On the Number of Pentagons in Triangle-Free Graphs

Hamed Hatami* Jan Hladký† Daniel Král'‡ Serguei Norine§
Alexander Razborov¶

Abstract

Using the formalism of flag algebras, we prove that the maximal number of copies of C_5 in a triangle-free graph with $5\ell + a$ vertices ($0 \leq a \leq 4$) is $\ell^{5-a}(\ell+1)^a$, and we show that the set of extremal graphs for this problem consists precisely of almost balanced blow-ups of a single pentagon. This settles a conjecture made by Erdős in 1984. For the transition from an asymptotic version of our result to the exact one, we introduce a new technique based on replacing finite objects by their infinite blow-ups which we expect to have further applications.

1. Introduction

Triangle-free graphs need not be bipartite. But how exactly far from being bipartite can they be?

In 1984, Erdős [Erd84, Questions 1 and 2] considered three quantitative refinements of this question. More precisely, he proposed to measure "non-bipartiteness" by:

---

*School of Computer Science, McGill University, hatami@cs.mcgill.ca.

†Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University in Prague *and* DIMAP, University of Warwick, honzahladky@gmail.com. Supported by EPSRC award EP/D063191/1.

‡Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University in Prague, kral@kam.mff.cuni.cz. The work leading to this invention has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 259385.

§Department of Mathematics, Princeton University, snorin@math.princeton.edu.

¶University of Chicago, razborov@cs.uchicago.edu. Part of this work was done while the author was at Steklov Mathematical Institute, supported by the Russian Foundation for Basic Research, and at Toyota Technological Institute, Chicago.

(i) the minimal possible number of edges in a subgraph spanned by half of the vertices;

(ii) the minimal possible number of edges that have to be removed to make the graph bipartite;

(iii) the number of copies of pentagons (cycles of length 5) in the graph.

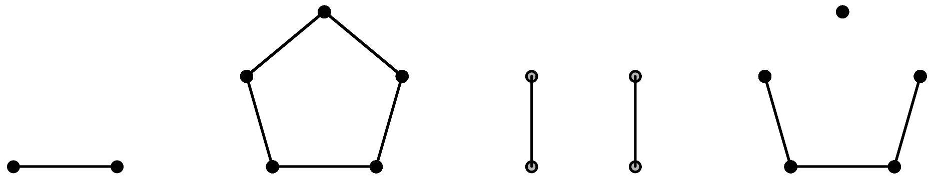
All these parameters vanish on bipartite graphs, and Erdős conjectured that in the class of triangle-free graphs every one of them is maximized by balanced blow-ups of the pentagon. Simonovits (referred to in [Erd84]) observed that another example which attains the conjectured extremum for **(i)** is provided by balanced blow-ups of the Petersen graph.

The first two of the above Erdős's questions have been investigated in [EFPS88, Kri95, KS06]. Győri investigated the third question in [Győ89]. In terms of densities, Erdős's conjecture regarding **(iii)** states that the density of pentagons in any triangle-free graphs is at most $\frac{5!}{5^5}$. Győri proved an upper bound of $\frac{3^3 \cdot 5!}{5 \cdot 2^{14}}$ that is within a factor 1.03 of the optimal. Füredi (private communication) refined Győri's approach and obtained an upper bound within a factor 1.001 of the optimal.

In this paper we completely settle this conjecture. Moreover, we prove that an (almost) balanced blow-up of the pentagon is the only extremal configuration, both in asymptotic and exact sense.

The proof of the asymptotic version of this result (Theorem 3.1) is a rather standard Cauchy-Schwarz calculation in flag algebras (introduced in [Raz07]). Furthermore, we obtain the asymptotic uniqueness by a relatively simple argument in Theorem 3.2. But then in the proof of the exact result our path substantially deviates from the standard approach based on stability, removal lemmas, etc. (see e.g. [KS05, Pik09]). Instead of trying to argue directly about finite graphs, we convert them into certain limit objects (using the theory of flag algebras) using infinite blow-ups and then apply to the resulting object the same analytical methods that were used in the proof of Theorem 3.2. This allows us to obtain the exact result from the asymptotic one basically for free. We feel that this general approach might turn out to be interesting in its own right.

We assume certain familiarity with the theory of flag algebras from [Raz07] (for the proof of the central Theorem 3.1 only the most basic notions are required). Thus, instead of trying to duplicate definitions, we occasionally give pointers to relevant places in [Raz07] and some subsequent papers.



ρ C_5 M_4 C_5^-

Figure 1: Models

2. Preliminaries

2.1. Notation

We denote vectors with bold font, e.g. $\mathbf{a} = (\mathbf{a}(1), \mathbf{a}(2), \mathbf{a}(3))$ is a vector with three coordinates. For every positive integer k, let $[k]$ denote the set $\{1, \ldots, k\}$. Following [Raz07, Definition 1], for two graphs H and G, the density of H in G as an induced subgraph is denoted by $p(H, G)$. That is $p(H, G)$ is the probability that the subgraph induced on $|V(H)|$ randomly chosen vertices of G is isomorphic to H.

Except for the stand-alone Section 2.2, we exclusively work [Raz07, §2] in the theory $T_{\mathrm{TF-Graph}}$ of triangle-free graphs. Recall from [Raz07] that for a theory T and a positive integer n, the set of all finite models of T up to an isomorphism is denoted by $\mathcal{M}_n[T]$. We work with the notion of types, flags, and flag algebras, and use the same notation as in [Raz07, §2.1] where this terminology is introduced. Let us list those models, types and flags that will be needed in this paper.

Let $\rho \in \mathcal{M}_2[T_{\mathrm{TF-Graph}}]$ and $C_5 \in \mathcal{M}_5[T_{\mathrm{TF-Graph}}]$ respectively denote the edge and the pentagon. These two graphs along with two other graphs that will be needed for proving the uniqueness and the exact result are illustrated in Figure 1.

We denote the *trivial type* of size 0 by 0. Let P denote the type of size 5 based on C_5 (see Figure 2). For $i = 0, 1, 2$, let σ_i denote the type of size 3 with i edges where the labeling is chosen in such a way that the permutation of 1 and 2 is an automorphism (see Figure 2).

For a type σ of size k and an independent set of vertices $V \subseteq [k]$ in σ, let F_V^σ denote the flag $(G, \theta) \in \mathcal{F}_{k+1}^\sigma$ in which the only unlabeled vertex v is connected to the set $\{\theta(i) \,|\, i \in V\}$. Note that since we are working in the theory of triangle-free graphs, we have

$$\mathcal{F}_{k+1}^\sigma = \{F_V^\sigma \mid V \subseteq [k] \text{ is an independent set in } \sigma\}.$$

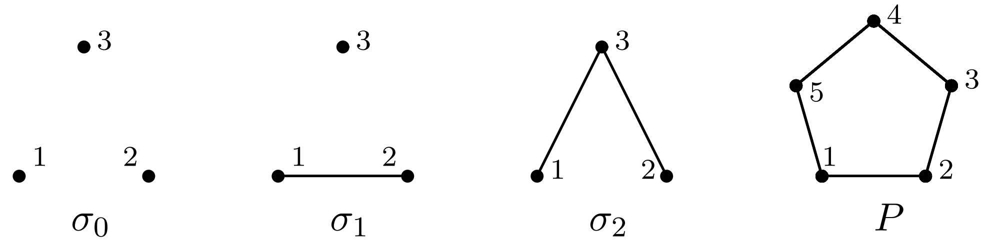


Figure 2: Types

2.2. Infinite blow-ups and $\mathrm{Hom}^+(\mathcal{A}^0[T_{\mathrm{Graph}}], \mathbb{R})$

In order to convert the asymptotic result into the exact one, we need to explore a little bit more the connection between blow-ups of a graph and the corresponding algebra homomorphism from $\mathrm{Hom}^+(\mathcal{A}^0[T_{\mathrm{Graph}}], \mathbb{R})$ already used in a similar context in [Raz08, Theorem 4.1].

For a finite graph G and a positive integer vector $\mathbf{k} = (\mathbf{k}(v) \mid v \in V(G))$, we define the *blow-up* $G^{(\mathbf{k})}$ of G as the graph with

$$V(G^{(\mathbf{k})}) \stackrel{\mathrm{def}}{=} \bigcup_{v \in V(G)} \{v\} \times [\mathbf{k}(v)]$$

$$E(G^{(\mathbf{k})}) \stackrel{\mathrm{def}}{=} \{((v,i),(w,j)) \mid v \neq w \ \wedge \ (v,w) \in E(G)\}.$$

When all $\mathbf{k}(v)$ are equal to some positive integer k, the corresponding blow-up is called *balanced* and denoted simply by $G^{(k)}$.

For every graph H, it is easy to see that the sequence $\{p(H, G^{(k)})\}_{k \in \mathbb{N}}$ is convergent. It follows [Raz07, §3] that there exists a homomorphism $\phi_G \in \mathrm{Hom}^+(\mathcal{A}^0[T_{\mathrm{Graph}}], \mathbb{R})$ such that for every graph H, we have

$$\lim_{k \to \infty} p(H, G^{(k)}) = \phi_G(H).$$

Note that since the blow-up of a triangle-free graph is also triangle-free, if G is triangle-free, then actually $\phi_G \in \mathrm{Hom}^+(\mathcal{A}^0[T_{\mathrm{TF-Graph}}], \mathbb{R})$.

Let us now give a combinatorial description of ϕ_G. For a finite graph H, let us denote by $s(H,G)$ the number of *strong homomorphisms* from H to G that we define as those mappings $\alpha : V(H) \longrightarrow V(G)$ for which $(\alpha(v), \alpha(w)) \in E(G)$ if and only if $(v,w) \in E(H)$. This notion is a natural hybrid of induced embeddings and graph homomorphisms, but we have not seen it in the literature before. It is easy to check that

$$\phi_G(H) = \frac{m!}{|\mathrm{Aut}(H)|} \cdot \frac{s(H,G)}{n^m}, \tag{1}$$

where m and n respectively denote $|V(H)|$ and $|V(G)|$, and $\mathrm{Aut}(H)$ is the group of automorphisms of H.

Let us say that H is *twin-free* if no two vertices in H have the same set of neighbors. Every strong homomorphism of a twin-free graph into any other graph is necessarily an induced embedding. Therefore, for twin-free H, we have $s(H,G) = p(H,G)\binom{n}{m}|\mathrm{Aut}(H)|$ and (1) considerably simplifies to

$$\phi_G(H) = p(H,G)\cdot\frac{n(n-1)\dots(n-m+1)}{n^m}. \tag{2}$$

For the partial case $H = K_r$, this formula was already used in [Raz08, Section 4.1]), and in this paper we are interested in another partial case

$$\phi_{C_5}(C_5) = \frac{5!}{5^5}.$$

Our approach to extracting exact results from asymptotic ones heavily relies on the fact that ϕ_G is a graph invariant:

Theorem 2.1 *Let G_1 and G_2 be finite graphs with the same number of vertices and such that $\phi_{G_1} = \phi_{G_2}$. Then G_1 and G_2 are isomorphic.*

Below we present a self-contained proof, even though Theorem 2.1 follows from a celebrated result of Lovász [Lov67]. We believe that our proof is simpler.

Proof. Given a graph G, define the equivalence relation $\approx$ on $V(G)$ by letting $u \approx v$ if and only if u and v have the same sets of neighbors. Clearly, all classes of this relation are independent sets in G and we can form the factor-graph $\widetilde{G} \stackrel{\text{def}}{=} G/\approx$. Note that $\widetilde{G}$ is twin-free and $G = \widetilde{G}^{(\mathbf{k})}$, where $\mathbf{k}$ is the vector of cardinalities of equivalence classes. This factorization of every finite graph as a blow-up of a twin-free graph is unique.

Let G_1 and G_2 be as in the statement of the theorem, and $G_1 = \widetilde{G_1}^{(\mathbf{k}_1)}$ and $G_2 = \widetilde{G_2}^{(\mathbf{k}_2)}$ be their corresponding representations. Clearly $\phi_{G_1}(\widetilde{G_1}) > 0$ and thus $\phi_{G_2}(\widetilde{G_1}) > 0$ which by (1) implies $s(\widetilde{G_1}, G_2) > 0$. In particular there exists at least one strong homomorphism from $\widetilde{G_1}$ to G_2. Composing this strong homomorphism with the natural projection $G_2 \to \widetilde{G_2}$ (that is also a strong homomorphism), we obtain a strong homomorphism from $\widetilde{G_1}$ to $\widetilde{G_2}$. Since $\widetilde{G_1}$ is twin-free, this implies that it is an induced subgraph of $\widetilde{G_2}$. By the same token, $\widetilde{G_2}$ is an induced subgraph of $\widetilde{G_1}$, and hence they are isomorphic.

So far we have established that there exists a twin-free graph $\widetilde{G}$ such that $G_1 = \widetilde{G}^{(\mathbf{k}_1)}$ and $G_2 = \widetilde{G}^{(\mathbf{k}_2)}$ for two nonnegative vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ with $\|\mathbf{k}_1\|_1 = \|\mathbf{k}_2\|_1$. Next we need to show that the vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ also coincide up to an automorphism of $\widetilde{G}$. To this end, consider the blow-up $\widetilde{G}^{(\mathbf{x})}$, where $\mathbf{x}$ is an unknown vector of positive integers. By (1), we have

$$s\left(\widetilde{G}^{(\mathbf{x})}, \widetilde{G}^{(\mathbf{k}_1)}\right) = s\left(\widetilde{G}^{(\mathbf{x})}, \widetilde{G}^{(\mathbf{k}_2)}\right). \tag{3}$$

Since $\widetilde{G}$ is twin-free, it is easy to see that all strong homomorphisms from $\widetilde{G}^{(\mathbf{x})}$ to $\widetilde{G}^{(\mathbf{k}_1)}$ result from automorphisms of $\widetilde{G}$, and thus for $i = 1, 2$, we have

$$s(\widetilde{G}^{(\mathbf{x})}, \widetilde{G}^{(\mathbf{k}_i)}) = \sum_{\gamma \in \mathrm{Aut}(\widetilde{G})} \prod_{v \in V(\widetilde{G})} \mathbf{k}_i(\gamma(v))^{\mathbf{x}(v)}.$$

Hence (3) in particular implies that for all positive integer vectors $\mathbf{x}$, we have

$$\sum_{\gamma \in \mathrm{Aut}(\widetilde{G})} \prod_{v \in V(\widetilde{G})} \mathbf{k}_1(\gamma(v))^{\mathbf{x}(v)} - \sum_{\gamma \in \mathrm{Aut}(\widetilde{G})} \prod_{v \in V(\widetilde{G})} \mathbf{k}_2(\gamma(v))^{\mathbf{x}(v)} = 0. \tag{4}$$

We claim that then the left-hand side of (4) must be syntactically equal to zero, that is there exists a bijection $b : \mathrm{Aut}(\widetilde{G}) \to \mathrm{Aut}(\widetilde{G})$ such that $\mathbf{k}_1(\gamma(v)) = \mathbf{k}_2(b(\gamma)(v))$ for each $v \in V(\widetilde{G})$ and $\gamma \in \mathrm{Aut}(\widetilde{G})$. This indeed finishes the proof as then $b(\mathrm{id.})$ provides an isomorphism between G_1 and G_2. Now to prove the claim suppose to the contrary that such a bijection b does not exist. Then the left-hand side of (4) can be written as

$$\text{L.H.S. of (4)} = \sum_{i=1}^{h} \alpha_i \prod_{v \in V(\widetilde{G})} a_i(v)^{\mathbf{x}(v)} = \sum_{i=1}^{h} \alpha_i \exp\left(\sum_{v \in V(\widetilde{G})} \mathbf{x}(v) \ln \mathbf{a}_i(v)\right), \tag{5}$$

for a positive integer h, *distinct* positive integer vectors $\mathbf{a}_i$, and not all equal to zero coefficients $\alpha_i \in \mathbb{R}$. Choose an arbitrary positive integer vector $\mathbf{x}_0$ such that the sums $\sum_v \mathbf{x}_0(v) \ln \mathbf{a}_i(v)$ (where $1 \le i \le h$) are all distinct. Since all these sums are non-negative, setting $\mathbf{x} = t\mathbf{x}_0$ for a sufficiently large integer t creates a dominant term in (5). This shows that (5) is not always equal to zero; a contradiction.■

3. Main results

Recall [Raz07, Definitions 5 and 6] that for a non-degenerate type σ in a theory T, and $f, g \in \mathcal{A}^\sigma[T]$, the inequality $f \le_\sigma q$ means that $\phi(f) \le \phi(g)$

for every $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma[T], \mathbb{R})$: this is the class of all inequalities that hold asymptotically on flags of the given theory [Raz07, Corollary 3.4]. We abbreviate $f \leq_\sigma q$ to $f \leq q$ when σ is clear from the context. Our first theorem, which answers the question of Erdős, says that in the theory of triangle-free graphs, we have $C_5 \leq \frac{5!}{5^5}$. Note that while in the theory of general graphs, the flag C_5 corresponds to *induced* pentagons, in the theory of triangle-free graphs, every pentagon is induced.

Theorem 3.1 *In the theory* $T_{\mathrm{TF-Graph}}$*, we have*

$$C_5 \leq \frac{5!}{5^5}.$$

Proof. The proof is by a direct computation in the flag algebra $\mathcal{A}^0[T_{\mathrm{TF-Graph}}]$ (cf. [Raz10a, HKN09] and [Raz10b, Section 4.1]). We claim that

$$62500C_5 + \frac{1097}{12}M_4 + \frac{68}{3}C_5^- + \left(\sum_{i=0}^{2} [\![Q^+_{\sigma_i}(\mathbf{g}_i^+)]\!]_{\sigma_i}\right) + \\ 200\left(\rho - \frac{2}{5}\right)^2 + [\![Q^-_{\sigma_1}(\mathbf{g}_1^-)]\!]_{\sigma_1} + 158266[\![(F^{\sigma_2}_{\{1\}} - F^{\sigma_2}_{\{2\}})^2]\!]_{\sigma_2} \leq 2400. \quad (6)$$

The graphs M_4 and C_5^- are illustrated in Figure 1. For the definition of the algebra operations see [Raz07, Eq. (5)], and for the definition of the averaging operator $[\![\cdot]\!]$ see [Raz07, §2.2]. Let us now define the notations $\mathbf{g}_i^{+/-}$ and $Q_i^{+/-}$ in (6). For a type σ of size k and an integer $0 \leq j \leq k$, we let

$$f_j^\sigma \stackrel{\mathrm{def}}{=} \sum \{ F_V^\sigma \mid V \subseteq [k] \text{ an independent set of size } j \text{ in } \sigma \},$$

where F_V^σ are as defined in Section 2.1. The vectors $\mathbf{g}_i^{+/-}$ are the following tuples of elements from $\mathcal{A}_4^{\sigma_i}$:

$$\begin{aligned}
\mathbf{g}_0^+ &\stackrel{\mathrm{def}}{=} (f_1^{\sigma_0} - f_2^{\sigma_0},\ f_1^{\sigma_0} - 2f_2^{\sigma_0} + 3f_3^{\sigma_0});\\
\mathbf{g}_1^+ &\stackrel{\mathrm{def}}{=} (2f_0^{\sigma_1} - f_1^{\sigma_1},\ f_1^{\sigma_1} - f_2^{\sigma_1},\ F^{\sigma_1}_{\{3\}});\\
\mathbf{g}_2^+ &\stackrel{\mathrm{def}}{=} (6f_0^{\sigma_2} + f_1^{\sigma_2} - 4f_2^{\sigma_2},\ 2f_0^{\sigma_2} - 2f_2^{\sigma_2} + F^{\sigma_2}_{\{3\}});\\
\mathbf{g}_1^- &\stackrel{\mathrm{def}}{=} (F^{\sigma_1}_{\{1\}} - F^{\sigma_1}_{\{2\}},\ F^{\sigma_1}_{\{2,3\}} - F^{\sigma_1}_{\{1,3\}}),
\end{aligned}$$

and $Q_i^{+/-}$ are positive-definite quadratic forms represented by the following positive-definite matrices:

$$M_0^+ \stackrel{\rm def}{=} \begin{pmatrix} 9760 & 2252 \\ 2252 & 592 \end{pmatrix} \qquad M_1^+ \stackrel{\rm def}{=} \begin{pmatrix} 13900 & -671 & -12807 \\ -671 & 31334 & -51136 \\ -12807 & -51136 & 98157 \end{pmatrix}$$

$$M_2^+ \stackrel{\rm def}{=} \begin{pmatrix} 22708 & -40788 \\ -40788 & 78132 \end{pmatrix} \qquad M_1^- \stackrel{\rm def}{=} \begin{pmatrix} 1416 & -16452 \\ -16452 & 256488 \end{pmatrix}.$$

The inequality (6) can be checked by expanding the left-hand side as a linear combination of elements from $\mathcal{M}_5$ (that is, triangle-free graphs on 5 vertices – there are 14 of them) and checking that all coefficients are less or equal than 2400.

It follows from [Raz07, Theorem 3.14] that all the summands on the left-hand side of (6) are nonnegative as elements of $\mathcal{A}^0[T_{\rm TF-Graph}]$ which in turn implies that $C_5 \leq \frac{2400}{62500} = 5!/5^5$.■

Remark 1 An explanation of our usage of the +/− superscripts in the proof of Theorem 3.1 can be found in [Raz10a, Section 4]. Here, the particular choice of subspaces spanned by the vectors $\mathbf{g}_i^{+/-}$ is dictated by the same principles as in [Raz10a, Section 4].

Let us now turn to the question of uniqueness, and we begin with the asymptotic version. Note that while Theorem 3.1 is subsumed by our exact result (Theorem 3.3 below), Theorem 3.2 is *not*.

Theorem 3.2 *The homomorphism* ϕ_{C_5} *is the unique element in* $\mathrm{Hom}^+(\mathcal{A}^0[T_{\rm TF-Graph}], \mathbb{R})$ *that fulfills*

$$\phi(C_5) = \frac{5!}{5^5}. \tag{7}$$

Proof. Fix $\phi \in \mathrm{Hom}^+(\mathcal{A}^0[T_{\rm TF-Graph}], \mathbb{R})$ such that (7) holds. Then (6) implies that

$$\phi(M_4) = \phi(C_5^-) = 0. \tag{8}$$

Recall that P is the type of size 5 based on C_5 (see Figure 2). Pick an arbitrary $\phi^P \in \mathrm{Hom}^+(\mathcal{A}^P[T_{\rm TF-Graph}], \mathbb{R})$ from the support of the measure $\mathbf{P}^P$ (see [Raz07, Definition 8]). Let us examine $\phi^P(F_V^P)$ for flags $F_V^P \in \mathcal{F}_6^P$.

Since V is an independent set in C_5, we have $|V| \leq 2$, and moreover, if $|V| = 2$, then $V = \{i-1, i+1\}$ for some $i \in \mathbb{Z}_5$. Trivially $[\![F_\emptyset^P]\!]_P \leq C_5^-$, and there exists a constant $\alpha > 0$ such that $[\![F_{\{i\}}^P]\!]_P \leq \alpha M_4$ for every $i \in \mathbb{Z}_5$. Hence (8) implies that $\phi([\![F_\emptyset^P]\!]_P) = \phi([\![F_{\{i\}}^P]\!]_P) = 0$, and furthermore, since ϕ^P belongs to the support of the measure $\mathbf{P}^P$, it follows from [Raz07, Eq. (17)] that $\phi^P(F_\emptyset^P) = \phi^P(F_{\{i\}}^P) = 0$. In other words, $\phi^P(F_V^P)$ can be non-zero only when $V = \{i-1, i+1\}$ for some $i \in \mathbb{Z}_5$. Define $H_i^P \stackrel{\text{def}}{=} F_{\{i-1,i+1\}}^P$ and $p_i \stackrel{\text{def}}{=} \phi^P(H_i^P)$ so that $\sum_{i \in \mathbb{Z}_5} p_i = 1$.

Recall from [Raz07, §2.3.1] that $\phi(C_5) = \phi^P(\pi^P(C_5))$ where $\pi^P(C_5)$ can be represented as the sum of those $F = (G, \theta) \in \mathcal{F}_{10}^P$ for which the unlabeled vertices form a copy of C_5, say $V(G) \setminus \mathrm{im}(\theta) = \{v_1, v_2, \ldots, v_5\}$ where v_j is adjacent to v_{j-1} and v_{j+1}. By the above discussion, non-zero contributions to $\phi^P(\pi^P(C_5))$ can be made only by those F for which every v_j is adjacent to $\theta(i(j)-1)$ and $\theta(i(j)+1)$ for some choice of $i(j) \in \mathbb{Z}_5$. Since G is triangle-free, the mapping $j \mapsto i(j)$ defines a graph homomorphism of the pentagon into itself, and since there are no such graph homomorphisms other than isomorphisms, we may assume without loss of generality that every v_j is adjacent to $\theta(j-1)$ and $\theta(j+1)$. In other words $F = (C_5^{(2)})^P$, where $(C_5^{(2)})^P$ is the uniquely defined P-flag based on $C_5^{(2)}$, the blow-up of the pentagon.

Since

$$(C_5^{(2)})^P \ \leq_P \ 5! \cdot \prod_{i \in \mathbb{Z}_5} H_i^P, \tag{9}$$

we have

$$\frac{5!}{5^5} = \phi(C_5) = \phi^P(\pi^P(C_5)) = \phi^P((C_5^{(2)})^P) \leq 5! \prod_{i \in \mathbb{Z}_5} \phi^P(H_i^P).$$

Consequently, we obtain that $\prod_{i \in \mathbb{Z}_5} p_i \geq 5^{-5}$. By the inequality of arithmetic and geometric means, this implies $p_i = 1/5$ (for all $i \in \mathbb{Z}_5$) and that there is no slackness in (9):

$$\phi^P\left(5! \cdot \prod_{i \in \mathbb{Z}_5} H_i^P - (C_5^{(2)})^P\right) = 0. \tag{10}$$

This equality allows us to completely describe the behavior of ϕ^P also on $\mathcal{F}_7^P$. For $i, j \in \mathbb{Z}_5$, let $H_{ij}^P \in \mathcal{F}_7^P$ be defined by adding two unlabeled non-adjacent vertices to P and connecting one of them to $\theta(i-1)$ and $\theta(i+1)$

and the other to $\theta(j-1)$ and $\theta(j+1)$. Note that if $i=j$ or $(i,j)\notin E(P)$, then the product $H_i^P H_j^P$ is equal to $q_{ij}H_{ij}^P$, where $q_{ij}=1$ if $i=j$, and $q_{ij}=1/2$ otherwise. Hence, in this case $\phi^P(H_{ij}^P)=\frac{1}{25q_{ij}}$. On the other hand, if $(i,j)\in E(P)$, then

$$H_{ij}^P\cdot\prod_{k\in\mathbb{Z}_5\setminus\{i,j\}}H_k^P\ \leq_P\ 5!\cdot\prod_{i\in\mathbb{Z}_5}H_i^P-(C_5^{(2)})^P,$$

which together with (10) implies that $\phi^P(H_{ij}^P)=0$. It follows that $\phi^P(G_{ij}^P)=\frac{1}{25q_{ij}}$, where G_{ij}^P is defined similar to H_{ij}^P with the difference that now there is an edge between the unlabeled vertices. As $\sum\phi^P(H_{ij}^P)+\sum\phi(G_{ij}^P)=1$, we have $\phi(F)=0$ for any other flag $F\in\mathcal{F}_7^P$.

With this knowledge in hand, for any fixed graph H on n vertices we can compute $\phi(H)$ as follows. Similarly as above we can write $\phi(H)=\phi^P(\pi^P(H))$. We expand $\pi^P(H)$ in $\mathcal{A}_{5+n}^P$. For a homomorphism $h:H\to\mathbb{Z}_5$ of H to C_5 (with its vertices labeled cyclically), we write F_h^P for the P-flag $(G,\theta)\in\mathcal{F}_{5+n}^P$ where the unlabeled vertices $V'\stackrel{\text{def}}{=}V(G)\setminus\text{im}(\theta)$ induce a copy of H, and each vertex $v\in V'$ is adjacent only to $\theta(h(v)-1)$ and $\theta(h(v)+1)$. The same reasoning as in the special case $H=C_5$ gives that ϕ^P evaluates to zero at any term $F\in\mathcal{F}_{5+n}^P$ in the expansion of $\pi^P(H)$, unless $F=F_h^P$ for some homomorphism $h:H\to P$. Furthermore, as $\phi^P(H_{ij}^P)=0$ for $(i,j)\in E(P)$, we actually have that h must be a strong homomorphism in this case. Observe that

$$\frac{\phi^P(F_h^P)}{C_h}=\phi^P\left(\prod_{v\in V(H)}H_{h(v)}^P\right)=5^{-n}\tag{11}$$

for each strong homomorphism $h:H\to P$, where C_h is the multinomial coefficient,

$$C_h\stackrel{\text{def}}{=}\binom{n}{|h^{-1}(1)|,\ |h^{-1}(2)|,\ |h^{-1}(3)|,\ |h^{-1}(4)|,\ |h^{-1}(5)|}.$$

It is easy to see that (11) holds when ϕ^P is replaced by $\phi_{C_5}^P$. Therefore $\phi=\phi_{C_5}$ as claimed.■

For an integer $n\geq 5$, let $\mathcal{E}(n)$ be the class of almost balanced blow-ups of the pentagon with n vertices. In other words, $\mathcal{E}(n)$ consists of all graphs $C_5^{(k_1,\ldots,k_5)}$, where $\sum_{i\in[5]}k_i=n$ and $|k_i-k_j|\leq 1$ for all $i,j\in[5]$. If $n=5\ell+a$, then the number of pentagons in every $G\in\mathcal{E}(n)$ is exactly

$$\chi(n)\stackrel{\text{def}}{=}\ell^{5-a}(\ell+1)^a.$$

Theorem 3.3 below asserts that each triangle-free graph G on n vertices contains at most $\chi(n)$ pentagons. Observe that a weaker upper-bound of $(\frac{n}{5})^5$ follows by applying Theorem 3.1 on the infinite blow-up $\phi_G \in \mathrm{Hom}^+(\mathcal{A}^0[T_{\mathrm{TF-Graph}}], \mathbb{R})$. This bound is sharp when n is a multiple of five as $\chi(n) = (\frac{n}{5})^5$ in this case.

Theorem 3.3 *Every triangle-free graph G with n vertices has at most $\chi(n)$ pentagons, with the equality attained if and only if $G \in \mathcal{E}(n)$.*

Proof. Fix a triangle-free graph G with n vertices, and apply the argument in the proof of Theorem 3.2 to the homomorphism ϕ_G. Note that by the nature of ϕ_G, for any type σ with $\phi_G(\sigma) > 0$, the measure $\mathbf{P}^\sigma$ is concentrated in finitely many points, corresponding to strong homomorphisms from σ to G. Furthermore for every such point ϕ^σ and every flag $F \in \mathcal{F}^\sigma_{|\sigma|+1}$, the value $\phi^\sigma(F)$ is a multiple of $1/n$. In particular, this is true for the flags H_i^P (defined in the proof of Theorem 3.2) which allows us to draw a slightly better conclusion from (9). Namely, denoting again $\phi^P(H_i^P)$ by p_i, we see that $\frac{1}{5!}\phi^P((C_5^{(2)})^P)$ does not exceed the value of the following integer program:

$$\begin{array}{rl} \textbf{Max} & p_1 p_2 \ldots p_5 \\ \text{subject to} & p_1, \ldots, p_5 \in \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right\}; \\ & p_1 + p_2 + \ldots + p_5 = 1. \end{array}$$

Every optimum of this program must satisfy $p_j - p_i \leq \frac{1}{n}$ for any pair $i, j \in [5]$, as otherwise the value of the goal function could be increased by setting $p_i := p_i + \frac{1}{n}$ and $p_j := p_j - \frac{1}{n}$. Hence the optimal value is equal to $\frac{\chi(n)}{n^5}$, and the only extremal points are obtained by setting exactly $5 - a$ of $p_1, \ldots, p_5$ to $\frac{\ell}{n}$ and the other a to $\frac{\ell+1}{n}$. Consequently, $\phi_G(C_5) \leq \frac{5!\chi(n)}{n^5}$. Now from (2), we conclude that

$$p(C_5, G) = \frac{n^5}{n(n-1)\ldots(n-4)}\phi_G(C_5) \leq \binom{n}{5}^{-1} \chi(n)$$

which proves the first part of the theorem.

For the second part, once we know that there is no slackness in (9), we can literally repeat the rest of the proof with the difference that ϕ_G is simulated not by ϕ_{C_5}, but by $\phi_{G'}$ where $G' \stackrel{\mathrm{def}}{=} C_5^{(p_1 n, \ldots, p_5 n)} \in \mathcal{E}(n)$. Then $G \approx G'$ follows from Theorem 2.1.■

Acknowledgement

When preparing the final version of this paper we learned that Andrzej Grzesik proved independently Theorem 3.1 ([Grz11]).

We thank Oleg Pikhurko for pointing out the work of Zoltán Füredi to us.

References

[EFPS88] Paul Erdős, Ralph Faudree, János Pach, and Joel Spencer. How to make a graph bipartite. *J. Combin. Theory Ser. B*, 45(1):86–98, 1988.

[Erd84] Paul Erdős. On some problems in graph theory, combinatorial analysis and combinatorial number theory. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 1–17. Academic Press, London, 1984.

[Grz11] Andrzej Grzesik. On the maximum number of C_5's in a triangle-free graph. arXiv:1102.0962, 2011.

[Győ89] Ervin Győri. On the number of C_5's in a triangle-free graph. *Combinatorica*, 9(1):101–102, 1989.

[HKN09] Jan Hladký, Daniel Král, and Serguei Norine. Counting flags in triangle-free digraphs. arXiv:0908.2791, 2009.

[Kri95] Michael Krivelevich. On the edge distribution in triangle-free graphs. *J. Combin. Theory Ser. B*, 63(2):245–260, 1995.

[KS05] Peter Keevash and Benny Sudakov. The Turán number of the Fano plane. *Combinatorica*, 25(5):561–574, 2005.

[KS06] Peter Keevash and Benny Sudakov. Sparse halves in triangle-free graphs. *J. Combin. Theory Ser. B*, 96(4):614–620, 2006.

[Lov67] László Lovász. Operations with structures. *Acta Math. Acad. Sci. Hungar.*, 18:321–328, 1967.

[Pik09] Oleg Pikhurko. The minimum size of 3-graphs without a 4-set spanning no or exactly three edges. Manuscript, 2009.

[Raz07] Alexander A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007.

[Raz08] Alexander A. Razborov. On the minimal density of triangles in graphs. *Combin. Probab. Comput.*, 17(4):603–618, 2008.

[Raz10a] Alexander A. Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM J. Discrete Math.*, 24(3):946–963, 2010.

[Raz10b] Alexander A. Razborov. On the Fon-der-Flaass interpretation of extremal examples for Turán's (3,4)-problem. arXiv:1008.4707, 2010.