Weak saturation and related matroids

Mykhaylo Tyomkyn (Charles University)*

We study the following deterministic process, or one-player game, on the edges of a complete graph.

Definition 1 (Weak saturation. Bollobás '68) *For two graphs G and H we say that G is* weakly H-saturated *if there exists an ordering $e_1, \ldots, e_m$ of $\binom{V(G)}{2} \setminus E(G)$ such that for each $\ell \in [m]$ the graph $G_\ell = G \cup \{e_1, \ldots, e_\ell\}$ contains a copy of H not contained in $G_{\ell-1}$. The* weak saturation number $\mathrm{wsat}(n, H)$ *is the minimal number of edges in a weakly H-saturated graph on n vertices.*

That is, in a weakly H-saturated graph G the missing edges can be added one by one, with every new edge creating a new copy of H.

Example 1 *For $H = K_3$ a graph G is weakly saturated if and only if G is connected. Therefore,* $\mathrm{wsat}(n, K_3) = n - 1$*, and the extremal examples are the n-vertex trees.*

What about $H = K_k$ more generally? The graph $S_{k,n} = K_n \setminus K_{n-k+2}$ is weakly K_k-saturated, and so

$$\mathrm{wsat}(n, K_k) \le \binom{n}{2} - \binom{n-k+2}{2} = (k-2)n - \binom{k-1}{2}.$$

This turns out to be sharp.

Theorem 1 (Frankl '82, Kalai '84)

$$\mathrm{wsat}(n, K_k) = \binom{n}{2} - \binom{n-k+2}{2}.$$

But how to prove a matching lower bound, or any lower bound for that matter? In the $k = 3$ case we were 'lucky' to use graph connectivity, but this does not extend straightforwardly to larger k. In particular, we do not have a similarly neat characterization of extremal examples. Generalizing the $S_{k,n}$ example, one can quickly see that any graph obtained by starting with K_{k-1} and successively adding vertices of degree $k - 2$ is weakly K_k-saturated, and of the same size. However, for $k \ge 4$ this does not give the full family of extremal examples.

Exercise 1 *Find a graph of order $n = 7$, size $2n - 3$, weakly K_4-saturated, and of minimum degree 3.*

To date, all known proofs of Theorem 1 use tools from algebra or geometry; no purely combinatorial proof is known for $k \ge 7$. One possible explanation is that the extremal examples exhibit no apparent stability: they are dissimilar. Since combinatorial proofs tend to give structural information about the extremal cases as a by-product, so we are less likely to find one here. This does not mean it is not possible — after all, it *was* for the triangle, but one would need to work harder. Even the next case $H = K_4$ is already a challenge.

*REU lecture. Prague, 07/22/2026

Challenge 1 *Prove combinatorially that* $\mathrm{wsat}(n, K_4) = 2n - 3$.

So, how to prove Theorem 1, and lower bounds more generally? Kalai, based on prior work of Lovász, came up with an ingenious method that converts an a priori universal problem into a *constructive* one.

Theorem 2 (Kalai's principle, vector space form) *Let H be a graph and let V be a vector space over a field. Suppose there is a mapping $f : E(K_n) \to V$ such that for every isomorphic copy $H' \subseteq K_n$ of H the vectors $\{f_e = f(e) : e \in E(H')\}$ satisfy*

$$\sum_{e \in E(H')} \lambda_e f_e = 0, \tag{1}$$

*with $\lambda_e \neq 0$ for **all** $e \in E(H')$. Then*

$$\mathrm{wsat}(n, H) \geq \dim \mathrm{span}\{f_e : e \in E(K_n)\}.$$

Note that in practice we may assume that $V = \mathrm{span}\{f_e : e \in E(K_n)\}$.

Proof Let G be a weakly H-saturated graph on n vertices with $|E(G)| = \mathrm{wsat}(n, H)$ many edges. Let $e_1, \ldots, e_m$ be a saturating edge sequence for G, and let $G_i = G + \{e_1, \ldots, e_i\}$ for each i, so that $G_0 = G$ and $G_m = K_n$. We claim that for $i = m-1, \ldots, 0$ we have

$$f(e_{i+1}) \in \mathrm{span}\{f_e : e \in E(G_i)\}.$$

Indeed, adding e_{i+1} to G_i creates, by definition, a new copy H' of H, necessarily containing e_{i+1}. Then (1) implies that the vector $f(e_{i+1})$ is a linear combination of the vectors $\{f_e : e \in E(H') \setminus \{e_{i+1}\}\} \subseteq \{f_e : e \in E(G_i)\}$. Note that here we crucially used the fact that all coefficients in (1) are non-zero.

Therefore, we obtain

$$\mathrm{span}\{f_e : e \in E(G_i)\} = \mathrm{span}\{f_e : e \in E(G_{i+1})\},$$

and inductively

$$\mathrm{span}\{f_e : e \in E(G)\} = \mathrm{span}\{f_e : e \in E(G_0)\} = \mathrm{span}\{f_e : e \in E(G_m)\} = \mathrm{span}\{f_e : e \in E(K_n)\}.$$

It follows that

$$\mathrm{wsat}(n, H) = |E(G)| \geq \dim \mathrm{span}\{f_e : e \in E(G)\} = \dim \mathrm{span}\{f_e : e \in E(K_n)\}.$$

$\square$

In order to apply Theorem 2, we need to construct a vector space V and a mapping f satisfying (1). The larger the dimension of V, the better the bound we obtain. Ideally, $\dim V$ should match the size of a constructed weakly H-saturated n-vertex graph G, in which case we would have determined $\mathrm{wsat}(n, H)$ exactly.

This naturally leads to a discussion of matroids defined on graphs. To this end, let us first recall the definition(s) of a matroid.

Definition 2 (Matroid) *A* matroid *is a pair $\mathcal{M} = (E, \mathcal{I})$ where E is a finite set, called the set of* elements*, or the ground set, and $\mathcal{I} \subseteq 2^E$ is a family of subsets of E. The members of $\mathcal{I}$ are called the* independent sets*, and $\mathcal{I}$ satisfies the following axioms:*

(I1) $\emptyset \in \mathcal{I}$.

(I2) *If* $A \in \mathcal{I}$ *and* $B \subseteq A$ *then* $B \in \mathcal{I}$.

(I3) *If* $A, B \in \mathcal{I}$ *and* $|A| > |B|$ *then there exists* $x \in A \setminus B$ *such that* $B \cup \{x\} \in \mathcal{I}$.

Two main sources of examples of matroids are vector spaces and graphs.

Example 2 (Linear matroid) *Given a finite set* E *of vectors in some vector space* V*, set* $\mathcal{I}$ *to be the family of all linearly independent subsets of* E*. We know from Linear algebra that* $\mathcal{I}$ *satisfies (I1)-(I3). Often such a set of vectors presents itself as a set of rows (columns) of a matrix* M*. In such cases we speak of a* row matroid *(column matroid) of* M*.*

Example 3 (Graphic matroid) *Let* G *be a finite graph. Define* $E = E(G)$ *and* $\mathcal{I}$ *to be the set of all sub-forests of* G*. We know from Graph theory that* $\mathcal{I}$ *satisfies (I1)-(I3).*

Looking at both examples, one can probably guess that there is a lot more to matroids than just the independent sets. The following generalizes the notion of dimension of a vector space.

Definition 3 (Rank function) *Given a matroid* $\mathcal{M} = (E, \mathcal{I})$ *for every set* $A \subseteq E$ *its* rank $r(A)$ *is the size of the largest independent subset of* A*. The rank of* $\mathcal{M}$ *is defined as* $r(\mathcal{M}) = r(E)$*.*

The properties of the rank function implied by the axioms of independent sets are as follows.

(R1) $0 \leq r(A) \leq |A|$

(R2) If $A \subseteq B$ then $r(A) \leq r(B)$

(R3) $r(A \cup B) + r(A \cap B) \leq r(A) + r(B)$

It turns out that properties (R1)-(R3) of an integer-valued function on 2^E define a matroid uniquely, and thus can be taken as an alternative definition. To obtain $\mathcal{I}$ from r, set $\mathcal{I} = \{I \subseteq E : r(I) = |I|\}$. This is an example of a *cryptomorphism* between two different matroid axiomatizations.

There is also a generalization of the notion of a graph cycle.

Definition 4 (Circuit) *A circuit in a matroid* $(E, \mathcal{I})$ *is a set* $C \subseteq E$ *such that* $C \notin \mathcal{I}$ *but* $C \setminus \{x\} \in \mathcal{I}$ *for every* $x \in C$*.*

The family of circuits $\mathcal{C}$ has the following properties, which again describe the matroid uniquely.

(C1) $\emptyset \notin \mathcal{C}$.

(C2) For any distinct $C_1, C_2 \in \mathcal{C}$ we have $C_1 \nsubseteq C_2$ and $C_2 \nsubseteq C_1$.

(C3) For any distinct $C_1, C_2 \in \mathcal{C}$, if $e \in C_1 \cap C_2$, then there exists $C_3 \in \mathcal{C}$ such that $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\}$.

Exercise 2 (Fundamental circuit) *Prove that if* $A \in \mathcal{I}$ *and* $A \cup \{x\} \notin \mathcal{I}$ *then* $A \cup \{x\}$ *contains a unique circuit as a subset.*

This is only the tip of the iceberg: a matroid can be defined in a multitude of ways: using bases (maximal independent sets), flats (subspaces), hyperplanes, to name a few. All of these definitions are pairwise equivalent, via appropriate cryptomorphisms. As such, when speaking of a matroid, one usually writes $\mathcal{M}$ meaning *all* of the above: $E, \mathcal{I}, r, \mathcal{C}$ etc.

With matroids, Theorem 2 can be rephrased entirely in the language of set systems as follows.

Theorem 3 (Kalai's principle, general form) *Let H be a graph without isolated vertices, and let $\mathcal{M}$ be a matroid with $E(\mathcal{M}) = E(K_n)$. Suppose that*

$$E(H') \in \mathcal{C}(\mathcal{M})$$

for every copy $H' \subseteq K_n$ of H. Then

$$\mathrm{wsat}(n, H) \geq r(\mathcal{M}).$$

If $\mathrm{wsat}(n, H)$ can be exactly determined in this way, we say that H is *weak saturation matroidal* (wsm, for short). Let us now consider some examples.

It is easy to see that $H = K_3$ is wsm. The underlying matroid is the graphic matroid of the complete graph K_n, where every triangle is a circuit and the rank is $n - 1$. The next case $H = K_4$ is already much more interesting.

Definition 5 (Count matroid) *Let k, ℓ be integers with $k \geq 1$ and $-\infty < \ell \leq 2k-1$. A (k, ℓ)*-count matroid *$\mathcal{M}$ on $E = E(K_n)$ is defined by (interpreting edge-sets as graphs without isolated vertices) $G \in \mathcal{I}$ if and only if*

$$|E(G')| \leq k|V(G')| - \ell \quad \text{for all} \quad \emptyset \neq G' \subseteq G. \tag{2}$$

Graphs G satisfying (2) are called (k, ℓ)*-sparse*. If, in addition, equality in (2) holds for G itself, we say that G is (k, ℓ)*-tight*.

Example 4 *The $(1, 1)$-sparse and $(1, 1)$-tight graphs are the forests and trees, respectively.*

Example 5 *The $(2, 3)$-count matroid $\mathcal{R}$ is a witness that K_4 is* wsm. *Indeed, K_4 is a circuit since removing any edge would make it $(2, 3)$-tight. Hence, $\mathrm{wsat}(n, K_4) \geq r(\mathcal{R}) = 2n - 3$, which matches the upper bound constructions.*

The matroid $\mathcal{R}$ is known as the *generic plane rigidity matroid*. As the name suggests, it is closely linked to the question of whether a graph, realized as a *bar-joint framework* in $\mathbb{R}^2$, is rigid or flexible.

Exercise 3 *Prove that $\mathcal{R}$ is indeed a matroid. Hint: work with the circuit axioms.*

The bases (maximal independent sets) of $\mathcal{R}$, the $(2, 3)$-tight graphs, are also known as the *Laman graphs*. They have an alternative characterization, involving the so-called *Henneberg extensions*.

- K_2 is a Laman graph
- Adding a new degree-2 vertex to a Laman graph results in a Laman graph
- Adding a new degree-3 vertex to a Laman graph and deleting an edge between two of its neighbors results in a Laman graph.

Moreover, every Laman graph can be obtained from K_2 via a sequence of Henneberg extensions.

Exercise 4 *Prove that the Laman graphs defined via Henneberg extensions are precisely the* $(2,3)$*-tight graphs. One direction should be very easy; the other is a bit more challenging.*

There is a second matroid that can be used to prove that $\mathrm{wsat}(n, K_4) = 2n - 3$.

Definition 6 (Hyperconnectivity matroid) *Let* $\mathbf{p_1}, \dots, \mathbf{p_n}$ *be* n *points in* $\mathbb{R}^2$ *in generic position (all coordinates are algebraically independent over* $\mathbb{Q}$*). The* plane hyperconnectivity matroid $\mathcal{H}$ *is the row matroid of the* $\binom{n}{2} \times 2n$ *matrix*

$$\begin{pmatrix} \mathbf{p_2} & -\mathbf{p_1} & 0 & \dots & 0 & 0 \\ \mathbf{p_3} & 0 & -\mathbf{p_1} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \mathbf{p_n} & 0 & 0 & \dots & 0 & -\mathbf{p_1} \\ 0 & \mathbf{p_3} & -\mathbf{p_2} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{p_n} & -\mathbf{p_{n-1}} \end{pmatrix}$$

Exercise 5 *Show that* K_4 *and* $K_{3,3}$ *are* $\mathcal{H}$*-circuits and that adding degree-*2 *vertices preserves* $\mathcal{H}$*-independence, and therefore* $r(\mathcal{H}) = 2n - 3$*.*

Note that $K_{3,3}$ is $(2,3)$-tight, and therefore $\mathcal{R}$-independent, but it is an $\mathcal{H}$-circuit. Hence $\mathcal{H} \neq \mathcal{R}$. Defining the analogous matrix and matroid using points in $\mathbb{R}^{k-2}$ in fact gives a proof of Theorem 1 in full. This was essentially Kalai's original proof (using, instead of matrices, the language of exterior algebra – this way the theorem can be extended to r-uniform hypergraphs). The $\mathcal{R}$-proof can also be generalized to K_k, but one must depart from the count matroids and work with rigidity in $\mathbb{R}^{k-2}$ instead, so the proof is geometric!

In this light, we should ask if every graph is wsm. This was answered in the negative in 2025.

Theorem 4 (Limiting constant: Alon '85) *For every graph* H *there exists a limiting constant*

$$\lim_{n\to\infty} \frac{\mathrm{wsat}(n, H)}{n} =: c_H \geq 0.$$

Theorem 5 (Terekhov–Zhukovskii '25) *If* H *is* wsm*, then* c_H *is an integer.*

Exercise 6 *Let* H *be the Moser spindle.*

1. *Construct a graph showing that* $\mathrm{wsat}(n, H) \leq 5n/3 + o(n)$*.*
2. *Show that a weakly* H*-saturated graph cannot have two adjacent vertices of degree* 2*. Deduce that* $\mathrm{wsat}(n, H) \geq 6n/5 + o(n)$*.*

Lastly, let us remark that weak saturation equally makes sense on arbitrary *host graphs*: For $G \subseteq F$ we call G weakly H-saturated in F if the edges in $E(F) \setminus E(G)$ can be ordered such that adding each new edge creates a new copy of H. Define $\mathrm{wsat}(F, H)$ accordingly. It turns out that determining the weak saturation numbers in this generality is a fairly hopeless task even when H is the triangle.

Theorem 6 (Tancer–T. '25) *Given a graph* F *on* n *vertices as input, it is NP-hard to decide whether* $\mathrm{wsat}(F, K_3) = n - 1$*.*