1 Introduction

Given a graph G, how long of a cycle can we guarantee in it? This is one of the basic questions in graph theory, and is naturally related to many of the central topics in the field, e.g. the study of Hamiltonicity. Perhaps the most basic result on this question is the following simple observation.

Fact 1.1. *If G has minimum degree $d \geqslant 2$, then G contains a cycle of length $\geqslant d+1$.*

Proof. Start walking through the vertices of G, always trying to avoid visiting a vertex you've already seen. You will eventually get stuck (since G is finite), at which point you have traversed a path in G, and the last vertex you reached has all of its neighbors on the path. Since it has at least d neighbors, the furthest-back one along the path creates a cycle of length at least $d+1$. □

This simple result is best possible, as shown by the example of $G = K_{d+1}$ (or, more generally, a disjoint union of K_{d+1}).

Nevertheless, while Fact 1.1 is best possible in general, it is not best possible in many special cases we care about. A famous example is Dirac's theorem, which states that if G has n vertices and $d \geqslant n/2$, then in fact G contains a Hamiltonian cycle, i.e. a cycle of length n. In fact, in the course of proving his Hamiltonicity theorem, Dirac proved (essentially) that Fact 1.1 can be improved by a factor of 2 whenever G is 2-connected.

In this talk, I will focus on another strengthening of Fact 1.1.

Theorem 1.2 (Erdős–Gallai 1959)**.** *If G has average degree $d \geqslant 2$, then G contains a cycle of length $\geqslant \lfloor d \rfloor + 1$.*

Theorem 1.2 is not particularly difficult to prove, especially if one is willing to rely on the result of Dirac mentioned above (in brief: we may pass to a subgraph of G with minimum degree at least $d/2$, and then the result of Dirac essentially allows us to win back this factor of two that we lose). However, I will admit to having never fully internalized the proof; unlike other classical results in extremal graph theory, such as Turán's theorem or the Kővári–Sós–Turán theorem, whose proofs feel like second nature to me, I have always found the Erdős–Gallai theorem somewhat mysterious.

Because Theorem 1.2 strengthens Fact 1.1, it is also best possible, as witnessed by the same extremal example of K_{d+1}. However, rather interestingly, this is no longer the unique (asymptotic) extremizer. A very different example is the complete bipartite graph $K_{s,t}$, where $s \leqslant t$. This graph has no cycle of length longer than $2s$, since every cycle must spend at least half its length in the side of size s. Moreover, as $t \to \infty$, the average degree of $K_{s,t}$ tends to $2s$. So apart from the lower-order terms, this is another extremal example, which seems to be genuinely different: while K_{d+1} avoids a long cycle because there is not enough space to squeeze it in, $K_{s,t}$ avoids a long cycle because it has a too-small vertex cover, which any cycle must spend at least half its length in.

This discussion may explain why Theorem 1.2 is somewhat harder to prove than it might seem at first glance: unlike something like Turán's theorem, where the unique extremal

structure naturally guides you towards a proof that identifies it, the fact that very different global structures are roughly comparable with respect to their average degree and longest cycle length seems to preclude such a simple proof. In fact, in this talk, I hope to sketch a new, and much more complicated, proof of (a slightly weaker form of) Theorem 1.2, which proceeds by first analyzing the structure of G, proving that in many cases it must actually have a cycle of length much longer than d, and finally in the remaining cases identifying the obstruction to a longer cycle.

2 New results

The other advantage of this more complicated proof is that it also yields two new strengthenings of the Erdős–Gallai theorem, one a "robustness" statement and one a "hyperstability" statement.

2.1 Hyperstability

In order to explain these, let us first take a step back and discuss stability in extremal combinatorics. Stability is a phenomenon that is present in a large number of extremal combinatorial problems, which is both interesting for its own sake and has numerous applications. Generally speaking, in extremal combinatorics, one seeks to maximize a function f over some (discrete) domain X; for example, Mantel's theorem asks about the maximum number of edges in a triangle-free graph, for which we would set X to be the set of all triangle-free graphs on n vertices, and f to be the function counting the number of edges. The first step in solving an extremal problem is to determine the maximum value M of f on X and ideally to also determine the set of extremizers, i.e. classify all $x^* \in X$ such that $f(x^*) = M$. Having done this, a stability statement is a result saying that all $x \in X$ for which $f(x)$ is close to M are, in turn, close to some extremizer x^*. In order to make this precise, we also need to pick some natural metric d on X, at which point a stability statement takes the following form. For every $\varepsilon > 0$, there exists $\delta > 0$ such that if $f(x) \geqslant (1-\delta)M$, then there is some extremizer x^* with $d(x, x^*) \leqslant \varepsilon$. In the context of Mantel's theorem, the stability theorem (first proved by Erdős and Simonovits) states that if an n-vertex triangle-free graph has at least $(1-\delta)\lfloor n^2/4 \rfloor$ edges, then it can be turned into a balanced complete bipartite graph by flipping at most εn^2 edges into non-edges and vice versa.

In additive combinatorics, such stability theorems are usually called "99% inverse theorems". Here, the 99% refers both to the fact that the assumption is that $f(x)$ is very close to its maximum value M, and to the fact that the conclusion is also of a 99% variety: 99% of the structure of x agrees with that of an extremizer. However, in additive combinatorics, it is frequently much more valuable to have "1% inverse theorems", where the assumption is only that $f(x) \geqslant \delta M$. In general one should not expect strong conclusions from such a weak assumption, and hence a typical 1% inverse theorem also has a 1% conclusion, namely that x shares an ε-fraction of its structure with an extremizer.

A relatively simple example of such a 1% inverse theorem concerns the following question.

Let $A \subseteq \mathbb{F}_2^n$ be a set, and let $f(A)$ denote the fraction of triples $a, b, c \in A$ for which $a+b+c$ is also an element of A. It is not hard to check that the sets A maximizing f are affine subspaces, for which 100% of triples of elements satisfy this property. The Balog–Szemerédi–Gowers and Freiman–Ruzsa theorems, both fundamental result in additive combinatorics, together imply a 1% inverse theorem for this statement: if a δ-fraction of the triples in A sum to another element of A, then an ε-fraction of the elements of A all lie in some affine subspace of size at most $|A|$.

1% inverse theorems seem to be rather rare in extremal graph theory. Indeed, while the stability version of Mantel's theorem gives us a lot of information on the structure of triangle-free graphs with $(1-\delta)\lfloor n^2/4 \rfloor$ edges, we know how to say very little about the structure of triangle-free graphs with δn^2 edges. In particular, it is not the case that such graphs need to "correlate" well with complete bipartite graphs.

However, one of our new results yields a "hyperstability" version of the Erdős–Gallai theorem, which incorporates the best aspects of the 1% and 99% inverse theorem regimes. Namely, the assumption is of the 1% type, and the conclusion is of the 99% type.

Theorem 2.1 (Christoph–Müyesser–W. 2026)**.** *Let G be an n-vertex graph, and let d be an integer. Either G contains a cycle of length at least d, or else one can delete $o(dn)$ edges from G to obtain a graph in which every connected component has a vertex cover of order at most d.*

Critically, note that while we use the same letter d as above, it is no longer necessarily the average degree of G. Indeed, this theorem makes sense even when G has average degree, say, $d/1000$. Even in that case, either we are guaranteed a cycle of length at least d—a factor of 1000 better than given by Theorem 1.2—or else we can delete a negligible fraction of the edges to obtain a graph in which every connected component has a vertex cover of order at most d. Note that such a guarantee is best possible, in case G is a disjoint union of K_d (for example).

As mentioned, the assumption is a 1% type of assumption: the Erdős–Gallai theorem states that the maximum average degree we can have without containing a cycle of length at least d is d, and Theorem 2.1 operates even with graphs whose average degree is a very small constant multiple of this maximum. However, the conclusion is of 99% type: we can keep 99% of the edges of G to obtain essentially the structure of the extremal example, namely a disjoint union of pieces with vertex cover of order d.

The only other such result I am aware of in extremal combinatorics is due to Pokrovskiy, who termed it a *hyperstability* theorem (referring to the fact that it operates at a rather larger regime than the classical stability theorems). I won't state his theorem precisely, but it is a very similar statement, simply about bounded-degree trees rather than about cycles. And in fact, our proof builds on the tools that Pokrovskiy developed, as I will shortly explain.

Incidentally, it is worth remarking that Pokrovskiy's work lies in a long line of work on the so-called Erdős–Sós conjecture, which is about the extremal number of trees. Since the 1990s, it has been understood that the way to prove this conjecture is to develop a robust structural theory of sparse graphs, and Pokrovskiy's main contribution is indeed such a structural description. As it happens, an AI model recently found a short and simple

inductive proof of the Erdős–Sós conjecture, and it turns out that building such a structural theory is not necessary if your only goal is to narrowly prove this conjecture. However, I personally think that the structural insights that Pokrovskiy developed, as well as the remarkable and unexpected hyperstability phenomenon that he discovered, are far more interesting than the Erdős–Sós conjecture itself.

2.2 Robustness

A major trend in extremal combinatorics is to start with an existing result, typically of the form "property P implies property Q", and to ask how *robust* the result is. Namely, does property P (or a minor strengthening of it) imply something much beyond property Q? A particularly well-studied example concerns Dirac's theorem, which we should recall states that every n-vertex graph with minimum degree at least $n/2$ contains a Hamiltonian cycle. It is now known, for example, that every such graph must actually contain superexponentially many Hamiltonian cycles, as well as a large collection of edge-disjoint Hamiltonian cycles. That is, starting from the same assumption as Dirac's theorem, we end up with a much more "robust" version of the conclusion.

Among the most well-studied topics in this direction concerns robustness under random edge sampling. Namely, let G be a graph and let $p \in [0,1]$, and let G_p denote the random subgraph obtained by retaining each edge of G independently with probability p. In this setting, one seeks to ask whether, for an appropriate value of p, the random subgraph G_p typically also has property Q. This is another way of measuring the robustness of the implication $P \implies Q$: not only does the graph G have property Q, it also has it when we randomly delete a large number of its edges. Note that this is frequently even stronger than the counting results mentioned above. For example, it was shown by Krivelevich, Lee, and Sudakov that if G is a Dirac graph and $p \geqslant C \log n/n$ for some constant C, then G_p has a Hamiltonian cycle with probability tending to 1 as $n \to \infty$; since the expected number of Hamiltonian cycles in G_p is p^n times the number in G, one can easily deduce from this result that G contains at least $(n/(C \log n))^n$ Hamiltonian cycles.

In this direction, we prove the following robust version of the Erdős–Gallai theorem, resolving a conjecture of Glebov, Naves, and Sudakov.

Theorem 2.2 (Christoph–Müyesser–W. 2026)**.** *For every $c > 0$, there exists $K > 0$ such that the following holds. If G is an n-vertex graph of average degree d and $p \geqslant K/d$, then with probability tending to 1 as $n \to \infty$, G_p contains a cycle of length at least $(1-c)d$.*

In particular, if $p = \omega(1/d)$, then with high probability G_p contains a cycle of length at least $(1 - o(1))d$, i.e. asymptotically matching the Erdős–Gallai theorem. In particular, by taking $p = 1$, we obtain the Erdős–Gallai theorem asymptotically. Moreover, this result is best possible up to the dependence between c and K (which our proof gives very bad estimates for), as shown by the example of $G = K_{d+1}$.

3 Pokrovskiy's structure theorem

I will not really discuss the proof of either of our new theorems, but I hope to very roughly sketch our approach, focusing on how to use it to asymptotically prove the Erdős–Gallai theorem. Namely, we will aim show that every graph of average degree d contains a cycle of length at least $(1-o(1))d$ (where the $o(1)$ tends to 0 as $d \to \infty$). This proof will also establish Theorem 2.1 with little extra difficulty. However, some highly non-trivial additional ideas are needed to establish the probabilistic robustness statement, which I will not discuss.

The key input we need is Pokrovskiy's structure theorem for sparse graphs. This theorem (which is not explicitly stated or proved in his work, but follows fairly straightforwardly from his techniques) allows us to edge-decompose any n-vertex graph of average degree d into four subgraphs, each having some prescribed structure, as well as a fifth subgraph consisting of a small number of "junk" edges. Three of the four types are very special: if any of them has an appreciable number of edges, then automatically it must contain an extremely long cycle. Thus, the only remaining case is when almost all the edges of the graph lie in the fourth subgraph, whose defining property is that all of its components have a small vertex cover.

Unfortunately, Pokrovskiy's theorem is rather overwhelming, so I want to build up to it rather slowly. As such, I will begin by discussing the first type of subgraph it yields, and why any such graph must contain very long cycles.

3.1 Nowhere dense graphs

A key feature of Pokrovskiy's theorem is to distinguish between dense graphs—those whose number of edges is quadratic in their order—and on non-dense graphs. In fact, we will focus on *nowhere dense* graphs, which are those graphs that have no dense subgraph of order $O(d)$.

Definition 3.1. A graph H is called *nowhere dense* if every set of $10d$ vertices spans at most $d^2/500$ edges.

This is a fairly weak condition, in the sense that (for example) a graph with no edges trivially satisfies it. However, it turns out that if a nowhere dense graph *does* have a lot of edges, then it is suddenly a very special graph. In particular, the nowhere dense condition implies that such a graph is an extremely strong vertex- and edge-expander, at least on sets of order $O(d)$.

For our purposes, we will not need any of that; all we need is that nowhere dense graphs with a lot of edges contain long cycles. There are a number of ways of proving such a statement; the following proof is based on a trick of Bucić–Frederickson–Müyesser–Pokrovskiy–Yepremyan.

Lemma 3.2. *Let H be a nowhere dense graph, and suppose that the minimum degree of H is at least $d/10$. Then H contains a cycle of length at least $9.9d$.*

Note that we are actually assuming a minimum degree condition on H rather than that H has a lot of edges, but this is not really significant (because every graph of average degree

$d/5$ contains a subgraph of minimum degree at least $d/10$, and the nowhere dense condition is preserved under passing to subgraphs). Note too that Lemma 3.2 yields a much stronger conclusion than Fact 1.1 or Theorem 1.2: the cycle we get is a factor of 99 longer than our minimum degree assumption, thanks to the extra structure given by the nowhere density.

Proof of Lemma 3.2. The proof is a simple modification of that of Fact 1.1. As before, we will walk around the graph, trying to not return to a vertex we have already visited. However, we will also maintain throughout the stronger property that as we walk around, our current vertex has at least $d/20$ neighbors which we have not yet visited. We can certainly start this process because of the minimum degree assumption, and hence we can keep doing it until we eventually get stuck.

When we get stuck, we have traversed a path P in G, and have reached an endpoint v of P with the property that v has a set N of at least $d/20$ neighbors not on P, but every vertex in N has at most $d/20$ neighbors not in P. Consequently, every vertex in N has at least $d/20$ neighbors in P.

If any of these edges goes at least $9.9d$ steps behind the endpoint v, we automatically obtain a cycle of length at least $9.9d$. As a consequence, we may assume that all of these neighbors of all vertices in N lie among the $9.9d$ closest vertices on P. Thus, the total number of vertices that these edges touch is at most $9.9d + d/20 < 10d$. Moreover, the number of edges that we are counting here is at least $(d/20)^2 > d^2/500$. In other words, we have found a set of fewer than $10d$ vertices spanning more than $d^2/500$ edges, a contradiction to the nowhere density of H. □

Of course, all of the constant factors in this proof (and in the definition of nowhere density) are rather flexible, and one can obtain stronger or weaker guarantees under stronger or weaker assumptions by playing with the constants. But the point is that nowhere density is a very strong condition: it allows us to find cycles whose length is a huge constant multiple larger than what is guaranteed by the trivial Fact 1.1. In particular, if we can find a nowhere dense subgraph H of our given graph G which has any substantial number of edges, then we are automatically done proving Theorem 1.2.

How can we go about finding a nowhere dense subgraph of G? The obvious thing to try is the following algorithm:

1. If G is nowhere dense, stop.
2. If not, there must exist a set $S \subseteq V(G)$ of order at most $10d$ which spans at least $d^2/500$ edges.
3. Remove the edges of $G[S]$ from G and return to step 1.

Essentially, we keep pulling out dense pieces of G until we get stuck. Certainly, when this algorithm terminates, we have produced a nowhere dense subgraph of G. If this subgraph has any appreciable minimum (or average) degree, we are done by Lemma 3.2. So essentially the only case we need to worry about is when the dense pieces we have pulled out have, collectively, essentially all the edges of G.

3.2 Dense pieces and their cores

Already by doing almost nothing, we have made progress: we can now assume that G is essentially the edge-disjoint union of a bunch of *dense* pieces, i.e. a bunch of graphs with $O(d)$ vertices and $\Omega(d^2)$ edges. The reason this is good news is that we have a lot of powerful tools for dealing with dense graphs, such as Szemerédi's regularity lemma.

The problem is that each dense piece individually might not suffice for our purposes: because each piece might only have average degree $d/100$, there is no reason to expect any dense piece to contain a long cycle by itself. Thus, we need to keep track of further structure; it turns out that the following definition is a very useful one.

Definition 3.3. Let H be a dense graph. Its *core* is the largest subgraph $R \subseteq H$ which is $(d/100)$-regular (breaking ties arbitrarily).

I should remark that from this point forward, more or less every definition and lemma statement will not, strictly speaking, be true. In particular, the constant 100 in the definition of the core is not the correct one to choose—the correct choice depends in complicated ways on all the other implicit constants, including those in the definition of nowhere dense graphs and in the big-O in the definition of a dense graph. However, this is the last time I will mention this, and all future statements will be with explicit, and incorrect, constants.

Why is the core a useful notion? There are at least two good reasons, given by the next two lemmas.

Lemma 3.4. *If H is a dense graph with core R, and if H is $(d/10)$-connected, then H contains a cycle of length at least $(1 - o(1))|V(R)|$.*

This lemma is highly non-trivial to prove. However, at a high level, the idea is the following. First, it is not too hard to essentially decompose $V(R)$ into $O(1)$ paths, because it is a dense regular graph (e.g. this follows from Szemerédi's regularity lemma). Now, using the high connectivity of H, we can join up the ends of these $O(1)$ paths using vertices in $V(H) \setminus V(R)$, thus merging them into a cycle. As this cycle contains at least all the vertices in the path, which covered almost all of $V(R)$, we obtain the stated conclusion.

In particular, if we can find a dense graph H whose core has order d, we are done, even though H itself might only have average degree $d/10$: just as with Lemma 3.2, the assumption of having a large core allows us to find a much longer cycle than we should expect in general.

The other reason that cores are useful is that they are closely related to vertex covers.

Lemma 3.5. *If H is a dense graph and R is its core, then $V(R)$ is an approximate vertex cover of H: at most 1% of the edges of H do not have an endpoint in $V(R)$.*

This lemma follows easily from a well-known but non-trivial fact, which is that dense graphs contain dense regular subgraphs (this is again easiest to prove using the regularity lemma). If 1% of the edges of H do not have an endpoint in $V(R)$, we can find a $(d/100)$-regular graph among these edges. By taking its disjoint union with R, we obtain a larger $(d/100)$-regular subgraph of H than R, contradicting that R is the core of H.

3.3 Pokrovskiy's merging algorithm

We are now ready to describe the heart of Pokrovskiy's argument. As with everything else, this is actually a wildly oversimplified, and thus grossly incorrect, description of Pokrovskiy's approach, but the hope is that it captures the main ideas.

Recall that we were given a graph G of average degree d. Our first step was to repeatedly pull out of G dense pieces, continuing until we are left with a nowhere dense subgraph (which may well be empty). In fact, we may as well assume that this nowhere dense graph is essentially empty: if it has an appreciable number of edges, we can find within it a cycle of length at least d by Lemma 3.2, and conclude the proof of the Erdős–Gallai theorem.

The remainder of the edges of G are partitioned into a collection $\mathcal{H}$ of dense graphs, and by the above we may in fact assume that $\mathcal{H}$ contains almost all the edges of G. Let us further assume that every graph in $\mathcal{H}$ is $(d/10)$-connected; very roughly speaking, we are allowed to make this assumption by, if necessary, splitting each graph in $\mathcal{H}$ into its $(d/10)$-connected components.

We now run the following merging algorithm on $\mathcal{H}$: if there are two graphs $H, H' \in \mathcal{H}$ with $|V(H) \cap V(H')| \geqslant d/10$, then we merge them into a single graph. Formally, we replace $\mathcal{H}$ by the new collection $(\mathcal{H} \setminus \{H, H'\}) \cup \{H \cup H'\}$. Because the size of $\mathcal{H}$ decreases by one at every step, this process will eventually terminate with some collection $\mathcal{H}_{\text{final}}$.

Because the graphs in the initial collection were all $(d/10)$-connected, and because the merging rule only merges graphs who overlap on at least $d/10$ vertices, every graph in $\mathcal{H}_{\text{final}}$ is $(d/10)$-connected. If some graph $H \in \mathcal{H}_{\text{final}}$ has a core R of order d or greater, we are done by Lemma 3.4: we conclude that this H contains a cycle of length at least $(1 - o(1))d$, and hence so does G. So we may as well assume that all graphs in $\mathcal{H}_{\text{final}}$ have their core of order less than d.

Recall that the graphs in $\mathcal{H}_{\text{final}}$ are edge-disjoint, since that was true with our initial collection $\mathcal{H}$, and we maintained it throughout the process. Moreover, by the merging rule, they are "almost" pairwise vertex-disjoint: every pair share at most $d/10$ vertices. Suppose, by some miracle, that the graphs in $\mathcal{H}_{\text{final}}$ are truly vertex-disjoint. This is not obviously good for proving Theorem 1.2, but at least it's definitely good for proving Theorem 2.1: in this case, we have partitioned almost all the edges of G into a vertex-disjoint union of pieces, each of which has an approximate vertex cover of order at most d (by Lemma 3.5). And in fact, it's hard but not too hard to also prove Theorem 1.2 under this assumption: the key point is that because almost all edges of G are in the vertex-disjoint union of the graphs in $\mathcal{H}_{\text{final}}$, there must be some $H \in \mathcal{H}_{\text{fina}}$ with average degree at least $(1 - o(1))d$. From this, as well as the small approximate cover, one can pass to a dense subgraph of average degree $(1 - o(1))d$, which means that it suffices to prove the Erdős–Gallai theorem only for *dense* graphs. This is not obviously an easier task, but it turns out to be doable by again invoking powerful tools like the regularity lemma.

The only remaining issue is that the graphs in $\mathcal{H}_{\text{final}}$ are not necessarily *actually* vertex-disjoint: they merely pairwise share few vertices. And this can actually be a serious issue: you could imagine that the graphs of $\mathcal{H}_{\text{final}}$ are arranged like the lines of a finite projective plane, so that every pair of them intersect in only a single vertex, but the global interaction

between them is extremely complex, and we cannot cheaply delete a few edges to pass to a genuinely vertex-disjoint collection.

Dealing with this case is by far the hardest part of both Pokrovskiy's and our arguments, and I will not really say anything about it. Nevertheless, it turns out that such a "complexly intersecting" collection of graphs is still sufficient to find a long cycle, even if it has very low average degree.

I have presented Pokrovskiy's merging algorithm with our application of finding long cycles in mind, but in fact, this algorithm yields a decomposition of any graph into nicely structured pieces. To wrap up, let me state a (somewhat informal) version of this theorem.

Theorem 3.6 (Pokrovskiy's structure theorem)**.** *Let G be a graph and let d be sufficiently large. We can edge-decompose G into subgraphs of the following types:*

0. *At most $o(e(G))$ "junk" edges which do not fit anywhere else.*
1. *A nowhere dense graph.*
2. *A vertex-disjoint union of graphs, each with a vertex cover of order at most d.*
3. *An edge-disjoint union of $(d/10)$-connected graphs, each containing a $(d/100)$-regular subgraph of order at least d.*
4. *An edge-disjoint union of a collection $\mathcal{H}$ of graphs, with the following properties:*
   - *Each $H \in \mathcal{H}$ is $(d/10)$-connected and has diameter at most 10,*
   - *$|V(H) \cap V(H')| = o(d)$ for all $H, H' \in \mathcal{H}$, and*
   - *every $H \in \mathcal{H}$ has at least $d/100$ vertices which lie in some other $H' \in \mathcal{H}$.*

Items 1–3 should be familiar from the discussion above, and hopefully you believe that item 4 is a version of the "complexly intersecting" idea I described above. Although this decomposition is rather complex, this complexity is unfortunately necessary: one can cook up examples showing that each of the four types of structured subgraph is necessary in general.

The other thing I want to stress about this result, and about how we use it, is that item 2 is genuinely different from the rest. In each of the other cases, we are able to find a cycle of length at least d even if the subgraph we are investigating only has average degree $d/100$. That is, the extra structure given in cases 1,3,4, while rather complicated to understand, really is powerful, allowing us to get results much stronger than the Erdős–Gallai theorem gives for general graphs. As such, we are able to essentially assume that all the edges of G lie in the vertex-disjoint union of pieces with small cover (item 2). This already yields Theorem 2.1 more or less directly, and moreover can be used to deduce the existence of a long cycle.

Philosophically, this says something rather remarkable about the structure of sparse graphs, where we now think of G has having $n \gg d$ vertices and average degree d. Basically, G may have some parts that are secretly a vertex-disjoint union of dense pieces of average

degree d, and thus in these pieces is not *really* a sparse graph. The remainder of G might be genuinely sparse (and very compliated), but if it is, it has some powerful additional structure which we can exploit. But the only way for a sparse graph to *not* have any such structure is for it to actually be the union of dense components!