Faster Isomorphism Testing of Strongly Regular Graphs*

Daniel A. Spielman†
UC Berkeley & MIT

Abstract

We demonstrate that isomorphism of strongly regular graphs may be tested in time $n^{O(n^{1/3}\log n)}$. Our approach is to analyze the standard individualization and refinement algorithm in light of Neumaier's claw bound, which implies that low degree strongly regular graphs have a small second-largest eigenvalue, unless they are Steiner or Latin square graphs.

1. Introduction

The problem of finding an algorithm that is guaranteed to efficiently determine whether two graphs are isomorphic has troubled many researchers [RC77]. For many years, no known algorithm had a worst-case running time better than $n^{O(n)}$. While this problem is polynomial-time solvable on average in many senses of the word [BK79, Kuč87], there are many classes of graphs that do not seem amenable to the heuristics that work on "average" graphs. For a while, the class of strongly regular graphs seemed to capture much of the difficulty of the isomorphism problem. But, in 1980, Babai [Bab80] proved that a simple combinatorial algorithm would test isomorphism of strongly regular graphs in time $n^{O(\sqrt{n}\log n)}$. Far more efficient tests of isomorphism were found for Latin square graphs and Steiner triple graphs [Mil78], graphs with bounded color class [Bab79], graphs of bounded genus [Mil80, FM80, Lic80], graphs with bounded eigenvalue multiplicity [BGM82], and graphs of bounded degree [Luk82]. For generalizations of these classes, see [Mil83] and [Pon89]. By combining Zemlyachenko's degree reduction technique with Luks's algorithm for graphs of bounded degree, it is possible to solve the general graph isomorphism problem in time $2^{\sqrt{O(n\log n)}}$ (See [Bab81a, BL83, ZKT85]). Thus, the known complexity of determining isomorphism of strongly regular graphs became no better than that for arbitrary graphs.

We will show that naive, known, combinatorial algorithms can be used to test isomorphism of strongly regular graphs in time $n^{O(n^{1/3}\log n)}$. We begin by observing that the algorithm used in [Bab80] only takes a long time on strongly regular graphs of relatively low degree. We then apply a theorem of Neumaier [Neu79] to show that such graphs have small second-largest eigenvalue or are Latin square graphs or Steiner graphs. In either case, we can obtain an improved analysis. Our analysis of isomorphism testing of strongly regular graphs with small second-largest eigenvalue is inspired by the intuition that such graphs are "quasi-random" (See [CGW89] for related notions).

In Section 2, we recall the definitions of strongly regular graphs, Latin square graphs, and Steiner graphs as well as various facts about these structures that will aid our analysis. The reader may want to skim this section and return to it when necessary. In Section 3, we review the basic individualization and refinement approach to testing graph isomorphism. We analyze the performance of the individualization and refinement algorithm for strongly regular graphs with small second-largest eigenvalue in Section 4, and for Steiner graphs in Section 5. We then combine the results of these sections to obtain our main theorem.

2. Strongly Regular Graphs and Partial Geometries

*Errata to this paper will be available at `http://theory.lcs.mit.edu/~spielman`.

†Computer Science Division, U.C. Berkeley, CA 94720, `spielman@cs.berkeley.edu`. Supported in part by an NSF postdoc. After July 1996, Department of Mathematics, M.I.T.

Definition 1. A *strongly regular graph* (SRG) with parameters (n, k, λ, μ) is a graph on n vertices such that

- each vertex in G has degree k,
- each pair of neighbors in G have exactly λ common neighbors, and
- each pair of non-neighbors in G have exactly μ common neighbors.

As the complement of a strongly regular graph is also strongly regular, we can assume that $k \leq n/2$.

There is a wonderful theory of strongly regular graphs which we will not be able to present here (See Figure 1 for an example). We just state a few facts that we will need. For more information about strongly regular graphs, consult [Bos63, Sei79, Cam78, vLW92].

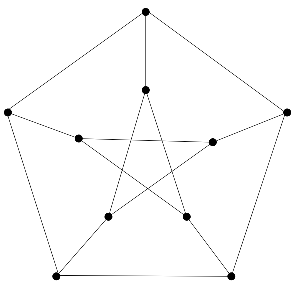

Figure 1: A strongly regular graph with parameters $(10, 3, 0, 1)$.

There are some *trivial* examples of strongly regular graphs. These are the disjoint unions of cliques, and the complements of these graphs. Isomorphism of such graphs is easy to determine, so we will not consider them in this paper. All other strongly regular graphs are connected and have degree $k \geq \sqrt{n-1}$.

Proposition 2. *Let G be a connected, non-trivial, strongly regular graph with parameters (n, k, λ, μ). Then,*

(a) $(n - k - 1)\mu = k(k - 1 - \lambda)$.

(b) The adjacency matrix of G has just three eigenvalues, $k > r \geq 0 \geq s$, of multiplicities 1, f, and g respectively, that satisfy

$$\begin{aligned} k + fr + gs &= 0, & \mu - rs &= k, \\ k^2 + fr^2 + gs^2 &= nk, \text{ and } & \lambda - \mu &= r + s. \end{aligned}$$

Definition 3. A *Steiner 2-system* with parameters (s, h, t) is a set system consisting of *points* and *lines* such that

1. any two points lie on exactly one line,
2. each line contains s points, and each point intersects h lines,
3. through any point x and any line l that does not contain x, there are exactly t lines through x that meet l.

Definition 4. A *Steiner graph* is the line graph of a Steiner 2-system—the graph whose vertices represent lines in the Steiner system such that two vertices are neighbors if their corresponding lines intersect.

Definition 5. An *s-net* is a set system consisting of *points* and *lines* such that the number of points is a square, say m^2, and the lines are divided into s classes of m lines each. Each line contains exactly m points, lines in the same class do not intersect, and lines in different classes intersect in exactly one point.

For example, a 3-net is a Latin square: The points can be indexed by pairs (i, j) where $1 \leq i, j \leq m$. The first set of lines can consist of collections of points of fixed first coordinate and the second set of points of fixed second coordinate. Lines in the third set then have exactly one point with each first coordinate and one point with each second coordinate (See Figure 2).

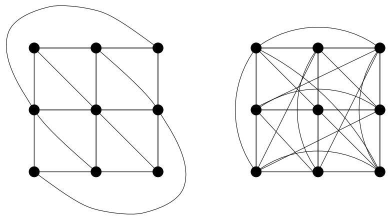

Figure 2: A Latin square with $m = 3$, and its graph

Definition 6. A *Latin square graph* is the point graph of an s-net: the graph contains a vertex for each point in the s-net and two vertices share an edge if there is a line through their corresponding points. (*warning:* the names of these objects vary throughout the literature.)

Proposition 7. *Steiner graphs and Latin square graphs are strongly regular.*

The cornerstone of this paper is the following theorem of Neumaier [Neu79].

Theorem 8 (Neumaier). *Let G be a strongly regular graph with parameters (n, k, λ, μ) and eigenvalues $k > r > s$. Then, at least one of the following conditions must hold*

(*a*) $r \leq \max\left\{2(-s-1)(\mu+1+s), \frac{s(s+1)(\mu+1)}{2} - s - 1\right\}$,

(*b*) $\mu = s^2$, *in which case G is a Steiner graph derived from a Steiner 2-system in which each line contains s points.*

(*c*) $\mu = s(s+1)$, *in which case G is a Latin square graph derived from an s-net.*

Condition (a) is known as *Neumaier's claw bound.* It tells us that the ratio of the first two eigenvalues of a strongly regular graph with $k = o(n)$ is small, unless it is a Latin square or Steiner graph. In either case, we can take advantage of this structure to test isomorphism.

Corollary 9. *Let G be a strongly regular graph with parameters (n, k, λ, μ) and eigenvalues $k > r > s$ that satisfies Neumaier's claw bound and such that $k = o(n)$. Then, $r = o(k)$.*

Proof: From part (a) of Proposition 2, we find that $\mu = o(k)$. From the claw bound and the fact that $2(-s-1)(\mu+1+s) \leq s^2(\mu+1)$, we find

$$r \leq s^2(\mu+1).$$

Following Neumaier [Neu79] (partially), we let $m = -s$, because s is negative. Thus, $k = \mu + rm$, from part (b) of Proposition 2, implies

$$k \leq rm(1 + o(1)).$$

Combining the last two inequalities, we obtain

$$\begin{aligned} k &\leq m^3(\mu+1)(1+o(1)) \Rightarrow \\ m &\geq \left(\frac{k}{(\mu+1)(1+o(1))}\right)^{1/3}, \text{ and} \\ r &\leq k^{2/3}\mu^{1/3}(1+o(1)), \end{aligned}$$

which, together with $\mu = o(k)$, implies $r = o(k)$. □

This corollary does not begin to take advantage of the power of Neumaier's theorem, but it is all that we will need for the results in Section 4.

If a strongly regular graph is of type (b) or (c) and if μ is not too large, then one can compute the Steiner 2-system or s-net from the graph in polynomial time. On the other hand, if μ is large, then we will show that r is small. The next two statements and proofs are analogous to a theorem of Miller [Mil78].

Proposition 10. *If G is a Steiner graph on n vertices derived from a Steiner 2-system in which each line has s points and $\sqrt{n}-2 > (s-1)^2$, then one can reconstruct the Steiner 2-system in time polynomial in n.*

Proof: Recall that n is the number of lines in the Steiner 2-system. Let v be the number of points in the system and let h be the number of lines through each point. Consider two lines, l_1 and l_2, of the Steiner 2-system that intersect at a point p. There are $h - 2 + (s-1)^2$ lines that intersect both l_1 and l_2: $h-2$ that also go through p and $(s-1)^2$ that go through other points on l_1 and l_2.

Now, examine the corresponding structures in the Steiner graph. The lines l_1 and l_2 correspond to vertices in the graph and p an edge between them. The $h - 2 + (s-1)^2$ lines that intersect l_1 and l_2 correspond to vertices that are neighbors of the vertices representing l_1 and l_2. Our algorithm needs to distinguish the $h-2$ vertices corresponding to the lines through p. To do this, note that these lines meet each other, and they cannot meet the other $(s-1)^2$ lines. Thus, if $h - 2 > (s-1)^2$, then they are distinguished by their degrees in the graph induced on the mutual neighbors of the vertices corresponding to l_1 and l_2.

We now complete the proof by showing that $h > \sqrt{n}$. By counting intersections of points with lines, we compute

$$\begin{aligned} \binom{v}{2} &= n\binom{s}{2}, \text{ and} && (1) \\ vh &= ns. && (2) \end{aligned}$$

Because there are more points than points per line, $v > s$. Together with (1), this implies that $v^2 < ns^2$. Plugging this into (2), we obtain $h > \sqrt{n}$. □

Proposition 11. *If G is a Latin Square graph on n vertices derived from an s-net and $n > (s-1)^4$, then one can reconstruct the s-net in time polynomial in n.*

Proof: Similar to the proof of Proposition 10. Also see [Mil78, Theorem 8]. □

In the cases where we cannot apply Proposition 10 or 11, we will apply the following lemma to obtain a situation like we have for graphs of type (a).

Lemma 12. *If G is a strongly regular graph of type (b) or (c) with $k = o(n^{2/3})$ and $s = \Omega\left(n^{1/4}\right)$, then $r = o(k)$.*

Proof: Since $k = o(n)$, we can use part (a) of Proposition 2 to show that $\mu = o(k)$. Combining this and $s = \Omega\left(n^{1/4}\right)$ with $\mu - rs = k$ from part (b) of Proposition 2 we find $r = o(k)$. $\square$

Miller [Mil78] has shown that isomorphism of s-nets can be decided in time $O\big(n^{\log n + O(1)}\big)$. In Section 5, we will show how to test isomorphism of Steiner 2-systems, and in Section 4 we will show how to test isomorphism of strongly regular graphs that satisfy Neumaier's claw bound or the conditions of Lemma 12.

3. Individualization and Refinement

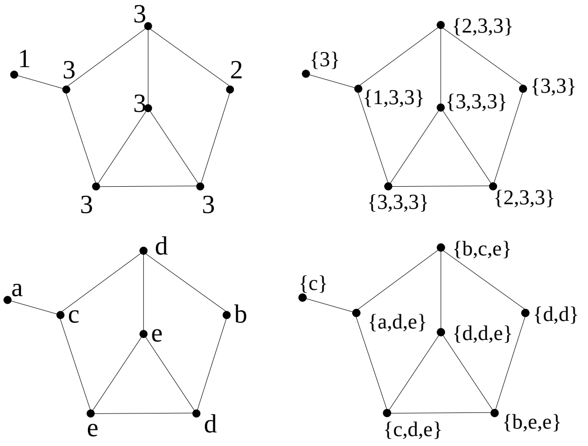

Figure 3: Animation of a refinement process that produced a canonical label.

A standard heuristic for testing isomorphism of graphs is to try to assign to each graph a *canonical label* so that two graphs are isomorphic if and only if they have the same label. One approach to finding a canonical label is to find a canonical way to assign a unique label to each vertex in a graph. For example, we could begin by labeling vertices by their degrees. We could then refine this labeling by creating, for each vertex, the list of labels of its neighbors, and assigning new labels to the vertices corresponding to the lexicographic order of these lists (See Figure 3). This process is called *refinement*. We keep refining the labeling until we cannot make any more progress. If, after refinement, every vertex has a distinct label, then we have produced a canonical label for the graph. Note that this process will terminate in polynomial time.

On the other hand, this process won't get very far if it begins with a graph in which every vertex has the same degree. A natural way to try to break the symmetry is

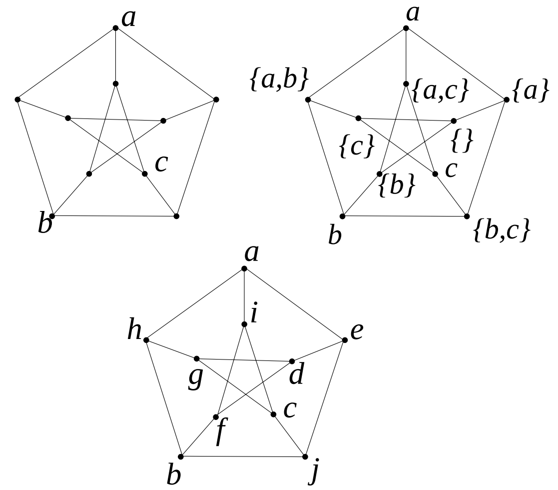

Figure 4: Nodes a, b, and c were individualized. After one refinement step, every vertex in the graph had a unique label. Note that it is important to replace the neighbor lists with shorter labels: if we don't, then the labels will quickly become too long to write down.

to *individualize* a particular vertex—assign it a unique label—and then repeatedly refine the labeling in hope of obtaining a unique label for each vertex (See Figures 4 and 5 for examples of this process). If this works, one could obtain a canonical label for the graph by enumerating over the n labels obtained by individualizing different vertices, and then using the lexicographically least label.

In fact, one could try individualizing all choices of k vertices at a cost of $O\big(n^k\big)$ time. Babai [Bab80] proved that a canonical label can be assigned to a strongly regular graph by individualizing $O\big(n^{1/2}\log n\big)$ vertices and refining once. In fact, it suffices to individualize $n(\log n)/k$ vertices [Bab81b]:

Theorem 13 (Babai). *Let G be a strongly regular graph with parameters (n, k, λ, μ). Then, there exists a set of $O(n(\log n)/k)$ vertices in G whose individualization will result in a unique labeling of every vertex in G after one refinement step. In fact, a randomly chosen set of $n(\log n)/k$ vertices usually suffices.*

4. Testing Isomorphism when $r = o(k)$

In this section, we will show that every strongly regular graph that has degree at most $k = o(n^{2/3})$ and second-largest eigenvalue $r = o(k)$ has a set of $O\big(n^{1/4}\sqrt{\log n}\big)$ vertices whose individualization will result in a unique labeling of the vertices in G after two refinement steps.

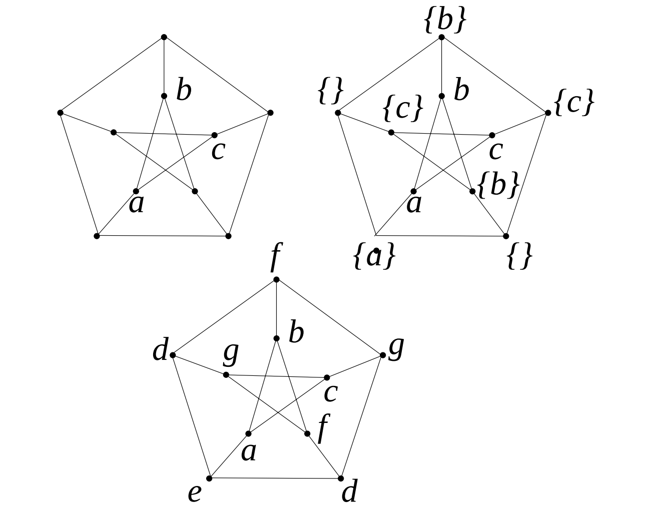

Figure 5: An attempt at individualizing a different set of nodes that did not result in a unique label for each vertex. Note that the current labels cannot be refined any further.

The basic idea is the following: consider what happens if we choose two non-neighbors at random. They will have μ common neighbors, and these common neighbors will each have k neighbors. Thus, these two non-neighbors define a set of roughly μk vertices. If a vertex x is in this set but y is not, then we say that these non-neighbors *distinguish* x from y, because the individualization of these non-neighbors will cause x and y to have different labels after two refinements (See Figure 6). The quasi-random property of the graph should imply that the event that x is in the set is only weakly correlated with the event that y is in the set. Thus, the probability that the set distinguishes x from y should be roughly $\mu k/n$. Because we individualize $O\big(n^{1/4}\sqrt{\log n}\big)$ vertices, we expect to get $O(n^{1/2}\log n)$ such sets. Again, the quasi-random property of the graph should imply that these sets act independently, so the probability that x is not distinguished from y should be

$$(1-\mu k/n)^{n^{1/2}\log n},$$

which is small for $n^{1/2} \leq k \leq o(n^{2/3})$, and $\mu \simeq k^2/n$.

We now make this intuitive argument formal.

Lemma 14. *Let G be a strongly regular graph with parameters (n,k,λ,μ) such that $r = o(k)$ and $k = o(n^{2/3})$. Then, individualization of a random set of $O\big(n^{1/4}\sqrt{\log n}\big)$ vertices of G followed by two refinement steps will give a unique name to every vertex in G with high probability.*

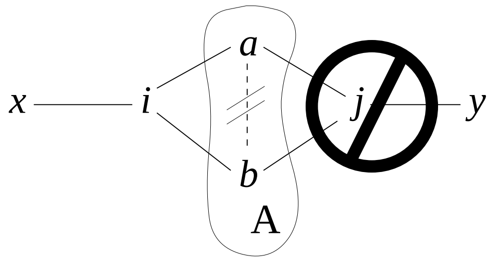

Figure 6: We want there to be vertices a and b in A that have a common neighbor with x, but do not have a common neighbor with y.

Proof: We begin by showing that each pair of vertices, x and y, will probably receive different labels. We will then sum over all pairs.

In the proof, we will make use of the inequalities $\mu^2 = o(k)$, $\lambda = o(k)$, and $\lambda k = o(\mu n)$, which follow, respectively, by combining $k = o(n^{2/3})$ with part (a) of Proposition 2, combining $\mu = o(k)$ with $r = o(k)$ and $\lambda - \mu = r + s$, and combining $\lambda = o(k)$ with part (a) of Proposition 2.

Let A be a randomly chosen subset of the vertices of G into which each vertex is placed independently with probability $cn^{-3/4}\sqrt{\log n}$, for some constant c. We will estimate the probability that there are two non-neighbors a and b in A such that a and b have a common neighbor i that is a neighbor of x, but have no common neighbor j that is a neighbor of y (see Figure 6).

Let $\{v_1,\ldots,v_k\}$ be the set of neighbors of x, and let S_i be the set of neighbors of v_i that are not neighbors of x. Each vertex z that is not a neighbor of x is in exactly μ of these sets, because x and z share μ common neighbors. Let $\{v_1,\ldots,v_l\}$ be the vertices that are common neighbors of x and y (clearly $l \leq \lambda$) and let W be the set of vertices in G that are in at least $\mu/2$ of the sets $\{S_{l+1},\ldots,S_k\}$. Corollary 9 implies that $lk \leq \lambda k = o(\mu n)$, so $|W| = n(1-o(1))$. We now apply Lemma 15 to find a set system

$$\{T_1,\ldots,T_{\epsilon k/\mu}\} \subseteq \{S_{l+1},\ldots,S_k\}$$

such that each set T_i contains at least δk elements of W that are in no other set T_j, for some constants ϵ and δ. Let U_i denote those elements that of W that are in T_i exclusively. We will show that, with high probability, some set U_i contains the elements a and b that we desire.

Write $p = n^{-3/4}\sqrt{\log n}$. The probability that U_i contains at least two elements of A is at least

$$1-(1-p)^{\delta k} - p\delta k\,(1-p)^{\delta k-1}$$

$$
\begin{aligned}
&> 1-\left(e^{-p}\right)^{\delta k} - p\delta k\left(e^{-p}\right)^{\delta k-1} \\
&> 1-\left(1-p\delta k+\frac{p^2(\delta k)^2}{2}\right) \\
&\quad -p\delta k\left(1-p(\delta k-1)+\frac{p^2(\delta k-1)^2}{2}\right) \\
&= \Omega\left(p^2(\delta k)^2\right).
\end{aligned}
$$

Given that U_i contains at least two elements of A, we want to compute the probability that they are non-neighbors and do not have a common neighbor that is also a common neighbor of y. Well, U_i contains at least $\delta k(\delta k-1)$ ordered pairs of distinct elements. Of these, $k\lambda$ are pairs that are neighbors. Lemma 16 tells us that, of the $\delta k(\delta k-1)-k\lambda$ pairs that are not neighbors, only $\mu(\mu-1)(k-\lambda)$ have a common neighbor that is also a neighbor of y. Thus, the fraction of ordered pairs of distinct elements of U_i that satisfy the desired conditions is

$$
\frac{\delta k(\delta k-1)-k\lambda-\mu(\mu-1)(k-\lambda)}{\delta k(\delta k-1)} = (1-o(1)),
$$

because $\lambda = o(k)$, $k = o(n)$, and $\mu^2 = o(k)$. So, the probability that U_i contains two elements of A that are non-neighbors and which do not have a common neighbor with y is

$$
\Omega\left(p^2(\delta k)^2\right)(1-o(1)) = \Omega\left(n^{-3/2}k^2c^2\log n\right).
$$

Thus, the probability that there is no set U_i that contains two such elements of A is at most

$$
\left(1-\Omega\left(n^{-3/2}k^2c^2\log n\right)\right)^{\epsilon k/\mu} = e^{-\Omega\left(c^2\log n\right)},
$$

because $n^{-3/2}k^3/\mu = \Omega(1)$ for $k > n^{1/2}$. That is, the probability that x and y receive different labels is at least $1-e^{-\Omega\left(c^2\log n\right)}$. Summing over the $n(n-1)/2$ possible choices for x and y, we see that each pair of vertices will receive different labels with probability at least $1-n^2e^{-\Omega\left(c^2\log n\right)}$, which is close to 1 for a sufficiently large constant c. Moreover, the probability that the size of A deviates from its expectation by more than a constant factor is very small. □

The rest of this section is devoted to proving the lemmas that we needed in the proof of Lemma 14.

Lemma 15. *There exist constants ϵ and δ such that for any set W and any family $\mathcal{A}$ of a subsets of W such that*

1. *each set in A has at most k elements, and*
2. *each $v \in W$ is contained in at least $\mu/2$ but fewer than μ of the sets in $\mathcal{A}$,*

for a/μ larger than some constant, there exists a subfamily $\mathcal{B} \subset \mathcal{A}$ such that

1. *$|\mathcal{B}| > \epsilon k/\mu$, and*
2. *for each $S \in \mathcal{B}$, there are at least δk elements of W that are contained in S but not in any other set in $\mathcal{B}$,*

Proof: Let n be the number of elements in W. By counting the number of containments of points in sets two ways, we determine

$$
ak \le \mu n \Rightarrow \frac{n\mu}{k} \ge a.
$$

For each set $S \in \mathcal{A}$, include S in $\mathcal{C}$ with probability $1/\mu$. Let U be the vertices in W that are contained in just one set in $\mathcal{C}$. For a $v \in W$ that is contained in p sets in $\mathcal{A}$, the probability that $v \in U$ is

$$
p\frac{1}{\mu}\left(1-\frac{1}{\mu}\right)^{p-1} \ge \frac{\mu}{2}\frac{1}{\mu}\left(1-\frac{1}{\mu}\right)^{\mu/2-1} \ge \frac{1}{2\sqrt{e}}.
$$

Because the expected size of U is at least $n/2\sqrt{e}$, the probability that U has size greater than $n/4\sqrt{e}$ is at least $1/4\sqrt{e}$. Since the size of $\mathcal{C}$ is the sum of independent random variables, one can use a Chernoff bound to show that the probability that $\mathcal{C}$ has size greater than twice its expectation, $2a/\mu$, is small, provided that a/μ is at least some large constant. Thus, we can assume that there exists a family of sets $\mathcal{C}$ of size at most $2a/\mu$ so that U has size at least $n/4\sqrt{e}$. Let $\epsilon_1 = 1/4\sqrt{e}$.

As each set in $\mathcal{C}$ has at most k elements, a simple counting argument shows that at least $\epsilon_1 a/2\mu$ sets in $\mathcal{C}$ must each have at least $\epsilon_1 k/4$ elements of U: Otherwise, we have at most $\epsilon_1 a/2\mu$ sets with at most k elements of U, and at most $2a/\mu$ sets with less than $\epsilon_1 k/4$ elements of U, for a total of fewer than

$$
\left(\epsilon_1\frac{a}{2\mu}\right)k+\frac{2a}{\mu}\left(\epsilon_1\frac{k}{4}\right) \le \frac{\epsilon_1 ka}{\mu} \le \epsilon_1 n
$$

elements of U, which would be a contradiction. These $\epsilon_1 a/2\mu$ sets are the subfamily $\mathcal{B}$ that we desire. Thus, we have proved the lemma with $\epsilon = \epsilon_1/2$ and $\delta = \epsilon_1/4$. □

Lemma 16. *Let G be a strongly regular graph with parameters (n,k,λ,μ). Let i and y be non-neighbors in G. Then, there are at most*

$$
2\mu(\mu-1)(k-\lambda)
$$

ordered triples (a,b,j) such that (See Figure 7)

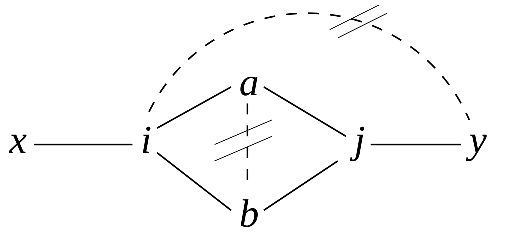

Figure 7: The configurations that we count in Lemma 16. x, i, and y are fixed.

(a) a and b are neighbors of i,

(b) a and b are neighbors of j,

(c) j is a neighbor of y, and

(d) a and b are non-neighbors.

Proof: We first count the number of triples that satisfy conditions (a) through (c). We will later remove those that violate (d).

There are μ j's such that j is a neighbor of both i and y. For each such j there are $\lambda(\lambda-1)$ ordered pairs (a, b) such that a and b are distinct common neighbors of i and j. Thus, from the j's which are neighbors of i, we obtain a contribution of $\mu\lambda(\lambda-1)$ triples.

Similarly, there are $(k-\mu)$ j's that are neighbors of y but not of i. For each such j there are $\mu(\mu-1)$ ordered pairs (a, b) such that a and b are distinct common neighbors of i and j. Thus, from the j's which are not neighbors of i, we obtain $(k-\lambda)\mu(\mu-1)$ triples.

We now observe that we have over-counted because we failed to exclude those triples in which a and b are neighbors. Many such triples are accounted for by those j that are neighbors of i. By Lemma 17, for each j that is a neighbor of i, there are at least

$$\lambda(\lambda-1)-(\mu-1)(k-\lambda)$$

ordered pairs (a, b) such that a and b are neighbors and are both neighbors of i and j as well. Thus, we can remove at least $\mu(\lambda(\lambda-1)-(\mu-1)(k-\lambda))$ triples from our count, for a total of

$$\begin{aligned} \mu\lambda(\lambda-1)+(k-\lambda)\mu(\mu-1)-\mu\left(\lambda(\lambda-1)-(\mu-1)(k-\lambda)\right) \\ = 2\mu(\mu-1)(k-\lambda). \end{aligned}$$

□

Lemma 17. *Let G be a strongly regular graph with parameters (n, k, λ, μ). Then, for every pair of neighbors i and j in G, there are at least $\lambda(\lambda-1)-(\mu-1)(k-\lambda)$ ordered pairs (a, b) such that a and b are neighbors and both a and b are common neighbors of i and j.*

Proof: Consider the ordered pairs of vertices (a, b) such that a is a neighbor of i and j, and b is a neighbor of a and j other than i. Because i and j have λ common neighbors and a and j have $\lambda-1$ common neighbors other than i, there will be $\lambda(\lambda-1)$ such pairs.

Now, consider a $b \in \overline{N(i)} \cap N(j)$. Such a b can account for at most $\mu-1$ ordered pairs, because b and i have only μ common neighbors, of which one is j. Thus, there are at least

$$\lambda(\lambda-1)-(\mu-1)(k-\lambda)$$

ordered pairs (a, b) such that (a, b) is an edge in $N(i) \cap N(j)$. □

5. Testing Isomorphism of Steiner Graphs

Steiner 2-systems in which each line contains two points are called *triangle graphs*. There is at most one triangle graph of any size, so determining isomorphism of these is trivial [Cam78].

Steiner 2-systems in which each line contains three points are the famous Steiner triple systems. Miller [Mil78] showed how to determine isomorphism of the line graphs of Steiner triple systems in $O\left(n^{\log n+O(1)}\right)$ time.

For general Steiner 2-systems, we will test isomorphism in time $n^{O\left(n^{1/4} \log n\right)}$. Our approach will again be that of individualization and refinement, but on the system of points and lines. To assign a unique name to every point in a Steiner 2-system, we first choose a collection of points at random. Since there is at most one line through any two points, this will give a unique name to every line through two of these points. Similarly, each point that lies on two lines with unique names can be given a unique name. We will show that $O\left(n^{1/4} \log n\right)$ points suffice to give enough lines unique names so that each point lies at the intersection of two of these lines.

Lemma 18. *Let S be a Steiner 2-system with n lines and $s>2$ points per line. Then there is a set, A, of $O\left(n^{1/4} \log n\right)$ points such that every point in the system lies on the intersection of two lines that each have at least two points in A.*

Proof: Let v be the number of points in the system and let h be the number of lines through each point. It is easy to see that

$$\binom{v}{2}=n\binom{s}{2}, \quad \text{and} \quad vh=ns.$$

We throw each point into the set A independently with probability

$$\frac{n^{1/4}\log n}{v} \geq \frac{n^{1/4}\log n}{s\sqrt{n}} = \frac{\log n}{sn^{1/4}}.$$

Since each line contains s points, the probability that a given line will contain at least two points of A is

$$1-\left(1-\tfrac{\log n}{sn^{1/4}}\right)^{s} - s\tfrac{\log n}{sn^{1/4}}\left(1-\tfrac{\log n}{sn^{1/4}}\right)^{s-1} = \theta\left(\left(\tfrac{\log n}{n^{1/4}}\right)^{2}\right).$$

For a given point x, the probability that there will be two lines through x that each contain at least two points of A is at least

$$1-\left(1-\theta\left(\tfrac{\log^2 n}{n^{1/2}}\right)\right)^{h} - h\tfrac{\log^2 n}{n^{1/2}}\left(1-\theta\left(\tfrac{\log^2 n}{n^{1/2}}\right)\right)^{h} \leq 1-e^{-\theta\left(\log^2 n\right)} - n^{1/2}(\log^2 n)e^{-\theta\left(\log^2 n\right)},$$

because h is at least $\sqrt{n}$ (as demonstrated in the proof of Proposition 10) and at most n. Thus, with very high probability, every point in S lies on at least two lines with two points in A. Moreover, it is highly unlikely that the number of elements in A will be a constant factor greater than $n^{1/4}\log n$. □

Theorem 19. *Isomorphism of strongly regular graphs can be determined in* $n^{O\left(n^{1/3}\log n\right)}$ *time.*

Proof: For graphs of degree greater than $o(n^{2/3})$, we use Babai's analysis of individualization and refinement, Theorem 13. For the remaining graphs, we use Theorem 8 to divide them into three classes. For those of class (a), those that meet Neumaier's claw bound, we apply Lemma 14 to show that a canonical labeling can be found in $n^{O\left(n^{1/4}\log n\right)}$ time. For the Latin square and Steiner graphs with $n = \Omega\left(s^4\right)$, we apply Lemma 12 and Lemma 14 to find the canonical labeling. For the remaining Latin square and Steiner graphs, we apply Proposition 11 or Proposition 10 to find the corresponding s-net or Steiner 2-System and then apply either Miller's algorithm [Mil78] or Lemma 18 to test for isomorphism. □

6. Conclusions

We have presented a better analysis of the performance of the individualization-and-refinement isomorphism test for strongly regular graphs. We suspect that it is possible to do better. For example, we have introduced a strange jump in the time complexity of testing isomorphism of strongly regular graphs: for graphs of degree $o(n^{2/3})$, we find a canonical label by individualizing $O\left(n^{1/4}\sqrt{\log n}\right)$ vertices, whereas we need to individualize $O\left(n^{1/3}\log n\right)$ vertices for graphs of degree $n^{2/3}$. We cannot believe that this jump is necessary.

Rather, we feel that it should be possible to take advantage of the quasi-random properties of low degree strongly regular graphs to analyze the occurrence of other structures in these graphs and thereby obtain a better time bound on isomorphism testing. Rather than just trying to find a unique name for every vertex in a graph, it should be possible to show that they will be broken into classes of limited size, and then apply techniques for testing isomorphism of graphs with bounded color classes. We note that if analyses such as these are going to be applied to general graphs, then it will be necessary to use such technology because Cai, Fürer, and Immerman [CFI92] have constructed graphs that require the individualization of $\Omega\left(n\right)$ vertices before refinement will give each vertex a unique name. We are encouraged by the idea that a separation of the first from the second eigenvalue of a graph might be able to aid in the analysis of refinement, because graphs that lack such a separation have small isoperimetric numbers [SJ89, Moh89], which we feel should be of assistance in testing isomorphism.

Acknowledgement: I would like to thank Shang-Hua Teng for many enlightening conversations about graph isomorphism, and for reading an early draft of this paper.

References

[Bab79] László Babai. Monte Carlo algorithms in graph isomorphism testing. Technical Report 79-10, Dép. Math et Stat., Univ. de Montréal, 1979.

[Bab80] László Babai. On the complexity of canonical labeling of strongly regular graphs. *SIAM J. Comput.*, 9(1):212–216, 1980.

[Bab81a] László Babai. Moderately exponential bound for graph isomorphism. In *Fundamentals of Computation Theory*, number 117 in Lecture Notes in Math, pages 34–50. Springer-Verlag, Berlin-Heidelberg-New York, 1981.

[Bab81b] László Babai. On the order of uniprimitive permuation groups. *Annals of Mathematics*, 113:553–568, 1981.

[BGM82] László Babai, D. Yu. Grigoryev, and David M. Mount. Isomorphism of graphs with bounded eigenvalue multiplicity. In *Proceedings of the 14th Annual ACM Symposium on Theory of Computing*, pages 310–324, 1982.

[BK79] László Babai and Ludik Kučera. Canonical labelling of graphs in linear average time. In *Proceedings of the 1920th IEEE Symposium on Foundations of Computer Science*, pages 39–46, 1979.

[BL83] László Babai and Eugene M. Luks. Canonical labeling of graphs. In *Proceedings of the 15th Annual ACM Symposium on Theory of Computing*, pages 171–183, 1983.

[Bos63] R. C. Bose. Strongly regular graphs, partial geometries and partially balanced designs. *Pacific J. Math.*, 13:389–419, 1963.

[Cam78] P. J. Cameron. Strongly regular graphs. In *Selected Topics in Graph Theory*, pages 337–380. Academic Press, London, 1978.

[CFI92] Jin-Yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, 1992.

[CGW89] F.R.K. Chung, R.L. Graham, and R.M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[FM80] I. S. Filotti and Jack N. Mayer. A polynomial-time algorithm for determining the isomorphism of graphs of fixed genus (working paper). In *Proceedings of the 12th Annual ACM Symposium on Theory of Computing*, pages 236–243, 1980.

[Kuč87] Luděk Kučera. Canonical labeling of regular graphs in linear average time. In *Proceedings of the 28th IEEE Symposium on Foundations of Computer Science*, pages 271–279, 1987.

[Lic80] David Lichtenstein. Isomorphism for graphs embeddable on the projective plane. In *Proceedings of the 12th Annual ACM Symposium on Theory of Computing*, pages 218–224, 1980.

[Luk82] Eugene M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *Journal of Computer and System Sciences*, 25:42–65, 1982.

[Mil78] Gary L. Miller. On the $n^{\log n}$ isomorphism technique: A preliminary report. In *Proceedings of the 10th Annual ACM Symposium on Theory of Computing*, pages 51–58, 1978.

[Mil80] Gary Miller. Isomorphism testing for graphs of bounded genus. In *Proceedings of the 12th Annual ACM Symposium on Theory of Computing*, pages 225–235, 1980.

[Mil83] Gary L. Miller. Isomorphism of k-contractible graphs. a generalization of bounded valence and bounded genus. *Information and Control*, 56:1–20, 1983.

[Moh89] B. Mohar. Isoperimetric numbers of graphs. *Journal of Combinatorial Theory, Series B*, 47:274–291, 1989.

[Neu79] A. Neumaier. Stronly regular graphs with smallest eigenvalue $-m$. *Arch. Math.*, 33:392–400, 1979.

[Pon89] Ilja N. Ponomarenko. The isomorphism problem for classes of graphs. *Soviet Math. Dokl.*, 39:119–122, 1989.

[RC77] R. C. Read and D. G. Corneil. The graph isomorphism disease. *J. Graph Theory*, 1:339–363, 1977.

[Sei79] J. J. Seidel. Strongly regular graphs. In B. Bollobás, editor, *Surveys in Combinatorics*, volume 38 of *London Mathematical Society Lecture Note Series*, pages 157–180. Cambridge University Press, 1979.

[SJ89] Alistair Sinclair and Mark Jerrum. Approximative counting, uniform generation and rapidly mixing Markov chains. *Information and Computation*, 82(1):93–133, July 1989.

[vLW92] J. H. van Lint and R. M. Wilson. *A Course in Combinatorics*. Cambridge University Press, 1992.

[ZKT85] V. M. Zemlyachenko, N. M. Kornienko, and R. I. Tyshkevich. Graph isomorphism problem. *Journal of Soviet Mathematics*, 29:1426–1481, 1985.