Broken-Cycle-Free Subgraphs and the Log-Concavity Conjecture for Chromatic Polynomials

P. H. Lundow and K. Markström

CONTENTS

This paper concerns the coefficients of the chromatic polynomial of a graph. We first report on a computational verification of the strict log-concavity conjecture for chromatic polynomials for all graphs on at most 11 vertices, as well as for certain cubic graphs.

In the second part of the paper we give a number of conjectures and theorems regarding the behavior of the coefficients of the chromatic polynomial, in part motivated by our computations. Here our focus is on $\varepsilon(G)$, the average size of a broken-cycle-free subgraph of the graph G, whose behavior under edge deletion and contraction is studied.

1. LOG-CONCAVITY

In a paper from 1912 aimed at proving the four-color theorem, G. D. Birkhoff [Birkhoff 12] introduced a function $P(G, x)$, defined for all positive integers x to be the number of proper x-colorings of the graph G. As it turns out, $P(G, x)$ is a polynomial in x and so is defined for all real and complex values of x as well. Of course, $P(G, x)$ is the by now well-known chromatic polynomial, and although Birkhoff's original hope that it would help resolve the four-color conjecture did not bear fruit, it has attracted a steady stream of attention through the years.

Most of the investigations regarding the chromatic polynomial have focused on the location of its zeros. An early example is the work of Tutte on the chromatic roots of triangulations and the so-called golden identity, nicely described in [Tutte 98]. More recently we have the results of Thomassen on zero-free intervals of minor closed graph families [Thomassen 97] and the influence of Hamiltonian paths on the zeros of the chromatic polynomial [Thomassen 00]. There has also been a recent influx of ideas from statistical physics due to the connection to the Potts model. Using this connection, Sokal [Sokal 01] has shown that the moduli of the zeros

are bounded by a function linear in the maximum degree of the graph. Another recent development is the results of Biggs accumulation points for the zeros of sets of chromatic polynomials [Biggs 02]. For recent surveys of results and conjectures about the zeros of chromatic polynomials see [Jackson 02] and [Sokal 05].

Another line of work has focused on the coefficients of the chromatic polynomial. For a graph G on n vertices we can express $P(G, x)$ as

$$P(G, x) = \sum_{i=0}^{n} (-1)^{n-i} a_i x^i,$$

where a_i are nonnegative integers. There are a number of results giving bounds on the coefficients; for a good survey see [Read and Tutte 88]. In 1968, Read [Read 68] made the following conjecture.

Conjecture 1.1. (The unimodality conjecture.) *For any chromatic polynomial the following statement is false for all j:*

$$a_{j-1} > a_j \text{ and } a_j < a_{j+1}.$$

This basically means that at first the coefficients are increasing with j and then possibly decreasing. A polynomial with this property is said to have unimodal coefficients. The conjecture was later given a stronger form by Hoggar [Hoggar 74], who made the following conjecture.

Conjecture 1.2. (The strict log-concavity conjecture.) *For any chromatic polynomial and any j,*

$$a_{j-1} a_{j+1} < a_j^2.$$

A polynomial satisfying this inequality is said to be strictly logarithmically concave, or strictly log-concave for short. Log-concavity is a stronger property than unimodality in the sense that it implies unimodality as well. Log-concavity is also preserved under multiplication of polynomials, which ties in nicely with the fact that the chromatic polynomial of a disconnected graph is the product of the chromatic polynomials of its components.

To our knowledge there has been basically no progress on either of these two conjectures since they were first stated. The corresponding conjectures for other ways of writing the chromatic polynomials, surveyed in [Brenti 92], have been shown not to be strictly log-concave; see the references in [Read and Tutte 88]. Conjectures 1.1

and 1.2 were verified for all graphs on at most nine vertices during the 1980s [Read and Tutte 88], and now we can report the following computational result:

Fact 1.3. *Conjecture 1.2 holds for all graphs on $n \leq 11$ vertices. Conjecture 1.2 also holds for all graphs on 12 vertices that have fewer than 20 or more than 45 edges.*

Using some simple properties of the chromatic polynomial [Read and Tutte 88], one can see that the conjecture holds for all graphs if it holds for 2-connected graphs. We used Brendan McKay's graph generator geng [McKay 84] to generate all 2-connected graphs on at most 12 vertices and the number of edges stated; then we used a simple FORTRAN-90 implementation of the basic deletion–contraction algorithm to compute the chromatic polynomials and test them for log-concavity. To give a feeling for the size of this undertaking, note that there are 900,969,091 2-connected graphs on 11 vertices. The polynomials were computed and tested for concavity as the graphs were generated, so no graphs or polynomials were saved on disk. The computation was done on 48 Sun workstations, each working for eight months. The computation of the chromatic polynomials could certainly have been done faster by using a more advanced algorithm, but the increased complexity of the code would also have meant a larger risk of programming errors. A more advanced program could probably manage the 12-vertex graphs with current computers as well.

We also made a smaller test on cubic graphs:

Fact 1.4. *Conjecture 1.2 holds for all cubic graphs on $n \leq 20$ vertices. Conjecture 1.2 also holds for all cubic graphs on 22 vertices that have girth at least 5; 24 vertices and girth at least 6; and 26, 28, or 30 vertices with girth at least 7.*

Here we used a version of the program that deleted edges until a spanning tree was reached, thereby making it faster for sparse graphs. This computation was done on a Linux cluster with 2.8-GHz Pentium 4 processors. Each 30-vertex graph used about 23 CPU hours. The high-girth graphs are of special interest, since Read and Royle found counterexamples among them to the conjecture that chromatic polynomials have only roots with nonnegative real parts; see, for example, [Read and Royle 91].

One of the main problems in trying to test Conjectures 1.1 and 1.2 is the fact that chromatic polynomials are notoriously hard to compute. It is one of the classical

#P-complete problems. There is a small number of graph classes for which explicit expressions for the chromatic polynomial are known, and the conjectures are known to hold, for example, for trees, cycles, and wheels; see [Read and Tutte 88] for a few more examples. One further class, considered by Read, should be mentioned. A graph is called a *broken wheel* if it can be constructed by deleting a subset of the radial edges in a wheel. Read proved [Read 86] that broken wheels satisfy Conjecture 1.2. Apart from where explicit expressions are known, there are few large graphs for which chromatic polynomials are known and these two conjectures have been verified.

One class of large graphs can be obtained using the transfer matrix methods developed by Biggs, starting with [Biggs 01]. Using this method the chromatic polynomials of what Biggs calls *bracelets* can be computed. For large graphs in this class the chromatic polynomials can be written as a short sum of high powers of small polynomials. Since powers of polynomials tend to make the coefficients more and more log-concave, this class seems unlikely to produce counterexamples to Conjecture 1.2.

There have been found isolated large graphs for which the chromatic polynomial has been computed using symmetries to reduce the number of graphs in the recursions. Here Haggard stands out, especially [Haggard and Mathies 99], with the computation of the chromatic polynomial of the truncated icosahedron, or buckyball, with 60 vertices. The fact that the graph is both very sparse and has a large automorphism group was essential for the computation. For graphs of even moderate density we know of no example of comparable size. A good computational challenge, even with the use of symmetries, is given in the following problem.

Problem 1.5. Compute the chromatic polynomial of a regular self-complementary graph on 40 vertices.

There is one more class in which the chromatic polynomials can be computed easily. Given graphs G_0, G_1, G_2, we say that G_0 is a *k-clique sum* of G_1 and G_2 if G_0 can be constructed by identifying the vertices of a clique of size k in G_1 with a clique of size k in G_2. Note that there are many ways of forming a k-clique sum of two graphs. One classical class of graphs that can be constructed as clique sums are the *chordal graphs*, i.e., the graphs in which any cycle of length greater than 3 has a chord. By a theorem of Dirac [Dirac 61] these graphs can be built by repeatedly taking the clique sum of a smaller chordal graph and a complete graph. Another well-known graph class constructed this way is that of the *outerplanar graphs*, i.e.,

planar graphs that can be drawn such that the outer face is a Hamiltonian cycle. The outerplanar graphs can be constructed by repeatedly taking 2-clique sums of cycles. Given a graph G that is a k-clique sum of G_1 and G_2, we can express the chromatic polynomial as

$$P(G, x) = \frac{P(G_1, x)P(G_2, x)}{P(K_k, x)}; \qquad (1\text{--}1)$$

see [Read and Tutte 88]. Thus for graphs that can be constructed by repeated clique sums we can compute the chromatic polynomial quite easily, in fact in polynomial time. This has already been observed for chordal graphs [Read and Tutte 88], for which it also follows that the chromatic polynomials have only positive-integer roots.

What can we say about Conjecture 1.2 for graphs of this last kind? Let us say that the chromatic polynomial $P(G, x)$ of a graph G has a *good factoring* if it can be written as

$$P(G, x) = P(K_\omega, x)Q(G, x),$$

where ω is the clique number of G and $Q(G, x)$ is a polynomial with strictly log-concave coefficients. We now have the following easy lemma.

Lemma 1.6. *If both G_1 and G_2 in (1–1) have chromatic polynomials with log-concave coefficients and at least one of them has a good factoring, then $P(G, x)$ has strictly log-concave coefficients.*

This follows immediately from the fact (see, for example, [Karlin 68]), that products preserve log-concavity. From the formulas for the chromatic polynomials of trees, cycles, and complete graphs it is easy to see that they all have good factorings. A more surprising fact is the following, which we have found by direct computation using Mathematica.

Fact 1.7. *The chromatic polynomials of all graphs on $n \leq 9$ vertices have good factorings.*

Thus Conjecture 1.2 holds for all graphs that can be built by repeated clique sums using complete graphs, cycles, trees, and graphs with at most nine vertices. Since the property of having a good factoring is stronger than being strictly log-concave, it is natural to ask the following question.

Problem 1.8. Do all chromatic polynomials have a good factoring?

The chromatic polynomial has a generalization to matroids as well, the so-called characteristic polynomial of a matroid; see, for example, [Oxley 92]. For this, polynomial analogues of Conjectures 1.1 and 1.2 have been posited to hold for all matroids. The conjectures have been shown to hold for some classes of matroids, but none of these include the graphic matroids, which would imply Conjecture 1.2. For a survey of these matroid connections see [Aigner 87].

2.   SUBGRAPHS WITHOUT BROKEN CYCLES

There are several different expansions for the chromatic polynomial of a graph in terms of its subgraphs; see [Biggs 93]. In 1932, Whitney [Whitney 32] gave the following characterization. Assume that the edges of a graph G have been labeled with the integers $1, \ldots, m$, where $m = |E(G)|$, in an arbitrary way. A path obtained from a cycle in G by removing the edge with the largest label among those in the cycle is called a broken cycle.

Theorem 2.1. [Whitney 32] *The coefficient a_i equals the number of spanning subgraphs of G with $n - i$ edges that do not contain a broken cycle.*

Here a subgraph is specified by its edge set. Note that the theorem implies that the number of broken-cycle-free subgraphs is independent of the labeling of the graph. So in light of Whitney's theorem and the deletion–contraction formulas for the chromatic polynomial, we see that Conjecture 1.2 really concerns how the number of broken-cycle-free subgraphs changes under edge deletion and contraction.

A subgraph that does not contain a broken cycle obviously cannot contain a cycle, and so must be a forest. This also implies that $a_0 = 0$ and $a_n = 1$. So apart from the alternating sign, the chromatic polynomial is the generating function for the broken-cycle-free subgraphs of G. In connection with our test of the log-concavity conjecture, we also made some further investigations into the behavior of the coefficients of chromatic polynomials for small graphs, and we will now discuss some of them and state a few observations and problems for future work. Let us first define

$$b_i = \frac{a_{n-i}}{\sum_j a_j}, \quad i = 0, \ldots, n-1.$$

The number b_i can be interpreted as the probability that a uniformly chosen broken-cycle-free subgraph has size i.
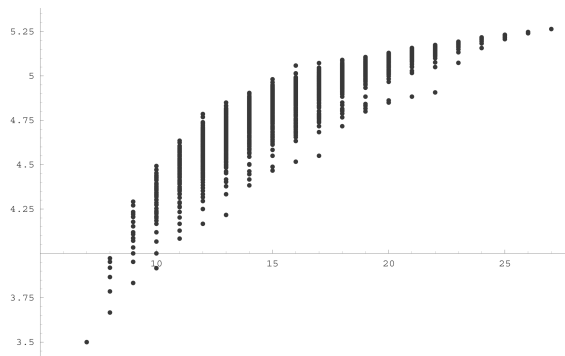

FIGURE 1. $\varepsilon(G)$ plotted for all connected graphs on eight vertices. The horizontal coordinate shows the number of edges in the graphs.

We say that two sequences α_i and β_i are *coconcave* if $\alpha_i + \beta_i$ is a log-concave sequence. Due to the deletion–contraction formulas for the chromatic polynomial, coconcavity is a key property for understanding the structure behind the log-concavity conjecture. If the two chromatic polynomials in the formulas could be shown to be coconcave, Conjecture 1.2 would follow. Here we would like to state the following conjecture.

Conjecture 2.2. *Let b_i be defined as before for a connected graph G of order n and let p_i be the probabilities of the binomial distribution on $n - 1$ events with expectation $\varepsilon(G)$. Then b_i and p_i are coconcave.*

We have verified the conjecture for all graphs on at most nine vertices.

Given a graph G, we can now calculate the mean size of a broken-cycle-free subgraph of G. Let us denote this size by $\varepsilon(G)$, that is, $\varepsilon(G) = \sum_i i \, b_i$. Let us look at two simple examples.

Example 2.3. The chromatic polynomial of a tree T on n vertices is just $x(x-1)^{n-1}$, and so the b_i's will equal the probabilities of the binomial distribution for $n - 1$ events with $p = \frac{1}{2}$ and mean $\frac{n-1}{2}$.

The chromatic polynomial of K_n is $\prod_{i=0}^{n-1}(x-i)$. Here we find that $b_i = \begin{bmatrix} n \\ i \end{bmatrix}/n!$, where $\begin{bmatrix} n \\ j \end{bmatrix}$ are the Stirling numbers of the first kind. The mean size of a broken-cycle-free subgraph here is $n - \sum_{i=1}^{n-1} i^{-1}$, and the b_i's converge to a Poisson distribution with mean $\varepsilon(K_n)$ [Moser and Wyman 58].

In Figure 1 we have plotted $\varepsilon(G)$ for all connected graphs on eight vertices. At the bottom left we find all the trees on eight vertices, all at the same point, and at the top right we find K_8. From our test on small graphs
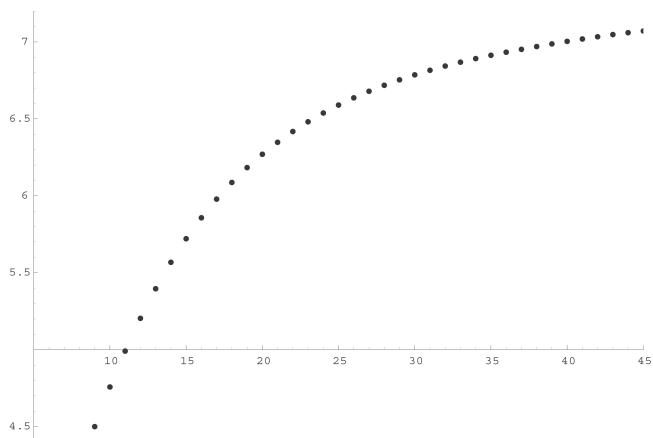
FIGURE 2. $\varepsilon\,(10,k)$.

we would like to pose a few problems and conjectures on the behavior of $\varepsilon\,(G)$.

Conjecture 2.4. *Let G be a connected graph on n vertices that is not complete or a tree. Then*

$$\varepsilon\,(P_n) < \varepsilon\,(G) < \varepsilon\,(K_n),$$

where P_n is the path on n vertices.

Of course P_n could be replaced by any tree on n vertices.

Problem 2.5. Given n and k, what are the maximum and minimum of $\varepsilon\,(G)$ among all connected graphs with n vertices and k edges?

In Figure 2 we have plotted the mean value of $\varepsilon\,(G)$ among the connected graphs on 10 vertices and k edges as a function of k. Let us denote the corresponding mean for a general n by $\varepsilon\,(n,k)$. We immediately know the values of $\varepsilon\,(n, n-1)$ and $\varepsilon\left(n, \binom{n}{2}\right)$ (see Example 2.3), and for a few k very close to $n-1$ and $\binom{n}{2}$ the value of $\varepsilon(n,k)$ can be calculated as well.

Problem 2.6. What is the asymptotic behavior of $\varepsilon\,(n,k)$ for large n?

3.   SOME RESULTS ON THE BEHAVIOR OF $\varepsilon\,(G)$

Apart from their inherent value, the interest in the problems and conjectures of the previous section really stems from the concept of coconcavity. The chromatic polynomial of a graph can be expressed in terms of chromatic polynomials of smaller graphs using the deletion–contraction formula

$$P\,(G,x) = P\,(G-e,x) - P\,(G/e,x)\,,$$

where $G-e$ denotes the graph obtained by removing the edge e from G, and G/e the graph obtained by contracting e. So if the coefficients of $P\,(G-e,x)$ and $-P\,(G/e,x)$ could be shown to be coconcave, the log-concavity conjecture would follow.

As a first step in this direction we would like to find out more about how both the sum of the a_i's and $\varepsilon\,(G)$ change when we form subgraphs in the above way. If these quantities were not well behaved, that would clearly reduce the chance of the involved polynomials having coconcave coefficients. However, as we shall see, there seems to be some nice structure to their behavior.

Let us first note that

$$\varepsilon\,(G) = n + \frac{P'\,(G,-1)}{P\,(G,-1)}. \qquad (3\text{–}1)$$

This makes the following definitions convenient:

$$\eta(G) = |P\,(G,-1)|\,,$$
$$\eta'(G) = |P'\,(G,-1)|\,.$$

As noted by Stanley [Stanley 73], $\eta(G)$ can also be interpreted as the number of acyclic orientations of the graph.

Let n_e denote the number of broken-cycle-free subgraphs of G containing the edge e, let n'_e denote the number of broken-cycle-free subgraphs of G *not* containing the edge e, and let n''_e be the number of broken-cycle-free subgraphs H of $G-e$ such that H considered as a subgraph of G contains a broken cycle.

Proposition 3.1. *Let G be a labeled connected graph on n vertices and e the edge with the highest label in G. Then*

(i) $n_e = n'_e$;

(ii) $\eta\,(G) = n_e + n'_e = 2n_e = \eta\,(G-e) + \eta\,(G/e)$;

(iii) $\eta\,(G) > \eta\,(G-e)$;

(iv) $\eta\,(G-e) = \eta\,(G) - n_e + n''_e = \frac{1}{2}\eta(G) + n''_e = n_e + n''_e$;

(v) $\eta\,(G/e) = \frac{1}{2}\eta(G) - n''_e = n_e - n''_e$;

(vi) $2^{n-2} \le n_e$, *and the bound is sharp if G is a tree.*

Proof: (i) Let us assume that the edges in G have been labeled and that e is the edge with the highest label. Now, $n'_e \geq n_e$, since from each graph counted by n_e we can obtain a unique graph counted by n'_e by removing e. We also find that $n_e \geq n'_e$, since if H is counted by n'_e and $H \cup e$ contains a broken cycle, then either H contains a broken cycle or $H \cup e$ contains a broken cycle for which the missing edge has a higher label than e. In both cases we have a contradiction.

(ii) The first equality follows directly from the definition of $\eta(G)$; the second equality follows from (i); and the last equality follows from the deletion–contraction formula together with (3–1):

$$\eta(G) = |P(G, -1)| = |P(G - e, -1) - P(G/e, -1)|$$
$$= \eta(G - e) + \eta(G/e).$$

The last equality holds thanks to the minus sign in the deletion–contraction formula, which ensures that the two terms have the same sign.

(iii) Assertion (iii) follows from (ii).

(iv) Let H be a subgraph counted by n_e. Then both H and $H - e$ will be broken-cycle-free subgraphs of G. However, none of the graphs counted by n_e are subgraphs of $G - e$, and so they should be subtracted when we count broken-cycle-free subgraphs of $G - e$. If H is a subgraph counted by n''_e, then H will be a broken-cycle-free subgraph of $G - e$ not counted by n_e and n'_e, and so should be added to the number of broken-cycle-free subgraphs of $G - e$. The rest follows from (i) and (ii).

(v) Assertion (v) follows from (iv) and (ii).

(vi) Every broken-cycle-free subgraph of G/e can be expanded into at least one subgraph of G counted by n_e. By (ii) we thus obtain that n_e will be at least $\frac{1}{2}\eta(T) = 2^{n-2}$, where T is a tree on $n - 1$ vertices. $\square$

Here we see that if we know how $\eta(G - e)$ relates to $\eta(G)$, we will also know what happens to $\eta(G/e)$.

Property (ii) in the proposition is nice, and together with similar reasoning for $\eta'(G)$ one might be tempted to make the following conjecture. Let $G_1 \prec G_2$ mean that G_1 can be obtained from G_2 by deleting edges. Together with this partial order the set of graphs on n vertices forms a lattice $G(n)$.

False Conjecture 3.2. *The function $\varepsilon(G)$ is increasing on chains in the lattice $G(n)$.*

However, as the alert reader might already have suspected, the conjecture is not true.

Example 3.3. A counterexample can be constructed from $K_{2,n}$, $n \geq 4$, by adding an edge e with endpoints in the smaller part of the bipartition. Let $G = K_{2,4} \cup e$. Then

$$P(G, x) = -16x + 48x^2 - 56x^3 + 32x^4 - 9x^5 + x^6,$$
$$P(K_{2,4}, x) = -15x + 44x^2 - 50x^3 + 28x^4 - 8x^5 + x^6,$$

giving us $\varepsilon(G) = \frac{13}{6} = 2.166\ldots$ and $\varepsilon(K_{2,4}) = \frac{319}{146} = 2.18\ldots$.

The critical property of the graphs in the example is that there exists one edge that is used by a very large number of short cycles and at the same time is a chord of all the longer cycles in the graph.

The rather weak Conjecture 2.4 is just a hint of what should be true. Experiment shows that most transitions in the lattice behave as the false conjecture claims. In fact, every graph on at most eight vertices contains many edges such that $\varepsilon(G - e) < \varepsilon(G)$. Due to the counterexamples, it is not clear what the right conjecture should be here. We can, however, prove the following result.

Theorem 3.4. *Let G be the union of two subgraphs G_1 and G_2 such that their intersection is a K_k. Then*

$$\varepsilon(G) = \varepsilon(G_1) + \varepsilon(G_2) - \varepsilon(K_k).$$

Proof: We first recall [Read and Tutte 88] that the chromatic polynomial of G can be expressed as

$$P(G, x) = \frac{P(G_1, x)P(G_2, x)}{P(K_k, x)}. \qquad (3\text{–}2)$$

If we now differentiate this and use the identity (3–1), we obtain

$$\varepsilon(G) = n + \frac{P'(G_1, -1)}{P(G_1, -1)} + \frac{P'(G_2, -1)}{P(G_2, -1)} - \frac{P'(K_k, -1)}{P(K_k, -1)},$$

as claimed. $\square$

Corollary 3.5. *If G is the union of two graphs G_1 and G_2 intersecting in at most one vertex, then*

$$\varepsilon(G) = \varepsilon(G_1) + \varepsilon(G_2).$$

The corollary implies that $\varepsilon(G)$ is determined by the blocks of G.

Corollary 3.6. *If G has a cut-edge e and the two components of $G - e$ are G_1 and G_2, then*

$$\varepsilon(G) = \varepsilon(G_1) + \varepsilon(G_2) + \frac{1}{2}.$$

Using Theorem 3.4 we can, for example, quickly compute $\varepsilon(K_n - e)$.

Example 3.7. Let e be any edge of K_n; $K_n - e$ is the union of two K_{n-1}'s, the intersection of which is a K_{n-2}. Using Theorem 3.4 and the value of $\varepsilon(K_n)$ from Example 2.3, we find that

$$\varepsilon(K_n - e) = 2\varepsilon(K_{n-1}) - \varepsilon(K_{n-2}) = \varepsilon(K_n) - \frac{1}{n(n-1)}.$$

In view of the two corollaries, it makes sense to study Conjecture 2.4 on the subset $G_c(n)$ of all connected graphs in $G(n)$. The poset $G_c(n)$ is not a lattice, since all trees are minimal elements, but all the minimal elements of $G_c(n)$ give the same value of $\varepsilon(G)$. So in order to prove Conjecture 2.4 one could try to prove that any graph in $G_c(n)$ belongs to a chain in which the statement of the false conjecture holds. This in turn reduces to showing that any 2-connected graph contains an edge e such that $\varepsilon(G - e) < \varepsilon(G)$, which we state as a conjecture.

Conjecture 3.8. *Every* 2-*connected graph* G *contains at least one edge* e *such that* $\varepsilon(G - e) < \varepsilon(G)$.

We have verified this property for all graphs on at most eight vertices.

From Theorem 3.4 we can obtain the following positive, but somewhat narrow, result.

Proposition 3.9. *Let* v *be a vertex of* G *such that the neighbors of* v *induce a* K_k *and let* e *be an edge incident with* v. *Then*

$$\varepsilon(G - e) = \varepsilon(G) - \frac{1}{k(k-1)}.$$

Proof: First split G into a K_{k+1} containing v and a new graph $G' = G - v$. Their intersection is a neighborhood of v that is a K_k. By Theorem 3.4 we have

$$\varepsilon(G) = \varepsilon(K_{k+1}) + \varepsilon(G') - \varepsilon(K_k)$$

and

$$\varepsilon(G - e) = \varepsilon(K_{k+1} - e) + \varepsilon(G') - \varepsilon(K_k),$$

whose difference, by Example 3.7, is $-\left(k(k-1)\right)^{-1}$. $\square$

This implies that our graph family from Example 3.3, which did not satisfy the false conjecture, does satisfy Conjecture 3.8.

We can also ask how much $\varepsilon(G)$ can be changed by removing an edge.

Theorem 3.10. *We have*

$$\varepsilon(G - e) \geq \varepsilon(G) - \frac{1}{2}.$$

Equality holds if e *is a cut-edge.*

Proof: Let n_e, n'_e, and n''_e be as before and recall that by Proposition 3.1, $n_e = \frac{1}{2}\eta(G)$. We now find that $\varepsilon(G - e)$ will be a linear combination $pl_1 + (1 - p)l_2$, $0 \leq p \leq 1$. Here l_1 is the average size of a broken-cycle-free subgraph of G counted by n'_e, and l_2 is the average size of the graphs counted by n''_e. The value of p depends on the relative sizes of n'_e and n''_e and is 1 when $n''_e = 0$.

By the same reasoning as in the proof of Proposition 3.1, we see that $l_1 = \varepsilon(G) - \frac{1}{2}$, since for every graph counted by n_e there is a graph with one edge fewer contributing to the average and n_e is $\frac{1}{2}\eta(G)$.

Similarly, we see that l_2 will be at least as large as the average size of the graphs counted by n_e, since each graph counted by n'_e has been obtained by adding at least one edge to a graph counted by n_e and then removing e.

Thus we see that $\varepsilon(G - e) \geq \varepsilon(G) - \frac{1}{2}$, and equality will hold only if $n'_e = 0$, which means that there are no cycles through e and so e is a cut-edge. $\square$

In a typical situation one would expect both l_1 and l_2 to contribute to the average and so make $\varepsilon(G - e)$ stay much closer to $\varepsilon(G)$. In fact, if Conjecture 2.4 is true, the function $\varepsilon(G)$ must increase by an amount of about $\frac{n}{2} - \log n$ along a maximal chain, corresponding to $\binom{n}{2} - (n - 1)$ edges, thus giving us an average increase in $\varepsilon(G)$ of

$$\frac{\frac{n}{2} - \log n}{\binom{n}{2} - (n - 1)} = \mathcal{O}(n^{-1}).$$

Theorem 3.10 also tells us that for graphs taken uniformly at random from $G(n, m)$, the set of all graphs on n vertices and m edges, the value of $\varepsilon(G)$ will be well concentrated.

Theorem 3.11. *Let* $X(n, m)$ *be the random variable given by* $\varepsilon(G)$ *when* G *is a connected graph taken uniformly at random from* $G(n, m)$. *Then*

$$Pr\left(|X(n, m) - \varepsilon(n, m)| \geq t\right) \leq 2e^{-8t^2/T},$$

where $T = \min\left(m - (n - 1), \binom{n}{2} - m\right)$.

Proof: When $m-(n-1) \leq \binom{n}{2}-m$, the result follows from Azuma's inequality (Theorem 3.12) together with Theorem 3.10 by considering the value of $\varepsilon(G)$ on a graph constructed by adding m random edges to an empty graph. We get a denominator of only $T = m - (n - 1)$ in the exponent, since the first $n-1$ edges can be taken to form a spanning tree in G, thereby giving no variation in ε.

For the dense case we can instead consider the graph as constructed by removing $\binom{n}{2} - m$ random edges from K_n.
□

Theorem 3.12. (Azuma's inequality.) [Azuma 67] *Let* $X_1, \ldots, X_n$ *be a martingale sequence such that* $|X_i - X_{i-1}| \leq c_i$. *Then*

$$\Pr\left(|X_n - \mathrm{E}(X_n)| > t\right) < 2e^{-t^2/\left(2\sum c_i^2\right)}.$$

We thus find that the value of $\varepsilon(G)$ should be well concentrated for sparse and very dense graphs, and possibly less so for graphs of intermediate densities. This agrees well with the observed behavior in Figure 1.

4. SOME HEURISTIC BOUNDS FOR PROBLEM 2.6

Problem 2.6 essentially boils down to finding the expectation of $\varepsilon(G)$ when G is drawn from $G(n, m)$, the set of all graphs on n vertices and m edges. We have not been able to solve this problem, but we can say something about the expectation of $\varepsilon(G)$ in $G(n, p)$, the graphs on n vertices with edges drawn independently with probability p. Let us denote the two expectations mentioned by $\varepsilon(n, m)$ and $\varepsilon(n, p)$.

Let $P(p, x)$ denote the expectation of $P(G, x)$ in $G(n, p)$. Grimmett [Grimmett 77] has found the following generating function for $P(p, x)$:

$$\sum_{n=0}^{\infty} P_n(p, x)\frac{t^n}{n!} = F(p, t)^x$$

with

$$F(p, t) = \sum_{n=0}^{\infty}(1-p)^{\binom{n}{2}}\frac{t^n}{n!}.$$

What we would like now to calculate is

$$\mathrm{E}\left(\frac{P'(G, x)}{P(G, x)}\right),$$

which we have not been able to do, but from the generating function above we can calculate

$$\frac{\mathrm{E}\left(P'(G, x)\right)}{\mathrm{E}\left(P(G, x)\right)}$$

for moderate values of n. Let us do so and at the same time say something about why we believe it to be a good approximation of the proper expectation. Let us introduce the following notation in order to simplify our writing:

$$\eta(p) = |P(p, -1)|,$$
$$\eta'(p) = |P'(p, -1)|.$$

If $\eta(G)$ and $\eta'(G)$ had been independent random variables, we would have had

$$\mathrm{E}\left(\frac{\eta'(G)}{\eta(G)}\right) = \mathrm{E}\left(\eta'(G)\right)\mathrm{E}\left(\frac{1}{\eta(G)}\right).$$

Next we can create a Taylor expansion of the distribution of $\frac{1}{\eta(G)}$ around $x = \frac{1}{\eta(p)}$, and we see that if $\eta(p)$ is reasonably large and $\eta(G)$ does not have too large variance, then

$$\mathrm{E}\left(\frac{1}{\eta(G)}\right) \approx \frac{1}{\eta(p)}.$$

So what can be said about $\eta(p)$ and the variance of $\eta(G)$? There are general upper and lower bounds for $\eta(G)$ in terms of the degree sequence of the graph G that come in handy, namely

$$\prod_{v \in V(G)} f(d_v + 1) \leq \eta(G) \leq \prod_{v \in V(G)} (d_v + 1),$$

where d_v is the degree of the vertex v and $f(x) = (x!)^{1/x}$. The lower bound is from [Goddard et al. 93] and the upper from [Graham et al. 80]. Further bounds can be found in [Kahale and Schulman 96]. Since the degree sequence of a graph from $G(n, p)$ is quite well concentrated [Bollobás 01a], $\eta(G)$ will be slightly less than $(pn)^n$ and does not have too large a variance,[1] and so our estimate for $\mathrm{E}\left(\frac{1}{\eta(G)}\right)$ will be quite good. Thus we should find that

$$\mathrm{E}\left(\frac{\eta'(G)}{\eta(G)}\right) \sim \frac{\mathrm{E}\left(\eta'(G)\right)}{\mathrm{E}\left(\eta(G)\right)},$$

where $\sim$ means that they are comparable for large n.

All of this is valid under the assumption that $\eta(G)$ and $\eta'(G)$ are independent, which of course is false. However, unless $\eta(G)$ and $\eta'(G)$ are strongly anticorrelated, the approximation should be reasonably good.

In Figure 3 we have plotted the exact values for $\varepsilon(10, k)$ together with $\eta'(10, p)/\eta(10, p)$. As can be seen in the figure, we have good agreement for large values of p, but for smaller values the curves grow apart. This

---

[1]In fact, one would expect $\eta(p)/(pn)^n$ to have a log-normal distribution.
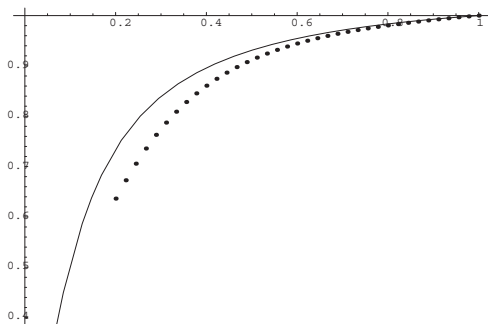
FIGURE 3. $\varepsilon(10, k)$ together with $\frac{\eta'(10,p)}{\eta(10,p)}$. Both axes have been rescaled to run between 0 and 1.
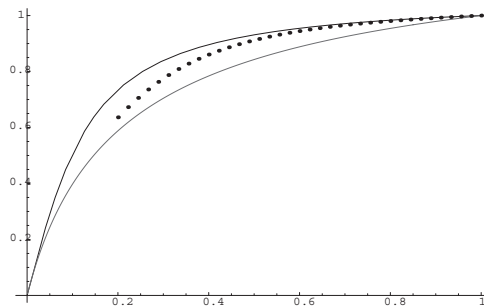

FIGURE 5. $\varepsilon(10, k)$ together with $\frac{\eta'(10,p)}{\eta(10,p)}$ and $\frac{S'_{10}(p,-1)}{S_{10}(p,-1)}$. Both axes have been rescaled to run between 0 and 1.
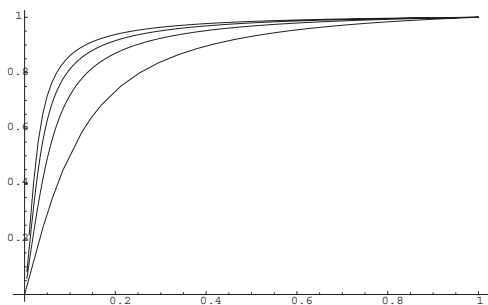

FIGURE 4. $\frac{\eta'(n,p)}{\eta(n,p)}$ for $n = 10, 20, 30, 40$ ($n = 10$ is the lowest curve). Both axes have been rescaled to run between 0 and 1.
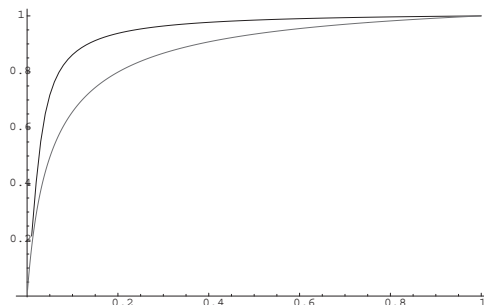

FIGURE 6. $\frac{\eta'(40,p)}{\eta(40,p)}$ and $\frac{S'_{40}(p,-1)}{S_{40}(p,-1)}$. Both axes have been rescaled to run between 0 and 1.

is not unexpected, since the exact values use only connected graphs, and for small p our approximation uses a large number of disconnected graphs, thus giving an overestimate for $\varepsilon(n,p)$. However, for a fixed $p > 0$ we expect to get better and better agreement between the curves as n increases and the proportion of disconnected graphs diminishes. Here we can also note that for $p = 1$ our estimate is actually exact,

$$\eta'(1) = \pm\left(n - \sum_{i=1}^{n-1} i^{-1}\right), \quad \eta(1) = \pm1,$$

and for $p = 0$ we also get the correct value $\varepsilon(n,0) = 0$.

In Figure 4 we have plotted our estimate for $n = 10, 20, 30, 40$, each estimate divided by its value at $p = 1$ in order to make them comparable.

Following the reasoning behind Proposition 3.10, we can also strive for a lower bound on $\varepsilon(n,p)$. Given a broken-cycle-free subgraph of K_n, the probability that it is also a subgraph of a graph from $G(n,p)$ is simply p^i, where i is the number of edges in the subgraph. Using the formula for $P(K_n, x)$, we find that the average generating function for these subgraphs is

$$S_n(p,x) = \sum_{i=0}^{n} \begin{bmatrix} n \\ n-i \end{bmatrix} p^i x^i = (px)^n (pn)^{\overline{n-1}},$$

where $x^{\bar{n}} = x(x+1)\cdots(x+(n-1))$. The subgraphs counted by this generating function are expected to be on average smaller than those present in a graph from $G(n,p)$, simply because many of the latter can be made larger by adding edges that would have created broken cycles in K_n. So we expect to get a lower bound for $\varepsilon(n,p)$ if we calculate $S'_n(p,-1)/S_n(p,-1)$. This can be done as follows:

$$\frac{S'_n(p,-1)}{S_n(p,-1)} = \frac{d}{dx}\log\left((px)^n (pn)^{\overline{n-1}}\right)\Big|_{x=-1}$$

$$= n - \sum_{i=0}^{n-1} \frac{1}{1+ip}.$$

We see that for $p = 1$ the bound coincides with the exact value. For $p = (n/2)^{-1}$, corresponding to trees, we get a value that is slightly lower than the exact $(n-1)/2$. Further, we see that for a fixed $p > 0$ the bound will be of the form $n - \mathcal{O}(\log n)$.

In Figure 5 we have plotted $\varepsilon(10, k)$, our previous upper bound, and our lower bound. In Figure 6 we have plotted both bounds for $\varepsilon(40, p)$.

REFERENCES

[Aigner 87] Martin Aigner. "Whitney Numbers." In *Combinatorial Geometries*, pp. 139–160. Cambridge: Cambridge Univ. Press, 1987.

[Azuma 67] Kazuoki Azuma. "Weighted Sums of Certain Dependent Random Variables." *Tôhoku Math. J. (2)* 19 (1967), 357–367.

[Biggs 93] Norman Biggs. *Algebraic Graph Theory*, second edition. Cambridge: Cambridge University Press, 1993.

[Biggs 01] Norman Biggs. "A Matrix Method for Chromatic Polynomials." *J. Combin. Theory Ser. B* 82:1 (2001), 19–29.

[Biggs 02] Norman Biggs. "Equimodular Curves." *Discrete Math.* 259:1-3 (2002), 37–57.

[Birkhoff 12] G. D. Birkhoff. "A Determinantal Formula for the Number of Ways of Colouring a Map." *Ann. of Math.* 14 (1912), 42–46.

[Bollobás 01a] Béla Bollobás. *Random Graphs*, second edition. Cambridge: Cambridge University Press, 2001.

[Brenti 92] Francesco Brenti. "Expansions of Chromatic Polynomials and Log-Concavity." *Trans. Amer. Math. Soc.* 332:2 (1992), 729–756.

[Dirac 61] G. A. Dirac. "On Rigid Circuit Graphs." *Abh. Math. Sem. Univ. Hamburg* 25 (1961), 71–76.

[Goddard et al. 93] Wayne Goddard, Claire Kenyon, Valerie King, and Leonard J. Schulman. "Optimal Randomized Algorithms for Local Sorting and Set-Maxima." *SIAM J. Comput.* 22:2 (1993), 272–283.

[Graham et al. 80] Ronald L. Graham, Andrew C. Yao, and F. Frances Yao. "Information Bounds Are Weak in the Shortest Distance Problem." *J. ACM* 27:3 (1980), 428–444.

[Grimmett 77] Geoffrey Grimmett. "Random Graph Theorems." In *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the Eighth European Meeting of Statisticians (Tech. Univ. Prague, Prague, 1974), vol. A*, pp. 203–209. Dordrecht: Reidel, 1977.

[Haggard and Mathies 99] Gary Haggard and Thomas R Mathies. "The Computation of Chromatic Polynomials." *Discrete Math.* 199:1-3 (1999), 227–231.

[Hajnal and Surányi 58] András Hajnal and János Surányi. "Über die Auflösung von Graphen in vollständige Teilgraphen." *Ann. Univ. Sci. Budapest. Eötvös. Sect. Math.* 1 (1958), 113–121.

[Hoggar 74] S. G. Hoggar. "Chromatic Polynomials and Logarithmic Concavity." *J. Combinatorial Theory Ser. B* 16 (1974), 248–254.

[Jackson 02] Bill Jackson. "Zeros of Chromatic and Flow Polynomials of Graphs." Available online (http://xxx. lanl.gov/abs/math.CO/0205047), 2002.

[Johnson and Goldschmidt 05] Oliver Johnson and Christina Goldschmidt. "Preservation of Log-Concavity on Summation." arXiv.org:math/0502548, 2005.

[Kahale and Schulman 96] Nabil Kahale and Leonard J. Schulman. "Bounds on the Chromatic Polynomial and on the Number of Acyclic Orientations of a Graph." *Combinatorica* 16:3 (1996), 383–397.

[Karlin 68] Samuel Karlin. *Total Positivity*, vol. I. Stanford: Stanford University Press, 1968.

[McKay 84] Brendan McKay. "Geng." Available online (http: //cs.anu.edu.au/people/bdm), 1984.

[Moser and Wyman 58] L. Moser and M. Wyman. "Asymptotic Development of the Stirling Numbers of the First Kind." *J. London Math. Soc.* 33 (1958), 133–146.

[Nijenhuis and Wilf 78] Albert Nijenhuis and Herbert S. Wilf. *Combinatorial Algorithms*, second edition. New York: Academic Press, 1978.

[Oxley 92] James G. Oxley. *Matroid Theory*. New York: Oxford University Press, 1992.

[Read 68] Ronald C. Read. "An Introduction to Chromatic Polynomials." *J. Combinatorial Theory* 4 (1968), 52–71.

[Read 86] Ronald C. Read. "Broken Wheels Are SLC." *Ars Combin.* 21 (1986) 123–128.

[Read and Royle 91] Ronald C. Read and Gordon F. Royle. "Chromatic Roots of Families of Graphs." In *Graph Theory, Combinatorics, and Applications*, vol. 2, pp. 1009–1029. New York: Wiley, 1991.

[Read and Tutte 88] R. C. Read and W. T. Tutte. "Chromatic Polynomials." In *Selected Topics in Graph Theory, 3*, pp. 15–42. San Diego: Academic Press, 1988.

[Sokal 01] Alan D. Sokal. "Bounds on the Complex Zeros of (Di)chromatic Polynomials and Potts-Model Partition Functions." *Combin. Probab. Comput.* 10:1 (2001), 41–77.

[Sokal 05] Alan D. Sokal. "The Multivariate Tutte Polynomial (Alias Potts Model) for Graphs and Matroids." In *Surveys in Combinatorics 2005*, pp. 173–227. Cambridge: Cambridge University Press, 2005.

[Stanley 73] Richard P. Stanley. "Acyclic Orientations of Graphs." *Discrete Math.* 5 (1973), 171–178.

[Thomassen 97] Carsten Thomassen. "The Zero-Free Intervals for Chromatic Polynomials of Graphs." *Combin. Probab. Comput.* 6:4 (1997), 497–506.

[Thomassen 00] C. Thomassen. "Chromatic Roots and Hamiltonian Paths." *J. Combin. Theory Ser. B* 80:2 (2000), 218–224.

[Tutte 98] W. T. Tutte. *Graph Theory as I Have Known It.* New York: Oxford University Press, 1998.

[Whitney 32] Hassler Whitney. "A Logical Expansion in Mathematics." *Bull. Amer. Math. Soc.* 38 (1932), 572–579.

P. H. Lundow, KTH Physics, AlbaNova University Center, SE-106 91, Stockholm, Sweden
        (phl@kth.se)

K. Markström, Department of Mathematics and Mathematical Statistics, Umeå University, SE-901 87,
        Umeå, Sweden (klas.markstrom@math.umu.se)