Short cycles in random regular graphs

Brendan D. McKay*

Department of Computer Science
Australian National University
Canberra ACT 0200, Australia

bdm@cs.anu.ed.au

Nicholas C. Wormald† and Beata Wysocka

Department of Mathematics and Statistics
University of Melbourne
Vic 3010, Australia

nwormald@uwaterloo.ca
beata@ms.unimelb.edu.au

Submitted: Aug 10, 2003; Accepted: May 20, 2004; Published: Sep 20, 2004
Mathematics Subject Classifications: 05C80, 05C38, 05C30

Abstract

Consider random regular graphs of order n and degree $d = d(n) \geq 3$. Let $g = g(n) \geq 3$ satisfy $(d-1)^{2g-1} = o(n)$. Then the number of cycles of lengths up to g have a distribution similar to that of independent Poisson variables. In particular, we find the asymptotic probability that there are no cycles with sizes in a given set, including the probability that the girth is greater than g. A corresponding result is given for random regular bipartite graphs.

1 Introduction

Let H be a fixed graph. The asymptotic distribution of the number of subgraphs of a random graph isomorphic to H has been studied in various places such as by Ruciński [9] for the random graph model $\mathcal{G}(n,p)$ and Janson [4] for the model $\mathcal{G}(n,m)$. In this paper we consider the distribution in a random d-regular graph. (Here, and henceforth in the paper, "random" refers to the uniform distribution on the set of all labelled graphs in the specified class.)

*Research supported by the Australian Research Council

†Research supported by the Australian Research Council. Current address: Department of Combinatorics and Optimization, University of Waterloo, Waterloo ON Canada N2L 3G1.

Many properties of random d-regular graphs on n vertices are known (see Bollobás [2] or Wormald [11] for details). For d fixed or growing slowly as a function of n, such a graph looks like a tree in the neighbourhood of almost all vertices; the expected number of cycles of any fixed length is small, and for any fixed graph H with more edges than vertices, the expected number of subgraphs isomorphic to H tends to 0 as $n \to \infty$. Thus, for subgraph enumeration questions, the most "interesting" subgraphs are the cycles. For this reason we consider only cycles in this paper. The girth of the graph is an interesting property which can be determined if enough is known about cycles. Our results apply for d, and the girth, both growing as functions of n, up to the point that small biconnected subgraphs with more than one cycle begin to proliferate.

Define $X_r = X_r(n)$ to be the number of cycles of length r in a random d-regular graph of order n. In [1], it was shown that the variables X_r for $3 \le r \le g$ are asymptotically distributed as independent Poisson variables with means $(d-1)^r/2r$, provided $d \le \sqrt{2 \log n} - 1$ with g fixed (and independently in [10] for fixed d). In this paper we allow $d = d(n)$ and $g = g(n)$ to increase with n, provided only that

$$(d-1)^{2g-1} = o(n). \tag{1.1}$$

We will show that, in a certain sense, the asymptotic behaviour as independent Poisson variables remains. In particular, our result implies the asymptotic probability that the girth is greater than g. (Note that this result reaches to approximately one quarter of the theoretical upper bound $(2 + o(1)) \log_{d-1} n$ on the girth of a d-regular graph.)

Assumption (1.1) can be motivated as follows. If (1.1) is satisfied, then only a vanishing fraction of d-regular graphs have two cycles of length at most g which share an edge, and the converse is also true. This makes (1.1) a natural boundary for our method, as will become apparent. We suspect, but did not prove, that it is also a boundary for our results in the sense that our main theorems are not true if (1.1) is violated.

The asymptotic distribution of the number of cycles of greater length was determined by Garmo [3], though not to the same accuracy, and not in a form that implies results about the girth.

Let $C = \{c_1, c_2, \ldots, c_t\}$ be a nonempty subset of $\{3, 4, \ldots, g\}$. For a random regular graph G of order n and degree d, define $M_C(G) = (m_1, m_2, \ldots, m_t)$, where m_i is the number of cycles of length c_i in G for $1 \le i \le t$. For $3 \le i \le g$, define

$$\mu_i = \frac{(d-1)^{c_i}}{2c_i}. \tag{1.2}$$

Our main results are the following two theorems. The first gives the asymptotic distribution, while the second gives the probability at 0.

Theorem 1 *Let S be a set of nonnegative integer t-tuples. Then, as $n \to \infty$, the probability that $M_C(G) \in S$ is*

$$\bigl(1 + o(1)\bigr)\Biggl(\sum_{(m_1, m_2, \ldots, m_t) \in S} \; \prod_{i=1}^{t} \frac{e^{-\mu_i} \mu_i^{m_i}}{m_i!} \Biggr) + o(1).$$

Note that, apart from the error terms, this is what holds for t independent Poisson variables with means $\mu_1, \mu_2, \ldots, \mu_t$ respectively.

In the special case where $S = \{(0, 0, \ldots, 0)\}$, we can leave off the additive error term.

Theorem 2 *The probability that a random d-regular graph of order n has no cycles of length c_i for $1 \le i \le t$ is*

$$\exp\Bigl(-\sum_{i=1}^{t} \mu_i + o(1)\Bigr)$$

as $n \to \infty$.

Corollary 1 *For $(d-1)^{2g-1} = o(n)$, the probability that a random d-regular graph has girth greater than g is*

$$\exp\Bigl(-\sum_{r=3}^{g} \frac{(d-1)^r}{2r} + o(1)\Bigr)$$

as $n \to \infty$.

Since (1.1) implies that $d = o(n^{1/5})$, we can take the total number of d-regular graphs from [7] or [8] to obtain the following.

Corollary 2 *For $(d-1)^{2g-1} = o(n)$, the number of d-regular graphs of order n with girth greater than g is*

$$\frac{(nd)!}{(nd/2)!\, 2^{nd/2} (d!)^n} \exp\Bigl(-\sum_{r=1}^{g} \frac{(d-1)^r}{2r} + o(1)\Bigr)$$

as $n \to \infty$.

To prove the main theorems we first show that the cycles whose lengths are in C are rarely more numerous than a certain bound and rarely share edges with each other even though sharing of vertices is common. Then we use a switching argument to estimate the distribution of the number of cycles when it is below that bound.

2 Bounding the numbers and overlaps of short cycles

In this section G denotes a random d-regular graph on n vertices and $N(n, d)$ denotes the total number of such graphs.

For $1 \le i \le t$, define $R_i = \lfloor \max\{2\mu_i, \log n\} \rfloor$. Let $\mathcal{R} = \mathcal{R}_C(n, d)$ be the set of d-regular graphs of order n such that the number of cycles of length c_i is at most R_i for $1 \le i \le t$, and furthermore that no cycle whose length is in C shares an edge with a different cycle whose length is at most g. First we show that $\mathcal{R}$ includes almost all d-regular graphs of order n.

We will make use of the following, which follows readily from McKay [5, Theorem 2.10].

Theorem 3 *For any d and n such that $N(n,d) \neq 0$, let $J \subseteq E(K_n)$. Then, with $[x]_m$ denoting the falling factorial and j_i the number of edges in J incident with vertex i.*

(a) if $|J| + 2d^2 \leq nd/2$ then

$$\mathbf{P}(J \subseteq E(G)) \leq \frac{\prod_{k=1}^{n} [d]_{j_k}}{2^{|J|}[nd/2 - 2d^2]_{|J|}} \;;$$

(b) if $2|J| + 4d(d+1) \leq nd/2$ then

$$\mathbf{P}(J \subseteq E(G)) \geq \frac{\prod_{k=1}^{n} [d]_{j_k}}{2^{|J|}[nd/2 - 1]_{|J|}} \left(\frac{n-2d-2}{n+2d}\right)^{|J|}.$$

We can now estimate $\mathbf{E}(X_r)$ and $\mathbf{Var}(X_r)$ for $r \in C$ where d and g satisfy (1.1). Note that (1.1) implies that $r = O(\log n)$ and $d = O(n^{1/5})$; this will ensure that both parts of Theorem 3 apply whenever $|J| = O(r)$.

Let J be the edge set of an r-cycle. Then $j_k = 2$ for exactly r values of k, and otherwise it is 0. So by Theorem 3,

$$\mathbf{P}(J \subseteq E(G)) = \frac{(d-1)^r}{n^r}\big(1 + O(rd/n)\big). \tag{2.1}$$

Hence, since $[n]_r = n^r\big(1 + O(r^2/n)\big)$,

$$\mathbf{E}(X_r) = \frac{(d-1)^r}{2r}\big(1 + O(r(r+d)/n)\big). \tag{2.2}$$

Next we estimate $\mathbf{E}(X_r^2)$ in order to find $\mathbf{Var}(X_r)$. Letting $\mathcal{C}$ denote the collection of all r-cycles in K_n (considered as sets of edges),

$$\mathbf{E}(X_r^2) = \sum_{C_1 \in \mathcal{C}} \sum_{C_2 \in \mathcal{C}} \mathbf{P}(C_1 \cup C_2 \subseteq G). \tag{2.3}$$

Partition the pairs (C_1, C_2) into three classes $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ as follows:

- $(C_1, C_2) \in \mathcal{C}_1$ if and only if $C_2 \cap C_1 = \emptyset$,
- $(C_1, C_2) \in \mathcal{C}_2$ if and only if $C_2 \cap C_1 \neq \emptyset$ and $C_2 \neq C_1$,
- $(C_1, C_2) \in \mathcal{C}_3$ if and only if $C_2 = C_1$.

Note that Theorem 3(a) implies that

$$\mathbf{P}(J \subseteq E(G)) \leq \left(\frac{d-1}{n}\right)^{|J|} \big(1 + O(d|J|/n)\big), \tag{2.4}$$

when $j_k \neq 1$ for all k, since then $[d]_{j_k} \leq d(d-1)^{j_k-1}$.

For $\mathcal{C}_1$, since $|C_2 \cup C_1| = 2r$ we have immediately from (2.4) that

$$\sum_{(C_1,C_2)\in\mathcal{C}_1} \mathbf{P}(C_1 \cup C_2 \subseteq G) \leq \mathbf{E}(X_r)^2\big(1 + O(r(r+d)/n)\big). \tag{2.5}$$

The contribution from $\mathcal{C}_3$ is trivially

$$\sum_{(C_1,C_2)\in\mathcal{C}_3} \mathbf{P}(C_1 \cup C_2 \subseteq G) = \mathbf{E}(X_r). \tag{2.6}$$

It remains to consider $\mathcal{C}_2$. This is rather more delicate. For later use we will generalize this calculation to allow C_1 and C_2 to have possibly different lengths, r and s respectively, with $r \in C$ and $3 \leq s \leq g$. Classify the graphs $H = C_1 \cup C_2$ according to the number of components p and edges j in the intersection graph $H' = \big(V(C_1) \cap V(C_2), E(C_1) \cap E(C_2)\big)$.

We proceed by bounding the number of possible isomorphism types of H, which has $r+s-p-j$ vertices and $r+s-j$ edges. Number and orient the components of H' in order of their appearance in C_1. The sizes of these components can be chosen in $\binom{p+j-1}{p-1}$ ways (the number of ordered partitions of j into p nonzero summands). The starting positions of these components in C_1 and C_2 (relative to the position of the first component) can be chosen in at most $\binom{r-1}{p-1}$ and $\binom{s-1}{p-1}$ ways, respectively, but we also have a factor of $2^{p-1}(p-1)!$ because the order and orientation of the components in C_2 can be different. In summary, the number of isomorphism types of H for given p, j is at most

$$\binom{p+j-1}{p-1}\binom{r-1}{p-1}\binom{s-1}{p-1}2^{p-1}(p-1)! \leq \frac{(2g^3)^{p-1}}{(p-1)!^2}.$$

Each can occur in G in $O(n^{r+s-p-j})$ possible positions, each with probability $O(1)\big((d-1)/n\big)^{r+s-j}$, by Theorem 3(a). Therefore,

$$\begin{aligned}
\sum_{(C_1,C_2)\in\mathcal{C}_2} \mathbf{P}(C_1 \cup C_2 \subseteq G) &\leq O(1)\sum_{j,p\geq 1} \frac{(2g^3)^{p-1}}{(p-1)!^2} n^{r+s-p-j}\left(\frac{d-1}{n}\right)^{r+s-j} && (2.7)\\
&= O\left(\frac{(d-1)^{r+s-1}}{n}\right), && (2.8)
\end{aligned}$$

where we have used (1.1) to infer that $g^3 = o(n)$.

Combining (2.3), (2.5) (2.6) and (2.7),

$$\mathbf{Var}(X_r) = \mathbf{E}(X_r) + O\big((r(r+d))/n\big)\mathbf{E}(X_r)^2.$$

It now follows, from Chebyshev's inequality applied separately to each X_r for $r \in C$, that

$$\mathbf{P}(X_{c_i} > R_i \text{ for some } 3 \leq i \leq t) = o(1). \tag{2.9}$$

Moreover, summing (2.7) over $r \in C, 3 \leq s \leq g$, we find that the probability that any cycle whose length is in C shares an edge with a different cycle of length at most g is

$$O\big((d-1)^{2g-1}/n\big) = o(1). \tag{2.10}$$

Applying (2.9) and (2.10), we have

Lemma 1 $N(d,n) = \big(1+o(1)\big)|\mathcal{R}|$. ∎

3 Switchings

Let $\mathcal{R}(m_1,\ldots,m_t)$ denote the subset of $\mathcal{R}$ such that the number of cycles of length c_i is m_i, for $1 \le i \le t$. In view of the definition of $\mathcal{R}$, we make the restrictions $0 \le m_i \le R_i$ for the rest of this section. Put $N(m_1,\ldots,m_t) = |\mathcal{R}(m_1,\ldots,m_t)|$. We will investigate the relative values of $N(m_1,\ldots,m_t)$ by means of a switching argument similar to that used in [8]. Define $C^+ = C \cup \{3,4,\ldots,\lfloor g/2 \rfloor\}$.

Let $Q = (dn)^{1/2}$ and $\delta = (dn)^{-1/2}$. The proof of the following lemma is deferred until later, as it relies on a special case of the result we will use it to prove. Fortunately, this does not create a circular argument, since the special case we will need is one where the lemma is vacuously true.

Lemma 2 *A random G in $\mathcal{R}(m_1,\ldots,m_t)$ has at most Q edges contained in cycles whose length is in $C^+ \setminus C$, with probability at least $1-\delta$.*

Lemma 3

$$\frac{N(m_1,\ldots,m_t)}{|\mathcal{R}|} = (1+o(1)) \prod_{i=1}^{t} \frac{e^{-\mu_i}\mu_i^{m_i}}{m_i!}. \tag{3.1}$$

Proof: Let $G_0 \in \mathcal{R}(m_1,\ldots,m_t)$ with some $m_j > 0$, and set $r = c_j$. Define a *forward r-switching* applied to G_0 as follows. Choose a cycle $Z = (v_0, v_1, \ldots, v_{r-1})$ of length r. Define $e_i = (v_i, v_{i+1})$ for $0 \le i \le r-1$, where subscripts are interpreted modulo r (as they will be henceforth without comment). Also choose r oriented edges $\{e'_i = (w_i, u_{i+1}) \mid 0 \le i \le r-1\}$ not incident with vertices in Z or with each other. Delete these $2r$ edges and add the $2r$ new edges $\{(v_i, w_i), (v_i, u_i) \mid 0 \le i \le r-1\}$. This must be done in such a way that no cycles other than Z whose length is in C may be either created or destroyed, so that the result is a d-regular graph G_1 in the set $\mathcal{R}(m_1,\ldots,m_{j-1}, m_j - 1, m_{j+1},\ldots,m_t)$.

Let F denote the average number of ways to apply a forward r-switching to G_0 if G_0 is chosen at random. As a naive upper bound, after choosing Z in m_j ways, we can choose each e'_i in nd ways. Thus

$$F \le m_j (nd)^r. \tag{3.2}$$

To investigate the sharpness of (3.2), consider the following conditions for all i, i'. When we speak of the *distance* between two edges, we mean the length of the shortest path that starts with a vertex of one of the edges and ends with a vertex of the other.

(a) e'_i does not lie in a cycle whose length is in C^+;
(b) the distance from e'_i to e_i is at least g;
(c) the distance from e'_i to $e'_{i'}$ is at least $g/2$;
(d) the distance from w_i to u_i is at least g.

We claim that any choice of $e'_1, \ldots, e'_r$ satisfying (a)–(d) gives a valid forward r-switching. Cycles other than Z whose length is in C can only be destroyed if they contain some e'_i (not some e_i, by the definition of $\mathcal{R}$), so condition (a) implies that no such cycles are destroyed. No cycles of length g or less are created either, as the following argument shows. Such a cycle Z' would consist of some nontrivial paths in $G_0 \cap G_1$ connected by new edges. These paths in $G_0 \cap G_1$ must have length at least $g/2$ for the following reasons. If they start and finish on Z, apply the definition of $\mathcal{R}$. If they start on Z and finish on $W = \{w_0, \ldots, w_{r-1}, u_0, \ldots, u_{r-1}\}$ (or vice-versa), apply (b). If they start and finish on W, apply (a) or (c). Thus, Z' can include only one nontrivial path in $G_0 \cap G_1$ or it would be too long. The remaining part of Z' must be an edge (v_i, w_i) or (v_i, u_i), eliminated by condition (b), or a path of the form $w_i v_i u_i$, eliminated by condition (d). Since no cycles of length g or less are created, the additional requirement on $\mathcal{R}$ that cycles of length in C cannot share an edge with cycles of length at most g is also preserved.

We can bound the average number of choices (out of $(nd)^r$) eliminated by (a)–(d), for a random $G_0 \in \mathcal{R}(m_1, \ldots, m_t)$. By Lemma 2, condition (a) eliminates

$$O(\delta)(nd)^r + O(r)Q(nd)^{r-1} + O(r)\big((d-1)^g + \log^3 n\big)(nd)^{r-1}$$

choices, since $\sum c_i R_i = O((d-1)^g + \log^3 n))$. Conditions (b) and (d) each eliminate $O(r)(d-1)^g(nd)^{r-1}$. Condition (c) eliminates $O(r^2)(d-1)^{g/2}(nd)^{r-1}$, which is smaller. Comparing these to (3.2), we have

$$F = m_j(nd)^r \left(1 + O\left(\delta + \frac{rQ + r(d-1)^g + r\log^3 n}{nd}\right)\right). \tag{3.3}$$

For $G_1 \in \mathcal{R}(m_1, \ldots, m_{j-1}, m_j - 1, m_{j+1}, \ldots, m_t)$, define a *backward r-switching* applied to G_1 as follows. (This will be the "inverse" operation of a forward switching.) Choose r mutually non-incident oriented 2-paths $u_i v_i w_i$ $(0 \le i \le r-1)$, where the $2r$ possible cyclic orderings including reversal are equivalent. Remove the $2r$ edges of all the paths and insert the edges $\{e_i = (v_i, v_{i+1}), e'_i = (w_i, u_{i+1}) \mid 0 \le i \le r-1\}$. This creates an r-cycle $Z = (v_0, v_1, \ldots, v_{r-1})$, but it is not permitted to create or destroy any other cycles whose length is in C. In fact, the resulting graph G_0 must be in $\mathcal{R}(m_1, \ldots, m_t)$.

Let B denote the average number of ways to apply a backward r-switching to G_1 if G_1 is chosen at random. As a naive upper bound, we can choose each oriented 2-path in $nd(d-1)$ ways, which achieves each cyclic ordering $2r$ times. Hence,

$$B \le \frac{\big(nd(d-1)\big)^r}{2r}. \tag{3.4}$$

To investigate the sharpness of (3.4), consider the following conditions for all i and $1 \le k \le g/2$:
(a) the edges (v_i, w_i) and (v_i, u_i) do not lie in any cycles whose length is in C^+;
(b) the distance between the 2-paths $u_i v_i w_i$ and $u_{i+1} v_{i+1} w_{i+1}$ is at least g;
(c) the distance between vertices v_i and v_k is at least $g - k + 1$.

We claim that conditions (a)–(c) are together enough to ensure that the backward r-switching is valid. Condition (a) ensures that no cycle whose length is in C is destroyed. Furthermore, except for Z, no cycle of length g or less is created as the following argument shows. Such a cycle Z' would consist of nontrivial paths in $G_0 \cap G_1$, portions of Z, and edges (w_i, u_{i+1}). Potential such paths in $G_0 \cap G_1$ have length at least $g/2$ by (a) and (c), so there can be only one such path. The remaining part of Z' is either an edge (w_i, u_{i+1}), which is eliminated by condition (b), or a segment of k edges of Z, which is eliminated by condition (c). Since no cycles of length g or less are created, in particular we do not create one that shares an edge with another. Thus the switching satisfies all the requirements.

We can bound the number of choices (out of $(nd(d-1))^r$) eliminated by (a)–(c) for a random $G_1 \in \mathcal{R}(m_1, \dots, m_j - 1, \dots, m_t)$. Condition (a) eliminates

$$O(\delta)(nd(d-1))^r + O(r)(nd(d-1))^{r-1}(d-1)\big((d-1)^g + \log^3 n + Q\big)$$

choices, by the same argument as for condition (a) of the forward switchings. Condition (b) eliminates $O(r)(nd(d-1))^{r-1}(d-1)^{g+1}$ choices. Finally, condition (c) eliminates

$$O(r)(nd(d-1))^{r-1} \sum_{k=1}^{\lfloor g/2 \rfloor} (d-1)^{g-k+2} = O(r)(nd(d-1))^{r-1}(d-1)^{\lceil g/2 \rceil + 1},$$

which is smaller. Comparing these to (3.4), we have

$$B = \frac{\big(nd(d-1)\big)^r}{2r}\left(1 + O\left(\delta + \frac{rQ + r(d-1)^g + r\log^3 n}{nd}\right)\right). \tag{3.5}$$

From (3.3) and (3.5), it follows that

$$\frac{N(m_1, \dots, m_t)}{N(m_1, \dots, m_{j-1}, m_j - 1, m_{j+1}, \dots, m_t)} \tag{3.6}$$

$$= \frac{B}{F} = \frac{(d-1)^r}{2rm_j}\left(1 + O\left(\delta + \frac{rQ + r(d-1)^g + r\log^3 n}{nd}\right)\right). \tag{3.7}$$

The values of δ and Q, together with the fact that $\sum_i c_i m_i = O((d-1)^g + \log^3 n)$, allow us to apply (3.6) repeatedly to obtain

$$\frac{N(m_1, \dots, m_t)}{N(0, \dots, 0)} = (1 + o(1)) \prod_{i=1}^{t} \frac{\mu_i^{m_i}}{m_i!}, \tag{3.8}$$

where the error term depends only on n. Summing over $0 \le m_i \le R_i$ for each i, with crude tail estimates, gives the lemma. ∎

In view of Lemma 1, this is very close to Theorem 1, but we have still to prove Lemma 3. This will rely on some crude bounds on the probability that there are very many short cycles, for which we begin with the following technical lemma.

Lemma 4 *Let $S_1, S_2, \ldots, S_q$ be finite sets of size at most k, with q finite. Define $W = \bigcup_{i=1}^{q} S_i$, and, for each $w \in W$, $W_w = \bigcup\{S_i \mid w \notin S_i\}$. Then at least half the elements $w \in W$ have the property that $1 \leq |W \setminus W_w| \leq 2k$.*

Proof: Note that $W \setminus W_w$ consists of all the elements v such that w lies in every S_i that v lies in. Let E be the set of pairs (v, w) of distinct elements of W such that $w \in \bigcap\{S_i \mid v \in S_i\}$ for each v. Each element of W can appear as v in at most $k-1$ pairs in E, since the sets have size at most k, and so $|E| \leq (k-1)|W|$. Therefore, the average number of times each element of W appears as w in a pair is at most $k-1$, so the average size of $|W \setminus W_w|$ is at most k. The result follows. ∎

Theorem 4 *Let $k = k(n) \geq 3$ and $d = d(n) \geq 3$ satisfy $k(d-1)^{k-1} = o(n)$. Let $M = M(n) = 20Ak(d-1)^k$ with $A = A(n) > c$ for some constant $c > 1$. Then the probability that a random d-regular graph of order n has exactly M edges which lie on cycles of length at most k is less than*

$$\left(e^{5(A-1)}A^{-5A}\right)^{(d-1)^k} = e^{-5(d-1)^k}(e/A)^{M/4k}$$

for sufficiently large n.

Proof: Write $D = (d-1)^k$. Let $X(G)$ be the number of edges of G that lie on cycles of length at most k, and let G_m be the set of d-regular graphs of order n such that $X(G) = m$. Also let $N_m = |G_m|$.

We will use a standard switching argument. Let $G \in G_m$ for $m > 1$. For each edge e, let $f(e) = X(G) - X(G-e) = m - X(G-e)$.

Choose an edge $e = (v, w)$ such that $1 \leq f(e) \leq 2k$. By Lemma 4, e can be chosen in at least $m/2$ ways. Now choose an edge $e' = (v', w')$, of distance at least $k-1$ from e, such that $f(e') \leq 2k$. Using Lemma 4 again, e' can be chosen in at least $nd/2 - m/2 - O(D) \geq nd/4 - O(D)$ ways. Now remove e, e' and insert either the two edges $(v, v'), (w, w')$ or the two edges $(v, w'), (v', w)$. In total, this switching operation can be performed in at least

$$\tfrac{1}{4}mnd(1 + o(1)) \tag{3.9}$$

ways. Let G' be the resulting graph. Since $f(e), f(e') \leq 2k$, we have $X(G') \geq m - 4k$. We also have $X(G') \leq m$, where equality is possible since the two new edges may lie together on some cycle of length at most k. (Edges of G lying on short cycles in G' must also lie on short cycles in G due to the distance between e and e', and so do not contribute to $X(G) - X(G')$.)

Now let G' be any d-regular graph of order n. To perform an operation inverse to the switching defined above, we need to first choose a path of length at most $k+1$ subject to some restrictions. (The inverse operation removes the first and last edges of the path and inserts two new edges.) The number of such paths is at most

$$ndD. \tag{3.10}$$

Now count the pairs (G, G') such that G' results by a switching from G and $G' \in G_{m-4k} \cup \cdots \cup G_{m-1} \cup G_m$. Considering that all the switchings from G land in the required place, (3.9) and (3.10) imply that

$$N_m \le \frac{(4+o(1))D}{m} \sum_{i=0}^{4k} N_{m-i},$$

which implies that

$$N_m \le \frac{5D}{m} \sum_{i=1}^{4k} N_{m-i} \tag{3.11}$$

if $m \ge 21D$ and n is large enough (since the $o(1)$ is independent of m).

One of the ways (3.11) can be used to bound N_m for large m is to notice that it implies

$$N_m \le \frac{20kD}{m} \max_{1 \le i \le 4k} N_{m-i}.$$

If $m > 20kAD$ with $A > c > 1$, we can apply this inequality repeatedly while the coefficient is at least 1. This gives

$$N_m \le \frac{(20kD)^\ell}{m_0 m_1 \cdots m_{\ell-1}} N_{m_\ell}$$

for some sequence $m = m_0 > m_1 > \cdots > m_\ell$ such that $m_i - m_{i+1} \le 4k$ for all i and $20kD - 4k \le m_\ell \le 20kD - 1$. It is easy to see that the weakest bound occurs when $m_i - m_{i+1} = 4k$ for all i. Using Stirling's formula, this gives $N_m \le \left(e^{5(A-1)} A^{-5A}\right)^D N_{m_\ell}$. This gives the required bound since $N_{m_\ell} \le N(d,n)$. ∎

Proof of Lemma 2: In the case that $C^+ \setminus C$ is empty, the lemma is vacuously true, so the proof of Lemma 3 is valid when C is replaced by C^+. This implies that $\mathcal{R}(m_1, \ldots, m_t)$ is a fraction at least

$$(1+o(1)) \sum_{i=1}^{t} e^{-\mu_i} \min\left\{1, \mu_i^{R_i}/R_i!\right\} = \exp\left(-O(1)(d-1)^g/g - O(1)\log^2 n\right) \tag{3.12}$$

of all d-regular graphs. Now apply Theorem 4 with $M > Q$ and $k = \lfloor g/2 \rfloor$ to find the probability $p(M)$ in the space of all d-regular graphs that exactly M edges lie in cycles of length at most k. Using (1.1), we have that $e/A \to 0$. Hence

$$p(M) = \exp\left(-\omega(n)(dn)^{1/2}/k\right)$$

for some $\omega(n) \to \infty$. Hence, by (3.12), the probability restricted to $\mathcal{R}(m_1, \ldots, m_t)$ is

$$\exp\left(-\omega(n)(dn)^{1/2}/k + O(1)(d-1)^g/g + O(1)\log^2 n\right).$$

From (1.1) we know that the first term dominates the others, and so the restricted probability is
$$\exp\bigl(-\tfrac{1}{2}\omega(n)(dn)^{1/2}/k\bigr) = O(e^{-n^{1/3}}),$$
which is smaller than δ even if summed over $M > Q$. ∎

Theorem 1 now follows from Lemmas 1, 2 and 3. To prove Theorem 2, note that the additive $o(1)$ term in Theorem 1 comes only from those d-regular graphs of order n that are not in $\mathcal{R}$. There are no such graphs without cycles whose lengths are in C, by the definition of $\mathcal{R}$, so the additive $o(1)$ term is 0 in that case.

4 Bipartite Regular Graphs

The same analysis can be done with the same method for the case of random bicoloured regular graphs, assuming that n and all cycle lengths are even. The only significant difference is that switchings must preserve the colour classes.

The results are almost the same. Define $\mu_i' = (d-1)^{c_i}/c_i$. Then Theorems 1 and 2 hold with μ_i' replacing μ_i. (Similarly, in the proofs, $R_i' = \lfloor \max\{2\mu_i', \log n\}\rfloor$.) The results corresponding to Corollaries 1 and 2 are as follows, where the total number of bipartite regular graphs comes from [6].

Corollary 3 *Let n and g be even, and $(d-1)^{2g-1} = o(n)$. Then the probability that a random d-regular bipartite graph has girth greater than g is*

$$\exp\left(-\sum_{s=2}^{g/2}\frac{(d-1)^{2s}}{2s} + o(1)\right)$$

as $n \to \infty$.

Corollary 4 *Under the same conditions, the number of d-regular bipartite graphs of order n with girth greater than g is*

$$\frac{(nd/2)!}{(d!)^n}\exp\left(-\sum_{s=1}^{g/2}\frac{(d-1)^{2s}}{2s} + o(1)\right)$$

as $n \to \infty$.

References

[1] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Combin.* **1** (1980) 311–316.

[2] B. Bollobás, *Random graphs*, Academic Press, London, 1985.

[3] H. Garmo, The asymptotic distribution of long cycles in random regular graphs. *Random Structures Algorithms* **15** (1999) 43–92.

[4] S. Janson, *Orthogonal Decompositions and Functional Limit Theorems for Random Graph Statistics*, Memoirs Amer. Math. Soc. 534, Amer. Math. Soc., Providence, R.I. (1994).

[5] B. D. McKay, Subgraphs of random graphs with specified degrees, *Congressus Numerantium* **33** (1981) 213–223.

[6] B. D. McKay, Asymptotics for 0-1 matrices with prescribed line sums, *in* Enumeration and Design, (Academic Press, 1984) 225–238.

[7] B. D. McKay, Asymptotics for symmetric 0-1 matrices with prescribed row sums, *Ars Combinatoria*, **19A** (1985) 15–26.

[8] B. D. McKay and N. C. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$, *Combinatorica* **11** (1991) 369–382.

[9] A. Ruciński, When are small subgraphs of a random graph normally distributed? *Probability Theory and Related Fields* 78 (1988), 1–10.

[10] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs, *J. Combin. Theory, Ser. B* **31** (1981) 168–182.

[11] N. C. Wormald, Models of random regular graphs. In J.D. Lamb and D.A. Preece (eds.), *Surveys in Combinatorics 1999*, *LMS Lecture Note Series* **267**, pp. 239–298. Cambridge University Press (1999).