INDISTINGUISHABLE TREES AND GRAPHS

STEPHAN WAGNER AND HUA WANG

Abstract. We show that a number of graph invariants are, even combined, insufficient to distinguish between nonisomorphic trees or general graphs. Among these are: the set of eigenvalues (equivalently, the characteristic polynomial), the number of independent sets of all sizes or the number of connected subgraphs of all sizes. We therefore extend the classical theorem of Schwenk that almost every tree has a cospectral mate, and we provide an answer to a question of Jamison on average subtree orders of trees. The simple construction that we apply for this purpose is based on finding graphs with two distinguished vertices (called pseudo-twins) that do not belong to the same orbit but whose removal yields isomorphic graphs.

1. Introduction

A number of graph-theoretical problems deal with the reconstruction of a graph from certain given information. By a *graph invariant*, we mean a map whose domain is the set of all graphs, and that is invariant under isomorphism. Examples include:

- The (ordered) *degree sequence*,
- The *characteristic polynomial* and the *spectrum*,
- The *chromatic number*, the *independence number*, and the *matching number*,

to name but a few. In all these instances, knowledge of the specific graph invariant is generally not sufficient to uniquely determine the underlying graph. It is a classical result of Schwenk [13], for instance, that a random tree almost surely (with probability tending to 1) has a cospectral mate, i.e., a non-isomorphic tree with the same spectrum. There are, however, many graphs that are uniquely determined by the spectrum.

On the other hand, the famous reconstruction conjecture [6] states that every graph of order at least three can be reconstructed from its deck, i.e., the multiset of its vertex-deleted subgraphs. It has been verified for various classes of graphs, among them trees.

2010 *Mathematics Subject Classification.* 05C60; 05C05, 05C31.

Key words and phrases. Nonisomorphic graphs, trees, characteristic polynomial, matchings, independent sets, connected subgraphs.

This material is based upon work supported financially by the National Research Foundation of South Africa under grant number 70560.

The question whether a graph invariant can be used to distinguish between non-isomorphic graphs also arises outside of mathematics. In chemistry, a number of graph invariants are used as molecular structure-descriptors [14] (and sometimes called *topological indices* in this context). The question how well such an invariant discriminates between non-isomorphic graphs has therefore been raised in the chemical literature as well (see for instance [10, 12]). The typical measure in this context is the ratio between the number of possible values and the number of non-isomorphic graphs in a certain class (e.g., trees of fixed order n). A popular example of a graph invariant in chemistry is the *Wiener index* [4], which is defined as the sum of all distances between vertices:

$$W(G) = \sum_{\{v,w\} \subseteq V(G)} d(v, w).$$

Equivalently, one can consider the average distance, which is $W(G)/\binom{|G|}{2}$.

Let us have a closer look at graph invariants that are polynomials. One of the best-known examples is the *Tutte polynomial* (see for instance [1, Section 9.1]), of which the *chromatic polynomial* is a special case. Other special values of the Tutte polynomial include the number of spanning trees, the number of forests and the number of connected spanning subgraphs. It can be defined via the rank polynomial or by a deletion-contraction process:

- If G has no edges, then $T(G, x, y) = 1$).
- If e is a bridge, then $T(G, x, y) = xT(G \setminus e, x, y)$. If e is a loop, then $T(G, x, y) = yT(G \setminus e, x, y)$.
- If e is neither a bridge nor a loop, then $T(G, x, y) = T(G \setminus e, x, y) + T(G/e, x, y)$. Here, G/e means the graph resulting from contracting e (i.e., identifying its ends).

The characteristic polynomial of a tree is intimately tied to the number of matchings – indeed, it is a well-known result that the characteristic polynomial of a tree T of order n is given by

$$\phi(T, x) = \sum_{k \geq 0} (-1)^k m(T, k) x^{n-2k},$$

where $m(T, k)$ is the number of matchings of order k in T. This polynomial is also called the *matching polynomial* (see for instance [3, Chapter 4]) for general graphs, and it is related to the matching-generating polynomial

$$M(T, x) = \sum_{k \geq 0} m(T, k) x^k.$$

In a similar fashion, one defines the independence polynomial [11] of a graph G by

$$I(G, x) = \sum_{k \geq 0} i(G, k) x^k,$$

where $i(G,k)$ is the number of independent sets of cardinality k in G. The *Hosoya polynomial* [7] is closely related to the aforementioned Wiener index and given by

$$H(G,x) = \sum_{\{v,w\}\subseteq V(G)} x^{d(v,w)}.$$

Clearly, one has $H'(G,1) = W(G)$. Yet another similar polynomial,

$$S(T,x) = \sum_{k\geq 1} s(T,k)x^k,$$

where $s(T,k)$ is the number of subtrees of cardinality k in a tree T, occurs in a paper of Jamison [8], who studies the average number of nodes in a subtree of a tree that can be expressed as the logarithmic derivative $S'(T,1)/S(T,1)$. Since the matching polynomial is not sufficient to determine a tree, what about the independence polynomial or the subtree polynomial? Indeed, Jamison poses the question (which still seems to be open) whether two non-isomorphic trees of the same order always have a different average number of nodes in a subtree. The aim of this paper is to give a simple construction which shows that even all three polynomials combined (plus possibly other information) are insufficient to reconstruct a tree. Therefore, in particular, the answer to Jamison's question is negative: there are two non-isomorphic trees T and T' for which $S(T,x) = S(T',x)$, hence the average order of a subtree is the same in T and T'.

Before we prove this statement, let us first review the idea of Schwenk in his proof of the aforementioned result on cospectral mates (see [2] for an excellent treatment as well as other constructions yielding cospectral graphs): the key is to take two trees T_1 and T_2 with distinguished vertices v_1 and v_2 such that the characteristic polynomials agree:

$$\phi(T_1,x) = \phi(T_2,x) \qquad \text{and} \qquad \phi(T_1 \setminus v_1,x) = \phi(T_2 \setminus v_2,x).$$

Figure 1 shows an example. Let R be an arbitrary tree and w be a vertex of R. We consider the two trees U_1 and U_2 obtained by taking the union of T_1 and R (T_2 and R, respectively) and identifying v_1 and w (v_2 and w, respectively), an operation known as *coalescence*. Then, making use of an identity for the characteristic polynomials of graphs [2, Theorem 2.2.3], one obtains

$$\begin{aligned}\phi(U_1,x) &= \phi(T_1 \setminus v_1,x)\phi(R,x) + \phi(T_1,x)\phi(R \setminus w,x) - \phi(T_1 \setminus v_1,x)\phi(R \setminus w,x)\\ &= \phi(T_2 \setminus v_2,x)\phi(R,x) + \phi(T_2,x)\phi(R \setminus w,x) - \phi(T_2 \setminus v_2,x)\phi(R \setminus w,x)\\ &= \phi(U_2,x),\end{aligned}$$

i.e., U_1 and U_2 have the same characteristic polynomial, regardless of the choice of R. Schwenk's theorem is now a result of the fact that a random tree of order n almost surely contains any given rooted subtree as $n \to \infty$ (and thus almost surely T_1, rooted at v_1, which can be replaced by T_2, rooted at v_2, to obtain a non-isomorphic tree with the same characteristic polynomial). In the following, we show that this applies to the independence polynomial and the subtree polynomial as well, and provide a common construction to find such pairs of trees (and other graphs).

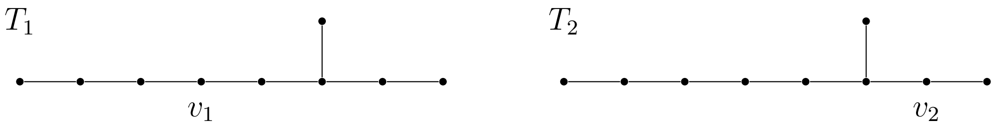

FIGURE 1. Construction of cospectral trees

2. PSEUDO-TWINS

Consider the pair of rooted trees of order 11 shown in Figure 2 below: it is obvious that the two are isomorphic as trees, but not as rooted trees. What is more striking, however, is the fact that the two remain isomorphic as forests if the roots are removed. This fact was was also exploited by Godsil and McKay [5] in their construction of cospectral trees whose complements are still cospectral. The pair is the smallest example of its kind. Let us first give a name to such pairs:

Definition 1. *Let G be a graph and v, w be two vertices such that $G \setminus v$ and $G \setminus w$ are isomorphic, but v and w do not belong to the same orbit of the automorphism group* $\operatorname{Aut} G$*. Then we call v and w* pseudo-twins *in G.*

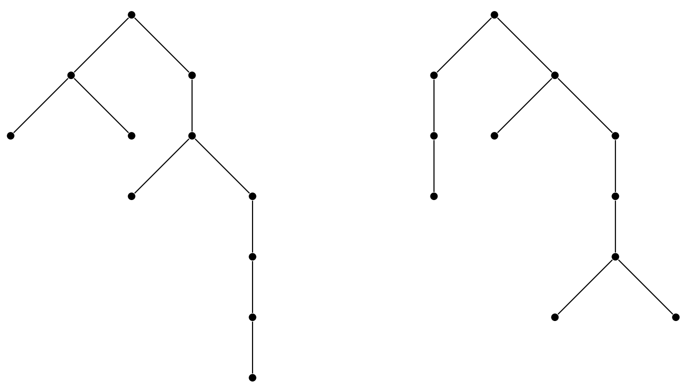

FIGURE 2. A pair of rooted trees.

If G is a tree, then rooting G at two pseudo-twins v and w yields nonisomorphic rooted trees. The importance of such pseudo-twins for our purposes comes from the following fact: let H be an arbitrary graph and u a vertex of H. Let L_v and L_w be the two graphs obtained from G and H by identifying u and v respectively u and w as in Schwenk's construction described in the introduction. Then L_v and L_w are not isomorphic, but they are indistinguishable in many ways: it is obvious, for instance, that they have the same degree sequence (since $G \setminus v$ and $G \setminus w$ have the same number of edges, v and w must have the same degree), and it is easy to see that they also have the same Tutte polynomials, since the vertex where G and H are "glued together" is a cut vertex of the new graph, and one can easily show that $T(L_v, x, y) = T(L_w, x, y) = T(G, x, y)T(H, x, y)$, without any conditions on v and w. As explained in the introduction, L_v and L_w also have the same characteristic polynomial. We show that L_v and L_w coincide in many other graph invariants as well. A number of such invariants will be discussed below, but before we proceed with these graph invariants, let us formulate the non-isomorphism of L_v and L_w as a theorem:

Proposition 2.1. *Suppose that G is a connected graph with two distinguished vertices v and w, and that H is another connected graph of order > 1 with a distinguished vertex u. We obtain the graphs L_v and L_w as the coalescence of G and H by identifying u and v respectively u and w. If L_v and L_w are isomorphic and one of the following two conditions holds:*

- $|H| \geq |G|$,
- *G is a tree,*

then v and w belong to the same orbit of $\mathrm{Aut}\, G$.

Remark 1. *This proposition is not as trivial as it seems, and indeed it is wrong if no additional conditions are imposed. Figure 3 shows a graph G with the property that an edge can be attached to two vertices that are not members of the same orbit, creating isomorphic graphs.*

Proof. Let $\phi : V(L_v) \to V(L_w)$ be an isomorphism, and let $v' = \phi(v)$ be the image of v (identified with u). If $v' = w$, then it is easy to see that the isomorphism also yields an automorphism of G that maps v to w: in this case, we view both L_v and L_w as the coalescence of a multiset of rooted graphs whose roots are identified. Some of these rooted graphs form H, rooted at u, the others form G, rooted at v, and G, rooted at w respectively. Hence there is an automorphism of G that maps v to w.

If $v' \neq w$, let us now assume that $|H| \geq |G|$. Since v is a cut vertex of L_v, v' has to be a cut vertex of L_w. If v' is in $G \setminus w$, then one of the components of $L_w \setminus v'$ contains H as a subgraph, while all components of $L_v \setminus v$ have cardinality $\leq |H| - 1$, a contradiction. Hence v' has to be in H.

Let $X_1, X_2, \dots, X_r$ be induced subgraphs of L_w such that $X_1 \setminus w, X_2 \setminus w, \dots, X_r \setminus w$ are connected components of $L_w \setminus w$ and that their union is G. Likewise, let $Y_1, Y_2, \dots, Y_s$ be induced subgraphs of L_w such that $Y_1 \setminus v', Y_2 \setminus v', \dots, Y_s \setminus v'$ are connected components of $L_w \setminus v'$

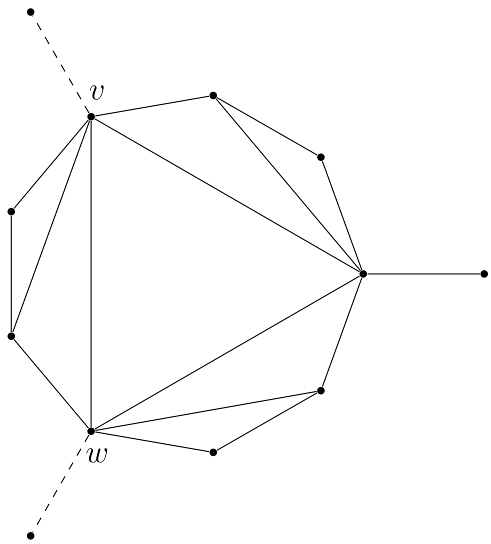

FIGURE 3. A counterexample to Proposition 2.1 if no additional conditions are imposed.

and their union is $\phi(G)$. Finally, let Z be the graph that remains if $X_1 \setminus w, \ldots, X_r \setminus w, Y_1 \setminus v', \ldots, Y_s \setminus v'$ are removed from L_w (see Figure 4).

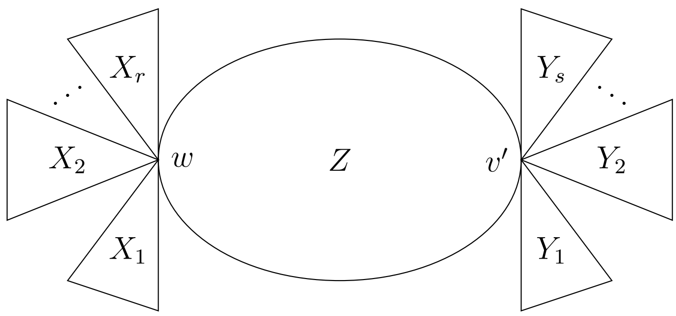

FIGURE 4. Illustration of the proof of Proposition 2.1.

Then the union

$$U_1 = Z \cup X_1 \cup X_2 \cup \cdots \cup X_r,$$

rooted at v', is isomorphic to H, rooted at u. Likewise,

$$U_2 = Z \cup Y_1 \cup Y_2 \cup \cdots \cup Y_s,$$

rooted at w, is also isomorphic to H, rooted at u. Let us now define *ends*: by a (A, a)-end with respect to a vertex b in a graph B, we mean an induced subgraph A' of B with a distinguished vertex a' such that there is an isomorphism from A to A' that maps a to a', and such that $A' \setminus a'$ is a component of $B \setminus a'$ that does not contain b. Since U_1, rooted at v', is isomorphic to U_2, rooted at w, U_2 has to contain the same number of (X_i, w)-ends with respect to w as there are (X_i, w)-ends in U_1 with respect to v' for any $i \in \{1, 2, \ldots, s\}$. Any such ends inside Z

are ends with respect to v' in U_1 as well, so the only possibility is that the set $\{(X_1, w), (X_2, w), \ldots, (X_r, w)\}$ of rooted graphs is isomorphic to the set $\{(Y_1, v'), (Y_2, v'), \ldots, (Y_s, v')\}$, which means that G, rooted at w, is isomorphic to $\phi(G)$, rooted at v', and so v and w belong to the same orbit of $\operatorname{Aut} G$.

Now suppose that $|H| < |G|$, but that G is a tree, and that v and w are not in the same orbit of $\operatorname{Aut} G$. In this case, v' must be in $G \setminus w$ by the same argument as above. H has to be mapped to a subset of G by ϕ, and so H has to be a tree as well. Similar to the ends in the previous part of the proof, we now look at isomorphic copies of H (rooted at u) as induced rooted subtrees in L_v and L_w. No two such copies with distinct roots can overlap, since they would have to contain each other's roots, and if this was the case, L_v would have to be the union of the two copies. But then $|L_v| \leq 2(|H| - 1)$, in contradiction with the assumption $|H| < |G|$.

Now let $v_1 = v$, $v_2 = v' = \phi(v)$. We already know that v_2 is in $G \setminus w$, so $v_3 = \phi(v_2)$ is well-defined, since we can restrict ϕ to G. Moreover, v_2 has to be the root of a copy of H, and so this also holds for v_3. Since copies of H with distinct roots do not overlap, v_3 cannot be in the H-part of L_w, so $\phi(v_3)$ is again well-defined. We repeat the argument to obtain a vertex $v_4 = \phi(v_3)$, etc. The sequence $v_1, v_2 = \phi(v_1), v_3 = \phi(v_2), \ldots$ must ultimately reach w. If not, it would have to return to v, which is impossible since the degree of v in L_v differs from the degree in L_w. If we remove the subtree that is isomorphic to H from each of the vertices $v = v_1, v_2, v_3, \ldots, v_k = w$ in L_v and L_w (in the case of v_1, only in L_v; in the case of v_k, only in L_w), we obtain a tree K on which ϕ becomes an automorphism.

To complete the proof, it remains to prove the following lemma:

Lemma 2.2. *Suppose that there is an automorphism ϕ of a tree K that maps v_i to v_{i+1}, $i = 1, 2, \ldots, k-1$. Then there exists an automorphism τ of K that maps v_1 to v_k and fixes the set $\{v_2, v_3, \ldots, v_{k-1}\}$.*

Proof. Let us assume that K has a single centroid x, the case of a double centroid being similar. Then x has to be a fixed point of ϕ. Let y be the first common vertex of the paths from v_1 to x and from v_2 to x, respectively. Then y has to be a fixed point of ϕ as well, and since $\phi(v_2) = v_3$, y also has to be on the path from v_3 to x, etc. Regarding K as a rooted tree with root x, y is a common ancestor of $v_1, v_2, \ldots, v_k$, and the subtrees rooted at $v_1, v_2, \ldots, v_k$ are all isomorphic as rooted trees. We prove the result by induction on the distance between y and the vertices v_i. If the distance is 1, then the statement is trivial (the automorphism that swaps v_1 and v_k and the subtrees rooted at these two vertices has the desired property).

Now consider the subtrees of K that are rooted at the children of y. All subtrees that contain some of the v_i have to be isomorphic as rooted trees, and they (and possibly others) are somehow permuted by ϕ. Let $\ell > 0$ be minimal with the property that $\phi^\ell(v_1) = v_{\ell+1}$ is in the same subtree as v_k. If v_1 and v_k are in the same subtree, then $k \equiv 1 \mod \ell$, and ϕ^ℓ is an automorphism that maps v_1 to $v_{\ell+1}$, $v_{\ell+1}$ to $v_{2\ell+1}$, etc., so we can apply the induction hypothesis to this subtree.

Otherwise, let R_1 and R_2 be the subtrees that contain v_1 and v_k respectively. There is an automorphism ψ of K that swaps these two subtrees and leaves everything else fixed. We can further choose ψ in such a way that it acts like ϕ^ℓ on R_1 and like $\phi^{-\ell}$ on R_2. For some $r > 0$ and $h \geq 0$, this automorphism maps $v_1, v_{r+1}, v_{2r+1}, \ldots, v_{hr+1}$ to $v_{\ell+1}, v_{r+\ell+1}, v_{2r+\ell+1}, \ldots, v_{hr+\ell+1} = v_k$, in this order. Here, r is the smallest positive integer $< k$ such that $\phi^r(v_1) = v_{r+1}$ is in the same subtree as v_1 (if such an r exists; if not, then this remains true with $h = 0$, and the automorphism ψ already has the desired property). Note that ϕ^r maps v_1 to v_{r+1}, v_{r+1} to v_{2r+1}, etc., so we can apply the induction hypothesis to find an automorphism τ_1 that only acts on the subtree R_1 and maps v_1 to v_{hr+1} while leaving the set $\{v_{r+1}, v_{2r+1}, \ldots, v_{(h-1)r+1}\}$ fixed (if $h = 0$, we may simply choose τ_1 to be the identity). Now we can combine τ_1 and ψ to an automorphism that acts like $\psi \circ \tau_1$ on R_1 and like $\tau_1^{-1} \circ \psi^{-1}$ on R_2 (leaving the other subtrees fixed). This automorphism τ has the desired property. ∎

Returning to the proof of Proposition 2.1, we now have an isomorphism on K that maps $v = v_1$ to $w = v_k$ and keeps the set $\{v_2, v_3, \ldots, v_{k-1}\}$ fixed. It therefore extends to an automorphism of G (once the copies of H that were deleted are added again) that maps v to w, which finally proves the statement. ∎

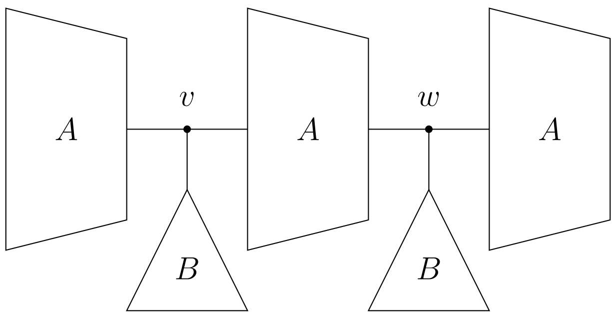

FIGURE 5. Construction of pseudo-twins I.

Having proven this result, the question remains how one can construct graphs with a pair of pseudo-twins. Figures 5 to 7 give three examples of possible constructions with the additional advantage that they yield trees if the "building blocks" are trees; parts of the same shape (and with the same letter) indicate isomorphic copies of the same graph, the two pseudo-twins are denoted by v and w.

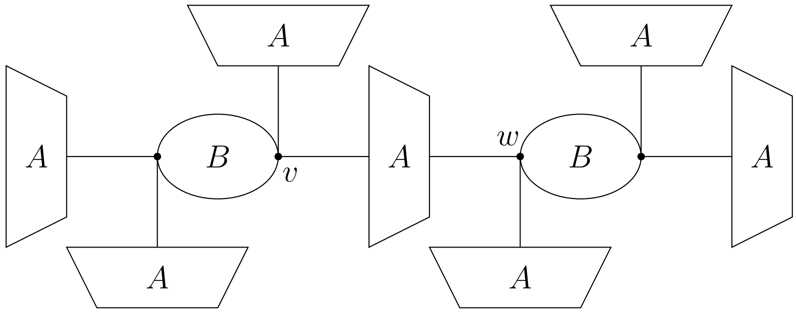

FIGURE 6. Construction of pseudo-twins II.

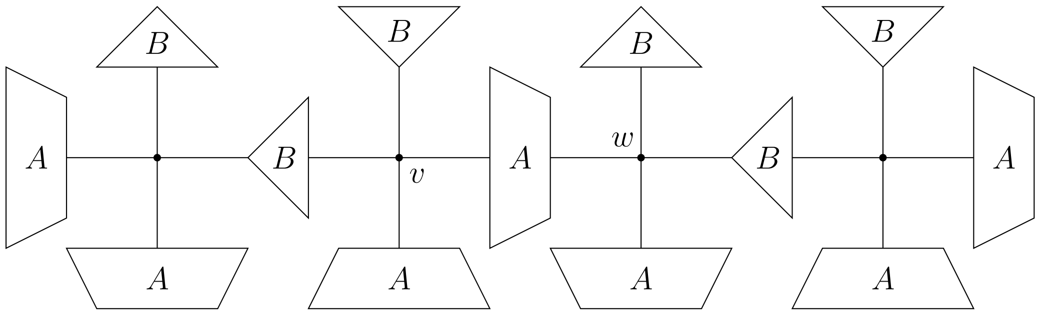

FIGURE 7. Construction of pseudo-twins III.

It is not necessary (though unavoidable for trees) that the pseudo-twins are cut vertices of the graph. Indeed, a simple way to construct such graphs is as follows: let G be any graph with two pseudo-twins v and w, and let H be any other graph. Take the graph join $G' = G \vee H$ (i.e., the graph that results from the union $G \cup H$ by connecting every vertex of G to every vertex of H). It is clear that v and w are still pseudo-twins of G', but no longer cut vertices (see Figure 8 for an example based on the trees in Figure 2).

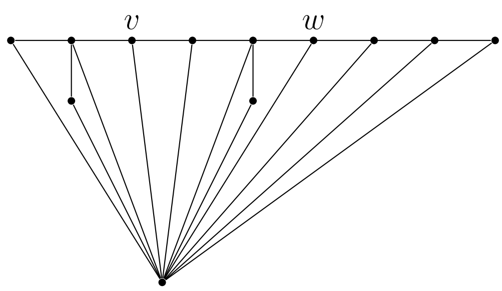

FIGURE 8. Construction of pseudo-twins IV.

Finally, a graph can have more than one pair of pseudo-twins, and a simple way to achieve this is to extend one of the three constructions above, see for example Figure 9, in which v_1, w_1 are pseudo-twins as well as v_2, w_2. It remains an interesting question, though, whether one can characterize exactly all graphs that have a pair of pseudo-twins.

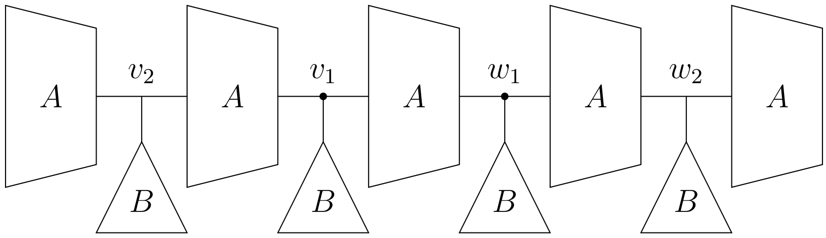

FIGURE 9. Construction of pseudo-twins V.

3. INDISTINGUISHABILITY

Let us now look at some graph invariants whose (even combined) knowledge is not sufficient to reconstruct a graph, and see how the construction of pseudo-twins proves this fact:

Theorem 3.1. *Let v and w be pseudo-twins of a graph G, and let L_v and L_w be defined as in the previous section by taking the coalescence with another graph H. Then*

- *L_v and L_w have the same number of matchings of any given cardinality (and thus in particular the same matching number),*
- *L_v and L_w have the same number of independent sets of any given cardinality (and thus in particular the same independence number),*
- *L_v and L_w have the same number of connected subgraphs (induced or non-induced) of any given order. In particular, if G and H are trees, then L_v and L_w have the same number of subtrees of any given order.*
- *L_v and L_w have the same characteristic polynomial and therefore the same eigenvalues.*

Proof. The four parts are all very similar and are based on recursive formulas for the stated invariants.

- Let us start with the number of matchings. A matching in L_v consists of a matching in G and a matching in H, with the additional requirement that the identified vertex $u = v$ is not covered in both. Let $m(G, k)$ denote the number of matchings of cardinality k in G. It is then easy to see (by means of the inclusion-exclusion principle) that

$$m(L_v, k) = \sum_{j=0}^{k} \Big[m(G, j)m(H \setminus u, k-j) + m(G \setminus v, j)m(H, k-j) - m(G \setminus v, j)m(H \setminus u, k-j) \Big].$$

Since $G \setminus v$ and $G \setminus w$ are isomorphic, it follows immediately that $m(L_w, k) = m(L_v, k)$.

- An independent set in L_v consists of independent sets in G and H, with the additional requirement that either both contain $u = v$ or both do not contain the identified vertex. Thus if $i(G, k)$ denotes the number of independent sets of cardinality k in G, we have

$$i(L_v, k) = \sum_{j=0}^{k} \Big[i(G \setminus v, j) i(H \setminus u, k-j) \\ + (i(G, j) - i(G \setminus v, j))(i(H, k-j) - i(H \setminus u, k-j)) \Big].$$

The same argument as before shows that $i(L_w, k) = i(L_v, k)$.

- Let $c(G, k)$ denote the number of (not necessarily induced) connected subgraphs of order k in G. Since a connected subgraph in L_v can either contain v or not, we obtain

$$c(L_v, k) = c(G \setminus v, k) + c(H \setminus u, k) \\ + \sum_{j=1}^{k} (c(G, j) - c(G \setminus v, j))(c(H, k-j+1) - c(H \setminus u, k-j+1)),$$

so that once again $c(L_w, k) = c(L_v, k)$ for all k. The above formula remains correct if the connected subgraphs are required to be induced.

- Let $\phi(G, x)$ denote the characteristic polynomial of G. Now the same identity [2, Theorem 2.2.3] that was used in the introduction to construct cospectral trees can be used again, and we obtain

$$\phi(L_v, x) = \phi(G, x)\phi(H \setminus u, x) + \phi(G \setminus v, x)\phi(H, x) - x\phi(G \setminus v, x)\phi(H \setminus u, x),$$

from which the statement follows immediately again.

∎

Remark 2. *Note the common form of the recursive relations for the four different parameters: the conclusion of Theorem 3.1 holds for any graph variant with the property that its value for the coalescence of two graphs G and H, where v in G and u in H are identified, can be determined from invariants of G, H, $G \setminus v$ and $H \setminus u$ only.*

Remark 3. *Theorem 3.1 can be generalized further to double coalescences: if H_1 and H_2 are two arbitrary graphs with distinguished vertices u_1 and u_2 respectively, then the graph L_v obtained by identifying u_1 and v as well as u_2 and w is indistinguishable (in the sense of the theorem, i.e., matching polynomial, independence polynomial, ... agree) from the graph L_w obtained by identifying u_1 and w as well as u_2 and v.*

The construction of pseudo-twins can be made in such a way that distance-based graph invariants are taken into account as well. Suppose that v and w are vertices of G such that the number of vertices whose distance from v is k equals the number of vertices whose distance from w is k for any $k \geq 1$. Then L_v and L_w, constructed as before, have the property that the number of pairs of vertices whose distance is k is the same in L_v and L_w for any $k \geq 1$. For the number of pairs within G or within H, this is trivial, and for pairs including one vertex of G and one vertex of H, this follows from the choice of v and w, since any shortest path has to pass through v in L_v and through w in L_w respectively in this case.

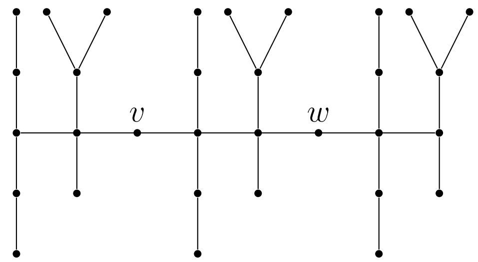

FIGURE 10. Construction of pseudo-twins VI.

It is not difficult to construct a graph with two-pseudo twins that satisfy this additional condition, see Figure 10 for an example of a tree with this property. It follows that L_v and L_w also agree in a number of distance-based graph invariants: the simplest of these is probably the diameter, another is the Wiener index that was mentioned in the introduction, and more generally the Hosoya polynomial. Generalizations of the Wiener index are included as well, such as

$$W_\lambda(G) = \sum_{\{v,w\} \subseteq V(G)} d(v,w)^\lambda,$$

which gives the λ-th moment of the distance upon division by $\binom{|G|}{2}$ (see [4] and the references therein). Another instance is the *hyper-Wiener index* [14]: for two vertices v, w of a tree T, let $n(v, w)$ be the number of vertices for which the unique path to w passes through v (this includes v itself). Then it is easy to prove that

$$WW(T) = \sum_{\{v,w\} \subseteq V(T)} n(v,w)n(w,v) = \sum_{\{v,w\} \subseteq V(T)} \frac{d(v,w)(d(v,w)+1)}{2}.$$

The right hand side makes sense for arbitrary graphs (see [9]) and defines the hyper-Wiener index. It can readily be seen that L_v and L_w, as constructed above, have the same hyper-Wiener index.

4. The special case of trees

Recall that a random tree of order n contains any given rooted subtree with probability tending to 1 as $n \to \infty$. Applying this to any tree with two pseudo-twins that also satisfies the additional condition on distances discussed in the preceding section (such as the tree in Figure 10), we immediately obtain the following result:

Corollary 4.1. *Almost every tree T has a mate T' such that T and T' have*

- *the same degree sequence,*
- *the same number of matchings of any given cardinality,*
- *the same number of independent sets of any given cardinality,*
- *the same number of subtrees of any given order,*
- *the same eigenvalues,*
- *the same number of vertex pairs whose distance is k, for any fixed k.*

As mentioned earlier, the second and fifth item in this list are intimately related, since the coefficients of the characteristic polynomial of a tree are exactly given by the number of matchings of different cardinalities. The list can probably be extended to other graph invariants.

This corollary also shows that the answer to Jamison's question mentioned in the introduction (do two non-isomorphic trees of the same order always have different average subtree order?) is negative, and that actually almost every tree T has a mate T' of the same order such that the average subtree orders of T and T' coincide.

References

[1] M. Aigner. *A course in enumeration*, volume 238 of *Graduate Texts in Mathematics*. Springer, Berlin, 2007.

[2] D. Cvetković, P. Rowlinson, and S. Simić. *An introduction to the theory of graph spectra*, volume 75 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2010.

[3] D. M. Cvetković, M. Doob, I. Gutman, and A. Torgašev. *Recent results in the theory of graph spectra*, volume 36 of *Annals of Discrete Mathematics*. North-Holland Publishing Co., Amsterdam, 1988.

[4] A. A. Dobrynin, R. Entringer, and I. Gutman. Wiener index of trees: theory and applications. *Acta Appl. Math.*, 66(3):211–249, 2001.

[5] C. Godsil and B. McKay. Some computational results on the spectra of graphs. In *Combinatorial mathematics, IV (Proc. Fourth Australian Conf., Univ. Adelaide, Adelaide, 1975)*, pages 73–92. Lecture Notes in Math., Vol. 560. Springer, Berlin, 1976.

[6] F. Harary. A survey of the reconstruction conjecture. In *Graphs and combinatorics (Proc. Capital Conf., George Washington Univ., Washington, D.C., 1973)*, pages 18–28. Lecture Notes in Math., Vol, 406. Springer, Berlin, 1974.

[7] H. Hosoya. On some counting polynomials in chemistry. *Discrete Appl. Math.*, 19(1-3):239–257, 1988.

[8] R. E. Jamison. On the average number of nodes in a subtree of a tree. *J. Combin. Theory Ser. B*, 35(3):207–223, 1983.
[9] D. J. Klein, I. Lukovits, and I. Gutman. On the Definition of the Hyper-Wiener Index for Cycle-Containing Structures. *J. Chem. Inf. Comput. Sci.*, 35(1):5052, 1995.
[10] M. Lepović and I. Gutman. A Collective Property of Trees and Chemical Trees. *J. Chem. Inf. Comput. Sci.*, 38:823–826, 1998.
[11] V. E. Levit and E. Mandrescu. The independence polynomial of a graph—a survey. In *Proceedings of the 1st International Conference on Algebraic Informatics*, pages 233–254, Thessaloniki, 2005. Aristotle Univ. Thessaloniki.
[12] M. Razinger, J. R. Chretien, and J. Dubois. Structural Selectivity of Topological Indexes in Alkane Series. *J. Chem. Inf. Comput. Sci.*, 25(1):23–27, 1985.
[13] A. J. Schwenk. Almost all trees are cospectral. In *New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich., 1971)*, pages 275–307. Academic Press, New York, 1973.
[14] R. Todeschini and V. Consonni. *Handbook of Molecular Descriptors*. Weinheim, 2000).

Stephan Wagner, Department of Mathematical Sciences, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa
E-mail address: swagner@sun.ac.za

Hua Wang, Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460, USA
E-mail address: hwang@georgiasouthern.edu