Broken cycle free subgraphs and the log-concavity conjecture for chromatic polynomials

Per Håkan Lundow, Klas Markström

Abstract. 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 behaviour 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 circuit free subgraph of the graph G, whose behaviour under edge deletion and contraction i studied.

1. Log-concavity

In a paper from 1912, [Bir12], aimed at proving the four-colour theorem, G. D. Birkhoff introduced a function $P(G,x)$, defined for all positive integers x to be the number of proper x-colourings 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. $P(G,x)$ is of course the by now well known chromatic polynomial and although Birkhoff's original hope that it would help resolve the four colour 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 [Tut98]. More recently we have the results of Thomassen on zero-free intervals of minor closed graph families [Tho97] and the influence of hamiltonian paths on the zeros of the chromatic polynomial [Tho00]. There has also been a recent influx of ideas from statistical physics due to the connection to the Potts model. Using this connection Sokal [Sok01] 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 [Big02] accumulation points for the zeros of sets of chromatic polynomials. For recent surveys of results and conjectures about the zeros of chromatic polynomials see [Jac02] and [Sok05].

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 is a number of results giving bounds on the coefficients, for a good survey see [RT88]. In 1968 Read [Rea68] 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 [Hog74] who conjectured,

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 have 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 [Bre92], have been shown not to be strictly log-concave, see references in [RT88]. Conjectures 1.1 and 1.2 were verified for all graphs on at most 9 vertices during the 1980's [RT88] 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 which have less than 20, or more than 45, edges.

Using some simple properties of the chromatic polynomial [RT88] 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 [McK] 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 900969091 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 disc. The computation was done on 48 SUN workstations, each working for 8 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 which 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 which deleted edges until a spanning tree was reached, thereby making it faster for sparse graphs. This computation was done on a linux cluster with 2.8GHz pentium4 processors. Each 30 vertex graph used about 23 CPU hours. The high girth graphs are of special interest since Read and Royle among them found counterexamples to the conjecture that chromatic polynomials have only roots with non-negative real parts, see e.g [RR91].

One of the main problem when 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 are a small number of graph classes for which explicit expressions for the chromatic polynomial is known, and the conjectures are known to hold, e.g. trees, cycles, wheels, see [RT88] 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. In [Rea86] Read proved 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 [Big01]. 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 has been isolated large graphs for which the chromatic polynomial has been computed by using symmetries to reduce the number of graphs in the recursions. Here Haggard stands out, especially [HM99] with the computation of the chromatic polynomial of the truncated icosahedron, or bucky-ball, with 60 vertices. The fact that the graph was both very sparse and had a large automorphism group was essential. For graphs of even moderate density we know of no example of comparable size. A good computational challenge, even with the use of symmetries, would be:

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 which can be constructed as clique sums ar the *chordal graphs*, i.e. the graphs in which any cycle of length greater than 3 has a chord. By a theorem of Dirac [Dir61] 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 the *outerplanar graphs*, i.e. planar graph which 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 which is a k-clique sum of G_1 and G_2 we can, see [RT88], express the chromatic polynomial as

$$P(G, x) = \frac{P(G_1, x)P(G_2, x)}{P(K_k, x)}. \tag{1.1}$$

Thus for graphs which 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 [RT88], 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 Formula 1.1 have chromatic polynomials with log-concave coefficients and at least one of them have a good factoring then $P(G,x)$ has strictly log-concave coefficients.*

This follows immediately from the fact, see e.g. [Kar68], that products preserve log-concavity. From the formulae 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 which can be built by repeated clique sums using complete graphs, cycles, trees, and graphs with at most 9 vertices. Since the property of having a good factoring is stronger than being strictly log-concave it is natural to ask

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

The chromatic polynomial has a generalisation to matroids as well, the so called characteristic polynomial of a matroid, see e.g. [Oxl92]. For this polynomial analogues of Conjecture 1.1 and Conjecture 1.2 have been posed 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 [Aig87].

2. Subgraphs without broken cycles

There are several different expansions for the chromatic polynomial of a graph in terms of its subgraphs, see [Big93]. In 1932 Whitney [Whi32] gave the following characterisation. Assume that the edges of a graph G have been labelled 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 greatest label, among those in the cycle, is called a broken cycle.

Theorem 2.1 ([Whi32])**.**
The coefficient a_i equals the number of subgraphs of G, with $n-i$ edges, which 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 labelling of the graph. So in light of Whitney's theorem and the deletion-contraction formulae for the chromatic polynomial we see that Conjecture 1.2 really concerns how the number of broken circuit free subgraphs change under edge deletion and contraction.

We say that two sequences a_i and b_i are *co-concave* if $a_i + b_i$ is a log-concave sequence. Due to the deletion-contraction formulae for the chromatic polynomial, co-concavity is a key property for understanding the structure behind the log-concavity conjecture. If the two chromatic polynomials in the formulae could be shown to be co-concave 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 co-concave.*

We have verified the conjecture for all graphs on at most 9 vertices.

A subgraph which does not contain a broken cycle obviously can not 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 behaviour 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.

First let us 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. Given a graph G we can now calculate the mean size of a broken cycle free subgraph of G, let us denote this size $\varepsilon(G)$, 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 $b_i = \frac{\left[{n \atop i}\right]}{n!}$, where $\left[{n \atop j}\right]$ are the Stirling numbers of the first kind. The mean size of a broken circuit free subgraph here is $n - \sum_{i=1}^{n-1} i^{-1}$, and the b_i's converge to a Poisson distribution with mean $\varepsilon(K_n)$ [MW58].

In Figure 1 we have plotted $\varepsilon(G)$ for all connected graphs on 8 vertices. At the bottom left we find all the trees on 8 vertices, all at the same point, and at the top right we find K_8. From our test on small graphs we would like to pose a few problems and conjectures on the behaviour of $\varepsilon(G)$.

Conjecture 2.4. *Let G be a connected graph on n vertices, which 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 is 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

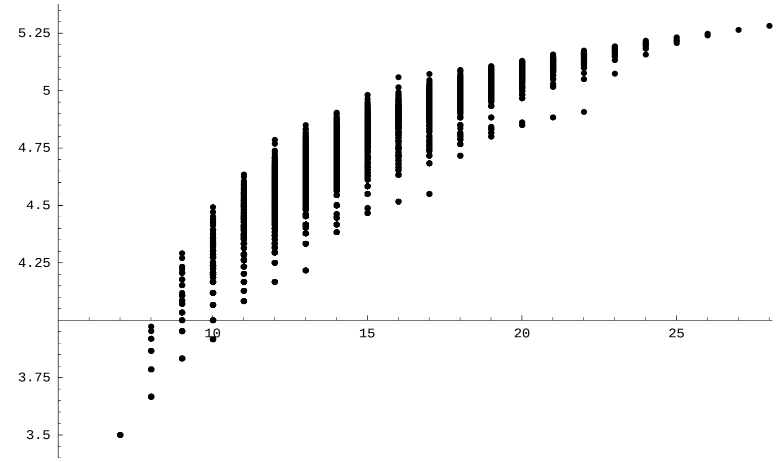

FIGURE 1. $\varepsilon(G)$ plotted for all connected graphs on 8 vertices. The horizontal coordinate show the number of edges in the graphs.

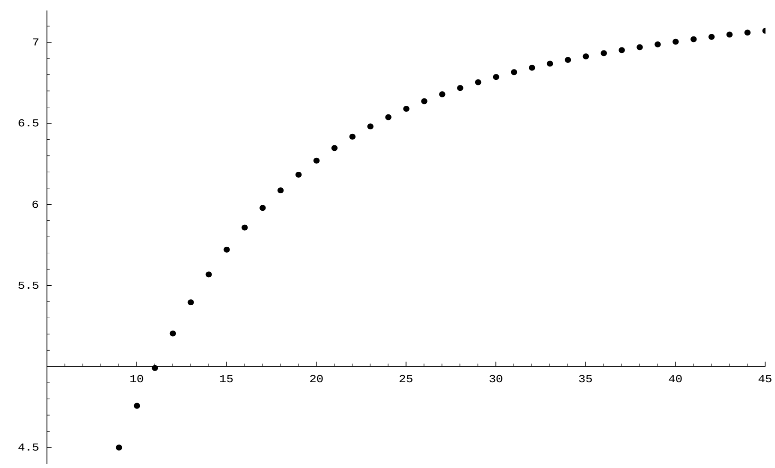

FIGURE 2. $\varepsilon(10, k)$

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 these two values it can be calculated as well.

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

3. Some results on the behaviour of $\varepsilon(G)$

Apart from their inherent value the interest in the problems and conjectures of the previous section really stem from the concept of co-concavity. 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 co-concave 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 $\varepsilon(G)$ changes 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 to have co-concave coefficients. However, as we shall see there seems to be some nice structure to their behaviour.

Let us first note that

$$\varepsilon(G) = n + \frac{P'(G,-1)}{P(G,-1)}. \tag{3.1}$$

This makes the following definitions convenient

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

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

As noted by Stanley [Sta73] $\eta(G)$ 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.

We now observe that

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

(1) $n_e = n'_e$.
(2) $\eta(G) = n_e + n'_e = 2n_e = \eta(G-e) + \eta(G/e)$.
(3) $\eta(G) > \eta(G-e)$.
(4) $\eta(G-e) = \eta(G) - n_e + n''_e = \frac{1}{2}\eta(G) + n''_e = n_e + n''_e$.
(5) $\eta(G/e) = \frac{1}{2}\eta(G) - n''_e = n_e - n''_e$.
(6) $2^{n-2} \le n_e$, *and the bound is sharp if G is a tree.*

Proof.

(1) Let us assume that the edges in G have been labelled and that e is the edge with the highest label. Now $n'_e \ge n_e$ 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 $n_e \ge n'_e$ since if H is counted by n'_e and $H \cup e$ contains a broken cycle then either H contained a broken cycle or $H \cup e$ contains a broken cycle where the missing edge has a higher label than e, in both cases we have a contradiction.

(2) The first equality follows directly from the definition of $\eta(G)$, the second equality follows from (1), and the last equality from the deletion-contraction formula together with Equation 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 makes sure that the two terms have the same sign.

(3) Follows from (2).

(4) Let H be a subgraph counted by n_e, then both H and $H-e$ will be a broken cycle free subgraph of G. However, none of the graphs counted by n_e are subgraphs of $G-e$ and so 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 (1) and (2).
(5) Follows from (4) and (2).
(6) Every broken cycle free subgraph of G/e can be expanded into at least one subgraph of G counted by n_e. By (2) we thus get that n_e will be at least $\frac{1}{2}\eta(T) = 2^{n-2}$, where T is a tree on $n-1$ vertices.

□

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 (2) 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 such that it is used by a very large number of short cycles and at the same time it 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 8 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

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, see [RT88], that the chromatic polynomial of G can be expressed as

$$P(G,x) = \frac{P(G_1,x)P(G_2,x)}{P(K_k,x)}. \tag{3.2}$$

If we now differentiate this and use the identity 3.1 we get,

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

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}, 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 8 vertices.

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

Proposition 3.9. *Let v be a vertex of G such that the neighbours of v induces 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 the neighbourhood of v which is a K_k. By Theorem 3.4 we have that,

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

the difference of which, 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 fulfill 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.

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

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 $\varepsilon\left(G-e\right)$ 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 $l_1 = \varepsilon\left(G\right) - \frac{1}{2}$, since for every graph counted by n_e there is a graph with one edge less 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 $\varepsilon\left(G-e\right) \geq \varepsilon\left(G\right) - \frac{1}{2}$ 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.

□

In a typical situation one would expect both l_1 and l_2 to contribute to the average and so make $\varepsilon\left(G-e\right)$ stay much closer to $\varepsilon\left(G\right)$. 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 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*

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

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

Proof. When $m-(n-1) \leq \binom{n}{2} - m$ the result follows from Azuma's inequality 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 only get a denominator of $T = m-(n-1)$ in the exponent since the first $n-1$ edges can be taken to form a spanning tree in G, and thereby not giving any 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. □

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 behaviour 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 [Gri77] 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

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

Now what we would like to calculate is

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

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

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

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 that

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

Next we can make 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

$$\mathsf{E}\left(\frac{1}{\eta(G)}\right) \sim \frac{1}{\eta(p)}.$$

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 which come in handy,

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

where d_v is the degree of the vertex v and $f(x) = (x!)^{\frac{1}{x}}$. The lower bound is from [GKKS93] and the upper from [GYY80], further bounds can be found in [KS96]. Since the degree sequence of a graph from $G(n,p)$ is quite well concentrated, see

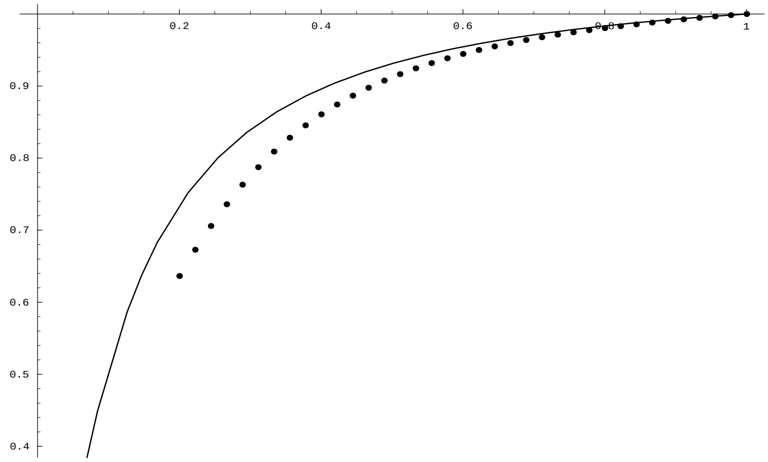

FIGURE 3. $\varepsilon\,(10,k)$ together with $\frac{\eta'(10,p)}{\eta(10,p)}$. Both axes have been rescaled to run between 0 and 1.

[Bol01], $\eta(G)$ will be slightly less than $(pn)^n$ and does not have too large variance[1] and so our estimate for $\mathsf{E}\left(\frac{1}{\eta(G)}\right)$ will be quite good. Thus we should find that

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

where the $\sim$ 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 anti-correlated the approximation should be reasonably good.

In Figure 3 we have plotted the exact values for $\varepsilon\,(10,k)$ together with $\frac{\eta'(10,p)}{\eta(10,p)}$. As can be seen we have a good agreement for large values of p but for lower values the curves grow apart. This 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, $\eta'(1) = \pm(n - \sum_{i=1}^{n-1} i^{-1}), \eta(1) = \pm 1$, 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

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

[1]In fact one would expect $\eta(p)/(pn)^n$ to have a log-normal distribution.

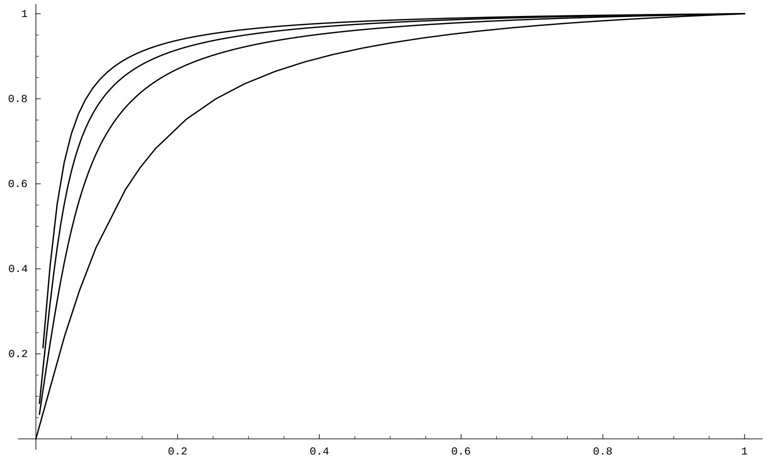

FIGURE 4. $\frac{\eta'(n,p)}{\eta(n,p)}$ for $n = 10, 20, 30, 40$ (n=10 is the lowest curve). Both axes have been rescaled to run between 0 and 1.

where $x^{\overline{n}} = x(x+1)\dots(x+(n-1))$. The subgraphs counted by this generating function are expected to be on the 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 which would have created broken cycles in K_n. So we expect to get a lower bound for $\varepsilon(n,p)$ if we calculate $\frac{S'_n(p,-1)}{S_n(p,-1)}$. 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 which 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

[Aig87] Martin Aigner, *Whitney numbers*, Combinatorial geometries, Cambridge Univ. Press, Cambridge, 1987, pp. 139–160.

[Big93] Norman Biggs, *Algebraic graph theory*, second ed., Cambridge University Press, Cambridge, 1993.

[Big01] Norman Biggs, *A matrix method for chromatic polynomials*, J. Combin. Theory Ser. B **82** (2001), no. 1, 19–29.

[Big02] Norman Biggs, *Equimodular curves*, Discrete Math. **259** (2002), no. 1-3, 37–57.

[Bir12] G.D. Birkhoff, *A detereminantal formula for the number of ways of colouring a map*, Ann. of Math **14** (1912), 42–46.

[Bol01] Béla Bollobás, *Random graphs*, second ed., Cambridge University Press, Cambridge, 2001.

[Bre92] Francesco Brenti, *Expansions of chromatic polynomials and log-concavity*, Trans. Amer. Math. Soc. **332** (1992), no. 2, 729–756.

[Dir61] G. A. Dirac, *On rigid circuit graphs*, Abh. Math. Sem. Univ. Hamburg **25** (1961), 71–76.

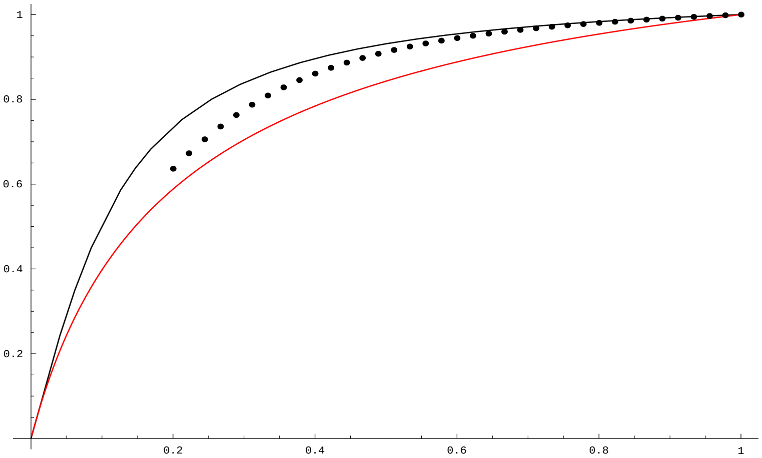

Figure 5. $\varepsilon(10,k)$ together with $\frac{\eta'(10,p)}{\eta(10,p)}$ and $\frac{S'_{10}(p,-1)}{S_{10}(p,-1)}$. Both axes have been rescaled to run between 0 and 1.

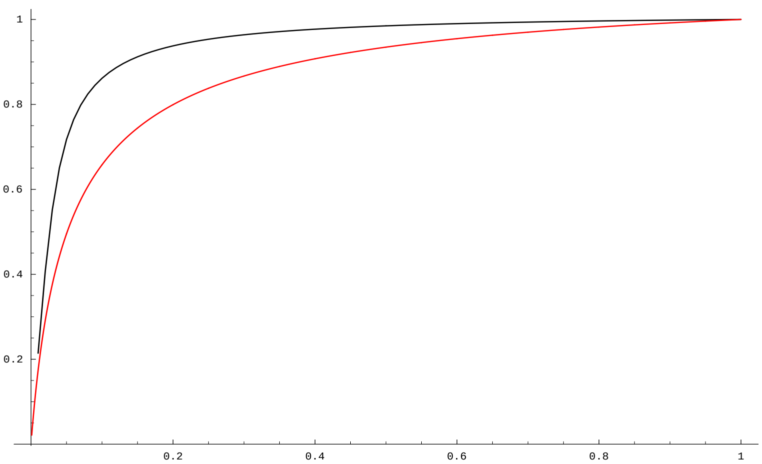

Figure 6. $\frac{\eta'(40,p)}{\eta(40,p)}$ and $\frac{S'_{40}(p,-1)}{S_{40}(p,-1)}$. Both axes have been rescaled to run between 0 and 1.

[GKKS93] Wayne Goddard, Claire Kenyon, Valerie King, and Leonard J. Schulman, *Optimal randomized algorithms for local sorting and set-maxima*, SIAM J. Comput. **22** (1993), no. 2, 272–283.

[Gri77] Geoffrey Grimmett, *Random graph theorems*, 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, Reidel, Dordrecht, 1977, pp. 203–209.

[GYY80] Ronald L. Graham, Andrew C. Yao, and F. Frances Yao, *Information bounds are weak in the shortest distance problem*, J. Assoc. Comput. Mach. **27** (1980), no. 3, 428–444.

[HM99] Gary Haggard and Thomas R. Mathies, *The computation of chromatic polynomials*, Discrete Math. **199** (1999), no. 1-3, 227–231.

[Hog74] S. G. Hoggar, *Chromatic polynomials and logarithmic concavity*, J. Combinatorial Theory Ser. B **16** (1974), 248–254.

[Jac02] Bill Jackson, *Zeros of chromatic and flow polynomials of graphs*, Available at http://xxx.lanl.gov/abs/math.CO/0205047, 2002.

[Kar68] Samuel Karlin, *Total positivity. Vol. I*, Stanford University Press, Stanford, Calif, 1968.

[KS96] Nabil Kahale and Leonard J. Schulman, *Bounds on the chromatic polynomial and on the number of acyclic orientations of a graph*, Combinatorica **16** (1996), no. 3, 383–397.

[McK] Brendan McKay, The progam geng is available from Brendan McKays home page at http://cs.anu.edu.au/people/bdm.

[MW58] L. Moser and M. Wyman, *Asymptotic development of the Stirling numbers of the first kind*, J. London Math. Soc. **33** (1958), 133–146.

[Oxl92] James G. Oxley, *Matroid theory*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York, 1992.

[Rea68] Ronald C. Read, *An introduction to chromatic polynomials*, J. Combinatorial Theory **4** (1968), 52–71.

[Rea86] Ronald C. Read, *Broken wheels are SLC*, Ars Combin. **21** (1986), no. A, 123–128, Thirteenth Australasian conference on combinatorial mathematics and computing (Sydney, 1985).

[RR91] Ronald C. Read and Gordon F. Royle, *Chromatic roots of families of graphs*, Graph theory, combinatorics, and applications. Vol. 2 (Kalamazoo, MI, 1988), Wiley-Intersci. Publ., Wiley, New York, 1991, pp. 1009–1029.

[RT88] R. C. Read and W. T. Tutte, *Chromatic polynomials*, Selected topics in graph theory, 3, Academic Press, San Diego, CA, 1988, pp. 15–42.

[Sok01] Alan D. Sokal, *Bounds on the complex zeros of (di)chromatic polynomials and Potts-model partition functions*, Combin. Probab. Comput. **10** (2001), no. 1, 41–77.

[Sok05] Alan D. Sokal, *The multivariate tutte polynomial (alias potts model) for graphs and matroids*, Surveys in Combinatorics 2005, Cambridge University Press, Cambridge Uk, 2005, pp. 173–227.

[Sta73] Richard P. Stanley, *Acyclic orientations of graphs*, Discrete Math. **5** (1973), 171–178.

[Tho97] Carsten Thomassen, *The zero-free intervals for chromatic polynomials of graphs*, Combin. Probab. Comput. **6** (1997), no. 4, 497–506.

[Tho00] C. Thomassen, *Chromatic roots and Hamiltonian paths*, J. Combin. Theory Ser. B **80** (2000), no. 2, 218–224.

[Tut98] W. T. Tutte, *Graph theory as I have known it*, The Clarendon Press Oxford University Press, New York, 1998, With a foreword by U. S. R. Murty.

[Whi32] Hassler Whitney, *A logical expansion in mathematics*, Bull. Amer. Math. Soc. **38** (1932), 572–579.

E-mail address: `phl@kth.se`

E-mail address: `Klas.Markstrom@math.umu.se`

KTH Physics, AlbaNova University Center,SE-106 91 Stockholm, Sweden

Department of Mathematics, UmeåUniversity , SE-901 87 Umeå, Sweden