

A TALE OF THREE COAUTHORS

BARRY SIMON^{1,2}

ABSTRACT. We tell the story of the discovery of an interesting bound on finite sums and its application to comparison of Ising models.

1. INTRODUCTION

In 2022, we celebrated Elliott Lieb's 90th birthday. On Friday, Jan 14, 2022, I had a draft of a single authored paper intended for a Festschrift to be published for Lieb. Six days later, that paper had three coauthors who I hadn't met before Jan 14, 2022 (indeed, even now, I've only met them on Zoom). This paper will explain the interesting story, expose some underlying machinery and sketch the proof of a lovely inequality on certain finite sums. It will include an improvement of 50 year old bounds of Griffiths [7] comparing transition temperatures on generalized Ising models for different spins. Because its a fun story and involves a charming inequality, I've given several talks on the material including at the Seminar on Analysis, Differential Equations and Mathematical Physics [20], an online seminar sponsored by the Institute of Mathematics, Mechanics and Computer Sciences of Southern Federal University, Rostov-on-Don, Russia. After the talk, I was invited to by one of the organizers, Alexey Karapetyants, to contribute an article telling this story for the special 50th anniversary issue of the Journal of Mathematical Sciences. I was originally reluctant but persuaded by the fact that I might convince young workers to react properly to rejections since a paper rejection is part of the story.

I am writing a book for Cambridge Press entitled *Phase Transitions in the Theory of Lattice Gases* [21]. It is, in many ways, the successor to my 1993 book *The Statistical Mechanics of Lattice Gases*, Vol. I [17], from Princeton University Press. That earlier book was mainly framework and largely left out all the most fun and beautiful elements

Date: December 25, 2023.

2020 Mathematics Subject Classification. 82B20, 62H99.

Key words and phrases. Ising model, Majorization.

¹ Departments of Mathematics and Physics, Mathematics 253-37, California Institute of Technology, Pasadena, CA 91125. E-mail: bsimon@caltech.edu.

² Research supported in part by Israeli BSF Grant No. 2020027.

of the theory: Correlation Inequalities, Lee-Yang, Peierls' Argument, BKT transitions, Infrared Bounds and Random Clusters and Currents which are the subjects of the new book. But since I decided to use a different publisher, this is certainly NOT volume 2 of the earlier work. The main mathematical focus of the story is an inequality which one can state and admire without knowing anything about Ising models so I will not bother to define what exactly what they are instead referring the reader in general to the two books or to the paper with the three coauthors [13], giving more precise references at appropriate points.

For our discussion we will only need to consider generalized Ising models in finite volume, a subset, Λ , of the lattice \mathbb{Z}^ν . For each site $j \in \Lambda$, one has a real valued "spin", σ_j . For $A \subset \Lambda$, one defines

$$\sigma^A = \prod_{j \in A} \sigma_j \quad (1.1)$$

One object needed to define the Gibbs measure is a function, $J : 2^\Lambda \rightarrow \mathbb{R}$ called a coupling which is called *ferromagnetic* if $J(A) \geq 0$ for all A . One then forms a Hamiltonian,

$$h = - \sum_{A \subset \Lambda} J(A) \sigma^A \quad (1.2)$$

The name ferromagnetic comes from the fact that when $J(A) \geq 0$, states with more spins parallel have lower energies and so according to Gibbs rules higher weights.

The other object one needs describes the uncoupled state which gives the $\{\sigma_j\}_{j \in \Lambda}$ the distribution of independent, identically distributed random variables with some common distribution, $d\mu$, a probability measure on \mathbb{R} called the *a priori measure*, which we will always suppose to be even. One then fixes $J(\cdot)$ but varies μ and forms Gibbs states, $\langle \cdot \rangle_{\mu, \Lambda}$, according to the standard prescription [17, Chap. III].

There are special choices of a priori measure that particularly concern us beginning with the spin- $\frac{1}{2}$ measure:

$$d\tilde{\mu}_{S=\frac{1}{2}} = \frac{1}{2}(\delta_1 + \delta_{-1}) \quad (1.3)$$

More generally, for any $T > 0$, we consider

$$db_T = \frac{1}{2}(\delta_T + \delta_{-T}) \quad (1.4)$$

b is for Bernoulli. Finally, for $S = \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$, we consider the measure $d\tilde{\mu}_S$ which takes $2S + 1$ values equally spaced between -1 and 1 , each with weight $1/(2S + 1)$. This is a scaled version of what is called spin S Ising (which has maximum spin value $2S$).

Conflict of Interest Statement: This is to certify that the author has no conflicts of interest.

Data Availability Statement: There is no data associated with this paper.

2. WELLS ORDERING

As I began to write about correlation inequalities, I wondered about a natural question. 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

$$\langle \sigma^B \rangle_{\mu, \Lambda} \leq \langle \sigma^B \rangle_{\nu, \Lambda} \quad (2.1)$$

In particular, for general μ of 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. To be explicit, if μ Ising dominates $b_{T_-(\mu)}$, and if $T_c(\mu)$ is a transition temperature for some fixed ferromagnetic pair interaction, one easily sees that

$$T_c(\mu) \geq T_-(\mu)^2 T_c(\text{classical Ising}) \quad (2.2)$$

For most even minor aspects of the subject of correlation inequalities, there are several papers, sometimes as many as a dozen. So I was surprised that I was unable to find a single published paper on the subject of what I just called Ising domination! Of course, it was unclear how to search for the subject in Google. Eventually, I did find one paper of van Beijeren and Sylvester [22] that is unsatisfactory in that, in their theory, the analog of what I call T_- is 0 if $0 \in \text{supp}(\mu)$. And I did find an appendix of a paper on another subject but that gets ahead of my story.

One of the pleasant things about writing a book on a subject that I once knew more about is that I get to rediscover things I've forgotten. With the question of Ising domination in the back of my mind, I found an interesting footnote in a 1980 paper of Aizenman and er, B. Simon [1] entitled *A comparison of plane rotor and Ising models*. The footnote said

then by results of Wells (D. Wells, Some moment inequalities for general spin Ising ferromagnets, Indiana Univ. preprint)

$$\langle s_j s_k \rangle_{\beta, 1} \leq 2 \langle \sigma_j^{(1)} \sigma_k^{(1)} \rangle_{\beta, 2} \quad (2.3)$$

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

So I set about finding this preprint. Google didn't help directly but did point me to a 1984 paper of Chuck Newman [15] that mentioned Wells' Indiana University PhD. thesis. So I wrote to Michael asking if he knew anything about our footnote and cced Chuck (who had been a grad student with me at Princeton) because I 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 delightful character, long dead. I then wrote to Kevin Pilgrim, the chair at Indiana, who located a copy of Wells thesis [23] for me on Proquest. But he had no luck on the preprint nor on locating Wells through Indiana University alumni records (to get way ahead of the story, when I eventually reached Wells, he was surprised by this remarking "gee, they don't have trouble finding me to send fund raising letters")! While the thesis did not have anything directly about the above inequality, it did have a general framework on what I called the Ising domination problem, lovely material that should have been published.

Given the beauty and relevance of the work in Wells thesis, I decided to include it in my book and to further develop it. Wells exploited tools that Ginibre had introduced [6] to prove GKS correlation inequalities which he applied instead to what I've called the Ising domination problem. I won't describe his work in detail here (but see [13, 21]) but will limit things to quoting the part of his major theorem relevant to our focus in this note.

Theorem 2.1 (Wells[23, 13]). *Let $d\mu$ be an even probability measure on \mathbb{R} with compact support that is not a point mass at 0. Then there is a strictly positive number, $T_-(\mu)$, so that μ Ising dominates b_S if and only if $S \leq T_-$. Moreover*

$$S \leq T_- \iff \forall_{n \in \mathbb{N}} \int_{\mathbb{R}} (x^2 - S^2)^n d\mu(x) \geq 0 \quad (2.4)$$

I should mention I happened to look at a 1981 paper of Bricmont, Lebowitz and Pfister [3] that includes in an appendix a proof (with attribution to Wells) of Wells result about the existence of $T_- > 0$.

One consequence of the theorem is

$$T_- \leq \left(\int_{\mathbb{R}} x^2 d\mu(x) \right)^{1/2} \quad (2.5)$$

It is an interesting question when one has equality in this inequality. I call the measure *canonical* if one does have it. One of my few new results was to show that if μ_D is the probability of distribution of a the first component of a D -component vector uniformly distributed on the unit sphere S^{D-1} in \mathbb{R}^D , then μ_D is canonical (it was clear from the quote from my paper with Aizenman that Wells had proven this for $D = 2$ in his preprint).

I also computed for $0 \leq \lambda \leq 1$, T_- for the probability measure supported by the three points $\{0, \pm 1\}$ given by

$$d\mu_\lambda = \frac{\lambda}{2}(\delta_1 + \delta_{-1}) + (1 - \lambda)\delta_0 \quad (2.6)$$

and found that

$$T_-(\lambda) = \begin{cases} \sqrt{\lambda}, & \text{if } \lambda \leq \frac{1}{2} \\ \sqrt{\frac{1}{2}}, & \text{if } \lambda \geq \frac{1}{2} \end{cases} \quad (2.7)$$

In particular, this measure is canonical if and only if $\lambda \leq \frac{1}{2}$. This shows some measures are canonical and others are not.

3. THE CONJECTURE

Consider the measure $d\tilde{\mu}_S$ discussed earlier - the scaled spin S Ising model of $2S + 1$ values equally spaced between -1 and 1 . We have just seen that for $S = 1$ ($\lambda = \frac{2}{3}$ in the above example), one has that $T_- = \sqrt{\frac{1}{2}} < \sqrt{\frac{2}{3}} = (\int_{\mathbb{R}} x^2 d\tilde{\mu}_{S=1}(x))^{1/2}$. So $T_- \neq (\langle x^2 \rangle_\mu)^{1/2}$ for spin 1 so that measure is *not* canonical!

But I quickly determined that one should expect equality in all other cases. I did spin $\frac{3}{2}$ by hand and used Mathematica to compute $\langle (x^2 - a_S)^{2n+1} \rangle_S$ where $a_S = (\int_{\mathbb{R}} x^2 d\tilde{\mu}_S(x))$ for $S = 2, \frac{5}{2}, 3$ and $m = 1, 2, \dots, 10$ and for $S = 20$ and $m = 1, \dots, 5$ and found them all positive which leads to a natural conjecture

$$\langle (x^2 - a_S)^{2n+1} \rangle_S \geq 0 \quad (3.1)$$

As explained earlier, because Wells domination implies Ising domination, one has that for pair interactions

$$T_c(S) \geq T_-(S)^2 T_c\left(\frac{1}{2}\right) \quad (3.2)$$

As it turns out, there is a result of this genre in the literature. In 1969, Griffiths wrote a famous paper [7] on obtaining spin S Ising spins by ferromagnetically coupling $2S$ spin $\frac{1}{2}$ spins together which lead to GKS and Lee Yang for spin S Ising systems. This is usually presented in terms of an elegant coupling discussed in the first part of the paper. Less attention is paid to the second part where he shows instead there

is such a coupling in which S of the spin $\frac{1}{2}$ spins are frozen together (for S half an odd integer, it's $S + \frac{1}{2}$) which he noted implies

$$T_c(S) \geq \frac{1}{4} T_c\left(\frac{1}{2}\right) \quad (3.3)$$

The quantity $a_S = \left(\int_{\mathbb{R}} x^2 d\tilde{\mu}_S(x)\right)$ of (3.1) is equal to $\frac{1}{3} + \frac{1}{3S}$. If one proves that this is T_-^2 for $S \neq 1$, one has for such S that

$$T_c(S) \geq \left(\frac{1}{3} + \frac{1}{3S}\right) T_c\left(\frac{1}{2}\right) \quad (3.4)$$

while for $S = 1$ where we know that one has that $T_-^2 = \frac{1}{2}$

$$T_c(1) \geq \frac{1}{2} T_c\left(\frac{1}{2}\right) \quad (3.5)$$

Not only is this an improvement of Griffiths by more than $\frac{4}{3}$ but in the result for $S \neq 1$, the improved constant is optimal!! For one has equality if T_c is replaced by its mean field values and, as noted by Dyson, Lieb and Simon [5], mean field theory is exact in the nearest neighbor infinite dimension limit.

Rescaling so the maximum spin value is S , the conjecture, (3.1), is the assertion that for $m = 1, 2, \dots$ and $S = \frac{3}{2}, 2, \frac{5}{2}, 3, \dots$, one has that

$$\sum_{j=-S}^S (3j^2 - S(S+1))^{2m+1} \geq 0 \quad (3.6)$$

For S an integer, this is the usual kind of sum. For $2S$ an odd integer, j takes the $2S + 1$ values $-S, -S + 1, \dots, S - 1, S$, i.e. $2j$ is an odd integer. Note, the constant $S(S + 1)$ is such that the sum is zero if $m = 0$

I found this conjecture fascinating and worked on it with no progress for about 7 months. I even got three of my coauthors from other papers to think about it with no luck.

Given that Lieb has a celebrated paper [11] on comparing Heisenberg models (admittedly classical vs. quantum and pressures, not correlations) and that I didn't want to bury in a long book this material which had already been buried for 45 years, it seemed natural to use this for an article when I was asked to contribute to a Festschrift for Elliott's 90th birthday. The paper was due on Jan 31, 2022 and on Friday, Jan 14, I had a first draft of the paper.

It seemed a shame not to make one more push to prove the conjecture so I did the obvious thing. Desperate situations call for desperate measures. At 11 AM on Friday, Jan 14, I sent an email entitled "A challenge" stating the conjectured inequality (and with the draft to explain its significance) to Terry Tao. When I logged on after Shabbat

the next evening I had an email dated Saturday at 1:30 PM with a proof of the conjecture!!!

But the scenario isn't quite as you image it. At 1:30 PM on Friday, Terry had emailed me back: "*I have a postdoc who works on some other inequalities vaguely of this type, I will forward this problem to him and see if he is interested.*" and it was the postdoc, José Madrid, who sent the proof.

4. MAJORIZATION

José's note had one wonderful idea (using Karamata's inequality) and 5 dense pages of calculation to implement it. We Zoomed several times, first for me to offer him a coauthorship (Terry had suggested an appendix) and to discuss simplifying the implementation. We discovered what we thought was a new criteria for majorization that led to a three line proof. OK, a proof with three long lines. We then discovered that the proof was only really simple in case S was half an odd integer. As I'll explain, the integer case is harder but we found a proof in that case that was only a little longer.

As indicated, the key notion is majorization, a set of ideas that go back to Schur [16] in 1923 and Hardy-Littlewood-Pólya [8]. A standard reference is Marshall-Olkin [14] which has been called a love poem to majorization; other references are Hardy-Littlewood-Pólya [9] and Simon [19, Chapters 14-15]. I suspect my coauthors hadn't seen this theory but I didn't have this excuse. My convexity book has a whole chapter on it!

If $\mathbf{x} \in \mathbb{R}_+^n$ (the set with $x_1, x_2 \dots x_n \geq 0$), we define, \mathbf{x}^* , its *decreasing rearrangement* to be the point in \mathbb{R}_+^n whose coordinates are a permutation of those of \mathbf{x} with $x_1^* \geq x_2^* \geq \dots \geq x_n^*$. If $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ we say that \mathbf{x} *majorizes* \mathbf{y} , written $\mathbf{x} \succ \mathbf{y}$ or $\mathbf{y} \prec \mathbf{x}$ if and only if

$$\sum_{j=1}^n x_j^* = \sum_{j=1}^n y_j^*; \quad S_k(\mathbf{x}) \equiv \sum_{j=1}^k x_j^* \geq \sum_{j=1}^k y_j^*, \quad k = 1, \dots, n-1 \quad (4.1)$$

which defines $S_k(\mathbf{x})$. Given $\pi \in \Sigma_n$ the group of permutations of $\{1, \dots, n\}$ and $\mathbf{x} \in \mathbb{R}_+^n$, one defines $\pi^*(\mathbf{x})$ by

$$\pi^*(\mathbf{x})_j = x_{\pi^{-1}(j)} \quad (4.2)$$

Majorization is the basis of a number of inequalities sometimes called rearrangement inequalities. Basic to most of them is

Proposition 4.1. $\mathbf{y} \prec \mathbf{x}$ if and only if \mathbf{y} is in the convex hull in \mathbb{R}^n of the (at most) $n!$ points $\{\pi^*(\mathbf{x})\}_{\pi \in \Sigma_n}$.

This is proven by slicing \mathbb{R}^n with specific hyperplanes; see Simon [18, Theorem 1.9] or Simon [19, Theorem 15.5]). An immediate consequence is

Theorem 4.2 (Karamata's Inequality[10]). *Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ with $\mathbf{x} \succ \mathbf{y}$ and let φ be an arbitrary continuous convex function on $[0, \max_j(x_j)]$. Then*

$$\sum_{j=1}^n \varphi(x_j) \geq \sum_{j=1}^n \varphi(y_j) \quad (4.3)$$

Even though this is widely referred to as Karamata's inequality after Karamata's 1932 paper [10], it or theorems that imply it appear in a 1923 paper of Schur [16] and a 1929 paper of Hardy-Littlewood-Pólya [8]. That said, we note that HLP [8] doesn't have a proof which may not have appeared until their 1934 book [9] and that Karamata proved a converse, namely, if $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ and the inequality holds for all convex φ , then $\mathbf{x} \succ \mathbf{y}$; see Simon [19, Theorem 15.5]).

The proof of Karamata's theorem is simple. One notes the function $\mathbf{w} \mapsto \sum_{j=1}^n \varphi(w_j)$ is convex and permutation symmetric and then uses Proposition 4.1.

Madrid and I found a simple criterion for majorization. Given the vast literature on the subject, we suspected it was already known but couldn't find it before [13] was published although after I gave the talk on this work, we were told that Astashkin et. al [2] had a continuum analog a few months earlier (and it may well appear even earlier somewhere!). Here is our criterion:

Proposition 4.3. *Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ with $\sum_{j=1}^n x_j = \sum_{j=1}^n y_j$ and that for some $\ell \in \{2, \dots, n-1\}$, one has that*

$$j < \ell \Rightarrow x_j^* > y_j^* \quad j \geq \ell \Rightarrow x_j^* \leq y_j^* \quad (4.4)$$

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

Proof. Without loss, we can suppose that $x = x^*, y = y^*$. 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 $\sum_{j=1}^k x_j \geq \sum_{j=1}^k y_j$. \square

With this result in hand, we can turn to the proof of (3.6) at least when S is half an odd integer. Let us begin by explaining the difference between this case and the case of S integral. The point is that as j runs through allowed values, there are degeneracies because $(-j)^2 = j^2$. When $2S$ is odd, the total number of points is even, $j = 0$ is not allowed

and every value occurs with multiplicity 2 so all values with the same weight. When S is integral, 0 is an allowed value of j^2 which has half the weight of every other allowed value of j^2 and this complicates the analysis. Here is the $2S$ odd result and its proof:

Theorem 4.4 (Madrid, Simon and Wells [13]). *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 $[-\|\psi\|_\infty, \|\psi\|_\infty]$ which is continuous, odd and whose restriction to $[0, \|\psi\|_\infty]$ is convex. Let*

$$\bar{\psi} = (N + 1)^{-1} \sum_{j=0}^N \psi\left(\frac{j}{N}\right) \quad (4.5)$$

Then

$$\sum_{j=1}^N \Phi\left(\psi\left(\frac{j}{N}\right) - \bar{\psi}\right) \geq 0 \quad (4.6)$$

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 $N+1$ half odd integers starting at $\frac{1}{2}$, i.e. $\frac{1}{2}, \frac{3}{2}, \dots, N + \frac{1}{2}$. The map $k \mapsto (k - \frac{1}{2})/N$ maps those $N+1$ half odd integers into the points of the theorem. Taking into account that the sum in (3.6) is twice the sum from $\frac{1}{2}$ to S , we see that because j^2 maps to a non-negative, continuous, strictly monotone increasing and convex function under $k \mapsto (k - \frac{1}{2})/N$ and $u \mapsto u^{2m+1}$ is continuous, odd and whose restriction to $[0, \infty)$ is convex, that we have the Corollary below.

Corollary 4.5 (Madrid, Simon and Wells [13]). *(3.6) holds for for $m = 1, 2, \dots$ and $S = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}, \dots$*

Remark. (3.6) continues to hold if u^{2m+1} is replaced by any function which is continuous, odd and whose restriction to $[0, \infty)$ is convex and if j^2 is replaced by any even, non-negative, continuous, function whose restriction to $[0, \infty)$ is strictly monotone increasing and convex.

We need two preliminaries for the proof of Theorem 4.4:

Lemma 4.6. *Let ψ be a convex function on $[0, 1]$ and suppose that*

$$0 \leq \tilde{b} \equiv 2c - b < \tilde{a} \equiv 2c - a \leq c \leq a < b \leq 1 \quad (4.7)$$

Then

$$\frac{1}{2}(\psi(b) + \psi(\tilde{b})) \geq \frac{1}{2}(\psi(a) + \psi(\tilde{a})) \geq \psi(c) \quad (4.8)$$

Moreover, the first inequality is strict unless $\psi'(s)$ is constant on (\tilde{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 $\tilde{b} = 0$ and $\tilde{a} = 1 - a$. 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 [19, Theorem 1.28])

$$\frac{1}{2}(\psi(1) + \psi(0)) - \frac{1}{2}(\psi(a) + \psi(\tilde{a})) = \frac{1}{2} \int_a^1 [\psi'(s) - \psi'(1-s)] ds \quad (4.9)$$

By convexity, the integrand is non-negative so we have proven (4.8). Moreover if $\psi'(s)$ is not constant on $(1-a, a)$, then the integral is strictly positive. \square

Proposition 4.7. *Let ψ , $\bar{\psi}$ and N be as in Theorem 4.4. Then*

$$n \equiv \#\{j \mid \psi\left(\frac{j}{N}\right) \leq \bar{\psi}\} \geq (N+1)/2 \quad (4.10)$$

and

$$\psi\left(\frac{1}{2}\right) \leq \bar{\psi} \leq \frac{1}{2}(\psi(0) + \psi(1)) \quad (4.11)$$

Moreover, the inequalities in (4.11) are strict if $N \geq 2$ and ψ is not an affine function on $[0, 1]$ (i.e. ψ' is not constant).

Proof. For any $j = 0, 1, \dots, N$, (4.8) implies that

$$\psi\left(\frac{1}{2}\right) \leq \frac{1}{2} \left(\psi\left(\frac{j}{N}\right) + \psi\left(1 - \frac{j}{N}\right) \right) \leq \frac{1}{2}(\psi(0) + \psi(1)) \quad (4.12)$$

Averaging over j yields (4.11). If ψ is not affine on $[0, 1]$, then the second inequality is strict for $1 \leq j \leq N-1$ so the second inequality in (4.11) is strict. Since $\psi\left(\frac{1}{2}\right) < \frac{1}{2}(\psi(0) + \psi(1))$ if ψ is not affine, we see that in the case the first inequality is always strict.

Since ψ is strictly monotone, the first inequality in (4.11) implies the unique $x \in [0, 1]$ with $\psi(x) = \bar{\psi}$ has $x \geq \frac{1}{2}$. This implies that $n = \#\{j \mid \frac{j}{N} \leq x\} \geq \#\{j \mid \frac{j}{N} \leq \frac{1}{2}\} \geq (N+1)/2$. \square

Proof of Theorem 4.4. Let $q = N+1-n \leq n$ by (4.9). Define

$$y_j = \bar{\psi} - \psi\left(\frac{j-1}{N}\right) \quad j = 1, \dots, n \quad (4.13)$$

$$x_j = \begin{cases} \psi\left(\frac{N+1-j}{N}\right) - \bar{\psi} & \text{if } j = 1, \dots, q \\ 0, & \text{if } j \geq q \end{cases} \quad (4.14)$$

Since ψ is monotone and n is defined by (4.10), we have that $\mathbf{x}, \mathbf{y} \in \mathbb{R}_{+, \geq}^n = \{\mathbf{x} \in \mathbb{R}_+^n \mid x_1 \geq \dots \geq x_n\}$. By the definition of $\bar{\psi}$, we have that

$$\sum_{j=1}^n x_j = \sum_{j=1}^n y_j \quad (4.15)$$

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 (4.1) holds. Thus henceforth we will suppose that $N \geq 2$ and ψ is not an affine function on $[0, 1]$, so, in particular, the inequalities in (4.11) are strict.

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

$$m < p \Rightarrow \psi\left(\frac{m+1}{N}\right) - \psi\left(\frac{m}{N}\right) \leq \psi\left(\frac{p+1}{N}\right) - \psi\left(\frac{p}{N}\right) \quad (4.16)$$

By the strict form of (4.11), $x_1 > y_1$. Because of (4.15), there must be a first ℓ so that $x_\ell \leq y_\ell$. We claim that if $\ell < n$, then $x_{\ell+1} \leq y_{\ell+1}$. If $\ell + 1 > q$, then $x_{\ell+1} = 0$ and the required inequality is immediate. If $\ell + 1 \leq q$, then (4.16) implies that $x_\ell - x_{\ell+1} \geq y_\ell - y_{\ell+1}$. Subtracting this from $x_\ell \leq y_\ell$ proves that $x_{\ell+1} \leq y_{\ell+1}$. Repeating this argument, proves that for all $j \geq \ell$ we have that $x_j \leq y_j$. Thus by Proposition 4.3, $\mathbf{x} \succ \mathbf{y}$.

By Karamata's inequality, (4.3), we conclude that $\sum_{j=1}^n \Phi(x_j) - \sum_{j=1}^n \Phi(y_j) \geq 0$. Since Φ is odd and $\Phi(0) = 0$, this is equivalent to (4.6). \square

Example 4.8. To understand why we need the extra condition to handle the case when S is integral, consider $d\tilde{\mu}_S$ for $S = 6$ scaled to have spacing 1, i.e. 13 pure points with weight $1/13$ at $0, \pm 1, \pm 2, \pm 3, \pm 4, \pm 5, \pm 6$. The average of the square is 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 14 and one sees that (ignore \mathbf{w} for now)

$$\begin{aligned} \mathbf{x} &= 22, 22, 11, 11, 2, 2, 0 \\ \mathbf{y} &= 14, 13, 13, 10, 10, 5, 5 \\ \mathbf{w} &= 22, 22, 0, 11, 11, 2, 2 \end{aligned} \quad (4.17)$$

One can verify that $\mathbf{x} \succ \mathbf{y}$ by hand (and, below, we will prove the result for all $S \geq 2$) but one can't use Proposition 4.3 as we did in our proof of Theorem 4.4 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 \mathbf{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 4.3 to prove the partial sums of \mathbf{w} dominate those of \mathbf{y} and note it is trivial that partial sums of \mathbf{x} dominate those of \mathbf{w} . The key is that by moving the 0, the pairs are no longer shifted.

We will need an extra condition that implies as in this example, the partial sum of the first two x_j 's dominates the partial sum of the first three y_k 's. The general theorem analogous to Theorem 4.4 is

Theorem 4.9 (Madrid, Simon and Wells [13]). *Fix 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 $\bar{\psi}$ be given by*

$$\bar{\psi} = (2N + 1)^{-1} \sum_{j=-N}^N \psi\left(\frac{j}{N}\right) \quad (4.18)$$

Suppose that

$$2\psi(1) + \psi(0) + 2\psi\left(\frac{1}{N}\right) \geq 5\bar{\psi} \quad (4.19)$$

If N is odd, suppose that

$$\psi\left(\frac{1}{2} + \frac{1}{2N}\right) \leq \bar{\psi} \quad (4.20)$$

Then

$$\sum_{j=-N}^N \Phi\left(\psi\left(\frac{j}{N}\right) - \bar{\psi}\right) \geq 0 \quad (4.21)$$

Remarks. 1. The condition (4.19) is exactly the condition that the partial sum of the first 2 x_j 's dominates the partial sum of the first three y_k 's.

2. It might be true that this theorem holds without the need for the condition (4.19) but it holds in the case we need so MSW didn't try hard to eliminate it. The example above shows why a naive extension of the proof of Theorem 4.4 doesn't work and led to the extra condition. We do note that (4.19) is a restriction. If we normalize ψ by $\psi(0) = 0, \psi(1) = 1$, then in the limit as $N \rightarrow \infty$, (4.19) becomes

$$\int_0^1 \psi(x) dx \leq \frac{2}{5} \quad (4.22)$$

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

3. On the other hand, (4.20) is quite natural independent of our method of proof; see Madrid, Simon and Wells [13].

The reader can check Madrid, Simon and Wells [13] for the proof but we note it combines the ideas of the proof of Theorem 4.4 and Example 4.8. We should give the details of checking that the Theorem proves (3.6) for $S = 2, 3, 4, \dots$. One needs to check (4.18) and (4.20), the latter for $S(= N)$ odd and $S \geq 3$. These are

$$2S^2 + 2 \geq \frac{5}{3}S(S + 1); S = 2, 3, 4, \dots \quad (4.23)$$

$$S^2\left(\frac{1}{2} + \frac{1}{2S}\right)^2 \leq \frac{1}{3}S(S + 1); S = 3, 5, 7, \dots \quad (4.24)$$

(4.23) is equivalent to $0 \leq S^2 - 5S + 6 = (S - 2)(S - 3)$ which holds for all integral S . (4.24) is equivalent to $0 \leq S^2 - 2S - 3 = (S - 3)(S + 1)$ which holds for all $S \geq 3$.

5. THE END OF THE STORY

In our first Zoom call, José also suggested it would be good to try again to locate Daniel Wells. I wasn't starting at ground zero. While I got nothing from Indiana University, I talked about this material during the conference in honor of my 75th birthday and Leonard Schulman, a computer scientist at Caltech (and son of a student of Arthur Wightman), heard my talk and did some Google searching. He found a short story available via Kindle on Amazon whose About the Author read

Daniel R Wells was born in Sterling, Illinois on March 15, 1945. He attended the local parochial schools and graduated from high school in 1963. In October of that year he enlisted in the United States Navy and served for four years. After the Navy, he started college in 1968, studying mathematics, eventually earning a PhD from Indiana University in 1977. He taught mathematics for two years at Texas A&M and then returned to school at the University of Illinois to study computer science. He achieved a PhD in 1982 and worked for various companies as a software engineer until he retired in 2004.

I wasn't clever enough to pull on the right threads of this fabric. Since I had friends at Texas A&M, I consulted them to see if they could find any record. Nope. I tried to leave a "review" of his book saying I wanted to contact the author about his thesis but Amazon said it wasn't a review and wouldn't post it. I bought his Kindle book hoping it might provide more information but it didn't. What I should have done is contact U of I computer science where he got his second PhD. and where he has continued to do some teaching.

Spurred by José, I posted a message on Facebook where I have a group of friends mainly mathematicians and theoretical physicists. The message gave some background and asked if anyone had any idea how to follow up. Joshua Paik, a math grad student at Penn State told me he regarded himself as an internet sleuth. The next morning I had a link in a private message from Mr. Paik to a Find a Person internet site with the right name, the right age who lived in the town where the Amazon profile said Wells was born. Shortly after that, Mr. Paik sent me what he though might be Wells' email address. I contacted the email address asking if the recipient was a Daniel Wells who got a math PhD in Indiana then sent him the current draft and asked him

to be a coauthor - after all, 2.5 out of 6 sections were from his thesis! He agreed, so in less than a week, I picked up two coauthors.

The next week, José and I zoomed with Daniel and I got some more background. Wells had gone to Texas A&M for a postdoc, written up his thesis with the addition of the rotor-Ising comparison theorem and sent the preprint that Aizenman and I referred to off to a journal where it should have been accepted. But it was rejected. At this point, his thesis advisor should have stepped in and explained the facts of life: just as there are bad papers, there are bad referees and one should send the paper off to another journal. But alas, Slim Sherman, his advisor, had passed away shortly before he took his oral exam and wasn't there to advise him. This was in the old days when postdocs didn't have formal advisors! Wells was so discouraged, he totally left mathematics even though he'd written a very good thesis. Sometimes the system doesn't work.

The moral is that young workers in particular need someone to say to them: *Jerk, submit it to another journal*. Your reaction might be that I'm the wrong one to say that since surely I've never had any of my papers rejected. But I've had many papers rejected. I thought I'd close with two amusing stories of rejections so that you might even respond to rejection with some humor.

One involves my paper with Elliott Lieb [12] whose main result was that Thomas-Fermi Theory was exact in a certain limit of large Z atoms and molecules. It was significant enough that on its 35th anniversary, some quantum chemists had a conference marking the occasion! This announcement was originally rejected by Physical Review Letters with a report that paraphrased began *This paper is one of the worst papers I have ever seen. It is a sequence of unproven assertions (true; it was an announcement after all) many of which are obviously wrong. For example, the author assert that the Thomas-Fermi density is C^∞ which would make it 0, 1 or ∞ depending on the value of C* . The referee was clearly unfamiliar with modern mathematical notation and incompetent to evaluate a paper in mathematical physics. We complained and asked for a second referee who accepted it. (General, I do not recommend resubmitting to the same journal).

The other involves my paper with Christiansen and Zinchenko [4] which settled a 40 year old conjecture of Widom [24] which was regarded as a major open question in the asymptotics of extremal polynomials. I thought it was good enough that we should submit it to one of the top three math journals but for various reasons, one of my coauthors wanted it to go to a journal just below those three. We got a quick rejection that paraphrased said *It is nice to have a 40 year old*

conjecture resolved but this paper should be rejected because the proof is too easy so it isn't up to the high standard of < journalname >. I was scandalized by this report so we submitted it after all to a top three journal where it was accepted!

REFERENCES

- [1] Aizenman, M., and Simon, B., *A comparison of plane rotor and Ising models*, Phys. Lett. **A76** (1980), 281–282.
- [2] Astashkin, S. V., Lykov, K. V., and Milman, M., *Majorization revisited: Comparison of norms in interpolation scales*, Math. Ann. (2023), <https://doi.org/10.1007/s00208-023-02606-w>.
- [3] Bricmont, J., Lebowitz, J. L., and Pfister, C. E., *Periodic Gibbs states of ferromagnetic spin systems*, J. Stat. Phys. **24** (1981), 269–277.
- [4] J. S. Christiansen, B. Simon, and M. Zinchenko, *Asymptotics of Chebyshev Polynomials, I. Subsets of \mathbb{R}* , Invent. Math. **208** (2017), 217–245.
- [5] Dyson, F., J., Lieb, E. H., and Simon, B., *Phase transitions in quantum spin systems with isotropic and nonisotropic interactions*, J. Stat. Phys. **18** (1978), 335–383.
- [6] Ginibre, J., *General formulation of Griffiths' inequalities*, Comm. Math. Phys. **16** (1970), 310–328.
- [7] Griffiths, R. B., *Rigorous results for Ising ferromagnets of arbitrary spin*, J. Math. Phys. **10** (1969), 1559–1565.
- [8] Hardy, G. H., Littlewood, J. E., and Pólya, G., *Some simple inequalities satisfied by convex functions*, Messenger of Math. **58** (1929), 145–152.
- [9] Hardy, G. H., Littlewood, J. E., and Pólya, G., *Inequalities*, Cambridge University Press, Cambridge, 1934.
- [10] Karamata, J., *Sur une inégalité relative aux fonctions convexes*, Publ. Math. Univ. Belgrade **1** (1932), 145–148.
- [11] Lieb, E. H., *The classical limit of quantum spin systems*, Comm. Math. Phys. **31** (1973), 327–340.
- [12] E. Lieb and B. Simon, *Thomas-Fermi theory revisited*, Phys. Rev. Lett. **31** (1973), 681–683.
- [13] Madrid, J., Simon, B., and Wells, D. R., *Comparison of Ising Models Under Change of Apriori Measure*, in *The physics and mathematics of Elliott Lieb—the 90th anniversary*, Vol. II., eds. Frank, R. L., Laptev, A., Lewin, M. and Seiringer, R. EMS Press, Berlin, 2022, pp. 47–71.
- [14] Marshall, A. W., and Olkin, I., *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York- London, 1979 (second edition with Arnold, B. C. in 2009)
- [15] Newman, C. M., *Asymptotic independence and limit theorems for positively and negatively dependent random variables*, in *Inequalities in statistics and probability*, IMS Lecture Notes Monogr. Ser., **5**, Inst. Math. Statist., Hayward, CA, 1984, pp. 127–140.
- [16] Schur, I., *Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie*, Sitzungsber. Berlin Math. Gesellschaft **22** (1923), 9–20.

- [17] Simon, B., *The Statistical Mechanics of Lattice Gases*, Princeton University Press, Princeton, 1993.
- [18] Simon, B., *Trace Ideals and Their Applications, second edition*, Mathematical Surveys and Monographs **120**, American Mathematical Society, 2005 (first edition, 1979).
- [19] Simon, B., *Convexity: An Analytic Viewpoint*, Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 2011.
- [20] Simon, B., *A Tales of Three Coauthors: Comparison of Ising Models*, Talk given at International Biweekly Online Seminar on Analysis, Differential Equations And Mathematical Physics on September 21, 2023, video at https://www.youtube.com/watch?v=zM_Ng8z7yWg.
- [21] Simon, B., *Phase Transitions in the Theory of Lattice Gases*, Cambridge University Press, Cambridge, in preparation.
- [22] Van Beijeren, H., and Sylvester, G. S., *Phase transitions for continuous-spin Ising ferromagnets*, J. Func. Anal. **28** (1978), 145–167.
- [23] Wells, D. R., *Some moment inequalities and a result on multivariate unimodality*, (unpublished) Indiana University Ph. D. Thesis, 1977.
- [24] H. Widom, *Extremal polynomials associated with a system of curves in the complex plane*, Adv. in Math. **3** (1969), 127–232.