THE LEE-YANG THEORY OF PHASE TRANSITIONS

PIYUSH SRIVASTAVA

Magnetic materials in nature exhibit a sharp *phase transition*: there exists a temperature (the *Curie temperature*) just above which the spontaneous magnetism of a magnetic material is lost. The modeling of such discontinuous behavior presented a major conundrum for statistical physics in the early part of the 20th century. One of the first models of magnets proposed to qualitatively model this phenomenon envisioned a magnet as a lattice of points, with a magnetic "spin"—that could be aligned in a "$+$" or "$-$" direction—at each vertex of the lattice. The spins were modeled as interacting with their neighbors so that neighbors favored having the same spin, and these interactions induced a probability distribution over the possible *configurations*, i.e. assignments of $+/-$ spins to the sites. The interactions themselves were modeled as becoming weaker with increase in a temperature parameter β.[1] In this picture, the magnetization $M(\beta)$ of the material was modeled as the fraction of the sites that could be expected to be assigned $+$, so that a fraction of $1/2$ signified no magnetism, and phase transitions were expected to manifest as discontinuities or singularities of the function $M(\beta)$ or its successive derivatives.

We shall formally describe the above model (which is known as the Ising model [Isi25], but was proposed by Lenz) in the next section. However, it should not be hard to imagine that in any reasonable version of the model, discontinuity in the behavior of $M(\beta)$ can not be expected to arise in finite graphs. One way around this obstacle is to consider the limit of $M(\beta)$ as the size of the lattice grows to infinity (this is referred to as the "infinite volume" limit). The first such calculation was done by Ising [Isi25], who considered the "one dimensional lattice" $\mathbb{Z}$, and showed that that the infinite volume limit of $M(\beta)$ was also a *real analytic* function in this case—there were *no* phase transitions. This engendered skepticism[2] about whether phase transitions could be expected at all in the Ising model, and whether it was a reasonable qualitative model of magnetism at all. This skepticism was not completely laid to rest till the celebrated *tour-de-force* of Onsager [Ons44], who explicitly calculated the infinite volume limit for the two dimensional lattice $\mathbb{Z}^2$, and showed that its second derivative indeed had a point of singularity.

However, another feature of magnetic materials is that their magnetization is influenced by external magnetic fields, and that the magnetization does *not* typically show a phase transition with respect to the magnetic field. It is not hard to account for an external magnetic field in the Ising model; this is modeled as a bias parameter λ which corresponds to the tendency of sites to align themselves to the $+$ direction. The previously cited results were, however, restricted to the "zero field" case, and shed little light on the behavior of the model under a magnetic field. In 1952, Lee and Yang [YL52, LY52] related this problem to the study of the location of the zeros of a polynomial called the *partition function* that is canonically associated with such models. Their work showed that the Ising model indeed does not show a phase transition with respect to the external field at any non-zero value of the field. In doing so, they also founded what proved to be a very fruitful branch of the stability theory of polynomials, one that continues to be an active area of research in both mathematics and statistical physics. In this short note, we survey how this connection to stability theory arose in the work of Lee and Yang.

Date: October, 2013. Minor corrections made in January 2016.

Computer Science Division, UC Berkeley. Email: `piyushsriva@gmail.com`. Some parts of the appendix to this informal note were also incorporated in primer written by Nisheeth Vishnoi for a workshop on "Zeros of Polynomials and their Applications" at the *IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, 2013.

[1] Our notation here is different from the standard notation in statistical physics, where β denotes the inverse temperature.

[2] A vote was apparently taken at the van der Waals Centenary Conference in 1937 about whether phase transitions could be demonstrated in such models through the use of infinite volume limits. Although the vote succeeded, it has been described as "close" [Sim93, pg. 6].

1. The Ising and Monomer-Dimer models

Partition functions are central objects of study in both combinatorics and statistical physics. Combinatorially, a partition function can be seen as a generating polynomial for a class of combinatorial objects. On the other hand, in statistical mechanics, the partition function encodes several important properties of the system under consideration. We begin by looking at the partition functions of two canonical models from statistical physics; the first of which is the *Ising model* [Isi25] described in the introduction. Given a graph $G = (V, E)$, the Ising model is a probability distribution over *spin* assignments $\sigma : V \to \{+, -\}$ to the vertices. The *temperature* of the system is modeled through a parameter β ($0 < \beta < 1$), and the preference for particular spins (the *magnetic field*) through another parameter λ in the following way: the probability of a spin assignment (or *configuration*) σ is proportional to its *weight* $w(\sigma)$ defined as

$$w(\sigma) := \beta^{d(\sigma)}\lambda^{m(\sigma)}, \tag{1}$$

where $d(\sigma)$ is the number of edges $e = \{u, v\} \in E$ such that $\sigma(u) \neq \sigma(v)$ and $m(\sigma)$ is the number of vertices which are assigned a '$+$' spin by σ. The partition function of the Ising model is then defined as

$$Z_\beta(\lambda) := \sum_{\sigma\in\{+,-\}^V} w(\sigma) = \sum_{\sigma\in\{+,-\}^V} \beta^{d(\sigma)}\lambda^{m(\sigma)}, \tag{2}$$

so that it is a polynomial of degree at most $|V|$ in λ. The parameter λ is also often referred to as the *vertex activity* in the literature.

Before proceeding further, we consider qualitatively the effect of the temperature parameter β. Note that since $0 < \beta < 1$, configurations with smaller number of "cut" edges tend to get larger weights. This corresponds to a *ferromagnetic* interaction: neighbors tend to have the same spins. The boundary case $\beta = 0$ corresponds to "zero temperature", in which the only possible configurations are those in which all sites have the same spins ("frozen" configurations), whereas the case $\beta = 1$ corresponds to "infinite temperature", so that the interactions between neighbors have no effect and the different sites behave like independent random variables.

The *monomer-dimer* model [HL72] is similarly a distribution over matchings in a graph $G = (V, E)$. The tendency of vertices to remain as monomers as opposed to participating in a matching is modeled through a parameter λ (which is sometimes referred to as the *fugacity* in the statistical physics literature). The weight of a matching M is given by

$$w(M) := \lambda^{u(M)}, \tag{3}$$

where $u(M)$ is the number of vertices left unmatched ("monomers") by M. The partition function, as before, is a polynomial in λ given by

$$Z(\lambda) := \sum_{M:\text{matching}} w(M) = \sum_{M:\text{matching}} \lambda^{u(M)}. \tag{4}$$

Qualitatively, the parameter λ represents the tendency of vertices to remain unmatched (or to stay as "monomers") as opposed to participating in a matching (forming "dimers"). The higher the value of λ, the more is the chance that a given vertex remains unmatched.

2. Free energy and phase transitions

The function $F_\beta(\lambda) = -\frac{1}{|V|}\log Z_\beta(\lambda)$ is called the *free energy* per unit volume of the model under consideration, and turns out be very important in the study of phase transitions. The reason is that important physical quantities (whose phase transition one might be interested in) can be written in terms of derivatives of the free energy. For example, consider the average fraction of monomers $U(\lambda)$ in the monomer dimer model. We can write

$$U(\lambda) := \frac{1}{V} \cdot \frac{\sum_{M:\text{matching}} u(m)\lambda^{u(M)}}{\sum_{M:\text{matching}} \lambda^{u(M)}} = \frac{1}{|V|Z(\lambda)}\, z\frac{\partial}{\partial z}Z(z)\bigg|_{z=\lambda} = -\, z\frac{\partial F(z)}{\partial z}\bigg|_{z=\lambda}. \tag{5}$$

Similarly, we can show that in the Ising model, $M_\beta(\lambda)$, the average "magnetization" of the graph under a "magnetic field" λ, normalized by the number of vertices (the "volume") can be written as

$$M_\beta(\lambda) := \frac{1}{|V|} \cdot \frac{\sum_{\sigma \in \{+,-\}^{|V|}} m(\sigma) \beta^{d(\sigma)} \lambda^{m(\sigma)}}{Z_\beta(\lambda)} = \frac{1}{|V| Z_\beta(\lambda)} \, z \frac{\partial}{\partial z} Z_\beta(z) \bigg|_{z=\lambda} = - \, z \frac{\partial F_\beta(z)}{\partial z} \bigg|_{z=\lambda}. \quad (6)$$

A *phase transition* of a physical system is defined as a discontinuity or singularity of the observable being considered (in our case, these are the quantities $M_\beta(\lambda)$ and $U(\lambda)$), or of one of its derivatives. Given the definitions of M and U above in terms of the free energy, we therefore see that:

(1) If the free energy $F(z)$ is an analytic function for $z \in D$, then, by definition, the system *cannot* have a phase transition in the region D of the parameter space.
(2) Since the only "physically" interesting region of the parameter space is $\lambda > 0$, and the free energy of any *finite* graph is clearly an analytic function for $\lambda > 0$, *neither* of the above systems can exhibit a phase transition on a *finite* graph. This was already hinted at informally in the introduction.

3. Infinite graphs and the zeros of the partition function

The fact that finite graphs cannot exhibit phase transitions led physicists in the second quarter of the 20th century to turn their attention towards infinite graphs. The most important such graph was the 2D integer lattice $\mathbb{Z}^2$. In order to define the free energy for such a graph, we take a increasing sequence of finite graphs G_n which "approach" the given graph "in the limit", and define the free energy $F(\lambda, G)$ of G as $\lim_{n\to\infty} F(\lambda, G_n)$. In the case of the lattice, the natural choice is to take G_n to be squares of increasing side length centered at the origin.[3] Phase transitions can indeed occur in this setting; in a celebrated paper [Ons44] Onsager explicitly calculated the free energy (as a function of β) of the Ising model on $\mathbb{Z}^2$ in the "zero field" case ($\lambda = 1$), and showed that there is a phase transition in the β parameter: the second derivative of the free energy (with respect to β) has a point of singularity.

As we saw in the introduction, this left open the question of phase transitions in the λ parameter. Yang and Lee [YL52] considered phase transitions in the λ parameter, and sought to identify regions where phase transitions cannot occur. They proved the following general result which relates the analyticty of the free energy of an infinite graph to the location of zeros of partition functions of finite graphs.

Theorem 3.1 (Yang and Lee [YL52]). *Let the free energy $F(z, G)$ of an infinite graph be defined as the limit $F(z, G) = \lim_{n\to\infty} F(z, G_n)$, where G_n is a sequence of* finite graphs. *If S is an open region in the complex plane such that the polynomials $Z(z, G_n)$ in z do not have any zeros in S, then $F(z, G)$ is analytic in S. In particular, this implies that there are no phase transitions on any open interval R of the positive real line contained inside S.*

We do not give a proof of this statement here, but give some intuitive reasons for why one can expect it to be true. Recall that the free energy $F(z, G_n)$ was defined as $-\frac{1}{|V_n|} \log Z(n, G_n)$, and phase transitions correspond to discontinuities or singularities of the limit $F(z, G)$ or its derivatives. However, we recall from complex analysis that if we can show that a function is complex analytic in a domain D, then all its derivatives exist (and hence, in particular, are continuous) in D. Since $F(z, G)$ is defined as the limit of the logarithms of polynomials $Z(z, G_n)$, we can reasonably expect that it would be analytic in regions in which the polynomials Z are zero-free (since the logarithm function is analytic everywhere except the non-positive part of the real line). The proof of Yang and Lee is a rigorous justification of this intuition.

4. Exploiting the connection: The Lee-Yang and Heilmann-Lieb theorems

Theorem 3.1 provides an excellent motivation for the study of zeros of partition functions. Its first application appeared in a subsequent paper of Lee and Yang [LY52]—published in the same journal

[3] Yang and Lee [YL52] showed that for integer lattices, the value $F(\lambda, G)$ calculated in this way does not depend upon the sequence G_n (their result also holds for more general classes of graphs).

issue—where they proved the following striking stability result for the partition function of the ferromagnetic Ising model.

Theorem 4.1 (Lee and Yang [LY52]). *Let G be any finite graph, and let $Z_\beta(z)$ be the partition function of the ferromagnetic Ising model on G with temperature parameter $\beta \in (0,1)$. Then, all zeros of $Z_\beta(z)$ lie on the circle $|z| = 1$ in the complex plane.*

Combining Theorem 4.1 with Theorem 3.1 immediately shows that the ferromagnetic Ising model exhibits *no* phase transitions except possibly at $\lambda = 1$ (since any open region of the positive real line not containing the point $\lambda = 1$ is contained in a open region of the complex plane not containing any partition function zeros). The proof of the theorem is not very hard; we will give almost a full proof (based on ideas in later work of Asano [Asa70]) in Appendix A.

Another classic example of the Lee-Yang program of exhibiting non-existence of phase transitions is a later result of Heilmann and Lieb [HL72] for the monomer-dimer model, who proved the following stability result.

Theorem 4.2 ([HL72]). *Let G be any graph and let $Z(z)$ be the partition function of the monomer-dimer model on G. Then, all zeros of $Z(z)$ lie on the imaginary line in the complex plane.*

Again, combined with Theorem 3.1, we immediately see that there are no phase transitions in the monomer dimer model except possibly at $\lambda = 0$.

5. Conclusion

We saw two classical examples of the Lee-Yang program of understanding the location of zeros of partition functions being used in the study of phase transitions in statistical physics. These examples are by no means isolated, and similar methods have been applied to a variety of models; see, e.g., Asano [Asa70], Suzuki and Fisher [SF71] and Newman [New74].

The Lee-Yang framework has also inspired work not immediately related to the study of phase transitions. Newman [New74] also studied correlation inequalities in the same framework, and in a later paper [New91], pointed at connections between such correlation inequalities and the Riemann hypothesis. Another line of work has looked at the location of zeros of other combinatorial polynomials; see, e.g., Choe et al. [COSW04], Sokal [Sok04], and Seymour and Chudnovsky [CS07]. In a somewhat different line of work, Ruelle [Rue10] characterized *all* multivariate polynomials that satisfy the conclusion of the multivariate version of the Lee-Yang theorem.

From a computer science perspective, the Lee-Yang framework has applications to the computational complexity of computing "averages" such as the magnetization in the ferromagnetic Ising model. For a discussion of an extension to the Lee-Yang theorem which proves useful for this problem, the reader is referred to [SS13].

6. Acknowledgments

I thank Nisheeth Vishnoi for helpful comments on the presentation in this note, and Alistair Sinclair for collaborating on related questions. I also thank Deepak Dhar for pointing out some inconsistencies between the use of the term "discontinuity" in an earlier version of this note and in the statistical mechanics literature.

References

[Asa70] T. Asano. Lee-Yang theorem and the Griffiths inequality for the anisotropic Heisenberg ferromagnet. *Physical Review Letters*, 24(25):1409–1411, June 1970.

[COSW04] Y.-B. Choe, J. G. Oxley, A. D. Sokal, and D. G. Wagner. Homogeneous multivariate polynomials with the half-plane property. *Advances in Applied Mathematics*, 32(1–2):88–187, February 2004.

[CS07] M. Chudnovsky and P. Seymour. The roots of the independence polynomial of a clawfree graph. *Journal of Combinatorial Theory, Series B*, 97(3):350–357, May 2007.

[HL72] O. J. Heilmann and E. H. Lieb. Theory of monomer-dimer systems. *Communications in Mathematical Physics*, 25(3):190–232, 1972.

[Isi25] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift fur Physik*, 31:253–258, February 1925.

[LY52] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. *Physical Review*, 87(3):410–419, August 1952.

[New74] C. M. Newman. Zeros of the partition function for generalized Ising systems. *Communications on Pure and Applied Mathematics*, 27(2):143–159, March 1974.

[New91] C. M. Newman. The GHS inequality and the Riemann hypothesis. *Constructive Approximation*, 7(1):389–399, December 1991.

[Ons44] L. Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Physical Review*, 65(3-4):117–149, 1944.

[Rue88] D. Ruelle. Is our mathematics natural? The case of equilibrium statistical mechanics. *Bulletin (New Series) of the American Mathematical Society*, 19(1):259–268, July 1988.

[Rue10] D. Ruelle. Characterization of lee-yang polynomials. *Annals of Mathematics*, 171(1):589–603, March 2010.

[SF71] M. Suzuki and M. E. Fisher. Zeros of the partition function for the Heisenberg, ferroelectric, and general Ising models. *Journal of Mathematical Physics*, 12(2):235–246, February 1971.

[Sim93] B. Simon. *The Statistical Mechanics of Lattice Gases*, volume 1. Princeton University Press, 1993.

[Sok04] A. D. Sokal. Chromatic roots are dense in the whole complex plane. *Combinatorics, Probability and Computing*, 13(2):221–261, 2004.

[SS13] A. Sinclair and P. Srivastava. Lee-Yang theorems and the complexity of computing averages. In *Proc. 45th Annual ACM Symposium on Theory of Computing (STOC)*, pages 625–634. ACM, 2013. To apear in *Communications in Mathematical Physics*.

[YL52] C. N. Yang and T. D. Lee. Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Physical Review*, 87(3):404–409, August 1952.

Appendix A. Asano's proof of the Lee-Yang theorem

The proof we give here is due to Asano [Asa70], who simplified the original proof of Lee and Yang [LY52] while proving a version of the Lee-Yang theorem for the quantum Ising model. As usual in proving stability results, it helps to consider the more general case of the multivariate partition function. Let G be a graph on n vertices. Instead of having the same vertex activity λ for each vertex (as we did in eq. (1)), we will have different vertex activities λ_i at the vertices v_i. The partition function Z_β is then a multivariate polynomial in the λ_i's, and is defined as

$$Z_\beta(\lambda_1, \lambda_2, \dots, \lambda_n) = \sum_{\sigma \in \{0,1\}^n} \beta^{d(\sigma)} \prod_{v:\sigma(v)=+} \lambda_v, \tag{7}$$

where, as before, $d(\sigma)$ denotes the number of edges which have different spins at their endpoints in the spin assignment σ.

Theorem A.1. *Let G be a any finite graph. Fix a temperature parameter $\beta \in (0,1)$. If we have $|\lambda_i| > 1$ for each i, then*

$$Z_\beta(\lambda_1, \lambda_2, \dots, \lambda_n) \neq 0.$$

Theorem 4.1 follows from the above theorem by first setting all λ_i to be equal (which already shows that $Z_\beta(\lambda) \neq 0$ for $|\lambda| > 1$). We then note that $Z_\beta(1/\lambda) = \frac{1}{\lambda^n} Z_\beta(\lambda)$ (this can be seen as a consequence of the symmetry between $+$ and $-$ spins), so that we also get $Z_\beta(\lambda) \neq 0$ for $|\lambda| < 1$. Before proceeding with the proof of the above theorem, we also remark that both Asano and Lee and Yang actually proved a slightly stronger multivariate version of the theorem under the additional constraint that the graph G is connected: in particular, they showed that the partition function is non-vanishing even under the weaker conditions $|\lambda_i| \geq 1$ for all i and $|\lambda_j| > 1$ for at least one j. However, that version requires a somewhat more careful induction, so we only prove the weaker version stated above. For the application to the univariate case, both versions are equally strong.

The main idea used in the proof is often called the *Asano contraction* (see for example, Ruelle's article [Rue88], whose presentation we loosely follow). Consider a graph H on n vertices, and identify two vertices v_1 and v_2 in G that are not connected by an edge. Let λ_1, λ_2 be variables representing the vertex activities at v_1 and v_2. The vertex activities of the other vertices in H are denoted $\lambda_3, \lambda_4, \dots, \lambda_n$.

We say that a graph G satisfies the *Lee-Yang property* if the partition functions of all the induced subgraphs of G satisfy the conclusion of the Lee-Yang theorem (that is, whenever all the vertex activities have magnitude greater than 1, the partition function is non-zero). The reason for considering the partition functions of all induced subgraphs is technical and will become clear later in the proof.

Now suppose that a graph H satisfies the Lee-Yang property. The Asano contraction lemma states then that the graph H' obtained by contracting the vertex v_1 and v_2 into a single vertex v also obeys the Lee-Yang property. To see this, we first ignore the issue of the induced subgraphs and consider the partition functions of H and H' themselves. We can write the partition function of H as

$$A\lambda_1\lambda_2 + B\lambda_1 + C\lambda_2 + D, \tag{8}$$

where A, B, C, D are polynomials in the other vertex activities $\lambda_3, \ldots, \lambda_n$. The first, and crucial, observation is that the partition function of the new graph H' is then simply

$$A\lambda + D, \tag{9}$$

which follows from a consideration of the definition of the Ising model partition function.

Now, consider any fixing of values of $\lambda_3, \ldots, \lambda_n$ such that they are all greater than 1 in magnitude. We would be done if we can show that this implies that the expression in (9) is non-zero for $|\lambda| > 1$. Now, since the expression in eq. (8) satisfies the Lee-Yang property, we see by substituting these values into (8) and setting $\lambda_1 = \lambda_2 = x$ that the quadratic equation $Ax^2 + (B+C)x + D = 0$ has no solution with $|x| > 1$. In particular, this implies that the product of its zeros, D/A, must have magnitude at most 1. But this implies that for the expression in (9) to be zero, we must have $|\lambda| = |D/A| \leq 1$, which is what we wanted to prove.[4] Essentially the same argument applies to the partition functions of the induced subgraphs of H'; we omit the details.

Armed with the Asano contraction, we now proceed to prove the Lee-Yang theorem for a given graph G. We first consider the case of a graph of just one vertex; in this case the partition function is simply $1 + \lambda_1$, and hence the Lee-Yang property is trivially satisfied. We now consider the first important case, that of a single edge. In this case, the partition function is $1 + \beta(\lambda_1 + \lambda_2) + \lambda_1\lambda_2$. Now suppose $|\lambda_1| > 1$. For the partition function to be 0, we must have

$$|\lambda_2| = \left|\frac{1 + \beta\lambda_1}{\beta + \lambda_1}\right| < 1, \tag{10}$$

where the last equation use the fact that for $\beta < 1$, the above fraction is a Möbius transformation that takes the exterior of the unit disk to its interior. These two observations show that a single edge has the Lee-Yang property.

We now notice that if two graphs G_1 and G_2 on disjoint sets of vertices have the Lee-Yang property, then so does the graph $G' = G_1 \cup G_2$, this follows simply because the partition function of G' is simply the product of the partition functions of G_1 and G_2 (and a similar argument applies to their subgraphs). Now, to prove that a given graph G of m edges has the Lee-Yang property, we start with a disjoint union of m edges, each corresponding to one of the edges of G; note that this "graph" has the Lee-Yang property. At each stage, we take the end-points of two edges which are supposed to be incident on the same vertex in G, and merge these end-points. Note that this process ultimately ends up with the graph G. By the Asano contraction lemma, the resulting graph after each of these steps continues to have the Lee-Yang property, and so G itself has the Lee-Yang property.

[4] We implicitly assumed in the above argument that for our fixing of values of $\lambda_3, \ldots, \lambda_n$, we have $A \neq 0$. This technicality can be avoided by noticing that A itself can be seen as the partition function of the Ising model on a induced subgraph of H obtained by removing v_1 and v_2, and hence satisfies the Lee-Yang property. We omit the details in this short note. This technicality, however, is the reason why we needed the stronger hypothesis involving all subgraphs.