A PROOF OF KALAI'S CONJECTURE

MICHAEL ZHENG

Abstract. These notes aim to give a proof of Kalai's conjecture based on a proof found by GPT-6 Astra in a more human-readable manner.

1. Introduction

Within extremal graph theory, one of the most studied problems are of *Turán-type* where the quantity $\text{ex}(n, F) = \max\{e(G) : v(G) = n \text{ and } F \not\subseteq G\}$ for an unlabeled graph F is studied. One of the most famous problems within this area which was recently proven by GPT-6 Astra (see [1]) is the Erdős–Sós conjecture from 1963. It predicts that, morally, the density at which an n-vertex graph contains a tree T only depends on $v(T)$ (or equivalently $e(T)$).

Conjecture 1.1 (Erdős–Sós, 1963)**.** Let T be a tree on $t \geqslant 1$ edges. Then any n-vertex, T-free graph G must satisfy $2e(G) \leqslant (t-1)n$. In particular, $\text{ex}(n, T) \leqslant (t-1)n/2$.

Note that this is tight for all n divisible by t as witnessed by $G = \frac{n}{t}K_t$. It is also worth pointing out that the factor of $(t-1)/2$ was of big interest. Indeed, if one were to replace it by $t-1$ in the statement, then by iteratively deleting vertices with degree at most $t-1$, one would obtain a subgraph with minimum degree t where T can be greedily embedded.

In 1984, Kalai posed a natural generalization, recorded in [3], to tight trees: Let $k \geqslant 2$. A k-uniform hypergraph H without isolated vertices is a *tight* k*-tree*, if there is an ordering of the edges $E(H) = \{e_1, \ldots, e_t\}$ such that for $i \geqslant 2$ there exists $v_i \in e_i$ and $s_i < i$ such that $v_i \notin \bigcup_{j<i} e_j$ and $e_i \setminus \{v_i\} \subseteq e_{s_i}$. Note that, in this case, $v(H) = t + k - 1$.

Kalai's conjecture, which is also now proven by GPT-6 Astra, states the following:

Theorem 1.2. Let T be a tight k-tree with $t \geqslant 1$ edges. Then any T-free hypergraph H must satisfy $ke(H) \leqslant (t-1)\,|\partial H|$ where ∂H denotes the *shadow*, the set of all $(k-1)$-sets of $V(H)$ with positive codegree. In particular, for $n \geqslant t + k - 1$

$$\text{ex}(n, T) \leqslant \left\lfloor \frac{t-1}{k} \binom{n}{k-1} \right\rfloor.$$

Department of Mathematics and Computer Science, Emory University, Atlanta, USA
Email address: xiangxiang.michael.zheng@emory.edu.
Date: September 16, 2026.

For more history and partial results such as those in [3, 4], we refer to [4]. Note that this bound tight for infinitely many values by adapting the construction in the graph setting: First, take a $(n, t+k-2, k-1)$-Steiner system S, which exists by the resolution of the existence conjecture of designs for infinitely many n, and then consider the hypergraph H where you replace each $b \in E(\mathbf{s})$ by a complete k-uniform hypergraph. By definition of S, any copy of T must be completely contained in an edge of S, which is impossible since $t+k-2 < v(T)$. Finally,

$$e(H) = \frac{\binom{n}{k-1}}{\binom{t+k-2}{k-1}} \binom{t+k-2}{k} = \frac{t-1}{k} \binom{n}{k-1}.$$

While expository papers for GPT-generated proofs of this former conjecture already exist, see for example [2], I wanted to write my own expository notes for the sake of my own understanding.

2. Proof overview

Fix a tight k-tree T with $t \geqslant 1$ edges. To make the induction go through, we will also fix a root $\mathbf{r} = r_1 \dots r_{k-1} \in \partial T$. For embedding purposes, $\mathbf{r}$ is technically an ordered tuple but we will still nevertheless refer to both the tuple and the underlying set by $\mathbf{r}$. Furthermore, let H be a k-uniform hypergraph on $[n]$, $n \geqslant t+k-1$. Denote by M the set of pairs $(\pi, i) \in S_n \times [k, n]$, $\pi = (\pi_1, \dots, \pi_n)$, such that $\pi_1\pi_2 \dots \pi_{k-1}\pi_i \in E(H)$. We will call such pairs *marked states* or sometimes just *states*. Occasionally, we will refer to π_i as the *mark*. Note that only permutations where the first $k-1$ vertices form a set in ∂H can be involved in pairs of M. We will call such permutations *eligible* and denote the set of eligible permutations by W. It is easy to see that

Proposition 2.1. $|W| = |\partial H|\,(k-1)!(n-k+1)!$ and $|M| = e(H)k!(n-k+1)!$.

Analogous to the 2-uniform case in [1], we will say that a state $(\pi, i) \in M$ is $(T, \mathbf{r})$-*good* if $H[\pi_1, \dots, \pi_i]$ supports a $\pi_1 \dots \pi_{k-1}$-rooted $(T, \mathbf{r})$-copy, i.e. there exists an injective homomorphism $\varphi\colon V(T) \to \{\pi_1, \dots, \pi_i\}$ such that $\varphi(r_j) = \pi_j$ for all $j \in [k-1]$. Otherwise, we call (π, i) $(T, \mathbf{r})$-*bad*. We denote the set of $(T, \mathbf{r})$-bad marked states by $B(T, \mathbf{r}) \subseteq M$ and denote its size by $b(T, \mathbf{r})$.

It is sufficient to show that

Theorem 2.2. $b(T, \mathbf{r}) \leqslant (t-1)\,|W|$.

Proof of Theorem 2.2 $\implies$ Theorem 1.2. If H is T-free, then $b(T, \mathbf{r}) = |M|$, regardless of the choice of $\mathbf{r}$. Using Proposition 2.1, we therefore get

$$e(H)k!(n-k+1)! \leqslant (t-1)\,|\partial H|\,(k-1)!(n-k+1)! \implies e(H) \leqslant \frac{t-1}{k}\,|\partial H|$$

as desired. □

The remainder of the proof is by induction and makes use of a simple decomposition of tight hypertrees into smaller tight hypertrees: Let $\mathfrak{I}(T)$ be the *edge-facet-incidence graph* of T, i.e. the graph with vertex set $E(T) \cup \partial T$ and with $ef \in E(\mathfrak{I}(T))$, $e \in E(T)$, $f \in \partial T$, if and only if $f \subseteq e$.

By induction, one can show the following

Proposition 2.3. $\mathfrak{I}(T)$ is a tree and for every vertex $v \in V(T)$, the vertices of $\mathfrak{I}(T)$ containing v form a subtree.

Thus, we may decompose T for $t \geqslant 2$ as follows: Consider the connected components of $\mathfrak{I}(T) - e_0$ where $e_0 = \{r_0\} \cup \mathbf{r} \in E(T)$ is an arbitrary edge containing $\mathbf{r}$. Each nontrivial component naturally corresponds to a tight subtree of T and by construction, these subtrees' edges partition $E(T)$. Rooting each of these subtrees at the unique $(k-1)$-subset of e_0 of positive codegree, we get

Lemma 2.4. Set $\mathbf{r}^{(i)} = (r_0, \ldots, r_{i-1}, r_{i+1}, \ldots, r_{k-1})$ for all $i \in [0, k-1]$. There exists an index set $I \subseteq [0, k-1]$ and nonempty tight k-trees S_i rooted at $\mathbf{r}^{(i)}$ such that

$$E(T) = \{e_0\} \cup \dot{\bigcup_{i \in I}} E(S_i),$$

$$V(S_i) \cap e_0 = \mathbf{r}^{(i)}, \qquad (i \in I)$$

$$V(S_h) \cap V(S_i) = \mathbf{r}^{(h)} \cap \mathbf{r}^{(i)}. \qquad (h, i \in I, h \neq i)$$

Note that in the notation above, $\mathbf{r}^{(0)} = \mathbf{r}$ and that $I \neq \emptyset$ provided that $t \geqslant 2$. Also, crucially, the last two equalities guarantee that given an embedding of $\mathbf{r}$, the remaining vertices of T can be embedded disjointly from one another; in particular if we are given "nice" embeddings of each subtree that agree on the image of $\mathbf{r}$ and otherwise map to disjoint subsets of $V(H)$, they naturally form a "nice" embedding of T.

Proof. It remains to show the latter two equalities. For the former, it suffices to show that $r_i \notin V(S_i)$: Assume for the sake of contradiction that $r_i \in V(S_i)$ and fix $r_i \in e \in E(S_i)$. Note that $e_0 \neq e$. Since the vertices in $\mathfrak{I}(T)$ containing r_i induce a subtree, all the vertices on the (unique) path between e and e_0 must contain r_i. But $\mathbf{r}^{(i)}$ is on that path and does not contain r_i by definition. ↯

For the latter, suppose that $r_h \in V(S_i)$; the case $r_i \in V(S_h)$ is handled similarly. Again, fix $r_h \in e \in E(S_i)$. Since the set of vertices in $\mathfrak{I}(T)$ containing r_h form a subtree, all the vertices on the (unique) path between e and $\mathbf{r}^{(i)}$ must contain r_h. But $\mathbf{r}^{(h)}$ must be on that path and does not contain r_h by definition. ↯ □

The main technical lemma thus is the following:

Lemma 2.5. In the setting of Lemma 2.4,

$$b(T, \mathbf{r}) \leqslant \sum_{i \in I} b(S_i, \mathbf{r}^{(i)}) + |I|\,|W|\,.$$

Proof of Lemma 2.5 $\implies$ Theorem 2.2. We proceed by induction on $t = e(T)$. If $t = 1$, then any pair $(\pi, i) \in M$ is good, so $b(T, \mathbf{r}) = 0$. For the induction step where $t \geqslant 2$, apply Lemma 2.4 on T and then Lemma 2.5 to get

$$b(T, \mathbf{r}) \leqslant \sum_{i \in I} b(S_i, \mathbf{r}^{(i)}) + |I|\,|W|\,.$$

By the induction hypothesis, $b(S_i, \mathbf{r}^{(i)}) \leqslant (e(S_i) - 1)\,|W|$ for all $i \in I$. Thus, we get

$$b(T, \mathbf{r}) \leqslant \left(\sum_{i \in I} (e(S_i) - 1) + |I| \right) |W| \leqslant |W| \sum_{i \in I} e(S_i) = (t-1)\,|W|$$ □

3. The proof of Lemma 2.5

For $x = x_0 \dots x_{k-1} \in E(H)$ and $i \in [0, k-1]$, we define $x^{(i)}$ as in the setting of Lemma 2.4. We will make use of two injection maps. The first one is straightforward: For distinct $h, i \in [0, k-1]$, let $\varphi_{\vec{ih}} \colon M \to M$ be the function defined by $\varphi_{\vec{ih}}(\pi, j) = (\pi', j)$ where $\pi = x^{(i)} A x_i B$, $\pi' = x^{(h)} A x_h B$, $\pi_j = x_i$ and $\pi'_j = x_h$. In other words, given (π, j), x_i and x_h get swapped such that, aside from x_h, the order on x's vertices in π' agrees with x. Note that the underlying k-tuple x is implicitly given through (π, j) and i; x_i is given by π_j and the first $k-1$ images of π make up the remaining entries of x in that order. That this is a map of type $M \to M$ follows immediately by construction, and injectivity follows from the simple observation that $\mathrm{id}_M = \varphi_{\vec{hi}} \circ \varphi_{\vec{ih}}$. In particular, $\varphi_{\vec{ih}}$ is even a bijection.

Fix now a rooted k-uniform hypergraph $(S, \mathbf{s})$. For the second injection, we partition the $(S, \mathbf{s})$-good states in $M \setminus B(S, \mathbf{s})$ into $(S, \mathbf{s})$-*first-good* states $\mathcal{F}(S, \mathbf{s})$, i.e. the set of $(S, \mathbf{s})$-good states (π, i) where i is the first index j such that (π, j) is $(S, \mathbf{s})$-good, and $(S, \mathbf{s})$-*late-good* states which we denote by $\mathcal{L}(S, \mathbf{s}) = M \setminus (B(S, \mathbf{s}) \cup \mathcal{F}(S, \mathbf{s}))$.

Lemma 3.1. For a rooted k-uniform hypergraph $(S, \mathbf{s})$ and distinct $i, h \in [0, k-1]$, define the mapping

$$\psi_{\vec{ih}} \colon \mathcal{L}(S, \mathbf{s}) \longrightarrow M$$
$$(x^{(i)} R X x_i Y, k + |R| + |X|) \longmapsto (x^{(h)} X x_h R Y, k + |X|)$$

where R ends in the first-good mark for $(S, \mathbf{s})$,[1] and the index $k + |R| + |X|$ points to x_i. Then, $\psi_{\vec{ih}}$ is injective.

It is worth emphasizing again that, since the indices i, h are fixed, any $(\pi, j) \in \mathcal{L}(S, \mathbf{s})$ can be written in the manner as given in the definition of $\psi_{\vec{ih}}$. In particular, the image is indeed an element in $S_n \times [k, n]$.

Intuitively, the map in Lemma 3.1 simply "sets aside" the first block R supporting an embedding of $(S, \mathbf{s})$ with $x^{(i)}$ as root and puts it behind the last vertex of the edge x. We will apply this injection in the setting where $(S, \mathbf{s})$ is one of the tight subtrees given by Lemma 2.4. By applying this injection whenever applicable, we iteratively "set aside" the blocks supporting those subtrees which in the end results in an embedding of T. However, since these subtrees have distinct ordered $(k-1)$-subsets of $\mathbf{r}$ as roots, the vertices in x are permuted by $\psi_{\vec{ih}}$ to accommodate that.

Proof. Note that R is nonempty and well-defined since the input state is in $\mathcal{L}(S, \mathbf{s})$. Furthermore, as $x^{(i)} \cup \{x_i\} = x^{(h)} \cup \{x_h\} \in E(H)$ and $k + |X|$ points to x_h in the output state, the map is well-defined.

To see that $\psi_{\vec{ih}}$ is injective, note that given the output state we can first reconstruct the ordered tuple x. (Since i, h are fixed, we know that the vertex our index points to must be the hth vertex in x.) Using the index $d + |X|$, we can then determine RY. From there, we can recover R by determining the smallest prefix P of RY such that $x^{(i)} P$ is a subsequence supporting an $x^{(i)}$-rooted copy for $(S, \mathbf{s})$ and ending in a mark. Finally, we can recover the input state using the output state, x and R. □

We are now ready to prove Lemma 2.5.

[1] In other words, $x^{(i)} R$ is the prefix corresponding to the state in $\mathcal{F}(S, \mathbf{s})$.

Proof of Lemma 2.5. Let $\ell = |I|$. Recall that after applying Lemma 2.4 to $(T, \mathbf{r})$, we get an edge-decomposition of T into an edge $e_0 \supseteq \mathbf{r}$, and nonempty tight k-trees $(S_i)_{i \in I}$, $I = \{i_1 < \cdots < i_\ell\}$, that are vertex-disjoint outside of $\mathbf{r}$. We root each S_i at $\mathbf{r}^{(i)}$.

Consider now the following multi-round process that "essentially" filters marked states for those that are $(T, \mathbf{r})$-good: Let $M_1 = M$. For $j = 1, \ldots, \ell$, discard from M_j all the states that are not $(S_{i_j}, \mathbf{r}^{(i_j)})$-late-good. Additionally, if $j < \ell$, given the set of remaining states M_j', define $M_{j+1} = \psi_{i_j \vec{i_{j+1}}}(M_j')$ where $\psi_{i_j \vec{i_{j+1}}} : \mathcal{L}(S_{i_j}, \mathbf{r}^{(i_j)}) \to M$ is the injection obtained from Lemma 3.1.

We will count $|M_\ell{}'|$ in two ways: First, note that by definition, every state discarded in the jth round is in $B(S_{i_j}, \mathbf{r}^{(i_j)}) \cup \mathcal{F}(S_{i_j}, \mathbf{r}^{(i_j)})$. Hence, we discard at most $b(S_{i_j}, \mathbf{r}^{(i_j)}) + \left|\mathcal{F}(S_{i_j}, \mathbf{r}^{(i_j)})\right|$ states. Note that $\left|\mathcal{F}(S_{i_j}, \mathbf{r}^{(i_j)})\right| \leqslant |W|$ since each eligible permutation can be paired with at most one index to form a $(S_{i_j}, \mathbf{r}^{(i_j)})$-first-good state. This gives

$$|M_\ell'| \geqslant |M| - \sum_{i \in I} (b(S_i, \mathbf{r}^{(i)}) + |W|) = |M| - \sum_{i \in I} b(S_i, \mathbf{r}^{(i)}) - \ell |W|.$$

On the other hand, we claim that the preimages that get mapped by the multi-round process to states in M_ℓ' "almost" form good states for $(T, \mathbf{r})$. First, as the composition of injective functions, note that the multi-round process indeed defines an injection if restricted to preimages of M_ℓ'. Denote by N the set of preimages and note that $|N| = |M_\ell'|$.

Let $(x^{(i_1)} A_1 x_{i_1} Y_0, k + |A_1|) \in N$. After the first round, we get $(x^{(i_2)} A_2 x_{i_2} R_1 Y_0, k + |A_2|)$ where R_1 ends in the initial input at the $(S_{i_1}, \mathbf{r}^{(i_1)})$-first-good mark. Iteratively, it becomes clear that after all the rounds we get

$$(x^{(i_\ell)} A_\ell x_{i_\ell} R_{\ell-1} \ldots R_2 R_1 Y_0, k + |A_\ell|)$$

where $x^{(i_j)} R_j$ ends in the $(S_{i_j}, \mathbf{r}^{(i_j)})$-first-good mark of the permutation at round j. Since this final state was not discarded, we know that $A_\ell = R_\ell A_{\ell+1}$ where R_ℓ ends in the $(S_{i_\ell}, \mathbf{r}^{(i_\ell)})$-first-good mark of the final permutation. Undoing all the switches from the rounds, we see that

$$A_1 = R_1 R_2 \ldots R_\ell A_{\ell+1}.$$

Since these R_j-blocks are disjoint and support an $x^{(i_j)}$-rooted copy of $(S_{i_j}, \mathbf{r}^{(i_j)})$, it becomes clear that applying the injection $\varphi_{i_1 \vec{0}}$ to states in N maps to good states for $(T, \mathbf{r})$ since $\mathbf{r} = \mathbf{r}^{(0)}$. Indeed, take the union of the S_i-embeddings and (possibly) extend via the map $r_i \mapsto x_i$. This is not only a well-defined embedding of $T - e_0$ since all these maps are compatible, but also note that e_0 gets mapped to $x \in E(H)$, thus forming a T-embedding that by our application of $\varphi_{i_1 \vec{0}}$ is $x^{(0)}$-rooted. In particular,

$$|M_\ell'| = |N| \leqslant |M| - b(T, \mathbf{r}).$$

Rearranging our bounds on $|M_\ell'|$, we get the desired inequality. □

Remark 3.2. A careful reading of this argument shows that Lemma 2.5 remains true even if we do not require each S_i to be a tight k-tree; as long as the equalities in the conclusion of Lemma 2.4 hold, Lemma 2.5 is valid.

Acknowledgment

I want to thank my advisor Cosmin Pohoata for sending me the GPT-generated proof / "cursing" me with this knowledge, and Sean Longbrake for his valuable feedback to improve the presentation.

References

[1] Tom Adamczewski and Thomas F. Bloom. *FrontierMath Erdős*. Tech. rep. Epoch AI, 2026. URL: https://epoch.ai/files/frontiermath-erdos.pdf (cit. on pp. 1, 2).

[2] Louis DeBiasio. *Tree embeddings in graphs, hypergraphs, and digraphs*. URL: https://github.com/louisdebiasio/erdos-sos (cit. on p. 2).

[3] Péter Frankl and Zoltán Füredi. "Exact solution of some Turán-type problems". In: *J. Combin. Theory Ser. A* 45.2 (1987), pp. 226–262. ISSN: 0097-3165,1096-0899. DOI: 10.1016/0097-3165(87)90016-1. URL: https://doi.org/10.1016/0097-3165(87)90016-1 (cit. on pp. 1, 2).

[4] Maya Stein. "Kalai's conjecture in r-partite r-graphs". In: *European J. Combin.* 117 (2024), Paper No. 103827, 7. ISSN: 0195-6698,1095-9971. DOI: 10.1016/j.ejc.2023.103827. URL: https://doi.org/10.1016/j.ejc.2023.103827 (cit. on p. 2).