DISTINGUISHING GRAPHS WITH ZETA FUNCTIONS AND GENERALIZED SPECTRA

CHRISTINA DURFEE AND KIMBALL MARTIN*

ABSTRACT. Conjecturally, almost all graphs are determined by their spectra. This problem has also been studied for variants such as the spectra of the Laplacian and signless Laplacian. Here we consider the problem of determining graphs with Ihara and Bartholdi zeta functions, which are also computable in polynomial time. These zeta functions are geometrically motivated, but can be viewed as certain generalizations of characteristic polynomials. After discussing some graph properties determined by zeta functions, we show that large classes of cospectral graphs can be distinguished with zeta functions and enumerate graphs distinguished by zeta functions on ≤ 11 vertices. This leads us to conjecture that almost all graphs which are not determined by their spectrum are determined by zeta functions.

Along the way, we make some observations about the usual types of spectra and disprove a conjecture of Setyadi and Storm about Ihara zeta functions determining degree sequences.

1. INTRODUCTION

A fundamental problem in spectral graph theory is: when can we distinguish (unlabeled simple) graphs by their spectra? If the answer were always, as was once conjectured, that would mean we could solve the graph isomorphism problem in polynomial time. However, many pairs of non-isomorphic cospectral graphs have since been found, and various constructions of cospectral pairs are known. Nevertheless, Haemers conjectured that almost all graphs are determined their spectra (DS), i.e., the fraction of graphs of order n which are DS goes to 1 as $n \to \infty$. In fact, from the numerical data for $n \leq 11$ in Haemers–Spence [HS04], it appears that more graphs are determined by their Laplacian spectra (L-DS), and even more by their signless Laplacian spectra ($|L|$-DS). We refer to surveys by van Dam and Haemers for more details [vDH03], [vDH09].

This question of which graphs are DS can be thought of geometrically. Knowing the adjacency spectrum of G is equivalent to knowing what one might call the *walk length spectrum* of G—the set $\{(\ell, w_G(\ell)) : \ell \geq 0\}$, where $w_G(\ell)$ is the number of closed walks of length ℓ on G—together with the order n of G. Hence the above question can be stated: when does the order and walk length spectrum determine G?

From the point of view of Riemannian geometry and covering space theory, it is more natural to look at *geodesics* on G rather than arbitrary walks. Roughly, geodesics are paths with no backtracking and they correspond to lines in the universal cover of G (see Section 2). The *geodesic length spectrum* of G is then the set $\{(\ell, a_G(\ell)) : \ell \geq 0\}$ of numbers of (primitive) closed geodesics of length ℓ for all ℓ. For graphs, the Ihara zeta function $Z_G(t)$ of G encodes the geodesic length spectrum, and knowing one is equivalent to knowing the other.

Date: April 21, 2015.

2010 *Mathematics Subject Classification.* 05C50, 05C30.

Key words and phrases. Ihara zeta function, Bartholdi zeta function, cospectral graphs.

*Partially supported by a Simons Collaboration Grant.

This is entirely analogous to the Selberg zeta function encoding the (geodesic) length spectrum for Riemann surfaces. For compact Riemann surfaces, Huber's theorem says that knowing the Selberg zeta function, i.e., the length spectrum, is equivalent to knowing the spectrum of the Laplacian (see, e.g., [Bus10]). Similarly, for connected regular graphs, knowing the Ihara zeta function is equivalent to knowing the (adjacency or Laplacian) spectrum (see Section 2.1). However, this is not the case for irregular graphs (cf. Table 5.2).

Here, we suggest that the Ihara zeta function—which is computable in polynomial time—provides a more effective way to differentiate (irregular) graphs than the usual spectra studied (Conjecture 5.1). One heuristic for why this should be the case is that geodesics capture much of the geometry of the graph better than arbitrary paths. Another is that the Ihara zeta function typically encodes more information—at least for md2 graphs (graphs with no vertices of degree ≤ 1), the reciprocal of the Ihara zeta function is a polynomial whose degree is twice the number of edges, and thus typically has more coefficients than the characteristic polynomial.

The obvious drawback of the Ihara zeta function is that it cannot detect leaves or isolated nodes. For this reason (at least in part), most studies of the Ihara zeta function restrict to md2 graphs. Most studies also restrict to connected graphs, which is convenient to discuss covering space theory, but this is less crucial for the problem of distinguishing graphs. (As with characteristic polynomials, the zeta function of a disconnected graph is the product of the zeta functions of its components.) Instead of making such restrictions here, we consider four methods of addressing this defect to distinguish graphs (undefined notation explained subsequently):

(M1) Use the order and the Ihara zeta $Z_{G^*}(t)$ function of the cone G^* of G. Equivalently use the order and the characteristic polynomial $\varphi_{T^*}(\lambda) = \det(\lambda I - T^*)$.
(M2) Use the order, size and the Ihara zeta functions $Z_G(t)$ and $Z_{\bar{G}}(t)$ of G and its complement $\bar{G}$. Equivalently, use the order and the characteristic polynomials $\varphi_T(\lambda) = \det(\lambda I - T)$ and $\varphi_{\bar{T}}(\lambda) = \det(\lambda I - \bar{T})$.
(M3) Use the order and the Bartholdi zeta function $\mathcal{Z}_G(t,u)$ of G. Equivalently, use the generalized characteristic polynomial $\varphi_{AD}(\lambda, x) = \det(\lambda I - A + xD)$.
(M4) Use the order and the Bartholdi zeta functions $\mathcal{Z}_G(t,u)$ and $\mathcal{Z}_{\bar{G}}(t,u)$ of G and $\bar{G}$. Equivalently, use the generalized characteristic polynomial $\varphi_{ADJ}(\lambda, x, y) = \det(\lambda I - A + xD + yJ)$.

Here I and J denote the identity and all-ones matrices of the appropriate sizes, and A and D denote the adjacency and degree matrices for a fixed ordering of vertices on G. Further, T (resp. T^*, $\bar{T}$) denotes Hashimoto's oriented edge matrix (see Section 2) of G (resp. G^*, $\bar{G}$)—whose spectrum contains the same information as Z_G together with the size of G (cf. (2.3)). The two-variable Bartholdi zeta function $\mathcal{Z}_G(t,u)$ is a generalization of the one-variable Ihara function $Z_G(t) : \mathbb{C} \to \mathbb{C}$ which counts "closed geodesics with r backtracks"—see Section 3, in particular Equation (3.3) for the relation with φ_{AD}.

Note that knowing φ_{AD} implies knowing the spectra of A, the Laplacian L, and the signless Laplacian $|L|$. Similarly knowing φ_{ADJ} implies knowing all of these spectra for both G and $\bar{G}$.

We will explain the precise motivation behind these choices as we consider each in turn in the body of the paper. The basic ideas are that the Ihara zeta function cannot detect "dangling edges," i.e., edges to a degree 1 node, but (i) taking cones or complements turns degree 1 nodes into higher degree nodes and (ii) the Bartholdi zeta function does detect dangling edges. In analogy with the DS terminology, we say that a graph G is DZ* (resp. DZ$\bar{\text{Z}}$, D$\mathcal{Z}$, D$\mathcal{Z}\bar{\mathcal{Z}}$) if it is uniquely determined by method (M1) (resp. (M2), (M3), (M4)).

We will see in Section 3.2 that (M4) is provably stronger than each of (M1), (M2) and (M3), i.e., being DZ* (resp. DZ$\bar{\text{Z}}$, D$\mathcal{Z}$) implies D$\mathcal{Z}\bar{\mathcal{Z}}$. However there are no obvious implications among (M1), (M2) and (M3) (but there are some indications that (M2) may be stronger than

(M1)—see the end of Section 5.). Nevertheless, a surprising outcome of our calculations in Table 5.1 is that for graphs on $n \leq 11$ vertices, methods (M1), (M2) and (M4) are always equivalent (whereas (M3) has slightly less discriminating power). This suggests one gains little by using the Bartholdi zeta function, and one is in practice justified in just considering the Ihara zeta function, which is both conceptually and computationally simpler. However, all of these methods run in polynomial time.

Now we outline the contents of the paper.

In Section 2, we first recall basic facts about the Ihara zeta function, and discuss some graph properties it determines. Then we consider some basic properties determined by methods (M1) and (M2). In particular, these zeta invariants force strong restrictions on the degree sequence of a graph. Setyadi–Storm [SS13], in their enumeration of pairs of connected md2 graphs with the same Ihara zeta function on $n \leq 11$ vertices, found that the Ihara zeta function of a connected md2 graph determines the degree sequence for $n \leq 11$, and conjectured this holds for all n, but we give a counterexample to this conjecture on 12 vertices (Example 2.2). Nevertheless, we show that knowing the zeta functions of sufficiently many cones of G algorithmically determines the degree sequence (Lemma 2.3).

In Section 3, we recall basic facts about the Bartholdi zeta function and discuss the relative strengths of methods (M1)–(M4). To show that (M4) is stronger than (M1), we show that two graphs have the same spectra with respect to $A + xD$ and $\bar{A} + x\bar{D}$ for fixed x, if and only if the same is true of their joins with another graph (Theorem 3.2). In particular, two graphs have the same A and $\bar{A}$ (or $|L|$ and $\overline{|L|}$) spectra if and only if the same is true for their cones (Corollary 3.3; cf. Table 5.4 for related data).

In Section 4, we show that methods (M1)–(M4) distinguish a large class of cospectral and Laplacian cospectral pairs coming from well known constructions: GM switching, coalescence and join. On the other hand, these methods will not distinguish graphs arising from the more restrictive GM* switching. Section 4.4 presents a new construction of pairs of graphs which cannot be distinguished by these methods. This construction is interesting because it constructs graphs G_1, G_2 such that the generalized adjacency matrices $A_1 + xD_1$ and $A_2 + xD_2$ are miraculously conjugate for all x but not "uniformly conjugate" (as happens in GM* switching) i.e., there is no invertible matrix P such that $P(A_2 + xD_2)P^{-1} = A_1 + xD_1$ for all x.

In Section 5, we enumerate all graphs on ≤ 11 vertices which are not DZ* (resp. DZ$\bar{\text{Z}}$, D$\mathcal{Z}$, D$\mathcal{Z}\bar{\mathcal{Z}}$) in Table 5.1. We state Conjecture 5.1, which asserts almost all graphs are DZ* (resp. DZ$\bar{\text{Z}}$, D$\mathcal{Z}$, D$\mathcal{Z}\bar{\mathcal{Z}}$), and further that almost all non-DS graphs are DZ* (resp. DZ$\bar{\text{Z}}$, D$\mathcal{Z}$, D$\mathcal{Z}\bar{\mathcal{Z}}$). Table 5.2 indicates that there is no essential difference in just using the Ihara zeta function Z_G compared to methods (M1)–(M4) when restricting to md2 graphs. This suggests Conjecture 5.2, which is the analogue of Conjecture 5.1 for determining md2 graphs G using only Z_G. We also compare the effectiveness of combining different kinds of spectra for $n \leq 11$ in Table 5.3. In particular, this suggests that using two of the usual spectra, such as A and L or A and $|L|$, is much more effective at distinguishing graphs than any single one.

Our calculations were done using Sage [Sage], including nauty [nauty], and standard Unix tools. In practice, we found that to check if two graphs have the same zeta function, it almost always sufficed to compute the two numbers $\det|L|$ and $\det(4D+2A-3I)$, which are essentially (the residue of) $Z_G(1)$ and $Z_G(-2)$. For convenience of the interested reader, when mentioning particular examples of graphs, we include (non-canonical) `graph6` strings to specify the graphs, which can be used to easily reconstruct the graphs in Sage.

As a final remark, we note that there are more general notions of zeta functions of graphs, such as path and edge zeta functions (see [ST96] or [Ter11]). We do not consider these here.

We thank C. Storm for discussions about [SS13], as well as the referee for helpful comments.

2. The Ihara zeta function

We begin by defining closed geodesics and the Ihara zeta function. This notion of a geodesic on a graph corresponds to a bi-infinite simple path in the universal cover, but we do not explain this here—see, e.g., [ST96] or [Ter11]. Ihara zeta functions are a special case of more general zeta functions of multigraphs considered by Hashimoto [Has89], generalizing the zeta functions originally defined by Ihara [Iha66].

Fix a finite (simple) graph $G = (V, E)$. Let $n = |V|$ and $m = |E|$. We will denote walks by sequences of adjacent oriented (or directed) edges e_i. For an oriented edge $e_i = (u, v)$, let $e_i^{-1} = (v, u)$ be the edge with reversed orientation. Suppose $\gamma = (e_1, e_2, \dots, e_\ell)$ denotes a closed walk of length ℓ in G. We say γ is a *closed geodesic* if $e_i \neq e_{i+1}^{-1}$ for $1 \leq i < \ell$ and $e_1 \neq e_\ell^{-1}$. The former condition is often expressed saying γ has no backtracking and the latter that γ has no tails. Write $k\gamma$ for the concatenation of k copies of γ. We say γ is *primitive* if γ is not of the form $k\delta$ for a closed geodesic δ and some $k \geq 2$.

Let $\sigma(\gamma) = (e_2, \dots, e_\ell, e_1)$. Then $\sigma(\gamma)$ is also a closed geodesic of length ℓ, and it is primitive if γ is. Thus the cyclic group $\langle\sigma\rangle$ of order ℓ generated by σ acts on all (primitive) closed geodesics of length ℓ. The $\langle\sigma\rangle$-orbits thus partition the set of (primitive) closed geodesics of length ℓ into equivalence classes. Let $a(\ell) = a_G(\ell)$ denote the number of primitive closed geodesics of length ℓ up to equivalence.

Note that, as G is simple, there are no closed geodesics of length < 3. The closed geodesics of length 3, 4 or 5 are just the cycles of length 3, 4 or 5. The closed geodesics of length 6 are just the cycles of length 6 together with the concatenations of 2 simple cycles of length 3 starting at a fixed base point. If there are two distinct cycles based at a vertex v_0, going around the first cycle k times and going around the second cycle once or more is a primitive closed geodesic. Hence, for connected md2 graphs, the lengths of primitive closed geodesics are unbounded unless G is a circuit.

The (Ihara) zeta function of G is

$$Z_G(t) = \prod_{\gamma} (1 - t^{\ell(\gamma)})^{-1} = \prod_{\ell > 2} (1 - t^\ell)^{-a(\ell)} = \exp\left(\sum_{\ell > 2} \sum_{k \geq 1} a(\ell) \frac{t^{\ell k}}{k} \right), \tag{2.1}$$

where, in the first product, γ runs over a set of representatives for the equivalences classes of primitive closed geodesics, and $\ell(\gamma)$ denotes the length of γ. Since no closed geodesics will involve degree 0 or degree 1 nodes, $Z_G(t) = Z_{G^\dagger}(t)$, where $G^\dagger$ is the "pruned graph" obtained by successively deleting degree 0 and degree 1 nodes until one is either left with an md2 graph or the null graph (the graph on 0 vertices).

Note if G is a disjoint union of two subgraphs, $G = G_1 \sqcup G_2$, then $a_G(\ell) = a_{G_1}(\ell) + a_{G_2}(\ell)$. Hence the zeta function of a graph is the product of the zeta functions of its connected components.

The zeta function is *a priori* an infinite product, but turns out to be a rational function and thus is meromorphic on $\mathbb{C}$. Namely, Hashimoto [Has89], [Has92] and Bass [Bas92] gave two determinant formulas for Z_G, which have been subsequently retreated many times (e.g., [ST96]). The Bass determinant formula [Bas92] is

$$Z_G(t) = (1 - t^2)^{n-m} \det(I - tA + t^2(D - I))^{-1}. \tag{2.2}$$

Note the right hand side is invariant under adding nodes of degree 0 or 1 to G.

Let $\{e_1, \dots, e_{2m}\}$ denote the set of oriented edges of G. The *oriented edge matrix* (with respect to this ordering of oriented edges) T is the $2m \times 2m$ matrix whose (i, j)-entry is 1 if $e_i = (u, v)$, $e_j = (v, w)$ and $u \neq w$; or 0 otherwise. The Hashimoto determinant formula [Has89], [Has92]

(cf. [ST96, Thm 3]) is

$$Z_G(t) = \det(I - tT)^{-1} = (t^{2m}\varphi_T(t^{-1}))^{-1}, \tag{2.3}$$

where $\varphi_T(\lambda) = \det(\lambda I - T)$ is the characteristic polynomial of T. Again, one can check that $\det(I - tT)$ is invariant under adding degree 0 or 1 nodes to G.

Since $G^\dagger \neq G$ in general, Z_G does not determine m (or n). This means knowing Z_G is not exactly the same as knowing the spectrum of T, but it almost is: φ_T determines both Z_G and m, and conversely. Indeed, the degree of φ_T is $2m$, so one can recover Z_G from φ_T. The converse is obvious. This observation relates to method (M2).

2.1. **Properties determined by the Ihara zeta function.** Here we summarize some elementary graph properties determined by the Ihara zeta function. This question has previously been considered mainly for md2 and connected md2 multigraphs (e.g., [Cza05] and [Coo09]).

First, examining the coefficients of t^k in the logarithm of (2.1) shows Z_G determines $a(\ell)$ for all ℓ, hence knowing Z_G is the same as knowing the primitive geodesic length spectrum. Note that the number of non-primitive geodesics of a given length l is the sum of the numbers of primitive geodesics of proper divisors of l. Hence the primitive geodesic length spectrum determines the full geodesic length spectrum and vice versa. Consequently Z_G determines the number of cycles of length 3, length 4 and length 5 in G. It is not true that Z_G determines the number of cycles of length 6.

Example 2.1. *The pair of graphs* `HheadXZ` *and* `Hhf@eS|` *pictured below, each with 9 vertices and 18 edges, have the same zeta functions but a different number of cycles of length 6 (46 and 50).*

We remark $Z_G(t) = 1$ if and only if G has no cycles, so Z_G can determine whether a graph is a forest or not. The girth g of G will be the smallest ℓ such that $a(\ell) \neq 0$. So Z_G determines g and, if G is connected, Scott and Storm [SS08] showed Z_G determines the number of cycles of length ℓ for any $\ell < 2g$.

From (2.3) it is clear that $Z_G(t)^{-1}$ is a polynomial in t of degree $2m^\dagger$, where $m^\dagger$ is the size of $G^\dagger$. In general, though Z_G does not determine m or n, it does determine $m = m^\dagger$ if G is md2. If we assume that G is connected (but not necessarily md2), then we can say Z_G determines $m - n$. Namely, if $Z_G(t) = 1$ then G is a tree and $m - n = -1$. Otherwise $G^\dagger$ is not the null graph so $m - n = m^\dagger - n^\dagger$, and $m^\dagger$ and $n^\dagger$ are determined by Z_G, as we will see below. Note that Z_G does not determine $m - n$ if we drop the assumption that G is connected, since all forests have trivial zeta function.

To get clean statements, for the rest of this section assume G is md2 and connected. For general connected G, the results here can be viewed as results about $G^\dagger$ when $G^\dagger$ is not the null graph. Note that if G is connected, then $G^\dagger$ is also connected.

In this case $m \geq n$, and Hashimoto [Has89] showed (i) $Z_G(t)$ has a pole at $t = 1$, which is of order $2 = m - n + 2$ if G is a circuit and $m - n + 1$ otherwise, and (ii) $Z_G(t)$ has a pole at $t = -1$ of order $m - n$, $m - n + 1$ or $m - n + 2$ according to whether G is non-bipartite, a bipartite non-circuit or a bipartite circuit. Hashimoto [Has90] later showed the residue at $t = 1$ for non-circuits is $2^{m-n+1}(n - m)\kappa(G)$, where κ denotes complexity of G, i.e., the number of spanning trees of G. Note that G is bipartite if and only if its geodesics all have even length, and recall G is a circuit if and only if $a(\ell) = 0$ for ℓ sufficiently large.

Consequently, Z_G determines n, m, $\kappa(G)$, whether G is bipartite, and whether G is a circuit. In fact, since $\det(I - tA + t^2(D - I))$ is simply $\det(D + A) = \det|L|$ at $t = -1$, (2.2) tells us $\lim_{t\to -1}(1-t^2)^{n-m}Z_G(t)^{-1} = \det|L|$, which is 0 if and only if G has a bipartite component. So $Z_G(t)$ also determines the product of eigenvalues of $|L|$.

Cooper [Coo09] also showed Z_G determines whether G is regular, and if so, the degree of regularity as well as the spectrum. In fact if G is $(q+1)$-regular, then the spectrum conversely determines Z_G by

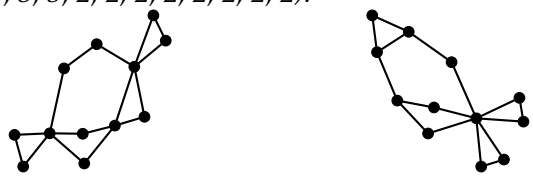
$$Z_G(t) = (1-t^2)^{n-m}\prod_i (1 - \lambda_i t + qt^2)^{-1} \tag{2.4}$$

where $\{\lambda_1, \dots, \lambda_n\}$ are the eigenvalues of A (attributed to A. Mellein [Cza05]).

Write $V = \{v_1, \dots, v_n\}$, $d_i = \deg(v_i)$ and $q_i = d_i - 1$. We see the leading term of $Z_G(t)^{-1}$ is $(\prod q_i)t^{2m}$. Consequently, for G md2, $Z_G(t)$ determines $\prod q_i$. One also knows $2m = \sum d_i$, so knowing $Z_G(t)$ places rather strong restrictions on the degree sequence for md2 graphs.

In fact, Setyadi–Storm [SS13], based on their enumeration of zeta functions of connected md2 graphs on ≤ 11 vertices, conjectured that connected md2 graphs with the same Ihara zeta function have identical degree sequences. However, we give a counterexample.

Example 2.2. *The graphs* `K??CA?_FEcdk` *and* `K??CA?_ccWNk` *on 12 vertices and 16 edges drawn below ("the crab and the squid") have the same zeta function, but their degree sequences are (5, 5, 4, 2, 2, 2, 2, 2, 2, 2, 2, 2) and (7, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2).*



2.2. **Zeta functions of cones—(M1).** The most obvious issue of using zeta functions to distinguish arbitrary graphs is that no closed geodesics will pass through "dangling links," i.e., paths not contained in cycles. To resolve this issue, an obvious thing to try is connecting all the degree 1 nodes to a new vertex. However, doing this to just the degree 1 nodes is not a nice operation on graphs (it is not injective), so it makes more sense to look at the *cone* G^* of G. This is just the join of G with a point: $G^* = G \vee K_1$, where $\vee$ denotes the operation of graph join. The new vertex in G^* is denoted v_{n+1}.

Clearly G^* has order $n+1$ and increases the degree of each vertex in G by 1. That is, G^* has degree sequence (d_i^*) where $d_i^* = d_i + 1$ for $1 \leq i \leq n$ and $d_{n+1}^* = n$. Hence if G is an md1 graph (a graph with no degree 0 vertices), then G^* is an md2 graph. A graph H on $n+1$ vertices will be the cone of some graph G if and only if there is a vertex of degree n, and G can be recovered by deleting any vertex of degree n.

We propose to use Z_{G^*} to study G. Now any edge in G will appear in some closed geodesic in G^*, so it is reasonable to expect that Z_{G^*} encodes much of the structure of G. Of course, degree 0 nodes in G become degree 1 nodes in G^* and still are not detected by Z_{G^*}. If one wanted to, say, count the degree 0 nodes only using zeta functions, one could also look at the zeta function of the double cone $G^{**} = G \vee K_2$. Instead, we will simply fix n, and see how to determine the number of degree 0 nodes from n and Z_{G^*}.

Hence, for the rest of this section, we will assume we know the order n of G (except where stated otherwise) and see what can be deduced about G from Z_{G^*}.

We first show that we can determine the number of degree 0 nodes from n and Z_{G^*}. Let G' be the graph obtained from G by removing all degree 0 nodes. Note G is the empty graph on n vertices if and only if $Z_{G^*} = 1$. Otherwise, $(G^*)^\dagger = (G')^*$ and $(G')^*$ is connected md2, so Z_{G^*}

determines the number of vertices and edges for $(G')^*$ and hence also for G'. Combining this with knowing n, we can determine G from G'. Hence in considerations below we may and will assume $G = G'$, i.e., G has no degree 0 nodes, thus G^* is connected md2.

From the previous section, we immediately see Z_{G^*} determines the following: n, m, $\prod d_i$, and the number of triangles in G, which equals $\frac{1}{2}a_{G^*}(3) - m$. The latter follows as a 3-cycle in G^* corresponds to either a 3-cycle in G or a triangle formed from an edge in G with v_{n+1}. Similarly, $a_{G^*}(4) = a_G(4) + 2\sum \binom{d_i}{2}$, since the latter term is the number of directed paths of length 2 in G. In particular, Z_{G^*} determines $a_G(4) + \sum d_i^2$. Also, one knows $a_{G^*}(5)$, which is $a_G(5)$ plus the number of directed paths of length 3 in G.

An elementary consequence of knowing n, $2m = \sum d_i$ and $\prod d_i$ is the following: if we fix a, b and c and know that each vertex has degree a, b or c, then Z_{G^*} determines the degree sequence. For general G, we remark that one can determine the degree sequence by looking at sufficiently many cones. Namely, let $G^{*(r)} = G \vee K_r$.

Lemma 2.3. *Let G be a graph of possibly unknown order $n \geq 1$. There exists a finite number r such that if $H_1, \ldots, H_r$ is any known sequence of graphs with distinct orders, then $Z_{G\vee H_1}, \ldots, Z_{G\vee H_r}$ determines the degree sequence of G. One may take r to be at most the order of G (or 3 for G with order less than 3). In particular, $Z_{G^*}, \ldots, Z_{G^{*(r)}}$ determines the degree sequence of G.*

Proof. We will take r to be at least 3, in order to guarantee that one of the H_j has order at least 2 (we allow one of the H_j's to be the null graph). We will assume that we know the degree sequences of the H_j and describe how the degree sequence of G can be computed from $Z_{G\vee H_1}, \ldots, Z_{G\vee H_r}$.

Let H be one of the H_j's of order $h \geq 2$. Then $G \vee H$ is connected and if $n \geq 2$ it is also md2. It follows that the order of $(G \vee H)^\dagger$ is less than $h + 2$ if and only if $n = 1$. Otherwise it has order $h + n$ and, in either case, we can find n from $Z_{G\vee H}$. Assume from now on that $n \geq 2$. Then $Z_{G\vee H}$ determines the number of edges of $G \vee H$ from which we can compute the number of edges of G, which we will call m.

Let x_i be the number of vertices of G of degree i for $0 \leq i \leq n-1$. From n and m we have the two linear equations $\sum_{i=0}^{n-1} x_i = n$ and $\sum_{i=0}^{n-1} i x_i = 2m$. Let h_j be the order of H_j for $1 \leq j \leq r$. First assume that each h_j is at least 2 so $G \vee H_j$ are all connected md2 graphs. Then $Z_{G\vee H_j}$ determines the products of the degrees of $G \vee H_j$ minus 1. The degrees coming from vertices of H_j are known and hence $Z_{G\vee H_j}$ will determine $\prod_{i=0}^{n-1}(i + h_j - 1)^{x_i}$, or equivalently, $\sum_{i=0}^{n-1} \ln(i+h_j-1)x_i$. Thus we have a linear system with $r+2$ equations and n unknowns. The coefficient matrix A of this linear system has $1, i, \ln(i+h_1-1), \ln(i+h_2-1), ..., \ln(i+h_r-1)$ as the entries of its i-th column, for $0 \leq i \leq n-1$. If $\mathbf{c} = \begin{bmatrix} c_1 & c_2 & \ldots & c_{r+2} \end{bmatrix}$ is a vector with $\mathbf{c}A = \mathbf{0}$, then the function $f(t) = c_1 + c_2 t + c_3 \ln(t + h_1 - 1) + c_4 \ln(t + h_2 - 1) + ... + c_{r+2} \ln(t + h_r - 1)$ has zeros at $0, 1, 2, ..., n-1$. As the h_j's are distinct, this function changes direction at most r times and thus has at most $r + 1$ zeros unless all c_i are zero. If we take $r = n - 2$ then the c_i must all be 0 so A is an invertible $n \times n$ matrix and the linear system has exactly one solution. If we drop the assumption that the h_j's are all at least 2, then at worst we can ignore Z_G and Z_{G^*} and take $r = n$. □

Remark 2.4. *If we allow $n \geq 0$ in the above proposition, the order of G can be determined using the $Z_{G\vee H_j}$'s, with one exception: if all the H_j's are empty graphs, then the $Z_{G\vee H_j}$ will not distinguish between the null graph and the graph on one vertex. Also, using fewer than n joins often suffices to determine the degree sequence of G. For instance, the proof shows $n-2$ works if all h_j's are ≥ 2. It also ignores the degree information gotten from Z_{G^*} if one of the H_j has order 1. If we also take into account that the x_i must be nonnegative integers, this could significantly decrease the number of joins needed in many cases.*

We also comment that for regular graphs, knowing Z_{G^*} is equivalent to knowing the spectrum of G^* [BS13].

2.3. **Zeta functions of graphs and their complements—(M2).** We begin by thinking of another way to encode the dangling nodes or links in the zeta function. By a dangling node (link), we mean a vertex (edge) not contained in any cycle, i.e., any node (edge) not in $G^\dagger$.

First note that if we fix n, we know the number $n-n^\dagger$ of dangling nodes, since Z_G determines the order $n^\dagger$ of $G^\dagger$. In particular, n and Z_G will tell us if G is connected md2.

Recall that knowing the spectrum of T gives slightly more information than just Z_G—it also tells us m. Since Z_G tells us $m^\dagger$, knowing φ_T tell us $m - m^\dagger$ which is the number dangling links. However, this tells us nothing about the structure of dangling links—e.g., φ_T cannot distinguish among a cycle C_n with leaves added, C_n with a path attached and C_n disjoint union a forest, provided the number of edges match.

If we want to somehow account for dangling nodes with zeta functions, we can also try looking at both G and $\bar{G}$. Degree 0 and degree 1 vertices in G now have degree $n-1$ and $n-2$ in $\bar{G}$, so for $n > 3$ they will now have degree at least 2. However, it is still possible that some of these vertices are dangling nodes in $\bar{G}$. Nevertheless, we can show the following.

Proposition 2.5. *Let G be a graph of order n. Then at least one of the following holds:*
(i) G is determined by φ_T and $\varphi_{\bar{T}}$ (or equivalently, by Z_G, $Z_{\bar{G}}$ and m); or
(ii) for any vertex $v \in V$, we have $v \in G^\dagger$ or $v \in (\bar{G})^\dagger$.

We remark this is not true if (i) is just replaced with "G is determined by Z_G and $Z_{\bar{G}}$." For instance, let $G_1 = K_3 \sqcup \{v\}$ and G_2 be the graph obtained from G_1 by adding an edge from v to one vertex in K_3. Then $\bar{G}_1$ is a tree and $\bar{G}_2$ is a forest, so $Z_{G_1} = Z_{G_2} = Z_{K_3}$ and $Z_{\bar{G}_1} = Z_{\bar{G}_2} = 1$. Thus we cannot distinguish G_1 and G_2 by looking at their zeta functions and their complements' zeta functions. Moreover, v does not lie in $G_i^\dagger$ or $(\bar{G}_i)^\dagger$ for $i = 1, 2$.

Before the proof, we give some numerical evidence that (M2)—using φ_T and $\varphi_{\bar{T}}$ (or Z_G, $Z_{\bar{G}}$ and m) to distinguish graphs of order n—is much better than just using Z_G and $Z_{\bar{G}}$, or just Z_G or φ_T. This data is presented in Table 2.1. The second column is the total number of graphs on n nodes. Subsequent columns contain the number of graphs G on n nodes which are not determined respectively by Z_G, by Z_G and $Z_{\bar{G}}$, by φ_T, and by both φ_T and $\varphi_{\bar{T}}$.

TABLE 2.1. Number of small graphs not distinguished by zeta invariants

n	# graphs	Z	$Z\bar{Z}$	T	$T\bar{T}$
2	2	2	2	0	0
3	4	3	2	0	0
4	11	8	4	4	0
5	34	23	8	15	0
6	156	94	22	75	0
7	1,044	534	68	449	0
8	12,346	4,889	312	4,297	0
9	274,668	76,807	350	68,708	2

The pair of graphs (G_1, G_2) on 9 vertices with the same φ_T and $\varphi_{\bar{T}}$ are the graphs with 18 edges pictured in Example 2.2. We remark that $G_1 \simeq \bar{G}_2$ in this example.

One explanation for why (M2) is so effective is that in most cases it restricts the problem of distinguishing arbitrary graphs with zeta functions to looking just at connected md2 graphs, where zeta functions give us a lot of information (cf. Section 2.1). Similarly we can determine a

lot about G if $\bar{G}$ is connected md2. And if neither G nor $\bar{G}$ is connected md2, this places strong constraints on G, as we will see in the proof.

We first treat a special case, where we can say something stronger.

Lemma 2.6. *Suppose G is a forest of order n. Then at least one of the following holds:*
(i) $\bar{G}$ is connected md2, or
(ii) φ_T and $\varphi_{\bar{T}}$ determine G among all graphs of order n.

Proof. By Table 2.1, we see φ_T and $\varphi_{\bar{T}}$ determine G for at least $n \leq 8$. Assume $n \geq 5$. Then G has at most 1 vertex of degree $\geq n-2$. If it has none, $\bar{G}$ is connected md2, so assume it does. Then G must be one of the following three graphs: a star graph on n nodes, a star-like graph on n nodes with exactly 1 node distance 2 from the hub (so the other $n-2$ non-hub nodes are adjacent to the hub), or a star graph on $n-1$ nodes union a point. The latter case can be distinguished from the previous two by counting edges. The former two can be distinguished by looking at the number of edges in $(\bar{G})^\dagger$. Since Z_G detects forests, φ_T and $\varphi_{\bar{T}}$ determine these 3 graphs among all graphs of order n. □

Proof of Proposition 2.5. If G or $\bar{G}$ is md2, then clearly (ii) holds, so we may assume neither is md2. Switching G and $\bar{G}$ if necessary, we may also assume G is connected. In light of the lemma, we may further assume $G^\dagger = (V^\dagger, E^\dagger)$ is not the null graph, and hence $G^\dagger$ has order $n^\dagger \geq 3$.

Let $H = (W, F)$ be the subgraph (which is a forest) of G induced from $W = V - V^\dagger$. Note that, in G, no $w \in W$ can be adjacent to more than 1 vertex in $V^\dagger$. Also, in any connected component of H, there is exactly one vertex which is adjacent to a vertex in $V^\dagger$. Suppose that H contains at least two vertices w_1, w_2. If w_1 and w_2 are not adjacent in G then there exists $v \in V^\dagger$ such that neither w_1 nor w_2 is adjacent to v in G and hence $\bar{G}$ contains the triangle w_1, w_2, v. If w_1 and w_2 are adjacent in G, then at most one of them is adjacent to a vertex in $V^\dagger$ so there exist $v_1, v_2 \in V^\dagger$ which are not adjacent to either w_1 or w_2. Then $\bar{G}$ contains the 4-cycle $\{w_1, w_2\} \vee \{v_1, v_2\}$. In either case, both w_1 and w_2 appear in $(\bar{G})^\dagger$.

It remains to consider the case where $|W| = 1$. Let w be the unique vertex in W and let v be the unique vertex in $V^\dagger$ which is adjacent to w in G. If any two vertices v_1, v_2 in $V^\dagger - \{v\}$ are not adjacent, then w, v_1, v_2 forms a triangle in $\bar{G}$. Also, if there exists $v_1, v_2 \in V^\dagger$ which are not adjacent to v then $\{w, v\} \vee \{v_1, v_2\}$ is a 4-cycle in $\bar{G}$. In either case, w appears in $(\bar{G})^\dagger$. We may therefore assume the graph induced by $V^\dagger - \{v\}$ is a complete graph and there is a most 1 vertex in $V^\dagger - \{v\}$ which does not connect to v. This leaves us with 2 possibilities for G, but we see that $\bar{G}$ for these two possibilities are a star-like graph on n nodes with exactly 1 node distance 2 from the hub or a star graph on $n-1$ nodes union a point. In the proof of the last lemma, we showed that these two graphs are distinguished by φ_T and $\varphi_{\bar{T}}$ among all graphs of order n. □

Note that the proofs of the lemma and proposition show that for $n \geq 5$, there are exactly three pairs of graphs $(G, \bar{G})$ of order n for which it is not true that every vertex v of G is in $G^\dagger$ or $\bar{G}^\dagger$. These are the following three graphs and their complements: the star graph on n nodes, the star-like graph on n nodes with exactly 1 node distance 2 from the hub, and the star graph on $n-1$ nodes union a point. Hence in all but these three cases, every vertex of G appears in a geodesic of G or $\bar{G}$ (or both).

3. The Bartholdi zeta function

An alternative to using zeta functions of graphs related to G in order to study G is to consider a more general notion of zeta function which involves dangling nodes and links. If we think

about the adjacency spectrum, or equivalently the closed walk spectrum, it can distinguish things like path graphs from star graphs because backtracking is allowed in closed walks. On the other hand, the closed walk spectrum loses a lot of information contained in the geodesic length spectrum. Bartholdi [Bar99] introduced a more general zeta function which encodes both the closed walk spectrum and the geodesic length spectrum.

Let $\gamma = (e_1, \dots, e_\ell)$ be a closed walk of length $\ell = \ell(\gamma)$. The number of backtracks in γ is the number of $1 \le i < \ell$ such that $e_{i+1} = e_i^{-1}$. We say γ has a tail if $e_\ell = e_1^{-1}$. The cyclic bump count $\mathrm{cbc}(\gamma)$ is the number of backtracks in γ plus 1 or 0, according to whether γ has a tail or not. The cyclic permutation group $\langle\sigma\rangle$ defined in Section 2 acts on closed walks of length ℓ and preserves the cyclic bump count. A closed walk is primitive if it is not of the form $k\gamma$ for $k > 1$. Let $a(\ell; c) = a_G(\ell; c)$ denote the number of $\langle\sigma\rangle$ orbits of primitive closed walks in G with $\ell(\gamma) = \ell$ and $\mathrm{cbc}(\gamma) = c$. Note $a(\ell; 0) = a(\ell)$ since $\mathrm{cbc}(\gamma) = 0$ means γ is a geodesic. The *Bartholdi zeta function* is

$$\mathcal{Z}_G(t,u) = \prod_{\gamma}(1-u^{\mathrm{cbc}(\gamma)}t^{\ell(\gamma)})^{-1} = \prod_{c,\ell}(1-u^c t^\ell)^{-a(\ell;c)} = \exp\left(\sum_{c,\ell} a(\ell;c)\sum_{k\ge 1}\frac{u^{ck}t^{\ell k}}{k}\right), \tag{3.1}$$

where γ runs over $\langle\sigma\rangle$ equivalence classes of primitive closed walks in G. Note this gives the Ihara zeta function when $u = 0$: $\mathcal{Z}_G(t, 0) = Z_G(t)$.

Bartholdi [Bar99] proved an analogue of the Bass determinant formula:

$$\mathcal{Z}_G(t,u) = (1-(1-u)^2t^2)^{n-m}\det(I - tA + (1-u)(D-(1-u)I)t^2)^{-1}. \tag{3.2}$$

3.1. **Properties determined by the Bartholdi zeta function—(M3).** First observe (3.1) tells us that knowing $\mathcal{Z}_G$ is equivalent to knowing all of the numbers $a(\ell; c)$. Since $a(2; 1) = 2m$, $\mathcal{Z}_G$ determines m. However, $\mathcal{Z}_G$ does not determine n as adding isolated vertices does not change $\mathcal{Z}_G$. We see $\mathcal{Z}_G$ determines the number of 3-, 4-, and 5- cycles in G since Z_G does. However it does not determine the number of 6-cycles, as the pair of graphs in Example 2.2 have the same Bartholdi zeta function.

[KL08] observed that one can rewrite (3.2) in terms of the generalized characteristic polynomial $\varphi_{AD}^G(\lambda, x) = \det(\lambda I - A + xD)$ by

$$\mathcal{Z}_G(t,u) = (1-(1-u)^2t^2)^{n-m}t^{-n}\varphi_{AD}^G(t^{-1}-(1-u)^2t, (1-u)t)^{-1}, \tag{3.3}$$

We will write $\varphi_{AD} = \varphi_{AD}^G$ if the graph G is clear from context. It is stated in [WLLX11] that $\mathcal{Z}_G(t, u)$ determines φ_{AD} and vice versa, but this is not true without further qualification in the same way that Z_G does not determine φ_T. Namely, φ_{AD} determines m and n, so also determines $\mathcal{Z}_G$. On the other hand, $\mathcal{Z}_G$ does not determine n. However, since $\mathcal{Z}_G$ determines m we see that $\mathcal{Z}_G$ and n determine φ_{AD} and vice versa.

From now on, we now consider (M3): what can be determined from n and $\mathcal{Z}_G$, or equivalently, φ_{AD}?

By specializing $\varphi_{AD}(\lambda, x)$ to $x = 0, \pm 1$, we see φ_{AD} determines the spectra of A, L and $|L|$. There is much literature about what these spectra individually determine about G, and various families of graphs that are determined by such spectra (e.g., see the books [BH12], [CRS10] and the survey articles [vDH03], [vDH09]). We just recall a few things determined by knowing all of these spectra: the number of edges, regularity, the number of components, the number of bipartite components, the complexity and the closed walk spectrum.

From Z_G, we also know whether G is connected md2, and all of the things discussed in Section 2.1. In addition, [WLLX11] proves that φ_{AD} determines the degree sequence of G.

3.2. **Bartholdi zeta functions of complements—(M4).** Finally, consider our last method (M4): what can be determined from n, $\mathcal{Z}_G$ and $\mathcal{Z}_{\bar{G}}$.

Equation (3.3) tells us these quantities determine the spectra of $A + xD$ and $A + xD - J$ for all x. When $x = 0$, [JN80] showed this determines the spectra of $A + yJ$ for all y, and the proof (see [vDH03] for a simpler proof) in fact works for $x \neq 0$, i.e., n, $\mathcal{Z}_G$ and $\mathcal{Z}_{\bar{G}}$ determine the generalized characteristic polynomial $\varphi_{ADJ}(\lambda, x, y) = \varphi^G_{ADJ}(\lambda, x, y) = \det(\lambda I - A + xD + yJ)$. The converse, that φ_{ADJ} determines n, $\mathcal{Z}_G$ and $\mathcal{Z}_{\bar{G}}$, is straightforward.

Consequently, (M4) determines everything (M2) and (M3) do. In fact, we show below that (M4) determines everything (M1) does.

Lemma 3.1. *Let X and Y be $n \times n$ and $m \times m$ matrices respectively and let J_{nm} denote the $n \times m$ all ones matrix. Let M be the block matrix $M = \begin{bmatrix} X & J_{nm} \\ J_{nm} & Y \end{bmatrix}$. We will write J for the square all ones matrix when the order is clear from context. The following are true:*

(1) The spectra of M and $J - M$ are determined by the spectra of $X, J - X, Y$, and $J - Y$.
(2) The spectra of X and $J - X$ are determined by the spectra of $M, J - M, Y$, and $J - Y$.

Proof. Note that, for an $r \times r$ matrix, one can determine the spectrum of the matrix from the traces of the first r powers of the matrix and, conversely, one can determine the traces of all powers from the spectrum.

We can prove using induction that M^k has the form $\begin{bmatrix} X^k + A_k & B_k \\ C_k & Y^k + D_k \end{bmatrix}$ where the matrices A_k, B_k, C_k, D_k are sums of matrices which are the product of k matrices coming from $\{X, Y, J_{nm}, J_{mn}\}$. For $k \geq 2$, each product of k matrices appearing in A_k and D_k is such that at most $k - 2$ of the k matrices are X's or Y's, and for B_k and C_k at most $k - 1$ are X's or Y's.

From the equation for M^k, we get that $\operatorname{tr}(M^k) = \operatorname{tr}(X^k) + \operatorname{tr}(Y^k) + \operatorname{tr}(A_k) + \operatorname{tr}(D_k)$. We also use the following properties: $\operatorname{tr}(XJ_{nm}YJ_{mn}) = (\sum_{i,j} x_{ij})(\sum_{i,j} y_{ij}) = \operatorname{tr}(XJ)\operatorname{tr}(YJ)$, trace is invariant under cyclic permutations of products, and the product of any two all ones matrices is a scalar multiple of an all ones matrix. From these properties, it follows that $\operatorname{tr}(A_k)$ and $\operatorname{tr}(D_k)$ can be determined from $\operatorname{tr}(X^iJ)$, and $\operatorname{tr}(Y^iJ)$ for $i = 0, 1, 2, .., k - 2$. This leads to a relation of the form $\operatorname{tr}((J-X)^k) = (-1)^k(\operatorname{tr}(X^k) - k\operatorname{tr}(X^{k-1}J) + \cdots)$, where the omitted terms are determined by $\operatorname{tr}(X^iJ)$ for $i = 0, 1, 2, ..., k - 2$, as well as a similar relation for Y.

We now prove the two statements. Note that in both cases we can determine n and m. Also note that $J - M$ is the block diagonal matrix with $J - X$ and $J - Y$ on the diagonals, so any two of the spectra of $J - M, J - X$, and $J - Y$ determines the third.

First suppose we know the spectra of the matrices $X, J - X, Y$, and $J - Y$. Then we know $\operatorname{tr}((J - X)^i)$ and $\operatorname{tr}(X^i)$ for all i. From this, we can recursively determine $\operatorname{tr}(X^iJ)$ for all i, and similarly for Y. We can therefore compute $\operatorname{tr}(M^k) = \operatorname{tr}(X^k) + \operatorname{tr}(Y^k) + \operatorname{tr}(A_k) + \operatorname{tr}(D_k)$ for all k and thus determine the spectrum of M.

Suppose now that we know the spectra of $Y, J - Y, M, J - M$. As mentioned above, this tells us the spectrum of $J - X$. It remains to show that we can compute $\operatorname{tr}(X^k)$ for $k = 1, 2, ..., n$. For $k = 1, 2$ we can find $\operatorname{tr}(X)$ and $\operatorname{tr}(X^2)$ from $\operatorname{tr}(M) = \operatorname{tr}(X) + \operatorname{tr}(Y)$ and $\operatorname{tr}(M^2) = \operatorname{tr}(X^2) + \operatorname{tr}(Y^2) + \operatorname{tr}(J_{mn}J_{nm}) + \operatorname{tr}(J_{nm}J_{mn})$. If $k > 2$ and we know $\operatorname{tr}(X^i)$ for $i < k$, then from $\operatorname{tr}((J-X)^i)$ and $\operatorname{tr}(X^i)$ for $i = 1, 2, .., k-1$ we can find $\operatorname{tr}(X^iJ)$ for $i = 1, 2, ..., k-2$, and therefore also $\operatorname{tr}(A_k)$ and $\operatorname{tr}(D_k)$. From these, we can compute $\operatorname{tr}(X^k)$ from $\operatorname{tr}(M^k) = \operatorname{tr}(X^k) + \operatorname{tr}(Y^k) + \operatorname{tr}(A_k) + \operatorname{tr}(D_k)$. □

If D is the degree matrix of G, we write $\bar{D}$ for the degree matrix of $\bar{G}$.

Theorem 3.2. *Fix $x \in \mathbb{C}$. Let G_1 and G_2 be two graphs of the same order and let H be any graph. The graphs $G_1 \vee H$ and $G_2 \vee H$ are cospectral with respect to $A + xD$ and $\bar{A} + x\bar{D}$ if and only if G_1 and G_2 are cospectral with respect to $A + xD$ and $\bar{A} + x\bar{D}$.*

Proof. Let $M = A + xD$ and note that, for an order r graph, $\bar{M} := \bar{A} + x\bar{D} = J - M + (x(r-1) - 1)I$. Thus if we know x and r, knowing the spectrum of $\bar{M}$ is equivalent to knowing the spectrum of $J - M$. Let n be the order of G_1 and G_2, let m be the order of H, and denote their respective matrices $M = A + xD$ by M_{G_i} and M_H. The join $G_i \vee H$ has matrix $M_{G_i \vee H} = \begin{bmatrix} M_{G_i} + xmI_n & J_{nm} \\ J_{mn} & M_H + xnI_m \end{bmatrix}$. We can thus apply the previous lemma to show that if G_1 and G_2 are cospectral with respect to M and $\bar{M}$, then so are $G_1 \vee H$ and $G_2 \vee H$ and vice versa. □

This is already well known in the case that $x = -1$, which corresponds to the Laplacian—in fact less is needed in this case since the Laplacian spectrum of G determines that of $\bar{G}$.

Corollary 3.3. *The graphs G_1 and G_2 have the same spectra with respect to A and $\bar{A}$ (or $|L|$ and $\overline{|L|}$), if and only if the same is true for G_1^* and G_2^*. Similarly, G_1 and G_2 have the same generalized characteristic polynomial φ_{ADJ} if and only if the same is true for G_1^* and G_2^*, i.e., $\varphi_{ADJ}^{G_1} = \varphi_{ADJ}^{G_2}$ if and only if $\varphi_{ADJ}^{G_1^*} = \varphi_{ADJ}^{G_2^*}$.*

4. Constructions

Here we discuss three well-known constructions of cospectral graphs, and show that in many cases, such graphs are distinguished by methods (M1)–(M4) using constraints on degree distributions. This provides some evidence for our conjectures in the next section that most cospectral graphs can be distinguished by any of (M1)–(M4).

In the last part of this section, we give a new construction for graphs which cannot be distinguished by (M1)–(M4), which generalizes the pair of graphs from Example 2.1.

4.1. **GM switching.** Godsil and McKay present a method for constructing cospectral pairs of graphs [GM82], which is now referred to as GM switching. We say G satisfies the $(k+1)$-GM (or just GM) condition if there is an ordering of the vertices such that the adjacency matrix can be written in the form

$$A = \begin{bmatrix} B_1 & B_{12} & \cdots & B_{1k} & N_1 \\ {}^tB_{12} & B_2 & \cdots & B_{2k} & N_2 \\ \vdots & & \ddots & & \vdots \\ {}^tB_{1k} & \cdots & {}^tB_{k-1,k} & B_k & N_k \\ {}^tN_1 & \cdots & {}^tN_{k-1} & {}^tN_k & C \end{bmatrix}$$

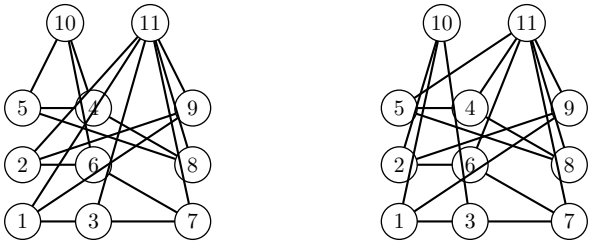
where (i) each G_{B_i} is regular, (ii) each $B_{i,j}$ has constant row and column sums, and (iii) each column of each N_i has exactly 0, $b_i/2$ or b_i 1's. Here b_i is the order of each square matrix B_i, and we assume at least one of the b_i's is even. Note $k+1$ is the number of diagonal blocks, so the ordering of the vertices and the size of the B_i's determines a partition of the vertex set into $k+1$ subsets $V_{B_1}, \dots, V_{B_k}, V_C$.

Let $\widetilde{N}_i$ be the matrix formed from N_i by replacing each column $\mathbf{v}$ of N_i which consists of $b_i/2$ ones by the column $J_{b_i,1} - \mathbf{v}$. The GM switch of G is the graph $\widetilde{G}$ with adjacency matrix $\widetilde{A}$ which is gotten from A by replacing each N_i with $\widetilde{N}_i$. Then G and $\widetilde{G}$ are cospectral with respect to the adjacency matrix. The proof is to exhibit a matrix Q which conjugates A to $\widetilde{A}$ and has constant row and column sums of 1, and hence commutes with the all ones matrix J. It follows that the complements of G and $\widetilde{G}$ are also cospectral with respect to the adjacency matrix. Also, by Corollary 3.3, the cones of G and $\widetilde{G}$ are cospectral with respect to the adjacency matrix.

Haemers and Spence introduced a special case of GM switching [HS04], which they called GM* switching, which gives cospectral pairs with respect to any matrix of the form $A + xD$. They only define GM* switching in the case where $k = 1$, but it works in the more general setup as well. We say G satisfies the $(k+1)$-GM* condition if G satisfies the $(k+1)$-GM condition with the additional requirement that each vertex in G_{B_i} has the same degree in G (i.e., each N_i has constant row sums). This condition guarantees that G and $\widetilde{G}$ have the same degree matrix D. Also D and J commute with Q, the matrix which conjugates A to $\widetilde{A}$, and hence $A+xD+yJ$ and $\widetilde{A}+xD+yJ$ are cospectral for any x, y so G and $\widetilde{G}$ have the same generalized characteristic polynomial φ_{ADJ}. Also, by Corollary 3.3, the cones of G and $\widetilde{G}$ have the same φ_{ADJ}. We call this $(k+1)$-GM* switching and say that $(G, \widetilde{G})$ is a $(k+1)$-GM* pair provided $G \not\cong \widetilde{G}$.

In the next section, we will find that GM* switching accounts for a significant percentage of graphs up to 11 vertices which have the same φ_{ADJ} (see Table 5.1). For the graphs on up to 10 vertices, 3-GM* switching does not produce any such examples that 2-GM* does not. However for 11 vertices, there are 108 3-GM* pairs which cannot be obtained by 2-GM* switching, even if one allows successive 2-GM* switching. (It happens that (G_1, G_2) can be a 3-GM* pair but not a 2-GM* pair, but that G_1 and G_2 are both obtained as 2-GM* switches from a third graph G_3.)

Example 4.1. *The graphs* `J?BD?oX[F[?` *and* `J?'CP''LE{?` *on 11 vertices and 18 edges drawn below are a 3-GM* pair, but not a 2-GM* pair. Further, neither of these form a 2-GM* pair with any other graph. The subgraphs G_{B_1}, G_{B_2}, and G_C are the subgraphs induced by vertices 1–6, 7–9, and 10–11 respectively.*



We now focus on the case of 2-GM switching. For simplicity, we omit the subscripts and write the adjacency matrix as

$$A = \begin{bmatrix} B & N \\ {}^tN & C \end{bmatrix}.$$

For $G \not\cong \widetilde{G}$, we need B to have even size ≥ 4. Further, the larger B is, the less likely it is that the GM condition will be satisfied. So Haemers and Spence [HS04, Thm 3] use GM switching with B of size 4 to get a lower bound on the number of non-DS graphs. The following shows that most cospectral pairs thus constructed are distinguished by zeta functions.

Theorem 4.2. *Suppose G_B is regular on 4 vertices and G_C is an md1 (resp. md2) graph on n vertices. Then the proportion of GM-admissible choices of N such that the pair $(G, \widetilde{G})$ of* labeled *graphs formed by GM switching on (B, C, N) which are distinguished by Z^* (resp. Z) goes to 1 as $n \to \infty$.*

Proof. For each column of N, there are $2 + \binom{4}{2} = 8$ possible choices, which we count with equal probability. The probability that at least 2 columns are all ones is

$$1 - \left[\binom{n}{0} \left(\frac{1}{8}\right)^0 \left(\frac{7}{8}\right)^n + \binom{n}{1} \left(\frac{1}{8}\right)^1 \left(\frac{7}{8}\right)^{n-1} \right] = 1 - \left(\frac{7+n}{8}\right) \left(\frac{7}{8}\right)^{n-1}.$$

When this is satisfied, G and $\widetilde{G}$ are md1 if G_C is md 1 and md2 if G_C is md2. Since this probability tends to 1 as $n \to \infty$, from now on, we will assume G and $\widetilde{G}$ are md1 in the case of Z^* or md2 in the case of Z.

If G and $\widetilde{G}$ are not distinguished by Z^* (resp. Z), then the products of the vertex degrees (resp. degrees minus 1) must be the same. Since all vertices coming from G_C have the same degree, this is equivalent to knowing that the products of the degrees of vertices (resp. degrees minus 1) from B are the same. Let k be the degree of a vertex in G_B plus the number of all ones columns in N in the case of Z^*, or this number minus 1 in the case of Z. Let N_0 be the submatrix of N formed by removing the columns consisting of all zeroes or all ones. Let x_i denote the number of 1's in the i-th row of N_0. Then G and $\widetilde{G}$ being distinguished by Z^* or Z implies

$$\prod_{i=1}^{4}(k+x_i) = \prod_{i=1}^{4}(k+n_0-x_i), \tag{4.1}$$

where n_0 is the number of columns in N_0.

View n_0 as fixed for now. Since $x_4 = 2n_0 - x_1 - x_2 - x_3$, for fixed x_1, x_2, the solutions to (4.1) in x_3 are the solutions to a degree 2 polynomial in x_3, of which there are at most 2. Now view the top 2 rows of N_0 as fixed. Let r be the number of columns in N_0 which have exactly one 1 in the first 2 entries. Then there are $r+1$ choices for x_3. There are 2^r choices for row 3 of N_0, hence for an integer $0 \le y \le r$, the probability that $x_3 = y$ is $\binom{r}{y}\frac{1}{2^r} \le \binom{r}{\lceil r/2 \rceil}\frac{1}{2^r}$. Given $r \ge 3$, the probability that x_3 avoids solving (4.1) is at least

$$1 - 2\binom{r}{\lceil r/2 \rceil}\frac{1}{2^r} \ge 1 - \frac{2}{\sqrt{3\lceil r/2 \rceil + 1}} \ge 1 - \sqrt{\frac{8}{3r}}.$$

(Here we use the inequality $\binom{2n}{n} \le \frac{2^{2n}}{\sqrt{3n+1}}$, which implies $\binom{r}{\lceil r/2 \rceil} \le \frac{2^r}{\sqrt{3\lceil r/2 \rceil + 1}}$.) Also note

$$P(width(N_0) = n_0) = \binom{n}{n_0}\left(\frac{3}{4}\right)^{n_0}\left(\frac{1}{4}\right)^{n-n_0}$$

and for a fixed n_0 the probability that the number of columns in N_0 which have exactly one 1 in the first 2 entries is r is

$$P(r|n_0) = \binom{n_0}{r}\left(\frac{2}{3}\right)^{r}\left(\frac{1}{3}\right)^{n_0-r}.$$

Fix $0 < \delta < 1$. By the law of large numbers, for any $\epsilon > 0$, the probability that $r \ge n_0(\frac{2}{3} - \epsilon)$ goes to 1 as $n_0 \to \infty$. Hence the probability that $r \ge n_0^\delta$ goes to 1 as $n_0 \to \infty$. Similarly, the probability that $n_0 \ge n^\delta$ goes to 1 as $n_0 \to \infty$. So the probability that x_3 avoids the solutions of (4.1) is at least

$$P(n_0 \ge n^\delta)P(r \ge n_0^\delta)\left(1 - \sqrt{\frac{8}{3n_0^\delta}}\right) \ge P(r \ge n^{\delta^2})\left(1 - \sqrt{\frac{8}{3n^{\delta^2}}}\right),$$

which goes to 1 as $n \to \infty$. □

Let g_n denote the number of simple graphs (up to isomorphism) of order n.

Corollary 4.3. *The number of graphs G on n vertices for which there exists a cospectral nonisomorphic graph $\widetilde{G}$ but $Z_G \ne Z_{\widetilde{G}}$ is at least $n^3 g_{n-1}(\frac{1}{24} - o(1))$. The same statement is true with $Z_G \ne Z_{\widetilde{G}}$ replaced by $Z_{G^*} \ne Z_{\widetilde{G}^*}$.*

Proof. Haemers and Spence [HS04, Thm 3] show there are at least $n^3 g_{n-1}(\frac{1}{24} - o(1))$ non-isomorphic pairs $(G, \widetilde{G})$ obtained from GM switching with B of size 4. Further, for almost all of these pairs, both graphs are md2. Since, in the argument above, we can replace the condition that C is md2 or md1 with knowing G and $\widetilde{G}$ are md2 or md1, we see almost all such pairs are distinguished by Z or Z^*. □

4.2. **Coalescence construction.** Suppose G_1 and G_2 are two graphs of order n with the same adjacency spectra but different degree sequences. Assume there are vertices x_1 of G_1 and x_2 of G_2 such that $G_1 - \{x_1\}$ and $G_2 - \{x_2\}$ are cospectral. Let U_i be the vertex set of $G_i - \{x_i\}$. Let Γ be any graph with a fixed vertex y. Let G_i' be the coalescence of (G_i, x_i) with (Γ, y), i.e., the union of G_i and Γ after identification of x_i with y.

Proposition 4.4. *With notation as above,* $\mathrm{Spec}(G_1') = \mathrm{Spec}(G_2')$ *for any* (Γ, y)*. If* G_1 *and* G_2 *are md2 and* $\prod_{v\in U_2}(\deg(v)-1) \neq \prod_{v\in U_2}(\deg(v)-1)$*, then* $Z_{G_1'} \neq Z_{G_2'}$*. Similarly, if* G_1 *and* G_2 *are md1 and* $\prod_{v\in U_1}\deg(v) \neq \prod_{v\in U_2}\deg(v)$*, then* $Z_{G_1'^*} \neq Z_{G_2'^*}$*.*

Proof. The cospectrality is due to Schwenk [Sch73].

For the distinction by zeta functions, note that $G_1 - \{x_1\}$ and $G_2 - \{x_2\}$ cospectral implies that $\deg_{G_1}(x_1) = \deg_{G_2}(x_2)$ as the spectrum determines the number of edges. Note $\deg_{G_i}(v) = \deg_{G_i'}(v)$ for any $v \in U_i$ and $\deg_{G_1'}(v) = \deg_{G_2'}(v)$ for any vertex v of Γ. The assertions follow as Z_G (resp. Z_{G^*}) determines the product of the degrees minus 1 (degrees) of the pruned graph $G^\dagger$. □

Example 4.5. *Let* G_1 *and* G_2 *be the graphs* `F?zPw` *and* `F@Rfo` *on 7 vertices and 10 edges pictured below, where* x_1 *and* x_2 *are the white vertices.*

Then G_1 *and* G_2 *are cospectral but not isomorphic, whereas* $G_1 - x_1$ *and* $G_2 - x_2$ *are cospectral because they are isomorphic. Here* x_1 *and* x_2 *have degree 2. The other vertex degrees are* $(4,4,4,2,2,2)$ *for* G_1 *and* $(5,3,3,3,2,2)$ *for* G_2*. It is clear that the products of the degrees and the products of the degrees minus 1 are different. Thus, for any coalescences* G_1' *and* G_2' *of* (G_1, x_1) *and* (G_2, x_2) *with any* (Γ, y)*, we have* $Z_{G_1'} \neq Z_{G_2'}$ *and* $Z_{G_1'^*} \neq Z_{G_2'^*}$*.*

We assumed that G_1 and G_2 are md2 or md1 in this proposition for simplicity, but this is not necessary. What one really needs is a condition on pruned subgraphs of G_1' and G_2'.

Note the coalescence construction includes the case of a disjoint union. If H_1 and H_2 are cospectral, we can take G_i to be H_i disjoint union a single vertex x_i. Then G_1 and G_2 are cospectral, as are $G_1 - \{x_1\} = H_1$ and $G_2 - \{x_2\} = H_2$ by assumption. Let Γ be any graph and y any vertex in Γ. Then the coalescence of (G_i, x_i) with (Γ, y) is simply the disjoint union $H_i \sqcup \Gamma$ of H_i with Γ. However in the case of disjoint unions, we already know the stronger statement that $Z_{H_1} \neq Z_{H_2}$ implies $Z_{G_1'} \neq Z_{G_2'}$ since zeta functions factor into products over their connected components (though this factorization is not true for the zeta of the cones $Z_{G_i'^*}$). E.g., if H_1 and H_2 are the unique pair of cospectral graphs on 5 vertices, then $Z_{H_1\sqcup\Gamma} \neq Z_{H_2\sqcup\Gamma}$ for any Γ, though G_1 and G_2 are not md2.

4.3. **Join construction.** Suppose G_1 and G_2 are two graphs of order n which have the same Laplacian spectra but different degree sequences. Let Γ be an arbitrary graph. If G_1 or G_2 has

an isolated vertex, then assume Γ has at least 2 vertices. Then the joins $G_1 \vee \Gamma$ and $G_2 \vee \Gamma$ are connected and md2.

Proposition 4.6. *With notation as above,* $\mathrm{Spec}_L(G_1 \vee \Gamma) = \mathrm{Spec}_L(G_2 \vee \Gamma)$ *for any* Γ. *However* $Z_{G_1 \vee \Gamma} \neq Z_{G_2 \vee \Gamma}$ *for all but finitely many* Γ. *Specifically, there is a finite set* S *consisting of at most* $n-1$ *integers such that* $Z_{G_1 \vee \Gamma} \neq Z_{G_2 \vee \Gamma}$ *for any* Γ *whose order* r *does not lie in* S.

Proof. The first part is true for any Γ.

Let $d_1, \dots, d_n$ be the degree sequence for G_1 and $d_1', \dots, d_n'$ be the degree sequence for G_2. Let $\delta_1, \dots, \delta_r$ be the degree sequence for Γ. Then the degree sequence for $G_1 \vee \Gamma$ is $d_1 + r, \dots, d_n + r, \delta_1 + n, \dots, \delta_r + n$ and similarly for $G_2 \vee \Gamma$. Assume $r \geq 2$ so $G_1 \vee \Gamma$ and $G_2 \vee \Gamma$ are md2. Consequently $Z_{G_1 \vee \Gamma}$ determines

$$\prod (d_i + r - 1) \cdot \prod (\delta_j + n - 1)$$

and similarly for $G_2 \vee \Gamma$. Hence $Z_{G_1 \vee \Gamma} \neq Z_{G_2 \vee \Gamma}$ if

$$\prod (d_i + r - 1) \neq \prod (d_i' + r - 1).$$

Consider the polynomial

$$f(x) = \prod (d_i + x) - \prod (d_i' + x),$$

which has degree $< n$. The previous equation holds if and only if $r - 1$ is not a root of $f(x)$, which can only happen for at most $n - 1$ values of r. □

Note, one can replace the bound on the size of S by the number of differing vertex degrees (counting multiplicity) with G_1 and G_2. Also, this proposition gives a non-algorithmic proof of Lemma 2.3 by showing that any two order n graphs with different degree sequences can by distinguished by the zeta functions of n joins with graphs of distinct orders.

Example 4.7. *Let* G_1 *and* G_2 *be* `ECZo` *and* `EEr_`. *These are graphs on 6 vertices with the same Laplacian spectra, but* G_1 *has degree sequence* $(4, 2, 2, 2, 2, 2)$ *and* G_2 *has degree sequence* $(3, 3, 3, 2, 2, 1)$.

For any (non-null) graph Γ, *the graphs* $G_1 \vee \Gamma$ *and* $G_2 \vee \Gamma$ *are automatically md2. The polynomial* $f(x)$ *in the previous proof is just* $f(x) = (4+x)(2+x)^5 - (3+x)^3(2+x)^2(1+x) = (2+x)^2(2x+5)$. *This has no positive roots, so* $Z_{G_1 \vee \Gamma} \neq Z_{G_2 \vee \Gamma}$ *for any* Γ.

The same conclusion is true if we replace G_1 *and* G_2 *by their complements.*

Corollary 4.8. *There are at least* $2g_{n-6}$ *pairs of graphs* (G_1, G_2) *of order* n *which have the same Laplacian spectra but different Ihara zeta functions.*

4.4. **A new construction.** There is one pair of non-isomorphic md2 graphs on 9 vertices which have the same generalized characteristic polynomial φ_{ADJ}, the pair from Example 2.1. This is the smallest such example. This pair does not occur as the result of GM* switching for the following reason: the matrices $A + xD$ are similar for all x, but there is no matrix P such that $P^{-1}(A_1 + xD_1)P = (A_2 + xD_2)$ for all x, as must be the case for GM* pairs. (In [SS13], it was mistakenly written that all pairs of connected md2 graphs on $n \leq 11$ vertices with the same zeta function, adjacency spectrum, and Laplacian spectra are obtained by GM* switching, but this is false as this example shows. The second author of that paper informed us that the conclusion of that sentence, "are obtained by GM* switching," should be "have the same φ_{ADJ}.")

There do however exist matrices P and R such that $(A_1+xD_1)(P+xR) = (P+xR)(A_2+xD_2)$ and $P+xR$ commutes with J for all x and is invertible for all real valued x. Here we prove this example is part of a more general construction of pairs of graphs with the same φ_{ADJ} which cannot be explained by GM* switching. The fact that the conjugating matrix $P + xR$ needs to depend on x makes this construction more delicate (and complicated) than GM* switching (and as far as we know, there are no other such constructions). So we first describe it in a concrete way, and then remark afterwards what are the necessary abstract conditions for this construction to work.

Construction 4.9. *Consider adjacency matrices A_1, A_2 of the following form:*

$$A_1 = \begin{bmatrix} B_1 & B_{12} & \cdots & B_{1k} & N_1 \\ {}^tB_{12} & B_2 & \cdots & B_{2k} & N_2 \\ \vdots & & \ddots & & \vdots \\ {}^tB_{1k} & \cdots & {}^tB_{k-1,k} & B_k & N_k \\ {}^tN_1 & \cdots & {}^tN_{k-1} & {}^tN_k & C \end{bmatrix}, A_2 = \begin{bmatrix} B_1 & {}^tB_{12} & \cdots & {}^tB_{1k} & N_1 \\ B_{12} & B_2 & \cdots & {}^tB_{2k} & N_2 \\ \vdots & & \ddots & & \vdots \\ B_{1k} & \cdots & B_{k-1,k} & B_k & N_k \\ {}^tN_1 & \cdots & {}^tN_{k-1} & {}^tN_k & C \end{bmatrix}$$

Here we allow the B_i's to be chosen from among the following 4×4 adjacency matrices given in 2×2 block form by:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & I \\ I & w \end{bmatrix}, \begin{bmatrix} 0 & J \\ J & 0 \end{bmatrix}, \begin{bmatrix} w & 0 \\ 0 & w \end{bmatrix}, \begin{bmatrix} w & J \\ J & w \end{bmatrix},$$

where $w = J - I$. The B_{ij}'s ($i < j$) are allowed to be arbitrarily chosen from among the following 0-1 matrices with constant row and column sums:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} J & J \\ J & J \end{bmatrix}, \begin{bmatrix} 0 & J \\ J & 0 \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & J \end{bmatrix}, \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \begin{bmatrix} w & 0 \\ 0 & w \end{bmatrix}, \begin{bmatrix} 0 & I \\ w & 0 \end{bmatrix}, \begin{bmatrix} 0 & w \\ I & 0 \end{bmatrix},$$

$$\begin{bmatrix} I & J \\ J & I \end{bmatrix}, \begin{bmatrix} w & J \\ J & w \end{bmatrix}, \begin{bmatrix} J & w \\ I & J \end{bmatrix}, \begin{bmatrix} J & I \\ w & J \end{bmatrix}, \begin{bmatrix} I & I \\ w & I \end{bmatrix}, \begin{bmatrix} I & w \\ I & I \end{bmatrix}, \begin{bmatrix} w & I \\ w & w \end{bmatrix}, \begin{bmatrix} w & w \\ I & w \end{bmatrix}.$$

We allow C to be any $m \times m$ adjacency matrix, and the N_i's are $4 \times m$ matrices such that each column of N_i consists of all zeros or all ones.

Then $\varphi_{ADJ}^{G_1} = \varphi_{ADJ}^{G_2}$, where G_i is the graph with adjacency matrix A_i.

Proof. Note that as each B_{ij} has constant row and column sums and each N_i has constant row sums, G_1 and G_2 have the same degree matrix, call this D. Consider the following 4×4 matrices:

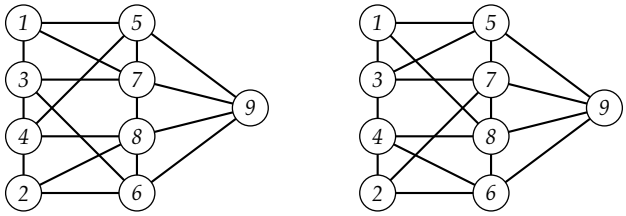
$$Q = \begin{bmatrix} w & w-I \\ w-I & I \end{bmatrix}, S = \frac{1}{2}\begin{bmatrix} I-w & 0 \\ 0 & w-I \end{bmatrix}$$

Let P be the block diagonal matrix whose first k blocks are the matrix Q and last block is the $m \times m$ identity matrix. Let R be the block diagonal matrix whose first k blocks are the matrix S and last block is the $m \times m$ zero matrix.

One can check that $(A_1 + xD)(P + xR) = (P + xR)(A_2 + xD)$ for all x. Note that P has constant row and column sums of 1 and R has constant row and column sums of 0 so $P + xR$ commutes with J for all x. Therefore $(A_1 + xD + yJ)(P + xR) = (P + xR)(A_2 + xD + yJ)$ for all x, y. Also, $\det(P + xR) = \det(Q + xS)^k = (-x^2 + 2x - 5)^k$ which is nonzero for all real x. It follows that $A_1 + xD + yJ$ and $A_2 + xD + yJ$ have the same eigenvalues for all x, y so G_1 and G_2 have the same φ_{ADJ}. □

Example 4.10. *The pair of graphs G_1 and G_2 from Example 2.1 can be obtained from the above construction as follows. Take $k = 2$ and $B_1 = B_2 = \begin{bmatrix} 0 & I \\ I & w \end{bmatrix}$ and $B_{12} = \begin{bmatrix} I & I \\ w & I \end{bmatrix}$. Take C to be the 1×1 zero matrix, N_1 to be the 4×1 zero vector and N_2 to be the 4×1 all ones vector. The resulting graphs*

are the graphs from Example 2.1, which we redraw below with the vertices labeled in the order that they appear in this description.



If in the previous example, we take B_1, B_2, B_{12} as before but let C be any adjacency matrix and $N_1 = 0$ and $N_2 = J$ (appropriately sized all zero and all ones matrices respectively), then this construction always results in a non-isomorphic pair of graphs. Hence this construction produces at least g_{n-8} non-isomorphic graphs with the same φ_{ADJ} on n vertices. On 10 vertices, there are 6 different non-isomorphic pairs of graphs that can be built using this construction.

This construction can be made more general. Given a set of $n \times n$ adjacency matrices B_i and a set of $n \times n$ matrices B_{ij} with 0 and 1 entries and constant row and column sums, then we can do an analogous construction if there exist $n \times n$ matrices Q and S with the following properties. First, Q and S should have constant row and column sums of 1 and 0 respectively and have $\det(Q + xS)$ not equal to the zero polynomial. If D_i is the degree matrix of M_i, then we need $QB_i = B_iQ$, $D_iS = SD_i$, and $D_iQ - QD_i = SB_i - B_iS$ for all i. Also, we need $B_{ij}Q = Q^tB_{ij}$, ${}^tB_{ij}Q = QB_{ij}$, $B_{ij}S = S^tB_{ij}$, and ${}^tB_{ij}S = SB_{ij}$. If all of these properties are satisfied, then any A_1 and A_2 built in the same manner as in the construction will result in graphs with the same φ_{ADJ}. Note that at least one of the B_{ij} must be non-symmetric for A_1 and A_2 to be different, but this does not guarantee that the resulting graphs will be non-isomorphic.

5. Calculations and Conjectures

In this section, we present our computational study of comparing graphs via the usual spectral methods (A, L and $|L|$) with methods (M1)–(M4). Our findings, together with the above results, lead us to several formal and informal conjectures.

In [HS04], Haemers and Spence enumerated the number of graphs on $n \leq 11$ vertices which are not DS, not L-DS, and not $|L|$-DS (see also [BS09] and [Spe] for some errata and additional data for $n = 12$). This data is summarized in the A, L and $|L|$ columns of Table 5.1. They also enumerated the number of non-DS graphs explained by 2-GM switching and 2-GM* switching. This data is recalled in the 2-GM and 2-GM* columns of Table 5.1. The 2-GM column is a lower bound for the A column, and the 2-GM* will be a lower bound for all other columns in the table.

Our main numerical results are a similar enumeration of the number of graphs on $n \leq 11$ vertices which are not determined by methods (M1)–(M4) (considering these methods separately—considering them in tandem is equivalent to just using (M4)). The $\mathcal{Z}$ column of Table 5.1 gives the number of order n graphs which are not D$\mathcal{Z}$, i.e., not determined by (M3). (Recall that non-D$\mathcal{Z}$ implies non-DS, non-L-DS, and non-$|L|$-DS.) It—surprisingly to us—turned out that for $n \leq 11$ being DZ*, DZ$\bar{\text{Z}}$ and D$\mathcal{Z}\bar{\mathcal{Z}}$ (i.e., not determined by (M1), (M2) or (M4)) are all equivalent. The number of graphs which do not satisfy these conditions is given in the $\mathcal{Z}\bar{\mathcal{Z}}$ column.

This data suggests several things: (i) there is not much difference among (M1)–(M4); (ii) any of (M1)–(M4) seem to do much better than A, L or $|L|$; (iii) most GM-pairs are distinguished by

TABLE 5.1. Counting graphs not determined by spectral invariants (M3) is in the $\mathcal{Z}$ column; (M1), (M2) and (M4) are in the $\mathcal{Z}\bar{\mathcal{Z}}$ column

n	# graphs	A	L	$\lvert L\rvert$	$\mathcal{Z}$	$\mathcal{Z}\bar{\mathcal{Z}}$	2-GM	2-GM*
≤ 5	51	2	0	6	0	0	0	0
6	156	10	4	16	0	0	0	0
7	1,044	110	130	102	0	0	40	0
8	12,346	1,722	1,767	1,201	0	0	1,054	0
9	274,668	51,038	42,595	19,001	2	2	38,258	0
10	12,005,168	2,560,606	1,412,438	636,607	10,146	10,140	2,047,008	9,480
11	1,018,997,864	215,331,676	91,274,836	38,966,935	1,353,426	1,353,402	176,895,408	1,297,220

(M1)–(M4); (iv) most graphs not distinguished by (M1)–(M4) are explained by 2-GM* switching. (We note that for $n = 11$, 216 additional non-D$\mathcal{Z}\bar{\mathcal{Z}}$ graphs are explained by 3-GM* switching, but none for $n \leq 10$.[1] Several more non-D$\mathcal{Z}\bar{\mathcal{Z}}$ graphs for $9 \leq n \leq 11$ may also be explained by Construction 4.9.) These suggestions are congruous with our work in previous sections. We just state one of them as a formal conjecture.

Conjecture 5.1. *We have: (i) almost all graphs are DZ*; and (ii) almost all non-DS graphs are DZ*. Further, (iii) if $\mathcal{H}_n$ (resp. $\mathcal{H}'_n$) is the set of non-DS (resp. non-DZ*) graphs of order n, then $\#\mathcal{H}'_n/\#\mathcal{H}_n \to 0$ as $n \to \infty$.*

The above statements are also true if we replace DZ with any of DZ$\bar{\text{Z}}$, D$\mathcal{Z}$, and D$\mathcal{Z}\bar{\mathcal{Z}}$.*

In light of Haemers' conjecture that almost all graphs are DS, (ii) seems much stronger than (i). Note (iii) is strictly stronger than (ii) as not all DS graphs are DZ*. One reason to believe (ii) is that it appears a significant fraction of non-DS graphs are explained by GM switching. Theorem 4.2 suggests that most of these are distinguished by (M1)–(M4). Further, most of the graphs not distinguished by (M1)–(M4) are explained by GM* switching, which appears to account for just a trivial fraction of graphs obtained by GM switching (see [HS04]). Note this reasoning does not apply to distinguishing graphs which are not L-DS or $|L|$-DS.

In fact, Theorem 4.2 suggests that just the Ihara zeta function Z_G by itself might differentiate most pairs of graphs with the same adjacency spectrum. However, if we are just looking at Z_G, it is more reasonable to restrict to md2 graphs. Let us say an md2 graph is DZ if there is no other md2 graph of the same order with the same Ihara zeta function.

Setyadi–Storm [SS13] enumerated all pairs of connected md2 graphs with the same Ihara zeta function for $n \leq 11$. This does not exactly tell us the number of non-DZ md2 graphs (even among connected graphs), so we did a similar enumeration to Table 5.1 for md2 graphs with $n \leq 10$. The results are in Table 5.2, with the column headings meaning the same things as in Table 5.1, except restricted to md2 graphs. The only new column is the Z column, which is the number of non-DZ md2 graphs on n vertices.

We remark that while computing this data, we discovered some small errors in the tables in [SS13] for $n = 11, 12$. Namely, in the cases where more than 2 connected md2 graphs have the same Ihara zeta function, [SS13] undercounts the number of pairs. For instance, for $n = 10$, the entries in the [SS13] tables should be augmented by 1 for $m = 20, 21, 24, 25$. In each of these cases, there is one triple of connected md2 graphs all with the same zeta function.

[1] To enumerate graphs obtained by 3-GM* switching, we observe that to get a non-isomorphic switch $\widetilde{G}$ from G, one needs at least one of the B_i's to be even of size ≥ 4, say B_1. One may also assume B_2 has size > 1, and the size of B_2 is not relatively prime to the size of B_1. Then for each G which is not distinguished by φ_{ADJ}, one can check if it satisfies the 3-GM* condition by iterating first through all "admissible" choices for V_{B_1}, then through choices for V_{B_2}.

TABLE 5.2. Counting md2 graphs not determined by spectral invariants

n	# graphs	A	L	$\lvert L\rvert$	Z	$\mathcal{Z}$	$\mathcal{Z}\bar{\mathcal{Z}}$
≤ 6	77	0	0	4	0	0	0
7	510	26	64	37	0	0	0
8	7,459	744	1,156	725	2	0	0
9	197,867	32,713	31,353	13,878	6	2	2
10	9,808,968	1,727,629	1,184,460	535,080	10,130	10,094	10,088

This data, together with our work above, suggests that, when restricting to md2 graphs, there is little difference between methods (M1)–(M4) and just using the Ihara zeta function. Hence we are led to the following conjecture.

Conjecture 5.2. *Almost all md2 graphs are DZ, and almost all non-DS md2 graphs are DZ. In fact, if $\mathcal{H}_n^{(2)}$ (resp. $\mathcal{H}_n'^{(2)}$) is the set of non-DS (resp. non-DZ) md2 graphs of order n, then $\#\mathcal{H}_n'^{(2)}/\#\mathcal{H}_n^{(2)} \to 0$ as $n \to \infty$.*

Since almost all graphs are md2 for n large, this first statement would imply that almost all graphs of order n are determined by their Ihara zeta function.

We performed the above calculations by generating the graphs with fixed n and m using nauty [nauty] in Sage 6.1.1 [Sage]. Then we used Sage to compute $\det|L| = \lim_{t\to -1}(1-t^2)^{n-m}Z_G(-1)^{-1}$ and $\det(4D+2A-3I) = (-3)^{n-m}Z_G(-2)^{-1}$ for each G, and similarly for G^* and $\bar{G}$. By sorting the graphs according to these values for G^* or G and $\bar{G}$, we made a list of candidate non-DZ* and non-DZ$\bar{\text{Z}}$ graphs. For each of these pairs, we check to see if the corresponding Ihara zeta functions match up. We compute the Ihara zeta functions by first constructing $G^\dagger$, then using the Hashimoto determinant formula. This is much faster than using Sage 6.1.1's built-in function to compute $Z_G(t)^{-1}$. To conserve memory, for $n = 10, 11$, we wrote out the data of the special values of the zeta function with the graph6 string to a file for fixed n and m, and sorted these files using the Unix `sort` tool. Then for each pair of graphs with the same Z_G and $Z_{\bar{G}}$, we compared their generalized characteristic polynomials φ_{ADJ}. This takes care of (M1), (M2) and (M4). We will return to (M3) momentarily.

This method of making a first pass by checking just 2 values of the zeta function $Z_G(-1)$ and $Z_G(-2)$ is based on the heuristic idea that it is unlikely that 2 graphs G_1 and G_2 will have the same spectra for several different random linear combinations of A and D, unless they have the same spectra for all linear combinations of A and D. (This is part of the reason we believe the Ihara zeta function has essentially the same distinguishing power as the Bartholdi zeta function.) From some computational experimentation, we discovered that just knowing n, m and the 2 values $Z_G(-1)$ and $Z_G(-2)$ almost always determines $Z_G(t)$ (for $n \leq 11$).

This heuristic also suggests that using 2 independent spectra to distinguish graphs is much better than a single one. To further test this idea, we enumerated graphs which are not determined by the following sets of spectra: A; A and L; A, L and $|L|$; all three of these plus the Ihara zeta function, and finally φ_{AD}. These numbers are respectively given in the 3rd through 7th columns of Table 5.3.

Note that even though the numbers of graphs not distinguished by A or L or $|L|$ individually are quite large (cf. Table 5.1), the number of graphs not distinguished by combining 2 or more of these shrinks drastically. (We did not compute the numbers for combining A and $|L|$ or L and $|L|$ for all $n \leq 11$, but we expect a similar phenomenon to hold in these cases also. For instance when $n = 9$, there are only 8 graphs not distinguished by A and $|L|$, and 4,405 graphs not distinguished by L and $|L|$.) This suggests that, say, using A and either L or $|L|$, to distinguish graphs is closer in effectiveness to using (M1)–(M4) than to using just A or just L or just $|L|$.

TABLE 5.3. Counting graphs not determined by combining spectral invariants

n	# graphs	A	AL	$AL\|L\|$	$AL\|L\|Z$	$\mathcal{Z}$
≤ 5	51	2	0	0	0	0
6	156	10	0	0	0	0
7	1,044	110	0	0	0	0
8	12,346	1,722	0	0	0	0
9	274,668	51,038	82	2	2	2
10	12,005,168	2,560,606	13,948	10,718	10,150	10,146
11	1,018,997,864	215,331,676	1,468,790	1,361,246	1,353,498	1,353,426

The calculations for Table 5.3 were obtained by comparing spectra via spectral moments, and successively sieving out graphs. E.g., once we know all pairs or graphs with the same A and L spectra, we search through these to see which also have the same $|L|$ spectra. To compute the last column (i.e., (M3)), we looked at all pairs with the same A, L, $|L|$ spectra and Ihara zeta functions, and checked their generalized characteristic polynomials φ_{AD}.

We conclude with a couple of final remarks about related calculations.

First, Brouwer and Spence [BS09] find that quite large families can have the same adjacency spectrum (46 graphs can have the same spectrum on 11 vertices, and this goes up to 128 on 12 vertices). It seems much rarer for a large family of graphs to have the same zeta functions. We found that on 10 vertices there are 4 triples of graphs with the same φ_{ADJ}, and no larger families. On 11 vertices, there are 1,442 triples of graphs with the same φ_{ADJ}, and 192 quadruples, but no larger families.

Second, one might wonder about using cones or complements with the usual spectra. Knowing the Laplacian spectrum of G is equivalent to knowing it for $\bar{G}$ or G^*, but what about A or $|L|$? From Corollary 3.3, we know that knowing the A or $|L|$ spectrum of G and $\bar{G}$ implies the same for G^* (and conversely knowing the A or $|L|$ spectrum of G^* and $\overline{G^*}$ implies the same for G). In Table 5.4, we enumerate the graphs G on $n \leq 10$ vertices which are not determined by the following: spectrum of G^*, spectra of G and $\bar{G}$, $|L|$-spectrum of G^* and $|L|$-spectrum of G and $\bar{G}$, listed respectively in columns 3 through 6. The data in the $A\bar{A}$ columns is already in [HS04]. We note there has been recent work towards showing almost all graphs are distinguished by the A- and $\bar{A}$- spectra—see, e.g., [Wan13] or [vDH09].

TABLE 5.4. Counting graphs not determined by cones and complements for A and $|L|$

n	# graphs	A^*	$A\bar{A}$	$\|L\|^*$	$\|L\|\overline{\|L\|}$
≤ 4	17	0	0	2	2
5	34	0	0	4	4
6	156	0	0	16	16
7	1,044	44	40	102	102
8	12,346	1,194	1,166	1,139	1,139
9	274,668	44,120	43,811	18,748	18,748
10	12,005,168	2,423,121	2,418,152	633,232	633,226

We remark that for A- or $|L|$- spectra, there appears to be little difference between using G and $\bar{G}$ versus G^* versus just G (cf. Table 5.1; see [HS04], [Spe] for the $A\bar{A}$ data when $n = 11, 12$). However, unlike the case of L-spectra, there is at least some difference. In particular, the A- or $|L|$- spectrum of G^* does not always determine that of G, but Table 5.4 suggests it usually does (particularly for $|L|$).

The coincidence of methods (M1) and (M2) on $n \leq 11$ vertices means that for graphs with $n \leq 11$ vertices of size m, Z_{G^*} always determines Z_G and $Z_{\bar{G}}$, and vice versa. Consequently, for $n \leq 11$ vertices, knowing Z_{G^*} is equivalent to knowing $Z_{(\bar{G})^*}$. By analogy with the A-, L-, and $|L|$- spectra, Corollary 3.3 suggests that Z_G and $Z_{\bar{G}}$ (at least almost) always determines Z_{G^*}, and Table 5.4 suggests that Z_{G^*} almost always determines Z_G and $Z_{\bar{G}}$. That is, we expect (M1) and (M2) to be almost always equivalent, but there seems to be no reason to expect this is always the case. Similarly, we have no reason to expect (M2) and (M4) always give the same results.

REFERENCES

[Bar99] Laurent Bartholdi, *Counting paths in graphs*, Enseign. Math. (2) **45** (1999), no. 1-2, 83–131.

[Bas92] Hyman Bass, *The Ihara-Selberg zeta function of a tree lattice*, Internat. J. Math. **3** (1992), no. 6, 717–797.

[BS13] Paymun Bayati and Marius Somodi, *On the Ihara zeta function of cones over regular graphs*, Graphs Combin. **29** (2013), no. 6, 1633–1646.

[BH12] Andries E. Brouwer and Willem H. Haemers, *Spectra of graphs*, Universitext, Springer, New York, 2012.

[BS09] A. E. Brouwer and E. Spence, *Cospectral graphs on 12 vertices*, Electron. J. Combin. **16** (2009), no. 1, Note 20, 3.

[Bus10] Peter Buser, *Geometry and spectra of compact Riemann surfaces*, Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, 2010. Reprint of the 1992 edition.

[CRS10] Dragoš Cvetković, Peter Rowlinson, and Slobodan Simić, *An introduction to the theory of graph spectra*, London Mathematical Society Student Texts, vol. 75, Cambridge University Press, Cambridge, 2010.

[Coo09] Yaim Cooper, *Properties determined by the Ihara zeta function of a graph*, Electron. J. Combin. **16** (2009), no. 1, Research Paper 84, 14.

[Cza05] Debra Czarneski, *Zeta functions of finite graphs*, ProQuest LLC, Ann Arbor, MI, 2005. Thesis (Ph.D.)–Louisiana State University and Agricultural & Mechanical College.

[GM82] C. D. Godsil and B. D. McKay, *Constructing cospectral graphs*, Aequationes Math. **25** (1982), no. 2-3, 257–268.

[HS04] Willem H. Haemers and Edward Spence, *Enumeration of cospectral graphs*, European J. Combin. **25** (2004), no. 2, 199–211.

[Has89] Ki-ichiro Hashimoto, *Zeta functions of finite graphs and representations of p-adic groups*, Automorphic forms and geometry of arithmetic varieties, Adv. Stud. Pure Math., vol. 15, Academic Press, Boston, MA, 1989, pp. 211–280.

[Has90] ———, *On zeta and L-functions of finite graphs*, Internat. J. Math. **1** (1990), no. 4, 381–396.

[Has92] ———, *Artin type L-functions and the density theorem for prime cycles on finite graphs*, Internat. J. Math. **3** (1992), no. 6, 809–826.

[Iha66] Yasutaka Ihara, *On discrete subgroups of the two by two projective linear group over* $\mathfrak{p}$*-adic fields*, J. Math. Soc. Japan **18** (1966), 219–235.

[JN80] Charles R. Johnson and Morris Newman, *A note on cospectral graphs*, J. Combin. Theory Ser. B **28** (1980), no. 1, 96–103.

[KL08] Hye Kyung Kim and Jaeun Lee, *A generalized characteristic polynomial of a graph having a semifree action*, Discrete Math. **308** (2008), no. 4, 555–564.

[nauty] Brendan D. McKay and Adolfo Piperno, *Practical graph isomorphism, II*, J. Symbolic Comput. **60** (2014), 94–112.

[Sage] W. A. Stein and others, *Sage Mathematics Software (Version 6.1.1)*, The Sage Development Team, 2014. `http://www.sagemath.org`.

[Sch73] Allen J. Schwenk, *Almost all trees are cospectral*, New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich., 1971), Academic Press, New York, 1973, pp. 275–307.

[SS08] Geoffrey Scott and Christopher Storm, *The coefficients of the Ihara zeta function*, Involve **1** (2008), no. 2, 217–233.

[SS13] A. Setyadi and C. K. Storm, *Enumeration of graphs with the same Ihara zeta function*, Linear Algebra Appl. **438** (2013), no. 1, 564–572.

[Spe] E. Spence, *Personal homepage*. `http://www.maths.gla.ac.uk/~es`, accessed Jul 15, 2014.

[ST96] H. M. Stark and A. A. Terras, *Zeta functions of finite graphs and coverings*, Adv. Math. **121** (1996), no. 1, 124–165.

[Ter11] Audrey Terras, *Zeta functions of graphs*, Cambridge Studies in Advanced Mathematics, vol. 128, Cambridge University Press, Cambridge, 2011. A stroll through the garden.

[vDH03] Edwin R. van Dam and Willem H. Haemers, *Which graphs are determined by their spectrum?*, Linear Algebra Appl. **373** (2003), 241–272. Special issue on the Combinatorial Matrix Theory Conference (Pohang, 2002).

[vDH09] ———, *Developments on spectral characterizations of graphs*, Discrete Math. **309** (2009), no. 3, 576–586.

[Wan13] Wei Wang, *Generalized spectral characterization of graphs revisited*, Electron. J. Combin. **20** (2013), no. 4, Paper 4, 13.

[WLLX11] Wei Wang, Feng Li, Hongliang Lu, and Zongben Xu, *Graphs determined by their generalized characteristic polynomials*, Linear Algebra Appl. **434** (2011), no. 5, 1378–1387.

(Christina Durfee) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OKLAHOMA, NORMAN, OK 73019
E-mail address: cdurfee@math.ou.edu

(Kimball Martin) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OKLAHOMA, NORMAN, OK 73019
E-mail address: kmartin@math.ou.edu