The Journal of Symbolic Logic
Volume 00, Number 0, XXX 0000

FLAG ALGEBRAS

ALEXANDER A. RAZBOROV

Abstract. Asymptotic extremal combinatorics deals with questions that in the language of model theory can be re-stated as follows. For finite models M, N of an universal theory without constants and function symbols (like graphs, digraphs or hypergraphs), let $p(M, N)$ be the probability that a randomly chosen sub-model of N with $|M|$ elements is isomorphic to M. Which asymptotic relations exist between the quantities $p(M_1, N), \ldots, p(M_h, N)$, where $M_1, \ldots, M_h$ are fixed "template" models and $|N|$ grows to infinity?

In this paper we develop a formal calculus that captures many standard arguments in the area, both previously known and apparently new. We give the first application of this formalism by presenting a new simple proof of a result by Fisher about the minimal possible density of triangles in a graph with given edge density.

§1. Introduction. A substantial part of modern extremal combinatorics (which will be called here *asymptotic* extremal combinatorics) studies *densities* with which some "template" combinatorial structures may or may not appear in unknown (large) structures of the same type[1]. As a typical example, let G_n be a (simple, non-oriented) graph on n vertices with m edges and t triangles. Then the edge density of G_n is $\rho(G_n) \stackrel{\mathrm{def}}{=} \frac{m}{\binom{n}{2}}$, its triangle density is $\mu(G_n) \stackrel{\mathrm{def}}{=} \frac{t}{\binom{n}{3}}$, and one can ask which pairs $(\rho, \mu) \in [0, 1]^2$ can be "asymptotically realized" (we will make this precise later) as $(\rho(G_n), \mu(G_n))$.

In the language of finite model theory, problems of this type can be formulated as follows. Let T be a universal theory in a first-order language without constants or function symbols. Then every set of elements of a model of T again induces a model of T. For two finite models M and N of T with $|M| < |N|$, let $p(M, N)$ be the density with which M appears as a sub-model of N. Fix finite models $M_1, \ldots, M_h$, and let the size of N grow to infinity. Which relations between the densities $p(M_1, N), \ldots, p(M_h, N)$ will then necessarily hold "in the limit"? The list of theories of interest includes theories of graphs, digraphs, tournaments,

Supported by the Charles Simonyi Endowment, NSF grant ITR-0324906 and Russian Foundation for Basic Research.

[1] One example of a famous open problem that does *not* fall into this category, and that does *not* seem to be amenable to our techniques, is the problem of estimating the maximal number of edges in C_4-free graphs. The reason is that this density is known to be asymptotically 0, so the question is actually about low-order terms.

0022-4812/00/0000-0000/$00.00

hypergraphs etc. In each case we can also append additional axioms forbidding certain structures, thus obtaining e.g. the theory of K_ℓ-free graphs (corresponding to the classical Turán theorem [Tur]), the theory of oriented graphs of girth $\geq g$ (featured in the Caccetta-Häggkvist conjecture [CaH]) or the theory of 3-hypergraphs in which every four vertices span at least one hyperedge (pertaining to the Turán hypergraph problem that is *the* most famous open problem in the whole area).

Asymptotic extremal combinatorics is full of ingenious and hard results. Surprisingly, however, there seem to exist only a handful of essentially different techniques in this area, and the difficulty (as well as beauty) lies in finding the "right" relations using instruments from this relatively small toolbox and combining them in the "right" way. And after trying this for a while, it very soon becomes clear that there is a rich algebraic structure underlying many of these techniques, and especially those that, besides induction, involve a non-negligible amount of counting. It is also more or less transparent that they can be arranged in the form of a formal calculus based on simply defined algebraic objects (that we will call *flag algebras*) associated with the theory in question.

We have found it extremely instructive to distill this "assumed" calculus in its pure form, and this is exactly what we attempt to do in this paper. Our arguments in favour of such a formalization (as opposed to "naive" exact calculations) are at least three-fold.

- In asymptotic extremal combinatorics lower-order terms that supposedly do not influence the final result are particularly annoying and in many cases bury the essence of the argument under technicalities. In fact, many authors give up and declare that they will ignore such terms from the outset (see e.g. one of the most interesting recent developments [CaF]). The danger with this radical approach, however, lurks in the proofs containing various inductive arguments. The more elaborate these arguments are, the more likely it becomes that low-order error terms will eventually accumulate interfering with the final asymptotic result. In our framework low-order terms do not exist in principle, and in this sense it is somewhat reminiscent of non-standard analysis (but, unlike the latter, we use only pretty standard mathematical concepts like commutative algebras, homomorphisms and some very basic notions of the measure theory). The statements that are responsible for taming error terms are proved as general facts once and for all. After this (admittedly, no less tedious than its naive counterparts) work is done, we simply use the results at yet further axioms or inference rules appended to our calculus. The best manifestation of this idea given in our paper is the differential structure explored in Section 4.3.
- The algebraic, topological and probabilistic structure introduced for our pragmatic purposes looks very much like the structure existing elsewhere in mathematics. This in particular allows us to draw upon "foundational" results from its different areas, sometimes quite deep (like Prohorov's theorem on the weak convergence of probability measures). It is also conceivable that on this way we will be able to draw upon *concrete calculations* performed in other areas for different purposes. We also feel that revealing

underlying structures of this kind should be especially beneficial for an area that, alas, is still sometimes viewed by some as somewhat isolated from "mainstream" mathematics.

- We put different techniques in extremal combinatorics to a common denominator by viewing them as linear operators (and often – algebra homomorphisms) acting between the same objects, flag algebras. This becomes extremely useful when trying to combine these techniques together. Also, this representation is very structured and thus very convenient to program, so that the search for "right" relations can be to a large degree computerized. Indeed, for the concrete proof included in this paper (Section 5) as well as for all other applications (to be given elsewhere) we have extensively used Maple and the CSDP package for semi-definite programming [Bor] to test various hypothesis and avoid hopeless directions.

There were, of course, previous attempts to achieve the goals we are pursuing here. Some underlying ideas can be traced back to [GGW], and in the context of the Caccetta-Häggkvist conjecture we would like to mention the paper [Bon]. However, the only *systematic* attempt we are aware of is the research on *graph homomorphisms* conducted by Lovász et. al. (see [LoSz] for the paper most closely related to our purposes, and [BCL*] for a general survey). Graph homomorphisms is a very promising direction that connects many different areas of mathematics, physics and theoretical computer science and draws upon them for motivation, research goals, ideas etc. It has already led to an extremely interesting body of results and connections, ranging from statistical physics to property testing. However, precisely due to its universality, when it comes to concrete implementations in extremal combinatorics, this approach suffers from certain deficiencies that, hopefully, are taken care of in our more focused treatment.

- Currently, the research on graph homomorphisms does not seem to incorporate arguments involving any kind of induction. Illustrations of its usefulness in extremal combinatorics consist so far of re-proving a few simple classical results, and, moreover, only those whose proofs are based entirely on the Cauchy-Schwarz inequality.
- The notion of a graph homomorphism appears to be more alien to extremal combinatorics than that of an induced substructure. First, it immediately restricts applicability of the theory to those structures for which this notion makes sense. But even for graphs, the asymmetry inherently contained in the notion of a homomorphism makes many standard arguments look rather unnatural. For example, suppose that we want to count the number of independent subgraphs on three vertices by averaging over all non-edges of a graph. Theoretically it is doable with graph homomorphisms, but the result would not look the way we would like it to.
- The primary semantic model in the theory of graph homomorphisms is made by (measurable weighted) graphs on an infinite measure space. This semantics is absolutely perfect for every conceivable purpose (in particular, all information contained in semantic models considered in our paper can be naturally retrieved from such graphs). But the problem with this semantics is that the proof of the corresponding "completeness theorem" is based on

> Szemerédi's Regularity Lemma and fells apart already for 3-hypergraphs. This is certainly not the feature we want in our theory, and we replace measurable graphs with "ensembles" of probability measures on homomorphisms from flag algebras to the reals. On the one hand, the latter objects still carry (apparently) all the information needed for the purposes of extremal combinatorics. On the other hand, now the "completeness theorem" is based on Prohorov's theorem (about weak convergence of probability measures) and can be applied to arbitrary combinatorial structures.

The paper is organized as follows. In Section 2 we develop the "syntactic" part of our calculus; we define flag algebras, and introduce linear operators (averaging, homomorphisms onto sub-algebras of constants, more general interpretation-based operations) that will serve as basic "inference rules". In Section 3 we consider various semantics for our calculus (homomorphisms into the reals, convergent sequences of finite models, ensembles of random homomorphisms), prove their equivalence, observe soundness of the inference rules introduced in the previous section, and prove a few further results of distinct "bootstrapping" nature. In Section 4 we study "extremal" homomorphisms and in particular introduce "differential operators" that, under certain conditions, allow us to come up with new and very useful axioms valid for such homomorphisms. In Sections 2-4 we will be trying to convey as much intuition as possible as to what and why we are doing, as well as provide examples illustrating our abstract notions with concrete calculations.

In Section 5 we give the first application of our machinery (further applications will appear elsewhere, and every one of them will use only a small fragment of the whole theory). Mantel's theorem [Man](generalized in Turán's classical paper [Tur] to cliques of arbitrary fixed size) asymptotically states that if the edge density ρ of a graph is $> 1/2$ then the density of triangles in this graph is > 0. The quantitative version of this question (that is, what *exactly* is the minimal possible density of triangles given ρ) has received much attention in the combinatorial literature, but so far only partial results are known [Goo, Bol, LoSi, Fish]. In particular, Fisher [Fish] solved this question for $\rho \in [1/2, 2/3]$. We give a totally new proof of Fisher's result which in our calculus amounts to a computation of several lines.

Lastly, in Section 6 we formulate a few open questions; all of them can be vaguely interpreted as attempts at asking if there is any sort of "compactness theorem" for our calculus. Or, in other words, can any true statement in extremal combinatorics be proved by "finite" methods appealing only to flags of bounded size?

§2. Syntax. Let T be a universal first-order theory with equality in a language L containing only predicate symbols; we assume that T has infinite models. Our assumptions imply that every set of elements of a model of T induces a model of T, and that T has at least one finite model of every given size. T will be almost always considered fixed (and dropped from notation); one notable exception will be Section 2.3 and a few other places related to it. As a reminder that we are eventually interested in combinatorial applications, the ground set of a model M will be denoted by $V(M)$, and its elements will be called *vertices*. For $V \subseteq V(M)$, $M|_V$ is the sub-model induced by V, and $M - V \stackrel{\text{def}}{=} M|_{V(M)\setminus V}$ is the result of removing vertices in V from the model M. $M - \{v\}$ will be abbreviated to $M - v$. A *model embedding* $\alpha : M \longrightarrow N$ is an injective mapping $\alpha : V(M) \longrightarrow V(N)$ that induces an isomorphism between M and $N|_{\mathrm{im}(\alpha)}$. $M \approx N$ means that M and N are isomorphic, and we let $\mathcal{M}_n$ denote the set of all finite models of T on n vertices up to an isomorphism.

We review some combinatorial notation. $[k] \stackrel{\text{def}}{=} \{1, 2, \ldots, k\}$. A collection $V_1, \ldots, V_t$ of finite sets is a *sunflower with center* C if $V_i \cap V_j = C$ for every two distinct $i, j \in [t]$. $V_1, \ldots, V_t$ are called the *petals* of the sunflower. Following the standard practice in discrete mathematics, we often visualize probability measures as random objects "picked", "drawn" or "chosen" according to them. And, trying to revive an extremely handy but unfortunately almost entirely forgotten convention from the classical book [ErSp], we always use $\boldsymbol{math\ bold\ face}$ for denoting random objects.

2.1. Flag algebras: definition. For $M \in \mathcal{M}_\ell$ and $N \in \mathcal{M}_L$ with $\ell \leq L$, let $p(M, N)$ be the probability of the event $M \approx N|_{\boldsymbol{V}}$, where $\boldsymbol{V}$ is a randomly chosen subset of $V(N)$ with ℓ vertices. As we explained in the Introduction, we are typically interested in the behaviour of $p(M, N)$ where M belongs to a fixed finite collection of "template" models, and $L \longrightarrow \infty$. It immediately turns out, however, that in order to *prove* anything intelligent about these quantities, one almost always needs a relativized version in which several distinguished vertices from M must attain prescribed values in N. So, we at once treat this more general case.

A *type* σ is a model M of the theory T with $V(M) = [k]$ for some non-negative integer k called the *size* of σ and denoted by $|\sigma|$. To every type σ of size k we can associate the universal theory T^σ in the extended language $L(c_1, \ldots, c_k)$ (c_i are new constants) by appending to T the open diagram of σ. T^σ is an extension of T in the language $L(c_1, \ldots, c_k)$, complete with respect to open closed formulas, and it uniquely determines σ. This logical representation will be extremely useful in Section 2.3, and for the time being it at least somewhat justifies our usurpation of the term "type" from model theory.

A *σ-flag* is a pair $F = (M, \theta)$, where M is a finite model and $\theta : \sigma \longrightarrow M$ is a model embedding (note that in the logical representation σ-flags are precisely finite models of T^σ). For small values of $k = |\sigma|$ we will sometimes write down the flag (M, θ) by explicitly listing all labeled vertices in the form $(M, \theta(1), \ldots, \theta(k))$. If $F = (M, \theta)$ is a σ-flag and $V \subseteq V(M)$ contains $\mathrm{im}(\theta)$, then the sub-flag $(M|_V, \theta)$ will be often denoted by $F|_V$. Likewise, if $V \cap \mathrm{im}(\theta) = \emptyset$, we use the notation $F - V$ for $(M - V, \theta)$. A *flag embedding* $\alpha : F \longrightarrow F'$, where $F = (M, \theta)$ and $F' = (M', \theta')$ are σ-flags, is a model embedding $\alpha : M \longrightarrow M'$ such that $\theta' = \alpha\theta$ ("label-preserving"). F and F' are isomorphic (again denoted $F \approx F'$) if there is a one-to-one flag embedding $\alpha : F \longrightarrow F'$. Let $\mathcal{F}^\sigma$ be the set of all σ-flags (up to an isomorphism), and $\mathcal{F}^\sigma_\ell \stackrel{\text{def}}{=} \{(M, \theta) \in \mathcal{F}^\sigma \mid M \in \mathcal{M}_\ell\}$ be the set of all σ-flags on ℓ vertices. In particular, $\mathcal{M}_\ell$ can (and often will) be identified with $\mathcal{F}^0_\ell$, where 0 is the only type of size 0. $\mathcal{F}^\sigma_{|\sigma|}$ consists of the single element (σ, id), where $\mathrm{id} : \sigma \longrightarrow \sigma$ is the identity embedding. We will denote this special σ-flag by 1_σ or even simply by 1 when σ is clear from the context. When $M \in \mathcal{M}_\ell$, a type σ with $|\sigma| \leq \ell$ is embeddable in M and has the property that all σ-flags resulting from such embeddings are isomorphic, we will denote this uniquely defined σ-flag by M^σ.

EXAMPLE 1 (see illustration on page 7). In the course of this paper, we will illustrate our abstract notions using two specific theories: the theory of undirected graphs T_{Graph} and the theory of directed graphs $T_{Digraph}$. In both cases we consider only *simple* graphs, that is we forbid loops and multiple edges (and in the oriented case we also forbid edges connecting any two vertices in opposite directions). In both cases $E(G)$ is the set of edges, $G - E$ is the result of removing the edges in E from G (without changing the vertex set), and for $e \in E(G)$, $G - e \stackrel{\text{def}}{=} G - \{e\}$.

Undirected case. For an undirected graph G we denote by $\bar{G}$ its complement (on the same vertex set). $K_\ell, P_\ell \in \mathcal{M}_\ell$ are an ℓ-vertex clique and an ℓ-vertex path (of length $\ell - 1$), respectively. $K_{1,l} \in \mathcal{M}_{\ell+1}$ is the *star* with ℓ rays (thus, $P_3 \approx K_{1,2}$).

We denote by 1 the (only) type of size 1, and by $E, \bar{E}$ types of size 2 corresponding to an edge [non-edge, respectively]. Then, according to our convention we have uniquely defined flags $K_\ell^1 \in \mathcal{F}_\ell^1$, $K_\ell^E \in \mathcal{F}_\ell^E$ and $\bar{P}_3^E \in \mathcal{F}_3^E$. The edge considered as a 0-flag $K_2 \in \mathcal{F}_2^0$ will be denoted by ρ, and the same edge considered as a 1-flag K_2^1 will be denoted by e.

In contrast, flags like P_3^1 or P_3^E are *not* uniquely defined. Call the vertex v_c of degree 2 in P_3 the *center* vertex, and two other vertices (of degree 1) *border vertices*; let v_b be one of them. Then we have *two* different versions of P_3^1: $P_3^{1,c} \stackrel{\text{def}}{=} (P_3, v_c)$ and $P_3^{1,b} \stackrel{\text{def}}{=} (P_3, v_b)$. Likewise, let $P_3^{E,c} \stackrel{\text{def}}{=} (P_3, v_c, v_b)$ and $P_3^{E,b} \stackrel{\text{def}}{=} (P_3, v_b, v_c)$. This can be further generalized to $K_{1,\ell}^{1,c}, K_{1,\ell}^{1,b}, K_{1,\ell}^{E,c}, K_{1,\ell}^{E,b}$ for any star $K_{1,\ell}$.

Directed case. Let $\vec{C}_n, \vec{T}_n \in \mathcal{M}_n$ be an oriented cycle and a transitive tournament on n vertices, respectively. $\vec{K}_{1,\ell}, \vec{K}_{\ell,1} \in \mathcal{M}_{\ell+1}$ are two orientations of the star $K_{1,\ell}$ in which all rays are oriented from the center [to the center, respectively]. 1 is again the only type of size 1, and A is the type of size 2 with $E(A) = \{< 1, 2 >\}$. A is our first example of a *non-symmetric* type, i.e. a type which is not preserved under the full group of permutations S_k. $\rho \in \mathcal{F}_2^0$ still has the same meaning as in the undirected case. No single element of $\mathcal{F}_2^1$, however, corresponds to e, and ρ gives rise to *two* different 1-flags $\alpha, \beta \in \mathcal{F}_2^1$, where θ labels the tail vertex in α, and in β it labels the head vertex. $\vec{C}_n^1, \vec{C}_n^A, \vec{K}_{1,\ell}^A, \vec{K}_{\ell,1}^A$ are uniquely defined, whereas there are many possibilities for turning $\vec{T}_n$ into a 1-flag or an A-flag, and *all* of them lead to pairwise different flags. $\vec{K}_{1,\ell}^{1,c}, \vec{K}_{1,\ell}^{1,b}, \vec{K}_{\ell,1}^{1,c}, \vec{K}_{\ell,1}^{1,b}$ are defined exactly as in the undirected case.

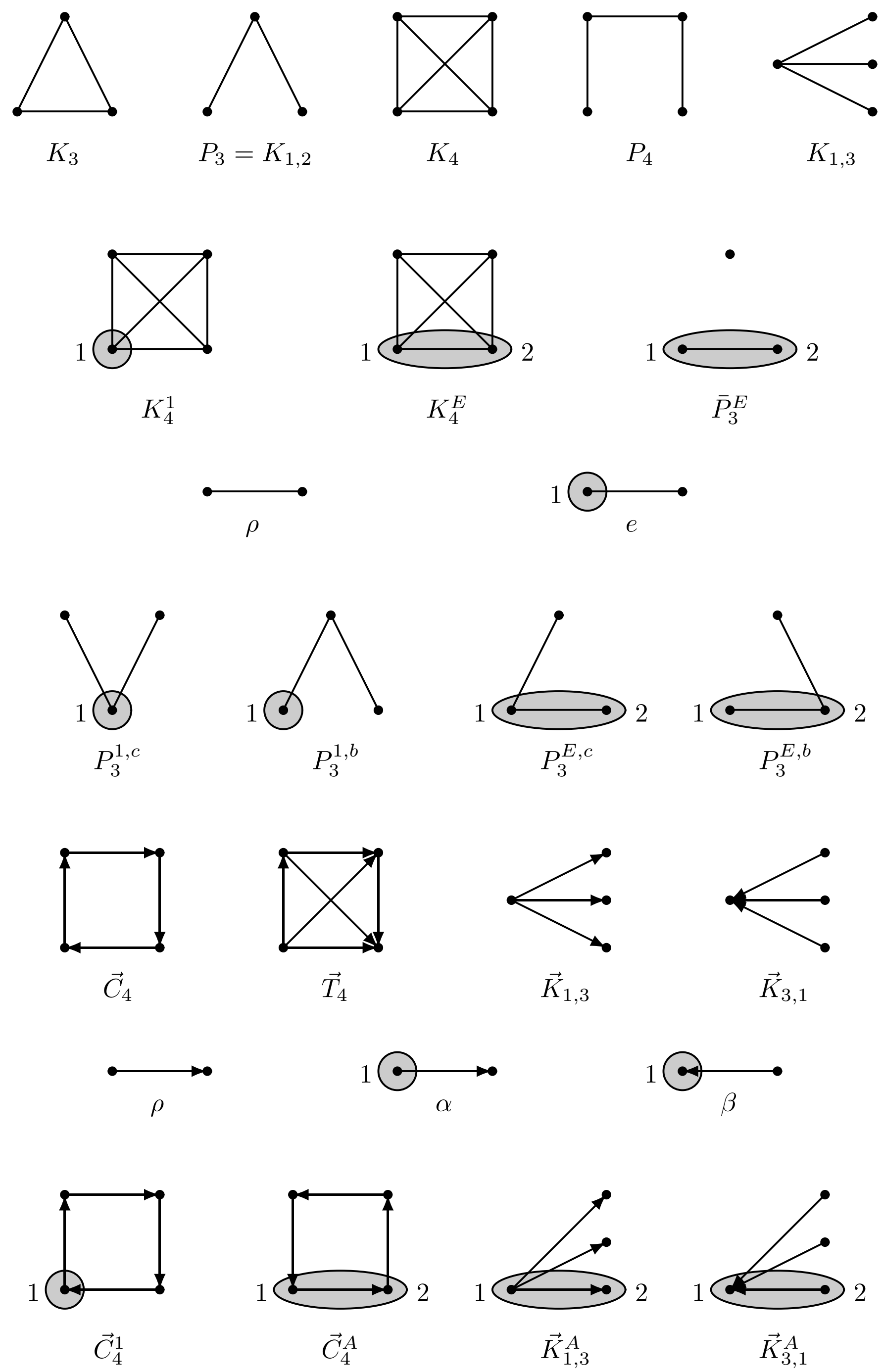

Figure 1. Examples of flags

Definition 1. Fix a type σ of size k, assume that integers $\ell, \ell_1, \ldots, \ell_t \geq k$ are such that

$$\ell_1 + \cdots + \ell_t - k(t-1) \leq \ell, \tag{1}$$

and $F = (M, \theta) \in \mathcal{F}^\sigma_\ell$, $F_1 \in \mathcal{F}^\sigma_{\ell_1}, \ldots, F_t \in \mathcal{F}^\sigma_{\ell_t}$ are σ-flags. We define the (key) quantity $p(F_1, \ldots, F_t; F) \in [0, 1]$ as follows. Choose in $V(M)$ uniformly at random a sunflower $(\boldsymbol{V_1}, \ldots, \boldsymbol{V_t})$ with center $\mathrm{im}(\theta)$ and petals of sizes $\ell_1, \ldots, \ell_t$, respectively (the inequality (1) ensures that such sunflowers do exist). We let $p(F_1, \ldots, F_t; F)$ denote the probability of the event "$\forall i \in [t](F|_{\boldsymbol{V}_i} \approx F_i)$". When $t = 1$, we use the notation $p(F_1, F)$ instead of $p(F_1; F)$.

Example 2 (undirected graphs). $p(\rho, G)$ is the edge density of G, and $p(K_3, G)$ is the density of triangles in G. $p(\underbrace{\rho, \ldots, \rho}_{t \text{ times}}; G)$ is the density of matchings with t edges (*not* necessarily induced). G is a complete t-partite graph for some t iff $p(\bar{P}_3, G) = 0$. In type 1, $p(e, (G, v)) = \frac{\deg_G(v)}{|V(G)|-1}$ is the relative degree of the labelled vertex v. $p(e, e; P_3^{1,b}) = 0$ whereas $p(e, e; P_3^{1,c}) = 1$. $p(e, e; K_\ell^1) = 1$ for every $\ell \geq 3$. In type E, $p(P_3^{E,b}, P_3^{E,b}; K_{1,3}^{E,b}) = p(P_3^{E,c}, P_3^{E,c}; K_{1,3}^{E,c}) = 1$ but $p(P_3^{E,b}, P_3^{E,c}; K_{1,3}^{E,b}) = p(P_3^{E,b}, P_3^{E,c}; K_{1,3}^{E,c}) = 0$.

Lemma 2.1.

a) *$p(1_\sigma, F_1, \ldots, F_t; F) = p(F_1, \ldots, F_t; F)$ and $p(1_\sigma, F) = 1$.*
b) *For $F, F' \in \mathcal{F}^\sigma_\ell$, $p(F, F') = 1$ if $F = F'$ and $p(F, F') = 0$ otherwise.*
c) *$p(F_1, \ldots, F_t; F) = p(F_{\gamma(1)}, \ldots, F_{\gamma(t)}; F)$ for any permutation $\gamma \in S_t$.*

Proof. Obvious. ⊣

Lemma 2.2 (chain rule). *Let $|\sigma| = k$, $F_i \in \mathcal{F}^\sigma_{\ell_i}$ $(1 \leq i \leq t)$, $1 \leq s \leq t$, $F \in \mathcal{F}^\sigma_\ell$ and $\tilde{\ell} \leq \ell$ be such that*

$$\begin{cases} \tilde{\ell} + \ell_{s+1} + \cdots + \ell_t - k(t-s) \leq \ell \\ \ell_1 + \cdots + \ell_s - k(s-1) \leq \tilde{\ell}. \end{cases} \tag{2}$$

Then

$$p(F_1, \ldots, F_t; F) = \sum_{\widetilde{F} \in \mathcal{F}^\sigma_{\tilde{\ell}}} p(F_1, \ldots, F_s; \widetilde{F}) p(\widetilde{F}, F_{s+1}, \ldots, F_t; F).$$

In particular $(s = t)$, for every $\tilde{\ell} \leq \ell$ satisfying the inequality

$$\ell_1 + \cdots + \ell_t - k(t-1) \leq \tilde{\ell},$$

we have

$$p(F_1, \ldots, F_t; F) = \sum_{\widetilde{F} \in \mathcal{F}^\sigma_{\tilde{\ell}}} p(F_1, \ldots, F_t; \widetilde{F}) p(\widetilde{F}, F). \tag{3}$$

Proof. Let $F = (M, \theta)$. We present another, two-step way of generating a random sunflower $(\boldsymbol{V_1}, \ldots, \boldsymbol{V_t})$ with the same distribution as the one appearing in Definition 1 (that is, uniform on the set of all possibilities). Namely, we first generate (uniformly at random) a sunflower $(\widetilde{\boldsymbol{V}}, \boldsymbol{V_{s+1}}, \ldots, \boldsymbol{V_t})$ with center $\mathrm{im}(\theta)$ and petals of sizes $\tilde{\ell}, \ell_{s+1}, \ldots, \ell_t$. Then we pick in $\widetilde{\boldsymbol{V}}$ (also uniformly at random) a sunflower $(\boldsymbol{V_1},\dots,\boldsymbol{V_s})$ with center $\mathrm{im}(\theta)$ and petals of sizes $\ell_1,\dots,\ell_s$. By symmetry, and due to the inequalities (2), this procedure also leads to the *uniform* distribution on the set of all sunflowers $(V_1,\dots,V_t)$ with center $\mathrm{im}(\theta)$ and petals of sizes $\ell_1,\dots,\ell_t$. Now, the identity we are proving becomes simply the formula of total probability, the right-hand side corresponding to the partition of the probability space according to the isomorphism type of $F|_{\widetilde{\boldsymbol{V}}} \in \mathcal{F}^\sigma_{\tilde{\ell}}$. $\dashv$

The following lemma states that when F is large, $p(F_1,\dots,F_t;F)$ becomes almost multiplicative in the first t arguments. It will not be needed until Section 3, but we present it here (for the purpose of orientation).

Lemma 2.3. *Let $F_i \in \mathcal{F}^\sigma_{\ell_i}$ $(1 \le i \le t)$ and $F \in \mathcal{F}^\sigma_\ell$. Then*

$$\left| p(F_1,\dots,F_t;F) - \prod_{i=1}^t p(F_i,F) \right| \le \frac{(\ell_1+\cdots+\ell_t)^{O(1)}}{\ell}.$$

Proof. Let $F=(M,\theta)$. Choose $\boldsymbol{V_i} \subseteq V(M)$ of size ℓ_i with $\mathrm{im}(\theta) \subseteq \boldsymbol{V_i}$ uniformly at random but *independently of one another*. Then $\prod_{i=1}^t p(F_i,F) = \mathbf{P}[A]$ and $p(F_1,\dots,F_t;F) = \mathbf{P}[\,A \mid B\,]$ where A is the event "$\forall i \in [t](F|_{\boldsymbol{V_i}} \approx F_i)$", and B is the event "$(\boldsymbol{V_1},\dots,\boldsymbol{V_t})$ is a sunflower with center $\mathrm{im}(\theta)$". We now have $|\mathbf{P}[A] - \mathbf{P}[\,A \mid B\,]\,| \le 1 - \mathbf{P}[B] \le \sum_{i\neq j} \mathbf{P}[(\boldsymbol{V_i} \cap \boldsymbol{V_j}) \supset \mathrm{im}(\theta)] \le \frac{(\ell_1+\cdots+\ell_t)^{O(1)}}{\ell}$. $\dashv$

Let $\mathbb{R}\mathcal{F}^\sigma$ be the linear space with the basis $\mathcal{F}^\sigma$, i.e. the space of all formal finite linear combinations of σ-flags with real coefficients. Let $\mathcal{K}^\sigma$ be its linear subspace generated by all elements of the form

$$\widetilde{F} - \sum_{F \in \mathcal{F}^\sigma_\ell} p(\widetilde{F},F)F, \tag{4}$$

where $\widetilde{F} \in \mathcal{F}^\sigma_{\tilde{\ell}}$ and $|\sigma| \le \tilde{\ell} \le \ell$. Let

$$\mathcal{A}^\sigma \stackrel{\mathrm{def}}{=} (\mathbb{R}\mathcal{F}^\sigma)/\mathcal{K}^\sigma.$$

Introduce the bilinear mapping $(\mathbb{R}\mathcal{F}^\sigma) \otimes (\mathbb{R}\mathcal{F}^\sigma) \longrightarrow \mathcal{A}^\sigma,\ f \otimes g \mapsto f \cdot g$ as follows. For two σ-flags $F_1 \in \mathcal{F}^\sigma_{\ell_1},\ F_2 \in \mathcal{F}^\sigma_{\ell_2}$ choose arbitrarily $\ell \ge \ell_1 + \ell_2 - |\sigma|$ and let

$$F_1 \cdot F_2 \stackrel{\mathrm{def}}{=} \sum_{F \in \mathcal{F}^\sigma_\ell} p(F_1,F_2;F)F. \tag{5}$$

Extend this mapping onto the whole $(\mathbb{R}\mathcal{F}^\sigma) \otimes (\mathbb{R}\mathcal{F}^\sigma)$ by linearity.

Definition 2. A type σ is *non-degenerate* if $\mathcal{F}^\sigma_\ell \neq \emptyset$ for all $\ell \ge |\sigma|$ (or, equivalently, if the theory T^σ has an infinite model).

In "reasonable" theories all types are non-degenerate (see Theorem 2.7 below for a much stronger property). For an example of a degenerate type, append to T_{Graph} the extra axiom "every *four* vertices span a triangle-free subgraph". Then the triangle on $\{1,2,3\}$ is a degenerate type.

Lemma 2.4.

- **a)** *The right-hand side of* (5) *does not depend on the choice of* ℓ (*modulo* $\mathcal{K}^\sigma$).
- **b)** (5) *induces a bilinear mapping* $\mathcal{A}^\sigma \otimes \mathcal{A}^\sigma \longrightarrow \mathcal{A}^\sigma$.

c) *Let $\ell \geq \ell_1 + \cdots + \ell_t - k(t-1)$. Then for $F_i \in \mathcal{F}^\sigma_{\ell_i}$ $(1 \leq i \leq t)$ we have the identity*

$$((F_1 \cdot F_2) \cdot F_3 \ldots) \cdot F_t = \sum_{F \in \mathcal{F}^\sigma_\ell} p(F_1, \ldots, F_t; F)F \pmod{\mathcal{K}^\sigma}.$$

d) *If σ is non-degenerate then the induced mapping $\mathcal{A}^\sigma \otimes \mathcal{A}^\sigma \longrightarrow \mathcal{A}^\sigma$ endows $\mathcal{A}^\sigma$ with the structure of a commutative associative algebra with the identity element 1_σ.*

Proof. **a).** Let $\ell \geq \tilde{\ell} \geq \ell_1 + \ell_2 - |\sigma|$. Then by Lemma 2.2 we have

$$\sum_{F \in \mathcal{F}^\sigma_\ell} p(F_1, F_2; F)F = \sum_{F \in \mathcal{F}^\sigma_\ell} \sum_{\widetilde{F} \in \mathcal{F}^\sigma_{\tilde{\ell}}} p(F_1, F_2; \widetilde{F}) p(\widetilde{F}, F)F$$
$$= \sum_{\widetilde{F} \in \mathcal{F}^\sigma_{\tilde{\ell}}} p(F_1, F_2; \widetilde{F}) \sum_{F \in \mathcal{F}^\sigma_\ell} p(\widetilde{F}, F)F = \sum_{\widetilde{F} \in \mathcal{F}^\sigma_{\tilde{\ell}}} p(F_1, F_2; \widetilde{F})\widetilde{F} \pmod{\mathcal{K}^\sigma}.$$

b). By Lemma 2.1 c), the operation $\cdot$ is symmetric, so we only have to show that $f_1 \in \mathcal{K}^\sigma,\ f_2 \in \mathbb{R}\mathcal{F}^\sigma$ implies $f_1 \cdot f_2 \in \mathcal{K}^\sigma$. By linearity, we may additionally assume that f_1 has the form (4) and $f_2 = F'$ is a σ-flag. That is, we want to prove

$$\widetilde{F} \cdot F' = \sum_{F \in \mathcal{F}^\sigma_\ell} p(\widetilde{F}, F)(F \cdot F') \pmod{\mathcal{K}^\sigma}.$$

By the already proven part **a)**, we may expand here $\widetilde{F} \cdot F'$ and $F \cdot F'$ as summations over $\widehat{F} \in \mathcal{F}^\sigma_L$ with the *same* $L \geq \ell$. Looking at the coefficients in front of every particular $\widehat{F} \in \mathcal{F}^\sigma_L$, we see $p(\widetilde{F}, F'; \widehat{F})$ in the left-hand side, and $\sum_{F \in \mathcal{F}^\sigma_\ell} p(\widetilde{F}, F)p(F, F'; \widehat{F})$ in the right-hand side. They coincide by Lemma 2.2.

c). By another straightforward application of Lemma 2.2, repeated $(t-1)$ times.

d). Commutativity and associativity follow from part **c)** and Lemma 2.1 c). The fact that 1_σ is the identity element follows from Lemma 2.1 a). Finally, we have to check that $0 \neq 1$ in this algebra, that is $1_\sigma \notin \mathcal{K}^\sigma$. Consider any *finite* set $\mathcal{R}$ of relations of the form (4), and let L be a common upper bound on the number of vertices in flags appearing in those relations. Since σ is non-degenerate, $\mathcal{F}^\sigma_L$ is non-empty; choose $\widehat{F} \in \mathcal{F}^\sigma_L$ arbitrarily. Then the linear functional on the direct sum $\bigoplus_{|\sigma| \leq \ell \leq L} \mathcal{F}^\sigma_\ell$ that maps every flag F to $p(F, \widehat{F})$ nullifies all relations in $\mathcal{R}$ and does not nullify 1_σ. Therefore, 1_σ does not belong to the linear subspace spanned by $\mathcal{R}$. ⊣

The algebras $\mathcal{A}^\sigma$ that we will call *flag algebras* make the backbone of our whole approach. We will denote by $\mathcal{A}^\sigma_\ell$ the linear subspace generated in $\mathcal{A}^\sigma$ by $\mathcal{F}^\sigma_\ell$. Also, when simultaneously working with several different theories (like in Section 2.3), we will be using the notation like $\mathcal{F}^\sigma_\ell[T]$ or $\mathcal{A}^\sigma[T]$ to indicate which theory these objects are related to.

Example 3 (see illsustration on page 11).

Undirected case.

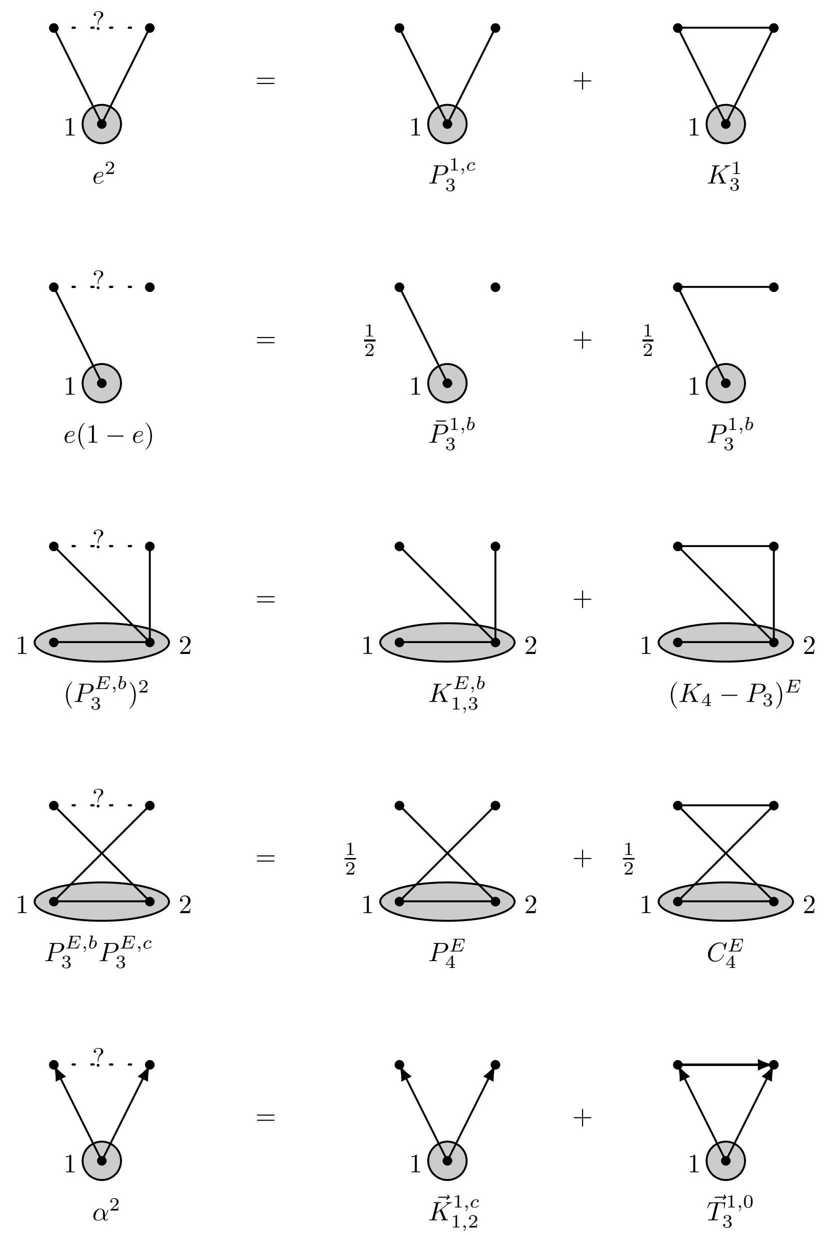

FIGURE 2. Multiplication

$\rho = \frac{1}{3}\bar{P}_3 + \frac{2}{3}P_3 + K_3$. $\rho^2 = \frac{1}{3}\sum_{G\in\mathcal{M}_4} m_2(G)G$, where $m_2(G)$ is the number of 2-matchings in G (*not* necessarily induced). $e^2 = P_3^{1,c} + K_3^1$. $e(1-e) = \frac{1}{2}(\bar{P}_3^{1,b} + P_3^{1,b})$. $(P_3^{E,b})^2 = K_{1,3}^{E,b} + (K_4 - P_3)^E$, whereas $P_3^{E,b}P_3^{E,c} = \frac{1}{2}(P_4^E + C_4^E)$ (here $K_4 - P_3$ and P_4 are turned into E-flags in an appropriate way).

Directed case.

We can now *define* the element $K_3 \in \mathcal{A}_3^0$ as $K_3 \stackrel{\mathrm{def}}{=} \vec{C}_3 + \vec{T}_3$, and define $e \in \mathcal{A}_2^1$ as $e \stackrel{\mathrm{def}}{=} \alpha + \beta$ (cf. Section 2.3 below). $\alpha^2 = \vec{K}_{1,2}^{1,c} + \vec{T}_3^{1,0}$, where $\vec{T}_3^{1,0}$ is the 1-flag obtained from $\vec{T}_3$ by labelling the vertex of in-degree 0.

Remark 1. Flag algebras are related to graph algebras (introduced in the context of graph homomorphisms in [FLS]) roughly as follows. While defining multiplication in that context, an inherent ambiguity arises for $k \geq 2$. Depending on the current goals, researchers in the area interchangeably work with the version in which multiple edges are allowed, and the version in which they are forbidden. The algebra structure seems to have been considered so far only for the version with multiple edges, which are strictly forbidden in our framework. However, there do not seem to be any principal obstacles to defining the graph algebras also in the context of *simple* graphs, and apparently the resulting algebra will be isomorphic to the product $\prod_{|\sigma|=k} \mathcal{A}^\sigma$. However, since graph algebras for simple graphs have apparently not been considered in the literature before, we prefer to be on the safe side and avoid definite statements in this remark.

We will be interested in the relations $f \geq g$ $(f, g \in \mathcal{A}^\sigma)$ that are "asymptotically true" (this will be made precise and further developed in Section 3). In the rest of this section we will define certain operators that will correspond to "inference rules" of our calculus; their soundness will again be shown in Section 3.

2.2. Averaging: downward operator. Almost all proofs in extremal combinatorics use, in one or another form, the following simple (and yet very powerful) idea. Suppose, say, that in the theory T_{Graph} we have proved some inequality $f(v_1, v_2) \geq 0$ for every pair of vertices $v_1, v_2 \in V(G)$ not connected by an edge. Then, averaging over all non-edges, we will get another inequality depending only on the graph under consideration. In our formalism this operation is described as a linear operator (*not* an algebra homomorphism) $\mathcal{A}^\sigma \longrightarrow \mathcal{A}^{\sigma'}$ between *different* flag algebras (and since $|\sigma'|$ will be always less than $|\sigma|$, we called this a downward operator in the title of the section).

Given a type σ of size k, $k' \leq k$ and an injective mapping $\eta : [k'] \longrightarrow [k]$, let $\sigma|_\eta$ be the naturally induced type of size k' (that is, for any predicate symbol $P(x_1, \ldots, x_r)$ in L and any $i_1, \ldots, i_r \in [k']$, $\sigma|_\eta \models P(i_1, \ldots, i_r)$ iff $\sigma \models P(\eta(i_1), \ldots, \eta(i_r))$). For a σ-flag $F = (M, \theta)$, the $\sigma|_\eta$-flag $F|_\eta$ is defined as $F|_\eta \stackrel{\mathrm{def}}{=} (M, \theta\eta)$. In particular, we have the $\sigma|_\eta$-flag $(1_\sigma)|_\eta = (\sigma, \eta)$, where in the right-hand side σ is considered as an *unlabelled* model of T.

Next, we define the normalizing factor $q_{\sigma,\eta}(F) \in [0, 1]$ as follows. For $F = (M, \theta)$ we generate an injective mapping $\boldsymbol{\theta} : [k] \longrightarrow V(M)$, uniformly at random subject to the additional restriction that it must be consistent with θ *on* $\mathrm{im}(\eta)$ (that is, $\boldsymbol{\theta}\eta = \theta\eta$). We let $q_{\sigma,\eta}(F)$ be the probability that $\boldsymbol{\theta}$ defines a model embedding $\sigma \longrightarrow M$ *and* the resulting σ-flag $(M, \boldsymbol{\theta})$ is *isomorphic* to F.

Finally, we let

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

and extend this mapping to a linear mapping $\mathbb{R}\mathcal{F}^\sigma \longrightarrow \mathbb{R}\mathcal{F}^{\sigma|_\eta}$. The most interesting case is when $k' = 0$ (thus, also $\sigma|_\eta = 0$), and we will abbreviate $[\![\cdot]\!]_{\sigma,0}$ to $[\![\cdot]\!]_\sigma$.

EXAMPLE 4.

Undirected graphs. $[\![e]\!]_1 = \rho$ and, more generally, $[\![K_\ell^1]\!]_1 = [\![K_\ell^E]\!]_E = K_\ell$. $[\![P_3^{1,c}]\!]_1 = \frac{1}{3}P_3$ and $[\![P_3^{1,b}]\!]_1 = \frac{2}{3}P_3$. Thus, $[\![e^2]\!]_1 = K_3 + \frac{1}{3}P_3$. $[\![P_3^{E,b}]\!]_E = [\![P_3^{E,c}]\!]_E = \frac{1}{3}P_3$. $[\![P_3^{E,b}]\!]_{E,1} = \frac{1}{2}P_3^{1,b}$, but $[\![P_3^{E,b}]\!]_{E,2} = P_3^{1,c}$ (here for $i = 1, 2$ we denoted by i the function $\eta : \{1\} \longrightarrow \{1, 2\}$ with $\eta(1) = i$).

Directed graphs. $[\![\alpha]\!]_1 = [\![\beta]\!]_1 = \rho/2$, so we still have $[\![e]\!]_1 = \rho$. $[\![\vec{C}_n^A]\!]_A = \frac{1}{n-1}\vec{C}_n$ and $[\![\vec{T}_n^A]\!]_A = \frac{1}{n(n-1)}\vec{T}_n$ for *every* A-flag $\vec{T}_n^A$ resulting from $\vec{T}_n$.

THEOREM 2.5. **a)** *$[\![\cdot]\!]_{\sigma,\eta}$ takes $\mathcal{K}^\sigma$ to $\mathcal{K}^{\sigma|_\eta}$ and thus defines a linear mapping $\mathcal{A}^\sigma \to \mathcal{A}^{\sigma|_\eta}$.*

b) *$[\![1_\sigma]\!]_{\sigma,\eta} = q_{\sigma,\eta}(1_\sigma) \cdot (\sigma, \eta)$.*

c) (chain rule) *Assume that we are additionally given an injective mapping $\eta' : [k''] \longrightarrow [k']$ for some $k'' \le k'$. Then*

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

PROOF. **a).** Apply $[\![\cdot]\!]_{\sigma,\eta}$ to the relation (4) and expand $\widetilde{F}|_\eta$ in the result as a linear combination of $F' \in \mathcal{F}_\ell^{\sigma|_\eta}$ using the respective relation in the algebra $\mathcal{A}^{\sigma|_\eta}$. Comparing coefficients in front of every particular $F' \in \mathcal{F}_\ell^{\sigma|_\eta}$, we only have to prove that for every fixed $\widetilde{F} \in \mathcal{F}_{\widetilde{\ell}}^\sigma$ and $F' \in \mathcal{F}_\ell^{\sigma|_\eta}$, we have

$$p(\widetilde{F}|_\eta, F') \cdot q_{\sigma,\eta}(\widetilde{F}) = \sum_{\substack{F \in \mathcal{F}_\ell^\sigma \\ F|_\eta = F'}} q_{\sigma,\eta}(F) p(\widetilde{F}, F). \tag{6}$$

For doing that, we (as in the proof of Lemma 2.2) calculate the probability of the same event in two different ways. Namely, let $F' = (M, \theta')$. Pick a pair $(\widetilde{\boldsymbol{V}}, \boldsymbol{\theta})$, where $\boldsymbol{\theta} : [k] \longrightarrow V(M)$ is an injective function and $\widetilde{\boldsymbol{V}} \subseteq V(M)$, $|\widetilde{\boldsymbol{V}}| = \ell$ uniformly at random, but subject to two additional restrictions $\boldsymbol{\theta}\eta = \theta'$ and $\widetilde{\boldsymbol{V}} \supseteq \mathrm{im}(\boldsymbol{\theta})$. Then both sides of (6) calculate the probability of the event "$\boldsymbol{\theta}$ is a model embedding $\sigma \longrightarrow M$ and the σ-flag $(M|_{\widetilde{\boldsymbol{V}}}, \boldsymbol{\theta})$ is isomorphic to $\widetilde{F}$" (the right-hand side splits this event according to the isomorphism type of $(M, \boldsymbol{\theta})$).

b) and **c)** are straightforward. ⊣

REMARK 2. The normalizing factor $q_{\sigma,\eta}(1_\sigma)$ in part b) of this theorem results from the dual treatment of σ as totally labelled and partially labelled model. Algebraically, it can be computed as follows. Let S be the subgroup in S_k stabilizing (pointwise) all points in $\mathrm{im}(\eta)$ and A be the automorphism group of σ. Then

$$q_{\sigma,\eta}(1_\sigma) = \frac{|S \cap A|}{|S|} = \frac{|S \cap A|}{(k - k')!} = (S : S \cap A)^{-1}.$$

2.3. Interpretations, upward operators and induction. Let us begin with three simple examples representing typical techniques in extremal combinatorics.

EXAMPLE 5. In the theory T_{Graph}, consider the element $K_3^1+\bar{P}_3^{1,c}+P_3^{1,b} \in \mathcal{A}_3^1$. This element represents the fact that the two unlabelled vertices are connected by an edge, and thus we expect this element to be equal to $\rho \in \mathcal{A}^0$ regardless of the choice of the labelled vertex. Thus, we have found in $\mathcal{A}^1$ an element "effectively equal" to an element of $\mathcal{A}^0$, and we can use it for further reasoning *within the algebra* $\mathcal{A}^1$.

EXAMPLE 6. Suppose we are trying to prove some inequality $f \geq 0$ for all undirected graphs. Given a graph G, for every $v \in V(G)$ we may consider its neighbourhood $N(v) \stackrel{\text{def}}{=} \{\, w \in V(G) \mid (v, w) \in E(G) \,\}$ and assume (by induction) that the inequality $f \geq 0$ holds for $G|_{N(v)}$. Then we can average the results over $v \in V(G)$, thus getting some new relation, and then we can use it as an "additional axiom" for proving $f \geq 0$ for the graph G itself.

EXAMPLE 7 (cf. [CaF]). Suppose now that we are working with 3-hypergraphs G. Then for every $v \in V(G)$ we may form its *link* as $\{\, (v', v'') \mid (v, v', v'') \in E(G) \,\}$. This is already an *ordinary* graph, and via this operation (followed again by averaging over v) we can use to our advantage every relation previously proved for ordinary graphs.

In our framework all three themes are treated simultaneously as special cases of an extremely general construction based on the logical notion of interpretation, and we begin with this construction in its full generality. The result (Theorem 2.6) is a bit technical, and in the three following subsections we indicate simpler partial cases roughly corresponding to the three examples above.

All interpretations considered in this paper will be open (that is, given by open formulas). To fix notation, and for the benefit of non-logic readership, we review the definition below.

DEFINITION 3. Let T_1 and T_2 be two universal first-order theories with equality in (possibly different) languages L_1 and L_2. We assume that L_1 and L_2 do not contain function symbols, but, along with predicate symbols, *we also allow constants*. Let $U(x)$ be an open formula in the language L_2 such that $T_2 \vdash U(c)$ for every constant $c \in L_2$, and I be a *translation* that takes every predicate symbol $P(x_1, \ldots, x_r) \in L_1$ to an open formula $I(P)(x_1, \ldots, x_r)$ in the language L_2, and takes every constant $c \in L_1$ to a constant $I(c) \in L_2$. I is extended to open formulas of the language L_1 by declaring that it commutes with logical connectives. The pair (U, I) is an *open interpretation* of T_1 in T_2, denoted $(U, I) : T_1 \rightsquigarrow T_2$ if for every axiom $\forall x_1, \ldots, x_n A(x_1, \ldots, x_n)$ of T_1 we have

$$T_2 \vdash \forall x_1, \ldots, x_n((U(x_1) \wedge \ldots \wedge U(x_n)) \Longrightarrow I(A)(x_1, \ldots, x_n)). \tag{7}$$

A model M of the theory T_2 will be called an *U-model* if $M \models \forall x U(x)$. Under the interpretation $(U, I) : T_1 \rightsquigarrow T_2$, every U-model M of T_2 gives rise, in a natural way, to a model $I(M)$ of T_1 that has the same ground set and *does not depend on* U (since the interpretation is open).

Now we are interested in the following, more specific set-up.

DEFINITION 4. Let T_1, T_2 be two universal theories in languages with equality, this time containing only predicate symbols (and no constants). Let σ_1, σ_2 be non-degenerate types in theories T_1, T_2 that have sizes k_1, k_2, respectively. The theories $T_1^{\sigma_1}, T_2^{\sigma_2}$ were introduced at the beginning of Section 2.1; recall that $\mathcal{F}^{\sigma_i}[T_i]$ is precisely the set of finite models of $T_i^{\sigma_i}$. Assume that we are given an open interpretation $(U, I) : T_1^{\sigma_1} \rightsquigarrow T_2^{\sigma_2}$. Then the condition (7) applied to the formulas $c_i \neq c_j$ $(1 \leq i < j \leq k_1)$ implies in particular that I acts injectively on the constants $c_1, \ldots, c_{k_1}$. Define the corresponding injective mapping $\eta : [k_1] \longrightarrow [k_2]$ by the property $I(c_i) = c_{\eta(i)}$ $(i \in [k_1])$. Denote $[k_2] \setminus \eta([k_1])$ by D, and its cardinality $k_2 - k_1$ by d.

Let $\mathcal{F}^{\sigma_2,U}[T_2]$ be the set of all σ_2-flags that correspond to U-models of the theory $T_2^{\sigma_2}$, and let $\mathcal{F}_\ell^{\sigma_2,U}[T_2] \stackrel{\text{def}}{=} \mathcal{F}^{\sigma_2,U}[T_2] \cap \mathcal{F}_\ell^{\sigma_2}[T_2]$. For $F \in \mathcal{F}_\ell^{\sigma_2,U}[T_2]$ we thus have a naturally defined σ_1-flag $I(F) \in \mathcal{F}_{\ell_1}^{\sigma_1}[T_1]$, and if $F = (M, \theta)$ then $I(F)$ has the form $(N, \theta\eta)$, where N is a model of T_1 with $V(N) = V(M)$. We also let

$$I'(F) \stackrel{\text{def}}{=} I(F) - \theta(D)$$

be the result of removing those labelled vertices that are "not in the image" of the interpretation I.

We now introduce the special element u that is the sum of all σ_2-flags on k_2+1 vertices whose only non-labelled vertex satisfies U:

$$u \stackrel{\text{def}}{=} \sum \left\{ F \;\middle|\; F \in \mathcal{F}_{k_2+1}^{\sigma_2,U}[T_2] \right\}, \tag{8}$$

and assume that

$$u \textit{ is not a zero divisor in } \mathcal{A}^{\sigma_2}[T_2]. \tag{9}$$

Then we may consider the *localization* $\mathcal{A}_u^{\sigma_2}[T_2]$ of the algebra $\mathcal{A}^{\sigma_2}[T_2]$ with respect to the multiplicative system $\left\{ u^\ell \mid \ell \in \mathbb{N} \right\}$ (every element of $\mathcal{A}_u^{\sigma_2}[T_2]$ has the form $u^{-\ell} f$ with $f \in \mathcal{A}^{\sigma_2}[T_2]$ and $\ell \geq 0$). Finally, for any flag $F_1 \in \mathcal{F}_{\ell_1}^{\sigma_1}[T_1]$ we define the element $\pi^{(U,I)}(F_1) \in \mathcal{A}_u^{\sigma_2}[T_2]$ as follows:

$$\pi^{(U,I)}(F_1) \stackrel{\text{def}}{=} \frac{1}{u^{\ell_1 - k_1}} \cdot \sum \left\{ F_2 \in \mathcal{F}_{\ell_1+d}^{\sigma_2,U}[T_2] \mid I'(F_2) \approx F_1 \right\}.$$

We extend $\pi^{(U,I)}$ to a linear mapping $\mathbb{R}\mathcal{F}^{\sigma_1}[T_1] \longrightarrow \mathcal{A}_u^{\sigma_2}[T_2]$.

The following is the main result of this section. Just as the definitions above, its proof is a little bit technical, so the reader may want to try it out on three simpler partial cases in the following subsections.

THEOREM 2.6. *$\pi^{(U,I)}(\mathcal{K}^{\sigma_1}[T_1]) = 0$, and the induced mapping*

$$\pi^{(U,I)} : \mathcal{A}^{\sigma_1}[T_1] \longrightarrow \mathcal{A}_u^{\sigma_2}[T_2]$$

is an algebra homomorphism.

PROOF. We first prove $\pi^{(U,I)}(\mathcal{K}^{\sigma_1}[T_1]) = 0$. Assume that $k_1 \leq \tilde{\ell} \leq \ell$, and let $\widetilde{F}_1 \in \mathcal{F}_{\tilde{\ell}}^{\sigma_1}[T_1]$. Applying $\pi^{(U,I)}$ to the relation $\widetilde{F}_1 - \sum_{F_1 \in \mathcal{F}_\ell^{\sigma_1}} p(\widetilde{F}_1, F_1) F_1$ (see (4)), we need to prove in the algebra $\mathcal{A}_u^{\sigma_2}[T_2]$ that

$$\frac{1}{u^{\tilde{\ell} - k_1}} \cdot \sum \left\{ \widetilde{F}_2 \in \mathcal{F}_{\tilde{\ell}+d}^{\sigma_2,U}[T_2] \;\middle|\; I'(\widetilde{F}_2) \approx \widetilde{F}_1 \right\}$$
$$= \frac{1}{u^{\ell - k_1}} \cdot \sum_{F_1 \in \mathcal{F}_\ell^{\sigma_1}[T_1]} p(\widetilde{F}_1, F_1) \sum \left\{ F_2 \in \mathcal{F}_{\ell+d}^{\sigma_2,U}[T_2] \mid I'(F_2) \approx F_1 \right\}.$$

Multiplying by $u^{\ell-k_1}$ and re-arranging the right-hand side, this amounts to proving

$$u^{\ell-\tilde{\ell}} \cdot \sum \left\{ \widetilde{F}_2 \in \mathcal{F}^{\sigma_2,U}_{\tilde{\ell}+d}[T_2] \,\middle|\, I'(\widetilde{F}_2) \approx \widetilde{F}_1 \right\} = \sum_{F_2 \in \mathcal{F}^{\sigma_2,U}_{\ell+d}[T_2]} p(\widetilde{F}_1, I'(F_2))F_2.$$

Now, using Lemma 2.4 c), expand the left-hand side as a linear combination of flags $F_2 \in \mathcal{F}^{\sigma_2}_{\ell+d}[T_2]$. Note that (since U is open), a σ_2-flag belongs to $\mathcal{F}^{\sigma_2,U}[T_2]$ if and only if all its induced subflags on k_2+1 vertices belong to $\mathcal{F}^{\sigma_2,U}_{k_2+1}[T_2]$. With this remark, it is clear that the coefficient in front of any particular $F_2 \in \mathcal{F}^{\sigma_2}_{\ell+d}[T_2]$ is equal to

$$\begin{cases} 0, & \text{if } F_2 \not\in \mathcal{F}^{\sigma_2,U}_{\ell+d}[T_2] \\ \displaystyle\sum_{\substack{\widetilde{F}_2 \in \mathcal{F}^{\sigma_2,U}_{\tilde{\ell}+d}[T_2] \\ I'(\widetilde{F}_2) \approx \widetilde{F}_1}} p(\widetilde{F}_2, F_2), & \text{if } F_2 \in \mathcal{F}^{\sigma_2,U}_{\ell+d}[T_2]. \end{cases}$$

By comparing coefficients, it only remains to prove the identity

$$\sum_{\substack{\widetilde{F}_2 \in \mathcal{F}^{\sigma_2,U}_{\tilde{\ell}+d}[T_2] \\ I'(\widetilde{F}_2) \approx \widetilde{F}_1}} p(\widetilde{F}_2, F_2) = p(\widetilde{F}_1, I'(F_2))$$

for every fixed $F_2 \in \mathcal{F}^{\sigma_2,U}_{\ell+d}[T_2]$. And this once more follows from the fact that both sides represent the probability of the same event described as follows. Let $F_2 = (M,\theta)$; pick uniformly at random an $\tilde{\ell}$-element subset $\boldsymbol{V} \subseteq V(M)$ subject to the only condition $\boldsymbol{V} \cap \mathrm{im}(\theta) = \mathrm{im}(\theta) \setminus \theta(D)$. Then the promised event is simply "$I'(F_2|_{\boldsymbol{V}}) \approx \widetilde{F}_1$" (and the summation variable $\widetilde{F}_2$ corresponds to the isomorphism type of $F_2|_{\mathrm{im}(\theta)\cup\boldsymbol{V}}$).

We now prove that $\pi^{(U,I)}$ respects multiplication. Let

$$F_1^{(1)} \in \mathcal{F}^{\sigma_1}_{\ell_1}[T_1],\ F_1^{(2)} \in \mathcal{F}^{\sigma_1}_{\ell_2}[T_1]$$

and $\ell \geq \ell_1 + \ell_2 - k$. We have to show (in the algebra $\mathcal{A}^{\sigma_2}_u[T_2]$) that

$$\pi^{(U,I)}(F_1^{(1)})\pi^{(U,I)}(F_1^{(2)}) = \sum_{F_1 \in \mathcal{F}^{\sigma_1}_{\ell}[T_1]} p(F_1^{(1)}, F_1^{(2)}; F_1)\pi^{(U,I)}(F_1).$$

By Lemma 2.2 and already proven fact $\pi^{(U,I)}(\mathcal{K}^{\sigma_1}[T_1]) = 0$, we may assume $\ell = \ell_1 + \ell_2 - k$. In this case the normalizing terms $\frac{1}{u^{\ell-k}}, \frac{1}{u^{\ell_1-k}}, \frac{1}{u^{\ell_2-k}}$ cancel out, and we again expand both parts as linear combinations of flags $F_2 \in \mathcal{F}^{\sigma_2,U}_{\ell+d}[T_2]$ and compare the coefficients in front of any particular F_2. This leaves us with proving the identity

$$\sum_{\substack{F_2^{(1)} \in \mathcal{F}^{\sigma_2,U}_{\ell_1+d}[T_2] \\ I'(F_2^{(1)}) \approx F_1^{(1)}}} \ \sum_{\substack{F_2^{(2)} \in \mathcal{F}^{\sigma_2,U}_{\ell_2+d}[T_2] \\ I'(F_2^{(2)}) \approx F_1^{(2)}}} p(F_2^{(1)}, F_2^{(2)}; F_2) = p(F_1^{(1)}, F_1^{(2)}; I'(F_2)).$$

This is true for (by now) standard reason: if $F_2 = (M,\theta)$, then both sides calculate the probability of the event "$I'(F_2|_{\boldsymbol{V^{(1)}}}) \approx F_1^{(1)}$ and $I'(F_2|_{\boldsymbol{V^{(2)}}}) \approx F_1^{(2)}$, where $(\boldsymbol{V^{(1)}}, \boldsymbol{V^{(2)}})$ is a random sunflower in $V(M)$ with center $\mathrm{im}(\theta)$ and petals of sizes $\ell_1 + d, \ell_2 + d$, respectively."

$\pi^{(U,I)}(1_{\sigma_1}) = 1_{\sigma_2}$ is obvious.

The proof of Theorem 2.6 is complete. $\dashv$

Before continuing with concrete examples, let us note one simple sufficient condition for (9). Given a type σ and two σ-flags F_1, F_2, represent them in the form $F_1 = (M_1, \theta)$, $F_2 = (M_2, \theta)$, where $V(M_1) \cap V(M_2) = \mathrm{im}(\theta)$, and define their *amalgam* $F_1 \sqcup_\sigma F_2$ as the structure (M, θ) with $V(M) \stackrel{\mathrm{def}}{=} V(M_1) \cup V(M_2)$ and $M \models P(v_1, \ldots, v_r)$ if and only if $\{v_1, \ldots, v_r\} \subseteq V(M_i)$ and $M_i \models P(v_1, \ldots, v_r)$ for some $i = 1, 2$. The theory T has the *amalgamation property* if the amalgam of every two σ-flags is also a σ-flag.

THEOREM 2.7. *Assume T has the amalgamation property. Then every type σ is non-degenerate, and $\mathcal{A}^\sigma$ is free* (*isomorphic to the algebra of polynomials in countably many variables*).

PROOF. Considering a given $\sigma \neq 0$ as an (unlabelled) model of T, and repeatedly taking its amalgam with itself $\sigma \sqcup_0 \sigma \sqcup_0 \ldots \sqcup_0 \sigma$ in type 0, we find arbitrarily large models of T that contain an induced copy of σ. This proves that σ is a non-degenerate.

For the second part, we associate to any σ-flag $F = (M, \theta)$ the (undirected) graph G_F with $V(G_F) \stackrel{\mathrm{def}}{=} V(M) \setminus \mathrm{im}(\theta)$ and

$$\begin{aligned} E(G_F) \stackrel{\mathrm{def}}{=} \{(v, w) \mid \exists P(x_1, \ldots, x_r) \in L\ \exists v_1, \ldots, v_r \in V(M) \\ (M \models P(v_1, \ldots, v_r) \wedge \{v, w\} \subseteq \{v_1, \ldots, v_r\})\}. \end{aligned}$$

Call F *connected* if G_F is so. Choose arbitrarily $F_0 \in \mathcal{F}^{|\sigma|}_{\sigma+1}$, and let $\widetilde{\mathcal{F}}^\sigma$ be the set of all connected flags *except for* 1_σ *and* F_0. We are going to prove that $\mathcal{A}^\sigma$ *is freely generated by* $\widetilde{\mathcal{F}}^\sigma$.

Note first that every σ-flag F allows a unique (up to isomorphism and permutations of components) decomposition $F = F_1 \sqcup_\sigma \ldots \sqcup_\sigma F_t$ into an amalgam of connected non-trivial flags. It is important for the following that if we have a flag embedding $F_1 \sqcup_\sigma \ldots \sqcup_\sigma F_t \longrightarrow F'_1 \sqcup_\sigma \ldots \sqcup_\sigma F'_{t'}$, and $F_1, \ldots, F_t$ are connected then every F_i is completely mapped into a *single* F'_j (different F_i can be mapped into the same F'_j).

We order all connected flags F arbitrarily, but in such a way that if F_1 has fewer vertices than F_2 then $F_1 < F_2$ and such that F_0 is the second least (after 1_σ) flag in this ordering. Extend this ordering to arbitrary σ-flags anti-lexicographically. More precisely, let $F = F_1 \sqcup_\sigma \ldots \sqcup_\sigma F_t$ and $F' = F'_1 \sqcup_\sigma \ldots \sqcup_\sigma F'_{t'}$ be decompositions of F and F' into amalgams of connected flags. We find the largest connected flag $\widetilde{F}$ appearing in these decompositions with different multiplicities and let $F < F'$ if the multiplicity of $\widetilde{F}$ in the decomposition of F is smaller than its multiplicity in the decomposition of F'.

Our claim almost immediately follows from the following easily checkable properties of the ordering $\leq$:

1. $\leq$ is consistent with amalgamation, that is $F_1 \leq F'_1$ and $F_2 \leq F'_2$ imply $F_1 \sqcup_\sigma F_2 \leq F'_1 \sqcup_\sigma F'_2$.

2. Let $F_i \in \widetilde{\mathcal{F}}^\sigma_{\ell_i} (\stackrel{\mathrm{def}}{=} \widetilde{\mathcal{F}}^\sigma \cap \mathcal{F}^\sigma_{\ell_i})$ for $1 \leq i \leq t$ and $\ell = \ell_1 + \cdots + \ell_t - |\sigma|(t-1) + d$, where $d \geq 0$. Then
$$p(F_1, \ldots, F_t; F_1 \sqcup_\sigma \ldots \sqcup_\sigma F_t \sqcup_\sigma \underbrace{F_0 \sqcup_\sigma \ldots \sqcup_\sigma F_0}_{d \text{ times}}) > 0.$$
Moreover:
   (a) $F_1 \sqcup_\sigma \ldots \sqcup_\sigma F_t \sqcup_\sigma \underbrace{F_0 \sqcup_\sigma \ldots \sqcup_\sigma F_0}_{d \text{ times}}$ is the *minimal* flag $F \in \mathcal{F}^\sigma_\ell$ for which $p(F_1, \ldots, F_t; F) > 0$;
   (b) for any other system $F'_i \in \widetilde{\mathcal{F}}^\sigma_{\ell'_i}$ $(1 \leq i \leq t')$ with $\ell = \ell'_1 + \cdots + \ell'_t - |\sigma|(t'-1) + d'$, $d' \geq 0$ such that
$$p(F'_1, \ldots, F'_t; F_1 \sqcup_\sigma \ldots \sqcup_\sigma F_t \sqcup_\sigma \underbrace{F_0 \sqcup_\sigma \ldots \sqcup_\sigma F_0}_{d \text{ times}}) > 0,$$
   we have
$$F'_1 \sqcup_\sigma \ldots \sqcup_\sigma F'_t \sqcup_\sigma \underbrace{F_0 \sqcup_\sigma \ldots \sqcup_\sigma F_0}_{d' \text{ times}} < F_1 \sqcup_\sigma \ldots \sqcup_\sigma F_t \sqcup_\sigma \underbrace{F_0 \sqcup_\sigma \ldots \sqcup_\sigma F_0}_{d \text{ times}}.$$

To see this, let $\mathcal{S}$ be the sub-algebra generated by $\widetilde{\mathcal{F}}^\sigma$. Then, since $\widetilde{\mathcal{F}}^\sigma_{|\sigma|+1} = \mathcal{F}^\sigma_{|\sigma|+1} \setminus \{F_0\}$, and due to the relation $\sum_{F \in \mathcal{F}^\sigma_{|\sigma|+1}} F = 1$ (which is a special case of (4)), we also have $F_0 \in \mathcal{S}$. For $\ell > |\sigma| + 1$ and $F \in \mathcal{F}^\sigma_\ell$, we prove $F \in \mathcal{S}$ by induction on ℓ. For a fixed ℓ we apply the reverse induction on F, and Property 2a takes care of the inductive step.

In the opposite direction, let $p(z_F | F \in \widetilde{\mathcal{F}}^\sigma)$ be any non-zero polynomial and let ℓ be a sufficiently large integer. Then for any monomial $z_{F_1} \ldots z_{F_t}$ occurring in p with a non-zero coefficient, we form the corresponding flag $F_1 \sqcup_\sigma \ldots \sqcup_\sigma F_t \sqcup_\sigma \underbrace{F_0 \sqcup_\sigma \ldots \sqcup_\sigma F_0}_{d \text{ times}} \in \mathcal{F}^\sigma_\ell$, and choose the *minimal* $F_{\min}$ of all these flags. Property 2b then implies that the expansion of $p(\vec{F})$ as a linear combination of $F \in \mathcal{F}^\sigma_\ell$ has non-zero coefficient in front of $F_{\min}$. Therefore, $p(\vec{F}) \neq 0$ in $\mathcal{A}^\sigma$. $\dashv$

This theorem can be applied e.g. to the theories of (directed and undirected) graphs and hypergraphs, undirected graphs with $\omega(G) \leq k$ for a fixed k, directed graphs without $\vec{C}_3$ or 3-hypergraphs without a complete subgraph on 4 vertices. Theorem 2.3 implies that if two theories T_1 and T_2 are isomorphic with respect to open interpretations such that $U(x) \equiv \top$ then the corresponding algebras $\mathcal{A}^\sigma$ are also isomorphic. This observation allows us to extend Theorem 2.7 to more theories such as the theory of graphs with *independence* number $\leq k$ or 3-hypergraphs in which every 4 vertices contain at least one edge.

2.3.1. *Algebras of constants: upward operator.* In the general set-up of Definition 4 assume that $T_1 = T_2 = T$, and that I acts trivially on predicate symbols from L. Then, denoting σ_2 simply by σ, we see that $\sigma_1 = \sigma|_\eta$, where η was also defined in Definition 4.

Assume now additionally that U is trivial (that is, $U(x) \equiv \top$). Then $u = 1$, the localization $\mathcal{A}^\sigma_u$ coincided with $\mathcal{A}^\sigma$, and we get an algebra homomorphism $\mathcal{A}^{\sigma|_\eta} \longrightarrow \mathcal{A}^\sigma$ that we will denote by $\pi^{\sigma,\eta}$. In view of its extremal importance in

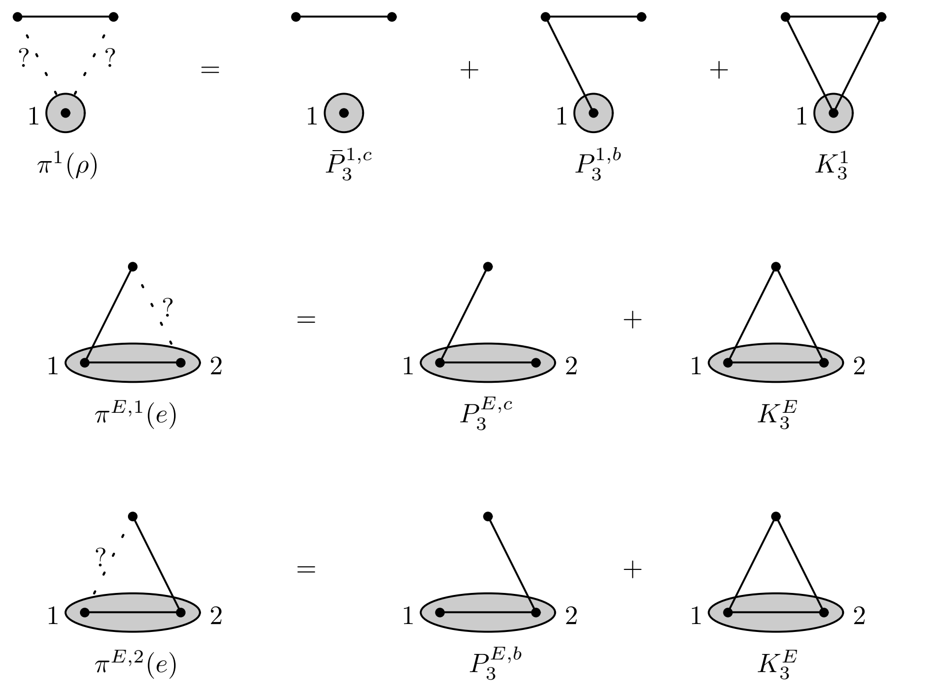

FIGURE 3. Algebras of constants

our framework, we give its independent description and prove two extra properties (the first of them also explains why we consider the elements in the subalgebra $\mathrm{im}(\pi^{\sigma,\eta})$ as "$\sigma|_\eta$-constants").

As in Section 2.2, let σ be a non-degenerate type of size k, $\eta : [k'] \longrightarrow [k]$ be an injective mapping, and $\sigma|_\eta$ be the induced type of size k'. As above, let $D \stackrel{\text{def}}{=} [k] \setminus \mathrm{im}(\eta)$ and $d \stackrel{\text{def}}{=} k - k' = |D|$. For a σ-flag $F = (M, \theta)$, we let

$$F{\downarrow_\eta} \stackrel{\text{def}}{=} F|_\eta - \theta(D)$$

(thus, the only difference between $F{\downarrow_\eta}$ and $F|_\eta$ is that we not only unlabel vertices in $\theta(D)$ but actually *remove* them from the flag). If $F \in \mathcal{F}_\ell^\sigma$ then $F{\downarrow_\eta} \in \mathcal{F}_{\ell-d}^{\sigma|_\eta}$.

In this notation, the homomorphism $\pi^{\sigma,\eta} : \mathcal{A}^{\sigma|_\eta} \longrightarrow \mathcal{A}^\sigma$ can be calculated as

$$\pi^{\sigma,\eta}(F) \stackrel{\text{def}}{=} \sum \left\{ \widehat{F} \in \mathcal{F}_{\ell+d}^\sigma \;\middle|\; \widehat{F}\Big\downarrow_\eta = F \right\} \; (F \in \mathcal{F}_\ell^{\sigma|_\eta}). \tag{10}$$

We will also abbreviate $\pi^{\sigma,0}$ to π^σ; thus, π^σ is a homomorphism from $\mathcal{A}^0$ to $\mathcal{A}^\sigma$.

EXAMPLE 8 (undirected graphs, see Figure 3).
$\pi^1(\rho) = P_3^{1,b} + \bar{P}_3^{1,c} + K_3^1$. $\pi^{E,1}(e) = K_3^E + P_3^{E,c}$, whereas $\pi^{E,2}(e) = K_3^E + P_3^{E,b}$.

THEOREM 2.8. **a)** *For every $f \in \mathcal{A}^{\sigma|_\eta}$ and $g \in \mathcal{A}^\sigma$,*

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

In particular $(g = 1)$,

$$[\![\pi^{\sigma,\eta}(f)]\!]_{\sigma,\eta} = f[\![1_\sigma]\!]_{\sigma,\eta} = q_{\sigma,\eta}(1_\sigma) \cdot (\sigma,\eta) \cdot f.$$

b) (chain rule) *If additionally* $\eta' : [k''] \longrightarrow [k']$ *is another injective mapping for some* $k'' \leq k'$ *then* $\pi^{\sigma,\eta\eta'}(f) = \pi^{\sigma,\eta}(\pi^{\sigma|_\eta,\eta'}(f))$.

Proof. **a).** By linearity, we may assume that $f = F_1 \in \mathcal{F}^{\sigma|_\eta}_{\ell_1}$ and $g = F_2 \in \mathcal{F}^{\sigma}_{\ell_2}$ are flags. As in similar previous proofs, we expand both parts as linear combinations of $F \in \mathcal{A}^{\sigma|_\eta}_{\ell}$ for sufficiently large ℓ and compare coefficients in front of any particular F. This leads us to the identity

$$\sum_{\substack{\widehat{F_1} \in \mathcal{F}^{\sigma}_{\ell_1+d} \\ \widehat{F_1}\downarrow_\eta \approx F_1}} \sum_{\substack{\widehat{F} \in \mathcal{F}^{\sigma}_{\ell} \\ \widehat{F}|_\eta \approx F}} p(\widehat{F_1}, F_2; \widehat{F}) q_{\sigma,\eta}(\widehat{F}) = p(F_1, F_2|_\eta; F) q_{\sigma,\eta}(F_2)$$

to be proven.

The event whose probability is calculated by both sides is constructed as follows. Let $F = (M, \theta')$ $(\theta' : [k'] \longrightarrow M)$. Pick in $V(M)$, uniformly at random, a sunflower $(\boldsymbol{V_1}, \boldsymbol{V_2})$ with center $\mathrm{im}(\theta')$ and petals of sizes ℓ_1, ℓ_2, respectively. Then pick at random an injective mapping $\boldsymbol{\theta} : [k] \longrightarrow \boldsymbol{V_2}$ subject to the condition $\boldsymbol{\theta}\eta = \theta'$. The desired event is "$\boldsymbol{\theta}$ is a model embedding $\sigma \longrightarrow M$, $F|_{\boldsymbol{V}_1} \approx F_1$ and $(M|_{\boldsymbol{V}_2}, \boldsymbol{\theta}) \approx F_2$" (in the left-hand side, the summation variable $\widehat{F_1}$ corresponds to the isomorphism type of $(M|_{\boldsymbol{V_1} \cup \mathrm{im}(\boldsymbol{\theta})}, \boldsymbol{\theta})$, and $\widehat{F}$ corresponds to the isomorphism type of $(M, \boldsymbol{\theta})$).

b) is obvious. ⊣

2.3.2. *Inductive arguments.* As in Section 2.3.1, let $T_1 = T_2 = T$ and assume that I acts identically on the predicates from L. But now we consider the opposite extreme case and assume that U represents the diagram of a single flag $F_0 \in \mathcal{F}^{\sigma}_{|\sigma|+1}$ (and, therefore, $u = F_0$). We will denote the resulting homomorphism $\mathcal{A}^{\sigma|_\eta} \longrightarrow \mathcal{A}^{\sigma}_{F_0}$ by $\pi^{F_0,\eta}$ and, again, when $\eta = 0$, abbreviate it to $\pi^{F_0} : \mathcal{A}^0 \longrightarrow \mathcal{A}^{\sigma}_{F_0}$.

Theorem 2.8 a) indicates that the operators $\pi^{\sigma,\eta}$ do not produce any non-trivial relations (they will be used mostly for structural purposes). This is no longer true for their relativized versions $\pi^{F_0,\eta}$, and, in fact, in combination with extremality conditions, we get a powerful tool for representing various inductive arguments in our framework. This will be further explored in Sections 3 and 4.

Example 9 (undirected graphs). $\pi^e(K_\ell) = K^1_{\ell+1}/e^\ell$; $\pi^e(\bar\rho) = P_3^{1,c}/e^2$ etc.

2.3.3. *"Genuine" interpretations.* And, finally, we consider the case when the theories T_1, T_2 are different (and typically in different languages), and we also assume that $U(x) \equiv \top$. Then we get a tool for transferring results about combinatorial objects of one kind to objects of another kind. There are two variants of this technique, both extensively used in the literature.

Global interpretations. By this we mean that $I : T_1^{\sigma_1} \rightsquigarrow T_2^{\sigma_2}$ is actually obtained from an interpretation of T_1 in T_2 (in other words, the formulas $I(P)$ $(P \in L_1)$ contain only predicate symbols from L_2 and no constants). Then for any non-degenerate type σ of the theory T_2 we have a uniquely defined non-degenerate type $I(\sigma)$ of the theory T_1 and a homomorphism $\mathcal{A}^{I(\sigma)}[T_1] \longrightarrow \mathcal{A}^\sigma[T_2]$. For example, the *orientation-erasing interpretation* $I : T_{Graph} \rightsquigarrow T_{Digraph}$ given by $I(E)(v_1, v_2) \stackrel{\text{def}}{=} E(v_1, v_2) \vee E(v_2, v_1)$ gives rise to homomorphisms $\mathcal{A}^{\sigma_1}[T_{Graph}] \longrightarrow \mathcal{A}^{\sigma_2}[T_{Digraph}]$ whenever σ_2 is any orientation of σ_1, and via this interpretation we can use all theorems proved about undirected graphs in the directed case. Another pivotal example is when T_2 is an extension of T_1 in the same language and I is the identity (for example, let $T_1 = T_{Graph}$ and let T_2 be obtained by forbidding certain induced subgraphs). In this case $\mathcal{A}^\sigma[T_2]$ is a quotient algebra of $\mathcal{A}^\sigma[T_1]$, and $\pi^{(U,I)}$ is the natural homomorphism onto it.

Local interpretations. Open formulas $I(P)$ may actually contain new constants added in the theory $T_2^{\sigma_2}$. An excellent example of this sort is the link construction ingeniously employed in [CaF].

Example 10. Let $T_1 = T_{Graph}$, and $T_2 = T_{3-Hypergraph}$ be the theory of 3-regular hypergraphs. Let $\sigma_1 = 0$, $\sigma_2 = 1$, and define the interpretation $I : T_{Graph} \rightsquigarrow T^1_{3-Hypergraph}$ by $I(E)(v_1, v_2) \equiv E(c_1, v_1, v_2)$. Then we get the *link homomorphism* $\mathcal{A}^0[T_{Graph}] \longrightarrow \mathcal{A}^1[T_{3-Hypergraph}]$.

§3. Semantics. The semantic model that we consider as the "base" one (and that will be considerably enhanced in Section 3.2) is suggested by the fact (Lemma 2.3) that when the size of the target flag F grows to infinity, the mappings $p(\cdot, F)$ look more and more like homomorphisms from $\mathcal{A}^\sigma$ to $\mathbb{R}$.

Definition 5. Let σ be a non-degenerate type. $\mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ is the set of all algebra homomorphisms from $\mathcal{A}^\sigma$ to $\mathbb{R}$ (in particular, $\phi(1_\sigma) = 1$ for every $\phi \in \mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$), and $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ consists of all those $\phi \in \mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ for which $\forall F \in \mathcal{F}^\sigma (\phi(F) \geq 0)$. Note that in $\mathcal{A}^\sigma$ we have

$$\sum_{F \in \mathcal{F}^\sigma_\ell} F = 1_\sigma \ (\ell \geq |\sigma|) \tag{11}$$

(which is a special case of (4)), therefore for every $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ and every $F \in \mathcal{F}^\sigma$ we actually have $\phi(F) \in [0, 1]$. The *semantic cone* $\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma) \subseteq \mathcal{A}^\sigma$ consists of all those $f \in \mathcal{A}^\sigma$ for which $\forall \phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ $(\phi(f) \geq 0)$.

$\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$ represents "true statements" in our framework, and we will indeed see in Section 3.1 that this is exactly the set of all "polynomial" relations in extremal combinatorics that hold asymptotically (Corollary 3.4).

Remark 3. The semantic cone can be of course defined in a more general situation, when $\mathcal{A}$ is an arbitrary commutative algebra and $\mathcal{F} \subset \mathcal{A}$ is an arbitrary subset. This is one of the main objects of study in real algebraic geometry [BCR]. When $\mathcal{A}$ is finitely generated, a constructive description of $\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$ is given by the *Positivstellensatz Theorem* [BCR, Section 4.4]. Unfortunately, flag algebras are not finitely generated, and this seems to be a very serious obstacle when trying to apply the corresponding techniques to our case.

Also, as we have seen in Theorem 2.7, for many interesting theories flag algebras are free and, therefore, $\mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ is totally trivial. It is the additional structure defined by the conditions $\phi(F) \geq 0$ that makes the objects $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ enormously complex.

The ordinary cone $\mathcal{C}(\mathcal{F}^\sigma)$ consists of all linear combinations with non-negative coefficients of elements of the form

$$f^2F_1F_2\dots F_h \quad (f\in\mathcal{A}^\sigma;\ F_1,\dots,F_h\in\mathcal{F}^\sigma),$$

and w.l.o.g. we may assume here $h=1$. Clearly, $\mathcal{C}(\mathcal{F}^\sigma)\subseteq\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$ so these elements represent the first non-trivial "axioms" in our calculus. And the operators introduced in Section 2 serve as the basic "inference rules":

Theorem 3.1. **a)** *Let σ be a non-degenerate type of size k, and $\eta:[k']\longrightarrow[k]$ be an injective mapping. Then*

$$[\![\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)]\!]_{\sigma,\eta}\subseteq\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^{\sigma|_\eta}).$$

b) *In the set-up of Definition 4, let $f\in\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^{\sigma_1}[T_1])$. Then for every integer ℓ such that*

$$u^\ell\cdot\pi^{(U,I)}(f)\in\mathbb{R}\mathcal{F}^{\sigma_2,U}[T_2], \tag{12}$$

we have $u^\ell\cdot\pi^{(U,I)}(f)\in\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^{\sigma_2}[T_2])$.

In particular, in the notation of Section 2.3.1,

$$\pi^{\sigma,\eta}(\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^{\sigma|_\eta}))\subseteq\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma), \tag{13}$$

and in the notation of Section 2.3.2,

$$\forall f\in\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^{\sigma|_\eta})\forall\ell\in\mathbb{Z}$$
$$(F_0^\ell\cdot\pi^{F_0,\eta}(f)\in\mathbb{R}\mathcal{F}^{\sigma,F_0}\Longrightarrow F_0^\ell\cdot\pi^{F_0,\eta}(f)\in\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)).$$

Although it is very easy to give an ad hoc proof of part a) of this theorem, we prefer to defer it until Section 3.2 (at which point it will become completely obvious).

Part b) is obvious already. Let $\phi\in\mathrm{Hom}^+(\mathcal{A}^{\sigma_2}[T_2],\mathbb{R})$.

Case 1. $\phi(u)=0$.

$\phi(F)=0$ for every flag $F\in\mathcal{F}^{\sigma_2,U}[T_2]$ (as $F\leq_{\sigma_2}Cu$ for some $C>0$). Therefore, $\phi(u^\ell\cdot\pi^{(U,I)}(f))=0$ due to assumption (12).

Case 2. $\phi(u)>0$.

ϕ can be extended to a homomorphism $\phi:\mathcal{A}_u^{\sigma_2}[T_2]\longrightarrow\mathbb{R}$, and $\phi(u^\ell\cdot\pi^{(U,I)}(f))=\phi(u)^\ell(\phi\pi^{(U,I)})(f)\geq 0$ since $\phi\pi^{(U,I)}\in\mathrm{Hom}^+(\mathcal{A}^{\sigma_1}[T_1],\mathbb{R})$ and $f\in\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^{\sigma_1}[T_1])$.

Definition 6. For $f,g\in\mathcal{A}^\sigma$, let $f\leq_\sigma g$ mean $(g-f)\in\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$. This is a partial preorder on $\mathcal{A}^\sigma$ that will be sometimes denoted simply by $f\leq g$ whenever σ is clear from the context.

Felix [Fel] and Podolski [Pod] independently proved that $\leq_\sigma$ is a partial order at least in the theory T_{Graph} (or, in other words, if $\forall\phi\in\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})(\phi(f)=0)$ then $f=0$). Their result indicates that the set of relations (4) is complete.

3.1. Convergent sequences. *The material in this subsection is borrowed, with minimal adaptations, from [LoSz].*

$\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})$ has been chosen as our "base" semantics mainly due to its simplicity and convenience (that was already witnessed by the proof of Theorem 3.1 b)). In this section we define the "intended" semantics and prove their equivalence.

Convention. For $F\in\mathcal{F}_\ell^\sigma$ and $F'\in\mathcal{F}_{\ell'}^\sigma$ with $\ell>\ell'$, we let $p(F,F')\stackrel{\mathrm{def}}{=}0$.

Definition 7. Let σ be a non-degenerate type. An infinite sequence

$$F_1, F_2, \dots, F_n, \dots$$

with $F_n \in \mathcal{F}^\sigma_{\ell_n}$ is *increasing* if $\ell_1 < \ell_2 < \dots < \ell_n < \dots$. An increasing sequence $\{F_n\}$ of σ-flags is *convergent* if $\lim_{n\to\infty} p(F, F_n)$ exists for every $F \in \mathcal{F}^\sigma$.

Every flag $F_0 \in \mathcal{F}^\sigma$ defines the point p^{F_0} in the (infinite-dimensional) space $[0,1]^{\mathcal{F}^\sigma}$ given by $p^{F_0}(F) \stackrel{\text{def}}{=} p(F, F_0)$, that will be also sometimes viewed as a linear functional on $\mathbb{R}\mathcal{F}^\sigma$. We endow $[0,1]^{\mathcal{F}^\sigma}$ with *product topology* (aka *pointwise convergence topology*).

An increasing sequence of σ-flags $\{F_n\}$ is convergent if and only if the sequence p^{F_n} is convergent in $[0,1]^{\mathcal{F}^\sigma}$.

Theorem 3.2. *Every increasing sequence of σ-flags contains a convergent subsequence.*

Proof. $[0,1]^{\mathcal{F}^\sigma}$ is compact. $\dashv$

From now on we will often be considering $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ as a subset in $[0,1]^{\mathcal{F}^\sigma}$. This subset is defined by countably many polynomial equations (4), (5) in the coordinate functions $x \mapsto x(F)$ $(F \in \mathcal{F}^\sigma,\ x \in [0,1]^{\mathcal{F}^\sigma})$. Therefore, it is a closed subset, and, as such, is also compact.

Theorem 3.3 ([LoSz]). **a)** *For every convergent sequence $\{F_n\}$ of σ-flags, $\lim_{n\to\infty} p^{F_n} \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$.*

b) *Conversely, every element $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ can be represented in the form $\lim_{n\to\infty} p^{F_n}$ for a convergent sequence $\{F_n\}$ of σ-flags.*

Proof. **a).** Every p^{F_n} satisfies the relation (4) as long as $\ell_n \geq \ell$. By Lemma 2.3 (and Lemma 2.2), p^{F_n} also satisfies (5) within the additive term $O(1/\ell_n)$. Taking the limit, we see that $\lim_{n\to\infty} p^{F_n}$ satisfies (5) exactly.

b). Let $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. (11) implies that for every ℓ the quantities $\phi(F)$ $(F \in \mathcal{F}^\sigma_\ell)$ define a probability measure on $\mathcal{F}^\sigma_\ell$. Consider the product measure on $\prod_{\ell=|\sigma|}^\infty \mathcal{F}^\sigma_\ell$, and choose an increasing sequence $\{\boldsymbol{F_n}\}$, where $\boldsymbol{F_n} \in \mathcal{F}^\sigma_{n^2}$, at random according to this measure (thus, $\boldsymbol{F_n}$ are independent for different n). It is sufficient to show that $\mathbf{P}\left[\lim_{n\to\infty} p^{\boldsymbol{F_n}} = \phi\right] = 1$. And, since $\mathcal{F}^\sigma$ is countable, it is enough to prove that for every *fixed* $F \in \mathcal{F}^\sigma_\ell$ and every *fixed* $\epsilon > 0$,

$$\mathbf{P}[\exists n_0 \forall n \geq n_0 |p(F, \boldsymbol{F_n}) - \phi(F)| \leq \epsilon] = 1. \tag{14}$$

For $n^2 \geq \ell$ we have

$$\mathbf{E}[p(F, \boldsymbol{F_n})] = \sum_{F_n \in \mathcal{F}^\sigma_{n^2}} p(F, F_n)\phi(F_n) = \phi(F)$$

(since ϕ satisfies (4)). Also,

$$\begin{aligned} Var[p(F, \boldsymbol{F_n})] &= \mathbf{E}\left[p(F, \boldsymbol{F_n})^2\right] - \phi(F^2) \\ &= \sum_{F_n \in \mathcal{F}^\sigma_{n^2}} p(F, F_n)^2 \phi(F_n) - \sum_{F_n \in \mathcal{F}^\sigma_{n^2}} p(F, F; F_n)\phi(F_n) \end{aligned}$$

(by (5)) and $|p(F, F_n)^2 - p(F, F; F_n)| \leq O(1/n^2)$ (uniformly over all choices of F_n) by Lemma 2.3. Thus, $Var[p(F, \boldsymbol{F_n})] \leq O(1/n^2)$, and (14) follows by a standard application of Chebyshev's inequality and Borel-Cantelli lemma. $\dashv$

The following immediate corollary of Theorem 3.3 is a rigorous formalization of why a large fragment of asymptotic extremal combinatorics can be identified with the study of homomorphisms from $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$.

COROLLARY 3.4. *Let $F_1, \ldots, F_h \in \mathcal{F}^\sigma$ be fixed σ-flags, $D \subseteq \mathbb{R}^h$ and $f : D \longrightarrow \mathbb{R}$ be a continuous function.*

If D is closed then

(15)
$$\left.\begin{array}{l} \forall \phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})((\phi(F_1), \ldots, \phi(F_h)) \in D \\ \quad \Longrightarrow f(\phi(F_1), \ldots, \phi(F_h)) \geq 0) \end{array}\right\}$$

implies

(16)
$$\left.\begin{array}{l} \liminf\limits_{\ell \to \infty} \min\{f(p(F_1, F), \ldots, p(F_h, F)) \mid F \in \mathcal{F}^\sigma_\ell \\ \quad \wedge (p(F_1, F), \ldots, p(F_h, F)) \in D\} \geq 0, \end{array}\right\}$$

and if D is open then (16) *implies* (15). *In particular, if $D = \mathbb{R}^h$ and f is a polynomial, then $f(F_1, \ldots, F_h) \in \mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$ if and only if*

$$\liminf_{\ell \to \infty} \min_{F \in \mathcal{F}^\sigma_\ell} f(p(F_1, F), \ldots, p(F_h, F)) \geq 0.$$

3.2. Ensembles of random homomorphisms. *Throughout this section, σ_0 will be a fixed non-degenerate type, and k_0 will denote its size.*

We will show that for every pair (σ, η) with $\sigma|_\eta = \sigma_0$ every homomorphism $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_0}, \mathbb{R})$ with $\phi_0(\sigma, \eta) > 0$ gives rise, in a canonical way, to a *random* homomorphism $\boldsymbol{\phi^{\sigma,\eta}}$ chosen according to some (uniquely defined!) probability measure on Borel subsets of $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. We will also establish some natural properties of the resulting ensemble $\{\boldsymbol{\phi^{\sigma,\eta}}\}$ of random homomorphisms suggested by obvious analogies with the discrete case. This becomes an indispensable tool in our framework when it comes to "discontinuous", "case-analysis" arguments like "consider the set of all vertices of degree at most cn". It is also worth noting that, due to uniqueness, these ensembles do not provide an independent semantics. Rather, they serve as convenient tools for extracting more useful information from our base objects $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_0}, \mathbb{R})$. And it should be also remarked that *this convenience does not have straightforward analogies in the discrete case*. Say, if we know exactly the densities of all subgraphs with at most 10 vertices in a given large graph, we still do not have enough information to determine its minimal degree. We will further elaborate on this in the concluding Section 6.

DEFINITION 8. Every pair (σ, η), where σ is a non-degenerate type of size $k \geq k_0$ and $\eta : [k_0] \longrightarrow [k]$ is an injective mapping such that $\sigma|_\eta = \sigma_0$, will be called an *extension of* σ_0. Assume that $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_0}, \mathbb{R})$ has the property[2] $\phi_0((\sigma,\eta)) > 0$. Let $\mathcal{B}^\sigma$ consist of all Borel subsets of $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. A probability measure $\mathbf{P}^{\sigma,\eta}$ on $\mathcal{B}^\sigma$ *extends* ϕ_0 if for any $f \in \mathcal{A}^\sigma$,

$$\int_{\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})} \phi(f)\mathbf{P}^{\sigma,\eta}(d\phi) = \frac{\phi_0([\![f]\!]_{\sigma,\eta})}{\phi_0(\langle\sigma,\eta\rangle)}, \tag{17}$$

where (to improve readability) we have introduced the notation

$$\langle\sigma,\eta\rangle \stackrel{\mathrm{def}}{=} [\![1_\sigma]\!]_{\sigma,\eta} = q_{\sigma,\eta}(1_\sigma)\cdot(\sigma,\eta) \in \mathcal{A}^{\sigma|_\eta}$$

(cf. Theorem 2.5 b) and Remark 2).

THEOREM 3.5. *For $\sigma_0, \sigma, \eta, \phi_0$ as in Definition 8 there exists a unique extension $\mathbf{P}^{\sigma,\eta}$ of ϕ_0.*

Before proving this theorem, we need some background from analysis, topology and (higher) probability theory (this material will also be used elsewhere in this section). Recall that $[0,1]^{\mathcal{F}^\sigma}$ endowed with product topology is compact. This topology is metrizable (e.g. by the metric

$$d(x,y) \stackrel{\mathrm{def}}{=} \sum_{n=0}^{\infty} 2^{-n}|x(F_n) - y(F_n)| \ (x, y \in [0,1]^{\mathcal{F}^\sigma}),$$

where $\{F_1, \ldots, F_n, \ldots\}$ is an arbitrary fixed enumeration of σ-flags). Therefore, its closed subspace $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ is also compact and metrizable. For a topological space X, $C(X)$ is the set of all $\mathbb{R}$-valued continuous functions on X. If X is compact then every $f \in C(X)$ is automatically bounded.

PROPOSITION 3.6 (Tietze Extension Theorem). *If X is a metrizable space and Y its closed subspace, then every function in $C(Y)$ can be extended to a function in $C(X)$.*

Every element $f \in \mathcal{A}^\sigma$ can be alternatively viewed as an element of $C(\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R}))$ (given by $f(\phi) \stackrel{\mathrm{def}}{=} \phi(f)$).

PROPOSITION 3.7. *$\mathcal{A}^\sigma$ is dense in $C(\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R}))$ with respect to ℓ_∞* (!) *norm.*

PROOF. Since $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ is compact and metrizable, this is a special case of the Stone-Weierstrass theorem. ⊣

PROPOSITION 3.8. *Let X be a metrizable topological space, $\mathcal{B}$ be the σ-algebra of its Borel subsets, and $\mathbf{P}, \mathbf{Q}$ be two probabilistic measures on $\mathcal{B}$ such that for any $f \in C(X)$,*

$$\int_X f(x)\mathbf{P}(dx) = \int_X f(x)\mathbf{Q}(dx).$$

Then $\mathbf{P}$ and $\mathbf{Q}$ coincide.

---

[2] Recall from Section 2.2 that by ignoring labels in σ, (σ,η) can be considered simply as a σ_0-flag.

A sequence $\{\mathbf{P}_n\}$ of probability measures on Borel subsets of a topological space X *weakly converges* to another probability measure $\mathbf{P}$ on the same space if for every $f \in C(X)$,

$$\lim_{n\to\infty} \int_X f(x)\mathbf{P}_n(dx) = \int_X f(x)\mathbf{P}(dx).$$

PROPOSITION 3.9. *Let X be a metrizable space, and $\{\mathbf{P}_n\}, \mathbf{P}$ be probability measures on the σ-algebra $\mathcal{B}$ of its Borel sets. Then the following are equivalent:*

- **a)** *$\{\mathbf{P}_n\}$ weakly converges to $\mathbf{P}$.*
- **b)** *For every $A \in \mathcal{B}$ with $\mathbf{P}(\partial A) = 0$ (∂A is the boundary of A),*

$$\lim_{n\to\infty} \mathbf{P}_n(A) = \mathbf{P}(A).$$

- **c)** *For every open A,*

$$\liminf_{n\to\infty} \mathbf{P}_n(A) \geq \mathbf{P}(A).$$

Now we recall a theorem due to Prohorov, but in a slightly less general form (the full version uses some relaxed notion of compactness).

PROPOSITION 3.10 (Prohorov's theorem). *Every sequence of probability measures on the σ-algebra of Borel subsets of a metrizable separable compact space contains a weakly convergent subsequence.*

We note that both spaces $[0,1]^{\mathcal{F}^\sigma}$ and $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ satisfy all the properties required in this proposition.

We now proceed to the proof of Theorem 3.5. Fix an extension (σ, η) of σ_0, and let $k = |\sigma|$.

DEFINITION 9. For a σ_0-flag $F = (M, \theta') \in \mathcal{F}_\ell^{\sigma_0}$ with $p^F((\sigma,\eta)) > 0$ define a (discrete) probability measure $\mathbf{P}_F^{\sigma,\eta}$ on Borel subsets of $[0,1]^{\mathcal{F}^\sigma}$ as follows. Choose, uniformly at random, a model embedding $\boldsymbol{\theta} : \sigma \longrightarrow M$ consistent with θ' (that is, $\boldsymbol{\theta}\eta = \theta'$). Then, for Borel $A \subseteq [0,1]^{\mathcal{F}^\sigma}$, we let

$$\mathbf{P}_F^{\sigma,\eta}(A) \stackrel{\text{def}}{=} \mathbf{P}\left[p^{(M,\boldsymbol{\theta})} \in A\right].$$

LEMMA 3.11. *Let $\{F_n\}$ be a convergent sequence of σ_0-flags, and $\phi_0 \stackrel{\text{def}}{=} \lim_{n\to\infty} p^{F_n}$. Assume that $\phi_0((\sigma,\eta)) > 0$. Then for any flag $F \in \mathcal{F}_\ell^\sigma$,*

$$\lim_{n\to\infty} \int_{[0,1]^{\mathcal{F}^\sigma}} x(F)\mathbf{P}_{F_n}^{\sigma,\eta}(dx) = \frac{\phi_0([\![F]\!]_{\sigma,\eta})}{\phi_0(\langle\sigma,\eta\rangle)}.$$

PROOF. $\phi_0((\sigma,\eta)) > 0$ implies $p^{F_n}((\sigma,\eta)) > 0$ for sufficiently large n. By Definition 9, as long as $\ell_n \geq \ell$ we have

$$\int_{[0,1]^{\mathcal{F}^\sigma}} x(F)\mathbf{P}_{F_n}^{\sigma,\eta}(dx) = \mathbf{E}[p(F, (M, \boldsymbol{\theta}))], \tag{18}$$

where $F = (M, \theta)$ and $\boldsymbol{\theta}$ is constructed as in Definition 9. On the other hand, $p^{F_n}([\![F]\!]_{\sigma,\eta})$ is the probability that a random injective mapping $[k] \longrightarrow M$ consistent with θ on $\mathrm{im}(\eta)$ defines a model embedding and the resulting σ-flag is isomorphic to F. Thus,

$$\int_{[0,1]^{\mathcal{F}^\sigma}} x(F)\mathbf{P}^{\sigma,\eta}_{F_n}(dx) = \frac{p^{F_n}(\llbracket F\rrbracket_{\sigma,\eta})}{p^{F_n}(\langle\sigma,\eta\rangle)}. \tag{19}$$

Taking the limit proves Lemma 3.11. ⊣

Proof of Theorem 3.5, existence.
By Theorem 3.3, there exists a convergent sequence $\{F_n\}$ of σ_0-flags with $\lim_{n\to\infty} p^{F_n} = \phi_0$. By Proposition 3.10, we can find in it a subsequence such that the corresponding measures $\mathbf{P}^{\sigma,\eta}_{F_n}$ weakly converge[3] to some probability measure $\mathbf{P}^{\sigma,\eta}$ on $[0,1]^{\mathcal{F}^\sigma}$. For this limit measure, (17) readily follows from Lemma 3.11, and we only have to check that $\mathbf{P}^{\sigma,\eta}$ is concentrated on homomorphisms; or, in other words, that $\mathbf{P}^{\sigma,\eta}[\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})] = 1$.

Indeed, $\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})$ is an algebraic subset of $[0,1]^{\mathcal{F}^\sigma}$ defined by a countable system of equations $\{f_i(x) = 0\}$ resulting from (4), (5). Let, as usual, ℓ_n be the number of vertices in F_n. Then the measure $\mathbf{P}^{\sigma,\eta}_{F_n}$ is concentrated on points of the form $p^{\widehat{F}}$, where $\widehat{F} \in \mathcal{F}^\sigma_{\ell_n}$. As we already remarked in the proof of Theorem 3.3, for every fixed relation f_i from the list we have the bound $|f_i(p^{\widehat{F}})| \leq O(1/\ell_n)$, *uniformly over all choices of* $\widehat{F} \in \mathcal{F}^\sigma_{\ell_n}$. This implies $\int_{x\in[0,1]^{\mathcal{F}^\sigma}} f_i(x)^2\mathbf{P}^{\sigma,\eta}_{F_n}(dx) \leq O(1/\ell_n)^2$ and thus, due to weak convergence, $\int_{x\in[0,1]^{\mathcal{F}^\sigma}} f_i(x)^2\mathbf{P}^{\sigma,\eta}(dx) = 0$. By a standard argument, we consecutively get from here $\mathbf{P}^{\sigma,\eta}(\{\,x \mid |f_i(x)| \leq \epsilon\,\}) = 1$ for every fixed $\epsilon > 0$, $\mathbf{P}^{\sigma,\eta}(\{\,x \mid f_i(x) = 0\,\}) = 1$ and $\mathbf{P}^{\sigma,\eta}(\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})) = 1$.

Uniqueness.
Assume that $\mathbf{Q}^{\sigma,\eta}$ is another probability measure for which (8) holds. Then

$$\int_{\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})} \phi(f)\mathbf{P}^{\sigma,\eta}(d\phi) = \int_{\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})} \phi(f)\mathbf{Q}^{\sigma,\eta}(d\phi)\ (f \in \mathcal{A}^\sigma). \tag{20}$$

By Proposition 3.7, (20) implies

$$\int_{\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})} f(\phi)\mathbf{P}^{\sigma,\eta}(d\phi) = \int_{\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})} f(\phi)\mathbf{Q}^{\sigma,\eta}(d\phi)$$

for *any* $f \in C(\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R}))$ (as it can be approximated by functions from $\mathcal{A}^\sigma$ within any fixed $\epsilon > 0$ in the ℓ_∞-norm). Since $\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})$ is metrizable, $\mathbf{P}^{\sigma,\eta} = \mathbf{Q}^{\sigma,\eta}$ follows by Proposition 3.8.

The proof of Theorem 3.5 is complete.

Remark 4. Lior Silberman (personal communication) observed that Theorem 3.5 possesses another proof using only basic facts from the functional analysis, and without any references to Prohorov's theorem. His argument goes along the following lines. First, we give an ad hoc proof of Theorem 3.1 a). Like the one based on extensions (and presented below), this proof will still use the random variable $\boldsymbol{\theta}$ from Definition 9. But now its only property needed will be the equality of the right-hand sides in (18), (19), which is a purely combinatorial fact. Next, Theorem 3.1 a) implies that the right-hand side of (17) defines a linear positive bounded functional μ on the subset $\mathcal{A}^\sigma \subseteq C(\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R}))$. Endow

[3] In fact, we will see in Theorem 3.12 that already the original sequence $\{F_n\}$ will do, but we do not need this fact for now.

$C(\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R}))$ with the ℓ_∞ metric; then μ is *uniformly* continuous w.r.t. this metric, and $\mathcal{A}^\sigma$ is dense in $C(\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R}))$ by Proposition 3.7. Therefore, μ has a unique (uniformly) continuous extension to the whole space $C(\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R}))$ that is easily seen to be again a linear positive bounded functional. Now Theorem 3.5 becomes a particular instance of Riesz Representation Theorem.

We, however, slightly prefer our proof as it includes a little bit more explicit way of generating the measures $\mathbf{P}^{\sigma,\eta}$ quite naturally extending the crucial Theorem 3.3. Let us formulate the corresponding result.

THEOREM 3.12. *Let σ_0 be a non-degenerate type, (σ,η) its extension, and let $\phi_0\in\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})$ be such that $\phi_0(\sigma,\eta)>0$. Fix an arbitrary convergent sequence $\{F_n\}$ of σ_0-flags such that $\lim_{n\to\infty}p^{F_n}=\phi_0$. Then $\{\mathbf{P}^{\sigma,\eta}_{F_n}\}$ weakly converges to the extension $\mathbf{P}^{\sigma,\eta}$ of ϕ_0.*

PROOF. We have to show that for any $f\in C([0,1]^{\mathcal{F}^\sigma})$,

$$\lim_{n\to\infty}\int_{[0,1]^{\mathcal{F}^\sigma}} f(x)\mathbf{P}^{\sigma,\eta}_{F_n}(dx)=\int_{\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})} f(\phi)\mathbf{P}^{\sigma,\eta}(d\phi).$$

For $f\in\mathcal{A}^\sigma$ this follows from Definition 8 and Lemma 3.11. The generalization to arbitrary $f\in C([0,1]^{\mathcal{F}^\sigma})$ is immediate by Proposition 3.7. ⊣

From now on we again resort to combinatorial parlance, and instead of the measures $\mathbf{P}^{\sigma,\eta}$ will be talking about "random homomorphisms" $\boldsymbol{\phi^{\sigma,\eta}}$ chosen according to these measures. Events and functions in which they will appear as arguments will always be obviously Borel.

DEFINITION 10. Let σ_0 be a non-degenerate type of size k_0 and $\phi\in\mathrm{Hom}^+(\mathcal{A}^{\sigma_0},\mathbb{R})$. An *ensemble of random homomorphisms rooted at ϕ* is a collection of random homomorphisms $\{\boldsymbol{\phi^{\sigma,\eta}}\}$, where (σ,η) runs over all extensions of σ_0 with $\phi((\sigma,\eta))>0$, and $\boldsymbol{\phi^{\sigma,\eta}}$ itself is chosen according to some probability measure on $\mathcal{B}^\sigma$ such that:

$$\mathbf{E}[\boldsymbol{\phi^{\sigma,\eta}}(f)]=\frac{\phi(\llbracket f\rrbracket_{\sigma,\eta})}{\phi(\langle\sigma,\eta\rangle)}\ (f\in\mathcal{A}^\sigma). \tag{21}$$

As always, in the most important case $\sigma_0=0$, η will be everywhere dropped from the notation.

Thus, Theorem 3.5 simply states that for every ϕ_0 there exists a unique ensemble of random homomorphisms rooted at ϕ_0.

As we promised above, Theorem 3.1 a) now becomes obvious. Indeed, let $f\in\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$, and $\phi\in\mathrm{Hom}^+(\mathcal{A}^{\sigma|_\eta},\mathbb{R})$. If $\phi((\sigma,\eta))=0$ then $\phi(F|_\eta)=0$ for every $F\in\mathcal{F}^\sigma$ and thus $\phi_0(\llbracket f\rrbracket_{\sigma,\eta})=0$. Otherwise, consider the extension $\boldsymbol{\phi^{\sigma,\eta}}$ of ϕ. Then $\mathbf{E}[\boldsymbol{\phi^{\sigma,\eta}}(f)]\geq 0$ (since $f\in\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$), therefore $\phi_0(\llbracket f\rrbracket_{\sigma,\eta})\geq 0$ by (21).

The ensembles become particularly useful, however, in the situations where we encounter "discontinuous" arguments, often involving random homomorphisms $\boldsymbol{\phi^{\sigma,\eta}}$ for different extensions (σ,η). At the first sight this looks a little bit paradoxical since Definition 10 has exactly the opposite spirit on both counts (every member of the ensemble is connected only to the root homomorphism ϕ, and only via expectations). In the next subsection we will give a few examples (suggested by obvious analogies with the discrete case) illustrating why this is not really an obstacle. But before that let us indicate one simple but very useful fact.

THEOREM 3.13. *Let σ_0 be a non-degenerate type, $\{\phi_n\}$ a convergent sequence in* $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, $\phi \stackrel{\mathrm{def}}{=} \lim_{n\to\infty} \phi_n$ *and* $\{\boldsymbol{\phi_n^{\sigma,\eta}}\}$, $\{\boldsymbol{\phi^{\sigma,\eta}}\}$ *be the ensembles rooted at* $\{\phi_n\}$, ϕ, *respectively. Then for every extension* (σ,η) *with* $\phi((\sigma,\eta)) > 0$, $\{\boldsymbol{\phi_n^{\sigma,\eta}}\}$ *weakly converges to* $\{\boldsymbol{\phi^{\sigma,\eta}}\}$.

PROOF. (cf. the proof of Theorem 3.12) We have to prove that

$$\lim_{n\to\infty} \mathbf{E}\left[f(\boldsymbol{\phi_n^{\sigma,\eta}})\right] = \mathbf{E}[f(\boldsymbol{\phi^{\sigma,\eta}})]$$

for any $f \in C(\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R}))$. For $f \in \mathcal{A}^\sigma$ this immediately follows from (21) and $\lim_{n\to\infty} \phi_n = \phi$. Proposition 3.7 once again reduces the case of general $f \in C(\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R}))$ to this one. ⊣

3.3. Bootstrapping. In this subsection we present several heterogeneous facts of "bootstrapping" nature elucidating some of the purposes for which we introduced ensembles of random homomorphisms, as well as how to work with them.

Let us begin with a hassle-free (and very natural) proof of the following most basic result.

THEOREM 3.14 (Cauchy-Schwarz inequality).

$$[\![f^2]\!]_{\sigma,\eta} \cdot [\![g^2]\!]_{\sigma,\eta} \geq_{\sigma|_\eta} [\![fg]\!]^2_{\sigma,\eta}.$$

In particular $(g = 1_\sigma)$,

$$[\![f^2]\!]_{\sigma,\eta} \cdot \langle\sigma,\eta\rangle \geq_{\sigma|_\eta} [\![f]\!]^2_{\sigma,\eta}, \tag{22}$$

which in turn implies

$$[\![f^2]\!]_{\sigma,\eta} \geq_{\sigma|_\eta} 0. \tag{23}$$

PROOF. By definition, we need to prove $\phi([\![f^2]\!]_{\sigma,\eta}) \cdot \phi([\![g^2]\!]_{\sigma,\eta}) \geq \phi([\![fg]\!]_{\sigma,\eta})^2$ for every $\phi \in \mathrm{Hom}^+(\mathcal{A}^{\sigma|_\eta}, \mathbb{R})$. If $\phi((\sigma,\eta)) = 0$, then both sides of this inequality evaluate to 0. Otherwise, let $\boldsymbol{\phi^{\sigma,\eta}}$ be the extension of ϕ. Then the inequality to be proven becomes

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

and this is just an ordinary instance of Cauchy-Schwarz.

In order to get (23) from (22), we only have to recall that $-C \cdot \langle\sigma,\eta\rangle \leq_{\sigma|_\eta} [\![f^2]\!]_{\sigma,\eta} \leq_{\sigma|_\eta} C \cdot \langle\sigma,\eta\rangle$ for some absolute constant $C > 0$, which implies that in the inequality $[\![f^2]\!]_{\sigma,\eta} \cdot \langle\sigma,\eta\rangle \geq_{\sigma|_\eta} 0$ the term $\langle\sigma,\eta\rangle$ can be removed. ⊣

EXAMPLE 11 (Goodman's bound [Goo]). In the theory T_{Graph}, $K_3 + \rho = \frac{1}{3}\bar{P}_3 + 2[\![e^2]\!]_1 \geq 2\rho^2$, therefore $K_3 \geq \rho(2\rho - 1)$.

Next, let us see how to use random homomorphisms for defining things like the minimal (or maximal) degree of a graph.

Given an arbitrary theory T, a non-degenerate type σ, $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ such that $\phi(\sigma) > 0$ and $F \in \mathcal{F}^\sigma$ we let

$$\delta_F(\phi) \stackrel{\mathrm{def}}{=} \max\{a \mid \mathbf{P}[\boldsymbol{\phi^\sigma}(F) < a] = 0\} = \inf\{a \mid \mathbf{P}[\boldsymbol{\phi^\sigma}(F) < a] > 0\}. \tag{24}$$

Theorem 3.15. *δ_F is an upper semi-continuous function on the open set* $\left\{\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R}) \mid \phi(\sigma) > 0\right\}$.

Proof. Let $\{\phi_n\}$ be a convergent sequence in $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$, and $\phi \stackrel{\text{def}}{=} \lim_{n\to\infty} \phi_n$. Let $\{\boldsymbol{\phi}_{\boldsymbol{n}}^{\boldsymbol{\sigma}}\}, \{\boldsymbol{\phi}^{\boldsymbol{\sigma}}\}$ be the ensembles rooted at $\{\phi_n\}, \phi$, respectively. By Theorem 3.13, $\{\boldsymbol{\phi}_{\boldsymbol{n}}^{\boldsymbol{\sigma}}\}$ weakly converges to $\boldsymbol{\phi}^{\boldsymbol{\sigma}}$. There could be at most countably many values a such that $\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma}}(F) = a] > 0$. For all other a, Proposition 3.9 implies

$$\lim_{n\to\infty} \mathbf{P}\left[\boldsymbol{\phi}_{\boldsymbol{n}}^{\boldsymbol{\sigma}}(F) < a\right] = \mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma}}(F) < a],$$

and, in particular, if $\mathbf{P}\left[\boldsymbol{\phi}_{\boldsymbol{n}}^{\boldsymbol{\sigma}}(F) < a\right] = 0$ for infinitely many n, then

$$\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma}}(F) < a] = 0.$$

$\delta_F(\phi) \geq \limsup \delta_F(\phi_n)$ follows. ⊣

We would like to stress that the meaning of the limit quantity $\delta_F(\phi)$ is slightly different from its finite analogue. Say, in the theory T_{Graph} the inequality $\delta_e(\phi) < a$ should be thought of not as "there exists a vertex of degree $< an$" but rather as "the density of such vertices is non-negligible". As a consequence, we can prove the analogue of Corollary 3.4 only in one direction (but, fortunately, it is that one which is important). With the following restriction, however, this difference becomes irrelevant and we can show the other direction as well.

Definition 11. A theory T is *vertex uniform* if it has only one model of size one (and, therefore, also only one type of size one that will be denoted by 1). Equivalently, T is vertex uniform if for every predicate symbol P either $T \vdash \forall x P(x, \ldots, x)$ or $T \vdash \forall x(\neg P(x, \ldots, x))$.

All theories we have mentioned so far are vertex uniform. If in Theorem 3.15 T is vertex uniform and $\sigma = 1$ then $\phi(\sigma) > 0$ holds automatically and δ_F is upper semi-continuous everywhere on $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$. In particular, it has a global maximum.

Theorem 3.16. *Let T be a vertex uniform theory and $F \in \mathcal{F}^1$. Then*

$$\max_{\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})} \delta_F(\phi) = \limsup_{\ell\to\infty} \max_{M \in \mathcal{F}_\ell^0} \delta_F(M),$$

where the minimal density $\delta_F(M)$ of F in M *is naturally defined as*

$$\delta_F(M) \stackrel{\text{def}}{=} \min_{v \in V(M)} p(F, (M, v)).$$

Proof. We begin with proving

$$\max_{\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})} \delta_F(\phi) \geq \limsup_{\ell\to\infty} \max_{M \in \mathcal{F}_\ell^0} \delta_F(M)$$

(this part holds for arbitrary types σ, and we do not need vertex uniformity). Let $a < \limsup_{\ell\to\infty} \max_{M \in \mathcal{F}_\ell^0} \delta_F(M)$; it is sufficient to show that $\max_{\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})} \delta_F(\phi) > a$. Fix an increasing sequence of models such that $\delta_F(M) > a$ for all its members, and find in it a convergent subsequence $\{M_n\}$. Then $\lim_{n\to\infty} p^{M_n} = \phi$ for some $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$, and by Theorem 3.12 the sequence of probability measures $\mathbf{P}_{M_n}^1$ in Definition 9 weakly converges to the extension $\mathbf{P}^1$ of

ϕ. Since $\mathbf{P}^1_{M_n}\left[\left\{x \in [0,1]^{\mathcal{F}^1} \mid x(F) < a\right\}\right] = 0$, from Proposition 3.9 we get $\mathbf{P}^1\left[\left\{x \in [0,1]^{\mathcal{F}^1} \mid x(F) < a\right\}\right] = 0$ which is exactly $\delta_F(\phi) \geq a$.

Next we prove

$$\max_{\phi \in \mathrm{Hom}^+(\mathcal{A}^0,\mathbb{R})} \delta_F(\phi) \leq \limsup_{\ell \to \infty} \max_{M \in \mathcal{F}^0_\ell} \delta_F(M).$$

Let $\phi \in \mathrm{Hom}^+(\mathcal{A}^0,\mathbb{R})$, $a_0 < a_1 < \delta_F(\phi)$, and $\{M_n\}$ be an increasing sequence of models such that $\{p^{M_n}\}$ converges to ϕ and $\mathbf{P}^1_{M_n}$ weakly converges to $\mathbf{P}^1$, where, again, $\mathbf{P}^1$ is the extension of ϕ. Then

$$\lim_{n\to\infty} \mathbf{P}[p(F,(M_n,\boldsymbol{v_n})) \leq a_1] = 0,$$

where $\boldsymbol{v_n}$ is picked uniformly at random from $V(M_n)$. In order to finish the argument, however, we need to show more, namely that there does not exist a *single* vertex $v \in V(M_n)$ for which $p(F,(M_n,v))$ is small. Whereas we do not know how to achieve this for larger types, when $\sigma = 1$ and T is vertex uniform, we can apply the following trick (that will also be used in Section 4.3). Namely, let V_n be the set of all "bad" vertices: $V_n \stackrel{\mathrm{def}}{=} \{v \in V(M_n) \mid p(F,(M_n,v)) \leq a_1\}$, and let $\widetilde{M_n} \stackrel{\mathrm{def}}{=} M_n - V_n$. Since the density of V_n tends to 0, $|p(F,(M_n,v)) - p(F,(\widetilde{M_n},v))|$ also tends to 0 as $n \longrightarrow \infty$, uniformly over all choices of $v \in V(M_n) \setminus V_n$. Therefore, for sufficiently large n no vertex $v \in V(M_n)$ with $p(F,(M_n,v)) > a_1$ may satisfy $p(F,(\widetilde{M_n},v)) \leq a_0$, and thus $\delta_F(\widetilde{M_n}) \geq a_0$. Since $|V(\widetilde{M_n})| \longrightarrow \infty$, $\limsup_{\ell\to\infty} \max_{M\in\mathcal{F}^0_\ell} \delta_F(M) \geq a_0$ follows. Since $a_0 < a_1 < \delta_F(\phi)$ were chosen arbitrarily, we are done. $\dashv$

In the examples given so far we have used ensembles of random homomorphisms for defining/arguing about objects that are "external" w.r.t their nature. The rest of this subsection is devoted to internal properties of ensembles.

Let *the kth level of the ensemble* $\{\boldsymbol{\phi}^{\boldsymbol{\sigma},\boldsymbol{\eta}}\}$ consist of all those $\boldsymbol{\phi}^{\boldsymbol{\sigma},\boldsymbol{\eta}}$ for which $|\sigma| = k$. First we note that every level completely determines all extensions belonging to lower levels.

Theorem 3.17. *Let σ_0 be a non-degenerate type, $\phi \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_0},\mathbb{R})$, and $\{\boldsymbol{\phi}^{\boldsymbol{\sigma},\boldsymbol{\eta}}\}$ be the ensemble rooted at ϕ. Let (σ_1,η_1) be an extension of σ_0 such that $\phi((\sigma_1,\eta_1)) > 0$, $k_1 \stackrel{\mathrm{def}}{=} |\sigma_1|$, $k_2 \geq k_1$ and $\eta : [k_1] \longrightarrow [k_2]$ be an arbitrary injective mapping. Denote by $Ext(\sigma_1,\eta)$ the set of all types σ_2 with $|\sigma_2| = k_2$ such that $\sigma_2|_\eta \approx \sigma_1$. Introduce non-negative weights on this set by*

$$w(\sigma_2) \stackrel{\mathrm{def}}{=} \phi(\langle\sigma_2,\eta\eta_1\rangle),$$

and choose a random type $\boldsymbol{\sigma_2} \in Ext(\sigma_1,\eta)$ according to this system of weights:

$$\mathbf{P}[\boldsymbol{\sigma_2} = \sigma_2] \stackrel{\mathrm{def}}{=} \frac{w(\sigma_2)}{\sum_{\sigma\in Ext(\sigma_1,\eta)} w(\sigma)} = \frac{\phi(\langle\sigma_2,\eta\eta_1\rangle)}{\phi(\langle\sigma_1,\eta_1\rangle)}.$$

Then the random homomorphism $\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta\eta_1}}\pi^{\boldsymbol{\sigma_2},\eta} \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_1},\mathbb{R})$ is equivalent (that is, corresponds to the same probability measure) to $\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}$.

PROOF. Due to uniqueness of ensembles, it suffices to show that $\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta}\boldsymbol{\eta_1}}\pi^{\boldsymbol{\sigma_2},\eta}$ satisfies (21), that is

$$\mathbf{E}[\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta}\boldsymbol{\eta_1}}\pi^{\boldsymbol{\sigma_2},\eta}(f)] = \frac{\phi([\![f]\!]_{\sigma_1,\eta_1})}{\phi(\langle\sigma_1,\eta_1\rangle)}\ (f\in\mathcal{A}^{\sigma_1}).$$

Applying the formula of total expectation to the left-hand side, we expand it as

$$\mathbf{E}[\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta}\boldsymbol{\eta_1}}\pi^{\boldsymbol{\sigma_2},\eta}(f)] = \frac{1}{\phi(\langle\sigma_1,\eta_1\rangle)}\cdot\sum_{\sigma_2\in Ext(\sigma_1,\eta)}\phi([\![\pi^{\sigma_2,\eta}(f)]\!]_{\sigma_2,\eta\eta_1})$$

$$= \frac{1}{\phi(\langle\sigma_1,\eta_1\rangle)}\cdot\sum_{\sigma_2\in Ext(\sigma_1,\eta)}\phi([\![\langle\sigma_2,\eta\rangle f]\!]_{\sigma_1,\eta_1}).$$

And now we only have to note that

$$\sum_{\sigma_2\in Ext(\sigma_1,\eta)}\langle\sigma_2,\eta\rangle = 1_{\sigma_1}.$$

$\dashv$

Finally, we show that the "inference rules" given by Theorem 3.1 are in fact "admissible" (and for (13) we will actually prove a much stronger statement). We will further elaborate on this topic in the concluding Section 6.

THEOREM 3.18. *Let σ_0 be a non-degenerate type, $\phi\in\mathrm{Hom}^+(\mathcal{A}^{\sigma_0},\mathbb{R})$, and $\{\boldsymbol{\phi}^{\boldsymbol{\sigma},\boldsymbol{\eta}}\}$ be the ensemble rooted at ϕ. Let $(\sigma_1,\eta_1),(\sigma_2,\eta_2)$ be two extensions of σ_0 of sizes k_1,k_2, respectively, such that $\phi((\sigma_i,\eta_i))>0$ $(i=1,2)$, and $\eta:[k_1]\longrightarrow[k_2]$ be an injective mapping such that $\sigma_2|_\eta=\sigma_1$ and $\eta_2=\eta\eta_1$.*

a) *For any $f\in\mathcal{A}^{\sigma_2}$ we have the following implication:*

$$\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}(f)\geq 0]=1\Longrightarrow\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}([\![f]\!]_{\sigma_2,\eta})\geq 0]=1.$$

b) *For any Borel set $A\subseteq\mathrm{Hom}^+(\mathcal{A}^{\sigma_1},\mathbb{R})$,*

$$\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}\in A]=0\Longrightarrow\mathbf{P}[(\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}\pi^{\sigma_2,\eta})\in A]=0.$$

c) *For any $f\in\mathcal{A}^{\sigma_1}$ we have*

$$\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}(f)\geq 0]=1\Longrightarrow\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}(\pi^{\sigma_2,\eta}(f))\geq 0]=1.$$

PROOF. **a).** Assume the contrary. Then for some $\epsilon,\delta>0$ we have

$$\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}([\![f]\!]_{\sigma_2,\eta})\leq-\epsilon]\geq\delta.$$

Also, there exists $C>0$ such that $\phi([\![f]\!]_{\sigma_2,\eta})\leq C$ for every $\phi\in\mathrm{Hom}^+(\mathcal{A}^{\sigma_1},\mathbb{R})$. By Proposition 3.6, there exists $\tilde g\in C(\mathrm{Hom}^+(\mathcal{A}^{\sigma_1},\mathbb{R}))$ such that

$$\phi([\![f]\!]_{\sigma_2,\eta})\leq-\epsilon\Longrightarrow\tilde g(\phi)=2,$$
$$\phi([\![f]\!]_{\sigma_2,\eta})\geq 0\Longrightarrow\tilde g(\phi)=0,$$

and by Proposition 3.7 we can approximate $\tilde g$ by some $g\in\mathcal{A}^{\sigma_1}$ with the property

$$\phi([\![f]\!]_{\sigma_2,\eta})\leq-\epsilon\Longrightarrow\phi(g)\geq 1,$$
$$\phi([\![f]\!]_{\sigma_2,\eta})\geq 0\Longrightarrow|\phi(g)|\leq(\epsilon\delta/(2C))^{1/2}.$$

Then, denoting the normalizing coefficient $\phi_0(\langle\sigma_i,\eta_i\rangle)>0$ by α_i, we get

$$\begin{aligned}\phi([\![g^2[\![f]\!]_{\sigma_2,\eta}]\!]_{\sigma_1,\eta_1}) &= \alpha_1\cdot \mathbf{E}\big[\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}(g)^2\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}([\![f]\!]_{\sigma_2,\eta})\big]\\ &\leq \alpha_1\left(-(\epsilon\delta)\cdot\mathbf{E}\big[\,\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}(g)^2\,|\,\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}([\![f]\!]_{\sigma_2,\eta})\leq-\epsilon\,\big]\right.\\ &\quad\left.+C\cdot\mathbf{E}\big[\,\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}(g)^2\,|\,\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}([\![f]\!]_{\sigma_2,\eta})\geq 0\,\big]\right)\leq-\frac{\alpha_1\epsilon\delta}{2}<0.\end{aligned}$$

On the other hand,

$$g^2[\![f]\!]_{\sigma_2,\eta}=[\![\pi^{\sigma_2,\eta}(g)^2f]\!]_{\sigma_2,\eta}$$

by Theorem 2.8 a) and, therefore,

$$[\![g^2[\![f]\!]_{\sigma_2,\eta}]\!]_{\sigma_1,\eta_1}=[\![\pi^{\sigma_2,\eta}(g)^2f]\!]_{\sigma_2,\eta_2} \tag{25}$$

by Theorem 2.5 c). Since $\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}(f)\geq 0]=1$, this implies

$$\phi([\![g^2[\![f]\!]_{\sigma_2,\eta}]\!]_{\sigma_1,\eta_1})=\alpha_2\cdot\mathbf{E}\big[\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}(\pi^{\sigma_2,\eta}(g))^2\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}(f)\big]\geq 0.$$

This contradiction proves part a) of the theorem.

b). By Theorem 3.17, $\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}\pi^{\sigma_2,\eta}$ is equivalent to $\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}$ conditioned by the event $\boldsymbol{\sigma_2}=\sigma_2$ of non-zero probability $\frac{\phi(\langle\sigma_2,\eta_2\rangle)}{\phi(\langle\sigma_1,\eta_1\rangle)}$.

c) follows from the already proven part b) applied to $A:=\{\phi\mid\phi(f)<0\}$ (it is also possible to give a more direct proof similar to the proof of part a)). ⊣

REMARK 5. Applying Theorem 3.18 to both f and $-f$, we also have its symmetric versions:

$$\begin{aligned}&\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}(f)=0]=1\Longrightarrow\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}([\![f]\!]_{\sigma_2,\eta})=0]=1\\ &\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}(f)=0]=1\Longrightarrow\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta_2}}(\pi^{\sigma_2\eta}(f))=0]=1\end{aligned} \tag{26}$$

((26) can be also obtained directly from part b) of that theorem).

COROLLARY 3.19. *Let σ_0 be a non-degenerate type, $\phi\in\mathrm{Hom}^+(\mathcal{A}^{\sigma_0},\mathbb{R})$ and $\boldsymbol{\phi}^{\boldsymbol{\sigma},\boldsymbol{\eta}}$ be the ensemble rooted at ϕ. Then for every $f\in\mathcal{A}^{\sigma_0}$,*

$$\mathbf{P}[\boldsymbol{\phi}^{\boldsymbol{\sigma},\boldsymbol{\eta}}(\pi^{\sigma,\eta}(f))=\phi(f)]=1.$$

PROOF. By applying (26) to $f-\phi(f)$. ⊣

Hopefully, results given in this subsection illustrate that (after a little bit of technical work) we *do not lose in our framework any customary finite arguments.* Let us exploit now what we clearly *gain* by the transfer to the limit case.

§4. Extremal homomorphisms. In extremal combinatorics (and in many other places) inductive arguments are often represented in the contrapositive form by casting the spell “consider a minimal counterexample”. This is exactly how we represent them in our framework, with the difference that now there is no such thing as “minimal” (as there is no such concept as “the number of vertices”). It is replaced by “the worst (or *extremal*) counterexample”.

4.1. Set-up and existence of extremal homomorphisms. Our basic set-up is quite easy (cf. Corollary 3.4), and for additional clarity we formulate it only for the case $\sigma = 0$. Let $M_1, \ldots, M_h \in \mathcal{F}^0$ be fixed models, $C \subseteq \mathbb{R}^h$ be a closed subset and $f : C \longrightarrow \mathbb{R}$ be a continuous function. We can view f as a continuous function on the compact set $\{\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R}) \mid (\phi(M_1), \ldots, \phi(M_h)) \in C\}$ given by

$$f(\phi) \stackrel{\text{def}}{=} f(\phi(M_1), \ldots, \phi(M_h)),$$

and we want to prove that $f(\phi) \geq 0$ on this set. Due to compactness, f attains its minimal value at some ϕ_0, and we will call any such ϕ_0 an *extremal homomorphism*. We fix this ϕ_0 and try to use to our advantage its extremality property. If, using this fact, we succeed in showing that $\phi_0(f) \geq 0$, then we are done. This simple scheme will be made more concrete in the following subsections (in Section 4.2 ϕ_0 must be a global minimum, and in Section 4.3 we will be content with local ones). Here we only remark that the general machinery developed in Section 3.2 allows us to treat in the same way more general problems. For example, by Theorem 3.16 the Caccetta-Häggkvist conjecture is equivalent to the inequality

$$\forall \phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R}) \left(\delta_\alpha(\phi) \leq \frac{1}{g-1}\right)$$

in the theory of oriented graphs with girth $\geq g$. By Theorem 3.15 we know[4] that δ_α attains its global maximum somewhere on $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$, and we again can concentrate on this extremal homomorphism only.

4.2. Inductive arguments. In this section we consider only those inductive arguments that are applied to a *substantially smaller* sub-model, usually defined by a selection criterion. "Continuous" induction, in which at every single step we change our model only "a little bit", will correspond to the differential structure explored in the next Section 4.3.

In the basic set-up (when C is a closed subspace in $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ and f is a continuous function on C) little can be added to what we already said in Section 2.3.2. If ϕ_0 is extremal and $f(\phi_0) = a$, then $f(\phi) \geq a$ for any other $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$, and we can translate this statement to a statement about elements from $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ by using the homomorphism $\pi^{F_0} : \mathcal{A}^0 \longrightarrow \mathcal{A}^\sigma_{F_0}$.

For more sophisticated settings (like the one with the minimal density δ_F) the induction in this style also becomes straightforward as long as we know that homomorphisms $\pi^{(U,I)} : \mathcal{A}^{\sigma_1}[T_1] \longrightarrow \mathcal{A}^{\sigma_2}_u[T_2]$ interact "nicely" with the construction of ensembles of random homomorphisms. In the rest of this section we develop the necessary formalism.

We continue working in the set-up of Definition 4, and we will be using all its notation.

Definition 12 (set-up continued). Fix $\phi_2 \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_2}[T_2], \mathbb{R})$ such that $\phi_2(u) > 0$. Then ϕ_2 can be extended to a homomorphism from the quotient algebra $\mathcal{A}^{\sigma_2}_u[T_2]$ to $\mathbb{R}$; let $\phi_1 \stackrel{\text{def}}{=} \phi_2 \pi^{(U,I)} \in \mathrm{Hom}^+(\mathcal{A}^{\sigma_1}[T_1], \mathbb{R})$. Fix an extension (σ_1^*, η_1) of σ_1 such that $\phi_1((\sigma_1^*, \eta_1)) > 0$, and let $\boldsymbol{\phi}_1^{\boldsymbol{\sigma}_1^*, \boldsymbol{\eta}_1}$ be the corresponding extension

[4]Since $\sigma = 1$, the condition $\phi(\sigma) > 0$ holds automatically.

of ϕ_1. Our goal is to describe $\boldsymbol{\phi_1^{\sigma_1^*,\eta_1}}$ in terms of the ensemble $\{\boldsymbol{\phi_2^{\sigma_2,\eta_2}}\}$ rooted at ϕ_2.

Let k_1^* be the size of σ_1^* and $k_2^* \stackrel{\text{def}}{=} k_1^* + (k_2 - k_1)$. Fix an injective $\eta^* : [k_1^*] \longrightarrow [k_2^*]$ such that $\mathrm{im}(\eta^*) = \mathrm{im}(\eta) \dot\cup \{k_2+1, \ldots, k_2^*\}$ and $\eta_2\eta = \eta^*\eta_1$, where $\eta_2 : [k_2] \longrightarrow [k_2^*]$ is the identical mapping: $\eta_2(i) \stackrel{\text{def}}{=} i$ $(i \in [k_2])$.

We now define the translation I^* of non-logical symbols in the language $L_1(c_1, \ldots, c_{k_1^*})$ to the language $L_2(c_1, \ldots, c_{k_2^*})$ as follows. On predicate symbols I^* acts in the same way as I (note that since η_2 is identical, old constants $c_1, \ldots, c_{k_2}$ that may appear in $I(P)$ retain their meaning). On constants $c_1, \ldots, c_{k_1^*}$, I^* acts accordingly to $\eta^*\eta_1$: $I^*(c_i) \stackrel{\text{def}}{=} c_{\eta^*\eta_1(i)}$.

Now consider the set $Ext(\sigma_1^*, \eta_1)$ consisting of those types σ_2^* of size k_2^* for which $\sigma_2^*|_{\eta_2} \approx \sigma_2$, $(\sigma_2^*, \eta_2) \in \mathcal{F}^{\sigma_2,U}[T_2]$ and $\eta^* : \sigma_1^* \longrightarrow \sigma_2^*$ is a model embedding (the last two conditions together are equivalent to the fact that $(U, I^*) : T_1^{\sigma_1^*} \rightsquigarrow T_2^{\sigma_2^*}$ is an open interpretation). Note that the element $u^* \in \mathcal{A}^{\sigma_2^*}[T_2]$ corresponding to this interpretation is computed as $u^* = \pi^{\sigma_2^*,\eta_2}(u)$; in particular, by Corollary 3.19 we have $\boldsymbol{\phi_2^{\sigma_2^*,\eta_2}}(u^*) = \phi_2(u) > 0$ with probability 1.

The following is analogous to Theorem 3.17 (and in fact they could be combined into one statement by allowing k_2^* in Definition 12 to be any integer $\geq k_1^* + (k_2 - k_1)$).

THEOREM 4.1. $\boldsymbol{\phi_1^{\sigma_1^*,\eta_1}}$ *is equivalent to the random homomorphism constructed by the following process. Introduce first non-negative weights on the set* $Ext(\sigma_1^*, \eta_1)$ *by letting* $w(\sigma_2^*) \stackrel{\text{def}}{=} \phi_2(\langle\sigma_2^*, \eta_2\rangle)$. *Next, choose* $\boldsymbol{\sigma_2^*} \in Ext(\sigma_1^*, \eta_1)$ *at random according to this system of weights:*

$$\mathbf{P}\big[\boldsymbol{\sigma_2^*} = \sigma_2^*\big] \stackrel{\text{def}}{=} \frac{w(\sigma_2^*)}{\sum_{\sigma_2^* \in Ext(\sigma_1^*,\eta_1)} w(\sigma_2^*)} = \frac{\phi_2(\langle\sigma_2^*, \eta_2\rangle)}{\phi_1(\langle\sigma_1^*, \eta_1\rangle)}.$$

Finally, output $\boldsymbol{\phi_2^{\sigma_2^*,\eta_2}}\pi^{U,I^*}$.

PROOF. Let $\boldsymbol{\psi^{\sigma_1^*,\eta_1}}$ be the random homomorphism constructed by the process described in the statement. Due to uniqueness of ensembles, we only have to show that for every $F \in \mathcal{F}^{\sigma_1^*}[T_1]$ we have

$$\mathbf{E}\Big[\boldsymbol{\psi^{\sigma_1^*,\eta_1}}(F)\Big] = \frac{\phi_1([\![F]\!]_{\sigma_1^*,\eta_1})}{\phi_1(\langle\sigma_1^*, \eta_1\rangle)}.$$

And after expanding definitions this amounts to yet another standard calculation. ⊣

4.3. Differential methods. If the goal function f is smooth enough, then the self-suggesting way to utilize the extremality condition is by using differential (or variational) techniques. This is perhaps the most rewarding outcome of all the technical work we have to do.

Let σ be a non-degenerate type and $\ell \geq |\sigma|$. We introduce the auxiliary linear mapping $\mu_\ell^\sigma : \mathbb{R}\mathcal{F}_\ell^0 \longrightarrow \mathbb{R}\mathcal{F}_\ell^\sigma$ by its action on models with ℓ vertices as follows:

$$\mu_\ell^\sigma(M) \stackrel{\text{def}}{=} \sum \{ F \in \mathcal{F}_\ell^\sigma \mid F|_0 \approx M \}.$$

Note that although μ_ℓ^σ differs from π^σ only in that it uses $F|_\eta$ instead of $F\downarrow_\eta$, μ_ℓ^σ does not possess any of its nice properties (in fact, it does not even satisfy $\mu^\sigma(\mathcal{K}^0) \subseteq \mathcal{K}^\sigma$). This is why we must explicitly indicate the subscript ℓ here.

Our first operator corresponds to vertex deletion in the finite world. We assume for simplicity that T is vertex uniform (that is, has only one singleton model), and define a linear mapping $\partial_1 : \mathbb{R}\mathcal{F}^0 \longrightarrow \mathbb{R}\mathcal{F}^1$ by its action on models as

$$\partial_1 M \stackrel{\text{def}}{=} \ell(\pi^1(M) - \mu^1(M)) \ (M \in \mathcal{M}_\ell).$$

EXAMPLE 12. In the theory T_{Graph}, $\partial_1 K_\ell = \ell(\pi^1(K_\ell) - K_\ell^1)$.

LEMMA 4.2. **a)** *Let $M \in \mathcal{M}_\ell$, $L \geq \ell + 1$, $N \in \mathcal{M}_L$ and $v \in V(N)$. Then*

$$p(M, N - v) = p(M, N) + \frac{1}{L} p^{(N,v)}(\partial_1 M).$$

b) *$\partial_1(\mathcal{K}^0) \subseteq \mathcal{K}^1$ and, therefore, ∂_1 defines a linear mapping from $\mathcal{A}^0$ to $\mathcal{A}^1$.*
c) *$[\![\partial_1 f]\!]_1 = 0$ for every $f \in \mathcal{A}^0$.*

PROOF. **a).** Pick uniformly at random an ℓ-subset $\boldsymbol{V}$ of $V(N)$. By the formula of total probability,

$$p(M, N) = \frac{\ell}{L}\mathbf{P}[\, N|_{\boldsymbol{V}} \approx M \mid v \in \boldsymbol{V}\,] + \left(1 - \frac{\ell}{L}\right)\mathbf{P}[\, N|_{\boldsymbol{V}} \approx M \mid v \not\in \boldsymbol{V}\,].$$

Now we only have to observe that $\mathbf{P}[\, N_{\boldsymbol{V}} \approx M \mid v \in \boldsymbol{V}\,] = p^{(N,v)}(\mu_\ell^1(M))$, whereas $\mathbf{P}[\, N|_{\boldsymbol{V}} \approx M \mid v \not\in \boldsymbol{V}\,] = p^{(N,v)}(\pi^1(M)) = p(M, N - v)$.

b). If $f = 0$ is a relation of the form (4) and L is sufficiently large, then by the already proven part **a)**, $p^{(N,v)}(\partial_1 f) = L \cdot (p^{N-v}(f) - p^N(f)) = 0$. Since $(N, v) \in \mathcal{F}_L^1$ is arbitrary, this implies $\partial_1 f = 0$.

c). For a model $M \in \mathcal{M}_\ell$, $[\![\mu^1(M)]\!]_1 = [\![\pi^1(M)]\!]_1 = M$; thus, $[\![\partial_1 M]\!]_1 = 0$. This is extended to arbitrary $f \in \mathcal{A}^0$ by linearity. ⊣

Assume now that $\vec{M} = (M_1, \ldots, M_h) \in \mathcal{M}$ are fixed models, $a \in \mathbb{R}^h$ and $f \in C^1(U)$, where $U \subseteq \mathbb{R}^h$ is an open neighbourhood of a (and $C^1(U)$ is the class of continuously differentiable functions on U). We let

$$Grad_{\vec{M},a}(f) \stackrel{\text{def}}{=} \sum_{i=1}^h \left.\frac{\partial f}{\partial x_i}\right|_{x=a} \cdot M_i \ \in \mathcal{A}^0$$

(thus, $Grad_{\vec{M},a}(f)$ is the inner product of the ordinary gradient $\nabla f(a)$ with the vector $\langle M_1, \ldots, M_h\rangle$).

THEOREM 4.3. *Let $\vec{M} = (M_1, \ldots, M_h) \in \mathcal{M}$ be fixed models of the theory T, $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$, and $f \in C^1(U)$, where U is an open neighbourhood of the point $a \stackrel{\text{def}}{=} (\phi_0(M_1), \ldots, \phi_0(M_h)) \in \mathbb{R}^h$. Assume that for any $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ such that $(\phi(M_1), \ldots, \phi(M_h)) \in U$ we have*

$$f(\phi(M_1), \ldots, \phi(M_h)) \geq f(a).$$

Then for the extension $\boldsymbol{\phi_0^1}$ of ϕ_0,

$$\mathbf{P}\left[\boldsymbol{\phi_0^1}(\partial_1 Grad_{\vec{M},a}(f)) = 0\right] = 1 \tag{27}$$

holds.

PROOF. Lemma 4.2 c) implies that

$$\mathbf{E}\Big[\boldsymbol{\phi}^{\mathbf{1}}_{\mathbf{0}}(\partial_1 Grad_{\vec{M},a}(f))\Big] = \phi_0(\llbracket\partial_1 Grad_{\vec{M},a}(f)\rrbracket_1) = 0.$$

Therefore, it suffices to prove (27) only in one direction, and we elect to prove that

$$\mathbf{P}\Big[\boldsymbol{\phi}^{\mathbf{1}}_{\mathbf{0}}(\partial_1 Grad_{\vec{M},a}(f)) \geq 0\Big] = 1. \tag{28}$$

Fix an increasing sequence $\{N_n\}$ of models such that $\lim_{n\to\infty} p^{N_n} = \phi_0$. Then by Theorem 3.12 the corresponding sequence of probability measures $\mathbf{P}^1_{N_n}$ weakly converges to the extension $\mathbf{P}^1$ of ϕ_0 (recall that $\mathbf{P}^1_{N_n}$ corresponds to the random element $p^{(N_n,\boldsymbol{v_n})} \in [0,1]^{\mathcal{F}^1}$, where $\boldsymbol{v_n}$ is chosen uniformly at random from $V(N_n)$).

Assume now that (28) fails, that is $\mathbf{P}\Big[\boldsymbol{\phi}^{\mathbf{1}}_{\mathbf{0}}(\partial_1 Grad_{\vec{M},a}(f)) < 0\Big] > 0$. Then for some $\epsilon > 0$ we have

$$\mathbf{P}\Big[\boldsymbol{\phi}^{\mathbf{1}}_{\mathbf{0}}(\partial_1 Grad_{\vec{M},a}(f)) < -\epsilon\Big] > 0,$$

and since the set $\Big\{\phi \in \mathrm{Hom}^+(\mathcal{A}^1, \mathbb{R}) \;\Big|\; \phi(\partial_1 Grad_{\vec{M},a}) < -\epsilon\Big\}$ is open, we can apply Theorem 3.9 and conclude that

$$\mathbf{P}\Big[p^{(N_n,\boldsymbol{v_n})}(\partial_1 Grad_{\vec{M},a}(f)) < -\epsilon\Big] \geq \epsilon'$$

for some absolute constants $\epsilon, \epsilon' > 0$, and for all sufficiently large n.

Fix now a sufficiently small positive constant $\delta < \epsilon'$ (to be specified later in the course of the proof). Let $\ell_n \stackrel{\text{def}}{=} |V(N_n)|$, $m_n \stackrel{\text{def}}{=} \lfloor \delta \ell_n \rfloor$, and choose arbitrarily an m_n-subset $V_n \subseteq V(N_n)$ such that

$$\forall v \in V_n \left(p^{(N_n,v)}(\partial_1 Grad_{\vec{M},a}(f)) < -\epsilon\right). \tag{29}$$

Let $\widetilde{N}_n \stackrel{\text{def}}{=} N_n - V_n$. We convert N_n to $\widetilde{N}_n$ by removing vertices in V_n one by one. That is, let $V_n = \{v_1, \ldots, v_{m_n}\}$ be an arbitrary enumeration and $N_{n,i} \stackrel{\text{def}}{=} N_n - \{v_1, \ldots, v_i\}$ $(0 \leq i \leq m_n)$. Let M be any fixed model and $\ell \stackrel{\text{def}}{=} |V(M)|$. We use Lemma 4.2 a) to calculate the difference $p^{\widetilde{N}_n}(M) - p^{N_n}(M)$ as follows:

$$\left.\begin{aligned} p^{\widetilde{N}_n}(M) - p^{N_n}(M) &= \sum_{i=1}^{m_n} p(M, N_{n,i}) - p(M, N_{n,i-1}) \\ &= \sum_{i=1}^{m_n} \frac{1}{\ell_n - i + 1} p^{(N_{n,i-1},v_i)}(\partial_1 M). \end{aligned}\right\} \tag{30}$$

Next, for every fixed i we want to approximate $\frac{1}{\ell_n - i+1} p^{(N_{n,i-1},v_i)}(\partial_1 M)$ by $\frac{1}{\ell_n} p^{(N_n,v_i)}(\partial_1 M)$. For a random ℓ-subset $\boldsymbol{V} \subseteq V(N_n)$ picked uniformly under the condition $v_i \in \boldsymbol{V}$ the probability that $\boldsymbol{V}$ will also contain at least one of the vertices $v_1, \ldots, v_{i-1}$ is $O(\delta)$ (remember that the model M is fixed, so the multiplicative constant assumed here may also depend on ℓ). Therefore,

$$\left|p^{(N_{n,i-1},v_i)}(\partial_1 M) - p^{(N_n,v_i)}(\partial_1 M)\right| \leq O(\delta). \tag{31}$$

Since also $\frac{1}{\ell_n - i+1} \leq O(1/\ell_n)$ and $\left|\frac{1}{\ell_n-i+1} - \frac{1}{\ell_n}\right| \leq O(\delta/\ell_n)$, we get

$$\left|\frac{1}{\ell_n - i + 1} p^{(N_{n,i-1},v_i)}(\partial_1 M) - \frac{1}{\ell_n} p^{(N_n,v_i)}(\partial_1 M)\right| \leq O(\delta/\ell_n),$$

which, along with (30), implies the bound

$$p^{\widetilde{N}_n}(M) - p^{N_n}(M) = \frac{1}{\ell_n}\sum_{i=1}^{m_n} p^{(N_n,v_i)}(\partial_1 M) \pm O(\delta^2). \tag{32}$$

Applying this to $M = M_1, \ldots, M_h$, taking the inner product with $\nabla f(a)$ and recalling (29), we finally get

$$p^{\widetilde{N}_n}(Grad_{\vec{M},a}(f)) - p^{N_n}(Grad_{\vec{M},a}(f)) = -\epsilon\delta \pm O(\delta^2) \leq -\frac{\epsilon\delta}{2} \tag{33}$$

provided δ is small enough.

By compactness, we can choose an increasing convergent subsequence in $\{\widetilde{N}_n\}$, and w.l.o.g. let us assume that $\{\widetilde{N}_n\}$ itself converges. Let $\phi \stackrel{\text{def}}{=} \lim_{n\to\infty} p^{\widetilde{N}_n}$. Taking limit in (32), we also see that $|\phi(M) - \phi_0(M)| \leq O(\delta)$ for every fixed M. In particular, if δ is small enough then $(\phi(M_1), \ldots, \phi(M_h)) \in U$. Taking limit in (33),

$$\phi(Grad_{\vec{M},a}(f)) - \phi_0(Grad_{\vec{M},a}(f)) \leq -\frac{\epsilon\delta}{2}.$$

But since f is a C^1-function, we have

$$\begin{aligned} f(\phi(M_1), \ldots, \phi(M_h)) &\leq f(a) + \sum_{i=1}^{h} \left.\frac{\partial f}{\partial x_i}\right|_{x=a} (\phi(M_i) - \phi_0(M_i)) + o(\delta) \\ &= f(a) + (\phi(Grad_{\vec{M},a}(f)) - \phi_0(Grad_{\vec{M},a}(f))) + o(\delta) < f(a) \end{aligned}$$

provided the constant δ is small enough. This contradiction proves the theorem. $\dashv$

Our second "variation principle" corresponds to edge-deletion, and certainly not all interesting theories admit this (or similar) operation. Although it should be straightforward to formulate general conditions on the theory T under which this principle works, we prefer to avoid here excessive generality and formulate it for undirected graphs only. Accordingly, we will be denoting models by more customary letters G, H rather than M, N; recall also that there are exactly two types of size 2: E (corresponding to an edge) and $\bar{E}$ (non-edge).

Denote by $Fill : \mathcal{A}^{\bar{E}} \longrightarrow \mathcal{A}^E$ the natural isomorphism defined by adding an edge between the two distinguished vertices, and for a graph $G \in \mathcal{M}_\ell$, let

$$\partial_E G \stackrel{\text{def}}{=} \frac{\ell(\ell-1)}{2}(Fill(\mu_\ell^{\bar{E}}(G)) - \mu_\ell^E(G)).$$

∂_E is extended to a linear mapping from $\mathbb{R}\mathcal{F}^0$ to $\mathbb{R}\mathcal{F}^E$ by linearity.

EXAMPLE 13. $\partial_E K_\ell = -\frac{\ell(\ell-1)}{2} K_\ell^E$.

LEMMA 4.4. **a)** *Let $H \in \mathcal{M}_\ell,\ L \geq \ell,\ G \in \mathcal{F}^0_L$ and $(v_1, v_2) \in E(G)$. Then*

$$p(H, G - (v_1, v_2)) = p(H, G) + \frac{2}{L(L-1)} p^{(G,v_1,v_2)}(\partial_E H).$$

b) *$\partial_E(\mathcal{K}^0) \subseteq \mathcal{K}^E$, and, therefore, ∂_E defines a linear mapping from $\mathcal{A}^0$ to $\mathcal{A}^E$.*

PROOF. **a).** Again, pick uniformly at random an ℓ-subset $\boldsymbol{V}$ of $V(G)$. Then

$$\begin{aligned} p(H, G) &= \frac{\ell(\ell-1)}{L(L-1)} \cdot \mathbf{P}[\, G|_{\boldsymbol{V}} \approx H \,|\, \{v_1, v_2\} \subseteq \boldsymbol{V}\,] \\ &+ \left(1 - \frac{\ell(\ell-1)}{L(L-1)}\right) \mathbf{P}[\, G|_{\boldsymbol{V}} \approx H \,|\, \{v_1, v_2\} \not\subseteq \boldsymbol{V}\,], \\ p(H, G - (v_1, v_2)) &= \frac{\ell(\ell-1)}{L(L-1)} \cdot \mathbf{P}[\, (G|_{\boldsymbol{V}} - (v_1, v_2)) \approx H \,|\, \{v_1, v_2\} \subseteq \boldsymbol{V}\,] \\ &+ \left(1 - \frac{\ell(\ell-1)}{L(L-1)}\right) \mathbf{P}[\, G|_{\boldsymbol{V}} \approx H \,|\, \{v_1, v_2\} \not\subseteq \boldsymbol{V}\,] \end{aligned}$$

(note that $\{v_1, v_2\} \not\subseteq \boldsymbol{V}$ implies that $G_{\boldsymbol{V}}$ and $G|_{\boldsymbol{V}} - (v_1, v_2)$ are the same),

$$\begin{aligned} p^{(G,v_1,v_2)}(\partial_E H) &= \frac{\ell(\ell-1)}{2} \cdot (\mathbf{P}[\, G|_{\boldsymbol{V}} - (v_1, v_2) \approx H \,|\, \{v_1, v_2\} \subseteq \boldsymbol{V}\,] \\ &-\mathbf{P}[\, G|_{\boldsymbol{V}} \approx H \,|\, \{v_1, v_2\} \subseteq \boldsymbol{V}\,]). \end{aligned}$$

b) is proved exactly as part b) in Lemma 4.2. $\dashv$

THEOREM 4.5. *Let $G_1, \ldots, G_h$ be fixed undirected graphs, $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ be such that $\phi_0(\rho) > 0$, and $f \in C^1(U)$, where U is an open neighbourhood of the point $a \stackrel{\mathrm{def}}{=} (\phi_0(G_1), \ldots, \phi_0(G_h)) \in \mathbb{R}^h$. Assume that for any $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ such that $(\phi(G_1), \ldots, \phi(G_h)) \in U$ we have*

$$f(\phi(G_1), \ldots, \phi(G_h)) \geq f(a).$$

Then for the extension $\boldsymbol{\phi_0^E}$ of ϕ_0,

$$\mathbf{P}\left[\boldsymbol{\phi_0^E}(\partial_E Grad_{\vec{G},a}(f)) \geq 0\right] = 1 \tag{34}$$

holds.

PROOF. We again fix an increasing sequence of graphs H_n, $|V(H_n)| = \ell_n$ converging to ϕ_0. By the same argument as in the proof of Theorem 4.3, we only have to show that

$$\mathbf{P}\left[p^{(H_n, \boldsymbol{e_n})}(\partial_E Grad_{\vec{G},a}(f)) \leq -\epsilon\right] \geq \epsilon'$$

leads to contradiction. Here ϵ, ϵ' are absolute constants, n is sufficiently large, and $\boldsymbol{e_n}$ is a random edge of H_n (note that since $\partial_E Grad_{\vec{G},a}(f) \in \mathcal{A}^E$ is invariant under the automorphism of $\mathcal{A}^E$ permuting the two distinguished vertices, the quantity $p^{(H_n,e_n)}(\partial_E Grad_{\vec{G},a}(f))$ does not depend on the orientation of the edge e_n in the E-flag (H_n, e_n)). Let

$$Bad \stackrel{\mathrm{def}}{=} \left\{ e_n \in E(H_n) \,\middle|\, p^{(H_n, \boldsymbol{e_n})}(\partial_E Grad_{\vec{G},a}(f)) \leq -\epsilon \right\}.$$

Now comes a slightly tricky point. As in the proof of Theorem 4.3, we want to choose in Bad a subset of appropriate density and delete these edges from G_n one by one. But now we have to be more careful and make sure that we do not choose too many edges adjacent to any particular vertex (since otherwise the analogue of (31) may fail). For this we need the following simple trick.

Pick, uniformly at random, a subset $\boldsymbol{V}$ of *vertices* of size $\lfloor \delta^{1/2}\ell_n + 1\rfloor$, where, again, $\delta > 0$ is a sufficiently small constant to be specified later. Let $Bad|_V$ be the set of all edges in Bad with both endpoints in V; then $|Bad|_{\boldsymbol{V}}| \leq \delta\ell_n^2$ (with probability 1). On the other hand, by an averaging argument, $\mathbf{E}[|Bad|_{\boldsymbol{V}}|] \geq \epsilon' \cdot \frac{\delta\ell_n^2}{2}$. Fix an arbitrary $\lfloor \delta^{1/2}\ell_n + 1\rfloor$-subset $V_n \subseteq V(H_n)$ for which

$$\epsilon' \cdot \frac{\delta\ell_n^2}{2} \leq |Bad|_{V_n}| \leq \delta\ell_n^2,$$

and let $m_n \stackrel{\mathrm{def}}{=} |Bad|_{V_n}|$.

Now we begin eliminating the edges from $Bad|_{V_n} = \{e_1, \ldots, e_{m_n}\}$ one by one in the same way as in the proof of Theorem 4.3; let

$$H_{n,0} = H_n, H_{n,1}, \ldots, H_{n,m_n} = \widetilde{H}_n$$

be the corresponding sequence of graphs. Similarly to (30), from Theorem 4.4 a) we get

$$p^{\widetilde{H}_n}(G) - p^{H_n}(G) = \frac{2}{\ell_n(\ell_n - 1)} \cdot \sum_{i=1}^{m_n} p^{(G_{n,i-1},e_i)}(\partial_E G)$$

($G \in \mathcal{M}_\ell$ any fixed graph). A random ℓ-subset $\boldsymbol{V} \subseteq V(H_n)$ picked uniformly at random under the condition that it contains both endpoints of e_i may contain any of the edges $e_1, \ldots, e_{i-1}$ only if it contains at least one more vertex from V_n. This observation gives us the analogue of (31), except that $O(\delta)$ in the right-hand side gets replaced by $O(\delta^{1/2})$. Then the upper bound $m_n \leq \delta\ell_n^2$ implies

$$p^{\widetilde{H}_n}(G) - p^{H_n}(G) = \frac{2}{\ell_n(\ell_n - 1)} \cdot \sum_{i=1}^{m_n} p^{(G_n,e_i)}(\partial_E G) \pm O(\delta^{3/2}).$$

The rest of the proof is the same as in Theorem 4.3. ⊣

Finally, we give a "light" version of these variational principles that does not refer to random homomorphisms at all.

Corollary 4.6. *Let T be a vertex uniform theory, $\vec{M} = (M_1, \ldots, M_h) \in \mathcal{M}$ be its fixed models, $\phi_0 \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$, and $f \in C^1(U)$, where U is an open neighbourhood of the point $a \stackrel{\mathrm{def}}{=} (\phi_0(M_1), \ldots, \phi_0(M_h)) \in \mathbb{R}^h$. Assume that for any other $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ such that $(\phi(M_1), \ldots, \phi(M_h)) \in U$ we have $f(\phi(M_1), \ldots, \phi(M_h)) \geq f(a)$.*

a) *For every $g \in \mathcal{A}^1$,*

$$\phi_0([\![(\partial_1 Grad_{\vec{M},a})g]\!]_1) = 0.$$

b) *Assume further that $T = T_{Graph}$. Then for any $g \in \mathcal{C}_{\mathrm{sem}}(\mathcal{A}^E)$ (and, in particular, for any E-flag) we have*

$$\phi_0([\![(\partial_E Grad_{\vec{M},a})g]\!]_E) \geq 0.$$

PROOF. Immediate from Theorems 4.3, 4.5 and Definition 8. ⊣

§5. Triangle density. *In this section we are exclusively working in the theory T_{Graph}.*

For a fixed $\rho \in [0,1]$, what is (asymptotically) the minimal possible density $g_3(\rho)$ of triangles in a graph with edge density ρ? More precisely, we want to compute the function $g_3(x)$ given by

$$g_3(x) \stackrel{\mathrm{def}}{=} \liminf_{n\to\infty} \min\left\{ p(K_3, G_n) \mid G_n \in \mathcal{M}_n \wedge p(\rho, G_n) \geq x \right\}.$$

$g_3(x)$ is clearly monotone in x, and (by a simple edge-adding argument), it is also continuous. Then by Corollary 3.4 it can be alternatively represented as

$$g_3(x) = \min\left\{ \phi(K_3) \;\middle|\; \phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R}) \wedge \phi(\rho) \geq x \right\}.$$

Next, $g_3(x) = 0$ if $x \leq 1/2$ and the asymptotic version of Mantel's theorem, combined with a general result from [ErSi], implies that $g_3(x) > 0$ as long as $x > 1/2$.

It is easy to see that for $x \in \left[1 - \frac{1}{t}, 1 - \frac{1}{t+1}\right]$ ($t \geq 2$ an integer) we have

$$g_3(x) \leq \frac{(t-1)\left(t - 2\sqrt{t(t - x(t+1))}\right)\left(t + \sqrt{t(t-x(t+1))}\right)^2}{t^2(t+1)^2}. \tag{35}$$

The homomorphism witnessing this inequality is the limit of the convergent sequence of $(t+1)$-partite graphs in which t parts are (roughly) equal and larger than the remaining part (and the edge density is roughly x). The question of whether this bound is tight turned out to be notoriously difficult. Goodman [Goo] proved that $g_3(\rho) \geq \rho(2\rho - 1)$ (see Example 11). This in particular shows that (35) is tight at the critical values $x = 1 - \frac{1}{t}$, $t = 2, 3, \ldots$ Further partial results were given in [Bol, LoSi, Fish]. In particular, Fisher [Fish] proved that the bound (35) is tight for $t = 2$:

$$g_3(x) = \frac{(1 - \sqrt{4-6x})(2 + \sqrt{4-6x})^2}{18},\ 1/2 \leq x \leq 2/3. \tag{36}$$

In this section we give another proof of (36). It was found independently of [Fish], and, as far as we can see, it is totally different from the one given there.

Denote the right-hand side of (36) by $g(x)$. In the setting of Section 4, let $h := 2$, $G_1 := \rho$, $G_2 := K_3$, $C := \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid 1/2 \leq x_1 \leq 2/3 \right\}$ and $f(x_1, x_2) := x_2 - g(x_1)$ so that $f(\phi) = \phi(K_3) - g(\phi(\rho))$. f attains its minimum on the compact space $\left\{ \phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R}) \mid 1/2 \leq \phi(\rho) \leq 2/3 \right\}$, and we fix any such minimum ϕ_0. It suffices to prove that $\phi_0(K_3) \geq g(\phi_0(\rho))$. If $\phi_0(\rho) = 1/2$ or $\phi_0(\rho) = 2/3$, we are done by Goodman's bound. If $1/2 < \phi_0(\rho) < 2/3$, we apply Corollary 4.6 (with $U := \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid 1/2 < x_1 < 2/3 \right\}$), where we choose $g := e$ for part a) and $g := \bar{P}_3^E$ for part b). Denoting $\phi_0(\rho)$ by a and $\phi_0(K_3)$ by b, we consecutively compute:

$$Grad_{\rho,K_3,a,b}(f) = K_3 - g'(a)\rho$$
$$\partial_1 Grad_{\rho,K_3,a,b}(f) = (3\pi^1(K_3) - 2g'(a)\pi^1(\rho)) - (3K_3^1 - 2g'(a)e)$$
$$\partial_E Grad_{\rho,K_3,a,b}(f) = g'(a)\cdot 1_E - 3K_3^E$$
$$\phi_0(3[\![eK_3^1]\!]_1 - 2g'(a)[\![e^2]\!]_1) = a(3b - 2ag'(a)) \tag{37}$$
$$\phi_0([\![\bar{P}_3^E K_3^E]\!]_E) \leq \frac{1}{9}g'(a)\phi_0(\bar{P}_3). \tag{38}$$

We relate the constraints (37) and (38) with the help of the following easy lemma (that does *not* use extremality).

Lemma 5.1. $3[\![eK_3^1]\!]_1 + 3[\![\bar{P}_3^E K_3^E]\!]_E \geq 2K_3$.

Proof of Lemma 5.1. Both sides of this inequality can be evaluated as linear combinations of those graphs in $\mathcal{M}_4$ that contain at least one triangle. There are only four such graphs, and the lemma is easily verified by computing coefficients in front of all of them. ⊣

Applying ϕ_0 to the inequality of Lemma 5.1, comparing the result with (37), (38) and re-grouping terms, we get

$$g'(a)\phi_0(\frac{1}{3}\bar{P}_3 + 2[\![e^2]\!]_1) + b(3a-2) \geq 2g'(a)a^2.$$

Next, $\frac{1}{3}\bar{P}_3 + 2[\![e^2]\!]_1 = K_3 + \rho$. This finally gives us

$$b(g'(a) + 3a - 2) \geq a(2a-1)g'(a).$$

But $g'(x) \geq 1$ on $[1/2, 2/3]$, which implies $g'(a) + 3a - 2 > 0$, and also $g(x)$ is a solution to the differential equation

$$g(x)(g'(x) + 3x - 2) = x(2x-1)g'(x).$$

Altogether this implies $b \geq g(a)$ and completes our proof.

§6. Conclusion. As we stated in Introduction, at the moment the formalism developed in this paper is considered by us mostly as a practical tool. Accordingly, the most interesting question is to which extent this calculus will turn out to be useful for solving concrete open problems in asymptotic extremal combinatorics.

However, this paper certainly raises at least one extremely interesting *general* issue. Typical proofs in extremal combinatorics use only finitely many flags of finitely many types. Do there exist results (or conjectures) in this area that are *in principle independent* of such finite methods? Although (for the reasons discussed below) we can not make this question completely precise, our formalism allows us to come rather close to this.

Recall that $\mathcal{A}_\ell^\sigma$ is the linear subspace in $\mathcal{A}^\sigma$ spanned by $\mathcal{F}_\ell^\sigma$. Given a type σ and $\ell \geq |\sigma|$, we can straightforwardly define the notion of a *partial* (σ,ℓ)-*homomorphism* as a linear functional ϕ on $\mathcal{A}_\ell^\sigma$ such that $\phi(F) \geq 0$ for every flag $F \in \mathcal{F}_\ell^\sigma$ and also $\phi(f_1 f_2) = \phi(f_1)\phi(f_2)$ *whenever* $f_1 \in \mathcal{A}_{\ell_1}^\sigma$, $f_2 \in \mathcal{A}_{\ell_2}^\sigma$ *with* $\ell_1 + \ell_2 - |\sigma| \leq \ell$. Then, given a partial $(0,\ell)$-homomorphism $\phi : \mathcal{A}_\ell^0 \longrightarrow \mathbb{R}$, for every type σ with $|\sigma| \leq \ell$ and $\phi(\sigma) > 0$ we can define its *partial* (σ, ℓ)-*extension* as a probability measure on the set of all partial (σ, ℓ)-homomorphisms $\phi^\sigma : \mathcal{A}_\ell^\sigma \longrightarrow \mathbb{R}$ such that (17) holds for all $f \in \mathcal{A}_\ell^\sigma$. Now we can also define *ensembles of random partial ℓ-homomorphisms* etc. Note, however, that the proof of Theorem 3.5 (as well as many other keystone results in Section 3.2) completely breaks down for partial homomorphisms and their extensions. In fact, it is easy to see that for partial $(0, \ell)$-homomorphisms ensembles may not exist, or they may exist but not be unique.

The first (well-defined) approximation to the question we are trying to capture is this:

Question 1. Assume that $f \in \mathcal{A}_\ell^\sigma$ and $f \geq_\sigma 0$. Does there exist $L \geq \ell$ such that $\phi(f) \geq 0$ for any partial $(0, L)$-homomorphism ϕ for which there exists (at least one) ensemble of partial L-homomorphisms rooted at ϕ?

This question makes more sense (and may turn out more difficult) than it may appear on the first glance. The reason is that this framework already captures all arguments solely based on Cauchy-Schwarz. More precisely, if $f \in \mathcal{A}_{\ell_1}^\sigma$ and $F \in \mathcal{F}_{\ell_2}^\sigma$ satisfy $2\ell_1 + \ell_2 - 2|\sigma| \leq \ell$ then every $(0, \ell)$-partial homomorphism ϕ can be evaluated at $[\![f^2 F]\!]_\sigma$ and, clearly,

$$\phi([\![f^2 F]\!]_\sigma) \geq 0 \tag{39}$$

if ϕ admits at least one partial (σ, ℓ)-extension. Thus, a weaker[5] version of Question 1 is this:

Question 0. Does there always exist $L \geq \ell$ such that $\phi(f) \geq 0$ for every partial $(0, L)$-homomorphism satisfying all conditions of the form (39)?

By analogy with research on algebraic proof systems (see e.g. [GHP]), it is also very natural to consider the dynamic version of this question. Define the *weak Cauchy-Schwarz calculus* that operates with statements of the form $f \geq_\sigma 0$ ($f \in \mathcal{A}^\sigma$), has axioms $F \geq_\sigma 0$ ($F \in \mathcal{F}^\sigma$), $f^2 \geq_\sigma 0$ ($f \in \mathcal{A}^\sigma$) and inference rules

$$\frac{f \geq_\sigma 0 \qquad g \geq_\sigma 0}{\alpha f + \beta g \geq_\sigma 0} \ (\alpha, \beta \geq 0),$$

$$\frac{f \geq_\sigma 0 \qquad g \geq_\sigma 0}{fg \geq_\sigma 0},$$

$$\frac{f \geq_\sigma 0}{[\![f]\!]_{\sigma,\eta} \geq_{\sigma|_\eta} 0},$$

$$\frac{f \geq_{\sigma|_\eta} 0}{\pi^{\sigma,\eta}(f) \geq_\sigma 0}.$$

It is not clear whether this calculus is powerful enough to prove Theorem 3.14 (whence the adjective "weak"). Let the *Cauchy-Schwarz calculus* be obtained from its weak version by explicitly appending all instances of Theorem 3.14 as new axioms.

[5]While comparing questions in this section in their strength, we always assume a negative answer to all of them.

The following looks[6] stronger than Question 0:

Question 2. Is (weak) Cauchy-Schwarz calculus complete?

From the perspective of partial homomorphisms, proofs of the theorems 3.17, 3.18 are also absolutely non-constructive and completely break down. If we want to incorporate these arguments into our framework, we must do so explicitly. Namely, let us call an ensemble of random partial ℓ-homomorphisms *regular* if it satisfies Theorem 3.17[7] and Theorem 3.18 a) (for $f \in \mathcal{A}_\ell^{\sigma_2}$). Then the following is stronger than both Question 1 and Question 2:

Question 3. Same as Question 1, with the difference that the ensemble is additionally required to be regular.

We, however, do not know if the answer to Question 3 is affirmative even for the result about triangle density from Section 5. We could keep defining more and more restrictions on ensembles of random partial L-homomorphisms, Theorems 4.3, 4.5 being the first candidates. But we feel that this would become more and more arbitrary, so instead we would like to finish with the following (which, unlike Questions 0-3 above is *not* well-defined).

Question 4. Is there any set of reasonable and efficient conditions on ensembles of random partial ℓ-homomorphisms such that:

Soundness: Every ensemble that is a projection of an ensemble of total homomorphisms satisfies these conditions.

Completeness: If $f \in \mathcal{A}_\ell^0$ and $f \geq 0$ then there exists $L \geq \ell$ such that $\phi(f) \geq 0$ for every $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ for which there exists at least one ensemble of random partial homomorphisms rooted at ϕ and satisfying these conditions.

Of course, it would be even better to answer the following:

Question 5. Completely describe, in reasonable and efficient terms, those ensembles of random partial homomorphisms that are projections of ensembles of total homomorphisms.

This, however, looks at the moment completely hopeless.

§7. Added in proof. In *On the Minimal Density of Triangles in Graphs* (manuscript available at `http://www.mi.ras.ru/~razborov/triangles.pdf`) we have proved that the bound (35) is tight for *any* t, thereby completely solving the problem of determining the minimal possible density of triangles in a graph with given edge density. The proof builds upon the easy case $t = 2$ from Section 5, and the novelty basically consists in a much more refined analysis of the extremal homomorphism ϕ_0. This analysis essentially uses both the homomorphism π^{F_0} from Section 2.3.2, as well as ensembles of random homomorphisms (Section 3.2).

[6] Formally they seem to be incomparable, since in Question 0 we can also take advantage of the fact that ϕ itself is partial $(0, L)$-homomorphism.

[7] More precisely, in its notation we require the equivalence of *restrictions* of $\boldsymbol{\phi}^{\boldsymbol{\sigma_2},\boldsymbol{\eta\eta_1}}\pi^{\boldsymbol{\sigma_2},\eta}$ and $\boldsymbol{\phi}^{\boldsymbol{\sigma_1},\boldsymbol{\eta_1}}$ *onto* $\mathcal{A}_{\ell+k_1-k_2}^{\sigma_1}$.

Acknowledgement. I am immensely grateful to Peter Winkler and Lior Silberman for carefully reading the paper and proposing many corrections greatly improving both its English and general readability. I am also thankful to Avi Wigderson and an anonymous referee for useful suggestions.

REFERENCES

[BCR] J. Bochnak, M. Coste, and M. Roy, ***Real algebraic geometry***, Springer-Verlag, 1998.

[Bol] B. Bollobás, *Relations between sets of complete subgraphs*, ***Proc. Fifth British Comb. Conference***, 1975, pp. 79–84.

[Bon] J. A. Bondy, *Counting subgraphs: A new approach to the Caccetta-Häggkvist conjecture*, ***Discrete Math.***, vol. 165/166 (1997), pp. 71–80.

[Bor] B Borchers, *CSDP, a C library for semidefinite programming*, ***Optimization Methods and Software***, vol. 11 (1999), no. 1, pp. 613–623.

[BCL*] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi, *Counting graph homomorphisms*, ***Topics in discrete mathematics*** (M. Klazar, J. Kratochvil, M. Loebl, J. Matousek, R. Thomas, and P. Valtr, editors), Springer, 2006, pp. 315–371.

[CaH] L. Caccetta and R. Häggkvist, *On minimal digraphs with given girth*, ***Congressus Numerantium***, vol. 21 (1978), pp. 181–187.

[GGW] F. R. K. Chung, R. L. Graham, and R. M. Wilson, *Quasi-random graphs*, ***Combinatorica***, vol. 9 (1989), pp. 345–362.

[CaF] D. De Caen and Z. Füredi, *The maximum size of 3-uniform hypergraphs not containing a Fano plane*, ***Journal of Combinat. Theory, Ser. B***, vol. 78 (2000), pp. 274–276.

[ErSp] P. Erdös and J. Spencer, ***Probabilistic methods in combinatorics***, Academic Press, 1974.

[ErSi] P. Erdös and M. Simonovits, *Supersaturated graphs and hypergraphs*, ***Combinatorica***, vol. 3 (1983), no. 2, pp. 181–192.

[Fel] D. Felix, *Asymptotic relations in flag algebras*, Manuscript, 2005.

[Fish] D. Fisher, *Lower bounds on the number of triangles in a graph*, ***Journal of Graph Theory***, vol. 13 (1989), no. 4, pp. 505–512.

[FLS] M. Freedman, L. Lovász, and A. Schrijver, *Reflection positivity, rank connectivity, and homomorphism of graphs*, Manuscript, 2004.

[Goo] A. W. Goodman, *On sets of acquaintances and strangers at any party*, ***American Mathematical Monthly***, vol. 66 (1959), no. 9, pp. 778–783.

[GHP] D. Yu. Grigoriev, E. A. Hirsch, and D. V. Pasechnik, *Complexity of semialgebraic proofs*, ***Proceedings of the* 19*th Annual Symposium on Theoretical Aspects of Computer Science***, 2002, pp. 419–430.

[LoSi] L. Lovász and M. Simonovits, *On the number of complete subgraphs of a graph, II*, ***Studies in pure mathematics***, Birkhaüser, 1983, pp. 459–495.

[LoSz] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, ***Technical Report TR-2004-79***, Microsoft Research, August 2004.

[Man] W. Mantel, *Problem 28, solution by H. Gouwentak, W. Mantel, J. Teixeira de Mattes, F. Schuh and W.A. Wythoff*, ***Wiskundige Opgaven***, vol. 10 (1907), pp. 60–61.

[Pod] V. V. Podolski, *(in Russian)*, ***Master's thesis***, Moscow State University, Moscow, 2006.

[Tur] P. Turán, *Egy gráfelméleti szélsöértékfeladatról*, ***Mat. és Fiz. Lapok***, vol. 48 (1941), pp. 436–453.

SCHOOL OF MATHEMATICS
INSTITUTE FOR ADVANCED STUDY
PRINCETON, NJ 08540, USA
and
STEKLOV MATHEMATICAL INSTITUTE
MOSCOW, RUSSIA
E-mail: razborov@math.ias.edu