Chapter 31

Comparison of Ising models under change of a priori measure

José Madrid, Barry Simon, and Daniel R. Wells

Dedicated with pleasure to Elliott Lieb on his 90th birthday

We study comparison of correlation functions for ferromagnetic generalized Ising models with two different a priori measures. One purpose of this note is to publicize some unpublished 45 year old work of Daniel Wells on the issue. We then prove results for the a priori measures associated to one component of D-vectors uniformly distributed on the unit sphere and also the case of spin S ($2S + 1$ equally spaced values symmetric about 0 and with equal weights) that improves some 50 year old bounds of Griffiths on transition temperatures.

1 Introduction

Besides Elliott Lieb's many major accomplishments, there are numerous gems that sparkle even though they are not among his most important. In this note, we want to discuss something related to his beautiful note [13] on the infinite spin limit of the pressure of quantum Heisenberg models which converges to a classical Heisenberg model, a work which motivated the second author's extension [17] to more general Lie groups than SU(2). Lieb proved comparison inequalities for partition functions that squeezed the spin S quantum partition function between the corresponding classical partition functions with slightly different coupling constants so that the difference of the coupling constants goes to zero as $S \to \infty$. In the totally anisotropic case (where only z components are coupled), Dyson, Lieb and the second author (we never published this work done in 1976 but it was included it in the 1993 book of Simon [18, Section II.9]) proved monotonicity (increasing) of the partition function in S and decreasing monotonicity if the coupling is scaled properly.

When we were working on Thomas–Fermi, Elliott taught the second author about the magic of convergence of convex functions – that convergence of convex functions implies convergence of derivatives at points where the limit is differentiable. This

implies convergence of certain correlation functions in the context that Lieb studied in [13]. But one loses for correlation functions inequalities like those that Dyson, Lieb and Simon found for partition functions. It is that question that we want to discuss here. Our framework will be less general in that we will only consider one component spins and more general in that we will allow general (even) a priori measures.

The second author is writing a book for Cambridge Press entitled *Phase transitions in the theory of lattice gases* [21]. It is in many ways the successor to the 1993 book [18] from Princeton University Press. That earlier book was mainly framework and largely left out all the most fun and beautiful elements of the theory: correlation inequalities, Lee–Yang, Peierls' argument, Berezinskii–Kosterlitz–Thouless transitions and infrared bounds which are the subjects of the new book. But since a different publisher is used, this is certainly *not* volume 2 of the earlier work.

The framework for much of the subject is to fix a finite set $\Lambda \subset \mathbb{Z}^\nu$, and an a priori *even* probability measure, $d\mu$, on $\mathbb{R}$, certainly with all moments finite and typically of compact support.

One considers the configurations in Λ, i.e. points $\boldsymbol{\sigma}$ in $\mathbb{R}^\Lambda$, indicated by $\{\sigma_j\}_{j\in\Lambda}$ and uncoupled measure with expectation

$$\langle f \rangle_{\mu,0} = \int f(\boldsymbol{\sigma}) \prod_{j\in\Lambda} d\mu(\sigma_j)$$

and one fixes a ferromagnetic Hamiltonian (i.e. $J(A) \geq 0$)

$$-H = \sum_{A\subset\Lambda} J(A)\sigma^A, \quad \sigma^A = \prod_{j\in A} \sigma_j$$

or more general over multiindices, i.e. assignments of an integer, $n_j \geq 0$ with then $\sigma^A = \prod_{j\in A} \sigma_j^{n_j}$ (and a finite sum or else ℓ^1 condition). One then considers the Gibbs state

$$\langle f \rangle_{\mu,\Lambda} = Z^{-1}\langle fe^{-H} \rangle_{\mu,0}, \quad Z = \langle e^{-H} \rangle_{\mu,0}. \tag{1}$$

One studies the infinite volume limit with translation invariant $J(A)$, typically by proving stuff about the finite volume expectations. The traditional case is the Ising model (aka spin $\frac{1}{2}$ Ising model) where $d\mu$ is a measure supported on ± 1 each point with weight $\frac{1}{2}$; more generally, we will refer to b_T with weights $\frac{1}{2}$ at $\pm T$ (b is for Bernoulli). While a lot of the literature is specific to the spin $\frac{1}{2}$ Ising model, there is considerable, mathematically interesting, literature on more general (even) a priori measures. Traditionally, one mainly considered the spin S measure (for $S = \frac{1}{2}, 1, \frac{3}{2}, \ldots$, the measure with $2S + 1$ pure points equally spaced symmetrically about 0 and with equal weights) but the work of Guerra, Rosen and Simon [9] and Griffiths and Simon [22] on discrete approximations to Euclidean Quantum Field Theory changed that.

As the second author began to write about correlation inequalities in his new book, he wondered about a natural question. We say that an a priori measure, ν, on $\mathbb{R}$ *Ising dominates* another measure μ if and only if for all $J(A) \geq 0$ and all B, one has that

$$\langle \sigma^B \rangle_{\mu, \Lambda} \leq \langle \sigma^B \rangle_{\nu, \Lambda}. \tag{2}$$

In particular, for general μ compact support, does one have that μ Ising dominates b_{T_-} and is Ising dominated by b_{T_+} for suitable $0 < T_- < T_+ < \infty$. That would imply phase transitions occur for one a priori measure if and only if they do for all and inequalities on transition temperatures.

For most, even minor, aspects of the subject of correlation inequalities there are several papers, sometimes even dozens. So it is surprised that the second author was unable to find a single published paper on the subject of what we just called Ising domination! Of course, it was unclear how to search for the subject in Google. Eventually, we did find one 1978 paper of van Beijeren and Sylvester [24] that we will mention later (see Remark 1 (2) after Theorem 3.1) although in one respect it is unsatisfactory. And we did also find an appendix of a paper on another subject but that gets ahead of our story (see Remark 1 (1) after Theorem 3.1).

One of the pleasant things about writing a book on a subject that one once knew more about is that one gets to rediscover things that they have forgotten. With the question of Ising domination in the back of his mind, the second author found an interesting footnote in a 1980 paper of Aizenman and Simon [1]. The footnote said:

then by results of Wells (D. Wells, Some moment inequalities for general spin Ising ferromagnets, Indiana Univ. preprint) $\langle s_j s_k \rangle_{\beta, 1} \leq 2 \langle \sigma_j^{(1)} \sigma_k^{(1)} \rangle_{\beta, 2}$.

The left hand side is an Ising expectation and the right with the a priori measure of the two-dimensional rotor with only couplings of the 1 components. So this was part of what seems to be an Ising domination result (the subscript 2 indicates the Ising measure should really be $b_{1/\sqrt{2}}$).

So the second author set about finding this preprint. Google did not help directly but did point him to a 1984 paper of Chuck Newman that mentioned Wells' Indiana University PhD. thesis. He wrote to Michael asking if he knew anything about our footnote and cced Chuck (who had been a grad student with the second author at Princeton) because the second author conjectured Wells had been his student. Chuck replied and said he remembered that Wells had been Slim Sherman's student. Sherman, the S of GKS and GHS was a delightful character, long dead.

So the second author wrote to Kevin Pilgrim, the chair at Indiana, who located a copy of Wells thesis [25] on Proquest. So far though, no one has had any luck on the preprint nor on locating Wells through Indiana University alumni records (but see later)! While the thesis did not have anything directly about the above inequality, it did have a general framework on what we called the Ising domination problem,

lovely material that should have been published. After an initial draft of this note was written, the first two authors got some help and located Dr. Wells who kindly agreed to be a coauthor which makes sense since much of this chapter publishes for the first time results from his thesis. Also, we convinced him to allow us to continue to use the term *Wells domination*.

Our main goal in this chapter is to describe Wells' framework in Section 2 and what we regard as his most significant theorem in Section 3. Since Wells extended a framework of Ginibre, we begin Section 2 by reminding (telling) you of that. Then the notion we call Wells' domination followed by his big theorem in Section 3 and the notion of canonical lower bound. We will note there that the approach of van Beijeren and Sylvester [24] has one big flaw in that there is no analog of the Wells Comparison Theorem, Theorem 3.1. Section 4 will then make explicit the bounds on transition temperatures implied by Wells Comparison Theorem, recall a result of Griffiths [8] on comparison of transition temperatures for different spins and note that the bounds when a measure is canonical are equalities in mean field theory and so optimal in the high dimension limit. Section 5 has one of our two new results here – that the distribution of one component of an D-vector spin has a canonical lower bound. From the footnote in [1], it is clear that the missing Wells preprint has the case $D = 2$. We will see that case is much easier to prove than $D \geq 3$. In Section 6, we will prove our most significant new result that the spin S measures for $S \neq 1$ have a canonical lower bound and its consequence for improving Griffiths bound on transition temperatures for arbitrary spins. Finally, Appendix A provides the proof of a technical inequality.

Elliott Lieb has been a major figure in all the fields he has worked in and contributed through his work, his students and his energy. We hope he enjoys this birthday bouquet.

2 Wells framework

As mentioned, Wells' approach is a slight modification of Ginibre's approach to the proof of GKS inequalities for Ising-type models. In a remarkable 1970 paper [7], Jean Ginibre (who alas passed away in March of 2020 at age 82) not only found a really simple proof of GKS inequalities but showed somewhat surprisingly that they held for all (even) a priori measures. If you are new to Ising models and have time for only one result, this one might be what you should know.

A *Ginibre system* is a triple $\langle X, \mu, \mathcal{F} \rangle$ of a compact Hausdorff space, X, a probability measure, μ, on X (with expectations $\langle \cdot \rangle_\mu$) and a class of continuous real-valued functions $\mathcal{F} \subset C(X)$ that obeys

$$\forall_{f_1,\ldots,f_n \in \mathcal{F}} \int_X f_1(x) \ldots f_n(x)\, d\mu(x) \geq 0 \qquad \text{(G1)}$$

and

$$\forall_{f_1,\dots,f_n \in \mathcal{F}} \int_{X \times X} \prod_{j=1}^{n} \left(f_j(x) \pm f_j(y) \right) \, d\mu(x) d\mu(y) \geq 0 \qquad \text{(G2)}$$

for all 2^n choices of the plus and minus sign.

When it is clear which measure is intended, we will drop the μ from $\langle \cdot \rangle_\mu$. We have restricted to compact Hausdorff spaces and so bounded functions for simplicity. But since all the arguments are essentially algebraic, all results extend to the case where X is only locally compact so long as all $f \in \mathcal{F}$ obey $\int |f(x)|^m \, d\mu(x) < \infty$ for all m since that condition assures that all integrals below are convergent.

Note that

$$\text{(G2)} \implies 2\langle f \rangle_\mu = \int_X (f(x) + f(y)) \, d\mu(x) d\mu(y) \geq 0$$

and

$$\int_{X \times X} (f(x) - f(y))(g(x) - g(y)) \, d\mu(x) d\mu(y)$$
$$= 2[\langle fg \rangle_\mu - \langle f \rangle_\mu \langle g \rangle_\mu] \geq 0.$$

We will see shortly that (G2) $\Rightarrow$ (G1). What makes the notion so powerful is that there are three theorems for getting new Ginibre systems from old ones.

Given a family of functions, $\mathcal{F} \subset C(X)$, we define the *Ginibre cone*, $\mathcal{C}(\mathcal{F})$, as the set of linear combinations with non-negative coefficients of products of functions from $\mathcal{F}$.

Theorem 2.1 (Ginibre Theorem 1). *If a triple $\langle X, \mu, \mathcal{F} \rangle$ obeys estimate (G2), so does $\langle X, \mu, \mathcal{C}(\mathcal{F}) \rangle$.*

It is trivial that (G2) holds for sums and positive multiples of functions for which it holds, so it suffices to prove it holds for products. By induction, we need only handle products of two functions. We note that

$$fg \pm f'g' = \frac{1}{2}(f + f')(g \pm g') + \frac{1}{2}(f - f')(g \mp g')$$

which allows us to prove (G2) for a single product when we have it for individual functions (and shows (G2) $\Rightarrow$ (G1)).

The following is trivial:

Theorem 2.2 (Ginibre Theorem 2). *Let $\{\langle X_j, \mu_j, \mathcal{F}_j \rangle\}_{j=1}^n$ be a family of Ginibre systems. Then $\langle \times_{j=1}^n X_j, \bigotimes_{j=1}^n \mu_j, \bigcup_{j=1}^n \mathcal{F}_j \rangle$ is also a Ginibre system*

And to add interactions, we use the following result.

Theorem 2.3 (Ginibre Theorem 3). *Let $\langle X, \mu, \mathcal{F} \rangle$ be Ginibre system. Let $-H \in \mathcal{F}$ and define a new measure μ_H by*

$$\langle f \rangle_{\mu_H} = \frac{\langle fe^{-H} \rangle_\mu}{\langle e^{-H} \rangle_\mu}.$$

Then $\langle X, \mu_H, \mathcal{F} \rangle$ is a Ginibre system.

The proof is easy. The normalization is irrelevant and we expand the exponential $\exp(-H(x) - H(y))$. Finally:

Theorem 2.4 (Ginibre Theorem 4). *Let X be $\mathbb{R}$ or a compact subset of the form $[-A, A]$ and let $d\mu$ be a probability measure which is invariant under $x \mapsto -x$ and so that (only non-trivial in case X is not compact) $\int x^{2n} \, d\mu(x) < \infty$ for all n. Let $\mathcal{F}$ contain the single function, $f(x) = x$. Then $\langle X, \mu, \mathcal{F} \rangle$ is a Ginibre system.*

The proof is easy! Condition (G2) says that for all non-negative integers, k and m, one has that

$$\int_{X \times X} (x + y)^k (x - y)^m \, d\mu(x) d\mu(y) \geq 0.$$

Interchanging x and y implies the integral is zero if m is odd and $(x, y) \mapsto (-x, -y)$ symmetry implies the integral is zero if $m + k$ is odd. Thus the only possible non-zero integrals are when m and k are even in which case the integrand is positive!

A little thought shows that for Hamiltonians of the form

$$- H = \sum_{A \subset \Lambda} J(A)\sigma^A, \quad \sigma^A = \prod_{j \in A} \sigma_j \tag{3}$$

with *any* (!) even a priori measure, one has positive expectations and positive correlations of the σ^A which is GKS inequalities for general even measures.

We would be remiss if we left the subject Ginibre's wonderful paper without mentioning two other examples he gives of Ginibre systems that are not relevant to Wells, although one will appear later. The first is to note that he proves that if $d\mu$ is a product of rotation invariant measures on circles, the set of functions $\cos(\sum_{j=1}^n m_j \theta_j)$ is a Ginibre system. This and some extensions are essentially half the correlation inequalities for plane rotors.

The second is related to an 1882 paper of Chebyshev [3] (which I do not think Ginibre knew about when he wrote his 1970 paper) which contained what is probably the earliest correlation inequality: Chebyshev proved that if f, g are two monotone functions on $[0, 1]$, then

$$\int_0^1 f(x)g(x) \, dx \geq \int_0^1 f(x) \, dx \int_0^1 g(x) \, dx.$$

Ginibre proved that for any (not necessarily even) positive probability measure on $\mathbb{R}$, the set $\mathcal{F}$ of all positive monotone functions is a Ginibre family. The proof is again very easy. This is a sort of poor man's FKG inequalities.

This completes our review of Ginibre, so we turn to Wells' work. There is a simple extension of Ginibre's method in Wells' thesis [25] that allows comparison of measures. Given two probability measures, μ and ν on a locally compact space, X, we say that μ *Wells dominates* ν, written $\mu \triangleright \nu$ or $\nu \triangleleft \mu$ with respect to a class of continuous functions $\mathcal{F}$ (with all moments of all $f \in \mathcal{F}$ finite with respect to both measures; not needed if X is compact) if for all n and all $f_1, f_2, \ldots, f_n$ and all 2^n choices of $\pm$, we have that

$$\int \int (f_1(x) \pm f_1(y)) \ldots (f_n(x) \pm f_n(y)) \, d\mu(x) \, d\nu(y) \geq 0.$$

We will be most interested in case $X = \mathbb{R}$, μ and ν are both even measures with all moments finite and $\mathcal{F}$ has the single function $f(x) = x$ in which case the condition takes the form

$$\int_{\mathbb{R}} \int_{\mathbb{R}} (x + y)^n (x - y)^m \, d\mu(x) \, d\nu(y) \geq 0 \qquad (4)$$

for all non-negative integers, n and m in which case we use the symbol $\triangleleft$ without being explicit about $\mathcal{F}$. Since the measures are even, one need only check this when $n + m$ is even. It is trivial if both are even, so we only need worry about the case that both are odd. Since the measures are different, we do not have the exchange symmetry that makes the integral vanish if both are odd but symmetry under $y \mapsto -y$ implies invariance under interchange of m and n, so we need only check for $m \geq n$. We will see examples later.

Extending the Ginibre machine is effortless. It is easy to prove that:

Theorem 2.5 (Wells [25]). *The following statements hold:*

(a) *If $\mu \triangleleft \nu$ for a set of functions $\mathcal{F}$, the same is true for the Ginibre cone $\mathcal{C}(\mathcal{F})$.*

(b) *If for $j = 1, \ldots, n$, $\mu_j \triangleleft \nu_j$ for probability measures on spaces X_j with respect to sets of functions $\mathcal{F}_j$ on X_j, then for the measures on $\times_{j=1}^{n} X_j$ and the set of functions $\bigcup_{j=1}^{n} \mathcal{F}_j$, one has that $\bigotimes_{j=1}^{n} \mu_j \triangleleft \bigotimes_{j=1}^{n} \nu_j$.*

(c) *If $\mu \triangleleft \nu$ for probability measures on a space X with respect to a set of functions $\mathcal{F}$ on X, if $-H \in \mathcal{F}$ and if μ_H, ν_H are Gibbs measures, then $\mu_H \triangleleft \nu_H$ for $\mathcal{F}$.*

(d) *If $\mu \triangleleft \nu$ with respect to a set of functions $\mathcal{F}$, then for every $f \in \mathcal{F}$, we have that*

$$\int f(x) \, d\mu(x) \leq \int f(x) \, d\nu(x).$$

This immediately implies that:

Corollary 2.6 (Wells [25]). *If for $j = 1, \ldots, n$, $\mu_j \lhd \nu_j$ for probability measures on spaces X_j with respect to sets of functions $\mathcal{F}_j$ on X_j, then if $-H \in \mathcal{C}(\bigcup_{j=1}^{n} \mathcal{F}_j)$ and if μ_H, ν_H are formed from the underlying product measures $\bigotimes_{j=1}^{n} \mu_j$ and $\bigotimes_{j=1}^{n} \nu_j$, then for all $F \in \mathcal{C}(\bigcup_{j=1}^{n} \mathcal{F}_j)$, one has that $\int f(x)\,d\mu_H(x) \leq \int f(x)\,d\nu_H(x)$. In particular, if each $X_j = \mathbb{R}$ (so implicitly F_j is the single function σ_j) and if H has the general ferromagnetic Ising form, (3) with all $J(A) \geq 0$, then for all $A \subset 2^{\{1,\ldots,n\}}$ one has that*

$$\langle \sigma^A \rangle_{\mu_H} \leq \langle \sigma^A \rangle_{\nu_H}.$$

Thus by the definition, (2), of Ising domination, we see that if $\mu \lhd \nu$, then ν Ising dominates μ.

Of course, $\lhd$ is a binary relation and it is tempting to think of it as a partial order on measures on $\mathbb{R}$ with all moments finite. Indeed, it is certainly reflexive. It is almost antisymmetric. It is easy to see that $\mu \lhd \nu$ and $\nu \lhd \mu$ if and only if μ and ν have the same moments. Thus it is antisymmetric among the measures of compact support or among measures obeying $\int e^{Ax^2}\,d\mu(x) < \infty$ for some $A > 0$ but not among all measures with finite moments because of the possibilities of measures non-unique for the moment problem. But we do not know the following

Question 1. Is Wells relation transitive among all even measures on $\mathbb{R}$? How about among all measures on a general topological space if $\mathcal{F}$ is rich enough?

Since Ising domination is trivially transitive, for applications, this lack is not so important.

3 The Wells Comparison Theorem

Given an even measure probability, μ, on $\mathbb{R}$ and $s > 0$, we define its scaling by

$$\mu^{(s)}[A] = \mu[s^{-1}A].$$

Then the Bernoulli measure b_S defined after (1) obeys $b_S = (b_1)^{(S)}$.

Even if it is not true that $\nu \lhd \mu$, it can happen that $\nu^{(s)} \lhd \mu$ for s sufficiently small. In the next section, we will see that this implies a bound on transition temperatures, so such comparison results are interesting. The main result of this section implies that any two non-trivial measures of compact support are comparable in this sense. Well's most important result is:

Theorem 3.1 (Wells Comparison Theorem [25]). *Let $d\mu$ be an even probability measure on $\mathbb{R}$ with compact support that is not a point mass at 0. Then there are two*

strictly positive numbers, $T_-(\mu)$ and $T_+(\mu)$, so that $\mu \lhd b_S$ if and only if $S \geq T_+$ and $b_S \lhd \mu$ if and only if $S \leq T_-$. Moreover,

$$T_+ = \sup\{s \mid s \in \text{supp}(\mu)\} \tag{5}$$

and

$$S \leq T_- \iff \forall_{n \in \mathbb{N}} \int_{\mathbb{R}} (x^2 - S^2)^n \, d\mu(x) \geq 0. \tag{6}$$

Remark 1. (1) Bricmont-Lebowitz-Pfister [2] state the existence of T_- part of this theorem, quoting Wells and providing his proof.

(2) There is a very different order from Wells order defined by van Beijeren and Sylvester [24] (discussed further in [21, Section 2.2]) that also implies Ising domination but it has the serious flaw that if $0 \in \text{supp}(\nu)$, then for no $T > 0$ does ν dominate b_T in their order.

(3) The proof below is essentially that of Wells.

Lemma 3.2. *Let μ be a positive measure on an interval $I \subset \mathbb{R}$ (either open or closed at each endpoint). Let $f, g \in L^2(d\mu)$ and suppose that g is monotone increasing on I and there is $c \in I$ so that $f(x) \leq 0$ (resp $f(x) \geq 0$) if $x \leq c$ (resp $x \geq c$). Then*

$$\int f(x)g(x)\,d\mu(x) \geq g(c) \int f(x)\,d\mu(x).$$

Proof. The function $f(x)[g(x) - g(c)]$ is positive so its integral is positive which is the claim. ∎

Proof of Theorem 3.1. We first prove the existence of T_+ and equation (5). To this end, if $S \geq \sup\{s \mid s \in \text{supp}(\mu)\}$, then, for the integrand in (4) to be positive, we need that $(S + y)^n (S - y)^m + (S + y)^m (S - y)^n \geq 0$ for all $y \geq 0$ in $\text{supp}(\mu)$. If $\mu(\{0\}) > 0$, there is an additional term of $S^{n+m}\mu(\{0\})$ on the right-hand side, but that is also positive, so for such S, we have that $\mu \lhd b_S$.

On the other hand, if $\mu \lhd b_S$, we have that

$$\int x^{2N} \, d\mu(x) \leq S^{2N},$$

so, taking $2N$th roots and then $N \to \infty$, we see that $S \geq \sup\{s \mid s \in \text{supp}(\mu)\}$ which proves the formula for T_+.

Next we will prove that

$$b_S \lhd \mu \iff \forall_{n \text{ odd}} \int_{\mathbb{R}} (x^2 - S^2)^n \, d\mu(x) \geq 0. \tag{7}$$

Taking $n = m$ in the basic integral, we see that

$$b_S \lhd \mu \implies \forall_{n \text{ odd}} \int_{\mathbb{R}} (x^2 - S^2)^n \, d\mu(x) \geq 0.$$

Now look at the basic integral when $\nu = b_S$ and $m > n$ with both odd. Since $(x \pm S)^n (x \mp S)^m = (x^2 - S^2)^n (x \mp S)^{m-n}$, we see that the integral in question is

$$\frac{1}{2} \int (x^2 - S^2)^n [(x + S)^{m-n} + (x - S)^{m-n}] \, d\mu(x)$$

$$= \int (x^2 - S^2)^n [(x + S)^{m-n} + (x - S)^{m-n}] \, d\tilde{\mu}(x),$$

where $\tilde{\mu}$ is the measure restricted to $(0, \infty)$ plus $\frac{1}{2}\mu(\{0\})\delta_0$. By the binomial theorem, the polynomial $Q_{2k}(y) = (y + S)^{2k} + (y - S)^{2k}$ only has even degree terms with only positive coefficients so the function in $[\cdot]$ in the last equation is monotone on $I = [0, \infty)$. Applying the lemma with $c = S$, we see that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} (x + y)^n (x - y)^m d\mu(x) d\nu(y) \geq (2S)^{m-n} \int_{\mathbb{R}} (x^2 - S^2)^n \, d\mu(x).$$

Thus, we have proven (7).

Finally, we show that $T_- > 0$. First, pick $a > 0$ so that $\mu([a, \infty)) > 0$. Pick $0 < b < a$ so small that

$$\frac{b^2}{a^2 - b^2} \leq \min(1, 2\mu([a, \infty)))$$

possible since the left side goes to zero as $b \downarrow 0$. Since the integrand is positive on $[b, a]$, we have that for all $k \in \mathbb{N}$,

$$\int (x^2 - b^2)^{2k+1} d\mu(x) \geq -(b^2)^{2k+1} + 2(a^2 - b^2)^{2k+1}\mu([a, \infty)),$$

$$= 2(a^2 - b^2)^{2k+1}\left[2\mu([a, \infty)) - \left(\frac{b^2}{a^2 - b^2}\right)^{2k+1}\right] \geq 0$$

by the choice of b. Thus $T_- \geq b > 0$. ∎

One consequence of the theorem is

$$T_- \leq \left(\int_{\mathbb{R}} x^2 \, d\mu(x)\right)^{\frac{1}{2}}. \tag{8}$$

It is an interesting question when one has equality. One would like as good a lower bound on T_- as possible which can yield good lower bounds on transition temperatures. Often one has equality in (8) in which case we will say that T_- is *canonical* for μ.

Example 3.3. We consider spins taking three values. For $0 \leq \lambda \leq 1$, consider the probability measure supported by the three points $\{0, \pm 1\}$ given by

$$d\mu_\lambda = \frac{\lambda}{2}(\delta_1 + \delta_{-1}) + (1 - \lambda)\delta_0.$$

For $\lambda = \frac{2}{3}$, which is equal weights, this is called (normalized) spin 1. Then

$$\langle (x^2 - T^2)^{2m+1} \rangle_\lambda = (1 - T^2)^{2m+1}\lambda - (1 - \lambda)T^{2(2m+1)}$$

$$\geq 0 \iff \left[\frac{1 - T^2}{T^2} \right]^{2m+1} \geq \frac{1 - \lambda}{\lambda}$$

$$\iff \frac{1 - T^2}{T^2} \geq \left(\frac{1 - \lambda}{\lambda} \right)^{\frac{1}{2m+1}}.$$

If $\lambda \leq \frac{1}{2}$, then $\frac{1-\lambda}{\lambda} \geq 1$ and the maximum on the right side of the last formula occurs for $m = 0$ while, if $\lambda \geq \frac{1}{2}$, then $\frac{1-\lambda}{\lambda} \leq 1$ and we get the maximum as $m \to \infty$. Thus, we find that

$$T_-(\lambda) = \begin{cases} \sqrt{\lambda} & \text{if } \lambda \leq \frac{1}{2}, \\ \sqrt{\frac{1}{2}} & \text{if } \lambda \geq \frac{1}{2}. \end{cases}$$

So we see there are cases where $T_- = \langle x^2 \rangle^{\frac{1}{2}} = \sqrt{\lambda}$ and other cases where the inequality is strict. Note also that at $\lambda = \frac{1}{2}$, the integral $\langle (x^2 - T_-^2)^{2m+1} \rangle_\lambda$ vanishes for all n, a sign that the distribution of $x^2 - T_-^2$ is symmetric about 0.

In the remainder of this chapter, we will discuss bounds on transition temperatures and then two interesting classes: in Section 5, the distribution of a single component of a D-vector model and in Section 6, the spin S spin.

4 Bounds on transition temperatures

Fix a translation invariant ferromagnetic interaction, $J(i - j) \geq 0$ and an even a priori measure, μ. Let $T_c(\mu)$ be the transition temperature for the model defined as the unique minimal temperature (which may be zero if there is no phase transition!) so that for larger temperatures, the two point infinite volume free boundary condition state $\langle \cdot \rangle$ has

$$\lim_{j \to \infty} \langle \sigma_j \sigma_0 \rangle = 0. \tag{9}$$

We want to see what $\nu^{(s)} \lhd \mu$ implies about the relation of $T_c(\mu)$ and $T_c(\nu)$ (a similar analysis holds with other possible definitions of transition temperature).

The arguments below while stated for Wells order only depend on Ising domination. Making the temperature and measure explicit, with $\langle \cdot \rangle_{T,\mu}$ the infinite volume free BC state, we note that by the definition of $\nu^{(s)}$, we have that (because we are assuming only pair interactions and because temperature appears as $\frac{\sigma_i \sigma_j}{T}$)

$$\langle \sigma^A \rangle_{T,\nu^{(s)}} = s^{|A|} \langle \sigma^A \rangle_{T/s^2, \nu}.$$

Thus, since Wells order implies Ising domination, we see that $\nu^{(s)} \lhd \mu$ implies that

$$\langle \sigma^A \rangle_{T,\mu} \geq s^{|A|} \langle \sigma^A \rangle_{T/s^2, \nu}. \tag{10}$$

Therefore, if $T \leq s^2 T_c(\nu)$, we see that (9) fails for $\langle \sigma_j \sigma_0 \rangle_{T/s^2, \nu}$ and so by (10) for $\langle \sigma_j \sigma_0 \rangle_{T, \mu}$. We have thus proven that:

Proposition 4.1. *Let $\nu^{(s)} \lhd \mu$ for two non-trivial even measures. Then*

$$T_c(\mu) \geq s^2 T_c(\nu).$$

In particular,

$$T_c(\mu) \geq T_-(\mu)^2 T_c(\text{classical Ising}),$$

so if T_- is canonical for μ, then

$$T_c(\mu) \geq \langle x^2 \rangle_\mu T_c(\text{classical Ising}). \tag{11}$$

This last putative inequality is especially interesting because the mean field transition temperature (see, for example, [18, Section II.13] or [21, Section 2.6]) is given by

$$T_{MF}(\mu) = \langle x^2 \rangle_\mu \sum_j J(j) \tag{12}$$

for a pair interacting ferromagnetic model. Thus one has equality in (11) if T_c is replaced by T_{MF}. It is known [5, 21] that mean field theory is exact in the infinite-dimensional limit of nearest neighbor generalized Ising models (in the sense that $T_c/T_{MF} \to 1$ as $d \to \infty$ for the model on $\mathbb{Z}^d$). The work [5] only discusses spin $\frac{1}{2}$ but to get that equality holds in (11), it suffices to get a MF lower bound on $T_c(\mu)$ and [5] get that from Fröhlich, Simon and Spencer [6] whose argument works for any a priori measure (for many models of interest including those of the next two sections, there is also a mean field upper bound on transition temperatures – see [21, Section 2.6] and [15, 23]). In event, we see that *if* (11) *holds for all ferromagnetic pair interactions, then the constant $\langle x^2 \rangle_\mu$ is best possible.* A major theme of the rest of the chapter is proving (11) in two classes of models.

One of these is the spin S measure and we want to end this section by noting what we believe is the best prior lower bound on $T_c(\widetilde{\mu}_S)$. To be explicit, for each value of $S = \frac{1}{2}, 1, \frac{3}{2}, \ldots$, consider the measure $\widetilde{\mu}_S$ which takes $2S + 1$ values equally spaced between -1 and 1, each with weight $\frac{1}{2S+1}$. It is interesting to find the square, $T_-(S)^2$, of the Wells T_- associated to $\widetilde{\mu}_S$ because if $T_c(S)$ is the transition temperature for a model with a priori measure $\widetilde{\mu}_S$ and some fixed two point ferromagnetic interaction, then Proposition 4.1 (and the easy $T_+(S) = 1$) implies that

$$T_-(S)^2 T_c\left(\frac{1}{2}\right) \leq T_c(S) \leq T_c\left(\frac{1}{2}\right) \tag{13}$$

(we will prove in Section 6 that $\widetilde{\mu}_S$ is canonical and compute $\langle x^2 \rangle_{\mu_S}$). So far as we know the best previous result of this genre in the literature is due to Griffiths [8, equation (4.23)] who proved the following.

Theorem 4.2 (Griffiths [8]). *One has that*

$$\frac{1}{4} T_c\left(\frac{1}{2}\right) \le T_c(S) \le T_c\left(\frac{1}{2}\right). \tag{14}$$

Remark 2. (1) This is what Griffiths proves for S an integer. For $2S = 2k + 1$ odd, he proves the slightly stronger result with $\frac{1}{4}$ replaced by $(k + \frac{1}{2k} + 1)^2$. Please note that what we call $T_c(S)$, Griffiths denotes $T_c(2S)$, e.g. T_c(*classical Ising*) which we denote $T_c(\frac{1}{2})$, he denotes as $T_c(1)$.

(2) This paper of Griffiths [8] is best known for proving Lee–Yang and GKS inequalities by realizing spin S (normalized so the maximum value is $2S$, a measure we call μ_S) by $2S$ spins with values ± 1 with finite ferromagnetic couplings (given by Figures 2 and 3 in his paper) but he notes that one can also realize them with S (if S is an integer or $S + \frac{1}{2}$ if that is an integer) frozen together and then GKS implies that $\widetilde{\mu}_S$ Ising dominates $b_{T=\frac{1}{2}}$.

5 Totally anisotropic D-vector model

We turn next to one of the two new results on this subject. It involves the interesting measure

$$d\mu_D(x) = \left[\frac{\Gamma\left(\frac{D}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{D-1}{2}\right)}\right](1 - x^2)^{\frac{1}{2}(D-3)} \chi_{[-1,1]}(x)\, dx.$$

This is the distribution of the component x_1 if one looks at a D-component unit vector, $\mathbf{x} = (x_1, \ldots, x_D)$, distributed with the rotation invariant probability measure on $\mathbb{S}^{D-1}$. Since, with respect to this measure, all x_j have the same distribution and $\sum_{j=1}^{D} x_j^2 = 1$, we clearly have that

$$\langle x^2 \rangle_D = \frac{1}{D}.$$

Theorem 5.1. *The number $T_-(\mu_D)$ is canonical, i.e. $T_-(\mu_D)^2 = \frac{1}{D}$, so, in particular, (11) holds for μ_D.*

The result for $D = 2$ is especially easy because, for all m,

$$\left\langle \left(x^2 - \frac{1}{2}\right)^{2m+1} \right\rangle_{D=2} = 0.$$

To prove this, note that it is equivalent to

$$\langle (2x^2 - 1)^{2m+1} \rangle_{D=2} = \langle (x_1^2 - x_2^2)^{2m+1} \rangle_{\text{rotor}} = 0$$

which follows by $x_1 \leftrightarrow x_2$. I note that this result for $D = 2$ is precisely the result that Aizenman and the second author say is in Wells' mystery preprint.

Lemma 5.2. *Let μ, ν be two measures of compact support on $[0, \infty)$. Then*

$$\int f(x)\, d\mu(x) \leq \int f(x)\, d\nu(x) \tag{15}$$

for all monotone functions with $f(0) = 0$ if and only if

$$\forall_{C>0}\ \mu([C, \infty)) \leq \nu([C, \infty)) \tag{16}$$

Proof. If (15) holds, we take f to be the characteristic function of $[C, \infty)$ to get (16).

Now suppose that we have (16). By a simple approximation argument, it suffices to prove (15) for functions f which are C^1 with $f(0) = 0$. For such functions f, we have that

$$f(x) = \int_0^\infty f'(C)\chi_{[C,\infty)}(x)\, dC$$

by doing the integral. Thus

$$\int f(x)\, d\mu(x) = \int_0^\infty f'(C)\mu([C, \infty))\, dC.$$

Monotonicity of f implies that $f' \geq 0$, so that (16) implies (15). ∎

Lemma 5.3. *Let $a, b > 0$ and suppose that g is positive on $(-b, a)$ with*

$$\int_{-b}^0 g(x)\, dx = \int_0^a g(x)\, dx, \ 0 \leq y < b \implies g(-y) \geq g(y). \tag{17}$$

Let f be an odd, monotone increasing, continuous function on $(-b, a)$. Then

$$\int_{-b}^a f(x)g(x)\, dx \geq 0. \tag{18}$$

Remark. The result is quite intuitive. The condition on g says that the measure $g(x)\, dx$ is concentrated on the right at larger x than on the left, so more concentrated where $|f|$ is larger.

Proof. Define the measures on $[0, \infty)$:

$$d\nu(x) = \chi_{[0,a]}(x)g(x)\, dx, \quad d\mu(x) = \chi_{[0,b]}(x)g(-x)\, dx.$$

We claim that for all $C > 0$, one has that

$$\nu([0, C)) \leq \mu([0, C)) \tag{19}$$

so, by (17), we have that (16) holds. Thus, since f is monotone, by Lemma 5.2, we have that

$$\int_0^b f(x)g(-x)\, dx \leq \int_0^a f(x)g(x)\, dx,$$

which is (18) since f is odd. So we need only prove (19).

If $C \leq b$, then (19) is immediate from the fact that for $x \geq 0$, $g(-x) \geq g(x)$. By (17), if $C > b$, then

$$\mu([0, C)) = \mu([0, \infty)) = \nu([0, \infty)) \geq \nu([0, C)). \qquad \blacksquare$$

Proof of Theorem 5.1. As noted after the theorem, $\frac{1}{D}$ is the second moment. Thus, by (8), $T_-^2 \leq \frac{1}{D}$, so, by (6), we need only prove that for all $m \in \mathbb{N}$, the have that

$$\int_{-1}^{1} \left(x^2 - \frac{1}{D} \right)^{2m+1} (1 - x^2)^{\frac{1}{2}(D-3)} \, dx \geq 0. \qquad (20)$$

Note first that, since the integrand is even in x, we can integrate only from 0 to 1 and then change variables from x to $y = x^2 - \frac{1}{D}$. One sees that (20) is equivalent to

$$\int_{-\frac{1}{D}}^{1-\frac{1}{D}} \frac{y^{2m+1} \left(1 - \frac{1}{D} - y \right)^{\frac{1}{2}(D-3)}}{\left(y + \frac{1}{D} \right)^{\frac{1}{2}}} \, dy \geq 0. \qquad (21)$$

If $D = 2$, the integral goes from $-\frac{1}{2}$ to $\frac{1}{2}$ and the integrand is $y^{2m+1}(y^2 - \frac{1}{4})^{-\frac{1}{2}}$ which is odd so the integral in (21) is 0 for all m, recovering what we saw above. If $D \geq 3$, define, for $m \geq 1$,

$$g(y) = \frac{|y| \left(1 - \frac{1}{D} - y \right)^{\frac{1}{2}(D-3)}}{\left(y + \frac{1}{D} \right)^{\frac{1}{2}}}, \qquad f(y) = \operatorname{sgn}(y) y^{2m}.$$

The fact that $\frac{1}{D}$ is the second moment of $d\mu_D$ implies the integral in (20) vanishes if $m = 0$ which means that g obeys the first equation in (17). Since $\frac{g(y)}{|y|}$ is monotone decreasing in y, the inequality in the second half of (17) holds. The function f is odd and monotone increasing, so (18) implies (21). $\qquad \blacksquare$

6 Spin S

For each value of $S = \frac{1}{2}, 1, \frac{3}{2}, \ldots,$ consider the measure $\tilde{\mu}_S$ which takes $2S + 1$ values equally spaced between -1 and 1, each with weight $\frac{1}{2S+1}$. We will prove here that except for $S = 1$, this measure is canonical which will lead to improvements in the first inequality in (14) by a factor of at least $\frac{4}{3}$ and which by discussion after (12) yields optimal constants. We begin by computing $\langle x^2 \rangle_{\tilde{\mu}_S}$.

Theorem 6.1. *We have*

$$a_S \equiv \int x^2 \, d\tilde{\mu}_S(x) = \frac{1}{3} \frac{S + 1}{S}. \qquad (22)$$

Remark 3. (1) Using $\sum_{j=1}^{S} j^2 = \frac{S(S+1)(2S+1)}{6}$, one easily gets this result for S integral. One can also use this formula to get the result for half odd integral S by

using the fact that this case can be rewritten as a sum over odd integers between 1 and $2S$ which one can realize as a sum over all integers minus the sum over even integers. That calculation is awkward and it is surprising it yields the same algebraic formula which is why we give the slick uniform proof below.

(2) We will often scale the measure by a factor of S so the $2S + 1$ values are equally spaced from $-S$ to S yielding a measure μ_S. If we interpret $\sum_{j=-S}^{S}$ as the sum over the $-S, -S + 1, \ldots, S - 1, S$ (the usual meaning if S is an integer but unusual if S is half and odd integer), then (22) is equivalent to

$$A_S \equiv \int x^2 \, d\mu_S(x) = \frac{1}{2S + 1} \sum_{-S}^{S} j^2 = \frac{S(S + 1)}{3}. \tag{23}$$

Proof. Use $\underline{\mathrm{Tr}}(\cdot)$ for the normalized trace on finite-dimensional spaces, i.e. the average of diagonal elements or $\mathrm{Tr}(\cdot)$ divided by the dimension. Consider what a physicist would call a quantum spin, σ, of spin S and a mathematician the generators of the irreducible representation of SU(2) of dimension $2S + 1$. The operator σ_z has eigenvalues $-S, -S + 1, \ldots, S - 1, S$ and the Casimir operator has the form

$$\sigma_x^2 + \sigma_y^2 + \sigma_z^2 = S(S + 1)\mathbf{1}.$$

Thus

$$A_S = \underline{\mathrm{Tr}}(\sigma_z^2) = \frac{1}{3}\underline{\mathrm{Tr}}((\sigma_x^2 + \sigma_y^2 + \sigma_z^2)) = \frac{1}{3}\underline{\mathrm{Tr}}(S(S + 1)\mathbf{1}) = \frac{S(S + 1)}{3}. \tag{24}$$

by the symmetry of the three directions. This proves (23).   ∎

We saw in Example 3.3 that for $S = 1$ ($\lambda = \frac{1}{3}$), one has that

$$T_- = \sqrt{\frac{1}{2}} < \sqrt{\frac{2}{3}} = \sqrt{a_{S=1}},$$

so $\tilde{\mu}_S$ is *not* canonical for that value of S. The main technical result of this section, which we will prove in Appendix A, is:

Theorem 6.2. *Let $S \neq 1$ be half a positive integer. Then*

$$\sum_{j=-S}^{S} (3j^2 - S(S + 1))^{2m+1} \geq 0. \tag{25}$$

We will also see that the opposite inequality is true when $S = 1$. By Theorem 3.1, this implies:

Corollary 6.3. *For any $S \neq 1$, we have that T_- is canonical for $\tilde{\mu}_S$.*

This, in turn, by (13) implies the following result.

Corollary 6.4. *For $S \neq 1$, one has that*

$$T_c(S) \geq \left(\frac{1}{3} + \frac{1}{3S}\right) T_c\left(\frac{1}{2}\right)$$

while for $S = 1$ one has that

$$T_c(1) \geq \frac{1}{2} T_c\left(\frac{1}{2}\right).$$

This improves Griffiths result, (14) by a factor of at least $\frac{4}{3}$. To be totally accurate, as we noted Griffith's lower bound (as proven, not as stated!) is slightly stronger than (14) when S is half an odd integer. In particular, for $S = \frac{3}{2}$ what (14) has as $\frac{1}{4}$, he actually proves is $\frac{4}{9}$ so our results are improvement but for this value only by a factor of $\frac{5}{4}$ rather than $\frac{4}{3}$.

One final remark. In the notation used in (24), (25) is equivalent to

$$\mathrm{Tr}\left((2\sigma_z^2 - \sigma_x^2 - \sigma_y^2)^{2m+1}\right) \geq 0.$$

We wonder if that could not be used for a matrix theoretic version of what we prove in the Appendix.

A A majorization bound

In this appendix, we will prove a general set of inequalities that includes (25). We will prove two theorems:

Theorem A.1. *Fix an integer $N \geq 1$, a function, ψ on $[0, 1]$, which is non-negative, continuous, strictly monotone increasing and convex and a function, Φ, on the interval $[-\|\psi\|_\infty, \|\psi\|_\infty]$ which is continuous, odd and whose restriction to $[0, \|\psi\|]$ is convex. Let*

$$\overline{\psi} = (N + 1)^{-1} \sum_{j=0}^{N} \psi\left(\frac{j}{N}\right).$$

Then

$$\sum_{j=1}^{N} \Phi\left(\psi\left(\frac{j}{N}\right) - \overline{\psi}\right) \geq 0.$$

Remark. By translation and scaling, this result can easily be generalized. For example, while stated for $N + 1$ equally spaced points between 0 and 1, we will apply it to $N + 1$ odd integers stating at 1. The map $k \mapsto \frac{k-1}{2N}$ maps those $N + 1$ integers into the points of the theorem.

Theorem A.2. *For an integer $N \geq 2$ and an even, continuous, convex function, ψ on $[-1, 1]$ and a function, Φ, on $[-\|\psi\|_\infty, \|\psi\|_\infty]$ which is continuous, odd and whose restriction to $[0, \|\psi\|]$ is convex. Let $\overline{\psi}$ be given by*

$$\overline{\psi} = (2N + 1)^{-1} \sum_{j=-N}^{N} \psi\left(\frac{j}{N}\right).$$

Suppose that

$$2\psi(1) + \psi(0) + 2\psi\left(\frac{1}{N}\right) \geq 5\overline{\psi}. \tag{26}$$

If N is odd, suppose that

$$\psi\left(\frac{1}{2} + \frac{1}{2N}\right) \leq \overline{\psi}. \tag{27}$$

Then

$$\sum_{j=-N}^{N} \Phi\left(\psi\left(\frac{j}{N}\right) - \overline{\psi}\right) \geq 0. \tag{28}$$

Remark 4. (1) We will see later (see (33)) that $2\psi(1) + 2\psi(0) \geq 4\overline{\psi}$ which puts (26) in perspective.

(2) It might be true that this theorem holds without the need for the condition (26) but it holds in the case we need so we did not try hard to eliminate it. We will explain later why a naive extension of the proof of Theorem A.1 does not work in a simple example and led to the extra condition. We do note that (26) is a restriction. If we normalize ψ by $\psi(0) = 0$, $\psi(1) = 1$, then in the limit as $N \to \infty$, (26) becomes

$$\int_0^1 \psi(x)\,dx \leq \frac{2}{5},$$

which for $\psi(x) = |x|^p$ requires $p \geq \frac{3}{2}$ while convexity only requires $p \geq 1$.

(3) On the other hand, (27), as we will see, is quite natural independent of our method of proof.

One key idea behind the proofs is the theory of majorization. We let $\mathbb{R}^n_{+,\geq}$ denote the set of n-tuples of reals, $\mathbf{x} = (x_1, \ldots, x_n)$ with

$$x_1 \geq x_2 \geq \cdots \geq x_n \geq 0. \tag{29}$$

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_{+,\geq}$, we say that $\mathbf{x}$ *majorizes* $\mathbf{y}$, written $\mathbf{x} \succ \mathbf{y}$ or $\mathbf{y} \prec \mathbf{x}$ if and only if

$$\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j, \qquad S_k(\mathbf{x}) \equiv \sum_{j=1}^{k} x_j \geq \sum_{j=1}^{k} y_j, \quad k = 1, \ldots, n-1,$$

which defines $S_k(\mathbf{x})$. A standard reference is Marshall and Olkin [14] which has been called a love poem to majorization; other references are Hardy, Littlewood and

Pólya [11] and Simon [20, Chapters 14–15]. We will rely on the following aspect of the theory:

Theorem A.3 (Karamata's inequality). *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_{+,\geq}$ *with* $\mathbf{x} \succ \mathbf{y}$ *and let* φ *be an arbitrary continuous convex function on* $[0, x_1]$. *Then*

$$\sum_{j=1}^n \varphi(x_j) \geq \sum_{j=1}^n \varphi(y_j). \tag{30}$$

Remark 5. (1) Even though this is widely referred to as Karamata's inequality (e.g., Cvetkovski[4] or Wikipedia [26]) after Karamata's 1932 paper [12], it or theorems that imply it appear in a 1923 paper of Schur [16] and a 1929 paper of Hardy, Littlewood and Pólya [10]. That said, we note that [10] does not have a proof which may not have appeared until [11] in 1934 and that Karamata proved a converse, namely, if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_{+,\geq}$ and (30) holds for all convex φ, then $\mathbf{x} \succ \mathbf{y}$.

(2) The idea of one proof is quite simple (for details, see, e.g., Simon [19, Theorem 1.9] or Simon [20, Theorem 15.5]): by slicing with specific hyperplanes, one proves that $\mathbf{y}$ is in the convex hull in $\mathbb{R}^n$ of the (at most) $n!$ points obtained from $\mathbf{x}$ by permuting the coordinates and then one notes the function $\mathbf{w} \mapsto \sum_{j=1}^n \varphi(w_j)$ is convex and permutation symmetric.

In the case of Theorem A.1, we will prove that $\mathbf{x} \succ \mathbf{y}$ using a simple observation:

Proposition A.4. *Suppose that* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_{+,\geq}$ *with* $\sum_{j=1}^n x_j = \sum_{j=1}^n y_j$ *and that for some* $\ell \in 2, \ldots, n-1$, *one has that*

$$j < \ell \implies x_j > y_j, \quad j \geq \ell \implies x_j \leq y_j.$$

Then $\mathbf{x} \succ \mathbf{y}$.

Proof. If $k < \ell$, it is immediate that

$$\sum_{j=1}^k x_j \geq \sum_{j=1}^k y_j$$

and similarly, it is immediate that if $k \geq \ell$, then

$$\sum_{j=k}^n x_j \leq \sum_{j=k}^n y_j.$$

Subtracting this from $\sum_{j=1}^n x_j = \sum_{j=1}^n y_j$, we see that also for $k \geq \ell$, one has that $\sum_{j=1}^k x_j \geq \sum_{j=1}^k y_j$. ∎

We need two preliminaries for the proof of Theorem A.1.

Lemma A.5. *Let ψ be a convex function on $[0,1]$ and suppose that*

$$0 \le \widetilde{b} \equiv 2c - b < \widetilde{a} \equiv 2c - a \le c \le a < b \le 1.$$

Then

$$\frac{1}{2}(\psi(b) + \psi(\widetilde{b})) \ge \frac{1}{2}(\psi(a) + \psi(\widetilde{a})) \ge \psi(c). \tag{31}$$

Moreover, the first inequality is strict unless $\psi'(s)$ is constant on $(\widetilde{a}, a)$.

Proof. If one takes $a = c$ and then replaces b by a, the first inequality becomes the second so it suffices to prove the first one. Without loss (by translation and scaling) we can take $c = \frac{1}{2}, b = 1$ so that $\widetilde{b} = 0$ and $\widetilde{a} = 1 - a$. By the fundamental theorem of calculus (a general convex function is not C^1 but it is differentiable with the possible exception of a countable set and the fundamental theorem of calculus holds; see Simon [20, Theorem 1.28])

$$\frac{1}{2}(\psi(1) + \psi(0)) - \frac{1}{2}(\psi(a) + \psi(\widetilde{a})) = \frac{1}{2}\int_a^1 [\psi'(s) - \psi'(1 - s)]\, ds.$$

By convexity, the integrand is non-negative so we have proven (31). Moreover, if $\psi'(s)$ is not constant on $(1 - a, a)$, then the integral is strictly positive. ∎

Proposition A.6. *Let ψ, $\overline{\psi}$ and N be as in Theorem A.1. Then*

$$n \equiv \#\left\{ j \;\middle|\; \psi\left(\frac{j}{N}\right) \le \overline{\psi} \right\} \ge \frac{N+1}{2} \tag{32}$$

and

$$\psi\left(\frac{1}{2}\right) \le \overline{\psi} \le \frac{1}{2}(\psi(0) + \psi(1)). \tag{33}$$

Moreover, the inequalities in (33) are strict if $N \ge 2$ and ψ is not an affine function on $[0,1]$ (i.e. ψ' is not constant).

Proof. For any $j = 0, 1, \ldots, N$, (31) implies that

$$\psi\left(\frac{1}{2}\right) \le \frac{1}{2}\left(\psi\left(\frac{j}{N}\right) + \psi\left(1 - \frac{j}{N}\right) \right) \le \frac{1}{2}(\psi(0) + \psi(1)).$$

Averaging over j yields (33). If ψ is not affine on the interval $[0,1]$, then the second inequality is strict for $1 \le j \le N - 1$ so the second inequality in (33) is strict. Since $\psi(\frac{1}{2}) < \frac{1}{2}(\psi(0) + \psi(1))$ if ψ is not affine, we see that in the case the first inequality is always strict.

As ψ is strictly monotone, the first inequality in (33) implies the unique $x \in [0,1]$ with $\psi(x) = \overline{\psi}$ has $x \ge \frac{1}{2}$. This implies that

$$n = \#\left\{ j \;\middle|\; \frac{j}{N} \le x \right\} \ge \#\left\{ j \;\middle|\; \frac{j}{N} \le \frac{1}{2} \right\} \ge \frac{N+1}{2}. \qquad ∎$$

Proof of Theorem A.1. Let $q = N + 1 - n \le n$ by (32). Define

$$y_j = \overline{\psi} - \psi\left(\frac{j-1}{N}\right), \quad j = 1, \ldots, n,$$

and

$$x_j = \begin{cases} \psi\left(\frac{N+1-j}{N}\right) - \overline{\psi} & \text{if } j = 1, \ldots, q, \\ 0 & \text{if } j \ge q. \end{cases}$$

Since ψ is monotone and n is defined by (32), we have that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_{+,\ge}$. By the definition of $\overline{\psi}$, we have that

$$\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j. \tag{34}$$

If $N = 1$ or ψ is affine on $[0, 1]$, it is easy to see that $x_j = y_j$ for all j, so, since Φ is odd, we have that (29) holds. Thus henceforth we will suppose that $N \ge 2$ and ψ is not an affine function on $[0, 1]$, so, in particular, the inequalities in (33) are strict.

Note next that because ψ is assumed convex, we have that

$$m < p \implies \psi\left(\frac{m+1}{N}\right) - \psi\left(\frac{m}{N}\right) \le \psi\left(\frac{p+1}{N}\right) - \psi\left(\frac{p}{N}\right). \tag{35}$$

By the strict form of (33), $x_1 > y_1$. Because of (34), there must be a first ℓ so that $x_\ell \le y_\ell$. We claim that if $\ell < n$, then $x_{\ell+1} \le y_{\ell+1}$. If $\ell + 1 > q$, then $x_{\ell+1} = 0$ and the required inequality is immediate. If $\ell + 1 \le q$, then (35) implies that $x_\ell - x_{\ell+1} \ge y_\ell - y_{\ell+1}$. Subtracting this from $x_\ell \le y_\ell$ proves that $x_{\ell+1} \le y_{\ell+1}$. Repeating this argument proves that for all $j \ge \ell$ we have $x_j \le y_j$. Thus by Proposition A.4, $\mathbf{x} \succ \mathbf{y}$.

By Karamata's inequality, (30), we conclude that $\sum_{j=1}^{n} \Phi(x_j) - \Phi(y_j) \ge 0$. As Φ is odd and $\Phi(0) = 0$, this is equivalent to (29). ∎

Example A.7. To understand why we need the extra condition (26) in Theorem A.2, we consider μ_S for $S = 6$, i.e. thirteen pure points with weight $\frac{1}{13}$ at $0, \pm 1, \pm 2, \pm 3, \pm 4, \pm 5, \pm 6$. By (23), the average of the square is $A_6 = 14$. The values of j^2 are $j^2 = 0, 1, 1, 4, 4, 9, 9, 16, 16, 25, 25, 36, 36$ so $n = 7$ values are less than A_6 and one sees that (ignore $\mathbf{w}$ for now)

$$\mathbf{x} = 22, 22, 11, 11, \ 2, 2, 0,$$
$$\mathbf{y} = 14, 13, 13, 10, 10, 5, 5,$$
$$\mathbf{w} = 22, 22, \ 0, 11, 11, 2, 2.$$

One can verify that $\mathbf{x} \succ \mathbf{y}$ by hand (and, below, we will prove the result for all $S \ge 2$) but one cannot use Proposition A.4 as we did in our proof of Theorem A.1 for $x_j - y_j$ shifts signs three times instead of one time. The problem is that the components of $\mathbf{x}$ and $\mathbf{y}$ are paired but shifted.

Look at **w** which we get by moving the 0 from position 7 to position 3. One can handle the first three partial sums by noting that $22 + 22 \geq 14 + 13 + 13$ and the remaining partial sums by noting that there is only one sign change after the third place and use Proposition A.4 to prove the partial sums of **w** dominate those of **y** and note it is trivial that partial sums of **x** dominate those of **w**. The key is that by moving the 0, the pairs are no longer shifted.

Proof of Theorem A.2. Notice that $2N + 1$ is odd, so either **x** or **y** needs to have a zero added. To be sure that it is added to the **x**, we need the analog of (32), namely that

$$n \equiv \#\left\{ j \ \middle| \ \psi\left(\frac{j}{N}\right) \leq \overline{\psi} \right\} \geq N + 1. \tag{36}$$

If $N = 2k$ is even, the $N + 1$ values of j among the $2N + 1$ overall values with the smallest values of $\psi(\frac{j}{N})$ are $0, \pm 1, \pm k$, so (36) follows from the first inequality in (33). If $N = 2k + 1$ is odd, we cannot take the max j to be k since that only yields $2k + 1 < N + 1$ values and therefore we need to go up to $j = k + 1$ yielding $N + 2$ values (leaving over $N - 1$), so we need (27) to hold. In that case, there are at least three more values $\leq \overline{\psi}$ than $> \overline{\psi}$! (This has to be because n is odd and $N - n$ even and we require that $n > N - n$.)

Define **x** to be the $2N + 1 - n$ values of $\psi - \overline{\psi}$ larger than 0 written in decreasing order plus $2n - 2N - 1$ zero values at the end and **y** the n non-negative values of $\overline{\psi} - \psi$ so $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. As in the proof of Theorem A.1, if we prove that $\mathbf{x} \succ \mathbf{y}$, then (28) follows from Karamata's inequality.

For $N \geq 2$ (so $n \geq 3$), define $\mathbf{w} \in \mathbb{R}^n_{+,\geq}$ by

$$w_j = \begin{cases} x_j & \text{if } j = 1, 2, \\ 0 & \text{if } j = 3, \\ x_{j-1} & \text{if } j \geq 3. \end{cases}$$

We claim first that

$$S_n(\mathbf{x}) = S_n(\mathbf{w}), \qquad S_j(\mathbf{x}) \geq S_j(\mathbf{w}), \quad j = 1, \ldots, n - 1 \tag{37}$$

(we do not write this as $\mathbf{x} \succ \mathbf{w}$, first because $\mathbf{w} \notin \mathbb{R}^n_{+,\geq}$ and we've only defined majorization for such sequences and also because it is usual to extend $\succ$ to $\mathbb{R}^n_+$ by demanding the relations for the decreasing rearrangements rather than by (37)). This follows because **w** is a rearrangement of **x** and the decreasing sequence, $S_j(\mathbf{x})$, is maximal among those sums for any arrangement or also, more directly, by noting that $S_j(\mathbf{w}) = S_j(\mathbf{x})$ for $j = 1, 2, n$ and $S_j(\mathbf{w}) = S_{j-1}(\mathbf{x})$ for $j \geq 3$.

Thus, if we prove that

$$S_n(\mathbf{w}) = S_n(\mathbf{y}), \qquad S_j(\mathbf{w}) \geq S_j(\mathbf{y}), \quad j = 1, \ldots, n - 1, \tag{38}$$

then $\mathbf{x} \succ \mathbf{y}$ and the proof is done.

Note that (38) for $j = n$ follows from the definition of $\overline{\psi}$. For $j = 1, 2$, we note that by the second inequality of (33), we have that $w_1 = x_1 \geq y_1$ and then that $S_2(\mathbf{w}) = S_2(\mathbf{x}) = 2x_1 \geq 2y_1 \geq y_1 + y_2$. For $j = 3$, the required inequality is the hypothesis (26).

Because of the equality for $j = n$ and inequality for $j = 3$, there must be a first $\ell \geq 4$ with $x_\ell \leq y_\ell$. We claim either $x_{\ell+1} = 0$ or $x_\ell - x_{\ell+1} \geq y_\ell - y_{\ell+1}$ because if $x_{\ell+1} > 0$ and ℓ is odd, we have that $x_\ell = x_{\ell+1}$ and $y_\ell = y_{\ell+1}$ or ℓ is even and we can use (35) to prove the inequality as in the proof of Theorem A.1. By induction we see that for $j \geq \ell$, we have that $x_j \leq y_j$. By following the argument used to prove Proposition A.4 we conclude that for all $j \geq 4$, we have (38). ∎

Proof of Theorem 6.2. If S is half an odd integer, we can use Theorem A.1 with $N = S - \frac{1}{2}$ and $\psi(x) = (\frac{1}{2} + Nx)^2$ which is non-negative, strictly monotone and convex and $\Phi(y) = (3y)^{2m+1}$.

If S is an integer, we will use Theorem A.2 with $N = S$, $\psi(x) = (Sx)^2$ and $\Phi(y) = (3y)^{2m+1}$. We need to check (26), which says that

$$2S^2 + 2 \geq 5A_S = \frac{5}{3}S(S+1) \tag{39}$$

and (27) which says that for $S \geq 3$ odd, one has that

$$S^2\left(\frac{1}{2} + \frac{1}{2S}\right)^2 \leq A_S = \frac{S(S+1)}{3}. \tag{40}$$

Estimate (39) is equivalent to

$$0 \leq S^2 - 5S + 6 = (S-2)(S-3)$$

which holds for all integral S (since the polynomial $x^2 - 5x + 6$ is negative precisely for $x \in (2, 3)$). (40) is equivalent to $3S^2 + 6S + 3 \leq 4S^2 + 4S$ or

$$0 \leq S^2 - 2S - 3 = (S-3)(S+1)$$

which holds for all $S \geq 3$ as required. ∎

We note that when $S = 1$ with the normalization used in this proof, one has that the values of j^2 are $0, 1, 1$ so $A_1 = \frac{2}{3}$ and $\mathbf{x} = (\frac{1}{3}, \frac{1}{3})$ while $y = (\frac{2}{3}, 0)$ so $\mathbf{x} \prec \mathbf{y}$ and for all m, (25) holds with the opposite sign!

As we mentioned, we would guess that results like Theorem A.2 hold without (26). To be explicit, consider the an analog of (25) with $3j^2$ replace by $|j|^p$ and $3A_S = S(S+1)$ replaced by the suitable average. We know this analog is valid for any $p > 1$ if S is half an odd integer but because (26) fails when $p < \frac{3}{2}$ and S is large we do not know the answer to the following question.

Question 2. Does the analog of (25) hold for $S \neq 1$ integral when $3j^2$ is replaced by $|j|^p$ for any $p > 1$ and with $3A_S = S(S+1)$ replaced by the suitable average?

This is an explicit example for what we hope might be a general result.

References

[1] M. Aizenman and B. Simon, A comparison of plane rotor and Ising models. *Phys. Lett. A* **76** (1980), no. 3–4, 281–282

[2] J. Bricmont, J. L. Lebowitz, and C. E. Pfister, Periodic Gibbs states of ferromagnetic spin systems. *J. Stat. Phys.* **24** (1981), no. 1, 269–277

[3] P. L. Chebyshev, On approximate expressions of some integrals in terms of others, taken within the same limits. *Proc. Math. Soc. Kharkov* **2** (1882), 93–98

[4] Z. Cvetkovski, *Inequalities*. Springer, Heidelberg, 2012

[5] F. J. Dyson, E. H. Lieb, and B. Simon, Phase transitions in quantum spin systems with isotropic and nonisotropic interactions. *J. Stat. Phys.* **18** (1978), no. 4, 335–383

[6] J. Fröhlich, B. Simon, and T. Spencer, Infrared bounds, phase transitions and continuous symmetry breaking. *Comm. Math. Phys.* **50** (1976), no. 1, 79–95

[7] J. Ginibre, General formulation of Griffiths' inequalities. *Comm. Math. Phys.* **16** (1970), 310–328

[8] R. B. Griffiths, Rigorous results for Ising ferromagnets of arbitrary spin. *J. Math. Phys.* **10** (1969), 1559–1565

[9] F. Guerra, L. Rosen, and B. Simon, The $\mathbf{P}(\phi)_2$ Euclidean quantum field theory as classical statistical mechanics. I, II. *Ann. of Math. (2)* **101** (1975), 111–189; ibid. (2) 101 (1975), 191–259

[10] G. H. Hardy, J. E. Littlewood, and G. Pólya, Some simple inequalities satisfied by convex functions. *Messenger of Math.* **58** (1929), 145–152

[11] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge Math. Libr., Cambridge University Press, Cambridge, 1934

[12] J. Karamata, Sur une inégalité rélative aux fonctions convexes. *Publ. Math. Univ. Belgrade* **1** (1932), 145–148

[13] E. H. Lieb, The classical limit of quantum spin systems. *Comm. Math. Phys.* **31** (1973), 327–340

[14] A. W. Marshall and I. Olkin, *Inequalities: Theory of majorization and its applications*. Math. Sci. Eng. 143, Academic Press, New York, 1979

[15] P. A. Pearce, Mean-field bounds on the magnetization for ferromagnetic spin models. *J. Stat. Phys.* **25** (1981), no. 2, 309–320

[16] I. Schur, Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie. *Sitzungsber. Berlin Math. Ges.* **22** (1923), 9–20

[17] B. Simon, The classical limit of quantum partition functions. *Comm. Math. Phys.* **71** (1980), no. 3, 247–276

[18] B. Simon, *The statistical mechanics of lattice gases. Vol. I*. Princeton Ser. Phys., Princeton University, Princeton, 1993

[19] B. Simon, *Trace ideals and their applications*. 2nd edn., Math. Surveys Monogr. 120, American Mathematical Society, Providence, 2005

[20] B. Simon, *Convexity*. Cambridge Tracts in Math. 187, Cambridge University Press, Cambridge, 2011

[21] B. Simon, *Phase transitions in the theory of lattice gases*. Cambridge University Press, Cmbridge, in preparation.

[22] B. Simon and R. B. Griffiths, The $(\phi^4)_2$ field theory as a classical Ising model. *Comm. Math. Phys.* **33** (1973), 145–164

[23] H. Tasaki and T. Hara, Mean field bound and GHS inequality. *J. Stat. Phys.* **35** (1984), no. 1–2, 99–107

[24] H. van Beijeren and G. S. Sylvester, Phase transitions for continuous-spin Ising ferromagnets. *J. Funct. Anal.* **28** (1978), no. 2, 145–167

[25] D. R. Wells, *Some moment inequalities and a result on multivariate unimodality*. Ph.D. Thesis, Indiana University, 1977

[26] Wikipedia article, Karamata's Inequality. https://en.wikipedia.org/wiki/Karamata%27s_inequality