Coloring triangle-free graphs on surfaces. Extended abstract.

Zdeněk Dvořák[*] Daniel Král'[†] Robin Thomas[‡]

Abstract

Gimbel and Thomassen asked whether 3-colorability of a triangle-free graph drawn on a fixed surface can be tested in polynomial time. We settle the question by giving a linear-time algorithm for every surface which combined with previous results gives a linear-time algorithm to compute the chromatic number of such graphs.

Our algorithm is based on a structure theorem that for a triangle-free graph drawn on a surface Σ guarantees the existence of a subgraph H, whose size depends only on Σ, such that there is an easy test whether a 3-coloring of H extends to a 3-coloring of G. The test is based on a topological obstruction, called the "winding number" of a 3-coloring. To prove the structure theorem we make use of disjoint paths with specified ends to find a 3-coloring.

If the input triangle-free graph G drawn in Σ is 3-colorable we can find a 3-coloring in quadratic time, and if G quadrangulates Σ then we can find the 3-coloring in linear time. The latter algorithm requires two ingredients that may be of independent interest: a generalization of a data structure of Kowalik and Kurowski to weighted graphs and a speedup of a disjoint paths algorithm of Robertson and Seymour to linear time.

1 Introduction

In this paper we are concerned with coloring graphs that embed in a fixed surface. This restriction often makes coloring problems tractable, and the purpose of this paper is to describe a new result along these lines. Furthermore, we restrict ourselves to triangle-free graphs. The following is a classical theorem of Grötzsch [10].

THEOREM 1.1. *Every triangle-free planar graph is* 3*-colorable.*

Thus deciding whether a triangle-free planar graph is 3-colorable is trivial. However, Theorem 1.1 does not generalize to any other surface Σ, and, in fact, the 3-colorability of triangle-free graphs embedded in Σ is an interesting problem. When Σ is the projective plane it was solved by Gimbel and Thomassen [9], but it was open for all other surfaces. Gimbel and Thomassen asked whether it is polynomial-time solvable for all surfaces. In this paper we settle the question, as follows.

THEOREM 1.2. *For every surface Σ there exists a linear-time algorithm that given an input triangle-free graph G embedded in Σ correctly determines whether G is* 3*-colorable.*

In fact, our result is more general in the sense that we allow a bounded number of vertices to have their color specified. That, in turn, is essential for the inductive argument to go through.

We prove Theorem 1.2 by means of the following structural result. We need some definitions first. If G is a graph embedded in a surface Σ and f is a face of a subgraph of G, then by $G[f]$ we denote the subgraph of G consisting of all vertices and edges of G embedded in the closure of f. Furthermore, we regard $G[f]$ as embedded in the surface $\Sigma[f]$ obtained from f by capping off each component of the boundary of f by a disk. Let $C_1, C_2, \ldots, C_k$ be the components of the boundary of f, and let us pretend that they are cycles. (Otherwise they are walks and can be turned into cycles by splitting vertices.) We say that a subgraph J of $G[f]$ is a *locally planar quadrangulation of* $\Sigma[f]$ if J includes $C_1 \cup C_2 \cup \cdots \cup C_k$, every face of J is bounded either by C_i for some $i = 1, 2, \ldots, k$ or by a null-homotopic cycle of length four, and J is w-representative and strongly boundary linked, as defined prior to Theorem 3.3, where w is as in Theorem 3.3 if $\Sigma[f]$ is orientable and Theorem 3.6 otherwise.

[*]Institute for Theoretical Computer Science (ITI), Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. E-mail: rakdver@kam.mff.cuni.cz.

[†]Institute for Theoretical Computer Science (ITI), Faculty of Mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic. E-mail: kral@kam.mff.cuni.cz. Institute for Theoretical Computer Science is supported as project 1M0545 by the Ministry of Education of the Czech Republic.

[‡]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332. E-mail: thomas@math.gatech.edu. Partially supported by NSF Grants No. DMS-0200595 and DMS-0354742.

THEOREM 1.3. *For every surface Σ there exists an integer N such that every triangle-free graph G embedded in Σ has a subgraph H on at most N vertices such that for every face f of H*

(i) *every 3-coloring of the boundary of f can be extended to a 3-coloring of $G[f]$, or*

(ii) *f is homeomorphic to a cylinder, or*

(iii) *the graph $G[f]$ has a subgraph that is a locally planar quadrangulation of $\Sigma[f]$.*

Finding the graph H is the most difficult part of the algorithm, and is described in Sections 4 and 5. Once the graph H is found, we test all 3-colorings of H (there are only constantly many) and for each 3-coloring c of H and each face f of H we test if c extends into $G[f]$. If some c extends for every f, then G is 3-colorable, and otherwise it is not.

Thus we need to explain how to test whether a 3-coloring of the boundary of f extends into $G[f]$. That is trivial if outcome (i) holds. Suppose now that (iii) holds, and let Q_f be the corresponding subgraph of $G[f]$. Using Theorem 3.3 or Theorem 3.6, depending on whether $\Sigma[f]$ is orientable or not, it is easy to decide whether the 3-coloring of the boundary of f extends to Q_f; but every 3-coloring of Q_f extends to a 3-coloring of $G[f]$ by [29]. Thus if outcome (iii) holds, then the extendability of a 3-coloring of the boundary of f into $G[f]$ is decided by Theorem 3.3 or Theorem 3.6. Finally, if outcome (ii) holds, then we break the problem into at most linearly many subproblems, each in a cylinder and each either satisfying the hypothesis of Theorem 3.3 or of bounded size, and use dynamic programming to combine the solutions on the smaller cylinders. We omit further details due to space limitations.

The paper is organized as follows. In the next section we survey related work. The proof of Theorem 1.3 and the attendant algorithm have two major ingredients, namely the special case of quadrangulations, and then the general case. We treat quadrangulations in Section 3, and the general case in Section 4. The proof of Theorem 1.3 can be easily converted to a polynomial-time algorithm, and in Section 5 we sketch the additional ideas needed to speed up the algorithm to linear time.

2 Related work

By a *surface* we mean a compact 2-dimensional manifold with empty boundary. The classification theorem of surfaces states that every surface is homeomorphic to either the surface S_g obtained from the sphere by adding g handles ("the orientable surface of genus g"), or to the surface N_k obtained from the sphere by adding k "cross-caps" ("the non-orientable surface of cross-cap number k"). We refer to [17] for background information on surfaces and graphs embedded in them.

The classical surface coloring theory studied the maximum chromatic number among graphs embeddable on a fixed surface Σ. When Σ is orientable of genus g this maximum is given by the Heawood formula $\lfloor(7+\sqrt{48g+1})/2\rfloor$. The upper bound was shown by Heawood [12], and the lower bound by Ringel and Youngs (see [24, 25]). The corresponding question for triangle-free graphs was studied by Gimbel and Thomassen [9], who proved that for triangle-free graphs the maximum is at least $\Omega(g^{1/3}/\log g)$ and at most $\Omega((g/\log g)^{1/3})$.

A more modern approach to surface coloring, initiated by the work of Thomassen [30, 31] is based on the fact that only very few graphs attain Heawood bound or even come close, while "most" graphs drawn on a surface have chromatic number at most five. Specifically, Thomassen formulated two related questions. The first one asks for the complexity status of the following decision problem for fixed integers k and g and a fixed surface Σ.

k-COLORING GIRTH $\geq q$ GRAPHS IN Σ
Input: a graph G of girth at least q embedded in Σ
Question: is G k-colorable?

The second question is closely related and concerns the finiteness of $(k+1)$-critical graphs among the instances of this decision problem. A graph G is $(k+1)$-*critical* if it is not k-colorable, but every proper subgraph is k-colorable. The second question Thomassen posed is

QUESTION 1. *Given integers k, q and a surface Σ, is the number of $(k+1)$-critical graphs of girth at least q that embed into Σ finite?*

If the answer is yes, then the decision problem k-COLORING GIRTH $\geq q$ GRAPHS IN Σ has an easy polynomial-time algorithm—simply test if the input graph G has a subgraph isomorphic to one of the finitely many $(k+1)$-critical graphs. In fact, this algorithm can be implemented to run in linear time [4]. Furthermore, in all instances when the number of $(k+1)$-critical graphs is finite we also have an explicit bound on the maximum size of such a graph, and hence we not only know that the algorithm exists, but we can actually construct it.

Question 1 is now completely settled. For $k=2$ the answer is negative, because of odd cycles, and for $k=3$ and $q\leq 4$ the answer is negative for all surfaces other than the sphere because of the graphs obtained from odd cycles by means of the Mycielski's construction [1, Section 8.5]. For $k = 4$, $q = 3$ the answer is negative because of a construction of Fisk [5]. For pairs k, q such that $q \geq 6$ and $k \geq 3$, or $q \geq 4$ and $k \geq 4$, or $k \geq 7$ the answer is positive by Euler's formula and a theorem of Gallai [6, 7]; see [9]. In the remaining cases $k = 5$, $q = 3$ and $k = 3$, $q = 5$ the answer is positive by two deep theorems of Thomassen [30, 31]. Let us state the one directly related to our work.

THEOREM 2.1. *For every surface Σ, the number of 4-critical graphs of girth at least five that embed into Σ is finite.*

It follows from the solution to Question 1 that decision problem k-COLORING GIRTH $\geq q$ GRAPHS IN Σ is polynomial-time solvable for all pairs (k, q), except the pairs $(3, 3)$, $(3, 4)$ and $(4, 3)$. In case of the last pair Problem 1 has a trivial solution when Σ is the sphere, but for all other surfaces its resolution is clouded by similar issues that make the Four-Color Problem so difficult. Thus the prospects for a solution are not very bright. In case of the pair $(3, 3)$ Problem 1 is NP-hard for all surfaces, because it is NP-hard even when Σ is the sphere [8, Theorem 4.2]. That leaves us with the case $k = 3$ and $q = 4$, which constitutes the main result of this paper. As mentioned earlier, the problem is trivial by Theorem 1.1 when Σ is the sphere, and an elegant solution was given by Gimbel and Thomassen [9] when Σ is the projective plane. The problem was open for all other surfaces.

3 Coloring Quadrangulations

The first step toward our ultimate goal was obtained by Hutchinson [11]. Let G be a graph embedded in a surface Σ. We say that a cycle C in G is *trivial* if it bounds a disk and *non-trivial* otherwise. We say that G is embedded with *edge-width at least w* if every nontrivial cycle in G has length at least w.

THEOREM 3.1. *For every orientable surface Σ there exists an integer w such that every graph G embedded in Σ with all faces even and edge-width at least w is 3-colorable.*

For our purposes we need a generalization to triangle-free graphs. Indeed, we were able to obtain such generalization, as follows.

THEOREM 3.2. *For every orientable surface Σ there exists an integer w such that every triangle-free graph G embedded in Σ with edge-width at least w is 3-colorable.*

A proof is omitted due to space restrictions. Theorem 3.2 is actually a corollary of our theory, rather than a logical next step, but it suggests what we should be doing. If G has large edge-width, then we are done by Theorem 3.2; otherwise we can cut the surface open along a non-trivial cycle of bounded length, and reduce the problem to a simpler surface. However, we reduce to a more general problem, namely one in which a bounded number of cycles, each of bounded length, are precolored. Thus we need a version of Theorem 3.2 for graphs with precolored cycles. We have obtained such version, but it is not completely straightforward, because of a topological obstruction that we now introduce.

To facilitate working with graphs with precolored cycles we introduce the following terminology. Let G be a graph and let $C_1, C_2, \ldots, C_k$ be distinct cycles in G. We say that G is a *rooted graph* with cycles $C_1, C_2, \ldots, C_k$. We define the *length* of the rooted graph as the sum of the lengths of the cycles $C_1, C_2, \ldots, C_k$. We say that a rooted graph G with cycles $C_1, C_2, \ldots, C_k$ is *embedded in a surface* Σ if G is embedded in Σ and each cycle C_i bounds a face. We say that a rooted graph G with cycles $C_1, C_2, \ldots, C_k$ embedded in a surface Σ *quadrangulates* the surface Σ if every face of G is either bounded by one of $C_1, C_2, \ldots, C_k$, or has length exactly four.

Theorem 3.1 does not have a straightforward extension to partially colored graphs because of the following topological obstruction. Let $W = v_1 v_2 \ldots v_k v_1$ be a closed walk in a graph G, and let $c : V(W) \to \{1, 2, 3\}$ be a 3-coloring of W. We can view c as mapping $V(W)$ to the vertices of a triangle, and speak of the *winding number of* c, defined as the number of indices $i \in \{1, 2, \ldots, k\}$ such that $c(v_i) = 1$ and $c(v_{i+1}) = 2$ minus the number of indices i such that $c(v_i) = 2$ and $c(v_{i+1}) = 1$, where v_{k+1} means v_1. We denote the winding number of W by $w_c(W)$ or simply $w(W)$ when the coloring is clear from the context. When G is embedded in an orientable surface Σ, then each boundary component of a face of G can be given a direction such that each edge e is traversed once in each direction by the two boundary walks containing e. We deduce the following important fact.

PROPOSITION 3.1. *Let G be a graph embedded in an orientable surface, and let $c : V(G) \to \{1, 2, 3\}$ be a 3-coloring. Then the sum of the winding numbers of all the face boundaries is zero.*

Since the winding number of every properly 3-colored 4-cycle is zero, we deduce the following. If G is a rooted graph with cycles $C_1, C_2, \ldots, C_k$ that quadrangulates an orientable surface Σ, and $c : V(G) \to$

$\{1, 2, 3\}$ is a 3-coloring, then

$$\sum_{i=1}^{k} w_c(C_i) = 0. \tag{3.1}$$

In other words, if c is a 3-coloring of $C_1 \cup C_2 \cup \cdots \cup C_k$, then (3.1) is a necessary condition for the coloring c to extend to a 3-coloring of G. We have shown a partial converse, that if G is embedded with high "representativity" and the cycles $C_1, C_2, \ldots, C_k$ are pairwise "far apart", then the 3-coloring c of $C_1 \cup C_2 \cup \cdots \cup C_k$ extends to a 3-coloring of G if and only if (3.1) holds. To state the result precisely we need a number of definitions.

Let G be a rooted graph with cycles $C_1, C_2, \ldots, C_k$ embedded in a surface Σ, and let $C = C_1 \cup C_2 \cup \cdots \cup C_k$. We say that a subgraph J of G is *essential* if there does not exist a closed disk containing J and at most one of the faces bounded by $C_1, C_2, \ldots, C_k$. We say that the rooted graph G is *w-representative* if every essential subgraph of G has at least w vertices. If $k \geq 2$ or Σ is not the sphere, then we say that the rooted graph G is *boundary-linked* if for every $i = 1, 2, \ldots, k$ and every cycle D in G bounding a disk that includes C_i and no other C_j we have $|V(D)| \geq |V(C_i)|$. We say that G is *strongly boundary-linked* if the inequality is strict for every such cycle D not equal to any C_j for $j = 1, 2, \ldots, k$. If $k = 1$ and Σ is the sphere, then G is *boundary-linked* and *strongly boundary-linked* if there is no path P with both ends, say u and v, on C_1 and otherwise disjoint from C_1, such that P is strictly shorter than both the u-v subpaths of C_1. Finally, if $k = 0$ then G is *boundary-linked* and *strongly boundary-linked* by definition.

Theorem 3.3. *For every orientable surface Σ and for every integer l there exists an integer w such that if G is a w-representative strongly boundary-linked rooted graph of length at most l with cycles $C_1, C_2, \ldots, C_k$ that quadrangulates Σ, then a 3-coloring c of $C_1 \cup C_2 \cup \cdots \cup C_k$ extends to a 3-coloring of G if and only if (3.1) holds.*

When the hypotheses of Theorem 3.3 are satisfied, then (3.1) gives an easy criterion for deciding whether a given precoloring extends to a 3-coloring of G (that is the "easy test" referred to in the abstract). If the rooted graph G is not w-representative, then we cut the surface along an offending essential subgraph to simplify the surface or the rooted graph, and apply the algorithm recursively to test all precolorings of the new cycle(s) resulting from the cutting.

The proof of Theorem 3.3 is algorithmic, but for deciding 3-colorability we only need to know that the theorem holds. The proof proceeds in two steps. If one of the cycles C_i has three consecutive vertices x, y, z such that $c(x) = c(z)$, then we contract all edges incident with y and proceed by induction. That simple argument allows us to reduce to the special case when the colors of the vertices on each cycle C_i are either $1, 2, 3, 1, 2, 3, \ldots$ ("positive cycle") or $1, 3, 2, 1, 3, 2, \ldots$ ("negative cycle"). Condition (3.1) implies that the sum of the lengths of positive cycles is equal to the sum of the lengths of negative cycles. For ease of exposition let us assume for the rest of this paragraph that Σ is the sphere. Let G' be the graph obtained from G by inserting a vertex into each face (except those bounded by the cycles C_i) and joining it to the four vertices on the boundary of the face. We then use the result of [27] to find suitable disjoint paths in the graph G' joining vertices on the positive cycles to like-colored vertices on the negative cycles. We need these paths be even, and that can be arranged using a result from [26]. Let H be the union of $C_1 \cup C_2 \cup \cdots \cup C_k$ and all the paths. For each face f of H the subgraph of G consisting of vertices and edges drawn in the closure of f is bipartite, and hence can be 2-colored. It is possible to combine these 2-colorings into a 3-coloring of G. Further details are omitted due to space limitations.

How about quadrangulations of non-orientable surfaces? Theorem 3.1 does not extend to non-orientable surfaces because of the following result of Youngs [32].

Theorem 3.4. *The chromatic number of every non-bipartite quadrangulation of the projective plane is exactly four.*

An analogue of Theorem 3.1 for non-orientable surfaces was obtained by Mohar and Seymour [16] and Nakamoto, Negami and Ota [18].

Theorem 3.5. *For every non-orientable surface Σ there exists an integer w such that if G is a graph embedded in Σ with all faces even and edge-width at least w, then G is 4-colorable. Furthermore, G is 3-colorable if and only if it has no subgraph H such that H is a quadrangulation of Σ and there is no odd cycle C in G such that cutting open along C produces an orientable surface.*

Our generalization to rooted graphs reads as follows. Let G be a rooted graph with cycles $C_1, C_2, \ldots, C_k$ that quadrangulates a non-orientable surface Σ, let c be a 3-coloring of $C_1 \cup C_2 \cup \cdots \cup C_k$, and let C be a cycle in G such that cutting Σ open along C results in an orientable surface Σ'. In Σ' we can speak of the winding numbers $w_c(C_i)$, and we say that c is *parity-compliant* if $\sum_{i=1}^{k} w_c(C_i)$ (which is even) is congruent to $2|V(C)|$ modulo 4. It can be shown that this notion is independent of the choice of C.

Theorem 3.6. *For every non-orientable surface* Σ *and for every integer* l *there exists an integer* w *such that if* G *is a* w*-representative boundary-linked rooted graph of length at most* l *with cycles* $C_1, C_2, \ldots, C_k$ *that quadrangulates* Σ*, then a* 3*-coloring* c *of* $C_1 \cup C_2 \cup \cdots \cup C_k$ *extends to a* 3*-coloring of* G *if and only if* c *is parity-compliant.*

Our proof of this is similar to the proof of Theorem 3.3.

4 Coloring general embedded graphs

In this section we explain how we reduce 3-coloring general triangle-free graphs on a surface Σ to 3-coloring quadrangulations of Σ. The main tool is the following theorem.

Theorem 4.1. *For every surface* Σ *and every integer* $l \geq 5$ *there exists an integer* M *with the following property. Let* G *be a triangle-free rooted graph of length at most* l *with cycles* $C_1, C_2, \ldots, C_k$ *embedded in* Σ *such that every* 4*-cycle in* G *bounds an open disk disjoint from* $C_1 \cup C_2 \cup \cdots \cup C_k$*. Then* G *has a subgraph* H *containing* $C_1 \cup C_2 \cup \cdots \cup C_k$ *such that*

(i) *a* 3*-coloring of* $C_1 \cup C_2 \cup \cdots \cup C_k$ *extends to a* 3*-coloring of* G *if and only if it extends to a* 3*-coloring of* H*,*

(ii) H *has at most* M *faces of size at least five and each face of* H *has size at most* M*, and*

(iii) *if* Σ *is the sphere and* $k = 1$*, then each face of* H *has size at most* $l - 2$*, and if equality holds, then* H *has exactly one face of size at least five.*

Let us remark that, unfortunately, the hypothesis about 4-cycles bounding disks is necessary.

Proof. Due to space limitations, we only sketch the proof. Let us say that G is *critical* (with respect to $(\Sigma, C_1, C_2, \ldots, C_k)$) if for every proper subgraph G' of G containing $C_1 \cup C_2 \cup \cdots \cup C_k$ some 3-coloring of $C_1 \cup C_2 \cup \cdots \cup C_k$ extends to G' but not to G. We may assume that G is critical, and will show that it satisfies (ii) and (iii). We proceed by induction on minus the Euler characteristic of Σ, subject to that on k, subject to that on l, and subject to that on the number of 4-cycles in G. If G has no 4-cycles, then the conclusion follows from [31, Theorem 6.2]. Otherwise let C be a 4-cycle in G. By hypothesis C bounds a disk Δ disjoint from $C_1 \cup C_2 \cup \cdots \cup C_k$. By [29] every 3-coloring of C extends into the subgraph of G inside Δ, and hence the criticality of G implies that Δ includes no vertices or edges of G. In other words, C is a facial cycle. If C includes vertices of two different cycles C_i and C_j, then we can cut open along $C \cup C_i \cup C_j$ and apply induction to a graph on the same surface but with fewer precolored cycles. A similar reduction applies if two non-consecutive vertices of C belong to the same cycle C_i.

To explain the main idea we will now assume that $k = 1$ and Σ is the sphere, and will indicate which parts of the argument need more care in the general case. We identify two diagonally opposite vertices of C, creating a graph G' and applying induction to it. (If $k \geq 2$ then there is a serious problem in that G' may have a 4-cycle violating the hypothesis, and the same may happen if we try to identify the other pair of diagonally opposite vertices on C. Fortunately, both can happen only in a very restricted way.) Let H' be a critical subgraph of G'; then H' satisfies (ii) and (iii) of the theorem. The graph H' can be converted to a subgraph L of G in a natural way. One face of L is bounded by C, and every other face f of L is in 1-1 correspondence with a face f' of H'. Let $|f|$ denote the length of the walk bounding the face f. We have $|f| \leq |f'| + 2$ for every f, except possibly for one face f that may conceivably satisfy $|f| = |f'| + 4$. However, that cannot happen, because H' is critical, and hence 2-connected. (This argument needs more care when $k \geq 2$.) Thus $|f| \leq |f'| + 2$ for every f.

Let f be a face of L other than the face bounded by C, let J be the subgraph of G drawn in the closed disk bounded by f, and let D be the boundary of f. Then J is critical, as is easily seen, and hence every face of J of length at least five other than the one bounded by D either

- has length at most $|f| - 3 \leq |f'| - 1$, or
- has length exactly $|f| - 2 \leq |f'|$ and it is the unique such face.

Now to deduce the theorem we actually need to prove something slightly stronger, namely that the sequence of face-lengths of a critical graph is "restricted" in a certain sense. However, it is clear from the above argument that if the face-lengths of H' form a restricted sequence, then so do the face-lengths of G (under an appropriate definition of "restricted", which we omit).

We now outline the main steps of the proof of Theorem 1.3.

Proof. Let G be a triangle-free graph embedded in Σ, and let H be initialized to be the null graph. We will proceed in a constant number of iterations, in each iteration enlarging H by adding a bounded size subgraph of G and simplifying the surfaces $\Sigma[f]$ or keeping them the same and simplifying the rooted graphs $G[f]$ (lexicographically over all faces f of H). Let f be a face of H that is not a cylinder, let $G' := G[f]$, let $\Sigma' := \Sigma[f]$ and let $C_1, C_2, \ldots, C_k$ be the components of the boundary of f. We will assume that $C_1, C_2, \ldots, C_k$ are cycles, for that can be arranged by splitting vertices of $C_1 \cup C_2 \cup \cdots \cup C_k$. Then G' is a rooted graph with cycles $C_1, C_2, \ldots, C_k$ embedded in Σ'.

Let w be an integer chosen so that Theorems 3.3 and 3.6 hold for Σ' and G'. If G' is not w-representative, then we add an offending essential subgraph to H and go to the next iteration. Thus we may assume that G' is w-representative, and similarly we may assume that it is strongly boundary linked (here we need that f is not a cylinder). In particular, we may assume that every 4-cycle in G' bounds an open disk disjoint from $C_1 \cup C_2 \cup \cdots \cup C_k$. Thus we may apply Theorem 4.1 to G'; let H' be the subgraph guaranteed therein. We may assume that some 3-coloring of $C_1 \cup C_2 \cup \cdots \cup C_k$ does not extend to G', for otherwise condition (i) of the theorem holds. Thus $H' \neq C_1 \cup C_2 \cup \cdots \cup C_k$. If all the faces of H' other than those bounded by $C_1, C_2, \ldots, C_k$ have length four, then H' witnesses that $G[f]$ satisfies condition (iii) of the theorem. Otherwise we add to H the face boundaries of all faces of H' of size at least five and proceed to the next iteration. By Theorem 4.1(ii) we are adding a graph of bounded size, and by Theorem 4.1(iii) it can be shown that this either simplifies Σ', or decreases k or the length of one the precolored cycles.

5 Linear-time algorithm

Let Σ be a fixed surface. It is easy to convert the proof of Theorem 1.3 to a polynomial-time algorithm to find the desired subgraph H. Furthermore, for each face f of H such that (iii) holds, the algorithm also finds the corresponding subgraph of $G[f]$ that quadrangulates $\Sigma[f]$. Once the subgraph H is found, we test all 3-colorings of H to check whether one of them extends into each face of H, as described in the paragraph immediately following Theorem 1.3. That part is relatively straightforward to implement in linear time.

It is somewhat less obvious how to find the subgraph H from Theorem 1.3 in linear time. Here is the main idea. We will be constructing the graph H in constantly many iterations, similarly as in the proof of Theorem 1.3. Let f be a face of H that is not a cylinder, and let G', Σ' and $C_1, C_2, \ldots, C_k$ be as before. If G' is not w-representative or strongly boundary linked, where w is as in Theorem 3.3 or 3.6, then we add the corresponding offending subgraph to H and go to the next iteration. Let G'' be initialized to be G'; we will proceed in constantly many iterations (for fixed f), shrinking G'' in each iteration, but letting it include $C_1 \cup C_2 \cup \cdots \cup C_k$ at all times.

We need a couple of definitions. Let C be a cycle in G'' bounding an open disk disjoint from $C_1 \cup C_2 \cup \cdots \cup C_k$. We say that C is *free* if there is no cycle $D \neq C$ in G'' of length at most that of C such that D bounds a closed disk in Σ' that includes C and has interior disjoint from $C_1 \cup C_2 \cup \cdots \cup C_k$. Given a cycle C of bounded length, it is easy to find (using bounded number of iterations of the augmenting paths algorithm) a free cycle of length at most the length of C that bounds a disk that includes C. We say that a cycle C in G'' is *extendable* if every 3-coloring of C extends to a 3-coloring of the subgraph of G drawn in the disk bounded by C. We will maintain the property that every face of G'' is either a face of G or is bounded by a free extendable cycle.

We will now construct a collection $\mathcal{C}$ of free extendable cycles of length at least five bounding disks with disjoint interiors. Suppose for a moment that some proper subset $\mathcal{C}'$ of $\mathcal{C}$ of size at least two can be separated off from $C_1 \cup C_2 \cup \cdots \cup C_k$ by a cycle D of length at most $g(\Sigma', |\mathcal{C}'|)$ bounding an open disk Δ, where the function g is to be chosen carefully. Then we may assume that D is free, by replacing it by a different cycle, if necessary. We delete from G'' all vertices and edges lying in Δ, replace $\mathcal{C}$ by $(\mathcal{C} - \mathcal{C}') \cup \{D\}$, and return for the next iteration. Let M be as in Theorem 4.1. This process lexicographically increases the sizes of cycles in $\mathcal{C}$, and hence after constantly many iterations it either

(i) runs out of facial cycles of length at least five, or

(ii) it finds a free cycle of length at least $M + 1$, or

(iii) it constructs a family $\mathcal{C}$ as above of size at least $M + 1$.

If (i) holds, then G'' is the subgraph of $G[f]$ desired for Theorem 1.3(iii), and if (ii) or (iii) holds, then it can be shown, using Theorem 4.1, that for some $i = 1, 2, \ldots, k$ the cycle C_i is irrelevant in the sense that a 3-coloring c of $C_1 \cup C_2 \cup \cdots \cup C_k$ extends to a 3-coloring of G if and only if the restriction of c to $C_1 \cup C_2 \cup \cdots C_{i-1} \cup C_{i+1} \cup \cdots \cup C_k$ extends to a 3-coloring of G. Thus if $k = 1$ then Theorem 1.3(i) holds, and otherwise we repeat the same algorithm for the graph G and cycles $C_1, C_2, \ldots C_{i-1}, C_{i+1}, \ldots, C_k$. This algorithm can be implemented to run in linear time but we omit details of its implementation due to space constraints.

6 Finding a 3-coloring

For a fixed surface Σ, we have described a linear-time algorithm that decides whether an input triangle-free graph embedded in Σ is 3-colorable. When combined with earlier results this gives a linear-time algorithm to compute the chromatic number of an input triangle-free graph embedded in Σ.

However, finding a 3-coloring in linear time (when it exists) seems to be harder. A difficulty arises already in the plane. By Theorem 1.1 every triangle-free planar graph is 3-colorable, but it was not known how to find a 3-coloring in linear time. The best previous result was due to Kowalik [14], who designed an $O(n \log n)$ algorithm to 3-color triangle-free planar graphs. Recently, two of us in joint work with Kawarabayashi [3] managed to find a linear-time algorithm to 3-color planar graphs.

For a general surface, we can find a 3-coloring of a quadrangulation in linear time, when it exists. There are two particular subproblems that we had to solve on our way to a linear time algorithm and which we address in the next two sections. In the proof of Theorem 3.3, we had to find disjoint paths in the auxiliary graph G' between the precolored vertices of the embedded graph. Robertson and Seymour [27] gave a polynomial-time algorithm to solve the disjoint connecting paths problem for graphs on a fixed surface. We speed up the running time of their algorithm to linear time in the next section.

Regarding the complexity of finding a 3-coloring of a triangle-free graph on a fixed surface, it is possible to design a quadratic algorithm based on the proofs of our structural theorems, and we believe that our approach also leads to a linear-time algorithm.

7 The disjoint tree algorithm

Let G be a rooted graph with cycles $C_1, C_2, \ldots, C_k$ embedded in a surface Σ. Let Σ' be the surface with boundary obtained from Σ by removing the interiors of the faces bounded by the cycles $C_1, C_2, \ldots, C_k$. Thus the cycles $C_1, C_2, \ldots, C_k$ form the boundary of Σ' and will be referred to as *cuffs*. We say that two trees S and T (not necessarily subgraphs of G) are *homotopic* if their intersection with the cuffs is the same and for every two points s, t that belong to the trees and also to some cuff, the unique s-t path in S is homotopic in Σ' to the unique s-t path in T. The following is the result from [27] we are using:

THEOREM 7.1. *For every surface Σ and for every integer ℓ there exists an integer $w(\Sigma, \ell)$ such that if G is a boundary-linked rooted graph of length at most ℓ with representativity at least $w(\Sigma, \ell)$ and $S_1, \ldots, S_r$ are disjoint trees embedded in Σ with leaves being distinct vertices on the cuffs, then there exists a homeomorphism φ of Σ preserving the cuffs such that G contains disjoint trees $T_1, \ldots, T_r$ and every tree T_i is homotopic to the tree $\varphi(S_i)$.*

Let us now state formally the algorithmic problems under consideration, whose names we further shorten to the DTP and the DTTP.

DISJOINT TREE PROBLEM

Input:	a rooted graph G embedded in a surface Σ subsets $X_1, \ldots, X_r$ of vertices on the cuffs
Output:	pairwise-disjoint trees $T_1, \ldots, T_r$ of G such that T_i contains the vertices of X_i, or a statement that no such trees exist

DISJOINT TOPOLOGICAL TREE PROBLEM

Input:	a rooted graph G embedded in a surface Σ pairwise disjoint trees $S_1, \ldots, S_r$ in Σ with leaves on the cuffs
Output:	pairwise-disjoint trees $T_1, \ldots, T_r$ of G such there is a homeomorphism φ of Σ' fixing the cuffs and T_i is homotopic to $\varphi(S_i)$, or a statement that no such trees exist

The trees S_i in the DTTP are given as subgraphs of a supergraph of a subdivision of G.

Let us recall two algorithmic results which are particular cases of Theorem 7.4.

THEOREM 7.2. (SUZUKI ET AL. [28]) *Let G and Σ be such that Σ' is the plane with at most two cuffs, and let λ be a fixed integer. Then the DTP can be solved in linear time in the order of G for inputs satisfying $|X_1 \cup \cdots \cup X_r| \leq \lambda$.*

THEOREM 7.3. (REED ET AL. [23]) *If Σ is the plane and λ is a fixed integer, then the DTP can be solved in linear time in the order of G for inputs satisfying $|X_1 \cup \cdots \cup X_r| \leq \lambda$.*

Though not explicitly stated in [23], their algorithm can be adapted to solve the DTTP in linear time for the plane with cuffs of bounded length.

Our theorem states that the DTP can be solved in linear time for an arbitrary surface. This result was also announced by the authors of [23] but never published; see also [22], which is an enhanced version of [23]. We also show that the DTTP can be solved in linear time. Problems related to the DTP and the DTTP are addressed in a recent paper [13], where the authors obtain linear-time algorithms. Unlike previous results, our approach is not based on the existence of a deletable vertex.

THEOREM 7.4. *If Σ is a fixed surface and ℓ is a fixed integer, then the DTP and the DTTP can be solved in linear time in the order of G for rooted graphs G of length at most ℓ.*

Proof. Our algorithm simultaneously solves the DTTP for all possible non-equivalent trees. Observe that there is only a bounded number of such non-equivalent trees joining them as Σ and ℓ are fixed. Since from the solution to the DTTP, we can read off a solution to the DTP (or lack thereof), it suffices to describe our algorithm for the DTTP.

Let $w(\Sigma, \ell)$ be the function from Theorem 7.1. Assume that we have already designed linear-time algorithms for surfaces Σ' of smaller genera and for the surface Σ with fewer cuffs, each of length at most $\ell + 2w(\Sigma, \ell)$.

Let G be the input graph and N the number of its vertices. We first test whether G is strongly-boundary linked and has representativity at least $w(\Sigma, \ell)$—if not, the problem is split into two smaller problems. Hence, G is a strongly boundary-linked graph with representativity at least $w(\Sigma, \ell)$. At this point, we already know that the problem always has a solution by Theorem 7.1.

Note that contracting some edges of G as long as G stays strongly boundary-linked and has representativity at least $w(\Sigma, \ell)$ does not change the fact that the problem has a solution. In the next section, we will present a linear-time algorithm (see Theorem 8.3) that shrinks the input graph G by contracting some of its edges either to a graph with αN vertices for $\alpha < 1$ that is strongly boundary-linked and has representativity at least $w(\Sigma, \ell)$, a graph that is boundary linked but not strongly boundary linked, or a graph that has an essential subgraph with $w(\Sigma, \ell)$ vertices. The latter two outcomes allow us to cut the surface and apply a recursive argument. In case of the former outcome, we keep applying the shrinking algorithm. Since each time, the number of vertices is decreased by a linear factor, the total running time of our algorithm remains linear.

8 Shrinking the input graph

We now present a generalization of a data structure of Kowalik and Kurowski for finding short paths in planar graphs. The data structure is formulated in the framework of classes of graphs with bounded expansion defined by Nešetřil and Ossona de Mendez [19, 20, 21]. Such classes of graphs include all proper minor-closed classes; in particular, graphs embedded on surfaces which are of interest to us.

The *grad with rank* r of a graph G is equal to the largest average degree of a graph G' that can be obtained from G by removing some of the vertices and then contracting subgraphs of radius at most r to single vertices (arising parallel edges are removed). The grad with rank r of G is denoted by $\nabla_r(G)$. A class $\mathcal{G}$ of graphs is a said to have *bounded expansion* if there exist absolute constants $\nabla_1, \nabla_2, \ldots$ such that $\nabla_r(G) \leq \nabla_r$ for every $G \in \mathcal{G}$. It is straightforward to show that any proper minor-closed class of graphs has bounded expansion.

Fix a class $\mathcal{G}$ of graphs with bounded expansion and let ∇_r be the supremum of $\nabla_r(G)$ for $G \in \mathcal{G}$. The edges of G can be oriented in such a way that the in-degree of every vertex is at most ∇_1. The results obtained by Nešetřil and Ossona de Mendez [20] in their series of papers imply the following.

Theorem 8.1. *There exists a function f with the following property. Let G be a graph with edges oriented in such a way that each vertex has in-degree at most D. Let G' be the graph obtained by adding all edges xy such that:*

- *there exists a vertex z such that G contains an edge oriented from x to z and an edge oriented from z to y, or*
- *there exists a vertex z such that G contains an edge oriented from x to z and an edge oriented from y to z.*

The grad with rank r of G' is at most $f(\nabla_r(G), D)$.

Consider the following series of graphs: G_0 is the graph G with edges oriented in such a way that the maximum in-degree of G_0 is at most $\nabla_0(G)$. The graph G_1 is obtained from the graph G_0 by adding the edges between vertices x and y with a common out-neighbor. The new edges are oriented in such a way that the maximum in-degree of G_1 is at most $\nabla_0(G) + \nabla_0(G')$; the edges present in G_0 preserve their orientation. In general, G_k is the graph obtained from G_{k-1} by adding edges between vertices with a common out-neighbor; the new edges are oriented in such a way that maximum in-degree is bounded by $\sum_{i=0}^{k} \nabla_0(G_i)$.

Kowalik and Kurowski [15] observed the above mentioned property for planar graphs and designed an algorithm for planar graphs that answers queries whether two vertices are joined by a path of length at most K_0 in constant time with a linear preprocessing time (when K_0 is fixed). Let us now sketch the idea of their algorithm.

Assume first we want to answer a query whether two vertices are joined by a path of length at most two. Construct the oriented graphs G_0 and G_1 that were described before. Let D_i be the maximum in-degree of a vertex in G_i, $i = 0, 1$. Checking the existence of a path of length one between two vertices u and v is easy: it is enough to verify whether u is one of the at most D_0 in-neighbors of v or v is one of the at most D_0 in-neighbors of u. If neither is the case, then there is no such path.

Checking the existence of a path of length two is little bit more involved. Let w be the middle vertex of such a path. If the edges uw and vw are both oriented to w, G_1 contains an edge corresponding to a path of length two between u and v. If both edges uw and vw are oriented from w, it is enough to check whether u

and v have a common in-neighbor in G_0 which can be done in time bounded by D_0. Finally, if the edge uw is oriented from u to w and the edge wv from w to v (or they both are oriented in the opposite way), it is enough to check whether u is an in-neighbor of an in-neighbor of v. Since there are at most D_0^2 such in-neighbors, this can be done in time $O(D_0^2)$.

It is possible to modify the above process for checking paths of length at most K_0 by considering the graph G_{K_0-1} and assigning to its edges weights corresponding to their "lengths". An analogous approach yields an algorithm for queries whether two vertices are joined by a path of length at most K_0 (K_0 is constant) with weight less than a given value. We extend this approach and obtain the following:

THEOREM 8.2. *Let $\mathcal{G}$ be a class of graphs with bounded expansion and K_0 a fixed integer. There exists an algorithm that given an N-vertex graph $G \in \mathcal{G}$ with integral weights of edges between* 0 *and* K_0

- *preprocesses the graph in time $O(N)$,*
- *answers a query whether two vertices are joined by a path of at most given weight and at most given length less than K_0 in $O(1)$ time,*
- *allows changes of weights of edges and updates its data structures in $O(1)$ time for a change of weight of any edge, and*
- *allows deletions of edges and updates the data structures in time $O(1)$ after a deletion of an edge.*

Our next aim is using the algorithm described in Theorem 8.2 to design an algorithm for finding short non-trivial cycles in graphs embedded in surfaces. The idea is similar to that used by Cabello and Mohar [2]. Fix a graph G embedded on the surface Σ and K_0 which is the length of non-trivial cycles whose existence we want to test. We assume that the input graph is embedded in the fundamental polygon of the surface in such a way that each edge crosses each edge of the polygon at most twice (such an embedding can be obtained from any other data structure describing an embedding of G in linear time).

Consider several copies of the embedding and glue them together along the sides of the polygon obtaining a tree-like structure (the copies of the embedding form a part of the universal cover of Σ). We consider as many copies of the embedding such that any non-trivial cycle of length at most K_0 is represented by a path between two copies of some of its vertices (since K_0 and Σ is fixed and every edge crosses each edge of the fundamental polygon at most twice, the number of copies needed is constant). Moreover, the way in which the copies should be glued together is determined by the fundamental group of Σ (in particular, the number of copies we consider is bounded), we can wire the recipe how to glue the copies into our algorithm.

Answering a question whether there exists a cycle of length at most K_0 in G with a given homotopy is easy—just check the existence of a path between a vertex u and its clone in the appropriate copy of G whose length at most K_0. Using the approach described before Theorem 8.2, the test for a short path of a given homotopy can be done in $O(N)$ time for a fixed surface Σ and K_0. In our application, we consider graphs G embedded in surfaces Σ with cuffs. To cope with this, we cut along curves from the cuffs to the boundary of the fundamental polygon and obtain a more complex surface which captures the closed curves surrounding the cuffs. When testing representativity, we ignore the non-trivial cycles arising just because of wrapping around a single cuff (since the surface is fixed and the number of cuffs is bounded, this can be wired into the algorithm).

THEOREM 8.3. *Let Σ be a fixed surface and L and w fixed integers. There exists α, $0 < \alpha < 1$, and a linear-time algorithm that given a boundary-linked N-vertex rooted graph G embedded in Σ with length at most L and representativity at least w, constructs boundary-linked graph G' with representativity at least w by contracting some of its edges and that*

- *is strongly boundary-linked and has less than αN vertices, or*
- *is not strongly boundary-linked in which case the algorithm outputs a closed curve surrounding a cuff with the same size, or*
- *has an essential subgraph with w vertices in which case the algorithm outputs such a subgraph.*

Proof. Let H be the graph obtained from G by adding a vertex inside each face and joining it to the vertices on that face. Since H is still embeddable on Σ, it is possible [21] to properly color the vertices of H in linear time with K colors, where K depends only on Σ, in such a way that the union of every two color classes induces a star forest in G. Consider now two color classes V_1 and V_2 that induce the largest number of edges contained in G (not necessarily of H). This number is linear in N as the number of color classes is bounded.

We use the algorithm given in Theorem 8.2 for H with the initial weights of edges of G equal to two and the weights of the new edges equal to one. Observe that H has no non-contractible cycle of length at most 2ℓ iff the representativity of G is at least ℓ. We construct the part of the universal cover as outlined before. One after another, we contract the edges of G with end-vertices in V_1 and V_2. This is achieved by setting the weight of the contracted edge to zero. After contracting each edge, we test (in constant time) for the existence of a short cycle through the contracted edge bounding a cuff or an essential subgraph. If no such subgraph is discovered, the order of G is shrunk by a linear factor after contracting all the edges between V_1 and V_2.

References

[1] J. A. Bondy and U. S. R. Murty, *Graph theory with applications*, North-Holland, New York, Amsterdam, Oxford, 1976.

[2] S. Cabello and B. Mohar, *Finding shortest non-separating and non-contractible cycles for topologically embedded graphs*, Discrete Comput. Geom., 37 (2007), pp. 213–235.

[3] Z. Dvořák, K. Kawarabayashi and R. Thomas, *Three-coloring triangle-free planar graphs in linear time*, to appear in the proc. of SODA 2009.

[4] D. Eppstein, *Subgraph isomorphism in planar graphs and related problems*, J. Algor. and Appl., 3 (1999), pp. 1–27.

[5] S. Fisk, *The nonexistence of colorings*, *J. Combin. Theory Ser. B*, 24 (1978), pp. 247–248.

[6] T. Gallai, *Kritische Graphen I*, Publ. Math. Inst. Hungar. Acad. Sci., 8 (1963), pp. 165–192.

[7] T. Gallai, *Kritische Graphen II*, Publ. Math. Inst. Hungar. Acad. Sci., 8 (1963), pp. 373–395.

[8] M. R. Garey and D. S. Johnson, *Computers and intractability. A guide to the theory of NP-completeness*, W. H. Freeman, San Francisco, 1979.

[9] J. Gimbel and C. Thomassen, *Coloring graphs with fixed genus and girth*, *Trans. Amer. Math. Soc.*, 349 (1997), pp. 4555–4564.

[10] H. Grötzsch, *Ein Dreifarbensatz für dreikreisfreie Netze auf der Kugel*, Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe, 8 (1959), pp. 109–120.

[11] J. Hutchinson, *Three-coloring graphs embedded on surfaces with all faces even-sided*, *J. Combin. Theory Ser. B*, 65 (1995), pp. 139–155.

[12] P. J. Heawood, *Map-color theorem*, Quart. J. Pure Appl. Math., 24 (1890), pp. 332–338.

[13] K. Kawarabayashi and K. Kobayashi, *An algorithm for finding an induced cycle in planar graphs and bounded genus graphs*, to appear in the proc. of SODA 2009.

[14] Ł. Kowalik, *Fast 3-coloring triangle-free planar graphs*, ESA 2004, Lecture Notes in Comput. Sci., 3221 (2004), pp. 436–447.

[15] L. Kowalik, M. Kurowski: *Short path queries in planar graphs in constant time*, STOC 2003, pp. 143–148.

[16] B. Mohar and P. D. Seymour, *Coloring locally bipartite graphs on surfaces*, *J. Combin. Theory Ser. B*, 84 (2002), pp. 301–310.

[17] B. Mohar and C. Thomassen, *Graphs on surfaces*, Johns Hopkins University Press, Baltimore, MD, 2001.

[18] A. Nakamoto, S. Negami and K. Ota, *Chromatic numbers and cycle parities of quadrangulations on nonorientable closed surfaces*, *Discrete Math.*, 285 (2004), pp. 211–218.

[19] J. Nešetřil, P. Ossona de Mendez: *Linear time low tree-width partitions and algorithmic consequences*, STOC 2006, pp. 391–400.

[20] J. Nešetřil, P. Ossona de Mendez: *Grad and classes with bounded expansion I. Decompositions*, Europ. J. Combin., 29 (2008), pp. 760–776.

[21] J. Nešetřil, P. Ossona de Mendez: *Grad and classes with bounded expansion II. Algorithmic aspects*, Europ. J. Combin., 29 (2008), pp. 777–791.

[22] B. Reed, *Rooted routing in the plane*, Discrete Appl. Math., 57 (1995), pp. 213–227.

[23] B. A. Reed, N. Robertson, A. Schrijver and P. D. Seymour, *Finding disjoint trees in planar graphs in linear time*, Contemp. Math., 147 (1993), pp. 295–301.

[24] G. Ringel, *Map Color Theorem*, Springer-Verlag, Berlin, 1974.

[25] G. Ringel and J. W. T. Youngs, *Solution of the Heawood map-coloring problem*, Proc. Nat. Acad. Sci. U.S.A., 60 (1968), pp. 438–445.

[26] N. Robertson and P. D. Seymour, *Graph Minors VI. Disjoint paths across a disc*, *J. Combin. Theory Ser. B*, 41 (1986), pp. 115–138.

[27] N. Robertson and P. D. Seymour, *Graph Minors VII. Disjoint paths on a surface*, *J. Combin. Theory Ser. B*, 45 (1988), pp. 212–254.

[28] H. Suzuki, T. Akama and T. Nishiseki, *An algorithm for fidning a forest in a planar graph—case in which a net may have terminals on the two specified face boundaries*, Electron. Comm. Japan Part III Fund Electron. Sci, 72 (1989), pp. 68–79.

[29] C. Thomassen, *Grötzsch's* 3*-color theorem and its counterparts for the torus and the projective plane*, *J. Combin. Theory Ser. B*, 62 (1994), pp. 268–279.

[30] C. Thomassen, *Color-critical graphs on a fixed surface*, *J. Combin. Theory Ser. B*, 70 (1997), pp. 67–100.

[31] C. Thomassen, *The chromatic number of a graph of girth 5 on a fixed surface*, *J. Combin. Theory Ser. B*, 87 (2003), pp. 38–71.

[32] D. A. Youngs, 4*-chromatic projective graphs*, *J. Graph Theory*, 21 (1996), pp. 219–227.