Münster J. of Math. **13** (2020), 205–218
DOI 10.17879/32109549631
urn:nbn:de:hbz:6-32109549978

Münster Journal of Mathematics
© by the author 2020

The Curie–Weiss model—an approach using moments

Werner Kirsch

(Communicated by Matthias Löwe)

Abstract. We prove a law of large numbers and a central and a noncentral limit theorem for the Curie–Weiss model using the method of moments.

1. Introduction

In this paper, we consider one of the easiest models for magnetism, the Curie–Weiss model. In this model, the elementary magnets can take values $+1$ (spin up) and -1 (spin down). Each spin interacts with all the other spins with the same strength. This interaction makes it more likely for two spins to have the same value than to assume opposite values.

More precisely, the spins $X_1, \dots, X_N$ are $\{-1,+1\}$-valued random variables. As typical in models of statistical mechanics, the (joint) probability distribution of the $X_1, X_2, \dots, X_N$ is defined via a function $H\colon \{-1,+1\}^N \to \mathbb{R}$, called the *energy* (or Hamiltonian), by the expression

$$\mathbb{P}\big(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N\big) = Z^{-1} e^{-\beta H(x_1, x_2, \dots, x_N)},$$

where Z is a normalization constant to make $\mathbb{P}$ a probability measure, i.e.,

$$Z = \sum_{(x_1, x_2, \dots, x_N) \in \{-1,+1\}^N} e^{-\beta H(x_1, x_2, \dots, x_N)}.$$

The parameter $\beta \geq 0$ plays the role of an inverse 'temperature' T, $\beta = \frac{1}{T}$. If $\beta = 0$, which means $T = \infty$, the random variables $X_1, X_2, \dots, X_N$ are actually independent. If $\beta > 0$, those $X_1, X_2, \dots, X_N$ which minimize H have higher probability. In other words, the system prefers states with low energy. This preference is more and more enhanced if β grows.

The details of the model under consideration are encoded in the energy function H. As a rule, H is of the form

$$H(X_1, X_2, \dots, X_N) = -\sum_{i,j=1}^{N} J_{i,j} X_i X_j.$$

If all $J_{i,j} \geq 0$ (and not all $= 0$) the minimum of the energy is attained if the X_i are 'aligned', i.e., all $X_i = 1$ or all $X_i = -1$. Thus, those 'configurations' with many $X_i = 1$ (or with many $X_i = -1$) are more likely than those with almost equal number of $+1$ and -1. Such models are called *paramagnetic*.

Presumably, the most famous example is the energy function of the Ising model. In this model the indices i of the random variables X_i come from a finite subset I of the lattice $\mathbb{Z}^d$, and the coupling constants $J_{i,j}$ are given by

$$J_{i,j} = \begin{cases} 1 & \text{if } \|i-j\| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

So, in the Ising model only spins which are nearest neighbors interact with each other.

In this paper, we consider the easiest non trivial model of magnetisms, the Curie–Weiss model. In this system every spin interacts with every other spin, more precisely, the spin X_i interacts with the average of all spins, namely,

$$H(x_1, x_2, \dots, x_N) = -\frac{1}{2}\sum_{i=1}^{N} x_i \cdot \left(\frac{1}{N}\sum_{j=1}^{N} x_j\right) = -\frac{1}{2N}\sum_{i,j=1}^{N} x_i x_j.$$

The Curie–Weiss model is interesting since it is accessible to mathematical methods (even not too sophisticated ones) and yet has a number of interesting properties physicists expect of a paramagnetic system, like a phase transition from a purely paramagnetic phase to a ferromagnetic phase. We will explain this in detail in the next section.

The *results* we describe and prove below are not new, but rather well-known to the community. However, the *proofs* we present are certainly not standard, and rather elementary. We use the moment method to prove both a 'law of large numbers' as well as a 'central limit theorem' and a 'non-central limit theorem'.

We remark that Ellis and Newman [9] mention that the method of moments can be used to prove such results. However, these authors do not carry out these arguments.

The Curie–Weiss model goes back to Pierre Curie and Pierre Weiss. A systematic mathematical treatment can be found in [19] and [7]. For the vast literature on the model, see the references in [7]. We refer in particular to [8] and [9].

Recently, there has been increasing interest in proving limit results for Curie–Weiss models with two or more groups, see [2, 3, 4, 5, 14, 15, 16, 17].

Besides describing magnetic systems, the Curie–Weiss model is also used to model voting behavior in various election models, where $X_i = 1$ (resp.

$X_i = -1$) means the voter i votes 'yes' (resp. 'no'). The basic idea is that voters tend to vote in a similar way as the other voters in their constituency (see [6, 11, 12, 13]).

2. Definitions and results

Definition 2.1. For $N \in \mathbb{N}$ and $x_1, x_2, \ldots, x_N \in \{-1, +1\}$, set

$$H_N(x_1, x_2, \ldots, x_N) = -\frac{1}{2N}\Big(\sum_{i=1}^{N} x_i\Big)^2.$$

The Curie–Weiss distribution $CW(\beta, N)$ is the probability measure $\mathbb{P}_{\beta,N}$ on $\{-1, +1\}^N$ defined by

$$\mathbb{P}_{\beta,N}(\{(x_1, x_2, \ldots, x_N)\}) = Z^{-1} e^{-\beta H_N(x_1,x_2,\ldots,x_N)} = Z^{-1} e^{\frac{\beta}{2N}(\sum_{i=1}^{N} x_i)^2}.$$

Here, $\beta \geq 0$ is called the inverse temperature and Z is a normalization constant so that $\mathbb{P}_{\beta,N}$ is a probability, i.e.,

$$Z = \sum_{x_1,x_2,\ldots,x_N=1}^{N} e^{\frac{\beta}{2N}(\sum_{i=1}^{N} x_i)^2}.$$

By $\mathbb{E}_{\beta,N}$, we denote the expectation with respect to the probability measure $\mathbb{P}_{\beta,N}$.

We say that a sequence $X_1, X_2, \ldots, X_N$ of $\{-1, +1\}$-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is *Curie–Weiss distributed* with inverse temperature $\beta = \frac{1}{T} \geq 0$ (or $CW(\beta, N)$-distributed) if

$$\mathbb{P}\big(X_1 = x_1, X_2 = x_2, \ldots, X_N = x_N\big) = Z^{-1} e^{\frac{\beta}{2N}(\sum_{i=1}^{N} x_i)^2}.$$

If $X_1, X_2, \ldots, X_N$ are $CW(\beta, N)$-distributed, we call

$$S_N := \sum_{i=1}^{N} X_i$$

the *total magnetization* of the $X_1, X_2, \ldots, X_N$.

Remark 2.2. Suppose $X_1, X_2, \ldots, X_N$ are $CW(\beta, N)$-distributed random variables. Since the function H_N is invariant under permutation of its arguments, the random variables $X_1, X_2, \ldots, X_N$ are exchangeable. In particular, $\mathbb{E}_{\beta,N}(X_i X_j) = \mathbb{E}_{\beta,N}(X_1 X_2)$ for $i \neq j$. Moreover, $\mathbb{E}_{\beta,N}(X_i) = 0$, as $\mathbb{P}_{\beta,N}(X_i = \pm 1) = \frac{1}{2}$ and $\mathbb{E}_{\beta,N}(X_i^2) = 1$; in fact, $X_i^2 = 1$.

In the following we will be concerned with a scheme of random variables

$$X_i^{(N)}, \quad \text{with } N = 1, 2 \ldots \text{ and } i = 1, 2, \ldots, N,$$

such that the sequence $X_1^{(N)}, X_2^{(N)}, \ldots, X_N^{(N)}$ is $CW(\beta, N)$-distributed.

We will be interested in the behavior of $S_N^{(N)} = \sum_{i=1}^{N} X_i^{(N)}$.

Note, that the joint distributions of, say, $X_1^{(N)}, X_2^{(N)}$ and of $X_1^{(M)}, X_2^{(M)}$ are *different* for $N \neq M$, since the distribution $\mathrm{CW}(\beta, N)$ depends explicitly on N. In fact, a priori, X_1^N and X_1^M are defined on different probability spaces, so that it does not make sense to speak of quantities like $\mathbb{E}(X_i^{(N)} X_j^{(M)})$.

With this being said, from now on we drop the superscript (N) and (M) and simply write

$$S_N = \sum_{i=1}^{N} X_i \quad \text{instead of} \quad S_N^{(N)} = \sum_{i=1}^{N} X_i^{(N)}$$

whenever it is clear which N is meant. This is an abuse of notation, but a very convenient one.

The first result is a kind of a 'law of large numbers'.

Theorem 2.3. *Suppose $X_1, X_2, \dots, X_N$ are CW(β, N)-distributed random variables and set $S_N = \sum_{i=1}^N X_i$.*

(i) *If $\beta \leq 1$, then*

$$\frac{1}{N} S_N = \frac{1}{N} \sum_{i=1}^{N} X_i \overset{\mathcal{D}}{\Longrightarrow} \delta_0,$$

where $\overset{\mathcal{D}}{\Longrightarrow}$ denotes convergence in distribution and δ_a is the Dirac measure in a.

(ii) *If $\beta > 1$, then*

$$\frac{1}{N} S_N = \frac{1}{N} \sum_{i=1}^{N} X_i \overset{\mathcal{D}}{\Longrightarrow} \frac{1}{2}(\delta_{-m(\beta)} + \delta_{-m(\beta)}),$$

where $m(\beta) > 0$ is the unique positive solution of the equation

$$x = \tanh(\beta x).$$

Theorem 2.3 shows that there is a phase transition at inverse temperature $\beta = 1$, in the sense that the Curie–Weiss system changes its behavior drastically at $\beta = 1$. Up to this point, a 'law of large numbers' holds: The arithmetic mean of the spins goes to zero (= the expectation value of X_i). Above $\beta = 1$, the limiting distribution of the normalized sum of the spins has two peaks.

We remark that the convergence for $\beta \leq 1$ can be strengthened to convergence in probability if we realize all random variables of the same probability space.

Given the law of large numbers in Theorem 2.3, one may hope that there is a central limit theorem for $\beta \leq 1$. This is indeed the case for $\beta < 1$.

Theorem 2.4. *Suppose $X_1, X_2, \dots, X_N$ are CW(β, N)-distributed random variables. If $\beta < 1$, then*

$$\frac{1}{\sqrt{N}} S_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i \overset{\mathcal{D}}{\Longrightarrow} \mathcal{N}\Big(0, \frac{1}{1-\beta}\Big),$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2.

It follows, in particular, that

$$\mathbb{E}_{\beta,N}\Big(\Big(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}X_i\Big)^2\Big) \to \frac{1}{1-\beta},$$

while $\mathbb{E}_{\beta,N}(X_i^2) = 1$.

The above result suggests that for $\beta = 1$, there is no 'standard' central limit theorem. Indeed, we have:

Theorem 2.5. *Suppose $X_1, X_2, \dots, X_N$ are $CW(1,N)$-distributed random variables. Then*

$$\frac{1}{N^{3/4}}S_N = \frac{1}{N^{3/4}}\sum_{i=1}^{N}X_i \overset{\mathcal{D}}{\Longrightarrow} \mu,$$

where μ is a measure with Lebesgue density $\rho(x) = Ce^{-\frac{1}{12}x^4}$.

Since for $\beta > 1$ the expression $\frac{1}{N}S_N$ converges to a distribution which is not concentrated in one point, there is no central limit theorem in the usual sense that for a suitable constant

$$\frac{1}{\sqrt{N}}(S_N - c) \overset{\mathcal{D}}{\Longrightarrow} \mu.$$

However, there is a 'conditional' version of the central limit theorem. For details, we refer to [10].

3. Strategy of the proofs

To prove convergence in distribution, we use the method of moments.

Theorem 3.1 (Method of moments)**.** *Suppose μ_n and μ are Borel measures on $\mathbb{R}$ such that all moments*

$$m_k(\mu_n) := \int x^k\, d\mu_n \quad \textit{and} \quad m_k(\mu) := \int x^k\, d\mu$$

are finite and such that

$$|m_k(\mu)| \le AC^k k!.$$

If for all k, $m_k(\mu_n) \to m_k(\mu)$, then $\mu_n \Longrightarrow \mu$.

For a proof, see, e.g., [1].

To employ Theorem 3.1 we got to estimate expressions of the form

$$\mathbb{E}_{\beta,N}\Big(\Big(\frac{1}{N^\alpha}\sum_{i=1}^{N}X_i\Big)^K\Big),$$

with $\alpha \in \{\frac{1}{2}, \frac{3}{4}, 1\}$.

We have

$$\mathbb{E}_{\beta,N}\Big(\Big(\sum_{i=1}^{N}X_i\Big)^K\Big) = \sum_{x_{1_1}, x_{i_2}, \dots, x_{i_K}=1}^{N} \mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K}). \tag{1}$$

Note that for pairwise *distinct* $j_1, \dots, j_\ell$

$$\mathbb{E}_{\beta,N}(X_{j_1} \cdot X_{j_2} \cdots X_{j_\ell}) = \mathbb{E}_{\beta,N}(X_1 \cdot X_2 \cdots X_\ell),$$

since the measure $\mathbb{P}_{\beta,N}$ is invariant under permutations of indices (exchangeability).

We observe that $X_i^\ell = X_i$ for odd ℓ and $X_i^\ell = 1$ for even ℓ. Thus,

$$\mathbb{E}_{\beta,N}(X_{i_1} \cdot X_{i_2} \cdots X_{i_K}) = \mathbb{E}_{\beta,N}(X_1 \cdot X_2 \cdots X_\ell), \tag{2}$$

where $\ell \leq K$ is the number of indices i_ν which occur an odd number of times among $i_1, \dots, i_K$.

In the following section, we estimate expectations of the form (2). It turns out that their behavior in N depends strongly on the parameter β. In Sections 5 to 7, we use this information to evaluate the moments (1), thus proving Theorems 2.3, 2.4 and 2.5.

4. Correlations

In this section, we estimate correlations of the form

$$\mathbb{E}_{\beta,N}(X_1 \cdot X_2 \cdots X_\ell). \tag{3}$$

To do so, it is convenient to write the probability distribution $\mathbb{P}_{\beta,N}$ in a form which is more suitable for sending N to infinity. The basic idea, known in physics as the Hubbard–Stratonovich transform, is to use the equality

$$e^{a^2/2} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2 + ax}\, dx, \tag{4}$$

which is nothing but $\frac{1}{\sqrt{2\pi}} \int e^{\frac{1}{2}(x-a)^2}\, dx = 1$.

This observation allows us to write the correlations (3) in the following form.

Proposition 4.1. *Define* $F_\beta(t) := \frac{1}{2\beta} t^2 - \ln\cosh(t)$ *and set*

$$\mathcal{Z}_N(\ell) := \int_{-\infty}^{+\infty} e^{NF_\beta(t)} \tanh^\ell(t)\, dt.$$

Then, for $\ell \leq N$,

$$\mathbb{E}_{\beta,N}(X_1 \cdot X_2 \cdots X_\ell) = \frac{\mathcal{Z}_N(\ell)}{\mathcal{Z}_N(0)}.$$

Proof. By (4), we have

$$\begin{aligned}
\mathcal{T}_N(\ell) &:= \frac{1}{2^N} \sum_{x_1,x_2,\dots,x_N\in\{-1,+1\}} x_1x_2\cdots x_\ell e^{\frac{\beta}{2N}(\sum_{i=1}^N x_i)^2} \\
&= \frac{1}{2^N\sqrt{2\pi}} \sum_{x_1,x_2,\dots,x_N\in\{-1,+1\}} x_1x_2\cdots x_\ell \int_{-\infty}^{\infty} e^{-\frac{1}{2}s^2+\frac{\sqrt{\beta}}{\sqrt{N}}(\sum_{i=1}^N x_i)s}\,ds \\
&= \frac{1}{2^N\sqrt{2\pi}} \sum_{x_1,x_2,\dots,x_N\in\{-1,+1\}} x_1x_2\cdots x_\ell \int_{-\infty}^{\infty} e^{-N\frac{1}{\beta}t^2}\,dt \\
&= \frac{1}{2^N\sqrt{2\pi}} \int_{-\infty}^{\infty} \sum_{x_2,\dots,x_N\in\{-1,+1\}} \Big(\sum_{x_1\in\{-1,+1\}} x_1e^{x_1t}\Big)x_2\cdots x_\ell \\
&\qquad \times e^{-N\frac{1}{\beta}t^2}\prod_{i=2}^N e^{x_it}\,dt \\
&= \frac{1}{2^{N-1}\sqrt{2\pi}} \int_{-\infty}^{\infty} \sum_{x_2,\dots,x_N\in\{-1,+1\}} \sinh(t)x_2\cdots x_\ell e^{-N\frac{1}{\beta}t^2}\prod_{i=2}^N e^{x_it}\,dt \\
&= \frac{1}{2^{N-\ell}\sqrt{2\pi}} \int_{-\infty}^{\infty} \sum_{x_{\ell+1},\dots,x_N\in\{-1,+1\}} \sinh^\ell(t)e^{-N\frac{1}{\beta}t^2}\prod_{i=\ell+1}^N e^{x_it}\,dt \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sinh^\ell(s)\cosh^{N-\ell}(s)e^{-N\frac{1}{\beta}t^2}\,dt \\
&= \frac{\sqrt{N}}{\sqrt{2\pi\beta}} \int_{-\infty}^{\infty} e^{-N(\frac{t^2}{2\beta}-\ln\cosh t)}\tanh^\ell(t)\,dt.
\end{aligned}$$

Consequently,

$$\mathbb{E}_{\beta,N}(X_1X_2\cdots X_\ell) = \frac{\mathcal{T}_N(\ell)}{\mathcal{T}_N(0)} = \frac{\mathcal{Z}_N(\ell)}{\mathcal{Z}_N(0)}. \qquad \square$$

By symmetry, we see that $\mathcal{Z}_N(\ell) = 0$ for odd ℓ. To estimate $\mathcal{Z}_N(\ell)$ for even ℓ, we use Laplace's method:

Theorem 4.2 (Laplace)**.** *Suppose the smooth function $F\colon \mathbb{R}\to\mathbb{R}$ has a unique global minimum at t_0 with $F^{(m)}(t_0) > 0$ for an even m and $F^{(r)}(t_0) = 0$ for all $0 \le r < m$, moreover, let φ be a bounded continuous function which is continuous at t_0 with $\varphi(t_0) \neq 0$.*

If $\int_{-\infty}^{+\infty} e^{-NF(t)}|t^\ell|\,dt$ is finite for all ℓ and all N large enough, then

$$\int_{-\infty}^{+\infty} e^{-NF(t)}t^\ell\varphi(t)\,dt \underset{N\to\infty}{\approx} \Big(\frac{1}{NF^{(m)}(0)}\Big)^{\frac{\ell+1}{m}}\varphi(0)\int_{-\infty}^{+\infty} e^{-\frac{1}{m!}t^m}t^\ell\,dt. \tag{5}$$

The Laplace theorem in the form we need it here can be deduced from [18]. For the reader's convenience, we give a rough sketch of a proof in Appendix A.

Propositions 4.1 and 4.2 allow us to compute the asymptotic behavior of the correlations (3).

Theorem 4.3. *Suppose $X_1, X_2, \ldots, X_\ell, X_{\ell+1}, \ldots, X_N$ are $CW(\beta,N)$-distributed variables. If ℓ is even, then as $N \to \infty$:*

(i) *If $\beta < 1$, then*

$$\mathbb{E}_{\beta,N}(X_1 \cdot X_2 \cdots X_\ell) \approx (\ell-1)!!\Big(\frac{\beta}{1-\beta}\Big)^{\frac{\ell}{2}} \frac{1}{N^{\frac{\ell}{2}}}. \tag{6}$$

(ii) *If $\beta = 1$, then*

$$\mathbb{E}_{1,N}(X_1 \cdot X_2 \cdots X_\ell) \approx \frac{1}{N^{\frac{\ell}{4}}} \frac{\int t^\ell e^{-\frac{1}{12}t^4}\,dt}{\int e^{-\frac{1}{12}t^4}\,dt}. \tag{7}$$

(iii) *If $\beta > 1$, then*

$$\mathbb{E}_{\beta,N}(X_1 \cdot X_2 \cdots X_\ell) \approx m(\beta)^\ell, \tag{8}$$

where $t = m(\beta)$ is the strictly positive solution of $\tanh \beta t = t$.

If ℓ is odd, then $\mathbb{E}_{\beta,N}(X_1 \cdot X_2 \cdots X_\ell) = 0$ for all β.

Remark 4.4. Up to the factor $N^{-\ell/2}$, (6) is the ℓth moment of the normal distribution $\mathcal{N}(0, \frac{\beta}{1-\beta})$, (7) are the moments of a probability measure with density proportional to $e^{-\frac{1}{12}t^4}$ up to the factor $N^{-\ell/4}$, and (8) are the moments of the measure $\frac{1}{2}(\delta_{-m(\beta)} + \delta_{m(\beta)})$.

Proof. We compute:

$$F_\beta'(t) = \frac{1}{\beta}t - \tanh t, \quad F_\beta''(t) = \frac{1}{\beta} - \frac{1}{\cosh^2 t}.$$

Thus, for $\beta < 1$, the function F_β is strictly convex and has a local minimum at $t = 0$. Consequently, this minimum is global and we can apply Proposition 4.2 to find

$$\begin{aligned}
\mathbb{E}_{\beta,N}(X_1 \cdots X_\ell) &= \frac{\mathcal{Z}_N(\ell)}{\mathcal{Z}_N(0)} \\
&= \int_{-\infty}^{+\infty} e^{NF_\beta(t)} \tanh^\ell(t)\,dt\Big(\int_{-\infty}^{+\infty} e^{NF_\beta(t)}\,dt\Big)^{-1} \\
&= \int_{-\infty}^{+\infty} e^{NF_\beta(t)} t^\ell \frac{\tanh^\ell(t)}{t^\ell}\,dt\Big(\int_{-\infty}^{+\infty} e^{NF_\beta(t)}\,dt\Big)^{-1} \\
&\approx \frac{1}{N^{\ell/2}}\Big(\frac{\beta}{1-\beta}\Big)^{\ell/2} \frac{1}{\sqrt{2\pi}} \int t^\ell e^{-t^2/2}\,dt \\
&= (\ell-1)!!\Big(\frac{\beta}{1-\beta}\Big)^{\ell/2} \frac{1}{N^{\ell/2}}.
\end{aligned}$$

For $\beta = 1$, we obtain $t = 0$ is still the unique solution of $F_1'(t) = 0$, $F_1^{(2)}(0) = F_1^{(3)}(0) = 0$ and $F_1^{(4)} = 2$. Thus, $t = 0$ is a global minimum of F_1 and the above reasoning gives (7).

For $\beta > 1$, we have $F_\beta'(0) = 0$ and $F_\beta''(0) = \frac{1}{\beta} - 1 < 0$, so 0 is a local maximum.

Since $F_\beta(t) = F_\beta(-t)$, we have, for r even,

$$\begin{aligned}\mathcal{Z}_N(r) &= \int_{-\infty}^{0} e^{NF_\beta(t)} \tanh^r(t)\,dt + \int_0^\infty e^{NF_\beta(t)} \tanh^r(t)\,dt \\ &= 2\int_0^\infty e^{NF_\beta(t)} \tanh^r(t)\,dt. \qquad (9)\end{aligned}$$

Thus, it suffices to estimate the integrals (9) for $r = \ell$ and $r = 0$.

Set $f(t) = \frac{1}{\beta}t$ and $g(t) = \tanh(t)$, so $F'_\beta(t) = f(t) - g(t)$.

We have $f(0) = g(0)$ and, due to $\beta > 1$, $f'(0) < g'(0)$, hence $f(t) < g(t)$ for small $t > 0$. Moreover, g is bounded and strictly concave (for $t > 0$). Consequently, there is a unique $t_0 > 0$ with $F'_\beta(t_0) = f(t_0) - g(t_0) = 0$. We have $g'(t_0) < f'(t_0)$ due to the concavity of g, hence $F''_\beta(t_0) > 0$.

By Proposition 4.2, we obtain

$$\mathbb{E}_{\beta,N}(X_1 \cdots X_\ell) \approx \Big(\frac{t_0}{\beta}\Big)^\ell =: m(\beta)^\ell.$$

We have

$$\tanh(\beta m(\beta)) = \tanh(t_0) = \beta t_0 = m(\beta).$$

This proves (8). □

5. Proof of Theorem 2.3

We estimate

$$\mathbb{E}_{\beta,N}\Big(\Big(\frac{1}{N}\sum_{i=1}^{N}\Big)^K\Big) = \frac{1}{N^K}\mathbb{E}_{\beta,N}\Big(\sum_{i_1,i_2,\dots,i_K=1}^{N} X_{i_1}\cdot X_{i_2}\cdots X_{i_K}\Big). \qquad (10)$$

Evaluating these sums is a combination of bookkeeping and correlation estimates as in Section 4. To do the bookkeeping, we define:

Definition 5.1. We set

$$\begin{aligned} W_{K,N} &:= \{\underline{\boldsymbol{i}} = (i_1, i_2, \dots, i_K) \mid 1 \le i_j \le N\}, \\ W_{K,N}(r) &:= \{\underline{\boldsymbol{i}} \in W_{K,N} \mid \text{exactly } r \text{ different indices occur once in } \underline{\boldsymbol{i}}\}.\end{aligned}$$

By $w_{K,N}$ and $w_{K,N}(r)$, we denote the number of multiindices in $W_{K,N}$ and $W_{K,N}(r)$, respectively.

Lemma 5.2. *We have*

$$w_{K,N}(r) \le K!N^{\frac{K+r}{2}}, \qquad (11)$$

$$w_{K,N}(K) = \frac{N!}{(N-K)!} \approx N^K. \qquad (12)$$

Proof. The multiindices in $W_{K,N}(r)$ contain at most $r + \frac{K-r}{2} = \frac{K+r}{2}$ different indices. There are at most $N^{\frac{K+r}{2}}$ ways to choose them and at most $K!$ ways to order them.

For $r = K$, we have $\frac{N!}{(N-K)!} \approx N^K$ possibilities to choose an ordered K-tuple from N indices (without repetition). □

We estimate

$$\begin{aligned}\mathbb{E}_{\beta,N}\Big(\Big(\frac{1}{N}\sum_{i=1}^{N}X_i\Big)^K\Big) &= \frac{1}{N^K}\sum_{r=0}^{K-1}\sum_{\underline{\boldsymbol{i}}\in W_{K,N}(r)}\mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K})\\ &\quad+\frac{1}{N^K}\sum_{\underline{\boldsymbol{i}}\in W_{K,N}(K)}\mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K})\\ &\approx \frac{1}{N^K}CN^{K-\frac{1}{2}}+\mathbb{E}_{\beta,N}(X_1\cdot X_2\cdots X_K)\\ &\approx \mathbb{E}_{\beta,N}(X_1\cdot X_2\cdots X_K).\end{aligned}$$

The last expression goes to 0 for $\beta \le 1$, by Theorem 4.3. For $\beta > 1$, it converges to $m(\beta)$ for even K and to 0 for odd K.

Together with Theorem 3.1, this proves Theorem 2.3.

6. Proof of Theorem 2.4

In our proof of Theorem 2.3, we realized that only terms with K distinct indices counted in the limit for (10). For the central limit theorem for *independent* random variables, the only important terms are those with all indices occurring exactly twice.

It will turn out that for the Curie–Weiss model with $\beta < 1$, both doubly occurring indices and those that occur only once play a role in the limit.

To do the bookkeeping we got to refine our definitions in Definition 5.1.

Definition 6.1. We set

$$\begin{aligned}W^0_{K,N}(r) &= \{\underline{\boldsymbol{i}}\in W_{K,N}(r)\mid \text{no index occurs more than twice}\},\\ W^+_{K,N}(r) &= W_{K,N}(r)\setminus W^0_{K,N}(r).\end{aligned}$$

and denote by $w^0_{K,N}(r)$ and the $w^+_{K,N}(r)$ the cardinality of $W^0_{K,N}(r)$ and $W^+_{K,N}(r)$, respectively.

Lemma 6.2. *We have*

$$w^+_{K,N}(r) \le K!N^{\frac{K+r}{2}-\frac{1}{2}}.$$

Proof. If the K-tuple $\underline{\boldsymbol{i}}$ contains r indices with only one occurrence and at least one index with three or more occurrences, there are at most $r-3$ places left for indices with (exactly) two occurrences. Therefore, a tuple in $w^+_{K,N_1}(r)$ contains at most $r+1+\frac{K-r-3}{2}$ *different* indices. Consequently, there are at most $K!N_1^{\frac{K+r}{2}-\frac{1}{2}}$ such tuples. □

Lemma 6.3. *We have*

$$w^0_{K,N}(r) = \begin{cases}\dfrac{N!}{(N-\frac{K+r}{2})!}\dfrac{K!}{r!(\frac{K-r}{2})!2^{\frac{K-r}{2}}} & \text{if } K-r \text{ is even,}\\ 0 & \text{otherwise.}\end{cases}$$

Proof. We choose an (ordered) r-tuple ρ of r indices to occur once and an ordered $(K-r)/2$-tuple λ of indices to occur twice in $\underline{\boldsymbol{i}}$. We have

$$\frac{N!}{(N-\frac{K+r}{2})!}$$

ways to do so.

Then we choose the r positions for those indices which occur once. We can do this in

$$\binom{K}{r} = \frac{K!}{r!(K-r)!}$$

ways. We fill these positions in $\underline{\boldsymbol{i}}$ with $\rho_1, \rho_2, \ldots, \rho_r$, starting with the left most open position.

Finally, we distribute the indices $\lambda_1, \ldots, \lambda_{(K-r)/2}$, twice each. The index λ_1 is put at the left most free place in $\underline{\boldsymbol{i}}$ and in one of the remaining $K-r-1$ positions, λ_2 is put at the then first free place in $\underline{\boldsymbol{i}}$ and in one of the $K-r-3$ remaining free places, and so on.

This gives

$$(K-r-1)!! = \frac{(K-r)!}{(\frac{K-r}{2})!2^{\frac{K-r}{2}}}$$

possibilities. □

We are now in a position to complete the proof of Theorem 2.4.

We split the sum into two parts:

$$\begin{aligned}
\mathbb{E}_{\beta,N}\Big(\Big(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Big)^{K}\Big) &= \frac{1}{N^{K/2}}\mathbb{E}_{\beta,N}\Big(\sum_{i_1,i_2,\ldots,i_K=1}^{N} X_{i_1}\cdot X_{i_2}\cdots X_{i_K}\Big) \\
&= \frac{1}{N^{K/2}}\sum_{r=0}^{K}\sum_{\underline{\boldsymbol{i}}\in W^0_{K,N}(r)} \mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K}) \\
&\quad + \frac{1}{N^{K/2}}\sum_{r=0}^{K}\sum_{\underline{\boldsymbol{i}}\in W^+_{K,N}(r)} \mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K}). \qquad (13)
\end{aligned}$$

We estimate the second term in (13) first. If $\underline{\boldsymbol{i}} \in W^+_{K,N}(r)$, then

$$\mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K}) = \mathbb{E}_{\beta,N}(X_1\cdot X_2\cdots X_r\cdots X_{r+s}), \qquad (14)$$

since $X_i^\ell = 1$ for even ℓ and $X_i^\ell = X_i$ for odd ℓ. (In (14), s may be 0.)

Consequently, for $\underline{\boldsymbol{i}} \in W^+_{K,N}(r)$, Theorem 4.3 (iii) gives

$$\mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K}) \le C_1 N^{-r/2}.$$

By Lemma 6.2, we conclude that

$$\frac{1}{N^{K/2}}\sum_{r=0}^{K}\sum_{\underline{\boldsymbol{i}}\in W^+_{K,N}(r)} \mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K}) \le C_2 N^{-1/2}.$$

The remaining, in fact leading, term is

$$\frac{1}{N^{K/2}}\sum_{r=0}^{K}\sum_{\underline{i}\in W^0_{K,N}(r)}\mathbb{E}_{\beta,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K})$$

$$= \frac{1}{N^{K/2}}\sum_{r=0}^{K}\sum_{\underline{i}\in W^0_{K,N}(r)}\mathbb{E}_{\beta,N}(X_1\cdot X_2\cdots X_r). \tag{15}$$

Since K is even and $\mathbb{E}_{\beta,N}(X_1\cdot X_2\cdots X_r)=0$ for odd r, we may set $K=2L$ and write (15) as

$$\begin{aligned}
&\frac{1}{N^{L}}\sum_{\ell=0}^{L}\sum_{\underline{i}\in W^0_{2L,N}(2\ell)}\mathbb{E}_{\beta,N}(X_1\cdot X_2\cdots X_{2\ell})\\
&\quad\approx \frac{1}{N^L}\sum_{\ell=0}^{L}\frac{N!}{(N-(L+\ell))!}\frac{(2L)!}{(2\ell)!(L-\ell)!2^{L-\ell}}(2\ell-1)!!\Big(\frac{\beta}{1-\beta}\Big)^{\ell}N^{-\ell}\\
&\quad\approx \sum_{\ell=0}^{L}\frac{(2L)!}{(2\ell)!(L-\ell)!2^{L-\ell}}(2\ell-1)!!\Big(\frac{\beta}{1-\beta}\Big)^{\ell}\\
&\quad= \frac{(2L)!}{L!2^L}\sum_{\ell=0}^{L}\frac{L!}{(L-\ell)!\ell!}\Big(\frac{\beta}{1-\beta}\Big)^{\ell}\\
&\quad= (2L-1)!!\Big(\frac{1}{1-\beta}\Big)^{L} = (K-1)!!\Big(\frac{1}{1-\beta}\Big)^{K/2},
\end{aligned}$$

which are the moments $m_K(\mathcal{N}(0,\frac{1}{1-\beta}))$ of a normal distribution with mean zero and variance $\frac{1}{1-\beta}$ for even K.

7. Proof of Theorem 2.5

To prove Theorem 2.5, we have to estimate

$$\begin{aligned}
\frac{1}{N^{\frac34 K}}\mathbb{E}_{1,N}\Big(\Big(\sum_{i=1}^{N}X_i\Big)^{K}\Big) &= \frac{1}{N^{\frac34 K}}\sum_{r=0}^{K-1}\sum_{\underline{i}\in W_{K,N}(r)}\mathbb{E}_{1,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K})\\
&\quad+\frac{1}{N^{\frac34 K}}\sum_{\underline{i}\in W_{K,N}(K)}\mathbb{E}_{1,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K}).
\end{aligned}\tag{16}$$

Due to Theorem 4.3, equation (7) and estimate (11), the first term in (16) goes to zero. The second term in (16) can be estimated by Theorem 4.3, equation (7) and (12):

$$\frac{1}{N^{\frac34 K}}\sum_{\underline{i}\in W_{K,N}(K)}\mathbb{E}_{1,N}(X_{i_1}\cdot X_{i_2}\cdots X_{i_K}) \approx \frac{1}{N^{\frac34 K}}N^{K}\frac{1}{N^{\frac14}}\frac{\int t^{\ell}e^{-\frac{1}{12}t^4}\,dt}{\int e^{-\frac{1}{12}t^4}\,dt}.$$

This gives the result.

APPENDIX A

In this appendix, we give a rough sketch of a proof of Theorem 4.2, details to justify the approximations made below can be found in [18] or [10].

Without loss of generality, we may assume that $t_0 = 0$. To approximate the left-hand side of (5), we make a Taylor expansion $F(t) \approx \frac{1}{m!}F^{(m)}(0)t^m$. We obtain

$$\int_{-\infty}^{+\infty} e^{-NF(t)}t^\ell\,\varphi(t)\,dt \approx \int_{-\infty}^{+\infty} e^{-N\,\frac{1}{m!}F^{(m)}(0)t^m}t^\ell\varphi(t)\,dt,$$

and by setting $s = (NF^{(m)}(0))^{1/m}t$, we get

$$\approx \frac{1}{(NF^{(m)}(0))^{1/m}}\int_{-\infty}^{+\infty} e^{-\frac{1}{m!}s^m}\Big(\frac{1}{(NF^{(m)}(0))^{1/m}}s\Big)^\ell\varphi\Big(\frac{1}{(NF^{(m)}(0))^{1/m}}s\Big)\,ds$$
$$\approx \frac{1}{(NF^{(m)}(0))^{(\ell+1)/m}}\varphi(0)\int_{-\infty}^{+\infty} e^{-\frac{1}{m!}s^m}s^\ell\,ds.$$

Acknowledgment. The author would like to thank the unknown referee for useful comments.

REFERENCES

[1] L. Breiman, *Probability*, Addison-Wesley Publishing Company, Reading, MA, 1968. MR0229267

[2] F. Collet, Macroscopic limit of a bipartite Curie–Weiss model: a dynamical approach, J. Stat. Phys. **157** (2014), no. 6, 1301–1319. MR3277768

[3] I. Gallo and P. Contucci, Bipartite mean field spin systems. Existence and solution, Math. Phys. Electron. J. **14** (2008), Paper 1, 21 pp. MR2407199

[4] M. Fedele, Rescaled magnetization for critical bipartite Curie–Weiss models, J. Stat. Phys. **155** (2014), no. 2, 223–236. MR3182530

[5] M. Fedele and P. Contucci, Scaling limits for multi-species statistical mechanics mean-field models, J. Stat. Phys. **144** (2011), no. 6, 1186–1205. MR2841921

[6] P. Contucci and S. Ghirlanda, Modelling society with statistical mechanics: an application to cultural contact and immigration, Quality and Quantity **41** (2007), 569–578.

[7] R. S. Ellis, *Entropy, large deviations, and statistical mechanics*, Grundlehren der Mathematischen Wissenschaften, 271, Springer-Verlag, New York, 1985. MR0793553

[8] R. S. Ellis and C. M. Newman, Limit theorems for sums of dependent random variables occurring in statistical mechanics, Z. Wahrsch. Verw. Gebiete **44** (1978), no. 2, 117–139. MR0503333

[9] R. S. Ellis and C. M. Newman, The statistics of Curie–Weiss models, J. Statist. Phys. **19** (1978), no. 2, 149–161. MR0503332

[10] W. Kirsch, A survey on the method of moments, available from `http://www.fernuni-hagen.de/stochastik/`.

[11] W. Kirsch, On Penrose's square root law and beyond, Homo Oeconomicus **24** (2007), 357–380.

[12] W. Kirsch, A mathematical view on voting and power, in *Mathematics and society*, 251–279, Eur. Math. Soc., Zürich, 2016. MR3497597

[13] W. Kirsch and J. Langner, The fate of the square root law for correlated voting, in *Voting power and procedures*, 147–158, Stud. Choice Welf, Springer, Cham, 2014. MR3290324

[14] W. Kirsch and G. Toth, Two groups in a Curie–Weiss model, arXiv:1712.08477v2 [math.PR] (2017).

[15] W. Kirsch and G. Toth, Two groups in a Curie–Weiss model with heterogeneous coupling, arXiv:1806.06708v3 [math.PR] (2019), to appear in: Journal of Theoretical Probability.

[16] H. Knöpfel, M. Löwe, K. Schubert, and A. Sinulis, Fluctuation results for general block spin Ising models, arXiv:1902.02080v1 [math.PR] (2019).

[17] M. Löwe and K. Schubert, Fluctuations for block spin Ising models, Electron. Commun. Probab. **23** (2018), Paper No. 53, 12 pp. MR3852267

[18] F. W. J. Olver, *Asymptotics and special functions*, Academic Press, New York, 1974. MR0435697

[19] C. J. Thompson, *Mathematical statistical mechanics*, The Macmillan Co., New York, 1972. MR0469020

Received September 11, 2019; accepted December 6, 2019

Werner Kirsch
Fakultät für Mathematik und Informatik,
Fernuniversität in Hagen, Germany
E-mail: `werner.kirsch@fernuni-hagen.de`