NOTES ON MATCHINGS IN CONVERGENT GRAPH SEQUENCES

M. ABÉRT, P. CSIKVÁRI, P. FRENKEL, AND G. KUN

ABSTRACT. These are notes on the paper "Matching in Benjamini-Schramm convergent graph sequences" by the above authors. We define Benjamini-Schramm convergence for a sequence of finite graphs, and show that for graphs in such a convergent sequence the respective matching measures converge to a limiting measure. This implies that the total number of matchings, suitably normalized, converges for graphs in such a sequence.

1. MOTIVATION

Earlier this semester, we studied matchings in graphs in the following context: suppose G_1, G_2, \ldots is a sequence of finite, d-regular graphs, whose girth $\to \infty$ as $n \to \infty$. Then the normalized matching polynomials converge, in the sense that

$$\lim_{n \to \infty} \frac{1}{|V_n|} \ln \text{MAT}_{G_n}(x)$$

converges to some power series in $\mathbb{R}[[x]]$, and the limit depends *only* on the degree d. (In fact, we showed this limit is

$$\frac{d}{2}\ln\left(\frac{1+\sqrt{1+4(d-1)x}}{2}\right) - \frac{d-2}{2}\ln\left(\frac{d-2+d\sqrt{1+4(d-1)x}}{2d-2}\right)$$

but this is not important in what follows.)

The idea behind this phenomenon is that any sequence of graphs $\{G_n\}$ with the above assumptions on degree and girth will "converge" to the same object T_d , the infinite d-regular tree. We cannot attach a matching polynomial "MAT $_{T_d}(x)$ " to the infinite graph T_d by the usual definition, but we can instead imagine taking the limiting object above as a sort of definition.

Example 1. The sequence of 2-regular graphs converges to the (bi-)infinite path T_2

$$\triangle$$
, \square , \diamondsuit , ... \rightarrow

Sometimes, by understanding the limiting object of the sequence $\{G_n\}$ sufficiently well, we can make conclusions about the finite graphs G_n .

Problem. Can we generalize this to other sequences of graphs, i.e. graphs with small girth or graphs which are not d-regular? In particular,

- (Q1) What does it mean for a sequence of finite graphs to "converge"?
- (Q2) When does a convergent sequence $\{G_n\}$ tell us anything about matchings?

Example 2. Here are some situations where we may want such a generalization:

(1) 2-dimensional square lattices:
$$\square$$
, \square , \square , $\cdots \rightarrow ?$
(2) "triangle paths": \longrightarrow , \longrightarrow ?

Notes written by Harry Richman, for Math 669 presentation.

(3) rooted binary trees:
$$\bigwedge$$
, \bigwedge , \bigwedge , \bigwedge , $\dots \rightarrow$?

(4) complete graphs: \bigwedge , \bigvee , \bigvee , $\dots \rightarrow$?

The main result of these notes says that that we can answer "yes" to (Q2) in the first three of the above examples. (The fourth one fails to converge by our definition.)

2. Benjamini-Schramm convergence

Here we answer (Q1) by defining a formal notion of convergence for a sequence of finite graphs, due to Benjamini and Schramm [3]. The idea is that the graphs should converge if they are "locally probably isomorphic"—by "locally" we mean look at finite-distance neighborhoods, and by "probably" we mean average over neighborhoods at all vertices.

2.1. **Local neighborhood probabilities.** Here we define how to measure neighborhood probabilities in a graph.

Given a finite graph G=(V,E), a vertex $v\in V$, and an integer distance $r\geq 0$, let $\mathcal{N}_r(v,G)$ denote the r-neighborhood subgraph of v, i.e.

$$\mathcal{N}_r(v,G) = \text{induced subgraph of } G \text{ on vertices } \{w \in V : \delta(v,w) \leq r\}$$

where $\delta(v, w)$ is the number of edges in the shortest path in G from v to w. (If no such path exists, $\delta(v, w) = +\infty$.) Note that $\mathcal{N}_r(v, G)$ is naturally a rooted graph, with root v.

Given an auxiliary finite rooted graph α , let

$$\mathbb{P}(G, r; \alpha) = \text{Prob}(\text{uniformly random vertex in } G \text{ has } r\text{-neighborhood } \cong \alpha)$$
$$= \frac{|\{v \in V : \mathcal{N}_r(v, G) \cong \alpha\}|}{|V|}.$$

Here " \cong " means isomorphic as rooted graphs, i.e. a graph isomorphism sending root vertex to root vertex.

Example 3. The graph $G = \bigcap$ has neighborhood probabilities $\mathbb{P}(G, 1; \bigcap) = \frac{1}{2}$ and $\mathbb{P}(G, 1; \bigcap) = 0$ since this does not match any *induced* 1-neighborhood.

2.2. Graph convergence. Here we define Benjamini-Schramm convergence of graphs. We will just call this "convergence" since it is the only notion of graph convergence which we consider. There is one additional assumption we make on the sequence of graphs (G_n) .

Definition. A sequence of (finite) graphs $(G_1, G_2, ...)$ is *sparse* if the degree of vertices in G_n is bounded $\leq d$, independent of n.

Definition ([3]). A sparse sequence of finite graphs (G_n) is (Benjamini-Schramm) convergent if, for all choices of distance $r \geq 1$ and rooted graph α ,

$$\lim_{n\to\infty} \mathbb{P}(G_n, r; \alpha)$$

converges to a limiting probability, necessarily $\in [0, 1]$.

Example 4. Going back to the Example 2 in the first section, all four sequences have converging probabilities $\mathbb{P}(G_n, r; \alpha)$ as $n \to \infty$ but only the first three sequences are sparse. Thus (1), (2) and (3) are convergent graph sequences.

Remark. In (4), all $\mathbb{P}(G_n, r; \alpha) \to 0$ as $n \to \infty$, since any fixed α has bounded degree but the degrees of G_n grow with n. These neighborhood probabilities do not tell us much about the limiting graph; this is the reason we restrict to sparse sequences $\{G_n\}$.

Remark. Note that sequence (3) of Example 2 does *not* converge to the infinite 3-regular tree T_3 , since each graph in the sequence has many degree-1 vertices, i.e. $\mathbb{P}(G_n, 1; \stackrel{\bullet}{\longrightarrow}) > \frac{1}{2}$ for all n.

2.3. **Estimability.** Now that we have a well-defined notion of convergence of graphs, we can ask when convergence respects the usual things we look at when studying finite graphs.

Definition. Let \mathcal{G} denote the set of finite graphs up to isomorphism. A graph parameter p(G): $\mathcal{G} \to \mathbb{R}$ is *estimable* if, for any convergent sequence (G_n) of finite graphs,

$$\lim_{n\to\infty} p(G_n) \quad \text{converges.}$$

Example 5. The graph parameter p(G) = (average degree of vertices) is estimable, since this is determined by the 1-neighborhood probabilities of G.

Example 6. The graph parameter p(G) = (# connected components) is not estimable, since taking disjoint unions of a graph with itself does not change neighborhood probabilities, so e.g.

$$\triangle$$
, \triangle \triangle , \triangle , \triangle , \triangle , ...

is convergent.

The work of Abért, Csikvári, Frenkel, and Kun (and others) shows that in fact many natural graph parameters are estimable. In [1] the focus is on independent sets of G and matchings of G, which are related by the line graph construction. The proofs rely on more general statements in [4], which covers e.g. many specializations of the Tutte polynomial. This was largely motivated by [2] in which the chromatic polynomial is considered.

In what follows, we focus on counting matchings in G.

3. Matching measure

Given a finite graph G = (V, E) on n = |V| vertices, recall that a matching in G is a subset of edges $M \subset E$ such that no edges in M share a vertex. The matching polynomial of G is

$$\mathrm{MAT}_G(x) = \sum_{k \ge 0} m_k(G) x^k$$

and the $matching\ defect\ polynomial\ of\ G$ is

(2)
$$q_G(x) = \sum_{k \ge 0} (-1)^k m_k(G) x^{n-2k}$$
$$= x^n - m_1(G) x^{n-2} + m_2(G) x^{n-4} - \cdots$$

where $m_k(G) = \#(k\text{-edge matchings in } G)$. These are related by $q_G(x) = x^n \operatorname{MAT}_G(-\frac{1}{x^2})$.

Definition. The matching measure ρ_G of a finite graph G is the uniform measure on the roots of its matching defect polynomial $q_G(x)$, i.e.

(3)
$$\rho_G = \frac{1}{|V|} \sum_{\text{roots } \lambda} \delta_{\lambda}$$

where the sum is over roots of q_G counted with multiplicity.

Note that ρ_G is a discrete probability measure on \mathbb{C} . The definition of ρ_G says that

$$\int_{\mathbb{C}} f(z) d\rho_G(z) = \frac{1}{|V|} \sum_{\text{roots } \lambda} f(\lambda)$$

for any continuous function $f: \mathbb{C} \to \mathbb{R}$.

Example 7. If $G = ^{\triangle}$ then $q_G(x) = x^3 - 3x$ has roots $\{\lambda\} \approx \{0, \pm 1.732\}$. If $G = ^{\square}$ then $q_G(x) = x^4 - 4x^2 + 2$ has roots $\{\lambda\} \approx \{\pm 0.765, \pm 1.868\}$. If $G = ^{\square}$ then $q_G(x) = x^5 - 4x^3$ has roots $\{\lambda\} = \{0, 0.0, \pm 2\}$.

Remark. In general, 0 if a root of q_G iff MAT_G has degree <|V|/2 iff G has no perfect matching. In other words, $\rho_G(\{0\}) = \frac{(\# \text{ roots } = 0)}{|V|}$ measures "how far" G is from having a perfect matching.

3.1. **Main result.** We restate (with minor modifications in notation) the main theorem of [1] in the section on matchings. The gist is that the answer to (Q2) from the introduction is *always* yes: convergence of graphs implies convergence of "matching data".

Theorem ([1], Theorem 3.5). (a) Suppose $\{G_n\}$ is a convergent sequence of finite graphs, and let ρ_{G_n} be the matching measure of G_n . Then the measures ρ_{G_n} converge weakly to a measure ρ , meaning that for any continuous function $f: \mathbb{C} \to \mathbb{R}$,

$$\int f(z)d\rho_{G_n}(z) \to \int f(z)d\rho(z).$$

(b) The "matching entropy per vertex"

$$\frac{1}{|V|} \ln \text{MAT}_G(1) = \frac{1}{|V|} \ln(\text{total } \# \text{ of matchings})$$

is an estimable graph parameter.

(c) The "matching ratio"

$$\frac{1}{|V|} \deg MAT_G(x) = \frac{1}{|V|} (\text{max size of matching})$$

is an estimable graph parameter.

Remark. By results of Heilmann-Lieb and Godsil on the matching polynomial, the measures ρ_{G_n} , ρ are in fact supported on the compact subset $K = [-2\sqrt{d-1}, 2\sqrt{d-1}]$ of the real line, where d is the degree bound. So in particular we do not have to worry about the individual integrals in (\star) converging for any continuous f.

Part (a) is the hard part, and we save the proof for the next section. To get parts (b) and (c), we just need to pick the right function f(z) to integrate against in (\star) .

Proof of $(a) \Rightarrow (b)$. Observe that the total number of matchings is

$$MAT_G(1) = \sum_{k>0} m_k(G) = q_G(i),$$

where $i = \sqrt{-1}$. Since q_G is monic, $q_G(x) = \prod_{\text{roots } \lambda} (x - \lambda)$. Thus

$$\begin{split} \frac{1}{|V|} \ln \mathrm{MAT}_G(1) &= \frac{1}{|V|} \ln |q_G(i)| \\ &= \frac{1}{|V|} \sum_{\mathrm{roots} \ \lambda} \ln |i - \lambda| = \int \ln |i - z| d\rho_G(z). \end{split}$$

The function $f(z) = \ln |i-z|$ is continuous on the support $K = [-2\sqrt{d-1}, 2\sqrt{d-1}]$ of the measures ρ_G , so we can apply (a).

Alternatively, by the symmetry of the roots of q_G occurring in $\pm \lambda$ pairs (or since $\lambda \in \mathbb{R}$), we have

$$\frac{1}{|V|} \ln \text{MAT}_G(1) = \frac{1}{2|V|} \sum_{\text{roots } \lambda} \ln |(i - \lambda)(i + \lambda)| = \frac{1}{2} \int \ln |1 + z^2| d\rho_G(z).$$

So
$$f(z) = \frac{1}{2} \ln |1 + z^2|$$
 also works.

Proof sketch of $(a) \Rightarrow (c)$. Here the appropriate function to integrate against is $f(z) = \frac{1}{2} - \frac{1}{2}\delta_0$, as suggested by the remark following Example 7. This is not continuous unfortunately, so an extra technical step needed is to show that each q_G does not have too many of its roots within a small neighborhood of 0. For details see Lemma 2.4 and Remark 3.2 of [1].

4. Proof: Convergence of Measure

The main steps of the proof of Theorem (a) are:

- (1) Weierstrass approximation of continuous functions $f : \mathbb{R} \to \mathbb{R}$ by polynomials implies that it suffices to show (\star) when $f(z) = z^k$, for all $k \ge 0$.
- (2) Newton's identities give explicit algebraic relations between power sums p_k and elementary symmetric sums e_k :

$$\{p_k(\lambda) = \lambda_1^k + \dots + \lambda_n^k : \text{ all } k \ge 1\} \leftrightarrow \{e_k(\lambda) = \sum_{i_1 < \dots < i_k} \lambda_{i_1} \cdots \lambda_{i_k} : \text{ all } k \ge 1\}$$

(3) Show coefficients $m_k(G)$ of the matching defect polynomial are expressible in terms of "subgraph counting" functions, to imply $\frac{1}{|V|}p_k(\lambda)$ is estimable

The connection $(1)\leftrightarrow(2)$ is that $\int z^k d\rho_G = \frac{1}{|V|} p_k(\lambda)$, and the connection $(2)\leftrightarrow(3)$ is that $e_k(\lambda) = (n-k)$ -th coefficient of q_G (up to sign), and this coefficient is 0 if k is odd and $m_{k/2}(G)$ if k is even:

$$q_G(x) = \prod_{\text{roots } \lambda} (x - \lambda) = x^n - e_1(\lambda)x^{n-1} + e_2(\lambda)x^{n-2} - e_3(\lambda)x^{m-3} + \cdots$$

$$= x^n - m_1(G)x^{n-2} + m_2(G)x^{n-4} - \cdots$$
(4)

Remark. To adapt this proof to other graph polynomials, steps (2) and (3) usually follow similarly but (1) may not work. If the graph polynomial has non-real roots, then convergence of "holomorhic moments" $\int z^k \rho_G$ does not imply convergence for all continuous (e.g. non-holomorphic) $f: \mathbb{C} \to \mathbb{R}$. See Theorem 1.10 of [4], or the Introduction of [2], for the appropriate generalized statement.

We omit the details in (1) and (2) and focus on step (3) which we restate as a lemma.

Lemma 1. For any $k \geq 0$, the normalized k-power sum of roots of the matching defect polynomial

$$\frac{1}{|V|}p_k(\lambda) = \frac{\lambda_1^k + \dots + \lambda_n^k}{n}, \quad \{\lambda_i\} = \text{ roots of } q_G(x)$$

is an estimable graph parameter.

(*CORRECTION: In my talk I stated the above lemma for the matching parameters $\frac{1}{|V|}m_k(G)$, but these are not estimable for $k \neq 1$.)

Example 8. For k = 0, the parameter $\frac{1}{|V|}p_0(\lambda) = 1$ is estimable.

For k = 1, since roots of $q_G(x)$ come in $\pm \lambda$ pairs $\frac{1}{|V|}p_1(\lambda) = 0$ is estimable.

For
$$k = 2$$
, $\frac{1}{|V|}p_2(\lambda) = \frac{1}{|V|}(e_1(\lambda)^2 - 2e_2(\lambda)) = \frac{2}{|V|}m_1(G) = 2\frac{|E|}{|V|}$ is estimable since $2\frac{|E|}{|V|} =$ (average degree).

The above discussion should indicate how Lemma 1 implies Theorem (a). The proof of this lemma will occupy the rest of these notes.

4.1. **Subgraph counting.** To prove Lemma 1 we will show that $m_k(G)$ is expressible in terms of functions which count connected subgraphs. These are evidently estimable by definition of graph convergence. We follow Csikvári and Frenkel [4], Sections 3 and 4. (A more direct argument is possible for $m_k(G)$, but the method here applies generally to many other graph polynomials.)

Definition. Given finite graphs H, G, the subgraph counting function

$$sub(H, G) = (\# subgraphs of G isomorphic to H).$$

Example 9. When $H = (\text{single vertex}), \text{sub}(\bullet, G) = |V(G)|.$

When H = (3 disjoint edges), sub $(\setminus \setminus \setminus, G) = m_3(G)$, the number of 3-matchings.

We think of sub(H, -) as an element of the infinite-dimensional \mathbb{C} -vector space

$$\mathbb{C}^{\mathcal{G}} = \{ \text{functions } f : \mathcal{G} \to \mathbb{C} \}.$$

Definition. Recall that $\mathcal{G} = \{\text{isomorphism classes of finite graphs}\}$. We also denote $\mathcal{C} = \{\text{isomorphism classes of } connected finite graphs}\}$. For any class $\mathcal{H} \subset \mathcal{G}$ of graphs up to isomorphism, let

$$\mathbb{C}\mathcal{H} = \{ \sum_{H \in \mathcal{H}} c_H \operatorname{sub}(H, -) : c_H \in \mathbb{C}, \text{ finitely many } c_H \neq 0 \}$$

denote the subspace of $\mathbb{C}^{\mathcal{G}}$ spanned by the *H*-subgraph counting functions from $H \in \mathcal{H}$.

Fact 1. The subgraph counting functions $\{\operatorname{sub}(H,-): H \in \mathcal{G}\}$ are linearly independent in $\mathbb{C}^{\mathcal{G}}$.

This follows from the "upper trinagular" property of subgraph counting:

$$sub(H,G) = \begin{cases} 0 & \text{if } |G| < |H| \\ 1 & \text{if } G = H \\ \ge 0 & \text{otherwise,} \end{cases}$$

where |G| < |H| means some appropriate combination of $|V(G)| \le |V(H)|$ and $|E(G)| \le |E(H)|$ and $G \ne H$.

Fact 2. The subspace $\mathbb{C}\mathcal{G}$ forms a subring of $\mathbb{C}^{\mathcal{G}}$ by pointwise multiplication:

(5)
$$\operatorname{sub}(H_1, G) \cdot \operatorname{sub}(H_2, G) = \sum_{H} c_{H_1, H_2}^{H} \operatorname{sub}(H, G)$$

for some integers c_{H_1,H_2}^H , non-zero for finitely many H.

Example 10. sub(\backslash, G)² = sub(\backslash, G) + 2 sub(\backslash, G) + 2 sub(\backslash, G).

Remark. Fact 1 and the above example imply $\mathbb{C}\mathcal{C}$ is NOT a subring of $\mathbb{C}^{\mathcal{G}}$.

Fact 3. The sequence (G_n) converges if and only if for all connected H,

$$\lim_{n\to\infty} \frac{1}{|V_n|} \operatorname{sub}(H, G_n)$$

converges.

This should be plausible since convergence is defined in terms of counting (rooted) neighborhoods of G_n , and neighborhoods are always connected. If we know all neighborhoods of radius r, then we know all connected subgraphs up to diameter r/2. (or up to diameter r?)

Remark. If H is disconnected then $\frac{1}{|V|}$ sub(H,G) may not be estimable. For example if H= (two vertices) then

$$\frac{1}{|V|}\operatorname{sub}(\bullet \bullet, G) = \frac{1}{|V|} \binom{|V|}{2} = \frac{|V| - 1}{2}$$

is not estimable for the same reason that (# connected components) is not estimable.

Returning to Lemma 1, we have the following problem: $m_k(G) \leftrightarrow p_k(\lambda)$ is obviously a subgraph-counting parameter since $m_k(G) = \text{sub}([k \text{ disjoint edges}], G)$, but H = (k disjoint edges) is NOT a connected subgraph c.f. Fact 3. The final step left here is to get around this problem.

4.2. Additivity and multiplicativity. Here we describe the argument to prove Lemma 1 as given in [4], Section 4. The goal is to show the power sums $p_k(\lambda)$ of the roots of $q_G(x)$ depend only on counting connected subgraphs of G, and not on counting disconnected subgraphs.

Definition. A graph parameter p(G) is additive if on disjoint unions of graphs

$$p(G_1 \sqcup G_2) = p(G_1) + p(G_2).$$

Definition. A graph polynomial $f_G(x)$ is multiplicative if on disjoint unions

$$f_{G_1 \sqcup G_2}(x) = f_{G_1}(x) f_{G_2}(x).$$

Lemma 2 ([4], Lemma 4.2). If a graph parameter $p \in \mathbb{CG}$, then p is additive if and only if $p \in \mathbb{CC}$.

Proof. The direction (\Leftarrow) is clear: if H is connected, then

$$sub(H, G_1 \sqcup G_2) = sub(H, G_1) + sub(H, G_2).$$

In the other direction (\Rightarrow) , suppose p is additive. Since $p \in \mathbb{C}\mathcal{G}$ write

$$p(G) = \sum_{H \in \mathcal{C} \text{ connected}} c_H \operatorname{sub}(H, G) + \sum_{H \notin \mathcal{C}} c_H \operatorname{sub}(H, G).$$

Additivity is linear, so without loss of generality suppose $c_H = 0$ for all connected H. Now let H_{\min} be the minimal graph with coefficient $c_H \neq 0$. By assumption this is not connected so $H_{\min} = H_1 \sqcup H_2$ for some non-empty subgraphs. Then by additivity

$$\begin{split} p(H_{\min}) &= p(H_1) + p(H_2) \\ &= \sum_{H \notin \mathcal{C}} c_H \operatorname{sub}(H, H_1) + \sum_{H \notin \mathcal{C}} c_H \operatorname{sub}(H, H_2) = 0 \end{split}$$

by the minimality of $H_{\min} > H_1, H_2$. But $p(H_{\min}) = \sum_{H \notin \mathcal{C}} c_H \operatorname{sub}(H, H_{\min}) = c_{H_{\min}}$, so $c_{H_{\min}} = 0$. This contradicts our choice of H_{\min} and implies $p \in \mathbb{C}\mathcal{C}$.

Lemma 3 ([4], Lemma 4.3). If $f_G(x)$ is a multiplicative graph polynomial such that the power sum of its roots $p_k(\lambda) = \sum_{\text{roots } \lambda} \lambda^k$ is $\in \mathbb{C}\mathcal{G}$ for some k, then $p_k(\lambda) \in \mathbb{C}\mathcal{C}$.

Proof. Since $f_G(x)$ is multiplicative, the power sum $p_k(\lambda)$ is an additive graph parameter. Then apply Lemma 2.

Proof of Lemma 1. We have shown the following:

$$\{ m_k(G) \in \mathbb{C}\mathcal{G} \text{ for all } k \} \Rightarrow \{ e_k(\lambda) \in \mathbb{C}\mathcal{G} \text{ for all } k \}$$

 $\Rightarrow \{ p_k(\lambda) \in \mathbb{C}\mathcal{G} \text{ for all } k \}$
 $\Rightarrow \{ p_k(\lambda) \in \mathbb{C}\mathcal{C} \text{ for all } k \}$

where $\lambda = \{\lambda_1, \dots, \lambda_n\}$ always refers to the roots of $q_G(x)$. The first implication is by equation (4), the second by Newton's identities and Fact 2, and the third by the fact that $q_G(x)$ is multiplicative and Lemma 3. By Fact 3, this implies $\frac{1}{|V|}p_k(\lambda)$ is estimable.

References

- [1] M. Abért, P. Csikvári, P. Frenkel, and G. Kun, *Matchings in Benjamini-Schramm convergent graph sequences*, Trans. Amer. Math. Soc., **368**, (2016), no. 6, 4197–4218, ISSN 0002-9947
- [2] M. Abért and T. Hubai, Benjamini-Schramm convergence and the distribution of chromatic roots for sparse graphs, Combinatorica, 35, (2015), no. 2, 127–151, ISSN 0209-9683
- [3] I. Benjamini and O. Schramm, Recurrence of Distributional Limits of Finite Planar Graphs, Electron. J. Probab., 6, (2001), no. 23, 13 pp. (electronic), DOI 10.1214/EJP.v6-96.
- [4] P. Csikvári and P. Frenkel, Benjamini-Schramm continuity of root moments of graph polynomials, European J. Combin., 52, (2016), part B, 302–320, ISSN 0195-6698