OXFORD

EXPONENTIALLY MANY GRAPHS ARE DETERMINED BY THEIR SPECTRUM

by ILLYA KOVAL[†] and MATTHEW KWAN[‡] (ORCID)

Institute of Science and Technology Austria (ISTA), Am Campus 1, Klosterneuburg 3400, Austria

[†]E-mail: illya.koval@ist.ac.at
[‡]Corresponding author. E-mail: matthew.kwan@ist.ac.at

ABSTRACT

As a discrete analogue of Kac's celebrated question on 'hearing the shape of a drum' and towards a practical graph isomorphism test, it is of interest to understand which graphs are determined up to isomorphism by their spectrum (of their adjacency matrix). A striking conjecture in this area, due to van Dam and Haemers, is that 'almost all graphs are determined by their spectrum', meaning that the fraction of unlabelled n-vertex graphs which are determined by their spectrum converges to 1 as $n \to \infty$.

In this paper, we make a step towards this conjecture, showing that there are exponentially many n-vertex graphs which are determined by their spectrum. This improves on previous bounds (of shape $e^{c\sqrt{n}}$). We also propose a number of further directions of research.

1. INTRODUCTION

A classical question, popularized in 1966 by Kac [26], is whether one can 'hear the shape of a drum': if we know the 'spectrum' of a planar domain $D \subseteq \mathbb{R}^2$ (formally, the eigenfrequencies of the wave equation on D, with Dirichlet boundary conditions), is this enough information to reconstruct D up to isometry? Famously (and perhaps surprisingly), the answer to this question is 'no': in 1992, Gordon, Webb and Wolpert [17] managed to construct two different 'drums' with the same spectrum.

A much shorter version of this story also took place in graph theory. In 1956, in a paper studying connections between graph theory and chemistry, Günthard and Primas [18] asked whether one can reconstruct a graph up to isomorphism given the eigenvalues of its adjacency matrix. (The *adjacency matrix* of a (simple) graph G, with vertices $v_1, \dots, v_n$, is the zero-one matrix $\mathrm{A}(G) \in \{0,1\}^{n \times n}$ whose (i,j)-entry is 1 if and only if G has an edge between v_i and v_j.) Due to the discrete nature of this question, the search for counterexamples is much easier than for Kac's question: only 1 year later, Collatz and Sinogowitz [8] exhibited a pair of graphs with the same spectrum. This has some rather important practical consequences: if it were the case that all graphs were determined by their spectrum, this would give rise to a very simple graph isomorphism test. It is an open problem to find a provably efficient graph isomorphism test, and spectral information is often used to distinguish graphs in practice.

A striking conjecture due to van Dam and Haemers [19, 33, 34] (also suggested somewhat later and seemingly independently by Vu [35]) is that graphs which cannot be uniquely identified by their spectrum are extremely rare, in the following natural asymptotic sense.

DEFINITION 1.1 The *spectrum* of a graph is the multiset of eigenvalues of its adjacency matrix. A graph G is *determined by its spectrum* (*DS* for short) if there is no other graph (non-isomorphic to G) which has the same spectrum as G.

CONJECTURE 1.2 *The fraction of unlabelled n-vertex graphs which are determined by their spectrum converges to 1 as $n \to \infty$. Equivalently [16], the number of (unlabelled) n-vertex graphs determined by their spectrum is*

$$(1 - o(1)) \frac{2^{n(n-1)/2}}{n!}.$$

(The number of labelled graphs on a particular set of n vertices is 2n(n–1)/2, and it is well known (see [16, Lemma 2.3.2]) that all but a vanishingly small fraction of these have a trivial automorphism group.)

REMARK 1.3 To elaborate on the attribution here: for a very long time, it has been an important question in spectral graph theory to understand the asymptotic proportion of graphs which are DS (for example, Schwenk [30] conjectured the *opposite* of Conjecture 1.2 in 1973, and Godsil and McKay [15] described this general question as 'one of the outstanding unsolved problems in the theory of graph spectra'). The possibility that Conjecture 1.2 might hold (supported by mounting computational evidence) was first suggested by van Dam and Haemers [33] in 2003 (also later in [34]), though they did not explicitly make a conjecture. It seems that Conjecture 1.2 first appeared explicitly in a paper of Haemers [19]. Vu seems to have arrived at Conjecture 1.2 via quite a different pathway: in [35], he presents it as a graph-theoretic variant of a similar conjecture in random matrix theory. We also remark that Garijo, Goodall and Nešetřil [14] and Noy [28] situated Conjecture 1.2 in (different) general frameworks which include a number of other questions about reconstructing graphs from various types of information.

Conjecture 1.2 is rather bold, on account of the fact that there are very few known examples of DS graphs. Indeed, to show that a graph G is DS (without exhaustively computing the spectra of all other graphs on the same number of vertices), it seems necessary to somehow translate information about the spectrum of G into information about the combinatorial structure of G. Spectral graph theory has a number of different tools along these lines, but all of them are rather crude, and essentially all known examples of DS graphs have very special structure. (For example, to prove that complete graphs are DS, one uses the fact that the n-vertex complete graph is the only n-vertex graph with exactly $\binom{n}{2}$ edges.)

To the best of our knowledge, the best lower bounds on the number of DS graphs are all of the form $e^{c\sqrt{n}}$ for some constant $c > 0$. Such a bound was first observed by van Dam and Haemers [33, Proposition 6], who proved that G is DS whenever every connected component of G is a complete subgraph (the number of graphs of this form is precisely the number of integer partitions of n, which is approximately $e^{c\sqrt{n}}$ for $c = \pi\sqrt{2/3}$ by the Hardy–Ramanujan theorem [23]). Several other families of graphs, similarly enumerated by integer partitions, have since been discovered (see [32, 38]). On the other hand, there has been much more progress in the *opposite direction* to Conjecture 1.2, proving lower bounds on the number of graphs which are *not* DS. For example, a famous result of Schwenk [30] says that only a vanishingly small fraction of trees are DS (meaning that almost all of the exponentially

many unlabelled n-vertex trees are non-DS), and, using an operation that is now known as *Godsil–McKay switching*, Godsil and McKay [15] (see also [20]) proved that the number of n-vertex graphs which are not DS is at least

$$(1 - o(1)) \frac{n^2}{12 \cdot 2^n} \cdot \frac{2^{\binom{n}{2}}}{n!}.$$

In this paper we prove the first exponential lower bound on the number of DS graphs, finally breaking the '$e^{c\sqrt{n}}$ barrier' (and thereby answering a question of van Dam and Haemers [33]).

THEOREM 1.4 *The number of (unlabelled) n-vertex graphs determined by their spectrum is at least e^{cn} for some constant $c > 0$.*

REMARK 1.5 Our proof shows that we can take $c = 0.01$ for large n, but we made no serious attempt to optimize this.

We will outline our proof strategy in Section 2, but to give a quick impression: we consider an explicit family of 'nice graphs', each consisting of a long cycle with leaves attached in various carefully chosen ways. Then, we consider a family of n-vertex graphs $\mathcal{Q}_n$ obtained by combining complete graphs with *line graphs* of nice graphs, in such a way that certain inequalities and number-theoretic properties are satisfied. (The line graph $\mathrm{line}(G)$ of a graph G has a vertex for each edge of G, and two vertices in $\mathrm{line}(G)$ are adjacent if the corresponding edges of G share a vertex). We then prove that there are exponentially many graphs in $\mathcal{Q}_n$ and that all graphs in $\mathcal{Q}_n$ are determined by their spectrum. We remark that there is an essential tension in the choice of $\mathcal{Q}_n$: in order to prove a strong lower bound we would like our families of graphs to be as 'rich' as possible, containing graphs with a wide variety of structure, but in order to reconstruct a graph using the limited information that is (legibly) available in its spectrum, we can only work with graphs with very special structure.

1.1. Further directions

It seems that significant new ideas would be required to go beyond the exponential bound in Theorem 1.4. Indeed, if we consider all the known combinatorial parameters that can be extracted from the spectrum of an n-vertex graph, then we end up with a list of about $2n$ integers (most notably, the first n spectral moments describe the number of closed walks of each length, and the n non-leading coefficients of the characteristic polynomial can be interpreted as certain weighted sums of subgraph counts). In order to use this combinatorial information to reconstruct say $\exp(n^{1+\varepsilon})$ different graphs, we would need to use a huge amount of information from each of the integers in our list: roughly speaking, the variation in each integer must correspond to about $\exp(n^{\varepsilon})$ different graphs. It is hard to imagine a natural combinatorial argument that could reconstruct so many different graphs from a single integer of information.

Instead, it seems that *non-constructive* methods may be necessary in order to prove Conjecture 1.2 or even to make much progress beyond Theorem 1.4. Is there some algebraic criterion which describes whether a graph is DS, without necessarily providing a combinatorial procedure to reconstruct the graph. (Some progress in this direction was made by Wang [36], who found an arithmetic criterion for a graph to be determined by its so-called "generalised spectrum"). Can one somehow show that the DS property is 'generic' without describing *which* graphs are DS?

We would also like to propose a number of other questions related to Conjecture 1.2.

- Consider two different n-vertex graphs G, G', chosen uniformly at random, and let Q_n be the probability that G and G' have the same spectrum. How large is this probability? It seems that one can obtain an exponential upper bound

$$Q_n \leq \mathbb{P}[\det(G) = \det(G')] \leq \sup_{d \in \mathbb{R}} \mathbb{P}[\det(G) = d] \leq e^{-cn}$$

for some $c > 0$, using powerful techniques in random matrix theory (see [7]).

- Conjecture 1.2 is equivalent to the statement that among all n-vertex graphs, there are

$$(1 - o(1)) \frac{2^{\binom{n}{2}}}{n!}$$

different spectra. What lower bounds can we prove on the number of different spectra realizable by n-vertex graphs? There are several different ways to prove an exponential lower bound: in particular, such a bound follows from Theorem 1.4, from the above bound $Q_n \leq e^{-cn}$ or from results on the range of possible determinants of $n \times n$ binary matrices (see [31]).

- Although it is known [30] that almost all trees are *not* DS, it would still be interesting to prove lower bounds on the number of DS trees. Could it be that there are exponentially many?
- In the continuous setting ('hearing the shape of a drum'), the *spectral rigidity* conjecture of Sarnak (see [29]) suggests that despite the fact that there are drums with the same spectrum, such drums are always 'isolated' from each other: for any drum, making a sufficiently small change to the shape of the drum always changes its spectrum. One can also ask similar questions for graphs. For example, as a weakening of Conjecture 1.2, we conjecture that for a $(1 - o(1))$-fraction of labelled graphs on n vertices, any nontrivial addition/deletion of at most $(1/2 - \varepsilon)n$ edges (for any constant $\varepsilon > 0$) results in a graph with a different spectrum. If this were true it would be best possible: for almost all n-vertex graphs G, one can exchange the roles of two vertices by adding and removing about $n/2$ edges (obtaining a graph which is isomorphic to G and therefore has the same spectrum).
- Apart from the adjacency matrix, there are several other matrices which can be associated with a graph. Perhaps the best-known examples are the *Laplacian* matrix and the *signless Laplacian* matrix (which are both actually used in this paper; see Definition 2.1). Such matrices give us different notions of graph spectra, with which we can ask variations on all the questions discussed so far. Actually, the Laplacian analogue of Theorem 1.4 has already been proved, taking advantage of the fact that the Laplacian spectrum is much better-behaved with respect to *complements*: Hammer and Kelmans [21] showed that all 2^n of the *threshold graphs* on n vertices (i.e., all n-vertex graphs which can be constructed from the empty graph by iteratively adding isolated vertices and taking complements) are determined by their Laplacian spectrum. In the course of proving Theorem 1.4, we actually end up giving new proofs of the analogous result for Laplacian and signless Laplacian spectra. It is still open (and not obviously easier or harder than for the adjacency spectrum) to prove better-than-exponential lower bounds on the number of n-vertex graphs determined by their Laplacian or signless Laplacian spectrum.

2. PROOF OVERVIEW

We start by defining the *Laplacian matrix* and the *signless Laplacian matrix*, two variations on the adjacency matrix.

DEFINITION 2.1 Consider a (simple) graph G with vertices $v_1, \ldots, v_n$. Let $\mathrm{D}(G)$ be the diagonal matrix whose (i, i)-entry is the degree of v_i and recall the adjacency matrix $\mathrm{A}(G)$ of G.

- The *Laplacian matrix* is defined as $\mathrm{L}(G) = \mathrm{D}(G) - \mathrm{A}(G)$.
- The *signless Laplacian matrix* is defined as $|\mathrm{L}(G)| = \mathrm{D}(G) + \mathrm{A}(G)$.

We sometimes refer to the spectra of $\mathrm{A}(G)$, $\mathrm{L}(G)$ and $|\mathrm{L}(G)|$ as the *adjacency spectrum*, *Laplacian spectrum* and *signless Laplacian spectrum* of G, respectively. We say that a graph G is *determined by its Laplacian spectrum* (respectively, *determined by its signless Laplacian spectrum*) if there is no other graph (non-isomorphic to G) which has the same Laplacian spectrum (respectively, signless Laplacian spectrum) as G.

While the adjacency matrix is the simplest and most natural way to associate a matrix to a graph, all three of the above notions of spectrum contain slightly different information about G, which can be useful for different purposes. For this paper, the crucial fact about the Laplacian spectrum is that it determines the number of *spanning trees* of a graph, via Kirchhoff's celebrated *matrix-tree theorem* (Theorem 3.13). In particular, the Laplacian spectrum tells us whether a graph is connected or not.

Fortunately, there are some connections between the above three notions of spectrum, which we will heavily rely on in this paper. For example, two simple observations are that:

- if a graph is bipartite, then its signless Laplacian spectrum is the same as its Laplacian spectrum (Fact 3.1);
- if two graphs have the same signless Laplacian spectrum, then their *line graphs* have the same adjacency spectrum (Proposition 3.15).

Unfortunately, there are some limitations to these connections. In general, neither the Laplacian spectrum nor the signless Laplacian spectrum of a graph contain enough information to actually determine whether the graph is bipartite (and it is *not* true that for a bipartite graph to be determined by its Laplacian spectrum is the same as for it to be determined by its signless Laplacian spectrum). Also, if a graph Q has the same adjacency spectrum as the line graph of some graph G, it does not necessarily follow that Q is the line graph of some graph with the same signless Laplacian spectrum as G (it does not even follow that Q is a line graph at all, though a deep structure theorem of Cameron, Goethals, Seidel and Shult [6], building on a previous slightly weaker theorem of Hoffman [25], shows that every connected graph which has the same adjacency spectrum as a line graph must be a so-called *generalized line graph*, with finitely many exceptions).

Despite these limitations, in our proof of Theorem 1.4, it is nonetheless extremely useful to move between the three different notions of graph spectra. Roughly speaking, our proof of Theorem 1.4 can be broken down into three parts. First, we describe an explicit family of graphs ('nice graphs') and prove that they are determined by their Laplacian spectrum (making crucial use of the matrix-tree theorem). Second, we prove that any graph which has the same signless Laplacian spectrum as a bipartite nice graph must be bipartite (from which we can deduce that in fact every bipartite nice graph is determined by its signless Laplacian spectrum). Finally, we define a family Q_n of exponentially many n-vertex graphs (which are essentially line graphs of bipartite nice graphs, with some small adjustments for number-theoretic reasons) and use the Cameron–Goethals–Seidel–Shult theorem to show that if a graph has the same adjacency spectrum as a graph in Q_n, then both graphs must have been constructed from line graphs with the same signless Laplacian spectrum. Putting everything together, we see that all of the exponentially many graphs in Q_n are determined by their adjacency spectrum.

We next outline each of the above three parts of the proof of Theorem 1.4 in more detail.

2.1. Nice graphs and the Laplacian spectrum

First, we define nice graphs and outline how to prove that they are determined by their Laplacian spectrum.

DEFINITION 2.2 Say that a graph is *sun-like* if it is connected, and deleting all degree-1 vertices yields a cycle. (The reason for this terminology is that the name "sun graph" is sometimes used in the literature to describe a graph obtained from a cycle by adding a leaf to each vertex). Equivalently, a sun-like graph can be constructed by taking a cycle C and attaching some leaves to some vertices of C. If a vertex of C has i leaves attached to it (equivalently, if the vertex has degree $i + 2$), we call it an *i-hub*. We simply call a vertex a *hub* if it is an i-hub for some $i \geq 1$ (equivalently, if its degree is at least 3).

For (integer) parameters $k \geq 1$ and $\ell \geq \max(12k, 15)$, say that a graph G is *(ℓ, k)-nice* if:

- G is a sun-like graph;
- the unique cycle C in G has length ℓ;

- there are exactly $k + 1$ hubs, one of which is a 1-hub and the others of which are 2-hubs;
- we can fix an orientation of C such that the following holds. Imagine starting at the 1-hub and walking clockwise around C. We should meet our first 2-hub after walking a distance of 4. Then, the second 2-hub should appear at distance 4 or 6 after the first. The third 2-hub should appear at distance 4 or 6 after the second, the fourth should appear at distance 4 or 6 after the third, and so on. (This freedom between 4 and 6 at each step is crucial; it ensures that there are many different nice graphs.)

See Figure 1 for an illustration of a $(46, 3)$-nice graph. We simply say that a graph is *nice* if it is (ℓ, k)-nice for some k, ℓ (satisfying $k \geq 1$ and $\ell \geq \max(12k, 15)$). We remark that the restriction $\ell \geq 12k$ is to ensure that all 2-hubs are closer to the 1-hub in the clockwise direction than the counterclockwise direction.

LEMMA 2.3 *Every nice graph is determined by its Laplacian spectrum.*

We will prove Lemma 2.3 in full detail in Section 4. As a brief outline: the first step in the proof of Lemma 2.3 is to prove that any graph G' with the same Laplacian spectrum as a nice graph G is itself nice (with the same parameters ℓ, k). This 'localizes' the problem: if we only have to consider nice graphs, we can give a much more explicit combinatorial meaning to certain spectral statistics (most crucially, we can give a combinatorial interpretation of the *Laplacian spectral moments* in terms of closed walks around the unique cycle C). This localization step crucially uses the matrix-tree theorem to show that G' is connected (once we know that G' is connected, certain spectral inequalities on various degree statistics allow us to deduce that G' has a single cycle, then that it is sun-like and then that it is nice). We remark that similar ideas were previously used by Boulet [4] to prove that so-called 'sun graphs' are determined by their Laplacian spectrum.

After localizing the problem, the second step is to show how to 'decode' a specific nice graph using spectral information: i.e., assuming that G' is nice, we use spectral information to discover which nice graph it is. The idea for this step is to 'inductively explore the graph around its 1-hub' using spectral moments: assuming we know the positions of all the 2-hubs up to distance d of the 1-hub, we can use the $(2d + 2)$th spectral moment to see whether there is a 2-hub at distance $d + 1$ from the 1-hub. Very roughly speaking, the reason this is possible is that the spectral moments can be interpreted as certain weighted sums over closed walks on C. If a closed walk 'interacts with 2-hubs' i times, then the weight of the walk is divisible by 2, so parity considerations allow us to distinguish closed walks involving the 1-hub from closed walks which only involve 2-hubs.

REMARK 2.4 For this 'decoding' step, there is no advantage of the Laplacian spectrum over the adjacency spectrum. In fact, it would have been much more convenient to work with the adjacency spectrum, as the spectral moments of the adjacency matrix have a much more direct combinatorial interpretation than the spectral moments of the Laplacian matrix. Indeed, the ith spectral moment of the adjacency matrix simply counts the number of closed walks of length i. For a nice graph, every nontrivial closed walk can be obtained by starting with a closed walk in the unique cycle C and then choosing some hubs in the walk at which we go in and out of a leaf. Every time we go in and out of a leaf at a 2-hub, we have an even
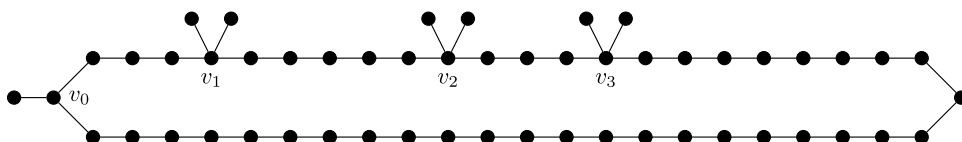
Figure 1. An example of a $(46, 3)$-nice graph. There is one 1-hub v_0 and three 2-hubs v_1, v_2 and v_3. The distances between v_0 and v_1, between v_1 and v_2 and between v_2 and v_3 are 4, 6 and 4, respectively.

number of choices, whereas every time we go in and out of a leaf at a 1-hub, we have an odd number of choices.

REMARK 2.5 There are some parallels between our 2-step strategy to prove Lemma 2.3 and a similar 2-step strategy that was recently applied with great success in the *continuous* case (i.e., in the 'hearing the shape of a drum' setting). Indeed, a recent breakthrough result of Hezari and Zelditch [24] is that ellipses with low eccentricity are determined by their spectrum. In their proof, the first step is to use certain spectral inequalities to 'localize' the problem, showing that any domain whose spectrum matches a low-eccentricity ellipse must be 'almost circular'. Then, the second step is to pin down the precise shape of the domain, taking advantage of the fact that the spectrum determines certain information about *closed billiard trajectories* inside the domain, and applying powerful results due to Avila, De Simoi and Kaloshin [2] to study such trajectories. There is some similarity between closed walks in graphs and closed billiard trajectories in a domain; it is not clear to us whether this connection runs deeper.

2.2. The signless Laplacian spectrum

As outlined, the next step is to prove an analogue of Lemma 2.3 for the signless Laplacian spectrum: we are able to do this with a mild condition on the length of the cycle ℓ, as follows.

LEMMA 2.6 *Let G be an (ℓ, k)-nice graph with $\ell \equiv 2 \pmod 4$. Then, G is determined by its signless Laplacian spectrum.*

Note that an (ℓ, k)-nice graph is bipartite if and only if ℓ is even, and as we have discussed, for bipartite graphs, the signless Laplacian spectrum is the same as the Laplacian spectrum. So, given Lemma 2.3, in order to prove Lemma 2.6, we just need to show that if $\ell \equiv 2 \pmod 4$, then every graph with the same signless Laplacian spectrum as an (ℓ, k)-nice graph must be bipartite.

The full details of the proof of Lemma 2.6 appear in Section 5, but to give a brief idea: the only spectral information we need is the product of nonzero eigenvalues. We observe that for every non-bipartite graph the product of nonzero eigenvalues is divisible by 4 and that the assumption $\ell \equiv 2 \pmod 4$ guarantees that the product of nonzero eigenvalues of G is *not* divisible by 4. For both of these facts, we use an explicit combinatorial description of the coefficients of the characteristic function of the signless Laplacian matrix, due to Cvetković, Rowlinson and Simić [10]. (These coefficients can be expressed as sums of products of eigenvalues via Vieta's formulas; in particular the nonzero coefficient with lowest degree tells us the product of nonzero eigenvalues.)

REMARK 2.7 Lemma 2.6 implies that if n is odd, then there are exponentially many n-vertex graphs which are determined by their signless Laplacian spectrum. However, there is no bipartite nice graph on an even number of vertices, so the analogous result for even n is not completely obvious. With a bit more work, we were nonetheless able to prove such a result, yielding a version of Theorem 1.4 for the signless Laplacian, as follows.

THEOREM 2.8 *The number of (unlabelled) n-vertex graphs determined by their signless Laplacian spectrum is at least e^{cn} for some constant $c > 0$.*

To prove Theorem 2.8, we combine Lemma 2.6 with some of the ideas described in the next subsection; the details appear in the appendix of the arXiv version of this paper.

2.3. Exponentially many graphs determined by their adjacency spectrum

As briefly mentioned earlier in this outline, there is a close connection between the signless Laplacian spectrum of a graph and the adjacency matrix of its line graph. To be a bit more specific, the nonzero eigenvalues of $|L(G)|$ are in correspondence with the eigenvalues of $A(\text{line}(G))$ different from -2.

One might (naively) hope that $\mathrm{line}(G)$ being determined by its adjacency spectrum is equivalent to G being determined by its signless Laplacian spectrum. If this were true, it would be easy to complete the proof of Theorem 1.4, by considering the family of all n-vertex graphs which are the line graph of some nice graph as in Lemma 2.6.

Unfortunately, this is too much to hope for in general, but quite some theory has been developed in this direction, and we are able to leverage this theory in the special case where G has a large prime number of vertices.

> LEMMA 2.9 *There is a constant n_0 such that the following holds. Let G be an (ℓ, k)-nice graph with $\ell \equiv 2 \pmod 4$, let $n = \ell + 2k + 1$ be its number of vertices, and suppose that n is a prime number larger than n_0. Then, $\mathrm{line}(G)$ is determined by its adjacency spectrum.*

The proof of Lemma 2.9 appears in Section 6. To give a rough idea of the strategy of the proof: recalling Lemma 2.6, in order to prove Lemma 2.9, it suffices to show that if a graph Q has the same (adjacency) spectrum as $\mathrm{line}(G)$, then

(1) $Q = \mathrm{line}(H)$ for some H, and
(2) H has the same signless Laplacian spectrum as G.

For (1), we have the Cameron–Goethals–Seidel–Shult theorem at our disposal, which we can use to show that Q is a so-called *generalized line graph* (except possibly for some 'exceptional' connected components with at most 36 vertices). Our main task is to rule out generalized line graphs which are not line graphs. For (2), our task is to show that in the signless Laplacian spectra of G and H, the multiplicities of the zero eigenvalue are the same (all nonzero eigenvalues are guaranteed to be the same). This amounts to showing that G and H have the same number of vertices.

For the first of these two tasks, we observe that if a generalized line graph is not a true line graph, then its adjacency matrix has a zero eigenvalue. So, it suffices to prove that $\mathrm{line}(G)$ does not have a zero eigenvalue, i.e., its adjacency matrix has nonzero determinant. We accomplish this by directly computing the determinant of $\mathrm{line}(G)$ (this is a little involved but comes down to a certain recurrence).

For the second of these two tasks, we recall that the adjacency spectrum of a line graph tells us the nonzero eigenvalues of the signless Laplacian spectrum and, in particular, tells us the product of these nonzero eigenvalues (this product was already discussed in Section 2.2). Via a direct computation on G, we observe that this product is divisible by n. For each connected component of H, the contribution to this product is always an integer, so if n is a prime number then there must be a single connected component which is 'responsible for the factor of n'. We are then able to deduce that this component has exactly n vertices and n edges, via a careful case analysis involving a combinatorial interpretation of the multiplicity of the eigenvalue -2.

Of course, even after proving Lemma 2.9, we are not yet done: every nice graph has the same number of edges as vertices, so Lemma 2.9 can only be directly used to prove Theorem 1.4 when n is prime. For general n, we consider graphs with two connected components, one of which is the line graph of a nice graph on a prime number of vertices and the other of which is a complete graph. The parameters of the nice graph and the size of the complete graph need to satisfy certain inequalities and number-theoretic properties; the details are a bit complicated and we defer the precise specification to Section 7.

In order to show that all relevant inequalities and number-theoretic properties can be simultaneously satisfied (by exponentially many graphs), we use a quantitative strengthening of Dirichlet's theorem on primes in arithmetic progressions. To actually show that all these graphs are determined by their adjacency spectrum, we proceed similarly to Lemma 2.9, but the details are more complicated. Roughly speaking, we identify the complete graph component using its single large eigenvalue and some number-theoretic considerations, and then we apply Lemma 2.9.

3. PRELIMINARIES

In this section, we collect a number of general tools and results that will be used throughout the paper. Where possible, we cite the original sources of each of these results, but we remark that many of these results can be found together in certain monographs on algebraic graph theory or graph spectra (see [3, 5, 9, 11, 16]).

3.1. Basic observations

First, in Section 2, we have already mentioned that the signless Laplacian and the Laplacian spectra coincide for bipartite graphs.

FACT 3.1 ([33, Section 2.3]) *If a graph is bipartite, then its signless Laplacian spectrum is the same as its Laplacian spectrum.*

Also, we record the near-trivial fact that for all notions of spectrum discussed so far, the spectrum of a graph can be broken down into the spectra of its connected components.

FACT 3.2 *For any graph G, the spectrum of G (with respect to the adjacency, Laplacian or signless Laplacian matrix) is the multiset union of the spectra of the connected components of G.*

3.2. Spectral inequalities

Spectral graph theory provides a range of powerful inequalities on various combinatorial parameters, usually in terms of the largest, second-largest or smallest eigenvalue of the adjacency or Laplacian matrix. In this paper, we will only need some simple inequalities concerning the numbers of vertices and edges and the degrees.

LEMMA 3.3 ([11, Section 3.2] and [37]) *Consider a graph G with n vertices, m edges and maximum degree Δ. Let $\lambda_{\max}$ be the largest eigenvalue of the adjacency matrix $\mathrm{A}(G)$. Then*

(1) $\lambda_{\max} \leq \Delta$,
(2) $\lambda_{\max} \leq \sqrt{2m - n + 1}$.

LEMMA 3.4 ([13, Theorem 3.7] and [1, Theorem 2]) *Let G be a graph, write V and E for its sets of vertices and edges and Δ for its maximum degree. Let $\rho_{\max}$ be the largest eigenvalue of the Laplacian matrix $\mathrm{L}(G)$. Then*

(1) $\rho_{\max} > \Delta$,
(2) $\rho_{\max} \leq \max\{\deg(u) + \deg(v) : uv \in E\}$.

3.3. Combinatorial interpretation of the spectral moments

As briefly mentioned in Section 2, in this paper, we use the term *spectral moments* to refer to sums of powers of eigenvalues.

DEFINITION 3.5 *For a matrix $M \in \mathbb{R}^{n \times n}$ with spectrum σ, the sth spectral moment of M is*

$$\sum_{\lambda \in \sigma} \lambda^s = \mathrm{trace}(M^s) = \sum_{i_1=1}^{n} \cdots \sum_{i_s=1}^{n} M_{i_1,i_2} M_{i_2,i_3} M_{i_3,i_4} \cdots M_{i_{s-1},i_s} M_{i_s,i_1}.$$

If M is the adjacency matrix of a graph G, then the product $M_{i_1,i_2} M_{i_2,i_3} M_{i_3,i_4} \cdots M_{i_{s-1},i_s} M_{i_s,i_1}$ is nonzero if and only if there is a *closed walk* in G running through the vertices indexed by $i_1, \ldots, i_s$ (in which case this product is exactly 1). So, spectral moments simply count closed walks of various lengths. For example, the second spectral moment is the number of closed walks of length 2, which is precisely

twice the number of edges in G (a closed walk of length 2 simply runs back and forth along an edge, starting at one of its two endpoints).

In our proof of Lemma 2.3, we will need to carefully study *Laplacian* spectral moments, which can also be interpreted in combinatorial terms (albeit in a more complicated way):

DEFINITION 3.6 An *s-route* in a graph is a sequence of vertices $\vec{v} = (v_1, \ldots, v_s)$, such that for each index j, either $v_j v_{j+1}$ is an edge or $v_j = v_{j+1}$ (where the subscripts should be interpreted modulo s). That is to say, a route consists of a sequence of s steps: at each step we may either walk along an edge or wait at the current vertex. Letting t be the number of 'waiting steps' in the s-route $\vec{v}$, we also define $w(\vec{v})$ to be the product of $\deg(v_j)$ over all waiting steps j, times $(-1)^{s-t}$.

FACT 3.7 *For any graph G, let $\mathcal{R}_s$ be the set of all s-routes in G. Then, the sth spectral moments of $L(G)$ and $|L(G)|$ are*

$$\sum_{\vec{v} \in \mathcal{R}_s} w(\vec{v}) \quad \text{and} \quad \sum_{\vec{v} \in \mathcal{R}_s} |w(\vec{v})|,$$

respectively.

We will repeatedly use Fact 3.7 (for many different s) in our proof of Lemma 2.3. For now, we just record some simple observations for $s \le 3$, which can be straightforwardly proved by considering all possible cases for a route of length s.

PROPOSITION 3.8 *Consider any graph G with n vertices and m edges, and write V for its set of vertices. Let $M = L(G)$ or $M = |L(G)|$, and let μ_s be the sth spectral moment of M. Then*

(1) $\mu_0 = n$;
(2) $\mu_1 = \sum_{v \in V} \deg(v) = 2m$;
(3) $\mu_2 = \sum_{v \in V} \deg(v)^2 + 2m$;
(4) If G has no triangles, then $\mu_3 = \sum_{v \in V} \deg(v)^3 + 3 \sum_{v \in V} \deg(v)^2$.

In particular, if we know that G has no triangles, then the spectrum of M is enough information to determine $\sum_{v \in V} \deg(v)^s$ for $s \in \{0, 1, 2, 3\}$.

3.4. Combinatorial interpretation of the characteristic coefficients

In addition to spectral moments, another very rich way to extract combinatorial structure from the spectrum is to consider the *coefficients of the characteristic polynomial* of our matrix of interest.

DEFINITION 3.9 Consider a matrix $M \in \mathbb{R}^{n \times n}$ with spectrum σ, and write its characteristic polynomial $\det(xI - M) = \prod_{\lambda \in \sigma}(x - \lambda) \in \mathbb{R}[x]$ in the form $\sum_{i=0}^{n}(-1)^i \zeta_i x^{n-i}$. Then, we define the *ith characteristic coefficient* to be

$$\zeta_i = \sum_{\Lambda \subseteq \sigma : |\Lambda| = i} \prod_{\lambda \in \Lambda} \lambda$$

(here we have used Vieta's formulas for the coefficients of a polynomial in terms of its roots).

Note that the nth characteristic coefficient ζ_n is the determinant of M. More generally, if we consider the largest s for which ζ_s is nonzero, then ζ_s is the product of nonzero eigenvalues of M. Recalling the definition $\det(xI - M)$ of the characteristic polynomial, we also have the following observation.

FACT 3.10 *If M is an integer matrix, then its characteristic coefficients are all integers.*

Now, the characteristic coefficients of the Laplacian, signless Laplacian and adjacency matrices all have different combinatorial interpretations, as follows.

DEFINITION 3.11 A connected graph is *unicyclic* if it has exactly one cycle (equivalently, if it has the same number of edges as vertices). If the length of this cycle is even it is *even-unicyclic*; otherwise, it is *odd-unicyclic*. Now, consider any graph G.

(1) A *spanning forest F* in G is a subgraph of G which is spanning (i.e., contains all the vertices of G) and whose connected components are trees. (Some authors define a spanning forest of G to have the same number of conected components as G. Here we have no such requirement.) Let $\alpha(F)$ be the product of the numbers of vertices in these trees.

(2) A *TU-subgraph H* of G is a spanning subgraph whose connected components are trees or odd-unicyclic. Generalizing the definition of α above, let $\alpha(H) = 4^c \prod_{i=1}^{s} n_s$, where c is the number of odd-unicyclic components in H, and the numbers of vertices in the tree components are $n_1, \dots, n_s$.

(3) An *elementary subgraph X* of G is a (not necessarily spanning) subgraph whose connected components are cycles and individual edges. Let $\beta(X) = (-1)^c(-2)^d$, where c and d are the number of edge-components and cycle-components in X, respectively.

Let $\Phi_i(G)$, $\Psi_i(G)$ and $\Xi_i(G)$ be the sets of spanning forests with i edges, TU-subgraphs with i edges and elementary subgraphs with i vertices, respectively, in G.

THEOREM 3.12 ([**3**, Theorem 7.5], [**10**, Theorem 4.4] and [**22**, Theorem 3]) *For any graph G, the i and* $A(G)$ *are*

$$\sum_{F \in \Phi_i(G)} \alpha(F), \quad \sum_{H \in \Psi_i(G)} \alpha(H) \text{ and } \sum_{X \in \Phi_i(G)} (-1)^i \beta(X),$$

respectively.

An immediate corollary (in the Laplacian case, considering the nth and $(n-1)$th characteristic coefficients) is Kirchhoff's celebrated *matrix-tree theorem*, as follows.

THEOREM 3.13 ([**27**]) *For any n-vertex graph G, the Laplacian* $L(G)$ *has a zero eigenvalue with multiplicity at least 1. G is connected if and only if the multiplicity of the zero eigenvalue is exactly 1, in which case the number of spanning trees in G is precisely the product of the nonzero eigenvalues divided by n.*

(One does not really need to make a connectedness case distinction here. Indeed, the matrix-tree theorem can be formulated as the statement that the number of spanning trees is equal to any cofactor of the Laplacian matrix; this number may be zero.)

Another corollary is as follows. (A very similar observation appears as [**10**, Proposition 2.1].)

PROPOSITION 3.14 *For any connected graph G:*

(1) If G is bipartite, then $|L(G)|$ *has a zero eigenvalue with multiplicity 1.*
(2) If G is not bipartite, then the determinant of $|L(G)|$ *is a positive integer divisible by 4.*

Proof. Let n be the number of vertices of G, so the determinant of $|L(G)|$ (i.e., its product of eigenvalues) is its nth characteristic coefficient. Note that a tree on at most n vertices has at most $n - 1$ edges, so in the description in Theorem 3.12, the only possible contributions to the nth characteristic coefficient of $|L(G)|$ come from spanning odd-unicyclic subgraphs. If G is bipartite, then clearly there is no such subgraph. On the other hand, if G is not bipartite then it has an odd cycle, and a suitable spanning odd-unicyclic subgraph can be found by iteratively removing edges outside this cycle. Each spanning odd-unicyclic subgraph H has $\alpha(H) = 4$.

Since every connected graph has a spanning tree, the $(n - 1)$th characteristic coefficient of G is always nonzero (so zero can never be an eigenvalue with multiplicity more than 1). $\qquad\square$

3.5. Line graphs

In Section 2, we mentioned a correspondence between the Laplacian spectrum of a graph G and the adjacency spectrum of its line graph $\text{line}(G)$. To elaborate on this: for a graph with vertices $v_1, \ldots, v_n$ and edges $e_1, \ldots, e_m$, consider the *incidence matrix* $N(G) \in \{0, 1\}^{n \times m}$, where the (i, j)-entry is 1 if and only if $v_i \in e_j$. Then, it is not hard to see that $|L(G)| = N(G)N(G)^T$ and $A(\text{line}(G)) = N(G)^T N(G) - 2I$. Since the nonzero eigenvalues of $N(G)N(G)^T$ are the same as the nonzero eigenvalues of $N(G)^T N(G)$ (including multiplicities), we have the following.

> PROPOSITION 3.15 *Consider any graph G and any $\lambda \neq 0$. Then, λ is an eigenvalue of $|L(G)|$ with multiplicity m if and only if $\lambda - 2$ is an eigenvalue of $A(\text{line}(G))$ with multiplicity m.*

If we know the signless Laplacian spectrum of a graph G, then Proposition 3.15 tells us the spectrum of $A(\text{line}(G))$, except the multiplicity of the eigenvalue -2. In order to determine this multiplicity, we just need to know the sum of multiplicities of all eigenvalues of $\text{line}(G)$, i.e., the number of vertices of $\text{line}(G)$, i.e., the number of edges of G. We have already seen that this information can be recovered from the signless Laplacian spectrum (Proposition 3.8(2)). So, the signless Laplacian spectrum of G fully determines the adjacency spectrum of $\text{line}(G)$. Unfortunately, as discussed in Section 2, it is not quite so easy to go in the other direction: there are examples of line graphs which share their adjacency spectrum with non-line-graphs, and there are examples of graphs G, G' which have different numbers of vertices (therefore different signless Laplacian spectra) but for which $\text{line}(G)$ and $\text{line}(G')$ have the same adjacency spectrum.

In this subsection, we collect a few results related to Proposition 3.15. First, $|L(G)| = N(G)N(G)^T$ is a positive semidefinite matrix, so we have the following corollary of Proposition 3.15.

> FACT 3.16 *For any graph G, the eigenvalues of $A(\text{line}(G))$ are all at least -2.*

Also, Proposition 3.14 gives us a combinatorial description of the multiplicity of the zero eigenvalue of G. Together with Proposition 3.15, this can be used to give a combinatorial description of the multiplicity of -2 as an eigenvalue of $A(\text{line}(G))$.

> LEMMA 3.17 ([9, Theorem 2.2.4]) *Let H be a connected graph with v vertices and e edges, and let μ_{-2} be the multiplicity of the eigenvalue -2 in $A(\text{line}(H))$. Then*
>
> $$\mu_{-2} = \begin{cases} e - v + 1 & \text{if } H \text{ is bipartite,} \\ e - v & \text{if } H \text{ is not bipartite.} \end{cases}$$

Finally, we state the Cameron–Goethals–Seidel–Shult theorem mentioned in Section 2: all but finitely many connected graphs which share their adjacency spectrum with a line graph are so-called *generalized line graphs*.

DEFINITION 3.18 Let K_n be the complete graph on n vertices. A *perfect matching* in K_{2m} is a collection of m disjoint edges (covering all the vertices of K_{2m}). The *cocktailparty graph* $\mathrm{CP}(m)$ is the graph obtained from K_{2m} by removing a perfect matching.

For a graph G with vertices $v_1, \dots, v_n$ and nonnegative integers $a_1, \dots, a_n$, the *generalized line graph* $\mathrm{line}(G; a_1, \dots, a_n)$ is defined as follows. First, consider the disjoint union of the graphs

$$\mathrm{line}(G),\ \mathrm{CP}(a_1), \dots,\ \mathrm{CP}(a_n).$$

(i.e., we include each of the above graphs as a separate connected component). Then, for each i, add all possible edges between the vertices of $\mathrm{CP}(a_i)$ and the vertices of $\mathrm{line}(G)$ corresponding to edges of G incident to v_i (this means $2a_i \deg(v_i)$ added edges for each i).

Note that for any graph G we have $\mathrm{line}(G; 0, \dots, 0) = \mathrm{line}(G)$.

THEOREM 3.19 ([**6**, Theorem 4.3 and 4.10]) *Suppose Q is a connected graph on more than 36 vertices, all of whose adjacency eigenvalues are at least -2. Then, Q is a generalized line graph.*

3.6. Primes in arithmetic progressions

As mentioned in Section 2, we will need a quantitative version of Dirichlet's theorem, counting primes in a given arithmetic progression.

THEOREM 3.20 *Fix coprime integers $a, d \geq 1$, and let $\varphi(d) > 0$ be the number of integers up to d which are relatively prime to d. Let $\pi_{a,d}(n)$ be the number of primes up to n which are congruent to $a \pmod{d}$. Then*

$$\lim_{n \to \infty} \left(\frac{\pi_{a,d}(n)}{n / \log n} \right) = \frac{1}{\varphi(d)}.$$

Theorem 3.20 was first proved by de la Vallée Poussin [**12**]. All we will need from Theorem 3.20 is the following (immediate) corollary.

COROLLARY 3.21 Fix $\varepsilon > 0$ and coprime integers $a, d \geq 1$. For any sufficiently large n, there is a prime number between $(1 - \varepsilon)n$ and $(1 + \varepsilon)n$ which is congruent to $a \pmod{d}$.

4. DISTINGUISHING NICE GRAPHS BY THEIR LAPLACIAN SPECTRUM

In this section, we prove Lemma 2.3: nice graphs are determined by their Laplacian spectrum. As discussed in Section 2.1, the first step is to 'localize' the problem, showing that any graph with the same Laplacian spectrum as a nice graph is itself nice. First, we adapt some ideas of Boulet [**4**, Theorem 9] to prove the following lemma, which provides some approximate structure (though does not yet completely determine niceness). Recall the definition of a sun-like graph from Definition 2.2.

LEMMA 4.1 *Let G be an (ℓ, k)-nice graph, and let H be a graph with the same Laplacian spectrum as G. Then H is a sun-like graph whose cycle has length ℓ. Moreover, H has exactly one 1-hub, k different 2-hubs, and no i-hubs for any $i > 2$.*

Proof. Let $n = \ell + 2k + 1$ be the number of vertices and edges in G. First of all, by Proposition 3.8(1) and (2), H also has n vertices and n edges, and by Kirchhoff's matrix-tree theorem (Theorem 3.13), H is connected. So, H is unicyclic. In a unicyclic graph, the number of spanning trees is equal to the length of the cycle, so by Kirchhoff's theorem again, the cycle in H has length ℓ.

Next, we study the degrees of vertices of H. Writing E for the set of edges of G, recall from Lemma 3.4(2) that the largest Laplacian eigenvalue $\rho_{\max}$ is at most $\max\{\deg(u) + \deg(v),\ uv \in E\} \le 6$ (in a nice graph, every hub has degree at most 4, every non-hub has degree at most 2, and no two hubs are adjacent). By Lemma 3.4(1), the maximum degree of H is strictly less than $\rho_{\max}$, so H can only have vertices of degree 1, 2, 3, 4 or 5.

Let n_i be the number of vertices of degree i in H. Since the definition of a nice graph includes the assumption that $\ell > 12k \ge 3$, there are no triangles in H, so by Proposition 3.8, the Laplacian spectrum determines the number of vertices, the sum of degrees, the sum of squares of degrees and the sum of cubes of degrees. In G, the numbers of vertices with degrees 1, 2, 3 and 4 are $2k + 1$, $\ell - k - 1$, 1 and k, respectively, so we have

$$n_1 + n_2 + n_3 + n_4 + n_5 = n = \ell + 2k + 1, \tag{4.1}$$

$$n_1 + 2n_2 + 3n_3 + 4n_4 + 5n_5 = 2n = 2\ell + 4k + 2,$$

$$n_1 + 4n_2 + 9n_3 + 16n_4 + 25n_5 = (2k+1) + 4(\ell - k - 1) + 9 + 16k = 4\ell + 14k + 6,$$

$$n_1 + 8n_2 + 27n_3 + 64n_4 + 125n_5 = (2k+1) + 8(\ell - k - 1) + 27 + 64k = 8\ell + 58k + 20.$$

This system of equations has a one-parameter family of solutions, given by

$$n_2 = -4n_1 + \ell + 7k + 3$$
$$n_3 = 6n_1 - 12k - 5$$
$$n_4 = -4n_1 + 9k + 4$$
$$n_5 = n_1 - 2k - 1. \tag{4.2}$$

Equations (4.1) and (4.2) together imply that $n_2 + n_3 + n_4 + n_5 = \ell - n_5$ (i.e., there are $\ell - n_5 \le \ell$ vertices with degree at least 2). But H has a cycle of length ℓ, and all the vertices on that cycle have degree at least 2, so we must have $n_5 = 0$ and all the vertices with degree at least 2 must lie on the cycle. This implies that H is sun-like.

There was only one degree of freedom in our system of equations: knowing that $n_5 = 0$ allows us to deduce the values of all n_i, and in particular $n_3 = 1$ and $n_4 = k$. That is to say, there is one 1-hub, k different 2-hubs and no i-hubs for $i > 2$, as desired. □

4.1. Decorated routes

Recall the definition of a *route* from Definition 3.6. The remainder of the proof of Lemma 2.3 proceeds by carefully studying routes in sun-like graphs. In this subsection, we introduce a convenient framework for working with such routes.

DEFINITION 4.2 Let G be a sun-like graph. A *decorated s-route* R consists of a route $\vec{v} = (v_1, \ldots, v_s)$ together with a label 'look' or 'wait' assigned to each j for which $v_j = v_{j+1}$ and v_j is a hub (here arithmetic is mod s). That is to say, recalling that we previously imagined a route $\vec{v}$ as a closed walk with some 'waiting steps', we are now reinterpreting some of the waiting steps as steps where we 'look at a hub'.

For a hub v, if $v_j = v_{j+1} = v$ and j has the label 'look', or if $v_j = v$ and v_{j+1} is one of the leaves attached to j, then we say that the decorated route *interacts with* v at step j (i.e., interacting with a hub means looking at it or entering one of the leaves attached to it).

For a decorated s-route R, define its *multiplicity* $\text{mult}(R)$ to be the number of different decorated routes that can be obtained by cyclically shifting or reversing R. For example, if R is a trivial route that repeatedly waits at a single vertex, then $\text{mult}(R) = 1$, but in general $\text{mult}(R)$ can be as large as $2s$.

Consider a decorated s-route R. Suppose that in this decorated route there are r_1 steps where we wait at leaf vertices, and r_2 steps where we wait at cycle vertices (not counting steps

in which we look at a hub). Suppose that for each i, there are t_i steps where we look at an i-hub. Then, we define the *weight* of R as

$$w(R) = 2^{r_2} \prod_{i=1}^{\infty} i^{t_i}.$$

That is to say, we accumulate a factor of 2 whenever we wait at some vertex on the cycle (not when we wait at a leaf vertex), and we accumulate a factor of i whenever we look at an i-hub.

EXAMPLE 4.3 Recall the $(46, 3)$-nice graph in Figure 1. Write a, b, c for the three vertices between the 1-hub v_0 and the 2-hub v_1, and let x be one of the leaf vertices attached to v_1. Then, an example of a route is

$$\vec{v} = (v_0, v_0, a, b, c, v_1, v_1, x, x, v_1, c, b, c, b, a, v_0).$$

This route has four 'waiting steps' (in the first and last steps we wait at v_0, at the sixth step we wait at v_1 and at the eighth step we wait at x).

In order to make this route into a decorated route, for each of the steps where we wait at a hub (i.e., the first, sixth and last step) we need to decide whether to reinterpret this step as a step where we 'look at the hub'. For example, say we label the first step as 'look' (and the sixth and last steps are labelled as 'wait'). This route interacts with v_0 and v_1, once each (we look at v_0, and enter a leaf attached to v_1). The weight of this decorated route is $2^2 \cdot 1 = 4$ (we wait twice at cycle vertices and look at a 1-hub once).

We then have the following consequence of Fact 3.7.

LEMMA 4.4 *Consider any sun-like graph G whose cycle has length ℓ. For $s < \ell$, let $\mathcal{D}_s$ be the set of all decorated s-routes in G. Then, the sth spectral moment of $\mathrm{L}(G)$ is*

$$\sum_{R \in \mathcal{D}_s} w(R).$$

Proof. In undecorated routes (as in Definition 3.6) we accumulate a factor of $\deg(v) = i + 2$ each time we wait at a 2-hub v. For our decorated routes, we have simply broken this down into 'waiting' and 'looking'; waiting accumulates a factor of 2 (just as it does for a non-hub vertex on the cycle) and looking contributes a factor of i.

Also, recall that in an undecorated route we accumulate a factor of -1 for each step we walk along an edge. We can ignore this factor if we only consider routes less than ℓ: such routes cannot make it all the way around the cycle so must 'retrace their steps' and therefore have an even number of 'walking steps'. □

The reason we have introduced the notion of a decorated route is that if we know the hub distribution of a graph, this is enough information to determine the contribution to the sth spectral moment from routes which interact with at most one hub. (So, we can focus on routes which interact with multiple hubs, which are key to understanding how the hubs are distributed around the cycle.)

LEMMA 4.5 *Let G be a sun-like graph whose cycle has length ℓ, and let k_i be the number of i-hubs in G. Let $\mathcal{D}_s^*$ be the set of decorated s-routes which interact with at most one hub (any number of times). Then, $\sum_{R \in \mathcal{D}_s^*} w(R)$ only depends on ℓ and $(k_i)_{i=1}^{\infty}$.*

Proof. Consider two different graphs G, H with the same statistics ℓ and $(k_i)_{i=1}^{\infty}$. We will show that the sum of weights under consideration is the same with respect to G and H. Roughly

speaking, the key observation will be that for any route involving a single hub in G, we can 'rotate the route around the cycle' to find a corresponding route in H.

The cycles C_G and C_H in G both have the same length ℓ, so we can fix an isomorphism $\phi : C_G \to C_H$. Fixing an orientation of C_H, let $\chi : C_H \to C_H$ be the automorphism that 'rotates one step clockwise around C_H'. Since G, H have the same hub distribution, we can also fix an bijection $\psi : C_G \to C_H$ such that v is an i-hub if and only if $\psi(v)$ is an i-hub. For each v in C_G, there is a unique $j \in \mathbb{Z}/\ell\mathbb{Z}$ such that $\psi(v) = \chi^{(j)}(\phi(v))$ (i.e., we 'make $\psi(v)$ line up with $\phi(v)$' by rotating it j steps around the cycle). Let $\phi_v = \chi^{(j)} \circ \phi$ for this j, so ϕ_v is an isomorphism $C_G \to C_H$ with $\phi_v(v) = \psi(v)$.

- Clearly, ϕ gives us a correspondence between decorated s-routes that do not interact with any hubs in G, and decorated s-routes that don't interact with any hub in H.
- For any i-hub v in C_G, the isomorphism ϕ_v (together with a bijection between the i leaves attached to v in C_G and the i leaves attached to $\psi(v)$ in C_H) gives us a correspondence between decorated s-routes which interact with the single hub v in G and s-routes which interact with the single hub $\psi(v)$ in H.

The above correspondences are weight-preserving, so the desired result follows. □

4.2. Localizing to nice graphs

Our first application of the framework in Section 4.1 is to finish the 'localization step' in the proof of Lemma 2.3: every graph with the same spectrum as a nice graph is itself nice. Given Lemma 4.1, this basically comes down to studying distances between hubs.

> **Lemma 4.6** *Let H be a graph with the same spectrum as an (ℓ, k)-nice graph G. Then H is an (ℓ, k)-nice graph.*

> *Proof.* In this proof, we will omit the word 'decorated' (we will have no reason to consider undecorated routes). First, we apply Lemma 4.1 to see that H is a sun-like graph whose cycle has length ℓ, with one 1-hub, k 2-hubs and no i-hubs for $i > 2$.
>
> Let $\eta_s(H)$ and $\eta_s(G)$ be the sum of weights of s-routes which interact with at least 2 hubs (with respect to H and G, respectively). By Lemmas 4.4 and 4.5 and the fact that $\ell \geq 15$ from Definition 2.2, we have $\eta_s(H) = \eta_s(G)$ for all $s \leq 14$ (so, we mostly just write 'η_s' to indicate this common value).
>
> Now, we use the parameters η_s to study the structure of H. We break this down into a sequence of claims. □

> **Claim 4.7** *In H, the closest pair of hubs is at distance 4.*

> *Proof.* Note that $\eta_{2d+2} > 0$ if and only if there are two hubs whose distance is at most d. Indeed, the shortest way for a route to interact with two hubs is to look at one hub, walk d steps to the next hub, look at it, and walk back; this takes $1 + d + 1 + d = 2d + 2$ steps.
>
> The closest pair of hubs in G are at distance 4, so the same is true in H. (Note that $2 \cdot 4 + 2 \leq 14$.) □

> **Claim 4.8** *In H:*

> *(1) the 1-hub has distance 4 from exactly one other hub, and*
> *(2) the number of pairs of hubs at distance 4 from each other is the same in G and H.*

> *Proof.* The only routes that contribute to η_{10} are those routes which walk back and forth between two different hubs at distance 4, looking once at each hub along the way. Each such

route contributes a weight of 4, unless one of the hubs is a 1-hub, in which case the route contributes a weight of 2. Also, each such route has multiplicity 10 (all 10 cyclic shifts yield different routes, but reversing the order does not yield any further routes).

So, $\eta_{10}/20$ can be interpreted as the number of hubs at distance 4 from the 1-hub plus two times the number of pairs of 2-hubs at distance 4 from each other.

In G, there is exactly one hub at distance 4 from the 1-hub. So, $\eta_{10}(G)/20 = \eta_{10}(H)/20$ is odd, meaning that there must be an odd number of hubs at distance 4 from the 1-hub in H. The only possible odd number here is 1, because there is simply no room to put three or more hubs at distance 4 from the 1-hub.

Then, in H and in G, the number of pairs of 2-hubs at distance 4 from each other is $(\eta_{10}/20 - 1)/2$. □

Now, Claims 4.7 and 4.8 show that the contributions to $\eta_s(H)$ and $\eta_s(G)$ from routes which interact with two hubs within distance at most 4 (and no other hubs) are the same. (Formally, this can be proved in a similar way to Lemma 4.5, considering a bijection between the set of pairs of hubs at distance 4 in G, and the set of pairs of hubs at distance 4 in H). Let $\eta'_s(H)$ and $\eta'_s(G)$ be obtained from $\eta_s(H)$ and $\eta_s(G)$ by subtracting these contributions, so $\eta'_s(H) = \eta'_s(G)$ for $s \leq 14$.

CLAIM 4.9 *In H, there are no hubs at distance 5 from each other.*

Proof. The only routes which can contribute to η'_{12} are routes which interact with two different hubs at distance 5 from each other. (By Claim 4.7, every pair of hubs is at distance at least 4 from each other, so routes of length 12 are much too short to interact with three different hubs.) Since G has no pair of hubs at distance 5, the same is true for H. □

CLAIM 4.10 *In H:*

(1) the 1-hub does not have distance 6 from any other hub, and
(2) the number of pairs of hubs at distance 6 from each other is the same in G and H.

Proof. Given Claim 4.9, the only routes which can contribute to η'_{14} are routes which interact with two different hubs at distance 6. (Routes of length 14 are still too short to interact with three different hubs.)

The same considerations as for Claim 4.8 show that $\eta'_{14}/28$ can be interpreted as the number of hubs at distance 6 from the 1-hub, plus two times the number of pairs of 2-hubs at distance 6 from each other.

In G, there is no hub at distance 6 from the 1-hub. So, $\eta'_{14}(G)/28 = \eta'_{14}(H)/28$ is even, meaning that there are an even number of hubs at distance 6 from the 1-hub v^* in H. The only possible even number here is zero, because if there were two hubs at distance 6 from v^* (one on either side), one of these 2-hubs would be at distance 2 from the hub guaranteed by Claim 4.8(1) at distance 4 from v^*, and this is ruled out by Claim 4.7.

Then, in H and in G, the number of pairs of 2-hubs at distance 6 from each other is $(\eta'_{14}/28)/2$. □

Now, Claims 4.7 to 4.10 together imply that H is a (k, ℓ)-nice graph. Indeed, imagine walking around the cycles of G and H and consider the distances between each pair of consecutive hubs. By Claims 4.7 and 4.9, these distances are either 4 or at least 6. By Claims 4.8(2) and 4.10(2), the number of consecutive pairs of hubs in H which are at distance 4 or 6 is the same as the number of consecutive pairs of hubs in G which are at distance 4 or 6; this number is exactly k. Recalling that H and G both have exactly $k + 1$ hubs, it follows that in H we can start from some hub v_0 and walk along the cycle, encountering a new hub every 4 or 6 steps until we reach a final hub v_k. By Claims 4.8(1) and 4.10(1), the

1-hub is either v_0 or v_k (with distance exactly 4 to its closest 2-hub). We have established that H is (ℓ, k)-nice.

4.3. Decoding a nice graph

Now, we complete the proof of Lemma 2.3, showing that we can decode a specific nice graph using its Laplacian spectrum.

Proof of Lemma 2.3 As in the proof of Lemma 4.6, we will omit the word 'decorated' (we will again have no reason to consider undecorated routes).

Suppose we know that G is an (ℓ, k)-nice graph (for some $k \geq 1$ and $\ell > 12k$), and suppose we know the spectrum of G. We will show how to use this information to determine exactly which (ℓ, k)-nice graph G is (this suffices to prove Lemma 2.3, by Lemma 4.6).

Specifically, it suffices to determine, for each $q \leq 3k - 1$, whether there is a hub at distance $2q$ from the 1-hub v^*. (In a nice graph, every hub is at even distance from v^*, and the furthest possible distance between hubs is $4 + 6(k - 1) = 6k - 2$.)

We proceed by induction. For some $q \leq 3k - 1$, suppose we know the positions of all hubs within distance $2q - 1$ of v^*. We would like to determine whether there is a hub at distance $2q$ from v^*.

Let η_s be the sum of weights of s-routes which interact with at least 2 hubs. By Lemma 4.5, we have enough information to determine η_s for $s < \ell$. We can refine this further: let η_s' be obtained from η_s by subtracting the contribution from all routes which interact only with hubs within distance $2q - 1$ of v^*. Since our inductive assumption is that we know the positions of all hubs within distance $2q - 1$ of v^*, we have enough information to determine η_s' (for $s < \ell$).

We focus in particular on the quantity η_{4q+2}' (note that $4q + 2 \leq 4(3k - 1) + 2 < 12k < \ell$, so we have enough information to determine this quantity). We break down η_{4q+2}' further:

- Let α_{4q+2} be the contribution to η_{4q+2}' from routes which interact with v^*, and
- Let $\beta_{4q+2}(t)$ be the contribution to η_{4q+2}' from routes which do not interact with v^* and interact with 2-hubs t times.

Note that $\eta_{4q+2}' = \alpha_{4q+2} + \sum_{t=2}^{\infty} \beta_{4q+2}(t)$. □

CLAIM 4.11 *We have*

$$\alpha_{4q+2} = \begin{cases} 8q + 4 & \text{if there is a hub at distance } 2q \text{ from } v^* \\ 0 & \text{otherwise} \end{cases}$$

Proof. If there is no hub at distance $2q$ from v^*, then a route of length $4q + 2$ is simply too short to interact with v^* and with one of the 2-hubs that is not within distance $2q - 1$ of v^*.

If there is a hub v at distance $2q$ from v^*, the only routes which contribute to α_{4q+2} are those routes which walk back and forth between v and v^*, looking once at v and v^* along the way. All these routes are cyclic shifts of each other (so, there are $4q + 2$ of them), and each such route contributes a weight of 2. □

CLAIM 4.12 $\beta_{4q+2}(t)$ *is divisible by 8 for all t.*

Proof. Consider a 2-hub v, and write x, y for the leaves attached to v. Consider a route R which at some step j enters x from v (then waits at x for some number of steps before returning to

v). We can slightly modify R by simply entering y instead of x at step j (and then waiting at y for the same number of steps before returning to v).

Say that two routes are *equivalent* if they can be obtained from one another by a sequence of modifications of this type. So, routes in an equivalence class have essentially the same structure, but they may visit different leaves. Now, consider a route R which looks at 2-hubs a times and enters leaves attached to 2-hubs b times (so, R interacts with 2-hubs $a + b$ times). The equivalence class of R has size 2^b, and the weight of each route in this equivalence class is divisible by 2^a. So, the total weight of this equivalence class is divisible by 2^{a+b}.

It immediately follows that $\beta_{4q+2}(t)$ is divisible by 2^t, so if $t \geq 3$ then $\beta_{4q+2}(t)$ is divisible by 8. It remains to consider $\beta_{4q+2}(2)$ in more detail.

The routes that contribute to $\beta_{4q+2}(2)$ are the routes which interact once each with two different 2-hubs u, v (and do not interact with v^*). Fix such a route R. As above, the equivalence class of R contributes weight divisible by 4, so we just need an additional factor of 2. This comes from the fact that $\mathrm{mult}(R)$ is equal to $4q + 2$ or $8q + 4$ (both of which are divisible by 2). Indeed, all $4q + 2$ cyclic shifts of R yield different routes, because there is a unique interaction-with-u step whose position changes with each cyclic shift. Reversing the order of R may or may not yield $4q + 2$ additional routes. □

Finally, given Claims 4.11 and 4.12, we can determine whether there is a 2-hub at distance 2 q from v^* simply by checking whether η'_{4q+2} is divisible by 8 or not. This completes the inductive step.

5. DETERMINING BIPARTITENESS WITH THE SIGNLESS LAPLACIAN SPECTRUM

In this section, we prove Lemma 2.6. This proof mostly comes down to the following two lemmas.

DEFINITION 5.1 For any graph G, let $f_{|L|}(G)$ be the product of nonzero eigenvalues of $|L(G)|$.

LEMMA 5.2 *If G is not bipartite, then $f_{|L|}(G)$ is divisible by 4.*

Proof. Let $G_1, \ldots, G_c$ be the connected components of a non-bipartite graph G, and suppose without loss of generality that G_1 is non-bipartite. By Fact 3.10, each $f_{|L|}(G_i)$ is an integer, and by Fact 3.2 we have $f_{|L|}(G) = f_{|L|}(G_1) \ldots f_{|L|}(G_c)$. By Proposition 3.14(2), $f_{|L|}(G_1)$ is divisible by 4. □

LEMMA 5.3 *If G is a connected bipartite unicyclic graph with n vertices, whose cycle has length ℓ, then $f_{|L|}(G) = n\ell$.*

Proof. Since G is bipartite, its signless Laplacian spectrum is the same as its Laplacian spectrum (by Fact 3.1), so by Kirchhoff's matrix tree theorem (Theorem 3.13), $f_{|L|}(G)$ is n times the number of spanning trees in G (which is ℓ, as we have already observed in the proof of Lemma 4.1). □

Now, we are ready to prove Lemma 2.6.

Proof of Lemma 2.6 Let G be an (ℓ, k) nice graph, for $\ell \equiv 2 \pmod 4$, and let H be a graph with the same signless Laplacian spectrum as G. As discussed in Section 2.2, given Lemma 2.3 and Fact 3.1, we just need to prove that H is bipartite.

Let $n = \ell + 2k + 1$ be the number of vertices in G. By Lemma 5.3, we have $f_{|L|}(G) = n\ell$. Since $\ell \equiv 2 \pmod 4$ and $n = \ell + 2k + 1$ is odd, $f_{|L|}(G)$ is not divisible by 4. Since G and H have the same spectrum, we have $f_{|L|}(G) = f_{|L|}(H)$, so Lemma 5.2 implies that H is bipartite. □

6. THE PRIME CASE OF THE MAIN THEOREM

In this section, we prove Lemma 2.9. As outlined in Section 2.3, we will use the Cameron–Goethals–Seidel–Shult theorem (Theorem 3.19), together with the following fact. Recall the definition of a generalized line graph from Definition 3.18.

LEMMA 6.1 *If a generalized line graph is not a line graph, then its adjacency matrix has a zero eigenvalue.*

Proof. Let G be a generalized line graph that is not a line graph. We will show that G has two vertices with the same set of neighbours, meaning that $A(G)$ has two equal rows so is not invertible and therefore has a zero eigenvalue.

By the definition of a generalized line graph, G contains a cocktailparty graph $CP(a)$ for some $a \geq 1$. This cocktailparty graph can be thought of as a complete graph K_{2a} with a perfect matching removed. Consider one of the edges of this removed perfect matching, and let u and v be its endpoints. Then, u and v have the same neighbourhood (in G), as desired. □

Now, crucially, line graphs of nice graphs as in Lemma 2.9 do not have zero eigenvalues.

LEMMA 6.2 *Let G be an (ℓ,k)-nice graph with $\ell \equiv 2 \pmod 4$. Then, $A(\mathrm{line}(G))$ does not have a zero eigenvalue.*

We will prove Lemma 6.2 by explicitly computing the determinant of $A(\mathrm{line}(G))$ using Theorem 3.12. We defer this computation to Section 6.1, as it is a little involved; first, we show how to use it to prove Lemma 2.9 (after stating a definition that will be used in the proofs of Lemma 2.9 and Theorem 1.4).

DEFINITION 6.3 For any graph G, let $f_A(G)$ be the product of nonzero eigenvalues of $A(G) + 2$. Equivalently, writing σ for the adjacency spectrum of G,

$$f_A(G) = \prod_{\substack{\lambda \in \sigma \\ \lambda \neq -2}} (\lambda + 2).$$

Proof of Lemma 2.9 assuming Lemma 6.2 Consider ℓ, k with $\ell \equiv 2 \pmod 4$, and let $n = \ell + 2k + 1$. Define

$$n_0 = \max\{f_{|A|}(Q) : Q \text{ is a graph on at most 36 vertices}\}, \tag{6.1}$$

and suppose n is a prime number larger than n_0.

Let G be an (ℓ,k)-nice graph, and suppose that Q is a graph with the same adjacency spectrum as $\mathrm{line}(G)$. Our objective is to prove that $Q = \mathrm{line}(H)$ for some graph H with n vertices. Indeed, if we are able to prove this, it will follow from Proposition 3.15 that H has the same nonzero signless Laplacian eigenvalues as G, and since H and G have the same number of vertices, the multiplicity of the zero eigenvalue will also be the same in H and G. It will then follow that H and G are isomorphic (hence Q and $\mathrm{line}(G)$ are isomorphic) by Lemma 2.6.

Write $Q_1, \ldots, Q_c$ for the connected components of Q. By Fact 3.10, each $f_A(Q_i)$ is an integer, and by Fact 3.2, we have $f_A(Q_1) \ldots f_A(Q_c) = f_A(Q)$. On the other hand, by Proposition 3.15 and Lemmas 5.3,

$$f_A(Q) = f_A(\mathrm{line}(G)) = f_{|L|}(G) = n\ell. \tag{6.2}$$

Recalling that n is a prime number, some $f_A(Q_i)$ must be divisible by n. Suppose without loss of generality that

$$f_A(Q_1) \text{ is divisible by } n. \tag{6.3}$$

By the Cameron–Goethals–Seidel–Shult theorem (Theorem 3.19), Lemma 6.2 and our assumption $n > n_0$ from the start of the proof, Q_1 is a line graph. We write $Q_1 = \mathrm{line}(H_1)$ for some (connected) graph H_1, with v_1 vertices and e_1 edges. Note that

$$e_1 \leq n, \tag{6.4}$$

because Q_1 has e_1 vertices and is a connected component of Q, which has n vertices (note that Q has the same number of vertices as $\mathrm{line}(G)$, which is n because G has n edges).

Now, by Proposition 3.15, we have $f_A(Q_1) = f_{|L|}(H_1)$. This cannot be divisible by 4, because $f_A(Q) = n\ell$ is not divisible by 4 (here we are recalling (6.2), and using that n is odd and $\ell \equiv 2 \pmod 4$). So, by Lemma 5.2, H_1 is bipartite.

By Lemma 3.17, $A(Q)$ has -2 as an eigenvalue with multiplicity 1, so (using Fact 3.2), either -2 is not an eigenvalue of $A(Q_1)$ or it is an eigenvalue with multiplicity 1.

Case 1: -2 is not an eigenvalue of $A(Q_1)$. In this case, Lemma 3.17 says that $e_1 = v_1 - 1$, and H_1 is a tree. The largest TU-subgraph of H_1 is H_1 itself, so by Theorem 3.12 and Proposition 3.15 we have $f_A(Q_1) = f_{|L|}(H_1) = v_1$. Then, (6.3) says that v_1 is divisible by n. (6.4) says that $v_1 - 1 \leq n$, so we must have $v_1 = n$. It follows that $Q_1 = \mathrm{line}(H_1)$ has $e_1 = n - 1$ vertices, meaning that Q only has room for one other component Q_2, consisting of a single isolated vertex. We then compute $f_A(Q_2) = 1$, so $f_A(Q) = f_A(Q_1)f_A(Q_2) = v_1 = n$. This is not consistent with the fact that $f_A(Q) = n\ell$ (as we observed in (6.2)), so this case cannot actually occur.

Case 2: -2 is an eigenvalue of $A(Q_1)$. In this case, Lemma 3.17 says that $e_1 = v_1$, and H_1 is an even-unicyclic graph. Let ℓ_1 be the length of the cycle in H_1, so by Lemma 5.3 we have $f_A(Q_1) = f_{|L|}(H_1) = v_1\ell_1$.

By (6.4), we have $\ell_1 \leq v_1 \leq n$, and (6.3) says that $v_1\ell_1$ is divisible by the prime number n. So, we must have $v_1 = n$. Since $Q_1 = \mathrm{line}(H_1)$ has $e_1 = v_1 = n$ vertices, there is no room for any other components: we have proved that $Q = Q_1 = \mathrm{line}(H_1)$ for some H_1 with n vertices, as desired. □

6.1. Computing the determinant of the line graph of a nice graph

In this subsection, we prove Lemma 6.2. First, we need some definitions that allow us to discuss the structure of the line graph of a nice graph.

DEFINITION 6.4 Let uv be an edge in a graph Q. To *add an i-house* to uv is to add a set S of i new vertices to Q, and to add all possible edges between vertices in $S \cup \{u, v\}$. Then, we say that the subgraph induced by $S \cup \{u, v\}$ (which is a complete graph on $i + 2$ vertices) is an *i-house*. The vertices u, v are *internal* and the vertices in S are *external*.

Note that the line graph of every (ℓ, k)-nice graph (as defined in Definition 2.2) can be obtained by starting with a cycle of length ℓ, then adding a 1-house to one edge and adding 2-houses to k other edges. The distances between pairs of consecutive i-houses are always 3 or 5 (except one longer distance around the cycle). See Figure 2 for an illustration.

Now, our objective is to compute the determinant of the line graph of a nice graph. We will be able to reduce this to computing the determinant of a slightly simpler type of graph, which can be studied recursively.

Recall from Definition 3.11 that a spanning elementary subgraph of a graph G is a spanning subgraph (covering all vertices) consisting of vertex-disjoint edges and cycles. For such a subgraph X, recall that $\beta(X)$ accumulates a factor of -1 for each edge-component, and a factor of -2 for each

cycle-component. By Theorem 3.12, the determinant of $A(G)$ is (up to sign) the sum of $\beta(X)$ over all spanning elementary subgraphs X of G.

DEFINITION 6.5 Consider $r \geq 0$ and $k \geq 0$ and $1 \leq a_1 \leq \cdots \leq a_k \leq r$ satisfying $a_i - a_{i-1} \geq 2$ for each $2 \leq i \leq k$. The graph $Q(r; a_1, \ldots, a_k)$ is defined by starting with a path of length r and adding a 2-house on the a_ith edge of this path for each i. (See Figure 3 for an illustration.) Let $q(r; a_1, \ldots, a_k)$ be the sum of $\beta(X)$ over all spanning elementary subgraphs X of $Q(r; a_1, \ldots, a_k)$.

LEMMA 6.6 *Let $r, k, a_1, \ldots, a_k$ be as in Definition 6.5. For inductive reasons, it is convenient to additionally allow $r = -1$ (in which case $Q(r)$ is the graph with no vertices).*

(1) Taking $k = 0$, we have $q(-1) = 1$ and $q(0) = 0$.
(2) If $r - a_k \geq 2$ (or if $r \geq 2$ and $k = 0$), then $q(r; a_1, \ldots, a_k) = -q(r-2; a_1, \ldots, a_k)$.
(3) If $r - a_k = 1$ then $q(r-2; a_1, \ldots, a_{k-1}) + 2q(r-3; a_1, \ldots, a_{k-1})$.
(4) If $r = a_k$ then $q(r; a_1, \ldots, a_{k-1}) = -2q(r-1; a_1, \ldots, a_{k-1}) - 3q(r-2; a_1, \ldots, a_{k-1})$.

Proof. First, (1) is an immediate observation.

If $a_k < r$ (or if $r \geq 1$ and $k = 0$), then the final vertex in $Q(r; a_1, \ldots, a_k)$ has degree 1. In an elementary spanning subgraph, this final vertex can only be contained in an edge-component, consisting of the final two vertices of $Q(a_1, \ldots, a_k)$.

In particular, if $r - a_k \geq 2$ (or if $r \geq 2$ and $k = 0$), the spanning elementary subgraphs of $Q(a_1, \ldots, a_k)$ can be obtained by taking a spanning elementary subgraph of $Q(a_1, \ldots, a_k - 2)$, and adding a single edge-component (see Figure 4). We deduce (2), recalling that each edge-component contributes a weight of -1.

If $r - a_k = 1$, then the aforementioned edge-component covers one of the internal vertices of the final 2-house. There are two different ways to cover the two external vertices in this 2-house by a spanning elementary subgraph: either we can cover them with a single edge or we can cover them, in addition to the remaining internal vertex, with a 3-cycle (see Figure 4). In the first case, we accumulate a factor of -2, and the remaining vertices of the spanning elementary subgraph can be interpreted as a spanning elementary subgraph of $Q(r-2; a_1, \ldots, a_{k-1})$. In the second case, we accumulate a factor of -1, and the remaining vertices of the spanning elementary subgraph can be interpreted as a spanning elementary subgraph of $Q(r-3; a_1, \ldots, a_{k-1})$. So,

$$q(r; a_1, \ldots, a_{k-1}) = (-1)^2 q(r-2; a_1, \ldots, a_{k-1}) + (-1)(-2)q(r-3; a_1, \ldots, a_{k-1}),$$

yielding (3).

If $r = a_k$, then the final vertex of $Q(r; a_1, \ldots, a_k)$ is an internal vertex of the final 2-house, and does not have degree 1. There are a few different ways to cover the final vertex and the two external vertices of the final house by a spanning elementary subgraph: we could cover just these three vertices with a 3-cycle, or we could cover the entire 2-house (there are three different ways to do this with two disjoint edges and three different ways to do this with a
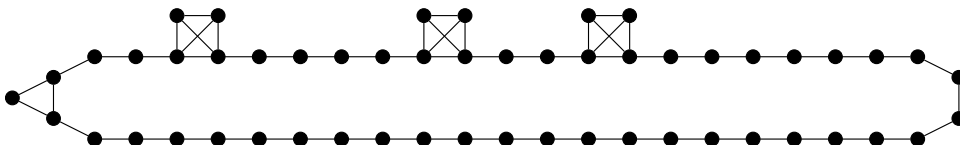
Figure 2. The line graph of the $(46, 3)$-nice graph in Figure 1.
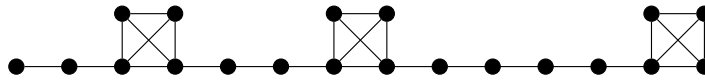
Figure 3. An illustration of the graph $Q(13; 3, 7, 13)$, with 2-houses on the third, seventh and thirteenth edges of the underlying path.
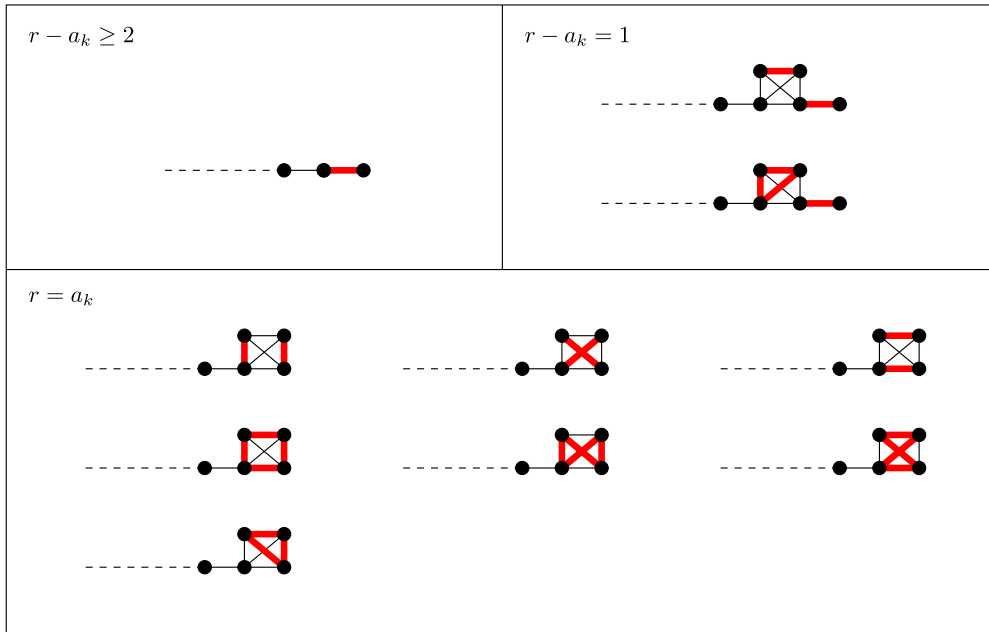
Figure 4. All the possible ways to cover the final vertex (and possibly the external vertices in the final 2-house) in a spanning elementary subgraph of a graph $Q(r; a_1, \ldots, a_k)$.

4-cycle; see Figure 4). Similar considerations as above yield

$$q(r; a_1, \ldots, a_{k-1}) = -2q(r - 1; a_1, \ldots, a_{k-1}) + (3(-1)^2 + 3(-2))q(r - 2; a_1, \ldots, a_{k-1}),$$

yielding (4). □

The recurrences described in Lemma 6.6 are sufficient to compute any $q(r; a_1, \ldots, a_k)$, but the general formulas are rather complicated. We consider a restricted class of choices of $a_1, \ldots, a_k$, which will be sufficient for the proof of Lemma 6.2.

COROLLARY 6.7 Suppose $a_1, \ldots, a_k$ are odd integers. Then

$$q(r; a_1, \ldots, a_k) = \begin{cases} 2k(-1)^{r/2+1} & \text{if } r \text{ is even,} \\ (2k+1)(-1)^{(r+1)/2} & \text{if } r \text{ is odd.} \end{cases}$$

Proof. We proceed by induction on k.

First, iterating Lemma 6.6(2), starting with Lemma 6.6(1), yields

$$q(a_1 - 2) = (-1)^{(a_1-1)/2}, \qquad q(a_1 - 1) = 0.$$

So, Lemma 6.6(3) and (4) give

$$q(a_1+1;a_1) = 2(-1)^{(a_1-1)/2} = 2(-1)^{(a_1+1)/2+1}, \qquad q(a_1;a_1) = -3(-1)^{(a_1-1)/2} = 3(-1)^{(a_1+1)/2},$$

respectively. Iterating Lemma 6.6(2) again yields the desired result for $k = 1$.

Now, consider $k \geq 2$ and assume that the desired statement holds for smaller k. Then, recalling that a_k is odd, our inductive assumption together with Lemma 6.6(3,4) yields

$$\begin{aligned}
q(a_k; a_1, \dots, a_k) &= -2q(a_k - 1; a_1, \dots, a_{k-1}) - 3q(a_k - 2; a_1, \dots, a_{k-1}) \\
&= -2(2k-2)(-1)^{(a_k-1)/2+1} - 3(2k-1)(-1)^{(a_k-1)/2} \\
&= (2k+1)(-1)^{(a_k+1)/2}, \\
q(a_k+1; a_1, \dots, a_k) &= q(a_k - 1; a_1, \dots, a_{k-1}) + 2q(a_k - 2; a_1, \dots, a_{k-1}) \\
&= (2k-2)(-1)^{(a_k-1)/2+1} + 2(2k-1)(-1)^{(a_k-1)/2} \\
&= 2k(-1)^{(a_k+1)/2+1}.
\end{aligned}$$

Iterating Lemma 6.6(2) proves the desired statement. □

Now, we are ready to prove Lemma 6.2.

Proof of Lemma 6.2 Let $b_1 < \cdots < b_k$ be the distances of the 2-hubs from the 1-hub in G (so in particular $b_1 = 4$, and all b_i are even). Let D be the sum of $\beta(X)$ over all spanning elementary subgraphs X of $\mathrm{line}(G)$.

Let u^* be the tip of the 1-house in $\mathrm{line}(G)$. There are four ways for an elementary subgraph to cover u^* (pictured in Figure 5):

(1) u^* could be covered by a long cycle that runs all the way around the nice graph.
(2) u^* could be covered by a 3-cycle covering the entire 1-house.
(3) u^* could be covered by a single edge, whose other vertex is at distance 3 from the 2-house.
(4) u^* could be covered by a single edge, whose other vertex is at distance 4 from the 2-house.

Let D_1, D_2, D_3 aand D_4 be the contributions to D from spanning elementary subgraphs that cover u^* in each of the above four ways (in that order). First, D_2, D_3 and D_4 can be handled with Corollary 6.7, as follows. Recall that $\ell \equiv 2 \pmod 4$.

For D_2: apart from the 3-cycle covering the 1-house, the rest of a spanning elementary subgraph corresponds to a spanning elementary subgraph of $Q(\ell - 3; b_1 - 1, \dots, b_k - 1)$, so

$$D_2 = -2q(\ell - 3; b_1 - 1, \dots, b_k - 1) = -2(2k+1) = -4k - 2. \tag{6.5}$$

For D_3: apart from the edge covering the tip of the 1-house, the rest of a spanning elementary subgraph corresponds to a spanning elementary subgraph of $Q(\ell - 2; b_1 - 1, \dots, b_k - 1)$, so

$$D_3 = -q(\ell - 2; b_1 - 1, \dots, b_k - 1) = -(-2k) = 2k. \tag{6.6}$$

For D_4: apart from the edge covering the tip of the 1-house, the rest of a spanning elementary subgraph corresponds to a spanning elementary subgraph of

$$Q(\ell - 2; 2, b_1, \dots, b_k) \cong Q(\ell - 2; \ell - b_k - 1, \ell - b_{k-1} - 1 \dots, \ell - b_1 - 1)$$

(we can describe the graph in 'two different directions'). Note that $\ell - b_k$ is even (as the difference of two even numbers), so

$$D_4 = -q(\ell - 2; \ell - b_k - 1, \ell - b_{k-1} - 1 \dots, \ell - b_1 - 1) = -(-2k) = 2k. \tag{6.7}$$
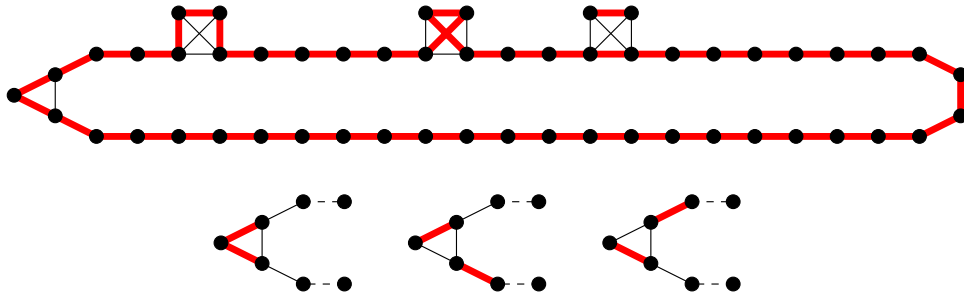
Figure 5. Four possible ways to cover the tip of the 1-house.

It remains to consider D_1. Suppose we have an elementary spanning subgraph which contains a long cycle C covering u^* and going around the ℓ-cycle of $\text{line}(G)$. There are three different ways that C can interact with each 2-house of $\text{line}(G)$ (all of which are pictured at the top of Figure 5). Specifically, there are two ways for C to pass through all 4 vertices of the 2-house, or alternatively C can simply pass through the internal vertices of the 2-house, leaving the remaining two external vertices to be covered by an edge-component.

So, there are 3^k spanning elementary subgraphs that contribute to D_1. To compute the weight of each such subgraph: first, start with a base weight of -2. For each 2-house, we have three choices; the first two (incorporating the 2-house in the cycle) do not affect the weight, but the third (leaving the external vertices for an edge-component) accumulates a factor of -1. So,

$$D_1 = (-2)(1 + 1 - 1)^k = -2. \tag{6.8}$$

Combining (6.5) to (6.8), we see that $D = -4$, so by Theorem 3.12, the determinant of $A(\text{line}(G))$ is nonzero (it has absolute value 4). □

7. AUGMENTING THE PRIME CASE

In this section, we show how to use line graphs of nice graphs to define a family of exponentially many graphs that are determined by their adjacency spectrum. This definition includes a number of inequalities and number-theoretic properties which will be used in a somewhat delicate case analysis in the proof of Theorem 1.4 (to rule out various possibilities for graphs which have the same spectrum as one of our graphs of interest, but have different structure).

DEFINITION 7.1 The *star graph* $K_{1,n}$ consists of n leaves attached to a single vertex. Note that $\text{line}(K_{1,n})$ is the complete graph K_n on n vertices.

Let $\mathcal{G}_n$ be the family of graphs G satisfying the following properties.

G1 G has two components. One of these components is an (ℓ, k)-nice graph G_1 (for some parameters ℓ, k satisfying $\ell \leq \max(12k, 15)$), and the other of these components is a star graph K_{1,n_2} (with some number of edges n_2).

G2 Writing $n_1 = \ell + 2k + 1$ for the number of edges and vertices of G_1, we have $n_1 + n_2 = n$ (i.e., G has n edges).

G3 n_1 is a sufficiently large prime number (larger than n_0 from (6.1)).

G4 $\ell = 2p$ for a sufficiently large prime number p (larger than n_0 from (6.1)).

G5 $n_2 \not\equiv 3 \pmod 4$.

G6 $n_1 < n_2$.

G7 $2n_1 + p - 2 > n$.

G8 $2n_1 - \ell + 2 < n_2 - 1$.

(Note that **G3** and **G4** imply that G_1 satisfies the properties in Lemma 2.9). Let $\mathcal{Q}_n = \mathrm{line}(\mathcal{G}_n)$ be the family of line graphs of graphs in $\mathcal{G}_n$.

Then, the following two lemmas imply Theorem 1.4.

LEMMA 7.2 *There is a constant $c > 0$ such that $|\mathcal{Q}_n| \geq e^{cn}$ for every sufficiently large n.*

LEMMA 7.3 *Every graph in $\mathcal{Q}_n$ is determined by its (adjacency) spectrum.*

It remains to prove these lemmas. First, Lemma 7.2 follows quite simply from Corollary 3.21.

Proof of Lemma 7.2 For sufficiently large n, Corollary 3.21 guarantees the existence of prime numbers p, n_1 such that $n - n_1 \not\equiv 3 \pmod 4$ and such that

$$|n_1 - 0.45n| \leq 0.001n, \quad |p - 0.2n| \leq 0.001n.$$

Let $\ell = 2p$, let $k = (n_1 - \ell - 1)/2$ (which is an integer since n_1 is an odd prime and ℓ is even), and let $n_2 = n - n_1$. Then, it is easy to check that $\ell \leq \max(12k, 15)$ (this is the condition for a nice graph in Definition 2.2), and that **G3** to **G8** all hold. We claim that there are exponentially many graphs in $\mathcal{Q}_n$ with this specific choice of parameters.

To see this, first note that different (ℓ, k)-nice graphs have different line graphs (as depicted in Figure 2, the ℓ-cycle in a nice graph G corresponds to an ℓ-cycle in $\mathrm{line}(G)$, and 1-hubs and 2-hubs in G correspond to 1-houses and 2-houses in $\mathrm{line}(G)$). So, it suffices to prove that there are exponentially many graphs in $\mathcal{G}_n$ with our specific choice of parameters.

An (ℓ, k)-nice graph is specified by a sequence of $k - 1$ binary choices (every pair of consecutive 2-hubs can be at distance 4 or 6). Each of the different ways to make these binary choices lead to different (non-isomorphic) graphs. So, there are 2^{k-1} different (ℓ, k)-nice graphs, meaning that

$$|\mathcal{Q}_n| \geq 2^k \geq 2^{((0.45-0.001)n-2(0.2+0.001)n-1)/2} \geq e^{0.01n}.$$

$\square$

Then, to prove Lemma 7.3, we need a more sophisticated version of the arguments used to prove Lemma 2.9. In particular, we will need the following more detailed version of the case distinction in the proof of Lemma 2.9.

LEMMA 7.4 *Let n_0 be as in (6.1) and let Q be a connected graph with more than n_0 vertices, such that all eigenvalues of $A(Q)$ are at least -2, and such that zero is not an eigenvalue of $A(Q)$. Then, we can write $Q = \mathrm{line}(H)$ for some connected H.*

1. *If -2 is not an eigenvalue of $A(Q)$, then one of the following holds.*
   A. *H is an odd-unicyclic graph, and $f_A(Q) = 4$.*
   B. *H is a tree, and $f_A(Q)$ is the number of vertices of H.*
2. *If -2 is an eigenvalue of $A(Q)$ with multiplicity 1, and if $f_A(Q)$ is not divisible by 8, then H is an even-unicyclic graph (with v vertices and a cycle of length ℓ, say), and $f_A(Q) = v\ell$.*

Proof. The initial part of the lemma (that Q is a line graph) follows from Theorem 3.19, Lemma 6.1. Then, the structural descriptions in **1A** and **1B** follow from Proposition 3.14 (specifically, **A** corresponds to the case where H is not bipartite, and **B** corresponds to the case where H is bipartite), and the statements about $f_A(Q)$ are immediate consequences of Theorem 3.12.

For **2**, we can similarly apply Proposition 3.14, considering the cases where H is or is not bipartite. We see that either H is an even-unicyclic graph (in which case the statement about

$f_A(Q)$ follows from Lemma 5.3) or H is a non-bipartite graph whose number of edges is one more than its number of vertices. We need to rule out this latter case (showing that whenever it occurs, $f_A(Q)$ is divisible by 8).

So, suppose that H is non-bipartite and its number of edges is one more than its number of vertices. Let H' be the *2-core* of H; its largest subgraph with minimum degree at least 2. One can obtain the 2-core by iteratively peeling off leaf vertices (in any order) until no leaves remain. There are two possibilities for the structure of H':

I. H' consists of two edge-disjoint cycles with a single path between them (this path may have length zero), or
II. H' is a 'theta graph', consisting of two vertices with three internally disjoint paths between them.

Case I. In the first case, write C_1 and C_2 for the two cycles, and let ℓ_1 and ℓ_2 be their lengths. For H to be non-bipatitite, at least one of ℓ_1 and ℓ_2 must be odd (suppose without loss of generality that ℓ_1 is odd).

- If ℓ_2 is even, then the largest TU-subgraphs of H are the odd-unicyclic subgraphs obtained by deleting a single edge from C_2. So, by Theorem 3.12, we have $f_{|L|}(H) = 4\ell_2$, which is divisible by 8.
- If ℓ_2 is odd, then the largest TU-subgraphs of H are the odd-unicyclic subgraphs obtained by deleting a single edge from C_1 or C_2 and the disconnected subgraphs (with two odd-unicyclic components) obtained by deleting an edge on the unique path between C_1 and C_2. Writing ℓ_3 for the length of the path between C_1 and C_2, by Theorem 3.12, we have $f_{|L|}(H) = 4(\ell_1 + \ell_2) + 4^2\ell_3$, which is divisible by 8.

Case II. In the second case, write P_1, P_2 and P_3 for the three internally disjoint paths, and let ℓ_1, ℓ_2 and ℓ_3 be their lengths. For H to be non-bipatitite, it cannot be the case that ℓ_1, ℓ_2 and ℓ_3 all have the same parity. Suppose without loss of generality that ℓ_1 is even and ℓ_2 is odd.

- If ℓ_3 is even, then the largest TU-subgraphs of H are the odd-unicyclic subgraphs obtained by deleting a single edge from P_1 or P_3. So, by Theorem 3.12 we have $f_{|L|}(H) = 4(\ell_1 + \ell_3)$, which is divisible by 8.
- If ℓ_3 is odd, then the largest TU-subgraphs of H are the odd-unicyclic subgraphs obtained by deleting a single edge from P_2 or P_3. So, by Theorem 3.12, we have $f_{|L|}(H) = 4(\ell_2 + \ell_3)$, which is divisible by 8. □

We also need the following consequence of Lemma 3.3(2), allowing us to recognize a complete graph by its number of vertices and its largest eigenvalue.

Lemma 7.5 *Let G be a graph with n vertices, such that $A(G)$ has $n - 1$ as an eigenvalue. Then, G is a complete graph.*

Proof. Let e be the number of edges of G, and let $\lambda_{\max}$ be the largest eigenvalue of $A(G)$. Then, Lemma 3.3(2) implies that $n - 1 \leq \lambda_{\max} \leq \sqrt{2e - n + 1}$ or equivalently that $e \geq n(n - 1)/2$; the only graph with this many edges is a complete graph. □

Now, we prove Lemma 7.3, completing the proof of Theorem 1.4.

Proof of Lemma 7.3 Let $G \in \mathcal{G}_n$ (with parameters ℓ, k, n_2, n_1, p as in Definition 7.1), and let Q be a graph with the same adjacency spectrum as $\text{line}(G)$. Our objective is to prove that Q

has the complete graph K_{n_2} as a connected component. Indeed, if we are able to prove this, then we can apply Lemma 2.9 to the graph that remains after removing this K_{n_2} component (here we are using Fact 3.2 to see that removing this K_{n_2} component has a predictable effect on the spectrum).

As is well known (see [5, Section 1.4.1]), the eigenvalues of a complete graph K_{n_2} are -1 (with multiplicity $n_2 - 1$) and $n_2 - 1$ (with multiplicity 1). So (by Fact 3.2), as in the proof of Lemma 2.9, we can see that in the spectrum of $A(Q)$ there is no zero eigenvalue and -2 appears as an eigenvalue with multiplicity 1. Also, by Fact 3.2, Proposition 3.15 and Lemma 5.3, we have

$$f_A(Q) = (n_2 + 1)n_1\ell = 2(n_2 + 1)n_1 p. \tag{7.1}$$

Recalling (7.1) and **G5**, we see that $f_A(Q)$ is not divisible by 8 (so by Fact 3.2, $f_A(Q_i)$ is not divisible by 8 for any connected component Q_i of Q)

Now, Fact 3.2 tells us that some connected component Q_2 of Q must have $n_2 - 1$ as an eigenvalue. Let Δ_2 be the maximum degree of Q_2, so by Lemma 3.3(1) we have

$$\Delta_2 \geq n_2 - 1. \tag{7.2}$$

In particular, Q_2 has at least $\Delta_2 + 1 \geq n_2$ vertices, so by Lemma 7.4 and the assumptions $n_1 > n_2 \geq n_0$ from **G3** and **G6**, we can write $Q_2 = \text{line}(H_2)$ for some graph H_2. Let v_2 be the number of vertices in H_2.

Now, we consider the cases in Lemma 7.4 (**1A**, **1B** and **2**) for the structure of H_2. We will show that all these cases lead to contradiction except **1B** (i.e., H_2 is a tree), and in that case we will prove that $v_2 = n_2 + 1$ vertices (so H_2 has n_2 edges and Q_2 has n_2 vertices; this suffices to show that Q_2 is our desired K_{n_2} component, by Lemma 7.5).

Case 1A: H_2 is odd-unicyclic. In this case, we have $f_A(Q_2) = 4$. Since $f_A(Q)$ is divisible by the prime number n_1, there must be some component $Q_1 \neq Q_2$ such that $f_A(Q_1)$ is divisible by n_1. Recall from (7.1) that $f_A(Q)$ is not divisible by 8, so $f_A(Q_1)$ must be odd.

By Lemma 7.4 (and the assumption $n_1 > n_0$ from **G3**), we can write $Q_1 = \text{line}(H_1)$ for some graph H_1. Let v_1 be the number of vertices in H_1. Considering all cases of Lemma 7.4, the only possibility that leads to $f_A(Q_1)$ being odd is the case where H_1 is a tree (whose number of vertices v_1 is odd and divisible by n_1). Now, we can proceed similarly to **Case 1** in the proof of Lemma 2.9.

Note that Q_1 has $v_1 - 1$ vertices and Q_2 has $v_2 \geq \Delta_2 + 1 \geq n_2$ vertices (for the latter inequality, we used (7.2)). So, $v_1 - 1 + n_2 \leq n$, or equivalently $v_1 \leq n_1 + 1$. Since v_1 is divisible by n_1, we must have $v_1 = n_1$, so Q only has room for one other component Q_3 (other than Q_1 and Q_2), consisting of a single isolated vertex. If this component exists, it has $f_A(Q_3) = 1$. We then compute $f_A(Q) = f_A(Q_1)f_A(Q_2) = 4n_1$, which is not consistent with (7.1). So, this case is impossible.

Case 1B: H_2 is a tree. In this case, $f_A(Q_2) = v_2$. Our objective is to prove that H_2 has $n_2 + 1$ vertices (this suffices, by Lemma 7.5). We need to carefully consider various possibilities for the connected components which are responsible for the large prime factors n_1 and p of $f_A(Q)$. The details will be a bit delicate.

First, note that Q_2 has $v_2 - 1$ vertices; recalling (7.2), we have

$$v_2 - 1 \geq \Delta_2 + 1 \geq n_2. \tag{7.3}$$

Now, suppose that v_2 is divisible by n_1 (we will show that this leads to contradiction). By (7.3) and **G6**, we have $v_2 > n_1 + 1$, so in order for n_1 to divide v_2 we must have $v_2 \geq 2n_1$. It cannot be the case that v_2 is divisible by p as well as n_1 (this would cause v_2 to be far too large, noting that Q_2 has $v_2 - 1 \leq n$ vertices), so there must be some component $Q^* \neq Q_2$ such that $f_A(Q^*)$ is divisible by p. Considering all cases in Lemma 7.4, we see that this is only

possible if Q^* has at least $p-1$ vertices (as the line graph of a graph with at least $p-1$ edges). But then Q^* and Q_2 together have at least $(2n_1 - 1) + (p - 1)$ vertices, which contradicts **G7**.

So, v_2 cannot be divisible by n_1, and there must be some component $Q_1 \neq Q_2$ such that $f_A(Q_1)$ is divisible by n_1. By Lemma 7.4, we can write $Q_1 = \text{line}(H_1)$ for some graph H_1.

Next, suppose that H_1 is a tree (we will show that this leads to contradiction). By (7.3), there are at most n_1 vertices in components other than Q_2. Since v_1 is divisible by n_1, we must have $v_1 = n_1$, meaning that Q_1 has $n_1 - 1$ vertices. So, Q only has room for one other component Q_3 (other than Q_1 and Q_2), consisting of a single isolated vertex, and $f_A(Q) = f_A(Q_1)f_A(Q_2) = n_1 v_2 \leq n_1(n_2 + 2)$ (here we used that Q_2 has at most $n_2 + 1$ vertices, so $v_2 \leq n_2 + 2$). This contradicts (7.1).

So, Q_1 is not the line graph of a tree. Considering all other cases in Lemma 7.4, we see that the only way for $f_A(Q_1)$ to be divisible by n_1 is for Q_1 to have at least n_1 vertices. Recalling (7.3), we deduce that Q_2 has exactly n_2 vertices, as desired.

Case 2: H_2 is even-unicyclic. Let $\ell_2 = 2q$ be the length of the cycle in H_2, so $f_A(Q_2) = v_2 \ell_2$. We will again need to consider various possibilities for the connected components which are responsible for the large prime factors n_1 and p of $f_A(Q)$ (in each case we need to reach a contradiction), but the details will be even more delicate.

- First, suppose that v_2 is divisible by n_1. Note that Q_2 has v_2 vertices. Recalling (7.2) and **G6**, we have $v_2 \geq \Delta_2 + 1 \geq n_2 > n_1$, and we also have $v_2 \leq n < 3n_1$ by **G7**, so in order for n_1 to divide v_2 we must have $v_2 = 2n_1$. We consider possibilities for the prime factor p.
  - Similarly to **Case 1B**, it cannot be the case that v_2 is divisible by p as well as n_1 (this would cause v_2 to be too large).
  - Also, similarly to **Case 1B**, it cannot be the case that there is another component $Q^* \neq Q_2$ such that $f_A(Q^*)$ is divisible by p (then Q^* would have to have at least $p-1$ vertices by Lemma 7.4, and Q^* and Q_2 together would have at least $2n_1 + (p-1)$ vertices, contradicting **G7**).
  - Recalling that $f_A(Q_2) = 2v_2 q$, the remaining case is that q is divisible by p. In this case we have $\ell_2 \geq \ell$, i.e., the cycle in H_2 has length at least ℓ. Each of the v_2 edges in H_2 can be incident to at most two of the edges of this cycle, so $\Delta_2 \leq v_2 - \ell + 2 = 2n_1 - \ell + 2$. But then (7.2) and **G8** are inconsistent with each other.
- So, v_2 is not divisible by n_1. Suppose next that q is divisible by n_1, so the cycle of H_2 has length at least $2n_1$ and $\Delta_2 \leq v_2 - 2n_1 + 2 \leq n - 2n_1 + 2$. But then (7.2) implies $p \leq n_1 \leq n_2 - 1 \leq n - 2n_1 + 2$ (using **G6**), which is inconsistent with **G7**.
- The only remaining possibility is that there is some component $Q_1 \neq Q_2$ such that $f_A(Q_1)$ is divisible by n_1. By Lemma 7.4, Q_1 has at least $n_1 - 1$ vertices, meaning that there are only $n_2 + 1$ vertices left for Q_2. By (7.2), Q_2 must have at least $\Delta_2 + 1 \geq n_2$ vertices.
  - If Q_2 has n_2 vertices, then some vertex in Q_2 is adjacent to all the other vertices in Q_2, meaning that some edge of H_2 is incident to all the other edges in H_2. This is not possible, recalling that H_2 is an even-unicyclic graph.
  - The only other possibility is that H_2 has $n_2 + 1$ vertices, meaning that Q_1 has $n_1 - 1$ vertices (and Q_1 and Q_2 are the only components of Q). This is only possible if H_1 is an n_1-vertex tree, recalling the cases in Lemma 7.4. Then, some vertex in Q_2 is adjacent to all but one of the other vertices in Q_2, meaning that some edge of H_2 is incident to all but one of the other edges in H_2. This can only happen if $\ell_2 = 4$. We deduce that $f_A(Q) = f_A(Q_1)f_A(Q_2) = n_1 \cdot 4(n_2 + 1)$, which is not consistent with (7.1).

$\square$

FUNDING

REFERENCES

1. W.N. Anderson and T.D. Morley, Eigenvalues of the Laplacian of a graph, *Linear Multilinear Algebra* **18** no. 2 (1985), 141–145.

2. A. Avila, J. De Simoi and V. Kaloshin, An integrable deformation of an ellipse of small eccentricity is an ellipse, *Ann. Math. (2)* **184** no. 2 (2016), 527–558. 10.4007/annals.2016.184.2.5

3. N. Biggs, *Algebraic graph theory*, 2nd edn, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1993.

4. R. Boulet, Spectral characterizations of sun graphs and broken sun graphs, *Discret. Math. Theor. Comput. Sci.* **11** no. 2 (2009), 149–160.

5. A.E. Brouwer and W.H. Haemers, *Spectra of Graphs*, Universitext, Springer, New York, 2012.

6. P.J. Cameron, J.M. Goethals, J.J. Seidel and E.E. Shult, Line graphs, root systems, and elliptic geometry, *J. Algebra* **43** no. 1 (1976), 305–327. 10.1016/0021-8693(76)90162-9

7. M., Campos M. Jenssen, M. Michelen and J. Sahasrabudhe, *The singularity probability of a random symmetric matrix is exponentially small*, preprint, arXiv:2105.11384.

8. L. Collatz and U. Sinogowitz, Spektren endlicher grafen, *Abh. Math. Semin. Univ. Hambg.* **21** no. 1 (1957), 63–77. 10.1007/BF02941924

9. D. Cvetković, P. Rowlinson and S. Simić, *Spectral Generalizations of Line Graphs: On Graphs with Least Eigenvalue-2*, London Mathematical Society Lecture Note Series, Vol, 314, Cambridge University Press, Cambridge, 2004.

10. D. Cvetković, P. Rowlinson and S.K. Simić, Signless Laplacians of finite graphs, *Linear Algebra Appl.* **423** no. 1 (2007), 155–171.

11. D.M. Cvetković, M. Doob and H. Sachs, *Spectra of Graphs: Theory and Applications*, 3rd edn, Johann Ambrosius Barth, Heidelberg, 1995.

12. C.-J. de la Vallée Poussin, Recherches analytiques sur la théorie des nombres; premiére partie: La function $\zeta(s)$ de Riemann et les nombres premiers en général, *Ann. Soc. Sci. Bruxelles* **20** (1896), 183–256.

13. M. Fiedler, Algebraic connectivity of graphs, *Czechoslov. Math. J.* **23** no. 2 (1973), 298–305. 10.21136/CMJ.1973.101168

14. D. Garijo, A. Goodall and J. Nešetřil, Distinguishing graphs by their left and right homomorphism profiles, *Eur. J. Combin.* **32** no. 7 (2011), 1025–1053. 10.1016/j.ejc.2011.03.012

15. C.D. Godsil and B.D. McKay, Constructing cospectral graphs, *Aequ. Math..* **25** no. 2-3 (1982), 257–268.

16. C. Godsil and G. Royle, *Algebraic Graph Theory*, Graduate Texts in Mathematics, Vol, 207, Springer, New York, 2001.

17. C. Gordon, D. Webb and S. Wolpert, Isospectral plane domains and surfaces via Riemannian orbifolds, *Invent. Math.* **110** no. 1 (1992), 1–22. 10.1007/BF01231320

18. H.H. Günthard and H. Primas, Zusammenhang von Graphentheorie und MO-Theorie von Molekeln mit Systemen konjugierter Bindungen, *Helv. Chim. Acta* **39** no. 6 (1956), 1645–1653.

19. W.H. Haemers, Are almost all graphs determined by their spectrum? *Not. S. Afr. Math. Soc.* **47** no. 1 (2016), 42–45.

20. W.H. Haemers and E. Spence, Enumeration of cospectral graphs, *Eur. J. Combin.* **25** no. 2 (2004), 199–211. 10.1016/S0195-6698(03)00100-8

21. P.L. Hammer and A.K. Kelmans, *Laplacian spectra and spanning trees of threshold graphs*, First International Colloquium on Graphs and Optimization (GOI), 1992 (Grimentz), Vol, 65, 1996, 255–273.

22. F. Harary, The determinant of the adjacency matrix of a graph, *SIAM Rev.* **4** no. 3 (1962), 202–210. 10.1137/1004057

23. G.H. Hardy and S. Ramanujan, Asymptotic formulae in combinatory analysis, *Proc. Lond. Math. Soc. (2)* **17** no. 1 (1918), 75–115. 10.1112/plms/s2-17.1.75

24. H. Hezari and S. Zelditch, One can hear the shape of ellipses of small eccentricity, *Ann. Math. (2)* **196** no. 3 (2022), 1083–1134. 10.4007/annals.2022.196.3.4

25. A.J. Hoffman, On graphs whose least eigenvalue exceeds $-1 - \sqrt{2}$, *Linear Algebra Appl.* **16** no. 2 (1977), 153–165.

26. M. Kac, Can one hear the shape of a drum? *Am. Math. Monthly* **73** no. 4 Part 2 (1966), 1–23. 10.1080/00029890.1966.11970915

27. G. Kirchhoff, Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird, *Annalen der Physik* **148** no. 12 (1847), 497–508. 10.1002/andp.18471481202

28. M. Noy, *Graphs determined by polynomial invariants*, Random generation of combinatorial objects and bijective combinatorics, Vol, 307, 2003, pp. 365–384.

29. P. Sarnak, Determinants of Laplacians; heights and finiteness, *Analysis, et Cetera*, Academic Press, Boston, MA, 1990, 601–622.

30. A.J. Schwenk, Almost all trees are cospectral, *New Directions in the Theory of Graphs*, Proceedings of the Third Ann Arbor Conference, University of Michigan, Ann Arbor, Michigan, 1971, (Ed. F. Harary), Academic Press, New York, 1973, 275–307.

31. R. Shah, Determinants of binary matrices achieve every integral value up to $\Omega(2^n/n)$, *Linear Algebra Appl.* **645** (2022), 229–236. 10.1016/j.laa.2022.03.024

32. X. Shen, Y. Hou and Y. Zhang, Graph Z_n and some graphs related to Z_n are determined by their spectrum, *Linear Algebra Appl.* **404** (2005), 58–68. 10.1016/S0076-6879(05)04006-1

33. E.R. van Dam and W.H. Haemers, *Which graphs are determined by their spectrum?* Special issue on the Combinatorial Matrix Theory Conference (Pohang, 2002), Vol, 373, 2003, 241–272.

34. E.R. van Dam and W.H. Haemers, Developments on spectral characterizations of graphs, *Discret. Math.* **309** no. 3 (2009), 576–586.

35. V.H. Vu, Recent progress in combinatorial random matrix theory, *Probab. Surv.* **18** (2021), 179–200.

36. W. Wang, A simple arithmetic criterion for graphs being determined by their generalized spectra, *J. Combin. Theory Ser. B* **122** (2017), 438–451. 10.1016/j.jctb.2016.07.004

37. H. Yuan, A bound on the spectral radius of graphs, *Linear Algebra Appl.* **108** (1988), 135–139.

38. J. Zhou and C. Bu, Spectral characterization of line graphs of starlike trees, *Linear Multilinear Algebra* **61** no. 8 (2013), 1041–1050.