Simple permutations: decidability and unavoidable substructures

Robert Brignall [a] Nik Ruškuc [a] Vincent Vatter [a,*,1]

[a] *University of St Andrews, School of Mathematics and Statistics, St Andrews, Fife, KY16 9SS, UK*

Abstract

We prove that it is decidable whether a finitely based permutation class contains infinitely many simple permutations, and establish an unavoidable substructure result for simple permutations: every sufficiently long simple permutation contains an alternation or oscillation of length k.

Key words: permutation class, restricted permutation, simple permutation
1991 MSC: 03B25, 03D05, 05A15, 68F10, 68Q45

1 Introduction

Simple permutations are the building blocks of permutation classes. As such, classes with only finitely many simple permutations, e.g., the class of 132-avoiding permutations, have nice properties. To name three: these classes have algebraic generating functions (as established by Albert and Atkinson [1]; see Brignall, Huczynska, and Vatter [9] for extensions), are partially well-ordered (see the conclusion), and are finitely based [1]. It is natural then to ask which finitely based classes contain only finitely many simple permutations. Our main result establishes that this can be done algorithmically:

Theorem 1 *It is possible to decide whether a permutation class given by a finite basis contains infinitely many simple permutations.*

* Corresponding author.
Email address: vince@mcs.st-and.ac.uk (Vincent Vatter).
1 Supported by EPSRC grants GR/S53503/01 and EP/C523229/1.

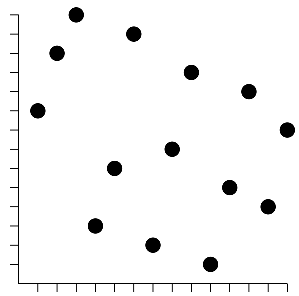

Fig. 1. The plot of a simple permutation.

Permutation classes. Two sequences $u_1, \ldots, u_k$ and $w_1, \ldots, w_k$ of distinct real numbers are said to be *order isomorphic* if they have the same relative comparisons, that is, if $u_i < u_j$ if and only if $w_i < w_j$. The permutation π is said to *contain* the permutation σ, written $\sigma \leq \pi$, if π has a subsequence that is order isomorphic to σ; otherwise π is said to *avoid* σ. For example, $\pi = 891367452$ contains $\sigma = 51342$, as can be seen by considering the subsequence 91672 $(= \pi(2), \pi(3), \pi(5), \pi(6), \pi(9))$. This pattern-containment relation is a partial order on permutations. We refer to downsets of permutations under this order as *permutation classes*. In other words, if $\mathcal{C}$ is a permutation class, $\pi \in \mathcal{C}$, and $\sigma \leq \pi$, then $\sigma \in \mathcal{C}$. We denote by $\mathcal{C}_n$ the set $\mathcal{C} \cap S_n$, i.e. the permutations in $\mathcal{C}$ of length n, and we refer to $\sum |\mathcal{C}_n| x^n$ as the *generating function for* $\mathcal{C}$. Recall that an *antichain* is a set of pairwise incomparable elements. For any permutation class $\mathcal{C}$, there is a unique (possibly infinite) antichain B such that $\mathcal{C} = \text{Av}(B) = \{\pi : \beta \not\leq \pi \text{ for all } \beta \in B\}$. This antichain B, which consists of the minimal permutations not in $\mathcal{C}$, is called the *basis* of $\mathcal{C}$. Permutation classes arise naturally in a variety of settings, ranging from sorting (see, e.g., Bóna's survey [6]) to algebraic geometry (see, e.g., Lakshmibai and Sandhya [15]).

It will also be useful to have a pictorial description of order isomorphism. Two sets S and T of points in the plane are said to be order isomorphic if the axes can be stretched and shrunk in some manner to map one of the sets onto the other, i.e., if there are strictly increasing functions $f, g : \mathbb{R} \to \mathbb{R}$ such that $\{(f(s_1), g(s_2)) : (s_1, s_2) \in S\} = T$. (As the inverse of a strictly increasing function is also strictly increasing, this is an equivalence relation). The *plot* of the permutation π is the point set $\{(i, \pi(i))\}$, and every finite point set in the plane in which no two points share a coordinate (often called a *generic* or *noncorectilinear* set) is order isomorphic to the plot of a unique permutation; in practice we simply say that a generic point set is order isomorphic to a permutation. This geometric viewpoint makes clear (if they were not already) several symmetries of the pattern-containment order. The maps $(x, y) \mapsto (-x, y)$ and $(x, y) \mapsto (y, x)$, which when applied to generic point sets correspond to reversing and inverting permutations, generate a dihedral group with eight elements.

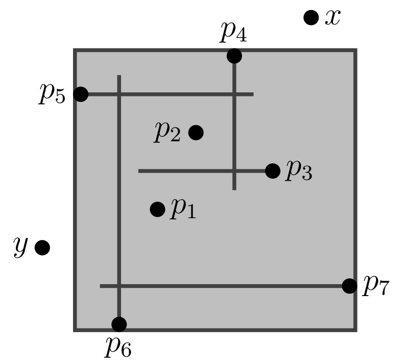

Fig. 2. The points $p_1, \ldots, p_7$ form a proper pin sequence and the gray box denotes $\mathrm{rect}(p_1, \ldots, p_7)$. The point x satisfies the externality and separation conditions for this pin sequence and thus could be chosen as p_8; y, however, fails the separation condition.

Simple permutations. An *interval* in the permutation π is a set of contiguous indices $I = [a, b]$ such that the set of values $\pi(I) = \{\pi(i) : i \in I\}$ also forms an interval of natural numbers. Every permutation π of $[n] = \{1, 2, \ldots, n\}$ has intervals of size 0, 1, and n; π is said to be *simple* if it has no other intervals. Note that simplicity is preserved under the eight symmetries mentioned above. Figure 1 shows the plot of a simple permutation.

We need several notions from [8]. Given points $p_1, \ldots, p_m$ in the plane, we denote by $\mathrm{rect}(p_1, \ldots, p_m)$ the smallest axes-parallel rectangle containing them. A *pin* for the points $p_1, \ldots, p_m$ is any point p_{m+1} not contained in $\mathrm{rect}(p_1, \ldots, p_m)$ that lies either horizontally or vertically amongst them. A *proper pin sequence* is a sequence of points in the plane satisfying two conditions for all i:

- *Externality condition*: p_{i+1} must lie outside $\mathrm{rect}(p_1, \ldots, p_i)$. (Note that this forces the pins to be distinct.)
- *Separation condition*: p_{i+1} must *separate* p_i from $\{p_1, \ldots, p_{i-1}\}$. That is, p_{i+1} must lie horizontally or vertically between $\mathrm{rect}(p_1, \ldots, p_{i-1})$ and p_i. (Note that this condition is vacuous for $i = 0, 1$.)

Figure 2 illustrates these definitions. The astute reader may note that we have replaced the maximality condition of [8] with the externality condition. This change reflects the differing viewpoints of the papers; while [8] was concerned with finding proper pin sequences in permutations, we will be building proper pin sequences from scratch, and in this context the externality and separation conditions together imply the maximality condition.

Proper pin sequences are intimately connected with simple permutations. In one direction, we have:

Theorem 2 (Brignall, Huczynska, and Vatter [8]) *If $p_1, \ldots, p_m$ is a proper*

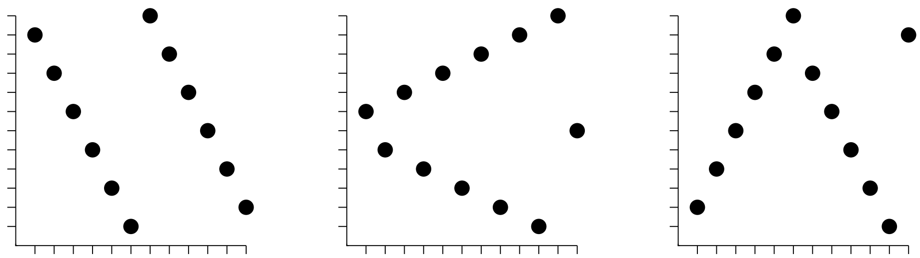

Fig. 3. From left to right: a parallel alternation (which in this case happens to be simple), a wedge simple permutation of type 1, and a wedge simple permutation of type 2.

pin sequence and $m \geq 5$ *then one of the sets of points* $\{p_1, \ldots, p_m\}$, $\{p_1, \ldots, p_m\} \setminus \{p_1\}$, *or* $\{p_1, \ldots, p_m\} \setminus \{p_2\}$ *is order isomorphic to a simple permutation.*

While proper pin sequences are simple or nearly so, there are other types of simple permutation. We call attention to three families, plotted in Figure 3. An *alternation* is a permutation in which every odd entry lies to the left of every even entry, or any symmetry of such a permutation. A *parallel alternation* is one in which these two sets of entries form monotone subsequences, either both increasing or both decreasing. A *wedge alternation* is one in which the two sets of entries form monotone subsequences pointing in opposite directions. Whereas every parallel alternation contains a long simple permutation (to form this simple permutation we need, at worst, to remove two points), wedge alternations do not. However, there are two different ways to add a single point to a wedge alternation to form a simple permutation (called *wedge simple permutations of types* 1 *and* 2).

These families of permutations capture, in a sense formalised below, the diversity of simple permutations.

Theorem 3 (Brignall, Huczynska, and Vatter [8]) *For any fixed* k, *every sufficiently long simple permutation contains either a proper pin sequence of length at least* k, *a parallel alternation of length at least* k, *or wedge simple permutation of length at least* k.

Theorems 2 and 3 show that Theorem 1 will follow if we can decide when a class has arbitrarily long parallel alternations, wedge simple permutations, and proper pin sequences.

2 The Easy Decisions

We begin by describing how to decide if a permutation class given by a finite basis contains arbitrarily long parallel alternations or wedge simple permutations. Consider first the case of parallel alternations, oriented \\, as in Figure 3.

These alternations nearly form a chain in the pattern-containment order; precisely, there are two such parallel alternations of each length, and each of these contains a parallel alternation with one fewer point and all shorter parallel alternations of the same orientation. Thus if the permutation class $\mathcal{C}$ has a basis element contained in any of these parallel alternations, it will contain only finitely many of them. Conversely, if $\mathcal{C}$ has no such basis element, it will contain all of these alternations. Therefore we need to characterise the permutations that are contained in any parallel alternation. To do so, we must first review juxtapositions of classes.

Given two permutation classes $\mathcal{C}$ and $\mathcal{D}$, their *horizontal juxtaposition*, $\begin{bmatrix}\mathcal{C} & \mathcal{D}\end{bmatrix}$, consists of all permutations π that can be written as a concatenation $\sigma\tau$ where σ is order isomorphic to a permutation in $\mathcal{C}$ and τ is order isomorphic to a permutation in $\mathcal{D}$, or in other words, those permutations π whose plot can be divided into two parts by a single vertical line so that the points to the left of the line are order isomorphic to a member of $\mathcal{C}$ while the points to the right of the line are order isomorphic to a member of $\mathcal{D}$.

Proposition 4 (Atkinson [4]) *Let $\mathcal{C}$ and $\mathcal{D}$ be permutation classes. The basis elements of the class $\begin{bmatrix}\mathcal{C} & \mathcal{D}\end{bmatrix}$ can all be written as concatenations $\rho\sigma\tau$ where either:*

- *σ is empty, ρ is order isomorphic to a basis element of $\mathcal{C}$, and τ is order isomorphic to a basis element of $\mathcal{D}$, or*
- *$|\sigma| = 1$, $\rho\sigma$ is order isomorphic to a basis element of $\mathcal{C}$, and $\sigma\tau$ is order isomorphic to a basis element of $\mathcal{D}$.*

(In particular, if two classes are finitely based then their juxtaposition is also finitely based.)

There is an obvious symmetry to this operation: the *vertical juxtaposition* of the classes $\mathcal{C}$ and $\mathcal{D}$, denoted $\begin{bmatrix}\mathcal{C} \\ \mathcal{D}\end{bmatrix}$, consists of those permutations π whose plot can be divided into two parts by a single horizontal line so that the points above the line are order isomorphic to a member of $\mathcal{C}$ while the points below the line are order isomorphic to a member of $\mathcal{D}$.

Proposition 4 is all we need to solve the parallel alternation decision problem.

Proposition 5 *The permutation class* $\mathrm{Av}(B)$ *contains only finitely many parallel alternations if and only if B contains an element of every symmetry of the class* $\mathrm{Av}(123, 2413, 3412)$.

Proof. The set of permutations that are contained in at least one (and thus,

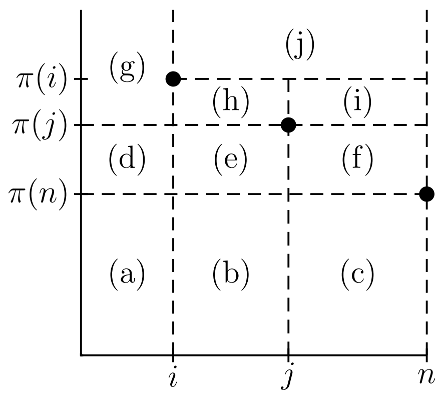

Fig. 4. The situation in the proof of Proposition 6.

all but finitely many) parallel alternation(s) oriented \\ is

$$\begin{bmatrix} \text{Av}(12) & \text{Av}(12) \end{bmatrix} = \text{Av}(123, 2413, 3412),$$

as desired. $\square$

Like parallel alternations, the wedge simple permutations of a given type and orientation also nearly form a chain in the pattern-containment order, and thus we are able to take much the same approach with them.

Proposition 6 *The permutation class* $\text{Av}(B)$ *contains only finitely many wedge simple permutations of type* 1 *if and only if* B *contains an element of every symmetry of the class*

$$\text{Av}(1243, 1324, 1423, 1432, 2431, 3124, 4123, 4132, 4231, 4312).$$

Proof. The wedge simple permutations of type 1 that are oriented $<$, as in Figure 3, are contained in

$$\begin{aligned} \begin{bmatrix} \begin{bmatrix} \text{Av}(21) \\ \text{Av}(12) \end{bmatrix} & \{1\} \end{bmatrix} &= \begin{bmatrix} \text{Av}(132, 312) & \text{Av}(12, 21) \end{bmatrix} \\ &= \text{Av}(1324, 1423, 1432, 2431, 3124, 4123, 4132, 4231). \end{aligned}$$

It is easy to see that these wedge simple permutations also avoid 1243 and 4312, and thus they are contained in the class stated in the proposition, which we call $\mathcal{D}$.

Now take a permutation $\pi \in \mathcal{D}$ of length n. We would like to show that π is contained in a wedge simple permutation. If $\pi \in \begin{bmatrix} \text{Av}(21) \\ \text{Av}(12) \end{bmatrix}$ then π is clearly contained in a wedge simple permutation, so suppose this is not the case. Thus $\pi(1)\cdots\pi(n-1)$ is order isomorphic to a permutation in $\begin{bmatrix} \mathrm{Av}(21) \\ \mathrm{Av}(12) \end{bmatrix}$, and it suffices to show that:

- the entries of π above $\pi(n)$ are increasing, and
- the entries of π below $\pi(n)$ are decreasing.

We prove the first of these items; the second then follows by symmetry because it can be observed from its basis that $\mathcal{D}$ is invariant under complementation, i.e., if the length n permutation π lies in $\mathcal{D}$ then so does the permutation π^c defined by $\pi^c(i) = n+1-\pi(i)$. Suppose to the contrary that there is a descent above $\pi(n)$. Thus there are indices $i < j < n$ such that $\pi(i) > \pi(j) > \pi(n)$. Choose these two indices to be lexicographically minimal with this property. There must be other entries of π as otherwise π is simply 321, which lies in the juxtaposition we have assumed π does not lie in. We now divide the entries above $\pi(n)$ into 7 regions as shown in Figure 4. About these regions we can state:

- regions (a)–(e) and (i) are empty because π avoids 1432, 4132, 4312, 2431, 4231, and 4231, respectively;
- the points in region (f) are decreasing because π avoids 4231;
- regions (g) and (h) are empty by the minimality of i and j, respectively;
- the points in region (j) are increasing because π avoids 2431.

This establishes that π lies in $\begin{bmatrix} \mathrm{Av}(21) \\ \mathrm{Av}(12) \end{bmatrix}$, a contradiction that completes the proof. $\square$

Proposition 7 *The permutation class* Av(B) *contains only finitely many wedge simple permutations of type* 2 *if and only if* B *contains an element of every symmetry of the class*

$$\mathrm{Av}(2134, 2143, 3124, 3142, 3241, 3412, 4123, 4132, 4231, 4312).$$

Proof. Let $\mathcal{D}$ denote the class in the statement of the proposition. It is clear that the wedge simple permutation of type 2 that are oriented Λ, as in Figure 3, lie in $\mathcal{D}$, and so it remains to show that every permutation $\pi \in \mathcal{D}$ is contained in one of these wedge simple permutations. Thus π is contained in

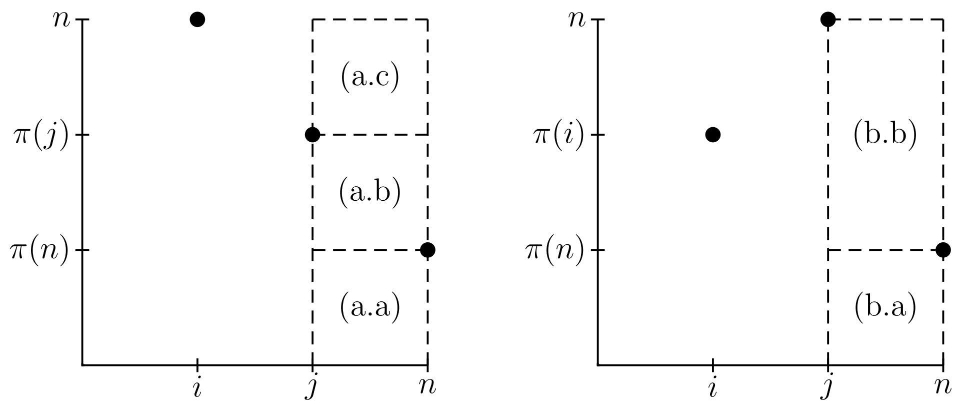

Fig. 5. The two situations in the proof of Proposition 7.

$$\left[\text{Av}(21)\ \text{Av}(12)\ \{1\}\right] = \left[\text{Av}(213,312)\ \text{Av}(12,21)\right]$$
$$= \text{Av}(2134, 2143, 3124, 3142, 3241, 4123, 4132, 4231),$$

and so in particular, the permutation obtained by removing the rightmost element of π, say $\pi(n)$, is contained in $\left[\text{Av}(21)\ \text{Av}(12)\right]$. It suffices to show that $\pi(n)$ is n or $n-1$. Suppose, to the contrary, that there are at least two entries of π above $\pi(n)$. Then we have one of the two situations depicted in Figure 5.

Again, we use the basis elements of $\mathcal{D}$ to derive the following about the labelled regions:

- regions (a.a), (a.c), and (b.a) are empty because π avoids 4312, 4231, and 3412, respectively;
- the points in regions (a.b) and (b.b) are decreasing because π avoids 4231.

These observations, combined with the fact that the permutation obtained from π by removing $\pi(n)$ lies in $\left[\text{Av}(21)\ \text{Av}(12)\right]$ shows that π itself lies in $\left[\text{Av}(21)\ \text{Av}(12)\right]$, and so π is contained in one of the desired wedge simple permutations, completing the proof. □

3 Pin Words

This leaves only proper pin sequences. Proper pin sequences, as well as subsets of proper pin sequences, can be described naturally, if not uniquely, by words over the eight-letter alphabet consisting of the numerals $\{1,2,3,4\}$ and directions $\{L,R,U,D\}$ (standing for left, right, up, and down). In this section

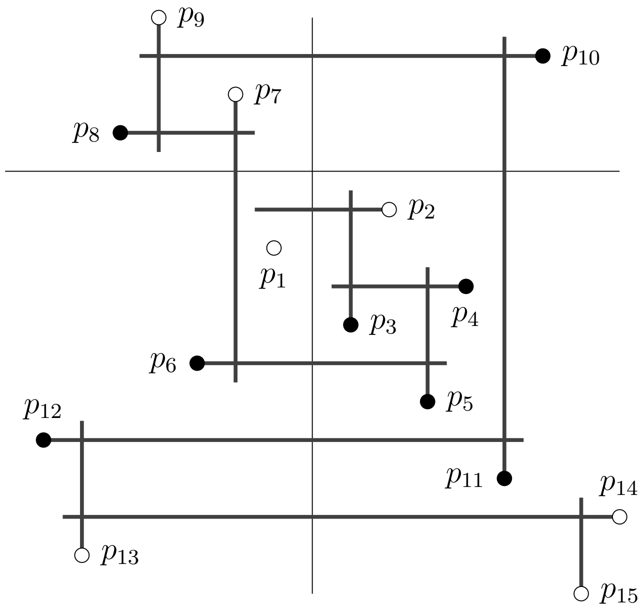

Fig. 6. The proper pin sequence $p_1, \ldots, p_{15}$ shown corresponds to the strict pin word $w = 3RDRDLULURDLDRD$. The filled points correspond to the pin word $u = 4RDL21DL$, the permutation corresponding to this word, i.e., the permutation order isomorphic to the filled points, is 27453618.

we study these words, laying the groundwork for the proof of our main result in the fourth.

The word $w = w_1 \cdots w_m \in \{1, 2, 3, 4, L, R, U, D\}^*$ is a *pin word* if it satisfies:

(W1) w begins with a numeral,
(W2) if $w_{i-1} \in \{L, R\}$ then $w_i \in \{1, 2, 3, 4, U, D\}$, and
(W3) if $w_{i-1} \in \{U, D\}$ then $w_i \in \{1, 2, 3, 4, L, R\}$.

Pin words with precisely one numeral, which we term *strict pin words*, correspond to proper pin sequences and it is this correspondence we describe first. Let $w = w_1 \cdots w_m$ denote a strict pin word and begin by placing a point p_1 in quadrant w_1. Next take p_2 to be a pin in the direction w_2 that separates p_1 from the origin, denoted 0. Continue in this manner, taking p_{i+1} to be a pin in the direction w_{i+1} that satisfies the externality condition and separates p_i from $0, p_1, \ldots, p_{i-1}$. See Figure 6 for an example. Upon completion, $0, p_1, \ldots, p_m$ is a proper pin sequence, and more importantly, $p_1, \ldots, p_m$ is as well; it is the latter pin sequence that we say *corresponds* to w. Note that not only is this sequence unique up to order isomorphism[2], but also the quadrant that point p_i lies in is determined by w (indeed, for $i \geq 2$, this quadrant is determined by w_{i-1} and w_i). We say that the *permutation corresponding to w* is the permutation that is order isomorphic to the set of points $p_1, \ldots, p_m$. Conversely,

[2] It is for this reason that we refer to it as *the* proper pin sequence corresponding to w.

we have the following result.

Lemma 8 *Every proper pin sequence corresponds to at least one strict pin word.*

Proof. Let $p_1, \ldots, p_m$ be a proper pin sequence in the plane. It suffices to place a point p_0 (corresponding to the origin) so that $p_0, p_1, \ldots, p_m$ form a proper pin sequence, as then the pin word for this sequence can be read. By symmetry, let us assume that p_1 lies below and to the right of p_2 and that p_3 is a left or right pin. Hence p_3 lies vertically between p_1 and p_2, and by the separation condition, p_3 is the only such pin. We place p_0 vertically between p_1 and p_3 and minimally to the left of p_2, i.e., so that no pin lies horizontally between p_2 and p_0. Clearly p_2 separates p_1 from p_0 while p_3 separates p_2 from $\{p_0, p_1\}$. Moreover, our placement of p_0 guarantees that no later pins separate $\{p_0, p_1, p_2\}$, so since p_{i+1} separates p_i from $\{p_1, \ldots, p_{i-1}\}$, it will also separate p_i from $\{p_0, p_1, \ldots, p_{i-1}\}$. $\square$

It remains to construct the permutations that correspond to nonstrict pin words. Letting $w = w_1 \cdots w_m$ denote such a word, we begin as before. Upon reaching a later numeral, say w_i, we essentially collapse $p_1, \ldots, p_{i-1}$ into the origin and begin anew. More precisely, we place p_i in quadrant w_i so that it does not separate any of $0, p_1, \ldots, p_{i-1}$. If w_{i+1} is a direction, we take p_{i+1} to be a pin in the direction w_{i+1} that satisfies the externality condition and separates p_i from $0, p_1, \ldots, p_{i-1}$; if w_{i+1} is a numeral then we again place p_{i+1} in quadrant w_{i+1} so that it does not separate any of the former points. In this process we build a *sequence of points corresponding to* w: $p_1, \ldots, p_m$. Again, see Figure 6 for an example. As is the case with strict pin words, this sequence of points is unique up to order isomorphism, and we define the *permutation corresponding to* w to be the permutation order isomorphic to this set of points.

We now define a partial order, $\preceq$, on pin words. Let u and w be two pin words. We define a *strong numeral-led factor* to be a sequence of contiguous letters beginning with a numeral and followed by any number of directions (but no numerals) and begin by writing u in terms of its strong numeral-led factors as $u = u^{(1)} \cdots u^{(j)}$. We then write $u \preceq w$ if w can be chopped into a sequence of factors $w = v^{(1)} w^{(1)} \cdots v^{(j)} w^{(j)} v^{(j+1)}$ such that for all $i \in [j]$:

(O1) if $w^{(i)}$ begins with a numeral then $w^{(i)} = u^{(i)}$, and
(O2) if $w^{(i)}$ begins with a direction, then $v^{(i)}$ is nonempty, the first letter of $w^{(i)}$ corresponds (in the manner described above) to a point lying in the quadrant specified by the first letter of $u^{(i)}$, and all other letters (which must be directions) in $u^{(i)}$ and $w^{(i)}$ agree.

(It is trivial to check that $\preceq$ is reflexive and antisymmetric; transitivity requires only slightly more effort.) Returning a final time to Figure 6, the division of

u into strong numeral-led factors is $(4RDL)(2)(1DL)$, while w can be written as $(3R)(DRDL)(U)(L)(U)(RDL)(DRD)$. We now match factors. Since w_3 corresponds to p_3 which lies in quadrant 4, $(4RDL)$ can embed as $(DRDL)$; because p_8 lies in quadrant 2, the (2) factor in u can embed as (L); lastly, p_{10} lies in quadrant 1, so the $(1DL)$ factor in u can embed as (RDL) in w. This verifies that $u \preceq w$.

This order is not merely a translation of the pattern-containment order on permutations (consider the words $11, 13, 1L, 1D, 21, 23, 2R, 2U, \ldots$, which are incomparable under $\preceq$ yet correspond to the same permutation), but $\leq$ and $\preceq$ are closely related:

Lemma 9 *If the pin word w corresponds to the permutation π and $\sigma \leq \pi$ then there is a pin word u corresponding to σ with $u \preceq w$. Conversely, if $u \preceq w$ then the permutation corresponding to u is contained in the permutation corresponding to w.*

Proof. If $w = w_1 \cdots w_m$ corresponds to the sequence of points $p_1, \ldots, p_m$ then the sequence $p_1, \ldots, p_{\ell-1}, p_{\ell+1}, \ldots, p_m$ corresponds to the pin word $w_1 \cdots w_{\ell-1} w'_{\ell+1} w_{\ell+2} \cdots w_m \preceq w$, where $w'_{\ell+1}$ is the numeral corresponding to the quadrant containing $p_{\ell+1}$. Iterating this observation proves the first half of the lemma.

The other direction follows similarly. Write u in terms of its strong numeral-led factors as $u = u^{(1)} \cdots u^{(j)}$ and suppose that the expression $w = v^{(1)} w^{(1)} \cdots v^{(j)} w^{(j)} v^{(j+1)}$ satisfies (O1) and (O2). Now delete every point in the sequence of points corresponding to w that comes from a letter in a $v^{(i)}$ factor. By conditions (O1) and (O2) and the remarks in the previous paragraph, it follows that the resulting sequence of points corresponds to u. Therefore the permutation corresponding to u is contained in the permutation corresponding to w. □

4 Brief Review of Regular Languages and Automata

The classic results mentioned here are covered more comprehensively in many texts, for example, Hopcroft, Motwani, and Ullman [14], so we give only the barest details.

A *nondeterministic finite automaton* over the alphabet A consists of a set S of *states*, one of which is designated the *initial state*, a *transition function* δ from $S \times (A \cup \{\varepsilon\})$ into the power set of S, and a subset of S designated as *accept states*. The *transition diagram* for this automaton is a directed graph on the vertices S, with an arc from r to s labelled by a precisely if $s \in \delta(r, a)$. The initial state is designated by an inward-pointing arrow. An automaton

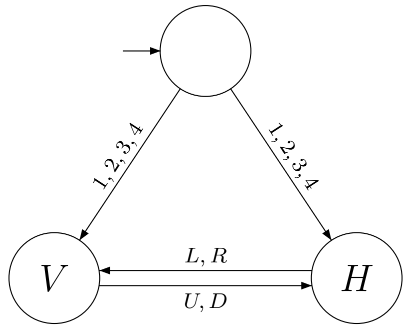

Fig. 7. An automaton that accepts the language of strict pin words (V and H are accept states).

accepts the word $w_1 \cdots w_m$ if there is a walk from the initial state to an accept state whose arcs are labelled (in order) by $w_1, \ldots, w_m$; the set of all such words is the *language accepted* by the automaton. For example, Figure 7 shows the transition diagram for an automaton that accepts strict pin words.

A language that is accepted by a finite automaton is called *recognisable*. By Kleene's theorem, the recognisable languages are precisely the *regular languages* [3], and they have numerous closure properties, of which we use two: the union of two regular languages and the set-theoretic difference of two regular languages are also regular languages. The other result we need about regular languages is below.

Proposition 10 *It can be decided whether a regular language given by a finite accepting automaton is infinite.*

Sketch of proof. A regular language is infinite if and only if one can find a walk in the given accepting automaton that begins at the initial state, contains a directed cycle, and ends at an accept state. □

A *finite transducer* is a finite automaton that can both read and write. Transducers also have states, S, one of which is designated the initial state and several may be designated accept states. The transition function for a transducer over the alphabet A is a map from $S \times (A \cup \{\varepsilon\}) \times (A \cup \{\varepsilon\})$ into the power set of A. In the transition diagram of a transducer we label arcs by pairs, so the transition $r \xrightarrow{a,b} s$ stands for "read a, write b". Empty inputs and outputs are allowed, both designated by ε, e.g., $r \xrightarrow{\varepsilon,b} s$ means "read nothing, write b". A word $w \in A^*$ is *produced* from the word $u \in A^*$ by the transducer T if there is a walk

$$s_1 \xrightarrow{u_1,w_1} s_2 \xrightarrow{u_2,w_2} s_3 \cdots \xrightarrow{u_m,w_m} s_{m+1}$$

[3] The reader unfamiliar with formal languages is welcomed to take this as the definition of regular languages.

in the transition diagram of T beginning at the initial state, ending at an accept state, and such that $u = u_1 \cdots u_m$ and $w = w_1 \cdots w_m$ (note that these u_i's and w_i's are allowed to be ε). We denote the set of words that the transducer T produces from set of input words $\mathcal{L}$ by $T(\mathcal{L})$.

Proposition 11 *If $\mathcal{L}$ is a regular language and T is a finite transducer then $T(\mathcal{L})$ is also regular, and a finite accepting automaton for $T(\mathcal{L})$ can be effectively constructed.*

Sketch of proof. Let M denote a finite accepting automaton for $\mathcal{L}$. Suppose that the states of M are R and the states of T are S. The states of an accepting automaton for $T(\mathcal{L})$ are then $R \times S$, where there is a transition $(r_1, s_1) \xrightarrow{b} (r_2, s_2)$ whenever there are transitions $r_1 \xrightarrow{a} r_2$ and $s_1 \xrightarrow{a,b} s_2$ in M and T, respectively. $\square$

5 Decidability

We are now in a position to prove our main result. We wish to decide whether the class $\mathrm{Av}(B)$, where B is finite, contains only finitely many simple permutations. Propositions 5–7 show how to decide if $\mathrm{Av}(B)$ contains arbitrarily long parallel alternations or wedge simple permutations, so by Theorem 3 it suffices to decide if $\mathrm{Av}(B)$ contains arbitrarily long proper pin sequences.

Consider a permutation π that is order isomorphic to a proper pin sequence and thus, by Lemma 8, corresponds to at least one strict pin word, say w. If $\pi \notin \mathrm{Av}(B)$ then $\pi \geq \beta$ for some $\beta \in B$. By Lemma 9, β corresponds to a pin word $u \preceq w$. Conversely, if $w \succeq u$ for some u corresponding to $\beta \in B$, then Lemma 9 shows that $\pi \geq \beta$. Therefore the set

$$\{\text{strict pin words } w : w \succeq u \text{ for some } u \text{ corresponding to a } \beta \in B\}$$

consists of all strict pin words which represent permutations not in $\mathrm{Av}(B)$, so by removing this set from the regular language of all strict pin words we obtain the language of all strict pin words corresponding to permutations in $\mathrm{Av}(B)$. In the upcoming lemma we prove that for any pin word u, the set $\{\text{strict pin words } w : w \succeq u\}$ forms a regular language, and thus the language of strict pin words in $\mathrm{Av}(B)$ is regular. It remains only to check if this language is finite or infinite, which can be determined by Proposition 10.

Lemma 12 *For any pin word u, the set $\{$strict pin words $w : w \succeq u\}$ forms a regular language, and a finite accepting automaton for this language can be effectively constructed.*

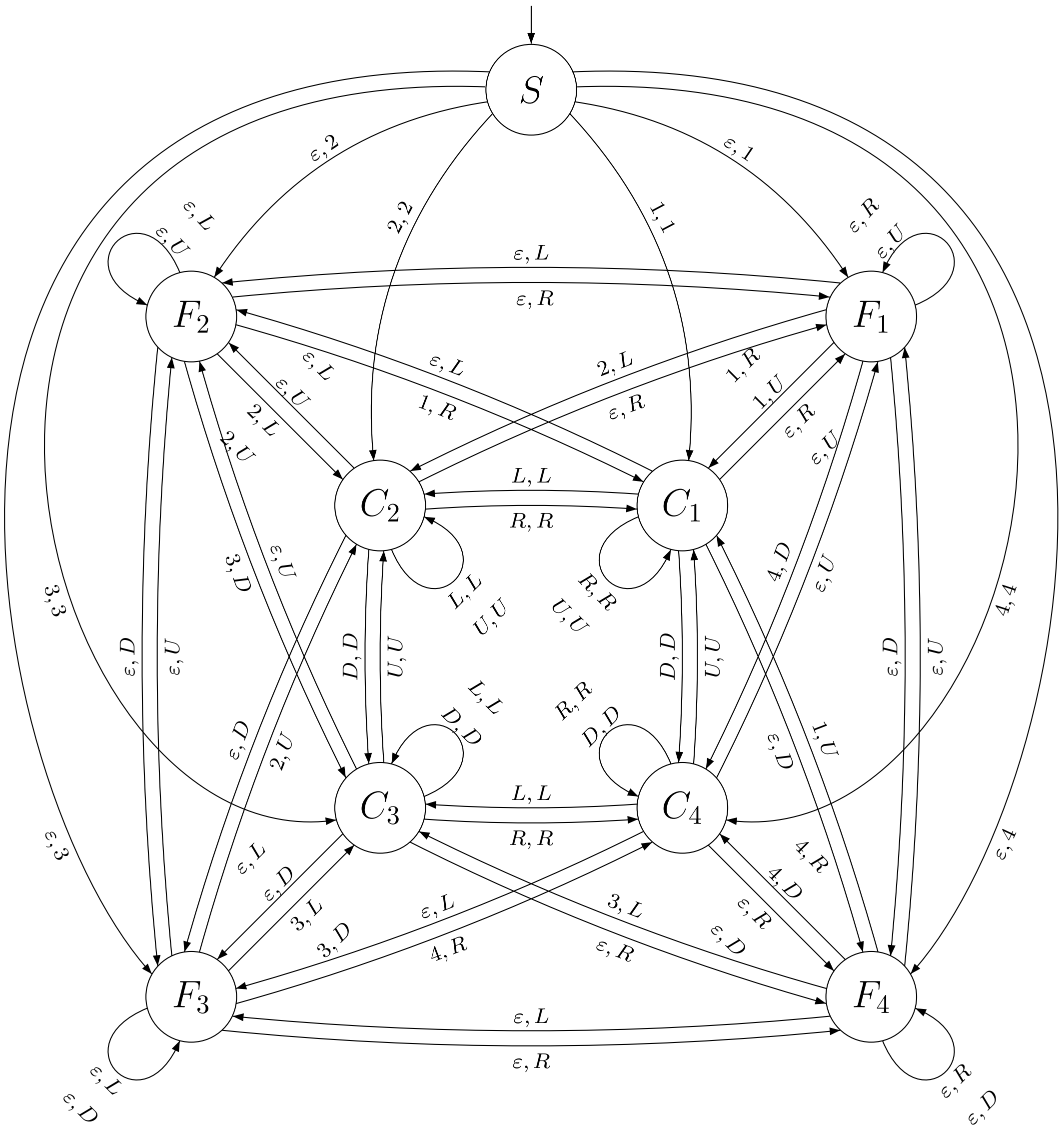

Fig. 8. The transducer that produces all strict pin words containing the input pin word.

Proof. Let T denote the transducer in Figure 8. We claim that a strict pin word w lies in $T(u)$ if and only if $w \succeq u$. The lemma then follows by intersecting $T(u)$ with the regular language of all strict pin words.

We begin by noting several prominent features of T:

(T1) Every transition writes a symbol.

(T2) Other than the start state S, the automaton is divided into two parts, the "fabrication" states F_i and the "copy" states C_i.

(T3) Every transition to a fabrication state has ε input.

(T4) Every transition from a fabrication state to a copy state reads a numeral and writes a direction, and except for the transitions from S, these are the only transitions that read a numeral.

(T5) All transitions between copy states read a direction and write the same direction, these are the only transitions that read a direction, and there is such a transition for every copy state and every direction.

(T6) From every fabrication and copy state, each direction can be output via a transition to a fabrication state with input ε.

(T7) The subscripts of the fabrication and copy states indicate quadrants: if the strict pin word $w_1 \cdots w_n$, corresponding to the pin sequence $p_1, \ldots, p_n$, has just been written by the transducer and the transducer is currently in state C_i or F_i, then p_n lies in quadrant i. Moreover, if the pin word $u_1 \cdots u_m$, corresponding to the pin sequence $q_1, \ldots, q_m$, has been read and the transducer currently lies in the copy state C_i, then q_m lies in quadrant i.

(T8) From any state, any copy state can be reached by two transitions, the first being a transition to a fabrication state; for example: $C_2 \xrightarrow{\varepsilon,D} F_3 \xrightarrow{4,R} C_4$.

First we prove that $w \succeq u$ for every strict pin word w produced from input u by this transducer. We prove this by induction on the number of strong numeral-led factors in u. The base case is when u consists of precisely one strong numeral-led factor. Suppose that the output right before the first letter of u is read is $v^{(1)}$. There are two cases. If $v^{(1)}$ is empty, then the transducer is currently in state S, and must both read and write the first letter of u, moving the transducer into state C_{u_1}. At this point, (T5) shows that the transducer could continue to transition between copy states, outputting a word $w = uv^{(2)} \succeq u$. The only other option available to the transducer (again, by (T5)) is to transition to a fabrication state, but then (T4) shows that the transducer can never again reach a copy state (because u has only one numeral), and thus by (T3), it can never finish reading u. In the other case, where $v^{(1)}$ is nonempty, the transducer lies in a fabrication state by (T4). The next transition must then by (T4) be into a copy state, and (T7) guarantees that the letter written corresponds to a point in quadrant u_1. The same argument as in the previous case shows that the transducer is now confined to copy states until the rest of u has been read, and thus the transducer will output $v^{(1)}w^{(1)}v^{(2)} \succeq u$.

Now suppose that u decomposes into $j \geq 2$ strong numeral-led factors as $u^{(1)} \cdots u^{(j)}$. By induction, at the point where $u^{(j-1)}$ has just been read, the transducer has output a word $v^{(1)}w^{(1)} \cdots v^{(j-1)}w^{(j-1)}$ and lies in a copy state. Since the first letter of $u^{(j)}$ is a numeral, the transducer is forced by (T4) to transition to a fabrication state, and this transition will write but not read by (T3). The transducer can then transition freely between fabrication states. Let us suppose that $v^{(1)}w^{(1)} \cdots v^{(j-1)}w^{(j-1)}v^{(j)}$ has been output at the moment just before the transducer begins reading $u^{(j)}$. As in our second base case above, the transducer must at this point transition to a copy state by (T4), which it will do by reading the numeral that begins $u^{(j)}$ and writing a letter that — by (T7) — corresponds to a point in this quadrant. The situation is then analogous to the base case, and the transducer will output $v^{(1)}w^{(1)} \cdots v^{(j-1)}w^{(j-1)}v^{(j)}w^{(j)}v^{(j+1)} \succeq u$.

Now we need to verify that the transducer produces every strict pin word w with $w \succeq u$. Break u into its strong numeral-led factors $u^{(1)} \cdots u^{(j)}$ and suppose that the factorisation $w = v^{(1)}w^{(1)} \cdots v^{(j-1)}w^{(j-1)}v^{(j)}w^{(j)}v^{(j+1)}$ satisfies (O1) and (O2). If $v^{(1)}$ is nonempty then it can be output immediately by a sequence of transitions to fabrication states by (T6); by (O2) and (T7), the first letter of $w^{(1)}$ (which must be a direction because w is a strict pin word) can then be output by transitioning to a copy state, from which (T5) shows that the rest of $u^{(1)}$ can be read and the rest of $w^{(1)}$ can be written. If $v^{(1)}$ is empty then $u^{(1)} = w^{(1)}$ by (O1). The transducer can, by (T5), read $u^{(1)}$ and write $w^{(1)}$ by transitioning from S to a copy state and then transitioning between copy states. Because w is a strict pin word, (O2) shows that $v^{(2)}$ must be nonempty, and (T6) shows that $v^{(2)}$ can be output without reading any more letters of u. We then must output $w^{(2)}$ whilst reading $u^{(2)}$. The only possible obstacle would be reaching the correct copy state, but (T8) guarantees that this can be done. The rest of u can be read, and the rest of w written, in the same fashion. $\square$

The proof of Theorem 1 now follows from the discussion at the beginning of the section.

6 Unavoidable Substructures in Simple Permutations

The pin words used to prove our main result allow us to prove an unavoidable substructures result for simple permutations. This provides an easy-to-check sufficient (but, n.b., not necessary) condition to guarantee that a permutation class contains only finitely many simple permutations.

We define the *increasing oscillating sequence* to be the infinite sequence

$$4, 1, 6, 3, 8, 5, \ldots, 2k+2, 2k-1, \ldots.$$

A plot is shown in Figure 9.

We define an *increasing oscillation* to be any simple permutation that is contained in the increasing oscillating sequence, a *decreasing oscillation* to be the reverse of an increasing oscillation, and an *oscillation* to be any permutation that is either an increasing oscillation or a decreasing oscillation.

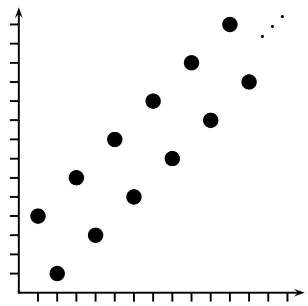

Fig. 9. A plot of the increasing oscillating sequence.

Theorem 13 *Every sufficiently long simple permutation contains an alternation of length k or an oscillation of length k.*

Proof. By Theorem 3, it suffices to prove that every sufficiently long proper pin sequence contains an alternation or oscillation of length k. Take a proper pin sequence $p_1, \ldots, p_m$. By Lemma 8, we may assume that these pins lie in the plane in such a way that $0, p_1, \ldots, p_m$ is also a proper pin sequence, where 0 denote the origin.

We say that this sequence crosses an axis whenever p_{i+1} lies on the other side of the x- or y-axis from p_i, and refer to $\{p_i, p_{i+1}\}$ as a *crossing*. First suppose that $p_1, \ldots, p_m$ contains at least $2k$ crossings, and so crosses some axis at least k times; suppose that this is the y-axis. Each of these y-axis crossings lies either in quadrants 1 and 2 or in quadrants 3 and 4. We refer to these as *upper crossings* and *lower crossings*, respectively. By the separation and externality conditions, both pins in an upper crossing lie above all previous crossings, while both pins in a lower crossing lie below all previous crossings. Thus we can find among the pins of these crossings an alternation of length at least k.

Therefore we are done if the pin sequence contains at least $2k$ crossings, so suppose that it does not, and thus that the pin sequence can be divided into at most $2k$ contiguous sets of pins so that each contiguous set lies in the same quadrant. Each of these contiguous sets is restricted to two types of pin (e.g., a contiguous set in quadrant 3 can only contain down and left pins) and thus since these two types of pin must alternate, these contiguous sets of pins must be order isomorphic to an oscillation (e.g., a contiguous set in quadrant 3 must be order isomorphic to an increasing oscillation). Thus we are also done if one of these contiguous sets has length at least k, which it must if the original pin sequence contains at least $m \geq 2k^2$ pins, proving the theorem. □

Thus a class without arbitrarily long alternations or arbitrarily long oscillations necessarily contains only finitely many simple permutations. First note that these strong conditions are not necessary; for example, the juxtaposition

$\left[\text{Av}(21)\ \text{Av}(12)\right]$ contains arbitrarily long (wedge) alternations, yet the only simple permutations in this class are 1, 12, and 21. The work of Albert, Linton, and Ruškuc [3] also attests to the strength of these conditions; they prove that classes without long alternations have rational generating functions.

Still, there are benefits to having such a straightforward sufficient condition. For example, such classes are guaranteed to be partially well-ordered. As we have already shown how to decide if $\text{Av}(B)$ contains arbitrarily long alternations, to convert Theorem 13 from a theorem about unavoidable substructures to an easily checked sufficient condition for containing only finitely many simple permutations we need to decide if $\text{Av}(B)$ contains arbitrarily long oscillations. As with the parallel alternations from Section 2, the increasing oscillations nearly form a chain in the pattern-containment order, so we need only compute the class of permutations that are contained in some increasing oscillation, or equivalently, order isomorphic to a subset of the increasing oscillating sequence. This computation is given without proof in Murphy's thesis [16], so we prove it below.

Proposition 14 *The class of all permutations contained in all but finitely many increasing oscillations is* $\text{Av}(321, 2341, 3412, 4123)$*.*

Proof. It is straightforward to see that every oscillation avoids 321, 2341, 3412, and 4123, so it suffices to show that every permutation avoiding this quartet is contained in the increasing oscillation sequence. We use the *rank encoding* [4] for this. The rank encoding of the permutation π of length n is the word $d(\pi) = d_1 \cdots d_n$ where

$$d_i = |\{j : j > i \text{ and } \pi(j) < \pi(i)\}|,$$

i.e., d_i is the number of points below and to the right of $\pi(i)$. It is easy to verify that a permutation can be reconstructed from its rank encoding. Now consider the rank encoding for some $\pi \in \text{Av}(321, 2341, 3412, 4123)$. Routinely, one may check:

- $d(\pi) \in \{0, 1, 2\}^*$,
- $d(\pi)$ does not end in 1, 2, or 20,
- $d(\pi)$ does not contain 21, 22, 111, 112, 2011, or 2012 factors.

We now describe how to embed a permutation with rank encoding satisfying these rules into the increasing oscillating sequence. Suppose that we have embedded $\pi(1), \ldots, \pi(i-1)$. If $d_i \geq 1$ then we embed $\pi(i)$ as the next even entry in the sequence. If $d_i = 0$ then we embed $\pi(i)$ as the next odd entry if it ends a 20, 110, or 2010 factor, and as the second next odd entry otherwise. See

[4] We refer the reader to Albert, Atkinson, and Ruškuc [2] for a detailed study of the rank encoding.

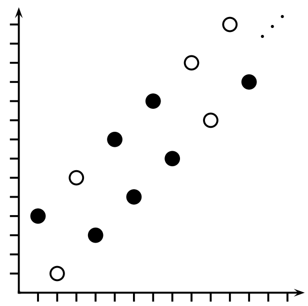

Fig. 10. The filled points show the embedding of 2153647, with rank encoding 1020100, given by the proof of Proposition 14.

Figure 10 for an example. It remains to show that this is indeed an embedding of π; to do this it suffices to verify that the number of points of this embedding below and to the right our embedding of $\pi(i)$ is d_i. This follows from the rules above. □

7 Concluding Remarks

Other contexts. Analogues of simplicity can be defined for other combinatorial objects, and such analogues have received considerable attention. For example, let T be a tournament (i.e., an oriented complete graph) on the vertex set $V(T)$ with (directed) edge set $E(T)$. For a set $A \subseteq V(T)$ and vertex $v \notin A$, we write $v \to A$ if $(v, a) \in E(T)$ for all $a \in A$ and similarly $v \leftarrow A$ if $(a, v) \in E(T)$ for all $a \in A$. An interval in T is a set $A \subseteq V(T)$ such that for all $v \notin A$, either $v \to A$ or $v \leftarrow A$. Clearly the empty set, all singletons, and the entire vertex set are all intervals of T, and T is said to be simple if it has no others. Crvenković, Dolinka, and Marković [11] survey the algebraic and combinatorial results concerning simple tournaments.

In the graph case the term "simple" is already taken; two correspondent terms are *prime* and *indecomposable*. An *interval* (also commonly, *module*) in the graph G is a set $A \subseteq V(G)$ such that every vertex $v \notin A$ is adjacent to every vertex in A or to none. We refer to Brandstädt, Le, and Spinrad's text [7] for a survey of simplicity in this context.

Simplicity has also, to some extent, been studied for relational structures in general, for example, by Földes [12] and Schmerl and Trotter [17].

To the best of our knowledge, no analogue of Theorem 1 is known for these other contexts. An approach similar to the one we have taken would require an analogue of Theorem 3 which, as remarked in [8], remains furtive.

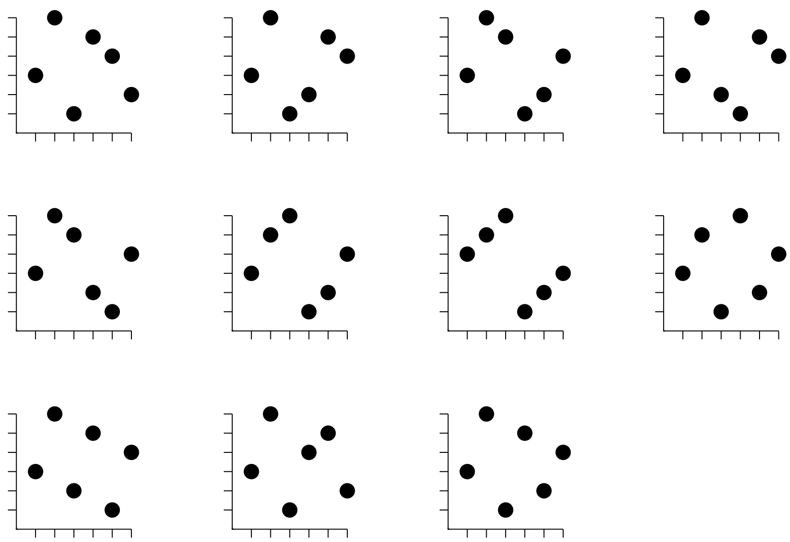

Fig. 11. The basis elements of length 6 for the pin class (up to symmetry).

Partial well-order. Recall that a partially ordered set is said to be *partially well-ordered* (*pwo*) if it contains neither an infinite strictly decreasing chain nor an infinite antichain. While permutation classes cannot contain infinite strictly decreasing chains, there are infinite antichains of permutations, see Atkinson, Murphy, and Ruškuc [5]. A permutation class with only finitely many simple permutations, on the other hand, is necessarily pwo (Albert and Atkinson [1] derive this from a result of Higman [13]). Thus Theorem 1 bears some resemblance to the pwo decidability question:

Question 15 *Is it possible to decide if a permutation class given by a finite basis is pwo?*

This question is considered in more generality by Cherlin and Latka [10].

The pin class. We close with a final, capricious, thought. The set of permutations that correspond to strict pin words forms a permutation class by Lemma 9. As this class arises from words, it has a distinctly "regular" feel, and thus we offer:

Conjecture 16 *The class of permutations corresponding to pin words has a rational generating function.*

The enumeration of this class begins $1, 2, 6, 24, 120, 664, 3596, 19004$. It is not even obvious that this "pin class" has a finite basis. Its shortest basis elements are of length 6, and there are 56 of these (see Figure 11). The class also has 220 basis elements of length 7.

Acknowledgements. We wish to thank Mike Atkinson for fruitful discus-

sions.

References

[1] Albert, M. H., and Atkinson, M. D. Simple permutations and pattern restricted permutations. *Discrete Math. 300*, 1-3 (2005), 1–15.

[2] Albert, M. H., Atkinson, M. D., and Ruškuc, N. Regular closed sets of permutations. *Theoret. Comput. Sci. 306*, 1-3 (2003), 85–100.

[3] Albert, M. H., Linton, S., and Ruškuc, N. The insertion encoding of permutations. *Electron. J. Combin. 12*, 1 (2005), Research paper 47, 31 pp.

[4] Atkinson, M. D. Restricted permutations. *Discrete Math. 195*, 1-3 (1999), 27–38.

[5] Atkinson, M. D., Murphy, M. M., and Ruškuc, N. Partially well-ordered closed sets of permutations. *Order 19*, 2 (2002), 101–113.

[6] Bóna, M. A survey of stack-sorting disciplines. *Electron. J. Combin. 9*, 2 (2003), Article 1, 16 pp.

[7] Brandstädt, A., Le, V. B., and Spinrad, J. P. *Graph classes: a survey.* SIAM Monographs on Discrete Mathematics and Applications. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[8] Brignall, R., Huczynska, S., and Vatter, V. Decomposing simple permutations, with enumerative consequences. arXiv:math.CO/0606186.

[9] Brignall, R., Huczynska, S., and Vatter, V. Simple permutations and algebraic generating functions. arXiv:math.CO/0608391.

[10] Cherlin, G. L., and Latka, B. J. Minimal antichains in well-founded quasi-orders with an application to tournaments. *J. Combin. Theory Ser. B 80*, 2 (2000), 258–276.

[11] Crvenković, S., Dolinka, I., and Marković, P. A survey of algebra of tournaments. *Novi Sad J. Math. 29* (1999), 95–130.

[12] Földes, S. On intervals in relational structures. *Z. Math. Logik Grundlag. Math. 26*, 2 (1980), 97–101.

[13] Higman, G. Ordering by divisibility in abstract algebras. *Proc. London Math. Soc. (3) 2* (1952), 326–336.

[14] Hopcroft, J. E., Motwani, R., and Ullman, J. D. *Introduction to automata theory, languages, and computation*, 2nd ed. Addison-Wesley Publishing Co., Reading, Mass., 2001.

[15] Lakshmibai, V., and Sandhya, B. Criterion for smoothness of Schubert varieties in $\mathrm{SL}(n)/B$. *Proc. Indian Acad. Sci. Math. Sci. 100*, 1 (1990), 45–52.

[16] Murphy, M. M. *Restricted Permutations, Antichains, Atomic Classes, and Stack Sorting*. PhD thesis, Univ. of St Andrews, 2002.

[17] Schmerl, J. H., and Trotter, W. T. Critically indecomposable partially ordered sets, graphs, tournaments and other binary relational structures. *Discrete Math. 113*, 1-3 (1993), 191–205.