Pak. J. Engg. & Appl. Sci. Vol. 18 January, 2016 (*p.* 50–55)

Cycle Discrepancy of *d*-Colorable Graphs

Laeeq Aslam*, Shahzad Sarwar*, Muhammad Murtaza Yousaf*, and Waqar ul Qounain*

* Punjab University College of Information Technology(PUCIT), University of the Punjab, Lahore, Pakistan
* **Corresponding Author:** E-mail: laeeq.aslam@pucit.edu.pk

Abstract

We show that cycle discrepancy of a 3-colorable graph, G, on at least five vertices is bounded by $2\times\lceil\lfloor n/3\rfloor/2\rceil$*; that is,* $\mathrm{cycdisc}(G)\leq 2\times\lceil\lfloor n/3\rfloor/2\rceil$. *We also show that this bound is best possible by constructing 3-colorable graphs, on at least five vertices for which cycle discrepancy is at least* $2\times\lceil\lfloor n/3\rfloor/2\rceil$. *Let* $\overline{G}_t$ *be the set of 3-colorable graphs on n ≥ 5 vertices with t vertices in the smallest color class. We show that for a graph, G from* $\overline{G}_t$, $\mathrm{cycdisc}(G)\leq 2\times\lceil t/2\rceil$. *Furthermore a graph G' exists in* $\overline{G}_t$ *with large cycle discrepancy, such that* $\mathrm{cycdisc}(G')\geq 2\times\lceil t/2\rceil$ *for t ≥ 1. We also construct such d-colorable graphs for d> 3 that have maximum possible cycle discrepancy.*

Key Words: Cycle discrepancy, Graph coloring

1. Introduction

Cycle discrepancy is a graph invariant which was first introduced in [1]. The inspiration of this invariant is from discrepancy theory which is a well known area in Combinatorics. In discrepancy theory the objective is to study irregularities or deviation from the absolute uniformity. In discrepancy theory a major question is to divide a given set, *T*, such that each subset in *S* is divided as equally as possible, where *S* is a given collection of subsets of *T*. The given set, *T*, is called ground set. Each element of, *T*, is assigned a label either red or blue such that each subset in *S* is equally red and blue (as much as possible). For each subset from *S* we note the difference between number of red and blue elements. The maximum difference noted is the discrepancy of the labeling assigned to *T*. If we consider all possible labeling of *T*, a labeling which has least discrepancy is called an optimal labeling. Discrepancy of *T* is the discrepancy of an optimal labeling. We may call this quantity as the discrepancy of the provided set system (*T*, *S*). There are many excellent investigations in Combinatorics regarding this quantity [8, 3]. A complete introduction of this area can be found in [2, 5, 7].

Given a graph, *G* = (*V*, *E*), we take *V*, as the ground set and the set of all cycles in *G*, call it C_A, is taken as the collection of subsets, each cycle being a subset of *V*. The discrepancy of (*V*, C_A) is named as cycle discrepancy and is denoted by *cycdisc* (*G*). Normally the nature of discrepancy problems is geometric having discrepancy bounded by square root of the input size. Cycle discrepancy is a concept of graph theory and it does not have any obvious geometric nature. Graph theoretic tools and techniques are used in [1] to show that cycle discrepancy of a cubic graph having *n* vertices is bounded by (*n* + 2)/6 and further this bound is shown to be the best possible.

In this paper we are looking at cycle discrepancy of 3-colorable graphs and we find upper and lower bounds on cycle discrepancy for this class of graphs. The following section consist some definitions and fundamental results. In Section 3 cycle discrepancy is bounded for the class of 3-colorable graphs. In Section 4 we discuss construction of 3-colorable graphs having large cycle discrepancy. We show the existence of *d*-colorable graphs, for *d*> 3, which have maximum possible cycle discrepancy in Section 5. Finally in Section 6 we briefly conclude this work.

2. Definitions and Preliminary Results

The graphs we talk about in this paper are loop less, undirected and without multiple edges. Standard notation of graph theory is used which is mostly adopted from Bollobás Monograph [4]. We call a graph3-colorable if it admits a 3-coloring, that is there exists $f: V \rightarrow \{red, green, blue\}$, such that for any two vertices *u* and *v* which are connected, $f(u) \neq f(v)$. A tri-coloring divides the set of vertices into three classes having vertices of same color. We can also describe these color classes as, $A = f^{-1}(red), B = f^{-1}(green)$, and $D = f^{-1}(blue)$. This tri-coloring can be characterized by permuting the colors of the partitions of set of vertices (color classes). Therefore we normally say that $f = (A, B, D)$ is a tri-coloring of a given graph G. Throughout this paper we assume that $|D| \leq |B| \leq |A|$ and further $|D| = t$.

A labeling, χ of a graph is to map its vertex set to the set {+1, -1}. We will simply use '+' and '-' instead of +1 and -1. The subsets of *A*, *B* and *D* which contains the vertices that have label '+', defines the sets *A*+, *B*+, and *D*+ respectively. The sets *A*-, *B*-, and *D*- are also defined similarly. For any subset, *J*, of vertices of a given graph *G*, define,

$$\chi(J) = \sum_{u \in J} \chi(u).$$

A cycle in a graph can also be viewed as a subset of the vertex set of that graph. If we define a set containing all the cycles as C_A, then the cycle discrepancy of a labeling, χ, can be defined as:

$$\text{cycdisc}(\chi) = \max_{J \in C_A(G)} |\chi(J)|.$$

On the same lines for a graph, *G*, the cycle discrepancy, denoted as, *cycdisc*(*G*), can be defined as:

$$\text{cycdisc}(G) = \min_{\chi: V \rightarrow \{+,-\}} \text{cycdisc}(\chi).$$

It is worth mentioning the fact that cycle discrepancy of a graph is zero if and only if the graph is bipartite. The cycle discrepancy of a graph is always greater than or equal to the cycle discrepancy of any of its subgraphs.

Note that if we want to establish a lower bound on cycle discrepancy then we have to present a construction for a graph such that for every possible labeling this graph has a cycle *C* for which $\chi(C)$ is greater than or equal to the lower bound. Similarly if we want to establish an upper bound then we present a labeling χ such that for every cycle *J*, $\chi(J)$ is less than or equal to the desired bound.

To elaborate the concept of cycle discrepancy, let us compute the cycle discrepancy of some famous small graphs. Two famous small graphs are shown in the following figure.

Fig. 1 Triangle graph C_3(Left) and Diamond graph G_D(right)

Fact 1

The cycle discrepancy of triangle graph, C_3 is one.

Proof

Take any labeling of C_3, at least two vertices will have the same label. That is *cycdisc*(C_3) ≥ 1. Without loss of generality if we label '+' to any two vertices and label '-' to the third; this labeling shows that *cycdisc*(C_3) ≤ 1.

Fact 2

The cycle discrepancy of diamond graph, G_D is one.

Proof

As, C_3 is a subgraph of diamond graph, *cycdisc*(G_D) ≥ 1. If we label two vertices of G_D with '+' and other two with '-', this labeling allows us to show that *cycdisc*(G_D) ≤ 1.

2.1 Cycle Discrepancy of Petersen Graph

Petersen graph, G_P, is a very famous cubic 3-colorable graph on ten vertices. G_P is not Hamiltonian however following two propositions are important to find out cycle discrepancy of this graph.

Fact 3

If we delete any vertex of G_P, the remaining graph is Hamiltonian.

Proof

If we delete the vertex *a*, a Hamiltonian cycle on remaining vertices is obtained by visiting the remaining vertices in the following order: *e, j, h, f, i,* *g*, *b*, *c*, *d*, *e* (see Fig. 2. (a)). Since the Petersen graph is vertex transitive, we can similarly obtain a Hamiltonian cycle after deleting any other vertex.

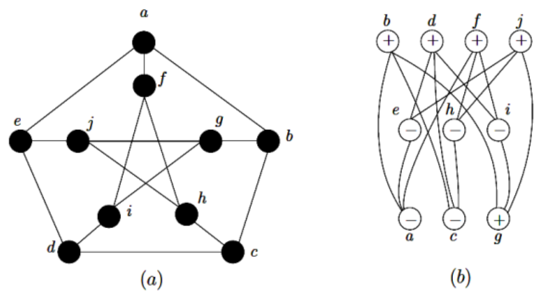

Fig.2 (a) Petersen graph (b) Tri-coloring and labeling of Petersen graph

Fact 4

If we delete any two connected vertices of G_P, the remaining graph is Hamiltonian.

Proof

If we delete the vertices {*a*, *b*}, a Hamiltonian cycle is obtained by visiting the remaining vertices in the following order: *e*, *j*, *g*, *i*, *f*, *h*, *c*, *d*, *e*. Since the Petersen graph is edge transitive, we can similarly obtain a Hamiltonian cycle after deleting any other edge (see Fig. 2. (a)).

Theorem 1

The cycle discrepancy of Petersen graph is 2. That is $cycdisc(G_P) = 2$.

Proof

Take any arbitrary labeling of G_P. Define the set of vertices labeled '+' as *VP+* and similarly set of vertices labeled '-' as *VP-*. If |*VP+* | = | *VP-*|, there will be an edge connecting two vertices having same label because G_P is not a bipartite graph. There will be a cycle without these two vertices due to Fact 4, and the discrepancy of this cycle is two. If all the vertices of G_P are of same label then there will be a cycle avoiding any vertex due to the Fact 3 with discrepancy nine. Otherwise, without loss of generality assume that |*VP+* | > | *VP-*| > 0. There will be a cycle avoiding a vertex having label '-' due to the Fact 3. This cycle will have discrepancy at least two.

As G_P is cubic, it is known from [1] that: $cycdisc(G_P) \le (10+2)/6 = 2$. We have also provided a labeling of G_P in Fig. 2. (b).

3. Cycle Discrepancy of 3-Colorable Graphs

For the purpose of bounding cycle discrepancy of 3-colorable graphs we will first define a labeling then we will show that this labeling is good enough to achieve the desired upper bound.

Define a labeling, χ, such that all the vertices are labeled '+' that belong to *A* and all the vertices are labeled '-'that belong to *B*. In *D*, label half of the vertices '+' and label the remaining half of the vertices '-'. If |*D*| = *t* is odd, either of the label is used an extra time but if *D* has only one vertex we label it '-'.

Lemma 1

Take a path *P* containing exactly two vertices from *D*, namely d_1 and d_2. The vertex $d_1$is the first and $d_2$is the last vertex on the path *P*. Then, $\chi(P) = \chi(d_1) + \chi(d_2) + h$.Where h ∈ {-1, 0, 1}.

Proof

Take a path, *P*, containing exactly two vertices from *D*, namely d_1 and d_2 such that d_1 is at the start of the path and d_2 is at the end of the path. *P* can be classified into three possible cases. If a path is not in one of these three cases, its reverse will be in one of these cases.

Case: 1.P = $d_1, b_1, a_1, \ldots, b_k, a_k, b_{k+1}, d_2$. Where $b_1, \ldots, b_{k+1} \in B$, $a_1, \ldots, a_k \in A$ and $k \ge 0$.

For this case, $\chi(P) = \chi(d_1) + \chi(d_2) - 1$ .Note that all other labels cancel out.

Case: 2.P = $d_1, b_1, a_1, \ldots, b_k, a_k, d_2$. Where $b_1, \ldots, b_k \in B$, $a_1, \ldots, a_k \in A$ and $k \ge 1$.

In this case, $\chi(P) = \chi(d_1) + \chi(d_2)$, as all other labels cancel out.

Case: 3.P = $d_1, a_1, b_1, \ldots, a_k, b_k, a_{k+1}, d_2$. Where $b_1, \ldots, b_k \in B$, $a_1, \ldots, a_{k+1} \in A$ and $k \ge 0$.

In this case, $\chi(P) = \chi(d_1) + \chi(d_2) + 1$, as all other labels cancel out.

In the next theorem we establish an upper bound for cycle discrepancy of the labeling χ.

Theorem 2

Let G be a 3-colorable graph. Take an arbitrary cycle, C,of$G$,

$$|\chi(C)| \le \begin{cases} t, & if \quad 3 \le n \le 4 \\ 2 \times \lceil t/2 \rceil, & if \quad n \ge 5. \end{cases}$$

If there is no vertex from D appearing on cycle C, then this cycle is alternating between A and B. Thus $\chi(C) = 0$.

If a cycle contains exactly one vertex from D, then the proof of Lemma 1 can be applied almost verbatim. Take $d_1 = d_2$ and in Case 1 and 3, $k \ge 1$.

If $n = 3$, only Case 2 can occur and for only possible cycle, C_1, we have $|\chi(C_1)| = t = 1$. For $n = 4$, Case 2 or Case 3 can occur and for any possible cycle C, $|\chi(C)| \le t = 1$. For $n = 5$, $|\chi(C)| \le 2$. Same holds for any other value of n with only one vertex of C in D.

Now, let us assume that $|C \cap D| = g > 1$. Let $d_0,\ldots,d_{g-1}$are the vertices on the cycle C and these vertices belong to D. Further we assume that$P_i$is the path starting from $d_i$and ending at d_{i+1}. The indices over here are modulo g. We have

$$\chi(C) = (\sum_{i=0}^{g-1} \chi(P_i)) - (\sum_{i=0}^{g-1} \chi(d_i))$$

$$= (\sum_{i=0}^{g-1} (\chi(d_i) + \chi(d_{i+1}) + h_i)) - (\sum_{i=0}^{g-1} \chi(d_i))$$

$$= \sum_{i=0}^{g-1} \chi(d_i) + \sum_{i=0}^{g-1} h_i \ .$$

For first equality observe that every d_i is exactly appearing on two paths, that is, P_i and P_{i-1}.While any other vertex of C is appearing exactly on one path. Apply Lemma 1 to each P_i to get the second equality. As indices are modulo g so we can write the last equality. At this point there are two cases:

Case: $g \le t/2$. In this case one can easily observe that $-t/2 \le \sum_{i=0}^{g-1} \chi(d_i) \le t/2$ *and* $-g \le \sum_{i=0}^{g-1} h_i \le g$.

Hence, $-t \le \chi(C) \le t$.

Case: $g > t/2$. In this case, note that at most $\lceil t/2 \rceil$ vertices are of the same label in D. Let us define $k = g - \lceil t/2 \rceil$. Then,

$$-\lceil t/2 \rceil + k \le \sum_{i=0}^{g-1} \chi(d_i) \le \lceil t/2 \rceil - k$$

and

$$-\lceil t/2 \rceil - k \le \sum_{i=0}^{g-1} h_i \le \lceil t/2 \rceil + k \ .$$

By combining both inequalities,

$$-2 \times \lceil t/2 \rceil \le \chi(C) \le 2 \times \lceil t/2 \rceil.$$

Hence, $|\chi(C)| \le 2 \times \lceil t/2 \rceil$.

As $t \le \lfloor n/3 \rfloor$, we can conclude the following corollary.

Corollary

Let G be a 3-colorable graph on at least three vertices,

$$cycdisc(G) \le \begin{cases} 1, & if \quad 3 \le n \le 4 \\ 2 \times \lceil \lfloor n/3 \rfloor / 2 \rceil, & if \quad n \ge 5. \end{cases}$$

4. 3-Colour Graphs with High Cycle Discrepancy

In this section we will construct graphs with high cycle discrepancy to establish a lower bound on cycle discrepancy of 3-colorable graphs.

For any 3-colorable graph, G, on $n \ge 3$ vertices let A, B, and D be three color classes of G, such that $|A| \ge |B| \ge |D| \ge 1$ and $|D| = t$. Let $\bar{G}_t$ is the set of 3-colorable graphs with smallest color class of size t. Note that $n \ge 3t$, that is why $\bar{G}_t$ can have graphs with different value of n but same value of t.

Lemma 2

Let G in $\bar{\mathcal{G}}_t$, be a graph on $n \geq 5$ vertices for $t \in \{1,2\}$ such that $|A|$-$|B| \leq 1$ and every vertex in G is connected to all the vertices which do not belong to the same color class. If we take any arbitrary labeling of G, there will be an induced path P on three vertices, all having the same label.

Proof

We first assume that A has two vertices x_1 and x_2 having same label. Without loss of generality take this label as '+'. If there is a vertex v with '+' label in B U D then x_1, v, x_2 will make the path P. If B U D does not have any vertex with '+' label then any two vertices from B and a vertex from D will make the path P.

If A does not have two vertices of the same label then A has only two vertices. In this case B will also have two vertices and if both of these are of the same label then they both are connected to a vertex with similar label in A. Otherwise B has two vertices which are not of the same label. Now D has at least one vertex which is connected to a vertex of the same label in A and a vertex of the same label in B.

Theorem 3

There is a 3-colorable graph, G in $\bar{\mathcal{G}}_t$, having n vertices and $t \geq 1$, such that

$$\mathrm{cycdisc}(G) \geq \begin{cases} t, & \text{if } \quad 3 \leq n \leq 4 \\ 2 \times \lceil t/2 \rceil, & \text{if } \quad n \geq 5. \end{cases}$$

Proof

We can make a 3-colorable graph G belonging to $\bar{\mathcal{G}}_t$ on $n \geq 3$ vertices such that $1 \leq t \leq n/3$. Put t vertices in D and distribute n-t vertices in both A and B as equally as possible. For any vertex v connect it to all the vertices which do not have the same color as v. Take any arbitrary labeling of G.

If, $3 \leq n \leq 4$ then G has C_3 as a subgraph that has discrepancy equal to one according to Fact 1. Now we may assume $n \geq 5$.

Case: $t < 3$. By Lemma 2 there will be an induced path, P, in G, on three vertices, say, v_1, v_2, and v_3, all having the same label. If v_1 and v_3 both belong to different color classes then they will be connected making a triangle having discrepancy three. If v_1 and v_3 both belong to same color class then both would be connected to a vertex in some other color class with possibly having different label making a cycle on four vertices having discrepancy at least two.

Case: $t \geq 3$. In D at least $\lceil t/2 \rceil$ vertices will be of same label. Without loss of generality consider this label as '+'. If there are $\lceil t/2 \rceil$ vertices or more in A U B with label '+' then there is a cycle with $2 \times \lceil t/2 \rceil$ vertices all with the same label. Otherwise, if the vertices with '+' label in A U B are less than $\lceil t/2 \rceil$, then the vertices with '-' label in B are at least $\lceil t/2 \rceil$ and same is the case with A. Hence there is a cycle with $2 \times \lceil t/2 \rceil$ vertices having the same label.

As the largest value of t can be $\lfloor n/3 \rfloor$, we can conclude the following corollary.

Corollary

There is a 3-colorable graph, G on at least three vertices, such that

$$\mathrm{cycdisc}(G) \geq \begin{cases} 1, & \text{if } \quad 3 \leq n \leq 4 \\ 2 \times \lceil \lfloor n/3 \rfloor /2 \rceil, & \text{if } \quad n \geq 5. \end{cases}$$

5. Cycle Discrepancy of d-Colorable Graphs

This section talks about cycle discrepancy of d-colorable graphs for $d > 3$. We start by noting the following fact about maximum cycle discrepancy.

Fact 5

A graph G on n vertices can have cycle discrepancy at most $\lceil n/2 \rceil$.

Proof

Let G be a graph having n vertices. We label any $\lceil n/2 \rceil$ vertices '+' and remaining $\lfloor n/2 \rfloor$ vertices '-'. Now in the worst case there may exist a cycle which only consist of all the '+' labeled vertices.

Theorem 4

There exists a d-colorable graph Gon n vertices for $d > 3$ with $\mathrm{cycdisc}(G) \geq \lceil n/2 \rceil$.

Proof

We start by constructing a d-colorable graph $G = (V, E)$ on $n = dk$ vertices with $k \geq 2$ and $d > 3$. Put k vertices in all d color classes and put an edge between every two vertices which do not belong to the same color class. Now assume an arbitrary labeling of the vertices. Without loss of generality we may assume that '+' labeled vertices are more than '-' labeled vertices. Let $G^+ = (V^+, E^+)$ is the subgraph containing only '+' labeled vertices with $|V^+| = n' \geq \lceil n/2 \rceil \geq n/2$.

Note that a color class can have at most $n/d \leq 2n'/d$ vertices which are labeled '+'. Hence for a vertex v in V^+, $d_{G^+}(v) \geq n'(d-2)/d$. As $d > 3$, we can write $d_{G^+}(v) \geq n'/2$. Hence, by applying Dirac's Theorem [6] G^+ is a Hamiltonian graph and the theorem follows.

6. Conclusion

In this paper, we show that the size of smallest color class, which can be at most $\lfloor n/3 \rfloor$, bounds the cycle discrepancy of a 3-colorable graph. We establish that the stated bound is best possible by showing the existence of such graphs which avail this bound. Knowing that a graph is 3-colorable gives us an upper bound on cycle discrepancy but in the case when the graph is d-colorable for $d > 3$, the cycle discrepancy can be maximum. That is one can construct a d-colorable graph($d > 3$) which has cycle discrepancy at least $\lceil n/2 \rceil$. At the end, we present some open problems in this area.

Problem 1:

Can we exactly compute the cycle discrepancy of a given 3-colorable graph?

Problem 2:

Can we characterize graphs with cycle discrepancy $k > 0$?

7. References

[1] S. Abbasi, L. Aslam, 2011, "The cycle discrepancy of three-regular graphs",*Graphs and Combinatorics*, Vol. 27, pp. 27–46.

[2] J. Beck, W. L. Chen, 1987, "Irregularities of distribution," Cambridge University press, New York.

[3] J. Beck, W. Fiala, 1981, "Integer making theorems", *Discrete Applied Mathematics*, Vol. 3, pp. 1-8.

[4] B. Bollobás, 1988, "Modern graph theory", Springer-Verlag.

[5] B. Chazelle, 2000, "The discrepancy method: randomness and complexity", Cambridge University press, New York.

[6] G. A. Dirac, 1952, "Some theorems on abstract graphs", *proc. London math. soc.*, Vol. s3-2, pp. 69-81.

[7] J. Matousek, 1999, "Geometric discrepancy: an illustrated guide. Algorithms and combinatorics", Springer, Berlin.

[8] J. Spencer, 1985, "Six standard deviations suffice", *Transactions of the American mathematical society*, Vol. 289, pp. 679-706.