J. Eur. Math. Soc. 19, 1811–1844

DOI 10.4171/JEMS/706

Péter Csikvári

Lower matching conjecture, and a new proof of Schrijver's and Gurvits's theorems

Received August 3, 2014

Abstract. Friedland's Lower Matching Conjecture asserts that if G is a d-regular bipartite graph on $v(G) = 2n$ vertices, and $m_k(G)$ denotes the number of matchings of size k, then

$$m_k(G) \geq \binom{n}{k}^2 \left(\frac{d-p}{d}\right)^{n(d-p)} (dp)^{np},$$

where $p = k/n$. When $p = 1$, this conjecture reduces to a theorem of Schrijver which says that a d-regular bipartite graph on $v(G) = 2n$ vertices has at least

$$\left(\frac{(d-1)^{d-1}}{d^{d-2}}\right)^n$$

perfect matchings. L. Gurvits proved an asymptotic version of the Lower Matching Conjecture, namely

$$\frac{\ln m_k(G)}{v(G)} \geq \frac{1}{2}\left(p \ln\left(\frac{d}{p}\right) + (d-p)\ln\left(1-\frac{p}{d}\right) - 2(1-p)\ln(1-p)\right) + o_{v(G)}(1).$$

In this paper, we prove the Lower Matching Conjecture. In fact, we establish a slightly stronger statement which gives an extra $c_p\sqrt{n}$ factor compared to the conjecture if p is separated away from 0 and 1, and is tight up to a constant factor if p is separated away from 1. We will also give a new proof of Gurvits's and Schrijver's theorems, and we extend these theorems to (a, b)-biregular bipartite graphs.

Keywords. Matchings, matching polynomial, Benjamini–Schramm convergence, infinite regular tree, infinite biregular tree, 2-lift

1. Introduction

Throughout this paper we use standard terminology, but the second paragraph of Section 2 might help the reader in the case of some concepts undefined in this introduction.

One of the best known theorem concerning the number of perfect matchings of a d-regular graph is due to A. Schrijver and M. Voorhoeve.

P. Csikvári: Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, and
Department of Computer Science, Eötvös Loránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary; e-mail: peter.csikvari@gmail.com

Mathematics Subject Classification (2010): Primary 05C35; Secondary 05C31, 05C70, 05C80

Theorem 1.1 (A. Schrijver [25] for general d, M. Voorhoeve [27] for $d = 3$). *Let G be a d-regular bipartite graph on $2n$ vertices and let $\mathrm{pm}(G)$ denote the number of perfect matchings of G. Then*

$$\mathrm{pm}(G) \geq \left(\frac{(d-1)^{d-1}}{d^{d-2}} \right)^n.$$

There are two different proofs of Theorem 1.1: the original one due to A. Schrijver [25], and another proof using stable polynomials due to L. Gurvits [13]; for a beautiful account of the latter see [18]. In this paper we will give a third proof which is essentially different from the previous ones.

S. Friedland, E. Krop and K. Markström [9] conjectured a possible generalization of this theorem which extends Schrijver's theorem to any size of matchings. This conjecture became known as Friedland's Lower Matching Conjecture:

Conjecture 1.2 (Friedland's Lower Matching Conjecture [9]). Let G be a d-regular bipartite graph on $v(G) = 2n$ vertices, and let $m_k(G)$ denote the number of matchings of size k. Then

$$m_k(G) \geq \binom{n}{k}^2 \left(\frac{d-p}{d} \right)^{n(d-p)} (dp)^{np}, \qquad \text{where } p = k/n.$$

They also proposed an asymptotic version of this conjecture which was later proved by L. Gurvits [14].

Theorem 1.3 (L. Gurvits [14]). *Let G be a d-regular bipartite graph on $v(G) = 2n$ vertices, and let $m_k(G)$ denote the number of matchings of size k. Then*

$$\frac{\ln m_k(G)}{v(G)} \geq \frac{1}{2} \left(p \ln\left(\frac{d}{p}\right) + (d-p) \ln\left(1 - \frac{p}{d}\right) - 2(1-p)\ln(1-p) \right) + o_{v(G)}(1),$$

where $p = k/n$.

When $p = 1$ this result almost reduces to Schrijver's theorem, but Gurvits used this special case to establish the general case. More precisely, Gurvits used the following result of Schrijver: Let $A = (a_{ij})$ be a doubly stochastic matrix, and $\tilde{A} = (\tilde{a}_{ij})$, where $\tilde{a}_{ij} = a_{ij}(1 - a_{ij})$. Then the permanent of $\tilde{A}$ satisfies the inequality

$$\mathrm{Per}(\tilde{A}) \geq \prod_{i,j} (1 - a_{ij}).$$

Note that Gurvits [14] proved an effective version of Theorem 1.3, but for our purposes any $o_{v(G)}(1)$ term would suffice, as we will make it "vanish". More details on Gurvits's results can be found in Remark 3.3.

It is worth introducing some notation for the function appearing in Theorem 1.3, and with some foresight we introduce another function with parameters a, b which will be important for us when we study (a, b)-biregular graphs.

Definition 1.4. Let $0 \le q \le 1$ and

$$H(q) = -(q \ln q + (1-q)\ln(1-q))$$

with the usual convention that $H(0) = H(1) = 0$. Furthermore, for a positive integer d and $0 \le p \le 1$ let

$$\mathbb{G}_d(p) = \frac{1}{2}\left(p \ln\left(\frac{d}{p}\right) + (d-p)\ln\left(1 - \frac{p}{d}\right) - 2(1-p)\ln(1-p)\right),$$

and for positive integers a and b let

$$\mathbb{G}_{a,b}(p) = \frac{a}{a+b} H\left(\frac{a+b}{2a} p\right) + \frac{b}{a+b} H\left(\frac{a+b}{2b} p\right) + \frac{1}{2} p \ln(ab) - \frac{ab}{a+b} H\left(\frac{a+b}{2ab} p\right),$$

where $0 \le p \le \min(\frac{2a}{a+b}, \frac{2b}{a+b})$.

Note that one can rewrite $\mathbb{G}_{a,b}(p)$ as follows:

$$\begin{aligned}\mathbb{G}_{a,b}(p) = \frac{1}{2}\Bigg(&p \ln\left(\frac{2ab}{(a+b)p}\right) + \left(\frac{2ab}{a+b} - p\right)\ln\left(1 - \frac{a+b}{2ab} p\right) \\ &- \left(\frac{2a}{a+b} - p\right)\ln\left(1 - \frac{a+b}{2a} p\right) - \left(\frac{2b}{a+b} - p\right)\ln\left(1 - \frac{a+b}{2b} p\right)\Bigg).\end{aligned}$$

From this form it is clear that for $a = b = d$, we have $\mathbb{G}_d(p) = \mathbb{G}_{a,b}(p)$. Later it will turn out that $\mathbb{G}_d(p)$ is the so-called entropy function of the infinite d-regular tree $\mathbb{T}_d$, and $\mathbb{G}_{a,b}(p)$ is the entropy function of the infinite (a, b)-biregular tree $\mathbb{T}_{a,b}$.

To show the connection between Conjecture 1.2 and Theorem 1.3, let us introduce one more parameter. Let $p = k/n$, and let p_μ be the probability that a random variable with Binomial(n, p) distribution takes its mean value $\mu = k$. In other words,

$$p_\mu = \binom{n}{k} p^k (1-p)^{n-k}.$$

With this new notation the function appearing in Conjecture 1.2 is

$$\binom{n}{k}^2 \left(\frac{d-p}{d}\right)^{n(d-p)} (dp)^{np} = p_\mu^2 \exp(2n\mathbb{G}_d(p)).$$

Hence Conjecture 1.2 claims that

$$m_k(G) \ge p_\mu^2 \exp(2n\mathbb{G}_d(p)).$$

It turns out that a slightly stronger statement is true.

Theorem 1.5. *Let G be a d-regular bipartite graph on $v(G) = 2n$ vertices, and let $m_k(G)$ denote the number of matchings of size k. Furthermore, let $p = k/n$, and p_μ be the probability that a random variable with* Binomial(n, p) *distribution takes its mean value $\mu = k$. Then*

$$m_k(G) \geq p_\mu \exp(2n\mathbb{G}_d(p)).$$

In particular, Conjecture 1.2 holds true. Furthermore, for every $0 \leq k < n$ there exists a d-regular bipartite graph G on $2n$ vertices such that

$$m_k(G) \leq \sqrt{\frac{1 - p/d}{1 - p}} \cdot p_\mu \exp(2n\mathbb{G}_d(p)).$$

Note that $p_\mu \approx 1/\sqrt{2\pi p(1-p)n}$, which means that if p is separated away from 0 and 1, then we can obtain an extra $c_p\sqrt{n}$ factor compared to Conjecture 1.2. Also note that always $p_\mu \geq \frac{1}{n+1} \geq \frac{1}{2n}$. This inequality might be easier to handle in some cases.

We will show that Theorem 1.3 implies Conjecture 1.2. The idea of the proof of Theorem 1.5 is to convert Gurvits's theorem to a statement on analytical functions arising from statistical mechanics. Then tools from analysis and probability theory together with a simple observation will enable us to replace the term $o_{v(G)}(1)$ in Gurvits's theorem with an effective one which is slightly better than the corresponding term in Gurvits's original theorem (see Remark 3.3).

We offer one more theorem for d-regular bipartite graphs.

Theorem 1.6. *Let G be a d-regular bipartite graph on $v(G) = 2n$ vertices, and let $m_k(G)$ denote the number of matchings of size k. Let $0 \leq p \leq 1$. Then*

$$\sum_{k=0}^{n} m_k(G)\left(\frac{p}{d}\left(1 - \frac{p}{d}\right)\right)^k (1-p)^{2(n-k)} \geq \left(1 - \frac{p}{d}\right)^{nd}.$$

When $p = 1$, Theorem 1.6 immediately yields Theorem 1.1. Indeed, when $p = 1$ only the term $m_n(G)\left(\frac{1}{d}\left(1 - \frac{1}{d}\right)\right)^n$ does not vanish on the left hand side, because of the term $(1-p)^{2(n-k)}$, and we get

$$m_n(G)\left(\frac{1}{d}\left(1 - \frac{1}{d}\right)\right)^n \geq \left(1 - \frac{1}{d}\right)^{nd},$$

which is equivalent to

$$m_n(G) \geq \left(\frac{(d-1)^{d-1}}{d^{d-2}}\right)^n.$$

As already mentioned, Theorem 1.3 implies Theorem 1.5, but the main goal of this paper is to give a new proof of Gurvits's and Schrijver's theorems with a novel method. This method will be used to prove Theorem 1.6 too. This new proof shows that the extremal graph is in some sense the d-regular infinite tree. Indeed, we will show that the function on the right hand side of Theorem 1.3 is nothing other than the so-called entropy function of the d-regular infinite tree; the entropy functions of finite and infinite graphs will be introduced in Section 2. This means that for a deeper understanding of these theorems, one needs to step out from the universe of finite graphs. We will do it by using the recently developed theory of Benjamini–Schramm convergence of bounded degree graphs. This new technique also enables us to extend these theorems to (a,b)-biregular bipartite graphs.

Theorem 1.7. *Let $G=(A,B,E)$ be an (a,b)-biregular bipartite graph on $v(G)$ vertices such that every vertex in A has degree a, and every vertex in B has degree b. Assume that $a \geq b$, i.e., $|A| \leq |B|$. Let $m_k(G)$ denote the number of matchings of size k, and $p = 2k/v(G)$. Furthermore, let $q = \frac{a+b}{2b}p$, and let p_μ be the probability that a random variable with* Binomial$(|A|,q)$ *distribution takes its mean value $\mu = k$. Then*

$$m_k(G) \geq p_\mu \exp(v(G) \cdot \mathbb{G}_{a,b}(p)).$$

Note that if $k=|A|$, then $p_\mu = 1$, and $p_\mu \geq \frac{1}{|A|+1} \geq \frac{1}{v(G)}$ for any p.

One can view Theorem 1.7 and the other results as extremal graph-theoretic problems where one seeks for the extremal value of a certain graph parameter $p(G)$ in a given family $\mathcal{G}$ of graphs. In extremal graph theory it is a classical idea to try to find some graph transformation φ such that $p(G) \leq p(\varphi(G))$ (or $p(G) \geq p(\varphi(G))$), and $\varphi(G) \in \mathcal{G}$ for every $G \in \mathcal{G}$. Then we apply this transformation as long as we can, and when we stop then we know that the extremal graph must be in a special subfamily of $\mathcal{G}$, where the optimization problem can be solved easily. See for instance the proof of Turán's theorem using Zykov's symmetrization [28]. In our case, the transformation φ will simply be any 2-lift of the graph (see Definition 4.1). The new ingredient in our proof is that the sequence of graphs obtained by repeatedly applying the 2-lifts will not stabilize, but instead converge to the infinite biregular tree. In fact, most of our work is related to the graph convergence part, and not the graph transformation part.

Organization. In the next section we introduce all the necessary tools including the density function $p(G,t)$, the entropy function $\lambda_G(p)$, and Benjamini–Schramm convergence, and we compute the entropy function of the infinite biregular tree. In this section we also give various results on the number of matchings of random (bi)regular graphs, which shows the tightness of our results. In particular, we prove the second half of Theorem 1.5 here. In Section 3 we show that Gurvits's theorem is equivalent to a certain (effective) statement on the entropy function. In Section 4 we give a new proof of Schrijver's and Gurvits's theorem together with the main part of the proof of Theorem 1.7. In Section 5 we deduce Theorem 1.5 from the new version of Gurvits's theorem, we prove Theorem 1.6, and we also finish the proof of Theorem 1.7.

How to read this paper. This paper is occasionally a bit technical, especially in Section 2. In order to make it easier to read we now roughly summarize its content and give a road map for a first reading. Assuming that the reader is mainly interested in the proof of Theorem 1.5, we first give an idea how the proof works.

Assume that $p(G)$ is some graph parameter related to matchings and it is normalized in such a way that we can compare graphs of different sizes, in particular it makes sense to compare two d-regular graphs. For instance

$$p(G) = \frac{\ln \mathrm{pm}(G)}{v(G)}$$

is such a graph parameter. We will prove that for a bipartite d-regular graph G we have

$$p(G) \geq p(\mathbb{T}_d),$$

where $p(\mathbb{T}_d)$ a priori does not make sense, but can be defined as a limit $\lim_{i\to\infty} p(H_i)$, where H_i is a sequence of graphs "locally converging" to $\mathbb{T}_d$. The plan is the following: We define a sequence of graphs G_i such that $G = G_0$ and

$$p(G) = p(G_0) \geq p(G_1) \geq p(G_2) \geq \cdots$$

and

$$\lim_{i\to\infty} p(G_i) = p(\mathbb{T}_d).$$

This clearly gives

$$p(G) \geq p(\mathbb{T}_d).$$

A technical difficulty arises from the fact that it is not really convenient to work with the parameter

$$q(G) = \frac{\ln m_k(G)}{v(G)}.$$

Instead we use the entropy function $\lambda_G(p)$ which is strongly related to $q(G)$, but is more amenable to analysis. So $\lambda_G(p)$ will play the role of $p(G)$. Of course, we need some tools to transfer our knowledge from $\lambda_G(p)$ to $q(G)$, but it is again just a technical problem.

So by keeping in mind our simple plan and the technical difficulty, we suggest the following road map for the first reading: (1) first read the alternative definition of $\lambda_G(p)$ (Remark 2.2), which is easy to understand, then take a quick look at its properties (Proposition 2.1) without reading the proof, (2) read the definition of Benjamini–Schramm convergence (Definition 2.5) and Example 2.7, (3) jump to Section 4, but read it only till the proof of Theorems 1.1 and 1.3, (4) finally, read Section 5. We believe that reading just this core of the paper will give a good impression of its content and the novel method applied.

Let us mention that if the reader is familiar with Gurvits's result (Theorem 1.3) and only wants to know how one can derive Theorem 1.5 from it, then after step (1) in the above plan one can jump immediately to Section 3 and then to Section 5.

2. Preliminaries and basic notions

This section is mostly reproduced from [2]. We could have simply cited that paper, but for the convenience of the reader, we also include some proofs.

Throughout, G denotes a finite graph with vertex set $V(G)$ and edge set $E(G)$. The number of vertices is denoted by $v(G)$. The *degree* of a vertex is the number of its neighbors. A graph is called *d-regular* if every vertex has degree exactly d. A *cycle C* is a sequence $v_1, \dots, v_k$ of vertices such that $v_i \neq v_j$ if $i \neq j$ and $(v_i, v_{i+1}) \in E(G)$ for $i = 1, \dots, k$, where $v_{k+1} = v_1$. The *length* of the cycle is k in this case. A *k-matching* is a set $\{e_1, \dots, e_k\}$ of edges such that for any i and j, the vertex sets of e_i and e_j are disjoint, in other words, $e_1, \dots, e_k$ cover $2k$ vertices between themselves. A *perfect matching* is a matching which covers all vertices. A graph is called *bipartite* if the vertices can be colored with two colors such that each edge connects two vertices of different colors. The standard notation for a bipartite graph is $G = (A, B, E)$, where A and B denote the vertex sets corresponding to the two color classes.

Let $G = (V, E)$ be a finite graph on $v(G)$ vertices, and let $m_k(G)$ denote the number of k-matchings ($m_0(G) = 1$). Let t be a non-negative real number; in statistical mechanics it is called the activity. Let

$$M(G,t) = \sum_{k=0}^{\lfloor v(G)/2 \rfloor} m_k(G) t^k, \qquad \mu(G,x) = \sum_{k=0}^{\lfloor v(G)/2 \rfloor} (-1)^k m_k(G) x^{v(G)-2k}.$$

We call $M(G,t)$ the *matching generating function*,[1] and $\mu(G,x)$ the *matching polynomial* [16, 10, 11]. Clearly, they encode the same information. Let

$$p(G,t) = \frac{2t \cdot \frac{d}{dt} M(G,t)}{v(G) \cdot M(G,t)}, \qquad F(G,t) = \frac{\ln M(G,t)}{v(G)} - \frac{1}{2} p(G,t) \ln(t).$$

We will call $p(G,t)$ the *density function*. Note that it has a natural interpretation: Assume that we choose a random matching M with probability proportional to $t^{|M|}$; then the expected number of vertices covered by a random matching is $p(G,t) \cdot v(G)$. Let

$$p^*(G) = \frac{2\nu(G)}{v(G)},$$

where $\nu(G)$ denotes the number of edges in the largest matching. If G contains a perfect matching, then clearly $p^* = 1$. The function $p = p(G,t)$ is a strictly increasing function which maps $[0, \infty)$ to $[0, p^*)$, where $p^* = p^*(G)$. Therefore, its inverse function $t = t(G,p)$ maps $[0, p^*)$ to $[0, \infty)$. (If G is clear from the context, then we simply write $t(p)$ instead of $t(G,p)$.) Let

$$\lambda_G(p) = F(G, t(p))$$

if $p < p^*$, and $\lambda_G(p) = 0$ if $p > p^*$. Note that we have not defined $\lambda_G(p^*)$ yet. We define it as a limit:

$$\lambda_G(p^*) = \lim_{p \nearrow p^*} \lambda_G(p).$$

We will show that this limit exists Proposition 2.1(c). We will call $\lambda_G(p)$ the *entropy function* of the graph G.

The intuitive meaning of $\lambda_G(p)$ is the following. Assume that we want to count the number of matchings covering a p fraction of the vertices. Assume that it makes sense: $p = 2k/v(G)$, and so we wish to count $m_k(G)$. Then

$$\lambda_G(p) \approx \frac{\ln m_k(G)}{v(G)}.$$

A more precise formulation of this statement is given below.

[1] In statistical mechanics, it is called the partition function of the monomer-dimer model.

Proposition 2.1 ([2]). *Let G be a finite graph.*

(a) *Let rG be the union of r disjoint copies of G. Then*

$$\lambda_G(p) = \lambda_{rG}(p).$$

(b) *If $p < p^*$, then*

$$\frac{d}{dp}\lambda_G(p) = -\frac{1}{2}\ln t(p).$$

(c) *The limit*

$$\lambda_G(p^*) = \lim_{p \nearrow p^*} \lambda_G(p)$$

exists.

(d) *Let $k \leq \nu(G)$ and $p = 2k/v(G)$. Then*

$$\left|\lambda_G(p) - \frac{\ln m_k(G)}{v(G)}\right| \leq \frac{\ln v(G)}{v(G)}.$$

(e) *Let $k = \nu(G)$. Then for $p^* = 2k/v(G)$ we have*

$$\lambda_G(p^*) = \frac{\ln m_k(G)}{v(G)}.$$

In particular, if G contains a perfect matching then

$$\lambda_G(1) = \frac{\ln \mathrm{pm}(G)}{v(G)}.$$

(f) *If for some function $f(p)$ we have*

$$\lambda_G(p) \geq f(p) + o_{v(G)}(1)$$

for all graphs G, then

$$\lambda_G(p) \geq f(p).$$

(g) *If for some graphs G_1 and G_2 we have*

$$\frac{\ln M(G_1, t)}{v(G_1)} \geq \frac{\ln M(G_2, t)}{v(G_2)} \quad \textit{for every } t \geq 0,$$

then

$$\lambda_{G_1}(p) \geq \lambda_{G_2}(p) \quad \textit{for every } 0 \leq p \leq 1.$$

Remark 2.2. Parts (a) and (d) of Proposition 2.1 together suggest an alternative definition for the entropy function $\lambda_G(p)$ for $p < p^*$: Let (k_r) be a sequence of integers such that

$$\lim_{r\to\infty} \frac{2k_r}{rv(G)} = p.$$

Then

$$\lambda_G(p) = \lim_{r\to\infty} \frac{\ln m_{k_r}(rG)}{rv(G)}.$$

In general when we have an infinite graph L, say $\mathbb{Z}^d$, then it is a natural idea to consider a graph sequence G_i converging to L and to take a sequence (k_i) such that

$$\lim_{i\to\infty} \frac{2k_i}{v(G_i)} = p,$$

and then to consider

$$\lambda_L(p) = \lim_{i\to\infty} \frac{\ln m_{k_i}(G_i)}{v(G_i)}.$$

In this sense, this alternative definition is nothing other than to consider the "G-lattice" of infinitely many disjoint copies of G and approximate it with $G_i = iG$, the union of i copies of G.

This alternative definition is much more natural, especially from the statistical physics point of view. On the other hand, it is hard to work with.

We will need some preparation to prove Proposition 2.1. First, we will need the following fundamental theorem about the matching polynomial.

Theorem 2.3 (Heilmann and Lieb [16]). *The zeros of the matching polynomial $\mu(G, x)$ are real, and if the largest degree D of G is greater than* 1*, then all zeros lie in the interval* $[-2\sqrt{D-1}, 2\sqrt{D-1}]$.

We will also use the following theorem of Darroch.

Lemma 2.4 (Darroch's rule [6]). *Let $P(x) = \sum_{k=0}^{n} a_k x^k$ be a polynomial with only positive coefficients and real zeros. If*

$$k - \frac{1}{n-k+2} < \frac{P'(1)}{P(1)} < k + \frac{1}{k+2},$$

then k is the unique number for which $a_k = \max(a_1, \ldots, a_n)$. If, on the other hand,

$$k + \frac{1}{k+2} < \frac{P'(1)}{P(1)} < k + 1 - \frac{1}{n-k+1},$$

then either a_k or a_{k+1} is the maximal element of $a_1, \ldots, a_n$.

Proof of Proposition 2.1. (a) Note that

$$M(rG,t)=M(G,t)^r,$$

implying that $p(rG,t)=p(G,t)$ and $\lambda_{rG}(p)=\lambda_G(p)$.

(b) Since

$$\lambda_G(p)=\frac{\ln M(G,t)}{v(G)}-\frac{1}{2}p\ln t,$$

we have

$$\frac{d\lambda_G(p)}{dp}=\left(\frac{1}{v(G)}\cdot\frac{\frac{d}{dt}M(G,t)}{M(G,t)}\cdot\frac{dt}{dp}-\frac{1}{2}\left(\ln t+p\cdot\frac{1}{t}\cdot\frac{dt}{dp}\right)\right)=-\frac{1}{2}\ln t,$$

since

$$\frac{1}{v(G)}\cdot\frac{\frac{d}{dt}M(G,t)}{M(G,t)}=\frac{p}{2t}$$

by definition.

(c) From $\frac{d}{dp}\lambda_G(p)=-\frac{1}{2}\ln t(p)$ we see that for $p>p(G,1)$, the function $\lambda_G(p)$ is decreasing. (We can also see that $\lambda_G(p)$ is a concave-down function.) Hence

$$\lim_{p\nearrow p^*}\lambda_G(p)=\inf_{p>p(G,1)}\lambda_G(p).$$

(d) First, assume that $k<\nu(G)$. For $k=\nu(G)$, we will slightly modify our argument. Let $t=t(G,p)$ be the value for which $p=p(G,t)$. The polynomial

$$P(G,x)=M(G,tx)=\sum_{j=0}^{n}m_j(G)t^jx^j,$$

considered as a polynomial in variable x, has only real zeros by Theorem 2.3. Note that

$$k=\frac{pv(G)}{2}=\frac{P'(G,1)}{P(G,1)}.$$

Darroch's rule says that in this case $m_k(G)t^k$ is the unique maximal element of the coefficient sequence of $P(G,x)$. In particular

$$\frac{M(G,t)}{v(G)}\le m_k(G)t^k\le M(G,t).$$

Hence

$$\lambda_G(p)-\frac{\ln v(G)}{v(G)}\le\frac{\ln m_k(G)}{v(G)}\le\lambda_G(p).$$

Hence for $k<\nu(G)$, we are done.

If $k=\nu(G)$, then let p be arbitrary such that

$$k-1/2<pv(G)/2<k.$$

Again by Darroch's rule,

$$\lambda_G(p)-\frac{\ln v(G)}{v(G)}\le\frac{\ln m_k(G)}{v(G)}\le\lambda_G(p).$$

Since this is true for all p sufficiently close to $p^* = 2\nu(G)/v(G)$, and $\lambda_G(p^*) = \lim_{p \nearrow p^*} \lambda_G(p)$, we have

$$\left| \lambda_G(p^*) - \frac{\ln m_k(G)}{v(G)} \right| \leq \frac{\ln v(G)}{v(G)}$$

in this case too.

(e) By (a) we have $\lambda_{rG}(p) = \lambda_G(p)$. Note also that if $k = \nu(G)$, then $m_{rk}(rG) = m_k(G)^r$. Applying the bound from (d) to the graph rG, we obtain

$$\left| \lambda_G(p^*) - \frac{\ln m_k(G)}{v(G)} \right| \leq \frac{\ln v(rG)}{v(rG)}.$$

Since $\ln v(rG)/v(rG) \to 0$ as $r \to \infty$, we get

$$\lambda_G(p^*) = \frac{\ln m_k(G)}{v(G)}.$$

(f) This is again a trivial consequence of $\lambda_{rG}(p) = \lambda_G(p)$.

(g) From the assumption it follows that for the relative sizes of the largest matchings, we have $\nu(G_1)/v(G_1) \geq \nu(G_2)/v(G_2)$, and if there is equality, then

$$\frac{\ln m_{\nu(G_1)}(G_1)}{v(G_1)} \geq \frac{\ln m_{\nu(G_2)}(G_2)}{v(G_2)}.$$

Thus the statement is trivial if $p \geq 2\nu(G_2)/v(G_2)$. So we can assume that $0 \leq p < 2\nu(G_2)/v(G_2)$. Let us consider the minimum of the function $\lambda_{G_1}(p) - \lambda_{G_2}(p)$ on the interval $[0, 2\nu(G_2)/v(G_2)]$. This minimum is either attained at some endpoints or inside the interval at a point where the derivative is 0. Note that $\lambda_{G_1}(0) = \lambda_{G_1}(0) = 0$. According to (b), the derivative of $\lambda_{G_1}(p) - \lambda_{G_2}(p)$ is

$$-\tfrac{1}{2} \ln t(G_1, p) + \tfrac{1}{2} \ln t(G_2, p).$$

If it is 0 at p_0 then $t(G_1, p_0) = t(G_2, p_0)$, but then with the notation $t = t(G_1, p_0) = t(G_2, p_0)$ we have

$$\lambda_{G_1}(p_0) = \frac{\ln M(G_1, t)}{v(G_1)} - \frac{1}{2} p_0 \ln(t) \geq \frac{\ln M(G_2, t)}{v(G_2)} - \frac{1}{2} p_0 \ln(t) = \lambda_{G_2}(p_0).$$

So at every possible minimum of $\lambda_{G_1}(p) - \lambda_{G_2}(p)$, the function is non-negative. So it is non-negative everywhere. □

2.1. Benjamini–Schramm convergence and the entropy function

In this part we extend the definition of the function $\lambda_G(p)$ to infinite lattices L, more precisely to certain *random rooted graphs*.

Definition 2.5. Let L be a probability distribution on (infinite) rooted graphs; we will call L a *random rooted graph*. For a finite rooted graph α and a positive integer r, let $\mathbb{P}(L, \alpha, r)$ be the probability that the r-ball centered at a random root vertex chosen from the distribution L is isomorphic to α.

For a finite graph G, a finite rooted graph α and a positive integer r, let $\mathbb{P}(G, \alpha, r)$ be the probability that the r-ball centered at a uniform random vertex of G is isomorphic to α.

We say that a sequence (G_i) of bounded degree graphs is *Benjamini–Schramm convergent* if for all finite rooted graphs α and $r > 0$, the probabilities $\mathbb{P}(G_i, \alpha, r)$ converge. Furthermore, we say that (G_i) *Benjamini–Schramm converges to* L if $\mathbb{P}(G_i, \alpha, r) \to \mathbb{P}(L, \alpha, r)$ for all positive integers r and finite rooted graphs α.

Note that Benjamini–Schramm convergence is also called local convergence. This refers to the fact that the finite graphs G_i look locally more and more like the infinite graph L.

Example 2.6. Consider a sequence of boxes in $\mathbb{Z}^d$ all of whose sides converge to infinity. This is a Benjamini–Schramm convergent graph sequence since for every fixed r, we can pick a vertex which is at least r away from the boundary with probability converging to 1. For all these vertices we will see the same neighborhood. This also shows that we can impose an arbitrary boundary condition, for instance the periodic boundary condition means that we consider the sequence of toroidal boxes. Boxes and toroidal boxes will be Benjamini–Schramm convergent even together, and they converge to a distribution which is a rooted $\mathbb{Z}^d$ with probability 1.

Example 2.7. Recall that a k-cycle of a graph H is a sequence of vertices $v_1, \dots, v_k$ such that $v_i \neq v_j$ if $i \neq j$, and $(v_i, v_{i+1}) \in E(H)$ for $1 \leq i \leq k$, where $v_{k+1} = v_1$. For a graph H, let $g(H)$ be the length of the shortest cycle in H; this is called the *girth* of the graph.

Let (G_i) be a sequence of d-regular graphs such that $g(G_i) \to \infty$. Then (G_i) Benjamini–Schramm converges to the rooted d-regular infinite tree $\mathbb{T}_d$. Note that if in a finite graph G the shortest cycle has length at least $2k + 1$ then the k-neighborhood of any vertex looks like the k-neighborhood of any vertex of an infinite d-regular tree.

Let (G_i) be a sequence of (a, b)-biregular graphs such that $g(G_i) \to \infty$. Then (G_i) Benjamini–Schramm converges to the following distribution: with probability $\frac{a}{a+b}$ it is the infinite (a, b)-biregular tree $\mathbb{T}_{a,b}$ with root vertex of degree b, and with probability $\frac{b}{a+b}$ it is the infinite (a, b)-biregular tree $\mathbb{T}_{a,b}$ with root vertex of degree a. With slight abuse of notation we will denote this random rooted tree by $\mathbb{T}_{a,b}$ as well.

Remark 2.8. Not every random rooted graph can be obtained as a limit of Benjamini–Schramm convergent finite graphs. A necessary condition is that the random rooted graph be *unimodular*, which is a certain reversibility property of the graph. On the other hand, it is not known whether every unimodular random graph can be obtained as a limit of Benjamini–Schramm convergent finite graphs. This is the famous Aldous–Lyons problem. The interested reader can consult the book [21].

The following theorem was known in many cases for the thermodynamic limit in statistical mechanics. We also note that a modification of the algorithm "CountMATCHINGS" in [3] yields an alternative proof of parts (a) and (b) of this theorem.

Theorem 2.9 ([2]). *Let (G_i) be a Benjamini–Schramm convergent graph sequence of bounded degree graphs. Then the sequences of functions*

(a) $p(G_i, t)$,
(b) $(\ln M(G_i, t))/v(G_i)$

converge to strictly increasing continuous functions on the interval $[0, \infty)$. *Let p_0 be a real number between* 0 *and* 1 *such that $p^*(G_i) \geq p_0$ for all n. Then*

(c) $t(G_i, p)$,
(d) $\lambda_{G_i}(p)$

are convergent for all $0 \leq p < p_0$.

Remark 2.10. H. Nguyen and K. Onak [22], and independently G. Elek and G. Lippner [8], proved that for a Benjamini–Schramm convergent graph sequence (G_i), the following limit exists:

$$\lim_{n\to\infty} p^*(G_i).$$

(Recall that $p^*(G_i) = 2\nu(G_i)/v(G_i)$.)

To prove Theorem 2.9, we need some preparation. We essentially repeat the argument from [1].

The following theorem deals with the behavior of the matching polynomial in Benjamini–Schramm convergent graph sequences. The matching measure was introduced in [1]:

Definition 2.11. The *matching measure* of a finite graph is defined as

$$\rho_G = \frac{1}{v(G)} \sum_{z_i : \mu(G, z_i)=0} \delta(z_i),$$

where $\delta(s)$ is the Dirac-delta measure at s, and we take every z_i into account together with its multiplicity.

In other words, the matching measure is the probability measure of uniform distribution on the zeros of $\mu(G, x)$.

Theorem 2.12 ([1, 2]). *Let (G_i) be a Benjamini–Schramm convergent bounded degree graph sequence. Let ρ_{G_i} be the matching measure of the graph G_i. Then the sequence (ρ_{G_i}) is weakly convergent, i.e., there exists some measure ρ_L such that for every bounded continuous function f, we have*

$$\lim_{i\to\infty} \int f(z)\, d\rho_{G_i}(z) = \int f(z)\, d\rho_L(z).$$

Remark 2.13. This theorem was first proved in [1]. The proof given there relied on a general result on graph polynomials given in [5]. To make this paper as self-contained as possible we sketch here a slightly different proof outlined in a remark in [1].

Proof of Theorem 2.12. For a graph G let $S(G)$ denote the multiset of zeros of the matching polynomial $\mu(G,x)$, and

$$p_k(G) = \sum_{\lambda \in S(G)} \lambda^k.$$

Then $p_k(G)/v(G)$ can be rewritten in terms of the measure ρ_G as follows:

$$\frac{p_k(G)}{v(G)} = \int z^k \, d\rho_G(z).$$

It is known that $p_k(G)$ counts the number of closed tree-like walks of length k in the graph G (see [10, Chapter 6]). Without going into the details of the description of tree-like walks, we only use the fact that these are a special type of walks that we can count by knowing all k-balls centered at the vertices of the graph G. Let $TW(\alpha)$ denote the number of closed tree-like walks starting at the root of α, and let $\mathcal{N}_k$ be the set of k-neighborhoods α. The size of $\mathcal{N}_k$ is bounded by a function of k and the largest degree of the graph G. Furthermore, let $N_k(G,\alpha)$ denote the number of vertices of G whose k-neighborhood is isomorphic to α. Then

$$p_k(G) = \sum_{\alpha \in \mathcal{N}_k} N_k(G,\alpha) \cdot TW(\alpha).$$

Therefore

$$\frac{p_k(G)}{v(G)} = \sum_{\alpha \in \mathcal{N}_k} \mathbb{P}(G,\alpha,k) \cdot TW(\alpha).$$

Hence, if (G_i) is Benjamini–Schramm convergent then for every fixed k, the sequence

$$\frac{p_k(G_i)}{v(G_i)} = \int z^k \, d\rho_{G_i}(z)$$

is convergent. Clearly, this implies that for every polynomial $q(z)$, the sequence

$$\int q(z) \, d\rho_{G_i}(z)$$

is convergent.

Assume that D is a general upper bound for all degrees of all graphs G_i. Then all zeros of $\mu(G_i,x)$ lie in the interval $[-2\sqrt{D-1}, 2\sqrt{D-1}]$. Since every continuous function on a bounded interval can be uniformly approximated by polynomials, we conclude that the sequence (ρ_{G_i}) is weakly convergent. □

Proof of Theorem 2.9. First we prove parts (a) and (b). For a graph G let $S(G)$ denote the set of zeros of the matching polynomial $\mu(G,x)$. Then

$$M(G,t) = \prod_{\substack{\lambda \in S(G) \\ \lambda > 0}} (1+\lambda^2 t) = \prod_{\lambda \in S(G)} (1+\lambda^2 t)^{1/2}.$$

Then

$$\ln M(G,t) = \sum_{\lambda \in S(G)} \tfrac{1}{2} \ln(1+\lambda^2 t).$$

By differentiating both sides with respect to t we get

$$\frac{\frac{d}{dt} M(G,t)}{M(G,t)} = \sum_{\lambda \in S(G)} \frac{1}{2} \frac{\lambda^2}{1+\lambda^2 t}.$$

Hence

$$p(G,t) = \frac{2t \cdot \frac{d}{dt} M(G,t)}{v(G) \cdot M(G,t)} = \frac{1}{v(G)} \sum_{\lambda \in S(G)} \frac{\lambda^2 t}{1+\lambda^2 t} = \int \frac{tz^2}{1+tz^2} \, d\rho_G(z).$$

Similarly,

$$\frac{\ln M(G,t)}{v(G)} = \frac{1}{v(G)} \sum_{\lambda \in S(G)} \tfrac{1}{2} \ln(1+\lambda^2 t) = \int \tfrac{1}{2} \ln(1+tz^2) \, d\rho_G(z).$$

Since (G_i) is a Benjamini–Schramm convergent bounded degree graph sequence, the sequence (ρ_{G_i}) weakly converges to some ρ_L by Theorem 2.12. Since both functions

$$\frac{tz^2}{1+tz^2} \quad \text{and} \quad \tfrac{1}{2} \ln(1+tz^2)$$

are continuous, we immediately see that

$$\lim_{n \to \infty} p(G_i,t) = \int \frac{tz^2}{1+tz^2} \, d\rho_L(z)$$

and

$$\lim_{n \to \infty} \frac{\ln M(G_i,t)}{v(G_i)} = \int \tfrac{1}{2} \ln(1+tz^2) \, d\rho_L(z).$$

Note that both $\frac{tz^2}{1+tz^2}$ and $\frac{1}{2}\ln(1+tz^2)$ are strictly increasing continuous functions of t. Thus their integrals are also strictly increasing continuous functions.

To prove part (c), let us introduce the function

$$p(L,t) = \int \frac{tz^2}{1+tz^2} \, d\rho_L(z).$$

We have seen that $p(L,t)$ is a strictly increasing continuous function, and $\lim_{i \to \infty} p(G_i,t) = p(L,t)$. Since $p^*(G_i) \geq p_0$ for all G_i, we have $\lim_{t \to \infty} p(G_i,t) \geq p_0$ for all i. This means that $\lim_{t \to \infty} p(L,t) \geq p_0$. Hence we can consider the inverse function of $p(L,t)$ which maps $[0, p_0)$ into $[0, \infty)$; let us call it $t(L,p)$.

We will show that

$$\lim_{i \to \infty} t(G_i,p) = t(L,p)$$

pointwise for $p < p_0$. Assume this is not the case. This means that for some p_1, there exists an ε and an infinite sequence n_i for which

$$|t(L, p_1) - t(G_{n_i}, p_1)| \geq \varepsilon.$$

We distinguish two cases:

(i) there exists an infinite sequence (n_i) for which $t(G_{n_i}, p_1) \geq t(L, p_1) + \varepsilon$,
(ii) there exists an infinite sequence (n_i) for which $t(G_{n_i}, p_1) \leq t(L, p_1) - \varepsilon$.

In the first case, let $t_1 = t(L, p_1)$, $t_2 = t_1 + \varepsilon$ and $p_2 = p(L, t_2)$. Clearly, $p_2 > p_1$. Note that

$$t(G_{n_i}, p_1) \geq t(L, p_1) + \varepsilon = t_2$$

and $p(G_{n_i}, t)$ are increasing functions, thus

$$p(G_{n_i}, t_2) \leq p(G_{n_i}, t(G_{n_i}, p_1)) = p_1 = p_2 - (p_2 - p_1) = p(L, t_2) - (p_2 - p_1).$$

This contradicts the fact that $\lim_{n_i \to \infty} p(G_{n_i}, t_2) = p(L, t_2)$.

In the second case, let $t_1 = t(L, p_1)$, $t_2 = t_1 - \varepsilon$ and $p_2 = p(L, t_2)$. Clearly, $p_2 < p_1$. Since $t(G_{n_i}, p_1) \leq t(L, p_1) - \varepsilon = t_2$ and $p(G_{n_i}, t)$ are increasing functions, we have

$$p(G_{n_i}, t_2) \geq p(G_{n_i}, t(G_{n_i}, p_1)) = p_1 = p_2 + (p_1 - p_2) = p(L, t_2) + (p_1 - p_2).$$

This again contradicts the fact that $\lim_{n \to \infty} p(G_{n_i}, t_2) = p(L, t_2)$. Consequently, we have $\lim_{i \to \infty} t(G_i, p) = t(L, p)$.

Finally, we show that $\lambda_{G_i}(p)$ converges for all p. Let $t = t(L, p)$, and

$$\lambda_L(p) = \lim_{i \to \infty} \frac{\ln M(G_i, t)}{v(G_i)} - \tfrac{1}{2} p \ln t.$$

Note that

$$\lambda_{G_i}(p) = \frac{\ln M(G_i, t_i)}{v(G_i)} - \tfrac{1}{2} p \ln t_i,$$

where $t_i = t(G_i, p)$. We have seen that $\lim_{i \to \infty} t_i = t$. Hence it is enough to prove that the functions $(\ln M(G_i, u))/v(G_i)$ are equicontinuous. Fix some positive u_0 and let

$$R(u_0, u) = \max_{z \in [-2\sqrt{D-1}, 2\sqrt{D-1}]} \left| \tfrac{1}{2} \ln(1 + u_0 z^2) - \tfrac{1}{2} \ln(1 + u z^2) \right|.$$

Clearly, if $|u - u_0| \leq \delta$ for some sufficiently small δ, then $R(u_0, u) \leq \varepsilon$, and

$$\left| \frac{\ln M(G_i, u)}{v(G_i)} - \frac{\ln M(G_i, u_0)}{v(G_i)} \right| = \left| \int \tfrac{1}{2} \ln(1 + u_0 z^2)\, d\rho_{G_i}(z) - \int \tfrac{1}{2} \ln(1 + u z^2)\, d\rho_{G_i}(z) \right|$$
$$\leq \int \left| \tfrac{1}{2} \ln(1 + u_0 z^2) - \tfrac{1}{2} \ln(1 + u z^2) \right| d\rho_{G_i}(z) \leq \int R(u, u_0)\, d\rho_{G_i}(z) \leq \varepsilon.$$

This completes the proof of the convergence of $\lambda_{G_i}(p)$. □

Now it is easy to define these functions for those random rooted graphs which can be obtained as a Benjamini–Schramm limit of finite graphs.

Definition 2.14. Let L be a random rooted graph which can be obtained as the Benjamini–Schramm limit of finite graphs (G_i) of bounded degree. Assume that $p^*(G_i) \geq p_0$ for all n. Let

$$p(L,t) = \lim_{n\to\infty} p(G_i,t), \qquad F(L,t) = \lim_{n\to\infty} \frac{\ln M(G_i,t)}{v(G_i)}$$

for all $t \geq 0$, and

$$t(L,p) = \lim_{n\to\infty} t(G_i,p), \qquad \lambda_L(p) = \lim_{n\to\infty} \lambda_{G_i}(p)$$

for all $p < p_0$. Finally, let

$$\lambda_L(p_0) = \lim_{p\nearrow p_0} \lambda_L(p).$$

Note that the functions $p(L,t), F(L,t), t(L,p)$ and $\lambda_L(p)$ are well-defined in the sense that if the sequences (G_i) and (H_i) both Benjamini–Schramm converge to L and $p^*(G_i), p^*(H_i) \geq p_0$ for all i, then they define the same functions on $[0,\infty)$ or $[0,p_0]$. Indeed, we can consider the two sequences together and apply Theorem 2.9 to find that the limits do not depend on the choice of the sequence. From the proof of Theorem 2.9, we also see that $p(L,t)$ and $F(L,t)$ can be expressed as integrals against a certain measure ρ_L.

2.2. *Entropy and density function of the infinite d-regular tree $\mathbb{T}_d$*

In this section we give the entropy and density functions of the d-regular and (a,b)-biregular trees.

Theorem 2.15. *Let $\mathbb{T}_d$ be the infinite d-regular tree. Then*

(a) $$p(\mathbb{T}_d,t) = \frac{2d^2t + d - d\sqrt{1+4(d-1)t}}{2d^2t+2},$$

(b) $$\int \tfrac{1}{2}\ln(1+tz^2)\,d\rho_{\mathbb{T}_d}(z) = \tfrac{1}{2}\ln S_d(t),$$

where

$$S_d(t) = \frac{1}{\eta_t^2}\left(\frac{d-1}{d-\eta_t}\right)^{d-2}, \qquad \eta_t = \frac{\sqrt{1+4(d-1)t}-1}{2(d-1)t},$$

(c) $$t(\mathbb{T}_d,p) = \frac{p(d-p)}{d^2(1-p)^2},$$

(d) $$\lambda_{\mathbb{T}_d}(p) = \mathbb{G}_d(p) = \frac{1}{2}\left(p\ln\left(\frac{d}{p}\right) + (d-p)\ln\left(1-\frac{p}{d}\right) - 2(1-p)\ln(1-p)\right).$$

Theorem 2.16. *Let $\mathbb{T}_{a,b}$ be the infinite (a,b)-biregular tree. Then for $0 \leq p \leq \min\left(\frac{2a}{a+b}, \frac{2b}{a+b}\right)$ we have*

(a) $$p(\mathbb{T}_{a,b}, t) = \frac{2abt + \frac{2ab}{a+b} - \frac{2ab}{a+b}\sqrt{1 + (2a+2b-4)t + (a-b)^2t^2}}{2abt + 2},$$

(b) $$t(\mathbb{T}_{a,b}, p) = \frac{a+b}{2ab} \frac{p\left(1 - \frac{a+b}{2ab}p\right)}{\left(1 - \frac{a+b}{2a}p\right)\left(1 - \frac{a+b}{2b}p\right)},$$

$$\lambda_{\mathbb{T}_{a,b}}(p) = \mathbb{G}_{a,b}(p) = \frac{a}{a+b}H\left(\frac{a+b}{2a}p\right) + \frac{b}{a+b}H\left(\frac{a+b}{2b}p\right) + \frac{1}{2}p\ln(ab) - \frac{ab}{a+b}H\left(\frac{a+b}{2ab}p\right),$$

where $H(q) = -(q\ln q + (1-q)\ln(1-q))$.

There are two essentially different proofs for Theorems 2.15 and 2.16. We detail the first proof, and in the next subsection we sketch a second one.

The first proof of Theorems 2.15 and 2.16 roughly follows the arguments of Section 4 of [1]. For an (infinite) tree, the spectral measure and the matching measure coincide. This can be proved via [1, Lemma 4.2], or an even simpler proof is that for trees, the number of closed walks and the number of closed tree-like walks are the same, so the moment sequences of the spectral measure and the matching measure coincide, and since they are supported on a bounded interval, they must be the same measure. For the d-regular tree $\mathbb{T}_d$, this is the Kesten–McKay measure given by the density function

$$f_d(x) = \frac{d\sqrt{4(d-1) - x^2}}{2\pi(d^2 - x^2)}\chi_{[-2\sqrt{d-1}, 2\sqrt{d-1}]}.$$

For the (a,b)-biregular infinite tree, the matching or spectral measure $\rho_{\mathbb{T}_{a,b}}$ is given by

$$d\rho_{\mathbb{T}_{a,b}} = \frac{|a-b|}{a+b}\delta_0 + \frac{ab\sqrt{-(x^2 - ab + (s-1)^2)(x^2 - ab + (s+1)^2)}}{\pi(a+b)(ab - x^2)|x|} \times \chi_{\{|\sqrt{a-1} - \sqrt{b-1}| \leq |x| \leq \sqrt{a-1} + \sqrt{b-1}\}}\, dx,$$

where $s = \sqrt{(a-1)(b-1)}$. As a next step one might try to compute the integral of the functions

$$\frac{tz^2}{1+tz^2} \quad \text{and} \quad \tfrac{1}{2}\ln(1+tz^2)$$

to obtain $p(\mathbb{T}_{a,b}, t)$ and $F(\mathbb{T}_{a,b}, t)$. We will slightly modify this argument to simplify it. Our modification shows that we do not need to compute these integrals—we can work directly with the moment sequences which are simply the numbers of closed walks in the corresponding trees. More precisely, in $\mathbb{T}_{a,b}$ we have to weight the number of closed walks starting and ending at a root vertex of degree a with weight $b/(a+b)$, and the number of closed walks starting and ending at a root vertex of degree b with weight $a/(a+b)$.

First of all, we need the following lemma on the number of closed walks in $\mathbb{T}_{a,b}$. We are sure that it is well-known, but since we have not been able to find any reference, we give its proof.

Lemma 2.17. *Let W_j^a and W_j^b be the number of closed walks of length j starting at and returning to a root vertex of $\mathbb{T}_{a,b}$ of degree a and of degree b, respectively. Then for the generating function we have*

$$G_a(z) := \sum_{j=0}^{\infty} W_j^a z^j = \frac{1}{1 - az^2 F_b(z)},$$

where

$$F_b(z) = \frac{1 + (b-a)z^2 - \sqrt{1 - (2a+2b-4)z^2 + (b-a)^2 z^4}}{2(a-1)z^2}.$$

Similarly,

$$G_b(z) := \sum_{j=0}^{\infty} W_j^b z^j = \frac{1}{1 - bz^2 F_a(z)},$$

where

$$F_a(z) = \frac{1 + (a-b)z^2 - \sqrt{1 - (2a+2b-4)z^2 + (b-a)^2 z^4}}{2(b-1)z^2}.$$

Proof. Consider the rooted tree $\mathbb{T}_{a,b}^a$, where the only difference compared to $\mathbb{T}_{a,b}$ is that the root vertex has degree $a-1$ and not a. Similarly, let $\mathbb{T}_{a,b}^b$ be the rooted tree where the only difference compared to $\mathbb{T}_{a,b}$ is that the root vertex has degree $b-1$ and not b. Let $\overline{W}_j^a$ be the number of closed walks of length j starting at and returning to the root vertex of $\mathbb{T}_{a,b}^a$. Furthermore, let $\overline{U}_j^a$ be the number of closed walks of length j starting at and returning to the root vertex of $\mathbb{T}_{a,b}^a$ such that the walk only visits the root at the beginning and at the end, and the walk has length at least 2, so it is not the empty walk. We can similarly define $\overline{W}_j^b$ and $\overline{U}_j^b$. Let

$$F_a(z) = \sum_{j=0}^{\infty} \overline{W}_j^a z^j, \qquad F_b(z) = \sum_{j=0}^{\infty} \overline{W}_j^b z^j,$$

and

$$R_a(z) = \sum_{j=1}^{\infty} \overline{U}_j^a z^j, \qquad R_b(z) = \sum_{j=1}^{\infty} \overline{U}_j^b z^j.$$

First of all,

$$F_a(z) = 1 + R_a(z) + R_a(z)^2 + R_a(z)^3 + \cdots = \frac{1}{1 - R_a(z)},$$

since every closed walk can be uniquely decomposed into walks which visit the root only at the beginning and at the end. Similarly,

$$F_b(z) = \frac{1}{1 - R_b(z)}.$$

Finally,

$$R_a(z) = (a-1)z^2 F_b(z) \quad \text{and similarly} \quad R_b(z) = (b-1)z^2 F_a(z),$$

since every closed walk which visits the root only at the beginning and at the end can be decomposed in the following way: we erase the first and last steps (we can choose these in $a-1$ different ways in $\mathbb{T}^a_{a,b}$), and we get a closed walk in $\mathbb{T}^b_{a,b}$. Solving these equations we get

$$F_a(z) = \frac{1+(b-a)z^2 - \sqrt{1-(2a+2b-4)z^2+(b-a)^2z^4}}{2(b-1)z^2},$$
$$F_b(z) = \frac{1+(a-b)z^2 - \sqrt{1-(2a+2b-4)z^2+(b-a)^2z^4}}{2(a-1)z^2},$$
$$R_a(z) = \tfrac{1}{2}\big(1+(a-b)z^2 - \sqrt{1-(2a+2b-4)z^2+(b-a)^2z^4}\big),$$
$$R_b(z) = \tfrac{1}{2}\big(1+(b-a)z^2 - \sqrt{1-(2a+2b-4)z^2+(b-a)^2z^4}\big).$$

(Note that at some point, we have to solve a quadratic equation, and we can choose only the minus sign, because of the evaluation of the generating function at $z=0$.)

Now let us go back to the original problem. Let U^a_j be the number of closed walks of length j starting at and returning to the root vertex of $\mathbb{T}_{a,b}$ of degree a such that the walk only visits the root at the beginning and at the end, and the walk has length at least 2, so it is not the empty walk. We define U^b_j similarly. Let

$$G_a(z) = \sum_{j=0}^{\infty} W^a_j z^j, \qquad G_b(z) = \sum_{j=0}^{\infty} W^b_j z^j,$$

and

$$H_a(z) = \sum_{j=1}^{\infty} U^a_j z^j, \qquad H_b(z) = \sum_{j=1}^{\infty} U^b_j z^j.$$

As before,

$$G_a(z) = \frac{1}{1-H_a(z)}, \qquad G_b(z) = \frac{1}{1-H_b(z)}.$$

Finally,

$$H_a(z) = az^2 F_b(z) \quad \text{and similarly} \quad R_b(z) = bz^2 F_a(z),$$

since every closed walk which visits the root only at the beginning and at the end can be decomposed in the following way: we erase the first and last steps (we can choose these in a different ways in $\mathbb{T}_{a,b}$), and we get a closed walk in $\mathbb{T}^b_{a,b}$. Hence

$$G_a(z) = \frac{1}{1-az^2F_b(z)}, \qquad G_b(z) = \frac{1}{1-bz^2F_a(z)}. \qquad \square$$

Proof of Theorems 2.15 and 2.16. Since Theorem 2.15 is a special case of Theorem 2.16, we concentrate on the proof of the latter. We only need to work with part (a), since then (b) follows immediately, and (c) follows from (b) by using

$$\frac{d}{dp}\lambda_{\mathbb{T}_{a,b}}(p) = -\tfrac{1}{2}\ln t(\mathbb{T}_{a,b}, p).$$

These are routine computations, left to the reader.

To prove (a), first assume that $|t| < 1/(4(\max(a,b)-1))$. Note that for such t, all subsequent series are converging. We have

$$\begin{aligned} p(\mathbb{T}_{a,b}, t) &= \int \frac{tu^2}{1+tu^2}\, d\rho_{\mathbb{T}_{a,b}}(u) = \int \sum_{j=1}^{\infty}(-1)^{j+1}t^j u^{2j}\, d\rho_{\mathbb{T}_{a,b}}(u) \\ &= \sum_{j=1}^{\infty}(-1)^{j+1}t^j \int u^{2j}\, d\rho_{\mathbb{T}_{a,b}}(u). \end{aligned}$$

Note that

$$\int u^{2j}\, d\rho_{\mathbb{T}_{a,b}}(u) = \frac{b}{a+b}W_{2j}^{a} + \frac{a}{a+b}W_{2j}^{b}.$$

Hence

$$p(\mathbb{T}_{a,b}, -z^2) = 1 - \left(\frac{b}{a+b}G_a(z) + \frac{a}{a+b}G_b(z)\right).$$

After some calculation we get

$$p(\mathbb{T}_{a,b}, t) = \frac{2abt + \frac{2ab}{a+b} - \frac{2ab}{a+b}\sqrt{1+(2a+2b-4)t+(a-b)^2t^2}}{2abt+2}$$

for $|t| < 1/(4(\max(a,b)-1))$. On the other hand, both functions appearing in the previous equation are holomorphic in $\{t \mid |\Im(t)| \le \Re(t)\}$, so they must coincide everywhere in this region. □

2.3. *Random graphs*

The goal of this subsection is twofold. On the one hand, we show that Theorems 1.5 and 1.7 are quite precise, for instance if p is separated away from 1 then Theorem 1.5 is the best possible up to a constant factor. On the other hand, we would also like to show a connection between random (bi)regular random graphs and the entropy function of an infinite (bi)regular tree.

An alternative way to obtain Theorems 2.15 and 2.16 is the following. We can use Theorem 2.9 to obtain the required functions by choosing an appropriate Benjamini–Schramm convergent graph sequence. It turns out that it is sufficient to consider random d-regular or (a,b)-biregular bipartite graphs. Indeed, one can compute the expected number of k-matchings of a random d-regular or (a,b)-biregular bipartite graph quite easily. Such a computation was carried out in [4, 9, 26, 12] for d-regular bipartite graphs and it easily generalizes to (a, b)-biregular bipartite graphs. We also note that a random (a, b)-biregular bipartite graph contains a very small number of short cycles. This is a classical result for random regular graphs, but it is also known for biregular bipartite graphs [7].

First of all, let us specify which biregular random graph model we use. Let the vertex set of the random graph be $V \cup W$, where $V = \{v_1, \dots, v_{an}\}$ and $W = \{w_1, \dots, w_{bn}\}$. Consider two random partitions of the set $\{1, \dots, abn\}$: the first one is $P_1 = \{A_1, \dots, A_{an}\}$ where each set has size b, and the second is $P_2 = \{B_1, \dots, B_{bn}\}$ where each set has size a. Then for every $k \in \{1, \dots, abn\}$ connect v_i and w_j if $k \in A_i \cap B_j$. This is the configuration model. Note that this model allows multiple edges, but this is not a problem for us. In the special case when $a = b = d$ we can choose V and W to be of size n. The following theorem was proved in [9].

Theorem 2.18 ([9]). *Let G be chosen from the set of labelled d-regular bipartite graphs on $v(G) = 2n$ vertices according to the configuration model. Then*

$$\mathbb{E}m_k(G) = \binom{n}{k}^2 d^{2k} \frac{1}{\binom{dn}{k}}.$$

The corollary of this theorem is the second part of Theorem 1.5.

Corollary 2.19. *Let $p = k/n$. There exists a d-regular bipartite graph G on $2n$ vertices such that*

$$m_k(G) \leq \sqrt{\frac{1 - p/d}{1 - p}} \cdot p_\mu \cdot \exp(2n\mathbb{G}_d(p)).$$

Proof. We will show that

$$E = \binom{n}{k}^2 d^{2k} \frac{1}{\binom{dn}{k}} \leq \sqrt{\frac{1 - p/d}{1 - p}} \cdot p_\mu \cdot \exp(2n\mathbb{G}_d(p)).$$

Note that

$$E = \binom{n}{k}^2 d^{2k} \frac{1}{\binom{dn}{k}} = \frac{\binom{n}{k}}{\binom{dn}{k}} d^{2k} \binom{n}{k} = \frac{\binom{n}{k}}{\binom{dn}{k}} d^{2k} \cdot \frac{p_\mu}{p^k(1-p)^{n-k}}.$$

For the first term we use Stirling's formula. Let Θ_m be defined by the following form of Stirling's formula:

$$m! = \sqrt{2\pi m}(m/e)^m e^{\Theta_m}.$$

It is known (see [24]) that

$$\frac{1}{12m+1} \leq \Theta_m \leq \frac{1}{12m}.$$

Then

$$E = \frac{1}{\sqrt{2\pi k(n-k)/n}} e^{\Theta_n - \Theta_k - \Theta_{n-k}} \sqrt{2\pi \frac{k(dn-k)}{dn}} e^{-\Theta_{dn} + \Theta_k + \Theta_{dn-k}} \cdot p_\mu \cdot \exp(2n\mathbb{G}_d(p))$$
$$= \sqrt{\frac{1 - p/d}{1 - p}} e^{\Theta_n - \Theta_{n-k} - \Theta_{dn} + \Theta_{dn-k}} \cdot p_\mu \cdot \exp(2n\mathbb{G}_d(p)).$$

Thus we only need to show that

$$\Theta_n - \Theta_{n-k} - \Theta_{dn} + \Theta_{dn-k} \leq 0.$$

This is indeed true:

$$\begin{aligned}\Theta_n - \Theta_{n-k} - \Theta_{dn} + \Theta_{dn-k} &\leq \frac{1}{12n} - \frac{1}{12(n-k)+1} - \frac{1}{12dn+1} + \frac{1}{12(dn-k)} \\ &= \frac{-(12k-1)}{12n(12(n-k)+1)} + \frac{12k+1}{(12dn+1)(12(dn-k)+1)} \leq 0\end{aligned}$$

if $d \geq 2$. □

The following lemma is a straightforward extension of the previous results to (a,b)-biregular bipartite graphs.

Lemma 2.20. *Let G be chosen from the set of labelled (a,b)-biregular bipartite graphs on $v(G) = (a+b)n$ vertices according to the configuration model. Then*

(a) $\mathbb{E}m_k(G) = \exp\big(v(G)(\mathbb{G}_{a,b}(p) + o_{v(G)}(1))\big)$, *where* $p = 2k/v(G)$.

(b) ([7]) *Let $c_{2j}(G)$ be the number of $2j$-cycles in the graph G. Then*

$$\mathbb{E}c_{2j}(G) = \frac{((a-1)(b-1))^j}{2j}(1 + o_{v(G)}(1)).$$

Proof. (a) Note that the number of all partition pairs (P_1, P_2) is

$$N = \frac{(abn)!}{a!^{bn}} \cdot \frac{(abn)!}{b!^{an}}.$$

The number of possible k-matchings is

$$U_k = \binom{abn}{k}\binom{an}{k}\binom{bn}{k}k!^2.$$

If we fix one k-matching then we need to repartition the remaining $abn-k$ elements into sets of sizes a and $a-1$, and b and $b-1$. This can be done in

$$V_k = \frac{(abn-k)!}{(a-1)!^k a!^{bn-k}} \cdot \frac{(abn-k)!}{(b-1)!^k b!^{an-k}}$$

ways. Hence

$$\mathbb{E}m_k(G) = \frac{1}{N}U_k V_k = \binom{an}{k}\binom{bn}{k}(ab)^k \frac{1}{\binom{abn}{k}}.$$

Then by the usual approximation of binomial coefficients we get

$$\mathbb{E}m_k(G) = \exp\big(v(G)(\mathbb{G}_{a,b}(p) + o_{v(G)}(1))\big),$$

where $p = 2k/v(G)$.

(b) We can choose the possible cycles in

$$T_j = \binom{abn}{2j}\binom{an}{j}\binom{bn}{j}(2j-1)!j!^2$$

different ways. (We can choose the "edges" and vertices in $\binom{abn}{2j}\binom{an}{j}\binom{bn}{j}$ ways, then we choose an ordering on the edges, and on each vertex set, and we connect the vertices and "edges" along the orderings. Finally, we divide by $2j$ since we counted each cycle in $2j$ ways.) Next we need to repartition the remaining $abn - 2j$ elements into sets of sizes a and $a-2$, and b and $b-2$. This can be done in

$$S_j = \frac{(abn-2j)!}{(a-2)!^j a!^{bn-j}} \cdot \frac{(abn-2j)!}{(b-2)!^j b!^{an-j}}$$

ways. Hence

$$\mathbb{E}c_{2j}(G) = \frac{1}{N} T_j S_j = \frac{((a-1)(b-1))^j}{2j}(1 + o_{v(G)}(1)). \qquad \square$$

Part (b) of Lemma 2.20 shows that the expected number of cycles of length $2j$ is bounded independently of the size of the graph. Note that the (a, b)-biregular graph sequence (G_i) Benjamini–Schramm converges to $\mathbb{T}_{a,b}$ if for all fixed j we have $c_{2j}(G_i) = o(v(G_i))$. Note that by Markov's inequality,

$$\mathbb{P}\big(m_k(G) > 3\mathbb{E}m_k(G)\big) \le \frac{1}{3} \quad \text{and} \quad \mathbb{P}\big(c_{2j}(G) > 3g\mathbb{E}c_{2j}(G)\big) \le \frac{1}{3g}$$

for $j = 1, \ldots, g$. Hence for any large enough n and fixed g, with probability at least $1/3$ we can choose a graph G_i on $(a+b)n$ vertices such that G_i has a bounded number of cycles of length at most $2g$ and $m_k(G_i) \le 3\exp(v(G)(\mathbb{G}_{a,b}(p) + o_{v(G)}(1)))$. This shows that we can choose a sequence (G_i) of graphs converging to $\mathbb{T}_{a,b}$ such that

$$\frac{\ln m_k(G_i)}{v(G_i)} + o_{v(G_i)}(1) = \lambda_{G_i}(p) \le \mathbb{G}_{a,b}(p) + o_{v(G_i)}(1).$$

This implies that

$$\lambda_{\mathbb{T}_{a,b}}(p) \le \mathbb{G}_{a,b}(p).$$

Note that we have only proved this inequality for rational p, but then it follows for all p by continuity.

Unfortunately, with this idea we have not been able to establish the inequality $\lambda_{\mathbb{T}_{a,b}}(p) \ge \mathbb{G}_{a,b}(p)$. The problem is the following. In principle, it can occur that a typical random graph has a much smaller (exponentially smaller) number of k-matchings than the expected value, and a large contribution to the expected value comes from graphs having a large number of short cycles and matchings. Note that Theorem 1.7 implies that this cannot occur, but we cannot use this result as it would result in a cycle in the proof of this theorem. Instead we propose a conjecture which would imply the inequality $\lambda_{\mathbb{T}_{a,b}}(p) \ge \mathbb{G}_{a,b}(p)$.

Conjecture 2.21. There exists a constant C independent of n and k such that

$$\mathbb{E}m_k(G)^2 \le C(\mathbb{E}m_k(G))^2.$$

Note that this conjecture is known to be true for perfect matchings in regular random graphs [4]. To show that this conjecture implies $\lambda_{\mathbb{T}_{a,b}}(p) \ge \mathbb{G}_{a,b}(p)$, we need the following proposition.

Proposition 2.22. *Let X be a non-negative random variable such that for some positive constant C we have*

$$\mathbb{P}(X > C\mathbb{E}X) \le \frac{1}{16C} \quad \textit{and} \quad \mathbb{E}X^2 \le C(\mathbb{E}X)^2.$$

Then

$$\mathbb{P}\big(\tfrac{1}{4}\mathbb{E}X \le X \le C\mathbb{E}X\big) \ge \frac{1}{2C}.$$

Proof. Let $A = \{\omega \mid X(\omega) < \frac{1}{4}\mathbb{E}X\}$, $B = \{\omega \mid \frac{1}{4}\mathbb{E}X \le X(\omega) \le C\mathbb{E}X\}$, and $D = \{\omega \mid X(\omega) > C\mathbb{E}X\}$. Then

$$\int_A X\,dP \le \tfrac{1}{4}\mathbb{E}X.$$

Furthermore,

$$\mathbb{P}(D)\cdot\mathbb{E}X^2 \ge \mathbb{P}(D)\cdot\int_D X^2\,dP = \int_D 1\,dP\cdot\int_D X^2\,dP \ge \left(\int_D X\,dP\right)^2.$$

Hence

$$\frac{1}{16C}C(\mathbb{E}X)^2 \ge \mathbb{P}(D)\cdot\mathbb{E}X^2 \ge \left(\int_D X\,dP\right)^2.$$

In other words, $\int_D X\,dP \le \frac{1}{4}\mathbb{E}X$. This implies that

$$\int_B X\,dP \ge \tfrac{1}{2}\mathbb{E}X.$$

Since

$$\int_B X\,dP \le \mathbb{P}(B)C\mathbb{E}X,$$

the claim of the proposition follows immediately. □

Fix a positive number g, and call a graph *typical* if

$$c_{2j}(G) < 16Cg\mathbb{E}c_{2j}(G)$$

for $j = 1, \dots, g$. Note that a typical graph has a bounded number of short cycles, and by Markov's inequality, the probability that a graph is typical is at least $1 - 1/(16C)$. First case: there is a typical graph G such that $m_k(G) > C\mathbb{E}m_k(G)$, then we are done, because $\lambda_G(p) \geq \mathbb{G}_{a,b}(p) + o(1)$. Second case: there is no typical graph with $m_k(G) > C\mathbb{E}m_k(G)$; then the proposition implies that

$$\mathbb{P}\big(\tfrac{1}{4}\mathbb{E}m_k(G) \leq m_k(G) \leq C\mathbb{E}m_k(G)\big) \geq \frac{1}{2C}.$$

Since the probability that a graph is typical is at least $1 - 1/(16C)$, we see that there are typical graphs for which

$$m_k(G) \geq \tfrac{1}{4}\mathbb{E}m_k(G),$$

implying again that $\lambda_G(p) \geq \mathbb{G}_{a,b}(p) + o(1)$. Hence we can choose a sequence of typical graphs to show that $\lambda_{\mathbb{T}_{a,b}}(p) \geq \mathbb{G}_{a,b}(p)$.

In spite of the fact that this proof did not lead to another proof of Theorem 2.16, we feel that it was instructive to carry out these computations as they showed that Theorems 1.5 and 1.7 are tight. This was known for perfect matchings of d-regular random graphs [4, 26], and for matchings of arbitrary size [9]. Our computation for biregular bipartite graphs is the natural counterpart of these results.

3. New version of Gurvits's theorem

In this section we prove the following theorem.

Theorem 3.1. *The following two statements are equivalent:*

(i) *For any d-regular bipartite graph G on $2n$ vertices, we have*

$$\frac{\ln m_k(G)}{v(G)} \geq \mathbb{G}_d(p) + o_{v(G)}(1),$$

where $p = k/n$ and $m_k(G)$ denotes the number of matchings of size k.

(ii) *For any d-regular bipartite graph G, we have*

$$\lambda_G(p) \geq \mathbb{G}_d(p).$$

Proof. First we show that (i) implies (ii). Since both functions $\lambda_G(p)$ and $\mathbb{G}_d(p)$ are continuous, it is enough to prove the claim for rational numbers p. Let $p = a/b$. Consider br copies of G, and consider the matchings of size $k = ar$. Then

$$\lambda_G(p) = \lambda_{brG}(p) \geq \frac{\ln m_k(brG)}{v(brG)} - \frac{\ln v(brG)}{v(brG)} \geq \mathbb{G}_d(p) + o_{v(brG)}(1) - \frac{\ln v(brG)}{v(brG)}.$$

The (first) equality follows from Proposition 2.1(a), the first inequality follows from Proposition 2.1(d), and the second inequality is the assumption of (i). As r tends to infinity, the last two terms disappear, and we get

$$\lambda_G(p) \geq \mathbb{G}_d(p).$$

Next we show that (ii) implies (i). We have

$$\frac{\ln m_k(G)}{v(G)} \geq \lambda_G(p) - \frac{\ln v(G)}{v(G)} \geq \mathbb{G}_d(p) - \frac{\ln v(G)}{v(G)}.$$

The first inequality follows from Proposition 2.1(d), and the second is the assumption of (ii). Since $-(\ln v(G))/v(G) = o_{v(G)}(1)$, we are done. □

Theorem 1.3 implies

Corollary 3.2.

$$\frac{\ln m_k(G)}{v(G)} \geq \mathbb{G}_d(p) - \frac{\ln v(G)}{v(G)}.$$

Proof. See the second part of the proof of Theorem 3.1. □

Remark 3.3. L. Gurvits actually proved much stronger results than Theorem 1.3. He showed that for all pairs (P, Q) of $n \times n$ matrices, where P is non-negative and Q is doubly stochastic, we have

$$\ln(\mathrm{Per}(P)) \geq \sum_{1 \leq i,j \leq n} (1 - Q(i,j)) \ln(1 - Q(i,j)) - \sum_{1 \leq i,j \leq n} Q(i,j) \ln\left(\frac{Q(i,j)}{P(i,j)}\right).$$

From this he deduced the following inequality: for any doubly stochastic matrix A we have

$$\mathrm{Per}(A) \geq \prod_{1 \leq i,j \leq n} (1 - A(i,j))^{1 - A(i,j)}.$$

Next he showed that this inequality implies that for a d-regular bipartite graph G we have

$$m_k(G) \geq \frac{(1 - p/d)^{(1-p/d)nd} (1 - 1/n)^{(1-1/n)2n^2(1-p)}}{(p/d)^{np} n^{-2n(1-p)} ((n(1-p))!)^2},$$

where $p = k/n$ as before. For fixed $p \in (0, 1)$ this gives

$$m_k(G) \geq \left(1 + O\left(\frac{1}{n}\right)\right) \frac{e^{1-p}}{2\pi n(1-p)} \exp(2n\mathbb{G}_d(p)).$$

Let us mention that M. Lelarge [19] was able to give new proofs of Gurvits's results and extend both Gurvits's results and the results of this paper by combining the methods of this paper with some new ideas.

4. New proof of Gurvits's and Schrijver's theorems

In this section we give a new proof of Gurvits's and Schrijver's theorems. We will show that for any d-regular bipartite graph G,

$$\lambda_G(p) \geq \mathbb{G}_d(p).$$

According to Theorem 3.1, this is equivalent to Gurvits's theorem. For $p = 1$ we recover Schrijver's theorem via Proposition 2.1(e). Note that the function on the right hand side is nothing other than $\lambda_{\mathbb{T}_d}(p)$ according to Theorem 2.15.

Definition 4.1. Let G be a graph. Then H is a 2-*lift* of G if $V(H) = V(G) \times \{0, 1\}$, and for every $(u, v) \in E(G)$, exactly one of the following two pairs are edges of H: $((u, 0), (v, 0))$ and $((u, 1), (v, 1))$ are in $E(H)$ or $((u, 0), (v, 1))$ and $((u, 1), (v, 0))$ are in $E(H)$. If $(u, v) \notin E(G)$, then none of $((u, 0), (v, 0))$, $((u, 1), (v, 1))$, $((u, 0), (v, 1))$ or $((u, 1), (v, 0))$ are edges in H.

Note that if G is bipartite then any 2-lift of G is bipartite too.

Lemma 4.2. *Let G be a bipartite graph, and H be a 2-lift of G. Then for any k,*

$$m_k(G \cup G) \geq m_k(H).$$

In particular, for any $t \geq 0$,

$$M(G, t)^2 \geq M(H, t).$$

Proof. Since $M(G \cup G, t) = M(G, t)^2$, the inequality $m_k(H) \leq m_k(G \cup G)$ would indeed imply the second part of the lemma. Note that $G \cup G$ can be considered as a trivial 2-lift of G. Let M be a matching of a 2-lift of G. Consider the projection of M to G. Then it will consist of disjoint unions of cycles of even lengths (here we use the fact that G is bipartite!), paths and "double-edges" (when two edges project to the same edge). Let $\mathcal{R}$ be the set of these configurations. Then

$$m_k(H) = \sum_{R \in \mathcal{R}} |\phi_H^{-1}(R)| \quad \text{and} \quad m_k(G \cup G) = \sum_{R \in \mathcal{R}} |\phi_{G \cup G}^{-1}(R)|,$$

where ϕ_H and $\phi_{G \cup G}$ are the projections from H and $G \cup G$ to G. Note that

$$|\phi_{G \cup G}^{-1}(R)| = 2^{k(R)},$$

where $k(R)$ is the number of connected components of R different from a double-edge. On the other hand,

$$|\phi_H^{-1}(R)| \leq 2^{k(R)},$$

since in each component if we know the inverse image of one edge then we immediately know the inverse images of all other edges. The only reason why there is not equality in general is that not necessarily every cycle can be obtained as a projection of a matching of a 2-lift: for instance, if one considers an 8-cycle as a 2-lift of a 4-cycle, then no matching will project to the whole 4-cycle. Hence

$$|\phi_H^{-1}(R)| \le |\phi_{G\cup G}^{-1}(R)|,$$

and consequently

$$m_k(H) \le m_k(G \cup G).$$ □

By Proposition 2.1(g) we get the following corollary.

Corollary 4.3. *If G is a bipartite graph, and H is a 2-lift of G, then $\lambda_G(p) \ge \lambda_H(p)$ for every $0 \le p \le 1$.*

Lemma 4.4 (Linial [20]). *For any graph G, there exists a graph sequence $(G_i)_{i=0}^{\infty}$ such that $G_0 = G$, G_i is a 2-lift of G_{i-1} for $i \ge 1$, and $g(G_i) \to \infty$, where $g(H)$ is the girth of the graph H, i.e., the length of the shortest cycle.*

Proof. We will show that there exists a sequence (G_i) of 2-lifts such that for any k, there exists an $N(k)$ such that for $j > N(k)$, the graph G_j has no cycle of length at most k. Clearly, if H has no cycle of length at most $k-1$, then neither does any 2-lift of it. So it is enough to prove that if H has no cycle of length at most $k-1$, then there exists an H' obtained from H by a sequence of 2-lifts without a cycle of length at most k. We show that if $g(H) = k$, then there exists a lift of H with fewer k-cycles than in H. Let X be the random variable counting the number of k-cycles in a random 2-lift of H. Every k-cycle of H lifts to two k-cycles or a $2k$-cycle with probability $1/2$ each, so $\mathbb{E}X$ is exactly the number of k-cycles of H. But $H \cup H$ has twice as many k-cycles as H, so there must be a lift with strictly fewer k-cycles than H has. Choose this 2-lift and iterate this step to obtain an H' with girth at least $k+1$. □

Corollary 4.5. (a) *For any d-regular graph G, there exists a graph sequence $(G_i)_{i=0}^{\infty}$ such that $G_0 = G$, G_i is a 2-lift of G_{i-1} for $i \ge 1$, and (G_i) is Benjamini–Schramm convergent to the d-regular infinite tree $\mathbb{T}_d$.*

(b) *For any (a,b)-biregular bipartite graph G, there exists a graph sequence $(G_i)_{i=0}^{\infty}$ such that $G_0 = G$, G_i is a 2-lift of G_{i-1} for $i \ge 1$, and (G_i) is Benjamini–Schramm convergent to the (a,b)-biregular infinite tree $\mathbb{T}_{a,b}$.*

Proof of Theorems 1.1 and 1.3. Let $0 \le p < 1$. Choose a graph sequence $(G_i)_{i=0}^{\infty}$ such that $G_0 = G$, G_i is a 2-lift of G_{i-1} for $i \ge 1$, and (G_i) is Benjamini–Schramm convergent to the d-regular infinite tree $\mathbb{T}_d$. Then by Corollary 4.3,

$$\lambda_{G_0}(p) \ge \lambda_{G_1}(p) \ge \lambda_{G_2}(p) \ge \cdots \quad \text{and} \quad \lim_{i\to\infty} \lambda_{G_i}(p) = \lambda_{\mathbb{T}_d}(p)$$

since G_i converges to $\mathbb{T}_d$ (see Theorem 2.9). Hence $\lambda_G(p) \ge \lambda_{\mathbb{T}_d}(p)$ for $0 \le p < 1$. Finally, for $p = 1$ we have

$$\lambda_G(1) = \lim_{p\to 1} \lambda_G(p) \ge \lim_{p\to 1} \lambda_{\mathbb{T}_d}(p) = \lambda_{\mathbb{T}_d}(1).$$

Note that by Proposition 2.1(e), the inequality $\lambda_G(1) \geq \lambda_{\mathbb{T}_d}(1)$ is equivalent to

$$\frac{\ln \mathrm{pm}(G)}{v(G)} \geq \frac{1}{2} \ln\left(\frac{(d-1)^{d-1}}{d^{d-2}}\right),$$

which completes the proof of Theorem 1.1. □

One can prove the following theorem the very same way.

Theorem 4.6. *For any (a, b)-biregular bipartite graph G we have*

$$\lambda_G(p) \geq \mathbb{G}_{a,b}(p) \quad \text{for every } 0 \leq p \leq \min\left(\frac{a}{a+b}, \frac{b}{a+b}\right).$$

With the same technique one can prove the following theorem.

Theorem 4.7. *Let G be a d-regular bipartite graph, and $t \geq 0$. Then*

$$\int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_G(z) \geq \int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_{\mathbb{T}_d}(z).$$

Proof. Note that

$$\frac{\ln M(G,t)}{v(G)} = \int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_G(z).$$

Choose a graph sequence $(G_i)_{i=0}^{\infty}$ such that $G_0 = G$, G_i is a 2-lift of G_{i-1} for $i \geq 1$, and (G_i) is Benjamini–Schramm convergent to the d-regular infinite tree $\mathbb{T}_d$. By Lemma 4.2,

$$\frac{\ln M(G_0,t)}{v(G_0)} \geq \frac{\ln M(G_1,t)}{v(G_1)} \geq \frac{\ln M(G_2,t)}{v(G_2)} \geq \cdots,$$

and by the weak convergence of the measures ρ_{G_i} (see Theorem 2.12),

$$\lim_{i\to\infty} \frac{\ln M(G_i,t)}{v(G_i)} = \lim_{i\to\infty} \int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_{G_i}(z) = \int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_{\mathbb{T}_d}(z).$$

Hence

$$\int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_G(z) \geq \int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_{\mathbb{T}_d}(z).$$ □

Next we prove Theorem 1.6 which is a direct consequence of the previous theorem.

Proof of Theorem 1.6. We can assume that $0 \leq p < 1$; for $p = 1$ the claim follows by continuity. We have seen that, for $t \geq 0$,

$$\frac{\ln M(G,t)}{v(G)} = \int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_G \geq \int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_{\mathbb{T}_d}.$$

Note that by Theorem 2.15 we have

$$\int \tfrac{1}{2} \ln(1+tz^2)\, d\rho_{\mathbb{T}_d} = \tfrac{1}{2} \ln S_d(t),$$

where

$$S_d(t) = \frac{1}{\eta_t^2}\left(\frac{d-1}{d-\eta_t}\right)^{d-2}, \qquad \eta_t = \frac{\sqrt{1+4(d-1)t}-1}{2(d-1)t}.$$

Hence

$$M(G,t) \geq S_d(t)^n$$

for all $t \geq 0$. Now let

$$t = t(\mathbb{T}_d, p) = \frac{p(d-p)}{d^2(1-p)^2}.$$

Then

$$\eta_t = \frac{1-p}{1-p/d}, \qquad S_d(t) = \frac{(1-p/d)^d}{(1-p)^2}.$$

Hence

$$M\left(G, \frac{p(d-p)}{d^2(1-p)^2}\right) \geq \frac{1}{(1-p)^{2n}}\left(1-\frac{p}{d}\right)^n.$$

After multiplying by $(1-p)^{2n}$, we get the claim of the theorem. □

We end this section with another corollary of Theorem 4.7. The so-called *matching energy*, introduced by I. Gutman and S. Wagner [15], is defined as follows:

$$ME(G) = \sum_{z_i:\,\mu(G,z_i)=0} |z_i|,$$

where all zeros are counted with their multiplicity. With our notation this reads

$$ME(G) = v(G)\int |z|\, d\rho_G(z).$$

The following theorem shows that if we normalize the matching energy by dividing by the number of vertices, then among d-regular bipartite graphs its "minimum" is attained at the infinite d-regular tree $\mathbb{T}_d$.

Corollary 4.8. *Let G be a d-regular bipartite graph. Then*

$$\int |z|\, d\rho_G(z) \geq \int |z|\, d\rho_{\mathbb{T}_d}(z).$$

Proof. For any z we have

$$|z| = \frac{1}{\pi}\int_0^\infty \frac{1}{t^2}\ln(1+t^2z^2)\, dt.$$

Hence

$$\begin{aligned}\int |z|\, d\rho_G &= \int\left(\frac{1}{\pi}\int_0^\infty \frac{1}{t^2}\ln(1+t^2z^2)\, dt\right) d\rho_G(z)\\ &= \frac{1}{\pi}\int_0^\infty \frac{1}{t^2}\left(\int \ln(1+t^2z^2)\, d\rho_G(z)\right) dt\\ &\geq \frac{1}{\pi}\int_0^\infty \frac{1}{t^2}\left(\int \ln(1+t^2z^2)\, d\rho_{\mathbb{T}_d}(z)\right) dt = \int |z|\, d\rho_{\mathbb{T}_d}.\end{aligned}$$

Since we have integrated a non-negative function, the interchange of the integrals was allowed. □

Remark 4.9. Note that

$$\int |z|\, d\rho_{\mathbb{T}_d}(z) = \frac{d}{\pi}\left(2\sqrt{d-1} - (d-2)\arctan\left(\frac{2}{d-2}\sqrt{d-1}\right)\right).$$

5. Proof of the Lower Matching Conjecture

In this section we prove Theorem 1.5. Here the main tool is that the matching polynomial has only real zeros—this gives sufficient information about its coefficients so that together with our results on the entropy function we can finish the proof of Theorem 1.5. The argument in this section is more or less standard; a survey of related methods and results can be found in [23].

Proof of Theorem 1.5. We can assume that $0 \le p < 1$, since for $p = 1$, the statement reduces to Schrijver's theorem. Choose t such that $p(G,t) = p = k/n$. Then

$$m_k(G) = \frac{m_k(G)t^k}{M(G,t)} \exp(v(G)\lambda_G(p)).$$

Let

$$a_j = \frac{m_j(G)t^j}{M(G,t)}.$$

Then the probability distribution $(a_0, a_1, \dots, a_n)$ has mean $\mu = k$. By the Heilmann–Lieb theorem, $\sum a_j x^j$ has only real zeros. Then it is known that it is the distribution of the number of successes in independent trials. Indeed, let

$$M(G,t) = \prod_{i=1}^{n}(1 + \gamma_i t),$$

where $\gamma_i = \lambda_i^2$ with our previous notation, and

$$p_j = \frac{\gamma_j t}{1 + \gamma_j t}.$$

If I_j is the indicator variable that takes the value 1 with probability p_j, and 0 with probability $1 - p_j$, then

$$\mathbb{P}(I_1 + \cdots + I_n = j) = a_j.$$

The advantage of this observation is that there is a powerful inequality for such distributions, namely Hoeffding's inequality.

Theorem 5.1 (Hoeffding's inequality [17]). *Let S be a random variable with probability distribution of the number of successes in n independent trials. Assume that $\mathbb{E}S = np$. Let b and c integers satisfying $b \le np \le c$. Then*

$$\mathbb{P}(b \le X \le c) \ge \sum_{j=b}^{c} \binom{n}{j} p^j (1-p)^{n-j}.$$

In the particular case when $np = k$, we get

$$a_k \geq \binom{n}{k} p^k (1-p)^{n-k} = p_\mu$$

with our previous notation.

Putting everything together we obtain

$$m_k(G) = \frac{m_k(G)t^k}{M(G,t)} \exp(v(G)\lambda_G(p)) \geq p_\mu \exp(2n\mathbb{G}_d(p)).$$

In the last step we have used the fact that $\lambda_G(p) \geq \mathbb{G}_d(p)$ by Theorem 3.1. □

Proof of Theorem 1.7. The proof is completely analogous to the previous one. We have to use the inequality $\lambda_G(p) \geq \mathbb{G}_{a,b}(p)$ (see Theorem 4.6). □

Acknowledgments. The author is grateful to the following people for helpful conversations: Miklós Abért, Péter Frenkel, Leonid Gurvits, Tamás Hubai, Gábor Kun and Brendan McKay.

The author is partially supported by National Science Foundation under grant no. DMS-1500219, the Hungarian National Research, Development and Innovation Office, NKFIH grant K109684 and NN114614, a Slovenian–Hungarian grant, and by the MTA Rényi "Lendület" Groups and Graphs Research Group, and by the ERC Consolidator Grant 648017.

References

[1] Abért, M., Csikvári, P., Frenkel, P. E., Kun, G.: Matchings in Benjamini–Schramm convergent graph sequences. Trans. Amer. Math. Soc. **368**, 4197–4218 (2016) Zbl 1331.05176 MR 3453369

[2] Abért, M., Csikvári, P., Hubai, T.: Matching measure, Benjamini–Schramm convergence and the monomer-dimer free energy. J. Statist. Phys. **161**, 16–34 (2015) Zbl 1327.82034 MR 3392506

[3] Bayati, M., Gamarnik, D., Katz, D., Nair, C., Tetali, P.: Simple deterministic algorithm for counting matchings. In: ACM Symp. on Theory of Computing, STOC'07, 122–127 (2007) Zbl 1232.68179 MR 2402435

[4] Bollobás, B., McKay, B. D.: The number of matchings in random regular graphs and bipartite graphs. J. Combin. Theory Ser. B **41**, 80–91 (1986) Zbl 0602.05050 MR 0854605

[5] Csikvári, P., Frenkel, P. E.: Benjamini–Schramm continuity of root moments of graph polynomials. Eur. J. Combin. **52**, 302–320 (2016) Zbl 1327.05159 MR 3425982

[6] Darroch, J. N.: On the distribution of the number of successes in independent trials. Ann. Math. Statist. **35**, 1317–1321 (1964) Zbl 0213.44402 MR 0164359

[7] Dumitriu, I., Johnson, T.: The Marčenko–Pastur law for sparse random bipartite biregular graphs. Random Structures Algorithms **48**, 313–340 (2016) Zbl 1332.05126 MR 3449601

[8] Elek, G., Lippner, G.: Borel oracles. An analytical approach to constant-time algorithms. Proc. Amer. Math. Soc. **138**, 2939–2947 (2010) Zbl 1200.68283 MR 2644905

[9] Friedland, S., Krop, E., Markström, K.: On the number of matchings in regular graphs. Electron. J. Combin. **15**, res. paper 110, 28 pp. (2008) Zbl 1181.05011 MR 2438582

[10] Godsil, C. D.: Algebraic Combinatorics. Chapman and Hall, New York (1993) Zbl 0784.05001 MR 1220704

[11] Godsil, C. D., Gutman, I.: On the matching polynomial of a graph. In: Algebraic Methods in Graph Theory, L. Lovász and V. T. Sós (eds.), Colloq. Math. Soc. János Bolyai 25, North-Holland, 241–249 (1981) Zbl 0476.05060 MR 0642044

[12] Greenhill, C., Janson, S., Ruciński, A.: On the number of perfect matchings in random lifts. Combin. Probab. Comput. **19**, 791–817 (2010) Zbl 1260.05144 MR 2726080

[13] Gurvits, L.: Van der Waerden/Schrijver–Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all. Electron. J. Combin. **15**, res. paper 66, 26 pp. (2008) Zbl 1182.15008 MR 2411443

[14] Gurvits, L.: Unleashing the power of Schrijver's permanental inequality with the help of the Bethe approximation. arXiv:1106.2844v11 (2012)

[15] Gutman, I., Wagner, S.: The matching energy of a graph. Discrete Appl. Math. **160**, 2177–2187 (2012) Zbl 1252.05120 MR 2954760

[16] Heilmann, O. J., Lieb, E. H.: Theory of monomer-dimer systems. Comm. Math. Phys. **25**, 190–232 (1972) Zbl 0228.05131 MR 0297280

[17] Hoeffding, W.: On the distribution of the number of successes in independent trials. Ann. Math. Statist. **35**, 713–721 (1964) Zbl 0073.13902 MR 0080391

[18] Laurent, M., Schrijver, A.: On Leonid Gurvits's proof for permanents. Amer. Math. Monthly **117**, 903–911 (2010) Zbl 1210.15008 MR 2759363

[19] Lelarge, M.: Counting matchings in irregular bipartite graphs and random lifts. arXiv:1507.04739 (2015)

[20] Linial, N.: Lifts of graphs (talk slides). http://www.cs.huji.ac.il/ nati/PAPERS/lifts_talk.pdf

[21] Lovász, L.: Large Networks and Graph Limits. Amer. Math. Soc. Colloq. Publ. 60, Amer. Math. Soc., Providence, RI (2012) Zbl 1292.05001 MR 3012035

[22] Nguyen, H. N., Onak, K.: Constant-time approximation algorithms via local improvements. In: 49th Annual IEEE Symposium on Foundations of Computer Science, 327–336 (2008)

[23] Pitman, J.: Probabilistic bounds on the coefficients of polynomials with only real zeros. J. Combin. Theory Ser. A **77**, 279–303 (1997) Zbl 0866.60016 MR 1429082

[24] Robbins, H.: A remark on Stirling's formula. Amer. Math. Monthly **66**, 26–29 (1955) Zbl 0068.05404 MR 0069328

[25] Schrijver, A.: Counting 1-factors in regular bipartite graphs. J. Combin. Theory Ser. B **72**, 122–135 (1998) Zbl 0905.05060 MR 1604705

[26] Schrijver, A., Valiant, W. G.: On lower bounds for permanents. Indag. Math. **42**, 425–427 (1980) Zbl 0451.15009 MR 0598000

[27] Voorhoeve, M.: A lower bound for the permanents of certain (0, 1)-matrices. Indag. Math. **41**, 83–86 (1979) Zbl 0401.05005 MR 0528221

[28] Zykov, A. A.: On some properties of linear complexes. Mat. Sb. **24**, 163–188 (1949) (in Russian) Zbl 0033.02602 MR 0035428