Turán's Graph Theorem

Nils Lauermann
Computer Laboratory, University of Cambridge CB3 0FD
Nils.Lauermann@cl.cam.ac.uk

February 6, 2026

Abstract

Turán's Graph Theorem [2] states that any undirected, simple graph with n vertices that does not contain a p-clique, contains at most $\left(1 - \frac{1}{p-1}\right) \frac{n^2}{2}$ edges. The theorem is an important result in graph theory and the foundation of the field of extremal graph theory.

The formalisation follows Aigner and Ziegler's [1] presentation of Turán's initial proof [2]. Besides a direct adaptation of the textbook proof, a simplified, second proof is presented which decreases the size of the formalised proof significantly.

Contents

Acknowledgements

This formalisation was developed as part of the author's MPhil dissertation (Submitted on 6 June 2022) of the Advanced Computer Science course at the University of Cambridge. The project was supervised by Prof Lawrence C. Paulson and Dr Angeliki Koutsoukou-Argyraki.

References

[1] M. Aigner and G. M. Ziegler. *Turán's graph theorem*, pages 285–289. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018.

[2] P. Turán. On an external problem in graph theory. *Mat. Fiz. Lapok*, 48:436–452, 1941.

theory *Turan*
  imports
    Girth-Chromatic.Ugraphs
    Random-Graph-Subgraph-Threshold.Ugraph-Lemmas
begin

1 Basic facts on graphs

lemma *wellformed-uverts-0* :
  assumes *uwellformed G* **and** *uverts G* = {}
  shows *card* (*uedges G*) = *0* ⟨*proof*⟩

lemma *finite-verts-edges* :
  assumes *uwellformed G* **and** *finite* (*uverts G*)
  shows *finite* (*uedges G*)
⟨*proof*⟩

lemma *ugraph-max-edges* :
  assumes *uwellformed G* **and** *card* (*uverts G*) = *n* **and** *finite* (*uverts G*)
  shows *card* (*uedges G*) ≤ *n* ∗ (*n*−*1*)/*2*
  ⟨*proof*⟩

lemma *subgraph-verts-finite* : ⟦ *finite* (*uverts G*); *subgraph G′ G* ⟧ ⟹ *finite* (*uverts G′*)
  ⟨*proof*⟩

2 Cliques

In this section a straightforward definition of cliques for simple, undirected graphs is introduced. Besides fundamental facts about cliques, also more specialized lemmata are proved in subsequent subsections.

definition *uclique* :: *ugraph* ⇒ *ugraph* ⇒ *nat* ⇒ *bool* **where**
  uclique C G p ≡ *p* = *card* (*uverts C*) ∧ *subgraph C G* ∧ *C* = *complete* (*uverts C*)

lemma *clique-any-edge* :
  assumes *uclique C G p* **and** *x* ∈ *uverts C* **and** *y* ∈ *uverts C* **and** *x* ≠ *y*
  shows {*x*,*y*} ∈ *uedges G*
  ⟨*proof*⟩

lemma *clique-exists* : ∃ *C p*. *uclique C G p* ∧ *p* ≤ *card* (*uverts G*)
  ⟨*proof*⟩

lemma *clique-exists1* :
  assumes *uverts G* ≠ {} **and** *finite* (*uverts G*)
  shows ∃ *C p*. *uclique C G p* ∧ *0* < *p* ∧ *p* ≤ *card* (*uverts G*)
⟨*proof*⟩

lemma *clique-max-size* : *uclique C G p* $\Longrightarrow$ *finite* (*uverts G*) $\Longrightarrow$ $p \leq$ *card* (*uverts G*)
⟨*proof*⟩

lemma *clique-exists-gt0* :
assumes *finite* (*uverts G*) *card* (*uverts G*) > *0*
shows ∃ *C p*. *uclique C G p* ∧ *p* ≤ *card* (*uverts G*) ∧ (∀ *C q*. *uclique C G q* ⟶ *q* ≤ *p*)
⟨*proof*⟩

If there exists a $(p+1)$-clique C in a graph G then we can obtain a p-clique in G by removing an arbitrary vertex from C

lemma *clique-size-jumpfree* :
assumes *finite* (*uverts G*) **and** *uwellformed G*
and *uclique C G* (*p+1*)
shows ∃ *C′*. *uclique C′ G p*
⟨*proof*⟩

The next lemma generalises the lemma *clique-size-jumpfree* to a proof of the existence of a clique of any size smaller than the size of the original clique.

lemma *clique-size-decr* :
assumes *finite* (*uverts G*) **and** *uwellformed G*
and *uclique C G p*
shows *q* ≤ *p* ⟹ ∃ *C*. *uclique C G q* ⟨*proof*⟩

With this lemma we can easily derive by contradiction that if there is no p-clique then there cannot exist a clique of a size greater than p

corollary *clique-size-neg-max* :
assumes *finite* (*uverts G*) **and** *uwellformed G*
and ¬(∃ *C*. *uclique C G p*)
shows ∀ *C q*. *uclique C G q* ⟶ *q* < *p*
⟨*proof*⟩

corollary *clique-complete* :
assumes *finite V* **and** *x* ≤ *card V*
shows ∃ *C*. *uclique C* (*complete V*) *x*
⟨*proof*⟩

lemma *subgraph-clique* :
assumes *uwellformed G subgraph C G C* = *complete* (*uverts C*)
shows {*e* ∈ *uedges G*. *e* ⊆ *uverts C*} = *uedges C*
⟨*proof*⟩

Next, we prove that in a graph G with a p-clique C and some vertex v outside of this clique, there exists a $(p+1)$-clique in G if v is connected to all nodes in C. The next lemma is an abstracted version that does not explicitly mention cliques: If a vertex n has as many edges to a set of nodes N as there are nodes in N then n is connected to all vertices in N.

lemma *card-edges-nodes-all-edges* :
fixes G :: *ugraph* **and** N :: *nat set* **and** E :: *nat set set* **and** n :: *nat*
assumes *uwellformed* G
and *finite* N
and $N \subseteq$ *uverts* G **and** $E \subseteq$ *uedges* G
and $n \in$ *uverts* G **and** $n \notin N$
and $\forall e \in E. \exists x \in N. \{n,x\} = e$
and *card* E = *card* N
shows $\forall x \in N. \{n,x\} \in E$
⟨*proof*⟩

2.1 Partitioning edges along a clique

Turán's proof partitions the edges of a graph into three partitions for a $(p-1)$-clique C: All edges within C, all edges outside of C, and all edges between a vertex in C and a vertex not in C.

We prove a generalized lemma that partitions the edges along some arbitrary set of vertices which does not necessarily need to induce a clique. Furthermore, in Turán's graph theorem we only argue about the cardinality of the partitions so that we restrict this proof to showing that the sum of the cardinalities of the partitions is equal to number of all edges.

lemma *graph-partition-edges-card* :
assumes *finite* (*uverts* G) **and** *uwellformed* G **and** $A \subseteq$ (*uverts* G)
shows *card* (*uedges* G) = *card* $\{e \in$ *uedges* $G.\ e \subseteq A\}$ + *card* $\{e \in$ *uedges* $G.\ e \subseteq$ *uverts* $G - A\}$ + *card* $\{e \in$ *uedges* $G.\ e \cap A \neq \{\} \wedge e \cap ($*uverts* $G - A) \neq \{\}\}$
⟨*proof*⟩

Now, we turn to the problem of calculating the cardinalities of these partitions when they are induced by the biggest clique in the graph.

First, we consider the number of edges in a p-clique.

lemma *clique-edges-inside* :
assumes *G1*: *uwellformed* G **and** *G2*: *finite* (*uverts* G)
and p: $p \leq$ *card* (*uverts* G) **and** n: n = *card*(*uverts* G)
and C: *uclique* C G p
shows *card* $\{e \in$ *uedges* $G.\ e \subseteq$ *uverts* $C\} = p * (p-1) / 2$
⟨*proof*⟩

Next, we turn to the number of edges that connect a node inside of the biggest clique with a node outside of said clique. For that we start by calculating a bound for the number of edges from one single node outside of the clique into the clique.

lemma *clique-edges-inside-to-node-outside* :
assumes *uwellformed* G **and** *finite* (*uverts* G)
assumes $0 < p$ **and** $p \leq$ *card* (*uverts* G)
assumes *uclique* C G p **and** $(\forall C\ p'.$ *uclique* C G $p' \longrightarrow p' \leq p)$
assumes y: $y \in$ *uverts* G − *uverts* C

shows *card* $\{\{x,y\}|\ x.\ x \in uverts\ C \wedge \{x,y\} \in uedges\ G\} \leq p - 1$
⟨*proof*⟩

Now, that we have this upper bound for the number of edges from a single vertex into the largest clique we can calculate the upper bound for all such vertices and edges:

lemma *clique-edges-inside-to-outside* :
assumes *G1*: *uwellformed G* **and** *G2*: *finite* (*uverts G*)
and *p0*: $0 < p$ **and** *pn*: $p \leq card\ (uverts\ G)$ **and** $card(uverts\ G) = n$
and *C*: *uclique C G p* **and** *C-max*: $(\forall\ C\ p'.\ uclique\ C\ G\ p' \longrightarrow p' \leq p)$
shows *card* $\{e \in uedges\ G.\ e \cap uverts\ C \neq \{\} \wedge e \cap (uverts\ G - uverts\ C) \neq \{\}\} \leq (p - 1) * (n - p)$
⟨*proof*⟩

Lastly, we need to argue about the number of edges which are located entirely outside of the greatest clique. Note that this is in the inductive step case in the overarching proof of Turán's graph theorem. That is why we have access to the inductive hypothesis as an assumption in the following lemma:

lemma *clique-edges-outside* :
assumes *uwellformed G* **and** *finite* (*uverts G*)
and *p2*: $2 \leq p$ **and** *pn*: $p \leq card\ (uverts\ G)$ **and** *n*: $n = card(uverts\ G)$
and *C*: *uclique C G* $(p-1)$ **and** *C-max*: $(\forall\ C\ q.\ uclique\ C\ G\ q \longrightarrow q \leq p-1)$
and *IH*: $\bigwedge G\ y.\ y < n \Longrightarrow finite\ (uverts\ G) \Longrightarrow uwellformed\ G \Longrightarrow \forall\ C\ p'.\ uclique\ C\ G\ p' \longrightarrow p' < p$
$\Longrightarrow 2 \leq p \Longrightarrow card\ (uverts\ G) = y \Longrightarrow real\ (card\ (uedges\ G)) \leq (1 - 1\ /\ real\ (p - 1)) * real\ (y^2)\ /\ 2$
shows *card* $\{e \in uedges\ G.\ e \subseteq uverts\ G - uverts\ C\} \leq (1 - 1\ /\ (p-1)) * (n - p + 1)\ \hat{}\ 2\ /\ 2$
⟨*proof*⟩

2.2 Extending the size of the biggest clique

In this section, we want to prove that we can add edges to a graph so that we augment the biggest clique to some greater clique with a specific number of vertices. For that, we need the following lemma: When too many edges have been added to a graph so that there exists a $(p+1)$-clique then we can remove at least one of the added edges while also retaining a p-clique

lemma *clique-union-size-decr* :
assumes *finite* (*uverts G*) **and** *uwellformed* (*uverts G*, *uedges G* $\cup$ *E*)
and *uclique C* (*uverts G*, *uedges G* $\cup$ *E*) $(p+1)$
and *card E* ≥ 1
shows $\exists\ C'\ E'.\ card\ E' < card\ E \wedge uclique\ C'\ (uverts\ G,\ uedges\ G \cup E')\ p \wedge uwellformed\ (uverts\ G,\ uedges\ G \cup E')$
⟨*proof*⟩

We use this preceding lemma to prove the next result. In this lemma we assume that we have added too many edges. The goal is then to remove some of the new edges appropriately so that it is indeed guaranteed that there is no bigger clique.

Two proofs of this lemma will be described in the following. Both fundamentally come down to the same core idea: In essence, both proofs apply the well-ordering principle. In the first proof we do so immediately by obtaining the minimum of a set:

lemma *clique-union-make-greatest* :
fixes $p\ n :: nat$
assumes $finite\ (uverts\ G)$ **and** $uwellformed\ G$
and $uwellformed\ (uverts\ G,\ uedges\ G \cup E)$ **and** $card(uverts\ G) \geq p$
and $uclique\ C\ (uverts\ G,\ uedges\ G \cup E)\ p$
and $\forall\, C'\ q'.\ uclique\ C'\ G\ q' \longrightarrow q' < p$ **and** $1 \leq card\ E$
shows $\exists\, C'\ E'.\ uwellformed\ (uverts\ G,\ uedges\ G \cup E')$
$\wedge\ (uclique\ C'\ (uverts\ G,\ uedges\ G \cup E')\ p)$
$\wedge\ (\forall\, C''\ q'.\ uclique\ C''\ (uverts\ G,\ uedges\ G \cup E')\ q' \longrightarrow q' \leq p)$
⟨*proof*⟩

In this second, alternative proof the well-ordering principle is used through complete induction.

lemma *clique-union-make-greatest-alt* :
fixes $p\ n :: nat$
assumes $finite\ (uverts\ G)$ **and** $uwellformed\ G$
and $uwellformed\ (uverts\ G,\ uedges\ G \cup E)$ **and** $card(uverts\ G) \geq p$
and $uclique\ C\ (uverts\ G,\ uedges\ G \cup E)\ p$
and $\forall\, C'\ q'.\ uclique\ C'\ G\ q' \longrightarrow q' < p$ **and** $1 \leq card\ E$
shows $\exists\, C'\ E'.\ uwellformed\ (uverts\ G,\ uedges\ G \cup E')$
$\wedge\ (uclique\ C'\ (uverts\ G,\ uedges\ G \cup E')\ p)$
$\wedge\ (\forall\, C''\ q'.\ uclique\ C''\ (uverts\ G,\ uedges\ G \cup E')\ q' \longrightarrow q' \leq p)$
⟨*proof*⟩

Finally, with this lemma we can turn to this section's main challenge of increasing the greatest clique size of a graph by adding edges.

lemma *clique-add-edges-max* :
fixes $p :: nat$
assumes $finite\ (uverts\ G)$
and $uwellformed\ G$ **and** $card(uverts\ G) > p$
and $\exists\, C.\ uclique\ C\ G\ p$ **and** $(\forall\, C\ q'.\ uclique\ C\ G\ q' \longrightarrow q' \leq p)$
and $q \leq card(uverts\ G)$ **and** $p \leq q$
shows $\exists\, E.\ uwellformed\ (uverts\ G,\ uedges\ G \cup E) \wedge (\exists\, C.\ uclique\ C\ (uverts\ G,\ uedges\ G \cup E)\ q)$
$\wedge\ (\forall\, C\ q'.\ uclique\ C\ (uverts\ G,\ uedges\ G \cup E)\ q' \longrightarrow q' \leq q)$
⟨*proof*⟩

3 Properties of the upper edge bound

In this section we prove results about the upper edge bound in Turán's theorem. The first lemma proves that upper bounds of the sizes of the partitions sum up exactly to the overall upper bound.

lemma *turan-sum-eq* :
fixes $n\ p :: nat$
assumes $p \geq 2$ **and** $p \leq n$
shows $(p-1) * (p-2) / 2 + (1 - 1 / (p-1)) * (n - p + 1)\ \hat{}\ 2 / 2 + (p - 2) * (n - p + 1) = (1 - 1 / (p-1)) * n\hat{}2 / 2$
⟨*proof*⟩

The next fact proves that the upper bound of edges is monotonically increasing with the size of the biggest clique.

lemma *turan-mono* :
fixes $n\ p\ q :: nat$
assumes $0 < q$ **and** $q < p$ **and** $p \leq n$
shows $(1 - 1 / q) * n\hat{}2 / 2 \leq (1 - 1 / (p-1)) * n\hat{}2 / 2$
⟨*proof*⟩

4 Turán's Graph Theorem

In this section we turn to the direct adaptation of Turán's original proof as presented by Aigner and Ziegler [1]

theorem *turan* :
fixes $p\ n :: nat$
assumes *finite* (*uverts G*)
and *uwellformed G* **and** $\forall C\ p'.\ uclique\ C\ G\ p' \longrightarrow p' < p$ **and** $p \geq 2$ **and** $card(uverts\ G) = n$
shows $card\ (uedges\ G) \leq (1 - 1 / (p-1)) * n\hat{}2 / 2$ ⟨*proof*⟩

5 A simplified proof of Turán's Graph Theorem

In this section we discuss a simplified proof of Turán's Graph Theorem which uses an idea put forward by the author: Instead of increasing the size of the biggest clique it is also possible to use the fact that the expression in Turán's graph theorem is monotonically increasing in the size of the biggest clique (Lemma *turan-mono*). Hence, it suffices to prove the upper bound for the actual biggest clique size in the graph. Afterwards, the monotonicity provides the desired inequality.

The simplifications in the proof are annotated accordingly.

theorem *turan′* :
fixes $p\ n :: nat$
assumes *finite* (*uverts G*)
and *uwellformed G* **and** $\forall C\ p'.\ uclique\ C\ G\ p' \longrightarrow p' < p$ **and** $p \geq 2$ **and** $card(uverts\ G) = n$
shows $card\ (uedges\ G) \leq (1 - 1 / (p-1)) * n\hat{}2 / 2$ ⟨*proof*⟩

end