Quantum Heisenberg models and their probabilistic representations

Christina Goldschmidt, Daniel Ueltschi, and Peter Windridge

ABSTRACT. These notes give a mathematical introduction to two seemingly unrelated topics: (i) quantum spin systems and their cycle and loop representations, due to Tóth and Aizenman-Nachtergaele; (ii) coagulation-fragmentation stochastic processes. These topics are nonetheless related, as we argue that the lengths of cycles and loops satisfy an effective coagulation-fragmentation process. This suggests that their joint distribution is Poisson-Dirichlet. These ideas are far from being proved, but they are backed by several rigorous results, notably of Dyson-Lieb-Simon and Schramm.

CONTENTS

1. Introduction

We review cycle and loop models that arise from quantum Heisenberg spin systems. The loops and cycles are geometric objects defined on graphs. The main goal is to understand properties such as their length in large graphs.

The cycle model was introduced by Tóth as a probabilistic representation of the Heisenberg ferromagnet [45], while the loop model is due to Aizenman and Nachtergaele and is related to the Heisenberg antiferromagnet [1]. Both models are built on the random stirring process of Harris [27] and have an additional geometric weight of the form $\vartheta^{\#\text{cycles}}$ or $\vartheta^{\#\text{loops}}$ with parameter $\vartheta = 2$. Recently, Schramm studied the cycle model on the complete graph and with $\vartheta = 1$ (that is, without this factor) [41]. He showed in particular that cycle lengths are generated by a split-merge process (or "coagulation-fragmentation"), and that the cycle lengths are distributed as Poisson-Dirichlet with parameter 1.

The graphs of physical relevance are regular lattices such as $\mathbb{Z}^d$ (or large finite boxes in $\mathbb{Z}^d$), and the factor $2^{\#\text{objects}}$ need be present. What should we expect in this case? A few hints come from the models of spatial random permutations, which also involve one-dimensional objects living in higher dimensional spaces. The average length of the longest cycle in lattice permutations was computed numerically in [24]. In retrospect, it suggests that the cycle lengths have the Poisson-Dirichlet distribution. In the "annealed" model where positions are averaged, this was proved in [8]; the mechanisms at work there (i.e., Bose-Einstein condensation and non-spatial random permutations with Ewens distribution) seem very specific, though.

We study the cycle and loop models in $\mathbb{Z}^d$ with the help of a stochastic process whose invariant measure is identical to the original measure with weight $\vartheta^{\#\text{cycles}}$ or $\vartheta^{\#\text{loops}}$, and which leads to an effective split-merge process for the cycle (or loop) lengths. The rates at which the splits and the merges take place depends on ϑ. This allows to identify the invariant measure, which turns out to be Poisson-Dirichlet

with parameter ϑ. While we cannot make these ideas mathematically rigorous, they are compatible with existing results.

As mentioned above, cycle and loop models are closely related to Heisenberg models. In particular, the cycle and loop geometry is reflected in some important quantum observables. The latter have been the focus of intense study by mathematical and condensed matter physicists, who have used imagination and clever observations to obtain remarkable results in the last few decades. Most relevant to us is the theorem of Mermin and Wagner about the absence of magnetic order in one and two dimensions [**35**], and the theorem of Dyson, Lieb, and Simon, about the existence of magnetic order in the antiferromagnetic model in dimensions 5 and more [**16**]. We review these results and explain their implications for cycle and loop models.

Many a mathematician is disoriented when wandering in the realm of quantum spin systems. The landscape of 2×2 matrices and finite-dimensional Hilbert spaces looks safe and easy. Yet, the proofs of many innocent statements are elusive, and one feels quickly lost. It has seemed to us a useful task to provide a detailed introduction to the Heisenberg models in both their quantum and statistical mechanical aspects. We also need notions of stochastic processes, split-merge, and Poisson-Dirichlet. The latter two are little known outside of probability and are not readily accessible to mathematical physicists and analysts, since the language and the perspective of those domains are quite different (see e.g. the dictionary of [**19**], p. 314, between analysts' language and probabilists' "dialect"). In these notes, we have attempted to introduce those different notions in a self-contained fashion.

1.1. Guide to notation. The following objects play a central role in these notes.

$\Lambda = (\mathcal{V}, \mathcal{E})$	A finite graph with undirected edges.
$\rho_{\Lambda, \beta}(d\omega)$	Probability measure for Poisson point processes on $[0, \beta]$ ($\beta > 0$) attached to each edge of Λ (defined in §3.1).
$\mathcal{C}(\omega), \mathcal{L}(\omega)$	Cycle and loop configurations constructed from the edges in ω (§3.1).
γ	A cycle (or loop) in $\mathcal{C}(\omega)$ (or $\mathcal{L}(\omega)$).
$\vartheta > 0$	A geometric weight for the number of cycles and loops in a configuration.
Δ_1	Countable partitions of $[0, 1]$ with parts in decreasing order, i.e. $\{p_1 \geq p_2 \geq \ldots \geq 0 : \sum_i p_i = 1\}$.
PD_θ	The Poisson-Dirichlet distribution with parameter $\theta > 0$ on Δ_1 (§7.2).
$(\mathfrak{X}_t, t \geq 0)$	A stochastic process with invariant measure given by our cycle and loop models (§8.2).

2. Hilbert space, spin operators, Heisenberg Hamiltonian

We review the setting for quantum lattice spin systems described by Heisenberg models. Spin systems are relevant for the study of electronic properties of condensed matter. Atoms form a regular lattice and they host localized electrons, that are characterized only by their spin. Interactions are restricted to neighboring spins. One is interested in equilibrium properties of large systems. There are two

closely related quantum Heisenberg models, that describe ferromagnets and anti-ferromagnets, respectively. The material is standard and the interested reader is encouraged to look in the references [**40, 43, 36, 17**] for further information.

2.1. Graphs and Hilbert space. Let $\Lambda = (\mathcal{V}, \mathcal{E})$ be a graph, where $\mathcal{V}$ is a finite set of vertices and $\mathcal{E}$ is the set of "edges", i.e. unordered pairs in $\mathcal{V} \times \mathcal{V}$. From a physical perspective, relevant graphs are regular graphs such as $\mathbb{Z}^d$ (or a finite box in $\mathbb{Z}^d$) with nearest-neighbor edges, but it is mathematically advantageous to allow for more general graphs. We restrict ourselves to spin-$\frac{1}{2}$ systems, mainly because the stochastic representations only work in this case.

To each site $x \in \mathcal{V}$ is associated a 2-dimensional Hilbert space $\mathcal{H}_x = \mathbb{C}^2$. It is convenient to use Dirac's notation of "bra", $\langle \cdot |$, and "ket", $| \cdot \rangle$, in which we identify

$$| \tfrac{1}{2} \rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad | - \tfrac{1}{2} \rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.1}$$

The notation $\langle f | g \rangle$ means the inner product; we use the convention that it is linear in the second variable (and antilinear in the first). Occasionally, we also write $\langle f | A | g \rangle$ for $\langle f | Ag \rangle$. The Hilbert space of a quantum spin system on Λ is the tensor product

$$\mathcal{H}^{(\mathcal{V})} = \bigotimes_{x \in \Lambda} \mathcal{H}_x, \tag{2.2}$$

which is the $2^{|\mathcal{V}|}$ dimensional space spanned by elements of the form $\otimes_{x \in \mathcal{V}} f_x$ with $f_x \in \mathcal{H}_x$. The inner product between two such vectors is defined by

$$\left\langle \otimes_{x \in \mathcal{V}} f_x \,\middle|\, \otimes_{x \in \mathcal{V}} g_x \right\rangle_{\mathcal{H}^{(\mathcal{V})}} = \prod_{x \in \Lambda} \langle f_x | g_x \rangle_{\mathcal{H}_x}. \tag{2.3}$$

The inner product above extends by (anti)linearity to the other vectors, which are all linear combinations of vectors of the form $\otimes_{x \in \mathcal{V}} f_x$.

The basis (2.1) of $\mathbb{C}^2$ has a natural extension in $\mathcal{H}^{(\mathcal{V})}$; namely, given $s^{(\mathcal{V})} = (s_x)_{x \in \mathcal{V}}$ with $s_x = \pm \frac{1}{2}$, let

$$| s^{(\mathcal{V})} \rangle = \bigotimes_{x \in \Lambda} | s_x \rangle. \tag{2.4}$$

These elements are orthonormal, i.e.

$$\langle s^{(\mathcal{V})} | \tilde{s}^{(\mathcal{V})} \rangle = \prod_{x \in \mathcal{V}} \langle s_x | \tilde{s}_x \rangle = \prod_{x \in \mathcal{V}} \delta_{s_x, \tilde{s}_x}, \tag{2.5}$$

where δ is Kronecker's symbol, $\delta_{ab} = 1$ if $a = b$, 0 otherwise.

2.2. Spin operators. In the quantum world physical relevant quantities are called *observables* and they are represented by self-adjoint operators. The operators for the observable properties of our spin 1/2 particles are called the Pauli matrices, defined by.

$$S^{(1)} = \tfrac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad S^{(2)} = \tfrac{1}{2} \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad S^{(3)} = \tfrac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.6}$$

We interpret $S^{(i)}$ as the spin component in the ith direction. The matrices are clearly Hermitian and satisfy the relations

$$[S^{(1)}, S^{(2)}] = \mathrm{i} S^{(3)}, \quad [S^{(2)}, S^{(3)}] = \mathrm{i} S^{(1)}, \quad [S^{(3)}, S^{(1)}] = \mathrm{i} S^{(2)}. \tag{2.7}$$

These operators have natural extensions of the spin operators in $\mathcal{H}^{(\mathcal{V})}$. Let $x \in \mathcal{V}$, and write $\mathcal{H}^{(\mathcal{V})} = \mathcal{H}_x \otimes \mathcal{H}^{(\mathcal{V} \setminus \{x\})}$. We define the operators $S_x^{(i)}$ indexed by $x \in \mathcal{V}$ by

$$S_x^{(i)} = S^{(i)} \otimes \mathrm{Id}_{\mathcal{V} \setminus \{x\}}. \tag{2.8}$$

The commutation relations (2.7) extend to the operators $S_x^{(i)}$, namely

$$[S_x^{(1)}, S_y^{(2)}] = \mathrm{i} \delta_{xy} S_x^{(3)}, \tag{2.9}$$

and all other relations obtained by cyclic permutations of (123). It is indeed not hard to check that the matrix elements $\langle s^{(\mathcal{V})} | \cdot | \tilde{s}^{(\mathcal{V})} \rangle$ of both sides are identical for all $s^{(\mathcal{V})} \in \{-\frac{1}{2}, \frac{1}{2}\}^{\mathcal{V}}$. It is customary to introduce the notation $\vec{S}_x = (S_x^{(1)}, S_x^{(2)}, S_x^{(3)})$, and

$$\vec{S}_x \cdot \vec{S}_y = S_x^{(1)} S_y^{(1)} + S_x^{(2)} S_y^{(2)} + S_x^{(3)} S_y^{(3)}. \tag{2.10}$$

Note that operators of the form $S_x^{(i)} S_y^{(j)}$, with $x \neq y$, act in $\mathcal{H}^{(\mathcal{V})} = \mathcal{H}_x \otimes \mathcal{H}_y \otimes \mathcal{H}^{(\mathcal{V} \setminus \{x,y\})}$ as follows

$$S_x^{(i)} S_y^{(j)} = S^{(i)} \otimes S^{(j)} \otimes \mathrm{Id}_{\mathcal{V} \setminus \{x,y\}}. \tag{2.11}$$

In the case $x = y$, and using $(S_x^{(i)})^2 = \frac{1}{4} \mathrm{Id}_{\mathcal{V}}$, we get

$$\vec{S}_x^2 = (S_x^{(1)})^2 + (S_x^{(2)})^2 + (S_x^{(3)})^2 = \frac{3}{4} \mathrm{Id}_{\mathcal{V}}. \tag{2.12}$$

LEMMA 2.1. *Consider $\vec{S}_x \cdot \vec{S}_y$ in $\mathcal{H}_x \otimes \mathcal{H}_y$. It is self-adjoint, and its eigenvalues and eigenvectors are*

- $-\frac{3}{4}$ *is an eigenvalue with multiplicity 1; the eigenvector is $\frac{1}{\sqrt{2}}(|\frac{1}{2}, -\frac{1}{2}\rangle - |-\frac{1}{2}, \frac{1}{2}\rangle)$.*
- $\frac{1}{4}$ *is an eigenvalue with multiplicity 3; the three orthonormal eigenvectors are*

$$|\tfrac{1}{2}, \tfrac{1}{2}\rangle, \qquad |-\tfrac{1}{2}, -\tfrac{1}{2}\rangle, \qquad \frac{1}{\sqrt{2}}\big(|\tfrac{1}{2}, -\tfrac{1}{2}\rangle + |-\tfrac{1}{2}, \tfrac{1}{2}\rangle\big).$$

The eigenvector corresponding to $-\frac{3}{4}$ is called a "singlet state" by physicists, while the eigenvectors for $\frac{1}{4}$ are called "triplet states".

PROOF. We have for all $a, b = \pm\frac{1}{2}$,

$$\begin{aligned} S_x^{(1)} S_y^{(1)} |a, b\rangle &= \tfrac{1}{4} |-a, -b\rangle, \\ S_x^{(2)} S_y^{(2)} |a, b\rangle &= -ab |-a, -b\rangle, \\ S_x^{(3)} S_y^{(3)} |a, b\rangle &= ab |a, b\rangle. \end{aligned} \tag{2.13}$$

The lemma follows from straightforward linear algebra. $\qquad\square$

2.3. Hamiltonians and magnetization. We can now introduce the Heisenberg Hamiltonians. These self-adjoint operators represent the energy of the system.

$$\begin{aligned} H_{\Lambda,h}^{\mathrm{ferro}} &= -\sum_{\{x,y\} \in \mathcal{E}} \vec{S}_x \cdot \vec{S}_y - h \sum_{x \in \mathcal{V}} S_x^{(3)}, \\ H_{\Lambda,h}^{\mathrm{anti}} &= +\sum_{\{x,y\} \in \mathcal{E}} \vec{S}_x \cdot \vec{S}_y - h \sum_{x \in \mathcal{V}} S_x^{(3)}. \end{aligned} \tag{2.14}$$

Next, let M_Λ be the operator that represents the magnetization in the 3rd direction.

$$M_\Lambda^{(3)} = \sum_{x \in \mathcal{V}} S_x^{(3)}. \tag{2.15}$$

LEMMA 2.2. *Hamiltonian and magnetization operators commute,*

$$[H_{\Lambda,h}, M_\Lambda] = 0.$$

PROOF. This follows from the commutation relations (2.9). Namely, using the fact that $S_x^{(i)}$ and $S_y^{(3)}$ commute for $x \neq y$,

$$[H_{\Lambda,h}, M_\Lambda] = \sum_{\{x,y\} \in \mathcal{E}, z \in \mathcal{V}} [\vec{S}_x \cdot \vec{S}_y, S_z^{(3)}]$$

$$= \sum_{\{x,y\} \in \mathcal{E}} \left([S_x^{(1)} S_y^{(1)}, S_x^{(3)}] + [S_x^{(1)} S_y^{(1)}, S_y^{(3)}] + [S_x^{(2)} S_y^{(2)}, S_x^{(3)}] + [S_x^{(2)} S_y^{(2)}, S_y^{(3)}] \right).$$

$$\tag{2.16}$$

The first commutator is

$$[S_x^{(1)} S_y^{(1)}, S_x^{(3)}] = [S_x^{(1)}, S_x^{(3)}] S_y^{(1)} = -\mathrm{i} S_x^{(2)} S_y^{(1)}, \tag{2.17}$$

and the others are similar. We get

$$[H_{\Lambda,h}, M_\Lambda] = \mathrm{i} \sum_{\{x,y\} \in \mathcal{E}} \left(-S_x^{(2)} S_y^{(1)} - S_x^{(1)} S_y^{(2)} + S_x^{(1)} S_y^{(2)} + S_x^{(2)} S_y^{(1)} \right) = 0. \tag{2.18}$$

$\square$

2.4. Gibbs states and free energy. The equilibrium states of quantum statistical mechanics are given by *Gibbs states* $\langle \cdot \rangle_{\Lambda,\beta,h}$. These are nonnegative linear functionals on the space of operators in $\mathcal{H}^{(\mathcal{V})}$ of the form

$$\langle A \rangle_{\Lambda,\beta,h} = \frac{1}{Z_\Lambda(\beta, h)} \operatorname{Tr} A \, \mathrm{e}^{-\beta H_{\Lambda,h}} , \tag{2.19}$$

where the normalization

$$Z_\Lambda(\beta, h) = \operatorname{Tr} \mathrm{e}^{-\beta H_{\Lambda,h}} . \tag{2.20}$$

is called the *partition function*. Here, Tr represents the usual matrix trace.

There are deep reasons for the Gibbs states to describe equilibrium states but we will not dwell on them here. We now introduce the *free energy* $F_\Lambda(\beta, h)$. Its physical motivation is that it provides a connection to thermodynamics. It is a kind of generating function and it is therefore useful mathematically. The definition of the free energy in our case is

$$F_\Lambda(\beta, h) = -\frac{1}{\beta} \log Z_\Lambda(\beta, h). \tag{2.21}$$

LEMMA 2.3. *The function $\beta F_\Lambda(\beta, h)$ is concave in $(\beta, \beta h)$.*

PROOF. We rather check that $-F_\Lambda$ is convex, which is the case if the matrix

$$\begin{pmatrix} \frac{\partial^2 \beta F_\Lambda}{\partial \beta^2} & \frac{\partial^2 \beta F_\Lambda}{\partial \beta \partial (\beta h)} \\ \frac{\partial^2 \beta F_\Lambda}{\partial \beta \partial (\beta h)} & \frac{\partial^2 \beta F_\Lambda}{\partial (\beta h)^2} \end{pmatrix}$$

is positive definite. Let us write $\langle \cdot \rangle$ instead of $\langle \cdot \rangle_{\Lambda,\beta,h}$. We have

$$\frac{\partial^2}{\partial \beta^2} \beta F_\Lambda(\beta, h) = -\big\langle (H_{\Lambda,0} - \langle H_{\Lambda,0}\rangle)^2 \big\rangle,$$

$$\frac{\partial^2}{\partial (\beta h)^2} \beta F_\Lambda(\beta, h) = -\big\langle (M_\Lambda - \langle M_\Lambda\rangle)^2 \big\rangle, \tag{2.22}$$

$$\frac{\partial^2}{\partial \beta \partial (\beta h)} \beta F_\Lambda(\beta, h) = \big\langle (H_{\Lambda,0} - \langle H_{\Lambda,0}\rangle)(M_\Lambda - \langle M_\Lambda\rangle) \big\rangle.$$

Then F_Λ is convex if

$$\big\langle (H_{\Lambda,0} - \langle H_{\Lambda,0}\rangle)(M_\Lambda - \langle M_\Lambda\rangle) \big\rangle^2 \le \big\langle (H_{\Lambda,0} - \langle H_{\Lambda,0}\rangle)^2 \big\rangle \big\langle (M_\Lambda - \langle M_\Lambda\rangle)^2 \big\rangle. \tag{2.23}$$

It is not hard to check that the map $(A, B) \mapsto \langle A^* B \rangle$ is an inner product on the space of operators that commute with $H_{\Lambda,h}$. Then

$$|\langle A^* B \rangle|^2 \le \langle A^* A \rangle \langle B^* B \rangle \tag{2.24}$$

by the Cauchy-Schwarz inequality, and this implies (2.23) in particular. $\square$

Concave functions are continuous. But it is useful to establish that $F_\Lambda(\beta, h)$ is uniformly continuous on compact domains.

LEMMA 2.4.

$$\big|\beta F_\Lambda(\beta, h) - \beta' F_\Lambda(\beta', h')\big| \le |\beta - \beta'|\big(\tfrac{3}{4}|\mathcal{E}| + \tfrac{|h|}{2}|\mathcal{V}|\big) + \tfrac{1}{2}\beta|h - h'||\mathcal{V}|.$$

PROOF. We have

$$\beta F_\Lambda(\beta, h) - \beta' F_\Lambda(\beta', h) = \int_{\beta'}^{\beta} \frac{\mathrm{d}}{\mathrm{d}s} s F_\Lambda(s, h) \mathrm{d}s = \int_{\beta'}^{\beta} \langle H_{\Lambda,h}\rangle_{\Lambda,s,h} \mathrm{d}s. \tag{2.25}$$

We can also check that $\beta F_\Lambda(\beta, h) - \beta F_\Lambda(\beta, h') = \int_{h'}^{h} \langle M_\Lambda\rangle_{\Lambda,\beta,s} \mathrm{d}s$. The result follows from $|\langle A \rangle_{\Lambda,\beta,h}| \le \|A\|$ for any operator A, and from $\|\vec{S}_x \cdot \vec{S}_y\| = \tfrac{3}{4}$ (cf Lemma 2.1) and $\|S_x^{(3)}\| = \tfrac{1}{2}$. $\square$

2.5. Symmetries. In quantum statistical mechanics, a symmetry is represented by a unitary transformation that leaves the Hamiltonian invariant. It follows that (finite volume) Gibbs states also possess the symmetry. However, infinite volume states may lose it. This is called *symmetry breaking* and is a manifestation of a phase transition. We only mention the "spin flip" symmetry here, corresponding to the unitary operator

$$U|s^{(\mathcal{V})}\rangle = |-s^{(\mathcal{V})}\rangle. \tag{2.26}$$

One can check that $U^{-1} S_x^{(i)} S_y^{(i)} U = S_x^{(i)} S_y^{(i)}$ and $U^{-1} S_x^{(3)} U = -S_x^{(3)}$. It follows that

$$U^{-1} H_{\Lambda,h} U = H_{\Lambda,-h}. \tag{2.27}$$

This applies to both the ferromagnetic and antiferromagnetic Hamiltonians. It follows that $F_\Lambda(\beta, -h) = F_\Lambda(\beta, h)$, so the free energy is symmetric as a function of h.

3. Stochastic representations

Stochastic representations of quantum lattice models go back to Ginibre, who used a Peierls contour argument to prove the occurrence of phase transitions in anisotropic models [**25**]. Conlon and Solovej introduced a random walk representation for the ferromagnetic model and used it to get an upper bound on the free energy [**12**]. A different representation was introduced by Tóth, who improved the previous bound [**45**]. Further work on quantum models using similar representations include the quantum Pirogov-Sinai theory [**10, 14**] and Ising models in transverse magnetic field [**28, 13, 26**].

A major advantage of Tóth's representation is that spin correlations have natural probabilistic expressions, being given by the probability that two sites belong to the same cycle (see below for details). A similar representation was introduced by Aizenman and Nachtergaele for the antiferromagnetic model, who used it to study properties of spin chains [**1**]. The random objects are a bit different (loops instead of cycles), but it also has the advantage that spin correlations are given by the probability of belonging to the same loop.

Both Tóth's and Aizenman-Nachtergaele's representations involve a Poisson process on the edges of the graph. The measure is modified by a function of suitable geometric objects ("cycles" or "loops"). We first describe the two models in Section 3.1, and we relate them later to the Heisenberg models in Sections 3.3 and 3.4.

3.1. Poisson edge process, cycles and loops. Recall that $\Lambda = (\mathcal{V}, \mathcal{E})$ is a finite undirected graph. We attach to each edge a Poisson process on $[0, \beta]$ of unit intensity (see §7.2.1 for the definition of a Poisson point process). The Poisson processes for different edges are independent. A realization of this "Poisson edge process" is a finite sequence of pairs

$$\omega = \big((e_1, t_1), \ldots, (e_k, t_k)\big). \tag{3.1}$$

Each pair is called a bridge. The number of bridges across each edge is Poisson distributed with mean $|\mathcal{E}|$ and the total number of bridges is Poisson with mean $\beta|\mathcal{E}|$. Conditional on there being k bridges, the times are uniformly distributed in $\{0 < t_1 < t_2 < \ldots < t_k < \beta\}$ and the edges are chosen uniformly from $\mathcal{E}$. The corresponding measure is denoted $\rho_{\Lambda,\beta}(\mathrm{d}\omega)$.

To each realization ω there is a configuration of cycles and configuration of loops. The mathematical definitions are a bit cumbersome but the geometric ideas are simpler and more elegant. The reader is encouraged to look at Figure 1 for an illustration.

We consider the cylinder $\mathcal{V} \times [0, \beta]_{\mathrm{per}}$. A *cycle* is a closed trajectory on this space; that is, it is a function $\gamma : [0, L] \to \mathcal{V} \times [0, \beta]_{\mathrm{per}}$ such that, with $\gamma(\tau) = (x(\tau), t(\tau))$:

- $\gamma(\tau)$ is piecewise continuous; if it is continuous in interval $I \subset [0, L]$, then $x(\tau)$ is constant and $\frac{\mathrm{d}}{\mathrm{d}\tau} t(\tau) = 1$ in I.
- $\gamma(\tau)$ is discontinuous at τ iff the pair (e, t) belongs to ω, where $t = t(\tau)$ and e is the edge $\{\gamma(\tau-), \gamma(\tau+)\}$.

We choose L to be the smallest positive number such that $\gamma(L) = \gamma(0)$. Then L is the length of the cycle; it corresponds to the sum of the vertical legs in Figure 1 and it is a multiple of β. Let us make the cycles semi-continuous by assigning
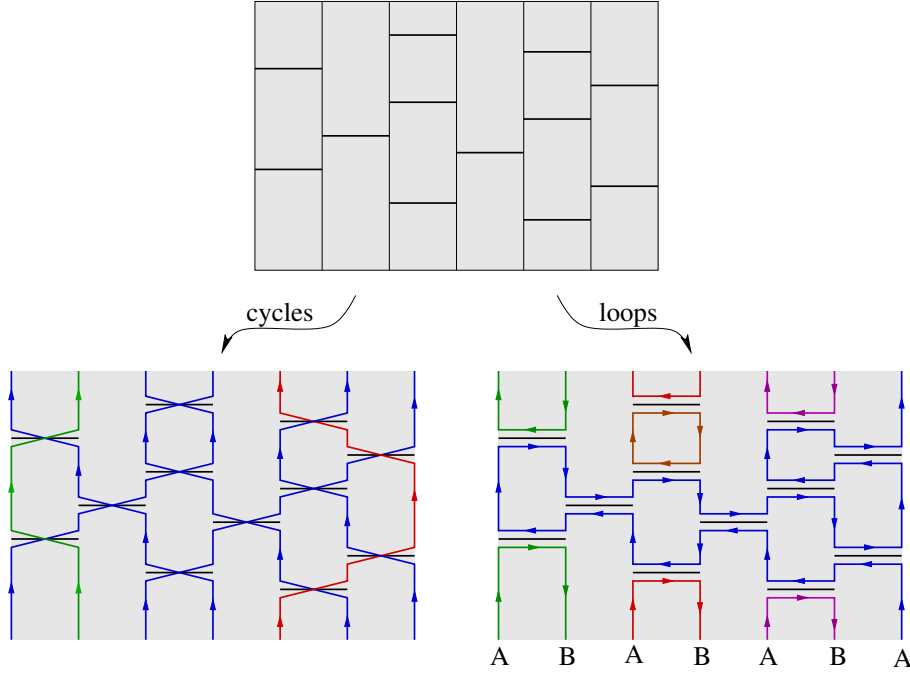
FIGURE 1. Top: an edge Poisson configuration ω on $\mathcal{V} \times [0, \beta]_{\mathrm{per}}$. Bottom left: its associated cycle configuration. Bottom right: its associated loop configuration. We see that $\mathcal{C}(\omega) = 3$ and $\mathcal{L}(\omega) = 5$.

the value $\gamma(\tau) = \gamma(\tau-)$ at the points of discontinuity. We identify cycles whose support is identical. Then to each ω corresponds a configuration of cycles $\mathcal{C}(\omega)$ whose supports form a partition of the cylinder $\mathcal{V} \times [0, \beta]_{\mathrm{per}}$. The number of cycles is $|\mathcal{C}(\omega)|$.

Loops are similar, but we now suppose that the graph is bipartite. The orientation is reversed on the B sublattice. We still consider the cylinder $\mathcal{V} \times [0, \beta]_{\mathrm{per}}$. A *loop* is a closed trajectory on this space; that is, it is a function $\gamma : [0, L] \rightarrow \mathcal{V} \times [0, \beta]_{\mathrm{per}}$ such that, with $\gamma(\tau) = (x(\tau), t(\tau))$:

- $\gamma(\tau)$ is piecewise continuous; if it is continuous in interval $I \subset [0, L]$, then $x(\tau)$ is constant and, in I,

$$\frac{\mathrm{d}}{\mathrm{d}\tau} t(\tau) = \begin{cases} 1 & \text{if } x(\tau) \text{ belongs to the A sublattice,} \\ -1 & \text{if } x(\tau) \text{ belongs to the B sublattice.} \end{cases} \tag{3.2}$$

- $\gamma(\tau)$ is discontinuous at τ iff the pair (e, t) belongs to ω, where $t = t(\tau)$ and e is the edge $\{\gamma(\tau-), \gamma(\tau+)\}$.

We choose L to be the smallest positive number such that $\gamma(L) = \gamma(0)$. Then L is the length of the loop; it corresponds to the sum of the vertical legs in Figure 1 (as for cycles), but it is not a multiple of β in general (contrary to cycles). We also make the loops semi-continuous by assigning the value $\gamma(\tau) = \gamma(\tau-)$ at the points of discontinuity. Identifying loops whose support is identical, to each ω corresponds a configuration of loops $\mathcal{L}(\omega)$ whose supports form a partition of the cylinder $\mathcal{V} \times [0, \beta]_{\mathrm{per}}$. The number of loops is $|\mathcal{L}(\omega)|$.

As we'll see, the relevant probability measures for the Heisenberg models (with $h = 0$) are proportional to $2^{|\mathcal{C}(\omega)|} \rho_{\mathcal{E}, \frac{\beta}{2}}(\mathrm{d}\omega)$ and $2^{|\mathcal{L}(\omega)|} \rho_{\mathcal{E}, \beta}(\mathrm{d}\omega)$.

3.2. Duhamel expansion. We first state and prove Duhamel's formula. It is a variant of the Trotter product formula that is usually used to derive stochastic representations.

PROPOSITION 3.1. *Let A, B be $n \times n$ matrices. Then*

$$
\mathrm{e}^{A+B} = \mathrm{e}^A + \int_0^1 \mathrm{e}^{tA} \, B \, \mathrm{e}^{(1-t)(A+B)} \, \mathrm{d}t
$$

$$
= \sum_{k \geq 0} \int_{0 < t_1 < \cdots < t_k < 1} \mathrm{d}t_1 \ldots \mathrm{d}t_k \, \mathrm{e}^{t_1 A} \, B \, \mathrm{e}^{(t_2 - t_1)A} \, B \ldots B \, \mathrm{e}^{(1-t_k)A} \, .
$$

PROOF. Let $F(s)$ be the matrix-valued function

$$
F(s) = \mathrm{e}^{sA} + \int_0^s \mathrm{e}^{tA} \, B \, \mathrm{e}^{(s-t)(A+B)} \, \mathrm{d}t. \tag{3.3}
$$

We show that, for all s,

$$
\mathrm{e}^{s(A+B)} = F(s). \tag{3.4}
$$

The derivative of $F(s)$ is

$$
F'(s) = \mathrm{e}^{sA} A + \mathrm{e}^{sA} B + \int_0^s \mathrm{e}^{tA} \, B \, \mathrm{e}^{(s-t)(A+B)} \, (A + B)\mathrm{d}t = F(s)(A + B). \tag{3.5}
$$

On the other hand, the derivative of $\mathrm{e}^{s(A+B)}$ is $\mathrm{e}^{s(A+B)} (A + B)$. The identity (3.4) clearly holds for $s = 0$, and both sides satisfy the same differential equation. They are then equal for all s.

We can iterate Duhamel's formula N times so as to get

$$
\mathrm{e}^{A+B} = \sum_{k=0}^N \int_{0 < t_1 < \cdots < t_k < 1} \mathrm{d}t_1 \ldots \mathrm{d}t_k \, \mathrm{e}^{t_1 A} \, B \, \mathrm{e}^{(t_2 - t_1)A} \, B \ldots B \, \mathrm{e}^{(1-t_k)A}
$$

$$
+ \int_{0 < t_1 < \cdots < t_N < 1} \mathrm{d}t_1 \ldots \mathrm{d}t_k \, \mathrm{e}^{t_1 A} \, B \, \mathrm{e}^{(t_2 - t_1)A} \, B \ldots B \Big[\mathrm{e}^{(1-t_N)(A+B)} - \mathrm{e}^{(1-t_N)A} \Big]. \tag{3.6}
$$

Using $\| \mathrm{e}^{tA} \| \leq \mathrm{e}^{t\|A\|}$, the last term is less than $2 \, \mathrm{e}^{\|A\|+\|B\|} \, \frac{\|B\|^N}{N!}$ and so it vanishes in the limit $N \to \infty$. The summand is less than $\mathrm{e}^{\|A\|} \frac{\|B\|^k}{k!}$ so that the sum is absolutely convergent. $\qquad\square$

Our goal is to perform a Duhamel's expansion of the Gibbs operator $\mathrm{e}^{-\beta H_{\Lambda, h}}$, where the Hamiltonian is given by a sum of terms indexed by the edges and by vertices. The following corollary applies to this case.

COROLLARY 3.2. *Let A and (h_e), $e \in \mathcal{E}$, be matrices in $\mathcal{H}^{(\mathcal{V})}$. Then*

$$
\mathrm{e}^{\beta(A+\sum_{e \in \mathcal{E}} h_e)} = \int \mathrm{d}\rho_{\mathcal{E}, \beta}(\omega) \, \mathrm{e}^{t_1 A} \, h_{e_1} \, \mathrm{e}^{(t_2 - t_1)A} \, h_{e_2} \ldots h_{e_k} \, \mathrm{e}^{(\beta - t_k)A} \, ,
$$

where $(t_1, e_1), \ldots, (t_k, e_k)$ are the bridges in ω.

PROOF. We can expand the right side by summing over the number of events k, integrate over $0 < t_1 < \cdots < t_k < \beta$ for the times of occurrence, and sum over edges $e_1, \ldots, e_k \in \mathcal{E}$. After the change of variables $t_i' = t_i/\beta$, we recognize the second line of Proposition 3.1. $\square$

3.3. Tóth's representation of the ferromagnet. It is convenient to introduce the operator $T_{x,y}$ that transposes the spins at x and y. In $\mathcal{H}_x \otimes \mathcal{H}_y$, the operator acts as follows:

$$T_{x,y}|a,b\rangle = |b,a\rangle, \qquad a,b = \pm\tfrac{1}{2}. \tag{3.7}$$

This rule extends to general vectors by linearity, and it extends to $\mathcal{H}^{(\mathcal{V})}$ by tensoring it with $\mathtt{Id}_{\mathcal{V}\backslash\{x,y\}}$. Using Lemma 2.1, it is not hard to check that

$$\vec{S}_x \cdot \vec{S}_y = \tfrac{1}{2}T_{x,y} - \tfrac{1}{4}\mathtt{Id}_{\{x,y\}}. \tag{3.8}$$

Recall that $\mathcal{C}(\omega)$ is the set of cycles of ω, and let $\gamma_x \in \mathcal{C}(\omega)$ denote the cycle that intersects $\{x\} \times \{0\}$ (which is henceforth abbreviated $x \times \{0\}$). Let $L(\gamma)$ denote the (vertical) length of the cycle γ; it is always a multiple of $\frac{\beta}{2}$ in the theorem below.

THEOREM 3.3 (Tóth's representation of the ferromagnet). *The partition function, the average magnetization, and the two-point correlation function have the following expressions.*

$$Z_\Lambda^{\mathrm{ferro}}(\beta,h) = \mathrm{e}^{-\frac{\beta}{4}|\mathcal{E}|} \int \mathrm{d}\rho_{\mathcal{E},\frac{\beta}{2}}(\omega) \prod_{\gamma \in \mathcal{C}(\omega)} \big(2\cosh(hL(\gamma))\big),$$

$$\mathrm{Tr}\, S_x^{(3)}\, \mathrm{e}^{-\beta H_{\Lambda,h}^{\mathrm{ferro}}} = \tfrac{1}{2}\,\mathrm{e}^{-\frac{\beta}{4}|\mathcal{E}|} \int \mathrm{d}\rho_{\mathcal{E},\frac{\beta}{2}}(\omega) \prod_{\gamma \in \mathcal{C}(\omega)} \big(2\cosh(hL(\gamma))\big)\tanh(hL(\gamma_x)),$$

$$\mathrm{Tr}\, S_x^{(3)}S_y^{(3)}\, \mathrm{e}^{-\beta H_{\Lambda,h}^{\mathrm{ferro}}} = \tfrac{1}{4}\,\mathrm{e}^{-\frac{\beta}{4}|\mathcal{E}|} \int \mathrm{d}\rho_{\mathcal{E},\frac{\beta}{2}}(\omega) \prod_{\gamma \in \mathcal{C}(\omega)} \big(2\cosh(hL(\gamma))\big)$$

$$\times \begin{cases} 1 & \text{if } \gamma_x = \gamma_y, \\ \tanh(hL(\gamma_x))\tanh(hL(\gamma_y)) & \text{if } \gamma_x \neq \gamma_y. \end{cases}$$
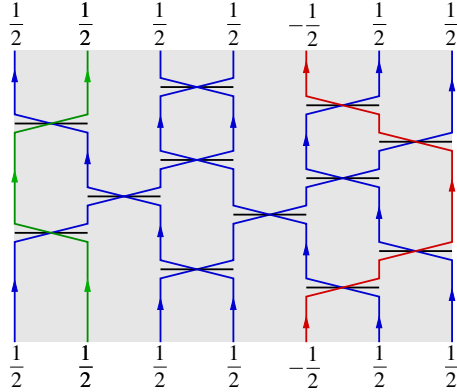

FIGURE 2. Each cycle is characterized by a given spin.

PROOF. The partition function can be expanded using Corollary 3.2 so as to get

$$Z_\Lambda^{\text{ferro}}(\beta, h) = e^{-\frac{\beta}{4}|\mathcal{E}|} \operatorname{Tr} e^{\frac{\beta}{2}(2hM_\Lambda + \sum_e T_e)}$$
$$= e^{-\frac{\beta}{4}|\mathcal{E}|} \int d\rho_{\mathcal{E}, \frac{\beta}{2}}(\omega) \sum_{s^{(\mathcal{V})}} \langle s^{(\mathcal{V})}| e^{2t_1 hM_\Lambda} T_{e_1} \dots T_{e_k} e^{2(\frac{\beta}{2} - t_k)hM_\Lambda} |s^{(\mathcal{V})}\rangle,$$

(3.9)

where $(e_1, t_1), \dots, (e_k, t_k)$ are the times and the edges of ω. Observe that the vectors $|s^{(\mathcal{V})}\rangle$ are eigenvectors of e^{tM_Λ}. It is not hard to see that the matrix element above is zero unless each cycle is characterized by a single spin value (see illustration in Figure 2). If the matrix element is not zero, then it is equal to

$$\langle s^{(\mathcal{V})}| e^{2t_1 hM_\Lambda} T_{e_1} \dots T_{e_k} e^{2(\frac{\beta}{2} - t_k)hM_\Lambda} |s^{(\mathcal{V})}\rangle = \prod_{\gamma \in \mathcal{C}(\omega)} e^{2hL(\gamma)s(\gamma)}$$

(3.10)

with $s(\gamma)$ the spin of the cycle γ. After summing over $s(\gamma) = \pm\frac{1}{2}$, each cycle contributes $e^{hL(\gamma)} + e^{-hL(\gamma)} = 2\cosh(hL(\gamma))$, and we obtain the expression for the partition function.

The expression that involves $S_x^{(3)}$ is similar, except that the cycle γ_x that contains $x \times \{0\}$ contributes $\frac{1}{2} e^{hL(\gamma_x)} - \frac{1}{2} e^{-hL(\gamma_x)} = \sinh(hL(\gamma_x))$. Since the factor $2\cosh(hL(\gamma_x))$ appears in the expression, it must be corrected by the hyperbolic tangent.

Finally, the expression that involves $S_x^{(3)} S_y^{(3)}$ has two terms, whether $x \times \{0\}$ and $y \times \{0\}$ find themselves in the same cycle or not. In the first case, we get $\frac{1}{2} \cosh(hL(\gamma_{xy}))$, but in the second case we get $\sinh(hL(\gamma_x)) \sinh(hL(\gamma_y))$, which eventually gives the hyperbolic tangents. □

It is convenient to rewrite a bit the cycle weights. Using $2\cosh(hL(\gamma)) = e^{hL(\gamma)}(1 + e^{-2hL(\gamma)})$ and $\sum_{\gamma \in \mathcal{C}(\omega)} L(\gamma) = \frac{\beta}{2}|\mathcal{V}|$, the relevant probability measure for the cycle representation can be written

$$\mathbb{P}_{\Lambda,\beta,h}^{\mathcal{C}}(d\omega) = Z_\Lambda^{\text{ferro}}(\beta, h)^{-1} e^{-\frac{\beta}{4}|\mathcal{E}| + \frac{\beta}{2}h|\mathcal{V}|} d\rho_{\mathcal{E}, \frac{\beta}{2}}(d\omega) \prod_{\gamma \in \mathcal{C}(\omega)} (1 + e^{-2hL(\gamma)})$$

(3.11)

This form makes it easier to see the effect of the external field $h \geq 0$. Notice that the product over cycles simplifies to $2^{|\mathcal{C}(\omega)|}$ when the external field strength vanishes (i.e. $h = 0$). We write $\mathbb{E}_{\Lambda,\beta,h}$ for the expectation with respect to $\mathbb{P}_{\Lambda,\beta,h}^{\mathcal{C}}$. Then, in terms of the cycle model, the expectation of the spin operators and correlations are given by

$$\langle S_x^{(3)} \rangle_{\Lambda,\beta,h} = \frac{1}{2} \mathbb{E}_{\Lambda,\beta,h}^{\mathcal{C}} (\tanh(hL(\gamma_x)))$$

(3.12)

and

$$\langle S_x^{(3)} S_y^{(3)} \rangle_{\Lambda,\beta,h} = \frac{1}{4} \mathbb{P}_{\Lambda,\beta,h}^{\mathcal{C}}(\gamma_x = \gamma_y) +$$
$$\frac{1}{4} \mathbb{E}_{\Lambda,\beta,h}^{\mathcal{C}} [1_{\gamma_x \neq \gamma_y} \tanh(hL(\gamma_x)) \tanh(hL(\gamma_y))].$$

(3.13)

In the case $h = 0$, we see that $\langle S_x^{(3)} \rangle_{\Lambda,\beta,0} = 0$, as already noted from the spin flip symmetry, and

$$\langle S_x^{(3)} S_y^{(3)} \rangle_{\Lambda,\beta,0} = \frac{1}{4} \mathbb{P}_{\Lambda,\beta,h}^{\mathcal{C}}(\gamma_x = \gamma_y).$$

(3.14)

That is, the spin-spin correlation of two sites x and y is proportional to the probability that the sites lie in the same cycle.

3.4. Aizenman-Nachtergaele's representation of the antiferromagnet.
The antiferromagnetic model only differs from the ferromagnetic model by a sign, but it leads to deep changes. As the transposition operator now carries a negative sign in the Hamiltonian, a possibility is to turn the measure corresponding to (3.11) into a signed measure, with an extra factor $(-1)^k$ where $k = k(\omega)$ is the number of transpositions. That means descending from the heights of probability theory down to... well, to measure theory. This fate can fortunately be avoided thanks to the observations of Aizenman and Nachtergaele [1].

Their representation is restricted to bipartite graphs. A graph is *bipartite* if the set of vertices $\mathcal{V}$ can be partitioned into two sets $\mathcal{V}_A$ and $\mathcal{V}_B$ such that edges only connect the A set to the B set:

$$\{x, y\} \in \mathcal{E} \implies (x, y) \in \mathcal{V}_A \times \mathcal{V}_B \text{ or } (x, y) \in \mathcal{V}_B \times \mathcal{V}_A. \tag{3.15}$$

This class contains many relevant cases, such as finite boxes in $\mathbb{Z}^d$; periodic boundary conditions are allowed provided the side lengths are even. In the following, we use the notation

$$(-1)^x = \begin{cases} 1 & \text{if } x \in \mathcal{V}_A, \\ -1 & \text{if } x \in \mathcal{V}_B. \end{cases} \tag{3.16}$$

Instead of the transposition operator, we consider the projection operator $P_{xy}^{(0)}$ onto the singlet state described in Lemma 2.1. Its action on the basis is

$$P_{xy}^{(0)}|a, a\rangle = 0, \qquad P_{xy}^{(0)}|a, -a\rangle = \tfrac{1}{2}|a, -a\rangle - \tfrac{1}{2}|-a, a\rangle, \tag{3.17}$$

for all $a = \pm\frac{1}{2}$. Further, it follows from Lemma 2.1 that

$$\vec{S}_x \cdot \vec{S}_y = \mathrm{Id}_{\{x,y\}} - \mathbb{P}_{xy}^{(0)}. \tag{3.18}$$

Recall that $\mathcal{L}(\omega)$ is the set of loops of ω. Let γ_x denote the loop that contains $x \times \{0\}$. We do not need notation for the loops that do not intersect the $t = 0$ plane. Also, it is not the *lengths* of the loops which are important but their winding number $w(\gamma)$.

THEOREM 3.4 (Aizenman-Nachtergaele's representation of the antiferromagnet). *Assume that Λ is a bipartite graph. The partition function, the average magnetization, and the two-point correlation function have the following expressions.*

$$Z_\Lambda^{\mathrm{anti}}(\beta, h) = \mathrm{e}^{-\beta|\mathcal{E}|} \int \mathrm{d}\rho_{\mathcal{E}, \frac{\beta}{2}}(\omega) \prod_{\gamma \in \mathcal{L}(\omega)} \big(2\cosh(\tfrac{1}{2}\beta h w(\gamma))\big),$$

$$\mathrm{Tr}\, S_x^{(3)}\, \mathrm{e}^{-\beta H_{\Lambda,h}^{\mathrm{anti}}} = \tfrac{1}{2}(-1)^x\, \mathrm{e}^{-\beta|\mathcal{E}|} \int \mathrm{d}\rho_{\mathcal{E}, \frac{\beta}{2}}(\omega)$$

$$\prod_{\gamma \in \mathcal{L}(\omega)} \big(2\cosh(\tfrac{1}{2}\beta h w(\gamma))\big) \tanh(\tfrac{1}{2}\beta h w(\gamma_x)),$$

$$\mathrm{Tr}\, S_x^{(3)} S_y^{(3)}\, \mathrm{e}^{-\beta H_{\Lambda,h}^{\mathrm{anti}}} = \tfrac{1}{4}(-1)^x(-1)^y\, \mathrm{e}^{-\beta|\mathcal{E}|} \int \mathrm{d}\rho_{\mathcal{E}, \frac{\beta}{2}}(\omega)$$

$$\prod_{\gamma \in \mathcal{L}(\omega)} \big(2\cosh(\tfrac{1}{2}\beta h w(\gamma))\big) \begin{cases} 1 & \text{if } \gamma_x = \gamma_y, \\ \tanh(\tfrac{1}{2}\beta h w(\gamma_x)) \tanh(\tfrac{1}{2}\beta h w(\gamma_y)) & \text{if } \gamma_x \neq \gamma_y. \end{cases}$$

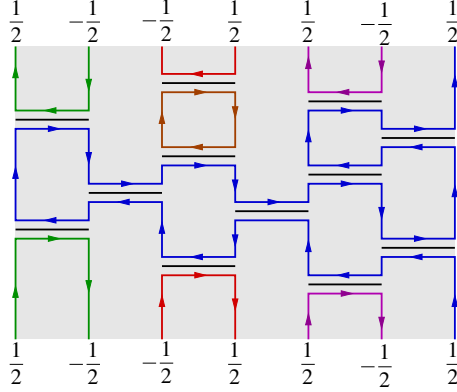When $h = 0$, we get the simpler factor $2^{|\mathcal{L}(\omega)|}$.

FIGURE 3. Each loop is characterized by a given spin, but the value alternate according to whether the site belongs to the A or B sublattice.

PROOF. As before, we expand the partition function using Corollary 3.2 and we get

$$Z_\Lambda^{\text{anti}}(\beta, h) = e^{-\beta|\mathcal{E}|} \operatorname{Tr} e^{\frac{\beta}{2}(2hM_\Lambda + \sum_e 2P_e^{(0)})}$$

$$= e^{-\beta|\mathcal{E}|} \int d\rho_{\mathcal{E}, \frac{\beta}{2}}(\omega) \sum_{s^{(\mathcal{V})}} \langle s^{(\mathcal{V})}| e^{2t_1 hM_\Lambda} 2P_{e_1}^{(0)} \dots 2P_{e_k}^{(0)} e^{2(\frac{\beta}{2} - t_k)hM_\Lambda} |s^{(\mathcal{V})}\rangle,$$

(3.19)

where $(e_1, t_1), \dots, (e_k, t_k)$ are the times and the edges of ω. Notice that

$$e^{tM_\Lambda} |s^{(\mathcal{V})}\rangle = e^{t\langle s^{(\mathcal{V})}|M_\Lambda|s^{(\mathcal{V})}\rangle}.$$

(3.20)

In Dirac's notation, the resolution of the identity is

$$\operatorname{Id}_\mathcal{V} = \sum_{s^{(\mathcal{V})} \in \{-\frac{1}{2}, \frac{1}{2}\}^\mathcal{V}} |s^{(\mathcal{V})}\rangle\langle s^{(\mathcal{V})}|.$$

(3.21)

We insert it at the right of each operator $P_e^{(0)}$ and we obtain

$$Z_\Lambda^{\text{anti}}(\beta, h) = e^{-\beta|\mathcal{E}|} \int d\rho_{\mathcal{E}, \frac{\beta}{2}}(\omega) \sum_{s_1^{(\mathcal{V})}, \dots, s_k^{(\mathcal{V})}} e^{2t_1 h\langle s_1^{(\mathcal{V})}|M_\Lambda|s_1^{(\mathcal{V})}\rangle} \langle s_1^{(\mathcal{V})}|2P_{e_1}^{(0)}|s_2^{(\mathcal{V})}\rangle$$

$$e^{2(t_2 - t_1)h\langle s_2^{(\mathcal{V})}|M_\Lambda|s_2^{(\mathcal{V})}\rangle} \langle s_2^{(\mathcal{V})}|2P_{e_2}^{(0)}|s_3^{(\mathcal{V})}\rangle \dots \langle s_k^{(\mathcal{V})}|2P_{e_k}^{(0)}|s_1^{(\mathcal{V})}\rangle e^{2(\beta - t_k)h\langle s_1^{(\mathcal{V})}|M_\Lambda|s_1^{(\mathcal{V})}\rangle}.$$

(3.22)

Let us see that this long expression can be conveniently expressed in the language of loops. We can interpret ω and $s_1^{(\mathcal{V})}, \dots, s_k^{(\mathcal{V})}$ as a spin configuration $\boldsymbol{s}$ in $\mathcal{V} \times [0, \beta]_{\text{per}}$. It is constant in time except possibly at (e_i, t_i). By (3.17), the product

$$\langle s_1^{(\mathcal{V})}|2P_{e_1}^{(0)}|s_2^{(\mathcal{V})}\rangle \dots \langle s_k^{(\mathcal{V})}|2P_{e_k}^{(0)}|s_1^{(\mathcal{V})}\rangle$$

differs from 0 iff the value of $(-1)^x \boldsymbol{s}(x, t)$ is constant on each loop (see illustration in Figure 3). In this case, its value is ± 1, as each bridge contributes $+1$ if the spins are constant, and -1 if they flip. Let us actually check that it is always $+1$. If the

bridge separates two loops with spins a and b, the factor is

$$(-1)^{a-b} = e^{i\frac{\pi}{2}a} \, e^{-i\frac{\pi}{2}b}. \tag{3.23}$$

Looking at the loop γ with spin a, there is a factor $e^{i\frac{\pi}{2}a}$ for each jump A→B (of the form $\sqcap$) and a factor $e^{-i\frac{\pi}{2}a}$ for each jump B→A (of the form $\sqcup$). Since there is an identical number of both types of jumps, these factors precisely cancel.

The product

$$e^{2t_1 h \langle s_1^{(\mathcal{V})}|M_\Lambda|s_1^{(\mathcal{V})}\rangle} \, e^{2(t_2-t_1)h\langle s_2^{(\mathcal{V})}|M_\Lambda|s_2^{(\mathcal{V})}\rangle} \, \cdots \, e^{2(\beta - t_k h \langle s_1^{(\mathcal{V})}|M_\Lambda|s_1^{(\mathcal{V})}\rangle}$$

also factorizes according to loops. The contribution of a loop γ with spin a is $e^{2hL_A(\gamma)a - 2hL_B(\gamma)a}$, where L_A, L_B are the vertical lengths of γ on the A and B sublattices. We have

$$L_A(\gamma) - L_B(\gamma) = \tfrac{\beta}{2}w(\gamma). \tag{3.24}$$

The contribution is therefore $e^{\beta h w(\gamma)a}$. Summing over $a = \pm\frac{1}{2}$, we get the hyperbolic cosine of the expression for the partition function of Theorem 3.4.

The expression that involves $S_x^{(3)}$ is similar; the only difference is that the loop that contains $x \times \{0\}$ contributes $(-1)^x \sinh(\frac{1}{2}\beta h w(\gamma))$ instead of $2\cosh(\frac{1}{2}\beta h w(\gamma))$, hence the hyperbolic tangent. Finally, the expression that involves $S_x^{(3)} S_y^{(3)}$ is similar but we need to treat separately the cases where $x \times \{0\}$ and $y \times \{0\}$ belong or do not belong to the same loop. □

4. Thermodynamic limit and phase transitions

Phase transitions are cooperative phenomena where a small change of the external parameters results in drastic modifications of the properties of the system. There was some confusion in the early days of statistical mechanics as to whether the formalism contained the possibility of describing phase transitions, as all finite volume quantities are smooth. It was eventually realized that the proper formalism involves a *thermodynamic limit* where the system size tends to infinity, in such a way that the local behavior remains largely unaffected. The proofs of the existence of thermodynamic limits were fundamental contributions to the mathematical theory of phase transitions, and they were pioneered by Fisher and Ruelle in the 1960's, see [40] for more references.

We show that the free energy converges in the thermodynamic limit along a sequence of boxes in $\mathbb{Z}^d$ of increasing size (Section 4.1). We discuss various characterizations of ferromagnetic phase transitions in Section 4.2, and magnetic long-range order in Section 4.3. In Section 4.4 we consider the relations between the magnetisation in the quantum models and the lengths of the cycles and loops.

4.1. Thermodynamic limit. Despite our professed intention to treat arbitrary graphs, we now restrict ourselves to a very specific case, namely that of a sequence of cubes whose side lengths tends to infinity. Since $F_\Lambda(\beta, h)$ scales like the volume of the system, we define the *mean free energy* f_Λ to be

$$f_\Lambda(\beta, h) = \frac{1}{|\mathcal{V}|} F_\Lambda(\beta, h). \tag{4.1}$$

We consider the sequence of graphs $\Lambda_n = (\mathcal{V}_n, \mathcal{E}_n)$ where $\mathcal{V}_n = \{1, \ldots, n\}^d$ and $\mathcal{E}_n$ is the set of nearest-neighbors, i.e., $\{x, y\} \in \mathcal{E}_n$ iff $\|x - y\| = 1$.

THEOREM 4.1 (Thermodynamic limit of the free energy). *The sequence $(f_{\Lambda_n}(\beta, h))_{n \geq 1}$ converges pointwise to a function $f(\beta, h)$, uniformly on compact sets.*
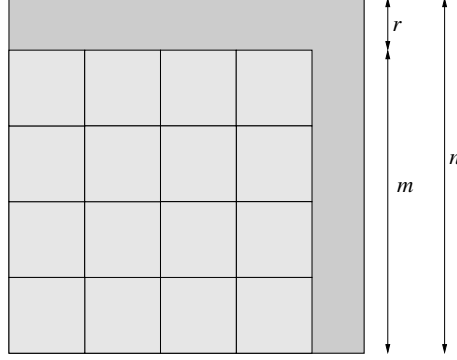

FIGURE 4. The large box of size n is decomposed in k^d boxes of size m; there are no more than drn^{d-1} remaining sites.

PROOF. We consider the ferromagnetic model, but the modifications for the antiferromagnetic models are straightforward. We use a subadditive argument. Notice the inequality $\text{Tr } e^{A+B} \geq \text{Tr } e^A$ that holds for all self-adjoint operators A, B with $B \geq 0$. (It follows e.g. from the minimax principle, or from Klein's inequality.) We also rewrite the Hamiltonian so as to have only positive definite terms. Namely, let

$$h_{x,y} = -\vec{S}_x \cdot \vec{S}_y + \tfrac{1}{4}\text{Id}. \tag{4.2}$$

Then

$$Z_\Lambda(\beta, h) = e^{-\frac{\beta}{4}|\mathcal{E}|} \text{Tr} \exp\Big(\beta \sum_{\{x,y\} \in \mathcal{E}} h_{x,y} + \beta h \sum_{x \in \mathcal{V}} S_x^{(3)}\Big). \tag{4.3}$$

Let m, n, k, r be integers such that $n = km + r$ and $0 \leq r < m$. The box $\mathcal{V}_n$ is the disjoint union of k^d boxes of size m, and of some remaining sites (less than $n^{d-1}r$), see Figure 4 for an illustration. We get an inequality for the partition function in Λ_n by dismissing all $h_{x,y}$ where $\{x, y\}$ are not inside a single box of size m. The boxes $\mathcal{V}_m$ become independent, and

$$Z_{\Lambda_n}(\beta, h) \geq e^{-\frac{\beta}{4}|\mathcal{E}_n|} \Big[\text{Tr}_{\mathcal{H}(\mathcal{V}_m)} \exp\Big(\beta \sum_{\{x,y\} \in \mathcal{E}_m} h_{x,y} + \beta h \sum_{x \in \mathcal{V}_m} S_x^{(3)}\Big)\Big]^{k^d}$$

$$= [Z_{\Lambda_m}(\beta, h)]^{k^d} e^{-\frac{\beta}{4}|\mathcal{E}_n|} e^{k^d \frac{\beta}{4}|\mathcal{E}_m|}. \tag{4.4}$$

We have neglected the contribution of $e^{\beta h S_x^{(3)}}$ for x outside the boxes $\mathcal{V}_m$, which is possible because their traces are greater than 1. It is not hard to check that

$$|\mathcal{E}_n| \leq k^d|\mathcal{E}_m| + k^d dm^{d-1} + dn^{d-1}r. \tag{4.5}$$

We then obtain a subbaditive relation for the free energy, up to error terms that will soon disappear:

$$f_{\Lambda_n}(\beta, h) \leq \frac{(km)^d}{n^d} f_{\Lambda_m}(\beta, h) + \frac{k^d dm^{d-1}}{n^d} + \frac{dr}{n}. \tag{4.6}$$

Then

$$\limsup_{n\to\infty} f_{\Lambda_n}(\beta, h) \leq f_{\Lambda_m}(\beta, h) + \frac{d}{m}. \tag{4.7}$$

Taking the lim inf over m in the right side, we see that it is larger or equal to the lim sup, and the limit necessarily exists.

Uniform convergence on compact intervals follows from Lemma 2.4 (which implies that (f_{Λ_n}) is equicontinuous) and the Arzelà-Ascoli theorem (see e.g. Theorem 4.4 in Folland [**19**]). $\square$

COROLLARY 4.2 (Thermodynamic limit with periodic boundary conditions). *Let $(\Lambda_n^{\mathrm{per}})$ be the sequence of cubes in $\mathbb{Z}^d$ of size n with periodic boundary conditions and nearest-neighbor edges. Then $(f_{\Lambda_n^{\mathrm{per}}}(\beta, h))_{n\geq 1}$ converges pointwise to the same function $f(\beta, h)$ as in Theorem 4.1, uniformly on compact sets.*

The proof follows from $|f_{\Lambda_n^{\mathrm{per}}}(\beta, h) - f_{\Lambda_n}(\beta, h)| \leq \frac{3d}{4n}$, which is not too hard to prove, and Theorem 4.1.

4.2. Ferromagnetic phase transition. In statistical physics, an *order parameter* is a quantity that allows one to identify a phase, typically because it vanishing in all phases except one. We consider three different order parameters, the first two will be shown to be equivalent, and the last one to be smaller than the first two.

- **Thermodynamic magnetization**. This is equal to (the negative of) the right-derivative of $f(\beta, h)$ with respect to h. We're looking for a jump in the derivative, which is referred to as a *first-order* phase transition. Let $m_{\mathrm{th}}^*(\beta)$ denote the corresponding order parameter, which, because f is concave, is equal to

$$m_{\mathrm{th}}^*(\beta) = -\sup_{h>0} \frac{f(\beta, h) - f(\beta, 0)}{h}. \tag{4.8}$$

- **Residual magnetization**. Imagine placing the ferromagnet in an external magnetic field, so that it becomes magnetized. Now gradually turn off the external field. Does the system still display global magnetization? Mathematically, the relevant order parameter is

$$m_{\mathrm{res}}^*(\beta) = \lim_{h\to 0+} \liminf_{n\to\infty} \frac{1}{n^d} \langle M_{\Lambda_n} \rangle_{\Lambda_n, \beta, h}. \tag{4.9}$$

  (We see below that the lim inf can be replaced by the lim sup without affecting m_{res}^*.)

- **Spontaneous magnetization** at $h = 0$. Since $\langle M_\Lambda \rangle = 0$ (because of the spin flip symmetry), we rather consider

$$m_{\mathrm{sp}}^*(\beta) = \liminf_{n\to\infty} \frac{1}{n^d} \langle |M_{\Lambda_n}| \rangle_{\Lambda_n, \beta, 0}. \tag{4.10}$$

Here, $|M_\Lambda|$ denotes the absolute value of the matrix M_Λ.

A more handy quantity, however, is the expectation of M_Λ^2, which can be expressed in term of the two-point correlation function, see below. It is equivalent to m_{sp}^*:

LEMMA 4.3.

$$\left\langle \frac{|M_\Lambda|}{|\mathcal{V}|} \right\rangle_{\Lambda, \beta, 0}^2 \leq \left\langle \left(\frac{M_\Lambda}{|\mathcal{V}|} \right)^2 \right\rangle_{\Lambda, \beta, 0} \leq \frac{1}{2} \left\langle \frac{|M_\Lambda|}{|\mathcal{V}|} \right\rangle_{\Lambda, \beta, 0}.$$

PROOF. For the first inequality, use $|M_\Lambda| = |M_\Lambda|\mathrm{Id}$ and use the Cauchy-Schwarz inequality (2.24). For the second inequality, observe that $|M_\Lambda| \leq \frac{1}{2}|\mathcal{V}|\mathrm{Id}$ implies that $M_\Lambda^2 \leq \frac{1}{2}|\mathcal{V}||M_\Lambda|$, and use the fact that the Gibbs state is a positive linear functional. □
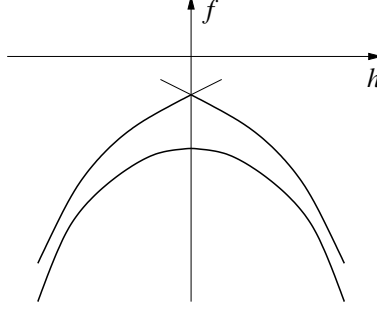

FIGURE 5. Qualitative graphs of the free energy $f(\beta, h)$ as a function of h, for β large (top) and β small (bottom).

The three order parameters above are related as follows:

PROPOSITION 4.4.
$$m_{\mathrm{th}}^*(\beta) = m_{\mathrm{res}}^*(\beta) \geq \tfrac{1}{2}m_{\mathrm{sp}}^*(\beta).$$

PROOF OF $m_{\mathrm{th}}^* = m_{\mathrm{res}}^*$. We prove that whenever f_n is a sequence of differentiable concave functions that converge pointwise to the (necessarily concave) function f, then

$$D_+f(0) = \lim_{h \to 0+} \limsup_{n \to \infty} f_n'(h) = \lim_{h \to 0+} \liminf_{n \to \infty} f_n'(h). \tag{4.11}$$

Up to the signs, the left side is equal to m_{th}^* and the right side to m_{res}^* and we obtain the identity in Proposition 4.4. The proof of (4.11) follows from the general properties

$$\limsup_i\big(\inf_j a_{ij}\big) \leq \inf_j\big(\limsup_i a_{ij}\big),$$
$$\liminf_i\big(\sup_j a_{ij}\big) \geq \sup_j\big(\liminf_i a_{ij}\big), \tag{4.12}$$

and from the following expressions for left- and right-derivatives of concave functions:

$$D_-f(h) = \inf_{s>0} \frac{f(h) - f(h-s)}{s}, \qquad D_+f(h) = \sup_{s>0} \frac{f(h+s) - f(h)}{s}. \tag{4.13}$$

With these observations, the proof is straightforward. For $h > 0$,

$$D_+f(0) \geq D_-f(h) = \inf_{s>0} \limsup_{n \to \infty} \frac{f_n(h) - f_n(h-s)}{s} \geq \limsup_{n \to \infty} f_n'(h)$$
$$\geq \liminf_{n \to \infty} f_n'(h) \geq \sup_{s>0} \liminf_{n \to \infty} \frac{f_n(h+s) - f_n(h)}{s} = D_+f(h). \tag{4.14}$$

Since right-derivatives are right-continuous, the last term converges to $D_+f(0)$ as $h \to 0+$. This proves Eq. (4.11). □

PROOF OF $m_{\mathrm{res}}^* \geq \frac{1}{2}m_{\mathrm{sp}}^*$. Let $h > 0$, and let $\{\varphi_j\}$ be an orthonormal set of eigenvectors of $H_{\Lambda_n,0}$ and M_{Λ_n} with eigenvalues e_j and m_j, respectively. Because of the spin flip symmetry, we have

$$
\begin{aligned}
\langle M_{\Lambda_n} \rangle_{\Lambda_n,\beta,h} &= \frac{\sum_{j:m_j>0} m_j \, \mathrm{e}^{-\beta e_j} \left(\mathrm{e}^{\beta h m_j} - \mathrm{e}^{-\beta h m_j} \right)}{\sum_{j:m_j>0} \mathrm{e}^{-\beta e_j} \left(\mathrm{e}^{\beta h m_j} + \mathrm{e}^{-\beta h m_j} \right) + \sum_{j:m_j=0} \mathrm{e}^{-\beta e_j}} \\
&\geq \frac{\sum_{j:m_j>0} m_j \, \mathrm{e}^{-\beta e_j + \beta h m_j} \left(1 - \mathrm{e}^{-2\beta h m_j} \right)}{2 \sum_{j:m_j>0} \mathrm{e}^{-\beta e_j + \beta h m_j} + \sum_{j:m_j=0} \mathrm{e}^{-\beta e_j}}.
\end{aligned}
\tag{4.15}
$$

After division by n^d, we only need to consider those j with $m_j \sim n^d$, in which case $\mathrm{e}^{-2\beta h m_j} \approx 0$. We can therefore replace the bracket by 1 under the limit $n \to \infty$. On the other hand, consider the function $G(h) = \frac{1}{\beta} \log \mathrm{Tr} \, \mathrm{e}^{-\beta H_{\Lambda_n,0} + \beta h |M_{\Lambda_n}|}$. One can check that it is convex in h, so $G'(h) \geq G'(0)$. Its derivative can be expanded as above, so that

$$
G'(h) = \frac{\sum_j |m_j| \, \mathrm{e}^{-\beta e_j + \beta h |m_j|}}{\sum_j \mathrm{e}^{-\beta e_j + \beta h |m_j|}}.
\tag{4.16}
$$

It is identical to twice the second line of (4.15) (without the bracket). Then

$$
m_{\mathrm{res}}^*(\beta) \geq \frac{1}{2} \lim_{h \to 0+} \liminf_{n \to \infty} \frac{1}{n^d} G'(h) \geq \frac{1}{2} \liminf_{n \to \infty} \frac{1}{n^d} G'(0) = \frac{1}{2} m_{\mathrm{sp}}^*(\beta).
\tag{4.17}
$$

$\square$

4.3. Antiferromagnetic phase transition.

While ferromagnets favor alignment of the spins, antiferromagnets favor staggered phases, where spins are aligned on one sublattice and aligned in the opposite direction on the other sublattice. The external magnetic field does not play much of a rôle. One could mirror the ferromagnetic situation by introducing a non-physical staggered magnetic field of the kind $h \sum_{x \in \mathcal{V}} (-1)^x S_x^{(3)}$, which would lead to counterparts of the order parameters m_{th}^* and m_{res}^*. We content ourselves with turning off the external magnetic field, i.e. we set $h = 0$, and with looking at magnetic long-range order. For $x,y \in \mathcal{V}$, we introduce the correlation function

$$
\sigma_{\Lambda,\beta}(x,y) = (-1)^x (-1)^y \langle S_x^{(3)} S_y^{(3)} \rangle_{\Lambda,\beta,0}.
\tag{4.18}
$$

One question is whether

$$
\bar{\sigma}^2 = \liminf_{n \to \infty} \frac{1}{|\mathcal{V}_n|^2} \sum_{x,y \in \mathcal{V}_n} \sigma_{\Lambda,\beta}(x,y)
\tag{4.19}
$$

differs from 0. A related question is whether the correlation function does not decay to 0 as the distance between x and y tends to infinity. One says that the system exhibits *long-range order* if this happens.

In $\mathbb{Z}^d$ and for β large enough, it is well-known that there is no long-range order and that the correlation function decays exponentially with $\|x - y\|$. Long-range order is expected in dimension $d \geq 3$ but not in $d = 1, 2$. This is discussed in more details in Section 5.

4.4. Phase transitions in cycle and loop models.

We clarify in this section the relations between the order parameters of the quantum systems and the nature of cycles and loops. This gives probabilistic interpretation to the quantum results. We introduce two quantities, that apply simultaneously to cycles and loops.

- The fraction of vertices in *infinite* cycles/loops:

$$\eta_\infty(\beta, h) = \lim_{K \to \infty} \liminf_{n \to \infty} \frac{1}{n^d} \mathbb{E}_{\Lambda_n, \beta, h}\big(\#\{x \in \mathcal{V}_n : L(\gamma_x) > K\}\big). \qquad (4.20)$$

- The fraction of vertices in *macroscopic* cycles/loops:

$$\eta_{\mathrm{macro}}(\beta, h) = \lim_{\varepsilon \to 0+} \liminf_{n \to \infty} \frac{1}{n^d} \mathbb{E}_{\Lambda_n, \beta, h}\big(\#\{x \in \mathcal{V}_n : L(\gamma_x) > \varepsilon n^d\}\big). \qquad (4.21)$$

It is clear that $\eta_\infty(\beta, h) \geq \eta_{\mathrm{macro}}(\beta, h)$. These two quantities relate to magnetization and long-range order as follows. The first two statements deal with cycles, the third statement for loops.

PROPOSITION 4.5.

(a) $m^*_{\mathrm{res}}(\beta) \geq \frac{1}{2} \lim_{h \to 0+} \eta_\infty(\beta, h)$.
(b) $m^*_{\mathrm{sp}}(\beta) > 0 \iff \eta_{\mathrm{macro}}(\beta, 0) > 0$.
(c) $\overline{\sigma}(\beta) > 0 \iff \eta_{\mathrm{macro}}(\beta, 0) > 0$.

PROOF. Let

$$m(\beta, h) = \liminf_{n \to \infty} \langle S_0^{(3)} \rangle_{\Lambda_n, \beta, h}. \qquad (4.22)$$

We use $\tanh x \geq \tanh K \cdot 1_{x > K}$, which holds for any K, and Theorem 3.3, so as to get

$$m(\beta, h) \geq \frac{1}{2} \tanh(\beta h K) \liminf_{n \to \infty} \mathbb{P}_{\Lambda_n, \beta, h}(L(\gamma_0) > K). \qquad (4.23)$$

Taking $K \to \infty$, we get $m(\beta, h) \geq \frac{1}{2} \eta_\infty(\beta, h)$. We now take $h \to 0+$ to obtain (a).

For (b), we observe that, since the vertices of Λ_n are exchangeable,

$$\frac{1}{n^{2d}} \langle M_{\Lambda_n} \rangle_{\Lambda_n, \beta, 0} = \frac{1}{2\beta} \mathbb{E}_{\Lambda_n, \beta, 0}\Big(\frac{L(\gamma_0)}{n^d}\Big). \qquad (4.24)$$

It follows from Lemma 4.3 that

$$m^*_{\mathrm{sp}}(\beta) > 0 \iff \liminf_{n \to \infty} \mathbb{E}_{\Lambda_n, \beta, 0}\Big(\frac{L(\gamma_0)}{n^d}\Big) > 0. \qquad (4.25)$$

On the other hand, we have

$$\eta_{\mathrm{macro}}(\beta, 0) = \lim_{\varepsilon \to 0+} \liminf_{n \to \infty} \mathbb{P}_{\Lambda_n, \beta, 0}\Big(\frac{L(\gamma_0)}{n^d} > \varepsilon\Big). \qquad (4.26)$$

The result is then clear.

The claim (c) is identical to (b), with loops instead of cycles. $\square$

The proposition suggests that m^*_{th} and m^*_{res} are related to the existence of infinite cycles, while m^*_{sp} is related to the occurrence of macroscopic cycles. The question is whether a phase exists where a positive fraction of vertices belongs to mesoscopic cycles or loops. Such a phase could have something to do with the Berezinskiĭ-Kosterlitz-Thouless transition [6, 33], which has been rigorously established in the classical XY model [23]. It is not expected in the Heisenberg model, though. The Mermin-Wagner theorem (Section 5.1) rules out any kind of infinite cycles or loops in one and two dimensions.

5. Rigorous results for the quantum models

Quantum lattice systems have seen a considerable amount of studies in past decades and the effort is not abating. Physicists are interested in properties of the ground state (i.e., the eigenvector of the Hamiltonian with lowest eigenvalue), in their dynamical behavior, and about the existence and nature of phase transitions. Out of many results, we only discuss two in this section, that have been chosen because of their direct relevance of the understanding of the cycle and loop models: the Mermin-Wagner about the absence of spontaneous magnetization in one and two dimensions, and the claim of Dyson, Lieb, and Simon, about the existence of long-range order in the antiferromagnetic model.

5.1. Mermin-Wagner theorem. This fundamental result of condensed matter physics states that a continuous symmetry cannot be broken in one and two dimensions [**35**]. In particular, there are no spontaneous magnetization or long-range order in the Heisenberg models.

THEOREM 5.1. *Let $(\Lambda_n^{\mathrm{per}})$ be the sequence of cubic boxes in $\mathbb{Z}^d$ with periodic boundary conditions and diverging size n. For $d = 1$ or 2, and for any β,*
$$m_{\mathrm{res}}^*(\beta) = 0.$$

By Proposition 4.4, all three ferromagnetic order parameters are zero, and there are no infinite cycles by Proposition 4.5 in the cycle model that corresponds to the Heisenberg ferromagnet. The theorem can also be stated for the staggered magnetic field discussed in Section 4.3. One could establish antiferromagnetic counterparts to Lemma 4.3 and Proposition 4.4, and therefore prove that $\eta_\infty(\beta)$ is also zero in the loop model that corresponds to the Heisenberg antiferromagnet.

An open quesetion is whether the theorem can be extended to more general measures of the form
$$\theta^{|\mathcal{C}(\omega)|}\mathrm{d}\rho_{\mathcal{E},\beta}(\omega) \quad \text{and} \quad \theta^{|\mathcal{L}(\omega)|}\mathrm{d}\rho_{\mathcal{E},\beta}(\omega)$$
(up to normalizations), for values of θ other than $\theta = 2$. The case $3^{|\mathcal{L}(\omega)|}$ can actually be viewed as the representation of a *ferro*magnetic model with Hamiltonian $-\sum_{\{x,y\}\in\mathcal{E}}(\vec{S}_x \cdot \vec{S}_y)^2$, see [**1**], and the Mermin-Wagner theorem certainly holds in that case.

The theorem may not apply when θ is too large, and the system is in a phase with many loops, similar to the one studied in [**11**].

We present the standard proof [**40**] that is based on Bogolubov's inequality.

PROPOSITION 5.2 (Bogolubov's inequality). *Let $\beta > 0$ and A, B, H be operators on a finite-dimensional Hilbert space, with H self-adjoint. Then*
$$\left|\mathrm{Tr}\,[A,B]\,\mathrm{e}^{-\beta H}\right|^2 \leq \tfrac{1}{2}\beta\mathrm{Tr}\,(AA^* + A^*A)\,\mathrm{e}^{-\beta H}\,\mathrm{Tr}\,[[B,H],B^*]\,\mathrm{e}^{-\beta H}.$$

PROOF. We only sketch the proof, see [**40**] for more details. Let $\{\varphi_i\}$ be an orthonormal set of eigenvectors of H and $\{e_i\}$ the corresponding eigenvalues. We introduce the following inner product:
$$(A, C) = \sum_{i,j:e_i\neq e_j} \langle\varphi_i, A^*\varphi_j\rangle\langle\varphi_j, C\varphi_i\rangle\frac{\mathrm{e}^{-\beta e_j} - \mathrm{e}^{-\beta e_i}}{e_i - e_j}. \tag{5.1}$$

One can check that
$$(A, A) \leq \tfrac{1}{2}\beta\mathrm{Tr}\,(AA^* + A^*A)\,\mathrm{e}^{-\beta}. \tag{5.2}$$

We choose $C = [B^*, H]$, and we check that

$$\operatorname{Tr}[A, B]\, \mathrm{e}^{-\beta H} = \overline{(A, C)} \tag{5.3}$$

and

$$\operatorname{Tr}\big[[B, H], B^*\big]\, \mathrm{e}^{-\beta H} = (C, C). \tag{5.4}$$

Inserting (5.3) and (5.4) in the Cauchy-Schwarz inequality of the inner product (5.1), and using (5.2), we get Bogolubov's inequality. □

PROOF OF THEOREM 5.1. Let $m_n(\beta, h) = n^{-d}\langle M_{\Lambda_n}\rangle_{\Lambda_n, \beta, h}$. Let

$$S_x^{(\pm)} = \tfrac{1}{\sqrt{2}}(S_x^{(1)} \pm \mathrm{i}S_x^{(2)}). \tag{5.5}$$

One easily checks that

$$[S_x^{(+)}, S_x^{(-)}] = S_x^{(3)}. \tag{5.6}$$

It is convenient to label the sites of Λ_n^{per} as follows

$$\mathcal{V}_n = \{x \in \mathbb{Z}^d : -\tfrac{n}{2} < x_i \le \tfrac{n}{2}, i = 1, \dots, d\}. \tag{5.7}$$

$\mathcal{E}_n$ is again the set of nearest-neighbors in $\mathcal{V}_n$ with periodic boundary conditions. For $k \in \frac{2\pi}{n}\mathcal{V}_n$, we introduce

$$S^{(\cdot)}(k) = \sum_{x \in \mathcal{V}_n} \mathrm{e}^{-\mathrm{i}kx} S_x^{(\cdot)}, \tag{5.8}$$

where kx denotes the inner product in $\mathbb{R}^d$. Then, using (5.6),

$$\langle [S^{(+)}(k), S^{(-)}(-k)]\rangle_{\Lambda_n, \beta, h} = \sum_{x, y \in \mathcal{V}_n} \mathrm{e}^{-\mathrm{i}kx}\, \mathrm{e}^{\mathrm{i}ky}\, \langle [S_x^{(+)}, S_y^{(-)}]\rangle_{\Lambda_n, \beta, h} = n^d m_n(\beta, h). \tag{5.9}$$

This will be the left side of Bogolubov's inequality. For the right side, tedious but straightforward calculations (expansions, commutation relations) give

$$\langle \big[[S^{(+)}(k), H_{\Lambda_n}], S^{(-)}(-k)\big]\rangle_{\Lambda_n, \beta, h}$$
$$= 2 \sum_{x, y: \{x, y\} \in \mathcal{E}_n} (1 - \mathrm{e}^{\mathrm{i}k(x-y)})\langle S_x^{(-)}S_y^{(+)} + S_x^{(3)}S_y^{(3)}\rangle_{\Lambda_n, \beta, h} + hn^d m_n(\beta, h). \tag{5.10}$$

Despite the appearances, this expression is real and positive for any k, as can be seen from (5.4). We get an upper bound by adding the same quantity, but with $-k$. This yields

$$4 \sum_{x, y: \{x, y\} \in \mathcal{E}_n} (1 - \cos k(x-y))\langle S_x^{(-)}S_y^{(+)} + S_x^{(3)}S_y^{(3)}\rangle_{\Lambda_n, \beta, h} + 2hn^d m_n(\beta, h).$$

From Cauchy-Schwarz's inequality and $\|S_x^{(\pm)}\| = \frac{1}{\sqrt{2}}$ (which is easy to check using Pauli matrices) we have

$$\big|\langle S_x^{(-)}S_y^{(+)} + S_x^{(3)}S_y^{(3)}\rangle_{\Lambda_n, \beta, h}\big| \le \tfrac{3}{4}. \tag{5.11}$$

Let us introduce the "dispersion relation" of the lattice

$$\varepsilon(k) = \sum_{i=1}^{d}(1 - \cos k_i). \tag{5.12}$$

Inserting all this stuff in Bogolubov's inequality, we get

$$\frac{n^d m_n(\beta, h)^2}{3\varepsilon(k) + |hm_n(\beta, h)|} \le \beta\langle S^{(+)}(k)S^{(-)}(-k) + S^{(-)}(-k)S^{(+)}(k)\rangle_{\Lambda_n, \beta, h}. \tag{5.13}$$

Summing over all $k \in \frac{2\pi}{n} \mathcal{V}_n$, and using $\sum_k \mathrm{e}^{-\mathrm{i}k(x-y)} = \delta_{x,y}$, we have

$$\sum_k \big\langle S^{(+)}(k)S^{(-)}(-k) + S^{(-)}(-k)S^{(+)}(k) \big\rangle_{\Lambda_n,\beta,h}$$

$$= n^d \sum_{x \in \mathcal{V}_n} \big\langle S_x^{(+)} S_x^{(-)} + S_x^{(-)} S_x^{(+)} \big\rangle_{\Lambda_n,\beta,h} = \tfrac{1}{2} n^d. \quad (5.14)$$

Then

$$m_n(\beta,h)^2 \frac{1}{n^d} \sum_{k \in \frac{2\pi}{n} \mathcal{V}_n} \frac{1}{3\varepsilon(k) + |h m_n(\beta,h)|} \leq \tfrac{1}{2}\beta. \quad (5.15)$$

As $n \to \infty$, we get a Riemann sum,

$$m(\beta,h)^2 \int_{[-\pi,\pi]^d} \frac{\mathrm{d}k}{3\varepsilon(k) + |h m_n(\beta,h)|} \leq \tfrac{1}{2}\beta. \quad (5.16)$$

Since $\varepsilon(k) \approx k^2$ around $k = 0$, the integral diverges when $h \to 0$, so $m(\beta,h)$ must go to 0. $\qquad\square$

Notice that the integral remains finite in $d \geq 3$; the argument only applies to $d = 1, 2$.

5.2. Dyson-Lieb-Simon theorem of existence of long-range order. Following the proof of Fröhlich, Simon, and Spencer, of a phase transition in the classical Heisenberg model [22], Dyson, Lieb, and Simon, proved the existence of long-range order in several quantum lattice models, including the antiferromagnetic quantum Heisenberg model in dimensions $d \geq 5$ [16]. Those articles use the "reflection positivity" method, which was systematized and extended in [20, 21]. We recommend the Prague notes of Tóth [46] and of Biskup [9] for excellent introductions to the topic. See also the notes of Nachtergaele [36].

Recall the definition of $\bar{\sigma}$ in Eq. (4.19).

THEOREM 5.3 (Dyson-Lieb-Simon). *Let* $(\Lambda_n^{\mathrm{per}})$ *be the sequence of cubic boxes in* $\mathbb{Z}^d$, $d \geq 5$, *with even side lengths and periodic boundary conditions. There exists* $\beta_0 < \infty$ *and* $\eta > 0$ *such that, for all* $\beta > \beta_0$, *the Heisenberg antiferromagnet has long-range order,*

$$\bar{\sigma}(\beta) > 0.$$

Clearly, this theorem has remarkable consequences for the loop model with weights $2^{|\mathcal{L}(\omega)|}$. Indeed, there are macroscopic loops, $\eta_{\mathrm{macro}}(\beta) > 0$, provided that β is large enough. The proof of [16] barely fails in $d = 4$; the result is expected to hold for all $d \geq 3$, though.

Despite many efforts and false hopes, there is no corresponding result for the Heisenberg ferromagnet, hence for the cycle model.

The proof of Theorem 5.3 is long and difficult and we are not discussing it here. The interested reader is invited to look at [16, 20, 46].

6. Rigorous results for cycle and loop models

The cycle and loop representations in Theorems 3.3 and 3.4 are interesting in their own right and can be studied using purely probabilistic techniques. Without the physical motivation, the external magnetic field is less relevant and more of an annoyance. We prefer to switch it off. The models in this simpler situation are defined below, with the small generalisation that the geometric weight on the

number of cycles or loops is arbitrary. This is analogous to how, for example, one obtains the random cluster or FK representation of the Ising model.

6.1. Cycle and loop models. As usual we suppose $\Lambda = (\mathcal{V}, \mathcal{E})$ is a finite undirected graph. Recall that the Poisson edge measure $\rho_{\Lambda,\beta}$, $\beta > 0$ for Λ is obtained by attaching independent Poisson point processes on $[0, \beta]$ to each edge.

On each realization ω of the Poisson edge process, we define cycles $\mathcal{C}(\omega)$ and loops $\mathcal{L}(\omega)$ as in §3.1. The random cycle and loop model is obtained via a change of measure in which the number of cycles or loops receives a geometric weight $\vartheta > 0$. That is, the probability measures of interest are

$$
\begin{aligned}
\mathbb{P}_{\Lambda,\beta}^{\mathcal{C}}(\mathrm{d}\omega) &= Z_\Lambda^{\mathcal{C}}(\beta)^{-1}\vartheta^{|\mathcal{C}(\omega)|}\rho_{\mathcal{E},\beta}(\mathrm{d}\omega), \\
\mathbb{P}_{\Lambda,\beta}^{\mathcal{L}}(\mathrm{d}\omega) &= Z_\Lambda^{\mathcal{L}}(\beta)^{-1}\vartheta^{|\mathcal{L}(\omega)|}\rho_{\mathcal{E},\beta}(\mathrm{d}\omega),
\end{aligned}
\tag{6.1}
$$

where $Z_\Lambda(\beta)$ are the appropriate normalisations. As remarked above, $\vartheta = 2$ is the physically relevant choice in both these measures.

The main question deals with the possible occurrence of cycles or loops of diverging lengths. Recall the definitions of the fraction of vertices in infinite cycles/loops, $\eta_\infty(\beta)$, and the fraction of vertices in macroscopic cycles/loops, $\eta_{\mathrm{macro}}(\beta)$, that were defined in Section 4.4. (We drop the dependence in h, since $h = 0$ here.) In the case where the graph is a cubic box in $\mathbb{Z}^d$ with periodic boundary conditions, and $\vartheta = 2$, the Mermin-Wagner theorem essentially ruled out infinite cycles in one and two dimensions, and the theorem of Dyson-Lieb-Simon showed that macroscopic loops are present in $d \geq 5$, provided that the parameter β is sufficiently large.

It is intuitively clear that there cannot be infinite cycles or loops when β is small. We find it useful to write a honest proof of this trivial fact, with an explicit condition that gives a lower bound on the critical value of β. The claim and its proof can be found in Section 6.2.

Two studies of the cycle model with $\vartheta = 1$ devote a mention. Angel considered the model on regular trees, and he proved the existence of infinite cycles (for large enough β) when the degree of the tree is larger than 5 [3]. Schramm considered the model on the complete graph and he obtained a fairly precise description of the distribution of cycle lengths [41]. We review this important result in Section 6.3.

6.2. No infinite cycles at high temperatures. We consider general graphs $\Lambda = (\mathcal{V}, \mathcal{E})$. We let κ denote the maximal degree of the graph, i.e., $\kappa = \sup_{x \in \mathcal{V}} |\{y : \{x, y\} \in \mathcal{E}\}|$. Recall that $L(\gamma_x)$ denotes the length of the cycle or loop that contains $x \times \{0\}$. Let a be the small parameter

$$
a = \begin{cases} \vartheta^{-1}(1 - \mathrm{e}^{-\beta}) & \text{if } \vartheta \leq 1, \\ 1 - \mathrm{e}^{-\beta} & \text{if } \vartheta \geq 1. \end{cases}
\tag{6.2}
$$

THEOREM 6.1. *For either the cycle of the loop model, i.e., either measures in (6.1), we have*

$$
\mathbb{P}_{\Lambda,\beta}(L(\gamma_x) > \beta k) \leq (a(\kappa - 1))^{-1}[a\kappa(1 - \tfrac{1}{\kappa})^{-\kappa+1}]^k.
$$

for every $x \in \mathcal{V}$.

Of course, the theorem is useful only if the right side is less than 1, in which case large cycles have exponentially small probability. This result is pretty reasonable on the square lattice with $\vartheta \leq 1$. When $\vartheta > 1$ configurations with many cycles are

favored, and the domain should allow for larger β. Our condition does not show it. The case $\vartheta \gg 1$ is close to the situation treated in [**11**] with phases of closely packed loops. In the case of the complete graph of N vertices and $\vartheta = 1$, the maximal degree is $\kappa = N - 1$ and the optimal condition is $\beta < 1/N$ (Erdös-Rényi). Using $a\kappa \leq \beta N$ and $(1 - \frac{1}{\kappa})^{-\kappa+1} \leq \mathrm{e}$, we see that our condition is off by a factor e.

As a consequence, we have $\eta_\infty(\beta) = 0$. It implies that $m_{\mathrm{sp}}^*(\beta) = \overline{\sigma}(\beta) = 0$ in the corresponding Heisenberg ferromagnet and antiferromagnet. One could extend the claim so that $m_{\mathrm{th}}^*(\beta) = 0$ as well.

PROOF. Given ω, let $G(\omega) = (V, E)$ denote the subgraph of Λ with edges

$$E = \{e_i : (e_i, t_i) \in \omega\}, \tag{6.3}$$

and $V = \cup_i e_i$ the set of vertices that belong to at least one edge. $G(\omega)$ can be viewed as the percolation graph of ω, where an edge e is open if at least one bridge of the form (e, t) occurs in ω. Then we denote $C_x(\omega) = (V_x, E_x)$ the connected component of $G(\omega)$ that contains x. It is clear that $L(\gamma_x) \leq \beta|V_x|$ for both cycles and loops. Then, using Markov's inequality,

$$\mathbb{P}_{\Lambda,\beta}(L(\gamma_x) > \beta k) \leq \mathbb{P}_\Lambda(|G_x| > k) \leq \eta^{-k}\mathbb{E}_{\Lambda,\beta}(\eta^{|V_x|}), \tag{6.4}$$

for any $\eta \geq 1$. Given a subgraph $G' = (V', E')$ of Λ, let

$$\phi(G') = \vartheta^{-|V'|} \int 1_{G(\omega)=G'} \vartheta^{|\mathcal{C}(\omega)|} \mathrm{d}\rho_{E',\beta}(\omega). \tag{6.5}$$

One can check that

$$\phi(G') \leq \begin{cases} [\vartheta^{-1}(1 - \mathrm{e}^{-\beta})]^{|E'|} & \text{if } \vartheta \leq 1, \\ (1 - \mathrm{e}^{-\beta})^{|E'|} & \text{if } \vartheta \geq 1. \end{cases} \tag{6.6}$$

The bound also holds for the loop model, i.e., if $\vartheta^{|\mathcal{C}(\omega)|}$ is replaced by $\vartheta^{|\mathcal{L}(\omega)|}$ in (6.5). Summing over all possible connected graphs $C_x' = (V_x', E_x')$ that contain x, and then over compatible subgraphs, we have

$$\mathbb{E}_{\Lambda,\beta}(\eta^{|V_x|}) = \sum_{C_x'} \phi(C_x')\eta^{|V_x'|} \frac{\sum_{G' \cap C_x' = \emptyset} \phi(G')}{\sum_{G'} \phi(G')} \leq \sum_{C_x'} \eta^{|V_x'|} a^{|E_x'|}, \tag{6.7}$$

Let $\delta(C_x')$ denote the "depth" of the connected graph C_x', i.e., the minimal number of edges of E_x' that must be crossed in order to reach any point of V_x'. Let

$$B(\ell) = \sum_{C_x', \delta(C_x') \leq \ell} \eta^{|V_x'|} a^{|E_x'|}. \tag{6.8}$$

We want an upper bound for $B(\ell)$ for any ℓ. We show by induction that $B(\ell) \leq b$ for a number b to be determined shortly. We proceed by induction on ℓ. The case $\ell = 0$ is $\eta \leq b$. For $\ell + 1$, we write the sum over graphs with depth less than $\ell + 1$, attached at x, as a sum over graphs of depth less than ℓ, attached at neighbors of x. Neglecting overlaps gives the following upper bound:

$$B(\ell+1) \leq \eta \prod_{y:\{x,\} \in \mathcal{E}} \left(1 + a \sum_{C_y', \delta(C_y') \leq \ell} \eta^{|V_y'|} a^{|E_y'|}\right)$$
$$\leq \eta(1 + ab)^\kappa. \tag{6.9}$$

This needs to be less than b; this condition can be written $a \leq b^{-1}((b/\eta)^{1/\kappa} - 1)$. The optimal choice that maximizes the possible values of a is $b = \eta(1 - \frac{1}{\kappa})^{-\kappa}$. A sufficient condition is then

$$a \leq \tfrac{1}{\eta\kappa}(1 - \tfrac{1}{\kappa})^{\kappa - 1} \tag{6.10}$$

We have obtained that

$$\mathbb{P}_{\Lambda,\beta}(L(\gamma_x) > \beta k) \leq \eta^{-k+1}(1 - \tfrac{1}{\kappa})^{-\kappa}, \tag{6.11}$$

and this holds for all $1 \leq \eta \leq \frac{1}{a\kappa}(1 - \frac{1}{\kappa})^{\kappa-1}$. Choosing the maximal value for η, we get the bound of the theorem. $\qquad\square$

6.3. Rigorous results for trees and the complete graph. The cycle model with $\vartheta = 1$ is known as random stirring and has been studied by several researchers.

Suppose $T_1, T_2, T_3, \ldots$ are independent random transpositions of $\{1, 2, \ldots, n\}$ and $\pi_k = T_1 \circ T_2 \circ \ldots \circ T_k$. Write $\lambda(\pi_k)$ for the vector of cycle lengths in π_k, sorted into order. So, $\lambda_i(\pi_k)$ is the size of the ith largest cycle and if there are less than i cycles in π_k, we take $\lambda_i(\pi_k) = 0$.

Note the simple connection between cycles here and the cycles in our model; if N is a Poisson random variable with mean $\beta n(n-1)/2$, independent of the T_i, then $\lambda(\pi_N)$ has exactly the distribution as the ordered cycle lengths in $\mathcal{C}$ under $\rho_{K_n,\beta}$, where K_n is the complete graph with n vertices.

Schramm proved that for $c > 1/2$, an asymptotic fraction $z(2c)$ of elements from $\{1, 2, \ldots, n\}$ lie in infinite cycles of $\pi_{\lfloor cn \rfloor}$ as $n \to \infty$. The (non-random) fraction $z(2c)$ turns out to be the asymptotic fraction of vertices lying in the giant component of the Erdos-Renyi random graph with edge probability c. Equivalently, $z(s)$ is the survival probability for a Galton-Watson process with Poisson offspring distribution with mean s. Berestycki [5] proved a similar result.

Furthermore, Schramm also showed that the normalised cycle lengths converge to the Poisson-Dirichlet(1) distribution.

THEOREM 6.2 (Schramm [41]). *Let $c > 1/2$. The law of $\lambda(\pi_{\lfloor cn \rfloor})/(nz(2c))$ converges weakly to* PD_1 *as $n \to \infty$.*

7. Uniform split-merge and its invariant measures

We now take a break from spin systems and consider a random evolution on partitions of $[0, 1]$ in which blocks successively split or merge. Stochastic processes incorporating the phenomena of coalescence and fragmentation have been much studied in the recent probability literature (see, for example, [2, 7] or Chapter 5 of [38], and their bibliographies). The space of partitions of $[0, 1]$ provides a natural setting for such processes. The particular model we will discuss here has the property that the splitting and merging can be seen to balance each other out in the long run, so that there exists a stationary (or invariant) distribution. Our aim is to summarise what is known about this invariant distribution. Only a basic familiarity with probability theory is assumed and we'll recall the essentials as we go. This section is self-contained and can be read independently of the first. As is the way among probabilists, we assume there is a phantom probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that hosts all our random variables. It is summoned only when needed.

7.1. Introduction. Let Δ_1 denote the space of (decreasing, countable) partitions of $[0, 1]$. Formally

$$\Delta_1 := \left\{ p \in [0, 1]^{\mathbb{N}} : \ p_1 \geq p_2 \geq \ldots, \ \sum_i p_i = 1 \right\}, \tag{7.1}$$

where the size of the ith part (or block) of $p \in \Delta_1$ is p_i. Each partition induces a natural distribution on its parts – the ith part is sampled with probability p_i. In practice this means sampling a uniform random variable on $[0, 1]$ and choosing whichever block it lands in. This is called size-biased sampling.

Define split and merge operators $S_i^u, M_{ij} : \Delta_1 \to \Delta_1$, $u \in (0, 1)$ as follows:
- $S_i^u p$ is the non-increasing sequence obtained by splitting p_i into two new parts of size $u p_i$ and $(1 - u)p_i$, and
- $M_{ij} p$ is the non-increasing sequence obtained by merging p_i and p_j into a part of size $p_i + p_j$.
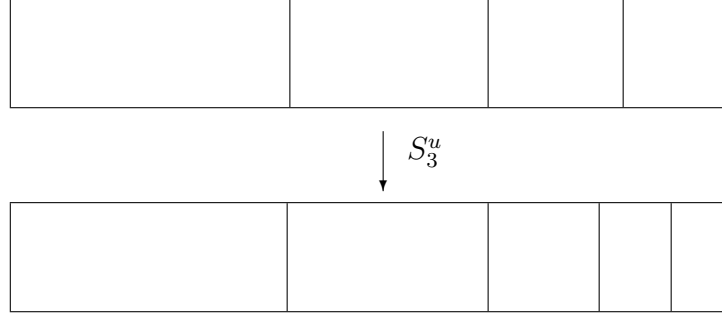

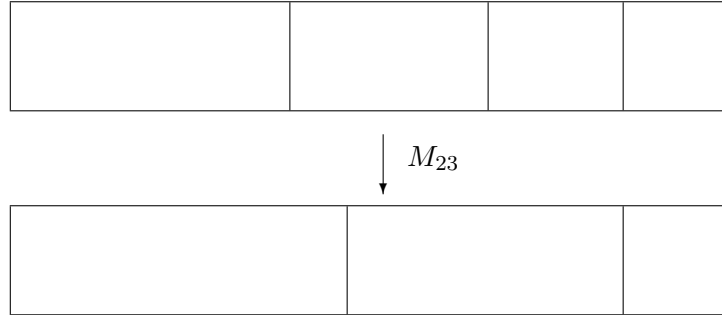FIGURE 6. Splitting the third part of the partition and reordering.

FIGURE 7. Merging second and third parts and reordering.

The basic uniform split-merge transformation of a partition p is defined as follows. Size-biased sample two parts, p_I and p_J say, of p. Each sample is made independently and we allow repetitions. If the same part is chosen twice, i.e. $I = J$, sample a uniform random variable U on $[0, 1]$ and split p_I into two new parts of size $U p_I$ and $(1 - U)p_J$ (i.e. apply S_I^U). If different parts are chosen, i.e. $I \neq J$, then merge them by applying M_{IJ}. This transformation gives a new (random)

element of Δ_1. Conditional on plugging a state $p \in \Delta_1$ into the transformation, the distribution of the new element of Δ_1 obtained is given by the so-called transition kernel

$$K(p, \cdot) := \sum_i p_i^2 \int_0^1 \delta_{S_i^u p}(\cdot) du + \sum_{i \neq j} p_i p_j \delta_{M_{ij} p}(\cdot). \tag{7.2}$$

Repeatedly applying the transformation gives a sequence $P = \left(P^k; \ k = 0, 1, 2, \ldots \right)$ of random partitions evolving in discrete time. We assume that the updates at each step are independent. So, given P^k, the distribution of P^{k+1} is independent of $P^{k-1}, \ldots, P^0$. In other words, P is a discrete time Markov process on Δ_1 with transition kernel K. We call it the basic split-merge chain.

Several authors have studied the large time behaviour of P, and the related issue of invariant probability measures, i.e. μ such that $\mu K = \mu$ (if the initial value P^0 is distributed according to μ, then P^k also has distribution given by μ at all subsequent times $k = 1, 2, \ldots$).

Recent activity begins with Tsilevich [48]. In that paper the author shows that the Poisson-Dirichlet(θ) distribution (defined in §7.2 below and henceforth denoted PD_θ) with parameter $\theta = 1$ is invariant. The paper contains the conjecture (of Vershik) that PD_θ is the only invariant measure.

Uniqueness within a certain class of analytic measures was established by Meyer-Wolf, Zerner and Zeitouni in [34]. In fact they extended the basic split-merge transform described above to allow proposed splits and merges to be rejected with a certain probability. In particular, splits and merges are proposed as above but only accepted with probability $\beta_s \in (0, 1]$ and $\beta_m \in (0, 1]$ respectively, independently at different times. The corresponding kernel is

$$K_{\beta_s, \beta_m}(p, \cdot) := \beta_s \sum_i p_i^2 \int_0^1 \delta_{S_i^u p}(\cdot) du + \beta_m \sum_{i \neq j} p_i p_j \delta_{M_{ij} p}(\cdot) +$$
$$\left(1 - \beta_s \sum_i p_i^2 - \beta_m \sum_{i \neq j} p_i p_j \right) \delta_p(\cdot). \tag{7.3}$$

We call this (β_s, β_s) split-merge (the basic chain, of course, corresponds to $\beta_s = \beta_m = 1$). The Poisson-Dirichlet distribution is still invariant, but with parameter $\theta = \beta_s / \beta_m$.

Tsilevich [47] provided another insight into the large time behaviour of the the basic split-merge process ($\beta_s = \beta_m = 1$). The main theorem is that if $P^0 = (1, 0, 0, \ldots) \in \Delta_1$, then the law of P, sampled at a random Binomial$(n, 1/2)$-distributed time, converges to Poisson-Dirichlet(1) as $n \to \infty$.

Pitman [37] studies a related split-merge transformation, and by developing results by Gnedin and Kerov, reproves PD invariance and refines the uniqueness result of [34]. In particular, PD is the only invariant measure under which Pitman's split-merge transformation composed with 'size-biased permutation' is invariant.

Uniqueness for the basic chain's invariant measure was finally established by Diaconis, Meyer-Wolf, Zerner and Zeitouni in [15]. They coupled the split-merge process to a discrete analogue on integer partitions of $\{1, 2, \ldots, n\}$ and then used representation theory to show the discrete chain is close to equilibrium before decoupling occurs.

Schramm [**42**] used a different coupling to give another uniqueness proof for the basic chain. His arguments readily extend to allow $\beta_s \in (0, 1]$ (although $\beta_m = 1$ is still required). In summary,

THEOREM 7.1.

   (a) *Poisson-Dirichlet(β_s/β_m) is invariant for the uniform split-merge chain with $\beta_s, \beta_m \in (0, 1]$. Furthermore,*
   (b) *If $\beta_m = 1$ then it is the unique invariant measure.*

We give a short proof of part (a) in Section 7.3 below.

7.2. The Poisson-Dirichlet random partition. Write $\mathcal{M}_1(\Delta_1)$ for the set of probability measures on Δ_1. The Poisson-Dirichlet distribution $\mathrm{PD}_\theta \in \mathcal{M}_1(\Delta_1)$, $\theta > 0$, is a one parameter family of laws introduced by Kingman in [**30**]. It has cropped up in combinatorics, population genetics, number theory, Bayesian statistics and probability theory. The interested reader may consult [**18, 31, 4, 39**] for details of the applications. We will simply define it and give some basic properties.

There are two important characterisations of PD_θ. Although both are useful, we'll only actually use one of them (Kingman's original definition). The other is the 'stick-breaking' construction, which is both easier to describe and gives a better intuitive picture. Let $T_1, T_2, \ldots$ be independent $\mathrm{Beta}(1, \theta)$ distributed random variables. (The $\mathrm{Beta}(a, b)$ distribution has density $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} t^a (1-t)^b$ on $[0, 1]$. Consequently, if U is uniform on $[0, 1]$ then $1 - U^{1/\theta}$ is $\mathrm{Beta}(1, \theta)$ distributed.) Form a random partition from the T_i by letting the kth block take fraction T_k of the unallocated mass. That is, the first block has size $P_1 = T_1$, the second $P_2 = T_2(1 - P_1)$ and $P_{k+1} = T_{k+1}(1 - P_1 - \ldots - P_k)$. One imagines taking a stick of unit length and breaking off a fraction T_{k+1} of what remains after k pieces have already been taken. A one-line induction argument shows that $1 - P_1 - \ldots - P_k = (1 - T_1)(1 - T_2) \ldots (1 - T_k)$, giving

$$P_{k+1} = T_{k+1}(1 - T_1)(1 - T_2) \ldots (1 - T_k). \tag{7.4}$$

In case there is any doubt that $\sum_{i=1}^\infty P_i = 1$ almost surely, note that

$$\mathbb{E}\left[1 - \sum_{i=1}^k P_i\right] = \mathbb{E}\left[\prod_{i=1}^k (1 - T_i)\right] = \left(\int_0^1 \theta t (1-t)^{\theta-1}\right)^k = (\theta + 1)^{-k} \to 0 \tag{7.5}$$

as $k \to \infty$. So, the vector $(P_{[1]}, P_{[2]}, \ldots)$ of the P_i sorted into decreasing order is an element of Δ_1. It determines a unique measure $\mathrm{PD}_\theta \in \mathcal{M}_1(\Delta_1)$. This is known as the Griffiths-Engen-McCloskey (GEM) construction. It is interesting to note that the original vector $(P_1, P_2, \ldots)$ is obtained from $(P_{[1]}, P_{[2]}, \ldots)$ by size-biased re-ordering. In other words, consider the interval $[0, 1]$ partitioned into lengths $(P_{[1]}, P_{[2]}, \ldots)$. Take a sequence $U_1, U_2, \ldots$ of i.i.d. uniform random variables on $[0, 1]$. Now list the blocks "discovered" by the uniforms in the order that they are found. The resulting sequence has the same distribution as $(P_1, P_2, \ldots)$.

7.2.1. *Poisson Point processes.* Kingman's characterisation of PD_θ is made in terms of a suitable random point process on $\mathbb{R}_+$. Here is a crash course in the general theory of such processes on a measurable space $(X, \mathcal{B})$. (The standard reference is [**32**].) Although we will only need this theory for $X = \mathbb{R}_+$, there is no extra cost for introducing it in general. Let $\mathcal{M}(X)$ denote the set of σ-finite measures on X.

Suppose that $\mu \in \mathcal{M}(X)$ and consider the special case $\mu(X) < \infty$. Thus $\mu(\cdot)/\mu(X)$ is a probability measure and we can sample, independently, points $Y_1, Y_2, \ldots$ according to this distribution. Let N_0 be Poisson$(\mu(X))$ distributed, which means $\mathbb{P}(N_0 = n) = \exp(-\mu(X))\mu(X)^n/n!$. Conceptually, the Poisson point process with intensity measure μ is simply the random collection $\{Y_1, \ldots, Y_{N_0}\}$.

Formally, the point process is defined in terms of the random counting measure $N(A) = \sum_{i=1}^{N_0} \mathbb{1}_{Y_i \in A}$, which counts the number of random points lying in $A \in \mathcal{B}$. Thus $N(A)$ is a random variable, which has Poisson$(\mu(A))$ distribution. Indeed,

$$
\begin{aligned}
\mathbb{P}(N(A) = k) &= \sum_{n=k}^{\infty} \mathbb{P}(N_0 = n) \, \mathbb{P}\left(\sum_{i=1}^{N_0} \mathbb{1}_{Y_i \in A} = k \,\Big|\, N_0 = n \right) \\
&= \sum_{n=k}^{\infty} \exp(-\mu(X))\frac{\mu(X)^n}{n!} \left(\frac{n!}{k!(n-k)!} \right) \left(\frac{\mu(A)}{\mu(X)} \right)^k \left(1 - \frac{\mu(A)}{\mu(X)} \right)^{n-k} \\
&= \exp(-\mu(X))\frac{\mu(A)^k}{k!} \sum_{n=k}^{\infty} \frac{1}{(n-k)!} (\mu(X) - \mu(A))^{n-k} \\
&= \exp(-\mu(A))\frac{\mu(A)^k}{k!}.
\end{aligned}
\tag{7.6}
$$

Similar calculations show that if $A_1, \ldots, A_k \in \mathcal{B}$ are disjoint then $N(A_1), \ldots, N(A_k)$ are independent. These properties turn out to be sufficient to completely specify the distribution of the random measure N.

DEFINITION 7.1 (Poisson Point Process). *A Poisson Point process on X with intensity $\mu \in \mathcal{M}(X)$ (or PPP(μ) for short) is a random counting measure $N : \mathcal{B}(X) \to \mathbb{N} \cup \{0\} \cup \{\infty\}$ such that*

- *for any $A \in \mathcal{B}(X)$, $N(A)$ has Poisson$(\mu(A))$ distribution. By convention, $N(A) = \infty$ a.s. if $\mu(A) = \infty$.*
- *if $A_1, A_2, \ldots, A_k \in \mathcal{B}$ are disjoint, then the random variables $N(A_1), \ldots, N(A_k)$ are independent.*

For general σ-finite intensity measures, we can construct N by superposition. Suppose that $X = \bigcup_i X_i$ where the X_i are disjoint and $\mu(X_i) < \infty$. Use the recipe given at the start of this section to construct, independently, a PPP$(\mu|_{X_i})$ N_i on each subspace X_i. Then $N(A) = \sum_{i=1}^{\infty} N_i(A)$ is the desired measure. It is purely atomic, and the atoms $Y_1, Y_2, \ldots$ are called the points of the process. In applications it is useful to know moments and Laplace transforms of functionals of the process.

LEMMA 7.2.
(1) *First moment: If $f \geq 0$ or $f \in L^1(\mu)$ then*

$$
\mathbb{E}\Big[\sum_i f(Y_i)\Big] = \int_X f(y)\mu(\mathrm{d}y)
$$

(we agree that both sides can be ∞).
(2) *Campbell's formula: If $f \geq 0$ or $1 - e^{-f} \in L^1(\mu)$ then*

$$
\mathbb{E}\Big[\exp\Big(-\sum_i f(Y_i)\Big)\Big] = \exp\Big(-\int_X (1 - e^{-f(y)})\mu(\mathrm{d}y)\Big)
$$

(we agree that $\exp(-\infty) = 0$).

(3) *Palm formula: Let $\tilde{\mathcal{M}}(X) \subset \mathcal{M}(X)$ denote the space of point measures on X, $G : X \times \tilde{\mathcal{M}} \to \mathbb{R}_+$ be a measurable functional of the points, and f as in (2). Then*

$$\mathbb{E}\Big[\sum_i f(Y_i)G(Y_i, N)\Big] = \int_X \mathbb{E}[G(y, \delta_y + N)]f(y)\mu(dy).$$

The formulation here is that of Lemma 2.3 of [**7**]. We include sketch proofs to give a flavor of the calculations involved.

PROOF. Let $f = \sum_{k=1}^n c_k \mathbb{1}_{A_k}$, be a simple function with $\mu(A_k) < \infty$.
(1)

$$\mathbb{E}\Big[\sum_i f(Y_i)\Big] = \mathbb{E}\Big[\sum_{k=1}^n c_k N(A_k)\Big] = \sum_{k=1}^n c_k \mu(A_k) = \int_X f(y)\mu(dy). \qquad (7.7)$$

(2)

$$\mathbb{E}\Big[\exp\Big(-\sum_i f(Y_i)\Big)\Big] = \mathbb{E}\Big[e^{-\sum_k c_k N(A_k)}\Big] = \prod_{k=1}^n \mathbb{E}\Big[e^{-\sum_k c_k N(A_k)}\Big]$$

$$= \prod_{k=1}^n \exp(-\mu(A_k)(1 - e^{-c_k})) = \exp\Big(-\int_X (1 - e^{-f(y)})\mu(dy)\Big). \qquad (7.8)$$

Both (1) and (2) extend to measurable $f \geq 0$ using standard arguments. Part (1) for $f \in L^1(\mu)$ follows immediately. Part (2) for $1 - e^{-f} \in L^1(\mu)$ is omitted.

(3) First suppose G is of the form $G(N) = \exp(-\sum_i g(Y_i))$ for some non-negative measurable g. Campbell's formula gives, for $q \geq 0$,

$$\mathbb{E}\Big[\exp\Big(-q\sum_i f(Y_i)\Big)G(N)\Big] = \exp\Big(-\int_X (1 - e^{-qf(y)-g(y)})\mu(dy)\Big). \qquad (7.9)$$

Differentiating this identity in q at 0 gives

$$\mathbb{E}\Big[\sum_i f(Y_i)G(N)\Big] = \int_X f(y)e^{-g(y)}\mu(dy)\exp\Big(-\int_X (1 - e^{-g(y)})\mu(dy)\Big)$$

$$= \int_X f(y)e^{-g(y)}\mu(dy)\,\mathbb{E}\Big[\exp\Big(-\sum_i g(Y_i)\Big)\Big]$$

$$= \int_X f(y)\,\mathbb{E}\Big[\exp\Big(-\sum_i g(Y_i) - g(y)\Big)\Big]\mu(dy) \qquad (7.10)$$

$$= \int_X f(y)\,\mathbb{E}[G(N + \delta_y)]\mu(dy),$$

where Campbell's formula is used to get the second and last lines.

Now, suppose $G(y, N) = \sum_{k=1}^n \mathbb{1}_{y \in A_k} \exp(-\sum_i g_k(Y_i))$ for $A_1, \ldots, A_n \in \mathcal{B}$ and measurable $g_k : X \to [0, \infty)$. By linearity, the preceding calculations give

$$\mathbb{E}\Big[\sum_i f(Y_i)G(Y_i, N)\Big] = \int_X \sum_{k=1}^n \mathbb{1}_{y \in A_k} f(y)\,\mathbb{E}\Big[\exp\Big(-\sum_i g_k(Y_i) - g_k(y)\Big)\Big]\mu(dy)$$

$$= \int_X f(y)\,\mathbb{E}[G(y, N + \delta_y)]\mu(dy).$$

The proof is completed by approximating arbitrary measurable G with functions of this form. We omit the details. □

7.2.2. *The Poisson-Dirichlet distribution via a PPP.* Consider the PPP with intensity measure given by $\eta(\mathrm{d}x) = \theta x^{-1} \exp(-x)\mathrm{d}x$ on $[0,\infty)$. (η is an infinite measure, but is σ-finite since $\eta(2^{-k-1}, 2^{-k}] \leq \theta$.) A practical way to construct this process is given in Tavaré [**44**]. Let $T_1 < T_2 < \ldots$ be the points of a Poisson process of rate θ (that is, the differences $T_{i+1} - T_i$ are independent exponential variables of rate θ) and $E_1, E_2, \ldots$ be exponentially distributed with rate 1. Then, the points in our PPP(η) can be enumerated as $\xi_i = \exp(-T_i)E_i$, $i \geq 1$.

The probability that all points are less than $K > 0$ is

$$\mathbb{P}(N(K, \infty) = 0) = \exp\left(-\int_K^\infty \theta x^{-1} \exp(-x)\mathrm{d}x\right) \to 1 \qquad (7.11)$$

as $K \to \infty$. Thus, there is a largest point and we can order the points in decreasing order so $\xi_1 \geq \xi_2 \geq \ldots \geq 0$. The sum $\sum_i \xi_i$ is finite almost surely. Indeed, we can say much more. Recall that the Gamma(γ, λ) distribution has density

$$\frac{1}{\Gamma(\gamma)} \lambda^\gamma x^{\gamma-1} \exp(-\lambda x).$$

LEMMA 7.3.

$$\sum_i \xi_i \sim \text{Gamma}(\theta, 1).$$

PROOF. Since $\sum_i \xi_i$ is a non-negative random variable, its distribution is determined by its Laplace transform. By Campbell's formula, this is given by

$$\mathbb{E}\left[\exp\left(-r\sum_i \xi_i\right)\right] = \exp\left(-\theta\int_0^\infty (1 - e^{-rx})x^{-1}\exp(-x)\mathrm{d}x\right)$$

$$= \exp\left(-\theta\int_0^r \int_0^\infty \exp(-x(1+r))\mathrm{d}x\mathrm{d}r\right)$$

$$= (1+r)^{-\theta},$$

for $|r| < 1$, implying that $\sum_i \xi_i$ is Gamma($\theta, 1$) distributed. □

The Poisson-Dirichlet(θ) distribution, $\text{PD}_\theta \in \mathcal{M}(\Delta_1)$, is the law of the ordered points, normalised by their sum, i.e.

$$\frac{1}{\sum_i \xi_i} (\xi_1, \xi_2, \xi_3, \ldots). \qquad (7.12)$$

In the next section, we will wish to appeal to various properties of Beta and Gamma random variables which are often known as the "Beta-Gamma algebra".

LEMMA 7.4. *Suppose that $\Gamma_\alpha^\lambda \sim \text{Gamma}(\alpha, \lambda)$ and $\Gamma_\beta^\lambda \sim \text{Gamma}(\beta, \lambda)$ are independent. Then*

- $\Gamma_\alpha^\lambda + \Gamma_\beta^\lambda \sim \text{Gamma}(\alpha + \beta, \lambda)$,
- $\Gamma_\alpha^\lambda/(\Gamma_\alpha^\lambda + \Gamma_\beta^\lambda) \sim \text{Beta}(\alpha, \beta)$,
- *The two random variables above are independent.*

Note that the converse also follows: if $B \sim \text{Beta}(\alpha, \beta)$ is independent of $\Gamma_{\alpha+\beta}^\lambda \sim \text{Gamma}(\alpha + \beta, \lambda)$ then $B\Gamma_{\alpha+\beta}^\lambda \sim \text{Gamma}(\alpha, \lambda)$, $(1 - B)\Gamma_{\alpha+\beta}^\lambda \sim \text{Gamma}(\beta, \lambda)$ and these last two random variables are independent.

PROOF. In order to simplify the notation, let $X = \Gamma_\alpha^\lambda$ and $Y = \Gamma_\beta^\lambda$. We will find the joint density of $S = X + Y$ and $R = X/(X+Y)$. We first find the Jacobian corresponding to this change of variables: we have

$$\frac{\partial x}{\partial s} = \quad r, \qquad \frac{\partial x}{\partial r} = s$$
$$\frac{\partial y}{\partial s} = 1 - r, \qquad \frac{\partial y}{\partial r} = -s$$

and so the Jacobian is $|-rs - (1-r)s| = s$. Noting that $X = RS$ and $Y = (1-R)S$, we see that S and R have joint density

$$s \frac{1}{\Gamma(\alpha)} \lambda^\alpha (rs)^{\alpha-1} e^{-\lambda rs} \frac{1}{\Gamma(\beta)} ((1-r)s)^{\beta-1} e^{-\lambda(1-r)s}$$
$$= \frac{1}{\Gamma(\alpha+\beta)} \lambda^{\alpha+\beta} s^{\alpha+\beta-1} e^{-\lambda s} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1} (1-r)^{\beta-1}. \tag{7.13}$$

Since this factorizes with the factors being the correct Gamma and Beta densities respectively, the result follows. $\qquad\square$

In the next lemma, we will see the power of the Beta-Gamma algebra. We use it to make a connection between our two different representations of the Poisson-Dirichlet distribution. This will serve as a warm up for the calculations in the next section.

LEMMA 7.5. *Suppose that $P = (P_1, P_2, \ldots) \sim \mathrm{PD}_\theta$. Let P_* be a size-biased pick from amongst $P_1, P_2, \ldots$. Then $P_* \sim \mathrm{Beta}(1, \theta)$.*

So P_* has the same distribution as the length of the first stick in the stick-breaking construction.

PROOF. Note that, conditional on $P_1, P_2, \ldots$, we have that

$$P_* = P_i \quad \text{with probability } P_i, i \geq 1. \tag{7.14}$$

In order to determine the distribution of P_*, it suffices to find $\mathbb{E}[f(P_*)]$ for all bounded measurable test functions $f : [0,1] \to \mathbb{R}_+$. (Indeed, it would suffice to find $\mathbb{E}[f(P_*)]$ for all functions of the form $f(x) = \exp(-qx)$ i.e. the Laplace transform. However, our slightly unusual formulation will generalize better when we consider random variables on Δ_1 in the next section.) Conditioning on $P_1, P_2, \ldots$ and using the Tower Law we see that

$$\mathbb{E}[f(P_*)] = \mathbb{E}[\mathbb{E}[f(P_*)|P_1, P_2, \ldots]] = \mathbb{E}\left[\sum_{i=1}^\infty P_i f(P_i)\right]. \tag{7.15}$$

Now use the representation (7.12) to see that this is equal to

$$\mathbb{E}\left[\sum_{i=1}^\infty \frac{\xi_i}{\sum_{j=1}^\infty \xi_j} f\left(\frac{\xi_i}{\sum_{k=1}^\infty \xi_k}\right)\right]. \tag{7.16}$$

This is in a form to which we can apply the Palm formula; we obtain

$$\mathbb{E}\left[\int_0^\infty \frac{y}{y + \sum_{i=1}^\infty \xi_i} f\left(\frac{y}{y + \sum_{j=1}^\infty \xi_j}\right) \theta y^{-1} e^{-y} \mathrm{d}y\right]. \tag{7.17}$$

After cancelling y and y^{-1}, we recognise the density of the $\mathrm{Exp}(1)$ ($=\mathrm{Gamma}(1,1)$) distribution and so we can write

$$\mathbb{E}\left[\frac{\theta}{\Gamma + \sum_{i=1}^{\infty} \xi_i} f\left(\frac{\Gamma}{\Gamma + \sum_{j=1}^{\infty} \xi_j}\right)\right], \tag{7.18}$$

where $\Gamma \sim \mathrm{Exp}(1)$ is independent of $\xi_1, \xi_2, \ldots$. Recall that $\sum_{i=1}^{\infty} \xi_i \sim \mathrm{Gamma}(\theta, 1)$. Then by Lemma 7.4, $\Gamma + \sum_{i=1}^{\infty} \xi_i$ is independent of $\Gamma/(\Gamma + \sum_{i=1}^{\infty} \xi_i)$ which has a $\mathrm{Beta}(1, \theta)$ distribution. Hence, we get

$$\mathbb{E}\left[\frac{\theta}{\Gamma + \sum_{i=1}^{\infty} \xi_i}\right] \mathbb{E}\left[f\left(B\right)\right], \tag{7.19}$$

where $B \sim \mathrm{Beta}(1, \theta)$. We conclude by observing that

$$\mathbb{E}\left[\frac{\theta}{\Gamma + \sum_{i=1}^{\infty} \xi_i}\right] = 1. \qquad \square$$

We close this section by noting an important property of the PPP we use to create the Poisson-Dirichlet vector.

LEMMA 7.6. *The random variable* $\sum_{i=1}^{\infty} \xi_i$ *is independent of*

$$\frac{1}{\sum_i \xi_i} \left(\xi_1, \xi_2, \xi_3, \ldots\right).$$

This is another manifestation of the independence in the Beta-Gamma algebra; see [**32**].

7.3. Split-merge invariance of Poisson-Dirichlet.

We use the same method that we exploited in the proof of Lemma 7.5 to prove part (a) of Theorem 7.1.

First define a random function $F : \Delta_1 \to \Delta_1$ corresponding to (β_s, β_m) split-merge as follows. Fix $p \in \Delta_1$ and let $I(p)$ and $J(p)$ be the indices of the two independently size-biased parts of p, that is

$$\mathbb{P}(I(p) = k) = \mathbb{P}(J(p) = k) = p_k, \quad k \geq 1. \tag{7.20}$$

Now let U and V be independent $\mathrm{U}(0,1)$ random variables, independent of $I(p)$ and $J(p)$. Let

$$F(p) = \begin{cases} S_i^U p & \text{if } I(p) = J(p) = i \text{ and } V \leq \beta_s \\ M_{ij} p & \text{if } I(p) = i \neq J(p) = j \text{ and } V \leq \beta_m \\ p & \text{otherwise.} \end{cases} \tag{7.21}$$

We wish to prove that if $P \sim \mathrm{PD}_\theta$ then $F(P) \sim \mathrm{PD}_\theta$ also. Let $g : \Delta_1 \to \mathbb{R}_+$ be a bounded measurable test function which is symmetric in its arguments (this means that we can forget about ordering the elements of our sequences). Then, conditioning on P, considering the different cases and using the Tower Law, we

have

$$
\begin{aligned}
\mathbb{E}[g(F(P))] =& \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{V\leq\beta_s}\sum_{i=1}^{\infty}\mathbb{1}_{I(P)=J(P)=i}\, g(S_i^U P)\,\Big|\,P\right]\right]\\
&+\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{V>\beta_s}\sum_{i=1}^{\infty}\mathbb{1}_{I(P)=J(P)=i}\, g(P)\,\Big|\,P\right]\right]\\
&+\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{V\leq\beta_m}\sum_{i\neq j}\mathbb{1}_{I(P)=i}\,\mathbb{1}_{J(P)=j}\, g(M_{ij}P)\,\Big|\,P\right]\right]\\
&+\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{V>\beta_m}\sum_{i\neq j}\mathbb{1}_{I(P)=i}\,\mathbb{1}_{J(P)=j}\, g(P)\,\Big|\,P\right]\right]
\end{aligned}
\tag{7.22}
$$

Note that, conditional on P, $I(P) = i$, $J(P) = j$ with probability P_iP_j, so that we get

$$
\begin{aligned}
\mathbb{E}[g(F(P))] =&\beta_s\,\mathbb{E}\left[\sum_{i=1}^{\infty}P_i^2 g(S_i^U P)\right]+(1-\beta_s)\,\mathbb{E}\left[\sum_{i=1}^{\infty}P_i^2 g(P)\right]\\
&+\beta_m\,\mathbb{E}\left[\sum_{i\neq j}P_iP_j g(M_{ij}P)\right]+(1-\beta_m)\,\mathbb{E}\left[\sum_{i\neq j}P_iP_j g(P)\right].
\end{aligned}
\tag{7.23}
$$

Now use the symmetry of g to write

$$
g(S_k^U P) = g\left((P_k U, P_k(1-U), (P_i : i\geq 1, i\neq k))\right)
$$

and

$$
g(M_{ij}P) = g\left((P_i + P_j, (P_k : k\geq 1, k\neq i, j))\right).
$$

Set $(P_1, P_2, \ldots) = \frac{1}{\sum_{i=1}^{\infty}\xi_i}(\xi_1, \xi_2, \ldots)$ as in (7.12) to obtain

$$
\begin{aligned}
\mathbb{E}[g(F(P))] =&\ \beta_s\,\mathbb{E}\left[\sum_{k=1}^{\infty}\frac{\xi_k^2}{\left(\sum_{i=1}^{\infty}\xi_i\right)^2}g\left(\frac{1}{\sum_{i=1}^{\infty}\xi_i}(\xi_k U, \xi_k(1-U), (\xi_i : i\geq 1, i\neq k))\right)\right]\\
&+(1-\beta_s)\,\mathbb{E}\left[\sum_{k=1}^{\infty}\frac{\xi_k^2}{\left(\sum_{i=1}^{\infty}\xi_i\right)^2}g\left(\frac{1}{\sum_{i=1}^{\infty}\xi_i}(\xi_i : i\geq 1)\right)\right]\\
&+\beta_m\,\mathbb{E}\left[\sum_{i\neq j}\frac{\xi_i\xi_j}{\left(\sum_{k=1}^{\infty}\xi_k\right)^2}g\left(\frac{1}{\sum_{k=1}^{\infty}\xi_k}(\xi_i+\xi_j, (\xi_k : k\geq 1, k\neq i, j))\right)\right]\\
&+(1-\beta_m)\,\mathbb{E}\left[\sum_{i\neq j}\frac{\xi_i\xi_j}{\left(\sum_{k=1}^{\infty}\xi_k\right)^2}g\left(\frac{1}{\sum_{i=1}^{\infty}\xi_i}(\xi_i : i\geq 1)\right)\right]
\end{aligned}
\tag{7.24}
$$

The Palm formula (Lemma 7.2, (3)) applied to each of the expectations above (twice for the double sums) gives that $\mathbb{E}[g(F(P))]$ is equal to

$$
\theta\beta_s\,\mathbb{E}\left[\int_0^\infty \frac{x^{-1}e^{-x}x^2}{(x+\sum_{k=1}^\infty \xi_k)^2}g\left(\frac{1}{x+\sum_{k=1}^\infty \xi_k}(xU, x(1-U), (\xi_i:i\geq 1))\right)\mathrm{d}x\right]
$$

$$
+\,\theta(1-\beta_s)\,\mathbb{E}\left[\int_0^\infty \frac{x^{-1}e^{-x}x^2}{(x+\sum_{k=1}^\infty \xi_k)^2}g\left(\frac{1}{x+\sum_{k=1}^\infty \xi_k}(x, (\xi_i:i\geq 1))\right)\mathrm{d}x\right]
$$

$$
+\,\theta^2\beta_m\,\mathbb{E}\left[\int_0^\infty\int_0^\infty \frac{x^{-1}e^{-x}y^{-1}e^{-y}xy}{(x+y+\sum_{k=1}^\infty \xi_k)^2}g\left(\frac{1}{x+y+\sum_{k=1}^\infty \xi_k}(x+y, (\xi_i:i\geq 1))\right)\mathrm{d}x\mathrm{d}y\right]
$$

$$
+\,\theta^2(1-\beta_m)\,\mathbb{E}\left[\int_0^\infty\int_0^\infty \frac{x^{-1}e^{-x}y^{-1}e^{-y}xy}{(x+y+\sum_{k=1}^\infty \xi_k)^2}g\left(\frac{1}{x+y+\sum_{k=1}^\infty \xi_k}(x, y, (\xi_i:i\geq 1))\right)\mathrm{d}x\mathrm{d}y\right].
$$

$$(7.25)$$

It helps to recognise the densities we are integrating over here (after cancellation). In the first two expectations, which correspond to split proposals, we have the density xe^{-x} of the Gamma(2,1) distribution. The other density to appear is $e^{-x}e^{-y}$, which corresponds to a pair of independent standard exponential variables. It follows that $\mathbb{E}[g(F(P))]$ is equal to

$$
\theta\beta_s\,\mathbb{E}\left[\frac{1}{(\Gamma+\sum_{k=1}^\infty \xi_k)^2}g\left(\frac{1}{\Gamma+\sum_{k=1}^\infty \xi_k}(\Gamma U, \Gamma(1-U), (\xi_i:i\geq 1))\right)\right]
$$

$$
+\,\theta(1-\beta_s)\,\mathbb{E}\left[\frac{1}{(\Gamma+\sum_{k=1}^\infty \xi_k)^2}g\left(\frac{1}{\Gamma+\sum_{k=1}^\infty \xi_k}(\Gamma, (\xi_i:i\geq 1))\right)\right]
$$

$$
+\,\theta^2\beta_m\,\mathbb{E}\left[\frac{1}{(\Gamma+\sum_{k=1}^\infty \xi_k)^2}g\left(\frac{1}{\Gamma+\sum_{k=1}^\infty \xi_k}(\Gamma, (\xi_i:i\geq 1))\right)\right]
$$

$$
+\,\theta^2(1-\beta_m)\,\mathbb{E}\left[\frac{1}{(\Gamma+\sum_{k=1}^\infty \xi_k)^2}g\left(\frac{1}{\Gamma+\sum_{k=1}^\infty \xi_k}(\Gamma U, \Gamma(1-U), (\xi_i:i\geq 1))\right)\right],
$$

$$(7.26)$$

where $\Gamma\sim\text{Gamma}(2,1)$, independently of $(\xi_i:i\geq 1)$. By Lemmas 7.4 and 7.6, $\Gamma+\sum_k \xi_k$ is Gamma$(2+\theta,1)$ distributed and independent of the argument of g in all of the above expectations. More calculation shows that

$$
\mathbb{E}\left[\frac{1}{(\Gamma+\sum_{k=1}^\infty \xi_k)^2}\right]=\frac{1}{\theta(\theta+1)},
$$

$$(7.27)$$

and so we are left with

$$
\frac{\theta\beta_s+\theta^2(1-\beta_m)}{\theta(\theta+1)}\,\mathbb{E}\left[g\left(\frac{1}{\Gamma+\sum_{k=1}^\infty \xi_k}(\Gamma U, \Gamma(1-U), (\xi_i:i\geq 1))\right)\right]
$$

$$
+\frac{\theta(1-\beta_s)+\theta^2\beta_m}{\theta(\theta+1)}\,\mathbb{E}\left[g\left(\frac{1}{\Gamma+\sum_{k=1}^\infty \xi_k}(\Gamma, (\xi_i:i\geq 1))\right)\right]
$$

$$(7.28)$$

Next use $\beta_s=\theta\beta_m$ to get

$$
\theta\beta_s+\theta^2(1-\beta_m)=\theta^2 \quad\text{and}\quad \theta(1-\beta_s)+\theta^2\beta_m=\theta.
$$

$$(7.29)$$

So the expression simplifies to

$$\frac{\theta}{(\theta+1)}\,\mathbb{E}\Big[g\Big(\frac{1}{\Gamma+\sum_{k=1}^{\infty}\xi_k}(\Gamma U,\Gamma(1-U),(\xi_i:i\geq 1))\Big)\Big]$$
$$+\frac{1}{(\theta+1)}\,\mathbb{E}\Big[g\Big(\frac{1}{\Gamma+\sum_{k=1}^{\infty}\xi_k}(\Gamma,(\xi_i:i\geq 1))\Big)\Big]$$

$$(7.30)$$

We can re-express this as a sum of expectations as follows:

$$\frac{1}{\theta(\theta+1)}\,\mathbb{E}\Big[\int_0^{\infty}\int_0^{\infty}\theta^2 e^{-x}e^{-y}g\Big(\frac{1}{x+y+\sum_{k=1}^{\infty}\xi_k}(x,y,(\xi_i:i\geq 1))\Big)\mathrm{d}x\mathrm{d}y\Big]$$
$$+\frac{1}{\theta(\theta+1)}\,\mathbb{E}\Big[\int_0^{\infty}\theta x e^{-x}g\Big(\frac{1}{x+\sum_{k=1}^{\infty}\xi_k}(x,(\xi_i:i\geq 1))\Big)\mathrm{d}x\mathrm{d}y\Big].$$

$$(7.31)$$

Using the Palm formula in the other direction gives

$$\frac{1}{\theta(\theta+1)}\,\mathbb{E}\Big[\sum_{i\neq j}\xi_i\xi_j g\Big(\frac{1}{\sum_{k=1}^{\infty}\xi_k}(\xi_k:k\geq 1)\Big)+\sum_{k=1}^{\infty}\xi_k^2 g\Big(\frac{1}{\sum_{k=1}^{\infty}\xi_k}(\xi_k:k\geq 1)\Big)\Big]$$
$$=\frac{1}{\theta(\theta+1)}\,\mathbb{E}\Big[\Big(\sum_{k=1}^{\infty}\xi_k\Big)^2 g\Big(\frac{1}{\sum_{k=1}^{\infty}\xi_k}(\xi_k:k\geq 1)\Big)\Big].$$

$$(7.32)$$

Once again, $\sum_{k=1}^{\infty}\xi_k$ is independent of the argument of g. Moreover, it is easily shown that

$$\mathbb{E}\Big[\Big(\sum_{k=1}^{\infty}\xi_k\Big)^2\Big]=\theta(\theta+1),\qquad(7.33)$$

since it is simply the second moment of a $\mathrm{Gamma}(\theta,1)$ random variable. Thus,

$$\mathbb{E}[g(F(P))]=\mathbb{E}[g(P)],\qquad(7.34)$$

from which the result follows. □

7.4. Split-merge in continuous time. In the next section all of our dynamics will be in continuous time. In preparation, we close this section by describing a continuous time version of the split-merge process. To warm up, consider the standard Poisson counting process (N_t, $t\geq 0$), perhaps the simplest continuous time Markov chain. Its trajectories take values in $\{0,1,2,\ldots\}$, are piecewise constant, increasing and right continuous. At each integer k, it is held for an exponentially distributed random time before jumping to $k+1$. Consequently, only finitely many jumps are made during each finite time interval. We say N_t increments at rate 1.

Continuous time split-merge is the process (P^{N_t}, $t\geq 0$) obtained by composing (P^k, $k=0,1,2,3,\ldots$) with an independent Poisson counting process. It is a Markov process in Δ_1 with the following dynamics. Suppose the present state is $p\in\Delta_1$ and attach to each part p_i an exponential alarm clock of rate $\beta_s p_i^2$ and to each pair (p_i,p_j) of distinct parts a clock of rate $2\beta_m p_i p_j$.

Wait for the first clock to ring. If p_i's clock rings first then split p_i uniformly (i.e. apply S_i^U with U uniform). If the alarm for (p_i,p_j) rings first then apply M_{ij}. In other words, part p_i splits uniformly at rate $\beta_s p_i$ and distinct parts p_i and p_j merge at rate $2\beta_m p_i p_j$.

More formally, define the rate kernel $Q : \Delta_1 \times \mathcal{B}(\Delta_1) \to [0, \infty)$ by

$$Q(p, \cdot) := \beta_s \sum_i p_i^2 \int_0^1 \delta_{S_i^u p}(\cdot) du + \beta_m \sum_{i \neq j} p_i p_j \delta_{M_{ij} p}(\cdot) \qquad (7.35)$$

and the (uniformly bounded) 'rate of leaving' $c : \Delta_1 \to [0, \infty)$

$$q(p) := Q(p, \Delta_1) = \beta_s \sum_i p_i^2 + \beta_m \sum_{i \neq j} p_i p_j. \qquad (7.36)$$

Using standard theory (e.g. Proposition 12.20, [**29**]), there exists a Markov process on Δ_1 that waits for an Exponential($q(p)$) amount of time in state p before jumping to a new state chosen according to $Q(p, \cdot)/q(p)$.

Furthermore, since

$$K_{\beta_s, \beta_m}(p, \cdot) = Q(p, \cdot) + (1 - q(p))\delta_p(\cdot), \qquad (7.37)$$

this process is constructed explicitly as $(P^{N_t}, \ t \geq 0)$. The coincidence of the invariant measures in discrete and continuous time is immediate.

LEMMA 7.7. *A measure $\nu \in \mathcal{M}(\Delta_1)$ is invariant for the continuous time process* $(P^{N_t}, \ t \geq 0)$ *if, and only if, it is invariant for* $(P^k, \ k = 0, 1, 2, 3, \ldots)$.

8. Effective split-merge process of cycles and loops

This section contains a heuristic argument that connects the loop and cycle models of section 6.1 and the split-merge process in section 7.4. The heuristic leads to the conjecture that the asymptotic normalized lengths of the cycles and loops have Poisson-Dirichlet distribution. By looking at the rates of the effective split-merge process, we can identify the parameter of the distribution.

Both loop and cycle models are treated with the same arguments. Thus for notational convenience and brevity we restrict attention to the cycle model. All statements made can be modified for loops, and we shall give some hints whenever the modification is non-trivial.

As hinted at in section 6.1, we conjecture that for β large enough, macroscopic cycles emerge as $\Lambda \nearrow \mathbb{Z}^d$.

CONJECTURE 8.1. *Suppose $d \geq 3$. There exists $\beta_c > 0$ such that for $\beta > \beta_c$,* $\eta_{\mathrm{macro}} = \eta_{\mathrm{macro}}(\beta) > 0$, *i.e. the vector*

$$\left(\frac{\lambda^{(1)}}{\eta_{\mathrm{macro}} n^d}, \frac{\lambda^{(2)}}{\eta_{\mathrm{macro}} n^d}, \ldots \right) \qquad (8.1)$$

converges to a point ξ in Δ_1 as $G \nearrow \mathbb{Z}^d$.

Assuming the conjectured result is true, what is the distribution of ξ? In some related models (the random-cluster model), ξ has found to be the trivial (and non-random!) partition $(1, 0, 0, \ldots)$.

However, we conjecture (8.2) that there are *many* macroscopic cycles in our model (rather than a unique giant cycle) and that their relative lengths can be described explicitly by the Poisson-Dirichlet distribution.

CONJECTURE 8.2. *The distribution of ξ in Conjecture 8.1 is PD_θ for an appropriate choice of θ.*

The rest of this section is concerned with justifying this conjecture. The reader may guess what the parameter θ should be. We will tease it out below and identify it in section 8.4.

Please see section 4 for a summary of rigorous results by Schramm to support this conjecture on the complete graph.

8.1. Burning and building bridges. We'll define an ergodic Markov process on Ω with $\mathbb{P}^{\mathcal{C}}_{\Lambda,\beta}$ as invariant measure. The process evolves by adding or removing bridges to the current configuration. Conveniently, the effect of such an operation is to either split a cycle or merge two cycles.

LEMMA 8.1. *Suppose $\omega \in \Omega$ and ω' is ω with either a bridge added (i.e. $\omega' = \omega \cup \{(e,t)\}$ for some $(e,t) \in E \times [0,\beta]$) or a bridge removed (i.e. $\omega' = \omega - \{(e,t)\}$ for some $(e,t) \in \omega$).*
Then $\mathcal{C}(\omega')$ is obtained by splitting a cycle or merging two cycles in $\mathcal{C}(\omega)$. Similarly, $\mathcal{L}(\omega')$ is obtained by a splitting or merging in $\mathcal{L}(\omega)$.

The point is that adding or removing a bridge never causes, for example, several cycles to join, a cycle to split into many pieces or the cycle structure to remain unchanged.

The Lemma is most easily justified by drawing pictures for the different cases. Suppose that we add a new bridge. Either both endpoints of the new bridge belong to the same cycle or two different cycles. In the former case, the cycle is split and in the latter case, the two cycles are joined. This is illustrated in Figure 8 for cycles and Figure 9 for loops.
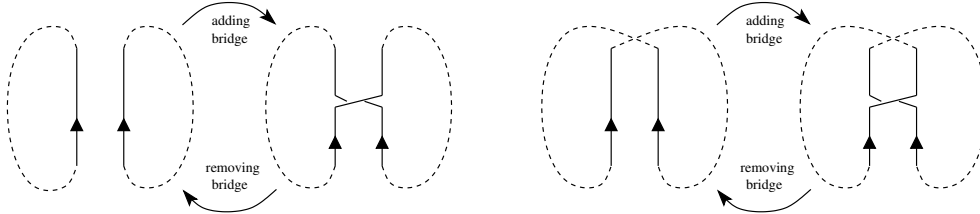

FIGURE 8. Adding or removing bridges always split or merge cycles. Up to topological equivalence, this figure lists all possibilities.

Suppose that we remove an existing bridge. Again, either both of the bridge's endpoints belong to the same cycle or they are in different cycles. In the former case, removal splits the cycle and in the latter, the two cycles are joined.

As this argument hints, it is helpful to formally define the 'contacts' between cycles. Suppose that $\gamma \in \mathcal{C}(\omega)$ is a cycle. Recall that precisely, this means $\gamma : [0,\infty) \to V$ is right continuous, piecewise constant with $\gamma(0) = \gamma(n\beta)$ (for some $n \in \mathbb{N}$) and a jump discontinuity at t across the edge $e = (\gamma(t-), \gamma(t)) \in E$ if, and only if, the bridge $(e, t - j\beta)$ is present in ω for some integer j. Such bridges are called self contact bridges, the set of which is denoted $\mathfrak{B}_\gamma$. Removing a bridge from $\mathfrak{B}_\gamma \subset \omega$ causes γ to split.

The self contact zone $\mathfrak{C}_\gamma$ of γ is the set of $(e,t) \in E \times [0,\beta]$ for which $e = (\gamma(t+j_1\beta), \gamma(t+j_2\beta))$ for some integer j_1, j_2, i.e. the (e,t) bridge touches different legs of γ's trajectory and so adding a bridge from $\mathfrak{C}_\gamma$ splits γ.
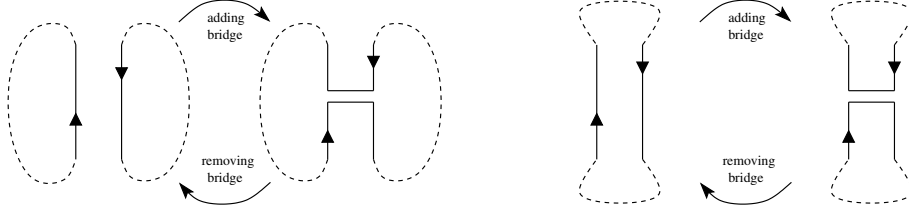
FIGURE 9. Same as Figure 8, but for loops instead of cycles.

The contact bridges $\mathfrak{B}_{\gamma,\gamma'}$ and zones $\mathfrak{C}_{\gamma,\gamma'}$ between distinct cycles $\gamma, \gamma' \in \mathcal{C}(\omega)$ are defined similarly. Specifically, $\mathfrak{B}_{\gamma,\gamma'} \subset \omega$ is comprised of bridges in ω that are traversed by γ and γ', i.e. $(e,t) \in \omega$ such that $e = (\gamma(t+j_1\beta), \gamma'(t+j_2\beta))$ for some j_1, j_2. Removal of a bridge in $\mathfrak{B}_{\gamma,\gamma'}$ causes γ and γ' to merge.

$\mathfrak{C}_{\gamma,\gamma'}$ is the set of $(e,t) \in E \times [0,\beta]$ such that $e = (\gamma(t+j_1\beta), \gamma'(t+j_2\beta))$ for some j_1, j_2, i.e. those bridges that would merge γ and γ'. Note that the contact (and self contact) zones partition $E \times [0,\beta]$ while the contact bridges partition ω.

8.2. Dynamics. The promised $\mathbb{P}^{\mathcal{C}}_{\Lambda,\beta}$-invariant Markov process, denoted $(\mathfrak{X}_t, \ t \geq 0)$ is defined as follows. Suppose that $\alpha > 0$.

- A new bridge appears in (e, dt) at rate $\vartheta^{\alpha}dt$ if its appearance causes a cycle to split and at rate $\vartheta^{-\alpha}dt$ if it causes two cycles to join.
- An existing bridge is removed at rate $\vartheta^{1-\alpha}$ if its removal causes a cycle to split and at rate $\vartheta^{-(1-\alpha)}$ if its removal causes two cycles to join.
- No other transitions occur.

The rates are not uniformly bounded, so a little effort is required to check $\mathfrak{X}$ is well behaved (does not 'explode'). Accepting this, we can show $\mathfrak{X}$ is actually reversible with respect to our cycle model.

LEMMA 8.2. *The unique invariant measure of $\mathfrak{X}$ is $\mathbb{P}^{\mathcal{C}}_{\Lambda,\beta}$.*

PROOF. The proof is straightforward but omitted. $\square$

In the sequel we take $\alpha = 1/2$ so that adding and removing bridges occurs at the same rates.

8.3. Heuristic for rates of splitting and merging of cycles. As we know, adding or removing bridges causes cycles to split or merge so the dynamics $(\mathcal{C}(\mathfrak{X}_t), \ t \geq 0)$ that $\mathfrak{X}$ induces on cycles is a kind of coagulation and fragmentation process. However they are not Markovian and depend on the underlying process in a complicated manner. Ideally we would like a simpler, more transparent description for the dynamics. The first step towards this is to rewrite the transition rates for $\mathfrak{X}$ in terms of the contact zones and bridges.

Suppose that $\mathfrak{X}$ is currently in state $\omega \in \Omega$. A cycle $\gamma \in \mathcal{C}(\omega)$ splits if either a bridge from $\mathfrak{C}_\gamma$ is added, or a bridge from $\mathfrak{B}_\gamma \subset \omega$ is removed. The total rate at which these transitions occur is

$$\sqrt{\vartheta} \left(|\mathfrak{B}_\gamma| + |\mathfrak{C}_\gamma| \right), \tag{8.2}$$

where $|\mathfrak{C}_\gamma| = \sum_{e \in E} \text{Leb}(\{t \in [0,\beta] : (e,t) \in \mathfrak{C}_\gamma\})$ is the (one dimensional) Lebesgue measure of the self contact zone. Two distinct cycles γ and γ' merge if a bridge

from $\mathfrak{C}_{\gamma,\gamma'}$ is added or one from $\mathfrak{B}_{\gamma,\gamma'}$ removed. The combined rate is

$$\sqrt{\vartheta}^{-1} \left(|\mathfrak{B}_{\gamma,\gamma'}| + |\mathfrak{C}_{\gamma,\gamma'}| \right),\tag{8.3}$$

where $|\mathfrak{C}_{\gamma,\gamma'}| = \sum_{e \in E} \mathrm{Leb}(\{t \in [0,\beta] : (e,t) \in \mathfrak{C}_{\gamma,\gamma'}\})$.

8.3.1. *Heuristics.* We believe that for suitably connected graphs and large enough β cycles should be macroscopic. The trajectories of these cycles should spread evenly over all edges and vertices in the graph. In particular, macroscopic cycles should come into contact with each other many times and we expect some averaging phenomenon to come into play. The longer a cycle is, on average, the more intersections with other cycles it should have. In particular, we believe the contact zone between two macroscopic cycles has size proportional to the cycles' length.

That is, if γ and γ' are cycles with lengths λ and λ' respectively then there is a 'law of large numbers'

$$|\mathfrak{C}_{\gamma}| \sim \frac{1}{2} c_2 \lambda^2, \quad |\mathfrak{B}_{\gamma}| \sim \frac{1}{2} c_1 \lambda^2 \tag{8.4}$$

and

$$|\mathfrak{C}_{\gamma,\gamma'}| \sim c_2 \lambda\lambda', \quad |\mathfrak{B}_{\gamma,\gamma'}| \sim c_1 \lambda\lambda', \tag{8.5}$$

for constants c_1 and c_2 (the notation $X \sim Y$ means that the ratio of the random variables converges to 1 in probability as G grows).

The constants may depend on ϑ and β and the graph geometry. We believe they are linear in β but do not depend on ϑ. Note that the size of the contact zones can be calculated easily for the complete graph. We get

$$|\mathfrak{C}_{\gamma,\gamma'}| = \beta\lambda\lambda', \quad |\mathfrak{C}_{\gamma}| = \tfrac{\beta}{2}\lambda(\lambda - 1). \tag{8.6}$$

In the case $\vartheta = 1$, we also have numeric support for

$$|\mathfrak{B}_{\gamma,\gamma'}| \sim \beta\lambda\lambda', \quad |\mathfrak{B}_{\gamma}| \sim \tfrac{\beta}{2}\lambda^2. \tag{8.7}$$

In light of this, our heuristic so far does not seem completely outrageous.

8.4. Connection to uniform split-merge. Continuing with the heuristic, $\mathcal{C}(\mathfrak{X})$ is 'nearly' a Markov process in which cycles split and merge. Substituting (8.4) into (8.2) and (8.5) into (8.3), and multiplying by $2\sqrt{\vartheta}(c_1 + c_2)$ (which just changes the speed of the process, not the invariant measure) we see that a cycle of length λ splits at rate $\vartheta\lambda^2$, while two cycles with lengths λ and λ' merge at rate $2\lambda\lambda'$. There seems no reason to suppose that splits are not uniform.

Suddenly there are many similarities between $\mathcal{C}(\mathfrak{X})$ and the continuous time split-merge process of section 7.4. This suggests that Poisson-Dirichlet PD_θ is lurking somewhere in the normalised cycle length distribution. What is the right choice of the parameter θ?

Write $\vartheta = \beta_s/\beta_m$, $\beta_s, \beta_m \in (0,1]$ and multiply the rates by β_m to see that a cycle of length λ splits uniformly at rate $\beta_s\lambda^2$, while two cycles with lengths λ and λ' merge at rate $2\beta_m\lambda\lambda'$. Up to the normalising factor (which is close to the constant $\nu|G|$), these are exactly the rates in 7.4. Thus, the parameter θ should be simply ϑ, which was not initially obvious.

References

[1] M. Aizenman and B. Nachtergaele. Geometric aspects of quantum spin states. *Comm. Math. Phys.*, 164(1):17–63, 1994.

[2] D. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.

[3] O. Angel. Random infinite permutations and the cyclic time random walk. In *Discrete random walks (Paris, 2003)*, Discrete Math. Theor. Comput. Sci. Proc., AC, pages 9–16 (electronic). Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2003.

[4] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003.

[5] N. Berestycki. Emergence of giant cycles and slowdown transition in random transpositions and k-cycles. *Electr. J. Probab.*, 16:152–173, 2011.

[6] V. L. Berezinskiĭ. Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group i. classical systems. *Soviet J. Exper. Theor. Phys.*, 32:493–500, 1971.

[7] J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[8] V. Betz and D. Ueltschi. Spatial random permutations and poisson-dirichlet law of cycle lengths. http://arxiv.org/abs/1007.2224, 2010.

[9] M. Biskup. Reflection positivity and phase transitions in lattice spin models. In *Methods of contemporary mathematical statistical physics*, volume 1970 of *Lecture Notes in Math.*, pages 1–86. Springer, Berlin, 2009.

[10] C. Borgs, R. Kotecký, and D. Ueltschi. Low temperature phase diagrams for quantum perturbations of classical spin systems. *Comm. Math. Phys.*, 181(2):409–446, 1996.

[11] L. Chayes, L. P. Pryadko, and K. Shtengel. Intersecting loop models on $\mathbf{Z}^d$: rigorous results. *Nuclear Phys. B*, 570(3):590–614, 2000.

[12] J. G. Conlon and J. P. Solovej. Upper bound on the free energy of the spin 1/2 Heisenberg ferromagnet. *Lett. Math. Phys.*, 23(3):223–231, 1991.

[13] N. Crawford and D. Ioffe. Random current representation for transverse field Ising model. *Comm. Math. Phys.*, 296(2):447–474, 2010.

[14] N. Datta, R. Fernández, and J. Fröhlich. Low-temperature phase diagrams of quantum lattice systems. I. Stability for quantum perturbations of classical systems with finitely-many ground states. *J. Statist. Phys.*, 84(3-4):455–534, 1996.

[15] P. Diaconis, E. Mayer-Wolf, O. Zeitouni, and M. P. W. Zerner. The Poisson-Dirichlet law is the unique invariant distribution for uniform split-merge transformations. *Ann. Probab.*, 32(1B):915–938, 2004.

[16] F. J. Dyson, E. H. Lieb, and B. Simon. Phase transitions in quantum spin systems with isotropic and nonisotropic interactions. *J. Statist. Phys.*, 18(4):335–383, 1978.

[17] W. G. Faris. Outline of quantum mechanics. In *Entropy and the quantum*, volume 529 of *Contemp. Math.*, pages 1–52. Amer. Math. Soc., Providence, RI, 2010.

[18] S. Feng. *The Poisson-Dirichlet distribution and related topics*. Probability and its Applications (New York). Springer, Heidelberg, 2010. Models and asymptotic behaviors.

[19] G. B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[20] J. Fröhlich, R. Israel, E. H. Lieb, and B. Simon. Phase transitions and reflection positivity. I. General theory and long range lattice models. *Comm. Math. Phys.*, 62(1):1–34, 1978.

[21] J. Fröhlich, R. B. Israel, E. H. Lieb, and B. Simon. Phase transitions and reflection positivity. II. Lattice systems with short-range and Coulomb interactions. *J. Statist. Phys.*, 22(3):297–347, 1980.

[22] J. Fröhlich, B. Simon, and T. Spencer. Infrared bounds, phase transitions and continuous symmetry breaking. *Comm. Math. Phys.*, 50(1):79–95, 1976.

[23] J. Fröhlich and T. Spencer. The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas. *Comm. Math. Phys.*, 81(4):527–602, 1981.

[24] D. Gandolfo, J. Ruiz, and D. Ueltschi. On a model of random cycles. *J. Stat. Phys.*, 129(4):663–676, 2007.

[25] J. Ginibre. Existence of phase transitions for quantum lattice systems. *Comm. Math. Phys.*, 14(3):205–234, 1969.

[26] G. R. Grimmett. Space-time percolation. In *In and out of equilibrium. 2*, volume 60 of *Progr. Probab.*, pages 305–320. Birkhäuser, Basel, 2008.

[27] T. E. Harris. Nearest-neighbor Markov interaction processes on multidimensional lattices. *Advances in Math.*, 9:66–89, 1972.

[28] D. Ioffe. Stochastic geometry of classical and quantum Ising models. In *Methods of contemporary mathematical statistical physics*, volume 1970 of *Lecture Notes in Math.*, pages 87–127. Springer, Berlin, 2009.

[29] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[30] J. F. C. Kingman. Random discrete distributions. *J. Roy. Statist. Soc. Ser. B*, 37:1–15, 1975. With a discussion by S. J. Taylor, A. G. Hawkes, A. M. Walker, D. R. Cox, A. F. M. Smith, B. M. Hill, P. J. Burville, T. Leonard and a reply by the author.

[31] J. F. C. Kingman. *Mathematics of genetic diversity*, volume 34 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1980.

[32] J. F. C. Kingman. *Poisson processes*, volume 3 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York, 1993. Oxford Science Publications.

[33] J. M. Kosterlitz and D. J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C: Solid State Physics*, 6(7):1181, 1973.

[34] E. Mayer-Wolf, O. Zeitouni, and M. P. W. Zerner. Asymptotics of certain coagulation-fragmentation processes and invariant Poisson-Dirichlet measures. *Electron. J. Probab.*, 7:no. 8, 25 pp. (electronic), 2002.

[35] N. D. Mermin and H. Wagner. Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic heisenberg models. *Phys. Rev. Lett.*, 17(22):1133–1136, Nov 1966.

[36] B. Nachtergaele. Quantum spin systems after DLS 1978. In *Spectral theory and mathematical physics: a Festschrift in honor of Barry Simon's 60th birthday*, volume 76 of *Proc. Sympos. Pure Math.*, pages 47–68. Amer. Math. Soc., Providence, RI, 2007.

[37] J. Pitman. Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformation of an interval partition. *Combin. Probab. Comput.*, 11(5):501–514, 2002.

[38] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

[39] J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, 25(2):855–900, 1997.

[40] D. Ruelle. *Statistical mechanics: Rigorous results*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[41] O. Schramm. Compositions of random transpositions. *Israel J. Math.*, 147:221–243, 2005.

[42] O. Schramm. Compositions of random transpositions. *Israel J. Math.*, 147:221–243, 2005.

[43] B. Simon. *The statistical mechanics of lattice gases. Vol. I*. Princeton Series in Physics. Princeton University Press, Princeton, NJ, 1993.

[44] S. Tavaré. The birth process with immigration, and the genealogical structure of large populations. *J. Math. Biol.*, 25(2):161–168, 1987.

[45] B. Tóth. Improved lower bound on the thermodynamic pressure of the spin 1/2 Heisenberg ferromagnet. *Lett. Math. Phys.*, 28(1):75–84, 1993.

[46] B. Tóth. Reflection positivity, infrared bounds, continuous symmetry breaking. http://www.math.bme.hu/∼balint/oktatas/statisztikus_fizika/jegyzet/, 1996.

[47] N. Tsilevich. On the simplest split-merge operator on the infinite-dimensional simplex. *Arxiv preprint math/0106005*, 2001.

[48] N. V. Tsilevich. Stationary random partitions of a natural series. *Teor. Veroyatnost. i Primenen.*, 44(1):55–73, 1999.

Departments of Mathematics and Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom

E-mail address: `c.a.goldschmidt@warwick.ac.uk`

E-mail address: `daniel@ueltschi.org`

E-mail address: `peter.windridge@warwick.ac.uk`