This is a translation of a paper published in:

Teoriya Veroyatnostei i ee Primeneniya, 48(2003), 411--416

One Probabilistic Equivalent of the Four Color Conjecture

Yu. V. Matiyasevich*

Abstract

For every two-connected planar threevalent graph we introduce in a natural way a probabilistic space and define two random events; the Four Color Conjecture turns out to be eqivalent to (positive) correlation of these events.

In 2002 the famous Four Color Conjecture became 150 years old. According to the experts on the history of mathematics (see, for example [8], [18]), the sole author of the Conjecture is Francis Guthrie. In 1852 he was coloring a map of England and found that 4 colors were sufficient for coloring the whole map in such a way that each two neighboring counties (To the Editor: sic! counties, not countries) were colored in different colors. Francis conjectured that 4 colors would be sufficient for any possible map as well (of course, areas for coloring should be connected). Being unable to prove it, he turned for help to his brother Frederick who was a student in mathematics. The latter in his turn addressed this question to his Professor, A. de Morgan, who began to put this question to others.

The first "proof" of Guthrie's conjecture was published by A. B. Kempe [15] a quoter of century later; after another ten years an error was found.

Kempe became the first but not the last author of an erroneous proof of Guthrie's conjecture. The situation changed in 1976 when K. Appel and W. Haken [2] annouced their proof based on intensive use of computers. The volume of required calculations was so big that the proof published in [3], [4] could not be checked by a human-being.

*St.Petersburg Department of Steklov Mathematical Institute, RAS, Fontanka St. 27, 191011 St. Petersburg, Russia URL: http://logic.pdmi.ras.ru/~yumat

Now, a quoter of century later, the state of the art essentially did not change. K. Appel and W. Haken published another, slightly simplified proof [5]. N. Robertson, D. Sanders, P. Seymour, and R. Thomas came to the conclusion that they cannot check even the part of this proof which was done without computers, wrote new programs from scratch and obtained further simplifications. Nevertheless their new proof [20] was based on the same ideas originating from the "proof" of Kempe and cannot not be checked by a human-being.

Thus, there is a "demand" to find a proof acceptible for a human-being. What could be the role of the theory of probabilities here? (In other words—how could one justify publication of a paper about graph colorings in this Journal?) One of attributes of an important mathematical problem is the presence of its restatements in the language of several areas of mathematics. In this respect the Four Color Conjecture is an outstanding problem—it has scores of reformulations in the languages of a dozen different areas of mathematics (see., for example, [1], [6], [7], [9], [10], [11], [12], [13], [14], [16], [19], [21], [22], [23]). In this paper the Four Color Conjecture is restated in the language of the theory of probabilities—with the notion of planar graph but without the notion of coloring. It remains open whether the ideas and methods of the theory of probabilities could be useful for a new proof, to say nothing about a proof acceptible for a human-being.

To begin with, we are to clarify the notion of map used in the Four Color Conjecture. We will assume that the map is presented via plane *image* of some two-connected graph. Graph image is a mapping of the set V of its vertices into pairwise different points of the plane and of the set E of its edges into simple Jordan curves having the images of corresponding verticies as their endpoints and having no other common points. Such an image split the plane into open areas called *countries*. Two different countries are considered to be neighboring if there is an edge the image of which belongs to the closure of each of these countries. (The curves are required to be Jordan in order to avoid pathological splitting of the plane into infinitely many pairwise neighboring countries; the two-connectness guaranties that the image of each edge belongs to the boundary of two different countries, i.e., no country is neighboring to itself.)

It is easy to understand that, while proving the Four Color Conjecture, we can restrict ourself to consideration of only threevalent graphs, i.e., graphs where every vertex is the endpoint of exactly three edges (otherwise, we could surround the image of every vertex by a small neighborhood and declare it a new country). Let $G = \langle V_G, E_G \rangle$ be such a graph with $3n$ edges $e_1, \dots, e_{3n}$ (clearly, the number of edges of a threevalent graph is always a multiple of 3).

Besides the graph G we can consider its *edge graph* $F_G = \langle E_G, L_G \rangle$. The verticies of graph F_G are just the edges of graph G. Two verticies, e_i and e_j, of the graph F_G are connected in it by an edge if and only if e_i and e_j, considered as edges of the graph G, are both incident to the same vertex of this graph.

Each of $3n$ verticies of graph F_G has degree 4, hence the graph F_G has $6n$ edges.

A non-directed graph with $6n$ edges can be made directed in 2^{6n} ways. We will assume that each of these ways has the same probability, in other words, each of the two possible orientations of every edge will be selected with probability $\frac{1}{2}$ independently of the orientations of other edges. More precisely, we are to consider simultaneously two random orientations $\overrightarrow{F'_G}$ and $\overrightarrow{F''_G}$ of the graph F_G, so our probabilistic space will consist of 2^{12n} pairs of orientations $\langle \overrightarrow{F'_G}, \overrightarrow{F''_G} \rangle$ with probability of any particular pair equal to 2^{-12n}.

We will introduce two notions of similarity of two orientations and establish their relationship with the Four Color Conjecture.

We say that two orientations of the graph F_G *have the same parity* if one of them can be obtained from the other via changing orientation of even number of edges.

In a directed graph every vertex is characterized by two numbers, namely, the *in-degree* and the *out-degree*, i.e., the number of incoming and outcoming edges. We say that two orientations of the graph F_G *are congruent modulo* 3 if for each of its vertex the out-degree in one orientation is congruent modulo 3 to the out-degree of the same vertex in the other orientation.

Let A_G denote the event "pair of randomly selected orientations has the same parity". Let B_G denote the event "pair of randomly selected orientations is congruent modulo 3".

It turns out that, for every two-connected planar threevalent graph G, the events A_G and B_G are not independent, moreover, the coefficient of correlation is closely related to $\chi_G(4)$, the number of 4-colorings of the map represented by the G.

Theorem *For every two-connected planar threevalent graph G with $3n$ edges*

$$\mathbf{P}(B_G \mid A_G) - \mathbf{P}(B_G) \quad = \quad \left(\frac{27}{4096}\right)^n \cdot \frac{\chi_G(4)}{4}. \tag{1}$$

Corollary *The Four Color Conjecture is equivalent to the statement that for every two-connected planar threevalent graph G the events A_G and B_G are not independent.*

Proof. In order to represent all 2^{12n} possible selections of orientations $\overrightarrow{F'_G}$ and $\overrightarrow{F''_G}$ we will use generating functions. They will be polynomials in formal variables $x_1, \ldots, x_{3n}$, each of which corresponds to one element of the set E_G (which is both the set of edges of the graph G and the set of vertices of the graph F_G).

When constructing the generating function for possible selections of the orientation $\overrightarrow{F'_G}$, we will represent the two possible directions of the edge connecting vertices e_i and e_j of the graph F_G via the polynomial

$$x_i - x_j, \tag{2}$$

where $i < j$, the monomial x_k symbolizing that edge goes from the vertex e_k. Under this agreement, the 2^{6n} possible selections of orientation $\overrightarrow{F'_G}$ are in a natural one-to-one correspondence with 2^{6n} monomials resulting from removing the brackets (but without combining similar terms) in

$$M' = \prod_{e_i e_j \in L_G} (x_i - x_j). \tag{3}$$

It is easy to see that orientations of equal parity correspond to monomials of equal signs, and vice versa.

To be able to deal with congruences modulo 3 we introduce operator $\mathcal{R}$ which replaces the exponent of every variable x_k by its remainder from dividing by 3. Respectively, orientations which are congruent modulo 3 are represented by equal (up to the sign) monomials in the polynomial $\mathcal{R}M'$.

Possible selections of orientation $\overrightarrow{F''_G}$ will be represented by a polynomial similar to polynomial M'. The distinction is as follows: instead of (2), the two possible selection of the direction of the edge connecting verticies e_i and e_j are now represented by polynomial

$$x_i^2 - x_j^2,$$

with monomial x_k^2 symbolizing that the edge goes from the vertex e_k. Respectively, polynomial M'' is defined by

$$M'' = \prod_{e_i e_j \in L_G} (x_i^2 - x_j^2). \tag{4}$$

Put

$$M = M'M'' \tag{5}$$

and let us calculate the free term m_0 of the polynomial

$$\mathcal{R}M = \mathcal{R}(M'M'') = \mathcal{R}(\mathcal{R}(M')\mathcal{R}(M''))$$

in two ways.

The first way of calculation of m_0 is connected with the left-hand side of (1). Clearly, $\mathbf{P}(A_G) = \frac{1}{2}$, so this left-hand side is equal to

$$2\mathbf{P}(A_G B_G) - \mathbf{P}(B_G). \tag{6}$$

Whatever is a pair of orientations $\langle \overrightarrow{F'_G}, \overrightarrow{F''_G} \rangle$, one of the three following cases holds.

I. These orientations are not congruent modulo 3.

II. These orientations are congruent modulo 3 and have equal parity.

III. These orientations are not congruent modulo 3 but have different parities.

In the case I the contribution of the pair of orientations $\langle \overrightarrow{F'_G}, \overrightarrow{F''_G} \rangle$ into (6) is equal to 0, and the contribution of corresponding monomials into m_0 is equal to 0 as well; in the case II the contribution of these pair of orientations into (6) is equal to 2^{-12n}, while the contribution of corresponding monomials into m_0

is equal to 1; at last, in the case III the contribution of the pair of orientations into (6) is equal to -2^{-12n}, while the contribution of the monomials into m_0 is equal to -1. Hence,

$$m_0 = 2^{12n}(\mathbf{P}(B_G \mid A_G) - \mathbf{P}(B_G)). \tag{7}$$

The second way of calculation of m_0 is connected with the right-hand side of (1).

Every coloring of the map presented by the graph G in four colors $\alpha, \beta, \gamma, \delta$ generates a coloring of the edges of this graph in three colors 1, 2, 3 known as *Tait coloring*. Namely, if two neighboring contries are colored either in colors α and β or in colors γ and δ then the edge separating these countries is colored in color 0; if two countries are colored either in colors α and γ or in colors β and δ then the edge is colored in color 1; at last, if the colors of the countries are either α and δ or β and γ, then the color of the edge is 2. It is easy to see that for every vertex the three edges incident to it are colored pairwise differently, i.e., in all three colors 1, 2 and 3.

Vice versa, every Tait coloring can be generated in this manner from a suitable coloring of the map. We can start by coloring an arbitrary country in any of the 4 colors and then proceed by coloring neighboring countries according to the above described rules (it is easy to prove that there will be no contradiction thanks to one-connectednes of the sphere). Thus, the number of Tait colorings is equal to $\frac{1}{4}\,\chi_G(4)$.

The degree of the polynomial $\mathcal{R}M$ in each of its $3n$ variables is at most 2, hence, this polynomial can be uniquely reconstructed from its values at suitable 3^{3n} choices of values of its variables. For these values we assign to each of the variables $x_1, \ldots, x_{3n}$ the three values of the cubic root of 1:

$$1, \quad \omega, \quad \omega^2, \tag{8}$$

where $\omega = (-1 + \sqrt{-3})/2$. By Lagrange interpolation

$$\mathcal{R}M = \sum_{\mu} (\mathcal{R}M)(\omega^{\mu(v_1)}, \ldots, \omega^{\mu(v_{3n})})\, P_\mu, \tag{9}$$

where

$$P_\mu = \prod_{k=1}^{3n} \frac{(x_k - \omega^{\mu(v_k)+1})(x_k - \omega^{\mu(v_k)+2})}{(\omega^{\mu(v_k)} - \omega^{\mu(v_k)+1})(\omega^{\mu(v_k)} - \omega^{\mu(v_k)+2})}, \tag{10}$$

and the summation is over all 3^{3n} maps of the form

$$\mu\colon\ E_G \longrightarrow \{0, 1, 2\}. \tag{11}$$

Thanks to our choice of values (8) we can eliminate the operator $\mathcal{R}$ from the right-hand side of (9):

$$\mathcal{R}M = \sum_{\mu} M(\omega^{\mu(v_1)}, \ldots, \omega^{\mu(v_3 n)})\, P_\mu. \tag{12}$$

It is easy to see that if for some edge $e_i e_j$ of the graph F_G the equality $\mu(e_i) = \mu(e_j)$ holds, then $M(\omega^{\mu(v_1)}, \ldots, \omega^{\mu(v_{3n})}) = 0$. Thus we can restrict summation in (12) to those $\frac{1}{4}\,\chi_G(4)$ maps μ, which are Tait colorings. Let us determine the value of $M(\omega^{\mu(v_1)}, \ldots, \omega^{\mu(v_{3n})})$ in this case.

According to (5), (3), and (4), the polynomial M is the product of $12n$ factors which can be groupped into $2n$ products of the form

$$(x_i - x_j)(x_j - x_k)(x_i - x_k)(x_i^2 - x_j^2)(x_j^2 - x_k^2)(x_i^2 - x_k^2) \qquad (13)$$

where $i < j < k$ and corresponding edges e_i, e_j, e_k of the graph G are incident to the same vertex. In a Tait coloring these three edges are colored in the three colors 0, 1 2, hence the value of the product (13) is always equal to

$$(1-\omega)(\omega-\omega^2)(1-\omega^2)(1-\omega^2)(\omega^2-\omega^4)(1-\omega^4) = 27$$

and, respectively,

$$\begin{aligned} M(\omega^{\mu(v_1)}, \ldots, \omega^{\mu(v_{3n})}) &= 3^{6n}, \\ m_0 = 3^{6n} \sum_{\mu} P_\mu(0, \ldots, 0), & \qquad (14) \end{aligned}$$

where the summation is taken over those $\frac{1}{4}\,\chi_G(4)$ maps μ which are Tait colorings.

According to (10),

$$\begin{aligned} P_\mu(0, \ldots, 0) &= \prod_{k=1}^{3n} \frac{\omega^{\mu(v_k)+1}\omega^{\mu(v_k)+2}}{\left(\omega^{\mu(v_k)} - \omega^{\mu(v_k)+1}\right)\left(\omega^{\mu(v_k)} - \omega^{\mu(v_k)+2}\right)} \\ &= \prod_{k=1}^{3n} \frac{1}{(1-\omega)(1-\omega^2)} = 3^{-3n}. \end{aligned}$$

Together with (14) and (7) we get the required equality (1).

References

[1] N. Alon and M. Tarsi, *A note on graph colorings and graph polynomials*, J. Combin. Theory, Ser. B, 70 (1997), pp. 197–201.

[2] K. Appel and W. Haken, *Every planar map is four colorable*, Bull. Amer. Math. Soc., 82 (1976), pp. 711–712.

[3] K. Appel and W. Haken, *Every planar map is four colorable. Part* I: *Discharging*, Illinois J. Math., 21 (1977), pp. 429–490.

[4] K. Appel, W. Haken, and J. Koch, *Every planar map is four colorable. Part* II: *Reducibility*, Illinois J. Math., 21 (1977), pp. 491–567.

[5] K. Appel and W. Haken, *Every Planar Map is Four Colorable*, Amer. Math. Soc., Providence, R.I., 1989, 741 p.

[6] Y. Colin de Verdiere, *On a new graph invariant and a criterion for planarity*, Contemp. Math., 147 (1993), pp. 137–147.

[7] T. Hoffman, J. Mitchem, and E. Schmeichel, *On edge-coloring graphs*, Ars Combin., 33 (1992), pp. 119–128.

[8] O. Hudry, *Le problème des quatre couleurs*, La recherche de la vérité, Ed. by M. Serfati. Kangourou, Paris, 1999, pp. 211–240.

[9] T. R. Jensen and B. Toft, *Graph Coloring Problems*, Wiley, New York, 1995.

[10] P. C. Kainen, *Is the four color theorem true*?, Geombinatorics, 3 (1993), pp. 41–56.

[11] L. H. Kauffman, *Map coloring and the vector cross product*, J. Combin. Theory, Ser. B, 48 (1990), pp. 145–154.

[12] L. H. Kauffman, *Combinatorial recoupling theory and* 3*-manifold invariants*, NATO Adv. Sci. Inst., Ser. B Phys., 315 (1993), pp. 1–17.

[13] L. H. Kauffman, *Spin networks, topology and discrete physics*, Adv. Ser. Math. Phys., 17 (1994), pp. 234–274.

[14] L. H. Kauffman and H. Saleur, *An algebraic approach to the planar coloring problem*, Comm. Math. Phys., 152 (1993), pp. 565–590.

[15] A. B. Kempe, *On the geographical problem of the four colors*, Amer. J. Math., 2 (1879), pp. 193–201.

[16] Yu. Matiyasevich, *Some arithmetical restatements of the four color conjecture*, Theoret. Comput. Sci., 257 (2001), pp. 167–183; `http://logic.pdmi.ras.ru/~yumat/Journal/jcontord.htm`, mirrored at `http://www.informatik.\hbox{uni-stuttgart}.de/ifi/ti/personen/Matiyasevich/Journal/jcontord.htm`.

[17] Yu. Matiyasevich, *Some algebraic methods for calculation of the number of colorings of a graph*, Zap. nauchn. seminarov POMI, 283 (2001), . 193–205; `http://www.pdmi.ras.ru/znsl/2001/v283.html`.

[18] K. O. May, *The origin of the four-color conjecture*, Isis, 56 (1965), pp. 346–348.

[19] Ø. Ore, *The Four-Color Problem*, Academic Press, New York, London, 1967, 259 p.

[20] N. Robertson, D. Sanders, P. Seymour, and R. Thomas, *The four-colour theorem*, J. Combin. Theory, Ser. B, 70 (1997), pp. 2–44.

[21] R. Thomas, *An update on the four-colour theorem*, Notices Amer. Math. Soc., 45 (1998), pp. 848–859; `http://www.ams.org/notices/199807/thomas.{ps,pdf}`.

[22] T. L. Saaty, *Thirteen colorful variations on Guthrie's four-color conjecture*, Amer. Math. Monthly, 79 (1972), pp. 2–43.

[23] T. L. Saaty and P. C. Kainen, *The Four-Color Problem. Assaults and Conquest*, McGraw-Hill, Düsseldorf etc., 1977.