The Asymptotic Distribution of Short Cycles in Random Regular Graphs

NICHOLAS C. WORMALD*

Department of Mathematics, Louisiana State University,
Baton Rouge, Louisiana 70803

Communicated by the Managing Editors
Received January 23, 1980

The probability that a random labelled $r$-regular graph contains a given number of cycles of given length is investigated asymptotically. Cycles of a finite number of different fixed lengths can be handled simultaneously. For $i \neq j$, the probabilities of an $i$-cycle and a $j$-cycle occurring are asymptotically independent. The results are actually derived in the more general setting of graphs which have any given degree sequence, as long as the maximum degree is bounded above by a constant. As a special case, an asymptotic formula results for the number of labelled $r$-regular graphs with a given girth.

1. INTRODUCTION

The main objective of this paper is an asymptotic formula for the number of labelled graphs with a given degree sequence and given numbers of cycles of given fixed lengths. The proofs are based on a formula obtained by Bender and Canfield [2] for the asymptotic number of symmetric non-negative integer matrices with given bounded row sums. As special applications, they obtain asymptotic formulae for the numbers of labelled graphs and multigraphs whose points have given bounded degrees. Their corresponding formula for pseudographs, in which loops are also allowed, only applies if each loop is counted just once in calculating the degree of a point. Usually a loop is regarded as contributing 2 to the degree of its adjacent point. The appropriate asymptotic formula under this more common definition of degree in a pseudograph is derived in Section 3. The main result is then treated in Section 4, using the result of Section 3.

* The results of this paper form a portion of the author's doctoral thesis, written at the University of Newcastle under the direction of Professor R. W. Robinson. The preparation of this paper was supported by the Natural Sciences and Engineering Research Council of Canada under Grants A8142 and A8235.
The results of the present paper appear in the author's doctoral thesis [7]. There, the proofs do not use [2] but instead begin with a formula of Békésy et al. [1] concerning the asymptotic enumeration of non-negative integer matrices with given bounded row and column sums. Some of the methods found in the present paper were used in [7] to obtain asymptotic results concerning the expected numbers of given submatrices of a certain type occurring in a random non-negative integer matrix with given row and column sums, and also the probability that a given special type of submatrix occurs. As a special case, this implies results concerning the occurrence of subgraphs in labelled pseudographs in which the lines are labelled independently of the points. The asymptotic effect of unlabelling the lines of such a pseudograph was then studied, thereby producing the results obtained herein as well as result (2.2) in Section 2. Related methods were also used to find asymptotic formulae for the numbers of certain types of bicoloured graphs with given girth.

2. PRELIMINARIES

For basic graph theoretic notation we refer to Harary [5]. Throughout this paper it is assumed, unless otherwise specified, that the \( p \) points of a pseudograph are labelled with the integers \( 1, \ldots, p \). Two such pseudographs are distinct if and only if there is no label-preserving isomorphism between them.

The degree sequence of a pseudograph \( G \) is the sequence \( d_1, \ldots, d_p \), where \( d_i \) is the degree of the point labelled \( i \). If \( d_i = r \) for each \( i \) then \( G \) is called \( r \)-regular or just regular. The set of pseudographs on \( p \) points with degree sequence \( d_1, \ldots, d_p \) is denoted by \( \mathcal{P}(d_1, \ldots, d_p) \) or simply \( \mathcal{P} \), and the set of multigraphs in \( \mathcal{P} \) is denoted by \( \mathcal{M}(d_1, \ldots, d_p) \) or just \( \mathcal{M} \). In other words, \( \mathcal{M} \) is the set of loopless pseudographs in \( \mathcal{P} \). The cardinality of \( \mathcal{M} \) is denoted by \( M(d_1, \ldots, d_p) \) or simply \( M \), and the cardinality of \( \mathcal{P} \) is denoted by \( P(d_1, \ldots, d_p) \) or simply \( P \). For the purposes of this article, we only consider multigraphs and pseudographs whose points have degrees at most \( D \), where \( D \) is any pre-chosen positive constant. Consequently, it is assumed throughout that \( d_1, \ldots, d_p \) is always a sequence of positive integers satisfying \( 0 \leq d_i \leq D \) for \( 1 \leq i \leq p \), whose sum is even.

Unless otherwise specified, all summations \( \sum \) and products \( \prod \) are for \( i = 1, \ldots, p \). By convention \( (j) = 0 \) if \( i < 0 \) or \( j < 0 \) or \( j > i \). The following notations are used throughout:

\[
q = q(d_1, \ldots, d_p) = \frac{1}{2} \sum d_i \quad \text{(this is the number of lines in a pseudograph with degree sequence } d_1, \ldots, d_p)\;.
\]
\[ \alpha = \alpha(d_1, \ldots, d_p) = \frac{\sum (d_i^2)}{\sum d_i} = \frac{\sum d_i(d_i - 1)}{4q}; \]

\[ N = N(d_1, \ldots, d_p) = \frac{(2q)! e^{\alpha + \alpha}}{q! 2^{\alpha} \prod d_i!}; \]

\[ \gamma = \gamma(d_1, \ldots, d_p) = \sum_{d_i > 2} 1 \quad \text{(this is the number of points of degree \( \geq 2 \)).} \]

For later use we note that \( 0 \leq \alpha \leq \frac{1}{2}(D - 1) \) and hence
\[ 1 \leq e^\alpha \leq e^{(D - 1)/2}. \] (2.1)

For \( l \geq 1 \), a cycle of length \( l \) is called an \( l \)-cycle and is denoted by \( C_l \), and a pseudograph consisting entirely of \( t \) disjoint \( l \)-cycles is denoted by \( tC_l \). By \( F \) we always mean an unlabelled pseudograph with \( m \) points, whose degrees are \( k_1, \ldots, k_m \) in some order, and \( n \) lines, where \( n \geq 1 \). Define
\[ W(F) = W(F; d_1, \ldots, d_p) = \sum \prod_{r=1}^{m} \prod_{j=0}^{k_r-1} (d_{s_r} - j), \]

where the summation is over all sequences \( s_1, \ldots, s_m \) of distinct integers in the range \( 1 \leq s_r \leq p \). The number of distinct automorphisms of \( F \) is denoted by \( A(F) \), and we put
\[ Y(F) = Y(F; d_1, \ldots, d_p) = \frac{(2q)^{-n} W(F)}{A(F)}. \]

A pseudograph \( G \) is \textit{rooted} at \( F \) if a subgraph of \( G \), called the \textit{root subgraph}, which is isomorphic to \( F \) is distinguished. By this it is meant that the points and lines in the root subgraph can be distinguished from the other points and lines in \( G \). Two pseudographs rooted at \( F \) are counted as the same only if there is an isomorphism from one to the other which preserves the root subgraph.

In all our asymptotic results it is assumed that \( D \) is fixed. More specifically, if \( A(d_1, \ldots, d_p) \) is a function of \( d_1, \ldots, d_p \),
\[ A(d_1, \ldots, d_p) = o(1) \]

denotes that for all \( \varepsilon > 0 \) there is some \( q_0(D, \varepsilon) \) such that \( A(d_1, \ldots, d_p) < \varepsilon \) whenever \( p > 0 \) and \( q > q_0(D, \varepsilon) \) and \( \sum d_i = 2q \) and \( 0 \leq d_i \leq D \) for \( 1 \leq i \leq p \). Similarly,
\[ A(d_1, \ldots, d_p) = O(1) \]

denotes that there is some \( K(D) \) such that \( A(d_1, \ldots, d_p) < K(D) \) whenever \( p > 0 \) and \( 0 \leq d_i \leq D \) for \( 1 \leq i \leq p \). Often \( o \) or \( O \) is subscripted by \( t \) or \( F \) to remind us that a quantity \( t \) or a pseudograph \( F \) also is being held fixed.
The special case of a result of [2] which will be used heavily is

\[ M = (1 + o(1))Ne^{-2a}. \]  

(2.2)

We close with two results to be exploited in other sections.

**Lemma 1.** If \( d_1, \ldots, d_p \) are restricted so that \( \gamma \to \infty \) as \( q \to \infty \) and if \( k_r = 2 \) for \( 1 \leq r \leq m \), then

\[ W(F) = (1 + o_f(1))(4qa)^m. \]

**Proof.** First observe that since \( k_r = 2 \) for each \( r \),

\[ |W(F) - (4qa)^m| \leq \binom{m}{2} (4qa)^{m-2} \sum d_r^2(d_1 - 1)^2. \]  

(2.3)

But \( \sum d_r^2(d_1 - 1)^2 \leq \gamma D^2(D - 1)^2 \) and \( (4qa)^2 > (2\gamma)^2 \). Thus, since \( \gamma \to \infty \) we have \( \sum d_r^2(d_1 - 1)^2 = o(4qa)^2 \), and hence (2.3) implies the stated asymptotic relation.

**Lemma 2.** If \( d_1, \ldots, d_p \) are restricted so that

\[ \gamma/q \to 0 \quad \text{as} \quad q \to \infty \]

then,

\[ W(F) = o_p(p^m). \]

**Proof.** At most \( \gamma p^{m-1} \) sequences \( s_1, \ldots, s_m \) of distinct integers from \( 1, \ldots, p \) satisfy

\[ \prod_{r=1}^m d_r(s_r) - 1 \neq 0. \]

It follows that \( W(F) \leq \gamma p^{m-1}(D!)^m \), which is \( o_p(p^m) \) since \( q \leq pD/2 \).

3. Asymptotic Enumeration of Locally Restricted Pseudographs

**Theorem 3.** The number of labelled pseudographs with \( q \) lines and with degree sequence \( d_1, d_2, \ldots, d_p \) is \( (1 + o(1))N \).

**Proof.** Assume \( d_i \neq 0 \) for \( 1 \leq i \leq p \), so that \( p \leq 2q \). Since \( P(d_1, \ldots, d_p) = P(d_1, \ldots, d_p, 0, 0, \ldots, 0) \) and \( N(d_1, \ldots, d_p) = N(d_1, \ldots, d_p, 0, 0, \ldots, 0) \), this affords no loss of generality.

Let \( \mathcal{P}' = \mathcal{P}'(d_1, \ldots, d_p) \) be the set of pseudographs in \( \mathcal{P} \) which contain some point incident with at least two loops, and let \( \mathcal{P}_t = \mathcal{P}_t(d_1, \ldots, d_p) \) be the
set of pseudographs in $\mathcal{S}\setminus \mathcal{S}'$ which contain precisely $t$ loops. Let $P' = P'(d_1, \ldots, d_p)$ denote the cardinality of $\mathcal{S}'$, and let $P_t = P_t(d_1, \ldots, d_p)$ denote the cardinality of $\mathcal{S}_t$. In particular, $P_0 = M$. Put

$$R(q) = \max_{\sum d_k = 2q} \frac{P}{N},$$

remembering $d_i \leq D$.

The number of pseudographs in $\mathcal{S}$ in which the first point is incident with at least two loops is precisely

$$P(d_1 - 4, d_2, \ldots, d_p) \leq R(q - 2) N(d_1 - 4, d_2, \ldots, d_p)$$

$$= O(1) R(q - 2) q^2 N/(2q - 3)^t \quad \text{(using (2.1))}$$

$$= O(1) R(q - 2) Nq^{-2}.$$

A pseudograph in $\mathcal{S}'$ has at least two loops incident with one of its $p \leq 2q$ points, so summing over all possibilities for the position of a double loop, we obtain

$$P' = O(1) R(q - 2) Npq^{-2} = O(1) R(q - 2) Nq^{-1}. \quad (3.1)$$

We require an asymptotic formula for $P_t$. If the loops of a pseudograph in $\mathcal{S}_t$ are incident with the points $1, 2, \ldots, t$, then the number of possibilities for the multigraph which results when all the loops are removed is given asymptotically by (2.2) as

$$(1 + o_t(1)) \frac{(2q - 2t)! e^{\beta_t - \beta}}{(q - t)! 2^{q-t} \prod d'_i!},$$

where $d'_i = d_i - 2$ for $i \leq t$ and $d'_i = d_i$ otherwise, and $\beta = \alpha(d_1 - 2, \ldots, d_t - 2, d_{t+1}, \ldots, d_p)$. By summing over all possibilities for the positions $s_1, \ldots, s_t$ of the loops in a pseudograph in $\mathcal{S}_t$, we obtain

$$P_t = (1 + o_t(1)) \sum^* \frac{(2q - 2t)! e^{\beta_t - \beta}}{(q - t)! 2^{q-t} \prod d'_i!},$$

where $*$ denotes that the summation ranges over all sequences $1 \leq s_1 < \cdots < s_t \leq p$ satisfying $d_{s_j} \geq 2$ for each $j$,

$$d'_i = d_i - 2 \quad \text{if} \quad i = s_j \text{ for some } j$$

$$= d_i \quad \text{otherwise}$$

and $\beta = \alpha(d'_1, \ldots, d'_p)$. 
Since \( d_i \leq D \), we have \( \beta - \alpha = o_i(1) \) and so
\[
e^{\beta - \alpha}/e^{\alpha^2 - \alpha} = 1 + o_i(1).
\] (3.2)

Hence
\[
P_t = (1 + o_i(1)) Ne^{-2\alpha} \frac{(2q - 2t)! q! 2^t W(tC_1)}{(2q)! (q - t)! t!},
\]
the \( t! \) arising since in \( \sum^* \) the condition \( s_1 < s_2 < \cdots < s_t \) holds, whereas in \( W(tC_1) \) the variables are unordered. Thus
\[
P_t = (1 + o_i(1)) Ne^{-2\alpha} \frac{W(tC_1)}{(2q)! t!}. \quad (3.3)
\]

In order to be able to ignore \( P_t \) for large \( t \), we need to find a suitable upper bound.

From (2.1) and (2.2) we have \( M = O(1)N \) and \( N = O(1)M \). With these variations of (2.2), and with (2.1) in place of (3.2), the fashion of argument which leads to (3.3) suffices to show
\[
P_t = O(1)N \frac{(2q - 2t)! q! 2^t W(tC_1)}{(2q)! (q - t)! t!}, \quad (3.4)
\]
where the bound denoted by \( O(\ ) \) is independent of \( t \). Note that \( p \leq 2q - t \) since a pseudograph in \( \mathcal{P}_t \) has \( t \) loops incident with \( t \) distinct points, and \( d_i \neq 0 \) for each \( i \). From this and the bound \( W(tC_1) \leq (D(D - 1))^t p!/p(t - t)! \), (3.4) now implies
\[
P_t = O(1)N(D(D - 1))^t/p!. \quad (3.5)
\]

We now have
\[
\sum_{t=0}^{\infty} P_t = O(1)N e^{\rho(D - 1)} = O(1)N.
\]
Since
\[
P = P' + \sum_{t=0}^{\infty} P_t, \quad (3.6)
\]
this and (3.1) imply
\[
P = O(1)(R(q - 2) Nq^{-1} + N),
\]
so that
\[
P/N = O(1)(R(q - 2) q^{-1} + 1)
\]
and hence
\[ R(q) = O(1)(R(q - 2)q^{-1} + 1). \]

It follows that \( R(q) = O(1) \), and so from (3.1) we have
\[
P' = o(1) N
= o(1) M \quad \text{by (2.1) and (2.2)}.\]

As \( P > P_0 = M \), it now follows from (3.6) that
\[
P = (1 + o(1)) \sum_{t=0}^{\infty} P_t,
\]
or
\[
P = (1 + o(1)) \left( \sum_{t=0}^{j} P_t + \sum_{t=j+1}^{\infty} P_t \right) \quad (3.7)
\]
for all \( j \geq 0 \). Note that from (3.5) we have that for \( j > 2D(D - 1) \),
\[
\sum_{t=j+1}^{\infty} P_t = O(1)N(D(D - 1))^{j/j!}.
\]

This shows that when \( j \) is large, \( P \) is close to \( \sum_{t=0}^{j} P_t \).

To complete the proof of the theorem we consider two separate cases. Firstly, restrict attention to sequences \( d_1, \ldots, d_p \) which satisfy \( \gamma \geq \sqrt{q} \). Then by Lemma 1,
\[
W(tC_1) = (1 + o_t(1))(4qa)^t,
\]
since if \( F = tC_1 \) then \( k_j = 2 \) for \( j = 1, \ldots, t \). By substituting this in (3.3), one obtains
\[
P_t = (1 + o_t(1)) Ne^{-2\alpha}(2\alpha)^t/t!.
\]

From (3.7) and (3.8), therefore,
\[
\left| P - Ne^{-2\alpha} \sum_{t=0}^{j} (2\alpha)^t/t! \right| = o_j(N) + O_j(1) N(D(D - 1))^{j/j!}.
\]
for \( j \) sufficiently large. It follows that
\[
P = (1 + o(1)) N. \quad (3.9)
\]
In the second case, restrict attention to sequences \( d_1, \ldots, d_p \) which satisfy \( \gamma < \sqrt{q} \), and take \( t > 0 \). Then by Lemma 2, \( W(tc_1) = o_t(p') \). As \( p \leq 2q \) this means \( W(tc_1) = o_t(q') \), and so (3.4) yields

\[
P_t = o_t(N).
\]

Since this holds for every \( t > 0 \), and since \( P_0 = M \), it now follows from (3.7) and (3.8) that

\[
P = (1 + o(1)) Ne^{-2\alpha}.
\]

(3.10)

As \( \gamma < \sqrt{q} \) we have

\[
\alpha \leq \sqrt{q} D(D - 1)/(4q) = o(1)
\]

and so (3.9) follows from (3.10). The combination of the two cases establishes the theorem.

4. Rootings and Cycle Restrictions

Asymptotically speaking, in order to count regular pseudographs with a specified number of \( l \)-cycles, we wish firstly to count regular pseudographs in which a number of \( l \)-cycles are distinguished. It is then possible to count regular pseudographs with precisely \( t(1) \) different \( l(1) \)-cycles and \( t(2) \) different \( l(2) \)-cycles by counting, firstly, regular pseudographs with precisely \( t(1) \) different \( l(1) \)-cycles in which a number of \( l(2) \)-cycles are distinguished. In the proof of the following theorem, this procedure is applied inductively, for many cycle lengths, to pseudographs with degree sequence \( d_1, \ldots, d_p \).

Throughout this section, when \( o \) and \( O \) are used it is always with the understanding that \( F \) and \( v, t, t(1), t(2), \ldots, t(v) \), \( l, l(1), l(2), \ldots, l(v) \) are all fixed. We define \( y_t = (2\alpha)^l/A(C_l) \) so that

\[
y_t = (2\alpha)^l/2l
\]

for \( l \geq 3 \) and

\[
y_t = (2\alpha)^l/l
\]

for \( l = 1 \) and 2. Also, put \( X_0 = 1 \) and

\[
X_j = X_j(d_1, \ldots, d_p) = \prod_{l=1}^j \frac{\exp(-y_{t(i)}) y_{t(i)}^{t(i)}}{t(i)!}.
\]
THEOREM 4. Let \( v, l(1), ..., l(v), t(1), ..., t(v) \) be non-negative integers with the \( l(i) \) all different, and suppose \( F \) contains no \( l(j) \)-cycle for \( j = 1, ..., v \). Then the number of pseudographs with degree sequence \( d_1, ..., d_v \) which contain precisely \( t(j) \) different \( l(j) \)-cycles for \( 1 \leq j \leq v \) is

(i) \( N(o(1) + X_v) \) for pseudographs which are not rooted, and
(ii) \( NY(F)(o(1) + X_v) \) for pseudographs rooted at \( F \).

Note. Depending on \( d_1, ..., d_v \), for any given \( p \) the number \( X_v \) can be less or greater than the value of the function of \( p \) which is denoted by \( o(1) \).

Proof. The two parts of the theorem are proved by a "two-legged" induction on \( v \). Firstly, when \( v = 0 \) part (i) is just Theorem 3. It will be shown that the truth of (i) for \( v = k \) implies the truth of (ii) for \( v = k \), and that the truth of (i) and (ii) for \( v = k \) implies the truth of (i) for \( v = k + 1 \).

Assume (i) holds for \( v = k \geq 0 \). Let \( G \) be a pseudograph with degree sequence \( d_1, ..., d_v \) and rooted at \( F \) such that for \( 1 \leq i \leq v \) there are precisely \( t(i) \) different \( l(i) \)-cycles in \( G \) which contain no line of the root subgraph of \( G \). Let \( P^*(F) \) be the number of possibilities for such a \( G \). We want to know the asymptotic value of \( P^*(F) \).

Since \( G \) is rooted at \( F \) there is a sequence \( r_1, ..., r_m \) of integers from \( 1, ..., p \), corresponding to the labels of the points in the root subgraph, such that \( d_{r_i} \geq k_i \) for \( 1 \leq i \leq m \). Let \( G' \) be the graph obtained by deleting the edges of the root subgraph from \( G \). Then \( G' \) contains precisely \( t(i) \) different \( l(i) \)-cycles for \( 1 \leq i \leq k \). With \( d'_i \) denoting the degree of the \( i \)th point of \( G' \), we therefore have by (i) and the definition of \( N \) that the number of possibilities for \( G' \) is given by

\[
N(d'_1, ..., d'_p)(o(1) + X_k(d'_1, ..., d'_p)). \tag{4.1}
\]

Here \( o(1) \) is valid because \( F \) is fixed and so

\[
\sum d'_i \to \infty \quad \text{as} \quad \sum d_i \to \infty.
\]

In fact, \( 2q = \sum d_i = 2n + \sum d'_i \), and it follows that

\[
\alpha(d'_1, ..., d'_p) = o_F(1) + \alpha(d_1, ..., d_p). \tag{4.2}
\]

Hence, in view of (2.1),

\[
X_k(d'_1, ..., d'_p) = o_F(1) + X_k(d_1, ..., d_p).
\]

As \( X_k \) is bounded above, it is thus seen that expression (4.1) can be written as

\[
N(2q)^{-"W"}(o(1) + X_k(d_1, ..., d_p)), \tag{4.3}
\]
where

\[ W' = \prod_{i=1}^{m} d_i! / \prod_{i=1}^{m} d'_i! = \prod_{l=1}^{k_i - 1} (d_{r_l} - j). \]

Note that (4.3) gives the number of possibilities for \( G \) such that the points of the root subgraph \( F \) correspond in a given order to the points \( r_1, \ldots, r_m \) of \( G \). Hence, if we sum (4.3) over all sequences \( r_1, \ldots, r_m \) satisfying \( d_{r_l} \geq k_l \), we will have counted each possibility for \( G \) rooted at \( F \) precisely \( A(F) \) times. Note that the sum of \( W' \) over all such sequences is just \( W(F) \).

Since the term \( o(1) \) in (4.3) is independent of the choice of \( r_1, \ldots, r_m \), this implies that the total number of possibilities for \( G \) rooted at \( F \) is given by

\[
P^*(F) = N(2q)^{-n} W(F)(A(F))^{-1}(o(1) + X_k) = NY(F)(o(1) + X_k).
\] (4.4)

Let \( P^\#(F) \) denote the number of possibilities for a pseudograph \( G \) with degree sequence \( d_1, \ldots, d_p \) and rooted at \( F \), containing precisely \( t(i) \) different \( l(i) \)-cycles for each \( i \). Then the difference between \( P^*(F) \) and \( P^\#(F) \) is at most the number of possibilities for \( G \) rooted at \( F \) such that there is a line of the root subgraph contained in an \( l(j) \)-cycle, \( C \) say, of \( G \) for some \( j \) (with no other restrictions on the numbers of \( l(i) \)-cycles in \( G \)). Let \( H \) be the pseudograph consisting of the points and lines in \( C \) or the root subgraph of \( G \). \( H \) is commonly called the union of \( C \) and the root subgraph. Given \( H \) as a subgraph of \( G \), the number of ways of recovering \( C \) and the root subgraph from \( H \) is bounded above by a number \( f_j \) depending only on \( F \) and \( l(j) \). It follows that

\[
|P^\#(F) - P^*(F)| \leq f \sum P(H),
\] (4.5)

where \( f \) denotes the maximum of \( f_i \) for \( 1 \leq i \leq v \), the sum is taken over all \( j \) and all possibilities for the pseudograph \( H \) as a union of \( F \) and an \( l(j) \)-cycle sharing a line, and \( P(H) \) denotes the number of possibilities for \( G \) rooted at \( H \).

Let \( m' \) and \( n' \) be the numbers of points and lines, respectively, in \( H \). Since \( F \) contains no \( l(j) \)-cycle, but \( C \) is an \( l(j) \)-cycle, we have

\[
m' - n' \leq m - n - 1.
\] (4.6)

Note that there are \( m \) points of \( H \) which have degrees at least \( k_1, \ldots, k_m \) (these come from \( F \)). Consequently, each sequence \( s_1, \ldots, s_m \) which contributes a non-zero term to \( W(F) \) corresponds to at most \( p^{m'-m} \) sequences \( s_1, \ldots, s_m' \), contributing non-zero terms to \( W(H) \). It follows that

\[
W(H) \leq D! m' p^{m'-m} W(F),
\]
and so
\[ Y(H) = O_H(1) Y(F)(2q)^{n-n'} p^{m'-m}. \]

It is fairly easy to see that the truth of (ii) whenever \( d_i \neq 0 \) for all \( i \) implies the truth of (ii) in general. Hence, we may assume without loss of generality that \( d_i \geq 1 \) for all \( i \), so that \( p \leq 2q \), and we have
\[ Y(H) = O_H(1) Y(F)(2q)^{n-m+m'-n'}. \quad (4.7) \]

From (4.4) with \( v = 0 \),
\[ P(H) = NY(H)(1 + o(1)), \]
and so (4.6) and (4.7) imply
\[ P(H) = O_H(1) NY(F) q^{-1}. \]

Since both the factor \( f \) and the number of terms in the summation in (4.5) are bounded above, (4.5) now yields
\[ |P^s(F) - P^*(F)| = O_f(1) NY(F) q^{-1}. \]

Hence, by (4.4),
\[ P^s(F) = NY(F)(o(1) + X_k), \]
which is (ii) for \( v = k \).

It must now be shown that the truth of (i) and (ii) for \( v = k \) implies the truth of (i) for \( v = k + 1 \). So suppose \( 0 \leq v = k \), and \( l(1), \ldots, l(k) \) and \( t(1), \ldots, t(k) \) satisfy the hypotheses of the theorem. Let \( t \geq 0 \) and \( l \geq 1 \) with \( l \neq l(i) \) for \( 1 \leq i \leq k \). Consider a pseudograph \( G \) with degree sequence \( d_1, \ldots, d_p \) and with precisely \( t(i) \) different \( l(i) \)-cycles for \( 1 \leq i \leq k \). As in the proof of Theorem 3 we may assume without loss of generality that \( d_i \geq 1 \) for \( 1 \leq i \leq p \), so that \( p \leq 2q \). Let \( P^s \) denote the number of possibilities for \( G \) with precisely \( s \) \( l \)-cycles, and let \( P^*_s \) denote the number of possibilities for \( G \) with \( s \) different \( l \)-cycles distinguished. These distinguished \( l \)-cycles may overlap. By inclusion–exclusion we have for all \( u > t \)
\[ \left| P^s_t - \sum_{s=t}^{u-1} (-1)^{s-t} \binom{s}{t} P^*_s \right| \leq \binom{u}{t} P^*_u. \quad (4.8) \]

We wish to find an asymptotic formula for \( P^*_s \). The number \( P^s(sC_p) \) is just the number of possibilities for \( G \) with \( s \) different \( l \)-cycles distinguished such that no two of these \( l \)-cycles share a common point. Suppose on the other hand that \( s \) distinct \( l \)-cycles of \( G \) are distinguished such that at least two
share a common point. Then the distinguished $l$-cycles of $G$ form a subgraph $E$ with more lines than points. The number of ways of recovering the $s$ $l$-cycles from $E$ is bounded above by a number $f_1$ depending only on $l$ and $s$. It follows that the number of possibilities for $G$ in this case is at most $f_1 \sum P^*(E)$, where the sum is taken over all relevant possibilities for $E$. The number of terms in this sum is bounded above by a number $f_2$ which also depends only on $l$ and $s$. Consider any term $P^*(E)$ in the sum, where $E$ has $m'$ points and $n'$ lines. Since $W(E) \leq D_{m'} p^{m'}$, we have

$$Y(E) = O_E(1)(2q)^{-n'} p^{m'}$$

and so (ii) with $v = k$ implies

$$P^*(E) = O_E(1) N(2q)^{-n'} p^{m'} = O_E(1) N q^{-1}$$

as $p \leq 2q$ and $n' > m!$. Hence

$$f_1 \sum P^*(E) = o_s(1) N.$$

We therefore have

$$P^* = P^*(sC_l) + o_s(N).$$

(4.9)

By (ii) with $v = k$ we also have

$$P^*(sC_l) = N(o_s(1) + X_k) Y(sC_l).$$

(4.10)

Suppose the sequence $d_1, \ldots, d_p$ is restricted so that $\gamma \geq \sqrt{q}$. Then by Lemma 1

$$W(sC_l) = (1 + o_s(1))(4qa)^{sl}.$$  

Since $sC_l$ has $sl$ lines and since $A(sC_l) = s! A(C_t)^s$, it now follows that

$$Y(sC_l) = (1 + o_s(1)) y_{l/s!}^s.$$

From (4.9) and (4.10) and the fact that $X_k$ is bounded, this gives $P^* = (N(o_s(1) + X_k) y_{l/s!}^s) + o_s(N)$. Putting this in (4.8), we get

$$\left| P^*_t - \sum_{s=t}^{u-1} (-1)^{s-t} \binom{s}{t} N(o_s(1) + X_k) y_{l/s!}^s \right|$$

$$\leq \binom{u}{t} N(o_s(1) + X_k) y_{l/t}^u / u! + o_u(N).$$
for all fixed $u > t$. As $y_i$ is also bounded we can make $({\text{\textsc{u}}}^n)_{y_i^n/u!}$ as small as we like by choosing $u$ large enough, and it follows that

$$P_t^\# = N(o(1) + X_k e^{-ny_i^\#/t!}).$$

Defining $t(k+1)$ as $t$ and $l(k+1)$ as $l$, this gives (i) with $v = k + 1$, in case $\gamma > \sqrt{q}$.

Now suppose $d_1, ..., d_p$ is restricted so that $\gamma < \sqrt{q}$. Then by Lemma 2, $W(C_i) = o(p^3)$, so since $p \leq 2q$ we have $Y(C_i) = o(1)$, and hence also $P^\#(C_i) = o(N)$ by (ii). We also have $\alpha = o(1)$ and thus also $y_i = o(1)$. The number of pseudographs in any class which contain precisely $t$ $l$-cycles is at most the number of pseudographs in that class rooted at $C_i$, and so if $t > 0$

$$P_t^\#/N = O(1) Y(C_i) = o(1).$$

On the other hand, in case $t = 0$ note that $P_0^\#$ must be within $P^\#(C_i)$ of the total number of possibilities for $G$, which is $N(o(1) + X_v)$. In either case, (i) is implied for $v = k + 1$ on putting $t(k+1) = t$ and $l(k+1) = l$.

Combining the different cases produces (i) with $v = k + 1$, regardless of the value of $\gamma$. This completes the proof of the theorem.

Theorem 4 can be specialised to obtain various results. In particular, by putting $l(1) = 1$, $l(2) = 2$ and $t(1) = t(2) = 0$, we find that the word "pseudographs" in the theorem can be replaced by "graphs" if $N$ is replaced by $Ne^{-2\alpha - 2\alpha^3}$ and the extra restrictions are made that $l(j) \geq 3$ and that $F$ is a graph. If all the $l(j)$ are zero, or if the sequence $d_1, ..., d_p$ is restricted so that $\gamma/p$ is bounded below by a positive constant, an actual asymptotic formula results, in the latter case because $a$ is then bounded below by a positive constant. This yields the following result.

**Corollary 1.** Suppose $l(1), ..., l(v)$, $t(1), ..., t(v)$ are fixed non-negative integers with the $l(i)$ all different and at least 3, and let $C$ and $D$ be positive constants. Then if $d_i \leq D$ for all $i$ and $d_i \geq 2$ for at least $Cp$ values of $i$, the probability that a graph chosen at random from the graphs with degree sequence $d_1, ..., d_p$ has precisely $t(j)$ different $l(j)$-cycles for $1 \leq j \leq v$ is asymptotic to $X_v$ as $p \to \infty$.

If we wish to consider $r$-regular graphs, the requirement that $\gamma/p > C$ can be omitted since $\gamma = p$. For $v = 1$ the result simplifies to the following.

**Corollary 2.** Let $l \geq 3$ and $t \geq 0$. The probability that a graph chosen at random from the $r$-regular graphs on $p$ points, where $p$ is even if $r$ is odd, contains precisely $t$ $l$-cycles is asymptotic to

$$\frac{\lambda e^{-\Lambda}}{t!} \quad (4.11)$$
as \( p \to \infty \) with \( t \), \( d \) and \( r \) fixed, where

\[
\lambda = \frac{(r - 1)^j}{2l}.
\]

This may be compared with a result of Erdős and Rényi [4] which states that the probability that a random graph with \( p \) points and \( \frac{1}{2}rp \) lines contains precisely \( t \) \( l \)-cycles is again asymptotically of the form (4.11), but this time with

\[
\lambda = \frac{r^j}{2l}.
\]

Putting \( t(i) = 0 \) for all \( i \) and \( l(j) = j \) for \( j = 1, \ldots, v \) in Theorem 4(i) yields an asymptotic formula for graphs with given degree sequence and girth at least \( v + 1 \). Subtracting from this the corresponding formula for \( v + 2 \) yields an asymptotic formula for graphs with given degree sequence and girth. Here there is no restriction on \( \gamma \). For labelled graphs in general, not quite so much is known in relation to girth. However, Erdős et al. [3] showed that the number of graphs with no triangles is asymptotic to the number of bipartite graphs as the number of points increases.

Our result for \( r \)-regular graphs may be stated as follows.

**Corollary 3.** The probability that a random \( r \)-regular graph on \( p \) points, where \( p \) is even if \( r \) is odd, has girth at least \( j \geq 4 \) is asymptotic to

\[
\exp \left( -\sum_{i=3}^{j-1} \frac{(r - 1)^i}{2i} \right)
\]

as \( p \to \infty \) with \( r \) and \( j \) fixed.

Using Theorem 4(ii) it is possible to give results corresponding to Corollaries 1–3 for graphs rooted at a subgraph. One of the most interesting is for \( r \)-regular graphs.

**Corollary 4.** Let \( j \geq 3 \). The expected number of \( j \)-cycles in a random \( r \)-regular graph on \( p \) points, with \( p \) even if \( r \) is odd, is asymptotic to

\[
\frac{(r - 1)^j}{2j}
\]

as \( p \to \infty \).

With \( r \) and \( j \) equal to 3, this provides a different proof of the result obtained in [6] that the mean number of triangles in a cubic graph with \( 2s \) points approaches \( \frac{4}{3} \) as \( s \to \infty \). The proof given in [6] was based on a recurrence relation for the exact, rather than asymptotic, number of triangles in cubic graphs on \( 2s \) points.
REFERENCES


