Discrete Mathematics 95 (1991) 1
North-Holland

Preface

This volume has arisen from a colloquium on infinite graph theory and combinatorics held at St. John's College, Cambridge, in July 1989. The meeting, perhaps the first of its kind, sought to bring together some of those graph theorists and combinatorialists most familiar with the infinite aspects of the theory. Its aim was to assess the state of the art in the discipline, to consider its links with other parts of mathematics, and to discuss possible directions for future development. Almost all of those invited were able to participate and have contributed to the present volume.

About half of the articles are research papers; the others are expository surveys of particular areas. They all seek to appeal to expert and non-expert alike. The research papers in particular are generally self-contained.

The Cambridge meeting benefited greatly from the generous support of St. John's College and its staff, early backing by the London Mathematical Society, and a special research grant from the SERC. I would like to thank the authors in this volume for making some of their contributions to the colloquium accessible to a wider audience; it remains to be hoped that the reader will derive as much enjoyment from the material as we did. Finally, my special thanks are due to C.St.J.A. Nash-Williams for contributing his own personal introduction.

Reinhard Diestel
Cambridge, June 1990

Elsevier Science Publishers B.V.

Discrete Mathematics 95 (1991) 3–4
North-Holland

Introduction

It has been reported that Dénes König, the author of the classic *Theorie der endlichen und unendlichen Graphen* (Leipzig, 1936), expressed a special liking for infinite graphs, which certainly receive substantial attention in his book. Nevertheless, the majority of combinatorialists seem to have concentrated on finite combinatorics, to the extent that it has almost seemed an eccentricity to think that graphs and other combinatorial structures can be either finite or infinite.

However, there seems to be no logical reason why combinatorial structures should 'usually' be finite, and indeed this would preclude many fascinating avenues of exploration. To a considerable extent, finite and infinite combinatorics are parts of the same subject. Most of the concepts of finite combinatorics and many of its results carry over (sometimes in more than one way) to the infinite case. Results and problems in infinite combinatorics often arise from seeking analogues of corresponding finite results, and sometimes the attempt to do this also leads to new ideas in finite combinatorics.

Nevertheless, infinite combinatorics has its own distinctive features. Some of its problems, such as certain ones involving *ends* of graphs, have no meaningful finite analogue but are closely related to other parts of mathematics. Sometimes problems which are difficult for finite structures become trivial or easy in the infinite case because infinite structures allow so much more room for manoeuvre. On the other hand, the passage from finite to infinite structures often introduces new difficulties, commonly, but not always, of a set-theoretic character. Sometimes these two phenomena occur together, i.e. passing to the infinite case of a problem may reduce some difficulties while introducing others. Increasingly, interactions between infinite combinatorics and mathematical logic are coming to light.

So far as I know, the very interesting and successful Cambridge conference on *Directions in Infinite Graph Theory and Combinatorics* was the first conference devoted wholly to infinite combinatorics. Therefore this conference, and the appearance of this volume, may be hailed as something of a landmark, showing increasing recognition of the far-reaching possibilities of infinite combinatorics after a period of comparative neglect. It provides an excellent introduction to the field and pointers to other relevant literature. Several of its papers show the increasing interaction of infinite combinatorics with mathematical logic and set

Elsevier Science Publishers B.V.

theory, as well as with aspects of topology and algebra. For all these reasons, the volume seems likely to play a major part in generating interest in infinite combinatorics, and deserves to be widely distributed and read.

C.St.J.A. Nash-Williams

Discrete Mathematics 95 (1991) 5–22
North-Holland

Infinite matching theory

Ron Aharoni

Department of Mathematics, Technion—Israel Institute of Technology, Haifa 32000, Israel

Received 3 October 1989
Revised 15 January 1990

Dedicated to C.St.J.A. Nash-Williams.

Abstract

Aharoni, R., Infinite matching theory, Discrete Mathematics 95 (1991) 5–22.

We survey the existing theory of matchings in infinite graphs and hypergraphs, with special attention to the duality between matchings and covers. Some results are presented which have not appeared elsewhere, mainly concerning Menger's theorem for infinite graphs.

1. Introduction

The study of a theory which encompasses, among other concepts, such a basic notion as that of injections between infinite sets, hardly needs further advocacy. But if one is required, an interesting observation is that infinite matching theory can boast a result which is probably as early as any result in finite matching theory, if not earlier.

Theorem 1.1. *If A and B are subsets of the two respective sides of a bipartite graph, and if there exist two matchings covering A and B respectively, then there is a matching covering $A \sqcup B$.*

This is a re-formulation of the famous Cantor-Bernstein Theorem (see, e.g. [20, p. 23]).

There is also no doubt that infinite matching theory, besides its inherent interest, can illuminate some points in finite matching theory. For example, a result saying that the infinite version of a theorem is 'hard' to prove (see [9]) probably reflects on the complexity of the finite problem, although no explicit connection has been established as yet.

Matching theory can be divided very roughly, but usefully, into two parts, circling around two themes: criteria for matchability (exemplified by the theorems of Hall and Tutte) and the duality between matchings and covers (the outstanding results here are the theorems of König and Menger). The two themes are very

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

closely knit together, and an important fact on the link between them should be mentioned here: whereas in the finite case duality results can be proved directly, it seems that in the infinite case they have to be proved via their counterparts on criteria for matchability. The reason is that in the finite case one can start from a maximal matching and produce the dual cover from it, an endeavour which is pretty meaningless in the infinite case (maximal matchings exist with respect to containment, but a stronger notion of maximality is needed here, see Theorem 5.3.)

Here is a somewhat brief historical survey. The first modern treatment of infinite matching theory is in the book of König [21]. It contains an extension due to Erdős of Menger's theorem to the infinite case. In 1935, P. Hall published his theorem [17] which quickly became famous, a fate which befell its predecessors, the theorems of König and Menger (of which it is an easy consequence) only much later.

Theorem 1.2 (P. Hall's Theorem). *In a finite bipartite graph with bipartition (M, W) there exists a matching of M if and only if every subset A of M is connected to at least $|A|$ vertices in W.*

It was well known that the theorem fails in the infinite case. There is a basic example showing this: a bipartite graph whose vertex set is $\{m_i: 0 \leqslant i < \omega\} \sqcup \{w_i: 1 \leqslant i < \omega\}$ and edge set $\{(m_i, w_i): 1 \leqslant i < \omega\} \sqcup \{(m_0, w_i): 1 \leqslant i < \omega\}$ (m_0 is sometimes called the 'playboy'). Every set A of m_i's is connected to at least $|A|$ elements w_i, but there is no matching of the m_i's. It was felt that the example is cannonical in some sense for the countable case. This was confirmed in the early 1970's on apparently two different lines of ideas. One was due to Nash-Williams, who proposed a somewhat complicated function which counts the difference between the size of a set of women (elements of W) and the number of men (elements of M) who must be married (matched) into this set. Damerell and Milner [13] proved a conjecture of Nash-Williams that, in the countable case, there is a marriage of all men if and only if this function is always nonnegative. Later [22] Nash-Williams simplified his own criterion. But there was another line of ideas, which culminated in Podewski and Steffens' criterion on 'critical sets' ([24], see Section 3). In [10] the two criteria were shown to be equivalent. In hindsight it seems that, because of its simplicity, the 'critical sets' criterion is the more useful one, in spite of its apparent circular requirement that the critical set itself should be matchable. For example, Podewski and Steffens were able to prove, using it, the countable case of Erdős' conjecture on the extension of König's theorem to the infinite case.

The general problem was settled in 1983, when a criterion for matchability was proved for general bipartite graphs [10]. Basically, it confirmed a feeling that the most general obstacle for the existence of an injective choice function in the

uncountable case is one which forces it to be regressive on a stationary set. This criterion was used to prove König's theorem for general graphs in [3].

As for general graphs, the finite criterion for matchability was found by Tutte [27]. Later he proved that his criterion is valid also for locally finite graphs [28]. (The corresponding extension of P. Hall's theorem, namely that P. Hall's condition is sufficient if the degree of every vertex in the set of 'men' is finite, was proved by M. Hall [18]. But this is nowadays an easy exercise in compactness.) Steffens [26] gave a criterion for matchability in countable graphs, which was later used to prove a Tutte-like result for countable graphs [4]. The general problem was solved in [6].

Although this is a survey paper, we have included a few new results. The survey is not intended in any way to be comprehensive. There is an excellent book on the bipartite case [19] which is much more detailed on this subject.

2. Definitions

A graph will be usually denoted by $G=(V, E)$, where V is the set of vertices and E the set of edges (i.e. $E \subseteq [V]^2$). If G is bipartite with bipartition $V = M \sqcup W$, it will be denoted also by (M, W, E). We then write: $M = M_G$, $W = W_G$, $E = E_G$. If X is a subset of V we write $G[X]$ for the subgraph of G induced by X. The graph $G[V \setminus X]$ is denoted by $G - X$. If $\Gamma = (M, W, E)$ is bipartite and $X \subseteq V$, we assume that $\Gamma[X]$ has the bipartition $(M \cap X, W \cap X)$. A *matching* in a graph is a set of disjoint edges. A *cover* is a set of vertices meeting all edges. For a given subset F of E, a vertex a and a subset A of V, we write

$$F\langle a\rangle = \{v \in V : \{a, v\} \in F\}, \qquad F[A] = \bigcup \{F\langle a\rangle : a \in A\} \quad \text{and}$$

$$F \lceil A = \{f \in F : f \cap A \neq \emptyset\}.$$

If $|F\langle a\rangle| = 1$ we write $F(a)$ for the single element of $F\langle a\rangle$. A matching F is said to *cover* a set A if $A \subseteq F[V]$. If such F exists, then A is said to be *matchable* and if, in such a case, $F[A] \subseteq B$, we say that A is *matchable into* B. For any subset X of V we write $D(X) = \{v \in V : E\langle v\rangle \subseteq X\}$. An *espousal* in a bipartite graph Γ is a matching covering M_Γ. If such a matching exists, then Γ is said to be *espousable*.

We need some terminology concerning paths. The vertex set of a path P is denoted by $V(P)$, and its edge set by $E(P)$. If P has a first vertex, it is denoted by $\mathrm{in}(P)$, and if it has a last vertex, it is denoted by $\mathrm{ter}(P)$. If $\boldsymbol{P}$ is a family of paths, we write $\boldsymbol{P}^f$ for the set of finite paths in $\boldsymbol{P}$. We also write

$$V[\boldsymbol{P}] = \bigcup \{V(P) : P \in \boldsymbol{P}\}, \qquad E[\boldsymbol{P}] = \bigcup \{E(P) : P \in \boldsymbol{P}\},$$

$$\mathrm{in}[\boldsymbol{P}] = \{\mathrm{in}(P) : P \in \boldsymbol{P} \text{ has a vertex}\} \quad \text{and}$$

$$\mathrm{ter}[\boldsymbol{P}] = \{\mathrm{ter}(P) : P \in \boldsymbol{P} \text{ has a last vertex}\}.$$

If P and Q are paths such that $V(Q) \cap V(P) = \{\text{in}(Q)\} = \{\text{ter}(P)\}$ then $P * Q$ denotes the concatenation of P and Q. If P and Q are two families of disjoint paths and $V[\boldsymbol{Q}] \cap v[\boldsymbol{P}] = \text{in}[\boldsymbol{Q}] \subseteq \text{ter}[\boldsymbol{P}]$ then $\boldsymbol{P} * \boldsymbol{Q}$ denotes the family $\{P * Q\colon P \in \boldsymbol{P},\ Q \in \boldsymbol{Q} \text{ and } \text{ter}(P) = \text{in}(Q)\}$.

3. Bipartite graphs

Some people regard Hall's Theorem as the cornerstone of finite matching theory, but others, the author included, feel that this title is due to the closely related Theorem of König (which has the advantages of being more symmetrical, of implying Hall's Theorem more easily than the other way round, and also of historical precedence!). Here it is, in a form which is true also in the infinite case:

Theorem 3.1 [3]. *In any bipartite graph there exist a matching F and a cover C such that C consists of the choice of precisely one vertex from each edge in F.*

A pair (F, C) as in the theorem is called *orthogonal.*

Let $\Gamma = (M, W, K)$ be a bipartite graph, and let F and C be as in the theorem. Let $X = C \cap W$ and $D = M \backslash C$. Since C is a cover, there is no edge from D to $W \backslash C$, which means that $K[D] \subseteq X$. Since obviously $X = F[D] \subseteq K[D]$ we have $X = K[D]$. Now assume that Γ is inespousable. Then D is unmatchable, for if I were a matching of D, then $I \sqcup (F \restriction (M \cap C))$ would be an espousal, since $I[D] \subseteq X \subseteq W \backslash F[M \cap C]$. Hence we have the following, which should be considered as the infinite version of the marriage theorem:

Theorem 3.2. *Γ is inespousable if and only if there exists an unmatchable subset D of M such that $K[D]$ is matchable into D.*

Note that in the finite case the conditions on D imply $|K[D]| < |D|$, and thus Hall's Theorem is a particular case of Theorem 3.2.

How does one go about proving Theorem 3.1? Here we are about to present one of the protagonists of the play: the *wave.* A *wave* $\boldsymbol{W}$ is a set of vertex disjoint paths such that $\text{in}[\boldsymbol{W}] \subseteq M$, and $\text{ter}[\boldsymbol{W}]$ is a cover. Note that although Theorem 3.1 is symmetrical in M and W, the notion of a 'wave' is not. In fact, we shall introduce an even greater a-symmetry—we are considering here Γ as a directed graph, in which the edges are directed from M to W. Thus the paths in $\boldsymbol{W}$ consist of either one edge or just one vertex. There exists at least one wave: $\{(m)\colon m \in M\}$. This is called the *trivial* wave. We define an order $\leqslant$ on the waves: $\boldsymbol{W} \leqslant \boldsymbol{U}$ if $\text{in}[\boldsymbol{U}] \subseteq \text{in}[\boldsymbol{W}]$ and every path in $\boldsymbol{U}$ is an extension of some path in $\boldsymbol{W}$. The trivial wave is minimal in this order. A wave $\boldsymbol{W}$ is called *tight* if there does not exist a set $\boldsymbol{V}$ of vertex disjoint paths such that $\text{in}[\boldsymbol{V}] = \text{in}[\boldsymbol{W}]$ and $\text{ter}[\boldsymbol{V}] \subsetneqq \text{ter}[\boldsymbol{W}]$. A wave $\boldsymbol{W}$ is called *full* if $\text{in}[\boldsymbol{W}] = M$. Otherwise it is called a

hindrance. A tight hindrance is called a *1-obstruction* (a slightly different definition is given in Section 4). If Γ contains a hindrance, it is called *hindered*, and if it contains a 1-obstruction, it is called *1-obstructed*. For a given wave $\boldsymbol{W}$ we denote by ess($\boldsymbol{W}$) the set $\boldsymbol{W}\backslash\{P \in \boldsymbol{W}: \mathrm{ter}[\boldsymbol{W}\backslash\{P\}]$ is a cover$\}$. It is easy to show that ess($\boldsymbol{W}$) is a wave. If $(\boldsymbol{W}_\alpha: \alpha < \theta)$ is a $\leqslant$-ascending sequence of full waves, then $\uparrow_{\alpha<\theta}\boldsymbol{W}_\alpha$ denotes the family of paths

$$\{(m, w): m \in \mathrm{in}[\boldsymbol{W}_\alpha] \text{ for all } \alpha \text{ and } (m, w) \in \boldsymbol{W}_\alpha \text{ for some } \alpha\}$$
$$\sqcup \{(m): (m) \in \boldsymbol{W}_\alpha \text{ for all } \alpha\}.$$

(This is just the result of 'growing' the waves along the sequence. The reason for the restriction of $\leqslant$ to full waves is that for general waves Lemma 3.3 is false. Another way of overcoming this difficulty is the introduction of "billows", as is done in Section 4.)

Lemma 3.3. *$\uparrow_{\alpha<\theta}\boldsymbol{W}_\alpha$ is a wave. Moreover, if each $\boldsymbol{W}_\alpha$ is tight, then so is $\uparrow_{\alpha<\theta}\boldsymbol{W}_\alpha$.*

For the proof of the first part, note that if $\theta = \zeta + 1$, then $\uparrow_{\alpha<\theta}\boldsymbol{W}_\alpha = \boldsymbol{W}_\zeta$, which is, of course, a wave. If θ is a limit ordinal, then $S = \mathrm{ter}[\uparrow_{\alpha<\theta}\boldsymbol{W}_\alpha]$ is a cover, since if $(m, w) \in K$ and $m \notin \mathrm{ter}[\boldsymbol{W}_\alpha]$ for some α, then $w \in \mathrm{ter}[\boldsymbol{W}_\beta]$ for all $\beta > \alpha$ (because $\boldsymbol{W}_\beta$ is a wave and $m \notin \mathrm{ter}[\boldsymbol{W}_\beta]$) and hence $w \in S$. The second part is left as an exercise.

By Lemma 3.3 there exists a $\leqslant$-maximal full wave $\boldsymbol{U}$. Let $\boldsymbol{W} = \mathrm{ess}(\boldsymbol{U})$. The set $\mathrm{ter}[\boldsymbol{W}]$ will serve as the set C of Theorem 3.1. (Note how easy it is to find C in comparison to the construction of F!) We also have at hand one part of F: the set of paths in $\boldsymbol{W}$ consisting of an edge. Let

$$A = \mathrm{ter}[\boldsymbol{W}] \cap M, \qquad B = \mathrm{ter}[\boldsymbol{W}] \cap W, \quad \text{and} \quad \Gamma' = \Gamma[A \sqcup W\backslash B].$$

Since $C = A \sqcup B = \mathrm{ter}[\boldsymbol{W}]$ is a cover, it suffices to show that Γ' is espousable in order to conclude the proof. For if H is an espousal of Γ', then the matching $(\boldsymbol{W} \lceil B) \sqcup H$ is obviously orthogonal to C. Now, the maximality of $\boldsymbol{U}$ implies that Γ' contains no nontrivial wave. Hence it suffices to show that if a bipartite graph contains no nontrivial wave, then it is espousable. In fact, the following somewhat stronger theorem is true.

Theorem 3.4. *An unhindered graph is espousable.*

Note that this theorem looks like a much weaker version of Theorem 3.2 (to see this assume that a graph Γ is inespousable. By Theorem 3.2 there exists an unmatchable subset D of M such that $K[D]$ has a matching I into D. Then $I \sqcup \{(m): m \in M\backslash D\}$ is a hindrance in Γ.) But in fact, by the preceding argument, it carries the whole weight of Theorem 3.2.

The proof of Theorem 3.4 depends on a criterion for espousability proved by Podewski and Steffens [24] in the countable case, and by Aharoni, Nash-Williams

and Shelah [10] in the uncountable case. Podewski and Steffens proved the following stronger result in the countable case.

Theorem 3.5. *A countable bipartite graph is espousable if and only if it is not 1-obstructed.*

(In [9] Theorem 3.4 is given, in the countable case, a more direct proof, although the two proofs are, at base, the same. They both have an alternating paths argument at their core.)

Here is a nice corollary of this theorem, which extends to the infinite case a well-known result of Birkhoff [12].[1]

Corollary 3.5a. *An infinite doubly-stochastic matrix contains a nonzero generalized diagonal.*

In terms of bipartite graphs, what the corollary says is this: if Γ is a bipartite graph and f a nonnegative real valued function on $E = E_\Gamma$ such that $\sum_{x \in e} f(e) = 1$ for every $x \in V$, then Γ has a perfect matching on which f is positive.

Outline of proof. Let $H = \{e \in E: f(e) > 0\}$ and let $\Gamma' = (V, H)$. Every vertex in Γ' has a countable degree, and hence every connected component of Γ' is countable. It suffices to show that every component Δ of Γ' is matchable, and by Theorem 1.1 one has to prove only that M_Δ is matchable. If not then, by the theorem, there exists in Δ a tight wave $\boldsymbol{W}$ and a vertex $a \in M_\Delta \backslash \mathrm{in}[\boldsymbol{W}]$. Let $F = E[\boldsymbol{W}]$ and let $U = F[W_\Delta] \sqcup \{a\}$ (i.e. U consists of a, together with the set of initial vertices of paths in $\boldsymbol{W}$ consisting of an edge.) Define a weight function w on the edges of the complete directed graph D_U on U by: $w(u, v) = f(u, F(v))$ $(u \in U, v \in U \backslash \{a\}, u \neq v)$ and $w(u, a) = 0$ for $u \neq a$. Then, for every vertex $v \in U \backslash \{a\}$ there holds

$$\sum_{u \in U \backslash \{v\}} w(u, v) \leqslant \sum_{u \in U \backslash \{v\}} w(u, v) \tag{*}$$

(the right-hand side is equal to $1 - f(v, F(v))$), while

$$\sum_{u \in U \backslash \{a\}} w(a, u) = 1, \quad \sum_{u \in U \backslash \{a\}} w(a, u) = 0 \tag{**}$$

This is easily seen to imply the existence of an infinite path $a, u_1, u_2, \ldots$ in D_U on whose edges w is positive. (A first step towards a proof of this is removing all positive circuits in D_U. This keeps both (*) and (**), and for acyclic graphs the claim is easy.) But then there exists in Δ an infinite F-alternating path starting at a, which contradicts the tightness of $\boldsymbol{W}$ (one sees this by replacing, in $\boldsymbol{W}$, the edges $\{(u_i, F(u_i)): i < \omega\}$ by $\{(u_i, F(u_{i+1})): i < \omega\}$. The resulting matching misses $F(u_1)$.)

[1] Note added in proof: the result was proved by a different method by J.R. Isbell, Birkhoff's problem 111, Proc. Amer. Math. Soc. 6 (1955) 217–218.

In the uncountable case there exist more complicated possible obstacles for espousability. Let Y be the set of all regular uncountable cardinals, together with 1. We define the notion of a *κ-obstruction* for all $\kappa \in Y$ by induction on κ. The notion of a 1-obstruction is already defined. If $\boldsymbol{W}$ is a 1-obstruction and $a \in M \backslash \mathrm{in}[\boldsymbol{W}]$ then, denoting $F = K \cap \boldsymbol{W}$, we call (by an abuse of language) also the graph $\Gamma[\{a\} \sqcup F[V]]$ a 1-obstruction in Γ. For $\kappa > 1$, a *κ-obstructive sequence* is a sequence $((M_\alpha, W_\alpha): \alpha < \kappa)$ of pairs of sets, such that $M_\alpha \subseteq M$, $W_\alpha \subseteq W$, $M_\alpha \cap M_\beta = W_\alpha \cap W_\beta = \phi$ for $\alpha \neq \beta$, $K[M_\alpha] \subseteq \bigcup \{W_\beta: \beta \leqslant \alpha\}$; for each $\alpha < \kappa$, either:

(i) $M_\alpha = \phi$ and $|W_\alpha| = 1$, or,

(ii) $\Gamma[M_\alpha, W_\alpha]$ is a μ-obstruction (in itself) for some $\mu < \kappa$,

and the set $\{\alpha$: (ii) occurs at $\alpha\}$ is stationary.

The graph $O = \Gamma[\bigsqcup_{\alpha<\kappa} M_\alpha \sqcup \bigsqcup_{\alpha<\kappa} W_\alpha]$ is then called a *κ-obstruction in Γ*. (This can also be taken as a special case of a definition in Section 4.) Γ is called *obstructed* if it contains a κ-obstruction for some $\kappa \in Y$.

Theorem 3.6 [10]. *A graph is espousable if and only if it is unobstructed.*

The 'only if' part follows quite easily from Fodor's Lemma. The 'if' part is harder.

It is also not easy to derive Theorem 3.4 from Theorem 3.6. But we can exhibit one main idea in a simple case. Our aim is to show that if Γ is inespousable, then it contains a hindrance. By Theorem 3.6, Γ contains some κ-obstruction T. If $\kappa = 1$, then T is a hindrance, and we are done. So, assume that κ is regular and uncountable. Let Ψ be the set of ordinals α of type (ii) in T. Here we shall greatly simplify matters and assume that, for each $\alpha \in \Psi$, the obstruction $\Gamma[M_\alpha \sqcup W_\alpha]$ is of a particularly simple type: $W_\alpha = \phi$ (i.e. $K[M_\alpha] \subseteq \bigcup \{W_\beta: \beta < \alpha\}$) and $|M_\alpha| = 1$. Let $M_\alpha = \{m_\alpha\}$ whenever $\alpha \in \Psi$ and $W_\alpha = \{w_\alpha\}$ for $\alpha \notin \Psi$. For $\alpha \notin \Psi$ define

$$\theta_\alpha = \{\beta \in \Psi: \beta > \alpha \quad \text{and} \quad (m_\beta, w_\alpha) \in K\}$$

Let

$$X = \{w_\alpha: \theta_\alpha \text{ is stationary}\}, \qquad Y = W_T \backslash X,$$

$$A = K[Y] \cap M_T \quad \text{and} \quad D = M_T \backslash A.$$

(The set X consists of those w_α connected to 'many' m_α's, and D is the set of m_α's connected only to X.)

Assertion 3.7. $\theta_Y := \{\beta \in \Psi: m_\beta \in A\}$ *is nonstationary.*

Proof. For each $m_\beta \in A$ choose $\alpha = f(\beta)$ such that $(m_\beta, w_\alpha) \in K$ and $w_\alpha \in Y$. Then $f(\beta) < \beta$ for all $\beta \in \theta_Y$, and hence if θ_Y is stationary, then by Fodor's lemma there exists some $w_\alpha \in Y$ such that $\{\beta: f(\beta)\}$ is stationary. But then θ_β is stationary, contradicting the fact that $w_\alpha \in Y$. $\square$

By Assertion 3.7, $\theta_\alpha \backslash \theta_Y$ is stationary for all $w_\alpha \in X$. Hence $|K\langle w_\alpha \rangle \cap D| = \kappa$ whenever $w_\alpha \in X$. Since $|X| \leqslant \kappa$, this easily implies that there exists a matching of X into D. The definitions imply that $K[D] \subseteq X$, and since $K\langle x \rangle \cap D \neq \phi$ for all $x \in X$, we have $K[D] = X$. Since $\Psi \backslash \theta_Y$ is stationary, $\Gamma[D \sqcup X]$ is a κ-obstruction, and thus D is unmatchable. Hence D satisfies the conditions in Theorem 3.2. As already noted, this proves also Theorem 3.4.

4. Webs

A *web* is a triple $\Gamma = (G, A, B)$ where $G = (V, E)$ is a directed graph, and $A, B \subseteq V(G)$. We write S_Γ for A (the S stands for 'source'). A finite path P is an *A–B path* if $V(P) \cap A = \{\text{in}(P)\}$, $V(P) \cap B = \{\text{ter}(P)\}$. A subset S of V is called *A–B separating* if $V(P) \cap S \neq \phi$ for all A–B paths P. Since we shall be mainly concerned with A–B paths and A–B separating sets, we may assume that there are no edges in Γ which go out of B or go into A, since the deletion of such edges does not alter the above mentioned objects. A web is called *bounded* if it contains no infinite paths.

A *ripple* in Γ is a set of vertex-disjoint paths, each starting at a vertex from A (in [5] this was called a 'warp'). A ripple $\boldsymbol{R}$ is called an *A–B ripple* if all paths in it are A–B paths. If, in addition, $\text{in}[\boldsymbol{R}] = A$, then $\boldsymbol{R}$ is called a *linkage*. If Γ contains a linkage, it is called *linkable*. A ripple $\boldsymbol{W}$ is called a *wave* if $\text{ter}[\boldsymbol{W}]$ is A–B separating. (Note that here we do not require finiteness of the paths—they may be unending. This is different from the terminology of [5]). A wave $\boldsymbol{W}$ is called *tight* if there does not exist a ripple $\boldsymbol{R}$ such that $\text{in}[\boldsymbol{R}] = \text{in}[\boldsymbol{W}]$ and $\text{ter}[\boldsymbol{R}] \subsetneqq \text{ter}[\boldsymbol{W}]$.

Given a wave $\boldsymbol{W}$ we write $\Gamma[\boldsymbol{W}]$ for the web $(G, A, \text{ter}[\boldsymbol{W}])$. If $\Gamma[\boldsymbol{W}]$ is not linkable, then $\boldsymbol{W}$ is called an *impediment*, and if Γ contains an impediment it is called *impeded*. Clearly, an impeded web is not linkable. If $\text{in}[\boldsymbol{W}] \neq A$ then $\boldsymbol{W}$ is called a *hindrance*, and if a hindrance exists Γ is called *hindered*.

The notions of 'linkage' and 'A–B separating set' generalize, respectively, the notions of 'espousal' and 'cover' in the bipartite case. Thus, a linkage is an espousal through a "medium" of vertices in $V \backslash (A \sqcup B)$. So, it is natural to ask for the generalizations of the main theorems of Section 3 to webs. The generalization of Theorem 3.1 corresponds to Menger's theorem in the finite case, and is the subject of a famous conjecture of Erdős:

Conjecture 4.1. In any web there exists an A–B ripple $\boldsymbol{R}$ and an A–B separating set S such that S consists of the choice of precisely one vertex from each path in $\boldsymbol{R}$.

As in the bipartite case, the first breakthrough was made by Podewski and Steffens [25]. They proved the conjecture for countable bounded webs. Since

then it has become clear that the main obstacle for the proof of Conjecture 4.1, or rather, what makes the conjecture harder than Theorem 3.1, is the possible existence of infinite paths. The difficulty is that trying to 'grow' A–B paths starting from A, one may end up with unending paths, rather than A–B paths. In fact, in [2] it was noted that in the absence of infinite paths, Conjecture 4.1 follows easily from Theorem 3.1 (so, the conjecture is now known for all bounded webs). The way to see this is to assign to Γ a bipartite graph $\hat{\Gamma}$ in the following way. To each vertex $a \in A$ assign a vertex a' of $\hat{\Gamma}$, to each $b \in B$ a vertex b'', and to each $v \in V \backslash (A \sqcup B)$ two vertices, v' and v''. Let

$$V' = \{v' : v \in A \sqcup (V \backslash B)\}, \qquad V'' = \{v'' : v \in B \sqcup (V \backslash A)\},$$

and

$$\hat{E} = \{(x', y'') : (x, y) \in E\} \sqcup J, \quad \text{where} \quad J = \{(x', x'') : x \in V \backslash (A \sqcup B)\}.$$

Let $\hat{\Gamma} = (V', V'', \hat{E})$. Choose F and C, a matching and a cover in $\hat{\Gamma}$, as in Theorem 3.1. Let

$$S = \{v \in V \backslash (A \sqcup B) : \{v', v''\} \subseteq C\} \sqcup \{a \in A : a' \in C\} \sqcup \{b \in B : b'' \in C\},$$

and let $\boldsymbol{R}$ be the set of maximal paths P in Γ satisfying $(x', y'') \in F$ whenever $(x, y) \in E(P)$, and P contains a vertex from S. It is not hard to show that S is A–B separating; that every $s \in S$ is contained in a path $Q \in \boldsymbol{R}$ such that $\mathrm{ter}(Q) \in B$ if $\mathrm{ter}(Q)$ exists and $\mathrm{in}(Q) \in A$ if $\mathrm{in}(Q)$ exists; and that every path in $\boldsymbol{R}$ contains precisely one vertex from S. If Γ is bounded, then all paths in $\boldsymbol{R}$ must be A–B paths, and thus $\boldsymbol{R}$ is an A–B ripple which satisfies, together with S, the conditions of Conjecture 4.1.

Things are even simpler, in the bounded case, with regard to criteria for linkability (the analogues of Theorems 3.2 and 3.6). For, the web Γ is linkable if and only if $\hat{\Gamma}$ is espousable. To see this, assume that Γ is linkable, and let $\boldsymbol{L}$ be a linkage of Γ. Then the set

$$\{(x', y'') : (x, y) \in E[\boldsymbol{L}]\} \sqcup \{(v', v'') : v \in V \backslash V[\boldsymbol{L}])\}$$

is an espousal in $\hat{\Gamma}$. Conversely, assume that F is an espousal of $\hat{\Gamma}$. Given any $a \in A$, let v_1 be such that $(a', v_1'') \in F$. If $v_1 \notin B$ then v_1' exists, and $(v_1', v_1'') \notin F$ (since $(a', v_1'') \in F$). Hence $(v_1', v_2'') \in F$ for some v_2. If $v_2 \notin B$ then $(v_2', v_3'') \in F$ for some vertex v_3. Continuing this way, since Γ is bounded, $v_k \in B$ for some k. Connect then a to B by the path $P_a = (a, v_1, \ldots, v_k)$. Since F is a matching, $V(P_{a_1}) \cap V(P_{a_2}) = \phi$ whenever $a_1 \neq a_2$, and thus we have constructed a linkage in Γ. (Note that we have used only the absence of unending paths!)

In order to formulate the criterion for linkability which the above observations yield (together with Theorems 3.2 and 3.6) we have to define "κ-obstructions" for webs, for $\kappa \in Y$. The definition is more involved here than in the special case of bipartite webs, discussed in Section 3. We really have to go into the inner structure of the obstructions.

First, we have to introduce a slight modification of the notion of a wave. A *billow* is a pair $\boldsymbol{B} = (\boldsymbol{W}, \boldsymbol{U} = \in(\boldsymbol{B}))$ of waves where $\mathrm{in}[\boldsymbol{W}] = A$ and $\boldsymbol{U} \subsetneq \boldsymbol{W}^f$ (the

set of finite paths in $\boldsymbol{W}$). If $\boldsymbol{C}=(\boldsymbol{Z}, \boldsymbol{Y})$ is another billow, we write $\boldsymbol{C} \leqslant \boldsymbol{B}$ if $\mathrm{in}[\boldsymbol{Y}] \subseteq \mathrm{in}[\boldsymbol{U}]$ and every path in $\boldsymbol{Z}$ is an extension of some path in $\boldsymbol{W}$. (The idea in these definitions is that although the part $\boldsymbol{W} \setminus \in (\boldsymbol{B})$ is 'redundant' in terms of $A-B$ separation, it is necessary for the definition of the relation $\leqslant$, since it keeps a path P in $\boldsymbol{Z}$ from intersecting the 'redundant' part of $\boldsymbol{W}$, except for the initial part of P which belongs to $\boldsymbol{W}$). We write $\Gamma/\boldsymbol{B}$ for the web $(G-(V[\boldsymbol{W}] \setminus \mathrm{ter}[\boldsymbol{U}]), \mathrm{ter}[\boldsymbol{U}], B)$. If $\boldsymbol{B}_\alpha=(\boldsymbol{W}_\alpha, \boldsymbol{U}_\alpha)$ $(\alpha<\theta)$ is a $\leqslant$-ascending sequence of billows, we write $\uparrow_{\alpha<\theta} \boldsymbol{B}_\alpha$ for the pair $(\boldsymbol{W}, \boldsymbol{U})$, where $\boldsymbol{W}$ is the ripple defined by $E[\boldsymbol{W}]=\bigsqcup\{E[\boldsymbol{W}_\alpha]: \alpha<\theta\}$ and $\boldsymbol{U}$ the sub-ripple defined by $\mathrm{in}[\boldsymbol{U}]=\bigcap\{\mathrm{in}[\boldsymbol{U}_\alpha]: \alpha<\theta\} \cap \mathrm{in}[\boldsymbol{W}^f]$.

Lemma 4.2. *$\uparrow_{\alpha<\theta} \boldsymbol{B}_\alpha$ is a billow.*

The *trivial billow*, denoted by $\square=\square_\Gamma$, is the pair $(\boldsymbol{W}, \boldsymbol{W})$, where $\boldsymbol{W}=\{(a): a \in A\}$.

We next define the *concatenation* of billows. Let $\boldsymbol{B}=(\boldsymbol{W}, \boldsymbol{U})$ be a billow in Γ, and let $\boldsymbol{C}=(\boldsymbol{Z}, \boldsymbol{Y})$ be a billow in $\Gamma/\boldsymbol{B}$. Then $\boldsymbol{B} * \boldsymbol{C}$ is the pair $(\boldsymbol{R}, \boldsymbol{S})$, where $\boldsymbol{R}=\{P * Q: P \in \boldsymbol{U}, Q \in \boldsymbol{Z}$ and $\mathrm{in}(Q)=\mathrm{ter}(P)\}$ and $\boldsymbol{S}=\{R \in \boldsymbol{R}: \mathrm{ter}(R) \in \mathrm{ter}[\boldsymbol{Y}]\}$.

Lemma 4.3. *$\boldsymbol{B} * \boldsymbol{C}$ is a billow.*

We can now define inductively the concatentation $*(\boldsymbol{B}_\alpha: \alpha<\theta)$ of a sequence $\boldsymbol{B}_\alpha=(\boldsymbol{W}_\alpha, \boldsymbol{U}_\alpha)$, $(\alpha<\theta)$ where each $\boldsymbol{B}_\alpha$ is a billow in $\Gamma/*(\boldsymbol{B}_\beta: \beta<\alpha)$. (So, we are already assuming that $*(\boldsymbol{B}_\beta: \beta<\alpha)$ is defined for all $\alpha<\theta$.) If $\theta=\gamma+1$ then $*(\boldsymbol{B}_\alpha: \alpha<\theta)$ is defined as $(*(\boldsymbol{B}_\alpha: \alpha<\theta)) * \boldsymbol{B}_\gamma$. If θ is a limit ordinal, then $*(\boldsymbol{B}_\alpha: \alpha<\theta)=\uparrow_{\beta<\theta}(*(\boldsymbol{B}_\alpha: \alpha<\beta))$. By Lemmas 4.2 and 4.3, $*(\boldsymbol{B}_\alpha: \alpha<\theta)$ is a billow. If $*(\boldsymbol{B}_\alpha: \alpha<\theta)$ is defined, we say that the sequence $(\boldsymbol{B}_\alpha: \alpha<\theta)$ is a *ladder*.

Some definitions needed at this point are the following: For any vertex v we write $e^-(v)$ for the set of edges going into v. If F is a set of vertices, we write $e^-[F]=\bigcup\{e^-(v): v \in F\}$.

If $F \subseteq V \setminus A$ then $\Gamma+(s)F$ denotes the web $(G, A \sqcup F, B)$, $\Gamma+(d)F$ denotes the web $(G, A, B \sqcup F)$, and $\Gamma-(d)F$ denotes the web (G', A, B) where $G'=(V, E \setminus e^-[F])$ (here (s) stands for 'source' and (d) for 'destination'. In $\Gamma-(d)F$ the vertices of F cannot be reached.) We also write $\Gamma \bullet F$ for $\Gamma-(d)F+(s)F$.

In our new terminology, a 1-*obstruction* is a pair $(\boldsymbol{B}, \phi)$, where $\boldsymbol{B}=(\boldsymbol{W}, \boldsymbol{U})$ is a billow, $\boldsymbol{W}$ is tight and $\boldsymbol{U} \subsetneq \boldsymbol{W}$. Assume now that $\kappa \in Y$ and that the notion of μ-obstruction is defined for all $\mu \in Y \cap \kappa$. A pair $(\boldsymbol{B}, F)$ is said to be a κ-*obstruction* if there exists a sequence $((\boldsymbol{B}_\alpha, F_\alpha): a<\kappa)$ of pairs, such that $\boldsymbol{B}_\alpha$ is a billow in $\Gamma_\alpha=(\Gamma \bullet \hat{F}_\alpha)/\hat{\boldsymbol{B}}_\alpha$ (where $\hat{F}_\alpha=\bigsqcup\{F_\beta: \beta<\alpha\}$ and $\hat{\boldsymbol{B}}_\alpha=*(\boldsymbol{B}_\beta: \beta<\alpha)$); $\boldsymbol{B}=*(\boldsymbol{B}_\alpha: \alpha<\kappa)$, $F=\bigsqcup\{F_\alpha: \alpha<\kappa\}$, and for each $\alpha<\kappa$ either:

(i) $(\boldsymbol{B}_\alpha, F_\alpha)$ is a μ-obstruction in Γ_α for some $\mu<\kappa$, or,

(ii) $\boldsymbol{B}_\alpha=\square_{\Gamma_\alpha}$, $|F_\alpha|=1$, $F_\alpha=\{x\}$ for some $x \in V(\Gamma_\alpha)-S_{\Gamma_\alpha}$,

and the set $\{\alpha<\kappa$: (i) occurs at $\alpha\}$ is stationary. We say that Γ is *obstructed* (or specifically κ-*obstructed*), if it contains a κ-obstruction for some $\kappa \in Y$.

Theorem 4.4. *A web containing no unending path is linkable if and only if it is unobstructed.*

Conjecture 4.5. The same is true for all webs.

For countable graphs the following weaker version of the conjecture is known.

Theorem 4.6. *A countable unhindered web is linkable.*

The main step in the proof of Theorem 4.6 is the following.

Theorem 4.7. *If Γ is countable and unhindered, then for every $a \in A$ there exists an a–B path P such that $\Gamma - V(P)$ is unhindered.*

This theorem should be true also for uncountable webs, but countability is so far used very essentially in the proof. Its proof in the uncountable case would probably clear the way to the proof of Conjecture 4.1. On the other hand, if the term 'unhindered' could be replaced in the theorem by 'not 1-obstructed' then Conjecture 4.5 would follow for countable webs.

Proof of Conjecture 4.1 for countable webs. Let $\boldsymbol{B} = (\boldsymbol{W}, \boldsymbol{U})$ be a $\leqslant$-maximal billow in Γ. By Lemma 4.3, $\Gamma/\boldsymbol{B}$ contains no nontrivial billow, and hence, by Theorem 4.6, it is linkable. Let $\boldsymbol{L}$ be a linkage in $\Gamma/\boldsymbol{B}$. Then the ripple $\boldsymbol{U} * \boldsymbol{L}$ and the A–B separating set $\mathrm{ter}[\boldsymbol{U}]$ satisfy the conditions of Conjecture 4.1. □

5. General graphs

A matching F in a graph $G = (V, E)$ is called *perfect* if $F[V] = V$. If G contains a perfect matching it is called *matchable*. In 1947 Tutte proved his celebrated theorem on a necessary and sufficient condition for matchability [27]. We shall bring it here in a stronger form, proved by Gallai [15].

A graph P is called *factor-critical* if it is unmatchable, but $P - x$ is matchable for every $x \in V(P)$. Clearly, a finite factor critical graph is of odd cardinality (for which reason it is called in [6] "peculiar"). Given a graph $G = (V, E)$ and a subset S of V, we write $\boldsymbol{P}(S) = \boldsymbol{P}(G, S)$ for the set of factor-critical connected components in $G - S$. We associate with S a bipartite graph $\Pi = \Pi(G, S) = (\boldsymbol{P}(G, S), S, H)$, where $(P, s) \in H$ (here $P \in \boldsymbol{P}(S)$ and $s \in S$) if $(v, s) \in E$ for some $v \in V(P)$. We can now state Gallai's theorem in a form which is true also for infinite graphs.

Theorem 5.1 [6]. *A graph G is matchable if and only if $\Pi(G, S)$ is espousable for every set of vertices S.*

If G is finite and unmatchable, then, by Theorem 5.1 and Hall's theorem, there exists a subset $\boldsymbol{P}'$ of $\boldsymbol{P}(S)$ connected in $\Pi(G, S)$ to a subset S' of S containing

fewer than $|\boldsymbol{P}'|$ vertices. Then $\boldsymbol{P}' \subseteq \boldsymbol{P}(S')$, and hence $|\boldsymbol{P}(S')| > |S'|$. This, in particular, implies Tutte's theorem, which states that if G is unmatchable, then it contains a set S' of vertices such that the number of odd components in $G - S'$ is larger than $|S'|$. It is also worth while noting that a locally finite factor-critical graph P is finite (this is proved by the use of König's Lemma and the fact that if I is a matching of $P - x$ for some vertex x, then there is no infinite I-alternating path.) This proves the result first proved by Tutte himself [28], that his theorem holds also for locally finite graphs.

We shall not attempt any description of the proof of this theorem. Instead, we give a structure theorem for graphs, which follows easily from it. (In the finite case this is the Gallai-Edmonds decomposition theorem.)

Theorem 5.2 [11]. *The vertex set V of any graph $G = (V, E)$ can be decomposed as a disjoint union $V = A \sqcup B \sqcup T$, where:*

(i) *$G[B]$ is matchable;*
(ii) *$A = \bigsqcup\{V[P]: P \in \boldsymbol{P}\}$ where $\boldsymbol{P} = \boldsymbol{P}(G, T)$; and*
(iii) *T is matchable into $\boldsymbol{P}$ in $\Pi(G, T)$.*

Proof. Let $\boldsymbol{Z}$ be the set of pairs (Y, I), where $Y \subseteq V$ and I is a matching of S into $\boldsymbol{P}(G, Y)$ in $\Pi(G, Y)$. Define an order $<$ on $\boldsymbol{Z}$ by $(Y, I) < (Z, H)$ if (1) $Y \subseteq Z \subseteq V \backslash V[\boldsymbol{P}(G, Y)]$ and (2) $I \subseteq H$ (note that (1) implies that $\boldsymbol{P}(G, Y) \subseteq \boldsymbol{P}(G, Z)$). Applying Zorn's Lemma, it is not hard to see that $\boldsymbol{Z}$ contains a maximal element (T, I). Let $A = V[\boldsymbol{P}(G, T)]$ and let $B = V \backslash (A \sqcup T)$. If $G' = G[B]$ were not matchable, then, by Theorem 5.1 there exists a subset S of B such that $\Pi = \Pi(G', S)$ is inespousable. By Theorem 3.2 there exists $X \subseteq S$ which has a matching J into $D_\Pi(X)$, and $D_\Pi(X)$ is unmatchable. Since $D_\Pi(X)$ is unmatchable and $D_\Pi(\phi) = \phi$, it follows that $X \neq \phi$. Then $T \sqcup X$ is matchable into $\boldsymbol{P}(G, T \sqcup X) = \boldsymbol{P}(G, T) \sqcup D_\Pi(X)$ by the matching $I \sqcup J$, contradicting the maximality of the pair (T, I). We have thus shown that $G[B]$ is matchable, as required in the theorem. □

In [11] Theorem 5.2 was used to prove a duality result on fractional covers and matchings in graphs (see Section 6). Here we shall use it to prove another result. A matching F is called strongly *maximal* if there does not exist a matching H such that $|H \backslash F| > |F \backslash H|$. Since for any matching H the set of edges in $F \Delta H$ forms a set of disjoint alternating paths and cycles, it is clear that F is strongly maximal if and only if there does not exist a finite improving F-alternating path, i.e. one starting and ending at $V \backslash F[V]$.

Theorem 5.3. *Every graph contains a strongly maximal matching.*

Proof. Let A, B, T and $\boldsymbol{P}$ be as in Theorem 5.2, let J be a perfect matching in $G[B]$, and let I be a matching of T into $\boldsymbol{P}$ in $\Pi(G, T)$. For each $t \in T$ choose a vertex $f(t) \in E\langle t \rangle \cap V(I(t))$, and a matching K_t of $I(t) - f(t)$. For each $P \in \boldsymbol{P}' =$

$\mathbf{P}\backslash I[T]$ choose a vertex $x = x(P) \in V(P)$, and a matching L_P of $P - x$. Let

$$F = J \sqcup \{(t, f(t)): t \in T\} \sqcup \bigcup \{K_t: t \in T\} \sqcup \bigsqcup \{L_P: P \in \mathbf{P}'\}.$$

Then the set $V\backslash F[V]$ of vertices not covered by F is $\{x(P): P \in \mathbf{P}'\}$. Suppose that $(y_1, y_2, \ldots, y_n)$ is an improving F-alternating path. Then $y_1 = x(P_1)$ for some $P_1 \in \mathbf{P}$. Since $E[V(P_1)] \subseteq V(P_1) \sqcup T$, there exists k_1 such that $y_i \in V(P_1)$ for $i < k_1$ and $y_{k_1} \in T$. Then $y_{k+1} = f(y_k) \in V(P_2)$, where $P_2 \in \mathbf{P}\backslash\mathbf{P}'$. Then there exists $k_2 > k_1 + 1$ such that $y_i \in V(P_2)$ for $k_1 < i < k_2$ and $y_{k_2} \in T$. Continuing this way we see that for no $i > 1$ does there hold $y_i \in V[\mathbf{P}']$. Hence y_n cannot belong to $V[\mathbf{P}']$, a contradiction. □

Problem 5.4. For which finite graphs H is it true that every infinite graph G contains a strongly maximal set of edges not spanning H? (In Theorem 5.3 H is a path of length 2.)

Problem 5.5. Does every hypergraph with finite edges contain a strongly maximal matching?

6. Hypergraphs and linear programming duality

So far there have been two (very closely related) main themes in our survey: conditions for matchability, and the duality between matchings and covers; all, of course, in graphs. When we turn to hypergraphs, we see that very little has been done on these two themes, even in the finite case. As is well known, one cannot expect a 'good characterization' of matchability in hypergraphs (which means that it is probably not a CO-NP problem). Yet it is quite possible that the theorems of König, Hall and Tutte can be subsumed in a more comprehensive theory on matchings in hypergraphs. Better understood in the finite case are fractional matchings and covers, and perhaps this is the first direction which should be investigated in the infinite case.

First let us define the above concepts. A *hypergraph* will mean here a pair $H = (V, E)$, where E is a set of subsets of V, called 'edges'. A *matching* in H is a set of disjoint edges, and a *cover* is a set of vertices (elements of V) which meets all edges. The subjects of investigation in matching theory are maximal matchings and minimal covers (it is easy to find large covers and small matchings—for example, the empty set of edges is a matching (see Fig. 1)).

Fig. 1. The empty matching.

In the finite case 'maximal' is with respect to cardinality. In the infinite case stronger notions are needed, which are given by the duality between matchings and covers. In this section we give another direction from which this duality can be viewed.

A *fractional matching* is a function $f: E \to \mathbb{R}^+$ such that $\sum \{f(e): v \in e\} \leqslant 1$ for every $v \in V$. A *fractional cover* is a function $g: V \to \mathbb{R}^+$, such that $\sum \{g(v): v \in e\} \geqslant 1$ for every $e \in E$. For any function $h: X \to \mathbb{R}^+$ write $|h| = \sum \{h(x): x \in X\}$ ($|h|$ may be an infinite cardinality). If f and g are a fractional matching and a fractional cover, respectively, then, by the definition of these objects,

$$|g| \geqslant \sum_{v \in V} g(v) \sum_{e \ni v} f(e) = \sum_{e \in E} f(e) \sum_{v \in e} g(v) \geqslant |f|. \tag{1}$$

Thus, if we write

$$\nu^* = \nu^*(H) = \sup\{|f|: f \text{ is a fractional matching in } H\}$$

and

$$\tau^* = \tau^*(H) = \inf\{|g|: g \text{ is a fractional cover}\},$$

there holds $\nu^* \leqslant \tau^*$. The duality theorem of linear programming implies the following.

Theorem 6.1. *In a finite hypergraph* $\nu^* = \tau^*$.

Of course, in a finite hypergraph ν^* and τ^* are attained, that is, there exist a fractional matching f and a fractional cover g such that $|f| = |g| = \nu^* = \tau^*$. By (1) it follows that:

(a) $\sum \{f(e): v \in e\} = 1$ whenever $g(v) > 0$; and

(b) $\sum \{g(v): v \in e\} = 1$ whenever $f(e) > 0$.

These are the so-called 'complementary slackness conditions' of linear programming. In the infinite case (a) and (b) are strictly stronger than the condition $\nu^* = \tau^*$. We say that a hypergraph H has the *strong duality* property if it has a fractional matching f and a fractional cover g satisfying (a) and (b), and that H has the *weak duality* property if $\nu^*(H) = \tau^*(H)$. We say that H has König's Property (or *integral strong duality*) if it has a matching F and a cover G such that G consists of the choice of precisely one vertex from each edge in F. Clearly, this means that H satisfies the strong duality property, with f and g being (0, 1) functions.

The question now arises—which hypergraphs satisfy the weak, strong and integral strong duality properties? As usual, the situation is much more pleasant in the case of graphs.

Theorem 6.2 [11]. *Every graph has the strong duality property, with f and g taking* $0, \frac{1}{2}, 1$ *values.*

The proof uses Theorem 5.2.

A hypergraph $H = (V, E)$ is said to be of *finite character* if each of its edges is finite. It is said to be *locally finite* if every vertex belongs to finitely many edges. Easy compactness arguments yield the following.

Theorem 6.3. (i) *If H is of finite character, then there exists a fractional cover g with $|g| = \tau^*$, and $\nu^* = \tau^*$.*

(ii) *If H is of finite character and locally finite, then it has the strong duality property.*

Here is an example (taken from [7]) of a hypergraph of finite character, which does not have the strong duality property.

Let $V = \{a_i: 0 \leqslant i < \omega\} \sqcup \{b_i: 0 \leqslant i < \omega\}$, and for each $1 \leqslant k < \omega$ let $e_k = \{a_0, a_1, \ldots, a_k, b_k\}$ and $d_k = \{b_0, b_1, \ldots, b_k, a_k\}$. Let $E = \{e_k: 1 \leqslant k < \omega\} \sqcup \{d_k: 1 \leqslant k < \omega\}$, and let $H = (V, E)$.

Assertion 6.4. *Every fractional matching f in H satisfies $|f| < 2$.*

Proof. Let $m = \min\{k: f(e_k) > 0 \text{ or } f(d_k) > 0\}$, say $\alpha = f(e_m) > 0$. Then each d_i for which $f(d_i) > 0$ contains b_m, and hence $\sum f(d_i) \leqslant 1 - \alpha$. Since each e_i contains a_1, we have $\sum f(e_i) \leqslant 1$. Thus $|f| \leqslant 2 - \alpha$. □

Assertion 6.5. $\nu^* = 2$.

Proof. For each $1 \leqslant n < \omega$ let $f = f_n$ be defined by: $f(e_k) = f(d_k) = (\frac{1}{2})^{n-k}$ for $k < n$, $f(e_k) = f(d_k) = 0$ for $k \geqslant n$. Then f_n is a fractional matching with $|f_n| = 2 - (\frac{1}{2})^{n-1}$. The assertion now follows from Assertion 6.4. □

By Theorem 6.3(i) $\tau^*(H) = \nu^*(H) = 2$. Since, by Assertion 6.4, ν^* is not attained, we see by (1) that H does not have the strong duality property.

This example can be adapted to yield also a locally finite hypergraph without the strong duality property (and in which ν^* is not attained). However, the following problem is as yet unsettled.

Problem 6.6. Does the strong duality property hold for hypergraphs whose edges are of *bounded* size (i.e., $|e| \leqslant k$ for all $e \in E$ for some $k \in N$)? Or does it hold for hypergraphs with bounded vertex valencies?

In [7] this was settled in the affirmative for the case that ν^* is finite.

Let us conclude this section with a basic open problem concerning König's Property.

Problem 6.7. Suppose that every finite subhypergraph of H (i.e. $H' = (V, E')$, where E' is a finite subset of E) satisfies König's Property. Does it necessarily follow that H satisfies König's Property?

7. Posets

It was probably Fulkerson who first used the trick of replacing a partially ordered set (or, for that matter, any directed graph) by a bipartite graph by splitting its vertices. The splitting is similar to that in Section 4, the difference being that you do not add the edges (x', x''). Thus the bipartite graph attained from a poset P is $\Gamma = (V', V'', E)$ where $V = V(P)$ and $(x', y'') \in E$ if and only if $x > y$ in P. A matching F in Γ corresponds in a natural way to a decomposition $\boldsymbol{C}_F$ of P into paths: an edge (x, y) will appear in some path in $\boldsymbol{C}_F$ if and only if $(x', y'') \in F$. (Thus, if $\{x', x''\} \cap \bigcup F = \emptyset$ then (x) is a single vertex path in $\boldsymbol{C}_F$). Fulkerson used this to derive Dilworth's theorem from König's theorem [14]. Oellrich and Steffens used the same method to prove a 'strong duality' version of Dilworth's theorem for posets containing no infinite chains.

Theorem 7.1 [23]. *If a poset $\boldsymbol{P}$ does not contain infinte chains then there exist a decomposition $\boldsymbol{C}$ of $\boldsymbol{P}$ into disjoint chains (= paths) and an independent set A such that $A \cap V(Q) \neq \emptyset$ for every $Q \in \boldsymbol{C}$.*

In [23] this was proved for countable posets, using the countable version of Theorem 3.1, but their proof applies also to the general case. The elegant proof is worth repeating.

Proof. Let F and C be a matching and a cover in Γ, as in Theorem 3.1. Let $\boldsymbol{C}_F = (Q_i : i \in I)$ and let $A = \{x \in V(P) : \{x', x''\} \cap C = \emptyset\}$. Then A is independent since if, say, $x > y$ for $x, y \in A$ then $(x', y'') \in E(\Gamma)$, contradicting the fact that x', y'' do not belong to the cover C.

Let $Q \in \boldsymbol{C}$, and suppose that Q and k edges, hence $k+1$ vertices. By the properties of C and F, the set $\{x', x'' : x \in V(Q)\}$ contains k elements from C (one for each edge in $E(Q)$.) Hence there is a vertex $x \in V(Q)$ such that $\{x', x''\} \cap C = \emptyset$, proving that $A \cap V(Q) \neq \emptyset$. □

In [8] the splitting trick was used to prove an infinite analogue of Greene–Kleitman's theorem, which, in the finite case, is a generalization of Dilworth's theorem.

Theorem 7.2 [8]. *If a poset $\boldsymbol{P}$ does not contain infinite chains and k is a positive integer, then there exist a decomposition $\boldsymbol{C}$ of $\boldsymbol{P}$ into chains and disjoint antichains $A_1, \ldots, A_k$ such that each path $Q \in \boldsymbol{C}$ meets* $\min(|V(Q)|, k)$ *antichains A_i.*

In the finite case Theorem 7.2 remains true if we replace everywhere the word 'chain' by 'antichain' and vice versa (this is a theorem of Greene [16]). We conjecture that this is true for Theorem 7.2 also in the infinite case, but this seems to be much harder to prove than Theorem 7.2. Even the case $k = 1$ which is

trivial for finite posets (and, in fact, for all well-founded posets) is unknown, and we state it here as follows.

Conjecture 7.3. If a poset P contains no infinite antichain then there exist a chain C and a decomposition of P into disjoint antichains $(A_i: i \in I)$ such that $A_i \cap C \neq \emptyset$ for all i.

In [8] this was proved for posets containing no antichain of size larger than 2.

References

[1] R. Aharoni, On the equivalence of two conditions for the existence of transversals, J. Combin. Theory. Ser. A. 34 (1983) 202–214.

[2] R. Aharoni, Menger's theorem for graphs containing no infinite paths, European J. Combin. 4 (1983) 201–204.

[3] R. Aharoni, König's duality theorem for infinite bipartite graphs, J. London Math. Soc. 29 (1984) 1–12.

[4] R. Aharoni, A generalization of Tutte's 1-factor theorem to countable graphs, J. Combin. Theory. Ser. B 37 (1984) 199–209.

[5] R. Aharoni, Menger's theorem for countable graphs, J. Combin. Theory. Ser. B 43 (1987) 303–313.

[6] R. Aharoni, Matchings in infinite graphs, J. Combin. Theory. Ser. B 44 (1988) 87–125.

[7] R. Aharoni, and R. Holzman, Fractional matchings and covers in infinite hypergraphs, Graphs Combin., to appear.

[8] R. Aharoni and V. Korman, Greene–Kleitman's theorem for infinite posets, submitted.

[9] R. Aharoni, M. Magidor and R. Shore, On the strength of König's Theorem for infinite bipartite graphs, J. Combin. Theory. Ser. B., to appear.

[10] R. Aharoni, C.S.J.A. Nash-Williams, and S. Shelah, A general criterion for the existence of tranversals, Proc. London Math. Soc. 47 (1983) 43–68.

[11] R. Aharoni and R. Ziv, LP duality in infinite hypergraphs, J. Combin. Theory. Ser. B 50 (1990) 82–92.

[12] G. Birkhoff, Trés observaciones sobre el algebra lineal, Revista Fac. Ciencias Univ. Tucuman 5 (1946) 147–159.

[13] M.R. Damerell and E.C. Milner, Necessary and sufficient conditions for transversals of countable set systems, J. Combin. Theory. Ser. A 17 (1974) 350–374.

[14] L.R. Ford and D.R. Fulkerson, Flows in Networks (Princeton Univ. Press, Princeton, NJ, 1962).

[15] T. Gallai, On factorization of graphs, Acta Math. Hungar. 1 (1950) 134–153.

[16] C. Greene, Some partitions associated with a partially ordered set, J. Combin. Theory. Ser. A 20 (1976) 41–68.

[17] P. Hall, On representatives of subsets, J. London Math. Soc. 10 (1935) 26–30.

[18] M. Hall, Distinct representatives of subsets, Bull. Amer. Math. Soc. 54 (1948) 922–926.

[19] M. Holz, K.-P. Podewski and K. Steffens, Injective choice functions, Lecture Notes in Math., Vol. 1238 (Springer, Berlin, 1988).

[20] T. Jech, Set Theory (Academic Press, New York, 1978).

[21] D. König, Theorie der Endlichen und Unendlichen Graphen (Chelsea, New York, 1950).

[22] C.S.J.A. Nash-Williams, Another criterion for marriage in denumerable societies, in: B. Bollobás, ed., Advances in Graph Theory, Ann. Discrete Math. 3 (North-Holland, Amsterdam, 1978) 165–179.

[23] H. Oellrich and K. Steffens, On Dilworth's decomposition theorem, Discrete Math. 15 (1976) 301–304.

[24] K.-P. Podewski and K. Steffens, Injective choice functions for countable families, J. Combin. Theory. Ser. B 21 (1976) 40–46.
[25] K.-P. Podewski and K. Steffens, Uber Translationen und der Satz von Menger in unendlichen Graphen, Acta Math. Acad. Si Hungar. 30 (1977) 69–84.
[26] K. Steffens, Matchings in countable graphs, Canad. J. Math. 29 (1976) 165–168.
[27] W.T. Tutte, The factorization of linear graphs. J. London Math. Soc. 22 (1947) 107–111.
[28] W.T. Tutte, The factors of graphs, Canad. J. Math. 4 (1952) 314–328.

Discrete Mathematics 95 (1991) 23–47
North-Holland

Gallai–Milgram properties for infinite graphs

J.-M. Brochet and M. Pouzet*

Lab. Algèbre Ordinale, Groupe L. M. D. I., Université Claude Bernard (Lyon 1), 69622-Villeurbanne cedex, France

Received 28 July 1989
Revised 25 October 1990

Abstract

Brochet, J.-M. and M. Pouzet, Gallai–Milgram properties for infinite graphs, Discrete Mathematics 95 (1991) 23–47.

We discuss extensions of the Gallai–Milgram theorem to infinite graphs. We define a path to be a directed graph whose transitive closure is a linear ordering. We show that an undirected graph with no infinite independent set is covered by finitely many pairwise disjoint paths; moreover for a given integer k this graph is covered by $\leqslant k$ (resp. $\leqslant k$ pairwise disjoint) paths if each finite set of vertices is contained in the union of $\leqslant k$ (resp. $\leqslant k$ pairwise disjoint) paths. Hence an undirected graph with no independent set of size $k+1$ is covered by $\leqslant k$ pairwise disjoint paths. We prove also that if to each edge (x, y) of a countable path is associated an element $\theta(x, y)$ of a finite group, then some edges can be deleted so that the new graph is still a path and the product $\theta(x_0, x_1) \bullet \theta(x_1, x_2) \bullet \cdots \bullet \theta(x_{n-1}, x_n)$ of the elements of the group along any finite path $(x_0, x_1, \ldots, x_n)$ of the new graph depends only upon the extremities x_0, x_n of the path.

1. Introduction

If $\mathcal{G} = (V, E)$, $(E \subseteq V \times V)$, is a finite directed graph with no independent set of size $k+1$, $(k < \omega)$, the Gallai–Milgram theorem [6] says that V can be covered by $\leqslant k$ pairwise disjoint paths, where a path is a finite sequence of distinct vertices $(x_0, \ldots, x_n)$ such that $(x_i, x_{i+1}) \in E$ for $i < n$. If the directed graph is an ordered set $\mathcal{P} = (P, \leqslant)$ we get Dilworth's theorem [5]: If $\mathcal{P}$ has no antichain of size $k+1$, $(k < \omega)$, then P can be covered by $\leqslant k$ (pairwise disjoint) chains. Dilworth's theorem can be extended in this form to any infinite ordered set, but we need the finiteness assumption on k. The proof, using a standard compactness argument, follows from the fact that a set is a chain if and only if all its finite subsets are chains. Hence the transitivity of the ordering is involved. For the Gallai–Milgram theorem the situation is different; elementary compactness

* The second author has been partially supported by C.N.R.S., P.R.C. Math-Info.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

arguments do not seem to yield any useful information about infinite directed graphs for the case when k is finite. In fact, we still do not know if the theorem is true or false in this case. Of course, in order that the question should make some sense, it is first necessary to extend the notion of a path in a directed graph which allows for the possibility of paths of arbitrary (countable or uncountable) order type. We adopt the following fairly natural definition: A path is a directed graph whose transitive closure is a linear (i.e. total) ordering; and a path in a directed graph is simply a subgraph (not necessarily induced) which is a path. The question is then whether the following extension of the Gallai–Milgram theorem to infinite graphs holds:

> If $1 \leqslant k < \omega$ and $\mathscr{G} = (V, E)$ is a directed graph with no independent set of size $k+1$, then V is covered by $\leqslant k$ pairwise disjoint paths. $(*)$

A necessary condition for the conclusion of $(*)$ is that any finite subset of V is contained in $\leqslant k$ pairwise disjoint finite paths. To conjecture that this condition is also sufficient is stronger than $(*)$ since, by the Gallai–Milgram theorem, the hypothesis of $(*)$ is equivalent to the fact that any finite subset of V is the union of $\leqslant k$ pairwise disjoint paths. We show in Section 3 that this stronger conjecture, and also plausible weakenings of it, are false in general. In Section 4.2 we prove (cf. Theorem 2) that this conjecture holds for the undirected graphs with no infinite independent set, where an undirected graph is defined to be a directed graph $\mathscr{G} = (V, E)$ such that $(x, y) \in E$ if and only if $(y, x) \in E$ for all $x, y \in V$. We also show in Section 4.2 (cf. Theorem 1) that the vertex set of an undirected graph with no infinite independent set is covered by finitely many pairwise disjoint paths. Some similar properties concerning the undirected graphs with no independent set of size ν, where ν is an arbitrary, finite or infinite, cardinal, are surveyed in Section 4.3 (cf. Theorems 2′ and 2″). The notions of path and chain that we use are studied in Section 2. In Section 5, we deal with a different problem, still related to the spirit of the Gallai–Milgram theorem since it concerns the linkage of different finite sets of the same size along a countable path. We prove (cf. Theorem 4) that if $\theta : E \to H$ is a function of the edge set of a countable path $\mathscr{G} = (V, E)$ to a finite group $\mathscr{H} = (H, \bullet)$, then there is a subset of edges $E' \subseteq E$ such that the graph $\mathscr{G}' = (V, E')$ is still a path and the element $\tilde{\theta}(p) = \theta(x_0, x_1) \bullet \theta(x_1, x_2) \bullet \cdots \bullet \theta(x_{n-1}, x_n) \in H$ depends only upon x_0 and x_n if $p = (x_0, x_1, \ldots, x_n)$ is a finite path of the graph $\mathscr{G}'$. The interpretation in terms of linking finite sets of the same size n corresponds to the case where $\mathscr{H}$ is the group of the permutations of a set of size n.

2. Notions of path and chain in infinite graphs

2.1. A (*directed*) *graph* is an ordered pair of sets $\mathscr{G} = (V, E)$ where $E \subseteq V \times V$. We say that $\mathscr{G}$ is an *undirected graph* if $E = E^{-1}$, with $E^{-1} = \{(y, x) \in$

$V \times V: (x, y) \in E\}$. As usual V is the *vertex set* of the graph $\mathcal{G}$, and E is its *edge set*. An undirected graph can be identified in an obvious way with an ordered pair of sets $\mathcal{G} = (V, E)$ where $E \subseteq [V]^1 \cup [V]^2 = \{e \subseteq V: 1 \leqslant |e| \leqslant 2\}$. The *degree* of an element $x \in V$ in an *undirected* graph $\mathcal{G} = (V, E)$ is the cardinal $|\{y \in V: x \neq y, (x, y) \in E\}|$. The undirected graph $\mathcal{G} = (V, E)$ is *bipartite* with sides X and Y if $V = X \cup Y$, $X \cap Y = \emptyset$, $E \subseteq (X \times Y) \cup (Y \times X)$. The graph $\mathcal{G}' = (V', E')$ is a *subgraph* of the graph $\mathcal{G} = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. It is an *induced subgraph* of $\mathcal{G}$ if $V' \subseteq V$ and $E' = E \cap (V' \times V')$. Hence an induced subgraph of a graph $\mathcal{G} = (V, E)$ is entirely determined by its vertex set $V' \subseteq V$ and $\mathcal{G} \restriction V' = (V', E \cap (V' \times V'))$ will denote the induced subgraph of $\mathcal{G}$ with vertex set V'. The size of a graph is the cardinality of its vertex set; in particular $\mathcal{G} = (V, E)$ is finite (respectively countable) if V is. A subset $X \subseteq V$ is an *independent set* (respectively a *clique*) of a graph $\mathcal{G} = (V, E)$ if $(x, y) \notin E$ (respectively $(x, y) \in E \cup E^{-1}$) for all distinct $x, y \in X$.

2.2. The graph $\mathcal{G} = (V, E)$ is *reflexive* if $(x, x) \in E$ for every $x \in V$, *transitive* if $(x, y) \in E$ and $(y, z) \in E$ imply $(x, z) \in E$ for every $x, y, z \in V$. $\mathcal{G}$ is a *preordering* if it is reflexive and transitive; it is an *ordering* if moreover at most one of the two relations $(x, y) \in E$, $(y, x) \in E$, holds for all distinct $x, y \in V$. If at least one of these two relations holds for all distinct $x, y \in V$ the graph $\mathcal{G}$ is *complete*; that means that V is a clique of $\mathcal{G}$. If exactly one of these two relations holds for all distinct $x, y \in V$, $\mathcal{G}$ is a *tournament*. A *linear ordering* is an ordering which is also a tournament. If the graph is an ordering, we use the terms *chain* and *antichain* for clique and independent set. An ordering is *well founded* if each non-empty subset of the vertex set contains a minimal element. A *well ordering* is a linear ordering which is well founded. A linear ordering is *scattered* if it does not contain an isomorphic copy of the natural ordering on the rational numbers. Finally $R \subseteq V \times V$ is an *equivalence relation* on V if the graph (V, R) is reflexive, undirected, and transitive; an *R-equivalence class* is a subset C of V such that $C = \{y \in V: (x, y) \in R\}$ for some $x \in V$. The *transitive closure* of a graph $\mathcal{G} = (V, E)$ is the graph $\mathcal{G}_t = (V, E_t)$, where E_t is the smallest set of edges containing E such that $\mathcal{G}_t$ is a preordering. For $x, y \in V$, we have $(x, y) \in E_t$ if there are an integer $n \geqslant 0$ and a sequence $(x_0, \ldots, x_n)$ of vertices such that $x_0 = x$, $x_n = y$ and $(x_i, x_{i+1}) \in E$ if $i < n$. A graph $\mathcal{G} = (V, E)$ is *acyclic* if its transitive closure is an ordering, which means there is no *cycle*, i.e. no sequence $(x_0, \ldots, x_n)$ of distinct vertices, where n is an integer $\geqslant 1$, such that $(x_n, x_0) \in E$ and $(x_i, x_{i+1}) \in E$ if $i < n$.

2.3. A graph is a *path* if its transitive closure is a linear ordering. The *order type* of a path is defined as the order type of its transitive closure and we say that a path is *well ordered*, (respectively *scattered*), if its transitive closure is. An *interval* of a path $\mathcal{P} = (P, E)$ is an interval of its transitive closure $(P, \leqslant)$, i.e. a subset I of P such that $I = \{z \in P: x \leqslant z \leqslant y$ for some $x, y \in I\}$. The *first vertex* (respectively the *last vertex*) of a path is the smallest (respectively the largest)

element of its transitive closure (when it exists) and the *extremities* of a path are its first and its last vertex (when they exist). $X \subseteq P$ is *cofinal*, (respectively *coinitial*), in the path $\mathcal{P} = (P, E)$ if X is cofinal, (respectively coinitial), in its transitive closure $\mathcal{P} = (P, \leqslant)$, i.e. if there is some $y \in X$ such that $x \leqslant y$, (respectively $y \leqslant x$), for any $x \in P$. A *path of a graph* $\mathcal{G}$ is a subgraph of $\mathcal{G}$ which is a path. We shall say that $W \subseteq V$ is contained in, (respectively $W \subseteq V$ is the union of), some paths of the graph $\mathcal{G} = (V, E)$ if W is contained in, (respectively W is), the union of their vertex sets. The intersection of two paths will in fact denote the intersection of their vertex sets. In particular, we shall say that two paths are disjoint if their vertex sets are. A *chain* of a graph $\mathcal{G} = (V, E)$ is a subset C of V such that any finite subset F of C is contained in a (finite) path of $\mathcal{G}$.

2.4. Examples; basic observations. In Figs. 1, 2, 3, we give examples of paths which are respectively well ordered, scattered, nonscattered, (and which are not linear orderings); (note that two vertices x, y are joined by an arrow $x \to y$ to indicate that $(x, y) \in E$). ω, ω^*, η denote respectively the order types of the natural orderings on the positive integers, on the negative integers, and on the

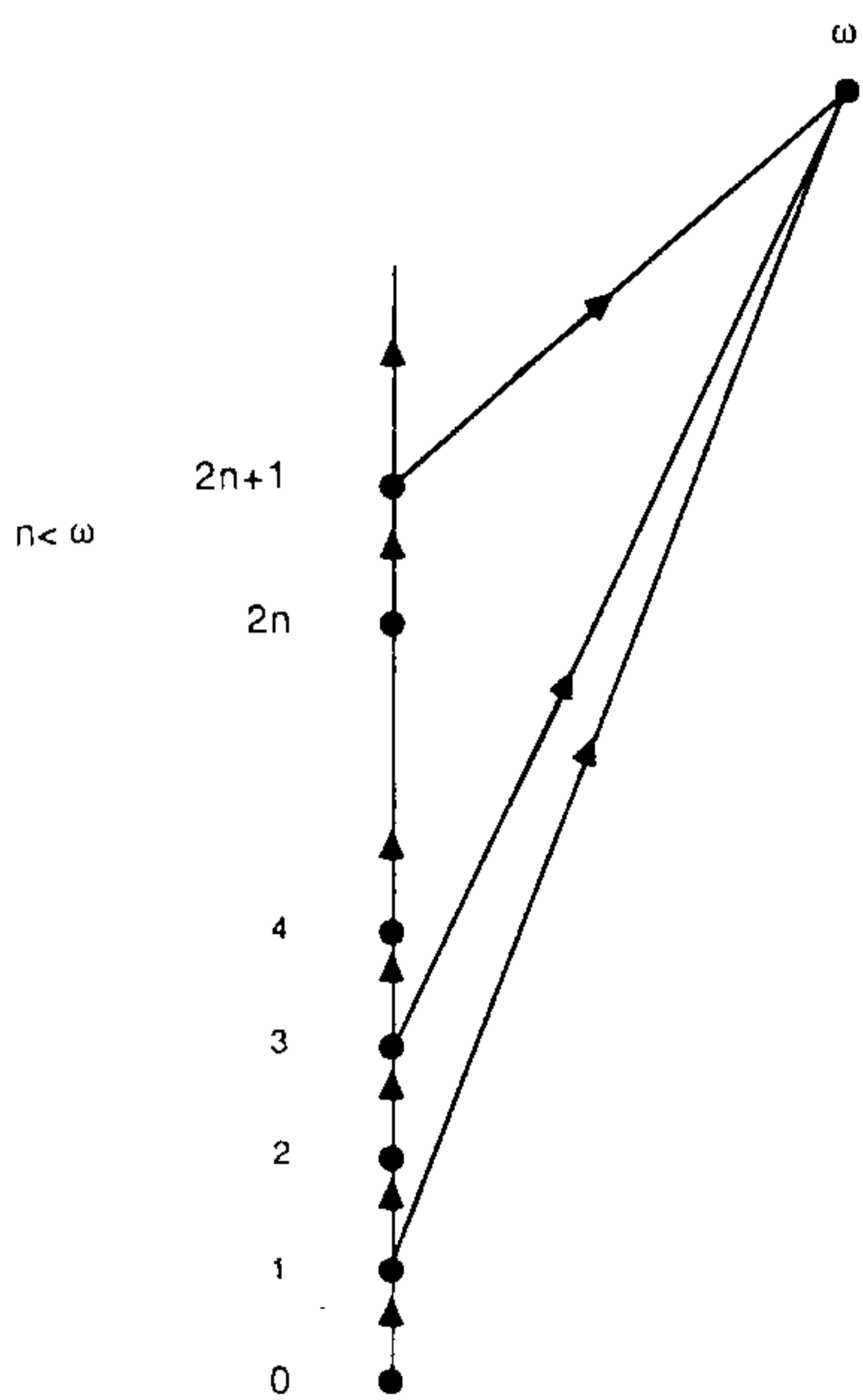


Fig. 1. A path of order type $\omega + 1$.

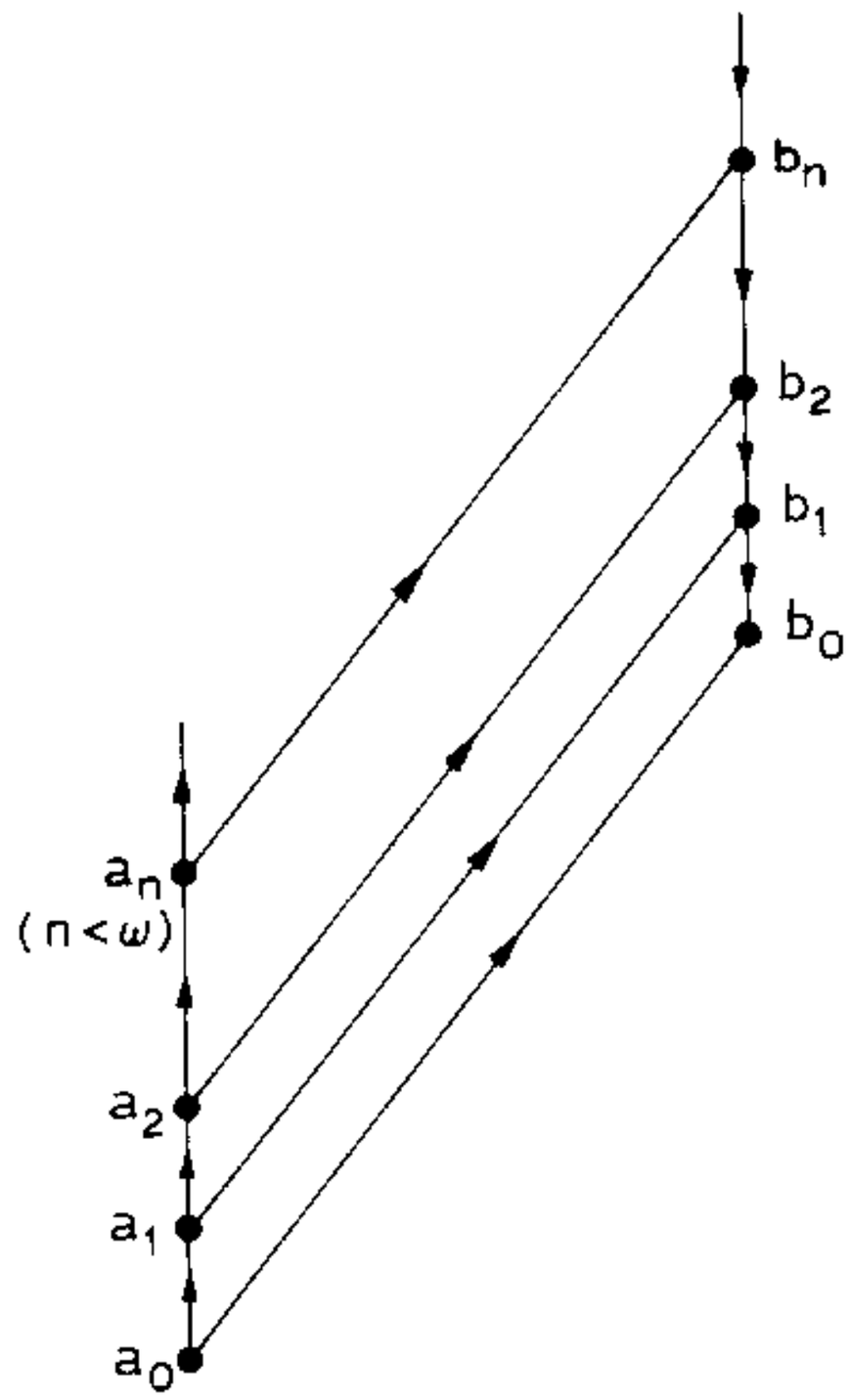


Fig. 2. A path of order type $\omega + \omega^*$.

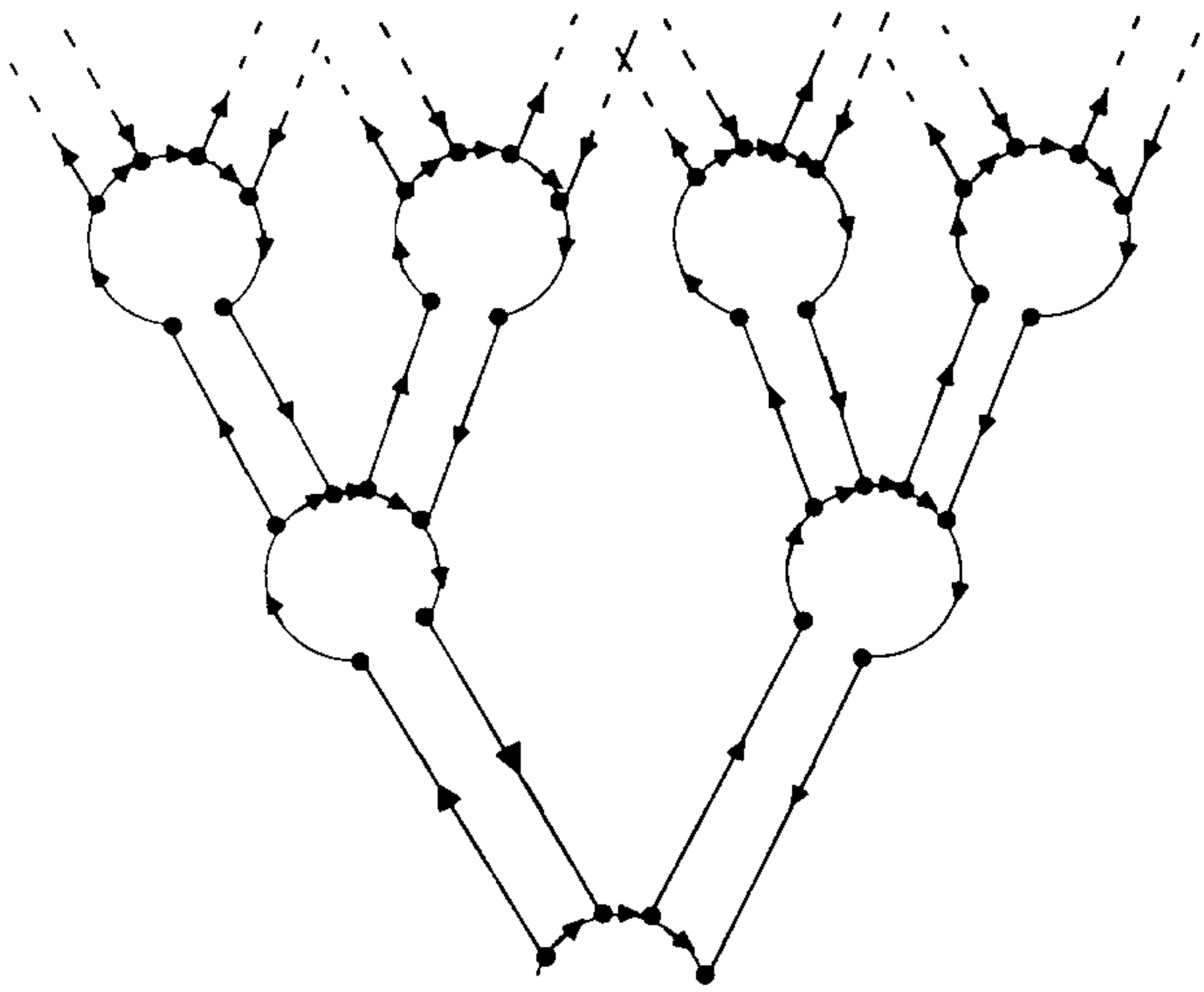


Fig. 3. A path of order type $\omega + (\omega^* + \omega) \cdot \eta + \omega^*$.

rational numbers. If $\mathscr{P}_0 = (P_0, \leqslant_0)$ and $\mathscr{P}_1 = (P_1, \leqslant_1)$ are linear orderings of order types α_0 and α_1, then $\alpha_0 + \alpha_1$ and $\alpha_0 \cdot \alpha_1$ denote respectively the order types of the linear orderings $\mathscr{P} = (P, \leqslant)$ and $\mathscr{P}' = (P', \leqslant')$ defined in the following way: $P = (P_0 \times \{0\}) \cup (P_1 \times \{1\})$ and $(x, i) \leqslant (y, j)$ if and only if $i < j$ or $i = j$ and $x \leqslant_i y$ in $\mathscr{P}_i$; $P' = P_0 \times P_1$ and $(x_0, x_1) \leqslant' (y_0, y_1)$ if and only if $x_1 <_1 y_1$, or $x_1 = y_1$ and $x_0 \leqslant_0 y_0$.

A well ordered path of the graph $\mathscr{G} = (V, E)$ can be identified with a sequence $(x_\beta)_{\beta<\alpha}$ of distinct vertices, where α is an ordinal $\geqslant 1$, such that $\{\gamma < \beta: (x_\gamma, x_\beta) \in E\}$ is cofinal in β whenever $\beta < \alpha$. In particular a finite path of $\mathscr{G}$ can be identified with a sequence $(x_0, \ldots, x_n)$ of distinct vertices, where n is an integer $\geqslant 0$, such that $(x_i, x_{i+1}) \in E$ if $i < n$.

The union of an up-directed family of chains is a chain, (a set $\mathscr{F}$ is *up-directed* if for every A, $B \in \mathscr{F}$ there is $C \in \mathscr{F}$ such that $A \cup B \subseteq C$). Therefore, by Zorn's lemma, each chain is contained in a maximal one. We shall see in Example 1 of Section 3 that this does not generally hold for vertex sets of paths. A subset of a chain is a chain and if each finite subset of a set C is a chain, then C is a chain.

By a standard compactness argument, *the following statements are equivalent*:

(i) *C is a chain of $\mathscr{G}$.*

(ii) *There is a linear ordering $\mathscr{C}$ on C, such that for every finite subset F of C there is a finite path $\mathscr{P}_F$ of $\mathscr{G}$ whose vertex sets contains F and is such that the ordering induced by $\mathscr{P}_F$ on F coincides with the induced ordering $\mathscr{C} \restriction F$.*

Indeed, let us assume that C is a chain of $\mathscr{G}$. For each finite subset F of C, let $\mathscr{U}_F = \{R \subseteq C \times C: (C, R) \restriction F$ is a linear ordering which is induced on F by the linear ordering defined by some finite path of $\mathscr{G}$ containing $F\}$. The $\mathscr{U}_F$'s, (F finite subset of C), are closed non-empty subsets of the compact space $2^{C \times C}$. They finitely intersect since $\mathscr{U}_{F_1} \supseteq \mathscr{U}_{F_2}$ for $F_1 \subseteq F_2$. Hence there is $R \in \bigcap\{\mathscr{U}_F: F$ finite subset of $C\}$, and $\mathscr{C} = (C, R)$ is a linear ordering which has the desired properties.

The notion of chain in an acyclic graph reduces to the classical notion of chain, (i.e. clique), of the transitive closure. For a transitive graph $\mathscr{G}$, the notions of chain and vertex set of a path both reduce to the notion of clique, (therefore there is no ambiguity for orderings in the notions of chain). But in general a clique of the transitive closure of a graph $\mathscr{G}$ is not a chain of $\mathscr{G}$, as shown by Fig. 4. The graph $\mathscr{G} = (V, E)$ (Fig. 4) is a finite undirected graph whose transitive closure is complete, but V is not a chain of $\mathscr{G}$.

The vertex set of a path is a chain. A finite chain is contained in a finite path, hence a maximal chain which is finite is a maximal vertex set of a path. Fig. 5 shows that a maximal chain of an acyclic graph is not necessarily (contained in) the vertex set of a path. The graph $\mathscr{G}$ (Fig. 5) is a countable acyclic graph with no independent set of size 3. $\{x_n: n < \omega\} \cup \{z\}$ is a maximal chain of $\mathscr{G}$ which is not the vertex set of a path.

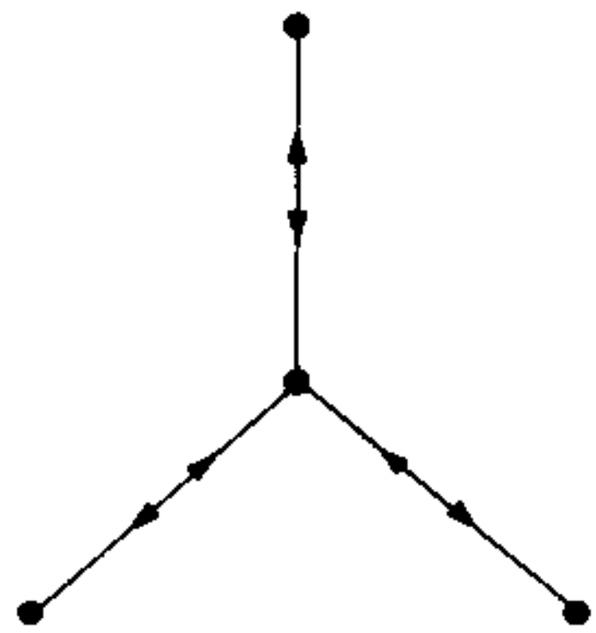


Fig. 4.

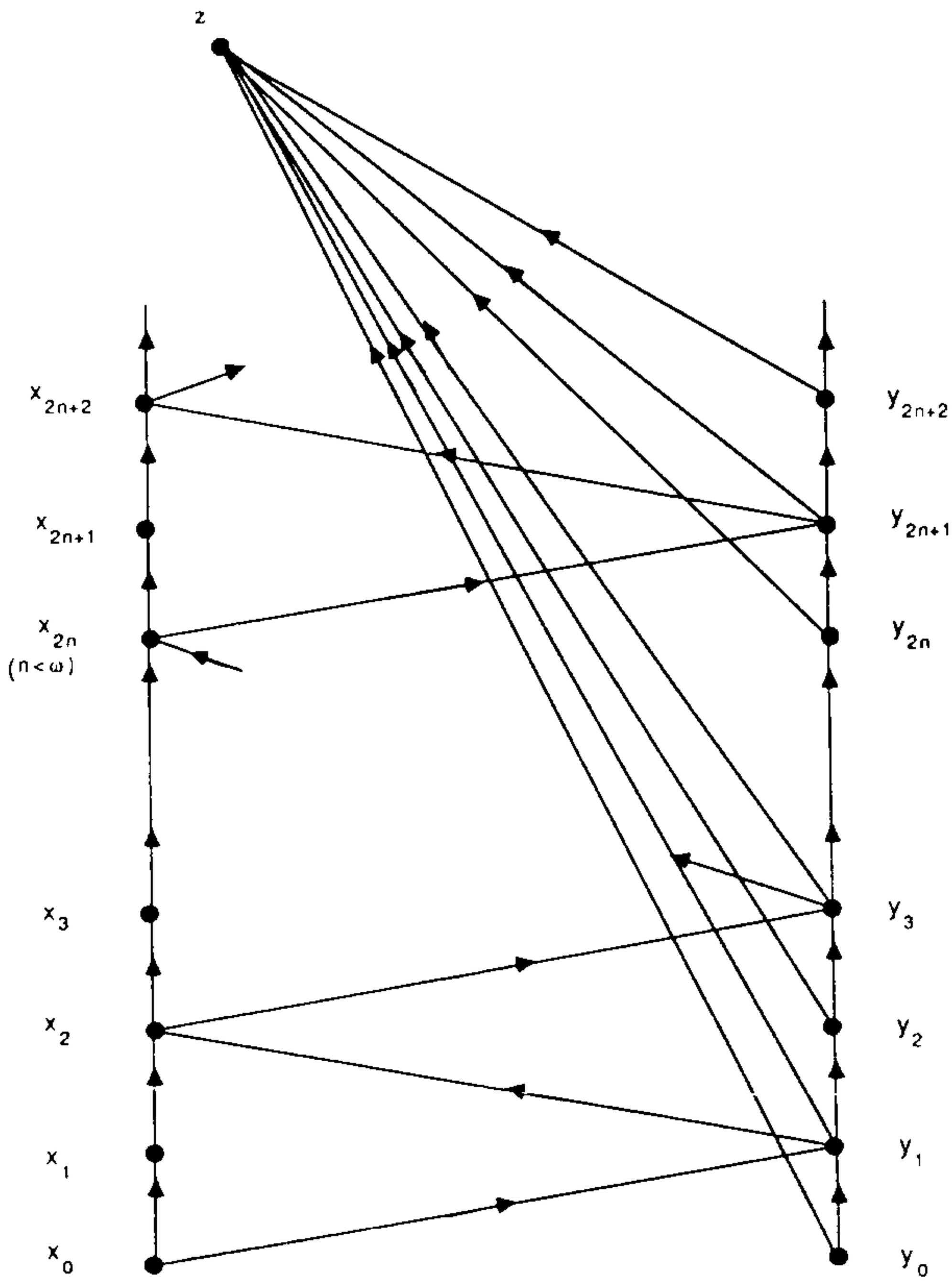


Fig. 5.

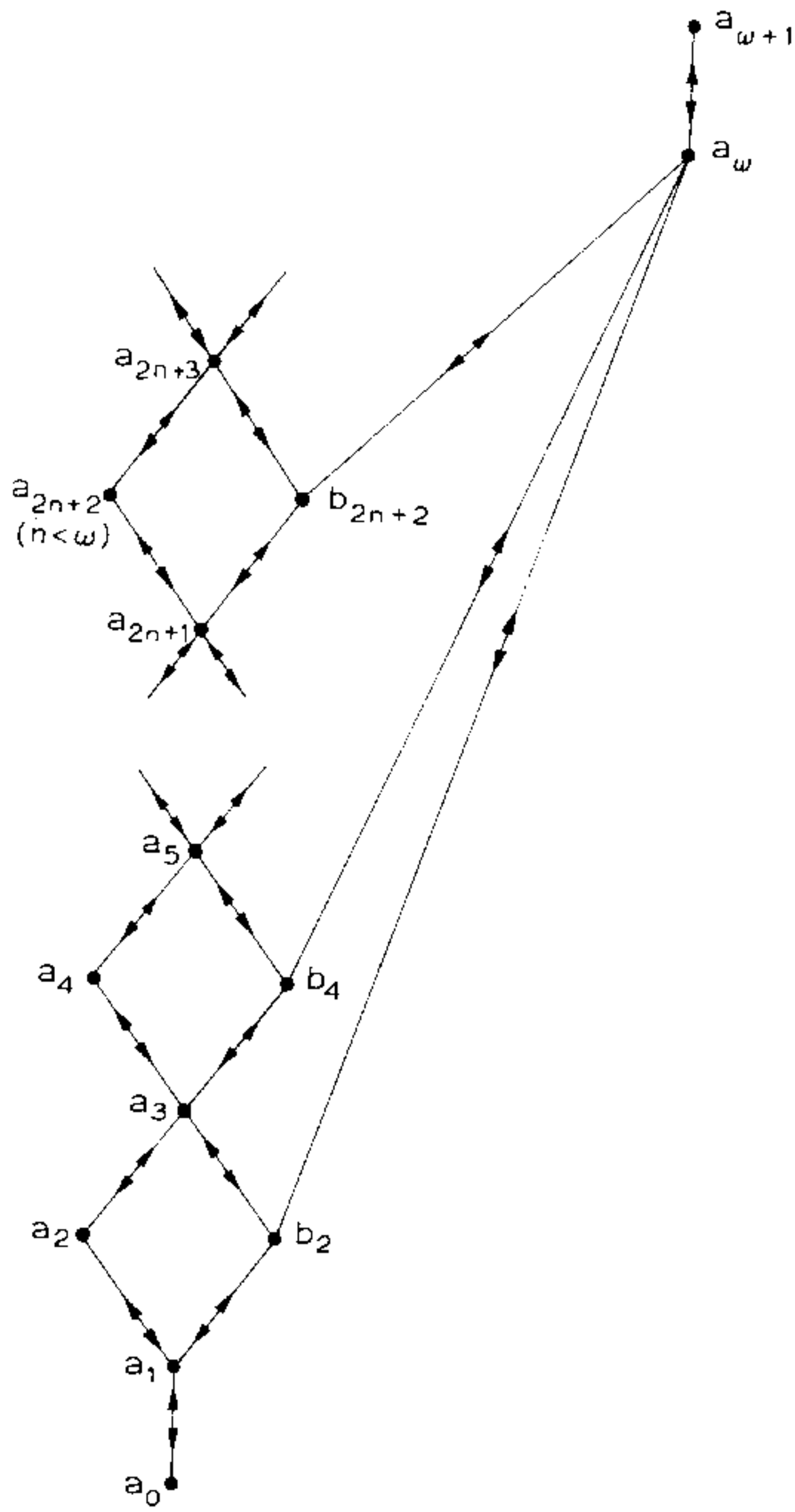


Fig. 6.

It is shown in [1] that a *maximal chain of an undirected graph with no infinite independent set is the vertex set of a path of this graph.* However a maximal chain of an arbitrary undirected graph is not necessarily the vertex set of a path, as shown by Fig. 6. The graph $\mathcal{G}$ (Fig. 6) is a countable undirected graph. $C = \{a_n : n \leqslant \omega + 1\}$ is a maximal chain of $\mathcal{G}$ which is not the vertex set of a path, (see [1] for the details), since it is not connected.

The two notions of maximal chain and maximal vertex set of a path are identical if and only if each chain is contained in a maximal vertex set of a path (this is the case for finite graphs), i.e. *if each maximal chain is the vertex set of a path.* This is for example the case *if each finite vertex set of a path is contained in a maximal vertex set of a path* (this condition always holds for acyclic graphs), *and if the set of the maximal vertex sets of the paths of the graph* $\mathcal{G} = (V, E)$ *is closed in* 2^V, *i.e. compact, for the product topology*: Indeed each finite subset of a chain is then contained in a maximal vertex set of a path and the result follows by compactness.

3. Conjectures and counter-examples

The main problem is whether the following extension of the Gallai–Milgram theorem to infinite graphs holds.

Problem 1. If $1 \leqslant k < \omega$ and $\mathscr{G} = (V, E)$ is a graph with no independent set of size $k+1$, then V is the union of $\leqslant k$ pairwise disjoint paths of $\mathscr{G}$.

The problem is still open even if $\mathscr{G}$ is countable, and if we do not require the paths to be disjoint.

Problem 1 has a positive answer for $k = 1$. If E' is a maximal subset of E such that $\mathscr{G}' = (V, E')$ is acyclic, then any independent set of the transitive closure of $\mathscr{G}'$ is an independent set of $\mathscr{G}$; hence $\mathscr{G}'$ is a path if $\mathscr{G}$ is complete. The following observation, which shows that one can say a little more, was suggested to us by Aharoni and Milner.

Fact 1. *If $\mathscr{G} = (V, E)$ is a complete directed graph, then there is $E' \subseteq E$ such that $\mathscr{G}' = (V, E')$ is a path with the additional property: For any distinct $x, y \in V$ such that x precedes y on $\mathscr{G}'$, either $(x, y) \in E'$ or there is $z \in V$ such that $(x, z) \in E'$ and $(z, y) \in E'$.*

Proof. Without loss of generality, we can assume that $\mathscr{G}$ is a tournament. Let $V = \{x_\alpha : \alpha < \kappa\}$ where $\kappa = |V|$. We inductively define an increasing sequence $(E_\alpha : \alpha \leqslant \kappa)$ of subsets of E such that for $\alpha \leqslant \kappa$ the transitive closure $\mathscr{P}_\alpha = (V, \leqslant_\alpha)$ of the graph $\mathscr{G}_\alpha = (V, E_\alpha)$ is an ordering for which two distinct elements x, y are incomparable if and only if $x, y \notin V_\alpha = \{x_\beta : \beta < \alpha\}$ and $I_\alpha(x) = I_\alpha(y)$ where $I_\alpha(z) = \{t \in V_\alpha : t <_\alpha z\}$ $(z \in V)$. We let $E_\alpha = \bigcup\{E_\beta : \beta < \alpha\}$ if $\alpha = 0$ or α is limit, and

$$E_\alpha = E_\beta \cup \{(x, y) \in E : x \neq y,\ x_\beta \in \{x, y\},\ x, y \notin V_\beta,\ I_\beta(x) = I_\beta(y)\}$$

if $\alpha = \beta + 1$ is successor. Now if we let $E' = E_\kappa$, we get clearly that $\mathscr{P}' = \mathscr{P}_\kappa$ is a linear ordering and therefore that $\mathscr{G}'$ is a path. Furthermore, for $\alpha < \beta < \kappa$, let $\gamma < \kappa$ be the least ordinal such that either $\gamma = \alpha$ or $I_{\gamma+1}(x_\alpha) \neq I_{\gamma+1}(x_\beta)$, (note also that the least ordinal $\delta < \kappa$ such that $I_\delta(x_\alpha) \neq I_\delta(x_\beta)$ is a successor ordinal). Assuming for example that $x_\alpha <_\kappa x_\beta$, we get in the first case that $(x_\alpha, x_\beta) \in E_{\gamma+1} \subseteq E'$ and in the second case that $(x_\alpha, x_\gamma), (x_\gamma, x_\beta) \in E_{\gamma+1} \subseteq E'$. $\square$

Problem 1 has a positive answer for transitive graphs.

Fact 2. *Let $k \geqslant 1$ be an integer and $\mathscr{G} = (V, E)$ be a transitive directed graph. Then $\mathscr{G}$ has no independent set of size $k+1$ if and only if V is the union of $\leqslant k$ (pairwise disjoint) paths of $\mathscr{G}$.*

Proof. The main point is that in a transitive graph the vertex set of path is a clique and conversely, as we already observed in Section 2.4. Thus a subset of the vertex set of a path is the vertex set of a path and $P \subseteq V$ is the vertex set of a path if each finite subset of P is the vertex set of a path. Therefore a standard compactness argument using the finite classical Gallai–Milgram theorem will prove the nontrivial implication. The reverse implication follows immediately from the transitivity of the graph $\mathcal{G}$. Note also that Fact 2 easily reduces to Dilworth's theorem. □

A natural generalization of Fact 2 to arbitrary graph is the following:

Fact 3. *Let $k \geqslant 1$ be an integer and $\mathcal{G} = (V, E)$ be a directed graph. Then V is the union of $\leqslant k$ chains if and only if each finite subset of V is the union of $\leqslant k$ chains.*

Proof. We have only to note that $C \subseteq V$ is a chain if only if each finite subset of C is a chain. One implication is obvious, the other one follows from a standard compactness argument: For each finite subset F of V, let $\mathcal{U}_F = \{R \subseteq V \times V : R \cap (F \times F)$ is an equivalence relation on F with at most k classes which are chains of $\mathcal{G}\}$. The $\mathcal{U}_F$'s, (F finite subset of V), are closed non-empty subsets of the compact space $2^{V \times V}$, and they finitely intersect. Hence there is $R \in \bigcap\{\mathcal{U}_F$: F finite subset of $V\}$, and R is an equivalence relation on V with at most k classes which are chains of $\mathcal{G}$. □

Remark. In Fact 3, we can replace the notion of 'chain', by the notion of 'clique in the transitive closure'; but the corresponding statement is then Fact 2, (as shown by the next sentence).

Now we can note that, using the classical Gallai–Milgram theorem, the hypothesis of Problem 1 that there is no independent set of size $k + 1$ in the graph $\mathcal{G} = (V, E)$ is equivalent to the fact that each finite subset of V is the union of $\leqslant k$ (pairwise disjoint) paths. But this hypothesis is not a necessary condition for the conclusion of Problem 1, (except if the graph $\mathcal{G}$ is transitive). A necessary condition for the conclusion of Problem 1 is that each finite subset is contained in $\leqslant k$ pairwise disjoint finite paths, (which is the same as the original condition if $\mathcal{G}$ is transitive). Note also that the hypothesis of Fact 3 that each finite subset of V is the union of $\leqslant k$ chains is equivalent to the fact that each finite subset of V is contained in $\leqslant k$ (finite) paths. From these considerations we can be tempted to strengthen both Problem 1 and Fact 3 in the following.

Problem 2. If $1 \leqslant k < \omega$ and $\mathcal{G} = (V, E)$ is a graph such that each finite subset of V is contained in $\leqslant k$, (respectively $\leqslant k$ pairwise disjoint), paths, $(1 \leqslant k < \omega)$, then V is the union of $\leqslant k$, (respectively $\leqslant k$ pairwise disjoint), paths.

We can weaken the conclusion of Problems 1 and 2 asking only that V be the union of $f(k)$ paths for some function $f:\omega\to\omega$, or even of finitely many paths, and also by removing the condition that these paths be pairwise disjoint. The answer to Problem 2 is trivially positive for $k=1$ if we consider acyclic graphs. In Section 4.2, we shall obtain a positive result in this direction, for the special case of an undirected graph which has no infinite independent set. But the next examples, (cf. [1]), which are given here without proof, show that the answer to Problem 2 is negative in general.

Example 1. An undirected (bipartite) graph $\mathcal{G}=(V, E)$ with no maximal vertex set of a path and such that any finite subset of V is contained in a finite path but V is not the union of fewer than $|V|$ paths.

Let $\mathcal{G}=K_{\lambda,\mu}=(V, E)$ with $V=X\cup Y$ and $X=\{x_\alpha: \alpha<\lambda\}$, $(x_\alpha=(\alpha, 0))$, $Y=\{y_\beta: \beta<\mu\}$, $(y_\beta=(\beta, 1))$, $E=(X\times Y)\cup(Y\times X)$, where λ and μ are two infinite cardinals. The graph $\mathcal{G}$ is a (bipartite) undirected graph (which is $(\min(\lambda, \mu))$-connected, as defined in Section 4.1) such that any finite subset of V is contained in a finite path. More precisely if $V'=X'\cup Y'$ with $X'\subseteq X$ and $Y'\subseteq Y$, then, for $E'\subseteq E$, $\mathcal{G}'=(V', E')$ is a path of $\mathcal{G}$ if and only if Y' is dense in X' and X' is dense in Y' in $\mathcal{G}'$, (i.e. in $\mathcal{G}'$ there is an element of Y' between any two distinct elements of X' and there is an element of X' between any two distinct elements of Y'). Therefore if V' is the vertex set of a path, then X' and Y' are either both finite and $|X'|\leqslant|Y'|+1$, $|Y'|\leqslant|X'|+1$, or both infinite and $|X'|\leqslant 2^{|Y'|}$, $|Y'|\leqslant 2^{|X'|}$; conversely if X' is infinite and $|Y'|=2^{|X'|}$, then V' is the vertex set of a path. Thus V is not the union of fewer than μ paths provided that $2^\lambda<\mu$. We can also note that there is no maximal vertex set of a path under this assumption that $2^\lambda<\mu$.

Example 2 (Fig. 7). A countable undirected (bipartite) graph $\mathcal{G}=(V, E)$ such that any finite subset of V is contained in a finite path but V is not the union of finitely many pairwise disjoint paths.

Independently, we can remark that $\{x_{2n+1}: n<\omega\}\cup\{y_n: n<\omega\}$ (Fig. 7) is the vertex set of a path but is not contained in a maximal one; however each vertex set of a finite path is contained in a maximal vertex set of a path. Note also that V is the union of the two paths $(x_n: n<\omega)$ and $(x_1, y_0, x_3, y_1, \ldots, x_{2n+1}, y_n, \ldots: n<\omega)$. We can in fact build a countable undirected (bipartite) graph $\mathcal{G}=(V, E)$ such that each finite subset of V is contained in a (finite) path but V is not the union of finitely many (even nondisjoint) paths, (see [1] or [3]).

Example 3. A countable acylic graph $\mathcal{G}=(V, E)$ with no infinite independent set, such that each finite subset of V is contained in the union of two disjoint finite paths, but V is not the union of finitely many paths.

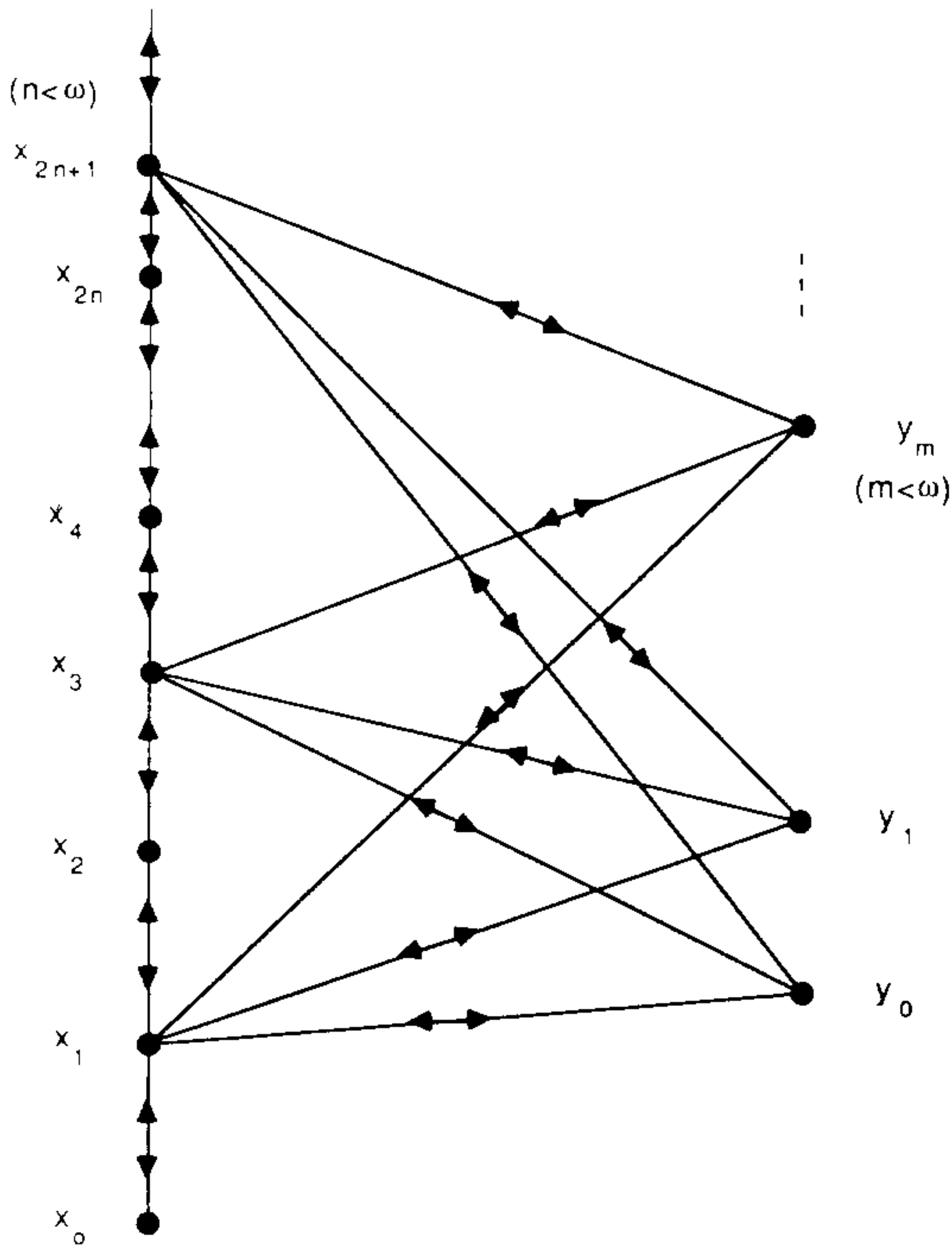


Fig. 7.

Let $\mathscr{P} = (P, <)$ be the ordering such that $P = X \cup Y$ with $X = \{x_n : n \in \omega^* + \omega\}$, $(x_n = (n, 0))$, and $Y = \{y_n : n \in \omega^* + \omega\}$, $(y_n = (n, 1))$, $x_n < x_m$, $y_n < y_m$, $y_n < x_m$ if $n < m$, y_n and x_m are incomparable if $m \leqslant n$. For $k \geqslant 1$, we define a directed graph $\mathscr{G}_k = (V_k, E_k)$ (Fig. 8) by

$$V_k = P \times \{l : 0 \leqslant l < k\}$$

and

$$E_k = \{((z, l), (z', l)) : 0 \leqslant l < k,\ z,\ z' \in P,\ z < z'\}$$
$$\cup \{((x, l), (y, l+1)) : 0 \leqslant l < k-1,\ x \in X,\ y \in Y\}.$$

The graphs $\mathscr{G}_k$ $(1 \leqslant k < \omega)$ are countable, acyclic, and the sizes of their independent sets are finitely bounded. More precisely the maximal size of an independent set of $\mathscr{G}_k$ is $k + 1$, $(1 \leqslant k < v)$. Each finite subset of V_k is contained in the union of two disjoint finite paths of $\mathscr{G}_k$. The minimal number of paths (not necessarily disjoint) of $\mathscr{G}_k$ whose union is V_k is $k + 1$, $(1 \leqslant k < \omega)$. We define then

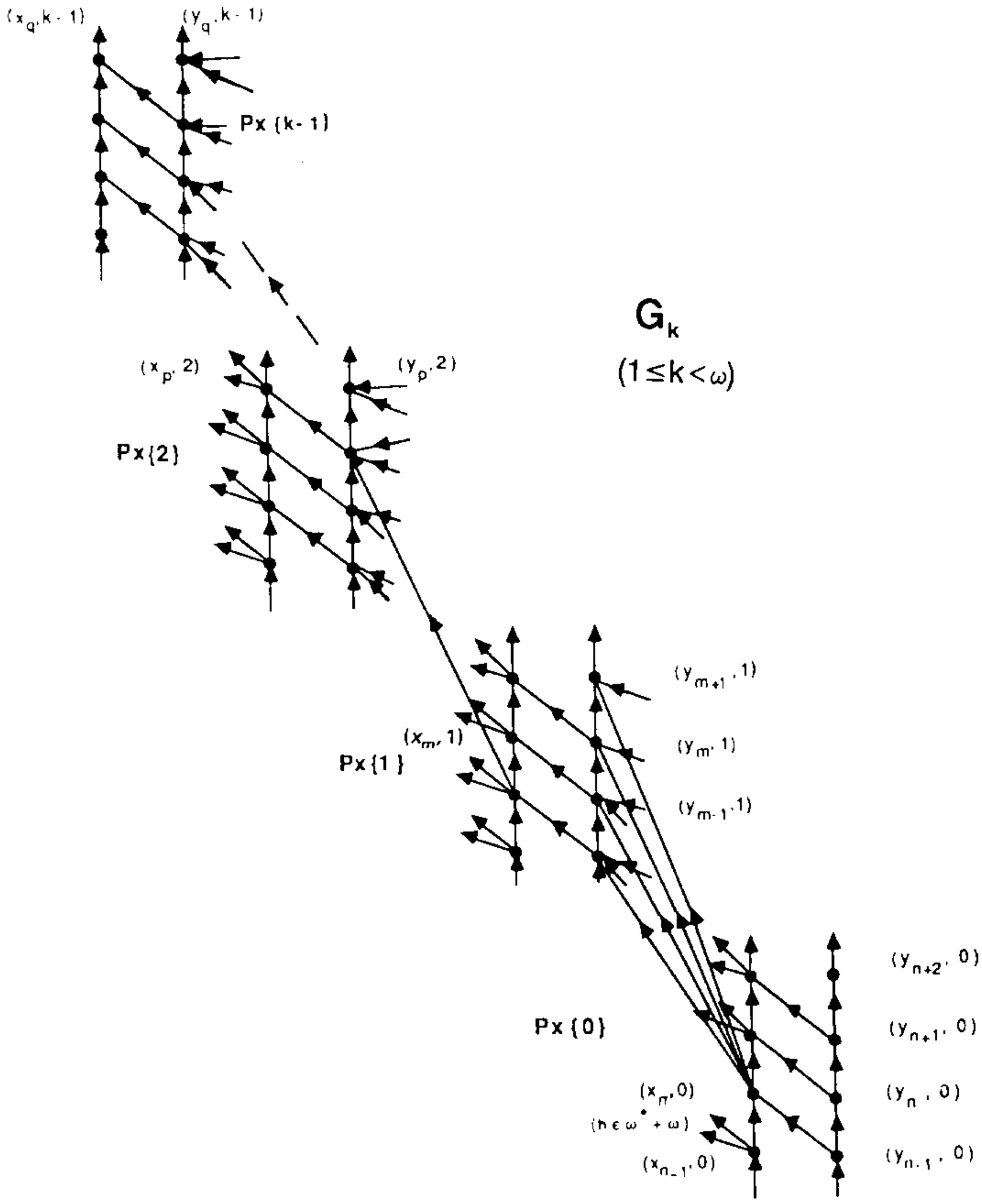


Fig. 8.

the directed graph $\mathcal{G} = (V, E)$ by

$$V = \bigcup \{V_k \times \{k\} : 1 \leq k < \omega\}$$

and

$$E = \{((z, k), (z', k)) : 1 \leq k < \omega,\ z, z' \in V_k,\ (z, z') \in E_k\}$$
$$\cup \{((z, k), (z', k')) : 1 \leq k < k' < \omega,\ z \in V_k,\ z' \in V_{k'}\}.$$

4. Positive results for undirected graphs

In Section 4.2, we show that Problem 2 of Section 3 holds for any undirected graph with no infinite independent set, (cf. Theorem 2). Such a graph has the following property: For any cardinal $\kappa \geqslant \omega$, the connected components of the induced graph obtained by removing fewer than κ vertices properly chosen from the original graph are κ-connected, i.e. they always remain connected when we remove again fewer than κ vertices, (Lemma 2). This leads us to study first the κ-connectivity in Section 4.1. The main tool, with Lemma 2, which we use to establish Theorem 2, is the following property: If the undirected graph $\mathcal{G} = (V, E)$ has size $\kappa \geqslant \omega$ and is κ-connected, then V is covered by a path with arbitrarily prescribed extremities (Lemma 1). In Section 4.3 we survey generalizations of some results of Section 4.2 to the case where the size of the independent sets is bounded by a given infinite cardinal which may be large. The case where the size of the independent sets is bounded by an integer is also investigated.

4.1. κ-Connected undirected graphs. We say that an undirected graph $\mathcal{G} = (V, E)$ is *connected* if $\{x, y\}$ is contained in a finite path of $\mathcal{G}$ for any $x, y \in V$, i.e. if the transitive closure of $\mathcal{G}$ is complete. A *connected set* of an undirected graph $\mathcal{G} = (V, E)$ is a subset C of V such that $\mathcal{G} \restriction C$ is connected. For an undirected graph $\mathcal{G} = (V, E)$, we define an equivalence relation R on V by $(x, y) \in R$ if $\{x, y\}$ is contained in a finite path of $\mathcal{G}$, $(x, y \in V)$. The R-equivalence classes are the *connected components* of the undirected graph $\mathcal{G}$. They are the maximal connected subsets of V.

If κ is a cardinal, the undirected graph $\mathcal{G} = (V, E)$ is *κ-connected* if $\mathcal{G} \restriction (V \setminus V')$ is connected for any subset V' of V of size $|V'| < \kappa$. A *κ-connected set* of an undirected graph $\mathcal{G} = (V, E)$ is a subset C of V such that $\mathcal{G} \restriction C$ is κ-connected. A complete graph is κ-connected for any cardinal κ. But the size of a κ-connected graph which is not complete is at least $\kappa + 2$.

For our purpose, the main property of κ-connected undirected graphs is Proposition 1 below.

Proposition 1. *Let κ be an infinite cardinal and $\mathcal{G} = (V, E)$ be a κ-connected undirected graph. Let $(W, <)$ be a linear ordering where $W \subseteq V$ and $|W| < \kappa$. Also, let $W' \subseteq V$, $|W'| \leqslant \kappa$. Then there is a path $\mathcal{G}' = (V', E')$ of $\mathcal{G}$ such that:*

(i) $W \cup W' \subseteq V'$;

(ii) *the linear ordering $(W, <)$ is induced by the linear ordering defined by $\mathcal{G}'$;*

(iii) *W is coinitial in $\mathcal{G}'$ if $W \neq \phi$ and cofinal if $|W| > 1$;*

(iv) *V' is finite or $|V'| = |W \cup W'|$.*

We shall need in Section 4.2 the following obvious corollary.

Lemma 1. *If $\mathcal{G}=(V, E)$ is a κ-connected undirected graph of size κ, where κ is an infinite cardinal, if x, y are two distinct elements of V, then there is $E' \subseteq E$ such that $\mathcal{G}' = (V, E')$ is a path with first vertex x and last vertex y.*

Remarks. (1) Proposition 1 and Lemma 1 can be improved. For example, we can ensure in Lemma 1 that $\mathcal{G}'$ is scattered (see [1] or [3]), and even well ordered of order type $\kappa + 1$ if $\kappa = \omega$ or if $\mathcal{G}$ has no infinite independent set (see [3]) (density conditions can also be ensured). However we cannot ensure that $\mathcal{G}'$ is well ordered in general (see [3]).

(2) Some additional basic properties on κ-connected graphs have been obtained, (in [1]). For example the following.

If $\mathcal{G}=(V, E)$ is a κ-connected graph of size $\kappa \geqslant \omega$, then V is the union of κ pairwise disjoint κ-connected sets of size κ.

Note also that *any κ-connected set of an undirected graph is contained in a maximal one.*

Proof of Proposition 1. We prove the result by induction on the cardinal $\lambda = |W \cup W'| \leqslant \kappa$. Let $\{x_\alpha : \alpha < \lambda\}$ be a well ordering of $W \cup W'$, such that $x_0 \in W$ if $W \neq \phi$ and $x_1 \in W$ if $|W| \geqslant 2$. We inductively define some paths $\mathcal{P}_\alpha = (P_\alpha, F_\alpha)$ of $\mathcal{G}$ $(1 \leqslant \alpha \leqslant \lambda)$, such that P_α is finite or $|P_\alpha| = |\alpha|$, the linear ordering defined by $\mathcal{P}_\alpha$ coincides with $<$ on $W \cap P_\alpha$, and $\{x_\beta : \beta < \alpha\} \subseteq W \cap P_\alpha$, in the following way. $\mathcal{P}_1 = (\{x_0\}, \phi)$. For α limit, $P_\alpha = \bigcup\{P_\beta : \beta < \alpha\}$ and $F_\alpha = \bigcup\{F_\beta : \beta < \alpha\}$. For $\alpha = \beta + 1$ successor $(\beta \geqslant 1)$, let $<_\beta$ be a linear ordering of $P_\beta \cup \{x_\beta\}$ which extends the linear ordering defined by $\mathcal{P}_\beta$, and also which coincides with $<$ on $W \cap (P_\beta \cup \{x_\beta\})$, and such that x_β is not the first element in this order if $x_\beta \notin W$, and is not the last element if $x_\beta \notin W$ and $\beta \geqslant 2$. If λ is infinite, then, by the induction hypothesis, there is a path $\mathcal{P} = (P, F)$ of the κ-connected graph $\mathcal{G} \restriction (V \setminus (W \setminus (P_\beta \cup \{x_\beta\})))$ such that $P_\beta \cup \{x_\beta\} \subseteq P$, the linear ordering defined by $\mathcal{P}$ coincides with $<_\beta$ on $P_\beta \cup \{x_\beta\}$, P is finite or $|P| = |P_\beta|$. Moreover, we can assume that $P_\beta \cup \{x_\beta\}$ is coinitial and cofinal in $\mathcal{P}$. We let $\mathcal{P}_\alpha = \mathcal{P}$. If λ is finite, then $\mathcal{P}_\beta$ is finite. Let x' be the last element lower than x_β for $<_\beta$ and x'' be the first element higher than x_β, (if x_β is an extremal element of $<_\beta$, the construction is similar but simpler). There is a finite path $\mathcal{P}' = (P', F')$ from x' to x_β in the connected graph $\mathcal{G} \restriction (V \setminus ((W \cup P_\beta) \setminus \{x', x_\beta\}))$ and there is a finite path $\mathcal{P}'' = (P'', F'')$ from x_β to x'' in the connected graph $\mathcal{G} \restriction (V \setminus ((W \cup P_\beta \cup P') \setminus \{x_\beta, x''\}))$. We let $P_\alpha = P_\beta \cup P' \cup P''$ and $F_\alpha = F_\beta \cup F' \cup F''$. Clearly $\mathcal{G}' = \mathcal{P}_\lambda$ satisfies the conditions of Proposition 1. □

4.2. Undirected graphs with no infinite independent set. The main results of this section are the following ones.

Theorem 1. *Let $\mathcal{G}=(V, E)$ be an indirected graph with no infinite independent set. Then V is the union of some pairwise disjoint paths such that there is an independent set with a vertex in each of these paths.*

Theorem 2. *Let $\mathcal{G}=(V, E)$ be an undirected graph with no infinite independent set, $1 \leqslant k < \omega$ and $f:[k]^2 \to [\omega]$ be a map which assigns a set of nonnegative integers to each (unordered) pair of nonnegative integers smaller than k. If each finite subset of V is contained in $\leqslant k$ finite paths $\mathcal{F}_i$ $(i < l \leqslant k)$ such that $|\mathcal{F}_i \cap \mathcal{F}_j| \in f(\{i, j\})$ for $i \neq j$, then V is the union of $\leqslant k$ paths $\mathcal{P}_i$ $(i < p \leqslant k)$ such that $|\mathcal{P}_i \cap \mathcal{P}_j| \in f(\{i, j\})$ for $i \neq j$.*

Theorem 1 shows in particular that V is the union of finitely many pairwise disjoint paths of $\mathcal{G}$ if the undirected graph $\mathcal{G}$ has no infinite independent set; and that V is the union of $\leqslant k$ pairwise disjoint paths if $\mathcal{G}$ has no independent set of size $k+1$ $(1 \leqslant k < \omega)$. Hence it shows that Problem 1 of Section 3 has a positive answer for the undirected graphs.

In Theorem 2, we can ensure in particular that the paths $\mathcal{P}_i$ $(i < p)$ are pairwise disjoint, provided that each finite subset of V is contained in $\leqslant k$ pairwise disjoint paths. This shows that Problem 2 of Section 3 has a positive answer for the undirected graphs with no infinite independent set.

Theorem 2, for $k=1$, shows immediately that the union of an up-directed family of vertex sets of paths of an undirected graph $\mathcal{G}$ with no infinite independent set is the vertex set of a path of $\mathcal{G}$. Hence, by Zorn's lemma, any vertex set of a path of $\mathcal{G}$ is contained in a maximal one. In fact, this follows from the fact that any chain of an undirected graph with no infinite independent set is contained in the vertex set of a path. This last property is easy to check directly by applying Lemma 3′ (and Lemma 2) in the same way as it will be used in the proof of Theorem 2 below. More precisely it is proved in [1] or [3] the following.

Theorem 3. *There are only finitely many maximal vertex sets of paths in an undirected graph with no infinite independent set, and hence the notions of maximal chain and of maximal vertex set of a path are identical in such a graph.*

Note finally that Theorem 1 is false for directed graphs in general. Any ordering (or transitive graph) with no infinite antichain but with arbitrarily large finite antichains will provide a counter-example. We recall a well known example, (see [8]).

Example. For an infinite cardinal κ, let the ordering $\mathcal{P}=(P, \leqslant)$ where $P=\{(\alpha, \beta): \alpha, \beta < \kappa\}$ and $\leqslant = \{((\alpha, \beta), (\alpha', \beta')) \in P \times P: \alpha \leqslant \alpha' \text{ and } \beta \leqslant \beta'\}$. $\mathcal{P}$ has no infinite antichain but P is not the union of fewer than κ chains.

Besides Lemma 1 of Section 4.1, the principal tool that we use to prove Theorems 1 and 2 is the following result.

Lemma 2. *If $\mathcal{G}=(V, E)$ is an undirected graph with no infinite independent set, then for any infinite cardinal κ, there is a subset V' of V of size $|V'| < \kappa$ such that*

any connected component of the graph $\mathcal{G} \restriction (V \setminus V')$ is κ-connected, (note that there are only finitely many such connected components).

Proof of Lemma 2. We shall use the following self-evident fact: If $\{V_i: i \in I\}$ is the set of the connected components of the undirected graph $\mathcal{G} = (V, E)$ then, for any set $\{W_i: i \in I\}$ such that $W_i \subseteq V_i$ for $i \in I$, the set of the connected components of the graph $\mathcal{G} \restriction (V \setminus (\bigcup\{W_i: i \in I\}))$ is $\{W_{i,j}: i \in I, j \in J_i\}$, where $\{W_{i,j}: j \in J_i\}$ is the set of the connected components of the graph $\mathcal{G} \restriction (V_i \setminus W_i)$, $(i \in I)$. Assume for contradiction that, for a given infinite cardinal κ, $\mathcal{G}$ fails to satisfy the conclusion of Lemma 2. Then this is also the case for $\mathcal{G} \restriction C$, for at least one connected component C of $\mathcal{G} \restriction (V \setminus V')$ whenever $V' \subseteq V$ and $|V'| < \kappa$. Hence we can inductively define V_n, $V'_n \subseteq V$ $(n < \omega)$ such that $V_0 = V$, $\mathcal{G} \restriction V_n$ does not satisfy the conclusion of Lemma 2, $|V'_n| < \kappa$ and $V_n \setminus V'_n$ is not connected in $\mathcal{G}$, V_{n+1} is a connected component of $\mathcal{G} \restriction (V_n \setminus V'_n)$, (such that $\mathcal{G} \restriction V_{n+1}$ does not satisfy the conclusion of Lemma 2). Choose $x_n \in V_n \setminus (V'_n \cup V_{n+1})$ $(n < \omega)$. The set $\{x_n: n < \omega\}$ so defined is an infinite independent set of $\mathcal{G}$ and this is a contradiction. □

Proof of Theorem 1. We first show by induction on the size of the graph $\mathcal{G} = (V, E)$ that V is the union of finitely many pairwise disjoint paths. If V is finite, the result is obvious. Otherwise, let $|V| = \kappa \geqslant \omega$. There is $V' \subseteq V$ such that $|V'| < \kappa$ and the connected components of $\mathcal{G} \restriction (V \setminus V')$ are κ-connected. They are only finitely many and they are vertex sets of paths of $\mathcal{G}$ by Lemma 1. To conclude, it is enough to apply the induction hypothesis to $\mathcal{G} \restriction V'$.

Assume now that n is the minimal integer such that V is the union of n pairwise disjoint paths $\mathcal{P}_i = (P_i, E_i)$ of $\mathcal{G}$ $(i < n)$. Then, for $i \neq j < n$, since $(P_i \cup P_j, E_i \cup E_j^{-1} \cup (E \cap (P_i \times P_j)))$ is not a path of $\mathcal{G}$, there are $x_{i,j} \in P_i$, $x_{j,i} \in P_j$ such that $(y, z) \notin E$ when y does not precede $x_{i,j}$ on $\mathcal{P}_i$ and z does not precede $x_{j,i}$ on $\mathcal{P}_j$. It is enough to note that if $x_i \in P_i$ is the last element amongst $\{x_{i,j}: j < n, j \neq i\}$ along $\mathcal{P}_i$ $(i < n)$, then $\{x_i: i < n\}$ is an independent set of $\mathcal{G}$. □

To prove Theorem 2, we first study the case where each finite subset of V is contained in a path of $\mathcal{G}$. If $\mathcal{G} = (V, E)$ is an undirected graph and $x, y \in V$, $W \subseteq V$, then we denote by $C(\mathcal{G}, x, y, W)$ the following assertion: *If $n < \omega$ and x_l, y_l $(l < n)$ are pairwise distinct elements of $V \setminus \{x, y\}$ such that $(x_l, y_l) \in E$ for $l < n$, then there is a path $\mathcal{G}' = (W', E')$ of $\mathcal{G}$ such that* (i) $W' \supseteq W$, (ii) *x and y are respectively the first and last vertices of $\mathcal{G}'$,* (iii) for each $l < n$, x_l and y_l *are consecutive vertices of $\mathcal{G}'$,* and (iv) W' is finite or $|W'| = |W|$.

We need the following compactness theorem.

Lemma 3. *Let $\mathcal{G} = (V, E)$ be an undirected graph having no infinite independent set and $x, y \in V$. If $C(\mathcal{G}, x, y, W)$ holds for every finite set $W \subseteq V$, then $C(\mathcal{G}, x, y, V)$ holds.*

Proof of Lemma 3. We prove by induction on $|W|$ that $C(\mathcal{G}, x, y, W)$ holds for $W \subseteq V$. Suppose that $|W| = \kappa \geqslant \omega$. By Lemma 2, there are $W' \subseteq W$, $|W'| < \kappa$ such that each connected component V_i $(i < p, p < \omega)$ of $\mathcal{G} \restriction (W \setminus W')$ is κ-connected and of size $|V_i| = \kappa$. Let $n < \omega$ and x_l, y_l $(l < n)$ be pairwise distinct elements of $V \setminus \{x, y\}$ such that $(x_l, y_l) \in E$, $(l < n)$. For each $i < p$ choose pairwise distinct elements x_{n+i} y_{n+i} of $V_i \setminus (\{x, y\} \cup \{x_l, y_l : l < n\})$ such that $(x_{n+i}, y_{n+i}) \in E$. By the induction hypothesis, $C(\mathcal{G}, x, y, W')$ holds and so there is a path $\mathcal{G}' = (W'', E'')$ such that $W'' \supseteq W'$, x and y are the first and last vertices of $\mathcal{G}'$ and x_l and y_l are consecutive vertices of $\mathcal{G}'$ for $l < n + p$, and $|W''| < \kappa$. Since $V_i \setminus (W'' \setminus \{x_{n+i}, y_{n+i}\})$ is κ-connected, it follows by Corollary 1 that there is $E_i \subseteq E$ such that $\mathcal{G}_i = (V_i \setminus (W'' \setminus \{x_{n+i}, y_{n+i}\}), E_i)$ is a path having x_{n+i} and y_{n+i} as first and last vertices if x_{n+i} precedes y_{n+i} in $\mathcal{G}'$, as last and first vertices otherwise, $(i < p)$. Then $\mathcal{G}' = (V', E')$, with $V' = W'' \cup W$ and $E' = E'' \cup \bigcup\{E_i : i < p\}$ is a witness that $C(\mathcal{G}, x, y, W)$ holds. □

Note that, by Proposition 1 of Section 4.1, if $\mathcal{G} = (V, E)$ is an ω-connected undirected graph, then $C(\mathcal{G}, x, y, W)$ holds for every $x, y \in V$ and every finite set $W \subseteq V$. Hence we have the following corollary of Lemma 3.

Lemma 3′. *If $\mathcal{G} = (V, E)$ is an ω-connected undirected graph with no infinite independent set, then $C(\mathcal{G}, x, y, V)$ holds for every $x, y \in V$.*

Proof of Theorem 2. For $|V| < \omega$ the result is obvious, and so we assume that $|V| \geqslant \omega$. By Lemma 2, there are a finite set $V' \subseteq V$ and a finite number of infinite sets V_i $(i < n)$ which are the distinct connected components of $\mathcal{G} \restriction (V \setminus V')$ and are moreover ω-connected. For each $i < n$, let X_i be a finite subset of V_i of size $|X_i| = k \bullet |V'| + k + 1$.

By the hypothesis of Theorem 2, there are a finite set $X \subseteq V$ and $p \leqslant k$ such that $V' \cup \bigcup\{X_i : i < n\} \subseteq X$, and X is the union of p paths $\mathcal{H}_j = (H_j, E_j)$ of $\mathcal{G}$ $(j < p)$ such that $|H_j \cap H_{j'}| \in f(\{j, j'\})$ if $j < j' < p$.

Let $i < n$. Since $X_i \subseteq X \subseteq \bigcup\{H_j : j < p\}$, $|X_i| = k \bullet |V'| + k + 1$ and $p \leqslant k$, it follows that there is some index $j_i < p$ such that $|X_i \cap H_{j_i}| \geqslant |V'| + 2$. Hence there is some element $x_i \in X_i \cap H_{j_i}$ which is not the last element of $\mathcal{H}_{j_i}$ and is such that its successor x_i' on $\mathcal{H}_{j_i}$ is not an element of V'. Since $(x_i, x_i') \in E$, $x_i \in X_i \subseteq V_i$ and $x_i' \notin V'$, and since V_i is a connected component of $\mathcal{G} \restriction (V \setminus V')$, it follows that x_i' also belongs to the set V_i. Put $V_i' = V_i \setminus (X \setminus \{x_i, x_i'\})$.

Since V_i' $(i < n)$ is ω-connected, it follows from Lemma 3′ that $C(\mathcal{G} \restriction V_i', x_i, x_i', V_i')$ holds. Hence there is a path $\mathcal{Q}_i = (V_i', E_i')$ having first vertex x_i and last vertex x_i' $(i < n)$. Put $P_j = H_j \cup \bigcup\{V_i' : j_i = j\}$ and $E_j^* = E_j \cup \bigcup\{E_i' : j_i = j\}$ $(j < p)$. Then $\mathcal{P}_j = (P_j, E_j^*)$ is a path of $\mathcal{G}$ for $j < p$ and $V = X \cup \bigcup\{V_i' : i < n\} = \bigcup\{P_j : j < p\}$. Moreover, since the sets V_i' $(i < n)$ are pairwise disjoint, it follows that $|P_j \cap P_{j'}| = |H_j \cap H_{j'}| \in f(\{j, j'\})$ for $j < j' < p$. □

Remark. (1) If the undirected graph $\mathcal{G}=(V, E)$ has no infinite independent set and if $W \subseteq V$ is the vertex set of a path of $\mathcal{G}$, then W is the vertex set of a well ordered path of $\mathcal{G}$ with order type smaller than $|W| \cdot \omega$, (see [3]). Hence in Theorems 1 and 2 we can ensure that the paths are well ordered. In Theorem 1 we can in fact even ensure that the order types of the paths are cardinals.

(2) It is easy to prove directly that Problem 1 of Section 3 has a positive answer for undirected graphs. Namely let $1 \leqslant k < \omega$ and $\mathcal{G}=(V, E)$ be an undirected graph with no independent set of size $k+1$. We inductively define $X_\alpha \subseteq V$, some 1-1 maps $\sigma_{\beta,\alpha}: X_\alpha \to X_\beta$ for $\beta < \alpha$ and a cofinal subset I_α of α, $(\alpha < \lambda)$, in the following way: If $\bigcup\{X_\beta: \beta < \alpha\} = V$, we let $\lambda = \alpha$; otherwise X_α is defined to be an independent set of $\mathcal{G} \restriction (V \setminus (\bigcup \{X_\beta : \beta < \alpha\})$ of maximal size. For $\beta < \alpha < \lambda$, by the classical Gallai–Milgram theorem applied to the finite directed graph $\mathcal{G}_{\beta,\alpha} = (X_\beta \cup X_\alpha, E \cap (X_\beta \times X_\alpha)$, there is a 1-1 map $\sigma_{\beta,\alpha}: X_\alpha \to X_\beta$ such that $(\sigma_{\beta,\alpha}(x), x) \in E$ for any $x \in X_\alpha$. Since all the sets X_β $(\beta < \alpha)$ have their size bounded by $k < \omega$, there is a cofinal subset I_α of α such that $\sigma_{\gamma,\alpha} = \sigma_{\gamma,\beta} \circ \sigma_{\beta,\alpha}$ for β, $\gamma \in I_\alpha$ and $\gamma \in I_\beta$. Let us define $R = \{(x, y) \in V \times V: x \in X_\beta, y \in X_\alpha$ for some α, $\beta < \lambda$ such that $\beta \in I_\alpha$ and $\sigma_{\beta,\alpha}(y) = x$, or $\alpha \in I_\beta$ and $\sigma_{\alpha,\beta}(x) = y\}$. The transitive closure of R is an equivalence relation on V which has $|X_0| \leqslant k$ classes and whose classes are obviously vertex sets of well-ordered paths of $\mathcal{G}$.

A slight modification of the above proof would show that V is the union of $\leqslant k$ pairwise disjoint well ordered paths whose order types are cardinals.

4.3. Undirected graphs with no independent set of size ν. The ideas used in Section 4.2 have been applied in [2] and [3] to study the class of the undirected graphs having no independent set of size ν, where ν is an arbitrary, finite in [2], or infinite in [3], cardinal.

For example we show in [2] the following finite version of Theorem 2.

Theorem 2′. *There is a map $f: \omega \to \omega$ such that for $1 \leqslant n, k < \omega$ and any undirected graph $\mathcal{G}=(V, E)$ with no independent set of size $n+1$, V is the union of $\leqslant k$ pairwise disjoint paths if and only if each subset of V of size $\leqslant f(n)$ is contained in $\leqslant k$ pairwise disjoint paths.*

For $k=1$ this gives a criterion for the existence of a hamiltonian path (or similarly a hamiltonian cycle), and enables to show that this existence problem is solvable in polynomial time for the class of the finite undirected graphs with no independent set of size $n+1$, $(1 \leqslant n < \omega)$, (whereas this problem is NP-complete in general).

To present the general infinite version of Theorem 2, we first need some definitions. A *tree* is a well founded ordering $\mathcal{T}=(T, <)$ which has a minimum element and is such that for any $x \in T$, the set $(\leftarrow x[= \{y \in T: y < x\}$ is a chain of $\mathcal{T}$. The *height* of x is the order type $\mathrm{ht}(x)$ of $(\leftarrow x[$, (which is an ordinal).

For an ordinal α, the αth level of $\mathcal{T}$ is the set $T_\alpha = \{x \in T: \mathrm{ht}(x) = \alpha\}$ and the height of $\mathcal{T}$ is the ordinal $\mathrm{ht}(\mathcal{T}) = \sup\{\mathrm{ht}(x) + 1: x \in T\}$, or the least α such that $T_\alpha = \phi$. We remind the reader that an ordinal is equal to the set of the ordinals smaller than it. A typical example of a tree is the following one, where, for two sets A, B, ${}^BA$ denotes the set of the maps $f: A \to B$.

Example. Let λ, μ be ordinals and let the ordering $\mathcal{T} = (T, <)$ where $T = \bigcup\{{}^\alpha\mu: \alpha < \lambda\}$ and, for $f \in {}^\alpha\mu$, $g \in {}^\beta\mu$, $f < g$ if $\alpha < \beta$ and $f = g \restriction a$. $\mathcal{T}$ is a tree of height λ and for $\alpha < \lambda$, $T_\alpha = {}^\alpha\mu$.

The *cofinality* $\mathrm{cf}(\kappa)$ of an infinite cardinal κ is the smallest cardinal of a cofinal subset of κ. κ is *regular* if $\mathrm{cf}(\kappa) = \kappa$. For an infinite regular cardinal κ, let $\mathrm{SH}(\kappa)$ denote the statement: *There is no Suslin tree of size* κ, i.e. *any tree of height* κ *has a chain or an antichain of size* κ. We show in [3] the following extension, (this is an extension since, by Ramsey's theorem [9], $\mathrm{SH}(\omega)$ is true), of Theorem 2.

Theorem 2″. *Let* $\nu \geqslant \omega$ *and* $\lambda \geqslant 2$ *be cardinals and let* $\mathcal{G} = (V, E)$ *be an undirected graph with no independent set of size* ν. *If each subset of* V *of size* $\leqslant \nu$ *is contained in* $<\lambda$ *pairwise disjoint paths, (respectively if* ν *is regular,* $\mathrm{SH}(\nu)$ *holds and each subset of* V *of size* $< \nu$ *is contained in* $<\lambda$ *pairwise disjoint paths), then* V *is the union of* $<\lambda$ *pairwise disjoint paths.*

As already observed, Theorem 1 shows that V is the union of finitely many pairwise disjoint paths of $\mathcal{G}$ whenever the undirected graph $\mathcal{G} = (V, E)$ has no infinite independent set. Note that the following obvious Corollary of Theorem 2″ generalizes this weak form of Theorem 1.

Corollary. *Let* ν *be an infinite cardinal and let* $\mathcal{G} = (V, E)$ *be an undirected graph with no independent set of size* ν. *Then* (i) V *is the union of* $\leqslant \nu$ *pairwise disjoint paths of* $\mathcal{G}$; (ii) V *is the union of* $<\nu$ *pairwise disjoint paths of* $\mathcal{G}$ *if* ν *is regular and* $\mathrm{SH}(\nu)$ *holds.*

The following extension of Lemma 2, whose proof involves a tree argument, is needed to establish Theorem 2″. It explains how the statement $\mathrm{SH}(\nu)$ occurs in Theorem 2″. Note that the above Corollary follows obviously (modulo Lemma 1) from Lemma 2″ by induction on the size of $\mathcal{G}$. We conclude this section by presenting the proof of Lemma 2″.

Lemma 2″. *Let* ν, κ *be two infinite cardinals and let* $\mathcal{G} = (V, E)$ *be an undirected graph with no independent set of size* ν. *If* $\nu < \mathrm{cf}(\kappa)$ *or* $\nu = \mathrm{cf}(\kappa)$ and $\mathrm{SH}(\mathrm{cf}(\kappa))$ *holds, then there is* $V' \subseteq V'$ *such that* $|V'| < \kappa$ *and the connected components of* $\mathcal{G} \restriction (V \setminus V')$ *are* κ*-connected.*

Proof of Lemma 2″. For an ordinal α and $f \in {}^{\alpha}\nu$, we define a subset V_f of V in the following way. $V_f = V$ if $\alpha = 0$. $V_f = \bigcap\{V_{f \restriction \beta}: \beta < \alpha\}$ if α is limit. Assume now that $\alpha = \beta + 1$ is a successor and let $f' = f \restriction \beta$. We define $V_f = \phi$ if $|V_{f'}| < \kappa$ or if $V_{f'}$ is κ-connected. Otherwise there is $W_{f'} \subseteq V_{f'}$ such that $|W_{f'}| < \kappa$ and $V_{f'} \setminus W_{f'}$ is not connected. Let $\{C_\alpha: \alpha < \lambda\}$ be the set of the connected components of $\mathcal{G} \restriction (V_{f'} \setminus W_{f'})$; then $2 \leqslant \lambda < \nu$ since there is no independent set of size ν. We put $V_f = C_{f(\beta)}$ if $f(\beta) < \lambda$, and $V_f = \phi$ if $\lambda \leqslant f(\beta) < \nu$. This completes the definition of the sets V_f. Assume that $f \in {}^{\alpha}\nu$ for some limit ordinal α and $V_{f \restriction \beta} \neq \phi$ for all $\beta < \alpha$. For each $\beta < \alpha$, choose

$$x_\beta \in V_{f \restriction \beta} \setminus (V_{f \restriction (\beta+1)} \cup W_{f \restriction \beta}).$$

$\{x_\beta: \beta < \alpha\}$ is clearly an independent set of $\mathcal{G}$ of size $|\alpha|$. In particular $V_f = \phi$ if $f \in {}^{\alpha}\nu$ and $\alpha \geqslant \nu$. Let

$$T = \{f \in \bigcup\{{}^{\alpha}\nu: \alpha < \nu\}: V_f \neq \phi\}$$

which is an initial segment of the tree $(\bigcup\{{}^{\alpha}\nu: \alpha < \nu\}, <)$ defined in the previous example, and let $\mathcal{T} = (T, <)$ be the induced tree. We just showed that $\mathcal{T}$ has no chain of size ν. Assume that f, $g \in T$ are incomparable in $\mathcal{T}$. Then V_f, V_g are non-empty and $\{x, y\}$ is an independent set of $\mathcal{G}$ if $x \in V_f$, $y \in V_g$. Indeed, if $\gamma < \min(\alpha, \beta)$ is the minimal ordinal such that $f(\gamma) \neq g(\gamma)$, where $f \in {}^{\alpha}\nu$, $g \in {}^{\beta}\nu$, then $V_{f \restriction (\gamma+1)}$ and $V_{g \restriction (\gamma+1)}$ are two distinct connected components of $\mathcal{G} \restriction (V_{f \restriction \gamma} \setminus W_{f \restriction \gamma})$ which contain respectively x and y. Hence to any antichain of $\mathcal{G}$ there corresponds an independent set of $\mathcal{G}$ of the same size. Therefore $\mathcal{T}$ has no antichain of size ν. Let $X = \{f \in T: f$ is a maximal element of $\mathcal{T}$ and $|V_f| \geqslant \kappa\}$, which is an antichain of $\mathcal{T}$. Hence if f, g are two distinct elements of X, and if $x \in V_f$, $y \in V_g$, we know that $(x, y) \notin E$. Moreover, from the definition of $\mathcal{T}$, it follows also that V_f is κ-connected if $f \in X$, and so $\{V_f: f \in X\}$ is the set of the connected components of $\mathcal{G} \restriction (\bigcup\{V_f: f \in X\})$. It is enough to show that $|V \setminus \bigcup\{V_f: f \in X\}| < \kappa$ to end the proof. Note that if $f' \in T \cap ({}^{\beta}\nu)$ is not a maximal element of $\mathcal{T}$, (i.e. if $f' \in {}^{\beta}\nu$ and $W_{f'}$ is defined), then $V_{f'} = W_{f'} \cup \bigcup\{V_f: f \in T \cap ({}^{\beta+1}\nu), f \restriction \beta = f'\}$. Thus $V \setminus \bigcup\{V_f: f \in X\} = \bigcup\{W_f: f \in T$ and f' is not maximal in $\mathcal{T}\} \cup \bigcup\{V_f: f \in T$ and $|V_f| < \kappa\}$. Since $|W_f| < \kappa$ when W_f is defined, it is enough to check that $|T| < \mathrm{cf}(\kappa)$. Since $\mathcal{T}$ is a tree with no chain or antichain of size ν, we get that $|T| \leqslant \nu$ and that $|T| < \nu$ if $\nu = \mathrm{cf}(\kappa)$ since $\mathrm{SH}(\mathrm{cf}(\kappa))$ holds. In any case, we get that $|T| < \mathrm{cf}(\kappa)$. □

5. A Gallai–Milgram property for 'regular' acyclic countable graphs

The Gallai–Milgram theorem for infinite acyclic directed (well-founded) graphs would generalize the Gallai–Milgram theorem that we proved for infinite undirected graphs (Corollary 2). Indeed, if $\mathcal{G} = (V, E)$ is an undirected graph and if $V = \{x_\alpha: \alpha < \kappa\}$ is a well ordering of V, then $\mathcal{G}' = (V, E')$, where $E' =$

$\{(x_\alpha, x_\beta) \in E\colon \alpha < \beta\}$, is an acyclic graph such that $\mathcal{G}$ and $\mathcal{G}'$ have the same independent sets and such that any path of $\mathcal{G}'$ is a path of $\mathcal{G}$, (but, of course, not conversely).

In this section we consider a Gallai–Milgram type property for acyclic graphs of a very special kind. We say that an ordering $\mathcal{P} = (P, <)$ is *n-regular*, $(n < \omega)$, if there is a linear ordering $\mathcal{I} = (I, <)$ such that $P = \bigcup\{C_i\colon i \in I\}$, where the sets C_i are pairwise disjoint finite antichains of $\mathcal{P}$ which have same size n and such that if $i, j \in I$, $i < j$, there is some bijection $\sigma_{i,j}$ from C_i to C_j such that $x < \sigma(x)$ for every $x \in C_i$. The width of $\mathcal{P}$ is clearly n, (i.e. n is the maximal size of an antichain of $\mathcal{P}$), and if $x \in C_i$, $y \in C_j$ for $i, j \in I$, $i < j$, then either $x < y$ or x and y are incomparable. Hence by Dilworth's theorem, P is the union of n pairwise disjoint maximal chains which intersect each C_i in a single point and which are canonically isomorphic to $\mathcal{I}$.

An acyclic graph $\mathcal{G} = (V, E)$ is *n-regular* if its transitive closure is n-regular. Although n is not necessarily the maximal size of an independent set of $\mathcal{G}$, we are interested in the question whether, by analogy with the above result for ordered sets, it is possible to cover V by n 'coherent' paths, (which will be necessarily pairwise disjoint). More precisely, for a given linear ordering $\mathcal{I} = (I, <)$ and a family $\{C_i; i \in I\}$ which witness the fact that $\mathcal{G}$ is n-regular, are there n paths $\mathcal{G}_l = (V_l, E_l)$ of $\mathcal{G}$, $(l < n)$, such that $V = \bigcup\{V_l\colon l < n\}$, (and hence $V_l \cap V_k = \phi$ if $l \neq k$), and such that $F_l = \{(i, j) \in I \times I\colon i < j$ and $(x, y) \in E_l$ for some $x \in C_i$, $y \in C_j\}$ does not depend on l? A necessary condition for this is that the graph (I, F) is a path, where $F = \{(i, j) \in I \times I\colon i < j$ and there is a bijection σ from C_i to C_j such that $(x, \sigma(x)) \in E$ for $x \in C_i\}$. We show now that, in the countable case, this condition is also sufficient. For any $i \in I$, we write $C_i = \{x_{i,l}\colon l < n\}$ and for $(i, j) \in F$, we choose a bijection $\sigma_{i,j}$ from n onto n such that $(x_{i,l}, x_{j,\sigma_{i,j}(l)}) \in E$ for $l < n$. We have to find $F' \subseteq F$ such that the graph (I, F') is a path and such that the composed permutation of $(\sigma_{i_0,i_1}, \sigma_{i_1,i_2}, \ldots, \sigma_{i_{k-1},i_k})$ along a finite path $p = (i_0, i_1, \ldots, i_k)$ of $\mathcal{G}'$ depends only upon the extremities i_0, i_k of p. This follows from Theorem 4 where we take $\mathcal{H} = \mathcal{S}_n$, the nth permutation group.

Theorem 4. *Let $\mathcal{G} = (V, E)$ be a countable path (such that $(x, x) \notin E$ for $x \in V$), let $\mathcal{H} = (H, \bullet)$ be a finite group, and let $\theta\colon E \to H$. For any finite path $p = (x_0, x_1, \ldots, x_k)$ of $\mathcal{G}$, we define $\tilde{\theta}(p) = \theta(x_0, x_1) \bullet \theta(x_1, x_2) \bullet \cdots \bullet (x_{k-1}, x_k) \in H$. Then there is $E' \subseteq E$ such that $\mathcal{G}' = (V, E')$ is a path and for any finite path p of $\mathcal{G}'$, $\tilde{\theta}(p)$ depends only upon the extremities of p.*

Remarks. (1) Since by Cayley's theorem any finite group is a subgroup of a permutation group, Theorem 4 is in fact equivalent to the result on regular acyclic graphs.

(2) We can define $\tilde{\theta}(p)$ for any finite sequence $p = (x_0, x_1, \ldots, x_k)$ such that $(x_i, x_{i+1}) \in E \cup E^{-1}$, $(i < k)$, if we first let $\tilde{\theta}(y, x) = (\theta(x, y))^{-1}$ when $(x, y) \in E$. Then Theorem 4 says that $\tilde{\theta}(p) = 1$ if $x_0 = x_k$.

(3) By Theorem 4, we can define an equivalence relation R on V by $(x, y) \in R$ if $x = y$ or if the value $\tilde{\theta}(x, y)$, ($\tilde{\theta}(y, x)$ if y precedes x in $\mathcal{G}$), calculated in $\mathcal{G}'$, is equal to the identity of $(H, \bullet)$. If we fix an origin $x_0 \in V$, we can index the R-equivalence classes by H. Conversely Theorem 4 says that there is a partition $\{V_\alpha: \alpha \in H\}$ of V such that $\mathcal{G}' = (V, E')$ is a path where

$$E' = \{(x, y) \in E: \theta(x, y) = \alpha^{-1} \bullet \beta \quad \text{if } x \in V_\alpha,\ y \in V_\beta\}.$$

In the case where $(H, \bullet) = (\mathbb{Z}/n\mathbb{Z}, +)$ is the cyclic group of order n, $(2 \leqslant n < \omega)$, and where $\theta(x, y) = \bar{1}$ for every $(x, y) \in E$, we deduce from Theorem 4 the weaker statement that for any countable path $\mathcal{G} = (V, E)$, there is a partition $\{V_l: l \in \mathbb{Z}/n\mathbb{Z}\}$ of V such that, if $l \in \mathbb{Z}/n\mathbb{Z}$ and $x \in V_l$, then $\{y \in V_{l+\bar{1}}: (x, y) \in E\}$ is coinitial in $\mathcal{G} \restriction \{y \in V: x \text{ precedes } y \text{ in } \mathcal{G}\}$ and $\{y \in V_{l+\overline{(n-1)}}: (y, x) \in E\}$ is cofinal in $\mathcal{G} \restriction \{y \in V: y \text{ precedes } x \text{ in } \mathcal{G}\}$. If $\mathcal{G} = \mathcal{I} = (I, <)$ is a countable linear ordering and $n = 2$, then Theorem 2 says that I is the union of two disjoint sets I_0, I_1 which are dense in $\mathcal{I}$, (in the sense that for every $i < 2$ and $x, y \in I$ such that $x < y$, there is $z_i \in I_i$ such that $x \leqslant z_i \leqslant y$). This very particular case is quite easy to check directly, without assuming that $\mathcal{I}$ is countable.

(4) It is not very difficult to check that the conclusion of Theorem 4 is also true if the path $\mathcal{G}$ is scattered, without the assumption that $\mathcal{G}$ is countable, (see [4]). However we prove in [4] that the conclusion of Theorem 4 is true if $\mathcal{G}$ has cardinality κ and MA(κ) holds, (see [7] for a definition of Martin's axiom MA), but false in general if $\mathcal{G}$ has same order type as the real line.

Proof of Theorem 4. We say that a non-empty interval I of $\mathcal{G}$ satisfies the condition $(*)$ if there is $E_I \subseteq E \cap (I \times I)$ such that $\mathcal{G}_I = (I, E_I)$ is a path and for any finite path p of $\mathcal{G}_I$, $\tilde{\theta}(p)$ depends only upon the extremities of p. Define a relation R on V by $(x, y) \in R$ if and only if $[x, y]$ satisfies $(*)$, where $[x, y]$ is the interval of $\mathcal{G}$ of the elements of V which are between x and y in $\mathcal{G}$, (possibly equal to x or y), $(x, y \in V)$. Let, for some $\lambda \in (\omega \setminus \{0\}) \cup \{\omega, \omega^*, \omega^* + \omega\}$, $(x_l: l \in \lambda)$ be a strictly increasing sequence of vertices of $\mathcal{G}$ such that $(x_l, x_{l+1}) \in R$, $(l, l+1 \in \lambda)$. Let E_l be a set of edges which witnesses the fact that $[x_l, x_{l+1}]$ satisfies $(*)$, $(l, l+1 \in \lambda)$. Then the set of edges $E_I = \bigcup\{E_l: l, l+1 \in \lambda\}$ witnesses the fact that I satisfies $(*)$ where I is the interval $\bigcup\{[x_l, x_{l+1}]: l, l+1 \in \lambda\}$ of $\mathcal{G}$. Indeed it is obvious that $\mathcal{G}_I = (I, E_I)$ is a path of $\mathcal{G}$. Now, let $p = (z_0, z_1, \ldots, z_k)$ be a finite path of $\mathcal{G}_I$, $(k \in \omega \setminus \{0\})$. There are some unique $l \in \lambda$ and $r \in \omega$, (depending only upon z_0 and z_k), such that $l + r + 1 \in \lambda$, $z_0 \in [x_l, x_{l+1}] \setminus \{x_{l+1}\}$ and $z_k \in [x_{l+r}, x_{l+r+1}] \setminus \{x_{l+r}\}$. Clearly $x_{l+s} \in \{z_i: i \in k+1\}$ for any $s \in (r+1) \setminus \{0\}$. Define $y_0 = z_0$, $y_{r+1} = z_k$, $y_s = x_{l+s}$ for $s \in (r+1) \setminus \{0\}$, and the finite path $p_s = p \restriction [y_s, y_{s+1}]$ of the graph $\mathcal{G}_s = ([x_{l+s}, x_{l+s+1}], E_{l+s})$ for $s \in r+1$. Then $\tilde{\theta}(p) = \tilde{\theta}(p_0) \bullet \cdots \bullet \tilde{\theta}(p_r)$ depends only upon z_0 and z_k. Hence R is an equivalence relation and any R-equivalence class is an interval I which satisfies $(*)$. Indeed if $(x_l: l \in \lambda)$ is a strictly increasing sequence of vertices of $\mathcal{G}$, $(\lambda \in \{1, 2, \omega, \omega^*, \omega^* + \omega\})$, such that $\{x_l: l \in \lambda\} \subseteq I$ is coinitial and cofinal in $\mathcal{G} \restriction I$, then

$I = \bigcup\{[x_l, x_{l+1}]: \ l, \ l+1 \in \lambda\}$ satisfies $(*)$ since $[x_l, x_{l+1}]$ satisfies $(*)$, $(l, l+1 \in \lambda)$.

Now it remains to prove that there is only one R-equivalence class. For each R-equivalence class I, let E_I be a set of edges which witnesses the fact that I satisfies $(*)$ and let $E' = \{(x, y) \in E$: x and y are in two different R-equivalence classes$\} \cup \{E_I$: I is an R-equivalence class$\}$. The graph $\mathcal{G}' = (V, E')$ is a path. Define the set $H(x, y) = \{\bar{\theta}(p)$: p is a finite path of $\mathcal{G}'$ with extremities x and $y\} \subseteq H$, and the positive integer $n(x, y) = |H(x, y)|$, $(x, y \in V, \ x \neq y)$. Clearly $1 \leqslant n(x, y) \leqslant |H|$ and $n(x, y) = 1$ if and only if x and y are in a same R-equivalence class. Also, since $\mathcal{H} = (H, \bullet)$ is a group (and since $\mathcal{G}'$ is connected), $n(x', y') \leqslant n(x, y)$ if x', y' are between x and y in $\mathcal{G}$, $(x' \neq y'$, possibly $\{x', y'\} \cap \{x, y\} \neq \phi)$. Assume that there are at least two R-equivalence classes and choose x and y in two different classes such that $n(x, y) = n$ is minimum amongst the values different from 1. We shall get the contradiction that $(x, y) \in R$. We can note that $n(x', y') = n$ if x' and y' are in two different R-equivalence classes and are between x and y in $\mathcal{G}$. Moreover if z is between x' and y' in $\mathcal{G}$ and if $h \in H(x', y')$, then there is a finite path p of $\mathcal{G}'$ from x' to y' (or y' to x') containing the vertex z and such that $\bar{\theta}(p) = h$. This is obvious if x' and y' are in the same R-equivalence class. Otherwise, assume for example that x' precedes y' in $\mathcal{G}$ and that z and y' are in two different R-equivalence classes, so that $n(z, y') = n$. $\mathcal{G}'$ is a path, and hence there is a finite path p_1 of $\mathcal{G}'$ from x' to z. Since $\{\bar{\theta}(p_1) \bullet g: g \in H(z, y')\}$ is a subset of $H(x', y')$ of size $n = |H(x', y')|$, we get that $h = \bar{\theta}(p_1) \bullet g$ for some $g \in H(z, y')$, which gives the desired conclusion. Let now $\{x_n: n \in \omega\}$ be a well-ordering of $[x, y]$, (which is infinite since two consecutive vertices of $\mathcal{G}$ are in the same R-equivalence class), and inductively define the finite paths $\mathcal{G}_n = (V_n, E_n)$ of $\mathcal{G}'$, $(n \in \omega)$, by

$$V_0 = \{x_0\}, \qquad E_0 = \phi, \qquad V_{n+1} = V_n \cup \{y_i: i \in m+1\},$$
$$E_{n+1} = E_n \cup \{(y_i, y_{i+1}): i \in m\},$$

$(n \in \omega)$, where $p = (y_0, \ldots, y_m)$ is a finite path of $\mathcal{G}'$ defined in the following way: Let s be the least integer such that $x_s \notin V_n$. If x_s precedes in $\mathcal{G}$ any element of V_n, then $y_0 = x_s$ and y_m is the first element of $\mathcal{G}_n$; if x_s follows in $\mathcal{G}$ any element of V_n, then $y_m = x_s$ and y_0 is the last element of $\mathcal{G}_n$; if x_s is between two elements z, z' of V_n which are consecutive in $\mathcal{G}_n$ and such that z precedes z', then $y_0 = z$, $y_m = z'$, $x_s \in \{y_i: i \in m \setminus \{0\}\}$ and, moreover, $\bar{\theta}(p) = \theta(z, z')$. Let $E_{[x,y]} = \bigcup\{E_n: n \in \omega\} \subseteq E'$. The graphs $\mathcal{G}_n$, $(n \in \omega)$, and hence the graph $\mathcal{G}_{[x,y]} = ([x, y], E_{[x,y]})$, are clearly paths of $\mathcal{G}'$ such that, for any finite subpath p, $\bar{\theta}(p)$ depends only upon the extremities of p. This proves the contradiction that $(x, y) \in R$. $\square$

Acknowledgement

The authors wish to express their thanks to R. Aharoni and very specially to E.C. Milner for his useful comments on the problems discussed here.

References

[1] J.-M. Brochet, Extremal problems in infinite graphs and orders, Ph.D. Thesis, University of Calgary, September 1988.
[2] J.-M. Brochet, Hamiltonian cycles in graphs with bounded independence numbers, preprint SFB 343, Universität Bielefeld, Germany, 1991.
[3] J.-M. Brochet, Cover of graphs by generalized infinite paths, preprint SFB 343, Universität Bielefeld, Germany, 1991.
[4] J.-M. Brochet, Coherent spanning sets of edges in a linear ordering, preprint 1991.
[5] R.P. Dilworth, A decomposition theorem for partially ordered sets, Amer. Math. Soc. 51 (1950) 161–166.
[6] T. Gallai and A.N. Milgram, Verallgemeinerung eines Graphentheoretischen Satzes von Rédei, Acta Sc. Math. 21 (1960) 181–186.
[7] K. Kunen, Set theory, Studies in logic, Vol. 102 (North-Holland, Amsterdam).
[8] M. Perles, on Dilworth's theorem in the infinite case, Israel J. Math. 1 (1963) 108–109.
[9] F.P. Ramsay, On a problem of formal logic, Proc. London Math. Soc. 30 (1930) 264–286.

Discrete Mathematics 95 (1991) 49–67
North-Holland

The age of a relational structure

Peter J. Cameron

School of Mathematical Sciences, Queen Mary and Westfield College, Mile End Road, London, UK, E1 4NS

Received 3 October 1989
Revised 26 January 1990

Abstract

Cameron, P.J., The age of a relational structure, Discrete Mathematics 95 (1991) 49–67.

The age of a relational structure is the class of its finite substructures. This paper surveys aspects of the relationship between structures and their ages, concentrating on three areas: universality, $\aleph_0$-categoricity and homogeneity; measure and Baire category applied to the class of countable structures with given age; and enumerative aspects (the profile of an age).

1. Introduction

In algebra and combinatorics (though not usually in topology), an object is determined by information about its finite subsets. Examples include whether or not the product of two given group elements is equal to a third, whether two vertices of a graph are adjacent, whether an element of an ordered set is less than another.

The age of a structure is the class of its finite substructures (up to isomorphism). This paper is a survey of results about ages of infinite relational structures. I will always assume that the infinite structures under discussion are countable. (In interesting cases, nothing is lost by this assumption, as we shall see).

The topics addressed are:

(i) How do we recognise an age? Of structures with a given age, is there one which is in some way distinguished, e.g. by properties of symmetry, universality or axiomatisability?

(ii) Can we 'measure' in some way the structures whose age is contained in a given age, so as to give meaning to the statement 'most such structures have property $\mathscr{P}$'?

(iii) How does the function enumerating isomorphism types of n-element structures in an age behave, and which functions can be realised? How do properties of the structure X affect this?

As these questions suggest, there are many and varied links with logic, combinatorics, permutation groups, measure theory, and topology.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

2. Ages and universality

A *relational structure* consists of a set X with a family $R = \{r_\alpha : \alpha \in A\}$ of finitary relations on X. That is, r_α is a subset of X^n, where $n = n(\alpha) \geqslant 1$ is the arity of r_α.

With any relational structure (X, R) is associated a language which has relation 'symbols' corresponding to the relations in R. As a notational convenience, I will not distinguish between relations and relation symbols. The language also contains the standard paraphernalia of first-order logic (variables, connectives, quantifiers, and parentheses), so that well-formed formulae can be built up in the usual way.

A *substructure* of (X, R) is a relational structure (Y, S), where Y is a subset of X and $S = \{s_\alpha : \alpha \in A\}$, $s_\alpha = r_\alpha \cap Y^n$, $n = n(\alpha)$. The relations correspond to those in R, and strictly speaking different symbols should be used, as here; but the notational convention above suggests using the same notation for corresponding relations. Of course, this is common mathematical practice; when we write $x < y$, or $x + y = z$, we do not specify whether the arguments of the relations lie in the natural numbers, integers, rationals or reals.

Most objects in combinatorics (graphs, directed or undirected, total or partial orders, equivalence relations) and algebra (groups, rings, etc.) can be regarded as relational structures. A group, for example, is described by a ternary relation which holds at (x, y, z) precisely when $xy = z$. Note that substructures in this case are much more general than subgroups; any subset carries a substructure! (If this situation is undesirable, the usual procedure is to allow function and constant symbols in the language. Then, however, most of what is said here will not apply.)

One small technical point must be mentioned. Often, especially in mathematical logic, relations of arity 0 are permitted; I have forbidden them here. A relation of arity n can be thought of as Boolean function on X^n; so a relation of arity 0 is just a Boolean variable, which may be true or false in a particular situation. If this were allowed, then the empty set would carry several different relational structures, which would be inconvenient. This problem does not arise in logic, where it is customary to forbid the empty set as the domain of a structure (to avoid paradoxes like 'nothing is better than a good book; a bad book is better than nothing; so . . .').

The domain of a relational structure in this paper will always be finite or countable. Again, following common practice, the name of the domain is often applied to the structure.

The *age* of the relational structure X, written $\mathrm{Age}(X)$, is the class of all finite structures which are isomorphic to substructures of X. The fact that the age, as thus defined, is a proper class rather than a set does not really give any trouble; but it will often be convenient to assume that the domain of a structure in the age

consists of the first n natural numbers, for some n. (In combinatorial terms, we are thinking of Age(X) as a class of labelled structures).

The terminology is due to Fraïssé (see [13]).

Knowing the age of X obviously gives us some information about X itself. We can tell, for example, whether X is a graph, whether directed or undirected, whether with or without loops, whether complete or not, bipartite or not, etc.; whether it is a totally or partially ordered set; and so on. (In all these cases, the appropriate axioms are universal statements involving only finitely many points). How much further information we get depends very sensitively on the particular case. If Age(X) consists of the finite complete graphs, then X is uniquely determined as the countably infinite complete graph. On the other hand, if Age(X) consists of all finite totally ordered sets, then X is any countable totally ordered set, and there are $2^{\aleph_0}$ possibilities for it. Note that, in either case, Age(X) contains a unique n-element structure (up to isomorphism).

Let Y be a relational structure with the same language as X. We say that Y is *younger than* X if $\mathrm{Age}(Y) \subseteq \mathrm{Age}(X)$. Clearly, any substructure of X is younger than X. The converse is false. (A two-way infinite path, and the disjoint union of three one-way infinite paths, have the same age; but neither is a substructure of the other.) We say that X is *universal* if every structure younger than X is isomorphic to a substructure of X. (In the terminology of Komjáth and Pach, in their article in this volume, such an X is universal in the class of structures younger than X. In special cases, e.g. universal graphs (Rado [31]), the unqualified word is already standard. Fraïssé's term is *rich for its age.*)

Cantor's celebrated theorem [6] asserts, in part, that $\boldsymbol{Q}$ (as totally ordered set) is universal. We will generalise this in various ways later.

Note that there are usually many universal structures with a given age, since any structure with a universal substructure is itself universal.

We now turn to the question: How do you recognise an age?

A class $\mathscr{K}$ of finite structures has the *joint embedding property* if for any B_1, $B_2 \in \mathscr{K}$, there exists $C \in \mathscr{K}$ and embeddings $g_i : B_i \to C$ for $i = 1, 2$. That is, any two members of $\mathscr{K}$ can be simultaneously embedded in a larger member of $\mathscr{K}$. This condition, together with trivial book-keeping requirements, is necessary and sufficient for $\mathscr{K}$ to be an age.

Theorem 2.1. *The class $\mathscr{K}$ of finite structures is the age of a countable structure X if and only if $\mathscr{K}$ is closed under isomorphism, closed under taking substructures, contains only countably many non-isomorphic members, and has the joint embedding property.*

Proof. The conditions on $\mathscr{K}$ are trivially necessary; so assume that they hold. Enumerate the isomorphism types in $\mathscr{K}$, say $A_0, A_1, \ldots$, and define a sequence $B_0, B_1, \ldots$ recursively as follows: $B_0 = \emptyset$; B_{n+1} is a structure in $\mathscr{K}$ jointly embedding B_n and A_n, $n \geqslant 0$.

We can regard B_n as a substructure of B_{n+1}. The let $X = \bigcup_{n \geq 0} B_n$. Then every A_n is a substructure of X; and any finite substructure of X is a substructure of B_n for some n, and hence is contained in $\mathcal{K}$ (since $B_n \in \mathcal{K}$ and $\mathcal{K}$ is substructure-closed). □

Remark. All conditions except the third are satisfied by the age of a relational structure of arbitrary cardinality. So, if the countability condition is satisfied, (and, in particular, if the language is finite), then there exists a countable structure with the same age. This justifies our restriction to countable structures.

3. First-order axioms

The theorem of Cantor alluded to above characterises $\boldsymbol{Q}$ as countable dense totally ordered set without endpoints. Thus $\boldsymbol{Q}$ is uniquely determined by the assumption of countability together with a collection of axioms, sentences (of first-order logic) in the language of ordered sets.

A structure X with this property is called $\aleph_0$-categorical. Thus the countable structure X is *$\aleph_0$-categorical* if there is a set Σ of sentences in the language of X such that any countable model of Σ is isomorphic to X. Note that countability has to be required as a 'nonlogical axiom' here: the upward Löwenheim-Skolem Theorem says that, if Σ has an infinite model, then it has arbitrarily large infinite models. See Chang and Keisler [7].

A remarkable characterisation of $\aleph_0$-categorical structures was found by Engeler [9], Ryll-Nardzewski [32] and Svenonius [35]. In this result, the language need not be assumed relational; functions and constants are permitted.

Theorem 3.1. *For the countable structure X, the following are equivalent:*

(i) *X is $\aleph_0$-categorical;*

(ii) *for every n, there are only finitely many n-types over* $\mathrm{Th}(X)$;

(iii) *for every n, the automorphism group of X has only finitely many orbits on X^n.*

Here an n-type is a maximal set of formulae in n variables consistent with $\mathrm{Th}(X)$, the theory of X (the set of all sentences true in X); that is, everything that can be said in first-order language about an n-tuple of points in a structure with the same theory. More is true: if X is $\aleph_0$-categorical, then every n-type over $\mathrm{Th}(X)$ is realised in X (i.e. there is an n-tuple in X satisfying it), and two n-tuples lie in the same orbit of $\mathrm{Aut}(X)$ if and only if they satisfy the same n-type.

The surprising fact is that axiomatisability of X is equivalent to the existence of a very large automorphism group. But perhaps, bearing in mind Felix Klein's philosophy (as expressed in the Erlanger Programm), this is not so unexpected.

As a simple consequence, we see that, if X is $\aleph_0$-categorical, then the number of n-element structures in $\mathrm{Age}(X)$ is finite for all n. So, if X has infinitely many

relations of arity at most n for some n, then most of them are 'redundant' (all are Boolean combinations of a finite number).

Problem. Find necessary and sufficient conditions for a class $\mathcal{H}$ to be the age of an $\aleph_0$-categorical structure. (The class of all finite graphs whose connected components are paths is not the age of any $\aleph_0$-categorical structure).

The theorem has a pretty corollary, which is presumably known to the experts, though I am not aware of a reference.

Theorem 3.2. *An $\aleph_0$-categorical structure is universal.*

Proof. Let X be $\aleph_0$-categorical, and let Y be younger than X; assume that the domain of Y is $\boldsymbol{N}$, the natural numbers. Let F_n be the substructure of Y on $\{0, \ldots, n-1\}$. We call two embeddings $f_1, f_2 : F_n \to X$ equivalent if there is an automorphism g of X such that $f_1 g = f_2$. By Theorem 3.1 and the fact that Y is younger than X, there are at least one but only finitely many equivalence classes for each n.

Construct a tree as follows: The nodes at level n are the equivalence classes of embeddings of F_n in X. Nodes on levels n and $n+1$ are adjacent in the tree if they have representatives f, f' respectively such that f is the restriction of f' to F_n.

König's Infinity Lemma [19] guarantees that the tree has an infinite path, starting at the root. It is easy to see that we can choose representatives of the nodes on this path consistently (that is, so that the embedding of F_n is the restriction of that of F_{n+1}). Now the union of all these embeddings is clearly a $1-1$ map from Y into X; and it is an isomorphism from Y onto a substructure of X, since an instance of a relation holds in Y if and only if it holds in F_n for some n. □

This is one generalisation of Cantor's Theorem. We will soon see another.

4. Homogeneity

The ordered set $\boldsymbol{Q}$ has a very strong symmetry property: whenever $x_1 < \cdots < x_n$ and $y_1 < \cdots < y_n$, there is an automorphism of $\boldsymbol{Q}$ carrying $(x_1, \ldots, x_n)$ to $(y_1, \ldots y_n)$. (Interpolate in the intervals (x_i, x_{i+1}) by linear functions, and at the two ends by appropriate shifts). Thus, $\boldsymbol{Q}$ is homogeneous, in the following sense: A structure X is *homogeneous* if, given any isomorphism $f : A \to B$ between finite substructures of X, there is an automorphism g of X whose restriction to A is f.

This is the strongest symmetry condition we can impose on a relational structure. Fraïssé [12] proved an important theorem which is both a characterisation and a construction of homogeneous structures, bearing a striking resemblance to Theorem 2.1.

A class $\mathcal{K}$ of finite structures has the *amalgamation property* if, whenever we have structures A, B_1, B_2, $C \in \mathcal{K}$ and $f_i : A \to B_i$ are embeddings ($i = 1, 2$), there exist $C \in \mathcal{K}$ and embeddings $g_i : B_i \to C$ ($i = 1, 2$) so that the diagram

$$\begin{tikzcd} A \arrow[r, "f_1"] \arrow[d, "f_2"'] & B_1 \arrow[d, "g_1"] \\ B_2 \arrow[r, "g_2"] & C \end{tikzcd}$$

commutes. In other words, any two structures with a common substructure can be jointly embedded so that their intersection contains at least this common substructure. By taking $A = \emptyset$, we see that the amalgamation property implies the joint embedding property. (This, however, depends on our conventions about the empty set.)

Theorem 4.1. *The class $\mathcal{K}$ of finite structures is the age of a countable homogeneous structure X if and only if $\mathcal{K}$ is closed under isomorphism, closed under taking substructures, contains only countably many non-isomorphic members, and has the amalgamation property. Moreover, if these conditions are fulfilled, then the countable homogeneous structure X is unique.*

The unique X is called the *Fraïssé limit* of the class $\mathcal{K}$.

For example, the class of all finite graphs satisfies the hypotheses; its Fraïssé limit was constructed explicitly by Rado [31], and implicitly earlier by Erdős and Rényi [11]. For reasons which will become clear, it is called the random graph. Henson [15] observed that, for any $n \geqslant 2$, the class $H(n)$ of graphs containing no complete subgraph of size n satisfies the hypotheses of Theorem 4.1. (If amalgamation is performed by adding no superfluous edges, then no new K_n will be constructed). The Fraïssé limits of these classes give further homogeneous graphs. Lachlan and Woodrow [22] proved that this is essentially all.

Theorem 4.2. *A countable homogeneous graph is one of the following:*

(a) *a disjoint union of m complete graphs of size n, where at least one of m and n is infinite;*

(b) *complement of* (a) (*a complete multipartite graph*);

(c) *the Fraïssé limit of the class of K_n-free graphs, $n \geqslant 3$;*

(d) *complement of* (c);

(e) *the random graph.*

This major theorem has been applied in Ramsey theory, see Nešetřil and Rödl [27]. Other classes of relational structures in which all homogeneous members

have been determined include tournaments (Lachlan [20]) and partial orders (Schmerl [33]).

A subsidiary result in the proof of Theorem 4.1 is useful in its own right.

Theorem 4.3. *The countable structure X is homogeneous if and only if it has the following property: if $A, B \in \mathrm{Age}(X)$ and $f: A \to B$, $h: A \to X$ are embeddings, then there is an embedding $g: B \to X$ such that $fg = h$. It suffices to require this when $|B| = |A| + 1$.*

If X is homogeneous and $\mathrm{Age}(X)$ has only finitely many n-element members for all n, then $\mathrm{Aut}(X)$ has only finitely many orbits on X^n for all n, since isomorphic structures lie in the same orbit. By Theorem 3.1, X is $\aleph_0$-categorical. But we can seen this more directly. Let X be a homogeneous structure with age $\mathscr{K}$. Then a countable structure Y is isomorphic to X if and only if it satisfies the following 'axioms':

(I) For each $A \in \mathscr{K}$, the statement that there exists a substructure of Y isomorphic to A. (This is a first-order existential sentence; the collection of such sentences ensures that $K \subseteq \mathrm{Age}(Y)$.)

(II) For each n, the statement that every n points of Y carry a substructure isomorphic to some member of $\mathscr{K}$. (This ensures that $\mathrm{Age}(Y) \subseteq \mathscr{K}$. It is not in general first-order; but if $\mathscr{K}$ has only finitely many n-element substructures, it is a disjunction over all of these, universally quantified).

(III) For each embedding $f: A \to B$ between members of $\mathscr{K}$, the statement of Theorem 4.3. (This is an $(\forall\exists)$ sentence, i.e. of the form $(\forall x_1 \cdots x_n)(\exists y)\phi$, where ϕ says 'if the substructure on $x_1, \ldots, x_n$ is A, then the substructure on $x_1, \ldots, x_n, y$ is B'.)

Note that (I) and (II) alone assert that $\mathrm{Age}(Y) = \mathscr{K}$.

We see that the axioms for a homogeneous $\aleph_0$-categorical structure can be taken to be $(\forall\exists)$ sentences (where one or the other of the types of quantifiers can be omitted). This implies, in particular, that such structures are model-complete (see [7]).

Another consequence of Theorem 4.3 is the following.

Theorem 4.4. *A homogeneous structure is universal.*

Proof. Let X be homogeneous, and Y younger than X, with the domain of Y equal to $\boldsymbol{N}$. For each n, let F_n be the substructure of Y on $\{0, \ldots, n-1\}$. By Theorem 4.3, any embedding of F_n in X can be extended to an embedding of F_{n+1}. Doing this for all n and taking the union gives an embedding of X. □

For more on homogeneity and universality, see the article by Komjáth and Pach in this collection.

One final remark. If X is homogeneous, then everything about the embedding of a finite structure A in X (in particular, any first-order formula satisfied by the points in its image) is determined by the isomorphism type of A. A theory is said to admit *quantifier-elimination* if any formula is equivalent, modulo the theory, to a formula without quantifiers. This suggests a relation between homogeneity and quantifier-elimination which is made precise in the following result.

Theorem 4.5. *A theory Σ admits quantifier-elimination if and only if any model of Σ has an elementary embedding into a homogeneous model. In particular, a $\aleph_0$-categorical theory admits quantifier-elimination if and only if its unique countable model is homogeneous.*

A slight variant on the amalgamation property has been studied. The *strong amalgamation property* is satisfied by $\mathcal{K}$ if, in the conclusion of the amalgamation property, we can assume that no unexpected identifications are made; that is, if $b_1 \in B_1$, $b_2 \in B_2$ satisfy $b_1 g_1 = b_2 g_2$, then there exists $a \in A$ such that $af_1 = b_1$ and $af_2 = b_2$.

Theorem 4.6. *Let X be a countable homogeneous structure. Then* $\mathrm{Age}(X)$ *has the strong amalgamation property if and only if the subgroup of* $\mathrm{Aut}(X)$ *fixing any finite tuple of points has no additional fixed points.*

The condition holds in $\boldsymbol{Q}$, for example: the orbits of the stabiliser of a tuple are the open intervals into which $\boldsymbol{Q}$ is divided by the points of the tuple. For an application of this concept (in mathematical psychology), see Cameron [2]. Proofs of the theorems in this section appear in the forthcoming book [3].

5. Absolute ubiquity

Informally, I will refer to a structure X as 'ubiquitous' if 'most' (in some sense) of the structures younger than X are actually isomorphic to X. It is quite remarkable, both that this often occurs, and that the property correlates well with other properties of interest. In particular, X is *absolutely ubiquitous* if every structure having the same age as X is isomorphic to X. We observed earlier that the complete graph is absolutely ubiquitous.

Conditions (I) and (II) preceding Theorem 4.4 ensure that a structure Y has prescribed age $\mathcal{K}$. Over a finite relational language, these conditions are first-order, as we noted. Now X is absolutely ubiquitous if and only if they (together with the assumption of countability) characterise X up to isomorphism. So, over a finite relational language, an absolutely ubiquitous structure is $\aleph_0$-categorical, and so has a large group of automorphisms. Just how large was made clear by Hodkinson and Macpherson [16], who gave what amounts to a complete description of these structures.

Theorem 5.1. *Let X be absolutely ubiquitous over a finite relational language. Then there is a partition of X into finitely many parts $Y_1, \ldots, Y_m$ such that the direct product of the symmetric groups on $Y_1, \ldots, Y_n$ is a group of automorphisms of X.*

Some of the parts may be finite. The content of the theorem is that whether or not an n-tuple satisfies a relational depends only on the parts of the partition containing its points and which pairs of points are equal. So a finite amount of information specifies the structure. The full automorphism group may be larger; consider, for example, a complete bipartite graph. A fairly typical absolutely ubiquitous graph is the disjoint union of an infinite complete bipartite graph, an infinite star, and a finite graph.

It follows easily from Theorem 5.1 that, if X is a relational structure over a finite language with the property that every structure younger than X is isomorphic to X, then X is trivial (i.e. $\mathrm{Aut}(X)$ is the symmetric group).

We could also consider properties $\mathscr{P}$ of X for which every structure with the same age as X satisfies $\mathscr{P}$. (Such properties might be called 'absolutely forced', but this term has a different meaning in logic.) The isomorphism type of an absolutely ubiquitous structure satisfies this condition. Clearly, any property of X expressed by a universal or existential sentence satisfies the condition; for example, the properties of being a graph, or of having no triangles.

6. The tree of an age

To investigate ages further, we provide them with some structure.

We take the 'normalised' version of $\mathrm{Age}(X)$, where the domain of any member consists of the first n natural numbers. We make $\mathrm{Age}(X)$ into a tree in the obvious way: nodes at level n are the n-element structures; nodes F_n, F_{n+1} at levels n, $n+1$ are adjacent if the restriction of F_{n+1} to $\{0, \ldots, n-1\}$ is F_n.

Now let $\mathrm{Younger}(X)$ be the set of structures than X, similarly 'normalised' so that their domains are all $\boldsymbol{N}$. Any structure $Y \in \mathrm{Younger}(X)$ defines a path in the tree $\mathrm{Age}(X)$, whose node at level n is the substructure of Y on $\{0, \ldots, n-1\}$. Conversely, any path in $\mathrm{Age}(X)$ specifies uniquely a member of $\mathrm{Younger}(X)$. Thus $\mathrm{Younger}(X)$ is identified with the set of infinite paths in the tree $\mathrm{Age}(X)$.

In the next two sections, we explore two ways of explicating the idea of 'most members of $\mathrm{Younger}(X)$', involving Baire category and measure respectively.

7. Baire category

Think of following a path in $\mathrm{Age}(X)$ as successively specifying more and more information about an infinite structure. Two paths are close together if a lot of

information is required to distinguish them. Accordingly, define a metric on Younger(X) (regarded as a set of paths rather than a set of structures) by the rules

$$d(Y, Y) = 0;$$

$$d(Y_1, Y_2) = \tfrac{1}{2}^n \quad \text{if the paths } Y_1, Y_2 \text{ agree to level } n \text{ but no further.}$$

(The actual choice of values is unimportant; any decreasing sequence of positive numbers tending to 0 would do.) The metric space axioms are easily verified; indeed, in place of the triangle inequality, we have the stronger *ultrametric inequality*

$$d(Y_1, Y_3) \leqslant \max\{d(Y_1, Y_2), (Y_2, Y_3)\},$$

which implies, that, if two open balls meet, then one contains the other.

This metric space is complete: a Cauchy sequence of paths has the property that its members agree on longer and longer initial segments; these initial segments define a path, which is the limit of the sequence.

The set of paths containing a fixed node F is an open ball, and every open ball has this form. Thus, a set S of paths is open if and only if, for every path $Y \in S$, there is a node F on Y with the property that every path containing F is in S. Similarly, a set S of paths is dense if and only if every node lies on a path in S.

A subset of a metric space is *residual* (or *comeagre*, or *the complement of a set of the first category*) if it contains a countable intersection of open dense sets. Residual sets are regarded as 'large'. The Baire category theorem asserts that, in a complete metric space, a residual set is non-empty. Of course, this immediately implies the apparently stronger statement that the intersection of countably many residual sets is non-empty.

Proving the Baire category theorem in our special case is much simpler than in the general case. Let $S_0, S_1, \ldots$ be open and dense. Define a sequence of nodes $F_0, F_1, \ldots$ by the rules: $F_0 = \emptyset$; F_{n+1} is a node above F_n with the property that every path through F_{n+1} lies in S_n (this exists because S_n is open and dense). Then the unique path containing all these nodes lies in $\bigcap_{n \geqslant 0} S_n$.

A slight modification of the argument shows that a residual set is dense. Given any node F, replace the start of the above recursion with $F_0 = F$; the resulting path passes through F.

Now we say that a structure X is *ubiquitous in category* if the set of structures isomorphic to X is residual in Younger(X); and a property $\mathscr{P}$ of X is *forced in category* if a residual subset of Younger(X) satisfies $\mathscr{P}$. (This relation between the theory of X and first-order sentences forced in category is known to logicians as finite forcing, following Abraham Robinson; but the present name seems more appropriate here.)

Theorem 7.1. (i) *Any* ($\forall\exists$) *sentence true in X is forced in category by X.*

(ii) *A homogeneous structure is ubiquitous in category.*

Proof. (i) Let the sentence ϕ be true in X, where ϕ is expressed in the form $(\forall x_1 \cdots x_n)(\exists y_1 \cdots y)\theta(x_1, \ldots, x_n, \quad y_1, \ldots, y_m)$. (For short we write $(\forall \bar{x})(\exists \bar{y})\theta(\bar{x}, \bar{y})$.) The set of structures in Younger(X) satisfying it is the intersection, over all choices of $a_1, \ldots a_n \in \boldsymbol{N}$, of the sets $S(\bar{a}) = S(a_1, \ldots, a_n)$ of structures satisfying $(\exists \bar{y})\theta(\bar{a}, \bar{y})$. So it suffices to show that $S(\bar{a})$ is open and dense. The openness is clear: if $Y \in S(\bar{a})$, so that $\theta(\bar{a}, \bar{b})$ holds in Y for some $\bar{b}(b_1, \ldots, b_m)$, and $k = \max\{\bar{a}, \bar{b}\}$, then any structure whose restriction to $\{0, \ldots, k\}$ agrees with that of Y is also in $S(\bar{a})$. To show denseness, suppose that the substructure F on $\{0, \ldots, k\}$ has been defined for some k, and that $a_1, \ldots, a_n \in F$. Find a copy F' of F inside X; let $a'_1, \ldots, a'_n$ be the corresponding points of X. Since ϕ holds in X, there exist $b'_1, \ldots, b'_m \in X$ such that $\theta(\bar{a}', \bar{b}')$ holds. Then extend F to a structure F'' on an initial segment of $\boldsymbol{N}$ isomorphic to $F|_{\{a'_1, \ldots, a'_n, b'_1, \ldots, b'_m\}}$. Any infinite structure extending F'' is in $S(\bar{a})$.

(ii) We observed in Section 4 that conditions (I)–(III) before Theorem 4.4 characterise the homogeneous structure X. Of these, (I) and (III) comprise countably many $(\forall\exists)$ sentences. (II), which may not even be first-order, simply ensures that $Y \in$ Younger(X), and so is not required here. Thus the isomorphism type of X is a countable intersection of residual sets in Younger(X), and so is residual. □

It follows from Theorem 7.1 that, for any structure X, almost all (i.e. a residual set of) structures younger than X actually have the same age as X, since this is ensured by condition (I) alone.

This is not the end of the story, however. Consider the two-way infinite path X. Age(X) consists of all finite unions of finite paths. It is easily checked that the following property is forced in category by X: *Any finite set is contained in the interior of a path.* Moreover, this property characterises X within Younger(X). So X is ubiquitous in category, though it is certainly not either homogeneous or $\aleph_0$-categorical. The above property, though not first-order, has the general flavour of a $(\forall\exists)$ sentence.

This example is instructive. There are two 'distinguished' structures having the same age: a single infinite path, and the disjoint union of countably many infinite paths. The first is ubiquitous in category, the second universal. They also have contrasting model-theoretic properties. (They have the same theory; they are the prime and saturated countable models of this theory respectively.)

8. Measure

In this section, I consider an alternative interpretation of 'most', using the notion of measure, which is closely related and in a sense dual to that of category (see Oxtoby [29]). We assign a positive real number $m(F)$ to each node F of the tree Age(X), this number being interpreted as the measure of the open ball

consisting of paths containing F. We must require that $m(F)$ is equal to the sum of the values of $m(F')$ over all nodes F' immediately above F. (For convenience I assume in this section that the number of n-element structures in Age(X) is finite for all n; this ensures that the sum in question is finite). Any function m satisfying these conditions defines in a standard way a measure on the σ-algebra of Borel subsets of Younger(X).

If the measure of the whole space (i.e. $m(\emptyset)$) is 1, as it will be in what follows, then m is a probability measure. The interpretation is that, in a random member of Younger(X), the probability that the substructure induced on $\{0, \ldots, n-1\}$ is F is equal to $m(F)$.

There are many functions satisfying this requirement. For example, we may take $m(\emptyset) = 1$, and then divide the measure of any node F equally among its children. However, a further condition is desirable:

(+) For any $F \in \mathrm{Age}(X)$ with $|F| = n$, and any distinct $x_0, \ldots, x_{n-1} \in \boldsymbol{N}$, the measure of the set $\{Y \in \mathrm{Younger}(X)$: the map $i \mapsto x_i$ $(i = 0, \ldots, n-1)$ is an embedding of F in $Y\}$ is equal to $m(F)$.

This condition is equivalent to the requirement that if F is isomorphic to F' then $m(F) = m(F')$. It is not satisfied by the measure defined above by equal division. Consider, for example, the age $H(3)$ consisting of all triangle-free graphs. The graph with vertex set $\{0, 1, 2\}$ and edges $\{0, 2\}$ and $\{1, 2\}$ has measure $\frac{1}{8}$, while that with edges $\{0, 1\}$ and $\{0, 2\}$ has measure $\frac{1}{6}$.

There is, however, a 'natural' way to attempt to define a measure, which (if it succeeds) ensures that (+) holds. This is motivated by the interpretation of probability as limiting frequency. Given a node F in Age(X), we define

$$m(F) = \lim_{n \to \infty} \frac{\text{number of nodes at level } n \text{ above } F}{\text{number of nodes at level } n}.$$

Informally, the probability that the substructure on $\{0, \ldots, n-1\}$ is F is equal to the limiting frequency of F among all n-element substructures of large labelled structures in Age(X). For example, if X is a graph and F an edge, $m(F)$ is the limiting average density of edges in large labelled subgraphs of X.

Problem. Does this limit exist for any choice of the relational structure X (over a finite language) and any finite substructure F of X?

(Of course, it is easy to construct trees in which the limit does not exist. But, for the comparatively few ages in which I have been able to decide the question, I have found no counterexamples.)

If $m(F)$, as defined above, exists for all $F \in \mathrm{Age}(X)$, it is straightforward to show that it is nonnegative, is additive over children of any node, and satisfies (+); so it defines a measure on Younger(X) satisfying all our requirements. Now, much as before, we say that a property $\mathcal{P}$ of X is *forced in measure* if it holds on a

set of measure 1 (the complement of a null set), and that X is *ubiquitous in measure* if its isomorphism type is forced in measure.

In the case where Age(X) consists of all finite graphs, the measure is precisely that defined by the informal statement 'choose edges independently with probability $\frac{1}{2}$'. Now the result of Erdős and Rényi [11] alluded to earlier is that (in the terminology introduced earlier) the random graph is ubiquitous in measure—hence its name. The ordered set $\boldsymbol{Q}$ is also ubiquitous in measure. This can be derived from a general result.

Theorem 8.1. *Let X be a countable homogeneous structure whose age has the following property: there is a function f on the natural numbers such that any n-element structure* Age(X) *is contained in exactly $f(n)$ $(n+1)$-element structures. Suppose further that $\sum 1/f(n)$ diverges. Then X is ubiquitous in measure.*

For $\boldsymbol{Q}$, the hypotheses of this theorem hold, with $f(n) = n + 1$. Ubiquity in measure has some pleasant consequences.

Theorem 8.2. *If X is ubiquitous in measure, then the substructure of X induced on any cofinite subset is isomorphic to X.*

This is proved by showing that the induced measure on the complement of a finite set corresponds to the original measure for structures on that set. The result shows that a structure ubiquitous in measure is sufficiently tough that removal of finitely many points leaves it unchanged. It is very likely true that ubiquity in measure implies other similar toughness results, perhaps even the existence of automorphisms; but little work has been done on this.

There are, however, some surprises. One remarkable fact is that the age $H(3)$ consisting of the finite triangle-free graphs forces in measure the property of being bipartite. (This is the content of a theorem of Erdös, Kleitman and Rothschild [10], according to which the proportion of triangle-free graphs on n vertices which are nonbipartite tends to 0 as $n \to \infty$. From this it follows that $m(F) = 0$ for any nonbipartite F, i.e. the probability that the subgraph on a given finite set is nonbipartite is equal to zero. Since there are only countably many finite subsets, and since a countable union of null sets is null, the random triangle-free graph almost surely has no nonbipartite finite subgraph).

It can be shown that the measure associated with $H(3)$ (which, from the above, is the same as that associated with the age consisting of all finite bipartite graphs) has the following simple description: Given a finite set F, divide it into two subsets by assigning points to one or the other independently with probability $\frac{1}{2}$; then choose edges independently with probability $\frac{1}{2}$ from the pairs with one point in each subset.

There is a countable bipartite graph B which is forced in measure by either of the ages described. (In particular, B is ubiquitous in measure). B is *almost*

homogeneous: any isomorphism between finite subgraphs which respects the bipartition in B is induced by an automorphism of B. B is also ubiquitous in category.

By contrast, Henson's homogeneous triangle-free graph is ubiquitous in category but not in measure.

9. Profiles

The profile of an age is the function on $\boldsymbol{N}$ which enumerates the isomorphism types of n-element structures in the age for each n. Throughout this section, we suppose that all values of the profile are finite (this holds, in particular, if the language is finite). The principal question is: which functions from $\boldsymbol{N}$ to $\boldsymbol{N}$ occur as profiles? Two specialisations of this question restrict attention to homogeneous structures, or structures over finite languages (for example, graphs).

If X is homogeneous, then the number of isomorphism types of n-element substructures of X is equal to the number of orbits of $\mathrm{Aut}(X)$ on n-element subsets of X.

The fundamental result is as follows.

Theorem 9.1. *The profile of an age is a nondecreasing function.*

Two entirely different proofs are known using linear algebra and Ramsey theory respectively. Both give additional information. For the first proof, let X be an infinite set, and k a positive integer. Let V_k denote the vector space of functions from the set of k-element subsets of X to $\boldsymbol{Q}$. For $k \leqslant l$, define a linear map θ from V_k to V_l as follows: for $f \in V_k$, $L \subseteq X$, $|L| = l$, set

$$(f\theta)(L) = \sum \{f(K) \colon K \subseteq L,\ |K| = k\}.$$

Theorem 9.2. *With the above notation,* $\mathrm{Ker}\,\theta = \{0\}$.

Now suppose that the numbers of isomorphism types of k- and l-element structures are n_k, n_l respectively. Let A be the $n_k \times n_l$ matrix whose (i, j) entry is the number of k-sets of the ith isomorphism type contained in an l-set of the jth isomorphism type. Then A represents the restriction of θ to the subspace spanned by the characteristic functions of the isomorphism types of k-sets; so A has rank n_k, and $n_l \geqslant n_k$. (In fact, this argument is valid also for finite sets X, provided that $|X| \geqslant k + l$.)

The second proof uses an extension of Ramsey's Theorem (which, however, is proved using Ramsey's Theorem).

Theorem 9.3. *Suppose that the k-subsets of the infinite set X are coloured with n colours, all of which are used. Then there exist an ordering $(c_1, \ldots, c_n)$ of the colours, and infinite subsets $X_1, \ldots, X_n$ of X, such that for each value of i, X_i contains a set of colour c_i but none of colour c_j for $j > i$.*

Now colour the k-subsets of X with $n = n_k$ colours according to their isomorphism types. Then, for $l \geqslant k$, at least n different l-sets can be distinguished by the isomorphism types of their k-subsets. This actually shows that the matrix A of the previous proof has its principal $n_k \times n_k$ submatrix upper triangular, if rows and columns are suitably ordered. (This proof also works for sufficiently large finite sets; but here, the lower bound obtained from the proof of a finite version of Theorem 9.3 is an iterated Ramsey number.)

How fast does the profile grow? The next result is an amalgam of results of Pouzet [30] and Macpherson [23].

Theorem 9.4. *Let f be the profile of a binary relational structure. Then either:*

(a) *f grows like a polynomial (that is, there exist $k \in \boldsymbol{N}$ and $c, d \in \boldsymbol{R}$ with $c > 0$, such that $cn^k \leqslant f(n) \leqslant dn^k$ for all n); or*

(b) *$f(n) \geqslant \exp(n^{1/2-\epsilon})$ for all sufficiently large n.*

All cases here occur. For a disjoint union of complete graphs of size $k+1$, the profile grows like a polynomial of degree k. For a disjoint union of infinitely many infinite complete graphs, the profile is the partition function, which is very roughly $\exp(n^{1/2})$ (see Hall [14]).

Pouzet showed that either (a) holds or f grows faster than any polynomial. His proof involves a delicate analysis of the sub-ages of an age, as partially ordered set. Macpherson showed that, in the latter case, it is possible to encode partitions of n into substructures of X of size cn for some fixed c, using Ramsey's Theorem.

Problem. Suppose that f is a profile which is *subexponential*, that is, $f(n) < \exp(n^{1-c})$ for sufficiently large n, where $c > 0$. Is it true that there exists $p \in \boldsymbol{N}$ such that, for any $\epsilon > 0$ and all sufficiently large n,

$$\exp(n^{p/(p+1)-\epsilon}) < f(n) < \exp(n^{p/(p+1)+\epsilon})?$$

(There are some reasons for thinking that such a strange spectrum of growth rates might be true, and examples having such growth are known for any $p \in \boldsymbol{N}$).

Considerably more detailed results are known for homogeneous structures. If X is homogeneous, then isomorphism types of n-subsets are orbits of $\mathrm{Aut}(X)$; our blanket assumption in this section just says that X is $\aleph_0$-categorical. On the other hand, given any permutation group G, there is a homogeneous relational structure X such that $G \leqslant \mathrm{Aut}(X)$ and G has the same orbits on X^n as $\mathrm{Aut}(X)$ (for all n). So results in the present case are equivalent to results about orbits of permutation groups, and are often phrased in this way.

We call X *transitive* if no nontrivial unary relation is definable (without parameters) in X, and *primitive* if no nontrivial equivalence relation is definable. (These conditions hold precisely when the permutation group $G = \mathrm{Aut}(X)$ is transitive, resp. primitive, in the usual sense). Macpherson [25] showed the following.

Theorem 9.5. *There is a constant $c > 1$ such that, if X is a primitive homogeneous structure with profile f, then either $f(n) = 1$ for all n, or $f(n) > c^n$ for all sufficiently large n.*

The proof is a considerable elaboration of his argument for Theorem 9.4, encoding trees rather than partitions. Examples where the profile grows no faster than exponentially are quite rare; most of them are 'treelike' objects (see [1]). It would be interesting to know more about the spectrum of possible 'exponential constants', i.e. values of $\lim_{n\to\infty} f(n)^{1/n}$, for profiles f of primitive homogeneous structures (and, in particular, whether this limit always exists).

Macpherson [26] has also found connections between growth rates faster than exponential for primitive homogeneous structures and model-theoretic properties such as stability.

The growth rate of the profile can be pushed up artificially by having relations in the language of arbitrarily large arity. However, for finite relational languages, there is an upper bound.

Theorem 9.6. *If X is a structure over a finite relational language L, then the profile of X is bounded by the exponential of a polynomial depending only on L.*

This follows by observing that, if $m_1, \ldots, m_k$ are the arities of the relations in L, then the total number of n-element L-structures is

$$2^{n^{m_1}+\ldots+n^{m_k}}.$$

Note that, by Fraïssé's Theorem 4.1, this bound is attained by a homogeneous structure over L.

There are structures with growth rates intermediate between these levels. Two examples follow.

(i) Consider the age consisting of finite sets carrying two independent total orders. (There is a homogeneous structure with this age.) The profile is the factorial function.

(ii) The class of finite line graphs is the age of the line graph of the infinite complete graph. Its profile grows a little faster than exponentially, but the precise asymptotic behaviour, surprisingly, seems to be unknown (despite detailed analysis by Wright [38], among others).

10. Further directions

One of the most important recent developments, on which I have touched only incidentally, is the connection with a class of model-theoretic concepts which had been introduced in connection with the question: given a theory, how many models of given infinite cardinality can it have? These concepts, such as stability, the strict order property, and the independence property, can be found in Shelah [34]. Stable $\aleph_0$-categorical infinite structures can be regarded as 'limits' of finite ones, and their behaviour is closely controlled by their finite substructures. One of the most striking results is that of Lachlan [21], which gives a kind of classification of all finite or stable infinite homogeneous structures over any finite relational language. The ω-stable, $\aleph_0$-categorical structures are quite well understood, due to work of Zil'ber [39], Cherlin, Harrington and Lachlan [8], and Hrushovski [17].

For structures of large infinite cardinality, the profile can be extended to infinite cardinal numbers. The second proof of Theorem 9.1 shows immediately that the number of countable structures is at least as great as the number of n-element substructures (if this is finite). Results on higher cardinalities have been obtained by Kierstead and Niykos, Pouzet and Woodrow [37], and (for orbits of permutation groups) by Neumann [28]. But the best result, due to Macpherson, Mekler and Shelah [26], extends the monotonicity of the profile to infinite cardinal numbers sufficiently much smaller than $|X|$. It is thought that the profile is nondecreasing for all cardinals below $|X|$. This is known to fail at $|X|$, however. Consider $X = \omega_1$, the first uncountable ordinal. The substructures of X are all the ordinals up to ω_1. There are thus uncountably many countable substructures (up to isomorphism), but only one of size $\aleph_1$.

The concepts of Baire category and measure, which were little more than descriptive tools in this paper, have been used in their more traditional roles in existence proofs. (To show that some object has property $\mathscr{P}$, show that almost all do.) These proofs mainly involve showing that certain structures admit groups of automorphisms with certain properties. Here is a single example. Among countable graphs which admit a cyclic automorphism, almost all (a residual set of measure 1) are isomorphic to the random graph R. Hence R admits cyclic automorphisms, and indeed admits $2^{\aleph_0}$ nonconjugate cyclic automorphisms. See Cameron and Johnson [4] for further results in this direction.

The generating function of a profile turns out to be a specialisation of a Pólya-type *cycle index function* associated with the age. Another specialisation of the same series is the exponential generating function for labelled structures in the age. In the case of a homogeneous structure X, this cycle index function can be associated with its automorphism group; the two specialisations mentioned count orbits of this group on n-sets and on n-tuples of distinct points respectively. There are rules for computing the cycle indices of direct and wreath products and of point stabilisers, and links with many topics in enumerative combinatorics (e.g.

Cameron and Taylor [5]). The cycle index itself can be regarded as a kind of specialisation of a 'combinatorial formal power series' in the sense of Joyal [18] (who would regard an age as a *species*).

The philosophy here is to approximate infinite structures by finite ones. This can be extended to approximate automorphisms of infinite structures by partial automorphisms of finite ones. For example, Truss (see [36] for a survey) has shown that in many cases (for example, the random graph), there is a 'generic' or residual conjugacy class of automorphisms, which have infinitely many cycles of each finite length (but, surprisingly, no infinite cycles). Further connections with permutation groups will appear in the forthcoming book [3].

Finally, I refer to Fraïssé's book [13] on the theory of relations, which contains a treasure trove of information in many areas related to the topic of this survey.

References

[1] P.J. Cameron, Some treelike objects, Quart. J. Math. Oxford Ser. (2) 38 (1987) 155–183.

[2] P.J. Cameron, Groups of order-automorphisms of the rational numbers with prescribed scale type, J. Math. Psych. 33 (1989) 163–171.

[3] P.J. Cameron, Oligomorphic Permutation Groups, London Math. Soc. Lecture Notes (Cambridge Univ. Press, Cambridge, 1991).

[4] P.J. Cameron and K.W. Johnson, An investigation of countable Burnside groups, Math. Proc. Cambridge Philos. Soc. 102 (1987) 223–231.

[5] P.J. Cameron and D.E. Taylor, Stirling numbers and affine equivalence, Ars Combin. 20 (B) (1985) 3–14.

[6] G. Cantor, Beiträge zur Begründung der transfiniten Mengenlehre, Math. Ann. 46 (1895) 481–512.

[7] C.C. Chang and H.J. Keisler, Model Theory (North-Holland, Amsterdam, 1973).

[8] G.L. Cherlin, L.A. Harrington and A.H. Lachlan, $\aleph_0$-categorical, $\aleph_0$-stable structures, Ann. Pure Appl. Logic 28 (1985) 103–135.

[9] E. Engeler, Äquivalenzklassen von n-Tupeln, Z. Math. Logik Grundlag. Math. 5 (1959) 126–131.

[10] P. Erdős, D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, Colloq. Internat. Teorie Combin. II (1977) 19–27.

[11] P. Erdős and A. Rényi, Asymmetric graphs, Acta Math. Acad. Sci. Hungar. 14 (1963) 295–315.

[12] R. Fraïssé, Sur certains relations qui généralisent l'ordre des nombres rationnels, C. R. Acad. Sci. Paris 237 (1953) 540–542.

[13] R. Fraïssé, Theory of Relations (North-Holland, Amsterdam, 1986).

[14] M. Hall Jr, Combinatorial Theory (Blaisdell, Waltham, MA 1967).

[15] C.W. Henson, A family of countable homogeneous graphs, Pacific J. Math. 38 (1971) 69–83.

[16] I.M. Hodkinson and H.D. Macpherson, Relational structures determined by their finite induced substructures, J. Symbolic Logic 53 (1988) 222–230.

[17] E. Hrushovski, to appear.

[18] A. Joyal, Une théorie combinatoire des series formelles, Adv. in Math. 42 (1981) 1–82.

[19] D. König, Theorie der Endlichen und Unendlichen Graphen (Chelsea, New York, 1950).

[20] A.H. Lachlan, Countable homogeneous tournaments, Trans. Amer. Math. Soc. 284 (1984) 431–461.

[21] A.H. Lachlan, On countable stable structures which are homogeneous for a finite relational language, Israel J. Math. 49 (1984) 69–153.

[22] A.H. Lachlan and R.E. Woodrow, Countable ultrahomogeneous undirected graphs, Trans. Amer. Math. Soc. 262 (1980) 51–94.

[23] H.D. Macpherson, Growth rates in infinite graphs and permutation groups, Proc. London Math. Soc (3) 51 (1985) 285–294.
[24] H.D. Macpherson, Orbits of infinite permutation groups, Proc. London Math. Soc. (3) 51 (1985) 246–284.
[25] H.D. Macpherson, Infinite permutation groups of rapid growth, J. London Math. Soc. (2) 35 (1987) 276–286.
[26] H.D. Macpherson, personal communication, 1989.
[27] J. Nešetřil, personal communication, 1988.
[28] P.M. Neumann, Homogeneity of infinite permutation groups, Bull. London Math. Soc. 20 (1988) 305–312.
[29] J.C. Oxtoby, Measure and Category (Springer, Berlin, 1971).
[30] M. Pouzet, Application d'une propriété combinatoire des partes d'un ensemble aux groupes et aux relations, Math. Z. 150 (1976) 117–134.
[31] R. Rado, Universal graphs and universal functions, Acta Arith. 9 (1964) 393–407.
[32] C. Ryll-Nardzewski, On the categoricity in power $\aleph_0$, Bull. Acad. Polon. Sci. Ser. Math. 7 (1959) 545–548.
[33] J. Schmerl, Countable homogeneous partially ordered sets, Algebra Universalis 9 (1979) 317–321.
[34] S. Shelah, Classification Theory and the Number of Non-isomorphic Models (North-Holland, Amsterdam, 1978).
[35] L. Svenonius, $\aleph_0$-categoricity in first-order predicate calculus, Theoria 25 (1959) 82–94.
[36] J.K. Truss, Infinite simple permutation groups—a survey, to appear.
[37] R.E. Woodrow, personal communication, 1986.
[38] E.M. Wright, The number of connected sparsely-edged graphs, III: Asymptotic results, J. Graph Theory 4 (1980) 393–407.
[39] B.I. Zil'ber, The structure of models of uncountably categorical theories, Internat. Congr. Math. 1 (1984) 359–368.

Discrete Mathematics 95 (1991) 69–89
North-Holland

Decomposing infinite graphs

Reinhard Diestel

Faculty of Mathematics (SFB 343), Bielefeld University, Postfach 8640, 4800 Bielefeld, Germany

Received 20 October 1989

Abstract

Diestel, R., Decomposing infinite graphs, Discrete Mathematics 95 (1991) 69–89.

This paper gives an introduction to the theory of simplicial and related decompositions of graphs as developed in [1]. It is intended for the non-specialist, and particular prominence is given to the presentation of open problems.

Introduction

In his classic paper *Über eine Eigenschaft der ebenen Komplexe,* Wagner [19] tackles the following problem. Kuratowski's theorem, in its excluded minor version, states that a finite graph is planar if and only if it has no minor isomorphic to K_5 or to $K_{3,3}$. (A *minor* of G is any graph obtained from some $H \subset G$ by contracting connected subgraphs.) If we exclude only one of these two minors, the graph may no longer be planar—but will it be *very* different from a planar graph? For example, can the non-planarity of an arbitrary finite graph without a K_5 minor be tied down to certain parts of it, the rest of the graph being planar?

Wagner's solution to this problem is based on the following observation. Suppose we take two graphs G_1 and G_2, neither of which has a K_5 minor, and paste them together along a complete subgraph. (Following Wagner, we shall use the term *simplex* for complete graphs. So here we let $G = G_1 \cup G_2$ and assume that $G_1 \cap G_2$ is a simplex.) Then the resulting graph G is again K_5-free (has no K_5 minor). For if $H_1, \ldots, H_5$ are connected subgraphs of some $H \subset G$ whose contraction yields a K_5, then either G_1 or G_2 must also have such subgraphs (Fig. 1), contrary to our assumption that these graphs are K_5-free.

Repeating this process, we can easily construct *simplicial decompositions* of arbitrarily large K_5-free graphs: just keep attaching new K_5-free graphs along simplices contained in the graph constructed so far (Fig. 2). And more importantly, the converse is also true: every K_5-free graph, however large, can be constructed in such a simplical decomposition from *prime* factors, i.e. from 'small'

Elsevier Science Publishers B.V.

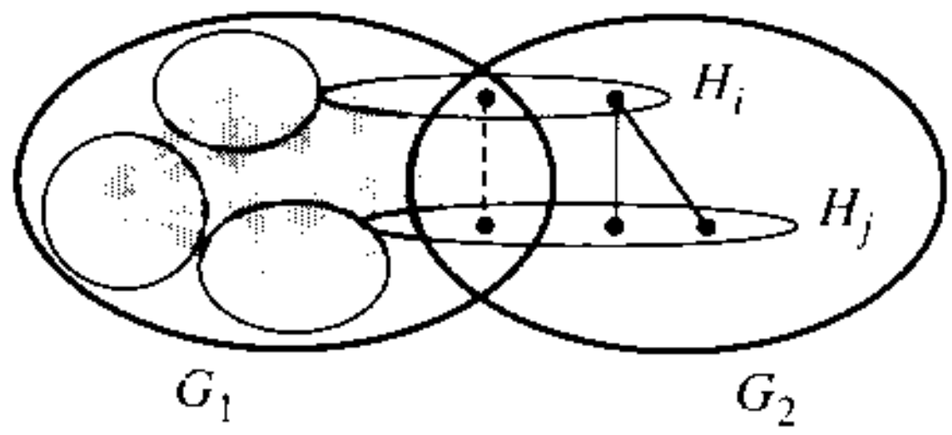


Fig. 1. Finding a K_5 minor in G_1.

graphs which do not themselves have a simplicial decomposition into more than one factor. (The general question of which graphs admit a prime decomposition is a fundamental problem in simplical decomposition theory; see Section 1.) Thus, we can characterize the finite K_5-free graphs by their decompositions if we succeed in drawing up a complete list of the prime factors needed to construct all these graphs.

Essentially, this is just what Wagner does in his paper—the only difference being that he keeps the list shorter by including only the factors of *edge-maximal* K_5-free graphs, those in which the addition of any new edge creates a minor isomorphic to K_5. As it turns out, this list, which he calls the 'homomorphism base' of K_5, contains only one non-planar graph W, while all its other graphs are planar (namely, the 4-connected plane triangulations). (More generally, the *homomorphism base* of a finite graph X is the class of all the graphs occurring as factors in prime decompositions of—finite or infinite—edge-maximal X-free graphs.) Our somewhat vague opening question thus has a surprisingly positive answer: the non-planarity of any finite K_5-free graph can be localized within parts of it that are either subgraphs of the one non-planar graph W, or else arise from at least three planar graphs pasted together along a triangle. (Note that pasting planar factors together along a simplex smaller than K_3 still yields a planar graph.)

Since Wagners original paper, homomorphism bases have been determined for

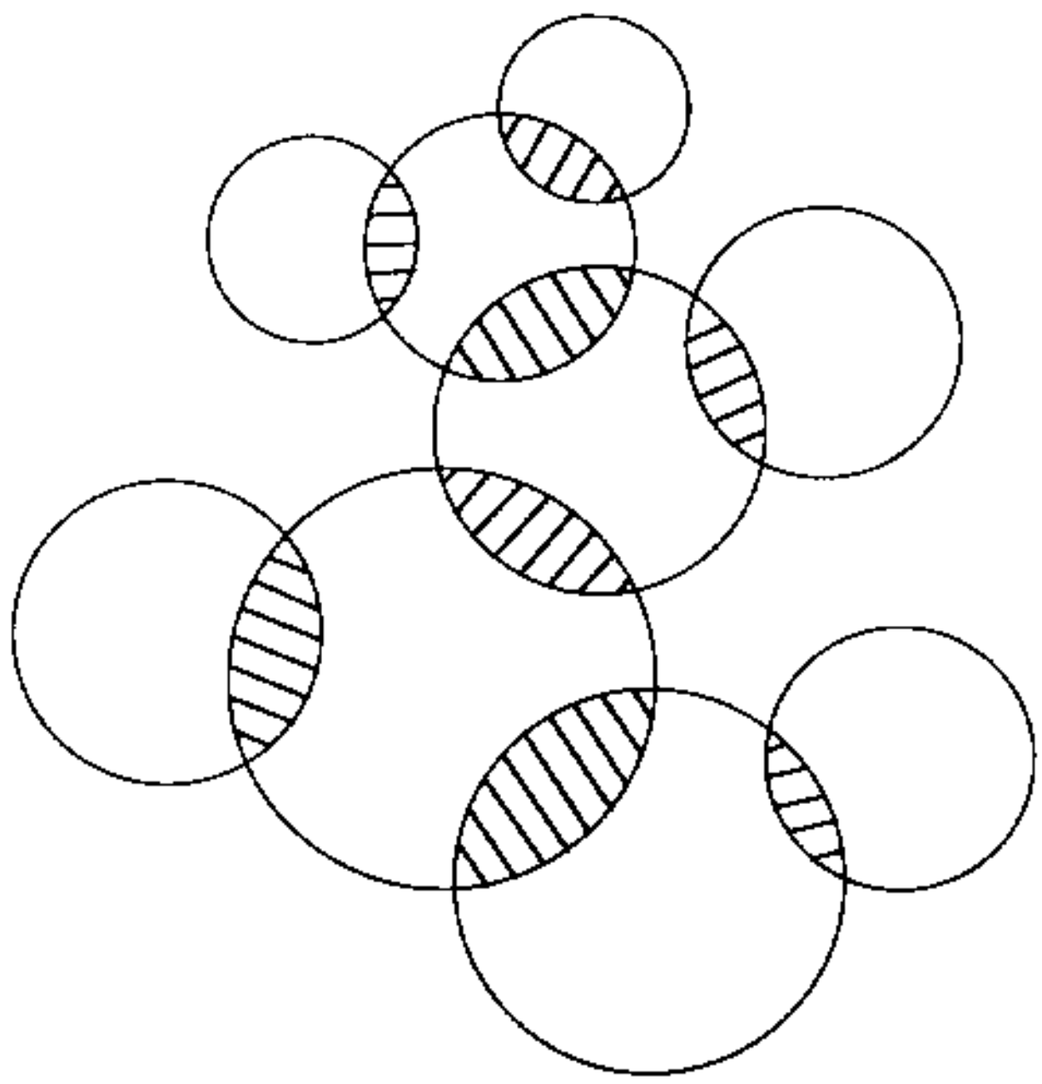


Fig. 2. A simplicial decomposition.

several other excluded minors, in each case giving rise to a similar structural characterization of the graphs without this minor (see Wagner [19–21], Halin [10–11, 13–14], or [1, Ch. 6.1] or [7] for a table of all known homomorphism bases.) A typical example: the edge-maximal K_5^--free graphs (where K_5^- denotes a K_5 minus an edge) are precisely the graphs which can be constructed in a simplicial decomposition with attachment simplices of order 2 from factors isomorphic to K_3, $K_{3,3}$, the prism ($K_3 \times K_2$) or wheels.

It is clear that excluded minor theorems in terms of homomorphism bases can be very powerful characterizations. In the case of K_5^-, for instance, the simplicity of the prime factors of the K_5^--free graphs enables us instantly to determine sharp bounds on their chromatic number, minimal degree and so on. Moreover, the graph properties definable by the exclusion of minors are important properties: they are precisely the properties that are closed under subcontraction (= taking minors), and include such natural properties as, say, the embeddability in a given surface.

However, not every homomorphism base offers as much information as does that of K_5^-. In the case of K_5, for example, we know that the base elements, with the exception of the graph W, are maximally planar. But how well do we really *know* an arbitrary maximally planar graph? We can hardly determine its chromatic number!

An important problem, therefore, is to learn to distinguish the minors whose exclusion gives rise to a simple homomorphism base from those where the base elements can be nearly as complicated as the graphs they serve to describe. At first glance, this notion of a 'simple' homomorphism base seems a difficult one to make precise. However, it so happens that as soon as we allow our graphs to be infinite, the simple and the complicated bases seem to fall neatly apart: into bases which are made up of finite graphs only (and are therefore countable, like the homomorphism base of K_5^-) and uncountable bases (like that of K_5, which contains all the—uncountably many—countable maximally planar graphs). By a beautiful theorem of Halin, homomorphism base elements are always themselves countable, whatever the cardinality of the graphs of which they are factors; see [1, Ch. 5].

Calling a homomorphism base *simple* if it is countable—or, alternatively, if all its members are finite—we are thus led to the following problem. (It is unknown whether the two suggested definitions of 'simple' coincide.)

Problem. For which excluded minors is the corresponding homomorphism base simple?

Although this problem in its full generality seems to be hard, so little is known about it that even the most basic results would mean progress. For example, if the homomorphism base of X is simple and X' is obtained from X by deleting an

edge, is the base of X' again simple? More such conjectures, including some farther reaching ones, can be found in [1, Ch. 6.1] or in [7].

Since their introduction by Wagner for the purpose of investigating the K_5-free graphs, simplicial decompositions have been applied to a wide range of problems, mainly in infinite graph theory. Moreover, the investigation of these and related decompositions has led to an interesting theory in its own right. The aim of this paper is to give an introduction to some of the central aspects of this theory, to state its main results (without proofs, but illustrated by examples), and to present its guiding open problems.

Sections 1 and 2 deal with the problem of the existence of simplicial decompositions into prime or otherwise 'small' factors. Section 3 gives a brief introduction to the problem of when such prime decompositions are unique. In Section 4 finally, we look at the 'structural essence' of simplical decompositions, which is neatly captured by another type of decomposition called *tree-decompositions*. For finite graphs, these tree-decompositions reduce to the by now familiar decompositions used by Robertson and Seymour for the proof of their well-quasi-ordering theorem (Wagner's Conjecture).

1. The existence of prime decompositions

The question of which graphs admit a simplicial decomposition into prime factors, already touched upon above, is perhaps the most fundamental and at the same time the most complex problem in simplicial decomposition theory. And while a good deal is now known about prime decompositions, existing results amount to no more than a partial solution of the general problem:

Problem. Which graphs admit a simplicial decomposition into primes?

Before we look into this problem further, let us give a precise definition of a simplicial decomposition. In order to make the definition suitable for infinite as well as for finite graphs, we do not follow the intuitive approach of 'decomposing' a graph into smaller and smaller pieces (a process which may never end), but build it up from below, adding one factor at a time.

Thus, let G be a graph, $\sigma > 0$ an ordinal, and let B_λ be an induced subgraph of G for every $\lambda < \sigma$. The family $F = (B_\lambda)_{\lambda<\sigma}$ is called a *simplicial decomposition* of G if the following three conditions hold:

(S1) $G = \bigcup_{\lambda<\sigma} B_\lambda$;
(S2) $(\bigcup_{\lambda<\mu} B_\lambda) \cap B_\mu =: S_\mu$ is a complete graph for each μ $(0 < \mu < \sigma)$;
(S3) no S_μ contains B_μ or any other B_λ $(0 \leqslant \lambda < \mu < \sigma)$.

(Condition (S3) is of lesser importance; its purpose is to avoid 'redundant' factors.)

The graph $\bigcup_{\lambda<\mu} B_\lambda$ in (S2) will be denoted by $G|_\mu$, and the simplex S_μ in (S2) will be called the *simplex of attachment* of B_μ. A graph is *prime* if it has no simplicial decomposition into more than one factor, and a *prime decomposition* is a simplicial decomposition in which every factor is prime. Furthermore, we shall call a subgraph $H \subset G$ *attached* to another subgraph $H' \subset G \setminus H$ if every vertex of H has a neighbour in H'. (These subgraphs H and H' will usually be induced and connected.) If H is not attached to any component of $G \setminus H$, we call H *unattached* in G.

We remark that a graph is prime if and only if it does not contain a separating simplex. This is not quite as trivial as it may seem, because one has to show that a simplex of attachment continues to separate the graph after the addition of further factors. This, however, is an easy consequence of (S2).

Proposition 1.1. *If $(B_\lambda)_{\lambda<\sigma}$ is a simplicial decomposition of G, and if $x \in G|_\mu \setminus S_\mu$ and $y \in B_\mu \setminus S_\mu$ for some $\mu < \sigma$, then S_μ separates x from y in G.*

Suppose we are given an arbitrary graph G and are asked to find a simplicial decomposition of G into primes. How shall we go about the problem? Clearly, there are two tasks involved; finding the right subgraphs of G to serve as the factors B_λ, and putting them together in accordance with (S1)–(S3).

The problem of which kinds of subgraph may be used as factors in a prime decomposition has been studied thoroughly and may be regarded as fully understood. Essentially, the prime factors of a graph (in any prime decomposition) are its smallest unattached convex subgraphs. (A subgraph $H \subset G$ is *convex* in G if any induced (or 'chordless') path in G connecting vertices of H is contained in H.) In fact, more is true: if F is a simplicial decomposition of G, then any subgraph of G which corresponds to a 'subtree' of the 'decomposition tree' of F (Fig. 2) is convex in G (see [1, Chs. 1.1 and 5.4] for details).

The more challenging part of the prime decomposition problem is the second of the two tasks mentioned: finding the correct order, if it exists, in which the potential prime factors B_λ may be assembled into a simplicial decomposition of G. Starting from a simple example, we shall now look into this problem in some detail. In order to concentrate on the essential, we shall assume that the *set* of potential factors B_λ has already been determined.

1.1. The construction of a prime decomposition

Consider the graph G shown in Fig. 3. The potential prime factors of G are the triangles $T_1, \ldots, T_4$. Trying to be as short-sighted as possible, let us begin to construct a prime decomposition $(B_\lambda)_{\lambda<\sigma}$ of G with the two bottom triangles, setting $B_0 := T_1$ and $B := T_3$, say. Then $S_1 = T_1 \cap T_3$, so S_1 is a simplex (of order 1) as required. For B_2, we can choose between the two remaining factors, the triangles T_2 and T_4. But neither of these choices is feasible: $S_2 = B_2 \cap (B_0 \cup B_1)$

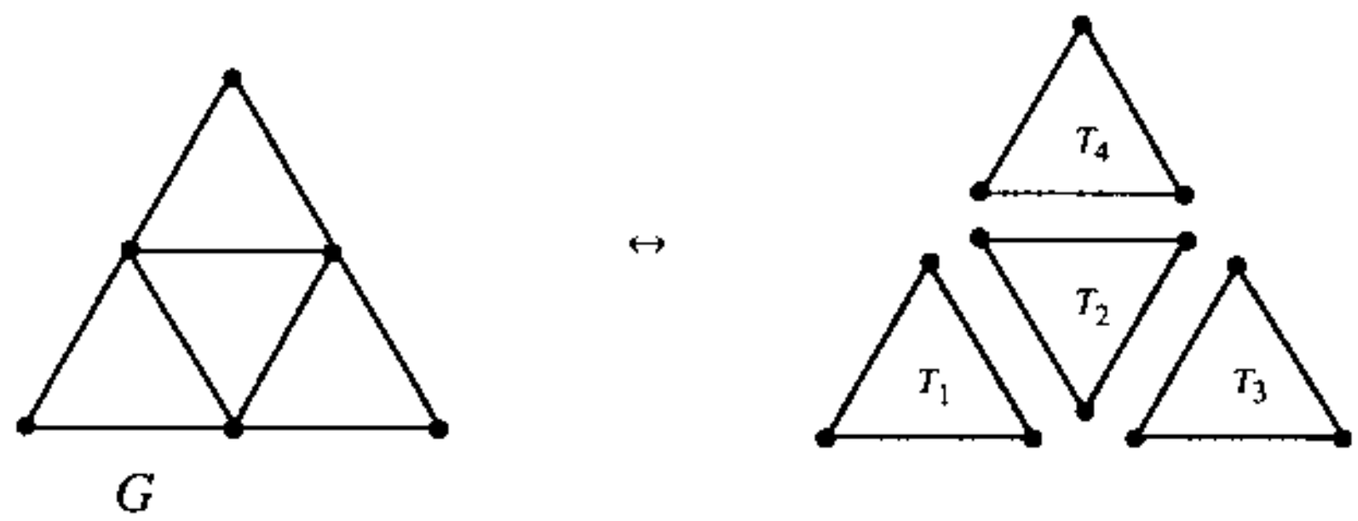


Fig. 3. Constructing a prime decomposition.

would be a path of length 2 or consist of two isolated vertices—thus in neither case would S_2 be a simplex.

What went wrong? Of course, we should not have chosen T_3 as the factor B_1: Proposition 1.1 requires that S_1 separate the vertices of $B_0 \setminus S_1$ from those of $B_1 \setminus S_1$ in G, which is clearly not the case with our choice of T_3 as B_1. And indeed, if we set $B_1 := T_2$ instead, we can easily complete our prime decomposition, e.g. to (T_1, T_2, T_3, T_4).

With the above example in mind, let us now consider a more abstract situation. Let G be a graph containing a simplex S which separates G into two components (making up $G \setminus S$), C and C' say. Let us assume that S is attached to C, and let $S' \subset S$ denote the simplex induced by those vertices of S that have a neighbour in C' (Fig. 4). Suppose we have started to build a prime decomposition of G, having chosen factors B_λ for all λ up to (but excluding) some ordinal μ. Suppose further that the part $G|_\mu$ of G we have covered lies entirely in $G[C' \cup S']$ (the subgraph of G induced by the vertices of C' and S'), so that $G|_\mu \cap C = \emptyset$ (but $G|_\mu \cap C' \neq \emptyset$). Let us consider the following question: under what assumptions can we choose B_μ so as to include a vertex from C?

Notice first that if B_μ is to include a vertex from C, it must lie entirely within $G[C \cup S]$: otherwise it would be separated by a subsimplex of S, which contradicts the requirement that it should be prime. Thus S_μ, the intersection of B_μ with $G|_\mu$, will be a part of S'. Now by Proposition 1.1, we can see that S_μ must in fact be equal to S': since S' is attached both to C and to C', S_μ could not otherwise separate $G|_\mu$, from B_μ in G.

This observation has two consequences, one practical, for the construction of concrete prime decompositions, the other more theoretical, for the problem of

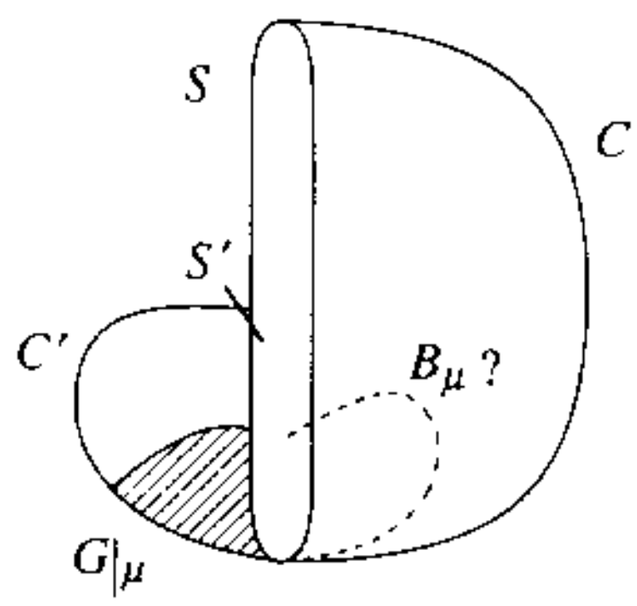


Fig. 4. Choosing B_μ.

when such a decomposition exists. The first consequence is that whenever we have S, C, C', $S' \subset G$ as in our example and we have started to construct a prime decomposition of G with factors inside $G[C' \cup S']$, we cannot move over into C until we have covered all of S' (because S' has to serve as the simplex of attachment for the first factor including a vertex from C). The impact of this restriction will be the greater the more separating simplices S there are in the graph G to be decomposed.

The other consequence of our observations sounds almost trivial, but it lies at the heart of the existence problem of prime decompositions: since B_μ is to be a prime factor and has to contain S', there must be at least one prime *subgraph* of G which contains S' and a vertex from C. We shall express this by saying that S' has a prime *extension into* C.

In the next section we shall see that such prime extensions do not always exist. This will lead us to the construction of a graph which does not admit a simplicial decomposition into primes, and on to a characterization of the graphs that do have prime decompositions in terms of the existence of prime extensions of separating simplices.

1.2. Prime extensions

When Halin [12] introduced simplicial decomposition for infinite graphs, he also proved the following first major theorem on the existence of prime decompositions.

Theorem 1.2. *Every graph not containing an infinite simplex admits a simplicial decomposition into primes.*

The key lemma in the proof of Halin's theorem asserts that a finite simple has a prime extension into any component to which it is attached.

Halin [12] also gave a construction of a graph that does not admit a prime decomposition. To get the flavour of why finding a prime decomposition may be difficult or impossible, consider the following variation of Halin's example: let x be a single vertex, $S = S[s_1, s_2, \ldots]$ an infinite simplex, $C = y_1 y_2 \ldots$ a one-way infinite path, and let H_1 be the graph obtained from the disjoint union of x, S and C by joining x to all the vertices of S and drawing the edges $y_i s_j$ for all $i \geqslant j$ (Fig. 5).

Here, S has no prime extension into C (observe that, in the notation of our earlier example, we would have $C' = \{x\}$ and hence $S' = S$): since every vertex $y \in C$ is separated in H_1 from some vertex $s \in S$ by a simplex T, there can be no prime subgraph in H_1 which contains both S and a vertex of C.

What does this mean for the construction of a prime decomposition of H_1? Certainly that if we approach S 'from the left', with x being contained in the first factor B_0, then the first factor B_μ containing a vertex from C would have to

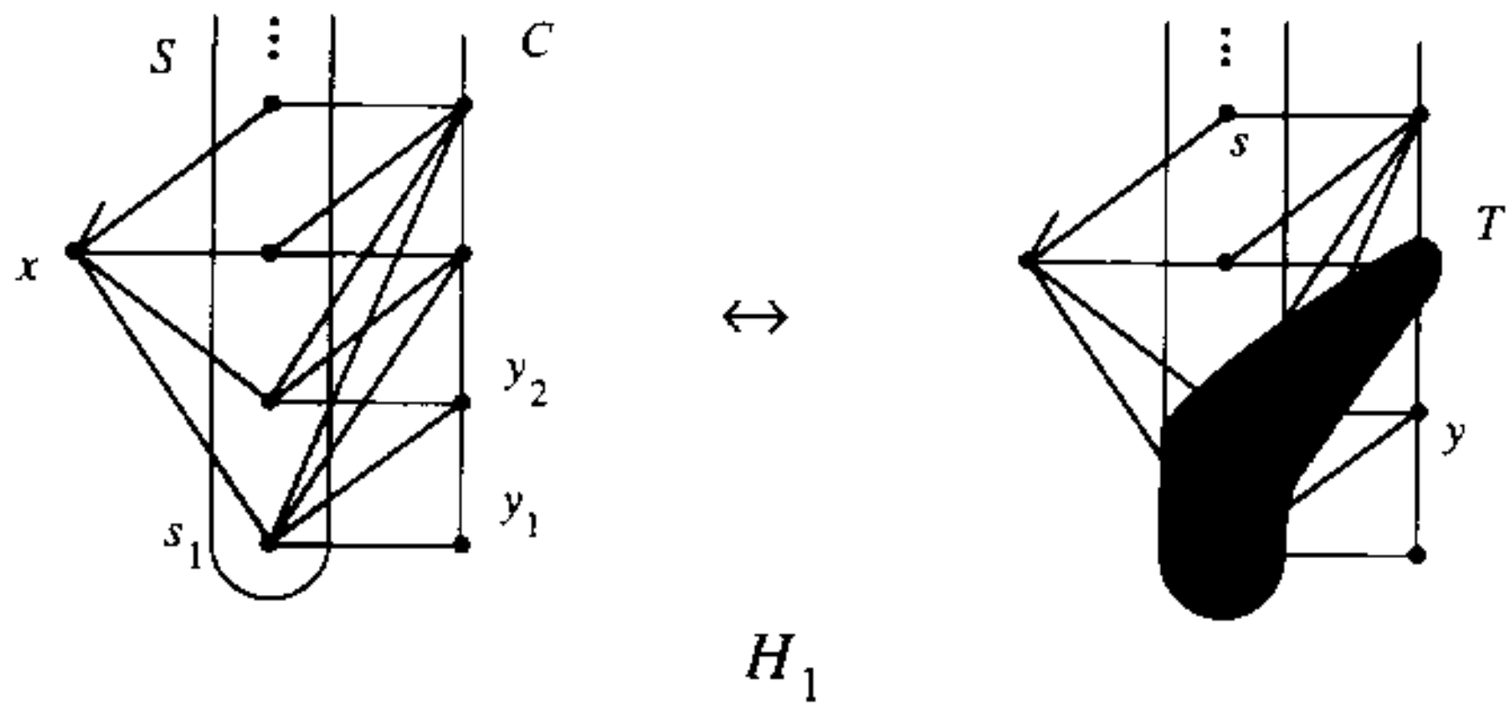


Fig. 5. S has no prime extension into C.

contain S as well (as its simplex of attachment S_μ), which is impossible if B_μ is to be prime. Hence, we have to approach S from the right: choosing factors from within $H_1[C \cup S]$ until S is covered, and then move across to cover x as well. Such a prime decomposition of H_1 does indeed exist; consider, for example, the decomposition (Fig. 6)

$$F = (Y_1, Y_2, Y_3, \ldots, X),$$

where

$$Y_i := H_1[y_i, y_{i+1}, s_1, \ldots, s_i], \qquad X := H_1[x, s_1, s_2, s_3 \ldots].$$

However, we only have to alter H_1 slightly to obtain a graph H_0 which does not admit a simplicial decomposition into primes: simply replace x with a copy C' of C (joined to S in the same way as C is). In this graph, which is very similar to Halin's original example, it makes no difference from which side we try to approach S in an attempt to construct a prime decomposition: the moment we try to cross S and move over into the other side we will have to find a prime extension of S (into that side), which does not exist.

Let us reformulate the indecomposability of the graph H_0 within the more general setting of our earlier prototype situation, involving an arbitrary separating simplex S and components C, C' of $G \backslash S$. We have seen that if S' (the part of S whose vertices have a neighbour in C') has no prime extension into C, then any prime decomposition of G has to start with factors containing vertices of C (and hence lying inside $G[C \cup S]$). Now there may be many such configurations of

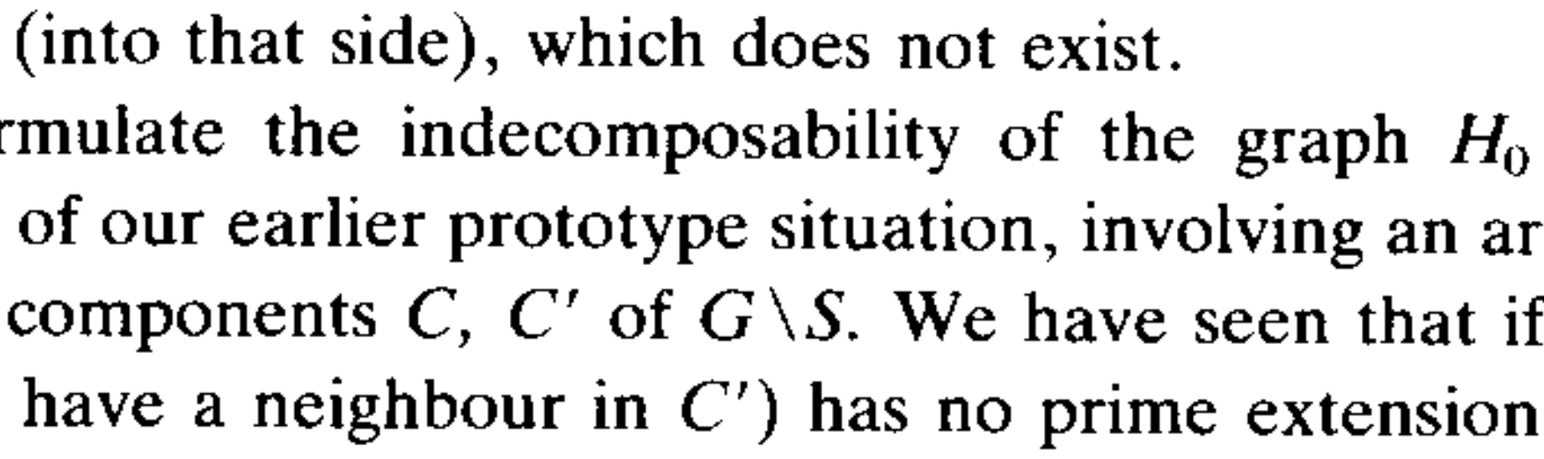


Fig. 6. A prime decomposition of H_1.

(say) S_i, C_i and C_i' in a graph, including several where S_i' fails to have a prime extension into C_i. As in each case any prime decomposition of G has to begin with vertices from C_i, these requirements are bound to conflict for different i.

One way in which this may happen led to the indecomposability of H_0: in a graph G which contains two critical configurations (S_1, C_1, C_1') and (S_2, C_2, C_2'), such that $C_1 \subset C_2'$ and $C_2 \subset C_1'$ (Fig. 7), a prime decomposition cannot begin with vertices from C_1 and at the same time with vertices from C_2. In this case, we shall say that G contains *simplices with opposite inaccessible sides*. (In the example of H_0, we have $S_1 = S_2 = S$, and $C_1 = C_2' = C$ (say), $C_2 = C_1' = C'$.)

Another way in which critical configurations may impose conflicting requirements on the order of factors in a prime decomposition is that a graph G contains an infinite series of critical configurations (S_i, C_i, C_i') such that $C_i \supset C_{i+1}$ for all i but $\bigcap_{i \in \mathbb{N}} C_i = \emptyset$ (Fig. 7): no matter how we choose the first factor B_0, there will be some $i \in \mathbb{N}$ for which B_0 lies in $G[C_i' \cup S_i']$, i.e. on the wrong side of S_i. In this case, we shall say that G contains *simplices with an infinite sequence of inaccessible sides*. (See [1, Ch. 2.5] for details.)

The following theorem of Dirac [9] summarizes our observations, providing a necessary condition for the existence of a prime decomposition.

Theorem 1.3. *A graph can only have a simplicial decomposition into primes if it contains neither simplices with opposite inaccessible sides nor simplices with an infinite sequence of inaccessible sides.*

The question of whether or not Theorem 1.3 has a direct converse is still an open problem:

Problem. Are the conditions in Dirac's theorem also sufficient for the existence of a prime decomposition?

While it is relatively easy to derive the necessary conditions for the existence of a prime decomposition as stated in Theorem 1.3—the formal proof goes hardly beyond the sketch we have seen above—a sufficiency proof is likely to require deeper arguments. However, it is quite possible that most of the necessary ideas

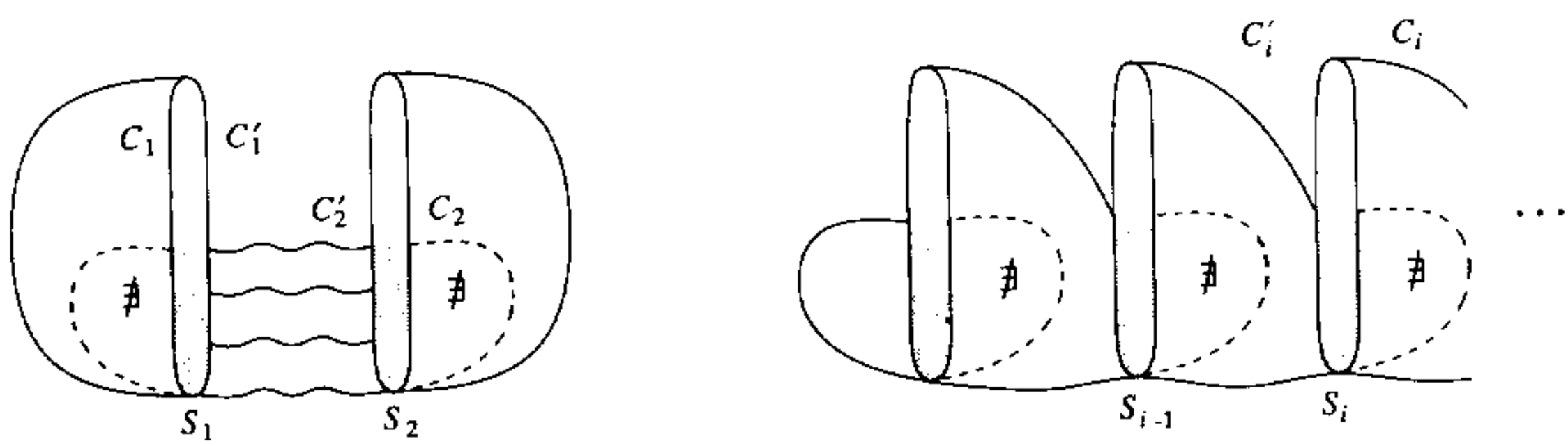


Fig. 7. The forbidden configurations in Dirac's theorem.

are already contained in the proof our next theorem, which solves the existence problem of prime decompositions for an important special case.

A simplicial decomposition $F = (B_\lambda)_{\lambda<\sigma}$ of G is called a *simplicial tree-decomposition* if, in addition to (S1)–(S3), it satisfies

(S4) each S_μ is contained in B_λ for some $\lambda < \mu$ $(\mu < \sigma)$.

Condition (S4) and the concept of simplicial tree-decompositions will be introduced properly in Section 4; for the moment it suffices to remark that these decompositions form a natural subclass of all simplicial decompositions, including (for example) all finite simplicial decompositions.

Returning to our prototype configuration of a separating simplex S with components C and C' of $G \backslash S$, notice that (S4) has the simplifying effect of making the situation symmetric: in order for G to admit a simplicial tree-decomposition into primes, S' must have prime extensions into *both* C and C'. For if we start our decomposition in C' (say) and B_μ is the first factor containing a vertex from C, then (S4) requires that S_μ $(= S')$ is also contained in some earlier factor B_λ, which would be a prime extension of S' into C'.

For countable graphs G, it was shown in [5] that the existence of prime extensions in such configurations is not only necessary but also sufficient for the existence of a simplicial tree-decomposition of G into primes (the proof draws on almost all aspects of simplicial decomposition theory; with additional analysis and motivation, it takes up most of a chapter in [1]).

Theorem 1.4. *A countable graph G has a simplical tree-decomposition into primes if and only if G satisfies the following condition.*

(†) *If $S \subset G$ is a simplex, C and C' are distinct components of $G \backslash S$, and S is attached to C, then the simplex $S' \subset S$ induced by the vertices of S with a neighbour in C' has a prime extension into C.*

Notice that our graph H_1, which has a (general) simplicial decomposition into primes, fails to satisfy (†) and therefore has no simplicial tree-decomposition into primes.

To conclude this section, let us see why Theorem 1.4 generally fails for uncountable G.

Let T be a the 'transitive closure' (or comparability graph) of the infinite dyadic tree. (For example, let $V(T) = \{0, 1\}^{<\omega}$, the set of all finite 0-1 sequences, and join $(a_0, \ldots, a_n)$ to $(b_0, \ldots, b_m)$ whenever $n < m$ and $a_i = b_i$ for $i = 0, \ldots, n$.) Let T' be obtained from T by adding its 'limit points': for each maximal simplex M in T add a new vertex $v(M)$, joining $v(M)$ to every vertex of M. Thus, $M' := T'[M \cup \{v(M)\}]$ is a maximal simplex in T'. (In the concrete model of T cited above, the new vertices $v(M)$ may be thought of as representing the 0-1 sequences of length ω.) We shall prove that T' satisfies (†) but has no simplicial tree-decomposition into primes.

To see that T' satisfies (†), notice that any separating simplex $S \subset T'$ attached to a component C of $T' \backslash S$ has a prime extension of the form M' into C.

In order to show that T' has no simplicial tree-decomposition into primes, notice first that there are $2^{\aleph_0}$ vertices of the form $v(M)$, and hence $2^{\aleph_0}$ simplices of the form M'. Moreover, these simplices are the only possible prime factors of T'. (Since we skirted round the subject of which subgraphs can occur as prime factors in simplicial decompositions, the reader must be asked to take this on faith. However, we did mention that simplicial factors are always unattached—and the simplices M' are clearly the only unattached and prime subgraphs of T'.)

Now suppose that $F = (B_\lambda)_{\lambda<\sigma}$ is a simplicial tree-decomposition of T' into primes. Consider a maximal simplex M in T, and assume that $M' = B_\mu$. Then $S_\mu \subset M$. Since M' is the only factor in F that contains the entire simplex M, but S_μ must also be contained in some other factor B_λ (by (S4)), S_μ cannot be equal to M; hence, $M \backslash S_\mu \neq \emptyset$. We may therefore associate with the factor $M' = B_\mu$ a vertex $w(M') \in M \backslash S_\mu$, noting that M' is the first factor in F that contains $w(M')$ (because $w(M') \notin S_\mu$ implies that $w(M') \notin G|_\mu$). We have thus obtained an injective map from the uncountable set of factors in F to the vertex set of T, which contradicts the countability of T.

1.3. *Simplicial minors*

It is a striking phenomenon that long after Halin had published his original example of a graph not admitting a simplicial tree-decomposition into primes, no essentially different such 'counterexample' had been found. Indeed, all graphs G without such a prime decomposition seem to have a structure very much like that of H_1: they all contain an infinite simplex S separating G into components C and C', where S is attached to C, $G[C \cup S]$ is covered by a family $(Y_1, Y_2, \ldots)$ of convex subgraphs that have the same intersection pattern with S and with each other as in the case of H_1, and S' (defined as earlier) is not contained in any one of the Y_i. It is then only a small step further to notice that contracting C' to a single vertex x and maybe shrinking the subgraphs Y_i a little gives one a contraction of $G[C' \cup S \cup C]$ onto H_1.

And indeed, the following can be shown:

Any graph not admitting a simplicial tree-decomposition into primes has a minor isomorphic to H_1. (*)

How about the converse of (*)—that is, how far is (*) from being a characterization of the graphs that admit simplicial tree-decompositions into primes?

Notice that the graph property of not containing H_1 as a minor is closed under taking minors. Thus we can only hope to characterize the prime-decomposable graphs by that property if they share this feature, i.e. if minors of decomposable graphs are again decomposable. Or more intuitively, if we seek to express our observation that H_1 is in a sense that the 'simplest' nondecomposable graph by

use of the minor relation, then this relation and our notion of 'simpler' should match: a minor of a graph with a prime decomposition should itself admit a prime decomposition, and this should be at most as complex as that of the original graph.

With the usual concept of a minor, however, this is far from true. Recall that a minor of a graph G is obtained in two steps:

(1) taking a subgraph of G, and
(2) contracting connected parts of the subgraph.

Clearly, both these steps are too general to meet our above requirement. That is, if G is a graph that has a relatively simple prime decomposition, and if H is obtained from G by either of the two steps, then H may only admit much more complex prime decompositions, or even none at all.

For example, if G is an infinite simplex and thus admits the trivial prime decomposition consisting only of itself, we can find subgraphs in G with arbitrarily complex or even no prime decompositions, including H_1. An only slightly more complicated example shows that even if we restrict step (1) to taking induced subgraphs it still allows us to obtain H_1 from a prime graph: if we add a new vertex to H_1 and join it to all old vertices by independent paths of length at least 2, the resulting graph is prime (because it has no separating simplex) and contains H_1 as an induced subgraph.

As an example for step (2), consider the graph G obtained by identifying two cycles of order 10 along 3 consecutive vertices x, y, z. Then G is prime, but contracting the edges xy and yz results in a graph that has a separating simplex (the contracted vertex), and therefore only a nontrivial prime decomposition into two factors. Or more extremely, if we subdivide every edge of H_1 once, the resulting graph will again be prime (and therefore admit the trivial prime decomposition), but we can reobtain H_1 from it only by contracting the appropriate edges.

Thus if we want the converse of (*) to hold, we have to restrict the definition of a minor by sharpening both of the above steps.

Let us call two vertices of a graph *simplicially close* if they are not separated by any simplex, and let us call H a *simplicial minor* of G if H is obtained from G by

(1′) taking a convex subgraph of G, and
(2′) contracting connected parts of this convex subgraph in such a way that simplicially close vertices remain simplicially close.

It is easily seen that this restricted concept of a minor no longer admits the examples we considered above. And indeed, simplicial minors satisfy the converse of (*):

If H_1 is a simplicial minor of G, then G admits no simplicial tree-decomposition into primes (**)

(More generally, one can show that if G has a simplicial tree-decomposition into

primes and H is a simplicial minor of G, then H has a simplicial tree-decomposition into primes; see [1, Ch. 3.2].)

In order to achieve (**), we had to come down a long way from the most general concept of a minor that formed the basis of (*). It would therefore not be surprising if we now had to pay for the gain of (**) with the loss of (*), that is, if not admitting a prime decomposition no longer implied the existence of certain (simplicial) minors like H_1.

The following main theorem of [2] asserts that this is not the case: the graphs admitting a simplicial tree-decomposition into primes are characterized by only two forbidden simplicial minors, H_1 and the graph H_2 obtained from H_1 by filling in all missing edges of the form $y_i y_j$.

Theorem 1.5. *A countable graph G admits a simplicial tree-decomposition into primes if and only if neither H_1 nor H_2 is a simplicial minor of G.*

Given the intuitive appeal of characterizations by forbidden configurations and the apparent suitability of simplicial minors in the context of the prime decomposition problem, it would be interesting to derive a simplicial minor version of Dirac's theorem (Theorem 1.3). Moreover, our example T' of an uncountable graph without a simplicial tree-decomposition into primes (see the end of the previous section) suggests the following conjecture.

Conjecture. A graph G (of any cardinality) admits a simplicial tree-decomposition into primes if and only if none of H_1, H_2 or T' is a simplicial minor of G.

2. Decompositions into small factors

The quest for prime decompositions, as discussed in Section 1, arises straight from the very notion of a decomposition: to break down a graph into factors which are as small as possible. In the case of prime decompositions, 'small' is defined locally: a factor is considered small enough (only) if it cannot be replaced with factors decomposing it further.

If we replace this local concept of 'small' with a global one, imposing a fixed bound on the size of each factor, we obtain a problem which is quite different in character but no less interesting: *given any cardinal κ, which graphs admit a simplicial decomposition into factors of order $<\kappa$?*

There are two obvious obstructions to the existence of a simplicial decomposition into such uniformly small factors. One is the existence of large complete subgraphs: if $G = K_\kappa$, for example, then G is prime and cannot be decomposed into factors of order $<\kappa$. (Note, however, that if we delete the vertex x from

our earlier graph H_1, we obtain a graph H_1' which contains an infinite simplex but still admits a simplicial decomposition into the finite factors Y_i.)

Another property which tends to be incompatible with decomposition into factors of order $<\kappa$ is high connectivity. For suppose that G has a simplicial decomposition into such factors, and that $x, y \in V(G)$ are non-adjacent. If x and y are in the same factor B, they cannot be joined by a or more independent paths. (These paths could without loss of generality be chosen induced, and would thus have to be contained in B, because as a simplicial factor B is a convex subgraph of G.) But if x and y are not in a common factor, they are separated by some simplex of attachment S_μ, so again they cannot be joined by κ or more independent paths (because $|S_\mu| < |B_\mu| < \kappa$).

The following theorem of Halin [15] says that if κ is regular and uncountable, then these two obstructions to the existence of a simplicial decomposition into factors of order $<\kappa$ are all there are: if we ban K_κ subgraphs and large systems of independent paths, the desired decomposition exists.

Theorem 2.1. *Let G be a graph and κ a regular uncountable cardinal. Suppose that $G \not\supset K_\kappa$, and that for any two non-adjacent vertices $x, y \in V(G)$ there are fewer than κ independent x–y paths in G. Then G admits a simplicial decomposition into factors of order $<\kappa$.*

Note that Theorem 2.1 does not extend down to $\kappa = \aleph_0$: the infinite grid, for example, is prime (and thus has no simplicial decomposition into finite factors), but it contains neither an infinite simplex nor an infinite set of independent paths joining any two vertices.

Thus, embarrassingly, the simplest and most intriguing case of our problem remains open:

Problem. Which graphs admit a simplicial decomposition into finite factors?

3. The uniqueness of prime decompositions

An aspect of simplicial decomposition theory which has been studied in some detail is that of the uniqueness of prime decompositions.

Problem. Which graphs have a simplicial decomposition into a unique set of prime factors?

Recall that, by Halin's theorem (Theorem 1.2), every graph not containing an infinite simplex admits a simplicial decomposition into prime factors. These

factors are uniquely determined; they are precisely those prime subgraphs that are convex and unattached. The unique set of primes of a finite graph can be obtained particularly easily, by iteratively splitting the graph along minimal separating simplices.

As soon as we admit infinite simplices as subgraphs, however, prime decompositions need no longer be unique. Consider, for example, our graph $H_1' = H_1 - x$.

H_1' admits two different prime decompositions,

$$F = (S, Y_1, Y_2, Y_3, \dots) \quad \text{and} \quad F' = (Y_1, Y_2, Y_3, \dots).$$

Thus, while F is a perfectly acceptable simplicial decomposition into primes, its factor S is in a sense redundant: if we omit it, the remaining factors still form a prime decomposition of H_1'.

Let us call a simplicial decomposition *reduced* if it has no such redundant factors. Reduced prime decompositions are by far the most 'common' kind; for example, all prime decompositions into finite factors are reduced [6].

The following result was obtained in [1].

Theorem 3.1. *Any two reduced prime decompositions of a graph have the same set of factors.*

The immediate question arising from this result is whether every graph that has some simplicial decomposition into primes also has a reduced such decomposition—in which case prime decompositions could in practice be taken reduced as a matter of course. However, this is not the case; our graph T, the 'transitive closure' of the infinite dyadic tree, has numerous simplicial decompositions into primes—select maximal simplices M in any order—but no reduced prime decomposition.

It would be interesting to know (and should not be too difficult to decide) whether T is essentially the only example of such a graph:

Problem. Is T in some sense contained in every graph that admits a simplicial decomposition into primes but no reduced such decomposition?

The problem of determining the graphs which admit a reduced simplicial decomposition into primes has another fascinating aspect: any solution would, on the basis of other known results, imply a solution to the last problem of the previous section, to determine which graphs admit a simplicial decomposition into finite factors. See [1, Ch. 4.4] for details.

4. Tree-decompositions

Simplicial decomposition theory does not only provide the theoretical basis for applications such as excluded minor theorems by homomorphism bases; it is also valuable for the study of a more general type of decompositions, called tree-decompositions.

The original idea behind the concept of a tree-decomposition, first introduced by Robertson and Seymour [16], was to make precise the apparent tree shape imposed on a graph by a simplicial decomposition; see Fig. 2. How could this be done?

The definition given by Robertson and Seymour refers directly to the tree $T(G)$ which the shape of the graph G is deemed to resemble: they call a family $(X_t)_{t\in T(G)}$ of subsets of $V(G)$ a tree-decomposition of G if:

(T1) $\bigcup_{t\in T(G)} X_t = V(G)$;
(T2) $\forall xy \in E(G): \exists t \in T(G): x, y \in X_t$;
(T3) if $t, t', t'' \in T(G)$ and t' lies on the $t - t''$ path in $T(G)$, then $X_t \cap X_{t''} \subset X_{t'}$.

For compatibility with simplicial decompositions, we shall here take a slightly different approach, closer to our definition of a simplicial decomposition. Given a simplicial decomposition $F = (B_\lambda)_{\lambda<\sigma}$, how should we associate a tree T_F with F, so as to express the tree shape of G imposed by F?

The obvious choice for the vertex set of T_F is the set of factors in F,

$$V(T_F) = \{B_\lambda \mid \lambda < \sigma\}.$$

The edges of T_F should, intuitively, correspond to the simplices of attachment in F, so let us choose them inductively for each $\mu < \sigma$. Having constructed a partial tree on $\{B_\lambda \mid \lambda < \mu\}$, how shall we join the next 'vertex' B_μ to this tree? In the example of Fig. 2, the obvious solution would be to join B_μ to the unique B_λ (with $\lambda < \mu$) which contains S_μ. Clearly, this can be done whenever F is such that

(S4) each S_μ is contained in B_λ for some $\lambda < \mu$ $(\mu < \sigma)$;

if B_λ is not unique, we simply choose λ to be minimal. We may therefore call F a *simplicial tree-decomposition* of G if, in addition to (S1)–(S3), it also satisfies (S4).

Note that the existence of the decomposition tree T_F, as defined above, is now independent of the conditions which originally gave rise to F as a simplicial decomposition, in particular of (S2). If the tree shape of G is all we are interested in, we may therefore discard (S2) (and, if we so wish, (S3)) and call any family $F = (B_\lambda)_{\lambda<\sigma}$ a *tree-decomposition* of G if it satisfies (S1) and (S4). It is not difficult to show that this definition of a tree-decomposition is equivalent to that given by Robertson and Seymour, except for the additional well-ordering of $V(T_F)$. (The latter is often convenient to have, especially for proofs by induction on μ.)

Tree-decompositions are well worth investigating for their own sake, just as simplicial decomposition are. The most natural questions, however, are different.

For example, there is not much point in asking for a tree-decomposition into primes: the only graphs that are prime (i.e. irreducible) with respect to tree-decompositions are the complete graphs, and only chordal graphs can have tree-decompositions into complete factors [3]. On the other hand, the value of a tree-decomposition can meaningfully be measured by the size of its factors. This raises questions like that of determining the *tree-width* of the graphs G with a given property—the tree-width of G is the smallest natural k, if one exists, such that G admits a tree-decomposition into factors of order at most $k+1$—or of how to obtain algorithmically a tree-decomposition of a graph that realizes its tree-width.

The following natural problem has not yet, to my knowledge, been investigated in its own right.

Problem. Which graphs admit a tree-decomposition into finite factors?

Robertson, Seymour and Thomas (see [18]) have recently shown that the exclusion of infinite complete minors (even of *subdivisions* of infinite simplices) is sufficient for the existence of such a decomposition: every $K_{\aleph_0}$-free graph has a tree-decomposition into finite factors. The converse of this, however, is false: our decomposition $F' = (Y_1, Y_2, Y_3, \dots)$ of the graph H_1' in Section 3 is a tree decomposition into finite factors, but H_1' contains an infinite simplex.

Even if a simplicial decomposition, by the example shown in Fig. 2, has inspired our definition of a tree-decomposition—we have not yet shown that every simplicial decomposition F does indeed satisfy (S4) and thereby defines a decomposition tree T_F. In other words: can a simplicial decomposition fail to satisfy (S4) and thus not be a simplicial tree-decomposition?

At first glance, it seems pretty obvious that (S4) should follow from (S2). For if a simplex of attachment S_μ is not contained in a single earlier factor B_λ, one should expect it to contain vertices that are separated by some other simplex of attachment—which of course is impossible if these vertices, being in the same simplex, are to be adjacent (Fig. 8). And indeed, it is not difficult to show that

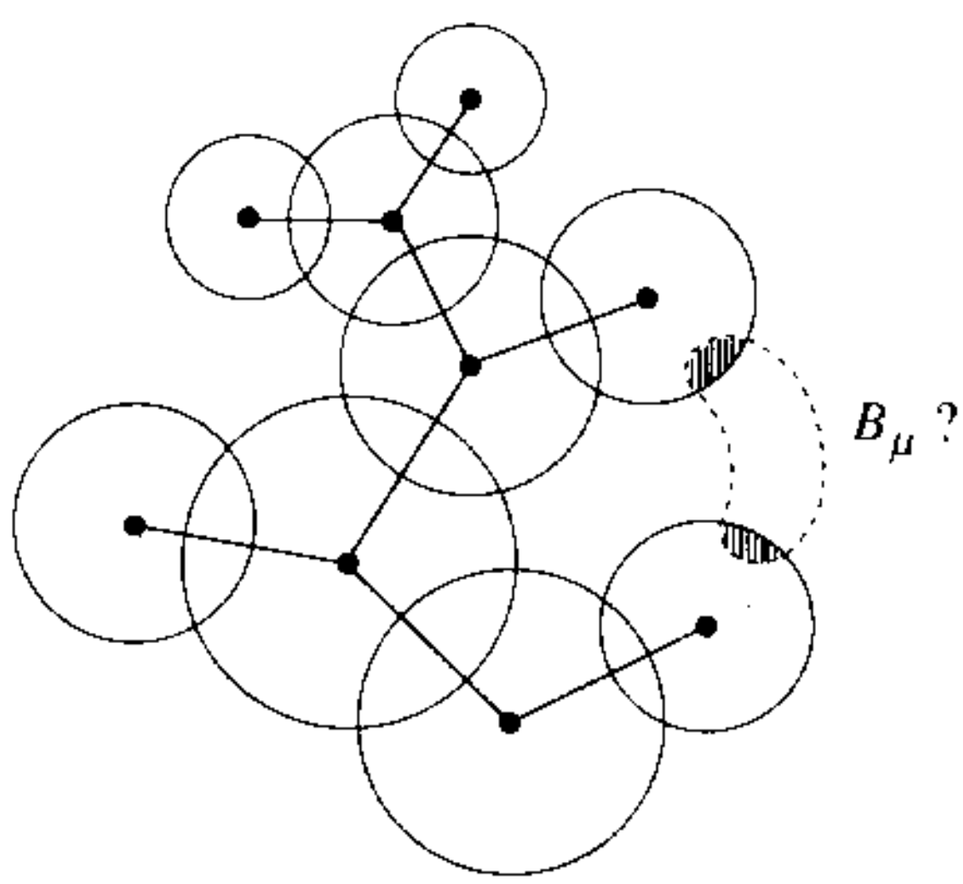


Fig. 8. How (S2) prevents cycles in the decomposition tree.

any simplicial decomposition is in fact a simplicial tree-decomposition—if all its simplices of attachment are finite.

However, we have already seen an example of a simplicial decomposition which is not a simplicial tree-decomposition: the decomposition of the graph H_1 given in Section 1.2. This decomposition does not satisfy (S4), because the (infinite) simplex of attachment of its last factor X, the simplex S, is not contained in any of the earlier factors Y_i.

Thus if we tried to define a decomposition tree for this decomposition of H_1 in the usual way, we would get stuck at the last factor X: this factor should somehow sit 'above' the infinite path $Y_1 Y_2 \ldots$, but it has no immediate predecessor among the vertices of this path. (Intuitively, we fail to obtain a tree not because the factor X would introduce a cycle—a danger averted effectively by condition (S2); see Fig. 8—but because we would lose the connectedness of the tree.)

Fortunately, such 'trees' with limit points are familiar objects: they are equivalent to *order theoretical trees,* partially ordered sets in which every set of predecessors is linearly ordered. (In our example, $\{Y_1, Y_2 \ldots\}$ would be the set of predecessors of X.) And indeed, it is not difficult to show that with any simplicial decomposition F which fails to satisfy (S4) one can associate such a (well-founded) order theoretical decomposition tree T_F, much in the same way as one associates a graph theoretical tree with a simplicial tree-decomposition. (The similarity can be made precise: the decomposition tree for general simplicial decompositions is defined in such a way that it coincides with the (natural order of the rooted) decomposition tree induced by (S4) if the decomposition happens to be a simplicial tree-decomposition; see [1, Ch. 5.4] for details.)

Let us return to the original theme of this section. We set out to find a concept of tree-decomposition which makes precise the apparent tree shape of a graph decomposed simplicially. We have partially succeeded in doing so by introducing the condition (S4); this condition, if satisfied, gives rise to a decomposition tree T_F which corresponds in the desired natural way to a given decomposition F. However, we have seen that (S4) is not satisfied by every simplicial decomposition. On the other hand, a general simplicial decomposition still gives rise to a generalized type of decomposition tree. The question we face, therefore, is this: how can (S4) be weakened, to another condition (S5), say, in such a way that every simplicial decomposition satisfies (S5) but (S5) is still strong enough to give rise to an order theoretical decomposition tree as associated with a general simplicial decomposition?

Our discussion of H_1 suggests that the desired condition (S5) should aim to capture the 'acyclicity' in the arrangement of factors in a simplicial decomposition, as shown in Fig. 8. For as we have seen, this acyclicity is precisely the tree-like property which is retained by order theoretical over graph theoretical trees (as opposed to their connectedness), while on the other hand this acyclicity does indeed follow from (S2), and is therefore a feature of general simplicial

decompositions. In other words, (S5) should express that a new factor B_μ cannot use vertices of $G|_\mu$ for its simplex of attachment if these vertices are taken from different 'branches' of the decomposition tree of $G|_\mu$. The following condition aims to express this:

(S5) $\mu < \sigma,\ x \in G|_\mu \setminus S_\mu,\ y \in B_\mu \setminus S_\mu \Rightarrow \nexists \lambda < \sigma : x, y \in B_\lambda.$

And indeed, it is not difficult to show that this condition (S5) has the desired properties: while being a consequence of (S2) and thus common to all simplicial decompositions, it gives rise to the same decomposition tree as that associated with a general simplicial decomposition. We are therefore justified in calling a family $F = (B_\lambda)_{\lambda<\sigma}$ a *generalized tree-decomposition* if it satisfies (S1) and (S5), and the order theoretical tree T_F associated with it its (*generalized*) *decomposition tree*. Then any tree-decomposition (as well as any simplicial decomposition) is also a generalized tree-decomposition—notice that (S4) implies (S5)—and the (order theoretical) decomposition tree of a generalized tree-decomposition F coincides with the (graph theoretical) decomposition tree of F if F happens to satisfy (S4).

Let us finally see how the concepts of these tree-decompositions can be applied. Recall that with the use of simplicial decompositions for the characterization of minor-closed graph properties (via homomorphism bases) we encountered a serious problem: in many cases the prime factors which were needed to construct all the graphs without a certain given minor were as numerous and as varied as these graphs themselves, a 'characterization' of these graphs by their prime factors therefore both difficult and pointless. It is therefore natural to relax the definition of the decomposition used: this should result in smaller and therefore fewer possible different factors.

The tree-decompositions defined above were inspired by this idea: they still aim to capture the overall tree structure of the graphs with a given property (in our case, without a given excluded minor), while making fewer requirements on the nature of the factors and their attachment graphs. Of course, the graphs with a given property can only be characterized by their tree shape if they all share a common tree structure. While this cannot be expected of an arbitrary minor-closed graph property, it is worth an effort trying to find out for which properties it can be done.

Problem. Which graph properties can be described in terms of the tree structure of their members?

For properties defined by the exclusion of a single finite minor X, this question is answered beautifully by a result of Robertson and Seymour [17]: the graphs without an X minor have tree-decompositions into factors of bounded finite order if and only if X is planar.

Graph properties defined by excluding infinite minors have only been studied very recently. For graphs without large infinite simplices as minors, a characterization purely in terms of tree-decompositions was obtained in [8] (or 1, Ch. 5.4)].

Theorem 4.1. *If κ is a regular uncountable cardinal, then a graph G has no K_κ minor if and only if G admits a generalized tree-decomposition $F = (B_\lambda)_{\lambda<\sigma}$ such that every B_λ and every chain in T_F has order $<\kappa$.*

The proof of this result is based on Theorem 2.1 and uses simplicial decompositions. Applying different methods, Robertson, Seymour and Thomas have recently obtained an equivalent theorem, which moreover has a natural extension to singular κ and to $\kappa = \aleph_0$. For a survey of this and other beautiful structure theorems in terms of generalized tree-decompositions see [18] in this volume.

References

[1] R. Diestel, Graph Decompositions: A Study in Infinite Graph Theory (Oxford Univ. Press, Oxford, 1990).

[2] R. Diestel, Simplicial minors and decompositions of graphs, Math. Proc. Camb. Phil. Soc. 103 (1988) 409–426.

[3] R. Diestel, Tree-decompositions, tree-representability and chordal graphs, Discrete Math. 71 (1988) 181–184.

[4] R. Diestel, Simplicial tree-decompositions of infinite graphs I, J. Combin. Theory Ser. B 48 (1990) 197–215.

[5] R. Diestel, Simplicial tree-decompositions of infinite graphs II—the existence of prime decompositions, J. Combin. Theory Ser. B 50 (1990) 93–116.

[6] R. Diestel, Simplicical tree-decompositions of infinite graphs III—the uniqueness of prime decompositions, J. Combin. Theory Ser. B 50 (1990) 117–128.

[7] R. Diestel, Simplicial decompositions of graphs: a survey of applications, Discrete Math. 75 (1989) 121–144.

[8] R. Diestel, The structure of TK_a-free graphs, J. Combin. Theory Ser. B, to appear.

[9] G.A. Dirac, Simplicial decompositions of infinite graphs, in: Combinatorica, Conv. Roma 1983, Symp. Math. 28 (1986) 159–195.

[10] R. Halin, Über einen graphentheoretischen Basisbegriff und seine Anwendung auf Färbungsprobleme, Diss. Köln, 1962.

[11] R. Halin, Über einen Satz von K. Wagner zum Vierfarbenproblem, Math. Ann. 153 (1964) 47–62.

[12] R. Halin, Simpliziale Zerfällungen beliebiger (endlicher oder unendlicher) Graphen, Math. Ann. 156 (1964) 216–225.

[13] R. Halin, Zur Klassifikation der endlichen Graphen nach H. Hadwiger und K. Wagner, Math. Ann. 172 (1967) 46–78.

[14] R. Halin, Ein Zerlegungssatz für unendliche Graphen und seine Anwendung auf Homomorphiebasen, Math. Nachr. 33 (1967) 91–105.

[15] R. Halin, Unterteilungen vollständiger Graphen in Graphen mit unendlicher chromatischer Zahl, Abh. Math. Sem. Univ. Hamburg 31 (1967) 156–165.

[16] N. Robertson and P.D. Seymour, Graph minors II: algorithmic aspects of tree-width, J. Algorithms 7 (1986) 309–322.

[17] N. Robertson and P.D. Seymour, Graph minors V: excluding a planar graph, J. Combin. Theory Ser. B 41 (1986) 92–114.
[18] N. Robertson, P.D. Seymour and R. Thomas, Excluding infinite minors, Discrete Math. 95 (this Vol.) (1991) 303–319.
[19] K. Wagner, Über eine Eigenschaft der ebenen Komplexe, Math. Ann. 114 (1937) 570–590.
[20] K. Wagner, Über eine Erweiterung eines Satzes von Kuratowski, Deutsche Math. 2 (1937) 280–285.
[21] K. Wagner, Bemerkungen zu Hadwigers Vermutung, Math. Ann. 141 (1960) 433–451.

Discrete Mathematics 95 (1991) 91–99
North-Holland

Bounded graphs

R. Halin

Mathematisches Seminar der Universität Hamburg, Bundestraße 55, W-2000 Hamburg 13, Germany

Received 3 April 1990
Revised 21 May 1990

Abstract

Halin, R., Bounded graphs, Discrete Mathematics 95 (1991) 91–99.

A graph G is called bounded if for every mapping $f: V(G) \to \mathbb{N}$ there exists a sequence $(c_n)_{n\in\mathbb{N}}$ of naturals such that for every ray $U = (u_1, u_2, u_3, \ldots)$ of G an n_0 can be found with $f(u_i) < c_i$ for all $i \geqslant n_0$. All locally finite connected graphs are bounded. Four basic types of unbounded graphs are given, and it is conjectured that one of these configurations must be present in every unbounded graph. This conjecture is proved for the graphs with rayless blocks and the graphs which do not contain an infinite system of disjoint rays.

The idea to introduce the class of graphs indicated in the title arose from Rado's study of universal graphs [6], especially de Bruijn's proof of the statement that there is no universal locally finite countable graph (i.e., no locally finite countable graph which contains a copy of every other such graph).

First some terminology has to be developed. Let $(a_i)_{i\in\mathbb{N}}$, $(b_i)_{i\in\mathbb{N}}$ be sequences of natural numbers. We will say that (b_i) *majorizes* (a_i) if there is an n_0 such that $b_i > a_i$ for all $i \geqslant n_0$.

Lemma 1. *If $(a_i^\nu)_{i\in\mathbb{N}}$, $\nu = 1, 2, 3, \ldots$, is a countable family of sequences of natural numbers, then there is a sequence $(b_i)_{i\in\mathbb{N}}$ which majorizes these sequences $(a_i^\nu)_{i\in\mathbb{N}}$ simultaneously.*

Proof. Let $b_i := 1 + \max(a_i^\nu : \nu \leqslant i)$. □

Now let G be a graph, $U = (u_0, u_1, u_2, \ldots)$ a ray (i.e., a one-way infinite path) in G. Here the u_i are the vertices of U in their natural order, which means that $[u_i, u_{i+1}]$ are the edges of U. Let $f: V(G) \to \mathbb{N}$ be a function and $b = (b_i)_{i\in\mathbb{N}}$ a sequence of natural numbers. We say that b majorizes f on U if b majorizes the sequence $f(u_0), f(u_1), f(u_2), \ldots$. Further we say that b *ray-majorizes* f (with respect to G, or on G), if b majorizes f on every ray of G.

Now we define the central notion of this article. A graph G is *bounded* if for every function $f: V(G) \to \mathbb{N}$ there is a sequence of natural numbers which

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

ray-majorizes f on G. The problem that will concern us is to try to determine the graphs which are bounded.

It is clear that every subgraph of a bounded graph is again bounded. Moreover, to check whether a connected graph is bounded it suffices to consider rays starting in some fixed vertex.

Lemma 2. *Let G be a connected graph, $f: V(G) \to \mathbb{N}$ a function, and assume that $(b_i)_{i \in \mathbb{N}}$ majorizes f on all those rays which start in some fixed vertex u_0. Then there is a sequence $(c_i)_{i \in \mathbb{N}}$ of natural numbers which ray-majorizes f on G.*

Proof. Without loss of generality we may assume (b_i) to be monotonically increasing; otherwise we replace b_i by $\max(b_1, \ldots, b_i)$. Now let $c_i = b_{2i}$. Let $V = (v_0, v_1, v_2, \ldots)$ be a ray in G. By the connectedness of G there is a ray $U = (u_0, u_1, u_2, \ldots)$ in G such that $U \cap V$ is a ray with the vertices $u_{i_0+\nu} = v_{j_0+\nu}$ $(\nu = 0, 1, 2, \ldots)$. By assumption there is a k such that $f(u_{i_0+\nu}) < b_{i_0+\nu}$ for all $\nu \geqslant k$. Then for all $\nu \geqslant \max(k, i_0)$ we have $f(v_{j_0+\nu}) < b_{i_0+\nu} \leqslant b_{2(j_0+\nu)} = c_{j_0+\nu}$. □

The following lemma is proved similarly.

Lemma 3. *If H is a bounded subgraph of G such that every ray of G has a subray in H, then G itself is bounded.*

As a consequence of Lemmas 1 and 3 we have the following.

Lemma 4. *If $F \subset G$ is finite and $G - F$ has at most countably many components which are all bounded, then G itself is bounded.*

Let us consider some examples of unbounded graphs which seem to be the prototypes for this class of graphs. Since there are no more than 2^ω distinct maps from the vertex set of a ray to $\mathbb{N}$ (ω denotes the smallest infinite cardinal), it is clear that the union of 2^ω disjoint rays is unbounded. We remark that, in the absence of the Continuum Hypothesis, Martin's Axiom (which is independent of ZFC) implies the existence of an uncountable system of disjoint rays forming a bounded graph. To avoid set theoretic intricacies, we shall therefore assume the Continuum Hypothesis in what follows.

Note further that the ω-regular tree T_ω is not bounded. For if any injective mapping $f: V(T_\omega) \to \mathbb{N}$ is given, every finite path can be extended by a vertex of arbitrarily large value under f. Similarly one establishes that every subdivision of T_ω is unbounded. By analogous reasoning one finds that also the *bundle graph* and the *fan graph* (Fig. 1) and their subdivisions are unbounded. (The bundle graph arises from a ray by replacing each edge in 'even position' with a $K_{2,\omega}$; the fan graph is obtained from a ray $(u_0, u_1, u_2, \ldots)$ by 'planting' disjoint rays R_n on u_{2n} and connecting u_{2n-1} to all the vertices of R_n by edges, for $n = 1, 2, \ldots$).

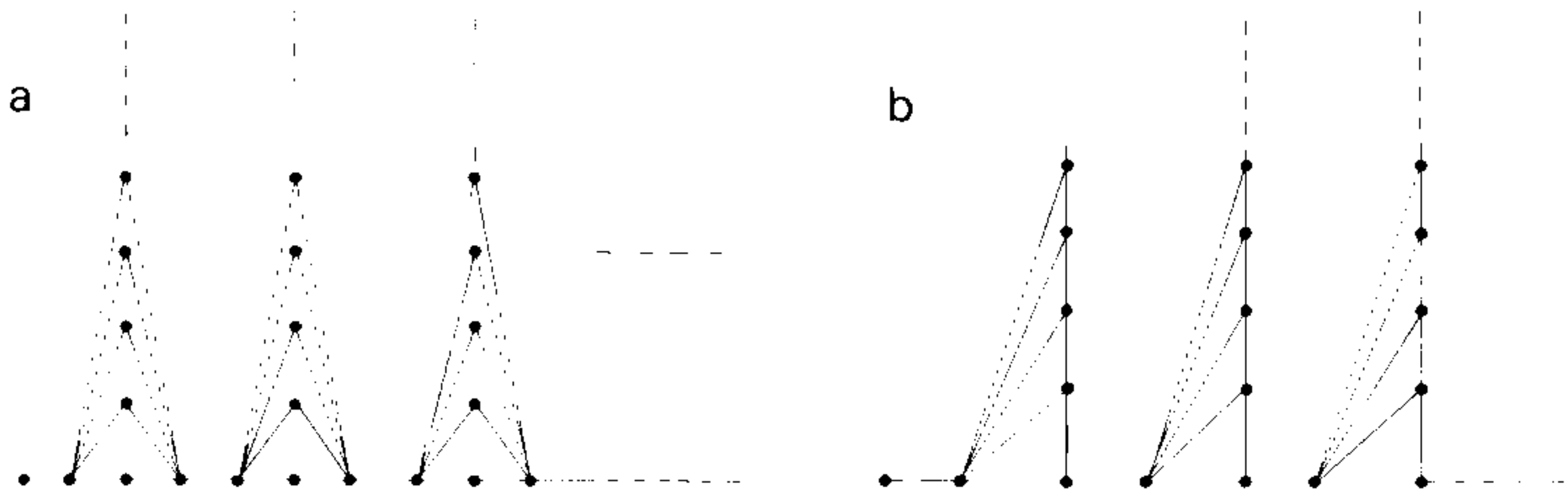


Fig. 1. The bundle graph (a) and the fan graph (b).

Note that the graph of Fig. 2, which looks so similar to the fan graph, is nevertheless bounded (apply Lemma 1 to the vertical rays, choosing the majorizing f_n monotonic).

A large variety of bounded graphs is given by the following theorem.

Theorem 1. *Every locally finite countable graph G is bounded.*

Proof. By Lemma 4 we may assume G to be connected. Let $V(G) = \{v_0, v_1, v_2, \ldots\}$ be an enumeration of the vertices of G. Consider an arbitrary function $f: V(G) \to \mathbb{N}$. Let V_i be the set of all $x \in V(G)$ with distance at most i from $\{v_0, \ldots, v_i\}$. Clearly every V_i is finite. Now let $b_i := 1 + \max(f(x): x \in V_i)$. If $U = (u_0, u_1, u_2, \ldots)$ is a ray in G, then $u_0 = v_k$ for some $k \in \mathbb{N}$. Then for $i \geqslant k$ every u_i belongs to V_i; hence $f(u_i) < b_i$ for all $i \geqslant k$. Thus the sequence (b_i) ray-majorizes f on G. □

Now it is easy to show de Bruijn's result [6] that the class of locally finite countable graphs has no universal element. Let G be a graph from this class and choose $f(v)$, for each $v \in V(G)$, as the degree of v in G. By Theorem 1 there is a sequence (b_i) of naturals which ray-majorizes f on G. Let H be any graph having a ray $U = (u_0, u_1, \ldots)$ such that the degree of u_i is $\geqslant b_i$ for each i. (For instance H can be chosen as a locally finite tree.) Then H cannot be isomorphic to a subgraph

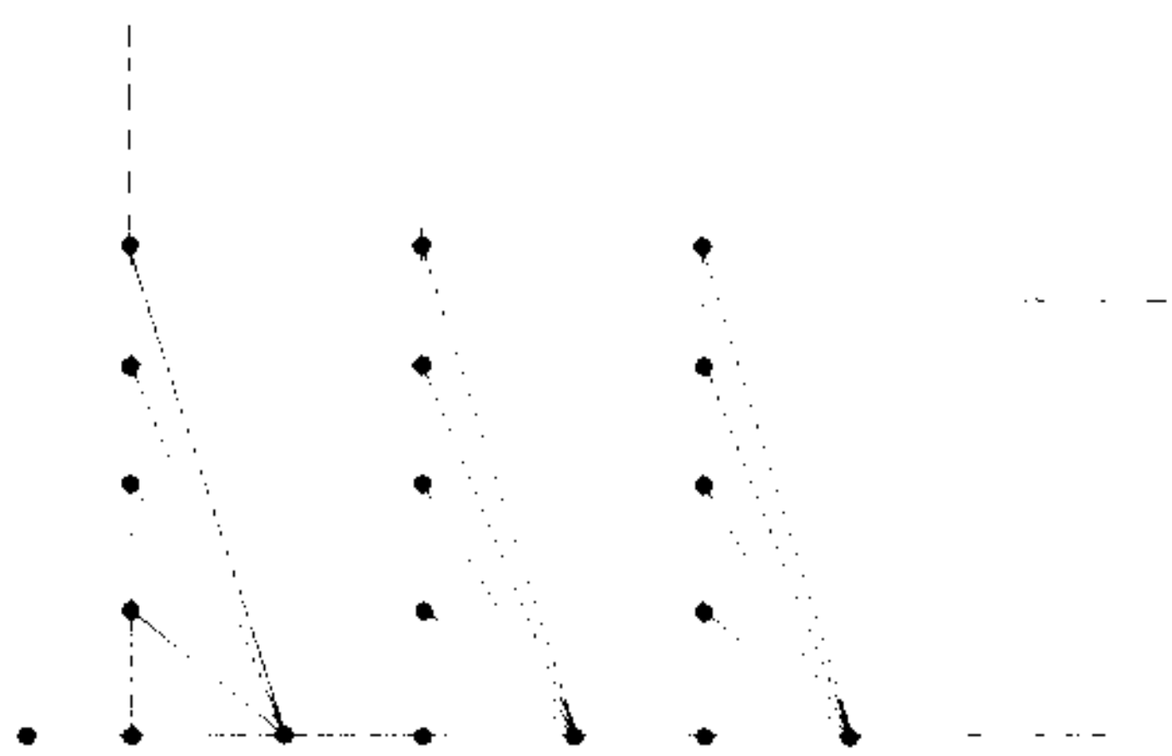


Fig. 2.

of G, by choice of the sequence (b_i). Hence G cannot be universal in the class of countable locally finite graphs.

In the class of all countable graphs the bounded ones form a proper sub-class, itself containing the class of locally finite countable graphs properly. The main problem in this context is, naturally, how the bounded graphs can be characterized, especially whether there is a criterion in terms of forbidden configurations.

As we have seen, the bundle graph, the fan graph, the ω-regular tree and the disjoint union of uncountably many rays are forbidden in a bounded graph. The following is an old conjecture of the author, first proposed (for countable graphs) in 1964, in connection with [6].

Bounded Graph Conjecture (BGC). Assume the Continuum Hypothesis. A graph G is bounded if and only if it contains neither an uncountable system of disjoint rays nor a subdivision of the bundle graph, the fan graph or the ω-regular tree T_ω.

BGC would imply that an uncountable graph which does not contain uncountably many disjoint rays is bounded if and only if every countable subgraph is bounded. Also it would follow from BGC that a countable graph G is already bounded if there exists at least one injective function $f: V(G) \to \mathbb{N}$ which is ray-majorized by some sequence of natural numbers. (This statement is easily derived from the fact that no injective $\mathbb{N}$-valued function on one of the forbidden configurations can be ray-majorized by a sequence of naturals.)

We shall prove BGC for two special classes of infinite graphs. First we state a further lemma.

Lemma 5. *Let R, S be disjoint finite sets of vertices in a graph G, and let H denote the union of all $R-S$ paths in G (i.e. of those paths which connect an $r \in R$ with an $s \in S$ and have only their endvertices in common with $R \cup S$). If H is infinite, then either it contains a ray, or there exists an $R-S$ path $P = r \cdots s$ with vertices $x, y \in P$ whose Menger (or local connectivity) number is infinite. That is to say, we can find infinitely many internally disjoint $x-y$ paths in H which have only x and y in common with P.* (Such a configuration will be called an *$R-S$ bundle* with endpoints r, s; see Fig. 3.)

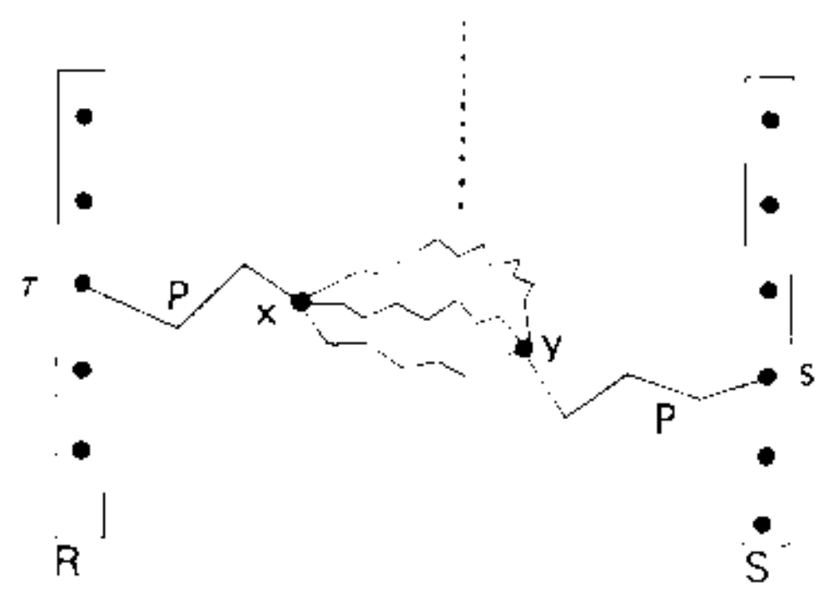


Fig. 3. An $R-S$ bundle.

Proof. Let G_{rs} be obtained from G by adding two new vertices r and s, joining r to the vertices of R and s to those of S. Similarly, let H_{rs} consist of H, the new vertices r and s, and the $r-H$ and $s-H$ edges of G_{rs}.

Of course, H_{rs} is connected. If it does not contain a ray (which would have a subray in H), there must exist a vertex $x \notin \{r, s\}$ of infinite degree (with respect to H_{rs}). Let N be the set of neighbours of x. For each $w \in N$ the edge $[x, w]$ lies on an $r-s$ path P_w which we consider to be oriented from r to s. There is an infinite subset N' of N such that, for all $w \in N'$, the vertices x, w have the same position with respect to the orientation of P_w; by symmetry we may assume that x preceeds w on P_w for all $w \in N'$. Let H' be the (infinite) union of the $x-s$ paths $P \subset P_w$ for $w \in N'$, and put $H'' := H' - x$. Note that H'' is still connected. If there is a finite $F \subset V(H'')$ such that $H'' - F$ has infinitely many components $C_\nu (\nu \in \mathbb{N})$ each containing a $w_\nu \in N'$, then choose paths Q_ν from w_ν to F. There must be a vertex $y \neq s$ of F which is the endvertex of infinitely many such Q_ν; let $w \in N'$ be such that $y \in P_w$, and let $\mu \in \mathbb{N}$ be such that $P_w \cap C_\nu = \emptyset$ for every $\nu > \mu$. The paths $[x, w_\nu] \cap Q_\nu$ for $\nu > \mu$, together with P_w, then give the desired configuration. On the other hand, if there is no such finite F, we choose an arbitrary vertex v_0 of H''. $H'' - v_0$ must have a component C_0 which contains infinitely many vertices from N'. Let v_1 be a neighbour of v_0 in C_0. Then again $C_0 - v_1$ must have a component C_1 containing infinitely many vertices from N', and we choose v_2 as a neighbour of v_1 in C_1. Continuing in this way we obtain a ray in H_{rs}, which has a subray in H. Our lemma follows. □

Theorem 2. *BGC holds for every graph G all blocks of which are rayless.*

Proof. We suppose that G is unbounded, and show that G contains one of the configurations listed in BGC. This is clear if G has uncountably many components each containing a ray; we may therefore assume that G has only countably many such components. By Lemma 4, then, one of these components must be unbounded. Since it suffices to find one of the desired configurations in this component, we may thus assume that G is connected. Let as arrange the blocks of G in such a way that they form a simplicial decomposition of G (see [1], or [2, p. 224]). This means that the blocks are indexed as subgraphs G_λ of G where λ runs through all ordinals less than some fixed ordinal σ, and for each τ (with $0 < \tau < \sigma$) G_τ has exactly one vertex a_τ (an articulation) in common with the union of the G_λ with $\lambda < \tau$. Let τ_- denote the smallest λ with $a_\tau \in V(G_\lambda)$. Note that unless $\tau_- = 0$, a_τ and a_{τ_-} are distinct vertices of G_τ , and hence joined by a nontrivial path in G_τ . Take the ordinals $< \sigma$ to be the vertices of graph $\mathcal{T}$ whose edges are the (unordered) pairs $[\tau, \tau_-]$. Each sequence $\tau, \tau_-, (\tau_-)_-, \ldots$ ends at 0 after finitely many steps; so $\mathcal{T}$ is connected. Further, if $\tau_1 < \tau_2 < \cdots < \tau_k$ are finitely many vertices of $\mathcal{T}$, then τ_k can have at most one neighbour with respect to $\mathcal{T}$ among $\tau_1, \ldots, \tau_{k-1}$; hence $\mathcal{T}$ is without circuits. We see that $\mathcal{T}$ is a tree. Consider the ordinal 0 as its root; we denote the natural order of the tree $\mathcal{T}$ (with

respect to the root 0) by $\ll$ and carry it over to the blocks of G (by the correspondence $\lambda \to G_\lambda$). We fix a vertex r in G_0. By Lemma 2, in order to test the boundedness of G, we need only consider rays starting in r. Each such ray R traces out a unique ray R^* of $\mathcal{T}$ starting in 0, and vice versa, if a ray T of $\mathcal{T}$ starting in 0 is given, we find at least one ray $R \subseteq G$ starting in r with $R^* = T$; all rays R of this kind traverse the same blocks, namely those corresponding to the vertices of T.

If, for some λ with $0 < \lambda < \sigma$, there is no ray $\subseteq \mathcal{T}$ starting in 0 and containing λ, then in G there is no ray starting in r and meeting G_λ in a vertex different from a_λ. So we may omit this block G_λ because it is not relevant for our problem of boundedness; we see that we may assume, without loss of generality, that $\mathcal{T}$ has no endvertex (with the possible exception of 0). Further, if $\mathcal{T}$ has a vertex of uncountable degree, we find uncountably many disjoint rays in $\mathcal{T}$ and therefore in G. So we can also assume that $\mathcal{T}$ is countable.

If there is a ray T in $\mathcal{T}$ such that infinitely many of the G_λ with $\lambda \in V(T)$ are infinite, then by applying Lemma 5 to each such G_λ we find a subdivision of the bundle graph in G. (Note that, since G_λ is 2-connected, every vertex or edge of G_λ lies on a path in G connecting the first and the last vertex in G_λ of any ray $R \subset G$ with $R^* = T$.) So we may assume that for every ray T of $\mathcal{T}$ all the G_λ with $\lambda \in V(T)$, apart from finitely many, are finite.

Let us show that $\mathcal{T}$ is unbounded. Suppose $\mathcal{T}$ is bounded; we shall prove that then G is also bounded, contrary to our assumption. Let $f: V(G) \to \mathcal{N}$ be given. For each finite G_λ put $f^*(\lambda) = \max(f(x): x \in V(G_\lambda))$; for infinite G_λ let $f^*(\lambda) = 1$. By assumption there is a monotonically increasing sequence (c_ν) ray-majorizing f^* on $\mathcal{T}$. Now let $U = (u_0, u_1, \ldots)$ with $u_0 = r$ be a ray in G. U determines a ray $U^* = (\lambda_0, \lambda_1, \ldots)$ in $\mathcal{T}$; each u_j is in a G_{λ_i} with $i \leqslant j$. There exists a k such that G_{λ_i} is finite and $f^*(\lambda_i) < c_i$ for each $i \geqslant k$. Since, for sufficiently large n, $u_n \in G_{\lambda_i}$ implies $i \geqslant k$, we therefore have $f(u_n) \leqslant f^*(\lambda_i) < c_i \leqslant c_n$ for all such n, showing that (c_ν) ray-majorizes G. As this implies that G is bounded (contrary to our assumption), we deduce that $\mathcal{T}$ is unbounded, as claimed.

Call a vertex α of $\mathcal{T}$ *essential* if the branch above α, i.e. the subtree induced by α and all $\lambda \gg \alpha$, is unbounded. The essential vertices induced a subtree $\mathcal{T}_0$ of $\mathcal{T}$, which includes the root 0. Since $\mathcal{T}$ is countable, $\mathcal{T} - \mathcal{T}_0$ has at most countably many components $\mathcal{C}_i$, which are all bounded. If $\mathcal{T}_0$ is also bounded, then let an arbitrary mapping $f: V(\mathcal{T}) \to \mathbb{N}$ be given. Since the $\mathcal{C}_i$ and $\mathcal{T}_0$ are all bounded, for each of these trees there is a sequence ray-majorizing the restriction of f to that tree, and by Lemma 1 we get a sequence which majorizes all these sequences simultaneously. Since each ray of G ends in $\mathcal{T}_0$ or in some $\mathcal{C}_i$, we find that the latter sequence ray-majorizes f on $\mathcal{T}$. As this contradicts the fact that $\mathcal{T}$ is unbounded, we find that $\mathcal{T}_0$ too is unbounded.

Hence, by Theorem 1, $\mathcal{T}_0$ cannot be locally finite; let τ_0 be a vertex of infinite degree with respect to $\mathcal{T}_0$. The neighbours $\gg \tau_0$ of τ_0 in $\mathcal{T}_0$ determine infinitely many branches of $\mathcal{T}$ which are themselves unbounded. In each of these branches

we analogously find a vertex of infinite degree, and repeating this procedure again and again we get a subdivision $\mathcal{U}$ of T_ω in $\mathcal{T}$; let ξ be the minimal vertex of $\mathcal{U}$ with respect to $\ll$.

If τ is a vertex of infinite degree in $\mathcal{U}$, then there are infinitely many λ in $\mathcal{U}$ such that a_λ is in G_τ. If $\tau \gg \xi$ then also a_{τ_-} is in G_τ. As G_τ is connected and rayless, we can find a vertex z_τ in G_τ and a system of paths (having pairwise only z_τ in common) from z_τ to infinitely many of the a_λ. (z_τ and some of the a_λ may coincide; then the paths in question are of length 0.) Further, by adding a $z_\tau - a_\tau$ path and 'sacrificing' finitely many of the other paths if necessary, we may assume that a_{τ_-} is among these a_λ. We now apply this observation repeatedly, starting with ξ and performing the same procedure step by step to subsequent vertices of infinite degree in U. We finally get a subdivision of T_ω in G, and our proof is complete. □

Theorem 2 is a sharpening of Theorem 6.6 in [5], in which BGC was proved for countable trees. In the light of the graphs in Figs. 1b and 2 it seems rather hopeless to apply the method of our last proof for a more general verification of BGC, e.g. for all graphs which have a simplicial decomposition with finite members.

In [4] it was proved that in every graph G there is a system of disjoint rays with maximal cardinality; the latter invariant is denoted by $m_1(G)$.

Theorem 3. *BGC holds for all graphs G with $m_1(G) < \infty$.*

Proof. Assume that $m_1(G) = n \geqslant 1$ (the case $n = 0$ is trivial). Suppose G is unbounded; our aim is to establish the existence of a subdivision of the bundle graph in G. Since there are only finitely many ends in G, there exists a finite $T \subset G$ such that each component of $G - T$ has at most one end (see [3] for the definition of an end). If each of the finitely many components of $G - T$ containing rays were bounded, their union would form a bounded subgraph of G in which each ray ends, and G itself would be bounded by Lemma 3. Therefore at least one of these components must be unbounded, so we may assume without loss of generality that G itself has only one end. By *Satz 2* in [4] there exists a finite $F \subset G$ such that $G - F$ is the union of an infinite sequence $G_0, G_1, G_2, \ldots$ of (induced) rayless subgraphs such that each G_i ($i \geqslant 1$) has a subgraph T_i of n vertices in common with G_{i-1} and empty intersection with each G_ν, $\nu \leqslant i-2$, and further there is a path matching of T_i and T_{i+1} in G_i (see Fig. 4).

Note that distinct T_i must be disjoint. Moreover, each T_i separates (in $G - F$) the vertices of $(G_0 \cup \cdots \cup G_{i-1}) - T_i$ from those of $(G_i \cup G_{i+1} \cup \cdots) - T_i$, and the union of the aforementioned path matchings forms a system of n disjoint rays $R_1, \ldots, R_n$. We shall assume that there is no rayless component of $G - F$; otherwise it could be omitted by Lemma 3.

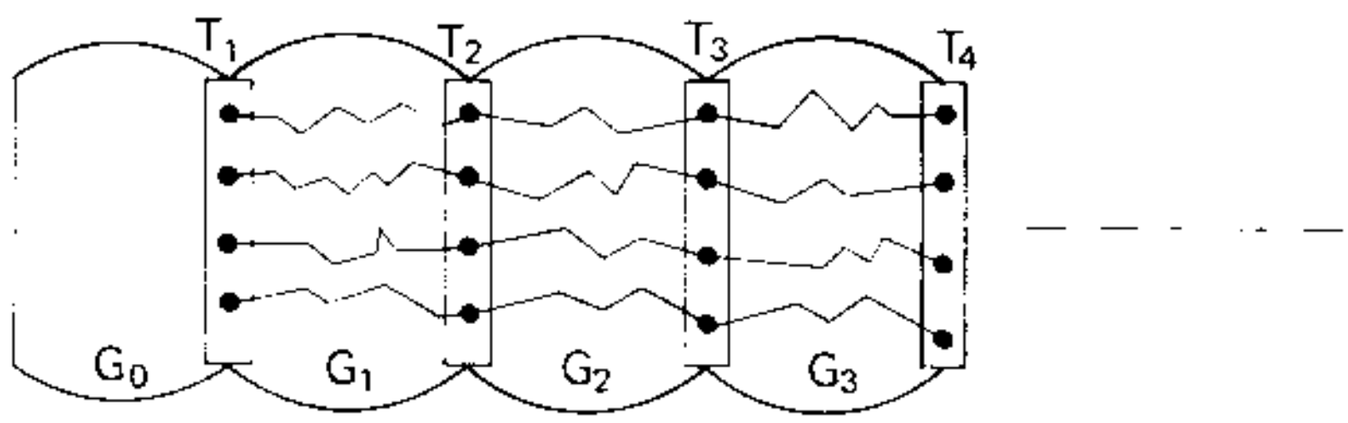


Fig. 4.

As $R_1, \ldots, R_n$ all belong to the same end, for any pair R_i, R_j there are infinitely many disjoint paths each connecting a vertex of R_i with a vertex of R_j, and we find an infinite sequence of integers $0 = i_0 < i_1 < i_2 < \cdots$ such that

$$G'_\nu := G_{i_\nu} \cup G_{i_\nu+1} \cup \cdots \cup G_{i_{\nu+1}-1}$$

is connected, for $\nu = 0, 1, 2, \ldots$.

We see that the G'_ν form an analogous representation of $G - F$; we may thus assume that already every G_i is connected. Choose a fixed vertex r of $G_0 - T_1$ (as a 'root'), and let V' be the set of vertices of $G - F$ which lie on a ray starting in r; put $V'_i = V(G_i) \cap V'$. Obviously $T_i \subseteq V'$ for all $i \in \mathbb{N}$, and every ray of G has a subray containing only vertices of V'.

If there is a $k \in \mathbb{N}$ such that all V'_j with $j \geq k$ are finite, then let

$$H := G_k \cup G_{k+1} \cup G_{k+2} \cup \cdots,$$

and let an arbitrary function $f: V(H) \to \mathbb{N}$ be given. Let r' be a fixed vertex ('root') of T_k, and put

$$c_i := 1 + \max(f(v) : v \in V'_k \cup \cdots \cup V'_{k+i})$$

for all $i \in \mathbb{N}$. If $U = (u_0, u_1, u_2, \ldots)$ is a ray $\subseteq H$ starting in $r' = u_0$, then obviously $u_i \in V'_k \cup \cdots \cup V'_{k+i}$ for each i, implying $f(u_i) < c_i$. Therefore H is bounded by Lemma 2. Since every ray of G ends in H, Lemma 3 then implies that G is bounded, a contradiction. Hence, there are infinitely many $i \in \mathbb{N}$ such that V'_i is infinite. Let V''_i denote the set of those vertices in G_i which lie on a $T_i - T_{i+1}$ path ($i \geq 1$). Then $V''_i \subseteq V'_i$, but there may be vertices in V'_i which do not belong to V''_i. Each such vertex x must lie on a path $P \subseteq G_i$ connecting a pair of vertices of T_i or a pair of vertices of T_{i+1}, and is separated from V''_i by T_i (or T_{i+1}, respectively). We then say that x is *attached* to T_i (or T_{i+1}, respectively).

For each i such that V''_i is infinite we may use Lemma 5 to obtain a $T_i - T_{i+1}$ bundle B_i with certain endpoints s_i in T_i and t_i in T_{i+1}. Now suppose that the set I of these i is infinite. We choose an infinite sequence i_ν ($\nu \in \mathbb{N}$) of elements of I such that $i_\nu + 2 \leq i_{\nu+1}$ for all $\nu \in \mathbb{N}$. Further we choose a path P_1 from r to s_{i_1} through $G_0 \cup \cdots \cup G_{i_1-1}$; then we add B_{i_1} and select a path P_2 from t_{i_1} to s_{i_2} through $G_{i_1+1} \cup \cdots \cup G_{i_2-1}$. Then we add B_{i_2} and select a path from t_{i_2} to s_{i_3} through $G_{i_2+1} \cup \cdots \cup G_{i_3-1}$. Continuing in this way we get a subdivision of the bundle graph in $G - F$.

If there are only finitely many i for which V_i'' is infinite, then there must be infinitely many i such that infinitely many elements of V_i' are attached to T_i or T_{i+1}. But every vertex of V_i' attached to T_i or T_{i+1} lies on a $T_{i-1} - T_{i+2}$ path. Putting $R := T_{i-1}$ and $S := T_{i+2}$ in Lemma 5 for infinitely many such i (sufficiently far apart), we again obtain a bundle graph. □

Acknowledgement

I express my thanks to Reinhard Diestel for several helpful comments resulting in a considerable improvement of the paper. Especially he drew my attention to a gap in the original proof of Theorem 3.

References

[1] R. Diestel, Graph Decompositions: A Study in Infinite Graph Theory (Oxford Univ. Press, Oxford, 1990).
[2] R. Halin, Über simpliziale Zerfällungen beliebiger (endlicher oder unendlicher) Graphen, Math. Ann. 156 (1964) 216–225.
[3] R. Halin, Über unendliche Wege in Graphen, Math. Ann. 157 (1964) 125–137.
[4] R. Halin, Über die Maximalzahl fremder unendlicher Wege in Graphen, Math. Nachr. 30 (1965), 63–86.
[5] R. Halin, Some problems and results on infinite graphs, Ann. Discrete Math. 41 (North-Holland, Amsterdam, 1989) 195–210.
[6] R. Rado, Universal graphs and universal functions, Acta Arithm. 9 (1964) 331–340.

Discrete Mathematics 95 (1991) 101–117
North-Holland

A survey on graphs with polynomial growth

W. Imrich and N. Seifter

Institut für Mathematik und Angewandte Geometrie, Montanuniversität Leoben, A-8700 Leoben, Austria

Received 6 December 1989
Revised 8 February 1990

Abstract

Imrich, W. and N. Seifter, A survey on graphs with polynomial growth, Discrete Mathematics 95 (1991) 101–117.

In this paper we give an overview on connected locally finite transitive graphs with polynomial growth. We present results concerning the following topics:

Automorphism groups of graphs with polynomial growth.
Groups and graphs with linear growth.
S-transitivity.
Covering graphs.
Automorphism groups as topological groups.

1. Introduction

The concept of growth was introduced to study finitely generated infinite groups. It first appeared in a paper of Adelson-Velsky and Shreider [1] which was published in 1957. The *growth function* $f_G(n)$ of a group G with respect to a finite generating set H is given by $f_G(0) = 1$ and

$$f_G(n) = |\{g \in G \mid g = h_1 \cdots h_n,\ h_i \in H \cup H^{-1} \cup \{e\}\}|$$

for $n \geqslant 1$. We say that a group G has *exponential growth* if there is a constant $c > 1$ such that $f_G(n) \geqslant c^n$ holds for all $n \in \mathbb{N}$. Otherwise G has *non-exponential growth*. In particular, G has *polynomial growth* if there are constants c and d such that $f_G(n) \leqslant cn^d$ for all $n \in \mathbb{N}$. Furthermore, we say that groups of non-exponential growth which grow faster than any polynomial have *intermediate growth*. Clearly these properties do not depend on the generating set.

Milnor [23] and Wolf [45] carried out the first investigations of growth conditions in groups. In [45] Wolf proved the important result that almost nilpotent groups have polynomial growth (A group G *almost* has property P if a normal subgroup of finite index of G has property P.) For groups G with

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

polynomial growth Milnor [25] conjectured that there always exist constants c_1, c_2 and an integer d such that

$$c_1 n^d \leqslant f_G(n) \leqslant c_2 n^d. \tag{1}$$

He also conjectured that the class of groups with polynomial growth coincides with the class of almost nilpotent groups. In 1972 Bass [3] proved Milnor's first conjecture, i.e. (1), for almost nilpotent groups. The second conjecture was settled by Gromov [14]. This deep result of Gromov [14], together with the above mentioned result of Wolf, is crucial for almost everything we carry out in this paper. For further references we therefore state it as the following theorem.

Theorem 1.1 (Gromov [14], Wolf [45]). *A finitely generated group has polynomial growth if and only if it is almost nilpotent.*

This result also implies the validity of Milnor's first conjecture. This means that (1) always holds for groups with polynomial growth. We call the thus well-defined integer d the growth degree d_G of G.

Wolf also conjectured that a group has exponential growth if its growth function dominates any polynomial, i.e. that there exist no groups with *intermediate* growth. This conjecture holds for solvable groups (cf. [45, 24]). In general it is false, as was first shown by Grigorchuk (cf. [12–13]). He found torsion-free groups as well as p-groups with intermediate growth. The growth functions of one class of groups with intermediate growth constructed by Grigorchuk satisfy e.g. the condition

$$2^{\sqrt{n}} < f_G(n) < 2^{n^{\log_{32} 31}}$$

for sufficiently large n. Further examples of groups with intermediate growth were given by Fabrykowski and Gupta [8].

For additional information about growth conditions in groups we refer to [39, Chapter 12].

2. Growth of graphs and preliminary results

The *growth function* of a locally finite graph X with respect to a vertex $x \in V(X)$ is given by

$$f_X(x, n) = |\{y \in V(X) \mid d(x, y) \leqslant n, 0 \leqslant n\}|,$$

where $d(x, y)$ denotes the distance between x and y. For transitive graphs the growth function obviously does not depend on a particular vertex x, hence we denote it by $f_X(n)$. Of course we can then define *polynomial*, *intermediate* and *exponential growth* analogously to the above definitions for groups. Even in the case of nontransitive graphs these properties do not depend on a particular vertex

x. Only the constants in the bounds for the growth function might be different if we consider the growth function with respect to different vertices. We also mention that we can identify the growth function of a group G, with respect to a generating set H, $e \notin H$, and that of its *Cayley graph* $C(G, H)$, where $C(G, H)$ is defined on G with the edge-set

$$E(C(G, H)) = \{(g, gh) \mid g \in G, h \in H \cup H^{-1}\}.$$

It is easily seen that G acts on $C(G, H)$ by left multiplication. This action is *regular*, i.e. transitive and fixed point free. Conversely a graph X is a Cayley graph of a group G if G acts regularly on X. This result is due to Sabidussi [29]. If the action of G on X is *semiregular*, i.e. fixed point free, one can contract X to a Cayley graph of G, as has been shown by Babai [2]. Both of these results are repeatedly used in the proofs of the results presented here.

By the above it is clear that the most natural examples of graphs satisfying special growth conditions are Cayley graphs of groups with polynomial, intermediate and exponential growth. In fact, one might be tempted to consider results about group growth as results about Cayley graphs. We note that the graphs in Fig. 1 have linear, quadratic and exponential growths, respectively.

The first question, concerning growth functions of locally finite transitive graphs with polynomial growth which immediately arises is, if a relation like (1) also holds for graphs which are not Cayley graphs. The answer to this question, as well as to other problems concerning graphs with polynomial growth, was given by Trofimov [38]. To state the result we need several definitions.

If a group G acts transitively on a graph X, then an *imprimitivity system* of G on X is a partition τ of $V(X)$ into subsets called *blocks*, such that every element of G is permutes the blocks of τ. Among imprimitivity systems we include the partition of $V(X)$ into singletons and into $V(X)$ itself. These are the so called *trivial imprimitivity systems*. The *quotient graph* X_τ is defined as follows: $V(X_\tau)$ is

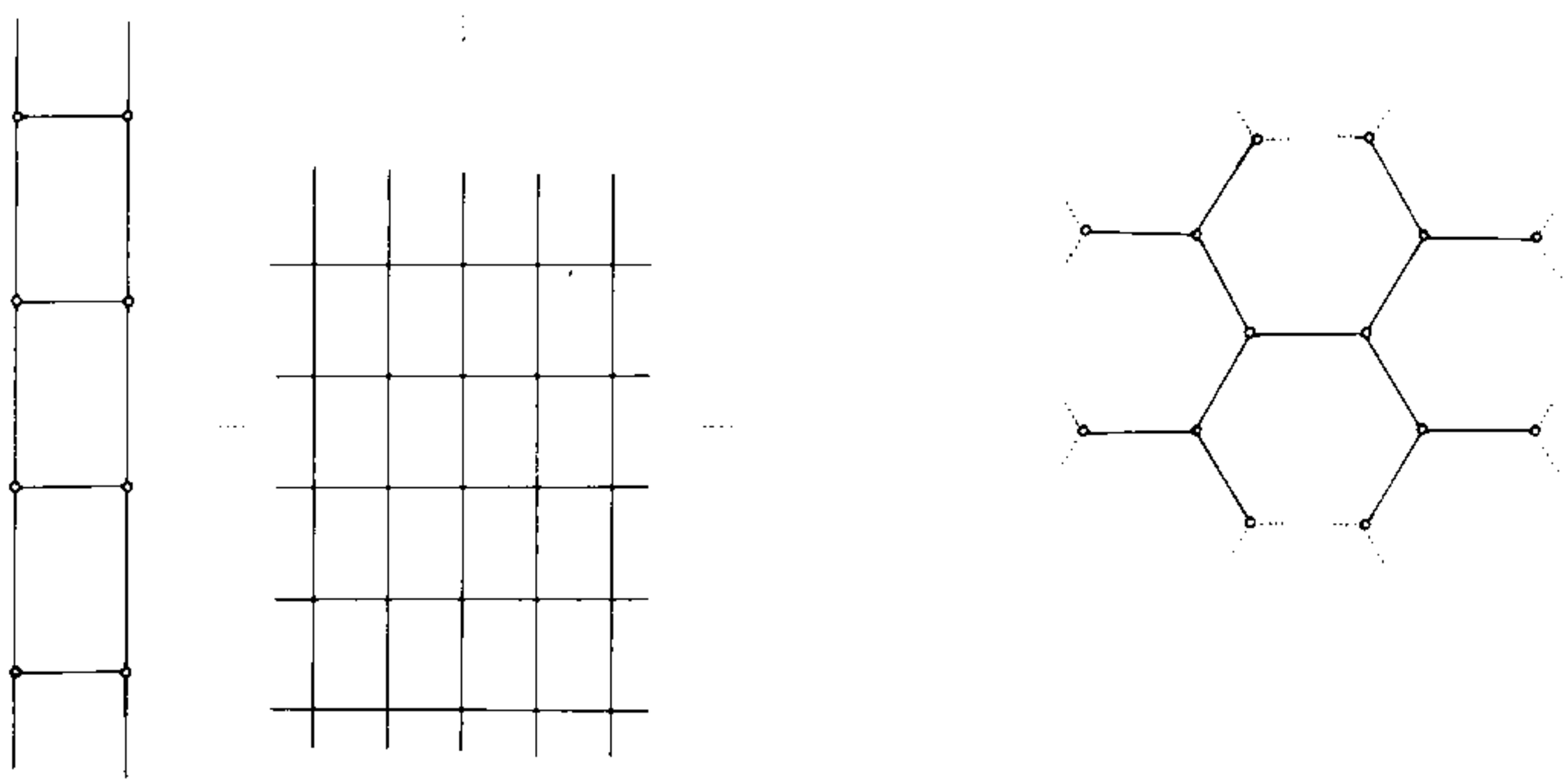


Fig. 1.

the set of blocks and two vertices v_τ, $w_\tau \in V(X_\tau)$ are adjacent in X_τ if and only if $(v, w) \in E(X)$ for at least two vertices $v \in v_\tau$, $w \in w_\tau$. By G_τ we denote the group acting on X_τ which is induced by G. It is a homomorphic image of G and $G_\tau \leqslant \mathrm{AUT}(X_\tau)$.

Theorem 2.1 (Trofimov [38]). *The following assertions are equivalent for an infinite transitive connected locally finite graph X:*

(i) *X has polynomial growth.*

(ii) *There exists an imprimitivity system τ of* $\mathrm{AUT}(X)$ *on $V(X)$ with finite blocks such that* $\mathrm{AUT}(X_\tau)$ *is a finitely generated almost nilpotent group and the stabilizer in* $\mathrm{AUT}(X_\tau)$ *of a vertex of the graph X_τ is finite.*

Theorem 2.1 combined with the following result of Sabidussi [30] shows that transitive graphs with polynomial growth are closely related to Cayley graphs of almost nilpotent groups.

First a definition. If X is a graph and n a cardinal, then the graph nX is defined on the vertex-set $V(nX) = V(X) \times N$, where N is a set of cardinality n, and

$$E(nX) = \{[(x, \mu), (y, \nu)] \mid [x, y] \in E(X), \mu, \nu \in N\}.$$

Theorem 2.2 (Sabidussi [30]). *Let X be a connected transitive graph, G a transitive subgroup of* $\mathrm{AUT}(X)$ *and let n be the cardinaltiy of the stabilizer in G of a vertex of X. Then nX is a Cayley graph of G.*

Now, let X be a connected transitive graph with polynomial growth. Then X and X_τ (as given by Theorem 2.1) obviously have the same growth degree. Also X_τ and nX_τ, where n is the cardinality of the stabilizer of a vertex of X_τ in $\mathrm{AUT}(X_\tau)$, have the same growth degree. Since, by Theorem 2.2, nX_τ is a Cayley graph of the almost nilpotent group $\mathrm{AUT}(X_\tau)$, the results mentioned in Section 1 imply that the growth degree of nX_τ is a well defined integer d. Hence, also the growth degree of X_τ is the same integer d and since the blocks of τ are finite, the same also holds for the growth degree of X.

Although these remarks show that the automorphism groups of graphs with polynomial growth are in some sense closely related to groups with polynomial growth, there are still many open questions. It is, for example, possible that the automorphism group of a connected locally finite graph X with polynomial growth is uncountable. A simple example being given by Fig. 2. The uncoun-

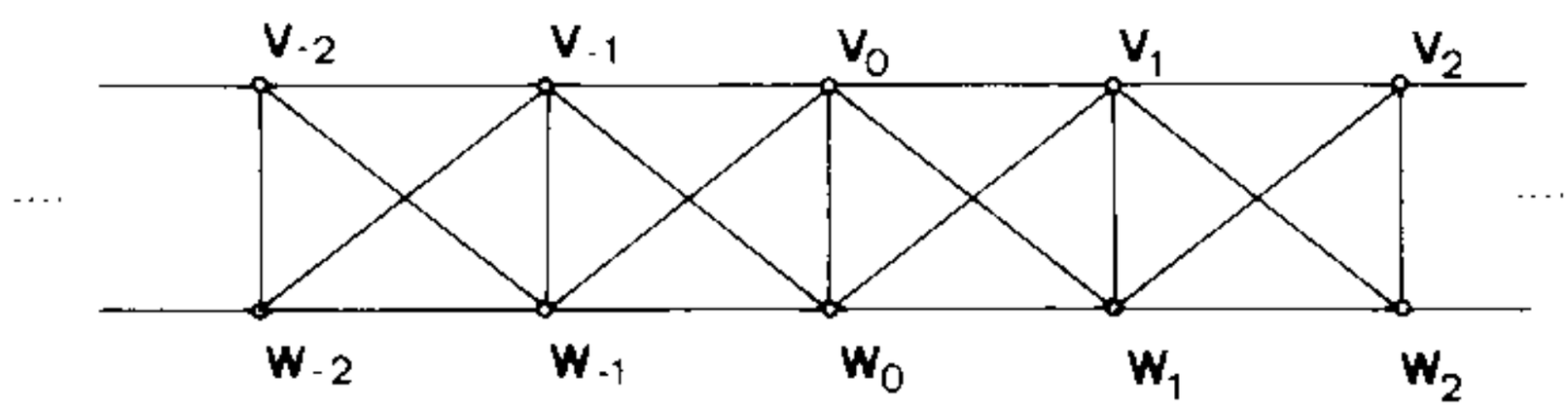


Fig. 2.

tability of the automorphism group of this graph is immediate. In general we have the following criterion by Halin.

Theorem 2.3 (Halin [15]). *The automorphism group of a locally finite connected graph X is uncountable if and only if for every finite subset $F \subset V(X)$ there exists a nontrivial automorphism that fixes F pointwise.*

The graph of Fig. 2 has another interesting property: Let a be the automorphism which permutes v_0 and w_0 but fixes all other vertices of X. By b we denote the automorphism which acts with the orbits $(\ldots, v_{-1}, v_0, v_1, \ldots)$ and $(\ldots, w_{-1}, w_0, w_1, \ldots)$ on X. We now show that the group G generated by a and b is metabelian (solvable of length two) and has exponential growth.

Let $G' = [G, G]$ be the commutator subgroup of G and let $g, h \in G'$. It is obvious that every element of G' leaves invariant the sets

$$\ldots, V_{-1} = \{v_{-1}, w_{-1}\}, \qquad V_0 = \{v_0, w_0\}, \qquad V_1 = \{v_1, w_1\}, \ldots$$

The action of g and h on those sets is induced by the action of a on V_0. Hence the restrictions of g and h to V_i commute for all V_i, $i \in \mathbb{Z}$. So G' is abelian and $[G', G'] = \{e\}$.

Let $g_j = b^j a b^{-j}$ for $j = 0, \ldots, n$, $n \in \mathbb{N}$. We now consider the automorphisms

$$h(\varepsilon_0, \ldots, \varepsilon_n) = g_0^{\varepsilon_0} \cdots g_n^{\varepsilon_n}$$

where $\varepsilon_j \in \{0, 1\}$. Clearly $h(\varepsilon_0, \ldots, \varepsilon_n) = h(\delta_0, \ldots, \delta_n)$ if and only if $\varepsilon_0 = \delta_0, \ldots, \varepsilon_n = \delta_n$. So there are 2^{n+1} different h's. Since the length of those automorphisms with respect to the generators a and b is less or equal to $(2n+1)(n+1)$ we obtain

$$f_G((2n+1)(n+1)) \geqslant 2^{n+1}$$

for $n = 0, 1, \ldots$. This implies that G does not have polynomial growth. Since G is solvable it thus has exponential growth.

This leads to the following problems:

(1) Groups with exponential growth can act transitively on graphs with polynomial growth. Does this also hold for groups with intermediate growth?

(2) Is it possible to give necessary and sufficient algebraic conditions such that the automorphism groups of graphs with polynomial growth are uncountable?

(3) Is it true that uncountable automorphism groups of graphs with polynomial growth always contain finitely generated solvable subgroups with exponential growth?

Before we present some answers to these questions, we continue with a characterization of growth conditions in the language of ends of graphs.

Two one-way infinite paths P and Q are *equivalent* in X, in symbols $\sim_x$, if there is a third path R which meets both of them infinitely often. The equivalence classes with respect to $\sim_x$ are called *ends*. Obviously the automorphisms of X also act on the set of ends of X. This definition of ends is due to Halin [16], but we

emphasize that the concept of ends first appears in papers of Hopf [17] and Freudenthal [9]. They introduced the concept of ends to study discrete groups. For us an end of a group G is an end of a Cayley graph of G. We note that the number of ends of a group (and a transitive graph in general) is either 0 (if the group is finite), 1, 2 or ∞. This number does not depend on the generating set although the number of disjoint one-way infinite paths in each end depends on it if this number is finite!

The fact that a transitive infinite locally finite graph has either 1, 2 or infinitely many ends immediately follows from results in [15, 21]. This, together with another result of Halin [16], readily implies that graphs with infinitely many ends have exponential growth. Also a transitive infinite locally finite graph has linear growth if and only if it has two ends (cf. [20]). (For a detailed description of those graphs see [20, 22].) If a graph has polynomial but nonlinear growth it thus has one end, but the converse is not true: All Cayley graphs of groups with intermediate growth clearly have one end and there also exist graphs with exponential growth and only one end.

Nevertheless, this relation between graphs with polynomial growth and graphs with one or two ends allows the application of many results which where formulated using the concept of ends rather than that of growth. Results of this kind can be found in [15–16, 20–22, 34, 40–41].

3. Automorphism groups of graphs with polynomial growth

An automorphism g of a graph X is called *bounded* if there is a constant k, depending upon g, such that

$$d(x, g(x)) \leqslant k \quad \text{for every } x \in V(X).$$

This concept is closely related to the concept of FC-groups. (A group G is an *FC-group* if the conjugacy class of every $g \in G$ is finite. For the theory of FC-groups we refer to [35].) For, let X be a Cayley graph of a group G such that every $g \in G$ acts as bounded automorphism on X. Then $d(x, gx) \leqslant k$ for every $x \in V(X) = G$ which is equivalent with $d(e, x^{-1}gx) \leqslant k$. But this only holds if $\{x^{-1}gx \mid x \in G\}$ is finite, i.e. if G is an FC-group.

We can then apply one of the results about FC-groups ([27, Theorem 5.1]) to prove the following useful theorem. We recall that a group G is *locally finite* if every finitely generated subgroup of G is finite.

Theorem 3.1 (Godsil, Imrich, Seifter, Watkins and Woess [10]). *Let X be a transitive connected locally finite graph with polynomial growth, and let $B(X)$ be the group of bounded automorphisms of X. Then the set $B_0(X)$ of elements of finite order in $B(X)$ forms a normal subgroup of* AUT(X). *It is locally finite, periodic and acts with finite orbits on $V(X)$.*

This result together with the following nice theorem of Rosset [28], enables us to solve the first problem posed in Section 2.

Theorem 3.2 (Rosset [28]). *If a finitely generated group G has non-exponential growth and H is a normal subgroup of G such that G/H is solvable, then H is finitely generated.*

Theorem 3.3 (Seifter [33]). *Finitely generated groups with intermediate growth cannot act transitively on connected locally finite graphs with polynomial growth.*

Proof. Suppose a finitely generated group G with intermediate growth acts transitively on a connected locally finite graph X with polynomial growth. By τ we denote the imprimitivity system of $\mathrm{AUT}(X)$ on X which is given by Theorem 2.1. Then a homomorphic image G_τ of G also acts transitively on X_τ. By φ we denote the homomorphism from G onto G_τ. From the remarks following Theorem 2.2 we also know that G_τ has the same growth degree as X_τ. Hence $\ker\varphi$ must be infinite for otherwise G also has polynomial growth of the same degree as G_τ. Theorem 1.1 now implies that G_τ contains a nilpotent normal subgroup N_τ of finite index. Since φ is a homomorphism, the subgroup N of G which is generated by all $g \in G$ with $\varphi(g) \in N_\tau$ also has finite index in G. Hence N also is a finitely generated group with intermediate growth. Since $\ker\varphi$ also is a normal subgroup of N, Theorem 3.2 implies that $\ker\varphi$ is finitely generated. But $\ker\varphi \leqslant B_0(X)$ also holds. Hence $\ker\varphi$ must be locally finite by Theorem 3.1, a contradiction. □

This result leads to the following necessary and sufficient algebraic condition for the uncountability of automorphism groups of graphs with polynomial growth.

Corollary 3.4 (Seifter [33]). *The automorphism group of a connected locally finite transitive graph X with polynomial growth is uncountable if and only if it contains a finitely generated subgroup with exponential growth which acts transitively on X.*

If a group with exponential growth acts transitively on X then it is clear that $\ker\varphi$ must be infinite, where φ is defined as in the proof of Theorem 3.3. Then Theorem 2.3 immediately implies that the automorphism group of X is uncountable. The 'only if' part of the proof is a little bit more complicated. We only mention that it mainly depends on Theorem 3.3. For the complete proof we refer to [33].

We recall (see also Section 1) that finitely generated solvable groups have either polynomial or exponential growth. Theorem 3.3 shows that the same holds for finitely generated groups which act transitively on graphs with polynomial growth. Hence it is quite natural to ask if every finitely generated group G which acts transitively on a graph with polynomial growth, contains a solvable subgroup

with the same growth properties. It clearly holds if G has polynomial growth and the example given in Section 2 indicates that this might also be true if G has exponential growth. But this turned out to be a quite difficult question. We only have the following partial answer which is a generalization of the example given by Fig. 2.

Theorem 3.5 (Seifter [32]). *Every finitely generated group with exponential growth which acts transitively on a locally finite connected graph X with linear growth contains a finitely generated metabelian subgroup with exponential growth which acts with finitely many orbits on X.*

For graphs with polynomial but nonlinear growth we still do not know if a similar result holds but we think that this should be the case. Hence the following.

Conjecture 3.6. Suppose the finitely generated group G with exponential growth acts transitively on a graph X with polynomial growth. Then G contains a finitely generated solvable subgroup with exponential growth which acts with finitely many orbits on X.

Graphs with linear growth are almost always easier to handle than graphs with higher growth degree. The reason, roughly spoken, is the fact that graphs with linear growth are the only transitive graphs with polynomial growth which have two ends. In the next paragraph we present a rather detailed analysis of groups and graphs with linear growth.

First we prove another interesting property of graphs with polynomial growth. We also give the proof of this result since the construction of suitable imprimitivity systems is one of the main techniques used for proving the results presented in this paper.

Theorem 3.7 (Godsil, Imrich, Seifter, Watkins and Woess [10]). *Let X be a locally finite connected transitive graph with polynomial growth. Then the action of* $\mathrm{AUT}(X)$ *on X is imprimitive.*

Proof. If the blocks of the imprimitivity system τ given by Theorem 2.1 contain more than one vertex, the assertion obviously holds. So we can assume that τ is trivial. By Theorem 2.1 this means that $\mathrm{AUT}(X)$ is almost nilpotent.

If we can find a nontrivial normal subgroup K of $\mathrm{AUT}(X)$ which does not act transitively on X, then the orbits of K on X give rise to a nontrivial imprimitivity system of $\mathrm{AUT}(X)$ on X.

By Theorem 3.1 $B_0(X)$ is a normal subgroup of $\mathrm{AUT}(X)$ which acts with finite orbits on X. Hence if $B_0(X) \neq \{e\}$ we already have a nontrivial imprimitivity system for $\mathrm{AUT}(X)$. Also $B(X)$ is normal in $\mathrm{AUT}(X)$. If $B(X) \neq \{e\}$ does not

act transitively on X we again have a nontrivial imprimitivity system. Since $B(X) \neq \{e\}$ always holds, we now have to consider the case that $B(X)$ acts transitively on X.

It follows from [37], Theorem 1, that $B(X) \cong \mathbb{Z}^d$ if $B_0(X) = \{e\}$ and $B(X)$ acts transitively on X, where d denotes the growth degree of X. Then $G = (2\mathbb{Z})^d$ is a subgroup of finite index in $B(X)$ which does not act transitively on X. Since $B(X)$ has the same growth degree as X, it also has the same growth degree as $\mathrm{AUT}(X)$ (cf. the remarks following Theorem 2.2). So $B(X)$ has finite index in $\mathrm{AUT}(X)$ which implies that G also has finite index in $\mathrm{AUT}(X)$. Furthermore, the intersection of all conjugates gGg^{-1}, $g \in \mathrm{AUT}(X)$, is a normal subgroup N of finite index of $\mathrm{AUT}(X)$. Clearly $N \leqslant G$ also does not act transitively on X, which completes the proof. □

4. Groups and graphs with linear growth

In [43] Wilkie and van den Dries proved the following interesting relation between the first difference function of the growth function of a group G and the index of a subgroup $\cong \mathbb{Z}$ in G.

Theorem 4.1 (Wilkie, van den Dries [43]). *Suppose G is a finitely generated infinite group, $k > 0$, and $f_G(k) - f_G(k-1) \leqslant k$. Set $c = f(k) - f(k-1)$. Then G contains a subgroup $\cong \mathbb{Z}$ of index $\leqslant c^4/2$.*

This result shows that a local condition like $f_G(k) - f_G(k-1) \leqslant k$ implies that G satisfies the global condition of linear growth. But as Wilkie and van den Dries themselves suggested, the bound for the index of a subgroup $\cong \mathbb{Z}$ in G is far from being sharp. This bound can be improved as follows.

Theorem 4.2 (Imrich, Seifter [19]). *Let G satisfy the assumptions of Theorem 4.1. Then G contains a subgroup $\cong \mathbb{Z}$ of index $\leqslant c$. This bound is sharp.*

The proof is based on the result of Wilkie and van den Dries [43] and the characterization of groups with two ends (cf. [34, 4.A.6.5]). By this characterization groups with two ends are exactly those with a homomorphism with finite kernel onto $\mathbb{Z}$ or onto the free produce $\mathbb{Z}_2 * \mathbb{Z}_2$. This implies that $H \cup H^2$ always contains an element of infinite order if H is a finite generating set of a group G with two ends. Based on these facts Theorem 4.2 was proved by a detailed discussion of properties of Cayley graphs with two ends.

To show that the bound is sharp we note that $\langle ab \rangle \cong \mathbb{Z}$ has index 2 in $\langle a, b \mid a^2 = b^2 = e \rangle = \mathbb{Z}_2 * \mathbb{Z}_2$. In this case $c = 2$ holds. Furthermore, let A be any finite group and let G denote the direct product of A by $\langle a, b \mid a^2 = b^2 = e \rangle$. If we

set $H = A \cup \{a, b\} \backslash e$ then H generates G and $c = 2\,|A|$ obviously holds. Also $\langle ab \rangle$ has index $2\,|A|$ in G.

Theorem 4.2 provides a rather detailed characterization of groups with linear growth. We now consider graphs with linear growth. The example given in Section 2, as well as Theorem 3.5, show that groups with exponential growth can act transitively on graphs with linear growth. On the other hand the remarks following Theorem 2.2 imply that a group with polynomial growth must have linear growth if it acts transitively on a graph with linear growth. Hence the question arises how close the connections between the growth functions of graphs with linear growth and finitely generated groups which have polynomial growth and act transitively on them must be. More precisely: Let X be a graph such that $f_X(n) \leqslant c_X n$ holds for some constant c_X. If a finitely generated group G with polynomial growth acts transitively on X, Theorem 2.1 implies that $f_G(n) \leqslant c_G n$ also holds for some constant c_G. If G acts regularly on X we know (cf. [29]) that X is a Cayley graph of G. Hence there is a generating set H of G such that $f_G(n) \leqslant c_X n$. But what happens if G does not act regularly on X? Is it still possible to find a generating set H of G and a function $h(c_X)$ such that the growth function of G is bounded by $h(c_X)n$?

Since we are now interested in the constant in the bound of the growth function, which clearly depends on the generating set, we write $f_G(H, n)$ for the growth function of a group. We also mention that a *translation* of a graph X is an automorphism which acts without finite orbits on X (sometimes (see e.g. [15]) translations are called automorphisms of type 2). This question can be answered as follows.

Theorem 4.3 (Seifter [31]). *Let the finitely generated group G with polynomial growth act transitively on X with $f_X(n) \leqslant c_X n$ for all $n \geqslant 1$. Then there exists a minimal integer k and a translation g which is central in $B(X) \cap G$ and $d(v, g(v)) \leqslant k$ for all $v \in V(X)$. Furthermore there is a generating set H of G such that*

$$f_G(H, n) < 2c_X(\lfloor c_X \rfloor !)^k n.$$

If, in addition, $\mathrm{AUT}(X)$ is a countable group the following characterization of $\mathrm{AUT}(X)$ can be given.

Theorem 4.4 (Seifter [31]). *Let X be a graph with $f_X(n) \leqslant c_X n$, $n \geqslant 1$, and countable automorphism group. Then there exists a finite generating set H of* $\mathrm{AUT}(X)$ such that

$$f_{\mathrm{AUT}(X)}(H, n) \leqslant 2c_X \left\lfloor \frac{c_X}{2} - 1 \right\rfloor ! \, n$$

holds for all $n \geqslant 1$.

Outline of the proofs of Theorems 4.3 and 4.4: One has to determine the cardinality of the stabilizer of a vertex of X in the considered group. Then an application of Theorem 2.2 immediately leads to these results.

Also, the following information about the structure of graphs with linear growth is available.

Proposition 4.5 (Imrich, Seifter [20]). *A transitive connected locally finite graph X has linear growth if and only if it has two ends. Then it is spanned by finitely many 2-way infinite paths and there is a translation which leaves these paths invariant.*

Even if a graph with linear growth is not transitive very much is known about its structure and automorphism group. We refer to [15, 22] but emphasize that the concept of growth is not used in these papers. In [22] graphs with linear growth are called *strips*, in [15] they are simply graphs with two ends and a translation acting on them.

Finally we emphasize that such detailed results cannot be proved for graphs with polynomial growth in general. Even groups with polynomial but nonlinear growth cannot be characterized as in Theorems 4.1 and 4.2. For, let G be the direct product of $\mathbb{Z}^d = \{a_1, \ldots, a_d \mid a_i a_j = a_j a_i \text{ for all } i, j \in \{1, \ldots, d\}, d \geqslant 2\rangle$, by the group $\mathbb{Z}_3 = \langle b \mid b^3 = e \rangle$. Then we have:

$$f_G(0) = 1, \qquad f_G(1) = 2d + 3, \qquad f_G(2) = 2d^2 + 6d + 3.$$

Hence, if d is sufficiently large, $f_G(2) - f_G(1) \leqslant 2^{d^*}$ always holds for some $d^* < d$. But this does not imply that G has growth degree d^*.

5. *S*-transitivity

A sequence $(v_0, \ldots, v_s)$ of $s+1$ vertices is called an *s-arc* if for each i, (v_{i-1}, v_i) is an edge of X and $v_{i-1} \neq v_{i+1}$. If a group acts transitively on the s-arcs of X we call X *s-transitive*. For completeness we mention that the so far used term 'transitive' in this context means 0-*transitive*.

In 1981 Weiss [42] showed that finite graphs with valency $\geqslant 3$ cannot be 8-transitive. Surprisingly the same holds for graphs with polynomial growth.

Theorem 5.1 (Seifter [33]). *Let X be a connected locally finite s-transitive graph with polynomial growth and valency at least 3. Then $s \leqslant 7$.*

Outline of the proof. Let X be an s-transitive graph with polynomial growth. Then the graph $Y = X_\tau$ (see Theorem 2.1) is also s-transitive. In addition we know that AUT(Y) is a finitely generated almost nilpotent group. It is then always possible to find an imprimitivity system σ of AUT(Y) on Y with infinite blocks such that Y_σ is finite and also s-transitive. With the exception of the case

when Y_σ is a cycle we can then apply Weiss's result [42]. Also the case that Y_σ is a cycle causes no difficulties (cf. [33]).

The following construction shows that the bound given in Theorem 5.1 is sharp.

Let Y be a finite graph and let G be a group. In [5, p. 127], a construction of a covering graph of Y with respect to G is given as follows: Each edge (u, v) of Y gives rise to two 1-arcs, (u, v) and (v, u). By $S(Y)$ we denote the set of 1-arcs and by $\varphi: S(Y) \to G$ a mapping such that $\varphi(u, v) = (\varphi(v, u))^{-1}$ for all $(u, v) \in S(Y)$. The covering graph $\tilde{Y} = \tilde{Y}(G, \varphi)$ of Y with respect to G is defined on the vertex-set $V(\tilde{Y}) = G \times V(Y)$ and two vertices $(g_1, u), (g_2, v) \in V(\tilde{Y})$ are joined by an edge if and only if $(u, v) \in S(Y)$ and $g_2 = g_1\varphi(u, v)$.

Let Y be a finite s-transitive graph with valency at least 3 where $0 \leqslant s \leqslant 7$. The fundamental group $\Pi(Y, u)$ of Y at u is a free group with the edges of $E(Y) \backslash E(T)$ as generators, where T is a spanning tree of Y (cf. [18]). We can also regard $\Pi(X, u)$ as a group generated by the closed walks with base point u which are associated with the edges of $E(Y) \backslash E(T)$. For details we again refer to [18].

We now consider the group $A = \Pi(X, u)/[\Pi(X, u), \Pi(X, u)]$. It is a free finitely generated abelian group. The generators $a_1, \ldots, a_\tau, a_1^{-1}, \ldots, a_\tau^{-1}$ of A are the images of the generators $g_1, \ldots, g_\tau, g_1^{-1}, \ldots, g_\tau^{-1}$ of $\Pi(X, u)$ under the homomorphism from $\Pi(X, u)$ onto A. Hence $A \cong \mathbb{Z}^\tau$. Let T be a spanning tree of Y and let φ, $\varphi: Y \to A$, be a mapping which maps those 1-arcs of Y which correspond to edges of $E(X) \backslash E(T)$ onto the generators of A and the edges of T onto the unit element of A. It is easy to see that the covering graph $\tilde{Y}(A, \varphi)$, constructed accordingly to the above rules is connected and locally finite. It also has polynomial growth of degree τ.

We can then apply Theorems 3 and 4 of [7] to show that $\tilde{Y}(A, \varphi)$ is s-transitive if Y is s-transitive. (For details we refer to [11, Proposition 2.3].) Hence, Theorem 5.1 gives a sharp bound for graphs with polynomial growth in general, but there are classes of graphs with polynomial growth with stronger restrictions on s.

Theorem 5.2 (Seifter [33]). *Let X be a connected locally finite graph with valence at least three and let $B(X)$ act transitively on X. Then X cannot be 3-transitive.*

In Theorem 5.2 one does not have to assume that X has polynomial growth since this follows from [37], Theorem 1, if $B(X)$ acts transitively on X. We also mention that it is possible to find a better bound for s if the considered graphs have linear growth (cf. [33]).

Not only s-transitivity is restricted by properties of the automorphism group, to some extent the converse also holds.

Theorem 5.3 (Seifter [33]). *Let X be a connected locally finite s-transitive graph with polynomial growth, valency at least three and let $s \geqslant 2$. Then* AUT(X) *is a finitely generated almost nilpotent group.*

Finally we wish to pose the following problem: The smallest known 7-transitive finite graph has 728 vertices and valency 4 (cf. [4], see also [5, p. 164]). Hence the rank of its fundamental group is equal to 729. By the above construction, we only know that a 7-transitive graph with polynomial growth of degree 729 exists. But we think that it should be possible to find e.g. 7-transitive graphs of much smaller growth degree.

Problem 5.4. *Let* $2 \leqslant s \leqslant 7$. Determine the minimal integer $d(s)$, depending upon s, such that there exists a connected locally finite s-transitive graph with polynomial growth of degree $d(s)$.

6. Covering graphs

As we have seen in Section 5 graphs with polynomial growth and high transitivity can be constructed as covering graphs of finite s-transitive graphs. As the next result shows, all graphs with polynomial growth are covering graphs of finite graphs.

Let $p, p: X_1 \to X_2$, denote a homomorphism and let $S(v)$, $v \in V(X_1)$, denote the star consisting of v and all edges incident to v. If $p(S(v))$ is isomorphic to $S(v)$, for every $v \in V(X_1)$, we call p *a covering map* and X_1 a covering graph of X_2.

Theorem 6.1 (Godsil, Seifter [11]). *Let X be a connected locally finite s-transitive, $s \geqslant 0$, graph with polynomial growth. Then there exist infinitely many finite graphs $Y_1, Y_2, \ldots$ such that:*

(1) *X is covering graph of every Y_i, $i \geqslant 1$, and*
(2) *each Y_k, $k \geqslant 2$, is covering graph of the graphs $Y_1, \ldots, Y_{k-1}$.*
(3) *If in addition $s \geqslant 2$ holds, then each Y_i also is at least s-transitive.*

We emphasize that Theorem 6.1 does not exclude the possibility that the finite graphs $Y_1, Y_2, \ldots$ are all s-transitive if $s \leqslant 1$. As the following example shows, some of those finite graphs can even be 'more' transitive than their covering graph with polynomial growth.

Let X be the Cayley graph $C(G, H)$ of the group $G \cong \mathbb{Z} \times \mathbb{Z}_5$ with respect to the generating set $H = \{a, c\}$ where $\langle a \rangle \cong \mathbb{Z}$ and $\langle c \rangle \cong \mathbb{Z}_5$. We denote the vertices of G by (a^j, c^i), $j \in \mathbb{Z}$, $0 \leqslant i \leqslant 4$ (see Fig. 3). Obviously every set $T_j = \{(a^j, c^0), \ldots, (a^j, c^4)\}$, $j \in \mathbb{Z}$, separates the two ends of X. Furthermore the T_j are minimal with respect to this property. We now assume that X is s-transitive for some $s \geqslant 1$. Then there is a $g \in \mathrm{AUT}(X)$ which maps the edge $((a, e), (a, c))$ onto the edge $((a, e), (a^2, e))$. But then the set $g(T_0)$ cannot separate the ends of X, a contradiction. Hence X is not 1-transitive.

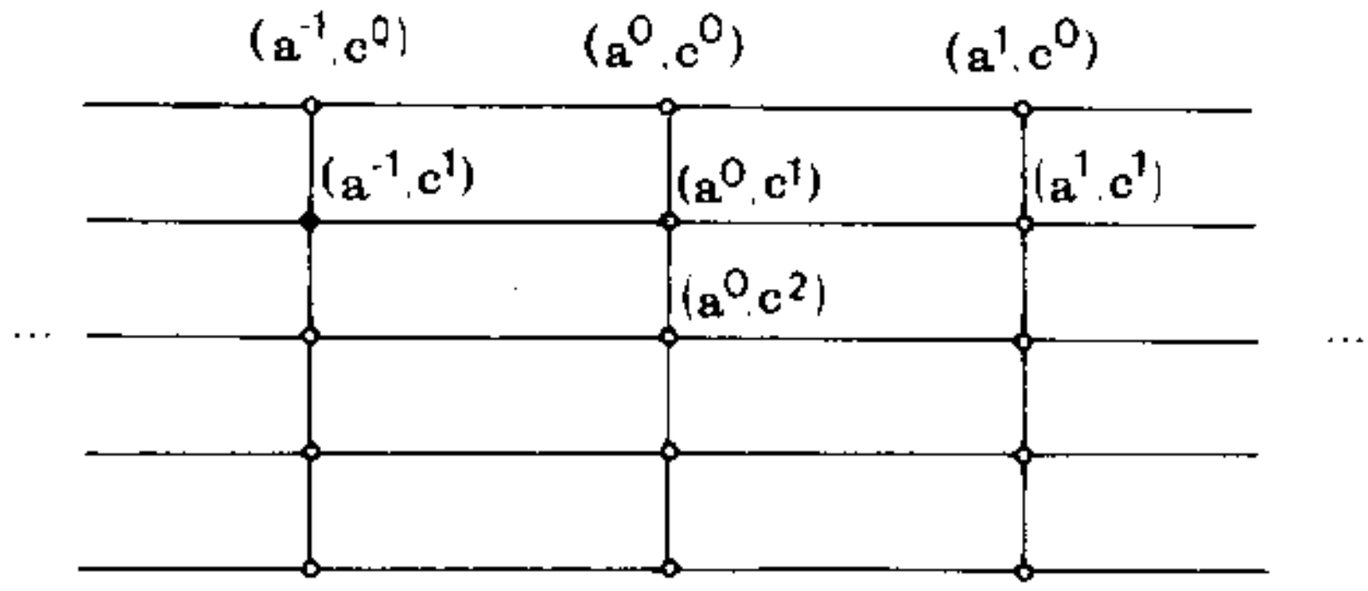


Fig. 3.

Let $\{v_0, \ldots, v_4\}$ denote the vertex set of the complete graph K_5 and let p, $p: X \to K_5$, be a map satisfying:

(a) $p(a^j, c^i) = v_k$ if $j \geqslant 0$ and $i \equiv (k + 2j)(5)$, and,

(b) $p(a^j, c^i) = v_k$ if $j < 0$ and $i \equiv (k + |3_j|)(5)$.

Obviously p is a covering map and K_5 is 2-transitive.

The graphs C_1 and C_2 given by Fig. 4 are Cayley graphs of $\mathbb{Z}_5 \times \mathbb{Z}_5$ and $\mathbb{Z}_3 \times \mathbb{Z}_5$, respectively. Obviously X also covers these graphs. The graph C_1 is clearly 1-transitive but the 2-arc (w_0, w_1, w_2) cannot be mapped onto the 2-arc (w_0, w_1, w_3) since there are two paths of length 2 from w_0 to w_2 but only one path of length 2 from w_0 to w_3. Hence it is not 2-transitive.

The edge $(v_0, v_1) \in E(C_2)$ is contained in the triangle (v_0, v_1, v_2, v_0) but there is no triangle which contains the edge (v_0, v_3). Hence C_2 is not 1-transitive.

If we now choose $Y_i = \mathbb{Z}_5 \times \mathbb{Z}_i$ for suitable i we again obtain an infinite sequence of finite graphs which all satisfy condition three. The reason is that the automorphism group of X is countable, which need not hold for s-transitive graphs with $s \leqslant 1$ in general. For $s \geqslant 2$ this always holds by Theorem 5.3.

As the above example shows, it can happen that some of the Y_i have higher transitivity than their covering graphs with polynomial growth. Of course this is impossible if X is 7-transitive. In this case the result of Weiss [42] implies that all Y_i are also 7-transitive. It would be interesting, but also rather difficult, to establish conditions such that the Y_i have the same transitivity as X if $s < 7$.

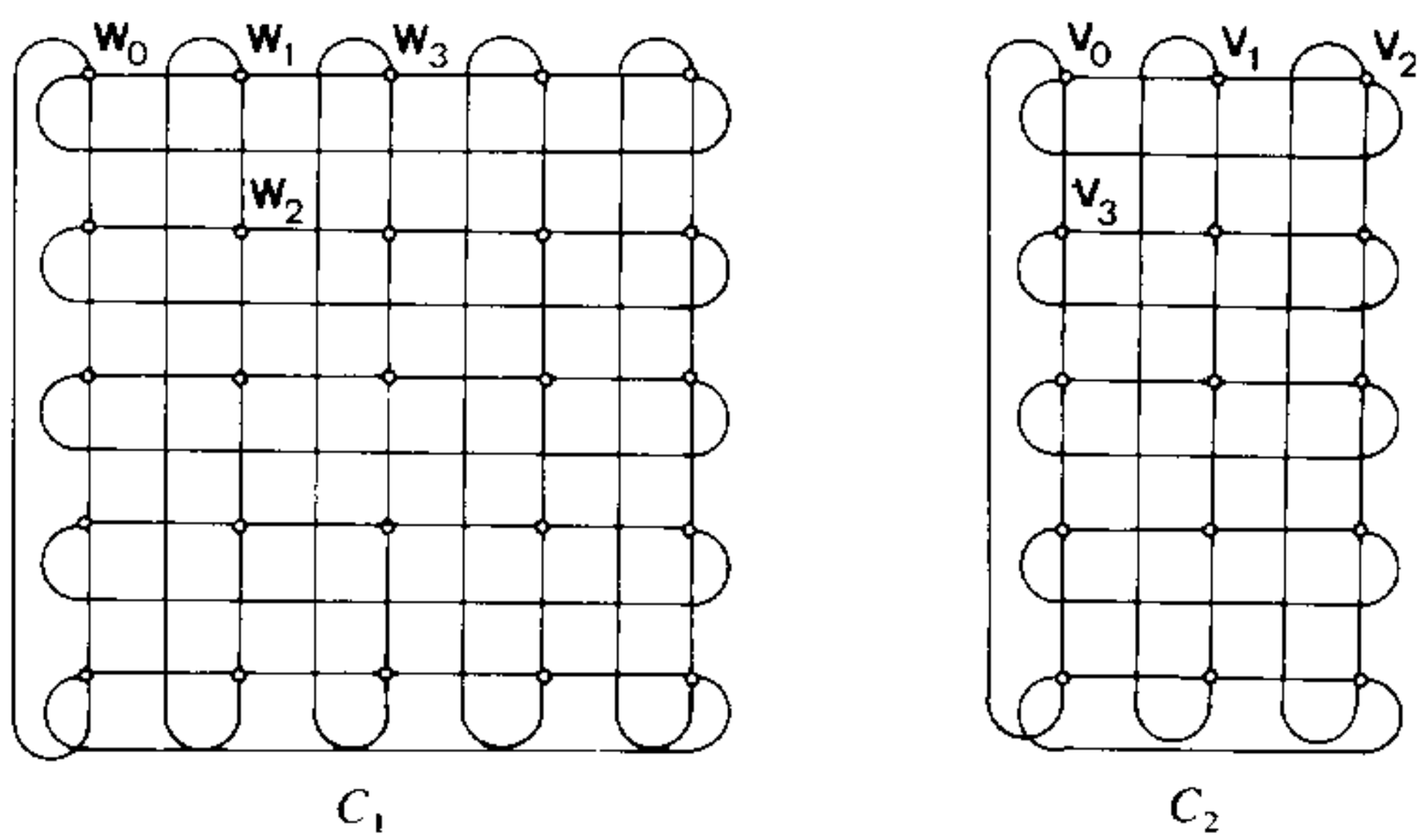


Fig. 4.

Problem 6.2. Find other assumptions than the countability of AUT(X) such that Theorem 6.1, (3), also holds for $s \leqslant 1$.

Problem 6.3. Let X be a locally finite connected s-transitive graph with polynomial growth and let $s \leqslant 6$. Is it always possible to find a sequence of finite graphs which satisfy all assertions of Theorem 6.1 and in addition have the property that they are all exactly s-transitive?

7. Automorphism groups as topological groups

In this paragraph we present parts of [36]. In that paper not only graphs with polynomial growth, but infinite connected locally finite graphs in general are considered. Since we assume that this topological approach might also be useful for the investigation of graphs with polynomial growth, we include some of the results in [36] in this survey.

Let X be a connected locally finite graph and let $g_1, g_2 \in \mathrm{AUT}(X)$, $x \in V(X)$. We set:

(a) $\rho_x(g_1, g_2) = 0$ if $g_1 = g_2$,

(b) $\rho_x(g_1, g_2) = n$ if $d(x, g_1^{-1}g_2(x)) = n \in \mathbb{N}$, and,

(c) $\rho_x(g_1, g_2) = 2^{-n-1}$ if the set $\{y \in V(X) \mid d(x, y) \leqslant n\}$ is fixed by $g_1^{-1}g_2$ and $g_1^{-1}g_2$ does not fix all vertices at distance $n+1$ from x, where $n \geqslant 0$.

This left-invariant metric defines a topology which is used in [36]. A neighbourhood basis of this topology consists of the pointwise stabilizers of finite subsets of $V(X)$.

Simple compactness arguments then immediately lead to a reformulation (and extension) of Halin's 'uncountability-theorem' ([15, Theorem 6]; see also Theorem 2.3 of this paper).

Theorem 7.1 (Trofimov [36]). *If G is a closed subgroup of* AUT(X), *then either the stabilizer G_x of a vertex of X in G is finite or G_x, and hence G, is uncountable.*

In view of Theorem 7.1 the next result supplies a useful test for the finiteness of the stabilizer of a vertex.

Let $P = (v_0, v_1, \ldots)$ be a one-way infinite path of X. We call P a *geodesic* if $d(v_i, v_j) = |i-j|$ holds for all $v_i, v_j \in V(P)$. The Buseman function α_P, $\alpha_P : V(X) \to \mathbb{Z}$, is defined by

$$\alpha_P(v) = \lim_{i \to \infty} (d(v, v_i) - i), \qquad v \in V(X).$$

Furthermore let $B(P) = \{v \in V(X) \mid \alpha_p(v) \leqslant 0\}$.

Theorem 7.2 (Trofimov [36]). *Let G act transitively on the connected locally finite graph X. If $\bar{G}$ denotes the closure of G in* $\mathrm{AUT}(X)$ *the following assertions are equivalent*:

(1) *The stabilizer of a vertex of X in G is infinite.*

(2) *There is a $g \in \bar{G}$ and a geodesic P in X such that g fixes $B(P)$ pointwise.*

For further combinatorial theorems based on this topological approach we refer to [36]. Also in [44] some interesting results are shown, using a topological approach. Since geodesics seem to play some role concerning properties of the automorphism group we want to refer to [41], where many conditions for the existence of geodesics in graphs are proved.

8. Final remarks

We mention that graphs with polynomial growth also play some role with respect to spectra of infinite graphs. Concerning this topic we only refer to "A survey on spectra of infinite graphs" by Mohar and Woess [26].

One of the reasons for the above considerations was the existence of transitive graphs that are not Cayley graphs. In the case of graphs of polynomial growth it turned out that they are closely related to Cayley graphs. In the case of graphs with nonpolynomial growth it is an open problem whether and how closely they are related to Cayley graphs. Even in the case of graphs with infinitely many ends (and hence exponential growth), which have been investigated thoroughly by Dicks and Dunwoody [6] such a close relationship as in the case of polynomial growth has not been established.

References

[1] G.M. Adelson-Velsky and Y. A. Shreider, The Banach mean on groups (in Russian), Uspechi Mat. Nauk 12 (78) (1957) 131–136.

[2] L. Babai, Some applications of graph contractions, J. Graph Theory 1 (1977) 125–130.

[3] H. Bass, The degree of polynomial growth of finitely generated nilpotent groups, Proc. London Math. Soc. 25 (1972) 603–614.

[4] C.T. Benson, Minimal regular graphs of girth eight and twelve, Canad. J. Math. 18 (1966) 1091–1094.

[5] N. Biggs, Algebraic graph theory, Cambridge Tracts in Math. 67 (Cambridge Univ. Press, Cambridge, 1974).

[6] W. Dicks and M.J. Dunwoody, Groups acting on graphs, Cambridge Studies in Advanced Math. 17 (Cambridge Univ. Press, Cambridge, 1989).

[7] D.Z. Djokovic, Automorphisms of graphs and coverings, J. Combin. Theory Ser. B 16 (1974) 243–247.

[8] J. Fabrykowski and N. Gupta, On groups with sub-exponential growth functions, J. Indian Math. Soc. 49 (1985) 249–256.

[9] H. Freudenthal, Über die Enden Diskreter Räume und Gruppen, Comm. Math. Helv. 17 (1944) 1–38.

[10] C. Godsil, W. Imrich, N. Seifter, M.E. Watkins and W. Woess, A note on bounded automorphisms of infinite graphs, Graphs Combin., to appear.
[11] C. Godsil and N. Seifter, Graphs with polynomial growth are covering graphs, submitted.
[12] R.I. Grigorchuk, Degrees of growth of finitely generated groups and the theory of invariant means, Math. USSR Izvestiya 25 (2) (1985) 259–300.
[13] R.I. Grigorchuk, On Milnor's problem of group growth, Soviet Math. Dokl. 28 (1) (1983) 23–26.
[14] M. Gromov, Groups of polynomial growth and expanding maps, Inst. Hautes Etudes Sci. Publ. Math. 53 (1981) 53–78.
[15] R. Halin, Automorphisms and endomorphisms of infinite locally finite graphs, Abh. Math. Sem. Univ. Hamburg 39 (1973) 251–283.
[16] R. Halin, Über unendliche Wege in Graphen, Math. Ann. 157 (1964) 125–137.
[17] H. Hopf, Enden offener Räume und unendliche diskontinuierliche Gruppen, Comm. Math. Helv. 15 (1943) 27–32.
[18] W. Imrich, Subgroup theorems and graphs, Comb. Math. 5, Lecture Notes in Math. 622 (Springer, Berlin, 1977) 1–27.
[19] W. Imrich and N. Seifter, A bound for groups of linear growth, Arch. Math. 48 (1987) 100–104.
[20] W. Imrich and N. Seifter, A note on the growth of transitive graphs, Discrete Math. 73 (1988/89) 111–117.
[21] H.A. Jung, A note on fragments of infinite graphs, Combinatorica 1 (3) (1981) 285–288.
[22] H.A. Jung and M.E. Watkins, Fragments and automorphisms of infinite graphs, European J. Combin. 5 (1984) 149–162.
[23] J. Milnor, A note on curvature and fundamental group, J. Diff. Geom. 2 (1968) 1–7.
[24] J. Milnor, Growth of finitely generated solvable groups, J. Diff. Geom. 2 (1968) 447–449.
[25] J. Milnor, Problem 5603, Amer. Math. Monthly 75 (1968) 685–686.
[26] B. Mohar and W. Woess, A survey on spectra of infinite graphs, Bull. London Math. Soc. 21 (3) (1989) 209–234.
[27] B.H. Neumann, Groups with finite classes of conjugate elements, Proc. London Math. Soc. 1 (3) (1951) 178–187.
[28] S. Rosset, A property of groups of nonexponential growth, Proc. Amer. Math. Soc. 54 (1976) 24–26.
[29] G. Sabidussi, On a class of fixed-point-free graphs, Proc. Amer. Math. Soc. 9 (1958) 800–804.
[30] G. Sabidussi, Vertex transitive graphs, Mh. Math. 68 (1964) 385–401.
[31] N. Seifter, Automorphism groups of graphs with linear growth, submitted.
[32] N. Seifter, Groups acting on graphs with polynomial growth, Discrete Math. 89 (1991) 269–280.
[33] N. Seifter, properties of graphs with polynomial growth, J. Combin. Theory Ser. B 52 (2) (1991) 222–235.
[34] J. Stallings, Group Theory and Three Dimensional Manifolds (Yale Univ. Press, New Haven, 1971).
[35] M.J. Tomkinson, FC-groups, Res. Notes in Math. 96 (Pitman, London, 1984).
[36] V.I. Trofimov, Automorphism groups of graphs as topological groups, Math. Notes 38 (1985) 717–720.
[37] V.I. Trofimov, Automorphisms of graphs and a characterization of lattices, Math. USSR Izvestiya 22 (1984) 379–391.
[38] V.I. Trofimov, Graphs with polynomial growth, Math. USSR Sbornik 51 (1985) 405–417.
[39] S. Wagon, The Banach–Tarski Paradox, Encyclopedia of Math. and its Appl. (Cambridge Univ. Press, Cambridge, 1985).
[40] M.E. Watkins, Computing the connectivity of circulant graphs, Congr. Numer. 49 (1985) 247–258.
[41] M.E. Watkins, Infinite paths that contain only shortest paths, J. Combin. Theory Ser. B 41 (3) (1986) 341–355.
[42] R. Weiss, The nonexistence of 8-transitive graphs, Combinatorica 1 (3) (1981) 309–311.
[43] A.J. Wilkie and L. van den Dries, An effective bound for groups with linear growth, Arch. Math. 42 (1984) 391–396.
[44] W. Woess, Topological groups and infinite graphs, Discrete Math. 95 (this Vol.) (1991) 373–384.
[45] J. Wolf, Growth of finitely generated solvable groups and curvature of Riemannian manifolds, J. Diff. Geom. 2 (1968), 421–446.

Some results on ends and automorphisms of graphs

H.A. Jung

FBB (Math.), Technical University, Berlin, Germany

Received 28 December 1989
Revised 9 July 1990

Abstract

Jung, H.A., Some results on ends and automorphisms of graphs, Discrete Mathematics 95 (1991) 119–133.

Automorphisms σ of a connected graph X satisfying $\sigma(F) \neq F$ for all finite non-empty subsets F of $V(X)$ are of particular interest in the theory of ends. Elementary properties of those automorphisms are studied and linked to the concept of a strip.

1. Introduction and notation

The central concepts in this paper are the notion of an *end* and the notion of a *translation* (type 2 automorphisms in [6]) of a graph X. We will distinguish *proper* translations (hyperbolic automorphisms in [16]) and nonproper translations (parabolic automorphisms in [16]).

First we relate some basic results of [6] to the concept of a *strip*, which was studied in [10]. Then characterizations of proper translations are supplied.

While the results in Section 2 are valid for arbitrary infinite graphs, the emphasis in Section 3 is on the difference between locally finite graphs and the general case.

The proofs are all elementary, and we do not elaborate on the topological aspects of the theory of ends (cf. [8, 11–12, 16]). Neither do we elaborate on the links of this theory to group-theoretical aspects, such as the growth of finitely generated groups. Here we refer the reader to [3, 13–16].

Given a graph X we denote for any subgraph H of X, by $X - H$ and $X - V(H)$ the maximum subgraph of X with vertex set $V(X) - V(H)$, and by ∂H the set of all vertices in $X - H$, which have at least one neighbor in H.

If R is a ray (that is, a one-way infinite path) in X and F a finite subset of $V(X)$, there is a unique component C of $X - F$, which contains all but finitely many vertices of R; we shall say that C *covers* R.

It is easy to see that the following relation $\sim$ is an equivalence relation on the set of rays in X: $R \sim R'$ if for each finite $F \subseteq V(X)$, the rays R and R' are covered by the same component C of $X - F$.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

The classes with respect to $\sim$ are called *ends* in X (see [2] and [4]). Since for any finite $F \subseteq V(H)$ all rays in an end l in X are covered by the same component C of $X - F$, we may say that C *covers* l. For distinct ends l and l' in X there exists a finite set $F \subseteq V(X)$ such that l and l' are covered by distinct components of $X - F$. In that event we say that F *separates* l and l'. Clearly distinct ends of X are disjoint subsets of the set of rays in X.

Any element σ of the automorphism group $\operatorname{Aut} X$ of X maps any end l in X onto the end $\sigma(l) = \{\sigma(R): R \in (l)\}$. Clearly σ acts as a permutation of the set $L(X)$ of ends in X. This action of $\operatorname{Aut} X$ on $L(X)$ defines a canonical group homomorphism from $\operatorname{Aut} X$ into the group of permutations of $L(X)$.

An automorphism of X is called a *translation* of X (automorphism of type 2 in [6]), if $\sigma(F) \neq F$ for every finite non-empty subset of $V(X)$.

The following preliminary result is a useful tool in the later proofs.

Lemma 1.1. *For any translation σ of a graph X and any finite set $F \subseteq V(X)$ there exists a positive integer n such that $\sigma^i(F) \cap \sigma^j(F) = \emptyset$, whenever $|j - i| \geqslant n$.*

It suffices to show, that there is an integer $n > 0$ such that $\sigma^i(F) \cap F = \emptyset$, whenever $i \geqslant n$. That assertion follows immediately from the fact that for every $x \in V(X)$ the elements $\sigma^i(x)$ $(i \in \mathbb{Z})$ are all distinct.

2. Fixed ends and strips

Given $\sigma \in \operatorname{Aut} X$ we call a set Σ of subgraphs of X *σ-invariant* if $\sigma(Y) \in \Sigma$ for all $Y \in \Sigma$; if in addition for all $Y \in \Sigma$ there is a nonzero integer m such that $\sigma^m(Y) = Y$, we call Σ *σ-periodic*. A σ-invariant end of X will be called a *fixed end* of σ.

By definition a *strip* is a connected locally finite graph with exactly two ends, whose automorphism group contains a translation.

Theorem 2.1. *Let σ be a translation of the connected graph X, and let H be a finite connected subgraph of X such that $\sigma(H) \cap H \neq \emptyset$. Then $S = \bigcup_{i \in \mathbb{Z}} \sigma^i(H)$ is a strip, and $\sigma \mid S$ fixes both ends of S. Each of these ends is a subset of some fixed end of σ. Conversely if some end l of X contains no ray of S then $\sigma^i(l) \neq \sigma^j(l)$ whenever $i \neq j$.*

Proof. By Lemma 1.1 we can determine a positive integer n such that $\sigma^i(H) \cap \sigma^j(H) = \emptyset$ whenever $|j - i| \geqslant n$. Suppose $x \in \sigma^j(H)$ and $y \in \sigma^i(H)$ are adjacent in S. Then xy is an edge of some $\sigma^k(H)$. From $x \in \sigma^i(H) \cap \sigma^k(H)$ and $y \in \sigma^j(H) \cap \sigma^k(H)$ we deduce $|i - k| \leqslant n - 1$ and $|j - k| \leqslant n - 1$, and hence $|i - j| \leqslant 2n - 2$. Therefore S is locally finite. Moreover, for $\bar{H} = H \cup \sigma(H) \cup \cdots \cup$

$\sigma^{n-2}(H)$ there exist distinct components C_0^+ and C_0^- of $S-\bar{H}$ such that C_0^+ contains all $\sigma^i(H)$ $(i \geqslant 2n-2)$ and C_0^- contains all $\sigma^i(H)$ $(i \leqslant -n)$. Since both C_0^+ and C_0^- contain respectively rays R^+ and R^-, S has at least two ends. For any finite set $F \subseteq V(S)$, there exist integers $p < q$ such that $F \cap \sigma^i(H) = \emptyset$ whenever $i \leqslant p$ or $i \geqslant q$. Hence $S - F$ has at most two infinite components. So far we have shown that S has exactly two ends; one end l_0^+ of S contains R^+ and the other l_0^- contains R^-.

Clearly $\sigma_0 = \sigma \mid S$ is a translation of S, and l_0^+ and l_0^- are subsets of certain ends l^+ and l^- of X respectively, possibly $l^+ = l^-$. Since R^+ meets infinitely many $\sigma^i(H)$ $(i \geqslant 0)$, so does $\sigma(R^+)$ and so C_0^+ covers $\sigma_0(l_0^+)$. Since C_0^+ does not cover l_0^- we must have $\sigma_0(l_0^+) = l_0^+$. From $\sigma(l^+) \cap l^+ \supseteq \sigma_0(l_0^+) \cap l_0^+ = l_0^+$ we infer $\sigma(l^+) = l^+$. Similarly l_0^- is a fixed end of σ_0 and $\sigma(l^-) = l^-$.

Finally let l be an end of X other than l^+ and l^-. We choose a finite set $F \subseteq V(X)$ which separates in X the ends l, l^+ and also separates l, l^-. Then we find a finite connected subgraph H_1 of X such that $V(H_1) \supseteq F \cup V(H)$. As shown above, $S_1 = \bigcup_{i \in \mathbb{Z}} \sigma^i(H_1)$ is a strip with ends $l_1^+ \subseteq l^+$ and $l_1^- \subseteq l^-$ By Lemma 1.1 there is an integer $n_1 \geqslant n$ such that $\sigma^i(H_1) \cap \sigma^j(H_1) = \emptyset$ whenever $|i-j| \geqslant n_1$.

We choose $R \in l$ and show first that R meets only finitely many $\sigma^i(H_1)$ $(i \in \mathbb{Z})$. Otherwise the component of $X - H_1$ which covers l would contain all but finitely many $\sigma^i(H_1)$ $(i \geqslant 0)$ or all but finitely many $\sigma^i(H_1)$ $(i \leqslant 0)$. Since in X the set $V(H_1)$ separates l from l^+ and l^- we arrive at a contradiction.

Hence all but finitely many vertices of R are contained in some component C of $X - S_1$. Since $V(H_1)$ separates in X the ends l from l^+ and l^- the set ∂C can intersect only finitely many $\sigma^i(H_1)$ $(i \in \mathbb{Z})$. In particular ∂C is finite and thus C covers l. Let i be any nonzero integer. For all $j \in \mathbb{Z}$ clearly $\sigma^{ij}(C)$ is another component of $X - S_1$ and $\partial\sigma^{ij}(C) = \sigma^{ij}(\partial C)$. If $\sigma^i(C) \cap C \neq \emptyset$, that is $\sigma^i(C) = C$ then for all $j \in \mathbb{Z}$, $\partial C = \partial\sigma^{ij}(C) = \sigma^{ij}(\partial C)$, which by Lemma 1.1 is impossible.

By the preceding arguments $\sigma^i(C) \cap C = \emptyset$ whenever $i \neq 0$. Therefore $\sigma^i(C) \cap \sigma^j(C) = \emptyset$ whenever $i \neq j$. Since for all $k \in \mathbb{Z}$, $\sigma^k(C)$ covers $\sigma^k(l)$ we have shown that $\sigma^i(l) \neq \sigma^j(l)$ whenever $i \neq j$. □

From Theorem 2.1 we can regain the following results from [6].

Corollary 2.2. *Let σ be a translation of the connected graph X. Then σ has at least one and at most two fixed ends. If $m \in \mathbb{Z}/\{0\}$ then σ^m is a translation, and σ and σ^m have the same fixed ends.*

Proof. Since X is connected we can choose a vertex $x \in X$ and a path H in X joining x to $\sigma(x)$. Hence by Theorem 2.1, σ has at least one and at most two fixed ends.

For any integer $m \neq 0$ clearly σ^m is a translation of X and fixes every fixed end of σ. If σ has only one fixed end l^+, then again by Theorem 2.1, σ^m has no fixed end other than l^+. If σ has two fixed ends l^+ and l^- then they are fixed by σ^m. As already shown a translation has no more than two fixed ends. □.

We remark that Halin's proof for the 'at least'-part is more direct and probably simpler. We also should note that Halin's proof for the 'at most'-part works in general, although he imposed locally finiteness in his statement of this fact.

A translation σ of a connected graph will be called *proper* (*hyperbolic* in [16]) if σ fixes two ends.

Let σ be a translation of the connected graph X. If H is a finite connected subgraph of X such that $\sigma(H) \cap H \neq \emptyset$ then, by Theorem 2.1, some fixed end l_σ^+ of σ contains all rays in $\bigcup_{i \geq 0} \sigma^i(H)$. Note that l_σ^+ is the same end for every choice of H. Indeed, if H_1 and H_2 are two finite subgraphs of X one can find a finite connected subgraph $H \supseteq H_1 \cup H_2$. Then all rays in $\bigcup_{i \geq 0} \sigma^i(H)$ belong to the same fixed end of σ. This fixed end l_0^+ of σ is called the *direction* of σ [6]. If σ has two fixed ends, the fixed end l_σ^- of σ, which is distinct from l_0^+, contains all rays in $\bigcup_{i \leq 0} \sigma^i(H)$. Here H is as in Theorem 2.1 and separates l^-, l^+. In that case $l_\sigma^- = l_{\sigma^{-1}}^+$ and $l_{\sigma^{-1}}^- = l_\sigma^+$.

Theorem 2.3. *Let σ be a proper translation of the connected graph X and H a finite connected subgraph of X such that $\sigma(H) \cap H \neq \emptyset$ and H separates the fixed ends of σ. Then*

$$V\left(\bigcup_{i \in \mathbb{Z}} \sigma^i(H)\right)$$

induces a strip $\bar{S} \subseteq X$. Further there exists a finite subgraph $\bar{H}$ of $\bar{S}$ such that $V(\bar{H})$ separates in X the fixed ends of σ and for all components C of $X - \bar{S}$

(i) *$\partial C \subseteq \sigma^k(\bar{H})$ for some $k \in \mathbb{Z}$ and*

(ii) *$\sigma^i(C) \cap \sigma^j(C) = \emptyset$ whenever $i \neq j$.*

Proof. Let $S = \bigcup(\sigma^i(H): i \in \mathbb{Z})$. Let n be as provided by Lemma 1.1 for $F = V(H)$. We first show that the graph $\bar{S}$ which is induced by $V(S)$ in X is a strip. Since $V(H)$ separates l_σ^+ and l_σ^-, these ends are covered by distinct components of $X - H$.

One of these components, say C^+, contains all $\sigma^i(H)$ $(i \geq n)$ and another, say C^-, contains all $\sigma^i(H)$ $(i \leq -n)$ as subgraphs.

In particular $\sigma^{-n}(H)$ has no neighbor in $\bigcup_{i \geq n} \sigma^i(H)$. By applying appropriate powers of σ we obtain that $x \in \sigma^i(H)$ and $y \in \sigma^j(H)$ cannot be adjacent in X unless $|j - i| \leq 2n - 1$.

Therefore $\bar{S}$ is locally finite. Since $V(H)$ separates l^+ and l^- in X, $\bar{S}$ has exactly two ends $l_l^+ \subseteq l_\sigma^+$ and $l_1^- \subseteq l_\sigma^-$. Therefore $\bar{S}$ is a strip.

Set $\bar{H} = H \cup \sigma(H) \cup \cdots \cup \sigma^{2n-1}(H)$. We now consider a component C of $X - S = X - \bar{S}$. For any $i \in \mathbb{Z}$, $\sigma^i(C)$ is a component of $X - \bar{S}$ and $\partial\sigma^i(C) = \sigma^i(\partial C)$. If ∂C intersects $\sigma^j(H)$ then $\sigma^{-j-n}(\partial C)$ intersects $\sigma^{-n}(H)$ and hence does not intersect $\bigcup_{i \geq n} \sigma^i(H)$. Therefore ∂C does not intersect $\bigcup_{i \geq j+2n} \sigma^i(H)$. From this we infer that ∂C is finite and, choosing j as the minimum integer satisfying $\sigma^j(H) \cap \partial C \neq \emptyset$, $\partial C \subseteq \sigma^j(H) \cup \cdots \cup \sigma^{j+2n-1} = \sigma^j(\bar{H})$.

If $\sigma^i(C) = \sigma^j(C)$ for some distinct integers i and j, clearly $\sigma^{k(j-i)}(\partial C) = \partial C$ for all $k \in \mathbb{Z}$. This is not possible, since $\partial C \neq \emptyset$ and σ is a proper translation. $\square$.

Corollary 2.4. *For $1 \leqslant i \leqslant m$ let l_i be a fixed end of some proper translation σ_i of the connected graph X. There exists a finite subset F of $V(X)$, such that two distinct ends l and l' of X can be separated by $\tau(F)$ for some $\tau \in \operatorname{Aut} X$, whenever l or l' is in $\{\varphi(l_i)\colon 1 \leqslant i \leqslant m,\ \varphi \in \operatorname{Aut} X\}$.*

Proof. Clearly there is a finite set $F_0 \subseteq V(X)$, which separates the elements of $\{l\colon l$ fixed end of some σ_i, $1 \leqslant i \leqslant m\}$ pairwise. We determine a connected finite subgraph H of X such that $\sigma_i(H) \cap H \neq \emptyset$ for $1 \leqslant i \leqslant m$ and $V(H) \supseteq F_0$.

By Theorem 2.3 the set $V(\bigcup(\sigma_i^j(H)\colon j \in \mathbb{Z}))$ induces a strip $\bar{S}_i$ $(1 \leqslant i \leqslant m)$. Choose a subgraph $\bar{H}_i$ which satisfies (i) and (ii) of Theorem 2.3 with respect to σ_i. We claim that $F = V(\bar{H}_1 \cup \cdots \cup \bar{H}_m) = F$ has the desired property.

Consider distinct ends l and l' of X, say $l = \varphi(l_1)$, for some $\varphi \in \operatorname{Aut} X$. As $\varphi^{-1}(l') \neq l_1$, the end $\varphi^{-1}(l')$ is the second fixed end of $\bar{S}_1$ or covered by a component C of $X - \bar{S}_1$. In either case, by Theorem 2.3, there is an integer k such that $\sigma_1^k(\bar{H}_1)$ separates l_1 and $\varphi^{-1}(l')$. Now $\varphi\sigma_1^k(\bar{H}_1)$ separates $\varphi(l_1) = l$ and $\varphi\varphi^{-1}(l') = l'$. As $V(\bar{H}_1) \subseteq F$, clearly $\varphi\sigma_1^k(F)$ separates l and l'. $\square$

We state another result of Halin [6] as a corollary.

Corollary 2.5. *Let σ be a proper translation of the connected graph X, and let the finite subset F of $V(X)$ separate l_σ^+ and l_σ^-. Then any $l \in \{l_\sigma^+, l_\sigma^-\}$ contains at most $|F|$ pairwise disjoint rays.*

Proof. We choose a finite connected subgraph H of X such that $\sigma(H) \cap H \neq \emptyset$ and $V(H) \supseteq F$. Let $R_1, \ldots, R_t$ be pairwise disjoint rays in l_σ^+. Using Lemma 1.1 we determine a positive integer n such that $\sigma^i(H) \cap \sigma^j(H) = \emptyset$ whenever $|j - i| \geqslant n$.

Let C^+ and C^- be the components of $X - F$ which respectively cover l_σ^+ and l_σ^-. Note that $\bigcup_{i \leqslant -n} \sigma^i(H) \subseteq C^-$ and $\bigcup_{i \geqslant n} \sigma^i(H) \subseteq C^+$. In particular $H \subseteq \sigma^n(C^-)$ and consequently $H \cap \sigma^n(C^+) = \emptyset$. Thus $\sigma^n(C^+ \cup H)$ spans a connected subgraph of $X - F$. From $\sigma^n(H) \subseteq C^+$ we deduce $\sigma^n(C^+ \cup H) \subseteq C^+$ which of course implies $\sigma^n(C^+ \cup \partial C^+) \subseteq C^+$. The last inequality will be used in the proof of Theorem 2.7.

For each positive integer i, $\sigma^{ni}(C^+)$ covers l_σ^+ and hence intersects every $R \in l_\sigma^+$. Choose $x \in \partial C^+$. One easily verifies that any vertex in $\sigma^{in}(C^+)$ has distance at least $i+1$ from x $(i \geqslant 0)$. Therefore $\bigcap_{i \geqslant n} \sigma^{ni}(C^+) = \emptyset$ and hence there is a positive integer i such that the terminal vertices of $R_1, R_2, \ldots, R_t$ are outside $\sigma^{ni}(C^+)$. For such an integer i each $R \in \{R_1, R_2, \ldots, R_t\}$ intersects $\partial\sigma^{ni}(C^+)$. Now

$$\partial\sigma^{ni}(C^+) = \sigma^{ni}(\partial C^+) \subseteq \sigma^{ni}(F)$$

hence $t \leqslant |\sigma^{ni}(F)| = |F|$.

Since $l_\sigma^- = l_{\sigma^{-1}}^+$ the proof of the corollary is complete. $\square$

The proof of Theorem 7 in [6] yields the following slightly stronger version of that result. For completeness we sketch the proof.

Proposition 2.6. *Let σ be a translation of the connected graph X. There exists a non-empty σ-periodic set of pairwise disjoint double rays in X.*

Proof. Among all paths P' such that $\sigma^m(P') \cap P' \neq \emptyset$ for some $m > 0$ choose a minimal one and call it P. Then $\sigma^m(P) \cap P \neq \emptyset$ for some $m > 0$, i.e. $\sigma^m(x) \in P$ for some $x \in P$. By the minimality of P we have $P = P[x, \sigma^m(x)]$. Moreover, $\sigma^i(P) \cap \sigma^j(P) = \emptyset$ whenever $m \neq |j - i|$, and $\sigma^{(j-1)m+i}(P) \cap \sigma^{jm+i}(P) = \sigma^{jm+i}(x)$.

Setting $Q_i = \bigcup_{j \in \mathbb{Z}} \sigma^{jm+i}(P)$ for all $i \in \mathbb{Z}$ we see that $\{Q_0, Q_1, \ldots, Q_{m-1}\}$ is a σ-invariant set of pairwise disjoint double rays. □

In the next result we establish a necessary and sufficient condition for an automorphism to be a proper translation.

Theorem 2.7. *An automorphism σ of a connected graph X is a proper translation of X if and only if there exists an induced connected subgraph C and a positive integer n such that $0 < |\partial C| < \infty$ and $\sigma^n(C \cup \partial C) \subseteq C$.*

Proof. If σ is a proper translation, we choose a finite subset F of $V(X)$ which separates l_σ^+ and l_σ^-. With n, H and C^+ as constructed in the proof of Corollary 2.5 we have $\sigma^n(C^+ \cup H) \subseteq C^+$ and $\partial C^+ \subseteq F \subseteq H$.

Now assume $\sigma^n(C \cup \partial C) \subseteq C$ for some positive integer n and some induced connected subgraph C of X with infinite ∂C. Abbreviate $\tau = \sigma^n$. If $i > 0$ then clearly $d(x, y) \geqslant i$ whenever $x \in \partial C$ and $y \in \tau^i(\partial C)$, hence $d(\partial C, \tau^i(\partial C)) \geqslant i$.

For any finite non-empty $F \subseteq V(X)$ and any positive integer i we have

$$d(\partial C, \tau^i(F)) \geqslant d(\partial C, \tau^i(\partial C)) - d(\tau^i(F), \tau^i(\partial C)).$$

From $d(\tau^i(F), \tau^i(\partial C)) = d(F, \partial C)$ we infer $d(\tau^i(F), \partial C) \geqslant i - d(F, \partial C)$. Therefore we cannot have $\tau(F) = F$. Thus we have shown that τ is a translation. Finally we choose some $x \in \partial C$ and some path H in X joining x to $\tau(x)$. Abbreviate $n = |V(H)| - 1$.

For any $i \in \mathbb{Z}$ we have $\tau^i(x) \in \tau^i(\partial C) \subseteq \tau^{i-1}(C) - \tau^i(C)$.

From $\tau^{i+1+m}(x) \in \tau^i(C)$ and $d(\tau^{i+m+1}(x), \partial\tau^i(C)) \geqslant m + 1$ we deduce that the path $\tau^{i+m}(H)$ of length m is a subgraph of $\tau^i(C)$.

From $\tau^{i-m}(x) \notin \tau^i(C)$ and $d(\tau^{i-m}(x), \partial\tau^i(C)) \geqslant m$ we similarly obtain $\tau^{i-m}(C) \supseteq \tau^{i-m-1}(H) \cap \tau^i(C) = \emptyset$. Therefore $\bigcup_{i \geqslant m} \tau^i(H) \subseteq C$ and C covers all rays of $l_{\tau/S}^+$, where S denotes the strip $\bigcup_{i \in \mathbb{Z}} \tau^i(H)$. Furthermore, $C \cap \bigcup_{i \leqslant -m-1} \tau^i(H) = \emptyset$ and so C does not cover any ray in $l_{\tau/S}^-$. Hence $l_\tau^+ \neq l_\tau^-$. By Corollary 2.2 also σ is proper. □

3. Nonproper translations and groups of automorphisms

Taking the viewpoint that an end l of X is a 'limit' vertex or a fictitious vertex, we may have the situation that ends l corresponds to real vertices. We call the vertex v of X a *main vertex* of l if for each finite $F \subseteq V(X) - \{v\}$ the component of $X - F$ which contains v covers l (see [7]). It is easy to see that v is a main vertex of l if and only if for each (equivalently some) $R \in l$ there exists an infinite system of openly disjoint paths joining v to distinct elements on R. Given a graph X_0, we call each graph of the form

$$X_0 \cup \bigcup(P_n : n = 1, 2, \ldots)$$

a *v-extension* of X_0 if $P_n = P_n[v, w_n]$ $(n = 1, 2, \ldots)$ are openly disjoint paths with distinct endvertices w_n in X_0 and $P_n(v, w_n) \cap X_0 = \emptyset$ $(n = 1, 2, \ldots)$. Note that v may be an element of $V(X_0)$. A union of v_n-extensions $\bar{X}_n$ of X_0 $(n = 1, 2, \ldots)$ is called an *infinite* extension of X_0 if all v_n $(n = 1, 2, \ldots)$ are distinct and the graphs $\bar{X}_n - V(X_0)$ $(n = 1, 2, \ldots)$ are pairwise disjoint.

If l is a fixed end of σ and x is a main vertex of l, then clearly $\sigma^i(x)$ is a main vertex of $\sigma^i(l) = l$ for all $i \in \mathbb{Z}$. Therefore Theorem 2.3 yields that the fixed ends of a proper translation in the connected graph X cannot have main vertices.

The following lemma is a useful tool for the construction of v-extensions.

Lemma 3.1. *Let $R_1, R_2, \ldots, R_k$ be rays in the end l of X, and $\bar{F} = V(R_1 \cup \cdots \cup R_k)$. Each component C of $X - \bar{F}$ with infinite ∂C contains a ray of l or a main vertex of l.*

Proof. Let C be a component of $X - \bar{F}$ with infinite ∂C. Without loss of generality we may assume that $V(R_1) \cap \partial C$ is infinite.

We first construct paths $P_n = P_n[v_n, w_n]$ such that $v_n \in V(C)$, $w_n \in V(R_1) \cap \partial C$ and $P_n[v_n, w_n) \subseteq C$ $(n = 1, 2, \ldots)$.

We start by choosing $v_1 \in V(C)$, $w_1 \in V(R_1) \cap \partial C$ and a path P_1 in X, which joins v_1 to w_1 and has inner vertices in C.

Suppose $P_1 = P_1[v_1, w_1], \ldots, P_n = P_n[v_n, w_n]$ have been constructed. We determine a vertex

$$w_{n+1} \in V(R_1) \cap \partial C - \{w_1, \ldots, w_n\}$$

and a path P_{n+1} joining w_{n+1} to a vertex in $V(P_1 \cup \cdots \cup P_n) \cap V(C)$. We can choose $P_{n+1} = P_{n+1}[v_{n+1}, w_{n+1}]$ such that

$$P_{n+1}[v_{n+1}, w_{n+1}) \subseteq C \quad \text{and} \quad P_{n+1} \cap (P_1 \cup \cdots \cup P_n) = v_{n+1}.$$

By construction $T = \bigcup(P_n : n = 1, 2, \ldots)$ is s tree. If some $v \in V(T)$ has infinite degree in T, then $v = v_n$ for infinitely many n and hence T contains a v-extension of R_1. Now we assume that T is locally finite and choose a ray R in T. For any $v \in V(R)$ let $n(v)$ be the minimum integer n such that v is on P_n. Clearly there is

a system $v_1', v_2', \ldots$ of vertices on R such that $n(v_1') < n(v_2') < \cdots$. By construction the paths $P_{n(v_i')}$ $(i = 1, 2, \ldots)$ are pairwise disjoint. Hence R is in l. □

In the following two lemmas we assume that a nonproper translation σ of a connected graph X is given, that Σ is a maximal σ-periodic set of pairwise disjoint double rays in X, and that

$$\bar{F} = \bigcup (V(Q) \colon Q \in \Sigma).$$

Lemma 3.2. *For each component C of $X - \bar{F}$ we have $\sigma^i(C) \cap \sigma^j(C) = \emptyset$ whenever i and j are distinct integers.*

Proof. If $\sigma^i(C) \cap \sigma^j(C) \neq \emptyset$ for some distinct i and j, then $\sigma^i(C) = \sigma^j(C)$ and hence there is a minimum positive integer q such that $\sigma^q(C) = C$. Then $\sigma^q \mid C$ is a translation of C and hence, by Proposition 2.6, there exists a σ^q-invariant set $\{Q_1, \ldots, Q_k\}$ $(k \geqslant 1)$ of pairwise disjoint double rays in C. Since $C, \sigma(C), \ldots, \sigma^{q-1}(C)$ are pairwise disjoint, the set

$$\Sigma \cup \{\sigma^j(Q_i) \colon 1 \leqslant i \leqslant k, 0 \leqslant j < q\}$$

would be a σ-periodic set of pairwise disjoint double rays which contains Σ as a proper subset. □

Lemma 3.3. *Let $Q \in \Sigma$ and let C be a component of $X - \bar{F}$. If $V(Q) \cap \partial C$ is infinite, then there exist an integer $q > 0$ and an infinite extension X_1 of Q such that $\sigma^q(X_1) = X_1$ and $V(X_1), \ldots, \sigma^{q-1}(V(X_1))$ are pairwise disjoint subsets of $V(Q \cup \cdots \cup \sigma^{q-1}(Q)) \cup (V(X) - \bar{F})$.*

Proof. Let q denote the minimum positive integer such that $\sigma^q(Q) = Q$. Choosing z on Q and setting $H = Q[z, \sigma^q(z)]$ we obtain $Q = \bigcup (\sigma^{iq}(H) \colon i \in \mathbb{Z})$. Assuming that $V(Q) \cap \partial C$ is infinite we can find a vertex x in H such that the set $S = \{i \in \mathbb{Z} \colon \sigma^{iq}(x) \in \partial C\}$ is infinite. Let $S' = \{\sigma^{iq}(x) \colon i \in S\}$.

We first construct a subtree T of X such that S' is the set of endvertices of T and $V(T) \subseteq V(Q \cup C)$. To this end we label $S' = \{w_1, w_2, \ldots\}$ and start the construction by choosing a path T_2 in X, which joins w_1 to w_2 and has all its inner vertices in C. Assuming that a finite tree $T_n \subseteq X$ with $V(T_n) \subseteq V(Q \cup C)$ and set of end vertices $\{w_1, \ldots, w_n\}$ has already been constructed, we add to T_n a minimal path joining w_{n+1} to a vertex of T_n to obtain a tree T_{n+1} with set of endvertices $\{w_1, \ldots, w_{n+1}\}$. Clearly $T = T_2 \cup T_3 \cup \cdots$ has the asserted properties. We distinguish two cases.

Case 1: *Some vertex $v \in V(T)$ has infinite degree in T.*

Since T-$\{v\}$ has infinitely many components we can determine openly disjoint paths $P_n = P_n[v, z_n]$ $(n = 1, 2, \ldots)$ in T such that $z_n \in \{\sigma^{iq}(x) \colon i \in S\}$ $(n = 1, 2, \ldots)$ and all z_n $(n = 1, 2, \ldots)$ are distinct. Then $z_n = \sigma^{i_n \cdot q}(x)$ where $i_n \in S$ $(n =$

$1, 2, \ldots$). Thus $Q' = Q \cup \bigcup (P_n : n = 1, 2, \ldots)$ is a v-extension of Q. Hence for all $j \in \mathbb{Z}$ the graph $\sigma^j(Q')$ is a $\sigma^j(v)$-extension of $\sigma^j(Q)$ such that $\sigma^j(Q') - \sigma^j(Q) \subseteq \sigma^j(C)$. Therefore by Lemma 3.2 the graph

$$X_1 = \bigcup (\sigma^{jq}(Q') : j \in \mathbb{Z})$$

is an infinite extension of Q. Clearly $\sigma^q(X_1) = X_1$, and for every integer r the graph $\sigma^r(X_1)$ is an infinite extension of $\sigma^r(Q)$. Consider integers r_1, r_2 such that $0 \leqslant r_1 < r_2 < q$. Since $\sigma^{r_1}(Q) \cap \sigma^{r_2}(Q) = \emptyset$,

$$\sigma^{r_1}(X_1 - Q) \subseteq \bigcup (\sigma^{jq + r_1}(C) : j \in \mathbb{Z})$$

and

$$\sigma^{r_2}(X_1 - Q) \subseteq \bigcup (\sigma^{jq + r_2}(C) : j \in \mathbb{Z})$$

we obtain $\sigma^{r_1}(X_1) \cap \sigma^{r_2}(X_1) = \emptyset$.

Case 2: *T is locally finite.*

We construct sequences i_n, j_n and P_n ($n = 1, 2, \ldots$), where P_n is a path in T, which joins $\sigma^{i_n q}(x)$ to $\sigma^{j_n q}(x)$.

We start by choosing i_1 and j_1 in S such that $i_1 < j_1$ and letting P_1 denote the path in T, which joins $\sigma^{i_1 q}(x)$ to $\sigma^{j_1 q}(x)$. Suppose that the path $P_1, P_2, \ldots, P_n$ in T have already been constructed. Since T is locally finite, $T - (P_1 \cup \cdots \cup P_n)$ has only finitely many components, and therefore some component T_{n+1} contains infinitely many elements of S'. Thus we can find i_{n+1} and j_{n+1} in S, such that $j_{n+1} - i_{n+1} \geqslant n + 1$. Let P_{n+1} denote the path in T_{n+1} joining $\sigma^{i_{n+1} q}(x)$ to $\sigma^{j_{n+1} q}(x)$.

By construction the paths P_n ($n = 1, 2, \ldots$) are pairwise disjoint, further $d(n) = j_n - i_n \geqslant n$ ($n = 1, 2, \ldots$). Hence we can find an infinite set M of positive integers such that all $d(n)$ ($n \in M$) are distinct and $1 \notin M$. Set $z = \sigma^{i_1 q}(x)$.

For any $n \in M$ the path $P'_n = \sigma^{(i_1 - i_n)q}(P_n)$ joins $\sigma^{i_1 q}(x) = z$ to $\sigma^{(i_1 - i_n + j_n)q}(x) = \sigma^{d(n)q}(z)$ and has inner vertices in $\sigma^{(i_1 - i_n)q}(C)$. Therefore, by Lemma 3.2, the graph

$$Q^* = Q \cup \bigcup (P'_n : n \in M)$$

is a z-extension of Q and $V(Q^*) - V(Q) \subseteq V(X) - \bar{F}$.

Suppose $\sigma^{aq}(P'_n)$ and $\sigma^{bq}(P'_m)$ have a common vertex outside Q, where $n, m \in M$. This means that $\sigma^{(a + i_1 - i_n)q}(P_n)$ and $\sigma^{(b + i_1 - i_m)q}(P_m)$ have a common vertex w outside Q. But

$$w \in \sigma^{(a + i_1 - i_n)q}(C) \cap \sigma^{(b + i_1 - i_m)q}(C)$$

hence, by Lemma 3.2, $a + i_1 - i_n = b + i_1 - i_m$. Consequently $P_n = P_m$ hence $n = m$ and $a = b$.

In particular the graph $X_1 = \bigcup (\sigma^{jq}(Q^*) : j \in \mathbb{Z})$ is an infinite extension of Q. Clearly $\sigma^q(X_1) = X_1$.

Now consider integers r_1 and r_2 such that $0 \leqslant r_1 < r_2 < q$. If $\sigma^{r_1}(X_1)$ and $\sigma^{r_2}(X_1)$ had a common vertex then X_1 and $\sigma^{r_2 - r_1}(X_1)$ would have a common vertex w.

Since $Q \cap \sigma^{r_2-r_1}(Q) = \emptyset$ the vertex w would be in $X - \bar{F}$ and hence be an inner vertex of some P'_n and some $\sigma^{r_2-r_1}(P'_m)$, where $n, m \in M$. Consequently

$$w \in \sigma^{(i_1-i_n)q}(C) \cap \sigma^{r_2-r_1+(i_1-i_m)q}(C)$$

and by Lemma 3.2, $(i_1 - i_n)q = r_2 - r_1 + (i_1 - i_m)q$. Since $0 < r_2 - r_1 < q$ this equality is absurd.

Thus we have shown that $\{X_1, \sigma(X_1), \ldots, \alpha^{q-1}(X)\}$ is a σ-invariant set of pairwise disjoint graphs. □

Theorem 3.4. *Let σ be a nonproper translation of the connected graph X. If Σ is a maximal finite σ-periodic set of pairwise disjoint double rays in X then there exist $q > 0$ and an infinite extension X_1 of some $Q \in \Sigma$ such that $\sigma^q(X_1) = X_1$ and $X_1, \sigma(X_1), \ldots, \sigma^{q-1}(X_1)$ are pairwise disjoint.*

Proof. Let $\Sigma = \{Q_1, \ldots, Q_k\}$. For any subgraph H of X and any subset T of $V(X)$ abbreviate

$$\bar{H} = \bigcup(\sigma^i(H): i \in \mathbb{Z}) \quad \text{and} \quad \bar{T} = \bigcup(\sigma^i(T): i \in \mathbb{Z});$$

clearly $\sigma(\bar{H}) = \bar{H}$ and $\sigma(\bar{T}) = \bar{T}$. For $1 \leqslant j \leqslant k$ let q_j denote the minimum positive integer such that $\sigma^{q_j}(Q_j) = Q_j$. Choosing for any j, $1 \leqslant j \leqslant k$, a vertex x_j on Q_j and setting

$$H_j = Q_j[x_j, \sigma^{q_j}(x_j)]$$

we obtain $Q_j = \bigcup(\sigma^{iq_j}(H_j): i \in \mathbb{Z})$. Setting

$$F = V(H_1 \cup \cdots \cup H_k)$$

we obtain $\bar{F} = V(Q_1 \cup \cdots \cup Q_k)$.

First we construct a v-extension $Q' = Q \cup \bigcup(P_n: n = 1, 2, \ldots)$ of some $Q \in \Sigma$ such that $V(Q') - V(Q) \subseteq V(X - \bar{F}) \cup \{v\}$.

As σ is nonproper, for any $n = 1, 2, \ldots$ the set $F_n = \bigcup(\sigma^i(F): -n < i < n)$ does not separate in X the ends of Q_1. Hence we can determine a path P'_n in $X - F_n$ having a terminal vertex y_n in $\bigcup(\sigma^i(F): i \leqslant -n)$ and a terminal vertex z_n in $\bigcup(\sigma^i(F): i \geqslant n)$ and such that the inner vertices of P'_n, if any, are in $X - \bar{F}$. Choosing i_n and j_n such that $y_n \in \sigma^{i_n}(F)$ and $z_n \in \sigma^{j_n}(F)$ we have $\sigma^{-i_n}(y_n), \sigma^{-j_n}(z_n) \in F$ $(n = 1, 2, \ldots)$. Therefore we can determine an infinite set M_0 of integers such that all $\sigma^{-i_n}(y_n)$ $(n \in M_0)$ are the same vertex $y \in F$ and all $\sigma^{-j_n}(z_n)$ $(n \in M_0)$ are the same vertex $z \in F$.

Let us abbreviate $d(n) = j_n - i_n$ $(n = 1, 2, \ldots)$. As $d(n) \geqslant 2n$ $(n = 1, 2, \ldots)$ we can determine an infinite subset M_1 of M_0 such that all $d(n)$ $(n \in M_1)$ are distinct. Exploiting the fact that

$$\sigma^{-i_n}(z_n) = \sigma^{d(n)}(z) \quad \text{and} \quad \sigma^{-j_n}(y_n) = \sigma^{d(n)}(y)$$

are elements of $\bar{F} = V(Q_1 \cup \cdots \cup Q_k)$ we can find an infinite set $M_2 \subseteq M_1$ and

elements Q_κ, Q_λ in $\{Q_1, \ldots, Q_k\}$, such that all $\sigma^{-d(n)}(y)$ $(n \in M_2)$ are in $V(Q_\kappa)$ and all $\sigma^{d(n)}(z)$ $(n \in M_2)$ are in $V(Q_\lambda)$.

Now observe that for all $n \in M_2$ the path $P_n'' = \sigma^{-i_n}(P_n')$ joins y to $\sigma^{d(n)}(z)$ and has inner vertices, if any, in a unique component C_n' of $X - \bar{F}$. If for some infinite subset M_3 of M_2 the paths P_n'' are openly disjoint then

$$Q_\lambda' = Q_\lambda \cup \bigcup (P_n'': n \in M_3)$$

is a y-extension of Q_λ.

In the remaining case all but finitely many $n \in M_2$ satisfy $|V(P_n)| \geqslant 3$. Moreover, there is an infinite subset M_3 of M_2 such that $|V(P_n)| \geqslant 3$ for all $n \in M_3$, and all C_n' $(n \in M_3)$ are the same component C' of $X - \bar{F}$. For all $n \in M_3$ the path $\sigma^{-j_n}(P_n') = \sigma^{i_n - j_n}(P_n'')$ joins $\sigma^{-j_n}(y_n) = \sigma^{d(n)}(y)$ to $\sigma^{-j_n}(z_n) = z$ and has inner vertices in $\sigma^{-d(n)}(C_n') = \sigma^{-d(n)}(C')$. Thus

$$Q_\kappa' = Q_\kappa \cup \bigcup (\sigma^{-d(n)}(P_n''): n \in M_3)$$

is a z-extension of Q_κ.

This completes the construction of the v-extension $Q' = Q \cup \bigcup (P_n: n = 1, 2, \ldots)$ of Q, where $Q \in \Sigma$. Let q denote the minimum positive integer such that $\sigma^q(Q) = Q$. In view of Lemma 3.3 we may assume that for any component C of $X - \bar{F}$ the boundary ∂C is finite. Let $P_n = P_n[v, w_n]$ $(n = 1, 2, \ldots)$.

We claim that for every positive integer n the set M_n of all positive integers m, such that $n \neq m$ and $(\bar{P}_n - \bar{F}) \cap (\bar{P}_m - \bar{F}) \neq \emptyset$, is finite. Assume to the contrary that for some positive integer n_0 the set M_{n_0} is infinite. Then for all $n \in M_{n_0}$ the paths P_{n_0} and P_n have inner vertices in respectively C_{n_0} and C_n. Abbreviate $P = P_{n_0}$, $M = M_{n_0}$ and $C = C_{n_0}$. By assumption there exist for each $n \in M$ integers i_n and j_n such that $\sigma^{i_n}(P)$ and $\sigma^{j_n}(P_n)$ have inner vertices in common, consequently $\sigma^{i_n}(C) = \sigma^{j_n}(C_n)$. Note that $v, w_n \in \partial C_n$ for all $n \in M \cup \{n_0\}$. Abbreviating $d(n) = j_n - i_n$ we have $C = \sigma^{d(n)}(C_n)$ and consequently

$$\sigma^{d(n)}(v),\ \sigma^{d(n)}(w_n) \in \sigma^{d(n)}(\partial C_n) = \partial C$$

for all $n \in M$. But clearly $\{\sigma^{d(n)}(v): n \in M\}$ or $\{\sigma^{d(n)}(w_n): n \in M\}$ is infinite, contrary to our assumption, that ∂C be finite.

Having proved the above claim, we can easily construct an infinite set M of positive integers such that $(\bar{P}_n - \bar{F}) \cap (\bar{P}_m - \bar{F}) = \emptyset$ whenever n and m are distinct elements of M. Set

$$Q^* = Q \cup \bigcup (P_n: n \in M).$$

As Q^* is a v-extension of Q, for all $i \in \mathbb{Z}$ the graph $\sigma^i(Q^*)$ is a $\sigma^i(v)$-extension of $\sigma^i(Q)$.

We claim that $X_1 = \bigcup (\sigma^{jq}(Q^*): j \in \mathbb{Z})$ is an infinite extension of Q. Assume to the contrary that $\sigma^{j_1 q}(Q^*) - Q$ and $\sigma^{j_2 q}(Q^*) - Q$ have a common vertex w_0, where $j_1 \neq j_2$, then w_0 is in $X - \bar{F}$ and hence an inner vertex of some $\sigma^{j_1 q}(P_{n_1})$ and of some $\sigma^{j_2 q}(P_{n_2})$, where $n_1, n_2 \in M$. Then $n_1 = n_2$ since otherwise $(\bar{P}_{n_1} - \bar{F}) \cap (\bar{P}_{n_2} - \bar{F}) \neq \emptyset$.

Since $w_0 \in \sigma^{j_1 q}(C_{n_1}) \cap \sigma^{j_2 q}(C_{n_1})$ we obtain a contradiction to Lemma 3.2.

Clearly $\sigma^q(X_1) = X_1$, and for all $0 \leqslant r < q$, the graph $\sigma^r(X_1)$ is an infinite extension of $\sigma^r(Q)$. It remains to show that $X_1, \sigma(X_1), \ldots, \sigma^{q-1}(X_1)$ are pairwise disjoint. Suppose $\sigma^{r_1}(X_1)$ and $\sigma^{r_2}(X_1)$ have a common vertex, where $0 \leqslant r_1 < r_2 < q$. Then X_1 and $\sigma^{r_2 - r_1}(X_1)$ have a common vertex w. As

$$Q \cap \sigma^{r_2 - r_1}(Q) = \emptyset \quad \text{and} \quad \{\sigma^{jq}(v) : j \in \mathbb{Z}\} \cap \{\sigma^{jq + r_2 - r_1}(v) : j \in \mathbb{Z}\} = \emptyset$$

the vertex w is an inner vertex of some P_{n_1} and of some $\sigma^{r_2 - r_1}(P_{n_2})$, where $n_1, n_2 \in M$. By the properties of M we must have $n_1 = n_2$, consequently $C_{n_1} \cap \sigma^{r_2 - r_1}(C_{n_1}) \neq \emptyset$ contrary to Lemma 3.2.

Thus $\{X_1, \sigma(X_1), \ldots, \sigma^{q-1}(X_1)\}$ is a σ-invariant set of pairwise disjoint infinite extensions of respectively $Q, \sigma(Q), \ldots, \sigma^{q-1}(Q)$. □

Corollary 3.5. *If σ is a nonproper translation of the connected graph X then the fixed end of σ in X contains infinitely many pairwise disjoint rays in X.*

Proof. Consider a maximal σ-invariant periodic set Σ of pairwise disjoint double rays. If Σ is infinite, clearly the ends of any $Q \in \Sigma$ are subsets of l_σ^+.

If $\Sigma = \{Q_1, \ldots, Q_n\}$ we determine q, Q and X_1 according to the theorem. It is easy to construct infinitely many pairwise disjoint rays in X_1. Clearly those rays are in l_σ^+. □

Theorem 3.6. *Let X be a connected graph with at least two ends, and let G be a subgroup of* Aut X *which has only finitely many orbits on $V(X)$. If G is vertex-transitive, or at least two ends of X have no main vertices, then G contains a proper translation.*

Proof. If G is a vertex transitive, we invoke Theorem 1 in [8], which says that there exist an induced subgraph C of X and an element σ of G satisfying $0 < |\partial C| < \infty$ and $\sigma(C \cup \partial C) \subseteq C$. Then σ is a proper translation by Theorem 2.7.

Now assume that there exist distinct ends l_1 and l_2, neither of which has a main vertex. Clearly there is a finite connected subgraph H of X, which separates l_1 and l_2 and contains additionally at least one element from each orbit of G on $V(X)$. Let C_1 and C_2 denote the components of $X - H$ which cover respectively l_1 and l_2.

We first show that, for $i = 1$ and 2, there exists a vertex x_i in C_i such that $d(x_i, H) > \operatorname{dia} H$, that is $\min(d(x_i, v) : v \in V(H)) > \max(d(v, w) : v, w \in V(H))$. Assuming the contrary we also may assume that C_1 contains no such vertex x_1. Setting $\operatorname{dia}(H) = k$ we obtain a decomposition

$$V(H) \cup V(C_1) = V_0 \cup V_1 \cup \cdots \cup V_k,$$

where $V_i = \{v \in V(H \cup C_1) : d(v, H) = i\}$ $(0 \leqslant i \leqslant k)$. For each $v \in V_i$ $(0 < i \leqslant k)$ we choose a unique edge which is incident with v and some vertex in V_{i-1}.

Clearly all chosen edges define a forest with vertex set $V(C_1 \cup H)$ and each component T of that forest contains a unique vertex $v(T)$ in $V(H)$. Further we determine a ray R in l_1 such that $R \subseteq C_1$. Some component T of the above forest contains infinitely many vertices of R, say the distinct vertices $w_1, w_2, \ldots$ For $n = 1, 2, \ldots$ let P_n denote the path in T which connects $v(T)$ to w_n. As T has finite diameter we can determine a vertex w in T at maximum distance from $v(T)$ in T, such that w is on infinitely many P_n, say on $P_{n_1}, P_{n_2}, \ldots (n_1 < n_2 < \cdots)$. By the choice of w there exists an infinite subsequence of $P_{n_1}[w, w_{n_1}], P_{n_2}[w, w_{n_2}], \ldots$ which defines an infinite set of openly disjoint paths. Therefore w is a main vertex of l_1, a contradiction.

Thus for $i = 1$ and 2 we can find x_i in C_i such that $d(x_i, H) > \operatorname{dia} H$, further an element $y_i \in V(H)$ and an element σ_i in G such that $\sigma_i(y_i) = x_i$. Then for $i = 1$ and 2,

$$x_i \in V(C_i) \cap \sigma_i(H) \quad \text{and} \quad d(x_i, H) > \operatorname{dia} H = \operatorname{dia} \sigma_i(H)$$

imply

$$H \cap \sigma_i(H) = \emptyset.$$

If for $i = 1$ and 2, $H \cap \sigma_i(C_i) = \emptyset$ then $\sigma_i(C_i \cup H)$ induces a connected subgraph of $X - H$ and hence $\sigma_i(C_i \cup H) \subseteq C_i$. In that case σ_1 or σ_2 is a proper translation by Theorem 2.7.

If for $i = 1$ and $i = 2$, $H \cap \sigma_i(C_i) \neq \emptyset$, then $H \subseteq \sigma_1(C_1) \cap \sigma_2(C_2)$. In that case both $\sigma_1(C_2 \cup H)$ and $\sigma_2(C_1 \cup H)$ induce connected subgraphs of $X - H$, hence

$$\sigma_1(C_2 \cup H) \subseteq C_1 \quad \text{and} \quad \sigma_2(C_1 \cup H) \subseteq C_2.$$

Then $\sigma_1\sigma_2(C_1 \cup H) \subseteq \sigma_1(C_2) \subseteq C_1$ and $\sigma_2\sigma_1(C_2 \cup H) \subseteq C_2$ and, by Theorem 2.7, both $\sigma_1\sigma_2$ and $\sigma_2\sigma_1$ are proper translations. □

It is not difficult to construct graphs X, such that Aut X has exactly two orbits on $V(X)$ and exactly two orbits on $L(X)$. Let X_1 and X_2 be two vertex-transitive 2-connected non-isomorphic graphs. For each $v \in V(X_1)$ let $X_2(v)$ be a copy of X_2. Let X be constructed from the disjoint union of X_1 and all $X_2(v)$ $(v \in V(X_1))$ by adding all edges $[v, w]$, where $v \in V(X_1)$ and $w \in V(X_2(v))$. Clearly Aut X has two orbits on $V(X)$. If each of X_1 and X_2 has exactly one end orbit, then Aut X has two end-orbits. If Aut X_1 and Aut X_2 are torsion groups, then Aut X contains no translation. Examples for graphs X_1 for which Aut X_1 is a vertex-transitive torsion group are given in [1]. X_1 can be chosen infinite, connected and locally finite. So in a certain sense Theorem 3.6 is sharp.

Let us call an end l of the graph X a *thick* end of X if l contains an infinite subset of pairwise disjoint rays. If Aut X is a vertex-transitive and $|L(X)| \geqslant 2$, by Theorem 3.6 the graph X allows a proper translation. But as noted above, fixed ends of proper translations are not thick. Therefore the ends of a vertex-transitive graph X with $|L(X)| \geqslant 2$ cannot all be thick. On the other hand it is easy to construct end-transitive locally finite graphs X with $|L(X)| = 2^{\omega}$ in which each end is thick.

Let T_d be the tree in which one vertex v_0 has degree $d-1$ and all other vertices have degree d $(d \geqslant 2)$. For $h = 0, 1, \ldots$ let $S(h)$ denote the set of vertices in $T = T_d$ at distance h from v_0. Let $p(1) < p(2) < \cdots$ be the sequence of prime numbers. For $n = 2, 3, 4, \ldots$ we add to T an edge between $v \in S(p(n))$ and $w \in S(p(n^2))$, whenever v is on the unique path $T[v_0, w]$ joining v_0 and w in T. Clearly the arising graph $X = X_d$ is locally finite. We first show that each end of X contains a (unique) ray in T with initial vertex v_0. Given a ray R in X we construct a sequence $x_1, x_2, \ldots$ in $V(R)$ such that

$$V(R) \cap V(T[v_0, x_m]) = \{x_1, x_2, \ldots, x_m\}$$

and the subray R_m of R with initial vertex x_m contains no element of $T[v_0, x_m)$. Let x_1 be the $\leqslant$ – minimum of $V(R_1)$, where $\leqslant$ is the order related to the rooted tree (T_d, v_0) and v_0 is the $\leqslant$-minimum of $V(T)$. Now assume that $x_1, \ldots, x_m$ have been constructed with the above properties. The second vertex y_m of R_m is not on $T[v_0, x_m]$. As y_m and x_m are $\leqslant$–comparable, x_m lies on $T[v_0, y_m]$. Let x_{m+1} be the first vertex on $T(x_m, y_m]$ in $V(R_m)$. Then $x_1, \ldots, x_{m+1}$ has the stipulated properties. Now

$$\bar{R} = \bigcup_{m=1}^{\infty} T[v_0, x_m]$$

is a ray in T. Let $v_0, v_1, \ldots$ be the vertices of $\bar{R}$, where $v_h \in S(h)$. For any positive integer n, which is not a perfect square, we have a ray $\bar{R}_n$ in X with vertices $v_{p(n)}, v_{p(n^2)}, v_{p(n^4)}, \ldots$ Those rays are pairwise disjoint. This shows that each end of X is thick. The edges of T_d are exactly the edges of X_d, which have at least one end vertex of degree d or $d-1$ in $X(d)$. As v_0 is the only vertex of degree $d-1$ in X and in T, we have $\operatorname{Aut} X_d = \operatorname{Aut} T_d$.

However question which groups $G \leqslant \operatorname{Aut} X$ (X connected) contain proper translations is in general unanswered.

References

[1] L. Babai and M.E. Watkins, Connectivity of infinite graphs having a transitive group action, Arch. Math. 34 (1980) 90–96.
[2] H. Freudenthal, Über die Enden diskreter Räume in Gruppen, Comment Math. Helv. 17 (1944) 1–38.
[3] M. Gromov, Groups of polynomial growth and expanding maps, Inst. Hautes Études Sci. Publ. Math. 53 (1981) 53–78.
[4] R. Halin, Über unendliche Wege in Graphen, Math. Ann. 157 (1964) 125–137.
[5] R. Halin, Die Maximalzahl zweiseitig unendlicher Wege in Graphen, Math. Nachr. 44 (1970) 119–127.
[6] R. Halin, Automorphisms and endomorphisms of infinite locally finite graphs, Abh. Math. Sem. Univ. Hamburg 39 (1973) 251–283.
[7] H.A. Jung, Wurzelbäume und unendliche Wege in Graphen, Math. Nachr. 41 (1969) 1–22.
[8] H.A. Jung, Connectivity in infinite graphs, in: L. Mirsky, ed., Studies in Pure Mathematics (Academic Press, New York 1971) 137–147.

[9] H.A. Jung, A note on fragments of infinite graphs, Combinatorica 1 (1981) 285–288.
[10] H.A. Jung and M.E. Watkins, Fragments and automorphisms of infinite graphs, European J. Combin. 5 (1984) 149–162.
[11] N. Polat, Aspects topologiques de la séparation dans les graphes infinis I, Math. Z. 165 (1979) 73–100.
[12] N. Polat, Développements terminaux des graphes infinis, Math. Nachr. 107 (1982) 283–314.
[13] N. Seifter, Properties of graphs with polynomial growth, J. Combin. Theory Ser. B, to appear.
[14] J. Stallings, Group theory and three-dimensional manifolds (Yale Univ. Press, New Haven, 1971).
[15] V.I. Trofimov, Graphs with polynomial growth, USSR Sbornik 51 (1985) 405–417.
[16] W. Woess, Graphs and groups with tree-like properties, J. Combin. Theory Ser. B 6 (1988).

Discrete Mathematics 95 (1991) 135–167
North-Holland

Analyzing Nash-Williams' partition theorem by means of ordinal types

Igor Kříž*

Department of Mathematics, The University of Chicago, 5734S University Ave., Chicago, IL 60637, USA

Robin Thomas**

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

Received 16 October 1989
Revised 20 March 1990

Abstract

Kříž, I. and R. Thomas, Analyzing Nash-Williams' partition theorem by means of ordinal types, Discrete Mathematics 95 (1991) 135–167.

There is a natural way of assigning ordinal-valued functions to certain Ramsey-type theorems. In particular, they can be regarded as an extension of the classical notion of Ramsey numbers. The purpose of this paper is to obtain an estimate of these functions for the Nash-Williams' partition theorem.

0. Motivation

This paper is a continuation of our earlier work [7]. The motivation for [7] was twofold. We wanted to generalize 'Ramsey numbers' and thus obtain means to quantatively measure various infinite Ramsey type results, and by doing so we wanted to capture metamathematical phenomena such as the unprovability of certain results in specified logical systems. In this section we further explain these two ideas. The rest of the paper is independent of this section.

Let us consider the following two statements. (Our notation is standard and is explained in the next section.)

For every coloring $r:[\omega]^2 \to \{1, 2\}$ there exists an infinite set $A \subseteq \omega$ such that $r \mid [A]^2$ is constant. (0.0)

* Work on this paper partially supported by: NSF grant DMS 8903132, The Ohio State University and Charles University, Prague, Czechoslovakia

** Work on this paper partially supported by: DIMACS Center (Hill Center, Busch Campus, Rutgers University, New Brunswick, NJ 08903) and Charles University, Prague, Czechoslovakia

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

For every coloring $r:[\omega]^4 \to \{1, 2\}$ there exists an infinite set $B \subseteq \omega$ such that $r \mid [B]^4$ is constant. (0.1)

Is there any reason to believe that (0.1) is 'harder' or 'stronger' than (0.0)? In [7] we defined 'Ramsey numbers' corresponding to these statements and showed that the 'Ramsey number' of (0.0) is smaller than the 'Ramsey number' of (0.1). But what are the 'Ramsey numbers'? To answer this question we first formulate our definition of the notion of a Ramsey number.

Let U be an infinite set. By $U^{<\omega}$ we denote the set of all non-empty finite sequences of elements of U. Let $T \subseteq U^{<\omega}$. We define the *type* of T to be the least ordinal γ such that there exists a mapping $f: T \to \gamma$ with the property that if a, $b \in T$ and a is a strict initial segment of b then $f(a) > f(b)$. The type is undefined if no such ordinal γ exists.

Let $r:[U]^n \to \{1, \ldots, k\}$ be a coloring, and let $l_1, \ldots, l_k$ be natural numbers. We say that a sequence $a \in U^{<\omega}$ is i-monochromatic if $r(M) = i$ for every n-element set M consisting of terms of a. Let A_i be the set of all i-monochromatic sequences $(i = 1, \ldots, k)$, and let $A = (A_1, \ldots, A_k)$. If $g = (g_1, \ldots, g_k)$ is a k-tuple of functions $g_i: U^{<\omega} \to \mathrm{Ord}$ $(i = 1, \ldots, k)$ we say that $a \in U^{<\omega}$ is (A, g)-bad (written $a \in \mathrm{Bad}(A, g)$) if the following condition is satisfied.

If $b_1, b_2 \in A_i$ and b_1 is a strict initial segment of b_2, then $g_i(b_1) > g_i(b_2)$. (0.3)

Let $g^{l_i}: U^{<\omega} \to \{0, \ldots, l_i - 1\}$ be defined by

$$g^{l_i}(a) = \max(0, l_i - \text{length of } a),$$

and let $\bar{g} = (g^{l_1}, \ldots, g^{l_k})$. Notice that a is $(A, \bar{g})$-bad if and only if every i-monochromatic subsequence of a has length $\leqslant l_i$ for all $i = 1, \ldots, k$. Now it can be verified (see [7]) that the supremum of the types of all $\mathrm{Bad}(A, g)$ (taken over all colorings $r:[U]^n \to \{1, \ldots, k\}$ and all k-tuples of functions $g = (g_1, \ldots, g_k)$ with $g_i: U^{<\omega} \to \{0, \ldots, l_i - 1\}$) is $R(n; l_1, \ldots, l_k) - 1$, where $R(n; l_1, \ldots, l_k)$ stands for the usual Ramsey number. (It follows from Ramsey's theorem that the types are well defined.)

The generalization is now obvious. If $\gamma_1, \ldots, \gamma_k$ are ordinals, we define the *R-function* $\rho_n(\gamma_1, \ldots, \gamma_n)$ as the supremum of the types of $\mathrm{Bad}(A, g)$ over all colorings $r:[U]^n \to \{1, \ldots, k\}$ and all k-tuples of functions $g = (g_1, \ldots, g_k)$ with $g_i: U^{<\omega} \to \gamma_i$. So in particular, if $\gamma_1, \ldots, \gamma_k$ are all finite, then $\rho_n(\gamma_1, \ldots, \gamma_n) = R(n; \gamma_1, \ldots, \gamma_k) - 1$. The following is a corollary of one of the results of [7].

Theorem 0.4.

$$\left.\omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega^k}}}}}}\right\}(n-2)\ \textit{times} \leqslant \rho_n\underbrace{(\omega, \ldots, \omega)}_{k\ \textit{times}} \leqslant \left.\omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega^k}}}}}}\right\}(n-1)\ \textit{times},$$

In the definition of the R-function we did not use the fact that A_i consisted of

i-monochromatic sequences. All that was needed was that:

(i) if $b \in A_i$ and a is an initial segment of b, then $a \in A_i$, and

(ii) the type of $\mathrm{Bad}(A, g)$ is well defined for all A and g.

Therefore we may define (and examine) the R-functions corresponding to other Ramsey type results for which (ii) is satisfied. This was done in [7] for the Erdős–Szekeres Theorem and its generalization and for the Canonical Ramsey Theorem of Erdős and Rado, and for well-partially-ordered sets. In this paper we investigate the Nash-Williams' Partition Theorem. The upper bound is reasonably easy to obtain, but it is the lower bound which makes the analysis of the Nash-Williams' theorem so complicated.

To understand the connection with logic let us consider the following concept. Let Q be a partially ordered set with a partial ordering $\leqslant$. A (finite or infinite) sequence $q_1, q_2, \ldots$ of elements of Q is called *good* if there are indices i, j such that $i<j$ and $q_i \leqslant q_j$, and is called *bad* otherwise. The set Q is called well-partially-ordered (wpo) if every bad sequence of elements of Q is finite. If Q is wpo we define the *type* of Q, denoted by γ_Q, to be the least ordinal γ for which there exists a mapping f from the set of all non-empty bad sequences of elements of Q into γ such that

$$f(q_1, \ldots, q_n) > f(q_1, \ldots, q_{n+1})$$

for every bad sequence $(q_1, \ldots, q_{n+1})$. It is worth noting that if we define B to be the set of all sequences $(q_1, q_2, \ldots, q_n) \in Q^{<\omega}$ with $q_1 \leqslant q_2 \leqslant \cdots \leqslant q_n$, and $\mathbf{0}$ to be the constant mapping which is zero everywhere then a sequence $(q_1, \ldots, q_n)$ is bad if and only if it is $((B), \mathbf{0})$-bad in the sense defined earlier. Therefore the types of well-partially-ordered sets are a special case of our more general concept of an R-function.

Harvey Friedman [2] discovered that by 'miniaturizing' the assertion "Q is wpo" for certain wpo sets Q one can obtain statements of finite mathematics unprovable in relatively strong fragments of second order arithmetic. Here is an example of such a miniaturization.

> [2] For any positive c, there exists a positive integer $n = n(c)$ such that the following holds. If $T_1, T_2, \ldots, T_n$ is a finite sequence of finite trees with $|V(T_i)| \leqslant c \cdot i$ for all $i \leqslant n$, then there exist indices i and j such that $i < j \leqslant n$ and T_i is homeomorphically embeddable into T_j. (0.5)

That (0.5) is true can be easily derived from a theorem of Kruskal [8] which states that the set of finite trees with the partial ordering "to be homeomorphically embeddable into" is well-partially-ordered (see e.g. [15]). What is not so easy is to establish the unprovability part. The way this is usually done is by:

(i) proving that (0.5) implies that a miniaturization of the statement that a specified ordinal γ is well ordered, and

(ii) applying a result of logic that the above statement is unprovable.

The combinatorial content is now extracted in the proof of (i) (see [15] for details). It turns out that there is a connection between the strength of (0.5) and the type of the underlying wpo set. Therefore the type of a wpo set is a combinatorial invariant which has something to say about metamathematics of the well-partially-ordered set.

Thus our second motivation was to define the R-functions in such a way that this connection with logic will be preserved. And indeed, for example, the Parris–Harrington principle [11] can be looked at as a finite miniaturization of Theorem 0.4. Since the lower bound in 0.4 tends to ε_0 as $n \to \infty$, this is in accordance with the results of [11], because ε_0 is the proof-theoretic ordinal of Peano arithmetic.

In this paper we show an analogous result for the Nash-Williams' Partition Theorem, which implies that here the 'critical' ordinal is Γ_0. The ordinal Γ_0 is an ordinal much bigger than ε_0, defined as follows. Let $\varphi_0(\beta) = \omega^\beta$ and for $\alpha > 0$ let

$$\varphi_\alpha(\beta) = \beta\text{th common fixed point of all } \varphi_{\alpha'}(\alpha' < \alpha).$$

Notice that $\varphi_1(0) = \varepsilon_0$. Now Γ_0 is the least ordinal with the property that if $\alpha, \beta < \Gamma_0$ then $\varphi_\alpha(\beta) < \Gamma_0$.

In [3] Friedman, McAloon and Simpson derived from Nash-Williams' theorem a statement of finite mathematics unprobable in a theory called ATR_0, which is much stronger than Peano arithmetic. The proof-theoretic ordinal of ATR_0 is Γ_0 and so the relation of the main theorem of this paper to [3] is the same as the relation of Theorem 0.4 to [11].

Finally, let us say explicitly that we do not derive any unprovability results. We merely concentrate on combinatorial computation of the 'R-functions.'

1. Introduction

Conventions and Notation 1.1. Let Ω be an infinite set. The symbols $\Omega^{<\omega}$, $[\Omega]^{<\omega}$, Ω^ω, Ω^n, $[\Omega]^n$ denote the sets of non-empty finite sequences in Ω, the set of non-empty finite subsets of Ω, the set of infinite sequences in Ω, the Cartesian product of n copies of Ω and the set of subsets of Ω of cardinality n, respectively. If $a \in \Omega^{<\omega}$ then $|a|$ is the length of a. For $a = (a_1, a_2, \ldots)$, $b = (b_1, b_2, \ldots) \in \Omega^{<\omega} \cup \Omega^\omega$ we write $a \subseteq b$ and say that a is a subsequence of b if there are $j_1 < j_2 < \ldots$ such that $(a_1, a_2, \ldots) = (b_{j_1}, b_{j_2}, \ldots)$ and $a < b$ if $a \neq b$ and there is an n such that $a = (a_1, \ldots, a_n) = (b_1, \ldots, b_n)$. We shall also write $a \leqslant b$ if $a = b$ or $a < b$. For $a \in \Omega^{<\omega}$ we put $\downarrow a = \{b \in \Omega^{<\omega} \mid b \subseteq a\}$. Put, also, $\bar{a} = \{a_1, \ldots, a_n\}$ for $a = (a_1, \ldots, a_n)$. For $T \subseteq \Omega^{<\omega}$, put $\downarrow T = \bigcup \{\downarrow a \mid a \in T\}$, $\bar{T} = \{\bar{a} \mid a \in T\}$. Further, it will be convenient to denote the ith element of an arbitrary sequence $a \in \Omega^{<\omega}$ by a_i. For $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_m) \in \Omega^{<\omega}$, $a \cdot b$ denotes the concatenation $(a_1, \ldots, a_n, b_1, \ldots, b_m)$. An element $x \in \Omega^{<\omega}$ is often identified with the one-element sequence (x). For a function $f: X \to Y$ and for $M \subseteq X$ the

restriction of f to M is denoted by $f \mid M$. The image of a mapping f is denoted by $\operatorname{Im} f$. The *disjoint union* of sets X, Y is the set $X \sqcup Y = (X \times \{0\}) \cup (Y \times \{1\})$. As a rule, however, in $X \sqcup Y$, $X \times \{0\}$ is identified with X and $Y \times \{1\}$ is identified with Y. The class of all ordinals is denoted by Ord. For a set $X \subset \mathrm{Ord}$ we put

$$MX = \sup\{\alpha + 1 \mid \alpha \in X\}.$$

For ordinal numbers α, β we define their *natural sum* by $0 \oplus \alpha = \alpha \oplus 0 = \alpha$, and, for α, $\beta > 0$,

$$\alpha \oplus \beta = \min\{\kappa \in \mathrm{Ord} \mid (\forall \alpha' < \alpha)(\forall \beta' < \beta)\kappa > \alpha' \oplus \beta \quad \text{and} \quad \kappa > \alpha \oplus \beta'\}$$

for a set X, $|X|$ denotes the cardinality of X. We define $\alpha + \beta$ as the type of α followed by β and $\alpha\beta$ as the type of $\alpha \times \beta$ with the lexicographical ordering

$$(x, y) < (z, t) \quad \text{if } x < z \quad \text{or} \quad (x = z \text{ and } y < t).$$

Definition 1.2. A subset $S \subseteq [\Omega]^{<\omega}$ is called a *Sperner system* in Ω if

$$(\forall a, b \in S)(\alpha \subseteq b \rightarrow a = b).$$

We now introduce the Nash-Williams' partition theorem, which is the central object of our study.

Theorem 1.3. (The Nash-Williams' Partition Theorem). *For any Sperner system $S \subseteq [\Omega]^{<\omega}$ and any partition $r : S \rightarrow \{1, \ldots, k\}$ there exists an infinite subset $\Omega' \subseteq \Omega$ such that $r \mid S \cap [\Omega']^{<\omega}$ is constant* (see [10]).

1.4. The language of ordinal types in Ramsey theory (see [7])

Definition 1.4.1. A *tree* is a couple $(T, \leqslant)$ where T is a set and $\leqslant$ is a partial ordering on T such that for every $t \in T$ the set $\{t' \in T \mid t' \leqslant t\}$ is a finite linearly ordered set. A tree $(T, \leqslant)$ is said to be rayless if T contains no infinite subset linearly ordered by $\leqslant$. Note that $(\Omega^{<\omega}, \preccurlyeq)$ is a tree. More generally, all subsets of $\Omega^{<\omega}$ will be regarded as trees with this ordering. A *character* on a tree $(T, \leqslant)$ is a function

$$\varphi : T \rightarrow \mathrm{Ord}$$

such that

$$(\forall x, y \in T)(x < y \rightarrow \varphi(x) > \varphi(y)).$$

A tree $(T, \leqslant)$ is rayless if and only if there is a character on $(T, \leqslant)$. In that case, we define the *ordinal type* γ_T by

$$\gamma_T = \min\{\gamma \mid \text{there is a character } \varphi : T \rightarrow \gamma\}.$$

The Main Definition 1.4.2. A *sheaf* (in Ω) is a subset $A \subseteq \Omega^{<\omega}$ such that ($a \in A$ and $b < a) \to b \in A$. A *k-sheaf* is a k-tuple of sheaves. By abuse of language, we identify 1-sheaves and sheaves. A *k-system* is a set of k-sheaves. A sheaf A is said to have the *Ramsey property* if

$$(\forall a \in \Omega^{\omega})(\exists b \in \Omega^{\omega})(b \subseteq a \text{ and } (\forall c < b)(c \in A)).$$

A k-sheaf $(A_1, \ldots, A_k)$ is said to have the Ramsey property if the sheaf $A_1 \cup \cdots \cup A_k$ does. A k-system $\mathcal{R}$ is said to have the *Ramsey property* if each $A \in \mathcal{R}$ does.

Let $A = (A_1, \ldots, A_k)$ be a k-sheaf. A $(\gamma_1, \ldots, \gamma_k)$*-testing* is a k-tuple $g = (g_1, \ldots, g_k)$ of functions $g_i : \Omega^{<\omega} \to \gamma_i$. A sequence $a \in \Omega^{<\omega}$ is called (A, g)-*bad* if each g_i is a character on $\downarrow a \cap A_i$. The subtree of $\Omega^{<\omega}$ of all (A, g)-bad sequences will be denoted by $\mathrm{Bad}(A, g)$.

The following is a result from [7].

Proposition 1.4.2.1. *A sheaf A in Ω has the Ramsey property if and only if* $\mathrm{Bad}(A, g)$ *is a rayless tree for all ordinals* $\gamma_1, \ldots, \gamma_k < |\Omega|^+$ *and every* $(\gamma_1, \ldots, \gamma_k)$*-testing* g.

If A has the Ramsey property we define the R-function $\varphi_A : \mathrm{Ord}^k \to \mathrm{Ord}$ by

$$\varphi_A(\gamma_1, \ldots, \gamma_k) = \sup\{\gamma_{\mathrm{Bad}(A,g)} \mid g \text{ is a } (\gamma_1, \ldots, \gamma_k)\text{-testing}\}.$$

Similarly, for a k-system $\mathcal{S}$ which has the Ramsey property we define

$$\varphi_{\mathcal{S}}(\gamma_1, \ldots, \gamma_k) = \sup\{\varphi_A(\gamma_1, \ldots, \gamma_k) \mid A \in \mathcal{S}\}.$$

1.5 Definition of the ψ-functions

Let $k \geqslant 1$ be an integer, S be a Sperner system and $r : S \to \{1, \ldots, k\}$ be a partition. Define a k-sheaf $A^{S,r} = (A_1^{S,r}, \ldots, A_k^{S,r})$ by putting

$$A_i^{S,r} = \{a \in \Omega^{<\omega} \mid \mathrm{Im}(r \mid (\overline{\downarrow a} \cap S)) = \{i\}\}.$$

The *Nash-Williams k-system* $\mathcal{R}_k$ is defined by

$$\mathcal{R}_k = \{A^{S,r} \mid S \text{ is a Sperner system and } r : S \to \{1, \ldots, k\}\}.$$

It follows from Theorem 1.3 that $\mathcal{R}_k$ has the Ramsey property. We put

$$\psi(\gamma_1, \ldots, \gamma_k) = \varphi_{\mathcal{R}_k}(\gamma_1, \ldots, \gamma_k).$$

Notation 1.6. Define functions:

$\varphi_0(\beta) = \omega^{\beta}$,

$\varphi_\alpha(\beta) =$ the βth common fixed–point of all $\varphi_{\alpha'}$, $\alpha' < \alpha$, for $\alpha > 0$,

$g(\alpha) = \sup\{\varphi_\gamma(\beta) \mid \gamma \geqslant 2 \text{ and } (\beta + 1)\omega^{2\gamma\omega} \geqslant \alpha\}$.

The Main Theorem 1.7. *Let $k \geqslant 2$ be an integer, let $\gamma_1, \ldots, \gamma_k$ be ordinals with $g(0) \leqslant \gamma_i < |\Omega|^+$ $(i = 1, \ldots, k)$, and let $\alpha = \min\{\gamma_1, \ldots, \gamma_k\}$. Then*

$$g(\alpha) \leqslant \psi(\gamma_1, \ldots, \gamma_k) \leqslant \varphi_\alpha(\gamma_1 \oplus \cdots \oplus \gamma_k).$$

Proof. Follows from Theorem 2.4 and Theorem 3.16 below. □

2. Prelude—the upper bound

2.1. Let S be a Sperner system and let $T \subseteq \Omega^{<\omega}$ be a rayless sheaf with $\downarrow T = T$. We define

$$C_T(S) = \{a \in T \mid \bar{a} \notin S \text{ and there is a } b \in \Omega^{<\omega} \text{ such that } \overline{a \cdot b} \in S \text{ and } a \cdot b \in T\}$$

and $c_T(S) = \gamma_{C_T(S)}$.
We remark that

$$\overline{C_T(S)} \cap S = \emptyset. \tag{2.1.1}$$

Let, for $\alpha \in \mathrm{Ord}$,

$$\psi_\alpha(\gamma_1, \ldots, \gamma_k) = \sup\{\gamma_T \mid T \text{ is rayless, } \downarrow T = T,\ c_T(S) \leqslant \alpha \text{ and there exists } (A_1, \ldots, A_k) \in R_k \text{ such that } \gamma_{T \cap A_i} \leqslant \gamma_i\ (i = 1, \ldots, k)\}.$$

Lemma 2.2. *Let $k \geqslant 1$ be an integer, let $\gamma_1, \ldots, \gamma_k$ be ordinals and let $\alpha = \min\{\gamma_1, \ldots, \gamma_k\}$. Then $\psi(\gamma_1, \ldots, \gamma_k) = \psi_\alpha(\gamma_1, \ldots, \gamma_k)$.*

Proof. We first prove '$\geqslant$'. Let T be a rayless sheaf with $\downarrow T = T$ and such that $c_T(S) \leqslant \alpha$ and $\gamma_{T \cap A_i} \leqslant \gamma_i$ $(i = 1, \ldots, k)$ for some $A = (A_1, \ldots, A_k) \in \mathscr{R}_k$. For $i = 1, \ldots, k$, let $g_i : T \cap A_i \to \gamma_i$ be characters and let $g = (g_1, \ldots, g_k)$. Then $T \subseteq \mathrm{Bad}(A, g)$, and hence $\gamma_T \leqslant \psi(\gamma_1, \ldots, \gamma_k)$.

To prove '$\leqslant$', let g be a $(\gamma_1, \ldots, \gamma_k)$-testing and let $A = (A_1, \ldots, A_k) \in \mathscr{R}_k$. By (2.1.1), $C_{\mathrm{Bad}(A,g)}(S) \subset \mathrm{Bad}(A, g) \cap A_i$ for every $i = 1, \ldots, k$, and hence

$$c_{\mathrm{Bad}(A,g)}(S) \leqslant \gamma_{\mathrm{Bad}(A,g) \cap A_i} \leqslant \gamma_i \ (i = 1, \ldots, k).$$

Therefore,

$$\gamma_{\mathrm{Bad}(A,g)} \leqslant \psi_a(\gamma_1, \ldots, \gamma_k). \qquad \square$$

Theorem 2.3. *Let $k \geqslant 1$ be an integer and let $\gamma_1, \ldots, \gamma_k$ be ordinals. Then:*

(a) $\psi_0(\gamma_1, \ldots, \gamma_k) \leqslant \gamma_1 \oplus \cdots \oplus \gamma_k$.

(b) *For $\alpha > 0$ we have*

$$\psi_\alpha(\gamma_1, \ldots, \gamma_k) \leqslant M\{\psi_{\alpha'}(\psi_\alpha(\gamma_1', \ldots, \gamma_k), \ldots, \psi_\alpha(\gamma_1, \ldots, \gamma_k')) \mid \alpha' < \alpha,\ \gamma_i' < \gamma_i\}.$$

Proof. To prove (a), observe that $c_T(S) = 0$ implies

$$S \supseteq \{\{x\} \mid (x) \in T\}.$$

Now (a) follows from Theorem 6.1 in [7].

We prove (b) by induction. Let α and $\gamma_1, \ldots, \gamma_k$ be ordinals and assume that (b) holds for all α' and $\gamma'_1, \ldots, \gamma'_k$ such that either $\alpha' < \alpha$, or $\alpha' = \alpha$ and $\gamma'_1 \leqslant \gamma_1, \ldots, \gamma'_k \leqslant \gamma_k$ and at least one of these inequalities is strict. Let T be a rayless sheaf with $\downarrow T = T$, and let $A^{S,r} = (A_1^{S,r}, \ldots, A_k^{S,r}) \in \mathscr{R}_k$ be such that $\gamma_{T \cap A_i^{S,r}} \leqslant \gamma_i$ $(i = 1, \ldots, k)$ and $c_T(S) \leqslant \alpha$. We must estimate γ_T. Let $x \in \Omega$ and let $T_x = \{a \in \Omega^{<\omega} \mid (x) \cdot a \in T\}$. Then T_x is a rayless sheaf with $\downarrow T_x = T_x$ and such that

$$T_x \subseteq T. \tag{2.3.1}$$

Let

$$S_x = \{m \in [\Omega]^{<\omega} \mid x \notin m \text{ and } m \cup \{x\} \in S\}$$

and for $m \in S_x$, let $r_x(m) = r(m \cup \{x\})$. It is easily seen that

$$c_{T_x}(S_x) < c_T(S). \tag{2.3.2}$$

By (2.3.1),

$$\{xa \mid a \in T_x \cap A_i^{S_x, r_x} \cap A_i^{S,r}\} \subseteq T \cap A_i^{S,r} \quad \text{for } i = 1, \ldots, k,$$
$$T_x \cap A_i^{S_x, r_x} \cap A_j^{S,r} \subseteq T \cap A_j^{S,r} \quad \text{for } i \neq j,\ i, j = 1, \ldots, k.$$

Thus,

$$\gamma_{T_x \cap A_i^{S_x, r_x} \cap A_i^{S,r}} < \gamma_i \quad \text{for } i = 1, \ldots, k,$$
$$\gamma_{T_x \cap A_i^{S_x, r_x} \cap A_j^{S,r}} \leqslant \gamma_j \quad \text{for } i, j = 1, \ldots, k.$$

Since, obviously, $c_{T_x \cap A_i^{S_x, r_x}}(S) \leqslant c_T(S) \leqslant \alpha$, it follows from the induction hypothesis that

$$\gamma_{T_x \cap A_i^{S_x, r_x}} \leqslant \psi_\alpha(\gamma_1, \ldots, \gamma'_i, \ldots, \gamma_k)$$

for some $\gamma'_i < \gamma_i$. By this, (2.3.2) and the induction hypothesis,

$$\gamma_{T_x} \leqslant \psi_{\alpha'}(\psi_\alpha(\gamma'_1, \gamma_2, \ldots, \gamma_k), \ldots, \psi_\alpha(\gamma_1, \ldots, \gamma_{k-1}, \gamma'_k)),$$

where $\alpha' = c_{T_x}(S_x)$. Since x was arbitrary, the result follows. □

Theorem 2.4. *Let $k \geqslant 1$ be an integer, let $\gamma_1, \ldots, \gamma_k < |\Omega|^+$ be ordinals and let $\alpha = \min\{\gamma_1, \ldots, \gamma_k\}$. Then we have*

$$\psi(\gamma_1, \ldots, \gamma_k) \leqslant \varphi_\alpha(\gamma_1 \oplus \cdots \oplus \gamma_k).$$

Proof. Observe that

$$\varphi_0(\beta) \geqslant \beta, \qquad \varphi_\alpha(\beta) \geqslant M\{\varphi_{\alpha'}(\varphi_\alpha(\beta')) \mid \alpha' > \alpha,\ \beta' < \beta\}.$$

Our result now follows from 2.2 and 2.3. □

3. Fugue

3.0. Outline of the proof

Let $tp\ X$ denote the ordinal type of a well-ordered set X. We are going to construct, for all ordinals α, γ, well-ordered sets $T_\gamma(\alpha)$ such that

$$\text{tp } T_0(\alpha) \geqslant 2^\alpha \quad \text{for all } \alpha \geqslant 0, \tag{3.0.1}$$

$$\text{tp } T_\delta(\text{tp } T_\gamma(\alpha)) \leqslant \text{tp } T_\gamma(\alpha) \quad \text{for all } \alpha \geqslant 0 \text{ and all } \gamma, \delta \geqslant 0 \text{ with } \gamma \geqslant \delta + \omega, \tag{3.0.2}$$

$$\text{tp } T_\gamma(\alpha) \leqslant \text{tp } T_\gamma(\beta) \quad \text{for all } \gamma \geqslant 0 \text{ and } 0 \leqslant \alpha \leqslant \beta. \tag{3.0.3}$$

We are going to construct a Sperner system S_γ in $T_\gamma(\alpha)$, a coloring $r_\gamma : S_\gamma \to 0, 1\}$ and an $((\alpha+1)\cdot\omega^{2^\gamma}, (\alpha+1)\cdot\omega^{2^\gamma})$-testing g_γ such that

$$\text{Dec } T_\gamma(\alpha) \subseteq \text{Bad}(A^{S_\gamma, r_\gamma}, g_\gamma) \cup \{\emptyset\}. \tag{3.0.4}$$

(Here Dec X denotes the set of all decreasing sequences in X.)

We claim that

$$\varphi_\gamma(\alpha) \leqslant \text{tp } T_{\gamma\cdot\omega}(\alpha) \quad \text{for all } \alpha \geqslant 0 \text{ and all } \gamma \geqslant 2. \tag{3.0.5}$$

For let, for $\gamma \in \text{Ord}$, $g_\gamma : \text{Ord} \to \text{Ord}$ be the function defined by $g_\gamma(\alpha) = \text{tp } T_{\gamma\cdot\omega}(\alpha)$. From (3.0.1)–(3.0.3) we deduce

$$g_0(\alpha) \geqslant 2^\alpha \quad \text{for all } \alpha \geqslant 0, \tag{3.0.6}$$

$$g_\gamma(\alpha) \text{ is a fixed point of } g_\delta \quad \text{for every } \delta < \gamma \text{ and every } \alpha \geqslant 0, \tag{3.0.7}$$

$$g_\gamma(\beta) < g_\gamma(\alpha) \quad \text{for all } \gamma \geqslant 0 \text{ and all } 0 \leqslant \beta < \alpha. \tag{3.0.8}$$

Hence, $g_\gamma(\alpha) \geqslant \varphi'_\gamma(\alpha)$ for all α, $\gamma \geqslant 0$ where $\varphi'_0(\alpha) = 2^\alpha$, $\varphi'_\gamma(\alpha)$ is the αth common fixed point of $\varphi_{\gamma'}$, $\gamma' < \gamma$ for $\gamma > 0$.
But $\varphi'_\gamma(\alpha) = \varphi_\gamma(\alpha)$ for all $\alpha \geqslant 0$ and $\gamma \geqslant 2$. (3.0.5) follows.

Now condition (3.0.4) implies

$$\gamma_{\text{Dec}\, T_\gamma(\alpha)} \leqslant \psi((\alpha+1)\cdot\omega^{2^\gamma}, (\alpha+1)\cdot\omega^{2^\gamma}). \tag{3.0.9}$$

(Recall that $T_\gamma(\alpha)$ is regarded as a tree under the relation $\preccurlyeq$.) Assume, without loss of generality,

$$T_\gamma(\alpha) \subseteq \Omega. \tag{3.0.10}$$

Since obviously $\text{tp } T_\gamma(\alpha) = \gamma_{\text{Dec}\, T_\gamma(\alpha)\setminus\{0\}}$, (3.0.5) and (3.0.9) imply the desired lower bound.

3.1. The language of category theory

The concepts used here can be found in any elementary text-book of category theory, e.g. [9]. We shall assume that the reader knows the concepts of a

category, *functor* and *natural transformation*. For a category C, let Obj C denote the class of objects of C and let Mor C denote the class of morphisms of C. For x, $y \in \text{Obj } C$, $C(x, y)$ denotes the set of morphisms from x to y in C. In the sequel, we shall mostly use the category Set of sets and mappings and the category $\mathcal{W}$ of linearly ordered sets and strictly monotone mappings.

Let Δ be a partially ordered set (regarded as a category in the usual way, i.e. objects are elements of Δ and, for x, $y \in \Delta$, $\Delta(x, y) = \emptyset$ if $x \nleqslant y$ and $|\Delta(x, y)| = 1$ otherwise), let C be an arbitrary category and let

$$F : \Delta \to C$$

be a functor. A *colimit* of F is an object $x \in \text{Obj } C$ together with a bunch of morphisms

$$(\varphi_z \in C(F(z), x) \mid z \in \Delta)$$

such that, for $\alpha \in \Delta(y, z)$,

$$\varphi_z \circ F(\alpha) = \varphi_y$$

and whenever there is an object $t \in \text{Obj } C$ together with a bunch of morphisms

$$(\psi_z \in C(F(z), t) \mid z \in \Delta)$$

such that, for $\alpha \in \Delta(y, z)$,

$$\psi_z \circ F(\alpha) = \psi_t$$

then there is a unique $g \in C(x, t)$ such that for each $z \in \text{Obj } \Delta$

$$\psi_z = g \circ \varphi_z.$$

A partially ordered set Δ is called *directed* if

$$(\forall x, y \in \text{Obj } \Delta)(\exists z \in \text{Obj } \Delta(x \leqslant z \text{ and } y \leqslant z).$$

A *directed colimit* is a colimit of a functor

$$F : \Delta \to C$$

where Δ is directed. We say that C *has directed colimits* if for each functor $F : \Delta \to C$ where Δ is directed there is a colimit.

Fact 3.1.1. *The categories* $\mathcal{W}$, Set *have directed colimits.*

Let $G : C_1 \to C_2$ be a functor. We say that G *preserves directed colimits* if for each functor $F : \Delta \to C_1$ where Δ is directed and for each colimit $(x, (\varphi_z \mid z \in \Delta))$ of $F(G(x), (G(\varphi_z) \mid z \in \Delta))$ is a colimit of GF.

Fact 3.1.2. *The forgetful functor* $W : \mathcal{W} \to$ Set *preserves directed colimits.*

Compositions 3.1.3. In this paper we shall generally denote compositions of morphisms by '$\circ$' and compositions of functors by '$\cdot$'. Let $S:C\to D$, $T:C\to D$ be functors. A *natural transformation* $\varphi:S\to T$ is a system of morphisms $(\varphi_x:Sx\to Tx \mid x\in \mathrm{Obj}\, C)$ such that for $f\in C(x,y)$ the following diagram commutes:

$$\begin{array}{ccc} Sx & \xrightarrow{Sf} & Sy \\ {\scriptstyle \varphi_x}\downarrow & & \downarrow{\scriptstyle \varphi_y} \\ Tx & \xrightarrow{Tf} & Ty. \end{array}$$

It is common to write φ instead of φ_x, φ_y if there is no danger of confusion. In natural transformations, we have two kinds of compositions: Let first $R:C\to D$, $S:C\to D$, $T:C\to D$ be functors and let $\sigma:R\to S$, $\tau:S\to T$ be natural transformations. The *vertical composition* of σ and τ is the natural transformation $\tau\circ\sigma:R\to T$ given by $(\tau\circ\sigma)_x=\tau_x\circ\sigma_x$.

Let, on the other hand, $S:C\to D$, $T:C\to D$, $S':D\to E$, $T':D\to E$ be functors and let $\tau:S\to T$, $\tau':S'\to T'$ be natural transformations. The *horizontal composition* of τ and τ' is the natural transformation $\tau'\cdot\tau:(S'\cdot S)\to(T'\cdot T)$ given by $(\tau'\cdot\tau)_x=\tau'_{Tx}\circ S'(\tau_x)$. Note that, by naturality, we also have $(\tau'\cdot\tau)_x=T'(\tau_x)\circ\tau'_{Sx}$ (Proof: For any z, $t\in\mathrm{Obj}\,D$ and any $g\in D(z,t)$, $\tau'_t\circ S'(g)=T'(g)\circ\tau'_z$. Put $z=Sx$, $t=Tx$, $g=\tau_x$.)

Put, in particular, $T'\cdot\tau=(\mathrm{Id}_{T'})\cdot\tau$, $\tau'\cdot T=\tau'\cdot(\mathrm{Id}_T)$. We see easily that $(T'\cdot\tau)_x=T'(\tau_x)$, $(\tau'\cdot T)_x=\tau'_{Tx}$. Thus, in general, $\tau'\cdot\tau=(T'\cdot\tau)\circ(\tau'\cdot S)=(\tau'\cdot T)\circ(S'\cdot\tau)$.

We would like to warn the reader that some authors use the symbols $\circ,\cdot$ in different meanings.

3.2. The functors $T_\beta:\mathcal{W}\to\mathcal{W}$ and natural transformations $\kappa^\beta:\mathrm{Id}\to T_\beta$, $\kappa^\beta_\alpha:T_\alpha\to T_\beta$, $\alpha<\beta\in\mathrm{Ord}$

3.2.1. We first define T_0 on objects. Let $\alpha\in\mathrm{Obj}\,\mathcal{W}$. We define $T_0(\alpha)$ as the set $[\alpha]^{<\omega}\sqcup\alpha$ together with the ordering $<$ given by the following conditions:

$$(\forall x\in\alpha)(\forall y\in T_0(\alpha))(x<\{x\}\text{ and }(y\leqslant x\text{ or }y\geqslant\{x\})) \tag{3.2.1.1}$$

$$(\forall x,y\in T_0(\alpha)\backslash\alpha)(x<y\text{ iff }(y\backslash x\neq\emptyset\text{ and }(\forall z\in x\backslash y)z<\max(y\backslash x))). \tag{3.2.1.2}$$

It is easily seen that conditions (3.2.1.1) and (3.2.1.2) indeed specify a unique linear ordering on $T_0(\alpha)$, whose restriction to α coincides with the original ordering on α. For α, $\beta\in\mathrm{Obj}\,\mathcal{W}$ and $\varphi\in\mathcal{W}(\alpha,\beta)$ define $T_0(\varphi)$ by

$$T_0(\varphi)(x)=\varphi(x)\quad\text{for } x\in\alpha,$$
$$T_0(\varphi)(m)=\{\varphi(x)\mid x\in m\}\quad\text{for } m\in[\alpha]^{<\omega}.$$

(It is easy to check that $T_0(\varphi)$ is strictly increasing if φ is.) Now define κ^0, $\mathrm{Id}\to T_0$ by $(\kappa^0)_\alpha(x)=x$ (for $x\in\alpha$).

3.2.2. Now assume T_γ already defined. Put $T_{\gamma+1} = (T_\gamma)^2$ (T^2 stands for $T \cdot T$). Put, further, for $\beta < \gamma$,

$$\kappa^{\gamma+1} = \kappa^\gamma \cdot \kappa^\gamma = (\kappa^\gamma \cdot T_\gamma) \circ \kappa^\gamma,$$
$$\kappa_\gamma^{\gamma+1} = T_\gamma \cdot \kappa^\gamma,$$
$$\kappa_\beta^{\gamma+1} = \kappa_\gamma^{\gamma+1} \circ \kappa_\beta^\gamma.$$

Note that

$$\kappa^{\gamma+1} = \kappa_\gamma^{\gamma+1} \circ \kappa^\gamma. \tag{3.2.2.1}$$

3.2.3. Now assume $T_{\gamma'}$, already defined for all $\gamma' < \gamma$ where γ is a limit ordinal. We define T_γ as the colimit of the (commutative) diagram of functors

$$\begin{tikzcd} T_0 \arrow[r, "\kappa_0^1"] & T_1 \arrow[r] & \cdots \arrow[r] & T_{\gamma'} \arrow[r] \arrow[rrr, bend right, "\kappa_{\gamma'}^{\gamma''}"'] & \cdots \arrow[r] & T_{\gamma''} \arrow[r] & \cdots . \end{tikzcd}$$

Let, also, κ_β^γ be the colimit mappings arising from the system $(\kappa_\beta^{\gamma'} \mid \gamma' < \gamma)$ for $\beta < \gamma$. Now define

$$\kappa^\gamma = \kappa_\beta^\gamma \circ \kappa^\beta.$$

Observe that this definition does not depend on the choice of $\beta < \gamma$.

3.3. The 'square' transformation $\iota_\gamma^\beta: T_\gamma \to T_\beta T_\gamma$, β, $\gamma \in$ Ord

We shall define natural transformations $\iota_\gamma^\beta: T_\gamma \to T_\beta T_\gamma$ by transfinite induction on γ in the following way.

3.3.1. Put, for $\gamma \leqslant \beta$, $\iota_\gamma^\beta = \kappa_\gamma^\beta \cdot \kappa^\gamma$.

3.3.2. For $\gamma = \beta + 1$, put $\iota_{\beta+1}^\beta = T_\beta \cdot \kappa_\beta^{\beta+1}$. (We have $\iota_{\beta+1}^\beta: T_{\beta+1} = T_\beta \cdot T_\beta \to T_\beta \cdot T_\beta \cdot T_\beta = T_\beta T_{\beta+1}$).

3.3.3. Let ι_γ^β be defined for $\gamma > \beta$. Put $\iota_{\gamma+1}^\beta = \iota_\gamma^\beta \cdot T_\gamma$.

3.3.4. Now let $\iota_{\gamma'}^\beta$ be already defined for all $\gamma' < \gamma$ where γ is a limit ordinal. First note that, by the naturality of $\iota_{\gamma'}^\beta$, the following diagram commutes:

$$\begin{tikzcd} T_{\gamma'} \arrow[r, "T_{\gamma'} \cdot \kappa^\gamma"] \arrow[d, "\iota_{\gamma'}^\beta"] & T_{\gamma'} \cdot T_{\gamma'} \arrow[d, "\iota_{\gamma'}^\beta \cdot T_{\gamma'}"] \\ T_\beta T_{\gamma'} \arrow[r, "T_\beta T_{\gamma'} \cdot \kappa^\gamma"] & T_\beta T_{\gamma'} T_{\gamma'}. \end{tikzcd} \tag{3.3.4.1}$$

Changing the notation, we get the commutative diagram (for $\gamma' \geqslant \beta + 1$):

$$\begin{tikzcd} T_{\gamma'} \arrow[r, "\kappa_{\gamma'}^{\gamma'+1}"] \arrow[d, "\iota_{\gamma'}^\beta"] & T_{\gamma'+1} \arrow[d, "\iota_{\gamma'+1}^\beta"] \\ T_\beta T_{\gamma'} \arrow[r, "T_\beta \kappa_{\gamma'}^{\gamma'+1}"] & T_\beta T_{\gamma'+1}. \end{tikzcd} \tag{3.3.4.2}$$

This allows us to define $\iota^{\beta}_{\gamma}: T_\gamma \to T_\beta T_\gamma$ as the colimit mapping of a (by induction, commutative) diagram

$$\begin{tikzcd} \cdots \arrow[r] & T_{\gamma'} \arrow[rr, "\kappa^{\gamma'}_{\gamma''}"] \arrow[dr, "(T_\beta \cdot \kappa^{\gamma}_{\gamma'}) \circ \iota^{\beta}_{\gamma'}"'] & & T_{\gamma''} \arrow[r] \arrow[dl, "(T_\beta \cdot \kappa^{\gamma}_{\gamma''}) \circ \iota^{\beta}_{\gamma''}"] & \cdots. \\ & & T_\beta T_\gamma & & \end{tikzcd}$$

The reader may feel that the restriction $\gamma' \geqslant \beta + 1$ in Diagram (3.3.4.2) violates the beauty. Indeed, it is not necessary. For $\gamma' = \beta$ we observe that $\iota^{\beta}_{\beta} = \kappa^{\beta+1}_{\beta} = T_\beta \kappa^{\beta}$, $\iota^{\beta+1}_{\beta} = T_\beta \kappa^{\beta+1}_{\beta} = T^2_\beta \kappa^{\beta}$. For $\gamma' < \beta$, we observe that

$$\begin{aligned} (T_\beta \kappa^{\gamma'+1}_{\gamma'}) \circ \iota^{\beta}_{\gamma'} &= (T_\beta \kappa^{\gamma'+1}_{\gamma'}) \circ (\kappa^{\beta}_{\gamma'} \cdot \kappa^{\gamma'}) = \kappa^{\beta}_{\gamma'} \cdot (\kappa^{\gamma'+1}_{\gamma'} \circ \kappa^{\gamma'}) \\ &= \kappa^{\beta}_{\gamma'} \cdot \kappa^{\gamma'+1} = (\kappa^{\beta}_{\gamma'+1} \cdot \kappa^{\gamma'+1}) \circ \kappa^{\gamma'+1}_{\gamma'} = \iota^{\beta}_{\gamma'+1} \circ \kappa^{\gamma'+1}_{\gamma'}. \end{aligned}$$

Summing up our results, we get the following commutative diagram valid generally for $\gamma' < \gamma$:

$$\begin{tikzcd} T_{\gamma'} \arrow[r, "\kappa^{\gamma}_{\gamma'}"] \arrow[d, "\iota^{\beta}_{\gamma'}"] & T_\gamma \arrow[d, "\iota^{\beta}_{\gamma}"] \\ T_\beta T_{\gamma'} \arrow[r, "T_\beta \kappa^{\gamma}_{\gamma'}"'] & T_\beta T_\gamma. \end{tikzcd} \tag{3.3.4.3}$$

3.4. The total-image transformation tim_γ

Denote by W the forgetful functor $W: \mathcal{W} \to \mathrm{Set}$ and by $K: \mathrm{Set} \to \mathrm{Set}$ the functor given by

$$K(m) = [m]^{<\omega},\ K(f) \text{ is the map induced by } f.$$

Note that there is a natural transformation

$$U: K^2 \to K,\ U(m) = \bigcup m,$$

called the *union*. We also have a natural transformation

$$S: \mathrm{Id} \to K \text{ given by } S(x) = \{x\}$$

called the *singleton*. It holds that

$$U \circ (K \cdot S) = \mathrm{Id}_K = U \circ (S \cdot K).$$

We shall now define natural transformations

$$\mathrm{tim}_\gamma: W \cdot T_\gamma \to K \cdot W$$

to satisfy

$$\mathrm{tim}_\gamma \circ (W \cdot K^\gamma) = S \cdot W.$$

3.4.1. Define tim_0 by $(\mathrm{tim}_0)_\alpha(\beta) = \{\beta\}$ for $\beta \in \alpha$ and $(\mathrm{tim}_0)_\alpha(m) = m$ else.

3.4.2. Now assume tim_γ already defined. To define $\mathrm{tim}_{\gamma+1}$, put

$$(\mathrm{tim}_{\gamma+1}: W \cdot T_\gamma \cdot T_\gamma \to K \cdot W) = (U \cdot W) \circ (\mathrm{K} \cdot \mathrm{tim}_\gamma) \circ (\mathrm{tim}_\gamma \cdot T_\gamma).$$

3.4.3. Now assume that $\mathrm{tim}_{\gamma'}$ has been already defined for all $\gamma' < \gamma$ where γ is a limit ordinal. By the naturality of $\mathrm{tim}_{\gamma'}$, we have the following commutative diagrams (the second one arises from the first one by change of notation):

$$\begin{tikzcd} WT_{\gamma'} \arrow[r, "WT_{\gamma'} \cdot \kappa^{\gamma'}"] \arrow[d, "\mathrm{tim}_{\gamma'}"'] & WT_{\gamma'}T_{\gamma'} \arrow[d, "\mathrm{tim}_{\gamma'} \cdot T_{\gamma'}"] \\ KW \arrow[r, "KW \cdot \kappa^{\gamma'}"'] & KWT_{\gamma'} \end{tikzcd} \qquad \begin{tikzcd} WT_{\gamma'} \arrow[r, "W\kappa_{\gamma'}^{\gamma'+1}"] \arrow[d, "\mathrm{tim}_{\gamma'}"'] & WT_{\gamma'+1} \arrow[d, "\mathrm{tim}_{\gamma'} \cdot T_{\gamma'}"] \\ KW \arrow[r, "KW \cdot \kappa^{\gamma'}"] & KWT_{\gamma'}. \end{tikzcd}$$

Now since

$$UW \circ (K \cdot \mathrm{tim}_{\gamma'}) \circ (KW_\kappa^{\gamma'}) = \mathrm{Id}_{KW},$$

we get the following commutative diagrams:

$$\begin{tikzcd} & WT_{\gamma'} \arrow[r, "W\kappa_{\gamma'}^{\gamma'+1}"] \arrow[dl, "\mathrm{tim}_{\gamma'}"'] & WT_{\gamma'+1} \arrow[d, "\mathrm{tim}_{\gamma'} \cdot T_{\gamma'}"] \\ KW & KKW \arrow[l, "U \cdot W"'] \arrow[r, "K\,\mathrm{tim}_{\gamma'}"] & KWT_{\gamma'}, \end{tikzcd}$$

and consequently

$$\begin{tikzcd} WT_{\gamma'} \arrow[rr, "W\kappa_{\gamma'}^{\gamma'+1}"] \arrow[dr, "\mathrm{tim}_{\gamma'}"'] & & WT_{\gamma'+1} \arrow[dl, "\mathrm{tim}_{\gamma'+1}"] \\ & KW. & \end{tikzcd}$$

This allows us to define tim_γ as the limit mapping of the (by induction, commutative) diagram

$$\begin{tikzcd} \arrow[r] & WT_{\gamma'} \arrow[r, "W\kappa_{\gamma'}^{\gamma''}"] \arrow[dr, "\mathrm{tim}_{\gamma'}"'] & WT_{\gamma''} \arrow[r] \arrow[dl, "\mathrm{tim}_{\gamma''}"] & \cdots \\ & & KW & \end{tikzcd}$$

(cf. Fact 3.1.2).

3.5. Sheaf functors

We first introduce a functor

$$\mathrm{Dec}: \mathcal{W} \to \mathrm{Set}$$

assigning to each linearly ordered set α the set of all finite strictly decreasing

sequences in α (including the empty sequence) and to each morphism $\beta \to \alpha$ the appropriate induced mapping. Observe that Dec preserves directed colimits. For two functors F_1, $F_2 : \mathcal{W} \to \mathrm{Set}$ we shall write $F_1 \subseteq F_2$ if there is a natural transformation $\iota : F_1 \to F_2$ such that for each $\alpha \in \mathrm{Obj}\ \mathcal{W}$, ι_α is an inclusion mapping. Now assume that we are given functors $F_1 : \mathcal{W} \to \mathrm{Set}$, $F_2 : \mathcal{W} \to \mathcal{W}$ such that $F_1 \subseteq \mathrm{Dec}\, F_2$ and, moreover, for each $\alpha \in \mathrm{Obj}\ \mathcal{W}$, $F_1(\alpha) \subseteq \mathrm{Dec}\, F_2(\alpha)$ is a sheaf in $\mathrm{Obj}\ F_2(\alpha)$. Then we call F_1 an *F_2-shear functor* and write

$$F_1 \lhd \mathrm{Dec}\, F_2.$$

Now let A and B be T_γ- sheaf functors. A natural transformation $\Delta : A \to B$ is called *(γ, λ)-regular*, $(\gamma, \lambda \in \mathrm{Ord})$, if the following conditions hold:

$$\begin{aligned} &\text{If } \alpha \in \mathcal{W} \text{ and } a,\ b \in A(\alpha) \text{ with } a \preccurlyeq b \text{ then } \Delta a \preccurlyeq \Delta b. \\ &\text{Also, } (\forall x \in \overline{\Delta a})(\exists y \in \bar{a})\, x \in \mathrm{tim}_\gamma(y). \end{aligned} \tag{3.5.1}$$

$$\begin{aligned} &\text{Put, for } \alpha \in \mathrm{Obj}\ \mathcal{W} \text{ and } a \in A(\alpha), \\ &t(a) = \{a \cdot (x_0, \ldots, x_k) \mid k \geqslant 0 \text{ and } \Delta(a \cdot (x_0, \ldots, x_{k-1})) = \Delta a\}. \end{aligned} \tag{3.5.2}$$

Then, for each $\alpha \in \mathrm{Obj}\ \mathcal{W}$ and $a \in A(v)$,

$$\gamma_{t(a)} \leqslant \lambda.$$

(In particular, the left-hand side exists.)

$$\begin{aligned} &\text{Let, under the notation of (3.5.2), } a \cdot (y) \preccurlyeq c \in t(a) \text{ and} \\ &(\Delta a) \cdot (x) \preccurlyeq \Delta c. \text{ Then there is a } z \in \overline{a \cdot (y)} \text{ such that } x \in \mathrm{tim}_\gamma z. \end{aligned} \tag{3.5.3}$$

There is a natural transformation $\bar{\kappa}^\gamma : \mathrm{Dec} \to A$ such that the following diagram commutes:

$$\begin{tikzcd} A \arrow[rr, "\subseteq"] & & \overline{\mathrm{Dec}}\, T_\gamma \\ & \overline{\mathrm{Dec}} \arrow[ul, "\bar{\kappa}^\gamma"] \arrow[ur, "\overline{\mathrm{Dec}}\, \kappa^\gamma"'] & \end{tikzcd} \tag{3.5.4}$$

$$\Delta \circ \bar{\kappa}^\gamma = \mathrm{Id}. \tag{3.5.5}$$

3.6. Extension of a $(0, \lambda)$-regular transformation $\Delta_0 : A_0 \to \mathrm{Dec}$

Let $A_0 \lhd \mathrm{Dec}\ T_0$ and let

$$\Delta_0 : A_0 \to \mathrm{Dec}$$

be a $(0, \lambda)$-regular transformation. We are going to construct T_γ-sheaf functors $A_\gamma \lhd \mathrm{Dec}\ T_\gamma$ and transformations

$$\Delta_\gamma : A_\gamma \to \mathrm{Dec}$$

such that the following diagrams of functors and natural transformations may be

completed to commute:

$$\begin{tikzcd}
 & \mathrm{Dec} & \\
\Delta_{\gamma'} \nearrow & & \nwarrow \Delta_{\gamma''} \\
A_{\gamma'} \arrow[rr, dashed] \arrow[d, dash] & & A_{\gamma''} \arrow[d, dash] \\
\mathrm{Dec}\, T_{\gamma'} \arrow[rr, "\mathrm{Dec}\,\kappa^{\gamma'}_{\gamma''}"] & & \mathrm{Dec}\, T_{\gamma''},
\end{tikzcd} \tag{3.6.1}$$

where $\gamma' < \gamma'' \in \mathbf{Ord}$;

$$\begin{tikzcd}
\mathrm{Dec}\, T_{\gamma'} \arrow[d, "i^{\beta}_{\gamma'}"] & \supseteq & A_{\gamma'} \arrow[d, dashed] \arrow[ddr, dashed] \arrow[ddrr, "\Delta_{\gamma'}"] & & \\
\mathrm{Dec}\, T_{\beta} T_{\gamma'} & \supseteq & A_{\beta} T_{\gamma'} \arrow[d, "\Delta_{\beta} T_{\gamma'}"] & & \\
 & & \mathrm{Dec}\, T_{\gamma'} & \supseteq & A_{\gamma'} \arrow[r, "\Delta_{\gamma'}"'] & \mathrm{Dec}
\end{tikzcd} \tag{3.6.2}$$

where $\beta,\ \gamma' \in \mathrm{Ord}$, and $i^{\beta}_{\gamma'} = \mathrm{Dec}\, \iota^{\beta}_{\gamma'}$;

$$\begin{tikzcd}
\mathrm{Dec} \arrow[d, "\mathrm{Dec}\,\kappa^{\gamma'}"'] \arrow[ddr, dashed] \arrow[ddrr, "\mathrm{Id}"] & & & \\
\mathrm{Dec}\, T_{\gamma'} & \supseteq & A_{\gamma'} \arrow[r, "\Delta_{\gamma'}"'] & \mathrm{Dec}
\end{tikzcd} \tag{3.6.3}$$

where $\gamma' \in \mathrm{Ord}$. Before preceding further, let us show that starting with a $(0, \lambda)$-regular transformation $\Delta_0 : A_0 \to \mathrm{Dec}$, we have (3.6.1)–(3.6.3) for $\gamma' = 0$.

First note that (3.6.1) is trivial. In (3.6.3), the missing map is $\bar{\kappa}^0$ (see (3.5.4) and (3.5.5)). In (3.6.2), the vertical missing map is $A_0\kappa^0$ and the diagonal one is $\bar{\kappa}^0 \circ \Delta_0$. The upper left square is commutative since A_0 is a sheaf functor (note that $\iota^0_0 = T_0\kappa^0$). The right triangle follows from the computation $\Delta_0 \circ (\bar{\kappa}^0 \circ \Delta_0) = (\Delta_0 \circ \bar{\kappa}^0) \circ \Delta_0 = \Delta_0$ (cf. (3.5.5)). The middle triangle is the circumference of the diagram, which is commutative by (3.5.4) and the naturality of Δ_0.

$$\begin{tikzcd}
A_0 \arrow[r, "\Delta_0"] \arrow[d, "A_0\kappa^0"] & \mathrm{Dec} \arrow[r, "\bar{\kappa}^0"] \arrow[d, "\mathrm{Dec}\,\kappa^0"] & A_0 \arrow[dl] \\
A_0 T_0 \arrow[r, "\Delta_0 T_0"'] & \mathrm{Dec}\, T_0 &
\end{tikzcd}$$

3.7. The construction

$$\begin{tikzcd}
A_{\gamma+1} \arrow[dr, dashed, "\bar{\Delta}_\gamma"] \arrow[dd, phantom, "\cap"] & & \\
& A_\gamma \arrow[dr, dashed, "\Delta_\gamma"] & \\
A_\gamma T_\gamma \arrow[dr, "\Delta_\gamma T_\gamma"] \arrow[dd, phantom, "\cap\|"] & & \\
& \mathrm{Dec}\ T_\gamma \arrow[r] & \mathrm{Dec} \\
\mathrm{Dec}\ T_\gamma^2 = \mathrm{Dec}\ T_{\gamma+1} & &
\end{tikzcd} \tag{3.7.1}$$

Let Δ_γ, A_γ be already defined. Let $A_{\gamma+1}$ be the pullback of Diagram (3.7.1). Let, further, $\Delta_{\gamma+1} = \Delta_\gamma \circ \bar{\Delta}_\gamma$. Now (3.6.1) and (3.6.3) for $\gamma' = \gamma + 1$ can be easily obtained by diagram-chasing (3.7.1) and (3.6.1), (3.6.2), (3.6.3) for $\gamma' \leqslant \gamma$. For example, to get (3.6.1) with $\gamma' = \gamma$, $\gamma'' = \gamma + 1$, consider the part

$$\begin{tikzcd}
& A_\gamma \arrow[dl] \arrow[dr] & \\
A_\gamma T_\gamma \arrow[dr, "\Delta_\gamma T_\gamma"'] & & A_\gamma \arrow[dl, phantom, "\cap\|"] \\
& \mathrm{Dec}\ T_\gamma &
\end{tikzcd}$$

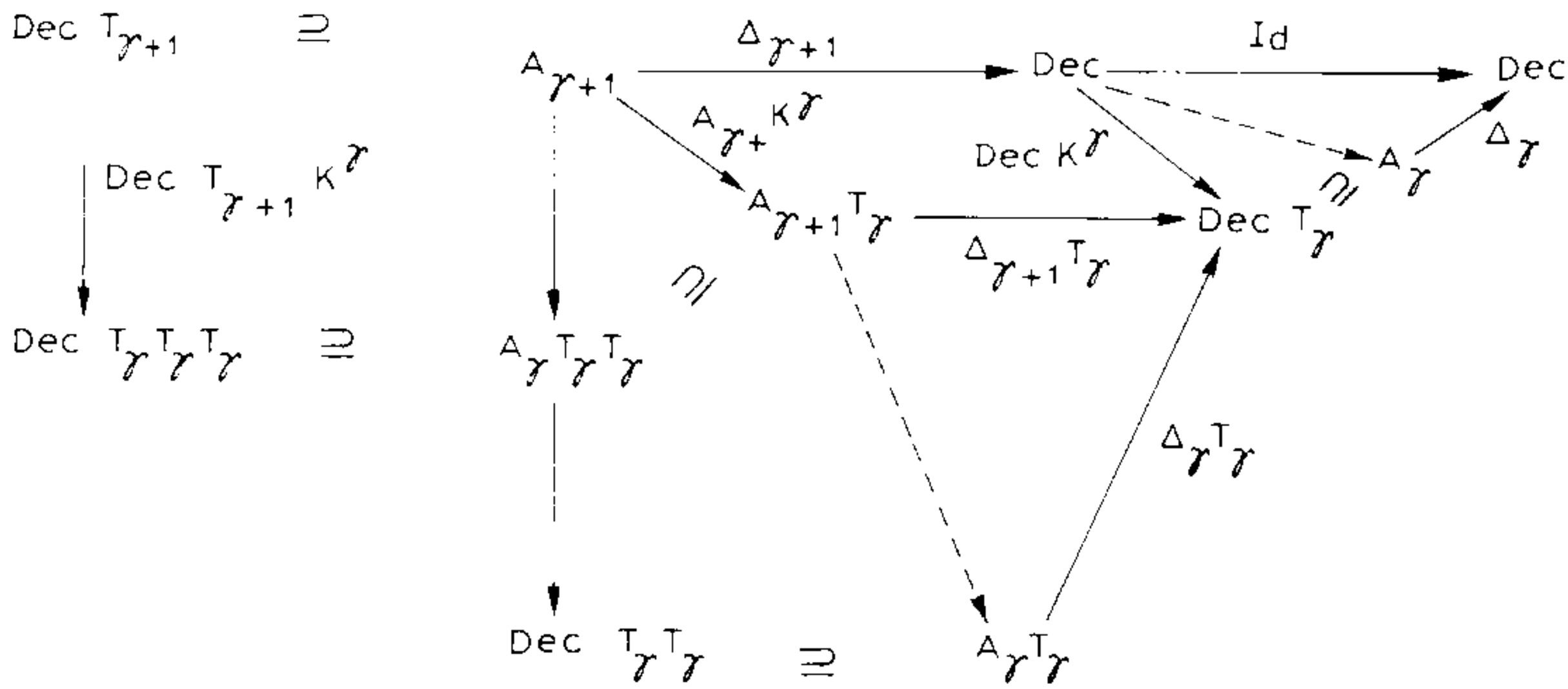
of Diagram (3.6.2) for $\beta = \gamma' = \gamma$. By the induction hypothesis, this diagram commutes. By pulling-back, we obtain a mapping $A_\gamma \to A_{\gamma+1}$ which satisfies the requisite properties (by 3.3.2 and the remaining parts of Diagram (3.6.2)).

The crucial step in the induction is to prove (3.6.2) for $\gamma' = \gamma + 1$. We distinguish the following two cases.

3.7.2 $\beta < \gamma$. By virtue of 3.3.3 and (3.6.2) with $\gamma' = \gamma$, we enjoy the following calculation (valid up to restrictions via inclusion mappings):

$$\begin{aligned}
\Delta_{\gamma+1} \mid A_{\gamma+1} &= \Delta_\gamma \circ \bar{\Delta}_\gamma = \Delta_\gamma \circ (\Delta_\gamma \cdot T_\gamma \mid A_{\gamma+1}) \\
&= \Delta_\gamma \circ ((\Delta_\gamma T_\gamma \circ (\Delta_\beta T_\gamma T_\gamma \circ i_\gamma^\beta T_\gamma) \mid A_\gamma T_\gamma) \mid A_{\gamma+1}) \\
&= \Delta_{\gamma+1} \circ \Delta_\beta T_{\gamma+1} \circ i_\gamma^\beta T_\gamma = \Delta_{\gamma+1} \circ \Delta_\beta T_{\gamma+1} \circ i_{\gamma+1}^\beta.
\end{aligned}$$

3.7.3. $\beta = \gamma$. We have a commutative diagram (see Fig. 1) which proves the



statement by 3.2.2 and 3.3.2. (The lower pentagon is the definition of $\Delta_{\gamma+1}$ composed with T_γ, the upper rhomboid is naturality of $\Delta_{\gamma+1}$ and the upper right triangle is (3.6.3) for $\gamma' = \gamma$.)

In view of (3.6.1), A_γ and Δ_γ for γ limits may be defined by passage to colimits. (3.6.1), (3.6.2) and (3.6.3) for $\gamma' = \gamma$ follow.

The Extension lemma 3.8. *Let* $A_0 \lhd \mathrm{Dec}\ T_0$ *and let* $\Delta_0 : A_0 \to \mathrm{Dec}$ *be a* $(0, \lambda)$ *regular transformation. Then* $A_\gamma \lhd \mathrm{Dec}\ T_\gamma$ *and the transformation* $\Delta_\gamma : A_\gamma \to \mathrm{Dec}$ *is* $(\gamma, \lambda^{2^\gamma})$*-regular.*

Proof. The fact that A_γ is a sheaf functor is proved by an easy induction together with (3.5.1). (3.5.4) and (3.5.5) follow from (3.6.3).

(3.5.3) follows from the definitions of Δ_γ and tim_γ by transfinite induction on γ. For example, the nonlimit step goes as follows: Assume the statement true for some $\gamma \in \mathrm{Ord}$. Now let

$$\alpha \in \mathrm{Obj}\ \mathcal{W}, \qquad a \in A_{\gamma+1}(\alpha), \qquad a \cdot (y) \preccurlyeq c \in t(a), \qquad (\Delta_{\gamma+1} a) \cdot (x) \preccurlyeq \Delta_{\gamma+1} c.$$

The last expression rewrites $(\Delta_\gamma(\Delta_\gamma a)) \cdot (x) \preccurlyeq \Delta_\gamma(\Delta_\gamma c)$. By definition, we conclude that $\Delta_\gamma(c) \in t(\Delta_\gamma(a))$ and by (3.5.1), there is a $\bar{y} \in T_\gamma(\alpha)$ such that

$$\Delta_\gamma(a) \cdot (\bar{y}) \preccurlyeq \Delta_\gamma(c). \tag{3.8.1}$$

(By abuse of notation, we usually drop the indices indicating at what object a natural transformation is applied. Here, Δ_γ stands for $(\Delta_\gamma)_{T_\gamma(\alpha)}$.)

This, by the induction hypothesis, implies that

$$(\exists \bar{z} \in \Delta_\gamma(a) \cdot (\bar{y})) x \in \mathrm{tim}_\gamma\ \bar{z}. \tag{3.8.2}$$

Also by the induction hypothesis and by (3.5.1), however, (3.8.1) implies that

$$(\exists \bar{\bar{z}} \in a \cdot (y)) \bar{z} \in \mathrm{tim}_\gamma \bar{\bar{z}}. \tag{3.8.3}$$

By the definition of $\mathrm{tim}_{\gamma+1}$, we conclude that $x \in \mathrm{tim}_{\gamma+1}\ \bar{\bar{z}}$.

We now turn to the proof of (3.5.2). We shall proceed by transfinite induction on γ. For the nonlimit step, we need to realize that $\lambda^{2^{\gamma+1}} = (\lambda^{2^\gamma})^2$ and to perform an easy deliberation essentially analogous to the one demonstrated above in order to prove (3.5.3).

The essential difficulty is in the limit step. Let, thus, γ be a limit ordinal and let for all smaller values of γ (3.5.2) hold true. Our aim is to construct a character

$$\chi : t(a) \to \lambda^{2^\gamma}.$$

For $b \in A_\gamma(\alpha)$, choose a $\gamma(b) < \gamma$ in such a way that

$$b \in (\mathrm{Dec}\, \kappa^\gamma_{\gamma(b)}) A_{\gamma(b)}(\alpha).$$

Now put

$$\chi(a \cdot (x_0, \ldots, x_k)) = \begin{cases} \lambda^{2^{\gamma(ax_0)}} & \text{when } k = 0, \\ \chi'(i_\gamma^{\gamma(ax_0)}(ax_0 \cdots x_k)) & \text{when } k > 0 \end{cases}$$

(recall that $i^\beta_\gamma = \mathrm{Dec}\, \iota^\beta_\gamma$) where

$$\chi' : t(i_\gamma^{\gamma(ax_0)}(a)) \to \lambda^{2^{\gamma(ax_0)}}$$

is the character corresponding to the (by the induction hypothesis, $(\gamma(ax_0), \lambda^{2^{\gamma(ax_0)}})$-regular) transformation $\Delta_{\gamma(ax_0)}$ applied at the object $T_\gamma(\alpha)$ of the category $\mathcal{W}$. (This means that we are considering the morphism $\Delta_{\gamma(ax_0)} : A_{\gamma(ax_0)}(T_\gamma(\alpha)) \to \mathrm{Dec}(T_\gamma(\alpha))$.)

It remains to show that χ is correctly defined and namely that

$$c \in t(a) \to i_\gamma^{\gamma(ax_0)}(c) \in t(b)$$

where $b = i_\gamma^{\gamma(ax_0)}(a)$. In the rest of the proof, we shall write γ' instead of $\gamma(ax_0)$. We will show that if, for some

$$i_\gamma^{\gamma'}(c) \in t(b) = i(i_\gamma^{\gamma'}(a)),$$

we have

$$\Delta_{\gamma'}(i_\gamma^{\gamma'}(a)) \neq \Delta_{\gamma'}(i_\gamma^{\gamma'}(c))$$

then we have

$$\Delta_\gamma(a) \neq \Delta_\gamma(c). \tag{3.8.4}$$

In effect, the choice of γ' implies

$$i_\gamma^{\gamma'}(a) \in \mathrm{Im}(\mathrm{Dec}\, T_{\gamma'} \kappa^\gamma : \mathrm{Dec}\, T_{\gamma'}(\alpha) \to \mathrm{Dec}\, T_{\gamma'} T_\gamma(\alpha)).$$

(We have $a \in (\mathrm{Dec}\, \kappa^\gamma_{\gamma'}) A_{\gamma'}(\alpha)$ since $A_{\gamma'}$ is a sheaf functor; by Diagram (3.3.4.3), we compute

$$\begin{aligned} i_\gamma^{\gamma'}(a) \in \mathrm{Im}(i_\gamma^{\gamma'} \circ \mathrm{Dec}\, \kappa^\gamma_{\gamma'}) &= \mathrm{Im}(\mathrm{Dec}(\iota_\gamma^{\gamma'} \circ \kappa^\gamma_{\gamma'})) = \mathrm{Im}(\mathrm{Dec}(T_{\gamma'} \kappa^\gamma_{\gamma'} \circ \iota_{\gamma'}^{\gamma'})) \\ &= \mathrm{Im}(\mathrm{Dec}(T_{\gamma'}(\kappa^\gamma_{\gamma'} \circ \kappa^{\gamma'}))) = \mathrm{Im}\, \mathrm{Dec}(T_{\gamma'} k^\gamma).) \end{aligned}$$

Thus, by (3.6.2) and (3.6.3), we have

$$(\Delta_{\gamma'} T_\gamma) \circ i_\gamma^{\gamma'}(a) = (\mathrm{Dec}\, \kappa^\gamma) \circ \Delta_\gamma(a). \tag{3.8.5}$$

In effect, by (3.6.3) we have

$$\Delta_\gamma \circ (\Delta_{\gamma'} T_\gamma) \circ i_\gamma^{\gamma'}(a) = \Delta_\gamma(a). \tag{3.8.6}$$

Putting $i_\gamma^{\gamma'}(a) = \mathrm{Dec}\, T_{\gamma'} \kappa^\gamma(b)$, compute

$$\begin{aligned}
(\mathrm{Dec}\, \kappa^\gamma) \circ \Delta_\gamma(a) &= (\mathrm{Dec}\, \kappa^\gamma) \circ \Delta_\gamma \circ (\Delta_{\gamma'} T_\gamma) \circ i_\gamma^{\gamma'}(a) \\
&= (\mathrm{Dec}\, \kappa^\gamma) \circ \Delta_\gamma \circ (\Delta_{\gamma'} \circ \mathrm{Dec}\, T_{\gamma'} \kappa^\gamma(b)) \quad \text{(by naturality)} \\
&= (\mathrm{Dec}\, \kappa^\gamma) \circ \Delta_\gamma \circ \mathrm{Dec}\, \kappa^\gamma \circ \Delta_{\gamma'}(b) \quad \text{(by (3.6.3))} \\
&= (\mathrm{Dec}\, \kappa^\gamma) \circ \Delta_{\gamma'}(b) \quad \text{(by naturality)} \\
&= \Delta_{\gamma'} \circ \mathrm{Dec}\, T_{\gamma'} \kappa^\gamma(b) = (\Delta_{\gamma'} T_\gamma) \circ i_\gamma^{\gamma'}(a).
\end{aligned}$$

(Again, by abuse of notation, a general transformation is occasionally identified with its specification to an object.)

Now let $\Delta_{\gamma'}(i_\gamma^{\gamma'}(a))(x) \preccurlyeq \Delta_{\gamma'}(i_\gamma^{\gamma'}(c))$. By (3.5.3), there is a $z \in a \cdot x_0$ with

$$x \in \mathrm{tim}_{\gamma' \iota_\gamma^{\gamma'}(z)}.$$

Again, by Diagram (3.3.4.3), we have $\iota_\gamma^{\gamma'}(z) \in \mathrm{Im}(T_{\gamma'} \cdot \kappa^\gamma)$. Thus, $\mathrm{tim}_{\gamma'} \iota_\gamma^{\gamma'}(z) \subseteq \mathrm{Im}\, \kappa^\gamma$ (by the naturality of $\mathrm{tim}_{\gamma'}$). Thus, for some $\bar{x} \in \alpha$,

$$x = \kappa^\gamma(\bar{x}). \tag{3.8.7}$$

Compute:

$$\begin{aligned}
&\Delta_\gamma(c) \quad \text{(by (3.8.6))} \\
&\quad = \Delta_\gamma \circ (\Delta_{\gamma'} T_\gamma) \circ i_\gamma^{\gamma'}(c) \quad \text{(by (3.5.1))} \\
&\quad \succcurlyeq \Delta_\gamma \circ (\Delta_{\gamma'}(i_\gamma^{\gamma'}(a)) \cdot (x)) \quad \text{(by (3.8.5))} \\
&\quad = \Delta_\gamma \circ (((\mathrm{Dec}\, \kappa^\gamma) \circ \Delta_\gamma(a)) \cdot (x)) \quad \text{(by (3.8.7))} \\
&\quad = \Delta_\gamma \circ (((\mathrm{Dec}\, \kappa^\gamma) \Delta_\gamma(a)) \cdot \kappa^\gamma(\bar{x})) \\
&\quad = \Delta_\gamma \circ \mathrm{Dec}\, \kappa^\gamma(\Delta_\gamma(a) \cdot \bar{x}) \quad \text{(by (3.6.3))} \\
&\quad = \Delta_\gamma(a) \cdot \bar{x}.
\end{aligned}$$

This concludes the proof of (3.8.4). □

3.9. The extension construction: A combinatorial input

In the sequel, we will be dealing with one particular set of T_γ-sheaf functors $A_\gamma \lhd \mathrm{Dec}\, T_\gamma$ and transformations $\Delta_\gamma : A_\gamma \to \mathrm{Dec}$.

Let $\alpha \in \mathrm{Obj}\, \mathcal{W}$ and $a > b \in T_0(\alpha)$. Then $a \backslash b \neq \emptyset$. Define $\delta(a, b) \in \alpha$ as

$$\begin{array}{ll}
\max(a \backslash b) & \text{if } a, b \in [\alpha]^{<\omega}, \\
a & \text{if } a \in \alpha, \\
\max a & \text{if } a \in [\alpha]^{<\omega} \text{ and } b \in \alpha.
\end{array}$$

The following observations for $a > b > c \in T_0(\alpha)$ are in order:

$$\delta(a, b) = \delta(b, c) \to a \in [\alpha]^{<\omega} \quad \text{and} \quad b = \max a, \tag{3.9.1}$$

$$\delta(a, c) = \max(\delta(a, b), \delta(b, c)), \tag{3.9.2}$$

$$b \in \alpha \to \delta(a, b) \geqslant b, \tag{3.9.3}$$

$$a \in \alpha \to \delta(b, c) < a. \tag{3.9.4}$$

For $a = (a_0, \ldots, a_n) \in \operatorname{Dec} T_0(\alpha)$, put $\delta^i = \delta(a_i, a_{i+1})$ and

$$\bar{\delta}(a) = \begin{cases} (\delta^0, \ldots, \delta^{n-1}) & \text{if } a_n \in [\alpha]^{<\omega}, \\ (\delta^0, \ldots, \delta^{n-1}, a_n) & \text{if } a_n \in \alpha. \end{cases}$$

Put, for $\alpha \in \operatorname{Obj} \mathcal{W}$

$$A_0(\alpha) = \{a \in \operatorname{Dec} T_0(\alpha) \mid \bar{\delta}(a) \in \operatorname{Dec}(\alpha)\}, \qquad (\Delta_0)_\alpha = \bar{\delta} \mid A_0(\alpha).$$

To make A_0 a functor, we define for $\varphi \in \mathcal{W}(\alpha, \beta)$

$$A_0(\varphi)((a_0, \ldots, a_n)) = (\varphi(a_0), \ldots, \varphi(a_n)).$$

Fact 3.9.5. *We have* $A_0 \lhd \operatorname{Dec} T_0$. *Moreover,* $\Delta_0 : A_0 \to \operatorname{Dec}$ *is a* $(0, 2)$*-regular transformation.*

Proof. We easily verify (3.5.1) and the first sentence of the statement follows. Also (3.5.2) is obvious. To see (3.5.3), note that $\delta(a, b) \in \operatorname{tim}_0(a)$. Concerning (3.5.4) and (3.5.5), we observe that, for $a = (a_0, \ldots, a_n) \in \operatorname{Dec} \alpha \subseteq \operatorname{Dec} T_0(\alpha)$, $\Delta_0(a) = a$. $\square$

At this point, we consider the sheaf functors $A_\gamma \lhd \operatorname{Dec} T_\gamma$ and the $(\gamma, 2^{2^\gamma})$-regular transformations $\Delta_\gamma : A_\gamma \to \operatorname{Dec}$ defined in 3.6 and 3.7.

3.10. The Transformations $\bar{\delta}_\gamma$, $\gamma \in \mathbf{Ord}$

Let $\operatorname{Seq} : \mathcal{W} \to \operatorname{Set}$ be the functor assigning to $\alpha \in \operatorname{Obj}(\mathcal{W})$ the set of all finite sequences in α and to $\varphi \in \operatorname{Mor}(\mathcal{W})$ the corresponding induced mapping. From 3.9, we have a natural transformation

$$\bar{\delta} : \operatorname{Dec} T_0 \to \operatorname{Seq}.$$

We are now going to define natural transformations

$$\bar{\delta}_\gamma : \operatorname{Dec} T_\gamma \to \operatorname{Seq} T_\gamma$$

in the following manner.

3.10.1. $\bar{\delta}_0 = (\operatorname{Seq} \cdot \kappa^0) \circ \bar{\delta}$.

3.10.2. Let $\bar{\delta}_\gamma$ be already defined. Let $\alpha \in \operatorname{Obj} \mathcal{W}$ and let $a \in \operatorname{Dec} T_{\gamma+1}(\alpha)$.

We define $\bar{\delta}_{\gamma+1}(a)$ to be

$$(\bar{\delta}_\gamma \cdot T_\gamma)(a) \quad \text{if } (\bar{\delta}_\gamma)_{\mathrm{Dec}\, T_\gamma(\alpha)}(a) \notin \mathrm{Im}(\mathrm{Dec}(\kappa^\gamma)_{T_\gamma(\alpha)})$$

and

$$(\mathrm{Seq}\, \kappa_\gamma^{\gamma+1})\bar{\delta}_\gamma(b) \quad \text{if } (\bar{\delta}_\gamma)_{\mathrm{Dec}\, T_\gamma(\alpha)}(a) = \mathrm{Dec}(\kappa^\gamma)_{T_\gamma(\alpha)}(b).$$

3.10.3. As usual, for γ limit we would like to define $\bar{\delta}_\gamma$ by passage to colimits. As usual, this requires a commutative diagram of the form:

$$\begin{array}{ccc} \mathrm{Dec}\, T_\gamma & \xrightarrow{\bar{\delta}_\gamma} & \mathrm{Seq}\, T_\gamma \\ \downarrow{\scriptstyle \mathrm{Dec}\, \kappa_\gamma^{\gamma+1}} & & \downarrow{\scriptstyle \mathrm{Seq}\, \kappa_\gamma^{\gamma+1}} \\ \mathrm{Dec}\, T_{\gamma+1} & \xrightarrow{\bar{\delta}_{\gamma+1}} & \mathrm{Seq}\, T_{\gamma+1}. \end{array} \tag{3.10.3.1}$$

However, using naturality of $\bar{\delta}_\gamma$ and the fact that $\kappa_\gamma^{\gamma+1} = T_\gamma \kappa^\gamma$, we only get a commutative diagram:

$$\begin{array}{ccc} \mathrm{Dec}\, T_\gamma & \xrightarrow{\bar{\delta}_\gamma} & \mathrm{Seq}\, T_\gamma \\ \downarrow{\scriptstyle \overline{\mathrm{Dec}}\, \kappa_\gamma^{\gamma+1}} & & \downarrow{\scriptstyle \overline{\mathrm{Seq}}\, \kappa_\gamma^{\gamma+1}} \\ \mathrm{Dec}\, T_{\gamma+1} & \xrightarrow{\bar{\delta}_\gamma T_\gamma} & \mathrm{Seq}\, T_{\gamma+1}. \end{array} \tag{3.10.3.2}$$

To obtain (3.10.3.1), we need to show that, for $a \in \mathrm{Dec}\, T_\gamma(\alpha)$,

$$(\bar{\delta}_\gamma)_{T_\gamma(\alpha)} \circ \mathrm{Dec}\, \kappa_\gamma^{\gamma+1}(a) = \mathrm{Dec}(\kappa^\gamma)_{T_\gamma(\alpha)}(b) \rightarrow (\bar{\delta}_\gamma(b) = \bar{\delta}_\gamma(a)). \tag{3.10.3.3}$$

This cannot be reached without some deeper insight into what the induction has done as yet. Comparing 3.10.2 with 3.7, we obtain the following key results.

$$(\bar{\delta}_\gamma(a) \in \mathrm{Dec}\, T_\gamma(\alpha)) \rightarrow (a \in A_\gamma(\alpha)) \text{ and } \bar{\delta}_\gamma(a) = (\mathrm{Dec}\, \kappa^\gamma)\Delta_\gamma(a)), \tag{3.10.4}$$

$$\bar{\delta}_\gamma \circ (\mathrm{Dec}\, \kappa^\gamma) = (\mathrm{seq}\, \kappa^\gamma) \circ (\subseteq : \mathrm{Dec} \rightarrow \mathrm{Seq}), \tag{3.10.5}$$

(this is a consequence of (3.10.4) and (3.6.3)) and

$$a \subseteq b \rightarrow \bar{\delta}_\gamma(a) \subseteq \bar{\delta}_\gamma(b) \tag{3.10.6}$$

(this follows from (3.9.2) by transfinite induction).

Now to prove (3.10.3.3), note that by the very fact that $(\kappa^\gamma)_{T_\gamma(\alpha)}$ is a $\mathcal{W}$-morphism, we have

$$\mathrm{Im}(\mathrm{Dec}(\kappa^\gamma)_{T_\gamma}(a)) \subseteq \mathrm{Dec}\, T_{\gamma+1}.$$

Thus, the precise of (3.10.3.3) implies

$$(\bar{\delta}_\gamma)_{T_\gamma(\alpha)}(\mathrm{Dec}\, \kappa_\gamma^{\gamma+1}(a)) \in \mathrm{Dec}\, T_{\gamma+1}$$

and hence, by (3.10.4) and the naturality of Δ_γ,

$$\begin{aligned}(\bar{\delta}_\gamma)_{T_\gamma(\alpha)}(\mathrm{Dec}\,\kappa_\gamma^{\gamma+1}(a)) &= (\mathrm{Dec}(\kappa^\gamma)_{T_\gamma(\alpha)})(\Delta_\gamma)_{T_\gamma(\alpha)}(\mathrm{Dec}\,\kappa_\gamma^{\gamma+1}(a))\\ &= (\mathrm{Dec}(\kappa^\gamma))_{T_\gamma(\alpha)})(\Delta_\gamma)_{T_\gamma(\alpha)}(\mathrm{Dec}\,T_\gamma\kappa^\gamma(a)) \qquad (3.10.7)\\ &= (\mathrm{Dec}(\kappa^\gamma)_{T_{\gamma(\alpha)}}(\mathrm{Dec}\,\kappa^\gamma)\Delta_\gamma(a) = (\mathrm{Dec}\,\kappa^{\gamma+1})\Delta_\gamma(a)\\ &= \mathrm{Dec}(\kappa^\gamma)_{T_\gamma(\alpha)}\circ(\mathrm{Dec}\,\kappa^\gamma)\Delta_\gamma(a).\end{aligned}$$

Since $\mathrm{Dec}(\kappa^\gamma)_\beta$ (as well as $(\kappa^\gamma)_\beta$) is obviously injective for any $\beta \in \mathrm{Obj}\,\mathcal{W}$ (by the mere fact that $(\kappa^\gamma)_\beta$ is a $\mathcal{W}$-morphism), we conclude that

$$b = (\mathrm{Dec}\,\kappa^\gamma)\Delta_\gamma(a) = \bar{\delta}_\gamma(a)$$

(the last equality following from (3.10.4)). Now (3.10.3.3) follows from (3.10.5). The induction is complete.

3.11. A partition

In the sequel, we shall identify $\mathrm{Dec}\,\beta$ with $[\beta]^{<\omega}$ via the obvious bijection $(a_0, \ldots, a_n) \mapsto \{a_0, \ldots, a_n\}$.
Define

$$\begin{aligned}Q_\gamma^1(\alpha) &= \{(q_1, q_2, q_3, q_4) \in \mathrm{Seq}\,T_\gamma(\alpha) \mid q_1 < q_2 \leqslant q_3 > q_4\},\\ Q_\gamma^2(\alpha) &= \{(q_1, q_2, q_3, q_4) \in \mathrm{Seq}\,T_\gamma(\alpha) \mid q_1 < q_2 > q_3 < q_4\},\\ Q_\gamma^3(\alpha) &= \{(q_1, q_2, q_3, q_4) \in \mathrm{Seq}\,T_\gamma(\alpha) \mid q_1 < q_2 > q_3 \geqslant q_4\},\\ Q_\gamma^4(\alpha) &= \{(q_1, q_2, q_3, q_4) \in \mathrm{Seq}\,T_\gamma(\alpha) \mid q_1 > q_2 < q_3 > q_4\},\\ Q_\gamma^5(\alpha) &= \{(q_1, q_2, q_3, q_4) \in \mathrm{Seq}\,T_\gamma(\alpha) \mid q_1 > q_2 < q_3 \leqslant q_4\},\end{aligned}$$

$$Q_\gamma(\alpha) = \bigcup_{i=1}^{5} Q_\gamma^i(\alpha), \qquad \bar{S}_\gamma(\alpha) = \{a \in \mathrm{Dec}\,T_\gamma(\alpha) \mid \bar{\delta}_\gamma(a) \in Q_\gamma(\alpha)\}$$

Observe that $i \neq j \to Q_\gamma^i(\alpha) \cap Q_\gamma^j(\alpha) = \emptyset$.
Next define a mapping $\bar{r}_\gamma : \bar{S}_\gamma(\alpha) \to \{0, 1\}$ by

$$\bar{r}_\gamma(a) = \begin{cases} 0 & \text{if } \bar{\delta}_\gamma(a) \in Q_\gamma^i \text{ with } i \leqslant 3,\\ 1 & \text{if } \bar{\delta}_\gamma(a) \in Q_\gamma^i \text{ with } i > 3,\end{cases}$$

It is quite clear that $\bar{S}_\gamma(\alpha)$ is usually not a Sperner system. On the other hand, by (3.10.6) we have

$$(a \subseteq b \quad \text{and} \quad a, b \in \bar{S}_\gamma(\alpha)) \to \bar{r}_\gamma(a) = \bar{r}_\gamma(b). \qquad (3.11.1)$$

We shall put

$$S_\gamma(\alpha) = \{a \in \bar{S}_\gamma(\alpha) \mid (\forall b \subsetneqq a) b \notin \bar{S}_\gamma(\alpha)\}, \qquad r_\gamma = \bar{r}_\gamma \mid S_\gamma(\alpha).$$

Now $S_\gamma(\alpha)$ is a Sperner system.

3.12. Some combinatorial properties of $\bar{\delta}$

Choose an $a = (a_0, \ldots, a_l) \in \operatorname{Dec} T_0(\alpha)$ and let $\bar{\delta}(a) = (q_0, \ldots, q_k)$ (it can be $k = l$ or $k = l - 1$). We call an $i \in \{0, \ldots, k\}$ *separable* (with respect to a) if one of the following conditions holds:

$$q_i = a_i \quad \text{(in particular, } a_i \in \alpha), \tag{3.12.1}$$

$$i < l \text{ and } a_{i+1} \in [\alpha]^{<\omega}. \tag{3.12.2}$$

We shall need the following facts.

3.12.3. *Let* $0 \leqslant i_1 < i_2 < i_3 < i_4 \leqslant k$. *Assume, further, that* i_4 *is separable. Put* $q = (q_i, q_{i_2}, q_{i_3}, q_{i_4})$. *Then we have:*

$$q \in Q_0^2(\alpha) \cup Q_0^3(\alpha) \to (\exists b \subseteq a)\bar{r}_0(b) = 0, \tag{3.12.3.1}$$

$$q \in Q_0^4(\alpha) \cup Q_0^5(\alpha) \to (\exists b \subseteq a)\bar{r}_0(b) = 1. \tag{3.12.3.2}$$

Proof. (3.9.2) allows us to construct b explicitly. The following eight cases need to be distinguished.

$$q \in Q_0^2(\alpha) \quad \text{and} \quad a_{i_4} \in [\alpha]^{<\omega}. \tag{3.12.3.3}$$

Then put $b = (a_{i_1}, a_{i_1+1}, a_{i_3}, a_{i_3+1}, a_{i_4+1})$, we have $\bar{\delta}(b) \in Q_0^2(\alpha)$.

$$q \in Q_0^2(\alpha) \quad \text{and} \quad a_{i_4} \in \alpha. \tag{3.12.3.4}$$

Since $a \in \operatorname{Dec} T_0(\alpha)$. Since $a \in \operatorname{Dec} T_0(\alpha)$, we must have $i_4 - 1 > i_3$ and $q_{i_4-1} > q_{i_4}$. Putting $b = (a_{i_1}, a_{i_1+1}, a_{i_4-1}, a_{i_4})$, we have $\bar{\delta}(b) \in Q_0^1(\alpha) \cup Q_0^3(\alpha)$.

$$q \in Q_0^3(\alpha) \quad \text{and} \quad a_{i_4} \in [\alpha]^{<\omega}. \tag{3.12.3.5}$$

Putting $b = (a_{i_1}, a_{i_1+1}, a_{i_3}, a_{i_3+1}, a_{i_4+1})$, we get $\bar{\delta}(b) \in Q_0^2(\alpha) \cup Q_0^3(\alpha)$.

$$q \in Q_0^3(\alpha) \quad \text{and} \quad a_{i_4} \in \alpha. \tag{3.12.3.6}$$

Putting $b = (a_{i_1}, a_{i_1+1}, a_{i_3}, a_{i_4})$, we get $\bar{\delta}(b) \in Q_0^1(\alpha) \cup Q_0^3(\alpha)$.

$$q \in Q_0^4(\alpha) \quad \text{and} \quad a_{i_4} \in [\alpha]^{<\omega}. \tag{3.12.3.7}$$

Putting $b = (a_{i_1}, a_{i_2}, a_{i_2+1}, a_{i_4}, a_{i_4+1})$, we get $\bar{\delta}(b) \in Q_0^4(\alpha)$.

$$q \in Q_0^4(\alpha) \quad \text{and} \quad a_{i_4} \in \alpha. \tag{3.12.3.8}$$

Putting $b = (a_{i_1}, a_{i_2}, a_{i_2+1}, a_{i_4})$, we get $\bar{\delta}(b) \in Q_0^4(\alpha)$.

$$q \in Q_0^5(\alpha) \quad \text{and} \quad a_{i_4} \in [\alpha]^{<\omega}. \tag{3.12.3.9}$$

Putting $b = (a_{i_1}, a_{i_2}, a_{i_2+1}, a_{i_4}, a_{i_4+1})$, we get $\bar{\delta}(b) \in Q_0^4(\alpha) \cup a_0^5(\alpha)$.

$$q \in Q_0^5(\alpha) \quad \text{and} \quad a_{i_4} \in \alpha. \tag{3.12.3.10}$$

Putting $b = (a_{i_1}, a_{i_2}, a_{i_2+1}, a_{i_4})$, we get $\bar{\delta}(b) \in Q_0^4(\alpha)$. □

3.12.4. *Let* $0 \leqslant i_0 < i_1 < \cdots < i_s \leqslant k$ *be all the separable indices. We have*

$$i_0 \leqslant 1 \quad \text{and} \quad i_{t+1} \leqslant i_t + 2. \tag{3.12.4.1}$$

Moreover, we have

$$(i_{t+1} = i_t + 2) \rightarrow (a_{i_{t+1}} \in [\alpha]^{<\omega} \quad \text{and} \quad a_{i_{t+1}} \in \alpha) \tag{3.12.4.2}$$

Proof: This is obvious. □

3.12.5. *Let* $0 \leqslant t < s$ *and let* $q_{i_t} \geqslant q_{i_t+1} \geqslant q_{i_{t+1}}$. *Then we have* $q_{i_t} > q_{i_{t+1}}$.

Proof. We distinguish the following two cases.

$$i_{t+1} = i_t + 1. \tag{3.12.5.1}$$

Then we have $a_{i_t+1}, a_{i_t+2} \in [\alpha]^{<\omega}$. Apply (3.9.1).

$$i_{t+1} = i_t + 2. \tag{3.12.5.2}$$

Then we have $a_{i_t+2} \in \alpha$, $a_{i_t+1} \in [\alpha]^{<\omega}$ (see 3.12.4.2). By (3.9.1), again, we have $q_{i_t} > q_{i_t+1}$. □

3.12.6. *Let* $0 \leqslant p < t \leqslant l$ *and let* $q_p \leqslant q_{p+1} \leqslant \cdots \leqslant q_t$. *Then*

$$\{a_p, a_{p+1}, \ldots, a_{t-1}\} \subseteq [\alpha]^{<\omega}. \tag{3.12.6.1}$$

$$q_p < q_{p+1} < \cdots < q_{t-1}, \tag{3.12.6.2}$$

$$\{q_p, q_{p+1}, \ldots, q_t\} \subseteq \mathrm{tim}_0(a_p). \tag{3.12.6.3}$$

Proof. (3.12.6.1) follows from the fact that $a \in \mathrm{Dec}\, T_0(\alpha)$. (3.12.6.2) is a consequence of (3.9.1). (3.12.6.3) follows from (3.9.2): For $p \leqslant i \leqslant t-1$, $q_i = \delta(a_i, a_{i+1}) = \delta(a_p, a_{i+1})$. For $i = t$ we either have $q_t = q_{t-1}$ or we can argue similarly. □

3.13. The sheaf functors $A^\epsilon_\gamma (\epsilon \in \{0, 1\})$ and $(\gamma, \omega^{2^\gamma})$-regular transformations Δ^ϵ_γ

3.13.1. Put

$$A^0_0(\alpha) = \{a \in \mathrm{Dec}\, T_0(\alpha) \mid \bar{\delta}(a) = (q_0, \ldots, q_k) \text{ and } q_0 \leqslant \cdots \leqslant q_i > q_{i+1} \geqslant \cdots \geqslant q_{k-2} \quad \text{for some } 0 \leqslant i \leqslant k-2\},$$

$$A^1_0(\alpha) = \{a \in \mathrm{Dec}\, T_0(\alpha) \mid \bar{\delta}(a) = (q_0, \ldots, q_k) \text{ and } q_0 \geqslant \cdots \geqslant q_i < q_{i+1} \leqslant \cdots \leqslant q_{k-2} \quad \text{for some } 0 \leqslant i \leqslant k-2\}.$$

No doubt, by taking the proper action on mappings, A^ϵ_0 become sheaf functors.

3.13.2. The transformation $\Delta^0_0 : A^0_0 \rightarrow \mathrm{Dec}$

For a given $a \in A^0_0(\alpha)$, choose the minimal $i \in \{0, \ldots, k-2\}$ such that $i = k-2$ or $q_i > q_{i+1}$. Let indices $i_1 < \cdots < i_s$ be defined in the following way.

$i_1 = i.$ (3.13.2.1)

For $1 \leqslant t < s$, $i_{t+1} \in \{i_t + 1, i_t + 2, \ldots, k\}$ is such that
$q_{i_t} = q_{i_t+1} = \cdots = q_{i_{t+1}-1} > q_{i_{t+1}}$. (3.13.2.2)

No i_{s+1} satisfying the condition of (3.13.2.2) for $t = s$ exists. (3.13.2.3)

Now define $\Delta_0^0(a)$ as

$$(q_{i_1}, \ldots, q_{i_s}) \quad \text{if } q_{i_s} \in \bar{a}$$

and

$$(q_{i_1}, \ldots, q_{i_s-1}) \quad \text{else.}$$

Lemma 3.13.3. *Δ_0^0 is $(0, \omega)$-regular.*

Proof. (3.5.1) is obvious. (3.5.2) follows from (3.12.4), (3.12.5) and (3.12.6). To prove (3.5.3), we have to distinguish the following three possibilities.

$\Delta_0^0 a = \emptyset.$ (3.13.3.1)

Then our statement follows from 3.12.6.

$\Delta_0^0 a \neq \emptyset$ and $q_{i_s} \notin a.$ (3.13.3.2)

Then the statement follows from $q_{i_s} \in \mathrm{tim}_0\, a_{i_s}$.

$\Delta_0^0 a \neq \emptyset$ and $q_{i_s} \in a.$ (3.13.3.3)

Extending the notation to c compatibly, we have by (3.9.1) $q_{i_{s+1}} = q_{i_s+1} \in \mathrm{tim}_0(c_{i_{s+1}})$. (3.5.4) and (3.5.5) are obvious. □

3.13.4. The transformation $\Delta_0^1 : A_0^1 \to$ Dec

Let $a \in A_0^1(\alpha)$. Let indices $i_1 < \cdots < i_s$ be defined in the following way.

$i_1 = 0$ (3.13.4.1)

For $1 \leqslant t < s$, $i_{t+1} \in \{i_t + 1, i_t + 2, \ldots, k\}$ is such that
$q_{i_t} = q_{i_t+1} = \cdots = q_{i_{t+1}-1} > q_{i_{t+1}}$. (3.13.4.2)

No i_{s+1} satisfying the condition of (3.13.4.2) for $t = s$ exists. (3.13.4.3)

Now define $\Delta_0^1(a)$ as

$$(q_{i_1}, \ldots, q_{i_s}) \quad \text{if } q_{i_s} \in \bar{a}$$

and

$$(q_{i_1}, \ldots, q_{i_s-1}) \quad \text{else.}$$

Lemma 3.13.5. *Δ_0^1 is $(0, \omega)$-regular.*

Proof. Follows in a similar way as the proof of 3.13.3. Moreover, observe that a case analogous to (3.13.3.1) does not occur. □

3.13.5. Now consider the sheaf functors $A_\gamma^\epsilon \lhd \mathrm{Dec}\, T_\gamma (\epsilon = 0, 1)$ and $(\gamma, \omega^{2^\gamma})$-regular transformations $\Delta_\gamma^\epsilon : A_\gamma^\epsilon \to \mathrm{Dec}$ constructed by the machinery of 3.6 and 3.7. Observe that by virtue of (3.9.2) and a transfinite induction, we have the following extra property.

Lemma 3.13.6. $(\forall a \in A_\gamma^\epsilon(\alpha))(\forall c \subseteq \Delta_\gamma^\epsilon(a)(\exists b \subseteq a)(b \in A_\gamma(\alpha)$ *and* $\Delta_\gamma(b) = c)$.

We shall next return to the circumstances of Nash-Williams' partition theorem. We shall observe the identification imposed at the beginning of 3.11. Thus, for $\epsilon \in \{0, 1\}$, we have sheaves

$$A_\epsilon^{S_\gamma, r_\gamma} \subseteq \mathrm{Seq}\ T_\gamma(\alpha).$$

Put $\bar{A}_\epsilon^{S_\gamma, r_\gamma}(\alpha) = A_\epsilon^{S_\gamma, r_\gamma} \cap \mathrm{Dec}\ T_\gamma(\alpha)$. Clearly, by taking the appropriate action on mappings, $\bar{A}_\epsilon^{S_\gamma, r_\gamma}$ turns into a sheaf functor.

The inclusion lemma 3.14. *We have*

$$\bar{A}_\epsilon^{S_\gamma, r_\gamma} \subseteq A_\gamma^\epsilon.$$

Proof. We shall proceed by a transfinite induction on γ. First we shall handle the case $\gamma = 0$. Let $a \notin A_0^0(\alpha)$. As we can easily check, then there are $0 \leqslant i_1 < i_2 < i_3 \leqslant k-2$ such that

$$q_{i_1} > q_{i_2} < q_{i_3}. \tag{3.14.1}$$

By 3.12.4, there is an $i_4 \in \{k-1, k\}$ such that q_{i_4} is separable. By (3.14.1), $(q_{i_1}, q_{i_2}, q_{i_3}, q_{i_4}) \in Q_0^4(\alpha) \cup Q_0^5(\alpha)$. Hence, by 3.12.3, there is a $b \subseteq a$ with $\bar{r}_0(b) = 1$.

Now let $a \notin A_0^1(\alpha)$. As we may easily check, then there are $0 \leqslant i_1 < i_2 < i_3 \leqslant k-2$ such that

$$q_{i_1} < q_{i_2} > q_{i_3}. \tag{3.14.2}$$

By 3.12.4, there is an $i_4 \in \{k-1, k\}$ such that q_{i_4} is separable. By (3.14.2), $(q_{i_1}, q_{i_2}, q_{i_3}, q_{i_4}) \in Q_0^2(\alpha) \cup Q_0^3(\alpha)$. Hence, by 3.12.3, there is a $b \subseteq a$ with $\bar{r}_0(b) = 0$.

Next we remark that the limit case is trivial, since both sides are obtained by passing to colimits. Thus, the nonlimit step remains.

Let the statement be true for a certain $\gamma \in \mathrm{Ord}$. By definition, we have

$$\bar{A}_\epsilon^{S_{\gamma+1}, r_{\gamma+1}}(\alpha) \subseteq \bar{A}_\epsilon^{S_\gamma, r_\gamma}(T_\gamma(\alpha)).$$

Thus, by the induction hypothesis,

$$\bar{A}_\epsilon^{S_{\gamma+1}, r_{\gamma+1}}(\alpha) \subseteq A_\gamma^\epsilon(T_\gamma(\alpha)).$$

Now we shall prove that for an $a \in \bar{A}_\epsilon^{S_{\gamma+1}, r_{\gamma+1}}(\alpha)$ we have

$$\Delta_\gamma^\epsilon(a) \in \bar{A}_\epsilon^{S_\gamma, r_\gamma}(\alpha). \tag{3.14.3}$$

In effect, let, say, $\epsilon = 0$ (the case of $\epsilon = 1$ is handled similarly). Should (3.14.3) be false, then there is a $c \subseteq \Delta^0_\gamma(a)$ with

$$\bar{\delta}_\gamma(c) \in Q^4_\gamma(\alpha) \cup Q^5_\gamma(\alpha).$$

By 3.13.6, there is a $b \subseteq a$ with

$$\Delta_\gamma(b) = c.$$

By (3.10.4), we obtain

$$\bar{\delta}_\gamma(b) = (\mathrm{Dec}\,\kappa^\gamma)_{T_\gamma(\alpha)}(c).$$

Now by definition 3.10.2 (the second part), we get

$$\begin{aligned}\bar{\delta}_{\gamma+1}(b) &= (\mathrm{Seq}\,\kappa^{\gamma+1}_\gamma)(\bar{\delta}_\gamma(c)) \in (\mathrm{Seq}\,\kappa^{\gamma+1}_\gamma)(Q^4_\gamma(\alpha) \cup Q^5_\gamma(\alpha)) \\ &\subseteq Q^4_{\gamma+1}(\alpha) \cup Q^5_{\gamma+1}(\alpha).\end{aligned}$$

We conclude $\bar{r}_{\gamma+1}(b) = 1$ which is a contradiction.

Thus, (3.14.3) is proved. By the induction hypothesis, we conclude that

$$\Delta^\epsilon_\gamma(a) \in A^\epsilon_\gamma(\alpha),$$

implying $a \in A^\epsilon_{\gamma+1}(\alpha)$ directly by the construction 3.7. □

3.15. Some computations

As yet, we did not assume that the argument α of $A^\epsilon_\gamma(\alpha)$, $T_\gamma(\alpha)$, $\bar{A}^{S_\gamma, r_\gamma}_\epsilon(\alpha)$ etc. would be an ordinal. Indeed, it was not even possible to restrict ourselves to working with ordinals. Our construction was based on the categorial behaviour of $\mathcal{W}$, which would substantially change by restricting to ordinals: For example, we would lose directed colimits.

Yet, the case of α being an ordinal is the most interesting one. We shall consider it throughout the rest of the paper. As before, we shall first make the assumption

$$\Omega \supseteq T_\gamma(\alpha). \tag{3.15.1}$$

Note that, by (3.5.2), for any (γ, λ)-regular transformation $\Delta : A \to \mathrm{Dec}$, we have ($A$ being considered a tree by the relation $\preccurlyeq$)

$$\gamma_{A(\alpha)} \leqslant (\alpha + 1) \cdot \lambda + 1. \tag{3.15.2}$$

Indeed, let $\chi : \mathrm{Dec}(\alpha) \to \alpha + 1$ be a character (the summand 1 comes from the empty sequence). Let, for each $b \in A(\alpha)$, $\chi_b : t(b) \to \lambda$ denote the character whose existence is stated by (3.5.2). Let, further, for $a \in A(\alpha)$,

$$b(a) \in A(\alpha)$$

denote the sequence satisfying

$$b(a) \preccurlyeq a, \qquad \Delta b(a) = \Delta a, \qquad c < b(a) \to \Delta c \neq \Delta a.$$

Now define a character $\bar{\chi}: A(\alpha) \to (\alpha+1)\cdot\lambda + 1$ by

$$\bar{\chi}(a) = \begin{cases} (\chi(\Delta a)\cdot\lambda) + \chi_{b(a)}(a) & \text{if } b(a) < a, \\ (\chi(\Delta a)+1)\cdot\lambda & \text{if } b(a) = a. \end{cases}$$

(3.15.2) is proved.

Now by (3.13) and (3.15.2) and $\emptyset \notin A^{\epsilon}_{\gamma}(\alpha)$ we conclude

$$\gamma_{A^{\epsilon}_{\gamma}(\alpha)} \leqslant (\alpha+1)\cdot\omega^{2^{\gamma}}, \quad (\epsilon = 0, 1), \tag{3.15.3}$$

and, by 3.14,

$$\gamma_{A^{S_{\gamma},r_{\gamma}}_{\epsilon}\cap \mathrm{Dec}\, T_{\gamma}(\alpha)} \leqslant (\alpha+1)\cdot\omega^{2^{\gamma}}, \quad (\epsilon = 0, 1). \tag{3.15.4}$$

Thus, there is an $((\alpha+1)\cdot\omega^{2^{\gamma}}, (\alpha+1)\cdot\omega^{2^{\gamma}})$-testing $g = (g_0, g_1)$ (recall that g_{ϵ} is a character on $A^{S_{\gamma},r_{\gamma}} \cap \mathrm{Dec}\, T_{\gamma}(\alpha)$) such that

$$\mathrm{Dec}\, T_{\gamma}(\alpha) \subseteq \mathrm{Bad}(A^{S_{\gamma},r_{\gamma}}_{\epsilon}, g) \cup \{\emptyset\}, \tag{3.15.5}$$

which is (3.0.4). From this, 1.4.2.1 and 1.3 we conclude that

$$\gamma_{\mathrm{Dec}\, T_{\gamma}(\alpha)} \leqslant \psi((\alpha+1)\cdot\omega^{2^{\gamma}}, (\alpha+1)\cdot\omega^{2^{\gamma}}). \tag{3.15.6}$$

Now we show that T_{γ} satisfies (3.0.1)–(3.0.3). Condition (3.0.1) follows directly from the definition of $T_0(\alpha)$. To prove (3.0.2), we first define a natural transformation

$$\lambda^{n}_{\beta}: T_{\beta}T_{\beta+n} \to T_{\beta+n+1}$$

for $\beta \in \mathrm{Ord}$, $n \in \omega$ by induction on n. For $n = 0$ let $\lambda^{0}_{\beta}: T_{\beta}T_{\beta} \to T_{\beta+1}$ be the identity. Is λ^{n}_{β} already defined, define λ^{n+1}_{β} as the composition

$$\begin{array}{ccc} T_{\beta}T_{\beta+n+1} = T_{\beta}T_{\beta+n}T_{\beta+n} & \xrightarrow{\lambda^{n}_{\beta}T_{\beta+n}} & T_{\beta+n+1}T_{\beta+n} \\ & & \Big\downarrow {\scriptstyle T_{\beta+n+1}\kappa^{\beta+n+1}_{\beta+n}} \\ T_{\beta+n+2} & = & (T_{\beta+n+1})^{2}. \end{array}$$

For $\gamma = \beta + n$, the naturality of λ^{n}_{β} yields a commutative diagram

$$\begin{array}{ccc} T_{\beta}T_{\gamma} & \xrightarrow{\lambda^{n}_{\beta}} & T_{\gamma+1} \\ {\scriptstyle T_{\beta}T_{\gamma}\kappa^{\gamma}}\Big\downarrow & & \Big\downarrow {\scriptstyle T_{\gamma+1}\kappa^{\gamma}} \\ T_{\beta}T_{\gamma+1} & \xrightarrow{\lambda^{n}_{\beta}T_{\gamma}} & T_{\gamma+1}T_{\gamma} \\ & & \Big\downarrow {\scriptstyle T_{\gamma+1}\kappa^{\gamma+1}_{\gamma}} \\ & & T_{\gamma+2} \end{array}$$

which is readily rephrased into

$$\begin{array}{ccc} T_\beta T_\gamma & \xrightarrow{\lambda_\beta^\gamma} & T_{\gamma+1} \\ {\scriptstyle T_\beta \kappa^\gamma}\downarrow & & \downarrow{\scriptstyle \kappa_\gamma^{\gamma+1}} \\ T_\beta T_{\gamma+1} & \xrightarrow{\lambda_\beta^{\gamma+1}} & T_{\gamma+2}. \end{array}$$

Since T_β obviously preserves directed colimits, we can pass to a colimit transformation

$$T_\beta T_{\beta+\omega} \xrightarrow{\mu_\beta^{\beta+\omega}} T_{\beta+\omega}.$$

Assumed we have already defined a natural transformation

$$\mu_\beta^\gamma : T_\beta T_\gamma \to T_\gamma$$

for some $\gamma \geqslant \beta + \omega$, we also have a natural transformation

$$\mu_\beta^{\gamma+1} = \mu_\beta^\gamma T_\gamma : T_\beta T_{\gamma+1} \to T_{\gamma+1}.$$

The naturality of μ_β^γ yields a commutative diagram

$$\begin{array}{ccc} T_\beta T_\gamma & \xrightarrow{\mu_\beta^\gamma} & T_\gamma \\ {\scriptstyle T_\beta \kappa^\gamma}\downarrow & & \downarrow{\scriptstyle \kappa_\gamma^{\gamma+1}} \\ T_\beta T_{\gamma+1} & \xrightarrow{\lambda_\beta^{\gamma+1}} & \mu_{\gamma+2} \end{array}$$

allowing us to pass to colimits and define a transformation

$$\mu_\beta^\gamma : T_\beta T_\gamma \to T_\gamma \quad \text{for any } \gamma \geqslant \beta + \omega. \tag{3.15.7}$$

Recall, however, that morphisms in $\mathcal{W}$ are strictly monotone mappings. Thus, we conclude that

$$\gamma_{\mathrm{Dec}(T_\beta T_\gamma(\alpha))} \leqslant \gamma_{\mathrm{Dec}(T_\gamma(\alpha))} \quad \text{for } \alpha, \beta, \gamma \in \mathrm{Ord},\ \gamma \geqslant \beta + \omega, \tag{3.15.8}$$

which is (3.0.2). To prove (3.0.3), we prove the following.

Lemma 3.15.9. *Let $\beta \subseteq \alpha \in \mathrm{Obj}\ \mathcal{W}$ and let $\rho \in \alpha$. We have*

$$(\forall \sigma \in \beta)\rho > \sigma) \to ((\forall \sigma \in T_\gamma(\beta))\kappa^\gamma(\rho) > \sigma)$$

where $T_\gamma(\beta)$ is considered a part of $T_\gamma(\alpha)$ via the identification $T_\gamma(\subseteq)$.

Proof. An induction on γ. For $\gamma = 0$, the statement is easily verified. On the other hand, for γ limit, we easily get it by passage to colimits. Thus, the non-limit step remains. Let, thus, the statement be true for a certain γ and let $(\forall \sigma \in \beta)\rho > \sigma$. Using the induction hypothesis, we get

$$(\forall \sigma \in T_\gamma(\beta))\kappa^\gamma(\rho) > \sigma$$

and

$$(\forall \sigma \in T_\gamma T_\gamma(\beta))(\kappa^\gamma)_{T_\gamma(\alpha)}(\kappa^\gamma(\rho)) > \sigma.$$

Now

$$(\kappa^\gamma)_{T_\gamma(\alpha)}\kappa^\gamma(\rho) = (\kappa^\gamma \cdot \kappa^\gamma)(\rho) = \kappa^{\gamma+1}(\rho),$$

proving the statement. □

Lemma 3.15.9 is summed up into

$$\gamma_{\mathrm{Dec}T_\gamma(\beta)} < \gamma_{\mathrm{Dec}T_\gamma(\alpha)} \quad \text{for } \beta < \alpha \in \mathrm{Ord}, \tag{3.15.10}$$

which is (3.0.3). This completes the construction outlined in 3.0. We have proved the following theorem.

Theorem 3.16. *For $\gamma \geqslant 2$, α, $\gamma < |\Omega|^+$, we have*

$$\psi((\alpha+1)\omega^{2^{\gamma\omega}}, (\alpha+1)\omega^{2^{\gamma\omega}}) \geqslant \varphi_\gamma(\alpha).$$

Remark 3.17. Defining ψ by means of Strong R-functions (see [7]), we could eliminate the cardinality restriction on Ω. Using the methods of [7], one can prove that on arguments $<|\Omega|^+$, the R-functions and the strong R-functions for the Nash-Williams' k-system coincide.

4. Concluding remarks

4.1. Restricting ourselves (say) to two colours, we may define the (classical) Ramsey number $R_n(s_0, s_1)$ as the minimal number K such that for each mapping

$$r : \binom{K}{n} \to 2$$

there is an $X \subseteq K$ and an $i \in 2$ such that

$$|X| = s_i \quad \text{and} \quad r^{-1}(i) \supseteq \binom{X}{n}.$$

The existence of the numbers $R_n(s_0, s_1)$ was first proved by Ramsey [12]. The first attempt to find an explicit upper bound was in Skolem's paper [18]. Roughly, Ramsey's and Skolem's methods lead to the following recursive upper bound:

$$R_n(s_0, s_1) \leqslant R_{n-1}(R_n(s_0 - 1, s_1), R_n(s_0, s_1 - 1)). \tag{4.1.1}$$

In other words, we obtain an upper estimate by a function $A_n(s_0, s_1)$ satisfying

$$A_1(s_0, s_1) = s_0 + s_1 - 1, \tag{4.1.2}$$

$$A_n(s_0, s_1) = A_{n-1}(A_n(s_0 - 1, s_1), A_n(s_0, s_1 - 1)). \tag{4.1.3}$$

This is an example of an Ackermann function (function growing faster than any primitively recursive function).

Later, Erdős and Rado found a better upper bound (see [1]). This upper bound can be expressed as a 'tower function' (iterated exponentiation (see [4, Chapter 4].

The results in our Section 2 indicate that the ordinal numbers $\psi_\alpha(\gamma_0, \gamma_1)$ may be regarded as generalizations of the numbers $R_n(s_0, s_1)$, where n, s_0, s_1 become ordinals. (More exactly, $\psi_\alpha(\gamma_0, \gamma_1)$ corresponds to $R_n(s_0+1, s_1+1)-1$.) Indeed, Theorem 2.2 is a precise analogy of the estimate (4.1.1). Also, the function $\varphi_\alpha(\gamma_0 \oplus \gamma_1)$ is an analogy of the Ackermann function $A_n(s_0, s_1)$. Indeed, we have, e.g.,

$$\varphi_{\alpha+1}(\gamma) = \sup\{\varphi_\alpha^n(\varphi_{\alpha+1}(\gamma')) \mid n \in \omega,\ \gamma' > \gamma\}$$

which is an analogy of (4.1.3).

What is interesting and surprising is that an analogy of Erdős and Rado's stronger upper bound [1] does not hold. Instead, we have shown that asymptotically, in the range of countable ordinals, the Ramsey–Skolem's upper bound is quite good.

We would also like to remark that in order to obtain our lower bound, we did use the classical Erdős-Hajnal's stepping-up lemma (see [4, Chapter 4]) as the first step.

4.2. Finally, we would like to make a short remark on the use of category theory in this paper. As the reader might have noticed, we did not use any actual results in category theory except a few elementary observations about the categories $\mathcal{W}$, Set. What is then the purpose of introducing all that 'abstract nonsense'?

We feel that it is the same as in the majority of other fields of mathematics: category theory is a very convenient language.

Recall that in order to make any achievement at all (a hyper-exponential lower bound), we needed to construct $T_\omega(\alpha)$. While $T_n(\alpha)$ can be still plausibly described as $(T_0)^{2^n}(\alpha)$, to define $T_\omega(\alpha)$, we need to identify $T_n(\alpha)$ with a part of $T_{n+1}(\alpha)$. The problem is that the identification can be made in many equally natural ways. The two coming up immediately are via $T_n\kappa^n$ and $\kappa^n \cdot T_n$. Note that they actually are different: for example, $T_1(\alpha)$ consists of α, two copies of $[\alpha]^{<\omega}$ and one copy of $[[\alpha]^{<\omega}]^{<\omega}$. The two mappings $T_0\kappa^0$, $\kappa^0 T_0$ both send α to α, but make a difference in the matter as to which of its copies to send $[\alpha]^{<\omega}$. Naturally, this has an impact on the ordering $\cdots$. Of course, we could define $T_\gamma(\alpha)$ in one blow as an appropriate set of sequences and introduce some peculiar ordering. Although we could carry out the whole proof this way, we would be introducing a lot more information than what we are actually using.

Thus, in the language of a computer programmer, we are using category theory as a *data structure* to keep exactly the amount of information we are using each time. Of course, as a computer programmer would understand, this requires

some extra work on *updating* our data structure even through the steps where it does not seem necessary. This is exactly the meaning of our auxillary diagram-chasing lemmas.

References

[1] P. Erdős and R. Rado, Combinatorial theorems on classifications of subsets of a given set, Proc. London Math. Soc. 2 (1952) 417–439.

[2] H. Friedman, unpublished manuscripts, Ohio State University.

[3] H. Friedman, K. McAloon and S.G. Simpson, A finite combinatorial principle which is equivalent to the 1-consistency of predicative analysis. in: G. Matakides, ed., Patras Logic Symposium (North Holland, Amsterdam, 1982) 197–230.

[4] R.L. Graham, B.L. Rothschild and J.H. Spencer, Ramsey Theory (Wiley, New York, 1980).

[5] D.H.J. de Jongh and R. Parikh, Well-partial-orderings and hierarchies, Indag. Math. 39 (1977) 195–207.

[6] J. Ketonen and R. Solovay, Rapidly growing Ramsey functions, Ann. of Math. 113 (2) (1981) 2, 267–314.

[7] I. Kříž and R. Thomas, Ordinal types in Ramsey Theory and well-partial-ordering theory, in: Mathematics of Ramsey Theory, Algorithms and Combinatorics 5 (Springer, Berlin, 1990) 57–95.

[8] J.B. Kruskal, Well-partial-ordering, the tree theorem, and Vázsonyi's conjecture, Trans. Amer. Math. Soc 95 (1960) 210–225.

[9] S. Mac Lane, Categories for a Working Mathematician, Graduate Texts in Mathematics 5 (Springer, Berlin).

[10] C.St. J.A. Nash-Williams, On well-quasi-ordering transfinite sequences, Proc. Camb. Phil. Soc. 61 (1965) 33–39.

[11] J. Paris and L. Harrington, A Mathematical Incompleteness in Peano Arithmetic, in: J. Barwise, ed., Handbook of Mathematical Logic, (North Holland, Amsterdam) 1133–1142.

[12] F.P. Ramsey, On a problem of formal logic, Proc. London Math. Soc. 30 (1930) 264–286.

[13] D. Schmidt, Well-partial ordering and their maximal order types, Habilitationsschrift, Heidelberg, 1979.

[14] K. Schütte and S.G. Simpson, Ein in der reinen Zahlen-Theorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen, Archiv für Math. Logik und Grundlagen der Math.

[15] S.G. Simpson, Nonprobability of certain combinatorial properties of finite trees, in: L.A. Harrington et. al., eds, Harvey Friedman's Research on the Foundations of Mathematics (Elsevier, Amsterdam, 1985).

[16] S.G. Simpson, Unprovable theorems and fast growing functions, in: Contemporary Mathematics, Vol. 65 (1987) 359–394.

[17] S.G. Simpson, Subsystems of second order arithmetic, to appear.

[18] T. Skolem, Eine Kombinatorische Satz mit Anwendung auf ein Logisches Entscheidungsproblem, Fundamenta Math. 20 (1933) 254–261.

Discrete Mathematics 95 (1991) 169–182
North-Holland

Matchings from a set below to a set above

Paul Erdős*

Mathematical Institute of the Hungarian Academy of Sciences, Budapest, Réaltanoda u. 13–17, H-1053 Hungary

Jean A. Larson**

Department of Mathematics, University of Florida, Gainesville, FL 32611, USA

Received 16 November 1989
Revised 15 February 1990

Abstract

Erdős, P. and J.A. Larson, Matchings from a set below to a set above, Discrete Mathematics 95 (1991) 169–182.

One way to represent a matching in a graph of a set A with a set B is with a one-to-one function $m : A \to B$ for which each pair $\{a, m(a)\}$ is an edge of the graph. If the underlying set of vertices of the graph is linearly ordered and every element of A is less than every element of B, then such a matching is a *down-up* matching. In this paper we investigate graphs on well-ordered sets of type α and in many circumstances find either large independent sets of type β or down-up matchings with the initial set of some prescribed size γ. In this case we write $\alpha \to (\beta, \gamma\text{-matching})$.

0. Introduction

In their triple paper on set mappings [8] Erdős, Hajnal and Milner investigated the partition relation

$$\alpha \to (\beta, \text{infinite path})^2.$$

To see if this relation holds for some ordinals α and β, one checks to see if every graph on a set of vertices of type α either has an independent set of type β or has an infinite path. It is easy to see that a successor ordinal $\alpha = \beta + 1$ does not satisfy the relation $\alpha \to (\alpha, \text{infinite path})$, since the graph which joins the first β points to

* The hospitality of St. John's College, Cambridge University and the Mathematics Department of the University of Florida is gratefully acknowledged.

** Research partially supported by NSF grand DMS-8600931. The hospitality of the Mathematical Sciences Research Institute, Berkeley and the Mathematics Department of the Pennsylvania State University is also gratefully acknowledged.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

the last point and has no other edges is a graph which has no independent set of type α and no path of length longer than two edges (paths are not allowed to use the same vertex more than once). Limit ordinals offer more challenge. Erdős, Hajnal and Milner [8] proved that limit ordinals α less than $\omega_1^{\omega+2}$ satisfy the partition relation $\alpha \to (\alpha$, infinite path). Unfortunately this pleasing situation does not continue to hold for larger ordinals. Under the assumption of the Diamond Principle, Baumgartner and Larson [5] have shown that ordinals α at least as large as $\omega_1^{\omega+2}$ and of cardinality ω_1 satisfy the negative partition relation $\alpha \to (\omega_1^{\omega+2}$, infinite path). By way of contrast, under the assumption of Martin's Axiom for ω_1 and the assumption that the continuum hypothesis fails, Larson [10] has shown that cofinally many ordinals α less than ω_2 satisfy the positive partition relation $\alpha \to (\alpha$, infinite path).

One of our goals in this paper is to find a partition relation which is true 'most' of the time that it makes sense. In the current paper we continue to investigate graphs whose underlying vertex set is linearly ordered, but we propose as an alternative configuration to the infinite path the *down-up matching*, which is a matching of a set A with a set B where $A < B$, that is, where every element $a \in A$ is less than every element $b \in B$. If every graph on a set of vertices of type α either has an independent set of type β or a down-up matching with initial set A of cardinality κ, then we write

$$\alpha \to (\beta, \kappa\text{-matching}).$$

We were inspired to look at matchings because of long term interest in matchings in graphs which dates back in particular to König's Duality Theorem on matchings and covers. We have also long been interested in its generalization, Menger's Theorem. Many people worked on criteria for matchability and the related question of the existence of transversals. For countable graphs see for example [7–13]. Aharoni, Nash-Williams and Shelah [4] developed general criteria for the existence of matchings in infinite graphs, not just countable ones. For more on these matters see the Springer volume by Holz, Podewski and Steffens [9]. The criterion of Podewski and Steffens [12] from 1976 combines with work of Brualdi [6] to give a proof of König's Duality Theorem in the countable case. In 1983–84 Aharoni generalized König's Duality Theorem to all infinite graphs (see [1–2]). The proof even in countable graphs has been shown to be necessarily quite difficult (see Aharoni, Magidor and Shore [3]).

We have looked at two different kinds of problems. The first problem seeks to classify the set of ordinals α of cardinality κ which satisfy $\alpha \to (\alpha, \kappa\text{-matching})$. In solution to this problem we have proved the following theorem whose proof is in Section 1.

Theorem 1.4. *For every cardinal κ and every ordinal $\theta < \kappa^+$, $\kappa \cdot \theta \to (\theta, \kappa$-matching).*

The second problem we consider is a variation on the first where we ask for more in the way of the down-up matching between a set A and a set B with $A < B$, namely that A have a prescribed order type γ.

Thus by $\alpha \to (\beta, \gamma\text{-matching})$ we mean that every graph with vertex set α either has an independent set of type β or a down-up matching $m : A \to B$ with A of type γ. We have only started an exploration of this partition relation, by looking at the case for countably infinite indecomposable ordinals. We develop positive results in Section 2; we construct counter-examples in Section 3 and draw conclusions in Section 4. We have one general result which is uniformly true.

Theorem 4.1. *If ω^b and $\gamma \geqslant \omega^2$ are countably infinite indecomposable ordinals, then*

$$\gamma^b \cdot \omega \to (\omega^b, \gamma\text{-matching}) \quad \textit{and} \quad \gamma^b \nrightarrow (\omega^b + 1, \gamma\text{-matching}).$$

The next theorem lists the cases in which our results are optimal.

Theorem 4.2. *Suppose ω^b and $\gamma \geqslant \omega^2$ are countably infinite indecomposable ordinals.*

(1) *If b is a limit ordinal, then $\gamma^b \to (\omega^b, \gamma\text{-matching})$, and for all $\alpha < \gamma^b$, $\alpha \nrightarrow (\omega^b, \gamma\text{-matching})$.*

(2) *If b is finite or if b is a successor ordinal and $\gamma \geqslant \omega^\omega$, then $\gamma^b \cdot \omega \to (\omega^b, \gamma\text{-matching})$ and $\gamma^b \nrightarrow (\omega^b, \gamma\text{-matching})$.*

We conclude the paper in Section 4 with directions for further study.

1. Matchings with regular cardinals

In this section we prove that for all cardinals κ and all sufficiently large ordinals α of power κ, every graph on α either has a matching of a set A of size κ with a set B for which $A < B$ or has a large independent set. We start with three lemmas on bipartite graphs. To simplify the statements, let us call a set $C \subseteq A$ *small* if it has cardinality less than κ where κ is the size of A, and let us say C is *almost all* of A if $A - C$ is small.

Bipartite Lemma 1.1. *Suppose $G \subseteq A \times B$ is any bipartite graph and A has cardinality κ. Then either G has a κ-matching or there is a small subset set $A' \subseteq A$ and a subset $B' \subseteq B$ which is almost all of B so that every edge from A into B' arises from A' and every point of A' is joined to κ many points of B'.*

Proof. Define by recursion as long as possible sequence $\langle q_\alpha \rangle$ of elements of B and a partial matching r as follows: at stage α, let $q_\alpha \in B$ be an element b which is different from q_β for $\beta < \alpha$ and which is joined to some element $a = r(q_\alpha)$ of

$A - \{r(q_\beta) \mid \beta < \alpha\}$. If the recursive definition continues for all κ steps, then r is the desired matching, or r^{-1} if one prefers a matching from a subset of A into B. So suppose that the recursion stops at δ. Let A' be the range of the matching defined up to δ, $A' = \{r(q_\beta) \mid \beta < \delta\}$ and let B' be the difference of B and the domain of the matching, $B' = B - \{q_\beta \mid \beta < \delta\}$. Clearly A' is small since it has the same cardinality as δ. Similarly B' is large, since it omits a set of the same cardinality as δ. Since the recursion does not continue, all edges from points of B' go to points of A', and the lemma follows. □

High Valence Points Lemma 1.2. *Suppose $G \subseteq A \times B$ is any bipartite graph and A has cardinality κ. If for all regular $\lambda \leqslant \kappa$, the set A_λ has cardinality at least λ, where A_λ is the set of all a in A with degree at least λ, then G has a κ-matching.*

Proof. Define the matching m by recursion on $\alpha < \kappa$. At stage α, assume $\langle p_\beta \mid \beta < \alpha \rangle$ and $\langle m(p_\beta) \mid \beta < \alpha \rangle$ have been defined with each p_β in A and each $m(p_\beta)$ in B. Let $\lambda(\alpha)$ be $|\alpha|^+$. Since $\alpha < \kappa$, it follows that $\lambda(\alpha)$ is a regular cardinal with $\lambda(\alpha) \leqslant \kappa$. By hypothesis, $A_{\lambda(\alpha)}$ has cardinality at least $\lambda(\alpha)$, so there is an element p_α in $A_{\lambda(\alpha)}$ which is different from p_β for $\beta < \alpha$. Since the degree of p_α is at least $\lambda(\alpha)$, there is some point $m(p_\alpha)$ in B different from $m(p_\beta)$ for $\beta < \alpha$ to which p_α is joined. After κ steps this recursion defines the desired matching m. □

The above lemmas are used to prove the next lemma which is a variant of the Bipartite Lemma.

Matchless Lemma 1.3. *Suppose $G \subseteq A \times B$ is any bipartite graph and A has cardinality κ. If G has no κ-matching then for all regular cardinals $\lambda \leqslant \kappa$, there are sets $C(\lambda) \subseteq A$ of power less than κ and $D(\lambda) \subseteq B$ which is almost all of B so that all edges from A into $D(\lambda)$ actually come from $C(\lambda)$ and every element of $C(\lambda)$ is joined to at least λ points of $D(\lambda)$.*

Proof. Start by using the Bipartite Lemma 1.1 to get sets $A' \subseteq A$ of power less than κ and $B' \subseteq B$ which is almost all of B so that all edges in G which end in B' arise in A'. For notational convenience let G' be the restriction of G to $A' \times B'$.

Since G' has no κ-matching, by the High Valence Points Lemma 1.2, there is some regular cardinal $\mu \leqslant \kappa$ so that A_μ has power less than μ where, for any regular cardinal ν, A_ν is the set of elements in A' having degree at least ν in G'.

Let ν be the maximum of μ, λ and $|A'|^+$. Since all these cardinals are regular and less than or equal to κ, it follows for ν also. Let $C(\lambda)$ be A_ν. Since $\mu \leqslant \nu$, the set A_ν is a subset of A_μ. Let $E(\lambda)$ be the set of all points in B' joined to some element of $A' - C(\lambda)$. Since each point of $A' - C(\lambda)$ is joined to less than ν points of B' and there are less than ν points in $A' - C(\lambda)$, the set $E(\lambda)$ has less than ν elements. Thus $D(\lambda) = B' - E(\lambda)$ is almost all of B' and hence almost all

of B. Moreover, every element in $C(\lambda)$ is joined to at least ν elements of $D(\lambda)$ as required. □

With these lemmas in hand, we are ready to prove the main theorem on cardinal matchings.

Theorem 1.4 (cardinal matching). *For all infinite cardinals κ and all ordinals $\theta < \kappa^+$, $\kappa \cdot \theta \to (\theta, \kappa\text{-matching})$.*

Proof. We prove the theorem for a fixed κ by induction on θ. For $\theta = 1$, the theorem is trivially true, since any point is an independent set.

Suppose θ is greater than 1 and the theorem is true up to θ. If θ is decomposable, set $\nu = 2$ and decompose θ into the sum of smaller ordinals as $\theta = \theta(0) + \theta(1)$. If θ is indecomposable, let ν be the cofinality of θ and express θ as the sum over $\alpha < \nu$ of a nondecreasing sequence $\theta(\alpha)$ of indecomposable ordinals. Consider a graph G on $\kappa \cdot \theta$. If it has a down-up κ-matching, we are done so we may assume that G has no such κ-matching. Express the set $\kappa \cdot \theta$ as the sum over $\alpha < \nu$ of sets $X(\alpha)$ of type $\kappa \cdot \theta(\alpha)$. For α with $\alpha + 1 < \nu$, let $\mathrm{Tail}(\alpha)$ be the union of $X(\beta)$ for β with $\alpha < \beta < \nu$. Use the High Valence Points Lemma 1.2 on each G restricted to $X(\alpha) \times \mathrm{Tail}(\alpha)$ to get a regular cardinal $\lambda(\alpha) \leq \kappa$ so that the set $A(\alpha)$ of points of $X(\alpha)$ of degree at least $\lambda(\alpha)$ in the subgraph has cardinality less than $\lambda(\alpha)$. If ν is a regular cardinal, let I be a subset of ν of power ν on which the function λ which assigns to each α the value $\lambda(\alpha)$ is nondecreasing. Since the union over α in I of $X(\alpha)$ is order-isomorphic to $\kappa \cdot \theta$, without loss of generality, we may assume that I is all of ν, and thereby simplify our notation. Otherwise $\nu = 2$ and only $\lambda(0)$ has been defined. For notational uniformity define $\lambda(1) = \lambda(0)$, so in both cases λ is nondecreasing and defined on all of ν.

For each α satisfying $0 < \alpha < \nu$, let $\mathrm{Init}(\alpha)$ be the union over $\beta < \alpha$ of $X(\beta)$. For each such α, express $X(\alpha) - A(\alpha)$ as the sum of $Y(\alpha, \zeta)$ for $\zeta < \theta(\alpha)$, where each $Y(\alpha, \zeta)$ has type κ. Apply the Matchless Lemma 1.3 to each $Y(\alpha, \zeta)$ for the graph G restricted to $\mathrm{Init}(\alpha) \times Y(\alpha, \zeta)$ for the cardinal $\lambda(\alpha)$ to get a small set $C(\alpha, \zeta) \subseteq \mathrm{Init}(\alpha)$ and a subset $D(\alpha, \zeta) \subseteq Y(\alpha, \zeta)$ which is almost all of it, so that all edges from $\mathrm{Init}(\alpha)$ into $D(\alpha, \zeta)$ arise from $C(\alpha, \zeta)$ and every element of $C(\alpha, \zeta)$ is joined to at least $\lambda(\alpha)$ elements of $D(\alpha, \zeta)$. It follows that $C(\alpha, \zeta)$ is a subset of the union of the sets $A(\gamma)$ for $\gamma < \alpha$. Let $U(0)$ be the difference $X(0) - A(0)$ and for α with $0 < \alpha < \nu$, let $U(\alpha)$ be the union over $\zeta < \theta(\alpha)$ of $D(\alpha, \zeta)$. Since $A(0)$ is small and each $D(\alpha, \zeta)$ is almost all of $Y(\alpha, \zeta)$, for all $\alpha < \nu$, the set $U(\alpha)$ has type $\kappa \cdot \theta(\alpha)$. Furthermore, if $\alpha < \beta$, then there are no edges from $U(\alpha)$ to $U(\beta)$ since any such edge must arise from points of $A(\alpha)$ and all of those have been omitted. Apply the induction hypothesis to each $U(\alpha)$ to get an independent set $W(\alpha)$ of type $\theta(\alpha)$. Finally W, the union of these sets, is the desired independent set of type θ.

2. Positive results for countable ordinals

In this section we prove that if b and γ are countable ordinals and γ is infinite and indecomposable, then $\gamma^b \cdot \omega \to (\omega^b, \gamma\text{-matching})$. We can sharpen the result if b is a limit or if b is infinite and $\gamma = \omega$ to $\gamma^b \to (\omega^b, \gamma\text{-matching})$. The first step is a lemma on bipartite graphs reminiscent of the lemmas in the previous section.

Countable Bipartite Lemma 2.1. *Let α, β be countable indecomposable ordinals with α infinite. If A has type $\alpha \cdot \beta$, B is infinite and $G \subseteq A \times B$ is any bipartite graph, then either there is a matching from a subset of A of type $\omega \cdot \beta$ with an infinite subset of B, or there is a set $C \subseteq A$ of type α and a finite subset $D \subseteq B$ so that $G \restriction C \times B = C \times D$.*

Proof. Since A has type $\alpha \cdot \beta$, it is the sum over $\delta < \beta$ of sets $A(\delta)$ of type α. Let $\langle \delta_n \mid n < \omega \rangle$ be a sequence of elements less than β chosen so that every $\delta < \beta$ is repeated infinitely often. Define by recursion as long as possible a matching m on a sequence of elements $p_n \in A(\delta_n)$ with $m(p_n) \in B \backslash \{m(p_k) \mid k < n\}$. If the recursive definition continues for ω steps, then the matching m is defined on $P = \{p_n \mid n < \omega\}$ and has range $Q = \{m(p_n) \mid n < \omega\}$. Since P has infinitely many elements from each $A(\delta)$, it has order type at least $\omega \cdot \beta$ as required. Otherwise, for some $n = n_0$, there is no choice of p_n from $A(\delta_n)$ that can be paired with some $m(p_n)$ in $Q_n := B \backslash \{m(p_k) \mid k < \omega\}$. That is, all edges from points of $A(\delta)$ go to points of Q_n. Since the type of $A(\delta)$ is indecomposable and Q_n is finite, there are a subset $C \subseteq A(\delta)$ of type α and a finite set $D \subseteq Q_n$ so that every point of C is joined to every point of D, and furthermore, no point of C is joined to any point of $B \backslash D$. The sets C and D are the sets required for the second alternative of the lemma. □

The second lemma is a recursion lemma that allows us to paste together various matchings.

Recursion Lemma 2.2. *Suppose that γ is an infinite indecomposable countable ordinal, and $\langle \alpha_n \mid n < \omega \rangle$ and $\langle \beta_n \mid n < \omega \rangle$ are nondecreasing sequences of countable indecomposable ordinals. Let*

$$\alpha := \sum_{n<\omega} \alpha_n \cdot \gamma \quad \text{and} \quad \beta := \sum_{n<\omega} \beta_n.$$

If $\alpha_n \to (\beta_n, \gamma\text{-matching})$ for all $n < \omega$, then $\alpha \to (\beta, \gamma\text{-matching})$.

Proof. First write α as the sum of sets $A(n)$ of type $\alpha_n \cdot \gamma$ for $n < \omega$ and suppose that G is a graph on α. Let $B(n) := \bigcup \{A(k) \mid n < k < \omega\}$ and let $G(n) := G \cap A(n) \times B(n)$. Apply the Countable Bipartite Lemma to each of the graphs $G(n)$. If some application yields a matching of a set of type γ with a countable set

completely above it, then we are done. Otherwise, for each $n < \omega$, we get a set $C(n) \subseteq A(n)$ of type α_n and a finite set $D(n) \subseteq B(n)$ so that all edges from $C(n)$ to $B(n)$ actually go to $D(n)$. By the hypothesis on each α_n, either there is a matching of the desired kind, or there is a set $E(n) \subseteq C(n)$ of type β_n which is independent for the graph G. By recursion define an infinite subset $N \subseteq \omega$ so that if $n < m$ are both in N, then $E(m)$ has no point of $D(n)$. Let E be the union over $n \in N$ of the sets $E(n)$. Clearly E is independent for G. Moreover, since $\langle \beta_n \mid n < \omega \rangle$ is nondecreasing, the order type of $E := \sum_{n \in N} \beta_n$ is β as desired. □

Basis Lemma 2.3. *For all infinite countable indecomposable ordinals γ, $\gamma \cdot \omega \to (\omega, \gamma$-matching).*

Proof. In a trivial way $1 \to (1, \gamma$-matching). Thus with $\alpha_n = \beta_n = 1$, the Basis Lemma follows from the Recursion Lemma. □

Lemma 2.4 (ω^k-matching). *Let α and $\beta = \omega^b$ be countable indecomposable ordinals with $b = 0$ or b a limit ordinal, and let $k \geqslant 2$ be a positive integer. If $\alpha \to (\beta, \omega^k$-matching), then for any positive integer j,*

$$\alpha \cdot \omega^{jk+1} \to (\beta \cdot \omega^j, \omega^k\text{-matching}).$$

Thus in particular,

$$\omega^{jk+1} \to (\omega^j, \omega^k\text{-matching}).$$

Proof. The second partition relation follows from the implication by choosing $\alpha = \beta = 1$, since in a trivial way $1 \to (1, \omega^k$-matching). The proof of the implication is by induction on j. The basis step is $j = 1$. Let $\varepsilon_n := \alpha$ for all $n < \omega$. Then $\varepsilon := \sum_{n<\omega} \varepsilon_n \cdot \omega^k = \alpha \cdot \omega^{k+1}$. Let $\theta_n := \beta$ for all $n < \omega$. By hypothesis, $\varepsilon_n \to (\theta_n, \omega^k$-matching) for all $n < \omega$. Hence by the Recursion Lemma, $\varepsilon = \alpha \cdot \omega^{k+1} \to (\beta \cdot \omega, \omega^k$-matching).

For the induction step, suppose that $j > 1$ and that the lemma is true for $j - 1$. Further suppose that G is a graph on $\alpha \cdot \omega^{jk+1}$. Write $\alpha \cdot \omega^{jk+1}$ as the sum over $n < \omega$ of sets $A(n)$ of type $\alpha \cdot \omega^{jk}$, let $B(n) := \bigcup \{A(i) \mid n < i < \omega\}$ and let $G(n) := G \cap A(n) \times B(n)$. Notice that ω^{jk} can be expressed as $\omega^{(j-1)\cdot k+1} \cdot \omega^{k-1}$. For each $n < \omega$, write $A(n)$ as the sum over $\xi < \omega^{k-1}$ of $A(n, \xi)$ where $A(n, \xi)$ has type $\alpha \cdot \omega^{(j-1)\cdot k+1}$. For each $n < \omega$ and $\xi < \omega^{k-1}$, let $G(n, \xi) := G \cap A(n, \xi) \times B(n)$. Apply the Countable Bipartite Lemma to each of the graphs $G(n, \xi)$ and ask either for a matching of a set $P(n, \xi) \subseteq A(n, \xi)$ of type $\omega \cdot 1$ with some subset of $B(n)$ via a mapping $m_{n,\xi}$, or for a subset $U(n, \xi) \subseteq A(n, \xi)$ of type $\alpha \cdot \omega^{(j-1)\cdot k+1}$ and a finite set $D(n, \xi) \subseteq B(n)$ so that all edges of $G(n, \xi)$ from $U(n, \xi)$ to $B(n)$ go to $D(n, \xi)$.

Suppose for some $n < \omega$ and every $\xi < \omega^{k-1}$, application yields a matching. By recursion, we can thin each of the sets $P(n, \xi)$ out to an infinite subset $R(n, \xi)$ so

that the various matchings have disjoint ranges on the thinned sets. The union $R = \bigcup \{R(n, \xi) \mid \xi < \omega^{k-1}\}$ is a subset of $A(n)$ of type ω^k matched via $m = \bigcup \{m_{n,\xi} \mid \xi < \omega^{k-1}\}$ with an infinite subset of $B(n)$, hence we are done.

Thus we may suppose that for every $n < \omega$, there is some $\xi(n)$ so that the Bipartite Lemma yields sets $U(n) = U(n, \xi(n))$ and $D(n) = D(n, \xi(n))$. For each $n < \omega$, by the induction hypothesis applied to $j - 1$, either there is a matching of the desired kind inside $U(n)$ and we are done, or there is a set $V(n) \subseteq U(n)$ of type $\beta \cdot \omega^{j-1}$ which is independent for the graph G. By recursion on $n < \omega$, define an infinite subset $N \subseteq \omega$ so that if $n < p$ are both in N, then $V(p)$ has no point of $D(n)$. Let V be the union over $n \in N$ of the sets $V(n)$. Clearly V is independent for G. Moreover, V has type $\beta \cdot \omega^{j-1} \cdot \omega = \beta \cdot \omega^j$ as desired.

Therefore by induction on j, the lemma follows. □

Theorem 2.5. *Suppose that $\omega^b > \omega$ and γ are countably infinite indecomposable ordinals. Then $\gamma^b \cdot \omega \to (\omega^b, \gamma$-matching) and if b is a limit or b is infinite and $\gamma = \omega$, then $\gamma^b \to (\omega^b, \gamma$-matching).*

Proof. The case for $\gamma = \omega$ is special so we treat it first. If b is finite, then the theorem is true by the Cardinal Matching Theorem with $\kappa = \omega$, since $\gamma^b \cdot \omega = \omega^b \cdot \omega = \omega^{b+1}$. If b is infinite, then we again use the Cardinal Matching Theorem, but now use the computation that $\omega \cdot \omega^b = \omega^b$.

The remainder of the proof is by induction on b. The Basis Lemma says that the theorem is true for $b = 1$ and any suitable choice for γ.

The induction step when b is a successor ordinal is divided into cases according to the choice of γ. We have treated the case $\gamma = \omega$. Next suppose $\gamma = \omega^k$ where $k > 1$ is a positive integer and write $b = d + j$ where d is a limit ordinal and j is a positive integer. Note that $\omega^d \to (\omega^d, \omega^k$-matching) either trivially when $\omega^d = 1$ or by the induction hypothesis. Thus by Lemma 2.4, $\gamma^d = (\omega^k)^d \to (\omega^d, \omega^k$-matching) and the induction follows in this case. Finally suppose that $\gamma = \omega^e$ for some infinite ordinal e and write $b = c + 1$. From the induction hypothesis and the fact that $\gamma^{c+1} \geqslant \gamma^c \cdot \omega$, we conclude that $\gamma^{c+1} \to (\omega^c, \gamma$-matching). For all $n < \omega$, let $\alpha_n = \gamma^c$ and let $\delta_n = \omega^c$. Then $\gamma^{c+1} \cdot \omega$ is the sum over $n < \omega$ of $\alpha_n \cdot \gamma$ and ω^{c+1} is the sum over $n < \omega$ of δ_n. Hence by the Recursion Lemma, $\gamma^{c+1} \cdot \omega \to (\omega^{c+1}, \gamma$-matching). This last case completes the induction step for successors.

Finally consider the induction step when b is a limit ordinal. Let $\langle b(n) \mid n < \omega \rangle$ be a strictly increasing sequence of ordinals whose limit is b. Then $\alpha = \gamma^b$ is the sum over $n < \omega$ of $\gamma^{b(n)}$. Since the sequence of $b(n)$'s is strictly increasing and $\gamma \geqslant \omega$, γ^b is also the sum over $n < \omega$ of $\alpha_n := \gamma^{b(n)} \cdot \omega$ and $\alpha_n \cdot \gamma$. Let $\beta_n := \omega^{b(n)}$. Then β is the sum over $n < \omega$ of β_n. By the induction hypothesis and the fact that α_n is greater than or equal to $\gamma^{b(n)}$, we conclude that $\alpha_n \to (\beta_n, \gamma$-matching). Hence by the Recursion Lemma, $\gamma \to (\omega^b, \gamma$-matching), which completes this induction step.

Therefore by induction on b, the theorem is true for all b and γ. □

3. Counterexamples for countable ordinals

In this section we present some arguments which give limitations on when the partition relation $\alpha \to (\beta, \gamma\text{-matching})$ can hold. Before we go into the main constructions, we present some simple results which indicate why we have focused on α indecomposable. The first is a simple monotonicity result.

Monotonicity Lemma 3.1. *Suppose α, β, γ, α'; β' and γ' are all ordinals.*

(1) *If $\alpha \to (\beta, \gamma\text{-matching})$ and $\alpha' > \alpha$, then $\alpha' \to (\beta, \gamma\text{-matching})$.*

(2) *If $\alpha \to (\beta, \gamma\text{-matching})$ and $\beta' < \beta$, then $\alpha \to (\beta' \gamma\text{-matching})$.*

(3) *If $\alpha \to (\beta, \gamma\text{-matching})$ and $\gamma' < \gamma$, then $\alpha \to (\beta, \gamma'\text{-matching})$.*

Lemma 3.2 (on decomposable ordinals). *Suppose $\alpha = \alpha_0 + \alpha_1$ is decomposable with $\alpha_1 \leqslant \alpha_0 < \alpha$. If β and γ are both indecomposable, then $\alpha \to (\beta, \gamma\text{-matching})$ if and only if either $\alpha_0 \to (\beta, \gamma\text{-matching})$ or $\alpha_1 \to (\beta, \gamma\text{-matching})$.*

Proof. One direction follows from the Monotonicity Lemma. To prove the contrapositive of the other direction assume one has witnesses to G_0 and G_1 to the negations of the two partition relations for α_0 and α_1 respectively. Then since α is the sum, one can define G on α as the required counterexample essentially as the union of G_0 and G_1. $\square$

For the purpose of smoothing inductive arguments, we extend the definition of the partition relation to include a case in which the desired independent set is bounded in our original set. To that end, write $\alpha \nrightarrow (\text{bounded } \beta, \gamma\text{-matching})$ to mean that there is no bounded independent set of type β and no γ-matching. Clearly the above monotonicity results extend to this partition relation.

Lemma 3.3 (negative lifting). *Suppose that α, β and γ are all countably infinite indecomposable ordinals and that $\gamma \geqslant \omega^2$.*

(1) (Small) *If $\alpha \nrightarrow (\beta, \gamma\text{-matching})$ then $\alpha \cdot \omega \nrightarrow (\text{bounded } \beta, \gamma\text{-matching})$.*

(2) (Medium) *Let δ be an indecomposable ordinal with $\gamma = \omega \cdot \delta$. If $\alpha \nrightarrow (\text{bounded } \beta, \gamma\text{-matching})$, then $\alpha \cdot \delta \nrightarrow (\beta + 1, \gamma\text{-matching})$.*

Proof. To start the proof, let $G \subseteq [\alpha]^2_<$ be a graph which has no γ-matching and for the first item no independent set of type β and for the second item no bounded independent set of type β. Let δ be as given in the second item with $\gamma = \omega \cdot \delta$ and let it be ω for the first item. The set $\delta \times \alpha$ ordered lexicographically has type $\alpha \cdot \delta$ Let $\mathrm{INDEX} : \delta \times \alpha \to \omega$ be a one-to-one onto mapping, and let $\mathrm{BLOCK} : \delta \times \alpha \to \omega$ be a mapping with the properties that for each $c < \delta$, the function $\mathrm{BLOCK}(c, \cdot)$ is monotonic and increases to infinity and if $c \neq c'$, then the ranges of $\mathrm{BLOCK}(c, \cdot)$ and $\mathrm{BLOCK}(c' \cdot)$ are disjoint. Note that INDEX simply gives the index associated with each element relative to some enumeration

of the set $\delta \times \alpha$. The function BLOCK is more complicated. It reflects the structure of $\alpha \cdot \delta$ as a sum over $c < \delta$ of sets of type α and of each these sets of type α as a sum of ω many pieces. These small pieces are what we think of as 'blocks' of the partition.

Let $F \subseteq [\delta \times \alpha]^2$ be the graph which has an edge between $\langle c, a \rangle$ and $\langle c', a' \rangle$ if and only if either $c = c'$ and a and a' are joined in G or $c < c'$ and $\mathrm{BLOCK}(c, a) > \mathrm{INDEX}(c', a')$. Call an edge *red* if it satisfies the first clause of this definition and *blue* if it satisfies the second clause.

To prove the lemma, it is enough to show that F has no γ-matching and no bounded independent set of type β for the first item, and no independent set of type $\beta + 1$ for the second item.

First consider the matchings. Assume by way of contradiction that a set A of type γ is matched with a set B via a mapping m and every element of A is below every element of B. Since any joined pair is joined either by a red edge or by a blue edge and since γ is indecomposable, we may assume without loss of generality that all edges of the matching have the same color. Let $\langle w, z \rangle$ be the least element of B.

If $\langle c, a \rangle$ is joined to $m(\langle c, a \rangle) = \langle c', a' \rangle$ by a red edge, then $c = c' \leqslant w \leqslant c'$. Thus, if all pairs from A are matched by red edges, then all pairs begin with the same element w. Since F restricted to the set of sequences that begin with w is isomorphic to G, and G has no γ-matching, it follows that all pairs from A are matched by blue edges.

If for some b the set A intersects the 'block' mapped to b in an infinite set $V(b)$, then we have the contradiction that an infinite set $V(b)$ is matched into a finite set, namely the set of all elements $\langle c', a' \rangle$ with $\mathrm{INDEX}(c', a') < b$. Thus for each b, A intersects the 'block' mapped to b in a finite set. Consequently, A intersects each $\{c\} \times \alpha$ in a set of type at most ω. Since every element of A is less than $\langle w, z \rangle$, the least element of B, it follows that the set A is a subset of $(w + 1) \times \alpha$. Therefore A has type at most $\omega \cdot (w + 1) < \omega \cdot \delta$ which contradicts our assumption that A has type γ. This contradiction shows that F has no γ-matching.

Next we consider independent sets. Assume by way of contradiction that X is an independent set of type β which is bounded by $\langle t, \tau \rangle$ for the first item and remains independent with the addition of $\langle t, \tau \rangle$ for the second item.

For the first item, t is finite. Since β is indecomposable and X is a subset of the union over $c \leqslant t$ of $\{c\} \times \alpha$, for some $s \leqslant t$, X intersects $\{s\} \times \alpha$ in a set of type β. However, on $\{s\} \times \alpha$, the graph F is isomorphic to the graph G, and G has no independent sets of type β. This contradiction completes the proof of the first item of the lemma.

Finally consider the second item. As in the previous case, for each c, the graph F on $\{c\} \times \alpha$ is isomorphic to the graph G, which now has no bounded independent sets of type β. Thus for each b, the set X intersects the 'block' mapped to b in a set $W(b) \subseteq \{c\} \times \alpha$ for some c where $W(b)$ has type less than β

since it is bounded in $\{c\} \times \alpha$. Since X has type β and β is indecomposable, the set B of b for which $W(b)$ is non-empty must be infinite. Let b in B be greater than INDEX(t, τ) and let $\langle c, a\rangle$ be in $W(b)$. Then $\langle c, a\rangle$ and $\langle t, \tau\rangle$ are joined by a blue edge contradicting the assumption that $X \cup \{\langle t, \tau\rangle\}$ is independent. This contradiction proves the second item of the lemma and completes the proof. □

For the first corollary we combine the parts of the above lemma in various ways with the monotonicity lemma.

Corollary 3.4 (more lifting). *Suppose that α, β and γ are all countably infinite indecomposable ordinals and that $\gamma \geqslant \omega^2$.*

(1) (Large) *If $\alpha \nrightarrow (\beta, \gamma\text{-matching})$ then $\alpha \cdot \gamma \nrightarrow (\beta + 1, \gamma\text{-matching})$.*

(2) (Alternate) *If $\alpha \nrightarrow (\beta + 1, \gamma\text{-matching})$, then $\alpha \cdot \gamma \nrightarrow (\beta \cdot \omega + 1, \gamma\text{-matching})$.*

Proof. The large lifting follows from small and medium liftings. The alternate lifting follows from monotonicity followed by small and medium liftings. □

The next corollary shows that for matchings of finite powers of ω in graphs defined on finite powers of ω, the positive results are sharp.

Corollary 3.5. *For all positive integers j, k with $k \geqslant 2$, $\omega^{jk} \nrightarrow (\omega^{j-1} + 1, \omega^k\text{-matching})$.*

Proof. The proof is by induction on j. The basis of the induction is $j = 1$. In this case, the complete graph is the required example. For $j > 1$ first use the induction hypothesis which gives $\omega^{(j-1)k} \nrightarrow (\omega^{j-2} + 1, \omega^k\text{-matching})$ and then use Corollary 3.5 (alternate lifting) to complete the induction step. □

The next corollary in which the desired size matching is again a finite power of ω is a bit of a disappointment since there is a gap between the counterexample it provides and the positive result of the previous section.

Corollary 3.6. *For any positive integers j, k with $k \geqslant 2$ and any countable indecomposable ordinal β of the form $\beta = \omega^b$ for some limit ordinal b,*

$$\beta \cdot \omega^{jk-1} \nrightarrow (\beta \cdot \omega^{j-1} + 1, \omega^k\text{-matching}).$$

Proof. The proof is by induction on j. For the basis ($j = 1$) start with the counterexample given by the empty graph which witnesses $\beta \nrightarrow$ (bounded β, ω^k-matching). Apply Lemma 3.3 (medium lifting) to conclude that $\beta \cdot \omega^{k-1} \nrightarrow$ ($\beta + 1$, ω^k-matching) as desired. For the induction step, apply Corollary 3.5 (alternate lifting). □

The final lifting lemma is designed for the limit stages of the theorem to follow which proves some partition relations of the form $\alpha \nrightarrow (\beta, \gamma\text{-matching})$ by induction on β.

Bounded Set Lifting Lemma 3.7. *Suppose β and γ are countable indecomposable ordinals. If $\alpha_n \nrightarrow (\beta, \gamma\text{-matching})$ for a nondecreasing sequence of indecomposable ordinals $\langle \alpha_n \mid n < \omega \rangle$ and $\alpha = \sum_{n<\omega} \alpha_n$,* then *$\alpha \nrightarrow$ (bounded β, γ-matching) and consequently $\alpha \nrightarrow (\beta + 1, \gamma\text{-matching})$.*

Proof. Write α as the sum over $n < \omega$ of sets $A(n)$ where $A(n)$ has type α_n, and let $G(n)$ be a graph on $A(n)$ witnessing $\alpha_n \nrightarrow (\beta, \gamma\text{-matching})$. Let G be the union over $n < \omega$ of the graphs $G(n)$. Since every edge of G lies entirely inside some $A(n)$, there can be no γ-matching.

Let I be an independent set for G. For each $n < \omega$ let $I(n)$ be the intersection of I with $A(n)$. By the hypothesis on each α_n, the sets $I(n)$ all have type less than β. Since β is indecomposable, for each $m < \omega$, the union of the sets $I(n)$ for $n < m$ also has type less than β. Hence I is not a bounded independent set of type β and it has type at most $\beta < \beta + 1$.

Since G has no bounded independent set of type β, no independent set of type $\beta + 1$ nor any γ-matching, it is the example needed to prove the lemma. □

Next we consider γ-matchings where γ is an indecomposable ordinal of infinite exponent.

Theorem 3.8. *Suppose that $\beta = \omega^b$ and $\gamma \geqslant \omega^2$ are countably infinite indecomposable ordinals.*

(1) *If b is a limit, then $\gamma^b \nrightarrow$ (bounded ω^b, γ-matching).*

(2) *If b is finite or if b is a successor and $\gamma \geqslant \omega^\omega$, then $\gamma^b \nrightarrow (\omega^{b-1} + 1, \gamma$-matching).*

(3) *If b is an infinite successor, $\gamma < \omega^\omega$ and $\gamma^b = \omega^\delta$, then $\omega^{\delta-1} \nrightarrow (\omega^{b-1} + 1, \gamma$-matching).*

Proof. First suppose that γ is less than ω^ω, namely that γ is ω^k for some finite $k \geqslant 2$. If b is a limit ordinal, the γ^b is the same as ω^b, so the empty graph is the desired counterexample. If b is finite, then the theorem follows from Corollary 3.5. If b is an infinite successor ordinal, $b = d + j$ where d is a limit ordinal and j is finite, then γ^b is $\omega^d \cdot \omega^{jk}$ and ω^{b-1} is $\omega^d \cdot \omega^{j-1}$, so this case follows by application of Corollary 3.6.

Fix $\gamma \geqslant \omega^\omega$ and prove the theorem induction on b. Notice that the empty graph shows $\gamma \nrightarrow (2, \gamma\text{-matching})$, and that $\omega^0 + 1 = 2$. Thus this graph starts the induction for $b = 1$. If $b = d + 1$ is a successor either of 0 or another successor ordinal, use Corollary 3.4 (alternate lifting). If $b = d + 1$ is the successor of a limit ordinal d, then use Lemma 3.3 (medium lifting). Finally if b is a limit ordinal,

first use monotonicity and the induction hypothesis to conclude for all $d < b$ that $\gamma^d \nrightarrow (\omega^b, \gamma\text{-matching})$ and then use the Bounded Set Lifting Lemma 3.7 to reach the desired conclusion. Thus by induction, the theorem holds for all b. □

The final theorem in this section is a corollary of the previous one.

Theorem 3.9. *For all infinite indecomposable ordinals* ω^b *and* $\gamma \geqslant \omega^2$, $\gamma^b \nrightarrow (\omega^b + 1, \gamma\text{-matching})$.

Proof. If b is finite or a limit ordinal or if it is a successor ordinal and $\gamma \geqslant \omega^\omega$, then Theorem 3.9 follows from Theorem 3.8 by monotonicity. Suppose that b is an infinite successor ordinal and $\gamma < \omega^\omega$. Further suppose that $\gamma^b = \omega^\delta$ for some ordinal δ. Notice in the proof of Theorem 3.8 that δ has the form $d + jk$ where $\gamma = \omega^k$, so that subtraction makes sense, and use that theorem to get $\omega^{\delta-1} \nrightarrow (\omega^{b-1} + 1, \gamma\text{-matching})$. By monotonicity, conclude that $\omega^{\delta-1} \nrightarrow (\omega^b, \gamma\text{-matching})$. Next apply Lemma 3.3 (small lifting) to get $\omega^\delta \nrightarrow (\text{bounded } \omega^b, \gamma\text{-matching})$. Finally use monotonicity to conclude the theorem. □

4. Conclusions and questions

In the final section we pull together the results of the previous sections to see what remains to be done. The section on cardinal matchings was satisfyingly complete and requires no further discussion here. The next theorem puts together a positive result from Section 2 (Theorem 2.5) and a counter-example from Section 3 (Theorem 3.9) to give a general statement true for all countably infinite indecomposable ordinals and nontrivial ordinal matchings.

Theorem 4.1. *If* ω^b *and* $\gamma \geqslant \omega^2$ *are countably infinite indecomposable ordinals, then*

$$\gamma^b \cdot \omega \rightarrow (\omega^b, \gamma\text{-matching})) \quad \textit{and} \quad \gamma^b \nrightarrow (\omega^b + 1, \gamma\text{-matching}).$$

The next theorem lists the cases in which our results are optimal.

Theorem 4.2. *Suppose* ω^b *and* $\gamma \geqslant \omega^2$ *are countably infinite indecomposable ordinals.*

(1) *If* b *is a limit ordinal, then* $\gamma^b \rightarrow (\omega^b, \gamma\text{-matching})$, *and for all* $\alpha < \gamma^b$, $\alpha \nrightarrow (\omega^b, \gamma\text{-matching})$.

(2) *If* b *is finite or if* b *is a successor ordinal and* $\gamma \geqslant \omega^\omega$, *then* $\gamma^b \cdot \omega \rightarrow (\omega^b, \gamma\text{-matching})$ *and* $\gamma^b \nrightarrow (\omega^b, \gamma\text{-matching})$.

Proof. The second item follows from Theorem 2.5 and Theorem 3.8 with some help from the Monotonicity Lemma.

The first item starts with a part of Theorem 2.5. To begin the proof of the remainder, notice that if b is a limit ordinal and $\langle b(n) \mid n < \omega \rangle$ is a cofinal sequence, then $\langle \gamma^{b(n)} \mid n < \omega \rangle$ is cofinal in γ^b. By Theorem 3.9, $\gamma^{b(n)} \nrightarrow (\omega^{b(n)} + 1, \gamma\text{-matching})$. Since $\omega^{b(n)} + 1$ is less than ω^b, the remainder of the first item follows by monotonicity. □

The above theorem gives a general result true for all countably infinite indecomposable ordinals ω^b and $\gamma > 2$. It also lists the case in which we have results which are optimal of this form. The following result obtained from Theorem 2.5 and Theorem 3.8 leaves an obvious open question.

Proposition 4.3. *Suppose that j and k are positive integers with $k \geqslant 2$ and η is a limit ordinal. Then $\omega^{\eta+jk+1} \rightarrow (\omega^{\eta+j}, \omega^k\text{-matching})$ and $\omega^{\eta+jk-1} \nrightarrow (\omega^{\eta+j}, \omega^k\text{-matching})$.*

Question 4.4. Suppose that j and k are positive integers with $k \geqslant 2$ and η is a limit ordinal. Is it true that $\omega^{\eta+jk} \rightarrow (\omega^{\eta+j}, \omega^k\text{-matching})$?

Another direction for study is the case for uncountable ordinals and uncountable ordinal matchings, but we have not yet had time to investigate it.

References

[1] R. Aharoni, On a duality principle in infinite bipartite graph, J. London Math. Soc. (2) 28 (1983) 385–392.

[2] R. Aharoni, König's duality theorem for infinite bipartite graphs, J. London Math. Soc. (2) 29 (1984) 1–12.

[3] R. Aharoni, M. Magidor and R.A. Shore, On the strength of König's duality theorem.

[4] R. Aharoni, C.St.J.A. Nash-Williams and S. Shelah, A general criterion for the existence of transversals, J. London Math. Soc. (3) 47 (1983) 43–68.

[5] J. Baumgartner and J. Larson, A diamond example of a graph with no infinite path, J Pure Appl. Logic, 47 (1) (1990) 1–10.

[6] R.A. Brualdi, Strong transfinite version of König's duality theorem, Monatsh. Math. 75 (1971) 106–110.

[7] R.M. Damerell and E.C. Milner, Necessary and sufficient conditions for transversals of countable set systems, J. Combin. Theory Ser. A 17 (1974) 350–379.

[8] P. Erdős, A. Hajnal and E.C. Milner, Set mappings and polarized partition relations, in: Combinatorial Theory and its Application, Colloq. Math. Soc. János Bolyai 4 (1969) 327–363.

[9] M. Holz, K.P. Podewski and K. Steffens, Injective choice functions, Lecture Notes in Math., Vol. 1238 (Springer, Berlin, 1987).

[10] J. Larson, Martin's axiom and infinite graphs: large independent sets or infinite paths, J. Pure Appl. Logic 47 (1) (1990) 31–39.

[11] C.St.J.A. Nash-Williams, Another criterion for marriage in denumerable societies, Ann. Discrete Math. 3 (1978) 165–179.

[12] K.P. Podewski and K. Steffens, Injective choice functions for countable families, J. Combin. Theory Ser. B 21 (1976) 40–46.

[13] S. Shelah, Notes on a partition calculus, in: Infinite and Finite Sets, Part III, Colloq. Math. Soc. János Bolyai 10 (1975) 1257–1276.

Discrete Mathematics 95 (1991) 183–191
North-Holland

A partition relation for triples using a model of Todorčević*

E.C. Milner
University of Calgary, Alta., Canada

K. Prikry
University of Minnesota, Minneapolis, MN, USA

Received 22 September 1989
Revised 29 June 1990

Abstract

Milner, E.C. and K. Prikry, A partition relation for triples using a model of Todorčević, Discrete Mathematics 95 (1991) 183–191.

Todorčević has shown that there is a *ccc* extension $\mathcal{M}$ in which $\mathrm{MA}_{\omega_1} + 2^{\omega} = \omega_2$ holds and also in which the partition relation $\omega_1 \to (\omega_1, \alpha)^2$ holds for every denumerable ordinal α. We show that the partition relation for triples

$$\omega_1 \to (\omega 2 + 1, 4)^3$$

holds in the model $\mathcal{M}$, and hence by absoluteness this is a theorem in ZFC.

1. Introduction

For an ordinal γ, a positive integer r and linear order types φ, ψ_i $(i < \gamma)$, the partition relation

$$\varphi \to (\psi_i)^r_{i<\gamma} \tag{1.1}$$

means that whenever $(S, <)$ is an ordered set of order type $\mathrm{tp}(S) = \gamma$ and $\{K_i : i < \gamma\}$ is a partition of $[S]^r = \{X \subseteq S : |X| = r\}$, then there are $i < \gamma$ and $T \subseteq S$ such that $[T]^r \subseteq K_i$ and $\mathrm{tp}(T) = \psi_i$. In the case when $\gamma = 2$ we write (1.1) as $\varphi \to (\psi_1, \psi_2)^r$, and the negation of this is expressed by replacing the arrow $\to$ by a nonarrow $\nrightarrow$.

Very few partition relations of this kind are known when $r \geqslant 3$ and the order types are not cardinal numbers. Such relations were discussed in some detail in [5], and we proved in that paper that

$$\varphi \to (\omega + k, 4)^3 \tag{1.2}$$

* Research supported by NSERC Grant No. A5198 and NSF Grant MCS 830361.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

holds for any finite k and linear type φ which satisfies

$$\varphi \to (\omega)^1_\omega. \tag{1.3}$$

It was conjectured in [5] that the more general relation

$$\varphi \to (\alpha, m)^3$$

holds for all countable ordinals α and finite m, but we were unable to extend our method of proof of (1.2) to establish either of the next simplest cases of the conjecture

$$\omega_1 \to (\omega 2, 4)^3, \tag{1.4}$$

or

$$\omega_1 \to (\omega + 2, 5)^3. \tag{1.5}$$

In this paper we use some heavier artillery from [8] in order to prove that

$$\omega_1 \to (\omega 2 + 1, 4)^3, \tag{1.6}$$

which is slightly stronger than (1.4).

Our proof of (1.2) for the case $\varphi = \omega_1$ in [5] used the same type of argument employed by Baumgartner and Hajnal in [1]. Since

$$\omega_1 \to (\omega + k, 4)^3 \tag{1.7}$$

is an absolute statement relative to a *ccc* extension, it is sufficient to prove this under the additional assumption that Martin's Axiom MA_{ω_1} holds. However, we could not prove (1.4) by using the same combinatorial tools. Here we will prove (1.6) using a model of Todorčević. He proved [8] that there is a *ccc* extension $\mathcal{M}$ in which MA_{ω_1} holds, $2^\omega = \omega_2$, and also in which the relation

$$\omega_1 \to (\omega_1, \alpha)^2 \tag{1.8}$$

holds for all $\alpha < \omega_1$. We prove that (1.6) holds in the model $\mathcal{M}$ and so, by absoluteness, (1.6) is a theorem of ZFC. Note that (1.8) is independent of the axioms of ZFC since, by an earlier result of Hajnal [3], CH implies that $\omega_1 \not\to (\omega_1, \omega + 2)^2$. Let us remark that we only use the special case of (1.8) when $\alpha = \omega 2 + 1$; we were unable to obtain anything better by using the full strength of (1.8). Also, we should point out that the proof used in [5] to prove (1.7) could, by an argument due to Baumgartner and Hajnal, be adapted to prove (1.2) by using MA in place of MA_{ω_1}. This argument does not allow us to extend in the same way our proof of (1.6) to the more general relation $\varphi \to (\omega 2 + 1, 4)^3$ for an order type satisfying (1.3). Thus, for example, whether or not the relation

$$\lambda \to (\omega 2, 4)^3$$

holds is still open, where λ is the order type of the reals.

2. Notation and preliminary lemmas

We use the standard notation $[X]^{<\omega}$ to denote the set of all finite subsets of X. If $(S, \leqslant)$ is a linearly ordered set and $x \in S$, then $S(\geqslant x) = \{y \in S: y \geqslant x\}$. Also, if X, Y are subsets of S, then we write $X < Y$ if $x < y$ holds for all $x \in X$ and $y \in Y$. If $\{K_i: i < \gamma\}$ is any partition of $[S]^r$, then we write

$$\psi \in \mathrm{homog}(K_i)$$

if there is a subset $T \subseteq S$ such that $\mathrm{tp}(T) = \psi_i$ and T is *homogeneous for the class* K_i, i.e. if $[T]^r \subseteq K_i$. For finite r, s and sets A, B, we denote by $[A]^r \otimes [B]^s$ the set of all subsets $X \subseteq A \cup B$ such that $|A \cap X| = r$ and $|B \cap X| = s$.

We need the following easily proved consequence of Ramsey's Theorem.

Lemma 2.1. *If r_i, k_i are finite and $f_i : [\omega]^{r_i} \to k_i$ $(i < \omega_1)$, then there is a uniform ultrafilter $\mathcal{U}$ on ω which contains an f_i-homogeneous set for each $i < \omega_1$.*

Proof. Let $\alpha < \omega_1$ and suppose that we have already constructed f_β-homogeneous sets U_β for $\beta < \alpha$ so that the intersection of any finite number of these is infinite. Then there is an infinite set X such tht $X \backslash U_\beta$ is finite for each $\beta < \alpha$, and by Ramsey's Theorem [6] there is an f_α-homogeneous set $U_\alpha \subseteq X$. The sets U_α $(\alpha < \omega_1)$ generate an ultrafilter $\mathcal{U}$. $\square$

We also need the following special case of Solovay's Lemma (see e.g. [4, p. 287]).

Lemma 2.2. *Assume* MA_{ω_1}. *If the sets $A_i \in [\omega]^\omega$ $(i < \omega_1)$ have the property that the intersection of any finite number of them is infinite, then there is an infinite set $X \subseteq \omega$ such that $X \backslash A_i$ is finite for all $i < \omega_1$.*

3. A proof of (1.6)

We will use the same convention that was used in [5]; the letters A and B (possibly with suffixes or superfixes or primed) will always denote subsets of ω_1 which have respectively order types ω and ω_1 under the induced ordering.

As already observed in Section 1, it will be enough to prove that (1.6) holds in the model $\mathcal{M}$, i.e. we may and do assume that MA_{ω_1} holds and also that (1.8) holds with $\alpha = \omega 2 + 1$. Let $K_0 \cup K_1$ be any partition of $[\omega_1]^3$. We have to show that either

$$\omega 2 + 1 \in \mathrm{homog}(K_0) \tag{3.1}$$

or

$$4 \in \mathrm{homog}(K_1) \tag{3.2}$$

Lemma 3.1. *Let $A < B$ and assume that*:

(i) $[A]^2 \otimes [B]^1 \subseteq K_0$,

(ii) *the set* $\{a \in A: \{a\} \cup s \in K_1\}$ *is finite for each* $s \in [B]^2$, *and*

(iii) $(\forall x \in A \cup B)(\forall B_1 \subseteq B)(\exists B_2 \subseteq B_1)(\forall s \in [B_2]^2)\{x\} \cup s \in K_0$.

Then either (a) $\omega \in \mathrm{homog}(K_1)$, *or* (b) *there is* $Z \subseteq A \cup B$ *such that* $\mathrm{tp}(A \cap Z) = \omega$, $\mathrm{tp}(B \cap Z) = \omega + 1$ *and* $[Z]^3 \subseteq K_0$ (*i.e.* (3.1) *or* (3.2) *holds*).

Proof. We will assume that (a) is false and deduce that (b) holds. Let $B = \{b_\alpha: \alpha < \omega_1\}$, where $b_0 < b_1 < \cdots$. We claim that there are $\alpha < \omega_1$, $X \in [A]^\omega$ and $Y \in [\{b_\beta: \beta < \alpha\}]^\omega$ such that

$$[X]^1 \otimes [Y \cup \{b_\alpha\}]^2 \subseteq K_0, \tag{3.3}$$

and

$$[Y]^2 \otimes [\{b_\alpha\}]^1 \subseteq K_0. \tag{3.4}$$

The lemma follows from the claim since, by Ramsey's Theorem and the assumption that (a) is false, we can assume that X and Y are both K_0-homogeneous. Then by (i), (3.3) and (3.4) the lemma holds with $Z = X \cup Y \cup \{b_\alpha\}$.

Let $\alpha < \omega_1$ be fixed. We try to construct the sets X, Y to satisfy (3.3) and (3.4) in ω steps as follows: Let $n < \omega$ and suppose that we have already constructed n-element sets $X_n \subseteq A$ and $Y_n \subseteq \{b_\beta: \beta < \alpha\}$ so that

$$[X_n]^1 \otimes [Y_n \cup \{b_\alpha\}]^2 \subseteq K_0, \tag{3.5}$$

and

$$[Y_n]^2 \otimes [\{b_\alpha\}]^1 \subseteq K_0 \tag{3.6}$$

both hold. If possible we now select $x_n \in A \backslash X_n$ and $y_n \in \{b_\beta: \beta < \alpha\} \backslash Y_n$ so that (3.5) and (3.6) remain true with X_n, Y_n replaced respectively by $X_n \cup \{x_n\}$ and $Y_n \cup \{y_n\}$. If it is not possible to choose suitable x_n and y_n, the construction terminates and we define

$$n_\alpha = n, \qquad X^\alpha = X_n, \qquad Y^\alpha = Y_n. \tag{3.7}$$

If, for some α, this construction continues for infinitely many steps, then our claim is established. So we can assume that n_α, X^α, and Y^α are defined as above for all $\alpha < \omega_1$. Now by Fodor's theorem [2] there are a stationary set $S \subseteq \omega_1$, an integer $n \in \omega$, and fixed n-element sets X_n and Y_n, so that (3.7) holds for each $\alpha \in S$.

By a finite number of applications of the hypothesis (iii), it follows that there is an uncountable set $T \subseteq S$ such that

$$[X_n \cup Y_n]^1 \otimes [\{b_\alpha: \alpha \in T\}]^2 \subseteq K_0. \tag{3.8}$$

Choose γ, $\alpha \in T$ with $\gamma < \alpha$. By the hypothesis (ii), the set

$$F = \{x \in A: \{x\} \cup s \in K_1 \quad \text{for some } s \in [Y_n \cup \{b_\gamma, b_\alpha\}]^2\}$$

is finite and so we can choose $a \in A \backslash (F \cup X_n)$. Since α belongs to S it follows that (3.5) and (3.6) both hold, and these also hold with b_γ in place of b_α since γ belongs to S. From these facts and (3.7) and by our choice of the element a, it is now a simple matter to check that (3.5) and (3.6) both hold with X_n replaced by $X_n \cup \{a\}$ and Y_n replaced by $Y_n \cup \{b_\gamma\}$. But this contradicts the fact that $n_\alpha = n$, since for α the above construction could be continued for at least one more step. □

The next lemma shows how we use the hypothesis that

$$\omega_1 \to (\omega_1, \omega 2 + 1)^2 \tag{3.9}$$

holds in $\mathcal{M}$.

Lemma 3.2. *Assume that* MA_{ω_1} *and* (3.9) *both hold. Let* A, $B \subseteq \omega_1$ *be such that* $A < B$ *and* $[A]^2 \otimes [B]^1 \subseteq K_0$. *Then either* (3.1) *or* (3.2) *holds.*

Proof. Let $\mathcal{U}$ be a uniform ultrafilter on A. For $s \in [B]^2$ and $i = 0$ or 1, define $A_i(s) = \{a \in A: \{a\} \cup s \in K_i\}$. Then either $A_0(s)$ or $A_1(s)$ belongs to $\mathcal{U}$. Suppose that there is a subset $Y \subseteq B$ having order type $\mathrm{tp}(Y) = \omega 2 + 1$ and such that $A_1(s) \in \mathcal{U}$ for all $s \in [Y]^2$. Then either $[Y]^3 \subseteq K_0$ and (3.1) holds, or there is $t \in [Y]^3 \cap K_1$. But in this latter case we can choose $a \in \bigcap \{A_1(s): s \in [t]^2\}$, and then $[\{a\} \cup t]^3 \subseteq K_1$ and (3.2) holds. Therefore, we may assume that, whenever $Y \subseteq B$ has order type $\omega 2 + 1$, then there is some pair $s \in [Y]^2$ such that $A_0(s)$ belongs to $\mathcal{U}$. By (3.9) it follows that there is $B' \subseteq B$ such that $A_0(s)$ belongs to $\mathcal{U}$ for all $s \in [B']^2$. Also, since MA_{ω_1} holds, it follows by Lemma 2.2 that there is an infinite set $A' \subseteq A$ such that $A' \backslash A_0(s)$ is finite for all $s \in [B']^2$. Thus, replacing A by A' and B by B', we may assume that condition (ii) of Lemma 3.1 is satisfied.

Fix an element $x \in A \cup B$ and a subset $B_1 \subseteq B$ and consider the partition $[B_1]^2 = L_0 \cup L_1$ in which $s \in L_0$ if and only if $\{x\} \cup s \in K_0$. If there is $Y \subseteq B_1$ such that Y has order type $\omega 2 + 1$ and $[Y]^2 \subseteq L_1$, then, by the same argument that was used in the preceding paragraph, it follows that either $[Y]^3 \subseteq K_0$, or there is $t \in [Y]^3$ such that $[\{x\} \cup t]^3 \subseteq K_1$; in either case the lemma follows. Therefore, we can assume that there is no such Y and so by (3.9) again, it follows that there is $B_2 \subseteq B_1$ such that $\{x\} \cup s \in K_0$ for all $s \in [B_2]^2$. In other words, the condition (iii) of Lemma 3.1 also holds. But condition (i) of that lemma holds by hypothesis, and so the result follows. □

Specker [7] proved that the partition relation

$$\omega^2 \to (\omega^2, m)^2 \tag{3.10}$$

holds for any integer m. The next lemma is a strengthening of this under the assumption that MA_{ω_1} holds.

Lemma 3.3. *Assume* MA_{ω_1}. *Let* W_i $(i<\omega)$ *be pairwise disjoint infinite sets,* $W=\bigcup\{W_i: i<\omega\}$ *and let* $f_\alpha:[W]^2\to 2(\alpha<\omega_1)$. *Then there is an uncountable set* $B\subseteq\omega_1$ *such that either* (i) *there are* $l_0<l_1<l_2<\cdots<\omega$ *and sets* $H_i\in[W_{l_i}]^\omega$ $(i<\omega)$ *such that* $f_\alpha(s)=0$ *for all* $\alpha\in B$ *and* $s\in\bigcup\{[H_i]^1\otimes[H_j]^1: i<j<\omega\}$, *or* (ii) *for any positive integer* m, *there are* m *integers* $l_0<l_1<\cdots<l_{m-1}$ *and* m *sets* $H_i\in[W_{l_i}]^m$ $(i<m)$ *such that* $f_\alpha(s)=1$ *for all* $\alpha\in B$ *and* $s\in\bigcup\{[H_i]^1\otimes[H_j]^1: i<j<m\}$.

Remark. Specker's Theorem corresponds to the case when $W_i=\{\xi: \omega i\leqslant\xi\leqslant\omega(i+1)\}$ $(i<\omega)$ and the f_α $(\alpha<\omega_1)$ are all equal, say $f_\alpha=f$. In this case, if (i) holds then, by Ramsey's Theorem, either there is an infinite set X contained in some H_i such that $f(s)=1$ for all $s\in[X]^2$ or, for each i there is an infinite set X contained in some H_i such that $f(s)=1$ for all $s\in[X]^2$ or, for each i there is a set $H_i'\in[H_i]^\omega$ such that $f(s)=0$ for all $s\in[H]^2$, where $H=\bigcup\{H_i: i<\omega\}$ has order type ω^2. On the other hand, if (ii) holds, then there is a set Y of cardinality m such that $|Y\cap H_i|=1$ $(i<m)$ and $[Y]^2\subseteq K_1$. Of course, the relation (3.10) is absolute and so it follows that this holds in ZFC. In fact, the proof of the lemma given below contains a ZFC proof of Specker's Theorem (see Remark (*) in the following proof).

Proof of Lemma 3.3. Without loss of generality we may assume that $W_i=\{\langle i,j\rangle: i<j<\omega\}$ $(i<\omega)$. For each $\alpha<\omega_1$ define a function $g_\alpha:[\omega]^4\to 2^3$ as follows: If $s=\{a,b,c,d\}$ where $a<b<c<d<\omega$, then we set $g_\alpha(s)=\langle g_\alpha^0(s), g_\alpha^1(s), g_\alpha^2(s)\rangle$, where

$$g_\alpha^0(s)=f_\alpha(\{\langle a,b\rangle,\langle c,d\rangle\}),$$
$$g_\alpha^1(s)=f_\alpha(\{\langle a,c\rangle,\langle b,d\rangle\}),$$
$$g_\alpha^2(s)=f_\alpha(\{\langle a,d\rangle,\langle b,c\rangle\}).$$

By Lemma 2.1, there is a uniform ultrafilter $\mathcal{U}$ on ω which contains a g_α-homogeneous set for each $\alpha<\omega_1$, i.e. there are $G_\alpha\in\mathcal{U}$ such that g_α is constant on $[G_\alpha]^4$. By MA_{ω_1} and Lemma 2.2, there is an infinite set $G\subseteq\omega$ such that $G\backslash G_\alpha$ is finite for each $\alpha<\omega_1$. It follows that there are a finite set $F\subseteq\omega$ and an uncountable set $B\subseteq\omega_1$ such that $G'=G\backslash F\subseteq G_\alpha$ holds for all $\alpha\in B$. (Remark (*): In the case when the f_α are all equal, then we may set $G=G'=G_\alpha$, and we do not need Lemma 2.2.) Replacing B by an uncountable subset if necessary, we may assume that the value of g_α on $[G']^4$ is constant for all $\alpha\in B$, say $g_\alpha(s)=\langle i_0,i_1,i_2\rangle$ for all $\alpha\in B$ and $s\in[G']^4$.

Let L_n $(n<\omega)$ be infinite pairwise disjoint subsets of G' and assume that $l_i<l_j$ holds for $i<j<\omega$, where $l_n=\min(L_n)$. Put $H_n=\{\langle l_n,y\rangle: y\in L_n\backslash\{l_n\}\}$ $(n<\omega)$.

If $i_0=i_1=i_2=0$, then (i) holds. For, if $h=\langle l_m,y_m\rangle\in H_m$, $h'=\langle l_n,y_n\rangle\in H_n$ and $m<n$, then $f_\alpha(\{h,h'\})=0$.

If one or more of i_0,i_1,i_2 is equal to 1, then (ii) holds. We will verify this for the case when $i_2=1$; the other cases are similar. Choose m^2 integers $y_r^s\in L_r\backslash\{l_r\}$

for $0 \leqslant r,s < m$ so that

$$l_{m-1} < y^0_{m-1} < \cdots < y^{m-1}_{m-1} < y^0_{m-2} < \cdots < y^{m-1}_{m-2} < \cdots < y^0_0 < y^1_0 < \cdots < y^{m-1}_0$$

Then for $\alpha \in B$, $s < m$, $t < m$ and $i < j < m$, we have

$$f_\alpha(\{\langle l_i, y^s_i\rangle, \langle l_j, y^t_j\rangle\} = g^2_\alpha(\langle l_i, l_j, y^t_j, y^s_i\rangle) = i_2 = 1. \quad \square$$

We need one additional lemma. This result follows from Lemmas 3.1, 3.2, and 5.2 of [5], but since it is essential for the present argument, and for the convenience of the reader, we give the proof.

Lemma 3.4. *Assume* MA_{ω_1}. *Suppose that*

$$[A]^2 \otimes [B]^1 \nsubseteq K_0, \tag{3.11}$$

whenever $A, B \subseteq \omega_1$. *Then there are subsets* A'_n $(n < \omega)$ *and* B' *of* ω_1 *such that* $A'_0 < A'_1 < \cdots < B'$ *and*

$$[A'_m]^2 \otimes [A'_n \cup B']^1 \subseteq K_1$$

holds for all $m < n < \omega$.

Proof. By Lemma 2.1, for any $A \subseteq \omega_1$, there is a uniform ultrafilter $\mathcal{U}$ on A such that, for each $\rho \in \omega_1$ there are $i(\rho) < 2$ and $A_\rho \in \mathcal{U}$ such that $[A_\rho]^2 \otimes [\{\rho\}]^1 \subseteq K_{i(\rho)}$. By Lemma 2.2 there is an infinite subset $X \subseteq A$ such that $X \backslash A_\rho$ is finite for all $\rho < \omega_1$. There is an uncountable set $B \subseteq \omega_1$ such that both $X \backslash A_\rho = F$ and $i(\rho) = i$ are constant for all $\rho \in B$. Thus, if we put $A' = X \backslash F$, then $[A']^2 \otimes [B]^1 \subseteq K_i$. From the hypothesis (3.11) it follows that $i = 1$. It follows from this and a simple induction argument that there are subsets A_n $(n < \omega)$ such that $A_0 < A_1 < \cdots$ and $[A_m]^2 \otimes [A_n]^1 \subseteq K_1$ holds for $m < n < \omega$.

By Lemma 2.1 again, there is a uniform ultrafilter $\mathcal{U}_n$ on A_n $(n < \omega)$ such that for each $\rho < \omega_1$ there are $i(n, \rho) < 2$ and $A(n, \rho) \in \mathcal{U}_n$ such that $[A(n, \rho)]^2 \otimes [\{\rho\}]^1 \subseteq K_{i(n,\rho)}$. By the same argument as above, for each n there are only countably many ρ such that $i(n, \rho) = 0$. Hence there is $B \subseteq \omega_1$ such that $i(n, \rho) = 1$ for all $n < \omega$ and all $\rho \in B$. By Lemma 2.2, there is an infinite set $X_n \subseteq A_n$ such that $X_n \backslash A(n, \rho)$ is finite for all $\rho \in B$.

For $j \in \omega$, $\rho \in B$ and $s \in [B]^{<\omega}$ define

$$F(j, \rho) = X_j \backslash A(j, \rho), \qquad F(j, s) = \bigcup\{F(j, \sigma) \colon \sigma \in s\}$$

Now consider the set P of all ordered pairs $\langle n, s\rangle$ such that $n \in \omega$, $s \in [B]^{<\omega}$ and

$$B(n, s) = \{\rho \in B \colon F(j, \rho) \subseteq F(j, s) \quad \text{for all } j \leqslant n\}$$

is uncountable. We order P by the rule that $\langle n, s\rangle \leqslant \langle n_1, s_1\rangle$ if and only if $n \leqslant n_1$, $s \subseteq s_1$ and $s_1 \backslash s \subseteq B(n, s)$. Note that this implies that

$$X_j \cap \bigcap\{A(j, \sigma) \colon \sigma \in s\} = X_j \cap \bigcap\{A(j, \sigma) \colon \sigma \in s_1\}$$

holds for all $j \leqslant n$.

We claim that P is *ccc*, i.e. if Q is an uncountable subset of P, then there are q_1, q_2 in Q and $p \in P$ such that $q_1 \leqslant p$ and $q_2 \leqslant p$. To see this, for any $p = \langle n, s \rangle \in P$, define $\bar{p} = \langle F_0, F_1, \ldots, F_n \rangle$, where $F_j = F(j, s)$. Since there are only countably many sequences of this kind, it follows that there are distinct elements $q_1 = \langle n, s_1 \rangle$ and $q_2 = \{n, s_2\rangle$ in Q which have a common first term and are such that $\bar{q}_1 = \bar{q}_2$. Put $s = s_1 \cup s_2$, $p = \langle n, s \rangle$. Since $F(j, s) = F(j, s_1) = F(j, s_2)$ for $j \leqslant n$, it follows that $p \in P$ and $q_1 \leqslant p$ and $q_2 \leqslant p$.

For $k < \omega$ and $\rho < \omega_1$, the sets

$$\mathscr{D}_k = \{\langle n, s \rangle \in P : n \geqslant k\} \quad \text{and} \quad \mathscr{E}_\rho = \{\langle n, s \rangle \in P : s \backslash \rho \neq \emptyset\}$$

are cofinal in $\langle P, \leqslant \rangle$ since, for a given element $\langle n, s \rangle \in P$ there is an uncountable set $B' \subseteq B(n, s)$ such that $F(n, \sigma) = F_j$ for all $j \leqslant m = \max(n, k)$ and for all $\sigma \in B'$, and so $\langle n, s \rangle \leqslant \langle m, s \cup \{\alpha\} \rangle \in \mathscr{D}_k \cap \mathscr{E}_\rho$ for $\alpha \in B'$ and $\alpha > \rho$. By MA_{ω_1} there is an ideal $\mathscr{T} \subseteq P$ which has a non-empty intersection with all the $\mathscr{D}_k$ and $\mathscr{E}_\rho$. Since $\mathscr{T}$ has non-empty intersection with all $\mathscr{E}_\rho$, it follows that $B' = \bigcup\{s : \langle n, s \rangle \in \mathscr{T}$ for some $n\}$ is uncountable.

For $k < \omega$ choose any $\langle n, s \rangle \in \mathscr{D}_k \cap \mathscr{T}$ and define

$$A'_k = X_k \cap \bigcap \{A(k, \sigma) : \sigma \in s\}.$$

Note that this definition of A'_k does not depend upon the particular choice of $\langle n, s \rangle$. For, if $\langle m, t\} \in \mathscr{D}_k \cap \mathscr{T}$, then there is a common upper bound of these two elements $\langle l, r \rangle \in \mathscr{D}_k \cap \mathscr{T}$, and then

$$X_k \cap \bigcap \{A(k, \sigma) : \sigma \in s\} = X_k \cap \bigcap \{A(k, \sigma) : \sigma \in t\} = X_k \cap \bigcap \{A(k, \sigma) : \sigma \in r\}$$

Now for any $k < \omega$ and $\rho \in B'$, there is some $\langle n, s \rangle \in \mathscr{T}$ such that $n \geqslant k$ and $\rho \in s$. Since $A'_k \subseteq A(k, \rho)$ and $i(k, \rho) = 1$, it follows that $[A'_k]^2 \otimes [\{\rho\}]^1 \subseteq K_1$. □

We now conclude the proof that (1.6) holds in $\mathscr{M}$.

If there are $A, B \subseteq \omega_1$ such that $[A]^2 \otimes [B]^1 \subseteq K_0$, then the result follows immediately from Lemma 3.2. Therefore, we may assume that

$$[A]^2 \otimes [B]^1 \nsubseteq K_0 \tag{3.12}$$

whenever $A, B \subseteq \omega_1$. Then by Lemma 3.4 there are sets A_n $(n < \omega)$ and B such that $A_0 < A_1 < \cdots < B$ and such that

$$[A_m]^2 \otimes [A_n \cup B]^1 \subseteq K_1 \tag{3.13}$$

holds for $m < n < \omega$. By Lemma 3.3 (with $m = 2$), we can assume that either (i) there are infinite subsets $A'_i \subseteq A_i$ $(i < \omega)$ such that $\{a_m, a_n, b\} \in K_0$ whenever $a_m \in A'_m$, $a_n \in A'_n$, $b \in B$ and $m < n$, or (ii) there are $H_0 \in [A_0]^2$, $H_1 \in [A_1]^2$ such that $\{h_0, h_1, b\} \in K_1$ whenever $h_i \in H_i$ $(i < 2)$ and $b \in B$. If (i) holds then we contradict (3.12) by choosing a set A such that $|A \cap A_i| = 1$ for all $i < \omega$. So we may suppose that (ii) holds. But in this case, since (3.13) holds, we have that $[Y]^3 \subseteq K_1$, where Y is a 4-element set such that $H_0 \subseteq Y$, $|Y \cap H_1| = 1$ and $|Y \cap B| = 1$, and so (3.2) holds. □

References

[1] J. Baumgartner and A. Hajnal, A proof (involving Martin's Axiom) of a partition relation, Fund. Math. 78 (1973) 193–203.
[2] G. Fodor, Ein Bermerkung zur Theorie der regressiven Functionen, Acta Sci. Math. 17 (1956) 139–142.
[3] A. Hajnal, A negative partition relation, Proc. Nat. Acad. Sci. USA 68 (1971) 142–144.
[4] K. Kunen, Set Theory: Studies in Logic and the Foundations of Mathematics, Vol. 102 (North-Holland, Amsterdam, 1980).
[5] E.C. Milner and K. Prikry, A partition theorem for triples, Proc. Amer. Math. Soc. 97 (1986) 488–494.
[6] F.P. Ramsey, On a problem in formal logic, Proc. London Math. Soc. (2) 30 (1930) 264–286.
[7] E. Specker, Teilmengen von Mengen mit Relationen, Comment. Math. Helv. 31 (1957) 302–314.
[8] S. Todorčević, Forcing positive partition relations, Trans. Amer. Math. Soc. 280 (1983) 703–720.

Discrete Mathematics 95 (1991) 193–219
North-Holland

Some relations between analytic and geometric properties of infinite graphs

Bojan Mohar*

Department of Mathematics, University of Ljubljana, Jadranska 19, 61111 Ljubljana, Yugoslavia

Received 9 October 1989
Revised 5 February 1990

Abstract

Mohar, B., Some relations between analytic and geometric properties of infinite graphs, Discrete Mathematics 95 (1991) 193–219.

For locally finite infinite graphs the following analytic invariants and properties are considered: the spectrum of the transition and difference Laplacian matrix, amenability and the Kazhdan property (T). They are related to several geometric invariants, such as the isoperimetric number, growth, the structure of the space of ends, etc. Usually, only the global behaviour of invariants is important. It is shown how each of the above properties has its 'essential' counterpart, e.g. the essential isoperimetric number, the essential spectrum, the essential maximum degree, etc. These invariants do not change if we add or delete finitely many edges in the graph.

1. Introduction

The main aim of this paper is to bring to attention some properties and invariants of infinite locally finite graphs (with possible loops and multiple edges). Several relations between some of these invariants are derived. We also introduce, for each invariant considered here, its essential counterpart, which is a similar quantity but with the property that it does not change if we add, or delete finitely many edges from the graph. These 'essential' invariants describe the global properties of graphs. If a finite piece of a graph looks arbitrarily ugly, it does not affect the essential properties.

The considered invariants are classified as analytic (algebraic), or combinatorial (geometric). In the first class we count the spectrum of several matrices associated to graphs, the amenability, and the Kazhdan property (T). The last two of these were originally introduced for locally compact topological groups, and it is

* Supported in part by the Research Council of Slovenia, Yugoslavia.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

suggested here how to define them for arbitrary graphs, so that they coincide on Cayley graphs of finitely generated discrete groups with the classical parameters. The related combinatorial invariants are the isoperimetric constants of graphs, some numbers based on vertex degrees and the growth of graphs, and the structure of the space of ends of graphs.

The paper intends to be something between introductory, expository, and research. We believe that some ideas presented here are worthwhile deeper investigation not only because of their applications in graph theory but also because of their close relationship with some other areas of mathematics, e.g. differential geometry, representation theory, random walks, stochastic processes, etc.

2. Some geometric invariants

Graphs in this paper are undirected and locally finite.

Graphs G and H are said to be *equivalent* (*by finite*), or *almost isomorphic*, if there is a bijection $\psi: V(G) \to V(H)$ which maps adjacent, or non-adjacent pairs of vertices of G to adjacent and non-adjacent pairs, respectively, in H, with finitely many exceptions. Equivalently, H can be obtained from G by first deleting and then adding finitely many edges, up to an isomorphism.

We shall denote by $\Delta(G)$ and $\delta(G)$ the maximal and minimal degrees of vertices in G, respectively. They have their 'essential' counterparts, the *essential maximal degree* $\Delta'(G)$ and the *essential minimal degree* $\delta'(G)$, which are defined by

$$\Delta'(G) := \inf\{\Delta(H) \mid H \text{ equivalent to } G\}$$

and

$$\delta'(G) := \sup\{\delta(H) \mid H \text{ equivalent to } G\}.$$

It may happen that $\Delta'(G) = \infty$ and $\delta'(G) = \infty$. It is clear that $\Delta'(G)$ is equal to the minimal value d for which there are only finitely many vertices of G with degree $> d$. Similarly, $\delta'(G)$ is the minimal d with infinitely many vertices of degree d. It is obvious that $\delta'(G) \leqslant \Delta'(G)$. These two quantities are defined in such a way that no finite perturbation of G changes their values, i.e., if G and H are equivalent then $\Delta'(G) = \Delta'(H)$ and $\delta'(G) = \delta'(H)$. Graph invariants with this property are said to be *essential*. So, $\Delta'(G)$ and $\delta'(G)$ are first examples of essential graph invariants we met.

Let $v \in V(G)$. Denote by $B_n(v)$ the set of all vertices of G at distance at most n from v (the *ball* of radius n centered at v), and let $b_n(v) := |B_n(v)|$. If necessary to expose the graph, we write $b_n(G, v)$, or $B_n(G, v)$. The graph G has *exponential growth* (from the vertex v) if $b_n(v) \geqslant Cq^n$ for some constants $C > 0$, $q > 1$, and each $n \geqslant 0$. It has *polynomial growth* if $b_n(v) \leqslant p(n)$ where $p(\ )$ is a polynomial. G has *subexponential growth* if it does not grow exponentially.

Note that the type of the growth is independent of v if G is connected. Let

$$\varepsilon(G, v) := \limsup_{n\to\infty} (b_n(G, v))^{1/n}$$

and

$$\tau(G, v) := \liminf_{n\to\infty} (b_n(G, v))^{1/n}.$$

If v and u are vertices in the same component of G then $b_n(u) \leqslant b_{n+d}(v) \leqslant b_{n+2d}(u)$ where $d = \mathrm{dist}(u, v)$. It follows that $\epsilon(G, v)$ and $\tau(G, v)$ are constant on each component of the graph. Define

$$\varepsilon(G) := \sup\{\varepsilon(G, v) \mid v \in V(G)\}$$

and

$$\tau(G) := \sup\{\tau(G, v) \mid v \in V(G)\}.$$

By the above conclusion, the suprema may be taken on representatives of each of the connected components only.

Lemma 2.1. *If e is an edge of the graph G then $\varepsilon(G-e) = \varepsilon(G)$ and $\tau(G-e) \leqslant \tau(G)$. If e is not a bridge then $\tau(G-e) = \tau(G)$.*

Proof. The proof of $\varepsilon(G-e) = \varepsilon(G)$ can be found in [13]. To verify the relations for τ, let $e = uv$. Clearly, $\tau(G, v) \geqslant \tau(G-e, v)$ and $\tau(G, v) = \tau(G, u) \geqslant \tau(G-e, u)$. Consequently, $\tau(G) \geqslant \tau(G-e)$. Assume now that e is not a bridge. Then u and v belong to the same component of $G-e$. Let d be the distance between u and v in $G-e$. Clearly, $b_n(G, v) \leqslant b_{n+d}(G-e, v)$. Therefore $\tau(G, v) \leqslant \tau(G-e, v)$. The same holds for u instead of v. This implies that $\tau(G) \leqslant \tau(G-e)$, so by the converse inequality proved above, $\tau(G) = \tau(G-e)$. □

Note. It may happen that $\tau(G-e) < \tau(G)$. There are also cases where $\tau(G) > 1$ and $\tau(G-e) = 1$. Specific examples are not hard to construct.

Corollary 2.2. *The numbers $\varepsilon(G)$ and $\tau'(G) := \inf\{\tau(H) \mid H$ equivalent to $G\}$ are essential numerical graph invariants.*

Proof. ϵ is essential by Lemma 2.1, since addition or deletion of edges does not affect its value. On the other hand, τ' is essential since for G and G' equivalent, the set of graphs which are equivalent to G is equal to the equivalence class of G'. □

Proposition 2.3. *Let G be a graph and $v \in V(G)$. Then $\tau(G, v) > 1$ if and only if G has exponential growth at v.*

Proof. First, if $b_n(v) \geqslant Cq^n$, $q > 1$, $C > 0$, then clearly $\tau(G) \geqslant \tau(G, v) \geqslant q > 1$. Conversely, if $\tau(G) > 1$, pick $v \in V(G)$ such that $\tau(G, v) > 1$. This imples that for an arbitrarily small $\alpha > 0$,

$$b_n(v)^{1/n} \geqslant \tau(G, v) - \alpha \tag{2.1}$$

for all but finitely many values of n. Take $\alpha := \frac{1}{2}(\tau(G, v) - 1)$ and $q := \tau(G, v) - \alpha > 1$. By (2.1), $b_n(v) \geqslant q^n$ for all but finitely many n, thus $b_n(v) \geqslant Cq^n$ for some small enough $C > 0$ and each n. □

The *isoperimetric number* $h(G)$ of G is the number

$$h(G) := \inf\left\{\frac{|\delta X|}{|X|} \,\middle|\, |X| < \infty,\ X \neq \emptyset,\ X \subset V(G)\right\}$$

where $\delta X = \{e \in E(G) \mid e = vu,\ v \in X,\ u \notin X\}$. Sometimes we will have various graphs on the same vertex set. To specify in which graph G the coboundary δX is taken we use the subscript, e.g. $\delta_G X$. The *essential isoperimetric number* of G is

$$h'(G) := \sup\{h(G') \mid G' \text{ equivalent to } G\}.$$

It is clear that $h'(G) \geqslant h(G)$. It is also immediate that $h(G) \leqslant \delta(G)$ and $h'(G) \leqslant \delta'(G)$. However, these inequalities need not be very tight since there are easily constructed examples where $h'(G) = 0$ and $\delta'(G) = \infty$. On the other hand also

$$h'(G) \leqslant \Delta'(G) - 2. \tag{2.2}$$

This can be shown as follows: Let G' be equivalent to G. Then it has only finitely many vertices of degree greater than $\Delta'(G)$. Therefore one can find an arbitrarily large finite $X \subset V(G')$ such that all vertices in X have degree at most $\Delta'(G)$ in G', and the subgraph $\langle X \rangle$ of G' induced on X is connected. Then, in G',

$$|\delta X| \leqslant \Delta'(G) \cdot |X| - 2\,|E(\langle X \rangle)| \leqslant \Delta'(G)\,|X| - 2(|X| - 1),$$

which easily implies (2.2). Notice that this bound is best possible since for the Δ-regular tree T_Δ, $h(T_\Delta) = h'(T_\Delta) = \Delta - 2$.

Theorem 2.4. *Let G be a connected locally finite graph with $\Delta'(G) < \infty$. Then*

$$h(G) \leqslant h'(G) \leqslant \Delta'(G)\,\frac{\tau(G) - 1}{\tau(G) + 1}. \tag{2.3}$$

Proof. The first inequality is obvious. By Lemma 2.1, $\tau(G)$ does not change by adding finitely many new edges to G since G is connected. Thus we may assume that $h(G)$ is arbitrarily close to $h'(G)$, and so it is sufficient to show that

$$h(G) \leqslant \Delta'(G) \cdot \frac{\tau(G) - 1}{\tau(G) + 1}. \tag{2.4}$$

Choose a vertex $v \in V(G)$ and denote by $s_n := b_n(v) - b_{n-1}(v)$. For n large enough all vertices at distance n from v have degree at most $\Delta'(G)$, thus

$$\Delta'(G)s_{n+1} \geqslant |\delta B_{n+1}(v)| + |\delta B_n(v)| \geqslant h(G)(b_{n+1} + b_n) = h(G)(2b_n + s_{n+1}).$$

It follows that $(\Delta'(G) - h(G))s_{n+1} \geqslant 2h(G)b_n$ and hence

$$b_{n+1} = b_n + s_{n+1} \geqslant b_n \frac{\Delta'(G) + h(G)}{\Delta'(G) - h(G)}$$

for each n large enough. Therefore

$$b_n \geqslant C \cdot \left(\frac{\Delta'(G) + h(G)}{\Delta'(G) - h(G)}\right)^n,$$

for some $C > 0$ and each $n \geqslant 1$. So $\tau(G) \geqslant (\Delta'(G) + h(G))/(\Delta'(G) - h(G))$. This is equivalent to (2.4), so we are done. □

Corollary 2.5. *If $\Delta'(G) < \infty$ and $h'(G) > 0$ then $\tau(G) > 1$, so G has exponential growth in some of its components.*

Note 2.6. There are graphs with exponential growth and with $h'(G) = 0$. For example, Cayley graphs of some soluble groups which are not nilpotent by finite are known to grow exponentially but having $h'(G) = 0$, see [18]. On the other hand, there are graphs with $\Delta'(G) = \infty$ and $h'(G) > 0$ and having polynomial growth. Take, for example, the graphs having as vertices all integers and between i and $i + 1$ having $|i| + 1$ parallel edges, $-\infty < i < +\infty$. This graph is easily seen to have $h(G) = h'(G) = 1$, but it grows linearly. If we want simple graphs, we may add a vertex of degree two on each edge, or replace each vertex i by a clique of order $|i| + 1$, and between any two consecutive cliques put a complete join. All of this shows that Corollary 2.5 is best possible.

There are other possibilities to define isoperimetric constants of graphs, by taking other distances in graphs. Let us introduce only the *transition isoperimetric number* $s(G)$ of a graph G. If G is a graph with no edges then $s(G) := 0$, and otherwise

$$s(G) := \inf\left\{\frac{|\delta X|}{S(X)} \;\middle|\; X \subset V(G),\ X \neq \emptyset,\ |X| < \infty\right\}$$

where $S(X) := \sum_{v \in X} \deg(v)$ is the sum of the degrees of the vertices in X. Clearly, $0 \leqslant s(G) < 1$. In contrast to $h(G)$, the transition isoperimetric number does not properly measure the connectivity of the graph since it may happen that adding an edge to a graph strictly decreases its value. It is easily seen that $\delta(G)s(G) \leqslant h(G) \leqslant \Delta(G)s(G)$. Therefore $s(G)$ has advantage over $h(G)$ in case $\Delta(G) = \infty$ only. The *essential transition isoperimetric number* $s'(G)$ is equal to the supremum of all $s(G')$, where G' is equivalent by finite to G. A similar result as Theorem 2.4 holds for $s(G)$.

Theorem 2.7. *Let G be a connected locally finite graph. Then:*

(a) *Regardless of the degrees in G*

$$\tau(G) \geqslant \sqrt{\frac{1+s'(G)}{1-s'(G)}}, \quad i.e. \quad s(G) \leqslant s'(G) \leqslant \frac{\tau^2(G)-1}{\tau^2(G)+1}. \tag{2.5}$$

(b) *For a vertex $v \in V(G)$ let D_n be the maximal degree of vertices at distance n from v. If $\limsup_{n\to\infty} D_n^{1/n} = 1$ then*

$$\tau(G) \geqslant \frac{1+s'(G)}{1-s'(G)}, \quad i.e. \quad s(G) \leqslant s'(G) \leqslant \frac{\tau(G)-1}{\tau(G)+1}. \tag{2.6}$$

(c) *If there is a constant $M < \infty$ such that each vertex of G is contained in at most M edge-disjoint cycles, then* (2.6) *holds.*

Proof. If $s'(G) = 0$ then all the statements of the theorem are clear. Thus we assume that $s'(G) > 0$. Choose an arbitrarily small $\eta > 0$, $\eta < s'(G)$. There is a graph G_1 equivalent to G such that $s(G_1) \geqslant s'(G) - \eta > 0$. G_1 has finitely many components and all of them are infinite. It is easily seen that one of the components, say G_2, has $s(G_2) = s(G_1)$. We shall prove that the claimed inequalities (2.5) and (2.6) hold in G_2, for $s(G_2)$ in place of $s'(G)$. Since $\tau(G) \geqslant \tau(G_2)$ and $s(G_2)$ is arbitrarily close to $s'(G)$, this will be enough. To make the notation easier we will do all the calculations with the graph G (instead of G_2), in particular we assume henceforth that G is connected.

Choose a vertex $v \in V(G)$. Let $S_n := B_n(v) \setminus B_{n-1}(v)$. Then

$$S(B_n) - S(B_{n-1}) = S(S_n) \geqslant |\delta B_n| + |\delta B_{n-1}| \geqslant s(G)(S(B_n) + S(B_{n-1}))$$

and so $S(B_n) \geqslant S(B_{n-1})(1+s(G))/(1-s(G))$. It follows that for some constant $C > 0$

$$S(B_n) \geqslant C \cdot \left(\frac{1+s(G)}{1-s(G)}\right)^n. \tag{2.7}$$

We now consider cases (a), (b), and (c).

(a) Since each edge with an end in B_n has its other end in B_{n+1},

$$S(B_n) \leqslant |B_n| \cdot |B_{n+1}| \leqslant |B_{n+1}|^2.$$

Combining this with (2.7) gives the lower bound (2.5) on $\tau(G)$.

(b) Choose any $\varepsilon > 0$. By the assumption on D_n, there is n_0 such that for each $n \geqslant n_0$, $D_n < (1+\varepsilon)^n$. Let $n_1 \geqslant n_0$ be such that $\varepsilon(1+\varepsilon)^{n_1} \geqslant S(B_{n_0})$. Then for each $n > n_1$

$$\begin{aligned} S(B_n) &< (1+\varepsilon)^n |S_n| + \cdots + (1+\varepsilon)^{n_0} |S_{n_0}| + S(B_{n_0}) \\ &\leqslant (1+\varepsilon)^n |S_n| + \cdots + (1+\varepsilon)\cdot(1+\varepsilon)^{n_1} |S_{n_1}| + \cdots + (1+\varepsilon)^{n_0} |S_{n_0}| \\ &\leqslant (1+\varepsilon)^n (|S_n| + \cdots + |S_{n_0}|) \leqslant (1+\varepsilon)^n b_n. \end{aligned}$$

Now (2.7) implies that

$$(1+\varepsilon)^n b_n \geqslant C \cdot \left(\frac{1+s(G)}{1-s(G)}\right)^n.$$

Since ϵ was arbitrary, this gives the required lower bound on $\tau(G)$.

(c) Let G_n be the subgraph of G induced on $B_n(v)$. Take a cycle C_1 in G_n and remove its edges. Repeat this by choosing another, edge-disjoint cycle C_2, remove its edges, etc. until an acyclic graph is obtained. By our condition, the total number of removed edges is at most $M\,|B_n|$, and therefore

$$|E(G_n)| \leqslant M\,|B_n| + (|B_n|-1) < (M+1)\,|B_n|. \tag{2.8}$$

We also have

$$S(B_n) \leqslant 2\,|E(G_{n+1})| \tag{2.9}$$

and now a combination of (2.7), (2.8), and (2.9) gives

$$|B_{n+1}| \geqslant \frac{C}{2(M+1)}\left(\frac{1+s(G)}{1-s(G)}\right)^n$$

which implies (2.6). □

Corollary 2.8. *If $s'(G) > 0$ then $\tau(G) > 1$, so G has exponential growth in some of its components.*

3. The spectrum

With a graph we may associate several matrices. For our purpose, only the adjacency, the transition and the Laplacian matrix will be introduced. These matrices have their rows and columns indexed by the vertices $V(G)$ of the graph. No order of $V(G)$ is assumed, so a matrix in fact means a whole class of permutation similar matrices with ordered rows and columns. Let $\ell^2(V)$ denote the Hilbert space of complex vectors $x = (x_v)_{v \in V(G)}$ with coordinates indexed by $V = V(G)$, and such that

$$\|x\|^2 = \sum_{v \in V} |x_v|^2 < \infty.$$

The inner product in $\ell^2(V)$ is defined $\langle x, y\rangle = \sum_{v \in V} x_v \bar{y}_v$. Our matrices act naturally on $\ell^2(V)$ as linear operators.

The *adjacency matrix* $A(G) = [a_{uv}]$ of the graph G has entries a_{uv} equal to the number of edges between vertices u and v. If $D(G)$ denotes the diagonal matrix with diagonal entries containing vertex degrees, $d_{uu} = \deg(u)$, $u \in V(G)$, then the matrix $L(G) := D(G) - A(G)$ is called the *difference Laplacian matrix* of G. The *transition matrix*, $P(G) = [p_{uv}]$ has entries $p_{uv} = a_{uv}/\deg(u)$. This one is assumed to

act on the Hilbert space $\ell^2_d(V)$ where the inner product is given by

$$\langle x, y\rangle_d = \sum_{v\in V} x_v \bar{y}_v \deg(v)$$

while the adjacency and the Laplacian matrix act on $\ell^2(V)$. With this convention all three matrices define symmetric linear operators on the corresponding Hilbert spaces. It will be assumed, whenever speaking about the Laplacian or the adjacency matrix, that the graphs have bounded degrees, i.e., $\Delta(G) < \infty$. This will not be required in case of the transition operator. Under this assumption, $A(G)$, $L(G)$ and $P(G)$ are everywhere defined and self-adjoint linear operators on the corresponding Hilbert space. Their spectra are real. If we denote the spectrum of $A(G)$, $L(G)$ and $P(G)$ by $\sigma_A(G)$, $\sigma_L(G)$, and $\sigma_P(G)$, respectively, then it is well known (cf., e.g. [15]) that

$$\sigma_A(G) \subseteq [-\Delta, \Delta], \qquad \sigma_L(G) \subseteq [0, 2\Delta], \quad \text{and} \quad \sigma_P(G) \subseteq [-1, 1],$$

where $\Delta = \Delta(G)$. The reader is referred to [15] for more details about the spectra of infinite graphs. Let us add that in each case, the spectrum is the *approximate point spectrum* in the corresponding Hilbert space; for example, $\lambda \in \sigma_L(G)$ if and only if there is a sequence of unit vectors $x^{(n)}$ such that

$$\|L(G)x^{(n)} - \lambda x^{(n)}\| \to 0 \quad \text{as } n \to \infty.$$

The vectors $x^{(n)}$ may be assumed to have only real entries $x_v^{(n)}$.

The introduced spectra are closely related if G is regular. It is easily seen that

$$k \cdot \sigma_P(G) = \sigma_A(G) = k - \sigma_L(G)$$

where k is the valency and the equalities hold as equalities between sets.

For us there are two particularly important numerical invariants based on the spectrum. Let us denote by

$$\lambda_1(G) := \inf \sigma_L(G)$$

and

$$\rho_1(G) := \sup \sigma_P(G).$$

They can also be expressed as

$$\lambda_1(G) = \inf\left\{ \frac{\langle L(G)x, x\rangle}{\langle x, x\rangle} \,\middle|\, x \in \ell^2(V), x \neq 0 \right\} \tag{3.1}$$

and

$$\rho_1(G) = \sup\left\{ \frac{\langle P(G)x, x\rangle_d}{\langle x, x\rangle_d} \,\middle|\, x \in \ell^2_d(V), x \neq 0 \right\}. \tag{3.2}$$

The 'essential' invariant corresponding to the spectrum is the *essential spectrum* of $A(G)$, $L(G)$, or $P(G)$, respectively. The essential spectrum of a self-adjoint linear operator B on a Hilbert space consists of all those elements from the

spectrum of B which are not isolated eigenvalues of finite multiplicity. It is well known (cf., e.g., [9, Section X.2]) that these are exactly those $\lambda \in \sigma(B)$ which remain in the spectrum if B is changed by any compact perturbation, i.e. $\lambda \in \sigma(B+K)$ for every compact K. It follows that the essential spectrum is an essential graph invariant since equivalence by finite perturbes the graphs matrices in finitely many entries only and thus presents only compact perturbations. The elements of the essential spectrum are also characterized as the approximate eigenvalues of infinite multiplicity [5] (cf. also [9]). This means that λ is in the essential spectrum of B if and only if there are *pairwise orthogonal* unit vectors $x^{(n)}$, $n = 1, 2, 3, \ldots$, such that

$$\|Bx^{(n)} - \lambda x^{(n)}\| \to 0 \quad \text{as } n \to \infty.$$

We shall denote by $\lambda_1'(G)$ the infimum of the essential spectrum of $L(G)$, and by $\rho_1'(G)$ the supremum of the essential spectrum of $P(G)$. Clearly, $\lambda_1(G) \leqslant \lambda_1'(G)$ and $\rho_1(G) \geqslant \rho_1'(G)$. Later we shall need the following inequality:

$$\lambda_1'(G) \leqslant \frac{\Delta'(G) + \delta'(G)}{2} - 1 < \Delta'(G) \tag{3.3}$$

if $\delta'(G) \geqslant 1$. The proof goes as follows: There are infinitely many pairs u_i, v_i of adjacent vertices, such that $\deg(u_i) \leqslant \delta'(G)$ and $\deg(v_i) \leqslant \Delta'(G)$. Let $y^{(1)}, y^{(2)}, \ldots$ be their characteristic vectors (i.e., $y^{(i)}$ has the coordinate of v_i and u_i equal to $1/\sqrt{2}$, all others 0). Then

$$\langle L(G)y^{(i)}, y^{(i)}\rangle \leqslant \frac{\deg(u_i) + \deg(v_i) - 2}{2} \leqslant \frac{\Delta'(G) + \delta'(G)}{2} - 1.$$

Since this happens for infinitely many pairwise orthogonal unit vectors $y^{(i)}$, an element of the essential spectrum must be $\leqslant (\Delta'(G) + \delta'(G))/2 - 1$. Since we were not able to find this last statement mentioned in the available literature, we give a short proof of it. We have to show that if $\langle Lx^{(i)}, x^{(i)}\rangle \leqslant t$ for infinitely many pairwise orthogonal unit vectors $x^{(i)}$ then there is some $\lambda \leqslant t$ in the essential spectrum of L. It is known [22] that for some compact symmetric operator K, the spectrum of $L + K$ is equal to the essential spectrum of L. Notice that since K is compact, $\langle Kx^{(i)}, x^{(i)}\rangle \to 0$. Consequently

$$\begin{aligned}\lambda_1'(L) &= \lambda_1(L+K) = \inf\{\langle (L+K)x, x\rangle \mid \|x\| = 1\}\\ &\leqslant \inf\{\langle Lx^{(i)}, x^{(i)}\rangle + \langle Kx^{(i)}, x^{(i)}\rangle\} \leqslant t.\end{aligned}$$

Theorem 3.1. *Let G be a locally finite graph, and let $\Delta = \Delta(G) < \infty$, $\Delta' = \Delta'(G)$. Then*

$$h(G) \geqslant \lambda_1(G)\frac{\Delta^2}{\Delta^2 - \Delta - \lambda_1(G)} \geqslant \lambda_1(G) \tag{3.4}$$

and

$$h'(G) \geqslant \lambda_1'(G) \frac{\Delta'^2}{\Delta'^2 - \Delta' - \lambda_1'(G)} \geqslant \lambda_1'(G). \tag{3.5}$$

Proof. Let U be an arbitrary finite subset of $V = V(G)$. Define $x \in \ell^2(V)$ by setting $x_v = 1$ if $v \in U$, and otherwise $x_v = (1/\Delta)n_v$, where n_v is the number of edges from v to vertices in U. Denote by W the set of those vertices of $V \setminus U$ which have a neighbour in U. By (3.1),

$$\lambda_1(G)\, \|x\|^2 \leqslant \langle L(G)x, x \rangle. \tag{3.6}$$

It is easily seen that

$$\langle L(G)x, x \rangle = \sum_{uv \in E(G)} (x_u - x_v)^2 \tag{3.7}$$

and since $x_u - x_v$ is nonzero only if v or u lies in W, we get from (3.6) and (3.7):

$$\begin{aligned} \lambda_1\, \|x\|^2 = \lambda_1\Big(|U| + \sum_{v \in W} \frac{n_v^2}{\Delta^2}\Big) &\leqslant \sum_{vu \in E(G)} (x_u - x_v)^2 \\ &\leqslant \sum_{v \in W} \left(n_v\Big(1 - \frac{n_v}{\Delta}\Big)^2 + (\Delta - n_v)\Big(\frac{n_v}{\Delta}\Big)^2\right) \\ &= \sum_{v \in W} n_v - \sum_{v \in W} \frac{n_v^2}{\Delta} = |\delta U| - \frac{1}{\Delta} \sum_{v \in W} n_v^2. \end{aligned} \tag{3.8}$$

By rearranging this inequality and using the fact $n_v^2 \geqslant n_v$, one immediately gets

$$|\delta U| - \lambda_1\, |U| \geqslant \frac{\Delta + \lambda_1}{\Delta^2} \sum_{v \in W} n_v \geqslant \frac{\Delta + \lambda_1}{\Delta^2} |\delta U|$$

and it follows that

$$\frac{|\delta U|}{|U|} \geqslant \lambda_1 \frac{\Delta^2}{\Delta^2 - \Delta - \lambda_1}$$

which implies (3.4) since U was arbitrary.

The proof of (3.5) needs some more care. Fix an arbitrarily small $\varepsilon > 0$. First we shall prove that there exists a graph G_1, which can be obtained from G by adding finitely many edges, such that $\lambda_1(G_1) \geqslant \lambda_1'(G) - \varepsilon$.

The only elements in the Laplacian spectrum of G which are smaller than $\lambda_1' = \lambda_1'(G)$ are isolated eigenvalues of finite multiplicity. Let N be the span of the eigenspaces of all those eigenvalues of $L(G)$ which are smaller than $\lambda_1' - \varepsilon/2$, and let $n := \dim N < \infty$. Fix an orthonormal basis $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$ of N. It should be mentioned at once that for any $x \in \ell^2(V)$ which is orthogonal to N

$$\langle L(G)x, x \rangle \geqslant \Big(\lambda_1' - \frac{\varepsilon}{2}\Big)\, \|x\|^2. \tag{3.9}$$

Let U be a finite subset of $V(G)$ such that for each $i = 1, 2, \ldots, n$,

$$\sum_{u \in U} (x_u^{(i)})^2 > 1 - \delta, \text{ i.e., } \sum_{u \notin U} (x_u^{(i)})^2 < \delta, \tag{3.10}$$

where $\delta > 0$ is some very small number to be fixed later. It will depend on ε, n, and G.

Let G_1 be the graph which is obtained from G by adding, for each vertex $u \in U$, m edges joining u to m new distinct neighbours which lie out of U. Denote by $L_1 := L(G_1)$ its Laplacian matrix. We shall prove that $\lambda_1(G_1) \geqslant \lambda_1' - \varepsilon$, assuming m is large enough. Notice that m may depend on ε, n, and G.

Take an arbitrary $x \in \ell^2(V)$, $\|x\| = 1$. We shall prove that $\langle L_1 x, x \rangle \geqslant \lambda_1' - \varepsilon$ which will imply, by (3.1) that $\lambda_1(G_1) \geqslant \lambda_1' - \varepsilon$. It is easy to see that it may be assumed that $x_v \geqslant 0$ for each $v \in V(G)$ (apply, for example, (3.7)). Write now $x = y + z$ where y agrees with x on U, and z agrees with x at other coordinates.

If $\|y\|^2 \leqslant \delta$ then we do the following calculation: Let $x = p + q$ where $p \in N$ and $q \perp N$. By (3.9), $\langle Lq, q \rangle \geqslant (\lambda_1' - \varepsilon/2)\, \|q\|^2$ while $\langle Lp, p \rangle$ can be estimated as follows: If $p = \sum_{i=1}^n \alpha_i x^{(i)}$ then $\|p\|^2 = \sum_{i=1}^n |\alpha_i|^2$ and

$$|\alpha_i| = |\langle x, x^{(i)} \rangle| = \sum_{u \in U} x_u x_u^{(i)} + \sum_{v \notin U} x_v x_v^{(i)} < \|y\| + \|x\| \sqrt{\delta} \leqslant 2\sqrt{\delta}. \tag{3.11}$$

In the last inequality we applied the Cauchy–Schwartz inequality and (3.10). Consequently, $\|p\|^2 < 4n\delta = \varepsilon/2\Delta(G)$ if we choose $\delta := \varepsilon/8n\Delta(G)$. Therefore $\|q\|^2 = \|x\|^2 - \|p\|^2 > 1 - \varepsilon/2\Delta(G)$. Using this estimate, the fact that $\langle Lp, q \rangle = \langle Lq, p \rangle = 0$, and (3.3) we get

$$\begin{aligned} \langle L_1 x, x \rangle &\geqslant \langle Lx, x \rangle = \langle Lp, p \rangle + \langle Lq, q \rangle \\ &\geqslant \left(\lambda_1' - \frac{\varepsilon}{2}\right) \|q\|^2 > \left(\lambda_1' - \frac{\varepsilon}{2}\right)\left(1 - \frac{\varepsilon}{2\Delta(G)}\right) \geqslant \lambda_1' - \varepsilon. \end{aligned}$$

The other possibility to consider is the case when $\|y\|^2 > \delta = \varepsilon/8n\Delta(G)$. In this case let $B := L_1 - L(G)$ and let $F := E(G_1) \backslash E(G)$. If $z = p + q$ where $p = \sum_{i=1}^n \alpha_i x^{(i)} \in N$, $q \perp N$, then we see as in (3.11) that

$$|\alpha_i| = |\langle z, x^{(i)} \rangle| < \sqrt{\delta}\, \|z\|$$

and hence $\|p\|^2 < n\delta\, \|z\|^2$, so $\|q\|^2 > (1 - n\delta)\, \|z\|^2$. Therefore

$$\langle L_1 z, z \rangle \geqslant \langle Lz, z \rangle = \langle Lp, p \rangle + \langle Lq, q \rangle \geqslant \left(\lambda_1' - \frac{\varepsilon}{2}\right) \|q\|^2 > (\lambda_1' - \varepsilon)\, \|z\|^2. \tag{3.12}$$

Finally,

$$\begin{aligned}\langle L_1 x, x\rangle &= \langle L_1 y, y\rangle + 2\langle L_1 y, z\rangle + \langle L_1 z, z\rangle \\ &\geqslant (\lambda_1' - \varepsilon)\,\|z\|^2 + \langle By, y\rangle + 2\langle L_1 y, z\rangle \\ &= (\lambda_1' - \varepsilon)\,\|z\|^2 + \sum_{uv\in F}(y_u - y_v)^2 + 2\sum_{uv\in E(G_1)}(y_u - y_v)(z_u - z_v) \\ &\geqslant (\lambda_1' - \varepsilon)\,\|z\|^2 + m\sum_{u\in U} y_u^2 - 2\sum_{uv\in E(G_1)} y_u z_v \\ &\geqslant (\lambda_1' - \varepsilon)\,\|x\|^2 + (m - \lambda_1')\,\|y\|^2 - 2\sum_{uv\in F_1} y_u z_v\end{aligned}$$

where $F_1 = F \cup \delta_G U \subseteq E(G_1)$, and it is assumed that $u \in U$, $v \notin U$. It remains to show that

$$(m - \lambda_1')\,\|y\|^2 - 2\sum_{uv\in F_1} y_u z_v \geqslant 0. \tag{3.13}$$

By the Cauchy–Schwartz inequality and assuming that each edge $uv \in F_1$, $u \in U$, has different endvertex $v \notin U$, we estimate

$$\Big(\sum_{uv\in F_1} y_u z_v\Big)^2 \leqslant \sum_{uv\in F_1} |y_u|^2 \sum_{uv\in F_1} |z_v|^2 \leqslant (m + \Delta(G))\,\|y\|^2 \cdot (\Delta(G) + 1)\,\|z\|^2.$$

This inequality and the assumption $\|y\|^2 > \delta$ imply that

$$\begin{aligned}&(m - \lambda_1')\,\|y\|^2 - 2\sum_{uv\in F_1} y_u z_v \\ &\quad\geqslant (m - \lambda_1')\sqrt{\delta} - 2\sqrt{m + \Delta(G)}\,\sqrt{\Delta(G) + 1}\,\sqrt{1 - \delta}\,.\end{aligned} \tag{3.14}$$

Obviously, the last expression is $\geqslant 0$ if m is large enough. It should be added that m depends on λ_1', $\Delta(G)$ and $\delta = \delta(\varepsilon, n, \Delta(G))$ but not on x. This finally establishes (3.13), and thus we succeeded to show $\langle L_1 x, x\rangle \geqslant \lambda_1' - \varepsilon$, and consequently that $\lambda_1(G_1) \geqslant \lambda_1' - \varepsilon$.

It is clear that adding edges to a graph can not decrease the isoperimetric number. Therefore there exists a graph G_2 which can be obtained from G_1 by adding finitely many edges so that $h(G_2) \geqslant h'(G) - \varepsilon$. Clearly, also $\lambda_1(G_2) \geqslant \lambda_1(G_1) \geqslant \lambda_1' - \varepsilon$.

Let $X \subset V(G_2)$ be the set containing all vertices of degree greater than $\Delta'(G)$ and all their neighbours. Let G_3 be the graph obtained from G_2 by adding m parallel edges between any two vertices of X, where $m = \Delta'(G)\,|X|$.

Choose a finite $U_1 \subset V(G_3)$ such that $|\delta_{G_3} U_1|/|U_1| \leqslant h(G_3) + \varepsilon$. If $U_1 \cap X = \emptyset$ then let $U := U_1$, otherwise let $U := U_1 \cup X$. So U either contains X, or it is disjoint from X. Therefore no vertex adjacent to U in G_2 has degree greater than

$\Delta'(G)$. It is easily seen that in each case

$$\frac{|\delta_{G_2}U|}{|U|}=\frac{|\delta_{G_3}U|}{|U|}\leqslant\frac{|\delta_{G_3}U_1|}{|U_1|}\leqslant h(G_3)+\varepsilon\leqslant h'(G)+\varepsilon \tag{3.15}$$

since $m=\Delta'(G)\,|X|$.

Now we may repeat the first part of the proof for the graph G_3, using $\Delta'(G)$ instead of Δ. Notice that in (3.8) we need the fact that $n_v\leqslant\Delta'(G)$ for each vertex v adjacent to U. This completes the proof since $h(G_3)$ and $\lambda(G_3)$ are arbitrarily close to their essential values. □

It is interesting that $h(G)$ is also bounded above by a function of $\lambda_1(G)$. This inequality, given in Theorem 3.2, is called the *strong discrete Cheeger inequality*, from its close relation to a well-known Cheeger inequality [4] (cf. also [3]) from the theory of differential operators on Riemannian manifolds. A weaker version of (3.16) is proved in [13].

Theorem 3.2. *For an arbitrary locally finite graph G with bounded degrees,*

$$h(G)\leqslant\sqrt{\lambda_1(G)(2\Delta(G)-\lambda_1(G))} \tag{3.16}$$

and

$$h'(G)\leqslant\sqrt{\lambda_1'(G)(2\Delta'(G)-\lambda_1'(G))}\,. \tag{3.17}$$

Proof. The spectrum is closed, so $\lambda_1=\lambda_1(G)\in\sigma_L(G)$. For each small enough $\varepsilon>0$ there is a vector $x\in\ell^2(V(G))$ with finite support (only finitely many x_v are nonzero) and with $\|x\|=1$ such that

$$\langle Lx,x\rangle\leqslant\lambda_1+\varepsilon\leqslant\Delta(G). \tag{3.18}$$

If

$$\eta:=\sum_{uv\in E(G)}|x_u^2-x_v^2|$$

then it can be shown by the summation per partes (see, e.g., [13]) that

$$\eta\geqslant h(G). \tag{3.19}$$

On the other hand

$$\begin{aligned}\eta^2&=\Big(\sum_{uv\in E}|x_u^2-x_v^2|\Big)^2=\Big(\sum_{uv\in E}|x_u+x_v|\,|x_u-x_v|\Big)^2\\&\leqslant\sum_{uv\in E}(x_u+x_v)^2\cdot\sum_{uv\in E}(x_u-x_v)^2\\&=\sum_{uv\in E}(2x_u^2+2x_v^2-(x_u-x_v)^2)\cdot\langle Lx,x\rangle\\&=\Big(2\sum_{v\in V}\deg(v)x_v^2-\langle Lx,x\rangle\Big)\cdot\langle Lx,x\rangle\\&\leqslant(2\Delta(G)\,\|x\|^2-\langle Lx,x\rangle)\langle Lx,x\rangle\\&\leqslant 2\Delta(G)(\lambda_1+\varepsilon)-(\lambda_1+\varepsilon)^2.\end{aligned} \tag{3.20}$$

In the first inequality we used the Cauchy–Schwartz inequality, while in the last one we needed (3.18). Since ε was arbitrarily small, it follows that $h^2(G) \leqslant \eta^2 \leqslant 2\Delta(G)\lambda_1 - \lambda_1^2$ and the proof of (3.16) is done.

To prove (3.17), let G_1 be a graph obtained from G by adding finitely many edges and such that $h(G_1) \geqslant h'(G) - \varepsilon$ where $\varepsilon > 0$ is arbitrarily small again. As we know from the proof of Theorem 3.1, there is a graph G_2 obtainable from G_1 by addition of finitely many edges such that $\lambda_1(G_2) \geqslant \lambda_1'(G) - \varepsilon$. Since $h(G_2) \geqslant h(G_1)$, we have $h(G_2) \geqslant h'(G) - \varepsilon$.

Let X be the set of vertices of G_2 having degree $> \Delta'(G)$. The set X is clearly finite. Since $\lambda_1' = \lambda_1'(G) = \lambda_1'(G_2)$ is in the essential spectrum, there are arbitrarily many unit vectors $x^{(i)}$, $i = 1, 2, \ldots$ such that $\langle L(G_2)x^{(i)}, x^{(i)}\rangle \leqslant \lambda_1' + \varepsilon$, and such that $\langle x^{(i)}, x^{(j)}\rangle = 0$ for $i \neq j$. Let us take $N = |X| + 1$ such vectors $x^{(i)}$. They are linearly independent, so any nontrivial linear combination is nonzero. On the other hand, there exist $\alpha_1, \alpha_2, \ldots, \alpha_N$ with some $\alpha_i \neq 0$ such that $\sum_{i=1}^N \alpha_i x^{(i)} =: x$ has coordinate $x_v = 0$ for each $v \in X$. Clearly, if we demand that $\sum \alpha_i^2 = 1$, then

$$\|x\|^2 = \left\langle \sum_{i=1}^N \alpha_i x^{(i)}, \sum_{i=1}^N \alpha_i x^{(i)} \right\rangle = \sum_{i=1}^N \alpha_i^2 = 1.$$

The vectors $x^{(i)}$ may also be assumed to satisfy $\|L(G_2)x^{(i)} - \lambda_1' x^{(i)}\| \leqslant \varepsilon/N$. Then

$$\begin{aligned}
\langle L(G_2)x, x\rangle &= \sum_{i=1}^N \sum_{j=1}^N \alpha_i\alpha_j \langle L(G_2)x^{(i)}, x^{(j)}\rangle \\
&= \sum_{i=1}^N \alpha_i^2 \langle \lambda_1' x^{(i)}, x^{(i)}\rangle + \sum_{i=1}^N \sum_{j=1}^N \alpha_i\alpha_j \langle L(G_2)x^{(i)} - \lambda_1' x^{(i)}, x^{(j)}\rangle \\
&\leqslant \lambda_1' + \sum_{i=1}^N \sum_{j=1}^N |\alpha_i\alpha_j| \, \|L(G_2)x^{(i)} - \lambda_1' x^{(i)}\| \cdot \|x^{(j)}\| \\
&\leqslant \lambda_1' + \frac{\varepsilon}{N} \sum_{i=1}^N \sum_{j=1}^N |\alpha_i\alpha_j| \leqslant \lambda_1' + \varepsilon.
\end{aligned}$$

Now we may carry out the same calculation as to obtain (3.19) and (3.20), this time for G_2. The inequality $\deg(v)x_v^2 \leqslant \Delta'(G)x_v^2$ is clearly satisfied, and the only remaning fact to verify is $\lambda_1' + \varepsilon \leqslant \Delta'(G)$. But we have shown this by (3.3) for any graph with $\delta'(G) \geqslant 1$. The remaining case $\delta'(G) = 0$ is unimportant since in this case $h'(G) = 0$ and $\lambda_1' = 0$, so (3.17) is trivial. □

There are similar relations between $s(G)$ and $\rho_1(G)$.

Theorem 3.3. *If G is a locally finite graph then*

$$s(G) \geqslant \frac{1 - \rho_1(G)}{1 - \nabla/\rho_1(G)} \tag{3.21}$$

where $\nabla = 1/\Delta(G) \geqslant 0$, and

$$s'(G) \geqslant 1 - \rho_1'(G). \tag{3.22}$$

Proof. Set $\rho_1 = \rho_1(G)$. Let A be a finite subset of $V(G)$. Define $x \in \ell^2(V)$ as follows: $x_v := 1$ if $v \in A$, and otherwise let $x_v := n_v/\rho_1 \deg(v)$, where n_v is the number of neighbours of v in the set A. Then a straightforward calculation gives

$$\langle x, x \rangle_d = S(A) + \sum_{v \notin A} \frac{n_v^2}{\rho_1^2 \deg(v)}$$

and

$$\langle P(G)x, x \rangle_d \geqslant S(A) - |\delta A| + \sum_{u \in A} \sum_{v \notin A} a_{uv} x_v$$

$$\geqslant S(A) - |\delta A| + 2 \sum_{v \notin A} x_v n_v.$$

Combining these inequalities with (3.2) gives

$$\rho_1 \langle x, x \rangle_d = \rho_1 S(A) + \sum_{v \notin A} \frac{n_v^2}{\rho_1 \deg(v)}$$

$$\geqslant \langle P(G)x, x \rangle_d \geqslant S(A) - |\delta A| + 2 \sum_{v \notin A} \frac{n_v^2}{\rho_1 \deg(v)}$$

and after rearranging

$$|\delta A| - S(A)(1 - \rho_1) \geqslant \sum_{v \notin A} \frac{n_v^2}{\rho_1 \deg(v)} \geqslant \frac{\nabla}{\rho_1} \sum_{v \notin A} n_v = \frac{\nabla}{\rho_1} |\delta A|.$$

Since A was arbitrary, this implies (3.21).

To prove (3.22), take G_1 to be a graph equivalent by finite to G and with $\rho_1(G_1) \leqslant \rho_1'(G) + \varepsilon$. Then, by (3.21)

$$s'(G) \geqslant s(G_1) \geqslant 1 - \rho_1(G_1) \geqslant 1 - \rho_1'(G) - \varepsilon.$$

This implies (3.22) since ε was arbitrarily small. □

Theorem 3.4. *If G is a locally finite graph then*

$$s(G) \leqslant \sqrt{1 - \rho_1(G)} \tag{3.23}$$

and

$$s'(G) < \sqrt{1 - \rho_1'(G)}. \tag{3.24}$$

Proof. The first inequality can be found in [13]. To prove (3.24), take any $\varepsilon > 0$, and let G_1 be a graph equivalent by finite to G with $s(G_1) \geqslant s'(G) - \varepsilon$. Then by (3.23)

$$s'(G) - \varepsilon \leqslant s(G_1) \leqslant \sqrt{1 - \rho_1(G_1)} \leqslant \sqrt{1 - \rho_1'(G)}.$$

The inequality (3.24) is now obvious. □

4. Amenability

In this section we shall shortly exhibit the notion of amenability of infinite graphs. This concept was classically introduced as a property of locally compact groups. We refer to [16, 17] for more information.

Let Γ be a locally compact group. By $L^\infty(\Gamma)$ we denote the set of all functions $\Gamma \to \mathbb{R}$ which are bounded a.e. with respect to the Haar measure $\lambda(.)$ on Γ. The group Γ is *amenable if* $L^\infty(\Gamma)$ has an *invariant mean*, i.e. a linear functional $m: L^\infty(\Gamma) \to \mathbb{R}$ which satisfies the following conditions:

(i) For $f \in L^\infty(\Gamma)$, if $f \geqslant 0$ (a.e.) then $m(f) \geqslant 0$.

(ii) $m(\chi_\Gamma) = 1$, where χ_Γ is the constant 1 function on Γ.

(iii) It is Γ-invariant, i.e., $m(g \cdot f) = m(f)$ for every $g \in \Gamma$. (Note that $(g \cdot f)(x) = f(g^{-1}x)$.)

By a theorem of Følner [6], the amenability is equivalent to the following condition:

(F) Given $\varepsilon > 0$ and a compact set $K \subseteq \Gamma$ there is a Borel set $U \subseteq \Gamma$ of positive finite measure $\lambda(U) < \infty$ such that

$$\lambda(kU \,\Delta\, U) < \varepsilon\lambda(U)$$

for all $k \in K$, where Δ denotes the symmetric difference of sets.

For a discrete group the Haar measure counts the number of elements of the set. Then it is easy to see that (F) is equivalent to

(F′) For every $\varepsilon > 0$ and finite $K \subseteq \Gamma$ one can find a finite set $U \subseteq \Gamma$ such that

$$|KU \,\Delta\, U| < \varepsilon\, |U|.$$

Lemma 4.1. *Let Γ be a group generated by a finite set of generators $S = \{g_1, \ldots, g_k\}$. Then Γ is amenable if and only if for every $\varepsilon > 0$ there is a finite set $U \subseteq \Gamma$ such that $|SU \setminus U| < \varepsilon |U|$.*

Proof. ($\Rightarrow$) Let $K := S \cup \{e\}$, where e is the identity in Γ. Then $KU \,\Delta\, U = SU \setminus U$. We are done by the amenability of Γ.

($\Leftarrow$) Choose $\varepsilon > 0$ and a finite $K \subseteq \Gamma$, $K = \{k_1, k_2, \ldots, k_t\}$. It may be assumed that $e \notin K$. If k_i can be written as a word of length $l(i)$ in terms of $S \cup S^{-1}$, let

$$\varepsilon_i := \frac{\varepsilon}{2t \cdot l(i)}$$

and let $U \subseteq \Gamma$ be a set for which $|SU \setminus U| < \varepsilon' |U|$, where $\varepsilon' = \min \varepsilon_i$. For each generator $g_j \in S$, $|g_j U \setminus U| \leqslant |SU \setminus U| < \varepsilon' |U|$. From this we also see that $|g_j U \cap U| > (1 - \varepsilon') |U|$, therefore also $|g_j^{-1} U \cap U| > (1 - \varepsilon') |U|$, so $|g_j^{-1} U \setminus U| < \varepsilon' |U|$. It follows that $|k_i U \setminus U| < l(i)\varepsilon' |U| \leqslant (\varepsilon/2t) |U|$. This implies that $|KU \setminus U| < (\varepsilon/2) |U|$. Since $|KU| \geqslant |U|$, we get $|KU \,\Delta\, U| < \varepsilon |U|$. $\square$

Corollary 4.2. *A finitely generated infinite group Γ is amenable if and only if the Cayley graph* $\mathrm{Cay}(\Gamma, S)$ *with respect to some (and hence for every) finite generating set S has the isoperimetric number* $h(\mathrm{Cay}(\Gamma, S)) = 0$.

By Corollary 4.2 it follows trivially that every finitely generated group with polynomial or subexponential growth is amenable. The converse is not true. For example, all soluble groups are amenable but some of them have exponential growth. Although finite graphs are out of our main concern, let us mention that every finite group is amenable (take $U = \Gamma$).

Corollary 4.2 was observed before by several authors. Its importance lies in the fact that it gives a possibility to extend the notion of amenability to graphs in such a way that a finitely generated group Γ is amenable if and only if it has an amenable Cayley graph. So, let us call a graph G *amenable* if $s(G) = 0$. We use $s(G) = 0$ instead of $h(G) = 0$ to cover the case of graphs with unbounded degrees.

The case of graphs with bounded degrees was investigated in detail by Gerl and we refer to his works [7–8] where several conditions equivalent to the amenability of graphs are derived.

The following result is a direct consequence of Corollary 2.8.

Proposition 4.3. *A graph with subexponential growth is amenable.*

At the end of this section we shall consider amenability of infinite covers of finite graphs. Let B be a given finite graph, and let $p : G \to B$ be a graph covering projection. Is there any relation between $h(G)$ and $h(B)$, or $\lambda_1(G)$ and $\lambda_1(B)$? It turns our that there are some obvious relations, but their converse depends on the fundamental groups of G and B. Let us mention that results of the same type as following here are known in the setting of Riemannian manifolds, cf. [1, 2].

Recall that a graph map $p : G \to B$ is a *covering projection* if it is a local isomorphism, i.e., the edges incident to any vertex $v \in V(G)$ are mapped bijectively onto the edges incident with $p(v)$. For each $b \in V(B)$, the set $p^{-1}(b) \subseteq V(G)$ is called the *fibre* of b. It can be shown that all fibres have equal cardinalities if B is connected (which we shall assume henceforth).

We need some more definitions concerning finite graphs. If K is a finite graph then we define its Laplacian and the transition matrix in the same way as we do in the infinite case. The Laplacian $L(K)$ is positive semidefinite and its smallest eigenvalue is $\lambda_0 = 0$ with a corresponding eigenvector $\mathbf{1}$ having all entries equal to 1. The second smallest eigenvalue of $L(K)$ plays the role of λ_1 in the infinite case, so we denote it by $\lambda_1(K)$. It is known that $\lambda_1(K) > 0$ if and only if K is connected. Similarly, let $\rho_1(K)$ be the second largest eigenvalue of $P(K)$, while the largest is always equal to 1 and has $\mathbf{1}$ as an eigenvector. Similarly, the isoperimetric numbers $h(G)$ and $s(G)$ of finite graphs are defined as for infinite graphs, as the minimum of the isoperimetric quotients $|\delta X|/|X|$ and $|\delta X|/S(X)$, respectively,

where in the first case X is limited by the condition that $|X| \leqslant \frac{1}{2}|V(G)|$, and for $s(G)$ we take only all those sets for which $|S(X)| \leqslant \frac{1}{2}|S(V(G))| = |E(G)|$.

Assume that B is a fixed connected finite graph. If G is a finite covering of B then the Laplacian spectrum $\sigma_L(G)$ contains $\sigma_L(B)$. See e.g. [14] for details. In particular, $\lambda_1(G) \leqslant \lambda_1(B)$. If G is infinite this is not necessarily true any more. It might happen that $\lambda_1(G) > \lambda_1(B)$. For example, the k-regular tree T_k covers any connected k-regular graph. It is well known that $\lambda_1(T_k) = k - 2\sqrt{k-1}$ but (finite) k-regular graphs may have their second smallest eigenvalue arbitrarily close to zero. Similarly, $h(G) \leqslant h(B)$ for finite covers G of B, while for G infinite it may happen that $h(G) > h(B)$. It is also possible that $h(G)$ is arbitrarily close to zero even for finite covers G of B. The next results make the situation clear.

Proposition 4.4. *Let $p: G \to B$ be a graph covering, where B is a finite connected graph, and G is an infinite graph. Pick a spanning tree T in B, and let H be the graph obtained from G by contracting each component of $p^{-1}(T)$ to a single point. Then*

$$h(G) \leqslant \frac{1}{|V(B)|} h(H) \leqslant 2(|E(B)| - |V(B)| + 1) \cdot h(G). \tag{4.1}$$

Proof. Choose an $\varepsilon > 0$. Let U be a finite subset of $V(H)$ such that $|\delta_H U| \leqslant (h(H) + \varepsilon)|U|$. For each $u \in V(H)$ let T_u be the set of vertices of G which are contracted to u, and let $\tilde{U} := \bigcup \{T_u \mid u \in U\}$. Then $|\tilde{U}| = |V(B)| \cdot |U|$ and $|\delta_G \tilde{U}| = |\delta_H U|$. Therefore

$$h(G)|\tilde{U}| \leqslant |\delta_G \tilde{U}| = |\delta_H U| \leqslant (h(H) + \varepsilon)|U| = \frac{|\tilde{U}|}{|V(B)|}(h(H) + \varepsilon)$$

which implies the first inequality of (4.1) since ε can be arbitrarily small.

Take now a finite $X \subset V(G)$ such that $|\delta_G X| \leqslant (h(G) + \varepsilon)|X|$. Then let $X_1 \subset V(H)$ be the set of all $u \in V(H)$ for which $T_u \cap X \neq \emptyset$. Clearly, $|X_1| \geqslant |X|/|V(B)|$. For each edge in $\delta_H X_1$ with a vertex $u \in X_1$ and the other end outside X_1, there is at least one edge in $\delta_G X$ which has one of its ends in T_u. This is clear if $T_u \subseteq X$, and if $T_u \nsubseteq X$ then such an edge lies in T_u. Since $\deg_H(U) \leqslant 2|E(B) \setminus E(T)|$, it follows that $|\delta_H X_1| \leqslant 2|\delta_G X| \cdot |E(B) \setminus E(T)|$. Consequently,

$$\begin{aligned} h(H) &\leqslant \frac{|\delta_H X_1|}{|X_1|} \leqslant \frac{2|V(B)| \cdot |\delta_G X| \cdot |E(B) \setminus E(T)|}{|X|} \\ &\leqslant 2(h(G) + \varepsilon)|V(B)| \cdot |E(B) \setminus E(T)|. \end{aligned}$$

which implies the second inequality of (4.1). □

Proposition 4.4 by itself is not a very surprising result, but it turns into a very useful result if we give another interpretation to the graph H. Denote by $\pi_1(.)$ the fundamental group of a graph. Fixing the spanning tree T of B, and a base vertex

$b \in V(B)$, $\pi_1(B)$ is generated by the *fundamental cycles* (with respect to T). These are defined as follows: For each edge $e \in E(B) \setminus E(T)$ there is a unique cycle in $T + e$. Let γ_e be the closed walk in $T + e$ which starts at our chosen base vertex b, leads to the cycle, goes once around it (in one of the possible directions), and returns to b. The corresponding elements in $\pi_1(B)$ are denoted by the same symbol γ_e and called fundamental cycles. Let $S_T := \{\gamma_e \mid e \in E(B) \setminus E(T)\}$.

The fundamental group $\pi_1(G)$ embeds naturally as a subgroup in $\pi_1(B)$. The embedding is just the induced monomorphism $p_* : \pi_1(G) \to \pi_1(B)$. Let $\Gamma := \mathrm{Cay}(\pi_1(B), S_T)$ be the Cayley graph of $\pi_1(B)$ with respect to the fundamental cycles as the generators. It is isomorphic to the infinite $2\,|S_T|$-regular tree. Let $\Gamma_1 := p_*(\pi_1(G)) \leqslant \pi_1(B)$, and denote by Γ/Γ_1 the quotient graph of Γ with respect to the action of Γ_1 on $V(\Gamma)$. Such quotients of Cayley graphs are called *Schreier coset graphs*. The vertices of Γ/Γ_1 are the right cosets of $\pi_1(B)/\Gamma_1$ and $\Gamma_1\alpha$ and $\Gamma_1\beta$ are joined by an edge if $\Gamma_1\alpha\gamma = \Gamma_1\beta$ (or vice versa) for some $\gamma \in S_T$. Notice that Γ/Γ_1 is finite if and only if Γ_1 has finite index in $\pi_1(B)$. If Γ_1 is a normal subgroup of $\pi_1(B)$ then Γ/Γ_1 is just the Cayley graph of $\pi_1(B)/\Gamma_1$ with respect to S_T.

Corollary 4.5. *Let G be a connected infinite graph and B a connected finite graph. If $p : G \to B$ is a graph covering projection then G is amenable if and only if the quotient graph $\Gamma/\Gamma_1 = \mathrm{Cay}(\pi_1(B), S_T)/p_*(\pi_1(G))$ is amenable.*

Proof. In view of Proposition 4.4 it suffices to show that the quotient graph $G_1 := \Gamma/\Gamma_1$ is isomorphic to the graph H. Notice that $\Delta(G) = \Delta(B) < \infty$, so $s(G) = 0$ is equivalent to $h(G) = 0$.

Let $\pi_1(G) = \pi_1(G, \bar{b})$ be the fundamental group of G with respect to the base vertex $\bar{b} \in p^{-1}(b)$. Take a closed walk W in B with base point b, and let $\tilde{W}$ denote its lift to G such that the initial point of $\tilde{W}$ is $\bar{b}$. The terminal point of $\tilde{W}$, say b_1, is by the homotopy lifting property of covering spaces equal for all walks W' which are homotopic to W in B. The mapping which assigns with each W the corresponding point $b_1 \in p^{-1}(b) \equiv V(H)$ therefore determines a map $\phi : \pi_1(B) \to V(H)$. It is easy to see that under this mapping closed walks W_1, W_2 based at b have the same image if and only if $[W_1][W_2]^{-1} = [W_1 W_2^{-1}] \in \Gamma_1$. Therefore ϕ induces a bijection $\bar{\phi} : \pi_1(B)/\Gamma_1 \to V(H)$. To prove that $\bar{\phi}$ determines a graph isomorphism between G_1 and H it suffices to see that the fundamental cycles $\gamma_e \in S_T$ determine, via $\bar{\phi}$, the edges in H.

The graphs H and G_1 are both $2\,|S_T|$-regular (loops are counted twice). Choose a vertex $\Gamma_1\alpha \in V(G_1)$ and an edge η in G_1 incident to $\Gamma_1\alpha$, which corresponds to the generator $\gamma_e \in S_T$. The other end of η is $\Gamma_1\alpha\gamma_e$. Let W be a closed walk in B in the homotopy class α, and $\tilde{W}$ its lift to G with initial point $\bar{b}$. If its terminal point is b_1 then $\bar{\phi}(\Gamma_1\alpha) = b_1$. The lift of $W\gamma_e$ to G leads through b_1 and after that it uses the lift of γ_e which has all its edges, except the edge $\tilde{e} \in p^{-1}(e)$, in copies of T. It is clear that the ends of $\tilde{e}$ are in copies of T labelled $b_1 = \bar{\phi}(\Gamma_1\alpha)$ and

$b_2 = \bar{\phi}(\Gamma_1 \alpha \gamma_e)$. If we extend now $\bar{\phi}$ to the edge map $\bar{\phi}(\eta) := \tilde{e}$, this obviously determines a graph isomorphism $G_1 \to H$. □

Let us notice that most of the arguments in the above proof can easily be deduced from the general theory of covering spaces [12, pp. 161–164].

It should be mentioned that Proposition 4.4 yields more than just the result stated in Corollary 4.5. It gives explicit lower bounds on $h(G)$ in terms of $h(\Gamma/\Gamma_1)$. This is important if we want to have some applications of it. For example, the result of the next section (Theorem 5.1) show that for certain groups all their quotients have the isoperimetric number uniformly bounded away from zero. Using these to play the role of $\mathrm{Cay}(\pi_1(B), S_T)/\Gamma_1$, this might give a method to produce many covers G over B with the isoperimetric number bounded below by a positive constant. We shall leave the detailed discussion of this phenomenon for further works. Cf. also [2].

5. The Kazhdan property (T)

Group representations are used not only to make abstract groups 'visible' but also to make analytic methods accessible to do the analysis on groups. Let Γ be a group. If H is a Hilbert space and $\rho: \Gamma \to \mathrm{Aut}(H)$ is a homomorphism (where $\mathrm{Aut}(H)$ are all invertible linear mappings $H \to H$), then the pair (H, ρ) is called a *representation* of Γ. A *trivial* representation sends each group element $\gamma \in \Gamma$ to the identity of $\mathrm{Aut}(H)$. The representation is *unitary* if $\rho(\gamma)$ is a unitary transformation for each $\gamma \in \Gamma$. It is *irreducible* if $\rho(\Gamma)$ has no nontrivial invariant subspaces.

A locally compact group Γ has the *Kazhdan property* (T) (or is said to be a *Kazhdan group*) if there exist an $\varepsilon > 0$ and a compact subset K of Γ such that for every nontrivial irreducible unitary representation (h, ρ) of Γ and every $v \in H \setminus \{0\}$, $\|\rho(k)v - v\| > \varepsilon \|v\|$ for some $k \in K$. To be more specific we also say that Γ is (K, ε)-*Kazhdan*, or *Kazhdan for* K and ε.

An equivalent definition of property (T) is: Γ is Kazhdan if and only if there are $\varepsilon > 0$ and a compact $K \subseteq \Gamma$ such that for every unitary representation (H, ρ) of Γ which has some (K, ε)-invariant vectors $(\exists v \in H: \|\rho(k)v - v\| < \varepsilon \|v\|, \forall k \in K)$ has in fact some nonzero Γ-invariant vectors $(\exists v \in H, v \neq 0$ and $\rho(g)v = v$ for every $g \in \Gamma)$.

The property (T) was introduced by Kazhdan in [10] as a tool for studying discrete subgroups of Lie groups of finite co-volume. Examples of infinite Kazhdan groups are all semi-simple Lie groups with all factors having $\mathbb{R}$-rank $\geqslant 2$, and all their cofinite-volume discrete subgroups, e.g. $\mathrm{SL}(n, \mathbb{Z})$, $n \geqslant 3$, but not $\mathrm{SL}(2, \mathbb{Z})$. Let us mention also that free groups are not Kazhdan groups.

We restrict ourselves to finitely generated (discrete) groups. (It can also be shown that if a countable discrete group has property (T), then it is finitely generated.)

(i) A finite group is Kazhdan. To see this, let (H, ρ) be a unitary representation such that there exists a $v \in H$, $\|v\| = 1$, and $\|\rho(k)v - v\| < 1$ for each $k \in K := \Gamma$. Let

$$\tilde{v} := \frac{1}{|\Gamma|} \sum_{k \in \Gamma} \rho(k)v.$$

Then

$$\|\tilde{v} - v\| = \frac{1}{|\Gamma|} \left\| \sum_{k \in \Gamma} (\rho(k)v - v) \right\| < 1,$$

so $\tilde{v} \neq 0$. But

$$\rho(g)\tilde{v} = \frac{1}{|\Gamma|} \sum_{k \in \Gamma} \rho(g)\rho(k)v = \frac{1}{|\Gamma|} \sum_{k \in \Gamma} \rho(gk)v = \tilde{v},$$

for every $g \in \Gamma$. So Γ is Kazhdan (for $K = \Gamma$ and $\varepsilon = 1$).

(ii) Let Γ have property (T) for ε, K. If $K' \supseteq K$ is finite (compact) and $0 < \varepsilon' \leqslant \varepsilon$ then Γ has property (T) for ε', K'.

(iii) A countable discrete group with property (T) is finitely generated. Therefore we may assume by (ii) that K generates Γ.

(iv) If Γ has property (T) for ε, K, where K generates Γ, then Γ has property (T) for ε', K' where K' is any other generating set and $\varepsilon' = \varepsilon'(K') > 0$ some constant depending on K', K and ε. To prove this, express each $k \in K$ as a word in the alphabet $K' \cup (K')^{-1}$ and denote by $l(k)$ its length. Let $L = \max\{l(k) \mid k \in K\}$. Let (H, ρ) be a unitary representation of Γ and assume that there exists a vector $v \in H$, $\|v\| = 1$, such that $\|\rho(k')v - v\| < \varepsilon/L$ for all $k' \in K'$. Take now an arbitrary $k \in K$ and let $k = k'_n k'_{n-1} \cdots k'_1$ where $k'_i \in K' \cup (K')^{-1}$ and $n \leqslant L$. It will be assumed for the sake of easier notation that $K' = (K')^{-1}$. This may be done since the representation is unitary and hence $\|\rho(k)v - v\| = \|\rho(k^{-1})v - v\|$. We will use induction on n to show that $\|\rho(k)v - v\| = \|\rho(k'_n \cdots k'_1)v - v\| < (\varepsilon/L)n$. This is true by our assumption if $n = 1$. Otherwise

$$\begin{aligned}
\|\rho(k'_n \cdots k'_1)v - v\| &= \|\rho(k'_{n-1} \cdots k'_1)v - \rho(k'^{-1}_n)v\| \\
&\leqslant \|\rho(k'_{n-1} \cdots k'_1)v - v\| + \|\rho(k'^{-1}_n)v - v\| \\
&\leqslant \frac{\varepsilon}{L}(n-1) + \frac{\varepsilon}{L} = \frac{\varepsilon n}{L}.
\end{aligned}$$

It follows that any (K', ε')-invariant vector v is (K, ε)-invariant, and hence Γ has property (T) with respect to K', ε' where $\varepsilon' := \varepsilon/L$.

One of the fundamental properties of Kazhdan groups determines that such groups are something opposite to amenable groups, in a very strong sense. Let us recall that λ_1, ρ_1, and the isoperimetric numbers h and s for finite graphs were introduced in Section 4.

Theorem 5.1. *Let Γ be an infinite group generated by a finite set S, and let $G = \mathrm{Cay}(\Gamma, S)$ be its Cayley graph with respect to S. If Γ is (S, ε)-Kazhdan and Γ_1 is a proper subgroup of Γ then the quotient graph G/Γ_1 has*

$$\rho_1(G/\Gamma_1) \leqslant 1 - \frac{\varepsilon^2}{2k}, \quad \textit{and} \quad \lambda_1(G/\Gamma_1) \geqslant \frac{1}{2}\varepsilon^2, \tag{5.1}$$

where $k := |S \cup S^{-1}|$. If no generator in S has order 2 then (5.1) can be improved to $\rho_1(G/\Gamma_1) \leqslant 1 - \varepsilon^2/k$ and $\lambda_1(G/\Gamma_1) \geqslant \varepsilon^2$.

Proof. Denote by $G_1 := G/\Gamma_1$, and assume first that G_1 is infinite. Let ρ be the right regular representation of Γ in $H := \ell^2(\Gamma/\Gamma_1) = \ell^2(V(G_1))$, i.e., for $\gamma \in \Gamma$, $\rho(\gamma)$ is a permutation matrix indexed by $V(G_1) = \Gamma/\Gamma_1$ and having entries

$$\rho(\gamma)_{\Gamma_1\alpha, \Gamma_1\beta} = \begin{cases} 1 & \text{if } \Gamma_1\alpha\gamma = \Gamma_1\beta, \\ 0 & \text{otherwise.} \end{cases}$$

It is easily seen that (H, ρ) is a unitary representation of Γ. If a vector $x \in H$ is fixed by each $\rho(\gamma)$, $\gamma \in \Gamma$, it must have all entries equal since Γ is transitive on $V(G_1)$, *so* $x = 0$. Since Γ is (S, ε)-Kazhdan we conclude that for each $x \neq 0$ there is some $s \in S$ such that

$$\|\rho(s)x - x\| \geqslant \varepsilon \, \|x\|. \tag{5.2}$$

The graph G_1 is regular of valency k. Its Laplacian matrix is easily seen to be equal to

$$L = L(G_1) = \sum_{t \in S \cup S^{-1}} (I - \rho(t)). \tag{5.3}$$

Let $x \in \ell^2(V(G_1))$ be a unit vector and s a corresponding element from S such that (5.2) holds. Since $\rho(s)$ is unitary, it follows by using (5.2) that

$$\begin{aligned} &\langle x - \rho(s)x, x\rangle + \langle x - \rho(s^{-1})x, x\rangle \\ &\quad = \langle x - \rho(s)x, x\rangle + \langle \rho(s)x - x, \rho(s)x\rangle \\ &\quad = \|x - \rho(s)x\|^2 \geqslant \varepsilon^2. \end{aligned} \tag{5.4}$$

Consequently, by (5.3)

$$\begin{aligned} \langle Lx, x\rangle &= \sum_{t \in S \cup S^{-1}} \langle (I - \rho(t))x, x\rangle \\ &\geqslant \tfrac{1}{2}(\langle x - \rho(s)x, x\rangle + \langle x - \rho(s^{-1})x, x\rangle) \geqslant \tfrac{1}{2}\varepsilon^2. \end{aligned} \tag{5.5}$$

Note that we need the factor $\frac{1}{2}$ to cover the case $s = s^{-1}$. Now (3.1) says that $\lambda_1(G_1) \geqslant \varepsilon^2/2$, and $\rho_1(G_1) = (k - \lambda_1(G_1))/k \leqslant 1 - \varepsilon^2/(2k)$.

If G_1 is finite we undertake a similar way. Matrices $\rho(\gamma)$ are the same as before but they are assumed to act as automorphisms of the subspace H^0 of $\ell^2(V(G_1))$, where

$$H^0 = \{x \in \ell^2(V(G_1)) \mid x \perp \mathbf{1}\},$$

i.e., the orthogonal complement of the eigensubspace of $\lambda_0 = 0$ as the eigenvalue of $L(G_1)$. Then (H^0, ρ) is a unitary representation. Relations (5.2)–(5.5) also hold if $x \in H^0$, $\|x\| = 1$, and we are done by the fact that $\lambda_1(G_1) = \inf\{\langle Lx, x\rangle \mid x \in H^0, \|x\| = 1\}$. □

Let us mention an important example of a non-Kazhdan group which has a similar property as stated in Theorem 5.1 for Kazhdan groups. The modular group $\Gamma = \mathrm{PSL}(2, \mathbb{Z})$ is a quotient of a non-Kahdan group $\mathrm{SL}(2, \mathbb{Z})$, and thus it is not Kazhdan. Its congruence subgroups

$$\Gamma_n = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}(2, \mathbb{Z}) \mid a, d \equiv 1 \pmod{n},\ b, c \equiv 0 \pmod{n} \right\}$$

have the property that for each $n \geqslant 2$, $\lambda_1(\mathrm{Cay}(\Gamma, S)/\Gamma_n) \geqslant \varepsilon > 0$ for some explicitly known ε, which follows from the Selberg's Theorem [19], cf. also [11]. This is an advantage in various applications over the examples of Kazhdan groups for which no explicit values of ε are known.

Corollary 5.2. *If $G = \mathrm{Cay}(\Gamma, S)$ is the Cayley graph of a Kazhdan group then G has exponential growth, and so do all the quotients G/Γ_1 where Γ_1 is any subgroup of Γ of infinite index. Moreover, $\tau(G/\Gamma_1)$ is uniformly bounded away from* 1.

Proof. The Kazhdan group Γ is (S, ε)-Kazhdan for some $\varepsilon > 0$. By the preceding theorem, $\lambda_1(G/\Gamma_1) \geqslant \frac{1}{2}\varepsilon^2$, thus by Theorem 3.1, $h(G/\Gamma_1) \geqslant \frac{1}{2}\varepsilon^2$. Theorem 2.4 now completes the proof. □

It seems that the property of G stated in Theorem 5.1 is characteristic for Cayley graphs of Kazhdan groups. We use it to define what a graph with Kazhdan property (T) will be.

Let G be a graph and $\Gamma \leqslant \mathrm{Aut}(G)$. The pair (G, Γ) is said to have *Kazhdan property* (T) if there exists an $\varepsilon > 0$ such that for every subgroup Γ_1 of Γ, which is not transitive on $V(G)$, the graph G/Γ_1 has $s(G/\Gamma_1) \geqslant \varepsilon$. Here G/Γ_1 is the *directed* graph whose vertices are the orbits $\Gamma_1 v$, $v \in V(G)$, of the action of Γ_1 on $V(G)$, and in the direction from the orbit $\Gamma_1 v$ to $\Gamma_1 u$ there are as many directed edges as the number of neighbours of the vertex v in $\Gamma_1 u$. Although $\lambda_1(G_1)$ or $\rho_1(G_1)$ make sense for directed quotients $G_1 := G/\Gamma_1$ of undirected graphs (but this is not so obvious), we use in our definition the requirement that $s(G_1)$ is uniformly bounded away from zero. By the results of Section 3 this is equivalent to the condition that $\rho_1(G_1)$ is uniformly bounded away from 1. It should be mentioned that for a directed graph G_1, the transition isoperimetric number $s(G_1)$ is defined in the same way as for undirected graphs. However, $|\delta X|$ counts the number of edges with the initial point in X and the terminal point outside X, and $S(X)$ is the sum of the out-degrees of vertices in X. We must require Γ_1 not to be transitive on $V(G)$ to exclude the quotients with one vertex, for which the isoperimetric number is undefined.

If Γ is a Kazhdan group for (S, ε) where S generates Γ, then the pair $(\mathrm{Cay}(\Gamma, S), \Gamma)$ has the Kazhdan property (T).

Note that the quotient graphs G/Γ_1 are regular quotients, i.e. the natural projection $p: G \to G/\Gamma_1$, $p(v) = \Gamma_1 v$, is a regular covering projection. This means that p is a covering projection (i.e. a local isomorphism between directed graphs), and it is *regular* (the group of covering transformations, which contains Γ_1, is transitive on each fibre $p^{-1}(\Gamma_1 v) = \Gamma_1 v \subset V(G)$).

A graph G is said to be *ε-Kazhdan* if for every regular covering $p: G \to G_1$ where G_1 is a directed graph with at least two vertices, $s(G_1) \geqslant \varepsilon$. Let us stop with some open problems.

(1) Suppose that we know that the pair (G, Γ) is Kazhdan. Give any results to estimate the corresponding ε from below. Notice that upper bounds on ε can be given in terms of $\lambda_1(G)$, or $\tau(G)$. The importance of this question lies in the fact that there are several known families of Kazhdan groups but no explicit estimates on ε are known.

(2) How can one 'see' the Kazhdan property? Infinite Kazhdan groups (and their quotients) have exponential growth. But the converse is far from being true. What other geometric properties are shared by Kazhdan pairs (G, Γ)?

(3) It would be extremely important to get nontrivial examples of Kazhdan graphs. At the moment it seems that any infinite Kazhdan graph with rich automorphism group should be classified as a 'nontrivial' example.

6. Almost symmetric graphs and ends of graphs

In this section we present a short proof of a result relating the amenability and the ends of graphs (Proposition 6.2). Such a result does not hold without further restrictions on graphs since there are easily constructed examples of amenable, or non-amenable graphs (even trees) having arbitrary space of ends. Let us mention that Proposition 6.1 below is somehow implicit already in the classical work of Stallings [21] (for Cayley graphs), but we have not found it stated in the available literature. It was proved by Soardi and Woess [20] that a locally finite vertex transitive graph G with infinitely many ends has $h(G) > 0$. Our Proposition 6.2 is an obvious generalization of this result to 'almost symmetric' graphs.

A locally finite graph G is said to be *almost (vertex) symmetric* if there is a constant $D < \infty$ and a vertex $v \in V(G)$ such that for each $u \in V(G)$ there is an automorphism $\sigma \in \mathrm{Aut}(G)$ such that $\mathrm{dist}(u, \sigma(v)) \leqslant D$. For a connected graph G this is equivalent to saying that $\mathrm{Aut}(G)$ has only finitely many orbits on $V(G)$. It is clear that an almost symmetric graph G has $\Delta(G) = \Delta'(G) < \infty$. Similarly, one can prove that $h(G) = h'(G)$, $\lambda_1(G) = \lambda_1'(G)$, and $\rho_1(G) = \rho_1'(G)$.

For an arbitrary graph G one may define the *space of ends* of G, $\mathrm{Ends}(G)$, as follows: Let $\Omega(G)$ be the set of all one-way-infinite paths in G. Call two such paths A and B *equivalent*, $A \sim B$, if for each finite subgraph C of G, the infinite

parts of $A \setminus C$ and $B \setminus C$ belong to the same connected component of $G \setminus C$. I.e., no finite subgraph can separate A and B. The equivalence classes $\Omega(G)/\sim$ are the *ends* of G, denoted by Ends(G). There is a natural topology in Ends(G). Pick a vertex in each of the components of G, and let C_n ($n \geqslant 0$) be the induced subgraph of G containing all those vertices which are at distance at most n from some of the chosen vertices. For an end $\omega \in \mathrm{Ends}(G)$, let U_n be the set of all those ends which have representative paths in the same component of $G \setminus C_n$ as ω. The sets U_n define then a decreasing sequence of basis neighbourhoods of ω. With this topology Ends(G) becomes a compact metrizable space. It can be shown that for a connected graph G the space Ends(G) is homeomorphic to a closed subset of the Cantor set, and conversely—for an arbitrary closed subset A of the Cantor set there are graphs with Ends(G) homeomorphic to A.

If $\omega_1, \omega_2, \ldots, \omega_k$ are distinct ends of G, then, since the Cantor set is totally disconnected, there is a finite set $C \subset V(G)$ such that $G \setminus C$ pairwise separates the ends ω_i. In particular, $G \setminus C$ has at least k infinite components. Let us denote the number of infinite components of $G \setminus C$ by $c_\infty(C)$. Notice that if $C' \supseteq C$ then $c_\infty(C') \geqslant c_\infty(C)$.

Proposition 6.1. *Let G be a locally finite connected almost vertex symmetric graph. Then* Ends(G) *contains either* 0, 1, 2, *or infinitely many ends. In the latter case,* Ends(G) *is homeomorphic to the Cantor set, and is thus uncountable.*

Proof. Assume that G has more than two ends. Fix $v \in V(G)$, and let $B_n := B_n(v)$ be the ball of radius n centered at v. Let D be the constant corresponding to G from the definition of almost symmetric graphs. For some $N > D$, $c_\infty(B_{N-D}) \geqslant 3$ since G has more than two ends. If u is a vertex at distance $2N - D + 1$ from v and lies in one of the infinite components of $G \setminus B_{N-D}$ then there is a vertex u' at distance d from v, $2N - 2D < d \leqslant 2N + 1$, and an automorphism σ of G such that $\sigma(v) = u'$. Clearly,

$$c_\infty(B_{N-D}(u')) = c_\infty(B_{N-D}(v)) \geqslant 3.$$

One of the components of $G \setminus B_{N-D}(u')$ contains $B_{N-D}(v)$. Therefore

$$c_\infty(B_{N-D}(u') \cup B_{N-D}(v)) \geqslant 4.$$

We repeat this for some u in each of the infinite components, and it follows that $c_\infty(B_{3N-D+1}) \geqslant 6$. By the same method we see that $c_\infty(B_{5N-D+2}) \geqslant 12$, etc. In general,

$$c_\infty(B_{(2t+1)N-D+t}) \geqslant 3 \cdot 2^t.$$

Thus G has infinitely many ends. Moreover, it follows that the space of ends of G is homeomorphic to the ends space of some locally finite infinite tree without vertices of degree two, and this is known to be homeomorphic to the Cantor set. $\square$.

Proposition 6.2. *Let G be a locally finite, connected almost vertex symmetric graph.*

(a) *If G has infinitely many ends then $h(G) > 0$.*

(b) *If G has exactly two ends then $h(G) = 0$.*

Proof. (a) Let $\beta_n := \max\{|B_n(v)| \mid v \in V(G)\}$, and let N, D be as in the above proof. Let X be an arbitrary finite subset of $V(G)$. If $|X| < 2\beta_N$ then

$$\frac{|\delta X|}{|X|} \geqslant \frac{1}{4\beta_{2N}}.$$

Otherwise let $X_1 := \{u \in X \mid B_N(u) \nsubseteq X\}$ and $X_2 := \{u \in X \mid B_N(u) \subseteq X\}$. If $|X_1| \geqslant |X|/2$ then we estimate $|\delta X|$ as follows: For each $u \in X_1$ there is an edge $e \in \delta X$ at distance at most N form u. Any edge $e \in \delta X$ corresponds in this way to at most β_N distinct vertices of X_1. Therefore

$$\frac{|\delta X|}{|X|} \geqslant \frac{|X_1|}{\beta_N |X|} \geqslant \frac{1}{2\beta_N}.$$

In the remaining case $|X_2| \geqslant |X|/2$. In this case we first make sure that X_2 contains at least $t := \lfloor |X_2|/\beta_{2N} \rfloor$ disjoint N-balls, i.e. there are vertices $u_1, u_2, \ldots, u_t$ in X_2 which are pairwise at distance at least $2N+1$. Next we see, by the same method as in the proof of Proposition 6.1, that $c_\infty(\bigcup_{i=1}^t B_N(u_i)) \geqslant t$. Consequently, $c_\infty(X) \geqslant c_\infty(X_2) \geqslant t$, and so

$$|\delta X| \geqslant t \geqslant \left\lfloor \frac{|X_2|}{\beta_{2N}} \right\rfloor \geqslant \left\lfloor \frac{|X|}{2\beta_{2N}} \right\rfloor \geqslant \frac{|X|}{4\beta_{2N}}.$$

In the last inequality we use the assumption that $|X| \geqslant 4\beta_{2N}$. It follows that in each case $|\delta X|/|X|$ is bounded away from 0, so $h(G) > 0$.

(b) Let U be a finite subset of $V(G)$ such that $c_\infty(U) = 2$. If σ is an automorphism of G which sends a vertex $u \in U$ (arbitrarily) far away from U, then a finite component of $G \setminus (U \cup \sigma(U))$ contains (arbitrary) many vertices but only constantly many edges on its boundary. Thus obviously $h(G) = 0$. □

References

[1] R. Brooks, The fundamental group and the spectrum of the Laplacian, Comment. Math. Helv. 56 (1981) 581–598.

[2] R. Brooks, Combinatorial problems in spectral geometry, in: Curvature and Topology of Riemannian Manifolds, Lecture Notes in Math., Vol. 1201 (Springer, Berlin, 1986) 14–32.

[3] I. Chavel, Eigenvalues in Riemannian Geometry (Academic Press, New York, 1984).

[4] J. Cheeger, A lower bound for the smallest eigenvalue of the Laplacian, in: R. C. Gunnig, ed., Problems in Analysis, (Princeton Univ. Press, Princeton, NJ, 1970) 195–199.

[5] P.A. Fillmore, J.G. Stampfli and J.P. Williams, On the essential numerical range, the essential spectrum, and a problem of Halmos, Acta. Math. Sci. (Szeged) 33 (1972) 179–192.

[6] E. Følner, On groups with full Banach mean values, Math. Scand. 3 (1955) 243–254.

[7] P. Gerl, Random walks on graphs with a strong isoperimetric property, J. Theoret. Probab. 1 (1988) 171–188.
[8] P. Gerl, Amenable groups and amenable graphs, in: P. Eymard and J.-P. Pier, eds., Harmonic Analysis, Lecture Notes in Math., Vol. 1359 (Springer, Berlin, 1988) 181–190.
[9] T. Kato, Perturbation Theory for Linear Operators (Springer, Berlin, 1966).
[10] D. A. Kazhdan, Connection of the dual space of a group with the structure of its closed subgroups, Functional Anal. Appl. 1 (1967) 63–65.
[11] A. Lubotzky, Discrete groups, expanding graphs and invariant measures, preprint, 1989.
[12] W. Massey, Algebraic Topology: An Introduction (Harcourt, Brace and World, 1967).
[13] B. Mohar, Isoperimetric inequalities, growth, and the spectrum of graphs, Linear Algebra Appl. 103 (1988) 119–131.
[14] B. Mohar, The Laplacian spectrum of graphs, in: Y. Alavi et al., eds., Graph Theory, Combinatorics and Applications, Vol. 2 (Wiley, New York, 1991) 871–898.
[15] B. Mohar and W. Woess, A survey on spectra of infinite graphs, Bull. London Math. Soc. 21 (1989) 209-234.
[16] A.L.T. Paterson, Amenability (Amer. Math. Soc., Providence, RI, 1988).
[17] J.-P. Pier, Amenable Locally Compact Groups (Wiley, New York, 1984).
[18] J.M. Rosenblatt, Invariant measures and growth conditions, Trans. Amer. Math. Soc. 193 (1974) 33–53.
[19] A. Selberg, On the estimation of Fourier coefficients of modular forms, Proc. Sympos. Pure Math. 8 (1965) 1–15.
[20] P.M. Soardi and W. Woess, Amenability, unimodularity, and the spectral radius of random walks on infinite graphs, Math. Z. 205 (1990) 471–486.
[21] J. Stallings, Group Theory and Three-Dimensional Manifolds, Yale Math. Monographs, Vol. 4 (Yale Univ. Press, New Haven, London, 1971).
[22] J.G. Stampfli, Compact perturbations, normal eigenvalues, and a problem of Salinas, J. London Math. Soc. 9 (1974) 165–175.

Discrete Mathematics 95 (1991) 221–229
North-Holland

Reconstruction of infinite graphs

C.St.J.A. Nash-Williams
Mathematics Department, Reading University, Reading, UK

Received 14 September 1989

Abstract

C.St.J.A. Nash-Williams, Reconstruction of infinite graphs, Discrete Mathematics 95 (1991) 221–229.

The paper recalls several known results concerning reconstruction and edge-reconstruction of infinite graphs, and draws attention to some possibly interesting unsolved problems.

1. Introduction

In this paper, graphs are understood to be simple, i.e. without loops or multiple edges, and the letters G, H always denote graphs. Digraphs are understood to have no loops and no pair of edges with the same tail and the same head, but two distinct vertices of a digraph may be joined by two edges if these are oriented in opposite directions.

If $\xi \in V(G)$ then $G - \xi$ is called a *vertex-deleted* subgraph of G. The 'Reconstruction Problem' asks whether a graph is determined up to isomorphism if we know its vertex-deleted subgraphs up to isomorphism, i.e. whether a graph is in this sense 'reconstructible' from its vertex-deleted subgraphs. To express this more carefully, we make the following definitions: A *hypomorphism* of G onto H is a bijection $\phi: V(G) \to V(H)$ such that $G - \xi \cong H - \phi(\xi)$ for every $\xi \in V(G)$ (where $\cong$ means 'is isomorphic to'). Two graphs G, H are *hypomorphic* (in symbols, $G \simeq H$) if there exists a hypomorphism of G onto H. A graph G is *reconstructible* if every graph hypomorphic to G is isomorphic to G. The terms 'hypomorphism', 'hypomorphic' and 'reconstructible' are defined in the same way for digraphs, but using the notion of isomorphism appropriate to digraphs. The well-known 'Reconstruction Conjecture' can be formulated in either of the following equivalent ways.

Reconstruction Conjecture (first version). If G, H are hypomorphic finite graphs with at least 3 vertices then $G \cong H$.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

Reconstruction Conjecture (second version). Every finite graph with at least 3 vertices is reconstructible.

The analogous conjecture concerning digraphs is false: Stockmeyer [19–20] has shown that there are infinitely many mutually non-isomorphic finite digraphs which are not reconstructible.

It is somewhat easier to see that the analogue for infinite graphs of the Reconstruction Conjecture is not in general true. Probably the simplest counterexample is the fact that $T_\alpha \simeq 2T_\alpha$ but $T_\alpha \not\simeq 2T_\alpha$ for any infinite cardinal α, where T_α denotes a tree in which every vertex has valency α and $2T_\alpha$ is the union of two disjoint trees isomorphic to T_α. Nevertheless, there are some challenging problems about reconstructibility of infinite graphs, such as the following.

Problem 1. Is every locally finite connected infinite graph reconstructible?

Problem 2. If two infinite trees are hypomorphic, are they necessarily isomorphic?

Problem 3. A conjecture of Halin: If two infinite graphs are hypomorphic, then each is isomorphic to a subgraph of the other.

We remark that two hypomorphic connected graphs need not be isomorphic: the complements of T_α, $2T_\alpha$ provide an immediate counterexample, and Andreae [2] has provided counterexamples in which the graphs concerned and their complements are all connected. We remark also that two hypomorphic locally finite forests need not be isomorphic: see [10].

Problems 1, 2 and 3 are probably very difficult, but this paper will suggest some further problems (all unsolved so far as I know), some of which might be more accessible.

For further background concerning reconstruction, including basic elementary results quoted here without specific references, the reader may consult [5, Section 10.2] [6, 9, 15] and [16, Section 11(a)].

2. Some problems about reconstruction of graphs which are not necessarily locally finite

Definitions. A *neighbour* of a vertex ξ of G is a vertex adjacent to ξ. The set of neighbours of ξ in G will be denoted by $N_G(\xi)$; and $v_G(\xi)$ will denote the *valency* $|N_G(\xi)|$ of ξ in G. We shall write $V_\alpha(G) = \{\xi \in V(G): v_G(\xi) = \alpha\}$; and $E_{\alpha,\beta}(G)$ will denote the set of those edges of G which join elements of $V_\alpha(G)$ to elements of $V_\beta(G)$. If G is finite, the *valency sequence* of G is the sequence obtained by listing the valencies of its vertices in nondecreasing order, the *valency sequence* of

a vertex ξ of G is the sequence of valencies of its neighbours listed in nondecreasing order and the *valency sequence sequence* of G is the sequence of sequences obtained by listing the valency sequences of the vertices of G in lexicographical order.

A graph is 2-*connected* if it is connected and has no cut-vertices. A *block* of G is a maximal 2-connected subgraph of G.

If ξ is a vertex of a digraph D then $\mathrm{ov}_D(\xi)$ will denote the outvalency of ξ in D, i.e. the number of edges of D with tail ξ, and $\mathrm{iv}_D(\xi)$ will denote the invalency of ξ in D, i.e. the number of edges with head ξ.

Partial results concerning the Reconstruction Conjecture often take one of two forms: either they assert reconstructibility of particular kinds of finite graphs or they assert that hypomorphic finite graphs (with at least 3 vertices) must have some common properties, e.g.

Proposition 1. *Hypomorphic finite graphs with at least* 3 *vertices have the same valency sequence.*

Proposition 2. *Hypomorphic finite graphs with at least* 3 *vertices have the same valency sequence sequence.*

Proposition 3. *Hypomorphic finite graphs with at least* 3 *vertices have the same number of components.*

These three propositions are all easy to prove. However, easily proved results concerning reconstruction of finite graphs may suggest analogous questions about infinite graphs which are appreciably harder. For example, the following theorem of Andreae [4] is substantially harder to prove than Proposition 1.

Theorem 1. *If G, H are hypomorphic infinite graphs then $|V_\alpha(G)| = |V_\alpha(H)|$ for every cardinal number α.*

I am not aware of any known result concerning infinite graphs on the lines of Proposition 2, and therefore suggest the following problem.

Problem 4. If G, H are hypomorphic infinite graphs, does there necessarily exist a bijection $\phi : V(G) \to V(H)$ such that $|N_G(\xi) \cap V_\alpha(G)| = |N_H(\phi(\xi)) \cap V_\alpha(H)|$ for every $\xi \in V(G)$ and every cardinal number α?

Since some infinite graphs are not reconstructible, a suitable counterexample might answer Problem 4 in the negative. In that event, one might ask the following somewhat less ambitious question.

Problem 5. If G, H are hypomorphic infinite graphs and α, β are cardinal numbers, are $|E_{\alpha,\beta}(G)|$ and $|E_{\alpha,\beta}(H)|$ necessarily equal?

If G, J are graphs and $\xi \in V(G)$ let $v_G(\xi, J)$ denote the number of subgraphs S of G such that $\xi \in V(S)$ and $S \cong J$. Thus $v_G(\xi) = v_G(\xi, K_2)$, where K_2 is a complete graph with two vertices, and therefore Proposition 1 is a special case of the following proposition, whose proof is also easy.

Proposition 4. *If G, H, J are finite graphs, $G \simeq H$ and $|V(J)| < |V(G)|$ then there exists a bijection $\phi: V(G) \to V(H)$ such that $v_G(\xi, J) = v_H(\phi(\xi), J)$ for every $\xi \in V(G)$.*

Problem 6. Is the conclusion of Proposition 4 necessarily true when G, H are hypomorphic infinite graphs and J is a graph such that $|V(J)| < |V(G)|$ (or in the more restricted case in which G, H are hypomorphic infinite graphs and J is a finite graph)?

Manvel [12] has proved the following theorem about finite digraphs on the lines of Proposition 1.

Theorem 2. *If C, D are hypomorphic finite digraphs with at least 5 vertices then there exists a bijection $\phi: V(C) \to V(D)$ such that $\mathrm{ov}_C(\xi) = \mathrm{ov}_D(\phi(\xi))$ and $\mathrm{iv}_C(\xi) = \mathrm{iv}_D(\phi(\xi))$ for every $\xi \in V(C)$.*

Problem 7. Is the conclusion of Theorem 2 necessarily true when C, D are hypomorphic infinite digraphs?

Problem 8. Is something on the lines of Proposition 2, but involving outvalencies and invalencies, true for finite and/or infinite digraphs?

The following theorem of Andreae [4] resembles Proposition 3 but is substantially harder to prove.

Theorem 3. *If G, H are hypomorphic infinite graphs each of which has at least one vertex of finite valency, then G and H have the same number of components.*

(Since $T_\alpha \simeq 2T_\alpha$ when α is infinite, the hypothesis that G, H have at least one vertex of finite valency is needed.)

This might prompt the following question.

Problem 9. Must two hypomorphic infinite graphs have the same number of blocks? If not, can this conclusion be drawn subject to some additional hypothesis like that of Theorem 3?

3. Locally finite graphs

Definitions. If ξ is a vertex of G, let $S_n(\xi, G)$ denote the subgraph of G induced by those vertices which are in the same component of G as ξ and at distance $\leqslant n$ from ξ. If (R, ρ) is a rooted graph (i.e. R is a graph and ρ is a vertex of R), let $\alpha_n(G, R, \rho)$ denote the number of vertices ξ of G such that there exists an isomorphism of $S_n(\xi, G)$ onto R which maps ξ to ρ.

We shall say that G is *homeomorphic to* T_3 if G is a tree, $V(G) = V_2(G) \cup V_3(G)$ and there is no infinite path P in G such that $V(P) \subseteq V_2(G)$. More informally, this means that G can be obtained from T_3 (a tree in which every vertex has valency 3) by inserting a finite number n_λ of 2-valent vertices into each edge λ (where n_λ may depend on λ and may be 0).

A type of condition on a graph which tends to play a role in infinite graph theory is one which says, loosely speaking, that a graph has a specified finite number p of infinite wings (commonly called 'ends') branching out of a finite centre. To make this idea precise, we define a graph G to be *p-coherent*, where p is a positive integer, if G can be expressed as the union of a finite subgraph and p disjoint infinite subgraphs but cannot be expressed as the union of a finite subgraph and $p+1$ disjoint infinite subgraphs. We shall only be concerned with applying this condition to locally finite graphs (i.e. graphs in which every vertex has finite valency). We observe that a locally finite tree T is p-coherent if and only if there are exactly p one-way infinite paths starting at each vertex in T. For infinite cardinal numbers α, it will be convenient to define an α-coherent graph to be one which can be expressed as the union of α but not fewer 1-coherent subgraphs.

Bondy and Hemminger [6] proved the following theorem.

Theorem 4. *If p is an integer $\geqslant 2$ then every p-coherent locally finite tree is reconstructible.*

This was proved by considering the cases $p \geqslant 3$, $p = 2$ separately, and the proof was somewhat longer and more difficult in the case $p = 2$.

Thomassen [23] and Andreae [1] proved the following theorems respectively.

Theorem 5. *Every 1-coherent locally finite tree is reconstructible.*

Theorem 6. *Every $\aleph_0$-coherent locally finite tree is reconstructible.*

Andreae [1] also pointed out that Theorems 4, 5 and 6 together have the following consequence.

Theorem 7. *A locally finite tree is reconstructible if it has no subgraph homeomorphic to T_3.*

In [17–18], I obtained the following generalisation of Theorem 4.

Theorem 8. *If p is an integer $\geqslant 2$ then every p-coherent locally finite connected graph is reconstructible.*

This was proved for $p \geqslant 3$ in [17] and for $p = 2$ in [18]. As in the case of Theorem 4, the proof was longer and more complicated when $p = 2$. We recall that a p-coherent graph might be thought of as having p 'infinite wings' branching out of a 'finite centre'. When $p \geqslant 3$, one can at least vaguely identify the position of the 'finite centre', but a 2-coherent graph could have a shape something like an infinitely long sausage, extending to infinity in both directions, and what one then regards as the location of the 'finite centre' of the graph (from which its two 'infinite wings' branch out) may be a more or less arbitrary choice. This inability to identify, even vaguely, the 'finite centre' in some 2-coherent graphs deprives us of a useful tool in proving reconstructibility and accounts for the more elaborate arguments needed when $p = 2$.

Theorem 7 followed from Theorems 4, 5 and 6. Therefore a tempting target might be to prove generalisations of Theorems 5 and 6 which, when combined with Theorem 8, would answer in the affirmative the following question.

Problem 10. Is a locally finite connected graph necessarily reconstructible if it has no subgraph homeomorphic to T_3?

Let us say that an infinite graph G is *narrow* if it can be expressed as the union of an infinite sequence $S_1, S_2, \ldots$ of finite subgraphs such that $S_i \cap S_j$ is empty when $|i - j| \geqslant 2$ and

$$|V(S_1 \cap S_2)| = |V(S_2 \cap S_3)| = |V(S_3 \cap S_4)| = \cdots.$$

This definition, which can also be expressed in other ways, embodies the idea that 'G does not get wider and wider as it goes off to infinity'. Clearly every narrow graph is locally finite. We observe also that, if p is a positive integer, a p-coherent locally finite tree is necessarily narrow since it can be expressed as the union of finite subgraphs $S_1, S_2, \ldots$ such that $S_i \cap S_j$ is empty when $|i - j| \geqslant 2$ and $|V(S_i \cap S_{i+1})| = p$ for every positive integer i.

Theorem 8 might suggest that we should try to generalise Theorem 5 by proving that every 1-coherent locally finite connected graph is reconstructible; but I suspect that this would be very difficult since 1-coherent graphs lack the kind of recognisable 'centre' which facilitates the proof of Theorem 8 when $p \geqslant 3$ and lack even those features of 2-coherent graphs which, somewhat less readily, enable Theorem 8 to be proved for $p = 2$. Indeed, proving reconstructibility of all

1-coherent locally finite connected graphs may well be of the same order of difficulty as Problem 1. However, since 1-coherent trees are narrow, another possible generalisation of Theorem 5 would be an affirmative answer to the following question.

Problem 11. Is every narrow connected graph reconstructible?

In fact, it would suffice to answer this question for 1-coherent narrow connected graphs, because it is easily proved that every narrow connected graph is p-coherent for some positive integer p, and when $p \geqslant 2$ Theorem 8 can be applied.

Although I have not yet found time to write out the details, it is probably easy to prove that, if p is a positive integer, every p-coherent locally finite graph that is not narrow has a subgraph homeomorphic to T_3. It may also be fairly easy to prove that a locally finite connected infinite graph has such a subgraph if, for all $p \leqslant \aleph_0$, it fails to be p-coherent. Therefore one might hope to obtain an affirmative answer to Problem 10 from affirmative answers to Problem 11 and the following problem.

Problem 12. Is every $\aleph_0$-coherent locally finite connected graph reconstructible?

Some of the questions raised in Section 2 can be answered for locally finite graphs. The following result is Theorem 1 of [3] (or perhaps, more accurately, a consequence of this theorem and the fact, proved in [6], that hypomorphic infinite graphs have the same finite components up to isomorphism).

Theorem 9. *If G, H are hypomorphic infinite locally finite graphs and (R, ρ) is a rooted graph and n is a positive integer then $\alpha_n(G, R, \rho) = \alpha_n(H, R, \rho)$.*

It is easily seen that this theorem implies affirmative answers to Problems 4 and 5 for locally finite graphs and to Problem 6 when G, H are locally finite and J is finite and connected. Moreover, the proof of Theorem 9 can be adapted to prove an analogous theorem concerning infinite locally finite digraphs, which implies affirmative answers to Problems 7 and 8 for such digraphs.

Theorem 9 can be generalised to some extent. Let $\mathscr{L}$ be the class of infinite locally finite graphs. Let a function f from $\mathscr{L}$ into the class of cardinal numbers be called *admissible* if it satisfies the conditions:

(i) if $G, H \in \mathscr{L}$ and $G \cong H$ then $f(G) = f(H)$;

(ii) if $G \in \mathscr{L}$ and $\omega \in V(G)$ then either $f(G) = f(G - \omega)$ or $f(G)$, $f(G - \omega)$ are both finite;

(iii) if $G \in \mathscr{L}$ and $\omega \in V(G)$ then for almost all $\alpha \in V(G - \omega)$ we have $f(G) + f(G - \omega - \alpha) = f(G - \omega) + f(G - \alpha)$.

(The expression 'almost all $\alpha \in V(G-\omega)$' means 'all but finitely many vertices $\alpha \in V(G-\omega)$'.) We can think of (ii) as saying that $f(G)$, $f(G-\omega)$ differ by at most a finite amount, which is true of many functions of interest since $G-\omega$ differs from G only by the absence of one vertex and finitely many edges. If the numbers involved are finite, the equation in (iii) can be written in the possibly more illuminating form

$$f(G)-f(G-\alpha)=f(G-\omega)-f(G-\omega-\alpha).$$

Thus (iii) may be thought of as saying that, with the possible exception of finitely many choices of α, removal of a vertex α has the same effect on $f(G-\omega)$ as on $f(G)$. This might seem likely to be true, for many naturally arising functions f, provided that α is not too near to ω in G. Thus both (ii) and (iii) seem fairly natural conditions. It is possible to prove the following theorem.

Theorem 10. *If f is an admissible function and G, $H \in \mathscr{L}$ and $G \simeq H$ then $f(G)=f(H)$.*

This contains Theorem 9, since $f(G)$ can be taken to be $\alpha_n(G, R, \rho)$. Moreover, Theorem 10 implies an affirmative answer to Problem 9 for infinite locally finite graphs because it is not hard to prove that f is admissible if $f(G)$ is the number of blocks of G for each $G \in \mathscr{L}$.

4. Edge-reconstruction

If $\lambda \in E(G)$ then $G-\lambda$ is called an *edge-deleted* subgraph of G. The 'Edge-Reconstruction Problem' asks whether a graph is 'reconstructible from its edge-deleted subgraphs'. To be more precise, an *edge-hypomorphism* of G onto H is a bijection $\phi: E(G) \to E(H)$ such that $G-\lambda \cong H-\phi(\lambda)$ for every $\lambda \in E(G)$. Two graphs G, H are *edge-hypomorphic* if there exists an edge-hypomorphism of G onto H. A graph G *is edge-reconstructible* if every graph edge-hypomorphic to G is isomorphic to G. The *Edge-Reconstruction Conjecture* is the conjecture that if G, H are edge-hypomorphic finite graphs with at least 4 edges then $G \cong H$, or, equivalently, that every finite graph with at least 4 edges is edge-reconstructible. A noteworthy result of Müller [13] (obtained by sharpening an idea of Lovász [11]) states that a finite graph with n vertices and more than $(n \log n)/\log 2$ edges is necessarily edge-reconstructible.

Infinite graphs are not all edge-reconstructible, although some ingenuity may be needed to prove this. Thomassen [21] has found examples of infinite graphs which are not edge-reconstructible and in [22] has found examples of infinite graphs without 0-valent vertices which are reconstructible but not edge-reconstructible, in contrast with the theorem [8] that every reconstructible finite graph with at least 4 edges and no 0-valent vertices is edge-reconstructible.

A graph G is said to be *bidegreed* if the set $\{v_G(\xi): \xi \in V(G)\}$ contains at most two distinct numbers. Myrvold, Ellingham and Hoffman [14] have proved that all

bidgegreed finite graphs with at least 4 edges are edge-reconstructible. Some but not all of their arguments seem likely to adapt fairly readily to infinite but locally finite graphs, and so it might be interesting to investigate the following question.

Problem 13. Is every bidegreed infinite locally finite graph edge-reconstructible?

References

[1] T. Andreae, On the reconstruction of locally finite trees, J. Graph Theory 5 (1981) 123–135.

[2] T. Andreae, Note on the reconstruction of infinite graphs with a fixed finite number of components, J. Graph Theory 6 (1982) 81–83.

[3] T. Andreae, On the reconstruction of locally finite, infinite graphs, J. Combin. Inform. System Sci. 7 (1982) 65–74.

[4] T. Andreae, Reconstructing the degree sequence and the number of components of an infinite graph, Discrete Math. 39 (1982) 1–7.

[5] M. Behzad, G. Chartrand and L. Lesniak-Foster, Graphs and Digraphs (Prindle, Weber and Schmidt, Boston, MA, 1979).

[6] J. A. Bondy and R. L. Hemminger, Reconstructing infinite graphs, Pacific J. Math. 52 (1974) 331–340.

[7] J. A. Bondy and R. L. Hemminger, Graph reconstruction—a survey, J. Graph Theory 1 (1977) 227–268.

[8] D. A. Greenwell, Reconstructing graphs, Proc. Amer. Math. Soc. 30 (1971) 431–433.

[9] F. Harary, On the reconstruction of a graph from a collection of subgraphs, in: M. Fiedler, ed., Proc. Symp on Theory of Graphs and its Applications (Czechoslovak Academy of Sciences, Prague, 1964) 47–52.

[10] F. Harary, A. J. Schwenk and R. L. Scott, On the reconstruction of countable forests, Publ. Inst. Math. (Beograd) 13 (1972) 39–42.

[11] L. Lovász, A note on the line reconstruction problem, J. Combin. Theory Ser. B 13 (1972) 309–310.

[12] B. Manvel, Reconstructing the degree-pair sequence of a digraph, J. Combin. Theory Ser. B 15 (1973) 18–31.

[13] V. Müller, The edge reconstruction hypothesis is true for graphs with more than $n \log_2 n$ edges, J. Combin. Theory Ser. B. 22 (1977) 281–283.

[14] W. J. Myrvold, M. N. Ellingham and D. G. Hoffman, Bidegreed graphs are edge-reconstructible, J. Graph Theory 11 (1987) 281–302.

[15] C. St. J. A. Nash-Williams, The reconstruction problem, in: L. W. Beineke and R. J. WIlson, eds., Selected Topics in Graph Theory (Academic Press, London, 1978) 205–236.

[16] C. St. J. A. Nash-Williams, A. glance at graph theory—Part II, Bull. London Math. Soc. 14 (1982) 294–328.

[17] C St. J. A. Nash-Williams, Reconstruction of locally finite connected graphs with at least three infinite wings, J. Graph Theory 11 (1987) 497–505.

[18] C. St. J. A. Nash-Williams, Reconstruction of locally finite connected graphs with two infinite wings, Discrete Math., to appear.

[19] P. K. Stockmeyer, The falsity of the reconstruction conjecture for tournaments, J. Graph Theory 1 (1977) 19–25.

[20] P. K. Stockmeyer, A census of nonreconstructible digraphs I. Six related families, J. Combin Theory Ser. B 31 (1981) 232–239.

[21] C. Thomassen, Counterexamples to the edge-reconstruction conjecture for infinite graphs, Discrete Math. 19 (1977) 293–295.

[22] C. Thomassen, Reconstructibility versus edge reconstructibility of infinite graphs, Discrete Math. 24 (1978) 231–233.

[23] C. Thomassen, Reconstructing 1-coherent locally finite trees, Comment. Math. Helv. 53 (1978) 608–612.

Discrete Mathematics 95 (1991) 231–254
North-Holland

f-Optimal factors of infinite graphs

F. Niedermeyer
Institut für Mathematik der Universität Hannover, Welfengarten 1, W-3000 Hannover 1, Germany

Received 4 April 1990

Abstract

Niedermeyer, F., f-Optimal factors of infinite graphs, Discrete Mathematics 95 (1991) 231–254.

A necessary and sufficient condition for the existence of perfect f-factors of countable graphs is proved. Further it is shown that every countable graph and every rayless graph possesses an f-optimal factor. Finally the structure of f-optimal factors of infinite graphs is examined.

Introduction

In this paper $G=(V, E)$ denotes always a graph, where $V \neq \emptyset$ is the set of vertices and $E \subseteq \{e \subseteq V: |e|=2\}$ is the set of edges of G, and f a function from V into ω, the set of natural numbers.

We will deal with the question of whether there is a perfect f-factor of G, i.e. whether there is a subset F of E such that $d_F(v)=f(v)$ for each vertex v. If this question is answered negatively, we ask whether there is at least some F such that d_F approximates f optimally. We will consider only those subsets F of E for which $d_F(v)$ is finite for each $v \in V$. Such a set of edges we call a factor of G. For comparing two factors F and H with respect to the property of being optimal, we call F *better* than H if $|d_F(v)-f(v)| \leqslant |d_H(v)-f(v)|$ for every vertex v, and $|d_F(v)-f(v)| < |d_H(v)-f(v)|$ for at least one v.

In this paper we prove for countable G that G possesses a perfect f-factor iff G is *unobstructed*, which means that for every nonperfect f-factor F and for every unsaturated vertex v there exists an F-alternating path starting at v such that the symmetric difference of F and the edge set which is traversed by the path is better than F. Moreover, it is shown for countable G that G possesses a perfect f-factor iff for every nonperfect factor F there exists a factor better than F. This result is equivalent to the fact that G contains an optimal factor if G is countable. Further we will see: if G is rayless, then there even exists, for every non-optimal factor F,

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

an optimal factor better than F. These results are generalizations of the corresponding statements for 1-factors which have been proved by Steffens [7] and Schmidt [6].

In Section 6 we will examine the structure of optimal factors. For finite graphs, Lovász and Plummer [4] introduced a decomposition of a graph's vertex set concerning optimal factors. It makes sense to consider this decomposition for arbitrary graphs. It will turn out that the properties of this decomposition which have been shown in [4] hold in infinite graphs, too.

1. Factors and f-factors

For any subset F of E,

$$d_F(v) := |\{e \in F : v \in e\}|$$

defines a function d_F from V into the class of cardinals. On the class of functions from V into the class of cardinals we define a weak order by

$$g \preccurlyeq h \;:\Leftrightarrow\; g(v) \leq h(v) \text{ for all } v \in V$$

and write $g \prec h$ if $g \preccurlyeq h$ and $g \neq h$. If misunderstandings are impossible, we denote by 1 the function with domain V and range $\{1\}$.

$F \subseteq E$ is said to be a *factor* of G if $d_F(v)$ is finite for each $v \in V$, and F is called an *f-factor* if $d_F \preccurlyeq f$. A factor F will be called *(f-)perfect* if $d_F = f$. If F is an f-factor, a vertex $v \in V$ is said to be *(f-F-)saturated* (*unsaturated*, resp.) if $d_F(v) = f(v)$ ($d_F(v) < f(v)$, resp.). If F is a factor and if $g : V \to \omega$ is a function, we define a function $\delta_{F,g} : V \to \omega$ by

$$\delta_{F,g}(v) := |d_F(v) - g(v)|.$$

If $g = f$, we shall write 'δ_F' instead of '$\delta_{F,f}$'. A factor F_1 is said to be *better* than a factor F_2 (with respect to f) if

$$\delta_{F_1} \prec \delta_{F_2},$$

i.e. if the function f is approximated better by d_{F_1} than by d_{F_2}. A factor is called *(f-)optimal* if there is no better factor.

Our aim is to characterize graphs which have an optimal factor, and to describe the structure of optimal factors. Indeed, the complete bipartite graph $K_{\aleph_0,\aleph_1}$ and the function $f = 1$ show that not every graph has an optimal factor. Another important example is the bipartite graph $G^* = (V_0 \cup V_1, E^*)$, where

$$V_0 = \{(\alpha, 0) : \alpha < \aleph_1 \text{ is a limit ordinal}\},$$
$$V_1 = \aleph_1 \times \{1\},$$
$$E^* = \{\{(\alpha, 0), (\beta, 1)\} : \beta < \alpha\}.$$

We will see later that there is no 1-optimal factor of G^*.

A measure for the 'goodness' of a factor F is in some cases the *defect* of F,

$$\Delta(F, f) := \sum_{v \in V} \delta_{F,f}(v).$$

In Lovász and Plummer [4] only finite graphs are investigated, and there a factor is called 'optimal' if its defect is minimal. We will see later that for finite graphs this definition coincides with our definition.

At the end of this section we give some definitions concerning the terminology used in this paper. A sequence $\mathscr{P} = (v_i)_{0 \leq i < k}$ $(0 < k \leq \omega)$ is called a *walk* if $\{v_{i-1}, v_i\} \in E$ for $0 < i < k$. We define

$$V(\mathscr{P}) := \{v_i : 0 \leq i < k\} \quad \text{and} \quad E(\mathscr{P}) := \{\{v_{i-1}, v_i\} : 0 < i < k\}.$$

If $k < \omega$, $\mathscr{P}$ is said to have the *length* $k-1$. $\mathscr{P}$ is called a *trail* (*path*, resp.) if no edge (vertex, resp.) will be traversed twice by $\mathscr{P}$. If $\mathscr{P} = (v_i)_{0 \leq i \leq k}$ and $\mathscr{Q} = (\omega_i)_{0 \leq i < l}$ are walks satisfying $k < \omega$, $l \leq \omega$, and $v_k = w_0$, then $\widehat{\mathscr{P}\mathscr{Q}}$ denotes the walk $(u_i)_{0 \leq i < k+l}$, where $u_i = v_i$ if $0 \leq i < k$, and $u_i = w_{i-k}$ if $k \leq i < k+l$. If $F_1, F_2 \subseteq E$, then a walk $\mathscr{P}$ is called F_1-F_2-*alternating* if the edges of $\mathscr{P}$ are alternately in $F_1 \setminus F_2$ and $F_2 \setminus F_1$. If $F \subseteq E$, let $F^C := E \setminus F$. A walk is called F-*alternating* if it is F-F^C-alternating. Note that an F_1-F_2-alternating walk is both F_1-alternating and F_2-alternating.

Given a function $g : V \to \omega$ and a vertex $v \in V$, we define new functions g_v and g^v on V as follows:

$$g_v(x) := \begin{cases} g(x) & \text{if } x \in V \setminus \{v\}, \\ g(x) - 1 & \text{if } x = v, \end{cases}$$

$$g^v(x) := \begin{cases} g(x) & \text{if } x \in V \setminus \{v\}, \\ g(x) + 1 & \text{if } x = v. \end{cases}$$

We can compose such functions; let $g^{uv} := (g^u)^v$, $g^u_v := (g^u)_v = (g_v)^u, \ldots$.

For any sets A, B, the symmetric difference $(A \setminus B) \cup (B \setminus A)$ will be denoted by $A \oplus B$.

2. F_1-F_2-trails

If F is a factor and $\mathscr{P} = (v_i)_{0 \leq i < k}$ an F-alternating trail, then $F \oplus E(\mathscr{P})$ is a factor, too. d_F and $d_{F \oplus E(\mathscr{P})}$ differ at most at the start- and the endpoint of $\mathscr{P}$. If for instance $1 < k < \omega$, $\{v_0, v_1\} \in F$, and $\{v_{k-2}, v_{k-1}\} \in F^C$, then $d_{F \oplus E(\mathscr{P})} = d_{F v_0}^{v_{k-1}}$. F-alternating trails may be useful to 'improve' a factor F. This can be illustrated by the example G^* from the previous section: Let $F \subseteq E^*$ be a factor of G^* and let $v_0 \in V_0$ be a vertex such that $d_F(v_0) = 0$. Now assume that, for $0 \leq i \leq l$, v_i is already defined such that $(v_i)_{0 \leq i \leq l}$ is an F-alternating trail. If l is even, then $v_l \in V_0$. Choose $v_{l+1} \in V_1 \setminus \{v_i : i < l\}$ such that $\{v_l, v_{l+1}\} \in F^C$. This is possible since $d_{E^*}(v_l) = \aleph_0$ and F is a factor. If l is odd, then $v_l \in V_1$. If $d_F(v_l) > 0$, choose $v_{l+1} \in V_0$ such that $\{v_l, v_{l+1}\} \in F$. If on the other hand $d_F(v_l) = 0$, let $k := l + 1$

and stop this recursive definition. If the definition does not stop, let $k := \omega$. Let $\mathscr{P} := (v_i)_{0 \leq i < k}$. $\mathscr{P}$ is an F-alternating trail, and we get

$$d_{F \oplus E(\mathscr{P})} = d_F^{v_0, v_{k-1}} \quad \text{if} \quad k < \omega, \quad \text{and} \quad d_{F \oplus E(\mathscr{P})} = d_F^{v_0} \quad \text{if} \quad k = \omega.$$

Therefore $\delta_{F \oplus E(\mathscr{P}),1} = (\delta_{F,1})_{v_0 v_{k-1}} < \delta_{F,1}$, if $k < \omega$, and $\delta_{F \oplus E(\mathscr{P}),1} = (\delta_{F,1})_{v_0} < \delta_{F,1}$, if $k = \omega$. So $F \oplus E(\mathscr{P})$ is better (with respect to $f = 1$) than F. By Fodor's Theorem (see e.g. [3, Theorem 11.2.6]) there exists, for every factor F of G^*, a vertex $v_0 \in V_0$ satisfying $d_F(v_0) = 0$. Thus no factor of G^* is 1-optimal.

The possibility of modifying a factor by an alternating trail will now be examined. The following lemma and definition will play a fundamental role.

Lemma 1. *Let F_1, F_2 be factors of G and let $\{v_0, v_1\} \in F_1 \setminus F_2$. Then there exists a trail $\mathscr{P} = (v_i)_{0 \leq i < k}$ with the following properties:*

(i) $k > 1$,

(ii) *$\mathscr{P}$ is F_1-F_2-alternating,*

(iii) *$k = \omega$ or*

$$k < \omega, \qquad \{v_{k-2}, v_{k-1}\} \in F_1 \setminus F_2 \text{ and } d_{F_1}(v_{k-1}) > d_{F_2}^{v_0}(v_{k-1})$$

or

$$k < \omega, \qquad \{v_{k-2}, v_{k-1}\} \in F_2 \setminus F_1 \text{ and } d_{F_1}(v_{k-1}) < d_{F_2}^{v_0}(v_{k-1}).$$

Such a trail we call an F_1-F_2-trail.

Proof. The vertices v_i of $\mathscr{P}$ will be defined recursively. Assume that $v_0, \ldots, v_l$ ($l \geq 1$) are already defined such that $\mathscr{P}_l := (v_i)_{0 \leq i \leq l}$ is an F_1-F_2-alternating trail.

If l is odd, then $\{v_{l-1}, v_l\} \in F_1 \setminus F_2$. If $d_{F_1}(v_l) > d_{F_2}^{v_0}(v_l)$, let $k := l + 1$, $\mathscr{P} := \mathscr{P}_l$, and stop the recursive definition. Otherwise we have $d_{F_1}(v_l) \leq d_{F_2}^{v_0}(v_l)$. Since

$$d_{F_1 \cap E(\mathscr{P}_l)}(v_l) = d_{F_2 \cap E(\mathscr{P}_l)}^{v_0}(v_l) + 1,$$

we can choose a vertex v_{l+1} such that $\{v_l, v_{l+1}\} \in (F_2 \setminus E(\mathscr{P}_l)) \setminus (F_1 \setminus E(\mathscr{P}_l)) = F_2 \setminus (E(\mathscr{P}_l) \cup F_1)$. Thus $\mathscr{P}_{l+1} = (v_i)_{0 \leq i \leq l+1}$ is an F_1-F_2-alternating trail.

If l is even, we stop the recursive definition if $d_{F_1}(v_1) < d_{F_2}^{v_0}(v_l)$, or we can find similarly to the first case an edge $\{v_l, v_{l+1}\} \in F_1 \setminus (E(\mathscr{P}_l) \cup F_2)$.

If the definition does not stop, let $k := \omega$ and $\mathscr{P} := (v_i)_{0 \leq i < \omega}$. □

The following lemma gives us some possibilities to apply the F_1-F_2-trails.

Lemma 2. *Let F_1, F_2 be factors of G and let $\mathscr{P} = (v_i)_{0 \leq i < k}$ be an F_1-F_2-trail.*

(a) *If $\delta_{F_2}(v_0) < \delta_{F_1}(v_0)$ and $d_{F_1}(v_0) > d_{F_2}(v_0)$, then*

$$\delta_{F_1 \oplus E(\mathscr{P})}(v_0) = \delta_{F_1 v_{k-1}}(v_0) - 1 \quad \text{if } k < \omega,$$

$$\delta_{F_1 \oplus E(\mathscr{P})}(v_0) = \delta_{F_1}(v_0) - 1 \quad \text{if } k = \omega.$$

(b) *If* $\delta_{F_1}(v_0) < \delta_{F_2}(v_0)$ *and* $d_{F_1}(v_0) > d_{F_2}(v_0)$, *then*

$$\delta_{F_2 \oplus E(\mathscr{P})}(v_0) = \delta_{F_2 v_{k-1}}(v_0) - 1 \quad \text{if } k < \omega,$$

$$\delta_{F_2 \oplus E(\mathscr{P})}(v_0) = \delta_{F_2}(v_0) - 1 \quad \text{if } k = \omega.$$

(c) *If* $k < \omega$, $v_0 \neq v_{k-1}$, *and* $\delta_{F_2}(v_{k-1}) \leqslant \delta_{F_1}(v_{k-1})$, *then*

$$\delta_{F_1 \oplus E(\mathscr{P})}(v_{k-1}) = \delta_{F_1}(v_{k-1}) - 1.$$

(d) *If* $k < \omega$, $v_0 \neq v_{k-1}$, *and* $\delta_{F_1}(v_{k-1}) \leqslant \delta_{F_2}(v_{k-1})$, *then*

$$\delta_{F_2 \oplus E(\mathscr{P})}(v_{k-1}) = \delta_{F_2}(v_{k-1}) - 1.$$

Before we present the simple proof of this lemma we want to give a corollary.

Corollary 1. *Let* $\mathscr{P} = (v_i)_{0 \leqslant i < k}$ *be an* F_1-F_2-*trail such that* $d_{F_1}(v_0) > d_{F_2}(v_0)$.

(a) *If* $\delta_{F_2}(v_0) < \delta_{F_1}(v_0)$ *and* $\delta_{F_2}(v_{k-1}) \leqslant \delta_{F_1}(v_{k-1})$, *then* $F_1 \oplus E(\mathscr{P})$ *is a factor better than* F_1.

(b) *If* $\delta_{F_1}(v_0) < \delta_{F_2}(v_0)$ *and* $\delta_{F_1}(v_{k-1}) \leqslant \delta_{F_2}(v_{k-1})$, *then* $F_2 \oplus E(\mathscr{P})$ *is a factor better than* F_2.

Note that the assumption $\delta_{F_2}(v_{k-1}) \leqslant \delta_{F_1}(v_{k-1})$ ($\delta_{F_1}(v_{k-1}) \leqslant \delta_{F_2}(v_{k-1})$, resp.) is fulfilled if F_2 is better than F_1 (F_1 is better than F_2, resp.).

Proof. (a) *Case* 1: $k = \omega$ *or* ($k < \omega$ *and* $v_0 \neq v_{k-1}$).

Since $\delta_{F_2}(v_0) < \delta_{F_1}(v_0)$ and $d_{F_1}(v_0) > d_{F_2}(v_0)$, we have $d_{F_1}(v_0) > f(v_0)$. Therefore

$$\begin{aligned}\delta_{F_1 \oplus E(\mathscr{P})}(v_0) &= |d_{F_1 \oplus E(\mathscr{P})}(v_0) - f(v_0)| = |d_{F_1}(v_0) - 1 - f(v_0)| \\ &= |d_{F_1}(v_0) - f(v_0)| - 1 = \delta_{F_1}(v_0) - 1.\end{aligned}$$

Case 2: $k < \omega$ *and* $v_0 = v_{k-1}$.

Case 2.1: $\{v_{k-2}, v_{k-1}\} \in F_1 \setminus F_2$ *and* $d_{F_1}(v_{k-1}) > d_{F_2}^{v_0}(v_{k-1}) = d_{F_2}(v_{k-1}) + 1$.

Since $\delta_{F_2}(v_0) < \delta_{F_1}(v_0)$, we conclude that $d_{F_1}(v_0) > f(v_0) + 1$. Therefore

$$\begin{aligned}\delta_{F_1 \oplus E(\mathscr{P})}(v_0) &= |d_{F_1 \oplus E(\mathscr{P})}(v_0) - f(v_0)| = |d_{F_1}(v_0) - 2 - f(v_0)| \\ &= |d_{F_1}(v_0) - f(v_0)| - 2 = \delta_{F_1 v_{k-1}}(v_0) - 1.\end{aligned}$$

Case 2.2: $\{v_{k-2}, v_{k-1}\} \in F_2 \setminus F_1$ *and* $d_{F_1}(v_{k-1}) < d_{F_2}^{v_0}(v_{k-1}) = d_{F_2}(v_{k-1}) + 1$.

This case cannot occur since $d_{F_1}(v_0) > d_{F_2}(v_0)$.

(b), (c), and (d) can be proved similarly. □

A first application of the 'technique of F_1-F_2-trails, we find in the proof of the following lemma.

Lemma 3. *Let F be an optimal factor and let $\mathcal{P} = (v_i)_{0 \le i < k}$ be an F-alternating trail satisfying one of the following properties:*

(i) $\{v_0, v_1\} \in F$ *and* $d_F(v_0) > f(v_0)$, *or*

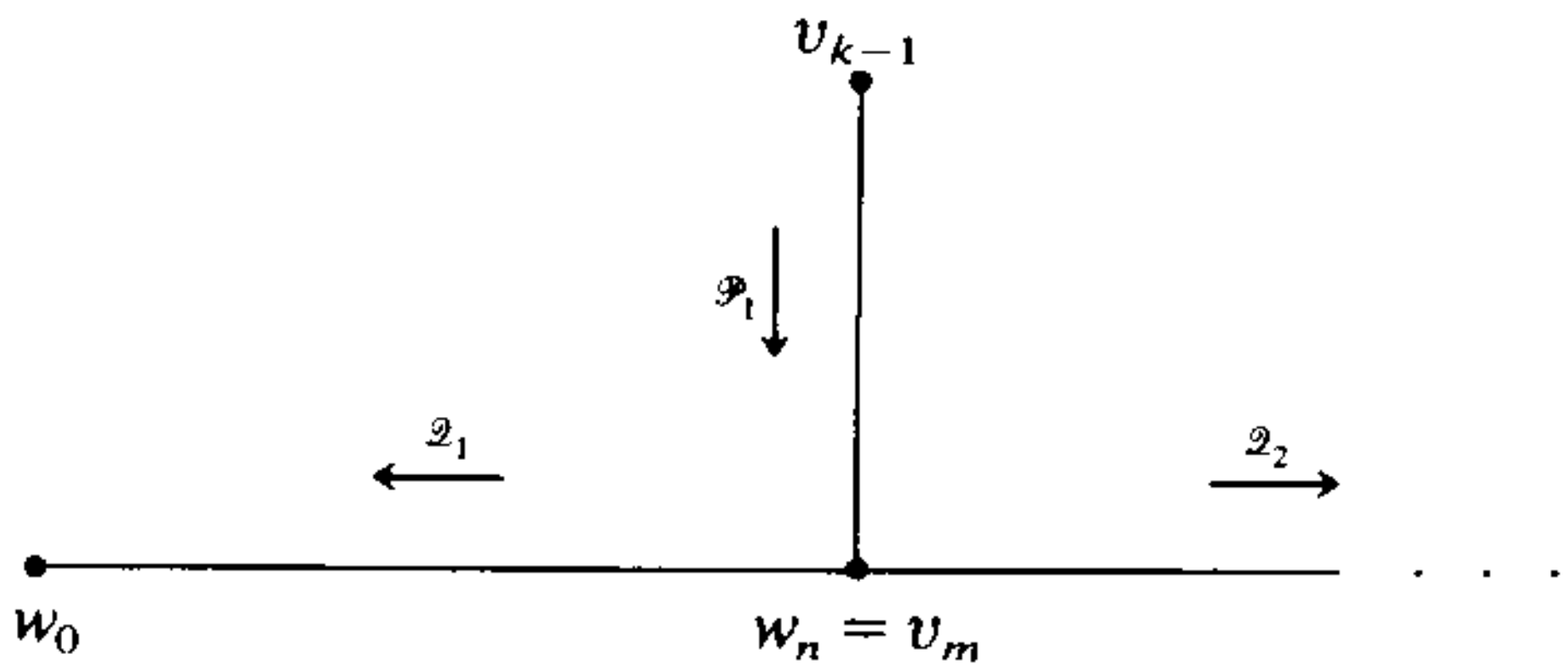
(ii) $\{v_0, v_1\} \in F^C$ *and* $d_F(v_0) < f(v_0)$.

Then $F \oplus E(\mathcal{P})$ is an optimal factor; further $k < \omega$ and $\delta_{F \oplus E(\mathcal{P})} = \delta_{Fv_0}^{v_{k-1}}$.

In other words: If an optimal factor is improved by an alternating trail at its startpoint it remains optimal, but it is debased at its endpoint.

Proof. Since F is optimal, $F \oplus E(\mathcal{P})$ is not better than F and therefore we conclude $k < \omega$ and $\delta_{F \oplus E(\mathcal{P})} = \delta_{Fv_0}^{v_{k-1}}$. To prove that $F \oplus E(\mathcal{P})$ is optimal, we assume for a contradiction that H is a factor better than $F \oplus E(\mathcal{P})$. Then there is some $w_0 \in V$ such that $\delta_H(w_0) < \delta_{F \oplus E(\mathcal{P})}(w_0)$. If $d_{F \oplus E(\mathcal{P})}(w_0) < d_H(w_0)$. let $\mathcal{Q} = (w_i)_{0 \le i < l}$ be an H-$(F \oplus E(\mathcal{P}))$-trail, and if $d_{F \oplus E(\mathcal{P})}(w_0) > d_H(w_0)$, let $\mathcal{Q} = (w_i)_{0 \le i < l}$ be an $(F \oplus E(\mathcal{P}))$-$H$-trail. We have $V(\mathcal{P}) \cap V(\mathcal{Q}) \neq \emptyset$; otherwise $d_F(v) = d_{F \oplus E(\mathcal{P})}(v)$ for each $v \in V(\mathcal{Q})$, $\mathcal{Q}$ would be an H-F-trail (an F-H-trail, resp.), and, by Lemma 2, we would get $\delta_{F \oplus E(\mathcal{Q})} = \delta_{Fw_0w_{l-1}}$, if $l < \omega$, and $\delta_{F \oplus E(\mathcal{Q})} = \delta_{Fw_0}$, if $l = \omega$. Both contradict the assumption that F is optimal. Consequently there exists $m := \max\{i : v_i \in V(\mathcal{Q})\}$. Choose n such that $w_n = v_m$ and let

$$\mathcal{Q}_1 := (w_{n-i})_{0 \le i \le n}, \qquad \mathcal{Q}_2 := (w_i)_{n \le i < l}, \qquad \mathcal{P}_1 := (v_{k-i})_{1 \le i \le k-m}.$$



Since F is optimal and, by Lemma 2,

$$\delta_{F \oplus E(\mathcal{P}) \oplus E(\mathcal{Q})} = \delta_{F \oplus E(\mathcal{P}) w_0 w_{l-1}} = \delta_{Fv_0w_0w_{l-1}}^{v_{k-1}} \quad \text{if } l < \omega,$$

$$\delta_{F \oplus E(\mathcal{P}) \oplus E(\mathcal{Q})} = \delta_{F \oplus E(\mathcal{P}) w_0} = \delta_{Fv_0w_0}^{v_{k-1}} \qquad \text{if } l = \omega,$$

we conclude that $v_{k-1} \notin \{w_0, w_{l-1}\}$. Therefore both $\widehat{\mathcal{P}_1\mathcal{Q}_1}$ and $\widehat{\mathcal{P}_1\mathcal{Q}_2}$ have positive length. We leave it to the reader to show that either $\mathcal{R} := \widehat{\mathcal{P}_1\mathcal{Q}_1}$ or $\mathcal{R} := \widehat{\mathcal{P}_1\mathcal{Q}_2}$ is an $(F \oplus E(\mathcal{P}))$-alternating trail such that $F \oplus E(\mathcal{P}) \oplus E(\mathcal{R})$ is a factor better than F. This is a contradiction to the fact that F is optimal. □

The following lemma will be applied in the proof of Theorem 1 and in Section 4.

Lemma 4. *Let F be a factor. Further let H be a factor such that at most one of the sets*

$$V_1(H) := \{v \in V : \delta_F(v) < \delta_H(v)\},$$
$$V_2(H) := \{v \in V : \delta_F(v) > \delta_H(v)\}$$

is infinite. Define

$$\vartheta(H) := \begin{cases} +\infty & \text{if } V_1(H) \text{ is infinite,} \\ -\infty & \text{if } V_2(H) \text{ is infinite,} \\ \sum_{v \in V} \delta_H(v) - \delta_F(v) & \text{else.} \end{cases}$$

If F is optimal, then $\vartheta(H) \geqslant 0$.

Proof. Let F be optimal. Suppose that there is a factor H with the property that at least one of the sets $V_1(H)$, $V_2(H)$ is finite and $\vartheta(H) < 0$. Then $\vartheta(H) \neq +\infty$ and therefore $V_1(H)$ is finite and

$$\vartheta_1(H) := \sum_{v \in V_1(H)} \delta_H(v) - \delta_F(v) = \sum_{v \in V \setminus V_2(H)} \delta_H(v) - \delta_F(v)$$

is nonnegative integer. Choose H such that $\vartheta_1(H)$ is minimal. Since $\vartheta(H) < 0$, we can choose $v_0 \in V_2(H)$. Then $d_H(v_0) \neq d_F(v_0)$.

Case 1: $d_H(v_0) < d_F(v_0)$.

Let $\mathscr{P} = (v_i)_{0 \leqslant i < k}$ be an F-H-trail. Since $v_0 \in V_2(H)$ and $d_H(v_0) < d_F(v_0)$, we have $d_F(v_0) > f(v_0)$ and hence $k < \omega$ by Lemma 3. Since F is optimal, we get $v_{k-1} \in V_1(H)$ by Corollary 1(a). In particular, $v_0 \neq v_{k-1}$ and from Lemma 2(d) follows

$$\delta_{H \oplus E(\mathscr{P})}(v_{k-1}) = \delta_H(v_{k-1}) - 1. \tag{1}$$

Therefore $v_{k-1} \notin V_2(H \oplus E(\mathscr{P}))$. Further it follows from

$$\delta_{H \oplus E(\mathscr{P})}(v_0) = \delta_H(v_0) \pm 1 \tag{2}$$

and $v_0 \in V_2(H)$, that $v_0 \notin V_1(H \oplus E(\mathscr{P}))$. Since, for each $v \notin \{v_0, v_{k-1}\}$, $v \in V_i(H \oplus E(\mathscr{P}))$ iff $v \in V_i(H)$ $(i = 1, 2)$, at most one of the sets $V_1(H \oplus E(\mathscr{P}))$, $V_2(H \oplus E(\mathscr{P}))$ is infinite and from (1) and (2) we get $\vartheta(H \oplus E(\mathscr{P})) \leqslant \vartheta(H) < 0$ and $\vartheta_1(H \oplus E(\mathscr{P})) = \vartheta_1(H) - \delta_H(v_{k-1}) + \delta_{H \oplus E(\mathscr{P})}(v_{k-1}) = \vartheta_1(H) - 1$. This contradicts the minimality of $\vartheta_1(H)$.

Case 2: $d_H(v_0) > d_F(v_0)$.

Analogously we can construct a contradiction. □

The following theorem characterizes in some cases the optimal factors of G—e.g. if G is finite.

Theorem 1. *Define*

$$\Delta_0 := \min\{\Delta(F, f) : F \subseteq E \text{ is a factor}\}.$$

If Δ_0 is finite, then a factor F is optimal iff $\Delta(F, f) = \Delta_0$.

Proof. We only proof the nontrivial implication. Let F be an optimal factor. We have to show that $\Delta(F, f) \leqslant \Delta_0$. Choose a factor H satisfying $\Delta(H, f) = \Delta_0$. Since $\Delta_0 \in \omega$, the set $V_1(H)$ from Lemma 4 is finite. By Lemma 4, we get $0 \leqslant \sum_{v \in V} \delta_H(v) - \delta_F(v) = \Delta_0 - \Delta(F, f)$. Thus $\Delta(F, f) \leqslant \Delta_0$. □

Corollary 2. *If G is a finite graph, then a factor F is optimal iff it is optimal in the sense of Lovász and Plummer* [4], *i.e.* $\Delta(F, f) = \Delta_0$.

3. f-F-trails, f-obstructions, and stable f-factors

In this section we are going to examine f-factors. We give a necessary and sufficient condition for the existence of an f-perfect factor if G is a countable graph (Theorem 3). This criterion is a generalization of Steffens' criterion for the existence of an 1-perfect factor of a countable graph [7]. By the aid of this criterion Steffens proves in [7] for countable graphs the existence of an $\subseteq$-maximal matchable set $V' \subseteq V$ or, in our terminology, the existence of an 1-optimal 1-factor. We want to generalize this for f-factors, too (Theorem 6). The role of 'F-augmenting pathes' in [7] will be played in this paper by the 'f-F-trails' which have been introduced by Steffens in [8] (and are called 'F-augmenting trails' there).

Let F be an f-factor of G and let $\mathscr{P} = (v_i)_{0 \leqslant i < k}$ be a trail satisfying $k > 1$. $\mathscr{P}$ is called an *f-F-trail* if the following properties are fulfilled:

(i) $\mathscr{P}$ is F-alternating,
(ii) $d_F(v_0) < f(v_0)$,
(iii) $\{v_0, v_1\} \in F^C$,
(iv) if $k < \omega$, then $\{v_{k-2}, v_{k-1}\} \in F^C$ and $d_F(v_{k-1}) < f_{v_0}(v_{k-1})$.

It is easy to see that the following holds.

Lemma 5. *Let F be an f-factor and let $\mathscr{P} = (v_i)_{0 \leqslant i < k}$ be an f-F-trail. Then $F \oplus E(\mathscr{P})$ is an f-factor better than F. If $k < \omega$, then $d_{F \oplus E(\mathscr{P})} = d_F^{v_0 v_{k-1}}$, and if $k = \omega$, then $d_{F \oplus E(\mathscr{P})} = d_F^{v_0}$.*

An easy consequence of Lemma 1 is the following.

Lemma 6. *Let F_1 be an f-factor better than an f-factor F_2 and let $v_0 \in V$ such that $d_{F_2}(v_0) < d_{F_1}(v_0)$. Then there exists an f-F_2-trail starting at v_0.*

Proof. Choose an F_1-F_2-trail starting at v_0. This is an f-F_2-trail. □

Let F be an f-factor and let v_0 be an unsaturated vertex such that there is no f-F-trail starting at v_0. Then (F, v_0) is called an *(f-)obstruction*. If there is no f-obstruction in G, G is called *(f-)unobstructed*. The condition of being

unobstructed corresponds to '*condition* (A)' for 1-factors in [7]. Since a perfect factor is better than any nonperfect factor, we have the following by Lemma 6.

Theorem 2. *If G possesses a perfect factor, then G is unobstructed.*

The converse of this theorem is trivial for finite graphs and false for arbitrary graphs. This is shown again by $K_{\aleph_0, \aleph_1}$ or G^*, described in Section 1. For countable graphs however the converse holds (Theorem 3). To show this, we need the concept of 'stable factors' which corresponds to that of 'independent matchings' in [7]. An f-factor F is called (f-)*stable* if there is no f-F-trail traversing edges of F. So, an f-factor is stable iff each f-F-trail has length 1. Note that $\emptyset$ and each perfect factor are stable. In the next lemmas we want to collect some properties of stable f-factors and obstructions.

Lemma 7. *Let F be a stable f-factor, let $H \supseteq F$ be an f-factor, and let $\mathcal{P}$ be an f-H-trail. Then $\mathcal{P}$ is an $(f - d_F)$-$(H \setminus F)$-trail in the graph $G \setminus F := (V, E \setminus F)$.*

Proof. Let $\mathcal{P} = (v_i)_{0 \leq i < k}$. Since $d_{H \setminus F} = d_H - d_F \leq f - d_F$, $H \setminus F$ is an $(f - d_F)$-factor of $G \setminus F$, and a vertex is f-H-saturated iff it is $(f - d_F)$-$(H \setminus F)$-saturated. It remains to show that there is no edge in $F \cap E(\mathcal{P})$. Suppose for a contradiction that $F \cap E(\mathcal{P}) \neq \emptyset$. Without loss of generality we can assume that $\{v_1, v_2\} \in F$; otherwise consider $H' = H \setminus \{\{v_{s-2}, v_{s-1}\}\}$ and $\mathcal{P}' = (v_i)_{s-1 \leq i < k}$, where $s := \min\{i: \{v_i, v_{i+1}\} \in F\}$. If there is some edge in $(H \setminus F) \cap E(\mathcal{P})$, let $t := \min\{i: \{v_i, v_{i+1}\} \in H \setminus F\}$. Otherwise let $t := k$. In both cases $(v_i)_{0 \leq i < t}$ is an f-F-trail which traverses an edge of F. This is a contradiction to the stability of F. $\square$

Lemma 8. *Let F be an f-stable f-factor of G and let $J \subseteq F^C$ be an $(f - d_F)$-stable $(f - d_F)$-factor of the graph $G \setminus F$. Then $F \cup J$ is an f-stable f-factor of G.*

Proof. Since $d_{F \cup J} = d_F + d_J \leq d_F + (f - d_F) = f$, $F \cup J$ is an f-factor. Let $\mathcal{P}$ be an f-$(F \cup J)$-trail. We have to show that there is no $(F \cup J)$-edge in $E(\mathcal{P})$. Because of Lemma 7, $\mathcal{P}$ is an $(f - d_F)$-J-trail in $G \setminus F$. Thus $\mathcal{P}$ traverses no edge of F. Furthermore, $\mathcal{P}$ traverses no edge of J by the stability of J. $\square$

Lemma 9. *If G is f-unobstructed and if F is a stable f-factor of G, then $G \setminus F$ is $(f - d_F)$-unobstructed.*

Proof. Let J be an $(f - d_F)$-factor of $G \setminus F$ and let $v_0 \in V$ be an $(f - d_F)$-J-unsaturated vertex. Since $J \cup F$ is an f-factor of G and v_0 is f-$(J \cup F)$-unsaturated and G is f-unobstructed, there exists an f-$(J \cup F)$-trail starting at v_0. Because of Lemma 7, this is already an $(f - d_F)$-J-trail in $G \setminus F$. Therefore (J, v_0) is no $(f - d_F)$-obstruction. $\square$

Lemma 10. *G contains an $\subseteq$-maximal stable f-factor.*

Proof. Let $\mathcal{M}$ be the set of stable f-factors of G. By Lemma 7, $\bigcup_{\lambda<\kappa} F_\lambda$ is a stable f-factor whenever $(F_\lambda)_{\lambda<\kappa}$ is a $\subseteq$-chain of stable f-factors. Since $\mathcal{M} \neq \emptyset$ ($\emptyset \in \mathcal{M}$) the assertion follows by Zorn's Lemma. □

Note that each optimal f-factor is $\subseteq$-maximal stable. The converse is false as $K_{\aleph_0,\aleph_1}$ and $\emptyset$ show. We will see later that the equivalence holds for countable graphs. Because of Lemma 8 we have the following.

Lemma 11. *Let F be an $\subseteq$-maximal f-stable f-factor of G. Then $\emptyset$ is the only $(f-d_F)$-stable $(f-d_F)$-factor of $G\backslash F$.*

In some cases, the following two lemmas can give us information about the existence of stable f-factors which are non-empty.

Lemma 12. *Let H be an f-factor of G and let $x_0 \in V$. We define $\mathcal{A} := \{\mathcal{P}$: $\mathcal{P}$ is an H-alternating trail starting at x_0 with an H-edge$\}$ and $F := \{e \in H$: There exists a trail $\mathcal{P} \in \mathcal{A}$ terminating at a vertex incident with $e\}$. If $\mathcal{A}$ contains no trail of infinite length and if no trail $\mathcal{P} \in \mathcal{A}$ terminates with an H^C-edge at an f-H-unsaturated vertex, then F is a stable f-factor.*

Proof. Suppose F is not stable. Then there exists an f-F-trail $\mathcal{Q} = (w_i)_{0\leq i<k}$ traversing an edge $e = \{x, y\} \in F$. By definition of F, there exists a trail $\mathcal{P} = (x_i)_{0\leq i\leq l} \in \mathcal{A}$ with endpoint $x_l = x$, without loss of generality.

Case 1: $\{x_0, x_1\} = \{w_s, w_{s+1}\}$ *for some* s.

If $x_0 = w_s$ let $\mathcal{R} := (w_i)_{s\leq i<k}$, and if $x_0 = w_{s+1}$ let $\mathcal{R} := (w_{s+1-i})_{0\leq i\leq s+1}$. $\mathcal{R}$ is an F-alternating trail starting at x_0 with an H-edge. $\mathcal{R}$ has infinite length or terminates with an F^C-edge at an F-unsaturated vertex. For the vertices $x_0 = w_s, w_{s+1}, w_{s+2}, \ldots,$ resp. $x_0 = w_{s+1}, w_s, w_{s-1}, \ldots$ it follows successively by the definition of F:

$$\{w_i, v\} \in H \text{ iff } \{w_i, v\} \in F.$$

So $\mathcal{R}$ is an H-alternating trail which starts at x_0 with an H-edge and has either infinite length or terminates with an H^C-edge at an H-unsaturated vertex. This is a contradiction to the assumption on $\mathcal{A}$.

Case 2: $\{x_0, x_1\} \notin E(\mathcal{Q})$.

Let $r := \min\{i \geq 1: x_i \in V(\mathcal{Q})\}$ and define

$$\mathcal{P}_1 := (x_i)_{0\leq i\leq r}, \qquad \mathcal{Q}_1 := (w_i)_{s\leq i<k}, \qquad \mathcal{Q}_2 := (w_{s-i})_{0\leq i\leq s},$$

where s is any index such that $w_s = x_r$. If either $\{x_{r-1}, x_r\} \in H$ and $s = 0$, or $\{x_{r-1}, x_r\} \in H$, $s > 0$, and $\{w_{s-1}, w_s\} \in F$, or $\{x_{r-1}, x_r\} \in H^C$, $s > 0$, and $\{w_{s-1}, w_s\} \in F^C$, let $\mathcal{R} := \widehat{\mathcal{P}_1\mathcal{Q}_1}$. Otherwise let $\mathcal{R} := \widehat{\mathcal{P}_1\mathcal{Q}_2}$. As in case 1 $\mathcal{R}$ is an H-alternating trail starting with an H-edge at x_0, and $\mathcal{R}$ has either infinite length or terminates with an H^C-edge at an f-H-unsaturated vertex. Again this is a contradiction to the assumption on $\mathcal{A}$. □

Lemma 13. *Let H be an f-factor and $x_0 \in V$. We define $\mathscr{B} := \{\mathscr{P}$: $\mathscr{P}$ is an H-alternating trail which starts at x_0, but not with an H-edge$\}$, $F := \{e \in H$: There exists a trail $\mathscr{P} \in \mathscr{B}$ terminating at a vertex incident with $e\}$. If $\mathscr{B}$ contains no trail of infinite length and if no trail $\mathscr{P} \in \mathscr{B}$ terminates with an H^C-edge at a vertex w satisfying $d_H(w) < f_{x_0}(w)$, then F is a stable f-factor.*

Note that the F in Lemma 12 and in Lemma 13 contains all H-edges incident with x_0.

Proof. As in the proof of Lemma 12 we assume that there is an f-F-trail $\mathscr{Q} = (w_i)_{0 \leq i < k}$ which traverses an edge $e = \{x, y\} \in F$. We conclude that $x_0 \notin \{w_0, w_{k-1}\}$ since otherwise $\mathscr{Q}$ (resp. $(w_{k-i})_{0 < i \leq k}$) would be an f-H-trail in contradiction to the assumption on $\mathscr{B}$. By the definition of F, there exists $\mathscr{P} = (x_i)_{0 \leq i \leq l} \in \mathscr{B}$ with endpoint $x_l = x$, without loss of generality. Let $r := \min\{i \geq 0: x_i \in V(\mathscr{Q})\}$ and define

$$\mathscr{P}_1 := (x_i)_{0 \leq i \leq r}, \qquad \mathscr{Q}_1 := (w_i)_{s \leq i < k}, \qquad \mathscr{Q}_2 := (w_{s-i})_{0 \leq i \leq s},$$

where s is any index satisfying $w_s = x_r$. As in the proof of Lemma 12, either $\mathscr{R} := \widehat{\mathscr{P}_1 \mathscr{Q}_1}$ or $\mathscr{R} := \widehat{\mathscr{P}_1 \mathscr{Q}_2}$ is an H-alternating trail starting at x_0 with an H^C-edge which has infinite length or terminates with an H^C-edge at an f-H-unsaturated vertex, which is f_{x_0}-H-unsaturated too since $x_0 \notin \{w_0, w_{k-1}\}$. This is a contradiction to the assumption on $\mathscr{B}$. $\square$

Lemma 14. *Let $x_0 \in V$ such that $f(x_0) > 0$. If G is f-unobstructed and if G contains no non-empty stable f-factor, then the following holds:*

(a) *G is f_{x_0}-unobstructed.*

(b) *G is f^{x_0}-unobstructed.*

Proof. (a) Suppose that there is some f_{x_0}-obstruction (J, v_0). Since G is f-unobstructed, there exists an f-J-trail $\mathscr{P} = (v_i)_{0 \leq i < k}$. We have $k < \omega$ and $d_J(v_{k-1}) \geq f_{x_0 v_0}(v_{k-1})$ because otherwise $\mathscr{P}$ still would be an f_{x_0}-J-trail. Since $d_J(v_{k-1}) < f_{v_0}(v_{k-1})$, we get $v_{k-1} = x_0$ and

$$d_J(x_0) = f_{v_0}(x_0) - 1. \tag{1}$$

Let $H := J \oplus E(\mathscr{P})$. H is an f-factor and so we conclude that

$$d_H = d_J^{v_0 v_{k-1}} \leq f. \tag{2}$$

By (1) and (2) we get

$$d_H(x_0) = f(x_0).$$

Now we consider the set $\mathscr{A}$ from Lemma 12. Suppose there is a trail $\mathscr{Q} = (x_i)_{0 \leq i < l} \in \mathscr{A}$ satisfying $l = \omega$ or $l < \omega$, $\{x_{l-2}, x_{l-1}\} \in H^C$, and $d_H(x_{l-1}) < f(x_{l-1})$.

Case 1: $l < \omega$.

Then $x_{l-1} \neq x_0$ and, by (2), $d_{H \oplus E(\mathcal{Q})} = d_{Hx_0}^{x_{l-1}} \leqslant f_{x_0}$. Further $d_J < d_J^{v_0 x_{l-1}} = d_{Hx_0}^{x_{l-1}} = d_{H \oplus E(\mathcal{Q})}$ by (2). Therefore $H \oplus E(\mathcal{Q})$ is an f_{x_0}-factor better than J and $d_J(v_0) < d_{H \oplus E(\mathcal{Q})}(v_0)$.

Case 2: $l = \omega$.

Similarly to case 1 one can prove that $H \oplus E(\mathcal{Q})$ is an f_{x_0}-factor better than J and $d_J(v_0) < d_{H \oplus E(\mathcal{Q})}(v_0)$.

Now Lemma 6 yields an f_{x_0}-J-trail starting at v_0. This is a contradiction to the fact that (J, v_0) is an f_{x_0}-obstruction. Therefore the conditions of Lemma 12 are fulfilled and we obtain an f-stable f-factor F. F is not empty since $d_H(x_0) = f(x_0) > 0$. This contradicts the requirement of the lemma.

(b) Suppose that there is an f^{x_0}-obstruction (H, v_0).

Case 1: $v_0 = x_0$.

Then $d_H(x_0) = f(x_0)$ since otherwise (H, x_0) would be an f-obstruction, too. Thus, the members of the set $\mathcal{B}$ in Lemma 13 which have infinite length or terminate with an H^C-edge in an f-H-unsaturated vertex are exactly the f^{x_0}-trails starting at x_0. Therefore, since (H, x_0) is an f^{x_0}-obstruction, the conditions on the set $\mathcal{B}$ in Lemma 13 are fulfilled. By Lemma 13 and $d_H(x_0) = f(x_0) > 0$, we obtain a non-empty f-stable f-factor. This contradicts the requirement of the lemma.

Case 2: $v_0 \neq x_0$.

Then $d_H(x_0) = f(x_0) + 1$ because otherwise (H, v_0) would be an f-obstruction, too. Choose $e = \{y, x_0\} \in H$ and define $J := H \setminus \{e\}$. J is an f-factor. By interated application of the fact that G is f-unobstructed and of Lemma 5, we obtain an f-factor K which is better than J and saturates v_0 and y. K is an f^{x_0}-factor and x_0 is f^{x_0}-unsaturated. There is no f^{x_0}-K-trail $\mathcal{P} = (x_i)_{0 \leqslant i < k}$, since otherwise $K \oplus E(\mathcal{P})$ is an f^{x_0}-factor better than H which saturates v_0 and Lemma 6 yields an f^{x_0}-H-trail starting at v_0, contrary to the assumption that (H, v_0) is an f^{x_0}-obstruction. Hence (K, x_0) is an f^{x_0}-obstruction. This is impossible by case 1. □

Lemma 15. *Let $x_0 \in V$. If G is f-unobstructed, then there exists an f-factor F saturating x_0 such that $G \setminus F$ is $(f - d_F)$-unobstructed.*

Proof. The lemma is proved by induction on $f(x_0)$. If $f(x_0) = 0$, let $F := \emptyset$. Now let $f(x_0) > 0$. By Lemma 9, 10, and 11, we can assume without loss of generality that G contains no non-empty stable f-factor. By the induction hypothesis and Lemma 14(a), we obtain an f_{x_0}-factor H such that $G \setminus H$ is $(f_{x_0} - d_H)$-unobstructed and $d_H(x_0) = f_{x_0}(x_0) = f(x_0) - 1$. Again we can assume without loss of generality that $G \setminus H$ contains no non-empty stable $(f_{x_0} - d_H)$-factor. Since x_0 is f-H-unsaturated and G is f-unobstructed, there exists an f-H-trail $\mathcal{Q} = (x_i)_{0 \leqslant i < k}$. If there is some $i < k$ such that x_i is f_{x_0}-H-unsaturated, let

$$r := \min\{i \colon d_H(x_i) < f_{x_0}(x_i)\} + 1.$$

Otherwise let $r := \omega$. Note that $r > 1$ and $x_0 \neq x_{r-1}$. Let

$$\mathcal{Q}_1 := (x_i)_{0 \leqslant i < r} \quad \text{and} \quad F := H \oplus E(\mathcal{Q}_1).$$

F is an f-factor which saturates x_0. We will show that $G\backslash F$ is $(f-d_F)$-unobstructed. For this let K be an $(f-d_F)$-factor of $G\backslash F$ and let v_0 be an $(f-d_F)$-K-unsaturated vertex. We must find an $(f-d_F)$-K-trail in $G\backslash F$ starting at v_0. Since

$$d_K(x_i)=0 \quad \text{for all } i<r-1,$$

it follows that $K\cap E(\mathcal{Q}_1)=\emptyset$. Hence K is a factor of $G\backslash H$.

Case 1: $r=\omega$.

Then $f-d_F=f-d_H^{x_0}=f_{x_0}-d_H$. Thus K is an $(f_{x_0}-d_H)$-factor of $G\backslash H$ and v_0 is $(f_{x_0}-d_H)$-K-unsaturated. Since $G\backslash H$ is $(f_{x_0}-d_H)$-unobstructed, there exists an $(f_{x_0}-d_H)$-K-trail $\mathcal{P}=(v_i)_{0\le i<k}$ in $G\backslash H$. Since $d_K(x_i)=0$ for all $i<\omega$, we conclude that $E(\mathcal{P})\cap E(\mathcal{Q})=\emptyset$. Thus $\mathcal{P}$ is an $(f-d_F)$-K-trail in $G\backslash F$.

Case 2: $r<\omega$ *and* $\{x_{r-2},x_{r-1}\}\in H$.

Then $f-d_F=f-d_{Hx_{r-1}}^{x_0}=(f_{x_0}-d_H)^{x_{r-1}}$. Thus K is an $(f_{x_0}-d_H)^{x_{r-1}}$-factor of $G\backslash H$ and v_0 is $(f_{x_0}-d_H)^{x_{r-1}}$-K-unsaturated. Since $f_{x_0}(x_{r-1})-d_H(x_{r-1})>0$, $G\backslash H$ is $(f_{x_0}-d_H)^{x_{r-1}}$-unobstructed by Lemma 14(b). Consequently there exists an $(f_{x_0}-d_H)^{x_{r-1}}$-K-trail $\mathcal{P}=(v_i)_{0\le i<k}$ in $G\backslash H$. Since $d_K(x_i)=0$ for all $i<r-1$, we have again $E(\mathcal{P})\cap E(\mathcal{Q}_1)=\emptyset$ and so $\mathcal{P}$ is an $(f-d_F)$-K-trail in $G\backslash F$.

Case 3: $r<\omega$ and $\{x_{r-2},x_{r-1}\}\in H^C$.

Similarly to case 2 we find with the help of Lemma 14(a) an $(f_{x_0}-d_H)_{x_{r-1}}$-K-trail in $G\backslash H$ starting at v_0 which is an $(f-d_F)$-K-trail in $G\backslash F$. □

Now we are able to prove the converse of Theorem 2 for countable graphs.

Theorem 3. *If G is countable, G possesses a perfect f-factor iff G is f-unobstructed.*

Proof. Let $V=\{v_i: i<\omega\}$. By Lemma 15 we are able to construct recursively f-factors $F_0\subseteq F_1\subseteq F_2\subseteq\cdots$ such that, for $i<\omega$, the vertices $v_0,\ldots,v_i$ are saturated by F_i and $G\backslash F_i$ is $(f-d_{F_i})$-unobstructed. Then $F:=\bigcup_{i<\omega}F_i$ is a perfect f-factor. □

The following theorem and its two corollaries are generalizations of Aharoni's Theorems 4, 3, and 1 in [1].

Theorem 4. *Let E be the disjoint union of E_1 and E_2, and let $f_1,f_2:V\to\omega$ be functions satisfying $f_1+f_2=f$ such that $G_1:=(V,E_1)$ is f_1-unobstructed and $G_2:=(V,E_2)$ is f_2-unobstructed. Then G is f-unobstructed.*

Proof. Let $F\subseteq E$ be an f-factor and v_0 be an f-F-unsaturated vertex. We have to find an f-F-trail starting at v_0. Let us define recursively a sequence $V_0\subseteq V_1\subseteq V_2\subseteq\cdots$ of finite subsets of V and, for $j=1,2$, sequences $F_0^{(j)}\subseteq F_1^{(j)}\subseteq F_2^{(j)}\subseteq\cdots$ of f_j-factors of G_j such that, for all $i<\omega$, each $x\in V_i$ is f_j-$F_i^{(j)}$-saturated and $G_j\backslash F_i^{(j)}$ is $(f_j-d_{F_i^{(j)}})$-unobstructed.

Let $V_0 := \{v_0\}$ and choose $F_0^{(1)}$, $F_0^{(2)}$ in agreement with Lemma 15. Now suppose that V_i, $F_i^{(1)}$, and $F_i^{(2)}$ have been defined. Let

$$V_{i+1} := \{x \in V : \exists y \in V_i : \{x, y\} \in F \cup F_i^{(1)} \cup F_i^{(2)}\} \cup V_i.$$

By induction hypothesis and since F, $F_i^{(1)}$, and $F_i^{(2)}$ are factors, V_{i+1} is finite. By iterated application of Lemma 15 and induction hypothesis, we obtain an $(f_j - d_{F_i^{(j)}})$-factor H_j saturating in $G_j \setminus F_i^{(j)}$ all vertices of V_{i+1} such that $G_j \setminus (F_i^{(j)} \cup H_j)$ is $(f_j - d_{F_i^{(j)}} - d_{H_j})$-unobstructed $(j = 1, 2)$. Let $F_{i+1}^{(j)} := F_i^{(j)} \cup H_j$.

For $j = 1, 2$, define $F_j := \bigcup_{0 \leq i < \omega} F_i^{(j)}$ and $H := F_1 \cup F_2$. H is an f-factor of G which saturates each vertex $x \in \bigcup_{0 \leq i < \omega} V_i$. Now let $\mathscr{P} = (v_i)_{0 \leq i < k}$ be an H-F-trail. It is easy to see that, for every $i < k$, $v_i \in V_i$ and therefore $d_H(v_i) = f(v_i)$. If either $k = \omega$ or $k < \omega$, $\{v_{k-2}, v_{k-1}\} \in H \setminus F$ and $d_H(v_{k-1}) > d_F^{v_0}(v_{k-1})$, then $\mathscr{P}$ is an f-F-trail and the lemma is proved. The case $k < \omega$, $\{v_{k-2}, v_{k-1}\} \in F \setminus H$, and $d_H(v_{k-1}) < d_F^{v_0}(v_{k-1})$ cannot occur since $d_F(v_{k-1}) \leq f_{v_0}(v_{k-1}) = d_{Hv_0}(v_{k-1})$. □

Corollary 3. *If G is f-unobstructed and if $G^* = (V, E^*)$ is a graph such that $E^* \supseteq E$, then G^* is f-unobstructed.*

Proof. Let $E_1 := E$, $E_2 := E^* \setminus E$, $f_1 := f$, $f_2 := 0$. □

Corollary 4. *If G is f-unobstructed and if x is a vertex satisfying $f(x) > 0$, then there exists a vertex y adjacent to x such that $f(y) > 0$ and $G \setminus \{\{x, y\}\}$ is f_{xy}-unobstructed.*

Proof. By Lemma 15, there exists an f-factor F such that $d_F(x) = f(x)$ and $G \setminus F$ is $(f - d_F)$-unobstructed. Choose $y \in V$ such that $\{x, y\} \in F$. Now let $E_1 := E \setminus F$, $E_2 := F \setminus \{\{x, y\}\}$, $f_1 := f - d_F$, $f_2 := d_{Fxy}$. □

4. Optimal f-factors

We will prove that G possesses an optimal f-factor if G is countable. For that we must prove the following lemma.

Lemma 16. *Let J be an $\subseteq$-maximal stable f-factor. Further let*

$$W := \{v \in V : d_J(v) < f(v) \wedge \exists w \in V (d_J(w) < f(w) \wedge \{v, w\} \in E \setminus J)\},$$

$$f^*(v) := \begin{cases} f(v) - d_J(v) & \text{if } v \in W, \\ 0 & \text{if } v \in V \setminus W. \end{cases}$$

Then $G \setminus J$ is f^*-unobstructed.

Proof. Let H be an f^*-factor of $G \setminus J$ and let $x_0 \in V$ be a vertex satisfying $d_H(x_0) < f^*(x_0)$. We have to find an f^*-H-trail in $G \setminus J$ starting at x_0. Consider the

set $\mathscr{B}$ and the factor F from Lemma 13, referring to the graph $G\backslash J$ and the function $f-d_J$. Since $x_0 \in W$, there exists $w \in W$ such that $\{x_0, w\} \in E\backslash J$. If $\{x_0, w\} \in H^C$ and $d_H(w) < f(w) - d_J(w)$, then (x_0, w) is an f^*-H-trail starting at x_0, and the lemma is proved. If on the other hand $\{x_0, w\} \in H$ or $d_H(w) = f(w) - d_J(w)$, then it follows from the definition of F that F is not empty. Hence F cannot be $(f-d_J)$-stable by Lemma 11. Therefore we conclude from Lemma 13 that $\mathscr{B}$ contains a trail $\mathscr{P}$ which has infinite length or terminates with an H^C-edge at an $(f-d_J)_{x_0}$-H-unsaturated vertex. So $\mathscr{P}$ is an $(f-d_J)$-H-trail in $G\backslash J$ starting at x_0. Clearly $\mathscr{P}$ cannot meet a vertex in $V\backslash W$. Thus $\mathscr{P}$ is an f^*-H-trail. □

Theorem 5. *Let F be an f-factor of G such that there is no f-factor better than F. Then F is already an optimal factor.*

Proof. Suppose there is some factor H better than F. Let $v_0 \in V$ such that $\delta_H(v_0) < \delta_F(v_0)$. Then we have $d_F(v_0) < d_H(v_0)$. Let $\mathscr{P} = (v_i)_{0 \leq i < k}$ be an H-F-trail. $F \oplus E(\mathscr{P})$ is better than F by Corollary 1(b). Thus $F \oplus E(\mathscr{P})$ is not an f-factor and therefore we get $k < \omega$ and

$$d_{F \oplus E(\mathscr{P})}(v_{k-1}) = d_F^{v_0 v_{k-1}}(v_{k-1}) > f(v_{k-1}).$$

Since $d_F(v_{k-1}) \leq f(v_{k-1})$ and $\delta_{F \oplus E(\mathscr{P})}(v_{k-1}) \leq \delta_F(v_{k-1})$, it follows that $v_0 = v_{k-1}$, $d_F(v_0) = f(v_0) - 1$, and $d_{F \oplus E(\mathscr{P})}(v_0) = f(v_0) + 1$. Hence $\delta_{F \oplus E(\mathscr{P})} = \delta_F$, and $F \oplus E(\mathscr{P})$ is not better than F. This is a contradiction. □

Theorem 6. *If G is countable, then G possesses an f-optimal factor. Moreover, each $\subseteq$-maximal stable f-factor is optimal.*

Proof. Let J be an $\subseteq$-maximal stable f-factor. Let W and f^* be defined as in Lemma 16. By Theorem 3 and Lemma 16, there exists a perfect f^*-factor H of $G\backslash J$. Assume that there is some f-$(J \cup H)$-trail $\mathscr{P} = (v_i)_{0 \leq i < k}$ in G. By Lemma 7, $\mathscr{P}$ is an $(f-d_J)$-H-trail in $G\backslash J$. Therefore we have $d_J(v_1) < f(v_1)$ and of course $d_J(v_0) < f(v_0)$ and $\{v_0, v_1\} \in J^C$. Thus $v_0 \in W$. Since $d_H(v_0) < f(v_0) - d_J(v_0) = f^*(v_0)$, we have a contradiction to the fact that H is a perfect f^*-factor.

Consequently, $J \cup H$ is stable. By the maximality of J, it follows that $H = \emptyset$. And therefore, no f-J-trail exists in G. By Lemma 6, there is no f-factor better than J. By Theorem 5, J is optimal. □

Corollary 5. *If G is countable, then G possesses a perfect f-factor iff for every nonperfect factor F there exists a factor better than F.*

For countable graphs it is not true that, for every factor F, there exists an f-optimal factor better than F. This is demonstrated by the following counterexample: Let $G = (V, E)$, where $V = \omega \times \{0, 1, 2\}$ and $E :=$

$\{\{(x,0),(x,1)\}: x \in \omega\} \cup \{\{(x,1),(y,2)\}: x, y \in \omega\}$. Further let $f = 1$ and $F := \{\{(x,1),(x,2)\}: x \in \omega\}$. Then there is no optimal factor better than F. But the following holds.

Lemma 17. *Let F be a factor of G. Further let $g: V \to \omega$ be a function such that $g(v) = f(v)$ for almost all $v \in V$ (i.e. $\{v \in V: g(v) \neq f(v)\}$ is finite), and let K be a g-optimal factor such that $d_K(v) = d_F(v)$ for almost all $v \in V$. Then there exists an f-optimal factor H satisfying $\delta_{H,f} \leqslant \delta_{F,f}$.*

Proof. To get a contradiction we assume that there is a sequence $(J_i)_{0 \leqslant i < \omega}$ of factors such that $\delta_{F,f} > \delta_{J_0,f} > \delta_{J_1,f} > \cdots$. If J is a factor satisfying $\delta_{J,f} \leqslant \delta_{F,f}$, it holds

$$\begin{aligned}
&\{v \in V: \delta_{K,g}(v) < \delta_{J,g}(v)\} \\
&\quad = \{v \in V: \delta_{F,f}(v) < \delta_{J,f}(v) \text{ and } f(v) = g(v) \text{ and } d_F(v) = d_K(v)\} \\
&\qquad \cup \{v \in V: \delta_{K,g}(v) < \delta_{J,g}(v) \text{ and } (f(v) \neq g(v) \text{ or } d_F(v) \neq d_K(v))\} \\
&\quad \subseteq \emptyset \cup \{v \in V: f(v) \neq g(v) \text{ or } d_F(v) \neq d_K(v)\}.
\end{aligned}$$

Hence the set $V_1(J)$ of Lemma 4, referring to g and K in place of f and F, is finite. Thus, by Lemma 4, triangle inequality, and $\delta_{F,f}(v) \geqslant \delta_{J,f}(v)$ for all $v \in V$, we conclude that

$$\begin{aligned}
0 &\leqslant \sum_{v \in V} \delta_{F,f}(v) - \delta_{J,f}(v) = \sum_{v \in V} |d_F(v) - f(v)| - |d_J(v) - f(v)| \\
&\leqslant \sum_{v \in V} |d_F(v) - d_K(v)| + |d_K(v) - g(v)| + |g(v) - f(v)| \\
&\qquad - |d_J(v) - g(v)| + |g(v) - f(v)| \\
&= 2 \sum_{v \in V} |g(v) - f(v)| + \sum_{v \in V} |d_F(v) - d_K(v)| - \sum_{v \in V} \delta_{J,g}(v) - \delta_{K,g}(v) \\
&\leqslant 2 \sum_{v \in V} |g(v) - f(v)| + \sum_{v \in V} |d_F(v) - d_K(v)| < \infty
\end{aligned}$$

Therefore $(\sum_{v \in V} \delta_{F,f}(v) - \delta_{J_i,f}(v))_{0 \leqslant i < \omega}$ is a strictly increasing and bounded sequence of natural numbers. This is impossible. □

5. Optimal factors of rayless graphs

We have seen that if G is countable, then G contains an optimal factor, but not necessarily an optimal factor better than a given factor. In 'rayless' graphs the situation is more pleasant. G is said to be *rayless* if no path in G has infinite length. We will show that if G is rayless and F a non-optimal factor of G, then there exists an optimal factor better than F. The corresponding result for 1-factors

has been proved by Schmidt [6] by transfinite induction on the 'order' of G. The order $\sigma(G)$ of a rayless graph will be defined by transfinite recursion as follows.

If G is finite, then G is said to have order 0. Let $\alpha > 0$ be an ordinal. G is said to have order α if there is a finite set $W \subseteq V$ such that each component of $G \setminus W := (V \setminus W, \{e \in E: e \cap W = \emptyset\})$ has order less than α. If G has order α for some α, then let $\sigma(G)$ be the least ordinal α such that G has order α.

Lemma 18 (Schmidt [6]). *G is rayless iff $\sigma(G)$ is defined.*

Proof. Let $\sigma(G)$ be undefined. We construct recursively a sequence $v_0, v_1, v_2, \ldots$ of vertices and a sequence $C_0, C_1, C_2, \ldots$ of subgraphs of G such that for all $n \in \omega$ the following holds: C_n is connected, $\sigma(C_n)$ is undefined, v_n is a vertex of C_n, C_{n+1} is a component of $C_n \setminus \{v_n\}$, $\{v_n, v_{n+1}\} \in E$. Then $(v_i)_{0 \leq i < \omega}$ is a path in G and hence G is not rayless. $n = 0$: Since $\sigma(G)$ is undefined, there exists a component C_0 of $G \setminus \emptyset = G$ such that $\sigma(C_0)$ is undefined. Choose a vertex v_0 of C_0. Now assume that v_n and C_n are already defined. Since $\sigma(C_n)$ is undefined by induction hypothesis, there exists a component C_{n+1} of $C_n \setminus \{v_n\}$ such that $\sigma(C_{n+1})$ is undefined. Since C_n is connected by induction hypothesis, there exists a vertex v_{n+1} of C_{n+1}, such that $\{v_n, v_{n+1}\} \in E$.

The other implication will be proved by transfinite induction on $\sigma(G)$. If $\sigma(G) = 0$, then G is finite and therefore rayless. If $\sigma(G) > 0$, there exists a finite set $W \subseteq V$ such that $\sigma(C) < \sigma(G)$ for each component of $G \setminus W$. Since by induction hypothesis each component of $G \setminus W$ is rayless, G is rayless, too. □

Theorem 7. *If G is rayless and if F is a non-optimal factor of G, then there exists an optimal factor better than F.*

Proof. We will prove the theorem by transfinite induction on $\sigma(G)$. If $\sigma(G) = 0$, then G is finite and the assertion holds trivially. For the induction step let $\sigma(G) > 0$. There exists a finite $W \subseteq V$ such that, for each component C of $G \setminus W$, $\sigma(C) < \sigma(G)$. Let $\{C_i = (V_i, E_i): i \in I\}$ be the set of components of $G \setminus W$. Define

$$F_W := F \cap \{e \in E: e \cap W \neq \emptyset\} \quad \text{and} \quad F_i := F \cap E_i \quad \text{for } i \in I.$$

F is the disjoint union of the F_i's and F_W and, for each $i \in I$, F_i is a factor of C_i. Further let us define

$$g(v) := \begin{cases} f(v) - d_{F_W}(v) & \text{if } v \notin W \text{ and } d_{F_W}(v) < f(v), \\ 0 & \text{otherwise.} \end{cases}$$

If $i \in I$, let g_i be the restriction of g to V_i. By induction hypothesis there exists, for each $i \in I$, a g_i-optimal factor K_i of C_i satisfying $\delta_{K_i, g_i} \preccurlyeq \delta_{F_i, g_i}$. Let

$$K := \bigcup_{i \in I} K_i \quad \text{and} \quad J := K \cup F_W.$$

An easy computation shows $\delta_{J,f} \preccurlyeq \delta_{F,f}$. Since $g(v)=0$ for all $v \in W$, K is a g-optimal factor of G. Since W and F_W are finite, we conclude that $g(v)=f(v)$ and $d_J(v)=d_K(v)$ for almost all $v \in V$. By Lemma 17, there exists an f-optimal factor H such that $\delta_{H,f} \preccurlyeq \delta_{J,f}$. It follows that $\delta_{H,f} \preccurlyeq \delta_{F,f}$. □

6. The Lovász-decomposition

For an arbitrary function $g: V \to \omega$ we define subsets $A(g)$, $B(g)$, $C(g)$, $D(g)$ of V as follows:

$$C(g) := \{v \in V: d_F(v) = g(v) \text{ for every } g\text{-optimal factor } F \text{ of } G\},$$
$$A(g) := \{v \in V \setminus C: d_F(v) \geqslant g(v) \text{ for every } g\text{-optimal factor } F \text{ of } G\},$$
$$B(g) := \{v \in V \setminus C: d_F(v) \leqslant g(v) \text{ for every } g\text{-optimal factor } F \text{ of } G\},$$
$$D(g) := V \setminus (A \cup B \cup C).$$

Let $A := A(f)$, $B := B(f)$, $C := C(f)$, $D := D(f)$. We call (A, B, C, D) the *Lovász-decomposition* of G. Note that $V = C$ iff G possesses a perfect factor or G contains no optimal factor. In Lovász and Plummer [4, pp. 388–404] this decomposition is considered for finite graphs.[1] It is a generalization of the 'Gallai–Edmonds-decomposition' for 1-factors. (See e.g. [4, pp. 93–102].) We will show that the properties of the Lovász-decomposition proved in [4] remain true for arbitrary graphs.

Lemma 19. *Let F be an optimal factor of G and let $x \in D$.*

(a) *If $d_F(x) > f(x)$, then there exists an F-alternating trail which starts and terminates at x with an F-edge and meets vertices of D only.*

(b) *If $d_F(x) < f(x)$, then there exists an F-alternating trail which starts and terminates at x with an F^C-edge and meets vertices of D only.*

Proof. (a) Since $x \in D$, there exists an optimal factor H satisfying $d_H(x) < f(x)$. Let $v_0 := x$ and let $\mathscr{P} = (v_i)_{0 \leqslant i < k}$ be an F-H-trail. Lemma 3 implies $k < \omega$ and

$$\delta_{F \oplus E(\mathscr{P})} = \delta_{Fv_0}^{v_{k-1}}, \qquad \delta_{H \oplus E(\mathscr{P})} = \delta_{Hv_0}^{v_{k-1}}. \tag{1}$$

Hence $v_0 = v_{k-1} = x$ because otherwise, by Lemma 2(c), (d),

$$\delta_{F \oplus E(\mathscr{P})}(v_{k-1}) = \delta_F(v_{k-1}) - 1 \quad \text{or} \quad \delta_{H \oplus E(\mathscr{P})}(v_{k-1}) = \delta_H(v_{k-1}) - 1$$

contrary to (1). Therefore $d_F(v_{k-1}) > d_H^{v_0}(v_{k-1})$ and $\{v_{k-2}, v_{k-1}\} \in F \setminus H$ by the definition of an F-H-trail. Now let $1 \leqslant i_0 \leqslant k-2$. We must show that $v_{i_0} \in D$. Assume that i_0 is odd. (The case that i_0 is even is handled analogously.) Thus $\{v_{i_0-1}, v_{i_0}\} \in F$. Let us define $\mathscr{P}' := (v_i)_{0 \leqslant i \leqslant i_0}$ and $\mathscr{P}'' := (v_{k-i})_{1 \leqslant i \leqslant k-i_0}$. By Lemma 3, $F \oplus E(\mathscr{P}')$ and $F \oplus E(\mathscr{P}'')$ are optimal factors. It follows that $d_F(v_{i_0}) \leqslant f(v_{i_0})$

[1] [4] deals with the more general case of the '(f, g)-factors'.

and $d_F(v_{i_0}) \geqslant f(v_{i_0})$ since otherwise $F \oplus E(\mathscr{P}')$ or $F \oplus E(\mathscr{P}'')$ is better than F, respectively. Hence $d_F(v_{i_0}) = f(v_{i_0})$, and therefore $d_{F \oplus E(\mathscr{P}')}(v_{i_0}) < f(v_{i_0})$ and $d_{F \oplus E(\mathscr{P}'')}(v_{i_0}) > f(v_{i_0})$. This implies $v_{i_0} \in D$.

(b) can be proved analogously. □

Theorem 8. *Let F be an optimal factor of G. Then the following holds:*

(a) *If $x \in A$ and $y \in A \cup C$, then $\{x, y\} \notin F$.*

(b) *If $x \in B$, $y \in B \cup C$, and $\{x, y\} \in E$, then $\{x, y\} \in F$.*

(c) *If $x \in D$, then $d_F(x) \in \{f(x) - 1, f(x), f(x) + 1\}$.*

(d) *If $x \in D$ and $y \in C$, then $\{x, y\} \notin E$.*

Proof. (a) Assume for a contradiction that $x \in A$, $y \in A \cup C$, and $\{x, y\} \in F$. By the definition of A and C, it follows that $d_F(x) \geqslant f(x)$ and $d_F(y) \geqslant f(y)$. We even have $d_F(x) = f(x)$ and $d_F(y) = f(y)$ because otherwise, by Lemma 3, $F \setminus \{\{x, y\}\}$ is an optimal factor which is better than F or which contradicts $x \in A$ or $y \in A \cup C$. Since $x \in A$, there exists an optimal factor H satisfying $d_H(x) > f(x)$. Then $\{x, y\} \notin H$ since otherwise $H \setminus \{\{x, y\}\}$ is an optimal factor which is better than H or which contradicts $y \in A \cup C$. Let $x_0 := y$, $x_1 := x$, and let $\mathscr{P} = (x_i)_{0 \leqslant i < k}$ be an F-H-trail. Define $\mathscr{P}' := (x_i)_{1 \leqslant i < k}$. By Lemma 3, $H \oplus E(\mathscr{P}')$ is optimal and $k < \omega$.

Case 1. *$\{x_{k-2}, x_{k-1}\} \in F \setminus H$ and $d_F(x_{k-1}) > d_H^y(x_{k-1})$.*

Then $x_{k-1} \notin \{x, y\}$. Since $H \oplus E(\mathscr{P}')$ is not better than H, we conclude that $d_H(x_{k-1}) \geqslant f(x_{k-1})$. Therefore $d_F(x_{k-1}) > f(x_{k-1})$. By Lemma 3 (referring to F and $(x_{k-i})_{1 \leqslant i \leqslant k}$), $F \oplus E(\mathscr{P})$ is optimal. Now $d_{F \oplus E(\mathscr{P})}(y) = d_F(y) - 1$ is a contradiction to $y \in A \cup C$.

Case 2: *$\{x_{k-2}, x_{k-1}\} \in H \setminus F$ and $d_F(x_{k-1}) < d_H^y(x_{k-1})$.*

Since $H \oplus E(\mathscr{P}')$ is not better than H, we conclude that $d_H(x_{k-1}) \leqslant f^x(x_{k-1})$. Therefore we get $x_{k-1} \notin \{x, y\}$ since otherwise $d_{H \oplus E(\mathscr{P}')}(x) = d_H(x) - 2 \leqslant f(x) - 1$ or $d_{H \oplus E(\mathscr{P}')}(y) = d_H(y) - 1 \leqslant f(y) - 1$ contrary to $x \in A$ and $y \in A \cup C$. Consequently we find $d_F(x_{k-1}) < f(x_{k-1})$. Now we get a contradiction as in case 1.

(b) Can be proved analogously.

(c) Let $x \in D$. If $d_F(x) > f(x)$, let $\mathscr{P}$ be an F-alternating trail in agreement with Lemma 19(a). Then $d_{F \oplus E(\mathscr{P})} = d_{Fxx}$. Since $F \oplus E(\mathscr{P})$ is not better than F, we conclude that $d_F(x) = f(x) + 1$. If $d_F(x) < f(x)$, analogously it follows that $d_F(x) = f(x) - 1$ from Lemma 19(b).

(d) Suppose that $\{x, y\} \in E$ for some $x \in D$ and some $y \in C$. By (c), there exists an optimal factor F such that $d_F(x) = f(x) - 1$. Then $\{x, y\} \in F$ since otherwise $F \cup \{\{x, y\}\}$ would be an optimal factor such that $d_{F \cup \{\{x, y\}\}}(y) = d_F(y) + 1 = f(y) + 1$, contradicting $y \in C$. Let $\mathscr{P}$ be an F-alternating trail in agreement with Lemma 19(b), and let $H := F \oplus E(\mathscr{P})$. Then $\{x, y\} \in H$ and $d_H(x) = f(x) + 1$. Hence $H \setminus \{\{x, y\}\}$ is an optimal factor, and we have $d_{H \setminus \{\{x, y\}\}}(y) = f(y) - 1$. This contradicts $y \in C$. □

We are going to examine the structure of optimal factors on the set D in detail. Obviously the following holds: If $x \in D$, then an f-optimal factor F satisfying $d_F(x) = f(x) + 1$ ($d_F(x) = f(x) - 1$, resp.) is f^x-optimal (f_x-optimal, resp.). The following lemma shows that the converse assertion holds as well.

Lemma 20. *Let $x \in D$. Then:*

(a) $x \in C(f^x)$ *and* $x \in C(f_x)$

(b) *Every f^x-optimal factor and every f_x-optimal factor is f-optimal.*

Proof. (a) To see that $x \in C(f^x)$, let F be an f^x-optimal factor. We have to show that $d_F(x) = f(x) + 1$.

(i) Assume $d_F(x) \leqslant f(x)$. Choose an f-optimal factor H satisfying $d_H(x) = f(x) + 1$ and an H-F-trail $\mathscr{P} = (v_i)_{0 \leqslant i < k}$ starting at $v_0 = x$. By Lemma 3, we have $k < \omega$ and

$$\delta_{H \oplus E(\mathscr{P}), f} = (\delta_{H,f})_{v_0}^{v_{k-1}} \quad \text{and} \quad \delta_{F \oplus E(\mathscr{P}), f^x} = (\delta_{F,f^x})_{v_0}^{v_{k-1}}. \tag{1}$$

If $v_0 \neq v_{k-1}$, then Lemma 2(c), (d) implies

$$\delta_{H \oplus E(\mathscr{P}), f}(v_{k-1}) = \delta_{H,f}(v_{k-1}) - 1 \quad \text{or} \quad \delta_{F \oplus E(\mathscr{P}), f}(v_{k-1}) = \delta_{F,f}(v_{k-1}) - 1,$$

contrary to (1). If $v_0 = v_{k-1}$, then Lemma 2(b) implies $\delta_{F \oplus E(\mathscr{P}), f^x}(x) = \delta_{F,f^x}(x) - 2$. Again this is a contradiction to (1).

(ii) Assume $d_F(x) > f(x) + 1$. By Lemma 17, there exists an f-optimal factor H satisfying $\delta_{H,f} < \delta_{F,f}$. By Theorem 8(c), we find $f(x) - 1 \leqslant d_H(x) < d_F(x)$. Let $\mathscr{P} = (v_i)_{0 \leqslant i < k}$ be an F-H-trail starting at $v_0 = x$. By Lemma 3 and Lemma 2(a), (c), we get

$$\delta_{F \oplus E(\mathscr{P}), f^x} = (\delta_{F,f^x})_x^{v_{k-1}} \quad \text{and} \quad \delta_{F \oplus E(\mathscr{P}), f}(v_{k-1}) < \delta_{F,f}(v_{k-1}).$$

Therefore $x = v_{k-1}$, $\delta_{F \oplus E(\mathscr{P}), f^x} = \delta_{F,f^x}$, and $d_{F \oplus E(\mathscr{P})}(x) \neq d_F(x)$. Thus $F \oplus E(\mathscr{P})$ is an f^x-optimal factor satisfying $d_{F \oplus E(\mathscr{P})}(x) = f(x)$. This is impossible by (i).

$x \in C(f_x)$ can be proved analogously.

(b) Suppose that there is some f^x-optimal factor F which is not f-optimal. By Lemma 17, there exists an f-optimal factor H such that $\delta_{H,f} < \delta_{F,f}$. Again by Lemma 17, there exists an f^x-optimal factor J such that $\delta_{J,f^x} \leqslant \delta_{H,f^x}$. Since $d_F(x) = d_J(x) = f(x) + 1$ by (a), it follows that $\delta_{J,f^x} \leqslant \delta_{F,f^x}$ and, by the f^x-optimality of F, $\delta_{J,f^x} = \delta_{F,f^x}$. Thus

$$\delta_{F,f}(y) = \delta_{H,f}(y) \text{ for } y \neq x \quad \text{and} \quad d_H(x) = f(x).$$

Now let K be an f-optimal factor satisfying $d_K(x) = f(x) + 1$ and let $\mathscr{P} = (v_i)_{0 \leqslant i < k}$ be a K-H-trail starting at $v_0 = x$. Then $k < \omega$ by Lemma 3. Since $\mathscr{P}$ is a K-H-trail, we have $d_K(v_{k-1}) \neq d_H^{v_0}(v_{k-1})$ and therefore $v_{k-1} \neq v_0$. By Corollary 1(a), it follows from the f-optimality of K that $\delta_{H,f}(v_{k-1}) > \delta_{K,f}(v_{k-1})$. By Lemma 2(d),

we get

$$\delta_{H\oplus E(\mathscr{P}),f^x}(v_{k-1}) = \delta_{H\oplus E(\mathscr{P}),f}(v_{k-1}) < \delta_{H,f}(v_{k-1}) = \delta_{F,f^x}(v_{k-1}).$$

Thus $\delta_{H\oplus E(\mathscr{P}),f^x} < \delta_{F,f^x}$ contrary to the f^x-optimality of F.

One can show analogously that every f_x-optimal factor is f-optimal. □

Corollary 6. *Let $x \in D$. Then*:

(a) $A \subseteq A(f^x) \cup C(f^x)$ *and* $A \subseteq A(f_x) \cup C(f_x)$.

(b) $B \subseteq B(f^x) \cup C(f^x)$ *and* $B \subseteq B(f_x) \cup C(f_x)$.

(c) $C \subseteq C(f^x)$ *and* $C \subseteq C(f_x)$.

(d) $D \subseteq D(f^x) \cup C(f^x)$ *and* $D \subseteq D(f_x) \cup C(f_x)$.

Proof. (a), (b), and (c) are immediate consequences of Lemma 20(b).

(d) Let y be a vertex of D and let F be an arbitrary f^x-optimal factor. If $d_F(y) < f^x(y)$ ($d_F(y) > f^x(y)$, resp.), then we must find an f^x-optimal factor J satisfying $d_J(y) > f^x(y)$ ($d_J(y) < f^x(y)$, resp.).

(i) $d_F(y) < f^x(y)$. By Lemma 20(a) we conclude that $y \neq x$ and therefore, by Lemma 20(b) and Theorem 8(c), $d_F(y) = f(y) - 1$. Lemma 19 yields an F-alternating trail $\mathscr{P}$ such that $d_{F\oplus E(\mathscr{P})} = d_F^{yy}$. Define $J := F \oplus E(\mathscr{P})$.

(ii) The case $d_F(y) > f^x(y)$ can be handled similarly.

$D \subseteq D(f_x) \cup C(f_x)$ can be proved analogously. □

If $W \subseteq V$, let us define $G[W] := (W, \{\{v, w\} \in E: v, w \in W\})$. A subset U of W is said to be a *component* of W if $G[U]$ is a component of $G[W]$. By Corollary 6(d), Lemma 20(a), and Theorem 8(d), we have the following.

Lemma 21. *If D^* is a component of D and if $x \in D^*$,* then *$D^* \subseteq C(f^x)$ and $D^* \subseteq C(f_x)$.*

Lemma 22. *Let D^* be a component of D, let $x, y \in D^*$, and let F be an optimal factor such that $d_F(x) \neq f(x)$. Then there exist optimal factors H_1, H_2 such that $d_{H_1}(y) = f(y) + 1$, $d_{H_2}(y) = f(y) - 1$, and*

$$H_i \oplus F \subseteq \{\{v, w\}: v, w \in D^*\} \quad (i = 1, 2).$$

Proof. Let $D' \subseteq D^*$ be the set of all vertices $z \in D^*$ such that there is some optimal factor H satisfying $d_H(z) \neq f(z)$ and

$$H \oplus F \subseteq \{\{v, w\}: v, w \in D^*\}. \tag{1}$$

Because of Theorem 8(c) and Lemma 19, it remains to prove that $D' = D^*$. Assuming the contrary, there are $v \in D'$ and $w \in D^* \setminus D'$ such that $\{v, w\} \in E$. Let H be an optimal factor satisfying $d_H(v) \neq f(v)$ and (1). By Lemma 19, we can assume without loss of generality that $d_H(v) = f(v) - 1$. Since $w \notin D'$, we have

$d_H(w) = f(w)$. Further $\{v, w\} \in H$ since otherwise $H \cup \{\{v, w\}\}$ would be an optimal factor which shows $w \in D'$. Choose an H-alternating trail $\mathscr{P} = (v_i)_{0 \leq i < k}$ in agreement with Lemma 19(b) (referring to H and v). $\mathscr{P}$ does not meet w since otherwise $H \oplus E(\mathscr{P}')$ is an optimal factor showing $w \in D'$, where $\mathscr{P}' = (v_i)_{0 \leq i \leq t}$ and $v_t = w$. Therefore $H \oplus E(\mathscr{P})$ is an optimal factor satisfying $d_{H \oplus E(\mathscr{P})}(v) = f(v) + 1$, (1), and $\{v, w\} \in H \oplus E(\mathscr{P})$. Thus $(H \oplus E(\mathscr{P})) \setminus \{\{v, w\}\}$ is an optimal factor showing $w \in D'$ contrary to $w \notin D'$. □

Theorem 9. *Let D^* be a component of D and let F be an f-optimal factor. Then exactly one of the following four statements holds:*

(i) $\forall x \in D^*$: $d_F(x) = f(x)$, $\{y, a\} \in F$ *for exactly one* $y \in D^*$ *and exactly one* $a \in A$, $\forall x \in D^* \forall b \in B$: $\{x, b\} \in E \Rightarrow \{x, b\} \in F$.

(ii) $\forall x \in D^*$: $d_F(x) = f(x)$, $\forall x \in D^* \forall a \in A$: $\{x, a\} \notin F$, $\{y, b\} \in E \setminus F$ *for exactly one* $y \in D^*$ *and exactly one* $b \in B$.

(iii) $\exists y \in D^* \forall x \in D^*$: $d_F(x) = f^y(x)$, $\forall x \in D^* \forall a \in A$: $\{x, a\} \notin F$, $\forall x \in D^* \forall b \in B$: $\{x, b\} \in E \Rightarrow \{x, b\} \in F$.

(iv) $\exists y \in D^* \forall x \in D^*$: $d_F(x) = f_y(x)$, $\forall x \in D^* \forall a \in A$: $\{x, a\} \notin F$, $\forall x \in D^* \forall b \in B$: $\{x, b\} \in E \Rightarrow \{x, b\} \in F$.

Proof. Obviously at most one of the four possibilities can occur.

Case 1: $d_F(y) \neq f(y)$ *for some* $y \in D^*$.

Then $d_F(y) = f(y) + 1$ or $d_F(y) = f(y) - 1$ by Theorem 8(c).

Case 1.1: $d_F(y) = f(y) + 1$.

Since F is f^y-optimal, it follows by Lemma 21 that, for all $x \in D^* \setminus \{y\}$, $d_F(x) = f(x)$.

Suppose there is some $x \in D^*$ and some $a \in A$ such that $\{x, a\} \in F$. By Lemma 21, Corollary 6(a), and Theorem 8(a), we get $a \in C(f^y)$. Hence $d_F(a) = f(a)$. Choose an f-optimal factor H satisfying $d_H(x) = f(x) + 1$ in agreement with Lemma 22. Then $\{x, a\} \in H$ and $d_H(a) = d_F(a) = f(a)$. By Lemma 3, $H \setminus \{\{x, a\}\}$ is f-optimal. Since $d_{H \setminus \{\{x, a\}\}}(a) = f(a) - 1 < f(a)$, this is a contradiction to $a \in A$. Thus $\forall x \in D^* \forall a \in A$: $\{x, a\} \notin F$.

One can show similarly that every edge $\{x, b\}$ with $x \in D^*$ and $b \in B$ is contained in F. Therefore (iii) holds.

Case 1.2: $d_F(y) = f(y) - 1$.

Analogously to case 1.1 (iv) can be shown.

Case 2: $d_F(x) = f(x)$ *for all* $x \in D^*$.

Case 2.1: *There is some* $y \in D^*$ *and some* $a \in A$ *such that* $\{y, a\} \in F$.

Since $y \in D$, there exists an f-optimal factor H satisfying $d_H(y) < f(y)$. By case 1.2, it follows that $d_H(x) = f_y(x)$ for all $x \in D^*$ and $\{y, a\} \notin H$. Let $v_0 := y$, $v_1 := a$ and let $\mathscr{P} = (v_i)_{0 \leq i < k}$ be an F-H-trail. Then, for $i > 0$, $v_i \notin D^*$.

Assuming the contrary, choose $j > 0$ minimal such that $v_j \in D^*$. If j is even, then $\{v_{j-1}, v_j\} \in H \setminus F$. By case 1.2, we get $v_{j-1} \notin A$. Therefore, by Theorem 8(d), $v_{j-1} \in B$. Define $\mathscr{P}' := (v_i)_{0 \leq i \leq j}$. $H \oplus E(\mathscr{P}')$ is f-optimal by Lemma 3. Further

$d_{H\oplus E(\mathscr{P}')}(v_j)=f(v_j)-1$ and $\{v_{j-1}, v_j\}\notin H\oplus E(\mathscr{P}')$. This is a contradiction to case 1.2 (ref. to $H\oplus E(\mathscr{P}')$ and v_{j-1}). If j is odd, we get similarly a contradiction to case 1.1.

In particular we get $v_{k-1}\neq y$. Now we want to show that $J:=F\oplus E(\mathscr{P})$ is f-optimal. By Lemma 3, we have $k<\omega$.

Case a: $\{v_{k-2}, v_{k-1}\}\in F\setminus H$ *and* $d_F(v_{k-1})>d_H^y(v_{k-1})=d_H(v_{k-1})$.

Since $H\oplus E(\mathscr{P})$ is not better than H, $d_H(v_{k-1})\geqslant f(v_{k-1})$. Therefore $d_F(v_{k-1})>f(v_{k-1})$ and J is f-optimal by Lemma 3. (Consider $(v_{k-i})_{0<i\leqslant k}$!)

Case b: $\{v_{k-2}, v_{k-1}\}\in H\setminus F$ *and* $d_F(v_{k-1})<d_H^y(v_{k-1})=d_H(v_{k-1})$.

This case will be treated similarly. Since $d_J(y)=f(y)-1\neq f(y)$ and J is f-optimal, it follows by case 1.2: $\forall x\in D^*\forall v\in A$: $\{x, v\}\notin J$ and $\forall x\in D^*\forall v\in B$: $\{x, v\}\in E\Rightarrow\{x, v\}\in J$. Since $v_i\notin D^*$ for $i>0$, an edge $\{x, v\}\neq\{y, a\}$ with $x\in D^*$ is in J iff it is in F. Consequently (i) holds.

Case 2.2: *There is some* $y\in D^*$ *and some* $b\in B$ *such that* $\{y, b\}\in E\setminus F$.

Analogously to case 2.1 one can show (ii).

Case 2.3: $\forall x\in D^*\forall a\in A$: $\{x, a\}\notin F$ *and* $\forall x\in D^*\forall b\in B$: $\{x, b\}\in E\Rightarrow\{x, b\}\in F$. We must show that this case cannot occur. Choose an arbitrary vertex $y\in D^*$ and an f-optimal factor H satisfying $d_H(y)<f(y)$. By case 1.2, we conclude that $d_H(x)=f_y(x)$ for all $x\in D^*$. Furthermore, $\forall x\in D^*\forall a\in A$: $\{x, a\}\notin H$ and $\forall x\in D^*\forall b\in B$: $\{x, b\}\in E\Rightarrow\{x, b\}\in H$. Let $\mathscr{P}=(v_i)_{0\leqslant i<k}$ be an F-H-trail starting at $v_0=y$. By Lemma 3, we have $k<\omega$. Since $\mathscr{P}$ traverses only edges of $F\oplus H$, $\mathscr{P}$ contains no edge joining D^* and $A\cup B$. Since there is no edge joining D^* and C by Theorem 8(d), $\mathscr{P}$ meets vertices of D^* only. In particular $v_{k-1}\in D^*$. Therefore $d_F(v_{k-1})=f(v_{k-1})=d_H^y(v_{k-1})$. This is a contradiction since v_{k-1} is the endvertex of an F-H-trail. □

G is said to be *f-critical* if G is connected and $V=D$. For $x\in C\cup D$ let us define

$$f_B(x):=f(x)-|\{b\in B: \{x, b\}\in E\}|.$$

It follows easily from Theorem 8(b) and Theorem 9 that f_B defines a function on $C\cup D$ into ω. From Theorem 8 and Theorem 9 we get the following corollary.

Corollary 7. *If G contains at least one f-optimal factor, then $G[C]$ possesses a perfect f_B-factor and, for each component D^* of D, $G[D^*]$ is f_B-*critical.

Acknowledgement

I wish to thank Professor K. Steffens who helped with many useful suggestions related to this work. In addition, I appreciate greatly the comments by M. Holz.

References

[1] R. Aharoni, On an obstruction for perfect matchings, Combinatorica 4 (1984) 1–6.
[2] M. Holz, K.-P. Podewski and K. Steffens, Injective Choice Functions, Lecture Notes in Math., Vol. 1238 (Springer, Berlin, 1987).
[3] K. Hrbacek and T. Jech, Introduction to Set Theory, Monographs Textbooks Pure Appl. Math., Vol. 85 (Marcel Dekker, New York, 1984).
[4] L. Lovász and M.D. Plummer, Matching theory, Ann. Discrete Math. 29 (North-Holland, Amsterdam, 1986).
[5] F. Niedermeyer, Optimale Faktoren in Graphen, Dissertation Hannover, 1990.
[6] R. Schmidt, Ein Reduktionsverfahren für Weg-endliche Graphen, Dissertation Hamburg, 1982.
[7] K. Steffens, Matchings in countable graphs, Canad. J. Math. 29 (1976) 165–168.
[8] K. Steffens, The f-factors of countable graphs, Grüne Reihe 233 (Hannover) (1989).
[9] W.T. Tutte, Selected Papers, Vol. 1 (Charles Babbage Research Center, Manitoba, Canada 1979).
[10] W.T. Tutte, The method of alternating paths, Combinatorica 2 (1982) 325–332.

Discrete Mathematics 95 (1991) 255–270
North-Holland

Universal elements and the complexity of certain classes of infinite graphs

Péter Komjáth*

Department of Computer Science, Eötvös University H-1088 Budapest, Hungary

János Pach**

Mathematical Institute of the Hungarian Academy H-1364 Budapest, P.O.B. 127, Hungary
Courant Institute of Mathematical Sciences, N.Y.U., 251 Mercer Street, New York, NY 10012, USA

Received 16 December 1989
Revised 20 March 1990

Abstract

Komjáth, P. and J Pach, Universal elements and the complexity of certain classes of infinite graphs, Discrete Mathematics 95 (1991) 255–270.

A class of graphs has a universal element G_0, if every other element of the class is isomorphic to an induced subgraph of G_0. In Sections 1–4 we give a survey of some recent developments in the theory of universal graphs in the following areas: (1) Graphs universal for isometric embeddings, (2) universal random graphs, (3) universal graphs with forbidden subgraphs, (4) universal graphs with forbidden topological subgraphs. Section 5 is devoted to the problem of deciding how far a class of graphs $\mathcal{G}$ is from having a universal element. We introduce a new measure of the complexity of the class $\mathcal{G}$, denoted by $\mathrm{cp}(\mathcal{G})$. This is defined to be the minimum cardinal κ such that there exist κ elements in $\mathcal{G}$ with the property that any other element of $\mathcal{G}$ can be embedded into at least one of them as an induced subgraph. $\mathcal{G}$ has a universal element if and only if $\mathrm{cp}(\mathcal{G}) = 1$. Among other theorems we prove that (i) the complexity of the class of all countable graphs without $n \geq 2$ independent edges is finite; (ii) for any cardinal κ, $\omega_1 \leq \kappa \leq 2^{\omega}$, it is consistent that the complexity of the class of all locally finite countable graphs is equal to κ. In Section 6 we consider some analogous questions for hypergraphs.

1. Isomorphic embeddings and isometric embeddings

Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. Given any $x, y \in V(G)$, let $d_G(x, y)$ denote the *distance* of x and y in G, i.e., the length of

* Research supported in part by Hungarian SF grant OTKA-1805 and MSRI, Berkeley, NSF-DMS 8505550.

** Research supported in part by Hungarian SF grant OTKA-1814, NSF grant CCR-8901484 and DIMACS-NSF Science and Technology Center, NSF-STC 88-09648.

Elsevier Science Publishers B.V.

the shortest path connecting them. An *isomorphic embedding* of G into another graph G_0 is an injection $f: V(G) \to V(G_0)$ with the property

$$f(x)f(y) \in E(G_0) \iff xy \in E(G)$$

for every $x, y \in V(G)$. If there exists such an embedding f, then we also say that G is (isomorphic to) an *induced subgraph* of G_0.

An *isometric embedding* of G into another graph G_0 is an injection $f: V(G) \to V(G_0)$ with the property that

$$d_{G_0}(f(x), f(y)) = d_G(x, y)$$

for every $x, y \in V(G)$. We say that H is an *isometric subgraph* of G_0 if $H \subseteq G_0$ and

$$d_H(x, y) = d_{G_0}(x, y)$$

for every $x, y \in V(H)$. Evidently, every isometric embedding is an isomorphic embedding, and every isometric subgraph is an induced subgraph.

Rado [25–26] made the following interesting observation, which raised a number of new questions.

Theorem 1.1 (Rado). *There exists a countable graph G_0 such that every countable graph is isomorphic to an induced subgraph of G_0.*

Proof. Let the vertex set of G_0 be the union of countably many disjoint sets $V_0 \cup V_1 \cup V_2 \cup \cdots$. Let $|V_0| = 1$, and assume that the part of G_0 induced by the set $W = \bigcup_{i<n} V_i$ has already been defined. For every subset $A \subseteq W$, take a different vertex $x_A \notin W$, and join it to all elements in A and to none in $W - A$. Let $V_n = \{x_A : A \subseteq W\}$. It is now clear that, for every countable graph G with $V(G) = \omega = \{0, 1, 2, \ldots\}$, there is an isomorphic embedding $f: V(G) \to V(G_0)$ such that $f(i) \in V_i$ for all i. Later we shall see that the simple idea behind this treelike construction can be developed to establish similar results for many classes of graphs. □

The question whether Theorem 1.1 can be generalized to isometric embeddings was raised by Howorka [17], and was answered by Pach [22] in the affirmative.

Theorem 1.2 (Pach). *There exists a countable graph G_0 such that every countable graph can be isometrically embedded into G_0.*

Proof. It is sufficient to construct a countable graph G^* such that every countable *connected* graph G can be isometrically embedded into G^*. Then the union of countably many vertex-disjoint copies of G^* will meet the requirements for G_0.

Let V_n denote the set of all $n+1$ by $n+1$ symmetric integer matrices $D = (d_{ij})_{i,j=0}^{n}$ for which one can find a connected graph G with $V(G) = \omega$ such that $d_G(i, j) = d_{ij}$ for all $0 \leq i, j \leq n$. Let $V(G^*) = \bigcup_{n<\omega} V_n$. Two vertices (matr-

ices) $D = (d_{ij})_{i,j=0}^{n}$ and $D' = (d'_{ij})_{i,j=0}^{m}$ $(m \leqslant n)$ will be joined by an edge in G^* if and only if $d_{ij} = d'_{ij}$ for all $0 \leqslant i, j \leqslant m$, and $d_{nm} = 1$.

It is not hard to see now that if G is any connected graph on $V(G) = \omega$, then

$$f(n) = (d_G(i, j))_{i,j=0}^{n}, \quad n < \omega,$$

is an isometric embedding of G into G^*. □

A countable graph G is called *isometrically constructible* if one can find a nested sequence of finite isometric subgraphs $G_1 \subset G_2 \subset G_3 \subset \cdots$ of G such that any $x \in V(G)$ is contained in some $V(G_i)$. It would be interesting to find a nice (meaningful) characterization of the isometrically constructible graphs. The only nontrivial result in this direction is due to Pach [22].

Theorem 1.3 (Pach). *Every countable planar graph is isometrically constructible.*

2. Universal graphs and zero-one laws

A completely different approach to Theorem 1.1 that leads to many far-reaching generalizations, was taken by Pósa and Fagin [9] (see also [8]). Roughly speaking, the idea is that a countable random graph almost surely contains every countable graph as an induced subgraph. Furthermore, the countable random graph is (almost surely) unique.

In order to be a little more precise, we need the following definition: A graph G is said to have *property* P_{ij} if for any distinct vertices $x_1, x_2, \ldots, x_i$, $y_1, y_2, \ldots, y_j \in V(G)$ one can find a vertex x adjacent to all of the x and none of the y. The proof of the following two statements is straightforward.

Lemma 2.1. *Let $0 < p < 1$ be fixed, and let $G_{n,p}$ denote a random graph on n vertices, whose edges are chosen independently with probability p. Then*

$$\lim_{n \to \infty} \text{Prob}[G_{n,p} \textit{ has property } P_{ij}] = 1$$

holds for any $i, j < \omega$.

Lemma 2.2 *Let G' and G'' be two countable graphs satisfying P_{ij} for all $i, j < \omega$. Then G' and G'' are isomorphic.*

Lemma 2.1 yields that the system $\mathcal{T}$ consisting of all statements P_{ij} $(i, j < \omega)$ is a *consistent* theory (every finite subsystem $\mathcal{T}' \subset \mathcal{T}$ has a finite model). Hence, by the Gödel Completeness Theorem, $\mathcal{T}$ has a countable model (which, of course, cannot be finite). Moreover, Lemma 2.2 implies that this countable model G_0 is uniquely determined. On the other hand, G_0 has the very strong property that the

embedding of any *finite* subgraph of any countable graph G into G_0 can be extended to an embedding of G into G_0. In particular, any countable graph G is isomorphic to an induced subgraph of G_0.

A similar argument leads to the following zero-one law.

Theorem 2.3 (Fagin). *For every first order statement Q, and for every $0<p<1$,*

$$\lim_{n\to\infty} \mathrm{Prob}[G_{n,p} \textit{ satisfies } Q] = 0 \textit{ or } 1.$$

Proof. It is easy to see that $\mathcal{T}$ is a *complete* theory, i.e., any statement S is either provable or disprovable from $\mathcal{T}$. For if not, then letting $\mathcal{T}' = \mathcal{T} \wedge S$ and $\mathcal{T}'' = \mathcal{T} \wedge (\neg S)$ we would obtain two consistent theories with two different countable models G' and G'', contradicting Lemma 2.2.

Assume now that Q is a first order statement which can be proved from $\mathcal{T}$. Then there exists a natural number k such that the proof uses only P_{ij} with $i, j<k$. Thus, $\mathrm{Prob}[G_{n,p}$ does not satisfy $Q] \leqslant \sum_{i,j<k} \mathrm{Prob}[G_{n,p}$ does not have property $P_{ij}]$, and this sum approaches 0, as n tends to infinity. Similarly, if the negation of Q can be proved from $\mathcal{T}$, then $\mathrm{Prob}[G_{n,p}$ satisfies $Q] \to 0$. □

The same plan can be followed when the edge probability p is not a constant but $p=p(n)$ is a function approaching 0. In this case, Shelah and Spencer [29] established the following result, suggesting that it makes certain sense to speak about a 'universal countable random graph with edge probability $p(n)$' (a rather strange looking notion, indeed). Let $p(n) \ll q(n)$ mean that $\lim_{n\to\infty} p(n)/q(n) = 0$.

Theorem 2.4 (Shelah, Spencer). *Assume that $n^{-1-1/k} \ll p(n) \ll n^{-1-1/(k+1)}$ for some integer $k \geqslant 1$, and Q is a first order statement. Then*

$$\lim_{n\to\infty} \mathrm{Prob}[G_{n,p(n)} \textit{ satisfies } Q] = 0 \textit{ or } 1.$$

As before, the proof gives a uniquely determined countable model (graph) G_0, which has the interesting feature that, if we wish to decide whether a given first order statement is true in $G_{n,p(n)}$ with probability tending to 1, then it is sufficient to check whether or not it holds for G_0. (However, one must admit that some of the most interesting graph-theoretic properties, like planarity, connectedness, etc., cannot be expressed by first order statements. In this sense, the 'universality' of G_0 is limited).

Shelah and Spencer can also prove that Theorem 2.4 remains valid for $p(n) = n^{-\alpha}$, where α can be any fixed irrational number. Yet the most surprising phenomenon in this field is that the above zero–one law is false for many reasonably smooth edge probability functions $p(n)$.

3. Universal graphs with forbidden subgraphs

A class of graphs $\mathcal{G}$ is said to have a *universal element* (or *universal graph*) G_0 if any other element of the class is isomorphic to an induced subgraph of G_0.

Using this terminology, Theorem 1.1 states that the class of all countable graphs possesses a universal element (or, in other words, there exists a universal countable graph). The idea of the proof given in Section 1 can readily be generalized to the class of all countable graphs containing no K_r, a complete subgraph with r vertices.

Proposition 3.1. *The class of all countable K_r-free graphs has a universal element for every $2 \leq r < \omega$.*

Moreover, if the Generalized Continuum Hypothesis (GCH) is assumed, then Proposition 3.1 remains valid for the class of all K_r-free graphs of any infinite cardinality.

Given an infinite cardinal γ and a family of so-called *forbidden subgraphs* $\mathcal{H} = \{H_1, H_2, \ldots\}$,let $\mathcal{G}_\gamma(\mathcal{H})$ denote the class of all graphs of size γ containing no subgraph isomorphic to any element of $\mathcal{H}$. (Note that, according to this definition, the elements of $\mathcal{G}_\gamma(\mathcal{H})$ avoid $\mathcal{H}$ in the strong sense: they must not contain any H_i even as a not necessarily induced subgraph). In the case when $\mathcal{H} = \{H\}$ consists of a single forbidden subgraph, we shall write $\mathcal{G}_\gamma(H)$ instead of $\mathcal{G}_\gamma(\{H\})$. The elements of $\mathcal{G}_\gamma(H)$ are often called H-free.

The easy proof of the following statement has been well known (in folklore) for a long time.

Proposition 3.2. *There is no universal element in $\mathcal{G}_\omega(K_\omega)$.*

Proof. Assume, for a contradiction, that there is a universal graph $G_0 \in \mathcal{G}_\omega(K_\omega)$. Let G_0' denote the graph obtained from G_0 by adding a new vertex v adjacent to all other points. Clearly, G_0' is also K_ω-free. Thus, by our assumption, there is an isomorphic embedding $f: V(G_0') \to V(G_0)$. However, in this case $\{f(v), f^2(v), f^3(v), \ldots\}$ induces an infinite complete subgraph in G_0, contradiction. □

Diestel, Halin and Vogler [7] generalized this argument to establish the following stronger statement.

Theorem 3.3 (Diestel, Halin, Vogler). *Let $\mathcal{H}$ be a non-empty class of countable graphs, each containing an infinite path. Then there is no universal element in $\mathcal{G}_\omega(\mathcal{H})$.*

Given any (finite or infinite) cardinals α, β, let $K_{\alpha,\beta}$ denote a complete bipartite graph with α resp. β elements in its classes. The $K_{1,\omega}$-free graphs, i.e., graphs without vertices of infinite degree, are often called *locally finite.*

De Bruijn (see [26]) showed that $\mathcal{G}_\omega(K_{1,\omega})$, the class of all countable locally finite graphs, does not possess a universal element. (In Section 5 we will prove a much stronger result, based on the original idea of de Bruijn.) Hajnal and Pach [18] proved that $\mathcal{G}_\omega(K_{2,2})$ has no universal element. On the other hand, using GCH, Rado was able to establish the following result for every regular cardinal $\gamma > \omega$, which was extended by Shelah [27] to every $\gamma > \omega$.

Theorem 3.4 (Rado, Shelah). *The class $\mathcal{G}_\gamma(K_{1,\gamma})$ has a universal element for every $\gamma > \omega$.*

We proved the following general theorem telling exactly in which cases $\mathcal{G}_\gamma(K_{\alpha,\beta})$ has a universal element and in which cases it does not, provided that α is finite [20]. Of course, it contains all the above mentioned results as special cases.

Theorem 3.5 (Komjáth, Pach). *Assume GCH. Let $1 \leqslant \alpha \leqslant \beta \leqslant \gamma$ be cardinals, α finite, γ infinite. Then $\mathcal{G}_\gamma(K_{\alpha,\beta})$ has a universal element if and only if*

(i) $\gamma > \omega$, *or*

(ii) $\gamma = \omega$, $\alpha = 1$ *and* $\beta \leqslant 3$.

The basic technique used for the construction of most universal graphs is a natural (nevertheless, powerful) extension of the (trivial) argument in Section 1 proving Theorem 1.1. To be a little more precise, we need some definitions.

A graph (or a structure) G is called *homogeneous* if any isomorphism between two finite subgraphs (substructures) of G can be extended to an automorphism of G. (Note that e.g. the countable universal graph constructed in Section 2 is clearly homogeneous. A complete list of countable homogeneous graphs is given in [21].) From this definition, we immediately obtain the following.

Claim 3.6. Let G be a homogeneous graph of size γ, and let $\mathcal{H}$ denote the family of all finite graphs not isomorphic to any induced subgraph of G. Then G is a universal element in $\mathcal{G}_\gamma(\mathcal{H})$.

A much stronger version of this claim is proved in [1; Theorem 3.2]. A class $\mathcal{G}$ of finite graphs (or structures) is called an *amalgamation class* if, for every pair $G_1, G_2 \in \mathcal{G}$ with a common induced subgraph (substructure) F, one can find an element $G^* \in \mathcal{G}$ such that $G^* \supseteq G_1 \cup G_2$.

Theorem 3.7 (Fraïssé). *Given any family $\mathcal{H}$ of finite graphs (structures), let $\mathcal{G}_\omega(\mathcal{H})$ be the class of all countable graphs (structures) that do not contain any element of $\mathcal{H}$ as a subgraph (not necessarily induced structure). Then $\mathcal{G}_\omega(\mathcal{H})$ has a homogeneous universal graph (structure) if and only if its finite elements form an amalgamation class.*

For more details about amalgamation classes, see [10]. If we want to show that a certain class of countable graphs possesses a universal element, then we can either use Theorem 3.7 directly, or we can apply it to some larger class of structures (obtained by adding some extra relations, coloring etc.). In the latter case, we obtain the universal graph by deleting these additional relations from the universal element in the larger class.

This technique was used by Komjáth Mekler and Pach [19] to establish the following results. Let C_k and P_k denote a cycle of length k and a path of length k, respectively.

Theorem 3.8 (Komjáth, Mekler, Pach). *For any positive integer k, the class $\mathcal{G}_\omega(\{C_3, C_5, \ldots, C_{2k+1}\})$ of all countable graphs containing no odd cycles of length at most $2k+1$ possesses a universal element.*

Theorem 3.9 (Komjáth, Mekler, Pach). *For any infinite cardinal γ and positive integer k, there is a universal element in the class*

(i) *$\mathcal{G}_\gamma(P_k)$ of all graphs of size γ containing no path of length k;*

(ii) *$\mathcal{G}_\gamma(\{C_k, C_{k+1}, \ldots\})$ of all graphs of size γ containing no cycles of length at least k.*

There are many further interesting related questions for uncountable cardinals. We do not discuss them here, because they lead to difficult model-theoretic problems. We only mention the following result, which is a special case of the existence theorem on saturated and special models (see e.g. [2]).

Theorem 3.10. *Assume GCH. If $\gamma > \omega$ and H is a finite graph, then $\mathcal{G}_\gamma(\mathcal{H})$ contains a universal element.*

The interested reader can find more results on universal graphs of size greater than or equal to 2^ω in [24, 16, 28].

4. Universal graphs with forbidden topological subgraphs

The following theorem of Pach [23] settled a problem of Ulam.

Theorem 4.1 (Pach). *There is no universal element in the class of all countable planar graphs.*

Given a graph H, let $\operatorname{top} H$ denote the family of graphs 'topologically equivalent' to H, i.e., the family of all subdivisions of H. (A *subdivision* of H is a graph arising from H by replacing its edges with independent paths.) Accordingly, for a system of graphs $\mathcal{H} = \{H_i\colon i \in I\}$, let $\operatorname{top} \mathcal{H} = \bigcup_{i \in I} \operatorname{top} H_i$. Using this notation, Theorem 4.1 asserts that $\mathcal{G}_\omega(\operatorname{top}\{K_5, K_{3,3}\})$ does not have a universal

element. The idea of the proof given in [23] stems from the observation that the class of finite planar graphs is far from having the amalgamation property discussed in the last section.

Using the same idea, one can easily prove that $\mathcal{G}_\omega(\text{top } K_{3,3})$ does not have a universal element either. Diestel [4] extended this result by showing the following.

Theorem 4.2 (Diestel). *Let* $2 \leqslant \alpha \leqslant \beta < \omega$, $\beta \geqslant 4$. *Then* $\mathcal{G}_\omega(\text{top } K_{\alpha,\beta})$ *has no universal element.*

A different approach was taken in [7] to prove the non-existence of universal elements in some classes of graphs with forbidden topological subgraphs. In particular, it enables us to give an alternative proof of Theorem 4.1.

To sketch their method we need some definitions. Let G be a countable graph, λ an ordinal. A family $\{G_\kappa\colon \kappa < \lambda\}$ of induced subgraphs of G is said to form a *simplicial decomposition* if

(i) $G = \bigcup_{\kappa<\lambda} G_k$,

(ii) $(\bigcup_{\kappa<\mu} G_\kappa) \cap G_\mu =: S_\mu$ is a complete graph (simplex) for every $0 < \mu < \lambda$,

(iii) no S_μ contains any G_κ, $\kappa \leqslant \mu$.

A graph is called *prime* if it has no simplicial decomposition into more than one subgraph, which is easily seen to be equivalent to the fact that it has no separating complete subgraph. It is well known (see [6, 11–14]) that every K_ω-free countable graph has a *prime decomposition*, i.e., a simplicial decomposition in which all elements are prime.

An element of $\mathcal{G}_\omega(\text{top } \mathcal{H})$ is called *maximal*, if the addition of any edge would result in a graph no longer in the class. The *subdivision base* $B(\mathcal{H})$ is defined as the class of all graphs that occur as a member of a prime decomposition of some maximal element of $\mathcal{G}_\omega(\text{top } \mathcal{H})$.

It is not hard to prove the following crucial result showing how these concepts relate to universal graphs. (See e.g. [5]).

Theorem 4.3 (Diestel, Halin, Vogler). *If the subdivision base* $B(\mathcal{H})$ *of some family* $\mathcal{H}$ *of finite graphs is uncountable, then* $\mathcal{G}_\omega(\text{top } \mathcal{H})$ *has no universal element.*

From this, one can deduce the following.

Theorem 4.4 (Diestel, Halin, Vogler). *Given* $2 \leqslant r < \omega$, *the class* $\mathcal{G}_\omega(\text{top } K_r)$ *has a universal element if and only if* $r \leqslant 4$.

It might be interesting to note that Theorem 3.3 immediately implies that $\mathcal{G}_\omega(\text{top } K_\omega)$ cannot have a universal graph.

Let $\mathcal{G}$ be a class of graphs. An element G_0 of $\mathcal{G}$ is called *weakly universal*, if every other element G of $\mathcal{G}$ can be embedded into G_0 as a not necessarily induced

subgraph. That is, there exists $f: V(G) \to V(G_0)$ such that

$$xy \in E(G) \;\Rightarrow\; f(x)f(y) \in E(G_0).$$

As a matter of fact, in all the above mentioned cases when we were able to prove the non-existence of a universal element in some class of graphs $\mathcal{G}$, it was also true that $\mathcal{G}$ had no weakly universal graph. However, the two notions do not always coincide, as is indicated by the following result of Diestel [4].

Theorem 4.5 (Diestel). $\mathcal{G}_\omega$ (top $K_{2,3}$) *has a weakly universal element, but it does not contain a universal one.*

Very similar results are true for classes of graphs of type $\mathcal{G}_\omega(\text{hom}\,\mathcal{H})$, where hom H is defined as the family of all graphs that can be contracted to $\mathcal{H}$.

5. The complexity of a class of graphs

Given a class of graphs $\mathcal{G}$, let $\text{cp}(\mathcal{G})$, the *complexity of* $\mathcal{G}$, be defined as the smallest cardinal κ such that there exist κ elements in $\mathcal{G}$ with the property that any other element is isomorphic to an induced subgraph of at least one of them. Obviously, $\text{cp}(\mathcal{G}) = 1$ means that $\mathcal{G}$ has a universal element.

For the sake of simplicity, in this section we will be concerned with the complexity of $\mathcal{G}_\omega(H)$, i.e., the class of all countable graphs with one forbidden subgraph. To simplify the notation, let

$$c(H) = \text{cp}(\mathcal{G}_\omega(H)).$$

If H is *connected,* then the union of disjoint H-free graphs is also H-free. This immediately implies that $c(H) \leqslant \omega$ if and only if $c(H) = 1$, i.e., $\mathcal{G}_\omega(H)$ has a universal graph.

However, if H is *disconnected* then $c(H)$ can be a natural number different from 1. For instance, if H consists of two disjoint edges, then K_3, $K_{1,\omega} \in \mathcal{G}_\omega(H)$, but K_3 has no proper extension belonging to this class. Hence, $c(H) \geqslant 2$. In fact, $c(H) = 2$, because any H-free graph is an induced subgraph of K_3 or $K_{1,\omega}$.

For any two graphs H_1 and H_2, let $H_1 + H_2$ denote the union of their vertex-disjoint copies. We write nH for $H + H + \cdots + H$, where the number of summands is n.

Theorem 5.1. *For* $2 \leqslant n < \omega$:

(i) $1 < c(nK_2) < \omega$,

(ii) $1 < c(K_n + K_2) < \omega$.

Proof. (i) To see that $\mathcal{G}_\omega(nK_2)$ does not possess a universal element, it is enough to observe that $(n-1)K_3$, $K_{1,\omega} \in \mathcal{G}_\omega(nK_2)$, but no proper extension of $(n-1)K_3$ belongs to this class.

The upper bound can be proved by induction on n. As we have seen before, it is true for $n=2$. Assume that we have already proved that there are finitely many elements in $\mathcal{G}_\omega((n-1)K_2)$ such that any other element can be embedded into at least one of them, i.e., $c((n-1)K_2)<\omega$. Since $\mathcal{G}_\omega((n-1)K_2)\subseteq\mathcal{G}_\omega(nK_2)$, it is sufficient to find finitely many elements in $\mathcal{G}_\omega(nK_2)$ which embed every $G\in\mathcal{G}_\omega(nK_2)-\mathcal{G}_\omega((n-1)K_2)$. Fix such a G, and let $X\subseteq V(G)$, $|X|=2n-2$, denote the vertex set of a maximum system of independent edges in G. We divide the remaining vertices of G into 2^{2n-2} disjoint classes (some of which might be empty) so that two vertices belong to the same class if and only if they are adjacent to the same elements of X. If a class contains at least n vertices, then we extend it to an infinite class by adding ω new vertices connected to X in the same way. Thus, we obtain a new graph $G^*\supseteq G$. Clearly, $G^*\in\mathcal{G}_\omega(nK_2)$, because any set of n independent edges in G^* could use only at most n vertices from the same class, therefore they would already occur in G. On the other hand, the number of different graphs that can be obtained as G^* is obviously finite. There are at most

$$2^{\binom{2n-2}{2}}$$

graphs induced by X, and the number of elements in each of the 2^{2n-2} classes of the remaining vertices of G^* is either an integer between 0 and $n-1$, or ω. Summarizing, we found at most

$$2^{\binom{2n-2}{2}}(n+1)^{2^{2n-2}}$$

elements in $\mathcal{G}_\omega(nK_2)$ such that any $G\in\mathcal{G}_\omega(nK_2)-\mathcal{G}_\omega((n-1)K_2)$ is an induced subgraph of at least one of them.

(ii) Clearly, K_{n+1}, $K_{1,\omega}\in\mathcal{G}_\omega(K_n+K_2)$, and K_{n+1} is a maximal element of this class. Hence, $c(K_n+K_2)>1$.

As for the upper bound, by Proposition 3.1 it is sufficient to show that there are finitely many elements in $\mathcal{G}_\omega(K_n+K_2)$ such that any $G\in\mathcal{G}_\omega(K_n+K_2)-\mathcal{G}_\omega(K_n)$ is contained in at least one of them as an induced subgraph. Fix such a G and let $X\subseteq V(G)$, $|X|=n$, denote the vertex set of a maximum complete subgraph of G. We can classify the remaining vertices of G, as before, and extend each class of size at least 2 to an infinite class, to obtain a (K_n+K_2)-free graph $G^*\supseteq G$. The number of different graphs that can be obtained as G^* is again finite, as required. □

As a matter of fact, the above argument proves the following slightly more general statement.

Theorem 5.2. *Let H be a finite graph with $c(H)<\omega$. Then $1<c(H+K_2)<\omega$.*

We conjecture that the same theorem holds for any K_r, $r\geqslant 3$. This would yield that if H is the sum of finitely many disjoint finite complete graphs, then $c(H)<\omega$.

To illustrate the difficulty of this question, we discuss another special case.

Theorem 5.3. $1 < c(2K_3) < \omega$.

Proof. The lower bound follows from the fact that $K_5 \in \mathcal{G}_\omega(2K_3)$ cannot be extended in the class by adding new triangles, hence there is no $2K_3$-free graph containing both K_5 and (say) the graph consisting of three edge-disjoint triangles with a point in common.

By Proposition 3.1, there exists a universal countable K_3-free graph. Let G_0 denote the graph obtained from this by replacing each vertex v be a set of two independent points $\{v_1, v_2\}$, and adding a new vertex which is connected to all v_1. Clearly, G_0 is a universal element in the class of all countable graphs whose triangles can be covered by one point. Thus, it is sufficient to show that there are finitely many elements in $\mathcal{G}_\omega(2K_3)$ such that any $2K_3$-free graph G whose triangles cannot be covered by one vertex is an induced subgraph of at least one of them.

Let us fix such a graph G.

Claim A. *There is a subset $A \subseteq V(G)$, $|A| \leqslant 7$ such that $|A \cap V(T)| \geqslant 2$ for every triangle $T \subseteq G$.*

Let v_1, v_2, v_3 be the vertices of a triangle $T_0 \subseteq G$. For every i, pick a triangle T_i (if it exists) such that $V(T_i) \cap \{v_1, v_2, v_3\} = \{v_i\}$. Let A be the union of the vertex sets of these triangles. Using the fact that G is $2K_3$-free, we get $|A| \leqslant 7$. Let $T \subseteq G$ be any triangle. Assume, in order to obtain a contradiction, that $V(T) \cap A = \{v_i\}$, and choose a triangle $S \subseteq G$ not covered by v_i. If $|V(S) \cap \{v_1, v_2, v_3\}| = 2$ then either S and T_i, or S and T are disjoint triangles. If $V(S) \cap \{v_1, v_2, v_3\} = \{v_j\}$, $j \neq i$, then T_j and T are disjoint. This contradiction proves Claim A.

Claim B. *There is a subset $B \subseteq V(G)$, $B \supseteq A$, $|B| \leqslant 8$ such that $B \cap V(T_1) \cap V(T_2) \neq \emptyset$ for any two triangles $T_1, T_2 \subseteq G$.*

If $B = A$ does not have the required property, then there are two triangles $T_1, T_2 \subset G$ with $V(T_1) \cap V(T_2) = \{x\}$ for some $x \notin A$. If $B = A \cup \{x\}$ does not satisfy the requirements, then one can choose two triangles $S_1, S_2 \subseteq G$ such that $V(S_1) \cap V(S_2) = \{y\}$ for some $y \notin A \cup \{x\}$. Since there are no two disjoint triangles in G, $(V(T_1) \cup V(T_2)) \cap A = (V(S_1) \cup V(S_2)) \cap A$ is a 4-element subset of A, all of whose points are connected to both x and y. But then T_1 and $(T_2 - \{x\}) \cup \{y\}$ are two disjoint triangles, a contradiction proving Claim B.

We define the *type* of G as the subgraph of G induced by B, and the collection $\mathcal{C}$ of those subsets $C \subseteq B$ for which there exists a vertex in G connected to all elements of C and none of $B - C$. Obviously, there are at most

$$2^{\binom{|B|}{2}} 2^{2^{|B|}}$$

different types.

Claim C. *The family of all graphs with a given type has a universal element.*

This can be shown by the usual treelike construction (see e.g. Theorem 3.7). We have to show only that the *amalgamation* property holds. Assume that G_1' and G_2' are finite induced subgraphs of two elements G_1 and G_2 of the class of type $(B, \mathscr{C})$. Let G_i'' denote the subgraph of G_i induced by $B \cup V(G_i')$, $i = 1, 2$. Then it is easy to see that $G_1'' \cup G_2''$ is $2K_3$-free and it can be extended to a $2K_3$-free graph of type $(B, \mathscr{C})$. This completes the proof of Claim C, and hence Theorem 5.3. □

It is obvious that $c(H) \leqslant 2^\omega$ for any forbidden subgraph H, because the total number of countable graphs is 2^ω. It is known that the following holds.

Theorem 5.4 (Hajnal, Komjáth). $c(K_\omega) = \omega_1$.

We will show here that $c(K_{1,\omega})$ can take any value between ω_1 and 2^ω.

Given two functions $f, g : \omega \to \omega$, we write $f \leqslant^* g$ if $f(n) \leqslant g(n)$ for all but finitely many n. A family of functions $\{f_\gamma : \gamma \in \Gamma\}$ is called *dominating* if for every $f : \omega \to \omega$ there exists $\gamma \in \Gamma$ with $f_\gamma \geqslant^* f$. The *domination number d* is defined as the smallest size $|\Gamma|$ of a dominating family. (See e.g. [3].)

As we mentioned in Section 3, $c(K_{1,\omega}) > 1$, i.e., the fact that the class of all countable, locally finite graphs does not have a universal element, was proved by de Bruijn. Now we prove the following stronger result.

Theorem 5.5. *$c(K_{1,\omega})$ is equal to the domination number d.*

Proof. Assume first that, for some $\kappa < d$, there is a system $\{G_\gamma : \gamma < \kappa\}$ of countable, locally finite graphs with $V(G_\gamma) = \omega$ and with the property that any other element of $\mathscr{G}_\omega(K_{1,\omega})$ is an induced subgraph of some G_γ. We can obviously assume that each G_γ is connected. Let $f_\gamma(n) < \omega$ denote the number of vertices of G_γ that can be reached for some vertex $i < n$ by a path of length at most n $(n = 1, 2, \ldots)$, and put $f_\gamma(0) = 1$. Since $\kappa < d$, these functions cannot form a dominating system. That is, one can find a function $g : \omega \to \omega$ such that for every $\gamma < \kappa$, $g(n) > f_\gamma(n)$ for infinitely many n. We can obviously suppose that $g(0) = 1$. Let G be a countable, locally finite graph on the vertex set ω such that the number of vertices that can be reached from 0 by a path of length n is at least $g(n)$, for every n. Assume that $h : V(G) \to V(G_\gamma)$ is an embedding of G into some G_γ, as an induced subgraph. However, if $h(0) = k$, then this implies that $f_\gamma(n) \geqslant g(n)$ for all $n > k$, contradicting the definition of G.

Next we prove that $c(K_{1,\omega}) \leqslant d$. It follows immediately from the definitions that we can also find a family of functions $\{f_\gamma : \gamma < d\}$ such that for every $g : \omega \to \omega$ there exists a γ with the (stronger) property that $g(n) \leqslant f_\gamma(n)$ for *all* $n < \omega$. The vertex set of any countable, locally finite graph G can be decomposed

into disjoint finite parts $V_0 \cup V_1 \cup V_2 \cup \cdots$ such that no two points $x \in V_i$, $y \in V_j$, $j > i+1$, are joined by an edge. Let $g_G(n) = |V_n|$, $n < \omega$. Then there exists $\gamma < d$ such that $g_G(n) \leqslant f_\gamma(n)$ for every n. In this case we shall say that the *type of* G is γ. However, an easy recursive construction shows that there is a countable, locally finite graph G_γ containing every countable, locally finite graph of type γ as an induced subgraph. □

Finally, we note that many earlier results for the non-existence of universal elements in certain classes of graphs (especially those whose proof was based on the idea presented in [23]) can be generalized to stronger statements about the complexity of these classes. For example, the following result partially generalizes Theorem 3.5.

Theorem 5.6. (i) $c(K_{1,r}) = 2^\omega$ *for every* $4 \leqslant r < \omega$,
(ii) $c(K_{r,s}) = 2^\omega$ *for every* $2 \leqslant r \leqslant s < \omega$.

6. The complexity of a class of hypergraphs

The *complexity* of a class of hypergraphs $\mathscr{H}$ can be defined as the smallest cardinal κ such that there exist κ elements $H_\alpha^* \in \mathscr{H}$ $(\alpha < \kappa)$ with the property that any other element of $\mathscr{H}$ can be embedded into some H_α^* as an induced subhypergraph. (An *induced subhypergraph* of H_α^* consists of all hyperedges contained in a given subset of $V(H_\alpha^*)$, the vertex set of H_α^*.)

In order to simplify the exposition, we shall only consider classes of countable, 3-*uniform* hypergraphs, i.e., *triplet systems*, although the results clearly generalize to k-uniform hypergraphs for any $k \geqslant 3$.

Given a triplet system F, let $\mathscr{H}_\omega^3(F)$ denote the class of all countable 3-uniform hypergraphs containing no subhypergraph isomorphic to F. (These hypergraphs are also called F-*free*.) Let $c(F)$ denote the complexity of $\mathscr{H}_\omega^3(F)$.

A triplet system F is said to be *weakly complete*, if for any two elements $x, y \in V(F)$ there is a hyperedge (triplet) $T \in E(F)$ with $x, y \in T$.

Proposition 6.1. *If F is a weakly complete, finite triplet system, then* $c(F) = 1$.

Proof. To prove that $\mathscr{H}_\omega^3(F)$ has a universal element, it is sufficient to show that its finite members form an amalgamation class (see Theorem 3.7). Let H_1 and H_2 be two (not necessarily disjoint) 3-uniform F-free hypergraphs which coincide on $V(H_1) \cap V(H_2)$. If $H_1 \cup H_2$ contains an isomorphic copy of F, then it must intersect both $V(H_1) - V(H_2)$ and $V(H_2) - V(H_1)$. However, no pair of points $x \in V(H_1) - V(H_2)$, $y \in V(H_2) - V(H_1)$ is contained in a triplet of $H_1 \cup H_2$, contradicting the assumption that F is weakly complete. □

Clearly, $c(F) \leqslant 2^\omega$ for any F. Our next assertion shows that this bound can be attained.

Theorem 6.2. *Let F consist of two triplets differing only in one point, i.e., $V(F)=\{a, b, c, d\}$ and $F=\{\{a, b, c\}, \{a, b, d\}\}$. Then $c(F)=2^{\omega}$.*

Proof. Let us define a 3-uniform hypergraph H, as follows:

$$V(H)=\{x_i\colon 1\leqslant i\leqslant 8\}\cup\{y_j\colon 1\leqslant j<\omega\},$$

$$E(H)=\{\{x_i, y_j, y_{j+1}\}\colon j\equiv i \pmod 8\}.$$

Obviously, H is F-free. Furthermore, if $f_1, f_2\colon V(H)\to V(H^*)$ are two embeddings of H into an F-free hypergraph H^* such that

$$f_1(y_1)=f_2(y_1), \qquad f_1(x_i)=f_2(x_i) \quad 1\leqslant i\leqslant 8,$$

then $f_1\equiv f_2$. Indeed, $f_1(y_j)=f_2(y_j)$ follows by induction on j from the fact that H^* is F-free.

For any function $g\colon\omega\to\{0, 1\}$, let H_g be defined as the hypergraph obtained from H by adding the triplets

$$\{y_{32k+1}, y_{32k+9}, y_{32k+17+8g(k)}\}$$

for every $\kappa<\omega$. It is clear that H_g is also F-free.

Assume now, in order to obtain a contradiction, that there is a system of fewer than 2^{ω} F-free hypergraphs H^*_{α} such that every H_g can be embedded into at least one of them. Then one can find an α such that H_g can be embedded into H^*_{α} for 2^{ω} different g's. Let $f_g\colon V(H_g)\to V(H^*_{\alpha})$ denote such an embedding. Since the number of 9-tuples in $V(H)$ is countable, there exist $g_1\neq g_2$ such that

$$f_{g_1}(y_1)=f_{g_2}(y_1), \qquad f_{g_1}(x_i)=f_{g_2}(x_i), \quad 1\leqslant i\leqslant 8.$$

This implies that $f_{g_1}\equiv f_{g_2}=:f$. Pick an integer k so that $g_1(k)\neq g_2(k)$. Obviously, both of $\{f(y_{32k+1}), f(y_{32k+9}), f(y_{32k+17})\}$ and $\{f(y_{32k+1}), f(y_{32k+9}), f(y_{32k+25})\}$ are contained in H^*_{α}, contradicting the assumption that H^*_{α} is F-free. □

Let K^3_n denote the system of all triplets of an n element set. If F_1 and F_2 are two triplet systems, then F_1+F_2 stands for the system obtained by taking the union of disjoint copies of F_1 and F_2. Similarly, $nF=F+F+\cdots+F$, where n is the number of terms.

Theorem 6.3. *For every $1<n<\omega$, $1<c(nK^3_3)<\omega$.*

Proof. The easy proof of the lower bound is left to the reader. To establish the upper bound, first we show that for any nK^3_3-free triplet system H one can find $A, B\subseteq V(H)$, $A\cap B=\emptyset$, $|A|=a<n$, $|B|\leqslant(3n-2)(3n-3)$, such that every triplet $T\in E(H)$ either intersects A or has at least two elements in common with B. Furthermore, there are no pairwise disjoint triples $T_1, T_2, \ldots, T_{n-a}\in E(H)$ with $T_i\cap A=\emptyset$ $(1\leqslant i\leqslant n-a)$.

Let $A \subseteq V(H)$, $|A| < n$ be a maximal subset satisfying the condition that H has no $n - |A|$ triplets disjoint from each other and from A. Let $T_1, T_2, \ldots, T_k \in E(H)$ be a maximal system of pairwise disjoint triplets in $V(H) - A$. If $|A| = n - 1$, then we are done. If $|A| < n - 1$ then, by the maximality of A, for any $t \in T_1 \cup T_2 \cup \cdots \cup T_k$ there exist $n - |A| - 1$ disjoint triplets $T_j^{(t)}$ $(1 \leqslant j \leqslant n - |A| - 1)$ in $V(H) - (A \cup \{t\})$. Set

$$B = \left(\bigcup_{1 \leqslant i \leqslant k} T_i \right) \cup \left(\bigcup_{t \in \bigcup T_i,\, 1 \leqslant j \leqslant n - |A| - 1} T_j^{(t)} \right).$$

Then $|B| \leqslant 3(n - |A| - 1) + 9(n - |A| - 1)^2 \leqslant (3n - 2)(3n - 3)$. On the other hand, if a triplet $T \in E(H)$ is disjoint from A then it has to intersect at least one T_i (by the maximality of k). If $|T \cap B| = 1$, then $T \cap B = \{t\}$ for some $t \in \bigcup T_i$, and T is disjoint from all $T_j^{(t)}$ $(1 \leqslant j \leqslant n - |A| - 1)$, contradicting the definition. Thus, A and B satisfy the requirements.

The points and the point pairs of $V(H) - A - B$ can now be classified (colored) according to the types of the triplets containing them. (The *type* of a triplet $T \in E(H)$ is its intersection with $A \cup B$.) This structure (coloured graph) has to omit a finite number of finite substructures (colored subgraphs). Now it is not difficult to build a universal structure of this kind, embedding all H having the same A, B and the same collection of types determined by the triplets of H. □

Theorem 6.4. $c(\omega K_3^3) = \omega$.

Proof. Let H be a countable, 3-uniform hypergraph containing no ωK_3^3. Let $T_1, T_2, \ldots, T_k \in E(H)$ be a non-extendable collection of triplets in H. Then H is $(3k + 1)K_3^3$-free, so $c(\omega K_3^3) \leqslant \omega$ follows from Theorem 6.3. The same argument shows that finitely many ωK_3^3-free graphs cannot embed nK_3^3 if n is large, hence $c(\omega K_3^3) \geqslant \omega$. □

References

[1] P.J. Cameron, The age of a relational structure, Discrete Math. 95 (this Vol.) (1991) 49–67.

[2] C.C. Chang and H.J. Kiesler, Model Theory (North-Holland, Amsterdam, 1973).

[3] E.K. van Douwen, The integers and topology, in: Kunen and Vaughan, eds., Handbook of Set-theoretic Topology (North-Holland, Amsterdam, 1984) 111–167.

[4] R. Diestel, On universal graphs with forbidden topological subgraphs, European J. Combin. 6 (1985) 175–182.

[5] R. Diestel, On the problem of finding small subdivision and homomorphism bases for classes of countable graphs, Discrete Math. 55 (1985) 21–33.

[6] R. Diestel, Graph Decompositions: A Study in Infinite Graph Theory (Oxford Univ. Press, Oxford, 1990).

[7] R. Diestel, R. Halin and W. Vogler, Some remarks on universal graphs, Combinatorica 5 (1985) 283–293.

[8] P. Erdős and J. Spencer, Probabilistic Methods in Combinatorics (Academic Press-Akadémiai Kiadó, New York, Budapest, 1974).
[9] R. Fagin, Probabilities on finite models, J. Symbolic Logic 41 (1976) 50–58.
[10] R. Fraïssé, Sur l'extension aux relations de quelques propriétés des ordres, Ann. Sci. École Norm. Sup. 71 (1954) 361–388.
[11] R. Halin, Graphentheorie II (Wissenschaftliche Buchgesellschaft, Darmstadt, 1981).
[12] R. Halin, Simplicial decompositions in infinite graphs, in: B. Bollobás, ed., Advances in Graph Theory, Ann. Discrete Math. 3 (North-Holland, Amsterdam, 1978).
[13] R. Halin, Simplicial decompositions: some new aspects and applications, in: B. Bollobás, ed., Graph Theory, Ann. Discrete Math. 13 (North-Holland, Amsterdam, 1982).
[14] R. Halin, Simplicial decompositions and triangulated graphs, in: B. Bollobás, ed., Graph Theory and Combinatorics (Academic Press, London, 1984).
[15] A. Hajnal and P. Komjáth, Embedding graphs into colored graphs, Trans. Amer. Math. Soc. 307 (1988) 395–409.
[16] H. Herre, A. Mekler and K. Smith, Superstable graphs, Fund. Math. 118 (1983) 75–79.
[17] E. Howorka, Graphs universal for isometric embeddings, in: Proc. 8th Southeastern Conf. on Combinatorics, Graph Theory and Computing, Baton Rouge (1977).
[18] A. Hajnal and J. Pach, Monochromatic paths in infinite graphs, in Finite and Infinite Sets, Colloq. Math. Soc. János Bolyai (North-Holland, Amsterdam, 1981) 359–369.
[19] P. Komjáth, A. Mekler and J. Pach, Some universal graphs, Israel J. Math. 64 (1988) 158–168.
[20] P. Komjáth and J. Pach, Universal graphs without large bipartite subgraphs, Mathematika 31 (1984) 282–290.
[21] A. Lachlan and R. Woodrow, Countable ultrahomogeneous undirected graphs, Trans. Amer. Math. Soc. 262 (1980) 51–94.
[22] J. Pach, On metric properties of countable graphs (in Hungarian, with an English summary), Mat. Lapok 26 (1975) 305–310.
[23] J. Pach, A problem of Ulam on planar graphs, European J. Combin. 2 (1981) 351–361.
[24] K. Podewski and M. Ziegler, Stable graphs, Fund. Math. 100 (1978) 101–107.
[25] R. Rado, Universal graphs and universal functions, Acta Arith. 9 (1964) 331–340.
[26] R. Rado, Universal graphs, in: Harary and Beineke, eds., A Seminar in Graph Theory (Holt, Rinehart and Winston, New York, 1967).
[27] S. Shelah, Notes in combinatorial set theory, Israel J. Math. 14 (1973) 262–277.
[28] S. Shelah, On universal graphs without instances of CH, Ann. Pure Appl. Logic 2 (1984) 75–87.
[29] S. Shelah and J. Spencer, Zero–one laws for sparse random graphs, J. Amer. Math. Soc. 1 (1988) 97–115.

Discrete Mathematics 95 (1991) 271–289
North-Holland

Asymmetrising sets in trees

Norbert Polat

I.A.E., Université Jean Moulin (Lyon III), 15, quai Claude Bernard, 69239 Lyon Cedex 2, France

Gert Sabidussi

Département de Mathématiques et de Statistique, Université de Montréal, C.P. 6128, Succ.A, Montréal, Que., Canada H3C 3J7

Received 15 September 1990

Abstract

Polat, N. and G. Sabidussi, Asymmetrising sets in trees, Discrete Mathematics 95 (1991) 271–289.

A set A of vertices of a graph G is *asymmetrising* if the identity is the only automorphism of G which stabilises A; G is *asymmetrisable* if it contains such a set. In this paper we investigate the existence of asymmetrising sets in trees.

We show that it is sufficient to consider rooted trees: given any (finite or infinite) tree T there is a vertex w such that T is asymmetrisable if and only if the rooted tree (T, w) is asymmetrisable. In the case of finite trees there is an n^3-algorithm for deciding the existence of an asymmetrising set and for determining the number of similarity classes of such sets. In the infinite case we obtain a characterisation of the asymmetrisable trees without endpoints. It follows from this characterisation that any regular tree, as well as any endpoint-free tree whose cardinal does not exceed the continuum, is asymmetrisable.

1. Introduction, definitions

A construction frequently encountered in graph theory is the extension of a graph by a new vertex (*one-point extension*). Let X be a graph, A a set of vertices of X, z a vertex not in X. The extension Y is obtained by adding to X all edges joining z to the vertices in A. Consider the automorphism group of Y. Clearly, the subgroup of Aut Y which fixes z is the same as the stabiliser $\mathrm{Aut}(X, A)$ of A in Aut X. Of particular interest are those extensions whose automorphism group coincides with $\mathrm{Aut}(X, A)$, in other words those for which z is a fixed point (*fixed point extensions*).

A general question which can be raised in this context is whether a given subgroup of Aut X can be the stabiliser of some set $A \subset V(X)$. In the situation which arises most frequently, the subgroup is Aut X itself and the existence of A is trivial. We shall deal with the opposite extreme, where the given subgroup is

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

trivial. Call a set $A \subset V(X)$ *asymmetrising* if the identity is the only automorphism of X which stabilises A. Clearly some graphs (e.g. asymmetric ones) have asymmetrising sets while others do not (e.g. the complete bipartite graphs K_{mn} with $m+n \geqslant 4$). One may therefore pose the *asymmetrisation problem*: Given a graph X determine whether it has an asymmetrising set. In other words, this is the same as asking whether the vertices of X admit a bicolouration such that the only colour-preserving automorphism of X is the identity.

In this paper we propose to investigate the asymmetrisation problem for trees. Questions of complexity will be of no concern as we shall deal primarily with infinite trees. One feature which makes the asymmetrisation problem particularly attractive in the case of trees is that—with the exception of one rather special situation—every one-point extension is a fixed point extension. In a tree, asking for an asymmetrising set therefore amounts to asking for an asymmetric one-point extension. The exception may occur when the tree T contains a vertex which pairwise separates the vertices in A. The subtree of T generated by A then consists of paths having an endpoint and nothing else in common, and A is called *starlike*. Apart from this case, the group of the extension, Aut Y, and the stabiliser Aut(T, A) are isomorphic.

In order to obtain an asymmetrising set for a tree T the natural approach is to root T at some vertex w and to try to construct an asymmetrising set for the rooted tree (T, w). It is therefore important to know whether one can conclude the existence of an asymmetrising set of T when such a set is known for (T, w). As is shown by the simple example of $K_{1,3}$ rooted at one of its endpoints, this need not be so when the root is arbitrary. Part of the paper is devoted to showing that in any tree one can always find a vertex w such that T has an asymmetrising set whenever there is one for (T, w) (Remark 2.4(i), Theorem 4.2, Proposition 5.2).

The basic distinction to make is between trees containing a ray (one-way infinite path) and those which are rayless. The latter are not considered in this paper, except the finite trees for which we obtain—not very surprisingly—a polynomial time algorithm solving the asymmetrisation problem (Section 3). Infinite rayless trees are dealt with in [5]. The main part of the paper is Section 4 which is concerned with trees that have no endpoints, i.e., are the union of their double rays (two-way infinite paths). We show that such a tree T either has no asymmetrising set or else the number of essentially different asymmetrising sets (which cannot be transformed into one another by automorphisms) is $2^{|T|}$. From this we obtain a characterisation of endpoint-free trees having an asymmetrising set from which in turn some significant sufficient conditions for the existence of such sets can be deduced. For example, all regular trees and all endpoint-free trees of order $\leqslant 2^{\aleph_0}$ have asymmetrising sets. For regular trees this result has already been obtained by Babai [1]. Finally, Section 5 deals with trees which contain a ray but no double ray. There remains the general case of trees which have double rays and endpoints. This case is not considered in this paper; for a treatment we refer the reader to [5].

Not unexpectedly, certain set theoretic hypotheses are related to the existence of asymmetrising sets. For example, the continuum hypothesis is equivalent to the statement that there is no asymmetrising set in the tree formed by $\aleph_2$ rays having precisely their endpoints in common. We have avoided such questions and are strictly working in ZFC.

Notation, Definitions 1.1. (i) Throughout, T will be a (finite or infinite) tree. We denote its vertex-set by V, occasionally by $V(T)$. The set of neighbours of $x \in V$ will be denoted by $V(x; T)$. If T' is a subgraph of T, we denote by $T \setminus T'$ the subgraph of T which consists of the edges of T not belonging to T', together with their incident vertices.

A *ray* is a 1-way infinite path, a *double ray* a 2-way infinite path. T is *rayless* if it contains no ray. A tree containing a ray but no double ray (i.e., having exactly one end in the sense of Halin [2]) will be called *one-ended.*

Given a tree T, the *core* T_* of T is defined as follows: if T is rayless, then $T_* = \emptyset$; if T is one-ended, then T_* is an arbitrary (but fixed) ray starting at some endpoint of T; if T has a double ray, then T_* is the union of all double rays. Note that any one-ended tree has an endpoint.

(ii) Given a rooted tree (T, w) denote by $\leqslant_w$ the natural partial order on V in which w is the least element. For $A \subset V$ and $x \in V$ put

$$A|_x := \{y \in A : x \leqslant_w y\}.$$

By T_x we denote the subtree of T induced by $V|_x$, called the *restriction of* T to x. With this notation, $A|_x = A \cap V(T_x)$. $U_x := V(x; T)|_x = V(x; T_x)$ is the set of *upper neighbours* of x. Of course, $U_w = V(w; T)$. Every $x \in V \setminus \{w\}$ has a unique *lower neighbour*, i.e., a vertex v such that $x \in U_v$.

The distance of a vertex $x \in V$ to the root w will be called the *height* of x, denoted by $h(x)$.

(iii) By Aut T we denote the *automorphism group* of T. If G is a subgroup of Aut T and $A \subset V$ then the *G-similarity class* (i.e., G-orbit) of A is $\{\sigma A : \sigma \in G\}$. Two subsets A, B of V are *G-similar* if there is a $\sigma \in G$ such that $\sigma A = B$. G-similarity of two vertices $x, y \in V$ is defined analogously. If $G = \text{Aut } T$ the reference to G will be omitted. The term 'similarity class of T' will be used to mean '(Aut T)-similarity class of some subset of V'. The *similarity number* of T, denoted by $s(T)$, is the cardinality of the set of all similarity classes of T.

The *stabiliser* of a set $A \subset V$ is the group $\text{Aut}(T, A) := \{\sigma \in \text{Aut } T : \sigma A = A\}$. A is *stable if* $\text{Aut}(T, A) = \text{Aut } T$. We shall say that A is an *asymmetrising set* of T if $\text{Aut}(T, A)$ consists of the identity alone. By $A(T)$ we denote the collection of all asymmetrising sets of T. T is *asymmetrisable* if it has an asymmetrising set. Note that $\text{Aut}(T, V \setminus A) = \text{Aut}(T, A)$ for any $A \subset V$. Hence the complement of an asymmetrising set is likewise asymmetrising.

The *asymmetrising number* of T, denoted by $a(T)$, is the maximum number of mutually nonsimilar asymmetrising subsets of V; in other words, $a(T)$ is the

number of similarity classes of asymmetrising sets in T. It is immediate from the definition that $a(T) \leqslant s(T)$.

(iv) Given a rooted tree (T, w) its *automorphism group* is

$$\mathrm{Aut}(T, w) := \{\sigma \in \mathrm{Aut}\, T : \sigma w = w\},$$

i.e., the stabiliser of $\{w\}$. We shall usually write G_w for $\mathrm{Aut}(T, w)$. The various concepts introduced in (iii) can be transferred to rooted trees by relativisation to G_w. Thus, two vertices or sets of vertices are *similar* in (T, w) if they are G_w-similar in T. A *similarity class of* (T, w) is the G_w-similarity class of some subset of V; the number of such similarity classes is the *similarity number* of (T, w), denoted by $s_w(T)$. $A \subset V$ is an *asymmetrising set* of (T, w) if $G_w \cap \mathrm{Aut}(T, A) = 1$. By $\mathscr{A}_w(T)$ we denote the set of asymmetrising sets of (T, w). (T, w) is *asymmetrisable* if $\mathscr{A}_w(T) \neq \emptyset$. The *asymmetrising number* of (T, w), $a_w(T)$, is the number of G_w-similarity classes of asymmetrising sets of (T, w). For $x \in V$ we abbreviate $s_x(T_x)$, $a_x(T_x)$ and $\mathscr{A}_x(T_x)$ by $s(x)$, $a(x)$ and $\mathscr{A}(x)$, respectively. In particular $s(w) = s_w(T)$, etc. In a rooted tree no set of vertices can be similar to its complement; hence $a_w(T)$ is never odd.

Let $x \in V \backslash \{w\}$. An important parameter of x is its *multiplicity* in (T, w), defined to be the number of vertices in the G_w-orbit of x which have the same lower neighbour as x. Clearly, G_w-similar vertices have the same multiplicity.

Observe that

$$\sigma[A|_x] = (\sigma A)|_{\sigma x} \quad \text{for any } x \in V,\ A \subset V,\ \sigma \in G_w. \tag{1}$$

Hence two upper neighbours x, y of a vertex u are similar in (T, w) if and only if (T_x, x) and (T_y, y) are isomorphic (as rooted trees). This means that the multiplicity $m(x)$ counts how many times the tree (T_x, x) appears among the upper neighbours of u.

(v) Given a set S and a cardinal number n, $\binom{S}{n}$ will denote the set of all n-element subsets of S. If S is itself a cardinal, $S = m$, say, then $\binom{S}{n}$ has the usual meaning as binomial coefficient, with the convention that $\binom{m}{n} = m^n$ if $n \leqslant m$ and m is infinite. Then $\binom{m}{n} = |\binom{S}{n}|$.

The set of all natural numbers will be denoted by ω. If κ is a cardinal, $\Pi(\kappa)$ denotes the set of cardinals $\leqslant \kappa$, and $\bar{\kappa} := |\Pi(\kappa)|$.

2. Preliminary results

Most of our results are based on the following simple observations.

Remarks 2.1. (i) *Let (T, w) be a rooted tree, T' a subtree of T such that $V(T') \cap V(T \backslash T')$ consists of a single vertex w'. Then given any asymmetrising set A of (T, w), $A \cap V(T')$ is an asymmetrising set of (T', w').*

This is a consequence of the fact that any automorphism of (T', w') can be extended to an automorphism of (T, w) by taking the identity on the complement of $V(T')$.

For any $x \in V$, $V(T_x) \cap V(T \setminus T_x) = \{x\}$. Hence we obtain that *if* $A \in \mathscr{A}_w(T)$, *then* $A|_x \in \mathscr{A}(x)$.

(ii) Given a rooted tree (T, w) let u, u' be two distinct upper neighbours of some $v \in V$. *If A is an asymmetrising set of (T, w), and u and u' are G_w-similar, then $A|_u$ and $A|_{u'}$ are not G_w-similar.*

The reason for this is that if $\sigma[A|_u] = A|_{u'}$ for some $\sigma \in G_w$, then $\sigma u = u'$ and the two sets $A|_u$ and $A|_{u'}$ are mapped into each other by the involution $\alpha_\sigma \in G_v \cap G_w$ defined by

$$\alpha_\sigma x := \begin{cases} \sigma x & \text{if } x \in V(T_u), \\ \sigma^{-1} x & \text{if } x \in V(T_{u'}), \\ x & \text{otherwise.} \end{cases}$$

Hence α_σ stabilises A without being the identity, contrary to the hypothesis that A is an asymmetrising set.

(iii) As an immediate consequence of the two preceding remarks we obtain that *if (T, w) is asymmetrisable, then*

$$m(x) \leqslant a(x) \quad \textit{for any } x \in V \setminus \{w\}. \tag{2}$$

(iv) Let T' be a stable subtree of T. Given $A \in \mathscr{A}(T)$ and $A' \in \mathscr{A}(T')$ then $B = A \setminus V(T') \cup A'$ is again an asymmetrising set of T, for if an automorphism σ of T stabilises B, then it stabilises A', hence $\sigma|_{T'} = \mathrm{id}_{T'}$, and consequently σ also stabilises A, so that $\sigma = \mathrm{id}_T$. In other words, *if T is asymmetrisable, then any asymmetrising set of T' can be extended to an asymmetrising set of T* (conditional extendability). Also it is clear that if two extensions

$$A \setminus V(T') \cup A' \quad \text{and} \quad A \setminus V(T') \cup A''$$

with $A', A'' \in \mathscr{A}(T')$ are (Aut T')-similar, then A' and A'' are (Aut T)-similar. Therefore

$$a(T) > 0 \text{ implies } a(T) \geqslant a(T'). \tag{3}$$

Conditional extendability of asymmetrising sets also occurs in an important situation where the subtree is *not* stable.

Lemma 2.2. *Suppose (T, w) is asymmetrisable, and let $u \in V$. Then any asymmetrising set of (T_u, u) can be extended to an asymmetrising set of (T, w).*

Proof. It suffices to consider the case where $u \in U_w$. Let A and A_u be asymmetrising sets of (T, w) and (T_u, u), respectively. If $A_u = \sigma(A|_x)$ for some $\sigma \in G_w$ and $x \in U_w$, then replace $A|_x$ by $A|_u$, and $A|_u$ by A_u; otherwise, replace $A|_u$ by A_u. The set so constructed is then clearly an asymmetrising set of (T, w), and its restriction to u is A_u. □

The above observations find a quantitative expression in the following recursion theorem which expresses the asymmetrising number $a_w(T)$ in terms of the asymmetrising numbers and multiplicities of the restrictions of (T, w) to the neighbours of w.

Theorem 2.3. *Given a rooted tree (T, w), let $\{M_i: i \in I\}$ be the set of G_w-similarity classes of the neighbours of w. For each $i \in I$ let $m_i = |M_i|$ and $a_i = a(x)$, $x \in M_i$. Then the asymmetrising number of (T, w) is $a_w(T) = 2\prod_{i\in I}\binom{a_i}{m_i}$.*

Proof. Let $i \in I$ and consider any representative v of the class M_i. Then by definition, a_i is the number of similarity classes of asymmetrising sets of the rooted tree (T_v, v). Let $\mathscr{A}_i$ be the set of all G_w-similarity classes of asymmetrising sets $X \in \mathscr{A}(x)$, where $x \in M_i$. Then clearly $|\mathscr{A}_i| = a_i$. Let $\mathscr{B} := \{A \in \mathscr{A}_w(T): w \notin A\}$. It is immediate that for any $A \in \mathscr{B}$ the set Q_i of G_w-similarity classes of the resrictions $A|_x$, $x \in M_i$, is a member of $\binom{\mathscr{A}_i}{m_i}$, and that the map $A \mapsto (Q_i)_{i\in I}$ is a bijection of $\mathscr{B}$ onto $\prod_{i\in I}\binom{\mathscr{A}_i}{m_i}$.

To complete the proof note that $|A_w(T)| = 2\,|\mathscr{B}|$ since $A \cup \{w\} \in \mathscr{A}_w(T)$ for any $A \in \mathscr{B}$. □

Remarks 2.4. (i) Suppose that a tree T has no fixed point but contains a fixed edge $e = [u, v]$. Then e is unique and the two components T', T'' of $T \backslash e$ (containing u and v, respectively) are isomorphic as rooted trees. Hence

$$a_u(T) = a_v(T) = a_u(T')a_v(T'') = a_u(T')^2.$$

Among the trees satisfying this hypothesis are the finite and rayless trees without fixed points [6].

(ii) If w is a fixed point or an endpoint of a fixed edge of T, then for any $x, y \in V$,

$$x \leqslant_w y \text{ implies } a_x(T) \leqslant a_y(T).$$

This follows immediately from the fact that any automorphism of (T, y) is also an automorphism of (T, x), whence $\mathscr{A}_x(T) \subset \mathscr{A}_y(T)$.

Proposition 2.5. *Suppose a fixed-point-free tree T has a fixed edge $e = [u, v]$. Then:*

(i) $a_u(T) = \min_{x\in V} a_x(T)$;

(ii) $a(T) = \binom{\sqrt{a_u(T)}}{2}$ *or* $a_u(T)$, *according as $a_u(T)$ is finite or infinite.*

Proof. (i) is an obvious consequence of Remark 2.4(ii).

(ii) Let T^+ be the tree obtained by subdividing the edge e by a new vertex w.

Then w is a fixed point of T^+, and the asymmetrising sets of T are clearly those of T^+ which do not contain the vertex w. Thus

$$a(T) = \tfrac{1}{2}a(T^+) = \tfrac{1}{2}a_w(T^+). \tag{4}$$

On the other hand, the neighbours u, v of w in T^+ form a single $\mathrm{Aut}(T^+, w)$-similarity class, hence by Theorem 2.3,

$$a(T^+) = a_w(T^*) = 2\binom{a_u(T_u^+)}{2} = 2\binom{a_u(T_u)}{2}.$$

Condition (ii) is then a consequence of Remark 2.4(i). □

3. Finite and rayless trees

Theorem 2.3 provides a simple recursive procedure for calculating the asymmetrising number of finite trees.

Algorithm 3.1. Let (T, w) be a finite rooted tree. Define a function $\tilde{a}(x)$, $x \in V$, by the following rules:

(i) Set $\tilde{a}(x) = 2$ for any endpoint x of T, $x \neq w$.

(ii) Select a vertex x such that $\tilde{a}(y)$ is already defined for every upper neighbour y of x. Let $M_1, \ldots, M_n$ be the G_w-similarity classes of upper neighbours of x, and set

$$\tilde{a}(x) = 2\prod_{i=1}^{n}\binom{a_i}{m_i},$$

where $m_i = |M_i|$, and $a_i = \tilde{a}(y)$, $y \in M_i$, $i = 1, \ldots, n$.

(iii) If $\tilde{a}(w)$ is defined, stop; otherwise go back to (ii).

Then $a_w(T) = \tilde{a}(w)$.

Determining the G_w-similarity classes M_i amounts to considering the upper neighbours $y \in U_x$ and testing the restrictions (T_y, y) for isomorphism. As this can be done in linear time [3–4], the algorithm for calculating the asymmetrising number is clearly polynomial.

Example 3.2. In this example, w is a fixed point of T (see Fig. 1). Therefore $G_w = \mathrm{Aut}\, T$ and we obtain the asymmetrising number of T itself: $a(T) = a_w(T) = 112$. Recall that any finite tree T either has a central vertex or a central edge and that these are fixed by $\mathrm{Aut}\, T$. Thus $a(T)$ may be calculated by rooting T at the central vertex or at a new vertex subdividing the central edge (cf. Remark 2.4(i)).

In the following, repeated use will be made of an operation which we call summation of rooted trees. Given a family (T_i, w_i), $i \in I$, of rooted trees, the *sum*

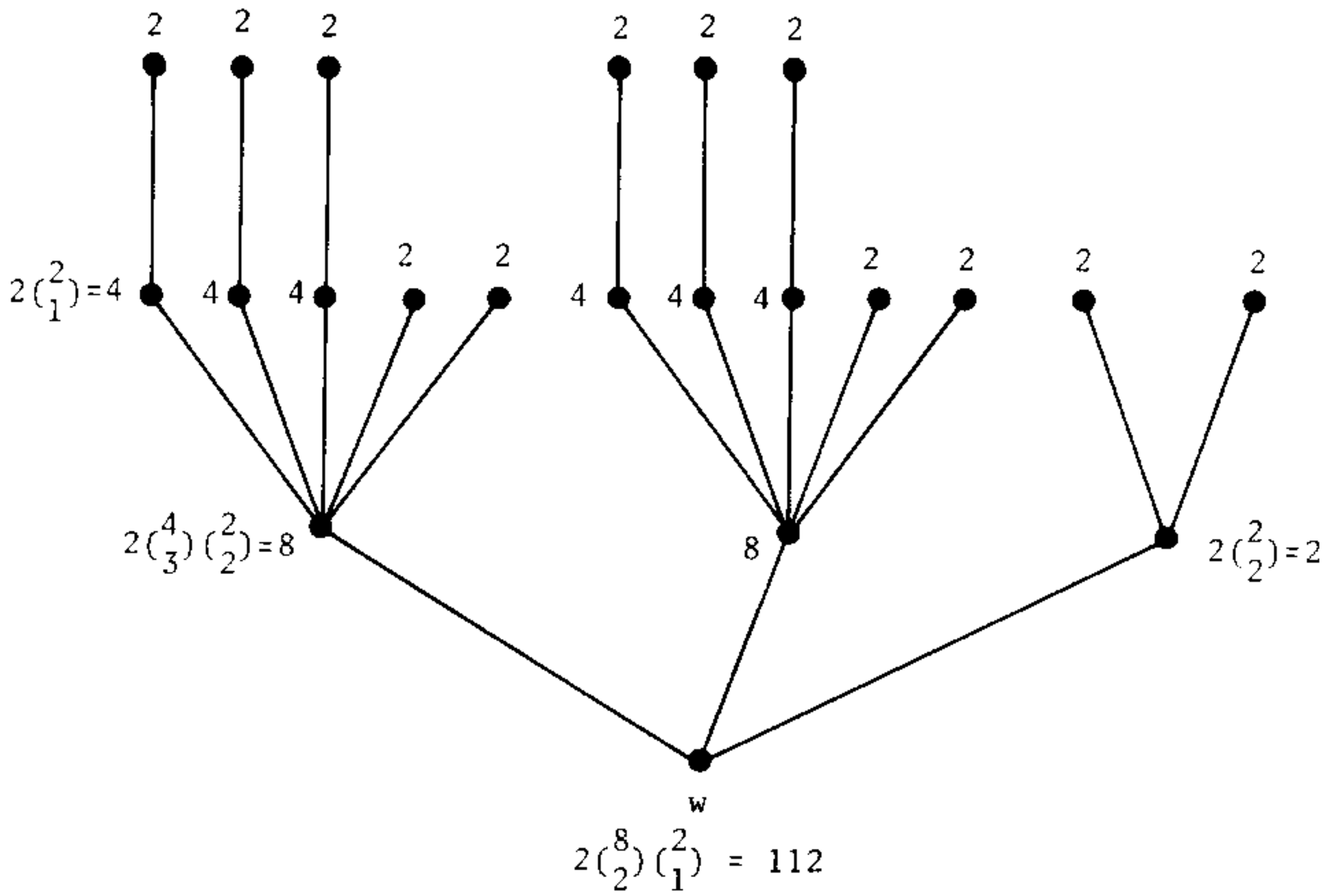


Fig. 1.

is the rooted tree (T, w) obtained by forming the disjoint union of the trees (T_i, w_i), and joining each w_i to a new vertex w. That is to say, the restrictions of the sum to the neighbours of w are precisely the given trees (T_i, w_i). Note that the sum of any family of rayless trees is rayless.

Remark 3.3. The minimal positive value for the asymmetrising number of a rooted tree is 2. Trees for which this value is attained are easily constructed. Let (T, w) be any finite asymmetrisable tree. Consider the sum, (T^*, w^*), of $r := a_w(T)$ copies $(T_1, w_1), \ldots, (T_r, w_r)$ of (T, w). In (T^*, w^*) the vertices $w_1, \ldots, w_r$ form a G_{w^*}-similarity class and $m^*(w_i) = a_w(T) = r$ so that by Theorem 2.3, $a_{w^*}(T^*) = 2$. Observe that the correspondence $(T, w) \mapsto (T^*, w^*)$ is one–one in the sense of isomorphism.

For unrooted trees the minimal positive value for the asymmetrising number is 1. In view of (4) this can be achieved by taking two disjoint copies (T_1, w_1), (T_2, w_2) of a rooted tree (T, w) with $a_w(T) = 2$ and joining the two roots w_1, w_2 by a new edge.

We conclude this section with an observation concerning the asymmetrising number of infinite rayless trees.

Remark 3.4. *There exist arbitrarily large rayless rooted trees whose asymmetrising number is* 2; *also, there exist rayless trees with arbitrarily large asymmetrising number. More precisely we have the following*:

(i) *For any ordinal α there is a set $\mathcal{M}_\alpha$ of $\beth_{\alpha+1}$ non-isomorphic rayless rooted trees (T, w) such that $|T| = \beth_\alpha$ and $a_w(T) = 2$.*

(ii) *Given any cardinal* $\kappa > 0$ *there is a rayless rooted tree* (T, w) *such that* $a_w(T) = 2^\kappa$.

Proof. Induction on α. For the sake of convenience we start at $\alpha = -1$, taking for $\mathcal{M}_{-1}$ the set of all trees (T^*, w^*) introduced in Remark 3.3, with (T, w) ranging over all finite asymmetrisable rooted trees. Suppose $\mathcal{M}_\alpha = \{(T_i, w_i): i \in I\}$ has already been constructed. For $J \subset I$ such that $|J| = \beth_{\alpha+1}$ let (T_J, w_J) be the sum of two copies of every (T_j, w_j) with $j \in J$. Then by Theorem 2.3, $a_{w_J}(T_J) = 2$. We may therefore take

$$\mathcal{M}_{\alpha+1} := \{(T_J, w_J): J \subset I, |J| = \beth_{\alpha+1}\}.$$

If α is a limit ordinal take any sequence $(T_\beta, w_\beta)_{\beta<\alpha}$, where $(T_\beta, w_\beta) \in \mathcal{M}_\beta$; again by Theorem 2.3, the sum of two copies of each (T_β, w_β) is a rayless rooted tree with asymmetrising number 2. Take $\mathcal{M}_\alpha$ to be the set of all such sums.

(ii) Given a cardinal $\kappa > 0$ choose α so large that $\beth_{\alpha+1} \geqslant \kappa$. Take a subset $\mathcal{K}$ of $\mathcal{M}_\alpha$ with $|\mathcal{K}| = \kappa$ and let (Z, z) be the sum of one copy of each of the rooted trees belonging to $\mathcal{K}$. Then by Theorem 2.3, $a_z(Z) = 2^\kappa$. □

4. Upward extendable trees

In this section we consider trees T which are (essentially) equal to their core T_* (Definition 1.1(i)). Call a rooted tree (T, w) *upward extendable* if $U_x \neq \emptyset$ for any $x \in V$. That is to say, the partially ordered set $(V, \leqslant_w)$ has no maximal element. One sees immediately that the following are equivalent:

(i) (T, w) *is upward extendable*;

(ii) *either* T *has no endpoint or* w *is the only endpoint of* T;

(iii) *every vertex of* T *lies on a ray originating at* w;

(iv) T *contains a ray, and* $T \backslash T_*$ *is either empty or a path one of whose endpoints is* w.

Clearly the property of being upward extendable is inherited by all restrictions (T_x, x). Any one-ended upward extendable tree is a ray.

For these trees the situation is particularly satisfactory in as much as the asymmetrising number can be given in closed form rather than by a recursive procedure (Corollary 4.3). We also show that the existence of an asymmetrising set is independent of the choice of the root (Theorem 4.2). A technical result (Theorem 4.5) provides a characterisation of the upward extendable trees which are asymmetrisable, by relating the multiplicities of the restrictions (T_x, x), $x \in V \backslash \{w\}$, to their similarity numbers. We then derive some results about the similarity number of upward extendable trees (Propositions 4.7, 4.9), and use these to obtain the main theorem which gives a necessary and sufficient condition for asymmetrisability in terms of the multiplicities and the orders of the restrictions of (T, w) (Theorem 4.10).

The principal tool for dealing with asymmetrising sets in upward extendable trees is the observation that in such a tree any set may be 'blown up' without changing its stabiliser.

Lemma 4.1. *Let (T, w) be an upward extendable rooted tree, and $f:\omega\to\omega$ a strictly increasing function. For $A\subset V$ define*

$$A_f := \bigcup_{x\in A} J_x, \tag{5}$$

where J_x is the set of all successors of x whose height is $f(h(x))$. Then:

(i) *the map $A\mapsto A_f$ is one–one;*

(ii) *$\sigma A_f = (\sigma A)_f$ for any $\sigma\in G_w$;*

(iii) *A and A_f have the same stabiliser in G_w; in particular, if A is an asymmetrising set of (T, w), then so is A_f.*

Proof. Note that J_x is a non-empty set because $f(h(x))\geqslant h(x)$.

(i) Suppose $A, B\subset V$ are such that $A_f = B_f$ and let $x\in A$. Then $\emptyset\neq J_x\subset B_f$ and hence $J_x\cap J_y\neq\emptyset$ for some $y\in B$, which implies that x and y are comparable with respect to $\leqslant_w$. Thus $A\subset B$ and by symmetry the two sets are equal.

(ii) Since $\sigma V(T_x) = V(T_{\sigma x})$ and $h(\sigma x) = h(x)$ for any $\sigma\in G_w$ and $x\in V$, we have that $\sigma J_x = J_{\sigma x}$ and therefore (ii).

(iii) This is an immediate consequence of (i) and (ii). □

We shall call A_f the *dilatation* of A by f. Note that if $f(n) = n$, then $A_f = A$ for any $A\subset V$.

The operation of dilatation may also be defined in a more general setting. Suppose (T, w) contains a ray, and let $w\in V(T_*)$, with the convention that if T is one-ended, then w is the unique endpoint of T_*. Abbreviate $V(T_*)$ by V_*. Given a strictly increasing $f:\omega\to\omega$, the dilatation of $A\subset V$ by f may be defined as

$$A_f := A\backslash V_* \cup \bigcup_{x\in A\cap V_*} (V_*\cap J_x). \tag{6}$$

Clearly (6) reduces to (5) if (T, w) is upward extendable. It is a straightforward exercise to check that the three statements of Lemma 4.1 still hold in the more general context. The essential point in the proof of (ii) is that the set V_* is stable with respect to G_w.

Theorem 4.2. *Let T be a tree containing a double ray, $w\in V(T_*)$. If (T, w) is asymmetrisable, then so is T, and $a(T)\geqslant 2^{|T_*|}$.*

Proof. Consider an asymmetrising set A of (T, w) such that $w\notin A$. By taking the complement if necessary, we may assume that $|V(T_*)\backslash A| = |T_*|$. We apply dilatation in the sense of (6) using the function $f(n) = 6n$. Since $w\in V(T_*)$ we

have that $V_* = V(T_*)$. Given $X \subset V_* \backslash A$ put

$$B_X = X_f \cup A_f \cup \bigcup_{x \in A_f \cap V_*} V(x; T_*) \cup D_2,$$

where D_2 is the distance 2 neighbourhood of w in T_*, i.e.,

$$D_2 = \{v \in V_* : \operatorname{dist}(v, w) \leqslant 2\}.$$

Note that $\{w\} \cup V(w; T_*) \subset D_2 \subset B_X$, and that the only part of B_X which depends on X is X_f.

By the dilatation all distances in T_* are multiplied by 6, i.e., if $S \subset V_*$, then S_f consists of vertices whose height is a multiple of 6. This implies that w, A_f, and X_f can be recognised within B_X by the following invariant description:

$$\begin{cases} w \text{ is the unique vertex in } B_X \cap V_* \text{ whose distance 2} \\ \text{neighbourhood in } T_* \text{ is contained in } B_X; \\ A_f = B_X \backslash V_* \cup \{x \in B_X \cap V_* : V(x; T_*) \subset B_X\} \backslash (\{w\} \cup V(w; T_*)); \\ X_f = \{x \in B_X \cap V_* : V(x; T_*) \cap B_X = \emptyset\}. \end{cases} \tag{7}$$

(6 is the smallest coefficient of dilatation that will guarantee the above characterisation of w in terms of D_2.)

Now suppose that $X, Y \subset V_* \backslash A$ are such that $\sigma B_X = B_Y$ for some $\sigma \in \operatorname{Aut} T$. Both B_X and B_Y contain w and A_f; since T_* is stable under $\operatorname{Aut} T$ it follows from (7) that $\sigma w = w$ and $\sigma A_f = A_f$. Hence by Lemma 4.1, $\sigma A = A$ and therefore $\sigma = 1$, A being an asymmetrising set of (T, w). Putting $X = Y$ this implies that B_X is an asymmetrising set of T. On the other hand, for $X \neq Y$ we obtain that B_X and B_Y are nonsimilar in T. We have therefore constructed $2^{|V_* \backslash A|} = 2^{|T_*|}$ nonsimilar asymmetrising sets of T. □

Corollary 4.3. *If T is a tree of infinite order κ having fewer than κ endpoints, then $a(T) = 0$ or 2^κ. In particular this holds for upward extendable trees.*

For the proof we need the following lemma.

Lemma 4.4. *If a tree T of infinite order κ has fewer than κ endpoints, then T contains a ray and $|T_*| = \kappa$.*

Proof. Since any rayless tree is generated by its endpoints, i.e., is the union of all paths joining two endpoints, it follows that an infinite rayless tree and its set of endpoints have the same cardinality. Therefore T has a ray.

Suppose $|T_*| < \kappa$. Then $|T \backslash T_*| = \kappa$. $T \backslash T_*$ has at most $|T_*|$ components. Let Z_i, $i \in I$, be the infinite components of $T \backslash T_*$, Z their union, and F the union of the finite components of $T \backslash T_*$. Clearly $|F| \leqslant |T_*|$.

Being rayless, each Z_i has $|Z_i|$ endpoints and all but at most one of them are endpoints of T. Therefore T has at least $\sum_{i \in I} |Z_i| = |Z|$ endpoints. But $|Z| + |F| = |T \backslash T_*| = \kappa$ and $|F| < \kappa$, so that $|Z| = \kappa$, contrary to the hypothesis that T has fewer than κ endpoints. □

Proof of Corollary 4.3. Assume that T is asymmetrisable.

If T is one-ended, then it is countable by Lemma 4.4. Hence the subtree T_0 of T generated by the endpoints of T is finite. Obviously it is stable. $T\backslash T_0$ is a ray and likewise stable. Hence by Remark 2.1(iv), $a(T) \geqslant a(T\backslash T_0) = 2^{\aleph_0} = 2^{\kappa}$.

If T contains a double ray then the corollary follows immediately from Theorem 4.2 and Lemma 4.4. □

Theorem 2.3 says that (T, w) is asymmetrisable if and only if $m(x) \leqslant a(x)$ for any $x \in V\backslash\{w\}$. We now show that for upward extendable trees, $a(x)$ may be replaced by the similarity number $s(x)$.

Theorem 4.5. *An upward extendable tree (T, w) is asymmetrisable if and only if $m(x) \leqslant s(x)$ for any $x \in V\backslash\{w\}$.*

Proof. The necessity follows at once from Remark 2.1(iii) since $a(x) \leqslant s(x)$.

Sufficiency. Let $x \in V$ and denote by M_i, $i \in I$, the similarity classes of the upper neighbours of x. By hypothesis, $|M_i| = m(y) \leqslant s(y)$ for any $y \in M_i$, $i \in I$. Hence we can choose a family of sets A_y, $y \in M_i$, such that $A_y \subset V(T_y)$ and A_y and A_z are nonsimilar in (T_x, x) whenever $y \neq z$.

We now use dilatation. Let N_k, $k \in \omega$, be pairwise disjoint infinite subsets of ω such that $r \geqslant k$ for any $r \in N_k$, and consider a family $(f_x)_{x \in V}$ of strictly increasing functions $f_x : \omega \to \omega$ such that

$$f_x(n) \in N_{h(x)} \quad \text{for any } n \in \omega,\ x \in V. \tag{8}$$

Put

$$Z := \bigcup_{x \in V} \bigcup_{y \in U_x} Z_{xy},$$

where Z_{xy} is the dilatation of A_y by f_x in (T_x, x), i.e., $Z_{xy} = (A_y)_{f_x}$ in the notation of Lemma 4.1. On the other hand, it is clear from (8) and the disjointness of the sets N_k that

$$Z_{xy} = Z|_y \cap H(x), \tag{9}$$

where $H(x)$ is the set of all vertices in T whose height belongs to $N_{h(x)}$.

We claim that Z is an asymmetrising set of (T, w). Suppose that $\sigma \in G_w$ stabilizes Z and that $\sigma x = x$ for some $x \in V$. Let $y \in U_x$. Then by (9)

$$\sigma_x[(A_y)_{f_x}] = \sigma Z_{xy} = Z|_{\sigma y} \cap H(\sigma x) = Z|_{\sigma y} \cap H(x) = Z_{x,\sigma y} = (A_{\sigma y})f_x,$$

where $\sigma_x = \sigma \mid T_x$. Applying Lemma 4.1(i)(ii) to (T_x, x) we obtain that $\sigma_x A_y = A_{\sigma y}$, i.e., that A_y and $A_{\sigma y}$ are similar in (T_x, x). By the choice of the A_y's this means that $\sigma y = y$. Induction on the height of the vertices of T then shows that σ is the identity. □

Remark 4.6. Suppose an upward extendable tree (T, w) has an asymmetrising set. By Remark 2.1(i) the same is true for any restriction (T_x, x). Hence by

Theorem 4.2,

$$s(x) = a(x) = 2^{|T_x|} \quad \text{for any } x \in V. \tag{10}$$

In other words, *if an upward extendable tree contains a vertex x such that $s(x) < 2^{|T_x|}$, then (T, w) is not asymmetrisable.*

In view of this observation it will be of interest to have some information about the similarity number of an upward extendable tree, in particular, under what conditions it is equal to $2^{|T|}$. We begin with a result, similar in spirit to Theorem 2.3, relating the similarity number of (T, w) to those of the restrictions of T to the neighbours of w.

Proposition 4.7. *Given an upward extendable tree (T, w) let $\{M_i : i \in I\}$ be the set of G_w-similarity classes of the neighbours of w. For each $i \in I$ let $m_i = |M_i|$, $s_i = s(u)$, $u \in M_i$, and $\mu_i = \min\{m_i, s_i\}$. Then the similarity number of (T, w) is*

$$s_w(T) = \prod_{i \in I} (\tilde{m}_i s_i)^{\mu_i}.$$

Recall that if κ is a cardinal, then $\tilde{\kappa}$ is the number of cardinals $\leqslant \kappa$.

Proof. For $i \in I$ let T_i be the subtree of T spanned by w and the trees T_x, $x \in M_i$. Each T_i is stable under G_w and clearly $s_w(T) = \prod_{i \in I} s_w(T_i)$. It therefore suffices to show that $s_w(T_i) = (\tilde{m}_i s_i)^{\mu_i}$ for any $i \in I$.

Given $A \subset V(T_i)$, $i \in I$, define an equivalence $\sim_A$ on M_i by $x \sim_A y$ if and only if $A|_x$ is G_w-similar to $A|_y$. Call the cardinal numbers of the $\sim_A$-classes the *multiplicities* associated with A.

Denote the set of $\mathrm{Aut}(T_u, u)$-similarity classes of (T_u, u) by S_i. It is clear that the $\mathrm{Aut}(T_i, w)$-similarity class of a set $A \subset V(T_i)$ may be identified with the function f, defined on S_i, which for each similarity class $\gamma \in S_i$ counts the number of vertices $x \in M_i$ for which $A|_x$ is of similarity type γ. The nonzero values of f are the multiplicities associated with A. Thus the sum of the values of f, S_f, equals m_i. Hence

$$s_w(T_i) = |F|, \quad \text{where } F = \{f \in \Pi(m_i)^{\Sigma_i} : S_f = m_i\}.$$

Note that since T is upward extendable, $s_i = |\Sigma_i|$ is infinite (indeed $\geqslant 2^{\omega}$).

If m_i is finite, then any $f \in F$ has finite support, hence $s_w(T_i) = s_i$. On the other hand, $\mu_i = \min\{m_i, s_i\} = m_i$ and therefore $(\tilde{m}_i s_i)^{\mu_i} = s_i^{m_i} = s_i$. Thus $s_w(T_i) = (\tilde{m}_i s_i)^{\mu_i}$.

Now assume that m_i is infinite. Then

$$\begin{aligned}
s_w(T_i) &= |\{f \in \Pi(m_i)^{\Sigma_i} : S_f = m_i\}| \\
&= |\{f \in \Pi(m_i)^{\Sigma_i} : S_f \leqslant m_i\}| && (s_i \text{ being infinite}) \\
&= |\{f \in \Pi(m_i)^{P} : P \subset \Sigma_i, |P| \leqslant \mu_i\}| && (m_i \text{ being infinite}) \\
&= \sum_{\alpha \in \Pi(\mu_i)} \sum \left\{ \tilde{m}_i^{|P|} : P \in \binom{\Sigma_i}{\alpha} \right\} \\
&= \sum_{\alpha \in \Pi(\mu_i)} \tilde{m}_i^{\alpha} \cdot s_i^{\alpha} = (\tilde{m}_i s_i)^{\mu_i}. \qquad \square
\end{aligned}$$

Corollary 4.8. *If (T, w) is upward extendable and $m(x) \leqslant s(x) = 2^{|T_x|}$ for any $x \in U_w$, then $s_w(T) = 2^{|T|}$.*

Proof. Using the notation of 4.7, $m(x) \leqslant s(x)$ for any $x \in U_w$ implies that $\mu_i = m_i$ for every $i \in I$. Hence $s_w(T_i) = s_i^{m_i} = 2^{|T_x| m_i} = 2^{|T_i|}$, and

$$s_w(T) = \prod_{i \in I} 2^{|T_i|} = 2^{\Sigma |T_i|} = 2^{|T|}. \qquad \square$$

Proposition 4.9. *Let (T, w) be an upward extendable tree such that $m(x) \leqslant 2^{|T_x|}$ for any $x \in V \backslash \{w\}$. Then $s_w(T) = 2^{|T|}$.*

Proof. Suppose that $s_w(T) < 2^{|T|}$. If it were the case that $m(x) \leqslant s(x)$ for any $x \in V \backslash \{w\}$, then by Theorem 4.5, T has an asymmetrising set, and by Theorem 4.2, $s_w(T) = a_w(T) = 2^{|T|}$, a contradiction. Hence there is an $x_0 \in V \backslash \{w\}$ such that $s(x_0) < m(x_0) \leqslant 2^{\kappa_0}$, where $\kappa_0 = |T_{x_0}|$. Now repeat this argument with T_{x_0} etc., to obtain a countable sequence of vertices $w <_w x_0 <_w x_1 <_w \cdots$ such that

$$s(x_j) < m(x_j) \leqslant 2^{\kappa_j}, \quad \text{where } \kappa_j = |T_{x_j}|,\ j = 0, 1, \ldots \tag{11}$$

Abbreviate $m(x_j)$ and $s(x_j)$ by m_j and s_j, respectively.

Clearly $s_0 \geqslant s_1 \geqslant \cdots$. Hence there is a k such that $s_k = s_{k+1}$. By Proposition 4.7,

$$s_k \geqslant (\bar{m}_{k+1} s_{k+1})^{\min\{m_{k+1}, s_{k+1}\}} = (\bar{m}_{k+1} s_k)^{\min\{m_{k+1}, s_k\}}$$

which implies that $m_{k+1} < s_k = s_{k+1}$, a contradiction to (11). $\square$

Theorem 4.10. *An upward extendable tree (T, w) is asymmetrisable if and only if*

$$m(x) \leqslant 2^{|T_x|} \quad \textit{for any} \quad x \in V \backslash \{w\}. \tag{12}$$

Proof. The necessity follows from Theorem 4.5 and (10).

Sufficiency. Assume (12). Let $x \in V \backslash \{w\}$, $y \in V(T_x) \backslash \{x\}$. Then the multiplicity of y when calculated with respect to (T_x, x) is the same as when calculated with respect to (T, w). Hence by applying Proposition 4.9 to (T_x, x) we obtain that $s(x) = 2^{|T_x|}$, i.e., $m(x) \leqslant s(x)$, and the result follows from Theorem 4.5. $\square$

Corollary 4.11. *Let (T, w) be an upward extendable tree such that $d_x \leqslant 2^{|T_y|}$ for any $x \in V$, $y \in U_x$, where d_x is the degree of x in T. Then T is asymmetrisable.*

Proof. Immediate, since $m(y) \leqslant d_x$ for any $x \in V$, $y \in U_x$. $\square$

Corollary 4.12. (i) *If T is an upward extendable tree of order at most $2^{\aleph_0}$, then T is asymmetrisable.*

(ii) *Let T be a tree of infinite order κ such that for any $x \in V$ all components of $T - x$ are of order $\geqslant \log \kappa$ (where $\log \kappa$ is the least cardinal λ such that $2^\lambda \geqslant \kappa$). Then T is asymmetrisable.*

Proof. (i) If $|T| \leqslant 2^{\aleph_0}$, then $d_x \leqslant |T| \leqslant 2^{\aleph_0} \leqslant 2^{|T_y|}$ for any $x, y \in V$.

(ii) Clearly T has no endpoint. Choose any root w and let $x \in V$, $y \in U_x$. Then T_y is a component of $T - x$, hence $d_x \leqslant \kappa \leqslant 2^{\log \kappa} \leqslant 2^{|T_y|}$. By Corollary 4.11, (T, w) has an asymmetrising set, hence by Theorem 4.2 so does T. □

Among the trees which satisfy condition (ii) are those which are generated by their vertices of degree $\geqslant \log \kappa$, in particular, all *regular* trees of infinite degree. Regular trees of finite degree are countable, hence they are asymmetrisable by 4.12(i).

Example 4.13. In view of Remark 4.6 it is meaningful to ask whether there exist trees T without endpoints having fewer than 2^κ nonsimilar subsets, where $\kappa = |T|$. Very simple examples show that the answer is affirmative and that, in fact, $s(T)$ can be as small as $\bar{\kappa}$ (of course, it cannot go below this).

For any infinite cardinal κ let (P_κ, w) be the sum of κ rays rooted at their endpoints, i.e., P_κ is the union of κ rays having exactly w in common. It will be seen that $s(P_\kappa)$ may attain extremal value, $\bar{\kappa}$ or 2^κ, for arbitrarily large κ.

The neighbours of w form a single similarity class whence by Proposition 4.7 (dropping the subscripts i),

$$s(P_\kappa) = s_w(P_\kappa) = (s\bar{m})^\mu = (2^{\aleph_0}\bar{\kappa})^\mu, \quad \text{where } \mu = \min\{\kappa, 2^{\aleph_0}\}. \tag{13}$$

Taking κ so large that $\bar{\kappa} \geqslant 2^{\aleph_0}$, (13) implies that

$$s(P_\kappa) = \bar{\kappa}^{2^{\aleph_0}}, \tag{14}$$

and therefore

$$\bar{\kappa} \leqslant s(P_\kappa) \leqslant 2^{\bar{\kappa}}. \tag{15}$$

Specifically, if one takes $\kappa = \aleph_{\omega_\alpha}$, where $\aleph_\alpha = 2^\nu$ and $\nu \geqslant 2^{\aleph_0}$, then $\bar{\kappa} = 2^\nu$ and (14) reduces to $s(P_\kappa) = \bar{\kappa}$.

Assuming the continuum hypothesis one can be slightly more precise. If $\aleph_1 \leqslant \kappa < \aleph_{\omega_1}$, then $\bar{\kappa} = \aleph_0$ and $\mu = \aleph_1$, hence by (13), $s(P_\kappa) = 2^{\aleph_1} > 2^{\bar{\kappa}}$. We can therefore say that (15) holds if and only if $\kappa \geqslant \aleph_{\omega_1}$.

Examples where the upper bound in (15) is reached are obtained by taking κ to be of cofinality ω and satisfying $\aleph_\kappa = \kappa$. Such cardinals may be arbitrarily large. Clearly they satisfy $\kappa = \bar{\kappa}$ and $\kappa < \kappa^{\aleph_1} \leqslant 2^\kappa$. Hence assuming the generalised continuum hypothesis, $\kappa^{\aleph_1} = 2^\kappa$ so that $s(P_\kappa) = 2^\kappa$.

In this context it is worth noting that because of the form of condition (12), the existence of an asymmetrising set in a given tree may depend on set theoretic assumptions. For example, the tree $P_{\aleph_2}$ is asymmetrisable if and only if the continuum hypothesis does not hold.

Remark 4.14. Is the asymmetrisability of a tree inherited by its core? The obvious approach, taking the part of an asymmetrising set of T that lies in the core, does not work because in general the automorphisms of the core are not

extendable to all of T. Indeed, the answer to the question is negative. This is of some interest inasmuch as the core provides a lower bound for the asymmetrising number of T (Theorem 4.2). Consider the following example.

Given a cardinal $\kappa > 2^{\aleph_0}$ let (P_κ, w) be the tree of Example 4.13. Choose a vertex $x_\alpha \neq w$ on each of the κ rays which make up P_κ. The subscripts α may be considered to be the ordinals $<\kappa$. By Remark 3.4(ii), for any α there is a rayless tree (Q_α, y_α) such that $a(Q_\alpha) = 2^{\aleph_\alpha}$ and y_α is a fixed point of Q_α. Form (T, w) by identifying each x_α with y_α. Clearly $T_* = P_\kappa$, and by Theorem 4.10, P_κ is not asymmetrisable. On the other hand, it is easily seen that taking an asymmetrising set of each Q_α one obtains an asymmetrising set of T.

5. One-ended trees

In the preceding sections we have repeatedly made use of the fact that for the trees under consideration there is a vertex w such that

$$a(T) > 0 \text{ if and only if } a_w(T) > 0. \tag{16}$$

If T has a fixed element (vertex or edge), then any fixed vertex or any vertex incident with a fixed edge may be chosen as the root w. As already pointed out in Remark 2.4(i), fixed elements exist whenever T is rayless. On the other hand, if T contains a double ray, then any vertex of the core will satisfy (16) (Theorem 4.2). We now consider the one remaining case, where T is one-ended and has neither a fixed vertex nor a fixed edge. Note that in a one-ended tree any endpoint of a fixed edge is a fixed point. The assumption that T has no fixed elements thus amounts to saying that there are no fixed points. An example of a tree with these properties is the *inverse binary tree* B which is defined as follows: Choose a ray $R = \langle x_0, x_1, \ldots \rangle$ in the 3-regular tree T_3; then B is the subgraph of T_3 induced by

$$\bigcup_{k \in \omega} \{y : \operatorname{dist}(x_k, y) \leqslant k\}$$

see Fig. 2.

An important property of an (unrooted) one-ended tree T is that its vertex-set can be endowed with a natural partial order by *rooting T at infinity*. T being one-ended we have that for any $x \in V$ there is a unique ray $R_x \subset T$ which starts at x. The order on V is then given by setting

$$x \leqslant y \text{ if and only if } R_x \subset R_y.$$

Lemma 5.1. *If a one-ended tree T contains a ray $R = \langle w_0, w_1, \ldots \rangle$ such that any (T, w_n) has an asymmetrising set, then $a(T) \geqslant 2^{\aleph_0}$.*

Proof. Let $R = \langle w_0, w_1, \ldots \rangle$ be a ray in T such that (T, w_n) has an asymmetrising set A_n, $n \geqslant 0$, and let a subset $I \subset V(R)$ be given. We shall construct an asymmetrising set B_I of T such that $B_I \cap V(R) = I$.

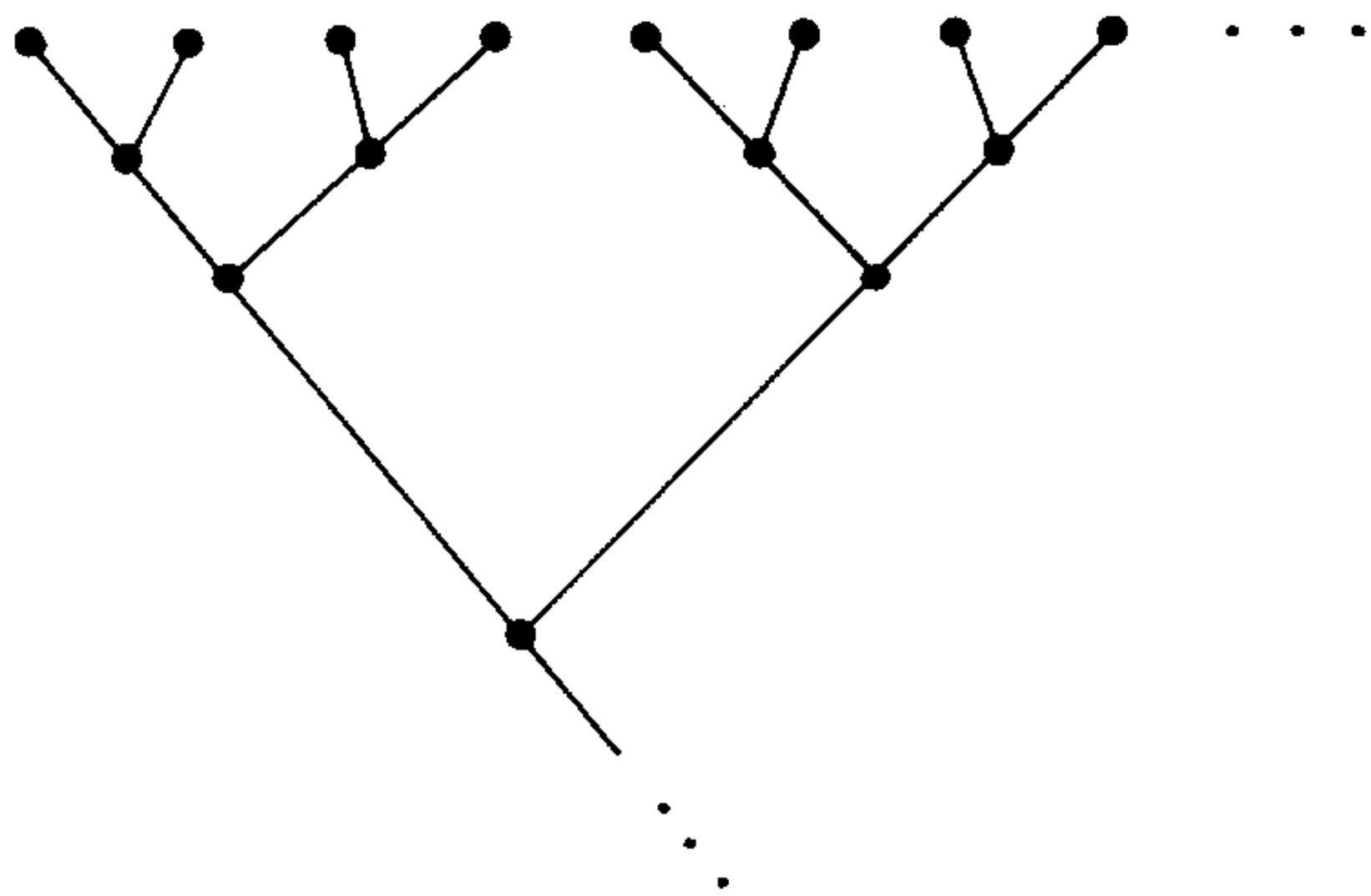


Fig. 2.

For $n \geqslant 0$ put $W_n := \{x \in V : x \geqslant w_n\}$ and let T_n be the subtree of T induced by W_n. We now construct an increasing sequence of sets $B_0, B_1, \ldots$ such that:

(i) $B_n \cap W_{n-1} = B_{n-1}$, $n \geqslant 1$;

(ii) $B_n \cap V(R) = I_n := \{w_i \in I : i \leqslant n\}$, $n \geqslant 0$;

(iii) B_n is an asymmetrising set of (T_n, w_n), $n \geqslant 0$.

Put $B_0 = \{w_0\}$ or $\emptyset$ according as w_0 does or does not belong to I, and suppose B_n has already been constructed. Since (T_n, w_n) is the restriction of (T, w_{n+1}) to w_n we can apply Lemma 2.2 (using the existence of A_{n+1}) to obtain an asymmetrising set A'_{n+1} of (T, w_{n+1}) such that

$$A'_{n+1} \cap W_n = B_n. \tag{17}$$

Put

$$B'_{n+1} := A'_{n+1} \cap W_{n+1}.$$

By Remark 2.1(i), B'_{n+1} is an asymmetrising set of (T_{n+1}, w_{n+1}) and by (17) it satisfies (i). Putting $B_{n+1} = B'_{n+1} \cup \{w_{n+1}\}$ or $B'_{n+1} \backslash \{w_{n+1}\}$ according as w_{n+1} does or does not belong to I we obtain a set satisfying (i)–(iii).

We claim that $B_I := \bigcup_{n \in \omega} B_n$ is an asymmetrising set of T. By one-endedness of T, given any $\sigma \in \operatorname{Aut} T$ there is a natural number N such that $\sigma \in \operatorname{Aut}(T, w_n)$ for all $n \geqslant N$, and therefore $\sigma_n := \sigma \mid W_n \in \operatorname{Aut}(T_n, w_n)$. Thus, if σ stabilises B, then σ_n stabilises B_n and hence is the identity for all $n \geqslant N$, i.e., $\sigma = 1$.

Since no automorphism of T can map a subset of $V(R)$ onto a different subset of $V(R)$, it follows that if I, I' are two distinct subsets of $V(R)$, then B_I and $B_{I'}$ are nonsimilar with respect to $\operatorname{Aut} T$. We have therefore constructed $2^{\aleph_0}$ nonsimilar asymmetrising sets of T. $\square$

Proposition 5.2. *Let T be a one-ended tree without fixed point. Then the following are equivalent*:

(i) *there is a $w \in V$ such that (T, w) is asymmetrisable*;

(ii) *T contains a ray $R = \langle w_0, w_1, \ldots \rangle$ such that any (T, w_n) is asymmetrisable*;

(iii) $a(T) \geqslant 2^{\aleph_0}$.

Proof. (iii) $\Rightarrow$ (i) is trivial; (ii) $\Rightarrow$ (iii) is Lemma 5.1.

(i) $\Rightarrow$ (ii). It suffices to show that if v and w are adjacent vertices such that $v < w$ and (T, w) has an asymmetrising set, then so does (T, v). This will imply that for any vertex x of the ray starting at w, (T, x) has an asymmetrising set.

For $x \in V \backslash \{v\}$ denote by $m_v(x)$ the multiplicity of x in (T, v). By Theorem 2.3, (T, v) is asymmetrisable provided

$$m_v(x) \leqslant a_x(T_{vx}) \quad \text{for any } x \in V(v; T), \tag{18}$$

where T_{vx} is the restriction of (T, v) to x, i.e., the subtree induced by $\{y \in V : x \leqslant_v y\}$.

If $x \in V(v; T) \backslash \{w\}$, then the restrictions of (T, v) and (T, w) to x coincide. Also, v is the lower neighbour of x in both (T, v) and (T, w). Hence (18) reduces to $m(x) \leqslant a(x)$, and this is satisfied because (T, w) is asymmetrisable.

To show that (18) also holds for $x = w$ we use that T is fixed point free. Thus there is a $\sigma \in \operatorname{Aut} T$ such that $\sigma v \neq v$. The key observation is that σT_{vw} is not only a restriction of (T, v) but also of (T, w); indeed, $\sigma T_{vw} = T_{v, \sigma w} = T_{\sigma w}$. Hence

$$m_v(w) = m(\sigma w) \leqslant a_{\sigma w}(T_{\sigma w}) = a_{\sigma w}(\sigma T_{vw}) = a_w(T_{vw}),$$

the inequality arising, as before, from the hypothesis that (T, w) is asymmetrisable. $\square$

Corollary 5.3. *If T is any asymmetrisable tree, then $a(T) \geqslant 2^{|T_*|}$.*

Proof. If T is rayless, then $T_* = \emptyset$, hence $a(T) \geqslant 2^{|T_*|}$ holds trivially. If T is one-ended, then $a(T) \geqslant 2^{\aleph_0} = 2^{|T_*|}$ by Lemma 5.1. If T has a double ray, this is Theorem 4.2. $\square$

Corollary 5.4. *The asymmetrising number of any tree is either finite or uncountable.*

Proof. Suppose that $a(T) = \aleph_0$ for some T. Then by Corollary 5.3, T is rayless. By [4] it has a fixed point or a fixed edge, and by subdividing the latter we may assume w.l.o.g. that T has a fixed point w (subdivision of a fixed edge does not alter an infinite asymmetrising number; see (4)). Thus $a_w(T) = a(T) = \aleph_0$.

We will obtain a contradiction by showing that T contains a ray $\langle w_0, w_1, \ldots \rangle$ such that $a(w_n) = \aleph_0$ for all n. Put $w_0 = w$ and suppose w_n has already been defined. Since T is asymmetrisable we have that $m(x) \leqslant a(x) \leqslant \aleph_0$ for any $x \in V$. Let $W := \{x \in U_{w_n} : m(x) < a(x) < \aleph_0\}$; that is to say, $x \in W$ if and only if

$$2 \leqslant \binom{a(x)}{m(x)} < \aleph_0.$$

If $a(x) < \aleph_0$ for all $x \in U_{w_n}$, then by Theorem 2.3 (applied to (T_{w_n}, w_n)), $a(w_n)$ is either finite or uncountable, depending on whether W consists of finitely or

infinitely many $\mathrm{Aut}(T_{w_n}, w_n)$-similarity classes, a contradiction. Hence w_n has an upper neighbour w_{n+1} with $a(w_{n+1}) = \aleph_0$. $\square$

In fact, it can be shown that if $a(T)$ is uncountable, then it is of the form 2^κ for some κ [5, Corollary 4.7].

Acknowledgements

The work for this paper was carried out during visits of each author at the institution of the other. These visits were supported by the Centre Jacques Cartier and the Université Claude Bernard (Lyon I), respectively. Support from the Natural Sciences and Engineering Research Council of Canada, Grant OGP 7315, is also acknowledged.

References

[1] L. Babai, Asymmetric trees with two prescribed degrees, Acta Math. Hungar. 29 (1977) 193–200
[2] R. Halin, Über unendliche Wege in Graphen, Math. Ann. 157 (1964) 125–137.
[3] J.E. Hopcroft and R.E. Trajan, A v^2 algorithm for determining isomorphism of planar graphs, Inform. Process. Lett. 1 (1971) 32–34.
[4] J.E. Hopcroft and J.K. Wong, Linear time algorithm for isomorphism of planar graphs (extended abstract), in Proc. 6th Ann. ACM Symp. on Theory of Computing, Seattle (May 1974).
[5] N. Polat, Asymmetrization of infinite trees, Discrete Math. 95 (this Vol.) (1991) 291–301.
[6] N. Polat and G. Sabidussi, Fixed elements of infinite trees (DMS 89-1, Université de Montréal, 1989, preprint), Discrete Appl. Math., to appear.

Discrete Mathematics 95 (1991) 291–301
North-Holland

Asymmetrization of infinite trees

Norbert Polat

I.A.E., Université Jean Moulin (Lyon III), 15, quai Claude Bernard, 62939 Lyon Cedex 2, France

Received 7 October 1989
Revised 1 October 1990

Abstract

Polat, N., Asymmetrization of infinite trees, Discrete Mathematics 95 (1991) 291–301.

An *asymmetrizing set* of a tree T is a set A of vertices of T such that the identity is the only automorphism of T which stabilizes A. The *asymmetrizing number* of T is the cardinality of the set of orbits of asymmetrizing sets of T. We complete the results of Polat and Sabidussi (1991) by characterizing the asymmetrizable trees containing a double ray, and by proving that the asymmetrizing number of such a tree T (resp. of any tree T containing a ray but no double ray) is the product of the asymmetrizing numbers of the components of $T \setminus T_*$, where T_* is the union of all double rays (resp. is some ray) of T. We show that, if the asymmetrizing number of a tree is infinite, then it is of the form 2^κ for some cardinal κ. Besides, given two cardinals κ and λ with $\kappa \leqslant \lambda$ we present, on the one hand, a set of 2^λ non-isomorphic rayless trees of cardinality λ and asymmetrizing number 2^κ, and on the other hand, a set of 2^κ non-isomorphic asymmetric (i.e. having no non-identity automorphism) rayless graphs of cardinality κ.

Introduction

Given a tree T, a set A of vertices of T is said to be *asymmetrizing* if the identity is the only automorphism of T which stabilizes A. The *asymmetrizing number* $a(T)$ of T is the cardinality of the set of orbits of asymmetrizing sets of T. In [3], also appearing in this volume, Polat and Sabidussi investigated the *asymmetrizing problem for trees,* i.e. to determine whether a tree has an asymmetrizing set.

The case of rayless trees (trees without infinite path) and that of trees which have no endpoints, i.e. which are the union of their double rays (two-way infinite paths), were completely settled in this paper. In particular, it was shown that an endpoint-free tree T is asymmetrizable, i.e. has an asymmetrizing set, if and only if its asymmetrizing number is $2^{|T|}$, and a very simple characterization of those asymmetrizable trees was given.

In this paper we solve the general case of trees containing a ray by reducing this case to that of rayless trees. We prove that the asymmetrizing number of a tree T containing a ray but no double ray (resp. containing a double ray and which is

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

asymmetrizable) is the product of the asymmetrizing numbers of the components of the forest obtained by deleting the edges of some ray (resp. of the double rays) of T. Besides we get a characterization of the asymmetrizable trees containing a double ray, quite similar to that of endpoint-free trees. These results (Theorems 3.1 and 3.5) complete in some way the study of the asymmetrization problem. Moreover, as several theorems in graph theory, Theorem 3.5 turns out to be 'self-refining', i.e. it yields a result (Proposition 3.8) stronger than itself. Finally, with this theorem and some new results about rayless trees (Section 3), we prove that if the asymmetrizing number of a tree is infinite, then it is of the form 2^κ for some cardinal κ.

1. Notation and definitions

The terminology and notation will be for the most part that used in [3] with the following differences.

1.1. Given a tree T, the *core* T_* of T is defined as follows: if T is rayless, then $T_* = \emptyset$; if T is one-ended, then T_* is an arbitrary (but fixed) ray of T starting at a vertex which must be chosen among the set of vertices fixed by any autmorphism of T if this set is non-empty (this definition is slightly different from that in [3]); if T has a double ray, then T_* is the union of all double rays. If $T_* \neq \emptyset$, then, for $x \in V(T_*)$, we denote by T^x the component of $T \setminus T_*$ containing x.

1.2. If G is a subgroup of Aut T and $A \subseteq V$ then the *G-similarity class* (i.e., G-orbit) of A is $G[A] := \{\sigma A : \sigma \in G\}$. We shall say that A is a *G-asymmetrizing set* of T if $G \cap \mathrm{Aut}(T, A)$ consists of the identity alone. By $\mathscr{A}_G(T)$ we denote the set of G-asymmetrizing sets of T. T is *G-asymmetrizable* if it has a G-asymmetrizing set. Note that $\mathrm{Aut}(T, V \setminus A) = \mathrm{Aut}(T, A)$ for any $A \subseteq V$. Hence the complement of a G-asymmetrizing set is likewise G-asymmetrizing.

The *G-asymmetrizing number* of T, denoted by $a_G(T)$, is the maximum number of mutually non-G-similar G-asymmetrizing subsets of V. It is immediate from the definitions that $a_G(T) \leq 2^{|T|}$.

If $G = \mathrm{Aut}\, T$ the reference to G will be omitted. Thus we will write similar, asymmetrizing, $\mathscr{A}(T)$ and $a(T)$ for G-similar, G-asymmetrizing, $\mathscr{A}_G(T)$ and $a_G(T)$, respectively.

1.3. Given a rooted tree (T, w) its *automorphism group* is $\mathrm{Aut}(T, w) := \{\sigma \in \mathrm{Aut}\, T : \sigma w = w\}$, i.e., the stabilizer of $\{w\}$. If G is a subgroup of Aut T we shall usually write G_w for $G \cap \mathrm{Aut}(T, w)$. The case $G = \mathrm{Aut}\, T$ (hence $G_w = \mathrm{Aut}(T, w)$) is particularly important. With this assumption we set the following definitions and notations: two vertices or sets of vertices are *similar in* (T, w) if they are G_w-similar in T; $A \subseteq V$ is an *asymmetrizing set* of (T, w) if it is a

G_w-asymmetrizing set of T; (T, w) is *asymmetrizable* if T is G_w-asymmetrizable; we denote $\mathscr{A}_{G_w}(T)$ and $a_{G_w}(T)$ by $\mathscr{A}_w(T)$ and $a_w(T)$, respectively. Notice that, *if w is a fixed point of T, then* $\mathrm{Aut}(T, w) = \mathrm{Aut}\, T$, *hence* $a(T) = a_w(T)$.

1.4. Let (T, w) be a rooted tree, G a subgroup of $\mathrm{Aut}(T, w)$, and $x \in V \setminus \{w\}$. The *G-multiplicity* of x in (T, w), denoted by $m_G(x)$, is the number of vertices in the G_w-orbit of x which have the same lower neighbor as x. Clearly, G-similar vertices have the same G-multiplicity. If $G = \mathrm{Aut}(T, w)$, we will write *multiplicity* and $m(x)$ for G-multiplicity and $m_G(x)$, respectively.

2. Rayless trees

Proposition 2.1. *Let (T, w) be a rayless rooted tree such that $a_w(T)$ is infinite. Then $a_w(T) = 2^\kappa$ for some cardinal κ.*

Proof. Suppose $a_w(T) \neq 2^\kappa$ for any cardinal κ. We will define by induction a sequence $(w_n)_{n \geqslant 0}$ of vertices of T such that $w_{n+1} \in V(w_n; T_{w_n})$ and $a_{w_n}(T_{w_n})$ is infinite and $\neq 2^\kappa$ for any cardinal κ.

Let $w_0 := w$. Suppose that w_n is defined for some $n \geqslant 0$. Since $a_{w_n}(T_{w_n})$ is infinite and $\neq 2^\kappa$ for any κ, there is, by [3, Theorem 2.3], a neighbor x of w_n in T_{w_n} such that $a_x(T_x)$ is infinite and $\neq 2^\kappa$ for any κ. Set $w_{n+1} := \kappa$.

$\langle w_0, w_1, \ldots \rangle$ is then a ray of T, a contradiction since T is rayless. Thus $a_w(T) = 2^\kappa$ for some cardinal κ. □

Corollary 2.2. *Let T be a rayless tree such that $a(T)$ is infinite. Then $a(T) = 2^\kappa$ for some cardinal κ.*

Proof. T has a fixed vertex or a fixed edge, since it is rayless (see [4]). Then let w be a fixed vertex of T if there is some, or an endpoint of its fixed edge otherwise. The result is then a consequence of 2.1 since $a(T) = a_w(T)$ by 1.3 and [3, Proposition 2.5]. □

Lemma 2.3. *For any infinite cardinal λ there is a set N_λ of 2^λ non-isomorphic rayless rooted trees (T, w) such that $|T| = \lambda$ and $a_w(T) = 2$.*

Proof. In this proof, as well as in that of Proposition 2.4, we will use the following operation. Given a family $(T_i, w_i)_{i \in I}$ of rooted trees, we define the *sum* of this family as the rooted tree (T, w) obtained by forming the disjoint union of the trees (T_i, w_i), and joining each w_i to a new vertex w. Note that the sum of any family of rayless trees is rayless.

We first recall Remark 3.4(i) of [3]: *For any ordinal α there is a set M_α of $\beth_{\alpha+1}$ non-isomorphic rayless rooted trees (T, w) such that $|T| = \beth_\alpha$ and $a_w(T) = 2$.*

This lemma is then a simple extension of that remark with $N_\lambda = M_\alpha$ when $\lambda = \beth_\alpha$. Suppose $\beth_\alpha < \lambda < \beth_{\alpha+1}$, and let $M_\alpha = \{(T_i, w_i): i \in I\}$. For $J \in \binom{I}{\lambda}$ let (T_J, w_J) be the sum of two copies of every (T_j, w_j) with $j \in J$. Then, by Theorem 2.2, $a_{w_J}(T) = 2$. We may therefore take $N_\alpha := \{(T_J, w_J): J \in \binom{I}{\lambda}\}$, since clearly $|N_\lambda| = 2^\lambda$. □

Proposition 2.4. *For any infinite cardinals κ and λ, with $\kappa \leqslant \lambda$, there is a set of 2^λ non-isomorphic rayless trees T such that $|T| = \lambda$ and $a(T) = 2^\kappa$.*

Proof. (a) We will first prove that, for any infinite cardinal κ there is a rayless tree T_κ such that $|T_\kappa| = \kappa$ and $a(T_\kappa) = 2^\kappa$.

Define (T_ω, w_ω) as the sum of the family $(P_n, a_n)_{n<\omega}$, where (P_n, a_n) is the path of length n rooted at one of its endpoints. Let $\kappa \geqslant \omega$. Suppose that (T_λ, w_λ) is defined for any cardinal $\lambda < \kappa$. Then:

- if κ is a limit cardinal, (T_κ, w_κ) is the sum of the family $(T_\lambda, w_\lambda)_{\omega \leqslant \lambda < \kappa}$;
- if $\kappa = \lambda^+$, then (T_κ, w_κ) is the sum of κ copies of (T_λ, w_λ).

In each case, T_κ is clearly rayless and of cardinality κ. Furthermore, $a(T_\kappa) = a_{w_\kappa}(T_\kappa)$ since w_κ is a fixed point of T_κ. We have then to prove that $a_{w_\kappa}(T_\kappa) = 2^\kappa$. The proof goes by induction on κ, using in each case Theorem 2.3 of [3].

$$a_{w_\omega}(T_\omega) = 2 \prod_{n<\omega} \binom{a_{w_n}(P_n)}{1} = 2 \prod_{n<\omega} 2^{n+1} = 2^\omega.$$

If κ is a limit ordinal $>\omega$, then

$$a_{w_\kappa}(T_\kappa) = 2 \prod_{\omega \leqslant \lambda < \kappa} a_{w_\lambda}(T_\lambda) = 2 \prod_{\omega \leqslant \lambda < \kappa} 2^\lambda = 2^\kappa.$$

If $\kappa = \lambda^+$, then

$$a_{w_\kappa}(T_\kappa) = 2a_{w_\lambda}(T_\lambda)^\kappa = 2(2^\lambda)^\kappa = 2^\kappa.$$

(b) Now let $\lambda \geqslant \kappa$, and let $N_\lambda = \{(T_i, w_i): i \in I\}$ be the set defined in the preceding lemma. Let $i \in I$, and denote by w_κ some vertex of T_κ. Now let $(T_{(i)}, w_{(i)})$ be the sum of (T_κ, w_κ) with two copies of (T_i, w_i). Then $|T_{(i)}| = \lambda$, and by [3, Theorem 2.3] and the fact that the vertex $w_{(i)}$ is a fixed point of $T_{(i)}$, $a(T_{(i)}) = a_{w_{(i)}}(T_{(i)}) = 2^\kappa$. Since $|I| = 2^\lambda$, the set $\{T_{(i)}: i \in I\}$ is then a solution to the statement. □

From the existence of asymmetrizing rayless trees with given cardinality and asymmetrizing number, we can easily obtain a result about the existence of asymmetric (i.e. having only the identity as automorphism, or equivalently such that the empty set is asymmetrizing) rayless graphs of given cardinality.

Proposition 2.5. *For any infinite cardinal κ there is a set of 2^κ non-isomorphic asymmetric rayless graphs of cardinality κ.*

Proof. By Proposition 2.4 there is an asymmetrizing rayless tree T of cardinality κ and asymmetrizing number 2^κ. Let $(A_\alpha)_{\alpha<2^\kappa}$ be a family of pairwise nonsimilar asymmetrizing sets of T. Denote by C a cycle $\langle w_0, w_1, w_2, w_3, w_4\rangle$ of length 5 with in addition the edge $[w_1, w_3]$. The identity is then the only automorphism of C fixing w_0. Now, for all $\alpha<2^\kappa$, let $(C_x^\alpha)_{x\in A_\alpha}$ be a family of pairwise disjoint graphs such that $V(C_x^\alpha\cap T)=\{x\}$ and (C_x^α, x) is isomorphic with (C, w_0); and finally let $G_\alpha := T\cup\bigcup_{x\in A_\alpha} C_x^\alpha$.

The graphs G_α are rayless, and pairwise non-isomorphic since so are the asymmetrizing sets A_α. Besides the restriction of any automorphism σ of G_α to the set $V(T)$ belongs to the stabilizer of A_α, and then it is the identity since A_α is asymmetrizing; hence $\sigma=1$ since $\sigma|\, V(C_x^\alpha)=1$ for any vertex x of T. □

3. Trees containing a ray

Theorem 2.3 of [3] completely settles the case of finite trees (see [3, Algorithm 3.1]), and in a least obvious way that of rayless infinite trees. We will now reduce the case of trees containing a ray to that of rayless trees. We will first improve substantially Proposition 5.2 of [3], which was the only result concerning one-ended trees without fixed point. Following what was done in [3], the vertex-set of a one-ended tree T will be endowed with the partial order defined by

$$x\leqslant y \text{ if and only if } R_x\subseteq R_y,$$

where, for a vertex x, R_x denotes the unique ray of T which starts at x.

Theorem 3.1. *Let T be a one-ended tree, and $R=\langle w_0, w_1, \ldots\rangle$ a ray of T. Then*

$$a_{w_0}(T)=\prod_{n\geqslant 0} a_{w_n}(T_n),$$

where, for $n\geqslant 0$, T_n denotes the component T^{w_n} of $T\backslash R$ containing w_n. Furthermore, $a(T)=a_{w_0}(T)$ if w_0 is a fixed point when T is not fixed-point free.

Proof. We introduce a few notations. For $n\geqslant 0$:

- T'_n is the subtree of T induced by the set of vertices $\geqslant w_n$;
- $T'_{n+1} := \bigcup\{T_{n+1}(x): x\in U_{w_{n+1}}\backslash G_{w_{n+1}}[w_n]\}$;
- $T''_0 := T'_0 = T_0$;
- $s_n := s_{w_n}(T_n)$, $s'_n := s_{w_n}(T'_n)$, $s''_n := s_{w_n}(T''_n)$, and $m_n := |G_{w_{n+1}}[w_n]|$.

Notice that, since the ray R is fixed by any element of G_{w_0}, if A is an asymmetrizing set of (T, w_0) then, for any $n\geqslant 0$, $A\cap V(T_n)\in\mathscr{A}_{w_n}(T_n)$; and conversely any set $\bigcup_{n\geqslant 0} A_n$ with $A_n\in\mathscr{A}_{w_n}(T_n)$, $n\geqslant 0$, is an asymmetrizing set of (T, w_0). Hence

$$a(T)\leqslant a_{w_0}(T)=\prod_{n\geqslant 0} a_n.$$

We are done if w_0 is a fixed point. Assume that T has no fixed point. We distinguish three cases.

Case 1: $a'_n = 0$ *for some* $n \geqslant 0$.

$a'_n = 0$ implies that there are $p \geqslant n$ and an automorphism s such that $s(w_i) = w_i$ if and only if $i \geqslant p$. Hence $a_p = 0$, and

$$0 \leqslant a(T) \leqslant a_{w_0}(T) = \prod_{n \geqslant 0} a_n = 0.$$

Case 2: $0 < a'_n < \omega$ *for every* $n \geqslant 0$.

Then clearly $0 < a_n < \omega$ for any $n \geqslant 0$. Thus $a_{w_0}(T) = \prod_{n \geqslant 0} a_n = 2^\omega$, since any positive asymmetrizing number is $\geqslant 2$. Besides, for any $n \geqslant 0$, $a_{w_n}(T) = a'_n \prod_{i > n} a_i \neq 0$, thus $a(T) \geqslant 2^\omega$ by [3, Lemma 5.1]; hence $a(T) = a_{w_0}(T) = 2^\omega$.

Case 3. *All* a'_n *are* >0 *and at least one of them is infinite.*

Then there is a p such that a'_n is infinite for every $n \geqslant p$. Without loss of generality we can suppose that a_n is infinite for any n. We will prove that $a(T) \geqslant \prod_{n \geqslant 0} a_n$. We define by induction sets $\boldsymbol{A}_n$ and $\boldsymbol{B}_n$ of cardinality a'_n and a_n of pairwise nonsimilar asymmetrizing sets of (T_n, w_n) and (T'_n, w_n), respectively.

Since a_0 is infinite, there are two disjoint sets $\boldsymbol{A}_0$ and $\boldsymbol{B}_0$ of cardinality a_0 such that their union is a set of pairwise nonsimilar asymmetrizing sets of (T_0, w_0). Suppose that $\boldsymbol{A}_n$ and $\boldsymbol{B}_n$ are defined for some $n \geqslant 0$. Let $\boldsymbol{C}_n$ be a set of cardinality a''_{n+1} of pairwise nonsimilar asymmetrizing sets of (T''_{n+1}, w_{n+1}).

If $m_n = 1$ then $a_{n+1} = a''_{n+1}$; let:

$$\boldsymbol{A}_{n+1} := \{A \cup C : A \in \boldsymbol{A}_n \text{ and } C \in \boldsymbol{C}_n\},$$

$$\boldsymbol{B}_{n+1} := \boldsymbol{C}_n.$$

Assume now that $m_n > 1$. Since a'_n is infinite, $a'_{n+1} = a'^{m_n}_n \cdot a''_{n+1}$. For every $x \in G_{w_{n+1}}[w_n]$ denote by σ_x an element of $G_{w_{n+1}}$ such that $\sigma_x(w_n) = x$. Since $|\boldsymbol{A}_n| = a'_n$ is infinite and $m_n \leqslant a'_n$ by Remark 2.1(iii) of [3], there is a family $(f_{\xi,n})_{\xi < a'^{m_n}_n}$ of injections from $G_{w_{n+1}}[w_n]$ into $\boldsymbol{A}_n$ such that their images are pairwise disjoint. Finally let

$$\boldsymbol{D}_n := \{\bigcup \{\sigma_x \circ f_{\xi,n}(x) : x \in G_{w_{n+1}}[w_n]\} \text{ and } \xi < a'^{m_n}_n\},$$

$$\boldsymbol{A}_{n+1} := \{C \cup D : C \in \boldsymbol{C}_n \text{ and } D \in \boldsymbol{D}_n\},$$

$$\boldsymbol{B}_{n+1} := \{A \cap V(T'_{n+1}) : A \in \boldsymbol{A}_{n+1}\}.$$

In both cases $\boldsymbol{A}_{n+1}$ and $\boldsymbol{B}_{n+1}$ are then sets of cardinality a'_{n+1} and a_{n+1} of pairwise nonsimilar asymmetrizing sets of (T'_{n+1}, w_{n+1}) and (T_{n+1}, w_{n+1}), respectively.

Let $B := \bigcup_{n \geqslant 0} B_n$, where $B_n \in \boldsymbol{B}_n$, $n \geqslant 0$. If $\sigma \in \mathrm{Aut}(T, B)$, then $\sigma B_0 = B_0$ since, by the definitions of $\boldsymbol{A}_0$ and $\boldsymbol{B}_0$, for any $n > 0$ and $x \in G_{w_n}[w_0]$, $B_n \cap V(T_x)$ is not similar with B_0. Thus $\sigma w_0 = w_0$, and this implies that σ fixes the ray R. Hence $\sigma B_n = B_n$ for any $n \geqslant 0$. Therefore the restriction of σ to $V(T'_n)$ is the identity, since B_n is an asymmetrizing set of T. Furthermore, by their definitions, two different such asymmetrizing sets are nonsimilar. Hence

$$a(T) \geqslant \prod_{n \geqslant 0} |\boldsymbol{B}_n| = \prod_{n \geqslant 0} a_n.$$

(c) Therefore

$$a(T) = a_{w_0}(T) = \prod_{n \geq 0} a_n.$$

And this completes the proof. □

We get immediately the following.

Corollary 3.2. *Let T be a one-ended tree without fixed point. Then $a(T) = a_w(T)$ for any vertex w.*

3.3. To get a similar result for trees containing a double ray, we have to extend in a way some results of [3] about upward extendable trees. The next lemma is obtained by relativization of different results of Section 4 of [3] to a particular subgroup of automorphisms. Its proof, which is exactly the same as those given in [3], with G instead of Aut T, is thus omitted. We recall that, for a rooted tree (T, w), if G is a subgroup of Aut T, then $G_w := G \cap \mathrm{Aut}(T, w)$.

Given a tree T we will denote by $\mathcal{G}(T)$ the set of subgroups G of Aut T satisfying the following condition: if $\sigma \in \mathrm{Aut}\, T$ is such that σx belongs to the G-orbit $G[x]$ of x for every $x \in V$, then $\sigma \in G$.

Lemma 3.4. *Let (T, w) be an upward extendable tree, and $G \in \mathcal{G}(T)$. The following are equivalent:*

(i) *T is G_w-asymmetrizable;*
(ii) *T is G-asymmetrizable;*
(iii) $a_{G_w}(T) = 2^{|T|}$;
(iv) $a_G(T) = 2^{|T|}$;
(v) $m_G(x) \leq 2^{|T_x|}$ *for every* $x \in V \setminus \{w\}$.

Theorem 3.5. *Let T be a tree containing a double ray, such that (T^x, x) is asymmetrizable for every $x \in V(T_*)$; and let $w \in V(T_*)$. The following are equivalent:*

(i) *(T, w) is asymmetrizable;*
(ii) *T is asymmetrizable;*
(iii) $a_w(T) = \prod_{x \in V(T_*)} a_x(T^x)$;
(iv) $a(T) = \prod_{x \in V(T_*)} a_x(T^x)$;
(v) $m(x) \leq \prod_{y \in V(T_* \cap T_x)} a_y(T^y)$ *for every* $x \in V(T_*) \setminus \{w\}$.

Proof. (a) Let $x \in V(T_*)$. The subtree T^x is rayless, and asymmetrizable by hypothesis. Define the cardinal κ_x in the following way: If $a_x(T^x)$ is infinite, then by Proposition 2.1 it is of the form 2^κ for some cardinal κ; put $\kappa_x := \kappa$. If, otherwise, $a_x(T^x)$ is finite, then put $\kappa_x := 1$.

Define then the family $(T^{(x)})_{x\in V(T_*)}$ of pairwise disjoint trees such that, $T^{(x)}$ is the tree reduced to the vertex x if $\kappa_x = 1$, and is a regular tree of degree κ_x with $V(T^{(x)} \cap T_*) = \{x\}$, otherwise. For any $x \in V(T_*)$, $|T^{(x)}| = \kappa_x$ and $a_x(T^{(x)}) = 2^{\kappa_x}$ (by [1, 3.2] or by [3, 4.12] any regular tree T is asymmetrizable with $a(T) = 2^{|T|}$, moreover, this is a simple consequence of Lemma 3.4 with $G = \operatorname{Aut} T$, since condition (v) is trivially satisfied).

Now let $T' := T_* \cup \bigcup_{x\in V(T_*)} T^{(x)}$. Clearly $(T')_* = T'$,

$$|T'| = \sum_{x\in V(T_*)} \kappa_x \quad \text{and} \quad |(T')_x| = \sum_{y\in V(T_*\cap T_x)} \kappa_y \quad \text{for any } x \in V(T_*).$$

Finally let $G := \{\sigma \in \operatorname{Aut} T' : \sigma | V(T_*) = \sigma' | V(T_*) \text{ for some } \sigma' \in \operatorname{Aut} T\}$. G is an element of $\mathcal{G}(T')$. Indeed, if an automorphism σ of T' is such that $\sigma x \in G[x]$ for every $x \in V(T')$, then on the one hand, by the definition of G, $V(T_*)$ is invariant with respect to σ; and on the other hand, (T^x, x) and $(T^{\sigma x}, \sigma x)$ are isomorphic for any $x \in V(T_*)$. Hence there is $\sigma' \in \operatorname{Aut} T$ such that $\sigma | V(T_*) = \sigma' | V(T_*)$.

(b) Notice that $a_w(T) \leqslant \prod_{x\in V(T_*)} a_x(T^x)$. This is obvious since, for any $x \in V(T_*)$, the intersection of any asymmetrizing set of T with $V(T^x)$ is an asymmetrizing set of (T^x, x). More generally

$$a_x(T^x) \leqslant \prod_{y\in V(T_*\cap T_x)} a_y(T^y) \quad \text{for every } x \in V(T_*).$$

(c) If (T, w) is asymmetrizable then T' is G-asymmetrizable:

Suppose that T' is not G-asymmetrizable. Then, by 3.4(v), there is $x \in V(T')\setminus\{w\}$ such that

$$m_{G_w}(x) > 2^{|(T')_x|} = \prod_{y\in V(T_*\cap T_x)} 2^{\kappa_y} = \prod_{y\in V(T_*\cap T_x)} a_y(T^y).$$

Thus $x \in V(T_*)$ since T^y is asymmetrizable for every $y \in V(T_*)$. Hence

$$m(x) = m_{G_w}(x) > \prod_{y\in V(T_*\cap T_x)} a_y(T^y) \geqslant a_x(T_x)$$

by (b). Thus, by [3, Remark 2.1(iii)], (T, w) is not asymmetrizable.

(d) Also, if T' is G-asymmetrizable, then T is asymmetrizable with

$$a(T) = a_G(T') = \prod_{x\in V(T_*)} a_x(T^x):$$

We will associate with every $x \in V(T_*)$ two sets $\boldsymbol{R}(x)$ and $\boldsymbol{R}'(x)$.

Case 1: $\kappa_x \neq 1$.

Let $\boldsymbol{R}(x)$ (resp. $\boldsymbol{R}'(x)$) be a set of one representative of each similarity class of asymmetrizing sets of (T^x, x) (resp. $(T^{(x)}, x)$) which do not contain x. Since $\boldsymbol{R}'(x)$ and $\boldsymbol{R}(x)$ have the same cardinality 2^{κ_x}, there is a bijection f_x between these two

sets. These bijections can be chosen so that, if x and y are similar in T, then, for any $A \in \boldsymbol{R}'(x)$ and $B \in \boldsymbol{R}'(y)$, A and B are G-similar in T' if and only if $f_x(A)$ and $f_y(B)$ are similar in T.

Case 2: $\kappa_x = 1$.

Let $\boldsymbol{R}'(x) = \{\emptyset\}$, and $\boldsymbol{R}(x)$ be the set having as single element an asymmetrizing set of (T^x, x) which does not contain x. And let f_x be the bijection between $\boldsymbol{R}'(x)$ and $\boldsymbol{R}(x)$. Let

$$\mathscr{A}' := \{A \in \mathscr{A}_G(T'): A \cap (V(T^{(x)}) \setminus \{x\}) \in \boldsymbol{R}'(x) \text{ for any } x \in V(T_*)\}.$$

For any G-asymmetrizing set A of T' there is obviously an $A' \in \mathscr{A}'$ which is G-similar with A. Hence $a_G(T')$ is equal to the cardinality of the set of G-similarity classes of elements of $\mathscr{A}'$.

For any $A \in \mathscr{A}'$ define

$$A_f := A \cap V(T_*) \cup \bigcup \{f_x(A \cap (V(T^{(x)}) \setminus \{x\})): x \in V(T_*)\}.$$

Let A, $B \in \mathscr{A}'$ be such that $B_f = \sigma[A_f]$ for some $\sigma \in \operatorname{Aut} T$. Since T_* is stable with respect to $\operatorname{Aut} T$, we have $B \cap V(T_*) = \sigma[A \cap V(T_*)]$. Let $x \in V(T_*)$ and $y = \sigma x$. Then

$$f_y(B \cap (V(T^{(y)}) \setminus \{y\})) = \sigma[f_x(A \cap (V(T^{(x)}) \setminus \{x\}))],$$

thus, by the definition of f_x and f_y, there is an automorphism σ_{xy} between $(T^{(x)}, x)$ and $(T^{(y)}, y)$. Hence

$$\sigma^* := \sigma | V(T_*) \cup \bigcup \{\sigma_{x,\sigma x}: x \in V(T_*)\}$$

is an element of G such that $B = \sigma^*[A]$.

Therefore, on the one hand, taking $A = B$, we get $\sigma^* = 1$ since A is G-asymmetrizing. Thus $\sigma | V(T_*) = \sigma^* | V(T_*) = 1$, and for any $x \in V(T_*)$, $\sigma | V(T^x) = 1$ since $A_f \cap (V(T^x)$ is an asymmetrizing set of (T^x, x). Hence $\sigma = 1$, and A_f is an asymmetrizing set of T.

On the other hand, if A and B are non-G-similar elements of $\mathscr{A}'$, then A_f and B_f are nonsimilar asymmetrizing sets of T. Thus

$$a(T) \geqslant a_G(T') = 2^{|T'|} = \prod_{x \in V(T_*)} 2^{\kappa_x} = \prod_{x \in V(T_*)} a_x(T^x)$$

by 3.4 and (a), hence $a(T) = \prod_{x \in V(T_*)} a_x(T^x) = a_G(T')$ by (b).

(e) We can now prove the theorem. The implications (iv) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Leftarrow$ (iii) are obvious. (i) $\Rightarrow$ (ii) $\Rightarrow$ (iv) are consequences of (c) and (d). (iv) $\Rightarrow$ (iii) is a consequence of (b) and of the inequality $a(T) \leqslant a_w(T)$. (iii) $\Rightarrow$ (v) since, by [3, Remark 2.1(iii)], if (T, w) is asymmetrizable, then by (b),

$$m(x) \leqslant a_x(T^x) \leqslant \prod_{y \in V(T_* \cap T_x)} a_y(T^y) \quad \text{for every } x \in V(T_*) \setminus \{w\}.$$

Conversely assume that

$$m(x) \leqslant \prod_{y \in V(T_* \cap T_x)} a_y(T^y) = 2^{|(T')_x|} \quad \text{for every } x \in V(T_*) \setminus \{w\}.$$

Since $(T(y), y)$ is asymmetrizing for every $y \in V(T_*)$, then $m_G(x) \leqslant 2^{|(T')_x|}$ for every $x \in V(T_*) \setminus \{w\}$. Hence, by Lemma 3.4, T' is G-asymmetrizable. Therefore T is asymmetrizable by (d). Thus (v) $\Rightarrow$ (ii). And this completes the proof. □

With this last result we are now able to prove the assertion announced at the end of [3].

Corollary 3.6. *If an asymmetrizable tree T contains a ray, then $a(T) = 2^\kappa$ for some cardinal κ.*

Proof. By Theorems 3.1 and 3.5 $a(T) = \prod_{x \in V(T_*)} a_x(T^x)$. The result is then a consequence of the facts that, if $a_x(T^x)$ is finite then $a_x(T^x) \geqslant 2$, and if $a_x(T^x)$ is infinite then $a_x(T^x) = 2^{\kappa_x}$ for some cardinal κ_x by Proposition 2.1. □

Remark 3.7. Theorem 3.5 could have been stated in a more general form by considering a subgroup $G \in \mathcal{G}(T)$ instead of Aut T itself. Obvious modifications in the given proof would have led to the extended result. But we will show that this generalization (Proposition 3.8) can easily be obtained by applying Theorem 3.5 to a modified tree. Notice that Theorem 3.5 itself was proved by applying Lemma 3.4 in that manner. Therefore, since Proposition 3.8 is also a generalization of that lemma, we can say that Lemma 3.4 and Theorem 3.5 are 'self-refining'.

Proposition 3.8. *Let T be a tree containing a double ray, and $G \in \mathcal{G}(T)$, such that T^x is G_x-asymmetrizable for every $x \in V(T_*)$; and let $w \in V(T_*)$. The following are equivalent:*

(i) *T is G_w-asymmetrizable;*
(ii) *T is G-asymmetrizable;*
(iii) $a_{G_w}(T) = \prod_{x \in V(T_*)} a_{G_x}(T^x)$;
(iv) $a_G(T) = \prod_{x \in V(T_*)} a_{G_x}(T^x)$;
(v) $m_{G_w}(x) \leqslant \prod_{y \in V(T_* \cap T_x)} a_{G_y}(T^y)$ *for every* $x \in V(T_*) \setminus \{w\}$.

Proof. (a) Let $(S^x)_{x \in V}$ be a family of pairwise disjoint rayless trees such that, $V(S^x \cap T) = \{x\}$, (S^x, x) is isomorphic with an element of the set N_λ defined in Lemma 2.3, where λ is a cardinal $> |T|$, and such that (S^x, x) and (S^y, y) are similar if and only if x and y are G-similar. Notice that $a_x(S^x) = 2$, and by the definition of N_λ, x is the only vertex of S^x whose degree is λ. Define then

$$T_G := T \cup \bigcup_{x \in V} S^x.$$

Clearly $(T_G)_* = T_*$, and $(T_G)^x := T^x \cup \bigcup_{y \in V(T_*)} S^y$ for every $x \in V(T_*)$. Besides the vertices of T are those of T_G whose degrees are λ. Hence, by the last property of the S^x, $\sigma x \in G[x]$ for any automorphism σ of T_G and any $x \in V$. Thus, by the definition of $\mathscr{G}(T)$, the restriction of σ to the set V is an element of G; and conversely any element σ of G can be extended to an automorphism σ_G of T_G by taking the identity on the complement of V.

(b) Let $A \in \mathscr{A}(T_G)$ and $\sigma \in G \cap \mathrm{Aut}(T, A \cap V)$. Then $\sigma_G \in \mathrm{Aut}(T_G, A)$, thus $\sigma_G = 1$; hence $\sigma = 1$ which implies that $A \cap V \in \mathscr{A}_G(T)$. Besides, if $B = \sigma A$ for some $\sigma \in \mathrm{Aut}\, T_G$ then $B \cap V = (\sigma | V)[A]$. Hence $a_G(T) \leqslant a(T_G)$.

Conversely if $A \in \mathscr{A}_G(T)$ and, for every $x \in V$, A_x is an asymmetrizing set of (S^x, x) which does not contain x, then

$$A_G := A \cup \bigcup_{x \in V} A_x \in \mathscr{A}(T_G).$$

Notice that, for any $x \in V$, $a_x(S^x) = 2$ implies that any two asymmetrizing sets of (S^x, x) not containing x are $\mathrm{Aut}(S^x, x)$-similar. Hence, if A, $B \in \mathscr{A}_G(T)$ are G-similar, then A_G and B_G are similar. Thus $a(T_G) \leqslant a_G(T)$.

Therefore we proved that T is G-asymmetrizable if and only if T_G is asymmetrizable, with $a_G(T) = a(T_G)$. Using exactly the same arguments we can prove that, on the one hand T is G_w-asymmetrizable if and only if (T_G, w) is asymmetrizable, with $a_{G_w}(T) = a_w(T_G)$; and on the other hand, for any $x \in V(T_*)$, $(T_G)^x$ is asymmetrizable since T^x is G_x-asymmetrizable, with $a_{G_x}(T^x) = a_x(T^x)$.

Consequently we get the equivalence of conditions (i), (ii), (iii) and (iv) of Proposition 3.8 by that of the respective conditions of Theorem 3.5.

Now T is G-asymmetrizable if and only if T_G is asymmetrizable, hence by Theorem 3.5 if and only if

$$m(x) \leqslant \prod_{y \in V((T_G)_* \cap (T_G)_x)} a_y((T_G)^y) \quad \text{for every } x \in V(T_*) \setminus \{w\};$$

thus, since $(T_G)_* \cap (T_G)_x = T_* \cap T_x$ and $m_{G_w}(x) = m(x)$, if and only if

$$m_{G_w}(x) \leqslant \prod_{y \in V(T_* \cap T_x)} a_{G_y}(T^y) \quad \text{for every } x \in V(T_*) \setminus \{w\}.$$

This proves the equivalence of conditions (i) and (v), and so completes the proof. □

References

[1] L. Babai, Asymmetric trees with two prescribed degrees. Acta Math. Hungar. 29 (1977) 193–200.

[2] R. Halin, Über unendliche Wege in Graphen, Math. Ann. 157 (1964) 125–137.

[3] N. Polat and G. Sabidussi, Asymmetrising sets in trees, Discrete Math. 95 (this Vol.) 1991 271–289.

[4] N. Polat and G. Sabidussi, Fixed elements of infinite trees (DMS 89-1, Université de Montréal, 1989, preprint), Discrete Appl. Math., to appear.

Discrete Mathematics 95 (1991) 303–319
North-Holland

Excluding infinite minors

Neil Robertson *

Department of Mathematics, Ohio State University, 231 West 18th Ave., Columbus, OH, 43210, USA

Paul Seymour

Bellcore, 445 South St., Morristown, NJ 07960, USA

Robin Thomas

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

Received 15 December 1989
Revised 4 April 1990

Abstract

Robertson, N., P. Seymour and R. Thomas, Excluding infinite minors, Discrete Mathematics 95 (1991) 303–319.

Let κ be an infinite cardinal, and let H be either a complete graph with κ vertices, or a tree in which every vertex has valency κ. What can we say about graphs G which (i) have no minor isomorphic to H, or (ii) contain no subgraph which is a subdivision of H?

These four questions are answered for each infinite cardinal κ. In each case we find that there corresponds a necessary and sufficient structural condition (or, in some cases, several equivalent conditions) for G not to contain H in the appropriate way. We survey these results and a number of related theorems.

1. Introduction

Intuitively, a graph H is a *minor* of a graph G if H can be obtained from a subgraph of G by contraction. (Graphs in this paper may be infinite, and may have loops or multiple edges.) More precisely, let us say that *H is isomorphic to a minor of G* (or G has an *H-minor*, for brevity) if for each vertex $v \in V(H)$ there is a non-null connected subgraph $\alpha(v)$ of G, and for each edge $e \in E(H)$ there is an edge $\alpha(e)$ of G, such that:

(i) for distinct $v_1, v_2 \in V(H)$, $\alpha(v_1)$ and $\alpha(v_2)$ are disjoint,
(ii) for distinct $e_1, e_2 \in E(H)$, $\alpha(e_1) \neq \alpha(e_2)$,
(iii) for $v \in V(H)$ and $e \in E(H)$, $\alpha(e) \notin E(\alpha(v))$,
(iv) if $e \in E(H)$ has distinct ends $v_1, v_2 \in V(H)$ then $\alpha(v_1)$, $\alpha(v_2)$ both contain

* This research was carried out under a consulting agreement with Bellcore.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

ends of $\alpha(e)$, while if $e \in E(H)$ is a loop with end v then $\alpha(v)$ contains both ends of $\alpha(e)$.
(In particular, the subgraphs $\alpha(v)$ may be infinite, and so the intuitive definition 'obtainable from a subgraph by contracting edges' must be handled with caution.)

The first two authors proved, in a long series of papers culminating in [12, 13], the following conjecture of Wagner.

1.1. *For any infinite sequence $G_1, G_2, \ldots$ of finite graphs there exist $j > i \geqslant 1$ such that G_i is isomorphic to a minor of G_j.*

The third author disproved the extension of 1.1 to infinite graphs [19], in the following.

1.2. *There is an infinite sequence $G_1, G_2, \ldots$ of graphs such that for all $j > i \geqslant 1$, G_i is not isomorphic to a minor of G_j.*

However, the graphs in this counterexample are uncountable; and it remains open to decide the following.

1.3. Problem. *Is there an infinite sequence $G_1, G_2, \ldots$ of countable graphs such that for all $j > i \geqslant 1$, G_i is not isomorphic to a minor of G_j?*

The results we wish to survey here were motivated by an attempt to answer 1.3 negatively. Although the attempt was unsuccessful, we have discovered a great number of new decomposition theorems for infinite graphs of independent interest.

The method of proof of 1.1 was the following: Suppose that $G_1, G_2, \ldots$ is as in 1.1. We may assume that none of $G_2, G_3, \ldots$ has a G_1-minor, and so there is a finite complete graph H say such that $G_2, G_3, \ldots$ all have no H-minor. But there is a theorem that for any finite complete graph H, all finite graphs with no H-minor have a restricted structure (they are tree-structures of pieces which more or less have bounded genus). It suffices then to show that if $G_2, G_3, \ldots$ is an infinite sequence of finite graphs each with this restricted structure, then there exist $j > i \geqslant 2$ such that G_i is isomorphic to a minor of G_j, and this can be done. Thus, the heart of the proof is the theorem about the structure of graphs with no H-minor. The analogous approach to 1.3 would require a theorem about the structure of graphs with no $K_{\aleph_0}$-minor; and we have indeed been able to obtain such a theorem, and a generalization to complete graph minors of all other cardinalities.

For κ a regular uncountable cardinal, a theorem of Jung [5] says that G has a K_κ-minor if and only if G topologically contains K_κ (defined later). For κ singular or $\kappa = \aleph_0$ this is not true, but in all cases there is a similar structure theorem for graphs not topologically containing K_κ. We also study the structure of graphs not

containing the tree T_κ of valency κ (either as a minor or topologically). These problems are closely related to a cops-and-robber game played on a graph, where $<\kappa$ cops try to corner a robber. It turns out, for κ uncountable, that the robber can survive if and only if the graph has a T_κ-minor; and there is a particularly simple kind of survival strategy if and only if there is a K_κ-minor.

The paper is organized as follows: In Section 2 we introduce the game and certain survival strategies (called 'escapes' and 'havens'), and describe their connection with minors of the graph. In Section 3 we observe that the non-existence of escapes is equivalent with the graph having a certain kind of decomposition. Sections 4–6 describe different kinds of decompositions, appropriate to different excluded objects. In Section 7 we discuss the structure corresponding to excluding $K_{\aleph_0}$-minors, which slipped through the net of the previous chapters. Then in Section 8, all the results are summarized.

2. Cops and a robber

Here is a game, played on a graph G. There are two players, one controlling the cops and the other the robber. There is a fixed cardinal κ (finite or infinite) and at any time the cardinality of the set of cops in play is constrained to be less than κ. The object of the cop player is to corner the robber, and the robber attempts to survive uncaptured. The robber occupies a vertex of the graph, and may at any time run at great speed along a finite path of the graph to another vertex. He is not permitted to run through a cop, however. Each cop is placed on a vertex of the graph or is temporarily removed from the graph. In the cop player's turn, he may either remove cops from the graph, or place new cops on arbitrary vertices of the graph, subject only to the constraint that fewer than κ cops may be in place in the graph at any time. The robber makes his response after the cop player has declared where he intends to place his new cops if any and before the new cops are actually in position. This is a full knowledge game, and in particular the cops know where the robber is; the problem is to capture him by landing a cop on the same vertex. (There is another version of the game on finite graphs, investigated by La Paugh [8] and Kirousis and Papadimitriou [6], where the robber is invisible, but that is different, and will not concern us. Our game can also be played transfinitely, where 'limit moves' are permitted, but we shall not consider this. For us, each turn is the tth turn for some integer t.)

For instance, if the graph G is a finite tree, then the cop player can win with only two cops at his disposal. He places cop 1 on some vertex v_1, and examines which component C of $G\backslash v_1$ contains the robber. He chooses a vertex v_2 of C adjacent to v_1 and places cop 2 there, and removes cop 1. By repeating this process the robber is eventually trapped in a leaf of the tree and captured. In general, if the cop player wins using $<\kappa$ cops, we say that '$<\kappa$ cops can search the graph'. Thus, <3 cops can search any finite tree.

On the other hand, if G is an infinite graph, containing a ray R (a *ray* is a 1-way infinite path) then $<\aleph_0$ cops cannot search the graph. For the robber can survive by remaining in the infinite part of R, further along R than any of the finitely many vertices occupied by cops. (When new cops land, he just runs further along R.)

We remark that when κ is infinite, there is no point in the cop player ever removing from a graph any cop which has already been positioned; for it costs nothing to create a new cop instead. In the finite case, however, it is evidently important to reuse the same cops.

Let G be a graph and $X \subseteq V(G)$. We call the vertex set of a component of $G \backslash X$ an *X-flap*. Let us state the game more formally. A *position* is a pair (X, F) where $X \subseteq V(G)$ with $|X| < \kappa$, and F is an X-flap. We set $X_1 = \emptyset$ and the game starts by the robber choosing a $\emptyset$-flap F_1. In general, at the beginning of the tth turn we have a position (X_{t-1}, F_{t-1}). (X_{t-1} is the set of vertices occupied by cops; and F_{t-1} is the X_{t-1}-flap containing the robber. Since he can run arbitrarily fast, all that matters is which X_{t-1}-flap contains him.) The cop player chooses a set $X_t \subseteq V(G)$ with $|X_t| < \kappa$ such that $X_t \subseteq X_{t-1}$ or $X_{t-1} \subseteq X_t$, and then the robber chooses an X_t-flap F_t such that $F_t \cap F_{t-1} \neq \emptyset$ if possible. (If there is no such choice of F_t, the cops have captured the robber.) Thus if $X_t \subseteq X_{t-1}$ then $F_t \supseteq F_{t-1}$ and F_t is uniquely determined; but if $X_t \supseteq X_{t-1}$ then the robber makes a choice. This completes the tth turn.

Of interest to us at the moment are survival strategies for the robber. (Later we shall study search strategies for the cops.) We denote by $[V(G)]^{<\kappa}$ the set of all subsets of $V(G)$ of cardinality $<\kappa$. An *escape of order* κ is a function β with domain $[V(G)]^{<\kappa}$, such that:

(i) if $X, Y \in [V(G)]^{<\kappa}$ and $X \subseteq Y$ then $\beta(X)$ is the union of all X-flaps which intersect $\beta(Y)$, and

(ii) $\beta(\emptyset) \neq \emptyset$ (and hence $\beta(X) \neq \emptyset$ for all $X \in [V(G)]^{<\kappa}$).

2.1. *$<\kappa$ cops can search G if and only if there is no escape of order κ.*

Proof. If $<\kappa$ cops cannot search G, then for each $X \in [V(G)]^{<\kappa}$, let $\beta(X)$ be the set of all vertices $v \in V(G) - X$ such that the robber can guarantee to survive if the game begins at position (X, F), where F is the X-flap containing v. Then β is an escape of order κ. Conversely, if there is an escape β of order κ, the robber can survive by always choosing $F_t \subseteq \beta(X_t)$, and so $<\kappa$ cops cannot search G. □

An escape β of order κ is a *haven of order* κ if $\beta(X)$ is an X-flap for every $X \in [V(G)]^{<\kappa}$. A haven β is *convex* if for all $X, Y \in [V(G)]^{<\kappa}$,

$$(X \cap \beta(Y)) \cup (Y \cap \beta(X)) \cup (\beta(X) \cap \beta(Y)) \neq \emptyset.$$

For G, κ finite these three concepts are closely related, as we see from the following rather difficult theorem of [18].

2.2. *Let G be a finite graph and κ a finite cardinal. Then the following are equivalent*:

(i) *there is an escape of order κ in G,*

(ii) *there is a haven of order κ in G,*

(iii) *there is a convex haven of order κ in G.*

The equivalence of (ii) and (iii) also holds if G, κ are infinite; indeed, if $\kappa \geqslant \aleph_0$ then it is easy to prove that *every* haven of order κ is convex. However, the equivalence of (i) and (ii) fails for $\kappa \geqslant \aleph_1$, as we shall see. For any cardinal $\kappa \geqslant 2$, let T_κ be the tree in which every vertex has valency κ. The connection between the game and the minors of interest to us is given by the following (from which in particular we see that T_κ has an escape of order κ but no haven of order κ, when $\kappa \geqslant \aleph_1$).

2.3 [14, 17]. *For $\kappa \geqslant \aleph_1$, G has an escape of order κ if and only if G has a T_κ-minor , and G has a haven of order κ if and only if G has a K_κ-minor.*

Proof. The 'only if' parts are difficult, and we omit them, but the 'if' parts are easy. Suppose that G has a T_κ-minor. For each $X \in [V(G)]^{<\kappa}$, let $\beta(X)$ be the union of all X-flaps F such that the restriction of G to F has a T_κ-minor; then β is an escape of order κ. If G has a K_κ-minor, let α be as in the definition of 'minor'. For each $X \in [V(G)]^{<\kappa}$, let $\beta(X)$ be the X-flap which includes $V(\alpha(v))$ for some $v \in V(K_\kappa)$. (This exists, because X is too small to intersect every $V(\alpha(v))$, and is unique, because any two distinct $V(\alpha(v))$'s are joined by an edge of G.) Then β is a haven of order κ. $\square$

The significance of this result is that in practice it is often easier to construct a haven or an escape than to construct the desired minor directly; and conversely by exhibiting a search strategy for the cops we can sometimes prove that no T_κ-minor exists.

What about escapes and havens of order $\aleph_0$? That is answered by the following.

2.4 [17]. *For a graph G, the following are equivalent*:

(i) *G has an escape of order $\aleph_0$,*

(ii) *G has a haven of order $\aleph_0$,*

(iii) *G has a ray.*

Proof. Let R be a ray of G and for each finite $X \subseteq V(G)$ define $\beta(X)$ to be the unique F with X-flap $F \cap V(R)$ infinite; then β is a haven of order $\aleph_0$. Thus (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i). To see (i) $\Rightarrow$ (iii), here is a strategy for the cops: place a new cop on every vertex just visited by the robber (including those on the paths he runs along). This forces the robber either to trace out a ray, or to be captured, and so (i) $\Rightarrow$ (iii). $\square$

In fact, a stronger version of the equivalence of (ii) and (iii) above holds. Let us say two rays R_1, R_2 in G are *parallel* if for every finite $X \subseteq V(G)$, the unique X-flap F with $F \cap V(R_1)$ infinite has $F \cap V(R_2)$ infinite. This is an equivalence relation on the rays of G, and its equivalence classes are called the *ends* of G (see [4]). We see that, if we construct a haven as in the proof of 2.4 starting from a ray R, then two rays yield the same haven if and only if they belong to the same end. Moreover, since every haven of order $\aleph_0$ arises from a ray, we have in this sense the following.

2.5. *There is a natural* $1-1$ *correspondence between the ends of* G *and the havens in* G *of order* $\aleph_0$.

Escapes and havens of finite order in finite graphs do not correspond so closely to minors, despite 2.2. The $n \times n$ *grid* is the graph with vertex set $\{(i, j)$: $1 \leqslant i, j \leqslant n\}$ where (i, j) and (i', j') are adjacent if $|i' - i| + |j' - j| = 1$. 2.6 follows from the theorems of [11, 16,18].

2.6. *Let* G *be a finite graph. For* $n \geqslant 1$, *if* G *has an* $n \times n$ *grid minor, then* G *has a haven of order* n. *Conversely, if* G *has a haven of order* 2000^{n^5} *then* G *has an* $n \times n$ *grid minor.*

Thus, we have two differences between the finite and infinite case. First, 2.3 gives a more exact relationship between escapes, havens and minors than 2.6 does; but perhaps more surprisingly, in the finite case the minors corresponding to the existence of escapes and havens are neither trees nor complete graphs but grids.

3. Escapes and rayless tree-decompositions

A *graph-theoretic tree-decomposition* of a graph G is a pair (T, W), where T is a tree and $W = (W_t\colon t \in V(T))$ is a family of subsets of $V(G)$, such that:

(i) $\bigcup(W_t\colon t \in V(T)) = V(G)$, and for every edge $e \in E(G)$ some W_t contains both ends of e,

(ii) if $t_1, t_2, t_3 \in V(T)$, and t_2 lies on the path of T between t_1 and t_3, then $W_{t_1} \cap W_{t_3} \subseteq W_{t_2}$.

It has *width* $<\kappa$, where κ is a cardinal, if $|W_t| < \kappa$ for each $t \in V(T)$, and

$$\left|\bigcup_{j \geqslant 1} \bigcap_{i \geqslant j} W_{t_i}\right| < \kappa$$

for every ray of T with vertices $t_1, t_2, \ldots$ in order. (There need not be any κ' such that the decomposition has width $<\kappa$ if and only if $\kappa' < \kappa$, and so we shall

not attempt to define the 'width' of a decomposition.) A *rayless tree-decomposition* is a graph-theoretic tree-decomposition (T, W) such that T has no ray. There is the following nice connection between rayless tree-decompositions and escapes.

3.1 [17]. *For any cardinal κ, finite or infinite, a graph G has no escape of order κ if and only if it has a rayless tree-decomposition of width $<\kappa$.*

Proof. The 'if' part is easy. For if (T, W) is a rayless tree-decomposition of width $<\kappa$ then $<\kappa$ cops can search the graph, as follows: Choose $t_1 \in V(T)$, and place cops on W_{t_1}. Choose $t \in V(T)$ such that the robber lies in W_t, and let t_2 be a neighbour of t_1 on the path between t_1 and t. Remove the cops in $W_{t_1} - W_{t_2}$, and place cops on $W_{t_2} - W_{t_1}$. Now repeat the process. Since T is rayless, the robber is eventually captured. The 'only if' part for κ finite is a difficult theorem of [18] for finite graphs G, and can be extended to infinite graphs G by a compactness argument, using the result of [8, 20]. For κ infinite, however, the 'only if' part is easy. For suppose that G has no escape of order κ. Define $\beta(X)$, for each $X \in [V(G)]^{<\kappa}$, to be the union of all X-flaps F such that the restriction of G to F has no rayless tree-decomposition of width $<\kappa$. Since κ is infinite it is easy to see that β satisfies the first escape axiom, and hence does not satisfy the second. Hence $\beta(\emptyset) = \emptyset$, and G has a rayless tree-decomposition of width $<\kappa$, as required. □

3.1 has several corollaries. From 3.1, 2.2, and 2.6, we deduce a sharpening of a result of [11].

3.2 [16]. *For any integer $n \geqslant 1$, every finite graph with no $n \times n$ grid minor has a graph-theoretic tree-decomposition of width $<2000^{n^5}$.*

From 3.1 and 2.4 we obtain a theorem of Halin [3].

3.3. *G has no ray if and only if G has a rayless tree-decomposition of width $<\aleph_0$.*

From 3.1 and 2.3 we obtain the first of our main theorems.

3.4 [17]. *For $\kappa \geqslant \aleph_1$, G has no T_κ-minor if and only if G has a rayless tree-decomposition of width $<\kappa$.*

What about $T_{\aleph_0}$-minors? Excluding them does not give us any kind of rayless tree-decomposition, but a more complicated structure, and we postpone it.

4. Dissections

A *separation* of G is a pair (A, B) of subsets of $V(G)$ such that $A \cup B = V(G)$ and no edge joins a vertex of $A - B$ to a vertex of $B - A$. Two separations (A, B) (C, D) *cross* unless either $A \subseteq C$ and $D \subseteq B$, or $A \subseteq D$ and $C \subseteq B$, or $B \subseteq C$ and $D \subseteq A$, or $B \subseteq D$ and $C \subseteq A$. A *dissection* of G is a set $\mathscr{D}$ of separations of G, such that:

(i) if $(A, B) \in \mathscr{D}$ then $(B, A) \in \mathscr{D}$,
(ii) if $(A, B) \in \mathscr{D}$ then $A \neq V(G)$,
(iii) if (A, B), $(C, D) \in \mathscr{D}$ and $A \neq C$ then $B \neq D$,
(iv) no two members of $\mathscr{D}$ cross.

Let (T, W) be a graph-theoretic tree-decomposition of G. For each $e \in E(T)$, let T_1, T_2 be the components of $T \backslash e$, and let $A_e = \bigcup(W_t\colon t \in V(T_1))$, $B_e = \bigcup(W_t\colon t \in V(T_2))$. The union over all $e \in E(T)$ of the sets $\{(A_e, B_e), (B_e, A_e)\}$ is certainly a set of mutually noncrossing separations, but it may not be a dissection because conditions (ii) and (iii) above may be violated. However, it is easy to show that any dissection of a *finite* graph G arises in this way from some graph-theoretic tree-decomposition of G, and we find that all the relevant results about tree-decompositions of finite graphs, like 3.1 and 3.2, can be reformulated in terms of dissections. Thus, dissections provide another way to generalize tree-decompositions of finite graphs to infinite graphs. It is not equivalent, for there are dissections which do not arise from any graph-theoretic tree-decomposition. (Indeed, one can characterize those dissections $\mathscr{D}$ which do arise from graph-theoretic tree-decompositions; it is necessary and sufficient that $\bigcup(B_i\colon i \geqslant 1) = V(G)$ for every infinite sequence (A_i, B_i) $(i = 1, 2, \ldots)$ of distinct members of $\mathscr{D}$ with $A_1 \supseteq A_2 \supseteq \cdots$ and $B_1 \subseteq B_2 \subseteq \cdots$.)

Let $\mathscr{D}$ be a dissection. An *orientation* of $\mathscr{D}$ is a subset $\mathscr{P} \subseteq \mathscr{D}$, such that:

(i) if $(A, B) \in \mathscr{D}$ then one of (A, B), $(B, A) \in \mathscr{P}$,
(ii) if (A, B), $(C, D) \in \mathscr{P}$ then $B \nsubseteq C$.

If $\mathscr{D}$ arises from a graph-theoretic tree-decomposition (T, W) with T finite, one can verify that the orientations of $\mathscr{D}$ are in $1-1$ correspondence with $V(T)$; the orientation corresponding to $t_0 \in V(T)$ is the set of all separations

$$\left(\bigcup(W_t\colon t \in V(T_1)), \bigcup(W_t\colon t \in V(T_2))\right)$$

taken over all $e \in E(T)$, where T_1, T_2 are the components of $T \backslash e$ and $t_0 \in V(T_2)$. If $\mathscr{D}$ arises from (T, W) and T is infinite, then $\mathscr{D}$ may have orientations not corresponding to the vertices; in fact, it has one for each vertex and one for each end of T. Nevertheless, orientations provide us with a generalization of the concept of a vertex of a tree, expressed in the language of dissections.

Motivated by this, let us say a dissection $\mathscr{D}$ has *width* $<\kappa$ if for every orientation $\mathscr{P}$ of $\mathscr{D}$,

$$\left|\bigcup(B\colon (A, B) \in \mathscr{P})\right| < \kappa.$$

If the dissection arises from a graph-theoretic tree-decomposition (T, W), one can verify that this definition agrees with our definition of 'width $<\kappa$' for such a decomposition. Indeed, if $\mathscr{P}$ is determined by a vertex $t \in V(T)$, then $\bigcap(B\colon (A, B) \in \mathscr{P}) = W_t$, while if $\mathscr{P}$ is determined by an end of T containing a ray with vertices $t_1, t_2, \ldots$ in order, then

$$\bigcap(B\colon (A, B) \in \mathscr{P}) = \bigcup_{i \geq 1} \bigcap_{j \geq i} W_{t_j}.$$

Dissections and havens work well together, because of the following easy lemma. (The *order* of a separation (A, B) is $|A \cap B|$.)

4.1 [14]. *Let β be a convex haven in G of order κ, and let $\mathscr{D}$ be a dissection of G each member of which has order $<\kappa$. Then the set $\{(A, B) \in \mathscr{D}\colon \beta(A \cap B) \subseteq B\}$ is an orientation of $\mathscr{D}$.*

Now we can begin to discuss the second of our structure theorems, concerning the structure of graphs not topologically containing K_κ. A graph G *topologically contains* H if there is a function α with domain $V(H) \cup E(H)$, such that:

(i) for each $v \in V(H)$, $\alpha(v) \in V(G)$ and $\alpha(v_1) \neq \alpha(v_2)$ for distinct $v_1, v_2 \in V(H)$,

(ii) for each $e \in E(H)$, $\alpha(e)$ is a finite path of G with $E(\alpha(e)) \neq \emptyset$ and with ends $\alpha(v_1)$, $\alpha(v_2)$ where e has ends v_1, v_2 in H,

(iii) for distinct $e_1, e_2 \in E(H)$, the paths $\alpha(e_1)$ and $\alpha(e_2)$ are disjoint except possibly for their ends,

(iv) for each $v \in V(H)$ and $e \in E(H)$, $\alpha(v)$ is not an internal vertex of $\alpha(e)$.

A preliminary form of our result is the following.

4.2 [14, 15]. *For any graph G and $\kappa \geq \aleph_0$, G does not topologically contain K_κ if and only if there is a dissection of G of width $<\kappa$.*

Proof. 'Only if' is complicated and we omit it, but let us prove 'if'. Let $\mathscr{D}$ be a dissection of width $<\kappa$; then every member of $\mathscr{D}$ has order $<\kappa$, as is easily seen, by considering the orientation $\{(A, B) \in \mathscr{D}\colon A_0 \subseteq B \text{ or } B_0 \subseteq B, \text{ and } A_0 \neq B\}$, for each $(A_0, B_0) \in \mathscr{D}$. Suppose, for a contradiction, that G topologically contains K_κ, and let α be the corresponding function. Let $Z = \{\alpha(v)\colon v \in V(K_\kappa)\}$. For each $X \in [V(G)]^{<\kappa}$, let $\beta(X)$ be the X-flap containing some member of Z. (This exists since X is too small to include Z, and is unique since any two members of Z are joined by κ mutually disjoint paths, not all of which meet X.) Then β is a convex haven of order κ. By 4.1,

$$\mathscr{P} = \{(A, B) \in \mathscr{D}\colon \beta(A \cap B) \subseteq B\}$$

is an orientation of $\mathscr{D}$. Let $X = \bigcap(B\colon (A, B) \in \mathscr{P})$; then $|X| < \kappa$ since $\mathscr{D}$ has width $<\kappa$. Choose $v \in Z - X$. Since $v \notin X$, there exists $(A, B) \in \mathscr{P}$ with $v \in A - B$. But since $v \notin A \cap B$ it follows from the definition of β that $v \in \beta(A \cap B)$, and so $\beta(A \cap B) \nsubseteq B$, a contradiction. $\square$

Given that G does not topologically contain K_κ, can we prove the existence of a more concrete structure than just a dissection of width $<\kappa$? An example of [7] shows that G need not have a graph-theoretic tree-decomposition of width $<\kappa$, if $\kappa > \aleph_0$. Nevertheless, we can improve 4.2 by the use of 'well-founded tree-decompositions', which are midway in generality between graph-theoretic tree-decompositions and dissections.

A *well-founded tree* is a non-null partially ordered set $T = (V(T), \leqslant)$ such that:

(i) for each $s \in V(T)$, the set $\{t \in V(T): t \leqslant s\}$ is well-ordered by $\leqslant$,

(ii) every non-empty $X \subseteq V(T)$ has an *infimum*, that is, an element $z \in V(T)$ such that for all $t \in V(T)$, $t \leqslant z$ if and only if $t \leqslant x$ for all $x \in X$.

It follows that there is a unique element $s \in V(T)$ such that $s \leqslant t$ for all $t \in V(T)$, which we call the *root*. If $t_1, t_2, t_3 \in V(T)$, we say that t_2 is *between* t_1 and t_3 if $t_0 \leqslant t_2$, where t_0 is the infimum of $\{t_1, t_3\}$, and either $t_2 \leqslant t_1$ or $t_2 \leqslant t_3$. A *well-founded tree-decomposition of a graph* G is a pair (T, W), where $T = (V(T), \leqslant)$ is a well-founded tree and $W = (W_t: t \in V(T))$ is a family of subsets of $V(G)$, such that:

(i) $\bigcup(W_t: t \in V(T)) = V(G)$, and every edge of G has both ends in some W_t,

(ii) if $t_1, t_2, t_3 \in V(T)$ and t_2 is between t_1 and t_3 then $W_{t_1} \cap W_{t_3} \subseteq W_{t_2}$,

(iii) if C is a chain of T with supremum s (that is, C is totally ordered by $\leqslant$, and for all $t \in V(T)$, $c \leqslant t$ for all $c \in C$ if and only if $s \leqslant t$), then $\bigcap(W_c: c \in C) \subseteq W_s$.

We say that (T, W) has *width* $<\kappa$ if for every chain C of T,

$$\left|\bigcup_{c \in C} \bigcap_{t \geqslant c,\, t \in C} W_t\right| < \kappa.$$

(We observe, by taking $|C| = 1$, that this implies that $|W_t| < \kappa$ for all $t \in V(T)$; but it is somewhat stronger, for some chains may not have suprema.) If (T, W) is a graph-theoretic tree-decomposition and we regard T as a well-founded tree by choosing a root in the natural way, this definition of 'width $<\kappa$' agrees with the earlier one.

A well-founded tree T has *height* $<\kappa$ if every chain in T has order type $<\kappa$, regarded as an ordinal. Then we can vary 4.2 as follows.

4.3 [15]. *For any graph G and cardinal $\kappa \geqslant \aleph_0$, the following are equivalent*:

(i) *G does not topologically contain K_κ,*

(ii) *G has a well-founded tree-decomposition of width $<\kappa$.*

If κ is regular with $\kappa > \aleph_0$, these are equivalent to

(iii) *G has a well-founded tree-decomposition of width $<\kappa$ and height $<\kappa$.*

If $\kappa = \aleph_0$, (i), (ii) *are equivalent to*

(iv) *G has a graph-theoretic tree-decomposition of width $<\aleph_0$.*

Diestel [2] has independently proved a structural characterization of the graphs not topologically containing K_κ, in the case when $\kappa > \aleph_0$ and is regular. His result

uses a different generalization of finite tree-decompositions, but is similar to (and interderivable with) the equivalence of (i) and (iii) in 4.3.

The need for our rather curious definition of width is shown by the following example: Let $G = K_{\aleph_0}$, with $V(G) = \{v_0, v_1, v_2, \ldots\}$. Let T be a ray with vertices $t_0, t_1, t_2, \ldots$ in order, and for each $i \geq 0$ let $W_{t_i} = \{v_0, \ldots, v_i\}$. Let $W = (W_t\colon t \in V(T))$; then (T, W) is a graph-theoretic tree-decomposition of G and each $|W_t| < \aleph_0$, yet G topologically contains $K_{\aleph_0}$.

We could replace the condition about chains in the definition of the width of a well-founded tree-decomposition by the weaker condition that each $|W_t| < \kappa$, together with an extra condition that every chain has a supremum; for one can always add to T any suprema that are missing. However, doing so can increase the height of T, and on this definition graph-theoretic tree-decompositions would not be well-founded tree-decompositions unless they were rayless; so we prefer the version given.

5. Adhesion

So far, we have discussed excluding T_κ-minors, and excluding K_κ topologically, because these are the simpler theorems to state. Next, let us consider excluding K_κ-minors. For regular $\kappa > \aleph_0$, this is equivalent to excluding K_κ topologically, for a theorem of Jung [5] states the following.

5.1. *For $\kappa > \aleph_0$ regular, a graph has a K_κ-minor if and only if it topologically contains K_κ.*

There is a similar result for T_κ when κ is regular and uncountable. However, for κ singular or $\kappa = \aleph_0$, these are false; indeed, in those cases there is a graph with a K_κ-minor in which all vertices have valency $<\kappa$. In particular then, if κ is singular or $\kappa = \aleph_0$, the structure of 4.2 (ii) or 4.3 (ii) is not sufficiently restrictive to exclude K_κ-minors, and we need to impose an additional condition on the dissection.

Let $\mathscr{D}$ be a dissection of a graph G. We say that $\mathscr{D}$ has *adhesion* $<\kappa$ if for every orientation $\mathscr{P}$ of $\mathscr{D}$ there exists $\kappa' < \kappa$ such that for every $(A, B) \in \mathscr{P}$ there exists $(A', B') \in \mathscr{P}$ of order $\leq \kappa'$ with $A \subseteq A'$ and $B' \subseteq B$. Then we have the following.

5.2 [14]. *For all $\kappa > \aleph_0$, a graph has no K_κ-minor if and only if it has a dissection of width $<\kappa$ and adhesion $<\kappa$.*

Even the 'easy' part of this, the 'if' half, is not very easy, and so we omit the proof completely.

There is a version of 5.2 in terms of well-founded tree-decompositions,

analogous to 4.3 but we omit details—see [14]. Also, 5.2 requires that $\kappa > \aleph_0$. We do have a structure theorem for excluding $K_{\aleph_0}$-minors, but postpone it because it is quite different.

6. Linear decompositions

Our results so far were motivated by attempts to find infinite analogues of 3.2. However, 3.2 concerns excluding finite grids, and we really want to exclude T_κ or K_κ. Thus, one might think that the following result is more likely to have an interesting infinite analogue.

6.1 [1, 10]. *For any finite tree H, every finite graph with no H-minor has a tree-decomposition (T, W) of width $<|V(H)|$ such that T is a path.*

Let us define a *linear decomposition* (T, W) of G to consist of a Dedekind complete, linearly ordered set $T = (V(T), \leqslant)$, and a family $W = (W_t\colon t \in V(T))$ of subsets of $V(G)$, such that:

(i) $\bigcup(W_t\colon t \in V(T)) = V(G)$, and every edge of G has both ends in some W_t,

(ii) if $t_1, t_2, t_3 \in V(T)$ and $t_1 \leqslant t_2 \leqslant t_3$ then $W_{t_1} \cap W_{t_3} \subseteq W_{t_2}$,

(iii) if $s \in V(T)$ is the supremum or infimum of a non-null subset $C \subseteq V(T)$ then $\bigcap(W_c\colon c \in C) \subseteq W_s$. It has *width* $<\kappa$ if $|W_t| < \kappa$ for each t.

Seese [21] extended 6.1 to infinite graphs by a compactness argument, to yield the following.

6.2 [21]. *For any finite tree H, every graph with no H-minor has a linear decomposition of width $<|V(H)|$.*

For countable trees a similar result holds, because of the following.

6.3 [17]. *A graph does not topologically contain $T_{\aleph_0}$ if and only if it has a linear decomposition of width $<\aleph_0$.*

But for larger cardinals, there is no similar result because T_κ has a linear decomposition of width $<\aleph_1$ (enumerate the rays from left to right). For larger regular cardinals, the appropriate excluded object is not a tree but a clique.

6.4 [15]. *For $\kappa > \aleph_0$, regular, a graph does not topologically contain K_κ if and only if it has a linear decomposition of width $<\kappa$.*

Proof. We only show the easy 'if' half. Suppose that (T, W) is a linear decomposition of width $<\kappa$ of G, and yet G topologically contains K_κ. Thus there exists $X \subseteq V(G)$ with $|X| = \kappa$ such that any two members of X are joined by κ internally disjoint paths. For each $x \in X$, let $I_x = \{t \in V(T)\colon x \in W_t\}$. Then I_x

is a closed interval. For distinct x, x', $I_x \cap I_{x'} \neq \emptyset$, for otherwise there exists $t_0 \in V(t)$ such that $t < t_0$ for all $t \in I_x$ and $t > t_0$ for all $t \in I_{x'}$ (or vice versa); and this is impossible, because then every path of G between x and x' meets W_{t_0}, and $|W_{t_0}| < \kappa$. Thus $I_x \cap I_{x'} \neq \emptyset$, and so from Dedekind completeness there exists $t \in V(T)$ such that $t \in I_x$ for all $x \in X$, that is, $X \subseteq W_t$. But that is impossible since $|W_t| < \kappa$. □

For $\kappa = \aleph_\omega$, 6.4 is false. Indeed, we give a graph in [15] which has no $K_{\aleph_\omega}$-minor, and yet has no linear decomposition of width $<\aleph_\omega$.

A *well-ordered decomposition* is a linear decomposition (T, W) such that T is a well-order. The third of our main results is the following.

6.5 [17]. *For all $\kappa \geqslant \aleph_0$, a graph does not topologically contain T_κ if and only if it has a well-ordered decomposition of width $<\kappa$.*

Proof. Again, we shall only prove the 'if' part. Let (T, W) be a well-ordered decomposition of G of width $<\kappa$, and suppose that G topologically contains T_κ. Thus, there exists $X \subseteq V(G)$ with $|X| = \kappa$ such that each $x \in X$ is joined by κ paths, mutually disjoint except for x, to κ other members of X. For each $x \in X$, choose $s(x) \in V(T)$ maximal such that $x \in W_{s(x)}$. Choose $x \in X$ with $s(x)$ minimal. (This is possible since T is a well-order.) Certainly $s(x)$ is not the supremum of $V(T)$, and so there exists t say with $s(x) < t$ such that there is no $t' \in V(T)$ with $s(x) < t' < t$. Since there are κ paths from x to other members of X, and $|W_{s(x)} \cup W_t| < \kappa$, there is a path P between x and some $x' \in X - \{x\}$ such that $V(P) \cap (W_{s(x)} \cup W_t) = \{x\}$. It follows that $s(x') < t$, and so $s(x') \leqslant s(x)$, and equality does not hold since $x' \notin W_{s(x)}$. But this contradicts the choice of x. □

Let us say a linear decomposition (T, W) has *adhesion* $<\kappa$ if for every $s \in V(T)$ there exists $\kappa' < \kappa$ such that for every $t \in V(T)$ with $t \neq s$ there exist $t_1, t_2 \in V(T)$ with $t_1 < t_2$ such that there is no $t_3 \in V(T)$ with $t_1 < t_3 < t_2$ and $|W_{t_1} \cap W_{t_2}| \leqslant \kappa'$, where $s \leqslant t_1 < t_2 \leqslant t$ if $s < t$ and $t \leqslant t_1 < t_2 \leqslant s$ if $t < s$.

6.6 [17]. *For all $\kappa \geqslant \aleph_0$, a graph has no T_κ-minor if and only if it has a well-ordered decomposition of width $<\kappa$ and adhesion $<\kappa$.*

Thus, 6.6 is related to 6.5 as 5.2 is to 4.2. Corresponding to 6.3, we have the following.

6.7 [17]. *A graph has no $T_{\aleph_0}$-minor if and only if it has a linear decomposition of width $<\aleph_0$ and adhesion $<\aleph_0$.*

We also have variant forms of these results. A linear decomposition (T, W) is *scattered linear* if the set of rational numbers cannot be monotonely embedded into T.

6.8 [17]. *For all $\kappa \geqslant \aleph_0$, a graph does not topologically contain T_κ if and only if it has a scattered linear decomposition of width $<\kappa$.*

6.9 [17]. *For all $\kappa \geqslant \aleph_0$, a graph has no T_κ-minor if and only if it has a scattered linear decomposition of width $<\kappa$ and adhesion $<\kappa$.*

6.9 fills a gap left by 3.4, which did not cater for excluding $T_{\aleph_0}$. Another way is given by 6.11 below. A graph-theoretic tree-decomposition (T, W) is *scattered* if T has no $T_{\aleph_0}$-minor (or equivalently, does not topologically contain T_3).

6.10 [17]. *For any cardinal κ with $\mathrm{cf}(\kappa) = \omega$, a graph does not topologically contain T_κ if and only if it has a scattered graph-theoretic tree-decomposition of width $<\kappa$.*

6.10 does not hold in general if $\mathrm{cf}(\kappa) > \omega$. (There is a form which holds in general using 'scattered' well-founded tree-decompositions—see [17].) But its analogue for T_κ minors does hold in general, as follows.

6.11 [17]. *For all $\kappa \geqslant \aleph_0$, a graph has no T_κ-minor if and only if it has a scattered graph-theoretic tree-decomposition of width $<\kappa$ and adhesion $<\kappa$.*

(A graph-theoretic tree-decomposition (T, W) has *adhesion* $<\kappa$ if $|W_s \cap W_t| < \kappa$ for every pair s, t of adjacent vertices of T, and for every ray of T with vertices $t_1, t_2, \ldots$ in order, there exists $\kappa' < \kappa$ such that $|W_{t_i} \cap W_{t_{i+1}}| \leqslant \kappa'$ for infinitely many integers $i \geqslant 1$.)

7. Countable clique and grid minors

There remains one case not yet covered; what is the structure of graphs with no $K_{\aleph_0}$-minor? By analogy with 5.2 one might guess that it was necessary and sufficient that the graph have a dissection of width $<\aleph_0$ and adhesion $<\aleph_0$. That turns out to be sufficient but not necessary; indeed, that structure is equivalent to excluding a 'half-grid' minor. The *half-grid* is the graph with vertex set all pairs of integers (x, y) with $y \geqslant 0$, where (x, y) and (x', y') are adjacent if $|x - x'| + |y - y'| = 1$.

7.1 [14]. *A graph has no half-grid minor if and only if it has a dissection of width $<\aleph_0$ and adhesion $<\aleph_0$.*

Let us say a graph-theoretic tree-decomposition (T, W) is *narrow* if $|W_s \cap W_t|$ is finite for every pair of adjacent vertices s, t of T, and for every ray of T with vertices $t_1, t_2, \ldots$ in order, infinitely many of the numbers $|W_{t_i} \cap W_{t_{i+1}}|$ are equal.

(This is equivalent to the decomposition having adhesion $<\aleph_0$.) We can sharpen 7.1 as follows.

7.2 [14]. *A graph has no half-grid minor if and only if it has a narrow graph-theoretic tree-decomposition of width* $<\aleph_0$.

This is an extension of a result of Halin [4]. We have not been able to determine the structure corresponding to excluding the 'full' grid.

So what then is the appropriate structure for excluding $K_{\aleph_0}$-minors? Let (T, W) be a graph-theoretic tree-decomposition of G. For each $t \in V(T)$, we define the *torso at* t to be the graph with vertex set W_t in which u, v are adjacent if either they are adjacent in G or $u, v \in W_t \cap W_{t'}$ for some neighbour t' of t in T.

7.3 [14]. *A graph has no* $K_{\aleph_0}$*-minor if and only if it has a narrow graph-theoretic tree-decomposition* (T, W) *such that for each* $t \in V(T)$ *there is an integer* k *such that the torso at* t *has no* K_k*-minor.*

8. Summary

Let us summarize all these results.

8.1. *The following are equivalent*:

(i) *G has no ray,*

(ii) $<\aleph_0$ *cops can search* G,

(iii) *G has a rayless tree-decomposition of width* $<\aleph_0$.

8.2. *For all* $\kappa \geqslant \aleph_0$, *the following are equivalent*:

(i) *G does not topologically contain* T_κ,

(ii) *G has a well-ordered decomposition of width* $<\kappa$,

(iii) *G has a scattered linear decomposition of width* $<\kappa$.

If $\mathrm{cf}(\kappa) = \omega$, *these are equivalent to*

(iv) *G has a scattered graph-theoretic tree-decomposition of width* $<\kappa$.

If $\kappa = \aleph_0$, *these are equivalent to*

(v) *G has a linear decomposition of width* $<\kappa$.

8.3. *The following are equivalent*:

(i) *G has no* $T_{\aleph_0}$*-minor,*

(ii) *G has a scattered graph-theoretic tree-decomposition of width* $<\aleph_0$ *and adhesion* $<\aleph_0$,

(iii) *G has a well-ordered decomposition of width* $<\aleph_0$ *and adhesion* $<\aleph_0$,

(iv) *G has a linear decomposition of width* $<\aleph_0$ *and adhesion* $<\aleph_0$,

(v) *G has a scattered linear decomposition of width* $<\aleph_0$ *and adhesion* $<\aleph_0$.

8.4. *For all $\kappa > \aleph_0$, the following are equivalent*:
(i) *G has no T_κ-minor,*
(ii) *$<\kappa$ cops can search G,*
(iii) *G has a rayless tree-decomposition of width $<\kappa$,*
(iv) *G has a scattered tree-decomposition of width $<\kappa$ and adhesion $<\kappa$,*
(v) *G has a well-ordered decomposition of width $<\kappa$ and adhesion $<\kappa$,*
(vi) *G has a scattered linear decomposition of width $<\kappa$ and adhesion $<\kappa$.*

8.5. *For all $\kappa \geqslant \aleph_0$, the following are equivalent*:
(i) *G does not topologically contain K_κ,*
(ii) *G has a dissection of width $<\kappa$,*
(iii) *G has a well-founded tree-decomposition of width $<\kappa$.*
If $\kappa > \aleph_0$ and is regular, then (i)–(iii) *are equivalent to*
(iv) *G has a well-founded tree-decomposition of width $<\kappa$ and height $<\kappa$,*
(v) *G has a linear decomposition of width $<\kappa$.*
If $\kappa = \aleph_0$ then (i)–(iii) *are equivalent to*
(vi) *G has a graph-theoretic tree-decomposition of width $<\aleph_0$.*

8.6. *A graph has no $K_{\aleph_0}$-minor if and only if it has a narrow graph-theoretic tree-decomposition (T, W), such that for each $t \in V(T)$ there is an integer k such that the torso at t has no K_k-minor.*

8.7. *For all $\kappa > \aleph_0$, G has no K_κ-minor if and only if G has a dissection of width $<\kappa$ and adhesion $<\kappa$.*

8.8. *The following are equivalent*:
(i) *G has no half-grid minor,*
(ii) *G has a dissection of width $<\aleph_0$ and adhesion $<\aleph_0$,*
(iii) *G has a narrow graph-theoretic tree-decomposition of width $<\aleph_0$.*

References

[1] D. Bienstock, N. Robertson, P.D. Seymour and R. Thomas, Quickly excluding a forest, J. Combin. Theory Ser. B. 52 (1991) 274–283.
[2] R. Diestel, The structure of TK_a-free graphs, submitted.
[3] R. Halin, Graphen ohne unendliche Wege, Math Nachr. 31 (1966) 111–123.
[4] R. Halin, Über die Maximalzahl fremder unendlicher Wege in Graphen, Math. Nachr. 30 (1965) 63–85.
[5] H.A. Jung, Zusammenzüge und Unterteilungen von Graphen, Math. Nachr. 35 (1967) 241–268.
[6] L.M. Kirousis and C.H. Papadimitriou, Interval graphs and searching, Discrete Math. 55 (1985) 181–184.
[7] I. Kříž and R. Thomas, Clique-sums, tree-decompositions and compactness, Discrete Math. 81 (1990) 177–185.
[8] I. Kříž and R. Thomas, The Menger-like property of tree-width of infinite graphs and related compactness results, J. Combin. Theory, Ser. B. 52 (1991) 86–91.

[9] A. La Paugh, Recontamination does not help to search a graph, 1982, manuscript.
[10] N. Robertson and P.D. Seymour, Graph minors I. Excluding a forest, J. Combin. Theory Ser. B. 35 (1983) 39–61.
[11] N. Robertson and P.D. Seymour, Graph minors V. Excluding a planar graph, J. Combin. Theory Ser. B 41 (1986) 92–114.
[12] N. Robertson and P.D. Seymour, Graph minors XV. Wagner's conjecture, manuscript.
[13] N. Robertson and P.D. Seymour, Graph minors XVI. Well-quasi-ordering on a surface, manuscript.
[14] N. Robertson, P.D. Seymour and R. Thomas, Excluding infinite clique minors, manuscript.
[15] N. Robertson, P.D. Seymour and R. Thomas, Excluding infinite clique subdivisions, Trans. Amer. Math. Soc., to appear.
[16] N. Robertson, P.D. Seymour and R. Thomas, Quickly excluding a planar graph, submitted.
[17] P.D. Seymour and R. Thomas, Excluding infinite trees, Trans. Amer. Math. Soc., to appear.
[18] P.D. Seymour and R. Thomas, Graph searching, and a minimax theorem for tree-width, submitted.
[19] R. Thomas, A counterexample to Wagner's conjecture for infinite graphs, Math. Proc. Cambridge Philos. Soc. 103 (1988) 55–57.
[20] C. Thomassen, Configurations in graphs of large minimum degree, connectivity or chromatic number, in: G.S. Bloom, R.L. Graham and J. Malkevitch, eds., Combinatorial Mathematics, Proc. 3rd International Conference 1985, Ann. New York Acad. Sci., Vol. 555 (New York Acad. Sci., New York, 1989) 402–411.
[21] D. Seese, Ordered tree representations of infinite graphs, 1988, preprint.

Discrete Mathematics 95 (1991) 321–330
North-Holland

An end-faithful spanning tree counterexample*

Paul Seymour
Bellcore, 445 South St., Morristown, NJ 07960, USA

Robin Thomas
DIMACS Center, Hill Center, Busch Campus, Rutgers University, New Brunswick, NJ 08903, USA
School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

Reprinted from the Proceedings of the AMS

Abstract

Seymour, P. and R. Thomas, An end-faithful spanning tree counterexample, Discrete Mathematics 95 (1991) 321–330.

We find an infinitely-connected graph in which every spanning tree has a 2-way infinite path. This disproves Halin's well-known 'end-faithful spanning tree' conjecture, and also disproves a recent conjecture of Širáň.

1. Introduction

A *ray* in a graph G is a 1-way infinite path. (Graphs in this paper may be infinite, and may have loops or multiple edges.) Two rays R_1, R_2 in G are *parallel* if for every finite $X \subseteq V(G)$, the unique component of $G \backslash X$ which has infinite intersection with R_1 also has infinite intersection with R_2. ($G \backslash X$ is the graph obtained from G by deleting X.) Parallelness is an equivalence relation, and its equivalence classes are called the *ends* of G. These were first investigated by Halin [4], who proposed the following, which is called the 'end-faithful spanning tree conjecture', and which we shall disprove.

Conjecture 1.1. In every connected graph G there is a spanning tree T such that each end of G includes a unique end of T.

Halin [3–4] proved that Conjecture 1.1 holds if G is countable, and also that it holds if G does not contain $K_{\aleph_0}$. (We denote by K_κ the complete graph with κ

* Reprinted from the Proceedings of the AMS, by permission of the American Mathematical Society.

Elsevier Science Publishers B.V.

vertices, when κ is a cardinal. A graph G *contains* a graph H if some subgraph of G is isomorphic to a subdivision of H, that is, a graph obtained from H by replacing its edges by internally disjoint paths.) However, we shall see that Conjecture 1.1 is false in general. A counterexample has independently been obtained by Thomassen [9].

Let us say that G is *infinitely connected* if $V(G)$ is infinite and $G \backslash X$ is connected for every finite $X \subseteq V(G)$. Since an infinitely connected graph has a unique end, a consequence of Conjecture 1.1 would be the following.

Conjecture 1.2. In every infinitely connected graph there is a spanning tree with a unique end.

We shall give a counterexample to Conjecture 1.2 and hence to Conjecture 1.1.

The following conjecture was proposed at a recent (1989) conference in Cambridge by Širáň. (A tree is *rayless* if it has no ray.)

Conjecture 1.3. Let G be a connected graph, and suppose that for every ray R there is a vertex v such that for every finite $X \subseteq V(G) - \{v\}$, the component of $G \backslash X$ containing v has infinite intersection with R. Then G has a rayless spanning tree.

Širáň [8] proved this for countable graphs G, but we shall see that it is false in general. Since every infinitely connected graph satisfies the hypothesis of Conjecture 1.3, a consequence of Conjecture 1.3 would be the following.

Conjecture 1.4. In every infinitely connected graph there is a rayless spanning tree.

Our counterexample to Conjectures 1.1 and 1.2 is also a counterexample to Conjecture 1.4 and hence to Conjecture 1.3. (Indeed, we shall show that a graph satisfies Conjecture 1.2 if and only if it satisfies Conjecture 1.4.)

Let $T_{\aleph_1}$ be the tree in which every vertex has valency $\aleph_1$. We shall show the following.

1.5. *There is an infinitely connected graph G with $|V(G)| = 2^{\aleph_0}$, such that every spanning tree contains $T_{\aleph_1}$.*

In particular, every spanning tree has $\geqslant \aleph_1$ ends, contrary to Conjectures 1.2 and 1.4. We shall also show the following.

1.6. *There is an infinitely connected graph G with $|V(G)| = 2^{\aleph_0}$ which does not contain $K_{\aleph_1}$, such that every spanning tree contains $T_{\aleph_1}$.*

We do not know whether 1.5 is true with $2^{\aleph_0}$ replaced by $\aleph_1$. However, we shall see that the existence of G as in 1.6 with $|V(G)| = \aleph_1$ rather than $2^{\aleph_0}$ is independent of ZFC. Finally, we shall show the following.

1.7. *Every infinitely connected graph which does not contain $T_{\aleph_1}$ has a rayless spanning tree.*

Before we begin the main proofs, let us see the equivalence of Conjectures 1.2 and 1.4.

1.8. *Let G be an infinitely connected graph. Then G has a rayless spanning tree if and only if it has a spanning tree with exactly one end.*

Proof. Let T be a rayless spanning tree, and let R be a ray. Extend R to a spanning tree T' of $R \cup T$; then T' has exactly one end (for any ray not parallel to R includes a ray disjoint from R, and so contains a ray in T). For the converse, let T be a spanning tree with only one end, and let R be a ray of T, with $V(R) \neq V(G)$. Choose $v \in V(G) - V(R)$. Since G is infinitely connected, there are infinitely many finite paths of G from v to $V(R)$, mutually disjoint except for v. Let the union of these paths be F, and extend F to a spanning tree T' of $F \cup T$. Suppose that R' is a ray of T'. Since F is connected and includes no ray, and $F \cup R'$ includes no circuit, it follows that R' includes a ray disjoint from F; and so we may assume that R' is disjoint from F. Hence $R' \subseteq T$, and so R, R' are parallel in T. Since T is a tree, it follows that $R'' = R \cap R'$ is a ray. But $R' \cap F$ is null, and $(V(R) - V(R'')) \cap V(F)$ is finite, and yet $V(R) \cap V(F)$ is infinite, a contradiction. Thus T' is rayless, as required. □

We state a stronger form of 1.4 for countable graphs, which we shall apply later. We omit the proof, which is easy. Let G be a graph, let $v \in V(G)$ and let $\{P_n\}_{n=1}^{\infty}$ be a collection of finite paths in G, each with at least one edge, with one endpoint v and otherwise disjoint. The tree $R = P_1 \cup P_2 \cup \cdots$ is called an *ω-star* in G with *centre* v.

1.9. *Every countable, infinitely connected graph has a spanning ω-star with centre any specified vertex.*

2. The counterexample

A *hypergraph* is a pair (V, M), where V is a set and M is a set of subsets of V. Two hypergraphs (V, M), (V', M') are *isomorphic* if there is a bijection $\alpha: V \to V'$ mapping M to M' (and α is an *isomorphism*).

Let (V, M) be a hypergraph such that:

(i) every member of M has cardinality $\aleph_0$,

(ii) for every partition $(X_1, X_2, \ldots)$ of V into countably many sets, some X_i includes some member of M,

(iii) $|V|, |M| \geq \aleph_1$.

(In fact condition (iii) is implied by (i) and (ii), as the reader may verify). For example, we could take $V = \aleph_1$ and let M be the set of all countable subsets of V. We shall specify (V, M) later.

Let Σ be the set of all finite sequences of members of M. If $\sigma = (\mu_1, \ldots, \mu_k) \in \Sigma$ and $\mu \in M$, we denote the sequence $(\mu_1, \ldots, \mu_k, \mu) \in \Sigma$ by $\sigma + (\mu)$. For each $\sigma \in \Sigma$, let (V_σ, M_σ) be a hypergraph isomorphic to (V, M), such that $V_\sigma \cap V_{\sigma'} = \emptyset$ for all distinct $\sigma, \sigma' \in \Sigma$, and let $\alpha_\sigma : V \to V_\sigma$ be an isomorphism. For each $\mu \in M$, we denote $\{\alpha_\sigma(v) : v \in \mu\}$ by μ_σ; thus, $\mu_\sigma \in M_\sigma$.

Let G be the graph with $V(G) = \bigcup (V_\sigma : \sigma \in \Sigma)$, in which $u, v \in V(G)$ are adjacent if $u \in \mu_\sigma$ and $v \in V_{\sigma'}$, for some $\sigma \in \Sigma$ and $\mu \in M$, where $\sigma' = \sigma + (\mu)$. We shall show that every spanning tree of G contains $T_{\aleph_1}$.

We shall need the following lemma, which is very similar to a result of Laver [5] and which can be proved similarly (we omit the proof).

2.1. *Let T be a tree which does not contain $T_{\aleph_1}$, and let $v_0 \in V(T)$. There is a function ϕ assigning an ordinal $\phi(v)$ to each $v \in V(T)$, such that:*

(i) *if $u, v \in V(T)$ are adjacent and u lies on the path of T between v_0 and v, then $\phi(v) \leq \phi(u)$,*

(ii) *for all $u \in V(T)$ the set of all $v \in V(T)$ as in* (i) *with $\phi(v) = \phi(u)$ has cardinality $\leq \aleph_0$.*

2.2. *With G as defined earlier, every spanning tree of G contains $T_{\aleph_1}$.*

Proof. Suppose that T is a spanning tree of G not containing $T_{\aleph_1}$. Choose $v_0 \in V_{\sigma_0}$, where σ_0 is the null sequence, and for $u, v \in V(G)$, let us say that u is *before* v if u lies on the path of T between v_0 and v.

For each $\sigma \in \Sigma$, let B_σ denote $\bigcup V_{\sigma'}$, the union being taken over all $\sigma' \in \Sigma$ of which σ is an initial subsequence. Let us say that $u \in V(G)$ *dominates* $\sigma \in \Sigma$ if u is before every $v \in B_\sigma$, and let us say that $u \in V(G)$ is *big* if it dominates some $\sigma \in \Sigma$. Thus, v_0 is big. Let ϕ be as in 2.1, and choose a big vertex $u \in V(G)$ with $\phi(u)$ minimum.

(1) *There are only countably many $v \in V(G)$ such that u is before v and $\phi(u) \leq \phi(v)$.*

For $\phi(u) = \phi(v)$ for every such v. Let X be the set of all such v, and let R be the minimal subtree of T with $X \subseteq V(R)$. Then $u \in V(R)$ and u lies before every other vertex of R. Since every vertex of R lies on a path between u and some vertex v of X, and the ϕ-values on such a path do not increase (by 2.1(i)), and $\phi(v) = \phi(u)$, it follows that every vertex of this path belongs to X, and in

particular $V(R) = X$. Thus every vertex of R has valency $\leqslant \aleph_0$ (by 2.1(ii)), and so $|V(R)| \leqslant \aleph_0$. This proves (1).

Since u dominates some $\sigma \in \Sigma$ and hence dominates all extensions of σ in Σ, we may choose $\sigma \in \Sigma$ such that u dominates σ and $u \notin B_\sigma$. Since M is uncountable, there are uncountably many 1-term extensions σ' of σ, and the corresponding sets $B_{\sigma'}$ are mutually disjoint. Thus by (1) we may choose $\sigma \in \Sigma$ such that in addition there is no $v \in B_\sigma$ with $\phi(u) \leqslant \phi(v)$. Choose $\mu \in M$, and let $\sigma' = \sigma + (\mu)$.

Let S be the minimal subtree of T with $\mu_\sigma \subseteq V(S)$. Since μ_σ is countable it follows that so is $V(S)$. For each $s \in V(S)$, let X_s be the set of all $v \in V_{\sigma'}$ such that there is a path of T between s and v with no vertex in $V(S)$ except s. Thus $(X_s : s \in V(S))$ is a partition of $V_{\sigma'}$ into countably many sets, and so there exists $s \in V(S)$ and $\mu' \in M$ such that $\mu'_{\sigma'} \subseteq X_s$. Let $\sigma'' = \sigma' + (\mu')$. We claim that

(2) *s dominates σ''.*

For let $v \in B_{\sigma''}$, and let P be the path of T between v_0 and v. Since $v_0 \in V_{\sigma_0}$, it follows that $V(P) \cap \mu'_{\sigma'} \neq \emptyset$. Let $x \in V(P) \cap \mu'_{\sigma'}$. Since $V(P) \cap \mu_\sigma \neq \emptyset$ and hence $V(P) \cap V(S) \neq \emptyset$, it follows that P includes the unique minimal path of T between x and $V(S)$. Since $x \in \mu'_{\sigma'} \subseteq X_s$ it follows that $s \in V(P)$. Thus s is before v, as required.

(3) $s \in B_\sigma$.

For since X_s is infinite and $X_s \cap V(S) \subseteq \{s\}$, there exists $v \in X_s - V(S)$. Thus $v \in V_{\sigma'}$. Let P be the path of T between v and s. Since no vertex of P is in $V(S)$ except s (because $v \in X_s$) and $\mu_\sigma \subseteq V(S)$, it follows that $V(P) \cap \mu_\sigma \subseteq \{s\}$, and so either $s \in B_{\sigma'}$ or $s \in \mu_\sigma$. In either case, $s \in B_\sigma$ as required.

Now we chose σ such that there is no $v \in B_\sigma$ with $\phi(u) \leqslant \phi(v)$, and so from (3) we deduce that $\phi(s) < \phi(u)$. But s is big by (2), contrary to the choice of u. This completes the proof. □

So far we have not specified the collection M of sets used in the construction of the graph G above. As we saw before, we can take $V = \aleph_1$ and M to be the collection of all countable subsets of V, and the graph G we construct satisfies 1.5. This proves 1.5. However, that graph contains $K_{\aleph_1}$; and in view of Halin's theorem that every counterexample to Conjecture 1.1 and 1.2 contains $K_{\aleph_0}$, it is natural to ask if every counterexample also contains $K_{\aleph_1}$. The answer is no, as we shall see by a more complicated choice of M.

A *well-founded tree* is a poset $T = (V(T), \leqslant)$, such that for every pair $t, t' \in V$ their infimum $inf(t, t')$ exists, and for every $t \in V(T)$ the set $\{t' \in V(T) : t' \leqslant t\}$ is well-ordered by $\leqslant$. It follows that every well-founded tree T has a minimum element, called the *root* of T and denoted by *root* (T). Let V be the set of all transfinite sequences of distinct positive integers, and let us say for such sequences s_1, s_2, that $s_1 \leqslant s_2$ if s_1 is an initial segment of s_2. It was shown in the Ph.D. Thesis of D. Kurepa (and later, independently, by R. Laver; see [1]) that

the well-founded tree $(V, \leqslant)$ cannot be partitioned in countably many antichains and $|V| = 2^{\aleph_0}$. Let M be the collection of all infinite chains of $(V, \leqslant)$; we claim that (V, M) satisfies the requirements at the start of this section. For let $(X_1, X_2, \ldots)$ be a partition of V; we must show that some X_i includes an infinite chain of $(V, \leqslant)$. Suppose not; then each X_i can be partitioned into countably many antichains, and hence so can V, a contradiction. Thus (V, M) satisfies our requirements, and so the corresponding graph G satisfies 2.2. Moreover, it follows from the results of [6] that G does not contain $K_{\aleph_1}$. This proves 1.6. □

3. An independence result

We have seen that Conjectures 1.2 and 1.4 are:

(i) true for all countable graphs and for all graphs which do not contain $K_{\aleph_0}$,

(ii) false for a graph with $2^{\aleph_0}$ vertices which does not contain $K_{\aleph_1}$.

What about graphs with $\aleph_1$ vertices? We do not know whether such a graph can satisfy 1.5, but for 1.6 we have an independence result. Let us consider the truth of the following statement.

3.1. In every infinitely connected graph G with $|V(G)| = \aleph_1$ which does not contain $K_{\aleph_1}$, there is a rayless spanning tree.

We shall see that 3.1 is independent of ZFC (Zermelo-Fraenkel set theory together with the axiom of choice). For we observe from 1.6 the following.

3.2. *If the continuum hypothesis holds then* (3.1) *is false.*

On the other hand, Baumgartner, Malitz and Reinhardt [2] proved that the following statement is consistent with ZFC (although not with the continuum hypothesis).

3.3. If $(V, \leqslant)$ is a well-founded tree with $|V| \leqslant \aleph_1$, and every chain of $(V, \leqslant)$ has order type $<\omega_1$, then V may be partitioned into countably many antichains.

In the rest of this section we shall prove that 3.3 implies 3.1. If $T = (V(T), \leqslant)$ is a well-founded tree and $t_1, t_2 \in V(T)$, we say that $t \in V(T)$ is *between* t_1 and t_2 if $inf(t_1, t_2) \leqslant t$, and either $t \leqslant t_1$, or $t \leqslant t_2$. We say that t_1 is a *predecessor* of t_2 if $t_1 \leqslant t_2$ and there is no $t \in V(T) - \{t_1, t_2\}$ such that $t_1 \leqslant t \leqslant t_2$. A *well-founded tree-decomposition* of a graph G is a pair (T, W), where T is a well-founded tree and $W = (W_t : t \in V(T))$ is a collection of sets such that:

(W1) $\bigcup (W_t : t \in V(T)) = V(G)$ and every edge of G has both its ends in some W_t,

(W2) if t' is between t and t'' in T, then $W_t \cap W_{t''} \subseteq W_{t'}$,

(W3) if t has no predecessor then $W_t \supseteq \bigcup_{t'<t} \bigcap_{t' \leqslant t'' < t} W_{t''}$.

We need the following structure theorem, theorem (2.7) of [6].

3.4. *Let G be an infinitely connected graph with $|V(G)| \leqslant \aleph_1$ which does not contain $K_{\aleph_1}$. Then there exists a well-founded tree-decomposition (T, W) of G such that:*

(i) $|V(T)| \leqslant \aleph_1$,

(ii) *every chain of T has order type $< \omega_1$, and*

(iii) *W_t induces a countable infinitely connected graph in G for every $t \in V(T)$.*

3.5. *If* 3.3 *holds then* 3.1 *holds.*

Proof. Let G be an infinitely connected graph with $|V(G)| = \aleph_1$ which does not contain $K_{\aleph_1}$, and let (T, W) be a well-founded tree-decomposition of G as in 3.4.

(1) *For every $v \in V(G)$ there exists a unique minimal element $t \in V(T)$ with $v \in W_t$.*

For there exists at least one such t by (W1), and the minimal one is unique by the existence of infima and (W2).

For $v \in V(G)$, the element $t \in V(T)$ as in (1) will be denoted by $t(v)$. From (i) and (ii) of 3.4 and 3.3, $V(T)$ can be partitioned into countably many antichains, say $A_1, A_2, \ldots$.

Let t_0 be the root of T. An *ideal* of T is a subset $S \subseteq V(T)$ such that $t_0 \in S$ and such that $s \in S$ for every $s \in V(T)$ with $s \leqslant t$ for some $t \in S$. Choose $v_0 \in W_{t_0}$. A *sprout* is a triple (S, R, ρ), where S is an ideal of T, R is a tree of G with $\bigcup_{s \in S} W_s = V(R)$, and ρ is a function from $V(R)$ into $\{1, 2, \ldots\}$ such that:

(2) if $v \neq v_0$ is between v_0 and v' in R then $\rho(v) \geqslant \rho(v')$, and the inequality is strict unless $t(v) = t(v')$,

(3) the set $\{v \in V(R) : t(v) = t \text{ and } \rho(v) \leqslant i\}$ is finite for every $t \in V(T)$ and every $i \in \{1, 2, \ldots\}$,

(4) $\rho(v) \geqslant i$ for every $v \in V(R)$ with $t(v) \in A_i$.

We first prove the following.

(5) *If (S, R, ρ) is a sprout, $t \in V(T)$ and $i \geqslant 1$ is an integer, then $\{v \in W_t \cap V(R) : \rho(v) \leqslant i\}$ is finite.*

We prove (5) by transfinite induction. The statement follows from (3) if $t = t_0$, so let $t \in V(T) - \{t_0\}$, let $i \geqslant 1$ be an integer and assume that (5) holds for all $t' < t$. There exists $t_1 \in V(T)$ with $t_1 < t$ such that $t' \notin A_1 \cup \cdots \cup A_i$ for every $t' \in V(T)$ with $t_1 < t' < t$. Hence

$$\{v \in W_t \cap V(R) : \rho(v) \leqslant i\} \subseteq \{v \in W_t \cap V(R) : \rho(v) \leqslant i,\ t(v) \leqslant t_1\} \cup \{v \in W_t \cap V(R) : \rho(v) \leqslant i,\ t(v) = t\}$$

by (4) and (W2). The first set is a subset of $\{v \in W_{t_1} \cap V(R) : \rho(v) \leqslant i\}$, and hence is finite by the induction hypothesis, and the second one is finite by (3). This proves (5).

(6) *There exists at least one sprout.*

For let $S = \{t_0\}$, and let $R = P_1 \cup P_2 \cup \cdots$ be the spanning ω-star of the subgraph of G induced by W_{t_0} with centre v_0, which exists by 1.9, and (iii) of 3.4.

Let $\rho: V(R) \to \{1, 2, \ldots\}$ be such that if $\rho(v) = i$ then $v \in V(P_i)$. Then (S, R, ρ) is a sprout, as desired.

We order sprouts by saying that $(S, R, \rho) \leqslant (S', R', \rho')$ if $S \subseteq S'$, R is a subtree of R' and $\rho(r) = \rho'(r)$ for every $r \in V(R)$. By Zorn's lemma there exists a maximal sprout (S, R, ρ).

(7) $S = V(T)$.

For suppose not; then there exists $t \in V(T) - S$, minimal. Let k be such that $t \in A_k$, let $F = \bigcup_{s \in S} W_s$, and let $u_k, u_{k+1}, \ldots$ be all the vertices of $W_t - F$. Since $W_t \cap V(R) = W_t \cap F$ and since $W_t \cap F$ is infinite (because $W_t \cap F$ separates the infinite sets W_t and W_{t_0} and G is infinitely connected) we deduce by repeated application of (5) and (iii) of 3.4 that there exist mutually disjoint finite paths $P_k, P_{k+1}, \ldots$ such that for every $i \geqslant k$, $V(P_i) \subseteq W_t$, P_i has one end some $w \in W_t \cap F$ with $\rho(w) > i$, and has no other vertex in F, and $u_i \in \bigcup_{k \leqslant i' \leqslant i} V(P_{i'})$. In particular,

$$W_t - F \subseteq \bigcup_{i \geqslant k} V(P_i) \subseteq W_t.$$

Let $S' = S \cup \{t\}$, let $R' = R \cup P_k \cup P_{k+1} \cup \cdots$, and let ρ be defined by

$$\rho'(r) = \begin{cases} \rho(r) & \text{if } r \in V(R), \\ i & \text{if } r \in V(P_i) - V(R). \end{cases}$$

Then (S', R', ρ') is a sprout, contradicting the maximality of (R, ρ, t). This proves (7).

From (7) and (W1) we deduce that R is a spanning tree of G, and from (2) and (3) it follows that R is rayless. $\square$

We deduce that both 3.1 and its negation are relatively consistent with ZFC, and hence 3.1 is independent of ZFC.

4. Excluding the $\aleph_1$-tree

Our final objective is to prove 1.7, which we restate.

4.1. *Every infinitely connected graph which does not contain $T_{\aleph_1}$ has a rayless spanning tree.*

If $X \subseteq V(G)$, an *X-flap* is the vertex set of a component of $G \setminus X$. We shall need the following, a consequence of theorem (2.3) of [7].

4.2. *Let G be a graph which does not contain $T_{\aleph_1}$. For each $X \subseteq V(G)$ with $|X| \leqslant \aleph_0$, let $\beta(X)$ be a union of X-flaps, such that if $X \subseteq Y \subseteq V(G)$ and $|Y| \leqslant \aleph_0$ then $\beta(X)$ includes precisely those X-flaps which meet $\beta(Y)$. Then $\beta(\emptyset) = \emptyset$.*

Proof of 4.1. Let us say that $X \subseteq V(G)$ is *good* if $X \subseteq V(T)$ for some rayless tree T of G (not necessarily a spanning tree).

(1) *Every countable subset of $V(G)$ is good.*

For let $X \subseteq V(G)$ be countable. Then $X \subseteq V(H)$ for some countable subgraph H of G which is infinitely connected, and by 1.9 H has a spanning ω-star. Hence X is good.

(2) *If $X \subseteq V(G)$ is good and $C_i (i \in I)$ are good X-flaps, then $X \cup \bigcup (C_i : i \in I)$ is good.*

For let T be a rayless tree of G with $X \subseteq V(T)$, and for each $i \in I$ let T_i be a rayless tree with $C_i \subseteq V(T_i)$. Since G is connected, we may assume that each T_i intersects T. For each $i \in I$, choose $S_i \subseteq T_i$ minimal such that $S_i \cup T$ is connected and $C_i \subseteq V(S_i \cup T)$. Then each component of S_i has exactly one vertex in $V(T)$, and $V(S_i \cup T) = C_i \cup V(T)$. Thus, $T \cup \bigcup (S_i : i \in I)$ is a rayless tree of G and its vertex set includes $X \cup \bigcup (C_i : i \in I)$, as required.

For each $X \subseteq V(G)$ with $|X| \leqslant \aleph_0$, let $\beta(X)$ be the union of all X-flaps which are not good.

(3) *If $X \subseteq Y \subseteq V(G)$ and $|Y| \leqslant \aleph_0$, then $\beta(X)$ includes just those X-flaps which meet $\beta(Y)$.*

For if an X-flap C meets a Y-flap $D \subseteq \beta(Y)$, then $D \subseteq C$; and so C is not good, since D is not good, and hence $C \subseteq \beta(X)$. Conversely, let C be an X-flap with $C \cap \beta(Y) = \emptyset$. Let the Y-flaps included in C be $C_i (i \in I)$. Since $C \cap B(Y) = \emptyset$, each C_i is good, and so by (1) and (2) $Y \cup \bigcup (C_i : i \in I)$ is good. Hence C is good, since $C \subseteq Y \cup \bigcup (C_i : i \in I)$, and so $C \notin \beta(X)$. This proves (3).

From (3) and 4.2, we deduce that $\beta(\emptyset) = \emptyset$. Hence G has a rayless spanning tree, as required. □

Acknowledgement

We are grateful to R. Diestel for pointing out an error in an earlier version of this paper.

References

[1] J.E. Baumgartner, Results and independence proofs in combinatorial set theory, Ph.D. Thesis, University of California, 1970.

[2] J.E. Baumgartner, J. Malitz and W. Reinhardt, Embedding trees in the rationals, Proc. Nat. Acad. Sci. 67 (1970) 1748–1753.

[3] R. Halin, Simplicial decompositions of infinite graphs, in: B. Bollobás, ed., Advances in Graph Theory, Ann. Discrete Math. 3 (North-Holland, Amsterdam, 1978).

[4] R. Halin, Über unendliche Wege in Graphen, Math. Ann. 157 (1964) 125–137.

[5] R. Laver, Better-quasi-ordering and a class of trees, in: Studies in Foundations and Combinatorics, Advances in Mathematics Supplementary Studies, Vol. 1.

[6] N. Robertson, P.D. Seymour and R. Thomas, Excluding subdivisions of infinite cliques, Trans. Amer. Math. Soc., to appear.
[7] P.D. Seymour and R. Thomas, Excluding infinite trees, Trans. Amer. Math. Soc., to appear.
[8] J. Širáň, Coterminal forests and spanning trees in infinite graphs, Proceedings of the Cambridge conference on infinite graphs, 1989, submitted.
[9] C. Thomassen, Infinite connected graphs with no end-preserving spanning trees, 1990, manuscript.

Discrete Mathematics 95 (1991) 331–340
North-Holland

End-faithful forests and spanning trees in infinite graphs

Jozef Širáň

Department of Mathematics, Faculty of Civil Engineering, Slovak Technical University, Radlinského 11, 813 68 Bratislava, Czechoslovakia

Received 12 October 1989
Revised 23 April 1990

Abstract

Širáň, J., End-faithful forests and spanning trees in infinite graphs, Discrete Mathematics 95 (1991) 331–340.

We present several sufficient conditions for an end-faithful forest to be extendable to an end-faithful spanning tree in an infinite graph. Results related to rayless spanning trees are also included. Finally, a common generalization of end-faithfulness and raylessness is discussed.

1. Introduction

In 1964, Halin [3] introduced the concept of an *end* of a graph. In the same paper, he proved that any connected countable graph contains an *end-faithful spanning tree* (i.e., a spanning tree which represents the structure of ends in the graph). One of the central questions about end-faithfulness—the problem of whether the cardinality restriction in Halin's result may be dropped—has very recently been answered in the negative by Seymour and Thomas [8]. On the other hand, numerous positive results on end-faithful spanning trees have been obtained by several authors (e.g. Halin [4], Jung [5], Polat [7], Diestel [2], Seymour and Thomas [8], etc.).

As pointed out by Polat [6–7], there is an intimate connection between end-faithful spanning trees and *rayless* spanning trees (those without infinite paths; see also [8]). We shall be interested here in another kind of relation between these two types of spanning trees, which arises from contracting end-faithful forests in graphs.

After introducing the necessary concepts, in Section 3 we characterize the countable graphs containing a rayless spanning tree. Although this result is implicitly contained in [6], we provide it with a new short proof using a method which is, in a way, dual to that used in [1]. Section 4 is devoted to the study of graphs that arise by contracting the components of end-faithful forests. Such

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

graphs are shown to be closely related to the rayless spanning tree problem as well as to the problem of extending an end-faithful forest to an end-faithful spanning tree. Finally, we propose a generalization of the approach used in Section 4, this time considering forests which are 'faithful' with respect to a chosen subset of ends, and 'rayless' with respect to the remaining ends of a graph.

2. Terminology

As usual, for a finite or infinite graph H we denote by $V(H)$ and $E(H)$ the set of vertices and edges of H, respectively. A finite path with end-vertices u and v in H will be called a *u–v path*. A u–v path P will be said to be *disjoint* from a set $W \subset V(H)$ if $V(P) \cap W = \emptyset$, and *internally disjoint* from W if $(V(P) - \{u, v\}) \cap W = \emptyset$. Note that, although only undirected graphs will be considered, it will sometimes be convenient to speak about u and v as the initial and terminal vertex of P, respectively (or to say that P emanates from u and terminates at v).

Let J and K be subgraphs of H. A u–v path Q will be called a J–K path if $u \in V(J)$, $v \in V(K)$, and Q is internally disjoint from both $V(J)$ and $V(K)$. In the case when $J = \{u\}$ we simply speak of a u–K path.

A one-way infinite path in H will be referred to as a *ray* in H. (If a graph contains no rays, it will be called *rayless*.) If $P = v_0 v_1 \cdots v_n v_{n+1} \cdots$ is a ray with initial vertex v_0, then any sub-ray $v_{n+1} v_{n+2} \cdots$ of P will be called a *tail* of P. The ray P will be said to be *dominated* by a vertex $u \in V(H)$ if there are infinitely many u–P paths in H no two of which have a common vertex except u. (Such a vertex u was called 'P-kritisch' in [5], and a 'neighbour of P' in [7].) Equivalently, u dominates P if there is no finite subset $W \subset V(H)$ such that u and a tail of P would lie in different components of the graph H-W; we say briefly that u and P *cannot be separated by a finite set of vertices*.

Two rays P and Q in H will be said to be *equivalent* (denoted by $P \sim_H Q$) if there is a third ray R in H which intersects both P and Q infinitely often. As was shown in [3], $P \sim_H Q$ if and only if there is an infinite set of mutually disjoint P-Q paths (some of which may be trivial one-vertex paths) in H. This is further equivalent to saying that the tails of P and Q cannot be separated by a finite set of vertices of H. It is easy to see that $\sim_H$ is an equivalence relation on the set of all rays in H [3]. An equivalence class under $\sim_H$ is called an *end* of H.

Note that if a ray R in H is dominated by a vertex u, then every ray equivalent to R is dominated by u as well. In this sense we often speak about an end being dominated by a vertex.

Ends of graphs have been studied from many points of view. We shall mainly be interested in forests (and sometimes in spanning trees) which reflect the structure of the ends of a graph. To be more precise, a forest F in a graph H will be said to be *end-faithful* ('coterminal' in [7]) if, for each end ε_H in H, there exists exactly one end ε_F in F such that $\varepsilon_F \subset \varepsilon_H$. Roughly speaking, a forest F is

end-faithful in H if each end of H is represented by a unique (up to an initial segment) ray in F.

Finally, if K is a subgraph of H, then the so-called *contracted graph* H/K is obtained by collapsing every component of K to a single vertex and removing all the loops and parallel edges that arose in the process of contraction. More formally, if H-K denotes the subgraph of H induced by the set $V(H)-V(K)$ and $C(K)$ is the set of components of K, then

$$V(H/K)=V(H-K)\cup C(K) \quad \text{and} \quad E(H/K)=E(H-K)\cup\{uB;u\in V(H-K)$$

and $B\in C(K)$ with a vertex $v\in B$ such that $uv\in E(H)\}\cup\{BD;B,D\in C(K)$ such that there are vertices $u\in B$ and $v\in D$ for which $uv\in E(H)\}$. In other words, the contracted graph H/K is the image of a *contraction map* $c:H\to H/K$, which is a graph homomorphism mapping each vertex of a component of K to that component, and is the identity on $H-K$. Observe that the contraction map $c:H\to H/K$ induces a similar map $T\to T/K$ whenever $K\subset T\subset H$, so we may refer to the contracted graph T/K as $c(T)$.

3. Ray domination and rayless spanning trees

As we shall see in Section 4, there is an interesting connection between rayless spanning trees and the extendability of end-faithful forests to end-faithful spanning trees in infinite graphs. In this section we shall concentrate on the implications of ray domination for the existence of rayless spanning trees, and vice versa. The main result, although implicitly contained in [6], is presented here with a new and short proof. Moreover, our method is in a natural way dual to that used in [1] for proving the existence of special end-faithful trees in connected countable graphs. We thus have one more reason to consider raylessness and end-faithfulness as dual properties (see also [7]).

Lemma 1. *Let G be a graph containing a rayless spanning tree. Then every ray in G is dominated.*

Proof. Let T be a rayless spanning tree in G and let P be a ray in G. Pick a vertex v in G and consider the subtree T_v of T consisting of all the paths in T emanating from v and terminating at a vertex of $V(P)-\{v\}$; note that we do not require these paths to be internally disjoint from P. Since for every vertex $u\in V(P)$ there is a unique u–v path in T (possibly the trivial one), every vertex of P belongs to T_v. So, T_v is an infinite subtree of T. The fact that T has been assumed to be rayless implies that T_v must be rayless as well. By König's *Unendlichkeitslemma* (see e.g. [10]) there exists a vertex $w\in V(T_v)$ which has infinite degree in T_v. According to the definition of the tree T_v, every edge $e\in E(T_v)$ incident with w (except, possibly, the one contained in the single w-v

path in T_v) belongs to a $w-P$ path $Q_e \subset T_v$. Clearly, any two Q_e's have only the vertex w in common. Therefore, the infinite set of w–P paths Q_e shows that P is dominated in G by the vertex w. □

Note that, in Lemma 1, no cardinality restriction has been made on the number of vertices of G. The following theorem of [6–7] asserts that the converse of Lemma 1 holds true for countable graphs.

Theorem 2. *Let G be a connected countable graph in which every ray is dominated. Then G contains a rayless spanning tree.*

Proof. Let $V(G) = \{v_0, v_1, v_2, \ldots\}$. Using induction, we shall first construct a nested sequence of trees $T_0 \subset T_1 \subset T_2 \cdots$ in G such that $v_n \in V(T_n)$ for every n. As the initial step, put $T_0 = v_0$. Now assume that the tree T_{n-1} has already been constructed. If $v_n \in V(T_{n-1})$, we simply set $T_n = T_{n-1}$. If $v_n \notin V(T_{n-1})$, let m be the smallest index i for which there exists a v_n-T_{n-1} path $v_n \cdots v_i$, $v_i \in V(T_{n-1})$; its existence is guaranteed by the connectedness of G. Adjoin one of such paths $Q_n = v_n \cdots v_m$ to T_{n-1} to obtain T_n, i.e., let $T_n = T_{n-1} \cup Q_n$.

Having thus finished our inductive construction, let us consider the subgraph

$$T = \bigcup_{n=0}^{\infty} T_n$$

of G. It is clear that T is a spanning tree of G. In what follows we shall prove that T is rayless.

Suppose the contrary, i.e. that T contains a ray P. According to the properties of G, P is dominated in G by a vertex v_k, say. Since only the tails of P will be of interest to us, we shall assume (without loss of generality) that P contains no vertex v_i such that $i \leqslant k$.

We begin with an important observation: For every n, the graph $T_n \cap P$ (if non-empty) is connected. To see this, suppose for a contradiction that $T_n \cap P$ has at least two components C and D. Then there is a C-D path $J \subset T_n$ as well as another C-D path $K \subset P$. Obviously $K \neq J$, since otherwise $K \cup C \cup D \subset T_n \cap P$ and C, D would not be distinct components of the graph $T_n \cap P$. From the facts that $T_n \subset T$ and $P \subset T$ we deduce that the union $C \cup D \cup K \cup J$ is a connected subgraph of T. However, it can readily be seen that the graph $C \cup D \cup K \cup J$ contains a cycle. Therefore, $T_n \cap P$ cannot have more than one component, as claimed.

Now let j be the smallest index for which $T_j \cap P \neq T_{j-1}P \neq \emptyset$; observe that necessarily $v_k \in V(T_{j-1})$. It follows that T_j must have been obtained from T_{j-1} by adding a v_j-T_{j-1} path $Q_j = v_j \cdots v_m$, $v_m \in V(T_{j-1})$. By the connectedness and inclusion of the paths $T_{j-1} \cap P \subset T_j \cap P$, the terminal vertex v_m of Q_j is easily seen to belong to P. Thus, according to the rules of our inductive construction,

$$m = \min M \quad \text{where } M = \{i;\ \text{there is a } v_j\text{-}T_{j-1} \text{ path } v_j \cdots v_i,\ v_i \in V(T_{j-1})\}.$$

Moreover, since v_k dominates P, one can readily find a v_j-v_k path Q_j' avoiding the (finite) set $V(T_{j-1}) - \{v_k\}$. Thus, $m = \min M \leqslant k$, which contradicts the fact that $m > k$ by our choice of P. This shows that our spanning tree T is rayless, as desired. □

Combining the result just proved with Lemma 1, we immediately obtain the following characterization of countable graphs containing a rayless spanning tree.

Corollary 3. *A connected countable graph G has a rayless spanning tree if and only if every ray in G is dominated.*

During the Infinite Graph Theory conference in Cambridge, 1989, I asked whether the above characterization extends to graphs with arbitrary cardinality. As was shown recently by Seymour and Thomas [8], the answer in general is no, so the problem has to be restated in broader terms: Which graphs among those in which every ray is dominated have rayless spanning trees?

4. Extending and contracting end-faithful forests

In this section we shall investigate the relationship between an end-faithful forest F in a graph G and the corresponding contracted graph G/F, obtained by shrinking every component of F into a single vertex (see Section 2). Our point of departure will be the following characterization theorem of those end-faithful forests that are extendable to end-faithful spanning trees.

Theorem 4. *Let F be an end-faithful forest in a connected graph G. Then F is contained in an end-faithful spanning tree of G if and only if the graph G/F contains a rayless spanning tree.*

Proof. Assume first that T is an end-faithful spanning tree of G such that $F \subset T$. Denote by $c: G \to G/F$ the contraction map and consider the subgraph $c(T)$ in G/F. Obviously, $c(T)$ is a spanning tree in G/F. Suppose that there is a ray Q in $c(T)$. Since T is a tree and F a subforest of T, there is a ray P in T such that $c(P) = Q$. From the fact that both F and T are end-faithful in G it follows that a tail of P is contained in a single component of F. However, this implies that $c(P)$ cannot be a ray in G/F, a contradiction. Thus, $c(T)$ is a rayless spanning tree in G/F.

Conversely, let T_0 be a rayless spanning tree in G/F. As before, let $c: G \to G/F$ be the contraction homomorphism. Let $U = \{c(D); D \text{ a component of } F\}$. For each edge $uv \in E(T_0)$ choose an edge $e_{uv} \in E(G)$ from the set $c^{-1}(uv)$. Obviously, if neither u nor v belongs to U then there is a single edge $e_{uv} = uv$ in $c^{-1}(uv)$. If $u = c(B) \in U$ and $v \neq U$, the edge e_{uv} joins the vertex v to a vertex in the

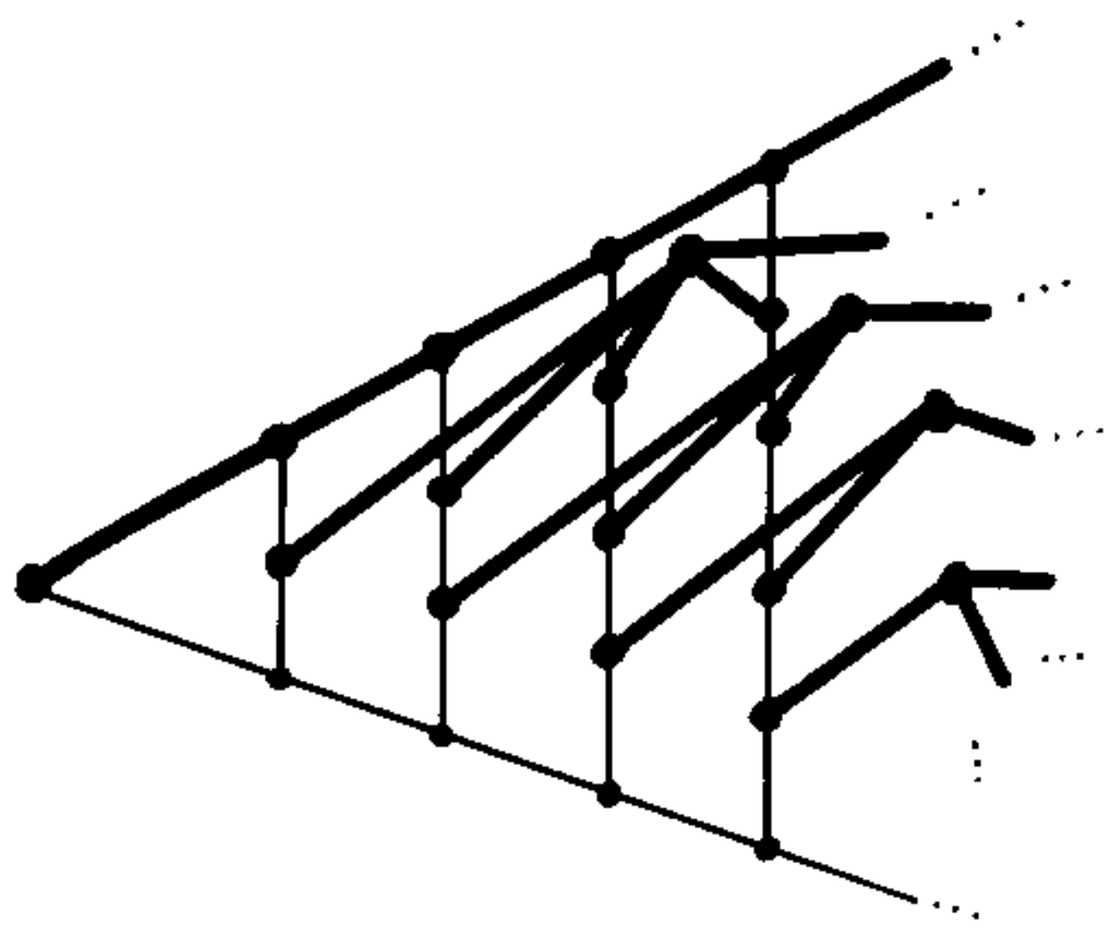


Fig. 1.

component B of F: if also $v = c(D) \in U$ then e_{uv} is one of the edges joining a vertex in B to a vertex in D.

Put $E' = \{e_{uv}; uv \in E(T_0)\}$, and consider the subgraph T of G generated by the edge set $E(F) \cup E'$. By our construction, T comprises every vertex of G. Moreover, since $c(T) = T_0$ is a tree and the contraction involves connected components of F only, the graph T is connected and cannot contain cycles. Thus, T is a spanning tree of G. (Roughly speaking, we have obtained T from T_0 by 'blowing up' every vertex $c(D) \in U$ to the original component D of F.) It remains to prove that T is end-faithful in G. If there were a ray $R \subset T$ which was not equivalent in T to any ray of F then, clearly, $c(R)$ would be a ray in $c(T) = T_0$ (note that $c(R)$ cannot be finite). This, however, would contradict the choice of T_0. As F ($\subset T$) represents every end of G, this completes the proof. □

It should be pointed out that, in general, an end-faithful forest need not extend to an end-faithful spanning tree. As an example, consider the graph G depicted in Fig. 1; the forest F is indicated by heavy lines. Note that, as asserted by Theorem 4, the graph G/F (Fig. 2) does not contain a rayless spanning tree.

We already saw in Section 3 that there is a relation between ray domination and the existence of rayless spanning trees. In the light of Theorems 2 and 4, it would be of interest to determine those end-faithful forests $F \subset G$ for which every ray in G/F is dominated. Towards this end we can offer only a sufficient condition. In order to state it, we shall introduce the concept of a ray-sensitive forest.

From now on, it will sometimes be useful to view any forest as a spanning forest. More explicitly, with each forest F in G we associated a new forest $\bar{F}$ by

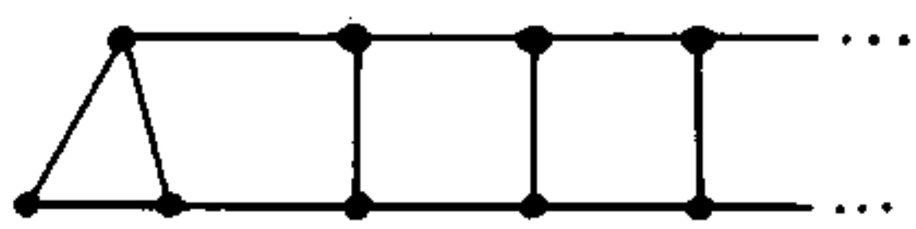


Fig. 2.

declaring every vertex not in F to be an isolated vertex of $\bar{F}$, i.e., setting $V(\bar{F}) = V(G)$ and $E(\bar{F}) = E(F)$.

Let F be an end-faithful forest in G. Then, for every ray P in G there exists a unique component C_P of F containing a ray P_F equivalent in G to P. The forest F will be called *ray-sensitive* if, for every ray P of G, either

(a) P intersects a component of F in an infinite number of vertices, or

(b) no finite set of components of $\bar{F}$ separates P from P_F (i.e., if S is any finite set of components of $\bar{F}$ other than C_P then there is a component of $G - \bigcup S$ which contains tails of both P and P_F).

Let us illustrate this concept by a simple observation.

Lemma 5. *Let F be a locally finite end-faithful forest in a graph G. Then F is ray-sensitive.*

Proof. Suppose the contrary, and let P be a ray in G which intersects no component of F in an infinite number of vertices. Let S be a finite set of components of $\bar{F}$ which separates P from P_F. Since $P \sim_G P_F$, the tails of P and P_F cannot be separated by a finite number of vertices of G. Consequently, if X is an infinite set of mutually disjoint P-P_F paths in G, then at least one component T of F, $T \in S$, intersects an infinite number of paths in X. Pick a vertex $v \in V(T)$ and let T_v be the subtree of T generated by all v-X paths. It follows that T_v is an infinite subtree of T; moreover, T_v is locally finite because of the assumed properties of F. Applying König's *Unendlichkeitslemma* we see that there is a ray R in $T_v \subset T$. But then, considering the paths in X which intersect T_v, it is easy to see that $R \sim_G P$. This, however, contradicts the fact that F is end-faithful in G, since obviously $T \neq C_P$. □

We are now ready to state and prove our sufficient condition for ray domination in contracted graphs.

Theorem 6. *Let F be a ray-sensitive end-faithful forest in a graph G. Then every ray in G/F is dominated.*

Proof. Let $c: G \to G/F$ be the contraction map and let P_0 be a ray in G/F. It is easy to see that there is a ray P in G such that $c(P) = P_0$. Obviously, P cannot intersect a component of F in an infinite number of vertices; otherwise $c(P)$ would not be a ray in G/F. As before, let C_P be the unique component of F which contains a ray equivalent (in G) to P. We claim that the vertex $u = c(C_P)$ dominates the ray $P_0 = c(P)$ in G/F. For suppose this is not the case. Then u can be separated from a tail of P_0 by a finite set S_0 of vertices of G/F. Considering the pre-images of these vertices under the map c, it follows that a tail of P can be separated from the component C_P by a finite number of components of $\bar{F}$,

namely, by the elements of the set $S = c^{-1}(S_0)$. But this contradicts the fact that F is ray-sensitive. □

The graph in Fig. 1 shows that if an end-faithful forest F in G is not ray-sensitive then G/F may be such that not every ray is dominated (and so, by Lemma 1, G/F has no rayless spanning tree, and therefore (Theorem 4) F cannot be extended to an end-faithful spanning tree of G). On the other hand, it is clear that the condition 'F is ray-sensitive' is not necessary for G/F to have every ray dominated.

Combining Lemma 5 with Theorem 6 we immediately obtain the following.

Corollary 7. *Let F be a locally finite end-faithful forest in a graph G. Then every ray in G/F is dominated.*

Applying now Theorem 2, the last result yields the following.

Corollary 8. *If F is a locally finite end-faithful forest in a connected countable graph G, then G/F contains a rayless spanning tree.*

By virtue of Theorem 4 we then obtain the following generalization of [9, Theorem 3].

Corollary 9. *Any locally finite end-faithful forest in a connected countable graph can be extended to an end-faithful spanning tree.*

Note that the assumption of local finiteness cannot be dropped from the last statement (see Figs. 1 and 2).

The following example (due to Diestel) shows that, in Theorem 6, the assertion that 'every ray in G/F is dominated' cannot be replaced with 'G/F has a rayless spanning tree' (so that the current version, although weaker than the potential alternative, is best possible, and it is indeed necessary for the proof of Corollaries 8–9 to invoke Theorem 2). Let H be the graph from [8] which shows that Theorem 2 does not extend to uncountable graphs; this graph H is infinitely connected (and thus every ray in it is dominated), but any spanning tree of H has uncountably many ends. Now let F be a ray disjoint from H, and let G be obtained from $H \cup F$ by inserting all edges between F and some fixed ray P in H. Clearly G and G/F are again infinitely connected, so F is an end-faithful and ray-sensitive forest in G (and every ray in G/F is dominated). However, G/F has no rayless spanning tree. For if T is such a tree and v is the vertex of G/F into which F was contracted, then $(T - v) \cup E(P)$ contains a spanning tree of H with at most one end, a contradiction.

We conclude our list of consequences with a result on extending end-faithful trees to end-faithful spanning trees [9, Theorem 2].

Corollary 10. *Let T be an arbitrary end-faithful tree in a connected countable graph G. Then T can be extended to an end-faithful spanning tree of G.*

Proof. By Theorems 6, 2 and 4, it is sufficient to notice that an end-faithful tree in G is necessarily ray-sensitive. Indeed, if P is a ray in G which intersects T in a finite number of vertices only, then a tail of P cannot be separated from a ray $R \subset T$ with $R \sim_G P$ by any finite number of vertices of G (i.e., single-vertex components of $\bar{T}$). □

5. Remarks on a common generalization of end-faithfulness and raylessness

The purpose of our last section is to introduce spanning trees which generalize both rayless and end-faithful ones. The proofs of the results can be obtained by appropriate modifications of the methods used in the previous section and are therefore left to the reader.

Let Ω be an arbitrary subset of the set of ends of a graph G. A forest $F \subset G$ will be said to be *Ω-faithful* if every ray of F belongs to an end in Ω, and for every end $\varepsilon \in \Omega$ there is precisely one end of F contained in ε. Thus an Ω-faithful forest in G is rayless if $\Omega = \emptyset$, and end-faithful in G if Ω comprises all ends of G.

Our result characterizing the end-faithful forests extendable to end-faithful spanning trees (Theorem 4) can now be stated in the following more general form.

Lemma 11. *Let Ω be an arbitrary collection of ends of a connected graph, G, and let F be an Ω-faithful forest in G. Then F can be extended to an Ω-faithful spanning tree of G if and only if the graph G/F contains a rayless spanning tree.*

In Section 4 we introduced the concept of a ray-sensitive forest. No doubt an analogous (but, unfortunately, more complicated) concept could be defined in order to generalize Theorem 6 to Ω-faithful forests. Instead of pursuing this idea, we present here a direct extension of Corollaries 7 and 9. Recall that an end is *dominated* if it contains a ray which is dominated.

Theorem 12. *Let Ω be a set of ends of a graph G, and let F be a locally finite Ω-faithful forest in G. Assume that every end of G which does not belong to Ω is dominated in G. Then every end of the graph G/F is dominated (in G/F).*

Combining the last result with Lemma 11 and Theorem 2 we immediately obtain the following.

Corollary 13. *Let Ω be a set of ends in a connected countable graph G such that every end not in Ω is dominated in G. Let F be an arbitrary locally finite Ω-faithful forest in G. Then F is contained in an Ω-faithful spanning tree of G.*

So far we have been dealing mostly with extending end-faithful or Ω-faithful forests to corresponding spanning trees. Let us conclude with a result refering directly to the existence of an Ω-faithful spanning tree.

Corollary 14. *Let Ω be a countable set of ends in a connected countable graph G. Assume that every end of G which does not belong to Ω is dominated in G. Then G contains an Ω-faithful spanning tree.*

Proof. Using induction it is easy to show that, since Ω is countable, G contains an Ω-faithful forest all of whose components are rays. Apply Corollary 13. □

In this context, it is natural to ask whether the last result extends to uncountable Ω. If so, it would be an interesting generalization of Halin's original result on end-faithful spanning trees in countable graphs [3] on the one hand, and of Theorem 2 on the other.

References

[1] R. Diestel, Strongly end-faithful spanning trees in countable graphs, 1988, preprint.
[2] R. Diestel, On end-faithful spanning trees in infinite graphs, Math. Proc. Cambridge Philos. Soc., in press.
[3] R. Halin, Über unendliche Wege in Graphen, Math. Ann. 157 (1964) 125–137.
[4] R. Halin, Simplicial decompositions of infinite graphs, in: B. Bollobás, ed., Advances in Graph Theory, Ann. Discrete Math. 3 (North Holland, Amsterdam, 1978).
[5] H.A. Jung, Wurzelbäume und unendliche Wege in Graphen, Math. Nachr. 41 (1969) 1–22.
[6] N. Polat, Développements terminaux des graphes infinis III. Arbres maximaux sans rayon, cardinalité maximaux des ensembles disjoints de rayons. Math. Nachr. 115 (1984) 337–352.
[7] N. Polat, Topological aspects of infinite graphs, in: Hahn, Sabidussi and Woodrow, eds., Cycles and Rays, NATO ASI Ser. C, Vol. 301 (1990) 197–220.
[8] P. Seymour and R. Thomas, An end-faithful spanning tree counterexample, submitted.
[9] J. Širáň and Ľ. Niepel, A refinement of Halin's end-faithful spanning tree theorem, Abh. Math. Sem. Univ. Hamburg 59 (1989) 85–88.
[10] C. Thomassen, Infinite graphs, in: Selected Topics in Graph Theory III (Acadademic Press, New York, 1983) 129–160.

Discrete Mathematics 95 (1991) 341–358
North-Holland

Fast growing functions based on Ramsey theorems

H.J. Prömel

Forschungsinstitut für Diskrete Mathematik, Universität Bonn, Nassestr. 2, W-5300 Bonn 1, Germany

W. Thumser

Fakultät für Mathematik, Universität Bielefeld, W-4800 Bielefeld 1, Germany

B. Voigt

Fakultät für Mathematik, Universität Bielefeld, W-4800 Bielefeld 1, Germany

Received 10 November 1989
Revised 17 June 1990

Abstract

Prömel, H.J., W. Thumser and B. Voigt, Fast growing functions based on Ramsey theorems, Discrete Mathematics 95 (1991) 341–358.

Introduction

Gödel's paper (1931) on formally undecidable propositions of first order Peano Arithmetic showed that any recursive axiomatic system which contains the axioms of Peano Arithmetic still admits propositions which are not decidable, i.e., which are neither provable nor refutable on the basis of the given axiomatic system. For the reader who is not used to work in Peano Arithmetic we mention that (for statements about natural numbers) Peano Arithmetic is equivalent to the result of replacing the axiom of infinity by its negation in the usual axioms of Zermelo-Fraenkel set theory (see, e.g., Jech [10] for these axioms). Gödel's original example of such a proposition was not that illuminating, as it merely formalized the well-known antinomy of the lyer. That raised the problem to find intuitively meaningful propositions which are valid in the 'real world' but which are not provable in Peano arithmetic (of course, such statements should be expressible in terms of first-order logic).

In this paper we present a general framework to define fast growing functions based on Ramsey theorems. This framework is suggested by the work of Ketonen and Solovay [13] and Kanamori and McAloon [12]. One advantage of our

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

approach is that we are dealing just with colorings of pairs. Such colorings may be interpreted as edge colorings of graphs. To appreciate the fast growingness of our functions we need certain sample functions to compare with. These sample functions are defined by ordinal recursion and are in fact generating functions for hierarchies of recursive functions.

In Section 1 we investigate a certain function $\mathrm{KM}:\omega\to\omega$, which has been introduced by Kanamori and McAloon [12]. Using a model theoretic argument they show that this function $\mathrm{KM}(m)$ fails to be primitive recursive. We use a purely combinatorial argument to achieve the same result, thereby obtaining explicit lower bounds. The ideas behind this argument are extended in Section 2, where an auxiliary example of a fast growing function is presented. A general framework which is extracted from the prototypical functions of Section 1 and 2 is provided in Section 3. Section 4, then, contains the main result as well as some applications. One might be tempted to object that the fast growingness of our Ramsey functions is merely due to the implicit use of ordinals in its definition. But this is not quite justified. Other basic approaches to fast growing functions, viz., via well-quasi-ordered sets (cf. Simpson [25]) also implicitly rely on the notion of ordinals.

1. A nonprimitive recursive Ramsey function

In this section we consider a rapidly growing Ramsey function which proves not to be primitive recursive. However, as it turns out, its growth rate is only slightly above all primitive recursively growing functions. The ideas behind our approach, based upon the work of McAloon and Kanamori [12], are generalized in the sequel in order to come up with faster and faster growing functions giving rise to noncollapsing hierarchies (cf., Rose [22]).

If we estimate functions and want to visualize their growth rates we are used to compare with certain well-understood functions which are supposed to be prototypical. So we speak, e.g., of polynomial or exponential growth rate. The kind of functions we are dealing with here are growing much faster. To appreciate this we provide a sample of fast growing functions.

We start with the so called Grzegorczyk hierarchy of primitive recursive functions.

For mappings $F:\omega\to\omega$ we denote by $F^n:\omega\to\omega$ the n-fold iterate of F, $F^{n+1}(x)=F(F^n(x))$ where $F^0(x)=x$. Consider the family $(F_k)_{k<\omega}$ of functions $F_k:\omega\to\omega$ which is defined by

$$F_0(n)=n+1,\qquad F_{k+1}(n)=F_k^n(n).$$

Note that the functions F_k speed up with a phantastic acceleration, e.g.,

$$F_1(n)=2\cdot n,\qquad F_2(n)=2^n\cdot n,\qquad F_3(n)\geqslant 2^{2^{\cdot^{\cdot^{2^n}}}\Big\}n\text{ twos}}.$$

F_2 is an exponential function, F_3 is called the stack-of-twos function, F_4 is an iteration of the stack-of-twos function, and so forth. Spencer suggests to call F_4 the *wow-function*, as its growth rate is already beyond imagination.

However, all these functions are still primitive recursive. Moreover, for every primitive recursive function $f:\omega\to\omega$ there exist k and n_0 such that $f(n)\leqslant F_k(n)$ for all $n\geqslant n_0$, i.e., the function F_k eventually dominates f. Thus we may define a function $F_\omega:\omega\to\omega$ which is recursive but not primitive recursive by $F_\omega(n)=F_n(n)$. This is an *Ackermann function* (though Ackermann's original definition is somewhat different, compare Peter [20]).

Lemma 1.1. *For every positive integer m there exists a smallest positive integer $n=\mathrm{KM}(m)$ such that for every regressive mapping $d:[{n \atop 2}]\to[n]$, meaning that $d(x,y)<x$, for all $0<x<y<n$, there exists an m-element set $M\in[{n \atop m}]$ such that $d(x,y)=d(x,z)$ for all $x<y<z$ in M.*

Proof. By the Erdős-Rado canonization theorem [4], compare also Graham, Rothschild and Spencer [9], for every mapping $d:[{\omega \atop 2}]\to\omega$ there exists an infinite set $F\in[{\omega \atop \omega}]$ such that $d\restriction[{\omega \atop 2}]$ is a canonical coloring. If d is regressive, i.e., $d(A)<\min A$ for $A\in[{\omega \atop 2}]$ it follows that $d(A)=d(B)$ for all $A,B\in[{F \atop 2}]$ with $\min A=\min B$. Here we use the infinity of F. Alternatively, this may also be established by iterated applications of the pigeon hole principle. Notice that for fixed $x<\omega$ the two element subsets $\{x,y\}$, $x<y$, are partitioned into just x classes. Hence the result follows by a compactness argument. □

Kanamori and McAloon [12] give a model-theoretic proof to show that the function $\mathrm{KM}(m)$ is not primitive recursive. We use a purely combinatorial argument.

Notation. $\mathrm{Ram}(l,m,k)$ is the least integer n such that $n\to(m)^l_k$, this is the ordinary Ramsey function, in other words, for every coloring $d:[{n \atop l}]\to[k]$ there exists an m-element set $M\in[{n \atop m}]$ such that $d(X)=d(Y)$ for all $X,Y\in[{M \atop l}]$.

Lemma 1.2. $\mathrm{KM}(\mathrm{Ram}(2,m+3,k))\geqslant F_k(m)$.

Proof. Let $m^*=\mathrm{Ram}(2,m+3,k)$ and define a regressive mapping $d:[{\mathrm{KM}(m^*) \atop 2}]\to[\mathrm{KM}(m^*)]$ by $d(x,y)=0$ if $F_k(x)\leqslant y$ and $d(x,y)=l$ otherwise, where the numbers $0\leqslant k^*<k$ and $1\leqslant l<x$ are defined by

$$F^l_{k^*}(x)\leqslant y<F^{l+1}_{k^*}(x).$$

Note that this is a proper definition as $F_{k^*}(x)<F_{k^*+1}(x)=F^x_{k^*}(x)$ by definition of the functions F_k.

Let $M^*\in[{\mathrm{KM}(m^*) \atop m^*}]$ be such that $d(x,y)=d(x,z)$ for all $x<y<z$ in M^*. We define a k-coloring $d^*:[{M^* \atop 2}]\to[k]$ by $d^*(x,y)=0$ if $F_k(x)\leqslant y$ and $d(x,y)=k^*$ otherwise, where k^* is as above. Let $M\in[{M^* \atop m+3}]$ be such that $d^*\restriction[{M \atop 2}]$ is a constant coloring and let $x<y<z$ be the three largest elements of M. Then $m\leqslant x$ and

thus it suffices to show that $F_k(x) \leqslant y$. Assume to the contrary that $F_k(x) > y$. Then also $F_k(x) > z$, as $d(x, y) = d(x, z) \neq 0$. Hence also $F_k(y) > z$, as $F_k(x) \leqslant F_k(y)$ (one readily sees that the functions F_k are increasing). Say, $d(x, y) = d(x, z) = l$ and $d^*(x, y) = d^*(x, z) = d^*(y, z) = k^*$. Then

$$F_{k^*}^l(x) \leqslant y < z < F_{k^*}^{l+1}(x).$$

Apply F_{k^*} to this inequality. Then $z < F_{k^*}^{l+1}(x) \leqslant F_{k^*}(y)$. But this contradicts $F_{k^*}(y) \leqslant z$. □

Corollary 1.3. *The function* KM(m) *is not primitive recursive.*

Proof. It is well known that $\mathrm{Ram}(2, m+3, k) \leqslant k^{(m+3)\cdot k}$ (cf. Graham, Rothschild and Spencer [9]), in particular, it is primitive recursive. But $\mathrm{KM}(\mathrm{Ram}(2, m+3, m)) \geqslant F_m(m) = F_\omega(m)$ by the lemma. As primitive recursive functions are closed under composition the assertion follows. □

Let us note that the growth rate of the function KM(m) is approximately that of F_ω, viz., it may be shown that (cf., Prömel and Voigt [21]):

Theorem 1.4. $\mathrm{KM}(m) < F_{m-1}(3) < F_\omega(m)$.

2. A prototypical fast growing Ramsey function

We extend the ideas of the previous section and define a Ramsey function which is growing much faster than all primitive recursive functions. For doing so we need additional sample functions to compare with. To enlarge the family $(F_k)_{k<\omega}$ into the transfinite we first require an effective system of notations for ordinals less than ϵ_0.

We assume that the reader is acquainted with the arithmetic of ordinals, compare, e.g., Bachmann [1].

In the following greek letters denote (countable) ordinals.

Lemma 2.1. *Every ordinal $\alpha > 0$ can be represented uniquely as*

$$\alpha = \omega^{\alpha_1} \cdot n_1 + \omega^{\alpha_2} \cdot n_2 + \cdots + \omega^{\alpha_k} \cdot n_k,$$

where $\alpha_1 > \alpha_2 > \cdots > \alpha_k \geqslant 0$ are ordinals and $n_1, \ldots, n_k$ are positive integers. Additionally, for $\alpha < \epsilon_0$ it follows that $\alpha_1 < \alpha$.

Recall that

$$\epsilon_0 = \omega^{\omega^{\cdot^{\cdot^{\cdot}}}} = \lim_{n\to\infty} \omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega}}}}\Big\}n \text{ times}}$$

is the first fixed point of the ordinal function $\alpha \to \omega^\alpha$.

The representation of α hinted at in Lemma 2.1 is the *Cantor normal form* of α. A coding of ordinals $\alpha < \epsilon_0$ with positive integers can be defined straightforwardly relying on the Cantor normal form. The details are somewhat technical. We refer, e.g., to Schütte [24].

Next we define *fundamental sequences* which are exploited in the definition of the sample functions F_α, $\alpha < \epsilon_0$. We need these fundamental sequences to handle limit ordinals properly.

To every limit ordinal $\alpha < \epsilon_0$ we associate a strictly monotone sequence $\alpha[n]$, $n < \omega$, which approaches α from below. For convenience we introduce $\alpha[n]$ for successor ordinals as well as for 0.

Let $0[n] = 0$ and let $(\alpha + 1)[n] = \alpha$ for all $n < \omega$. In general, if $\alpha < \epsilon_0$ is given in its Cantor normal form $\alpha = \alpha' + \omega^{\alpha_k} \cdot n_k$, where α_k is the minimal exponent, let

$$\alpha[n] = \begin{cases} \alpha' + \omega^{\alpha_k} \cdot (n_k - 1) + \omega^{\alpha_k[n]} \\ \quad \text{if } \alpha_k \text{ is a limit ordinal,} \\ \alpha' + \omega^{\alpha_k} \cdot (n_k - 1) + \omega^{\alpha_k - 1} \cdot n \\ \quad \text{if } \alpha_k \text{ is a successor ordinal,} \\ \alpha' + n_k - 1 \\ \quad \text{if } \alpha_k = 0. \end{cases}$$

For example, $\omega[n] = n$, $\omega^\omega[n] = \omega^n$, $\omega^{\omega^\omega}[n] = \omega^{\omega^n}$ and $\omega^{\omega^n}[n] = \omega^{\omega^{n-1} \cdot n}$.

With the aid of these fundamental sequences we define functions F_α, $\alpha < \epsilon_0$, by extending the definitions from Section 1:

$$F_0(n) = n + 1, \qquad F_{\alpha+1}(n) = F_\alpha^n(n), \qquad F_\alpha(n) = F_{\alpha[n]}(n)$$

for limit ordinals α.

Finally we define a function F_{ϵ_0} by

$$F_{\epsilon_0}(n) = F_{\gamma_n}(n), \quad \text{where } \gamma_n = \omega^{\omega^{\cdot^{\cdot^{\omega}}}\}n \text{ times}}.$$

This is the so called *Wainer hierarchy* of provably recursive functions (actually, it is a slight modification of Wainer's original approach which is due to Ketonen and Solovay [13]).

The significance of the Wainer hierarchy in connection with unprovability results is emphasized by a result of Wainer which relates the F_α's to the class of ordinal recursive functions. Futhermore, from Kreisel [17] it is known that the provably total recursive functions (provably total with respect to Peano arithmetic) can be characterized in terms of ordinal recursion up to ϵ_0. What we actually need is the following result, compare also Bucholz and Wainer [2].

Theorem 2.2. *Let $f : \omega \to \omega$ be a provably total recursive function (with respect to Peano arithmetic). Then f is eventually dominated by some F_α for an $\alpha < \epsilon_0$. Moreover, F_{ϵ_0} eventually dominates every provably total recursive function but it is itself not provably total recursive.*

The sample functions F_α are used to estimate the growth rate of rapidly growing functions. In particular, if such a function eventually dominates every provably total recursive function then the formula expressing its totality cannot be proved in Peano arithmetic.

Suppose we want to compute $F_\beta(n)$. The computation proceeds along the recursive definition of the F_α's. The set of names of previous functions we need to know is given by

$$N(\beta, n) = \{\beta[n]\underbrace{[n]\cdots[n]}_{i \text{ times}} \mid i < \omega\}.$$

For example,

$$\begin{aligned} N(\omega, n) &= \{n, n-1, \ldots, 0\}, \\ N(\omega^\omega, n) &= \{\omega^n, \omega^{n-1}\cdot n, \omega^{n-1}\cdot(n-1) + \omega^{n-2}\cdot n, \\ &\qquad \omega^{n-1}\cdot(n-1) + \omega^{n-2}\cdot(n-1) + \omega^{n-3}\cdot n, \ldots\} \\ &\subseteq \left\{\sum_{k=0}^{n} \omega^k \cdot n_k \mid n_k \in [n+1]\right\}. \end{aligned}$$

Properties of the fundamental sequences, as defined above, which are relevant in our proof, are summarized in the following observation.

Observation 2.3. (1) *$F_\alpha(k) \leqslant F_\alpha(l)$ for all $\alpha < \epsilon_0$ and positive integers $k \leqslant l$,*
(2) *$F_\alpha(k) \leqslant F_\beta(k)$ for all $\alpha \in N(\beta, k)$,*
(3) *$N(\alpha, k) \subseteq N(\alpha, l)$ for all positive integers $k \leqslant l$.*

Proof. The assertions (1), (2) and (3) resp. are proved by straightforward inductions on α. We show how to prove (3). If $\alpha = \beta + 1$ is a limit number then

$$N(\beta + 1, k) = \{\beta\} \cup N(\beta, k) \subseteq \{\beta\} \cup N(\beta, l) = N(\beta + 1, l),$$

as $N(\beta, k) \subseteq N(\beta, l)$ by induction. Next let α be a limit ordinal, say, $\alpha = \beta + \omega^\gamma$ with $\beta > \omega^\gamma$ and $\gamma > 0$. If $\gamma = \delta + 1$ is a successor then

$$N(\alpha, k) = \{\beta + \omega^\delta \cdot k\} \cup N(\beta + \omega^\delta \cdot k, k) \subseteq \{\beta + \omega^\delta \cdot l\} \cup N(\beta + \omega^\delta \cdot l, l)$$

as $N(\beta + \omega^\delta \cdot k, k) \subseteq N(\beta + \omega^\delta \cdot k, l) \subseteq N(\beta + \omega^\delta \cdot l, l)$ by induction and since

$$\beta + \omega^\delta \cdot k \in N(\beta + \omega^\delta \cdot l, l) \quad \text{for } k < l.$$

If γ is a limit number then

$$N(\alpha, k) = \{\beta + \omega^{\gamma[k]}\} \cup N(\beta + \omega^{\gamma[k]}, k) \subseteq N(\alpha, l)$$

as $N(\gamma, k) \subseteq N(\gamma, l)$ by induction. □

Recall that we assume that a primitive recursive coding of ordinals $\alpha < \epsilon_0$ into ω is available and thus we may talk about mappings into ϵ_0, knowing that these are actually mappings into ω.

Equipped with these tools we define a fast growing Ramsey function KS (Ketonen-Solovay) and then prove that it is actually growing as fast as F_{ϵ_0}.

Lemma 2.4. *Let k and m be positive integers. Then there exists a least positive integer $n = \mathrm{KS}_{\epsilon_0}(k, m)$ such that for every mapping $d:[{n \atop 2}] \to \epsilon_0 \times \omega$ with $d(x, y) \in N(\gamma_k, x) \times [x]$ there exists an m-element set $M \in [{n \atop m}]$ such that for all $x < y < z$ in M*

(1) $d(x, y) = d(x, z)$,

(2) $\alpha \leqslant \alpha'$, *where* $d(x, y) = (\alpha, l)$ *and* $d(y, z) = (\alpha', l')$.

Proof. Again, we use a compactness argument. So one first shows that for every mapping $d:[{\omega \atop 2}] \to \epsilon_0 \times \omega$ such that $d(x, y) \in N(\gamma_k, x) \times [x]$ for all $x < y$ there exists an infinite set $F \in [{\omega \atop \omega}]$ such that (1) and (2) hold for all $x < y < z$ in F.

As in the proof of Lemma 1.1 assertion (1) readily follows from the Erdős-Rado canonization theorem, since the sets $N(\gamma_k, x)$ are finite. Assertion (2), then follows by noting that there does not exist any infinite descending chain of ordinals. □

Lemma 2.5. $\mathrm{KS}_{\epsilon_0}(k, m+3) \geqslant F_{\gamma_k}(m)$.

Proof. Let the mapping $d:[{\omega \atop 2}] \to \epsilon_0 \times \omega$ be defined by $d(x, y) = (0, 0)$ if $F_{\gamma_k}(x) \leqslant y$ and $d(x, y) = (\alpha, l)$ otherwise, where $\alpha \in N(\gamma_k, x)$ and $l \in [1, x-1]$ satisfy

$$F_\alpha^l(x) \leqslant y < F_\alpha^{l+1}(x).$$

One readily sees that α and l are defined properly. Let $M \in [{\omega \atop m+3}]$ satisfy (1) and (2) of Lemma 2.4 and let $x < y < z$ be the three largest elements of M. We show that $d(x, y) = (0, 0)$ or $d(y, z) = (0, 0)$ from which $\mathrm{KS}_{\epsilon_0}(m+3) \geqslant F_{\gamma_k}(m)$ follows. Assume to the contrary that $d(x, y) = d(x, z) = (\alpha, l)$ and $d(y, z) = (\beta, l')$ with $l, l' > 1$ and $\alpha \leqslant \beta$. Then

$$F_\alpha^l(x) \leqslant y < z < F_\alpha^{l+1}(x).$$

We apply F_α to this inequality. By assertion (1) of Observation 2.3 it follows that

$$z < F_\alpha^{l+1}(x) \leqslant F_\alpha(y). \tag{$*$}$$

By definition of $d(y, z)$ we know that

$$F_\beta^{l'}(y) \leqslant z. \tag{$**$}$$

But

$$F_\alpha(y) \leqslant F_\beta(y) \leqslant F_\beta^{l'}(y) \tag{$***$}$$

by Observation 2.3 and as $l' \geqslant 1$. Now $(*)$, $(**)$ and $(***)$ produce the obvious contradiction that $z < z$. □

Corollary 2.6. $\mathrm{KS}_{\epsilon_0}(m, m+3) \geqslant F_{\epsilon_0}(m)$.

This lower bound gives essentially the right order of magnitude, as the following result of Thumser [27] shows.

Theorem 2.7. $\mathrm{KS}_{\epsilon_0}(m, m) \leqslant F_{\omega^2_{m+1}}(g(m))$, *where g is an appropriate primitive recursive function and*

$$\omega^2_m = \omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega^2}}}}\Big\}m\ \omega\text{'s}}.$$

3. A general framework for developing fast growing Ramsey functions

In this Section we set up a general framework which is sufficient to exploit the ideas introduced so far. What we require are notational systems for initial segments of ordinals, in particular, appropriate selections of fundamental sequences.

3.1. *Some remarks about the second number class, fundamental sequences and hierarchies*

Throughout this chapter we are working within some fixed initial segment Δ of the second number class of ordinals, i.e., all ordinals strictly less than a given countable ordinal. To develop constructively noncollapsing hierarchy classes $\mathscr{F}^\alpha$ ($\alpha \in \Delta$) of functions some requirements concerning the ordinals in Δ must be met. We tacitly suppose all ordinals in Δ to be smaller than the least nonconstructive ordinal ω_1 in the sense of Church and Kleene [3]; otherwise there would even arise problems in ordinal notation. Although there is no problem in defining the initial functions on which our hierarchy is based at successor ordinals (diagonalization will do!) there still remains the question of choosing appropriate fundamental sequences at limit stages. Kleene in his seminal paper [16] succeeded in defining function classes based on his system $\mathcal{O}$ of ordinal notations; nevertheless he did not touch on the problem mentioned before. Despite the facts that recursively well-ordered relations on ω already possess primitive recursive well-orderings of the same order type (Kleene [15]) and that for all known ordinal notations (cf. Schütte [24]) there seems to be a natural choice of fundamental sequences, a general method for obtaining such sequences has not been found yet (and it is most unlikely to be found).

3.2. *The Bachmann property*

3.2.1. *Fundamental sequences for $\lambda < \Delta$*

A sequence $(\lambda(n))_{n<\omega}$ of ordinals in Δ is called a fundamental sequence for λ if (1) $\lambda(n) < \lambda$, (2) $\lambda(n) < \lambda(n+1)$, *and* (3) $\lim_{n<\omega}\lambda(n) = \lambda$. We extend the

definition of fundamental sequences to successor ordinals by $(\alpha+1)(n)=\alpha$ for $n<\omega$. For convenience we write $\beta(0)^n$ for $\beta(0)(0)\cdots(0)$, the n-fold interate of the fundamental sequence evaluated at zero.

3.2.2. *Bachmann property*

An assignment of fundamental sequences to all $\lambda<\Delta$, λ a limit ordinal, has the Bachmann property if for all limit ordinals $\lambda, \mu<\Delta$, and all positive integers $n<\omega$ the inequality $\lambda(n)<\mu\leqslant\lambda(n+1)$ implies that $\lambda(n)\leqslant\mu(0)$.

3.2.3. *Relation* $<_A$

For $\alpha<\beta<\Delta$ we let $\alpha<_A\beta$ if $\alpha=\beta(0)^n$ for some positive $n<\omega$.

3.2.4.

The relation $<_A$ is *built up* if for all limit ordinals $\lambda\in\Delta$, and all positive integers $n<\omega$ it follows that $\lambda(n)<_A\lambda(n+1)$.

Comment. The Bachmann property has been studied in Bachmann [1]. It requires the first term of the fundamental sequence for λ to be as large as possible and enables us to prove certain monotonicity properties of the hierarchy-generating functions thus assuring its noncollapsibility. The concept of built up relations $<_A$ is due to Schmidt [23]. She also discovered its equivalence to the Bachmann property.

Proposition 3.2.1 (Schmidt [23]). *An assignment of fundamental sequences to Δ has the Bachmann property if and only if $<_A$ is built up.*

Proof. By transfinite induction (see Rose [22]). □

Although there is no way to assign fundamental sequences to the whole second number class satisfying the Bachmann property, Schmidt [24] succeeded in constructing such fundamental sequences for each proper initial segment of the second number class.

Theorem 3.2.2 (Schmidt [23]). *For every proper initial segment Δ of the second number class there exists an assignment of fundamental sequences with the Bachmann property.*

It is well known that the Bachmann property implies certain monotonicity properties.

Proposition 3.2.3. *Let an initial segment Δ of the second number class and an assignment of fundamental sequences to Δ satisfying the Bachmann property be*

given. Let $(H_\alpha)_{\alpha\in A}$ *be a sequence of functions* $H_\alpha:\omega\to\omega$ *with the following properties*:

(i) H_0 *is strictly monotone,*

(ii) $H_{\alpha+1}$ *is strictly monotone whenever* H_α *is strictly monotone,*

(iii) $H_\alpha(0)\leqslant H_{\alpha+1}(0)$ *and* $H_\alpha(x+1)<H_{\alpha+1}(x+1)$,

(iv) $H_\lambda(x)=H_{\lambda(x)}(x)$ *if* λ *is a limit ordinal.*

Then if follows for all $\alpha,\beta\in\Delta$ *that*:

(a) H_α *is strictly monotone,*

(b) *if* $\beta<_A\alpha$ *then* $H_\beta(0)\leqslant H_\alpha(0)$ *and* $H_\beta(x+1)<H_\alpha(x+1)$,

(c) *if* $\beta<\alpha$ *then* H_β *is eventually dominated by* H_α.

Proof. By induction on α this is fairly obvious, where we use Proposition 3.2.1 for part (b). □

3.3. *A strictly increasing hierarchy*

Let Δ be an initial segment of the second number class together with an assignment of fundamental sequences having the Bachmann property. According to Theorem 3.2.2 this is always achievable.

We define the family $(H_\alpha)_{\alpha<\Delta}$ of mappings $H_\alpha:\omega\to\omega$ as follows:

$$H_0(x)=x+1,\qquad H_{\alpha+1}(x)=H_\alpha^{x+1}(x),\qquad H_\alpha(x)=H_{\alpha(x)}(x)$$

for a limit ordinal α.

Proposition 3.3.1. *The functions* H_α *satisfy the conditions of Proposition* 3.2.3 *as well as* $H_\alpha(x)>x$.

Proof. (i) and (iv) are obvious, as is the first half of (iii) because by induction on α it follows that $H_\alpha(0)=1$. To prove (ii) and the second part of (iii) we note that by transfinite induction on α it follows that $H_\alpha(x)>x$. For α a limit ordinal or $\alpha=0$ this is obvious, so assume $\alpha=\lambda+1$. By induction on $z\in\omega\backslash\{0\}$ we have $H_\lambda^z(x)>x$ because $H_\lambda^1(x)>x$ and

$$H_\lambda^{z+1}(x)=H_\lambda(H_\lambda^z(x))>H_\lambda^z(x)>x,$$

by our main induction. Thus especially for $z=x+1$, $H_\alpha(x)=H_{\lambda+1}(x)=H_\lambda^{x+1}(x)>x$. Now we prove (ii) as follows:

$$\begin{aligned}H_{\alpha+1}(x+1)&=H_\alpha^{x+2}(x+1)=H_\alpha(H_\alpha^{x+1}(x+1))\\&>H_\alpha^{x+1}(x+1)>H_\alpha^{x+1}(x)=H_{\alpha+1}(x),\end{aligned}$$

where the last inequality follows from the assumption.

Finally, (iii) part two is proved by noting that

$$H_{\alpha+1}(x+1)=H_\alpha^{x+2}(x+1)=H_\alpha^{x+1}(H_\alpha(x+1))>H_\alpha(x+1),$$

by our previous observation. □

Remark. The hierarchy corresponding to the family $(H_\alpha)_{\alpha<\Delta}$ is obtained by letting $\mathscr{H}^0 = \mathscr{H}^1$ be the set of elementary functions and by letting $\mathscr{H}^\alpha = E\{H_\beta \mid \beta < \alpha\}$, the elementary closure of functions preceding H_α. Properties (i)–(iv) guarantee that this hierarchy is strictly increasing in the sense that $\mathscr{H}^\alpha$ is a proper subset of $\mathscr{H}^\beta$ for all $2 \leqslant \alpha < \beta < \Delta$, cf., Rose [22, p. 45].

Recall that the families $(F_\alpha)_{\alpha<\omega}$ resp., $(F_\alpha)_{<\epsilon_0}$ introduced in Section 1 and 2 slightly differ from the family of the H_α's as defined above. However, it turns out that $H_\alpha(n) = F_\alpha(n+1) - 1$. So we define the F-functions in a general setting, viz.

$$F_0(n) = n + 1, \qquad F_{\alpha+1}(n) = F_\alpha^n(n), \qquad F_\alpha(n) = F_{\alpha[n]}(n),$$

α a limit ordinal, $\alpha[n] := \alpha(n-1)$ and $n \geqslant 1$. We define $F_\alpha(0) = 1$ for convenience.

Lemma 3.3.2. *For all $n < \omega$ and $\alpha < \Delta$ we have $H_\alpha(n) = F_\alpha(n+1) - 1$.*

Proof. By transfinite induction on α we have the following:

$\alpha = 0$: $F_0(n+1) = n + 2 = (n+1) + 1 = H_0(n) + 1$.

$\alpha = \lambda + 1$: we have $F_\lambda(n+1) = H_\lambda(n) + 1$ and using induction on z:

$$\begin{aligned} F_\lambda^{z+1}(n+1) &= F_\lambda(F_\lambda^z(n+1)) = F_\lambda(H_\lambda^z(n) + 1) \\ &= H_\lambda H_\lambda^z(n) + 1 = H_\lambda^{z+1}(n) + 1, \end{aligned}$$

especially

$$\begin{aligned} F_\alpha(n+1) &= F_{\lambda+1}(n+1) = F_\lambda^{n+1}(n+1) = H_\lambda^{n+1}(n) + 1 \\ &= H_{\lambda+1}(n) + 1 = H_\alpha(n) + 1. \end{aligned}$$

α a limit ordinal:

$$\begin{aligned} F_\alpha(n+1) &= F_{\alpha[n+1]}(n+1) = F_{\alpha(n)}(n+1) \\ &= H_{\alpha(n)}(n) + 1 = H_\alpha(n) + 1. \quad \square \end{aligned}$$

Remark. Generally $\alpha[n]$ fails to have the Bachmann property. However, from the lemma it follows that, in any case, the hierarchies belonging to the families $(H_\alpha)_{\alpha<\Delta}$ and $(F_\alpha)_{\alpha<\Delta}$ resp., are the same. For convenience we state our results in terms of the F_α's. This is possible as the family $(F_\alpha)_{\alpha<\Delta}$ satisfies properties (1), (2) and (3) of Observation 2.3.

Lemma 3.3.3. *$F_\alpha(k) \leqslant F_\alpha(l)$ for all $\alpha < \Delta$ and all positive integers $k \leqslant l$.*

Proof. By Propositions 3.3.1 and 3.2.2 H_α is strictly monotone and $F_\alpha(x) = H_\alpha(x-1) + 1$ for $x \geqslant 1$ by Lemma 3.3.2, so F_α has the property required in the lemma. $\square$

As in the previous section let us define for $\alpha < \Delta$ and $k \geqslant 1$,

$$N(\alpha, k) := \{\alpha[k]^i \mid i \geqslant 1\}.$$

Lemma 3.3.4. *Clearly, $N(\alpha, k)$ is a finite set, moreover, $F_\alpha(k) \leqslant F_\beta(k)$ for all $\alpha \in N(\beta, k)$, $\beta < \Delta$ and $k \geqslant 1$.*

Proof. Observe, that $F_\alpha(k) = H_\alpha(k-1) + 1 > k - 1 + 1 = k$ because of Lemma 3.3.2 and Proposition 3.3.1. As $\alpha = \beta[k]^i$, $i > 0$ we use induction on i.

$$i = 1: \qquad F_\beta(k) = \begin{cases} F_{\beta[k]}(k) \\ F^k_{\beta[k]}(k) \end{cases} \geqslant F_\alpha(k)$$

according to whether β is a limit ordinal or not ($\beta = 0$ being trivial). For the inductive step the same argument works. □

Lemma 3.3.5. *$N(\alpha, k) \subseteq N(\alpha, l)$ for all $\alpha \in \Delta$, $1 \leqslant k \leqslant l$.*

Proof. It suffices to prove that $N(\alpha, k) \subseteq N(\alpha, k+1)$. We use transfinite induction on α: $\alpha = 0$ is trivial. Suppose $\alpha = \beta + 1$.

$$\begin{aligned} N(\alpha, k) &= N(\beta + 1, k) = N(\beta, k) \cup \{\beta\} \subseteq N(\beta, k+1) \cup \{b\} \\ &= N(\beta + 1, k+1) = N(\alpha, k+1). \end{aligned}$$

Finally, let α be a limit ordinal. The inductive assumption gives

$$N(\alpha[k], k) \subseteq N(\alpha[k], k+1).$$

So it suffices to show that $\alpha[k] \in N(\alpha, k+1)$. Since $\alpha[k] < \alpha[k+1]$ choose $i \geqslant 1$ maximal with respect to $\alpha[k] < \alpha[k+1]^i \leqslant \alpha[k+1]$. $\alpha[k+1]^i$ being a successor ordinal implies either $\alpha[k+1]^{i+1} = \alpha[k]$ and we are through or otherwise contradicts the maximality of i. So suppose $\alpha[k+1]^i =: \lambda$ is a limit number. We show, that this case cannot occur. Otherwise, using the Bachmann property we would get

$$\alpha[k] \leqslant \lambda(0) = \lambda[1] = \alpha[k+1]^i[1] < \alpha[k+1]^{i+1},$$

which again would contradict the maximality of i. □

Let us summarize the preceding lemmas in the following theorem.

Theorem 3.3.6. *Let an arbitrary initial segment Δ of the second number class together with an assignment of fundamental sequences satisfying the Bachmann property be given. Define $\alpha[n+1] := \alpha(n)$, for $n \geqslant 0$ and define the family $(F_\alpha)_{\alpha<\Delta}$ by*

$$F_0(n) = n + 1, \qquad F_{\alpha+1}(n) = F^n_\alpha(n), \qquad F_\alpha(n) = F_{\alpha[n]}(n),$$

where $F_\alpha(0) = 1$. *Then* $(F_\alpha)_{\alpha<\Delta}$ *defines a strictly increasing hierarchy, extending Grzegorczyk's hierarchy, and satisfies*:

(1) $F_\alpha(k) \leqslant F_\alpha(l)$ *for all* $1 \leqslant k \leqslant l < \omega$ *and* $\alpha < \Delta$.

(2) $F_\alpha(k) \leqslant F_\beta(k)$ *for all* $k < \omega$ *and* $\alpha \leqslant \beta < \Delta$.

(3) $N(\alpha, k) \subseteq N(\alpha, l)$ *for all* $1 \leqslant k \leqslant l < \omega$ *and* $\alpha < \Delta$.

4. Main theorem and applications

4.1. Main theorem

In order to state our theorem let us fix some segment $\Delta = \lambda + 1$ of the second number class together with an assignment of fundamental sequences $(\lambda_n := \lambda[n])$. Appealing to a coding system (recall our tacit assumption that Δ should be a constructive ordinal), thus all following mappings actually map into ω.

Lemma 4.1.1. *Let* k *and* m *be positive integers. Then there exists a smallest positive integer* $n = \mathrm{KS}_\lambda(k, m)$ *such that for every mapping* $d : [{n \atop 2}] \to \lambda \times \omega$ *with* $d(x, y) \in N(\lambda_k, x) \times [x]$ *there exists an* m*-element set* $M \in [{n \atop m}]$ *such that for all* $x < y < z$ *in* M *it follows that*:

(1) $d(x, y) = d(x, z)$,

(2) $\alpha \leqslant \alpha'$ *for* $d(x, y) = (\alpha, l)$, $d(y, z) = (\alpha', l')$.

Proof. Exactly as in Section 2. □

Theorem 4.1.2. $\mathrm{KS}_\lambda(k, m+3) \geqslant F_{\lambda_k}(m)$, $\mathrm{KS}_\lambda(m, m+3) \geqslant F_\lambda(m)$.

Proof. The essential conditions used in the proof of Lemma 2.5 are established in Theorem 3.3.6. We omit to repeat the arguments. □

4.2. Applications

4.2.1. KS_ω

Let $\lambda = \omega$ and define $\lambda(n) = \lambda[n+1] = n + 1$. The function KS_ω fails to be primitive recursive. Comparing the definitions one readily observes that $\mathrm{KS}_\omega(1, m) = \mathrm{KM}(m)$. From Corollary 1.3 it follows that already $\mathrm{KS}_\omega(1, m)$ is not primitive recursive, whereas Theorem 4.1.2 just asserts that $\mathrm{KS}_\omega(1, m) \geqslant 2 \cdot m$. Thus one may wonder about the real growth rate of the function $\mathrm{KS}_\omega(k, m)$. It turns out that the lower bound of Theorem 4.1.2 is not too bad, as one may show that $\mathrm{KS}_\omega(k, m) \leqslant F_{\omega^{1+k}}(k \cdot m)$ (Thumser [27]).

4.2.2. KS_{ϵ_0}

Let $\lambda = \epsilon_0$ and associate fundamental sequences with λ as in Section 2 using the Cantor normal form; however, we let

$$(\alpha' + \omega^{\alpha_k} \cdot n_k)(n) = \alpha' + \omega^{\alpha_k} \cdot (n_k - 1) + \omega^{\alpha_k - 1}(n + 1)$$

if α_k is a successor ordinal. A straightforward calculation shows that the above defined assignment has the Bachmann property, cf. also Rose [22, p. 78]. The sample functions $(F_\alpha)_{\alpha<\epsilon}$ as defined in Section 3 rely on the shifted assignment $\alpha[n+1]=\alpha(n)$. This square bracket assignment proves to be the assignment used in Section 2, as $\omega^{\beta+1}[n]=\omega^\beta\cdot(n-1)=\omega^\beta\cdot n$.

Although the predicate '$\mathrm{KS}_{\epsilon_0}(k,m)=n$' is primitive recursive (the number of color-checks can be bounded by an elementary function in k, m, n) and therefore is expressible in Peano arithmetic, there is no way of proving $\forall k, m\,\exists n\,\mathrm{KS}_{\epsilon_0}(k,m)=n$ in PA because of the rapid growth of KS_{ϵ_0}. As we mentioned in Section 2 no problems arise in coding ordinals less than ϵ_0 to get a system of notations satisfying the conditions of the last section: we simply use Cantor's normal form. Recalling $\epsilon_{0_k}=\epsilon_0[k]=\gamma_k$ we see in this particular case that

$$\mathrm{KS}_{\epsilon_0}(k, m+3)\geqslant F_{\gamma_k}(m).$$

We also know from Parsons [19], that the class of functions provably recursive in PA_{k-1}, where induction is restricted to formulas having at most $(k-1)$-nested quantifiers, is identical to the class $\mathscr{F}^{\gamma_k}$. So it is impossible to prove Lemma 4.1.1 within this restricted theory where k is supposed to be fixed.

4.2.3. KS_{Γ_0}

To extend our results further into the transfinite, let us choose a notational system sufficient to provide expressions for all ordinals less than the first strongly critical ordinal Γ_0. This ordinal was introduced by Feferman [6] and is analyzed in great detail in Schütte [24]. Γ_0 is also known to be the proof theoretical ordinal of Feferman's system of predicative analysis [5] as well as of Friedman's system ATR_0. It is shown in Friedman [8] that both systems are able to prove the same Π^1_1 sentences and a great deal of classical mathematics can be developed in either of them. The notational system to be employed is based upon some basic facts, which we want the reader to recall. Let $\Phi_\alpha:\mathrm{Ord}\to\mathrm{Ord}$ (Ord denotes the set of countable ordinals in any of the usual set theories) be given by

$$\Phi_0(\alpha)=\omega^\alpha$$

$$\Phi_\beta(\alpha)=\text{the } \alpha\text{th common fixed point of all } \Phi_\gamma,\ \gamma<\beta$$

(We have $\Phi_1(\alpha)=\epsilon_\alpha$).

Γ_0 can be characterized as the least ordinal satisfying $\Phi_{\Gamma_0}(0)=\Gamma_0$.

We want to make some comments concerning Fig. 1. All entries are so called principle ordinals, i.e. those $\alpha\in\mathrm{Ord}\,(\alpha\neq 0)$ which satisfy $\beta+\alpha=\alpha$ for all $\beta<\alpha$. They are listed in increasing order in the first row by the function $\Phi_0(\gamma)=\omega^\gamma$. Principal ordinals are indecomposable with respect to ordinal addition and each ordinal $\gamma\neq 0$ may be uniquely written as a decreasing, finite ordinal sum built up by principal ordinals (Cantor's normal form theorem). The

	0	1	2	3	$\cdots$	ω	$\cdots$	ϵ_0	$\cdots$
0	1	ω	ω^2	ω^3	$\cdots$	ω^ω	$\cdots$	ϵ_0	$\cdots$
1	ϵ_0	ϵ_1	ϵ_2	$\cdots$					
$\vdots$	$\vdots$								
ω	$\Phi_\omega(0)$	$\cdots$							
$\vdots$	$\vdots$								
Γ_0	$\Phi_{\Gamma_0}(0) = \Gamma_0$								

Fig. 1.

next theorem gives an appropriate choice and tells us which entries should be used in order to avoid infinite regress. To be more precise we take a closer look at the diagram and list all occurrences of some principal ordinal α, see Fig. 2. Up to some point δ we have $\Phi_\gamma(\alpha) = \alpha$ for all $\gamma < \delta$. These entries are of no help if we want a notational system, they would be denoted by themselves. So we take hold on the circled $\alpha = \phi_\delta(\beta)$ hoping $\delta < \alpha$ in order to continue. But

$$\alpha = \phi_\delta(\beta) \geqslant \phi_\delta(0) \geqslant \delta,$$

and $\delta < \alpha$ firstly may fail as soon as $\alpha = \delta$, which implies $\beta = 0$ and $\alpha = \phi_\alpha(0) = \Gamma_0$ by definition. Up to Γ_0 the process requires no infinite regress and can be used to prove Theorem 4.3.1. The existence of the circled α entry follows from the fact that $\phi_\alpha(\alpha) > \phi_\alpha(0) \geqslant \alpha$, where we use the normality (strict monotonicity and continuity) of $\phi_\alpha(\cdot)$ and $\phi_{(\cdot)}(\alpha)$. It should be noted however that generally $\phi_{(\cdot)}(\alpha)$, $\alpha \neq 0$, fails to be continuous: otherwise the inequalities $\phi_n(0) < \phi_n(1) < \phi_{n+1}(0)$ would imply $\phi_\omega(0) = \phi_\omega(1)$, contradicting the strict monotonicity of ϕ_ω.

The following theorem can be used to establish a convenient notational system and coding into ω, cf. Schütte [24].

	0	1	1	$\cdots$	β	$\cdots$	α	$\cdots$
0							α	
1							α	
2							α	
$\vdots$					α		$\vdots$	
δ								
$\vdots$								
α								
$\vdots$								

Fig. 2.

Theorem 4.3.1. *For each ordinal* $0 \neq \gamma < \Gamma_0$ *there exist uniquely determined ordinals* $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ $(n \geqslant 1)$ *where* $\alpha_i, \beta_i < \Phi_{\alpha_i}\beta_i$ $(i = 1, \ldots, n)$ *and* $\Phi_{\alpha_1}\beta_1 \geqslant \cdots \geqslant \Phi_{\alpha_n}\beta_n$ *and* $\gamma = \Phi_{\alpha_1}\beta_1 + \cdots + \Phi_{\alpha_n}\beta_n$.

Using this kind of Cantor normal form, we are now able to assign fundamental sequences to limit ordinals $\gamma < \Gamma_0$ by concentrating the definition always on the rightmost term $\Phi_{\alpha_n}\beta_n$. In detail we proceed as follows by transfinite induction (λ, μ are supposed to denote limit ordinals):

$$\begin{aligned}
&\Phi_0(\beta+1)(n) = \Phi_0(\beta)\cdot(n+2),\\
&\Phi_0(\lambda)(n) = \Phi_0(\lambda(n)),\\
&\Phi_{\alpha+1}(0)(n) = \Phi_\alpha^{n+1}(0),\\
&\Phi_{\alpha+1}(\beta+1)(n) = \Phi_\alpha^{n+1}(\Phi_{\alpha+1}(\beta)+1),\\
&\Phi_{\alpha+1}(\lambda)(n) = \Phi_{\alpha+1}(\lambda(n)),\\
&\Phi_\lambda(0)(n) = \Phi_{\lambda(n)}(0),\\
&\Phi_\lambda(\beta+1)(n) = \Phi_{\lambda(n)}(\Phi_\lambda(\beta)+1),\\
&\Phi_\lambda(\mu)(n) = \Phi_\lambda(\mu(n)).
\end{aligned}$$

Because of the normality of the functions $\Phi_{\lambda(\cdot)}$ the last definition provides us with a fundamental sequence converging to $\Phi_{\lambda(\mu)}$ assuming already one converging to μ. Let us justify the second-last definition (the other definitions may be established in a similar way).

Lemma 4.3.2. $\Phi_{\lambda(n)}(\Phi_\lambda(\beta)+1) \to \Phi_\lambda(\beta+1)$ *whenever* $\lambda(n) \to \lambda$, *where* $\to$ *denotes convergence.*

Proof. Let $\gamma := \sup_{n<\omega} \Phi_{\lambda(n)}(\Phi_\lambda(\beta)+1)$. We claim that $\gamma \leqslant \phi_\lambda(\beta+1)$. Certainly $\Phi_\lambda(\beta)+1 < \Phi_\lambda(\beta+1)$ since $\Phi_\lambda(\beta+1)$ is principal. It follows that

$$\Phi_{\lambda(n)}(\Phi_\lambda(\beta)+1) < \Phi_{\lambda(n)}(\Phi_\lambda(\beta+1)) = \Phi_\lambda(\beta+1)$$

for all $n < \omega$ because $\Phi_\lambda(\beta+1)$ is a fixed point of all $\Phi_\delta(\delta < \lambda)$.

By definition of γ it follows that $\gamma \leqslant \Phi_\lambda(\beta+1)$ establishing our claim. On the other hand $\gamma > \Phi_\lambda(\beta)$ and it remains to prove that γ is a fixed point of all $\Phi_\delta(\delta < \lambda)$. But

$$\begin{aligned}
\Phi_\delta(\gamma) &= \Phi_\delta(\sup \Phi_{\lambda(n)}(\Phi_\lambda(\beta)+1))\\
&= \sup_{n<\omega} \Phi_\delta \Phi_{\lambda(n)}(\Phi_\lambda(\beta)+1) = \sup_{n<\omega,\lambda(n)>\delta} \Phi_{\lambda(n)}(\Phi_\lambda(\beta)+1) = \gamma,
\end{aligned}$$

where the second equality holds since Φ_δ is continuous. Therefore γ is a fixed point of all $\Phi_\delta(\cdot)$, $\delta < \lambda$, larger than their βth fixed point yielding $\Phi_\lambda(\beta+1) \leqslant \gamma$, which together with our claim finishes the proof. □

Schmidt [23] gives a slightly different, although equivalent definition.

Theorem 4.3.3. *The assignment of fundamental sequences to limit ordinals $\lambda < \Gamma_0$ given above determines a built up system of ordinal notation and therefore has the Bachmann property.*

It is straightforward but somewhat tedious to check the other requirements of notational systems we defined above. We would nevertheless like to mention that reasonable bounds for the cardinality of $N(\lambda_k, x)$ (where λ_k denotes the k-fold first argument iterate of $\Phi(\alpha, 0) := \Phi_\alpha(0)$ evaluated at 0) can be found. To deduce the following corollary one needs the analogue of Theorem 2.2 with respect to the system ATR_0, cf., Friedman, McAloon and Simpson [8].

Corollary 4.3.4. *The Ramsey function K_{Γ_0} grows faster than any recursive function f for which $\mathrm{ATR}_0 \vdash f$ is total.*

Analogous examples of even faster grower Ramsey functions may be constructed following the patterns of this paper.

References

[1] H. Bachmann, Transfinite Zahlen (2nd Ed.) (Springer, Berlin, 1967).

[2] W. Buchholz and S. Wainer, Provably computable functions and the fast growing hierarchy, in: S.G. Simpson, ed., Logic and Combinatorics, Contemporary Mathematics, Vol. 65 (Amer. Math. Soc., Providence 1987) 179–198.

[3] A. Church and S. Kleene, On notation for ordinal numbers, J. Symbolic Logic 3 (1983) 150–155.

[4] P. Erdős and R. Rado, A combinatorial theorem, J. London. Math. Soc. 25 (1950) 249–255.

[5] S. Feferman, Systems of predicative analysis, J. Symbolic Logic 29 (1964) 1–30.

[6] S. Feferman, Systems of predicative analysis II, J. Symbolic Logic 33 (1968) 193–220.

[7] H. Friedman, Harvey Friedman's Research on the Foundations of Mathematics, L.A. Harrington, M.D. Morley, A. Ščedrov and S.G. Simpson, eds. (North-Holland, Amsterdam, 1985).

[8] H. Friedman, K. McAloon and S. Simpson, A finite combinatorial principle, which is equivalent to the 1 consistency of predicative analysis, in: G. Metakides, ed., Patras Logic Symposium, (North-Holland, Amsterdam, 1982) 197–230.

[9] R.L. Graham, B.L. Rothschild and J.H. Spencer, Ramsey Theory (Wiley, New York, 1980).

[10] T. Jech, Set Theory (Academic Press, New York 1978).

[11] C.G. Jockusch, Ramsey's theorem and recursion theory, J. Symbolic Logic 37 (1972) 268–280.

[12] A. Kanamori and K. McAloon, On Gödel incompleteness and finite combinatorics, Ann. Pure Appl. Logic 33 (1987) 23–41.

[13] J. Ketonen and R. Solovay, Rapidly growing Ramsey functions, Ann. Math. 113 (1981) 267–314.

[14] S.C. Kleene, On the forms of the predicates in the theory of constructive ordinals, Amer. J. Math. 66 (1944) 41–58.

[15] S.C. Kleene, On the forms of the predicates in the theory of constructive ordinals (second paper), Amer. J. Mathematics 77 (1955) 405–428.

[16] S.C. Kleene, Extension of an effectively generated class of functions by enumeration, Colloq. Math. 6 (1985) 67–78.

[17] G. Kreisel, On the interpretation of nonfinitist proofs II, J. Symbolical Logic 17 (1952) 43–58.

[18] J. Paris and L. Harrington, A mathematical incompleteness in Peano arithmetic, in: J. Barwise, ed., Handbook of Mathematical Logic (North-Holland, Amsterdam, 1977) 1133–1142.

[19] C. Parsons, On n-quantifier induction, J. Symbolic Logic 37 (1972) 466–482.

[20] R. Peter, Recursive Functions (Academic Press, New York, 1967).

[21] H.J. Prömel and B. Voigt, Aspects of Ramsey theory I: Sets, report no 87495, Forschungsinstitut für Diskrete Mathematik, Universität Bonn, 1989.
[22] H.E. Rose, Subrecursion: functions and hierarchies (Clarendon Press, Oxford, 1984).
[23] D. Schmidt, Built-up-systems of fundamental sequences and hierarchies of number-theoretic functions, Arch. Math. Logik 18 (1976/77) 47–53, 145–146.
[24] K. Schütte, Proof Theory (Springer, Berlin, 1977).
[25] S.G. Simpson, Unprovable theorems and fast growing functions, in: S.G. Simpson, ed., Logic and Combinatorics, Contemporary Mathematics, Vol. 65 (Amer. Math. Soc., Providence, 1987) 359–394.
[26] C. Smorynski, The incompleteness theorems. in: J. Barwise, ed., Handbook of Mathematical Logic (Elsevier, Amsterdam, 1977) 821–865.
[27] W. Thumser, On upper bounds for Kanamori-McAloon functions, doctoral dissertation, Bielefeld, 1989.
[28] F. Zemke, *P.R.*-regulated systems of notation and the subrecursive hierarchy equivalence property, Trans. Amer. Math. Soc. 234 (1977) 89–118.
[29] K. Gödel, Über formal unendscheidbare Sätze der Principia Mathematica und verwandte Systeme I, Monatsh. Math. Phys. 38 (1931) 173–198.

Discrete Mathematics 95 (1991) 359–372
North-Holland

Edge-transitive strips

Mark E. Watkins

Mathematics Department, Syracuse University, Syracuse, NY 13244-1150, USA

Received 23 November 1989
Revised 29 June 1990

Abstract

Watkins, M.E., Edge-transitive strips, Discrete Mathematics 95 (1991) 359–372.

A *strip* may be characterized as an infinite locally finite connected graph admitting an automorphism with finitely many orbits. First we show that all edge-transitive strips have even valence(s). We then characterize the edge-transitive planar strips as follows. For each integer $k \geqslant 2$, there exists a unique flag-transitive example with connectivity k; it is 4-valent and 4-covalent. By introducing multiple edges to these graphs or to the 2-way infinite path and then subdividing each edge, we construct all examples that are not vertex-transitive.

1. Introduction

Highly symmetric infinite graphs of bounded valence, that is to say, locally finite graphs with a small number of orbits, such as those that are vertex- or edge-transitive, fall into classes according to their having one, two, or uncountably many ends. A now classical result due to Erdős [1] states that a countably infinite graph Γ is planar if and only if every finite subgraph is planar. Thus Γ is planar if and only if no finite subgraph is a subdivision of one of the two forbidden Kuratowski graphs K_5 and $K_{3,3}$.

Grünbaum and Shephard have studied extensively those edge-transitive infinite graphs that admit an imbedding in the plane with a unique limit point at infinity (see, for example, [4]). Their examples generally are 1-ended. J.E. Graver and the author are currently investigating infinitely-ended, vertex-transitive planar graphs (of finite valence). Some of their examples are edge-transitive, too. The latter results will be published elsewhere.

The present paper is concerned with graphs possessing exactly two ends. H.A. Jung and the author have given the name 'strip' to such a graph when it possesses finitely many orbits and have given various characterizations of strips [9]. The attention in this paper will be upon edge-transitive strips. In Section 3 it will be shown that edge-transitive strips have only even valence(s). In the final section we turn to edge-transitive planar strips and give a detailed description of all of them.

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

In particular, for every integer $b \geqslant 3$, there is a unique triconnected, edge-transitive planar strip Σ with $\kappa_\infty(\Sigma) = b$.

2. Preliminaries

Capital Greek letters will be reserved for graphs and their subgraphs. The graphs we consider may be finite or infinite, but all infinite graphs will be presumed to be *locally finite*, that is, they are infinite and every vertex has finite valence. The symbols $V(\Gamma)$, $E(\Gamma)$, $G(\Gamma)$, $\rho(\Gamma)$, $\rho_{\min}(\Gamma)$ and $\kappa(\Gamma)$ will denote, respectively, the vertex set, the edge set, the (full) automorphism group, the valence (when constant on $V(\Gamma)$), the minimum valence, and the connectivity of Γ. The graph-argument will be suppressed when no ambiguity is likely.

We say that Γ is a *flag-transitive* (also called '1-transitive' in the literature) if given $\{x_i, y_i\} \in E(\Gamma)$ $(i = 1, 2)$, there exists an automorphism $\alpha \in G(\Gamma)$ such that $\alpha(x_1) = x_2$ and $\alpha(y_1) = y_2$. If $S \subseteq V(\Gamma)$, then $\partial(S)$ denotes the set of vertices in $V(\Gamma) \backslash S$ which are adjacent to at least one vertex in S. We abbreviate $\partial(\{x\})$ to $\partial(x)$ for $x \in V(\Gamma)$. Thus $\rho(x) = |\partial(x)|$ for $x \in V(\Gamma)$.

Following convention, K_n denotes the complete graph on n vertices; its complement is the empty graph Ω_n. The complete bipartite graph with $m + n$ vertices and mn edges is denoted by $K_{m,n}$. The symbol Δ_∞ will denote the graph consisting of a 2-way-infinite path.

If R is a region of a planar imbedding of Γ, then $\rho^*(R)$, called the *covalence* of R, denotes the number of edges in the boundary of R. If Γ has a uniquely determined set of regions for all planar imbeddings and if ρ^* is a constant k on that set, then we may write $\rho^*(\Gamma) = k$ and say that Γ is *k-covalent.*

By a *lobe* of Γ (also called a 'block' in the literature) we mean either a maximal biconnected subgraph of Γ or an isthmus together with its incident vertices.

The set of integers (positive integers) is denoted by $\mathbb{Z}$ $(\mathbb{Z}^+)$, while $\mathbb{Z}_n$ denotes the ring of integers modulo n.

Given graphs Γ_1 and Γ_2, the *lexicographic product* of Γ_1 by Γ_2 is the graph $\Gamma_1[\Gamma_2]$ having vertex set $V(\Gamma_1) \times V(\Gamma_2)$ while $\{(x_1, y_1), (x_2, y_2)\} \in E(\Gamma_1[\Gamma_2])$ whenever either $x_1 = x_2$ and $\{y_1, y_2\} \in E(\Gamma_2)$ or $\{x_1, x_2\} \in E(\Gamma_1)$ (cf. [11]).

We use the notion of an 'end' in a graph as formulated by Halin [5]. In the special case of a locally finite graph Γ, the number of ends $\varepsilon(\Gamma)$ of Γ may be defined as the supremum of the number of infinite components of $\Gamma - S$, where S ranges over all finite subsets of $V(\Gamma)$. By combining [6, Corollary 15] and [7, Theorem 1], one has that $\varepsilon(\Gamma)$ equals 1 or 2 or is uncountable whenever Γ is infinite, locally finite, connected, and vertex-transitive. The group $G(\Gamma)$ induces a permutation group on the set of ends of Γ.

Suppose that Γ is connected. The cardinality of a smallest separating set $T \subseteq V(\Gamma)$ such that $\Gamma - T$ has at least one finite component is denoted by $\kappa_f(\Gamma)$. It is obvious that $\kappa_f(\Gamma) \leqslant \rho_{\min}(\Gamma)$. The cardinality of a smallest set $T \subseteq V(\Gamma)$ such

that $\Gamma - T$ has at least two infinite components is denoted by $\kappa_\infty(\Gamma)$. Unless the appropriate 'smallest set' exists, the corresponding parameter is not defined. Otherwise, clearly $\kappa = \min\{\kappa_f, \kappa_\infty\}$. It follows from the definitions that for Γ infinite, $\varepsilon(\Gamma) > 1$ if and only if $\kappa_\infty(\Gamma)$ is defined and finite.

Proposition 2.1. *If Γ is edge-transitive, then $\kappa_f(\Gamma) = \rho_{\min}(\Gamma)$.*

Although Proposition 2.1 was originally proved for finite graphs [14, Corollary 1A], the same proof is valid for infinite graphs if one invokes [8, Corollary 1B].

The proof of the following useful proposition is often given to students as an exercise.

Proposition 2.2. *If Γ is edge-transitive but not vertex-transitive, then Γ is bipartite, admitting the bipartition $V(\Gamma) = V_0 \cup V_1$, where V_0 and V_1 are the two orbits of $G(\Gamma)$ in $V(\Gamma)$.*

If Ψ is a subgraph of Γ, we say that an end $\mathscr{E}$ 'is contained in Ψ' if for every 1-way-infinite path $\Pi \in \mathscr{E}$, a 1-way-infinite subpath of Π is contained in Ψ. We say that a set $S \subseteq V(\Gamma)$ 'separates (distinct) ends $\mathscr{E}_1$ and $\mathscr{E}_2$' if there exist distinct infinite components Θ_1 and Θ_2 of $\Gamma - S$ such that $\mathscr{E}_i$ is contained in Θ_i $(i = 1, 2)$. If $|S| = \kappa_\infty(\Gamma) < \infty$ and S separates some two ends, then S will be called a *κ_∞-separator.*

Proposition 2.3 [10]. *Let S be a κ_∞-separator of a locally finite, edge-transitive graph Γ. Then $\Gamma - S$ has no finite component.*

In [9, p. 153], H.A. Jung and the author defined a connected graph Γ to be a *strip* if there exist a set $S \subseteq V(\Gamma)$ such that S induces a connected subgraph with $0 < |\partial(S)| < \infty$ and an automorphism $\alpha \in G(\Gamma)$ such that $\alpha[S \cup \partial(S)] \subseteq S$ and $|S \setminus \alpha[S]| < \infty$. We proceeded in [9, Theorem 5.13] to characterise a strip as a locally finite graph admitting an automorphism with finitely many orbits (all of them necessarily infinite). By [9, Lemma 5.3], all strips have exactly two ends.

If d denotes the distance function in Γ and $\beta \in G(\Gamma)$, let us call β *bounded* if there exists an integer d_β such that $d(x, \beta(x)) \leqslant d_\beta$ for all $x \in V(\Gamma)$. The set $B(\Gamma)$ of bounded automorphisms is a normal subgroup of $G(\Gamma)$ by [9, Lemma 5.9]. If Γ is a strip, then $B = B(\Gamma)$ is the set of automorphisms fixing the two ends of Γ, and so $[G:B] \leqslant 2$. (Cf. [9, Theorem 5.10].) The stabilizer in G of a vertex x is denoted by G_x. We let $B_x = B(\Gamma) \cap G_x$. By the Second Isomorphism Theorem for groups, we thus have

$$1 \leqslant [G_x : B_x] = [G_x B : B] \leqslant [G:B] \leqslant 2. \tag{2.4}$$

We require the following result in the last section of this paper.

Lemma 2.5. *Let Γ be a planar graph, let T be a finite set of vertices of Γ, and let Θ_1 and Θ_2 be distinct components of $\Gamma - T$. Let $x \in T$ and let u_i, $v_i \in V(\Theta_i)$ $(i = 1, 2)$ be neighbors of x. Then for no planar imbedding of Γ can the neighbors of x, in the counterclockwise cyclic ordering, form the cyclic subsequence u_1, u_2, v_1, v_2.*

Proof. For each $i = 1, 2$, there exists a $u_i v_i$-path Π_i contained in Θ_i. Form the circuit $\Xi_i = \Pi_i + \{u_i, x\} + \{x, v_i\}$. Since Ξ_1 and Ξ_2 have only the vertex x in common, these two circuits cannot cross at x, by the so-called 'Jordan Curve Theorem for Multigraphs' (cf. [3, p. 83, Theorem E22]). □

Let $d(k)$ denote the number of vertices at distance k from some fixed vertex of a connected, locally finite graph Γ. We say that Γ *has exponential growth* if there exist numbers $a > 1$ and $c > 0$ such that $d(k) \geqslant ca^k$ holds for all but finitely many k. The following result gives a sufficient condition for exponential growth for a directed graph (where $d(k)$ is determined using undirected distance).

Proposition 2.6 [13, Lemma 2]. *Let Δ be a directed graph with constant in-valence ρ^- and constant out-valence ρ^+, where $1 \leqslant \rho^- < \rho^+$ and $\rho = \rho^+ + \rho^-$, and assume that Δ has no 2-cycle. Then, measuring $d(k)$ from any vertex whatever, one has*

$$d(k) \geqslant 2[(\rho + 1)/(\rho - 1)]^{k-1}$$

for all $k \geqslant 1$.

One immediately obtains the following.

Theorem 2.7. *Let Γ be a connected infinite graph. If the edges of Γ can be so oriented that the in-valence ρ^- and the out-valence ρ^+ are constant on $V(\Gamma)$ and $1 \leqslant \rho^- < \rho^+$, then Γ has exponential growth.*

Corollary 2.8. *Let Γ be a locally finite, vertex-transitive, edge-transitive, nonflag-transitive graph, and let $\{x, y\} \in E(\Gamma)$. Then for every vertex $u \in V(\Gamma)$, there exists a partition $\{\partial^+(u), \partial^-(u)\}$ of $\partial(u)$ such that*

$$\partial^+(u) = \{v \in \partial(u)\colon \gamma(x) = u \text{ and } \gamma(y) = v \text{ for some } \gamma \in G(\Gamma)\}$$

and

$$\partial^-(u) = \{v \in \partial(u)\colon \gamma(x) = v \text{ and } \gamma(y) = u \text{ for some } \gamma \in G(\Gamma)\}.$$

If $|\partial^+(u)| \neq |\partial^-(u)|$ for some $u \in V(\Gamma)$, then Γ has exponential growth.

Proof. Since Γ is edge-transitive but not flag-transitive, it is elementary that for each $\{u, v\} \in E(\Gamma)$, either $\gamma(x) = u$ and $\gamma(y) = v$ for some $\gamma \in G(\Gamma)$ or $\gamma(x) = v$ and $\gamma(y) = u$ for some $\gamma \in G(\Gamma)$, but not both.

If $|\partial^+(u)| \neq |\partial^-(u)|$, we may suppose $|\partial^-(u)| < |\partial^+(u)|$; otherwise one may interchange x and y at the outset. The images of $\{x, y\}$ induce an orientation of Γ whereby $\rho^+ = |\partial^+(u)|$ and $\rho^- = |\partial^-(u)|$. The result now follows from Theorem 2.7. □

If Σ is a strip, then the sequence $\{d(k)\}$ is bounded [9, Lemma 5.4], and so the growth of Σ is certainly not exponential. Hence, in the language of the preceding corollary, for any vertex x of a vertex-transitive, edge-transitive, nonflag-transitive strip, one has $|\partial^+(x)| = |\partial^-(x)|$, and so x has even valence.

3. Applications of Halin's theorem to strips

In order that the present paper be more self-contained, we include a seminal result from Halin's substantial paper [6]. The statement of this result requires some preliminary definitions. We then apply Halin's result for locally finite graphs to the special case of strips.

Let Γ be any graph. An automorphism $\gamma \in G(\Gamma)$ is *of type* 1 if there exists a non-empty finite set $S \subset V(\Gamma)$ such that $\gamma[S] = S$. Otherwise γ is *of type* 2. If Γ is locally finite, then γ is of type 2 if and only if $\gamma[U] \subset U$ (n.b. proper containment) for some subset $U \subset V(\Gamma)$, (cf. [9, Lemma 2.4]).

If γ is of type 2, then the number of ends fixed by γ is either 1 or 2 (cf. [6, Theorem 8]). Thus if Σ is a strip and $\gamma \in G(\Sigma)$ is of type 2, then γ fixes both ends of Σ. It follows that all the type 2 automorphisms of a strip are contained in $B(\Sigma)$. Moreover, by their very definition, all strips admit type 2 automorphisms. Type 1 automorphisms may or may not belong to $B(\Sigma)$.

For an end $\mathscr{E}$, Halin denotes by $m_1(\mathscr{E})$ the maximum number of disjoint 1-way infinite paths belonging to $\mathscr{E}$. If γ fixes ends $\mathscr{E}_1$ and $\mathscr{E}_2$, then $m_1(\mathscr{E}_1) = m_1(\mathscr{E}_2)$. Moreover, $\mathscr{E}_1 \neq \mathscr{E}_2$ if and only if $m_1(\mathscr{E}_1) < \infty$. If $\mathscr{E}$ is an end of a strip Σ then clearly

$$m_1(\mathscr{E}) \leqslant \kappa_\infty(\Sigma) < \infty \tag{3.1}$$

Proposition 3.2 [6, Theorem 9e]. *Let γ be a type 2 automorphism of a locally finite connected graph Γ such that $m_1(\mathscr{E}) < \infty$ for the ends $\mathscr{E}$ fixed by γ. Then there exist γ-invariant 2-way infinite paths $\Pi_1, \ldots, \Pi_m$, a positive integer n, and a finite subset T of $V(\Gamma)$ such that the following hold:*

(a) $\gamma^n[\Pi_i] = \Pi_i$ $(i = 1, \ldots, m)$.

(b) *T separates the two ends fixed by γ.*

(c) *$|T| = m$, and for each $i = 1, \ldots, m$, there is a unique vertex $t_i \in T \cap V(\Pi_i)$.*

(d) *The sets in the family $\{\gamma^{ni}[T]: i \in \mathbb{Z}\}$ are pairwise disjoint.*

(e) *If $v \in V(\Pi_i)$, $w \in V(\Pi_j)$, and $v < t_i$, $t_j < w$, where $<$ denotes the natural order of Π_i (resp., Π_j) given by γ, then T separates v and w.*

(f) *If* $p, q, r \in \mathbb{Z}$ *and* $p < q < r$, *then* $\gamma^{nq}[T]$ *separates* $\gamma^{np}[T]$ *from* $\gamma^{nr}[T]$, *and no two vertices in* $\gamma^{nq}[T]$ *are separated by a set of the form* $\gamma^{nk}[T]$, $k \in \mathbb{Z}$.

We now apply Halin's theorem to strips.

Lemma 3.3. *Let* Σ *be a strip. Let* $\gamma \in G(\Sigma)$ *be of type* 2. *Then* Σ *contains a maximal set of* 2-*way-infinite,* γ-*invariant paths* $\Pi_1, \ldots, \Pi_k$ *such that the following statements hold*:

(a) $k = \kappa_\infty(\Sigma)$.

(b) *For each* $i = 1, \ldots, k$, Π_i *is the union of two* 1-*way infinite paths, each from one of the two ends of* Σ.

(c) *If* S *is any* κ_∞-*separator and* $\alpha \in G(\Sigma)$, *then each vertex in* $\alpha[S]$ *belongs to a different path* Π_i.

(d) *If* Σ *is edge-transitive, then for any* κ_∞-*separator* S *of* Σ,

(i) S *is an independent set*;

(ii) *If* Σ *is not vertex-transitive, then* Σ *admits a bipartition* $\{V_0, V_1\}$ *such that* $V_0 \subseteq \bigcup_{i=1}^{k} V(\Pi_i)$.

Proof. Let $\Pi_1, \ldots, \Pi_k$ be the paths and T the k-set postulated by Proposition 3.2. Since T separates the two ends of Σ, we have $\kappa_\infty(\Sigma) \leqslant |T| = k$. From this and (3.1), we now infer (a).

By 3.2(b,e), any two 1-way-infinite paths making up a path Π_i belong to different ends. Part (b) is now immediate.

It follows from (b) that since every κ_∞-separator separates the two ends of Σ, every κ_∞-separator of Σ meets every path Π_i. If S is a κ_∞-separator and $\alpha \in G(\Sigma)$, then $\alpha[S]$ is also a κ_∞-separator. By (a), $|\alpha[S]| = k$, and (c) follows.

Now suppose that Σ is edge-transitive and suppose e is an edge with both of its incident vertices in S. Let $e' \in E(\Pi_1)$ and pick $\alpha \in G(\Sigma)$ such that $\alpha[e] = e'$. But then Π_1 would contain two vertices of $\alpha[S]$, contrary to (c). Hence S is an independent set.

Finally, suppose that Σ is edge-transitive but not vertex-transitive. Let S be a κ_∞-separator and let $x_0 \in S$. Now Σ admits a bipartition $\{V_0, V_1\}$ where $x_0 \in V_0$. If $y_0 \in V_0$, then $\alpha(x_0) = y_0$ for some $\alpha \in G(\Sigma)$, and so $y_0 \in \alpha[S]$. By (c), $y_0 \in V(\Pi_i)$ for some $i = 1, \ldots, k$, as required. □

Definition. A 2-way-infinite path Π in a strip Σ will be called *translatable* if it is a member of a $\kappa_\infty(\Sigma)$-set of γ-invariant paths for some type 2 automorphism γ satisfying the conditions of Lemma 3.3.

If the strip Σ is edge-transitive, then clearly for any given edge there exists a translatable path of Σ containing it.

Lemma 3.4. *Let Σ be an edge-transitive or vertex-transitive strip, and let the vertex x lie between vertices w and y on some translatable path Π in Σ. Let $\alpha \in G_x$. If $\alpha(y) = w$, then $\alpha \notin B$.*

Proof. Let Π^+ denote the one-way infinite subpath of Π with initial vertex x that passes through y. Suppose that $\alpha(y) = w$ for some $\alpha \in B_x$. If T is a κ_∞-separator, then T does not separate Π^+ from $\alpha[\Pi^+]$. If $x \in T$, then $\alpha[\Pi^+]$ must meet T at some vertex other than x, contrary to Lemma 3.3(c).

Now suppose x belongs to no κ_∞-separator, and let $\Pi^- = \Pi - \Pi^+$. The neighbor x' of x in Π^- belongs to some κ_∞-separator T. Since T must not separate Π^- from $\alpha[\Pi^-]$, we have that $\alpha[\Pi^-]$ meets T at some vertex other than w, providing a contradiction. □

The author has proved [15, Theorem 5.4] that if a strip is both vertex- and edge-transitive, then it has even valence. The following theorem generalizes that result by showing *inter alia* that the hypothesis of vertex-transitivity may be dropped.

Theorem 3.5. *Let Σ be an edge-transitive strip. Then all vertices of Σ have even valence. More specifically, let T be a κ_∞-separator of Σ, and let F^+ and F^- be the components of $\Sigma - T$. Let $x \in T$.*

(A) *If Σ is not vertex-transitive or if Σ is flag-transitive, then:*
  (1) $[G:B] = 2$;
  (2) *$\partial(x)$ is a G_x-orbit;*
  (3) *$\partial(x) \cap F^+$ and $\partial(x) \cap F^-$ are B_x-orbits of equal size.*

(B) *If Σ is vertex-transitive but not flag-transitive, then either:*
  (1) *$G = B$, in which case $\partial(x) \cap F^+$ and $\partial(x) \cap F^-$ are orbits of G_x of equal size, or*
  (2) *$[G:B] = 2$, in which case:*
    (a) *Each of $\partial(x) \cap F^+$ and $\partial(x) \cap F^-$ comprises exactly two B_x-orbits, all of the same size, implying that $\rho(\Sigma)$ is divisible by 4, and*
    (b) *$\partial(x)$ comprises exactly two G_x-orbits, each being the union of one B_x-orbit in $\partial(x) \cap F^+$ and one B_x-orbit in $\partial(x) \cap F^-$.*

Proof. By Proposition 2.3, $\Sigma - T$ has no finite component. Since $\partial(x) \cap T = \emptyset$ by Lemma 3.3(d), $\{\partial(x) \cap F^+, \partial(x) \cap F^-\}$ is a partition of $\partial(x)$. There exist vertices $w \in \partial(x) \cap F^-$ and $y \in \partial(x) \cap F^+$ such that the edges $\{w, x\}$ and $\{x, y\}$ lie on some translatable path Π. Thus $\{x\} = T \cap V(\Pi)$. By Lemma 3.4, B_x admits at least two orbits in $\partial(x)$, and no orbit of B_x meets both F^+ and F^-. If r denotes the number of G_x-orbits in $\partial(x)$, then the number of B_x-orbits in $\partial(x)$ is r if $[G:B] = 1$ and is r or $2r$ if $[G:B] = 2$ by (2.4).

If Σ is not vertex-transitive, then by Proposition 2.2 G_x acts transitively on $\partial(x)$, even when x does not belong to a κ_∞-separator. Similarly $\partial(x)$ is an orbit of

G_x when Γ is flag-transitive. Hence $\partial(x)$ comprises $[G_x:B_x]=2$ orbits of B_x, namely $\partial(x)\cup F^+$ and $\partial(x)\cap F^-$. Any automorphism in the coset $G_x\backslash B_x$ interchanges these two orbits.

We now assume that Σ is vertex-transitive but not flag-transitive. Thus Σ is bipartite, and by the final paragraph of Section 2, $\rho(\Sigma)$ is even.

Suppose there exists a vertex $z\in\partial(x)\cap F^+$ that belongs to a B_x-orbit different from that of y. Since Σ is edge-transitive, there exists $\sigma\in G$ such that $\sigma(z)=x$ and $\sigma(x)=y$. At least one of y and z does not share a G_x-orbit with w; for definiteness, let us assume that w and y are in different G_x-orbits. Again by edge-transitivity, one may fix $\tau\in G$ such that $\tau(w)=x$ and $\tau(x)=y$. But then $\sigma^{-1}\tau$ fixes x while $\sigma^{-1}\tau(w)=z$, implying that w and z belong to the same G_x-orbit. This orbit must thus contain at least two — and hence exactly two — B_x-orbits, one from $\partial(x)\cap F^+$ and one from $\partial(x)\cap F^-$. By (2.4), $[G:B]=[G_x:B_x]=2$ in this case.

What we have proved is that given any two B_x-orbits in $\partial(x)\cap F^+$, one of them is contained in the G_x-orbit that includes w. Hence $\partial(x)\cap F^+$ comprises only two B_x-orbits O_1^+ and O_2^+. By symmetry, $\partial(x)\cap F^-$ comprises two B_x-orbits O_1^- and O_2^-. If $\gamma\in G_x\backslash B_x$, then γ interchanges O_i^+ with O_i^- $(i=1,2)$. It follows from the final paragraph of Section 2 that $|O_1^+\cup O_2^+|=|O_1^-\cup O_2^-|$. We conclude that the four B_x-orbits in $\partial(x)$ are all of equal size.

The remaining situation is the one in which $\partial(x)\cap F^+$ and $\partial(x)\cap F^-$ are G_x-orbits. It follows that $[G_x:B_x]=1$. That the orbits are of equal size follows from Corollary 2.8.

It remains only to prove that $G=B$. Were this not the case, (2.4) would imply that $[G:G_xB]=2$. For arbitrary $z\in V(\Sigma)$, suppose $\eta_1(x)=\eta_2(x)=z$ for some $\eta_1,\eta_2\in G$. Then $\eta_2^{-1}\eta_1\in G_x\leqslant G_xB=B$, and so η_1 and η_2 belong to the same B-coset. This yields a partition $\{V_0,V_1\}$ of $V(\Sigma)$ whereby a vertex $s\in V_1$ if and only if $\eta(x)=s$ for some $\eta\in G\backslash B$. Thus $x\in V_0$. It is immediate that V_0 and V_1 are each fixed under B and are interchanged by each automorphism in $G\backslash B$. In fact, $\{V_0,V_1\}$ is the bipartition of Σ.

Pick $w\in\partial(x)$ such that $\{w,x\}$ lies on a translatable path Π. Since $w\in V_1$ and Σ is vertex-transitive, there exists $\alpha\in G\backslash B$ such that $\alpha(w)=x$. Let $y=\alpha(x)$. Suppose there existed $\beta\in G_x$ such that $\beta(w)=y$. Then $\beta^{-1}\alpha(x)=w$ and $\beta^{-1}\alpha(w)=x$, although Σ is not flag-transitive. Hence w and y belong to different G_x-orbits: say, $w\in\partial(x)\cap F^+$ and $y\in\partial(x)\cap F^-$. In the notation of the proof of Lemma 3.4, Π^- and $\alpha[\Pi^+]$, belonging to the same end, could not be separated by T. Hence $\alpha[\Pi^+]$ would meet $T\backslash\{x\}$, contrary to Lemma 3.3c. □

In the final paragraph of [13], Thomassen and the author describe a vertex- and edge-transitive strip that is not flag-transitive. It satisfies $\rho=\kappa_f=4$ and $\kappa_\infty=9$ and exemplilfies subcase B1 of Theorem 3.5. This strip has girth 6, which implies by Theorem 4.2 below that it is nonplanar.

A strip Σ exemplifying the conditions of Subcase B2 may be constructed as follows. In [2], Folkman presented a family of finite edge-transitive nonvertex-transitive graphs of constant valence. The smallest member of this family is a 4-valent graph having 20 vertices; let Γ denote this graph. Following Proposition 2.2, let $\{A, B\}$ denote the bipartition of $V(\Gamma)$. Let $A = \{a_1, \ldots, a_{10}\}$ and $B = \{b_1, \ldots, b_{10}\}$. Note that A and B are the two vertex-orbits of $G(\Gamma)$. Now let S be the set $\{x_1, \ldots, x_{10}, y_1, \ldots, y_{10}\}$ and let

$$V(\Sigma) = S \times \mathbb{Z},$$

and for each edge $\{a_i, b_j\} \in E(\Gamma)$ we have for each $n \in \mathbb{Z}$ the following four edges in Σ:

$$\{(x_i, n), (x_j, n+1)\},$$
$$\{(y_i, n), (y_j, n+1)\},$$
$$\{(x_i, n), (y_j, n-1)\},$$
$$\{(y_i, n), (x_j, n-1)\}.$$

Thus $\rho(\Sigma) = 4\rho(\Gamma) = 16$. One verifies straightforwardly that Σ has the desired properties.

4. Planar edge-transitive strips

Let Σ be a planar strip. It is not hard to see that Σ admits an imbedding on an infinite cylinder in such a way that its two ends are separated by every circumference of the cylinder. Consider an automorphism $\alpha \in G(\Sigma)$ like the one in the definition of 'strip' in Section 2. Some properties of α are that $\alpha \in B(\Sigma)$ and that α has infinite order and a finite number m of orbits, all of them infinite. For any given $d_0 \in \mathbb{Z}^+$, we can pick $n \in \mathbb{Z}^+$ such that $d(x, \alpha^n(x)) \geqslant d_0$ for all $x \in V(\Sigma)$. For $x, y \in V(\Sigma)$ write $x \cong y$ if $y = \alpha^{kn}(x)$ for some $k \in \mathbb{Z}$. Obviously $\cong$ is an equivalence relation. The quotient graph $\Sigma/\cong$ is finite, having mn vertices, and is imbeddable on the surface of a torus. Hence $\rho_{\min}(\Sigma/\cong) \leqslant 6$. By choosing d_0 sufficiently large, we can be assured that for all $x \in V(\Sigma)$, $\rho(x)$ in Σ equals $\rho([x])$ in $\Sigma/\cong$, where $[x]$ denotes the equivalence class containing x. We proved the following.

Lemma 4.1. *Let Σ be a planar strip. Then $\rho_{\min}(\Sigma) \leqslant 6$.*

Theorem 4.2. *Let Σ be a planar, edge-transitive strip. Then Σ has the following characterization:*

(A) *If $\kappa(\Sigma) \geqslant 3$, let $b \geqslant 3$, and let*

$$W = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : 0 \leqslant x + y \leqslant 2b\}$$

where $(-x, x)$ *has been identified with* $(-x+b, x+b)$ *for all* $x \in \mathbb{Z}$. *Let* $V((\Sigma) = W$, *and when* $0 \leqslant x+y < 2b$, *let* (x, y) *be adjacent to* $(x+1, y)$ *and* $(x, y+1)$. *In this case,* Σ *is flag-transitive,* $\rho = \rho^* = \kappa_f = 4$, $\kappa_\infty = b$, *and* $\kappa = \min\{4, b\}$.

(**B**) *If* $\kappa(\Sigma) = 2$, *then*:

(1) $\Sigma = \Delta_\infty[\Omega_2]$, *or*

(2) Σ *is obtained from one of the graphs in* (**A**) *or from* $\Delta_\infty[\Omega_2]$ *by constructing the multigraph wherein each edge is replicated* $r \geqslant 1$ *times and subdividing each edge once.*

(**C**) *If* $\kappa(\Sigma) = 1$, *then*:

(1) Σ *is a 2-way-infinite path, or*

(2) *for some given* $m \geqslant 2$, *all the lobes of* Σ *are isomorphic to* $K_{2,m}$ *and all the separating vertices are incident with exactly two lobes and belong to the 2-vertex side of each copy of* $K_{2,m}$.

Proof. Let Σ be a planar, edge-transitive strip.

(**A**) Suppose that $\kappa(\Sigma) \geqslant 3$. Thus the imbedding of Σ in the plane is essentially unique; we will be able to speak of the 'regions' of Σ as well as of its dual. (See [12] for an extension to infinite graphs of Whitney's classical result [16] concerning the uniqueness of planar imbeddings of finite graphs.)

We first show that Σ is 4-valent. Let $x \in V(\Sigma)$ be given.

Subcase 1: $x \in T$ *for some* κ_∞*-separator* T.

By Theorem 3.5, $\rho(x)$ is even. If $\rho(x) \neq 4$, then $\rho(x) \geqslant 6$ and we may write $\partial(x) \cap F^+ = \{u_1, \ldots, u_m\}$ and $\partial(x) \cap F^- = \{u_{m+1}, \ldots, u_{2m}\}$, where $m \geqslant 3$ and F^+ and F^- are the vertex sets of the two infinite components of $\Sigma - T$. By Lemma 2.5 we may assume that the elements of $\partial(x)$ have been indexed in such a way that as one proceeds in counterclockwise fashion around x, one encounters in cyclic order the incident edges $\{x, u_1\}, \{x, u_2\}, \ldots, \{x, u_{2m}\}$.

Note that the restriction of G_x (and hence of B_x) to $\partial(x)$ is a subpermutation group of the full dihedral group on the cyclic list of symbols $u_1, \ldots, u_{2m}$. Suppose for some $\beta \in B_x$ that $\beta(u_1) = u_2$. By Theorem 3.5, $\beta(u_{2m}) \neq u_1$. This forces $\beta(u_2) = u_1$ and, in turn, $\beta(u_3) = u_{2m}$, contrary to Theorem 3.5. Hence u_1 and u_2 belong to different B_x-orbits, and so Σ is subject to the conditions of Case B2 of Theorem 3.5, whence Σ is vertex-transitive and so $\rho(x) = \rho_{\min}(\Gamma)$, a number divisible by 4. By Lemma 4.1 together with our initial assumption in this subcase, $\rho_{\min}(\Gamma) \leqslant 6 \leqslant \rho(x)$, giving a contradiction.

Subcase 2: x *belongs to no* κ_∞*-separator.*

In this case Σ is not vertex-transitive; Σ is bipartite, and every neighbor of x belongs to a κ_∞-separator. By edge-transitivity, some translatable path Π passes through x. By Lemma 3.3, x has two neighbors $t_0, t_1 \in V(\Pi)$ and no κ_∞-separator contains both t_0 and t_1. Let T_0 be a κ_∞-separator that contains t_0. Since $t_1 = \gamma_1(t_0)$ for some $\gamma_1 \in G_x \setminus B_x$, we let $T_1 = \gamma_1[T_0]$. For $i = 0, 1$, let F_i be the vertex set of

the component of $\Sigma - T_i$ that contains x. Since Π cannot 'recross' T_0 or T_1, x is the only vertex of Π in $F_0 \cap F_1$. Since Σ has only two ends, it follows that $F_0 \cap F_1$ is finite. We may assume that T_0 has been chosen so as to minimize $|F_0 \cap F_1|$.

We claim that $\partial(x) \subseteq T_0 \cup T_1$. Otherwise let $t_2 \in \partial(x) \setminus (T_0 \cup T_1)$. For some $i = 0, 1$, $\gamma_2(t_i) = t_2$ for some $\gamma_2 \in B_x$. Let $T_2 = \gamma_2[T_i]$ and let F_2 be the component of $\Sigma - T_2$ that contains x. Thus $F^* = F_i \cap F_2$ is infinite. By the proof of [8, Theorem 2], $\partial(F^*)$ is a κ_∞-separator and $\partial(F^*) = (F_i \cap T_2) \cup (T_i \cap T_2) \cup (T_i \cap F_2)$. For $j \in \{0, 1\} \setminus \{i\}$, $|F_j \cap F^*| < |F_0 \cap F_1|$ since $t_2 \notin V(F^*)$. But $x \in F_j \cap F^*$, contrary to the minimality assumption.

Were x to have a neighbor $t \in T_0 \cap T_1$, then a translatable path through $\{x, t\}$ would meet T_0 or T_1 twice. Hence $T_0 \cap T_1 = \emptyset$. For $\alpha \in G_x$, $\alpha(t_0) \in T_0$ if and only if $\alpha \in B_x$. It follows that $\partial(x) \cap T_0$ and $\partial(x) \cap T_1$ are the two B_x-orbits in $\partial(x)$ and hence are of equal size.

[We remark that thus far in the present subcase, no assumption about planarity, connectivity, or valence has yet been used. We note this because the foregoing argument will be cited below in the proof of Part (B) of this theorem.]

Letting $m = |\partial(x) \cap T_0|$, we have that $2m = \rho(x) \geqslant \kappa(\Sigma) \geqslant 3$, and so $m \geqslant 2$. We observe that $\Sigma - (\partial(T_0) \cap F_0)$ has two infinite components and that $\partial(x) \cap T_1 \subseteq \partial(\partial(T_0) \cap F_0) \cap F_0$. Thus we may repeat the argument of Subcase 1 by replacing T, $\partial(x) \cap F^-$, and $\partial(x) \cap F^+$ by $F_0 \cap F_1$, T_0, and T_1, respectively. The argument this time is abbreviated by the fact that $\partial(x)$ is a G_x-orbit. We readily deduce that $m \leqslant 2$. Thus Σ is a 4-valent graph.

Now consider the dual graph Σ^*, which must similarly be 3-connected and edge-transitive. Also, Σ and Σ^* have the same number of ends, i.e., Σ^* is a strip. Applying the foregoing argument to Σ^*, we obtain that Σ^* is 4-valent. Thus Σ is 4-covalent. In fact, Σ must be bipartite, whether it is vertex-transitive or not.

Let us now investigate the structure of Σ. Let $x_1 \in V(\Sigma)$ and let T be a κ_∞-separator such that $x_1 \in T$. Let F^+ and F^- denote the vertex sets of the two components of $\Sigma - T$. Let $T' = \partial(T) \cap F^+$. In the event that Σ is not vertex-transitive, an enumeration of the $2\kappa_\infty$ TT'-edges proves that $|T| = |T'|$, and so every vertex belongs to a κ_∞-separator. Proceeding counterclockwise about x_1, let its neighbors encountered in cyclic order be u_1, v_1, v_2, u_2, where $u_1, u_2 \in F^+$ and $v_1, v_2 \in F^-$. The edges $\{x_1, u_2\}$ and $\{x_1, v_2\}$ are two of the edges on the boundary of a region R_1. There exists a vertex x_2 such that $\{x_1, u_2, x_2, v_2\}$ determines R_1. Since T separates u_2 from v_2, it follows that $x_2 \in T$. There are vertices $u_3 \in V(F^+)$, $v_3 \in V(F^-)$ such that the neighbors of x_2 in counterclockwise cyclic order are u_2, v_2, v_3, u_3. If $u_3 = u_1$, then $\{u_1, x_1, v_2, x_2, u_1\}$ determines a circuit that separates u_2 from other vertices such as v_1, and so $\{x_1, x_2\}$ is a separating set although $\kappa(\Sigma) > 2$. Hence u_1, u_2, u_3 are distinct, and similarly so are v_1, v_2, v_3. Now the edges $\{x_2, u_3\}$ and $\{x_2, v_3\}$ lie on the boundary of a region R_2, whose fourth boundary vertex is a vertex $x_3 \in T$. If $x_3 = x_1$, then $\{x_1, x_2\}$ would be a κ_∞-separator. Hence x_1, x_2, x_3 are distinct.

Continuing in this way, the process terminates with $T = \{x_1, x_2, \ldots, x_b\}$,

where $b = \kappa_\infty$ and the neighbors of x_i in counterclockwise cyclic order are u_i, v_i, v_{i+1}, u_{i+1}, where subscripts are understood modulo b.

Passing to Σ^*, we see that $T^* = \{R_1, \ldots, R_b\}$ is a κ_∞-separator. One constructs the regions of Σ^* adjacent to those in T^*, then passes back to $\Sigma^{**} = \Sigma$. In this way, all of Σ is forced. (Note that Σ is isomorphic to Σ^*. See Fig. 1 for a cylindrical representation of the unique edge-transitive planar strip with connectivity 3.)

(B) Suppose that $\kappa(\Sigma) = 2$.

Case 1: $\kappa_f(\Sigma) > 2$.

Thus $\kappa_\infty = 2$. Since all valences are even, all vertices have valence at least 4. Let x be any vertex adjacent to a κ_∞-separator of Σ. As noted in the proof of Subcase 2 above, there exist κ_∞-separators T_0 and T_1 of Σ such that $\partial(x) \cap T_0$ and $\partial(x) \cap T_1$ are the two orbits of B_x in $\partial(x)$. Since $|T_0| = |T_1| = 2$ and $|\partial(x)| \geqslant 4$, we have that $\rho(x) = 4$ and $\partial(x) = T_0 \cup T_1$.

Let $T_0 = \{w_1, w_2\}$ and let F_i be the vertex set of the component of $\Sigma - T_i$ that contains x, $(i = 0, 1)$.

We assert that if $z \in F_0$, then $\{w_1, z\}$ is not a κ_∞-separator. By Lemma 3.3(d), $z \notin \partial(T_0)$, and clearly $z \notin V(\Sigma) \backslash (T_1 \cup F_1)$. But then some translatable path passes through w_2, x, and one vertex in T_1, thus avoiding $\{w_1, z\}$. Similarly, $\{w_2, z\}$ is not a κ_∞-separator.

Suppose there were a vertex $x' \in \partial(w_1) \cap F_0$ but $x' \notin \partial(w_2)$. There exists $\beta \in B_{w_1}$ such that $\beta(x) = x'$. But $\beta(w_2) \neq w_2$ and so $\beta[T_0]$ is not a κ_∞-separator. We infer that every vertex in $T' = \partial(T_0) \cap F_0$ is adjacent to both w_1 and w_2.

If $|T'| > 2$, then one can construct a homeomorph of $K_{3,3}$ in Σ, the nodes of which are w_1, w_2, three of the vertices in T' and one vertex in $F_0 \backslash T'$, contrary to the planarity of Σ. Hence $|T'| = 2$, and it is now routine to show that the structure of Σ is forced; $\Sigma = \Delta_\infty[\Omega_2]$, the copies of Ω_2 being the κ_∞-separators.

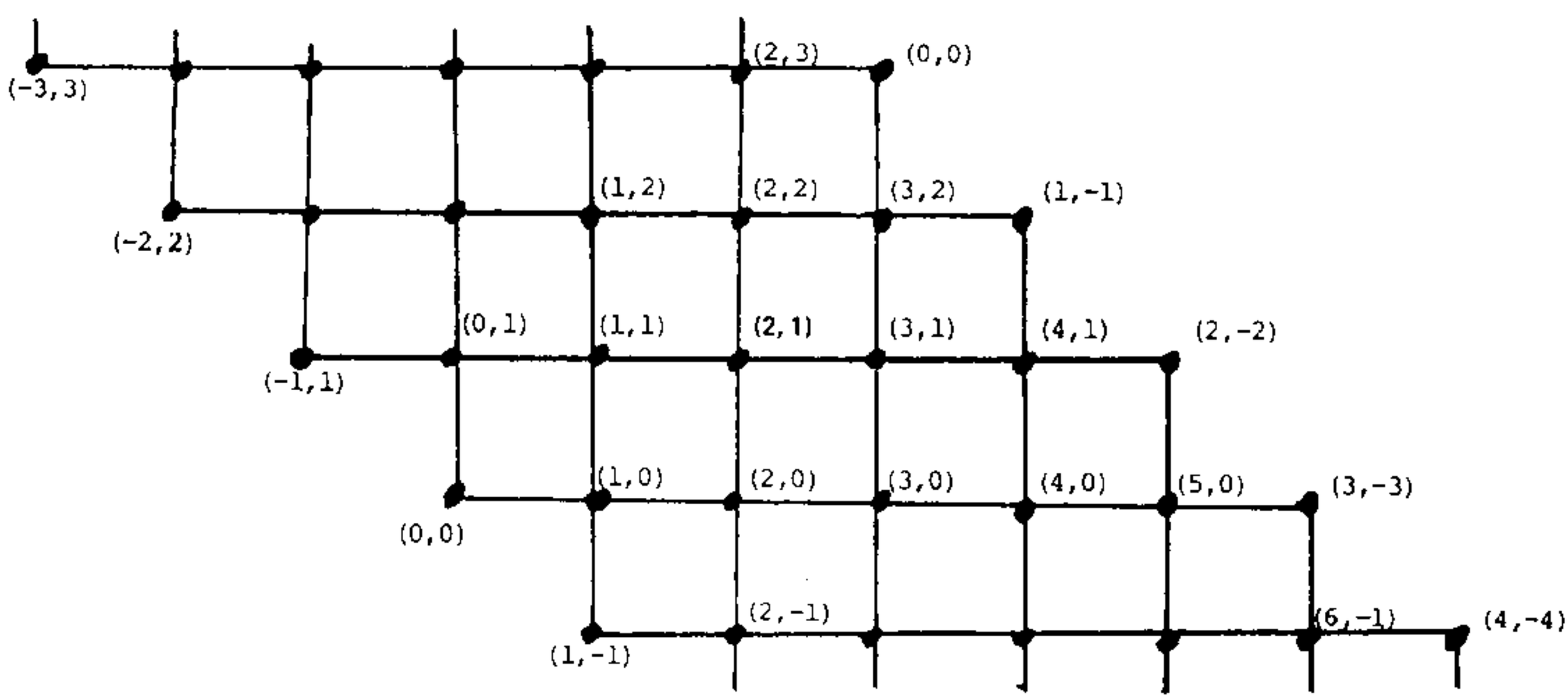


Fig. 1. Edge-transitive strip with $\kappa = 3$.

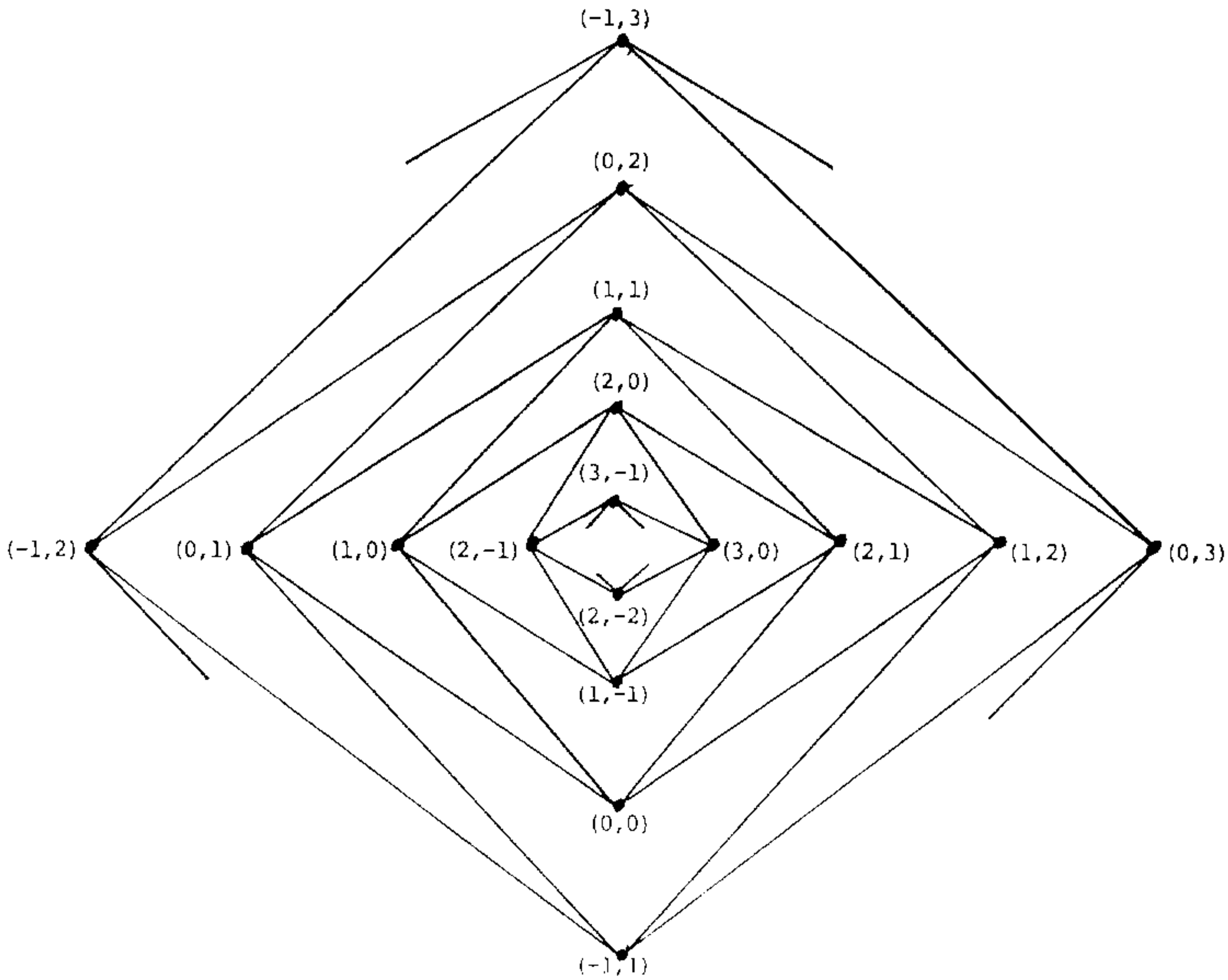


Fig. 2. $\Delta_\infty[\Omega_2]$.

[We note that were we to let $b=2$ in part (A), the same graph $\Delta_\infty[\Omega_2]$ would be produced. See Fig. 2 for a planar representation of this graph.]

Case 2: $\kappa_f(\Sigma)=2$.

Thus $\kappa_\infty \geqslant 2$. By Proposition 2.1, Σ admits 2-valent vertices, although Σ is not itself 2-valent. The situation is that Σ is not vertex-transitive; it admits a bipartition $\{V_0, V_1\}$ where V_1 is the set of all 2-valent vertices. Consider the strip Ψ with $V(\Psi)=V_0$ and $E(\Psi)$ consisting of pairs of vertices of V_0 having a common neighbor in Σ. Since $\kappa_f(\Psi)=\rho_{\min}(\Psi)>2$ and Ψ is edge-transitive and planar, either $\Psi=\Delta_\infty[\Omega_2]$ or Ψ has been characterized in part (A). By replicating each edge of Ψ some fixed number of times and then subdividing all the edges once, one recovers the original strip Σ.

(C) Suppose $\kappa(\Sigma)=1$. Since Σ admits a separating vertex, every edge is incident with a separating vertex. Since Σ is locally finite, it admits infinitely many separating vertices and infinitely many lobes.

Every lobe Λ of Σ is itself an edge-transitive graph, for if $e, e' \in E(\Lambda)$, we have $\alpha(e)=e'$ for some $\alpha \in G(\Sigma)$, whence $\alpha[\Lambda]=\Lambda$ and so $\alpha_{|\Lambda} \in G(\Lambda)$. By extension, all the lobes of Σ are pairwise isomorphic.

If a lobe Λ were incident with three or more separating vertices of Σ, then the

deletion of any three of these vertices would leave at least three infinite components. On the other hand, edge-transitivity precludes that Λ contain a single separating vertex of Σ. Hence each lobe of Σ is incident with exactly two separating vertices. If some edge of Λ is incident with both of the separating vertices in Λ, then $\Lambda = K_2$ and Σ is a 2-way-infinite path.

The situation remains in which each edge of Λ is incident with one separating vertex of Σ. The two separating vertices in Λ belong to the same orbit of G and hence of $G_{|\Lambda}$. It follows that Λ is bipartite and hence is a subgraph of $K_{2,m}$ for some m. By definition of a lobe, Λ is biconnected, and so $\Lambda = K_{2,m}$ for some $m \geqslant 2$. [The graph defined in part (A) with $b = 1$ is precisely this graph when $m = 2$.] The proof is now complete. □

References

[1] G.A. Dirac and S. Schuster, A theorem of Kuratowski, Indag. Math. 16 (1954) 343–348.
[2] J. Folkman, Regular line-symmetric graphs, J. Combin. Theory 3 (1967) 215–232.
[3] J.E. Graver and M.E. Watkins, Combinatorics with Emphasis on the Theory of Graphs (Springer, Berlin, 1977).
[4] G. Grünbaum and G.C. Shephard, Edge-transitive planar graphs, J. Graph Theory 11 (1987) 141–155.
[5] R. Halin, Über unendliche Wegen in Graphen, Math. Ann. 157 (1964) 125–137.
[6] R. Halin, Automorphisms and endomorphisms of infinite locally finite graphs, Abh. Math. Sem. Univ. Hamburg 39 (1973) 251–283.
[7] H.A. Jung, A note on fragments of infinite graphs, Combinatorica 1 (1981) 285–288.
[8] H.A. Jung and M.E. Watkins, On the connectivities of finite and infinite graphs, Monatsh. Math. 83 (1977) 121–131.
[9] H.A. Jung and M.E. Watkins, Fragments and automorphisms of infinite graphs, European J. Combin. 5 (1984) 149–162.
[10] H.A. Jung and M.E. Watkins, Finite separating sets in locally finite graphs, submitted.
[11] G. Sabidussi, The lexicographic product of graphs, Duke Math. J. 28 (1961) 573–578.
[12] C. Thomassen, Infinite graphs, in: L.W. Beineke and R.J. Wilson, eds., Selected Topics in Graph Theory 2 (Academic Press, London, 1983) 129–160.
[13] C. Thomassen and M.E. Watkins, Infinite vertex-transitive, edge-transitive, non-1-transitive graphs, Proc. Amer. Math. Soc. 105 (1989) 258–261.
[14] M.E. Watkins, Connectivity of transitive graphs, J. Combin. Theory 8 (1970) 23–29.
[15] M.E. Watkins, Vertex-transitive graphs that are not Cayley graphs, in: G. Hahn, G. Sabidussi and R. Woodrow, eds., Cycles and Rays, NATO Advanced Studies Institute, Ser. C, (Kluwer Academic Publishers, Dordrecht, 1990) 243–256.
[16] H. Whitney, Non-separable and planar graphs, Trans. Amer. Math. Soc. 34 (1932), 339–362.

Discrete Mathematics 95 (1991) 373–384
North-Holland

Topological groups and infinite graphs

Wolfgang Woess

Dipartimento di Matematica, Universitá di Milano, Via C. Saldini, 50, 20133 Milano, Italy

Received 17 November 1989
Revised 20 March 1990

Abstract

Woess, W., Topological groups and infinite graphs, Discrete Mathematics 95 (1991) 373–384.

We show how results concerning infinite, locally finite, vertex-symmetric graphs can be related with the structure theory of topological groups, when the latter is applied to automorphism groups of the graphs. In particular, we discuss polynomial growth, bounded automorphisms and infinite expanders. In an appendix, we present three problems on infinite graphs, not necessarily linked with topological considerations.

1. Introduction

The automorphism group of an infinite, locally finite graph is a topological group with the topology of pointwise convergence. In this note we show how the structure theory of topological groups can be applied to prove results concerning vertex-symmetric graphs. In particular, we discuss a series of results due to Trofimov [19–21].

After presenting preliminaries and basic facts (Section 2), we show in Section 3 how Trofimov's theorem on graphs with polynomial growth [20] can be easily proved by combining Gromov's famous theorem on discrete groups with polynomial growth [6] with a result of Losert [11] concerning topological groups with this property. In Section 4, we discuss the connection of another theorem of Trofimov [19], concerning lattice-type graphs, with the theory of FC^--groups of Grosser and Moskowitz [7]. However, deriving Trofimov's characterization of lattices directly from the results of [7] includes a difficulty which seems hard to overcome. On the other hand, one obtains from [19] that the group of bounded automorphisms of a vertex-symmetric graph is closed [21] and hence an FC^--group. Applying [7] we get that the bounded automorphisms of finite order constitute a closed group in the topology of pointwise convergence: this completes partial results by Godsil et al. [5].

0012-365X/91/$03.50 © 1991 — Elsevier Science Publishers B.V. All rights reserved

Finally in Section 5, we report a result of Soardi and Woess [16] which characterizes infinite vertex-symmetric expanders in terms of properties of the automorphism group, viewed as a topological group. The properties in question are amenability and unimodularity, and the result can be applied to show that every vertex-symmetric graph with infinitely many ends is an expander.

It should be pointed out that the spirit of this note is mainly that of an introductory survey with the aim to illustrate the use of topological groups in the study of infinite graphs. As this paper is principally addressed to readers working in the field of discrete mathematics, some space is given to the explanation of the relevant topological prerequisites.

2. Basic facts

In the sequel, $\Gamma(X, E)$ will always denote an infinite, connected, locally finite graph with vertex set X and (unoriented) edge set E. The *automorphism group* $\mathrm{AUT}(\Gamma)$ of Γ is the group of isometries of the vertex set X with respect to the discrete metric d: $d(x, y)$ is the smallest number of edges on a path in Γ connecting x and y. We shall always assume that Γ is *vertex-symmetric*, i.e., $\mathrm{AUT}(\Gamma)$ acts transitively on the vertex set.

We introduce *pointwise convergence* of a sequence (g_n) in $\mathrm{AUT}(\Gamma)$:

$$g_n \to g \in \mathrm{AUT}(\Gamma), \text{ if for every } x \in X,\ g_n x = gx \text{ for all } n \geqslant n_x.$$

Recall that a *topological group* is a group G equipped with a topology such that the maps $(g, h) \mapsto gh$ and $g \mapsto g^{-1}$ are continuous on $G \times G$ and on G, respectively. A good introduction to the fundamentals of topological groups is given, for example, by Hewitt and Ross [10]. The topology of a topological group is completely determined by a neighbourhood base at the identity, see [10, (4.5)]. For the topology of pointwise convergence in $G = \mathrm{AUT}(\Gamma)$, a neighbourhood base at the identity is given by the family of pointwise stabilizers of finite subsets of X. We write G_x for the stabilizer of x. In other words, if we fix a reference vertex o and, for every $x \in X$, an automorphism g_x such that $g_x o = x$ (here we use transitivity!), then the family of subgroups

$$g_x\, \mathrm{AUT}(\Gamma)_o\, g_y^{-1}, \quad x, y \in X$$

is a subbasis of the topology. Thus, $\mathrm{AUT}(\Gamma)$ is a Hausdorff group with countable base, and it is locally compact, as the following well-known lemma shows.

Lemma 1. *The stabilizer* $\mathrm{AUT}(\Gamma)_x$ *is compact.*

Proof. Let (g_n) be a sequence in $\mathrm{AUT}(\Gamma)_x$, and let $\{x_0 = x, x_1, x_2, \ldots\}$ be an enumeration of X. As $g_n x = x$ for every n, and as Γ is locally finite and connected, the set $\{g_n x_k \mid n \geqslant 0\}$ is finite for every k. Hence there is a

subsequence $(\tau_1(n))$ of (n) such that all $g_{\tau_1(n)}x_1$ coincide; write gx_1 for this common image. Repeating this argument inductively, we get a sub-subsequence $(\tau_k(n))$ of the preceding subsequence $(\tau_{k-1}(n))$, such that all $g_{\tau_k(n)}$, $n \geqslant 0$, send x_k to the same element of X, denoted gx_k. Thus, $g_{\tau_n(n)} \to g \in \mathrm{AUT}(\Gamma)$ pointwise. □

We now see that the identity has a neighbourhood base consisting of compact-open subgroups, so that $\mathrm{AUT}(\Gamma)$ is totally disconnected. Next, we describe compactness in $\mathrm{AUT}(\Gamma)$.

Lemma 2. *A subset U of $\mathrm{AUT}(\Gamma)$ has compact closure if and only if the orbit Ux is finite for every $x \in X$.*

Proof. If U is relatively compact then it is contained in a finite union of sets $g_i\,\mathrm{AUT}(\Gamma)_x$, $i = 1, \ldots, r$. Hence, $Ux \subset \{g_i x \mid i = 1, \ldots, r\}$.

Conversely, if $Ux = \{y_1, \ldots, y_r\}$, then there are $g_i \in U$ such that $g_i x = y_i$, $i = 1, \ldots, r$. But then

$$U \subset \bigcup_{i=1}^{r} g_i\,\mathrm{AUT}(\Gamma)_x.$$

The latter set is compact by Lemma 1, so that A has compact closure. □

Indeed, by local finiteness, in Lemma 2 it is enough to check finiteness of Ux for one $x \in X$. Now let G be a closed subgroup of $\mathrm{AUT}(\Gamma)$ acting transitively on X. In the relative topology, G inherits all properties of $\mathrm{AUT}(\Gamma)$ discussed so far. Choose a reference vertex $o \in X$, and define

$$V = \{g \in G \mid d(go, o) \leqslant 1\}.$$

For the following lemma, recall that $V^n = \{g_1 g_2 \cdots g_n \mid g_i \in V\}$.

Lemma 3. *V is a compact, symmetric neighbourhood of the identity in G. If $g \in G$ and $n \geqslant 0$, then $g \in V^n$ if and only if $d(go, o) \leqslant n$.*

Proof. Clearly, $V = V^{-1}$. For every $x \in X$ with $d(x, o) \leqslant 1$ choose $g_x \in G$ with $g_x o = x$. Then

$$V = \bigcup_{d(x,o)\leqslant 1} g_x G_o,$$

and V is compact by local finiteness and Lemma 2. By definition of the topology, V is also open.

The last statement is true for $n = 1$ by definition of V. Suppose it is true for n. Observe that $V^n \subset V^{n+1}$. Let $g \in G$, $go = y$ with $d(y, o) = n + 1$. Then there is a neighbour w of y such that $d(w, o) = n$. By transitivity, $ho = w$ for some $h \in G$,

and by the induction hypothesis, $h \in V^n$. Now, $d(h^{-1}go, o) = d(w, y) = 1$, so that $h^{-1}g \in V$ and $g \in hV \subset V^{n+1}$. Conversely, if $g \in V_{n+1}$, then obviously $d(go, o) \leqslant n+1$. $\square$

3. Graphs with polynomial growth

A famous theorem of Gromov [6] (in combination with a result of Wolf [27]) states that a finitely generated discrete group has polynomial growth if and only if it has a nilpotent subgroup of finite index. If one views a finitely generated group in terms of its Cayley graph with respect to some finite symmetric set of generators, then the question naturally arises if and how Gromov's Theorem generalizes to an arbitrary locally finite, vertex-symmetric graph Γ with polynomial growth. Recall that Γ is said to have *polynomial growth* if

$$\beta(n) = |\{x \in X \mid d(x, o) \leqslant n\}| \leqslant Cn^d$$

for some finite $C, d > 0$. The answer to the above question is as follows.

Theorem 1 [20, Theorem 2]. *Let $\Gamma(X, E)$ have polynomial growth, and let G be a group of automorphisms of Γ which acts transitively on X. Then there is an imprimitivity system σ of G on X with finite blocks, such that G^σ is a finitely generated nilpotent-by-finite group with finite vertex stabilizers on Γ^σ.*

For the understanding of this remarkable theorem, recall that an *imprimitivity system* of G on X is a partition of X into subsets called *blocks*, which is preserved by the action of G. By x^σ we denote the block of $x \in X$. The factor graph Γ^σ has vertex set $X^\sigma = \{x^\sigma \mid x \in X\}$ and edge set $E^\sigma = \{[x^\sigma, y^\sigma] \mid [x, y] \in E\}$. In other words, two blocks are adjacent in the factor graph if and only if they have some pair of representatives which are adjacent in Γ. The induced action of G gives rise to a homomorphic image G^σ in $\mathrm{AUT}(\Gamma^\sigma)$, which acts transitively on X^σ.

For a detailed survey on graphs with polynomial growth, see the article by Imrich and Seifter in this volume.

The purpose of this section is to provide a new, short proof of Theorem 1, using a result of Losert [11] concerning *topological* groups with polynomial growth. To do so, we need more preliminaries.

If G is any locally compact, Hausdorff topological group, then it carries a *left Haar measure*. This is a sigma-additive (Radon) measure λ defined on the Borel sigma-algebra of G (*i.e.*, the sigma-algebra generated by the open sets), whose important features are the following:

(1) $\lambda(K) < \infty$ if $K \subset G$ is compact,
(2) $\lambda(U) > 0$ if $U \subset G$ is open,
(3) $\lambda(gU) = \lambda(U)$ for every Borel set $U \subset G$ and every $g \in G$, and
(4) $\lambda(G) < \infty$ if and only if G is compact.

Up to multiplication with a positive constant, λ is unique. Once more, a good introduction and all relevant properties of Haar measure can be found in [10].

A *compactly generated* group G is said to have polynomial growth, if for some (equivalently, every) compact symmetric neighbourhood V of the identity,

$$\lambda(V^n) \leqslant Cn^d.$$

In the special case when G is a finitely generated discrete group, λ is the counting measure, and if V is chosen to be a finite symmetric set of generators containing the identity, then $\lambda(V^n)$ is the size of the n-ball in the corresponding Cayley graph. Thus, the above definition is the direct extension to topological groups of the notion of polynomial growth. If G is the abelian group $\mathbf{R}^d$, then λ is the Lebesgue measure in the corresponding dimension, and the growth of G is polynomial with degree d.

In a significant extension of Gromov's Theorem, Losert [11] describes completely the structure of topological groups with polynomial growth. As an initial step, the following important proposition is proved.

Proposition 1 [11]. *If G is a locally compact Hausdorff group with polynomial growth, then there is a compact normal subgroup K of G such that G/K is a Lie group.*

The topology on G/K is of course the one induced by the natural projection. As K is compact, G/K has the same polynomial growth as G. The *Lie groups* are a very important class of topological groups. In this paper, all we need to know is that every Lie group is locally homeomorphic with d-dimensional Euclidean space, $d \geqslant 0$. We emphasize that in Proposition 1 the dimension d may be zero; the corresponding 0-dimensional space has one point, so that G is discrete in this case. For the fundamentals concerning Lie groups, see e.g. Varadarajan [22].

We can use Proposition 1 to give a short and transparent proof of Theorem 1.

Proof of Theorem 1. If the Theorem is true for the closure of G in $\mathrm{AUT}(\Gamma)$ then it is true for G. Hence, we may assume that G is closed and thus locally compact. As the stabilizer G_o of our reference vertex o is open-compact, we may normalize the Haar measure of G such that $\lambda(G_o) = 1$.

Let V be as in Lemma 3. Once more, for every $x \in X$ choose $g_x \in G$ with $g_x o = x$. Then, by Lemma 3,

$$V^n = \bigcup_{d(x,o) \leqslant n} g_x G_o,$$

a disjoint union, and left invariance of λ yields $\lambda(V^n) = \beta(n)$ for the growth coefficients $\beta(n)$ of Γ. Thus, the group G has polynomial growth. By Proposition 1, G/K is a Lie group for some compact normal subgroup K of G. By Lemma 2,

the imprimitivity system

$$\sigma = \{Kx \mid x \in X\}$$

has finite blocks, and $G^\sigma = G/K$ is a closed vertex-transitive subgroup of $\mathrm{AUT}(\Gamma^\sigma)$. Hence, in view of the properties of the automorphism group described in Section 2, G^σ is a compactly generated, totally disconnected Lie group: it must be zero-dimensional. In other words, G^σ is a finitely generated, discrete group with polynomial growth. By [6], it is nilpotent-by-finite.

Furthermore, by Lemma 2, $H = \{g \in G \mid go \in Ko\}$ is a compact subgroup of G. The stabilizer of σ^σ is $H^\sigma = H/K$. This is a compact subgroup of G^σ and as such must be finite. □

At this point we remark that Trofimov did not have the result of [11] (which was published later) at his disposal. We also remark that Trofimov uses the above Theorem 1 to deduce a slightly stronger result [20, Theorem 1]: working with $G = \mathrm{AUT}(\Gamma)$, the imprimitivity system σ can be constructed such that not only $\mathrm{AUT}(\Gamma)^\sigma$, but even the—possibly larger—group $\mathrm{AUT}(\Gamma^\sigma)$ is discrete, nilpotent-by-finite and has finite vertex stabilizers on Γ^σ. An important step in this deduction is [20, Proposition 2.3], see below.

4. Bounded automorphisms and FC^--groups

An automorphism g of Γ is called *bounded*, if there is a constant $M = M(g) < \infty$ such that

$$d(gx, x) \leqslant M \quad \text{for every } x \in X.$$

It is obvious that the bounded automorphisms constitute a normal subgroup of $\mathrm{AUT}(\Gamma)$, denoted by $B(\Gamma)$.

An element g in a topological group G is called an FC^--element, if its conjugacy class has compact closure: $\{g^h \mid h \in G\}^-$ is compact. (For a set A, A^- denotes its closure.) Recall that $g^h = h^{-1}gh$. An FC^--*group* is a group consisting of FC^--elements only.

Lemma 4. *$g \in B(\Gamma)$ if and only if g is an FC^--element of $\mathrm{AUT}(\Gamma)$.*

Proof. Let $g \in B(\Gamma)$, and let $h \in \mathrm{AUT}(\Gamma)$. Then $d(g^h x, x) = d(ghx, hx) \leqslant M$ for every $x \in X$. Thus,

$$\{g^h x \mid h \in \mathrm{AUT}(\Gamma)\} \subset \{y \in X \mid d(y, x) \leqslant M\}.$$

By local finiteness, the latter set is finite, so that the conjugate class of g in $\mathrm{AUT}(\Gamma)$ is compact by Lemma 2.

Conversely, suppose that $\{g^h \mid h \in \mathrm{AUT}(\Gamma)\}$ has compact closure. By Lemma 2, $\{g^h o \mid h \in \mathrm{AUT}(\Gamma)\}$ is finite. Hence, there is $M \leqslant \infty$ such that $d(g^h o, o) \leqslant M$

for every $h \in \mathrm{AUT}(\Gamma)$. If $x \in X$, choose $h \in \mathrm{AUT}(\Gamma)$ with $ho = x$. Then

$$d(gx, x) = d(gho, ho) \leqslant M. \qquad \square$$

In his remarkable series of papers, Trofimov obtains the following two results.

Theorem 2 [19]. *$B(\Gamma)$ acts transitively on X if and only if there is an imprimitivity system σ of $B(\Gamma)$ on X, with finite blocks, such that $B(\Gamma)^\sigma$ is a free finitely generated abelian group.*

Theorem 3 [21]. *If Γ is any vertex-symmetric graph, then $B(\Gamma)$ is a closed subgroup of* $\mathrm{AUT}(\Gamma)$.

We remark that in Theorem 2, the factor graph Γ^σ must be a Cayley graph of $B(\Gamma)^\sigma$, which is isomorphic with $\boldsymbol{Z}^d$ for some $d \geqslant 1$.

One would hope that Theorem 2 could be deduced from Theorem 3 in a similar way as Theorem 1 was deduced from Proposition 1: if $B(\Gamma)$ acts transitively on X and is closed, then by Lemmas 3 and 4 it is a locally compact FC^--group; the structure theory of Grosser and Moskowitz [7, (3.13) and (3.17)] yields existence of a compact normal subgroup K of $B(\Gamma)$ such that $B(\Gamma)/K$ is discrete, torsionfree and abelian; as $B(\Gamma)$ is compactly generated, $B(\Gamma)/K$ is finitely generated, and the imprimitivity system σ can be chosen to be the one induced by K as in the proof of Theorem 1.

However, we cannot apply this simple argument, because Trofimov *uses* Theorem 2 in order to prove Theorem 3. (Indeed, I have spent—in vain—quite some time trying to find a more simple and direct proof of Theorem 3 which does not make use of Theorem 2.) In general, it is not true that the normal subgroup of FC^--elements in a locally compact group is closed.

Theorem 2 (whose proof is long and complicated, we refer to [19]) is used to deduce the following proposition, which is also of interest in itself and a key tool in deriving the stronger version of Theorem 1 mentioned at the end of Section 3.

Proposition 2 [20, Proposition 2.3]. *Let Γ be any vertex-symmetric graph. Then there is an imprimitivity system σ of* $\mathrm{AUT}(\Gamma)$ *with finite blocks, such that the stabilizer of a vertex of Γ^σ in* $\mathrm{AUT}(\Gamma^\sigma)$ *contains no bounded automorphism different from the identity.*

Thus, if $g \in B(\Gamma)$ stabilizes some $x \in X$, then it permutes each block of σ. In particular, g has finite order.

We now give an application of Theorem 3. Motivated by a problem raised by Watkins for strips (vertex-symmetric graphs with linear growth), Godsil et al. [5] show that the set $B_0(\Gamma)$ of bounded automorphisms with finite order is a *subgroup* of $B(\Gamma)$ if Γ is a vertex-symmetric graph with polynomial growth, while

$B_0(\Gamma) = B(\Gamma)$ if Γ has infinitely many ends. As a byproduct of Theorem 3, one obtains the following completion of this result.

Corollary 1. *Let Γ be a vertex-symmetric graph. Then $B_0(\Gamma)$ is a closed normal subgroup of* $\mathrm{AUT}(\Gamma)$, *and $B(\Gamma)/B_0(\Gamma)$ is a torsion free discrete abelian group.*

Proof. By Theorem 3 and Lemma 3, $B(\Gamma)$ is a locally compact FC^--group. Thus, [7, 3.13] applies and yields the result. □

We remark that $B(\Gamma)$ need not be compactly generated (if it does not act vertex-transitively), so that a priori $B(\Gamma)/B_0(\Gamma)$ is not necessarily finitely generated. It would be interesting to know whether this may really happen.

5. Expanders, amenability and unimodularity

In this final section, we present without proofs another result which links a structural property of vertex-symmetric graphs with their automorphism groups, seen as topological groups.

If F is a finite subset of the vertex set X of Γ, then ∂F denotes the set of vertices in F which have a neighbour in $X \setminus F$. The graph Γ is called an *infinite expander* if there is a number $\kappa > 0$ such that

$$|\partial F| \geqslant \kappa \cdot |F| \quad \text{for every finite } F \subset X.$$

See, for example, Bien [1], Dodzuik [2], and Gerl [4] for various reasons why expanders are of interest. A theorem of Soardi and Woess [16] characterizes infinite vertex-symmetric (non-)expanders in terms of their automorphism groups.

A topological group G is called *amenable*, if there is a nonnegative measure μ, defined on the Borel sets of G, with the following properties:

(1) $\mu(G) = 1$,

(2) μ is finitely additive, and

(3) $\mu(gU) = \mu(U)$ for every Borel set $U \subset G$ and every $g \in G$.

If G is compact, then we may take the Haar measure, but in general, not every group is amenable. Examples of amenable groups are abelian and solvable groups. Examples of (discrete) non-amenable groups are free groups with at least two free generators. The class of amenable groups is of interest in many respects, see the books by Pier [13] and Wagon [23].

Next, we turn to the Haar measure: the modular function Δ on a locally compact group G is defined by

$$\Delta(g) = \frac{\lambda(Ug)}{\lambda(U)}, \quad g \in G,$$

where $U \subset G$ is open with compact closure (Δ is independent of such U). A group is called *unimodular* if $\Delta \equiv 1$. Discrete and abelian groups are unimodular, but there are non-unimodular solvable groups. For details, see e.g. [10].

Now consider a graph Γ, and let G be a closed subgroup of $\mathrm{AUT}(\Gamma)$ which acts transitively on X. Then the modular function of G can be easily determined.

Lemma 5 ([14, 21]). *If $g \in G$ and $go = x$ then $\Delta(g) = |G_o x|/|G_x o|$.*

Non-expanders can be characterized as follows.

Theorem 4 [16]. *Let Γ be a vertex-symmetric graph. Then the following statements are equivalent:*

(a) *Γ is a non-expander.*

(b) *Some closed, vertex-transitive subgroup of* $\mathrm{AUT}(\Gamma)$ *is amenable and unimodular.*

(c) *Every closed, vertex-transitive subgroup of* $\mathrm{AUT}(\Gamma)$ *is amenable and unimodular.*

To conclude, we point out that this theorem is applied in [16] to prove that every vertex-symmetric graph with infinitely many ends is an infinite expander. For the definition and basic features of the space of ends, see Freudenthal [3], Halin [8] or Woess [25]. In [25], the relation between amenability of a closed group G of automorphisms of an arbitrary locally finite graph Γ and the action of G on the space of ends of Γ is described in full detail. This completes results by Tits [17] and Nebbia [12] for trees and by Seifter [15] for arbitrary graphs and nilpotent groups. Based on [25], the following is proved in [16].

Proposition 3 [16]. *Assume that Γ has infinitely many ends and that G is a closed, vertex-transitive group of automorphisms of Γ. Then G is amenable if and only if it fixes an end of Γ. In this case, G is non-unimodular.*

Appendix: Three problems on infinite graphs

The three problems presented here concern locally finite, infinite, connected, vertex-symmetric graphs.

First problem

"Are there any vertex-symmetric graphs which do not look like Cayley graphs?"

First of all, let us recall the definition of a *Cayley graph*. If G is a finitely generated group and A is a finite symmetric set of generators, then the Cayley graph of G with respect to A has vertex set G; $[g, h]$ is an (unoriented) edge if and only if $h = ga$ for some $a \in A$. By left multiplication, G acts on its Cayley graph(s) as a vertex-transitive group of graph automorphisms.

The crucial point in the above question is the definition of 'look like'. Let $\Gamma_1(X_1, E_1)$ and $\Gamma_2(X_2, E_2)$ be connected, locally finite graphs, and let d_i be the

discrete metric on X_i induced by the respective graph structure, $i = 1, 2$. We say that Γ_1 and Γ_2 are *quasi-isometric* if the following holds: there are maps $\varphi: X_1 \to X_2$ and $\psi: X_2 \to X_1$ and constants $C, D > 0$ such that for every $x_1, y_1 \in X_1$ and $x_2, y_2 \in X_2$ one has:

(a) $d_2(\varphi x_1, \varphi y_1) \leq C d_1(x_1, y_1) + D$,

(b) $d_1(\psi x_2, \psi y_2) \leq C d_2(x_2, y_2) + D$,

(c) $d_1(\psi\varphi x_1, x_1) \leq C$, and

(d) $d_2(\varphi\psi x_2, x_2) \leq C$.

Conditions (a) and (b) say that φ and ψ are 'quasi Lipschitz', while (c) and (d) say that they are 'quasi inverse' to each other. Thus, the two graphs are metrically equivalent up to bounded deviation. Now we can formulate our problem more precisely.

Problem 1. Is there a vertex-symmetric graph which is not quasi-isometric with some Cayley graph?

We remark that by [20], every vertex-symmetric graph with polynomial growth is quasi-isometric with the Cayley graph of a nilpotent-by-finite group in a rather strong sense, see Theorem 1 above. Furthermore, the non-Cayley graphs exhibited in [18] and [16] are quasi-isometric with homogeneous trees.

Second problem

This and the third problem regard vertex-symmetric graphs with infinitely many ends. We recall the definition of the *space of ends* of an infinite graph Γ: an infinite path in Γ is a sequence $\pi = [x_0, x_1, x_2, \ldots]$ of successively adjacent vertices in Γ without repetitions. Two infinite paths are equivalent if there is a third one which meets each of the two infinitely often. An end ω is an equivalence class of infinite paths under this relation. For more details, see [3, 8]. Let Ω denote the space of ends. Every automorphism (isometry) of Γ acts on Ω in an obvious way. If Γ is vertex-symmetric, then it is easy to see that Γ has one, two or infinitely many ends [3]. From the structural viewpoint, the first case is the most difficult and the second the simplest one. We are interested in the third.

Problem 2. Classify all vertex-symmetric graphs Γ with infinitely many ends, for which the automorphism group $\mathrm{AUT}(\Gamma)$ acts transitively on the space of ends Ω.

Typical examples are homogeneous trees and more generally, infinite distance-regular graphs. If Γ is a graph where $\mathrm{AUT}(\Gamma)$ acts transitively on Ω, then this is also true for Γ^k (the graph with the same vertex set, where two vertices are connected by an edge if their distance in Γ is bounded by k) and for $\Gamma \times \Gamma_1$ (Cartesian product), where Γ_1 is a finite, vertex-symmetric graph. In particular, all ends of Γ must have the same finite size (maximal number of disjoint equivalent paths, see [9]) and diameter (see [25] for the definition), so that Γ is

quasi-isometric with a tree by [24]. Of course there are many examples of vertex-symmetric graphs whose automorphism group does not act transitively on Ω (for example, take the Cayley graph of the free product of two one-ended groups).

Third problem

The homogeneous tree T with degree at least three has another interesting property. Let ω_0 be an end of T and let G be the group of automorphisms of T which fix ω_0. Then G acts transitively on the vertex set.

In general, if Γ has infinitely many ends and $G \leqslant \mathrm{AUT}(\Gamma)$ acts transitively on the vertex set, then it is not hard to see that either $|G\omega| = \infty$ for every $\omega \in \Omega$ or G fixes an end ω_0, and $|G\omega| = \infty$ for every other end ω. The second case is quite particular and of interest in various contexts, see e.g. [12, 24, 25].

Problem 3. Classify all vertex-symmetric graphs Γ with infinitely many ends with the following property: there is a vertex-transitive group of automorphisms of Γ (not necessarily the whole automorphism group) which fixes an end of Γ.

Once more, every infinite distance-regular graph has this property. If Γ admits a vertex-transitive group of automorphisms which fixes an end, then so do Γ^k ($k \geqslant 1$) and $\Gamma \times \Gamma_1$, where Γ_1 is finite and vertex-symmetric. My impression is that one cannot get far beyond these possibilities. In particular, I conjecture that every such graph is quasi-isometric with a tree.

Notes added in proof

(1) G. Schlichting (München) has told me that he has found a direct 'topological' proof of Theorem 3, as discussed in Section 4.

(2) V.I. Trofimov (Sverdlovsk) has informed me that group theorists have well-known examples of discrete groups (Cayley graphs) with $B(\Gamma)/B_0(\Gamma)$ not finitely generated.

(3) Problem 2 has been solved recently and independently by R. Möller (Oxford) and A. Nevo (Jerusalem).

(4) More recently, R. Möller has also given a solution of Problem 3.

(5) Problem 1 remains open and seems to be difficult.

References

[1] F. Bien, Constructions of telephone networks by group representations, Notices Amer. Math. Soc. 36 (1989) 5–22.

[2] J. Dodzuik, Difference equations, isoperimetric inequality and transience of certain random walks, Trans. Amer. Math. Soc. 294 (1984) 787–792.

[3] H. Freudenthal, Über die Enden diskreter Räume und Gruppen, Comment. Math. Helv. 17 (1944) 1–38.
[4] P. Gerl, Random walks on graphs with a strong isoperimetric inequality, J. Theoret. Probab. 1 (1988) 171–187.
[5] C.D. Godsil, W. Imrich, N. Seifter, M.E. Watkins and W. Woess, A note on bounded automorphisms of infinite graphs, Graphs Combin. 5 (1989) 333–338.
[6] M. Gromov, Groups of polynomial growth and expanding maps, Publ. Math. IHES 53 (1981) 53–78.
[7] S. Grosser and M. Moskowitz, Compactness conditions in topological groups, J. Reine Angew. Math. 246 (1971) 1–40.
[8] R. Halin, Über unendliche Wege in Graphen, Math. Ann. 157 (1964) 125–137.
[9] R. Halin, Die Maximalzahl zweiseitig unendlicher Wege in Graphen, Math. Nachr. 44 (1970) 119–127.
[10] E. Hewitt and K.A. Ross, Abstract Harmonic Analysis, Vol. I (Springer, Berlin, 1963).
[11] V. Losert, On the structure of groups with polynomial growth, Math. Z. 195 (1987) 109–117.
[12] C. Nebbia, Amenability and Kunze–Stein property for groups acting on trees, Pacific J. Math. 135 (1988) 371–380.
[13] J.P. Pier, Amenable Locally Compact Groups (Wiley, New York, 1984).
[14] G. Schlichting, Polynomidentitäten und Permutationsdarstellungen lokalkompakter Gruppen, Invent. Math. 55 (1979) 97–106.
[15] N. Seifter, On the action of nilpotent and metabelian groups on infinite, locally finite graphs, European J. Combin 10 (1989) 41–45.
[16] P.M. Soardi and W. Woess, Amenability, unimodularity, and the spectral radius of random walks on infinite graphs, Math. Z. 205 (1990) 471–486.
[17] J. Tits, A "theorem of Lie–Kolchin" for trees, in: Contributions to Algebra, a Collection of Papers Dedicated to Ellis Kolchin (Academic Press, New York, 1977) 337–388.
[18] C. Thomassen and M.E. Watkins, Infinite vertex-transitive, edge-transitive, non-1-transitive graphs, Proc. Amer. Math. Soc. 105 (1989) 258–261.
[19] V.I. Trofimov, Automorphisms of graphs and a characterization of lattices, Math. USSR-Izv. 22 (1984) 379–391.
[20] V.I. Trofimov, Graphs with polynomial growth, Math. USSR-Sb. 51 (1985) 405–417.
[21] V.I. Trofimov, Automorphism groups of graphs as topological groups, Math. Notes 38 (1985) 717–720.
[22] V.S. Varadarajan, Lie Groups, Lie Algebras and their Representations (Prentice-Hall, Englewood Cliffs, 1974).
[23] S. Wagon, The Banach–Tarski Paradox (Cambridge Univ. Press, Cambridge, 1985).
[24] W. Woess, Graphs and groups with tree-like properties, J. Combin. Theory Ser. B 47 (1989) 361–371.
[25] W. Woess, Amenable group actions on infinite graphs, Math. Ann. 284 (1989) 251–265.
[26] W. Woess, Boundaries of random walks on graphs and groups with infinitely many ends, Israel J. Math. 68 (1989) 271–301.
[27] J.A. Wolf, Growth of finitely generated solvable groups and curvature of Riemannian manifolds, J. Differential Geom. 2 (1968) 421–446.

Discrete Mathematics 95 (1991) 385
North-Holland

Author index to volume 95 (1991)

Aharoni, R., Infinite matching theory (1–3) 5– 22
Brochet, J.-M. and M. Pouzet, Gallai–Milgram properties for infinite graphs (1–3) 23– 47
Cameron, P.J., The age of a relational structure (1–3) 49– 67
Diestel, R., Decomposing infinite graphs (1–3) 69– 89
Erdős, P. and J.A. Larson, Matchings from a set below to a set above (1–3) 169–182
Halin, R., Bounded graphs (1–3) 91– 99
Imrich, W. and N. Seifter, A survey on graphs with polynomial growth (1–3) 101–117
Jung, H.A., Some results on ends and automorphisms of graphs (1–3) 119–133
Kříž, I. and R. Thomas, Analyzing Nash-Williams' partition theorem by means of ordinal types (1–3) 135–167
Komjáth, P. and J. Pach, Universal elements and the complexity of certain classes of infinite graphs (1–3) 255–270
Larson, J.A., see Erdős, P. (1–3) 169–182
Milner, E.C. and K. Prikry, A partition relation for triples using a model of Todorčević (1–3) 183–191
Mohar, B., Some relations between analytic and geometric properties of infinite graphs (1–3) 193–219
Nash-Williams, C.St.J.A., Reconstruction of infinite graphs (1–3) 221–229
Niedermeyer, F., f-Optimal factors of infinite graphs (1–3) 231–254
Pach, J., see Komjáth, P. (1–3) 255–270
Polat, N. and G. Sabidussi, Asymmetrising sets in trees (1–3) 271–289
Polat, N., Asymmetrization of infinite trees (1–3) 291–301
Pouzet, M., see Brochet, J.-M. (1–3) 23– 47
Prikry, K., see Milner, E.C. (1–3) 183–191
Prömel, H.J., W. Thumser and B. Voigt, Fast growing functions based on Ramsey theorems (1–3) 341–358
Robertson, N., P. Seymour and R. Thomas, Excluding infinite minors (1–3) 303–319
Sabidussi, G., see Polat, N. (1–3) 271–289
Seifter, N., see Imrich, W. (1–3) 101–117
Seymour, P. and R. Thomas, An end-faithful spanning tree counterexample (1–3) 321–330
Seymour, P., see Robertson, N. (1–3) 303–319
Širáň, J., End-faithful forests and spanning trees in infinite graphs (1–3) 331–340
Thomas, R., see Kříž, I. (1–3) 135–167
Thomas, R., see Robertson, N. (1–3) 303–319
Thomas, R., see Seymour, P. (1–3) 321–330
Thumser, W., see Prömel, H.J. (1–3) 341–358
Voigt, B., see Prömel, H.J. (1–3) 341–358
Watkins, M.E., Edge-transitive strips (1–3) 359–372
Woess, W., Topological groups and infinite graphs (1–3) 373–384