Part III Essay

32. Flag Algebras in Graph Theory

Contents

1 Introduction

1.1 Essay overview

Flag algebras are a method of establishing bounds in asymptotic extremal combinatorics. They were formally introduced by Alexander Razborov in [Raz07]. The method works by creating objects called flags, and then equating "convergent" sequences of graphs with special homomorphisms from flag algebras to $\mathbb{R}$. Thus, if you can show that a certain inequality on flags holds under the application of any such homomorphism, then the corresponding inequality of flag densities must hold for every convergent sequence. In that way, concrete asymptotic results can be obtained from syntactic reasoning.

In this essay, my primary goal will be to explain the necessary formality and intuition with which to apply flag algebras to prove results in extremal combinatorics. I will focus particularly on the following three questions:

i) To what classes of structure can the method of flag algebras be applied?
ii) How do you construct a proof using semi-definite programming (SDP) over flag algebras?
iii) How can you apply asymptotic results to prove results about finite objects?

As a side goal, I will build up to the exposition of the paper *On the number of pentagons in triangle-free graphs* [Hat+13] in section 7.

I give a (hopefully intuitive) overview of the general idea in section 2. Model Theory is introduced in section 3, in an attempt to answer question i). Section 4 formally introduces the method, first looking at the syntax and tools to construct and manipulate the flag algebras, and then looking at how these manipulations are used to yield asymptotic results. Full, explained proofs underlying the core of the method will be given. Section 5 discusses how the method can be applied to construct proofs in practice, and will look at common techniques for extracting results from them, including examining questions ii) and iii).

The essay will then turn to Razborov's paper *On 3-hypergraphs with forbidden 4-vertex configurations* [Raz10] in section 6 to demonstrate some of the SDP techniques, and finish with an exposition of the paper *On the number of pentagons in triangle-free graphs* [Hat+13] (Hatami et al) in section 7.

With regard to attribution, I will attribute the source/s of material either inline, or at the start of each section or subsection as appropriate. All images are my own. My own proofs are highlighted where necessary so as to claim responsibility for any errors they may contain.

1.2 Notation and common definitions

Below is some general notation, and some notation that will be used in section 2.

General combinatorial definitions and notation
Let $[k] \stackrel{\text{def}}{=} \{1, 2, \dots, k\}$, and let $A^{(k)} \stackrel{\text{def}}{=} \{S \subset A : |S| = k\}$. A collection of sets $V_1, \dots, V_k$, is a *sunflower with centre* C if for all $i \neq j$, $V_i \cap V_j = C$. The V_i are called the *petals* of the sunflower. Following Razborov, I will use ***math bold face*** to denote random objects, which will be assumed to be picked uniformly unless otherwise mentioned.

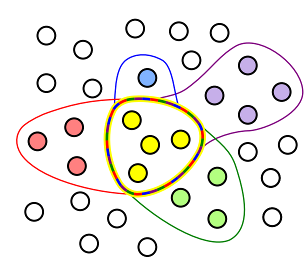


Figure 1: A sunflower with yellow centre.

2-graph notation and definitions
Now, I will give some notation for working with standard 2-graphs in the next section. The same notation will be generalized in a natural way in section 4 for a much larger class of combinatorial structures.

I will consider graphs G on vertex set $V(G)$. I will let the number of vertices of G be called the *size* or *order* of G, and denote this $|G|$. For some subset $V \subset V(G)$, let $G|_V$ denote the *induced* subgraph on V. Let $\mathcal{G}_n$ denote the family of all graphs on n vertices, up to isomorphism. Similarly, let $\mathcal{G} = \cup_{n \in \mathbb{N}} \mathcal{G}_n$.

For two graphs, H and G, we let $p(H, G)$ denote the probability that $H \cong G|_{\boldsymbol{V}}$, where $\boldsymbol{V}$ is a randomly chosen subset of $V(G)$ of size $|H|$. Define $p(H, G) = 0$ if $|H| > |G|$.

2 Flag Algebras: An intuitive overview

In this section, I will restrict explanation to 2-graphs. This will extend in an obvious manner to a larger class of structures introduced in the next section.

Densities in convergent sequences Consider a sequence (G_n) of graphs growing in size, such that $\lim_{n\to\infty} p(H, G_n)$ exists for all $H \in \mathcal{G}$. Such a sequence is known as a *convergent sequence.* (These are readily found: For any sequence (H_n) of graphs growing in size, and a fixed graph H, $p(H, H_n)$ is a bounded real sequence, so we can find a subsequence (H_{n_i}) such that $p(H, H_{n_i})$ converges. By a compactness argument, we can repeat this for all countably many H, giving a convergent sequence). For any convergent sequence, we can store these values in a *graph parameter* $\Phi_{(G_n)} : \mathcal{G} \to \mathbb{R}$ taking $H \mapsto \lim_{n\to\infty} p(H, G_n)$. These turn out to be very important, and tell us everything we could know about the asymptotics of such a sequence (more on that later).

Convergent sequences and their graph parameters are the main objects of study in both the topics of graph limits and flag algebras. The topic of graph limits focuses on what Razborov describes as the "semantics", ie the form of the limit object. For graphs, this turns out to be a graphon (a symmetric, measurable function on $[0,1]^2$), for hypergraphs, a hypergraphon (a subset of $[0,1]^{2^{[r]}}$) [Lov12, Chapters 7, 23.3]. Flag algebras deal instead with the "syntactics", ie with the language and manipulations that can be used to calculate quantities [Raz13a, Page 4].

A simple game Let's consider the following situation: An evil overlord has a graph G, with $|G| \geq 3$, and will only let you go if you can tell him the edge density of G. The catch is that the only information he gives you is the density of the different induced graphs on 3 vertices. That is, (denoting an edge by ρ, a triangle by C_3, a path of length two by P_3, its complement by $\bar{P}_3$, and a set of independent vertices by I_3), he wishes you to calculate $p(\rho, G)$, but gives you only $p(C_3, G)$, $p(P_3, G)$, $p(\bar{P}_3, G)$ and $p(I_3, G)$.

Thankfully the answer is pretty obvious: we have that C_3 has an edge density of 1, ie $p(\rho, C_3) = 1$, similarly, $p(\rho, P_3) = \frac{2}{3}$, $p(\rho, \bar{P}_3) = \frac{1}{3}$, and $p(\rho, I_3) = 0$. We can multiply these densities by the densities of the corresponding subgraphs to get our answer. This gives $p(\rho, G) = p(C_3, G) + \frac{2}{3}p(P_3, G) + \frac{1}{3}p(\bar{P}_3, G)$. It's worth quickly considering if the fact we're counting each edge/non-edge multiple times is problematic. A little thought confirms we're okay: each pair of vertices is contained in an equal number of induced graphs on 3 vertices, so our conclusion holds.

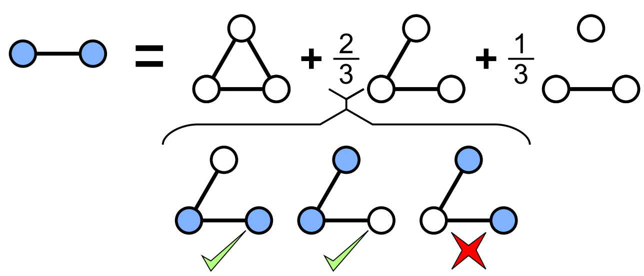


Figure 2: Demonstration of $\rho = C_3 + \frac{2}{3}P_3 + \frac{1}{3}\bar{P}_3$. The colouring is to aid understanding. Ticks represent when the induced graph on the blue vertices is isomorphic to an edge.

Thus, in sufficiently large graphs G, (well of order ≥ 3), we have that an edge, ρ, is equivalent to $C_3 + \frac{2}{3}P_3 + \frac{1}{3}\bar{P}_3$, as demonstrated in figure 2. This idea can be readily generalised: consider some graph H, with $|H| \leq l \leq |G|$. Then, by similar arguments:

$$H - \sum_{\tilde{H} \in \mathcal{G}_l} p(H, \tilde{H})\tilde{H} = 0 \tag{†}$$

It is clear that taking linear combinations of graphs is a useful concept, so we can consider the vector space $\mathbb{R}\mathcal{G}$. An element of $\mathbb{R}\mathcal{G}$ is known as a *quantum graph.* We can also extend $\Phi_{G_n} : H \mapsto$

$\lim_{n\to\infty} p(H, G_n)$ (defined at the start of this section) linearly onto $\mathbb{R}\mathcal{G}$. Then, for all $H \in \mathcal{G}$ and $l \geq |H|$, every Φ has $\Phi\left(H - \sum_{\tilde{H}\in\mathcal{G}_l} p(H, \tilde{H})\tilde{H}\right) = 0$. In other words, defining:

$$\mathcal{K} = \operatorname*{span}_{\substack{H\in\mathcal{G}\\ l\geq|H|}} \left\{ H - \sum_{\tilde{H}\in\mathcal{G}_l} p(H, \tilde{H})\tilde{H} \right\}$$

Then $\Phi(k) = 0$ for all $k \in \mathcal{K}$. The Φ are all we care about, and as this is space is in the kernel of every Φ, it makes sense to consider Φ on the quotient space, $\mathcal{A} \overset{\text{def}}{=} \mathcal{G}/\mathcal{K}$. This space captures all asymptotic density relations.

Flags We can extend this a little further, and consider labelling some vertices of our graphs (with distinct labels). This will allow us to do more intelligent constructions by, eg, allowing us to combine graphs together in a meaningful way. We will call such partially labelled graphs *flags*. To be useful, we will want similar arguments to work as in the graphs case above. A little thought yields the following conclusions:

i) To combine graphs, we will require that the labels are all from the same set. This set will be $[k]$.
ii) We require that the induced subgraph on the labelled vertices is the same for each graph. Let such a graph be σ, defined on vertex set $[k]$.
iii) We require that the labelling is always fixed in our arguments. That is, we define $p(F, G)$ to be the probability that $F \cong \boldsymbol{H}$ for $\boldsymbol{H}$ a random subgraph of G of size $|F|$ and containing all the labelled vertices (ie, all of σ). The isomorphism must also ensure the labels are in the correct place. This is known as a flag isomorphism.
iv) We sum in (†) over all σ-flags $\mathcal{F}_l^\sigma$ on l vertices, not just all graphs $\mathcal{G}_l$.

With the above sorted, we can then define $\mathcal{F}^\sigma$ and correspondingly $\mathcal{K}^\sigma$ and $\mathcal{A}^\sigma$.

It turns out that we can define a commutative multiplication operation on flags, mapping $(\mathcal{F}^\sigma)^2 \to \mathbb{R}\mathcal{F}^\sigma$ (extended by linearity onto $\mathbb{R}\mathcal{F}^\sigma$, and by extension, $\mathcal{A}^\sigma$). We create $F_1 \cdot F_2$ by taking the union of F_1 and F_2, but identifying identically labelled vertices, and taking the weighted sum of flags F' covering all possibilities for the edges between. The weights are necessary as we are considering equivalence classes (of isomorphisms): the weight of F' is proportional to the number of ways we can decompose the graph it into F_1 and F_2 (strictly, it is the probability that by randomly choosing disjoint sets $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ of *unlabelled* vertices of size $|\boldsymbol{V}_i| = |F_i| - |\sigma|$ from F', we have $F'|_{\boldsymbol{V}_i\cup\sigma} \cong F_i$ for each i). This probability is denoted $p(F_1, F_2; F')$. An explanation for why this is a useful weighting is given in remark 2 below. The construction is demonstrated in figure 3.

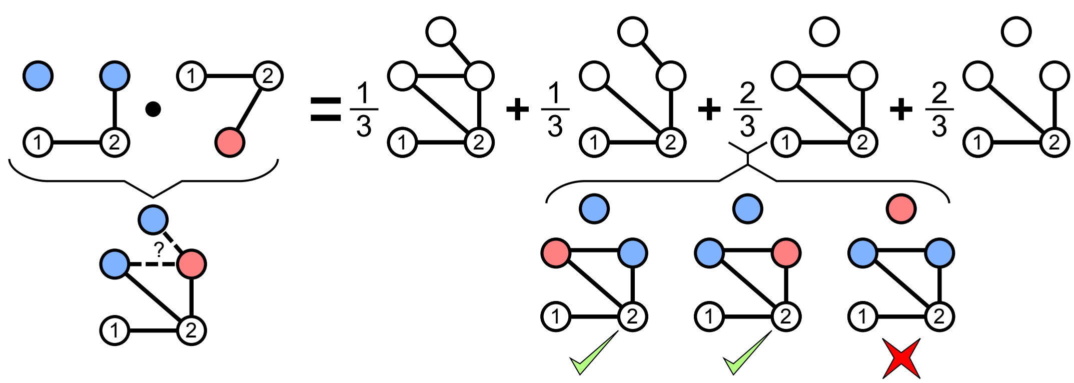


Figure 3: Demonstration of multiplication. The colouring is to aid understanding. Ticks represent where the induced graphs on blue/red and labelled vertices are flag isomorphic to the originals.

This can be extended in a natural way to an associative product, $F_1 \cdot \cdots \cdot F_k$ on $\mathcal{F}^\sigma$, and further extended to take the sum over flags of any fixed size $l \geq |\sigma| + \sum_i |F_i - \sigma|$, in such a way that it agrees with the $\mathcal{A}^\sigma$ construction. This is perhaps most clear by combining the ideas in figures 2 and 3: we sum over all flags in $F \in \mathcal{F}_l^\sigma$, with coefficients calculated by looking over all blue/red/.../colourless partitions of the unlabelled vertices of F into the correct sized classes. This multiplication turns $\mathcal{A}^\sigma$ into a commutative algebra, with identity 1^σ, defined as the unique σ-flag on $|\sigma|$ vertices.

Now instead consider an analogous probability, the only change being that we allow the $\boldsymbol{V}_i$ to overlap (ie in figure 3, this corresponds to allowing a vertex to be coloured both red and blue). It's clear that the selection of the $\boldsymbol{V}_i$ are now independent, hence $p_{allow-overlap}(F_1, F_2, \ldots, F_k; G) = p(F_1, G)p(F_2, G) \cdots p(F_k, G)$. But as $|G| \to \infty$, the probability of any overlap tends to 0, ie

$$\begin{aligned}
\lim_{|G|\to\infty} p(F_1, F_2, \ldots, F_k; G) &= \lim_{|G|\to\infty} p_{allow-overlap}(F_1, F_2, \ldots, F_k; G) \\
&= \lim_{|G|\to\infty} p(F_1, G)p(F_2, G) \cdots p(F_k, G)
\end{aligned}$$

This yields that $\Phi(F_1 \cdot F_2 \cdots F_k) = \Phi(F_1)\Phi(F_2) \cdots \Phi(F_k)$. Noting also that $\Phi(0) = 0$ and $\Phi(1^\sigma) = 1$, we have that Φ *is an algebra homomorphism* $\mathcal{A}^\sigma \to \mathbb{R}$*, which is non-negative on all flags* F.

Remark 1. Observe, by using the flag analogue of (†), that knowing the densities of all flags of size l in a graph or limit, also tells us about all densities of flags of size less than l in that graph/limit. Ie, the flags of size l form a basis over the quotient space $\mathcal{A}^\sigma_{\leq l}$. This will be very useful for later computations.

Remark 2. An asymptotic Φ captures density information about all finite objects. Analogously, knowing the densities of all σ-flags of some finite size l, tell us everything about densities of smaller flags. These reflect that Flag Algebras concern themselves with substructures and their densities. It is for this reason that weightings in flag algebras have to be given in a "looking-down" fashion. Ie, we don't ask "When multiplying F_1 by F_2, and assigning undefined edges at random, what is the probability of getting this resultant flag?" because that approaches the question in the wrong way. The more useful question is "What is the probability that this resultant flag came from a multiplication of F_1 and F_2?".

The ability to use such probabilities recursively underlies the structure of the Flag Algebra. Thus, weightings appear in many places in a number of definitions. In general, the weightings answer the question "what is the probability that a random substructure with a consistent labelling and correct size comes from the given inputs".

Overall goal The aim of flag algebras is to develop tools to find *flag relations*, $f \in \mathcal{A}^\sigma$, with $\Phi(f) \geq 0$ for all Φ. Interpreted a slightly different way, this says if we treat flags as their asymptotic densities, such fs are "asymptotically true relations". For example, the famous Goodman bound on triangle densities is equivalent to saying $f = \rho \cdot (2\rho - 1) - K_3$ satisfies the condition, and so we also write it as $K_3 \geq \rho(2\rho - 1)$. It interprets as "in the limit, the triangle density is greater than or equal to 2 times the edge density squared, minus the edge density".

Section 4 will present the formalism of the ideas in this section, and will also present some tools for helping to prove such combinations have $\Phi(f) \geq 0$ for all Φ (ie, are asymptotically true).

Remark 3. Normally our final result concerns unlabelled graphs, ie with $\sigma = 0$. During calculations, we typically consider larger σs, and then perform an unlabelling operation.

3 A brief introduction to finite model theory

Note: The definitions in this section are attributed to Marker's book [Mar02]. The examples and interpretation are my own. To be consistent with literature, I will temporarily adopt the notational conventions of model theory, and return to Razborov's notation upon leaving this section.

Flag algebras can be explained just in the context of standard 2-graphs, but they are truly defined on a much more general class of structures using first order model theory. I believe that understanding this formalism is important to be able to appreciate the applicability of the method, and to confidently apply the method to these interesting structures (including examples such as coloured graphs, hypergraphs and "X-free" graphs - of which the latter two are covered in this essay). It also allows a formalization of a flag (the labels become constants in model theory). The basic model theory required is not at all complex, and I present it (with relevant examples) over the next couple of pages.

3.1 Basic definitions

We start with a *language*, a (possibly infinite) set $\mathcal{L}$, consisting of a set of function symbols $f \in \mathcal{F}$ (each with an associated number n_f, their number of inputs, or *arity*), a set of relation (or predicate) symbols, $R \in \mathcal{R}$ (again, each with an associated number of inputs, n_R), and a set of constant symbols, $c_i \in \mathcal{C}$ (these can be thought of as labels).

Using our language, we can write *formulae*. These are (well-formed) strings built using symbols from $\mathcal{L}$, as well as the equality symbol $=$ (which, for all intents and purposes is treated as a relation), the symbols $(,), \vee, \wedge, \rightarrow, \neg, \exists, \forall$ (with their traditional meanings), and variable symbols $v_1, v_2, \ldots$.

If a variable symbol in a formula is not covered by a $\forall$ or $\exists$, it is known as a *free variable*. A formula with no free variables is known as a *closed formula* or *sentence*. An *$\mathcal{L}$-theory*, T, is a (possibly infinite) set of $\mathcal{L}$-sentences (also called its axioms).

Example 1. The theory T_{Graph}

Let my language be $\mathcal{L}_{\text{Graph}} = \{E\}$, where E is a relation taking two inputs. Conceptually, this relation marks when two vertices are joined by an edge.

Define a theory T_{Graph} as the set of sentences:

i) $\forall v_1 \; \neg E(v_1, v_1)$ *(the edge relation is irreflexive)*

ii) $\forall v_1 \forall v_1 \; (E(v_1, v_2) \rightarrow E(v_2, v_1))$ *(the edge relation is symmetric)*

Example 2. Graphs with an induced square, and $T^{\square}$, graphs with a labelled induced square

Define the theory $T_{\text{Graph with square}}$ over the language $\mathcal{L}_{\text{Graph}}$ to be the set of sentences T_{Graph}, together with the sentence $\exists x_1 \exists x_2 \exists x_3 \exists x_4 \; (\phi_1 \wedge \phi_2 \wedge \phi_3)$, where ϕ_i are formulae over the x_j (which will represent our square's corners) defined as follows:

- $\phi_1 = \bigwedge_{1 \le i < j \le 4} (\neg x_i = x_j)$ *(the x_i are distinct)*
- $\phi_2 = E(x_1, x_2) \wedge E(x_2, x_3) \wedge E(x_3, x_4) \wedge E(x_4, x_1)$ *(adjacent corners are connected)*
- $\phi_3 = \neg E(x_1, x_3) \wedge \neg E(x_2, x_4)$ *(opposite corners are not connected)*

Alternatively, if we wish to label the vertices of the square, to be able to reference them in formulae, we need to represent them in our language. We can do this by extending the language of graphs to $\mathcal{L}_4 = \{E, c_1, c_2, c_3, c_4\}$, where the c_i are constant symbols. Form ϕ_i' from ϕ_i by replacing the variables x_i with the constants c_i. Then over $\mathcal{L}_4$, we can define $T^{\square} \stackrel{\text{def}}{=} T_{\text{Graph}} \cup \{\phi_1', \phi_2', \phi_3'\}$.

Example 3. Graphs with no induced squares

Take $\mathcal{L} = \mathcal{L}_{\text{Graph}}$. Then define the theory, $T_{\text{SF-Graph}}$ to be the set of sentences T_{Graph}, together with the following sentence, formed from the ϕ_i from example 2 above:

$$\forall x_1 \forall x_2 \forall x_3 \forall x_4 \, (\phi_1 \rightarrow \neg(\phi_2 \wedge \phi_3))$$

An $\mathcal{L}$-*structure*, or *model*, $\mathcal{M}$, is given by a non-empty set, M, the *ground set* of $\mathcal{M}$, together with *interpretations* of symbols in the language. Specifically:

i) A function, $f^{\mathcal{M}} : M^{n_f} \to M$ for each $f \in \mathcal{F}$.
ii) A set, $R^{\mathcal{M}} \subset M^{n_R}$, for each $R \in \mathcal{R}$. This is the set of inputs on which the relation is true.
iii) An element, $c^{\mathcal{M}} \in M$ for each $c \in \mathcal{C}$. Ie each constant is interpreted as an element of M.

Informally, $\mathcal{M}$ satisfies a closed formula, ϕ (written $\mathcal{M} \models \phi$), if ϕ is satisfied when the formulae is interpreted in $\mathcal{M}$. That is, if we plug in the appropriate constant symbols, range the quantifiers over $m \in M$, calculate functions, find the truths of relations, and inductively combine these parts together, then we output a true result. We can then say a model $\mathcal{M}$ satisfies a theory, T, (also written $\mathcal{M} \models T$) if $\mathcal{M}$ satisfies all sentences in T.

We define $\mathcal{N}$ to be a *substructure* of $\mathcal{M}$ if it is an $\mathcal{L}$-structure with ground set $N \subset M$, and whose interpretations of the symbols in $\mathcal{L}$ are consistent with $\mathcal{M}$ on N.

Example 4. A model of $T_{\text{Graph with square}}$ and $T^{\square}$

The graph in figure 4 can be modelled as an $\mathcal{L}_{\text{Graph}}$-structure, $\mathcal{G}$, with ground set $V = \{A, B, C, D, E\}$, and $E^{\mathcal{G}} = \{\{A, B\}, \{B, A\}, \{B, C\}, \{C, B\}, \ldots$ $\ldots \{C, D\}, \{D, C\}, \{D, E\}, \{E, D\}, \{E, B\}, \{B, E\}\}$.

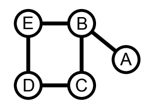


Figure 4

Clearly E is irreflexive and symmetric, so as an $\mathcal{L}_{\text{Graph}}$ structure, $\mathcal{G} \models T_{\text{Graph}}$. It is also clearly observed that $\mathcal{G} \models T_{\text{Graph with square}}$, as $x_1 = B$, $x_2 = C$, $x_3 = D$, $x_4 = E$ satisfies the sentence $\exists x_1 \exists x_2 \exists x_3 \exists x_4\ (\phi_1 \wedge \phi_2 \wedge \phi_3)$.

We can also make it into an $\mathcal{L}_4$ structure, by assigning the labels $c_1^{\mathcal{G}} = B$, $c_2^{\mathcal{G}} = C$, $c_3^{\mathcal{G}} = D$, $c_4^{\mathcal{G}} = E$. As an $\mathcal{L}_4$-structure, we have $\mathcal{G} \models T^{\square}$. This formalizes the notion "$\mathcal{G}$ *is a* $\square$*-flag*".

3.2 Finite model theory applied to flag algebras

Flag algebras are introduced with respect to a first-order theory T (over a language $\mathcal{L}$) satisfying 3 properties:

i) T must be **universal**. A *universal* first order theory T is one for which we can find an equivalent axiomatisation, T', (ie such that that $\mathcal{M} \models T \Leftrightarrow \mathcal{M} \models T'$) only having axioms written in the following form (where ϕ is quantifier-free):

$$\forall x_1 \ldots \forall x_n\ \phi$$

It can be easily observed that for such a T, if $\mathcal{M} \models T$, and $\mathcal{N}$ is a substructure of $\mathcal{M}$, then $\mathcal{N} \models T$. It turns out the converse also holds [Mar02, Theorem 2.3.9], ie a theory T is universal if and only if the substructure of any model of T is also a model of T. Razborov calls this the *hereditary property* in [Raz13a].

ii) $\mathcal{L}$ must contain **only relation symbols** (ie no functions or constants). Observe that this implies that $\forall N \subset M$, we get an induced substructure $\mathcal{N}$ on N. Universality then implies that $\mathcal{N} \models T$.

This also makes embedding and isomorphism easy to define. Let $f : M \to N$, be such that for each relation $R \in \mathcal{L}$ we have $R^{\mathcal{M}}(m_1, \ldots, m_k)$ if and only if $R^{\mathcal{N}}(f(m_1), \ldots, f(m_k))$.

If f is also a bijection, then it is an *isomorphism*, and $\mathcal{M} \cong_{\mathcal{L}} \mathcal{N}$. If f is an injection, if is called a *model embedding*. Observe that $f(\mathcal{M}) \cong_{\mathcal{L}} \mathcal{N}|_{\text{im}(f)}$.

iii) The theory is required to have an infinite model. In combination with the second property, this implies that the theory must have **models of all finite cardinality**. Again, model theory provides a converse: a compactness argument [Mar02, Prop 2.1.12] shows that if a first order theory T has models of arbitrarily large finite size, then it has an infinite model.

These can be captured exactly in an intuitive summary as follows:

Summary. Flag algebras can be used to study classes $\mathcal{C}$ of all finite structures composed of relations over a ground set V, which satisfy a set of rules restricting how the relations behave on subsets of V of bounded size (ie the rules are first-order). These rules must be such that $\mathcal{C}$ is closed under taking arbitrary (induced) substructures, and contains structures of every finite cardinality.

Examples 5. Applicable classes of structures

Below are some examples of classes of structures that are covered. In some cases, I have given an example of an implementation of the theory. In many cases, there are multiple possible formulations of such a theory:

i) All n-uniform hypergraphs (some fixed $n \in \mathbb{N}$).

ii) All hypergraphs containing edges of size in some set $S \subset \mathbb{N}$.

Implementation. Define relations E_i of arity i, for each edge-size $i \in S$, and adding the correct symmetry/irreflexivity axioms to T.

iii) All directed graphs.

iv) All graphs edge-coloured with colours in $S \subset \mathbb{N}$.

Implementation. Define relations for each edge colour, ie E_i of arity 2 for each $i \in S$, with the standard edge axioms, as well as for each pair $(i, j) \in S^2$:

i) $\forall x \forall y\ \neg(E_i(x, y) \wedge E_j(x, y))$ *(Each edge is only one colour)*

ii) $\forall x \forall y \forall z\ (E_i(x, z) \wedge E_i(y, z) \to x = y)$ *(Only one edge of each colour is connected to z)*

v) All graphs with a vertex-colouring with $\leq k$ colours.

Implementation. Take the normal edge relation, E, together with an equivalence relation, C, indicating that vertices are coloured the same:

i) $\forall x \forall y\ \neg(E(x, y) \wedge C(x, y))$ *(Connected vertices aren't coloured the same)*

ii) $\forall x\ C(x, x)$ *(Reflexitivity)*

iii) $\forall x \forall y\ (C(x, y) \to C(y, x))$ *(Symmetry)*

iv) $\forall x \forall y \forall z\ (C(x, y) \wedge C(y, z) \to C(x, z))$ *(Transitivity)*

v) $\forall x_1 \ldots \forall x_{k+1}\ \bigvee_{1 \leq i < j \leq k+1} C(x_i, x_j)$ *(≥ 2 of any $k+1$ vertices are in the same class)*

vi) All vertex-coloured graphs which either have size ≤ 10 or are 2-coloured.

vii) All graphs with no induced subgraphs isomorphic to $H \in \mathcal{H}$ for some (possibly infinite) set of graphs $\mathcal{H}$.

Remark 4. You must ensure that there are indeed arbitrarily large models. Eg, if $\mathcal{H}$ contains the empty graph on 3 vertices and the triangle, then there is no model satisfying this theory with size > 5 (as $R(3, 3) = 6$).

Implementation. As in example 3 with $T_{\text{SF-Graph}}$.

viii) All graphs with max degree $\leq k$.

Examples 6. Non-applicable classes of structures

Here are some examples that don't qualify. Unless otherwise stated, they satisfy all of the other properties (ie universality, definable in a first order theory, having arbitrarily large models).

i) All graphs with no isolated vertex. (Not universal)

ii) All graphs containing no induced triangle or set of 3 independent edges. (No large models)

iii) All graphs which have less than K non-isomorphic subgraphs of each degree. (Not first order)

iv) All graphs that are trees. (Not first order, not universal)

v) All bipartite graphs with equal vertex classes. (Not first order, not universal)

4 Flag Algebras: A formal introduction

Note: Unless otherwise stated, the material in this section is based upon or borrows heavily from material in Razborov's introductory paper, [Raz07].

4.1 Definition of flags and flag spaces

Take a first order theory, T, satisfying the conditions in section 3.2. Let models satisfying T be denoted now by roman capitals. The ground-set of a model M is denoted $V(M)$ and called its *vertices*. Define the *size* of M as $|V(M)|$, and henceforth will also denoted $|M|$. Let $\mathcal{M}_n$ now stand for the set of all finite models of T on n vertices up to isomorphism, and let $\mathcal{M} \stackrel{\text{def}}{=} \cup_{n\in\mathbb{N}} \mathcal{M}_n$.

Following our definitions in section 1.2, for $V \subset V(M)$, we define $M|_V$ to be the induced substructure on V. Similarly, define $M - V \stackrel{\text{def}}{=} M|_{V(M)\setminus V}$. We note the definition of model embedding from part (ii) of section 3.2, as an injective, structure-preserving map.

We can also define $p(M, N)$ analogously to the definition for graphs. For $M \in \mathcal{M}_l$, $N \in \mathcal{M}_L$, and $l \leq L$, define $p(M, N)$ as the probability that $M \cong N|_{\boldsymbol{V}}$ for $\boldsymbol{V} \in V^{(l)}$ chosen at random.

Formally introducing flags We define a *type* to be a model, σ of T on vertex set $[k]$ for some $k = |\sigma|$, the *size* of σ. Razborov also defines T^σ as a theory of models T containing an induced and labelled copy of σ. Formally, he defines this by introducing constants to form a language $\mathcal{L}_k$, and adding axioms defining the structure σ on these constants, analagously to examples 2 and 4 in section 3.

A *σ-flag* is a pair $F = (M, \theta)$, where M is a finite model, and θ is a model embedding $\sigma \to M$. Thus, as in section 2, F is a partially labelled graph, with labelled component $\operatorname{im}(\theta) \cong \sigma$. Again, we can define the *size* of a σ-flag to be $|F| = |V(M)|$, and define $\mathcal{F}_l^\sigma$ as the set of all flags on l vertices up to isomorphism, and $\mathcal{F}^\sigma = \cup_{l\in\mathbb{N}} \mathcal{F}_l$. Note that $\mathcal{F}^0$ (where 0 denotes the unique type on 0 vertices) is equivalent to $\mathcal{M}$. Also observe that $\mathcal{F}^\sigma_{|\sigma|}$ has only one element, the flag (σ, id). This flag is denoted 1^σ, or just 1 if it's clear from context. It will act as the identity in our algebra defined later.

Let $V \subset V(M)$. If $\operatorname{im}(\theta) \subset V$, then we can form the σ-flag $F|_V = (M|_V, \theta|_V)$. Similarly, if $V \cap \operatorname{im}(\theta) = \varnothing$ then we can form the σ-flag, $F - V = F|_{V(M)\setminus V} = (M|_{V(M)\setminus V}, \theta)$. I will refer to the model induced on the unlabelled vertices as $F - \sigma \stackrel{\text{def}}{=} M|_{V(M)\setminus \operatorname{im}(\theta)} \in \mathcal{M}$.

A *σ-flag embedding* is a T^σ model embedding (in the language $\mathcal{L}_k$). More concretely, for flags $F = (M, \theta)$ and $F' = (M', \theta')$, it's a model embedding $\alpha : V(F) \to V(G)$ which is label preserving, that is, satisfies $\theta' = \alpha\theta$. α is an isomorphism if it's also surjective (and thus bijective). A demonstration of a σ-flag embedding α is given in figure 5.

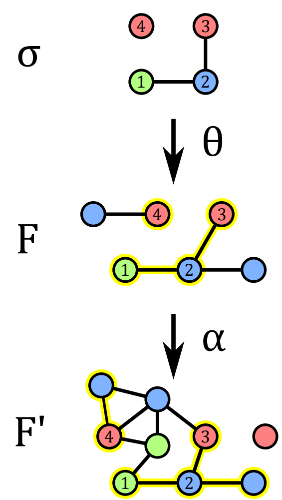


Figure 5: A demonstration of a type embedding, $\theta : \sigma \to F$, and a flag embedding $\alpha : F \to F'$, in the theory of 3-coloured graphs.

Remark 5. In case the reader is interested about the origin of the term *flag*, the following is a quote from Razborov published in his recent article for *Notices of the AMS* [Raz13b, Page 1326]:

> *The choice of the term "flag" to stand for "a partially labeled combinatorial structure in which labeled vertices span a prescribed model" is admittedly somewhat arbitrary. It is largely suggested by a visual association: a few vertices are fixed rigidly while many more are "free" and "waving" through the model we are studying. It has very little to do with other usages of this term in mathematics... incidentally, I have never seen a good explanation of what increasing sequences of linear spaces have to do with corporeal flags, either.*

4.2 Formal definition of $p(F_1, \ldots, F_n; F)$ and $p(F, F')$

Note: Before continuing, I'd recommend flicking back to the general combinatorial definitions in section 1.2, in particular, the definition of a sunflower.

Let F and F_i ($i \in [k]$) be σ-flags, such that:

$$|F| \geq |\sigma| + |F_1 - \sigma| + |F_2 - \sigma| + \cdots + |F_k - \sigma| \tag{4.2.1}$$

We can define the quantity $p(F_1, F_2, \cdots, F_n; F)$ to be the probability that a randomly chosen sunflower, $(\boldsymbol{V}_1, \boldsymbol{V}_2, \cdots, \boldsymbol{V}_k)$ over $V(F)$ with centre $\mathrm{im}(\theta)$ and petal sizes $|\boldsymbol{V}_i| = |F_i|$ satisfies:

$$F|_{\boldsymbol{V}_i} \cong F_i$$

Inequality (4.2.1) is required simply to allow such a sunflower to exist. If it doesn't hold, we define $p(F_1, F_2, \cdots, F_n; F)$ to be 0. We define $p(F_1, F)$ as an alternative notation for $p(F_1; F)$. It is easily observed that this is the probability than a random label-preserving injection, $\boldsymbol{\alpha} : F_1 \to F$ has $\mathrm{im}(\boldsymbol{\alpha}) \cong F_1$. When $\sigma = 0$, the definition of $p(F_1, F)$ reduces to the standard ones for models.

Remark 6. This is a formalization of the explanation preceding figure 3 in section 2.

Remark 7. The quantity $p(F_1, F_2, \cdots, F_n; F)$ is important, because it will soon be equated with $p(F_1 \cdot F_2 \cdots F_n; F)$ when we define the flag algebra.

The definition of $p(F_1, F_2, \cdots, F_n; F)$ has a fairly obvious property, called the chain rule, which allows us to break up the calculation into sub-calculations over flags of some size $l \leq |\mathcal{F}|$.

Lemma 1 (Chain Rule)**.** *Let F and F_i ($i \in [k]$) be σ-flags, and let $s \leq k$ and $l \leq |F|$, satisfy:*

$$|\sigma| + |F_1 - \sigma| + |F_2 - \sigma| + \cdots + |F_s - \sigma| \leq l$$
$$l + |F_{s+1} - \sigma| + |F_{s+2} - \sigma| + \cdots + |F_k - \sigma| \leq |F|$$

Then:

$$p(F_1, F_2, \cdots, F_k; F) = \sum_{\tilde{F} \in \mathcal{F}_l^\sigma} p(F_1, \cdots, F_s; \tilde{F})\, p(\tilde{F}, F_{s+1}, \cdots, F_k; F) \tag{4.2.2, the chain rule}$$

Proof.

In this proof, all sunflowers should be taken to have centre $\mathrm{im}(\theta)$ implicitly. Consider the following method of generating a uniformly random sunflower $(\boldsymbol{V}_1, \ldots, \boldsymbol{V}_k)$ on $V(F)$ with petal sizes $|\boldsymbol{V}_i| = |F_i|$.

First, generate a sunflower $(\tilde{\boldsymbol{V}}, \boldsymbol{V}_{s+1}, \ldots, \boldsymbol{V}_k)$ on $V(F)$ with petal sizes l, $|F_{s+1}|$, …, $|F_t|$. Then, generate a sunflower $(\boldsymbol{V}_1, \ldots, \boldsymbol{V}_s)$ on $\tilde{\boldsymbol{V}}$, with petal sizes $|F_1|$, …, $|F_s|$.

It is clear by symmetry that $(\boldsymbol{V}_1, \ldots, \boldsymbol{V}_k)$ is a uniformly random sunflower on $V(F)$. So the RHS of (4.2.2, the chain rule) is simply the formula for total probability, conditioned on the isomorphism class of $F|_{\tilde{\boldsymbol{V}}}$. □

Remark 8. This is most commonly used by taking $s = k$:

$$p(F_1, F_2, \cdots, F_k; F) = \sum_{\tilde{F} \in \mathcal{F}_l^\sigma} p(F_1, \cdots, F_k; \tilde{F})\, p(\tilde{F}; F) \tag{4.2.3, simple chain rule}$$

4.3 Definition of $\mathcal{A}^\sigma$ and the flag algebra

Note: Section 2 provides some intuition for this section. In particular, remark 1 considers $\mathcal{A}^\sigma_{\leq l}$ as an illustrative example, remark 2 addresses the presence of the $p(F_1, F_2; F)$ factor in the product defined below, and figure 3 demonstrates the meaning of the product pictorially.

Defining $\mathcal{K}^\sigma$ and $\mathcal{A}^\sigma$ Consider the vector space, $\mathbb{R}\mathcal{F}$, of (finite) real linear combinations of flags. We will denote such linear combinations of flags by lower case roman letters such as f. We can extend the definitions of $p(F_1, \ldots, F_k; F)$ to $p(F_1, \ldots, F_k; f)$ by linearity.

Taking arbitrary flags F_i, $\tilde{F}$ and l such that they satisfy the existence conditions $|\sigma| + \sum |F_i - \sigma| \leq |\tilde{F}| \leq l$, we can subtract the RHS from both sides of (4.2.3, simple chain rule) to get:

$$p\left(F_1, \cdots, F_k \ ; \ \tilde{F} - \sum_{F \in \mathcal{F}^\sigma_l} p(\tilde{F}, F)F\right) = 0 \tag{4.3.1}$$

But this is true for arbitrary $F_1, \ldots, F_k$, so as in the explanation in section 1, such sums of flags act as 0 as far as asymptotic densities are concerned. So we may consider the subspace,

$$\mathcal{K}^\sigma = \operatorname*{span}_{\substack{\tilde{F} \in \mathcal{F} \\ l \geq |\tilde{F}|}} \left\{ \tilde{F} - \sum_{F \in \mathcal{F}^\sigma_l} p(\tilde{F}, F)F \right\} \tag{4.3.2, the definition of $\mathcal{K}^\sigma$}$$

And from that, create the quotient space $\mathcal{A}^\sigma = \mathbb{R}\mathcal{F}^\sigma / \mathcal{K}^\sigma$, on which the flag algebra will live. We define $\mathcal{A}^\sigma_l$, as the linear subspace of $\mathcal{A}^\sigma$ generated by $\mathcal{F}^\sigma_l$.

Defining the product For some l large enough for $p(F_1, F_2; F)$ to exist (ie, $l \geq |\sigma| + |F_1 - \sigma| + |F_2 - \sigma|$), we can define a product on σ-flags $\cdot : \mathcal{F}^\sigma \times \mathcal{F}^\sigma \to A^\sigma$ as:

$$F_1 \cdot F_2 = \sum_{F \in \mathcal{F}^\sigma_l} p(F_1, F_2; F)F \mod \mathcal{K}^\sigma \tag{4.3.3, the definition of $\cdot$}$$

Note: Henceforth, when an equality is clearly in $\mathcal{A}^\sigma$, I will drop the $\mod \mathcal{K}^\sigma$ *to improve the notation.*

Remark 9. To ease calculation, l is usually taken to be equal to $|\sigma| + |F_1 - \sigma| + |F_2 - \sigma|$.

We can extend the product bilinearly to $\cdot : \mathbb{R}\mathcal{F}^\sigma \otimes \mathbb{R}\mathcal{F}^\sigma \to \mathcal{A}^\sigma$. This can be used to define a commutative, associative algebra $(\mathcal{A}^\sigma, \cdot)$, using the following lemma:

Lemma 2 (Defining the Flag Algebra)**.** *Multiplication over $\mathcal{A}^\sigma$ has the following properties:*

i) *$F_1 \cdot F_2$ is independent of the choice of l.*

ii) *For $k_1, k_2 \in \mathcal{K}^\sigma$,*

$$(F_1 + k_1) \cdot (F_2 + k_2) = F_1 \cdot F_2$$

Hence $\cdot$ induces a bilinear mapping on $\mathcal{A}^\sigma$.

iii) *For $l \geq |\sigma| + |F_1 - \sigma| + \cdots + |F_k - \sigma|$,*

$$((F_1 \cdot F_2) \cdot F_3 \cdots) \cdot F_k = \sum_{F \in \mathcal{F}^\sigma_l} p(F_1, F_2, \ldots, F_k; F)F$$

Hence $\cdot : \mathcal{A}^\sigma \otimes \mathcal{A}^\sigma \to \mathcal{A}^\sigma$ is associative and commutative.

iv) *1^σ acts as the identity of this multiplication.*

v) *If σ is non-degenerate (ie, there exist arbitrarily large finite models of σ in T), then $1^\sigma \neq 0$*, **and thus $(\mathcal{A}^\sigma, \cdot)$ is an associative, commutative algebra, with identity 1^σ.**

Proof. The proofs of are all reasonably straightforward applications of (4.2.2, the chain rule).

i) Let $l \geq \tilde{l}$ both satisfy the "large enough" condition. We can then apply the chain rule to $p(F_1, F_2; F)$:

$$\begin{aligned}
\sum_{F \in \mathcal{F}^\sigma_l} p(F_1, F_2; F)F &= \sum_{F \in \mathcal{F}^\sigma_l} \left(\sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} p(F_1, F_2; \tilde{F}) p(\tilde{F}, F) \right) F \\
&= \sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} p(F_1, F_2; \tilde{F}) \left(\sum_{F \in \mathcal{F}^\sigma_l} p(\tilde{F}, F) F \right) \\
&\equiv \sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} p(F_1, F_2; \tilde{F}) \tilde{F} \mod \mathcal{K}^\sigma
\end{aligned}$$

Where in the last line, we used (4.3.2, the definition of $\mathcal{K}^\sigma$).

ii) For $k_1, k_2 \in \mathcal{K}^\sigma$, by bilinearity,

$$(F_1 + k_1) \cdot (F_2 + k_2) = F_1 \cdot F_2 + F_1 \cdot k_2 + k_1 \cdot F_2 + k_1 \cdot k_2$$

So, if we can show that $k \cdot F \in \mathcal{K}^\sigma$ for arbitrary k in the generating span of $\mathcal{K}^\sigma$, and arbitrary $F \in \mathcal{F}^\sigma$, then by linearity and commutativity, we are done. But, for l large enough:

$$\begin{aligned}
k \cdot F &= \left(\tilde{F} - \sum_{F' \in \mathcal{F}^\sigma_{l'}} p(\tilde{F}, F') F' \right) \cdot F \\
&= \sum_{G \in \mathcal{F}^\sigma_l} \left(p(\tilde{F}, F; G) - \sum_{F' \in \mathcal{F}^\sigma_{l'}} p(\tilde{F}, F') p(F', F; G) \right) G \\
&= \sum_{G \in \mathcal{F}^\sigma_l} \left(p(\tilde{F}, F; G) - p(\tilde{F}, F; G) \right) G \\
&= 0
\end{aligned}$$

Where, the deduction from line 2 to line 3 uses the chain rule.

iii) For the l_k appropriately sized, we simply apply the chain rule repeatedly:

$$\begin{aligned}
((F_1 \cdot F_2) \cdot F_3 \cdots) \cdot F_k &= \sum_{G_k \in \mathcal{F}^\sigma_{l_k}} \cdots \sum_{G_3 \in \mathcal{F}^\sigma_{l_3}} \sum_{G_2 \in \mathcal{F}^\sigma_{l_2}} p(F_1, F_2; G_2) p(G_2, F_3; G_3) \cdots p(G_{k-1}, F_k, G_k) G_k \\
&= \sum_{G_k \in \mathcal{F}^\sigma_{l_k}} \cdots \sum_{G_3 \in \mathcal{F}^\sigma_{l_3}} p(F_1, F_2, F_3; G_3) p(G_3, F_4; G_4) \cdots p(G_{k-1}, F_k, G_k) G_k \\
&\;\;\vdots \\
&= \sum_{G_k \in \mathcal{F}^\sigma_{l_k}} p(F_1, F_2, \cdots, F_k; G_k) G_k
\end{aligned}$$

iv) To show that 1^σ acts as the identity, consider $F \in \mathcal{F}^\sigma$.

$$F \cdot 1^\sigma = \sum_{\tilde{F} \in \mathcal{F}^\sigma_{|F|}} p(F, 1^\sigma; \tilde{F}) \tilde{F} = \sum_{\tilde{F} \in \mathcal{F}^\sigma_{|F|}} p(F, \tilde{F}) \tilde{F} = F$$

because, for $\tilde{F} \in \mathcal{F}^\sigma_{|F|}$, we have $p(F, \tilde{F}) = \delta_{F, \tilde{F}}$. Commutativity and linearity gives $f \cdot 1^\sigma = 1^\sigma \cdot f = f$ for arbitrary $f \in \mathcal{A}^\sigma$.

v) We must show that $1^\sigma \notin \mathcal{K}^\sigma$ (ie, that $\mathcal{A}^\sigma$ is non-trivial). Well, take arbitrary $k \in \mathcal{K}^\sigma$,

$$k = \sum_{i=1}^{n} a_i \left(\tilde{F}_i - \sum_{F \in \mathcal{F}^\sigma_{l_i}} p(\tilde{F}_i, F) F \right)$$

As σ is a non-degenerate type over T, take a σ-flag G of size at least $\max_i(l_i)$. Define $p'(f, g)$ to be the linear extension of $p(F, g)$ to its first argument. Then:

$$\begin{aligned} p'(k, G) &= \sum_{i=1}^{n} a_i \left(\tilde{p}(F_i, G) - \sum_{F \in \mathcal{F}^\sigma_{l_i}} p(\tilde{F}_i, F) p(F, G) \right) \\ &= 0 \end{aligned}$$

By the chain rule. But $p'(1^\sigma, G) \neq 0$ as G is a σ-flag. So $1^\sigma \neq k$. But $k \in \mathcal{K}^\sigma$ arbitrary, hence $1^\sigma \notin \mathcal{K}^\sigma$.

□

Henceforth, when considering $\mathcal{A}^\sigma$, I will implicitly assume σ is a non-degenerate type over T.

Remark 10. For most commonly used theories T, all σ are non-degenerate, and $\mathcal{A}^\sigma$ is free [Raz07, Theorem 2.7]. The theorem requires that T has the amalgamation property, (that is if $\mathcal{M}_1$, $\mathcal{M}_2 \models T$, then $\mathcal{M}' \stackrel{\text{def}}{=} \mathcal{M}_1 \sqcup \mathcal{M}_2 \models T$. We define $\mathcal{M}'$, known as the *amalgam* of $\mathcal{M}_1$ and $\mathcal{M}_2$, over the base set $M_1 \sqcup M_2$ (for $\sqcup$ the disjoint union), with predicates $P^{\mathcal{M}'} = P^{\mathcal{M}_1} \cup P^{\mathcal{M}_2}$. That is, there is no link between $\mathcal{M}_1$ and $\mathcal{M}_2$ in the amalgam).

Being free means that $\mathcal{A}^\sigma$ is isomorphic to the algebra of polynomials with countably many variables. Such a freely generating set of variables turns out to be $\tilde{\mathcal{F}}^\sigma$, the set of all connected flags except 1^σ and F_0, for some fixed $F_0 \in \mathcal{F}^\sigma_{|\sigma|+1}$. We define a flag to be *connected* if we can't split its base set into $M_1 \sqcup M_2$ where each predicate P of arity r is a subset of ${M_1}^{(r)} \cup {M_2}^{(r)}$ (ie doesn't link between the two disconnected halves).

4.4 The averaging operator $[\![\cdot]\!]_\sigma$

The most widely used syntactic tool for flag algebras is the averaging operator, which unlabels some of the vertices of σ. That is, if $|\sigma| = k$, we end up with α-flags, where $\alpha \cong \sigma|_V$ for some $V \subset [k]$.

Formally, let σ be a type of size k. Let $\eta : [k'] \to [k]$ be an injective mapping. Let $\sigma|_\eta$ be the type on $[k']$ induced by η. (ie, $\sigma|_\eta$ is $\sigma|_{\mathrm{im}(\eta)}$, but with its vertex labels replaced by their pre-images under η). For a σ-flag $F = (M, \theta)$, we can define a $\sigma|_\eta$-flag $F|_\eta \stackrel{\text{def}}{=} (M, \theta\eta)$. (That is, for each labelled vertex in F, if it's in the image of η, we relabel it with its pre-image under η, else we unlabel it.)

In the spirit of remark 2 on weightings, we will need to weight the averaging operation by the probability that the larger object (F) is an extension of the smaller object ($F|_\eta$). That is, the probability that assigning the labels $[k] \setminus \mathrm{im}(\eta)$ at random onto the unlabelled vertices of $F|_{\mathrm{im}(\eta)}$ gives us F. Formally, we define this normalizing factor $q_{\sigma,\eta}(F) \in [0, 1]$ as follows: Take a random injective mapping (ie labelling) $\boldsymbol{\theta} : [k] \to V(M)$ which satisfies $\boldsymbol{\theta}\eta = \theta\eta$ (ie, is consistent on the already labelled vertices). Then $q_{\sigma,\eta}(F)$ is the probability that $(M, \boldsymbol{\theta})$ is a σ-flag, and is flag isomorphic to F.

So, given the above, we define:

$$[\![F]\!]_{\sigma,\eta} \stackrel{\text{def}}{=} q_{\sigma,\eta}(F)\ F|_\eta$$

Remark 11. This is often used to derive an end result about unlabelled models. For this, we take η as the empty function, ie $\sigma|_\eta = 0$. In this case, we simplify the notation to $[\![F]\!]_\sigma$ and $q_\sigma(F)$.

The averaging operation can be extended linearly to a map $[\![\cdot]\!]_{\sigma,\eta} : \mathbb{R}\mathcal{F}^\sigma \to \mathbb{R}\mathcal{F}^{\sigma|_\eta}$. We show in the following lemma that it gives an induced map $\mathcal{A}^\sigma \to \mathcal{A}^{\sigma|_\eta}$.

Lemma 3. *$[\![\cdot]\!]_{\sigma,\eta}$ defines an induced mapping $\mathcal{A}^\sigma \to \mathcal{A}^{\sigma|_\eta}$. That is, $[\![\cdot]\!]_{\sigma,\eta}$ takes $\mathcal{K}^\sigma$ to $\mathcal{K}^{\sigma|_\eta}$.*

Proof. By linearity, we can reduce this to proving it for arbitrary $k = \tilde{F} - \sum_{F\in\mathcal{F}_l^\sigma} p(\tilde{F}, F)F$ in the set generating $\mathcal{K}^\sigma$ from (4.3.2, the definition of $\mathcal{K}^\sigma$). Thus,

$$\begin{aligned}[\![k]\!]_{\sigma,\eta} &= [\![\tilde{F}]\!]_{\sigma,\eta} - \sum_{F\in\mathcal{F}_l^\sigma} p(\tilde{F}, F)[\![F]\!]_{\sigma,\eta} \\ &= q_{\sigma,\eta}(\tilde{F})\ \tilde{F}|_\eta - \sum_{F\in\mathcal{F}_l^\sigma} p(\tilde{F}, F)q_{\sigma,\eta}(F)\ F|_\eta \end{aligned} \tag{4.4.1}$$

Claim. *Let $|\sigma| = k$, $\eta : [k'] \to [k]$, and F' be a $\sigma|_\eta$-flag of size l. Let $\tilde{F}$ be a σ-flag of size $\tilde{l} \le l$. (That is, F' is a larger flag than $\tilde{F}$, but has fewer labels.) Then*

$$\sum_{\substack{F\in\mathcal{F}_l^\sigma \\ F|_\eta = F'}} p(\tilde{F}, F)q_{\sigma,\eta}(F) = p(\tilde{F}|_\eta, F')q_{\sigma,\eta}(\tilde{F})$$

Proof of Claim. Take a random sunflower $\{\boldsymbol{v}_{k'+1}\}, \{\boldsymbol{v}_{k'+2}\}, \dots, \{\boldsymbol{v}_k\}, \boldsymbol{U}$ with empty centre, over the unlabelled vertices of $F' = (M, \theta')$ with petal sizes 1, 1, $\cdots$, 1, $\tilde{l} - k$. These will form an extension of the labelling of F', and a choice of some other unlabelled vertices, on which to form a new σ-flag $\boldsymbol{G} = (M|_{\boldsymbol{V}}, \boldsymbol{\theta})$. (Where we denote $\boldsymbol{V} = \mathrm{im}(\boldsymbol{\theta}) \cup \boldsymbol{U}$ where $\boldsymbol{\theta} : [k] \to M$ is the extension of the labelling $\theta' : [k'] \to M$ defined by taking $i \mapsto \boldsymbol{v}_i$.) The equation in the claim expresses the probability that $\boldsymbol{G}$ is a σ-flag, and that $\boldsymbol{G} \cong \tilde{F}$.

The LHS conditions over $\boldsymbol{\theta}$, through forms of $\boldsymbol{F} = (M, \boldsymbol{\theta})$. That is, it multiplies the probability that $\boldsymbol{F}|_{\boldsymbol{V}} = \boldsymbol{G} \cong \tilde{F}$ by the conditioning probability that an extension $\boldsymbol{\theta}$ of θ' gives $\boldsymbol{F}$. The RHS breaks down the calculation by multiplying the probability that $(M|_{\mathrm{im}(\theta')\cup\boldsymbol{U}}, \theta') \cong \tilde{F}|_\eta$, with the probability that the extension $\boldsymbol{\theta}$ of the labelling θ' takes $\tilde{F}|_\eta$ to $\tilde{F}$. ■

Multiplying the claim by the flag $F' = F|_\eta$, then summing over all $F' \in \mathcal{F}_l^{\sigma|_\eta}$, yields:

$$\sum_{F'\in\mathcal{F}_l^{\sigma|_\eta}} \sum_{\substack{F\in\mathcal{F}_l^\sigma \\ F|_\eta = F'}} p(\tilde{F}, F)q_{\sigma,\eta}(F)\ F|_\eta = q_{\sigma,\eta}(\tilde{F}) \sum_{F'\in\mathcal{F}_l^{\sigma|_\eta}} p(\tilde{F}|_\eta, F')F'$$

Noting that the double sum reduces to a sum over $F \in \mathcal{F}_l^\sigma$, we can substitute into (4.4.1) to give:

$$[\![k]\!]_{\sigma,\eta} = q_{\sigma,\eta}(\tilde{F})\left(\tilde{F}|_\eta - \sum_{F'\in\mathcal{F}_l^{\sigma|_\eta}} p(\tilde{F}|_\eta, F')F'\right)$$

This is a multiple of an element from the set generating $\mathcal{K}^{\sigma|_\eta}$. □

Remark 12. The proof of the claim made important use of the form of the normalizing factor, $q_{\sigma,\eta}(F)$. The normalizing factor is indeed required for the statement of lemma 3 to hold.

Remark 13. The averaging operation satisfies a chain rule, in the sense that:

$$[\![f]\!]_{\sigma,\eta\eta'} = \left[\!\!\left[[\![f]\!]_{\sigma,\eta}\right]\!\!\right]_{\sigma|_\eta,\eta'}$$

This is clear by the observation that $q_{\sigma,\eta\eta'}(F) = q_{\sigma,\eta}(F)q_{\sigma|\eta,\eta'}(F)$.

Remark 14. $[\![\cdot]\!]$ appears to be common notation in other graph algebras [Lov12, Sec 6.1.3], such as the glue algebra [Lov12, Sec 6.1.1]. Care should be taken however, as in these cases, the operation usually refers to a straight unlabel (ie $[\![F]\!]$ corresponds to $F|_\sigma$ in our notation), and thus differs from our definition by the absence of the normalizing factor $q_\sigma(F)$.

4.5 Interpretations and upwards operators $\pi^{\sigma,\eta}$

Note: The upwards operator will be used later in the essay, but the abstration to an interpretation is noted briefly for completeness, and is a hopefully slightly more natural explanation/introduction than in [Raz07]

Interpretations [Raz07] also introduces a formal way of moving from σ_1-flag algebras over a theory T_1 in language $\mathcal{L}_1$ to a σ_2-flag algebras in another (perhaps different) theory T_2, perhaps even over another language $\mathcal{L}_2$. (For example, reducing coloured graphs to normal graphs, or simply extending σ in the same theory).

Essentially, if you can think of a natural way of representing one theory inside another, then you can define an "interpretation" $I : T_1^{\sigma_1} \rightsquigarrow T_2^{\sigma_2}$ in model theory which applies this representation. Strictly, I maps each predicate in $\mathcal{L}_1$ to some formula over predicates in $\mathcal{L}_2$, such that $I(\varphi)$ holds in T_2 for each axiom in $\varphi \in T_1$; and I takes each distinct constant in $\mathcal{L}_1^{\sigma_1}$ (a labelled σ_1 vertex) to a distinct constant in $\mathcal{L}_2^{\sigma_2}$ (a labelled σ_2 vertex). (For example, the theory of normal graphs can be represented in the theory of 3-edge-coloured graphs with $I(E(x,y)) = E_{\text{red}}(x,y) \vee E_{\text{green}}(x,y) \vee E_{\text{blue}}(x,y)$). We can then define $I^*(F_2)$ which interprets a σ_2-flag back into T_1 (there is always a natural way to do this), and completely removes the labelled vertices of σ_2 not in the injection from σ_1. We can use this to form the map $\pi^{(I)}$ as follows:

$$\begin{aligned} \pi^{(I)} \,:\, & \mathcal{A}^{\sigma_1}[T_1] \to \mathcal{A}^{\sigma_2}[T_2] \\ & F_1 \mapsto \sum \left\{ F_2 \in \mathcal{F}^{\sigma_2}_{|F_1|+|\sigma_2|-|\sigma_1|} \,:\, I^*(F_2) \cong F_1 \right\} \end{aligned}$$

It turns out that $\pi^{(I)}$ is an algebra homomorphism. The proof of this isn't particularly long or hard, but is similar to previous double-counting arguments, and for the sake of relevance and brevity, has been left out. For reference, see [Raz07, Theorem 2.6, Pages 15-17].

Note: Razborov defines a more general type of interpretation (U, I) with U a formula in 1 free variable telling us whether a vertex in a model of T_2 is allowed to be in a representation. But this adds little for most uses, and complicates things a lot.

Upwards Operator In particular, we can take $T_1 = T_2$, and I the identity map; but with $\sigma_2 = \sigma$ extending $\sigma_1 = \sigma|_\eta$. From this, we get a homomorphism $\pi^{\sigma,\eta} : \mathcal{A}^{\sigma|_\eta} \to \mathcal{A}^{\sigma}$, which is named the *upward operator*. $\pi^{\sigma,\eta}(F)$ takes a flag $F \in \mathcal{A}^{\sigma|_\eta}$ and gives the sum of all flags formed by adding $|\sigma| - |\text{im}(\eta)|$ new vertices to F, and labelling them such that we get a σ flag. That is, for $F = (M, \theta) \in \mathcal{A}^{\sigma|_\eta}$,

$$\pi^{\sigma,\eta}(F) = \sum \{F' \in \mathcal{A}^{\sigma} \text{ such that } F'|_\eta - (\text{im}(\theta) \setminus \text{im}(\theta\eta)) \cong F\}$$

Remark 15. The definition is the first to use a straight sum with unit weighting factors. Whilst atypical, this is actually consistent with the principle of remark 2 on weightings. Because we are adding labelled vertices, the only possible induced substructure with consistent labelling is F, so the probability of such a substructure being F is 1.

The upwards operator acts as the opposite of the averaging map $[\![\cdot]\!]_{\sigma,\eta}$, in the sense that, for any $f \in \mathcal{A}^{\sigma|_\eta}$, $g \in \mathcal{A}^{\sigma}$:

$$[\![\pi^{\sigma,\eta}(f) \cdot g]\!]_{\sigma,\eta} = f \cdot [\![g]\!]_{\sigma,\eta}$$

Again, the proof is very similar in structure to the previous double-counting arguments, and for the sake of brevity/relevance has been left out. For reference, see [Raz07, Theorem 2.8, Pages 19-20].

4.6 Homomorphisms from $\mathcal{A}^\sigma$ to $\mathbb{R}$ and densities in convergent sequences

$\Phi_{(G_n)}$ was introduced as the type of object we wish to study in section 2. It was defined as a function $\mathcal{G} \to \mathbb{R}$, giving $\Phi(H) = \lim_{n\to\infty} p(H, G_n)$ for a convergent sequence G_n. It was noted there, that extending Φ linearly to $\mathbb{R}\mathcal{G}$, we have that $\Phi(k) = 0$ for all $k \in \mathcal{K}$. It was then observed at the end of the section that Φ is a homomorphism $\mathcal{A}^\sigma \to \mathbb{R}$, which is positive on all σ-flags.

It turns out the converse is also true, and that each such homomorphism is represented by the densities in some convergent sequence. This equivalence is what enables flag manipulations to tell us about asymptotic density. This section will formalize these ideas rigorously.

Homomorphism Definitions For σ a non-degenerate type, $\mathcal{A}^\sigma$ is an algebra by lemma 2. Let the set of algebra homomorphisms $\mathcal{A}^\sigma \to \mathbb{R}$ be denoted by $\mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$. These will be denoted by lower case greek letters. By the definition of an algebra homomorphism, $\phi \in \mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ satisfies:

i) $\phi(0) = 0$
ii) $\phi(1^\sigma) = 1$
iii) $\phi(f + g) = \phi(f) + \phi(g)$
iv) $\phi(f \cdot g) = \phi(f)\phi(g)$

We define the subset $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R}) \subset \mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ by:

$$\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R}) = \{\phi \in \mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R}) \text{ such that } \phi(F) \geq 0 \text{ for all flags } F \in \mathcal{F}^\sigma\}$$

Observe that, invoking (4.3.2, the definition of $\mathcal{K}^\sigma$):

$$1 = \phi(1^\sigma) = \sum_{F \in \mathcal{F}_l^\sigma} p(1^\sigma, F)\phi(F) = \sum_{F \in \mathcal{F}_l^\sigma} \phi(F)$$

In particular, if $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, then $\phi(F) \in [0, 1]$ for all $F \in \mathcal{F}^\sigma$.

Remark 16. As noted in remark 10, often $\mathcal{A}^\sigma$ is free, so $\mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ has a very straightforward description. It is the added constraints of $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ that grant the necessary complexity.

Sequence Definitions For any σ-flag F, we can define the map $p^F \stackrel{\text{def}}{=} p(\cdot, F) : \mathcal{F}^\sigma \to [0, 1]$ giving the densities of all flags inside F. An infinite sequence of flags $F_n \in \mathcal{F}^\sigma_{l_n}$ is *increasing* if l_n is a strictly increasing sequence. An increasing sequence (F_n) of σ-flags is *convergent* if for all $H \in \mathcal{F}^\sigma$, $\lim_{n\to\infty} p^{F_n}(H)$ exists.

We denote the function space $\mathcal{F}^\sigma \to [0, 1]$ as $[0, 1]^{\mathcal{F}^\sigma}$, and equip it with the pointwise convergence (ie product) topology. By Tychonoff's Theorem, $[0, 1]^{\mathcal{F}^\sigma}$ is compact. Observe that it is metrizable: index $\mathcal{F}^\sigma$ with $\mathbb{N}$, then can take metric $d(p, q) = \sum_{i\in\mathbb{N}} 2^{-i} |p(F_i) - q(F_i)|$. As it is metrizable, it is also sequentially compact. In this notation, a sequence F_n is convergent if and only if p^{F_n} converges in $[0, 1]^{\mathcal{F}^\sigma}$.

Remark 17. By sequential compactness of $[0, 1]^{\mathcal{F}^\sigma}$, it is clear that any increasing sequence of flags has a subsequence that is convergent.

We consider the space $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})|_{\mathcal{F}^\sigma} = \{\phi|_{\mathcal{F}^\sigma} : \phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})\} \subset [0, 1]^{\mathcal{F}^\sigma}$. Observe that the value of a homomorphism on flags defines it uniquely by linearity. $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})|_{\mathcal{F}^\sigma}$ can be defined by as a subset of $[0, 1]^{\mathcal{F}^\sigma}$ by finite polynomial equalities in the values of ϕ on its co-ordinates, $F \in \mathcal{F}^\sigma$. Hence it is a closed subset of $[0, 1]^{\mathcal{F}^\sigma}$, and as such, it is also compact and metrizable. (Specifically, it is defined by equalities of the form $\phi(\tilde{F}) = \sum_{F\in\mathcal{F}_l^\sigma} p(\tilde{F}, F)\phi(F)$ and $\phi(F_1)\phi(F_2) = \sum_{F\in\mathcal{F}_l} p(F_1, F_2; F)\phi(F)$). The remainder of the section is devoted to proving the equivalence of convergent sequences and positive homomorphisms. We start by proving the following obvious lemma:

Lemma 4. *For $|\sigma| = k$, $F_1 \in \mathcal{F}^\sigma_{l_1}$, $F_2 \in \mathcal{F}^\sigma_{l_2}$, $F = (M, \theta) \in \mathcal{F}^\sigma_l$. Take l_1, l_2 fixed, and let $l \geq l_1 + l_2 - k$ tend to ∞ then:*

$$|p(F_1, F_2; F) - p(F_1, F)p(F_2, F)| = O(l^{-1})$$

Proof of Lemma. Choose independent random sunflowers $\boldsymbol{V}_1$ over $V(M)$ with centre $\operatorname{im}(\theta)$ and petal size l_1, and $\boldsymbol{V}_2$ over $V(M)$ with centre $\operatorname{im}(\theta)$ and petal size l_2. Let X denote the event that $F|_{\boldsymbol{V}_1} \cong F_1$ and $F|_{\boldsymbol{V}_2} \cong F_2$. Observe $\operatorname{im}(\theta) \subset \boldsymbol{V}_1 \cap \boldsymbol{V}_2$; let Y denote the event $\boldsymbol{V}_1 \cap \boldsymbol{V}_2 = \operatorname{im}(\theta)$. Then:

$$\begin{aligned} p(F_1, F_2; F) - p(F_1, F)p(F_2, F) &= \mathbb{P}[X \mid Y] - \mathbb{P}[X] \\ &= \frac{\mathbb{P}[X] - \mathbb{P}[X \mid Y^c]\,\mathbb{P}[Y^c]}{\mathbb{P}[Y]} - \frac{\mathbb{P}[X]\,\mathbb{P}[Y]}{\mathbb{P}[Y]} \\ |p(F_1, F_2; F) - p(F_1, F)p(F_2, F)| &= \frac{\mathbb{P}[Y^c]}{\mathbb{P}[Y]}\Big|\mathbb{P}[X] - \mathbb{P}[X \mid Y^c]\Big| \leq 2\frac{\mathbb{P}[Y^c]}{\mathbb{P}[Y]} \end{aligned}$$

Now observe that $\mathbb{P}[Y] = \mathbb{P}[\text{Each choice of vertex for } \boldsymbol{V}_2 \text{ misses } \boldsymbol{V}_1] = \sum_{i=0}^{l_2-k-1} \frac{l-k-l_1-i}{l-k-i} = 1 - \sum_{i=0}^{l_2-k-1} \frac{l_1}{l-k-i} \geq 1 - \frac{l_1 l_2}{l-k-l_2}$. So $\mathbb{P}[Y^c] = O(l^{-1})$. And as $x/(1-x) = x + O(x^2)$ as $x \to 0$, we have:

$$|p(F_1, F_2; F) - p(F_1, F)p(F_2, F)| \leq 2\mathbb{P}[Y^c] + O(\mathbb{P}[Y^c]^2) = O(l^{-1}) \qquad \square$$

Theorem 5. *For each convergent sequence (F_n) of flags, $\lim_{n\to\infty} p^{F_n}$ lies in $\operatorname{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})|_{\mathcal{F}^\sigma}$.*

Proof. Extend $\lim_{n\to\infty} p^{F_n} \overset{\text{def}}{=} \Phi_{\mathcal{F}^\sigma} \in [0,1]^{\mathcal{F}^\sigma}$ linearly to $\Phi_{\mathbb{R}\mathcal{F}^\sigma} : \mathbb{R}\mathcal{F}^\sigma \to \mathbb{R}$. By the chain rule, for $|F_n| \geq l$, $p^{F_n}(\tilde{F}) = \sum_{F \in \mathcal{F}^\sigma_l} p(\tilde{F}, F)p^{F_n}(F)$, so taking the limit gives $\Phi_{\mathcal{F}^\sigma}(\tilde{F}) - \sum_{F \in \mathcal{F}^\sigma_l} p(\tilde{F}, F)\Phi_{\mathcal{F}^\sigma}(F) = 0$, thus $\Phi_{\mathbb{R}\mathcal{F}^\sigma}(k) = 0$ for $k \in \mathcal{K}^\sigma$ by linearity. Hence $\Phi_{\mathbb{R}\mathcal{F}^\sigma}(f + k) = \Phi_{\mathbb{R}\mathcal{F}^\sigma}(f)$. So $\Phi_{\mathbb{R}\mathcal{F}^\sigma}$ can be induced on $\mathcal{A}^\sigma$, which I will denote $\Phi : \mathcal{A}^\sigma \to \mathbb{R}$. I now claim $\Phi \in \operatorname{Hom}(\mathcal{A}^\sigma, \mathbb{R})$:

i) $\Phi(0) = \Phi_{\mathbb{R}\mathcal{F}^\sigma}(0) = 0$.
ii) For any σ-flag F_n, $p^{F_n}(1^\sigma) = p(1^\sigma, F_n) = 1$. So $\Phi(1^\sigma) = 1$.
iii) By linearity, $\Phi(f + g) = \Phi(f) + \Phi(g)$.
iv) We require $\Phi(f)\Phi(g) = \Phi(f \cdot g)$. As f, g are finite weighted sums of flags, we can expand $f \cdot g$ into a sum of $a_{ij} F_i \cdot G_j$. Applying lemma 4 for each such, we have that $p^{H_n}(f \cdot g) = p^{H_n}(f) \cdot p^{H_n}(g) + O(|H_n|^{-1})$. Taking the limit gives $\Phi(f \cdot g) = \Phi(f)\Phi(g)$.

Furthermore, $\Phi(F) = \lim_{n\to\infty} p^{F_n}(F) \in [0,1]$. So $\Phi \in \operatorname{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. Hence, $\lim_{n\to\infty} p^{F_n} = \Phi_{\mathcal{F}^\sigma} = \Phi|_{\mathcal{F}^\sigma} \in \operatorname{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})|_{\mathcal{F}^\sigma}$ as required. $\square$

Theorem 6. *The converse to theorem 5 is also true. That is, each $\phi \in \operatorname{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ has $\phi|_{\mathcal{F}^\sigma} = \lim_{n\to\infty} p^{F_n}$ for some convergent sequence (F_n) of σ-flags.*

Proof. For each $n \geq |\sigma|$, pick $\boldsymbol{F}_n \in \mathcal{F}^\sigma_n$ at random with probability $\mathbb{P}[\boldsymbol{F}_n = F] = \phi(F)$. Let $\boldsymbol{G}_n = \boldsymbol{F}_{n^2}$. Fix $F \in \mathcal{F}^\sigma$. Then, as $\phi(k) = 0$ for $k \in \mathcal{K}^\sigma$:

$$\mathbb{E}\left[p^{\boldsymbol{G}_n}(F)\right] = \mathbb{E}\left[p(F, \boldsymbol{G}_n)\right] = \sum_{G \in \mathcal{F}_{n^2}} p(F, G)\phi(G) = \phi(F)$$

Also,

$$\begin{aligned} \operatorname{Var}\left[p^{\boldsymbol{G}_n}(F)\right] &= \mathbb{E}\left[p(F, \boldsymbol{G}_n)^2\right] - \phi(F)^2 \\ &= \sum_{G \in \mathcal{F}_{n^2}} p(F, G)^2\phi(G) \;-\; \phi(F)^2 \\ &= \sum_{G \in \mathcal{F}_{n^2}} \left(p(F, F; G) + O(n^{-2})\right)\phi(G) \;-\; \phi(F)^2 && \textit{(by lemma 4 above)} \\ &= \phi(F \cdot F) - \phi(F)^2 + O(n^{-2}) = O(n^{-2}) && \textit{(by linearity of } \phi \textit{ and def of product)} \end{aligned}$$

So, for each F and any $\varepsilon > 0$, let $X_{F,\varepsilon,n}$ be the event that $|p^{\boldsymbol{G}_n}(F) - \phi(F)| > \varepsilon$, then Chebyshev gives:

$$\mathbb{P}\left[X_{F,\varepsilon,n}\right] = \mathbb{P}\left[\left|p^{\boldsymbol{G}_n}(F) - \mathbb{E}\left[p^{\boldsymbol{G}_n}(F)\right]\right| > \varepsilon\right] \leq \frac{\mathrm{Var}\left[p^{\boldsymbol{G}_n}(F)\right]}{\varepsilon^2} = \frac{O(n^{-2})}{\varepsilon^2} \implies \sum_n \mathbb{P}\left[X_{F,\varepsilon,n}\right] < \infty$$

So, letting $Y_{F,\varepsilon}$ be the event that $\limsup_{n\to\infty} |p^{\boldsymbol{G}_n}(F) - \phi(F)| > \varepsilon$:

$$\mathbb{P}\left[Y_{F,\varepsilon}\right] = \mathbb{P}\left[X_{F,\varepsilon,n} \text{ infinitely often}\right] = \mathbb{P}\left[\limsup_{n\to\infty} X_{F,\varepsilon,n}\right] = 0$$

Where Borel-Cantelli gives the last equality. So,

$$\begin{aligned}\mathbb{P}\left[p^{\boldsymbol{G}_n} \text{ doesn't converge pointwise to } \phi\right] &= \mathbb{P}\left[p^{\boldsymbol{G}_n}(F) \text{ doesn't converge to } \phi(F) \text{ for some } F\right] \\ &= \mathbb{P}\left[\exists\, k, F \text{ such that } Y_{F,k^{-1}}\right] \leq \sum_{\substack{F\in\mathcal{F}^\sigma \\ k\in\mathbb{N}}} \mathbb{P}\left[Y_{F,k^{-1}}\right] = 0\end{aligned}$$

Hence $\mathbb{P}\left[\lim_{n\to\infty} p^{\boldsymbol{G}_n} = \phi\right] = 1 > 0$, so there must exist a (G_n) with $\phi = \lim_{n\to\infty} p^{G_n}$. □

4.7 The ordering on $\mathcal{A}^\sigma$ and asymptotically true relations

Note: The proof of theorem 7 is my own, as it was stated without proof in [Raz07].

We define a pre-ordering on $\mathcal{A}^\sigma$ as follows:

$$f \geq_\sigma g \iff \phi(f) \geq \phi(g) \text{ for all } \phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$$

This should already seem like a useful definition: if we can show with flag algebra/homomorphism arguments that $f \geq_\sigma g$, then we know the equivalent inequality (replacing flags with their asymptotic densities) holds for *any* convergent sequence. Even better, the following theorem and corollary extend this equivalence from convergent sequences to arbitrary asymptotics.

Remark 18. To reflect their usage, Razborov often calls the $f \in \mathcal{A}^\sigma$ *(flag) relations* (particularly in his more recent papers such as [Raz13a], [Raz13b]).

Remark 19. At least in T_{Graph}, it is known that $\leq_\sigma$ is actually a partial order. That is, it also satisfies antisymmetry, meaning that if $\phi(f) = 0$ for all $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, then $f = 0$. In other words, $\mathcal{K}^\sigma$ captures all forms of asymptotic equivalence.

Theorem 7 (Corollary to theorems 5 and 6)**.**
Let f be a real polynomial in h variables, and let $F_1, \cdots, F_h \in \mathcal{F}^\sigma$. Then:

$$f(F_1, \cdots, F_h) \geq_\sigma 0 \iff \liminf_{n\to\infty} \min_{F\in\mathcal{F}_n^\sigma} f(p^F(F_1), \cdots, p^F(F_h)) \geq 0$$

Proof of $\implies$ direction of theorem 7. We prove the contrapositive. So take a sequence (G_n) satisfying:

$$\liminf_{n\to\infty} \min_{G_n\in\mathcal{F}_n^\sigma} f(p^{G_n}(F_1), \cdots, p^{G_n}(F_h)) = C < 0$$

Then, $\min_{G_n\in\mathcal{F}_n} f(p^{G_n}(F_1), \cdots, p^{G_n}(F_h)) < \frac{1}{2}C$ occurs for infinitely many n, say in some sub-sequence (G_{n_i}). From this, by remark 17, we can extract a convergent subsequence, (G_{n_j}). Theorem 5 then gives that $\lim_{j\to\infty} p^{(G_{n_j})} = \phi|_{\mathcal{F}^\sigma}$ for some $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. Now consider

$$\begin{aligned}f(\phi(F_1), \cdots, \phi(F_h)) &= \lim_{j\to\infty} f(p^{G_{n_j}}(F_1), \cdots, p^{G_{n_j}}(F_h)) && \text{(move limit outside by continuity of } f) \\ &\leq \tfrac{1}{2}C < 0 && \text{(as each term in the sequence is } < \tfrac{1}{2}C)\end{aligned}$$

That is, $\phi(f(F_1, \cdots, F_h)) < 0$. This gives that $f(F_1, \cdots, F_h) \not\geq_\sigma 0$. □

Proof of* $\Longleftarrow$ *direction of theorem 7. Let $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ arbitrary. Then, by theorem 6, $\phi|_{\mathcal{F}^\sigma} = \lim_{n\to\infty} p^{G_n}$ for some convergent sequence (G_n). Hence,

$$\begin{aligned}\phi(f(F_1, \cdots, F_h)) &= f(\phi(F_1), \cdots, \phi(F_h)) \\ &= \lim_{n\to\infty} f(p^{G_n}(F_1), \cdots, p^{G_n}(F_h)) \qquad \text{(move limit outside by continuity of } f) \\ &\geq \liminf_{n\to\infty} \min_{F\in\mathcal{F}^\sigma_n} f(p^F(F_1), \cdots, p^F(F_h) \geq 0\end{aligned}$$

This holds for all $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, hence the polynomial $f(F_1, \cdots, F_h) \geq_\sigma 0$. ☐

Remark 20. Razborov defines the set of relations $\{f \in \mathcal{A}^\sigma : f \geq_\sigma 0\}$ as the *semantic cone*, which he denotes by $\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$, and encapsulates the concept of an "asymptotically true relation" in his deductive calculus. I will stick to $\geq_\sigma$ which I believe is a slightly clearer notation.

Remark 21. There is a slight generalisation of theorem 7 given in [Raz07, Corrolary 3.4], in which f is generalised to a continuous function over $\phi(F_i)$, and $f : D \to \mathbb{R}$ has domain restricted to some $D \subset \mathbb{R}^h$. Its proof is no harder, but its statement adds unnecessary complexity without much gain.

4.8 Common constructs which have $f \geq_\sigma 0$

Note: The proof of (ii) below is my own; its triviality means it is usually be stated without proof in the literature.

By manipulating flag algebras, we wish to be able to find elements $f \in \mathcal{A}^\sigma$ which have $f \geq_\sigma 0$. Some important examples are below:

i) The *quadratic, or ordinary cone*, $\mathcal{C}(\mathcal{F}^\sigma) = \{F_1F_2\cdots F_k f^2 \text{ for } F_1, \cdots, F_k \in \mathcal{F}^\sigma \text{ and } f \in \mathcal{A}^\sigma\}$ trivially has $g \geq_\sigma 0$ for all $g \in \mathcal{C}(\mathcal{F}^\sigma)$.

   Proof. For any $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, $\phi(F_1F_2\cdots F_k f^2) = \phi(F_1)\cdots\phi(F_k)\phi(f)^2 \geq 0$. ☐

ii) $Q(f_1, \cdots, f_r) \geq_\sigma 0$ for $f_i \in \mathcal{A}^\sigma$, and Q a real positive semi-definite quadratic form.

   Proof. Take an orthonormal eigenbasis of Q with the k^{th} eigenvector being $\vec{e}^k = (e^k_1, \cdots, e^k_r)$, with corresponding eigenvalue $\lambda_k \geq 0$. Then Q is the sum of the weighted projectors, ie, $Q_{ij} = \sum_{k=1}^r \lambda_k e^k_i e^k_j$. Hence,

$$Q(f_1, \cdots, f_r) = \sum_{1\leq i,j\leq r} Q_{ij}\ f_i \cdot f_j = \sum_{1\leq i,j\leq r} \sum_{k=1}^r \lambda_k (e^k_i f_i) \cdot (e^k_j f_j) = \sum_{k=1}^r \lambda_k \left(\sum_{i=1}^r e^k_j f_j\right)^2 \geq_\sigma 0$$

   Where the last inequality follows as a sum of positive terms. ☐

iii) $[\![f]\!]_{\sigma,\eta} \geq_{\sigma|\eta} 0$ if $f \geq_\sigma 0$.

   Proof. See corollary 9 in the next section. ☐

iv) $\pi^{(I)}(f) \geq_{\sigma_2} 0$ if $f \geq_{\sigma_1} 0$, where I is an interpretation taking T_1 over $\mathcal{L}_1^{\sigma_1}$ to T_2 over $\mathcal{L}_2^{\sigma_2}$. In particular, $\pi^{\sigma,\eta}(f) \geq_\sigma 0$ if $f \geq_{\sigma|_\eta} 0$.

   Proof. Take $\phi \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_2}[T_2], \mathbb{R})$ arbitrary. Then $\phi \circ \pi^{(I)} \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_1}[T_1], \mathbb{R})$ because $\pi^{(I)}$ is also an algebra homomorphism, and moreover sends flags to a positive sum of flags, giving $(\phi \circ \pi^{(I)})(F) > 0$ for all flags $F \in \mathcal{F}^{\sigma_2}[T_2]$. Hence $\phi(\pi^{(I)}(f)) = (\phi \circ \pi^{(I)})(f) \geq 0$. ☐

Remark 22. As we shall see later, most SDP proofs combine (*ii*) and (*iii*), and use terms of the form $[\![Q(\vec{f})]\!]_\sigma$.

4.9 Extremal homomorphisms

For some continuous function $f : C \to \mathbb{R}$ (with $C \subset \mathbb{R}^h$ closed), and fixed σ-flags F_i, we can define f over $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ by taking $f(\phi) = f(\phi(F_1), \cdots, \phi(F_h))$.

By compactness of $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, the infimum/supremum of $f(\phi)$ is attained for some (possibly many) extremal homomorphism/s $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. That is:

$$f(\phi_0) = \inf_{\substack{\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R}) \\ (\phi(F_1), \cdots, \phi(F_h)) \in C}} f(\phi)$$

For example, we could take $T = T_{\triangle\text{-free Graph}}$, $\sigma = 0$, $f(x) = x$, $C = [0, 1]$, $F_1 = Pentagon$, and take a sup instead of an inf to find a homomorphism having the maximum density of pentagons in a triangle free graph (see section 7). In [Raz07], Razborov also introduces some tools (such as a form of differentiation) which can be used to derive information about ϕ_0, but these aren't needed in this essay.

4.10 Ensembles of random homomorphisms extending a labelling

Note: This section is slightly outside the core understanding necessary to apply the theory, but its results (corollaries 9 and 10) underpin most of the SDP proofs, and moreover it is instrumental in the main proof in the pentagons paper (section 7). Razborov's proof of the main result in this section makes use of advanced probability theory and isn't particularly relevant to the rest of the paper, so only a sketch proof will be given.

Fix some (non-degenerate) type σ_0, and let the pair σ, η be such that $\sigma|_\eta = \sigma_0$. Let such a pair σ, η be known as an *extension* of σ_0. Take some homomorphism $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_0}, \mathbb{R})$, over the smaller type, which has $\phi([\![1^\sigma]\!]_{\sigma,\eta}) > 0$ (ie has positive density of the extended type).

The goal of this section is to be able to lift ϕ_0 to some natural $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ over the larger type σ, but a little thought shows there may be many such ϕs. It turns out there is instead a unique probability measure on (Borel subsets of) $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, giving a *random* extension homomorphism $\boldsymbol{\phi}^{\sigma,\eta}$, which captures the lifting of ϕ_0 in an exact sense.

We define a probability measure $\mathbb{P}^{\sigma,\eta}$ on the Borel sets of $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ to *extend* ϕ_0 if, for any $f \in \mathcal{A}^\sigma$,

$$\phi_0([\![f]\!]_{\sigma,\eta}) = \phi_0([\![1^\sigma]\!]_{\sigma,\eta}) \int_{\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})} \phi(f)\ \mathbb{P}^{\sigma,\eta}(d\phi)$$

Theorem 8. *For $\sigma_0, \sigma, \eta, \phi_0$ as above, there is a* unique *extension* $\mathbb{P}^{\sigma,\eta}$ *of* ϕ_0.

Proof Sketch. The uniqueness part follows from purely measure-theoretic arguments. A sketch of the (more important) existence part is presented here.

For $\hat{F}_0 = (M, \theta_0)$ a σ_0-flag with $p([\![1^\sigma]\!]_{\sigma,\eta}, \hat{F}_0) > 0$, define a (discrete) probability measure $\mathbb{P}^{\sigma,\eta}_{\hat{F}}$ on Borel subsets $[0, 1]^{\mathcal{F}^\sigma}$ as follows: choose a uniformly random flag $\hat{\boldsymbol{F}} = (M, \boldsymbol{\theta}) \in \mathcal{F}^\sigma$ by extending the labelling θ_0 of F to some $\boldsymbol{\theta}$, so that we get a σ-flag. Then define:

$$\mathbb{P}^{\sigma,\eta}_{\hat{F}}(A) \stackrel{\text{def}}{=} \mathbb{P}\left[p^{\hat{\boldsymbol{F}}} \in A\right]$$

For any σ-flag F, and arbitrary σ_0-flag $\hat{F}_0 = (M, \theta_0)$ with $p^{\hat{F}_0}([\![1^\sigma]\!]_{\sigma,\eta}) > 0$ and $|\hat{F}_0| \geq |F|$, we have (where the second equality considers the definition of $\boldsymbol{\theta}$ above):

$$\frac{p^{\hat{F}_0}([\![F]\!]_{\sigma,\eta})}{p^{\hat{F}_0}([\![1^\sigma]\!]_{\sigma,\eta})} = \frac{q_{\sigma,\eta}(F)}{q_{\sigma,\eta}(1^\sigma)} p(F|_\eta, \hat{F}_0) = \sum_{\hat{F}} \mathbb{P}\left[\hat{\boldsymbol{F}} = \hat{F}\right] p(F, \hat{F}) = \mathbb{E}\left[p^{\hat{\boldsymbol{F}}}(F)\right] = \int_{[0,1]^{\mathcal{F}^\sigma}} \psi(F) \mathbb{P}^{\sigma,\eta}_{\hat{F}_0}(d\psi)$$

Now take $\hat{F}_0 = F_n$ a convergent sequence with $\phi_0|_{\mathcal{F}^\sigma} \stackrel{\mathsf{def}}{=} \lim_{n\to\infty} p^{F_n}$. By Prohorov's Theorem, it's possible to find a subsequence with $\mathbb{P}^{\sigma,\eta}_{F_{n_i}}$ converging weakly to some $\mathbb{P}^{\sigma,\eta}$. So taking the limit of the above:

$$\phi_0(\llbracket F \rrbracket_{\sigma,\eta}) = \phi_0(\llbracket 1^\sigma \rrbracket_{\sigma,\eta}) \int_{[0,1]^{\mathcal{F}^\sigma}} \psi(F)\mathbb{P}^{\sigma,\eta}(d\psi)$$

We nearly have our result. The rest of the proof is to show we can instead integrate over $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R}) \subset [0,1]^{\mathcal{F}^\sigma}$, ie we are required to show that $\mathbb{P}^{\sigma,\eta}(\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})) = 1$. Observe that $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ is defined as $\psi \in [0,1]^{\mathcal{F}^\sigma}$ satisfying a countable set of relations $r_i(\psi) = 0$ (which restrict ϕ to have $\phi(\mathcal{K}^\sigma) = 0$ and $\phi(f_1 \cdot f_2) = \phi(f_1)\phi(f_2)$; see the comments preceding theorem 5).

By observing $|r_i(p^F)| \leq O(|F|^{-1})$, we can argue that $\int_{\psi\in[0,1]^{\mathcal{F}^\sigma}} r_i(\psi)^2 \mathbb{P}^{\sigma,\eta}_{F_n}(d\psi) \leq O(|F_n|^{-2})$, and hence that $\int_{\psi\in[0,1]^{\mathcal{F}^\sigma}} r_i(\psi)^2 \mathbb{P}^{\sigma,\eta}(d\psi) = 0$. So $\mathbb{P}^{\sigma,\eta}(\{\psi : r_i(\psi) \neq 0\}) = 0$ for all i, giving that $\mathbb{P}^{\sigma,\eta}(\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})) = 1$ as required. □

We can now talk about a *random homomorphism rooted at ϕ_0, with extension σ, η* as $\boldsymbol{\phi}^{\sigma,\eta}$ chosen uniquely according to that probability measure above, such that, for any $f \in \mathcal{A}^\sigma$,

$$\mathbb{E}\left[\boldsymbol{\phi}^{\sigma,\eta}(f)\right] = \frac{\phi(\llbracket f \rrbracket_{\sigma,\eta})}{\phi(\llbracket 1^\sigma \rrbracket_{\sigma,\eta})} \qquad \text{(4.10.1, Extension Homomorphisms)}$$

If we look over all pairs σ, η, we get an *ensemble of random homomorphisms, rooted at ϕ_0*, $\{\boldsymbol{\phi}^{\sigma,\eta}\}$.

Corollary 9 (Asymptotic relations are conserved under averaging)**.** *If $f \geq_\sigma 0$ then $\llbracket f \rrbracket_{\sigma,\eta} \geq_{\sigma|_\eta} 0$.*

Proof. Take $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^{\sigma|_\eta}, \mathbb{R})$ arbitrary. If $\phi_0(\llbracket 1^\sigma \rrbracket_{\sigma,\eta}) = 0$, then $\phi_0(\llbracket f \rrbracket_{\sigma,\eta}) = 0$ so we're done.

Else, by (4.10.1, Extension Homomorphisms), $\phi_0(\llbracket f \rrbracket_{\sigma,\eta}) = \phi_0(\llbracket 1^\sigma \rrbracket_{\sigma,\eta})\mathbb{E}\left[\boldsymbol{\phi}^{\sigma,\eta}(f)\right]$. This is necessarily non-negative because all $\phi^{\sigma,\eta} \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ have $\phi^{\sigma,\eta}(f) \geq 0$ as $f \geq_\sigma 0$. □

Corollary 10 (Cauchy-Schwarz)**.**

$$\llbracket f^2 \rrbracket_{\sigma,\eta} \cdot \llbracket g^2 \rrbracket_{\sigma,\eta} \geq_{\sigma|_\eta} \llbracket fg \rrbracket^2_{\sigma,\eta}$$

Proof. Again, if $\phi \in \mathrm{Hom}^+(\mathcal{A}^{\sigma|_\eta}, \mathbb{R})$ has $\phi(\llbracket 1^\sigma \rrbracket_{\sigma,\eta}) = 0$, both sides are 0 so it's trivial. Else, consider extending ϕ_0 to $\boldsymbol{\phi}^{\sigma,\eta} \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, then, applying (4.10.1, Extension Homomorphisms), the statement simply becomes Cauchy-Schwarz for random variables:

$$\mathbb{E}\left[\boldsymbol{\phi}^{\sigma,\eta}(f)^2\right] \mathbb{E}\left[\boldsymbol{\phi}^{\sigma,\eta}(g)^2\right] \geq \mathbb{E}\left[\boldsymbol{\phi}^{\sigma,\eta}(f)\boldsymbol{\phi}^{\sigma,\eta}(g)\right]^2$$ □

Razborov continues with some more results concerning the consistency of the extension ensembles with respect to averaging and upwards operators; and includes a result (as is common in flag algebras) that says that knowing the extensions to all σ with $|\sigma| = k$ captures the information in extensions for all smaller σ.

In this essay, the only additional result from these which is required is [Raz07, Corollary 3.19], which is stated here without proof:

Lemma 11. *Let $\boldsymbol{\phi}^{\sigma,\eta}$ be an ensemble rooted at $\phi \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_0}, \mathbb{R})$. Then, for every $f \in \mathcal{A}^{\sigma_0}$,*

$$\mathbb{P}\left[\boldsymbol{\phi}^{\sigma,\eta}(\pi^{\sigma,\eta}(f)) = \phi(f)\right] = 1$$

In other words, the value of ϕ is captured by its extension through use of the upwards operator.

The proof is given in [Raz07, Page 33], as a corollary to some of the other consistency results concerning upwards operators.

5 Applying the Flag Algebras method

5.1 SDP methods

5.1.1 General approach

Note: This section is an amalgamation of many papers I have read, in particular [Raz10], [Hat+13], [Hir13] and [Cum+13].

In many problems in asymptotic extremal combinatorics, we wish to prove that the weighted sum of some set of model densities is asymptotically less than or equal to some constant. By section 4.7, this is equivalent to $g \leq_0 a$, for some $g \in \mathcal{A}^0$, $a \in \mathbb{R}$. For this section, we'll stick to graphs $g \in \mathcal{G} = \mathcal{F}^0$ satisfying some theory T, with the obvious extension to other combinatorial structures left implicit. The general approach to solving this is as follows:

Choose/find types $\tau_1, \ldots, \tau_n$; real positive semi-definite (PSD) quadratic forms (QFs) Q_1 of dimension $m_1, \cdots, Q_n$ of dimension m_n; and $f_i^{\tau_k} \in \mathcal{A}^{\tau_k}$ and express the following as a sum over graphs of size l chosen large enough to contain all the products, ie $l \geq |f_i^{\tau_k} \cdot f_j^{\tau_k}|$ for all i, j, k. (This is always possible as, per remark 1, the flags $F \in \mathcal{F}_l^\sigma$ form a basis for flags of size $\leq l$ in $\mathcal{A}^\sigma$):

$$g + [\![Q_1(f_1^{\tau_1}, \cdots, f_{m_1}^{\tau_1})]\!]_{\tau_1} + \cdots + [\![Q_n(f_1^{\tau_n}, \cdots, f_{m_n}^{\tau_n})]\!]_{\tau_n} = \sum_{G \in \mathcal{G}_l} \alpha_G G$$

Our aim is to choose types τ_k, forms Q_k and flag relations $f_i^{\tau_k}$ such that $\alpha_G \leq a$ for all $G \in \mathcal{G}_l$. Recalling that $a = a1^0 = \sum_{G \in \mathcal{G}_l} aG$ for any l, this would give $\sum_{G \in \mathcal{G}_l} \alpha_G G \leq_0 a$. Then, using the results of section 4.8 (on constructions giving $f \geq_\sigma 0$), we would have $g \leq_0 a$ as required.

5.1.2 How can this be represented as a semi-definite problem?

Using flag algebra methods, we can represent and save the coefficients of g and $[\![f_i^{\tau_k} \cdot f^{\tau_k}]\!]_{\tau_k}$ in the basis of $\mathcal{G}_l$. Let these be $g \stackrel{\text{def}}{=} \sum_{G \in \mathcal{G}_l} B_G G$, and $[\![f_i^{\tau_k} \cdot f^{\tau_k}]\!]_{\tau_k} \stackrel{\text{def}}{=} \sum_{G \in \mathcal{G}_l} A_{k;(i,j);G} G$. In particular, it is possible to calculate $[\![f_i^{\tau_k} \cdot f^{\tau_k}]\!]_{\tau_k}$ directly in the basis $\mathcal{G}_l$ instead of calculating in $\mathcal{A}_l^\sigma$ and averaging; for the details, please see [Raz10, Page 952].

Now if we let the quadratic form Q_k be represented by a matrix with entries $M_{k;(i,j)}$, then the term $[\![Q_k(f_1^{\tau_k}, \cdots, f_{m_k}^{\tau_k})]\!]_{\tau_k}$ is equal to $\sum_G (\sum_{i,j} A_{k;(i,j);G}\, M_{k;(i,j)})G$. So we can pose our task as a semi-definite problem over the reals as follows:

Choose $M_{k;(i,j)}$ in order to:

$$\text{Minimise } a = \min_{G \in \mathcal{G}_l} \alpha_G = \min_{G \in \mathcal{G}_l} \left(B_G + \sum_{k=1}^{n} \sum_{1 \leq i,j \leq m_k} A_{k;(i,j);G}\, M_{k;(i,j)} \right)$$

Subject to M_k being positive, semi-definite matrices.

Whilst general-purpose SDP solvers can help with this, there are still issues of long run times if $|\mathcal{G}_l|$ is too large or the dimension of the matrices m_k is too large. And of course, we still need to choose our types τ_k and our relations $f_i^{\tau_k} \in \mathcal{A}^{\tau_k}$. The next few sections give some advice on how to appropriately focus these choices.

5.1.3 Typical simplifications of the general approach

Firstly, we normally take the (majority of) types τ_k to have the same size s. Because they'll eventually be unlabelled, they should all be distinct as models. (ie, we choose distinct models in $\mathcal{M}_s$ and label them as we wish).

Secondly, the $f_i^{\tau_k} \in \mathcal{A}^{\tau_k}$ are also all typically taken to be the same size, ie $f_i^{\tau_k} \in \mathcal{A}_t^{\tau_k}$, and hence we can take $l = s + 2(t - s)$. As l increases, the dimension of our optimization space ($|\mathcal{M}_l|$) increases rapidly, so we wish to try and keep it as small as possible to aid computation.

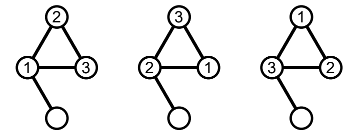
5.1.4 Invariant and anti-invariant parts

Note: The ideas in this section and the following section are courtesy of [Raz10], but were given there without proof. The proofs here are my own.

Given a fixed type σ, let Γ_σ be its group of automorphisms (you can imagine it acting by relabelling vertices of σ). For $\gamma \in \Gamma_\sigma$ and $F \in \mathcal{F}^\sigma$, we can define γF by applying γ to the labelling in F. We can extend this to γf for $f \in \mathcal{A}^\sigma$ by linearity.

Now take an $f \in \mathcal{A}^\sigma$. Define f^+, the *invariant* part of f as:

$$f^+ \stackrel{\text{def}}{=} |\Gamma_\sigma|^{-1} \sum_{\gamma \in \Gamma_\sigma} \gamma f$$



$F_1^\triangle$ $F_2^\triangle$ $F_3^\triangle$

Figure 6: As an example, $(F_1^\triangle)^+ = \frac{1}{3}(F_1^\triangle + F_2^\triangle + F_3^\triangle)$, $(F_1^\triangle)^- = \frac{1}{3}(2F_1^\triangle - F_2^\triangle - F_3^\triangle)$

Observe that:

$$f^{++} = |\Gamma_\sigma|^{-2} \sum_{\gamma, \gamma' \in \Gamma_\sigma} \gamma\gamma' f = |\Gamma_\sigma|^{-1} \sum_{\hat{\gamma} \in \Gamma_\sigma} \hat{\gamma} f = f^+$$

Hence the $\circ^+$ operator is idempotent. It is the part of f unchanged by relabelling. We define $f^- \stackrel{\text{def}}{=} f - f^+$ as the *anti-invariant* part of f. Observe that it has $(f^-)^+ = f^+ - f^{++} = 0$. Hence, we can split $\mathcal{A}^\sigma$ into a sum of its invariant and anti-invariant parts:

$$\mathcal{A}^\sigma = \mathcal{A}^{\sigma,+} \oplus \mathcal{A}^{\sigma,-}$$

Lemma 12. $[\![\mathcal{A}^{\sigma,+} \cdot \mathcal{A}^{\sigma,-}]\!]_\sigma = 0$

Proof. Observe that if $F, G \in \mathcal{F}^\sigma$ and $\gamma \in \Gamma_\sigma$, then as unlabelled models, $(\gamma F) \cdot G$ and $F \cdot (\gamma^{-1} G)$ are identical (when taking the product, they identify the same vertices, just with permuted labels). We observe this also gives that the relevant q_σ factors are equal, hence $[\![(\gamma F) \cdot G]\!]_\sigma = [\![F \cdot (\gamma^{-1} G)]\!]_\sigma$. Summing over all $\gamma \in \Gamma_\sigma$ gives that $[\![F^+ \cdot G]\!]_\sigma = [\![F \cdot G^+]\!]_\sigma$, so by linearity, $[\![f^+ \cdot g]\!]_\sigma = [\![f \cdot g^+]\!]_\sigma$ for all $f, g \in \mathcal{A}^\sigma$.

In particular, taking $f \in \mathcal{A}^{\sigma,+}$, $g \in \mathcal{A}^{\sigma,-}$, then $[\![f \cdot g]\!]_\sigma = [\![f^+ \cdot g]\!]_\sigma = [\![f \cdot g^+]\!]_\sigma = 0$. ☐

Hence, if we are constructing a positive semi-definite quadratic form Q over $f_i \in \mathcal{A}^\sigma$, then we can instead split it into $Q = Q^+ + Q^-$, where Q^+ acts over $f_i \in \mathcal{A}^{\sigma,+}$, and Q^- acts over $f_i \in \mathcal{A}^{\sigma,-}$.

5.1.5 Making use of a conjectured extremal homomorphism

Often with these problems we have in mind an extremal convergent sequence with homomorphism ϕ, achieving $\phi(g) = k$ (or even a set Φ of them). So, if we wish to show $\psi(g) \le k$ for all $\psi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$, then in particular, all our inequalities must be tight for each such ϕ.

Denote by $\Delta^\sigma \subset \mathcal{A}^\sigma$ the ideal of all relations $f \in \mathcal{A}^\sigma$ which are exactly true for all $\phi \in \Phi$. This can be defined rigorously in terms of the labelling-extension measure as:

$$\Delta^\sigma \stackrel{\text{def}}{=} \{f \in \mathcal{A}^\sigma \,:\, \mathbb{P}[\boldsymbol{\phi}^\sigma(f) = 0] = 1 \text{ for all } \phi \in \Phi\}$$

For ease, it can be shown that $[\![\Delta^\sigma]\!]_\sigma \subset \Delta^0$ which is a more easily usable condition. The lack of slackness for such ϕ means that our quadratic forms Q for a type σ must be supported on Δ^σ. This can be combined with the invariance result by taking:

$$\Delta^\sigma = \Delta^{\sigma,+} \oplus \Delta^{\sigma,-}$$

Where $\Delta^{\sigma,+} \subset \mathcal{A}^{\sigma,+}$ and $\Delta^{\sigma,-} \subset \mathcal{A}^{\sigma,-}$. Then Q^+ is supported on $\Delta^{\sigma,+}$ and Q^- supported on $\Delta^{\sigma,-}$. That is, if we use for example $[\![Q^+(f_1^\sigma, \cdots, f_n^\sigma)]\!]_\sigma$, then we must have $f_i^\sigma \in \mathcal{A}^{\sigma,+}$ and $\phi([\![f_i^\sigma]\!]_\sigma) = 0$.

5.1.6 Other considerations

Programs to solve semi-definite problems typically work with floating point numbers. Due to rounding errors, these can't give rigorous proofs. So typically at the end of a calculation, the matrices M representing the quadratic forms Q have to have their coefficients replaced with close rational points, such that the inequality still holds (and such that the eigenvalues remain positive).

General purposes solvers frequently mentioned in literature include *CSDP* and *SDPA*. More relevantly, the *Flagmatic* program [Fla], available at `http://flagmatic.org/`, is a Flag-Algebras specific set of tools developed to work in the *Sage* mathematics environment, with support for 2-graphs, 3-graphs and oriented 2-graphs. It attempts to solve the problem automatically, including using the simplifications included in this section. More information about how it works, and the intricacies of the SDP calculation, is given in its introductory paper, [FRV13] (Victor Falgas-Ravry, Emil R. Vaughan).

5.2 Overview of useful densities and how they relate to flag algebras

5.2.1 Different types of embedding/homomorphism

Note: The information in this section comes from [Lov12], [Hat+13], [Raz13a] and [Hir13].

Results in asymptotic extremal combinatorics generally deal with a variety of closely related probabilities over homomorpisms, embeddings and substructures. I'll quickly review these, most of which will be useful in the next section and/or the analysis of the papers in the following sections.

Let $M, N \in \mathcal{M}$, with $m = |M|$, $n = |N|$. A *homomorphism* describes a function $\alpha : V(M) \to V(N)$, which satisfies $P(v_1, \ldots, v_n) \implies P(\alpha(v_1), \cdots, \alpha(v_n))$ for each predicate $P \in \mathcal{L}$. (Eg in T_{Graph}, these preserve graph edges). A homomorphism is described as *induced* if it also preserves falsehood of predicates, ie, for all $P \in \mathcal{L}$, $P(v_1, \ldots, v_n) \iff P(\alpha(v_1), \cdots, \alpha(v_n))$. (Eg in T_{Graph}, these preserve edges and non-edges).

	Not necessarily injective			**Injective**		
	Name	*Amount*	*Density*	*Name*	*Amount*	*Density*
Non-induced	Graph	$\hom(M,N)$	$t(M,N)$	Injective	$\mathrm{inj}(M,N)$	$t_{\text{inj}}(M,N)$
Induced	Strong	$\mathrm{s}(M,N)$	$t_{\text{str}}(M,N)$	Induced	$\mathrm{ind}(M,N)$	$t_{\text{ind}}(M,N)$

Table 1: A table defining the different types of graph homomorphisms. The *Name* column relates to the term preceding "homomorphism" in the name of such objects in T_{Graph}. The *Amount* relates to the number of such functions. For non-injective homomorphisms, *Density* = Amount / n^m, and for injective homomorphisms, *Density* = Amount / $[n(n-1)\cdots(n-m+1)]$.

For comparison, the quantity of interest in flag algebras is $p(M, N)$, the *induced substructure density*, (the amount of such [Raz13a], or the density itself [Hir13], is sometimes also called the *inducibility*, $i(M, N)$).

Observe that if we fix some $V \subset V(M)$ with $M|_V \cong N$, then of the $m!$ injections $\alpha : M \to V$, there are exactly $|\mathrm{Aut}(M)|$ induced homomorphisms (where $\mathrm{Aut}(M)$ is the group of automorphisms of M). Hence $p(M, N)$ is related to $t_{\text{ind}}(M, N)$ by:

$$p(M,N) = \frac{m!\, t_{\text{ind}}(M,N)}{|\mathrm{Aut}(M)|} = \frac{\mathrm{ind}(M,N)}{|\mathrm{Aut}(M)|\binom{n}{m}} \qquad \text{(5.2.1, Formula for } p \text{ in terms of } t_{\text{ind}})$$

Calculating the non-induced versions So from $p(H, G)$, we can get $t_{\text{ind}}(H, G)$. But can we get the non-induced versions? Yes, the non-induced versions in the above table can be calculated from the induced versions by simply summing over all models $M' \models T$ with $V(M') = V(M)$ and $P^{M'} \supset P^M$ for each predicate $P \in \mathcal{L}$. (ie, in graphs, we form M' by adding edges to M). Eg,

$$t_{\text{inj}}(M,N) = \sum_{\substack{M' \supset M \\ V(M')=V(M)}} t_{\text{ind}}(M',N) \qquad (5.2.2,\ t_{\text{inj}} \text{ from } t_{\text{ind}})$$

Remark 23. In many common theories, we can invert this with a Möbius style inversion to get, eg, in T_{Graph}:

$$t_{\text{ind}}(H,G) = \sum_{\substack{H' \supset H \\ V(H')=V(H)}} (-1)^{e(H')-e(H)}\, t_{\text{inj}}(H',G) \qquad (5.2.3,\ t_{\text{ind}} \text{ from } t_{\text{inj}})$$

Calculating the non-injective versions But can we get the non-injective versions? For most purposes, the theory of flag algebras deals with asymptotics, and thus the injective/non-injective densities into convergent sequences are equal, by analogues of lemma 4, so this question is moot. But for the purposes of the section on blow-ups, we quickly consider if we can derive the number of strong homomorphisms from the number of induced homomorphisms.

This is possible. Define vertices v_1, v_2 to be *twins* if they are connected in the same way to every other vertex of M (formally, if $P(w_1, \cdots, w_k, v_1, w_{k+1}, \cdots w_n) \iff P(w_1, \cdots, w_k, v_2, w_{k+1}, \cdots w_n)$ for all predicates $P \in \mathcal{L}$ and vertices w_i). Then, for P any partition of $V(M)$ into classes of twins, we can form a quotient model M/P. Summing over all such P gives:

$$s(M,N) = \sum_P \text{ind}(M/P,N) \qquad (5.2.4,\ s \text{ from ind})$$

Remark 24. (5.2.4, s from ind) can also be inverted in many common theories (such as graphs), by a Möbius sum over a lattice (see [Lov12, Eq 5.18] for related details) which determines $\text{ind}(H,G)$ as a weighted sum over $s(H/P,G)$.

Remark 25. It is possible to capture the same asymptotics as flag algebras, using homomorphism densities t, t_{ind} instead of the induced substructure densities p (see, eg, Hirst's paper on inducibility, [Hir13]). However, the arguments concerning densities in finite objects no longer hold, so the limit objects of T, such as graphons, have to be used instead. Inductive results can be applied by equating each graph H with a certain Möbius style sum of other graphs. It turns out that the description of the product and $[\![\cdot]\!]$ operator is more straightforward in this algebra (as we can just use disjoint unions of the unlabelled vertices, and probabilistic weightings can be neglected), but the conceptual leap and use of graphons makes it more difficult to introduce.

5.2.2 Turán densities

A common quantity of study in asymptotic extremal combinatorics is Turán densities. These are very closely linked to flag algebras.

Let $\mathcal{F}$ be a family of r-graphs. An r-graph G is *$\mathcal{F}$-free* if it does not contain any (not necessarily induced) subgraphs isomorphic to any $F \in \mathcal{F}$. Similarly, I will define an r-graph G to be *$\mathcal{F}$-induced-free* if it does not contain any induced subgraphs isomorphic to $\mathcal{F}$.

Take some target r-graph H, the Turán density, $\pi_H(\mathcal{F})$ is traditionally defined as the limit of the maximal density of H in an $\mathcal{F}$-free graph G as $|G| \to \infty$. Similarly, I'll define $\pi_{\text{ind},H}(\mathcal{F})$ to be the same but require G to be $\mathcal{F}$-induced-free. Note that by adding all graphs formed by adding possible sets of edges to graphs in $\mathcal{F}$ to form $\mathcal{F}'$, then $\pi_H(\mathcal{F}) = \pi_{\text{ind},H}(\mathcal{F}')$, so we don't lose any power with this notation. If H is excluded, it is implied to be an r-edge.

We also define $\pi_{\min,H}$ to be the same as $\pi_{\text{ind},H}$, except we look for the *minimal* density of H. Observe that, by taking graph complements, it captures the same ideas: $\pi_{\text{ind},H}(\mathcal{F}) = \pi_{\text{ind},\bar{H}}(\bar{\mathcal{F}}) = 1 - \pi_{\min,H}(\bar{\mathcal{F}})$, where $\bar{\mathcal{F}} = \{\bar{F} : F \in \mathcal{F}\}$. But it is still a useful notation to have.

In flag algebras, taking $T = \mathcal{F}$-induced-free r-graphs, we have:

$$\pi_{\text{ind},\mathcal{F}} = \max_{\phi \in \text{Hom}^+(\mathcal{A}^0,\mathbb{R})} \phi(H)$$

So showing $H \leq_0 k$ over this theory is equivalent to showing $\pi_{\text{ind},\mathcal{F}} \leq k$.

5.3 Finite results from asymptotics

To quote Razborov's flag algebra survey ([Raz13a, Page 15], with a few minor additions in brackets):

> This was also one of the first papers to demonstrate the three-step program for converting asymptotic flag-algebraic results into exact ones:
>
> 1) Describe the set of all extremal elements in $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ (often a single element).
> 2) Prove stability, that is that the convergence in the pointwise topology can be strengthened to convergence in the edit distance (also known as the cut distance).
> 3) Move from stability to exact results using combinatorial techniques.

In particular, the second half of the pentagons in triangle-free graphs paper (Hatami et al) uses this technique. The extremal construction takes the form of a blow-up (see the following section).

5.3.1 Blow-ups

Note: Material in this subsection is mostly taken (with modifications) from [Hat+13], the Pentagons in Triangle-Free graphs paper, and will be used in the exposition of said paper in section 7.

Extremal homomorphisms typically take the form of either a randomised construction, or a more ordered construction, typically a blow-up. I will introduce these for graphs for ease, but the definition readily extends to hypergraphs. It should be noted that extremal constructions for hypergraphs sometimes involve a more complex generalized *iterative blow-up* construction [FRV13].

A *blow-up* of a graph G by a vector of positive integers $\vec{k} = \{k_v : v \in V(G)\}$ is a graph $G^{\vec{k}}$ which replaces each vertex $v \in G$ with k_v twin nodes. Formally, we construct it by taking $V(G^{\vec{k}}) := \{(v, i) : v \in V(G), i \in [k_v]\}$ and $E(G^{\vec{k}}) := \{((v, i), (w, j)) : (v, w) \in E(G)\}$.

A blow-up is *balanced* if all the k_v are equal, a balanced blow-up with $\vec{k} = (k, \cdots, k)$ is denoted $G^{(k)}$. If the k_v vary by at most 1, this is often denoted *almost balanced*. As a reminder of a definition at the end of the previous section, vertices which have the same connectivity to all other vertices in the graph are known as *twin nodes*. In particular, (v, i) and (v, j) are twin nodes in $G^{(k)}$.

It is easily observed, that for any graph H, $\lim_{k\to\infty} p(H, G^{(k)})$ converges, hence by theorem 5, there exists $\phi_G|_{\mathcal{F}^\sigma} \stackrel{\text{def}}{=} \lim_{k\to\infty} p^{G^{(k)}}$ with $\phi_G \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. So blow-ups are a useful way of representing a finite graph as a convergent graph sequence. It also turns out that many extremal asymptotic constructions take this form. The rest of this section is devoted to understanding ϕ_G.

Now what is $\phi_G(H)$? Let $|H| = m$, $|G| = n$. Consider a random injection $\boldsymbol{\alpha}_k : V(H) \to V(G^{(k)})$ with $h \mapsto (v_h, i_h)$. Define $\boldsymbol{\alpha}'_k : V(H) \to V(G)$ taking $h \mapsto v_h$ (note: this map is not injective). Observe that $E(\boldsymbol{\alpha}_k(h_1), \boldsymbol{\alpha}_k(h_2)) \iff E(\boldsymbol{\alpha}'_k(h_1), \boldsymbol{\alpha}'_k(h_2))$, also, in the limit, $\boldsymbol{\alpha}'_\infty$ maps each $h \in V(H)$ to each $v \in V(G)$ uniformly and independently. Thus,

$$\lim_{k\to\infty} t_{\text{ind}}(H, G^{(k)}) = \lim_{k\to\infty} \mathbb{P}\left[\boldsymbol{\alpha}_k \text{ is an induced hom}\right] = \mathbb{P}\left[\boldsymbol{\alpha}'_\infty \text{ is a strong hom}\right] = t_{\text{str}}(H, G)$$

Hence, by (5.2.1, Formula for p in terms of t_{ind}),

$$\phi_G(H) = \lim_{k\to\infty} \frac{m!}{|\mathrm{Aut}(H)|} \cdot t_{\text{ind}}(H, G^{(k)}) = \frac{m!}{|\mathrm{Aut}(H)|} \cdot \frac{s(H, G)}{n^m} \qquad \text{(5.3.1, Formula for } \phi_G\text{)}$$

In particular, if H is twin-free, then (5.2.4, s from ind) gives that $s(H, G) = \mathrm{ind}(H, G)$, so, by (5.2.1, Formula for p in terms of t_{ind}), $s(H, G) = p(H, G)|\mathrm{Aut}(H)|\binom{n}{m}$, giving:

$$\phi_G(H) = \frac{n(n-1)\cdots(n-m+1)}{n^m} p(H, G) \qquad \text{(5.3.2, Formula for } \phi_G(H) \text{ with } H \text{ twin-free)}$$

To finish off this section, we look at the uniqueness of ϕ_G. The following proofs are adapted from related (albeit slightly different) proofs in [Lov12, Chapter 5.4].

Lemma 13. *If $|G_1| = |G_2|$ and $\phi_{G_1} = \phi_{G_2}$ then $G_1 \cong G_2$.*

Proof. By (5.3.1, Formula for ϕ_G), we have $s(H, G_1) = s(H, G_2)$ for every graph H. By remark 24 in the previous section, we can calculate $\mathrm{ind}(H, G)$ as a Möbius style sum over terms $s(H/P, G)$, and thus we must have $\mathrm{ind}(H, G_1) = \mathrm{ind}(H, G_2)$ for all H. In particular, $\mathrm{ind}(G_1, G_1) = \mathrm{ind}(G_2, G_1) > 0$ and $\mathrm{ind}(G_2, G_1) = \mathrm{ind}(G_2, G_2) > 0$, so there are induced homomorphisms $G_1 \to G_2$ and $G_2 \to G_1$, hence $G_1 \cong G_2$. $\square$

Corollary 14. *If $\phi_{G_1} = \phi_{G_2}$, then G_1 and G_2 are both balanced blow-ups of a mutual graph, G.*

Proof. Take l the lowest common multiple of $|G_1|$ and $|G_2|$, with $l = k_1|G_1| = k_2|G_2|$ and k_1, k_2 coprime. By considering the definition of $\phi_G(H)$, it is clear that $\phi_G = \phi_{G^{(k)}}$ for any k. Hence $\phi_{G_1^{(k_1)}} = \phi_{G_1} = \phi_{G_2} = \phi_{G_2^{(k_2)}}$. And thus by lemma 13, $G_1^{(k_1)} = G_2^{(k_2)}$. So, in this graph, the number of elements in every class of twin nodes is divisible by both k_1 and k_2, and so by $k_1 k_2$. Hence, $G_1^{(k_1)} = G_2^{(k_2)} = G^{(k_1 k_2)}$ for some graph G, and thus $G_1 = G^{(k_2)}$ and $G_2 = G^{(k_1)}$. Ie, G_1 and G_2 are both blow-ups of G. $\square$

6 On 3-hypergraphs with forbidden 4-vertex configurations

This 2010 paper by Razborov [Raz10] is presented here as a useful demonstration of the plain SDP approach. It was "the first application of the SDP method in its genuinely plain form" [Raz13a, Page 14], presenting the techniques for many that followed.

6.1 Introduction to the problem and overview of the paper

The famous Turán problem concerns determining $\pi(K_l^r)$, ie the maximum asymptotic edge-density in r-graphs which contain no complete r-graph on l vertices. This is still unknown (as of the flag algebras survey, [Raz13a]) for any $l > r > 2$. We note here that $\pi(K_l^r) = \pi_{\text{ind}}(K_l^r) = 1 - \pi_{\min}(I_l^r)$ for I_l^r the independent set of size l. Henceforth, we restrict to the theory of 3-graphs, and look specifically at the $(r, l) = (3, 4)$ case, sometimes known as *Turán's Tetrahedron Problem.* Turán proved many years ago that

$$\pi_{\min}(I_4) \leq 4/9$$

Where I_4 denotes the independent set of size 4. He did this by an asymptotic construction achieving this bound. He conjectured that $\pi_{\min}(I_4) = 4/9$ exactly, but the progress on this has been slow, and the current (as of [Raz13a]) best lower bound is $\pi_{\min}(I_4) \geq 0.438334$ (also given in [Raz10] as a result of an SDP flag algebras calculation). Razborov noted however that Turán's construction was also missing the (unique) 3-graph on 4 vertices with 3 edges. Denoting the unique 3-graph on with k edges on 4 vertices as G_k, Razborov observed (see table 2 below) that Turán's construction exhibits:

$$\pi_{\min}(I_4, G_3) \leq 4/9$$

In this paper, Razborov showed with a "plain" SDP flag algebras calculation, that $\pi_{\min}(I_4, G_3) = 4/9$ exactly, by showing that $\pi_{\min}(I_4, G_3) \geq 4/9$. This is equivalent (cf section 5.2.2) to showing that $\rho \geq_0 4/9$ (where ρ is the model of a 3-edge in $\mathcal{F}^0$), over the theory $\hat{T}$ of 3-graphs with no induced I_4 or G_3. We will call a 3-graph (respectively 3-flag) *admissable* if it satisfies $\hat{T}$.

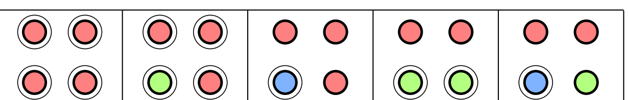
6.2 Turán's construction

As per the simplifications in section 5.1, the form of the extremal construction is important, to help us to choose the correct flags so that our inequality is tight on the extremal construction.

The construction is as follows: Colour the vertices of the graph red/green/blue uniformly at random. We form a 3-edge over a set of 3 vertices if EITHER all three are the same colour OR two are red, one is green OR two are green, one is blue OR two are blue, one is red.

In this paper, I will denote by ϕ the corresponding homomorphism in $\text{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. Tables 2 and 3 show the representations (as per remark 26) of the possible induced four and five vertex graphs, and their relevant probabilities in the construction. The possible induced graphs will be known as *regular.*

Remark 26. To represent a 3-graph G on four vertices pictorially, its complement vertex set will be marked, that is $E \subset V(G)$ is a 3-edge in G if and only if the vertex $V(G) \setminus E$ is circled. To represent a 3-graph G on five vertices pictorially, its complement 2-graph will be drawn, that is $E \subset V(G)$ is a 3-edge in G if and only if the edge $V(G) \setminus E$ is drawn.



	4, 0, 0	3, 1, 0	3, 0, 1	2, 2, 0	2, 1, 1
Probability	$\frac{1}{27}$	$\frac{4}{27}$	$\frac{4}{27}$	$\frac{6}{27}$	$\frac{12}{27}$

Table 2: The different possible colourings of 4 vertices, up to the symmetric cyclic re-ordering of colours; and the representations of the corresponding 3-graphs under Turán's construction. This gives $\phi(I_4) = 0$, $\phi(G_1) = 16/27$, $\phi(G_2) = 6/27$, $\phi(G_3) = 0$, $\phi(K_4) = 5/27$.

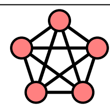
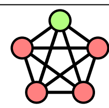
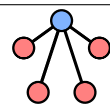
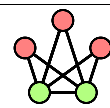
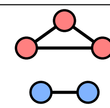
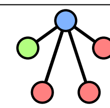
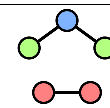
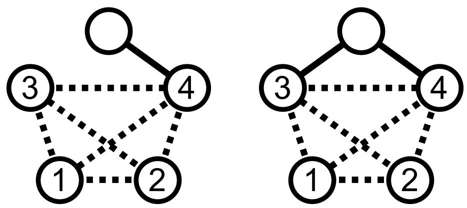
	 5, 0, 0	 4, 1, 0	 4, 0, 1	 3, 2, 0	 3, 0, 2	 3, 1, 1	 2, 2, 1
Probability	$\frac{1}{81}$	$\frac{5}{81}$	$\frac{5}{81}$	$\frac{10}{81}$	$\frac{10}{81}$	$\frac{20}{81}$	$\frac{30}{81}$

Table 3: The different possible colourings of 5 vertices, up to the symmetric cyclic re-ordering of colours; and the representations of the corresponding 3-graphs under Turán's construction.

6.3 The relevant types, and regular flags

Razborov's SDP proof uses two types, τ_1 and τ_2, which are labellings of G_1 (with a 3-edge on $\{1,2,3\}$) and G_2 (with 3-edges on $\{1,2,3\}$, $\{1,2,4\}$) respectively. These are represented as graphs on 5 vertices in figure 7. Observe that $\Gamma_{\tau_1} \cong \mathbb{Z}_3$, $\Gamma_{\tau_2} \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.



τ_1 τ_2

Figure 7: The main types used in the paper.

The SDP will operate over flags of size 5. There are 17 flags in $\mathcal{F}_5^{\tau_1}$ and 15 flags in $\mathcal{F}_5^{\tau_2}$. The τ_1 flags are shown in tables 4 and 5, and the τ_2 flags are shown in tables 6 and 7.

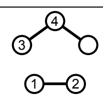
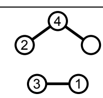
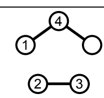
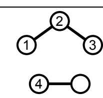
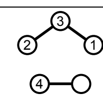
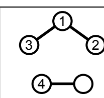
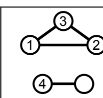
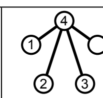
	 $F_1^{\tau_1}$	 $F_2^{\tau_1}$	 $F_3^{\tau_1}$	 $F_4^{\tau_1}$	 $F_5^{\tau_1}$	 $F_6^{\tau_1}$	 $F_7^{\tau_1}$	 $F_8^{\tau_1}$
$\phi(F_i^{\tau_1}\|_{\tau_1})$	$\frac{30}{81}$	$\frac{30}{81}$	$\frac{30}{81}$	$\frac{30}{81}$	$\frac{30}{81}$	$\frac{30}{81}$	$\frac{10}{81}$	$\frac{25}{81}$
$q_{\tau_1}(F_i^{\tau_1})$	$\frac{4}{125}$	$\frac{4}{125}$	$\frac{4}{125}$	$\frac{4}{125}$	$\frac{4}{125}$	$\frac{4}{125}$	$\frac{12}{125}$	$\frac{24}{125}$
$\phi(\llbracket F_i^{\tau_1}\rrbracket_{\tau_1})$	$\frac{120}{10125}$	$\frac{120}{10125}$	$\frac{120}{10125}$	$\frac{120}{10125}$	$\frac{120}{10125}$	$\frac{120}{10125}$	$\frac{120}{10125}$	$\frac{600}{10125}$

Table 4: The regular τ_1 flags. $F_1^{\tau_1}+F_2^{\tau_1}+F_3^{\tau_1}$ is invariant, as are $F_4^{\tau_1}+F_5^{\tau_1}+F_6^{\tau_1}$, $F_7^{\tau_1}$ and $F_8^{\tau_1}$.

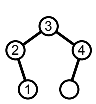
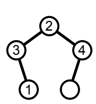
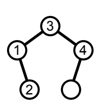
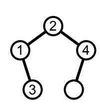
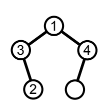
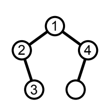
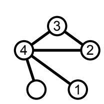
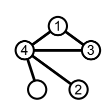
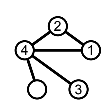
								
$F_9^{\tau_1}$	$F_{10}^{\tau_1}$	$F_{11}^{\tau_1}$	$F_{12}^{\tau_1}$	$F_{13}^{\tau_1}$	$F_{14}^{\tau_1}$	$F_{15}^{\tau_1}$	$F_{16}^{\tau_1}$	$F_{17}^{\tau_1}$

Table 5: The remaining τ_1 flags. $F_9^{\tau_1}+F_{12}^{\tau_1}+F_{13}^{\tau_1}$ is invariant, as are $F_{10}^{\tau_1}+F_{11}^{\tau_1}+F_{14}^{\tau_1}$ and $F_{15}^{\tau_1}+F_{16}^{\tau_1}+F_{17}^{\tau_1}$.

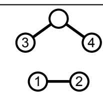
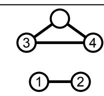
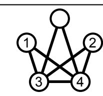
	 $F_1^{\tau_2}$	 $F_2^{\tau_2}$	 $F_3^{\tau_2}$
$\phi(F_i^{\tau_2}\|_{\tau_2})$	$\frac{30}{81}$	$\frac{10}{81}$	$\frac{10}{81}$
$q_{\tau_2}(F_i^{\tau_2})$	$\frac{4}{125}$	$\frac{4}{125}$	$\frac{4}{125}$
$\phi([\![F_i^{\tau_2}]\!]_{\tau_2})$	$\frac{120}{10125}$	$\frac{40}{10125}$	$\frac{40}{10125}$

Table 6: The regular τ_2 flags. $F_1^{\tau_2}$, $F_2^{\tau_2}$ and $F_3^{\tau_2}$ are all invariant.

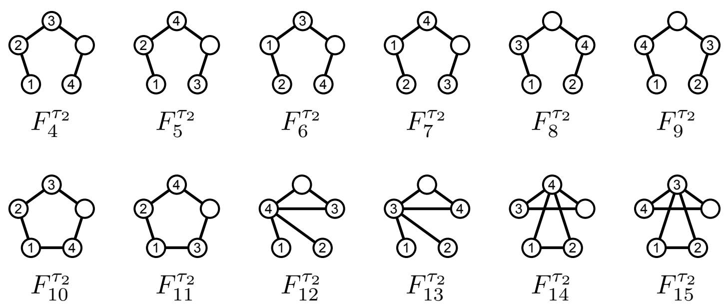


Table 7: The remaining τ_2 flags. $F_4^{\tau_2} + F_5^{\tau_2} + F_6^{\tau_2} + F_7^{\tau_2}$ is invariant, as are $F_8^{\tau_2} + F_9^{\tau_2}$, $F_{10}^{\tau_2} + F_{11}^{\tau_2}$, $F_{12}^{\tau_2} + F_{13}^{\tau_2}$ and $F_{14}^{\tau_2} + F_{15}^{\tau_2}$.

6.4 Setting up the SDP problem

The choice of τ_1 and τ_2 was by experimentation. It was then apparent that whilst the complete graph on n vertices is regular ($\phi(K_n) > 0$), there is no regular τ_1 or τ_2 flag on K_n. This causes some difficulties, as we need an optimal bound to be tight over all regular subgraphs. To be concrete: we will compare both sides by using the coefficients α_F over flags in $F \in \mathcal{F}_6^0$. Imagine we could show for all $F \in \mathcal{F}_6^0$ that $\alpha_{F,\mathrm{LHS}} \geq \alpha_{F,\mathrm{RHS}}$ in

$$\rho - \tfrac{4}{9} \geq [\![Q_1(\vec{f}^{\tau_1})]\!]_{\tau_1} + [\![Q_2(\vec{f}^{\tau_2})]\!]_{\tau_2}$$

But then observe that the coefficient for K_6 is positive ($\frac{5}{9}$) on the left hand side, but 0 on the right hand side. So $\phi(\mathrm{LHS} - \mathrm{RHS}) \geq \phi(\frac{5}{9}K_6) > 0$. So $\phi(\rho) > 4/9 + \phi(\mathrm{RHS}) \geq 4/9$, contradiction. The easiest way to remedy this is to multiply by a term like $1 - \rho$ which is 0 if and only if we have a complete graph. The multiplication was done over the type 1, using the flag e, a 3-edge with one vertex labelled. So, to try and prove $\rho - \frac{4}{9} \geq 0$, Razborov considered:

$$\begin{aligned}\rho - \tfrac{4}{9} = [\![e - \tfrac{4}{9}]\!]_1 &\geq [\![e - \tfrac{4}{9}]\!]_1 - \tfrac{9}{5}[\![(e - \tfrac{4}{9})^2]\!]_1 \\ &= \tfrac{9}{5}[\![(e - \tfrac{4}{9})(\tfrac{5}{9} - [e - \tfrac{4}{9}])]\!]_1 \\ &= \tfrac{9}{5}[\![(e - \tfrac{4}{9})(1 - e)]\!]_1\end{aligned}$$

And then, we can take our semi-definite problem to be to find $Q_i^{\pm}$ and $\vec{f}^{\tau_i,\pm} \in \Delta^{\tau_i,\pm}$ such that:

$$[\![(e - \tfrac{4}{9})(e - 1)]\!]_1 \geq [\![Q_1^+(\vec{f}^{\tau_1,+})]\!]_{\tau_1} + [\![Q_1^-(\vec{f}^{\tau_1,-})]\!]_{\tau_1} + [\![Q_2^+(\vec{f}^{\tau_2,+})]\!]_{\tau_2} + [\![Q_2^-(\vec{f}^{\tau_2,-})]\!]_{\tau_2}$$

A set of forms satisfying this is presented in the next section, and was derived by Razborov with the help of the *CSDP* package and *Maple*. The inequality is verified by comparing coefficients over the 34 flags in $\mathcal{F}_6^0$. The calculated coefficients are shown in [Raz10, Table 3], but they haven't been replicated in this essay as they are not particularly instructive.

Remark 27. This is equivalent to taking $\tau_3 = e$ and $Q_3(\vec{f}^{\tau_3}) = \frac{9}{5}(e - \frac{4}{9})^2$ in the general approach.

6.5 The solution to the SDP problem

Forms over τ_1 It turns out we only need the invariant part (ie, we take $Q_1^- = 0$). We represent Q_1^+ by the real positive definite matrix ($\lambda_{\min} = 0.01417\ldots$)

$$M_1 \stackrel{\text{def}}{=} \frac{1}{24}\begin{pmatrix} 32 & -4 & -22 & 25 \\ -4 & 16 & 1 & -13 \\ -22 & 1 & 17 & -17 \\ 25 & -13 & -17 & 27 \end{pmatrix}$$

The form acts on $\vec{f}^{\tau_1,+} = (f_1, f_2, f_3, f_4)$, where:

$$\begin{aligned}
f_1 &\stackrel{\text{def}}{=} (F_1^{\tau_1} + F_2^{\tau_1} + F_3^{\tau_1}) - (F_4^{\tau_1} + F_5^{\tau_1} + F_6^{\tau_1}) \\
f_2 &\stackrel{\text{def}}{=} (F_1^{\tau_1} + F_2^{\tau_1} + F_3^{\tau_1}) + 2(F_7^{\tau_1}) - (F_8^{\tau_1}) \\
f_3 &\stackrel{\text{def}}{=} (F_9^{\tau_1} + F_{12}^{\tau_1} + F_{13}^{\tau_1}) + (F_{10}^{\tau_1} + F_{11}^{\tau_1} + F_{14}^{\tau_1}) \\
f_4 &\stackrel{\text{def}}{=} (F_{15}^{\tau_1} + F_{16}^{\tau_1} + F_{17}^{\tau_1})
\end{aligned}$$

The representations of the relevant flags appear in tables 4 and 5. The individual invariant parts have been surrounded in brackets for clarity; it is clear that each $f_i \in \mathcal{A}^{\tau_1,+}$. They are also all in Δ^{τ_1}: $f_3^{\tau_1}$ and $f_4^{\tau_1}$ are trivially so, as they consist of non-regular flags. For $f_1^{\tau_1}$ and $f_2^{\tau_1}$, observe that $\phi(\llbracket f_i^{\tau_1} \rrbracket_{\tau_1}) = 0$ by summing the relevant entries from the third row of numbers in table 4.

Forms over τ_2 We need both parts this time. We take the anti-invariant form to simply be the quadratic $\frac{1}{2}g_0^2$, where g_0 is an element over anti-invariant flag relations,

$$g_0 \stackrel{\text{def}}{=} (-F_4^{\tau_2} + F_5^{\tau_2} + F_6^{\tau_2} - F_7^{\tau_2}) + 2(F_8^{\tau_2} - F_9^{\tau_2}) + 2(F_{10}^{\tau_2} - F_{11}^{\tau_2})$$

The flags over the same model have been bracketed. Razborov notes that g_0 does actually have an underlying structure: if we let e^* be the 2-flag on an edge on 3 vertices, and make use of the upwards operator briefly introduced in section 4.5 on page 16:

$$g_0 = \sum_{\substack{i \in \{1,2\} \\ j \in \{3,4\}}} (-1)^{i+j} \pi^{\tau_2, \eta:[1\mapsto i,\, 2\mapsto j]}(e^*)$$

We represent Q_2^+ by the real positive definite matrix ($\lambda_{\min} = 0.0325\ldots$)

$$M_2 \stackrel{\text{def}}{=} \frac{1}{24}\begin{pmatrix} 64 & -20 & 6 & -42 & 44 \\ -20 & 49 & 25 & -41 & -19 \\ 6 & 25 & 28 & -55 & -1 \\ -42 & -41 & -55 & 131 & -18 \\ 44 & -19 & -1 & -18 & 33 \end{pmatrix}$$

The form acts on $\vec{g}^{\tau_2,+} = (g_1, g_2, g_3, g_4, g_5)$, where:

$$\begin{aligned}
g_1 &\stackrel{\text{def}}{=} (F_1^{\tau_2}) - (F_2^{\tau_2}) \\
g_2 &\stackrel{\text{def}}{=} (F_1^{\tau_2}) - (F_3^{\tau_2}) \\
g_3 &\stackrel{\text{def}}{=} (F_4^{\tau_2} + F_5^{\tau_2} + F_6^{\tau_2} + F_7^{\tau_2}) \\
g_4 &\stackrel{\text{def}}{=} (F_8^{\tau_2} + F_9^{\tau_2}) \\
g_5 &\stackrel{\text{def}}{=} (F_{12}^{\tau_2} + F_{13}^{\tau_2})
\end{aligned}$$

The representations of the relevant flags appear in tables 6 and 7. Just as for τ_1, we have $\phi(\llbracket g_i \rrbracket_{\tau_2}) = 0$ (non-trivial for the regular flags in g_1, g_2), and $g_i \in \mathcal{A}^{\tau_2,+}$.

7 On the number of pentagons in triangle-free graphs

This essay finishes with an exposition of the 2013 paper by Hatami, Hladký, Král, Norine and Razborov [Hat+13]. It combines most of the content of this essay to show how the technique of flag algebras can be used to also prove finite results.

7.1 Introduction to the problem and overview of the method

Every bipartite graph is triangle free, but the converse clearly doesn't hold. In 1984, Erdős [Erd84] proposed 3 measures of the bipartiteness of a triangle-free graph, and conjectured they were all maximised by balanced blow-ups of the pentagon.

One such measure is the number of pentagons in the graph (observe here that all pentagons are necessarily induced in a triangle-free graph). Hatami et al showed in this paper that Erdős' hypothesis for this measure is correct. That is, that the density of pentagons in any triangle free graph on n vertices is at most $(n/5)^5$, and that this is achieved uniquely by balanced blow-ups of the pentagon. They do this via the following results (where henceforth we use the theory of triangle-free graphs, $T_{\text{TF-Graph}}$, implicitly):

i) Firstly, they show using SDP methods that, for the 5-cycle/pentagon $C_5 \in \mathcal{G}_5$,

$$C_5 \leq_0 \frac{5!}{5^5}$$

ii) Secondly, they show that the extremal homomorphism achieving this bound is shown to be uniquely given by the balanced blow-up of the pentagon. This is shown with an in-depth flag algebra argument.

iii) By appealing to blow-up constructions, they conclude that every triangle free graph on n vertices contains at most $(n/5)^5$ pentagons (that is, has pentagon density at most $\frac{n^5}{n\cdots(n-4)} \cdot \frac{5!}{5^5}$), and that the bound is only achieved for balanced blow-ups of the pentagon.

The rest of the paper is devoted to showing that if $5 \not| n$ and n is sufficiently large, then the maximal number of pentagons is the same as in an almost-balanced blow-up of the pentagon, and moreover that this is the unique construction achieving that bound. This utilises a theorem concerning convergence in the cut distance (as discussed in Razborov's remark at the start of section 5.3 on page 27). But the argument employed isn't particularly instructive, and doesn't involve flag algebras, so I have decided to leave it out of the essay.

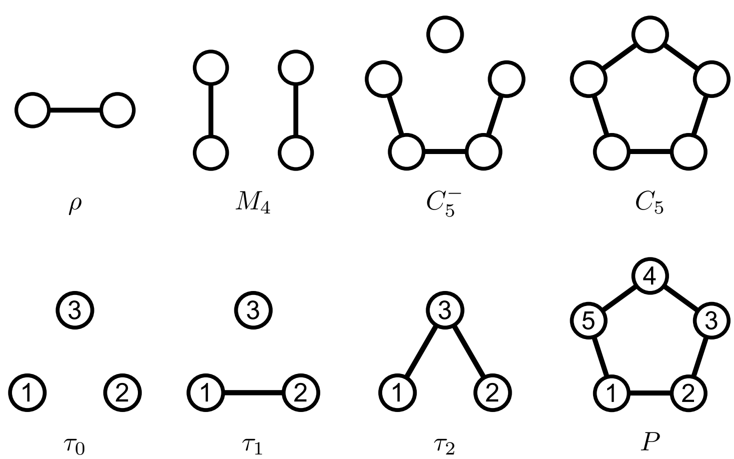


Figure 8: Relevant models and types in the paper.

7.2 Flag notation

Let σ be a type of size k. For $V \subset [k]$, we define F_V^σ to be the σ-flag on $|\sigma|+1$ vertices, with the unlabelled vertex connected to the vertices with labels $v \in V$. Note that, in order for F_V^σ to be triangle free, $\sigma|_V$ must be an independent set. So we introduce the notation:

$$f_l^\sigma \overset{\text{def}}{=} \sum \{F_V^\sigma \text{ such that } |V| = l \text{ and } \sigma|_V \text{ is independent}\}$$

Now, we observe the types τ_i and their symmetries from figure 8.

For τ_0 we have $\Gamma_{\tau_0} = S_{\{1,2,3\}}$, and so get the following invariant flag relations:

$$\begin{aligned} f_0^{\tau_0} &= F_\varnothing^{\tau_0} \\ f_1^{\tau_0} &= F_{\{1\}}^{\tau_0} + F_{\{2\}}^{\tau_0} + F_{\{3\}}^{\tau_0} \\ f_2^{\tau_0} &= F_{\{2,3\}}^{\tau_0} + F_{\{1,3\}}^{\tau_0} + F_{\{1,2\}}^{\tau_0} \\ f_3^{\tau_0} &= F_{\{1,2,3\}}^{\tau_0} \end{aligned}$$

For τ_1 we have $\Gamma_{\tau_1} = S_{\{1,2\}}$, and so get the following invariant flag relations:

$$f_0^{\tau_1} = F_\varnothing^{\tau_1}, \quad F_{\{1\}}^{\tau_1} + F_{\{2\}}^{\tau_1}, \quad F_{\{3\}}^{\tau_1}, \quad f_2^{\tau_1} = F_{\{1,3\}}^{\tau_1} + F_{\{2,3\}}^{\tau_1}$$

For τ_2 we also have $\Gamma_{\tau_2} = S_{\{1,2\}}$, giving the following invariant flag relations:

$$f_0^{\tau_2} = F_\varnothing^{\tau_2}, \quad F_{\{1\}}^{\tau_2} + F_{\{2\}}^{\tau_2}, \quad F_{\{3\}}^{\tau_2}, \quad f_2^{\tau_2} = F_{\{1,2\}}^{\tau_2}$$

7.3 The SDP solution

As the relevant theory and demonstration has already been covered in sections 5 and 6, I will just present the solution here, straight from [Hat+13]. By comparing over coefficients over the 14 admissible graphs in $\mathcal{F}_5^0[T_{\text{TF-graph}}]$, the following inequality can be verified:

$$\begin{aligned} 62500\, C_5 + \tfrac{1097}{12} M_4 + \tfrac{68}{3} C_5^- + 200\left(\rho - \tfrac{2}{5}\right)^2 + [\![Q_0^+(\vec{g}_0^{\,+})]\!]_{\tau_0} + \\ [\![Q_1^+(\vec{g}_1^{\,+})]\!]_{\tau_1} + [\![Q_2^+(\vec{g}_2^{\,+})]\!]_{\tau_2} + [\![Q_1^-(\vec{g}_1^{\,-})]\!]_{\tau_1} + [\![Q_2^-(\vec{g}_2^{\,-})]\!]_{\tau_2} \end{aligned} \leq_0 2400 \qquad (7.3.1,\ C_5 \text{ inequality})$$

Where:

$$\begin{aligned} \vec{g}_0^{\,+} &\overset{\text{def}}{=} (\, f_1^{\tau_0} - f_2^{\tau_0}, \quad -2f_0^{\tau_0} + f_2^{\tau_0} + 3f_3^{\tau_0} \,) \\ \vec{g}_1^{\,+} &\overset{\text{def}}{=} (\, 2f_0^{\tau_1} - f_1^{\tau_1}, \quad f_1^{\tau_1} + f_2^{\tau_1}, \quad F_{\{3\}}^{\tau_1} \,) \\ \vec{g}_2^{\,+} &\overset{\text{def}}{=} (\, 6f_0^{\tau_2} + f_1^{\tau_2} - 4f_2^{\tau_2 2}, \quad 2f_0^{\tau_2} - 2f_2^{\tau_2} + F_{\{3\}}^{\tau_2} \,) \\ \vec{g}_1^{\,-} &\overset{\text{def}}{=} (\, F_{\{1\}}^{\tau_1} - F_{\{2\}}^{\tau_1}, \quad F_{\{2,3\}}^{\tau_1} - F_{\{1,3\}}^{\tau_1} \,) \\ \vec{g}_2^{\,-} &\overset{\text{def}}{=} (\, F_{\{1\}}^{\tau_2} - F_{\{2\}}^{\tau_2} \,) \end{aligned}$$

With $Q_i^\pm$ denoted by the matrices:

$$M_0^+ \overset{\text{def}}{=} \begin{pmatrix} 9760 & 2252 \\ 2252 & 592 \end{pmatrix} \quad M_1^+ \overset{\text{def}}{=} \begin{pmatrix} 13900 & -671 & -12807 \\ -671 & 31334 & -511136 \\ -12807 & -51136 & 98157 \end{pmatrix} \quad M_2^+ \overset{\text{def}}{=} \begin{pmatrix} 22708 & -40788 \\ -40788 & 78132 \end{pmatrix}$$

$$M_1^- \overset{\text{def}}{=} \begin{pmatrix} 1416 & -16452 \\ -16452 & 256488 \end{pmatrix} \qquad M_2^- \overset{\text{def}}{=} \begin{pmatrix} 158266 \end{pmatrix}$$

7.4 Proving the uniqueness of the extremal homomorphism

Note: The results in this section are from [Hat+13]. The proof of corollary 16 is my own, as it was stated without proof in the paper.

The theorem that forms the backbone of the the paper is the following.

Theorem 15. *The homomorphism ϕ_{C_5} is the unique $\phi \in \mathrm{Hom}^+(\mathcal{A}^0[T_{TF\text{-}Graph}], \mathbb{R})$ that satisfies:*

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

To demonstrate its relevance, I will quickly prove the following corollary, before giving the proof of theorem 15 itself.

Corollary 16. *Every n-vertex triangle-free graph G contains at most $(n/5)^5$ pentagons. Moreover, the equality is attained only when n is divisible by five and G is the balanced blow-up of the pentagon.*

Proof of Corollary 16. Assume a graph G on n vertices contains $\geq (n/5)^5$ pentagons, ie $p(C_5, G) \geq (n/5)^5/\binom{n}{5}$. Then, as C_5 is twin-free, we have by (5.3.2, Formula for $\phi_G(H)$ with H twin-free) from section 5.3.1 on blow-ups, that:

$$\phi_G(C_5) = \frac{n(n-1)\cdots(n-4)}{n^5} p(C_5, G) \geq \frac{n!}{n^5(n-5)!} \cdot \frac{(n/5)^5}{\binom{n}{5}} = \frac{5!}{5^5}$$

But we know from our SDP calculation, (7.3.1, C_5 inequality), that $C_5 \leq_0 5!/5^5$, hence $\phi_G(C_5) > 5!/5^5$ is not possible. So we conclude that $\phi_G(C_5) = 5!/5^5$, and hence, by theorem 15, $\phi_G = \phi_{C_5}$. Then, we use corollary 14, also from section 5.3.1, which tells us that G and C_5 are balanced blow-ups of a mutual graph. Noting that C_5 is only a balanced blow-up of itself gives the result. □

We now proceed with the proof of theorem 15. It is a rather long proof, though not too difficult, and it combines many aspects of flag algebras introduced in this essay. The core of the proof is due to Hatami et al, but I have filled in some gaps, corrected some minor errors, re-structured some parts of the proof, and added some comments to aid understanding. Before starting, the reader is recommended to briefly recap the definitions of the upwards operator (section 4.5, page 16), and the extension measure (section 4.10, page 21).

Proof of Theorem 15. Fix $\phi \in \mathrm{Hom}^+(\mathcal{A}^0[T_{\text{TF-Graph}}], \mathbb{R})$ such that $\phi(C_5) = 5!/5^5$. We will use the type P of the 5-labelled pentagon (figure 8), and consider the extension measure $\mathbb{P}^P \stackrel{\text{def}}{=} \mathbb{P}^{P,0}$ of ϕ. This satisfies:

$$\int_{\mathrm{Hom}^+(\mathcal{A}^P[T_{\text{TF-Graph}}],\mathbb{R})} \psi(f)\mathbb{P}^P(d\psi) = \frac{\phi([\![f]\!]_P)}{\phi([\![1^P]\!]_P)} \qquad \text{(7.4.1, Extension Measure)}$$

In particular, let $S^P(\phi)$ denote the support of this integration, ie the smallest closed set $A \subset \mathrm{Hom}^+(\mathcal{A}^P[T_{\text{TF-Graph}}], \mathbb{R})$ such that $\mathbb{P}^P(A) = 1$.

Proof overview We will consider some $\phi^P \in S^\sigma(\phi)$ arbitrary, ie some ϕ^P extending ϕ. In particular, we will use that $\phi^P(\pi^P(f)) = \phi(f)$. The result follows from showing that, for any triangle-free graph H, $\phi^P(\pi^P(H))$ satisfies (5.3.1, Formula for ϕ_G) for $\phi_{C_5}(H)$ from section 5.3.1 on blow-ups.

Initially the proof extracts results over ϕ from our SDP inequality, and transfers them to ϕ^P. We then discover as much as we can about the densities in ϕ^P, firstly by adding one vertex to P, and then two vertices. We use the larger 2-balanced blow-up $C_5^{(2)}$ to help us with determining these values.

Finally, we can use the possible realized P-flags with 2 unlabelled vertices to find $\phi^P(H)$ for general graphs H, relating it to strong homomorphisms $H \to C_5$, to show $\phi(H)$ equals $\phi_{C_5}(H)$.

We proceed now with the proof. The form of SDP inequality (7.3.1, C_5 inequality) is important to this argument. To get a tight bound, $C_5 = 2400/62500 = 5!/5^5$ requires that $\phi(M_4) = \phi(C_5^-) = 0$. Now we observe that C_5^- occurs with positive density in $[\![F_\varnothing^P]\!]_P$ (C_5 with an additional isolated vertex), and $[\![F_{\{i\}}^P]\!]_P$ (C_5 with another vertex connected to exactly one corner of the pentagon) for each $i \in [5]$, hence $\phi(C_5^-) = 0$ implies that $\phi([\![F_\varnothing^P]\!]_P) = \phi([\![F_{\{i\}}^P]\!]_P) = 0$. This extends to the extension measure:

Claim. *If $\psi \in S^P(\phi)$, then $\psi(F_\varnothing^P) = \psi(F_{\{i\}}^P) = 0$ for all $i \in [5]$.*

Proof. Let F be any of $F_\varnothing^P$ or $F_{\{i\}}^P$ for $i \in [5]$. So by the above, $\phi([\![F]\!]_P) = 0$. Hence, by (7.4.1, Extension Measure),

$$\int\limits_{S^P(\phi)} \psi(F)\mathbb{P}^P(d\psi) = 0$$

Now F is a flag, hence $\psi(F) \geq 0$ for $\psi \in S^P(\phi)$. Define $Y \subset \mathrm{Hom}^+(\mathcal{A}^P[T_{\text{TF-Graph}}], \mathbb{R})$ as those ψ having $\psi(F) > 0$. Then we must have $1 = \mathbb{P}^P\left[\{\psi : \psi(F) = 0\}\right] = \mathbb{P}^P\left[S^P(\phi) \setminus Y\right]$. Y is clearly open, so $S^P(\phi) \setminus Y$ is closed, but $S^P(\phi)$ was defined as the smallest closed set on which $\mathbb{P}^P\left[S^P(\phi)\right] = 1$, so $S^P(\phi) = S^P(\phi) \setminus Y$, hence $S^P(\phi) \cap Y = \varnothing$. That is, $\psi(F) = 0$ for all $\psi \in S^P(\phi)$. ■

Now consider an arbitrary extension $\phi^P \in S^P(\phi)$ of ϕ. Define a P-flag F to be *realized* if $\phi^P(F) > 0$. To start, we consider which $F \in \mathcal{F}_6^P$ are realized. That is, which $V \subset [5]$ could give $\phi^P(F_V^P) > 0$? Well, V must label independent vertices in P, else F_V^P is not triangle free. Also, $|V| > 1$ by the claim. So, denoting $H_i^P \stackrel{\text{def}}{=} F_{\{i-1,i+1\}}^P$ (where the addition is done mod 5), these are the only flags $F \in \mathcal{F}_6^P$ which could be realized. Hence we get:

$$1 = \phi^P(1^P) = \sum_{F \in F_6^P} \phi^P(F) = \sum_{i=1}^{5} \phi^P(H_i^P)$$

So, by the AM-GM inequality, with equality if and only if $\phi^P(H_i^P) = \frac{1}{5}$ for all i, we have:

$$\prod_{i \in \mathbb{Z}_5} \phi^P(H_i^P) \leq \left(\frac{\sum_{i=1}^5 \phi^P(H_i^P)}{5}\right)^5 = 5^{-5} \tag{7.4.2}$$

Now we recall the upwards operator from section 4.5. Also recall the result of lemma 11, page 22 that, applied in this case, says that, if $\phi^\sigma \in S^\sigma(\phi)$, and $f \in \mathcal{A}^0$, then $\phi^\sigma(\pi^\sigma(f)) = \phi(f)$.

In particular, $\phi(C_5) = \phi^P(\pi^P(C_5))$. So what is $\pi^P(C_5)$? It is simply the sum of all P-flags (up to isomorphism) whose unlabelled vertices form a C_5. So which of those flags F could have non-zero $\phi^P(F)$? Well, for each vertex v_i (for $i \in \mathbb{Z}_5$) in the unlabelled C_5 part of F, we must have $F|_{\{v_i\} \cup \mathrm{im}(\theta)} \cong H_j^P$ (some j). Represent the isomorphic pairs in the map $\alpha : i \mapsto j$. By considering possibilities for α which keep F triangle-free, we have that (if the v_i are labelled cyclically around C_5), then $\alpha : i \mapsto k \pm i \pmod 5$ for some fixed choice on $\pm$ and $k \in \mathbb{Z}_5$. In other words, $F \cong (C_5^{(2)})^P$, the (only possible) P flag on the 2-balanced blow-up of C_5, and so $\phi^P(\pi^P(C_5)) = \phi^P(F) = \phi^P((C_5^{(2)})^P)$.

Now, consider multiplying all the H_i^P together as flags, then:

$$\prod_{i \in \mathbb{Z}_5} H_i^P = \sum_{F \in \mathcal{F}_{10}^P} p(H_0^P, \cdots, H_4^P; F)F \;\; \geq_P \;\; p(H_0^P, \cdots, H_4^P; (C_5^{(2)})^P)(C_5^{(2)})^P = \tfrac{1}{5!}(C_5^{(2)})^P \tag{7.4.3}$$

Hence, combining (7.4.2), the comment $\phi^P(\pi^P(C_5)) = \phi^P(F) = \phi^P((C_5^{(2)})^P)$ and (7.4.3) gives:

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

So this must be tight. In particular, $\phi^P((C_5^{(2)})^P) = \frac{5!}{5^5}$, and by (7.4.2), $\phi^P(H_i^P) = \frac{1}{5}$ for all i.

Define H_{ij}^P as P with the addition of two independent unlabelled vertices, the first connected to $\theta(i-1)$ and $\theta(i+1)$, and the second connected to $\theta(j-1)$ and $\theta(j+1)$. Define G_{ij}^P similarly, but with an edge between the unlabelled vertices.

Claim. *The only realized $F \in \mathcal{F}_7^P$ are $F = G_{ij}^P$ for $(i,j) \in E(P)$, or $F = H_{ij}^P$ for $(i,j) \notin E(P)$. Moreover, $\phi^P(G_{ij}^P) = \frac{2}{25}$, $\phi^P(H_{ii}^P) = \frac{1}{25}$ and $\phi^P(H_{ij}^P) = \frac{2}{25}$ for $i \neq j$.*

Proof of Claim. First, we note that, due to the H_i being the only realized flags on 6 vertices, the only possible realized flags on 7 vertices are H_{ij}^P and G_{ij}^P.

Case 1. Take $(i,j) \in E(P)$.

In this case, we have that $H_i^P \cdot H_j^P = \frac{1}{2}H_{ij}^P + \frac{1}{2}G_{ij}^P$, so:

$$\prod_{k\in\mathbb{Z}_5} H_k^P = (\tfrac{1}{2}G_{ij}^P + \tfrac{1}{2}H_{ij}^P) \cdot \prod_{k\in\mathbb{Z}_5\setminus\{i,j\}} H_k^P \geq_P \tfrac{1}{5!}(C_5^{(2)})^P + \tfrac{1}{2}H_{ij}^P \cdot \prod_{k\in\mathbb{Z}_5\setminus\{i,j\}} H_k^P \tag{7.4.4}$$

Where the inequality follows in a similar manner to (7.4.3), by noting that

$$p(G_{12}^P, H_3^P, H_4^P, H_5^P; (C_5^{(2)})^P) = \tfrac{2}{5!}$$

Now, we apply ϕ^P to (7.4.4), and use that $\phi^P((C_5^{(2)})^P) = \frac{5!}{5^5}$, and that $\phi^P(H_P^k) = \frac{1}{5}$ for all k. We conclude that $\phi^P(H_{ij}^P) = \frac{1}{25}$, and that $\phi^P(G_{ij}^P) = \frac{2}{25}$.

Case 2. Take $(i,j) \notin E(P)$.

This case is easier: we must have either $i = j$, or $i - j \equiv \pm 2 \pmod 5$. Either way, observe that G_{ij}^P is not triangle-free. If $i = j$, then $H_i^P \cdot H_j^P = H_{ii}^P$, hence $\phi^P(H_{ii}^P) = 0$. If $i \neq j$, then $H_i^P \cdot H_j^P = \frac{1}{2}H_{ij}^P$, so $\phi^P(H_{ij}^P) = \frac{2}{25}$.

■

Now, for a triangle-free graph H, and a strong homomorphism $\alpha : H \to C_5$, define F_α^P as the following P-flag: Take the unlabelled vertices to be an induced H, and join each vertex $v \in H$ to the labelled vertices $\alpha(v) + 1$ and $\alpha(v) - 1$ (performing the addition cyclically around the pentagon).

Claim. *For any triangle-free graph H, define $\mathcal{S} \subset \mathcal{F}_{5+|H|}^P$ to be*

$$\pi^P(H) = \sum_{F\in\mathcal{S}} F$$

Then each $F \in \mathcal{S}$ is realized if and only if $F \cong F_\alpha^P$ for some $\alpha : H \to C_5$ a strong homomorphism. Moreover,

$$\phi^P(F_\alpha^P) = 5^{-|H|}c_\alpha \quad \textit{where} \quad c_\alpha \stackrel{\text{def}}{=} \binom{|H|}{|\alpha^{-1}(1)|\ |\alpha^{-1}(2)|\ |\alpha^{-1}(3)|\ |\alpha^{-1}(4)|\ |\alpha^{-1}(5)|}$$

Proof of Claim. Let $F = (M, \theta) \in \mathcal{S}$ have $\phi^P(F) > 0$ (that is, let F be realized), and take two distinct unlabelled vertices v, w in F. Then $F|_{\mathrm{im}(\theta)\cup\{v,w\}}$ is also realized. So, by the previous claim, $(v,w) \in E(F)$ if and only if $(\alpha(v), \alpha(w)) \in E(C_5)$. This gives that $F \cong F_\alpha^P$ for some $\alpha : H \to C_5$ a strong homomorphism. Now, consider:

$$\prod_{v\in V(H)} H_{\alpha(v)}^P = \sum_{F\in\mathcal{F}_{5+|H|}^P} p(H_{\alpha(v_1)}^P, \cdots, H_{\alpha(v_{|H|})}^P; F)F$$

Which of these F are realized? In fact, only F_α^P is: the presence of an edge in F between v and w is exactly determined by the presence of an edge between $\alpha(v)$ and $\alpha(w)$, as above. Also, note that $p(H_{\alpha(v_1)}^P, \cdots, H_{\alpha(v_{|H|})}^P; F_\alpha^P) = c_\alpha^{-1}$. So, applying ϕ^P, and using $\phi(H_i^P) = \frac{1}{5}$ gives:

$$\phi^P(F_\alpha^P) = 5^{-|H|}c_\alpha$$

■

Observe that $F^P_\alpha \cong F^P_{\alpha'}$ is possible. Define an equivalence relation $\sim$ over $\alpha \in \mathrm{Hom}_{\mathrm{str}}(H, C_5)$ by $\alpha \sim \alpha' \iff F^P_\alpha \cong F^P_{\alpha'}$. Denote an equivalence class by $[\alpha]$, and its size by $\big|[\alpha]\big|$.

Claim.

$$\big|[\alpha]\big| = \frac{c_\alpha \cdot |\mathrm{Aut}(H)|}{|H|!}$$

Proof of Claim. Rewriting the multinomial coefficient c_α in terms of factorials turns the required form into:

$$\big|[\alpha]\big| = \frac{|\mathrm{Aut}(H)|}{|\alpha^{-1}(1)|!\,|\alpha^{-1}(2)|!\cdots|\alpha^{-5}(1)|!} \tag{7.4.5}$$

Now, consider the action of $\mathrm{Aut}(H)$ on $[\alpha]$ given by $\gamma(\beta) = \beta \circ \gamma$. Observe that the orbit of α is the whole of $[\alpha]$. The stabiliser of α is $S(\alpha) = S_{\alpha^{-1}(1)} \times \cdots \times S_{\alpha^{-1}(5)}$. So, (7.4.5) is simply a statement of the orbit-stabilizer theorem. ■

Hence, taking the sum over equivalence classes $[\alpha]$ in $\mathrm{Hom}_{\mathrm{str}}(H, C_5)/\sim$, we get, for an arbitrary triangle-free graph H:

$$\begin{aligned}
\phi(H) &= \phi^P(\pi^P(H)) && \text{(By lemma 11, page 22)}\\
&= \sum_{F\in\mathcal{S}} \phi^P(F) && \text{(By the definition of } \mathcal{S})\\
&= \sum_{[\alpha]} \phi^P(F^P_\alpha) = 5^{-|H|}\sum_{[\alpha]} c_\alpha && \text{(By the second-last claim)}\\
&= \frac{|H|!}{|\mathrm{Aut}(H)|}\cdot 5^{-|H|}\cdot\sum_{[\alpha]}|[\alpha]| && \text{(By the last claim)}\\
&= \frac{|H|!}{|\mathrm{Aut}(H)|}\cdot\frac{s(H, C_5)}{5^{|H|}}\\
&= \phi_{C_5}(H) && \text{(By (5.3.1, Formula for } \phi_G\text{), page 27)}
\end{aligned}$$

□

8 Conclusion

In this essay, I have addressed the common theory and typical use cases associated with flag algebras. The semi-definite programming approach not only helps to prove asymptotic results, but the form of bounds acquired helps nail down which other models are disallowed in an extremal construction, when tightness of the inequality is required. But even if tight results cannot be proven in a given case, the systematic approach can still be applied to improve inexact bounds. To quote Razbarov [Raz13a] “let me cautiously suggest that I am not aware of a *single* example of a *non-exact* bound in asymptotic extremal combinatorics that could not be improved by a plain application of flag algebras.” For the interested reader, results over recent years have been compiled into a flag algebras survey by Razborov in [Raz13a], and some important SDP results are listed on the Flagmatic website [Fla].

There has also been some research into the limits of the associated methodology, in particular, into the limits of the SDP method, the so called “Cauchy-Schwarz Calculus” [Raz07].

Practical comments, on the feasibility of SDP calculations over models on vertex sets larger than 7 are highlighted in [FRV13], and they also discuss that, particularly for hypergraphs, there are often many families of extremal homorphisms for a given problem. An exact bound using flag algebras would need to be tight on all of them, which becomes increasingly improbable as the number of extremal asymptotic constructions grows.

This concern was also made rigorous, in the sense that, even for 2-graphs, the problem of determining if an arbitrary $f \in \mathcal{A}^\sigma$ has $f \geq_\sigma 0$ using arguments over a finite number of flags is undecidable [HN10], as are many other common theories [Raz13a].

Still, the formalization that the theory of flag algebras offers seems to have worth in many places. In particular, it has uses beyond its plain SDP applications, as demonstrated in the proof of theorem 15, and in many other papers.

It should also be mentioned that formal methods similar to flag algebras (which build a mathematical calculus over objects of study) have been extended successfully to other areas. For example, the ideas been extended to prove results about sparse graph limits (as opposed to dense graph limits), in the particular case of subgraphs of the hypercube Q_n [Bab11, Chapter 2.5].

9 Acknowledgements

I am grateful to Prof Andrew Thomason for introducing me to this very interesting topic. Further, I am very appreciative of his feedback with regards to the planning of this essay.

10 References

[Bab11] Rahil Baber. "Some Results in Extremal Combinatorics". PhD thesis. Department of Mathematics, UCL, Mar. 2011.

[Cum+13] James Cummings et al. "Monochromatic Triangles in three-coloured graphs". In: *Journal of Combinatorial Theory, Series B* 103.4 (2013), pp. 489–503.

[Erd84] Paul Erdős. "On some problems in graph theory, combinatorial analysis and combinatorial number theory". In: *Graph Theory and Combinatorics, Cambridge, 1983*. Academic Press, London, 1984, pp. 1–17.

[Fla] *Flagmatic*. http://flagmatic.org/. Accessed: 15-04-2015.

[FRV13] Victor Falgas-Ravry and Emil R. Vaughan. "Applications of the semi-definite method to the Turán density problem for 3-graphs". In: *Combinatorics, Probability and Computing* 22.1 (2013), pp. 21–54.

[Hat+13] Hamed Hatami et al. "On the number of pentagons in triangle-free graphs". In: *Journal of Combinatorial Theory, Series A* 120.3 (Apr. 2013), pp. 722–732.

[Hir13] James Hirst. "The Inducibility of Graphs on Four Vertices". In: *Journal of Graph Theory* 75.3 (2013), pp. 231–243.

[HN10] Hamed Hatami and Serguei Norine. "Undecidability of linear inequalities in graph homomorphism densities". In: *Journal of the American Mathematical Society* 24.2 (2010), pp. 547–565.

[Lov12] László Lovász. *Large Networks and Graph Limits*. Vol. 60. Colloquiam Publications. American Mathematical Society, 2012.

[Mar02] David Marker. *Model Theory: An Introduction*. Vol. 217. Graduate Texts in Mathematics. Springer, 2002.

[Raz07] Alexander Razborov. "Flag Algebras". In: *Journal of Symbolic Logic* 72 (2007), 1239–1282.

[Raz10] Alexander Razborov. "On 3-hypergraphs with forbidden 4-vertex configurations". In: *SIAM Journal on Discrete Mathematics* 24.3 (2010), pp. 946–963.

[Raz13a] Alexander Razborov. "Flag Algebras: An Interim Report". In: *The Mathematics of Paul Erdős II*. Ed. by Ronald L. Graham, Jaroslav Nešetřil, and Steve Butler. Springer New York, 2013, pp. 207–232.

[Raz13b] Alexander Razborov. "What is... a Flag Algebra?" In: *Notices of the AMS* (Nov. 2013), pp. 1324–1327.