The minimum number of edges in a pancyclic graph

A short proof for Erdős Problem #1016

KNT · Astra · Fable 5.2

Draft of 12 November 2026

A graph on n vertices is pancyclic if it has a cycle of every length from 3 to n. Let $n + h(n)$ be the least number of edges of such a graph. We prove that

$$\log_2 n + \log^* n - O(1) \;\le\; h(n) \;\le\; \lfloor \log_2 n \rfloor + \log^* n + 1.$$

This answers Problem #1016 of Erdős. The lower bound follows from one estimate for random even subgraphs, which the companion paper [3] proves. The upper bound is a construction. Both bounds are formalised in Lean 4. Astra and Fable 5.2 are language models. Section 5 states the part of each author.

The proof in five sentences

1. A pancyclic graph with n vertices and $n + h$ edges has rank $h + 1$, and its cycles have $n - 2$ different lengths. A graph of rank r has fewer than 2^r cycles. This gives only $h \ge \log_2 n - 1$.
2. Take a small subgraph W that already contains cycles of all short lengths. A cycle through an edge of W is determined by its part outside W, up to at most $|E(W)|$ lengths, and this part is a forest.
3. By the *forest estimate*, at most $\frac{1}{2} + o(1)$ of all outside parts are forests. So the largest possible density of cycle lengths decreases by a factor of almost two when the rank grows from R to 2^{256R}.
4. The halving repeats at each scale of a tower of exponentials. This gives the term $\log^* n$ of the lower bound.
5. For the upper bound, k chords with path lengths $1, 2, 4, \dots$ give 2^k consecutive cycle lengths, and $\log^* k$ further chords supply the short lengths.

1 The result

A simple graph on n vertices is *pancyclic* if it contains a cycle of every length $3, 4, \dots, n$. Write $n + h(n)$ for the least number of edges of such a graph. Throughout, $\log$ is the binary logarithm, and $\log^{(j)}$ is its j-fold iterate. The number $\log^* x$ is the least nonnegative integer j with $\log^{(j)} x \le 1$.

Theorem 1.1 (the minimum excess of a pancyclic graph)

There is a constant A such that, for every $n \geq 3$,

$$\log n + \log^* n - A \;\leq\; h(n) \;\leq\; \lfloor \log n \rfloor + \log^* n + 1.$$

	lower bound for $h(n)$	**upper bound for $h(n)$**
Bondy [2], stated without proof	$\log(n-1) - 1$	$\log n + \log^* n + O(1)$
Theorem 1.1	$\log n + \log^* n - A$	$\lfloor \log n \rfloor + \log^* n + 1$

A count of cycles gives the lower bound of Bondy: a connected graph with $n + h$ edges has at most $2^{h+1} - 1$ cycles, and a pancyclic graph needs $n - 2$ of them. Griffin [4] gave a written proof and computed small values. Problem #1016 of Erdős [1] asks if the lower bound can increase to $\log n + \log^* n - O(1)$, and if at least $h(n) - \log n \to \infty$. Theorem 1.1 answers both questions.

The theorem and its proof are from [5]. This paper gives the proof in a short form. It uses one result of the companion paper [3], the forest estimate, which we state in Section 2.

2 The forest estimate

NOTATION | **even subgraphs and outside parts**

- For a graph G, an *even subgraph* is a set of edges that meets every vertex an even number of times. With symmetric difference as addition, the even subgraphs form a vector space $\mathcal{Z}(G)$ over $\mathbb{F}_2$.
- Its dimension is the *rank* $\beta(G) = |E(G)| - |V(G)| + \mathrm{comp}(G)$, where $\mathrm{comp}(G)$ is the number of components. A cycle is a nonzero element of $\mathcal{Z}(G)$, so a graph of rank r has at most $2^r - 1$ cycles.
- Let W be a subgraph of G, with $r = \beta(G)$, $t = \beta(W)$ and $m = |E(W)|$. For an even subgraph ξ we call $\xi \setminus E(W)$ its *outside part*. The map $\xi \mapsto \xi \setminus E(W)$ is linear on $\mathcal{Z}(G)$ with kernel $\mathcal{Z}(W)$. So there are 2^{r-t} outside parts.
- $p(G, W)$ is the proportion of the outside parts that are forests. Each outside part comes from 2^t even subgraphs. So $p(G, W)$ is also the probability that $\xi \setminus E(W)$ is a forest, for a uniformly random $\xi \in \mathcal{Z}(G)$.

Theorem 2.1 (Forest estimate, [3, Corollary 1.2])**.** There is r_0 such that the following holds for all $r \geq r_0$. Let G be a connected simple graph of rank r. Let $W \subseteq G$ be a nonempty union of cycles with at most $r/\log^2 r$ edges and at most $(\log r)/100$ components. Then

$$p(G, W) \;\leq\; \frac{1}{2} + \frac{20}{\log^{(3)} r}.$$

The proof in [3] is a second-moment argument. It reduces the estimate to a multigraph in which a small set of vertices is marked and all other vertices have degree three. In this multigraph it counts the cycles of length about $\log r$ that are components of a random even subgraph. The union W can have several components, and Section 3 needs this, because the cycles of

Lemma 3.1 can be disjoint.

Lemma 2.2 (Lengths). The cycles of G have at most $2^{r-t}\bigl(p\,m+1-p\bigr)$ different lengths, where $p = p(G,W)$.

Proof. Let x be an outside part, and let C be a cycle with the outside part x. If x is not a forest, then $x \subseteq C$ contains a cycle, so $C = x$. Such an x accounts for at most one length. If x is a forest, then C has an edge in W, and the length of C is $|x| + |C \cap E(W)|$ with $1 \le |C \cap E(W)| \le m$. Such an x accounts for at most m lengths. There are $p\,2^{r-t}$ outside parts of the second kind and $(1-p)2^{r-t}$ of the first. ∎

3 The lower bound

IDEA | outline of the proof

A number $\Phi(R) \le 1$ measures how well a graph of rank at least R can fill an interval of cycle lengths. In a graph with many cycle lengths we find a small subgraph W that already contains the short cycles. By the forest estimate for W, the number Φ decreases by a factor of almost two when the rank grows from R to 2^{256R} (Figure 1).

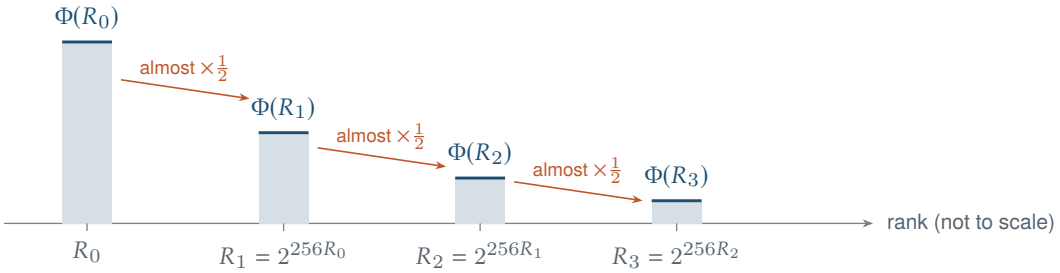


Figure 1. The density Φ of cycle lengths at the scales $R_{j+1} = 2^{256R_j}$. Each step almost halves it (Proposition 3.2). The number of steps below a rank r is $\log^* r - O(1)$.

For an integer $R \ge 1$ let

$$\Phi(R) = \sup \frac{L-2}{2^{\beta(G)}},$$

where the supremum is over all simple graphs G with $\beta(G) \ge R$ and all integers $L \ge 2$ such that G contains a cycle of each length $3, \dots, L$. A graph of rank β has at most $2^\beta - 1$ cycles, so $\Phi(R) \le 1$. The function Φ is non-increasing.

Lemma 3.1 (Witness). Let $R \ge 1$ and $\varphi = \Phi(R)$, and let G contain a cycle of each length $3, \dots, L$, where $L \ge 2^{2R}$. Then G has a subgraph W that is a union of at most $2R$ cycles, with $\beta(W) \ge 2R$ and $|E(W)| \le \varphi\, 2^{2R} + 2^R + 5R$.

Proof. Put $W_0 = \emptyset$ and $W_{i+1} = W_i \cup C_i$, where C_i is a cycle of G whose length ℓ_i is the least integer ≥ 3 that is not the length of a cycle of W_i. Stop as soon as the rank is at least $2R$. Consider a step, and let $s = \beta(W_i) < 2R$. The graph W_i contains cycles of the $\ell_i - 3$ lengths $3, \dots, \ell_i - 1$. Therefore $\ell_i - 3 \leq 2^s - 1$. This gives $\ell_i \leq 2^{2R-1} + 2 \leq L$, so C_i exists. If $s \geq R$, then also $\ell_i - 3 \leq \varphi\, 2^s$, by the definition of Φ. The cycle C_i is not a subgraph of W_i, so the rank increases at every step. So there are at most $2R$ steps, their ranks s are distinct, and

$$|E(W)| \;\leq\; \sum_i \ell_i \;\leq\; \sum_{s=0}^{R-1} \bigl(2^s + 2\bigr) + \sum_{s=R}^{2R-1} \bigl(\varphi\, 2^s + 3\bigr) \;\leq\; \varphi\, 2^{2R} + 2^R + 5R. \qquad \blacksquare$$

Proposition 3.2 (Near-halving). For every sufficiently large integer R,

$$\Phi\bigl(2^{256R}\bigr) \;\leq\; \Bigl(\frac{1}{2} + \frac{20}{\log^{(2)} R}\Bigr)\Phi(R) + 2^{2-R}.$$

Proof. Let G be a simple graph of rank $r \geq 2^{256R}$ with a cycle of each length $3, \dots, L$. We bound $(L-2)\,2^{-r}$. Identifying a vertex of one component with a vertex of another changes neither the rank nor the set of cycles, so we may assume that G is connected. If $L < 2^{2R}$, then $(L-2)\,2^{-r} < 2^{2-R}$, and we are done.

Otherwise let W be as in Lemma 3.1, with m edges and rank $t \geq 2R$, and let $\varphi = \Phi(R)$. Then $m \leq 2^{2R+1} \leq \sqrt{r} \leq r/\log^2 r$ and $\mathrm{comp}(W) \leq 2R \leq (\log r)/100$. So Theorem 2.1 gives

$$p(G,W) \;\leq\; \frac{1}{2} + \frac{20}{\log^{(3)} r} \;\leq\; \frac{1}{2} + \delta, \qquad \text{where } \delta = \frac{20}{\log^{(2)} R}.$$

The second inequality holds because $\log^{(3)} r \geq \log(8 + \log R) \geq \log^{(2)} R$. The graph G has cycles of at least $L - 2$ lengths. So Lemma 2.2, with $p = p(G,W)$, and the bounds on t and m give

$$\frac{L-2}{2^r} \;\leq\; \frac{pm+1}{2^t} \;\leq\; \Bigl(\frac{1}{2} + \delta\Bigr)\bigl(\varphi + 2^{-R} + 5R\,2^{-2R}\bigr) + 2^{-2R} \;\leq\; \Bigl(\frac{1}{2} + \delta\Bigr)\varphi + 2^{2-R}. \qquad \blacksquare$$

Corollary 3.3. There is a constant c such that $\Phi(r) \leq c\, 2^{-\log^* r}$ for all $r \geq 1$.

Proof. Fix $R_0 \geq 512$ so large that Proposition 3.2 holds for all $R \geq R_0$. Let $R_{j+1} = 2^{256R_j}$ and $\delta_j = 20/\log^{(2)} R_j$. By Proposition 3.2, the numbers $a_j = 2^j \Phi(R_j)$ satisfy

$$a_{j+1} \;\leq\; (1 + 2\delta_j)\, a_j + 2^{\,j+3-R_j}.$$

For $j \geq 2$ we have $\log^{(2)} R_j = 8 + 256R_{j-2}$, so $\sum_j \delta_j < \infty$. Also $\sum_j 2^{\,j-R_j} < \infty$. Hence the numbers a_j are bounded, say $\Phi(R_j) \leq c_1 2^{-j}$.

The numbers R_j grow like a tower of exponentials. To bound $\log^* R_j$, we compare them with

an exact tower. Let $U_0 = R_0^2$ and $U_{j+1} = 2^{U_j}$, so that $\log^* U_j = j + \log^* U_0$. Then $R_j \le \sqrt{U_j}$ for all j. Indeed, if this holds for j, then

$$R_{j+1} = 2^{256R_j} \;\le\; 2^{256\sqrt{U_j}} \;\le\; 2^{U_j/2} = \sqrt{U_{j+1}},$$

because $\sqrt{U_j} \ge R_0 \ge 512$. Hence $\log^* R_j \le j + \log^* U_0$.

Finally let $r \ge R_0$, and let j be largest with $R_j \le r$. Then $\Phi(r) \le \Phi(R_j) \le c_1 2^{-j}$ and $\log^* r \le \log^* R_{j+1} \le j + 1 + \log^* U_0$. Together these give $\Phi(r) \le c\, 2^{-\log^* r}$ with $c = c_1\, 2^{1+\log^* U_0}$. For $r < R_0$ use $\Phi(r) \le 1$, and enlarge c if necessary. ∎

Proof of the lower bound in Theorem 1.1. Let G be a pancyclic graph with n vertices and $n + h$ edges. It is connected, so its rank is $r = h + 1$. By the definition of Φ and by Corollary 3.3,

$$n - 2 \;\le\; \Phi(r)\, 2^r \;\le\; c\, 2^{\,r - \log^* r}.$$

The trivial bound $n - 2 \le 2^r - 1$ gives $n \le 2^{r+1} \le 2^{2^r}$, hence $\log^* n \le \log^* r + 2$. Now take logarithms in the display:

$$h = r - 1 \;\ge\; \log(n-2) + \log^* r - \log c - 1 \;\ge\; \log n + \log^* n - \log c - 5.$$ ∎

4 The construction

IDEA | outline of the construction

The construction is a chain of k segments of one long cycle. Segment i has a chord between its two ends. So a path can pass it in two ways, and the two ways differ in length by 2^i. The chain therefore has 2^k paths of consecutive lengths, and it uses k chords. A few more chords close shorter and shorter initial parts of the chain, and these give the short cycles. The construction follows [5] (Figure 2).

Proposition 4.1. For every $n \ge 3$ there is a pancyclic graph with n vertices and at most $n + \lfloor \log n \rfloor + \log^* n + 1$ edges.

Proof. For $3 \le n \le 6$ take a cycle $u_0u_1 \cdots u_{n-1}$, with the chord u_0u_2 if $n \ge 4$ and also the chord u_0u_3 if $n = 6$. Now let $n \ge 7$.

The necklace. Let $k \ge 2$ be the largest integer with $2^k + k + 1 \le n$, so that $n \le 2^{k+1} + k + 1$. Let H be a cycle of length n, and let $Q \subseteq H$ be a path that is composed of consecutive segments $Q_0, \dots, Q_{k-1}$, where Q_i has $2^i + 1$ edges and joins v_i to v_{i+1}. Thus Q has $2^k + k - 1 \le n - 2$ edges. Add the chords $e_i = v_iv_{i+1}$ for $0 \le i < k$. For $1 \le j \le k$ let N_j be the union of $Q_0, \dots, Q_{j-1}$ and $e_0, \dots, e_{j-1}$ (Figure 2).

A path from v_0 to v_j in N_j uses either Q_i or e_i for each $i < j$, and it is shorter by 2^i if it uses e_i. So N_j has 2^j such paths, exactly one of each length from j to $2^j + j - 1$. The rest $H - Q$ of the outer cycle is a path from v_k to v_0 with $n - 2^k - k + 1$ edges. Together with N_k it contains a cycle of each length in the interval $[\, n - 2^k + 1,\ n \,]$.

Closing chords. For $2 \le j \le k$ the chord $f_j = v_0 v_j$ closes the paths of N_j to cycles of all lengths in the interval $I_j = [\, j + 1,\ 2^j + j \,]$. Add the chord f_j for $j = k$, then for $j = \lceil \log k \rceil$, and so on: as long as the last value of j is larger than 2, the next value is $\lceil \log j \rceil$. The last value is 2. If $j' = \lceil \log j \rceil$, then $2^{j'} \ge j$, so the intervals $I_{j'}$ and I_j overlap or are adjacent. Since $I_2 = [3, 6]$, these chords give a cycle of each length in $[\, 3,\ 2^k + k \,]$.

All lengths. The intervals $[\, 3,\ 2^k + k \,]$ and $[\, n - 2^k + 1,\ n \,]$ cover all lengths from 3 to n, except for the length $2^k + k + 1$ in the case $n = 2^{k+1} + k + 1$. In that case add one more chord, which joins v_k to the neighbour of v_0 on $H - Q$. This chord, the edge from that neighbour to v_0, and the path Q form a cycle of length $2^k + k + 1$.

The number of chords. For an integer $j \ge 3$ we have $\log^* \lceil \log j \rceil = \log^* j - 1$, and $\log^* 2 = 1$. So exactly $\log^* k$ chords f_j were added. The number of all chords is therefore at most $k + \log^* k + 1 \le \lfloor \log n \rfloor + \log^* n + 1$. ∎

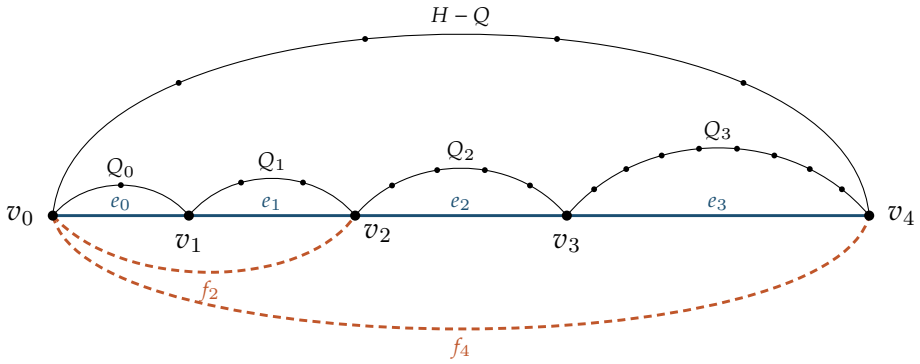


Figure 2. The construction for $n = 24$, where $k = 4$. The segments $Q_0, \dots, Q_3$ of the outer cycle H have $2, 3, 5, 9$ edges and are doubled by the chords $e_0, \dots, e_3$. The chord f_4 gives the cycle lengths $5, \dots, 20$, the rest $H - Q$ of the outer cycle gives $9, \dots, 24$, and f_2 gives $3, \dots, 6$.

5 Authorship and verification

KNT is a person. Astra and Fable 5.2 are the language models GPT-6 Astra and Claude Fable 5.2. KNT obtained Theorem 1.1, the forest estimate and their proofs in September 2026 with Astra [5], and directed the work at each step. Fable 5.2 wrote this short form.

Theorem 1.1 has a formal proof in Lean 4. The statement `Erdos1016.mainTheorem` has both bounds, with absolute additive constants, and `Erdos1016.Problem1016.upper_bound_integer` is the upper bound $\lfloor \log n \rfloor + \log^* n + 1$. The formal proof uses the forest estimate for $|E(W)| \le \sqrt{r}$, which is `Erdos1016.ShortProof.fewComponentForestEstimate`; Section 3 needs no more. The Lean kernel accepts the proofs with the axioms `propext`,

`Classical.choice` and `Quot.sound` only. The files are in [6], at commit `a5cb5fa`.

The kernel check covers these formal statements, and it does not cover the text of this paper. No referee has read the text.

References

[1] T. F. Bloom, Erdős Problem #1016, https://www.erdosproblems.com/1016, accessed November 2026.

[2] J. A. Bondy, Pancyclic graphs I, *J. Combin. Theory Ser. B* **11** (1971), 80–84.

[3] KNT, Astra and Fable 5.2, The forest proportion of a random even subgraph, companion manuscript, November 2026.

[4] S. Griffin, Minimal pancyclicity, arXiv:1312.0274 (2013).

[5] KNT, The minimum number of edges in a pancyclic graph, manuscript, September 2026, https://jehlp.net/erdos1016/.

[6] KNT, Astra and Fable 5.2, Lean 4 formalisation of the results of [5], https://github.com/JWKNT/erdos1016.