibit ferromagnetic properties (Fe, Ni...etc), magnetic moments are due to electrons in partially filled bands. This is the 3d band of transition metals. The wave-functions of these electrons overlap allowing the jump (movement) of the 3d electrons from one site to another one. Itinerant magnetism in these systems is due to the competition between screened Coulomb interaction and kinetic energy.

The Hubbard model offers the simplest possible description for such a system, and a simple way to understand insulating, magnetic and even superconducting states in solids. It was proposed in the early 1960's, and it is the amazing feature of this model that despite its simplicity, it explains

qualitatively well many subtle properties of solid-state systems. Its exact solution is possible only in some very limited cases, but it has been studied by many approximation methods of statistical physics ranging from simple mean-field approach to field-theoretical methods. It is also extensively studied by numerical methods such as numerical diagonalization or quantum Monte Carlo methods.

The model takes into account that there is a regular array of nuclear positions in a solid, so it is defined on a lattice. It is based on a lattice of atoms on which the electrons can move. A single atom being a very complex structure, with many different energy levels would already lead to a complicated model. The simplest possible atom one would imagine would have only one energy level. Due to the Pauli exclusion principle in such case only two electrons with oppositely oriented spins are allowed on one site. This is what the Hubbard model considers. Electrons can also jump on nearby sites, providing that the Pauli exclusion principle is not violated.

In a real solid where electrons can freely move the electrons are interacting via Coulomb potential. This interaction is however strongly screened, and can be approximated by a form:

$$V(r) = -\frac{e^2 \exp(-qr)}{r} \quad (74)$$

Due to this screening the Coulomb interaction is relevant only for atoms on the same site. For simplicity reasons in the Hubbard model we consider only interactions between electrons on the same site, and neglect Coulomb interactions between electrons on neighboring sites.

The Hubbard Hamiltonian is written up in a second-quantized form, using fermionic creation, annihilation and number operators. It contains a kinetic energy and an interaction term. It considers the interacting electron system in a quantum-statistical model. In the absence of a magnetic field the Hamiltonian operator of the system writes as:

$$\hat{H} = -t \sum_{\langle j,l \rangle \sigma} \hat{c}_{j\sigma}^\dagger \hat{c}_{l\sigma} + U \sum_j \hat{n}_{j\uparrow} \hat{n}_{j\downarrow} \quad (75)$$

The first term is the kinetic energy term which governs the hopping of the electrons from one site i to a neighboring j site. The energy scale t governs the barrier needed to overcome for such a jump. The operators \hat{c}_i^\dagger and \hat{c}_i are creation and annihilation operators for the site i . In the second quantized description the state of the quantum-mechanical system is given by the occupation numbers on the lattice sites. If there are N lattice sites available a state of the system can be written as:

$$|\psi\rangle = |n_{1\uparrow}, n_{1\downarrow}, n_{2\uparrow}, n_{2\downarrow} \dots n_{N\uparrow}, n_{N\downarrow}\rangle \quad (76)$$

For fermionic systems $n_{i\sigma} \in \{0,1\}$. A system with no fermions on any site would be described thus by: $|0,0,\dots,0\rangle$, a state with one fermion on lattice site 1 and spin up would have the state vector: $|1,0,\dots,0\rangle$. The action of the creation and annihilation operators of the states are the following:

$$\begin{aligned} \hat{c}_{1\uparrow}^\dagger |0,0,\dots,0\rangle &= |1,0,\dots,0\rangle \\ \hat{c}_{1\uparrow}^\dagger |0,0,\dots,0\rangle &= 0 \\ \hat{c}_{1\downarrow}^\dagger |1,0,\dots,0\rangle &= 0 \\ \hat{c}_{1\downarrow}^\dagger |0,0,\dots,0\rangle &= |0,1,\dots,0\rangle \\ \hat{c}_{1\downarrow}^\dagger |0,1,\dots,0\rangle &= |0,0,\dots,0\rangle \end{aligned} \quad (77)$$

The fermionic creation and annihilation operators satisfy the following anticommuting relations:

$$\begin{aligned} \{\hat{c}_{i\sigma}, \hat{c}_{j\sigma'}\} &\equiv \hat{c}_{i\sigma}\hat{c}_{j\sigma'} + \hat{c}_{j\sigma'}\hat{c}_{i\sigma} = 0 \\ \{\hat{c}_{i\sigma}^+, \hat{c}_{j\sigma'}^+\} &= 0 \\ \{\hat{c}_{i\sigma}, \hat{c}_{j\sigma'}^+\} &= \delta_{ij}\delta_{\sigma\sigma'} \end{aligned} \quad (78)$$

The creation and annihilation operator for the same site applied one after the other, acts as a number operator:

$$\hat{c}_{i\sigma}^+ \hat{c}_{i\sigma} |n_{1\uparrow}, n_{1\downarrow}, n_{2\uparrow}, n_{2\downarrow}, \dots, n_{i\uparrow}, n_{i\downarrow}, \dots, n_{N\uparrow}, n_{N\downarrow}\rangle = n_{i\sigma} |n_{1\uparrow}, n_{1\downarrow}, n_{2\uparrow}, n_{2\downarrow}, \dots, n_{i\uparrow}, n_{i\downarrow}, \dots, n_{N\uparrow}, n_{N\downarrow}\rangle \quad (79)$$

We will study the problem in a grand canonical ensemble, fixing the (T, V, μ) values.

The problem is solved if the grand-partition function is calculated:

$$\Omega = \text{Tr} [e^{-\beta(\hat{H} - \mu N)}] \quad (80)$$

with $\beta = \frac{1}{kT}$ and μ the chemical potential. Thermal expectation value of a quantity A will be computed as:

$$\langle A \rangle = \frac{1}{Z} \text{Tr} [\hat{A} e^{-\beta(\hat{H} - \mu N)}] \quad (81)$$

For example, if one considers a single spinless fermionic state with energy: ε , $\hat{H} = \varepsilon \hat{c}^+ \hat{c} = \varepsilon \hat{n}$, the allowed states are $|0\rangle$ or $|1\rangle$. The grand partition function is then:

$$\Omega = \left\langle 0 \middle| e^{-\beta(\hat{H} - \mu \hat{n})} \right| 0 \rangle + \left\langle 1 \middle| e^{-\beta(\hat{H} - \mu \hat{n})} \right| 1 \rangle = 1 + e^{-(\beta - \mu)} \quad (82)$$

The average occupation number of such a state is given then as:

$$\langle n \rangle = \frac{1}{\Omega} \left(\left\langle 0 \middle| \hat{n} e^{-\beta(\hat{H} - \mu \hat{n})} \right| 0 \rangle + \left\langle 1 \middle| \hat{n} e^{-\beta(\hat{H} - \mu \hat{n})} \right| 1 \rangle \right) = \frac{1}{\Omega} [0 + e^{-(\beta - \mu)}] = \frac{1}{e^{\beta(\varepsilon - \mu)} + 1} \quad (83)$$

which is the Fermi-Dirac distribution....

The notations and methods being clarified, let us study now the case of the Hubbard model. We consider first the problem without the hopping term ($t=0$), which implies independent sites. The sites being independent we can calculate the grand-partition function for each of them. The Hamiltonian on one site writes as:

$$\hat{H} = U \hat{n}_\uparrow \hat{n}_\downarrow \quad (84)$$

the possible states for the considered site are: $|0\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle$. The grand partition function writes as:

$$\Omega = \sum_{\eta} \left\langle \eta \middle| e^{-\beta(\hat{H} - \mu \hat{n})} \right| \eta \rangle = 1 + e^{\beta\mu} + e^{\beta\mu} + e^{-\beta U + 2\beta\mu} \quad (85)$$

and the average particle number on one site:

$$\begin{aligned} \langle n \rangle &= \frac{1}{\Omega} \sum_{\eta} \left\langle \eta \middle| (\hat{n}_\uparrow + \hat{n}_\downarrow) e^{-\beta(\hat{H} - \mu \hat{n})} \right| \eta \rangle = \frac{1}{\Omega} [0 + e^{\beta\mu} + e^{\beta\mu} + 2e^{-\beta U + 2\beta\mu}] = \\ &= \frac{2(e^{\beta\mu} + e^{-\beta U + 2\beta\mu})}{1 + 2e^{\beta\mu} + e^{-\beta U + 2\beta\mu}} \end{aligned} \quad (86)$$

For different temperatures and $U=1$ value the $\langle n \rangle(\mu)$ curves are plotted on Figure 12. From here we learn, that there is a gap in the chemical potential of value U to go from the half-filling state ($n=1$) to the total filling ($n=2$). Since the chemical potential is related to the energy per particle

$\mu = \left(\frac{\partial U}{\partial N} \right)_{V,T}$, this means that the energy necessary to add a particle at half-filling is of the order U .

This is called the Hubbard gap, and for the Hubbard model with half-filling and $t \ll U$ is responsible for the insulator state of the model.

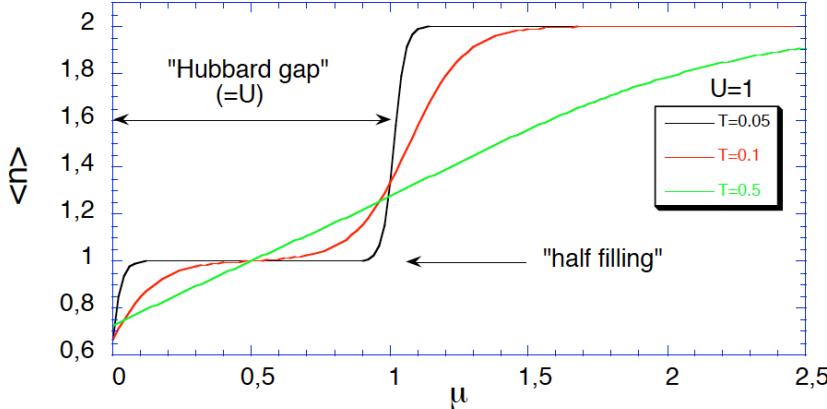


Figure 12. The average occupancy of the sites in $t=0$ Hubbard model with $U=1$ (various temperatures)

Another possibility is to consider the non-interacting limit ($U=0$), where only the hopping term exists. Let us assume that we have a 1D lattice with periodic boundary conditions (sketched in Figure 13):

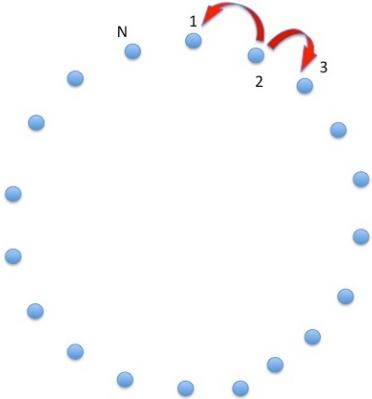


Figure 13. The considered 1D lattice topology and possible hopping from site 2.

In this case we transform the problem in reciprocal Fourier space, by introducing the $\hat{c}_{k\sigma}$ annihilation and $\hat{c}_{k\sigma}^+$ creation operators:

$$\begin{aligned}\hat{c}_{k\sigma}^+ &= \frac{1}{\sqrt{N}} \sum_{l=1}^N e^{ikl} \hat{c}_{l\sigma}^+ \\ \hat{c}_{k\sigma} &= \frac{1}{\sqrt{N}} \sum_{l=1}^N e^{-ikl} \hat{c}_{l\sigma}\end{aligned}\quad (87)$$

The summation in equations (87) is going for $k_n = \frac{2\pi n}{N}$ ($n=0, \dots, N-1$). The Hamiltonian in the Fourier space takes a simple form:

$$\hat{H} = -t \sum_{\langle j,l \rangle \sigma} \hat{c}_{j\sigma}^+ \hat{c}_{l\sigma} = -\frac{t}{N} \sum_{k,k'} \sum_{\langle j,l \rangle \sigma} e^{ikj} e^{-ik'l} \hat{c}_{k\sigma}^+ \hat{c}_{k'\sigma} \quad (88)$$

For the 1d chain one gets:

$$\begin{aligned}\hat{H} &= -\frac{t}{N} \sum_{k,k',\sigma} \hat{c}_{k\sigma}^+ \hat{c}_{k'\sigma} \sum_{j\sigma} (e^{ikj} e^{-ik'(j+1)} + e^{ikj} e^{-ik'(j-1)}) = -\frac{t}{N} \sum_{k,k',\sigma} \hat{c}_{k\sigma}^+ \hat{c}_{k'\sigma} [e^{-ik'} + e^{-ik'}] \sum_j e^{i(k-k')j} = \\ &= -t \sum_{k,k',\sigma} \hat{c}_{k\sigma}^+ \hat{c}_{k'\sigma} (2 \cos(k')) \delta_{k,k'} = \sum_{k,\sigma} \varepsilon_k \hat{n}_{k,\sigma}\end{aligned}\quad (89)$$

with

$$\varepsilon_k = -2t \cos(k) \quad (90)$$

The eigenvalues of the Hubbard Hamiltonian are thus obtained by the Fourier transform, there are given by: (87) , where the k value can be: $k_n = \frac{2\pi n}{N}$ ($n = 0, 1, 2, \dots, N$). The obtained spectrum is plotted on Figure 14:

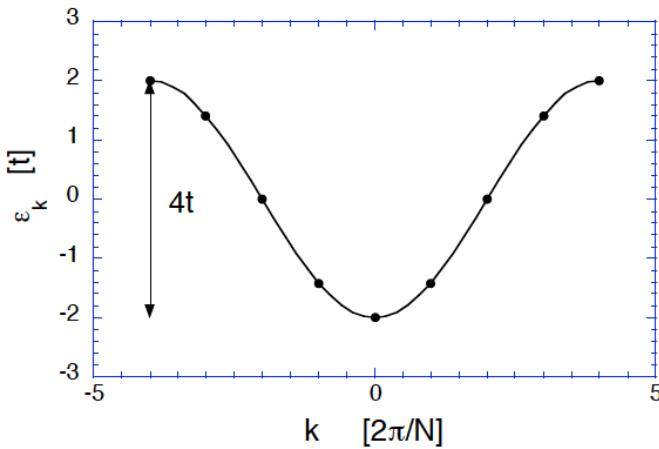


Figure 14 . Energy spectrum of the 1D Hubbard model with $U=0$.

In the thermodynamic limit of $N \rightarrow \infty$ the spectrum becomes continuous, so the $U=0$ Hubbard model describes a conducting state. It is believed thus that for $t \gg U$ the Hubbard model describes a conducting state.

Since for ($t \ll U$) we have an insulating state and for ($t \gg U$) we have a conducting state, it is natural to believe that the Hubbard model exhibits a conductor-isolator transition, which is called as the Mott transition.

9.1 Magnetism of the ground-state in the 1D half-filled Hubbard model (mean-field description)

Let us consider now a mean-field approach for the $T=0$ case when both $t \neq 0$ and $U \neq 0$. The problem with an exact calculation of the ground state is that the Hamiltonian is non-diagonal in the direct and Fourier space. Calculation of the stationary states (energy spectrum) is thus complicating. Diagonalizing the Hamiltonian is a heavy task, so we will apply the mean-field approximation, which leads us to a diagonalized Hamiltonian. We start with the Hubbard model:

$$\hat{H} = -t \sum_{\langle j,l \rangle \sigma} \hat{c}_{j\sigma}^+ \hat{c}_{l\sigma} + U \sum_j \hat{n}_{j\uparrow} \hat{n}_{j\downarrow} \quad (91)$$

The problem is the second term which contains four creation/annihilation operators, and the mixing of the creation and annihilation operators for different sites in the first term. For a one-dimensional

chain the Hamiltonian would be:

$$\hat{H} = -t \sum_{j,\sigma} (\hat{c}_{j,\sigma}^+ \hat{c}_{j+1,\sigma} + \hat{c}_{j,\sigma}^+ \hat{c}_{j-1,\sigma}) + U \sum_j \hat{n}_{j\uparrow} \hat{n}_{j\downarrow} \quad (92)$$

The approximation we will consider is very similar with the one made in the case of localized spins, namely we neglect the second order fluctuation terms. We write:

$$n_{l\uparrow} = \langle n_{l\uparrow} \rangle + [n_{l\uparrow} - \langle n_{l\uparrow} \rangle] = \langle n_{l\uparrow} \rangle + \delta_{l\uparrow} \quad (93)$$

where the averages denote the average occupation number per site. In the second, interaction term we then neglect the correlation term:

$$\delta_{l\uparrow} \delta_{l\downarrow}$$

The last term becomes thus:

$$n_{l\uparrow} n_{l\downarrow} \approx n_{l\uparrow} \langle n_{l\downarrow} \rangle + n_{l\downarrow} \langle n_{l\uparrow} \rangle - \langle n_{l\uparrow} \rangle \langle n_{l\downarrow} \rangle \quad (94)$$

With this approximation the Hamiltonian can be diagonalized in the reciprocal space in analogous manner as we did for the case $U=0$. In the reciprocal space one would get:

$$\begin{aligned} \hat{H} &= \sum_k \left[-2t \cos(k) (\hat{n}_{k\uparrow} + \hat{n}_{k\downarrow}) + U (\hat{n}_{k\uparrow} \langle n_{k\downarrow} \rangle + \hat{n}_{k\downarrow} \langle n_{k\uparrow} \rangle) \right] - UN \langle n_{\uparrow} \rangle \langle n_{\downarrow} \rangle = \\ &= \sum_k (\epsilon_{k\uparrow} \hat{n}_{k\uparrow} + \epsilon_{k\downarrow} \hat{n}_{k\downarrow}) - UN \langle n_{\uparrow} \rangle \langle n_{\downarrow} \rangle \end{aligned} \quad (95)$$

with:

$$\epsilon_{k,\uparrow\downarrow} = -2t \cos(k) + U \langle n_{\downarrow\uparrow} \rangle \quad (96)$$

This is now in a diagonal term and one can calculate the ground-state energy, assuming that all states are occupied up to a “Fermi level”, given by:

$$k_{F\uparrow\downarrow} = \frac{\pi \langle n_{\uparrow\downarrow} \rangle}{N} \quad (97)$$

We replace the summation by an integral:

$$\sum_{k,\uparrow\downarrow} \dots \rightarrow \frac{N}{\pi} \int_0^{k_{F\uparrow\downarrow}} \dots dk \quad (98)$$

Performing these integral one gets at $T=0$:

$$E_{\text{ground}} = -\frac{2Nt}{\pi} [\sin(\pi \langle n_{\uparrow} \rangle) + \sin(\pi \langle n_{\downarrow} \rangle)] + U \langle n_{\downarrow} \rangle \langle n_{\uparrow} \rangle \quad (99)$$

Now let us consider the case of half filling, which means:

$$\langle n_{\uparrow} \rangle + \langle n_{\downarrow} \rangle = 1 \rightarrow \langle n_{\downarrow} \rangle = 1 - \langle n_{\uparrow} \rangle \quad (100)$$

In such case we obtain:

$$\frac{\pi E_{\text{ground}}}{4Nt} = -\sin(\pi \langle n_{\uparrow} \rangle) + \frac{U\pi}{4t} \langle n_{\uparrow} \rangle (1 - \langle n_{\uparrow} \rangle) \quad (101)$$

We can study thus the shape of the

$$E_{\text{ground}}(\langle n_{\uparrow} \rangle)$$

curves, and identify the position of the minima as a function of $\langle n_{\uparrow} \rangle$ for different values of the parameter $p = \frac{U\pi}{4t}$ (Figure 15)

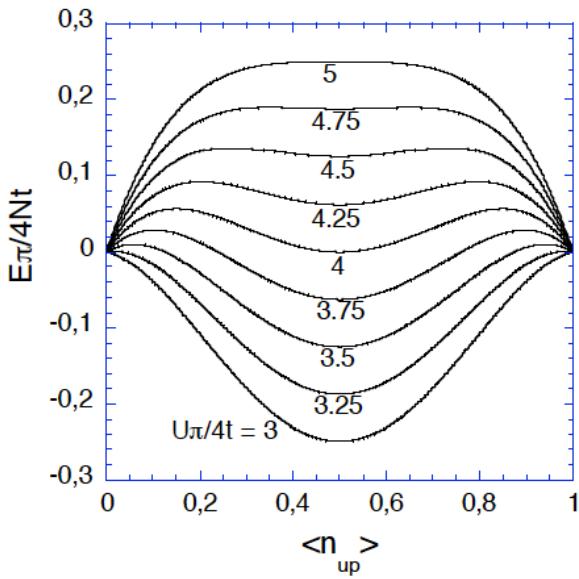


Figure 15. The ground-state energy of the 1D Hubbard model at half filling in the mean-field approximation.

Figure 15 suggests that the ground state of this system can be both ferromagnetic and non-magnetic (paramagnetic or antiferromagnetic) depending on the value of p . For small U values we get a non-magnetic ground-state since the minima of the curves are at :

$$\langle n_{\uparrow} \rangle = 1/2$$

On the contrary, when U is large we get the minimum at:

$$\langle n_{\uparrow} \rangle = \begin{cases} 0 \\ 1 \end{cases}$$

suggesting a ferromagnetic ordering. Mean-field theory predicts thus a ferromagnetic ground-state for the 1D Hubbard model at half filling for strong coupling (large U). Similarly with many other mean-field results in 1D this result is however also wrong! An exact solution base on the Bethe anstaz given by Lieb and Wu (PRL , vol.20, 1445 (1998) proves that the ground is antiferromagnetic.

10. Conclusions

Studying two and three-dimensional locally interacting models is a hard problem in statistical physics. Usually no exact calculation of the partition function is possible, one has to consider thus approximations. For the elementary models of magnetic ordering (Heisenberg, Ising, Hubbard, etc...) this is also the case. The simplest attempt for an analytical solution is the mean-field approximation, where fluctuations are assumed small, and correlation effects are neglected. This simple approximation is useful to get a hint about the qualitative behavior of the model, however the results are usually trustable only for high dimensional systems. (The results become exact for infinite dimensional systems). For low dimensional systems one has to be very careful with interpreting the results of the mean-field approximations. According to the Mermin-Wagner-Hohenberg theorem the continuous symmetry cannot be spontaneously broken at finite temperature in systems with short-range interactions in dimensions $d \leq 2$. Phase-transitions obtained at low dimensional systems ($d=1$) for locally interacting systems by mean-field approach, are an artifact. Thus although the mean-field approach is a very elegant way of studying models of magnetic

ordering, the results has to be interpreted critically, and other approximations should be used, to get a hint about the validity of the obtained results. For high dimensional systems, mean-field results are however valuable and gives a first description of the systems behavior.