On vertex-disjoint paths in regular graphs

Jie Han*

Instituto de Matemática e Estatística
Universidade de São Paulo
Rua do Matão 1010, 05508-090, São Paulo, Brazil

jhan@ime.usp.br

Submitted: Jun 20, 2017; Accepted: Apr 9, 2018; Published: Apr 27, 2018

Abstract

Let $c \in (0, 1]$ be a real number and let n be a sufficiently large integer. We prove that every n-vertex cn-regular graph G contains a collection of $\lfloor 1/c \rfloor$ paths whose union covers all but at most $o(n)$ vertices of G. The constant $\lfloor 1/c \rfloor$ is best possible when $1/c \notin \mathbb{N}$ and off by 1 otherwise. Moreover, if in addition G is bipartite, then the number of paths can be reduced to $\lfloor 1/(2c) \rfloor$, which is best possible.

Mathematics Subject Classifications: 05C38

1 Introduction

Paths and cycles are fundamental objects in graph theory. The *path cover number* is the minimum number of vertex-disjoint paths whose union covers all vertices of G. Note that we allow paths of length 0 (single vertices) in the definition above. Trivially the path cover number of a graph G is upper bounded by the independence number of G, because the set of the (arbitrary one out of the two) end vertices of the paths in a minimal path cover form an independent set in G. It is evident that for general graphs, determining the path cover number is NP-hard, because deciding if the path cover number equals 1 is equivalent to the decision problem for a Hamiltonian path, which is NP-complete. For bounds on the path cover number for general graphs, see e.g. [3, 6]. For regular graphs, Magnant and Martin [5] made the following conjecture and confirmed it for $0 \leqslant k \leqslant 5$.

Conjecture 1. If G be a k-regular graph of order n, then the path cover number of G is at most $n/(k+1)$.

*Supported by FAPESP (2013/03447-6, 2014/18641-5).

If true, the bound in Conjecture 1 would be tight as seen by disjoint copies of complete graph K_{k+1} (if $n \equiv j$ modulo $k+1$ and $j \neq 0$, then we change one copy of K_{k+1} to a k-regular subgraph of K_{k+1+j}). By the celebrated Dirac theorem on Hamiltonian paths [2], Conjecture 1 is true for $k \geqslant (n-1)/2$. As far as we know, Conjecture 1 is open for all other cases. To provide more evidence on the validity of the conjecture, in this note we prove the following result for dense regular graphs.

Theorem 2. *For any $c, \alpha > 0$, there exists $n_0 \in \mathbb{N}$ such that the following holds for every integer $n \geqslant n_0$.*

1. *Every $\lceil cn \rceil$-regular graph of order n contains a collection of at most $\lfloor 1/c \rfloor$ vertex-disjoint paths whose union covers all but αn vertices.*
2. *Every bipartite $\lceil cn \rceil$-regular graph of order n contains a collection of at most $\lfloor 1/(2c) \rfloor$ vertex-disjoint paths whose union covers all but αn vertices.*

Note that Part (2) of the theorem corresponds to the bipartite version of Conjecture 1: if G is a bipartite k-regular graph of order n, then the path cover number of G can be as large as $n/(2k)$, as seen by the vertex-disjoint copies of $K_{k,k}$. Note that

$$\frac{-2}{c(\lceil cn \rceil + 1)} \leqslant \frac{n}{\lceil cn \rceil + 1} - \frac{1}{c} = \frac{cn - \lceil cn \rceil - 1}{c(\lceil cn \rceil + 1)} < 0.$$

So when n is large, if $1/c \notin \mathbb{N}$, then $\lfloor \frac{n}{\lceil cn \rceil + 1} \rfloor = \lfloor 1/c \rfloor$, i.e., the number of paths in Theorem 2 matches the quantity in Conjecture 1; however if $1/c \in \mathbb{N}$, then the quantity $\lfloor 1/c \rfloor$ is off by 1. On the other hand, the quantity $\lfloor 1/(2c) \rfloor$ in Part (2) is optimal.

At last, we remark that the bound in Conjecture 1 is not tight if we restrict the problem on connected regular graphs, see [7] for connected cubic graphs.

2 A weaker result

We first prove the following weaker result. For reals a, b, c, we write $a = (1 \pm b)c$ if there exists a real $x \in (1-b, 1+b)$ such that $a = xc$.

Theorem 3. *Given any reals $c, \alpha > 0$, there exists $\epsilon > 0$ and integer $C > 0$ such that the following holds for sufficiently large integer n. Let G be a graph of order n such that $\deg(v) = (1 \pm \epsilon)cn$ for every $v \in V(G)$. Then there exists a collection of C vertex-disjoint cycles in G whose union covers all but αn vertices of G.*

Our main tools for embedding the cycles are the Regularity Lemma of Szemerédi [11] and the Blow-up Lemma of Komlós et al. [4]. For any two disjoint vertex-sets A and B of a graph G, the density of A and B is defined as $d(A, B) := e(A, B)/(|A||B|)$, where $e(A, B)$ is the number of edges with one end vertex in A and the other in B. Let ϵ and δ be two positive real numbers. The pair (A, B) is called *ϵ-regular* if for every $X \subseteq A$ and $Y \subseteq B$ satisfying $|X| > \epsilon|A|$, $|Y| > \epsilon|B|$, we have $|d(X, Y) - d(A, B)| < \epsilon$. Moreover, the pair (A, B) is called *(ϵ, δ)-super-regular* if (A, B) is ϵ-regular and $\deg_B(a) > \delta|B|$ for all $a \in A$ and $\deg_A(b) > \delta|A|$ for all $b \in B$.

Lemma 4 (Regularity Lemma – Degree Form)**.** *For every $\epsilon > 0$ there is an $M = M(\epsilon)$ such that for any graph $G = (V, E)$ and any real number $d \in [0, 1]$, there is a partition of the vertex set V into $t+1$ clusters $V_0, V_1, \dots, V_t$, and there is a subgraph G' of G with the following properties:*

- $t \leqslant M$,
- $|V_i| \leqslant \epsilon |V|$ *for* $0 \leqslant i \leqslant t$ *and* $|V_1| = |V_2| = \cdots = |V_t|$,
- $\deg_{G'}(v) > \deg_G(v) - (d+\epsilon)|V|$ *for all* $v \in V$,
- $G'[V_i] = \varnothing$ *for all* i,
- *each* (V_i, V_j), $1 \leqslant i < j \leqslant t$, *is* ϵ*-regular with* $d(V_i, V_j) = 0$ *or* $d(V_i, V_j) \geqslant d$ *in* G'.

The Blow-up Lemma allows us to regard a super regular pair as a complete bipartite graph when embedding a graph with bounded degree. Since we will always use it to embed a cycle, we state it in the following special form.

Lemma 5. *For every $\delta > 0$, there exists an $\epsilon > 0$ such that the following holds for sufficiently large integer N. Let (X, Y) be an (ϵ, δ)-super-regular pair with $|X| = |Y| = N$. Then (X, Y) contains a spanning cycle (a cycle of length $2N$).*

A *fractional matching* is a function f that assigns to each edge of a graph a real number in $[0, 1]$ so that, for each vertex v, we have $\sum f(e) \leqslant 1$ where the sum is taken over all edges incident to v. The *fractional matching number* $\mu_f(G)$ of a graph G is the supremum of $\sum_{e \in E(G)} f(e)$ over all fractional matchings f. We use the following so-called 'fractional Berge-Tutte formula' of Scheinerman and Ullman [10, Theorem 2.2.6]. Note that it is also proved in [10] that (see Theorem 2.1.5) in a graph G, the maximum fractional matching, i.e., with weight $\mu_f(G)$, can be achieved with weights only chosen from $\{0, 1/2, 1\}$.

Theorem 6 ([10])**.** *For any graph G,*

$$\mu_f(G) = \frac{1}{2}\left(|V(G)| - \max\{i(V(G) \setminus S) - |S|\}\right),$$

where $i(X)$ denotes the number of isolated vertices in $G[X]$, and the maximum is taken over all $S \subseteq V(G)$.

Proof of Theorem 3. Given $c, \alpha > 0$, let $d = \alpha c/9$. We apply Lemma 5 with $\delta = d/2$ and obtain $\epsilon_1 > 0$. Let $\epsilon = \min\{\epsilon_1, d/6, 3d/(2c)\}$. We then apply Lemma 4 with ϵ and obtain $M = M(\epsilon)$. Let $n \in \mathbb{N}$ be sufficiently large. Let $G = (V, E)$ be a graph of order n such that $\deg(v) = (1 \pm \epsilon)cn$ for every $v \in V$. We apply the Regularity Lemma (Lemma 4) on G with the constants ϵ, d chosen as above and obtain a partition of V into $V_0, V_1, \dots, V_t$ for some $t \leqslant M$, and a subgraph G' of G with the properties as described in Lemma 4. By moving at most one vertex from each V_i, $i \in [t]$ to V_0, we can assume that $m := |V_i|$ is even. Thus we have $|V_0| \leqslant \epsilon n + t \leqslant 2\epsilon n$. Now for any $v \in V$,

$$\deg_{G'}(v) - |V_0| > \deg_G(v) - (d+\epsilon)n - 2\epsilon n \geqslant (c - 2d)n.$$

Let $\beta = 3d/c$. Let H be the graph on $[t]$ such that $ij \in E(H)$ if and only if $d(V_i, V_j) \geqslant d$. We first assume that there exists a set $S \subseteq [t]$, such that $i(V(H) \setminus S) - |S| \geqslant \beta t$. In particular, let T be the collection of $|S| + \beta t$ isolated vertices in $H - S$. Thus we have

$$e_G(T, S) \geqslant |T|m(\deg_{G'}(v) - |V_0|) \geqslant |T|m(c - 2d)n.$$

However, by averaging, this implies that there exists a vertex $v \in V$ such that

$$\deg_G(v) \geqslant \deg_{G'}(v) \geqslant \frac{|T|m(c-2d)n}{|S|m} \geqslant \frac{t}{t-\beta t}(c-2d)n > (1+\epsilon)cn,$$

by the definition of β and ϵ, a contradiction. Thus we have $i(V(H) \setminus S) - |S| \leqslant \beta t$ for any $S \subseteq [t]$. So by Theorem 6, we get $\mu_f(H) \geqslant (1-\beta)t/2$. Moreover, there exists a fractional matching f such that $\sum_{e \in E(H)} f(e) = \mu_f(H) \geqslant (1-\beta)t/2$ and $f(e) \in \{0, 1/2, 1\}$ for every edge $e \in E(H)$.

For each $i \in [t]$ we arbitrarily split V_i into V_i^1 and V_i^2 each of size $m/2$. Thus the existence of f implies that we can partition $V \setminus V_0$ into at least $(1-\beta)t/2$ pairs of sets each of form (V_i^a, V_j^b) with density at least d, where $i, j \in [t]$, $i \neq j$ and $a, b \in [2]$, and a set of at most $2\beta t \cdot m$ vertices. Note that here (to simplify the argument) even if an edge $ij \in E(H)$ receives weight 1, we still split it, e.g., as (V_i^1, V_j^1) and (V_i^2, V_j^2). Thus every vertex of H is in at most two pairs so there are at most t pairs.

We will show that each such pair contains a cycle that covers all but at most $2\epsilon m$ vertices. Indeed, fix any pair (V_i^a, V_j^b), let A be the set of vertices in V_i^a whose degree to V_j^b is less than $(d-\epsilon)|V_j^b|$. Since $d(A, V_j^b) < d - \epsilon$ and $|V_j^b| > \epsilon m$, the regularity of (V_i, V_j) implies that $|A| \leqslant \epsilon m$. Similarly let B be the set of vertices in V_j^b whose degree to V_i^a is less than $(d-\epsilon)|V_i^a|$ and we have $|B| \leqslant \epsilon m$. Let $A' \supseteq A$ and $B' \supseteq B$ be arbitrary subsets of V_i^a and V_j^b, respectively, of size exactly ϵm. Now let $X = V_i^a \setminus A'$ and $Y = V_j^b \setminus B'$, we get that (X, Y) is $(\epsilon, d - 3\epsilon)$-super-regular with density at least $d - 3\epsilon$, and $|X| = |Y| = m - \epsilon m$. Since $d - 3\epsilon \geqslant d/2$, by Lemma 5, (X, Y) contains a spanning cycle and we are done.

Let $C = M$. Thus we obtain a set of at most $t \leqslant M = C$ vertex-disjoint cycles in G that covers all but at most

$$t \cdot 2\epsilon m + |V_0| + 2\beta t \cdot m \leqslant 3\beta n = \alpha n$$

vertices, completing the proof. □

3 Proof of Theorem 2

In the proof of Theorem 2 we use the trick of a 'reservoir lemma' from [8, 9]. Roughly speaking, we will reserve a random set R of vertices at the beginning of the proof, and use them to connect the paths returned by applying Theorem 3 on $G - R$. We first recall the following Chernoff's bounds (see, e.g., [1]) for binomial random variables and for $x > 0$:

$$\mathbb{P}[Bin(n', \zeta) \geqslant n'\zeta + x] < e^{-x^2/(2n'\zeta + x/3)}$$
$$\mathbb{P}[Bin(n', \zeta) \leqslant n'\zeta - x] < e^{-x^2/(2n'\zeta)}.$$

Lemma 7. *Given any $c, \gamma, \epsilon > 0$, the following holds for sufficiently large integer n. Let G be a $\lceil cn \rceil$-regular graph of order n. Then there exists a set $R \subseteq V(G)$ such that $|R| = (1 \pm \epsilon)\gamma n$ and every vertex of G has degree $(1 \pm \epsilon)c\gamma n$ in R.*

Proof. We select the set R by including each vertex of G independently and randomly with probability γ. Note that $|R|$ and $\deg(v, R)$ for each $v \in V(G)$ are both binomial random variables with expectation γn and $\gamma \lceil cn \rceil$, respectively. By Chernoff's bounds, we get

$$\mathbb{P}[|R| > (1+\epsilon)\gamma n] < e^{-\epsilon^2 \gamma n/3}, \ \mathbb{P}[|R| < (1-\epsilon)\gamma n] < e^{-\epsilon^2 \gamma n/3},$$
$$\mathbb{P}[x_v > (1+\epsilon)c\gamma n] < e^{-\epsilon^2 c\gamma n/3}, \ \mathbb{P}[x_v < (1-\epsilon)c\gamma n] < e^{-\epsilon^2 c\gamma n/3},$$

where $x_v := \deg(v, R)$ for all $v \in V(G)$. Since $(2n+2)e^{-\epsilon^2 c\gamma n/3} < 1$ because n is large enough, there is a choice of R with the desired properties. □

Proof of Theorem 2. Given $c, \alpha \in (0, 1)$, let $\gamma = \alpha/4$. We apply Theorem 3 with c and $\alpha/2$ in place of α and obtain ϵ_1 and $C \in \mathbb{N}$. Let $\epsilon = \min\{\epsilon_1, ((\lfloor 1/c \rfloor + 1)c - 1)/3\}$. Let G be a $\lceil cn \rceil$-regular graph of order n. We first pick the set R by Lemma 7. Let $G_1 = G - R$ and $n_1 = n - |R|$. Thus for every vertex $v \in V(G_1)$, we know that $\deg_{G_1}(v) = \lceil cn \rceil - (1 \pm \epsilon)c\gamma n$. Since $|R| = (1 \pm \epsilon)\gamma n$ we know that $\deg_{G_1}(v) = (1 \pm \epsilon)cn_1$. Indeed, for the upper bound we have

$$\deg_{G_1}(v) \leqslant c(n_1 + |R|) + 1 - (1-\epsilon)c\gamma n \leqslant cn_1 + 1 + 2\epsilon c\gamma n \leqslant (1+\epsilon)cn_1,$$

where in the last inequality we use $n < 2n_1$ and $\gamma \leqslant 1/4$; and the lower bound can be shown similarly. By applying Theorem 3 with $\alpha/2$ in place of α, we obtain a collection of at most C vertex-disjoint paths whose union covers all but at most $\alpha|V(G_1)|/2 \leqslant \alpha n/2$ vertices of G_1.

Next we iteratively use the property of R to connect some pair of paths. We first explain the general case. Suppose there are at least $\lfloor 1/c \rfloor + 1$ paths left. Indeed, let $v_1, v_2, \ldots, v_{\lfloor 1/c \rfloor + 1}$ be the (arbitrary one out of the two) ends of the $\lfloor 1/c \rfloor + 1$ paths. Note that throughout the iteration there are at most C vertices in R that have been used for connecting and thus removed from R. So for each i, by $\deg(v_i, R) - C \geqslant (1-\epsilon)c\gamma n - C \geqslant (1-2\epsilon)c|R|$, we have

$$(\lfloor 1/c \rfloor + 1)(\deg(v_i, R) - C) \geqslant (1+3\epsilon)(1-2\epsilon)|R| > |R|,$$

by the definition of ϵ. Thus there exist two vertices v_i, v_j which have a common neighbor w in R so that we can connect the corresponding two paths by w. At the end, we obtain a collection of at most $\lfloor 1/c \rfloor$ vertex-disjoint paths whose union covers all but at most $\alpha n/2 + (1+\epsilon)\gamma n \leqslant \alpha n$ vertices in G.

Second, assume that there are at least $\lfloor 1/(2c) \rfloor + 1$ paths left and in addition that G is bipartite with bipartition X and Y. Note that since G is regular we have $|X| = |Y| = n/2$ (so in particular n must be even). Fix $\lfloor 1/(2c) \rfloor + 1$ paths. By throwing away at most one vertex from each path we can assume that each path has exactly one end vertex in X and one in Y. Let $v_1, v_2, \ldots, v_{\lfloor 1/(2c) \rfloor+1}$ be the end vertices in X. By the similar calculation, we can find a vertex $w \in R \cap Y$ which connects some pair of paths. At the end, we obtain a collection of at most $\lfloor 1/(2c) \rfloor$ vertex-disjoint paths whose union covers all but at most $\alpha n/2 + (1+\epsilon)\gamma n + C^2 \leqslant \alpha n$ vertices in G, because the iteration has at most C steps and in each step we threw away at most C vertices from the current paths. □

Acknowledgement

The author would like to thank Yoshiharu Kohayahawa, Phablo Moura and Yoshiko Wakabayashi for discussions and valuable comments on the manuscript. He would also like to thank an anonymous referee for the helpful comments.

References

[1] N. Alon and J. Spencer. *The probabilistic method.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[2] G. A. Dirac. Some theorems on abstract graphs. *Proc. London Math. Soc. (3)*, 2:69–81, 1952.

[3] I. Hartman. Variations on the Gallai-Milgram theorem. *Discrete Math.*, 71(2):95–105, 1988.

[4] J. Komlós, G. Sárközy, and E. Szemerédi. Blow-up lemma. *Combinatorica*, 17(1):109–123, 1997.

[5] C. Magnant and D. Martin. A note on the path cover number of regular graphs. *Australas. J. Combin.*, 43:211–217, 2009.

[6] O. Ore. Arc coverings of graphs. *Ann. Mat. Pura Appl. (4)*, 55:315–321, 1961.

[7] B. Reed. Paths, stars and the number three. *Combin. Probab. Comput.*, 5(3):277–295, 1996.

[8] V. Rödl, A. Ruciński, and E. Szemerédi. A Dirac-type theorem for 3-uniform hypergraphs. *Combin. Probab. Comput.*, 15(1-2):229–251, 2006.

[9] V. Rödl, A. Ruciński, and E. Szemerédi. An approximate Dirac-type theorem for k-uniform hypergraphs. *Combinatorica*, 28(2):229–260, 2008.

[10] E. Scheinerman and D. Ullman. *Fractional graph theory.* Dover Publications, Inc., Mineola, NY, 2011. A rational approach to the theory of graphs, With a foreword by Claude Berge, Reprint of the 1997 original.

[11] E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.