# Density, convergence and limits

Zdeněk Dvořák

November 27, 2020

It is natural to view the stability results we have seen so far as a kind of limit statements. For example, the stability version of the Erdős-Stone theorem can be re-stated as follows. Let $F$ be a graph of chromatic number $r+1$ and consider any sequence $G_1$, $G_2$, …of graphs such that for each $i$, $|G_i| = i$ and $F \not\subseteq G_i$. If $\|G_n\|/\binom{n}{2} \to 1 - 1/r$ as $n \to \infty$, then the sequence “converges to the balanced complete $r$-partite graph”. We now aim to develop a theory that will enable us to make such statements precise.

For graphs $H$ and $G$, let

$$p(H;G) = \frac{|\{S \subseteq V(G) : G[S] \simeq H\}|}{\binom{|G|}{|H|}}.$$

In other words, $p(H;G)$ is the probability that a subset of $|H|$ vertices of $G$ chosen uniformly at random induces a subgraph isomorphic to $H$. For example, $p(K_2;G) = \|G\|/\binom{|G|}{2}$ is the density of $G$.

Note that if we know $p(H;G)$ for all graphs $H$ with $m$ vertices, we can also determine it for all graphs with less than $m$ vertices, as follows. Let $\mathcal{H}_m$ denote the set of all pairwise non-isomorpic graphs with $m$ vertices.

**Lemma 1.** *For any graph $F$ with at most $m$ vertices and any graph $G$, we have*

$$p(F;G) = \sum_{H \in \mathcal{H}_m} p(F;H) \cdot p(H;G).$$

*Proof.* To choose a set $S$ of $|F|$ vertices of $G$, we can first choose a set $S_1$ of $m$ vertices, then choose $S$ as a subset of $S_1$. Hence, we have

$$\begin{aligned} p(F;G) = \Pr[G[S] \simeq F] &= \sum_{H \in \mathcal{H}_m} \Pr[G[S] \simeq F \mid G[S_1] \simeq H] \cdot \Pr[G[S_1] \simeq H] \\ &= \sum_{H \in \mathcal{H}_m} p(F;H) \cdot p(H;G). \end{aligned}$$

□

This fact (together with the obvious equality $\sum_{H\in\mathcal{H}_m} p(H;G) = 1$) can be used to obtain some bounds on the extremal functions (but usually not tight ones).

**Example 2.** *What can we say about the density of graphs without triangles? Let $G$ be an $n$-vertex triangle-free graph (so $p(K_3;G) = 0$). Let $N_3$ denote the graph consisting of three isolated vertices and $S_3$ the 3-vertex graph with one edge. We have*

$$\begin{aligned}
\frac{\|G\|}{\binom{n}{2}} &= p(K_2;G) = \sum_{H\in\mathcal{H}_3} p(K_2;H)p(H;G) \\
&= p(K_2;N_3)p(N_3;G) + p(K_2;S_3)p(S_3;G) + p(K_2;K_{1,2})p(K_{1,2};G) + p(K_2;K_3)p(K_3;G) \\
&= 0\cdot p(N_3;G) + \tfrac{1}{3}p(S_3;G) + \tfrac{2}{3}\cdot p(K_{1,2};G) + 1\cdot 0 \\
&\le \tfrac{2}{3}(p(N_3;G) + p(S_3;G) + p(K_{1,2};G)) = \tfrac{2}{3}.
\end{aligned}$$

*Recall that Mantel's theorem gives an asymptotically much better bound $\|G\| \le n^2/4 \approx \frac{1}{2}\binom{n}{2}$.*

To get an improvement, we need a more general notion. A flag $\mathbf{H}$ with $k$ roots is a pair $(H, \lambda_{\mathbf{H}})$, where $H$ is a graph and $\lambda_{\mathbf{H}} : \{1,\ldots,k\} \to V(H)$ is an injective function; i.e., a flag is a graph with some of its vertices assigned labels 1, …, $k$, where each label appears on exactly one vertex. We say two flags $\mathbf{H_1}$ and $\mathbf{H_2}$ are isomorphic and write $\mathbf{H_1} \simeq \mathbf{H_2}$ if they have the same number $k$ of roots and there exists an isomorphism $f$ of $H_1$ and $H_2$ such that for $i = 1,\ldots,k$, $f(\lambda_{\mathbf{H_1}}(i)) = \lambda_{\mathbf{H_2}}(i)$, i.e., the isomorphism respects the labels. The type of the flag $\mathbf{H}$ is the graph with vertex set $\{1,\ldots,k\}$, where $ij$ is an edge iff $\lambda_{\mathbf{H}}(i)\lambda_{\mathbf{H}}(j) \in E(H)$; i.e., the subgraph of $H$ induced by the labelled vertices. Clearly, two isomorphic flags have the same type.

For a flag $\mathbf{H}$ with $k$ roots, a graph $G$, and an injective function $\theta : \{1,\ldots,k\} \to V(G)$, let

$$p(\mathbf{H};G,\theta) = \frac{|\{S \subseteq V(G)\setminus \mathrm{im}(\theta) : (G[S\cup \mathrm{im}(\theta)],\theta) \simeq \mathbf{H}\}|}{\binom{|G|-k}{|H|-k}};$$

i.e., the probability that a random flag in $G$ with $|H|$ vertices and with labels on vertices $\theta(1)$, …, $\theta(k)$ in order is isomorphic to $\mathbf{H}$. For example, letting $\mathbf{K^1_m}$ be the flag with one root and the graph $K_m$, we have $p(\mathbf{K^1_2};G,\theta) = \deg(\theta(1))/(|G|-1)$.

Let us note that $p(\mathbf{H};G,\theta)$ is related to $p(H,G)$ by averaging. More precisely, for an expression $X(G,\theta)$ depending on a graph $G$ and an injective function $\theta : \{1,\ldots,k\} \to V(G)$, we define

$$E_\theta[X(\theta)] = \frac{\sum\{X(\theta) : \theta : \{1,\ldots,k\} \to V(G) \text{ injective}\}}{|G|(|G|-1)\cdots(|G|-k+1)}.$$

For example,

$$E_\theta[p(\mathbf{K_2^1}; G, \theta)] = \frac{\sum_{v \in V(G)} \deg(v)/(|G|-1)}{|G|},$$

and thus $E_\theta[p(\mathbf{K_2^1}; G, \theta)] \cdot (|G|-1)$ is the average degree of $G$.

**Lemma 3.** *For a flag* $\mathbf{H}$*, we have* $E_\theta[p(\mathbf{H}; G, \theta)] = E_\theta[p(\mathbf{H}; H, \theta)] p(H; G)$.

*Proof.* Let $k$ be the number of roots of $\mathbf{H}$. Observe that $E_\theta[p(\mathbf{H}; G, \theta)]$ is the probability that, after choosing uniformly at random an injective function $\theta : \{1, \ldots, k\} \to V(G)$ and a set $S \subseteq V(G) \setminus \mathrm{im}(\theta)$ of size $|H| - k$, the flag arising from the subgraph of $G$ induced by $\theta$ and $S$ is isomorphic to $\mathbf{H}$. The right hand side computes the same probability in a different way, first selecting a set $T$ of $|H|$ vertices, then an injective function from $\{1, \ldots, k\}$ to $T$. □

Next, we consider a combination of flags. Suppose $\mathbf{H_1}$ and $\mathbf{H_2}$ are flags of the same type, with $k$ roots. For a graph $G$, and an injective function $\theta : \{1, \ldots, k\} \to V(G)$, let us define

$$p(\mathbf{H_1}, \mathbf{H_2}; G, \theta) = \frac{|\{S_1, S_2 \subseteq V(G) \setminus \mathrm{im}(\theta) : S_1 \cap S_2 = \emptyset, (G[S_i \cup \mathrm{im}(\theta)], \theta) \simeq \mathbf{H_i} \text{ for } i \in \{1,2\}\}|}{\binom{|G|-k}{|H_1|-k, |H_2|-k, |G|-|H_1|-|H_2|+k}}.$$

For example, let $\mathbf{N_m^1}$ be the flag with one root and the graph consisting of $m$ isolated vertices. Then

$$p(\mathbf{K_2^1}, \mathbf{N_2^1}; G, \theta) = \frac{\deg(\theta(1)) \cdot (|G| - 1 - \deg(\theta(1)))}{(|G|-1)(|G|-2)}$$
$$p(\mathbf{K_2^1}, \mathbf{K_2^1}; G, \theta) = \frac{\deg(\theta(1)) \cdot (\deg(\theta(1)) - 1)}{(|G|-1)(|G|-2)}.$$

We can express this combined probability in terms of larger flags similarly to Lemma 1. For a type $\sigma$ and integer $m$, let $\mathcal{H}_{\sigma,m}$ denote the set of all flags of type $\sigma$ with $m$ vertices. For a flag $\mathbf{H}$, by $p(\mathbf{F_1}, \mathbf{F_2}; \mathbf{H})$ we mean $p(\mathbf{F_1}, \mathbf{F_2}; H, \lambda_{\mathbf{H}})$.

**Lemma 4.** *Suppose* $\mathbf{F_1}$ *and* $\mathbf{F_2}$ *are flags of the same type* $\sigma$*, with* $k$ *roots, and let* $m \geq |F_1| + |F_2| - k$ *be an integer. Then for any* $G$ *and* $\theta$*, we have*

$$p(\mathbf{F_1}, \mathbf{F_2}; G, \theta) = \sum_{\mathbf{H} \in \mathcal{H}_{\sigma,m}} p(\mathbf{F_1}, \mathbf{F_2}; \mathbf{H}) \cdot p(\mathbf{H}; G, \theta).$$

*Proof.* On the left-hand side, we calculate the probability that if we choose disjoint sets $S_1, S_2 \subseteq V(G) \setminus \mathrm{im}(\theta)$ of sizes $|F_1| - k$ and $|F_2| - k$, respectively, uniformly at random, then the flag induced by $\theta$ and $S_i$ in $G$ is isomorphic to $\mathbf{F_i}$ for $i \in \{1, 2\}$. On the right-hand side, we compute the same probability by first selecting a set $S \subseteq V(G) \setminus \mathrm{im}(\theta)$ of size $m - k$ uniformly at random, then choosing disjoint $S_1, S_2 \subseteq S$ uniformly at random. □

Let us now relate $p(\mathbf{F_1}, \mathbf{F_2}; G, \theta)$ to $p(\mathbf{F_1}; G, \theta) \cdot p(\mathbf{F_2}; G, \theta)$. The latter calculates the probability that, if we choose sets $S_1, S_2 \subseteq V(G) \setminus \mathrm{im}(\theta)$ of the appropriate size independently uniformly at random, then the flag induced by $\theta$ and $S_i$ in $G$ is isomorphic to $\mathbf{F_i}$ for $i \in \{1, 2\}$. Note that if $|G|$ is large, then the independently chosen sets $S_1$ and $S_2$ will almost surely be disjoint, and thus this probability will be close to $p(\mathbf{F_1}, \mathbf{F_2}; G, \theta)$. The following lemma gives this more precisely.

**Lemma 5.** *Suppose $\mathbf{F_1}$ and $\mathbf{F_2}$ are flags of the same type, with $k$ roots. Let $G$ be a graph with $n \geq |F_1| + |F_2| - k$ vertices and let $\theta : \{1, \ldots, k\} \to V(G)$ be an injective function. Then*

$$|p(\mathbf{F_1}, \mathbf{F_2}; G, \theta) - p(\mathbf{F_1}; G, \theta) \cdot p(\mathbf{F_2}; G, \theta)| \leq \frac{|F_1||F_2|}{n}.$$

*Proof.* Let $a = p(\mathbf{F_1}, \mathbf{F_2}; G, \theta)$ and $b = p(\mathbf{F_1}; G, \theta) \cdot p(\mathbf{F_2}; G, \theta)$. Let $m = \binom{n-k}{|F_1|-k, |F_2|-k, n-|F_1|-|F_2|+k}$ and $q = \binom{n-k}{|F_1|-k}\binom{n-k}{|F_2|-k}$. By the definition, $am$ is the number of pairs of disjoint sets $S_1$ and $S_2$ extending $\theta$ in $G$ to flags isomorphic to $\mathbf{F_1}$ and $\mathbf{F_2}$, while $bq$ is the same quantity without the constraint that $S_1$ and $S_2$ are disjoint. Moreover, $q - m$ is the number of ways how to choose a pair of non-disjoint subsets of $V(G) \setminus \mathrm{im}(\theta)$ of the appropriate size, and $0 \leq a, b \leq 1$. Hence,

$$\begin{aligned} am \leq bq &\leq am + q - m \\ -a(q-m)/q \leq b - a &\leq (1-a)(q-m)/q \\ -(q-m)/q \leq b - a &\leq (q-m)/q, \end{aligned}$$

and thus $|a - b| \leq (q - m)/q$. Recall $q - m$ is the number of ways how to choose a pair of non-disjoint subsets of $V(G) \setminus \mathrm{im}(\theta)$ of of sizes $|F_1| - k$ and $|F_2| - k$, and thus it is upper-bounded by $(n-k)\binom{n-k-1}{|F_1|-k-1}\binom{n-k-1}{|F_2|-k-1}$. Hence,

$$\frac{q-m}{q} \leq \frac{(n-k)\binom{n-k-1}{|F_1|-k-1}\binom{n-k-1}{|F_2|-k-1}}{\binom{n-k}{|F_1|-k}\binom{n-k}{|F_2|-k}} = \frac{(|F_1|-k)(|F_2|-k)}{n-k} \leq \frac{|F_1||F_2|}{n},$$

as required. □

We now have the tools for the applications of this framework. As a very simple example, let us prove an asymptotic version of Mantel's theorem.

**Example 6.** *Let $G$ be a triangle-free graph with $n$ vertices. In the calculation below, we use the following abbreviations (for flags $\mathbf{F_1}, \mathbf{F_2}$ and a graph $F$):*

$$\begin{aligned} \mathbf{F_1} &\equiv p(\mathbf{F_1}; G, \theta) \\ \mathbf{F_1} \circ \mathbf{F_2} &\equiv p(\mathbf{F_1}, \mathbf{F_2}; G, \theta) \\ F &\equiv p(F; G) \end{aligned}$$

*Since $G$ is triangle-free, we have (in this notation) $K_3 = 0$ and for every $\theta$, $\mathbf{K_3^1} = 0$. Let $\mathbf{K_{1,2}^1}$ be the flag with graph $K_{1,2}$ and the label 1 on one of the leaves, and $\mathbf{K_{1,2}^m}$ the flag with the same graph and the label 1 on the vertex of degree two. Let $\mathbf{S_3^1}$ be the flag with graph $S_3$ and the label 1 on one of the leaves, and $\mathbf{S_3^m}$ the flag with the same graph and the label 1 on the isolated vertex. By Lemma 4,*

$$\begin{aligned} \mathbf{K_2^1} \circ \mathbf{K_2^1} = & \; p(\mathbf{K_2^1}, \mathbf{K_2^1}; \mathbf{N_3^1}) \cdot \mathbf{N_3^1} \\ & + p(\mathbf{K_2^1}, \mathbf{K_2^1}; \mathbf{S_3^m}) \cdot \mathbf{S_3^m} \\ & + p(\mathbf{K_2^1}, \mathbf{K_2^1}; \mathbf{S_3^1}) \cdot \mathbf{S_3^1} \\ & + p(\mathbf{K_2^1}, \mathbf{K_2^1}; \mathbf{K_{1,2}^1}) \cdot \mathbf{K_{1,2}^1} \\ & + p(\mathbf{K_2^1}, \mathbf{K_2^1}; \mathbf{K_{1,2}^m}) \cdot \mathbf{K_{1,2}^m} \\ & + p(\mathbf{K_2^1}, \mathbf{K_2^1}; \mathbf{K_3^1}) \cdot \mathbf{K_3^1} \\ = & \; 0 \cdot \mathbf{N_3^1} + 0 \cdot \mathbf{S_3^m} + 0 \cdot \mathbf{S_3^1} + 0 \cdot \mathbf{K_{1,2}^1} + 1 \cdot \mathbf{K_{1,2}^m} + 1 \cdot 0 \\ = & \; \mathbf{K_{1,2}^m}. \end{aligned}$$

*Similarly,*

$$\begin{aligned} \mathbf{K_2^1} \circ \mathbf{N_2^1} &= \tfrac{1}{2}\mathbf{S_3^1} + \tfrac{1}{2}\mathbf{K_{1,2}^1} \\ \mathbf{N_2^1} \circ \mathbf{N_2^1} &= \mathbf{N_3^1} + \mathbf{S_3^m} \end{aligned}$$

*Furthermore, using Lemma 3, we have*

$$\begin{aligned} E_\theta[\mathbf{K_2^1} \circ \mathbf{K_2^1}] &= E_\theta[\mathbf{K_{1,2}^m}] \\ &= E_\theta[p(\mathbf{K_{1,2}^m}; K_{1,2}, \theta)] \cdot K_{1,2} = \tfrac{1}{3}K_{1,2}, \end{aligned}$$

*and similarly*

$$\begin{aligned} E_\theta[\mathbf{K_2^1} \circ \mathbf{N_2^1}] &= \tfrac{1}{3}S_3 + \tfrac{1}{3}K_{1,2} \\ E_\theta[\mathbf{N_2^1} \circ \mathbf{N_2^1}] &= N_3 + \tfrac{1}{3}S_3 \end{aligned}$$

*Using Lemma 5, we have*

$$\begin{aligned}
0 &\le E_\theta[(\mathbf{K_2^1} - \mathbf{N_2^1})^2] = E_\theta[(\mathbf{K_2^1})^2] - 2E_\theta[\mathbf{K_2^1}\cdot\mathbf{N_2^1}] + E_\theta[(\mathbf{N_2^1})^2] \\
&\le E_\theta[\mathbf{K_2^1}\circ\mathbf{K_2^1}] - 2E_\theta[\mathbf{K_2^1}\circ\mathbf{N_2^1}] + E_\theta[\mathbf{N_2^1}\circ\mathbf{N_2^1}] + \tfrac{16}{n} \\
&= \tfrac{1}{3}K_{1,2} - 2(\tfrac{1}{3}S_3 + \tfrac{1}{3}K_{1,2}) + (N_3 + \tfrac{1}{3}S_3) + \tfrac{16}{n} \\
&= N_3 - \tfrac{1}{3}S_3 - \tfrac{1}{3}K_{1,2} + \tfrac{16}{n} \qquad (1)
\end{aligned}$$

*Recall from Exercise 2 that* $K_2 = \frac{1}{3}S_3 + \frac{2}{3}K_{1,2}$. *Adding to this half of (1), we obtain*

$$\begin{aligned}
K_2 &\le \tfrac{1}{2}N_3 + \tfrac{1}{6}S_3 + \tfrac{1}{2}K_{1,2} + \tfrac{8}{n} \\
&\le \tfrac{1}{2}(N_3 + S_3 + K_{1,2}) + \tfrac{8}{n} = \tfrac{1}{2} + \tfrac{8}{n}.
\end{aligned}$$

*Hence, we have* $\|G\| \le \left(\frac{1}{2} + \frac{8}{n}\right)\binom{n}{2} \le \frac{n^2}{4} + 4n$ *for every triangle-free graph on* $n$ *vertices. Moreover, note that the inequality could be improved if* $S_3 > 0$; *hence, in any extremal graph, the density of* $S_3$ *must be very close to 0 (and from this, one can see that the extremal graphs are close to being bipartite).*

*Let us remark that we can easily improve this bound to the optimal one: Suppose* $G$ *is a triangle-free graph with* $n$ *vertices and* $cn^2$ *edges. Let* $G'$ *be the graph obtained from* $G$ *by blowing up each vertex into an independent set of* $k$ *vertices (turning edges of* $G$ *into complete bipartite subgraphs in* $G'$*). Clearly,* $G'$ *is also triangle-free. Moreover,* $G'$ *has* $nk$ *vertices and* $cn^2k^2$ *edges, and using the inequality from the previous paragraph, we have* $cn^2k^2 \le \frac{n^2k^2}{4} + 4nk$, *and thus* $c \le \frac{1}{4} + \frac{4}{nk}$. *Since this holds for every* $k$, *we have* $c \le 1/4$. *Consequently, every triangle-free graph with* $n$ *vertices has at most* $n^2/4$ *edges.*

Let us remark that generally, we do not need to care about the lower-order term $O(1/n)$ arising from the usage of Lemma 5; we can just ignore it throughout the calculations and add it to the final result. There are two (basically equivalent) approaches how to deal with this formally.

- Razborov introduced the notion of flag algebras, whose elements are formal linear combinations of flags and the multiplication is defined via the identities from Lemma 4 and Lemma 5, factorized by the identities given by Lemma 1. The elements of the algebra are then given a semantics (assigning to each flag $\mathbf{F}$ the mapping $p(\mathbf{F};\bullet)$) and it is argued in the natural way that all true statements in the flag algebra are asymptotically true in this interpretation.

- Lovász introduced the notion of convergent sequences. A sequence $\vec{G} = G_1, G_2, \ldots$ is convergent if for every graph $F$, there exists a limit

$\lim_{n\to\infty} p(F; G_n)$; we denote this limit by $p(F; \vec{G})$. The identities we obtain in the limit are then exact, i.e., we have $p(N_3; \vec{G}) - \frac{1}{3}p(S_3; \vec{G}) - \frac{2}{3}p(K_{1,2}; \vec{G}) = 0$ for any convergent sequence $\vec{G}$ of triangle-free graphs.

Suppose we in such a way show that for any convergent sequence $\vec{G}$ of $F$-free graphs, we have $p(K_2; \vec{G}) \leq a$. This implies that $\overline{\mathrm{ex}}(F; \infty) \leq a$: For contradiction assume that for some $\varepsilon > 0$, there exist arbitrarily large graphs $G$ such that $p(K_2, G) \geq a + \varepsilon$, and we let $\vec{A}$ be a sequence of such graphs with $|A_i| \to \infty$ as $i \to \infty$. It is easy to show that from any infinite sequence of graphs, we can select an infinite convergent subsequence. Letting $\vec{G}$ be an infinite convergent subsequence of $\vec{A}$, we obtain the contradiction.

In the latter approach, we have defined a notion of convergence of a sequence of graphs. It is natural to ask whether there exists a limit object towards which the sequence converges. One can indeed define such a natural object, a graphon (a symmetric measurable function $g : [0, 1]^2 \to [0, 1]$, where $g(x, y)$ can be intuitively interpreted as the probability that the vertices $x$ and $y$ are joined by an edge). Thus we can similarly interpret the identities as exact statements on graphons (with $p(F; g)$ defined appropriately for a graph $F$ and a graphon $g$).