Egerváry Research Group
on Combinatorial Optimization

Technical reportS

TR-2019-17. Published by the Egerváry Research Group, Pázmány P. sétány 1/C, H–1117, Budapest, Hungary. Web site: www.cs.elte.hu/egres . ISSN 1587–4451.

A novel approach to graph isomorphism

Alpár Jüttner and Péter Madarasi

December 29, 2019

[1]The project has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00002).

A novel approach to graph isomorphism

Alpár Jüttner⋆ and Péter Madarasi⋆

Abstract

This paper presents the concept of *walk-labeling* that can be used to design polynomial algorithm for solving the graph isomorphism problem for various graph classes. For example, all non-cospectral graph pairs can be distinguished by the proposed combinatorial method. Furthermore, even non-isomorphic cospectral graphs might be distinguished assuming certain properties of their eigenspaces.

The concept of k*-strong walk-labeling* is a refinement of the aforementioned labeling, which has both theoretical and practical applications. Its applications include the generation of graph fingerprints, which uniquely identify all the graphs in the considered databases – including all strongly regular graphs on at most 64 nodes and all graphs on at most 12 nodes. They provably identify all trees and 3-connected planar graphs up to isomorphism, which – as a byproduct – gives a new isomorphism algorithm for both graph classes. The practical importance of this fingerprint lies in significantly speeding up searching in graph databases and graph matching algorithms, which are commonly required in biological and chemical applications.

Keywords: graph isomorphism, graph fingerprint, graph hash, searching in graph databases, strongly regular graphs, isomorphism invariant, planar graph

1 Introduction

The graph isomorphism problem is one of the few natural problems in NP that are neither known to be in P nor NP-Complete. At the same time, polynomial-time isomorphism algorithms have been developed for various graph classes, like trees and planar graphs [1], bounded valence graphs [2], interval graphs [3] or permutation graphs [4]. Furthermore, an FPT algorithm has recently been presented for the colored hypergraph isomorphism problem [5]. The most efficient practical graph isomorphism algorithms include Nauty [6], VF2 [7] and its variants [8].

Many applications require more than just verifying if two given graphs are isomorphic — in most cases an isomorphic copy of a given graph G is to be found in a large graph database. Instead of solving the graph isomorphism problem between G and each graph in the database, one might generate so-called fingerprints for all graphs s.t. if two fingerprints are different, then the corresponding graphs can not be isomorphic.

After this preprocessing step, one can omit each graph having different fingerprint from that of G.

Graph fingerprints are widely used, and multiple schemes have been proposed to generate them. For example, graph fingerprints were generated by considering the node labels of short paths in [9]. Another graph isomorphism invariant, the spectrum has been theoretically studied in [10], [11] and combined with heat-kernels in [12]. The number of graphs determined (i.e. distinguished from the non-isomorphic graphs) by their spectrum was numerically examined up to 12 nodes in [13], and around 80% of the graphs were found to be determined by their spectrum.

Recently, various algorithms have been developed based on discrete time quantum walks (DTQW) or continuous time quantum walks (CTQW), aiming at distinguishing non-isomorphic graph pairs. It is well known that neither standard single-particle DTQW nor CTQW can distinguish Strongly regular graphs (SRG) of the same parameters, furthermore a constant-particle CTQW without interaction can distinguish no SRG pairs of the same parameters, see [14] and [15]. However, the distinguishing power of a variant of single-particle DTQW presented in [14] turned out to be larger than that of a standard DTQW. Namely, it generates different signatures for certain non-isomorphic SRG pairs of the same parameters, but there are still SRG pairs that it fails to distinguish. In [16], CTQW were shown to be less powerful than DTQW as far as the graph isomorphism is concerned. On the other hand, a state-of-the-art quantum walk method using interacting bosons turned out to distinguish all SRG's on at most 64 nodes [17]. This compares to the easy-to-compute fingerprint introduced in Section 3, which distinguishes all the mentioned SRG's, in addition, it provides a compact description of the graphs.

This work presents the concept of walk-labeling, which can be used to solve the graph isomorphism problem in polynomial time under certain conditions — which hold for a wide range of the graph pairs. All non-cospectral graph pairs are proved to be distinguished by the proposed combinatorial method (without computing the graph spectra). Furthermore, even if the graphs are cospectral and non-isomorphic, various conditions are shown that ensure that the graphs are distinguished.

A refinement of the aforementioned labeling called k-strong walk-labeling is also introduced. Its applications include speeding up any backtracking-based graph matching algorithm, and a fingerprint generation method, which uniquely identifies all the graphs in the considered graph databases — including all known strongly regular graphs. Therefore, it is competitive with the state-of-the-art quantum walk algorithms. In addition, it compresses all information about the graph to a short fingerprint. The fingerprint is a promising Co-NP characterisation candidate for the graph isomorphism problem, since strongly regular graphs — which it manages to uniquely identify on up to 64 nodes — are known as possibly the hardest instances of the graph isomorphism problem.

The rest of the paper is structured as follows. Section 1.1 introduces the most important notations. Section 2 defines the so called *walk-labeling*, and presents some spectral-based result. A refinement of walk-labeling is introduced in Section 3, which is proved to identify trees and 3-connected planar graphs up to isomorphism.

1.1 Notation

As usual, sets are described in curly brackets, and multisets are described in curly brackets followed by a superscript hash character. For example, $\{1, 2, 3\}$ denotes the set consisting of the numbers 1,2,3, and $\{1, 1, 2, 3\}^{\#}$ denotes the multiset consisting of numbers 1,1,2,3. Let $\mathbb{N}$ denote the non-negative integer numbers. For a positive integer n, let $[n]$ denote the set $\{i \in \mathbb{N} : 1 \leq i \leq n\}$.

Throughout the thesis $G = (V, E)$, $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ denote three arbitrary loop-free undirected graphs with $n > 1$ nodes, where V, V_1, V_2 denotes the node sets and E, E_1, E_2 the edge sets, respectively. For the sake of simplicity, the node sets are assumed to be $[n]$, that is $V = V_1 = V_2 = [n]$. The adjacency matrices of these graphs are $A, A_1, A_2 \in \{0, 1\}^{n\times n}$, respectively. Let $\Gamma_G(i)$ denote the set of the neighbors of node i in graph G.

Matrices A_1 and A_2 denote the adjacency matrices of G_1 and G_2, respectively. Let $\lambda_1 \geq \lambda_2 \geq .. \geq \lambda_n$ and $\mu_1 \geq \mu_2 \geq .. \geq \mu_n$ denote the corresponding eigenvalues. G_1 and G_2 are **cospectral** if the multisets of the eigenvalues of A_1 and A_2 are equal. Let $U, V \in \mathbb{R}^{n\times n}$ orthogonal matrices (i.e. $U^T U = I$ and $V^T V = I$) s.t. $A_1 U = U \operatorname{diag}(\lambda_1, \lambda_2, .., \lambda_n)$ and $A_2 V = V \operatorname{diag}(\mu_1, \mu_2, .., \mu_n)$. U and V are called the **eigenmatrices** of G_1 and G_2, respectively. Let $u_1, u_2, ..u_n$ and $v_1, v_2, ..v_n$ denote the column vectors of U and V, respectively. Note that V denotes both the eigenmatrix of G_2 and the node set of G, but this will not cause ambiguity. Please note that u_{ij} denotes the j^{th} entry of eigenvector u_i, i.e. it is the entry of U in the j^{th} row and i^{th} column, where $i, j \in [n]$. For a matrix Q, let $Q|_k$ denote the first k columns of Q. Finally, let δ_{ij} denote the Kronecker delta.

2 Counting walks

Let $\ell_G : V_G \longrightarrow \mathbb{N}^{n\times\mathbb{N}}$ be s.t. $\ell_G(i)_{jl}$ denotes the number of walks of length l between node i and node j for $l \geq 0$. In other words, column l of matrix $\ell_G(i)$ is $A^{l-1}e_i$, where e_i is the incidence vector of node $i \in V_G$ and $l \geq 1$. The function ℓ_G will be referred to as **(infinite) walk-labeling**.

Two matrices Q_1 and Q_2^w are said to be **permutation-equal** if there exists a permutation matrix P for which $PQ_1 = Q_2$. This equivalence relation is denoted by $Q_1 \overset{\text{p}}{=} Q_2$.

Claim 2.1. *If $\ell(u) \overset{\text{p}}{\neq} \ell(v)$ for two nodes $u \in V_1$ and $v \in V_2$, then there is no isomorphism between G_1 and G_2 that maps node u to node v.*

The definition of walk-isomorphism follows, which plays an important role in Section 2.2.

Definition 2.2. G_1 and G_2 are **walk-isomorphic** if the nodes can be relabeled s.t. $\ell_{G_1}(i) \overset{\text{p}}{=} \ell_{G_2}(i)$ for each node i.

Claim 2.3. *If two graphs are isomorphic, then they are walk-isomorphic.*

Later on, it will be shown in important special cases, that the reverse direction holds as well.

2.1 Only short walks matter

The matrices that ℓ assigns to the nodes are infinite long, therefore there is no straightforward way of checking whether two such matrices are permutation-equal or not. In what follows, it turns out that it is sufficient to consider the first $n+1$ columns of the label matrices.

Definition 2.4. For given column vectors $q_0, q_1, ..$ over a field, let $\text{span}(q_0, q_1, ..)$ denote the linear subspace spanned by the column vectors $q_0, q_1, ...$

The following lemma will be useful in the proof of Theorem 2.6.

Lemma 2.5. *For an arbitrary real square matrix $M \in \mathbb{R}^{n\times n}$ and $q_0 \in \mathbb{R}^n$ column vector,* $\text{span}(q_0, q_1, q_2, ..) = \text{span}(q_0, q_1, .., q_{n-1})$*, where $q_i := M^i q_0$ for all $i \geq 0$.*

Proof. By induction, one may show that if $\text{span}(q_0, q_1, ..q_i) = \text{span}(q_0, q_1, ..q_{i+1})$, then $\text{span}(q_0, q_1, ..q_i) = \text{span}(q_0, q_1, q_2, ..)$ for all i. Therefore, columns $q_0, q_1, \dots, q_n$ generates $\text{span}(q_0, q_1, q_2, ..)$. □

The following theorem shows that it is sufficient to consider the first few columns of the labels, i.e. only the number of short walks matters. Recall that $\ell_G|_k(i)$ denotes the first k columns of matrix $\ell_G(i)$.

Theorem 2.6. *For every graph pair G_1, G_2 with n nodes and for all $i_1 \in V_1, i_2 \in V_2$*

$$\ell_{G_1}(i_1) \overset{\text{p}}{=} \ell_{G_2}(i_2) \iff \ell_{G_1}|_{n+1}(i_1) \overset{\text{p}}{=} \ell_{G_2}|_{n+1}(i_2).$$

Proof. Let Q_1, Q_2, Q_1' and Q_2' denote the matrices $\ell_{G_1}(v_1), \ell_{G_2}(v_2), \ell_{G_1}|_{n+1}(v_1)$ and $\ell_{G_2}|_{n+1}(v_2)$, respectively. If $Q_1 \overset{\text{p}}{=} Q_2$, then, by definition, there exists a permutation matrix P for which $PQ_1 = Q_2$. Clearly, $PQ_1 = Q_2 \Rightarrow PQ_1' = Q_2'$. To show the other direction, suppose that $Q_1' \overset{\text{p}}{=} Q_2'$, and the columns of Q_1 and Q_2 are $q_0, q_1, q_2..$ and $q_0', q_1', q_2'..$, respectively. Let A_1, A_2 denote the adjacency matrices of G_1 and G_2, respectively. Since $Q_1' \overset{\text{p}}{=} Q_2'$, there exists a permutation matrix P s.t. $PQ_1' = Q_2'$, thus it is sufficient to prove that $Pq_i = q_i'$ hols for all $i \geq n+1$.
By induction, suppose that $k < i \Longrightarrow Pq_k = q_k'$ for all k. The existence of coefficients $\alpha_0, ..\alpha_{n-1}$ s.t. $q_{i-1} = \sum_{j=0}^{n-1} \alpha_j q_j$ and $q_{i-1}' = \sum_{j=0}^{n-1} \alpha_j q_j'$ is an immediate consequence of Lemma 2.5. Therefore,

$$Pq_i = PA_1 q_{i-1} = P\sum_{j=0}^{n-1} \alpha_j A_1 q_j = \sum_{j=0}^{n-1} \alpha_j P q_{j+1} = \sum_{j=0}^{n-1} \alpha_j q_{j+1}' = \sum_{j=0}^{n-1} \alpha_j A_2 q_j' = A_2 q_{i-1}' = q_i' \tag{1}$$

holds for all $i \geq n+1$, which had to be shown. □

The following example shows that the previous theorem is tight in the sense that it is not always sufficient to consider the first n columns of the walk labels.

Example 2.7. Let P_n denote the path of n nodes, and let P'_n denote the path of n nodes with a loop on one of its endpoints. To distinguish two loop-free endpoints of the two graphs, indeed $n+1$ columns are necessary, since their labels do not turn out to be different earlier.

From now on, ℓ_G might refer to $\ell_G|_{n+1}$ or the infinite walk-labeling. Note that the walk label $\ell_G|_{n+1}(i)$ of a given node i can be computed in $O(nm)$ operations using a simple dynamic programming method. Furthermore, one might prove that the occurring numbers consist of polynomial many bits in the size of the graph. Therefore it takes $O(n^2m + n^3log(n))$ steps to decide whether two graphs are walk-isomorphic by sorting the labels of both graphs.

2.2 Spectral results

Simple observations follow for later reference.

Claim 2.8. *If $\ell_{G_1}(i) \overset{\text{p}}{=} \ell_{G_2}(i')$, then the number of closed walks of length l starting from $i \in V_1$ and $i' \in V_2$ are the same for all $l \geq 0$.*

Proof. By definition, there exists a permutation matrix P s.t. $P\ell_{G_1}(i) = \ell_{G_2}(i')$. Notice that the first column of $\ell_{G_1}(i)$ and $\ell_{G_2}(i')$ enforces that P maps the i^{th} row of $\ell_{G_1}(i)$ to the i'^{th} row of $\ell_{G_2}(i')$, which means that the number of closed walks from $i \in V_1$ and $i' \in V_2$ are the same for all $l \geq 0$. □

Lemma 2.9. *For all $i, j \in [n]$ and for all $l \geq 1$, $(A^l)_{ij} = \sum_{k=1}^{n} u_{ki}u_{kj}\lambda_k^l$ holds, where $\lambda_1, \lambda_2, ..\lambda_n$ are the eigenvalues of G. The right-hand side of this equation will be referred to as the **eigen decomposition**.*

Proof. $U \in \mathbb{R}^{n\times n}$ is an orthonormal matrix s.t. $AU = U\,\mathrm{diag}(\lambda_1, \lambda_2, ..\lambda_n)$. Clearly, $U^{-1}A^lU = \mathrm{diag}(\lambda_1^l, \lambda_2^l, ..\lambda_n^l)$ holds, hence $A^l = U\,\mathrm{diag}(\lambda_1^l, \lambda_2^l, ..\lambda_n^l)U^{-1}$. Therefore, $(A^l)_{ij} = \sum_{k=1}^{n} u_{ki}u_{kj}\lambda_k^l$ for any node pair $i, j \in [n]$. □

The following observation is an immediate consequence of Lemma 2.9.

Corollary 2.10. *For all $i, j \in [n]$ and $l \geq 1$, there exist $\beta_1^{ij}, \beta_2^{ij}, ..\beta_p^{ij} \in \mathbb{R}$ s.t. $(A^l)_{ij} = \sum_{m=1}^{p} \beta_m^{ij}\tilde{\lambda}_m^l$, where $\tilde{\lambda}_1, \tilde{\lambda}_2, ..\tilde{\lambda}_p$ are the distinct non-zero eigenvalues of G. The the right-hand side of this equation will be referred to as the **aggregated eigen decomposition**.*

Proof. By Lemma 2.9, $(A^l)_{ij} = \sum_{k=1}^{n} u_{ki}u_{kj}\lambda_k^l$ for $l \geq 1$ and $i, j \in [n]$. Clearly, $\beta_m^{ij} := \sum_{k:\lambda_k=\tilde{\lambda}_m} u_{ki}u_{kj}$ is a proper choice, where $i, j \in [n]$ and $m \in [p]$. □

The following theorem shows that non-cospectral graphs are not walk-isomorphic.

Theorem 2.11. *If G_1 and G_2 are walk-isomorphic, then the spectra of G_1 and G_2 are the same.*

Proof. The proof consists of two steps.

Step 1: We prove that the set of non-zero eigenvalues of G_1 and G_2 are the same.

Lemma 2.12. *Coefficient β_k^{ii} in the aggregated eigen decomposition is zero if it corresponds to a non-zero eigenvalue of exactly one of G_1 and G_2 for all $i, k \in [n]$.*

Proof. Let $\tilde{\lambda}_1, \tilde{\lambda}_2, .., \tilde{\lambda}_r, \tilde{\theta}_{r+1}, ..\tilde{\theta}_p$ and $\tilde{\lambda}_1, \tilde{\lambda}_2, .., \tilde{\lambda}_r, \tilde{\mu}_{r+1}, ..\tilde{\mu}_q$ denote all the distinct non-zero eigenvalues of G_1 and G_2, respectively, where $\tilde{\lambda}_1, \tilde{\lambda}_2, .., \tilde{\lambda}_r$ are the mutual non-zero eigenvalues of the two graphs and $\tilde{\theta}_{r+1}, ..\tilde{\theta}_p, \tilde{\mu}_{r+1}, ..\tilde{\mu}_q$ are pairwise distinct.

For the sake of simplicity, suppose that the nodes are reindexed s.t. the identity mapping is a walk-isomorphism, i.e. $\ell_{G_1}(i) \overset{\mathrm{p}}{=} \ell_{G_2}(i)$ for all node i.

By Corollary 2.10, there exist coefficients $\alpha_1, \alpha_2, .., \alpha_p, \beta_1, \beta_2, ..\beta_q$ for any i, j s.t.

$$(A_1^l)_{ij} = \sum_{k=1}^{r} \alpha_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{p} \alpha_k \tilde{\theta}_k^l \tag{2}$$

and

$$(A_2^l)_{ij} = \sum_{k=1}^{r} \beta_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{q} \beta_k \tilde{\mu}_k^l \tag{3}$$

for all $l \geq 1$.

The two graphs being walk-isomorphic, one gets that

$$\sum_{k=1}^{r} \alpha_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{p} \alpha_k \tilde{\theta}_k^l = (A_1^l)_{ii} = (A_2^l)_{ii} = \sum_{k=1}^{r} \beta_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{q} \beta_k \tilde{\mu}_k^l \tag{4}$$

holds for all $i \in [n]$ and $l \geq 1$, where the second equation follows from Claim 2.8.

Subtracting the right-hand side, one gains the the following equations from (4)

$$\sum_{k=1}^{r} (\alpha_k - \beta_k) \tilde{\lambda}_k^l + \sum_{k=r+1}^{p} \alpha_k \tilde{\theta}_k^l - \sum_{k=r+1}^{q} \beta_k \tilde{\mu}_k^l = 0 \tag{5}$$

for all $l \geq 1$. Let $m := p + q - r$, and consider the following linear equations for $l \in [m]$.

$$\sum_{k=1}^{r} x_k \tilde{\lambda}_k^l + \sum_{k=r+1}^{p} x_k \tilde{\theta}_k^l + \sum_{k=r+1}^{q} x_{p+k-r} \tilde{\mu}_k^l = 0, \tag{6}$$

where

$$x_s := \begin{cases} \alpha_s - \beta_s, & \text{if } 1 \leq s \leq r \\ \alpha_s, & \text{if } r+1 \leq s \leq p \\ -\beta_{r+s-p}, & \text{if } p+1 \leq s \leq p+q-r, \end{cases} \tag{7}$$

for all $s \in [m]$. The matrix of this linear equation system is

$$M := \begin{bmatrix} \tilde\lambda_1^1 & \dots & \tilde\lambda_r^1 & \tilde\theta_{r+1}^1 & \dots & \tilde\theta_p^1 & \tilde\mu_{r+1}^1 & \dots & \tilde\mu_q^1 \\ \tilde\lambda_1^2 & \dots & \tilde\lambda_r^2 & \tilde\theta_{r+1}^2 & \dots & \tilde\theta_p^2 & \tilde\mu_{r+1}^2 & \dots & \tilde\mu_q^2 \\ \tilde\lambda_1^3 & \dots & \tilde\lambda_r^3 & \tilde\theta_{r+1}^3 & \dots & \tilde\theta_p^3 & \tilde\mu_{r+1}^3 & \dots & \tilde\mu_q^3 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \tilde\lambda_1^m & \dots & \tilde\lambda_r^m & \tilde\theta_{r+1}^m & \dots & \tilde\theta_p^m & \tilde\mu_{r+1}^m & \dots & \tilde\mu_q^m \end{bmatrix}. \tag{8}$$

Observe that $M = M' \operatorname{diag}(\tilde\lambda_1^1, \dots, \tilde\lambda_r^1, \tilde\theta_{r+1}^1, \dots, \tilde\theta_p^1, \tilde\mu_{r+1}^1, \dots, \tilde\mu_q^1)$, where M' denotes the following Vandermonde matrix.

$$M' := \begin{bmatrix} 1 & \dots & 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ \tilde\lambda_1^1 & \dots & \tilde\lambda_r^1 & \tilde\theta_{r+1}^1 & \dots & \tilde\theta_p^1 & \tilde\mu_{r+1}^1 & \dots & \tilde\mu_q^1 \\ \tilde\lambda_1^2 & \dots & \tilde\lambda_r^2 & \tilde\theta_{r+1}^2 & \dots & \tilde\theta_p^2 & \tilde\mu_{r+1}^2 & \dots & \tilde\mu_q^2 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \tilde\lambda_1^m & \dots & \tilde\lambda_r^m & \tilde\theta_{r+1}^m & \dots & \tilde\theta_p^m & \tilde\mu_{r+1}^m & \dots & \tilde\mu_q^m \end{bmatrix} \tag{9}$$

Therefore $\det(M) = \det(M') \prod_{k=1}^{r} \tilde\lambda_k \prod_{k=r+1}^{p} \tilde\theta_k \prod_{k=r+1}^{q} \tilde\mu_k \neq 0$, thus the only solution is $x \equiv 0$, that is

$$\begin{cases} \alpha_s = \beta_s, & \text{if } 1 \le s \le r \\ \alpha_s = 0, & \text{if } r+1 \le s \le p \\ \beta_{r+s-p} = 0, & \text{if } p+1 \le s \le p+q-r \end{cases} \tag{10}$$

follows for all $s \in [m]$. □

Let $\lambda^* \neq 0$ denote an eigenvalue which corresponds to exactly one of the graphs, say to G_1. Next we argue that there exists a node $i \in V_1$ s.t. λ^* has non-zero coefficient in the aggregated eigen decomposition given by Corollary 2.10 for $(A_1^l)_{ii}$ – contradicting Lemma 2.12. Let $\tilde m$ denote the unique index for which $\tilde\lambda_{\tilde m} = \lambda^*$. By Corollary 2.10, the coefficient of $\tilde\lambda_{\tilde m}$ in the case of the number of closed walks from node i is $\beta_{\tilde m}^{ii} := \sum_{k:\lambda_k=\tilde\lambda_{\tilde m}} u_{ki}u_{ki}$. Let m be an index such that $\lambda_m = \tilde\lambda_{\tilde m}$, and let i be s.t. $u_{mi}u_{mi} > 0$ (there exists at least one index like this, since $u_m u_m = 1$). Observe that $\beta_{\tilde m}^{ii} \ge u_{mi}u_{mi} > 0$ holds, therefore node i meets the requirements, contradicting Lemma 2.12.

Step 2: We show that the multiplicities of the eigenvalues are the same in G_1 and G_2. It is sufficient to show that the multiplicities of the non-zero eigenvalues are the same, because this implies that the multiplicities of zero are the same in G_1 and G_2. Let $\tau_i^{(k)}$ denote the multiplicity of $\tilde\lambda_i$ in G_k $(k = 1, 2)$, where $\tilde\lambda_1, .., \tilde\lambda_p$ are the mutual eigenvalues of G_1 and G_2.

As a consequence of Lemma 2.9, the sum of the numbers of closed walks of G_k of length l is $\sum_{j=1}^{p} \tau_j^{(k)} \tilde\lambda_j^l$, $(l \ge 1)$. Since G_1 and G_2 are walk-isomorphic, Claim 2.8 applies, thus the sum of the numbers of closed walks of length l in the two graphs are the same for all l, i.e. $\sum_{j=1}^{p} \tau_j^{(1)} \tilde{\lambda}_j^l = \sum_{j=1}^{p} \tau_j^{(2)} \tilde{\lambda}_j^l$ for all $l \geq 1$. Subtracting the right-hand side provides for all $l \geq 1$ that

$$\sum_{j=1}^{p} (\tau_j^{(1)} - \tau_j^{(2)}) \tilde{\lambda}_j^l = 0 \tag{11}$$

Consider these equations for $l \in [p]$, and let $x_j := \tau_j^{(1)} - \tau_j^{(2)}$ for all $j \in [p]$. Similarly to step 1, the matrix of this equation system has non-zero determinant, thus the only solution is $x \equiv 0$, i.e. $\tau_j^{(1)} = \tau_j^{(2)}$ for all $j \in [p]$. Therefore each non-zero eigenvalue has the same multiplicities in the two graphs, which implies that the multiplicities of eigenvalue 0 is the same, as well. This means that the multisets of the eigenvalues are indeed equal.

□

Theorem 2.13. *Let G_1 and G_2 be cospectral with single eigenvalues. If one of the eigenmatrices has a row which contains non-zero elements only, then the walk-isomorphism is equivalent to the graph isomorphism.*

Proof. Clearly, it suffices to show that if G_1 and G_2 are walk-isomorphic, then they are isomorphic.

It suffices to show a permutation matrix Π s.t. $\Pi A_1 \Pi^T = A_2$. Recall that $U = (u_1, u_2, .., u_n)$ and $V = (v_1, v_2, .., v_n)$ denote the eigenmatrices of G_1 and G_2, respectively, i.e. $A_1 = U diag(\lambda_1, .., \lambda_n) U^T$ and $A_2 = V diag(\lambda_1, .., \lambda_n) V^T$. A permutation matrix Π corresponds to an isomorphism if and only if $\Pi U diag(\lambda_1, .., \lambda_n) U^T \Pi^T = V diag(\lambda_1, .., \lambda_n) V^T$, which holds if and only if $\Pi U = VS$ for some matrix $S = diag(\sigma_1, .., \sigma_n)$, where $\sigma_i \in \{-1, 1\}$. Therefore it is sufficient to show such matrices Π and S.

Without loss of generality, assume that row i^* of U consists of non-zero elements. By the definition of walk-isomorphism, there is a permutation π s.t. $(A_1^l)_{i^*j} = (A_2^l)_{\pi(i^*)\pi(j)}$, thus $u_{ki^*} u_{kj} = v_{k\pi(i^*)} v_{k\pi(j)}$ for all $j \in [n]$. Clearly, row $\pi(i^*)$ of V consists of non-zero elements. Let $S := \mathrm{diag}(\sigma_1, .., \sigma_n)$, where $\sigma_k := \mathrm{sgn}(u_{ki^*}) \, \mathrm{sgn}(v_{k\pi(i^*)}) \in \{-1, 1\}$, and let $\Pi = \begin{cases} 1, & \text{if } \pi(j) = i \\ 0, & \text{otherwise} \end{cases}$. The following claim completes the proof.

Claim 2.14. $\Pi U = VS$

Proof. The values in position (j, k) of the left and the right side are $u_{k\pi^{-1}(j)}$ and $\sigma_k v_{kj}$, respectively. $\forall j, k \in [n] : u_{k\pi^{-1}(j)} = \sigma_k v_{kj} \iff \forall j, k \in [n] : u_{kj} = \sigma_k v_{k\pi(j)} \iff \forall j, k \in [n] : u_{ki^*} u_{kj} = v_{k\pi(i^*)} v_{k\pi(j)}$, where the last equivalence holds because $u_{ki^*} = \sigma_k v_{k\pi(i^*)}$ and $\sigma_k^2 = 1$, and indeed, π was chosen s.t. $u_{ki^*} u_{kj} = v_{k\pi(i^*)} v_{k\pi(j)}$ for all $j, k \in [n]$. □

□

Lemma 2.15. *If G_1 and G_2 are walk-isomorphic graphs and the nodes of G_2 are reindexed s.t. $\ell_{G_1}(i) \stackrel{\mathrm{P}}{=} \ell_{G_2}(i)$ for all $i \in [n]$, then for any single eigenvalue, the corresponding normalized eigenvectors in the two graphs are element-wise the same up to sign.*

Proof. By definition, $\ell_{G_1}|_{n+1}(i) \overset{\mathrm{P}}{=} \ell_{G_2}|_{n+1}(i)$ implies that

$$\sum_{k=1}^{n} u_{ik}u_{ik}\lambda_k^l = (A_1^l)_{ii} = (A_2^l)_{ii} = \sum_{k=1}^{n} v_{ik}v_{ik}\lambda_k^l, \tag{12}$$

thus $\forall i \in [n] : u_{ik}u_{ik} = v_{ik}v_{ik}$. That is, $|u_{ik}| = |v_{ik}|$ for all $i, k \in [n]$. Notice that this proof works even if 0 is an eigenvalue. □

Theorem 2.16. *Let G_1 and G_2 be cospectral with single eigenvalues. If $\{u_{ik} : k \in [n]\}^{\#} \neq \{-u_{ik} : k \in [n]\}^{\#}$ and $\{v_{ik} : k \in [n]\}^{\#} \neq \{-v_{ik} : k \in [n]\}^{\#}$ for all $i \in [n]$, then the walk-isomorphism is equivalent to the graph isomorphism.*

Proof. If G_1 and G_2 are isomorphic, then they are clearly walk-isomorphic. On the other hand, let $w_1, w_2 \in \mathbb{R}^n$ be arbitrary vectors, s.t. $\{w_{1k} : k \in [n]\}^{\#} \neq \{w_{2k} : k \in [n]\}^{\#}$. Let $w_1 \overset{\mathrm{L}}{\succ} w_2$ mean that after non-increasingly ordering their coordinates, w_1 is lexicographically strictly larger than w_2.

Without loss of generality, it can be assumed that $u_i \overset{\mathrm{L}}{\succ} -u_i$ and $v_i \overset{\mathrm{L}}{\succ} -v_i$ holds for all $i \in [n]$. Assume that there is a walk-isomorphism realized by $\pi : V_1 \longrightarrow V_2$. By Lemma 2.15, $|u_{ki}| = |v_{k\pi(i)}|$ holds for all $k \in [n]$ and $i \in [n]$. By contradiction, suppose that there is an index k^* and i^* s.t. $u_{k^*i^*} \neq v_{k^*\pi(i^*)}$. Clearly, $u_{k^*i^*} = -v_{k^*\pi(i^*)}$. Let π^* denote the bijection of i^*, i.e. $u_{ki^*}u_{kj} = v_{k\pi^*(i^*)}v_{k\pi^*(j)}$ holds for all $j, k \in [n]$, even if 0 is an eigenvalue. π^* can be prescribed to satisfy $\pi^*(i^*) = \pi(i^*)$. Thus one gets that $u_{k^*i^*}u_{k^*j} = v_{k^*\pi^*(i^*)}v_{k^*\pi^*(j)}$ for all $j \in [n]$, which implies $-u_{k^*} = \pi v_{k^*}$. But $u_{k^*} \overset{\mathrm{L}}{\succ} -u_{k^*} = \pi^* v_{k^*} \overset{\mathrm{L}}{\succ} -v_{k^*} = \pi^{*-1}u_{k^*}$. Therefore G_1 and G_2 are indeed isomorphic. □

Definition 2.17. A graph is **friendly** [18] if each of its eigenvalues has multiplicity one and $\mathbb{1}U$ has no zero coordinates, where U is the eigenmatrix of the graph.

Corollary 2.18. *If G_1 and G_2 are friendly, then the walk-isomorphism is equivalent to the graph isomorphism.*

Proof. For all $i \in [n] : \mathbb{1}u_i \neq 0$ and $\mathbb{1}v_i \neq 0$ implies that $\{u_{ik} : k \in [n]\}^{\#} \neq \{-u_{ik} : k \in [n]\}^{\#}$ and $\{v_{ik} : k \in [n]\}^{\#} \neq \{-v_{ik} : k \in [n]\}^{\#}$, thus Theorem 2.16 can be applied. □

Theorem 2.19 (Perron-Frobenius)**.** *Let graph G be connected and have at least two nodes. The largest eigenvalue λ_1 of the adjacency matrix of G is positive, has multiplicity one, and $\lambda_1 \geq |\lambda|$ for every eigenvalue λ. In addition, the eigenvector corresponding to λ_1 can be chosen strictly positive.*

The positive normalized eigenvector corresponding to the largest positive eigenvalue in Theorem 2.19 will be referred to as the **Perron-Frobenius eigenvector** of G. The Perron-Frobenius eigenvector of a graph determines the invariant distribution with respect to infinite random walks. Therefore, the following theorem states that if the invariant distributions of two graphs are different, then they are not walk-isomorphic.

Theorem 2.20. *Let G_1 and G_2 be connected cospectral graphs on at least two nodes. If the Perron-Frobenius eigenvectors of G_1 and G_2 are different, then G_1 and G_2 are not walk-isomorphic.*

Proof. Let λ_1 denote the unique largest eigenvalue, and let u_1, v_1 denote the corresponding Perron-Frobenius eigenvectors in G_1 and G_2, respectively. Theorem 2.19 implies u_1 and v_1 are strictly positive.

Clearly, $\lambda_1 > 0$, since the sum of the eigenvalues is the number of closed walks of length one, thus it is non-negative, therefore $\lambda_1 \leq 0$ would imply that each eigenvalue is zero. It is easy to see that a graph having zero eigenvalues only must be the empty graph, which contradicts the assumption of the theorem.

In any walk-isomorphism, $\sum_{k=1}^{n} u_{ki}u_{ki}\lambda_k^l = \sum_{k=1}^{n} v_{ki}v_{ki}\lambda_k^l$ holds for all $l \geq 1$ after reindexing the nodes. With the multiplicity of λ_1 being one, $u_{1i}u_{1i} = v_{1i}v_{1i}$ holds for all $i \in [n]$. Since both u_1 and v_1 are strictly positive, indeed $u_1 = v_1$. $\square$

3 Structure of walks

This section introduces a refined version of walk-labeling.

Notation 3.1. *Let $\mathfrak{s}_G^k(i_1, \ldots, i_k)$ be an $n \times \mathbb{N}$ matrix whose position (j, l) describes the structure of walks of length l between nodes $\{i_1 \ldots i_k\}$ and j. Formally, let*

$$\mathfrak{s}_G^k(i_1, \ldots, i_k)_{il} := \begin{cases} (\emptyset, \{\sum_{q=1}^{k} q\delta_{ii_q}\}), & \text{if } l = 0 \\ (\mathfrak{s}_G^k(i_1, \ldots, i_k)_{il-1}, \{\mathfrak{s}_G^k(i_1, \ldots, i_k)_{i'l-1} : i' \in \Gamma_G(i)\}^{\#}), & \text{otherwise} \end{cases} \tag{13}$$

for all $j \in V$ and for all $l \geq 0$.

Note that the first column of matrix $\mathfrak{s}_G^k(i_1, \ldots, i_k)$ corresponds to walks of length zero, therefore its index is zero. Column l will be denoted by $\mathfrak{s}_G^k(i_1, \ldots, i_k)_{\bullet l}$. Recall that $\mathfrak{s}_G^k(i_1, \ldots, i_k)\big|_q$ denotes the first q columns of matrix $\mathfrak{s}_G^k(i_1, \ldots, i_k)$, that is, it describes the walks upto length $q - 1$.

Simple inductive proof shows that it suffices to consider the first $n + 1$ columns, likewise in the case of walk-labeling.

Claim 3.2. *For every graph pair G_1, G_2 and for any distinct $i_1, \ldots, i_k \in V_1, j_1, \ldots, j_k \in V_2$*

$$\mathfrak{s}_{G_1}^k(i_1, \ldots, i_k) \stackrel{\text{P}}{=} \mathfrak{s}_{G_2}^k(j_1, \ldots, j_k) \iff \mathfrak{s}_{G_1}^k(i_1, \ldots, i_k)\big|_{n+1} \stackrel{\text{P}}{=} \mathfrak{s}_{G_2}^k(j_1, \ldots, j_k)\big|_{n+1},$$

where $n = |V_1| = |V_2|$ and $k \geq 1$ is arbitrary.

Observe that $\mathfrak{s}_{G_1}^k(i_1, \ldots, i_k)\big|_{n+1} \stackrel{\text{P}}{=} \mathfrak{s}_{G_2}^k(j_1, \ldots, j_k)\big|_{n+1}$ if and only if $\mathfrak{s}_{G_1}^k(i_1, \ldots, i_k)_{\bullet n} \stackrel{\text{P}}{=} \mathfrak{s}_{G_2}^k(j_1, \ldots, j_k)_{\bullet n}$. From now on, $\mathfrak{s}_G^k(i_1, \ldots, i_k)$ might refer to $\mathfrak{s}_G^k(i_1, \ldots, i_k)\big|_{n+1}$, as well.

Notation 3.3. *For $q = 1 \dots k$, let $\mathfrak{s}^k_G(i_1, \dots, i_{k-q}) := \{\mathfrak{s}^k_G(i_1, \dots, i_{k-q}, i) : i \in V \setminus \{i_1, \dots, i_{k-q}\}\}^{\#}$.*

The following two claims easily follow by definition.

Claim 3.4. *For all $k \geq 1$, if $\mathfrak{s}^k_{G_1}(i_1) \neq \mathfrak{s}^k_{G_2}(i_2)$ for two nodes $i_1 \in V_1$ and $i_2 \in V_2$, then there is no isomorphism between G_1 and G_2 that maps node i_1 to node i_2.*

Claim 3.5. *For all $k \geq 1$ and any $i_1 \in V_1$ and $i_2 \in V_2$, $\mathfrak{s}^k_{G_1}(i_1) \neq \mathfrak{s}^k_{G_2}(i_2) \implies \mathfrak{s}^{k+1}_{G_1}(i_1) \neq \mathfrak{s}^{k+1}_{G_2}(i_2)$.*

Definition 3.6. G_1 and G_2 are **k-strongly walk-isomorphic** if $\mathfrak{s}^k_{G_1} = \mathfrak{s}^k_{G_2}$.

Remark 3.7. *For any given k, one can verify in polynomial time whether two graphs are k-strongly walk-isomorphic or not.*

Claim 3.8. *If G_1 and G_2 are 1-strongly walk-isomorphic, then they are walk-isomorphic.*

The previous claim implies that if two graphs can be distinguished by walk-isomorphism, then they can be distinguished by k-strong walk isomorphism ($k \geq 1$).

Example 2.7 shows that the previous claim is tight in the sense that considering the first n columns would not be sufficient. Note that the size of $\mathfrak{s}^k_G(i_1, \dots, i_k)_{jl}$ may be exponentially large in n. Practically, one may address this issue by hashing the occurring data using SHA512 – this also enables the generation of graph fingerprints that we found to distinguish all *strongly regular graphs* on at most 64 nodes considering $\mathfrak{s}^k_G$, where $k \geq 2$. The hash function also identifies all graphs on at most 12 nodes for $k \geq 2$. In fact, it remains open whether there exists any non-isomorphic graph pairs that it fails to distinguish (assuming that there are no hash collisions, i.e. a perfect hash function). Note that it is possible to give a perfect hash function for this specific problem by building a dictionary dynamically throughout the labelling process.

The rest of this section investigates the distinguishing power of the above notion on trees and planar graphs.

Theorem 3.9. *The 1-strong walk-isomorphism is equivalent to the graph isomorphism on trees.*

Proof. Let $\mathfrak{s}_G$ denote $\mathfrak{s}^1_G$ in this proof. Given two strongly walk-isomorphic trees $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$, we show that they are isomorphic. For an edge $(r, p) \in E_i$, let $T_i(r, p) = (V_i(r, p), E_i(r, p))$ denote the subtree of G_i obtained as the connected component of $(V, E_i \setminus \{(r, p)\})$ containing node r.

By induction, we prove that for any edges $(r_1, p_1) \in E_1$ and $(r_2, p_2) \in E_2$ if $\mathfrak{s}_{G_1}(r_1)|_{k+1} \stackrel{\mathrm{p}}{=} \mathfrak{s}_{G_2}(r_2)|_{k+1}$ and $\mathfrak{s}_{G_1}(p_1)|_{k+1} \stackrel{\mathrm{p}}{=} \mathfrak{s}_{G_2}(p_2)|_{k+1}$ for $k = |V_1(r_1, p_1)|$, then $T_1(r_1, p_1)$ and $T_2(r_2, p_2)$ are isomorphic. Clearly, if $k = 1$ — i.e. r_1 is a leaf node in G_1 — then r_2 must also be a leaf node in G_2.

Otherwise, one gets that $\{\mathfrak{s}_{G_1}(i)|_k : i \in \Gamma_{G_1}(r_1)\}^{\#} = \{\mathfrak{s}_{G_2}(i)|_k : i \in \Gamma_{G_2}(r_2)\}^{\#}$. Thus r_1 and r_2 have the same number of neighbors and there is a one-to-one mapping $\phi : \Gamma_{G_1}(r_1) \longrightarrow \Gamma_{G_2}(r_2)$ so that v and $\phi(v)$ have same label up to the first k columns for each $v \in \Gamma_{G_1}(r_1)$. Therefore, from the induction hypothesis, $T_1(v, r_1)$ and $T_2(\phi(v), r_2)$

are isomorphic subtrees for all $v \in \Gamma_{G_1}(r_1) \setminus p_1$. The isomorphism of $T_1(r_1, p_1)$ and $T_2(r_2, p_2)$ follows from this immediately.

In order to complete the proof of the theorem, let us choose an arbitrary leaf node $r_1 \in V_1$ and a node $r_2 \in V_2$ with $\mathfrak{s}_{G_1}(r_1) \overset{\text{p}}{=} \mathfrak{s}_{G_2}(r_2)$. Node r_2 is also a leaf node and $\mathfrak{s}_{G_1}(p_1)|_{|V_1|-1} \overset{\text{p}}{=} \mathfrak{s}_{G_2}(p_2)|_{|V_1|-1}$ for their neighbors $p_1 \in V_1$ and $p_2 \in V_2$. Applying the above claim to r_1, p_1, r_2, p_2 proves the isomorphism of G_1 and G_2. □

Note that the above proof provides a new polynomial time algorithm for trees. In fact, there exists a linear time algorithm to decide whether two trees are isomorphic [1].

In what follows, we prove that two 3-connected planar graphs are isomorphic if and only if they are 3-strongly walk isomorphism.

Lemma 3.10. *Let G be a 3-connected planar graph. If $i_1, i_2, i_3 \in V$ are three distinct nodes sharing a common face, then $\mathfrak{s}^3_G(i_1, i_2, i_3)_i \neq \mathfrak{s}^3_G(i_1, i_2, i_3)_j$ for all distinct $i, j \in V$.*

Proof. For all $k \in \mathbb{N}$, let γ_k be a $V \to \mathbb{R}^2$ function defined as follows. If $k = 0$, let

$$\gamma_0(i) := \begin{cases} (0,0), & \text{if } i = i_1 \\ (0,1), & \text{if } i = i_2 \\ (1,0), & \text{if } i = i_3 \\ (1,1), & \text{otherwise,} \end{cases} \tag{14}$$

for $k \geq 1$, let

$$\gamma_k(i) := \begin{cases} \gamma_{k-1}(i), & \text{if } i \in \{i_1, i_2, i_3\} \\ \frac{1}{\delta_G(i)} \sum\limits_{i' \in \Gamma_G(i)} \gamma_{k-1}(i'), & \text{otherwise.} \end{cases} \tag{15}$$

As k goes to infinity, γ_k converges to a planar embedding [19], therefore γ_k is an injection for sufficiently large k. Therefore it suffices to show that

$$\gamma_k(i) \neq \gamma_k(j) \Longrightarrow \mathfrak{s}^3_G(i_1, i_2, i_3)_{ik} \neq \mathfrak{s}^3_G(i_1, i_2, i_3)_{jk} \tag{16}$$

holds for all $i, j \in V$, which we prove by induction on k.

The base case, $\gamma_0(i) \neq \gamma_0(j) \Longrightarrow \mathfrak{s}^3_G(i_1, i_2, i_3)_{i0} \neq \mathfrak{s}^3_G(i_1, i_2, i_3)_{j0}$, easily follows by definition. By induction, suppose that (16) holds for $k - 1$, where $k \geq 1$.

If $i \in \{i_1, i_2, i_3\}$ or $j \in \{i_1, i_2, i_3\}$, then (16) holds, since the rows of $\mathfrak{s}^3_G(i_1, i_2, i_3)|_k$ corresponding to nodes $\{i_1, i_2, i_3\}$ are unique. Assume that $i, j \notin \{i_1, i_2, i_3\}$. By definition, $\gamma_k(i) \neq \gamma_k(j)$ means that

$$\frac{1}{\delta_G(i)} \sum_{i' \in \Gamma_G(i)} \gamma_{k-1}(i') \neq \frac{1}{\delta_G(j)} \sum_{j' \in \Gamma_G(j)} \gamma_{k-1}(j'). \tag{17}$$

If $\delta_G(i) \neq \delta_G(j)$, then (16) holds by the definition of $\mathfrak{s}^3_G(i_1, i_2, i_3)$. Otherwise, (17) implies that

$$\{\gamma_{k-1}(i') : i' \in \Gamma_G(i)\}^{\#} \neq \{\gamma_{k-1}(j') : j' \in \Gamma_G(j)\}^{\#} \tag{18}$$

which, by induction, means that

$$\{\mathfrak{s}^3_G(i_1, i_2, i_3)_{i'\,k-1} : i' \in \Gamma_G(i)\}^\# \neq \{\mathfrak{s}^3_G(i_1, i_2, i_3)_{j'\,k-1} : j' \in \Gamma_G(j)\}^\# \tag{19}$$

holds, and therefore $\mathfrak{s}^3_G(i_1, i_2, i_3)_{ik} \neq \mathfrak{s}^3_G(i_1, i_2, i_3)_{jk}$. □

Theorem 3.11. *Two 3-connected planar graphs, G_1 and G_2 are isomorphic if and only if $\mathfrak{s}^3_{G_1} = \mathfrak{s}^3_{G_2}$.*

Proof. It suffices to show that if $\mathfrak{s}^3_{G_1} = \mathfrak{s}^3_{G_2}$, then G_1 and G_2 are isomorphic. Let $i_1, i_2, i_3 \in V_1$ be three distinct nodes on a common face in some planar embedding of G_1. By definition, $\mathfrak{s}^3_{G_1} = \mathfrak{s}^3_{G_2}$ means that $\{\mathfrak{s}^3_{G_1}(i) : i \in V_1\}^\# = \{\mathfrak{s}^3_{G_2}(j) : j \in V_2\}^\#$, therefore there exists $j_1 \in V_2$ s.t. $\mathfrak{s}^3_{G_1}(i_1) = \mathfrak{s}^3_{G_2}(j_1)$. Similarly, one gets that there exists $j_2 \in V_2$ s.t. $\mathfrak{s}^3_{G_1}(i_1, i_2) = \mathfrak{s}^3_{G_2}(j_1, j_2)$, and there exists $j_3 \in V_2$ s.t. $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3) \overset{\mathrm{p}}{=} \mathfrak{s}^3_{G_2}(j_1, j_2, j_3)$. The following claim provides the sought bijection.

Claim 3.12. *There is a unique bijection $\pi : V_1 \longrightarrow V_2$ for which $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3)_i \overset{\mathrm{p}}{=} \mathfrak{s}^3_{G_2}(j_1, j_2, j_3)_{\pi(i)}$ holds for all $i \in V_1$, and this π is edge-preserving.*

Proof. By Lemma 3.10, the labels in G_1 are unique, that is

$$\mathfrak{s}^3_{G_1}(i_1, i_2, i_3)_i \overset{\mathrm{p}}{=} \mathfrak{s}^3_{G_1}(i_1, i_2, i_3)_{i'} \iff i = i' \tag{20}$$

follows. Since $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3) \overset{\mathrm{p}}{=} \mathfrak{s}^3_{G_2}(j_1, j_2, j_3)$, the labels in G_2 are unique too, i.e. one has that

$$\mathfrak{s}^3_{G_2}(j_1, j_2, j_3)_j \overset{\mathrm{p}}{=} \mathfrak{s}^3_{G_2}(j_1, j_2, j_3)_{j'} \iff j = j'. \tag{21}$$

Given that $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3) \overset{\mathrm{p}}{=} \mathfrak{s}^3_{G_2}(j_1, j_2, j_3)$, the unique existence of π easily follows from (20) and (21). In order to show that π is edge-preserving, observe that (20) and (21) hold even for the first $n+1$ columns of matrices $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3)$ and $\mathfrak{s}^3_{G_2}(j_1, j_2, j_3)$ by Claim 3.2. Accordingly, no two rows turn out to be different in column $(n+2)$. More precisely,

$$\{\mathfrak{s}^3_{G_1}(i_1, i_2, i_3)_{i'\,n+1} : i' \in \Gamma_{G_1}(i)\}^\# = \{\mathfrak{s}^3_{G_2}(j_1, j_2, j_3)_{j'\,n+1} : j' \in \Gamma_{G_2}(\pi(i))\}^\# \tag{22}$$

hold for all nodes $i \in V_1$. Observe that for all nodes $i \in V_1$

$$\{\pi(i') : i' \in \Gamma_{G_1}(i)\}^\# = \{j' : j' \in \Gamma_{G_2}(\pi(i))\}^\# \tag{23}$$

follows from (22), since the rows of matrices $\mathfrak{s}^3_{G_1}(i_1, i_2, i_3)\big|_{n+1}$ and $\mathfrak{s}^3_{G_2}(j_1, j_2, j_3)\big|_{n+1}$ uniquely identify the corresponding nodes. Equation (23) means that π is edge-preserving, which completes the proof. □

□

References

[1] J. E. Hopcroft and J. K. Wong. Linear time algorithm for isomorphism of planar graphs. *Proceeding STOC '74 Proceedings of the sixth annual ACM symposium on Theory of computing, Pages 172-184*, April 1974.

[2] Eugene M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *Journal of Computer and System Sciences, Volume 25, Issue 1, Pages 42-65*, August 1982.

[3] George S. Lue and Kellogg S. Booth. A linear time algorithm for deciding interval graph isomorphism. *Journal of the ACM (JACM), Volume 26, Issue 2, Pages 183-195, 1979*, April 1979.

[4] Charles J. Colbourn. On testing isomorphism of permutation graphs. *Networks, Volume 11, Issue 1, Pages 13-21*, March 1981.

[5] V. Arvind, B. Das, J. Köbler, and S. Toda. Colored hypergraph isomorphism is fixed parameter tractable. *Algorithmica Volume 71, Pages 120-138*, January 2015.

[6] B. D. McKay. Practical graph isomorphism. *Congressus Numerantium, Volume 30, Pages 45-87*, 1981.

[7] L. P. Cordella, P. Foggia, C. Sansone, and M. Vento. A (sub)graph isomorphism algorithm for matching large graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 26 Issue 10, Page 1367-1372*, 2004.

[8] Alpár Jüttner and Péter Madarasi. Vf2++an improved subgraph isomorphism algorithm. *Discrete Applied Mathematics*, 2018.

[9] Dennis Shasha, Jason T. L. Wang, and Rosalba Giugno. Algorithmics and applications of tree and graph searching. In *Proceedings of the Twenty-first ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, PODS '02, pages 39–52, New York, NY, USA, 2002. ACM.

[10] Edwin R. van Dam and Willem H. Haemers. Which graphs are determined by their spectrum? *Linear Algebra and its Applications*, 373(Supplement C):241 – 272, 2003. Combinatorial Matrix Theory Conference (Pohang, 2002).

[11] Richard C. Wilson and Ping Zhu. A study of graph spectra for comparing graphs and trees. *Pattern Recognition*, 41(9):2833 – 2841, 2008.

[12] D. Raviv, R. Kimmel, and A. M. Bruckstein. Graph isomorphisms and automorphisms via spectral signatures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1985–1993, Aug 2013.

[13] A E. Brouwer and Edward Spence. Cospectral graphs on 12 vertices. *Electr. J. Comb.*, 16, 06 2009.

[14] Brendan L Douglas and Jingbo B Wang. A classical approach to the graph isomorphism problem using quantum walks. *Journal of Physics A: Mathematical and Theoretical*, 41(7):075303, 2008.

[15] Kenneth Rudinger, John King Gamble, Mark Wellons, Eric Bach, Mark Friesen, Robert Joynt, and Susan Coppersmith. Noninteracting multiparticle quantum random walks applied to the graph isomorphism problem for strongly regular graphs. *Phys. Rev. A*, 86, 08 2012.

[16] A Mahasinghe, J A Izaac, J B Wang, and J K Wijerathna. Phase-modified ctqw unable to distinguish strongly regular graphs efficiently. *Journal of Physics A: Mathematical and Theoretical*, 48(26):265301, 2015.

[17] John King Gamble, Mark Friesen, Dong Zhou, Robert Joynt, and S. N. Coppersmith. Two-particle quantum walks applied to the graph isomorphism problem. *Physical Review A*, 81:052313, May 2010.

[18] Yonathan Aflalo, Alexander Bronstein, and Ron Kimmel. On convex relaxation of graph isomorphism. *Proceedings of the National Academy of Sciences*, 112(10):2942–2947, 2015.

[19] W. T. Tutte. How to Draw a Graph. *Proceedings of the London Mathematical Society*, s3-13(1):743–767, 01 1963.