WHAT IS...

a Flag Algebra?

Alexander A. Razborov

Before attempting to answer the question from the title, it would be useful to say a few words about another question: what kind of problems have flag algebras been invented for?

Let us consider three similar combinatorial puzzles. Assume that we have a (simple, undirected) graph with n vertices. What is the minimal number of edges m (as a function in n) that guarantees the existence of a triangle? Also, assuming that m is above this threshold, *how many* triangles are guaranteed to exist? Let us now offset everything by one, and instead of graphs consider (simple) 3-graphs, i.e., sets of unordered triples (called 3-*edges*) on n vertices. We again ask, what is the minimal number m of 3-edges that guarantees the existence of four vertices such that all four possible triples spanned by these vertices are in the set of 3-edges?

The subarea of discrete mathematics that deals with questions of this sort is called *extremal combinatorics*, and it is very strategically located at a crossroads between "pure" mathematics and its applications. One good way to describe flag algebras is as an attempt to expose and emphasize some common mathematical structure underlying many standard techniques in extremal combinatorics, and a survey of *concrete* results obtained in this way can be found in [1]. Before going into more detail, however, let me encourage the reader to put this article aside and try to predict the current status of the three problems from the previous paragraph.

Ready? The first problem (on the threshold value $m(n)$) was solved in a classic paper by Mantel published in 1907. The second problem (on the minimal number of triangles beyond the threshold) had been open for some forty years. It was asymptotically solved only recently using flag algebras. The generalization to 3-graphs was suggested by Turán in another classic paper written in 1941; he conjectured that $m \leq (\frac{5}{9} + o(1))\binom{n}{3}$ and gave the first construction attaining this bound. This conjecture remains unsolved despite repeated attempts by many strong researchers, and it has greatly stimulated the development of the whole field. Some partial results toward Turán's $(3,4)$-conjecture (by the way, its analogues are open for any values of the parameters 4 and 3 as long as $4 > 3 > 2$), though, were obtained with the help of flag algebras. As Sidorenko, one of the leading experts in the area, put it in his survey dated 1995, "The general problem of Turán having an *extremely* simple formulation but being extremely hard to solve, has become one of the most fascinating *extremal problems* in combinatorics."

Now, let us do one of the many proofs of Mantel's result: it will serve as a motivating running example for our definitions. Let $d_1, \ldots, d_n$ be vertex degrees in our graph, and let $e(v) \approx d_v/n$ be the *relative* degree of the vertex $v \in \{1, 2, \ldots, n\}$. (Strictly speaking, $e(v) = d_v/(n-1)$, but systematically ignoring low-order terms is one of the most basic principles of the theory we are discussing.) We

Alexander A. Razborov is the Andrew MacLeish Distinguished Service Professor in the Department of Computer Science at the University of Chicago. Part of this work was done while the author was at the Steklov Mathematical Institute, supported by the Russian Foundation for Basic Research, and at Toyota Technological Institute, Chicago. His email address is razborov@cs.uchicago.edu.

DOI: http://dx.doi.org/10.1090/noti1051

have

(1)

$$0 \le \frac{1}{n}\sum_{v=1}^{n}(e(v)-\frac{1}{2})^2 \approx \frac{1}{n^3}\sum_v d(v)^2 - \frac{1}{n^2}\sum_v d(v) + \frac{1}{4}.$$

The term $(1/n^2)\sum_v d(v)$ is easy to interpret: it is simply (remember that we are ignoring low-order terms!) the edge density $\rho\ (= m/\binom{n}{2})$. To calculate $(1/n^3)\sum_v d(v)^2$, we use double counting and we look at configurations in our graph spanned by three vertices. We can see there are 0, 1, 2, or 3 edges, and let us denote by $I_3, \bar{P}_3, P_3, K_3$ the respective densities or probabilities with which these configurations occur ("I" stands for "independent", "P" stands for "path", "$\bar{P}$" stands for "complement of a path", and "K" stands for a misspelled "clique"). Then a moment's reflection reveals that $(1/n^3)\sum_v d(v)^2 \approx (1/3)P_3 + K_3$: only these two cases contribute to the sum, and K_3 contributes thrice as much since we have three different choices of v in it. But ρ can also be expressed in these terms as $\rho = \frac{1}{3}\bar{P}_3 + \frac{2}{3}P_3 + K_3$: we generate a random pair of vertices in two steps, first by picking a random *triple* and then by selecting a random pair within this triple. Plugging all this into (1), after simple calculations we conclude that $\rho \le (1/2) - \bar{P}_3 + K_3$. Thus, $\rho > 1/2$ implies the existence of triangles, and, moreover, their density K_3 satisfies $K_3 \ge \rho - 1/2$. If we are slightly more careful and instead of $(1/n)\cdot\sum_{v=1}^{n}(e(v)-1/2)^2$ compute $(1/n)\cdot\sum_{v=1}^{n}(e(v)-\rho)^2$ (that is, the *actual* variance of the degree sequence), we will get a better bound $K_3 \ge \rho(2\rho - 1)$ proved by Goodman in 1959. This latter bound had remained the best known for the second problem on our list until it was superseded in a beautiful paper by Bollobás (1975).

Let us now see which kind of structure we can extract from this template (all this material can be found in [2]). Our target graph G is large and unknown. Thus, for every fixed graph H we introduce a formal real-valued variable with the same name and the intuitive meaning "the density of induced copies of H in G" (in the argument above, H was one of $\rho, I_3, \bar{P}_3, P_3$, or K_3). As we often need to sum these quantities with real coefficients, we form the linear space of formal (finite) linear combinations of these variables. The identity $\rho = \frac{1}{3}\bar{P}_3 + \frac{2}{3}P_3 + K_3$ that we used above can be widely generalized: the density of any fixed graph H can be expressed in terms of densities of graphs with a fixed but larger number of vertices. We factor our space out by these relations. Multiplication is also available: for example, ρ^2 can be expressed as a linear combination of [the densities of] graphs on four vertices; this is done by double counting similar to our calculation of $(1/n^3)\sum_v d(v)^2$. All these developments give us a commutative associative algebra $\mathcal{A}^0$.

Furthermore, when the size of the target graph G tends to infinity, the densities of induced copies of H converge to an algebra homomorphism ϕ from $\mathcal{A}^0$ to $\mathbb{R}$. These algebra homomorphisms possess an extra property that $\phi(H)$, being a density, is nonnegative for any graph H. Then it turns out that we also have an important "completeness result": every *abstract* algebra homomorphism from $\mathcal{A}^0$ to $\mathbb{R}$ with this nonnegativity property (let us call their set $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$) can be obtained from a convergent sequence $\{G_n\}$. In other words, the object $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ defined in the best mathematical traditions quite abstractly nonetheless corresponds *exactly* to the class of extremal problems we intend to study. For example, the second problem on our list can be reformulated like this: given $x \in [0,1]$, compute the minimal possible value of $\phi(K_3)$, where $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ satisfies $\phi(\rho) = x$, and, yes, we do mean min here, not inf, since $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ is compact. The computation *does* begin with the words "let us fix once and for all an extremal ϕ."

Having thus reformulated the questions of study in the appropriate language, the rest of the theory is basically devoted to developing useful syntactic tools for *proving* theorems about the behavior of the limit densities $\phi(H)$. Most of these tools have evolved from analogous methods employed in finite arguments, but, again, the mathematical structure allows us to search for the desired proofs either completely automatically or in an interactive computer-human mode. This allows us to expand the search space by an order of (literally!) hundreds or thousands. We once more refer to [1] for a survey of concrete results that have been obtained with this method, and now we review some of these tools.

Our account above was tailored to the case of ordinary graphs, but this was done only for simplicity of exposition. The theory of flag algebras was deliberately set up in such a way that it applies in a uniform way to *arbitrary* combinatorial structures: hypergraphs, directed graphs, mixtures of these, colored versions of these, you name it. In logical terms, the set-up can be described as a "universal theory in a language containing only predicate symbols," but the only property that is actually needed is that a subset of vertices of a model spans an (induced) submodel. This is paramount since structures other than simple graphs are where the most important *open* problems in the area reside, Turán's $(3,4)$-problem being just the tip of the iceberg.

Next, our definitions can be readily generalized to the case when all our models are required to contain k distinguished "base" vertices (for an analogy, the reader should think of base points in algebraic topology) that must be preserved by all mapping involved. This clearly makes sense only when we also specify what we see on $c_1, \ldots, c_k$ itself; for example, if $k = 2$, then in the graph theory we should specify whether (c_1, c_2) is an edge or not. Thus, we say that a *type* σ is simply a model with the ground set $\{1, 2, \ldots, k\}$, a *flag*[1] of type σ is a model with k base vertices respecting the structure of σ, and then (finally!) the *flag algebra* $\mathcal{A}^\sigma$ and $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ are defined just as before. As an example, let 1 be the (only) type of size 1, and let $e \in \mathcal{A}^1$ be an edge with one distinguished vertex in it. Then, instead of computing the sum $(1/n^3)\sum_{v=1}^n d(v)^2$ in (1), we could alternatively express first $e^2 = K_3^1 + P_3^{1,c}$, where K_3^1 and $P_3^{1,c}$ are 1-flags that are obtained from K_3, P_3, respectively, by adding one base vertex which is the center of the path in the case of P_3.

Every flag algebra $\mathcal{A}^\sigma$ has a good supply of elements that are guaranteed to be evaluated to a nonnegative value by any $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$: these are the squares f^2 and their positive linear combinations. We can also define a linear *averaging operator* $[\![\cdot]\!]_\sigma$ that generalizes the summation in (1). For example, $[\![K_3^1]\!]_1 = K_3$, $[\![P_3^{1,c}]\!]_1 = (1/3)P_3$, which gives us another proof of what we did above differently, $[\![e^2]\!]_1 = (1/3)P_3 + K_3$. Now we have a healthy supply of nonnegative elements in $\mathcal{A}^0$, too: these are $[\![f^2]\!]_\sigma$, where $f \in \mathcal{A}^\sigma$ for some type σ, and their positive linear combinations, possibly mixing up relations coming from different types.

It already turns out that these simple ideas (called in [1] "plain methods") can solve many open problems if you look deep enough, i.e., for $f^2 \in \mathcal{A}^\sigma$ that involve flags on sufficiently many vertices (from four to six in a typical application). It is obvious from the look of most of these results that they could hardly be obtained by hand and that they do require computer assistance. Fortunately, the question of "how to represent a specific f in the form $\sum_i \alpha_i [\![f_i^2]\!]_{\sigma_i}$ $(\alpha_i \geq 0)$" is completely answered by semidefinite programming (SDP), and, equally fortunately for flag algebras, for the latter we have not only theoretical results (polynomial time algorithms) but also good noncommercial packages that really do the job. In my own work I most often use CSDP, a package developed by Brian Borchers (`https://projects.coin-or.org/Csdp`); it is also used in the publicly available *flagmatic software* (`http://www.maths.qmul.ac.uk/~ev/flagmatic/`) by Emil R. Vaughan. One thing that greatly hinders these developments is the absence (to the best of my knowledge) of generic SDP-solvers that also provide certificates of feasibility/unfeasibility: many actual calculations in flag algebras are performed so close to the border between them that numerical results often cannot be trusted.

Besides purely computational convenience, flag algebras lead to more sophisticated structures and objects that also have found applications in concrete proofs; due to lack of space we can only name some of them here (see [2] for all missing details).

There are many useful constructions that allow us to convert combinatorial objects of one sort into objects of another sort. For example, given a directed graph, we can view it as an ordinary graph by erasing its orientation, or, given a 3-graph and a vertex in it, we can look at its link, which is again an ordinary graph. Constructions of this kind are captured by the logical notion of an interpretation, and in the language of flag algebras they lead to algebra homomorphisms between $\mathcal{A}^\sigma$ for different theories. This allows us to conveniently move around theorems (i.e., statements of the form $f \geq 0$) from one context to another.

The space $\mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ is compact, which implies that, for every extremal problem, there exists an *individual* optimal solution to it. Extremality can be exploited in a variety of ways: for example, one can write (an analogue of) a functional derivative according to some intuitive changes and use its equality to zero as a new useful relation not necessarily possessed by an arbitrary $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$.

Tuples of densities $\phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R})$ encode a surprising amount of information about the intended object. For example, if we "pick" a base vertex (or, more generally, a copy of a more complicated type σ) at random, then it "should" give rise to a probability distribution over $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. It turns out that this information can already be uniquely retrieved from ϕ even when we do not really have a good idea which space we are sampling from. For example, we can straightforwardly (say, avoiding Szemeredi's Regularity Lemma) determine the fraction of vertices in a graph that have relative degree at least $1/3$ using ϕ only.

It is tempting to employ a similar axiomatic approach in other areas where the Cauchy-Schwarz inequality is used. The difficulties arising here are of a more technical and practical nature: since the group of symmetries is not as rich in those

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

situations, it is more difficult to come up with a calculus that is good not only theoretically but also allows us to get new concrete results. Some work in this direction, however, has already been done; see [1, Section 4.2].

Last, but definitely not least, we are often interested not only in the properties of the limit densities ϕ itself, but also in what is the *actual* limit object these densities correspond to (or, in more logical terms, in the associated *model theory*). This leads to the deep and beautiful theory of *graph limits* with many connections to other disciplines, and we highly recommend Lovász's recent book [3] for an introduction to the subject. We strongly feel that emphasizing more connections between the syntactical (flag algebras) and semantical (graph limits) approaches to the same class of objects should be very beneficial for both.

References

[1] A. Razborov, Flag algebras: An interim report, 2013, to appear in the *Erdős Centennial Volume*, preliminary text available at http://people.cs.uchicago.edu/~razborov/files/flag_survey.pdf.

[2] ———, Flag algebras, *Journal of Symbolic Logic*, 72(4):1239–1282, 2007.

[3] L. Lovász, *Large Networks and Graph Limits*, American Mathematical Society, 2012.