ELSEVIER
Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®
Theoretical Computer Science 299 (2003) 179–210
Theoretical Computer Science
www.elsevier.com/locate/tcs

Structuring the elementary components of graphs having a perfect internal matching

Miklós Bartha*,[1], Miklós Krész[2]

Department of Computer Science, Memorial University of Newfoundland St. John's, Newfoundland, Canada A1B 3X5

Received 22 May 2001; accepted 29 October 2001
Communicated by G. Ausiello

Abstract

Graphs with perfect internal matchings are decomposed into elementary components, and these components are given a structure reflecting the order in which they can be reached by external alternating paths. It is shown that the set of elementary components can be grouped into pairwise disjoint families determined by the "two-way accessible" relationship among them. A family tree is established by which every family member, except the root, has a unique father and mother identified as another elementary component and one of its canonical classes, from which the given member is two-way accessible. It is proved that every member of the family is only accessible through a distinguished canonical class of the root by external alternating paths. The families themselves are arranged in a partial order according to the order they can be covered by external alternating paths, and a complete characterization of the graph's forbidden and impervious edges is elaborated. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Graph matching; Alternating path; Elementary graph; Canonical partition

1. Introduction

The results reported in this paper contribute to the research on soliton automata, which has been active for more than a decade now. The aim of this research is to

* Corresponding author.
E-mail address: bartha@cs.mun.ca (M. Bartha).

[1] Partially supported by Natural Science and Engineering Research Council of Canada Operating Grant 335591, and by National Scientific Research Fund of Hungary Operating Grant T 014202.
[2] Author on leave from Juhász Gyula Teacher Training College, University of Szeged, Hungary. Work partially supported by the International G. Soros Foundation and by the APPOL 14084 Thematic Network project within the Fifth European Community Framework Program (FP5).

PII: S0304-3975(02)00048-8

explore the mathematical aspects of some molecular switching devices suggested in [6]. Although the possibility of actually building a molecular computer is still rather remote, results on soliton automata are very encouraging.

The underlying object of a soliton automaton is a so-called open graph, which is simply an undirected graph having at least one vertex with degree 1. Such vertices are called external, and their role is to provide an interface by which soliton automata can be controlled from the outside world. The states of a soliton automaton are matchings of the underlying graph that cover all vertices, except possibly the external ones. Such a matching is called a perfect internal matching, and a graph having a perfect internal matching is a soliton graph. A state change of a soliton automaton is carried out by selecting an alternating walk connecting two external vertices of the underlying graph with respect to the current state, and exchanging the status of each edge along the walk regarding its being present or not present in that state (i.e., perfect internal matching). The reader is referred to [7,8,9,10,11] for some early results on soliton automata.

The concept "perfect internal matching" emerged directly from the study of soliton automata, therefore relatively little has been done so far in order to adopt even the most fundamental results in matching theory [12] on perfect matchings, and find out their usefulness concerning soliton automata. Tutte's well-known theorem on the structure of maximum matchings has been generalized in [1] for maximum internal matchings, and the counterpart of the Gallai–Edmonds structure theorem was worked out in [2]. Other basic results directly related to soliton automata have been gathered in [3]. An algebraic approach to study open graphs and perfect internal matchings has been outlined in [2,4].

The present paper makes a significant step towards the decomposition of soliton automata into elementary ones. For technical reasons, namely space restrictions, the actual decomposition has been carried out in a separate paper [5], while this work concentrates on matching theoretic issues only. Implications on soliton automata are never spelled out, although the traces of these implications should be conspicuous even for a reader completely unfamiliar with soliton automata.

The elementary components of soliton graphs are grouped into pairwise disjoint families based on the so-called "two way accessible" relationship among them. A family tree is then established in each of these groups, which reflects the order in which family members can be reached by external alternating paths. In addition, a complete characterization of the graph's impervious and forbidden edges is given.

The paper is organized as follows. Section 2 introduces the notation and terminology relating to graphs and matchings, and puts forward four simple claims for the sections to follow. Section 3 provides a link between perfect matchings and perfect internal matchings by elaborating a framework in which the latter constructs can be studied in terms of the former ones. Section 4 introduces hidden edges to soliton graphs, and shows that the addition of these edges does not change the elementary decomposition of the graph. The grouping of elementary components into families is carried out in Section 5. It is proved that each viable family contains a unique one-way component, called the root, and that all external alternating paths targeting any member in that family must enter the family at the principal canonical class of the root. Section 6 establishes a family tree within each viable family, characterizes forbidden edges connecting elementary components inside a family and between two viable families, and identifies impervious edges of the graph as ones that are incident with a viable family only at vertices belonging to the principal canonical class of the root of that family. Finally, Section 7 is a summary of the results obtained.

2. Preliminaries

In this section we review some of the basic concepts concerning graphs and matchings, and state a few claims that will often be used in later sections. Our notation and terminology will be compatible with that of [12], except that "point" and "line" will be replaced by the more conventional terminology "vertex" and "edge", respectively.

By a graph we shall mean a finite undirected graph in the most general sense, i.e., with multiple edges and loops allowed. For a graph G, $V(G)$ and $E(G)$ will denote the set of vertices and the set of edges of G, respectively. An edge $e\in E(G)$ connects two vertices $v_1, v_2\in V(G)$, which are said to be *adjacent* in G. The vertices v_1 and v_2 are called the *endpoints* of e, and we say that e is *incident with* v_1 and v_2. If $v_1=v_2$, then e is called a *loop* around v_1.

The *degree* of a vertex v in graph G is the number of occurrences of v as an endpoint of some edge in $E(G)$. According to this definition, every loop around v contributes two occurrences to the count. The vertex v is called *external* if its degree is one, *internal* if its degree is greater than one, and *isolated* otherwise. An edge $e\in E(G)$ is said to be an *external edge* if one of its endpoints is an external vertex. *Internal edges* are those that are not external. The sets of external and internal vertices of G will be denoted by $\mathrm{Ext}(G)$ and $\mathrm{Int}(G)$, respectively. Graph G is said to be *open* if it has at least one external vertex, and G is *closed* if all vertices of G are internal.

A *matching* M of graph G is a subset of $E(G)$ such that no vertex of G occurs more than once as an endpoint of some edge in M. Again, it is understood by this definition that loops are not allowed to participate in M. The endpoints of the edges contained in M are said to be *covered* by M. A matching is called *perfect* if it covers all vertices of G. A *perfect internal matching* is one that covers all of $\mathrm{Int}(G)$. Clearly, the notions perfect matching and perfect internal matching coincide for closed graphs.

By the usual definition, a *subgraph* G' of G is just a collection of vertices and edges of G. Since in our treatment we are particular about external vertices, we do not want to allow that new external vertices (i.e., ones that are not present in G) emerge in G'. Therefore, whenever this happens, so that vertex $v\in\mathrm{Int}(G)$ becomes external in G', we shall augment G' with a loop edge around v. This augmentation will be understood automatically in all subgraphs of G. The subgraph of G determined by a set of vertices $X\subseteq V(G)$ will be denoted by $G[X]$, or just by $[X]$ if G is understood.

Assume, for the rest of this section, that G is a graph having a perfect internal matching. An edge $e\in E(G)$ is called *allowed* if e is part of some perfect internal matching of G, and e is *forbidden* if this is not the case. Edge e is *mandatory* if it is present in all perfect internal matchings of G, and e is *constant* if it is either forbidden or mandatory. Graph G is *elementary* if its allowed edges form a connected subgraph covering all the external vertices, and G is *1-extendable* if all of its edges, except the loops if any, are allowed.

A subgraph G' of G is *nice* if it has a perfect internal matching, and every perfect internal matching of G' can be extended to a perfect internal matching of G. In this case, a perfect internal matching of G is *G'-permissible* if it is the extension of an appropriate perfect internal matching of G'. Obviously, not all perfect internal matchings of G must be G'-permissible. Take, for example, a single non-constant internal edge e in G (with a loop around both endpoints) as G'. Clearly, G' is nice, but a perfect internal matching M of G is G'-permissible iff $e \in M$.

In general, the subgraph of G determined by its allowed edges has several connected components, which are called the *elementary components* of G. An elementary component C is *external* if it contains external vertices of G, otherwise C is *internal*. Notice that an elementary component can be as small as a single external vertex of G. Such a component is the only exception from the general rule that each elementary component is an elementary graph. A *mandatory elementary component* is a single mandatory edge $e \in E(G)$ with a loop around one or both of its endpoints, depending on whether e is external or internal. Note that an edge connecting two external vertices is not mandatory in G, therefore it is not a mandatory elementary component either.

A *walk* in graph G is an alternating sequence of vertices and edges, starting and ending with a vertex, such that each edge in the sequence is incident with the vertex immediately preceding and following it. A *trail* is a walk in which no edge occurs more than once, and a *path* is a trail with no repetition in the sequence of vertices. A *cycle* is a trail that returns to its starting point after covering a path, and then stops. A trail is called *external* if one of its endpoints is such, otherwise the trail is *internal*.

Let M be a perfect internal matching of G. A trail $\alpha = v_0, e_1, \ldots, e_n, v_n$ is *alternating* with respect to M (or M-alternating, for short) if for every $1 \leqslant i \leqslant n-1$, $e_i \in M$ iff $e_{i+1} \notin M$. Notice that an alternating trail can return to itself only at its endpoints. Therefore we shall specify alternating trails just by giving the set of their edges, indicating the starting point and other particulars of the trails only in words if this is necessary. If α is an M-alternating path and $e_1 \in M$ ($e_1 \notin M$), then we say that α is *positive* (respectively, *negative*) at its v_0 end. An external alternating path leading to an internal vertex is positive (negative) if it is such at its internal endpoint. An internal alternating path is positive (negative) if it is such at both ends. A *positive M-alternating fork* is a pair of disjoint positive external M-alternating paths leading to two different internal vertices. Although it sounds somewhat confusing, we say that a positive alternating fork *connects* its two internal endpoints.

A perfect internal matching of G is often called a *state*. For any state M, an M-alternating path connecting two external vertices of G is called a *crossing*. An *M-alternating loop* around vertex v is an odd M-alternating cycle starting from v. Clearly, the first and the last edge of any M-alternating loop must not be in M. Since we now have a distinct name for odd alternating cycles, we shall reserve the term "alternating cycle" for even length ones. An *M-alternating unit* α is either a crossing or an (even length) alternating cycle with respect to M. *Making* the unit α in state M means creating a new state $M' = S(M, \alpha)$ in which for every edge e in α, $e \in M'$ iff $e \notin M$, and for every edge e not in α, $e \in M'$ iff $e \in M$. It is easy to see that M' is indeed a state. An *M-alternating network* Γ is a set of pairwise disjoint M-alternating units. Again, by making Γ in state M we mean creating a new state $S(M, \Gamma)$ by making

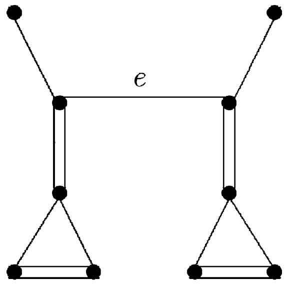

Fig. 1. An impervious edge e.

the units in Γ one by one in an arbitrary order. It was proved in [4] that for every two states M and M' there exists an M-alternating network Γ such that $M' = S(M, \Gamma)$ and $M = S(M', \Gamma)$. This network Γ is called the *mediator* alternating network between states M and M'. An immediate consequence of this result is that an edge e is not constant iff there exists an M-alternating unit passing through e in every state M.

An internal vertex v of G is called *accessible* in state M if there exists a positive external M-alternating path leading to v. An edge e is *impervious* in state M if neither of its endpoints are accessible in M. Edge e is *viable* if it is not impervious. See Fig. 1 for a graph containing an impervious edge e. In this figure, as well as in some of the forthcoming ones, double lines connecting two vertices indicate edges that belong to the given matching M.

Claim 2.1. *An internal vertex v is accessible in state M iff v is accessible in all states of G.*

Proof. Let us augment G by a new external edge at v, that is, by an edge $e = (v, v')$, where $v' \notin V(G)$. If $G + e$ denotes the augmented graph, then $G + e$ still has a perfect internal matching, moreover, G is a nice subgraph of $G + e$. Obviously, there is only one way to extend any perfect internal matching of G to $G + e$, i.e., by not including the edge e in that matching. We shall therefore identify each state of G by its unique extension to $G + e$. By assumption, there exists an M-alternating crossing α in $G + e$ passing through the edge e. Consider the state $S(M, \alpha)$, and switch to any G-permissible state M' of $G + e$ by making the mediator alternating network Γ between $S(M, \alpha)$ and M'. It is clear that Γ contains a unique crossing β going through e. Stripping β from the edge e results in the desired positive external M'-alternating path in G leading to vertex v. $\square$

By virtue of Claim 2.1 we can say that an internal vertex v is *accessible* in G without specifying the state M relative to which this concept was originally defined.

Corollary 2.2. *An edge e is impervious in some state of G iff e is impervious in all states of G.*

Claim 2.3. *Every internal vertex of an open elementary graph G is accessible.*

Proof. It was proved in [4] that, for every two allowed edges e_1, e_2 of an elementary graph, there exists a state M such that both e_1 and e_2 are contained in an appropriate M-alternating unit. Let v be an arbitrary internal vertex of G. Clearly, there exists an edge $e \in M$ incident with v. If e is external, then we are through. Otherwise, since e is allowed, for any external edge e' of G there exists a state M' and a crossing α with respect to M' such that α goes through e and e'. Thus, v is indeed accessible (e.g. in state M'). □

Claim 2.4. *Let C_1 and C_2 be two different external elementary components of G. There exists no alternating path β with respect to any state M connecting C_1 and C_2 in such a way that the two endpoints of β, but no other vertices, lie in C_1 and C_2.*

Proof. Assume, by contradiction, that there exists an M-alternating path β connecting vertex v_1 in C_1 with vertex v_2 in C_2 as described in the claim. Clearly, β must be negative at both ends. Moreover, v_i $(i = 1, 2)$ can be external only if $C_i = \{v_i\}$. Take a positive external M-alternating path α_i leading to v_i inside C_i if v_i is internal, otherwise let α_i be the empty path. The path α_i exists by Claim 2.3 above. Combining α_1, β, and α_2 then results in a crossing through both components C_1 and C_2, which contradicts that $C_1 \neq C_2$. □

Now we recall the definition of *canonical equivalence* from [12,3]. Let G be elementary, and define the relation $\sim$ on $\mathrm{Int}(G)$ by $v_1 \sim v_2$ if an extra edge e connecting v_1 with v_2 becomes forbidden in $G + e$. It is well-known that, in case G is closed, $\sim$ is an equivalence relation that determines the so-called canonical partition of $V(G)$. It was proved in [3] that, for open graphs, too, $\sim$ is an equivalence relation on $\mathrm{Int}(G)$.

Claim 2.5. *If v_1 and v_2 are two internal vertices of an elementary graph G, then $v_1 \nsim v_2$ iff one of the following conditions are met in any state M of G:*
(a) *there exists a positive M-alternating path connecting v_1 and v_2,*
(b) *there exists a positive M-alternating fork connecting v_1 and v_2.*

Proof. Consider the extra edge $e = (v_1, v_2)$ in the graph $G + e$. Since G is a nice subgraph of $G + e$, the edge e cannot be mandatory. Therefore e is not forbidden iff there exists an M_e-alternating unit passing through e in any state M_e of $G + e$. Identifying the G-permissible states of $G + e$ with those of G, this is equivalent to saying that e is not forbidden in $G + e$ iff there exists an M-alternating unit passing through e in any state M of G. The claim is now obvious. □

3. The closure of open graphs

In order to prove a result on open graphs and perfect internal matchings it is sometimes useful to start reasoning about some related closed graphs with perfect matchings, and then deduce the desired result by reopening these graphs. The closure operation introduced in this section allows a deduction mechanism of this nature. Throughout this section, unless otherwise stated, G will denote an open graph.

Definition 3.1. The *closure* of graph G is the closed graph G^* for which:
- $V(G^*)=V(G)$ if $|V(G)|$ is even, and $V(G^*)=V(G)\cup\{c\}$, $c\notin V(G)$ if $|V(G)|$ is odd;
- $E(G^*)=E(G)\cup\{(v_1,v_2)\mid v_i\in \mathrm{Ext}(G)\cup\{c\}\}$.

Intuitively, G^* is obtained from G by connecting all of its external vertices with each other in all possible ways. If $|V(G)|$ happens to be odd, then a new vertex c is added to G, and edges are introduced from c to all the external vertices. The edges of G^* belonging to $E(G^*)-E(G)$ will be called *marginal*, and the vertex c will be referred to as the *collector*. Edges incident with the collector vertex will also be called collector edges.

Notice that, in the specification of $E(G^*)$, it is not required that $v_1\neq v_2$. Consequently, in G^*, we are going to have a loop around each external vertex of G. These loops have no specific role if G has at least two external vertices, although their introduction as trivial forbidden edges is harmless. If there is only one external vertex in G, however, the loop is essential to make G^* closed.

Proposition 3.2. *Graph G has a perfect internal matching iff G^* has a perfect matching.*

Proof. If G^* has a perfect matching M^*, then deleting the marginal edges from G^* and M^* will leave G with a perfect internal matching. Conversely, if G has a perfect internal matching M, then it is always possible to extend M to a perfect matching of G^* by matching up the external vertices of G not covered by M in an arbitrary way, using the collector vertex c if necessary. Obviously, the use of c is necessary if and only if $|V(G)|$ is odd. □

Lemma 3.3. *Every M-alternating crossing of G can be turned into an M^*-alternating cycle of G^* by any extension of M to a perfect matching M^*. Conversely, for an arbitrary perfect matching M^* of G^*, every M^*-alternating cycle of G^* containing at least one marginal edge opens up to a number of alternating crosses with respect to the restriction of M^* to $E(G)$ when the marginal edges are deleted from G^*.*

Proof. Straightforward, using the same argument as in Proposition 3.2. □

Corollary 3.4. *For every edge $e\in E(G)$, e is allowed in G iff e is allowed in G^*.*

Proof. Indeed, by Lemma 3.3, e is allowed in G
iff there exists a M-alternating unit through e in G for some M,
iff there is an M^*-alternating cycle through e in G^* for some M^*,
iff e is allowed in G^*. □

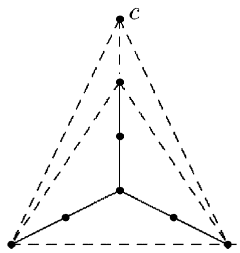

Fig. 2. Marginal edges that are forbidden in G^*.

Corollary 3.5. *A connected graph G is elementary iff G^* is elementary.*

Proof. If G is elementary, then its allowed edges form a connected subgraph G_e of G covering all the external vertices. By virtue of Corollary 3.4, G_e is part of an elementary component in G^*, which must be the only one as the collector vertex alone cannot form an elementary component in the closed graph G^*. Conversely, let G^* be elementary, and assume by way of contradiction that G has more than one elementary component. All these components must be external, because any internal elementary component of G would also be an elementary component of G^* according to Corollary 3.4. Since G is connected, there must be two elementary components in G that are connected by a forbidden edge, which is in contradiction with Claim 2.4. □

By Corollary 3.4, if the closure G^* of a connected graph G is 1-extendable, then so is G. Conversely, if G is 1-extendable, then only the marginal edges of G^* might be forbidden in G^*. Among these, however, the collector edges are ruled out for the following reason. Let v be an arbitrary external vertex of G, and consider a state M of G by which v is left uncovered. Such a state M can always be found, because if a randomly chosen M' does cover v, then switching to state $M = S(M', \alpha)$ for an appropriate crossing α starting from v will do the job. (Crossing α will exist, for G cannot be a single mandatory external edge if the collector vertex is present.) Now we can extend M to a perfect matching M^* of G^* by first putting in the edge (v, c), then matching up the remaining uncovered external vertices of G in an arbitrary way. This proves the edge (v, c) allowed. Thus, only those marginal edges can be forbidden in G^* that connect the external vertices of G directly. Fig. 2 shows a simple example where all these edges are indeed forbidden.

If G is not elementary, then several of its external elementary components may be amalgamated in G^*. The internal elementary components of G, however, will remain intact in G^* as every forbidden edge of G is still forbidden in G^*. The mandatory external elementary components of G, too, will remain mandatory in G^*. We claim that the union of all non-mandatory external elementary components of G, together with the collector vertex if that is present, forms one elementary component in G^*, called the *amalgamated* elementary component. Indeed, as we have already seen, every collector edge not adjacent to a mandatory external edge of G is allowed in G^*. Similarly, if e is an edge in G^* connecting two external vertices of G belonging to different non-mandatory elementary components, then it is always possible to find a state M of G by which the two endpoints of e are not covered. Then M can be extended to a perfect matching M^* of G^* by putting in the edge e first, so verifying it to be allowed in G^*.

The observations of the previous paragraph are summarized in Theorem 3.6 below, which provides a characterization of the elementary decomposition of G^*.

Theorem 3.6. *The set of elementary components of G^* consists of*:
(i) *the internal elementary components of G*;
(ii) *the mandatory external elementary components of G*;
(iii) *the amalgamated elementary component, which is the union of all non-mandatory external elementary components of G and the collector vertex, if that is present.*

4. Canonical equivalence

Recall from Section 2 that the canonical partition of an elementary graph G is determined by the equivalence relation $\sim$ on $V(G)$. We generalize this relation for non-elementary graphs in the following natural way.

Definition 4.1. For any two internal vertices $u, v \in V(G)$, $u \sim v$ if u and v belong to the same elementary component of G and the edge $e = (u, v)$ becomes forbidden in $G + e$.

One might think that the relation $\sim$, when restricted to a particular elementary component C, results in the equivalence $\sim_C$, which is canonical equivalence on C alone in the usual sense. In general this fails to hold, and we shall see that $\sim|C$—the restriction of $\sim$ to C—is just a refinement of $\sim_C$. At the moment, however, we do not even know that $\sim$ is an equivalence relation for non-elementary graphs. All we know is that $\sim$ is reflexive and symmetric, and that $u \not\sim_C v$ implies $u \not\sim v$, i.e., $\sim|C \subseteq \sim_C$. Claim 2.5, too, remains true under the current more general conditions.

In the light of Claim 2.5, Lemma 3.3 and Corollary 3.4 it is easy to see that for any two internal vertices u and v belonging to the same elementary component of G, $u \sim v$ holds in G iff $u \sim v$ holds in G^*. Furthermore, if u and v are arbitrary vertices belonging to different non-mandatory external elementary components, then $u \not\sim v$ holds in G^*. Indeed, by Claim 2.3, if u (v) is internal, then there exists a positive external alternating path leading to that vertex in its elementary component with respect to any state M of G, which path (paths) will give rise to a positive alternating path connecting u with v in G^* with respect to any extension of M. Finally, by the same argument, $c \not\sim v$ holds for the collector vertex c and any other vertex v in the amalgamated elementary component of G^*. Thus, we have proved the following characterization of the relation $\sim_{G^*}$ in terms of $\sim_G$.

Theorem 4.2. *Let u and v be vertices of an elementary component C in G^*.*
(i) *If u and v are both internal in G, then, irrespective of the choice of C, $u \sim_{G^*} v$ iff u and v are in the same elementary component of G, too, and $u \sim_G v$.*
(ii) *If C is a mandatory external elementary component of G, then $u \sim_{G^*} v$ iff $u = v$.*

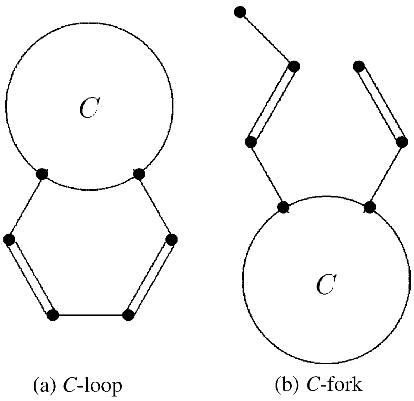

Fig. 3. A C-loop and a C-fork.

(iii) *For C being the amalgamated elementary component, $u \not\sim_{G^*} v$ whenever u and v belong to different external elementary components of G, or exactly one of them is the collector vertex. If u and v are external vertices of the same elementary component in G, then either of $u \sim_{G^*} v$ and $u \not\sim_{G^*} v$ is possible.*

(iv) *Statements* (i)–(iii) *remain true if we replace $\sim_G$ and $\sim_{G^*}$ in them by the local relations $\sim_C$ in G and G^*, respectively.*

Corollary 4.3. *For every elementary component C of G,*

$$\sim_G |C = \sim_{G^*} |C \qquad \textit{and} \qquad \sim_C \; = \; \sim_{C^*} |C,$$

where C^* is the elementary component of G^* containing C.

Proof. Straightforward by Theorem 4.2(i) and (iv). □

Let C be a nice elementary subgraph of G, and consider a C-permissible perfect internal matching M in G. An *M-alternating C-loop* (or just C-loop if M is understood) is a negative internal M-alternating path or loop in G having both endpoints, but no other vertices, in C, see Fig. 3a. If C is closed, then an *M-alternating C-fork* is a pair of edge-disjoint negative external M-alternating paths such that their internal endpoints, but no other vertices, are in C, see Fig. 3b. A C-loop (fork) is said to *connect* its internal endpoints even if this does not in fact happen in the case of forks. Notice that Claim 2.4 excludes the possibility of having a C-fork with C being external.

Definition 4.4. A *hidden edge* of G is an edge $e = (v_1, v_2)$, not necessarily in $E(G)$, for which v_1 and v_2 are the endpoints of an M-alternating C-loop or C-fork for some elementary component C and state M of G. The word "*shortcut*" will sometimes be used as a synonym for "hidden edge".

Note that, by definition, every forbidden edge in an elementary component C of G is a C-loop, and hence becomes a hidden edge of G. Reversing the argument one can see that hidden edges always become forbidden in their respective elementary components. Indeed, suppose that $v_1 \not\sim_C v_2$ for the two endpoints v_1 and v_2 of an M-alternating C-loop or C-fork α. Then there exists a positive M-alternating path or fork β connecting v_1 with v_2 running entirely in C. (See Claim 2.5.) Notice that a fork α cannot be coupled with a fork β in one case, since an alternating C-fork exists only if C is closed. Combining the negative α with the positive β then results in an M-alternating unit in G containing α, which contradicts the fact that C is an elementary component.

Let us now have a closer look at the composition of an alternating C-loop α for some elementary component C. Intuitively, α starts out from an internal vertex of C and, after traversing a forbidden edge of G, enters another elementary component C_1. After making a positive alternating path in C_1 the whole process is iterated, so that by the time α returns to C, a sequence $C_1,\ldots,C_n$ of elementary components will have been visited. Note that the case $n=0$ is possible, indicating the presence of a single forbidden edge in C as a C-loop. Also notice that there might be repetitions in the sequence $C_1,\ldots,C_n$, as any of these components can be left and reentered subsequently. We say that the components $C_1,\ldots,C_n$ are *covered* by the C-loop α.

The following proposition shows that the particular matching M, relative to which α is defined, has no bearing on the existence and composition of C-loops covering internal components only.

Proposition 4.5. *Let α be a C-loop connecting vertices v_1 and v_2 of an elementary component C with respect to some state M of G, and assume that all components covered by C are internal. Then, for every state M', there exists an M'-alternating C-loop connecting v_1 and v_2 that goes through the same forbidden edges as α and covers the same set of elementary components, too.*

Proof. Let $\mathscr{C}_\alpha$ be the set of elementary components covered by α, and consider the subgraph $G[\cup\mathscr{C}_\alpha]$ of G determined by the union of these components. Augment $G[\cup\mathscr{C}_\alpha]$ by the two forbidden edges e_1 and e_2 of α originally incident with v_1 and v_2, and consider them as external edges. Denote the resulting graph having two external vertices by G_α, and let M_α (M'_α) be the restriction of M (respectively, M') to G_α. Clearly, G_α is elementary, since the opening of the loop α—being an M_α-alternating crossing in this graph—connects the components in $\mathscr{C}_\alpha$ to each other. Consider the state $S(M_\alpha,\alpha)$ of G_α. Making the crossing α in this state and then switching to state M'_α determines an alternating network N with respect to state M'_α. The network N will consist of several cycles within the components belonging to $\mathscr{C}_\alpha$ and one crossing α' connecting the two external vertices. Clearly, the crossing α' determines a C-loop in G with respect to state M'. All the forbidden edges of G traversed by α will also be traversed by α', as none of these edges are present in either M or M'. Thus, α' covers exactly the same elementary components as α, not necessarily in the same order, though. Nevertheless, it certainly covers each one with the same multiplicity as α. □

Lemma 4.6. *The hidden edges of G^* different from the forbidden marginal edges are exactly the hidden edges of G.*

Proof. Let α be a C-loop or C-fork in G for some elementary component C with respect to state M. If C is internal, then obviously α determines a C-loop α^* in G^* with respect to any extension of M to a perfect matching M^*. If C is external, then Claim 2.4 implies that α is a loop that will not reach any other external elementary component of G. Therefore α^* becomes an A-loop in G^*, where $A = C$ if C is mandatory, and A is the amalgamated elementary component otherwise. Thus, every hidden edge in G is one in G^*.

Now let α be a C-loop connecting vertices v_1 and v_2 of an elementary component C in G^* with respect to some perfect matching M^*. By Theorem 4.2, neither v_1 nor v_2 is the collector. If either v_1 or v_2, say v_1, is external in G, then v_2 is external, too, belonging to the same elementary component of G as v_1. Indeed, by Theorem 4.2, there are no forbidden edges in G^* incident with v_1 other than the marginal ones. Let therefore v_1 and v_2 be both internal in G. By Claim 2.4, these two vertices are in the same elementary component of G even if $C = A$ is the amalgamated elementary component. Therefore there exists an elementary component C' of G such that either α is a C'-loop or it opens up to a C'-fork with respect to the restriction of M^* to G. Thus, every hidden edge of G^* that is not a forbidden marginal edge is a hidden edge of G. □

For every elementary component C of G, let C_h denote the enhancement of C with all the hidden edges belonging to C. Similarly, denote by G_h the graph obtained from G by adding all of its hidden edges.

Corollary 4.7.

$$(G_h)^* = (G^*)_h.$$

Proof. Straightforward by Lemma 4.6. □

Corollary 4.8. *Let α be a C-loop or C-fork connecting vertices v_1 and v_2 of some elementary component C with respect to state M of G. Then for every state M' there exists an M'-alternating C-loop or C-fork connecting v_1 and v_2 that covers the same forbidden edges and elementary components as α.*

Proof. By Proposition 4.5 it is enough to prove the statement in the case when α is either a fork or it is a loop covering an external elementary component D. Claim 2.4 then implies that C is internal and D is unique. Let M^* and $(M')^*$ be any extensions of M and M' to perfect matchings in G^*. Following the argument in the first paragraph of the proof of Lemma 4.6, α determines an appropriate C-loop α^* in G^* with respect to M^*. Using Proposition 4.5 again, there exists a C-loop $(\alpha')^*$ with respect to $(M')^*$ in G^* covering the same forbidden edges and elementary components as α^*. Reopening G^* then determines a C-loop or C-fork α' with respect to M' in G. Since the external

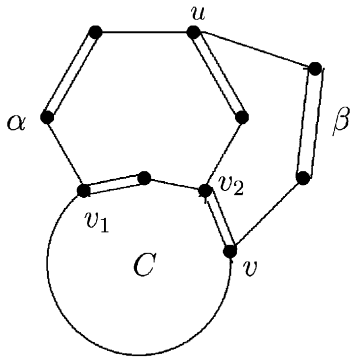

Fig. 4. The proof of Lemma 4.9.

component D that might affect the opening of $(\alpha')^*$ into α' is unique, α' will cover the same forbidden edges and elementary components as α. □

Our goal is to show that the elementary decomposition of G_h is the same as that of G, and all the hidden edges of G remain forbidden in G_h. Although this fact might seem obvious to the reader already at this point, its formal proof poses a technical challenge, which will be dealt with in Lemma 4.9 and Theorem 4.11 below.

Let C be any elementary subgraph of G, and assume that a negative alternating trail α is such that none of its vertices, except possibly the endpoints, are in C. We shall refer to this situation by saying that α runs *essentially outside* C.

Lemma 4.9. *Let C be a nice elementary subgraph of G, and let $v, v_1, v_2 \in V(C)$ be such that $v_1 \sim_C v_2$ but $v \not\sim_C v_i$ for $i = 1, 2$. Moreover, for some C-permissible state M of G, let α be an M-alternating C-loop or C-fork connecting v_1 and v_2, and β be a negative M-alternating path running essentially outside $C + \alpha$, connecting v with a vertex u lying on α. Then there exists an M-alternating unit in $C + \alpha + \beta$ containing β.*

Proof. (i) Assume first that G is closed, so that α is a C-loop. The situation is depicted by Fig. 4. The edge $e \in M$ on α incident with u acts like a valve for β in the sense that it points to either v_1 or v_2. Say the valve points to v_2 as in Fig. 4. Let γ be the M-alternating path that starts out from v on β, then switches to α at u, and ends in v_2. Since $v_2 \not\sim_C v$, there exists a positive M-alternating path connecting v_2 with v inside C. Combining this path with the negative alternating path γ results in the desired M-alternating cycle.

(ii) If G is open, then consider the closure $[C]^*$ of the subgraph $[C]$ $(= G[C])$, and observe that $[C]^*$ is a nice elementary subgraph of G^*. This is obvious if the collector vertex is present in G^*. If it is not, but the collector is needed for $[C]^*$, then any external vertex of G not in C is suitable for this purpose. Such a vertex will always exist, otherwise the collector would not be necessary in $[C]^*$ either. Clearly, $v_1 \sim_{[C]^*} v_2$ and $v \not\sim_{[C]^*} v_i$ for $i = 1, 2$. Moreover, α determines a $[C]^*$-loop α^* in G^* with respect to any $[C]^*$-permissible extension of M to a perfect matching. This is true because at most one of $[C]^* \neq C$ and $\alpha^* \neq \alpha$ can hold, keeping α^* essentially outside $[C]^*$. (Remember that C must be internal for any C-fork.) Now the statement follows easily from (i). □

Corollary 4.10. *Let α be an M-alternating C-loop or C-fork for some elementary component C of G connecting vertices v_1 and v_2, and let β be an M-alternating path starting out from a vertex v in C, but running essentially outside C. If $v \not\sim_C v_i$ for both $i = 1, 2$, then β must avoid all the elementary components covered by α.*

Proof. Assume, on the contrary, that there exists a negative M-alternating path β satisfying the conditions of the corollary in such a way that the other endpoint u of β lies on an elementary component C' covered by α, but β runs essentially outside $C \cup C'$. By switching to G^* we can assume, without loss of generality, that α is a loop. (See Lemma 4.6.) According to Lemma 4.9, β and α cannot have a vertex in common. Let u_1 and u_2 be two vertices of C' where α enters and subsequently leaves this component. Clearly, $u_1 \not\sim_{C'} u_2$, so that u and at least one of u_1, u_2 are in different canonical classes by $\sim_{C'}$. The path β can therefore be continued from u inside C' in an M-alternating way to reach u_1 or u_2. In either way this continuation will eventually hit the loop α, which contradicts Lemma 4.9. □

Theorem 4.11. *For an elementary component C of G, let $e_1, \ldots, e_n$ be any number of hidden edges in C. Then, for the elementary graph $C_n = C + e_1 + \cdots + e_n$, each edge e_i remains forbidden in C_n, and $\sim |C \subseteq \sim_{C_n}$.*

Proof. (i) Again, assume first that G is closed. The proof is an induction argument on n. For $n = 0$ the statement is trivial. Assume it holds for any choice of hidden edges $e_1, \ldots, e_n$, $n \geqslant 0$, and let e_{n+1} be a further hidden edge. Let β be an arbitrary positive alternating path or alternating cycle in C_{n+1} with respect to some state M, and try to replace the edges e_i on β by appropriate C-loops one-by-one, until an overlap occurs between two of them in G. Note that such loops always exist by Corollary 4.8. We claim that the process of unfolding the hidden edges in β will be successful all the way, that is, all newly introduced C-loops will be pairwise disjoint. On the contrary, let us assume that we encounter an overlap when introducing a C-loop for edge e_i with the one that has been substituted for e_j previously, and this is the first time an overlap occurs. Without loss of generality we can assume that the hidden edges that have already been successfully replaced are $e_1, \ldots, e_{i-1}$, and $j = 1$, see Fig. 5.

In the way described above, we will have an instance of the situation captured by Lemma 4.9 with C in that lemma being $C^i = C + e_2 + \cdots + e_{i-1}$ now, α being the loop that replaced e_1 with endpoints v_1, v_2, and β being an appropriate subpath of the loop attempted to be substituted for e_i starting out from vertex v. Note, however, that the base graph G in that lemma is now $G + e_2 + \cdots + e_{i-1}$, in which we do not know yet if C^i is an elementary component. But it certainly is a nice elementary subgraph. To verify the conditions of the lemma, observe that $v_1 \sim_{C^i} v_2$, since e_1 is still forbidden in $C^i + e_1 = C_{i-1}$ by the induction hypothesis. Moreover, $v_2 \not\sim_{C^i} v$, since there exists a

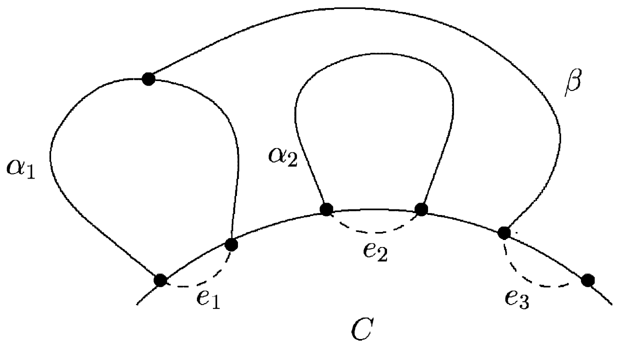

Fig. 5. Unfolding the loops in Theorem 4.11.

positive alternating path connecting v_2 with v in G using the pairwise disjoint C-loops introduced for $e_2,\dots,e_{i-1}$, therefore there exists one in $C+e_2+\cdots+e_{i-1}$ without using them. The application of Lemma 4.9 then results in an M-alternating cycle γ in $C^i+\alpha+\beta$ containing β. As β does not overlap with the previously introduced loops for e_k, $2\leqslant k\leqslant i-1$, these loops can be reintroduced in γ to obtain a M-alternating cycle already in G containing β, which is a contradiction.

Having made the above powerful argument, the induction started in (i) can now be finished easily. Suppose e_{n+1} becomes allowable in the graph $C+e_1+\cdots+e_{n+1}$. Then there is an M-alternating cycle γ containing some (in fact all) of the edges e_i, $1\leqslant i\leqslant n+1$. Replacing these edges by appropriate pairwise disjoint C-loops yields an M-alternating cycle in G covering forbidden edges, which is impossible. The proof of $\sim|C\subseteq\sim_{C_{n+1}}$ follows exactly the same argument, and is left to the reader.

(ii) If G is open, then switch to the graph G^*, and apply part (i) for this graph and its elementary component C^* containing C. Theorem 3.6 and Lemma 4.6 ensure that all the required conditions are met. Thus, the edges $e_1,\dots,e_n$ are forbidden in C_n^*, and $\sim_{G^*}|C^*\subseteq\sim_{C_n^*}$. Coming back to the graph G it follows immediately that $e_1,\dots,e_n$ are forbidden in C_n. Furthermore,

$$
\begin{aligned}
&\sim_G|C\\
&\quad=\sim_{G^*}|C\quad\text{(by Corollary 4.3)}\\
&\quad=(\sim_{G^*}|C^*)|C\\
&\quad\subseteq\sim_{C_n^*}|C\\
&\quad=\sim_{C_n}\quad\text{(by Corollary 4.3)}.\quad\square
\end{aligned}
$$

Corollary 4.12. *For every elementary component C,*

$$\sim|C=\sim_{C_h}.$$

Proof. Notice that $\sim_{C_h}\subseteq\sim|C$, because every positive alternating path or fork β in G connecting two vertices of C can be turned into a path or fork β_h in C_h by making the appropriate shortcuts. This fact is obvious unless C is external and β is a fork. But in this case, too, Claim 2.4 implies that β_h remains in the elementary component C_h. On the other hand, $\sim|C \subseteq \sim_{C_h}$ follows from Corollary 4.12. □

Corollary 4.13. *The elementary decomposition of G is the same as that of G_h.*

Proof. It is sufficient to prove that the addition of just one hidden edge e to G does not change the elementary decomposition of G. This is equivalent to saying that e is forbidden in $G+e$. Suppose, by contradiction, that for any hidden edge e connecting vertices v_1 and v_2 in elementary component C there exists an inter-elementary alternating unit γ in $G+e$ with respect to some state M of $G+e$ going through e. Without loss of generality we can assume that $e \notin M$, i.e., M is a state of G, too. The unit γ puts v_1 and v_2 in different canonical classes according to $\sim|C$. But then, by Corollary 4.12, v_1 and v_2 cannot be in the same canonical class according to $\sim_{C_h}$ either, which is in contradiction with Theorem 4.11. □

The key observation made in the proof of Theorem 4.11 is now generalized and stated as a separate principle.

Theorem 4.14 (Shortcut Principle). *For any state M, let γ be an arbitrary M-alternating trail in G_h. Then any number of the shortcuts along γ can be unfolded into appropriate M-alternating loops or forks without the chance of creating any intersections. Moreover, γ either remains a trail or becomes a pair of external trails after the unfolding, the latter only if γ is internal.*

Proof. It is sufficient to prove that the unfolding of just one shortcut $e=(v_1,v_2)$ in some elementary component C into a C-loop or C-fork α does not create an intersection with the rest of γ, and that the unfolding of γ has the desired properties.

(i) *G is closed*. Assume, by contradiction, that α intersects with γ. Setting out on γ from v_1 or v_2 in a positive M-alternating way (i.e., on an edge belonging to M) we must encounter a vertex that lies on α. Let u be the first such vertex, starting out from say v_1. On the interval from v_1 to u there is a last vertex v at which γ leaves component C. Making the appropriate shortcuts in C on the interval of γ from v_1 to v results in a positive M-alternating path connecting these two vertices in C_h, indicating that $v_1 \not\sim_{C_h} v$. A contradiction is now immediate by Lemma 4.9. Obviously, γ is a single trail after the unfolding.

(ii) *G is open*. Consider the graph $(G^*)_h$ and the elementary component C^* in G^* containing C. In this setting γ determines an alternating trail γ^* in $(G^*)_h$, and α determines an alternating C^*-loop α^* with respect to any extension of M to a perfect matching. (See Theorem 3.6, Lemma 4.6, and Corollary 4.8.) Knowing from (i) that α^* and γ^* do not intersect, it follows that their subtrails α and γ do not intersect either. If α is a fork, then γ cannot be external, because in that case one of the two trails arising from the unfolding would be an inter-elementary crossing. This observation proves the second statement of the theorem. □

Corollary 4.15. *An edge $e \in E(G)$ is impervious in G iff e is impervious in G_h.*

Proof. Straightforward by the Shortcut Principle. □

Let $\mathscr{P}(G)$ denote the canonical partition of Int(G) determined by the equivalence $\sim$.

Corollary 4.16. $\mathscr{P}(G) = \mathscr{P}(G_h)$.

Proof. Immediate by Claim 2.5 and the Shortcut Principle. □

Let $\mathscr{F}(G)$ and $\mathscr{H}(G)$ denote the sets of forbidden and hidden edges of G.

Corollary 4.17. $\mathscr{F}(G_h) = \mathscr{F}(G) \cup \mathscr{H}(G)$.

Proof. For any graph G, the set of forbidden edges consists of:
(a) the edges connecting two different elementary components in G;
(b) the forbidden edges of the elementary components themselves.

By Corollary 4.13, edges in (a) are common for G and G_h. Moreover, by Theorem 4.11, the forbidden edges of G_h belonging to (b) are exactly the hidden edges of G. □

In the sequel, by a canonical class of some elementary component C we shall mean a class by the partition $\mathscr{P}(G) = \mathscr{P}(G_h)$, rather than one by the partition associated with the equivalence $\sim_C$. According to Corollary 4.12, $\mathscr{P}(G)$ is determined locally by the equivalence relations $\sim_{C_h}$.

5. Structuring the elementary components

In the previous section we were concerned with the behavior of one particular elementary component of G when placed in the global environment determined by the surrounding elementary components. In this section we look at the global environment itself, and investigate the structure of all elementary components in G. Elementary components will be related to each other according to their accessibility from external vertices by alternating paths. Unlike in the previous sections, we shall use the phrase "external alternating path γ *enters* elementary component C" in the strict sense, meaning that γ enters C for the first time. Obviously, the path γ must then be negative.

Definition 5.1. An elementary component of G is *viable* if it does not contain impervious allowed edges. A viable internal elementary component C is *one-way* with respect to some state M of G if all external M-alternating paths enter C in the same canonical class of C. This unique class is called *principal* in C. Further to this, every external elementary component is a priori one-way by the present definition (with no principal canonical class, of course). A viable elementary component is *two-way* if it is not one-way. An *impervious* elementary component is one that is not viable.

It is easy to see that an impervious elementary component consists of impervious edges only. On the contrary, let us assume that there exists a positive external alternating path α leading to some vertex of an impervious elementary component C. Let v be the vertex of C where α enters this component, and denote by β the prefix of α up to v. By Claim 2.3, C is internal. Moreover, if e is an arbitrary allowed edge of C, then, using the argument in Claim 2.3, there exists a positive alternating path γ in C connecting v with one endpoint of e in some state of C. Thus, the positive external alternating path $\beta\gamma$ in an appropriate state of G proves e to be viable. Since e was arbitrary, this contradicts the fact that C is impervious.

Proposition 5.2. *The one-way property is matching invariant with the principal canonical class preserved.*

Proof. Consider a negative external alternating path γ entering C in state M, and let M' be any other state. As in the proof of Proposition 4.5, restrict G and M to the elementary components visited by γ, and designate the last edge of γ incident with C as an external edge. In the resulting graph G_γ, γ becomes an M_γ-alternating crossing. Make the crossing γ in state $S(M_\gamma,\gamma)$, and then switch to state M'_γ. Apply the argument in Proposition 4.5 to conclude that there exists an M'-alternating crossing γ' in G_γ with the same endpoints and visiting the same elementary components as γ. Thus, γ' determines an M'-alternating external path in G entering C at the very same vertex as γ. In this way we have shown that the entry points of external alternating paths in C are the same with respect to all states of G. □

Proposition 5.3. *Let C be a viable internal elementary component of G. Then a C-fork exists in any state M only if C is one-way, and the internal endpoints of the fork are in the principal canonical class of C. The corresponding hidden edge in C_h is impervious in G_h.*

Proof. Let (α_1,α_2) be an M-alternating C-fork in G connecting vertices v_1 and v_2 belonging to a canonical class P of C. Suppose, by contradiction, that there exists a negative external M-alternating path γ entering C_h in a vertex v belonging to a canonical class different from P. By Lemma 4.9, γ must avoid the fork (α_1,α_2). But then a crossing would be obtained in G_h through γ, a positive M-alternating path in C_h from v to v_1 (v_2) and α_1 (respectively, α_2). We conclude that C is one-way with the class P being principal. Observe that all vertices v in any principal canonical class P are inaccessible. Indeed, if there was a positive external alternating path γ leading to v, then $v \not\sim u$ would hold for the vertex u where γ enters C. This is impossible, however, since u is also in class P. The edge (v_1,v_2) is therefore impervious. □

Proposition 5.4. *Component C is one-way in G iff C_h is one-way in G_h, and the principal canonical class of C is the same as that of C_h.*

Proof. It is sufficient to prove that if C is one-way and internal, then C_h is also one-way and its principal canonical class is that of C. On the contrary, assume that C is one-way with principal canonical class P, yet, there exists a negative external alternating path γ in G_h with respect to some state M of G_h that enters C_h at a vertex v belonging to a class different from P. Using the Shortcut Principle (Theorem 4.14), let us unfold the hidden edges on γ one-by-one, starting from the external vertex, into pairwise disjoint M-alternating loops until an intersection occurs with C at some vertex u. This intersection will indeed occur, otherwise the unfolding of γ would enter C at vertex v. At the vertex u, an appropriate external subpath of the unfolding of γ enters C, therefore u is in class P. Extend γ by a positive M-alternating path inside C_h up to the vertex u to obtain an M-alternating path γ'. It is now obvious that the Shortcut Principle fails to work for γ', which is a contradiction. □

Definition 5.5. Component C' is *two-way accessible* from component C with respect to some state M, in notation $C\rho C'$, if C' is covered by an appropriate M-alternating C-loop α. It is required, though, that if C is one-way and internal, then the endpoints of α *not* be in the principal canonical class of C.

Let C' be two-way accessible from C via loop α. The endpoints of α in C are called the *domain* vertices of α, while the *range* vertices of α (on C') are the vertices at which α first hits C' from both ends. The common canonical class of the domain vertices in C is also called domain, and the classes of the range vertices in C' are called range as well. Clearly, the two range classes are different. The two negative alternating paths connecting the domain and range vertices within α are called the (C'-) branches of α.

According to Definition 5.5, if C is internal and $C\rho C'$ via loop α, then there exists an external alternating path entering C in a vertex belonging to a canonical class different from the domain of α. This observation will often be used in the sequel.

Lemma 5.6. *If $C\rho C'$ with respect to M, then C' cannot be one-way.*

Proof. Let α be a C-loop covering C' from domain class P. Suppose first that C is viable. By Claim 2.4, at most one of C and C' can be external. If C' were external, then C, being internal, could be entered by an external M-alternating path γ in a vertex belonging to a canonical class different from P. By Lemma 4.9 and Corollary 4.10, γ avoids the loop α and component C', which contradicts Claim 2.4 again. We conclude that C' is internal. In this case, however, regardless of C being internal or external, C' can be entered by an external M-alternating path through C and the loop α in both range vertices of α, which proves that C' is two-way.

Now let C be impervious, and assume by way of contradiction that C' is viable, let alone one-way. Let γ be an external M-alternating path entering C' at some vertex u. Clearly, there exists a positive M-alternating path β connecting u with at least one of the range vertices of α inside C'. If γ does not intersect with α, then C could be entered through γ, β, and an appropriate branch of α, contradicting that C is impervious. The same contradiction arises if γ does overlap with α, since in this case one can simply switch from γ to α at the first overlap to reach C from one direction. □

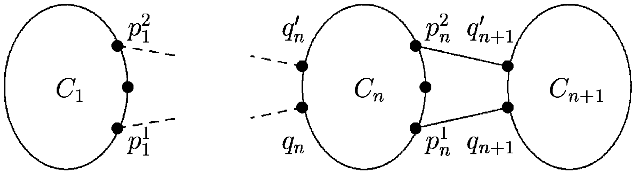

Fig. 6. The proof of Lemma 5.8.

Proposition 5.7. *The relation ρ is matching invariant.*

Proof. By Definition 5.5, if $C\rho C'$ via some C-loop α, then $C\rho C''$ holds for every elementary component C'' covered by α. Lemma 5.6 then implies that all elementary components covered by α are internal. Now the statement follows directly from Proposition 4.5. □

Let us fix a state M for the rest of Section 5. All alternating paths, C-loops, etc., will be meant with respect to this state. Since all the concepts to be dealt with are matching invariant, the choice of M is irrelevant.

Lemma 5.8. *Let $C_1,\dots,C_n$ $(n\geqslant 2)$ be elementary components such that $C_i\rho C_{i+1}$ for all $1\leqslant i\leqslant n-1$ by appropriate C_i-loops α_i with domain vertices p_i^1, p_i^2 and range vertices q_i, q_i'.*

(i) *The components $C_1,\dots,C_n$ are all different.*

(ii) *For either choice $q\in\{q_n,q_n'\}$ there exists $j\in\{1,2\}$ such that p_1^j is connected to q by a negative alternating path β in G_h running essentially outside $C_1\cup C_n$. Moreover, every edge of β is either on a loop α_i or belongs to some elementary component $(C_i)_h$.*

(iii) *If v is a vertex in C_1 such that $v\not\sim p_1^j$ $(j=1,2)$, then there exists no alternating path β in G_h running essentially outside C_1 and connecting v with any vertex in C_n.*

Proof. Induction on n. For $n=2$ statements (i) and (ii) are straightforward, while (iii) is equivalent to Corollary 4.10. Assume that all three statements hold for some $n\geqslant 2$, and proceed to $n+1$. See Fig. 6 for an illustration.

(i) Assume, by contradiction, that $C_{n+1}=C_m$ for some $1\leqslant m\leqslant n$. Without loss of generality we can take $m=1$. Then at least one C_1-branch of α_{n+1} violates part (iii) of the induction hypothesis when that branch is taken for β.

(ii) By (i) above we already know that C_{n+1} is different from all C_i, $1\leqslant i\leqslant n$. Choose $q\in\{q_{n+1},q_{n+1}'\}$ arbitrarily, and let q be connected to p_n^k by the branch α_n^k of α_n, where $k\in\{1,2\}$. Since $q_n\not\sim q_n'$ holds in $(C_n)_h$, either $q_n\not\sim p_n^k$ or $q_n'\not\sim p_n^k$. Say $q_n\not\sim p_n^k$. Then there exists a positive alternating path β' in $(C_n)_h$ between p_n^k and q_n. On the other hand, the induction hypothesis provides an appropriate negative alternating path β_j between p_1^j and q_n for some $j\in\{1,2\}$, and Lemma 4.9 ensures that β_j does not overlap with α_n^k. Moreover, Corollary 4.10 ensures that β_j does not reach C_{n+1} either.

In this way $\beta_j\beta'\alpha_n^k$ becomes a negative alternating path, which connects p_1^j with q in G_h, running essentially outside $C_1 \cup C_{n+1}$ with the desired edge composition.

(iii) Contrary to the statement, assume that an undesired alternating path β exists. By the induction hypothesis, β connects v with a vertex u in C_{n+1} in such a way that it avoids all the components C_i, $1 \leqslant i \leqslant n$, and loops α_i, $1 \leqslant i \leqslant n-1$. Without loss of generality we can also assume that β runs essentially outside C_{n+1}, so that it is at vertex u where β first hits any elementary component along the loop α_n. Clearly, $u \not\sim q_{n+1}$ or $u \not\sim q'_{n+1}$ holds in C_{n+1}, say $u \not\sim q_{n+1}$. A big alternating unit will then show up in G_h containing β, a positive alternating path in $(C_{n+1})_h$ connecting u with q_{n+1}, a positive alternating path or fork in $(C_1)_h$ connecting v with p_1^1 (p_1^2), and a negative alternating path connecting q with p_1^1 (respectively, p_1^2) according to (ii). □

Corollary 5.9. *With the parameters of Lemma* 5.8, *if v is an arbitrary vertex in C_n, then there exists an alternating path β in G_h connecting v with one of p_1^1 and p_1^2 in such a way that*

(a) *β is positive at the v end and negative at the other end,*

(b) *every edge of β is either on a loop α_i, $1 \leqslant i \leqslant n-1$, or belongs to $(C_i)_h$ for some $2 \leqslant i \leqslant n$.*

Proof. Since $v \not\sim q_n$ or $v \not\sim q'_n$, the statement follows directly from Lemma 5.8(ii). □

Corollary 5.10. *The transitive closure of ρ is asymmetric.*

Proof. Immediate by Lemma 5.8(i). □

Corollary 5.11. *The connection $C\rho C'$ holds in G iff $C_h\rho C'_h$ holds in G_h.*

Proof. It is sufficient to prove that $C_h\rho C'_h$ implies $C\rho C'$. Let α_h be a C_h-loop covering C'_h, and unfold α_h using the Shortcut Principle. By Lemma 5.6, none of the components covered by α_h are one-way, and by definition, the loop α_h itself cannot be a single hidden edge connecting two vertices belonging to the principal canonical class of a one-way component either. Therefore, by Proposition 5.3, α_h unfolds into a trail α. We claim that α is a C-loop, and therefore $C\rho C'$. To this end we need to verify that α avoids C. Should α overlap with C, there would be a component D along α such that $C\rho D$ and $D\rho C$, which contradicts Corollary 5.10. □

Lemma 5.12. *For every two-way C' there exists a viable C such that $C\rho C'$.*

Proof. Assuming that C' is two-way, let γ_1 and γ_2 be two external alternating paths entering C' in different canonical classes. Clearly, γ_1 and γ_2 must overlap. If e is the last overlapping allowed edge along γ_1 and γ_2, then it is easy to see that $C_h\rho C'_h$ holds for the elementary component C containing e. Thus, by Corollary 5.11, $C\rho C'$. □

Let ρ^* denote the reflexive and transitive closure of ρ. By Corollary 5.10, ρ^* is a partial order.

Lemma 5.13. *Let C_1 and C_2 be two different elementary components of G such that $C_1\rho^*C$ and $C_2\rho^*C$ for some elementary component C. Then C_1 and C_2 cannot both be one-way.*

Proof. Based on Proposition 5.4 and Corollary 5.11 we can change the present setting from graph G to graph G_h. Let $C_1=C_1^1\rho C_1^2\rho\ldots\rho C_1^n=C$ and $C_2=C_2^1\rho C_2^2\rho\ldots C_2^m=C$ for appropriate components C_j^i, $j=1,2$, $1\leqslant i\leqslant n$ (m) via some loops α_j^i. By Lemma 5.8(ii) there exists a negative alternating path β_1 connecting a domain vertex v_1 in C_1^1 with a range vertex v_n in C_1^n, running essentially outside $C_1\cup C_1^n$ with an appropriate edge composition. If β_1 covers C_2, then $C_1\rho^*C_2$, therefore C_2 is not one-way by Lemma 5.6. Otherwise follow β_1 starting from v_1, and let C' be the first among those elementary components covered by β_1 that are also covered by some of the C_2^i-loops α_2^i. Note that C' exists, as C is always a candidate to be chosen for C' at last. Clearly, $C_2\rho^*C'$ via the column of loops $\alpha_2^1,\ldots,\alpha_2^i$ for some $1\leqslant i\leqslant m-1$. Let v be the vertex in C' where β_1 enters this component, and let β_1' denote the subpath of β_1 from v_1 to v. Apply Corollary 5.9 to obtain an alternating path β_2 connecting v with a domain vertex v_2 in C_2, so that β_2 is positive at the v end and negative at the v_2 end. By the choice of C', $\beta=\beta_1'\beta_2$ is a negative alternating path between v_1 and v_2 running essentially outside $C_1\cup C_2$. We shall make use of the path β in the next paragraph.

Let the vertices v_1 and v_2 belong to canonical classes P_1 and P_2, and assume by way of contradiction that both C_1 and C_2 are one-way. According to Claim 2.4, one of C_1 and C_2, say C_1, is internal. Then there exists an external alternating path γ_1 entering C_1 at some vertex u_1 belonging to its principal class R_1. Clearly, $P_1\neq R_1$ and P_2 is not principal either, for P_j, $j=1,2$ are the domain classes of the C_j-loops α_j^1. Without loss of generality we can assume that γ_1 does not reach C_2. Indeed, if γ_1 reached C_2, then C_2 would also be internal and we could continue the proof with C_2 and the prefix of γ_1 that enters C_2. If γ_1 and β overlap, then it is straightforward to assemble an external alternating path from parts of γ_1 and β which enters C_1 or C_2 in the non-principal canonical class P_1 (respectively, P_2). This contradicts both of these components being one-way. Assume therefore that γ_1 and β are edge-disjoint. Then γ_1, a suitable positive alternating path in C between u_1 and v_1, and β will form an external alternating path entering C_2 in class P_2, which is again a contradiction. □

Lemma 5.14. *Let C be one-way, and suppose that $C\rho^*C'$. Then every external alternating path entering C' must enter C first.*

Proof. Let $C=C_1\rho\ldots\rho C_n=C'$ via a column of C_i-loops α_i, $1\leqslant i\leqslant n-1$. Contrary to the statement of the lemma, assume that there exists an external alternating path γ entering C' at vertex v without having visited C first. Without loss of generality we can assume that γ does not overlap with any of the loops α_i. But then it is possible to enter C at a domain vertex of α_1 through γ and an appropriate continuation from v that is available by Corollary 5.9. This is a contradiction, since the canonical class of any domain vertex is not supposed to be principal. □

Definition 5.15. A *family* of elementary components in G is a block of the partition determined by the equivalence relation $(\rho \cup \rho^{-1})^*$. A family $\mathscr{F}$ is *viable* if every elementary component in $\mathscr{F}$ is such. An *impervious* family is one that is not viable.

As we observed in the proof of Lemma 5.6, for elementary components C and C' such that $C\rho C'$, C is viable iff C' is viable. Thus, any impervious family will consist of impervious elementary components only.

Theorem 5.16. *Every viable family contains a unique one-way elementary component, called the root of the family. Every member of the family is only accessible through the vertices belonging to the principal canonical class of the root by external alternating paths.*

Proof. By Corollary 5.10 and Lemma 5.12, each viable family does contain a one-way elementary component. Let C_1 and C_2 be one-way elementary components in a family $\mathscr{F}$. By Lemma 5.6, there is no elementary component D in $\mathscr{F}$ such that $D\rho C_i$ for either $i=1$ or 2. Thus, there exists $D \in \mathscr{F}$ such that $C_1\rho^* D$ and $C_2\rho^* D$. Lemma 5.13 then implies that $C_1 = C_2$. The second statement of the theorem is equivalent to Lemma 5.14. □

6. Families

We are now ready to further improve the results of Section 5, and provide an amusing description of the exciting world of elementary components in G. As in the second half of Section 5, let us fix a state M of G for reasonings involving alternating paths.

Lemma 6.1. *Let $D\rho C$ via some D-loop α, and let γ be an alternating path in G running essentially outside $C \cup D$, connecting any vertex $u \in V(G)$ with a vertex v in C. Then there exists an alternating path β in G_h connecting u with a domain vertex of α in such a way that it also covers C. The path β consists of edges belonging to α, γ and C_h only, and it ends in a suffix that contains one of the C-branches of α in full.*

Proof. Let v_1 and v_2 be the C-range vertices of α. Fix one of v_1 and v_2 as the *designated* range vertex v_d with the property that $v_d \not\sim v$, and denote by α_d the designated C-branch of α, i.e., the branch that leads to v_d. Denote by u_d and u_n the corresponding designated and non-designated domain vertices of α in D. Furthermore, let $C_{\gamma/\alpha}$ denote the elementary component containing the last allowed edge e on α_d before it reaches v_d such that e is also on γ, and denote $v_{\gamma/\alpha}$ the vertex of $C_{\gamma/\alpha}$ where α_d leaves this component after traversing e on its way to v_d. Observe that the valve e points γ to the direction v_d on α, otherwise α, starting from v_d, could be continued after e in an alternating way on γ to reach the vertex $v \not\sim v_d$ in C, see Fig. 7. If γ does not overlap with α_d, then take $C_{\gamma/\alpha} = D$ and $v_{\gamma/\alpha} = u_d$. In order to prove the statement of the lemma,

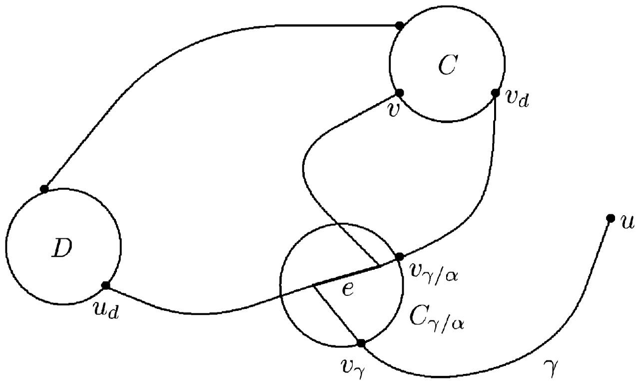

Fig. 7. The proof of Lemma 6.1.

we are going to further strengthen it by imposing the following two restrictions on the composition of β.

(i) β contains the interval of α between $v_{\gamma/\alpha}$ and v_{d},

(ii) β contains edges from C_h not on α only if γ does not overlap with α_{d}, in which case these edges constitute an arbitrary positive alternating path from v to v_{d}.

Now the proof is an induction argument on the number n of common edges to α and γ. If $n=0$, then the statement is obvious. Let $n\geqslant 1$, and suppose that the strengthened statement holds in all possible situations where the number of common edges to α and γ is less than n. If γ does not overlap with α_{d}, then we are through. If it does, then consider the component $C_{\gamma/\alpha}$ and the vertex v_γ where γ arrives at this component before traversing e, see Fig. 7. In the case when $u\in C_{\gamma/\alpha}$ and γ does not leave $C_{\gamma/\alpha}$ before arriving at e, the desired path β is easy to assemble from α and γ. Therefore we can assume that $u\notin C_{\gamma/\alpha}$. The vertex v_γ is then well-defined. As we have already seen, $v_\gamma \not\sim v_{\gamma/\alpha}$, because the valve e on α points to v_{d}. In this way we have reproduced the situation described in the lemma with C being $C_{\gamma/\alpha}=C'$, v being $v_\gamma=v'$, v_{d} being $v_{\gamma/\alpha}=v'_{\mathrm{d}}$ and γ being the prefix γ' of the original γ from u to v_γ. Notice that, in the new arrangement, the vertex $v_{\gamma'/\alpha}$ will be somewhere on the non-designated C-branch of α. In order for this, one must only select the interval of α from v_{d} all the way to u_n as the designated C'-branch.

Obviously, the path γ' has fewer edges in common with α than γ, so the induction hypothesis can be applied to find an appropriate path β' satisfying the strengthened statement of the lemma. Since the interval of α between $v_{\gamma'/\alpha}$ to $v'_{\mathrm{d}}=v_{\gamma/\alpha}$ covers the interval between $v_{\gamma/\alpha}$ and v_{d}, together with component C, the path $\beta=\beta'$ from u to v will also cover C, and will satisfy (i). If γ' does not overlap with the designated C'-branch of α, then, capitalizing on (ii) for β', we can achieve that the part of β' inside $C'=C_{\gamma/\alpha}$ becomes the obvious positive alternating path from $v'=v_\gamma$ to $v'_{\mathrm{d}}=v_{\gamma/\alpha}$ through the valve e. (See Fig. 7.) In this way β will consist of edges in γ and α only. On the other hand, the assumption that γ' does overlap with the designated C'-branch of α yields the same result directly by (ii) of the induction hypothesis. Thus, β satisfies

(ii) as well. Finally, if β' ends in a suffix that contains the entire designated C'-branch of α, then $\beta=\beta'$ will end in a suffix that contains the entire non-designated C-branch of α. Conversely, if the appropriate suffix for β' is the non-designated C'-branch, then (ii) ensures that β' consists of edges in α and γ only. This fact, together with (i), guarantees that the designated C-branch will do for β as the desired suffix. The proof is now complete. □

Theorem 6.2. *Let C_1, C_2 and C be viable elementary components of G such that $C_1\rho C$, $C_2\rho C$ and $C_1\neq C_2$. Then one of the following two statements holds.*
(a) *There exists a viable elementary component D and a D-loop α such that α covers all three of C_1, C_2 and C.*
(b) *There exists $i\in\{1,2\}$ and a C_i-loop α such that α covers both C_{3-i} and C.*

Proof. We follow the idea of the proof of Lemma 5.13, working in the graph G_h rather than in G. The switch is justified by the fact that the unfolding of any C_h-loop α_h in the spirit of Corollary 5.11 results in a C-loop α that covers the elementary components covered by α_h.

Let $C_i\rho C$ via some C_i-loops α_i with domain classes P_i. If C_i is covered by α_{3-i} for either of $i\in\{1,2\}$, then we are done. Suppose therefore that this is not the case, and, as a further initial assumption, let C_1 and C_2 be both internal. Consider two external alternating paths γ_1 and γ_2 entering C_1 and C_2 at some vertices w_1 and w_2 belonging to canonical classes $R_1\neq P_1$ and $R_2\neq P_2$. By Corollary 4.10, γ_i cannot reach any of the elementary components covered by α_i. Then the following two cases are possible.

Case a: for both $i=1,2$, γ_i does not overlap with either C_{3-i} or α_{3-i}.

As we observed in the proof of Lemma 5.13, there exists a negative alternating path β connecting a domain vertex v_1 of α_1 with a domain vertex v_2 of α_2, so that β consists of edges in α_1, α_2 and C_h only. Notice that, according to Lemma 6.1, we can now assume that β does in fact cover C. Clearly, γ_1 and γ_2 must overlap, otherwise there would be an alternating crossing in G_h containing γ_1, β and γ_2. Choose D to be the elementary component containing the last overlapping allowed edge on say γ_1 before it arrives in w_1. This component will obviously satisfy (a).

Case b: there exists $i\in\{1,2\}$ such that γ_i overlaps with either C_{3-i} or α_{3-i}.

Assume, without loss of generality, that $i=2$, and let D be the elementary component containing the last edge e on γ_2 before it reaches w_2 such that e is also on α_1 or C_1, see Fig. 8.

If $D=C_1$, then consider the vertex u where γ_2 finally leaves C_1 before reaching w_2. The suffix of γ_2 from u to w_2 can then be taken as one C_2-branch of a C_1-loop. The other C_2-branch of this loop is an appropriate path from either domain vertex of α_1 to a domain vertex of α_2 through component C. The existence of such a path is ensured by Lemma 6.1.

If $D\neq C_1$, then let D be located along the C-branch α_1^1 of α_1 associated with domain vertex v_1^1, and let the other branch, α_1^2, originate from domain vertex v_1^2. Furthermore, denote by u_1 and u_2 the vertices in D where α_1^1 arrives at, and subsequently leaves D when traversing e, and let u be the vertex where γ_2 leaves D after e. We claim that the valve e points γ_2, when coming from u, to the direction u_1 on α_1^1, so that $u\not\sim u_1$.

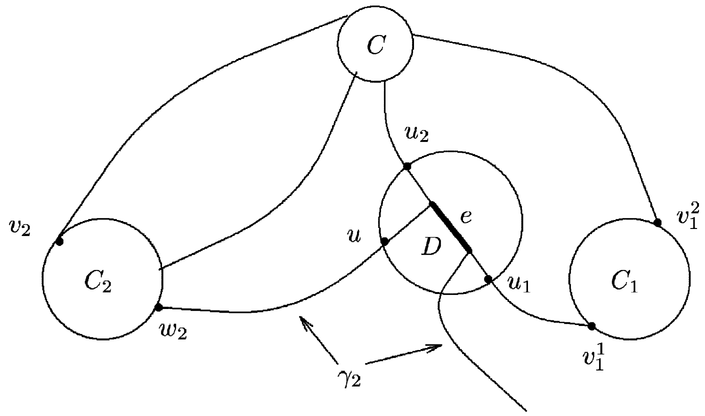

Fig. 8. The proof of Theorem 6.2.

Indeed, by Lemma 6.1, there exists a negative alternating path β_2 connecting u_2 with a domain vertex v_2 of α_2 in such a way that β_2 does not overlap with the interval γ of γ_2 between u and w_2. The assumption that the valve e on α_1^1 points γ_2 to u_2 would then give rise to an alternating cycle in G_h through β_2 and γ, after filling in the gaps with two appropriate positive alternating paths in C_2 and D. This verifies our claim that $u \not\sim u_1$.

Apply Lemma 6.1 again to establish a negative alternating path β from v_1^2 to a domain vertex v_2 of α_2, such that β covers C. Let δ_{C_2} denote a suitable positive alternating path connecting v_2 and w_2 in $(C_2)_h$. Furthermore, let δ_D be a positive alternating path in D connecting u_1 with u through the valve e, consisting of edges in α_1^1 and γ_2 only. We claim that β and the prefix α_1' of α_1^1 from v_1^1 to u_1 are edge-disjoint. If this were not the case, then let D' denote the elementary component containing the last edge f on β before it arrives in C_2 such that f is also on α_1'. Clearly, D' must be along the loop α_2, so that $C_2 \rho D'$. On the other hand, by appending the appropriate subpath of α_1^1 from D' to D and the path δ_D to γ, this path extends to a negative alternating path γ' connecting D' and C_2. The path γ', coupled with the suffix of β between D' and C_2 as D'-branches, will then determine a D'-loop covering C_2, which contradicts $C_2 \rho D'$.

It is now clear that a C_1-loop α satisfying (b) can be constructed by joining β, the paths γ and δ_D, and the paths α_1' and δ_{C_2}.

To finish the proof, we now drop our initial assumption that both C_1 and C_2 are internal. It is clear that only one of them, say C_1, is external. Consider the external alternating path γ_2 as described above, and observe that γ_2 must overlap with either C_1 or α_1. (See Claim 2.4.) Thus, the argument in Case b applies, and the proof is complete. □

For an elementary component D, the *scope* of D is the set $\{C \mid D \rho C\}$.

Theorem 6.3. *For every two-way elementary component C there exists a unique component $f(C)$ with the property that $f(C)\rho C$, and for every C' such that C is in the scope of C', either $f(C)=C'$ or C' is in the scope of $f(C)$.*

Proof. The uniqueness of $f(C)$ is obvious by Corollary 5.10. To prove its existence, let C_1 be an arbitrary elementary component such that $C_1\rho C$. Such a component exists by Lemma 5.12. If C_1 has the desired property of $f(C)$, then we are through. Otherwise there exists a component D such that $D\rho C$ but not $C_1\rho D$. By Theorem 6.2, there exists C_2 such that either $C_2=D$ and $C_2\rho C_1$, or $C_2\rho D$, $C_2\rho C_1$ and $C_2\rho C$. In both cases, $C_2\rho C_1$ and $C_2\rho C$. Continuing in this way, a sequence $C_1, C_2, \ldots, C_n, \ldots$ of elementary components can be constructed, so that for every $1\leqslant i\leqslant n$, $C_i\rho C$ and $C_{i+1}\rho C_i$. By Corollary 5.10 this sequence must be finite, therefore the last element of the sequence is $f(C)$. □

Definition 6.4. For every two-way elementary component C the *father* of C is the component $f(C)$ in Theorem 6.3.

Theorem 6.5. *Every two-way elementary component C is only accessible through its father by external alternating paths. Furthermore, for every external alternating path γ leading to C, the last vertex of γ that is in $f(C)$ belongs to a unique canonical class of that component, which is the common domain class of all $f(C)$-loops covering C.*

Proof. Corollary 5.11 implies that the father–son relationship in G_h is the same as that in G. Therefore we can carry out the proof in G_h.

By way of contradiction, assume that γ is an external alternating path entering C without visiting $f(C)$ first. Let $f(C)\rho C$ via an $f(C)$-loop α with domain class P. Then, according to Lemma 6.1, there exists an external alternating path β leading to a domain vertex v of α such that β covers C and it consists of the edges in γ, C_h and α only. Concerning $f(C)$, there exists an external alternating path β' entering $f(C)$ at vertex v' in canonical class $R\neq P$. As we have already observed several times, β and β' must overlap. Consider the elementary component D containing the last allowed edge on β in common with β' before it reaches v. By Lemma 6.1 and Corollary 4.10, D comes before C on β, so that both C and $f(C)$ are covered by an appropriate D-loop. This, however, contradicts that $f(C)$ is the father of C. The second statement of the theorem follows directly from Corollary 4.10. □

We say that two distinct elementary components C_1 and C_2 are *distant cousins* if they are in the same distance from their closest common ancestor in the family tree. Component C_1 is a *distant uncle* of component C_2 if the distance of C_1 from the closest common ancestor of C_1 and C_2 is one greater than that of C_2. Note that, according to these definitions, brothers are distant cousins and fathers are distant uncles, too.

Proposition 6.6. *If $D\rho C$ holds for viable elementary components D and C, then either D is a distant uncle of C, or C and D are distant cousins.*

Proof. By definition, $D\rho C$ implies that either $D=f(C)$ or $f(C)\rho D$. In the latter case we have either $f(C)=f(D)$ or $f(D)\rho f(C)$. Continuing in this fashion we obtain that there exists a smallest $i\geqslant 1$ such that either $f^{i-1}(D)=f^i(C)$ or $f^i(C)=f^i(D)$, where $f^0(D)=D$, and for every $j\geqslant 0$, $f^{j+1}(D)=f(f^j(D))$. If $f^{i-1}(D)=f^i(C)$, then D is a distant uncle of C, while $f^i(C)=f^i(D)$ means that C and D are distant cousins. □

Corollary 6.7. *If $D\rho C_1$ and $D\rho C_2$ holds for two distinct viable elementary components C_1, C_2, then either C_1 and C_2 are distant cousins, or C_i is a distant uncle of C_{3-i} for one of $i\in\{1,2\}$.*

Proof. Immediate by Proposition 6.6. □

We can look at the members of a viable family $\mathscr{F}$ as individuals belonging to a strange species with the following reproduction rules. A male individual is an elementary component of $\mathscr{F}$, and a female individual is a canonical class of some elementary component. All females are therefore dependents on a particular male for their lives. Males, on the other hand, are born together with their potential mates. (Observe polygamy.) A mating process initiated by male C with female P is associated with a C-loop α from domain P. The potential offsprings arising from this process are the elementary components covered by α. Note that, by Corollary 4.10, each son of C will have a unique mother, as one would normally expect. Not all elementary components along α are, however, offsprings of the couple (C,P). It may be the case that component D along α is already "alive" as a distant cousin or distant nephew of C. (See Proposition 6.6.) In this case we say that D is a stillborn son of the couple (C,P). Component D can be a stillborn son of several other components, but the transitive closure of the stillborn relationship is asymmetric by Corollary 5.10. That is, if components $C=D_1,D_2,\ldots,D_n$ for any $n\geqslant 2$ are such that D_{i+1} is a stillborn son of D_i, then $D_n\neq C$. The root, denoted $r(\mathscr{F})$, being the unique one-way component in the family, is the ultimate forefather of $\mathscr{F}$, the root of the family tree. The family $\mathscr{F}$ is called *external* if $r(\mathscr{F})$ is such.

Now let us have a closer look at the arrangement of the forbidden edges inside a family $\mathscr{F}$.

Theorem 6.8. *An edge e in a viable family $\mathscr{F}$ is impervious iff both endpoints of e are in the principal canonical class of the root. Every forbidden edge e connecting two different elementary components in $\mathscr{F}$ is part of a C-loop for some $C\in\mathscr{F}$.*

Proof. By Corollary 5.9, for every vertex v of a two-way elementary component C there exists an alternating path γ from a vertex u of the root to v inside $\mathscr{F}$ that is positive at the v end and negative at the u end. We also know from the construction in Lemma 5.8 that u is a domain vertex (i.e., non-principal), therefore γ can be extended to a positive external alternating path leading to v. The same holds true if v is an internal vertex of the root, but belongs to a non-principal canonical class. Thus, every edge incident with v is viable. This leaves room for any impervious edges inside $\mathscr{F}$ only between two vertices belonging to the principal canonical class of the root. As we have seen in Lemma 5.3, such edges are indeed impervious. This proves the first statement of the theorem. Observe that if one endpoint of a forbidden edge e is principal in the root, then the other must be such, too. Indeed, otherwise either e would connect two vertices in different canonical classes of the root, or e, as a negative alternating path, would violate Lemma 5.8(iii).

As to the second statement of the theorem, let e connect a vertex v_1 in C_1 with a vertex v_2 in $C_2 \neq C_1$. We already know that neither v_1 nor v_2 is principal. Assuming that C_i, $i \in \{1,2\}$, is different from $r(\mathscr{F})$, let γ_i be a path connecting a suitable vertex in $r(\mathscr{F})$ with v_i such that γ_i is negative at $r(\mathscr{F})$ and positive at v_i. (See Corollary 5.9.) Since γ_i is positive at v_i, γ_i does not pass through e. Therefore, if either $C_1 = r(\mathscr{F})$ or $C_2 = r(\mathscr{F})$, then we are done. By the same token, if both C_1 and C_2 are different from the root, but γ_{3-i} covers C_i for either $i = 1$ or 2, then e will be part of an appropriate C_i-loop. Otherwise, let C be the elementary component containing the last allowed edge on γ_1 before it reaches C_1 that is also on γ_2. Clearly, it is now this component for which there exists a C-loop containing the edge e. □

Although Theorem 6.8 provides much information about the nature of forbidden edges inside a viable family $\mathscr{F}$, we would like to be yet more specific as to which elementary components can and which cannot be connected by a forbidden edge.

Theorem 6.9. *Let e be a forbidden edge connecting two different elementary components C_1 and C_2 of a viable family $\mathscr{F}$. Then one of the following two conditions must be met.*
(i) *C_1 and C_2 are distant cousins,*
(ii) *C_1 and C_2 are in a distant uncle-nephew relationship with each other.*

Proof. As we observed in the proof of Theorem 6.8, one of the following two statements holds:
(1) $C_1 \rho C_2$ or $C_2 \rho C_1$;
(2) there exists $C \neq C_i$, $i = 1,2$, such that $C \rho C_1$ and $C \rho C_2$.
In any case, the statement of the theorem follows from Proposition 6.6 and Corollary 6.7. □

Theorems 6.8 and 6.9 provide a satisfactory description of the forbidden edges inside a family $\mathscr{F}$, so now we concentrate on the ones that connect different families.

Proposition 6.10. *Let e be a viable forbidden edge of G connecting two different families $\mathscr{F}_1$ and $\mathscr{F}_2$. Then both $\mathscr{F}_1$ and $\mathscr{F}_2$ are viable, and exactly one endpoint of e belongs to the principal canonical class of the root of either $\mathscr{F}_1$ or $\mathscr{F}_2$.*

Proof. As we have seen in Theorem 6.8, for every internal vertex v of a family $\mathscr{F}$, there exists a positive external alternating path leading to v—that is, v is accessible—iff

v is not a principal vertex of $r(\mathscr{F})$. By definition, at least one endpoint of the viable edge e is accessible, thus falls into a viable family. Since this endpoint is not principal, the other endpoint also marks a viable family, even if that endpoint is principal (as we wish to prove). Suppose now, by contradiction, that both endpoints v_1 and v_2 of e are non-principal, and let α_1 and α_2 be positive external alternating paths leading to v_1 and v_2, respectively. The paths α_1 and α_2 must overlap, so that there exists a C-loop α for an appropriate elementary component C, which loop contains e. This is a contradiction, for the endpoints of e are in different families. □

If e is a viable edge connecting families $\mathscr{F}_1$ and $\mathscr{F}_2$, then we write $e:\mathscr{F}_1 \mapsto \mathscr{F}_2$ to indicate that the principal endpoint of e is in $\mathscr{F}_2$.

Lemma 6.11. *Let $e_1:\mathscr{F}_1 \mapsto \mathscr{F}_2, \ldots, e_n:\mathscr{F}_n \mapsto \mathscr{F}_{n+1}$ $(n \geqslant 1)$ be viable edges among families $\mathscr{F}_i$, $1 \leqslant i \leqslant n+1$. Then $\mathscr{F}_1 \neq \mathscr{F}_{n+1}$.*

Proof. Assume, by contradiction, that $\mathscr{F}_{n+1} = \mathscr{F}_1$. Without loss of generality we can assume that the families $\mathscr{F}_1, \ldots, \mathscr{F}_n$ are all different. Then, using Corollary 5.9 and Proposition 6.10, we can construct a negative alternating path γ in G_h starting from a vertex u of $r(\mathscr{F}_1)$, going through the edges $e_1, \ldots, e_{n-1}$ and families $\mathscr{F}_2, \ldots, \mathscr{F}_n$, and returning to a vertex v of $\mathscr{F}_1$ via e_n, so that γ runs essentially outside $r(\mathscr{F}_1)$. We know that the vertex v is principal in $r(\mathscr{F}_1)$, while u is not. The path γ can then be closed inside $r(\mathscr{F}_1)$ to an inter-elementary alternating cycle, which is a contradiction. □

By Lemma 6.11, if $e:\mathscr{F}_1 \mapsto \mathscr{F}_2$ for some families $\mathscr{F}_1$ and $\mathscr{F}_2$, then $e':\mathscr{F}_1 \mapsto \mathscr{F}_2$ for all viable edges e' connecting $\mathscr{F}_1$ and $\mathscr{F}_2$. This establishes and justifies $\mapsto$ as a binary relation between viable families. Let $\overset{*}{\mapsto}$ denote the reflexive and transitive closure of $\mapsto$.

Theorem 6.12. *The relation $\overset{*}{\mapsto}$ is a partial order on the collection of all viable families of G, by which the external families are maximal elements.*

Proof. Immediate by Lemma 6.11. □

Our closing theorem characterizes the relationship between the viable and impervious parts of G. An impervious edge $e \in E(G)$ is called *principal impervious* if at least one of its endpoints belongs to the principal canonical class of the root of some viable family.

Theorem 6.13. *Removing the principal impervious edges from G disconnects the viable families from the impervious ones.*

Proof. Indeed, as we have seen earlier, any edge e incident with a viable family is impervious iff e is a principal impervious edge. Thus, the removal of these edges from G will leave no connection between the viable and the impervious families of G. □

7. Conclusion

We have given a complete description of the structure of elementary components in a graph G having a perfect internal matching. As a first step we proved that the augmentation of G by its hidden edges does not change the elementary decomposition of the graph. We also generalized the notion of canonical equivalence, and showed that $\sim|C = \sim_{C_h}$ holds for every elementary component C of G.

Viable elementary components have been classified as one-way or two-way, depending on whether they could be accessed by external alternating paths in one or more than one canonical class. It was demonstrated that every two-way elementary component C' is indeed two-way accessible from another viable elementary component C via an appropriate C-loop covering C'. The reflexive and transitive closure of the "two-way accessible" relationship between elementary components (relation ρ^*) was proved to be a partial order.

Elementary components have been grouped into families according to the partition determined by the smallest equivalence relation containing ρ. It was shown that each viable family contains a unique one-way elementary component, called the root of the family, and that every member of the family is only accessible through the principal canonical class of the root by external alternating paths.

A more sophisticated analysis of the relation ρ showed that the members of each viable family can be arranged in a family tree, reflecting the order in which they can be reached by external alternating paths. The father of elementary component C is the component $f(C)$ having the property that C is in the scope of $f(C)$, and whenever C is in the scope of any elementary component D, D is in the scope of $f(C)$. The mother of C is the unique canonical class of $f(C)$ from which all $f(C)$-loops covering C originate. Forbidden edges connecting two elementary components of the same viable family have been characterized in terms of the family relationship, and a partial order has been defined on the collection of viable families along the lines of forbidden edges connecting elementary components belonging to different families. Finally, impervious edges have been identified as ones that are incident with a vertex v in a viable elementary component C only if C is one-way and v lies in the principal canonical class of C.

References

[1] M. Bartha, E. Gombás, A structure theorem for maximum internal matchings in graphs, Inform. Process. Lett. 40 (1991) 289–294.

[2] M. Bartha, E. Gombás, The Gallai-Edmonds algebra of graphs, Research Report # 9105, Department of Computer Science, Memorial University, St. John's, Canada, 1991.

[3] M. Bartha, E. Gombás, On graphs with perfect internal matchings, Acta Cybernet. 12 (1995) 111–124.

[4] M. Bartha, H. Jürgensen, Characterizing finite undirected multigraphs as indexed algebras, Research Report # 252, Department of Computer Science, The University of Western Ontario, London, Canada, 1989.

[5] M. Bartha, M. Krész, Elementary decomposition of soliton automata, Acta Cybernet. 14 (2000) 631–652.

[6] F.L. Carter, Comformational switching at the molecular level, in: F.L. Carter (Ed.), Molecular Electronic Devices, Marcel Dekker, Inc., New York, 1982, pp. 51–72.
[7] J. Dassow, H. Jürgensen, Soliton automata, J. Comput. System Sci. 40 (1990) 154–181.
[8] J. Dassow, H. Jürgensen, Soliton automata with a single exterior node, Theoret. Comput. Sci. 84 (1991) 281–292.
[9] J. Dassow, H. Jürgensen, The transition monoids of soliton trees, Research Report # 331, Department of Computer Science, The University of Western Ontario, London, Canada, 1992.
[10] J. Dassow, H. Jürgensen, Soliton automata with at most one cycle, J. Comput. System Sci. 46 (1993) 155–197.
[11] F. Gécseg, H. Jürgensen, Automata represented by products of soliton automata, Theoret. Comput. Sci. 74 (1990) 163–181.
[12] L. Lovász, M.D. Plummer, Matching Theory, North-Holland, Amsterdam, 1986.