Erdős-Gallai Conjecture

Juncheng Wan, jorgenwan@gamil.com

This is a note of graph decomposition of cycles and edges.

Pre-requisites of this note include a knowledge of the basic concepts of linear algebra, probability theory.

Contents

1 Introduction

In 1966, Erdős and Gallai made the following conjecture [4]:

Conjecture 1.1. *Any n-vertex graph can be decomposed into $O(n)$ cycles and edges.*

Following Theorem 1.1 in [5], Erdős and Gallai greedily removed cycles of longest length $O(n)$ and got the bound $O(n \log n)$, following from a simple iteration.

Theorem 1.1. *Every graph with n nodes and more than $(n-1)l/2$ edges ($l \geq 2$) contains a circle with more than l edges.*

In 2014, Fox, Conlon and Sudakov [3] made the first major breakthrough on this problem, showing that such a decomposition for general graph with only $O(n \log \log d)$ cycles and edges always exists, as in Theorem 1.2.

Theorem 1.2. *Every graph on n vertices with average degree d can be decomposed into $O(n \log \log d)$ cycles and edges.*

They also proved that the conjecture holds asymptotically almost surely, or a.a.s. for short, for random graph and for graphs of linear minimum degree.

Theorem 1.3. *There exists a constant $c > 0$ such that for any probability $p \triangleq p(n)$ the random graph $G(n,p)$ a.a.s. can be decomposed into at most cn cycles and edges.*

Theorem 1.4. *Every graph G on n vertices with minimum degree cn can be decomposed into at most $O(c^{-12}n)$ cycles and edges.*

In 2022, Bucić and Montgomery improved the previous bound of general graph to $O(n \log^* d)$ [2], as in Theorem 1.5. The iterated logarithm function $\log^* n$ is the minimum number of times we need to apply the logarithm function to n until it becomes at most one.

Theorem 1.5. *Every graph on n vertices with average degree d can be decomposed into $O(n \log^* d)$ cycles and edges.*

2 General Graphs with $O(n \log n)$ Decompositions

2.1 Proof

Following Theorem 1.1 in [5], Erdős and Gallai greedily removed cycles of longest length $O(n)$ times, that the graph that remains will be acyclic or have at most half the edges, and got the bound $O(n \log n)$, following from a simple iteration.

Proof. Recall that we need to prove that every graph G on n vertices can be decomposed into $O(n \log n)$ cycles and edges.

Delete the longest cycle of G at each iteration until there is no cycle in G. Let m_i $(i \geq 0)$ be the remaining edges of G at the i-th iteration. Initially, $m_0 = E(G)$. Let l_i $(i \geq 0)$ be the length of the longest cycle at the i-th iteration. Thus, $m_{i+1} = m_i - l_i$. By Theorem 1.1, we have $m_{i+1} \leq (n-1)l_i/2$.

Consider the process, where the length l_i decreases strictly to $l_j <= l_i/2$. The number of deleted cycles is at most $\frac{m_i}{l_i/2} \leq n-1$. This process can be repeated for $\log l_0 \leq \log n$ times at most. Thus, the number of deleted cycles is at most $n \log n$. The remaining graph is disjoint trees and thus have at most $n-1$ edges. In total, G is decomposed with $O(n \log n)$ cycles and $O(n)$ edges. □

2.2 Proof of Theorem 1.1 (to be done)

3 General Graphs with $O(n \log \log n)$ Decompositions

3.1 Proof sketch

Assume that the graph G has average degree d.

1. Delete the longest cycles in G until the cycles remained having length below d/c_1, getting G'. In this stage, it deletes at most $\frac{dn/2}{d/c_1} = \frac{c_1}{2}n$ cycles.
2. Partition edges of G' and get $G'_1, \dots, G'_k$, where $|V(G'_i)| \leq d/c_1 + 2$, $\sum_{i=1}^{s} |V(G'_i)| \leq 3n$.
3. Delete around $4|V(G'_i)|^{2-1/9} \log |V(G'_i)|$ edges in each G'_i. It gets disjoint expanders $G'_{i1}, \dots, G'_{ir_i}$, where each component has expanding property.
   (a) The expansion property is that for all $X \subset G'_{ij}$ with $|X| \leq G'_{ij}/2$, $e(X, X^c) \geq s|X|$, where $s = 3|V(G'_i)|^{8/9}$.
   (b) For each expander G'_{ij}, it has some robustness property. Specifically, there is a set $U_{ij} \subset V(G'_{ij})$, such that for each $x, y \in V(G'_{ij}) \backslash U_{ij}$, avoiding at most $2\sqrt{|V(G'_{ij}|}$ vertices and $\frac{1}{2}|V(G'_{ij}|^{4/3}$ edges of U_{ij}, x, y is still connected through U_{ij} with length at most $|V(G'_{ij}|^{2/9}$. Besides, U_{ij} is not too large, with $|U_{ij}| \leq 3|V(G'_{ij})|^{8/9}$.
   (c) Using path decomposition theorem, decompose $G'_{ij} \backslash U_{ij}$ into $|V(G'_{ij})|/2$ paths and cycles.
   (d) Using pigeonhole principle, delete at most $2 \cdot |V(G'_{ij})|/2$ edges from those paths and the remained paths have vertices sharing at most $\sqrt{2 \cdot |V(G'_{ij})|/2}$ endpoints.
   (e) With robustness property, connect those paths using U and get $(1/2 + 1/2)|V(G'_{ij})|$ cycles with at most $|V(G'_{ij})| + 3|V(G'_{ij})|^{8/9} \cdot |V(G'_{ij})|$ edges left. Then, sum over i and j to get the total cycles and remained edges.
4. In this way, the iteration deletes $3n$ cycles and remains at most $16n^{2-1/9} \log n \leq O(n^{2-1/10})$ edges. Thus, the average degree decreases from d to $d^{0.9}$.

3.2 Longest path and its cycle

The longest path P in a graph G has some structures, as the neighborhoods of endvertices in P can only fall into its internal vertices, which is shown in Lemma 2.6 and Lemma 2.7 in [1] as the following Lemma 3.1 and Lemma 3.2.

Lemma 3.1. *Let G be a graph, let $P = u \dots v$ be a longest path in G, and put $S := S(P)$. Then $N_G(S) \subseteq S^- \cup S^+$.*

Proof. Let $x \in S$ and $y \in N(x)$ be given; we show that $y \in S^- \cup S \cup S^+$. As $x \in S$ there is a path $Q = x \dots v$ derived from P. Then $y \in V(Q) = V(P)$, because Q (like P) is a longest path; let z denote the predecessor of y on Q.

Suppose that $y \notin S^- \cup S \cup S^+$. Then each of the (one or two) edges $e \in P$ at y lies on every path derived from P (and in particular on Q), because in any elementary exchange in which e is first deleted, its two ends (including y) would have become members of S and of $S^- \cup S^+$, respectively. Hence $z \in \{y^-, y^+\}$. But $Q + xy - yz$ is obtained from Q by an elementary exchange, which puts z in S and y in $S^- \cup S^+$. $\square$

Lemma 3.2. *Let G be a graph, let $P = u \dots v$ be a longest path in G, and put $S := S(P)$. Then G has a cycle containing $S \cup N(S)$.*

Proof. Let y be the last vertex of P in $N(S)$. Then all the vertices from $S \cup N(S)$ lie on Py, because any vertex of yP in S would differ from v and hence have its successor on P in $N(S)$. Let $x \in S$ be a neighbour of y in G, and let $Q = x \dots v$ be derived from P. As in the proof of Lemma 3.1, all the edges of yP are still edges of Q, so $yQ = yP$. Thus $S \cup N(S) \subseteq V(Py) = V(Qy)$, and Qyx is a cycle in G. □

Using Lemma 3.1 and Lemma 3.2, we have the following Pósa-type statement.

Lemma 3.3. *If a graph G contains no cycle of length greater than $3t$, then there is a subset S of size at most t such that $|N_G(S)| \leq 2|S|$.*

Proof. Otherwise, for all subset S of size at most t, we have $|N_G(S)| > 2|S|$. Consider the longest path P in G. Consider the subset $S \subset V(P)$ defined in Lemma 3.1, which has $N_G(S) \subset S^- \cup S^+$. Assume $|S| \leq t$. As both $|S^-|, |S^+| \leq |S|$, we have $|N_G(S)| \leq |S^- \cup S^+| \leq |S^-| + |S^-| \leq 2|S|$, which is a contradiction.

Thus, $|S| > t$. We can pick a subset $T \subset S$, such that $|T| = t$. Thus, by the assumption, we have $|N_G(S)| > |N_G(T)| > 2t$. By Lemma 3.2, G has a cycle C containing $S \cup N(S)$. Thus $|C| \geq |S \cup N(S)| = |S| + |N(S)| > 3t$. □

3.3 Path decomposition with low endpoint coincidence

Lemma 3.4. *Suppose a graph G has a decomposition into h paths. Then G can also be decomposed into h subpaths of these paths and at most $2h$ edges so that each vertex is an endpoint of at most $\sqrt{2h}$ of the paths.*

Proof. Consider the vertex v, which is the endpoint of t paths $\mathcal{P}_v$, where $t > \sqrt{2h}$. Consider those vertices $N_G(v)$, by pigeonhole principle, there is a vertex $u \in N_G(v)$ such that u is the endpoint of at most $(2h - t)/t < \sqrt{2h} - 1$ paths. Delete edge uv in G. Repeat this process until no vertex is the endpoint of more that $\sqrt{2h}$ paths. As each path will be deleted at most two edges, the total deleted number of edges is at most $2h$. □

3.4 Questions

Question 3.1. *Why Lemma 3.2 in [3] use $2n/s$ for each M_i? Why Claim in Lemma 3.2 in [3] use $4n/s$?*

Answer 3.1. *First, for each vertex, it is $s/2n$, then for almost all vertices it is $s/2$. As the coefficient of expansion property is s, expanding with $s/2$ is highly reachable.*

Second, it is almost the same.

Question 3.2. *Why Lemma 3.2 in [3] use $n^{2/3}$ for vertices in E in the proof?*

Answer 3.2. *The proof could be understood in this way. As $y_0 = y \notin U$ and $2\sqrt{n} \leq 2n^{2/3}$, we could assume that in each step, y_i has at least $4n^{2/3} - 2n^{2/3} = 2n^{2/3}$ neighbors not in $G \backslash W - E$.*

4 General Graphs with $O(n \log^* d)$ Decompositions

4.1 Proof sketch

Assume that the graph G has average degree d.

1. Delete the longest cycles in G until the cycles remained having length below d, getting G'. In this stage, it deletes at most $\frac{dn/2}{d} = \frac{1}{2}n = O(n)$ cycles.
2. Let $\epsilon = 0, s = 2^{-5}$. Delete at most $4sn \log n = 0$ edges from G' and partition the remaining edges to get $G'_1, \dots, G'_k$, where each G'_i is an (ϵ, s)-expander, i.e. $(2^{-5}, 0)$-expander, $|V(G'_i)| = O(d \log^4 d)$, $\sum_{i=1}^{k} |V(G'_i)| \leq 2n$.
3. For each G'_i, let $\epsilon = 2^{-5}, s = \log^{273} |V(G'_i)|$. Delete at most $4s|V(G'_i)| \log |V(G'_i)| = 4|V(G'_i)| \log^{274} |V(G'_i)|$ edges from G'_i and partition the remaining edges to get $G'_{i1}, \dots, G'_{ir_i}$, where each G'_{ij} is an (ϵ, s)-expander, i.e. $(2^{-5}, \log^{273} |V(G'_i)|)$-expander, $|V(G'_{ij})| = O(d \log^4 d)$ and $\sum_{j=1}^{r_i} |V(G'_{ij})| \leq 2|V(G'_i)|$.
   (a) We can only consider those G'_{ij} relative large, with $|V(G'_{ij})| \geq 2^{12}$. Otherwise deleting edges of those expanders with $|V(G'_{ij})| \leq 2^{12}$, instead of utilizing those edges to compose cylces, will only delete at most $\sum_{j=1}^{r_i} 1/2|V(G'_{ij})|^2 \leq 2^{11} \sum_{j=1}^{r_i} |V(G'_{ij})| \leq 2^{12}|V(G'_i)|$ edges.
   (b) For each expander G'_{ij} with $|V(G'_{ij})| \geq 2^{12}$, it has some robustness property, which makes it decomposed into $2|V(G'_{ij})|$ cycles and at most $3 \cdot 2^9 |V(G'_{ij})| \log^{10} |V(G'_{ij})|$ edges.
   (c) Specifically, the robustness property is as follows.
      i. We can decompose G'_{ij} into three edge disjoint $(\frac{\epsilon}{4}, \frac{\sqrt{s\epsilon}}{24 \log |V(G'_i)|})$-expander, i.e. $(2^{-7}, \geq \log^{135} |V(G'_i)|)$-expander, $G'_{ij1}, G'_{ij2}, G'_{ij3}$. Note that $V(G'_{ijl}) = V(G'_{ij})$, for all $l \in \{1, 2, 3\}$.
      ii. Let $V \subset V(G'_{ij})$ be chosen by including each vertex independently at random with probability $\frac{1}{3}$. Then, with high probability, there is a subgraph $G''_{ijl} \subset G'_{ijl}$ with at most $2^9 |V(G'_{ij})| \log^{10} |V(G'_{ij})|$ edges which is $(\log^7 |V(G'_{ij})|, 2)$-path connected through V.
      iii. We can view G''_{ijl} as an edge subset of G''_{ij}, with some relative dense edges in V and some edges in $e(v, V)$ for each $v \in V(G'_{ij}) \backslash V$.
   (d) With the above robustness property, we can find a partition of $V(G'_{ij}) = V_{ij1} \cup V_{ij2} \cup V_{ij3}$, such that there is $G''_{ijl} \subset G'_{ij}$ which is $(\log^7 |V(G'_{ij})|, 2)$-path connected through V_{ijl}. Let $G''_{ij} = G'_{ij} - G''_{ij1} - G''_{ij2} - G''_{ij3}$ and $H_{ijl} = G''_{ij}[V_{ij(l+1)}] + e(V_{ij(l+1)}, V_{ij(l+2)})$, where $+$ is in the sense of appropriately modulo 3.
   (e) Decompose each H_{ijl} with a collection of paths $\mathcal{P}_{ijl}$ with the property that no vertex is an endvertex of more than two of the paths in $\mathcal{P}_{ijl}$, which also means that $|\mathcal{P}_{ijl}| \leq |V(H_{ijl})|$. As V_{ijl} not contained in $V(H_{ijl})$, we can use G''_{ijl} to connect those paths, with at most $|V(H_{ijl})| \leq |V(G'_{ij})| - |V(G'_{ijl})|$ cycles and $2^9 |V(G'_{ij})| \log^{10} |V(G'_{ij})|$ edges left. Thus, in total we are left with $2|V(G'_{ij})|$ cycles and $3 \cdot 2^9 |V(G'_{ij})| \log^{10} |V(G'_{ij})|$ edges.
4. In this way, the iteration deletes at most $0.5n + \sum_i \sum_j 3|V(G'_{ij})| \leq 0.5n + \sum_i 3 \cdot 2|V(G'_i)| \leq 0.5n + 12n = 12.5n$ cycles and remains $\sum_i (4|V(G'_i)| \log^{274} |V(G'_i)| + 2^{12}|V(G'_i)| + \sum_j 2^{11}|V(G'_{ij})| \log^{10} |V(G'_{ij})|) \leq \sum_i 2^{14}|V(G'_i)| \log^{274} |V(G'_i)| = O(\sum_i |V(G'_i)| \log^{274} d) = O(n \log^{274} d)$ edges. Thus, the average degree decreases from d to $\log^{274} d$ with $O(n)$ cycles deleted. Repeat this iteration until no cycles exists and the number of iteration is $O(\log^* d)$. Thus, there are $O(n \log^* d)$ cycles deleted in total and $O(n)$ edges left.

4.2 Questions

Question 4.1. *Why find a partition of* $V(G'_{ij}) = V_{ij1} \cup V_{ij2} \cup V_{ij3}$*?*

Answer 4.1. *Three partitions is a must.*

1. *If we only have one partition of* $V(G'_{ij})$*, then let the subgraph* $G'_{ij1} \subset G'_{ij}$ *path connected through* $V(G'_{ij})$*. It is problematic to complete the decomposed paths set* $\mathcal{P}_{ij}$ *in* $G'_{ij} - G'_{ij1}$ *to cycle by using paths* $\mathcal{P}'_{ij}$ *in* G'_{ij1}*, as both paths* $\mathcal{P}_{ij}$ *and* $\mathcal{P}'_{ij}$ *use the same vertex set* $V(G'_{ij})$*.*

2. *If we only have two partitions of* $V(G'_{ij})$*, then let the subgraphs* $G'_{ij1}, G'_{ij2} \subset G'_{ij}$ *path connected through* V_{ij1}, V_{ij2} *respectively, where* $V(G'_{ij}) = V_{ij1} \cup V_{ij2}$*. For* $G''_{ij} = G'_{ij} - G'_{ij1} - G'_{ij2}$*, it is hard to manage the edges* $e(V_{ij1}, V_{ij2})$*, as those edges have endvertex in* V_{ij1} *and* V_{ij2}*.*

References

[1] Stephan Brandt et al. "Global connectivity and expansion: long cycles and factors in f-connected graphs". In: *Combinatorica* 26 (2006), pp. 17–36.

[2] Matija Bucić and Richard Montgomery. "Towards the Erdős-Gallai cycle decomposition conjecture". In: *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*. 2023, pp. 839–852.

[3] David Conlon, Jacob Fox, and Benny Sudakov. "Cycle packing". In: *Random Structures & Algorithms* 45.4 (2014), pp. 608–626.

[4] Paul Erdös, Adolph W Goodman, and Louis Pósa. "The representation of a graph by set intersections". In: *Canadian Journal of Mathematics* 18 (1966), pp. 106–112.

[5] T Gallai et al. "On maximal paths and circuits of graphs". In: *Acta Math. Acad. Sci. Hungar* 10 (1959), pp. 337–356.