

## Mean Field Approximation (MFA) in the theory of magnetism.

MFA was constructed by P. Weiss and P. Curie for explaining the behavior of ferromagnetic materials. Such systems behave like paramagnets at high temperatures, but below certain critical temperature  $T_c$  they exhibit “spontaneous magnetization” — i.e., a sizable macroscopic magnetic moment even in the absence of an external magnetic field.

In its original version MFA was a phenomenological theory. Microscopic justification for the model postulates was added later, when the nature of interactions between individual magnetic atoms (Heisenberg exchange) became understood on the grounds of quantum mechanics.

MFA is often regarded as a quite “primitive” model. In fact, some of its predictions — especially, concerning the system behavior at  $T \rightarrow 0$  and in the region close to  $T_c$  (the so-called “critical region”) — appear to be in “less than perfect” agreement with the experiment. It would be too much to say that the MFA results are *completely wrong* in these two regions; however, they definitely cannot be used for fine-tuned quantitative data interpretation. The main reason for the poor model performance for  $T \rightarrow 0$  is that simple averaging is not a good tool for describing properties associated with propagating excitations, especially in the case of long wavelengths — and it is the long- $\lambda$  propagating spin wave modes that play the principal role in the ferromagnet behavior at the lowest temperature region<sup>1</sup>. The same applies to fluctuations: averaging, for obvious reasons, “smoothens out” any fluctuations. Consequently, a method based on averaging cannot describe absolutely correctly any phase transition phenomena — because, as we know, fluctuations play a crucial role in them.

Models that produce a much better agreement with experimental observations in those two regions have been developed. However, they were never able to fully remove the MFA from the stage because they all lack “comprehensiveness” — none of them works in BOTH regions. MFA, in contrast, offers solutions for the entire region between  $T \rightarrow 0$  and  $T \rightarrow \infty$ . So, if used with caution, this simple model may still offer much help in understanding the properties of magnetic materials.

From the general viewpoint of statistical physics MFA represents a simple approach to systems consisting of strongly interacting particles. Historically, it has been developed for studying magnets — however, the same approach can be used for modelling a number of other physical effects not involving magnetism.

MFA is based on a few assumptions, some of them we have already discussed.

**1. Ideal paramagnet.** It consists of non-interacting magnetic atoms, each of which has a magnetic moment:

$$\vec{\mu} = -g\mu_B\vec{J} \quad (1)$$

where  $\vec{J}$  is the total angular momentum,  $\mu_B$  is the Bohr magneton, and  $g$ , the Lande factor, is given by:

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L-1)}{2J(J+1)} \quad (2)$$

where  $S$ ,  $L$  and  $J$  are the spin, orbital momentum, and total angular momentum numbers, respectively. The magnetization  $M$  (i.e., the magnetic moment per unit volume) of an ideal paramagnet,

---

<sup>1</sup>Note that the Einstein’s model of solid is neither very successful for  $T \rightarrow 0$  — it’s for analogous reasons, because the model ignores the propagating nature of excitations in crystals.

induced by an external magnetic field  $H$  is (see Notes, page 22 ):

$$M = NgJ\mu_B B_J(x), \quad \text{with} \quad x \equiv \frac{gJ\mu_B\mu_0 H}{kT} \quad (3)$$

where  $N$  is the number of atoms per unit volume, and  $B_J(x)$  is the Brillouin function:

$$B_J(x) = \frac{2J+1}{2J} \coth \left[ \frac{(2J+1)x}{2J} \right] - \frac{1}{2J} \coth \left( \frac{x}{2J} \right) \quad (4)$$

The fact that  $B$  is the conventionally used symbol for the Brillouin function forces us to use  $H$ , not  $B$ , for the magnetic field. Also, it is not very convenient that  $J$  appears in the Brillouin function because  $J$  is also the conventionally used symbol for the exchange parameter. We will start using the exchange parameter in our math pretty soon, and the exchange parameter and the total angular momentum number will both appear in the same equations. In order to avoid confusion, one  $J$  has to be renamed. The simplest way seems to be starting using  $S$  in the Brillouin function instead of  $J$ :

$$B_S(x) = \frac{2S+1}{2S} \coth \left[ \frac{(2S+1)x}{2S} \right] - \frac{1}{2S} \coth \left( \frac{x}{2S} \right) \quad (5)$$

It makes much sense, because in many magnetic atoms the orbital momentum is zero, so  $J = S$ . Hence, one can safely use the  $S$  symbol, only keeping in mind that for certain atoms  $J \neq S$  and in such cases  $S$  in the final formulae should be replaced by the total angular momentum number encoded by a symbol clearly distinguishable from the exchange parameter  $J$

One of the principal objects of interest in the MFA modelling of magnetics is the magnetic susceptibility  $\chi = dM/dH$ . Since  $M(H)$ , in general, is not a linear function, the value of  $\chi$  is  $H$ -dependent. But if we do not specify  $H$ , it means that we are talking about  $\chi$  in the region where  $M(H)$  is approximately linear – i.e., in the “low field region”. Then, indeed, one can assume that  $\chi \neq \chi(H)$ .

Taking the series expansion of the  $\coth x$  function:

$$\coth x = \frac{1}{x} + \frac{x}{3} - \frac{x^3}{45} + \dots \quad (6)$$

and applying it to Eq. (5), we obtain, after dropping  $x^3/45$  and higher-order terms and some uncomplicated algebra:

$$B_S(x) \cong \frac{S+1}{3S} x \quad (7)$$

Combining this with the  $x$  definition in Eq. (3), we get:

$$B_S(x) = \frac{NS(S+1)g^2\mu_B^2}{3kT} \mu_0 H \quad (8)$$

which is valid for  $x \ll 1$ , or  $\mu_0 H \ll kT/g\mu_B S$  (one can also write  $\mu_0 H \ll kT/\mu_B$ , because the value of  $gS$  is always  $\geq 1$ , but it is rarely larger than 10 — rarely, as it happens only for some atoms belonging to the Rare Earth Metals family). In this region  $\chi$  is field-independent:

$$\chi = \frac{C}{T} \quad \text{where} \quad C = \frac{NS(S+1)g^2\mu_B^2\mu_0}{3k} \quad (9)$$

This relation expresses the Curie Law for paramagnets, and  $C$  is the Curie-Weiss constant.

**2. The Curie-Weiss Law.** Weiss based his model on the empirical observations that in the region above the critical temperature  $T_c$  a ferromagnet behaves very much as a paramagnet, except that its magnetic susceptibility does not obey the  $\chi \propto 1/T$  law, but rather:

$$\chi \propto \frac{1}{T + T_c}. \quad (10)$$

In order to explain this behavior, Weiss assumed that the field “seen” by an average atom in a ferromagnet is the sum of the external field  $\vec{H}$  and some sort of “internal” or “effective” field  $\vec{H}_E$  that he also called the “molecular field”. One crucial point in this assumption was that *the effective field is proportional to the magnetization*:

$$\vec{H}_E = \lambda \vec{M} \quad (11)$$

where  $\lambda$  is the “effective field constant”.

Next, he reasoned that if the magnetization of a paramagnet depends on  $\vec{H}$  as:

$$\vec{M} = \chi \vec{H} = \frac{C}{T} \vec{H}, \quad (12)$$

then, by replacing the external field by a sum of the external field and the effective field, one should obtain the magnetization law for a ferromagnet:

$$\vec{M} = \frac{C}{T} (\vec{H} + \vec{H}_E) = \frac{C}{T} (\vec{H} + \lambda \vec{M}). \quad (13)$$

By solving that for  $M$ , Weiss obtained the magnetization formula:

$$M = \frac{CH}{T - C\lambda} \quad (14)$$

(since  $M$  and  $H$  are parallel vectors, we dropped the vector symbols), and the susceptibility law (commonly, referred to as the Curie-Weiss Law):

$$\chi = \frac{C}{T - C\lambda}. \quad (15)$$

By comparing that with the empirical formula given in Eq. (10), one can write:

$$\chi = \frac{C}{T - T_c} \quad \text{with} \quad T_c = C\lambda. \quad (16)$$

Using Eq.(9), we get the following formula for the “effective field constant”  $\lambda$ :

$$\lambda = \frac{T_c}{C} = \frac{3kT_c}{Ng^2\mu_B^2\mu_0S(S+1)}. \quad (17)$$

**3. Microscopic justification for the Weiss' effective field.** Of course, in his considerations Weiss could get only as far as to Eqs. (15) and (16), but not to Eq. (17). In 1907, he could not know anything yet about the Brillouin function. His theory was phenomenological because he did not have enough information for explaining the microscopic nature of the effective field. This could be done only years later, after Heisenberg formulated the famous law of interaction between two magnetic atoms with spins  $\vec{S}_1, \vec{S}_2$ :

$$\mathcal{H} = -2J\vec{S}_1 \cdot \vec{S}_2 \quad (18)$$

where  $J$  is the exchange constant.

Let's consider an individual atom  $i$  with magnetic moment  $\vec{\mu}_i$  in a ferromagnetic crystal. The moment interacts with the external magnetic field and with the Weiss' effective field, so that the total interaction energy is:

$$E_i = -\vec{\mu}_i \cdot (\mu_0\vec{H} + \mu_0\vec{H}_{\text{eff}}), \quad (19)$$

From the viewpoint of the Heisenberg model, however, the energy  $E_i$  is the sum of interaction energy with the external field, and of the exchange interactions with all relevant neighbors to atom  $i$ . Let's assume that each atom has  $z$  such "relevant" neighbors. Also, for simplicity, let's assume that all those  $z$  neighbors interact with spin  $i$  with the same force, i.e., that exchange parameter has the same value for all  $z$  neighbors. Then, the energy  $E_i$  can be written as:

$$E_i = -\vec{\mu}_i \cdot (\mu_0\vec{H}) - \sum_k^z 2J\vec{S}_i \cdot \vec{S}_k, \quad (20)$$

where the index  $k$  in the sum runs over all  $z$  neighbors of the atom  $i$ .

It will be convenient to express the Heisenberg energy in the above equation in terms of the atomic moments  $\vec{\mu}$  rather than in terms of the spins  $\vec{S}$ . It can be readily done, considering that the relation between the spin and the atomic magnetic moment is  $\vec{\mu} = -g\mu_B\vec{S}$ . So, Eq. (20) can be rewritten as:

$$E_i = -\vec{\mu}_i \cdot \left( \mu_0\vec{H} + \frac{2J}{(g\mu_B)^2} \sum_k^z \vec{\mu}_k \right), \quad (21)$$

Note that the two expressions for the energy  $E_i$ , Eq. (19) and Eq. (21), are pretty similar. However, the sum term in Eq. (21) is not the same as  $\mu_0 H_{\text{eff}}$ , because the former is the interaction experienced by the individual spin  $i$ , whereas  $H_{\text{eff}}$  — in the spirit of the Weiss model — is the *time average* of all such interaction terms in the entire system. So, in order to obtain the proper  $H_{\text{eff}}$ , we should perform such averaging over all  $N$  atoms:

$$\mu_0\vec{H}_{\text{eff}} = \left\langle \frac{2J}{(g\mu_B)^2} \sum_k^z \vec{\mu}_k \right\rangle = \frac{2J}{(g\mu_B)^2} \times \frac{1}{N} \left\langle \sum_i^N \sum_k^z \vec{\mu}_k \right\rangle \quad (22)$$

The question of the indices in the above expression is little "tricky": at the first glance, it may look formally incorrect because the moments that are summed are not labelled with the index  $i$ . Actually, they are — but in a "hidden way". To explain more clearly what is going on, let's rewrite the whole thing in a more elaborate fashion:

$$\mu_0\vec{H}_{\text{eff}} = \left\langle \frac{2J}{(g\mu_B)^2} \sum_k^z \vec{\mu}_k \right\rangle = \frac{2J}{(g\mu_B)^2} \times \frac{1}{N} \left\langle \sum_i^N \underbrace{\sum_k^z \vec{\mu}_k}_{\substack{\text{with } k \\ \text{labeling} \\ \text{all the } z \\ \text{neighbors} \\ \text{of the } i \\ \text{atom}}} \right\rangle \quad (23)$$

Calculating the right side double-sum average is a very simple task if we introduce  $z$ , which denotes the number of neighbors with which each individual atom  $i$  interacts. Conversely, we can say that each atom is a neighbor to  $z$  other atoms. Hence, during the summation process, the magnetic moment of each atom will be taken  $z$  times; in other words, the value of the double sum is simply  $z$  times the vector sum of all individual atomic magnetic moments in the system! And the time average of this sum, of course, is nothing else than the average magnetization of the system.

The above conclusions can be written down in the equation language:

$$\left\langle \sum_i^N \sum_k \vec{\mu}_k^i \right\rangle = z \left\langle \sum_i^N \vec{\mu}_i \right\rangle.$$

and

$$z \left\langle \sum_i^N \vec{\mu}_i \right\rangle = z \vec{M}.$$

Combining the above results with Eq. (21) leads to a new equation for the magnetic energy of an individual atomic moment:

$$E_i = -\vec{\mu}_i \cdot \left( \vec{H}_{\text{ext}} + \frac{2Jz}{(g\mu_B)^2 N} \vec{M} \right). \quad (24)$$

Note that this is exactly what we wanted — the equation contains a term proportional to the magnetization, in perfect agreement with the Weiss' postulate. We can write now:

$$\mu_0 \vec{H}_{\text{eff}} = \frac{2Jz}{(g\mu_B)^2 N} \vec{M}, \quad (25)$$

from which we immediatly obtain the formula for the Weiss' "effective field constant":

$$\lambda = \frac{2Jz}{(g\mu_B)^2 \mu_0 N}. \quad (26)$$

Comparing this equation with Eq. (17), i.e., the phenomenological expression for  $\lambda$ , we obtain the solution for the critical temperature:

$$T_c = \frac{2zJS(S+1)}{3k} \quad (27)$$

This solution, note, does not contain any phenomenological parameters. It is often used for getting an estimate of  $J$  in a given system. In order to make the formula more "realistic", one has to analyze the interactions with neighbors in greater detail because even in the case of the simplest magnetic materials it is not possible to describe these interactions in terms of a single exchange parameter. But it is rather a straightforward task to construct an improved version of the model that properly takes into account all relevant neighbors with different  $J$ s. In many cases very accurate values of exchange constants for different neighbors can be obtained from dispersion relations of spin wave excitations measured by inelastic neutron scattering. It offers a chance to test the MFA results; such test show that for many systems the  $T_c$  values given by MFA may be accurate within a percent or two, even though the model is so simple.

**4. The region below  $T_c$ .** Another advantage of the Weiss' theory is that it offers a solution for  $M$  the region below the transition temperature as well. But here  $M$  is no longer small!

We cannot use the truncated power series for the Brillouin function any more. Instead, we have to use the equation for magnetization in its full full form, as given by Eq. (3):

$$M = NgS\mu_B B_S(x). \quad (28)$$

(again, we have changed  $J$  to  $S$ ). In the case of a paramagnet, the argument  $x$  was:

$$x = \frac{gS\mu_B\mu_0 H}{kT}.$$

To describe the situation in a ferromagnet in the spirit of the Weiss' theory, all we have to do is to replace the external field by a sum of the external field and the effective field. So, now the argument is:

$$x = \frac{gS\mu_B\mu_0(H + H_{\text{eff}})}{kT} = \frac{gS\mu_B\mu_0(H + \lambda M)}{kT} \quad (29)$$

Eq. (28) with the argument in the form of Eq. (29) describes the ferromagnet magnetization in all temperature regions for any external field value. However, the most interesting subject for us is the *spontaneous* magnetization, the one that occurs without any "aid" from the external field. To obtain the spontaneous  $M$ , we simply put  $H = 0$  in the equation and obtain:

$$M = NgS\mu_B B_S\left(\frac{gS\mu_B\mu_0\lambda M}{kT}\right) \quad (30)$$

By putting in  $\lambda$  given by Eq. (17) we obtain the final equation form:

$$M = NgS\mu_B B_S\left(\frac{3M}{(S+1)Ng\mu_B} \times \frac{T_c}{T}\right). \quad (31)$$

We can now introduce the reduced temperature  $\tau$ :

$$\tau \equiv \frac{T}{T_c} \quad (32)$$

and

$$m \equiv \frac{M}{Ng\mu_B S}. \quad (33)$$

Note that the expression in the denominator,  $Ng\mu_B S$ , is the *maximum value* the magnetization can attain, when all magnetic moments are completely parallel. So,  $m$  is the reduced magnetization which can take values between 0 and 1. With the new symbols, Eq. (31) takes the form:

$$m = B_S\left(\frac{3S}{S+1} \cdot \frac{m}{\tau}\right) \quad (34)$$

The only thing that has to be done now is to solve this equation for  $M$  to obtain the  $M(T)$  dependence! Well, but it's easy to say that, and more difficult to do that because this equation is a nontrivial transcendental one, whose solutions cannot be expressed in terms of elementary functions. The only way is to solve it numerically. With a computer it is not a very difficult task.

It's instructive to find the solution for the simplest possible case of  $S = \frac{1}{2}$ . Then the whole equation reduces to:

$$m = \tanh\left(\frac{m}{\tau}\right) \quad (35)$$

If one doesn't need a very high accuracy, this equation can be solved even without a digital computer, by a simple construction on a piece of graph paper.