Journal of Combinatorial Theory, Series B 100 (2010) 671–682

Contents lists available at ScienceDirect

Journal of Combinatorial Theory, Series B

www.elsevier.com/locate/jctb

Spectra of symmetric powers of graphs and the Weisfeiler–Lehman refinements

Afredo Alzaga, Rodrigo Iglesias, Ricardo Pignol

Departamento de Matemática, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina

ARTICLE INFO

Article history:
Received 18 December 2007
Available online 24 July 2010

Keywords:
Graph isomorphism problem
Graph spectra
Weisfeiler–Lehman algorithm
Random walks

ABSTRACT

The k-th power of an n-vertex graph X is the iterated cartesian product of X with itself. The k-th symmetric power of X is the quotient graph of certain subgraph of its k-th power by the natural action of the symmetric group. It is natural to ask if the spectrum of the k-th power – or the spectrum of the k-th symmetric power – is a complete graph invariant for small values of k, for example, for $k = O(1)$ or $k = O(\log n)$.

In this paper, we answer this question in the negative: we prove that if the well-known $2k$-dimensional Weisfeiler–Lehman method fails to distinguish two given graphs, then their k-th powers – and their k-th symmetric powers – are cospectral. As it is well known, there are pairs of non-isomorphic n-vertex graphs which are not distinguished by the k-dim WL method, even for $k = \Omega(n)$. In particular, this shows that for each k, there are pairs of non-isomorphic n-vertex graphs with cospectral k-th (symmetric) powers.

1. Introduction

Many fundamental graph invariants arise from the study of random walks of a particle on a graph. Most of these invariants can be described in terms of the spectrum of the adjacency or the Laplacian matrix. Since the graph spectrum fails to distinguish many non-isomorphic graphs, it is interesting to study the properties of walks (or quantum walks) of k particles, as a means to construct more powerful invariants.

This led Audenaert et al. [1] to define the k-th symmetric power $X^{\{k\}}$ of a graph X: each vertex of $X^{\{k\}}$ represents a k-subset of vertices of X, and two k-subsets are joined if and only if their sym-

E-mail address: iglesiasrodrigo@gmail.com (R. Iglesias).

0095-8956/$ – see front matter
doi:10.1016/j.jctb.2010.07.001

metric difference is an edge of X. They show that the spectra of these graphs is a family of invariants stronger than the ordinary graph spectra. For $k = 2$, they provide examples of cospectral graphs X and Y such that $X^{\{2\}}$ and $Y^{\{2\}}$ are not cospectral. On the other hand, they prove that if X and Y are strongly-regular cospectral graphs then $X^{\{2\}}$ and $Y^{\{2\}}$ are cospectral. For $k = 3$, the authors reported computational evidence suggesting that the spectra of the symmetric cube may be a strong invariant. They did not find any pair of non-isomorphic graphs with cospectral 3-symmetric powers, upon inspection of all strongly regular graphs of up to 36 vertices.

In this paper we prove that for each k there are pairs of non-isomorphic graphs such that their k-th symmetric powers are cospectral by showing how these invariants are related to the well-known k-dimensional Weisfeiler–Lehman (WL) algorithm.

The automorphism group of the graph acts on the set of k-tuples of vertices. The k-WL method is a combinatorial algorithm that attempts to find the associated orbit partition (see, for example, [3,6]). It starts by classifying the k-tuples according to the isomorphism type of their induced graphs, then an iteration is performed attaching to the previous color of a k-tuple the multiset of colors of the neighboring k-tuples. In this way, the partition of the k-tuples is refined in each step until a stable partition is reached. The multiset of colors of the stable partition is a graph invariant.

Our main result is the following theorem.

Theorem 1. *If the 2k-dim Weisfeiler–Lehman algorithm fails to distinguish two given graphs, then their k-th symmetric powers are cospectral.*

In fact, the result remains true if we consider k-th powers of graphs (associated to walks of k labelled particles), instead of symmetric powers.

In [3], Cai, Immerman and Fürer showed how to construct pairs of non-isomorphic n-vertex graphs which are not distinguished by the k-WL method, even for $k = \Omega(n)$. Then, our result implies that

Theorem 2. *If we require the k-th symmetric power spectrum to determine all n-vertex graphs then, necessarily, $k = \Omega(n)$.*

Nevertheless, the spectrum of the k-symmetric power of a graph – we refer to it here as the k-spectrum – is a strong invariant with remarkable computational features. It is determined by the characteristic polynomial of a matrix of polynomial size with entries in $\{0, 1\}$. Then, the problem of determining if two graphs are k-cospectral is in the complexity class NC, that is, it can be computed in polylogarithmic time by a polynomial number of processors running in parallel.

This attribute contrasts with the inherently sequential nature of the k-dim WL algorithm. In [5], Grohe proved that finding the k-dim WL stable partition is a P-complete problem ($k \geqslant 2$). This fact strongly suggests that the k-spectrum is strictly weaker than the 2k-dim WL coloring as a graph invariant.

In effect, if we assume that any two graphs are k-cospectral if and only if they are not distinguished by the 2k-dim WL, then determining graph k-cospectrality is a P-complete problem, according to Grohe's result. This implies that $P = NC$, but this is widely believed to be not true.

Thus, it is interesting to point out where the k-spectrum is strictly weaker than the 2k-dim WL invariant. For $k = 1$, these two invariants have very different performances on trees. Almost all n-vertex trees are 1-cospectral – this is Schwenk's theorem [7] – while non-isomorphic trees are always distinguished by the h-dim WL algorithm for $h \geqslant 1$.

For arbitrary k, there is at least one natural family of candidates: those graphs for which the 2k-WL requires a large number of iterations to reach the stable partition. Actually, in [4], Fürer showed that for any h there are pairs of graphs which are distinguished by the h-WL algorithm, but for which a linear number of iterations are required. We expect that, indeed, a logarithmic number of iterations of the 2k-WL are sufficient to distinguish any pair of graphs with different k-spectra. If this is true, then Fürer's pairs for $h = 2k$ are k-cospectral but distinguished by the 2k-WL invariant.

Besides power graph spectra, there are other families of graph invariants in the literature for which it is not known whether they distinguish any pair of non-isomorphic graphs. As it turns out, the

WL-refinements provide a natural benchmark to compare other graph invariants and it is reasonable to expect that arguments of the kind we use in this work would show the limitations of some of them.

The paper is organized as follows. In Section 2 we define the k-th power X^k and the k-th symmetric power $X^{\{k\}}$ of a graph X. In Section 3 we recall the general notion of quotient of a graph by the action of a group, and we describe the k-th symmetric power as a quotient of the restricted k-th power $X^{(k)}$. For later use, we prove some formulas concerning the walk generating function of quotient graphs. In Section 4 we define precisely the k-Weisfeiler–Lehman algorithm. The heart of the proof of Theorem 1 is in Section 6. Essentially, we show that the $2k$-WL method is stronger than the spectra of the k-th power X^k. Since the idea of the proof is easier to exhibit in the case $k = 1$, we write this special case separately in Section 5. Finally, the proof of Theorem 1 is given in Section 7, by passing to the quotient $X^{\{k\}}$. In order to achieve this, we exploit the structure of the set of k-tuples and the formulas for quotient graphs presented in Section 3.

2. Powers of graphs

In this section we present the notion of the k-th symmetric power of a graph, as introduced in [1], and some other related constructions.

Through the paper, a *graph* G is a finite set V_G of vertices together with a set E_G of unordered pairs (v, w) of vertices with $v \neq w$. We denote by A_G the *adjacency matrix of* G. Since we do not assume an order on V, we consider A_G as a function $A_G : V_G \times V_G \to \mathbb{Z}$, defined by $A_G(v, w) = 1$ if $(v, w) \in E$, and $A_G(v, w) = 0$ otherwise.

A *k-tuple* $(i_1 \dots i_k)$ of vertices is a function from $\{1, \dots, k\}$ to V_G. Let $\mathcal{U}_{G,k}$ be the set of all k-tuples and let $\mathcal{D}_{G,k} \subset \mathcal{U}_{G,k}$ denote the set of those k-tuples of pairwise distinct vertices. The symmetric group S_k acts naturally on $\mathcal{D}_{G,k}$ by $\sigma(i_1 \dots i_k) = (i_{\sigma^{-1}(1)} \dots i_{\sigma^{-1}(k)})$, for $\sigma \in S_k$. The orbits are identified with the *k-subsets* of vertices.

The k-th symmetric power of G, denoted by $G^{\{k\}}$, has the k-subsets of V_G as its vertices; two k-subsets are adjacent if their symmetric difference – elements in their union but not in their intersection – is an edge of G. The picture behind this construction is borrowed from the physical realm: start with k undistinguishable particles occupying k different vertices of G and consider the dynamics of a walk through the graph in which, for each step, any single particle is allowed to move to an unoccupied adjacent vertex. In this way, a k-walk on G corresponds to a 1-walk on $G^{\{k\}}$. The connection between symmetric powers and quantum mechanics exchange Hamiltonians is further explored in [1].

Likewise, one can define the *cartesian product* $G \times H$ of two graphs as follows

$$A_{G\times H}(i_1 i_2, j_1 j_2) = \begin{cases} 1 & \text{if } A_G(i_1, j_1) = 1 \text{ and } i_2 = j_2 \\ & \text{or else } A_G(i_2, j_2) = 1 \text{ and } i_1 = j_1, \\ 0 & \text{otherwise.} \end{cases}$$

The k-th *power* G^k of a graph is defined as the iterated cartesian product of G with itself. The set of its vertices is $\mathcal{U}_{G,k}$ and its adjacency matrix A_{G^k} is given by

$$A_{G^k}(i_1 i_2 \dots i_k, j_1 j_2 \dots j_k) = \begin{cases} 1 & \text{if there exists } u \in \{1, \dots, k\} \text{ such that} \\ & A_G(i_u, j_u) = 1 \text{ and } i_l = j_l \text{ for } l \neq u, \\ 0 & \text{otherwise.} \end{cases}$$

In the physical cartoon of the particles, the k-th power correspond to the situation in which the k particles are labeled, and more than one particle is allowed to occupy the same vertex at the same time.

Given a graph G, the *walk generating function* of G is the power series

$$\sum_{r=0}^{\infty} t^r (A_G)^r.$$

The coefficient of t^r in the (i, j)-entry counts the number of paths of length r from the vertex i to the vertex j. See [1] for further properties. The trace of the walk generating function is a graph invariant, and we denote it by

$$F(G,t) = Tr\sum_{r=0}^{\infty} t^r (A_G)^r.$$

Since the spectrum of two matrices A and B coincides if and only if $Tr(A^r) = Tr(B^r)$ for all r, two graphs G and H are cospectral if and only if $F(G,t) = F(H,t)$. In particular, they cannot be distinguished by the spectrum of their k-th symmetric powers if and only if $F(G^{\{k\}},t) = F(H^{\{k\}},t)$.

3. Quotient graphs

The k-th symmetric power $G^{\{k\}}$ can be constructed from G^k in two steps. First, we cut G^k, deleting all those vertices which are not in $\mathcal{D}_{G,k}$. In this way we obtain the *restricted k-th power*, denoted by $G^{(k)}$, defined as the subgraph of $G^{\{k\}}$ whose vertices are the k-tuples in $\mathcal{D}_{G,k}$. Second, we take the quotient of $G^{\{k\}}$ by the natural action of S_k on the restricted k-th power $G^{(k)}$.

Let us give the general definition of a quotient graph and discuss some properties. Given a graph X and a group Γ acting on X by automorphisms, the quotient X/Γ is a directed graph, in general with multiple edges and loops, defined as follows. The vertices of X/Γ are the orbits of the vertices of X, and given two orbits U and W, there are as many arrows from U to W as edges in X connecting a fixed element $u \in U$ with vertices in W.

We are interested in the case where this quotient has no loops and no multiple edges; we say that the quotient X/Γ is *simply laced* if

1. $(u,v) \in E$ implies that u and v are not in the same orbit.
2. $(u,v) \in E$ and $(u,w) \in E$ implies that v and w are not in the same orbit.

If X/Γ is simply laced, we can consider it an ordinary graph, where (U,W) is an edge if and only if there is an arrow in X/Γ connecting them.

In the simply laced case, every path on X/Γ can be lifted to an essentially unique path on X. This fact simplifies the task of path-counting, and allows to derive a simple formula for the walk generating function of a quotient graph. We apply it to the symmetric power $G^{\{k\}}$ to obtain a formula that will be useful later.

Proposition 1. *Let X be a graph, X/Γ a simply laced quotient, and let U and W be two orbits. Then, the r-th power of the adjacency matrix of X/Γ is given by*

$$A^r_{X/\Gamma}(U,W) = \frac{1}{|U|}\sum_{u\in U}\sum_{w\in W} A^r_X(u,w).$$

Proof. The entry $A^r_{X/\Gamma}(U,W)$ equals the number of paths of length r on X/Γ from U to W. Fix an element $u_0 \in U$ and let $V_0, V_1, V_2, \dots, V_r$ be a path of length r on X/Γ, with $U = V_0$ and $V_r = W$. Since there is at most one edge in X connecting a vertex in X to a vertex in a different orbit, there is a unique path $v_0, v_1, v_2, \dots, v_r$ in G such that $v_0 = u_0$ and $v_j \in V_j$ for $0 \leqslant j \leqslant r$. Then,

$$A^r_{X/\Gamma}(U,W) = \sum_{w\in W} A^r_X(u_0,w).$$

The set of paths of length r from u_0 to W is carried bijectively to the set of paths from any $u \in U$ to W via some automorphism in Γ. Then, the sum

$$\sum_{w\in W} A^r_X(u,w)$$

does not depend on u, and this proves the formula of the proposition. □

Observe that this formula implies that if X/Γ is a connected, simply laced quotient, then all the orbits have the same size.

Let $M_{X/\Gamma}$ be the matrix with rows and columns indexed by the vertices of X, defined by

$$M_{X/\Gamma}(v,w) = \begin{cases} |U| & \text{if } v \text{ and } w \text{ are in the same orbit } U, \\ 0 & \text{otherwise.} \end{cases}$$

From Proposition 1 it follows:

Proposition 2. *Let X/Γ be simply laced quotient, and let $M_{X/\Gamma}$ be defined as above. Then,*

$$Tr\big(A^r_{X/\Gamma}\big) = Tr\big(A^r_X M_{X/\Gamma}\big).$$

Now we set $X = G^{(k)}$ and $\Gamma = S_k$, acting in the natural way on $G^{(k)}$. The quotient $G^{(k)}/S_k$ is isomorphic to the k-th symmetric power $G^{\{k\}}$, and it is easily seen to be a simply laced quotient. In this case, the matrix $M_{X/\Gamma}$ is the matrix M_k, with rows and columns indexed by k-tuples in $\mathcal{D}_{G,k}$, given by

$$M_k(i_1 \dots i_k, j_1 \dots j_k) = \begin{cases} k! & \text{if } \{i_1 \dots i_k\} \text{ and } \{j_1 \dots j_k\} \text{ are equal as sets,} \\ 0 & \text{otherwise.} \end{cases}$$

From Proposition 2 we obtain:

Proposition 3. *Let $G^{(k)}$ and $G^{\{k\}}$ be the restricted k-th power and the k-th symmetric power of a graph G, respectively. Let M_k be the matrix defined as above. Then,*

$$Tr\big(A^r_{G^{\{k\}}}\big) = Tr\big(A^r_{G^{(k)}} M_k\big).$$

4. The Weisfeiler–Lehman algorithm

A natural approach to graph isomorphism testing is to develop algorithms to compute the vertex orbits of the automorphism group of a graph. In particular, if the orbits of the union of two graphs are known, one can decide if there is an isomorphism between them. As a first approximation to the orbit partition of a given graph, one can assign different colors to the vertices according to their degrees. We can refine this partition iteratively, by attaching to the previous color of a vertex, the multiset of colors of its neighbors. After at most $n = |V|$ steps, the partition stabilizes. For most graphs, this method distinguishes all the vertices [2], but it does not work in general. For example, it clearly fails if the vertex degrees are all equal.

A more powerful method, generalizing the previous one, is obtained by coloring the k-tuples of vertices (single vertices are implicit as k repetitions of the same vertex). We start classifying the k-tuples according to the isomorphism type of their induced labelled graphs. Next, we apply an iteration attaching to the previous color of a k-tuple, the multiset of colors of the neighboring k-tuples. This is the so-called k-dimensional Weisfeiler–Lehman refinement. For fixed $k \geqslant 1$ the partition of the k-tuples is no longer refined after n^k steps, so the algorithm runs in polynomial time.

This type of combinatorial methods have been investigated since the seventies, and for some time there was hope in solving the graph isomorphism problem provided that $k = O(\log n)$ or $k = O(1)$. In [3], Cai, Immerman and Fürer, disposed of such conjectures; they proved that, for large n, k must be greater than cn for some constant c, if we require the k-WL refinement to reach the orbit partition of any n-vertex graph. Despite of this limitation, the method works with k constant when restricted to some important families, such as planar or bounded genus graphs [6].

Let us define the k-WL algorithm more precisely. We define an equivalence relation on the set of all k-tuples of all graphs. Let $(i_1 \dots i_k)$ be a k-tuple of vertices of a graph I and $(j_1 \dots j_k)$ a k-tuple of vertices of a graph J. The graphs I and J can be the same. We say that $(i_1 \dots i_k)$ and $(j_1 \dots j_k)$ are equivalent if

1. $i_l = i_{l'}$ if and only if $j_l = j_{l'}$.
2. $(i_l, i_{l'}) \in E_I$ if and only if $(j_l, j_{l'}) \in E_J$.

We define the *type* of a k-tuple as its equivalence class, and we denote it by $tp(i_1 \dots i_k)$. Let S_1 be the set of all different types of k-tuples. This is the initial set of colors. We define the set S of colors by

$$S = \bigcup_{k=1}^{\infty} S_k$$

where elements of S_{r+1} are finite sequences or finite multisets of elements of $\bigcup_{k=0}^{r} S_k$. In practice, it suffices to work with as many colors as k-tuples: in order to preserve the length of their names, the colors can be relabelled in each round (using a rule not depending on the graph). Nevertheless, this relabelling plays no role in our arguments.

We denote the color assignment of the k-WL iteration in its r-th round, applied to the graph G, by $W^r_{G,k} : \mathcal{U}_{G,k} \to S$. Evaluated at the k-tuple $(i_1 \dots i_k)$ it gives the color $W^r_{G,k}(i_1 \dots i_k) \in S$. Initially, for $r = 1$, it is defined by

$$W^1_{G,k}(i_1 \dots i_k) = tp(i_1 \dots i_k).$$

The iteration is given by

$$W^{r+1}_{G,k}(i_1 \dots i_k) = \sum_{m \in V_G} \bigl(tp(i_1 \dots i_k m), S^r_{G,k}(i_1 \dots i_k m)\bigr) \tag{1}$$

where $S^r_{G,k}(i_1 \dots i_k m)$ is the sequence

$$\bigl(W^r_{G,k}(i_1 \dots m), \dots, W^r_{G,k}(i_1 \dots m \dots i_k), \dots, W^r_{G,k}(m \dots i_k)\bigr).$$

The summation symbol in (1) must be interpreted as a formal sum, so that it denotes a multiset. For example, if $x_1 = x_3 = x_4 = a$ and $x_2 = x_5 = b$, then $\sum_{i=1}^{5} x_i$ is the multiset $\{a, a, a, b, b\}$.

For each round, a certain number of different colors is attained. We say that the coloring scheme *stabilizes* in the r-th round if the number of different colors does not increase in the $r + 1$-th iteration.

In order to compare the invariant $F(G^{\{k\}}, t)$ with the k-Weisfeiler–Lehman refinement, we define a graph invariant $I_{G,k}$ which captures the result of the k-WL coloring and, at the same time, it is a combinatorial analogue of $F(G^k, t)$. For each round r, we collect all the resulting colors in the multiset

$$M^r_{G,k} = \sum_{(i_1 \dots i_k) \in \mathcal{U}_{G,k}} W^r_{G,k}(i_1 \dots i_k).$$

Then we define the formal power series

$$I_{G,k}(t) = \sum_{r=0}^{\infty} t^r M^r_{G,k}.$$

The following technical proposition will be used later.

Proposition 4. *Let G and H be two graphs with n vertices. Then, $I_{G,k}(t) = I_{H,k}(t)$ if and only if there is a bijection σ from $\mathcal{U}_{G,k}$ to $\mathcal{U}_{H,k}$ such that*

$$W^r_{G,k}(i_1 \dots i_k) = W^r_{H,k}\bigl(\sigma(i_1 \dots i_k)\bigr)$$

for all $r \geqslant 1$. In particular,

$$tp(i_1 \dots i_k) = tp\bigl(\sigma(i_1 \dots i_k)\bigr).$$

Proof. The *if* part is immediate. Conversely, assume $I_{G,k}(t) = I_{H,k}(t)$. The coefficient of t^r, when $r = n^k$, implies the existence of a bijection σ from $\mathcal{U}_{G,k}$ to $\mathcal{U}_{H,k}$ such that

$$W^{n^k}_{G,k}(i_1 \dots i_k) = W^{n^k}_{H,k}\big(\sigma(i_1 \dots i_k)\big). \tag{2}$$

Whenever Eq. (2) holds for some particular round r_0, it holds for all $1 \leqslant r \leqslant r_0$. Then,

$$W^r_{G,k}(i_1 \dots i_k) = W^r_{H,k}\big(\sigma(i_1 \dots i_k)\big) \tag{3}$$

for all $1 \leqslant r \leqslant n^k$. In addition, since the WL refinement stabilizes after the n^k round, we see that Eq. (3) is true for $r \geqslant n^k$. The last assertion is obtained by setting $r = 1$ in Eq. (3). $\square$

5. Graph spectrum is weaker than the 2-WL refinement

As a warm-up we start by showing that the spectrum of a graph is a weaker invariant than the 2-Weisfeiler–Lehman coloring algorithm. This case displays the essential ingredients of the proof for arbitrary k.

Theorem 3. *Let G and H be two graphs with adjacency matrices A_G and A_H, respectively. If $W^r_{G,2}(i,j) = W^r_{H,2}(p,q)$ then $A^r_G(i,j) = A^r_H(p,q)$.*

Proof. We use induction on the number of rounds r. The base case $(r = 1)$ is trivial. Assume the statement is valid for r, and suppose that

$$W^{r+1}_{G,2}(i,j) = W^{r+1}_{H,2}(p,q).$$

Then, by the definition of the WL coloring,

$$\sum_m \big(tp(i,j,m), W^r_{G,2}(i,m), W^r_{G,2}(m,j)\big) = \sum_m \big(tp(p,q,m), W^{r+1}_{H,2}(p,m), W^{r+1}_{H,2}(m,q)\big).$$

This is an equality of multisets. This means that there exists a permutation σ of $\{1, 2, \dots, n\}$ such that

$$\begin{cases} tp(i,j,m) = tp\big(p,q,\sigma(m)\big), \\ W^r_{G,2}(i,m) = W^r_{H,2}\big(p,\sigma(m)\big), \\ W^r_{G,2}(m,j) = W^r_{H,2}\big(\sigma(m),q\big). \end{cases}$$

By the induction hypothesis, this implies

$$\begin{cases} A_G(i,m) = A_H\big(p,\sigma(m)\big), \qquad A_G(m,j) = A_H\big(\sigma(m),q\big), \\ A^r_G(i,m) = A^r_H\big(p,\sigma(m)\big), \\ A^r_G(m,j) = A^r_H\big(\sigma(m),q\big). \end{cases}$$

Summing over m, we have

$$\sum_m A_G(i,m) A^r_G(m,j) = \sum_m A_H(p,m) A^r_H(m,q),$$

that is, $A^{r+1}_G(i,j) = A^{r+1}_H(p,q)$. $\square$

Theorem 4. *Let G and H be two graphs. If $I_{G,2}(t) = I_{H,2}(t)$, then G and H are cospectral.*

Proof. Assume $I_{G,2}(t) = I_{H,2}(t)$. By Proposition 4, there is a bijection σ between the respective sets of 2-tuples, such that for every 2-tuple ij,

$$W^r_{G,2}(ij) = W^r_{H,2}\big(\sigma(ij)\big)$$

for $r \geqslant 1$. When $r = 1$, this is

$$tp(ij) = tp\big(\sigma(ij)\big).$$

In particular, σ sends the diagonal of $W^r_{G,2}$ to the diagonal of $W^r_{H,2}$, that is,

$$\sigma(ii) = pp$$

for some element p. Then, collecting all the colors in the diagonal, we have

$$\sum_i W^r_{G,2}(ii) = \sum_i W^r_{H,2}\big(\sigma(i)\sigma(i)\big).$$

By Theorem 3, this implies

$$\sum_i A^r_G(i,i) = \sum_i A^r_H\big(\sigma(i),\sigma(i)\big)$$

that is, $Tr\, A^r_G = Tr\, A^r_H$ for $r \geqslant 1$. Then, $F(G,t) = F(H,t)$ and this means that G and H are cospectral. $\square$

6. Spectra of *k*-th powers

For each round r, we think of the $2k$-WL coloring as a matrix of colors: the rows and columns are indexed by k-tuples, with the color $W^r_{G,k}(i_1 \dots i_k\ j_1 \dots j_k)$ in the entry $(i_1 \dots i_k, j_1 \dots j_k)$.

Theorem 5. *Let G^k and H^k be the k-th powers of two graphs G and H, respectively. Let $A^r_{G^k}$ and $A^r_{H^k}$ be the r-th powers of their adjacency matrices. If*

$$W^r_{G,2k}(i_1 \dots i_k\ j_1 \dots j_k) = W^r_{H,2k}(p_1 \dots p_k\ q_1 \dots q_k),$$

then

$$A^r_{G^k}(i_1 \dots i_k, j_1 \dots j_k) = A^r_{H^k}(p_1 \dots p_k, q_1 \dots q_k).$$

Proof. The proof goes along the lines of Theorem 3. Let $r = 1$. Suppose that

$$A_{G^k}(i_1 \dots i_k, j_1 \dots j_k) = 1.$$

Then $i_l = j_l$ for all l except for a unique value l_0, for which $A_G(i_{l_0}, j_{l_0}) = 1$. By hypothesis,

$$W^1_{G,2k}(i_1 \dots i_k\ j_1 \dots j_k) = W^1_{H,2k}(p_1 \dots p_k\ q_1 \dots q_k),$$

that is,

$$tp(i_1 \dots i_k\ j_1 \dots j_k) = tp(p_1 \dots p_k\ q_1 \dots q_k).$$

By the definition of type, this implies that $p_l = q_l$ for $l \neq l_0$ and $A_H(p_{l_0}, q_{l_0}) = 1$. Then $A_{H^k}(p_1 \dots p_k, q_1 \dots q_k) = 1$. The argument can be reversed, proving that

$$A_{G^k}(i_1 \dots i_k, j_1 \dots j_k) = A_{H^k}(p_1 \dots p_k, q_1 \dots q_k).$$

This prove the case $r = 1$. Now assume the statement is valid for r, and suppose that

$$W^{r+1}_{G,2k}(i_1 \dots i_k\ j_1 \dots j_k) = W^{r+1}_{H,2k}(p_1 \dots p_k\ q_1 \dots q_k).$$

By the definition of the WL coloring,

$$\sum_{m \in V_G} \bigl(tp(i_1 \dots i_k\ j_1 \dots j_k\ m), S^r_{G,2k}(i_1 \dots i_k\ j_1 \dots j_k\ m)\bigr)$$
$$= \sum_{m \in V_H} \bigl(tp(p_1 \dots p_k\ q_1 \dots q_k\ m), S^r_{H,2k}(p_1 \dots p_k\ q_1 \dots q_k\ m)\bigr).$$

Therefore there exists a permutation σ of $\{1, 2, \dots, n\}$ such that

$$\begin{cases} tp(i_1 \dots i_k\ j_1 \dots j_k\ m) = tp\bigl(p_1 \dots p_k\ q_1 \dots q_k\ \sigma(m)\bigr), \\ W^r_{G,2k}(i_1 \dots i_k\ j_1 \dots j_{k-1}\ m) = W^r_{H,2k}\bigl(p_1 \dots p_k\ q_1 \dots q_{k-1}, \sigma(m)\bigr), \\ \dots \\ W^r_{G,2k}(m\ i_2 \dots i_k\ j_1 \dots j_k) = W^r_{H,2k}\bigl(\sigma(m)\ p_2 \dots p_k\ q_1 \dots q_k\bigr). \end{cases}$$

The induction hypothesis implies

$$\begin{cases} A_G(i_t, m) = A_G\bigl(p_t, \sigma(m)\bigr) \quad \text{for } t = 1, \dots, k, \\ A^r_{G^k}(i_1 \dots m \dots i_k, j_1 \dots j_k) = A^r_{H^k}\bigl(p_1 \dots \sigma(m) \dots p_k, q_1 \dots q_k\bigr). \end{cases}$$

Our goal is to show that

$$A^{r+1}_{G^k}(i_1 \dots i_k, j_1 \dots j_k) = A^{r+1}_{H^k}(p_1 \dots p_k, q_1 \dots q_k).$$

We have

$$A^{r+1}_{G^k}(i_1 \dots i_k, j_1 \dots j_k) = \sum_{s_1 \dots s_k} A_{G^k}(i_1 \dots i_k, s_1 \dots s_k) A^r_{G^k}(s_1 \dots s_k, j_1 \dots j_k). \tag{4}$$

Observe that $A_{G^k}(i_1 \dots i_k, s_1 \dots s_k) = 0$ unless there exists an index t such that $A_G(i_t, s_t) = 1$ and $i_l = s_l$ for all $l \neq t$. Hence

$$\begin{aligned} A^{r+1}_{G^k}(i_1 \dots i_k, j_1 \dots j_k) &= \sum_{m \in V_G} \sum_{t=1}^{k} A_G(i_t, m) A^r_{G^k}(i_1 \dots m \dots i_k, j_1 \dots j_k) \\ &= \sum_{m \in V_H} \sum_{t=1}^{k} A_H\bigl(p_t, \sigma(m)\bigr) A^r_{H^k}\bigl(p_1 \dots \sigma(m) \dots p_k, q_1 \dots q_k\bigr) \\ &= A^{r+1}_{H^k}(p_1 \dots p_k, q_1 \dots q_k). \qquad \square \end{aligned}$$

Theorem 6. *Let G and H be two graphs. If $I_{G,2k}(t) = I_{H,2k}(t)$, then*

$$F\bigl(G^k, t\bigr) = F\bigl(H^k, t\bigr).$$

In other words, if the $2k$-th WL refinement cannot distinguish G from H, then their k-th powers are cospectral.

Proof. Assume $I_{G,2k}(t) = I_{H,2k}(t)$. By Proposition 4, there is a bijection σ from the set of $2k$-tuples of G to the set of $2k$-tuples of H, such that for every $2k$-tuple $i_1 \dots i_k\ j_1 \dots j_k$ of G,

$$W^r_{G,2k}(i_1 \dots i_k\ j_1 \dots j_k) = W^r_{H,2k}\bigl(\sigma(i_1 \dots i_k\ j_1 \dots j_k)\bigr)$$

for $r \geqslant 1$. When $r = 1$, this is

$$tp(i_1 \dots i_k\ j_1 \dots j_k) = tp\big(\sigma(i_1 \dots i_k\ j_1 \dots j_k)\big).$$

In particular, σ sends the diagonal of $W^r_{G,2k}$ to the diagonal of $W^r_{H,2k}$, that is,

$$\sigma(i_1 \dots i_k\ i_1 \dots i_k) = p_1 \dots p_k\ p_1 \dots p_k$$

for some k-tuple $p_1 \dots p_k$. Then, collecting all the colors in the diagonal, we have

$$\sum_{i_1 \dots i_k} W^r_{G,2k}(i_1 \dots i_k\ i_1 \dots i_k) = \sum_{i_1 \dots i_k} W^r_{H,2k}\big(\sigma(i_1 \dots i_k)\sigma(i_1 \dots i_k)\big).$$

By Theorem 5, this implies

$$\sum_{i_1 \dots i_k} A^r_{G^k}(i_1 \dots i_k, i_1 \dots i_k) = \sum_{i_1 \dots i_k} A^r_{H^k}\big(\sigma(i_1 \dots i_k), \sigma(i_1 \dots i_k)\big)$$

that is, $Tr\, A^r_{G^k} = Tr\, A^r_{H^k}$ for $r \geqslant 1$. Then, $F(G^k, t) = F(H^k, t)$. $\square$

Our goal is to prove the analogue of Theorem 6 for the k-th symmetric powers. As an intermediate step, we prove analogues of Theorems 5 and 6 for the restricted k-th powers.

Theorem 7. *Let $G^{(k)}$ and $H^{(k)}$ be the k-th restricted powers of two graphs G and H. Let $A^r_{G^{(k)}}$ and $A^r_{H^{(k)}}$ be the r-th powers of their adjacency matrices. Assume that $i_1 \dots i_k$ and $j_1 \dots j_k$ are k-tuples in $\mathcal{D}_{G,k}$, and that $p_1 \dots p_k$ and $q_1 \dots q_k$ are k-tuples in $\mathcal{D}_{H,k}$. If*

$$W^r_{G,2k}(i_1 \dots i_k\ j_1 \dots j_k) = W^r_{H,2k}(p_1 \dots p_k\ q_1 \dots q_k),$$

then

$$A^r_{G^{(k)}}(i_1 \dots i_k, j_1 \dots j_k) = A^r_{H^{(k)}}(p_1 \dots p_k, q_1 \dots q_k).$$

Proof. The proof mimics that of Theorem 5. The case $r = 1$ is unaltered, so we assume the proposition is valid for r and we suppose that

$$W^{r+1}_{G,2k}(i_1 \dots i_k\ j_1 \dots j_k) = W^{r+1}_{H,2k}(p_1 \dots p_k\ q_1 \dots q_k).$$

This means that there is a permutation σ of $\{1, 2, \dots, n\}$ such that

$$\begin{cases} tp(i_1 \dots i_k\ j_1 \dots j_k\ m) = tp\big(p_1 \dots p_k\ q_1 \dots q_k\ \sigma(m)\big), \\ W^r_{G,2k}(i_1 \dots i_k\ j_1 \dots j_{k-1}\ m) = W^r_{H,2k}\big(p_1 \dots p_k\ q_1 \dots q_{k-1}, \sigma(m)\big), \\ \dots \\ W^r_{G,2k}(m\ i_2 \dots i_k\ j_1 \dots j_k) = W^r_{H,2k}\big(\sigma(m)\ p_2 \dots p_k\ q_1 \dots q_k\big). \end{cases}$$

From the first of these equations, we observe that $m = i_t$ implies $\sigma(m) = p_t$. Therefore, the k-tuple $(i_1 \dots i_{l-1}\ m\ i_{l+1} \dots i_k)$ is in $\mathcal{D}_k$ if and only if

$$\big(p_1 \dots p_{l-1}\ \sigma(m)\ p_{l+1} \dots p_k\big)$$

is in $\mathcal{D}_k$.

This observation shows that, if we assume $m \neq i_t$ for $t = 1, \dots, k$, we are allowed to apply the induction hypothesis to obtain

$$\begin{cases} A_G(i_t, m) = A_H\big(p_t, \sigma(m)\big) \quad \text{for } t = 1, \dots, k, \\ A^r_{G^k}(i_1 \dots m \dots i_k, j_1 \dots j_k) = A^r_{H^k}\big(p_1 \dots \sigma(m) \dots p_k, q_1 \dots q_k\big). \end{cases}$$

Then

$$
\begin{aligned}
A^{r+1}_{G^{(k)}}(i_1 \dots i_k, j_1 \dots j_k) &= \sum_{(s_1 \dots s_k) \in \mathcal{D}_k} A_{G^{(k)}}(i_1 \dots i_k, s_1 \dots s_k) A^r_{G^{(k)}}(s_1 \dots s_k, j_1 \dots j_k) \\
&= \sum_{m \notin \{i_1, \dots, i_k\}} \sum_{t=1}^{k} A_G(i_t, m) A^r_{G^{(k)}}(i_1 \dots m \dots i_k, j_1 \dots j_k) \\
&= \sum_{\sigma(m) \notin \{p_1, \dots, p_k\}} \sum_{t=1}^{k} A_H\big(p_t, \sigma(m)\big) A^r_{H^{(k)}}\big(p_1 \dots \sigma(m) \dots p_k, q_1 \dots q_k\big) \\
&= A^{r+1}_{G^{(k)}}(p_1 \dots p_k, q_1 \dots q_k). \qquad \square
\end{aligned} \tag{5}
$$

Theorem 8. *If the* 2*k-th WL refinement fails to distinguish* G *from* H*, then their restricted k-th powers are cospectral.*

Proof. The proof is analogous to that of Theorem 6. Assume $I_{G,2k}(t) = I_{H,2k}(t)$. Let σ be the bijection given by Proposition 4. Since σ preserves the type of the $2k$-tuples, if $i_1 \dots i_k$ is in $\mathcal{D}_{G,k}$, then

$$\sigma(i_1 \dots i_k\, i_1 \dots i_k) = p_1 \dots p_k\, p_1 \dots p_k$$

for some k-tuple $p_1 \dots p_k \in \mathcal{D}_{H,k}$. Then,

$$\sum_{(i_1 \dots i_k) \in \mathcal{D}_{G,k}} W^r_{G,2k}(i_1 \dots i_k\, i_1 \dots i_k) = \sum_{(i_1 \dots i_k) \in \mathcal{D}_{G,k}} W^r_{H,2k}\big(\sigma(i_1 \dots i_k)\sigma(i_1 \dots i_k)\big).$$

By Theorem 7, this implies

$$\sum_{(i_1 \dots i_k) \in \mathcal{D}_{G,k}} A^r_{G^{(k)}}(i_1 \dots i_k, i_1 \dots i_k) = \sum_{(i_1 \dots i_k) \in \mathcal{D}_{G,k}} A^r_{H^{(k)}}\big(\sigma(i_1 \dots i_k), \sigma(i_1 \dots i_k)\big)$$

that is, $Tr\, A^r_{G^{(k)}} = Tr\, A^r_{H^{(k)}}$ for $r \geqslant 1$. Then, $F(G^{(k)}, t) = F(H^{(k)}, t)$. $\square$

7. Proof of Theorem 1

We can restate Theorem 1 as follows:

Theorem 9. *Let* G *and* H *be two graphs. If* $I_{G,2k}(t) = I_{H,2k}(t)$*, then*

$$F\big(G^{\{k\}}, t\big) = F\big(H^{\{k\}}, t\big).$$

Proof. Assume $I_{G,2k}(t) = I_{H,2k}(t)$. Again, by Proposition 4, there is a bijection σ from $\mathcal{U}_{G,k}$ to $\mathcal{U}_{H,k}$ such that

$$W^r_{G,2k}(i_1 \dots i_{2k}) = W^r_{H,2k}\big(\sigma(i_1 \dots i_{2k})\big) \tag{6}$$

for all $r \geqslant 1$. Since

$$tp(i_1 \dots i_{2k}) = tp\big(\sigma(i_1 \dots i_{2k})\big),$$

we can restrict σ in the following way. If θ is a permutation of $\{1, \dots, k\}$, we denote by $\theta(i_1 \dots i_k)$ the k-tuple $(i_{\theta(1)} \dots i_{\theta(k)})$. Let us write the $2k$-tuples as pairs of k-tuples: $(i_1 \dots i_k, j_1 \dots j_k)$. Observe that if a $2k$-tuple is of the form

$$\big(i_1 \dots i_k, \theta(i_1 \dots i_k)\big)$$

where $(i_1 \dots i_k) \in \mathcal{D}_{G,k}$ and $\theta \in S_k$, then – due to the conservation of the type – σ sends it to a $2k$-tuple of the form $(j_1 \dots j_k, \theta(j_1 \dots j_k))$, for some $(j_1 \dots j_k) \in \mathcal{D}_{G,k}$. Thus, there is a bijection ω from $\mathcal{D}_{G,k}$ to $\mathcal{D}_{H,k}$ such that for every $(i_1 \dots i_k) \in \mathcal{D}_{G,k}$

$$W^r_{G,2k}\big(i_1 \dots i_k, \theta(i_1 \dots i_k)\big) = W^r_{H,2k}\big(\omega(i_1 \dots i_k), \theta\big(\omega(i_1 \dots i_k)\big)\big). \tag{7}$$

By Theorem 7, it follows that

$$A^r_{G^{(k)}}\big(i_1 \dots i_k, \theta(i_1 \dots i_k)\big) = A^r_{H^{(k)}}\big(\omega(i_1 \dots i_k), \theta\big(\omega(i_1 \dots i_k)\big)\big). \tag{8}$$

In particular,

$$\sum_{(i_1 \dots i_k) \in \mathcal{D}_{G,k}} \sum_{\theta \in S_k} A^r_{G^{(k)}}\big(i_1 \dots i_k, \theta(i_1 \dots i_k)\big)$$
$$= \sum_{(i_1 \dots i_k) \in \mathcal{D}_{G,k}} \sum_{\theta \in S_k} A^r_{H^{(k)}}\big(\omega(i_1 \dots i_k), \theta\big(\omega(i_1 \dots i_k)\big)\big).$$

Since ω is a bijection, we can drop it from this last equation, and we have

$$\sum_{(i_1 \dots i_k) \in \mathcal{D}_{G,k}} \sum_{\theta \in S_k} A^r_{G^{(k)}}\big(i_1 \dots i_k, \theta(i_1 \dots i_k)\big) = \sum_{(i_1 \dots i_k) \in \mathcal{D}_{G,k}} \sum_{\theta \in S_k} A^r_{H^{(k)}}\big(i_1 \dots i_k, \theta(i_1 \dots i_k)\big).$$

Let M^k be the matrix of Proposition 3. This last equation can be written as

$$Tr\big(A^r_{G^{(k)}} M_k\big) = Tr\big(A^r_{H^{(k)}} M_k\big).$$

By Proposition 3, this is equivalent to

$$Tr\big(A^r_{G^{\{k\}}}\big) = Tr\big(A^r_{H^{\{k\}}}\big). \tag{9}$$

Since this is true for all r, then $F(G^{\{k\}}, t) = F(H^{\{k\}}, t)$. $\square$

References

[1] K. Audenaert, C. Godsil, G. Royle, T. Rudolph, Symmetric squares of graphs, J. Combin. Theory Ser. B 97 (1) (January, 2007) 74–90.

[2] L. Babai, P. Erdös, S. Selkow, Random graph isomorphism, SIAM J. Comput. 9 (1980) 628–635.

[3] J.-Y. Cai, M. Fürer, N. Immerman, An optimal lower bound on the number of variables for graph identification, Combinatorica 12 (4) (December, 1992) 389–410.

[4] M.M. Fürer, Weisfeiler–Lehman refinement requires at least a linear number of iterations, in: Automata, Languages and Programming, in: Lecture Notes in Comput. Sci., vol. 2076, Springer, Berlin/Heidelberg, January 2001, pp. 322–333.

[5] M. Grohe, Equivalence in finite-variable logics is complete for polynomial time, in: Proceedings of the 37th Annual Symposium on Foundations of Computer Science, FOCS, 1996, p. 264.

[6] M. Grohe, Isomorphism testing for embeddable graphs through definability, in: Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, STOC, 2000, pp. 63–72.

[7] A.J. Schwenk, Almost all trees are cospectral, in: New Directions in the Theory of Graphs, Academic Press, New York, 1973, pp. 275–307.