1 Introduction

This talk will be an extended application of what we did last time, which was a concentration result using exchangeable pairs. The object of our study is the Curie-Weiss model, which is a mean-field magnet model. Think of this model as a simplified version (or a special case) of the Ising model; we have a network with spins at the vertices and interactions across the edges. The Ising model is usually thought of on $\mathbb{Z}^d$ or a finite grid, but the Curie-Weiss model lives on the complete graph, so every spin interacts with every other spin, and there are no geometrical considerations to consider. This is all based on §7.2 of Ross's notes about Stein's method.

2 Definition of the Model

The state space is $\Omega = \{-1, +1\}^n$, and the probability of any given spin configuration $\sigma \in \Omega$ is

$$\mathbb{P}[\sigma] = \frac{1}{Z} \exp\left(\frac{\beta}{n} \sum_{i<j} \sigma_i \sigma_j + \beta h \sum_i \sigma_i \right),$$

where $\beta > 0$ is the inverse temperature, $h \in \mathbb{R}$ is the strength of an external magnetic field, and Z is a normalization constant (the partition function). This $\mathbb{P}$ is called the *Gibbs measure*.

Notice that if $\beta = 0$, which corresponds to infinite temperature, then this is the uniform distribution on Ω, which means that each site will have an independent Rademacher spin. If β is small (high temperature) we expect to see similar behavior: mostly disorder. However, if β is large (low temperature), then the system is heavily penalized for having sites with disagreeing spins, so we expect to see that all sites will share the same spin (either $+1$ or -1).

In this case, if $h > 0$ then the system will also be penalized for having negative spins, so the system should concentrate around the "all $+1$" configuration. Similarly, if $h < 0$ then the system should concentrate around the "all -1" configuration. If $h = 0$, then the spins will all tend to agree, but with equal chance of being near the "all $+1$" or "all -1" configurations.

The "magnetization" of the system is just the average of all of the spins at all of the sites:

$$M = M(\sigma) = \frac{1}{n} \sum_{i=1}^{n} \sigma_i.$$

From the previous paragraphs, if β is small then we expect M to be close to zero. If β is large we expect $|M|$ to be close to 1, with the sign depending on h.

3 Theorem Statement and Discussion

This theorem makes (most) of the previous section's heuristics quantitative:

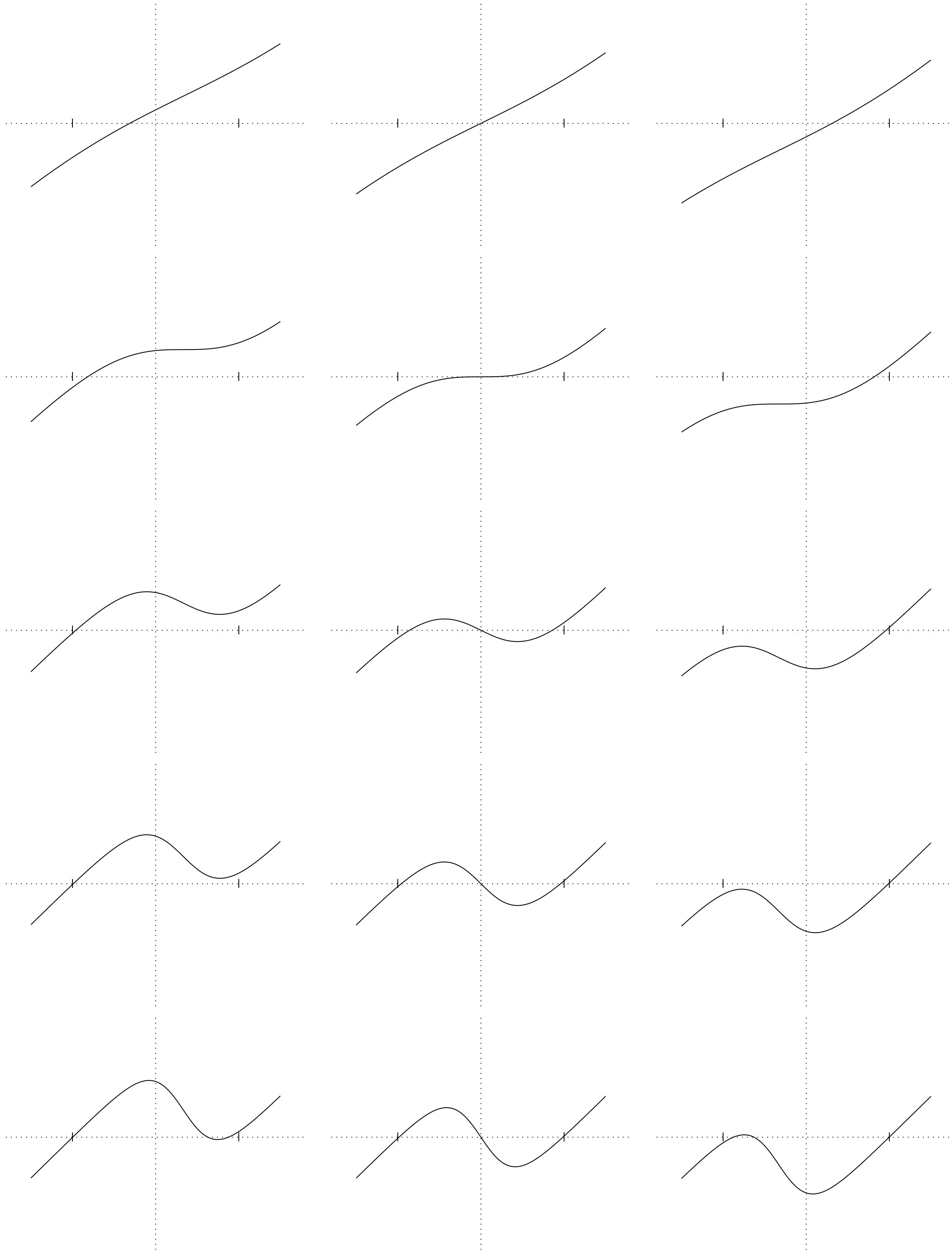
Theorem 1. *For any $t \geq 0$,*

$$\mathbb{P}\left[|M - \tanh(\beta M + \beta h)| \geq \frac{\beta}{n} + \frac{t}{\sqrt{n}} \right] \leq 2 \exp\left(-\frac{t^2}{4(1+\beta)} \right),$$

where M is the magnetization of the Curie-Weiss model with parameters n, β, h.

At first, this seems a bit mysterious: what is that tanh function actually doing? It is very helpful to look at some plots of the expression $M - \tanh(\beta M + \beta h)$ as a function of M; the next page is full of them.

For some context, recall that $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, so it approaches ± 1 as $x \to \pm\infty$, is strictly increasing, and has an inflection point at $x = 0$ with derivative 1.



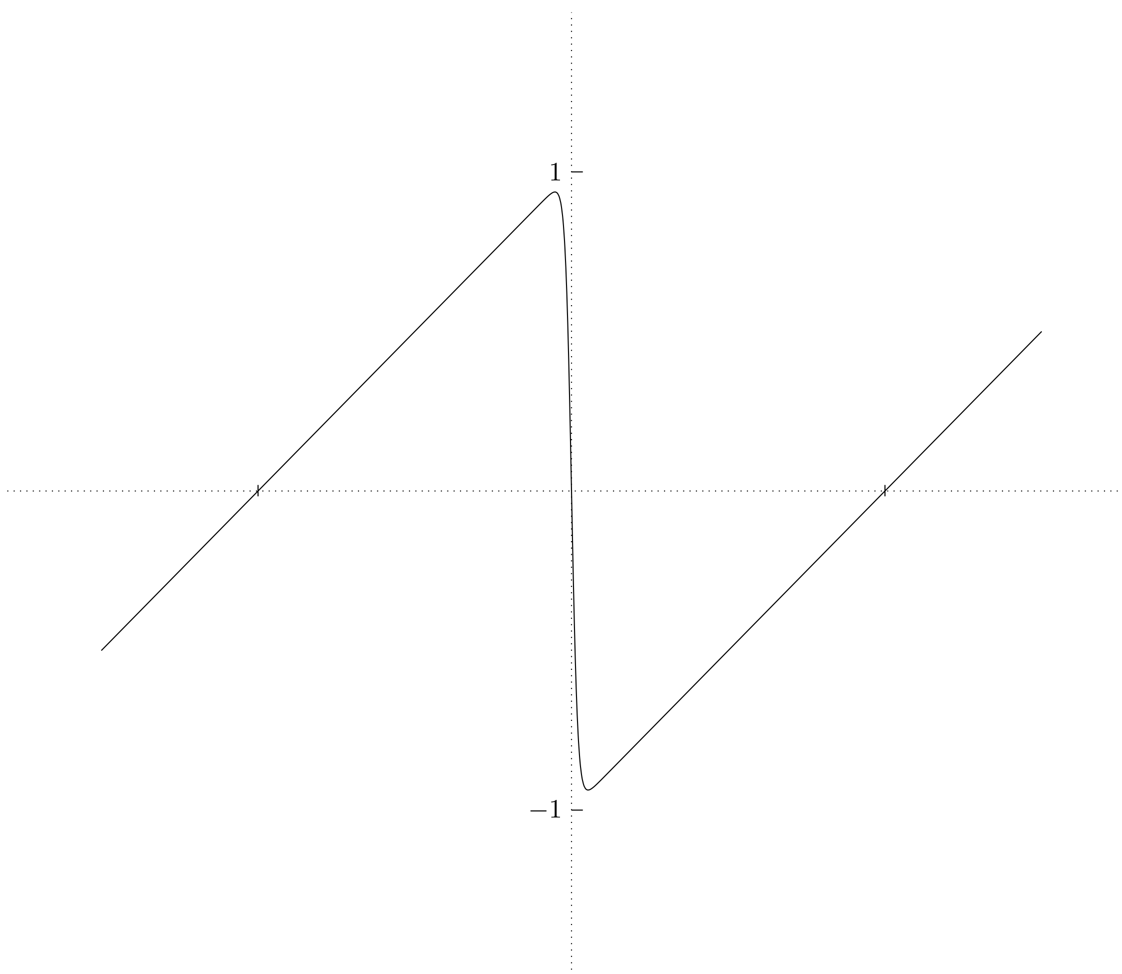
Plots of the function $M \mapsto M - \tanh(\beta M + \beta h)$, with tick marks at $x = \pm 1$. The columns (in order from left to right) have $h = -\frac{1}{3}, 0, \frac{1}{3}$. The rows (in order from top to bottom) have $\beta = \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}$. Notice that increasing β adds a "twist" to the curve, and changing h moves the twist up or down along the curve.

The theorem says that the magnetization M is concentrated around the zeroes of this function. For $\beta \leq 1$, there is exactly one zero of the function, so the magnetization will concentrate at that value. If $h = 0$, this value is zero, so the magnetization will be 0 and thus there will be about the same number of positive and negative spins. For positive values of h, there will be a higher proportion of positive spins in the ground state(s), and vice versa for negative values of h. This is a sort of "law of large numbers" type result for the magnetization at high temperatures.

As soon as $\beta > 1$, the twist in the function becomes extreme enough that there is a chance for more than one zero. When $h = 0$, there are two new zeroes, and they quickly approach -1 and $+1$ as β increases. However, there is still a zero at 0, and this means that the above theorem does not tell the whole story. In fact, it is known that for $\beta > 1$ and $h = 0$, the magnetization will concentrate around just the two nonzero zeroes of the function.

Now, if $\beta > 1$ and h is nonzero, the twist in the curve will be shifted up or down. This usually means that there will be only one root again, and it will be close to either -1 or $+1$, depending on the sign of h. This makes sense, since the external magnetic field will force one of the spins to be much more likely. However, as can be seen in the last row of plots on the previous page, if the external field h is sufficiently weak, then there is a chance for the opposite spin to win, and since constant spin is much more favorable (due to low temperature), if we run a markov chain simulation, the opposite spin will still be stable.

I'm not sure what happens with the third root in this situation, since it seems that by some sort of continuity, at least two of the roots should be possible, but there is some merging of roots at some values of h. It should also be mentioned that the above behavior can only possibly hold when the external field h is less than 1 in absolute value. This is because, even for very low temperatures (i.e. high β), the twist in the curve cannot make the total variation of the curve increase by more than 1 (as compared to a straight diagonal line), see the plot below with $\beta = 50$:



4 Proof of Theorem

We will use a slight generalization of the theorem from last time on exchangeable pairs, and I won't discuss the proof of this result at all.

Theorem 2. *Let (X, X') be an exchangeable pair of random variables (on a Polish space). Let F be an antisymmetric function and define*

$$f(X) = \mathbb{E}[F(X, X')|X].$$

If $\mathbb{E}[e^{\theta f(X)}|F(X, X')|] < \infty$ for all $\theta \in \mathbb{R}$ and there are constants $B, C \geq 0$ such that

$$\frac{1}{2}\mathbb{E}\left[|(f(X) - f(X'))F(X, X')|\,|\,X\right] \leq Bf(X) + C,$$

then for all $t > 0$,

$$\mathbb{P}[f(X) > t] \leq \exp\left(\frac{-t^2}{2C + 2Bt}\right) \qquad \text{and} \qquad \mathbb{P}[f(X) \leq -t] \leq \exp\left(\frac{-t^2}{2C}\right).$$

The main point of this generalization is that X doesn't have to be a real-valued random variable. We want to apply this to the exchangeable pair (σ, σ') corresponding to a step in the Glauber dynamics markov chain for the Curie-Weiss model. Then we will apply a function f, and we hope that

$$f(\sigma) = M - \tanh(\beta M + \beta h).$$

(In fact, this will not quite be true, but it will be close).

More precisely, let σ be chosen according to the Gibbs measure, and define σ' as follows: choose one of the n sites uniformly at random (call it I), and then resample the spin σ_I at that site, conditional on all of the spins at all the other sites. This is one step of the Glauber dynamics markov chain which is known to be reversible, so the pair (σ, σ') is exchangeable. Ross's notes call this the *Gibbs sampler* markov chain.

Now we need to decide on an antisymmetric function F. We want to study the magnetization $M = \frac{1}{n}\sum_{i=1}^n \sigma_i$, so we should try to make sure that $f(\sigma) = \mathbb{E}[F(\sigma, \sigma')|\sigma]$ involves M. Naïvely, if we wanted $f(\sigma) = M(\sigma)$, we could take $F(\sigma, \sigma') = M(\sigma)$. The problem is that this is not antisymmetric, so let's take

$$F(\sigma, \sigma') = M(\sigma) - M(\sigma') = \frac{1}{n}\sum_{i=1}^n (\sigma_i - \sigma_i').$$

But actually this has another problem: the ith summand will be zero unless the randomly chosen index I is i. This has probability $\frac{1}{n}$, so the conditional expectation of the above sum will actually involve $\frac{1}{n^2}M(\sigma)$, and thus it is off by a factor of $\frac{1}{n}$, and we should actually take

$$F(\sigma, \sigma') = \sum_{i=1}^n (\sigma_i - \sigma_i').$$

This is the final version of F that will work for us.

Let's check what function f this actually gives us. If the chosen site I is i, then

$$\mathbb{P}[\sigma_i' = \pm 1|I = i, \sigma] = \frac{\mathbb{P}[\sigma_i' = \pm 1, I = i, \sigma]}{\mathbb{P}[I = i, \sigma]} = \frac{\mathbb{P}[\sigma_i' = \pm 1, I = i, \sigma]}{\mathbb{P}[\sigma_i' = 1, I = i, \sigma] + \mathbb{P}[\sigma_i' = -1, I = i, \sigma]}. \qquad (*)$$

Now note that

$$\mathbb{P}[\sigma_i' = \pm 1, I = i, \sigma] = \frac{1}{Z}\exp\left(\frac{\beta}{n}\left(\sum_{k<j, j\neq i}\sigma_j\sigma_k \pm \sum_{j\neq i}\sigma_j\right) + \beta h\sum_{j\neq i}\sigma_j \pm \beta h\right).$$

The only place the $+$ version differs from the $-$ is where we see the $\pm$ sign above, so in computing the fraction $(*)$, everything that doesn't have a $\pm$ sign will cancel, and we obtain

$$\mathbb{P}[\sigma_i' = \pm 1|I = i, \sigma] = \frac{\exp\left(\pm\frac{\beta}{n}\sum_{j\neq i}\sigma_j \pm \beta h\right)}{\exp\left(\frac{\beta}{n}\sum_{j\neq i}\sigma_j + \beta h\right) + \exp\left(-\frac{\beta}{n}\sum_{j\neq i}\sigma_j - \beta h\right)}.$$

Therefore

$$\begin{aligned}
f(\sigma) &= \mathbb{E}[F(\sigma,\sigma')|\sigma] \\
&= \sum_{i=1}^{n} \mathbb{P}[I=i]\cdot \mathbb{E}[F(\sigma,\sigma')|I=i,\sigma] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\sigma_i - \sigma_i'|I=i,\sigma] \\
&= \frac{1}{n}\sum_{i=1}^{n}\sigma_i - \frac{1}{n}\sum_{i=1}^{n}\tanh\left(\frac{\beta}{n}\sum_{j\neq i}\sigma_j + \beta h\right),
\end{aligned}$$

recalling that $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. Writing $M_i = M_i(\sigma) = \frac{1}{n}\sum_{j\neq i}\sigma_j$, this is

$$f(\sigma) = M - \frac{1}{n}\sum_{i=1}^{n}\tanh(\beta M_i + \beta h).$$

Again, this is not exactly what we want, but notice that $|\tanh(x) - \tanh(y)| \le |x-y|$ which follows from the fact that the derivative of $\tanh(x)$ is always between 0 and 1, we have

$$\left|\frac{1}{n}\sum_{i=1}^{n}\left(\tanh(\beta M_i + \beta h) - \tanh(\beta M + \beta h)\right)\right| \le \sum_{i=1}^{n}|\beta M_i - \beta M| \le \frac{\beta}{n}. \qquad (**)$$

This is where the $\frac{\beta}{n}$ in the statement of Theorem 1 comes from; we will use the triangle inequality after applying Theorem 2 to our function f.

Now we need to check that the conditions of Theorem 2 are satisfied. The condition involving the MGF is obvious (according to Ross) because all of the quantities involved are finite. Thus we just need to find $B, C > 0$ such that

$$\frac{1}{2}\mathbb{E}\left[|(f(\sigma) - f(\sigma'))F(\sigma,\sigma')|\,|\,\sigma\right] \le Bf(\sigma) + C.$$

Since at most one of the spins changes during a Glauber dynamics step, we have $|F(\sigma,\sigma')| \le 2$. Also, using the inequality $|\tanh(x) - \tanh(y)| \le |x-y|$ again, we have

$$|f(\sigma) - f(\sigma')| \le |M(\sigma) - M(\sigma')| + \frac{\beta}{n}\sum_{i=1}^{n}|M_i(\sigma) - M_i(\sigma')| \le \frac{2}{n} + \frac{2\beta}{n^2} \le \frac{2(1+\beta)}{n}.$$

Thus the condition in Theorem 2 is satisfied with $B = 0$ and $C = \frac{2(1+\beta)}{n}$, and Theorem 2 now says that

$$\mathbb{P}\left[|f(\sigma)| > \frac{t}{\sqrt{n}}\right] \le 2\exp\left(-\frac{t^2}{4(1+\beta)}\right).$$

Now applying the triangle inequality with $(**)$ gives

$$\mathbb{P}\left[|M - \tanh(\beta M + \beta h)| \ge \frac{\beta}{n} + \frac{t}{\sqrt{n}}\right] \le 2\exp\left(-\frac{t^2}{4(1+\beta)}\right),$$

which finishes the proof of Theorem 1.