Graph Homomorphisms and Phase Transitions

Graham R. Brightwell [*]
Department of Mathematics
London School of Economics
Houghton St.
London WC2A 2AE England

Peter Winkler[†]
Bell Laboratories 2C-379
Lucent Technologies
700 Mountain Ave.
Murray Hill, NJ 07974 U.S.A.

December 18, 1998

Abstract

We model physical systems with "hard constraints" by the space $\mathrm{Hom}(G,H)$ of homomorphisms from a locally finite graph G to a fixed finite constraint graph H. For any assignment λ of positive real activities to the nodes of H, there is at least one Gibbs measure on $\mathrm{Hom}(G,H)$; when G is infinite, there may be more than one.

When G is a regular tree, the simple, invariant Gibbs measures on $\mathrm{Hom}(G,H)$ correspond to node-weighted branching random walks on H. We show that such walks exist for every H and λ, and characterize those H which, by admitting more than one such construction, exhibit phase transition behavior.

1 Introduction

In what follows we are concerned with discrete models which exhibit what physicists sometimes call "hard constraints"—forbidden configurations, in which (for example) adjacent particles are not permitted to have certain pairs of spins. In the classical (ferromagnetic) Ising model, adjacent particles are discouraged from having opposing spins, since such opposition increases the energy of a configuration, making it a less likely state; this is a "soft" constraint. In contrast, the hard-core lattice gas model studied e.g. by Dobrushin [6] and Van den Berg and Steif [2] strictly forbids configurations in which adjacent sites are both occupied.

Our framework generalizes the gas model above, as well as the symmetric nearest-neighbor systems of Burton and Steif [3], but is still a very special case within the general study of Gibbs measures introduced by Dobrushin [6] and Lanford and Ruelle [11]. Our interest in this special case stems only partly from a desire to explore the notion of hard constraint; we also aspire to combine techniques from graph theory and combinatorics with the traditional methods of statistical mechanics, with the object of obtaining some relatively tangible phase transitions. Finally, we are intrigued by the attractive central role played by branching random walks when the sites are arrayed in a regular tree.

Suppose we are given a (possibly infinite) graph G whose nodes represent particles, or "sites". We model hard constraints by means of a fixed finite graph H, whose nodes may be thought of as

*Author's net address: g.r.brightwell@lse.ac.uk
†Author's net address: pw@lucent.com

different spins, or, as we prefer, different "colors". Adjacent sites of G may receive colors i and j only if i and j are adjacent nodes of H; in particular both sites may have color i only when in H there is a loop at node i. Thus a legal configuration is no more or less than a homomorphism from G to H: that is, a map φ from the sites of G to the nodes of H such that if u is adjacent to v in G (written $u \sim v$) then $\varphi(u) \sim \varphi(v)$ in H.

Among legal configurations, relative likelihood is determined by positive reals, called "activities", assigned to the colors. Thus, suppose two legal configurations differ only at site u; if the activity of the color of u in the first configuration is twice that of the color of u in the second, then the first configuration is twice as likely. When the activities are all equal and G is finite, we have the uniform distribution; in fact, whenever the activities have rational ratios, the distribution obtained is equivalent to the uniform distribution on some "blow-up" of the constraint graph H.

Typically, in statistical mechanics, G is an infinite, locally finite graph with many symmetries, such as a regular d-dimensional crystal structure. We shall be especially interested in the case where G is a regular tree, also known as a "Cayley tree" or "Bethe lattice". Here the connection between Gibbs measures and branching random walks facilitates analysis, and indeed quantitative analysis of the Ising model and various generalizations has been made on the Bethe lattice (see e.g. [15, 9, 8]).

When G is a regular tree, for any H and any set of activities on the nodes of H, certain well-behaved measures on the space of homomorphisms from G to H correspond precisely to branching, or "tree-indexed", random walks on H. Our goal in this work is a theorem which says roughly that, unless H is a complete multipartite graph, there are activities for H for which the branching random walk is uniquely determined, and others for which there are more than one.

2 Notation and Preliminaries

We will frequently abuse notation by confusing a graph with its set of nodes. There are two roles for graphs in our framework: the "lattice" G and the "constraint graph" H. Each is an undirected graph, possibly with some loops (edges with both ends at the same node), but no multiple edges.

The lattice G will always be countable and locally finite; for our purposes G will usually be the regular r-branching tree T_r, for some $r \geq 1$, or some finite part thereof.

The nodes of the lattice G will be called "sites" to distinguish them from the nodes of H. We will tend to use letters u, v, w, possibly primed or with subscripts, to denote sites of G. A finite set U of sites of G will be called a *patch* and, again, we will deliberately confuse U with the subgraph of G induced by U. The (exterior) *boundary* ∂U of U is the set of sites in $G \backslash U$ which are adjacent to at least one site of U.

The tree T_r is uniquely defined by being connected, $r+1$-regular and cycle-free; thus T_1 is the two-way infinite path and T_2 the complete binary tree. It is often useful to fix a root $w \in T_r$ and let the patch $T_r(k)$ be the set of all sites at distance at most k from w. (Here distance is graph distance, that is, number of edges in a shortest path. In T_r, of course, there is only one path between any two sites.)

In contrast to the lattice, the constraint graph H will be finite, with nodes $1, 2, \ldots, n$ represented by variables i, j or k. H will often have loops at some or all of its nodes; the loops are important. To avoid degeneracies we assume always that H is connected and has at least two nodes.

We denote by $\mathrm{Hom}(G, H)$ the graph whose nodes are homomorphisms from G to H, with $\alpha \sim \beta$ when α and β differ on at most one site of G. (There is no requirement that the two values at that site be adjacent in H.) We denote the image under α of u by $\alpha(u)$, so that for $\alpha \in \mathrm{Hom}(G, H)$, $u \sim v$ in G implies $\alpha(u) \sim \alpha(v)$ in H.

Example 2.1 If H is the graph K_n^{loop} on $\{1,\ldots,n\}$ containing all possible loops and edges then there are no constraints, and $\mathrm{Hom}(G,H)$ contains all maps from G to $\{1,\ldots,n\}$.

Example 2.2 If H is the loopless complete graph K_n, then $\mathrm{Hom}(G,H)$ is the set of all proper n-colorings of G, a coloring being proper if no two adjacent sites get the same color. Since trees are bipartite and connected, $\mathrm{Hom}(T_r,K_2)$ has just two elements.

Example 2.3 Let H be the graph on nodes $0,1$ with edge $\{0,1\}$ and a loop at node 0. Then for $\alpha \in \mathrm{Hom}(G,H)$, the set of sites u for which $\alpha(u)=1$ is an independent set in G, that is, a set containing no edges (or loops); and conversely. Thus $\mathrm{Hom}(G,H)$ is essentially the set of independent subsets of the graph G. This constraint graph is precisely the one in effect for the hard-core lattice gas model of [2].

We shall be studying probability distributions on $\mathrm{Hom}(G,H)$. When G is finite there is a very useful and natural class of such measures, sometimes called "multiplicative" measures, defined as follows by positive reals attached to the nodes of H.

A *set of activities* for a constraint graph H is a function $\lambda: H \to \mathbb{R}_+$ from the nodes of H to the positive reals, two such being regarded as equivalent if they differ by a constant factor. The value λ_i of λ at a node i is called its "activity", and will represent the relative probability of i as an image.

When G is finite and H and λ are given, we define the probability measure m_G on $\mathrm{Hom}(G,H)$, with respect to a random map φ, by

$$\Pr_{m_G}(\varphi=\alpha) := m_G(\{\alpha\}) := \frac{1}{Z}\prod_{u\in G}\lambda_{\alpha(u)}$$

where Z is the necessary normalizing constant,

$$Z := \sum_{\varphi\in\mathrm{Hom}(G,H)}\prod_{u\in G}\lambda_{\varphi(u)}\ .$$

If all the λ_i's are equal then m_G is, of course, the uniform distribution on $\mathrm{Hom}(G,H)$. In fact, the multiplicative distributions are not a vast generalization of the uniform distribution. If λ is a rational point we can clear denominators so that each λ_i is an integer, then blow up each node i of H to λ_i copies; looped nodes are replaced by fully looped complete graphs, unlooped nodes by independent sets. Then the uniform distribution on homomorphisms from G to the new, inflated H will behave just as m_G does for the old H and λ.

We now ask the question: is there a natural extension of the class of multiplicative measures to cover the case where G is infinite?

3 Gibbs Distributions

Suppose that H and λ are given, along with a lattice G. If U is a patch (finite subset of G) and $\varphi \in \mathrm{Hom}(G,H)$, we denote by $\varphi\restriction U$ the restriction of φ to U; thus $\varphi\restriction U \in \mathrm{Hom}(U,H)$. If A is an event of the form

$$A = \{\varphi \in \mathrm{Hom}(G,H):\ \varphi\restriction U \in F\}$$

for some patch U and some $F \subset \mathrm{Hom}(U,H)$, then we call A a "patch event". We equip $\mathrm{Hom}(G,H)$ with the σ-field (denoted by $\mathcal{F}$) generated by the patch events, and consider henceforth only measures μ on $\langle \mathrm{Hom}(G,H),\mathcal{F}\rangle$ such that $\mu(\mathrm{Hom}(G,H))=1$.

Let $U_1, U_2, \ldots$ be a nested sequence of patches whose union is G. We can define a compact topology on the space $\mathcal{M}$ of all probability measures on $\mathrm{Hom}(G,H)$ via the metric ρ defined as follows:

$$\rho(\mu,\nu) = \sum_{i=1}^{\infty} 2^{-i} \|\mu \upharpoonright U_i, \nu \upharpoonright U_i\|$$

where $\mu \upharpoonright U$ is the measure induced by μ on $\mathrm{Hom}(U,H)$, and $\|\cdot\|$ is the total variation metric between probability distributions (a.k.a. half the ℓ^1 distance). Of course this topology is not sensitive to the choice of the U_i's, but it is useful to fix them anyway.

Let us now consider what properties we might demand of a measure μ on $\mathrm{Hom}(G,H)$ in order that it be a candidate for generalizing the notion of a multiplicative measure. Of greatest importance by far is the following condition, variously called the "Gibbs" or "Markov" or "DLR" condition, which ensures that the conditional behavior of μ on patches is exactly what it should be. We fix and suppress reference to G, H and λ, and define $U^+ := U \cup \partial U$ for any patch U.

Definition 3.1 A measure μ on $\mathrm{Hom}(G,H)$ is a *Gibbs measure* if for any finite $U \subset G$, and almost every $\psi \in \mathrm{Hom}(G,H)$,

$$\Pr_{\mu}\left(\varphi \upharpoonright U = \psi \upharpoonright U \,\middle|\, \varphi \upharpoonright (G-U) = \psi \upharpoonright (G-U)\right) = \Pr_{m_{U^+}}\left(\varphi \upharpoonright U = \psi \upharpoonright U \,\middle|\, \varphi \upharpoonright \partial U = \psi \upharpoonright \partial U\right) .$$

In other words, the probability distribution of a random φ inside a patch U depends only on its value on the boundary of U, and is the same as if U and its boundary comprised all of the lattice. We remark here that, for some constraint graphs H (constituting an easily defined class known as the "cop-win", or "dismantlable", graphs) the Gibbs condition can be verified merely by checking patches U which consist only of a single site. The role played by these special constraint graphs is explored in [5].

Example 3.2 It is evident by taking $U = G$ that m_G itself is the only Gibbs measure on $\mathrm{Hom}(G,H)$ when G is finite.

Example 3.3 If, as in Example 2.1, $H = K_n^{\text{loop}}$, then the Gibbs condition guarantees that $\varphi \upharpoonright U$ is independent of $\varphi \upharpoonright \partial U$ (since that is the case for finite G). Thus the only Gibbs measure is the i.i.d. coloring obtained by choosing $\varphi(u)$ for each site u independently, with probabilities proportional to the λ_i's.

Example 3.4 This example shows that a Gibbs measure may look quite different from a finite multiplicative measure. Let $H = K_3$, with any set of activities, and $G = T_3$. Fix a root w for G. We color G by first choosing a color (say, uniformly) for w, then randomly coloring two of w's neighbors with one of the other colors and the other two neighbors with the third color. We continue coloring out from the root, always making sure that every site has two neighbors of one color and two of another.

Now let $\varphi \in \mathrm{Hom}(G,H)$. Any patch $U \subset G$ has at least one site u with at least three neighbors in ∂U; the color of u is then forced by $\varphi \upharpoonright \partial U$, and proceeding inductively, the whole coloring $\varphi \upharpoonright U$ is forced. Since the same thing happens in the finite space $\mathrm{Hom}(U^+,H)$, the Gibbs condition is satisfied regardless of λ. Note that even the measure which is concentrated at a single instance of such a 2-neighbor coloring is a Gibbs measure.

Gibbs measures do exist for any G, H and λ, and indeed in far more general circumstances. Following is a theorem of Dobrushin [6], specialized to our framework. We give a proof here, as it is quite easy, and the ideas and notation employed therein will be useful later.

Theorem 3.5 *Let H be a constraint graph with set of activities λ, and G any countable, locally finite graph for which* $\mathrm{Hom}(G,H)$ *is nonempty. Then there is a Gibbs measure on* $\mathrm{Hom}(G,H)$.

Proof. Let α be *any* fixed homomorphism from G to H. For each i, let μ_i be the discrete measure which is positive only on homomorphisms that agree with α on $G \setminus U_i$, and which satisfies

$$\Pr_{\mu_i}(\varphi \restriction U_i^+ = g) = \Pr_{m_{U_i^+}}\left(\varphi = g \,\middle|\, \varphi \restriction \partial U_i = g \restriction \partial U_i\right)$$

for every $g \in \mathrm{Hom}(U_i^+, H)$ that agrees with α on ∂U_i.

Since $\mathcal{M}$ is compact, it remains only to choose a convergent subsequence of $\mu_1, \mu_2, \ldots$ and to observe that its limit is indeed a Gibbs distribution. ∎

Definition 3.6 Let $\mathcal{A}(G)$ be the automorphism group of the lattice G, and for any subset $S \subset \mathrm{Hom}(G,H)$ and $\kappa \in \mathcal{A}(G)$ let $S \circ \kappa := \{\varphi \circ \kappa : \varphi \in S\}$. We say that a measure μ on $\mathrm{Hom}(G,H)$ is *invariant* if, for any μ-measurable $S \subset \mathrm{Hom}(G,H)$ and any $\kappa \in \mathcal{A}(G)$, we have $\mu(S \circ \kappa) = \mu(S)$.

It is immediate that for finite G any Gibbs measure is invariant, since multiplicative measures are invariantly defined. In the infinite case, however, especially when the lattice is highly symmetric (as in the case $G = T_r$), there will often be far more non-invariant than invariant Gibbs measures on $\mathrm{Hom}(G,H)$.

The third and last condition we will consider is stated here only for the case where the lattice G is a tree (roughly speaking, it takes the place of various "regularity" conditions found in the literature). For any site u in a tree T, let $d(u)$ be the number of edges incident with u and let $C_1(u), C_2(u), \ldots, C_{d(u)}(u)$ be the connected components of $T \setminus \{u\}$.

Definition 3.7 A Gibbs measure μ on $\mathrm{Hom}(T,H)$, where T is a tree, is *simple* if, for any site $u \in T$ and any node $i \in H$, the μ-distributions of

$$\varphi \restriction C_1(u), \ldots, \varphi \restriction C_{d(u)}(u)$$

are mutually independent given $\varphi(u) = i$.

Like the others, this condition is trivially satisfied when the lattice is finite. Hence, in a nutshell: for a finite tree T, every Gibbs measure is a multiplicative measure and is both simple and invariant. For measures on $\mathrm{Hom}(T_r,H)$, the three conditions add up to a requirement that the measure mimic very closely the behavior of multiplicative measures. Indeed there is often just one measure on $\mathrm{Hom}(T_r,H)$ which passes all three tests. In fact, we will see that for any H, r and λ there is at least one such measure, and frequently several.

The general problem of determining whether there is more than one Gibbs measure for a given set of "specifications" is a central problem of statistical mechanics, often called the *DLR problem* after Dobrushin, Lanford and Ruelle (see e.g. [13]). The term "phase transition" is often associated with the existence of multiple Gibbs measures, so that a statistical physicist might say that the subject of this article is phase transitions for hard-constraint systems on the Bethe lattice.

There is no universal agreement about the exact meaning of phase transition. We will say here that a lattice T_r and constraint graph H *exhibit a phase transition* if there are values of the activity vector λ to which correspond more than one simple, invariant Gibbs measure on $\mathrm{Hom}(T_r,H)$. Since in fact there are always activities for which only one such measure exists, it is unimportant whether the phase transition is regarded as something which turns one Gibbs measure into another for the same set of activities, or something which changes activities from the region of unique Gibbs measures to a region where there is more than one.

The importance of our "invariance" and "simplicity" conditions to our definition of phase transition is evident from the example of the next section.

4 The Hard-Core Lattice Gas Model

In the hard-core lattice gas model of statistical mechanics, sites are discrete positions in space and are either "occupied" (by a gas molecule) or "empty". The hard constraint is that particles may not be too close together; in the model of Dobrushin [6] and Van den Berg and Steif [2], this takes the form of forbidding occupation of two neighboring sites (typically in a cubical grid). Thus, from the graph-theoretic point of view, a configuration is legal iff its occupied sites are an independent set of the lattice. As we saw in Example 2.3, this means that the constraint graph H consists of a looped node 0 connected by an edge to an unlooped node 1. We will always take $\lambda_0 = 1$ so that the set of activities for H is specified by $\lambda := \lambda_1$.

Identifying $\mathrm{Hom}(G, H)$ with the set $\mathcal{I}$ of independent sets in G, we have that $\mathcal{I}$ inherits the uniform distribution when $\lambda = 1$ and G is finite. Fig. 1 illustrates (with sites mapped to node 1 of H blackened) a uniformly random independent set I in the tree $T_2(8)$, which we may think of as the patch of sites in the tree T_2 at distance at most 8 from the root u.

In fact, the recursive structure of T_r makes quantitative analysis of the hard-core gas model on $T_r(k)$ quite straightforward; see e.g. [15], or even [10, pp. 383–384] where the motivation is models of communications systems with call blocking. It is convenient to replace $T_r(k)$ by the tree $T'_r(k)$ whose root u has degree only r, then let $p_k(\lambda) := \Pr(u \in I)$, where the probability is taken with respect to the measure $m_{T'_r(k)}$.

We then have

$$p_0(\lambda) = \frac{\lambda}{1+\lambda}$$

and

$$p_{k+1}(\lambda) = \frac{\lambda(1-p_k(\lambda))^r}{1+\lambda(1-p_k(\lambda))^r} \ .$$

When $\lambda \le r^r/(r-1)^{r+1}$, $p_k(\lambda)$ oscillates and converges to the unique solution $p(\lambda)$ of

$$p(\lambda) = \frac{\lambda(1-p(\lambda))^r}{1+\lambda(1-p(\lambda))^r} \ .$$

Otherwise $p_k(\lambda)$ has two limit points, with

$$\lim_{k\to\infty} p_{2k+1}(\lambda) < p(\lambda) < \lim_{k\to\infty} p_{2k}(\lambda) \ .$$

In Fig. 1 one can see a tendency to have many black sites among the leaves (sites at distance 8 from the root u) of $T_2(8)$, where the constraint is less restrictive. Thus there are relatively few black sites at level 7, then more again at level 6 and so forth. When $r = 2$ the critical value for λ is $2^2/1^3 = 4$, which is greater than 1; hence this oscillation dies out as $k \to \infty$, for $\lambda = 1$, as indicated by the graph in Fig. 2.

When λ is at or below the threshold, there is only one Gibbs measure on $\mathrm{Hom}(T_r, H)$, and for any site v the *a priori* probability that v is occupied will be this same $p(\lambda)$. We will see later that this Gibbs measure can be obtained by taking a branching random walk on H.

Fig. 3 illustrates a sample from the $\lambda = 5$ distribution on $\mathrm{Hom}(T_2(8), H)$. As indicated in Fig. 4, the oscillation no longer dies out. The measure $m_{T_2(k)}$ is the same as the restriction to $T_2(k)$ of the measure $m_{T_2(k+1)}$ conditioned on none of the leaves being black, hence, letting α of the proof of Theorem 3.5 correspond to the empty independent set in T_2, we see that there are at least two Gibbs measures on $\mathrm{Hom}(T_2, H)$.

The behavior of $\Pr(u \in I)$ as a function of λ is illustrated in Fig. 5.

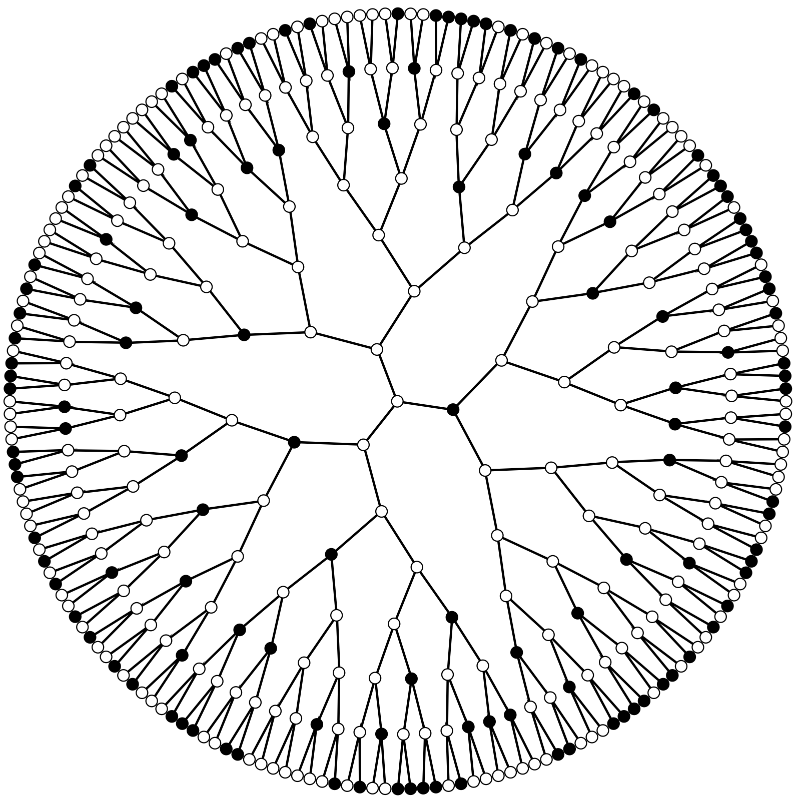

Figure 1: Uniform random independent set in a finite tree

Figure 2: Probability that the root is in I, uniform case

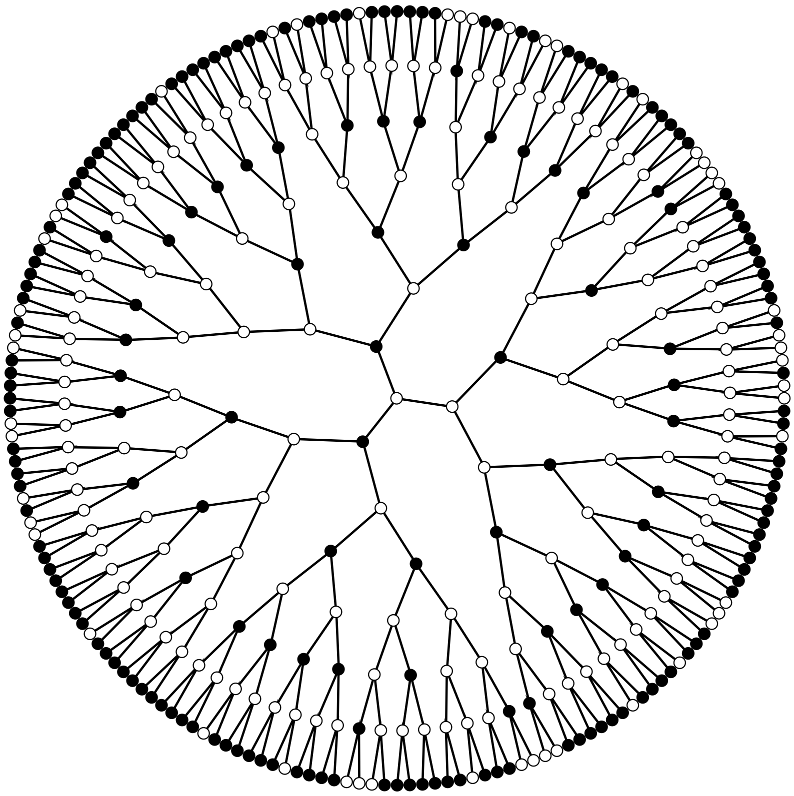

Figure 3: Random independent set in a finite tree, with $\lambda = 5$

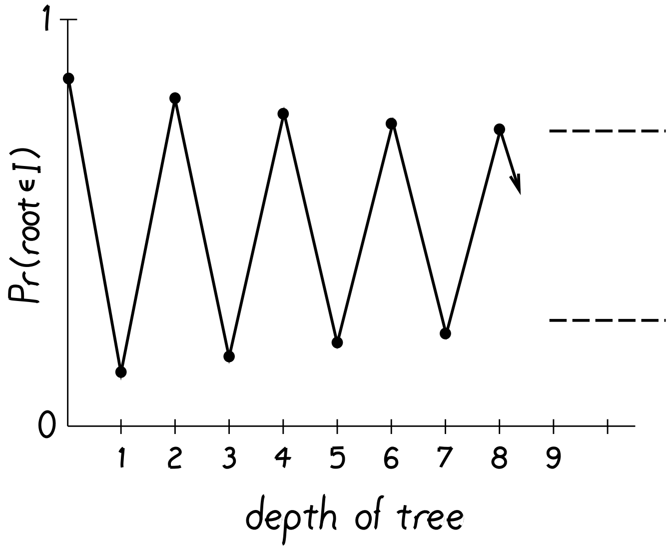

Figure 4: Probability that the root is in I, with $\lambda = 5$

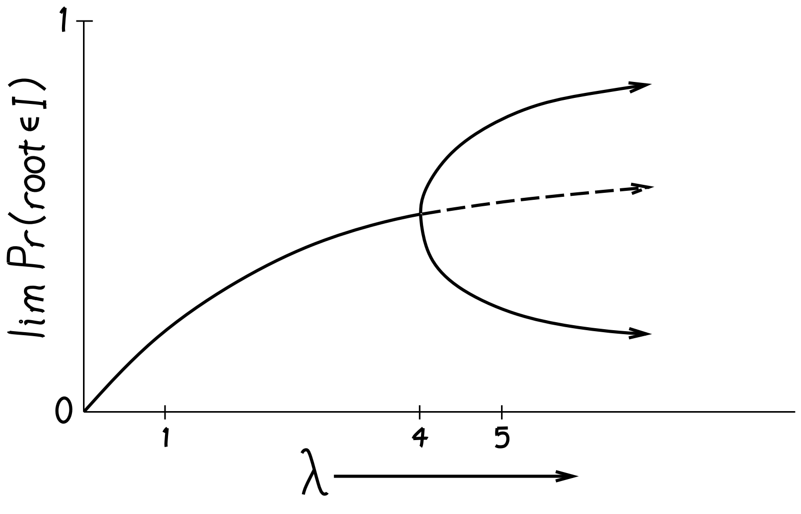

Figure 5: Upper and lower limits of $\Pr(u \in I)$ as a function of λ, for $k = 2$

In fact, the limit measures

$$\mu_{\text{even}} = \lim_{k\to\infty} m_{T_r(2k)}$$

and

$$\mu_{\text{odd}} = \lim_{k\to\infty} m_{T_r(2k+1)}$$

are both well-defined Gibbs measures on T_r. Above threshold, (i.e. when $\lambda > r^r/(r-1)^{r+1}$) these measures differ and neither is invariant; μ_{even} favors sites at even distance from the root as elements of the random independent set I, while μ_{odd} favors the odd sites.

An invariant Gibbs measure may be constructed by taking an equal mix $\frac{1}{2}\mu_{\text{even}} + \frac{1}{2}\mu_{\text{odd}}$. However, this measure will not be simple when $\mu_{\text{even}} \neq \mu_{\text{odd}}$, since the event of some particular even point on one component of $T_r \setminus \{u\}$ being in I would then be positively correlated to similar events on other components even when the membership of u in I is fixed.

It is a consequence of our main result (below) that, regardless of the values of r and λ, there is precisely one simple, invariant Gibbs measure for the hard-core model on T_r, and that measure is given by a node-weighted branching random walk on H.

5 Node-weighted Random Walks

In this section we make an excursion into what at first may appear to be quite a different subject, namely random walks on graphs. Since we are interested in ergodic (irreducible, aperiodic) random walks, it is useful to call a graph *ergodic* if it is connected and non-bipartite (equivalently, if it is connected and contains a loop or odd cycle).

Let H be a fixed ergodic graph on nodes $\{1, 2, \ldots, n\}$ with adjacency matrix A, and let $P = \{p_{ij}\}$ be the transition matrix of a Markov chain M whose states are the nodes of H. We say that M is *H-based* if P is positive exactly where A is; in other words, if $i \to j$ is an allowable transition of M exactly when $\{i, j\}$ is an edge of H. In that case M will indeed be an ergodic Markov chain, with a unique stationary distribution π.

A natural subclass of the H-based Markov chains is obtained by assigning positive real weights w_{ij} to the edges of H, then putting

$$p_{ij} = \frac{w_{ij}}{\sum_{k \sim i} w_{ik}}$$

for $i \sim j$, and $p_{ij} = 0$ otherwise. This produces a *reversible* Markov chain ($\pi_i p_{ij} = \pi_j p_{ji}$) and indeed every reversible H-based Markov chain arises in this manner.

It turns out that of greatest interest here is an even more restrictive class, obtained by attaching real weights w_i to each *node* of H, then putting

$$p_{ij} = \frac{w_j}{\sum_{k \sim i} w_k} \ .$$

for $i \sim j$, and $p_{ij} = 0$ otherwise. Such a Markov chain will be called a *node-weighted random walk* on H with weights $w = \langle w_1, \ldots, w_n \rangle$.

Setting $w_{ij} = w_i w_j$ shows that node-weighted random walks are also reversible, but it is easily verified that the converse fails; in fact the class of node-weighted random walks enjoys only $n-1$ degrees of freedom while the edge-weights live in $\mathbb{R}^{m-1}$, where m is the number of edges of H. In both cases, two sets of weights give the same Markov chain if and only if they differ by a constant factor; this relies on H being ergodic. Henceforth, we regard two weight vectors differing by a constant multiple as equivalent.

Given a set w of weights for the nodes of H, it is convenient to define

$$z_i := \sum_{j \sim i} w_j$$

so that $p_{ij} = w_j / z_i$ for any $i \sim j$. We then have for any $j \in H$,

$$\sum_{i \sim j} p_{ij} z_i w_i = \sum_{i \sim j} \frac{w_j}{z_i} z_i w_i = w_j \cdot \sum_{i \sim j} w_i = z_j w_j \ ,$$

so that $\pi_i = c z_i w_i$ for some constant c.

The stationary distributions for H-based Markov chains live in the hyperplane

$$\{s = \langle s_1, \ldots, s_n \rangle : \ \sum s_i = 1\}$$

of dimension $n-1$, as do the node-weighted random walks, so it is natural to ask how the stationary distributions of node-weighted random walks compare to those of the more general H-based Markov chains. The answer is as nice as it could be.

Theorem 5.1 *Let H be an ergodic graph and M an H-based Markov chain. Then there is exactly one node-weighted random walk on H with the same stationary distribution as M.*

Proof. We begin by supposing that there are two different node-weighted random walks on H, given by weights w and w', with the same stationary distribution π; we wish to derive a contradiction.

We denote by $f : \mathbb{R}^n \to \mathbb{R}^n$ the map whose coordinates are given by $f_i(w) := z_i w_i$. Let A be the set of nodes j of H for which w_j / w'_j is minimal; A cannot be all of H else w and w' would be equivalent.

Since H is not bipartite there is an edge e contained either in A or $H \setminus A$; and since H is connected, we can include e in a path $\mathcal{P}$ which begins in A and ends in $H \setminus A$. In $\mathcal{P}$ we can find a node j_2 which is simultaneously adjacent to some $j_1 \in A$ and some $j_3 \notin A$. (Since e might be a loop, these nodes are not necessarily all distinct.)

Let $p_{i,j}$ denote the transition probability from i to j in H weighted by w, and similarly p'_{ij} for the weights w'. Then $p_{i,j} \leq p'_{i,j}$ for any $j \in A$, and $p_{j_2,j_1} < p'_{j_2,j_1}$ since j_2 has a neighbor outside A. It follows that

$$\sum_{j \sim j_1} p_{j,j_1} \pi_j < \sum_{j \sim j_1} p'_{j,j_1} \pi_j$$

but this is impossible because both sides are equal to π_{j_1}.

To show that f is onto, we need to consider which stationary distributions are achievable by H-based Markov chains. Let P be the (unbounded) polyhedron in $\mathbb{R}^n$ consisting of all s such that $s_i \geq 0$ for each node i of G, and $s_I \leq s_{N(I)}$ for each independent set I of G, where $s_A := \sum_{i \in A} s_i$ for any set of nodes A, and $N(A) := \{j : \ j \sim i \text{ for some } i \in A\}$.

Let π be the stationary distribution π of an H-based Markov chain, and c any positive real. We claim $c\pi$ lies in the interior $\hat{P}$ of P. Of course, $\pi > 0$ for each node i. For any independent set $I \subset H$, $\pi_I \leq \pi_{N(I)}$ since in a random walk, every visit to I is followed immediately by a visit to $N(I)$. Equality is not possible because the ergodicity of H implies that $N(I)$ can also be visited from $H \setminus I$.

Let us now choose any $s \in \hat{P}$, say at ℓ^1-distance ε from the boundary $P \setminus \hat{P}$. It suffices to show that there is a set of weights w for H with $f(w) = s$.

It is convenient at this time to introduce the notion of a "push". Suppose w is a set of weights for which

$$|f(w), s|_1 = \sum_{i=1}^{n} |z_i w_i - s_i| > 0$$

and j is a node for which $z_j w_j > s_j$. It turns out that w_j can be lowered without increasing $|f(w), s|_1$.

To see this, let $\delta > 0$ be tiny and put $w'_j = w_j - \delta$, $w'_i = w_i$ for $i \neq j$. Then if j has no loop,

$$|f_j(w') - s_j| = |f_j(w) - s_j| - \delta z_j \ ;$$

otherwise

$$\begin{aligned} |f_j(w') - s_j| &= |f_j(w) - s_j| - \delta(z_j + w_j - \delta) \\ &< |f_j(w) - s_j| - \delta z_j \ . \end{aligned}$$

In any case,

$$\sum_{i \in N(j)} |f_i(w') - s_i| \leq \sum_{i \in N(j)} (|f_i(w) - s_i| + w_i \delta) = \left(\sum_{i \in N(j)} |f_i(w) - s_i| \right) + \delta z_j$$

with equality only when $f_i(w) \leq s_i$ for every $i \in N(j)$. Thus

$$|f(w'), s|_1 \leq |f(w), s|_1$$

and the inequality is strict if j is looped or if *any* neighbor i of j has $z_i w_i > s_i$.

Now we are ready to proceed with the proof of the theorem. Let R be the range of f and let x be a nearest point in the closure R^{Cl} of R to s (again in the ℓ^1 metric).

Among the weight sets w for which $f(w) = x$, we choose one which maximizes the size of the set

$$\Delta := \{i \in H : \ f_i(w) \neq s_i\} \ .$$

Then $\Delta \neq \emptyset$, of course, since $f(w) \neq s$, and in fact we claim Δ is all of H. Were this not so, then on account of the connectedness of H, there would be an edge $\{j,k\}$ of H with $j \in \Delta$ but $k \notin \Delta$. Supposing that $f_j(w) > s_j$, we push w_j down by a tiny quantity δ to get a weight set w' with $\Delta' = \Delta \cup N(j)$ and $|f(w'),s|_1 \leq |f(w),s|_1$. Since $k \in N(j) \setminus \Delta$, this contradicts the choice of w. A similar argument applies if $f_j(w) < s_j$.

Now that we have $\Delta = H$ we may again use the ergodicity of H to conclude that there are two neighboring (but not necessarily distinct) nodes j and k with $f_j(w) > s_j$ and $f_k(w) > s_k$, or with $f_j(w) < s_j$ and $f_k(w) < s_k$; we may assume the former. Then a push down on w_j decreases $|f(w),s|_1$, this time contradicting the choice of x.

We are reduced now to the case where x, the nearest point in R^{Cl} to s, is not in R itself but is instead the ℓ^1-limit of a sequence of points in R. Since f is continuous we may (by taking subsequences n times) select a sequence $w(1), w(2), \ldots$ of weight vectors so that $f(w(t)) \to x$ and each $w(t)_i$ tends to a limit u_i in the compact interval $[0,\infty]$ as $t \to \infty$.

Let B (for "big") be the set of nodes i for which $u_i = \infty$ and $S := \{i: \ u_i = 0\}$. Clearly B and S cannot both be empty since then $w(t)$ has a legal weight vector w as its limit, and $f(w) = x \in R$.

We claim next that $x \notin \hat{P}$ (and therefore x lies on the boundary ∂P of P). This will certainly follow if $B = \emptyset$, because then $f_i(w(t)) \to 0$ for any $i \in S$.

We now observe that $N(B) \subset S$, because if there are $i \in B$, $j \notin S$ with $i \sim j$ then $f_i(w(t)) \to \infty$ (an impossibility since x is only finitely far from s). In particular, B is an independent set. Moreover, every node $i \in S$ does have a neighbor in B else $f_i(w(t)) \to 0$ as before; thus the neighborhood of B is exactly S.

We then have

$$\begin{aligned}
&\sum_{i \in N(B)} f_i(w(t)) - \sum_{j \in B} f_j(w(t)) \\
&= \sum_{i \in S} z(t)_i w(t)_i - \sum_{j \in B} z(t)_j w(t)_j \\
&= \sum_{\{i,j:\ i \in S,\ j \sim i\}} w(t)_i w(t)_j - \sum_{\{i,j:\ i \in S,\ j \in B,\ j \sim i\}} w(t)_i w(t)_j \\
&= \sum_{\{i,j:\ i \in S,\ j \notin B,\ j \sim i\}} w(t)_i w(t)_j \to 0 \ ,
\end{aligned}$$

thus $\lim f(w(t))$ fails one of the constraints and must lie in ∂P as claimed.

We have shown that for any $s \in \hat{P}$, either $s \in R$ or the nearest point of R^{Cl} to s is in ∂P. We claim the latter case cannot in fact arise.

Assume the contrary and let $s \in \hat{P} \setminus R$, with nearest point $x \in R^{\mathrm{Cl}}$, $x \in \partial P$. Let π be any point of R, e.g. the stationary distribution for the simple (uniform weights) random walk on H. Let $a = |s, \partial P|_1 \leq |s,x|_1$ and $b = |\pi, \partial P|_1$, and consider the line segment

$$L := \{\rho s + (1-\rho)\pi : \ 0 \leq \rho \leq 1\}$$

from s to π. We claim that $|L, \partial P|_1 \geq \min\{a,b\}$.

To see this suppose that u is a point in $\mathbb{R}^n$ with $|u,v|_1 = c < \min\{a,b\}$ for some $v = \rho s + (1-\rho)\pi \in L$; then

$$|(s+u-v), s|_1 = |(\pi + u - v), \pi|_1 = c$$

so $s+u-v$ and $\pi+u-v$ are both points of $\hat{P}$. But

$$u = \rho(s+u-v) + (1-\rho)(\pi+u-v)$$

so $y \in \hat{P}$ by convexity of P, proving the claim.

Finally, let y and y' be two points on L with $|y, y'|_1 < \min\{a, b\}$ and $y \in R$, $y' \notin R$. Then supposedly the nearest point of R^{Cl} to y' is on ∂P at distance at least $\min\{a, b\}$ from y' but this is nonsense, since y is closer. This contradiction completes the proof of the theorem. ∎

A constraint graph H may fail to be ergodic, by virtue of being bipartite (with, say, parts U and V); then a set of node weights for H has only $n-2$ degrees of freedom since multiplying the weights on the nodes in U by a constant has no effect on the random walk. The random walk still has a unique stationary distribution π satisfying $\pi_j = \sum_{i \sim j} \pi_i p_{ij}$ for all j. The stationary distributions of H-based Markov chains now live in the subspace

$$\sum_{i \in U} \pi_i = \sum_{i \in V} \pi_i = \frac{1}{2}$$

of $\mathbb{R}^n$ but the essentials of the proof of Theorem 5.1 go through, yielding the following addendum.

Theorem 5.2 *Let H be a connected, bipartite graph, and M an H-based Markov chain. Then there is exactly one node-weighted random walk on H with the same stationary distribution as M.*

6 Branching Random Walks

The object of this section is to show that a node-weighted random walk on H yields a simple, invariant Gibbs measure on $\mathrm{Hom}(G, H)$ when G is the infinite r-branching tree T_r; and conversely, that every simple, invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$ can be obtained in this way.

In statistical mechanics T_r is sometimes called the r-branching "Cayley tree" or "Bethe lattice". Baxter [1] makes a subtle distinction between the Cayley tree and the Bethe lattice, by declaring that statistical properties of the latter are to be determined by averaging only over those sites in a finite r-ary tree which are far from the leaves. This is necessary because for $r > 1$ the size of the boundary of any finite subset of T_r is at least a constant fraction of the size of the subset. Although the tree T_r is a planar graph it is for some purposes infinite in dimension, since (for $r > 1$) the number of sites at distance k from a given site, namely $(r+1)r^{k-1}$, exceeds any power of k.

Let S_r be the set of finite strings of integers in $\{1, \ldots, r\}$, except that 0 is permitted as the first entry in a string; the empty string $\emptyset$ is also allowed. An *r-branching* random walk on H is a function σ from S_r to H defined recursively as follows: $\sigma(\emptyset)$ is a node of H drawn from the distribution π (so that $\Pr(i)$ is proportional to $z_i w_i$), and if $\sigma(s) = i$ is defined and s' is an extension of s by one symbol, then $\sigma(s')$ is chosen randomly (and independently) from the neighbors of i with probabilities proportional to their weights.

We may imagine that a branching random walk is like an ordinary, stationary random walk on H except that the usual token is replaced by an amoeba which divides into r independent amoebas at each step. The amoeba starts at a node chosen from the stationary distribution, and divides $r+1$ ways at the first step only. In the case $r = 1$ the branching random walk begins (say) at the origin of the number line and walks both left and right from there; this produces an ordinary stationary random walk on account of reversibility.

Since S_r has the structure of the tree T_r it is evident that the r-branching random walk on H, together with a labeling of T_r by S_r, produces a random "coloring" φ of T_r by H. A moment's reflection will show that $\varphi \in \mathrm{Hom}(T_r, H)$, and thus the branching random walk produces a measure μ on $\mathrm{Hom}(T_r, H)$. Our interest in branching random walks is explained by the following result.

Theorem 6.1 *Let H be a fixed connected constraint graph with node-weights w and let r be a positive integer; then the measure μ induced on $\mathrm{Hom}(T_r, H)$ by the r-branching node-weighted random walk on H is a simple, invariant Gibbs measure, for some set λ of activities on H. Conversely, if H, r and λ are given, then every simple, invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$ is given by the r-branching random walk on H with nodes weighted by some w.*

Proof. Let H, r and w be given, and let us first check that μ does not depend on the labeling of T_r by S_r, and is therefore an invariant measure. On account of the symmetry of the random walk definition and the fact that T_r is connected, we need only check that for two neighboring sites u and v of T_r, μ is the same whether u is chosen as root or v is.

But, either way we may choose $\varphi(u)$ and $\varphi(v)$ as the first two "colors" and the rest of the procedure is the same; so it suffices to check that for any nodes i and j of H, the probability that $\varphi(u) = i$ and $\varphi(v) = j$ is the same with either root choice. But these two probabilities are

$$\pi_i p_{ij} = z_i w_i \frac{w_j}{z_i} = w_i w_j = z_j w_j \frac{w_i}{z_j} = \pi_j p_{ji}$$

as desired. In fact, we are using only reversibility here; any edge-weighted random walk would also yield an invariant (and simple) measure. It is the Gibbs condition which fails when the branching random walk is not of the node-weighted variety.

To show that μ is simple is, indeed, simple: if we condition on $\varphi(u) = i$ then, using invariance to put the root at u, the independence of φ on the $r+1$ components of $T_r \setminus \{u\}$ is evident from the definition of the branching random walk. (In fact any branching Markov chain, even a non-reversible one, yields a simple measure.)

Let us now define a mapping $\Phi : \mathbb{R}_+^{n-1} \to \mathbb{R}_+^{n-1}$ by

$$\Phi(w)_i := w_i / z_i^r \ .$$

We claim that μ is a Gibbs measure on $\mathrm{Hom}(T_r, H)$ with set of activities $\lambda = \Phi(w)$. Let U be any finite set of sites in T_r, with exterior boundary ∂U. On account of invariance of labeling, we may assume that the root x (corresponding to the empty string in S_r) does not lie in $U^+ := U \cup \partial U$.

Let $g \in \mathrm{Hom}(U^+, H)$; we want to show that the probability that a branching random walk φ matches g on U, given that it matches on ∂U, is the same as the corresponding conditional probability for the measure m_{U^+} defined from λ.

Let T be the subtree of T_r induced by U^+ and the root x; for any $f \in \mathrm{Hom}(T, H)$,

$$\Pr(\varphi \restriction T = f) = \pi_{f(x)} \cdot \prod_{u \to v} p_{f(u), f(v)}$$
$$= z_{f(x)} w_{f(x)} \prod_{u \to v} \frac{w_{f(v)}}{z_{f(u)}}$$

where $u \to v$ means that v succeeds u in the tree construction. The factors $z_{f(u)}$ corresponding to sites u in U each occur as denominator r times in the above expression, since each site in U has all of its r successors in T; and of course each $w_{f(u)}$ occurs once as a numerator as well. It follows that if we compare $\Pr(\varphi \restriction T = f)$ with $\Pr(\varphi \restriction T = f')$, where f' differs from f only on U, then the value of the first is proportional to

$$\prod_{u \in U} \frac{w_{f(u)}}{z_{f(u)}^r} = \prod_{u \in U} \lambda_{f(u)}$$

which means that μ coincides with the finite measure, as desired.

Now let us assume that μ is a simple, invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$ with set of activities λ, with the intent of showing that μ arises from a node-weighted branching random walk on H.

We start by constructing a μ-random φ, site by site. Fix a labeling of T_r by S_r and choose $\varphi(x) = \varphi(\emptyset)$ from the *a priori* distribution σ of colors of the root x (and therefore, by invariance, of any other site). We next color the neighbor $y = \langle 0 \rangle$ of x according to the conditional distribution matrix $P = \{p_{ij}\}$ given by

$$p_{ij} := \Pr\left(\varphi(y) = j \,\middle|\, \varphi(x) = i\right) \ ;$$

again, by invariance of μ, P is the same for any pair of neighboring sites. It follows that $\sigma = \sigma \cdot P$, and moreover that P is the transition matrix of a reversible Markov chain, since the roles of x and y can be interchanged.

Next we proceed to the rest of the neighbors of x, then to the sites at distance 2 from x (i.e. strings of length 2 in S_r), etc., coloring each conditionally according to all sites so far colored.

We claim, however, that the distribution of the colors $\varphi(s)$ of the string $s = \langle s_1, s_2, \ldots, s_k \rangle$ depends only on $\varphi(\langle s_1, \ldots, s_{k-1} \rangle)$; this is so because μ is simple and all sites so far colored are in components of $T_r \setminus \{\langle s_1, \ldots, s_{k-1} \rangle\}$ other than the component containing s. Thus the value of $\varphi(u)$ is given by P for *every* site $u \neq x$, and it follows that μ arises from an r-branching Markov chain on H, starting at distribution σ.

Evidently for any (not necessarily distinct) nodes i, j of H, there will be pairs (u, v) of adjacent sites with $\varphi(u) = i$ and $\varphi(v) = j$ if and only if $i \sim j$ in H. Hence P is an H-based Markov chain, and there is a unique distribution π satisfying $\pi \cdot P = \pi$; thus $\sigma = \pi$.

It remains only to show that P is a node-weighted random walk, and it turns out that a special case of the Gibbs condition for one-site patches suffices. Let j and j' be nodes of H which have a common neighbor i, and suppose that all of the neighbors of the root x have color i. Such a configuration will occur with positive probability and according to the Gibbs condition for $U = \{x\}$,

$$\frac{\Pr(\varphi(x) = j')}{\Pr(\varphi(x) = j)} = \frac{\lambda_{j'}}{\lambda_j}$$

but

$$\Pr(\varphi(x) = j) = \frac{\pi_j p_{ji}^{r+1}}{\sum_{k \sim i} \pi_k p_{ki}^{r+1}}$$

and similarly for j', so

$$\frac{\Pr(\varphi(x) = j')}{\Pr(\varphi(x) = j)} = \frac{\pi_{j'} p_{j'i}^{r+1}}{\pi_j p_{ji}^{r+1}} \ .$$

Thus the ratio

$$\frac{p_{j'i}}{p_{ji}} = \left(\frac{\lambda_{j'} \pi_j}{\lambda_j \pi_{j'}}\right)^{\frac{1}{r+1}}$$

is independent of i.

Since P is reversible we have $p_{ij} = \pi_j p_{ji} / \pi_i$, hence

$$\frac{p_{ij'}}{p_{ij}} = \frac{\pi_{j'} p_{j'i}}{\pi_j p_{ji}}$$

is also independent of i, and it follows that P is a node-weighted random walk on H. ∎

If φ is a coloring in $\mathrm{Hom}(T_r, H)$ and there is a vector $\rho = \langle \rho_1, \ldots, \rho_n \rangle$ such that for any root of T_r and any node i of H

$$\lim_{k\to\infty} \frac{|\{x \in T_r(k) : \varphi(x) = i\}|}{|T_r(k)|} = \rho_i$$

then we say φ has *density* ρ. If a measure μ on $\mathrm{Hom}(T_r, H)$ has the property that with μ-probability 1 every coloring φ has the same density ρ, then we say μ itself has density ρ.

Since the stationary distribution of a random walk is unaffected by branching, the measure induced by one of our node-weighted branching random walks on an ergodic graph has density π; we may thus put together Theorems 6.1 and 5.1 to get the following.

Corollary 6.2 *If H is ergodic then any simple, invariant Gibbs measure on* $\mathrm{Hom}(T_r, H)$ *has a density ρ and moreover is uniquely determined by ρ.*

Suppose now that H, r and λ are given; is there a simple, invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$? Theorem 6.1 reduces the problem to the technical one of showing that for any $\lambda \in \mathbb{R}_+^n$, there is a set of weights $w \in \mathbb{R}_+^n$ such that

$$\lambda_i = w_i / z_i^r$$

for each node i of H.

In the one-dimensional case ($r = 1$) there is always exactly one simple invariant Gibbs measure. We will deal with this special case separately, in the next section.

Theorem 6.3 *For every $r \geq 2$, every constraint graph H and every set λ of activities for H, there is a node-weighted branching random walk on T_r which induces a simple, invariant Gibbs measure on* $\mathrm{Hom}(T_r, H)$.

Proof. For the purposes of this proof, we wish to assume that both the weights and the activities sum to 1 on H. We may do this, since if we have a weight function w (summing to 1) with $\lambda_i = cw_i/z_i^r$ for all i, then multiplying all weights by $c^{1/(r-1)}$ gives us a new weight function inducing λ.

Let S be the simplex $\{x \in \mathbb{R}^n : x \geq 0, \sum_i x_i = 1\}$, and let $\hat{S}$ be its ($(n-1)$-dimensional) interior. For $w \in \hat{S}$, we define $(g(w))_i = w_i / C z_i^r$, where $C = \sum_i w_i / z_i^r$. Thus $g(w) \in \hat{S}$ for every $w \in \hat{S}$; we need to prove that g is a surjection.

For $t > 0$, we define the function $g_t : \hat{S} \to \hat{S}$ by

$$g_t(w) = \frac{tw + (w_1 w_2 \cdots w_n) g(w)}{t + w_1 w_2 \cdots w_n}\ .$$

Note that, for $w \in \hat{S}$, $g_t(w)$ is indeed in $\hat{S}$. Also, g_t is continuous on $\hat{S}$.

For w on the boundary of S (i.e., with at least one entry 0), we define $g_t(w) = w$. We claim that this constitutes a continuous extension of g_t to the boundary of S. Indeed, if w is on the boundary of S with, say, $w_i = 0$, and v is in $\hat{S}$, then

$$g_t(w) - g_t(v) = \frac{w(t + v_1 v_2 \cdots v_n) - tv - (v_1 v_2 \cdots v_n) g(v)}{t + v_1 v_2 \cdots v_n}\ .$$

Therefore

$$\|g_t(w) - g_t(v)\|_1 \leq \frac{t\|w - v\|_1 + 2 v_1 v_2 \cdots v_n}{t} \leq \|w - v\|_1 + \frac{2v_i}{t},$$

since $\|w - g(v)\|_1 \leq 2$. If $v^{(k)}$ is a sequence of points in $\hat{S}$ converging to w, then $v_i^{(k)}$ converges to 0, and so $g_t(v^{(k)})$ converges to $g_t(w)$.

Thus each g_t ($t > 0$) is a continuous function from the simplex S to itself, equal to the identity on the boundary of S. If $x \in S$ were not in the range of g_t, then the map obtained by composing g_t with projection from x to ∂S would be a retraction of S to ∂S, contradicting Brouwer's Fixed-Point Theorem (see, for instance, Dugundji [7], Chapter XVI.2). Thus each g_t is a surjection.

Now, given any $\lambda \in \hat{S}$, and any sequence t_k tending to 0, choose points $w^{(k)} \in \hat{S}$ such that $g_{t_k}(w^{(k)}) = \lambda$. By taking a subsequence if necessary, we may assume that the sequence $w^{(k)}$ tends to a limit $w \in S$. We aim to show that $w \in \hat{S}$, and that $g(w) = \lambda$. For each k, set $z_i^{(k)} = \sum_{j\sim i} w_j^{(k)}$, as normal, and $C^{(k)} = \sum_i w_i^{(k)}/(z_i^{(k)})^r$, so that $(g(w^{(k)}))_i = w_i^{(k)}/(z_i^{(k)})^r C^{(k)}$.

Let us suppose first that some w_i is equal to 0. Of course, not all the w_j are 0, so we may assume that i is adjacent to a node j with $w_j \neq 0$. Now λ_i/λ_j is a fixed positive quantity, which is equal to $(g_{t_k}(w^{(k)}))_i/(g_{t_k}(w^{(k)}))_j$ for every k. This ratio is given by

$$\frac{t_k w_i^{(k)} + w_1^{(k)} w_2^{(k)} \cdots w_n^{(k)} \frac{w_i^{(k)}}{\left(z_i^{(k)}\right)^r C^{(k)}}}{t_k w_j^{(k)} + w_1^{(k)} w_2^{(k)} \cdots w_n^{(k)} \frac{w_j^{(k)}}{\left(z_j^{(k)}\right)^r C^{(k)}}}$$

$$= \frac{w_i^{(k)}}{w_j^{(k)}} \left[\frac{t_k C^{(k)} + w_1^{(k)} w_2^{(k)} \cdots w_n^{(k)} / \left(z_i^{(k)}\right)^r}{t_k C^{(k)} + w_1^{(k)} w_2^{(k)} \cdots w_n^{(k)} / \left(z_j^{(k)}\right)^r} \right].$$

The first fraction above tends to zero as k tends to infinity, so the second must tend to infinity. This means that $z_j^{(k)}/z_i^{(k)}$ tends to infinity as $k \to \infty$. But $z_i^{(k)}$ is at least $w_j^{(k)}$, which is bounded away from 0 as $k \to \infty$, while $z_j^{(k)}$ is at most 1. This is a contradiction, so we conclude that $w \in \hat{S}$.

Now we consider $g(w^{(k)}) - \lambda = g(w^{(k)}) - g_{t_k}(w^{(k)})$. This is equal to

$$\frac{t_k (g(w^{(k)}) - w^{(k)})}{t_k + w_1^{(k)} w_2^{(k)} \cdots w_n^{(k)}}.$$

As $k \to \infty$, the numerator tends to 0, while the denominator tends to $w_1 w_2 \cdots w_n > 0$, so we have that $g(w^{(k)}) \to \lambda$ as $k \to \infty$. But, by continuity of g, $g(w^{(k)}) \to g(w)$, so $g(w) = \lambda$, as claimed. ∎

The fact that for each r, H and λ there is a simple invariant Gibbs measure is deducible as a special case of a Theorem 4.2 of Zachary [18]. Zachary associates his Gibbs measures to "entrance laws" instead of branching random walks, but the general ideas are similar.

7 Two Cases of Uniqueness

In this section we show that for any H there is a unique simple invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$ when $r = 1$, or, for arbitrary r and sufficiently unbalanced sets of activities λ.

The well-known dictum that "you don't get phase transitions in one dimension" is false if taken in sufficient generality, but true as expected for our hard-constraint model.

Theorem 7.1 *For any connected constraint graph H and any set of activities λ, there is a unique simple invariant Gibbs measure on* $\mathrm{Hom}(T_1, H)$.

Proof. We show that there is precisely one set of weights (up to equivalence) for H which yield the activity set λ. Let $A = \{a_{ij}\}$ be the adjacency matrix of H, and let $B = \{b_{ij}\}$ where $b_{ij} := \lambda_i a_{ij}$.

It follows from the Perron-Frobenius Theorem (see, e.g., [14]) that B has a positive left eigenvector w which is unique up to a constant multiple; say $w \cdot B = cw$. But then, for any node i of H,

$$\frac{w_i}{z_i} = \frac{w_i}{(w \cdot A)_i} = \frac{\lambda_i w_i}{(w \cdot B)_i} = \frac{1}{c}\lambda_i$$

so w is a valid set of weights. Reversing the argument shows that any w for which w_i/z_i is proportional to λ_i is a left eigenvector for B.

In the case where H is bipartite with parts U and V, and w and w' are two equivalent sets of weights which do *not* differ by a constant factor, we seem to have proved that different sets of activities emerge. Of course this cannot be the case; in the bipartite case the activities also lose a degree of freedom. Multiplying the activities of the nodes in U by a constant factor does not change the specification. ∎

We now move to the case $r \geq 2$, with the intent of showing that there is always a region of activities where the simple invariant Gibbs measure is unique.

Theorem 7.2 *For any r and H there is a set of activities λ for which there is only one simple invariant Gibbs measure on* $\mathrm{Hom}(T_r, H)$.

Proof. Set $n = |V(H)|$, and let M be a large constant. Choose one node of H, a looped node if there is one, and denote this node as 1. Now, for each node i of H, set $\lambda_i = M^{-d(i,1)}$, where d denotes the graph distance in H. Our aim is to show that, for sufficiently large M (depending on n and r), this set of activities corresponds to just one r-branching random walk, and hence just one simple invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$.

Suppose then that w is a weight-vector for H such that $\lambda_i = w_i/z_i^r$ for every node i of H. Our first aim is to show that every node other than 1 has a neighbor of much higher weight. For any constant $R > 1$ such that $M > R^{r+1}n^r$, suppose that node $i \neq 1$ has the property that $w_j \leq Rw_i$ for all $j \sim i$ in H. Then $z_i \leq nRw_i$, whereas $z_j \geq w_i$ for all $j \sim i$. Now there is some vertex $k \sim i$ with $d(k,1) = d(i,1) - 1$, so $\lambda_k/\lambda_i = M$. But

$$\frac{\lambda_k}{\lambda_i} = \frac{w_k}{w_i}\frac{z_i^r}{z_k^r} \leq R\frac{n^r R^r w_i^r}{w_i^r} = R^{r+1}n^r < M,$$

a contradiction.

So every node i other than 1 has a neighbor of weight greater than Rw_i. In particular, node 1 has the largest weight and $w_1 > Rw_i$ for every $i \neq 1$. Hence also $w_i/z_i < 1/R$ for every $i \neq 1$.

Consider now a neighbor i of 1. We have that $\lambda_i = M^{-1}$, and that z_i lies between w_1 and $w_1\alpha$, where $\alpha = 1 + (n-1)/R$. Thus $w_i = \lambda_i z_i^r$ lies between w_1^r/M and $w_1^r\alpha^r/M$.

Turning back to node 1, we have to consider two cases depending on whether or not it is looped.

If node 1 is looped, then all we will need is that $w_1 < z_1 \leq w_1\alpha$, where, as before, $\alpha = 1 + (n-1)/R$.

If node 1 is unlooped, let ρ denote its degree in H, so that

$$\rho w_1^r/M \leq z_1 \leq \rho w_1^r \alpha^r/M.$$

Again we have $w_1 = z_1^r$, which yields

$$\left(\frac{M}{\rho\alpha^r}\right)^{r/(r^2-1)} \leq w_1 \leq \left(\frac{M}{\rho}\right)^{r/(r^2-1)}.$$

Combining these with the estimates for w_i, $i \sim 1$, gives us (after a short calculation, and being somewhat crude with powers of α) the following inequalities:

$$\left(\frac{M}{\rho}\right)^{1/(r+1)} \alpha^{-r-2} < \frac{w_1}{z_1} < \left(\frac{M}{\rho}\right)^{1/(r+1)} \alpha^{r+1},$$

$$\left(\frac{M}{\rho}\right)^{r/(r+1)} \frac{1}{M}\alpha^{-2} < \frac{w_i}{z_i} < \left(\frac{M}{\rho}\right)^{r/(r+1)} \frac{1}{M}\alpha^{r+2},$$

for every node $i \sim 1$.

We will make use of the following lemma, which is a straightforward consequence of the multi-dimensional Mean Value Theorem (see, for instance, [12], Theorem 12.2).

Lemma 7.3 *Let E be an open convex set in $\mathbb{R}^n$, and let $\Phi : E \to \mathbb{R}^m$ be a function whose components $\Phi_1, \ldots, \Phi_m$ are differentiable on E. Suppose that, for every n-tuple of points $x^{(1)}, \ldots, x^{(n)}$ in E, the determinant of the matrix A given by*

$$a_{ij} = \frac{\partial \Phi_i}{\partial x_j}\Big|_{x^{(i)}}$$

is non-zero. Then Φ is 1-1 on E.

Note that the condition of the lemma is stronger than merely saying that the Jacobian is non-zero throughout the region E, which would not be strong enough to guarantee that Φ is 1-1 on E.

Our aim is of course to apply this lemma to the function Φ given by $\Phi_i(w) = w_i/z_i^r$, with some suitable region E containing all possible solutions w with $\Phi(w) = \lambda$.

For this function Φ, we have:

$$\frac{\partial \Phi_i}{\partial w_j} = \begin{cases} 0 & i \neq j, i \not\sim j \\ -r w_i z_i^{-r-1} & i \neq j, i \sim j \\ z_i^{-r} & i = j, i \text{ not looped} \\ z_i^{-r} - r w_i z_i^{-r-1} & i = j, i \text{ looped.} \end{cases}$$

Now consider any n-tuple of weight vectors $w^{(1)}, \ldots, w^{(n)}$, and let $z^{(i)}$ be the corresponding z-vectors. The matrix A of the lemma, evaluated at these vectors, is thus given by

$$A = \begin{pmatrix} (z_1^{(1)})^{-r} & & 0 \\ & \ddots & \\ 0 & & (z_n^{(n)})^{-r} \end{pmatrix} B$$

where

$$B = I - r \begin{pmatrix} w_1^{(1)}/z_1^{(1)} & & 0 \\ & \ddots & \\ 0 & & w_n^{(n)}/z_n^{(n)} \end{pmatrix} A_H,$$

where A_H is the adjacency matrix of the graph H.

The diagonal matrix in the expression for A is non-singular whenever all the $z_i^{(i)}$ are non-zero, so we need only consider the matrix B.

We first consider the (easier) case where the node 1 of H is looped. Set $\varepsilon = 1/(2r^2 n!)$, and let E be the region of $\mathbb{R}^n$ where: $w_i > 0$ for all i, $w_i < \varepsilon z_i$ for all $i \neq 1$, and $z_1 > w_1 > (2/r)z_1$. This is clearly an open convex region. Our earlier calculations showed that all solutions w of $\Phi(w) = \lambda$ lie in this region, provided $R > 1/\varepsilon$ and $\alpha = [1 + (n-1)/R] < r/2$.

For any set of points $w^{(i)}$ inside E, the entries of the matrix B satisfy:

$$\begin{array}{ll} b_{11} < 1 - r\left(\frac{2}{r}\right) = -1 & \\ b_{ii} > 1 - r\varepsilon > 1/2 & i \neq 1 \\ |b_{1i}| < r & i \neq 1 \\ |b_{ji}| < r\varepsilon & 1 \neq j \neq i. \end{array}$$

The expansion of the determinant of B contains the diagonal term $\prod b_{ii}$, and $n!-1$ terms of modulus at most $2r^2\varepsilon \prod b_{ii}$. (Choosing a term from the first row instead of b_{11} multiplies the product by a factor of at most r; all other terms involve at least one other off-diagonal entry, which is smaller than the diagonal entries by a factor of at least $2r\varepsilon$.)

Thus, since $2r^2\varepsilon(n!-1) < 1$, the determinant of B is non-zero. By Lemma 7.3, this means that, for M sufficiently large, the function Φ is 1-1 on E, and hence there is at most one solution to $\Phi(w) = \lambda$ in E and none elsewhere. This completes the proof in the case where node 1 is looped (and indeed in the case where H has a looped node).

Now we move to the case where node 1, along with all the other nodes, is unlooped. In this case, we need the more precise estimates for the ratios w_i/z_i derived earlier. In this case, we set $\varepsilon = 1/(2r^3 n!)$, and define E as the region of $\mathbb{R}^n$ such that: $w_i > 0$ for all i, $w_i < \varepsilon z_i$ for all $i \neq 1$,

$$\left(\frac{M}{\rho}\right)^{1/(r+1)} (1+\varepsilon)^{-1} < \frac{w_1}{z_1} < \left(\frac{M}{\rho}\right)^{1/(r+1)} (1+\varepsilon),$$

$$\left(\frac{M}{\rho}\right)^{r/(r+1)} \frac{1}{M}(1+\varepsilon)^{-1} < \frac{w_i}{z_i} < \left(\frac{M}{\rho}\right)^{r/(r+1)} \frac{1}{M}(1+\varepsilon),$$

for every neighbor i of 1.

Evidently E is a convex open region, and all solutions of $\Phi(w) = \lambda$ lie inside E, provided $R > 1/\varepsilon$ and $\alpha^{r+2} < 1+\varepsilon$, both of which can be guaranteed by choosing M sufficiently large. Again, we choose an n-tuple of weight-vectors $w^{(1)}, \ldots, w^{(n)}$ inside E, and evaluate the matrix B at this n-tuple.

The matrix B has 1s on the leading diagonal (since H has no loops). All other entries, except those on the first row, are at most $r\varepsilon$ in modulus. For each $i \sim 1$, the entry b_{1i} is within a multiplicative factor $(1+\varepsilon)$ of $-r(M/\rho)^{1/(r+1)}$, and the entry b_{i1} is within a multiplicative factor $(1+\varepsilon)$ of $-r(M/\rho)^{r/(r+1)}/M$. Other entries in the first row and column are 0.

The expansion of the determinant of B has one term $\prod b_{ii} = 1$, and ρ terms $b_{1i}b_{i1}\prod_{j\neq 1,i} b_{jj}$, $i \sim 1$, which are all within a multiplicative factor $(1+\varepsilon)^2$ of r^2/ρ, and all occur with signature -1 in the expansion. All other terms have absolute value at most $r\varepsilon \max(1, (1+\varepsilon)^2 r^2/\rho) < 2r^3\varepsilon$. The determinant of B is thus at most $1 - r^2/(1+\varepsilon)^2 + 2r^3\varepsilon n! < 0$. As before, Lemma 7.3 now tells us that, for M sufficiently large, there is at most one solution to $\Phi(w) = \lambda$ inside E, which completes the proof. ∎

We are finally ready for the main question: for which constraint graphs can we get *multiple* simple invariant Gibbs measures?

8 Sterile and Fertile Graphs

We have seen that for a fixed constraint graph H and positive activity set λ, each simple, invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$ corresponds precisely to a branching random walk on H with certain positive weights w on the nodes. We know also that for any $r \geq 1$ and any ergodic H, there are activity sets for which the branching random walk is unique; and that when $r = 1$, it is always unique.

It is thus natural to ask: when $r > 1$, for which H is there a set of activities yielding multiple branching random walks? The answer, which turns out not to depend on r, is given by the theorem below.

Theorem 8.1 *Fix $r > 1$ and let H be any constraint graph. Suppose that H satisfies the following two conditions:*

(a) *Every looped node of H is adjacent to all other nodes of H;*

(b) *With its loops deleted, H is a complete multipartite graph.*

Then for every positive activity set on H, there is a unique invariant Gibbs measure on the space $\mathrm{Hom}(T_r, H)$.

If H fails either condition (a) *or condition* (b) *then there is a set of activities λ on H for which* $\mathrm{Hom}(T_r, H)$ *has at least two simple, invariant Gibbs measures, and therefore λ can be obtained by more than one branching random walk.*

A graph H which satisfies (a) and (b) above will be called *sterile*, otherwise we will say that H is *fertile*.

Proof. We proceed in three stages. First we show that for sterile graphs H, the map Φ described in the proof of Theorem 6.1, taking a set of weights w to its corresponding activity set λ, is injective. Then we show that any fertile H must admit at least one of a certain list of seven small graphs as an induced subgraph. Finally, we show that for any $r > 1$, a graph H containing one of the seven small induced subgraphs can be provided with two sets of weights giving the same set of positive activities. (In view of Theorems 5.1 and 5.2, these sets of weights give different stationary probabilities, hence very different Gibbs measures.)

Recall that the map Φ is defined by

$$\Phi(w)_i := \lambda_i := w_i / z_i^r$$

and suppose that H is a graph for which Φ is injective; let H' be obtained from H by splitting an unlooped node j of H into two new unlooped nodes, j_1 and j_2, each with the same neighborhood previously enjoyed by j. Then $z'_{j_1} = z'_{j_2}$ for any set of weights w' on H', so if u' and v' are distinct sets of weights for H' with the same image λ' under Φ, then we must have $u'_k \neq v'_k$ for some node $k \neq j$ of H. But then, if we define u on H by $u_i = u'_i$ for $i \neq j$ and $u_j = u_{j_1} + u_{j_2}$, and similarly for v and λ, we have $\Phi(u) = \Phi(v) = \lambda$, a contradiction. Hence Φ is also injective for H'.

We are thus reduced to showing injectivity of Φ when H is a complete graph on nodes $\{1, 2, \ldots, n\}$ where (say) nodes 1 through m are looped. Putting $\lambda = \Phi(w)$ and

$$Z := \sum_{i=1}^{n} w_i$$

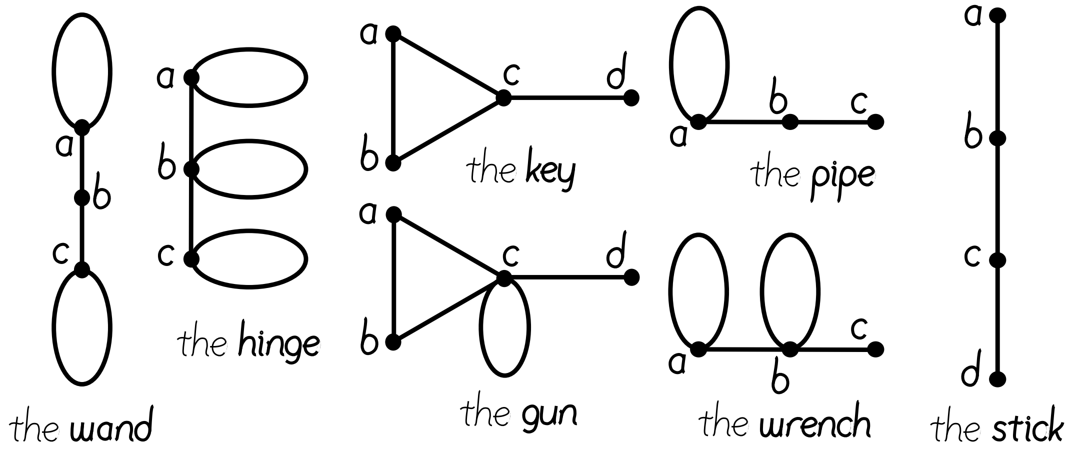

Figure 6: The seven minimal fertile graphs

we have

$$Z = \left(\frac{w_i}{\lambda_i}\right)^{\frac{1}{r}} = \left(\frac{w_j}{\lambda_j}\right)^{\frac{1}{r}} + w_j$$

for every i, j with $1 \leq i \leq m < j \leq n$. Since each of these expressions for Z is strictly increasing in its corresponding variable w_i or w_j, we must have $w < w'$ or $w > w'$ (say, the former) if w and w' are distinct vectors with $\Phi(w) = \Phi(w') = \lambda$.

But this cannot be, since if i minimizes w'_i/w_i then $z'_i/z_i \geq w'_i/w_i$ hence $\Phi(w')_i < \Phi(w)_i$. Thus the first part of the proof is complete.

For the second, we begin by providing the seven minimal fertile graphs, illustrated in Fig. 6. Each of them has three or four nodes, with node set, edges and loops as follows:

the *stick*: $\{a, b, c, d\}$; $\{a, b\}, \{b, c\}, \{c, d\}$; no loops

the *pipe*: $\{a, b, c\}$; $\{a, b\}, \{b, c\}$; loop at a

the *wrench*: $\{a, b, c\}$; $\{a, b\}, \{b, c\}$; loops at a and b

the *wand*: $\{a, b, c\}$; $\{a, b\}, \{b, c\}$; loops at a and c

the *hinge*: $\{a, b, c\}$; $\{a, b\}, \{b, c\}$; loops at a, b and c

the *key*: $\{a, b, c, d\}$; $\{a, b\}, \{a, c\}, \{b, c\}, \{c, d\}$; no loops

the *gun*: $\{a, b, c, d\}$; $\{a, b\}, \{a, c\}, \{b, c\}, \{c, d\}$; loop at c

The following observation will be used twice.

Lemma 8.2 *Let J be the graph on three nodes with one edge and no loops. Then any (finite) unlooped graph H which does not induce a copy of J is complete multipartite.*

Proof. Probably the easiest way to see this is to note that the complement of J is a path P_3 on three nodes, and every component of a graph with no induced P_3 must have diameter 1. Hence the complement of any H which does not induce J must be a disjoint union of complete graphs. ∎

Now let us suppose that H is a connected graph which fails condition (a); then H has a looped node, which we may call a, at distance 2 from some node c. Then a, c and any node b adjacent to both induce either a pipe, a wand, a wrench or a hinge depending upon which of b and c are looped.

Suppose next that H satisfies condition (a) but not (b), and has some looped nodes, say 1 through m, which must therefore constitute a clique. The remaining, unlooped nodes of H cannot constitute a multipartite graph else H would satisfy (b), hence by Lemma 8.2, they induce a copy of the graph J. But J plus any of the looped nodes of H constitute an induced gun.

Finally, suppose H has no loops and is not multipartite. Since H induces a copy of J we can find an edge $\{a, b\}$ and a node d of distance precisely 2 from the set $\{a, b\}$; we may assume there is a node c adjacent to both b and d. Then $\{a, b, c, d\}$ induces either a stick or a key.

Now we must show that any graph H containing one of the seven small graphs admits more than one simple invariant Gibbs measure. Our approach will be to show first that, for each of the seven graphs, there is a set of activities that is induced by more than one node-weighted random walk. Then we show that this property is preserved when moving to a larger graph.

We break the seven graphs into three groups; we deal with all the "asymmetric" graphs—the wrench, pipe, key and gun—together, then the wand and hinge together, and finally the stick separately.

The asymmetric graphs

For all the four asymmetric graphs, we define two different sets of node-weights, depending on two parameters s and γ, as well as on the branching factor r of the tree, and then show that certain values of γ give rise to node-weights that induce the same activity sets—and hence the same Gibbs measure—in both cases. This then gives a family of examples, parameterized by s, for each of the four graphs.

For given real $s > 1$ and $\gamma > 0$, and integer $r \geq 2$, set:

$$w_1 = s^r, w_2 = 1, w_3 = \gamma,$$

$$w_1' = 1, w_2' = \beta \equiv \frac{s^r + 1}{s} - 1, w_3' = \gamma\beta^r.$$

Observe here that, if $s > 1$, then $s(\beta - 1) = (s^r - s^2) + (s^2 - 2s - 1) > 0$, so $\beta > 1$. For the wrench and the pipe, we attach weight w_1 (w_1') to a, w_2 (w_2') to b and w_3 (w_3') to c. For the key and the gun, we attach weight w_1 (w_1') to both a and b, w_2 (w_2') to c and w_3 (w_3') to d.

As usual, we have $\lambda_i = w_i/z_i^r$, and $\lambda_i' = w_i'/(z_i')^r$, where z_i (z_i') is the sum of w_j (w_j') over all neighbors j of i. We wish to show that there are conditions ensuring that $\lambda_i = \lambda_i'$ for all i.

In all cases, all bar one of the conditions is automatically satisfied by our specification of the w_i and w_i'. To be precise, at node a, we have

$$\lambda_a = \frac{w_1}{(w_1 + w_2)^r} = \frac{s^r}{(1 + s^r)^r} = \frac{w_1'}{(w_1' + w_2')^r} = \lambda_a'$$

(of course, this also means that the condition is satisfied at node b in the key and gun), and at c for the wrench and pipe we have

$$\lambda_c = w_3/w_2^r = \gamma = w_3'/(w_3')^r = \lambda_c',$$

and similarly for d in the key and gun.

For the remaining equation, we need to consider each of the four graphs separately.

For the wrench, we require that

$$\lambda_b = \frac{w_2}{(w_1 + w_2 + w_3)^r} = \frac{w_2'}{(w_1' + w_2' + w_3')^r},$$

which translates to

$$1 + \beta + \gamma\beta^r = \beta^{1/r}(s^r + 1 + \gamma),$$

or

$$\gamma = \frac{\beta^{1/r}(s^r + 1) - (\beta + 1)}{\beta^r - \beta^{1/r}}.$$

We need to check that this specifies a positive value of γ; for this we recall that s and β are both greater than 1, and note in addition that $\beta^{1/r}(s^r + 1) = (\beta + 1)s\beta^{1/r} > \beta + 1$.

Therefore, for each $s > 1$, we have a value of $\gamma = \gamma(s)$ giving two different sets of weights, but the same set of activities. These two sets of weights thus give rise to two different Gibbs measures.

The three other asymmetric graphs are handled in exactly the same way; we obtain the values for γ as follows.

$$\text{Pipe}: \ \gamma = \frac{\beta^{1/r}s^r - 1}{\beta^r - \beta^{1/r}},$$

$$\text{Key}: \ \gamma = \frac{2(\beta^{1/r}s^r - 1)}{\beta^r - \beta^{1/r}},$$

$$\text{Gun}: \ \gamma = \frac{\beta^{1/r}(2s^r + 1) - (\beta + 2)}{\beta^r - \beta^{1/r}}.$$

In each case, the specified value of γ is positive, so we obtain two different node-weighted random walks producing the same set of activities.

The wand and the hinge

In both of these cases, we give weight 1 to node a, weight $s > 1$ to node c, and weight β to node b. If we can do this so that $\lambda_a = \lambda_c$, then exchanging the weights on a and c will give us a different set of node-weights and the same set of activities. It is evidently immaterial whether b has a loop, so the two graphs—wand and hinge—can be treated identically.

The condition we have to satisfy is

$$\lambda_a = \frac{1}{(1 + \beta)^r} = \frac{s}{(s + \beta)^r} = \lambda_c.$$

This can be rearranged to give $s + \beta = s^{1/r}(1 + \beta)$, or

$$\beta = \frac{s - s^{1/r}}{s^{1/r} - 1}.$$

We note that this specifies a positive value of β. Thus, for any $s > 1$, this asymmetric system of weights does indeed produce a symmetric set of activities, as required.

Incidentally, the choice $s = 2^r$ gives $\beta = 2^r - 2$; a particularly simple example of the phenomenon of multiple Gibbs measures.

The stick

Again, it is enough to exhibit an asymmetric system of node-weights inducing a symmetric set of activities, and again we actually produce a one-parameter family of such node-weightings.

Set $w_b = 1$, $w_c = s > 1$, $w_a = \alpha$, and $w_d = \alpha s^r$. This choice already ensures that $\lambda_d = w_d/w_c^r = \alpha = w_a/w_b^r = \lambda_a$, so it remains to choose α to ensure that $\lambda_b = \lambda_c$. The condition is that

$$\frac{1}{(\alpha + s)^r} = \frac{s}{(1 + \alpha s^r)^r} \ ,$$

or $1 + \alpha s^r = s^{1/r}(\alpha + s)$. This is true when

$$\alpha = \frac{s^{1+1/r} - 1}{s^r - s^{1/r}} ,$$

which is positive since $s > 1$.

Since (unlike the other six graphs) the stick is bipartite, we need to observe that our allegedly asymmetrical system of weights is not accidentally equivalent to its reverse. It suffices to note that

$$\frac{w_c}{w_a} = \frac{s}{\alpha} \neq \frac{s^{-r}}{\alpha} = \frac{w_b}{w_d} .$$

The final part of the proof is to show that, if J is a graph with a set of activities $\lambda^{(J)}$ induced by more than one node-weighted random walk, and H contains J as an induced subgraph, then H also has a set of activities λ induced by more than one node-weighted random walk. The basic idea is very simple; set $\lambda = \lambda^{(J)}$ on J, and λ very small elsewhere; then any system of weights on J inducing $\lambda^{(J)}$ should be close to a system inducing λ on H. Unfortunately we are not allowed to take either weights or activities equal to 0, and it is necessary that we obtain exactly the same λ for two different systems of weights, so we need to show that, if the weights w give rise to a set of activities near the target λ, then a nearby system of weights w' hits λ exactly. In essence, we wish to show that the map Φ is locally a bijection. Of course, we are in the process of proving exactly that the map is not globally 1-1, so we will need some condition to ensure that we are not near a "fold" in Φ. In short, we should expect a certain amount of complication at this point, and the proof below does indeed become somewhat messy.

We start with a routine inequality.

Lemma 8.3 *For real numbers a and b, and positive integer s,*
(i) $|a^s - b^s| \le |a - b| s (\max\{|a|, |b|\})^{s-1}$.
(ii) $|a^s - b^s - s b^{s-1}(a - b)| \le |a - b|^2 \frac{s(s-1)}{2} (\max\{|a|, |b|\})^{s-2}$.

Proof. Part (i) follows on writing out

$$a^s - b^s = (a - b)(a^{s-1} + a^{s-2}b + \cdots + b^{s-1}).$$

Similarly, part (ii) is immediate from

$$a^s - b^s - s b^{s-1}(a - b) = (a - b)^2 (a^{s-2} + 2a^{s-3}b + \cdots + (s - 1)b^{s-2}).$$

∎

We next give a somewhat technical lemma that is closely related to what we need. The general theme is that, if λ and μ are two sets of activities, close to each other, and λ is induced by a weight function w, then μ is also induced by a nearby weight function v. For the purposes of this lemma, we will allow the various activities and weights to take the value 0 (or even, in the case of the function v we seek, negative values).

In what follows, $\|.\|$ will always denote the ℓ_1 norm, i.e., the sum of the absolute values of the coordinates; and the norm of a matrix B is defined as $\max \|Bx\|/\|x\|$, which turns out to be the maximum of the norms of the columns of B.

Lemma 8.4 *Let H be a graph with adjacency matrix A. Suppose that we are given an integer $r \geq 2$, two functions λ and μ from H to the non-negative reals, a positive real number ε, and a function w from H to the non-negative reals such that $w_i = \lambda_i (Aw)_i^r$. (Here we treat w as a vector, so that Aw is just the function we have been denoting z.) Let D be the diagonal matrix whose ith diagonal entry is $\mu_i r(Aw)_i^{r-1}$. Suppose that the matrix $I - DA$ is invertible, with inverse B, and set*

$$K = 3r(r-1)\|\mu\| \, \|B\|^2 \|w\|^r (\|w\| + 2\varepsilon\|B\| \, \|w\|^r)^{r-2}.$$

Finally suppose that $2K\varepsilon < 1$, and that $\|\mu - \lambda\| < \varepsilon$. Then there is a function v from H to $\mathbb{R}$ with $v_i = \mu_i (Av)_i^r$, such that $\|v - w\| \leq 2\|w\|^k \|B\|\varepsilon$.

Proof. For any vectors x and y, we apply the previous lemma to obtain that

$$|(Ax)_i^r - (Ay)_i^r - r(Ay)_i^r (A(x-y))_i| \leq |(A(x-y))_i|^2 \frac{r(r-1)}{2} (\max\{(Ax)_i, (Ay_i)\})^{r-2}.$$

Now note that $|(Ax)_i| \leq \|x\|$, for any i, since A is an adjacency matrix; thus we have

$$|(Ax)_i^r - (Ay)_i^r - r(Ay)_i^r (A(x-y))_i| \leq \|x - y\|^2 \frac{r(r-1)}{2} (\max\{\|x\|, \|y\|\})^{r-2}.$$

Similarly the first part of the previous lemma tells us that, for any vector y,

$$|(Ay)_i^{r-1} - (Aw)_i^{r-1}| \leq \|y - w\|(r-1)(\max\{\|y\|, \|w\|\})^{r-2}.$$

Now define a function $g : \mathbb{R}^n \to \mathbb{R}^n$ by $(g(x))_i = \mu_i (Ax)_i^r$, so we are looking for a non-negative vector v with $g(v) = v$. (A first attempt to prove this result might be to iterate the function g, starting from w, and hope to converge; in fact we have to proceed a little more delicately.)

For x and y arbitrary vectors, and w the specific vector we are given, we have:

$$\|g(x) - g(y) - DA(x-y)\| = \sum_{i \in H} \mu_i \left| (Ax)_i^r - (Ay)_i^r - r(Aw)_i^{r-1}(A(x-y))_i \right|$$

$$\leq \sum_{i \in H} \mu_i \Big[\left| (Ax)_i^r - (Ay)_i^r - r(Ay)_i^{r-1}(A(x-y))_i \right| + r|(A(x-y))_i| \left| (Ay)_i^{r-1} - (Aw)_i^{r-1} \right| \Big]$$

$$\leq \sum_{i \in H} \mu_i \left[\|x-y\|^2 \frac{r(r-1)}{2} (\max\{\|x\|, \|y\|\})^{r-2} + r\|x-y\| \, \|y-w\|(r-1)(\max\{\|y\|, \|w\|\})^{r-2} \right]$$

$$\leq \|\mu\| \frac{r(r-1)}{2} (\max\{\|x\|, \|y\|, \|w\|\})^{r-2} \left(\|x-y\|^2 + 2\|x-y\| \, \|y-w\| \right).$$

We are now in a position to describe the iterative procedure that gets us from w to our required v. Set $v^{(0)} = w$ and, for $j \geq 1$,

$$v^{(j)} = v^{(j-1)} + B\left(g(v^{(j-1)}) - v^{(j-1)} \right).$$

We claim that the following hold for all j:

(i) $\|g(v^{(j)}) - v^{(j)}\| \leq \|w\|^r \varepsilon 2^{-j}$,
(ii) $\|v^{(j)} - v^{(j-1)}\| \leq \|B\| \, \|w\|^r \varepsilon 2^{1-j}$,
(iii) $\|v^{(j)} - w\| \leq 2\varepsilon \|B\| \, \|w\|^r$.

Once we have verified (i)–(iii) for every j, we will be done; condition (ii) implies that the $v^{(j)}$ tend to a limit v; condition (i), plus the continuity of g, tells us that $g(v) = v$; condition (iii) tells us that v is close to w, as in the statement.

For $j = 0$, we have that

$$|(g(w) - w)_i| = |\mu_i - \lambda_i|\, |(Aw)_i|^r \leq |\mu_i - \lambda_i|\, \|w\|^r,$$

so $\|g(w) - w\| \leq \|\mu - \lambda\|\, \|w\|^r < \varepsilon \|w\|^r$. This means that (i) is true for $j = 0$ as, trivially, is (iii).

Supposing that (i) holds for $j-1$, we immediately have that

$$\|v^{(j)} - v^{(j-1)}\| = \|B\left(g(v^{(j-1)}) - v^{(j-1)}\right)\| \leq \|B\|\, \|g(v^{(j-1)}) - v^{(j-1)}\| \leq \|B\| \leq \|w\|^{r-1}\varepsilon 2^{1-j},$$

so (ii) holds for j.

Supposing that (ii) holds up to j, we have

$$\|v^{(j)} - w\| \leq \sum_{l=1}^{j} \|v^{(l)} - v^{(l-1)}\|$$

$$\leq \|B\|\, \|w\|^r \varepsilon (1 + 2^{-1} + \cdots + 2^{1-j}) \leq 2\|B\|\, \|w\|^r \varepsilon,$$

i.e., (iii) holds for j.

Finally, supposing that (ii) and (iii) hold for j, and (i) holds for $j-1$, we have

$$g(v^{(j)}) - v^{(j)} = \left[g(v^{(j)}) - g(v^{(j-1)}) - DA(v^{(j)} - v^{(j-1)})\right]$$

$$-(I - DA)(v^{(j)} - v^{(j-1)}) + [g(v^{(j-1)}) - v^{(j-1)}].$$

But we have

$$(I - DA)(v^{(j)} - v^{(j-1)}) = B^{-1}B(g(v^{(j-1)}) - v^{(j-1)}),$$

by definition of $v^{(j)}$ and of B. So the final two terms above cancel, and we have

$$\|g(v^{(j)}) - v^{(j)}\| = \|g(v^{(j)}) - g(v^{(j-1)}) - DA(v^{(j)} - v^{(j-1)})\|$$

$$\leq \|\mu\| \frac{r(r-1)}{2} (\max\{\|v^{(j)}\|, \|v^{(j-1)}\|, \|w\|\})^{r-2} \left(\|v^{(j)} - v^{(j-1)}\|^2 + 2\|v^{(j)} - v^{(j-1)}\|\, \|v^{(j-1)} - w\|\right)$$

$$< \|\mu\| \frac{r(r-1)}{2} (\|w\| + 2\varepsilon\|B\|\, \|w\|^r)^{r-2}\, \|B\|\, \|w\|^r \varepsilon 2^{1-j} \cdot 6\varepsilon \|B\|\, \|w\|^r$$

$$= \varepsilon^2 2^{1-j} \|w\|^r K < \varepsilon 2^{-j} \|w\|^r.$$

This verifies (i) for j, completing the inductive proof of (i)–(iii), and thus the proof of the lemma. ∎

We continue to home in on what we need with the next lemma.

Lemma 8.5 *Suppose that J, with adjacency matrix A_J, is an induced subgraph of a graph H, with adjacency matrix A_H. Suppose that we are given an integer $r \geq 2$, a positive real number δ, a set of activities λ on J, and a weight vector w on J inducing λ, i.e., with $w_i = \lambda_i (A_J w)_i^r$. Set D to be the diagonal matrix whose ith diagonal entry is $\lambda_i r (A_J w)_i^{r-1}$, and suppose that $I - DA_J$ is invertible. Set λ^+ to be the function on the nodes of H which agrees with λ on J and is 0 elsewhere, and w^+ similarly. Then there is some positive real number ε such that, whenever μ is a set of activities on H such that $\|\mu - \lambda^+\| < \varepsilon$, there is a node-weighting v on H inducing μ, i.e., with $v_i = \mu_i (A_H v)_i^r$, such that $\|v - w^+\| < \delta$.*

Proof. Suppose that the node set of H is $[n]$ and, without loss of generality, that the node set of the subgraph J is $[m]$. This means that A_J is the top $m \times m$ submatrix of the $n \times n$ matrix A_H. We shall use I_n and I_m to denote identity matrices of sizes n and m.

Note first that we have $w_i^+ = \lambda_i^+(A_H w^+)_i^r$ for every $i \in H$. For the moment, we leave ε unspecified, and take any set of activities μ on H with $\|\mu - \lambda^+\| < \varepsilon$. Our intention is of course to apply the previous lemma to w^+, λ^+ and μ. To do this, we need to consider the diagonal matrix D_μ of the lemma, whose ith entry is $\mu_i r(A_H w^+)_i^{r-1}$; we need a bound on $\|(I_n - D_\mu A_H)^{-1}\|$ that does not depend on the particular choice of μ.

Set D^+ to be the $n \times n$ diagonal matrix whose top $m \times m$ submatrix is D and which is 0 elsewhere; in other words the ith diagonal entry of D^+ is $\lambda_i^+ r(A_H w^+)_i^{r-1}$. The ith diagonal entry of $D_\mu - D^+$ is then equal to $(\mu_i - \lambda_i^+) r(A_H w^+)_i^{r-1}$, which has absolute value at most $\varepsilon r\|w\|^{r-1}$. This is also therefore an upper bound on $\|D_\mu - D^+\|$.

Our hypothesis states that $I_m - DA_J$ is invertible, with inverse B_J, say, not depending on μ. Suppose A_H has the structure

$$\begin{pmatrix} A_J & C \\ C^t & E \end{pmatrix} ;$$

then one may verify that the inverse of $I_n - D^+ A_H$ is

$$B_H = \begin{pmatrix} B_J & B_J DC \\ 0 & I_{n-m} \end{pmatrix} .$$

Next observe that $(I_n - D_\mu A_H)B_H = I_n - (D_\mu - D^+)A_H B_H$. It is well-known and easy to check that, for $\|X\| < 1$, the inverse of $I - X$ is given by the convergent sum $I + X + X^2 + \cdots$, of norm at most $(1 - \|X\|)^{-1}$. If we choose ε so that $\varepsilon r\|w\|^{r-1}\|A_H\|\,\|B_H\| < 1/2$, then we have that $X = (D_\mu - D^+)A_H B_H$ has norm less than $1/2$, so $I_n - X$ has inverse Y with norm at most 2. Then $I_n - D_\mu A_H$ has inverse $B_\mu = B_H X$ of norm at most $2\|B_H\|$.

Set

$$K_H = 3r(r-1)(2\|\lambda\|)(2\|B_H\|)^2\|w\|^r(\|w\| + 2\|B_H\|\|w\|^r)^{r-2},$$

and note that, provided $\varepsilon \le \|\lambda\|$ and $\varepsilon < 1/2$, this is at least the quantity K defined in the previous lemma. If we also demand that $2K_H\varepsilon < 1$, then we may apply that lemma to obtain that there is a function v from the nodes of H to $\mathbb{R}$ such that $v_i = \mu_i(A_H v)_i^r$ and $\|v - w^+\| \le 4\|w\|^r\|B_H\|\varepsilon$.

If we also choose ε so that $4\|w\|^r\|B_H\|\varepsilon < \delta$, then we are done, except that we must be sure that all the v_i are positive. This is trivial if r is even; if r is odd it is true whenever $\varepsilon < n^{-r}/\delta^{r-1}$, and ε is less than any of the w_i for $i \in J$. Indeed, suppose i (necessarily not in J) has the largest negative value of v_i, then we have $(A_H v)_i = (v_i/\mu_i)^{1/r}$, $v_i > -\delta$ and $\mu_i < \varepsilon$, so there is some neighbour j of i with $v_j \le \frac{1}{n}(v_i/\mu_i)^{1/r} < \frac{1}{n}(v_i/\varepsilon)^{1/r} < v_i\frac{1}{n}\delta^{1/r-1}\varepsilon^{-1/r} < v_i$, contradicting the choice of i. ∎

We are now almost done. If H contains one of the seven small graphs J as a subgraph, take two different node-weightings w and w' on J inducing the same set of activities λ, producing invertible matrices $I - DA_J$ and $I - D'A_J$ respectively, and choose $\delta < \frac{1}{2}\|w - w'\|$. The previous lemma guarantees the existence of positive ε and ε' such that, if μ is any set of activities on H with $\|\mu - \lambda^+\| < \min\{\varepsilon, \varepsilon'\}$, then there are node-weightings v and v' on H, both inducing μ, with $\|v - w\|$ and $\|v' - w'\|$ both at least δ. The choice of δ then guarantees that v and v' are different.

All that remains is to check that, in each of the seven cases, and for each r, the node-weightings we gave earlier do indeed produce invertible matrices $I - DA_J$ and $I - D'A_J$, at least for some values of the parameter s. It turns out to be easiest to check that this is the case as s tends to ∞.

Observe that, in the definition of the matrix D from the weight function w, the ith diagonal entry d_i is $\lambda_i r(Aw)_i^{r-1} = rw_i/z_i$, where $z_i = (Aw)_i$, as usual. In all our cases, most of the d_i will tend to 0 as $s \to \infty$.

We start with the four asymmetric graphs, and the weighting w. As $s \to \infty$, it is easy to see that the parameter γ tends to 0, for any value of r. Thus $d_i = rw_i/z_i$ tends to 0 if i is a node with weight 1 (adjacent to a node of weight s^r) or weight γ (adjacent to a node of weight 1). If i is a node of weight s^r, then it is adjacent to one (other) such node, and so $d_i \to r$ as $s \to \infty$. The matrices $I - DA_J$ thus tend to

$$\begin{pmatrix} 1-r & -r & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ for the pipe and wrench, and } \begin{pmatrix} 1-r & -r & -r & 0 \\ -r & 1-r & -r & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ for the key and gun,}$$

both of which are non-singular for $r \geq 2$.

Exactly the same argument goes through for the wand and hinge; the node-weightings we gave there tend to 0 except on the one weighted node c, where they tend to r, so the limit matrix will simply be a relabeling of the first matrix above.

The case of the weighting w' on the four asymmetric graphs is a little more complex. Let us take the pipe as an example; the others are similar. Here we have

$$|I - DA_J| = \begin{vmatrix} 1-d_a & -d_a & 0 \\ -d_b & 1 & -d_b \\ 0 & -d_c & 1 \end{vmatrix} = \begin{vmatrix} 1-d_a & -d_a & 0 \\ -d_b & 1-d_b d_c & -d_b \\ 0 & 0 & 1 \end{vmatrix},$$

doing a column operation. The point is that, in this case, $d_a = r/(\beta+1)$ and $d_b = r\beta/(1+\gamma\beta^r)$ both tend to 0 as $s \to \infty$, and although $d_c = r\gamma\beta^r/\beta$ tends to infinity, $d_b d_c$ tends to r^2. The above calculation makes it clear that the matrix $I - DA_J$ is non-singular for sufficiently large s.

Finally, the same technique works for the stick; again we have that all the d_i tend to 0 except for d_d, which tends to infinity, but with $d_c d_d$ tending to r^2, the same column operation then shows that the matrix is non-singular for s sufficiently large.

This completes the proof of the main theorem. ∎

9 The Hinge

In this final section we look at a phase transition exhibited by $\mathrm{Hom}(T_2, H)$ where H is the hinge. We associate nodes a, b, c of the hinge with colors red, yellow and green, the only non-adjacency in H being between red and green. Thus $\mathrm{Hom}(T_2, H)$ is the space of all red-yellow-green colorings of the sites of T_r in which no red site is adjacent to a green one.

The hinge constraint in fact corresponds to a discrete version of the Widom-Rowlinson model, in which two gases (whose particles are represented by red and green) compete for space and are not permitted to occupy adjacent sites; see e.g. [17, 16, 4]. In fact our methods can be used to obtain a result of [16], namely that when $\lambda_a = \lambda_c = \gamma$ and $\lambda_b = 1$, then there is a unique simple invariant Gibbs measure just when

$$\gamma \leq \frac{1}{r-1}\left(\frac{r+1}{r}\right)^r .$$

For any lattice G, and H the hinge, putting $a < b < c$ imposes on $\mathrm{Hom}(G, H)$ the structure of a partially ordered set, in fact a distributive lattice (with "lattice" now used in the algebraic sense). This nice structure can be used to show that if there is only one simple invariant Gibbs measure, then there is only one Gibbs measure of any kind; see [4]. Here, the hinge is chosen for our example just because it is particularly easy to visualize.

Consider the weights $w_a = 1$, $w_b = 2$ and $w_c = 4$, giving

$$\lambda_a = \frac{1}{(1+2)^2} = \frac{1}{9},\ \lambda_b = \frac{2}{(1+2+4)^2} = \frac{2}{49},\ \lambda_c = \frac{4}{(2+4)^2} = \frac{1}{9}\ .$$

It is perhaps worthwhile to compute just two conditional probabilities, to see how it can be that asymmetric weights give rise to symmetric activities. Suppose the predecessor and both successors to some site $v \in T_2$ are all yellow; then from Bayes' formula the probability that x is red is

$$\frac{\frac{1}{1+2+4} \cdot \left(\frac{2}{1+2}\right)^2}{\frac{1}{1+2+4} \cdot \left(\frac{2}{1+2}\right)^2 + \frac{2}{1+2+4} \cdot \left(\frac{2}{1+2+4}\right)^2 \frac{4}{1+2+4} \cdot \left(\frac{2}{2+4}\right)^2} = 49/116$$

while the probability that v is green is

$$\frac{\frac{4}{1+2+4} \cdot \left(\frac{2}{2+4}\right)^2}{\frac{1}{1+2+4} \cdot \left(\frac{2}{1+2}\right)^2 + \frac{2}{1+2+4} \cdot \left(\frac{2}{1+2+4}\right)^2 \frac{4}{1+2+4} \cdot \left(\frac{2}{2+4}\right)^2} = 49/116$$

The stationary probabilities for this node-weighted random walk are, in contrast to the activities, even less balanced than the weights:

$$\pi_a = \frac{1(1+2)}{1(1+2) + 2(1+2+4) + 4(2+4)} = \frac{3}{41} \approx 7\%$$

and similarly

$$\pi_b = \frac{14}{41} \approx 34\% \text{ and } \pi_c = \frac{24}{41} \approx 59\%\ .$$

We call the measure induced by this random walk the "September" phase, a sample coloring from which is illustrated in the top third of the color plate. In the illustration a portion of T_2 has been laid out on the plane, and colored according to a computer-generated random walk.

Of course the reverse weighting, $w_a = 4$, $w_b = 2$ and $w_c = 1$, yields the same set of activities and the reverse stationary probabilities; this is the "November" phase, illustrated in the **bottom** third of the color plate.

In addition to the above weighting and its reverse, there is a third, symmetric, weighting which generates the same set of activities. Its values are

$$w_a = w_c \approx 2.5866,\ w_b \approx 2.2383$$

giving stationary distribution

$$\pi_a = \pi_c \approx 30\%, \pi_b \approx 40\%.$$

A sample from this "October" phase is provided in the center of the color plate.

Acknowledgments

This work was begun during a visit by the second author to the London School of Economics, and continued with visits by the first author to New Jersey under the auspices of DIMACS. The authors are indebted to a meeting of the Institute for Elementary Studies (Pinecrest Lodge, California, August 1995) for inspiration, and to all of the following for insight: Eric Bach, Rob van den Berg, Christian Borgs, Jennifer Chayes, Olle Häggström, Russ Lyons, Robin Pemantle, Yuval Peres, Jim Propp, Len Schulman, Peter Shor, Jeff Steif and David Wilson.

References

[1] R.J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, London (1982).

[2] J. van den Berg and J.E. Steif, Percolation and the hard-core lattice gas model, *Stochastic Proc. and their Appls.* **49** (1994), 179–197.

[3] R. Burton and J.E. Steif, Nonuniqueness of measures of maximal entropy for subshifts of finite type, preprint.

[4] G.R Brightwell, O. Häggström and P. Winkler, Nonmonotonic behavior in hard-core and Widom–Rowlinson models, CDAM Research Report LSE-CDAM-98-13 (June 1998); shorter version to appear in *J. Stat. Physics*.

[5] G.R. Brightwell and P. Winkler, Gibbs measures and dismantlable graphs, CDAM Research Report LSE-CDAM-97-17 (December 1997); submitted for publication.

[6] R.L. Dobrushin, The description of a random field by means of conditional probabilities and conditions of its regularity, *Thy. of Prob. and its Appls.* **13** #2 (1968), 197–224.

[7] J. Dugundji, *Topology*, Allyn & Bacon, Boston (1966).

[8] W. Evans, C. Kenyon, Y. Peres and L.J. Schulman, A critical phenomenon in a broadcast process, preprint (Sept. 1996).

[9] H.-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter, Berlin (1988).

[10] F.P. Kelly, Stochastic models of computer communication systems, *J. R. Statist. Soc. B* **47** #3 (1985), 379–395.

[11] O.E. Lanford and D. Ruelle, Observables at infinity and states with short range correlations in statistical mechanics, *CMP* **9** (1969), 327–338.

[12] G.B. Price, *Multivariate Analysis*, Springer (1984).

[13] B. Prum and J.C. Fort, *Stochastic Processes on a Lattice and Gibbs Measures*, Kluwer, Dordrecht (1991).

[14] E. Seneta, *Non-Negative Matrices*, George Allen and Unwin (1973).

[15] F. Spitzer, Markov random fields on an infinite tree, *Ann. Prob.* **3** (1975), 387–398.

[16] J.C. Wheeler and B. Widom, Phase equilibrium and critical behavior in a two-component Bethe-lattice gas or three-component Bethe-lattice solution, *J. Chem. Phys.* **52** (1970), 5334–5343.

[17] B. Widom and J.S. Rowlinson, New model for the study of liquid-vapor phase transition, *J. Chem. Phys.* **52** (1970), 1670–1684.

[18] S. Zachary, Countable state space Markov random fields and Markov chains on trees, *Ann. Prob.* **11** (1983), 894–903.