Path homomorphisms, graph colourings and boolean matrices

Jaroslav Nešetřil∗
Department of Applied Mathematics
and
Institute for Theoretical Computer Science (ITI)
Charles University
Malostranské nám.25, 11800 Praha 1
Czech Republic
Email: nesetril@kam.ms.mff.cuni.cz

Claude Tardif†
Department of Mathematics and Computer Science
Royal Military College of Canada
PO Box 17000, Station “Forces”
Kingston, Ontario K7K 7B4 Canada
Email: Claude.Tardif@rmc.ca

July 2007

Keywords: Graph colourings, paths, Hasse-Gallai-Roy-Vitaver Theorem
AMS 1991 Subject Classification: 05C15.

Abstract

We investigate bounds on the chromatic number of a graph $G$ derived from the nonexistence of homomorphisms from some path $\vec{P}$ into some orientation $\vec{G}$ of $G$. The condition is often efficiently verifiable using boolean matrix multiplications. However, the bound associated to a path $\vec{P}$ depends on the relation between the “algebraic length” and “derived algebraic length” of $\vec{P}$. This suggests that paths yielding efficient bounds may be exponentially large with respect to $G$, and the corresponding heuristic may not be constructive.

∗Partially supported by the Project LN00A056 of the Czech Ministry of Education and by CRM Barcelona, Spain.
†Supported by grants from NSERC and ARP
1 Introduction

In this paper we will investigate oriented paths $\vec{P}$ and integers $n$ for which the following statement holds:

If a graph $G$ admits an orientation $\vec{G}$ such that there is no homomorphism from $\vec{P}$ to $\vec{G}$, then $G$ is $n$-colourable.

The best known result of this type states that if a graph $G$ admits an orientation with no directed path with $n$ forward arcs, then $G$ is $n$-colourable. This is usually called the “Gallai-Roy” theorem, but in fact in the sixties it was published independently in four languages by four different authors: Gallai [3] in english, Hasse [4] in german, Roy [14] in french and Vitaver [16] in russian. To restore the balance we will call it the Hasse-Vitaver theorem in this paper.

Some extensions of the Hasse-Vitaver theorem departed from the context of graph colouring. In the seventies, Neˇ setˇ ril and Pultr [10] interpreted this result as a “homomorphism duality” between the directed path $\vec{P}_n$ with $n$ forward arcs and the transitive tournament $\vec{T}_n$ with $n$ vertices. All singleton dualities for digraphs were later characterised in Komárek’s Ph. D. thesis [6] in czech. The finite dualities were characterised in [11] for digraphs and general relational structures, and later drew attention in descriptive complexity as they correspond to the first-order definable constraint satisfaction problems (see [1, 8, 9]). Here we link these developments back to their original context of graph colouring.

Every path $\vec{P}$ admits a dual $\vec{D}$ such that for every digraph $\vec{G}$ there exists a homomorphism from $\vec{G}$ to $\vec{D}$ if and only if there is no homomorphism from $\vec{P}$ to $\vec{G}$. Thus the following holds:

If a graph $G$ admits an orientation $\vec{G}$ such that there is no homomorphism from $\vec{P}$ to $\vec{G}$, then $G$ is $\chi(\vec{D})$-colourable. When $\vec{D}$ admits a homomorphism to the transitive tournament $\vec{T}_{\chi(\vec{D})}$, the corresponding bound is weaker than or equivalent to the Hasse-Vitaver bound. However it has been noted in [2, 12] that some paths yield bounds that are stronger or independent. For the moment, a complete classification of the paths with $n$-chromatic duals seems out of reach.

From the point of view of heuristics, we have the following situation: Given a $n$-chromatic graph $G$, finding an orientation $\vec{G}$ for which the Hasse-Vitaver theorem gives a tight upper bound amounts to finding a $n$-colouring of $G$, so these orientations are computationally hard to find. Now suppose that there exists a large path $\vec{P}$ containing $\vec{P}_n$, such that the dual of $\vec{P}$ is...
still $n$-chromatic. Then the orientations of $G$ avoiding homomorphisms from $\vec{P}$ are more numerous than the orientations avoiding $\vec{P}_n$; for instance, some orientations derived from $(n + 1)$-colourings may be suitable. Thus it would seem that in some cases these orientations could be easier to find, yielding better heuristics. On the other hand, when a $n$-colouring is derived from an orientation of $G$ avoiding $\vec{P}$, the orientation heuristic turns out to be just a colouring heuristic in disguise, so it cannot be an improvement. Therefore it seems that heuristics cannot be improved by such strenghtenings of the Hasse-Vitaver theorem.

However there is a flaw in this argument, in that it assumes that finding an orientation of $G$ avoiding $\vec{P}$ necessarily yields a $n$-colouring of $G$. Indeed it is true that the known $n$-colourability heuristics yield $n$-colourings, contrary to the good compositeness heuristics in number theory which are nonconstructive. Could there exist good nonconstructive heuristic to determine that a graph is $n$ colourable? Our main results indicate that the promising paths candidates using our approach may be exponential in size. Using recursive boolean matrix multiplications, it is sometimes possible to tell in polynomial time that a digraph $\vec{G}$ admits no homomorphism from an exponentially large path $\vec{P}$; however this information may not be sufficient to find a $n$-colouring of $G$ in polynomial time, so that the corresponding heuristics could be nonconstructive.

Our concrete results are in Sections 2, 3, 4. In Sections 2 and 3, we give examples of paths whose duals have relatively low chromatic number, and in Section 4 we use shift graphs to provide a lower bound on the chromatic number of the dual of these paths. In Sections 5, 6, we construct recursive variants of the “Pell paths” respecting the constraints of Section 4; we cannot prove that their duals have bounded chromatic numbers, but we prove that in some case their duals have double exponential size. We conclude with open problems in Section 7.

2 Algebraic length

A homomorphism is an arc-preserving map between two digraphs. We write $\vec{G} \rightarrow \vec{H}$ if there exists a homomorphism from $\vec{G}$ to $\vec{H}$, and $\vec{G} \not\rightarrow \vec{H}$ otherwise. An orientation $\vec{G}$ of an undirected graph $G$ is obtained by selecting one of the two arcs $(u, v), (v, u)$ for every edge $[u, v]$. However when no orientation of $G$ is given, we consider that both arcs $(u, v), (v, u)$ corresponding to an edge $[u, v]$ are present. In particular, the chromatic number of a digraph is defined by ignoring the orientation, that is, $\chi(\vec{G}) = \min\{n : \vec{G} \rightarrow K_n\}$. 

3
Let \( \tilde{P}_n \) be the path with vertices 0, 1, \ldots, \( n \) and arcs \((0, 1), (1, 2), \ldots, (n-1, n)\). We will use these paths as building blocks for larger paths using inversion and concatenation: We denote \( \tilde{P}^{-1} \) the path obtained from \( \tilde{P} \) by reading it backwards, and \( \tilde{P} \circ \tilde{Q} \) the concatenation of \( \tilde{P} \) and \( \tilde{Q} \) obtained by identifying the last vertex of \( \tilde{P} \) with the first vertex of \( \tilde{Q} \). (Thus informally we view paths as drawn horizontally with a “first” vertex on the left and a “last” vertex on the right; though formally \( \tilde{P}^{-1} \) is isomorphic to \( \tilde{P} \).) For every directed path \( \tilde{P} \), the least integer \( n \) such that \( \tilde{P} \to \tilde{P}_n \) is called the **algebraic length** of \( \tilde{P} \); it is denoted \( \text{al}(\tilde{P}) \). The homomorphism \( h : \tilde{P} \to \tilde{P}_{\text{al}(\tilde{P})} \) is unique; for a vertex \( u \) of \( \tilde{P} \) we call \( h(u) \) the **height** of \( u \).

We know from \([6, 11, 12, 13]\) that every path \( \tilde{P} \) admits a dual \( \tilde{D}(\tilde{P}) \) such that for every digraph \( \tilde{G} \), we have \( \tilde{G} \to \tilde{D}(\tilde{P}) \) if and only if \( \tilde{P} \not\to \tilde{G} \). In particular \( \tilde{D}(\tilde{P}_n) \) is the transitive tournament \( \tilde{T}_n \) with \( n \) vertices. In general, for \( n = \text{al}(\tilde{P}) \), we have \( \tilde{P} \to \tilde{P}_n \) whence \( \tilde{P}_n \neq \tilde{D}(\tilde{P}) \) and \( \tilde{D}(\tilde{P}) \to \tilde{T}_n \), thus we get the trivial upper bound \( \chi(\tilde{D}(\tilde{P})) \leq n = \text{al}(\tilde{P}) \). The trivial lower bound is the clique number of (the symmetrisation of) \( \tilde{D}(\tilde{P}) \), that is, the largest \( m \) such that \( \tilde{P}_m \to \tilde{P} \). Indeed, \( \tilde{P}_m \to \tilde{P} \) implies \( \tilde{P} \not\to \tilde{T}_m \) whence \( \tilde{T}_m \to \tilde{D}(\tilde{P}) \). Both bounds can be off the mark, as is implicit in the following:

**Lemma 1** ([12, 2]) For any integers \( i, j \geq 2 \), we have \( \text{al}(\tilde{P}_i \circ \tilde{P}_1^{-1} \circ \tilde{P}_j) = i + j - 1 \) and \( \chi(\tilde{D}(\tilde{P}_i \circ \tilde{P}_1^{-1} \circ \tilde{P}_j)) \leq i + j - 2 \). In particular \( \chi(\tilde{D}(\tilde{P}_3 \circ \tilde{P}_1^{-1} \circ \tilde{P}_3)) = 4 \) while \( \text{al}(\tilde{P}_3 \circ \tilde{P}_1^{-1} \circ \tilde{P}_3) = 5 \).

Our main result of this section is that for some paths the ratio of the algebraic length over the chromatic number of the dual can be arbitrarily large.

**Theorem 2** There exists a sequence \( \tilde{S}_n \) of paths such that

\[
\lim_{n \to \infty} \frac{\text{al}(\tilde{S}_n)}{\chi(\tilde{D}(\tilde{S}_n))} = \infty.
\]

Theorem 2 will be a consequence of Lemma 3 below. We define the sequence \( \{\tilde{S}_n\}_{n \in \mathbb{N}} \) of “staircase” paths recursively by

\[
\begin{align*}
\tilde{S}_1 &= \tilde{P}_2, \\
\tilde{S}_{n+1} &= \tilde{S}_n \circ \tilde{P}_1^{-1} \circ \tilde{P}_2.
\end{align*}
\]

Thus the algebraic length of \( \tilde{S}_n \) is \( n + 1 \). Also, since \( \tilde{S}_n \to \tilde{S}_{n+1} \), we have \( \tilde{D}(\tilde{S}_n) \to \tilde{D}(\tilde{S}_n+1) \), whence the sequence \( \{\chi(\tilde{S}_n)\}_{n \in \mathbb{N}} \) is nondecreasing. We prove the following:

**Lemma 3** \( \chi(\tilde{D}(\tilde{S}_{2n})) \leq \chi(\tilde{D}(\tilde{S}_n)) + 4 \).
Proof. Since $\vec{S}_{2n} \not\rightarrow \vec{S}_n$, we can consider the set of all homomorphisms from $\vec{S}_n$ to $D(\vec{S}_{2n})$ and define the following sets of vertices:

A: The set of all images of the first vertex of $\vec{S}_n$ in a homomorphism to $D(\vec{S}_{2n})$.

B: The set of all images of the last vertex of $\vec{S}_n$ in a homomorphism to $D(\vec{S}_{2n})$.

C: The set of all vertices in $D(\vec{S}_{2n}) \setminus (A \cup B)$ that have an outneighbour or an inneighbour in $A$ or $B$.

Since $\vec{S}_n \rightarrow \vec{P}_1 \circ \vec{S}_{n-1}$, for every arc $(u, v)$ of $D(\vec{S}_{2n})$ with $v \in A$ we have $u \in A$, and since $\vec{S}_n \rightarrow \vec{S}_{n-1} \circ \vec{P}_1$, for every arc $(u, v)$ of $D(\vec{S}_{2n})$ with $u \in B$ we have $v \in B$. Also, since $\vec{S}_{2n} \rightarrow \vec{S}_n \circ \vec{S}_n$, there is no homomorphism from $\vec{S}_n$ to $A$ or to $B$ as this would induce a homomorphism from $\vec{S}_{2n}$ to $D(\vec{S}_{2n})$. Similarly, since $\vec{S}_{2n} = \vec{S}_n \circ \vec{P}^{-1} \circ \vec{S}_n$, there is no arc from $A$ to $B$ and thus no arc between $A$ and $B$.

Thus we can use the same $\chi(D(\vec{S}_n))$ colours to colour both $A$ and $B$. Also, $D(\vec{S}_{2n}) \setminus (A \cup B \cup C)$ admits no homomorphism from $\vec{S}_n$, and is not connected to $A$ and $B$, so we can again use the same $\chi(D(\vec{S}_n))$ colours for that part. Thus we need new colours only for $C$. However $C$ cannot contain a path with arcs $(u_0, u_1), (u_1, u_2), (u_2, u_3), (u_3, u_4)$ for then an arc from $A$ to $u_2$ would imply that $u_0$ is in $A$, and an arc from $u_2$ to $B$ would imply that $u_4$ is in $B$. Thus $\vec{P}_4 \not\rightarrow C$ whence by the Hasse-Vitaver theorem, four colours suffice for $C$. Therefore $\chi(D(\vec{S}_{2n})) \leq \chi(D(\vec{S}_n)) + 4$. \hfill \qed

Proof of Theorem 2. Since $\vec{S}_4 \rightarrow \vec{P}_3 \circ \vec{P}_1^{-1} \circ \vec{P}_3$, we have $\chi(D(\vec{S}_4)) \leq 4$ by Lemma 1. Thus by Lemma 3, $\chi(S_{2k}) \leq 4(k - 1)$ for $k \geq 2$. Therefore, $\frac{\text{al}(\vec{S}_n)}{\chi(D(\vec{S}_n))} \geq (n + 1)/4\lg(n)$ for all $n \geq 4$. \hfill \qed

The tests provided by the paths $\vec{S}_n$ are independent from the Hasse-Vitaver theorem: For every orientation $\vec{G}$ of a graph $G$ containing a triangle, we have $\vec{S}_n \rightarrow \vec{G}$ for every $n$, hence these tests are inconclusive. However for some oriented triangle-free graphs, the paths $\vec{S}_n$ provide better bounds than the Hasse-Vitaver theorem. We will see in Section 4 that the order of growth of $\chi(D(\vec{S}_n))$ is indeed logarithmic. But first in the next section we will discuss variations of Lemma 3.

3 Digression: self-feeding

In the proof of Lemma 3, we used the Hasse-Vitaver theorem to bound the number of additional colours needed to colour a graph. Since our end
purpose is to find strong variants of the Hasse-Vitaver theorem, we should instead use our partial results to refine our bounds. We are not able to do this with the paths \( \tilde{S}_n \) of the previous section, but here we present another family of paths where such a refinement is possible.

We define the sequence \( \{ \tilde{Z}_n \}_{n \in \mathbb{N}} \) of “long staircases” by

\[
\begin{align*}
\tilde{Z}_1 &= \tilde{P}_2, \\
\tilde{Z}_{n+1} &= \tilde{Z}_n \circ \tilde{P}_1^{-1} \circ \tilde{P}_1 \circ \tilde{P}_1^{-1} \circ \tilde{P}_2.
\end{align*}
\]

**Lemma 4** \( \chi(\tilde{D}(\tilde{Z}_{2n})) \leq \chi(\tilde{D}(\tilde{Z}_n)) + 3 \).

**Proof.** The proof is similar to that of Lemma 3; so we will omit many details. Consider the set of all homomorphisms from \( \tilde{Z}_n \) to \( \tilde{D}(\tilde{Z}_{2n}) \) and define the following sets of vertices:

- \( A \): The set of all images of the first vertex of \( \tilde{Z}_n \) in a homomorphism to \( \tilde{D}(\tilde{Z}_{2n}) \).
- \( B \): The set of all images of the last vertex of \( \tilde{Z}_n \) in a homomorphism to \( \tilde{D}(\tilde{Z}_{2n}) \).
- \( C \): The set of all vertices in \( \tilde{D}(\tilde{Z}_{2n}) \setminus (A \cup B) \) that have an outneighbour or an inneighbour in \( A \) or \( B \).

There is no homomorphism from \( \tilde{Z}_n \) to \( A \), \( B \), or \( \tilde{D}(\tilde{Z}_{2n}) \setminus (A \cup B \cup C) \), and no arcs leaving \( B \) or entering \( A \). Thus we can use the same \( \chi(D(Z_n)) \) colours for \( A \) and \( B \), and \( \tilde{D}(\tilde{Z}_{2n}) \setminus (A \cup B \cup C) \). We need new colours only for \( C \). However \( C \) cannot contain a path with arcs \((u_0, u_1), (u_1, u_2), (u_2, u_3), (u_4, u_5), (u_5, u_6)\), for then an arc from \( A \) to \( u_2 \) would imply that \( u_0 \) is in \( A \), and an arc from \( u_2 \) to \( B \) would imply that \( u_6 \) is in \( B \). Thus \( \tilde{P}_3 \circ \tilde{P}_1^{-1} \circ \tilde{P}_2 \not\to C \), whence by Lemma 1, three new colours suffice for \( C \). \( \blacksquare \)

**Corollary 5** \( \chi(\tilde{D}(\tilde{Z}_{2k})) \leq 3(k - 1) \) for all \( k \geq 2 \).

**Proof.** By a result of [2], we have \( \chi(\tilde{D}(\tilde{P})) = 3 \) where

\[
\tilde{P} = \tilde{P}_3 \circ \tilde{P}_1^{-1} \circ \tilde{P}_2 \circ \tilde{P}_1^{-1} \circ \tilde{P}_1 \circ \tilde{P}_1^{-1} \circ \tilde{P}_2.
\]

Since \( \tilde{Z}_4 \to \tilde{P} \), we have \( \chi(\tilde{D}(\tilde{Z}_4)) = 3 \), hence by Lemma 4, \( \chi(\tilde{D}(\tilde{Z}_{2k})) \leq 3(k - 1) \) for all \( k \geq 2 \). \( \blacksquare \)

Thus for the paths \( \tilde{S}_n \) of the previous section we get \( \chi(\tilde{D}(\tilde{S}_n)) \leq 4 \lg(n) \), while \( \chi(\tilde{D}(\tilde{Z}_n)) \leq 3 \lg(n) \). We might get even better bounds for chromatic numbers of duals of “longer staircases”. The next section looks at lower bounds in terms of homomorphisms from shift graphs.
4 Shift graphs

For integers $1 \leq d < n$, the shift graph $\partial^d(\mathcal{T}_n)$ is the digraph whose vertices are the sequences $A = (a_0, a_1, \ldots, a_d)$ with $0 \leq a_0 < a_1 < \cdots < a_d \leq n - 1$, and whose arcs are the couples of “consecutive” sequences, that is, the couples $(A, B)$ such that $b_0 = a_1, b_1 = a_2, \ldots, b_{d-1} = a_d$. We use the following folklore result.

**Theorem 6 (see [5])** $\chi(\partial^d(\mathcal{T}_n)) \geq \lg^d(n)$, where $\lg^d$ means $\lg \circ \lg \circ \cdots \circ \lg$.

For the paths $S_n, Z_n, n \geq 1$ defined in the previous sections, it is not hard to see that $Z_n \not\sim \partial(T_{n+1})$, thus $\partial(T_{n+1}) \to \bar{D}(Z_n)$. Therefore $\chi(\bar{D}(Z_n)) \geq \chi(\bar{D}(Z_n)) \geq \lg(n + 1)$. We will prove this in more generality after we introduce a relevant parameter.

Every path $\bar{P}$ contains a subpath $S = u_0, u_1, \ldots, u_m$ such that $h(u_0) = 0, h(u_m) = \text{al}(\bar{P})$ and no other vertex of the subpath has height 0 or $\text{al}(\bar{P})$. Then there exists a maximum integer $d$ such that there exist indices $i \leq j \leq k$ with $h(u_i) = h(u_k) = h(u_j) + d$. If $\bar{P}$ contains many such subpaths the minimum of the corresponding $d$’s is called the derived algebraic length $\text{al}'(\bar{P})$ of $\bar{P}$. For instance, for every $n$ we have $\text{al}'(\bar{P}_n) = 0$ and $\text{al}'(\bar{S}_n) = \text{al}'(\bar{Z}_n) = 1$.

**Lemma 7** For $d = \text{al}'(\bar{P})$ and $n = \text{al}(\bar{P}) + \text{al}'(\bar{P})$ we have $\partial^d(\mathcal{T}_n) \to \bar{D}(\bar{P})$, and therefore $\chi(\bar{D}(\bar{P})) \geq \lg^d(n)$.

**Proof.** It suffices to show that $\bar{P} \not\sim \partial^d(\mathcal{T}_n)$. Suppose for a contradiction that there exist a homomorphism $f : P \to \partial^d(\mathcal{T}_n)$. For every vertex $u$ of $\bar{P}$, $f(u)$ is an array $(a_0, a_1, \ldots, a_d)$; we put $f_i(u) = a_i$ for $i = 0, \ldots, d$. Consider the restriction of the height function $h$ on the subpath $S = u_0, u_1, \ldots, u_m$ used to define $d = \text{al}'(P)$ above. On this subpath the value $h(u_i)$ will vary from $h(u_0) = 0$ to $h(u_m) = \text{al}(\bar{P})$ in steps of $\pm 1$.

**Claim:** Whenever $h$ reaches a new maximum $x$ on a vertex $u_i$, we have $f_0(u_i) \geq h(u_i)$.

The contradiction follows easily from the Claim. Indeed, the height $\text{al}(\bar{P})$ is reached for the first time at $u_m$, so that

$$f_d(u_m) \geq f_0(u_m) + d \geq \text{al}(P) + \text{al}'(P) = n,$$

which is impossible.

**Proof of Claim.** We use induction on $x$. The first two arcs of $S$ are $(u_0, u_1), (u_1, u_2)$ so the Claim is clearly true for $x = 0, 1, 2$. Suppose that the Claim is true for $x$, where the height $x$ is reached for the first time at vertex
Let $u_{k+1}$ be the first vertex where the next height $x+1$ is reached. Thus $h(u_k) = h(u_i) = x$, and by definition of $d = \text{al}'(\vec{P})$, for $i \leq j \leq k$ we have $x - d \leq h(u_j) \leq x$. We use the following.

**Subclaim:** For $i \leq j \leq k$, we have $f_{x-h(u_j)}(u_j) = f_0(u_i)$. The Claim follows easily from the Subclaim. Indeed, we then get

$$f_0(u_{k+1}) = f_1(u_k) > f_0(u_k) = f_0(u_i) \geq x,$$

therefore $f_0(u_{k+1}) \geq x + 1 = h(u_{k+1})$.

**Proof of Subclaim.** We use induction on $j$. For $j = i$, the result is given. Now suppose that $f_{x-h(u_j)}(u_j) = f_0(u_i)$. If $(u_j, u_{j+1})$ is an arc of $\vec{P}$, then $h(u_{j+1}) = h(u_j) + 1$ and $(f(u_j), f(u_{j+1}))$ is an arc of $\partial^d(\vec{T}_n)$. Thus by definition of these arcs,

$$f_{x-h(u_{j+1})}(u_{j+1}) = f_{x-h(u_j)-1}(u_{j+1}) = f_{x-h(u_j)}(u_j) = f_0(u_i).$$

Similarly, if $(u_{j+1}, u_j)$ is an arc of $\vec{P}$, then $h(u_{j+1}) = h(u_j) - 1$ and $(f(u_{j+1}), f(u_j))$ is an arc of $\partial^d(\vec{T}_n)$. Thus by definition of these arcs,

$$f_{x-h(u_{j+1})}(u_{j+1}) = f_{x-h(u_j)+1}(u_{j+1}) = f_{x-h(u_j)}(u_j) = f_0(u_i).$$

Thus we are able to deduce a further restriction on sequences of paths whose duals have bounded chromatic numbers.

**Corollary 8** Let $\{\vec{Q}_n\}_{n \in \mathbb{N}}$ be paths such that $\lim_{n \to \infty} \text{al}(\vec{Q}_n) = \infty$ and $\{\text{al}'(\vec{Q}_n)\}_{n \in \mathbb{N}}$ is bounded. Then $\lim_{n \to \infty} \chi(\vec{Q}_n) = \infty$.

5 **Pell paths**

Perhaps the simplest construction of paths $\{\vec{Q}_n\}_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} \text{al}(\vec{Q}_n) = \lim_{n \to \infty} \text{al}'(\vec{Q}_n) = \infty$ is the sequence $\{\text{Pell}_n\}_{n \in \mathbb{N}}$ defined by

$$\text{Pell}_1 = \vec{P}_1$$

$$\text{Pell}_2 = \vec{P}_2$$

$$\text{Pell}_{n+1} = \text{Pell}_n \circ \text{Pell}_{n-1}^{-1} \circ \text{Pell}_n.$$

Thus $\text{al}(\text{Pell}_n) = n$ and for $n \geq 2$, $\text{al}'(\text{Pell}_n) = n - 2$. Hence Corollary 8 does not imply that $\{\chi(\vec{D}(\text{Pell}_n))\}_{n \in \mathbb{N}}$ is unbounded; perhaps it is even bounded by 3, but we have no proof of this.
The number of arcs in $\vec{P}_{\text{ell}}_n$ is the $n$-th Pell number $p_n$, where $p_1 = 1$, $p_2 = 2$ and $p_{n+1} = 2p_n + p_{n-1}$. Thus the size of $\vec{P}_{\text{ell}}_n$ is exponential in $n$; in fact, it is well known that $p_n = \frac{1}{\sqrt{2}}((1 + 2\sqrt{2})^n - (1 - 2\sqrt{2})^n)$. Nonetheless, for a fixed digraph $\vec{G}$, the existence of a homomorphism from $\vec{P}_{\text{ell}}_n$ to $\vec{G}$ can be tested efficiently, using boolean matrix multiplications.

Let $\{u_1, \ldots, u_g\}$ be the vertex set of $\vec{G}$. The boolean adjacency matrix $A = (a_{ij})_{1 \leq i,j \leq g}$ of $\vec{G}$ is defined by

$$a_{ij} = \begin{cases} 1 & \text{if } (u_i, u_j) \text{ is an arc of } \vec{G}, \\ 0 & \text{otherwise.} \end{cases}$$

Boolean products of matrices are performed by replacing products by minimums and sums by maximums. For a path $\vec{P}$ with $p$ arcs, consider the product $B = \prod_{k=1}^{p} B_k$, where $B_k = A$ if the $k$-th arc of $\vec{P}$ is forward (left to right), and $B_k = A^T$ if the $k$-th arc of $\vec{P}$ is backward (right to left). Then $b_{ij} = 1$ if and only if there exists a homomorphism from $\vec{P}$ to $\vec{G}$ mapping the first vertex of $\vec{P}$ to $u_i$ and the last to $u_j$. Thus $\vec{P} \rightarrow \vec{G}$ if and only if $B \neq 0$.

**Proposition 9** There exists a polynomial algorithm which tests, for an input digraph $\vec{G}$, which is the minimum integer $n$ such that $\vec{G} \rightarrow \vec{D}(\vec{P}_{\text{ell}}_n)$ if such a $n$ exists.

**Proof.** We can test whether $\vec{P}_{\text{ell}}_n \rightarrow \vec{G}$ recursively with the following calculations. Let $A_1 = A$, the adjacency matrix of $\vec{G}$, $A_2 = A^2$, and $A_{n+1} = A_n \cdot A_{n-1}^T \cdot A_n$. Then $\vec{G} \rightarrow \vec{D}(\vec{P}_{\text{ell}}_n)$ if and only if $\vec{P}_{\text{ell}}_n \not\rightarrow \vec{G}$, that is, $A_n = 0$. We will show that if there exists an integer $n$ such that $A_n = 0$, then the smallest such integer is at most $2|V(\vec{G})|$. Consider the following sets of vertices:

- $F_n$: The set of all images of the first vertex of $\vec{P}_{\text{ell}}_n$ in a homomorphism to $\vec{G}$.
- $L_n$: The set of all images of the last vertex of $\vec{P}_{\text{ell}}_n$ in a homomorphism to $\vec{G}$.

By the structure of the Pell paths, we have $F_n \subseteq F_{n-1}$ and $L_n \subseteq L_{n-1}$, so there exists an integer $n \leq 2|V(\vec{G})|$ such that $F_n = F_{n-1}$ and $L_n = L_{n-1}$. If $F_n = \emptyset$, then $A_n = 0$. Otherwise for every vertex $a$ of $F_n$ there exists a homomorphic image of $\vec{P}_{\text{ell}}_n$ linking $a$ to some $b \in L_n$, and since $L_n = L_{n-1}$ there exists a homomorphic image of $\vec{P}_{\text{ell}}_{n-1}$ linking $b$ to some $c \in F_{n-1} = F_n$. 


and again a homomorphic image of $\tilde{Pell}_n$ linking $c$ to some $d \in L_n$. Thus $F_{n+1} = F_n$, and similarly $L_{n+1} = L_n$, and for all $k \geq n$, $F_k = F_n$ and $L_k = L_n$. We can then conclude that $A_k \neq 0$ for all $k$.

Therefore the question as to whether an input digraph $\bar{G}$ admits a homomorphism to some $\tilde{D}(\tilde{Pell}_n)$ can be tested polynomially in terms of $\bar{G}$ without reference to $n$. We now turn to the question as to whether such a homomorphism can be described efficiently.

We first note that the various constructions given in [6, 11, 12] give different duals of a same path. Nonetheless, the various duals of a given path are all “equivalent” in the following way. If $\tilde{D}$ and $\tilde{D}'$ are both duals of a given path $\tilde{P}$, then $\tilde{P} \not\to \tilde{D}$ implies $\tilde{D} \to \tilde{D}'$ and $\tilde{P} \not\to \tilde{D}'$ implies $\tilde{D} \to \tilde{D}$. Two digraphs $\tilde{G}$, $\tilde{H}$ are called homomorphically equivalent when $\tilde{G} \to \tilde{H}$ and $\tilde{H} \to \tilde{G}$. It is well known (see [5, 11]) that among all the digraphs homomorphically equivalent to a given finite digraph $\tilde{G}$, the one with the smallest number of vertices is unique up to isomorphism; we call it the core of $\tilde{G}$. Thus we can restrict the notation $\tilde{D}(\tilde{P})$ to denote the core of the dual of $\tilde{P}$, which is well defined and unique.

The dual constructions in [6, 11, 12] each give a dual of exponential size to a tree, but the example $\tilde{D}(\bar{P}_n) = \bar{T}_n$ shows that the core of the dual can be much smaller, and suggests that perhaps more efficient constructions exist. However, examples in [13] show that sometimes the core of a dual can have exponential size. In fact, in [7] it is shown that most paths have exponential-sized dual cores.

Now $\tilde{Pell}_n$ has more than $2^n$ arcs, so the core $\tilde{D}(\tilde{Pell}_n)$ may have in the order of $2^{2^n}$ vertices. If this is the case each of these vertices needs an exponential representation, and we arrive at the strange conclusion that we cannot efficiently describe a homomorphism $\phi : \tilde{G} \to \tilde{D}(\tilde{Pell}_n)$ even when we can easily prove that such a homomorphism exists. For the moment, we cannot prove that $\tilde{D}(\tilde{Pell}_n)$ is indeed double exponential, but in the next section we present a variant of the Pell paths for which this holds.

6 Large paths with very large duals

We first introduce variants of Pell paths with blue-red coloured arcs: Put

\begin{align*}
\bar{B}_1 &= \bar{P}_1 \text{ with its arc coloured blue,} \\
\bar{R}_1 &= \bar{P}_1 \text{ with its arc coloured red,} \\
\bar{B}_2 &= \bar{P}_2 \text{ with its arcs coloured blue,} \\
\bar{R}_2 &= \bar{P}_2 \text{ with its arcs coloured red,}
\end{align*}
\[
\tilde{B}_{n+1} = \tilde{B}_n \circ \tilde{R}_{n-1}^{-1} \circ \tilde{B}_n,
\]
\[
\tilde{R}_{n+1} = \tilde{R}_n \circ \tilde{B}_{n-1}^{-1} \circ \tilde{R}_n.
\]

Algebraically, digraphs with blue-red coloured arcs are relational structures with two binary relations. The homomorphisms between these structures are the maps that preserve the colour and the orientation of the arcs. It is known (see [11, 13]) that every such (finite) structure has a core which is unique up to isomorphism and that every relational “tree” admits a dual.

**Lemma 10** For \( n \geq 4 \), the cores \( \tilde{D}(\tilde{B}_n) \) and \( \tilde{D}(\tilde{R}_n) \) each have at least \( 2^{2n}/n \) vertices.

**Proof.** We will prove the result for \( \tilde{D}(\tilde{B}_n) \), the proof is identical for \( \tilde{D}(\tilde{R}_n) \). The algebraic length of \( \tilde{B}_n \) is \( \text{al}(\tilde{P}_{\text{Pell}}) = n \), but the number of vertices of these paths grows exponentially, so there exists a height \( k_n \) such that the number of vertices of height \( k_n \) in \( \tilde{B}_n \) grows exponentially. It is easy to see that for \( n \geq 6 \), \( \tilde{B}_n \) has at least \( 2^n + 2 \) vertices and that the largest level has at least \( 2^n/n \) vertices, and it can be checked that the latter holds also for \( n = 4, 5 \).

For a vertex \( x \) of \( \tilde{B}_n \) we will denote \( [0, x] \) the subpath of \( \tilde{B}_n \) containing all its vertices from the first to \( x \), and \( [x, \infty] \) the subpath of \( \tilde{B}_n \) containing all its vertices from \( x \) to the last. Let \( M = \{ x_1, \ldots, x_m \}, m \geq 2^n/n \) be the set of vertices of \( \tilde{B}_n \) of height \( k_n \). For \( S \subseteq M \), we construct the (relational) tree \( \tilde{T}_S \) with root \( r_S \) as follows:

- For every \( x_i \in S \) we put a copy of \( [0, x_i] \) in \( \tilde{T}_S \), identifying the copy of \( x_i \) to \( r_S \),
- For every \( x_i \notin S \) we put a copy of \( [x_i, \infty] \) in \( \tilde{T}_S \), identifying the copy of \( x_i \) to \( r_S \).

We can then prove that \( \tilde{B}_n \not\rightarrow \tilde{T}_S \). Indeed, if \( S = \emptyset \) or \( S = M \), the algebraic length of \( \tilde{T}_S \) is \( n-1 \), which smaller than that of \( \tilde{B}_n \). For other values of \( S \), the algebraic length of \( \tilde{T}_S \) is exactly \( n \). Therefore a homomorphism \( \phi : \tilde{B}_n \rightarrow \tilde{T}_S \) would have to map the first vertex of \( \tilde{B}_n \) (of height 0) to the first vertex of some \( [0, x_i] \) in \( \tilde{T}_S \), and the last vertex of \( \tilde{B}_n \) (of height \( n \)) to the last vertex of some \( [x_j, \infty] \) in \( \tilde{T}_S \). However, in \( \tilde{B}_n \) all the blue arcs are left to right, and all the red arcs are right to left. Thus \( \phi \) cannot identify vertices and induces an isomorphism between \( \tilde{B}_n \) and its copy \( \phi(\tilde{B}_n) = [0, x_i] \cup [x_j, \infty] \) in \( \tilde{T}_S \). Therefore \( x_i = x_j \), and this vertex is both in \( S \) and out of \( S \), which is impossible.
Thus, $\tilde{B}_n \not\rightarrow \tilde{T}_S$ hence there exists a homomorphism $\psi_S : \tilde{T}_S \rightarrow \tilde{D}(\tilde{B}_n)$ for all $S$. We show that for $S \neq S'$, we have $\psi_S(r_S) \neq \psi_{S'}(r_{S'})$. Indeed, without loss of generality, there exists $x_i \in S' \setminus S$, so that $S'$ contains a copy of $[0, x_i]$ with $x_i$ identified to $r_{S'}$, and $S$ contains a copy of $[x_i, \infty]$ with $x_i$ identified to $r_S$. If $\psi_S(r_S) = \psi_{S'}(r_{S'})$, then $\psi_S|_{[0, x_i]} \cup \psi_{S'}|_{[x_i, \infty]}$ is a homomorphism from a copy of $\tilde{B}_n$ to $\tilde{D}(\tilde{B}_n)$, which is impossible.

Therefore, the vertices $\psi_S(r_S) : S \subseteq M$ are all distinct in $\tilde{D}(\tilde{B}_n)$, whence the number of vertices of $\tilde{D}(\tilde{B}_n)$ is at least $2^m \geq 2^{2^n/n}$.

Using the arrow (or indicator) construction, we can then use these examples to find directed graphs with the same properties. For instance, to every structure $\tilde{X}$ with blue-red coloured arcs, we can associate a digraph $g(\tilde{X})$ by replacing every blue arc by a copy of $\tilde{P}_3 \circ \tilde{P}_1^{-1} \circ \tilde{P}_2$ and every red arc by a copy of $\tilde{P}_2 \circ \tilde{P}_1^{-1} \circ \tilde{P}_3$.

**Corollary 11** Let $\tilde{Q}_n = g(\tilde{B}_n)$. Then for $n \geq 4$, $\tilde{D}(\tilde{Q}_n)$ has at least $2^{2^n/n}$ vertices.

**Proof.** For every tree $\tilde{T}_S : S \subseteq M$ of the proof of Lemma 10, we have $\tilde{B}_n \not\rightarrow \tilde{T}_S$. It is not hard to infer from this that $\tilde{Q}_n = g(\tilde{B}_n) \not\rightarrow g(\tilde{T}_S)$. Indeed, a homomorphism $\phi : g(\tilde{B}_n) \not\rightarrow g(\tilde{T}_S)$ would have to map vertices whose height is a multiple of 4 into vertices whose height is a multiple of 4, and these correspond to vertices of $\tilde{B}_n$ and $\tilde{T}_S$ respectively. Thus $\phi$ would induce a map $\phi' : \tilde{B}_n \rightarrow \tilde{T}_S$, easily seen to be a homomorphism.

Thus, for every $S \subseteq M$ there exists a homomorphism $\phi_S : \tilde{T}_S \rightarrow \tilde{D}(\tilde{Q}_n)$, and as in the proof of Lemma 10 we have $\phi_S(r_S) \neq \phi_{S'}(r_{S'})$ when $S \neq S'$. Therefore $\tilde{D}(\tilde{Q}_n)$ has at least $2^{\vert M \vert}$ vertices.

As in Proposition 9, we can efficiently test the existence of a homomorphism from $\tilde{G}$ to $\tilde{D}(\tilde{Q}_n)$ using the matrices $A^3 \cdot A^T \cdot A^2$ and $A^2 \cdot A^T \cdot A^3$ as building blocks for recursive multiplication. Corollary 11 shows that we cannot efficiently describe the corresponding homomorphisms when we prove that they exist.

### 7 Concluding comments and problems

We still do not know whether there exists an infinite sequence $\{\tilde{B}_n\}_{n \geq 1}$ of paths such that $\{\text{al}({\tilde{B}_n})\}_{n \geq 1}$ is unbounded and $\{\chi(\tilde{B}_n)\}_{n \geq 1}$ is bounded by a constant $b$. Perhaps this is already true of the Pell paths of Section 5. In any case, by Corollary 8, for any such sequence $\{\tilde{B}_n\}_{n \geq 1}$ of paths, the sequence $\{\text{al}({\tilde{B}_n})\}_{n \geq 1}$ would also be unbounded. Furthermore the condition would need to be hereditary: every sequence of subpaths of $\{\tilde{B}_n\}_{n \geq 1}$
with unbounded algebraic length would also have unbounded derived algebraic length. Sequences of paths fulfilling these conditions tend to have exponential growth. If such a sequence \( \{\vec{B}_n\}_{n \geq 1} \) of exponential growth exists, Section 5 suggests that the condition \( \vec{B}_n \not\rightarrow \vec{G} \) may still be verifiable in polynomial time, while Section 6 indicates that the corresponding homomorphism \( \phi : \vec{G} \rightarrow \vec{D}(\vec{B}_n) \) may not be describable in polynomial time. Thus if the colour of a vertex \( \phi(u) \in \vec{D}(\vec{B}_n) \) depends on the whole exponential information describing it, we end up with a polynomial \( b \)-colourability certificate for \( G \) which does not yield a \( b \)-colouring. However it is still the first step in this chain of speculations that is outstanding:

**Problem 1.** Let

\[
f(n) = \min \{ \chi(\vec{D}(\vec{P})) : \text{al}(\vec{P}) = n \}.
\]

Is \( f \) bounded or unbounded?

Perhaps the fractional version of this question is easier to solve. In fact, since the shift graphs have bounded fractional chromatic number, we cannot even conclude that the duals of the paths \( \vec{S}_n \) and \( \vec{Z}_n \) of Sections 2 and 3 have unbounded fractional chromatic numbers. Linear programming duality has proved to be a powerful tool and often progress on the fractional version of a problem is much faster than on the nonfractional version. Therefore it seems reasonable to consider the following.

**Problem 2.** Does there exist a polynomial algorithm which inputs a path \( \vec{P} \) and outputs the value \( \chi_f(\vec{D}(\vec{P})) \)?

**References**


